UNIVERSITY OF SOUTHAMPTON

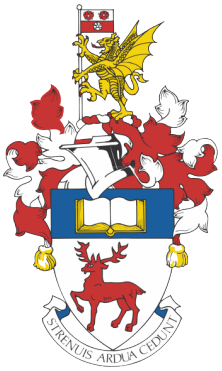
FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Centre for Biological Sciences

Is there a conserved function for the GTG/GPHR family of membrane proteins?

by

Nancy Wong



Thesis for the degree of Doctor of Philosophy

October 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Cellular and Molecular Biology

Thesis for the degree of Doctor of Philosophy

IS THERE A CONSERVED FUNCTION FOR THE GTG/GPHR FAMILY OF MEMBRANE PROTEINS?

Nancy Wong

The G protein-coupled receptor (GPCR)-type G protein/Golgi pH regulator (GTG/GPHR) proteins are a conserved family of membrane proteins in eukaryotes, but are yet to be fully characterised. So far, two possible functions have been described: anion channels for Golgi pH regulation in animals or plasma membrane abscisic acid receptors in plant signalling. Here, the role GTG/GPHRs has been explored using plant (*Arabidopsis thaliana*), animal (*Caenorhabditis elegans*) and fungal (*Saccharomyces cerevisiae*) models. There are two *Arabidopsis thaliana GTG/GPHR* genes, At *GTG1* and At *GTG2*. Knocking out both in the *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants results in shorter root and hypocotyl growth under certain conditions, and expression of either At *GTG* can restore these defects (Jaffé et al., 2012). In this thesis, these growth defects are confirmed in another *gtg1 gtg2* double mutant (*gtg1-1 gtg2-1*). GTG/GPHRs have a conserved glycine in their domain of unknown function, DUF3735. Mutating this glycine (G166) to leucine in At *GTG1* negates its rescue of *gtg1-3 gtg2-3* defects, indicating its importance for function.

There are also two *C. elegans* (*Cel*-) *GTG/GPHR* genes, *Cel-gphr-1* and *Cel-gphr-2*. In *C. elegans*, GFP-tagged *Cel*-GPHR-1 shows an intracellular punctate pattern consistent with Golgi localisation. Backcrossed single mutants (*Cel-gphr-1(ok1579)* and *Cel-gphr-2(tm4228)*) and a double mutant (*Cel-gphr-1(ok1579) gphr-2(tm4228)*) are generated in this project. Single mutants display similar characteristics to wild-type *C. elegans* whereas the double mutant, lacking both *gphr* genes, shows abnormal egg-laying, egg development and hatching, as well as larval arrest and reduced pharyngeal pumping. When expressed in Arabidopsis, *Cel*-GPHR-1 shows a similar localisation pattern to the endogenous At GTGs and rescues the root and hypocotyl defects observed in Arabidopsis *gtg1-3 gtg2-3*. Localisation studies also show that both Arabidopsis and *C. elegans* GTG/GPHRs are Golgi/ER localised in *S. cerevisiae* following heterologous expression. All results are consistent with the GTG/GPHRs having a conserved function across kingdoms.

Table of Contents

List of tables

List of figures

DECLARATION OF AUTHORSHIP

I, Nancy Wong

declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Is there a conserved function for the GTG/GPHR family of membrane proteins?

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission.

Signed: .....

Date:.....

Acknowledgements

Thank you to all past and present L. E. Williams lab members (especially Kerry, Emily, Nick, Paloma, Ilectra, Mike and Tania) who have provided such an organised and friendly environment to work in. From hoarding boxes, lab lunches, detailed risk assessments to an anthology of protocols. Thank you to Prof. Yuichi Iino (University of Tokyo, Japan) for supplying pDEST-EGFP and Prof. H. Kuroyanagi (Tokyo Medical and Dental University, Japan) for pDEST-myo-2p and pDEST-eft-3p plasmids for generating *C. elegans* expression vectors. To Prof. Lindy Holden-Dye and Prof. Vincent O'Connor for help with *C. elegans*. Nico Dallière and Dr. James Dillon for microinjection training. Thank you to the Gerald Kerkut Trust for funding and opportunity to go to the international Annual SEB conference (part-funded by SEB). Data shown in Figure 4.9, Figure 4.12 and Figure 4.13 was collected in collaboration with A. Anzai-Jackson and K. Peaston. But mostly I would like to thank you my supervisors Neil Smyth, Matthew Terry and Lorraine Williams. Your support and guidance has been second to none and this thesis would not be as thorough and long without you! Matthew, thank you for all the treats from Japan! Lorraine I appreciate your genuine care and helping me through it all. Thanks a bunch!

A thesis acknowledgements would not be complete without a mushy myriad of standing ovations on top of a "poignant" narrative to juxtapose the requisite serious tributes. Challenging, breathtaking, exhausting, heartbreaking, soul-destroying, ecstasy, opportunistic, competitive, entertaining, agonising, exuberating, (un)motivating, passionate, rewarding. Blunt, abrasive word association to summarise my life the past four years. When I decided to do a PhD I was naïve, young and quite frankly arrogant. My inquisitive nature, thirst for experimenting and experiences drove me to this pit. Boy, did I speak too soon; I am nostalgic for the sweet bliss of ignorance I had for further education but I would never take back embarking on a PhD. This chronicle of my life is over (for now); watch the public space for "Wong et al. (2015)"! I honestly could not have finished this without many of you, whom I will name for several reasons: (1) I want you to know who you are. (2) If this is the end of our friendship, I want to remember your relevance when I am using this to prop up a door and (infrequently) flicking through. (3) I dabble in corny bullocks. (4) Because this may well be the only thing people read. On y va! Kerry Peaston, you made my lab life: your nervous laughter, training and friendship. Dr Ben Mulcahy and Nico Dallière for taking a candid interest in my work and inspiring me intellectually. Flora O'Brien for being that breath of fresh air. Your kindness, compassion and fitness will endure a trusty long friendship between us. Nash Matinyarare: reminding me of being "single and ready to

mingle"; I could not have got through it without you. The Mighty Cheesy Circle will always be close to my heart (and no doubt a glass of Pomerol). Drs Gregor Elliot, Simon Willcock and Kandela for providing lunchtime entertainment, geography lessons and common sense. Hazel Smith, for being a great friend, always there and thoughtful. Chris Ford: reminding me of how fragile life is, and how we should cherish and embrace every moment, opportunity and person important to us. Matthew Davies: always singing Sisqó and giving the best hugs! Katy Stubbs: always (over)enthusiastic, supportive and a great shoulder to cry on post micros-*coping*. Filip Markiewicz, our friendship was a little late to the party but better late than never. Banterous, mischievous and ten pounds of fun in a five pound bag; with that my friend you will go far in life! All the new postgrads: great diverse personalities that I have been lucky to meet before my departure. Thanks for making this year tolerable, keeping it fresh. Ben Yarnall keep up the hardwork, I look forward to being referenced by you. The AU and Vixens for keeping me well-procrastinated, injecting the endorphins essential to keep me going! Megan Sealey and Tom Bright for keeping a roof over my head and looking after me. Dr David Rusling, for reminding me of my worth, intellect and strength. Knowing you has changed my life, thank you for being that torch. To everyone else in B85 (especially Zoe Harris, Indri Causon, Pzemeck Ociepa, Christine Reitmeyer, Annabelle Damerum, Karolina Mirowska, Dr Tom Secker). My darling Laurence Regan the voice of reason, my soulmate, thank you for the ongoing support. Ste-fanny Payne, Faye England, Sharley Smith, Amie Horne, Sarah Bellis, Carley Quigley and Chiara Pelizzari: you beauties thanks for the sassy fun, chinwags and events to look forward to! Alex Watson-Lazowski, where do I begin? The best house mate: cheers for wining, dining and doing bins. Picking me up (including off the streets), being a pillow to hurl abuse at and obviously the finest faux husband; the divorce has been hard! Dame Helen Hyde for serving the best, priceless secondary education at the one and only Watford Grammar School for Girls. It has indeed geared me up to being the "polished diamond"; living the mantra "*sperati perati*" in true nature. The Watford crew (you know who you are): keeping me afloat, reminding me of what life is missing; the light at the end of the tunnel. The ones I can't wait to celebrate with, I'm coming for you! Michael Hutchinson, you brave (possibly foolish) soul, see you on the flipside – good luck. Finally the family who have provided the financial support required to live like a queen. Dear Wongs how skinny I would be without your sustenance but how tedious my life would be without the extravagant, privileged lifestyle you have so generously supplied. Chez-moi et Le Chinaillon oh comment je vous ai manqué. I love you all: friends, family and Southampton. It's been a pleasure. Adios. Enfin, the world is my oyster.

Definitions and Abbreviations

(+)	WT
(+/-)	heterozygous
(-/-)	homozygous for mutation
ºC	degrees Celsius
ΔG	change in free energy
MΩ	megaOhms
1°	primary
2°	secondary
2D	two-dimensional
3D	three-dimensional
5-FOA	5′-fluoroorotic acid
7TM	seven transmembrane
TM3	third transmembrane segment
TM5	fifth transmembrane segment
TM6	sixth transmembrane segment
ABA	abscisic acid
ABA-GPCR	abscisic acid G-protein coupled receptor domain
ABRC	Arabidopsis Biological Resource Center
ADP	adenosine diphosphate
AL	*T. urartu* chromosome A, long arm
Ala	alanine
AMAN-2	α-mannosidase-II
ANOVA	analysis of variance
APC/C	anaphase-promoting complex/cyclosome
At	*Arabidopsis thaliana* prefix for gene/protein symbols

ATP	adenosine triphosphate
βAR	β_2-adrenergic receptor
β-ME	2-mercaptoethanol
BL	brassinolide
BODIPY	boron-dipyrromethene
bp	basepair
BR	brassinosteroid
BSA	bovine serum albumin
Ca^{2+}	calcium ion(s)
$CaCl_2$	calcium chloride
cADPR	cyclic ADP-ribose
CaMV	cauliflower mosaic virus
Car ^{r}	carbenicillin resistance gene
CDF	Cation Diffusion Facilitator
cDNA	complementary DNA
CDS	coding sequence
Cel-	*Caenorhabditis elegans* prefix for gene/protein symbols
Cg	*Cricetulus griseus* prefix for gene/protein symbols
CGC	*Caenorhabditis* Genetics Center
CHO	Chinese hamster ovary
Cl^-	chloride ion(s)
CLC	chloride ion channels
COPII	coat of protein 2
Col	Columbia
CHX	cation exchangers
Cys	cysteine

D	day(s)
dATP	2′-deoxyadenosine-5′-triphosphate
dCTP	2′-deoxycytidine-5′-triphosphate
ddH_2O	double-distilled water
d.f.	degrees of freedom
dGTP	2′-deoxyguanosine-5′-triphosphate
DL	*T. tauschii* chromosome D, long arm
DIC	differential interference contrast
dicot	dicotyledon
DIDS	4,4′-diisothiocyanatostilbene-2,2′-disulphonic acid
Dm	*Drosophila melanogaster* prefix for gene/protein symbols
DMSO	dimethyl sulfoxide
DNA	deoxyribose nucleic acid
dNTPs	2′-deoxynucleotide-5′-triphosphate
dsRNA	double-stranded RNA
dTTP	2′-deoxythymidine-5′-triphosphate
DUF3735	domain of unknown function 3735
ECR	environmentally controlled room
e.g.	*exemplī grātiā*, “for example”
EGCG	epigallocatechin gallate
EGFP	enhanced GFP
EMS	ethyl methyl sulphonate
ER	endoplasmic reticulum
EST	expressed sequence tag
et al.	*et alii* meaning “and others”

EUROSCARF	EUROpean *Saccharomyces Cerevisiae* ARchive for Functional Analysis
F_1	first generation
F_2	second generation
FLP	FMRF amide-like peptide
FSEC	fluorescence-detection size-exclusion chromatography
g	unit of gravitational acceleration
G	glycine
GDP	guanosine-5′-diphosphate
GEFs	guanine nucleotide exchange factors
GFP	green fluorescent protein
Gln	glutamine
GOLAC	Golgi anion channel
GPCR	G protein-coupled receptor
GPHR	Golgi pH regulator
GPI	glycosylphosphatidylinositol
GPR	orphan G protein-coupled receptor
GTG	GPCR-type G protein
GTP	guanosine-5′-triphosphate
GTPase	GTP-hydrolase
GUS	β-glucuronidase
h	hour
H^+	proton
H_2O	water
HCl	hydrogen chloride
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HOG	high osmolarity glycerol
Hs	*Homo sapiens* prefix for gene/protein symbols
Hyg ^{r}	hygromycine resistance gene
i.e.	*id est*, “that is”
IPTG	isopropyl β-D-1-thiogalactopyranoside
IS pathway	insulin signalling pathway
K^+	potassium ion(s)
kb	kilobase(s)
K_D	dissociation constant
kDa	kilodalton
KH_2PO_4	monopotassium phosphate
knockology	possible outcomes from the random integration of T-DNA into plant chromosomal DNA
KO	knockout
KOH	potassium hydroxide
L	litre(s)
L	leucine
LANCL2	lanthionine synthetase component C-like protein 2
LB medium	Luria-Bertani broth medium
LB	left border
LZM	zinc-limiting media
Mb	megabases
MES	2-(N-morpholino)ethanesulfonic acid
Mg_2Cl_2	magnesium chloride
$MgSO_4$	magnesium sulphate
MD	Meta-Disorder

MSC	mesenchymal stem cells
min (unit)	minute(s)
Mm	*Mus musculus* prefix for gene/protein symbols
mmBCFAs	monomethyl branched-chain fatty acids
monocots	monocotyledons
MOPS	3-(N-morpholino)propanesulfonic acid
$MnCl_2$	manganese chloride
mRNA	messenger RNA
MS medium	Murashige and Skoog medium
MTP1	metal tolerance protein 1
mV	millivolts
N.A.	numerical aperture
Na^+	sodium ion(s)
NaAc	sodium acetate
NaCl	sodium chloride
Na_2 EDTA	ethylenediaminetetraacetic acid, disodium salt
Na_2HPO_4	disodium phosphate
NaOH	sodium hydroxide
NASC	Nottingham (European) Arabidopsis Stock Centre
NBP	National Bioresource Project
NGM	nematode growth medium
NLP	neuropeptide-like protein
NMR	nuclear magnetic resonance
NO	nitric oxide
NP-$_{40}$	tergitol-type nonyl phenoxypolyethoxylethanol detergent
OAT	organic anion transporter

O.D.	outside diameter
OD	optical density
ODCase	orotidine 5′-phosphate decarboxylase
ORF	open reading frame
Os	*Oryza sativa* (rice) prefix for gene/protein symbols
OTL	*Oryza* Tag Line
PAPS	3′-phosphoadenosine 5′-phosphosulfate
PBS	Phosphate buffered saline
PCR	polymerase chain reaction
PCS	phytochelatin synthase
PEG	polyethylene glycol
PheP	phenylalanine and tyrosine transporter
PIPES	piperazine-1,2-bis[2-ethanesulfonic acid]
PLATE solution	PEG, lithium acetate and TE buffer solution
PME	pectin methylesterase
ppm	pumps per minute
Pro	proline
PTGS	post-transcriptional gene silencing
PTX	pertussis toxin-sensitive receptor
PYR/PYL/RCAR	pyrabactin resistant receptor/PYR-like/regulatory components of ABA receptors
qRT-PCR	quantitative reverse-transcriptase polymerase chain reaction
RAM	root apical meristem
RB	right border
RdRP	RNA-dependent RNA polymerase
RI	reliability indices

Rif^r	rifampicin resistance gene
ROI	region of interest
ROS	reactive oxygen species
rpm	rotations per minute
RGS	regulators of G protein signalling
RISC	RNA-independent Silencing Complex
RNA	ribonucleic acid
RNAi	RNA interference
s	second(s)
SAM	shoot apical meristem
Sc	*Saccharomyces cerevisiae* (yeast) prefix for gene/protein symbols
SC medium	synthetic complete medium
SC-ura medium	SC medium without uracil
SDM	site-directed mutagenesis
SDS	sodium dodecyl sulfate
SGD	*Saccharomyces* genome database
SLAC1	slow anion channel-associated 1
siRNA	short interfering RNA
S.O.C. medium	Super Optimal Broth with Catabolite repression medium
START domain	StAR-related transfer domain
ST-mRFP	sialyl-transferase-mRFP
T_0	parental plant
T_1	first-generation (plant/seed)
T_2	second-generation (plant/seed)
T_3	third-generation (plant/seed)

T_4	fourth-generation (plant/seed)
TAE	Tris-acetate-EDTA
TAIR	The Arabidopsis Information Resource
T-DNA	transfer-DNA
TE	Tris-EDTA
TGN	trans-Golgi network
Ti plasmid	tumour-inducing plasmid
TM	transmembrane
TUA6	α-tubulin 6
URGI	Unité de Recherches en Génomique Info
UTR	untranslated region
UV/TMP	utraviolet/trimethylpsoralen
V	volts
V	valine
V-ATPase	vacuolar-type H^+-ATPase
$VSVG^{ex}$	extracellular domain of temperature-sensitive vesicular stomatitis virus G protein
$VSVG^{full}$	full-length temperature-sensitive vesicular stomatitis virus G protein
v/v	volume/volume
w/o	without
w/v	weight/volume
WL	white light
Ws	Wassilewskija
YLS8	Yellow-Leaf-Specific 8

CHAPTER 1

General Introduction

1.1 Membrane proteins

Membrane proteins have fundamental roles in various principal physiological functions in all organisms. Some are key players in signal transduction cascades such as receptors or enzymes; or structurally important, for instance in vesicle tethering components in synaptic signalling; or involved in transport mechanisms by acting as channel proteins. There are some membrane proteins that are highly conserved, for example, the vacuolar-type H^+-ATPase (V-ATPase) is a proton (H^+) pump (Marshansky et al., 2014) that helps regulate acidification of intracellular organelles through maintaining a H^+ gradient (Scheel et al., 2005). Approximately 20 – 30 % of all genes in sequenced genomes of various organisms encode for membrane proteins (Krogh et al., 2001). For example in both *Arabidopsis thaliana* (Arabidopsis from herein) and *Saccharomyces cerevisiae* (*S. cerevisiae* from herein) 18 % of their genomes are estimated to encode for membrane proteins, while this is predicted to be higher (27 %) in *Caernorhabditis elegans* (*C. elegans* from herein) (Ward, 2001). Membrane proteins are essential to organisms and broadly researched, but there are some yet to be fully characterised (Krogh et al., 2001).

1.2 The GTG/GPHRs are a family of conserved membrane proteins found in eukaryotes

The **G** protein-coupled receptor-**t**ype **G** proteins/**G**olgi **pH** **r**egulators (GTG/GPHRs, to be used interchangeably in accordance with the referred function), have been described as a novel family of membrane proteins that are highly conserved in eukaryotes (Maeda et al., 2008, Pandey et al., 2009, Jaffé et al., 2012). *In silico* analysis identified the GTG/GPHRs, including the *H. sapiens* (human) GTG/GPHR orthologue, GPR89, was annotated as an orphan G protein-coupled receptor (GPCR) in the NCBI database (http://www.ncbi.nlm.nih.gov/gene/) (Maeda et al., 2008). GPCRs whose endogenous ligand are unknown, are referred to as “orphan GPCRs” and typically named “GPR” with a number suffix (Libert et al., 1991, Wain et al., 2002).

The GTG/GPHRs in *Cricetulus griseus* (hamster) was identified in a screen and both the hamster and Arabidopsis GTG/GPHRs share high similarity to the human GPR89, as well as other plant and animal GTG/GPHR orthologues (Maeda et al., 2008, Pandey et al., 2009, Jaffé et al., 2012). There are differing roles suggested for the GTG/GPHRs: an abscisic acid (ABA) receptor through a mechanism involving GPCR- and G protein-like features (Pandey et al., 2009), or a Golgi apparatus (Golgi from herein) pH regulator with anion channel activity (Maeda et al., 2008).

1.2.1 GTG/GPHRs are described as GPCR-type G proteins functioning as ABA receptors in Arabidopsis

There are two *GTG/GPHR* genes in Arabidopsis that encode for GTG/GPHR proteins. These have been described as plasma membrane ABA receptors and abbreviated as "GTGs" since the regulatory mechanism interestingly involves characteristics of both GPCRs and G proteins (Pandey et al., 2009). Therefore, these proteins in Arabidopsis will be referred to as GTGs throughout this thesis.

In eukaryotes, GPCRs and G proteins have evolved to facilitate a mechanism for organisms to sense and respond to their environment. In G protein-dependent pathways, GPCRs and G proteins function together; GPCRs receive stimuli and pass the signals onto G proteins to orchestrate downstream events and biological responses. The GPCRs are a family of conserved proteins with seven transmembrane (7TM) segments (Rosenbaum et al., 2009). G proteins are heterotrimeric structures made up of Gα, Gβ and Gγ subunits (Ross and Gilman, 1980). The Gβ and Gγ subunits are denoted as the Gβγ dimer since they exist in an associated dimeric form (Morris and Malbon, 1999, Jones and Assmann, 2004). GPCR signalling is well-studied and resolving the three-dimensional (3D) structure of the β_2-adrenergic receptor (βAR) GPCR in complex with agonist and G protein has furthered our insight into GPCR regulation (Rasmussen et al., 2011).

GPCRs are guanine nucleotide exchange factors (GEFs), which release guanosine-5′-diphosphate (GDP) to allow for guanosine-5′-triphosphate (GTP) binding in the Gα subunit of G proteins (Cherfils and Zeghouf, 2013). As a result of GEF activity, the GTP-bound G protein is in its activated state (Morris and Malbon, 1999, Temple and Jones, 2007, Cherfils and Zeghouf, 2013). In the GDP-bound inactive state, Gα interacts with Gβγ through its N-terminal helix and three switch regions

(Offermanns, 2003). When GTP is bound however, the monomer Gα and Gβγ dimer dissociate from one another forming active substrates that target downstream effectors (Ross and Gilman, 1980). The Gα subunit contains a Ras-like domain harbouring GDP/GTP nucleotide-binding and GTPase activity sites. The intrinsic GTPase activity of Gα stimulates GTP hydrolysis, which is accelerated by GTPase activating proteins (GAPs), leaving Gα bound to GDP once again; this inactivates the G protein and Gβγ reassociates with Gα (Fung and Stryer, 1980, Pedersen and Ross, 1982, Cherfils and Zeghouf, 2013). This process of G protein activation upon perception of GPCR activation is a conserved process (De Lean et al., 1980). In plants, accessory proteins such as regulators of G protein signalling (RGS) also help to regulate the G protein cycle (Chen et al., 2003).

1.2.1.1 Evidence that GTG/GPHRs function as GPCR-type G proteins

The Arabidopsis GTGs (At GTGs) demonstrate both GTP-binding and GTPase activity as shown by real-time assays using boron-dipyrromethene (BODIPY) dyes (Pandey et al., 2009). Fluoresence occurs upon binding of fluorescently labelled nucleotides, indicative of GTP-binding; and decreases upon GTP hydrolysis, indicative of GTPase activity (Willard et al., 2005). These bound fluorescent labels (nonhydrolysable GTPγS-BODIPY FL) can be competed with nonfluorescently labelled nucleotides to show specific binding, resulting in a reduction of fluorescence. The results indicated that both At GTGs demonstrated specific GTP-binding ability, shown by the initial binding of fluorescently labelled nucleotide (increase in fluorescence) that was subsequently efficiently competed with nonfluorescent GTP or GDP (reduction in fluorescence) but not ADP (maintained high fluorescence) (Pandey et al., 2009). The At GTGs demonstrated GTPase activity by a similar method but using a hydrolysable fluorescent label, GTP-BODIPY FL. The At GTGs exhibited an initial increase in fluorescence to show GTP-binding activity, followed by a subsequent reduction of fluorescence which was indicative of GTP hydrolysis (Pandey et al., 2009). In addition, the At GTGs were shown to interact with the canonical plant Gα subunit, GPA1, using a split ubiquitin system and coimmunoprecipitation (Pandey et al., 2009). G protein-binding is a feature of GPCRs, while GTP binding and GTPase activity are classic hallmarks of G proteins (Figure 1.1). This unique mechanism of GTG action is consistent with other molecular accounts of ABA receptor signalling, which indicate

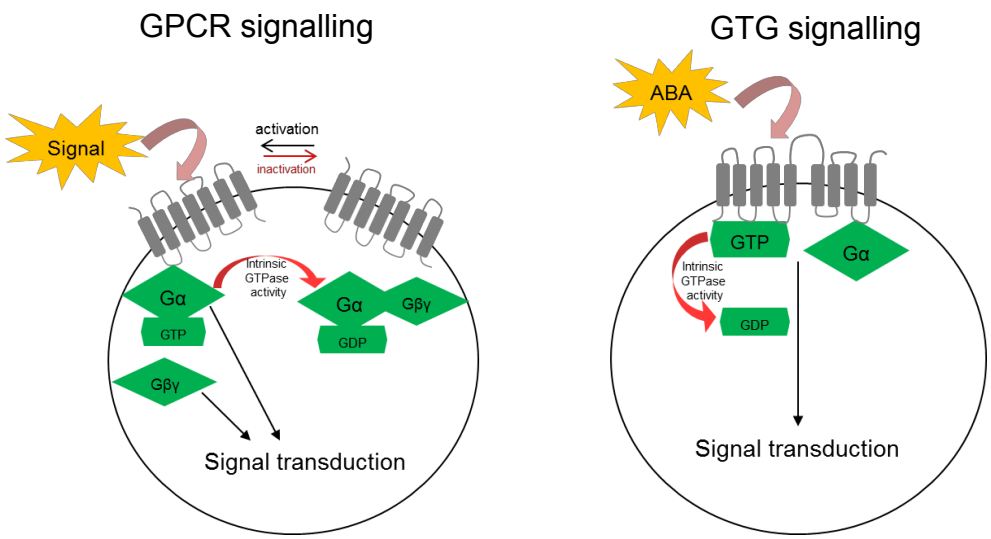


Figure 1.1 GPCR and GTG signalling

Diagrams showing GPCR (**A**) and GTG signalling (**B**). GPCR represented by 7TM (grey) and G proteins (Gα and Gβγ subunits) are shown in green.

that G proteins are involved in ABA signalling (Pandey and Assmann, 2004, Alvarez et al., 2011).

1.2.1.2 ABA and the role of GTGs in plant hormone signalling

The phytohormone ABA plays an integral role in plant biological processes. It is one of the five classic plant hormones and is ubiquitously synthesised in *Viridiplantae* (green plants). Plants respond to fluctuations in the environment by adapting metabolic and developmental biological processes. For instance, ABA is released as an endogenous signal in response to environmental constraints including drought, salt stress and wounding (Zhu, 2002, Christmann et al., 2006, Bassaganya-Riera et al., 2010). When plants suffer from a low water status for example, ABA stimulates stomatal closure in guard cells through increased intracellular calcium pathways and further prevents dehydration by promoting gene expression of for example, *RAB18* (Webb et al., 2001, Christmann et al., 2006, Bassaganya-Riera et al., 2010, Bisht et al., 2011). In the absence of stress ABA also plays a vital role in plant development such as hormonal regulation of seed development, including phases of maturation and dormancy (Umezawa et al., 2010). ABA is therefore an important stress hormone for plant adaptive responses to a changing environment.

To date, there are several candidates for plant ABA receptors including the GTG/GPHRs (Pandey et al., 2009, Alvarez et al., 2013), the pyrabactin resistant receptor (PYR)/PYR-like (PYL)/regulatory components of ABA receptors (RCAR) family of START domain proteins (Ma et al., 2009, Park et al., 2009a, Park et al., 2009b, Santiago et al., 2009, Klingler et al., 2010, Nishimura et al., 2010, Weiner et al., 2010) and the magnesium chelatase subunit H (CHLH) (Wu et al., 2009). Many other candidates have also been proposed as ABA receptors based upon their ability to interact with ABA. For instance, FCA made its debut as an ABA receptor in 2006 (Razem et al., 2006), however subsequent data showed that FCA was unable to bind ABA (Risk et al., 2009). Controversy remains concerning another proposed ABA receptor, GCR2, where a number of *gcr2* mutants failed to demonstrate loss of ABA-sensitivity and lacked ABA-binding affinity (Gao et al., 2007, Guo et al., 2008, Risk et al., 2009). There is also discrepancy in the literature for the CHLH aproposed ABA receptor including its inability to bind ABA and normal ABA responses to postgermination growth and stomatal aperture exhibited by CHLH mutants (Müller and

Hansson, 2009). However, recently it has been reported that CHLH affects ABA signalling in guard cells but is not an ABA receptor (Tsuzuki et al., 2011). In contrast, there is strong evidence supporting the PYR/PYL/RCAR proteins as ABA receptors (Ma et al., 2009, Park et al., 2009b, Santiago et al., 2009, Yin et al., 2009, Nishimura et al., 2010, Wang et al., 2013).

Extensive research has been carried out on plant ABA and its receptors. However this hormone is not only found in plants, but also has also been reported in marine sponges (Puce et al., 2004), fungi (Bassaganya-Riera et al., 2010) and in higher animals (Bruzzone et al., 2007, Bruzzone et al., 2008, Magnone et al., 2009, Scarfì et al., 2009, Sturla et al., 2009, Li et al., 2011, Sturla et al., 2011). In lower animals such as in hydroid sponges, ABA is involved in light-induced tissue regeneration (Puce et al., 2004). Much work has now been conducted using mammalian cells indicating ABA plays an important role in immune responses in animals, as well as having potential medicinal applications for several human diseases including cancer (Li et al., 2011). ABA has recently been described as a pro-inflammatory hormone, produced by human and/or murine granulocytes, β-pancreatic cells and mesenchymal stem cells (MSC) to stimulate cell-specific responses including production of cytotoxic reactive oxygen species (ROS) (Bruzzone et al., 2007, Bruzzone et al., 2008). This was found to involve interaction with a human ABA receptor, lanthionine synthetase component C-like protein 2 (LANCL2) (Sturla et al., 2009, Sturla et al., 2011) through activating a membrane pertussis toxin-sensitive receptor (PTX)/G protein complex. This leads to activation of the second messenger cyclic ADP-ribose (cADPR) and in turn an increase in intracellular Ca^{2+} concentration (Bruzzone et al., 2007, Bruzzone et al., 2008).

In addition, ABA stimulates the release of stress signals resulting in human haematopoiesis progenitor proliferation. A series of downstream signalling events evoked by ABA ultimately leads to the production of immunomodulatory chemokines and cytokines (Scarfì et al., 2009). This suggests that like plant ABA signalling, animal ABA may also function as a stress hormone (Scarfì et al., 2009). Heatshock can also induce ABA release, which showed this hormone as a key component in crosstalk between platelets, monocytes and aortic smooth muscle cells mediating secretory and proliferative responses (Magnone et al., 2009). With recent advances in ABA responses in animals, it has been proposed that an evolutionary common ABA signalling pathway exists for both plants and animals (Tossi et al., 2012). It is suggested that UV-B

stimulates ABA release, promoting an increase in intracellular Ca^{2+} levels leading to nitric oxide (NO) release promoting adaptive responses in plants and animals (Bruzzone et al., 2007, Bruzzone et al., 2012, Tossi et al., 2012). Since ABA signalling is establised in not only plants but also animals and fungi, the conserved GTG/GPHR proteins could act as ABA receptors in all eukaryotes; to date this has only been explored in plants.

In plants, reverse genetics approach was employed to investigate the roles of the At GTGs. Single T-DNA mutants, *gtg1-1* and *gtg2-1*, were isolated in the Wassilewskija (Ws) ecotype and used to generate double *gtg1-1 gtg2-1* mutants (Pandey et al., 2009). While *gtg1-1* or *gtg2-1* single mutants were indistinguishable from wild type (WT), the double *gtg1-1 gtg2-1* mutants exhibited hyposensitivity to ABA in seed germination and post-germination growth assays compared to WT (Pandey et al., 2009). For example, on 1 μM ABA ~50 % of the WT seeds germinated, whilst this was ~80 % for *gtg1-1 gtg2-1*. For WT seeds, the ABA inhibition of seed germination was dose-dependent up to 2 μM ABA; at this concentration no seed germinated. However, at each ABA concentration *gtg1-1 gtg2-1* seed germination was significantly higher than WT. A similar trend was observed in root length assays where ABA inhibited root growth in increasing ABA concentrations in WT seedlings. In contrast, *gtg1-1 gtg2-1* exhibited ABA hyposensitivity with only a small reduction in root length in rising ABA concentrations. The phenotypes exhibited by *gtg1-1 gtg2-1* were complemented by the expression of either At GTG1 or At GTG2 in *gtg1-1 gtg2-1*. This suggested that the ABA hyposensivity was characteristic of *gtg1-1 gtg2-1* mutants (Pandey et al., 2009). Subsequently, additional mutant alleles were isolated, *gtg1-2*, *gtg1-3*, *gtg2-3* and *gtg2*-3, where the *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* double mutants exhibited normal responses to ABA (Jaffé et al., 2012).

The At GTGs also displayed ABA-binding capacity, which was shown to be both specific and saturable (Pandey et al., 2009). In ABA-binding assays using ^{3}H-ABA and purified recombinant At GTGs, the proteins bound to ^{3}H-ABA with dissociation constant (K_D) values of 35.8 nM and 41.2 nM, respectively. In addition, recombinant At GTG1 expressed in *S. cerevisiae* exhibited ABA-binding capacity *in vitro* with a K_D value of 80 nM (Kharenko et al., 2013). This study reinforces the ability of At GTG1 to interact with ABA.

A quantitative proteomics approach was conducted in Arabidopsis roots to corroborate a role for the At GTGs in ABA responses (Alvarez et al., 2013). Quantitative changes in protein abundance comparing WT and *gtg1-1 gtg2-1* were analysed to assess the effect of ABA on the root proteome and whether this was affected by the loss of At GTGs. The results revealed that functional At GTGs are needed for ABA responses of more than two-thirds of ABA-regulated proteins; this further supports their role in ABA signalling. Remarkably the PYL11, PYL12 and PYL13 subclass of PYR/PYL receptors were significantly downregulated by ABA in WT but not in *gtg1-1 gtg2-1*; this suggests possible crosstalk between different ABA receptor subclasses (Alvarez et al., 2013). The data comparing the *gtg1-1 gtg2-1* mutant with its WT further support that the At GTGs are ABA receptors in plants.

1.2.2 The GTG/GPHRs are described as Golgi pH regulators and have voltage-dependent anion channel activity

The GTG/GPHRs have also been described as Golgi pH regulators and have been shown to have voltage-dependent anion channel activity in mammalian cells (Maeda et al., 2008). The human, *Mus musculus* (mouse), hamster and *Drosophila melanogaster* (*Drosophila* from herein) GTG/GPHRs have been studied and denoted as GPHRs (Maeda et al., 2008, Tarutani et al., 2012, Charroux and Royet, 2014); for this reason, these animal GTG/GPHRs will be referred to as the GPHRs throughout this thesis.

In eukaryotes, it is imperative to maintain the appropriate luminal pH of organelles within the secretory and endocytic pathways for correct cellular function such as *N*-glycosylation (Weisz, 2003, Rivinoja et al., 2009). These pathways are involved in correct protein targeting, sorting and modifications. Any deviation from the optimal luminal pH for these pathways will impair cargo protein loading, lipid processing and trafficking (Tartakoff et al., 1978, Axelsson et al., 2001), as well as compromise the morphological integrity of the Golgi (Kellokumpu et al., 2002). Furthermore, congenital diseases such as Alzheimer's and cancer have been shown to be correlated to defective pH regulation of the Golgi lumen (Huse et al., 2002, Kellokumpu et al., 2002, Rivinoja et al., 2006, Rivinoja et al., 2009, Rivinoja et al., 2011, Demaegd et al., 2013).

To date, the mechanisms underlying luminal acidification remain incomplete. However, it is broadly assumed that moderation of luminal pH involves an equilibrium

between the rates of intraluminal H^+ distribution. This implicates the V-ATPase H^+ pump, a counterion conductance, i.e. chloride (Cl^-) ions, and intrinsic proton leakage (Weisz, 2003, Paroutis et al., 2004, Maeda et al., 2008). In animals, all these processes are implicated and V-ATPase transports H^+ into the Golgi lumen, increasing the membrane potential (Nishi and Forgac, 2002). This stimulates the requirement for counterion conductance to decrease the membrane potential; as this declines it drives V-ATPase to transfer more H^+ into the Golgi lumen. Intracellular organelles possess various Cl^- ion channels including the Cl^- intracellular channel, members of the voltage-dependent chloride channel, mid-1-related Cl^- channel and Golgi anion channel families (GOLAC) (Nordeen et al., 2000, Thompson et al., 2002b, Ohgaki et al., 2010). There have been reports that suggest involvement of a few of these channels in endosomal acidification. None however, have been shown to demonstrate luminal acidification of the Golgi.

In plants, a number of proteins are thought to regulate organelle pH including V-ATPase, sodium/proton NHX-type antiporters (Martinière et al., 2013, Shen et al., 2013) and AP-3 β-adaptin (Niñoles et al., 2013). As in animals, the V-ATPase appears to be responsible, in part, for pH regulation in the trans-Golgi network (TGN) and vacuole. This was shown using pH sensors such as the plant-solubility-modified ecliptic pHluorin, which was used to demonstrate a drastic rise in pH when the V-ATPase inhibitor concanamycin A was applied (Shen et al., 2013). With gaps in our knowledge regarding Golgi pH regulation, the GPHR may be a candidate Golgi pH regulator.

A mutant Chinese hamster ovary (CHO) cell line, C27, was established displaying a primary defect of delayed protein transport in flow cytometric assays compared to the parental WT cell line, FF8 (Maeda et al., 2008). The C27 cells transported two non-glycosylphosphatidylinositol-(GPI) reporter proteins, $VSVG^{ex}$-Flag-EGFP-GPI and Flag-$VSVG^{full}$-EGFP; and a transmembrane (TM) protein, endoglycosidase-H resistant VSVG, more slowly than the parental FF8 cells. Therefore, transport of both GPI-anchored and TM cargos from the ER to the plasma membrane are impaired in C27 cells. In addition, the mutant C27 cells processed cargo proteins with impaired glycosylation, as shown by smaller relative molecular masses compared to WT. Electron microscopy demonstrated that C27 cells also displayed Golgi disorganisation, characterised by the loss of Golgi stacks, swelling and vesiculation of cisternae (Maeda et al., 2008). The luminal pH of the Golgi and TGN in the mutant C27 cells was

significantly higher than WT, as shown by three Golgi pH-sensing proteins. The functions of lysosomes and endosomes in C27 mutant cells were unaffected by defective GPHR, suggesting that GPHR activity was specific to the Golgi (Maeda et al., 2008). These findings were consistent with the phenotypes caused by alkalinsation of Golgi pH in other studies (Tartakoff et al., 1978, Chapman and Munro, 1994, Axelsson et al., 2001, Puri et al., 2002). Therefore acidification is indispensable for normal Golgi function.

DNA sequencing using C27 mRNA identified that the hamster *GPHR* gene had a nonsense mutation that would result in a small truncated protein; this was confirmed by reverse transciptase polymerase chain reaction (RT-PCR) and immunoprecipitation indicating the absence of GPHR protein in C27 cells (Maeda et al., 2008). The hamster *GPHR* was transfected into the mutant C27 cells, which restored the observed phenotypes. RNA interference (RNAi) was also used to confirm that *GPHR* knockdown could cause the characteristic mutant phenotypes, indicating that these phenotypes were specific to the loss of *GPHR* (Maeda et al., 2008).

The human GPHR was expressed in an artificial lipid bilayer to show that it exhibited a typical single channel activity, but with complicated gating behaviour, having at least four and possibly five – six subconductance states. At low membrane potentials (less than 30 mV), the channel was mostly open. However, at higher membrane potentials the channel appeared to transition into subconductance states, which was indicative of voltage-dependence. The GPHR is anion selective, as demonstrated by the relative halide anion permabilities measured (I^- (1.5) > Cl^- (1.0) = Br^- (1.0) > F^- (0.9) relative to Cl^-), as well as exhibiting inhibition using 4,4′-diisothiocyanatostilbene-2,2′-disulphonic acid (DIDS; anion transport inhibitor). These results suggested that the mammalian GPHR functions as an anion channel to regulate organelle pH (Maeda et al., 2008).

1.2.2.1 The loss of animal *GPHR* causes abnormal growth and development

The importance of GPHR in growth and development has also been shown using tissue-specific and whole organism *GPHR* knockout (KO) mutants. In mice, GPHR-dependent Golgi functions are crucial for lamellar body and skin barrier function (Tarutani et al., 2012). Skin-specific *GPHR*-KO mice (K5-Cre:$GPHR^{f/f}$) exhibited growth retardation and impaired development in their earlobes and external genitals in

addition to the skin. Immediately after birth, there was no significant difference between the K5-Cre:*GPHR*$^{f/f}$ and WT neonatal mice. However, 4 d after birth the K5-Cre:*GPHR*$^{f/f}$ mice showed hypopigmentation and after a week this manifested into scaliness, skin inflammation and the loss of hair (Figure 1.2 A and B). Histological examination of the newborn K5-Cre:*GPHR*$^{f/f}$ mice demonstrated distended basal cells and follicular dysplasia. Keratinocyte maturation was also affected in K5-Cre:*GPHR*$^{f/f}$ mice. The results showed expression of keratin 6, which is an abnormal differentiation marker as well as several lamellar body markers with reduced expression; these included filaggrin, ceramide, TGN46, TGN and Kallikrein 7a. These data suggested that lamellar body formation was impaired by keratinocyte GPHR deficiency. Furthermore about half of the K5-Cre:*GPHR*$^{f/f}$ mice died within 1 month, showing that even a tissue-specific KO of *GPHR* is very detrimental to life (Tarutani et al., 2012).

In another study, the complete loss of GPHR function in a whole organism was analysed. *Drosophila GPHR* null mutants (denoted Dm *GPHR*^{k34} and Dm *GPHR*^{LL03674} by Charroux and Royet (2014), collectively Dm *GPHR* mutants in this thesis) exhibited a range of growth defects as well as disrupted endoplasmic reticulum (ER) and Golgi organisation (Charroux and Royet, 2014). For example, both Dm *GPHR* mutants exhibited lethality at larval stages of development, but this was considered sublethal since there were some that escaped this arrest (escapers). These Dm *GPHR* mutant escapers were smaller in appearance (dwarf flies) compared to controls (Figure 1.2 C), with dramatically impaired developmental growth as shown by reduced weight and smaller wings, as well as significantly smaller wing cell size and reduced number of wing cells. These phenotypes were restored by Dm GPHR expression indicating that defects were characteristic of knocking out the Dm *GPHR* gene. The reduced larval growth rate and developmental arrest are hallmarks of mutants defective in the insulin signalling (IS) pathway (Hietakangas and Cohen, 2009). However, Dm GPHR did not appear to directly act via the IS pathway, as shown by unaffected expression levels of IS pathway transcriptional targets and similar phalloidin staining of wing imaginal discs to WT (Charroux and Royet, 2014).

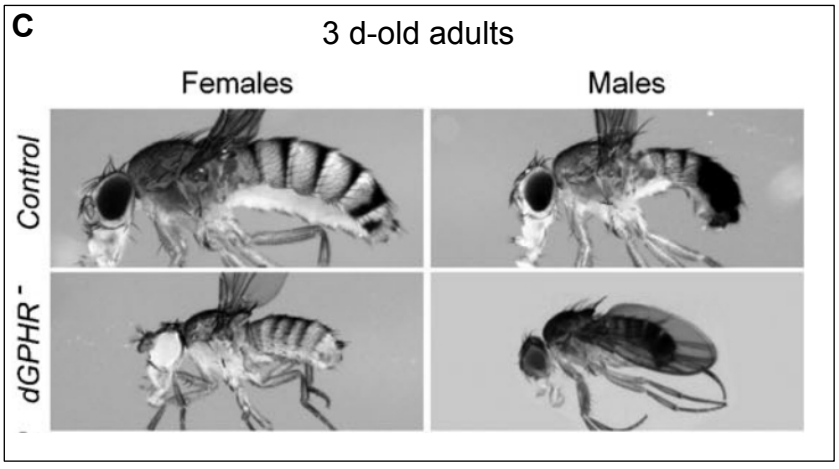


Figure 1.2 GPHR is required for normal growth and development in animals

Mouse skin-specific *GPHR* KO mutants (K5-Cre:*GPHR*$^{f/f}$) at 5 d (**A**) and 1 month (**B**) post birth. Images taken from Tarutani et al. (2012). **C**. *Drosophila* whole organism *GPHR* KO mutants (*dGPHR* $^{-}$) at 3 d. Images taken from Charroux and Royet (2014).

1.3 Other roles for the GTG/GPHRs

1.3.1 The Arabidopsis *gtg1-1 gtg2-1* mutant exhibits phenotypes related to other phytohormones

Following the Arabidopsis root proteomic analysis, differential abundance of many proteins in the *gtg1-1 gtg2-1* mutant compared to WT were found. These included those associated with signalling or biosynthesis of plant hormones including jasmonic acid, ethylene and BRs. For this reason, the effect of these phytohormones on *gtg1-1 gtg2-1* and WT seedling growth was tested (Alvarez et al., 2013). The *gtg1-1 gtg2-1* mutant also exhibited less sensitivity to ethylene in root and hypocotyl inhibition assays, while BL hyposensitivity was only observed in root inhibition assays. Contrastingly, *gtg1-1 gtg2-1* was hypersensitive to jasmonic acid in hypocotyl inhibition assays. These results indicated that the At GTGs may be more than just ABA receptors and are also involved in plant hormone crosstalk (Alvarez et al., 2013).

1.3.2 The Arabidopsis *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants exhibit phenotypes not related to ABA

In another study, different single *gtg1* (*gtg1-2* and *gtg1-3*) and *gtg2* (*gtg2-2* and *gtg2-3*) mutants were isolated (Jaffé et al., 2012), distinct from the *gtg1-1* and *gtg2-1* alleles used previously by Pandey et al. (2009). Single *gtg1-2* and *gtg2-2* mutants in the Columbia (Col) ecotype and Ws *gtg1-3* and *gtg2-3* mutants, were used to generate the two distinct double *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants. For both ecotypes, segregation ratios were distorted whilst isolating homozygous *gtg1(-/-) gtg2(-/-)* double mutants from heterozygous *gtg1(+/-) gtg2(+/-)* progeny. Very few double homozygous *gtg1 gtg2* mutants were isolated from double heterozygous plants, which was significantly different to the expected Mendelian segregation ratio of 1 in 16 for both ecotypes (Jaffé et al., 2012).

As previously indicated, the *gtg1-1 gtg2-1* mutant exhibited hyposensitivity to ABA (Pandey et al., 2009). However, both *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants displayed normal ABA responses as well as exhibited other interesting phenotypes suggesting that the At GTGs play an important role in light-dependent seedling growth and fertility but not ABA signalling (Jaffé et al., 2012). Both *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants exhibited short root length and reduced fresh weight compared to WT in

the absence of sucrose. Under low light conditions, hypocotyl growth was also reduced in the *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants and interestingly this was exacerbated by the presence of sucrose. In the dark however, the hypocotyl lengths were comparable to the WT. Observations of the hypocotyl and cotyledon surface showed large distended cells in both *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3*. Transverse sections through these double mutant hypocotyls and cotyledons revealed irregular and disrupted cells (Figure 1.3) (Jaffé et al., 2012). The *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants displayed defective fertility phenotypes (Jaffé et al., 2012). The silique lengths and seed yield were reduced in *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* compared to WT; this was more marked in *gtg1-2 gtg2-2*. Reduced silique length could result from the lack of pollen tube extension since there is a strong correlation between seed number and silique length in Arabidopsis (Cox and Swain, 2006). Therefore, *in vitro* and *in vivo* assays were conducted to investigate whether the distorted segregation ratios, silique and seed phenotypes resulted from defective pollen. *In vitro* growth of pollen showed that both *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants exhibited reduced pollen germination and pollen tube growth. These pollen tubes displayed no growth further than approximately half way along the pistil compared to WT pollen tubes, which grew along the whole length of the pistil. *In vivo*, *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* pollen tubes displayed less extension compared to WT and again this was was more noticeable in the *gtg1-2 gtg2-2* mutant. The defective pollen functionality in *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants was suggested to cause the fertility defects observed: reduced silique lengths, seed yield and distorted segregation ratios (Jaffé et al., 2012).

These studies show there are differing reports for the function of the GTG/GPHRs, it is therefore important to clarify what this role may be. Extended characterisation of the Arabidopsis *gtg1 gtg2* double mutants would help elucidate the role for plant GTG/GPHRs; this would be critical for either supporting or rejecting the proposed role as ABA receptors.

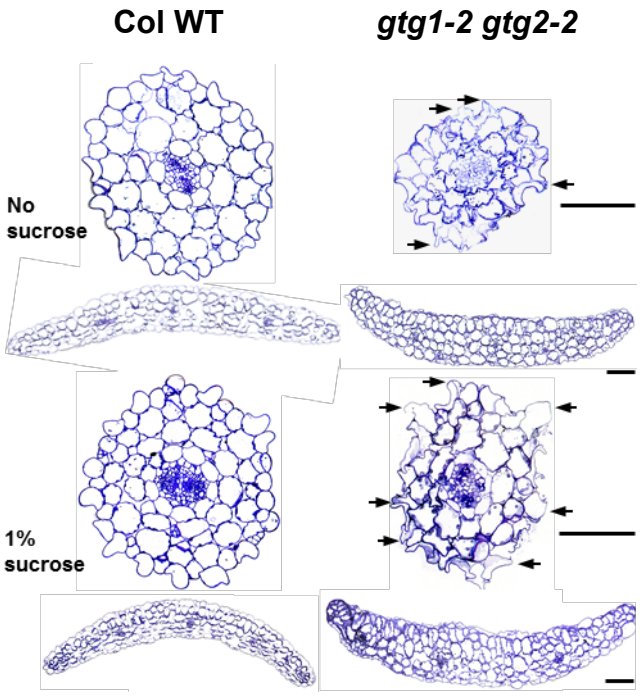


Figure 1.3 Arabidopsis *gtg1-2 gtg2-2* exhibits deformed and disrupted cells

Electron microscopy showing the hypocotyls (upper) and cotyledons (lower) transverse sections of Arabidopsis *gtg1-2 gtg2-2* and Col WT seedlings. The appearance of irregular, disrupted and deformed cells is exacerbated in the presence of 1 % sucrose. Black arrows indicate distended cells. Scale bar: 200 µm. Figure taken from Jaffé et al. (2012).

1.4 The localisation of the GTG/GPHRs

1.4.1 Spatial expression

In two separate reports, the At *GTG* genes consistently displayed broad expression at the tissue level (Pandey et al., 2009, Jaffé et al., 2012). Quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR) revealed that At *GTGs* are widely expressed throughout the plant, and this was also supported using β-glucuronidase (GUS) staining which showed expression in guard cells, roots, stems, flowers and cotyledons (Pandey et al., 2009). Jaffé et al. (2012) carried out an independent GUS reporter analysis showing that both At *GTG1* and At *GTG2* are not only expressed throughout young seedlings (cotyledons, hypocotyls and roots with particularly strong expression in the root tip and vascular system) but also mature plant stages (Figure 1.4). In mature plants they were highly expressed in the vascular system of mature leaves and hydathodes. In the flowers high levels of expression were found in the stigma, stamen, and pollen as well as at the base of the carpel (Figure 1.4) (Jaffé et al., 2012).

1.4.2 Cellular localisation

Protein subcellular localisation can be predicted computationally (section 3.1.2) and determined experimentally. Fluorescent microscopy is one of the most common experimental techniques for defining the localisation of proteins (Yuste, 2005). Green fluorescent protein (GFP)-derived from jellyfish *Aequorea victoria* is now a widely used tool for studying biological processes (Chudakov et al., 2010). Protein localisation can be investigated using stable and transient expression methods. For example, transient expression of Arabidopsis chloride channels (CLCs) have shown the localisation of the Arabidopsis CLC-e and CLC-f in thylakoid membranes in chloroplasts and Golgi membranes, respectively (Marmagne et al., 2007). In addition, slow anion channel-associated 1 (SLAC1) was transiently expressed in not only onion epidermis, but also tobacco protoplasts, to show colocalisation of this channel with the FM4-H4 plasma membrane stain in guard cells (Vahisalu et al., 2008).

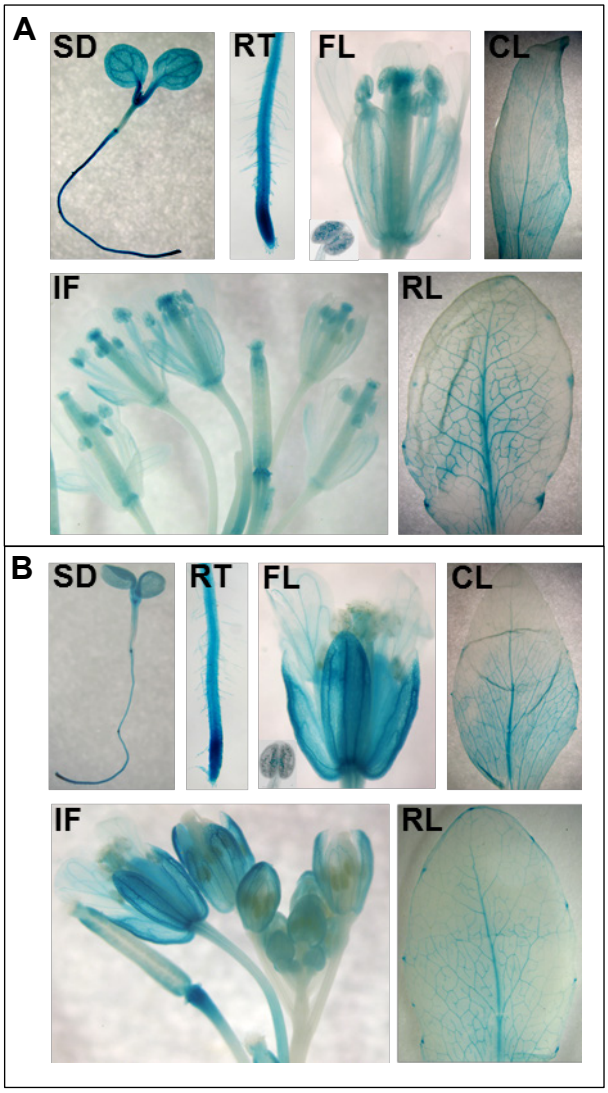


Figure 1.4 The Arabidopsis *GTG* genes are widely expressed

GUS staining (blue) to show the broad spatial expression of the Arabidopsis *GTG* genes. **A**. Expression of At *GTG1* genomic reporter, *PAt GTG1:At GTG1:GUS*. **B**. Expression of At *GTG2* genomic reporter, *PAt GTG2:At GTG2:GUS*. CL, cauline leaf; FL, flower and insert photograph of anther showing GUS-stained pollen; IF, inflorescence; RL, rosette leaf; RT, root tip; SD, seedling. Figure taken from Jaffé et al. (2012).

There are different reports for the subcellular localisation of the GTG/GPHRs. The GTG/GPHRs have been shown to be localised at the Golgi membrane in CHO cells (Maeda et al., 2008); Arabidopsis and tobacco leaf cells (Figure 1.5 A – C) (Jaffé et al., 2012), and in *Drosophila* (Charroux and Royet, 2014). These GTG/GPHRs have also been reported to reside in the ER membrane in *Drosophila* (Charroux and Royet, 2014), Arabidopsis and tobacco leaf cells (Figure 1.5 A and B). However, another study shows that the GTGs are expressed at the plasma membrane in Arabidopsis protoplasts (Figure 1.5 D – E) (Pandey et al., 2009). In addition to the intracellular localisation of the At GTGs by Jaffé et al. (2012) it was also demonstrated that At GTG GFP-tagged proteins localised to both the periphery and interior of mesophyll protoplasts (Jaffé et al., 2012). Therefore, the inconsistencies in the literature not only encompass the function but also the localisation of the plant GTGs. Since the localisation of proteins is often linked to their function, determining the subcellular localisation of the GTG/GPHRs will be critical for either supporting or rejecting the proposed functions for these proteins. On the other hand, the subcellular localisation may even suggest an alternative role for these proteins.

1.5 Studying protein function using model organisms

1.5.1 Reverse genetics to study gene function

Reverse genetics is an approach used to study gene function. Mutant analysis and gene silencing are used to investigate the phenotypes that may arise from a particular genetic sequence. Mutants from different model organisms are available for reverse genetics. For example, phenotypic analysis of plant T-DNA mutants (described in section 4.1.1.1) can be used to study the effects of the loss of a (or many) gene(s). Gene silencing is also a widely used technique for studying gene function, and is discussed in section 6.1.3.2. In this thesis, a number of model organisms will be used to study GTG/GPHR function, including Arabidopsis (plant), *C. elegans* (animal) *S. cerevisiae* (fungus).

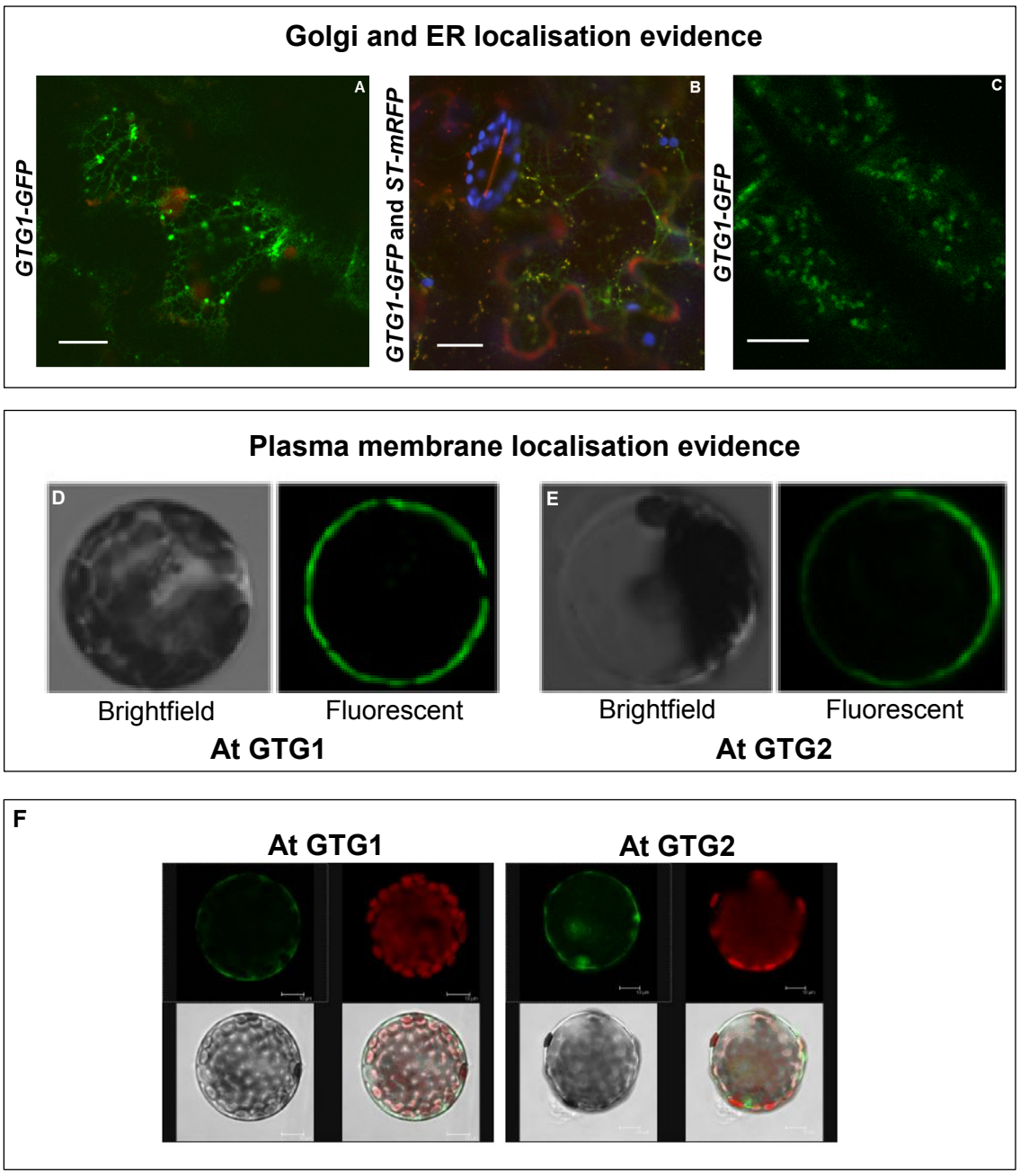


Figure 1.5 There are different reports for the localisation of *Arabidopisis* GTGs

Confocal fluorescent microscopy showing Arabidopsis GTG1 expressed in the Golgi and ER network (**A – C**), and both Arabidopsis GTGs expressed at the plasma membrane (**D – E**). **A**. *P35S:At GTG1:GFP* (*GTG1-GFP*) transient expression in ER and Golgi of tobacco leaf cells. **B**. *P35S:At GTG1:GFP* (*GTG1-GFP*) transient expression showing colocalisation (yellow) with Golgi marker, sialyl-transferase-mRFP (ST-mRFP) (red). **C**. *PAt GTG1:At GTG1:GFP* (*GTG1-GFP*) stable expression in unexpanded Arabidopsis root cells showing Golgi localisation. **D** and **E**. Brightfield and confocal fluorescent images showing At GTG:GFP expression at the plasma membrane of mesophyll protoplasts. **F**. GFP-tagged GTG1 and GTG2 at the periphery and in the interior of mesophyll protoplasts. Different images of protoplasts transfected with pART7

containing *GTG1-GFP*,*GTG2-GFP* or GFP alone. Protoplasts are shown as a set of four grouped images: GFP fluorescence (top left), autoflorescence of chloroplasts (top right), transmission light (bottom left) and combined (bottom right). White scale bar indicates 10 μm. Images taken from Jaffé et al. (2012) (**A – C**, **F**) and Pandey et al. (2009) (**D – E**).

1.5.2 Arabidopsis as a plant model organism

Arabidopsis is a member of the *Brassicaceae* (mustard) family. Arabidopsis is a dicotyledon (dicot) and was the first plant species to have its genome sequenced and since then it has become the organism of choice for a variety plant science studies (AGI, 2000). The genome of Arabidopsis constitutes 125 Mb (AGI, 2000) and databases such as The Arabidopsis Information Resource (TAIR: http://www.arabidopsis.org/) and the European/Nottingham Arabidopsis Stock Centre (NASC: http://arabidopsis.info/) offer fundamental genomic information. Research using Arabidopsis as a model organism has contributed to an extensive understanding into all aspects of modern biology. For instance, analysing Arabidopsis mutants has led to the identification of plant hormone receptors (Inoue et al., 2001) and components in hormone signalling (Santner and Estelle, 2009), as well as advancing biochemical and cellular insights including ion transport (Krysan et al., 1996) and cell wall formation (Brown et al., 2005). Young Arabidopsis seedlings, specifically the roots, are relatively translucent; this makes the plant suitable for light microscopy analysis. Live cell imaging using both fluorescence and confocal laser scanning microscopy can also be conducted on Arabidopsis seedlings through wet-mounting in water or culture media (Shaw, 2006). Plasmids expressing fluorescent protein fusions can be used to study protein localisation (section 4.1.1.2) by stable integration into the plant genome. For all these reasons, Arabidopsis serves as a useful plant model system for studying genes.

1.5.3 The life cycle of Arabidopsis

One special feature of plant development is that new organs are formed at the growing apices during the life of the plant; plants can adapt to the changing local environment. Not all organs that make up a plant are visible in a mature embryo; roots, additional inflorescences, leaves and flowers for instance, are progressively grown throughout the plant life in response to environmental cues. A mature Arabidopsis plant and its main organs are shown in Figure 1.6. Plants progress through three stages of development: embryogenesis, vegetative and reproductive development to ultimately form a mature plant that self-fertilises. At the end of embryogenesis there are two groups of meristem cells from which the entire plant structure is subsequently derived. Meristems enable plants to adapt to a changing environment through regulated proliferation. The root and shoot apical meristems (RAM and SAM respectively) give

rise to the root and shoot organs of the plant. The root can be categorised into separate zones that possess distinct characteristics. The boundaries are not strictly discrete but generally can be divided into four developmental regions within the root tip: the root cap, meristematic zone, elongation zone and the maturation zone Figure 1.7 (Zeiger, 2006).

Arabidopsis has a rapid life cycle of ~8 weeks from germination to mature seed and this is depicted in Figure 1.8. Embryogenesis occurs in the ovule upon fertilisation, which is the process of pollen sperm cell fusing with the egg cell to generate a zygote. The single cell develops into a mature embryo with defined sequential phases: zygotic, globular, heart, torpedo and mature stages. Within this framework, groups of cells become specialised to form epidermal, cortical and vascular tissues. When embryogenesis is completed food reserves are accumulated in the mature embryo and the seed enters a phase of dormancy (Raz et al., 2001, Le et al., 2010, van Zanten et al., 2011). Seed dormancy is important for seeds to regulate germination when environmental conditions are favourable (Finch-Savage and Leubner-Metzger, 2006). After embryogenesis the next stage of plant development is vegetative growth. This is marked by seed germination, which is the growth of the plant embryo from the seed and forming of a seedling, with the first embryonic leaves (cotyledons) emerging from the hypocotyl (Figure 1.8). After a period of vegetative growth, the specialised floral meristems develop that give rise to flowers; the reproductive organs of angiosperms. Flowers comprise four organs: sepals, stamens, petals and carpels (Figure 1.6). The stamen produces pollen and is the male reproductive organ. The carpel constitutes a basal portion containing the ovaries, an elongated structure called the style and an apical structure, the stigma; these make up the female reproductive organ (Zeiger, 2006). Pollination is the pollen transfer to the stigma by the process of pollen grain germination giving rise to a pollen tube which extends down the style towards the ovary where pollen fertilises an ovule, which develops into a seed (Vivian-Smith and Koltunow, 1999). Silique (fruit) development begins when ovules become fertilised. Siliques elongate from the centre of a flower within 24 h post-fertilisation and can reach a mature length merely 4 d later. The silique length ranges according to both developmental age and the number of seeds. For example, siliques with fewer seeds are shorter in length. During silique development the petals, sepals and stamens detach. As seeds ripen the silique browns and dessication of the seed typically occurs within two

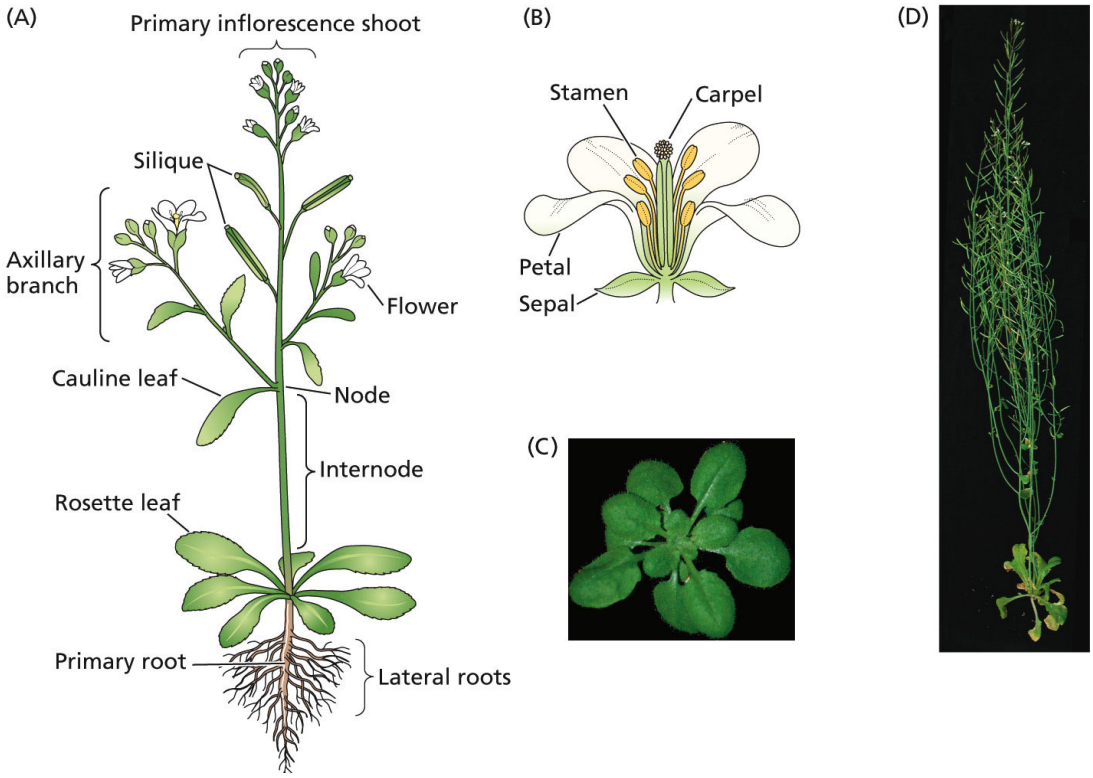


Figure 1.6 The anatomy of the Arabidopsis plant

A. Diagram of a mature Arabidopsis plant showing various organs. **B**. Diagram of flower showing the four main floral organs. **C**. An immature vegetative plant comprising the basal rosette leaves. **D**. The mature plant after the flower maturity and silique development. Figure taken from Zeiger (2006).

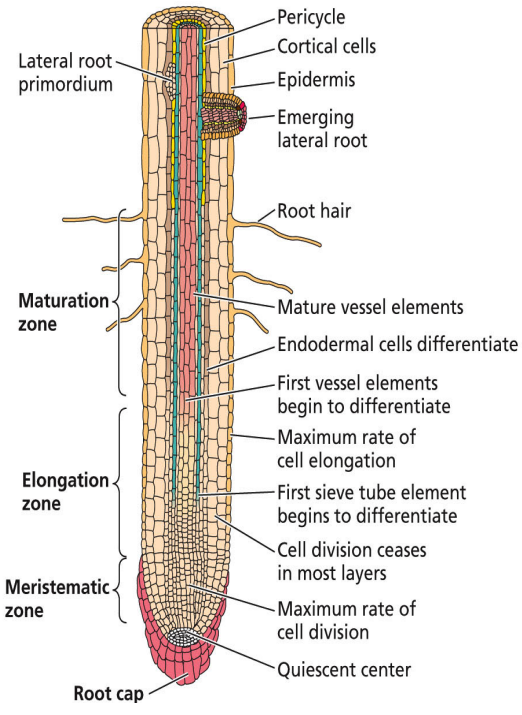


Figure 1.7 Arabidopsis primary root

Diagram of a primary root showing the root cap, meristematic, elongation and mature zones. The root cap protects the apical meristem from potential mechanical injury. Root cap stem cells are produced in this region and are responsible for generating new cells; older cells are gradually displaced and ultimately slough off at the root tip. The meristematic zone generates the primary root. The elongation zone is the site of rapid and extensive cell elongation. After division and elongation have ceased, cells enter the maturation zone where they acquire differentiated characteristics. Figure taken from Zeiger (2006).

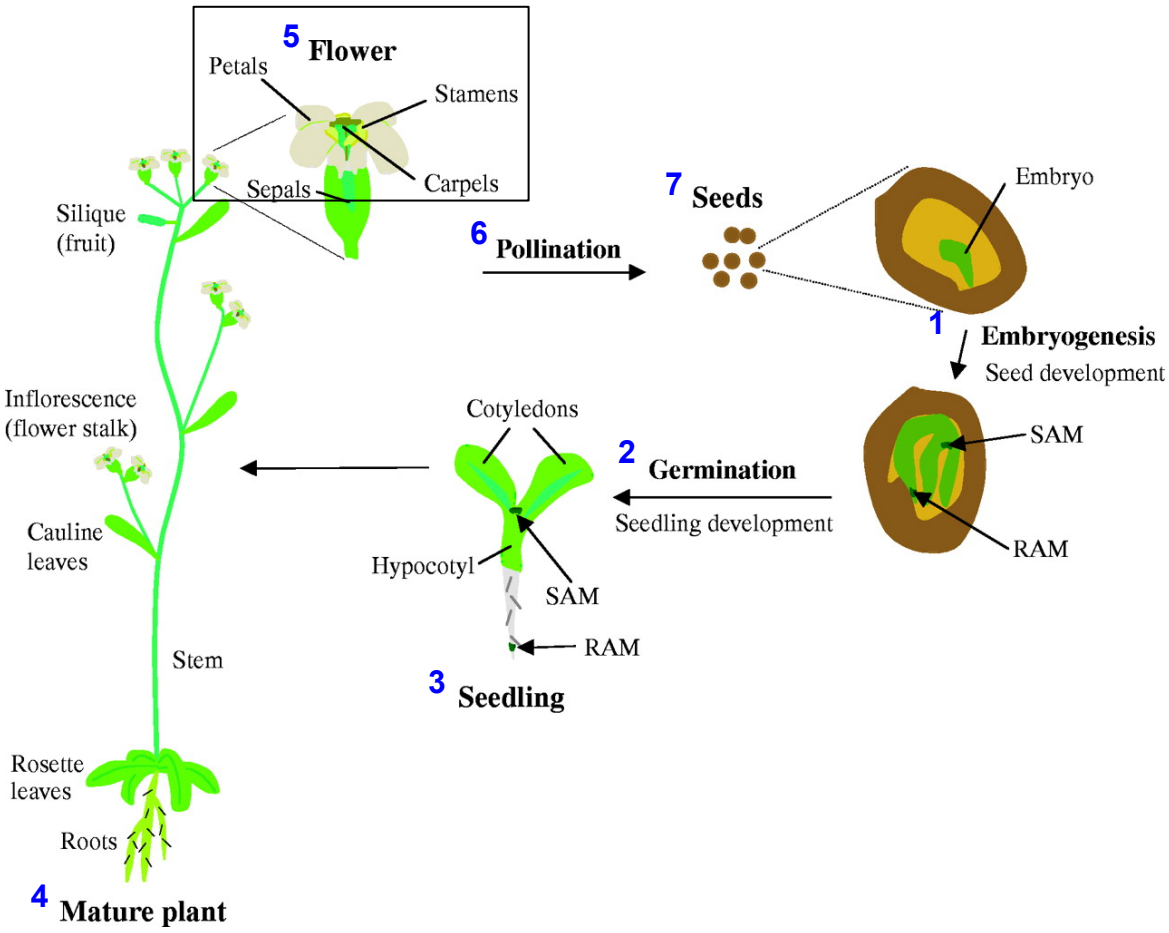


Figure 1.8 The Arabidopsis life cycle

Simplified diagram of the Arabidopsis life cycle. The mature plant has roots, rosette and cauline leaves; stems (only one shown) and inflorescences where flowers develop. Flowers are composed of sepals, petals, stamens (male reproductive organs) and carpels (female reproductive organs). (**1**) Embryogenesis occurs giving rise to an embryo having two meristems: SAM and RAM. (**2**) Seed germination produces a seedling with embryonic hypocotyl and cotyledons (**3**). (**4**) Vegetative development occurs and a mature plant develops. (**5**) Reproductive development follows with the formation of flowers; the reproductive organ of Arabidopsis. (**6**) After pollination the fertilised egg develops into an embryo inside the silique (**7**) After ripening the seed can be collected/dispersed and the life cycle of Arabidopsis can begin once again. Figure modified from Diévart and Clark (2004).

weeks post-fertilisation. The seeds can be collected for planting or in a natural environment are dispersed to commence the life cycle once more (Figure 1.8) (Zeiger, 2006).

1.5.4 *C. elegans* as an animal model organism

C. elegans is a nematode (roundworm) approximately 1 mm in length (Figure 1.9) (Brenner, 1974). They have a small genome size of 97 megabases (Mb) (The *C. elegans* Sequencing Consortium, 1998), a short life cycle (Byerly et al., 1976), are optically transparent and are easy to grow and maintain in the laboratory. *C. elegans* have become a powerful tool for studies in genetics (Fire et al., 1998), neuroscience (Chalfie et al., 1994, Kaletta and Hengartner, 2006, Kosinski and Zaremba, 2007) and toxicology (Leung et al., 2008).

The transparency of *C. elegans* (Figure 1.9 B) renders it useful for studying expression patterns of membrane proteins through genetic bioengineering (Bamps and Hope, 2008) using fluorescent proteins (Chalfie et al., 1994, Shaner et al., 2004, Chudakov et al., 2010). Gene expression reporter analysis can be used alongside phenotypic data to determine potential physiological functions for genes. For example, the *C. elegans* gene *pst-1* encodes for a 3′-phosphoadenosine 5′-phosphosulfate (PAPS) transporter, PST-1. The deletion mutant show phenotypic defects in cuticle formation, vulval morphogenesis, cell migration and embryogenesis; gene expression results showed this protein localised to the Golgi membrane in the hypodermis. Together these data revealed that PST-1 is an essential transporter for *C. elegans* development (Dejima et al., 2010). Localisation studies are therefore important to further our understanding regarding gene function.

1.5.4.1 The life cycle of *C. elegans* and embryogenesis

The *C. elegans* life cycle is very short; hatching from an egg through to adult development occurs in a mere 3 – 4 d, depending on temperature (Figure 1.11). There are two *C. elegans* sexes, the hermaphoridite and male, the former are introduced here (males are discussed in section 6.1.2). The life cycle of *C. elegans* begins with embryogenesis (described in Figure 1.10), and ~14 h after being laid the egg hatches (Figure 1.10) (von Ehrenstein G., 1980, Sulston et al., 1983, Bird, 1991). Subsequently, four larval stages (L1 – L4) ensue before development into an adult (Figure 1.11). A

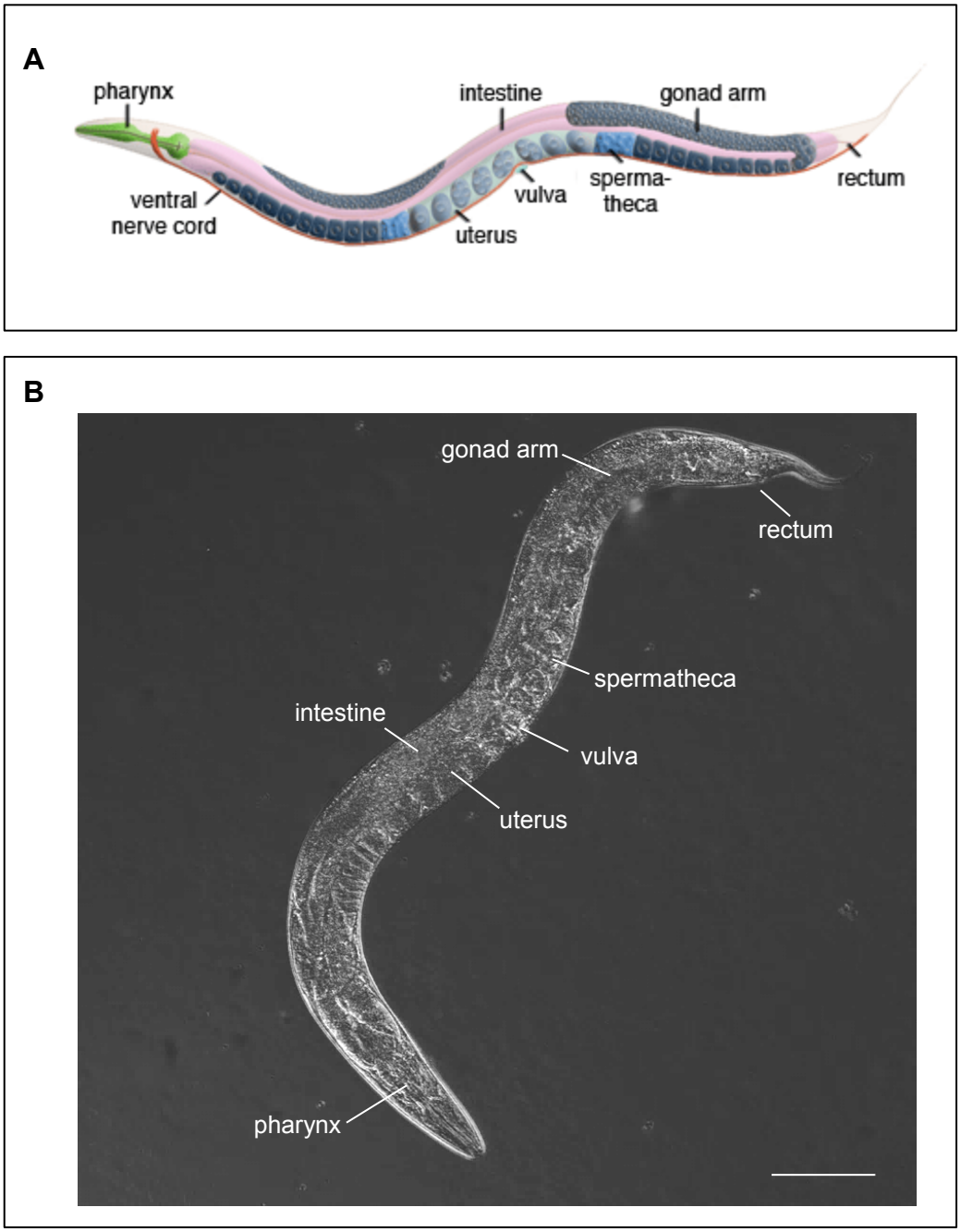


Figure 1.9 *C. elegans* adult worm anatomy

A. Diagram showing the anatomy of the adult *C. elegans* worm. Figure taken from WormAtlas (Altun, 2006). **B**. DIC image of the *C. elegans* adult worm. Scale bar: 200 µm.

molt characterises the end of each larval stage; the old cuticle is shed and a new, stage-specific cuticle containing different protein composition and ultrastructure is synthesised (Byerly et al., 1976, Wood, 1987). Molting involves: apolysis, old cuticle separation from the hypodermis; new cuticle formation; ecdysis, old cuticle shedding. Prior to apolysis, the animal enters a brief lethargus; locomotion stops and the cuticle is loosened from the lips, buccal cavity and around the tail (Bird, 1991). Before ecdysis, the pharyngeal cuticle lining breaks down into the intestine and the larva pushes against the old cuticle to make a hole at the head for it to emerge through (Byerly et al., 1976, Altun, 2006).

1.5.4.2 The *C. elegans* pharynx is required for feeding

C. elegans sample their environment and forage for bacterial food. The pharynx is the organ that *C. elegans* uses to accomplish this. The pharynx is myogenic and often considered as the "heart" of the worm. This organ provides the means for *C. elegans* feeding: bacterial food uptake, concentration and grinding, then transportation to the intestine through pumping and isthmus peristalsis (Doncaster, 1962, Seymour et al., 1983). There are 20 muscle cells in the pharynx forming eight muscle layers (Franks et al., 2006) making up the three fundamental components of the pharynx; these are the corpus, isthmus and terminal bulb (Figure 1.12) (Avery and Shtonda, 2003). The muscle cells are radially orientated forming a tri-radiate symmetry; upon muscle contraction, this serves to open the lumen into a triangular cross-section (Figure 1.12 A).

Pumping is the action that gets the food into the worm by cycles of contraction and relaxation to suck liquid from the environment along with suspended particles, followed by expelling the liquid to trap the particles. A near-simultaneous contraction of the corpus, anterior isthmus and terminal bulb muscles begins the cycle of a pump. Contraction of the muscles pulls the lumen open; this sucks in liquid and suspended bacteria from the outside. Following this is a near-simultaneous relaxation, resulting in the expulsion of liquid but trapping bacteria. In this way the pharynx concentrates bacteria and therefore the food can be stored in the intestine more efficiently,aiding nutrient uptake (Avery and Shtonda, 2003).

With such a short life cycle, *C. elegans* culture, including crossing to generate mutants, requires much less time than other animal models such as mice. The larval molts can be distinguished relatively easily, especially L4-stage animals, which is

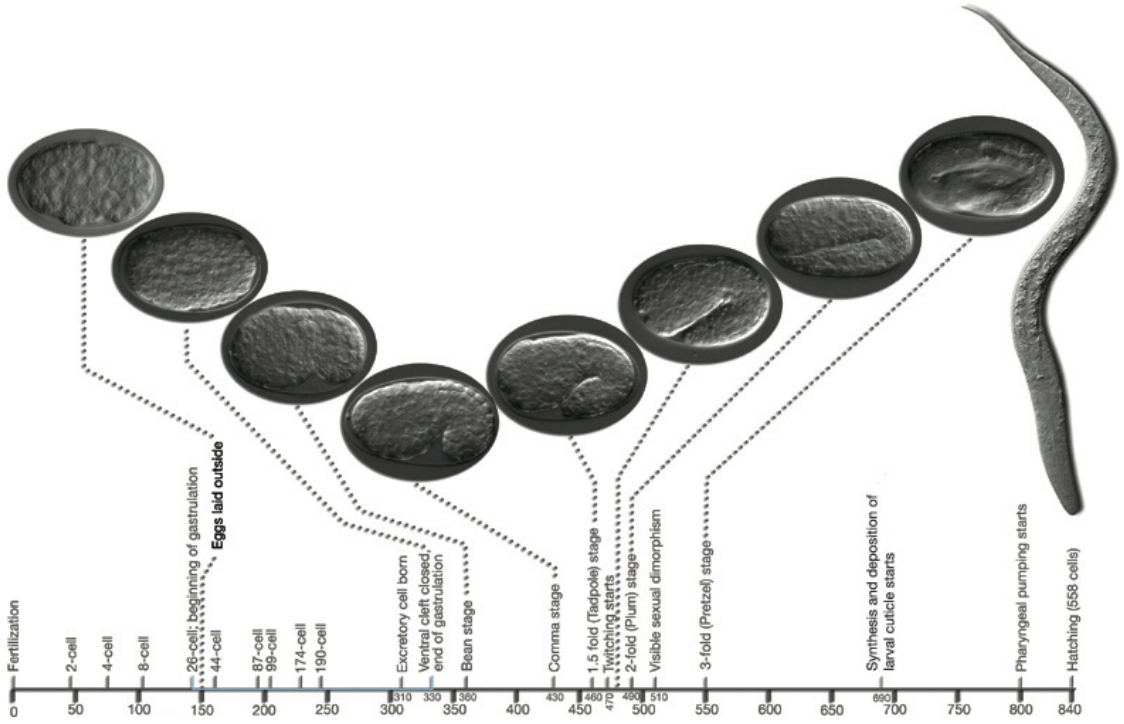


Figure 1.10 Embryonic stages of egg development

Embryonic stages of development. The numbers below the horizontal axis show approximate time in minutes after fertilisation at 22 ºC. The stages, marker events/stage names and DIC images of the embryos and a newly hatched larva is shown above the horizontal axis.

Embryogenesis begins in the uterus when an oocyte is fertilised in the hermaphrodite. Subsequently a series of developmental stages proceeds, categorised into two main stages: proliferation and organogenesis/morphogenesis (Sulston et al., 1983). At 22 °C, the proliferation stage occurs between 0 to 330 – 350 min post-fertilisation (Sulston et al., 1983) and involves cell division from a single cell to 558 undifferentiated cells (von Ehrenstein G., 1980, B., 1988b). The initial 150 min of proliferation occurs in the uterus, when the embryo reaches the 30-cell stage (gastrulation) the egg is laid and the rest of embryogenesis occurs outside the worm (Figure 1.10). The end of proliferation is marked by the formation of the spheroid of cells organised into three germ layers: ectoderm, mesoderm and endoderm, which gives rise to the hypodermis and neurons; pharynx and muscle; and the germline and intestine respectively. Organogenesis (morphogenesis) makes up the rest of the embryogenesis time. The embryo elongates three-fold and an animal takes form with fully differentiated tissues and organs. The lima bean stage marks the start of morphogenesis and first muscle twitches can be seen ~7 h (430 min) post first-cell cleavage (Figure 1.10). In the late three-fold stage, the larva can roll around inside the egg, indicative of advanced motor system development. The egg hatches ~14 h after being laid (von Ehrenstein G., 1980, Sulston et al., 1983, Bird, 1991). Figure taken from WormAtlas (Altun, 2006).

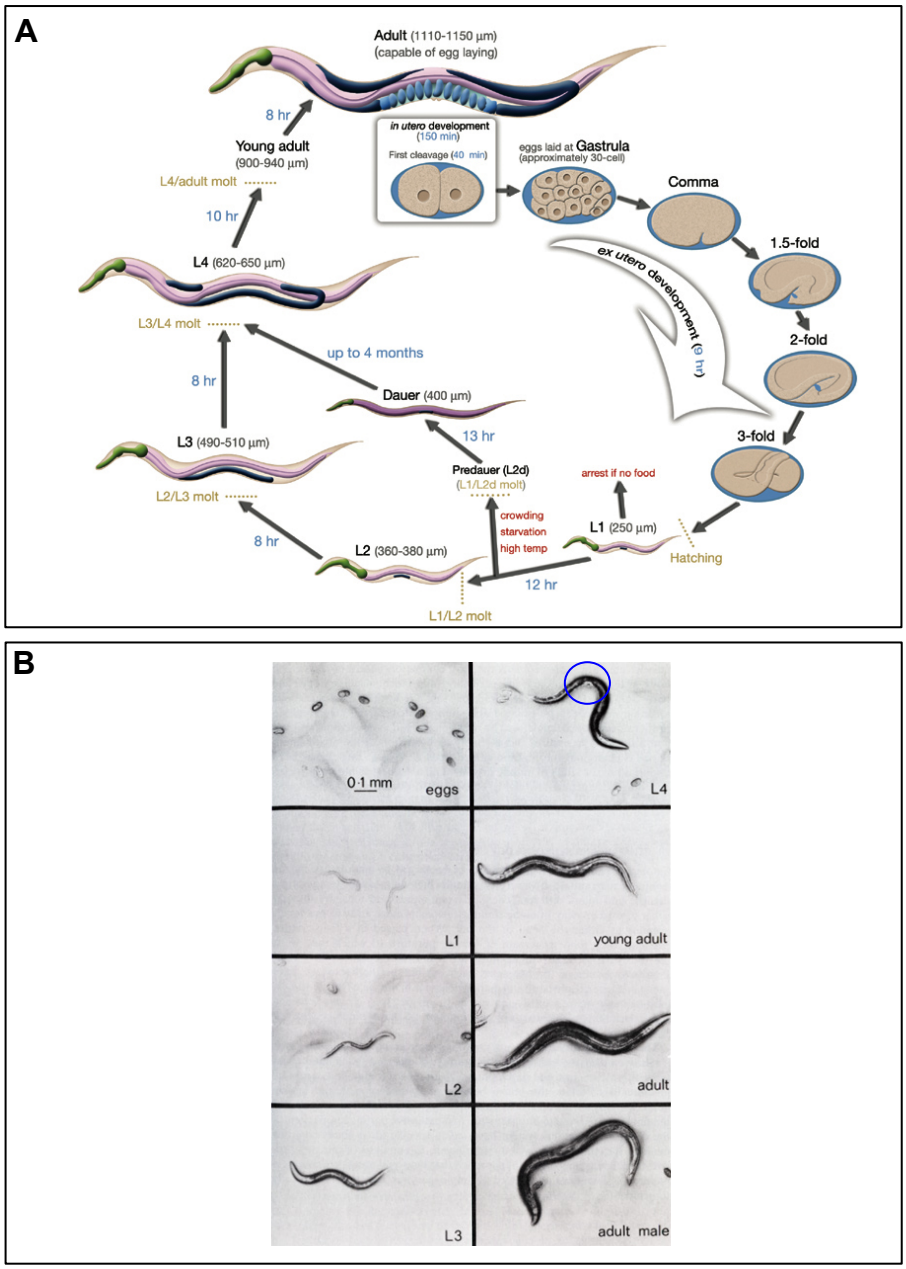
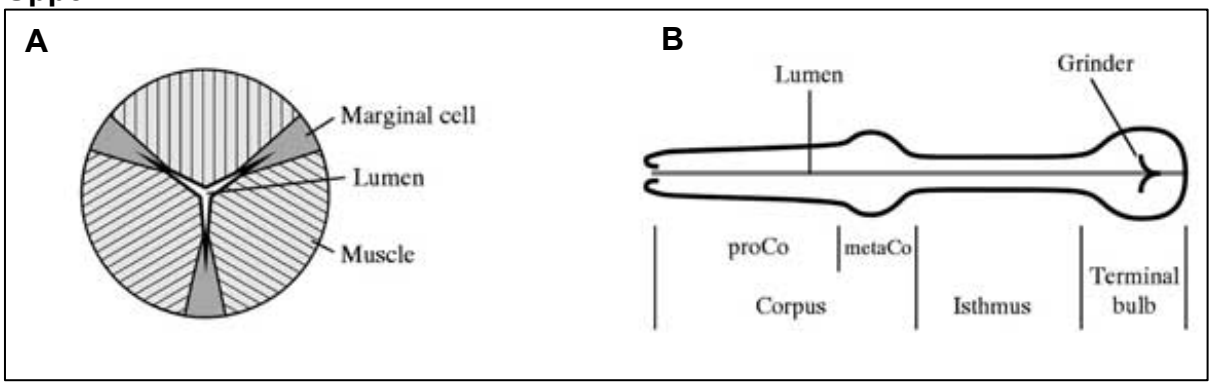


Figure 1.11 The life cycle of *C. elegans* at 22 ºC

A. The life cycle of *C. elegans* from fertilisation (0 min) to development into an adult. Numbers (blue) along the arrows show the length of time the animal spends at each stage. The length of the animal at each stage is shown in brackets next to the developmental stage name (beige). First cleavage occurs *in utero* at ~40 min post-fertilisation. At ~150 min post-fertilisation, the egg is laid outside (*ex utero*). On food, *C. elegans* undergoes embryogenesis (Figure 1.10), followed by L1 – L4 molts before reaching adulthood. Under crowded, starved or high temperature conditions, prior to the L1/L2 molt they enter a dauer phase where they can live up to 4 months. If *C. elegans* hatch in the absence of food, L1 larva will arrest at L1-stage. Figure taken from WormAtlas (Altun, 2006). **B**. Bright field images to show various stages of *C. elegans*. L4-staged worms are distinguishable by a white crescent at the vulva site (circled). (Images from George L. Sutphin and Matt Kaeberlein at University of Washington, Seattle).

Upper



Lower

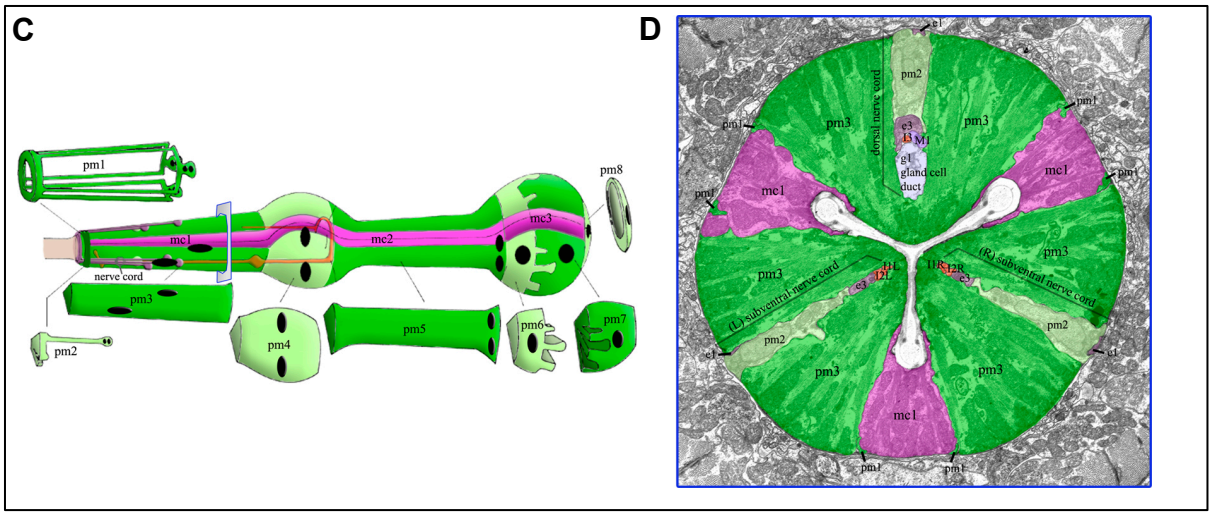


Figure 1.12 The anatomy of the pharynx

Upper: **A**. Simplified generic cross-section through the pharynx to illustrate tri-radiate symmetry. Shown are the geometric relationships between muscle cells, marginal cells and pharyngeal lumen. **B**. Side-view of the pharynx. The three fundamental components of the pharynx: corpus, isthmus and terminal bulb. Further subsections of the corpus are also labelled: pro- and metacorpus. **Lower**: **C**. Side-view of the 8 **p**haryngeal **m**uscle layers (pm 1 – 8, green) separated by structural **m**arginal **c**ells (mc 1 – 8, pink) in the pharynx. **D**. Cross-section through the pharynx to further illustrate the tri-radiate symmetry between the different cell types and layers. Figures taken from WormBook (Mango, 2007).

characterised by a white crescent shape at its vulva site (Figure 1.11 B). Egg development is well-characterised, thus this process can also be followed as part of phenotype analysis, for example. Their amenability in the laboratory and other well-studied behavioural aspects renders it a practical animal model for reverse genetic studies (section 6.1.3).

1.5.5 *S. cerevisiae* as a fungal model organism

S. cerevisiae is a unicellular organism and as a eukaryote it has a similar cell structure to animals and plants (Kurtzman, 2005). With a fully sequenced genome comprising 16 chromosomes (Goffeau et al., 1996), a wide range of resources available and its ease of manipulation and culture, it is commonly used as a model system for functional analysis of genes (Qi et al., 2013), protein-protein interactions (Botstein et al., 1997, Schwikowski et al., 2000, Kharenko et al., 2013) or for studying biological processes such as ageing (Kaeberlein et al., 1999, Guarente and Kenyon, 2000, Piper, 2006, Fontana et al., 2010). *S. cerevisiae* reproduces approximately every 2 h at 30 °C so it can be easily cultured and maintained (Botstein et al., 1997, Zeyl, 2000, Boekhout and Robert, 2003, Karathia et al., 2011). There are many processes that are conserved throughout eukaryotes, for example ER molecular processes such as *N*-glycosylation (Aebi, 2013). Therefore simpler eurkaryotic organisms can be used to study complex processes of higher organisms. Since *S. cerevisiae* is a simple organism it can be employed for recombinant expression of genes. In this way, biological processes can be studied using a simpler eukaryotic system such as *S. cerevisiae*. For instance, *S. cerevisiae* was the first organism used to identify the importance of the COPII proteins in formation and budding of ER-derived vesicles (Novick et al., 1980, Novick et al., 1981) and used as a model organism for determining the functionality of the Arabidopsis COPII proteins. This was shown by the Arabidopsis COPII proteins rescuing the thermo-sensitivity and secretion phenotypes in a temperature-sensitive *S. cerevisiae* mutant (De Craene et al., 2014). Therefore *S. cerevisiae* serves as an excellent expression system for advancing our insight into biological systems.

1.6 Aims of this thesis

The GTG/GPHR family of membrane proteins are found in all eukaryotes, but is there a conserved function? So far, the GTG/GPHRs have been studied in Arabidopsis (Pandey et al., 2009, Jaffé et al., 2012, Alvarez et al., 2013), a CHO cell line (Maeda et al., 2008), mice (Tarutani et al., 2012) and, very recently, *Drosophila* (Charroux and Royet, 2014). There are two fundamental roles proposed for these proteins but as yet there is no study to compare and critically analyse the data supporting either hypothesis: anion channels that regulate Golgi pH (Maeda et al., 2008) or plasma membrane ABA receptors (Pandey et al., 2009). Furthermore, there are differing reports for the Arabidopsis *gtg1 gtg2* double mutant responses to ABA as well as their localisation in plants, Therefore, a number of aims have been set out to resolve the controversy related to the GTG/GPHR function:

1. To conduct a comprehensive computational analysis of the GTG/GPHRs. This information could reveal any conserved features of these proteins and serve as a platform for experimental design.
2. To elucidate the GTG function in plants by:
    a. Further characterising Arabidopsis *gtg1 gtg2* mutants and undertaking a phenotypic comparison of all three available double mutants (*gtg1-1 gtg2-1*, *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3*).
    b. Initiating the study of the GTGs in the monocotyledon (monocot) *Oryza sativa* (*O. sativa*).
3. To investigate whether there is a conserved function for the GTG/GPHRs. Orthologue GTG/GPHRs will be expressed in the Arabidopsis *gtg1 gtg2* double mutant, which has well-characterised phenotypes that can be rescued by At GTG1 or At GTG2 expression. Restoration of mutant phenotypes by an orthologue GTG/GPHR would demonstrate conservation of function.
4. To further investigate the roles of the GTG/GPHRs in eukaryotes:
    a. The first whole animal *gphr* KO model will be generated in *C. elegans* for characterisation.
    b. An *S. cerevisiae gtg* mutant will be isolated for phenotypic analysis.
5. The localisation of the GTG/GPHRs will be investigated using three organisms (Arabidopsis, *C. elegans* and *S. cerevisiae*). The results could either support or

reject whether the GTG/GPHRs are plasma membrane localised or Golgi and ER localised.

All these aims will address whether there is a conserved function for the GTG/GPHR family of membrane proteins and elucidate a role in ABA receptor signalling or Golgi pH regulation.

CHAPTER 2

Material and Methods

2.1 Species-specific nomenclature

For the purpose of consistency in this thesis, when referring to the *GTG/GPHR* genes collectively i.e. not species-specific the following formats will be applied: *GTG/GPHR* genes, GTG/GPHR proteins and *gtg/gphr* mutants. Otherwise, in this thesis the species-specific convention will be applied as outlined below.

Species symbols are two-lettered prefixes to genes/proteins with first letter in uppercase only, followed by a space and will be used when required to distinguish between orthologues (Table 2.1). For Arabidopsis the gene symbols are italicised in uppercase, protein symbols are uppercase and are not italicised, and mutant symbols are in all lowercase and italicised (Table 2.1). All Arabidopsis nomenclature used follows the Arabidopsis Biological Resource Center (ABRC: http://abrc.osu.edu/) guidelines found in TAIR (http://www.arabidopsis.org/). For *C. elegans* the gene symbols are three- or four-lettered and italicised in lowercase, protein symbols are uppercase and are not italicised, and mutant symbols are specific for the strain and are shown as the gene name with the mutant allele in brackets. WT alleles are indicated by a plus symbol in brackets (Table 2.1). Heterozygotes, with allelic differences between chromosomes, are designated by separating mutations on the two homologous chromosomes with a slash. New animal lines i.e. *C. elegans* transformed with expression vectors are named according to whether it is an extrachromosomal array or integrated into the genome. Extrachromosomal arrays or integrated constructs are italicised names of the allele prefix, the two letters *Ex* or *Is,* respectively, and a number, followed by the genotypic information of the construct transformed in square brackets e.g. N2*Ex*1[*Pmyo-2::Cel-gphr-1::egfp*]. All *C. elegans* nomenclature is according to Horvitz et al. (1979). For *S. cerevisiae* the gene symbols are three-lettered and italicised (lowercase for recessive and upper for dominant). The WT or mutant allele is indicated by suffix superscript + or – signs to the gene symbol. Since we do not know whether the GTG^{-} mutant is a recessive or dominant allele, the gene will be denoted in uppercase, *GTG*. Protein symbols have their first letter in uppercase only and with the suffix “p” non-italicised (Table 2.1). Alleles created by recombinant DNA technology are named by use of the

symbol for the gene that is altered followed by a double colon (::) for disruption, -Δ for deletion or Δ:: for a replacement. E.g. *gtg*-Δ::*kanMX4* is the replacement of the *GTG* ORF by the geneticin/kanamycin resistance gene. Genotypes of strains typically have the mating-type loci listed first (either *MAT a* or *MAT* α). If the cell is haploid, just one copy of each gene is listed, if the cell is diploid then both copies of each gene are listed, separated by a slash. Non-mendelian genotypes i.e. strains transformed by plasmids are distinguished by square brackets. All *S. cerevisiae* nomenclature is outlined according to the SGD (Chervitz et al., 1999).

For constructs using Arabidopsis genes a single colon indicates gene fusions; lowercase "p" is indicative for plasmids, e.g. pMDC32; promoters are designated with uppercase "P", e.g. *P35S*; the name may be followed by a reporter, all italicised, e.g. *P35S:At GTG1:GFP*. For constructs using *C. elegans* genes lowercase "p" is indicative for plasmids e.g. pLT61; promoters are designated with uppercase "P" e.g. *Pmyo-2;* the gene name is followed by two colons and the reporter, all italicised. All constructs use the nomenclature specific to the organism of the gene inserted, e.g. *C. elegans* gene inserted into a plant expression vector will follow the *C. elegans* construct nomenclature.

2.2 Plant growth conditions

2.2.1 Light treatments

For all physiological treatments, unless otherwise stated, all mature Arabidopsis plants (section 2.2.2) were grown in long-day white light (WL) conditions: 16 h light/8 h dark; light intensity 100 – 120 $\mu mol.m^{-2}.s^{-1}$, 23/18 °C, 60/55 % humidity provided by environmentally controlled rooms (ECRs), Fitotron Plant Growth Chambers (Weiss Gallenkamp, UK). Where stated, Arabidopsis plants (section 2.2.2) were also grown in short-day conditions: 8 h light/16 h dark; light intensity 100 – 120 $\mu mol.m^{-2}.sec^{-2}$, 23/18 °C, 60/55 % humidity. *O. sativa* plants were grown in long-day WL conditions: 16 h light/8 h dark; light intensity 100 – 120 $\mu mol.m^{-2}.sec^{-2}$, 28/28 °C, 80/85 % humidity. All Arabidopsis seedlings (section 2.2.3) were grown under long-day WL treatment conditions: 16 h light /8 h dark; light intensity 120 $\mu mol.m^{-2}.sec^{-2}$, 22/18 °C provided by plant growth cabinets (Percival Scientific Inc., Boone, IA, USA); unless otherwise

Table 2.1 Species-specific symbols for the *GTG/GPHR* genes

The studied GTG/GPHRs are listed in this table. Available mutants for these genes are also listed according to species-specific mutant symbols. For example, *Drosophila* mutants are referred to according to specific mutations such as Dm $GPHR^{P91L}$, while Arabidopsis mutant genes are referred as e.g. At *gtg1* (or specific alleles e.g. *gtg1-1*).

Gene ID	Species	Prefix symbol	Example species-specific symbols		
			Gene	Protein	Mutant(s)
AT1G64990	*Arabidopsis thaliana*	At	At *GTG1*	At GTG1	At *gtg1*
AT4G27630	*Arabidopsis thaliana*	At	At *GTG2*	At GTG2	At *gtg2*
Y75B8A.16	*Caenorhabditis elegans*	*Cel-*	*Cel-gphr-1*	*Cel*-GPHR-1	*Cel-gphr-1(ok1579)*
C11H1.2	*Caenorhabditis elegans*	*Cel-*	*Cel-gphr-2*	*Cel*-GPHR-2	*Cel-gphr-2(tm4228)*
CG8090	*Drosophila melanogaster*	Dm	Dm *GPHR*	Dm GPHR	Dm $GPHR^{P91L*}$ or Dm $GPHR^{LL03647}$
100689387	*Cricetulus griseus*	Cg	Cg *GPHR*	Cg GPHR	#
HGNC:13840	*Homo sapiens*	Hs	Hs *GPHR*	Hs GPHR	#
MGI:1914799	*Mus musculus*	Mm	Mm *GPHR*	Mm GPHR	K5-Cre:$GPHR^{f/f}$
YHR078W	*Saccharomyces cerevisiae*	Sc	Sc *GTG*	Sc Gtgp	Sc *gtg*-Δ

*originally denoted $dGPHR^{k34}$ by Charroux and Royet (2014) but renamed Dm $GPHR^{P91L}$ in this thesis as this mutant contains a point-mutation changing proline-91 to leucine.

#No mutant has been generated for the *H. sapiens GPHR* gene; a *GPHR* mutant CHO cell line has been created (C27) but not a whole animal or tissue-specific KO.

stated i.e. hypocotyl assays (section 2.3.1). The broad-band WL was provided by fluorescent tubes (MASTER TL-D Reflex 58W/840 1SL, Philips, NL).

2.2.2 Growth of Arabidopsis and *O. sativa* plants

For growth of mature Arabidopsis and *O. sativa* plants, the soil composition was made to a 1: 1: 1 ratio of Levington (Green-Tech, UK), John Innes No. 2 (John Innes Manufacturers Association, UK) and Sinclair Vermiculite Medium Grade (Scot Plants Direct, UK). Intercept-5 GR insecticide (Bayer Environmental Sciences SAS, UK) was added to a final concentration of 0.28 g/L of soil.

To grow Arabidopsis, four seeds were sown onto moist soil contained in 7 x 7 x 8 cm DESCH7B square black pots (Desch, UK). The pots were covered in clear film and placed in ECRs (section 2.2.1). Following germination and the emergence of cotyledons, the film was removed and three seedlings were removed to leave one seedling per pot. As the first inflorescence appeared the plants were bagged in 60 x 30 x 10 cm clear micropunched flower sleeves (Geerings, Kent, UK) to prevent cross-pollination. Siliques were harvested after approximately 2 – 3 months and transferred to 30 mL universal containers (Sterillin, UK), with the lid loosely screwed to ventilate the seeds allowing them to dry completely. For growth of Arabidopsis plants for transformation (section 2.3.3), the first flowering inflorescence was snipped to the rosette level to induce secondary bolting for increased transformation efficiency.

To grow *O.* sativa, seeds were heat-treated at 50 °C for 3 d. To germinate *O. sativa* seeds, the outer coat (pericarp) was removed and seeds were imbibed on moist blue roll in a petri dish for ~1 week in the ECR under conditions for growth of *O. sativa* outlined in section 2.2.1. Germinated seedlings were transplanted onto moist soil contained in 11 x 11 round pots (Desch, UK).

2.2.3 Growth of Arabidopsis seedlings

Arabidopsis seeds were sterilised using 15 % bleach for 15 min, washed with sterile ddH_2O before being sown individually onto plates containing 0.5 Murashige and Skoog (MS) medium (specified for different experiments in section 2.3). Plates were sealed using 1.25 cm 3M™ Micropore™ (3M, UK), wrapped in aluminium foil and the

plated seeds were stratified at 4 ˚C for 48 h prior to transfer to a plant growth cabinet (section 2.2.1).

2.3 Arabidopsis phenotype assays

2.3.1 Seedling root and hypocotyl growth analysis assays

For root experiments, seeds were sterilised and individually sown onto 0.5 x MS media containing 0.8 % (w/v) agar (section 2.2.3). Any additional supplements are specified in figure legends e.g. 1 % (w/v) sucrose. Unless otherwise stated all reagents used were analytical grade in this project. After stratification, seeds were transferred to a plant growth cabinet (section 2.2.1) and grown vertically for 14 d. Photographs of the plates (from the reverse side) were taken at 7, 10, 12 and 14 d to monitor root growth, except at 14 d where an open-plate photo was also taken. At 14 d the fresh weight of the seedlings was measured. For hypocotyl experiments, seeds were sterilised and individually sown onto 0.5 MS media containing 0.8 % (w/v) agar (section 2.2.3). Any additional supplments are specified in figure legends e.g. 1 % (w/v) sucrose. These were transferred to a controlled-environment cabinet (WL 100 – 120 $\mu mol.m^{-2}.s^{-1}$) for 2 h prior to horizontal growth in low WL (8 – 10 $\mu mol.m^{-2}.s^{-1}$) using a neutral density filter 211/210 (Lee Filters, Hampshire, UK) for 5 d under 16 h low WL/8 h dark cycles at 22 ˚C/18 ˚C conditions. To analyse hypocotyl cell growth, the seedlings were placed onto wet-mounted glass slides and covered with a coverslip and visualised using light microscopy. Images were made using a dissecting microscope and digital images taken using a digital camera. For all root and hypocotyl length analysis, photographic images were taken and ImageJ software (http://rsbweb.nih.gov/ij/) was used to measure the seedling root and hypocotyl lengths from photographic images.

2.3.2 Measuring mature plant phenotypes

For measuring rosette growth, Arabidopsis seeds were grown (section 2.2.2) under short-day conditions (section 2.2.1). Each week the rosette diameter was measured and the pots were randomised. For measurement of flowering time, seed yield, seed number and seed per silique, Arabidopsis seeds were grown under long-day conditions (section 2.2.1). The flowering time of the first inflorescence was measured as the number of days from planting was calculated using this date. Plants were left to grow and measurements of the siliques when the plant reached full maturity i.e. plant

beginning to brown (seed ripening) but with plenty of green siliques but few flowers left. One bolt was taken from each plant using scissors and photographed using a digital camera. The siliques were measured from the ninth silique down (inclusive) using ImageJ (http://rsbweb.nih.gov/ij/). To determine seed number, siliques were removed from the bolt and transferred to 1.5 mL eppendorf tubes (eppendorfs from herein), these were immersed in 70 % ethanol (v/v) overnight and then 1 mL of 100 % methyl salicylate (Sigma, Missouri, USA) also overnight. The siliques were photographed for seed counting. For seed yield measurements, the plants were matured and seed left to ripen. The seed was collected (section 2.2.2), and dried for at least 1 week before the weight was measured.

2.3.3 Arabidopsis transformation using the floral dip method

To transform Arabidopsis, the construct of interest was first transformed into *Agrobacterium tumefaciens* (*Agrobacterium*). One µL of the expression clones generated in section 2.7.15 was transferred to 50 µL of electrocompetent *Agrobacterium* GV3850 cells (section 2.7.1). After gentle mixing, the cells were transferred to precleaned and prechilled 0.2 cm electroporation cuvettes (Bio-Rad Laboratories, CA, USA) and pulsed at 1.8 V for 5 ms using the Biorad MicroPulser ^{TM} Electroporation Apparatus (Bio-Rad, CA, USA). One mL of Luria-Bertani (LB) medium (1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1 % (w/v) NaCl) was added to the cuvette and the bacterial suspension was transferred to a chilled 15 mL culture tube. The culture was incubated at 30 °C with vigorous agitation (250 rpm) in a Innova ^{TM} 4300 Incubator Shaker (New Brunswick Scientific Ltd., UK) for 2 h before being spread on freshly prepared LB agar (LB medium with 1.5 % (w/v) agar) plates containing 50 µM carbenicillin, 50 µM kanamycin and 50 µM rifampicin; these were cultured at 30 °C for 3 d. Colony PCR was performed on a selection of colonies (section 2.7.11).

Arabidopsis plants were transformed with constructs of interest (Table 4.3 and Table 5.3) using the floral dip method adapted from Clough and Bent (1998). After carrying out colony PCR and obtaining positive PCR results for *Agrobacterium* containing the plasmid of interest, a single colony was and inoculated in 5 mL LB medium containing 50 µM carbenicillin, 50 µM kanamycin and 50 µM rifampicin and incubated at 30 °C with gentle agitation (190 rpm) overnight. The small cultures were then poured into 500 mL of LB medium containing 50 µM carbenicillin, 50 µM

kanamycin and 50 μM rifampicin and incubated at 30 °C with vigorous agitation overnight. Acetosyringone (3′,5′-dimethoxy-4′-hydroxyacetophenone) was added to the cultures to a final concentration of 100 μM and incubated for a further 3 h. Cultures were then centrifuged at 3700 *g* for 15 min. The supernatant was removed and the cells were resuspended in 250 mL of ddH_2O containing 5 % (w/v) sucrose. Silwet L-77 (van Meeuwen Chemicals BV, NL) was added to the *Agrobacterium* to a final concentration of 0.05 % (v/v) prior to dipping the plant material. The aerial parts of approximately 6-week old plants (T_0) that had had their first inflorescence clipped (section 2.2.2) to promote secondary bolts were dipped in *Agrobacterium* culture for 30 s with gentle agitation (Figure 2.1 step 1). The dipped plants were then covered with a plastic bag and sealed with tape to maintain high humidity and placed in dimmed lighting for 24 h, subsequently these were placed into long-day conditions (section 2.2.1). The seeds were collected when ripened.

To select for positive transformants, the T_1 seed from these transformed plants were plated onto 0.5 MS media containing 1 % (w/v) sucrose and 1 % (w/v) agarose with antibiotic selection: 50 μM hygromycin and 200 μM augmentin (amoxicillin sodium and potassium clavulanate; Melford Laboratories Ltd., UK) (Figure 2.1 step 3). Plates were initially incubated at 4 °C for 48 h in the dark and then placed in long-day conditions (section 2.2.1). Positive transformants (indicating indpenedent transformation events) became visible after ~2 weeks. Positive transformants were transplanted onto soil to grow into mature T_1 plants (Figure 2.1 step 4). Seeds collected from the T_2 plant were then plated on 0.5 MS media containing 1 % (w/v) sucrose, 1 % (w/v) agar and 50 μM hygromycin, and the survival ratio noted (Figure 2.1 steps 5 – 6). A selection (8 – 12) of seedlings from positive transformants with a survival ratio of ~75 % (±5 %) were transferred to soil and grown to maturity (Figure 2.1 step 7). Seeds (T_3) collected from the T_2 plant were then plated on 0.5 MS media containing 1 % (w/v) sucrose, 1 % (w/v) agar and 50 μM hygromycin, and the survival ratio noted. Any T_3 seed lines that gave a 100 % survival ratio were considered as homozygous lines for the transformed construct. It would be expected that 1 in 4 of the sister-lines of an independent line would be homozygous for the insert (Figure 2.1 step 9). At least three independent lines for each construct transformed into each plant background was identified for Arabidopsis phenotype assays (section 2.3). Subsequently plants were confirmed as positive transformants by PCR (section 2.7.10).

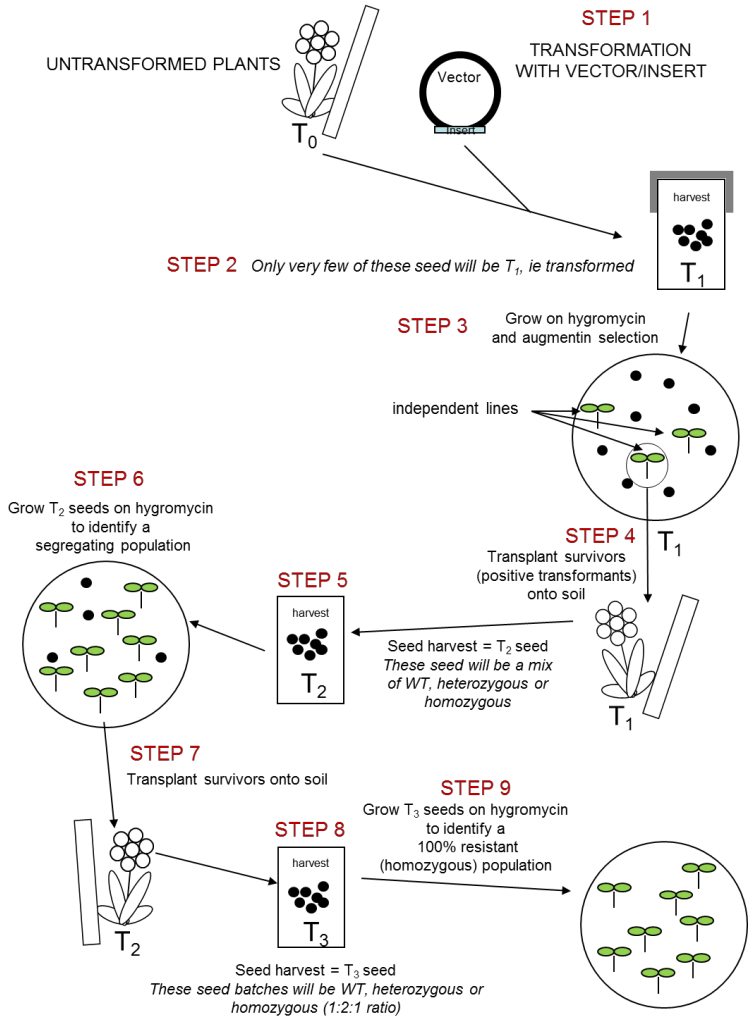


Figure 2.1 Scheme used to generate transgenic Arabidopsis lines

Transform 4 – 6 of each genotype e.g. *gtg1 gtg2* mutant plants with construct (**step 1**). Harvest T_1 seed (**step 2**). All plant destination vectors used here (Table 4.3 and Table 5.3) have hygromycin resistance gene as a positive selection marker. Grow T_1 seed grown on hygromycin and augmentin selection to identify independent transformants (**step 3**). Transplant survivors onto soil (**step 4**). Grow T_1 plants and harvest T_2 seed; all T_2 seed are a mix of WT, heterozygous or homozygous for the insert (**step 5**). Analyse segregation ratios by growing T_2 seeds on hygromycin; ~75 % resistance indicates a single insert into the genome i.e. a segregating line (**step 6**). Transplant 8 – 10 survivors from each segregating line (independent lines) onto soil (**step 7**). Allow T_2 plants to mature for T_3 seed collection; all seed batches are either WT, heterozygous or homozygous for the insert (**step 8**). Identify homozygous seed batches by growing T_3 seed on hygromycin; 100 % resistance is indicative of homozygous for the insert (**step 9**). At least three independent lines for each construct transformed was identified in this way.

2.4 Growth and maintenance of *C. elegans* strains

2.4.1 *C. elegans* culture

C. elegans were grown on 11 mL nematode growth medium (NGM) in 50 mm plates at 20 ˚C as described by Brenner (1974). NGM containing 50 mM NaCl, 2 % (w/v) agar and 0.25 % (w/v) peptone was autoclaved. Once cooled to ~55 ˚C the NGM was supplemented with 1 mM $CaCl_2$, 1 mM $MgSO_4$, 1 mM KH_2PO_4 and 0.1 % (w/v) 5 M cholesterol (in ethanol). The NGM plates were left to dry prior to seeding in the centre of the plate with 50 µL *Escherichia coli* (*E. coli*) OP50 food source. OP50 was streaked and cultured in LB agar plates. Ten mL LB medium inoculates were grown overnight (section 2.7.1) from single OP50 colonies for plate seeding. Seeded plates were left at room temperature for at least 24 h before use. Worm transfer (single or up to ten at once) was conducted using a platinum wire. To maintain *C. elegans* strains, prior to bacterial depletion (due to worm feeding), the area of NGM containing worms was "chunked" by cutting this area out using a sterile scalpel and transferred to a new NGM plate; typically this was carried out once every 3 – 4 d.

2.4.2 *C. elegans* strains and maintenance

C. elegans strains were obtained from the *Caenorhabditis* Genetics Center (CGC), which is funded by the NIH Office of Research Infrastructure Program (P40 OD010440), (http://www.cgc.cbs.umn.edu/) and the National Bioresource Project (NBP) for the Experimental Animal "Nematode *C. elegans*" (http://www.shigen.nig.ac.jp/c.elegans/). Table 2.2 outlines all the strains used and generated (section 2.5.3) for this thesis; the single *gphr-1* and *gphr-2* mutant strains RB1390 and FXO4228 were also crossed to generate a double *gphr-1 gphr-2* mutant but could not be maintained (section 6.2.3). *C. elegans* strains were stored at -80 °C for long-term storage. The population of worms from a plate was frozen when the food source was depleted and a large population of L1/L2 worms was evident. Two mL of M9 buffer (22 mM KH_2PO_4, 42 mM Na_2HPO_4, 85.6 mM NaCl and 1 mM $MgSO_4$) was used to wash off the worms from a plate and these were pipetted into a sterile 30 mL universal container. Two mL freezing buffer (10 M NaCl, 50 mM KH_2PO_4, 3.26 M glycerol, 5.6 mM NaOH and 0.4 % agar (w/v)) was added and the tube was vortexed

Table 2.2 *C. elegans* strains obtained and generated for use in this study

Strain	Genotype	Backcrossed	Source
N2	WT	N/A	CGC[1]
RB1390	*Cel-gphr-1(ok1579)*	0	CGC[1]
FXO4228	*Cel-gphr-2(tm4228)*	0	NRP[2]
RB1390x6	*Cel-gphr-1(ok1579)*	x6	Generated in this project
FXO4228x6	*Cel-gphr-2(tm4228)*	x6	Generated in this project
N2*Ex*1[*Pmyo-2::gphr-1::egfp; Peft-4::mCherry::sp12*]	WT	N/A	Generated in this project
N2*Ex*2[*Pmyo-2::gphr-1::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*1[*Pmyo-2::gphr-2::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*2[*Pmyo-2::gphr-2::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*3[*Pmyo-2::gphr-2::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*1[*Pgphr-2::gphr-2::egfp; Paman-2::aman-2::mCherry*]	WT	N/A	Generated in this project
N2*Ex*2[*Pgphr-2::gphr-2::egfp; Paman-2::aman-2::mCherry*]	WT	N/A	Generated in this project
N2*Ex*1[*Pgphr-1::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*2[*Pgphr-1::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*3[*Pgphr-1::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*4[*Pgphr-1::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*1[*Pgphr-2::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*2[*Pgphr-2::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*3[*Pgphr-2::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project
N2*Ex*4[*Pgphr-2::egfp; Peft-4::mCherry:: sp12*]	WT	N/A	Generated in this project

[1] *Caenorhabditis* Genetics Center, [2] National Bioresource Project

briefly before aliquoting into four sterile cryovials and frozen at -80 °C. A test cryovial was thawed after 24 h to check for viability. If more than ten worms survived successfully, these stocks were moved to long-term storage at -80 °C.

2.5 Isolating and phenotyping *C. elegans gphr* mutants

2.5.1 Generating and propagating male nematode strains

L4 hermaphrodite nematodes were heat shocked for 5.5 h at 31 ˚C in a water bath. These were then propagated at 20 ˚C. A few males were produced in the F_1 generation and in order to maintain a male line these males were crossed with an excess of L4 hermaphrodites of its respective strain. Subsequently, male stocks were maintained by setting up a new cross: selecting 12 young adult males with four L4 hermaphrodites.

2.5.2 Backcrossing *C. elegans gphr* single mutants

Four L4 single mutant homozygous hermaphrodites were mated with 12 male young adult N2 nematodes. After 4 d 12 L4 male young adult cross progeny (heterozygous or hemizgyous) were mated with four L4 N2 hermaphrodites. After 4 d, 16 L4 hermaphrodites were cloned on individual plates (all should be either homozygous, heterozygous or WT; corresponding to a 1: 2: 1 ratio). After a further 4 d, genomic DNA was extracted from these candidate nematodes (section 2.7.6) and used for genotyping by PCR (section 2.7.10) using appropriate primers (Table 2.3). Animals carrying the mutation of interest have been backcrossed twice: twice on the autosomes and once on the X chromosome (specific for the FXO4228 strain, only once as *Cel-gphr-2* gene is on the X chromosome). Twenty-four progeny at the L4 stage were individually mated with four young adult N2 male nematodes. After 4 d, the original candidate was genotyped by PCR as above. From a candidate homozygous plate, the backcross was repeated as detailed from the beginning of this section until six backcrosses were carried out. RB1390 strain was backcrossed six times. FXO4228 was backcrossed 12 times on the autosomes and six times on its X chromosome (where *Cel-gphr-2* is situated). Strains backcrossed six times were given the nomenclature “x6” after the strain name: RB1390x6 and FXO4228x6. This scheme is outlined in Figure 6.6.

2.5.3 Generating a *C. elegans gphr-1 gphr-2* double mutant strain

Twelve male RB1390x6 (*Cel-gphr(ok1579)* mutant) at the young adult stage were picked and crossed with 4 L4 FXO4228x6 hermaphrodite nematodes. After 4 d, ten candidate cross progeny (hermaphrodites) were picked at the L4 stage to individual plates and left to self-fertilise for 4 d. These were genotyped by single worm lysis (section 2.7.6) and PCR (section 2.7.10) to identify a *gphr-1(+/-) gphr-2(+/-)* candidate using appropriate primers (Table 2.3). After identifying a plate whose original hermaphrodite mother was *gphr-1(+/-) gphr-2(+/-)*, 32 individual worms were picked to new individual plates. These candidates were left to self-fertilise for 4 d. The original 32 candidates were genotyped as above to identify *gphr-1(-/-) gphr-2(-/-)* nematodes. If no *gphr-1(-/-) gphr-2(-/-)* candidates were identified, 16 worms were transferred to new individual plates from a *gphr-1(+/-) gphr-2(-/-)* candidate plate and left to self-fertilise for 4 d. These candidates were genotyped to isolate a *gphr-1(-/-) gphr-2(-/-)* mutant strain. This scheme is outlined in Figure 6.8.

Table 2.3 Primers used for genotyping *Cel-gphr-1* and *Cel-gphr-2*

Primer name	Primer sequence
CeGTG1 1F	5′ AATGCAATGAGGAAGGCAAG 3′
CeGTG1 1R	5′ TAGGACGCTCGAAACGAAGT 3′
CeGTG1 2F	5′ ACTTCGCCAAATGTTGAATCG 3′
CeGTG1 2R	5′ TAATCTGTGCCATCAAAAGTGC 3′
CeGTG1 3F	5′ AGTTCGAGGACGCAAAATCC 3′
CeGTG1 3R	5′ GGGTTGCAAATGAATCTCG 3′
CeGTG1 4F	5′ GAGACCCTCGCCCGATTCC 3′
CeGTG1 4R	5′ CTCGTAAATCGACACGAAGCAAGC 3′
CeGTG1 5R	5′ GTCATTGTGATCAGAAGACCTC 3′
CeGTG2 1F	5′ CCATATCAATGCCAACGTAG 3′
CeGTG2 1R	5′ CGGATATTTTGGATCCAACG 3′

2.5.3.1 Simultaneous generation and phenotyping of the *C. elegans gphr-1 gphr-2* double mutant

Twelve male RB1390x6 worms at the young adult stage were picked onto a plate and crossed with four L4 FXO4228x6 hermaphrodite worms. After 4 d, ten candidate cross hermaphrodites at the L4-stage were moved to individual plates and left to self-fertilise for 4 d. These were genotyped by single worm lysis (section 2.7.6) and PCR (section 2.7.10) to identify a *gphr-1(+/-) gphr-2(+/-)* worm (using primers from Table 2.3). After identifying a plate whose original hermaphrodite mother was *gphr-1(+/-) gphr-2(+/-)*, 20 individual late L4-staged worms were picked to new individual plates. Also, 20 individual late L4-staged N2 worms were picked to new individual plates. Worms were incubated at 20 ˚C for 24 h then transferred to new individual plates and incubated at 20 ˚C for 1 h. After 1 h, each worm was transferred to a new plate and the number of eggs laid was scored and the number of hatched eggs was scored 2 d after. All plates were incubated at 20 ˚C for 24 h and the process was repeated for a further 4 d. On the fourth day, rather than transferring the worms to new plates, genomic DNA was extracted from them for genotyping to identify a *gphr-1(-/-) gphr-2(-/-)* worm. To score total number of eggs and progeny produced by the candidate *gphr-1(-/-) gphr-2(-/-)* and N2 worms, the number of adult worms were picked off individually and scored using a counter until no worms were left on each plate in order to accurately count the progeny. If no *gphr-1(-/-) gphr-2(-/-)* candidates were identified, the process was repeated using 20 candidate *gphr-1(-/-) gphr-2(-/-)* L4-staged worms from a *gphr-1(+/-) gphr-2(-/-)* plate and 20 N2 worms as the control. This scheme is outlined in Figure 6.13. Images of the eggs were taken to follow embryogenesis using the eggs laid on the 1-h plates (as these eggs are laid within 1 h of one another i.e. synchronised) by the L4+1 animals. Images of the eggs and worms were taken on a Zeiss Axioplan 2 (Carl Zeiss Micrscopy, Oberkochen, Germany) on the NGM plate in order to not disturb them. Whole-worm images were taken using either Zeiss 10x 0.3 N.A. or Zeiss 20x 0.5 N.A. objectives. Images of eggs were taken using a Zeiss 40x 0.75 N.A. objective.

2.5.4 RNAi treatment of *C. elegans*

Feeding bacteria were inoculated in LB medium (section 2.7.1) containing the appropriate antibiotic selection (Table 2.4) e.g. tetracycline (12.5 μM) and ampicillin (100 μM) for HT115(DE3) pL4440 empty vector. This was achieved by picking a

single well-isolated colony and dropping it into LB containing selection (Table 2.4). The inoculate was agitated overnight at 225 rpm at 37 °C using an Innova ™ 4300 Incubator Shaker (New Brunswick Scientific Ltd., UK). 150 µL bacterial culture was seeded directly onto NGM plates (section 2.4.1) containing 1 µM IPTG and 50 µM ampicillin (induced), or NGM plates containing only 50 µM ampicillin (uninduced). Seeded plates were dried at room temperature overnight. Single L4-stage RB1390x6 hermaphrodite worms were transferred onto NGM plates.

Table 2.4 RNAi feeding bacteria and vectors used in this thesis

Bacterial background	Vector	Selection	Explanation
HT115(DE3)	None	12.5 µM tetracyclin	HT115(DE3) *E. coli* – used as a negative control
HT115(DE3)	pL4440	50 µM ampicillin + 12.5 µM tetracyclin	Empty vector transformed into HT115(DE3) *E. coli* – used as positive control
HT115(DE3)	pL4440 C11H1.2	50 µM ampicillin + 12.5 µM tetracyclin	pL4440 vector carrying fragment of *Cel-gphr-2* designed to knockdown *Cel-gphr-2* expression
HT115(DE3)	pL4440 C11H1.2 ORF	50 µM ampicillin + 12.5 µM tetracyclin	pL4440 vector carrying whole ORF of *Cel-gphr-2* designed to knockdown *Cel-gphr-2* expression
HT115(DE3)	pLT61	50 µM ampicillin + 12.5 µM tetracyclin	Used as a positive control for RNAi experiments; specific knockdown of *unc-22* gene causes a twitching phenotype (Fire et al., 1998, Timmons and Fire, 1998, Timmons et al., 2001).

2.5.5 Generating transgenic *C. elegans* by microinjection to produce extrachromosomal arrays

A coinjection marker was used in all microinjections of the plasmids of interest for positive selection of successful *C. elegans* transformation; these are outlined in Table 2.5. A 30 µL injection mix was prepared containing 60 ng of the plasmid of interest and either 20 or 2 ng of the *Paman-2::aman-2::mCherry* or *Peft-4::mCherry::sp12* marker respectively, made up to final volume using sterile ddH_2O. The injection mix was centrifuged for 20 min at 8200 *g* (to remove physical impurities) and 20 µL of the supernatant was transferred to a new tube; this spin was repeated and 10 µL of the supernatant was transferred to a new tube. Injection needles were made from 1 mm outside diameter (O.D.) x 0.5 mm aluminosilicate glass capillaries (Harvard Apparatus Ltd., UK) and pulled on a Sutter Instrument Model P-2000 micropipette puller (Intracel LTD, Hertfordshire, UK) using program: Heat = 450, Fil=4, vel = 50, Del = 220, Pul = 170, Heat = 470, Fil = -, Vel = -, Del = -, Pul = -, Line = 1.

The open-end of the needle was placed into the injection mix for at least 1 h to fill by capillarity. The needle was then secured to a mount on the ECLIPSE TE2000 inverted microscope (Nikon, Tokyo, Japan) and moved into position using a TransferMan NK2 micromanipulator (Eppendorf, Hamburg, Germany). The needle was broken at its tip by gently brushing against a broken glass coverslip immersed in Halocarbon 700 oil (Halocarbon Products Corporation, South Carolina, USA). L4+1 N2 worms were transferred to a 2 % (w/v) agarose pad (prepared on glass microscope slides and baked overnight in a 100 °C oven) and covered with sufficient Halocarbon 700 oil to prevent drying out. DNA was injected into the worm gonad through the needle by using the Femtojet microinjection system (Eppendorf, Hamburg, Germany). A drop of sterile M9 buffer (22 mM KH_2PO_4, 42 mM Na_2HPO_4, 85.6 mM NaCl and 1 mM $MgSO_4$) was then applied onto the worm to dislodge it from the pad. A platinum wire was used to transfer the worm from the pad to a new plate. Individually injected worms were transferred to individual plates and incubated at 20 ˚C and allowed to self-fertilise for 3 – 5 d. The progeny of each individually injected worm were screened for positive transformants using an mCherry/mRFP filter on a dissecting microscope. Individual F_1 worms positive for the coinjection marker (red fluorescence) were transferred to individual plates to generate independent lines. The progeny of each independent line were then screened 4 d later for red fluorescence indicating transmittance of

extrachromasomal arrays. Any lines that transmitted the extrachromasomal arrays to its progeny were deemed to be stable extrachromasomal lines and were maintained by picking 4 – 5 marker positive worms per line to a new plate each week.

Table 2.5 Plasmids and coinjection markers used in this study for *C. elegans* microinjection

Plasmid of interest	Coinjection marker
Pmyo-2::Cel-gphr-1::egfp	*Paman-2::aman-2::mCherry*
Pmyo-2:: Cel-gphr-1::egfp	*Peft-4::mCherry::sp12*
Pmyo-2::Cel-gphr-2::egfp	*Paman-2::aman-2::mCherry*
Pmyo-2:: Cel-gphr-2::egfp	*Peft-4::mCherry::sp12*
P-eft-3:: Cel-gphr-1::egfp	*Paman-2::aman-2::mCherry*
P-eft-3:: Cel-gphr-1::egfp	*Paman-2::aman-2::mCherry*
Pgphr-1::egfp	*Peft-4::mCherry::sp12*
Pgphr-2::egfp	*Peft-4::mCherry::sp12*
Pgphr-2::gphr-2::egfp	*Peft-4::mCherry::sp12*

2.6 Growth of *Saccharomyces cerevisiae*

2.6.1 Culturing *S. cerevisiae* strains

*S. cerevi*siae strains were typically grown on synthetic complete (SC) medium and SC media without uracil (SC-ura) was prepared using reagents in the concentrations as outlined in Table 2.6. The media was adjusted to pH 5.3 (unless otherwise stated) and autoclaved on a sugar cycle for sterilisation. Where required, uracil was added to the medium to a final concentration of 76 mM before either culturing on 90 mm plates or for inoculation for yeast drop testing (section 2.6.2).

2.6.2 Transformation of *S. cerevisiae*

A single *S. cerevisiae* colony was resuspended in 1 mL sterile ddH_2O. The cells were pelleted in a microcentrifuge at 15,700 *g* for 1 min and the supernatant was removed. 250 µL of filter-sterilised PLATE solution (8:1:1 ratio of 50 % PEG: 1 M lithium acetate, pH 7.5: 10x TE buffer at pH 7.5) was added to the cells. DNA (sodium salt from salmon testes; Sigma, Missouri, USA) solution (1 µM salmon sperm; 10x TE buffer, pH 7.5) was boiled for 5 min to produce single-stranded DNA. Five µL of boiled salmon sperm DNA solution was added and gently vortexed. One µg of the plasmid of interest was resuspended in 30 µL of ddH_2O prior to addition to the cell solution mix. This was vortexed to resuspend the pellet prior to heatshock at 42 °C in a waterbath for 2 h. The cells were pelleted in a microcentrifuge for 1 min at 15,700 *g*. The supernatant was removed and the pellet resuspended in 200 µL of sterile ddH_2O. The solution was vortexed to resuspend the pellet prior to spreading on solid selective media (SC-ura glucose media as all plasmids used contained the *URA3* gene). Plates were incubated at 30 °C for 3 d to allow for positive transformants to grow. Table 7.2 indicates all the *S. cerevisiae* transformants generated in this way.

2.6.3 *S. cerevisiae* drop testing for growth assays

A single colony was inoculated overnight in 10 mL of SC glucose media containing uracil (Table 2.6), with shaking at 200 rpm at 30 °C in an Innova ™ 4300 Incubator Shaker (New Brunswick Scientific Ltd., UK). The overnight culture was corrected to an OD_{600} of 0.42 – 0.45 and the inoculate was diluted 1 in 10 and 1 in 100. Seven µL of each of the three dilutions was dropped onto SC plates containing appropriate nutrients (any additional components for various growth assays are specified in the figure legends). Plates were incubated at 30 °C for 3 d and photographs were taken to document growth. The pH growth assay was set up by adjusting the pH using 0.5 M KOH or acetic acid prior to autoclaving.

Table 2.6 SC-ura media composition for *S. cerevisiae* culture

Liquid media was made without the addition of agar for inoculating colonies.

Reagent	SC Glucose	SC Galactose
Yeast nitrogen base w/o amino acids and ammonium sulphate (Difco 233520)	0.17 %	0.17 %
Ammonium sulphate (Sigma, Missouri, USA)	0.5 %	0.5 %
Amino acids minus uracil (Sigma, Missouri, USA)	0.192 %	0.192 %
Agar	2 %	2 %
Glucose (Sigma, Missouri, USA)	2 %	-
Galactose (Sigma, Missouri, USA)	-	2 %

2.7 Standard molecular biology methods

2.7.1 Bacterial preparations and cultures

Various strains of *E. coli* and *Agrobacterium* were used. Bacterial cells to be used for transforming with plasmid DNA were required to be competent. Both electrocompetent *Agrobacterium* and chemically competent *E. coli* cells were generated. The different bacterial strains used in this study are shown in Table 2.7. For all cultures of bacterial strains, inoculation was carried out by picking a single well-isolated colony using a sterile metal loop and agitation was provided by the Innova ™ 4300 Incubator Shaker (New Brunswick Scientific Ltd., UK). *E. coli* strains were incubated in the Classic Incubator (LEEC, UK) at 37 °C, while *Agrobacterium* strains were incubated in the Gallenkamp Cooled Illuminated Incubator at 30 °C. Unless otherwise stated, LB medium (1 % (w/v) tryptone, 0.5 % (w/v) yeast extract and 1 % (w/v) NaCl, pH 7.0) and LB agar (LB medium with 1.5 % (w/v) agar) were used for culturing bacterial strains. These were prepared in ddH_2O and autoclaved (121 °C for 20 min) for sterilisation. LB agar was melted using a microwave and cooled to ~55 °C before adding any antibiotic selection and pouring into 90 mm plates. The Sorvall legend RT table-top centrifuge (Thermo Fisher Scientific, DE, USA) was used to prepare and harvest bacterial cells.

To generate electrocompetent *Agrobacterium*, the GV3850 strain was streaked onto LB agar plates with 50 μM carbenicillin and 50 μM rifampicin and cultured by incubating at 30 °C for 3 – 4 d. A single well-isolated colony was picked using a sterile pipette tip and inoculated in 1 mL of LB medium with antibiotic selection and incubated at 30 °C with vigorous agitation (225 rpm) overnight. The culture was poured into 100 mL of LB medium containing antibiotic selection and incubated at 30 °C with vigorous agitation for 16 h. The culture was chilled on ice for 10 min before harvesting the cells by centrifugation at 4000 *g* at 4 °C for 10 min. The supernatant was discarded and the cells were resuspended in 100 mL of ice-cold sterile ddH_2O. The resuspension steps were repeated twice under the same centrifugation conditions, but after the first centrifugation it was resuspended in 40 mL of ice-cold sterile ddH_2O and after the second centrifugation the cells were resuspended in 10 mL of ice-cold sterile ddH_2O. The cells were pelleted for a final time by centrifuging at 4000 *g* for 10 min at 4 °C and resuspending in 1 mL of 10 % (v/v) ice-cold sterile glycerol. Electrocompetent cells were dispensed into 50 μL aliquots, snap-frozen in liquid nitrogen and stored at -80 °C.

To generate chemically competent *E. coli* cells, the strain was streaked out onto LB agar plates and a single well-isolated colony was inoculated in 5 mL LB medium at 37 °C with vigorous agitation overnight. The bacterial suspension was poured into 250 mL of LB medium and cultured at 25 °C with moderate agitation (160 rpm) for 8 – 14 h, until an OD_{600} of 0.5 was obtained. The culture was cooled on ice for 10 min before harvesting the cells by centrifugation at 2500 *g* at 4 °C for 10 min. The supernatant was removed and the cells were resuspended in 20 mL filter-sterilised ice-cold transformation buffer (55 mM $MnCl_2$, 15 mM $CaCl_2$, 250 mM KCl and 10 mM PIPES (piperazine-1,2-bis[2-ethanesulfonic acid], pH 6.7). The cells were harvested again by centrifugation at 2500 *g* at 4 °C for 10 min and the supernatant discarded. The cells were resuspended using a small volume of ice-cold transformation buffer (~0.8 mL). DMSO (1.5 mL DMSO / 20 mL transformation buffer) was added and mixed gently to the bacteria suspension. Competent cells were dispensed into 50 μL aliquots, snap-frozen by immersing in liquid nitrogen and stored at -80 °C.

Table 2.7 Bacterial strains used for cloning in this thesis

Bacterial strain	Usage	Source
E. coli DB3.1	*ccdB* gene expressing vector propagation	In-house (section 2.7.1)
E. coli DH5α	Entry and expression vector propagation	In-house (section 2.7.1
E. coli TOP10	TOPO entry vector transformation	Invitrogen (CA, USA)
E. coli XL Gold	Propagation	Stratagene Agilent Technologies (CA, USA)
Agrobacterium GV3850	Arabidopsis expression vector transformation	Homemade (section 2.7.1)

2.7.2 Transformation of chemically-competent *E. coli*

Plasmid DNA (1 µL) was transferred into 50 µL chemically-competent *E. coli* cells and incubated on ice for 30 min. The cells were then heatshocked for 30 s at 42 °C and transferred immediately back to ice for 2 min. Next 800 µL of Super Optimal Broth with Catabolite repression (S.O.C.) or LB medium was added to the cells before incubating at 37 °C with vigorous agitation (225 rpm) for 1 h. The cells were then spread on freshly prepared LB agar plates containing 100 µg/mL antibiotic selection and incubated at 37 °C overnight.

2.7.3 Genomic DNA extraction from Arabidopsis and *O. sativa* tissue

The DNAmite plant genomic DNA extraction kit (Microzone, UK) was used to extract genomic DNA from Arabidopsis and *O. sativa* tissue according to the manufacturer's instructions. A few leaves from the were removed from Arabidopsis or one leaf from *O. sativa* and 300 µL of cell lysis (LA) solution was added before the tissue was ground using a pestle in an eppendorf. 30 µL of protein denaturation (PA) solution was added before the plant suspension was centrifuged at 9300 *g* at 4 °C for 5 min. The supernatant (~200 µL) was transferred to a new eppendorf and centrifuged at 9300 *g* at 4 °C for 5 min. 175 µL of the supernatant was transferred to a new tube containing 175 µL of microCLEANg (CA) solution and this was incubated at room temperature for 5 min. DNA was pelleted at 15,700 *g* for 7 min at 4 °C. The supernatant was removed and the DNA pellet resuspended in 20 µL of sterile 18 MΩ H_2O. The

DNA concentration was determined (section 2.7.8) and 1 µL of genomic DNA was used in a 10 µL PCR reaction (section 2.7.10).

2.7.4 RNA extraction from Arabidopsis tissue

Prior to use, all tips, eppendorfs, pestle and mortars, glassware, buffers and solutions were autoclaved for 90 min at 121 °C and gloves were worn at all stages. Arabidopsis tissue was harvested and frozen immediately in liquid nitrogen before it was ground using a sterile and prechilled pestle and mortar. The plant material was homogenised in 500 µL RNA Miniprep buffer (100 mM NaCl,10 mM Tris/HCl (pH 7.0), 1 mM Na_2 EDTA, pH 8.0 with NaOH, 1 % sodium dodecyl sulfate (SDS)) in a prechilled eppendorf before adding 30 % (v/v) phenol and 50 % (v/v) chloroform. This extract was centrifuged at 15,700 *g* at 4 °C for 5 min. The upper phase was extracted to a new tube and an equal volume of 4 M LiCl was added before the RNA was precipitated for 2 h at 4 °C. RNA was pelleted by centrifuging at 15,700 *g* at 4 °C for 20 min and the supernatant discarded. The RNA pellet was resuspended in 300 µL DNase buffer (10 mM Tris/HCl (pH 7.0), 2.5 mM $MgCl_2$, 0.5 mM $CaCl_2$) and 2 Units (1µL) of Turbo ^{TM} DNase (Ambion, CA, USA) before incubating at 37 °C for 60 min to remove any genomic DNA. Next, 83 % (v/v) phenol and 83 % (v/v) chloroform:isoamyl alcohol were added before centrifuging at 15,700 *g* at 4 °C for 5 min to remove protein and DNA remnants. The upper layer was transferred to a new eppendorf and 2.5x the volume of 95 % (v/v) ethanol containing 5 % NaAc (3 M, pH 5.5) was added before precipitating at -20 °C for 1 h to overnight. The RNA was pelleted by centrifuging at 15,700 *g* at 4 °C for 5 min and the supernatant was removed. The pellet was resuspended in 75 % (v/v) ethanol before centrifuging at 7600 *g* at 4 °C for 5 min and the 75 % (v/v) ethanol wash was repeated using the same conditions. The supernatant was removed and the RNA was finally resuspended in 30 µL TE buffer (pH. 7.0) (10 mM Tris/HCl , 1 mM Na_2 EDTA). The RNA concentration was quantified (section 2.7.8) and the quality was confirmed by running 1 µg of RNA in a 1 % agarose (w/v) gel (section 2.7.12).

2.7.5 RNA extraction from *O. sativa* tissue

Prior to use, all tips, eppendorfs, pestle and mortars, glassware, buffers and solutions were autoclaved for 90 min at 121 °C and gloves were worn at all stages. *O.*

sativa tissue was harvested and frozen immediately in liquid nitrogen before grinding using a sterile and prechilled pestle and mortar. The plant material was homogenised in 1 mL Trizol reagent (Invitrogen, CA, USA) and 25 % (v/v) chloroform, vortexed briefly and incubated at room temperature for 3 min before centrifuging at 12,000 *g* at 4 °C for 15 min. The upper phase was transferred to a new tube and 50 % (v/v) isopropyl alcohol was added. The mixture was vortexed briefly before incubating at room temperature for 10 min. RNA was pelleted at 12,000 *g* at 4 °C for 10 min and the supernatant removed. The pellet was resuspended in 1 mL 75 % (v/v) ethanol and vortexed briefly, before centrifuging at 7600 *g* at 4 °C for 5 min. The 75 % (v/v) ethanol wash was repeated using the same conditions. The supernatant was removed and the pellet air-dried for 5 min at room temperature prior to resuspending in 30 µL TE buffer (pH 7.0). The RNA concentration was determined (section 2.7.8) and the quality was confirmed by running 1 µg of RNA in a 1 % agarose (w/v) gel (section 2.7.12).

2.7.6 Genomic DNA extraction from *C. elegans*

To extract genomic DNA from a single *C. elegans* worm, lysis buffer (10 mM Tris HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 0.45 % NP40 (v/v), 0.45 % Tween 20 (v/v) and 10X gelatin) was used with 1 mM proteinase K. Single worms were picked into thin-walled PCR tubes (STARLAB Ltd, UK) containing 5 µL of lysis buffer and lysed using the following cycle: 60 °C for 1 h, 95 °C for 15 min and 4 °C indefinitely using a peqSTAR 96 Universal PCR machine (peqlab, DE). The genomic DNA sample was stored at -20 °C for immediate or future use.

To extract genomic DNA from a population of worms, worms were removed from an NGM plate using 1 mL M9 buffer (22 mM KH_2PO_4, 42 mM Na_2HPO_4, 85.6 mM NaCl and 1 mM $MgSO_4$). The worms suspended in M9 were transferred to a sterile eppendorf, and left to settle on ice. The worms were collected by pelleting briefly in a microcentrifuge at 4 °C. The supernatant was removed and the worm pellet was washed using 1 mL M9, before spinning briefly to pellet the worms; this wash step was repeated. The supernatant was removed and 100 µL lysis buffer containing 1 mM proteinase K was added. This was incubated at -80 °C for 15 min prior to incubation at 65 °C for 1 h, vortexing every 15 min. After 1 h, the mix was incubated at 95 °C for 15 min before placing on ice and adding 300 µL sterile ddH_2O. The genomic DNA was used immediately or stored at -20 °C for future use.

2.7.7 Genomic DNA extraction from *S. cerevisiae*

To extract genomic DNA from a single colony of *S. cerevisiae*, a single colony was resuspended in 100 µL 200 mM LiAc, 1 % SDS solution. The suspension was vortexed and incubated at 70 °C for 15 min. DNA was precipitated by adding 300 µL 96 % (v/v) ethanol. To pellet the DNA, the sample was briefly vortexed before centrifuging at 15,000 *g* for 3 min. The supernatant was removed and the pellet was washed using 500 µL 70 % (v/v) ethanol and centrifuged at 15,000 *g* for 1 min. The supernatant was removed and the pellet was resuspended in 100 µL sterile ddH_2O.

2.7.8 Quantification of DNA/RNA concentration using the NanoDrop

The quantity and purity of DNA and RNA samples was measured on a NanoDrop ND-1000 spectrophotometer (Thermo Fischer Scientific, DE, USA) using the program ND-1000 V3.7.1. One µL of the RNA/DNA sample was loaded onto the sample pedestal and the absorbance was measured over the range of 200 – 300 nm. The program utilises the Beer-Lambert equation (A = E x b x c) to correlate the absorbance reading at 260 nm to calculate nucleic acid concentration.

2.7.9 cDNA synthesis by reverse transcription

Total RNA was used to synthesise first strand complementary DNA (cDNA). One µg of RNA was added to 5 µM oligo dT_{18} primer (Sigma, Missouri, USA) in a total of 3.6 µL sterile 18 MΩ H_2O. This was incubated at 72 °C for 5 min and then chilled to 4 °C. To the denatured sample (3.6 µL), 1x buffer (Promega, WI, USA), 8 mM $MgCl_2$ (Promega, WI, USA), 2 nM dNTPs (containing 2 nM each of dATP, dTTP, dCTP and dGTP) and 1 µL ImProm-II™ reverse transcriptase (Promega, WI, USA) was added to a final volume of 20 µL. The sample was incubated at 25 °C for 5 min, 42 °C for 60 min and finally 70 °C for 15 min using a peqSTAR 96 Universal PCR machine (peqlab, DE). The sample was chilled to 4 °C and the DNA concentration determined (section 2.7.8) before use or storage at -20 °C.

2.7.10 RT-PCR amplification of DNA

RT-PCR reactions were used to amplify DNA using cDNA or genomic DNA templates. The reactions were carried in PCR tubes and cycled in a peqSTAR 96 Universal PCR machine (peqlab, DE). The concentration of PCR reagents and the

cycling conditions used for the different reactions are given in Table 2.8 and Table 2.9. All the primers (Sigma, Missouri, USA) used are listed in Table 2.10.

2.7.11 Colony PCR

Colony PCR was employed to check colonies for successful bacterial transformation. The PCR reaction was set up as described previously (section 2.7.10) but using one colony instead of 1 µL DNA. To do this a sterile pipette tip was touched on a colony before dipping and swirling into the 10 µL PCR reaction mix. The PCR reaction reagents and cycle conditions are outlined in Table 2.8 and Table 2.9.

Table 2.8 PCR reagent concentrations used for different reactions

	Amplifying products for cloning using proofreader *Pfu*	**RT-PCR + Colony PCR using BioMix™**
Reagent	**Concentration**	
Template DNA	10 nM	1 nM /colony trace
2X BioMix™	-	1x
10X *Pfu* Buffer	1x	-
10 mM dNTPs	200 µM	-
10 µM forward primer	1 µM	200 nM
10 µM reverse primer	1µM	200 nM
Pfu polymerase	1.25 Units	-
Sterile 18 MΩ H2O	Up to 50 µL	Up to 10 µL

Table 2.9 Cycling conditions used for different PCR reactions

	Amplifying products for cloning using proofreader *Pfu*	**RT-PCR + Colony PCR using BioMix™**
Step	**Temperatures and Duration**	
Initial denaturation	95 °C, 2 min	94 °C, 2 min
Denaturation	95 °C, 1 min	94 °C, 30 s
Annealing	54 °C*, 30 s	55 °C, 1 min
Extension	72 °C, 1 min/500 bases	72 °C, 1 min
Repeat cycle	40 cycles	35 cycles
Final extension	72 °C, 7 min	72 °C, 3 min
Storage	4 °C, indefinite	4 °C, indefinite

*Annealing temperature is specific to primers used for cloning (Table 2.13).

Table 2.10 All primers used in this study

Primer Name	Primer Sequence
CeAct-2 F	5′ TGCTGATCGTATGCAGAAGG 3′
CeAct-2 R	5′ GAAGCATTTGCGATGAACAA 3′
Actin2 F	5′ GGTAACATTGTGCTCAGTGGTGG 3′
Actin2 R	5′ CTCGGCCTTGGAGATCCACATC 3′
AtGTG1 F	5′ CACCAAAAAAATGAGTTACGGATGGGCG 3′
At GTG1 R	5′ TTACTCGATGGCGTGCTTATC 3′
AtGTG2 F	5′ CACCAAAAAAATGGGTTACGGATGGGGG 3′
At GTG2 R	5′ TTAGTCTATTGGGTGCTTATC 3′
AtGTG1.ex12 qF1	5′ GGGAAGTGGGTCTTCAAGCA 3′
AtGTG1.ex13 qR1	5′ GCACTCGCCACAAAAATTGC 3′
AtGTG1-G166L F	5′ GTGACCTTAATGGCTGTCTTATCACT TTTTGGAGCTGTAAATTTACCCT 3′
AtGTG1-G166L R	5′ AGGGTAAATTTACAGCTCCAAATGA AGATAAGACAGCCATTAAGGTCAC 3′
CeGTG1 F	5′ CACCATGGACGGATTAACAGAT 3′
CeGTG1 R	5′ TTATGCCCTAAACATTGAAGA 3′
CeGTG1 1F	5′ AATGCAATGAGGAAGGCAAG 3′
CeGTG1 1R	5′ TAGGACGCTCGAAACGAAGT 3′
CeGTG1 2F	5′ ACTTCGCCAAATGTTGAATCG 3′
CeGTG1 2R	5′ TAATCTGTGCCATCAAAAGTGC 3′
CeGTG1 3F	5′ AGTTCGAGGACGCAAAATCC 3′
CeGTG1 3R	5′ GGGTTGCAAATGAATCTCG 3′
CeGTG1 4F	5′ GAGACCCTCGCCCGATTCC 3′
CeGTG1 4R	5′ CTCGTAAATCGACACGAAGCAAGC 3′
CeGTG1 5R	5′ GTCATTGTGATCAGAAGACCTC 3′
CeGTG2 F	5′ CACCATGGAAGGAAGCTAT 3′
CeGTG2 R	5′ TCAGTGTTTGAATCTAGTGTCTC 3′
CeGTG 2 Forward	5′ TGCTATTACATCAGTCAGAGGT 3′
CeGTG2 1F	5′ CCATATCAATGCCAACGTAG 3′
CeGTG2 1R	5′ CGGATATTTTGGATCCAACG 3′
CeEGFP-R	5′ TCAAAAATAGGAGATCTCTATTGTAC 3′
CeGTG1 NO STOP R	5′ TGCCCTAAACATTGAAGAACCG 3′
CeGTG2 NO STOP R	5′ GTGTTTGAATCTAGTGTCTCC 3′
Fr1 F	5′ ATGAGTTACGGATGGGCG 3′

FR1 RT R	5′ TAAGAAAGCCCCAACAGAGG 3′
FR2 RT F	5′ GGCTATGCGTGTTCCATTTAC 3′
GTG1 2R	5′ ACTGACCAGCTGAGGCATGGTAAA 3′
GTG2 1R	5′ AAGTGAAAGCCTCCGTCTA 3′
GTG2 2F	5′ TGCGGTCGGTTCAAGAT 3′
Hyg7	5′ GTCGATGCGACGCAATCGTCCGATC 3′
Hyg8	5′ GTCTGGACCGATGGCTGTGTAGAAG 3′
JL202	5′ CATTTTATAATAACGCTGCGGACATCTAC 3′
OsGTG1F	5′ GCGGCTGTACAAGGAGTACG 3′
OsGTG1R	5′ CTTCTGATCCTTGAATCCTCTCC 3′
OsGTG1 TOPO F1	5′ CACCATGGGTTGGGGC 3′
OsGTG1 NO STOP R	5′ GACAAGCACCCAATTGATTGA 3′
Os UBQ F	5′ AACCAGCTGAGGCCCAAGA 3′
Os UBQ R	5′ ACGATTGATTTAACCAGTCCATGA 3′
pCeGTG1-1intron-R	5′ CTGAATGAAAAATTTGAG 3′
pCeGTG2-1intron-R	5′ CTGAAAAAAAACGAATGGG 3′
pro:CeGTG1-TOPO-F	5′ CACCGTGCTGAAATATTAATTTTG 3′
pro:CeGTG1-R	5′ ATTTTTTTGCTAAAAAATCGCTCTG 3′
pro:CeGTG2-TOPO-F	5′ CACCAAAGTCTTTTATTTTTGTT 3′
pro:CeGTG2-R	5′ ATTATTTTAAATTATTGTGCAGCAA 3′
pDEST-P-Forward	5′ CAGGAGGACCCTTGGCTCACAAGTT 3′
pDEST-P-Reverse	5′ GGGGGTGGGAGCACAGGGAGAA 3′
pYTV-R	5′ CTACGGCTTCATCGTGTTGC 3′
MDCnosR	5′ AAGACCGGCAACAGGATTC 3′
M13 F (-20)	5′ GTAAAACGACGGCCAG 3′
M13 R	5′ CAGGAAACAGCTATGAC 3′
pMDC35S	5′ CATTTGGAGAGGACCTCGACTCT 3′
p426Gal1.F	5′ GCGAAGCGATGATTTTTGATCTATT 3′
p426CYC1.R	5′ TCCTTCCTTTTCGGTTAGAGCG 3′
SC ACT1 F	5′ GATTCTGAGGTTGCTGCTTTGG 3′
SC ACT1 R	5′ GGGAAGACAGCACGAGGAG 3′
ScGTG F1	5′ CACCATGGAGGCACTTAT 3′
ScGTG R1	5′ TATTCAACACTGACTTTTTTGG 3′

2.7.12 qRT-PCR

Quantitative RT-PCR (qRT-PCR)was carried out using a SYBR Green kit (Primerdesign, Southampton, UK) according to the manufacturer's instructions. The concentration of sample components are shown in Table 2.12. These were loaded into 96-well plates and sealed with clear caps (MJ Research, MA, USA). The whole plate was vortexed briefly and spun at 2000 *g* for 2 min. The PCR conditions are outlined in Table 2.11. The qPCR reaction was run on an Opticon DNA Engine Continous Fluorescene Detector (GRI Braintree, UK) using the Opticon Monitor III program.

Table 2.11 Reagent concentrations used for qPCR

Reagent
2.5 ng template DNA
0.3 μM forward primer
0.3 μM reverse primer
1x SYBR-Green Master Mix
Sterile 18Ω H_2O to a final volume of 10 μL

Table 2.12 Cycling parameters used for the qPCR reaction

Step	**Conditions**
Initial denaturation	95 °C, 1 min
Denaturation	95 °C, 50 s
Annealing	65 °C, 50 s
Extension	68 °C, 4 min
Repeat cycle	40 cycles
Final extension	69 °C, 7 min
Storage	4 °C, indefinite

2.7.13 DNA/RNA gel electrophoresis

PCR and restriction enzyme digest products were assessed on a 1 % (w/v) agarose/TAE (40 mM Tris/acetate, pH 8.0, 1 mM EDTA) gel containing 3 % (v/v) Gel Red (Biotium Cambridge Biosciences, UK) for DNA detection at room temperature. Ten µL of DNA sample with 1 µL of DNA loading buffer (Bioline, MA, USA) was loaded into the gel wells. Additionally, 5 µL of DNA Hyperladder I (Bioline, MA, USA) was loaded to assess the product size (Figure 2.2). Gels were run at 120 V using Power Pac 200/2.0 (Bio-Rad, CA, USA) for 1 h in 1x TAE buffer, before visualising under UV light.

2.7.14 Purification of DNA from agarose gel

DNA was extracted from agarose gels for cloning (section 2.7.15) or sequencing (section 2.7.17). DNA was purified using the Qiagen Gel Extraction kit (Qiagen, CA, USA) according to the manufacturer's instructions. The DNA fragment was visualised under UV light and extracted from the gel using a clean, sharp scalpel blade. Buffer QG was added to the gel piece in a 3:1 ratio (v/v) and incubated at 50 °C for 10 min. One volume of 100 % isopropanol was added before applying the sample to a column to bind the DNA by spinning at 15,700 *g* for 1 min. The flow-through was discarded and 0.5 mL of Buffer QG was added to the column and centrifuged at 15,700 *g* for 1 min to remove traces of agarose. The column was washed by adding 0.75 mL Buffer PE to the column and centrifuging the sample at 15,700 *g* for 1 min. The column was placed into a new eppendorf and the DNA was eluted with 20 µL sterile 18 MΩ H_2O by centrifuging at 15,700 *g* for 1 min. The DNA was determined (section 2.7.8) and stored at -20 °C for future use.

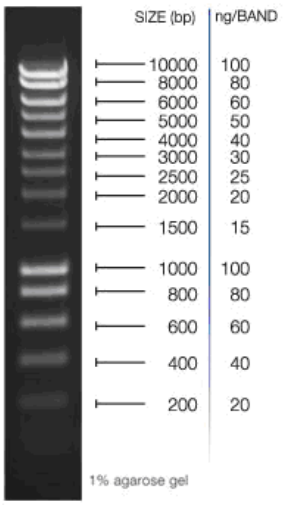


Figure 2.2 Hyperladder I DNA molecular marker used for gel electrophoresis

The range of molecular sizes shown on the ladder with higher intensity bands as 1000 bp and 10,000 bp. The Hyperladder I is supplied in a ready-to-use format and each lane of 5 µL provides 720 ng of DNA.

2.7.15 Gateway® cloning to generate Arabidopsis, *C. elegans* and *S. cerevisiae* expression vectors

The TOPO® cloning system (Invitrogen, CA, USA) was employed to produce entry vectors. All forward cloning primers were therefore designed to contain the sequence (5′ – 3′) CACC immediately prior to the start of the desired sequence. Primer pairs used for cloning and their product sizes are outlined in Table 2.13. For all cloning reactions the proofreading enzyme, *Pfu* polymerase, was used in the PCR reaction to amplify the product of interest; the cycling conditions are outlined in Table 2.8 and Table 2.9. Gel electrophoresis (section 2.7.13) was used as to show the PCR products were the correct size and the DNA was purified from the gel (section 2.7.14). The DNA concentration was determined (section 2.7.8) and the sequences of interest (Table 2.13) were cloned into the TOPO® pENTR/D entry vector (Invitrogen, CA, USA) according to the manufacturer's instructions (Invitrogen, CA, USA): 35 ng of PCR product was combined with 1 µL salt solution, 1 µL of TOPO® pENTR/D entry vector and sterile 18 MΩ H_2O to a final volume of 6 µL. The reaction was incubated at room temperature for 30 min and then 2 µL of the reaction was added to a vial of thawed One Shot® TOP10 chemically-competent *E. coli* cells whilst mixing gently. The cells were incubated on ice for 30 min and then heatshocked for 30 s at 42 °C and chilled on ice for 2 min before adding 250 µL of S.O.C. medium (Invitrogen, CA, USA). The cells were then vigorously agitated (225 rpm) at 37 °C for 1 h using an Innova ™ 4300 Incubator Shaker (New Brunswick Scientific Ltd., UK) before spreading on freshly prepared LB agar plates containing 100 µM kanamycin and incubated at 37 °C overnight. Colony PCR was performed on a selection of colonies from any successful transformants (section 2.7.11), to confirm the sequence had inserted into the TOPO® pENTR/D entry vector in the correct orientation. Any accurate clones (up to six) were inoculated in 5 mL LB media containing 100 µM kanamycin by incubating the culture at 37 °C overnight. The plasmid DNA was extracted using the Qiagen Mini-prep kit (Qiagen, CA, USA) as described in the manufacturer's instructions. Correct insert location and orientation was confirmed using both restriction enzyme digestion and sequencing (section 2.7.15 and 2.7.17 respectively). The restriction enzymes outlined in Table 2.15 were used to confirm successful generation of entry vectors by producing expected product sizes.

Prior to the LR recombination reaction, 500 ng of the entry vector was digested with the *Sfa*I restriction enzyme (section 2.7.15). The LR Clonase™ reaction was carried out based on the manufacturer's instructions (Invitrogen, CA, USA). At room temperature, 75 ng digested entry vector was combined with 75 ng destination vector and TE buffer (pH 8.0) to a final volume of 4 µL. One µL of LR Clonase ™ II enzyme was added to the reaction mix. The mixture was vortexed briefly and incubated at 25 °C for 10 h. To terminate the reaction, 0.5 µL of proteinase K was added followed by incubation at 37 °C for 10 min. One µL of LR reaction was used to transform chemically-competent DH5α *E. coli* cells (section 2.7.2). The cells were then spread on LB agar plates containing 100 µM antibiotic selection (Table 2.14) and incubated at 37 °C overnight. Colony PCR was performed on a selection of colonies from any successful transformants (section 2.7.11), to confirm the gene had inserted into the destination vector in the correct orientation. Any accurate clones (up to six) were inoculated in 5 mL LB broth containing 100 µM kanamycin by incubating the culture at 37 °C overnight with vigorous agitation (250 rpm). The plasmid DNA was extracted using the Qiagen Mini-prep kit. Correct insert location and orientation was confirmed using both restriction enzyme digestion and sequencing (section 2.7.15 and 2.7.17 respectively). The restriction enzymes outlined in Table 2.16 were used to confirm successful generation of expression vectors by producing expected product sizes. Table 2.14 outlines all the expression vectors generated in this way and Figure 2.3 shows vector maps of *S. cerevisiae* destination vectors used to highlight C-terminal tags.

2.7.16 Diagnostic restriction enzyme analysis of plasmids

Restriction enzyme digestion was conducted according to manufacturer's instructions (Promega, WI, USA) in a 10 µL reaction: 0.2 – 1 µg of DNA was added to 5 units of restriction enzyme (10 units if using two enzymes) with final concentrations of 1x buffer and 1 mM BSA made up to the final volume using sterile 18 MΩ H_2O. The reaction was incubated at 37 °C for 4 h followed by enzyme denaturation at 65 °C for 15 min. Samples were analysed by gel electrophoresis in a 1 % (w/v) agarose gel (section 2.7.12). The restriction digest product sizes were predicted using the pDRAW32 program (http://www.acaclone.com/).

Table 2.13 Primers and products used for cloning

Product	Primer set	Primer sequence	Product size (bp)	Annealing temperature (°C)
Cel-gphr-1 CDS with stop codon	CeGTG1 F CeGTG1 R	5′ CACCATGGACGGATTAACAGAT 3′ 5′ TTATGCCCTAAACATTGAAGA 3′	1402	54
Cel-gphr-2 CDS with stop codon	CeGTG2 F CeGTG2 R	5′ CACCATGGAAGGAAGCTAT 3′ 5′ TCAGTGTTTGAATCTAGTGTCTC 3′	1387	56
Cel-gphr-1 CDS without stop codon	CeGTG1 *F* CeGTG1 NO STOP R	5′ CACCATGGACGGATTAACAGAT 3′ 5′ TGCCCTAAACATTGAAGAACCG 3′	1399	58
Cel-gphr-2 CDS without stop codon	CeGTG2 *F* CeGTG2 NO STOP R	5′ CACCATGGAAGGAAGCTAT 3′ 5′ GTGTTTGAATCTAGTGTCTCC 3′	1384	55
Cel-gphr-1 promoter	pro:CeGTG1-TOPO-F pCeGTG1-1intron-R	5′ CACCGTGCTGAAATATTAATTTTG 3′ 5′ CTGAATGAAAAATTTGAG 3′	602	52
Cel-gphr-2 promoter	pro:CeGTG2-TOPO-F pCeGTG2-1intron-R	5′ CACCAAAGTCTTTTATTTTTGTT 3′ 5′ CTGAAAAAAAACGAATGGG 3′	1189	52
Cel-gphr-1 genomic sequence	pro:CeGTG1-TOPO-F CeGTG1 NO STOP R	5′ CACCGTGCTGAAATATTAATTTTG 3′ 5′ TGCCCTAAACATTGAAGAACCG 3′	5451	59
Gel-gphr-2 genomic sequence	pro:CeGTG2-TOPO-F CeGTG2 NO STOP R	5′ CACCAAAGTCTTTTATTTTTGTT 3′ 5′ GTGTTTGAATCTAGTGTCTCC 3′	3325	55
Cel-gphr-1 CDS with EGFP tag	CeGTG1 F CeEGFP-R	5′ CACCATGGACGGATTAACAGAT 3′ 5′ TCAAAAATAGGAGATCTCTATTGTAC 3′	2197	55
Cel-gphr-2 CDS with EGFP tag	CeGTG2 F CeEGFP-R	5′ CACCATGGAAGGAAGCTAT 3′ 5′ TCAAAAATAGGAGATCTCTATTGTAC 3′	2182	55

CDS = coding sequence

Table 2.14 The destination vectors used in LR recombination reactions in this thesis

Destination vector	Antibiotic resistance	Organism for expression
pMDC32	Kanamycin	Arabidopsis
pMDC83	Kanamycin	Arabidopsis
pAG426GAL-EGFP	Ampicillin	*S. cerevisiae*
pAG426GAL	Ampicillin	*S. cerevisiae*
pDEST-myo-2p	Ampicillin	*C. elegans*
pDEST- eft-3p	Ampicillin	*C. elegans*
pDEST-EGFP	Ampicillin	*S. cerevisiae*
pYTV	Ampicillin	*S. cerevisiae*

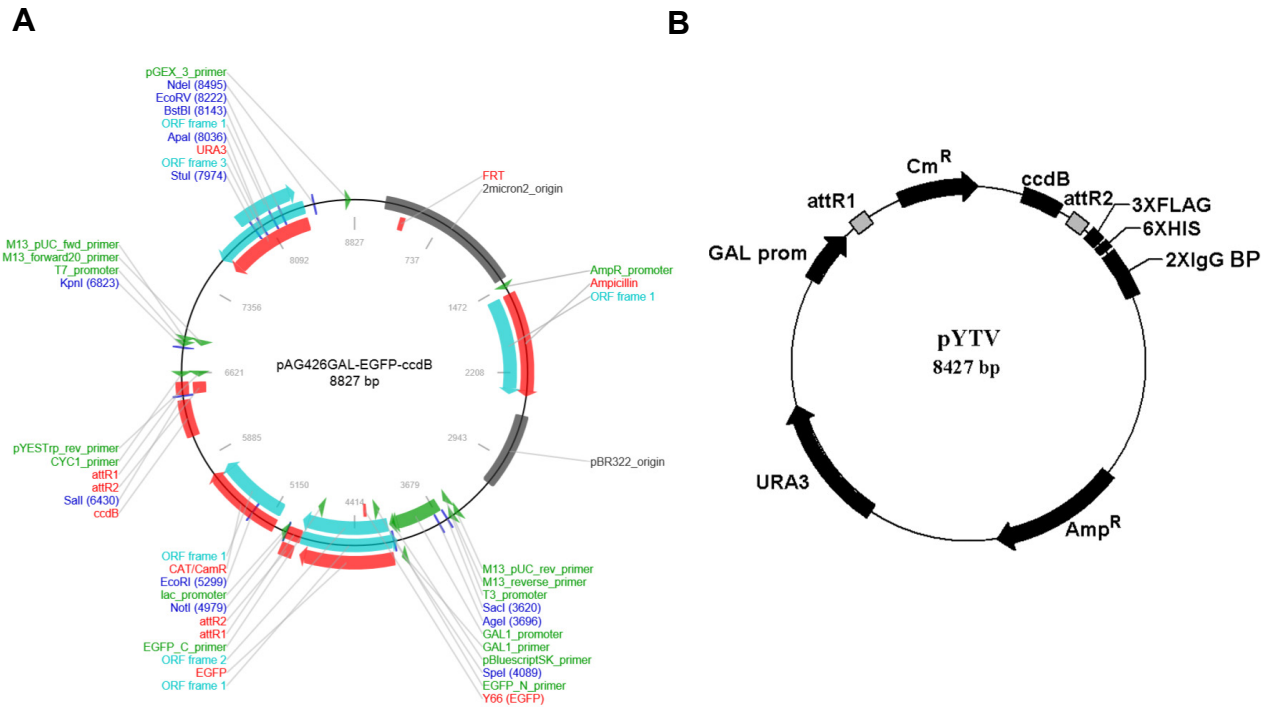


Figure 2.3 Vector maps of *S. cerevisae* destination vectors for C-terminal tagging

Vector maps of *S. cerevisae* destination vectors pAG426GAL-ccdB-EGFP (**A**) and pYTV (**B**). **A**. pAG426GAL-ccdB-EGFP plasmid allows for C-terminal EGFP tagging for localisation studies. B. pYTV plasmid features various C-terminal tags for protein purification: 3XFLAG (a triple FLAG epitope sequence), 6XHis (polyhistidine-tag of 6 hisitine residues), a 3C cleavage (a cysteine proteinase for recognising LEVLFQGP cleavage site for tag removal e.g. 6XHis; not shown) and 2XIgG binding protein (for tandem affinity purification). Figures taken from Alberti et al. (2007) (**A**) and Gong et al. (2004) (**B**).

Table 2.15 Restriction enzymes used to digest entry vectors and the resulting predicted product sizes

Virtual restriction analyses were conducted using pDRAW32 program (http://www.acaclone.com) and predicted product sizes for each vector is shown.

Vector	Restriction enzyme(s)	Predicted product sizes
pENTR/D *Cel-gphr-1* with stop codon	*Nsi*I + *Xho*I	2528, 1163, 266
	*Dra*I + *Pvu*II	1942, 937, 607, 471
pENTR/D *Cel-gphr-2* with stop codon	*Bam*HI + *Hind*III	3459, 483
	*Eco*RV + *Xho*I	2885, 1057
pENTR/D *Cel-gphr-1* without stop codon	*Dra*I + *Pvu*II	1942, 937, 607, 471
pENTR/D *Cel-gphr-2* without stop codon	*Eco*RV + *Xho*I	2885, 1057
pENTR/D *Cel-gphr-1* promoter	*Eco*RV + *Nsi*I	2459, 432 and 266
pENTR/D *Cel-gphr-2* promoter	*Eco*RV + *Nsi*I	3046, 432 and 266
pENTR/D genomic *Cel-gphr-2*	*Nsi*I + *Xho*I	3495, 1914 and 266
pENTR/D At $GTG1^{G166L}$	*Nsi*I + *Xba*I	1949, 1056, 651 and 266
pENTR/D *Cel-gphr-1::egfp*	*Eco*RV and *Xho*I	3226 and 1526
pENTR/D *Cel-gphr-1::egfp*	*Eco*RV and *Xho*I	2885 and 1852

Table 2.16 Restriction enzymes for confirmation of expression vectors and the resulting predicted product sizes

Virtual restriction analyses were conducted using pDRAW32 program (http://www.acaclone.com) and predicted product sizes for each vector is shown.

Expression vector	Restriction enzyme(s)	Predicted product sizes (bp)
pMDC32 *Cel-gphr-1*	*Eco*RV + *Xho*I	6059, 1852, 1395, 1094, 842, 135, 133
pMDC32 *Cel-gphr-2*	*Eco*RV + *Xho*I	6059, 2154, 1395, 1094, 483, 135, 133
pMDC32 At GTG1^{G166L}	*Nsi*I + *Xba*I	6051, 3766, 651, 569, 293 and 153
pMDC83 *Cel-gphr-1*	*Eco*RV + *Xho*I	6059, 1929, 1399, 1094, 1004, 656 and 131
pMDC83 *Cel-gphr-2*	*Eco*RV + *Xho*I	6059, 1929, 1399, 1094, 982, 663 and 131
pAG426GAL-EGFP *Cel-gphr-1*	*Eco*RV + *Xho*I	5531, 1661, 731 and 661
pAG426GAL-EGFP *Cel-gphr-2*	*Pvu*I + *Xho*I	3810, 2624, 1718 and 317
pDEST-myo-2p *Cel-gphr-1*	*Eco*RV + *Xho*I	5740, 1017 and 235
pDEST-myo-2p *Cel-gphr-2*	*Eco*RV + *Xho*I	6066, 664 and 235
pDEST-EGFP genomic *Cel-gphr-2*	*Dra*I	1911, 1558, 1293, 1023, 692, 636, 415, 225 and 130
pDEST-EGFP *Cel-gphr-1* promoter	*Dra*I	1911, 1089, 1023, 293 and 638
pDEST-EGFP *Cel-gphr-2* promoter	*Dra*I	199, 1989, 1023, 692, 624, 415 and 225
pDEST-eft-3p *Cel-gphr-1*	*Eco*RV + *Xho*I	6215
pDEST-eft-3p *Cel-gphr-2*	*Eco*RV + *Xho*I	6188
pAG426GAL-EGFP At *GTG1*	*Eco*RV + *Xho*I	6204, 1661 and 731
pYTV *Cel-gphr-1*	*Pvu*I + *Xho*I	2986, 2293, 1901, 481 and 173
pYTV *Cel-gphr-2*	*Pvu*I + *Xho*I	2645, 2293, 1901 and 980

2.7.17 DNA sequencing and analysis

Plasmid DNA or DNA purified from gels were sent off for Sanger sequencing (Source Bioscience Life Sciences, UK). In all cases forward and reverse reactions were obtained and this was repeated twice. The sequence received was visualised using BioEdit 7.0.9.0 Sequence Alignment Editor Program (http://www.mbio.ncsu.edu/bioedit/bioedit.html) and aligned using the EMBL-EBI EMBOSS Needle - Pairwise Sequence Alignment Program, which uses the Needleman-Wunsch algorithm (http://www.ebi.ac.uk/Tools/psa/).

2.7.18 Site-directed Mutagenesis (SDM) to generate At GTG1^{G166L}

The QuikChange® Primer Design Program was used to design primers for SDM (http://www.genomics.agilent.com/primerDesignProgram.jsp; Stratagene Agilent Technologies, CA, USA). The primers used for SDM of At *GTG1*^{G166L} are outlined in Table 2.17. Site-directed mutagenesis was carried out using the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene Agilent Technologies, CA, USA) according to manufacturer's instructions. The mutant strand synthesis reaction was set up in a PCR tube containing the reagents listed in Table 2.18 under using the cycling conditions outlined in Table 2.19. XL10-Gold Ultracompetent cells (Stratagene Agilent Technologies, CA, USA) were thawed on ice and 2 µL of β-ME was added to the cells and mixed gently. Cells were incubated on ice for 10 min with swirling at 2 min intervals. Two µL of the *Dpn*I-treated DNA was transferred to the ultracompetent cells and the mixed gently before incubating the reaction on ice for 30 min. The transformation reaction was then heatshocked at 42 °C for 30 s and transferred back to ice for 2 min. The S.O.C. medium was preheated to 42 °C and 0.5 mL was added to the reaction and the culture was incubated at 37 °C for 1 h with vigorous agitation (250 rpm). Cells were spread onto LB agar plates containing 100 µM kanamycin. Plates were incubated inverted at 37 °C overnight and a selection of colonies containing any successful transformants were used to carry out colony PCR (section 2.7.11) using primers FR1 F and M13 R (Table 2.1). Ten units (1 µL) of *Dpn*I restriction enzyme was added directly to the amplification reaction and mixed thoroughly prior to centrifuging for 1 min at 15,700 *g*. The reaction mixture was incubated at 37 °C for 2 h to digest the parental supercoiled plasmid DNA.

Following colony PCR confirmation (section 2.7.11), any accurate clones (up to six) were inoculated in 5 mL LB medium containing 50 µM kanamycin by incubating the culture at 37 °C overnight. The plasmid DNA was extracted using the Qiagen Mini-prep kit as described by the manufacturer's instructions. DNA extracted from one colony was sequenced and analysed (section 2.7.17). Following confirmation of successful SDM (At *GTG1*^{G166L}), the LR recombination reaction (section 2.13.4) was performed to insert At *GTG1*^{G166L} into pMDC32 to generate a plant expression vector. The expression vector was sequenced (section 2.7.17) and transformed into *Agrobacterium* for Arabidopsis transformation (section 2.3.3).

Table 2.17 Primers sets used for SDM of At GTG1^{G166L}

Mutated codon in bold and underlined

Primer	Primer sequence
AtGTG1-G166L F	5′ GTGACCTTAATGGCTGTCTTATCA**CTT**TTTGGAGCTGTAAATTTACCCT 3′
AtGTG1-G166L R	5′ AGGGTAAATTTACAGCTCCAAAT**GAA**GATAAGACAGCCATTAAGGTCAC 3′

Table 2.18. Reagents used for SDM

Reagent
1x reaction buffer
10 ng plasmid DNA
125 ng forward primer
125 ng feverse primer
1 µL dNTP mix
3 µL QuikSolution Reagent
Sterile 18 MΩ H_2O to a final volume of 50 µL
2.5 units *PfuUltra* High Fidelity DNA polymerase

Table 2.19 The cycling conditions used for mutant strand synthesis during SDM

Step	Conditions
Initial denaturation	95 °C, 1 min
Denaturation	95 °C, 50 s
Annealing	60 °C, 50 s
Extension	68 °C, 4 min
Repeat cycle	18 cycles
Final extension	69 °C, 7 min
Storage	4 °C, indefinite

2.8 Confocal fluorescent microscopy

All confocal fluorescent microscopy was conducted using a Spectral confocal Leica TCS SP2 laser scanning system (Leica Microsystems, Wetzlar, Germany). mCherry was excited at 594 nm; detection, 605 – 680 nm. EGFP and GFP were excitated at 488 nm; detection, 505 – 530 nm. Lambda scans were conducted to check for green autofluorescence across wavelengths 500 – 560 nm; excitation, 488 nm. D.I.C. or bright field images were also taken using the Leica system. Different organisms were prepared for imaging as described below.

Arabidopsis root cells were imaged by preparing T_2 seedlings grown on 0.5 MS containing 1 % (w/v) sucrose and 0.8 % (w/v) agar for 5 d (as described in section 2.3.1). Whole seedlings were placed on the slide in water and covered with a glass coverslip for imaging using a Leica 63x 1.2 N.A. water immersion lens.

Live-adult *C. elegans* were imaged by transferring to 2 % (w/v) agarose pads on glass slides. A drop of 10 mM levamisole (Sigma, Missouri, USA) was used to paralyse the worm and a glass coverslip was placed on top. Nail varnish was used to seal the coverslip prior to imaging using Leica 10x 0.4 N.A. and Leica 20x 0.7 N.A. objective for whole-worm imaging, and a Leica 63x 1.4 N.A. oil immersion objectives for higher magnification images.

To image live-cell *S. cerevisiae*, a single well-isolated colony was inoculated overnight in SC glucose media. The cells were centrifuged at 15,700 *g* and washed twice with SC galactose media. The cells were induced by resuspending in 5 mL SC-ura galactose media at 30 °C with 200 rpm agitation for 24 h. Following induction, the cells were pelleted at 15,700 *g* for 1 min, supernatant removed and resuspended in 1 mL sterile PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4). Seven µL of cells were mounted onto a glass slide sealed with a glass coverslip and nail varnish for imaging using Leica 100x 1.4 N.A and Leica 63x 1.4 N.A. oil immersion objectives.

2.9 Bioinformatic analysis of the GTG/GPHR sequences

The BLAST algorithm was used to identify homologous GTG/GPHR protein sequences in different species. Homologues of the Arabidopsis*, Brachypodium distachyon*, *C. elegans*, human, mouse, *S. cerevisiae*, *Pichia* and *Candida* GTG/GPHRs were found by searching with their protein sequence in the BLAST database (http://blast.ncbi.nlm.nih.gov/Blast.cgi). For phylogenetic analysis, MEGA6 was used to construct a Neighbour-Joining phylogenetic tree with 1000 replicates (http://www.megasoftware.net/). The multiple sequence alignments were made with the ClustalW module within MEGA6 using default parameters: gap opening penalty = 11; gap extension penalty = 1; protein weight matrix = BLOSUM with residue specific and hydrophilic penalties; gap separation distance = 5 and a 30 % delay divergent cutoff. This multiple alignment was used to identify completely conserved residues (Table 3.3). The sequences used to generate the trees are in Appendix 1. CDSs were translated into protein sequences using the Translate tool from ExPASy Bioinformatics Resource Portal (http://web.expasy.org/translate/). All computational analyses carried out and databases used are summarised in Table 2.20. Typically the sequences were entered into the program/web server and the output was saved. If this option was unavailable screenshots were taken instead.

Table 2.20 Programs/databases used for various computational analyses in this thesis

Analysis	Program/database	Website	Reference(s)
Conserved domains	Pfam database	http://pfam.xfam.org	Bateman et al. (2004)
	PROSITE	http://prosite.expasy.org/	Hulo et al. (2006)
C. elegans genome information	WormBase	http://www.wormbase.org/#01-23-6	Schwarz et al. (2006)
Arabidopsis genome information	TAIR	http://www.arabidopsis.org/	Huala et al. (2001), Rhee et al. (2003), Swarbreck et al. (2008) and Lamesch et al. (2012)
S. cerevisiae genome information	SGD	http://www.yeastgenome.org/	Chervitz et al. (1999) and Hirschman et al. (2006)
O. sativa genome information	Rice Genome Annotation Project	http://rice.plantbiology.msu.edu/	Ouyang et al. (2007)
Multiple sequence alignments	ClustalW and GeneDoc	http://www.clustal.org/	Thompson and Gibson (2002)
Phylogenetic analysis	MEGA6	http://www.megasoftware.net/	Tamura et al. (2013).
Protein disorder prediction	PredictProtein (Meta-Disorder)	http://ppopen.informatik.tu-muenchen.de/	Yachdav et al. (2014)
Secondary structure prediction	PredictProtein (PROFphd)		
Binding sites prediction	PredictProtein (ISIS)		
Functional changes due to single nucleotide polymorphism prediction	PredictProtein (SNAP2)		

Localisation prediction	LocTree2		
TM regions prediction	PredictProtein (PHDhtm)		
	Phobius	http://phobius.sbc.su.se/index.html	Käll et al. (2004)
	TopPred	http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::toppred	Claros and von Heijne (1994)
	SOSUI	http://harrier.nagahama-i-bio.ac.jp/sosui/sosui_submit.html	Hirokawa et al. (1998) and Mitaku and Hirokawa (1999)
	TMHMM	http://www.cbs.dtu.dk/services/TMHMM/	Sonnhammer et al. (1998) and Krogh et al. (2001)
	TMPred	http://www.ch.embnet.org/software/TMPRED_form.html	Hofmann (1993)
	MPEX	http://blanco.biomol.uci.edu/mpex/)	Snider et al. (2009)
	TOPCONS	http://topcons.cbr.su.se/	Bernsel et al. (2009)
	ZPRED		Granseth et al. (2006)
TM region prediction	ARAMEMNON	http://aramemnon.botanik.uni-koeln.de/	Schwacke et al. (2003)
Localisation prediction	WoLF PSORT	http://psort.hgc.jp/	Horton et al. (2006),
	TargetP 1.1	http://www.cbs.dtu.dk/services/TargetP/	Emanuelsson et al. (2007)
Signal peptide prediction	SignalP	http://www.cbs.dtu.dk/services/SignalP/	Petersen et al. (2011)
Gene expression analysis	Genevestigator	https://www.genevestigator.com	Hruz et al. (2008)
	BAR *A. thaliana* eFP Browser	http://bar.utoronto.ca/welcome.htm	Winter et al. (2007)

2.10 Statistical analysis

All statistical analysis was conducted using Minitab 16.0 (www.minitab.com/). For parametric analysis of variance (ANOVAs), all data must be normally distributed. The *Anderson-Darling* algorithm was used to analyse for normality. For data sets requiring normalisation, the *Johnson Normalisation* method was used. Following normalising of data, ANOVAs were conducted using a 95 % confidence level. ANOVAs are reported as an *F*-statistic and its associated degrees of freedom (d.f.) and *p*-value i.e. ($F_{\text{d.f. between groups, total d.f. within groups}}$ = *F*-value; $p <$ value). Where there was a significant difference ($p < 0.05$), Tukey *post hoc* tests were used to analyse the means that were significantly different between groups. Lowercase letters are used to show significant difference between means. Krustal-Wallis tests (non-parametric analogue of one-way ANOVA) were conducted when data could not be normalised. Mann-Whitney tests (non-parametric analogue of the *t*-test) were conducted as *post hoc* analysis for pairs of data sets to analyse groups with a significant difference. These were reported as the *H*-statistic of the Kruskal–Wallis test, d.f. and *p*-value, with mean ranks for each category; adjusted for ties indicates adjusted values when identical values were recorded. All statistical tests are reported in figure legends.

CHAPTER 3

Sequence analysis of the GTG/GPHRs

3.1 Introduction

3.1.1 Using bioinformatics and databases to analyse sequences

Bioinformatics is a scientific discipline applying computer science and information technology to biology (Pevzner, 2000). Besides genome and genetic analysis by large-scale DNA sequencing (Mount, 2004, Stenger and Gregory, 2006), bioinformatics predicts, analyses and interprets a variety of data. For example, DNA and protein sequences, protein domains and structures, gene ontology and gene expression can be analysed (Persson, 2000). Algorithms and statistics assess relationships between members of large data sets, predicting protein structure for instance, or aligning protein sequences within a family. For example, sequence alignment of different potassium (K^+) channels showed that they have a conserved signature sequence in the pore helix, and conserved residues in the inner helix (Heginbotham et al., 1994, Doyle et al., 1998).

Bioinformatics is a valuable tool used to explore protein structure, function and regulation, especially when these data are unavailable experimentally. There is a vast spectrum of methods used to predict characteristics of genes, proteins and regulatory sequences. Ultimately, *in silico* analysis allows the discovery of new biological insights to gain a global perspective that can be applied to experimental design (Persson, 2000). For instance, multiple sequence alignments can identify conserved residues in a protein that may be functionally relevant, and subsequent mutagenesis studies can determine such structure-function relationships. For instance, SNAP is a method for evaluating the effects of amino acid substitutions on protein function and could be useful for designing mutagenesis studies (Bromberg and Rost, 2007). This neural-network based method uses *in silico*-derived protein data including secondary structure, conservation and solvent accessibility. A score for independent amino acid substitutions of a protein sequence is translated into predictions of effect (present/absent) and reliability indices (RI) (Bromberg and Rost, 2007). For example, human organic anion transporter 4 (OAT4) variants, such as $OAT4^{L29P}$, were designed using SNAP to study functional

effects of amino acid substitutions. The $OAT4^{L29P}$ variant resulted in loss-of-function presented by reduced oestrone sulphate uptake (Shima et al., 2010).

Bioinformatics is useful for mapping and analysing DNA and protein sequences such as determining targeting signals of proteins. For example, the peroxisome targeting signals, type 1 and type 2, of plant peroxisomes were discovered by bioinformatics (Reumann, 2004). Bioinformatics is often used in combination with experimental data to identify or study proteins. For example, mass spectrometry alongside bioinformatics was used to identify cell wall proteins in apoplastic fluids of Arabidopsis rosettes (Boudart et al., 2005). In addition, gene expression profiling and bioinformatics analyses of Arabidopsis and cereals identified novel stress-regulated membrane proteins that were homologous to *Triticum aestivum* (bread wheat) cold-responsive genes of unknown function (Breton et al., 2003).

The Basic Local Alignment Search Tool (BLAST) algorithm is commonly used to compare primary sequences. A query sequence can be compared with a library of sequences to identify any resembling the query above a specific threshold (Altschul et al., 1990). Sequence alignments using algorthims such as ClustalW (Thompson et al., 1994, Thompson and Gibson, 2002) and MUSCLE (Edgar, 2004) to compare sequence(s) of interest with others of known structure or function are one of the most frequently used methods to assist with predicting gene families or related sequences. BLAST searches use sequence databases like UniProt (http://www.uniprot.org) and NCBI (http://www.ncbi.nlm.nih.gov/), which contain uploaded information from sequenced genomes and annotated sequences. However, not all organism genomes are fully sequenced or well-annotated, including *Triticum aestivum*. Therefore other methods of identifying similar/orthologue proteins are required in these cases (section 3.2.1).

Sequence alignments are often used as precursors for phylogenetic analyses to evaluate evolutionary relationships between sequences that may not be apparent from alignments alone. Phylogenetic studies using the software MEGA6 (Tamura et al., 2013), for example, were employed to show that there are two types of Arabidopsis caleosins, lipid-droplet surface proteins (Shen et al., 2014). Phylogenetics is conducted on similar proteins to designate whether they are part of a protein family or orthologue proteins (homologous proteins from different species). For example, phylogenetic relationships of the Arabidopsis cation transporter families were analysed to reveal that

there are several members with closely related isoforms in addition to separate subfamilies, thus suggesting that there are likely redundancies as well as diverse functions (Mäser et al., 2001). Typically full-length protein sequences are used for phylogenetic analyses although, sometimes only highly similar regions identified by alignments are selected for further analysis to gain a better insight into sequence relationships (Castresana, 2000). However, these analysis methods only show similarity between primary sequences, therefore experimental evidence would be required to support predictions of similarity. Furthermore, members of a gene family do not always cluster according to divergence of speciation, and could cluster according to function. For example, members of the ubiquitous Cation Diffusion Facilitator (CDF) family of metal ion transporters cluster into groups of the same specificity towards zinc, iron/zinc or manganese metal transport (Montanini et al., 2007).

3.1.2 Protein sequence annotation based on prediction programs

Prediction programs that annotate structural elements can be useful when the protein structure is unknown. Algorithms are available for predicting a wide range of structural elements including TM regions, secondary structure, domain architecture, protein disorder and binding sites. For example, PredictProtein is an open web server for annotating protein structure based on integrated prediction methods for secondary structure, solvent accessibility, TM helices, globular regions, disorder regions, intra-residue contacts, protein-protein, protein- and DNA-binding sites, subcellular localisation, β-barrels, cysteine bonds, metal binding sites and disulphide bridges (http://ppopen.informatik.tu-muenchen.de/; Yachdav et al. (2014)).

The ARAMEMNON database is a resource for plant membrane proteins found in both dicots such as Arabidopsis, and *Solanum lycopersicum*, and monocots such as *O. sativa* and *Zea mays* (http://aramemnon.botanik.uni-koeln.de/; Schwacke et al. (2003)). Information on sequences, gene expression, membrane protein topology predictions and relationships between orthologue genes can be found from a number of resources. Both Arabidopsis and *O. sativa GTG/GPHR* genes are annotated in this database. The membrane topology predictor in this database utilises up to 18 individual programs to contribute to the prediction of the TM helices including, SOSUI (Hirokawa et al., 1998), Phobius (Käll et al., 2004), TopPred (Claros and von Heijne, 1994) and TMHMM (Sonnhammer et al., 1998, Krogh et al., 2001). The prediction programs PHDhtm (Rost,

1996), MemSat (Jones et al., 1994) and Minnou (Cao et al., 2006) also take into account homologous proteins for their predictions. Therefore, ARAMEMNON is a very concise database for predicting plant membrane protein topology. In addition consensus topology predictions are generated using ConPred II (Arai et al., 2004) and the ARAMEMNON built-in TMultiConsensus and TConsensus methods (Schwacke et al., 2003). TConsensus uses the Bayes' theorem and generates a TM helix prediction based on individual predictions combined together. TMultiConsens is an extended consensus TM helix prediction performed by combining consensus predictions of homologous proteins. The algorithm maps TM helix consensus predictions to a multiple alignment of homologous protein sequences and projects the new consensus.

Computational methods may be used to predict whether a protein contains TM regions, characteristic of membrane proteins. Predictions are made based on the amino acid sequence of a protein and may be a useful precursor to functional and structural studies. Furthermore, some protein families have characteristic protein structures; for example, GPCRs are 7TM proteins characterised by an extracellular N-terminus and intracellular C-terminus with the tertiary structure forming a barrel-like conformation and the helices forming a cavity within the membrane (Rosenbaum et al., 2009). Therefore, an uncharacterised protein with this predicted structure can be annotated as a putative GPCR. For example, homology searches were used to identify putative GPCRs in the human genome (Takeda et al., 2002).

Membrane topology predictions can be informative as an initial model for uncharacterised protein structures; however, these need to be supported by experimental evidence. There are several membrane topology prediction methods that use different algorithms. Many methods, for instance, are based on the commonly used hidden Markov model such as in SOSUI and TopPredII (Claros and von Heijne, 1994, Hirokawa et al., 1998, Sonnhammer et al., 1998). Output data indicate the number of TM regions and the residues that are predicted to be within the membrane. Furthermore, most outputs suggest locations of the N- and C-termini in relation to the interior and exterior of the cell (Von Heijne, 1992, Claros and von Heijne, 1994, Hirokawa et al., 1998, Mitaku and Hirokawa, 1999, Takeda et al., 2002, Käll et al., 2004). These programs can be accessed as web servers or downloadable software. As with alignment and phylogenetic analyses, TM predictions are not based on experimental evidence and are not always completely accurate. For instance C. *elegans* and *Drosophila* proteins of

known topology were analysed using the TMHMM (a hidden Markov model-based method) and indicated 97-98 % of these proteins were correctly predicted (Krogh et al., 2001). Given the variation in results, in general several prediction and/or consensus prediction methods should be conducted.

Regions of protein disorder are flexible in solution and do not remain in a single equilibrium structure. Proteins can exhibit structural disorder in their native, functional state (intrinsically disordered proteins) as well as in specific regions (Iakoucheva et al., 2002, Dyson and Wright, 2005). Predicting protein disorder can be useful for linking function to structure. For example, calsequestrin is an intrinsically disordered, sarcoplasmic reticulum protein, which regulates intracellular calcium through C-terminal polyanion tails (Dunker et al., 2001). Protein disorder can be essential to function for example, phosphorylation ability (Iakoucheva et al., 2004). Disorder is encoded by the amino acid sequence and various methods have been developed for predicting this (Romero et al., 1997, Li et al., 1999, Obradovic et al., 2003). DISOPRED, for example, is a server for predicting protein disorder by using a knowledge-based method from an amino acid sequence (Ward et al., 2004a). There is also an integrated method used by the PredictProtein web server. This predicts regions of disorder and solubility determined by Meta-Disorder (MD), which is a consensus prediction of protein disorder (Schlessinger et al., 2009). MD is based on a resource of neural networks, combining outputs from several original prediction methods including NORSnet (Schlessinger et al., 2007a), DISOPRED2 (Ward et al., 2004a, Ward et al., 2004b), PROFbval (Schlessinger et al., 2006) and Ucon (Schlessinger et al., 2007b). With a wide range of prediction methods available, these approaches would extend structural and functional annotation of the GTG/GPHRs.

3.1.3 Gene expression analysis

Gene expression analysis is important in revealing spatial and developmental gene expression, which is useful for linking protein function and localisation. Understanding the function of proteins and their expression patterns is central for studying cellular responses to environment and regulatory gene defects in for example disease and cancer. For example, gene expression analysis of membrane proteins in a human mesenchymal stem cell (MSC) line undergoing osteoblast differentiation was important for finding novel molecular markers and provided insight into the

proliferation and differentiation of MSCs, which provides potential therapeutic targets (Foster et al., 2005).

Databases containing such information can therefore be useful for linking protein function to gene expression. The Bio-Analytic Resource for Plant Biology (http://bar.utoronto.ca/welcome.htm) is a database that can be outsourced from ARAMEMNON, containing gene expression data for predominantly Arabidopsis, but also other plant species including *O. sativa*. Genevestigator is a search engine for gene expression, which utilises manually curated, microarray and RNAseq experiments. The program is available for a number of organisms including Arabidopsis, *O. sativa*, human and mouse. Gene expression data is presented under biological contexts such as developmental stages and neoplasms. There is also a tool to indicate how the genes repond to various diseases, chemicals and experimental conditions (pertubations) (Hruz et al., 2008).

3.1.4 Aims

The structure of the GTG/GPHR proteins is yet to be characterised. Protein sequence alignments have been carried out suggesting that the GTG/GPHRs are a family of conserved membrane proteins with 8-9 TM domains (Maeda et al., 2008, Pandey et al., 2009, Jaffé et al., 2012). Although some bioinformatics has been conducted, there is still little known relating structure and function of the GTG/GPHRs. *In silico* analyses of the GTG/GPHR sequences will provide further insight into the current understanding of the structural and functional features and this could help experimental design. To do this, a number of aims have been set out for this chapter:

1. To identify GTG/GPHR orthologues which have not already been detected.
2. To conduct phylogenetic analyses on the GTG/GPHR protein sequences to evaluate the evolutionary relationships between these proteins.
3. To identify similarities between the GTG/GPHRs using multiple sequence alignments to find conserved residues and identify a domain architecture.
4. To conduct computational structural analysis of the GTG/GPHRs using several prediction methods.
5. To analyse expression profiles of the GTG/GPHRs using meta-analysis databases, which may reveal information regarding their function and localisation.

3.2 Results

3.2.1 Searching for and identifying GTG/GPHRs

The BLAST database was used to identify GTG/GPHR proteins from different species (section 2.9). Using this method 179 full-length GTG/GPHR proteins were identified in eukaryotes with no orthologues in prokaryotes. Several GTG/GPHRs from different organisms will be discussed in this chapter; to avoid confusion the genera and species will be specified in full in this chapter from herein (unless otherwise abbreviation already mentioned e.g. *Homo sapiens* → *H. sapiens*).

The Wheat Portal of Unité de Recherches en Génomique Info (URGI) is a resource for databases and tools to study genetics and genomics of wheat (http://wheat-urgi.versailles.inra.fr). The wheat *Triticum aestivum* (*T. aestivum)* genome is yet to be fully sequenced and annotated. *T. aestivum* is hexaploid (AABBDD) (Snape and Pánková, 2007, Shewry, 2009). The *Brachypodium distachyon* (*B. distachyon*) genome has been sequenced and is more closely related to cereals wheat and barley than the *A. thaliana* genome (Draper et al., 2001, Huo et al., 2009, Larre et al., 2010). *T. aestivum* orthologues can be searched by using *B. distachyon* coding sequences (CDSs) as query sequences in *T. aestivum* databases such as the Wheat Portal. *T. aestivum* contigs may be returned from the search and subsequently used to obtain *T. aestivum* CDSs and translated for the protein sequence. The *B. distachyon* GTG/GPHR CDS (LOC100825374) was searched in the database and results indicated a contig from chromosome 2BL matched the full-length of the CDS of *B. distachyon GTG*. Using this contig sequence, a predicted CDS of *T. aestivum GTG/GPHR* 2BL was obtained and translated for the protein sequence. However, since the hexaploid genome of *T. aestivum* is yet to be fully sequenced; only partial sequences of homeologues *T. aestivum GTG/GPHR* AL and *GTG/GPHR* DL were acquired.

So far, the GTG/GPHRs of *A. thaliana* (Pandey et al., 2009, Jaffé et al., 2012, Alvarez et al., 2013, Kharenko et al., 2013), *H. sapiens*, *C. griseus* (Maeda et al., 2008), and *M. musculus* (Tarutani et al., 2012) have been studied. To further investigate this family of proteins the GTG/GPHRs in *C. elegans*, *S. cerevisiae* and *O. sativa* will be considered in this thesis. For this reason, the analysis will be based on these GTG/GPHRs, and their gene structures are displayed in Figure 3.1. The gene structures

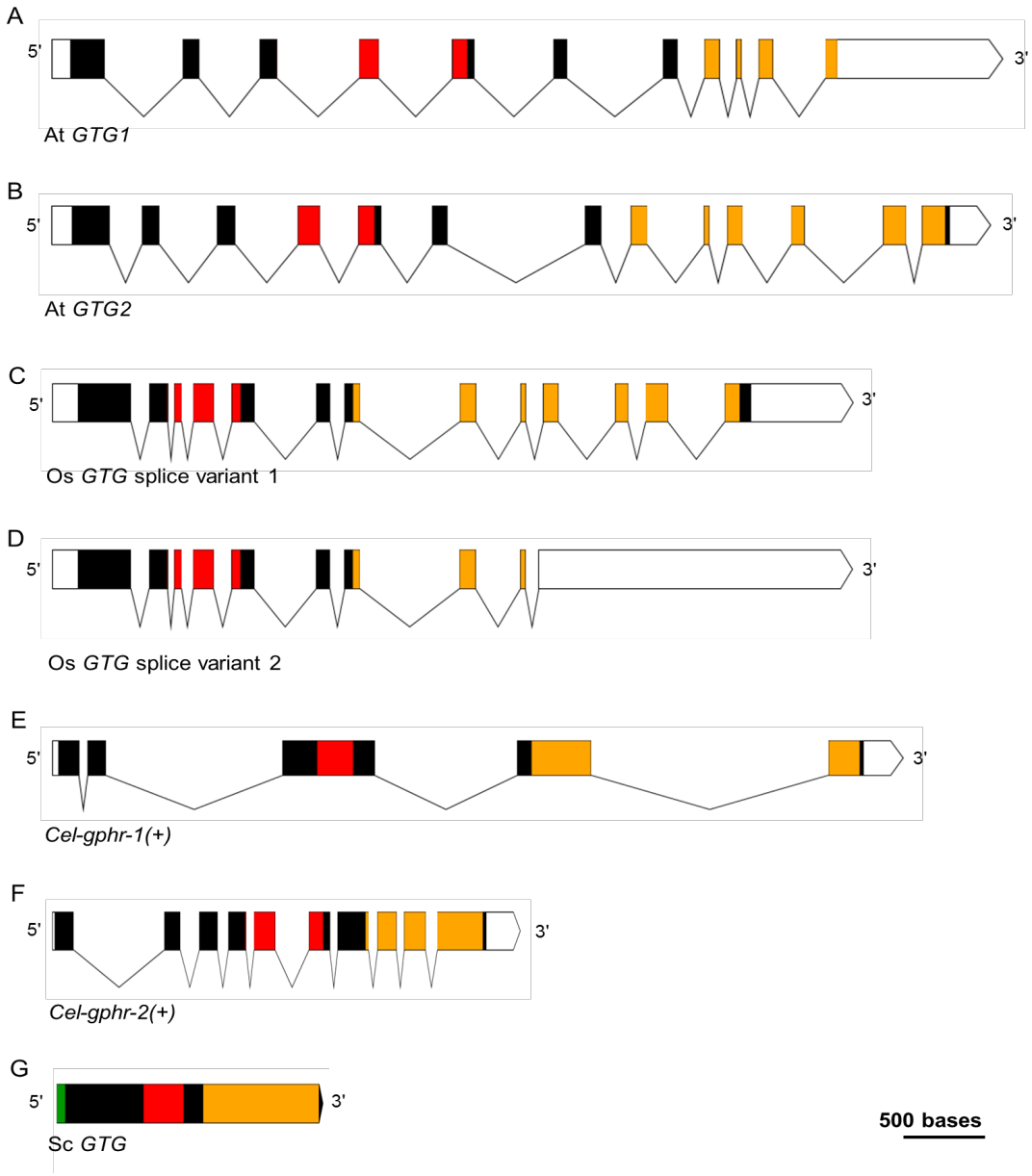


Figure 3.1 The gene structures of *A. thaliana*, *C. elegans*, *S. cerevisiae* and *O. sativa* GTG/GPHRs

Genomic structures *GTG/GPHR* genes in *A. thaliana*, *O. sativa*, *C. elegans* and *S. cerevisiae*, drawn to scale (500 bases) using Exon-Intron Graphic Maker (http://wormweb.org/exonintron). Black bar = exon, introns shown by black lines joining exons, red bar = conserved DUF3735, orange bar = conserved ABA-GPCR domain, white bar = UTRs green = signal peptide. **A**. *A. thaliana GTG1* (At *GTG1*) (5992 bp). **B**. *A. thaliana GTG2* (At *GTG2*) (5506 bp). **C**. *O. sativa GTG* (Os *GTG*) (4786 bp), splice variant 1 (LOC_Os02g06840.1) and splice variant 2 (LOC_Os02g06840.2) (**D**). **E**. *C. elegans gphr-1* (*Cel-gphr-1(+)*) (5050 bp). **F**. *C. elegans gphr-2* (*Cel-gphr-2(+)*) (2834 bp). **G**. *S. cerevisiae GTG* (Sc *GTG*) (1659 bp). Sequence information shown in (Table 3.1).

Table 3.1 Gene information of *GTG/GPHRs* from different organisms

Gene (ID)	Sequence information					
	Genomic	**Chromosome**	**Exons**	**Introns**	**CDS**	**Protein**
At *GTG1* (AT1G64990)	5992 bp	I	13	12	1407 bp	469 aa
At *GTG2* (AT4G27630)	5506 bp	IV	15	14	1404 bp	468 aa
Os *GTG* (LOC_Os04g51180.1)	4785 bp	IV	13	12	1407 bp	469 aa
Cel-gphr-1 (Y75B8A.16)	5050 bp	III	5	4	1398 bp	465 aa
Cel-gphr-2 (C11H1.2)	2834 bp	X	10	9	1383 bp	460 aa
Sc *GTG* (YHR078W)	1659 bp	VIII	1	0	1659 bp	553 aa

bp = base pairs, aa = amino acids

show that At *GTG* gene structures are similar, each with 13 relatively small exons across ~5 kB (Figure 3.1). At GTG1 and At GTG2 are highly similar in sequence (Table 3.2) and the presence of these two *GTG* genes is hypothesised as a result of gene duplication (Jaffé et al., 2012). The Os *GTG* contains 13 exons with two splice variants, LOC_Os02g06840.1 and LOC_Os02g06840.2 (Figure 3.1). Two splice variants exist for At *GTG1* (At1g64990.1 and At1g64990.2) and three for At *GTG2* (At4g27630.1, At4g27630.2 and At4g27630.3). The gene models At1g64990.1 and At4g27630.2 were confirmed for At *GTG1* and At *GTG2* respectively (Jaffé et al., 2012). The gene structures for *C. elegans gphrs* are quite dissimilar, with no splice variants predicted. *Cel-gphr-1* is ~5 kB containing five exons, while *Cel-gphr-2* is ~3 kb and contains ten exons. In contrast, Sc *GTG* is composed of only one exon (Figure 3.1).

Despite the differences in gene structure, the CDSs of these genes are similar sizes, ~1400 bp, producing protein sequences containing about 460 residues (Table 3.1). *S. cerevisiae* is most divergent in sequence, with a genomic sequence and CDS of 1659 bp and protein sequence of 553 amino acids (Table 3.1). Interestingly, the *H. sapiens GPHR* has three isoforms (GPR89A, GPR89B and GPR89C) predicted in the human genome found in the UniProt database (http://www.uniprot.org/) (Figure 3.2). The *H. sapiens* GPHR isoform, GPR89B, was cloned and expressed in Sf9 insect cells. Hs GPHR not only displayed channel activity (Maeda et al., 2008), but also was shown as a functional orthologue to *D. melanogaster GPHR* (Charroux and Royet, 2014); therefore, this was the isoform used to conduct phylogenetic analysis.

3.2.2 Phylogenetic analysis of the GTG/GPHRs

Traditionally, the kingdoms have been split into the five major groups: animals, plants, protists, fungi and monera (prokaryotes) (Whittaker, 1969). However with the continual advances in phylogenetics there are many organism which do not fit into one kingdom such as the “plant-like protista” chromalveolates (Simpson and Roger, 2004). Therefore in this study, the six-kingdom classification of organisms will be used (Cavalier-Smith, 1998). Additional GTG/GPHR protein sequences to those found by Jaffé et al. (2012) have been identified such as those from *T. aestivum*, *Pichia pastoris* and *Loa loa*. These GTG/GPHR sequences were aligned using ClustalW and phylogenetic analysis was conducted.

Table 3.2 Percentage identities and similarities between GTG/GPHR protein sequences from different organisms

Results obtained using the ClustalW algorithm (Thompson et al., 2002a, Li, 2003). Percentage identities shown in black, percentage similaries shown in light grey.

Sequence	At GTG2	Os GTG	*Cel*-GPHR-1	*Cel*-GPHR-2	Dm GPHR	Mm GPHR	Cg GPHR	Hs GPHR	Sc GTG
At GTG1	90 95	80 90	39 62	39 62	37 59	40 66	40 66	41 66	13 29
At GTG2		79 90	38 62	38 62	36 60	40 66	40 66	40 66	13 29
Os GTG			38 62	38 62	36 60	39 66	39 66	40 66	13 30
***Cel*-GPHR-1**				69 83	48 67	54 72	53 72	54 72	13 29
***Cel*-GPHR-2**					49 70	52 72	52 73	52 72	14 29
Dm GPHR						55 74	54 74	55 74	13 29
Mm GPHR							98 100	96 99	13 31
Cg GPHR								96 99	13 32
Hs GPHR									13 32

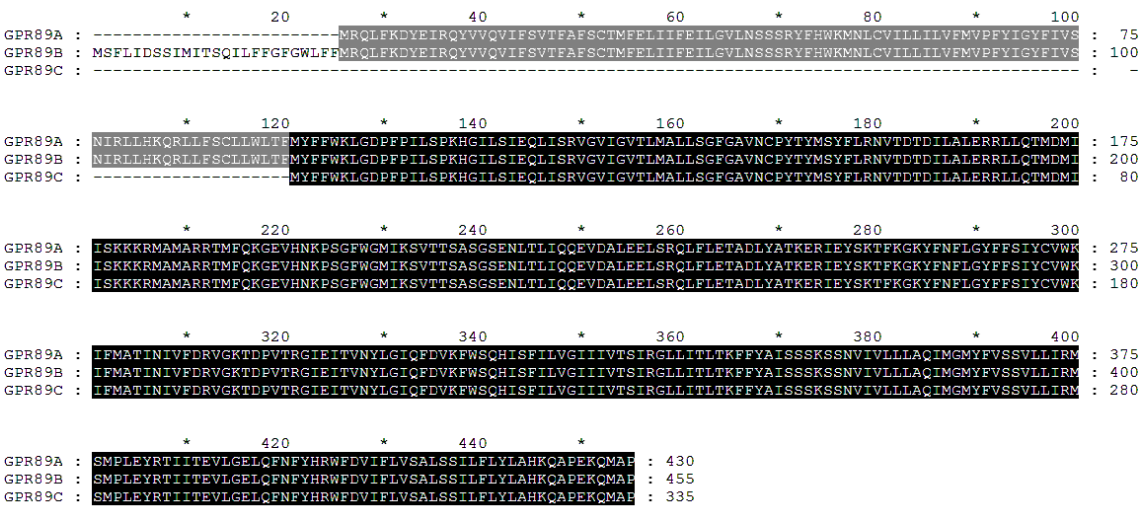
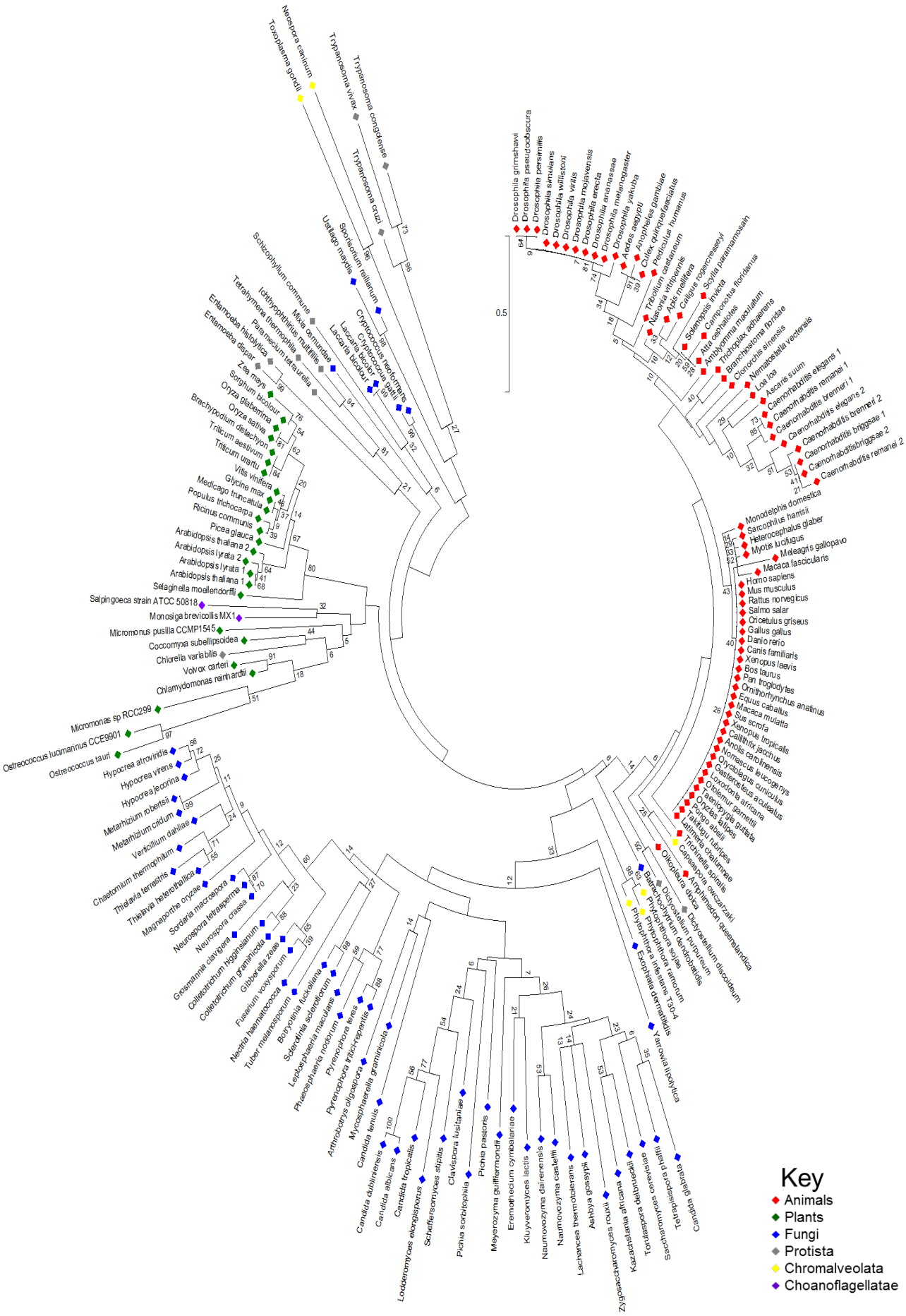


Figure 3.2 There are three isoforms for *H. sapiens* GTG/GPHR

Multiple protein sequence alignment using ClustalW (Thompson et al., 1994) and presented in GeneDoc (Nicholas and Nicholas Jr, 1997), showing the three protein isoforms for the GTG/GPHR found in *H. sapiens* according to the UniProt database (http://www.uniprot.org/). Black = identical residues in all three sequences aligned, grey = residues found in two out of three sequences aligned.


Drosophila grimshawi
Drosophila pseudoobscura
Drosophila persimilis
Drosophila simulans
Drosophila willistoni
Drosophila virilis
Drosophila mojavensis
Drosophila erecta
Drosophila ananassae
Drosophila melanogaster
Drosophila yakuba
Aedes aegypti
Anopheles gambiae
Culex quinquefasciatus
Pediculus humanus
Tribolium castaneum
Nasonia vitripennis
Apis mellifera
Caligus rogercresseyi
Solenopsis invicta
Scylla paramamosain
Camponotus floridanus
Atta cephalotes
Amblyomma maculatum
Trichoplax adhaerens
Branchiostoma floridae
Clonorchis sinensis
Nematostella vectensis
Loa loa
Ascaris suum
Caenorhabditis elegans 1
Caenorhabditis remanei 1
Caenorhabditis brenneri 1
Caenorhabditis elegans 2
Caenorhabditis brenneri 2
Caenorhabditis briggsae 1
Caenorhabditis briggsae 2
Caenorhabditis remanei 2
0.5
Monodelphis domestica
Sarcophilus harrisii
Heterocephalus glaber
Myotis lucifugus
Meleagris gallopavo
Macaca fascicularis
Homo sapiens
Mus musculus
Rattus norvegicus
Salmo salar
Cricetulus griseus
Gallus gallus
Danio rerio
Canis familiaris
Xenopus laevis
Bos taurus
Pan troglodytes
Ornithorhynchus anatinus
Equus caballus
Macaca mulatta
Sus scrofa
Xenopus tropicalis
Callithrix jacchus
Anolis carolinensis
Ictalurus punctatus
Oryzias latipes
Oreochromis niloticus
Gasterosteus aculeatus
Oryzias melastigma
Oreochromis aureus
Takifugu rubripes
Latimeria chalumnae
Tetraodon nigroviridis
Capsaspora owczarzaki
Amphimedon queenslandica
Dictyostelium discoideum
Dictyostelium purpureum
Batrachochytrium dendrobatidis
Oikopleura dioica
Phytophthora ramorum
Phytophthora sojae
Phytophthora infestans T30-4
Esophilia dermatitidis
Yarrowia lipolytica
Candida glabrata
Tetrapisispora phaffii
Saccharomyces cerevisiae
Tortaspora delbrueckii
Kazachstania africana
Zygosaccharomyces rouxii
Ashbya gossypii
Lachancea thermotolerans
Naumovozyma castellii
Naumovozyma dairenensis
Kluyveromyces lactis
Eremothecium cymbalariae
Meyerozyma guilliermondii
Pichia pastoris
Pichia sorbitophila
Clavispora lusitaniae
Scheffersomyces stipitis
Lodderomyces elongisporus
Candida tropicalis
Candida albicans
Candida dubliniensis
Candida tenuis
Mycosphaerella graminicola
Arthrobotrys oligospora
Pyrenophora tritici-repentis
Pyrenophora teres
Phaeosphaeria nodorum
Leptosphaeria maculans
Setosphaeria turcica
Botryotinia fuckeliana
Nectria haematococca
Tuber melanosporum
Fusarium oxysporum
Gibberella zeae
Colletotrichum graminicola
Colletotrichum higginsianum
Gnomoniopsis castanea
Neurospora crassa
Neurospora tetrasperma
Sordaria macrospora
Magnaporthe oryzae
Thielavia heterothallica
Thielavia terrestris
Chaetomium thermophilum
Verticillium dahliae
Metarhizium anisopliae
Metarhizium robertsii
Hypocrea jecorina
Hypocrea virens
Hypocrea atroviridis
Ostreococcus tauri
Ostreococcus lucimarinus CCE9901
Micromonas sp RCC299
Chlamydomonas reinhardtii
Volvox carteri
Chlorella variabilis
Coccomyxa subellipsoidea
Micromonus pusilla CCMP1545
Monosiga brevicollis MX1
Salpingoeca strain ATCC 50818
Selaginella moellendorffii
Arabidopsis thaliana 1
Arabidopsis lyrata 1
Arabidopsis lyrata 2
Arabidopsis thaliana 2
Picea glauca
Ricinus communis
Populus trichocarpa
Medicago truncatula
Glycine max
Vitis vinifera
Triticum urartu
Triticum aestivum
Brachypodium distachyon
Oryza sativa
Oryza glaberrima
Sorghum bicolour
Zea mays
Entamoeba dispar
Entamoeba histolytica
Tetrahymena thermophila
Paramecium tetraurelia
Ichthyophthirius multifiliis
Mixia osmundae
Schizophyllum commune
Laccaria bicolor
Cryptococcus gattii
Cryptococcus neoformans
Sporisorium reilianum
Ustilago maydis
Trypanosoma cruzi
Trypanosoma congolense
Trypanosoma vivax
Toxoplasma gondii
Neospora caninum
Key
Animals
Plants
Fungi
Protista
Chromalveolata
Choanoflagellatae

B

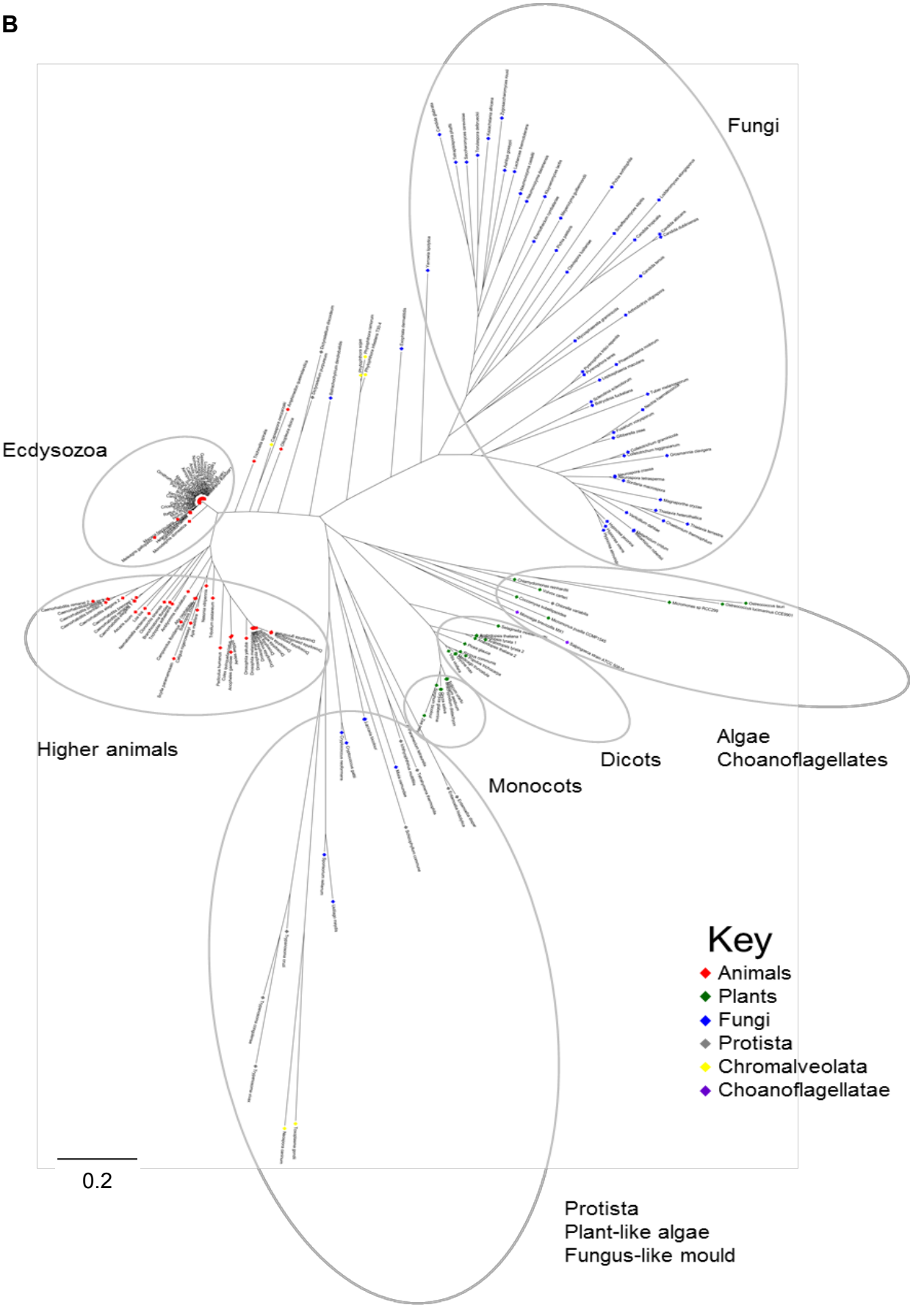

Fungi
Ecdysozoa
Higher animals
Dicots
Monocots
Algae
Choanoflagellates
Protista
Plant-like algae
Fungus-like mould
0.2
Key
Animals
Plants
Fungi
Protista
Chromalveolata
Choanoflagellatae

Figure 3.3 Phylogenetic analysis of predicted full-length GTG/GPHRs from a range of organisms from all eukaryotic kingdoms

Phylogenetic analysis showing non-rooted, bootstrapped plots (1000 replicates) constructed using a multiple alignment of 179 full-length GTG/GPHR protein sequences in circle (**A**) and radiation (**B**) branch styles. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The optimal tree with the sum of branch length = 32.74104615 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). Diamonds represent kingdoms or class: red = animals, green = plants, blue = fungi, grey = protists, yellow = chomalveolates, purple = choanoflagellates. The accession numbers can be found in Appendix 1.

The GTG/GPHRs are found in all the eukaryotic kingdoms (Figure 3.3). For example, they are found in lower animals like the simplest metazoan (multicellular animal), *Trichoplax adhaerens*, as well as the Ecdysozoa superphylum of the animal kingdom; including arthropoda (e.g. insects and crustaceans) like *Drosophila*, and nematoda (roundworms) such as *Caenorhabditis* (Figure 3.3). The GTG/GPHRs exist in higher chordate animals including *H. sapiens* and *M. musculus* (Figure 3.3). In plants the *GTG/GPHRs* are found in monocots and dicots as well as lower plant species such as algae (Figure 3.3). Besides metazoa, the GTG/GPHRs are also found in single-celled organisms such as the green algae *Chamydomonas reinhardtii*, *Ostreococcus tauri* and *Osterococcus lucimarinus* (Figure 3.3). In protista, GTG/GPHRs exist in amoebozoans (a type of protozoan) *Dictyostelium discoideum* and *Dictyostelium purpureum* (Figure 3.3). These species begin life as unicellular amoebae before subsequent transition into multicellular organisms. Other representative fungal GTG/GPHRs include those found in *Neurospora crassa, Lachancea thermotoleran, Gibberella zeae* and *Zygosaccharomyces rouxii* (Figure 3.3). Therefore, the GTG/GPHRs are broadly present in eukaryotes.

Generally, the GTG/GPHRs separate into various clades according to kingdoms (Figure 3.3). For example, the animal GTG/GPHRs cluster in higher vertebrate animals, such as *H. sapiens* and *C. griseus*; or the Ecdysozoa such as *C. elegans* and *D. melanogasta* (Figure 3.3). However, *Amphimedon queenslandica* (sponge), *Trichinella spiralis* (nematode parasite), and tunicate, *Oikopleura dioica* (sea-squirt) do not group with either the higher animals or the lower animals (Figure 3.3). These animal GTG/GPHRs are more similar to some protists or fungal GTG/GPHRs that also do not group with their own kingdom clades (Figure 3.3). For example, the *Amphimedon queenslandica* GTG/GPHR is most closely related to the single-celled protist *Capsaspora owczarzaki* GTG/GPHR. These GTG/GPHRs are more closely related to one another than to the lower and higher animals (Figure 3.3). Plant GTG/GPHRs (green; Figure 3.3) group into monocots and dicot clusters as well as a less defined group containing plant algae such as species belonging to the *Micromonus* genus, and unicellular choanoflagellate class (purple; Figure 3.3) that are organisms closely related to animals (red; Figure 3.3). Most of the fungal GTG/GPHRs form one clade, however some including *Laccaria bicolour*, *Mixia osmundae*, *Sporisorium reilianum*, *Ustilago*

maydis, *Cryptococcus neoformans* and *Cryptococcus gatti* group with protists, plant-like algae and fungus-like slime moulds (Figure 3.3).

The animal and plant GTG/GPHRs are closely related within their own taxa with little divergence. However, although the fungal and protist GTG/GPHRs tend to cluster together, they are more evolutionarily diverse, showing larger distance between sequences compared to any other kingdom (Figure 3.3 B and C). Furthermore, some fungal GTG/GPHRs are more similar to GTG/GPHRs from a different kingdom. For example, protist *Chlorella variabilis* groups with the algae plants and choanoflagellates; this is not surprising as it is a single-cell green algae. A group of GTG/GPHRs separate into one clade of protists, fungi-like slime moulds and plant-like algae: protists of the *Trypanosoma* genus, the *Cryptococcus* yeast and the chromalveolatae *Neospora caninium* and *Toxoplasma gondii* (Figure 3.3). Fungal GTG/GPHR sequences tend to be longer than the sequences from other kingdoms. For example, the longest GTG/GPHR sequence is the GTG/GPHR of the fungus *Toxoplasma gondii* (751 amino acids), followed by *Cryptococcus gattii*, *Cryptococcus neoformans*, *Thielavia terrestris* and *Thielavia heterothallica*. The shortest GTG/GPHR sequence is found in *Ostreoccocus tauri* (269 amino acids), closely followed by *Ostreococcus lucimarinus* (332 amino acids) and plant GTG/GPHR of *Zea mays* (367 amino acids) as well as plant algae of the *Micromonus* genus.

To explore regions of conservation between sequences from certain organisms, multiple alignments of these protein sequences were conducted. Since the *A. thaliana*, *C. elegans*, *O. sativa* and *S. cerevisiae* GTG/GPHRs will be studied throughout this thesis, these sequences were considered, along with the GTG/GPHRs in *H. sapiens*, *C. griseus*, *M. musculus* and *D. melanogaster* as these sequences have been previously studied (Maeda et al., 2008, Tarutani et al., 2012, Charroux and Royet, 2014). The protein sequence alignment demonstrates that there are conserved regions across the length of each GTG/GPHR from these organisms, with completely conserved residues in all ten organisms (Figure 3.4). As mentioned, the fungal GTG/GPHRs tend to be longer sequences than GTG/GPHRs from animals and plants. Therefore, alignments were also generated comparing GTG/GPHRs from: monocots and dicots (Figure 3.5 A); monocots *O. sativa*, *B. distachyon* and *T. aestivum* (Figure 3.5 B); *A. thaliana* and *C. elegans* (Figure 3.6); and fungal sequences *S. cerevisiae*, *Pichia pastoris* and *Candida albicans* (Figure 3.7).

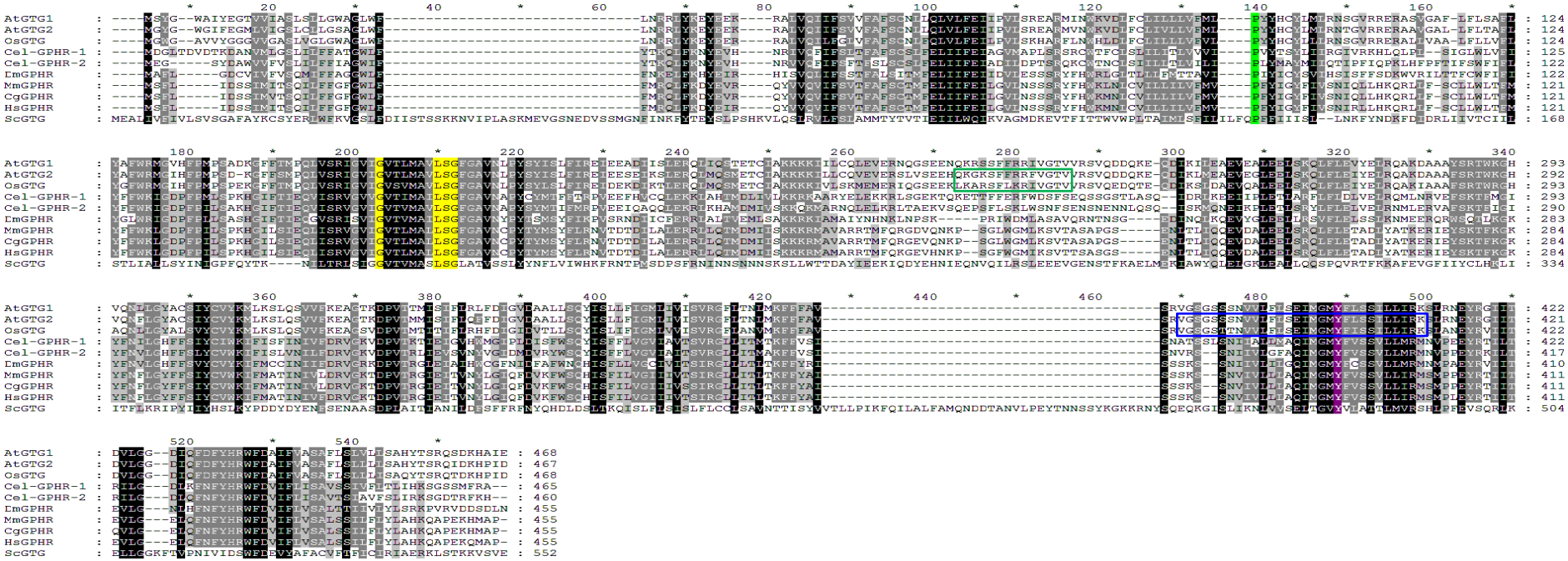


Figure 3.4 Protein sequence alignment of the GTG/GPHR proteins studied to date

Multiple alignment of the GTG/GPHRs in *A. thaliana* (AtGTG1 and AtGTG2), *O. sativa* (OsGTG), *C. elegans* (Cel-GPHR-1 and Cel-GPHR-2), *H. sapiens* (HsGPHR), *D. melanogaster* (DmGPHR), *C. griseus* (CgGPHR), *M. musculus* (MmGPHR) and *S. cerevisiae* (ScGTG). Sequences were aligned using the ClustalW algorithm (Thompson et al., 1994) and presented using GeneDoc (Nicholas and Nicholas Jr, 1997). For the sequences aligned here: black = conserved residues, dark grey = conserved in 8 – 9 of 10 organisms, light grey = conserved in 6 – 7 of 10 organisms; conserved residues found in all GTG/GPHRs (Table 3.3) are also highlighted: green = proline in predicted third TM region, yellow = in the DUF3735 or purple = in ABA-GPCR domain. The predicted Protein kinases ATP-binding motif (ATP-/GTP-binding site) (blue) (PROSITE, http://prosite.expasy.org) and Ras-GTPase-activating protein domain (green) (Pandey et al., 2009) are also shown by boxes for the At GTG sequences.

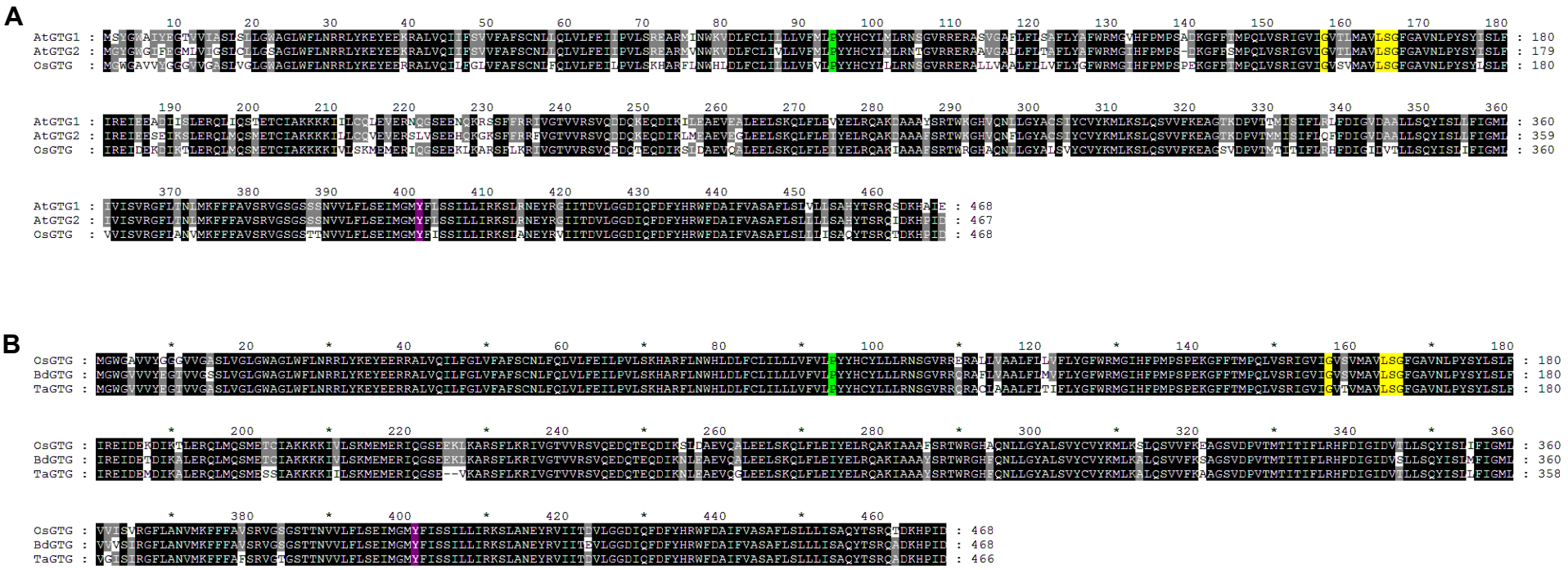


Figure 3.5 Plant GTG/GPHR proteins are closely related

A. Alignment comparing dicot and monocot GTG/GPHR protein sequences of *A. thaliana* (AtGTG1 and AtGTG2) and *O. sativa* (OsGTG). **B**. Alignment comparing monocot GTG/GPHR protein sequences from *O. sativa* (OsGTG), *B. distachydon* (BdGTG and *T. aesitvum* (TaGTG). Sequences were aligned using the ClustalW algorithm (Thompson et al., 1994) and presented using GeneDoc (Nicholas and Nicholas Jr, 1997). For the sequences aligned here: black = conserved residues, grey = conserved in 2 of 3 organisms; conserved residues found in all GTG/GPHRs (Table 3.3) are also highlighted: green = proline in predicted third TM region, yellow = in the DUF3735 or purple = in ABA-GPCR domain.

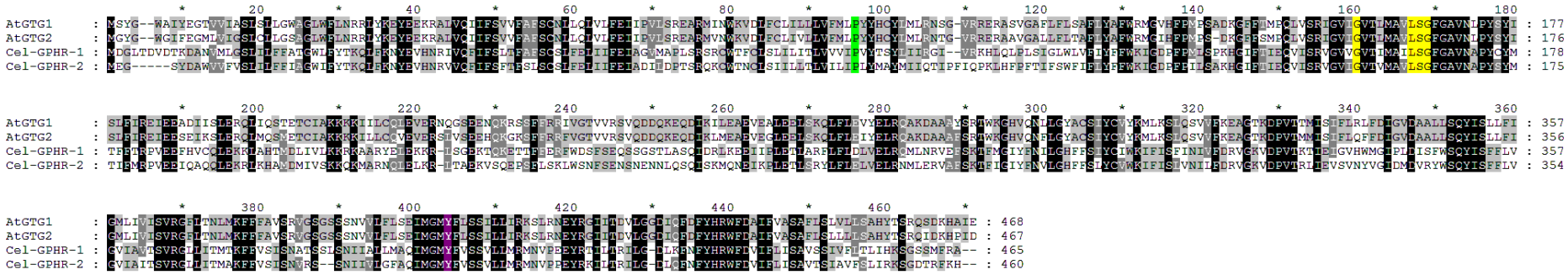
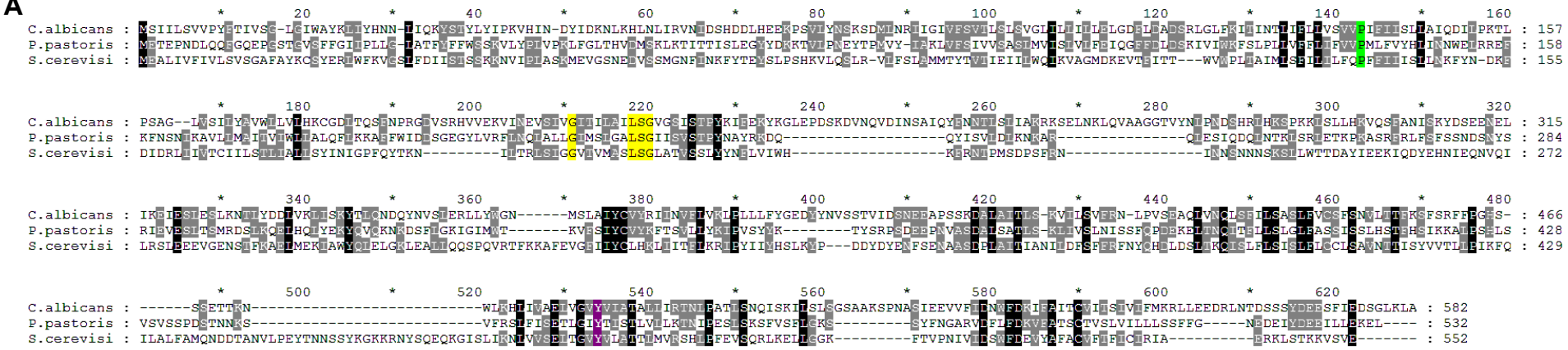


Figure 3.6 Protein sequence alignment of the *A. thaliana* and *C. elegans* GTG/GPHR proteins

Multiple alignment comparing the plant *A. thaliana* and animal *C. elegans* GTG/GPHRs. Sequences were aligned using the ClustalW algorithm (Thompson et al., 1994) and presented using GeneDoc (Nicholas and Nicholas Jr, 1997). For the sequences aligned here: black = conserved residues, dark grey = conserved in 3 of 4 organisms, light grey = conserved in 2 of 4 organisms; conserved residues found in all GTG/GPHRs (Table 3.3) are also highlighted: green = proline in predicted third TM region, yellow = in the DUF3735 or purple = in ABA-GPCR domain.

A



B

Sequence	*P. pastoris*	*S. cerevisiae*
C. albicans	23 46	17 36
P. pastoris		17 39

Figure 3.7 Fungal GTG/GPHR sequences are more diverse at the amino acid level

A. Multiple alignment comparing the plant *Candida albicans*, *Pichia pastoris* and *S. cerevisiae* funal GTG/GPHRs. Sequences were aligned using the ClustalW algorithm (Thompson et al., 1994) and presented using GeneDoc (Nicholas and Nicholas Jr, 1997). For the sequences aligned here: black = conserved residues, grey = conserved in 2 of 3 organisms; conserved residues found in all GTG/GPHRs (Table 3.3) are also highlighted: green = proline in predicted third TM region, yellow = in the DUF3735 or purple = in ABA-GPCR domain. **B**. Table of percentage identities and similarities between fungal GTG/GPHR protein sequence. Results obtained using ClustalW algorithm (Thompson et al., 2002a, Li, 2003). Percentage identities shown in black, percentage similaries shown in light grey.

The plant GTG/GPHRs are very closely related indicated by regions of high homology from sequence alignments (Figure 3.5). While *A. thaliana* and *O. sativa* are from different angiosperm groups, the GTG/GPHR protein sequences are very similar at the protein level (Figure 3.5 A), with At GTGs being 90 % identical and 95 % similar and sharing ~80 % identity and 90 % similarity with Os GTG (Table 3.2). Monocot GTG/GPHRs in *O. sativa*, *B. distachyon* and *T. aestivum* share higher homology with one another (Figure 3.5 B) (~95 % identity and 97 % similarity). The GTG/GPHRs from angiosperms are therefore closely related at the evolutionary level (Figure 3.3). Database searching indicate that most organisms contain only one copy of the *GTG/GPHR* gene, but there are some that contain two copies each (Figure 3.3); for example, nematode and dicot GTG/GPHRs of the *Caenorhabditis* and *Arabidopsis* genera. This may signify that there has been a recent duplication event resulting in two proteins produced in species arising from these genera.

To compare the similarity of plant and animal GTG/GPHRs at the protein level, the *A. thaliana* and *C. elegans* GTG/GPHRs were compared since both contain two copies of the *GTG/GPHR* genes. The alignment results show that both GTG/GPHRs share regions of homology along the whole length of their sequences (Figure 3.6). The percentage identities and similarities demonstrate that *A. thaliana* and *C. elegans* GTG/GPHRs are ~39 % identical and 62 % similar to one another (Table 3.2). Furthermore the two *A. thaliana* GTGs are more similar than the two *C. elegans* GPHRs, which are only 69 % identical and 83 % similar at the amino acid level (Table 3.2). This suggests that GTG/GPHR gene duplication events from the *Arabidopsis* genus have occurred more recently than the *Caenorhabditis* genus.

The plant and animal GTG/GPHR sequences are closely related within their own kingdoms (Figure 3.3 and Table 3.2). In contrast, the fungal GTG/GPHR sequences are more diverse from one another within this kingdom as shown by both the multiple sequence alignment (Figure 3.7) and phylogenetic trees (Figure 3.3). Table 3.2 shows that the *S. cerevisiae*, *C. albicans* and *P. pastoris* GTG/GPHR sequences are only ~20 % identical and ~40 % similar to one another, at the amino acid level. This is only slightly higher than its evolutionary relationships to the GTG/GPHRs from animal and plant kingdoms. For instance, Sc GTG is only ~13 % identical and ~30 % similar to the other studied GTG/GPHRs (Table 3.2).

3.2.3 The GTG/GPHRs contain conserved domains

The alignments demonstrate plant and animal GTG/GPHRs are highly conserved at the amino acid level. The *S. cerevisiae* GTG demonstrates less similarity to both the animal and plant GTG/GPHRs and only exhibits ~13 % identity to the plant and animal GTG/GPHRs (Figure 3.4 – Figure 3.6 and Table 3.2). Whilst building the phylogenetic tree, some completely conserved residues were identified in 179 GTG/GPHRs (with a few exceptions shown in Table 3.3). These residues in the *A. thaliana*, *O. sativa*, *T. aestivum*, *B. distachydon*, *C. elegans*, *H. sapiens*, *M. musculus*, *C. griseus* and *S. cerevisiae* GTG/GPHRs have been highlighted (Figure 3.4 – Figure 3.6). The prefixes for these organisms will be used when referring to these GTG/GPHRs from herein. The GTG/GPHRs share regions of homology throughout the protein sequences (Figure 3.4 – Figure 3.6). The Pfam database was used to identify conserved domains found in GTG/GPHRs: DUF3735 (**d**omain of **u**nknown **f**unction 3735) and ABA-GPCR (**ab**scisic **a**cid **G**-**p**rotein **c**oupled **r**eceptor). The conserved residues identified are found in either of these conserved domains of the GTG/GPHRs, except a conserved proline, which is predicted to be in the third TM segment (TM3) of plant GTG/GPHRs (section 3.2.3, Figure 3.10, Table 3.3).

DUF3735 is approximately 70 amino acids and exclusive to eukaryotes. It is found in 266 sequences across 220 species, many of which are predicted GTG/GPHRs while others are putative, uncharacterised proteins (Bateman et al., 2004). There are 239 sequences that have a domain architecture containing both DUF3735 and ABA-GPCR and are predicted to be GTG/GPHRs. The exception is *Sorghum bicolour*, whose predicted GTG/GPHR orthologue only contains an ABA-GPCR (using sequence obtained from the genome and ESTs). Amongst the 239 sequences that contain this protein architecture are the *A. thaliana*, *O. sativa*, *C. elegans*, *H. sapiens*, *M. musculus*, *C. griseus*, *D. melanogaster* and *S. cerevisiae* GTG/GPHR proteins.

There are 14 sequences containing DUF3735 alone; most are putative uncharacterised proteins or fragments of proteins (Bateman et al., 2004). Similarly, other protein domain architectures containing DUF3735 (Table 3.4) are mostly putative uncharacterised proteins (Bateman et al., 2004). DUF3735 has a characteristic conserved LSG sequence motif, with a single completely conserved glycine residue (Bateman et al., 2004). However, through constructing the phylogenetic tree, requiring alignment of all the input GTG/GPHR sequences, it has been found that this glycine is

replaced with an alanine in the *Zygosaccharamyces rouxii* GTG/GPHR, which is backed up by expressed sequence tag (EST) data. Since alanine has only a methyl group as a side-chain this substitution may not have an effect on protein function, though experimental evidence would be required to confirm this. Other conserved residues within DUF3735 were identified (Table 3.3); therefore, the signature sequence of DUF3735 will be denoted as GxxxxxxLSG from herein. Excluding exceptions (Table 3.3), the GxxxxxxLSG motif is conserved in all GTG/GPHR sequences shown in the phylogenetic tree (Figure 3.3) and found in the fifth TM segment (TM5) of all GTG/GPHRs analysed (Figure 3.13).

The ABA-GPCR domain, typically 177 – 216 amino acids in length is a domain found in eukaryotes (Bateman et al., 2004). There is a single conserved tyrosine found in all 179 GTG/GPHRs analysed (Table 3.3). ABA-GPCR is found in 301 sequences across 234 species, with 239 sequences also containing DUF3735 as predicted GTG/GPHRs and 62 others annotated as putative uncharacterised proteins (Bateman et al., 2004). There is a predicted GPCR (based on predicted topology) in *Ostreococcus tauri* (UniProt ID Q01A66) which contains an ABA-GPCR and serine/threonine protein kinase SAPK4 domain (Pfam ID PF00069). However, the majority containing only ABA-GPCR or in combination with another domain are putative, uncharacterised proteins such as the *Zea mays* protein with UniProt ID C0PMR5.

Pandey et al. (2009) indicated that the At GTGs have an ATP-/GTP-binding site and a Ras GTPase-activating protein domain, as predicted by PROSITE (http://prosite.expasy.org) (Figure 3.4). In this study however, the At GTGs are not predicted to have an Ras GTPase-activating protein domain but do contain a Protein kinase ATP-binding region signature (PROSITE accession, PS00107). This corresponds to the region of both At GTG1 and At GTG2 protein sequences which were predicted to be the ATP-/GTP-binding site identified by Pandey et al. (2009) (Figure 3.4 residues 382 – 411 and 381 – 410 inclusive for At GTG1 and At GTG2 respectively). Furthermore, none of the other studied GTG/GPHR sequences were predicted to contain any annotated motifs including these using PROSITE. The Protein kinase ATP-binding region signature lies within the predicted ABA-GPCR domain; this is residues 282 – 460 and 281 – 459 in At GTG1 and At GTG2 according to Pfam (Bateman et al., 2004).

Table 3.3 Conserved residues in the GTG/GPHRs and any exceptions

Conserved residue	Predicted TM	Domain Location	Exceptions to this conservation		Example residue in At GTG1
			Organism	Base change	
P	3	-	*Trypanosoma cruzi* *Trichinella spiralis, Trypanosoma vivax and Trypanosoma congolense* *Ostreococcus tauri, Ostreococcus lucimarinus, Micromonus pusilla and Micromonus SP*	V A -	P94
G	5	DUF3735 (**G**xxxxxxLSG motif)	*Sorghum bicolour* *Trichinella spiralis* *Trypanosome cruzi* and *Trypanosome congolense,*	- A L	G157
L	5	DUF3735 (**G**xxxxxxLSG motif)	*Trypanosome vivax* *Sorghum bicolour*	V -	L164
S	5	DUF3735 (**G**xxxxxxLSG motif)	*Sorghum bicolour* *Leptosphaeria maculans, Metarhizium cridium, Metarhizium robertsii, Neospora caninum, Neurospora crassa and Phaeosphaeria nodorum*	- A	G165
G	5	DUF3735 (**G**xxxxxxLSG motif)	*Sorghum bicolour* *Zygosaccharamyces rouxii* and *Tetrapisispora phaffil*	- A	G166
Y	8 or 9	ABA-GPCR	-	-	Y401

- = not found in this/these organism(s).

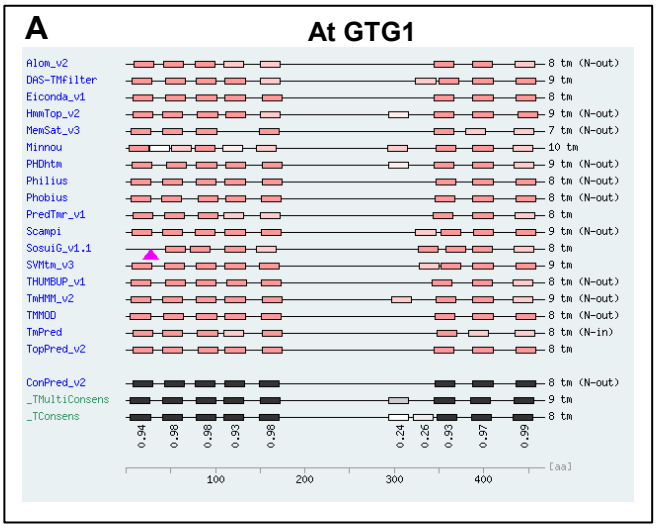
Table 3.4 Domain architecture of proteins containing DUF3735 and/or ABA-GPCR

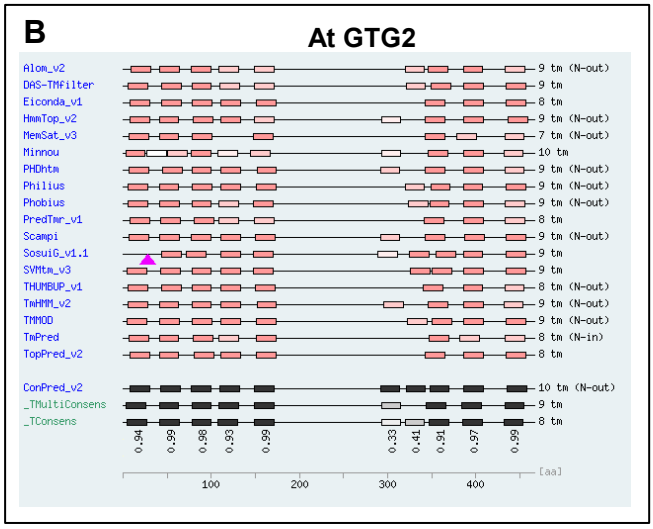
	Domain architecture of proteins containing DUF3735 or ABA-GPCR								
	DUF3735	**ABA-GPCR**	**DUF3735 2xABA-GPCR**	**DUF3735, ABA-GPCR, DUF842**	**DUF3735, ABA-GPCR, PNK3P, AAA-33**	**Homeobox, DUF3735, ABA-GPCR**	**DUF3735, ABA-GPCR, FA-hydroxylase**	**Sel1 x 6, DUF3735, ABA-GPCR**	**ABA-GPCR, Pkinase**
Number of Sequences	14	39	8	1	1	1	1	1	1
Protein classification	Putative uncharacterised or fragments of putative uncharacterised	Putative uncharacterised or fragments of putative uncharacterised	Putative, uncharacterised	Putative, uncharacterised	*Ichthyophthirius multifiliis (strain G5)* Sap DNA-binding domain protein	Putative, uncharacterised	Putative, uncharacterised	Putative, uncharacterised	*Ostreococcus tauri* Serine/threonine protein kinase SAPK4 (ISS)

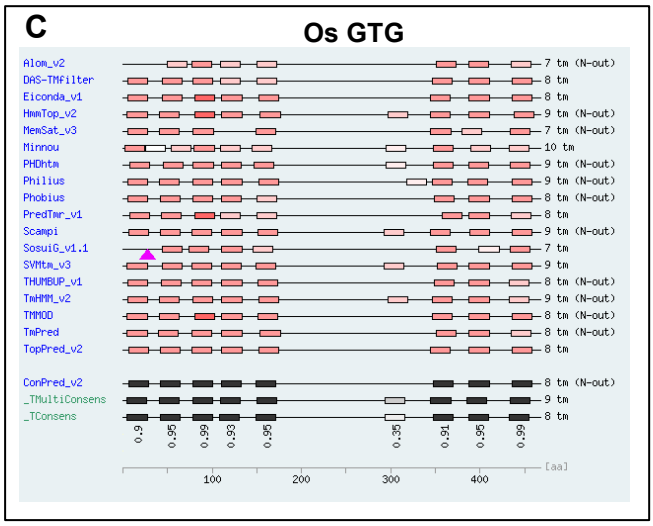
3.2.4 The GTG/GPHRs are predicted to be membrane proteins with similar protein architecture

Several protein topology methods were used to predict the number of TM helices in GTG/GPHRs. ARAMEMNON was used as a resource for the plant GTGs in the first instance. The plant GTGs are predicted to contain seven to ten TM segments (Figure 3.8 A – C). Overall, the methods predict five TM regions within the first ~170 amino acids of sequence, except for MemSat_v3 (Jones et al., 1994), Minnou (Cao et al., 2006), SosuiG_v1.1 (Hirokawa et al., 1998) and Alom_v2 (Nakai and Kanehisa, 1991), which predict four TM domains. SosuiG_v1.1 also predicts a cleavable N-terminal signal peptide in the *O. sativa* GTG (Figure 3.8 A – C). All predictions indicate that there is a large, intracellular portion approximately between residues 170 – 350 (Figure 3.8 A – C). Differences in predictions occur from ~300 amino acids until the end of the sequences, in which most programs predict three or four more TM domains, while there are some showing five more TM domains within this C-terminal region of the protein (Figure 3.8 A – C). The TConsensus prediction, which only takes into account individual predictions, shows that plant GTGs contain eight TM domains (Figure 3.8 A – C). However, nine TM domains are predicted if taking into account other homologous protein predictions projected from a multiple alignment (TMultiConsens). The consensus prediction by ConPred_v2 (ConPredII) is similar to TConsensus (predicted eight TM domains) except for At GTG2, which is predicted to have ten TM regions. According to these predictions the N-terminus is on the outside of the membrane, however the position of the C-terminus differs according to the number of predicted TM helices.

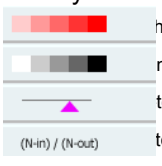
The membrane topology predictions by ConPred II were used to generate 2-dimensional (2D) illustrations of the plant GTGs (Figure 3.10 D – F). However, since ARAMEMNON is a database solely for plant membrane proteins, the animal and fungal GTG/GPHRs were not annotated by the resource. For this reason, membrane topology prediction of the GTG/GPHRs were further analysed by out-sourcing other topology prediction programs (Table 3.5).







Key

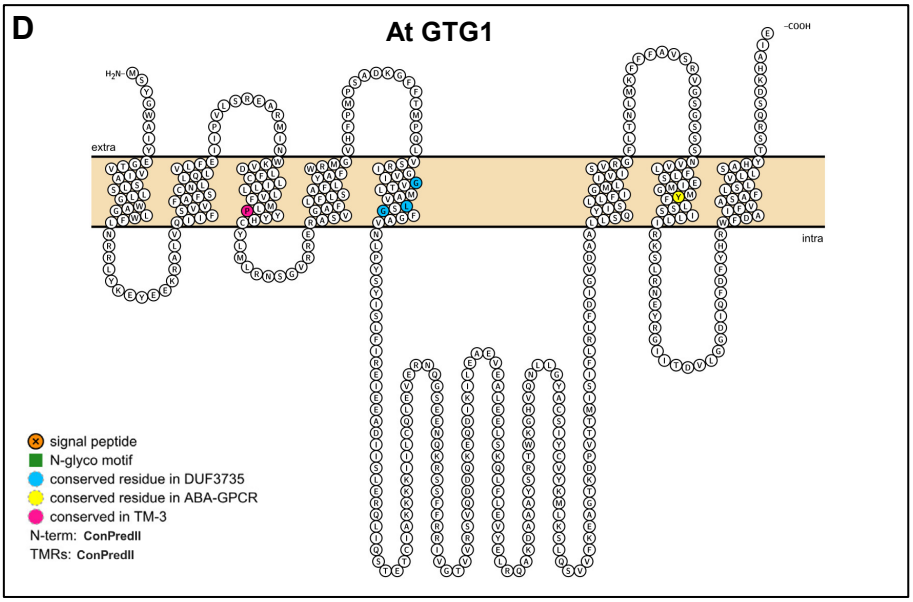


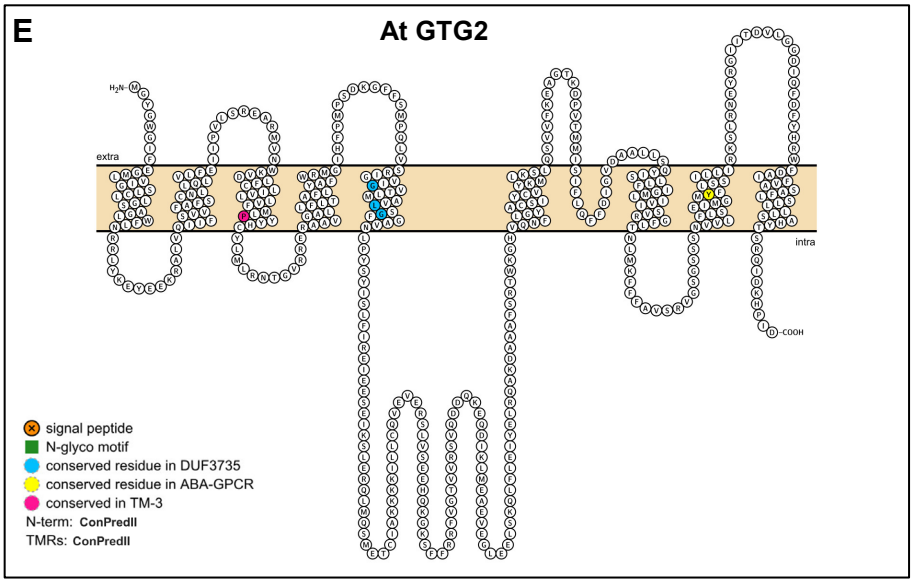
helix TMs: low → high hydrophobicity

nsensus α-helix TMs: weak → strong score

terminal cleavable signal peptide

toplasmic/non-cytoplasmic orientation of N-terminal


D
At GTG1
extra
intra
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: ConPredII
TMRs: ConPredII


E
At GTG2
extra
intra
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: ConPredII
TMRs: ConPredII

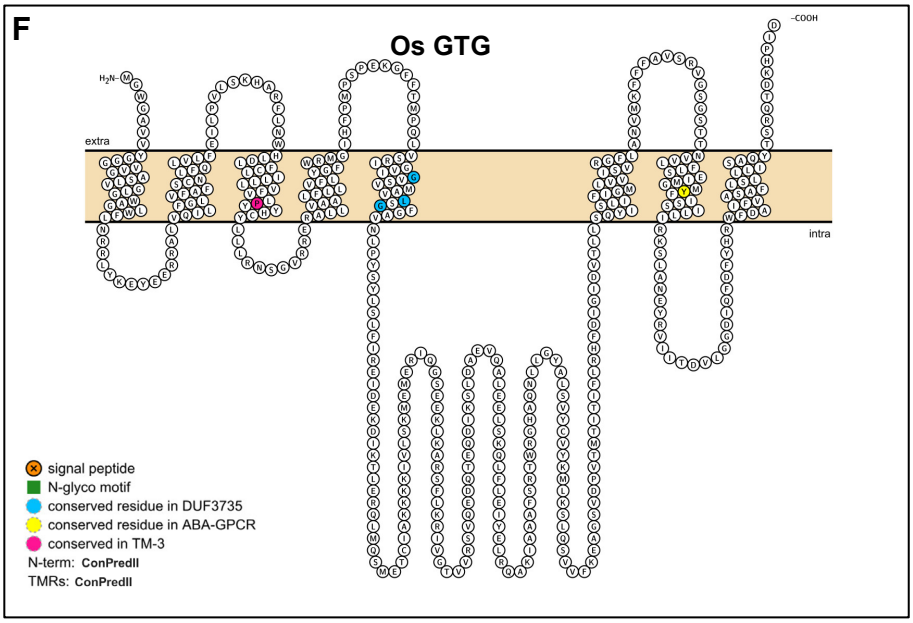

F
Os GTG
extra
intra
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: ConPredII
TMRs: ConPredII

Figure 3.8 Plant GTG/GPHRs contain seven to ten TM domains

The predicted membrane protein topology of the plant GTG/GPHRs: At GTG1 (**A** and **D**), At GTG2 (**B** and **E**) and Os GTG (**C** and **F**). **A** – **C**. Individual and consensus topology prediction methods showing the predicted number of TM domains (red shades) and the orientation of the N-terminus (brackets) using 18 different programs (blue). The consensus membrane topology predictions are shown below the 18 predictions with the scores (black/grey shades). Scale bar shown on figure. Data shown as screenshots obtained from the ARAMEMNON database (Schwacke et al., 2003). **D** – **F**. Membrane protein topology diagrams showing 2D structure of the GTG/GPHRs based on consensus topology prediction using ConPred II (Arai et al., 2004). Diagrams drawn using Protter (Omasits et al., 2013). Highlighted residues are predicted to be in the signal peptide (orange), *N*-glycosylated (N-glyco motif; green), conserved in DUF3735 (blue), conserved in ABA-GPCR domain (yellow) or conserved in the third TM region (TM-3; pink). N-term = N-terminus predicted by the specified method; TMRs = TM regions predicted by the specified method; extra = extracellular and intra = intracellular. These sequences are not predicted to contain a signal sequence.

Table 3.5 The number of TM domains predicted by individual methods

Sequence	Number of transmembrane domains obtained by each method					
	Phobius[1]	**TopPred[2]**	**SOSUI[3]**	**TMHMM[4]**	**TMPred[5]**	**MPEX[6]**
Cel-GPHR-1	9	9	9	9	9	9
Cel-GPHR-2	9	9	9	9	9	9
At GTG1	8	8	9	9	8	9
At GTG2	9	8	10	9	8	9
Sc GTG	7	8	9	8	10	9
Os GTG	8	8	8	9	9	9
Mm GPHR	9	9	9	9	9	9
Cg GPHR	9	9	9	9	9	9
Hs GPHR	9	9	9	9	9	9
Dm GPHR	8	8	8	8	9	9

[1] Phobius: http://phobius.sbc.su.se/index.html (Käll et al., 2004)

[2] TopPred: http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::toppred (Claros and von Heijne, 1994)

[3] SOSUI: http://harrier.nagahama-i-bio.ac.jp/sosui/sosui_submit.html (Hirokawa et al. (1998), Mitaku and Hirokawa, 1999)

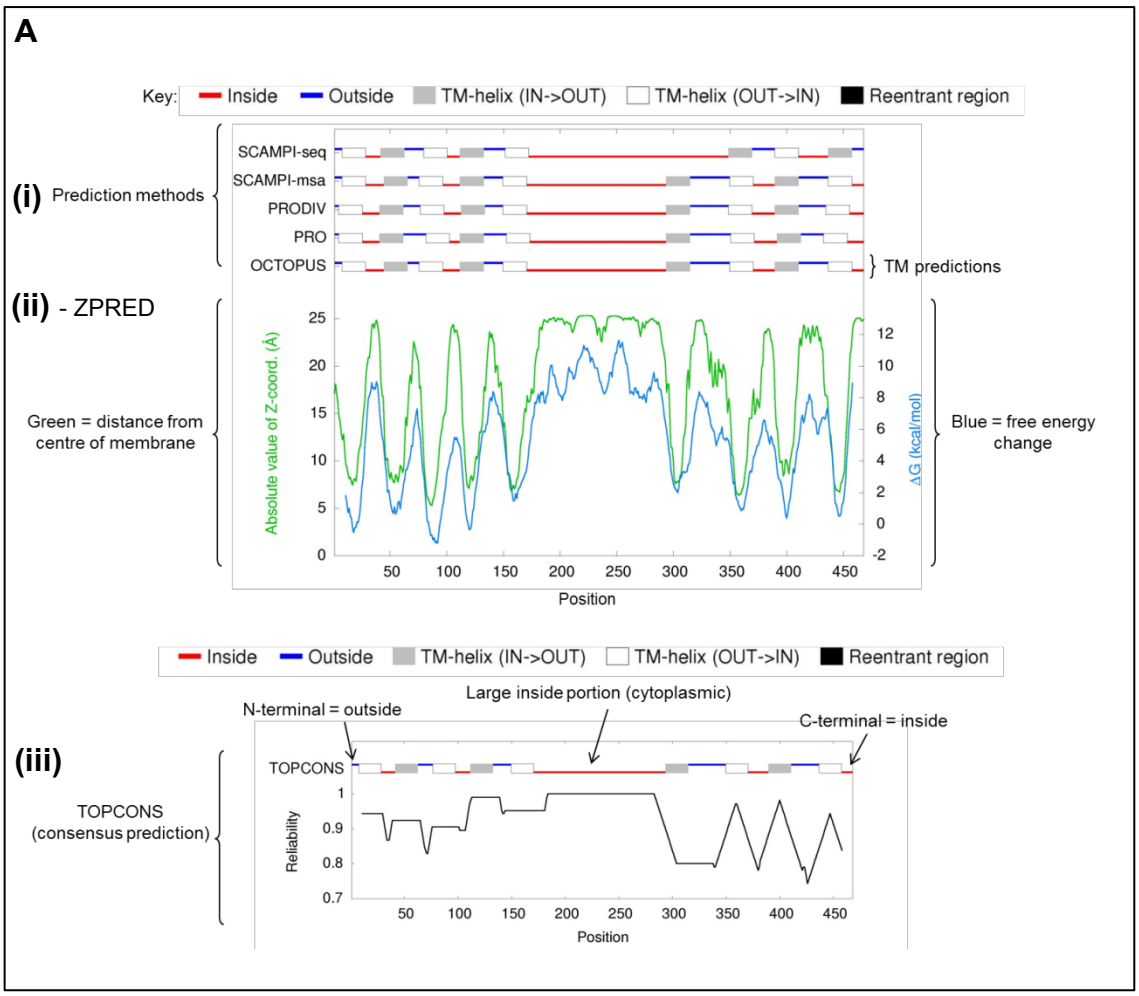
[4] TMHMM: http://www.cbs.dtu.dk/services/TMHMM/ (Sonnhammer et al., 1998, Krogh et al. (2001))

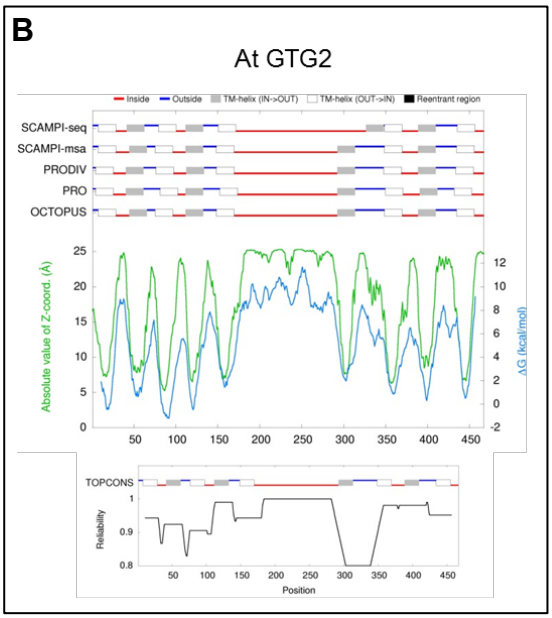
[5] TMPred: http://www.ch.embnet.org/software/TMPRED_form.html (Hofmann, 1993)

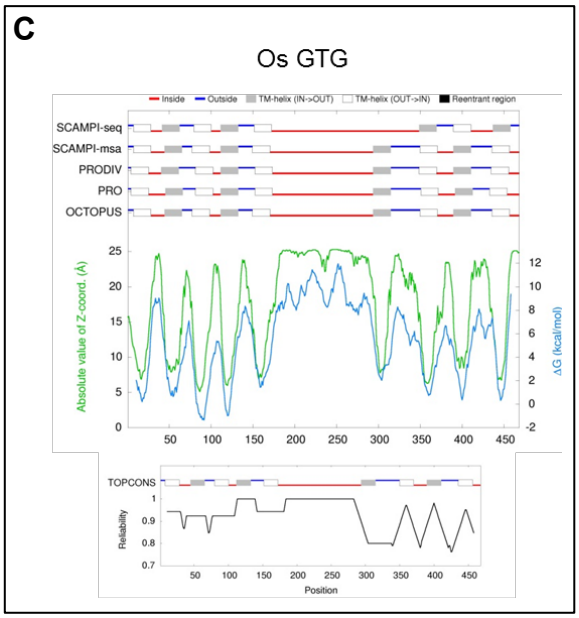
[6] MPEX: http://blanco.biomol.uci.edu/mpex/ (Snider et al., 2009)

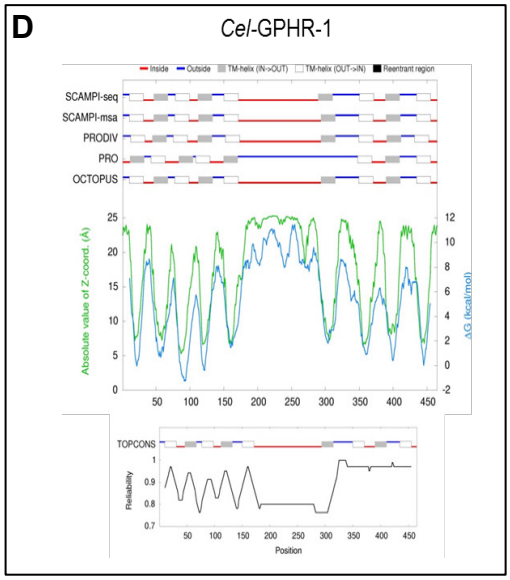
Ultimately, the GTG/GPHRs were analysed using the consensus membrane topology web server TOPCONS because only plant proteins are annotated in ARAMEMNON and TOPCONS is widely used by the community for both prokaryotic and eukaryotic membrane protein predictions (Bernsel et al., 2009). The results indicate that according to this consensus prediction, the GTG/GPHRs are generally predicted to contain nine TM domains (Figure 3.9). Figure 3.9 A shows an example output from TOPCONS for At GTG1. Similar to the analysis of the plant GTGs using ARAMEMNON, the topology predictions using individual methods (Figure 3.9 (i)) generally indicated that the GTG/GPHRs contain five TM segments followed by a large intracellular portion (depending on the N-terminus location) followed by either three to five TM segments (with one exception described below). The ZPRED results further illustrate this, showing the change in free energy (ΔG) in blue and distance from the centre of the membrane (absolute value of Z-coordinates) in green across the protein sequences (Figure 3.9 (ii)). The PRO/PRODIV-TMHMM method predicts that the Sc GTG contains only five TM domains; however this outcome is only predicted once in all the various methods used (Figure 3.9 I). The animal GTG/GPHRs demonstrate fewer differences between the numbers of TM domains predicted by individual programs than the plant and fungal GTG/GPHRs. In individual and consensus predictions the animal GTG/GPHRs are consistently predicted to have nine TM regions except for Dm GPHR, which is mostly predicted to contain eight (Table 3.5 and Figure 3.9). To show that the GTG/GPHRs have similar TM predictions, a multiple alignment of the plant At GTGs and Os GTG was conducted with the predicted consensus topology predictions by TOPCONS highlighted (Figure 3.10). This alignment indicates the predicted topologies are very similar.

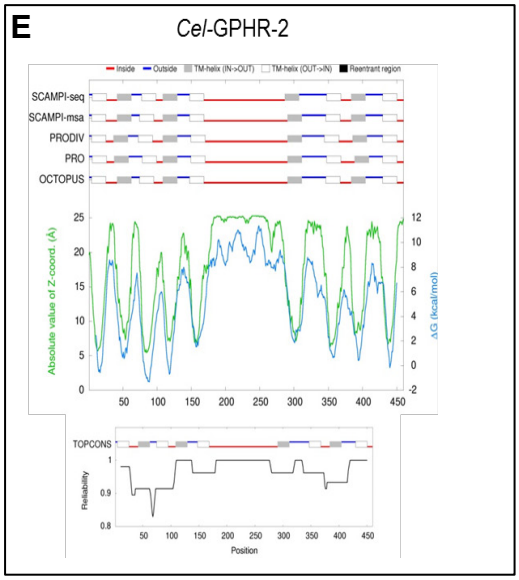
PredictProtein was used to predict other structural annotations based on prediction methods. Figure 3.11 A shows an example of the output given for At GTG1. The GTG/GPHRs are predicted to contain many regions of disorder (yellow; Figure 3.11); these cover the regions not predicted to be TM domains (purple; Figure 3.11). Moreover, the predicted secondary structures (red; Figure 3.11) coincide with the predicted TM regions (purple; Figure 3.11). Also predicted are a few (1 – 3) DNA- and many (7 – 11) protein-binding sites throughout the GTG/GPHR sequences in regions predicted to be exposed/not transmembrane (Figure 3.11). These binding sites

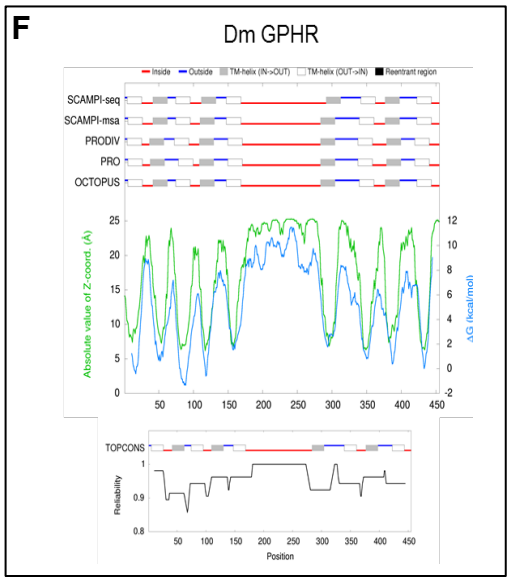

A
Key:
Inside
Outside
TM-helix (IN->OUT)
TM-helix (OUT->IN)
Reentrant region
(i) Prediction methods
SCAMPI-seq
SCAMPI-msa
PRODIV
PRO
OCTOPUS
TM predictions
(ii) - ZPRED
Green = distance from centre of membrane
Absolute value of Z-coord. (Å)
ΔG (kcal/mol)
Blue = free energy change
Position
(iii)
TOPCONS (consensus prediction)
N-terminal = outside
Large inside portion (cytoplasmic)
C-terminal = inside
TOPCONS
Reliability
Position

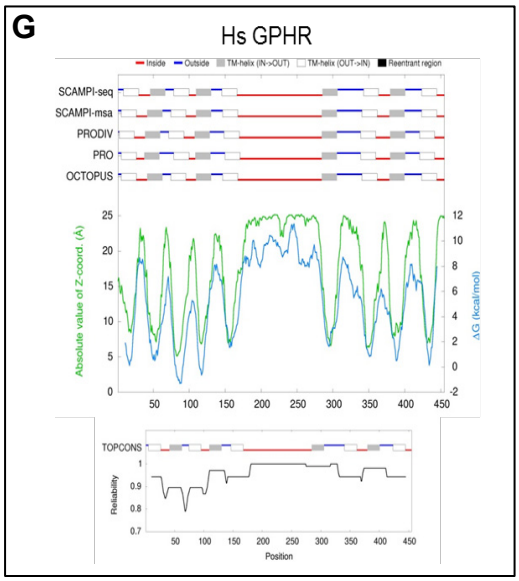

B
At GTG2
SCAMPI-seq
SCAMPI-msa
PRODIV
PRO
OCTOPUS
TOPCONS
Reliability
Position

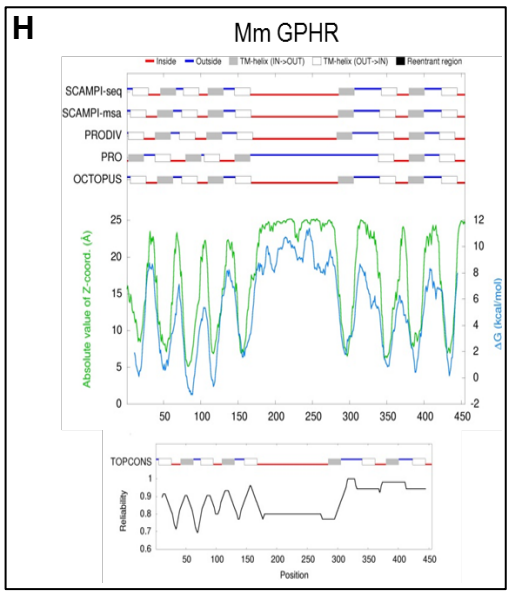

C
Os GTG
SCAMPI-seq
SCAMPI-msa
PRODIV
PRO
OCTOPUS
TOPCONS
Reliability
Position

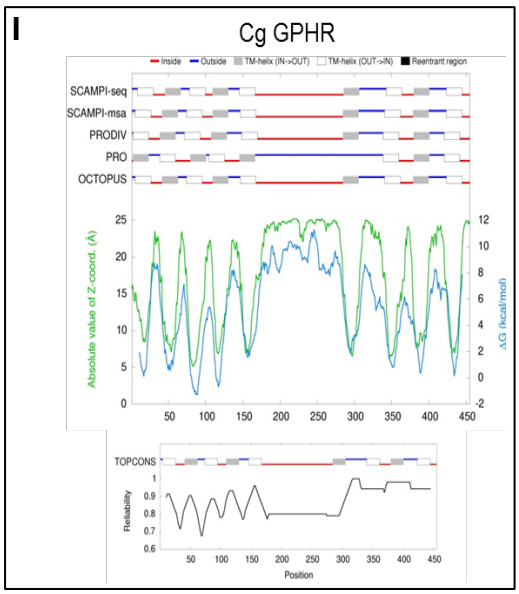

D
Cel-GPHR-1
Inside
Outside
TM-helix (IN->OUT)
TM-helix (OUT->IN)
Reentrant region
SCAMPI-seq
SCAMPI-msa
PRODIV
PRO
OCTOPUS
Absolute value of Z-coord. (Å)
ΔG (kcal/mol)
TOPCONS
Reliability
Position


E
Cel-GPHR-2
Inside
Outside
TM-helix (IN->OUT)
TM-helix (OUT->IN)
Reentrant region
SCAMPI-seq
SCAMPI-msa
PRODIV
PRO
OCTOPUS
Absolute value of Z-coord. (Å)
ΔG (kcal/mol)
TOPCONS
Reliability
Position


F
Dm GPHR
Inside
Outside
TM-helix (IN->OUT)
TM-helix (OUT->IN)
Reentrant region
SCAMPI-seq
SCAMPI-msa
PRODIV
PRO
OCTOPUS
Absolute value of Z-coord. (Å)
ΔG (kcal/mol)
TOPCONS
Reliability
Position


G
Hs GPHR
Inside
Outside
TM-helix (IN->OUT)
TM-helix (OUT->IN)
Reentrant region
SCAMPI-seq
SCAMPI-msa
PRODIV
PRO
OCTOPUS
Absolute value of Z-coord. (Å)
ΔG (kcal/mol)
TOPCONS
Reliability
Position


H
Mm GPHR
Inside
Outside
TM-helix (IN->OUT)
TM-helix (OUT->IN)
Reentrant region
SCAMPI-seq
SCAMPI-msa
PRODIV
PRO
OCTOPUS
Absolute value of Z-coord. (Å)
ΔG (kcal/mol)
TOPCONS
Reliability
Position


I
Cg GPHR
Inside
Outside
TM-helix (IN->OUT)
TM-helix (OUT->IN)
Reentrant region
SCAMPI-seq
SCAMPI-msa
PRODIV
PRO
OCTOPUS
Absolute value of Z-coord. (Å)
ΔG (kcal/mol)
TOPCONS
Reliability
Position

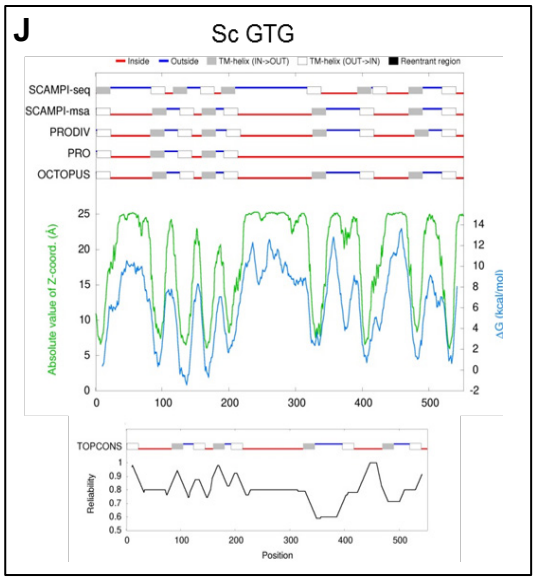


Figure 3.9 The consensus topology predictions for the GTG/GPHRs contain 9-TM domains

Topologies predicted by individual methods (**i**), predicted distance from the centre of the membrane, the Z-coordinates (green); and predicted free energy change, ΔG-values (blue) using the ZPRED method (**ii**); the consensus prediction and reliability score by TOPCONS method (**iii**) for: At GTG1 (**A**), At GTG2 (**B**), Os GTG (**C**), *Cel*-GPHR-1 (**D**), *Cel*-GPHR-2 (**E**), Dm GPHR (**F**), Hs GPHR (**G**), Mm GPHR (**H**), Cg GPHR (**I**) and Sc GTG (**J**). **A**. Annotated example showing the outputs for each method of topology prediction(s) in different sections (i – iii).

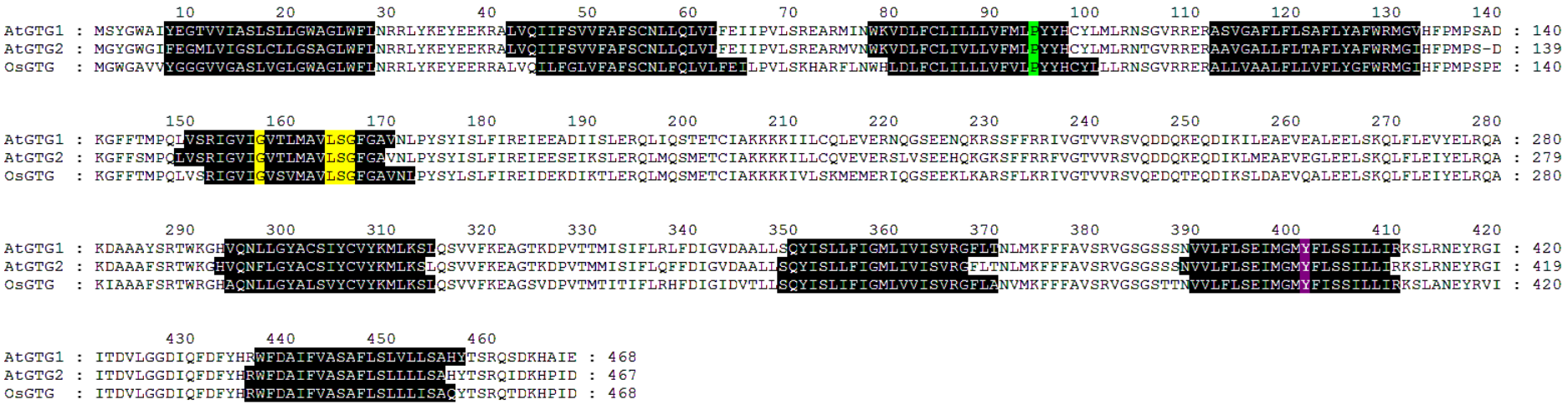


Figure 3.10 The plant GTG/GPHRs have similar predicted membrane topology

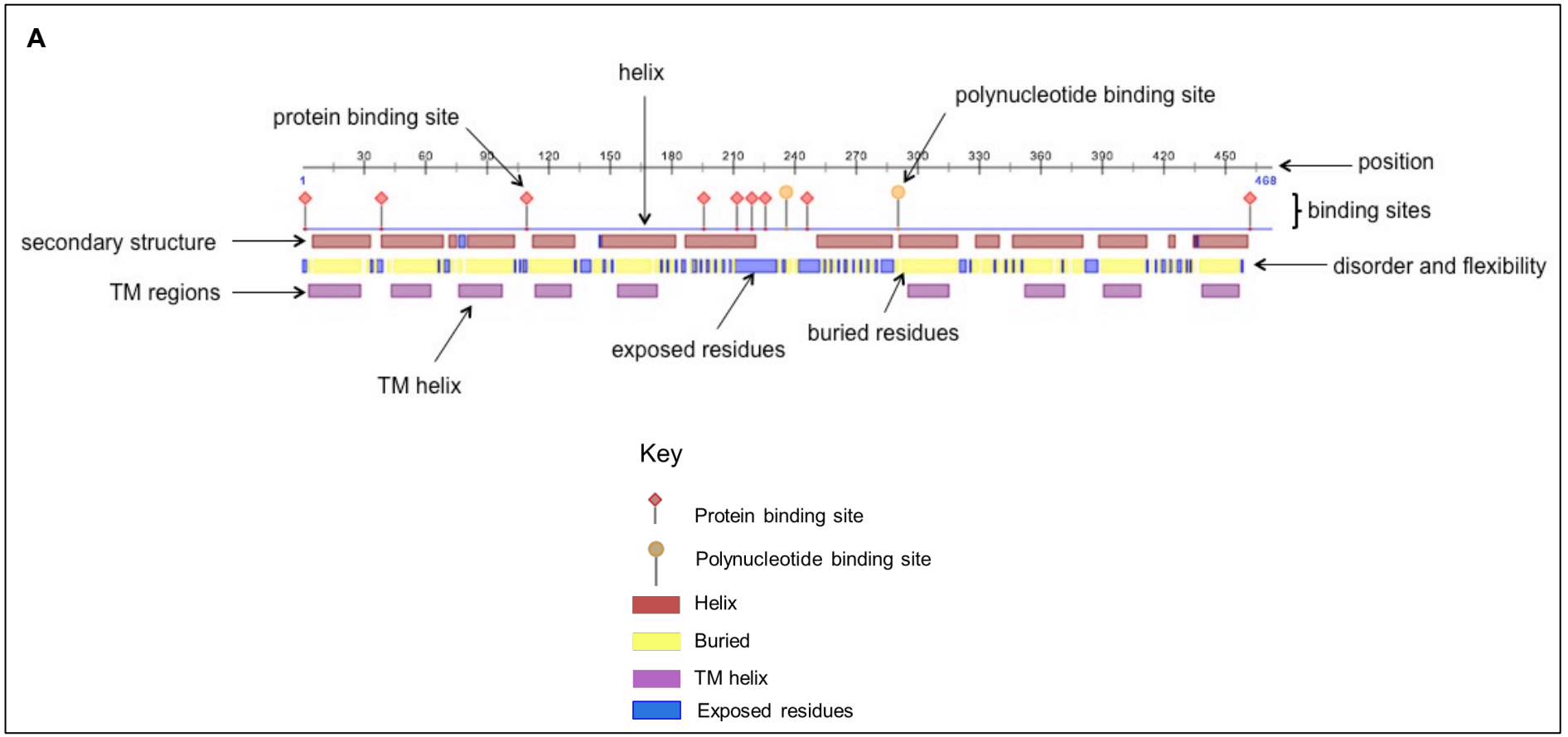
Protein sequence alignment of the GTG/GPHRs in *A. thaliana* (AtGTG1 and AtGTG2) and *O. sativa* (OsGTG) showing the predicted consensus transmembrane regions (black) using TOPCONS method. Sequences were aligned using the ClustalW (Thompson et al., 1994) and presented using GeneDoc (Nicholas and Nicholas Jr, 1997). Black = predicted TM regions, conserved residues found in all GTG/GPHRs (Table 3.3) are also highlighted: green = proline in predicted third TM region, yellow = in the DUF3735 or purple = in ABA-GPCR domain.

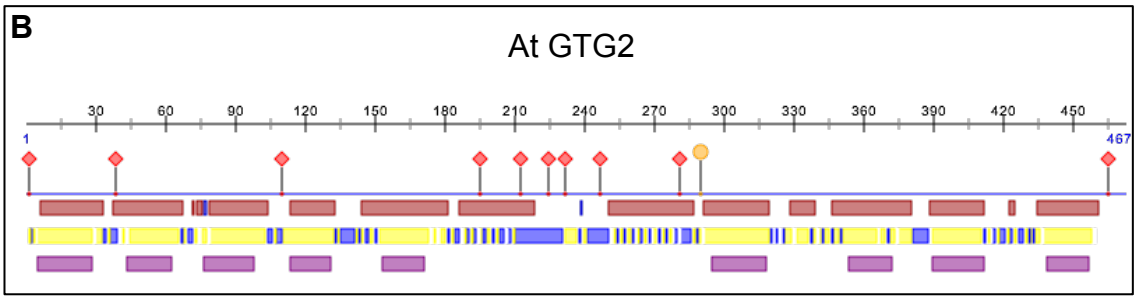
are more concentrated in the predicted large intracellular segment between TM5 and the sixth TM segment (TM6) (Figure 3.11).

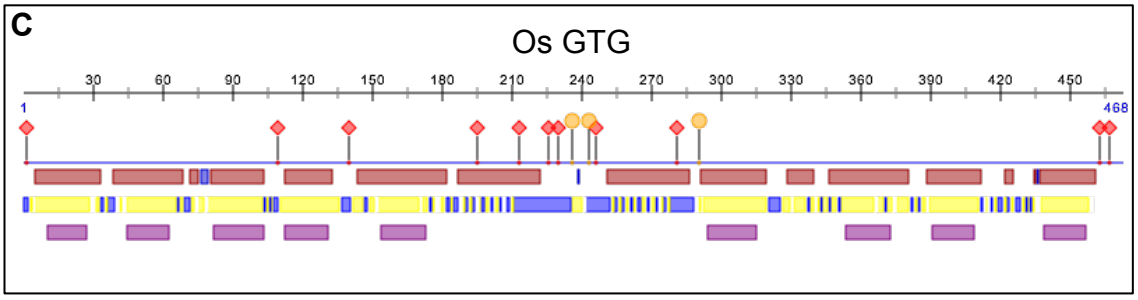
For instance, At GTG1 is predicted to contain nine protein-binding sites, with five in the large intracellular region (between TM5 – TM6) and two DNA-binding sites also in this region (Figure 3.11 A). *Cel*-GPHR-2 is predicted to contain many protein-binding sites but no DNA-binding sites (Figure 3.11 E). All GTG/GPHRs have predicted protein-binding sites in the N- and/or C-terminus (Figure 3.11).

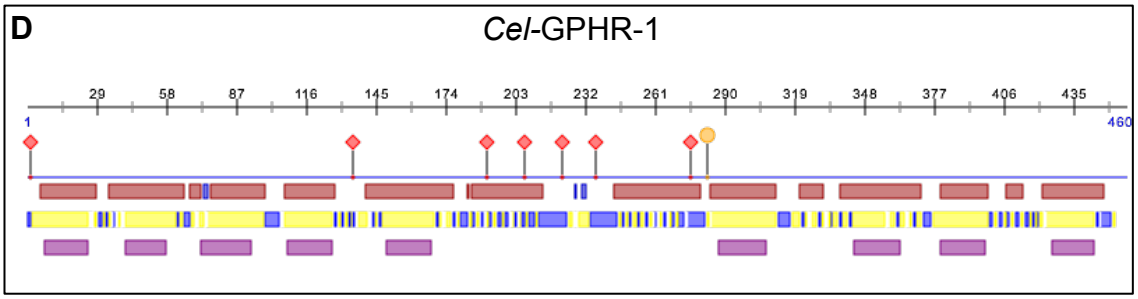
SignalP is a web server (http://www.cbs.dtu.dk/services/SignalP/) used for predicting the presence of a cleavable signal peptide of potential secretory proteins (Petersen et al., 2011). The first 70 residues of the *A. thaliana*, *O. sativa*, *S. cerevisiae*, *C. elegans*, *H. sapiens*, *M. musculus* and *C. griseus* GTG/GPHR protein sequences were analysed to see if they contain signal peptides for secretory targeting. The results indicate that Sc GTG contains a signal peptide that is cleaved between position 17 and 18 (Figure 3.12 A), while signal peptides were not predicted for the other GTG/GPHRs analysed. Figure 3.12 B demonstrates an example output obtained from the SignalP web server demonstrating this for *Cel*-GPHR-1. Furthermore, Phobius, a combined signal peptide prediction and membrane topology prediction method (Käll et al., 2004), also identified *S. cerevisiae* containing a signal peptide sequence whilst the others did not (Figure 3.13 J).

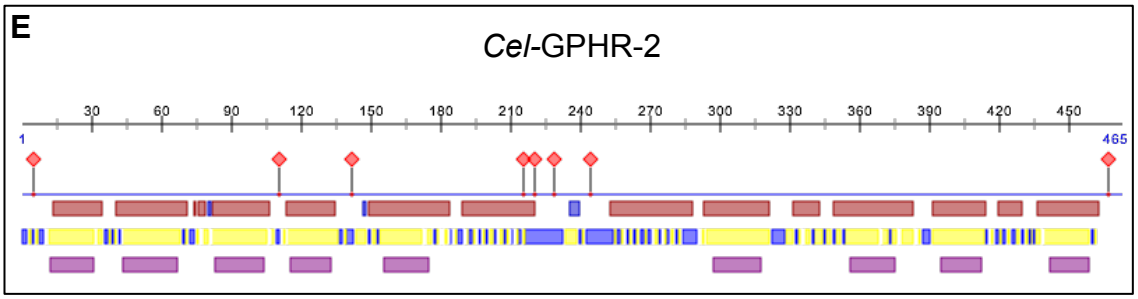
To visualise the predicted membrane topologies, conserved residues in the DUF3735 and ABA-GPCR domains and predicted signal peptide (for Sc GTG), membrane topology diagrams were drawn to illustrate the predicted, consensus topologies for the GTG/GPHRs (Figure 3.13). Phobius and SignalP methods predicted Sc GTG to contain a signal peptide. The TOPCONS and Phobius methods indicate that the predictions are quite different compared to one another. TOPCONS predicts Sc GTG to contain nine TM helices while Phobius only predicts seven (Figure 3.13).

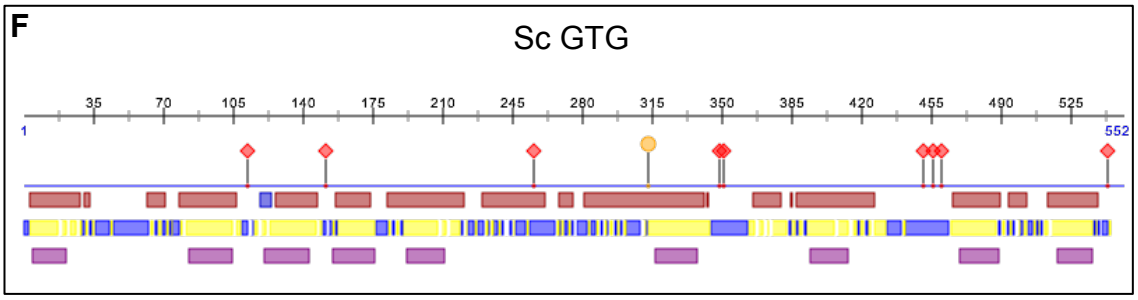

A
helix
protein binding site
polynucleotide binding site
30
60
90
120
150
180
210
240
270
300
330
360
390
420
450
position
1
468
binding sites
secondary structure
disorder and flexibility
TM regions
TM helix
exposed residues
buried residues
Key
Protein binding site
Polynucleotide binding site
Helix
Buried
TM helix
Exposed residues


B
At GTG2
30 60 90 120 150 180 210 240 270 300 330 360 390 420 450
1
467


C
Os GTG
30 60 90 120 150 180 210 240 270 300 330 360 390 420 450
1
468


D
Cel-GPHR-1
29 58 87 116 145 174 203 232 261 290 319 348 377 406 435
1
460


E
Cel-GPHR-2
30 60 90 120 150 180 210 240 270 300 330 360 390 420 450
1
465


F
Sc GTG
35 70 105 140 175 210 245 280 315 350 385 420 455 490 525
1
552

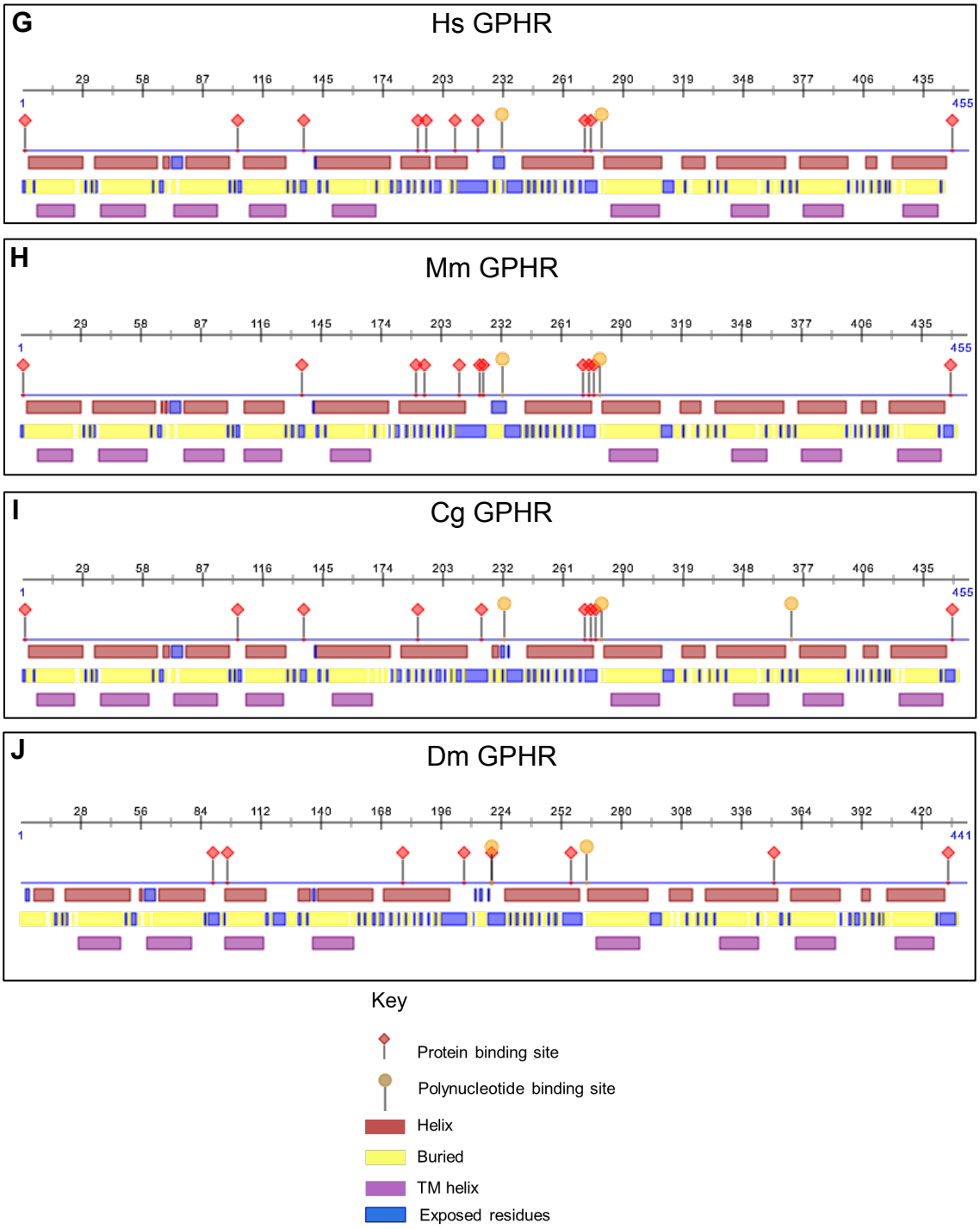


Figure 3.11 Structural annotations predicted for the GTG/GPHRs

The GTG/GPHRs contain many predicted regions of disorder (blue), secondary structure (dark red) including transmembrane domains (purple) and binding sites for DNA (polynucleotide) and protein-protein interaction sites, as predicted by PredictProtein (Yachdav et al., 2014). **A**. At GTG1. Annotated example of the structural elements predicted. **B**. At GTG2. **C**. Os GTG. **D**. *Cel*-GPHR-1. **E**. *Cel*-GPHR-2. **F**. Sc GTG. **G**. Hs GPHR **H**. Mm GPHR. **I**. Cg GPHR. **J**. Dm GPHR.

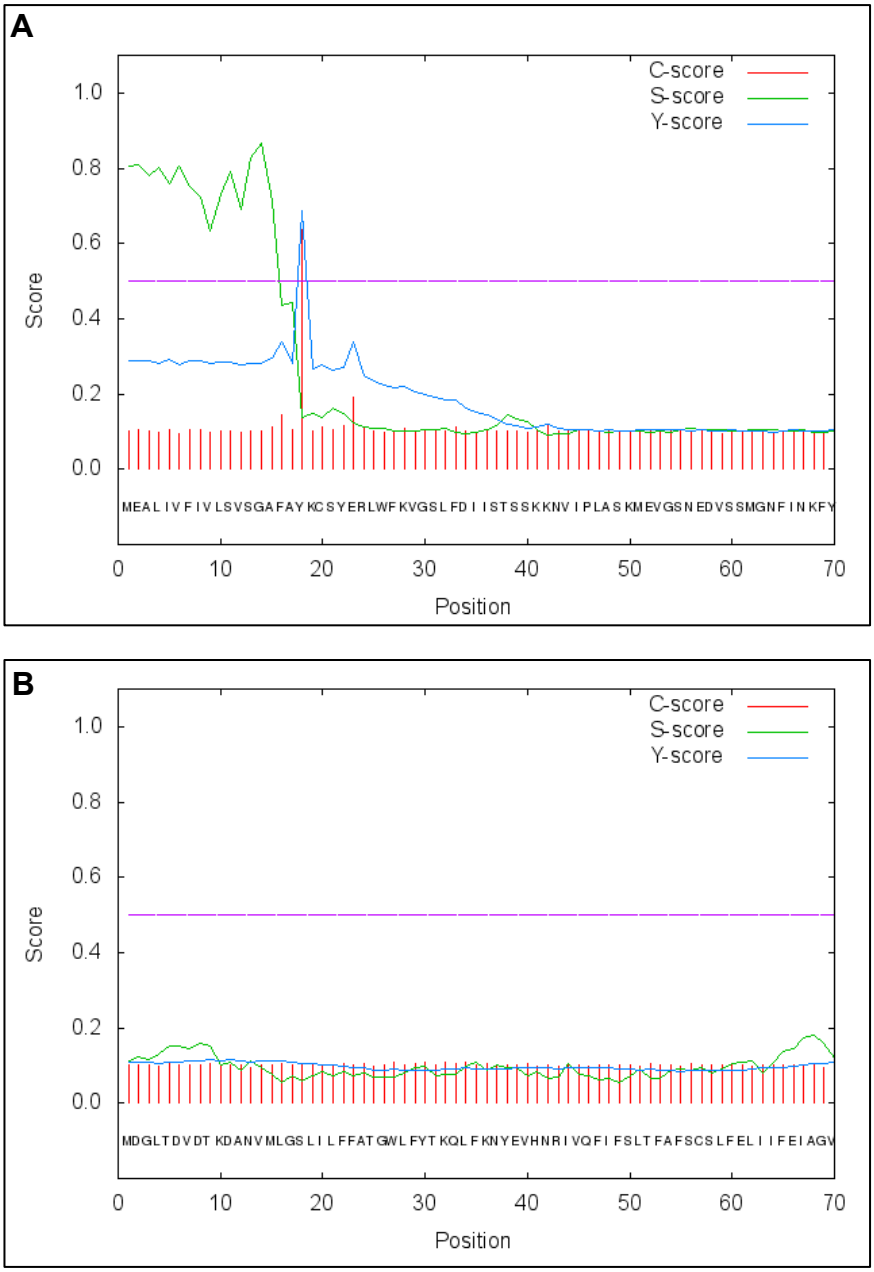
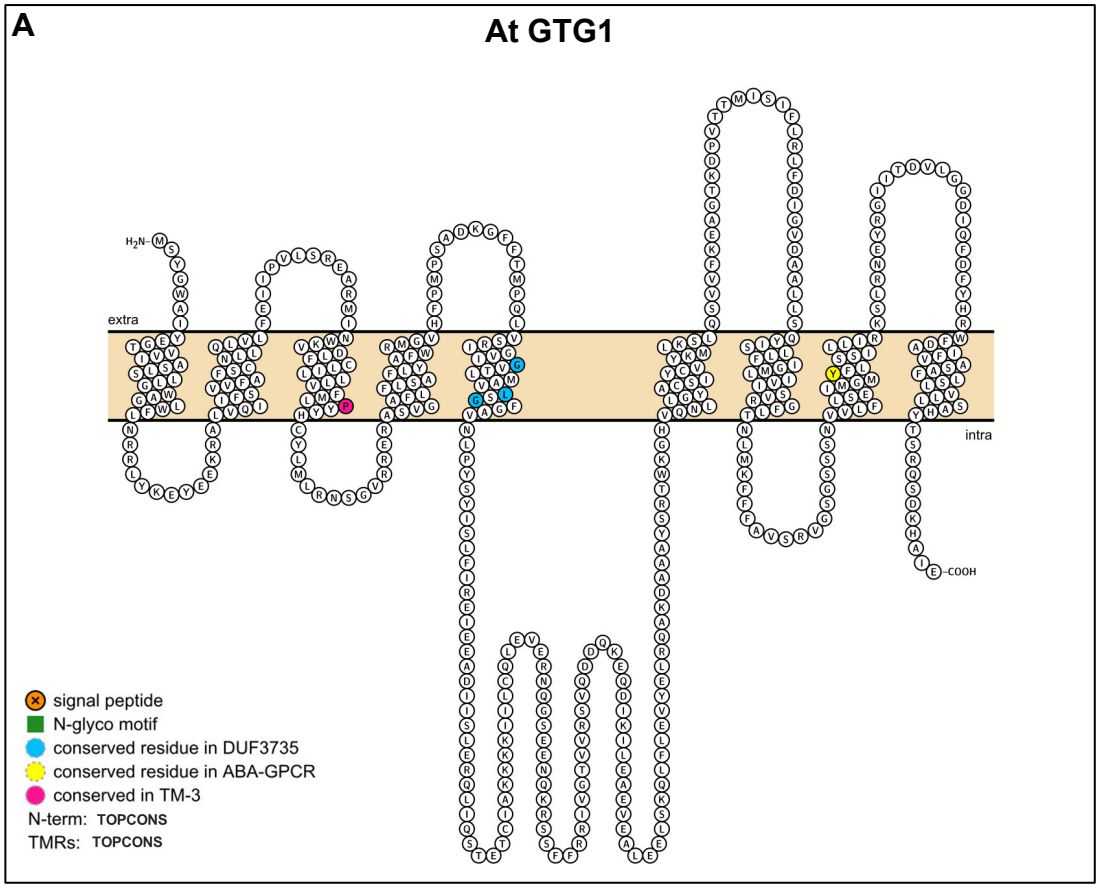
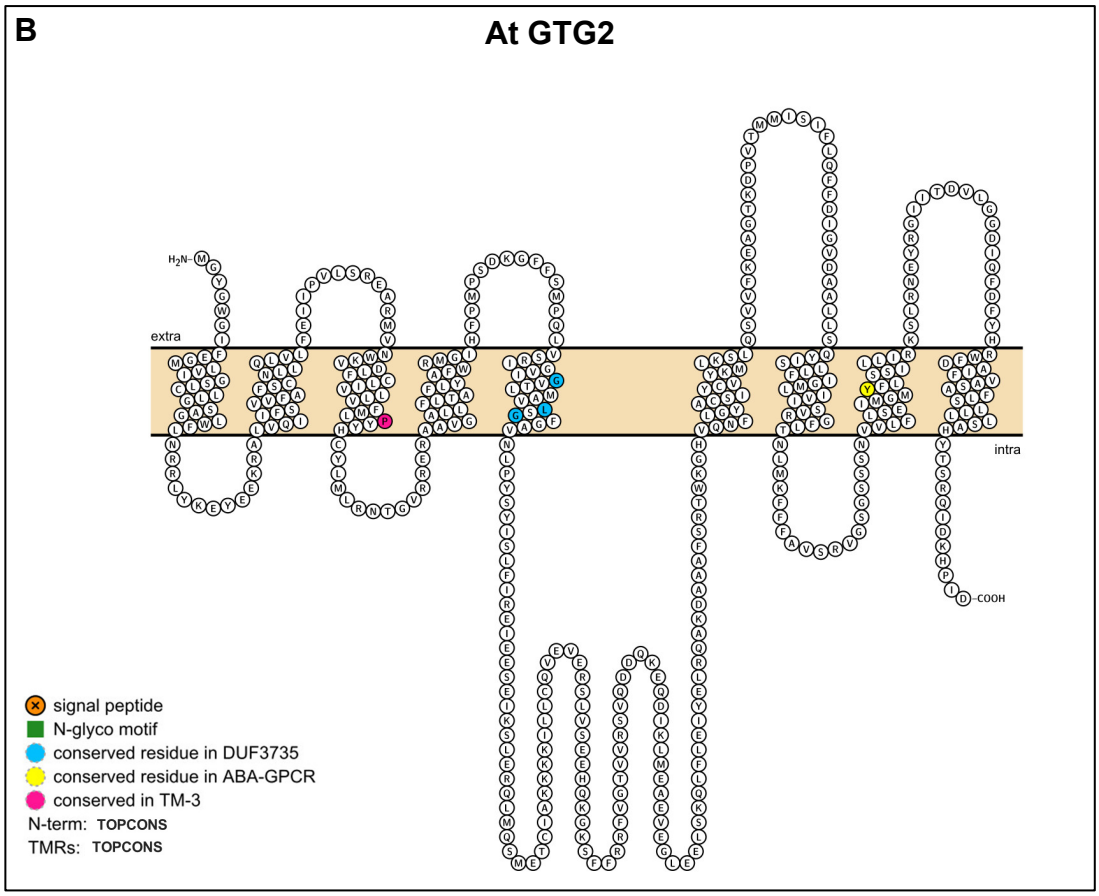
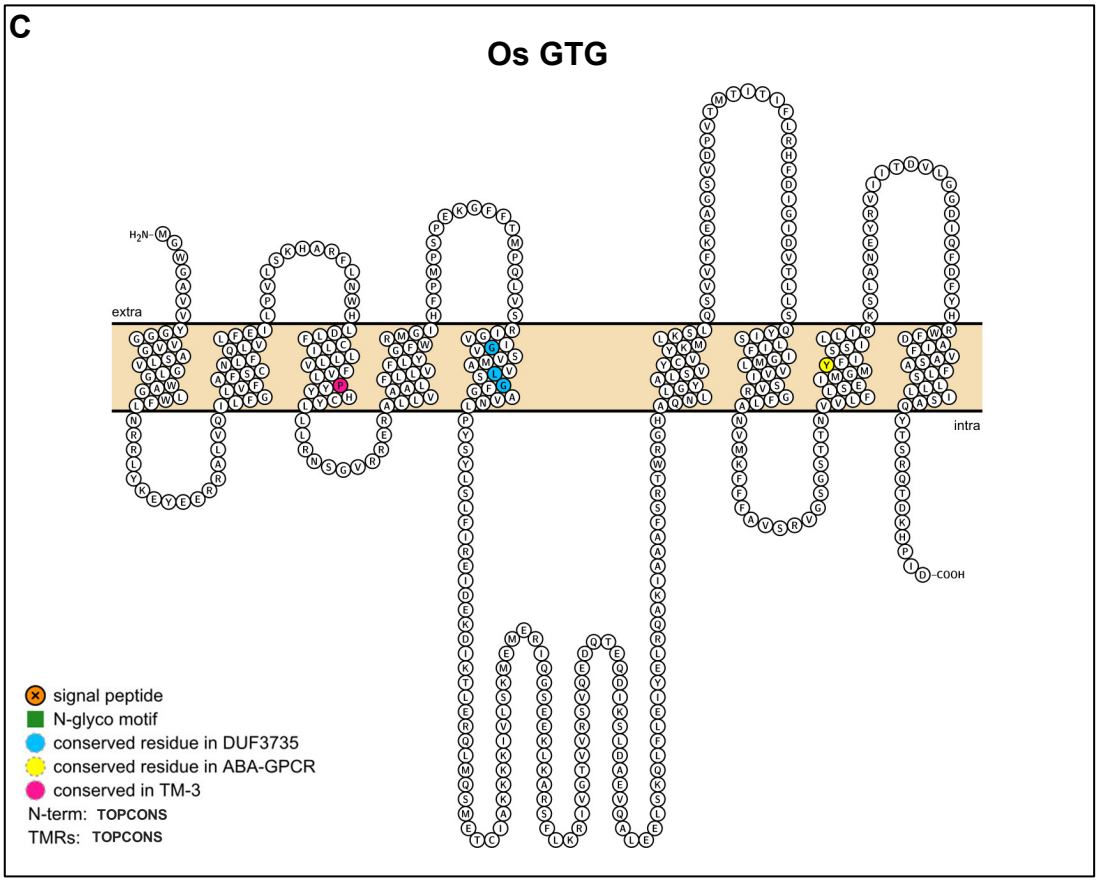


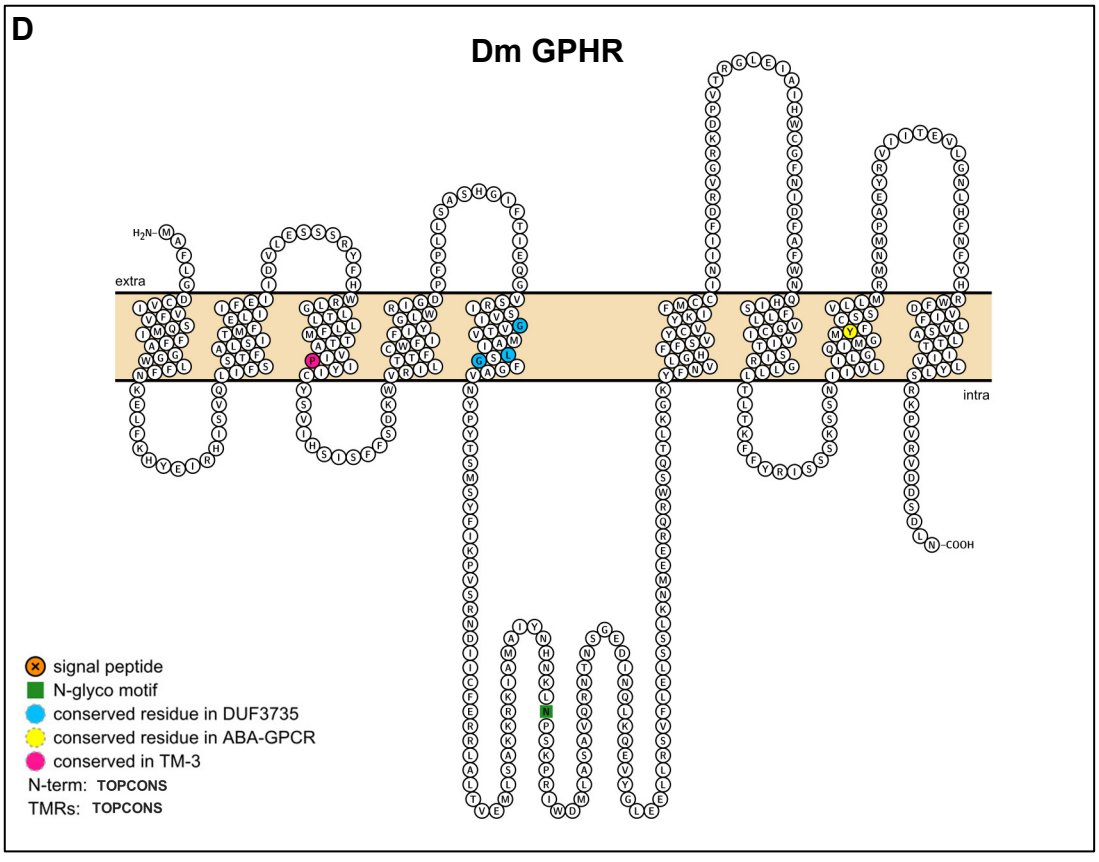
Figure 3.12 *S. cerevisiae* GTG is predicted to contain a signal peptide

Signal peptide prediction prediction showing the presence (**A**) and absence (**B**) of a signal peptide in the first 70 amino acids of Sc GTG (**A**) and *Cel*-GPHR-1 (**B**) sequences respectively, using SignalP (Petersen et al., 2011). **A**. Signal sequence predicted in Sc GTG. Maximum C value: position 18 (0.636), maximum Y value (0.688) at position 18, maximum S value (0.866) at position 14 and mean S value (0.729) mean D value (0.709) between residues 1 – 17.

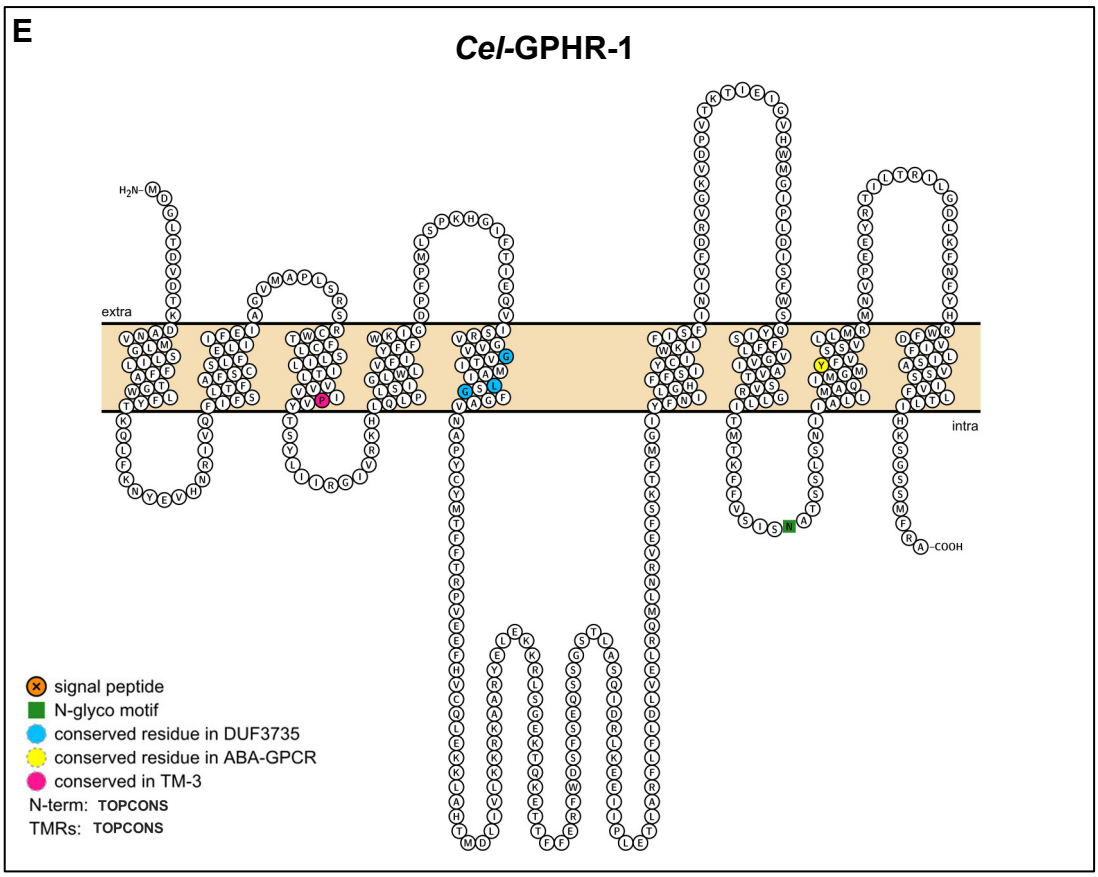

A
At GTG1
extra
intra
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: TOPCONS
TMRs: TOPCONS


B
At GTG2
extra
intra
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: TOPCONS
TMRs: TOPCONS

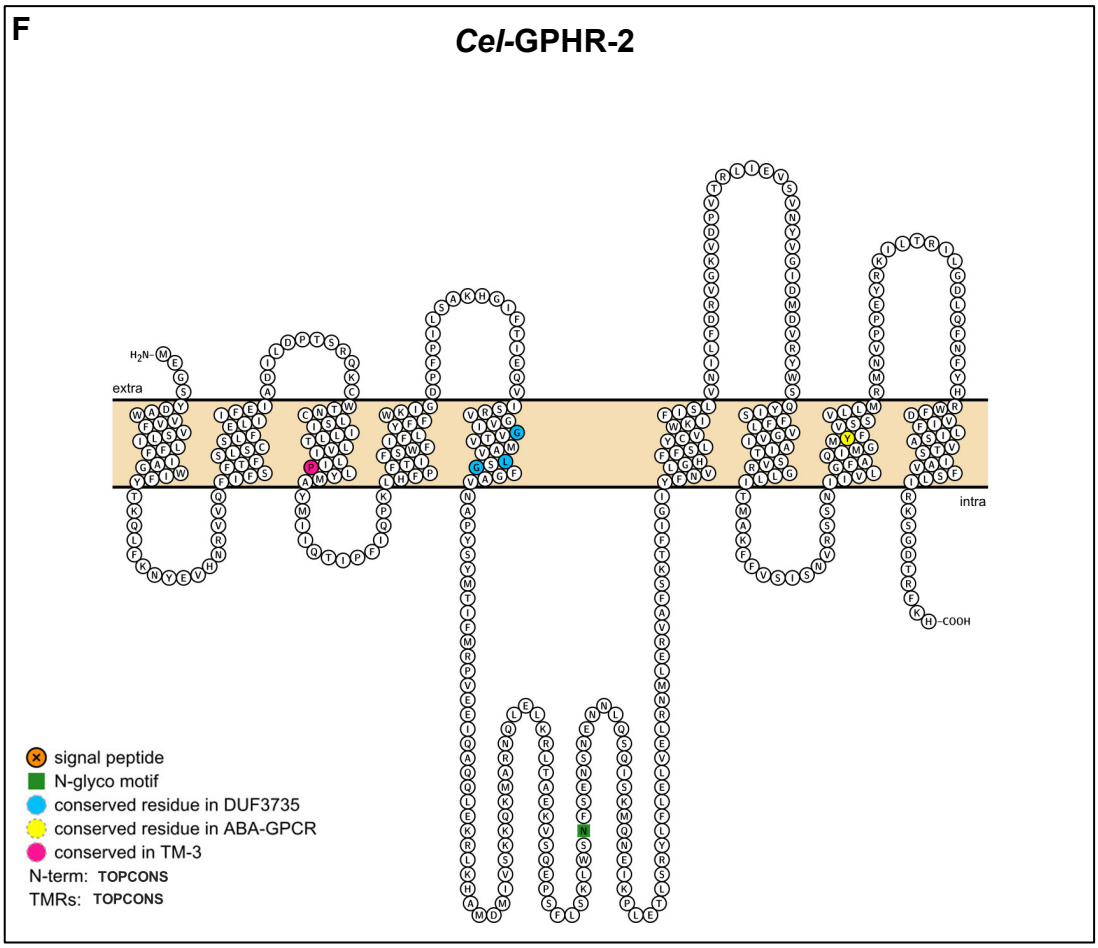

C
Os GTG
extra
intra
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: TOPCONS
TMRs: TOPCONS

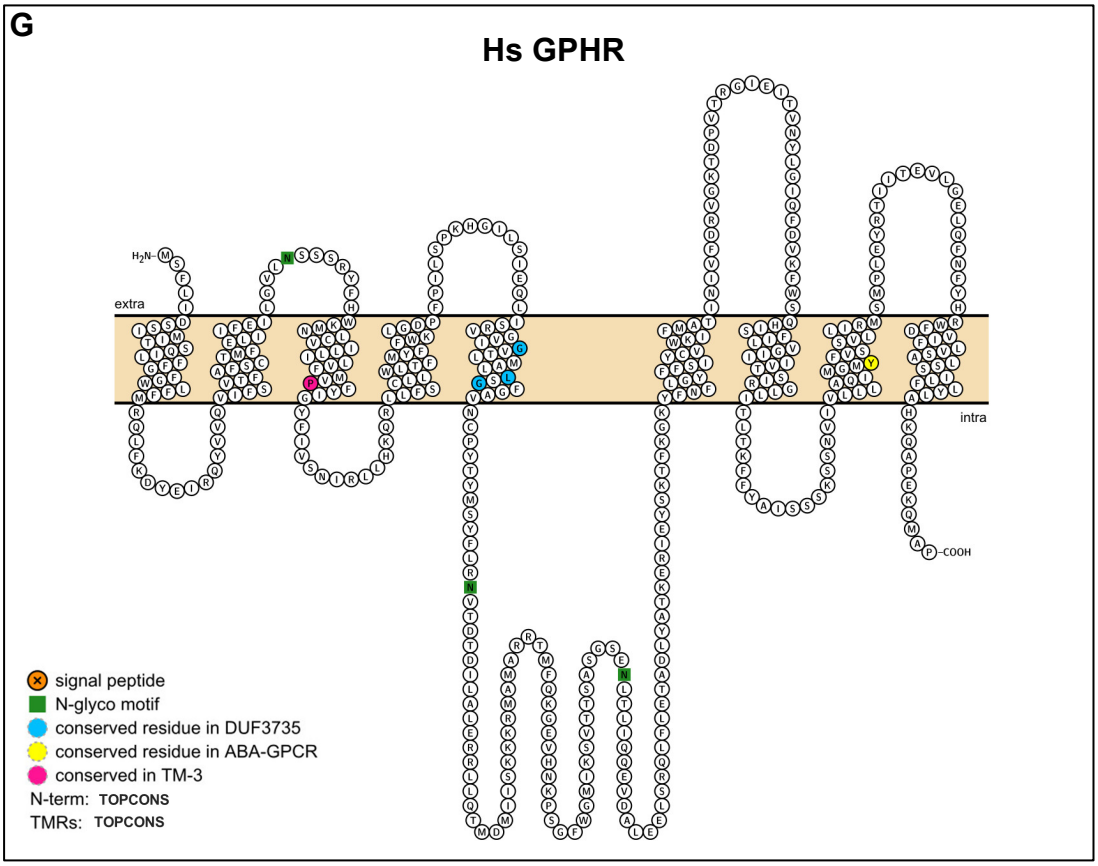

D
Dm GPHR
extra
intra
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: TOPCONS
TMRs: TOPCONS

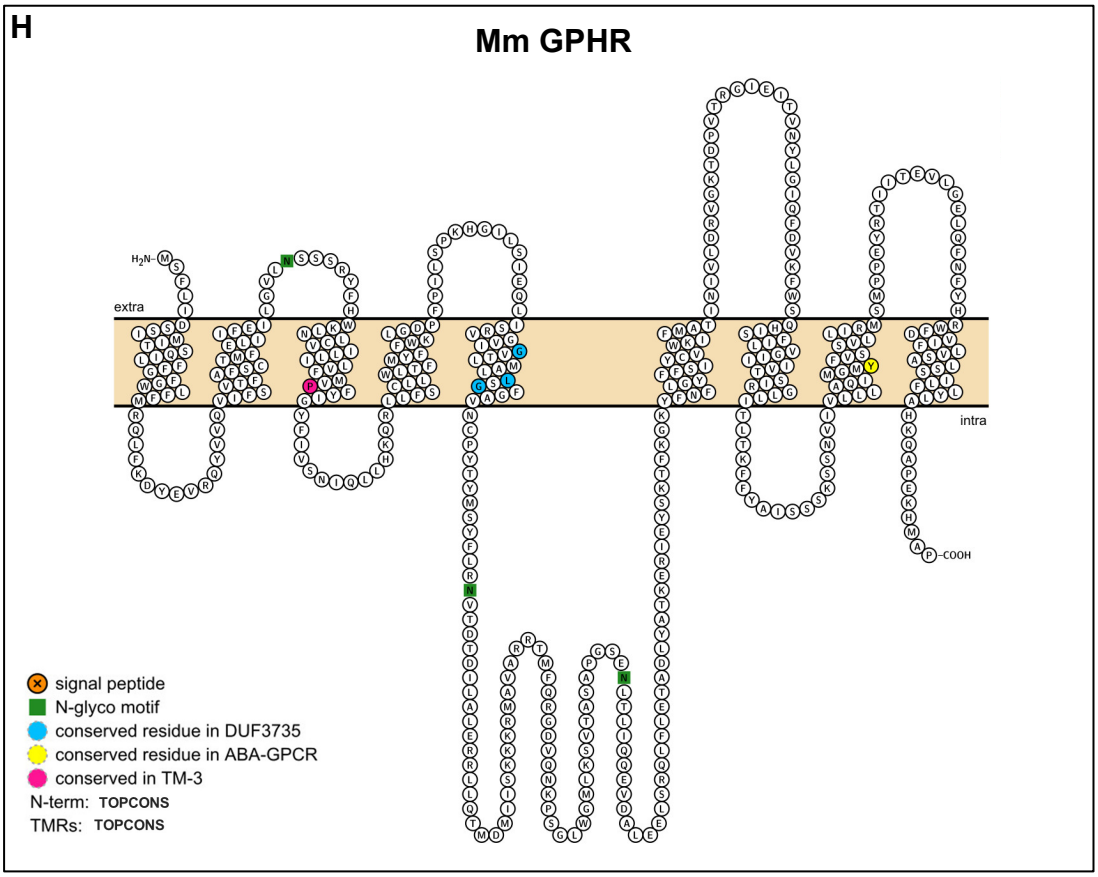
E


Cel-GPHR-1
H2N-
extra
intra
-COOH
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: TOPCONS
TMRs: TOPCONS

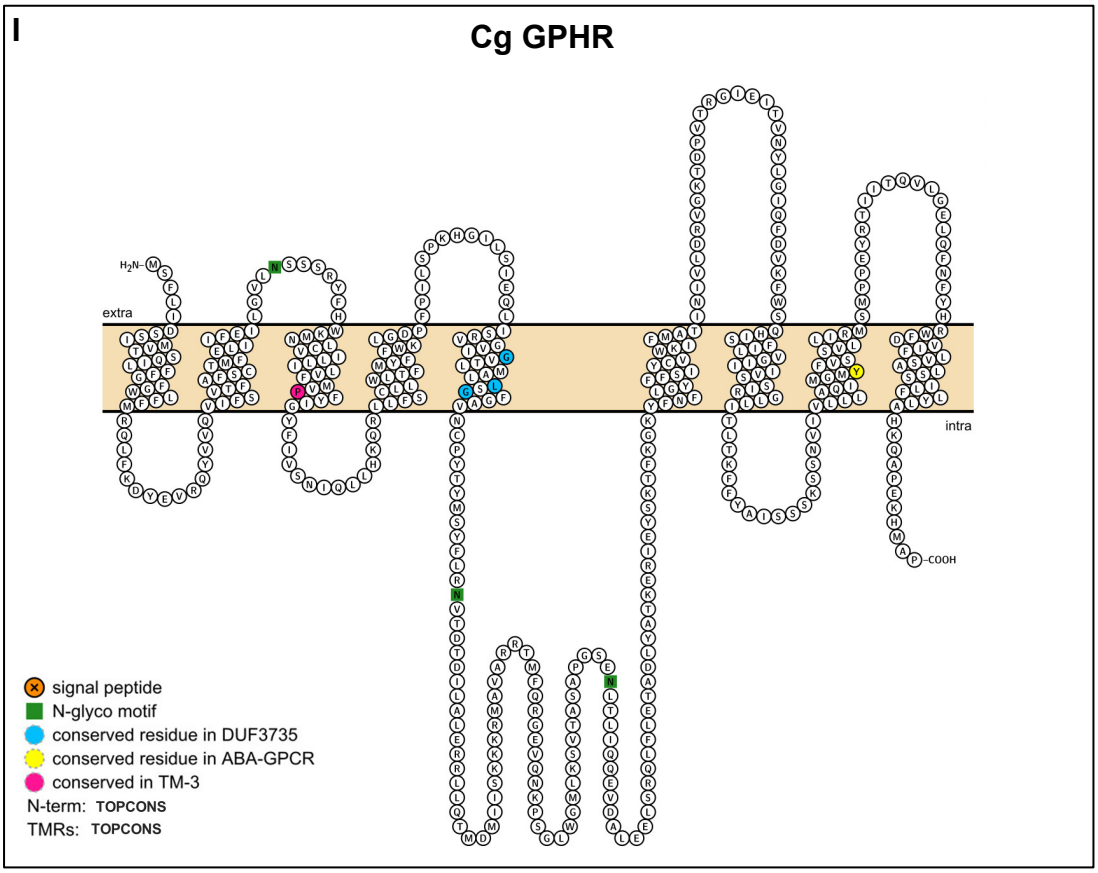
F


Cel-GPHR-2
H2N-
extra
intra
-COOH
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: TOPCONS
TMRs: TOPCONS

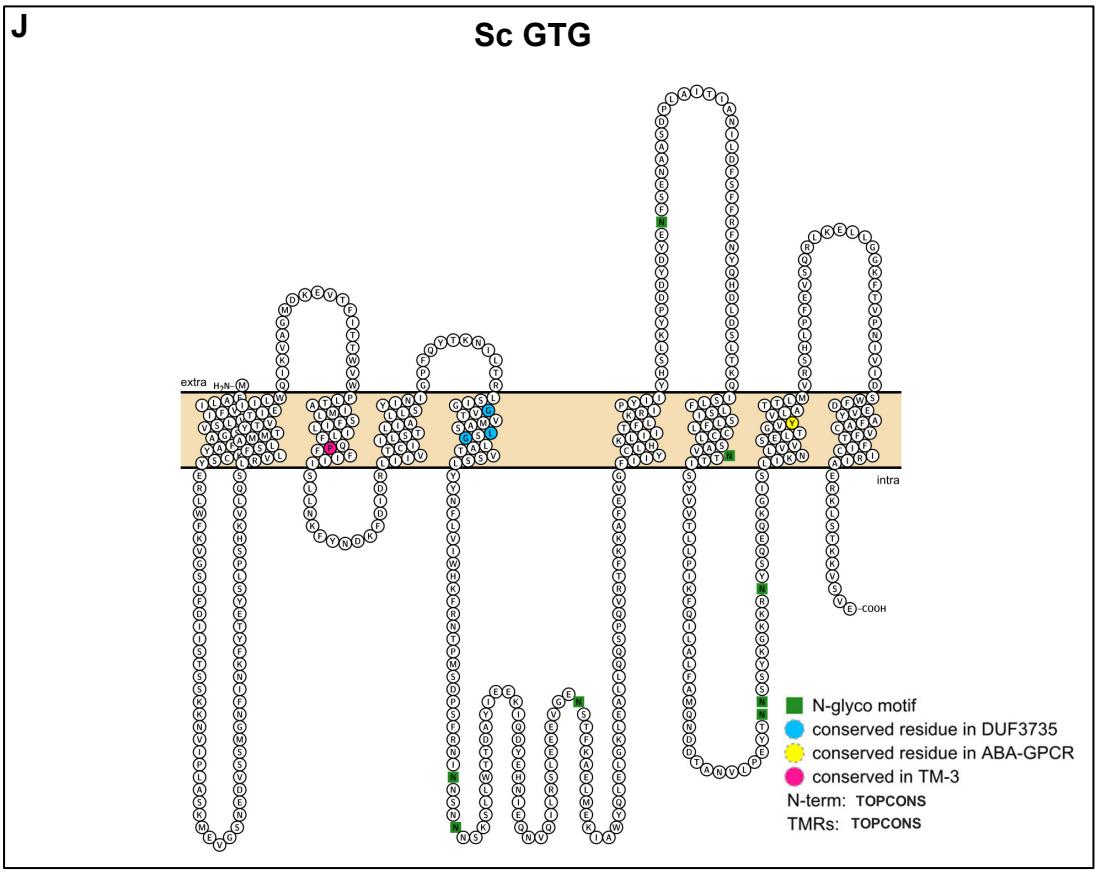

G
Hs GPHR
extra
intra
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: TOPCONS
TMRs: TOPCONS


H
Mm GPHR
extra
intra
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: TOPCONS
TMRs: TOPCONS

I


Cg GPHR
extra
intra
signal peptide
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: TOPCONS
TMRs: TOPCONS

J


Sc GTG
extra
intra
N-glyco motif
conserved residue in DUF3735
conserved residue in ABA-GPCR
conserved in TM-3
N-term: TOPCONS
TMRs: TOPCONS

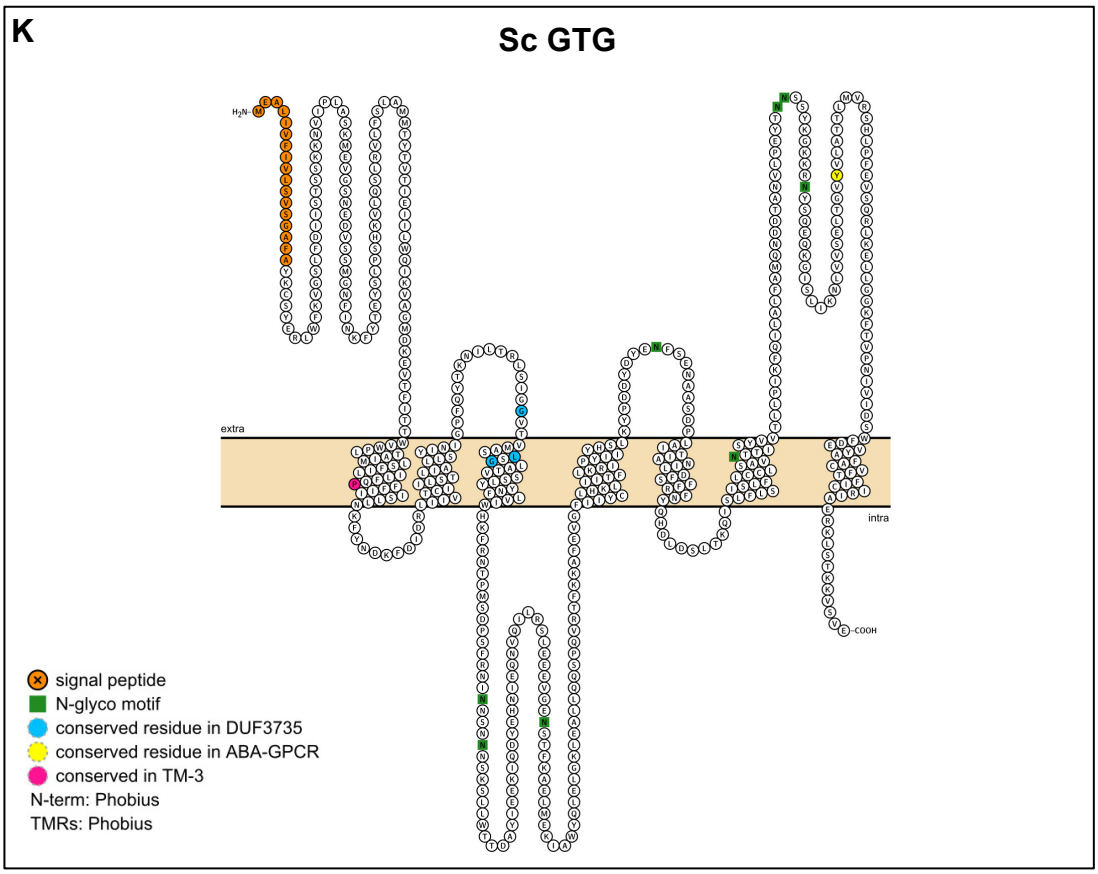


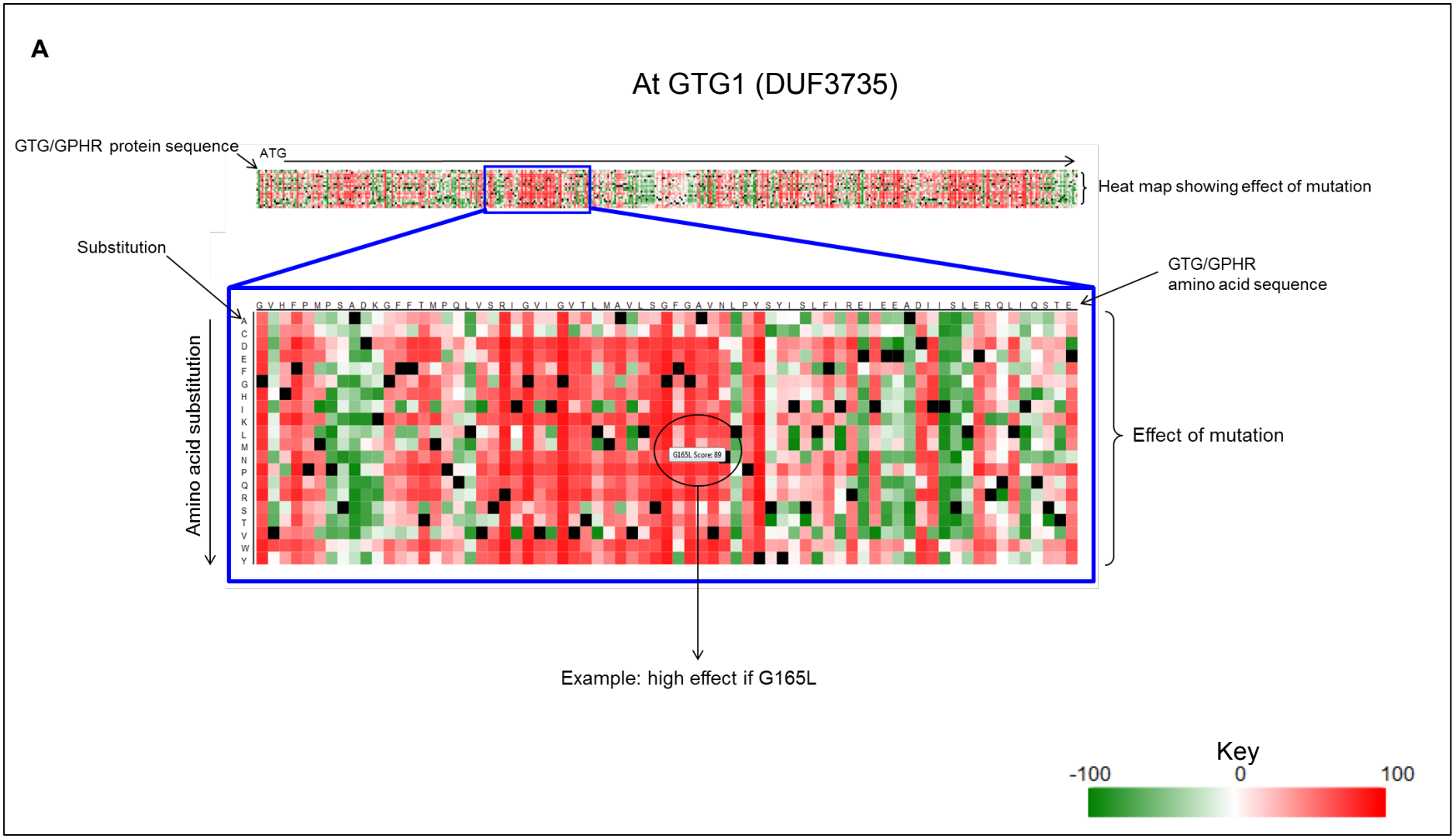
Figure 3.13 Predicted consensus membrane topology diagrams of the GTG/GPHRs

Membrane protein topology diagrams showing 2D structure predicted using TOPCONS (Bernsel et al., 2009) (**A – J**) and Phobius (Käll et al., 2004) (**K**) methods, drawn in Protter (Omasits et al., 2013) for: At GTG1 (**A**), At GTG2 (**B**), Os GTG1 (**C**), Dm GPHR (**D**), *Cel*-GPHR-1 (**E**), *Cel*-GPHR-2 (**F**), Hs GPHR (**G**), Mm GPHR (**H**), Cg GPHR (**I**) and Sc GPHR (**J – K**). Highlighted residues are predicted to be in the signal peptide (orange), N-glysolyated (N-glyco motif; green), conserved in DUF3735 (blue), conserved in ABA-GPCR domain (yellow) or conserved in the third TM region (TM-3; pink). N-term = N-terminus predicted by the specified method; TMRs = TM regions predicted by the specified method; extra = extracellular and intra = intracellular.

3.2.5 Predicted functional effects of mutations

The GTG/GPHRs are conserved proteins with high homology in specific regions and predicted to have a protein architecture containing DUF3735 and ABA-GPCR. As yet there has been little consideration of the functional effects of mutagenesis on conserved residues. PredictProtein can be used to explore the effect of mutagenesis on amino acids along the entire query sequence (Yachdav et al., 2014). Figure 3.14 A is an example output from this tool for At GTG1.

The results suggest that mutations in concentrated regions along the GTG/GPHR protein sequences are predicted to have a functional effect; since all these are in similar regions along the sequence of all the GTG/GPHRs, only At GTG1 is shown in detail (Figure 3.14). These regions, where a mutation into any amino acid may cause an effect, are estimated to correspond to the described DUF3735 and ABA-GPCR domains (Figure 3.14). The results suggest that even a substitution of leucine for glycine, or a more conserved substitution to alanine, in the LSG motif in DUF3735 will have a strong functional effect in At GTG1, for example (Figure 3.14). In fact, substitutions of any residues in the DUF3735 GxxxxxxLSG motif are predicted to have a functional effect. In addition, substitutions in residues that are estimated to be in the ABA-GPCR domain also cause a strong effect, including the conserved tyrosine residue found in the GTG/GPHRs (Figure 3.14). These results indicate substitutions within the predicted DUF3735 and ABA-GPCR of each could have an effect on function. Therefore these results are useful for mutagenesis studies when investigating structure-function relationships.


A
At GTG1 (DUF3735)
GTG/GPHR protein sequence
ATG
Heat map showing effect of mutation
Substitution
GTG/GPHR
amino acid sequence
Amino acid substitution
Effect of mutation
G165L Score: 89
Example: high effect if G165L
Key
-100
0
100

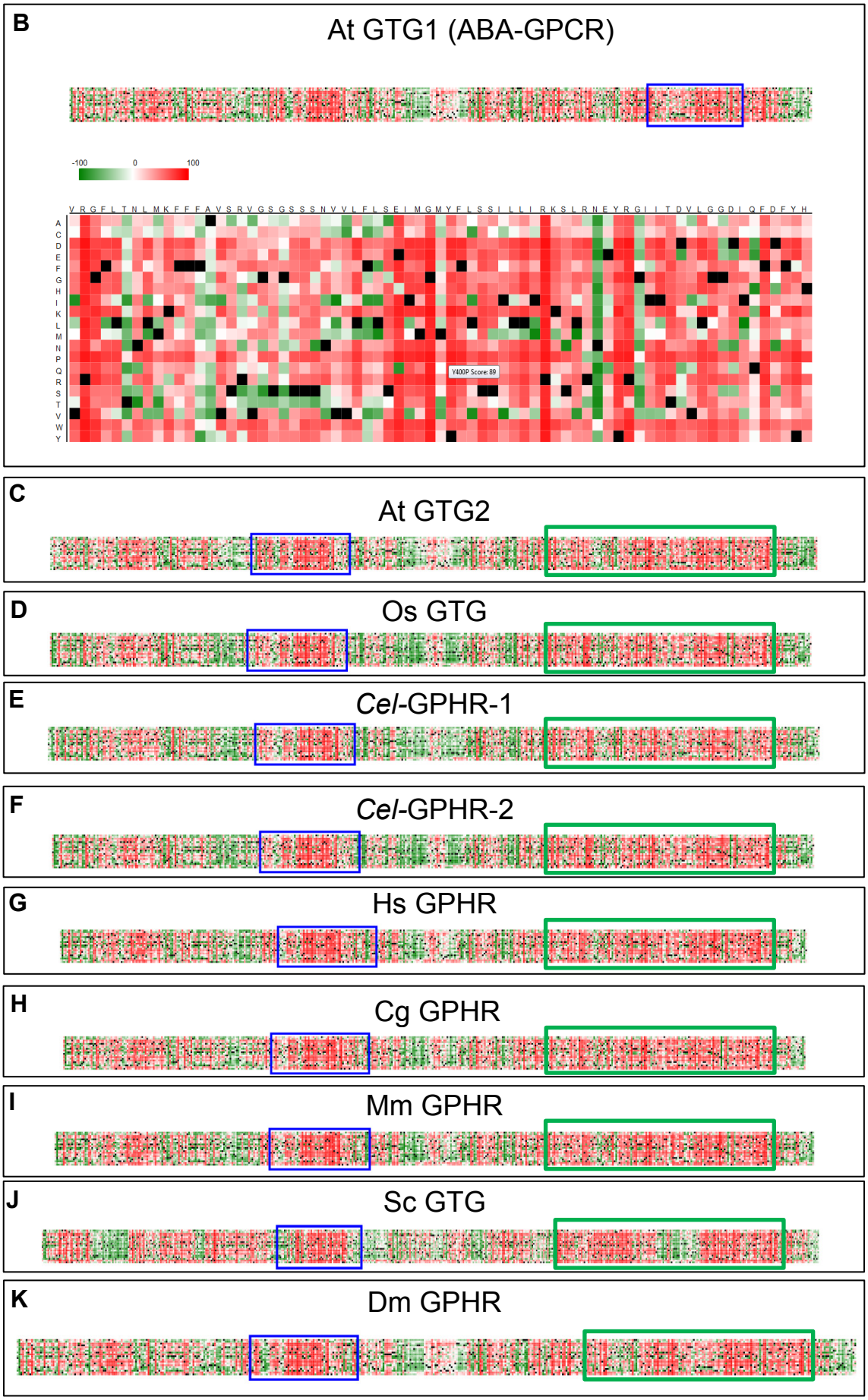

B
At GTG1 (ABA-GPCR)
-100
0
100
C
At GTG2
D
Os GTG
E
Cel-GPHR-1
F
Cel-GPHR-2
G
Hs GPHR
H
Cg GPHR
I
Mm GPHR
J
Sc GTG
K
Dm GPHR

Figure 3.14 Effect of point mutations in the GTG/GPHR protein sequences

Screenshots of the functional effects of different amino acid substitutions displayed in a heat map representation. PredictProtein was used to analyse the effect of independent substitutions for each position along the GTG/GPHR protein sequences, with blue insets on the panel indicating magnified area (labelled in brackets) showing regions of high functional effect (red) in the heat map presented below (only shown for **A** and **B**). At GTG1 (**A** and **B**), At GTG2 (**C**), Os GTG (**D**), *Cel*-GPHR-1 (**E**), *Cel*-GPHR-2 (**F**), Hs GPHR (**G**), Cg GPHR (**H**), Mm GPHR (**I**), Sc GTG (**J**) and Dm GPHR (**K**) (Yachdav et al., 2014). **A**. At GTG1; annotated example of the functional effects of point mutations in DUF3735 shown in a heat map representation. Dark red indicates a high score (score > 50, strong signal for effect), white indicates weak signals (-50 < score < 50), and green a low score (score < -50, strong signal for neutral/no effect). Black marks the corresponding wild type residues. Note: The program counts methionine (start) as amino acid residue "0". For example, G165L substitution in diagram A corresponds to G166L if Methionine is counted as residue 1. **C – K**. Heat map panels along the whole length of the sequence (not drawn to scale) to show that regions of high functional effect are concentrated to the predicted DUF3735 (blue rectangle) and the 3′ end, corresponding to the ABA-GPCR domain (green rectangle).

3.2.6 The GTG/GPHRs are predicted to be either plasma membrane localised or Golgi membrane-localised proteins

Protein localisation prediction methods were used to identify whether GTG/GPHRs from different organisms localise to similar locations of the cell. The ARAMEMNON database was used to predict the plant GTG/GPHRs in the first instance, which uses up to 16 individual methods as well as generating a built-in consensus method prediction, based on the Bayes theorem for localisation prediction (Schwacke et al., 2003). Overall, the plant GTG/GPHRs are predicted to be localised to the secretory pathway, as shown by the consensus prediction (Figure 3.15). However, this database only provided localisation predictions for proteins that are intended for the chloroplast, mitochondria or secretory pathway; there were no localisation predictions specifically for the plasma membrane, for example. The At GTGs were also searched for in several proteome databases including the Arabidopsis Nucleolar Protein database (http://bioinf.scri.sari.ac.uk/cgi-bin/atnopdb/home), the Plastid Proteome Database (http://ppdb.tc.cornell.edu/), the Arabidopsis Mitochondrial Protein Database (http://www.plantenergy.uwa.edu.au/applications/ampdb/index.html), the Arabidopsis Subcellular Proteomic Database and the Plant Organellome Database (http://podb.nibb.ac.jp/Organellome/); however none listed At GTG localisations. Since ARAMEMNON only provides information for plant proteins, other prediction methods had to be used to explore the predicted localisation of the other GTG/GPHRs analysed in this chapter. The results indicate that according to PredictProtein which uses the LocTree2 method for localisation prediction (Goldberg et al., 2012), the plant GTGs are predicted to be plasma membrane localised, while the animal and *S. cerevisiae* GTG/GPHRs are predicted to be Golgi membrane localised (Table 3.6). However, the WoLF PSORT localisation prediction method (Horton et al., 2007), shows that all GTG/GPHRs are plasma membrane localised (Table 3.6). Since the predicted localisation results are so varied suggesting either plasma membrane or Golgi/secretory localisation for the GTG/GPHRs, experimental data is necessary to confirm their native localisations.

A

At GTG1

B

At GTG2

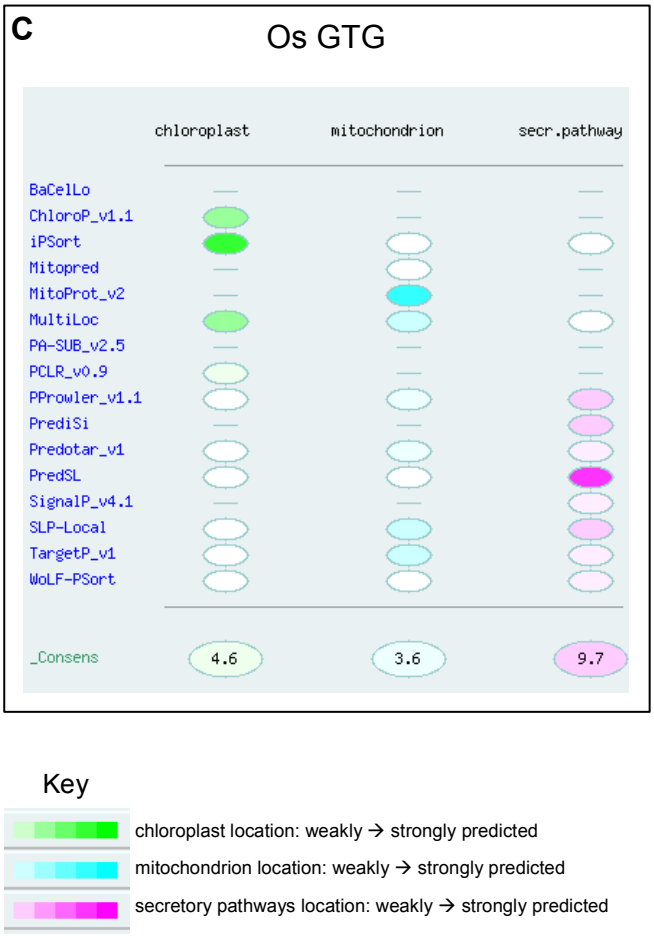


Figure 3.15 The plant GTG/GPHRs are predicted to localise to the secretory pathway

Screenshots displaying localisation predictions showing At GTG1 (**A**), At GTG2 (**B**) and Os GTG (**C**) are localised to the secretory pathway using information collated by the ARAMEMNON database (Schwacke et al., 2003).

Table 3.6 Predicted localisation for GTG/GPHRs

Protein	Sequence information			
	PredictProtein[1]	**TargetP 1.1[2]**	**Consensus[3]**	**WoLF PSORT[4]**
At GTG1	Plasma membrane	Secretory pathway	Secretory pathways	Plasma membrane
At GTG2	Plasma membrane	Secretory pathway	Secretory pathways	Plasma membrane
Os GTG	Plasma membrane	Any other location	Secretory pathways	Plasma membrane
Cel-GPHR-1	Golgi membrane	Secretory pathway	-	Plasma membrane
Cel-GPHR-2	Golgi membrane	Secretory pathway	-	Plasma membrane
Sc GTG	Golgi membrane	Secretory pathway	-	Plasma membrane
Hs GPHR	Golgi membrane	Secretory pathway	-	Plasma membrane
Mm GPHR	Golgi membrane	Secretory pathway	-	Plasma membrane
Cg GPHR	Golgi membrane	Secretory pathway	-	Plasma membrane
Dm GPHR	Golgi membrane	Secretory pathway	-	Plasma membrane

[1] PredictProtein: http://ppopen.informatik.tu-muenchen.de/ (Yachdav et al., 2014)

[2] TargetP 1.1: http://www.cbs.dtu.dk/services/TargetP/ (Emanuelsson et al., 2007)

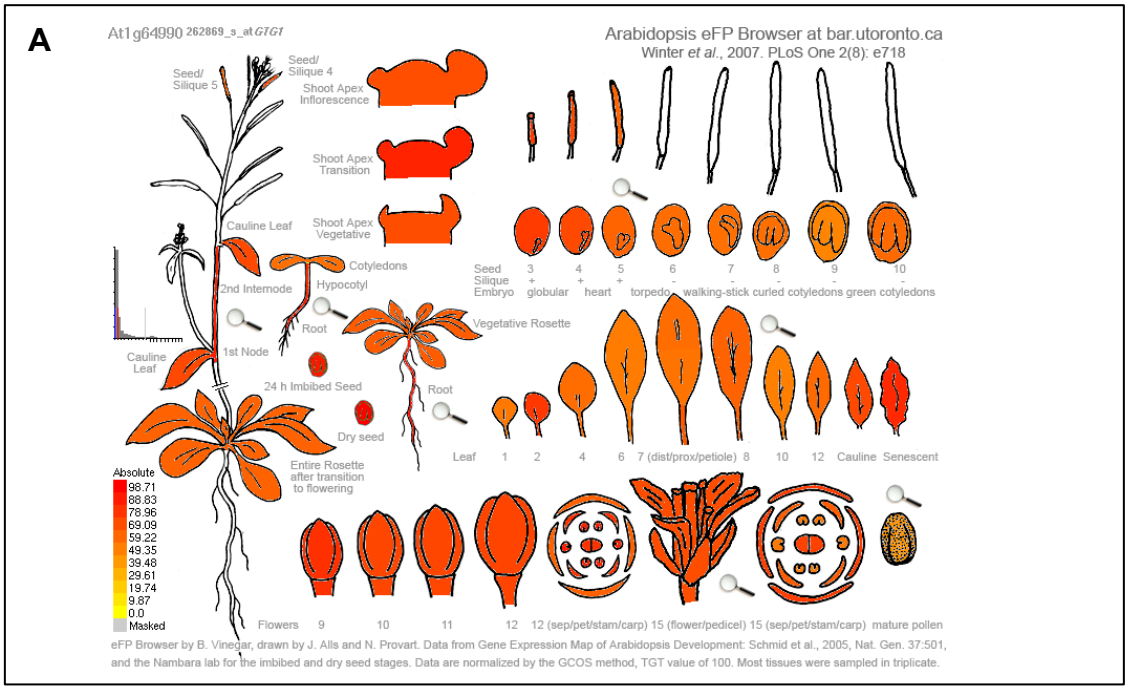
[3] Consensus: http://aramemnon.botanik.uni-koeln.de/ (Schwacke et al., 2003)

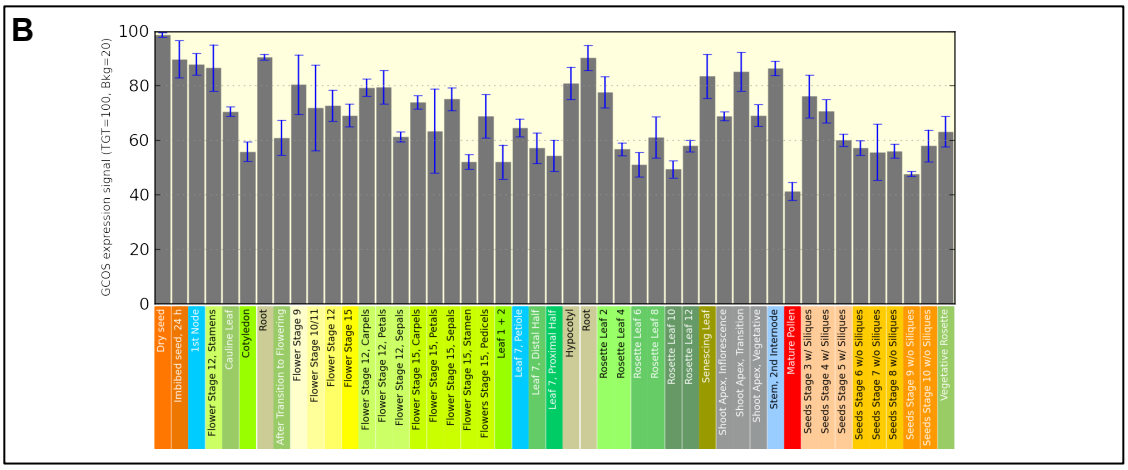
[4] WoLF PSORT: http://psort.hgc.jp/ (Horton et al., 2007)

3.2.7 The *GTG/GPHR* genes are widely expressed

Gene expression data from the BAR *A. thaliana* eFP Browser indicate that the At *GTGs* are highly expressed throughout all stages of plant development (Figure 3.16 A and B). The probe set used for this data indicate detect the expression of both At *GTG1* and At *GTG2* genes. This was shown by large sample sizes (93 – 2208) at different developmental stages, except for seeds where only 18 samples were used. However, the Genevestigator data indicates that At *GTG* expression is only at a medium level throughout developmental stages, including germination, seedling, young rosettes, developed rosette, bolting, young flowering, developed flower, flower and silique, siliques and senescence stages (Figure 3.16 C). The BAR *A. thaliana* eFP Browser shows that At *GTGs* are most expressed in dry seed and least in mature pollen (almost 100 % and approximately 40 % signal, respectively) (Figure 3.16 A – B). Other areas of very high (around 80 – 90 %) At GTG expression include imbibed seeds, first node, flower stage 12 in the stamens, carpels and petals, rosettes after flowering, roots, hypocotyls, and rosette leaf stage 2, senescing leaves and in the transition stage of the apex (Figure 3.16 A – B). While At *GTG* expression is around 50 % in the cotyledons, rosette leaves in stages 1 and 2 and between stages 4 – 12, as well as in seeds without siliques (Figure 3.16 A – B). The At *GTG* genes are expressed in developing siliques but not in mature siliques (Figure 3.16 A); this is consistent with the Genevestigator data which also indicates very low levels of At *GTG* expression in siliques (Figure 3.18). Anatomical gene expression analysis using Genevestigator indicates that the At *GTGs* are highly expressed in the seed coat and generally very highly expressed in the shoot apex, giant cell, chalazal endosperm, root xylem protoplast, micropylar endosperm, columella protoplast, suspensor and embryo. Medium expression of the At *GTGs* occurs in roots, cotyledons, hypocotyl, roots, elongation zone, root tip, seeds and siliques, amongst other tissues. The lowest expression (although still classed as medium level) occurs in the phloem, pollen, carpel and mesophyll cell protoplast (Figure 3.18).

In *O. sativa*, the Os *GTG* gene is highly expressed (Figure 3.17). The highest-level expression was in the seedling roots, young and mature leaves, and lowest in the seed stage 1. In the seeds and inflorescence development Os *GTG* is expressed at lower levels at first, but during development, this increases before declining in expression again at later stages (Figure 3.17).


A
At1g64990 262869_s_at GTG1
Arabidopsis eFP Browser at bar.utoronto.ca
Winter et al., 2007. PLoS One 2(8): e718
Seed/ Silique 5
Seed/ Silique 4
Shoot Apex Inflorescence
Shoot Apex Transition
Shoot Apex Vegetative
Cauline Leaf
Cotyledons
Hypocotyl
2nd Internode
Root
1st Node
Cauline Leaf
24 h Imbibed Seed
Dry seed
Entire Rosette after transition to flowering
Seed Silique Embryo
3 + globular
4 + heart
5 + torpedo
6 - walking-stick
7 - curled cotyledons
8 - green cotyledons
9 -
10 -
Vegetative Rosette
Root
Leaf 1 2 4 6 7 (dist/prox/petiole) 8 10 12 Cauline Senescent
Absolute
98.71
88.83
78.96
69.09
59.22
49.35
39.48
29.61
19.74
9.87
0.0
Masked
Flowers 9 10 11 12 12 (sep/pet/stam/carp) 15 (flower/pedicel) 15 (sep/pet/stam/carp) mature pollen
eFP Browser by B. Vinegar, drawn by J. Alls and N. Provart. Data from Gene Expression Map of Arabidopsis Development: Schmid et al., 2005, Nat. Gen. 37:501, and the Nambara lab for the imbibed and dry seed stages. Data are normalized by the GCOS method, TGT value of 100. Most tissues were sampled in triplicate.


B
GCOS expression signal (TGT=100, Bkg=20)
100
80
60
40
20
0
Dry seed
Imbibed seed, 24 h
1st Node
Flower Stage 12, Stamens
Cauline Leaf
Cotyledon
Root
After Transition to Flowering
Flower Stage 9
Flower Stage 10/11
Flower Stage 12
Flower Stage 15
Flower Stage 12, Carpels
Flower Stage 12, Petals
Flower Stage 12, Sepals
Flower Stage 15, Carpels
Flower Stage 15, Petals
Flower Stage 15, Sepals
Flower Stage 15, Stamen
Flowers Stage 15, Pedicels
Leaf 1 + 2
Leaf 7, Petiole
Leaf 7, Distal Half
Leaf 7, Proximal Half
Hypocotyl
Root
Rosette Leaf 2
Rosette Leaf 4
Rosette Leaf 6
Rosette Leaf 8
Rosette Leaf 10
Rosette Leaf 12
Senescing Leaf
Shoot Apex, Inflorescence
Shoot Apex, Transition
Shoot Apex, Vegetative
Stem, 2nd Internode
Mature Pollen
Seeds Stage 3 w/ Siliques
Seeds Stage 4 w/ Siliques
Seeds Stage 5 w/ Siliques
Seeds Stage 6 w/o Siliques
Seeds Stage 7 w/o Siliques
Seeds Stage 8 w/o Siliques
Seeds Stage 9 w/o Siliques
Seeds Stage 10 w/o Siliques
Vegetative Rosette

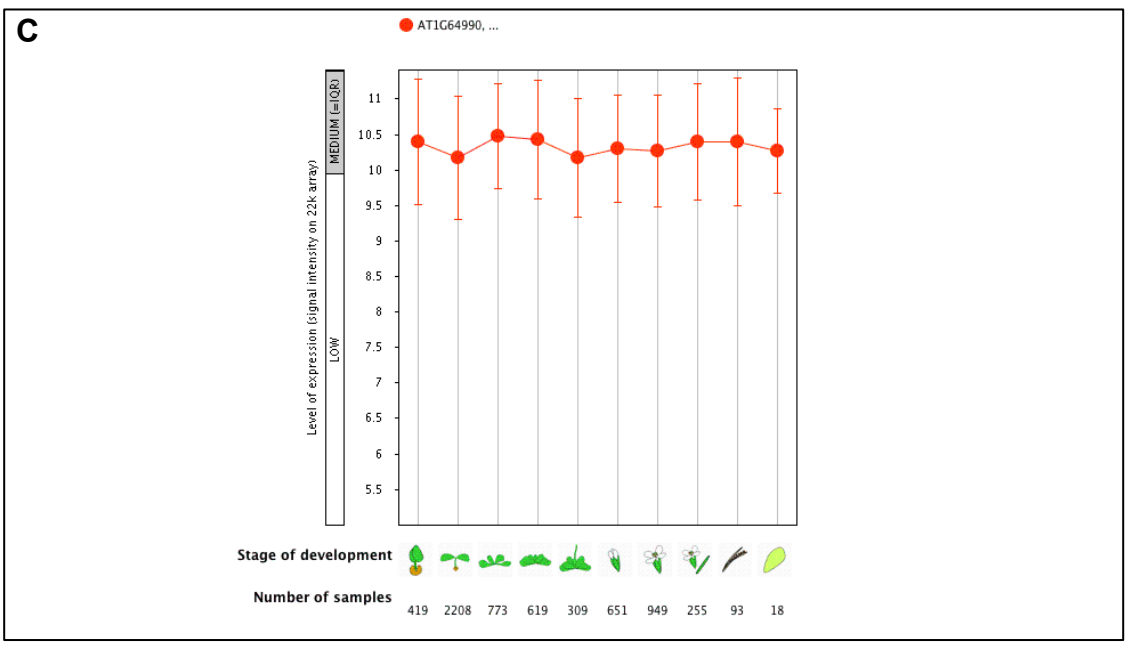

C
AT1G64990, ...
MEDIUM (=IQR)
LOW
Level of expression (signal intensity on 22k array)
11
10.5
10
9.5
9
8.5
8
7.5
7
6.5
6
5.5
Stage of development
Number of samples
419 2208 773 619 309 651 949 255 93 18

Figure 3.16 At *GTG* genes are expressed at all Arabidopsis developmental stages

A – B. Data from BAR *A. thaliana* eFP Browser showing the level of expression of the At *GTG* genes. **A**. Level of gene expression shown in sampled tissues shown using different intensities of red colour (scale: red = high expression → yellow = low/no expression). **B**. Graphical depiction showing the expression signal in different tissue samples (±S.E.). Data obtained using BAR *A. thaliana* eFP Browser (http://bar.utoronto.ca/welcome.htm) (Winter et al., 2007), using results from Schmid et al. (2005). 262869_s_at was used as the probe set identifier for both At *GTG1* (At1g64990) and At *GTG2* (At4g27630); secondary gene used was At3g27340. Data shown is representative of At *GTG2* gene expression levels since the same probe set identifier and secondary gene was used. **C**. Scatterplot showing the expression of the At *GTGs* at various developmental stages. The stage of development is shown with the number of samples used to quantify expression levels for each stage. ATH1: 262869_s_at probe set used in the *A. thaliana* 22k array platform. Data obtained using Genevestigator (Hruz et al., 2008).

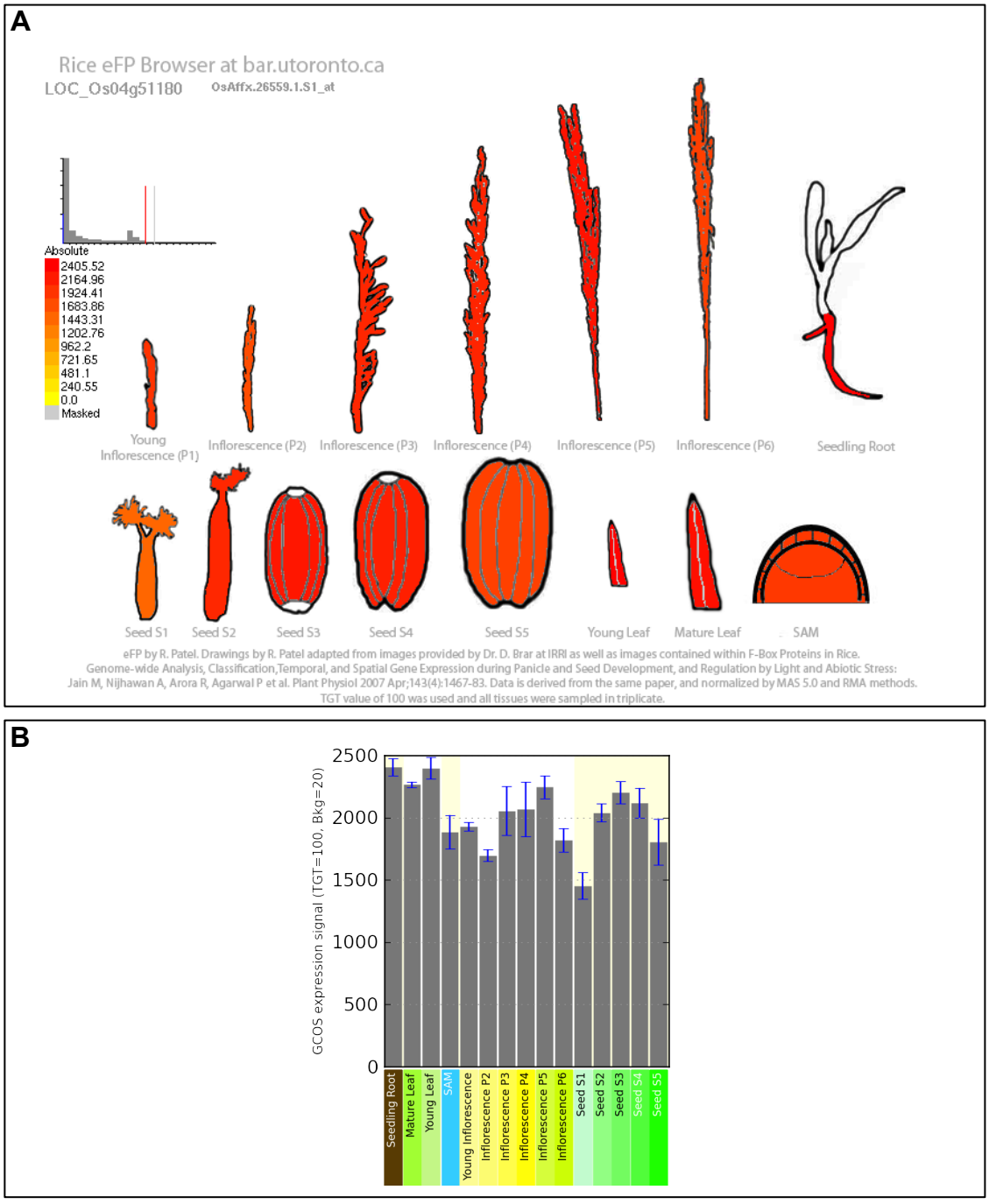


Figure 3.17 Os *GTG* expression occurs at different *O. sativa* developmental stages

Data from BAR Rice eFP Browser showing the level of expression of the Os *GTG*. **A**. Level of gene expression shown in sampled tissues shown using different intensities of red colour (scale: red = high expression → yellow = low/no expression). **B**. Graphical depiction showing the expression signal in different tissue samples (±S.E.). Data obtained using BAR Rice eFP Browser (Winter et al., 2007) (http://bar.utoronto.ca/welcome.htm), using results from (Jain et al., 2007). OsAffx.26559.1.S1_at was used as the probe set identifier Os *GTG* (LOC_Os04g51180); secondary gene used was LOC_Os01g51540.1.

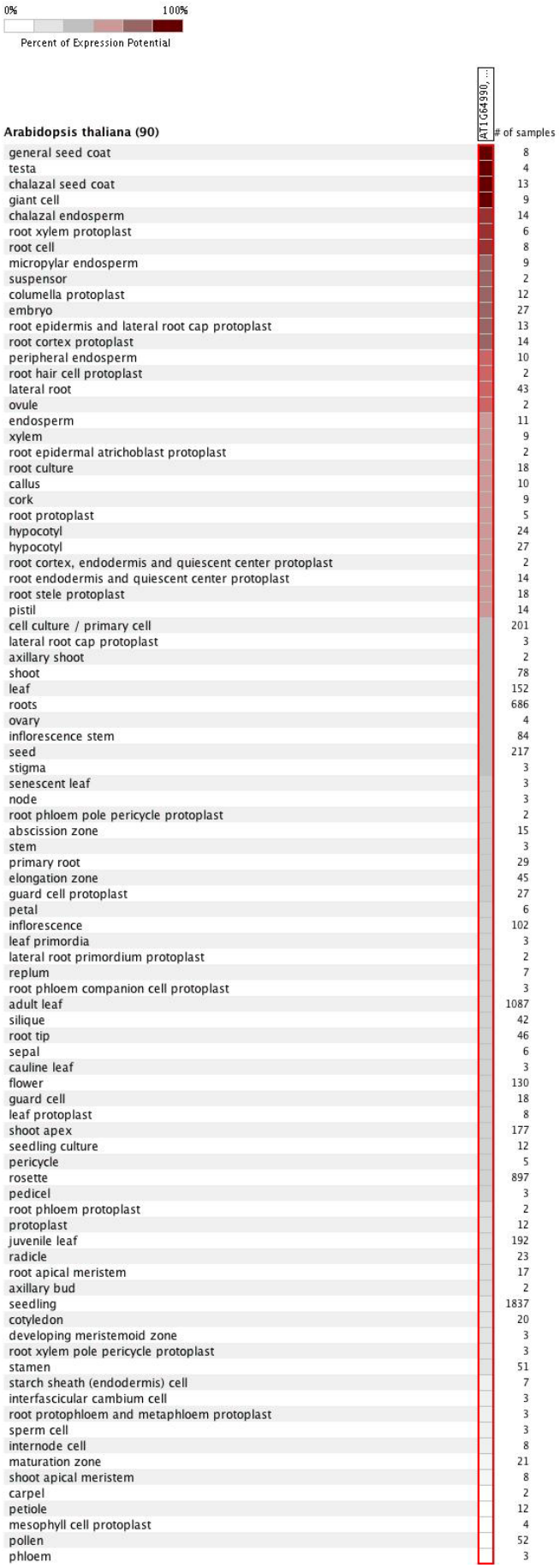


Figure 3.18 Anatomical expression of At *GTG* genes

Heat map showing the expression of At *GTG* genes in various organs, tissues and cells. The number of sample and percentage expression potential scale is shown on figure. ATH1: 262869_s_at probe set used to identify At *GTG1* and At *GTG2* expression in the *A. thaliana* 22k array platform. Data obtained using Genevestigator (Hruz et al., 2008).

The expression of At *GTG* genes alters under different experimental conditions (pertubations). Figure 3.19 displays the pertubations under which the At *GTG* expression is upregulated or downregulated. At *GTG* expression was upregulated when treated with brassinolide (BL) and boric acid in combination by almost two-fold in 4, 6 or 8 d treatments compared to untreated samples. There was an increased At *GTG* expression when plants were desiccated in senescence stage, compared to when germinated. Interestingly, At *GTG* expression is upregulated in regulator of G protein signalling (*rgs1-2*) mutants compared to WT; this indicates that At GTGs may be involved in G protein signalling. At *GTG* genes are both upregulated and downregulated in different iron-deficiency experiments. In roots, iron-deficiency promotes At *GTG* upregulation, however in root protoplasts these genes were downregulated. At *GTG* expression was also downregulated in a hypocotyl growth experiment comparing *det3* mutants to WT. They are also downregulated in long-day conditions compared to short-day (Figure 3.19).

Genevestigator results show that Sc *GTG* expression occurs at all developmental stages, with highest expression occurring in the stationary phase and the lowest expression in the mid-log phase (Figure 3.20 A and B). Anatomical expression analysis indicates that Sc *GTG* expression occurs most highly in haplophase, diplophase and pre-sporulation *S. cerevisiae* while there is less expression of Sc *GTG* at the sporulation stage (Figure 3.20 C).

In animals, Mm *GPHR* is expressed in all stages of development and is slightly higher in the embryonic blastocyst stage (Figure 3.21). Information is not available regarding the stage of developmental feature for *H. sapiens*, but Hs *GPHR* is expressed in 288 organs, tissues or cells showing the highest level of expression in the kidneys and lowest in immune cells (Figure 3.23). The Hs and Mm *GPHRs* are highly expressed in 960 and 29 neoplasms, respectively (Figure 3.22 and Figure 3.24). In *H. sapiens*, *GPHR* expression is found to be high in a number of metastatic and/or malignant myelomas or carcinomas including Hodgekin's disease and plasma cell tumours and multiple myelomas. However, there is low expression of Hs GPHR in other neoplasms such as renal cell carcinoma, malignant insulinoma and primary cutaneous $CD30^+$ T-cell lymphoproliferative disorders (Figure 3.24). Mm *GPHR* is most highly expressed in body fluid neoplasms but is also highly expressed in digestive, alimentary and nervous system neoplasms as well as in the eye (Figure 3.22).

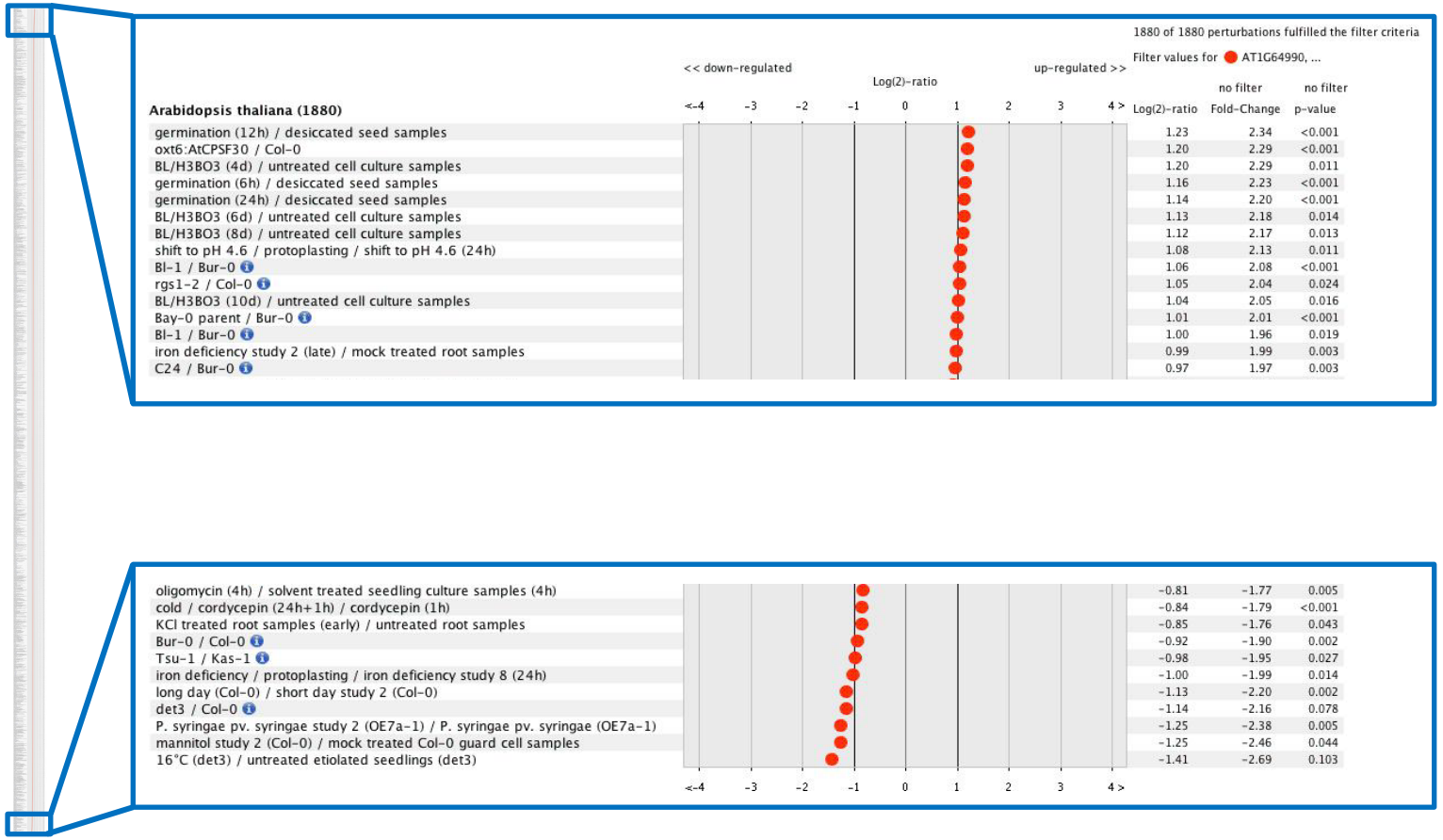
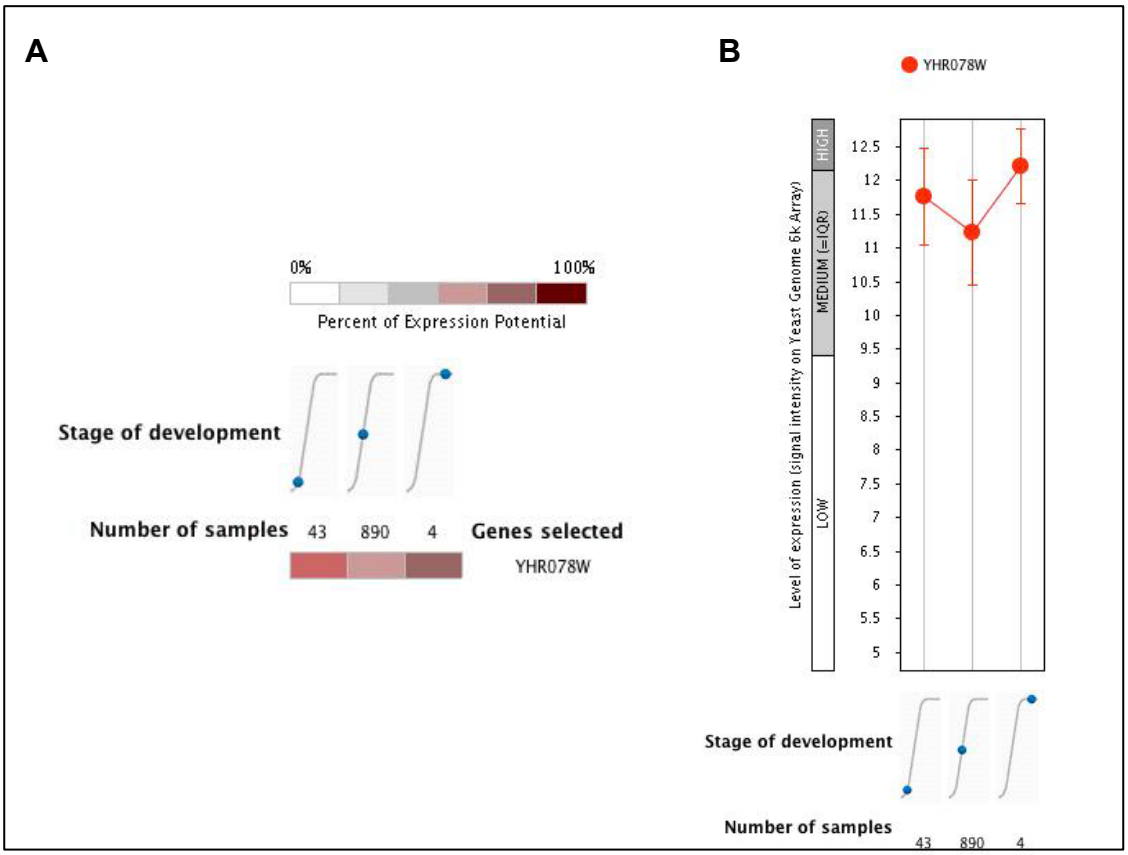


Figure 3.19 At *GTG* gene expression is altered under different pertubations

Scatterplot showing the expression of At *GTG* genes in various pertubations. The number of fold-change, log(2)-ratio and *p*-values are shown on figure. ATH1: 262869_s_at probe set used to identify At *GTG1* and At *GTG2* in the *A. thaliana* 22k array platform. The highest level of up and downregulation are magnified and indicated by blue insets showing these regions. Data obtained using Genevestigator (Hruz et al., 2008).



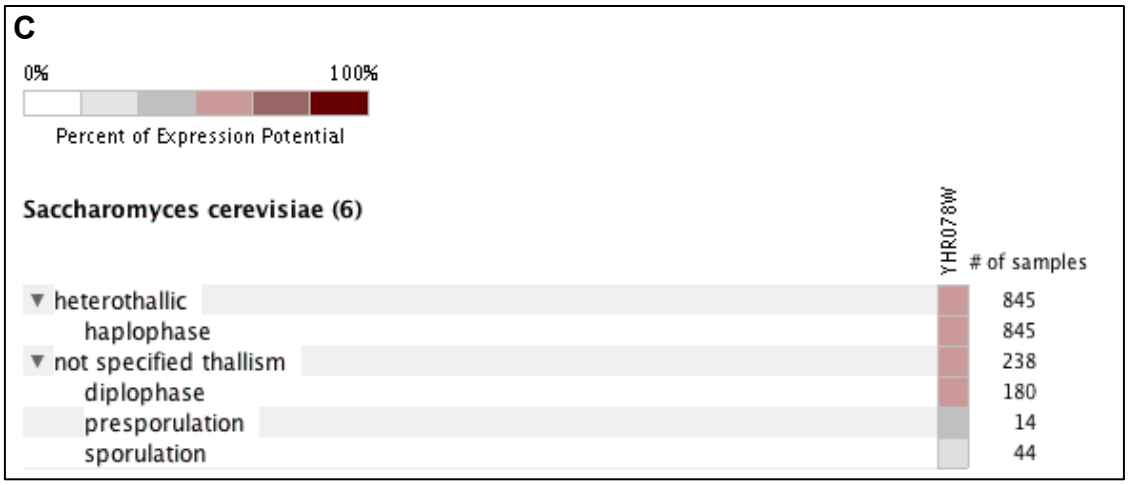


Figure 3.20 Sc *GTG* is expressed at all stages of development

Heat maps (**A** and **C**) and scatterplot (**B**) showing Sc *GTG* (YHR078W) gene expression in all stages phases of development (**A** and **B**) as well as in both heterothallic and non-specific thallic phases (**C**). The number of sample and percentage expression potential scale is shown on figure. 4458_at probe set used in the YG-S98: Yeast Genome 6k Array platform. Data obtained using Genevestigator (Hruz et al., 2008).

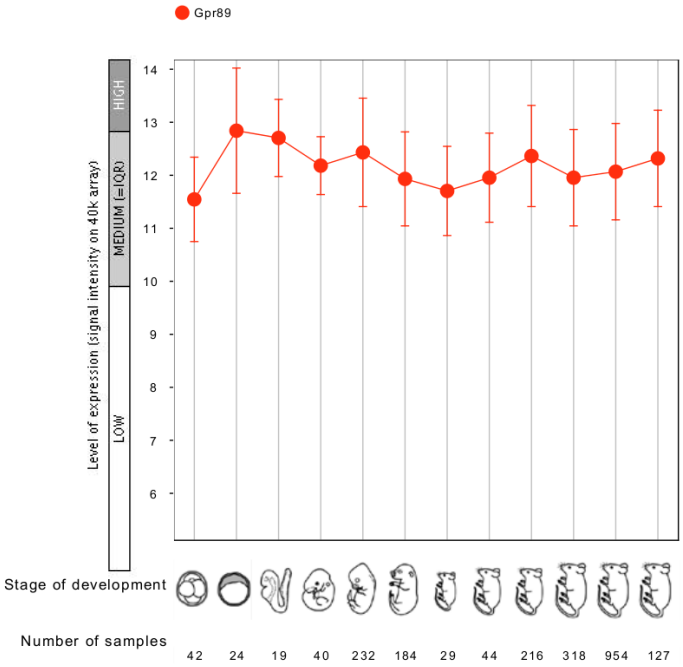


Figure 3.21 Mm *GPHR* expression is similar at all developmental stages

Scatterplot showing Mm *GPHR* gene expression levels at various developmental stages of the mouse. The number of samples used for each stage is shown on figure. Data obtained using Genevestigator (Hruz et al., 2008). 1424696_at probe set used in the Mouse430_2: 40k array platform.

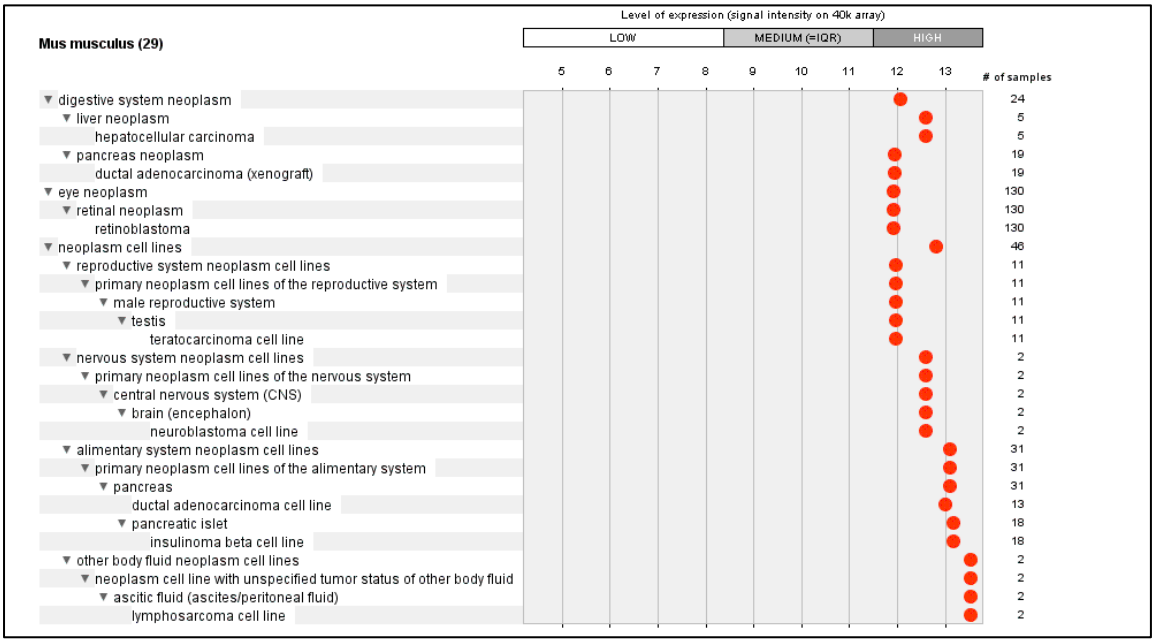


Figure 3.22 Expression of Mm *GPHR* in different neoplasms

Scatterplot showing Mm *GPHR* expression in 29 categories of neoplasm. Expression level and number of samples shown on figure. The 1424696_at probe set used in the Mouse430_2: 40k array platform. Data obtained using Genevestigator (Hruz et al., 2008).

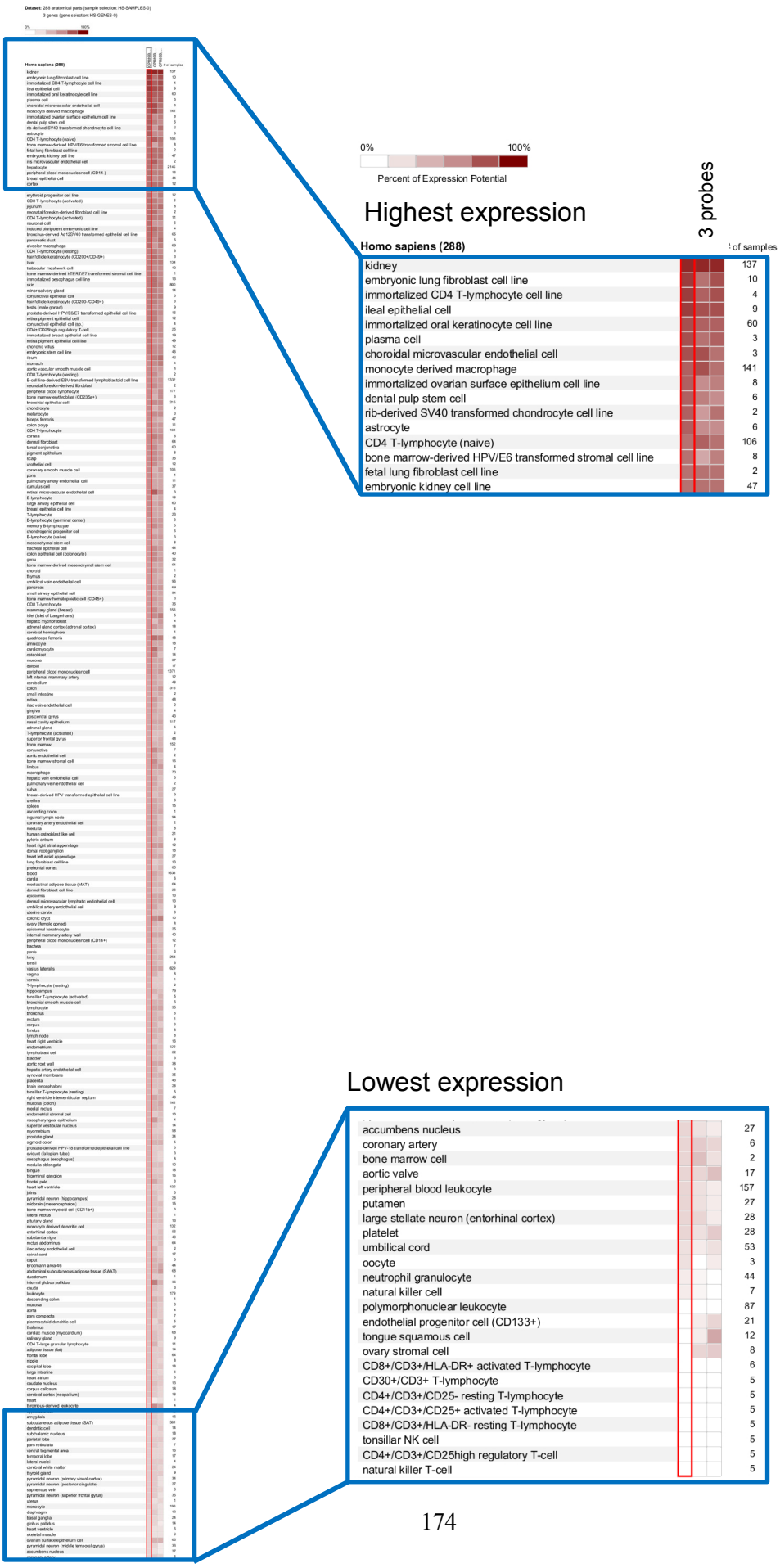

0%
100%
Percent of Expression Potential
Highest expression
3 probes
Homo sapiens (288)
# of samples
kidney 137
embryonic lung fibroblast cell line 10
immortalized CD4 T-lymphocyte cell line 4
ileal epithelial cell 9
immortalized oral keratinocyte cell line 60
plasma cell 3
choroidal microvascular endothelial cell 3
monocyte derived macrophage 141
immortalized ovarian surface epithelium cell line 8
dental pulp stem cell 6
rib-derived SV40 transformed chondrocyte cell line 2
astrocyte 6
CD4 T-lymphocyte (naive) 106
bone marrow-derived HPV/E6 transformed stromal cell line 8
fetal lung fibroblast cell line 2
embryonic kidney cell line 47
Lowest expression
accumbens nucleus 27
coronary artery 6
bone marrow cell 2
aortic valve 17
peripheral blood leukocyte 157
putamen 27
large stellate neuron (entorhinal cortex) 28
platelet 28
umbilical cord 53
oocyte 3
neutrophil granulocyte 44
natural killer cell 7
polymorphonuclear leukocyte 87
endothelial progenitor cell (CD133+) 21
tongue squamous cell 12
ovary stromal cell 8
CD8+/CD3+/HLA-DR+ activated T-lymphocyte 6
CD30+/CD3+ T-lymphocyte 5
CD4+/CD3+/CD25- resting T-lymphocyte 5
CD4+/CD3+/CD25+ activated T-lymphocyte 5
CD8+/CD3+/HLA-DR- resting T-lymphocyte 5
tonsillar NK cell 5
CD4+/CD3+/CD25high regulatory T-cell 5
natural killer T-cell 5

Figure 3.23 Anatomical expression of Hs *GPHR*

Heat map showing Hs *GPHR* expression in different organs, tissues and cells using three different probe sets (220642_x_at, 223531_x_at, 222140_s_at) using the Human Genome 47k array platform. The number of sample and percentage expression potential scale is shown on figure. The highest and lowest expression found are magnified and indicated by blue insets showing these regions. Data obtained using Genevestigator (Hruz et al., 2008).

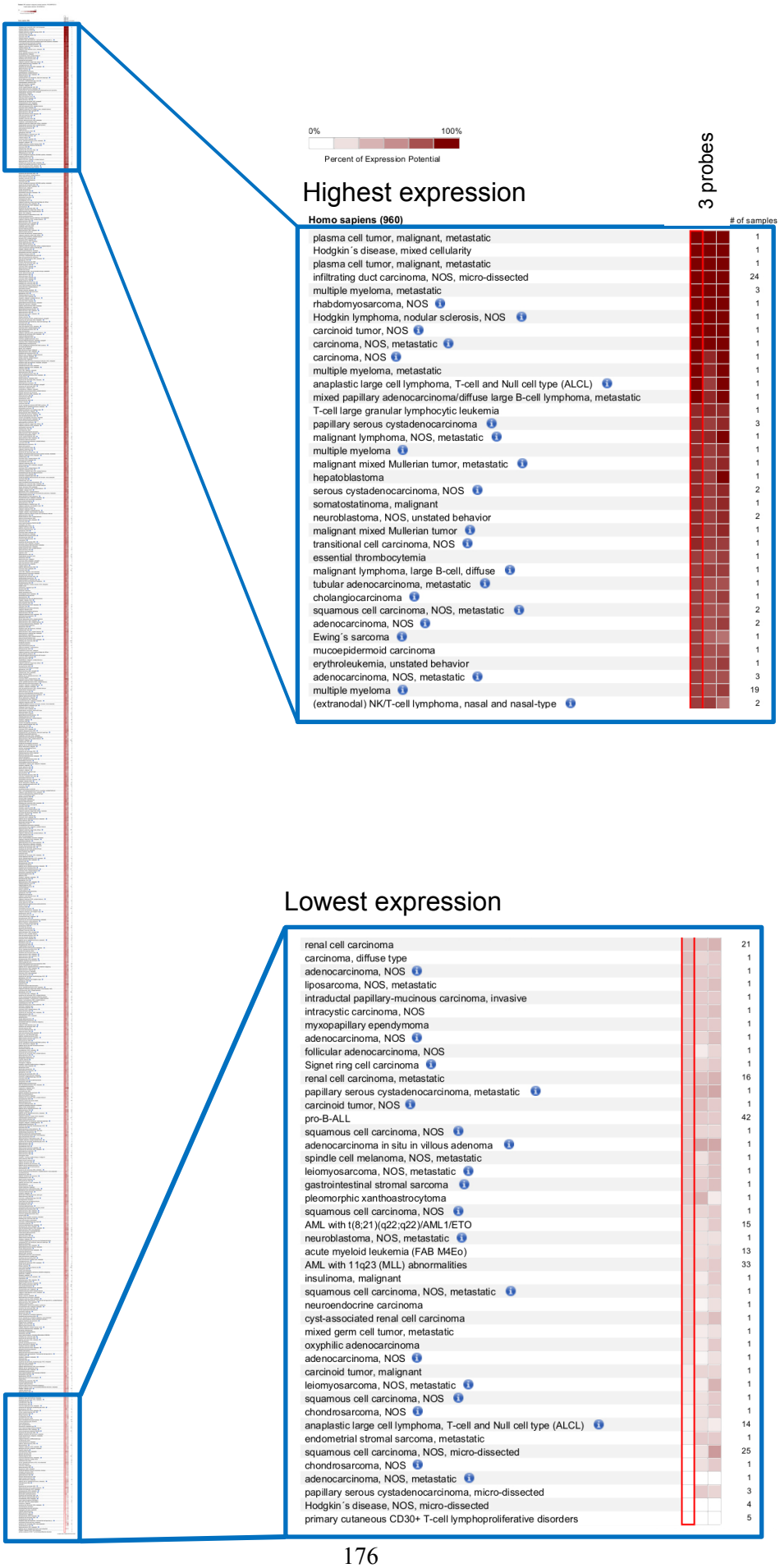

0%
100%
Percent of Expression Potential
Highest expression
3 probes
Homo sapiens (960)
# of samples
plasma cell tumor, malignant, metastatic 1
Hodgkin´s disease, mixed cellularity 1
plasma cell tumor, malignant, metastatic 1
infiltrating duct carcinoma, NOS, micro-dissected 24
multiple myeloma, metastatic 3
rhabdomyosarcoma, NOS 1
Hodgkin lymphoma, nodular sclerosis, NOS 1
carcinoid tumor, NOS 1
carcinoma, NOS, metastatic 1
carcinoma, NOS 1
multiple myeloma, metastatic 1
anaplastic large cell lymphoma, T-cell and Null cell type (ALCL) 1
mixed papillary adenocarcinoma/diffuse large B-cell lymphoma, metastatic 1
T-cell large granular lymphocytic leukemia 1
papillary serous cystadenocarcinoma 3
malignant lymphoma, NOS, metastatic 1
multiple myeloma 1
malignant mixed Mullerian tumor, metastatic 1
hepatoblastoma 1
serous cystadenocarcinoma, NOS 2
somatostatinoma, malignant 1
neuroblastoma, NOS, unstated behavior 2
malignant mixed Mullerian tumor 1
transitional cell carcinoma, NOS 1
essential thrombocytemia 1
malignant lymphoma, large B-cell, diffuse 1
tubular adenocarcinoma, metastatic 1
cholangiocarcinoma 1
squamous cell carcinoma, NOS, metastatic 2
adenocarcinoma, NOS 2
Ewing´s sarcoma 1
mucoepidermoid carcinoma 1
erythroleukemia, unstated behavior 1
adenocarcinoma, NOS, metastatic 3
multiple myeloma 19
(extranodal) NK/T-cell lymphoma, nasal and nasal-type 2
Lowest expression
renal cell carcinoma 21
carcinoma, diffuse type 1
adenocarcinoma, NOS 1
liposarcoma, NOS, metastatic 1
intraductal papillary-mucinous carcinoma, invasive 1
intracystic carcinoma, NOS 1
myxopapillary ependymoma 1
adenocarcinoma, NOS 1
follicular adenocarcinoma, NOS 1
Signet ring cell carcinoma 1
renal cell carcinoma, metastatic 16
papillary serous cystadenocarcinoma, metastatic 1
carcinoid tumor, NOS 1
pro-B-ALL 42
squamous cell carcinoma, NOS 1
adenocarcinoma in situ in villous adenoma 1
spindle cell melanoma, NOS, metastatic 1
leiomyosarcoma, NOS, metastatic 1
gastrointestinal stromal sarcoma 1
pleomorphic xanthoastrocytoma 1
squamous cell carcinoma, NOS 1
AML with t(8;21)(q22;q22)/AML1/ETO 15
neuroblastoma, NOS, metastatic 1
acute myeloid leukemia (FAB M4Eo) 13
AML with 11q23 (MLL) abnormalities 33
insulinoma, malignant 1
squamous cell carcinoma, NOS, metastatic 1
neuroendocrine carcinoma 1
cyst-associated renal cell carcinoma 1
mixed germ cell tumor, metastatic 1
oxyphilic adenocarcinoma 1
adenocarcinoma, NOS 1
carcinoid tumor, malignant 1
leiomyosarcoma, NOS, metastatic 1
squamous cell carcinoma, NOS 1
chondrosarcoma, NOS 1
anaplastic large cell lymphoma, T-cell and Null cell type (ALCL) 14
endometrial stromal sarcoma, metastatic 1
squamous cell carcinoma, NOS, micro-dissected 25
chondrosarcoma, NOS 1
adenocarcinoma, NOS, metastatic 1
papillary serous cystadenocarcinoma, micro-dissected 3
Hodgkin´s disease, NOS, micro-dissected 4
primary cutaneous CD30+ T-cell lymphoproliferative disorders 5

Figure 3.24 Hs *GPHR* is highly expressed in a number of neoplasms

Heat map showing the expression of Hs *GPHR* in several neoplasms using three different probe sets (220642_x_at, 223531_x_at, 222140_s_at) using the Human Genome 47k array platform. The number of sample and percentage expression potential scale is shown on figure. The highest and lowest expression found are magnified and indicated by blue insets showing these regions. Data obtained using Genevestigator (Hruz et al., 2008).

3.3 Discussion

3.3.1 The GTG/GPHRs are found only in eukaryotes and have a conserved domain structure

The GTG/GPHRs exist solely in eukaryotes including animals, plants, fungi, chromalveolata and protists (Figure 3.3). Phylogenetic analyses show that these proteins generally group into clades with members of their own kingdoms. This suggests that these proteins diverged as a result of speciation and orthologues are similar due to a common origin such as a common ancestral structure or function. Convergent evolution in contrast, arises due to ecological or physical drivers towards a similar result, with structure and function arising independently (Reece et al., 2011). For example, three families of sugar kinases have distinct 3D folds, but catalyse chemically equivalent reactions on identical substrates (Bork et al., 1993). Since the function of the GTG/GPHRs has not been characterised in many of these organisms, it cannot be concluded that these proteins are evolutionary grouped according to function. The animal GTG/GPHRs cluster into lower animals, higher animals or an undefined group (Figure 3.3). The plant GTG/GPHRs group into monocot, dicot and less-defined algal clades (Figure 3.3). This supports the hypothesis that the GTG/GPHRs have converged evolutionary due to speciation since monocots such as *O. sativa* are more closely related to other monocot species compared to those in the dicot group such as *A. thaliana*.

Multiple sequence alignments also demonstrate that the GTG/GPHRs have two conserved domains, DUF3735 and ABA-GPCR (Figure 3.1), both containing conserved residues (Table 3.3) that could be functionally important (Figure 3.4 – Figure 3.6 and Table 3.2). These DUF3735 and ABA-GPCR domains are predicted to be in the N- and C-terminal halves of GTG/GPHRs (Bateman et al., 2004). Conserved domains in membrane proteins are relatively common, for example, the fasciclin-like adhesion domain is conserved in GPI-anchored membrane proteins (Kawamoto et al., 1998, Elortza et al., 2003). Since the GTG/GPHR domain architecture is conserved across many species it is unlikely that these proteins have converged evolutionarily. Furthermore, convergent evolution of domain architecture has been reported as a rare occurrence (Gough, 2005). Phylogenetic analysis, sequence alignments, mutation rates and comparative genomics were conducted on domains of known structure across 62 genomes. Results indicated that divergence was driven by evolutionary descent rather

than functional necessity (Gough, 2005). Therefore, it is likely that the GTG/GPHRs are diverged through speciation rather than convergence.

The DUF3735 contains a signature sequence GxxxxxxLSG, which is always predicted to be in TM5 (Table 3.3). This conserved signature sequence found in 179 GTG/GPHRs suggests it may be functionally important. Glycines are important for helix-helix interactions since they do not possess a side chain. This property of glycines confers its indispensability in maintaining protein structures. GxxxG and GxxxxxxG are conserved motifs for helix-helix associations (Javadpour et al., 1999, Russ and Engelman, 2000, Liu et al., 2002). Although not identical to these motifs, the glycine residues may still be functionally important as they are conserved in so many species.

A completely conserved tyrosine residue lies in the ABA-GPCR domain of all GTG/GPHRs identified (Table 3.3). Tyrosine residues can be phosphorylated and therefore important in kinase signalling pathways (Nelson et al., 2008). The conserved residues in GTG/GPHRs could be functionally important and provide scope for mutagenesis studies. This was also highlighted by the SNAP2 method (Bromberg and Rost, 2007, Bromberg et al., 2008, Hecht et al., 2013), which was used to predict functional effects of amino acid substitutions in several GTG/GPHRs (Figure 3.14). The results showed that there are many amino acid substitutions that could cause a functional effect and many of these amino acids lie within the estimated DUF3735 and ABA-GPCR domain sequences (Figure 3.14). This method has been used to identify variants of OAT4 to show the importance of conserved residues (Shima et al., 2010), hence it could be a useful prediction system for mutagenesis studies to study structure-function relationships in the GTG/GPHRs. Furthermore SNAP2 has an estimated ~82 % accuracy when predicting functional effects of substitutions, as shown by mutagenesis studies of *E. coli Lac*I repressor (Hecht et al., 2013).

There is also a conserved proline residue in the GTG/GPHRs, predicted to reside in TM3 (Figure 3.10). This residue has been shown to functionally important in the *D. melanogaster GPHR*^{P91L} KO mutant since several abnormal growth phenotypes were observed. This proline residue must have a functionally significant role, which could be in regulating selectivity of channels. For instance, the At CLCa transports nitrate coupled to protons and contains a conserved proline 160. When proline was mutated to serine, it resulted in loss of nitrate selectivity but unaltered anion proton exchange activity. Furthermore, in contrast to At CLCa expression, At CLCa^{P160S} failed to

complement either the At *CLCa* KO plants or Sc *CLC*-Δ mutant, signifying channel nitrate selectivity is essential to function (Wege et al., 2010). Therefore, mutagenesis studies could be designed bearing the SNAP2 results in mind.

In addition to the predicted DUF3735 and ABA-GPCR domains, it was previously reported by Pandey et al. (2009) that PROSITE predicts additional motifs for the At GTGs (Sigrist et al., 2012). Both At GTGs were predicted to have an Ras GTPase-activating protein motif within the intracellular segment after TM5, and an ATP-/GTP-binding site in the ABA-GPCR domain (these are highlighted in Figure 3.4). These predicted motifs and domains coincide with the proposed roles of the At GTGs as GPCR-type G proteins that have GTP-binding capacity and GTPase activity (Pandey et al., 2009). In this study, the At GTGs are not predicted to have a Ras GTPase-activating protein motif using PROSITE. However, they do contain a predicted Protein kinase ATP-binding region signature (Figure 3.4), which actually corresponded to the previously predicted ATP-/GTP-binding site identified by Pandey et al. (2009). Furthermore the nucleotide-binding/Protein kinases ATP-binding motif coincidently is predicted to lie within the ABA-GPCR domain. The multiple sequence alignment shows that there are high regions of similarity between the At GTG and other sequences aligned for the predicted nucleotide-binding/Protein kinases ATP-binding motif (Figure 3.4). Therefore it is unusual that none of the other GTG/GPHR sequences are predicted to contain this predicted domain except for Hs GPHR, which did not exhibit GTP binding or GTPase activity (Pandey et al., 2009). The PROSITE predictions could be inaccurate and experimental evidence would be required to support that these sequences correspond to the predicted domains. For example, these domains could be cloned and tested to see if they carry out their proposed function such as GTP-binding or ABA-binding by using fluorescent dyes for example. However, this would be more difficult with the DUF3735 which has an unknown function.

3.3.2 The GTG/GPHRs are predicted to be membrane proteins with several structural elements

Computational analyses were conducted to predict structural features of the GTG/GPHRs. Known datasets of membrane proteins were analysed using different methods and indicated that TM predictions were not 100 % accurate (Krogh et al., 2001). For this reason, many different prediction methods were used to analyse

membrane protein topology of the GTG/GPHRs. The results indicate that the GTG/GPHRs contain multiple (7 – 11) TM domains using individual methods (Table 3.5), and the consensus membrane protein topology predictions indicate these proteins have nine TM segments (Figure 3.9 and Figure 3.13), which is consistent with the predictions previously obtained (Maeda et al., 2008, Pandey et al., 2009, Jaffé et al., 2012). The ConPred II consensus topology prediction method was used to analyse the plant GTG/GPHRs since it is used as an outsourced prediction method component of ARAMEMNON and has been used to accurately predict protein topology. For example, ConPred II was used to show that an Arabidopsis Golgi-localised P_{2A}-Type ATPase, ECA3, was predicted to contain ten TM domains (Mills et al., 2008), which was consistent with crystallisation studies for this protein family (Palmgren and Nissen, 2011). These results suggest that the GTG/GPHRs are membrane proteins with more than seven membrane-spanning regions, which is not typical of classical GPCR structure (7TM proteins) (Rosenbaum et al., 2009). Moreover, there are inconsistencies between prediction methods; for example the *A. thaliana* and *O. sativa* GTG/GPHRs have very similar predicted TM segments (see Figure 3.10), which contrasts with the prediction by PredictProtein (Figure 3.11) showing slightly more variation. Therefore experimental evidence would be needed to clarify the number of TM regions, which cannot be predicted using computational methods alone.

The GTG/GPHRs are predicted to contain many protein binding sites and a few DNA-binding sites located in similar regions (Figure 3.11), suggesting that these could be regions of functional importance. Protein-binding sites could correspond to binding regions for G proteins for example, which have been shown to bind to the Arabidopsis GTGs (Pandey et al., 2009). The predicted DNA-binding sites are unusual since there is no experimental or computational evidence suggesting a role as a DNA-binding protein such as a transcription factor. Since these are are only predictions, no firm conclusion can be drawn on these predicted binding sites without experimental evidence. Furthermore the GTG/GPHRs have regions that correspond to protein disorder (Figure 3.11), which could be important for their biological role. For example, many proteins that are predicted to be disordered are involved in signalling and regulatory responses (Bracken et al., 2004). The large intracellular region between TM5 and TM6 is an area in GTG/GPHR sequences predicted to be intrinsically disordered with and

could be functionally important. Again, experimental evidence would be required to support these predictions.

3.3.3 The *GTG/GPHR* genes are widely expressed

The *GTG/GPHR* genes are widely expressed at all stages of development for plant, animal and fungal models. For example, in *A. thaliana* and *O. sativa* the *GTG/GPHR* genes are highly expressed in all stages of development and anatomy (Figure 3.16 and Figure 3.17). This data coincides with the spatial expression of the At *GTG* genes demonstrating these are expressed in cotyledons, stems, roots, flowers, leaves and guard cells (Pandey et al., 2009, Jaffé et al., 2012). The At *GTG* genes were either up or downregulated under certain pertubations. Many experiments indicated that there was an approximate two-fold increase in At *GTG* expression in BL assays (Figure 3.19). Since phytohormone sensitivity is a reoccurring theme in GTG/GPHR function, conducting BL assays would be important to clarify whether the *A. thaliana gtg1 gtg2* mutants exhibit any hyposensitivity similar to the ABA hyposensitivity previously reported (Pandey et al., 2009). Interestingly, At *GTG* genes are also upregulated in the *rgs1-2* mutant (Figure 3.19); this supports the idea that the GTG/GPHRs have a role in GPCR signalling. The *rgs1-2* mutant exhibits hyposensitivity to ABA and sucrose; this was not related to osmotic stress since seed germination was similar to WT when treated with mannitol and sorbitol (Chen et al., 2006). The *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants also exhibited a normal response to mannitol (Jaffé et al., 2012). Therefore these expression data cannot exclude the concept that the At GTGs do not have a role in phytohormone and/or GPCR signalling. In addition, At *GTG* expression was also downregulated in long-day treatments compared to short-day treatments (Figure 3.19). Previously it has been shown that in long-day conditions, the *gtg1-1 gtg2-1* mutant exhibited earlier flowering compared to WT. Therefore, short-day condition experiments could be conducted to test whether there are any other effects observed in *A. thaliana gtg1 gtg2* mutant plants.

Sc *GTG* and Mm *GPHR* are expressed at all stages of development (Figure 3.20 and Figure 3.21). Hs *GPHR* is widely expressed, particularly in the kidneys (Figure 3.23). This data is in accordance with the analysis of conducted by Maeda et al. (2008), which indicated that the Hs *GPHR* is predicted to be ubiquitously expressed. These data indicate that animal, plant and fungal *GTG/GPHR* genes are widely

expressed from expression databases. Spatial expression of the At GTGs has been shown using GUS staining, however the tissue expression is yet to be carried out experimentally using an animal model. Furthermore, if these proteins function as proposed Golgi pH regulators (Jaffé et al., 2012), these proteins would be expected to be relatively ubiquitous. Interestingly, the *GTG/GPHRs* are highly expressed in a number of *H. sapiens* and *M. musculus* neoplasms (Figure 3.22 and Figure 3.24). For example in *M. musculus*, the *GPHR* expression is highest in ascitic body fluid neoplasm cell lines and lymposarcoma cell lines, although this was only conducted using two samples. It is also very highly expressed in the alimentary system neoplasm cell lines including primary neoplasm cell lines of the alimentary system and pancreas cell lines. In many samples Mm *GPHR* is also expressed in liver and nervous system neoplasms. In *H. sapiens*, *GPHR* is highly expressed in malignant and metastatic myelomas and carinomas as well as in Hodgkin's disease. In light of these data, the GTG/GPHRs could be suggested as a novel drug target.

Computational analysis of the GTG/GPHRs indicates that these proteins are highly conserved but experimental data would be required to confirm whether there is a conserved role for these proteins. Furthermore, a number of predicted structural annotations suggest that these proteins are membrane proteins with conserved domains and some predicted binding sites. These are only predictions and would need to be backed up using experimental data. However, these predictions are useful for designing experiments to analyse their roles in plants, animals and fungi. The GTG/GPHRs are highly conserved at the protein level, therefore there could be a conserved function for these proteins; this could be investigated using functional complementation assays. Several conserved residues have been identified for the GTG/GPHRs, therefore mutagenesis studied could be conducted to investigate their roles in these proteins. Lastly, the localisation predictions for the GTG/GPHRs indicate differing locations (Figure 3.15 and Table 3.6), however these are only predictions and not always accurate. For example WoLF PSORT predicts the animal GPHR proteins to be plasma membrane localised; however there are independent reports indicating that these are localised to the Golgi/ER network (Maeda et al., 2008, Charroux and Royet, 2014). It is therefore important in this study to confirm this by conducting independent localisation studies since there are also inconsistencies within the literature regarding this (section 1.4.2).

CHAPTER 4

Functional characterisation of plant GTGs

4.1 Introduction

4.1.1 Using Arabidopsis for transgenic research

The plant model Arabidopsis (introduced in section 1.5.2) has been used extensively to reveal information about the physiological roles of plant proteins such as transporters (Krysan et al., 1996) and receptors (Inoue et al., 2001). The use of transgenic Arabidopsis including mutants has greatly enhanced such discoveries. Transgenic Arabidopsis may have specific genes knocked out, knocked down or overexpressed. To confirm that mutant phenotypes result from a specific gene mutation, functional complementation assays can be carried out using the gene of interest. For example, *CDP* in Arabidopsis encodes cytochrome P450 and a *cdp* transfer-DNA (T-DNA) mutant exhibited a number of mutant phenotypes including dwarfism and male sterility. Ectopic overexpression of *CPD* cDNA rescued these mutant phenotypes, suggesting that these phenotypes were both caused by the absence of *CDP* (Szekeres et al., 1996). Heterologous expression of genes from another organism can be carried out using transgenic Arabidopsis to study gene families. Additionally, reporter constructs can be expressed in plants to study organ, tissue and subcellular localisations (section 1.4). Therefore a range of gene functional analyses has been conducted using transgenic plants. In Arabidopsis, mutagenesis can be induced using a range of methods including ethyl methyl sulphonate (EMS) (Maple and Møller, 2007), radiation (Shirley et al., 1992, Østergaard and Yanofsky, 2004) and introduction of either T-DNA or transposon insertions (Krysan et al., 1999, Parinov and Sundaresan, 2000). In this thesis, plant T-DNA insertional mutants (T-DNA mutants from herein) are used for functional analysis.

4.1.1.1 T-DNA insertional mutants

In Arabidopsis*, Agrobacterium tumefaciens* (*Agrobacterium* from herein) is used widely for transferring foreign DNA into plant genomes, and can be used for generating plant mutants (for explanation see section 4.1.1.2). The T-DNA of *Agrobacterium*

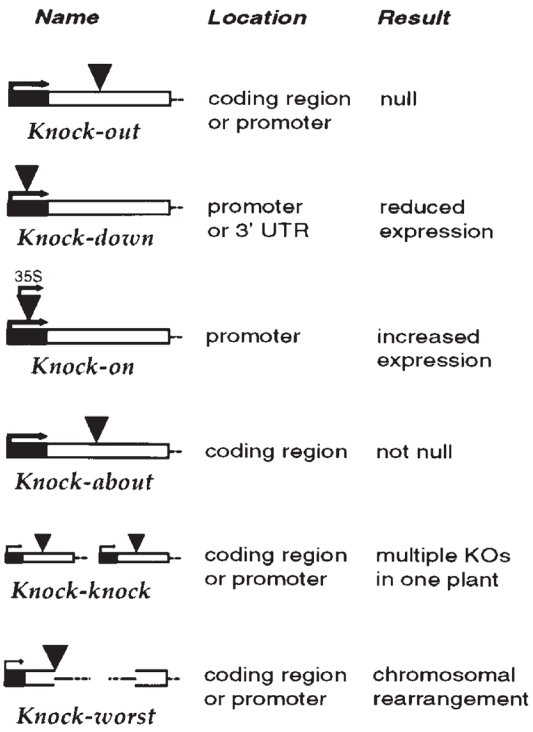


Figure 4.1 The outcomes of T-DNA insertions into the Arabidopsis chromosome

The standard nomenclature, or "knockology", describes the several possible outcomes of T-DNA mutants. White box = coding region, black region with an arrow = promoter, black triange = T-DNA insert. KO, knockout; UTR, untranslated region. Figure taken from Krysan et al. (1999).

randomly integrates into chromosomal DNA and in the context of generating mutants a diverse number of outcomes may result (Figure 4.1). Using this technology T-DNA mutants have been generated in Columbia (Col) (Alonso et al., 2003) and Wassilewskija (Ws) ecotypes. There are two collections used in this project for the latter: the Wisconsin (Sussman et al., 2000) and Flag T-DNA lines (http://publiclines.versailles.inra.fr/), both generated using different accessions, Ws-2 and Ws-4 respectively. There are also Col T-DNA mutant libraries but only the SALK collection has been used in this study.

T-DNA express (http://signal.salk.edu/cgi-bin/tdnaexpress) is an Arabidopsis gene mapping database used to identify T-DNA mutants for gene functional analysis. Crossing single T-DNA mutants can be used to create a double mutant. Comparative functional studies of genes from a family can be used to quantify relative functional contributions that may not be otherwise evaluated by expression studies alone. For instance, the growth and membrane properties of both single and double mutants of AKT1 and AKT2 K^+ channels in WT and mutant seedlings were compared. The results showed that AKT1 facilitated growth-sustaining uptake of K^+ into roots and cotyledons, whereas AKT2 had a larger role in K^+ uptake in leaf mesophyll cells (Dennison et al., 2001). Creating double mutants can either exacerbate existing single mutant phenotypes or reveal phenotypes of functionally redundant genes.

4.1.1.2 The binary vector system for Arabidopsis transformation

Agrobacterium are routinely used for transgenic technology by exploiting their tumour-inducing (Ti) plasmid containing the T-DNA and the virulence region harbouring the *vir* genes. The *vir* genes are necessary for integrating T-DNA into the host genome (Chilton et al., 1977, Garfinkel and Nester, 1980, Ooms et al., 1980, Ooms et al., 1982). To produce transgenic plants a binary vector system has been developed, based upon the use of a pair of plasmids. One plasmid contains the *vir* genes (Ti-plasmid) and the other carries the T-DNA (Figure 4.2; binary vector) (Hoekema et al., 1983). Since the T-DNA region integrates into the plant genome, this region is replaced by the gene of interest. The binary vector is subsequently transformed into *Agrobacterium* harbouring the Ti-plasmid. The floral dip method is commonly used with *Agrobacterium* to mediate Arabidopsis transformation (Clough and Bent, 1998). The *vir* genes of the Ti-plasmid facilitate the integration of the manipulated T-DNA

(now carrying the gene of interest) into the plant genome. This method successfully transforms many Arabidopsis ecotypes including Ws and Col (Clough and Bent, 1998). In this way, the binary vector system can genetically engineer transgenic plants.

A set of Gateway-compatible binary destination vectors has been developed for a range of gene functional analysis in plants (Curtis and Grossniklaus, 2003). The destination vectors pMDC32, pMDC83, pMDC107 and pMDC99 have the Gateway cassette (Karimi et al., 2002) positioned adjacent to a dual 35S CaMV promoter (Figure 4.3). The CaMV 35S promoter (Franck et al., 1980) is highly active in transgenic cells allowing constitutive gene expression (Odell et al., 1985). Genes driven by duplicated 35S promoter sequences (2x 35S from herein) are highly expressed (Kay et al., 1987), as in the pMDC series of destination vectors. Transformed Arabidopsis must undergo a positive selection process in order to isolate transgenic Arabidopsis homozygous for the T-DNA insertion (Figure 2.1). Therefore, selection markers such as kanamycin and gentamycin, are essential for identifying positive transformants. In this study, the destination vectors use hygromycin for selection (Figure 2.1 and Figure 4.3) (Curtis and Grossniklaus, 2003).

The pMDC vector series are used for a variety of functional analyses. For example, *ROF2* is an Arabidopsis chaperone gene encoding a peptidyl-prolyl *cis-trans* isomerase induced upon intracellular acid stress (Bissoli et al., 2012). *ROF2* overexpression under the 35S promoter conferred tolerance to intracellular acidification, revealing that *ROF2* induction and intracellular acidification are typical effects of many stresses (Bissoli et al., 2012). Furthermore, the 35S promoter is commonly used to drive constitutive expression of GFP fusion proteins. The localisation of two members in the Arabidopsis CLC family was revealed by transient expression of a 35S driven protein-GFP fusion in onion epidermal cells. The results show that At CLC-e and At CLC-f target to thylakoid membranes in chloroplasts and Golgi membranes, respectively (Marmagne et al., 2007). The 35S promoter can therefore drive gene expression for a variety of functional studies.

pMDC83 is a useful expression vector for functional localisation studies since it permits C-terminal GFP tagging. For example, the slow anion channel-associated 1 (SLAC1) was cloned into pMDC83 for transient expression in onion epidermis and tobacco protoplasts to demonstrate its expression in the plasma membrane of guard

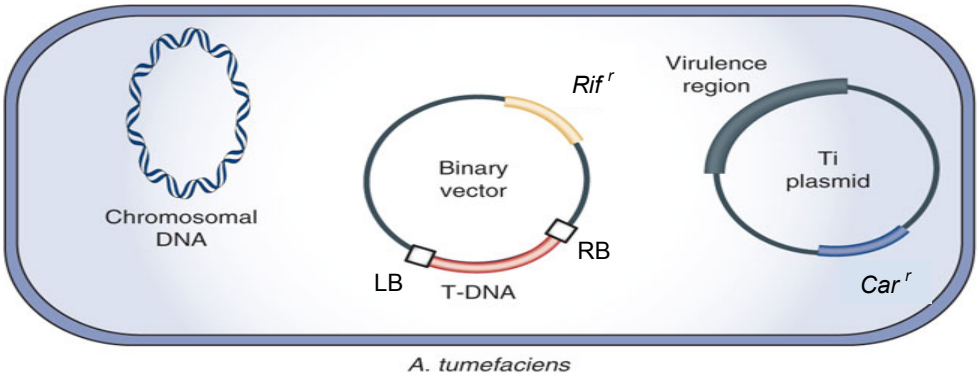


Figure 4.2 The binary vector system in *Agrobacterium* for plant transformation

Diagram showing the binary vector system that carries two plasmids in *Agrobacterium* strain GV3850. The Ti-plasmid contains the carbenicillin resistance gene (*Car* ^{r}) and carrys the *vir*-region. The manipulated T-DNA region (containing DNA of interest) is on the binary vector containing the rifampicin resistance gene (*Rif* ^{r}). DNA sequences of interest are cloned between the left (LB) and right border (RB) sequences of the T-DNA region, which is transferred to the host. Figure adapted from Michielse et al. (2008).

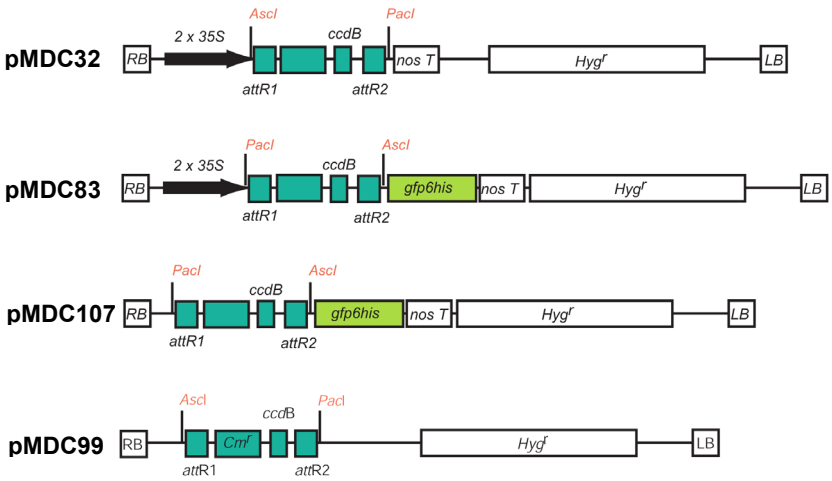


Figure 4.3 Schematics of the Gateway-compatible Arabidopsis destination vectors

Structure of the destination vectors pMDC32, pMDC83, pMDC107 and pMDC99 showing the recombination sites flanked by *Asc*I and *Pac*I recognition sites, with hygromycin resistance genes (*Hyg* ^{r}) for selection in Arabidopsis. LB = left border, RB = right border sequences of the T-DNA region (Figure 4.2). pMDC32 harbours 2x 35S promoter for constitutive expression. pMDC83 harbours 2x 35S promoter with a C-terminal GFP-6His-tag. pMDC107 is promoterless with a C-terminal GFP-6-His-tag. pMDC99 is promoterless with no C-terminal tags. pMDC107 and pMDC99 are promoterless for genomic construct insertion to allow for gene expression under native promoters. Figure taken from Curtis and Grossniklaus (2003).

cells (Vahisalu et al., 2008). pMDC83 has also been used for stable, heterologous expression of *O. sativa* metal tolerance protein 1 (Os MTP1) in Arabidopsis. The At *mtp1-1* mutant exhibits zinc hypersensitivity, presenting stunted growth and chlorosis (Kobae et al., 2004). The expression of Os *MTP1:GFP* functionally complemented the zinc sensitivity phenotype as well as demonstrated vacuolar membrane localisation (Menguer et al., 2013). Therefore, pMDC83 is a very versatile plant expression protein that can be used for both transient and stable expression of GFP fusion proteins as well as functional complementation assays. Gene expression driven by their own promoters is also important, especially for studying the native localisation of proteins. For example, the cation exchangers, CHX21 and CHX23, in Arabidopsis are essential for pollen tube guidance. These localised to the ER when gene expression was driven under their native promoters via the pMDC107 vector (Lu et al., 2011). Therefore native promoters are integral for protein function and localisation studies.

4.1.2 Localisation of the Arabidopsis GTGs

The subcellular localisation of the plant GTG/GPHRs is yet to be fully determined as there are contrasting reports (section 1.4.2). Further localisation studies are therefore required to support either report. Previous studies have focussed on At GTG1 and therefore it is important to determine if At GTG2 also shows a similar localisation. The pMDC107 vector can facilitate this analysis as it is promoterless with a C-terminal GFP tag; a gene of interest can be inserted into this vector driven by its own promoter (Figure 4.3).

4.1.3 The role of GTG/GPHRs in plants

As mentioned in section 1.2, the GTG/GPHRs have been studied in the plant model Arabidopsis (Pandey et al., 2009, Jaffé et al., 2012). Different T-DNA alleles for Arabidopsis *GTG1* and *GTG2* null single mutants were isolated and used to generate different Arabidopsis *gtg1 gtg2* double mutants (Table 4.1). In both reports, phenotypic analyses were carried out whereby the single and double mutant phenotypes indicate that the At GTGs function redundantly since no overt phenotypes were observed in the single mutants (Pandey et al., 2009, Jaffé et al., 2012). One of the proposed roles for the GTG/GPHRs is that they are ABA receptors as shown by the hyposensitivity of the *gtg1-1 gtg2-1* double mutants (Pandey et al., 2009). However, this hyposensivity to

ABA was not observed using different *gtg1 gtg2* mutant alleles (*gtg1-2 gtg2-2* and *gtg1-3 gtg2-3*), suggesting that there may be alternative roles for these proteins (Jaffé et al., 2012). The *gtg1-1 gtg2-1* mutants have also been reported to exhibit hyposensitivity to brassinolide (BL). In the dark, WT and *gtg1-1 gtg2-1* roots and hypocotyls were comparable in length. However in 10 nM BL, the *gtg1-1 gtg2-1* mutant displayed a lack of root growth inhibition compared to WT seedlings. A similar but less severe response to BL was reported in the hypocotyls, but this was not actually significant (Alvarez et al., 2013). Pandey et al. (2009) and Alvarez et al. (2013) predominantly focussed on ABA- and other phytohormone-related studies while a more extensive analysis of mutant growth was carried out by Jaffé et al. (2012). The *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants exhibited growth, development and fertility defects, manifested by reduced root, hypocotyl and fresh weight growth; reduced silique length, seed yield and seed number per silique; and large distended cell growth of hypocotyls (Jaffé et al., 2012). To extend the phenotypic analysis of the GTG/GPHRs in plants, in this study the mutant phenotypes for all three different *gtg1 gtg2* mutants are compared.

Table 4.1 Arabidopsis *gtg1 gtg2* double mutants used in this project

The different Arabidopsis *gtg1 gtg2* double mutants isolated from null single mutants available in different ecotypes: Wassilewskija (Ws), Columbia (Col). Two Ws WTs are used: Ws-2 (ABRC ID: CS2360) and Ws-4: (NASC ID: N5390) as these are the accessions used for generating the Wisconsin (Sussman et al., 2000) and FLAG (http://publiclines.versailles.inra.fr/) T-DNA lines respectively.

Ecotype	Double mutant name	Corresponding WT (symbol used in this thesis)	Reference
Ws	*gtg1-1 gtg2-1*	Ws-2	Pandey et al. (2009)
Col	*gtg1-2 gtg2-2*	WT	Jaffé et al. (2012)
Ws	*gtg1-3 gtg2-3*	Ws-4	Jaffé et al. (2012)

The GTG/GPHRs have been studied so far only in the dicot plant Arabidopsis. This study goes further by initiating GTG/GPHR study in monocot *O. sativa* as T-DNA mutants are becoming more available. *O. sativa* is an important crop in both agriculture and biology. The genome of the *Oryza sativa L. ssp japonica* has been sequenced and found to constitute 389 Mb (Project, 2005). Since the genome sequence release, a number of resources are becoming increasingly available including the Rice Genome Annotation project database containing various resources of sequence and annotation data for the rice genome (http://rice.plantbiology.msu.edu/).

4.1.4 Brassinosteroids and their role in plant physiology

Brassinosteroids (BRs) are a class of polyhydroxylated, steroid hormones important for plant developmental responses including cell elongation (Azpiroz et al., 1998), cell division (Cheon et al., 2010) and seed germination (Steber and McCourt, 2001), as well as pollen tube growth and root inhibition (Mandava, 1988). There are different types of BRs such as BL and (homo)castasterone (Grove et al., 1979, Adam and Marquardt, 1986). The functions of BRs in plants have been studied by characterising several Arabidopsis BR mutants such as *dwf1* (Choe et al., 1999), *dwf4* (Choe et al., 1998) and *det2* (Li et al., 1996). For example, in the light *det2* and *cpd* mutants exhibit dwarfism (stunted, smaller plant phenotype). However, in dark conditions they have short hypocotyls, which is characteristic of light-grown seedlings. The dwarfism and hypocotyl defects resulted from reduced longitudinal growth as shown by sections of these seedlings, and thus revealed BRs role in cell elongation (Asami et al., 2000). *det* and *cdp* mutant phenotypes can also be rescued by the application of BRs, which indicated that these phenotypes were caused by the lack of BR (Szekeres et al., 1996). Examples of different types of BR mutants and their defect are summarised in Table 4.2.

BR inhibitors are often used to study the physiological roles of BR in plants. For example, brassinozole is a BR inhibitor that caused short hypocotyl growth in WT Arabidopsis seedlings grown in the dark; this was morphologically similar to BR-deficient mutants. The phenotype was rescued upon addition of BL, indicating that the short hypocotyl growth was due to BR biosynthesis inhibition and that BRs play a role in light-dependent responses (Asami et al., 2000). BRs also function in pectin-dependent cell wall homeostasis. Pectins are components of the cell wall, which can be

modified by pectin methylesterase (PME). PMEs are important for root growth, as shown by the use of a potent PME inhibitor, epigallocatechin gallate (EGCG), which inhibited root elongation (Lewis et al., 2008). Altering PME genetically or pharmacologically resulted in modifications in growth behaviour. It was shown that BR signalling regulates cell wall feedback mechanisms involving pectin-dependent homeostasis of the cell wall (Wolf et al., 2012).

As mentioned (section 4.1.3) *gtg1-1 gtg2-1* mutants exhibited hyposensitivity to BL (Alvarez et al., 2013). Therefore, BL will be used in phenotypic assays in this study to examine BR responses of different *gtg1 gtg2* mutants. BL is also commonly used in many other BR response experiments; for example, *exo* mutants exhibited reduced BL-induced hypocotyl growth. *EXO* is important for cell expansion and plays a role in controlling BR responses in the root (Schröder et al., 2009).

Table 4.2 Examples of brassinosteroid mutants

BR mutant	Example phenotypes	Defective in	Reference
dwf1	Dwarfism	BR synthesis	Choe et al. (1999)
dwf4	Dwarfism		Choe et al. (1998)
det2	Dwarfism		Li et al. (1996)
det3	Dwarfism		Azpiroz et al. (1998)
cpd	Dwarfism		Asami et al. (2000)
bri1	Dwarfism, insensitive to BR in root and hypocotyl inhibition assays	BR signalling	Clouse et al. (1996)
bin2			Li et al. (2001)

4.1.5 Mutagenesis studies for defining structure-function relationships

Orthologue proteins sharing similar roles may have functionally conserved residues within a domain or region of the protein. Mutagenesis can be employed for this kind of study. In plants, mutagenesis has been used to study protein function. For example, lysine-40 (K40) of Arabidopsis α-tubulin 6 (TUA6) was found to be an important residue for cell division and expansion. Mutated forms of TUA6 were constitutively expressed in Arabidopsis, which was used to analyse the importance of K40 in TUA6 function. The replacement of K40 with either alanine or glutamine, hydrophobic and polar in biochemical nature respectively, resulted in several growth

defects (Xiong et al., 2013). This indicates that substituting residues of a dissimilar biochemical nature can cause loss-of-function phenotypes. Sometimes however proteins are able to tolerate non-conservative mutations, as seen for the *Shaker* K^+ channel where three non-conservative substitutions in the signature sequence had no effect on K^+ selectivity (Heginbotham et al., 1994). Therefore, amino acids substitutions should be carefully considered.

A conserved glycine was identified in TM5 of GTG/GPHRs (section 3.2.3) and its possible role in helix-helix interactions was discussed (section 3.3.1). The GTG/GPHRs have been described as an anion channel to regulate Golgi pH (Maeda et al., 2008). This glycine could have a structural role in the protein or have a key function such as ion selectivity, like in the K^+ channel for example (Heginbotham et al., 1994, Doyle et al., 1998). In this chapter, the function of this conserved glycine in Arabidopsis GTG function is explored.

4.1.6 Aims

A number of aims have been set out to further investigate the physiological roles of the GTGs in plants.

1. To confirm the T-DNA insert positions in previously isolated *gtg1-1 gtg2-1* mutants and to test whether *gtg1-1 gtg2-1* mutants exhibit similar phenotypes to previously isolated *gtg1-3 gtg1-3*.
2. To extend the phenotypic analysis conducted to date on the *gtg1 gtg2* mutants.
3. To investigate the role of a conserved glycine residue in GTG/GPHR function by generating an At GTG1 mutant variant. This will be expressed in *gtg1 gtg2* mutants to observe the effects on the complementation capability of At GTG1.
4. To investigate the cellular localisation of At GTG2 by generating several independent transgenic Arabidopsis lines expressing At GTGs under native promoters; these can also be used for testing the rescue effect of the At GTGs in phenotypic assays.
5. To isolate an *O. sativa gtg* mutant for future phenotypic characterisation.

4.2 Results

4.2.1 Confirmation of the T-DNA *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants

Jaffé et al. (2012) showed data for two *gtg1 gtg2* mutants (*gtg1-2 gtg2-2* and *gtg1-3 gtg2-3*) which were different mutants to the previously isolated *gtg1-1 gtg2-1* mutant. The Ws reference line corresponding to *gtg1-3 gtg2-3* used by Jaffé et al. (2012) is Ws-4, but the reference line corresponding to *gtg1-1 gtg2-1* used by Pandey et al. (2009) is Ws-2 (Table 4.1). In order to compare the phenotypes of different *gtg1 gtg2* mutants, a homozygote *gtg1-1 gtg2-1* mutant was obtained from a segregating T_2 population as well as Ws-2 (supplied by Prof. S. Assman, Penn State University, USA). The independent T-DNA *gtg1 gtg2* mutants isolated are shown in Table 4.1. Since both Ws *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants have different corresponding WTs these will be referred to as Ws-2 or Ws-4 respectively throughout this chapter. First, it was important to confirm the genotype of these plants.

The *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants isolated by Jaffé et al. (2012), were generated from single *gtg1* and *gtg2* mutant alleles isolated from the SALK (Alonso et al., 2003) and FLAG (http://publiclines.versailles.inra.fr/) T-DNA lines in the Col and Ws ecotypes, respectively. The T-DNA inserts are in distinct sites of both At *GTG1* and At *GTG2* in all *gtg1 gtg2* mutants (Figure 4.4). The *gtg1-1 gtg2-1* mutant isolated by Pandey et al. (2009) is contains T-DNA inserts in intron 2 and exon 7 of At *GTG1* and At *GTG2* respectively. *gtg1-2 gtg2-2* contains T-DNA inserts in exon 12 and intron 10 of At *GTG1* and At *GTG2* respectively. *gtg1-3 gtg2-3* contains T-DNA inserts in intron 1 and intron 7 of At *GTG1* and At *GTG2*, respectively. Since *gtg1-2 gtg2-2* is in the Col ecotype, different to the other *gtg1 gtg2* mutants (Ws), most of the comparisons here were conducted between *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* alone.

Genomic DNA was extracted from *gtg1-1 gtg2-1*, *gtg1-3 gtg2-3* and respective WTs for genotyping (Figure 4.5). Primers across the predicted T-DNA insert sites (Figure 4.4) confirmed WT or mutant At *GTG* alleles. The presence of 1813 bp (using primers FR1 F and GTG1 2R; Figure 4.4 A) and 2202 bp products (using primers GTG2 2F and GTG2 1R; Figure 4.4 B) are indicative of WT At *GTG1* (Figure 4.5 A lane 7) and At *GTG2* respectively (Figure 4.5 B lane 9). The T-DNA border primer (JL202; Figure 4.4) which anneals to the pD991-AP3 plasmid used for Wisconsin T-

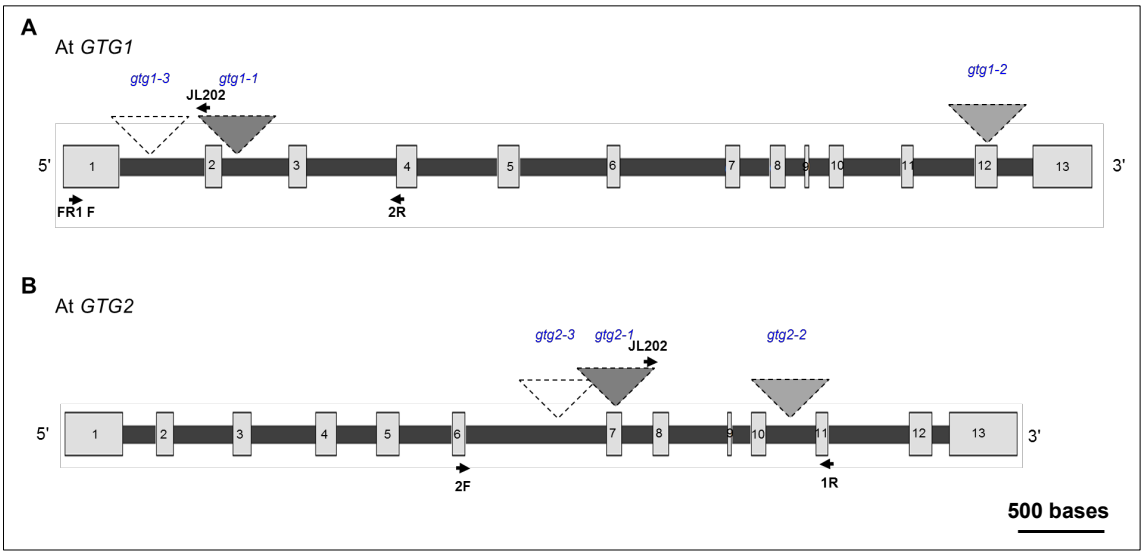


Figure 4.4 Genomic structures and the T-DNA insert sites of At *GTG*

The At *GTG* genomic structures showing the exons and introns of the genes and T-DNA insertion sites drawn to scale (scale bar: 500 bases). Exons are shown in light shade grey; introns are shown in dark shade grey. Symbols below the structures indicate primers used for genotyping; arrows indicate direction of primers (primer sequences outlined in Table 2.10). The JL202 primer annealing to the T-DNA insert is shown above the dark grey triangle. T-DNA insertion sites indicated by triangles above the gene and the mutant allele generated is shown in blue above the triangle: dark grey = *gtg1-1*, light grey = *gtg1-2* and white = *gtg1-3*. **A**. At *GTG1* gene structure with T-DNA insertion sites indicated. **B**. At *GTG2* gene structure with T-DNA insertion sites indicated.

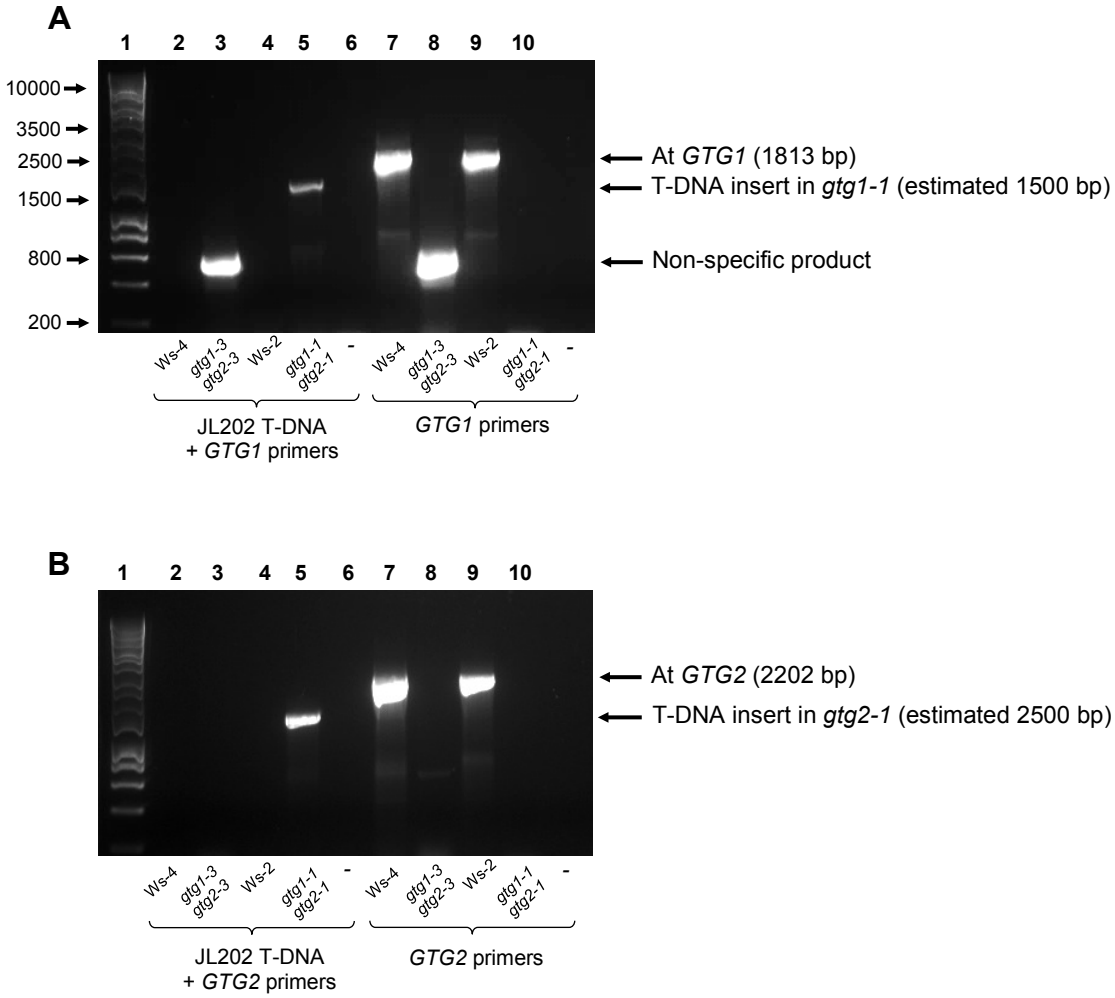


Figure 4.5 *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants are distinct T-DNA mutants

PCR and gel electrophoresis using genomic DNA extracted from different WT and *gtg1 gtg2* mutants (*gtg1-3 gtg2-3* and *gtg1-1 gtg2-1*). Ws-2 and Ws-4 are the corresponding WTs for *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* respectively. Predicted and estimated products shown on figure (primer sequences outlined in Table 2.10). Lane 1: molecular markers. Non-specific products produced by primer mispairing are shown on figure. Dash (-) indicates negative control using water instead of DNA in PCR mix. **A**. FR1 F and JL202 primers amplify an estimated 1000 bp fragment of the T-DNA insert in At *GTG1* in lane 5 but not in lane 3 (although there is a predicted 600 bp unspecific product). FR1 F and GTG1 2R primers (At *GTG1* primers) amplify a fragment of At *GTG1*. Ws-4 and Ws-2 show At *GTG1* presence in lanes 7 and 9 respectively (predicted 1813 bp). **B**. JL202 and GTG2 1R primers amplify an estimated 1000 bp fragment of the T-DNA insert in At *GTG2* in lane 5 but not in lane 3. GTG2 2F and GTG2 1R primers amplify a fragment of At *GTG2*. Ws-4 and Ws-2 show the presence of At *GTG2* in lanes 7 and 9 respectively (predicted 2202 bp).

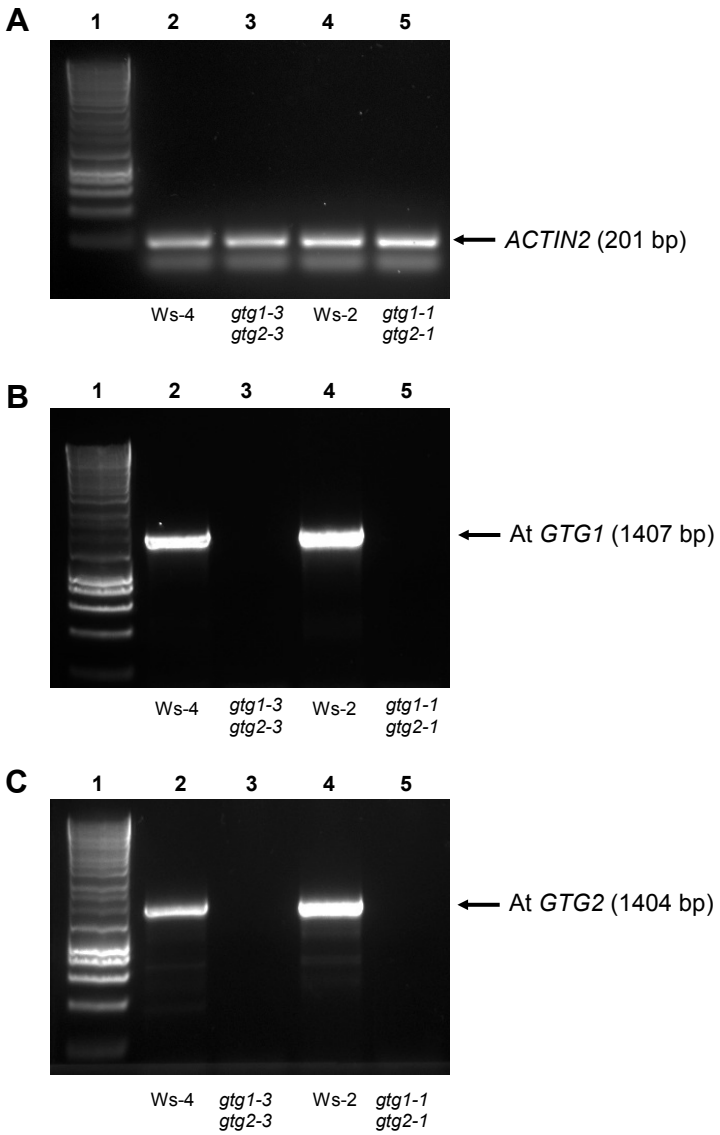


Figure 4.6 *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* are null mutants

PCR and gel electrophoresis using cDNA synthesised from RNA extracted from different WT and *gtg1 gtg2* mutants (*gtg1-3 gtg2-3* and *gtg1-1 gtg2-1*). Ws-2 and Ws-4 are the corresponding WTs for *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* respectively. Predicted and estimated products shown on figure. Lane 1: molecular markers (see Figure 4.5 A for predicted molecular marker sizes and primer sequences are outlined in Table 2.10). **A**. Actin2 F and Actin2 R primers amplify a fragment of *ACTIN2*. These plant lines show a consistent level of *ACTIN2* expression at the RNA level (201 bp predicted product). **B**. Ws-4 and Ws-2 show At *GTG1* expression in lanes 2 and 4 respectively using primers AtGTG1 F and AtGTG1 R (predicted 1407 bp), while *gtg1-3 gtg2-3* (lane 3) and *gtg1-1 gtg2-1* (lane 5) do not express At *GTG1*. **C**. Ws-4 and Ws-2 show At *GTG2* expression in lanes 2 and 4 respectively using primers AtGTG2 F and AtGTG2 R (predicted 1404 bp), while *gtg1-3 gtg2-3* (lane 3) and *gtg1-1 gtg2-1* (lane 5) do not express At *GTG2*.

DNA mutants (Sussman et al., 2000), was used with either a primer annealing to At *GTG1* (FR1 F; Figure 4.4 A) or At *GTG2* (GTG2 1R; Figure 4.4 B). The results show the estimated product for the T-DNA in either At *GTG1* or At *GTG2* of *gtg1-1* and *gtg2-1* respectively (Figure 4.5 lane 5 of A and B). *gtg1-3 gtg2-3* does not contain this T-DNA as shown by the absence of the estimated product (Figure 4.5 lane 3). Therefore, *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* are distinct mutants and their genomic structures are displayed in Figure 4.4.

To test whether these T-DNA inserts generate KO (null) mutants, RNA was extracted for cDNA synthesis to confirm by PCR the At *GTG* transcipt expression levels in the *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants. To check that the RNA was good quality and there was no genomic DNA contamination, primers amplifying a fragment of *ACTIN2* were used (*ACTIN2* = AT3G18780). All plant lines express *ACTIN2* (lanes 2 – 5) as shown by amplification of the predicted 201 bp products (Figure 4.6 A). As expected, WT (lane 2 and 4) plant lines express full-length At *GTG1* and At *GTG2*, (using primers AtGTG1 1F + AtGTG1 STOP 1R and AtGTG2 1F + AtGTG2 STOP 1R respectively). This was shown by the amplification of the predicted 1407 and 1404 bp products, respectively (Figure 4.6 B and C). The results indicate that both *gtg1-1 gtg2-1* (lane 5) and *gtg1-3 gtg2-3* (lane 3) do not contain full-length At *GTG* transcripts (Figure 4.6 B and C). This was shown by the absence of the predicted 1407 and 1404 bp products for At *GTG1* and At *GTG2*, respectively. These results confirmed that the *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* are distinct double *GTG* KO mutants.

4.2.2 Different *gtg1 gtg2* T-DNA mutants share similar phenotypes

The different T-DNA *gtg1 gtg2* double mutants isolated are listed in Table 4.1. *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* exhibit several mutant phenotypes including reduced root and hypocotyl growth; defective cellular growth in the hypocotyls; reduced silique length and seed number. Either of the At GTGs can functionally complement *gtg1 gtg2* mutant defects, as shown by restored root, hypocotyl and cellular growth. Therefore mutant phenotypes are characteristic of knocking out both At *GTG1* and At *GTG2* and these genes function redundantly (Jaffé et al., 2012). Here it is shown that these mutant phenotypes are not only seen in the two *gtg1 gtg2* mutants (*gtg1-2 gtg2-2* and *gtg1-3 gtg2-3*) isolated by Jaffé et al. (2012) but also the *gtg1-1 gtg2-1* mutant isolated by Pandey et al. (2009).

4.2.2.1 *gtg1-1 gtg2-1* exhibits a similar root growth and fresh weight defect to *gtg1-3 gtg2-3*

In the absence of sucrose, root growth and fresh weight Arabidopsis *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants in the Col and Ws background respectively, are significantly reduced (Jaffé et al., 2012). In the absence of sucrose, the *gtg1-1 gtg2-1* mutant also displayed short root lengths as well as reduced fresh weight (Figure 4.7). The root growth was significantly reduced in both *gtg1-1 gtg-2-1* and *gtg1-3 gtg2-3*, compared to corresponding WTs (Ws-2 and Ws-4 respectively; Figure 4.7 A and C). There was also no significant difference in the root lengths between the two WTs or the two *gtg1 gtg2* mutants (Figure 4.7 A and C). Similarly, the fresh weight per seedling was reduced in both *gtg1-1 gtg-2-1* and *gtg1-3 gtg2-3*, compared to their corresponding WTs (Figure 4.7 B and D). Again, there was no significant difference in the fresh weight per seedling between the two WTs or the two *gtg1 gtg2* mutants (Figure 4.7 B and D).

4.2.2.2 *gtg1-1 gtg2-1* exhibits a similar hypocotyl growth defect to *gtg1-3 gtg2-3*

In low white light, the *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants have shorter hypocotyls and large, distended epidermal cells on the hypocotyl surface, which is exacerbated in the presence of sucrose (Jaffé et al., 2012). On sucrose, the *gtg1-1 gtg2-1* mutant also displayed short hypocotyl lengths as well as deformed epidermal cells (Figure 4.8). The hypocotyl length was significantly reduced in both *gtg1-1 gtg-2-1* and *gtg1-3 gtg2-3*, compared to corresponding WTs (Figure 4.8 D). Interestingly, Ws-2 and *gtg1-1 gtg2-1* had slightly longer hypocotyl lengths compared to the Ws-4 and *gtg1-3 gtg2-3* lines (Figure 4.8 D). Since different Ws WT accessions were used to generate the *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* lines (Table 4.1), this could be a contributing factor.

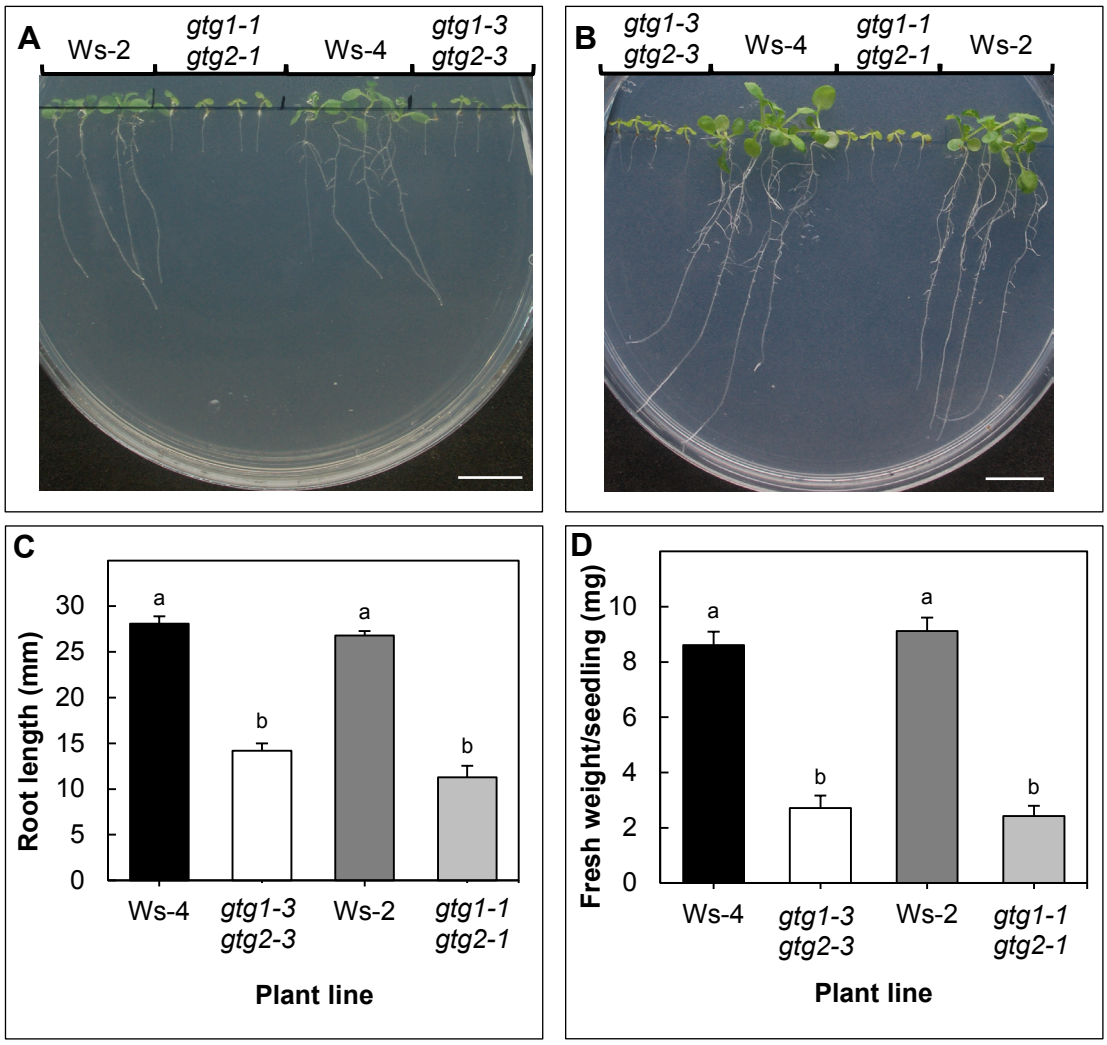


Figure 4.7 Different *gtg1 gtg2* mutants exhibit similar root and fresh weight defects

Root growth and fresh weight per seedling of *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants on 0.5 MS containing 0 % sucrose. Data shows the mean root length (**A** and **C**) and fresh weight per seedling (**B** and **D**) calculated for eight plates (+S.E.) with four seedlings per plant line, per plate, each plate containing four plant lines. Scale bar: 1 cm. Results represent one of two independent experiments using two distinct mutant lines: *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3*. Ws-2 and Ws-4 are the corresponding WTs for *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* respectively. **A**. Image displaying root growth at 10 d. **B**. Image displaying root growth at 14 d. **C**. Mean root length at 10 d (+S.E.). There is a significant difference in root lengths between plant lines as determined by a one-way ANOVA ($F_{3, 28} = 18.60$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* test. **D**. Mean fresh weight per seedling at 14 d (+S.E.). There is a significant difference in fresh weight per seedling between plant lines; one-way ANOVA ($F_{3, 28} = 23.38$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* test.

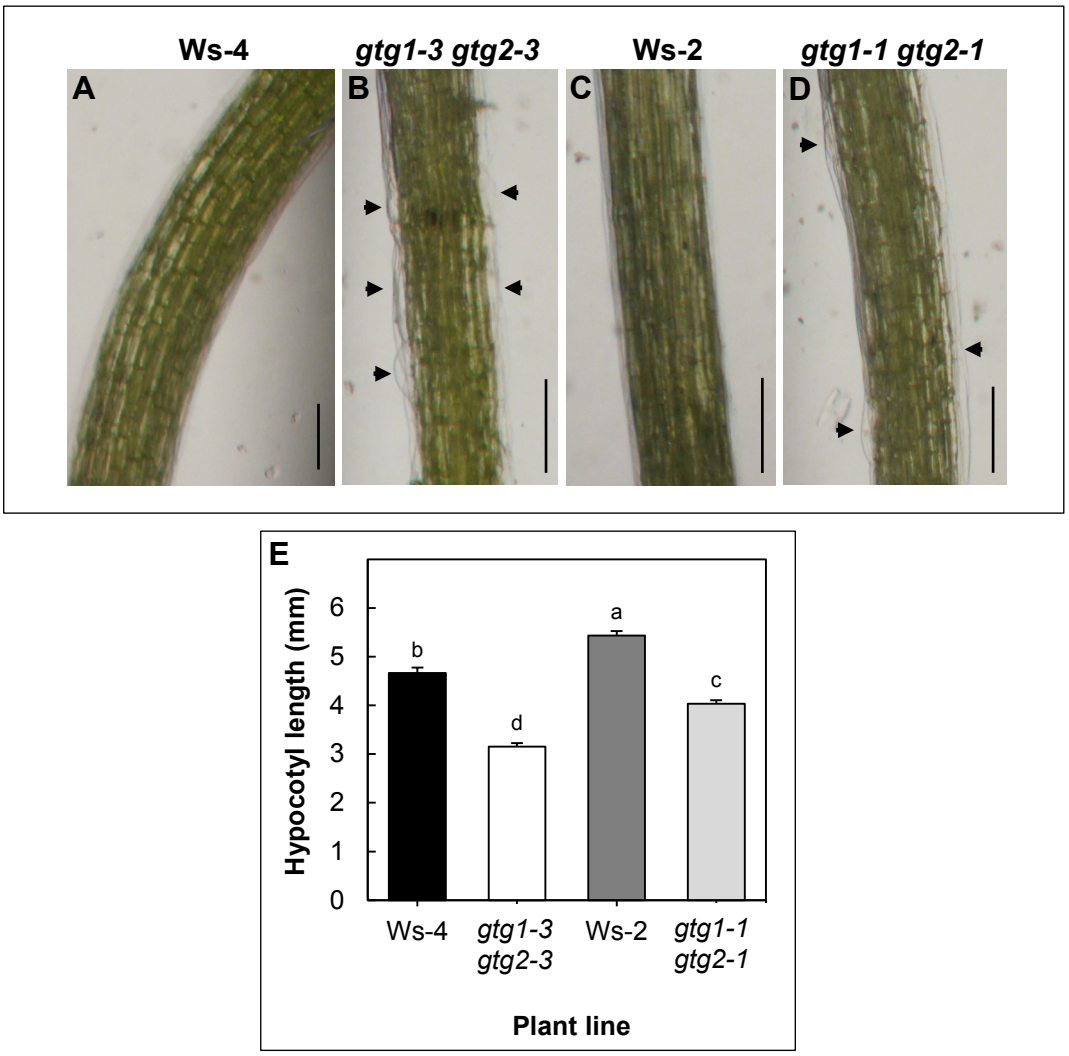


Figure 4.8 Different *gtg1 gtg2* mutants exhibit similar hypocotyl defects

Hypocotyl growth of *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants on 0.5 MS containing 1 % sucrose at 5 d. Data shows the mean hypocotyl length per seedling calculated for six plates (+S.E.) with 15 seedlings per plant line per plate, each plate containing four plant lines. **A – D**. Distended cell growth of independent *gtg1 gtg2* mutants. Light microscopy showing hypocotyls of WT (**A**), *gtg1-3 gtg2-3* (**B**), Ws-2 (**C**) and *gtg1-1 gtg2-1* (**D**). Scale bars: 200 µm, arrowheads indicate ballooning cells. **E**. Mean hypocotyl length per seedling (+S.E.) in 1 % sucrose. There is a significant difference in hypocotyl length per seedling between plant lines; one-way ANOVA ($F_{3, 20} = 65.93$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* test. Results represent one of two independent experiments using two distinct mutant lines: *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3*. Ws-2 and Ws-4 are the corresponding WTs for *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* respectively.

4.2.2.3 *gtg1 gtg2* mutants have reduced rosette growth compared to corresponding WT in short-day conditions

Genevestigator data indicated that the At *GTG* genes are downregulated in long-day conditions compared to short-day conditions (Figure 3.19). To explore whether there is an effect of short-day conditions using different *gtg1 gtg2* mutants, the rosette growth of the *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants and the *gtg1-2 gtg-2* and *gtg1-3 gtg2-3* mutants were measured over a course of 24 weeks (155 d) and 93 d respectively, in short-day conditions (8 h light/16 h darkness; light intensity 120 $\mu mol.m^{-2}.s^{-1}$). The results indicate that under short-day conditions across the whole data set, *gtg1-1 gtg2-1*, *gtg1-2 gtg-2* and *gtg1-3 gtg2-3* had smaller rosettes than their respective WTs (Figure 4.9). Overall, Ws-2 showed a significantly larger rosette growth compared to the other plant lines (Figure 4.9 C). There was no significant difference in rosette diameter between plant lines up to ~70 d however, from at 78 d onwards there was a difference in rosette diameter between plant lines (Figure 4.9 C). In general, Ws-2 had a larger rosette diameter compared to the rest of the plant lines. Both *gtg1 gtg2* mutants exhibited similar, rosette diameters throughout and these were smaller than corresponding WTs.

4.2.3 *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* flower earlier than corresponding WTs

It was previously reported that the under long-day conditions (16 h light/8 h darkness; light intensity 120 $\mu mol.m^{-2}.s^{-1}$), the *gtg1-1 gtg2-1* mutant flowers earlier than the WT (Pandey et al., 2009). In this study, both Ws *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants flower slightly earlier (~17 d) than corresponding WTs (~20 – 28 d), which varied more in time. However, this was not observed in the Col *gtg1-2 gtg2-2* mutant, which showed no significant difference between flowering time compared to its WT (Figure 4.10). Interestingly, *gtg1-3 gtg2-3* and Ws-2 took significantly longer to flower than *gtg1-1 gtg2-1* and Ws-4; this response was more marked between the *gtg1-3 gtg2-3* mutant and Ws-4 than *gtg1-1 gtg2-1*. Furthermore, the Ws-2 flowered earlier (~20 d post sowing) than the other two WT lines (~28 d for Ws-4 and ~25 for Col WT; Figure 4.10). Although both Ws WT lines are in the same ecotype, they are from different accessions (Ws-2 and Ws-4), as mentioned, which could account for the difference observed.

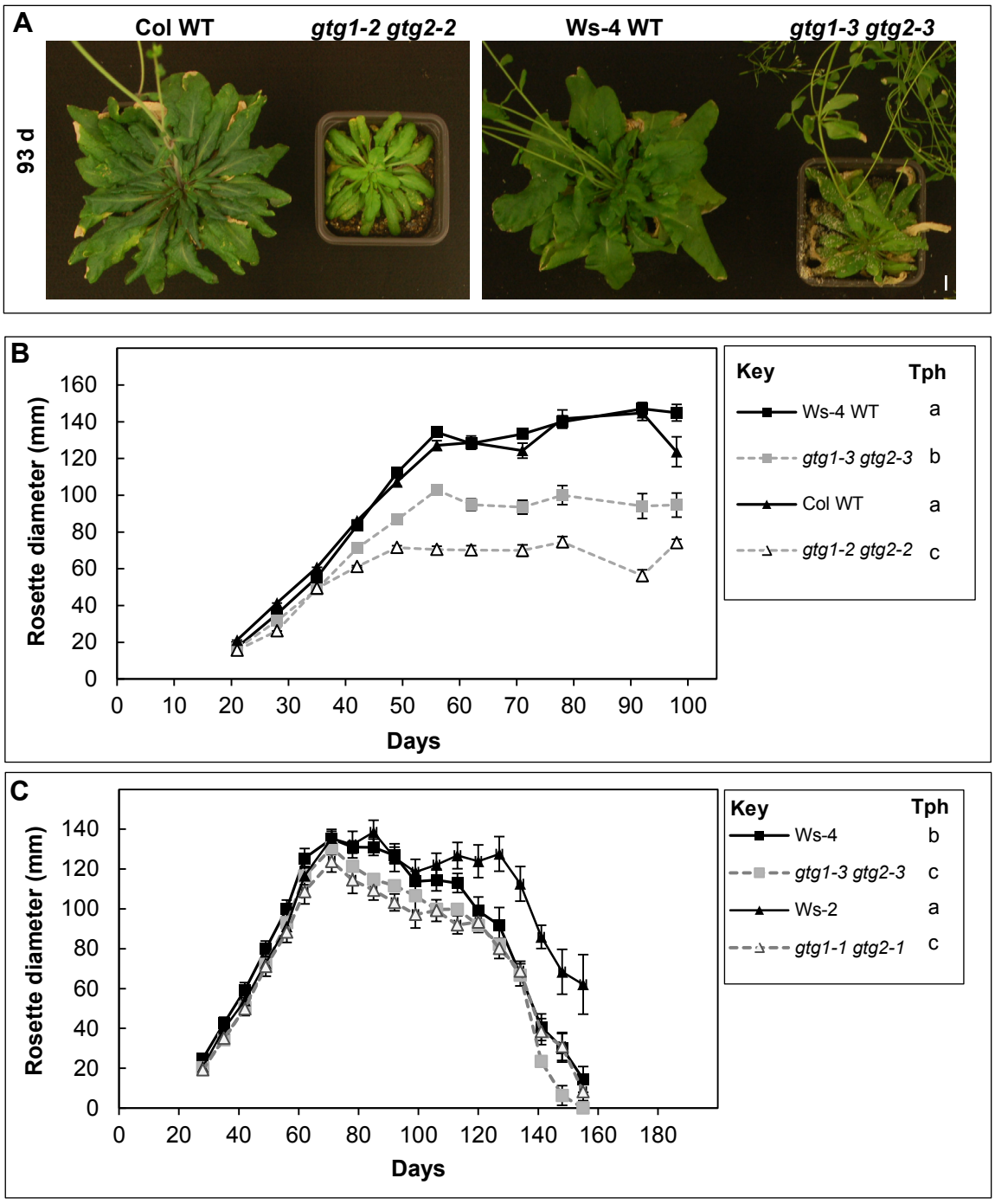


Figure 4.9 Different *gtg1 gtg2* mutants show a smaller rosette phenotype than corresponding WTs

Rosette growth of distinct *gtg1 gtg2* mutants over time under short-day conditions (section 2.2.1). **A**. Images showing the rosette growth of the Col and Ws-4 WT and *gtg1 gtg2* mutants at 93 d. Scale bar: 1 cm. **B**. Mean rosette diameter per plant of Col *gtg1-2 gtg2-2* and Ws-4 *gtg1-3 gtg2-3* mutants (+S.E.) calculated from ten plants per line, per week, over 98 d. **C**. Mean rosette diameter per plant of *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants (+S.E.) calculated for 12 plants per line, per week, over 158 d. Two-way ANOVAs indicate there is a significant

difference in rosette diameter between lines **B**: ($F_{3, 520} = 9.38$, $p < 0.0001$), **C**: ($F_{3, 907} = 15.78$, $p < 0.0001$); in rosette diameter between time **B**: ($F_{1, 520} = 915.58$, $p < 0.0001$), **C:** ($F_{1, 907} = 10.72$, $p < 0.001$) and an interaction between time and plant lines **B**: ($F_{3, 520} = 39.19$, $p = 0.0001$), **C**: ($F_{3, 907} = 12.03$, $p < 0.0001$). Different letters indicate significantly different across the entire data set; Tukey *post hoc* (Tph) test, adjacent to plant line legend. Data is representative of two (**A – B**) and one (**C**) experiment using two distinct mutant lines **B**: *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3*, **C**: *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3*. Ws-2 and Ws-4 are the corresponding WTs for *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3*. The data was collected in collaboration with A. Anzai-Jackson (**A**) and K. Peaston (**B**) (University of Southampton, UK).

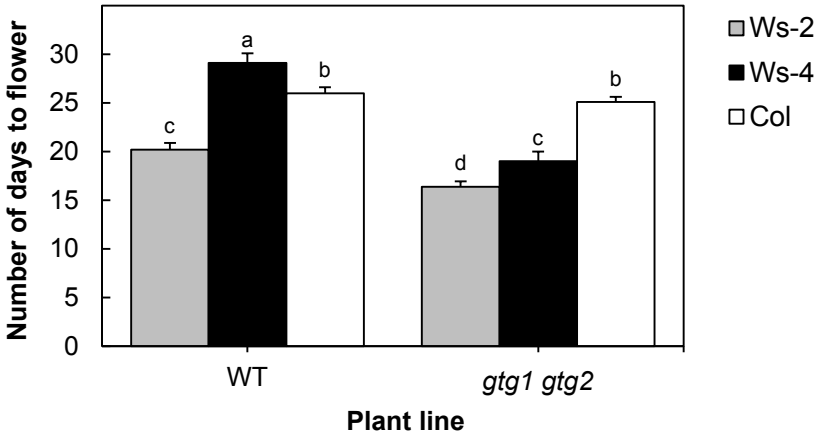


Figure 4.10 *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants flower earlier than corresponding WTs

Time taken to flower of *gtg1-1 gtg2-1*, *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants under long day conditions. Data shows the mean number of days to flower per plant (+S.E.) calculated from ten plants per line. There is a significant difference in time to flower between some plant lines; one-way ANOVA ($F_{5, 54} = 50.35$, $p < 0.0001$). WT and *gtg1 gtg2* correspond to ecotypes and respective mutants: Ws-2 (Pandey et al., 2009) = *gtg1-1 gtg2-1*, Col = *gtg1-2 gtg2-2*, Ws-4 (Jaffé et al., 2012) = *gtg1-3 gtg2-3*. Means not sharing a letter are significantly different; Tukey *post-hoc*. Data from one experiment using three distinct mutant lines: *gtg1-1 gtg2-1*, *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3*.

4.2.4 *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* exhibit a similar silique and seed defect

To investigate whether different *gtg1 gtg2* mutants exhibit a similar fertility defect observed previously by *gtg1-2 gtg2-2* (Col) and *gtg1-3 gtg2-3* (Ws) mutants (Jaffé et al., 2012), silique length, seed number and yield were measured. Similar to the *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants, *gtg1-1 gtg2-1* also exhibited shorter siliques compared to WT (Figure 4.11). Furthermore, the ecotype had an effect on the severity of this short silique phenotype in the *gtg1 gtg2* double mutants. The *gtg1-2 gtg2-2* mutant in the Col background was significantly shorter than both the Ws mutants (*gtg1-1 gtg-2-1* and *gtg1-3 gtg2-3*; Figure 4.11 A).

The total seed yield per plant was reduced in both *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants compared to their respective WT plant lines. However, there was no significant difference between the *gtg1-1 gtg2-1* mutant and Ws-2 (Figure 4.11 B). The total seed yield of the *gtg1-1 gtg2-1* mutant was similar to both the other two mutants (*gtg1-2 gtg2-2* and *gtg1-3 gtg2-3*), as well as significantly different to the Col and Ws WTs (Figure 4.11 B). Despite not displaying a significant seed yield defect, the number of seed per silique was reduced in the *gtg1-1 gtg2-1* mutant compared to Ws-2 (Figure 4.11 C and D). This suggests that the *gtg1-1 gtg2-1* had more siliques, although this was not measured. Further studies would need to be completed since this experiment was only carried out once.

4.2.5 BL has an effect on hypocotyl growth in all three *gtg1 gtg2* mutants

The response of different *gtg1 gtg2* mutants to BL was investigated. First, the hypocotyl length of *gtg1-2 gtg2-2* (Col) and *gtg1-3 gtg2-3* (Ws) mutants on varying BL concentrations in the absence of sucrose was measured (Figure 4.12). The results showed that the hypocotyl length increases in rising BL concentrations up to 1 μM. At concentrations of 0 – 100 nM BL, both *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* had shorter hypocotyl lengths than their corresponding WTs. However, at BL concentrations 250 nM and above, the hypocotyl growth of both *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* exceeded that of their WTs. A two-way ANOVA demonstrates that there was a significant difference in hypocotyl lengths between plant lines, and this was also significantly affected by BL concentration. Furthermore, there was a significant interaction between the BL concentration and the plant lines. At 250 nM BL the short hypocotyl defect in double mutants was rescued and no significant difference was seen between WT and

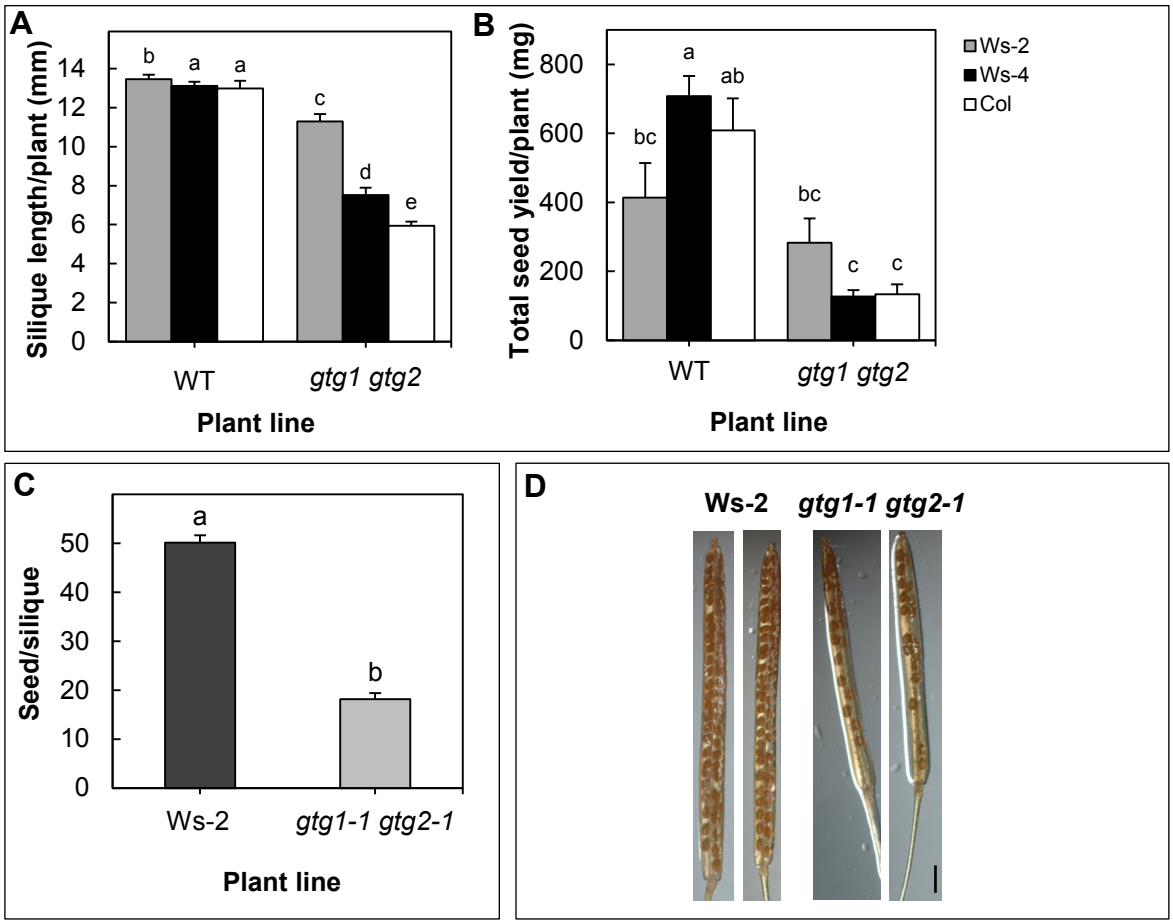


Figure 4.11 Different *gtg1 gtg2* mutants exhibit defects in silique growth and seed yield per plant

WT and *gtg1 gtg2* correspond to ecotypes and respective mutants: Ws-2 (Pandey et al., 2009) = *gtg1-1 gtg2-1*, Col = *gtg1-2 gtg2-2*, Ws-4 (Jaffé et al., 2012) = *gtg1-3 gtg2-3*. **A** – **B** using three *gtg1 gtg2* mutants (*gtg1-1 gtg2-1*, *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3*); **C** – **D** using *gtg1-1 gtg2-1*. **A**. Mean silique length per plant (+S.E.) calculated from eight siliques/plant, with ten plants per line. There is a significant difference in silique length between lines; Kruskal-Wallis test (H = 44.31, d.f. = 5, $p = 0.0001$); mean ranks = 44.9 (Ws-2), 41.7 (WT), 38.6 (Col WT), 26.4 (*gtg1-1 gtg2-1*), 13.7 (*gtg1-3 gtg2-3*) and 6.7 (*gtg1-2 gtg2-2*). Means not sharing a letter are significantly different; Mann-Whitney tests. **B**. Mean seed yield per plant (+S.E.) from ten plants. There is a significant difference in total seed yield between lines; one-way ANOVA ($F_{5, 54} = 7.10$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc*. **C**. Mean seeds per silique (+S.E.) calculated from five plants, using three siliques per plant. There is a significant difference in seeds per silique; one-way ANOVA ($F_{1, 8} = 22.35$, $p < 0.001$). Means not sharing a letter are significantly different; Tukey *post-hoc*. **D**. Images show *gtg1-1 gtg-2-1* have reduced silique length (**A**) and seed/silique (**C**) compared to Ws-2. Scale bar: 1 mm. Results are from one experiment.

double mutants. Above this concentration, the mutants had longer hypocotyls than their respective WTs (Figure 4.12).

In the next experiment, Ws mutants *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* were grown with either 0 nM or 250 nM BL, but this time with sucrose present. Both *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* exhibited a hypocotyl growth defect in the absence of BL, as previously shown (Figure 4.8 E). Interestingly although still significantly different, there was not as marked difference in hypocotyl length between the WT and mutants (Figure 4.13 I) as in Figure 4.8 E. The only difference was the presence of 0.0002 % ethanol which was used to make up BL, but whether this was responsible would need further investigation. The presence of 250 nM BL resulted in similar hypocotyl lengths between all plant lines, showing no significant difference between the WTs or either *gtg1 gtg2* mutants (Figure 4.13 I). There was also a significant effect between 0 and 250 nM BL, as well as a significant interaction between the concentration of BL and the plant line (Figure 4.13 I). The percentage growth stimulated by 250 nM BL was calculated (Figure 4.13 J). The results indicate that BL stimulated hypocotyl growth more in the *gtg1-3 gtg2-3* mutant compared to *gtg1-1 gtg2-1* and respective WTs, but growth was limited to ~6 mm for each plant line, suggesting there maybe a limit for hypocotyl growth under these light conditions (Figure 4.13 I and J).

gtg1 gtg2 mutants showed deformed cells in the hypocotyl epidermis (Figure 4.8 A – D and Figure 4.13 A – D). In 250 nM BL, the hypocotyl cells displayed some variation; some *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* mutants exhibited deformed cells (Figure 4.13 E – H), while there were some that showed cells similar to WT. This experiment was only conducted once and would need to be repeated to draw a firmer conclusion. Furthermore, it would be interesting to test a higher concentration of BL such as 500 nM. This would test whether there is a BL concentration limit needed to rescue the deformed cellular phenotype.

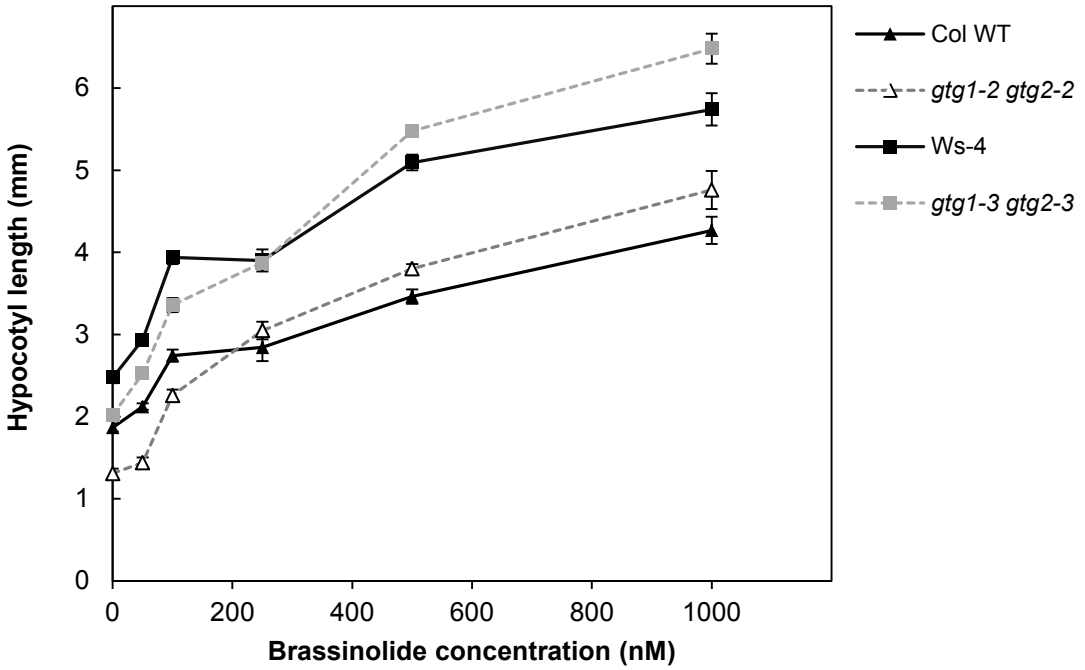
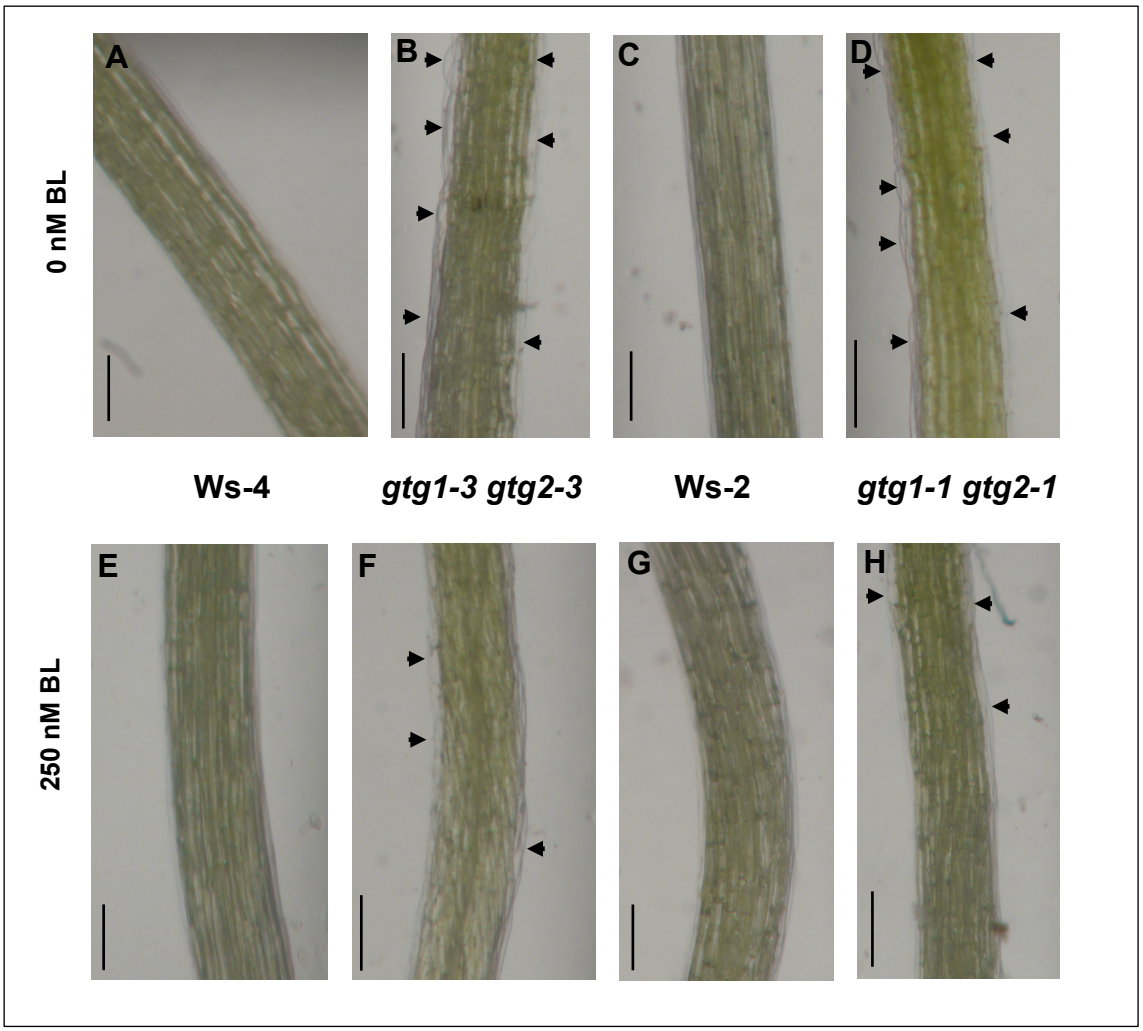


Figure 4.12 Comparing the effect of varying BL concentrations on hypocotyl growth of WT and *gtg1 gtg2* mutants in the absence of sucrose

Col *gtg1-2 gtg2-2* and Ws *gtg1-3 gtg2-3* were grown on 0.5 MS containing 0 % sucrose with either 0, 50, 100, 250, 500 or 1000 nM BL. As the BL concentration increases, the hypocotyl length of all plant lines increase. There is a significant difference between all plant lines ($F_{3, 72}$ = 72.19, $p < 0.0001$) and in BL concentrations between plant lines ($F_{5, 72}$ = 140.78, $p < 0.0001$), as well as a significant interaction between the BL concentrations and plant lines ($F_{15, 72}$ = 2.88, p = 0.001); two-way ANOVA. The data is from one experiment and was collected in collaboration with K. Peaston (University of Southampton, UK).


A
B
C
D
0 nM BL
Ws-4
gtg1-3 gtg2-3
Ws-2
gtg1-1 gtg2-1
E
F
G
H
250 nM BL

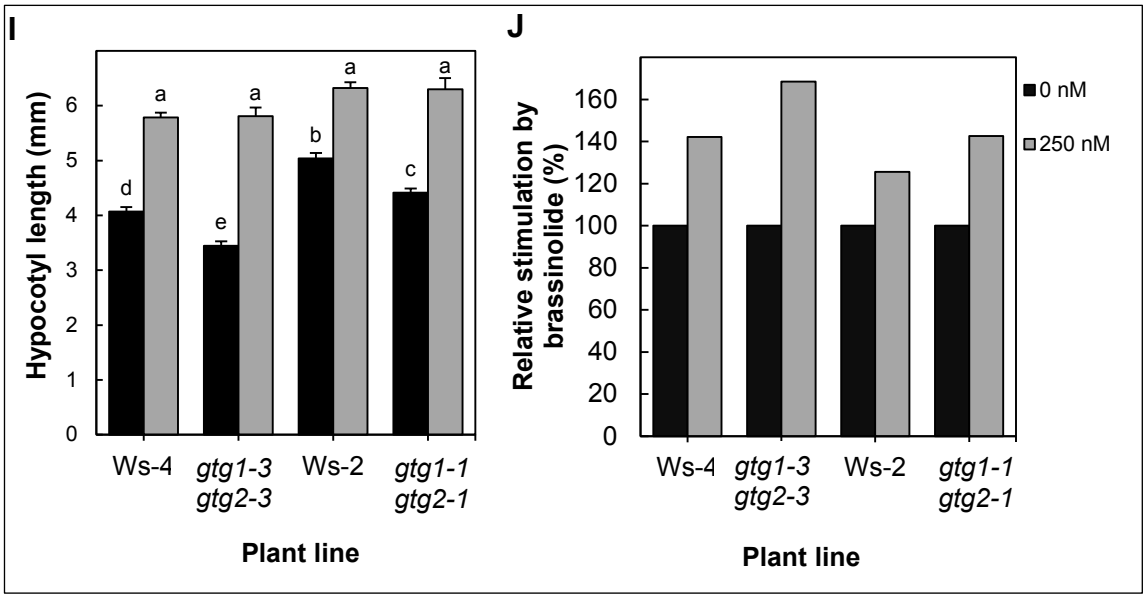

I
Hypocotyl length (mm)
0
1
2
3
4
5
6
d
a
e
a
b
a
c
a
Ws-4
gtg1-3 gtg2-3
Ws-2
gtg1-1 gtg2-1
Plant line
J
Relative stimulation by brassinolide (%)
0
20
40
60
80
100
120
140
160
0 nM
250 nM
Ws-4
gtg1-3 gtg2-3
Ws-2
gtg1-1 gtg2-1
Plant line

Figure 4.13 Comparing the effect of BL on hypocotyl growth of WT and *gtg1 gtg2* mutants in the presence of sucrose

Hypocotyl growth of *gtg1 gtg2* mutants on 0.5 MS containing 1 % sucrose and 0.0002 % ethanol, in 0 or 250 nM BL at 5 d. **A** – **H**. Distended cell growth of different *gtg1 gtg2* mutants. Light microscopy showing hypocotyls of Ws-4 (**A** and **E**), *gtg1-3 gtg2-3* (**B** and **F**), Ws-2 (**C** and **G**) and *gtg1-1 gtg2-1* (**D** and **H**); **A – D** containing 0 nM BL; **E – H** containing 250 nM BL. Scale bars: 200 μm, arrowheads indicate ballooning cells. **I**. Data shows the mean hypocotyl length per seedling calculated for six plates (+S.E.), with 15 seeds per plant line per plate. There is a significant difference in hypocotyl length per seedling between plant lines ($F_{3, 40} = 14.90$, $p < 0.0001$), in hypocotyl length between BL concentrations ($F_{1, 40} = 232.60$, $p < 0.0001$) and a significant interaction between plant lines and concentration of BL ($F_{3, 40} = 3.46$, $p < 0.025$); two-way ANOVA. Means not sharing a letter are significantly different; Tukey *post-hoc* test. **J**. The relative percentage growth stimulated by BL in hypocotyls on 1 % sucrose. The hypocotyl growth in 250 nM BL is expressed as a percentage of the growth in 0 nM BL. This is to compare the amount of growth stimulated by 250 nM BL compared to 0 nM BL (100 % growth). Results are from one experiment and collected in collaboration with A. Anzai-Jackson and K. Peaston (University of Southampton, UK).

4.2.6 Using At *GTG* genomic constructs to study functional complementation and localisation

To investigate the localisation and functionality of the At *GTG* genes, the genomic sequences including the promoters of the At *GTG* genes were transformed into Arabidopsis WT (Ws-4; this will be referred to as WT in this section) and *gtg1-3 gtg2-3* plants. The root and hypocotyl growth of transgenic *gtg1-3 gtg2-3* mutants expressing genomic At *GTG* sequences can be used to determine whether these sequences can be translated into functional proteins as well as targeted to the correct locations in transgenic Arabidopsis. The genomic At *GTG* sequences were also tagged with GFP for localisation studies.

4.2.6.1 Generating transgenic Arabidopsis expressing At *GTG* in *gtg1-3 gtg2-3* and WT plants

pMDC99 is a promoterless vector that can be used to express genes driven by their own promoters. pMDC107 is also a promoterless vector but contains a C-terminal GFP tag, which can be used in localisation studies to observe gene expression under the native promoter (Figure 4.3). Previously in the lab, At *GTG* genes were cloned into pMDC107 (Table 4.3), transformed into Arabidopsis *gtg1-3 gtg2-3* and the resulting T_2 seeds were provided for this project. Segregating plant lines were identified here and used to isolate at least three independent T_3 plant lines transformed with pMDC107 At *GTG1* and pMDC107 At *GTG2* (Table 4.4); these were used in phenotypic and localisation studies (section 4.2.7 and 4.2.7.1 respectively).

In addition, pMDC107 At *GTG1*, pMDC107 At *GTG2*, pMDC99 At *GTG1* and pMDC99 At *GTG2* were also used for Arabidopsis transformation since not all constructs had been transformed into Ws plant lines. WT and *gtg1-3 gtg2-3* were transformed with both pMDC99 At *GTG1* and pMDC99 At *GTG2* (Table 4.5). WT was also transformed with pMDC107 At *GTG1* and pMDC107 At *GTG2* (Table 4.5). All of these transgenic T_2 lines are currently growing on soil for T_3 seed collection to isolate independent, homozygous lines for future phenotypic analysis including silique length analysis (discussed in section 4.3.2).

Table 4.3 The expression vectors used for expressing At *GTG* genes in Arabidopsis

The construct specifies the promoter the gene is expressed under and the reporter, and were supplied for this project by Dr. L. E. Williams (University of Southampton, UK).

Plasmid	Construct	Explanation
pMDC107 At *GTG1*	*PAt GTG1:At GTG1:GFP*	At *GTG1* genomic sequence C-terminally tagged with GFP under At *GTG1* promoter
pMDC107 At *GTG2*	*PAt GTG2:At GTG2:GFP*	At *GTG2* genomic sequence C-terminally tagged with GFP under At *GTG2* promoter
pMDC99 At *GTG1*	*PAt GTG1:At GTG1*	At *GTG1* genomic sequence under At *GTG1* promoter
pMDC99 At *GTG2*	*PAt GTG2:At GTG2*	At *GTG2* genomic sequence under At *GTG2* promoter

Table 4.4 The transgenic Arabidopsis lines created that express At *GTG* genes

Plasmid	Construct	Transformed into background	# of independent lines
pMDC107 At *GTG1*	*Pat GTG1:At GTG1:GFP*	*gtg1-3 gtg2-3*	3
pMDC107 At *GTG2*	*Pat GTG2:At GTG2:GFP*	*gtg1-3 gtg2-3*	4

Table 4.5 Transgenic Arabidopsis currently at the T_2 stage for T_3 seed isolation

Plasmid	Construct	Transformed into background
pMDC107 At *GTG1*	*Pat GTG1:At GTG1:GFP*	WT (Ws-4)
pMDC107 At *GTG2*	*PAt GTG2:At GTG2:GFP*	WT (Ws-4)
pMDC99 At *GTG1*	*PAt GTG1:At GTG1*	WT (Ws-4)
pMDC99 At *GTG1*	*PAt GTG1:At GTG1*	*gtg1-3 gtg2-3*
pMDC99 At *GTG2*	*PAt GTG2:At GTG2*	WT (Ws-4)
pMDC99 At *GTG2*	*PAt GTG2:At GTG2*	*gtg1-3 gtg2-3*

4.2.7 *PAt GTG1:At GTG1:GFP* and *PAt GTG2:At GTG2:GFP* rescue root and hypocotyl growth defects seen in the *gtg1-3 gtg2-3* mutant

P35S driven At *GTG* expression rescued *gtg1 gtg2* mutant phenotypes including root, hypocotyl, fresh weight and cellular growth defects (Jaffé et al., 2012). To investigate whether the promoter and C-terminal tag have an effect on its ability to rescue the *gtg1 gtg2* mutant phenotypes, GFP-tagged At *GTG* were expressed in *gtg1-3 gtg2-3* under their own promoters. *gtg1-3 gtg2-3* exhibit reduced root length and fresh weight in the absence of sucrose. This defect was rescued by the expression of either *PAt GTG1:At GTG1:GFP* or *PAt GTG2:At GTG2:GFP* (Figure 4.14 and Figure 4.15). On sucrose, the short hypocotyl length and deformed cells observed in *gtg1-3 gtg2-3* was also rescued by the expression of either *PAt GTG1:At GTG1:GFP* or *PAt GTG2:At GTG2:GFP* (Figure 4.16). This indicates that under their own promoters, either At GTG1 or At GTG2 can functionally complement the loss of both At *GTG* genes in *gtg1-3 gtg2-3*, consistent with the gene functional redundancy observed (Jaffé et al., 2012).

4.2.7.1 At GTG2 localises intracellularly in *gtg1-3 gtg2-3*

The results suggest that the GFP tag does not affect At GTG functionality (section 4.2.7). Therefore, these plant lines are ideal for determining the native localisation of the At GTGs. At *GTG1* expressed under its native promoter was previously shown to be Golgi and ER localised in Arabidopsis (Jaffé et al., 2012). In this study, preliminary results (using a segregating T_2 line) show that At GTG2 under the native promoter also localised to intracellular, punctate structures (Figure 4.17), consistent with Jaffé et al. (2012).

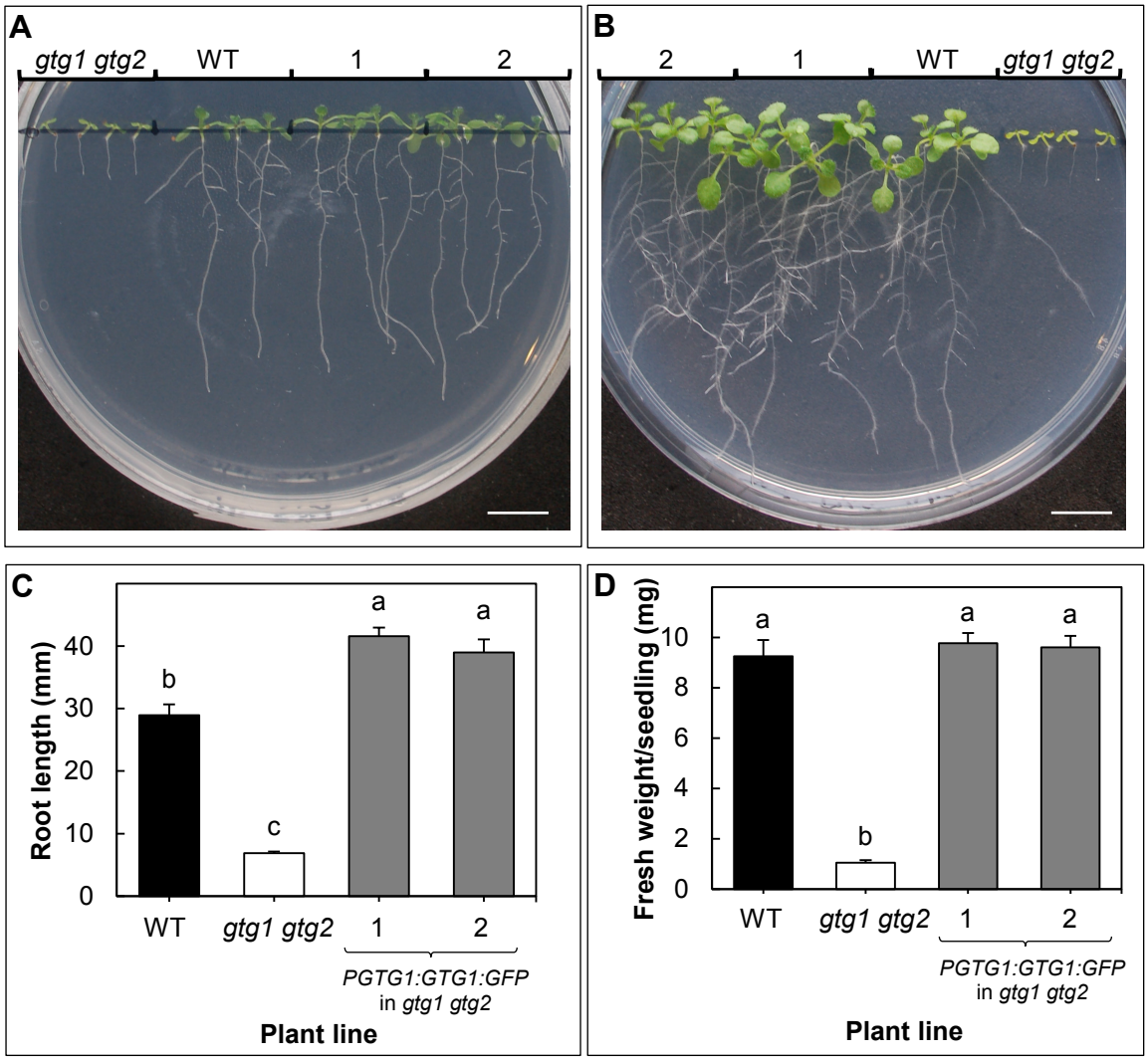


Figure 4.14 *PGTG1:GTG1:GFP* functionally complements root length and fresh weight defects seen in the Arabidopsis *gtg1-3 gtg2-3* mutant

Root growth and fresh weight per seedling of *gtg1-3 gtg2-3* (*gtg1 gtg2*) 0.5 MS containing 0 % sucrose. Mean root length (**A** and **C**) and fresh weight per seedling (**B** and **D**) calculated for six plates (+S.E.) with four seedlings per plant line per plate, each plate containing four plant lines. Scale bar: 1 cm. Results represent one of two independent experiments using three independent lines expressing *PGTG1:GTG1:GFP* (lines 1 and 2 shown). Images displaying root growth at 10 d (**A**) and 14 d (**B**). (1 and 2 = *PGTG1:GTG1:GFP* in *gtg1 gtg2* lines). **C**. Mean root length at 10 d (+S.E.). There is a significant difference in root lengths between plant lines; one-way ANOVA ($F_{3, 20} = 30.00$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* test. **D**. Mean fresh weight per seedling at 14 d (+S.E.). There is a significant difference in fresh weight per seedling between plant lines; one-way ANOVA ($F_{3, 20} = 7.63$, $p < 0.001$). Means not sharing a letter are significantly different; Tukey *post-hoc* test.

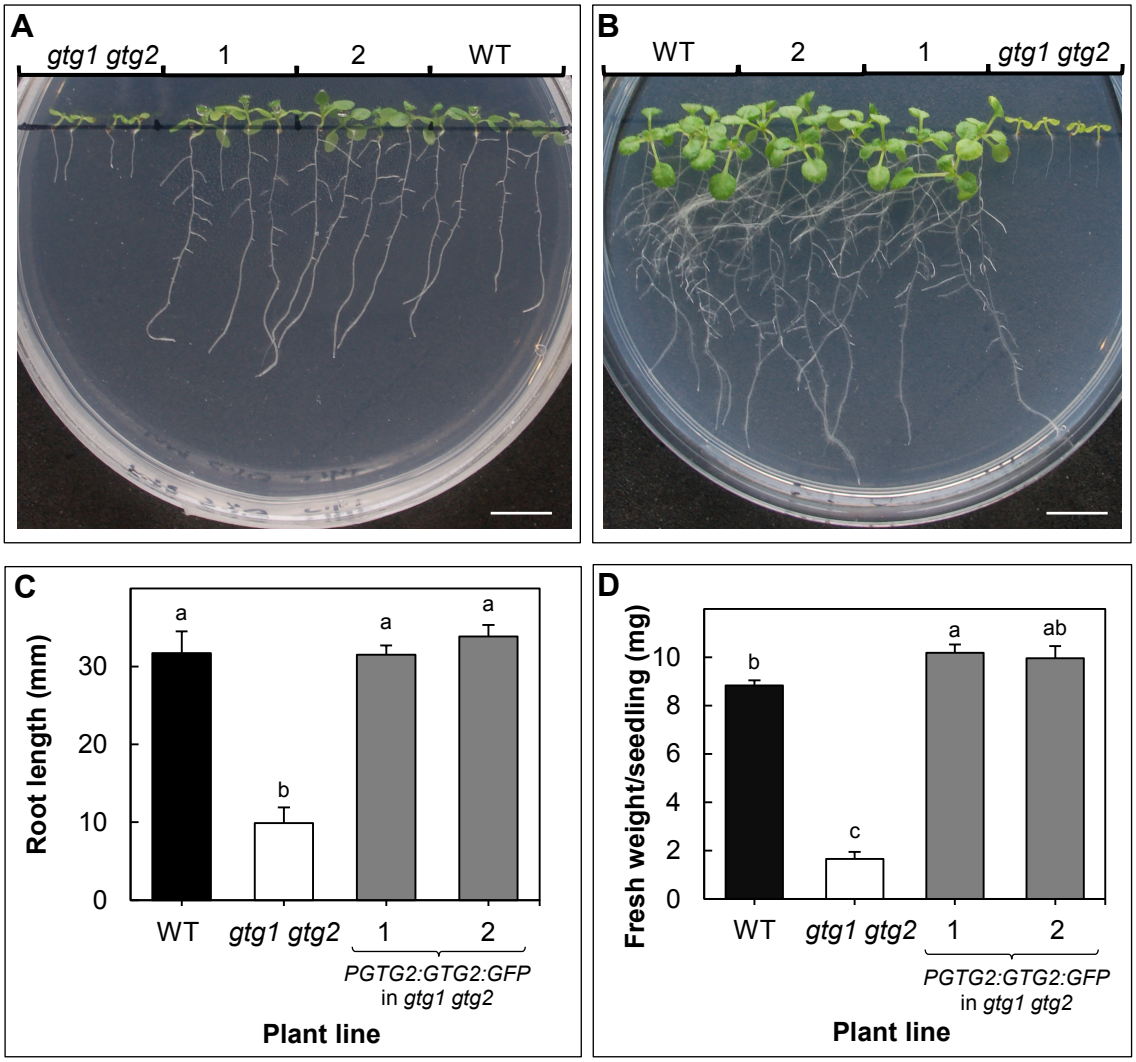


Figure 4.15 *PGTG2:GTG2:GFP* functionally complements root length and fresh weight defects seen in the Arabidopsis *gtg1-3 gtg2-3* mutant

Root growth and fresh weight per seedling of *gtg1-3 gtg2-3* (*gtg1 gtg2*) 0.5 MS containing 0 % sucrose. Mean root length (**A** and **C**) and fresh weight per seedling (**B** and **D**) calculated for six plates (+S.E.) with four seedlings per plant line per plate, each plate containing four plant lines. Scale bar: 1 cm. Results represent one of two independent experiments using four independent lines expressing *PGTG2:GTG2:GFP* (lines 1 and 2 shown). **A**. Image shows root growth at 10 d (1 and 2 = *PGTG2:GTG2:GFP* in *gtg1 gtg2* lines). **B**. Image shows root growth at 14 d (1 and 2 = *PGTG2:GTG2:GFP* in *gtg1 gtg2* lines). **C**. Mean root length at 10 d (+S.E.). There is a significant difference in root lengths between plant lines; one-way ANOVA ($F_{3, 20} = 21.97$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* test. **D**. Mean fresh weight per seedling at 14 d (+S.E.). There is a significant difference in fresh weight per seedling between plant lines; one-way ANOVA ($F_{3, 20} = 20.44$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* test.

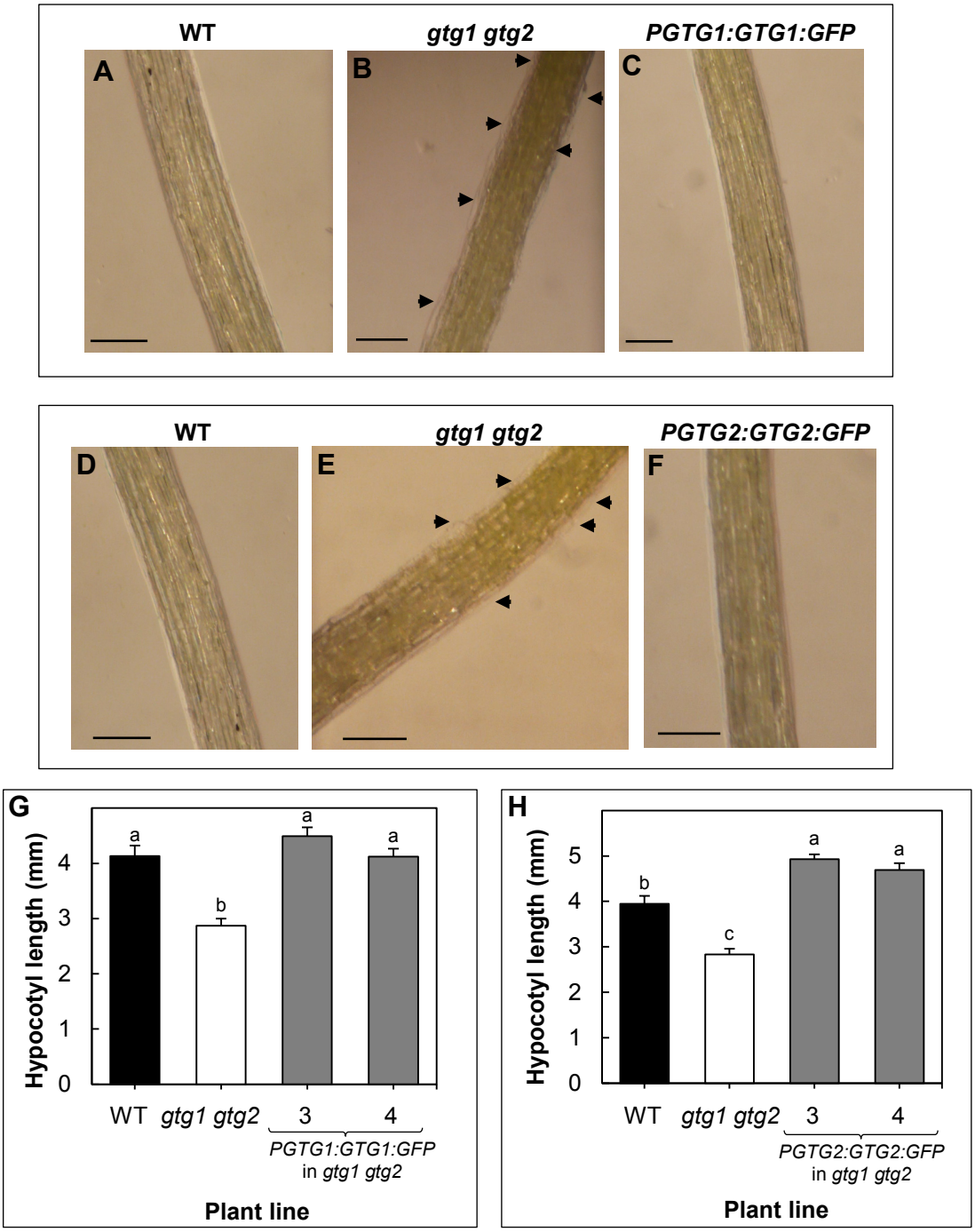


Figure 4.16 *PGTG1:GTG1:GFP* and *PGTG2:GTG2:GFP* functionally complement hypocotyl defects of the Arabidopsis *gtg1 gtg2* mutant

Hypocotyl growth of *gtg1-3 gtg2-3* (*gtg1 gtg2*) mutants on 0.5 MS containing 1 % sucrose at 5 d. Data shows the mean hypocotyl length per seedling calculated for six plates (+S.E.) with 15 seedlings per plant line per plate, each plate containing four plant lines. **A – F**. Distended cell growth of *gtg1 gtg2* is rescued by *PGTG1:GTG1:GFP* (**A – C**) and *PGTG2:GTG2:GFP* (**D – F**) expression. Light microscopy showing hypocotyls of WT (**A** and **D**), *gtg1 gtg2* (**B** and **E**), *PGTG1:GTG1:GFP* in *gtg1 gtg2* (line 3) (**C**), and *PGTG2:GTG2:GFP* in *gtg1 gtg2* (line 4) (**F**).

Scale bars: 200 μm, arrowheads indicate ballooning cells. **G – H**. The mean hypocotyl length per seedling (+S.E.) on 1 % sucrose. Results represent one of two experiments using three independent *gtg1 gtg2* lines expressing *PGTG1:GTG1:GFP* (lines 3 and 4 shown) (**G**) and four independent *gtg1 gtg2* lines expressing *PGTG2:GTG2:GFP* (lines 3 and 4 shown) (**H**). There is a significant difference in hypocotyl length per seedling between plant lines; one-way ANOVAs: (**G**) *PGTG1:GTG1:GFP* ($F_{3, 20} = 20.15$, $p < 0.0001$) and (**H**) *PGTG2:GTG2:GFP* ($F_{3, 20} = 44.61$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* tests.

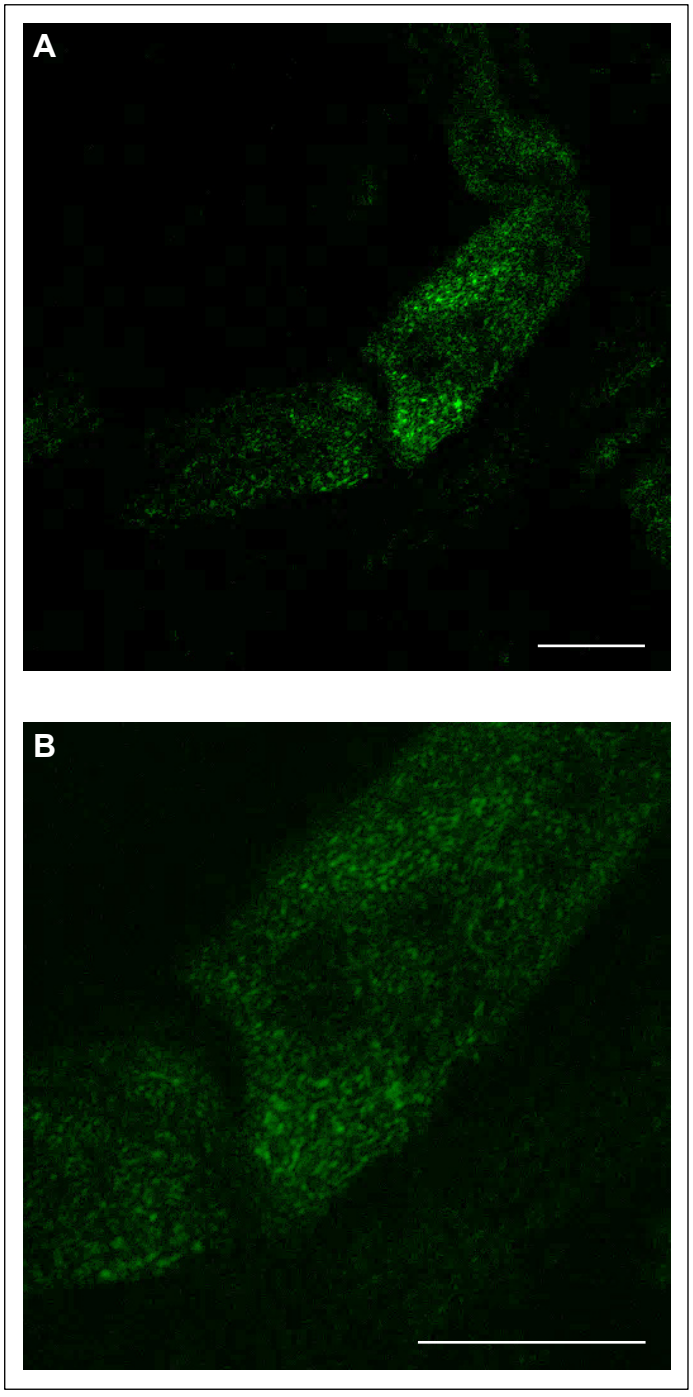


Figure 4.17 *PGTG2:GTG2:GFP* localises intracellularly in *gtg1-3 gtg2-3*

Confocal microscopy to show *PGTG2:GTG2:GFP* expression in the root tip. *PGTG2:GTG2* localises intracellularly to punctate structures. Image is representative of three seedlings from one segregating (T_2) line expressing *PGTG2:GTG2:GFP* in *gtg1-3 gtg2-3*. Images show one section of the root tip, scale bars: 10 µm. **A**. *PGTG2:GTG2:GFP* expression in a few cells. **B**. Magnified area of (**A**), centred on one cell.

4.2.8 Glycine-166 in At GTG1 is important for protein function

Root length and hypocotyl defects of Arabidopsis *gtg1 gtg2* mutants can be functionally complemented by At GTG1 expression (Jaffé et al., 2012). In section 3.2.3, a conserved glycine residue in the DUF3735 was identified in GTG/GPHRs, and this corresponded to AtGTG1^{G166}. To investigate the importance of this residue, a point-mutation of G166 in At GTG1 was created and the resulting At GTG1 mutant variant (AtGTG1^{G166L}) was expressed in Arabidopsis *gtg1-3 gtg2-3* to test its functionality.

4.2.8.1 Generating an expression vector for At GTG1^{G166L} expression in Arabidopsis *gtg1-3 gtg2-3* mutant

Site-directed mutagenesis (SDM) was employed for mutating G166 of At GTG1 to leucine (G166L). Primers were designed based on the requirements of the Stratagene Quikchange II XL kit and pENTR/D At GTG1 was used to synthesise pENTR/D At GTG1 with the G166L mutation (pENTR/D At GTG1^{G166L}). pENTR/D At GTG1 plasmid DNA was used as a template with the designed primers, for producing the mutant strand pENTR/D At GTG1^{G166L}. After mutant strand synthesis, *Dpn*I was used to digest the parental DNA (pENTR/D At GTG1) before transformation into XL10-Gold Ultracompetent cells and the plasmid DNA was extracted from a few of the colonies and sequenced in using primers M13 F (-20) and M13 R as outlined in section 2.13.4 (Table 2.5). The sequence results received confirmed that the At *GTG1* corresponded to the CDS provided by TAIR (At *GTG1* = AT1G64990, TAIR Accession # Locus:2010796), and that the codon for residue 166 corresponded to leucine instead of glycine.

The pENTR/D At GTG1^{G166L} entry vector was subjected to an overnight restriction digestion with *Pvu*I; this was to linearise the entry vector by cutting within the kanamycin resistance gene to aid insertion into and selection of the destination vectors. At GTG1^{G166L} was inserted into the destination vector pMDC32 via the LR recombination reaction using the linearised entry vector followed by transformation into *E. coli* DH5α cells. To check that At GTG1^{G166L} was inserted into pMDC32 correctly, diagnostic restriction analysis and gel electrophoresis was performed on plasmid DNA preparations extracted from individual colonies containing successful transformants. Results indicated that At GTG1^{G166L} was inserted into pMDC32, as shown by the predicted digest product sizes (Figure 4.18 A). The expression vector was sequenced in

both forward and reverse directions using primers pMDC35S and MDCnosR outlined in section 2.13.4 (Table 2.5). The sequence results received confirmed that the At *GTG1* corresponded to the CDS provided by TAIR, and that the codon for residue 166 corresponded to leucine instead of glycine (Figure 4.18 B). pMDC32 At GTG1^{G166L} was transformed into *Agrobacterium* for subsequent transformation of Arabidopsis *gtg1-3 gtg2-3*. Several independent *gtg1-3 gtg2-3* lines expressing At GTG1^{G166L} were generated following the scheme outlined in Figure 2.1. Table 4.6 shows the transgenic lines generated and used to study the structure-function relationship of At GTG1.

4.2.8.2 At GTG1^{G166L} was unable to functionally complement root, hypocotyl and fresh weight growth defects of Arabidopsis *gtg1-3 gtg2-3*

P35S:At GTG1^{G166L} expression was unable to rescue *gtg1-3 gtg2-3* phenotypes. Measurements comparing the root growth of WT, *gtg1-3 gtg2-3* and *gtg1-3 gtg2-3* lines expressing *P35S:At GTG1* or *P35S:At GTG1^{G166L}* were conducted in the absence of sucrose. *P35S:At GTG1* expression rescued the root growth defect seen in *gtg1-3 gtg2-3* (Figure 4.19 A and C). However, *P35S:At GTG1^{G166L}* expression did not restore root growth in three replicate experiments using three independent *P35S:At GTG1^{G166L}* expressing lines, although the roots were slightly longer than the *gtg1-3 gtg2-3* mutant in two of the lines but this was not significant (Figure 4.19 A and C). The fresh weight defect of *gtg1-3 gtg2-3* was also functionally complemented by *P35S:At GTG1* expression but not by *P35S:At GTG1^{G166L}* (Figure 4.19).

Similarly, while *P35S:At GTG1* expression was able to restore normal hypocotyl and cellular growth in the *gtg1-3 gtg2-3* mutant in sucrose, the expression of *P35S:At GTG1^{G166L}* did not show functional complementation (Figure 4.20). On sucrose however, GTG^{G166L}-2 line showed a slightly longer hypocotyl length than the other At GTG1^{G166L} expressing lines, but this was not observed in the absence of sucrose (Figure 4.20). The At GTG1^{G166L} expressing lines did not restore normal hypocotyl cell growth in any of the independent experiments (Figure 4.20 A – D). These results suggest that the G166L had a significant effect on the ability of At GTG1 to functionally complement *gtg1-3 gtg2-3* mutant phenotypes and therefore G166 must be functionally important.

The At *GTG1* expression levels of the plant lines were analysed by qRT-PCR in order to determine whether these effects observed in these plant lines were due to

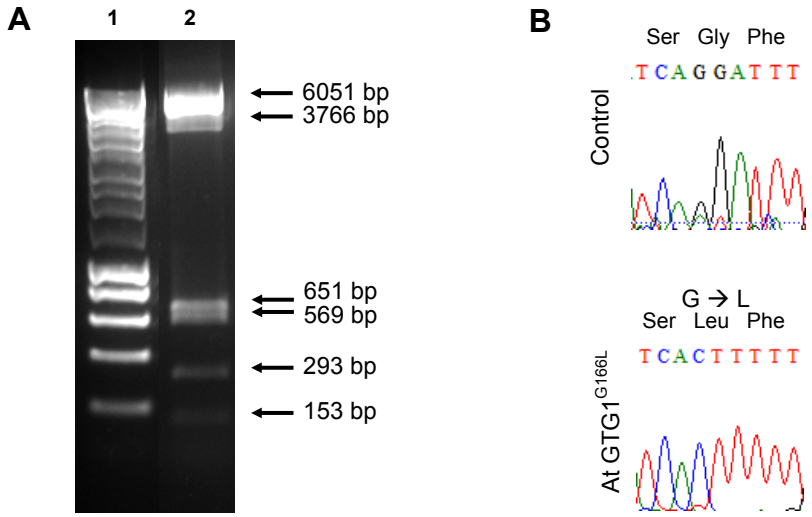


Figure 4.18 Diagnostic restriction analysis of pMDC32 At GTG1^{G166L} expression vector

A. Analysis of At GTG1^{G166L} cloned into pMDC32 destination vector using restriction enzymes *Nsi*I and *Xba*I (lane 2). The predicted product sizes are shown on figure. Lane 1: molecular markers (see Figure 4.5 A for predicted molecular marker sizes). **B**. Sequencing of pMDC32 At GTG1^{G166L} plasmid DNA reveals the presence of the mutation (GGA → CTT) in codon G166.

Table 4.6 Transgenic Arabidopsis lines created for expressing At *GTG1*

Plasmid	Construct	Transformed into	# of independent lines
pMDC32 At GTG1*	*P35S:At GTG1*	*gtg1-3 gtg2-3*	2
pMDC32 At GTG1^{G166L}	*P35S:At GTG1^{G166L}*	*gtg1-3 gtg2-3*	3

*previously created by Jaffé et al. (2012); # = number.

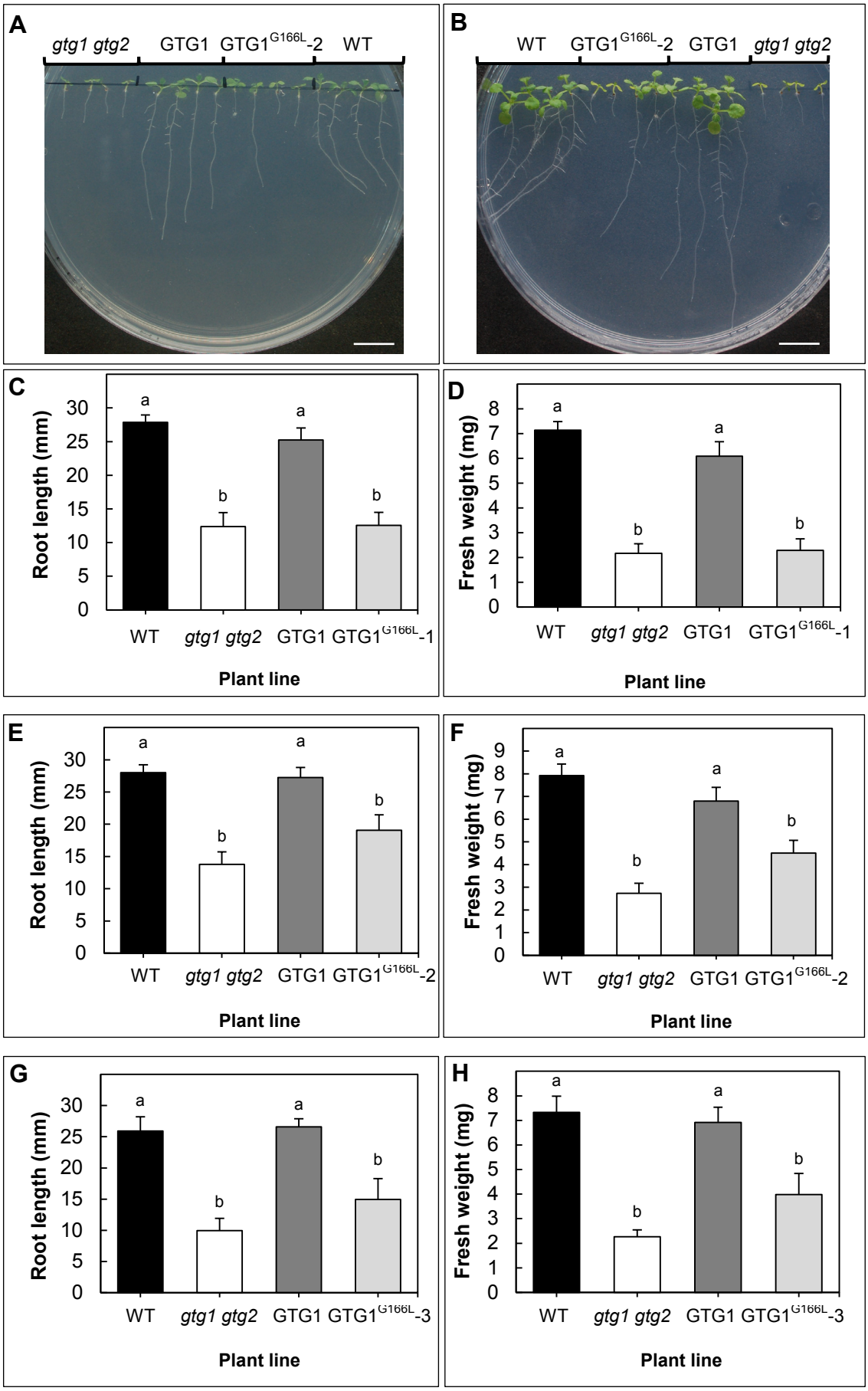

A
gtg1 gtg2
GTG1
GTG1G166L-2
WT
B
WT
GTG1G166L-2
GTG1
gtg1 gtg2
C
Root length (mm)
0
5
10
15
20
25
30
a
b
a
b
WT
gtg1 gtg2
GTG1
GTG1G166L-1
Plant line
D
Fresh weight (mg)
0
1
2
3
4
5
6
7
8
a
b
a
b
WT
gtg1 gtg2
GTG1
GTG1G166L-1
Plant line
E
Root length (mm)
0
5
10
15
20
25
30
a
b
a
b
WT
gtg1 gtg2
GTG1
GTG1G166L-2
Plant line
F
Fresh weight (mg)
0
1
2
3
4
5
6
7
8
9
a
b
a
b
WT
gtg1 gtg2
GTG1
GTG1G166L-2
Plant line
G
Root length (mm)
0
5
10
15
20
25
30
a
b
a
b
WT
gtg1 gtg2
GTG1
GTG1G166L-3
Plant line
H
Fresh weight (mg)
0
1
2
3
4
5
6
7
8
a
b
a
b
WT
gtg1 gtg2
GTG1
GTG1G166L-3
Plant line

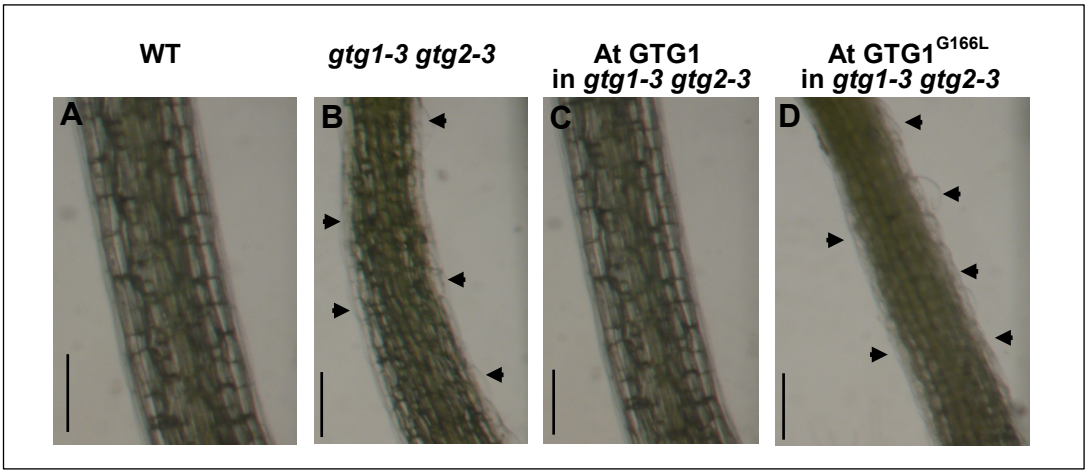
Figure 4.19 At GTG1^{G166L} does not functionally complement root growth or fresh weight defect of the Arabidopsis *gtg1-3 gtg2-3*

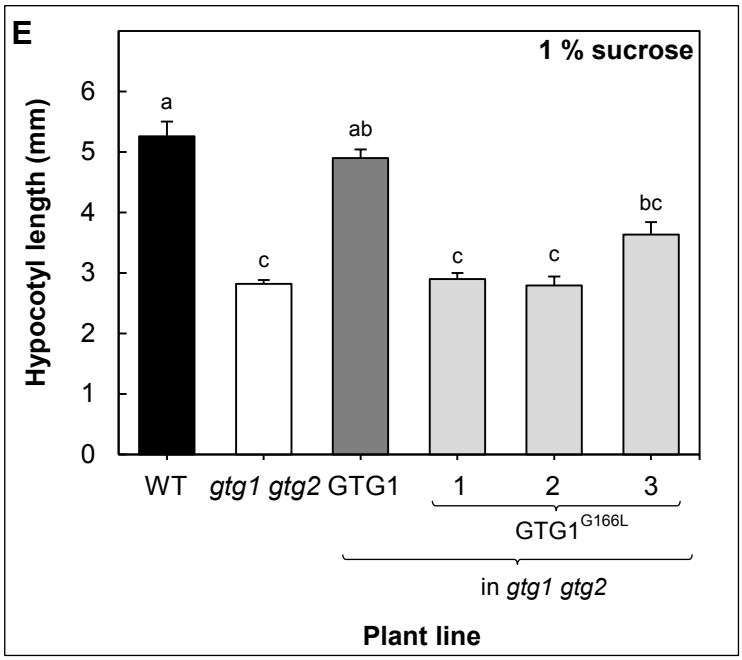
Root growth and fresh weight per seedling of *gtg1-3 gtg2-3* (*gtg1 gtg2*) 0.5 MS containing 0 % sucrose. Mean root length (**A**, **C**, **E** and **G**) and fresh weight per seedling (**B**, **D**, **F** and **H**) calculated for six plates (+S.E.) with four seedlings per plant line per plate, each plate containing four plant lines. Scale bar: 1 cm. Results represent one of three independent experiments using three independent *gtg1 gtg2* lines expressing *P35S:At GTG1^{G166L}* (GTG1^{G166L}-1 – -3) and two independent *gtg1 gtg2* lines expressing *P35S:At GTG1* (GTG1; only one line used). **A**. Image displaying root growth at 10 d (GTG1^{G166L}-1 and GTG1 = *P35S:At GTG1^{G166L}* and *P35S:At GTG1* in *gtg1 gtg2* lines, respectively). **B**. Image displaying root growth at 14 d (GTG1^{G166L}-1 and GTG1 = *P35S:At GTG1^{G166L}* and *P35S:At GTG1* in *gtg1 gtg2* lines, respectively). **C**, **E** and **G**. Mean root length at 10 d (+S.E.). There is a significant difference in root lengths between plant lines; one-way ANOVAs: **C**. ($F_{3, 20} = 13.53$, $p < 0.0001$), **E**. ($F_{3, 20} = 13.68$, $p < 0.0001$) and **G**. ($F_{3, 20} = 12.32$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* tests. **D** and **F**. Mean fresh weight per seedling at 14 d (+S.E.). There is a significant difference in fresh weight per seedling between plant lines; one-way ANOVAs: **D**. ($F_{3, 20} = 15.37$, $p < 0.0001$), **F**. ($F_{3, 20} = 19.28$, $p < 0.0001$) and **H**. ($F_{3, 20} = 17.16$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* tests.

differences in At *GTG1* expression levels. For example, the lack of phenotype rescue in *gtg1 gtg2* could be because of a silenced or low level of At *GTG1*^{G166L} expression at the RNA level or at the protein level, rather than due to the effect of G166L mutation on GTG activity. Therefore, the level of At *GTG1* with *ACTIN 2* reference gene expression was analysed. Since only one codon was altered in At GTG1^{G166L}, the primers AtGTG1.ex12 qF1 and AtGTG1.ex12 qR1 used to detect At *GTG1* transgene expression could be used for both At GTG1 and At GTG1^{G166L} expressing lines.

qRT-PCR results indicate that the expression of At *GTG1* is very low in the *gtg1-3 gtg2-3* mutant (Figure 4.21). The T-DNA insert for *gtg1-3* is in the first intron (Figure 4.4 A), therefore a low level of At *GTG1* expression could be detected as a result of low 3′ transcript levels produced since qRT-PCR primers were designed for the 3′ end and not across the T-DNA insert at the 5′ end. However, the transgenic Arabidopsis *gtg1-3 gtg2-3* transformed with At GTG1 (GTG1) or At GTG1^{G166L} (GTG1^{G166L}-1 – -3) have similar levels of expression, which is approximately two-fold higher than the WT (Figure 4.21). These results indicate that there are high levels of At *GTG1* and At *GTG1*^{G166L} expression under *P35S* in *gtg1-3 gtg2-3* (Figure 4.21).

At *GTG1* expression was able to rescue root, fresh weight and hypocotyl defects, while At *GTG1*^{G166L} did not exhibit this rescue (Figure 4.19 and Figure 4.20). G166L appears to cause a disruption in At GTG1 function as shown by the lack of functional complementation in root and hypocotyl assays. However, further testing would be required to indicate whether this is a complete loss-of-function. The qRT-PCR would also need to be repeated and with possibly a different reference gene such as *Yellow-Leaf-Specific gene8* (*YLS8*), since the primer efficiencies for *ACTIN2* and At *GTG1* were only 1.2521 and 1.4283 respectively for these experiments (Figure 4.21). It is also important to show At GTG1 protein expression levels.


WT
gtg1-3 gtg2-3
At GTG1 in gtg1-3 gtg2-3
At GTG1G166L in gtg1-3 gtg2-3
A
B
C
D


E
1 % sucrose
Hypocotyl length (mm)
0
1
2
3
4
5
6
a
c
ab
c
c
bc
WT
gtg1 gtg2
GTG1
1
2
3
GTG1G166L
in gtg1 gtg2
Plant line

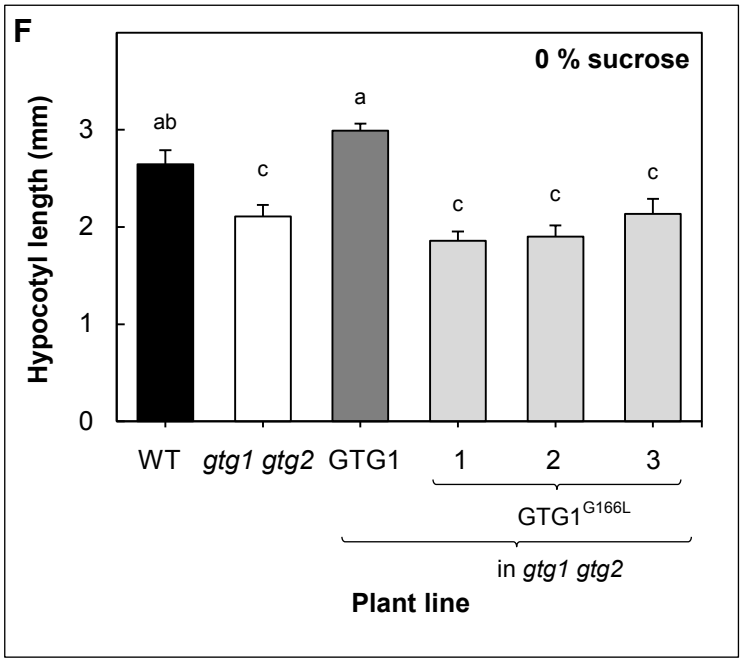

F
0 % sucrose
Hypocotyl length (mm)
0
1
2
3
ab
c
a
c
c
c
WT
gtg1 gtg2
GTG1
1
2
3
GTG1G166L
in gtg1 gtg2
Plant line

Figure 4.20 At GTG1^{G166L} does not functionally complement the hypocotyl growth defects seen in the Arabidopsis *gtg1-3 gtg2-3* mutant

Hypocotyl growth of *gtg1-3 gtg2-3* (*gtg1 gtg2*) mutants on 0.5 MS containing 1 % (**A – E**) and 0 % (**F**) sucrose at 5 d. Short hypocotyl defect of *gtg1 gtg2* is exacerbated in 1 % sucrose (**E**) compared to 0 % sucrose (**F**). Data shows the mean hypocotyl length per seedling calculated for six plates (+S.E.) with 15 seedlings per plant line per plate, each plate containing six plant lines. **A – D**. Distended cell growth in *gtg1 gtg2* is rescued by At GTG1 (**C**) but not At GTG1^{G166L} (**D**) expression. Light microscopy showing hypocotyls of WT (**A**), *gtg1 gtg2* (**B**), At GTG1 in *gtg1 gtg2* (GTG1) (**C**), and At GTG1^{G166L} in *gtg1 gtg2* (GTG1^{G166L}-1 line shown) (**F**). Scale bars: 200 μm, arrowheads indicate ballooning cells. **E – F**. The mean hypocotyl length per seedling (+S.E.) in 1 % sucrose (**E**) and 0 % sucrose (**F**). There is a significant difference in hypocotyl length per seedling between plant lines as determined by one-way ANOVAs: (**E**) 1 % sucrose ($F_{5, 30} = 3.15$, $p = 0.021$) and (**F**) 0 % sucrose ($F_{5, 30} = 14.02$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* tests. Results represent one of three (**E**) and one (**F**) independent experiment(s) using three independent *gtg1 gtg2* lines expressing *P35S:At GTG1^{G166L}* and two independent *P35S:At GTG1* expressing lines.

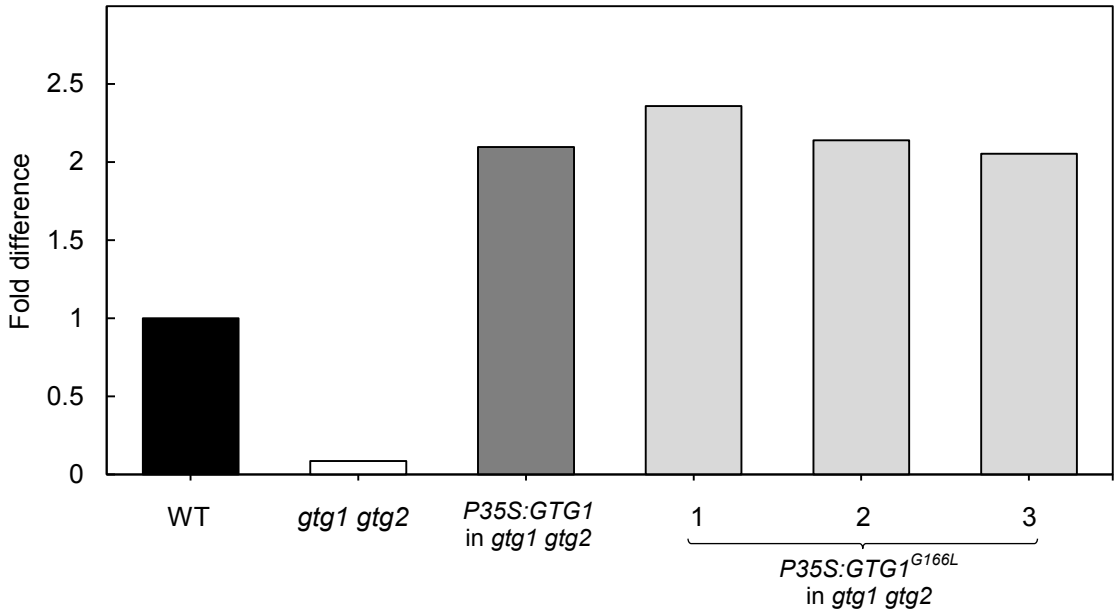


Figure 4.21 Transgenic Arabidopsis *gtg1-3 gtg2-3* expressing At *GTG1* transcript

qRT-PCR showing the fold differences of At GTG1 RNA transcript expression. *gtg1-3 gtg2-3* (*gtg1 gtg2*) mutant display negligible levels of At *GTG1* expression, while the At *GTG1* and three At *GTG1*^{G166L} (GTG1^{G166L} -1 – -3) expressing lines show an approximate two-fold difference to WT. Results represent one reaction using *ACTIN2* as the reference gene. Primers AtGTG1.ex12 qF1 and AtGTG1.ex12 qR1 used to amplify At *GTG1*, and primers Actin2 F and Actin2 R used to amplify *ACTIN2*.

4.2.9 Isolating an *O. sativa gtg* mutant

The Arabidopsis *gtg1 gtg2* mutants exhibit defective root, hypocotyl, fresh weight and fertility defects (Jaffé et al., 2012), suggesting that the At GTGs have an important function in plant growth and development. There are some differences in plant biological processes between dicots and monots. To study the impact of GTGs in a monocot plant species, investigation of the GTGs in *O. sativa* was initiated. *O. sativa* contains one copy of the *GTG/GPHR* gene, Os *GTG* (Accession: LOC_Os04g51180). The genomic structure of Os *GTG* is shown in Figure 4.22 (displaying the Os04g51180.1 gene model, one of the two predicted splice variants Os04g51180.1 and Os04g51180.2 in *O. sativa*). Similar to Arabidopsis, T-DNA mutants in *O. sativa* can be used for studying gene functionality. *O. sativa* T-DNA mutants can be identified from the *Oryza* Tag Line (OTL) database (http://oryzatagline.cirad.fr/). An Os *GTG* mutant in the Nipponbare background was identified in the OTL, and the seed was obtained (ALH H03 mutant allele symbol; FST accession SAM4G08). T_2 segregating seed for the T-DNA insertion into Os *GTG* were grown on soil for T_3 seed collection.

Genomic DNA was extracted from the T_2 plants for genotyping and identifying a homozygote for the insertion. Primers were designed to anneal to the T-DNA plasmid inserted into Os *GTG*, p4978, as well as to WT Os *GTG* (Figure 4.22). The presence of a predicted 960 bp product (lanes 1, 5, 8 and 11) using primers OsGTG 1F (1F) and OsGTG 1R (1R) confirms the presence of the WT allele for Os *GTG* (Figure 4.23). The presence of a predicted 844 bp product (lane 6 and 9) using primers Hyg7 and 1R, and a 870 bp product using primers Hyg8 and 1R (lanes 7 and 10) confirms the presence of the T-DNA insert. Heterozygous plants for the insert will show amplified products for each of the three primer sets (lanes 5 – 10), while WT plants will only show a product for Os *GTG* (lanes 2 – 4 and 11 – 13). Homozygous plants for the insert will show no WT product and two products for the T-DNA insert. There was some difficulty in isolating a homozygous Os *gtg* mutant plant as many (12 out of 21) of the segregating T_2 seed did not germinate. Results confirmed three WT and four heterozygous T_2 plants out of the seven tested but no homozygous plants for the insertion. Of the four heterozygous plants identified, two died so only two heterozygous T_3 plants were obtained from the T_2 segregation population. To identify the exact location of the T-DNA insert site, the DNA from lane 6 and 7 of Figure 4.23 was purified for sequencing

and analysis of the T-DNA insert. The insert was found to be in the second intron, after nucleotide 574 from the start codon (Figure 4.22).

With the T-DNA insert site confirmed, the T_3 seed from a heterozygous T_2 plant (candidate b) were grown on soil (Figure 4.25) to try to identify a homozygote in this generation. However, only two seeds germinated and matured on soil from the heterozygous parent. Genomic DNA was extracted from these T_3 plants; PCR and gel electrophoresis results indicate that the two candidates, b-1 and b-2, are homozygous and heterozygous for the insertion, respectively (Figure 4.24 A). To test whether the T-DNA insert generates a null mutant, RNA was extracted and cDNA was synthesised to confirm by PCR the RNA transcript levels of Os *GTG* in the candidates b-1 and b-2. The results confirmed that candidate b-1 is a null Os *gtg* mutant (homozygous for the T-DNA insert) which does not contain Os *GTG* transcript, as shown by the absence of the full-length Os *GTG* (Figure 4.24 B; lane 3). Candidate b-2 contained the predicted 1408 bp product for the Os *GTG* transcript (Figure 4.24 B; lane 4), as well as the predicted products for the T-DNA inserts at the genomic level (Figure 4.24 A; lanes 9 and 10). These results confirm that candidate b-2 is heterozygous for the T-DNA insert.

Images displaying the growth after 3 months indicate that candidates b-1 and b-2 are smaller, have less leaves and there appears to be less seed compared to the WT. These observed phenotypes are exacerbated in the homozygous (candidate b-1) compared to heterozygous (candidate b-2) plant (Figure 4.25). At 4 months growth, candidate b-1 (homozygous) is smaller in height than WT, but with substantially more leaves than at 3 months. Candidate b-2 (heterozygous) however, appears more like WT with many leaves and a similar height (Figure 4.25). Therefore, preliminary observations indicate that the homozygous Os *gtg* mutant is a smaller plant than WT.

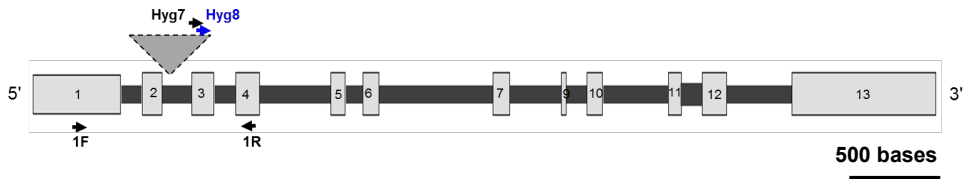


Figure 4.22 Genomic structure of Os *GTG* with T-DNA insert

The Os *GTG* genomic structure displaying the exons and introns of the gene and T-DNA insertion site drawn to scale (scale bar: 500 bases). Exons are shown in light shade grey and numbered (exon 8 and intron 9 are too short to label); introns are shown in dark shade grey; Grey triangle = T-DNA insert (not drawn to scale) displays the insert site after nucleotide 574 from the start codon. Symbols below the structures indicate primers used for genotyping; arrows show primer direction (primer sequences outlined in Table 2.10). OsGTG1 1F (1F) and OsGTG1 1R (1R) anneal to exons 1 and 4 respectively. The T-DNA primers, Hyg7 and Hyg8, shown in black and blue respectively. Figure shows the Os04g51180.1 gene model, one of two splice variants in *O. sativa* (shown in Figure 3.1).

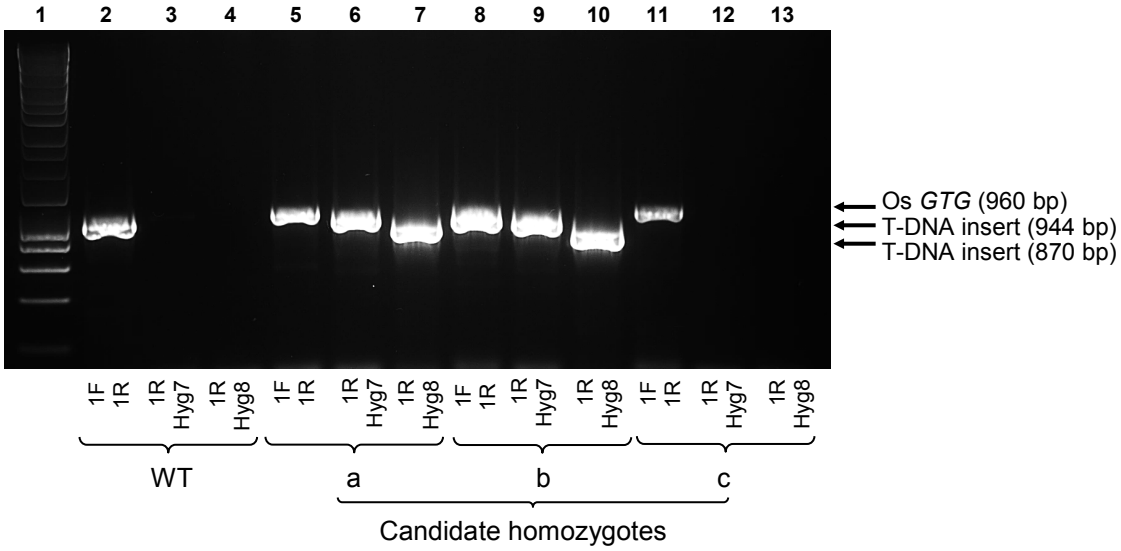


Figure 4.23 Genotyping T_2 *O. sativa* to identify a homozygous Os *gtg* mutant

PCR and gel electrophoresis using genomic DNA showing the genotypes of candidate homozygous Os *gtg* mutants. Predicted product sizes and lanes are shown on figure (see Figure 4.5 A for predicted molecular marker sizes and Table 2.10 for primer sequences). Three primer sets were used to identify WT Os *GTG* and the T-DNA insert using three candidate plants shown (a – c). Primers 1F and 1R (Figure 4.22) amplify the Os *GTG* gene (lanes 2, 5 and 8). Primers Hyg7 and 1R (Figure 4.22) amplify a fragment of the T-DNA insert in Os *GTG*, displaying presence (lanes 6 and 9) or absence (lane 12) of an insert. Primers Hyg8 and 1R (Figure 4.22) amplify a fragment of the T-DNA insert in Os *GTG*, displaying presence (lanes 7 and 10) or absence (lane 13) of an insertion in Os *GTG*.

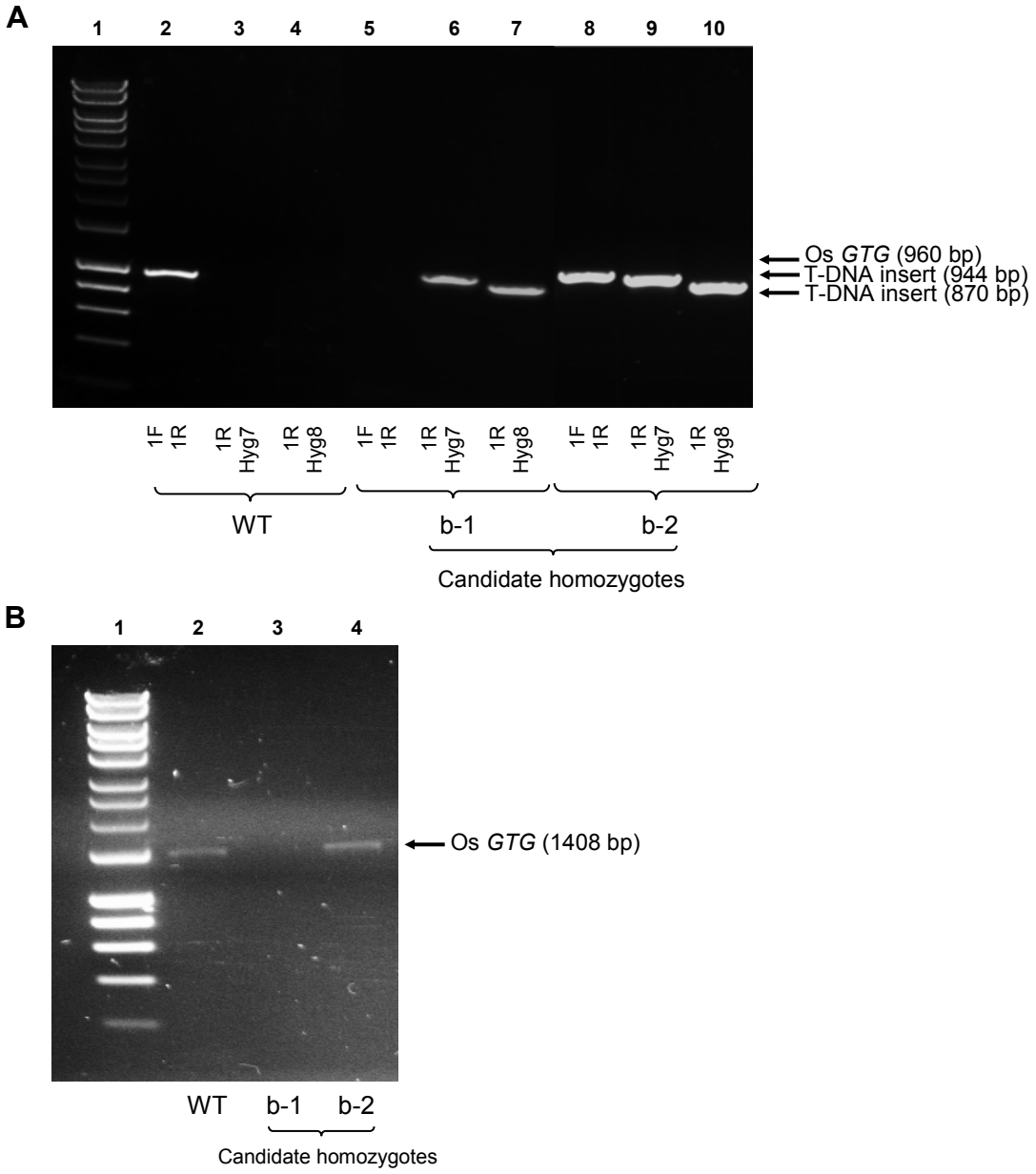


Figure 4.24 Genotyping T_3 *O. sativa* to identify a homozygous Os *gtg* mutant

PCR and gel electrophoresis using genomic DNA (**A**) and cDNA synthesised from RNA (**B**) showing the genotypes of candidate homozygous Os *gtg* mutants. Predicted product sizes and lanes are shown on figure (see Figure 4.5 A for predicted molecular marker sizes and Table 2.10 for primer sequences). **A**. Using genomic DNA and three primer sets (Figure 4.22) to identify WT Os *GTG* and the T-DNA insert of two candidate plants (b-1 – b-2). Primers 1F and 1R amplify fragment of the Os *GTG* gene (lane 2). Primers Hyg7 and 1R amplify a fragment of the T-DNA, displaying presence (lanes 6 and 9) of an insert in Os *GTG*. Primers Hyg8 and 1R amplify a fragment of the T-DNA, displaying presence (lanes 7 and 10) of an insertion in Os *GTG*. **B**. Using cDNA synthesised from RNA and primers OsGTG TOPO 1F and OsGTG TOPO NO STOP 1R (primer sequences outlined in Table 2.10) to identify Os *GTG* using two candidate plants shown (b-1 – b-2). Product of 1408 bp indicates presence of full-length Os *GTG* CDS without stop codon (lane 2).

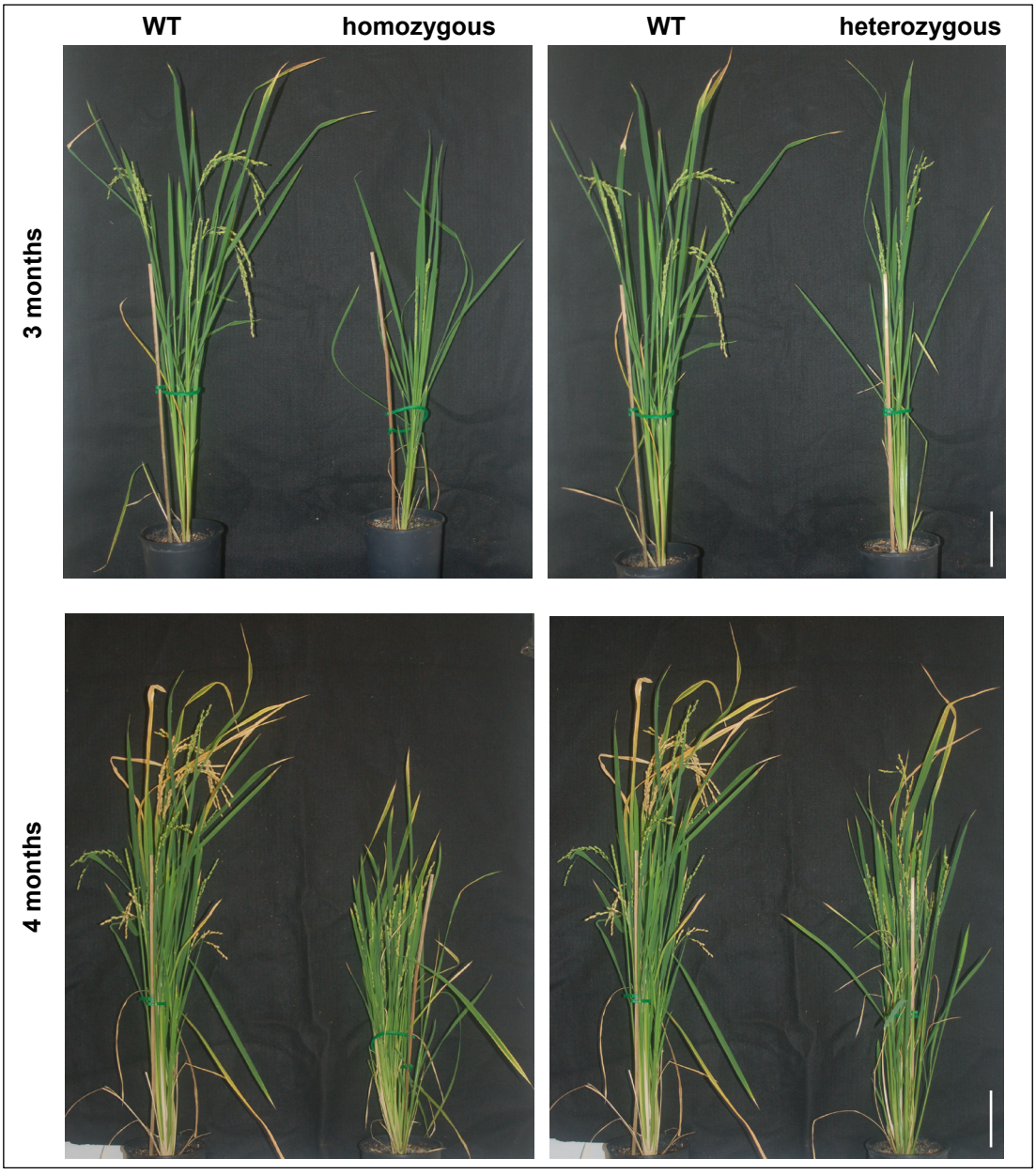


Figure 4.25 Candidate T_3 homozygous Os *gtg* mutants

Images showing T_3 candidate homozygous Os *gtg* mutants grown on soil at 3 and 4 months. Images display homozygous and heterozygous plants grown from seed originating from a T_2 heterozygous plant for the Os *gtg* mutant allele (ALH H03). Scale bar: 10 cm.

4.3 Discussion

4.3.1 All *gtg1 gtg2* double mutants show similar phenotypes

The *gtg1 gtg2* mutants isolated by Pandey et al. (2009) and Jaffé et al. (2012) were compared to investigate whether different T-DNA double mutants exhibit similar phenotypes. The genotyping by PCR of the *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* shows that these mutants contain T-DNA inserts in different sites (Figure 4.4 and Figure 4.5). These were confirmed as null mutants, since full-length At *GTG* transcripts were not expressed (Figure 4.6). The *gtg1-1 gtg2-1* mutants were previously reported to be hyposensitive to ABA (Pandey et al., 2009), but this sensitivity was not observed in different *gtg1 gtg2* mutants (*gtg1-2 gtg2-3* and *gtg1-3 gtg2-3*) (Jaffé et al., 2012). Although ABA hyposensitivity was not observed in either *gtg1-2 gtg2-3* or *gtg1-3 gtg2-3*, other interesting phenotypes were observed. Therefore it was important to ascertain whether *gtg1-1 gtg2-1* also showed this range of phenotypic characteristics.

The reduced root, fresh weight, hypocotyl growth, seed yield, silique length, and seed number per silique as well as cellular defects exhibited by the *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants (Jaffé et al., 2012), were also observed in the *gtg1-1 gtg2-1* mutant (Figure 4.7 – Figure 4.11). A novel defect in rosette growth was also observed in all the *gtg1 gtg2* mutants, which has not been previously reported. In *gtg1-1 gtg2-1* and *gtg1-3 gtg2-3* the rosette diameter measured over 155 d (24 weeks) was significantly smaller than their respective WTs in short-day conditions (Figure 4.9). However, no apparent difference in rosette growth between the *gtg1 gtg2* mutants and WT have been observed in previous experiments conducted in the laboratory in long-day conditions (M. J. Terry and L. E. Williams, unpublished results). This is an interesting observation as it coincides with the Genevestigator results indicating that At *GTG* expression is reduced in long-day conditions compared to short-day (Figure 3.19). Since At *GTG* genes are upregulated in short-day conditions, this could be the reason for a rosette growth defect manifesting in *gtg1 gtg2* mutants in short-day conditions when At *GTG* expression is higher than in long-day conditions. The *gtg1-1 gtg2-1* mutant was reported to exhibit ABA hyposensitivity (Pandey et al., 2009); however, the *gtg1-1 gtg2-1* mutant displayed no hyposensitivity to ABA in assays conducted in our lab (K. Peaston, M. J. Terry and L. E. Williams, unpublished results). In all, these data support that the

different *gtg1 gtg2* mutants exhibit similar defects and this includes not being ABA-insensitive.

In Arabidopsis, heterotrimeric G proteins GPA1 and AGB1 are the canonical Gα and Gβ subunits, while AGG1 and AGG2 are the Gγ subunits, respectively. Interestingly, *gpa1-4* and *agb1-2* mutants produce fewer seeds per silique, while AGG1- and AGG2-deficient mutants showed WT levels of seeds (Chakravorty et al., 2011). This is an interesting finding since the GTG/GPHRs have been reported to be GPCR-type G proteins which have GTPase activity, characteristic of G proteins. However, seed yield is not solely associated with GPCRs or G protein genes.

BRs stimulate growth through cell elongation (Azpiroz et al., 1998) and cell division (Cheon et al., 2010). All *gtg1 gtg2* mutants display a similar response to BL (Figure 4.12 and Figure 4.13). WT and *gtg1 gtg2* mutants exhibit increased hypocotyl growth with increasing concentrations of BL (Figure 4.12). Interestingly, all *gtg1 gtg1* mutants exhibited a stronger response to BL treatment compared to WT, as shown by a higher percentage stimulation of hypocotyl growth (Figure 4.13 J). This indicates that these mutants can sense and respond to BRs; unlike BR-insensitive mutants such as *bri1*, which do not exhibit growth stimulation in the presence of increasing levels of BR (Clouse et al., 1996). In 250 nM BL there was no difference between WT and *gtg1 gtg2* hypocotyl growth (Figure 4.13 I). For both WT and *gtg1 gtg2* plant lines the hypocotyl length was ~6 mm in 250 nM BL, therefore suggesting the hypocotyl growth may be at its optimum. *gtg1 gtg2* mutants have large, distended cells at their hypocotyl surface (Jaffé et al., 2012). In 250 nM BL, the hypocotyl cells showed some variation between *gtg1 gtg2* hypocotyls; there were some comparable to WT while others were deformed. Therefore this rescue defect was inconsistent suggesting that *gtg1 gtg2* mutants could exhibit some BR signalling impairement but further testing would be required. These data suggest that *gtg1 gtg2* mutants are hypersensitive to BL as a higher relative percentage stimulation of hypocotyl growth is observed. BL-hypersensitivity could result from BR-deficiency. This could be caused by Golgi/ER impairment since many BR synthesis enzymes are located within the secretory pathway such as cytochrome P450 monooxygenase in the ER (Werck-Reichhart and Feyereisen, 2000). Therefore a defect in the secretory pathway due to for example the removal of GTG/GPHRs, could result in a dysfunctional secretory pathway. Alvarez et al. (2013) and Jaffé et al. (2012) observed that there is no effect of hypocotyl growth reduction in the dark.

Genevestigator results indicate that At *GTG* genes are upregulated when BL is applied, as well as downregulated in *det3* mutants (BR synthesis-defective; Table 4.2) compared to WT (Figure 3.19). These expression data suggest that At GTGs are not directly involved in the BR synthesis pathway but their expression is affected by changing BR levels.

These phytohormone results differ to those reported by Alvarez et al. (2013), although the experiments were conducted in different ways. The *gtg1-1 gtg2-1* exhibited hyposensitivity to BL as shown by reduced root and hypocotyl growth inhibition compared to the WT (Alvarez et al., 2013). However, in this project, BL increased the hypocotyl length in both WT and *gtg1 gtg2*, and stimulation of hypocotyl growth was more pronounced in *gtg1 gtg2* than the WT (Figure 4.13). However, it is important to note the differences in the way these experiments were carried out. Alvarez et al. (2013) used 10 nM BL and light treated seeds for 1 d followed by growth in the dark for 5 d. Whereas in this report, a range of BL concentrations between 0 – 1000 nM were used; the light treatment of seeds lasted only 2 h prior to being transferred to low light (10 $\mu mol.m^{-2}.sec^{-1}$) conditions for 5 d. The difference in BL hyposensitivity could result from the difference in light conditions used. When grown in the dark, the *gtg1 gtg2* mutants do not display significant differences to WT, indicating that *gtg1 gtg2* are required for light-dependent responses (Jaffé et al., 2012). In addition, BR signalling is altered in accordance to light conditions (Li et al., 1996). For these reasons, further BR-sensitivity studies would need to be carried out. For example, different *gtg1 gtg2* mutants could be grown in the dark in the presence of 10 nM BL to see whether a similar effect to Alvarez et al. (2013) is observed.

4.3.2 At *GTG1:GFP* and At *GTG2:GFP* functionally complements *gtg1-3 gtg2-3* mutants, and At GTG2 localises to intracellular structures

The *gtg1-3 gtg2-3* mutants exhibit defects in root, hypocotyl and cell development, which was rescued by the constitutive expression of either At *GTGs* (Jaffé et al., 2012). The genomic sequences of both At *GTG* genes were cloned into pMDC99 and pMDC107 plant destination vectors; the latter contains a C-terminal GFP tag (Jaffé et al., 2012). Several independent *gtg1-3 gtg2-3* lines expressing either At *GTG1:GFP* or At *GTG2:GFP* under their native promoters, were isolated for phenotypic analysis (Table 4.4). The expression of *PAt GTG1:At GTG1:GFP* or *PAt GTG2:At GTG2:GFP*

in *gtg1-3 gtg2-3* functionally complemented the root growth, fresh weight and hypocotyl defects (Figure 4.14 – Figure 4.16). This suggests that under their native promoters, either At GTGs functionally complement defects of double At *GTG* KO mutants. GFP is a 27 kDa protein made up of 237 amino acids (Inouye and Tsuji, 1994), and therefore its fusion may affect functionality of proteins. In *gtg1-3 gtg2-3* lines expressing *PAt GTG1:At GTG1:GFP* or *PAt GTG2:At GTG2:GFP*, the functionality of the At GTGs was not affected by the C-terminal GFP fusion, as shown by the ability to rescue root, hypocotyl, fresh weight and cellular defects of the *gtg1-3 gtg2-3* mutant (Figure 4.14 – Figure 4.16). In addition, At GTG1 was previously shown to be Golgi localised when *PAt GTG1:At GTG1:GFP* was stably expressed in Arabidopsis (Jaffé et al., 2012). The functional complementation shown here with this construct demonstrates that full functionality can be restored.

PAt GTG2:At GTG2:GFP expression in the *gtg1-3 gtg2-3* mutant also functionally complemented the root, fresh weight, hypocotyl and cellular defects, and so preliminary localisation studies of At GTG2 were performed, which indicated punctate fluorescent foci (Figure 4.17). Since *PAt GTG1:At GTG1:GFP* was shown to be Golgi and ER localised and either At GTG1 or At GTG2 are able to functionally complement *gtg1 gtg2* mutant defects, it can be hypothesised that *PAt GTG2:At GTG2:GFP* and *PAt GTG1:At GTG1:GFP* localise to the same place i.e. Golgi and ER network. However, further localisation studies are needed to finally conclude whether At GTG2 also localises to the Golgi and ER network, as seen for At GTG1 (Jaffé et al., 2012).

Transgenic T_3 Arabidopsis expressing *PAt GTG1:At GTG1* or *PAt GTG2:At GTG2* i.e. without GFP tags, are still to be isolated (Table 4.5). These lines are important for investigating the reduced silique length phenotype observed by the *gtg1 gtg2* mutants (Jaffé et al., 2012), which is not rescued by At GTGs under the 35S promoter. Although *P35S* drives ectopic expression (Benfey et al., 1990), there is also a report indicating that expression is weak in siliques (Ge et al., 2008). If this were the case, it can be concluded that *P35S* does not drive expression to rescue the short silique phenotype of the *gtg1 gtg2* mutants. This experiment has not been carried out yet because only *gtg1-3 gtg2-3* lines expressing At GTG tagged with GFP have been isolated (*PAt GTG1:At GTG1:GFP* and *PAt GTG2:At GTG2:GFP*, refer to Table 4.3). Ideally an experiment would compare *gtg1-3 gtg2-3* lines expressing At GTGs with and without GFP fusions (*PAt GTG1:At GTG1* and *PAt GTG2:At GTG2* refer to Table 4.3)

as GFP could interfere with the function of the At GTGs in silique development. Therefore, results from the non-tagged lines would not have GFP as a factor that could affect function. Furthermore the *PAt GTG1:At GTG1:GFP* and *PAt GTG2:At GTG2:GFP* constructs can be used to follow the At *GTG* expression pattern during silique development.

The GTG/GPHRs are found in a number of plant species but so far only studied in Arabidopsis. To initiate research of these proteins in a monocot, the *O. sativa* crop model species was selected for this project as T-DNA mutants have become available. An Os *gtg* T-DNA mutant was identified in the OTL database, and a homozygous mutant was obtained (Figure 4.23). The sequencing data indicated that the T-DNA insertion is in the second intron (Figure 4.22). The isolation of this Os *gtg* mutant can be used in various phenotypic assays to further our insight into the function of the GTG/GPHRs in monocots. Preliminary observations indicate that Os *gtg* could be smaller than the WT and produce less seed (Figure 4.25), however additional testing would be required to confirm this. *O. sativa* genes related to seed yield is a particularly important field of study since rice is a commonly consumed crop of high demand. Since the At GTGs have exhibited defects in seed size, number and yield (Jaffé et al., 2012), determining if GTG/GPHRs have a role in regulating grain size or yield in the plant could provide a platform for potential genetic manipulation for agricultural purposes.

It was difficult to isolate a homozygous Os *gtg* (-/-) mutant as not many of the segregating T_2 Os *gtg* (+/-) seeds germinated. The segregation ratios were distorted during the isolation of the Arabidopsis *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* double mutants (Jaffé et al., 2012). However, it is not possible to estimate if the plants segregate normally as out of 21 plants only seven were genotyped (three WT, four heterozygous and no homozygous), while the others either did not germinate (12) or died (two). With a T_3 homozygous Os *gtg* mutant confirmed (Figure 4.24), phenotypic analyses may be carried out to characterise the Os *gtg.* These could include plant size measurements (root and shoot parameters) and seed characterisation such as seed yield and weight.

4.3.3 G166 in At GTG1 plays an important role in protein function

Mutations in proteins causing loss-of-function have been used to identify key residues important for function. For example, in voltage-gated K^+ channels the selectivity pore has several key residues important for discriminating between K^+ and

sodium (Na^+) ions. This was shown by mutagenesis studies in the signature sequence of the *Shaker* K^+ channel. In this protein, the conserved glycine residues in the pore helix are essential for ion selectivity and fast conductance through the pore (Heginbotham et al., 1994, Doyle et al., 1998).

The GTG/GPHRs have a conserved DUF3735 domain, in which a conserved LSG motif exists (section 3.2.3). Highly conserved glycine residues are often found in transport proteins, for example the *E. coli* phenylalanine and tyrosine transporter (PheP) (Dogovski et al., 2003). The structure of glycine allows it to reside in regions of proteins that other amino acids cannot, for example in tight turns (Betts and Russell, 2003). Therefore glycine possesses attributes that are absolutely indispensable when relating structure to function. For this reason, even conservative substitutions of glycine can be detrimental to function in certain proteins. For example, substituting conserved glycine residues with increasing side chain volume in PheP resulted in progressive loss of transport activity (Dogovski et al., 2003). Glycine residues are often found in conserved motifs of membrane-spanning helices of membrane proteins, suggesting a structural role (Javadpour et al., 1999). Glycine in hydrophobic membrane sequences permits strong hydrogen-binding interactions (Li and Deber, 1992), which is important for specific helix packing interactions (Lemmon and Engelman, 1994). Substitutions of glycine residues by more hydrophobic residues can result in loss of dimerisation capacity (Lemmon et al., 1992, Lemmon et al., 1994). GxxxG and GxxxxxxG are conserved motifs for transmembrane helix-helix associations (Russ and Engelman, 2000, Liu et al., 2002). For instance in PheP, GxxxG mediates high-affinity homo-oligomerisation for close helix packing conferring transport activity (Dogovski et al., 2003). Furthermore, glycine residues are able to bind phosphates (Schulze-Gahmen et al., 1996), which is useful for enzmyes such as protein kinases (Hanks et al., 1988). Therefore conserved glycine residues have a variety of structural roles associated with protein function.

The conserved glycine residue of DUF3735 in At GTG1 is found at residue-166 (G166). This residue is predicted to reside in the TM5 by several topology prediction programs (Figure 3.13) and could play a significant role in At GTG1 structure and function. Taking into account the mutagenesis studies described above and the predicted functional effects of different amino acid substitutions (Figure 3.14), G166 was mutated to leucine. Leucine, like glycine, is a hydrophobic amino acid but with an aliphatic isobutyl side chain. Leucine was chosen to substitute glycine since it is not only a

hydrophobic amino acid but also has been suggested to minimise disruption to α-helix formation (Lyu et al., 1991). Since At GTG1^{G166} is predicted to reside in a transmembrane domain, a leucine whose akyl side chains have the potential to stabilise α-helices would be a suitable substitution for observing the effects of a conservative substitution whilst with a larger side chain (Lyu et al., 1991, Betts and Russell, 2003).

Using three independent lines (Table 4.6) it was shown that the single point-mutation At GTG1^{G166L} suppressed the ability of At GTG1 to functionally complement root, fresh weight, hypocotyl and cellular defects of *gtg1-3 gtg2-3* (Figure 4.19 and Figure 4.20). Preliminary qRT-PCR results showed that all transgenic Arabidopsis *gtg1-3 gtg2-3* lines expressed either At *GTG1* or At *GTG1*^{G166L} (Figure 4.21), suggesting that the lack of rescue was due to the G166L rather than a lack of At *GTG1* expression. This would also need to be confirmed at the protein level before any firm conclusions can be made. Together these results suggest that At GTG1^{G166} is a key residue for protein function, but further work is required to determine the exact role for this residue. As mentioned, glycine may be involved in ion selectivity, for example, glycine residues within the selectivity pore of the K^+ channels. These play important roles in specific helix packing and amino acid orientations, providing exposed carbonyl groups facing the pore for not only K^+ selectivity but also fast ion conductance (Heginbotham et al., 1994, Doyle et al., 1998). The importance of the glycine residues was shown by mutagenesis studies where no conservative mutations (Ala, Cys, Pro, Gln, or Ser) in the K^+ selectivity filter gave rise to a K^+ selective pore (Heginbotham et al., 1994). In addition, glycine is very flexible and can adopt an extensive range of angles at specific points in protein structures. This property gives glycine the ability for hinge involvement of channels, for example G83 in MthK and G99 in KcsA, through a mechanism which causes a kink, redirecting the inner helix such that the pore broadens at the cytoplasmic face (Irizarry et al., 2002). Therefore, glycine has been shown to be functionally important in channel proteins.

As described above, replacing residues with another of a similar biochemical nature may or may not have an effect. Therefore, further mutagenesis studies could be conducted. For example, G166A and G166Q would be conservative mutations whereby G166A would be more conservative (due to its smaller side chain) than either G166L or G166Q (Betts and Russell, 2003). In addition to the conserved glycine in DUF3735, the GxxxxxxLSG motif identified in section 3.2.3 is a well-conserved motif in DUF3735

found in 178 of 179 GTG/GPHR sequences analysed (Table 3.3). In the ABA-GPCR domain there is a conserved tyrosine, which could be functionally important (Table 3.3). Interestingly, GPCRs also contain a highly conserved tyrosine needed for agonist-mediated sequestering of βAR (Barak et al., 1994). This is interesting since the GTG/GPHRs have been described as GPCR-type G proteins. However, tyrosine residues are not always indispensable for function. For example the conserved tyrosine of PheP was replaced using several residues which caused no apparent loss in activity (Dogovski et al., 2003). Lastly, a proline in the predicted TM3 was also found to be conserved (Table 3.3), and has recently been shown to be functionally important. When this proline was mutated to a leucine, GPHR function in *Drosophila* was abolished resulting in systemic growth defects (Charroux and Royet, 2014). This was shown using the denoted $GPHR^{k34}$ mutant generated by Charroux and Royet (2014) (renamed $GPHR^{P91L}$; Table 2.1), which exhibited similar loss-of-function phenotypes to the other isolated *GPHR* mutant, $GPHR^{LL03674}$ suggesting that these growth defects were characteristic of a *GPHR* null mutant. To investigate whether these conserved residues have an important role in GTG/GPHR function, further mutagenesis studies would be required. These studies would serve as a platform to identify important residues for GTG/GPHR function, especially when the structure of these proteins has yet to be characterised.

CHAPTER 5

Expressing the *C. elegans gphr* genes in Arabidopsis

5.1 Introduction

5.1.1 The conserved GTG/GPHR family of membrane proteins

The GTG/GPHR family of membrane proteins are conserved in eukaryotes (section 3.3.1). The two *C. elegans gphr* genes are 69 % identical and 83 % similar at the amino acid level while Arabidopsis and *C. elegans* GTG/GPHRs share 39 % identity to one another (Table 3.2). There are various regions of conservation throughout the length of the protein sequences (Figure 3.6), but is their function conserved? For instance, the CLC family of membrane proteins are evolutionary conserved and function as voltage-gated Cl^- channels from prokaryotes to eukaryotes (Jentsch et al., 1999, Estévez and Jentsch, 2002). They have been widely studied in prokaryotes such as *E. coli* (Iyer et al., 2002), as well in animals such as *C. elegans* (Schriever et al., 1999), fungi such as *S. cerevisiae* (Gaxiola et al., 1998) and plants such as Arabidopsis (Hechenberger et al., 1996, Geelen et al., 2000, Guo et al., 2014).

The GTG/GPHRs are highly conserved at the amino acid level but so far, there has not been one defined, conserved role for this family of membrane proteins. In fact, there are differing roles proposed for the GTG/GPHRs (section 1.2). The localisation of these proteins also appears to differ (Table 5.1) (Maeda et al., 2008, Pandey et al., 2009, Jaffé et al., 2012, Charroux and Royet, 2014). For this reason, it is important to verify their function and investigate whether this role is conserved for the GTG/GPHRs.

5.1.2 The role of GTGs in Arabidopsis

Functional analysis of the At GTGs has involved isolating and characterising single and double mutants. The phenotypic analyses carried out thus far would therefore serve as a good basis for investigating conservation of function. When both At *GTG1* and At *GTG2* genes are knocked out in double mutants plant lines (Col: *gtg1-2 gtg2-2* and Ws: *gtg1-1 gtg2-1 gtg1-3 gtg2-3*), a number of characteristic phenotypes manifest. These phenotypes include inhibition of root and hypocotyl growth, reduction in silique length and seed generation, deformed cellular formation and reduced fresh weight

Table 5.1 Proposed functions and localisations for the GTG/GPHR family of membrane proteins

Principal function proposed	Mechanism	Localisation	Reference
ABA receptor	GPCR-type G protein	Plasma membrane	Pandey et al. (2009)
Golgi pH regulator	Anion channel	Golgi	Maeda et al. (2008)
Endomembrane pH regulator	Not specified	Golgi and ER	Jaffé et al. (2012)
Golgi pH regulator	Not specified	Golgi and ER	Charroux and Royet (2014)

(section 4.2.2) (Jaffé et al., 2012). From this it was suggested that At GTGs have a role in fertility, growth and cell structure/integrity (Jaffé et al., 2012). Functional complementation assays observing a selection of these phenotypes indicated that the root and hypocotyl defects can be rescued by either At GTG1 or At GTG2 expression in the double mutants; that is, normal development is restored (Jaffé et al., 2012). The root length for instance is significantly shorter in the *gtg1 gtg2* mutants but can be rescued with either At GTG1 or At GTG2, restoring the root growth (Jaffé et al., 2012). Therefore At GTGs are able to complement *gtg1 gtg2* KO plants, but would GTG/GPHRs from a different species be able to accomplish this effect? For instance, would an animal GTG/GPHR rescue the phenotypes of the plant *gtg1 gtg2* mutant in a similar fashion to the At *GTGs*? If the phenotypes of the *gtg1 gtg2* can be rescued by heterologous expression of a GTG/GPHR from another species, this would suggest that there is conservation of function between this family of proteins. The *gtg1-3 gtg2-3* mutant was used for all experiments here so will be referred to as *gtg1 gtg2* throughout this chapter.

5.1.3 Nematode gene expression in plants

The role of *Cel*-GPHRs has not yet been characterised. As a starting point to characterising the *Cel*-GPHRs, functional complementation assays could test whether the *Cel*-GPHRs are able function in a similar way to the At GTGs and rescue the phenotypes in the Arabidopsis *gtg1 gtg2* mutant. If this were the case, it would suggest

that the *Cel*-GPHRs have a similar function to the At GTGs and would imply conservation of function of the GTG/GPHR family of membrane proteins.

Nematode gene expression in plants is an unusual and uncommon approach. However, there have been some reports utilising this methodology. The plant-parasitic nematode, *Meloidogyne incognita* expresses a dorsal gland protein, 7E12 that accelerated the formation of the feeding site (dos Santos de Lima e Souza et al., 2011). Constitutive expression of 7E12 stimulated faster feeding and cell development of *Meloidogyne incognita* in transgenic tobacco plants and consequently accelerated root-knot nematode egg hatching (dos Santos de Lima e Souza et al., 2011). Another study showed that *C. elegans* phytochelatin synthase (PCS) transformation into tobacco plants caused increased cadmium and arsenic tolerance (Wojas et al., 2008, Wojas et al., 2010). Another study investigated the *Hetero schachtii 4F01* (*Hs4F01*) gene of the root-parasitic soybean cyst nematode. *Hs4F01* encodes an annexin-like effector that functions in host infection. Hs4F01 has 33 % identity to Arabidopsis annexin-1 (annAt1) at the amino acid level. The *annAt1* mutant was hypersensitive to 75 mM NaCl, which was rescued by constitutively expressing Hs4F01. This implied a similar role of the Hs4F01 annexin-like effector to annAt1 in plant cell stress responses (Patel et al., 2010). This study aimed to investigate whether expression of the nematode *Cel-gphr* genes in Arabidopsis could be used to assess conservation of function.

5.1.4 Localisation of the GTG/GPHRs

As mentioned previously, the subcellular localisation of the GTG/GPHRs is yet to be fully determined (section 1.4.2). If the *Cel*-GPHRs are able to functionally complement *gtg1 gtg2* mutant phenotypes, these *Cel*-GPHRs are likely to localise to the same site as the At GTGs. Furthermore if transgenic plant lines expressing either *P35S::Cel-gphr-1::gfp* or *P35S::Cel-gphr-2::gfp* are able to rescue *gtg1 gtg2* phenotypes and the localisation is disclosed simultaneously, it signifies that the GFP is not interfering with functionality of the GPHRs. Since the role for the GTG/GPHRs is yet to be fully resolved, determining the subcellular localisation for these proteins would therefore be important for supporting either of their proposed functions: ABA receptor at the plasma membrane (Pandey et al., 2009) or Golgi pH regulator at the Golgi (Maeda et al., 2008). On the other hand, the subcellular localisation could reveal an alternative role for these proteins.

5.1.5 Aims

In order to investigate the conservation of function of the GTG/GPHR family of membrane proteins, there are a number of aims in this chapter.

1. Generate expression constructs containing the *Cel-gphr* CDSs, which will be used to generate transgenic Arabidopsis lines.
2. To isolate several independent lines expressing *Cel-gphr* genes in Arabidopsis WT and *gtg1 gtg2*.
3. To conduct phenotypic assays (root growth, fresh weight and hypocotyl growth) on these lines to explore the conservation of function between Arabidopsis and *C. elegans* GTG/GPHRs. This will reveal whether *C. elegans* GTG/GPHRS can complement the defects in Arabidopsis *gtg1 gtg2* mutants.
4. To examine the cellular localisation of the *Cel*-GPHRs by generating GFP-tagged expression constructs and producing transgenic Arabidopsis lines. In addition to determine whether these tagged *Cel*-GPHR constructs will rescue the Arabidopsis *gtg1 gtg2* phenotypes.

These approaches will help clarify whether there is a conserved function for the GTG/GPHR family of membrane proteins.

5.2 Results

5.2.1 Cloning *Cel-gphr* genes into Arabidopsis expression vectors

To express the *Cel-gphr* genes in Arabidopsis to study conservation of function, the *Cel-gphr* CDSs were cloned into Arabidopsis expression vectors. The online database, WormBase (http://www.wormbase.org/#012-3-6), is a *C. elegans* resource with genome and biology information for nematode research, including information on CDSs (Schwarz et al., 2006). The Directional TOPO® Cloning System (Invitrogen, CA, USA) was employed for *C. elegans* gene cloning. To clone the *Cel-gphr* CDSs, primers for cloning the *Cel-gphr* CDSs ± stop codons (without stop codon allows for C-terminal tagging) were designed based on the the requirements of the TOPO® Cloning System and *Cel-gphr* sequences provided by WormBase (*Cel-gphr-1* = Y75B8A.16, WBGene00013551; *Cel-gphr-2* = C11H1.2, WBGene00007528). Total RNA was used to synthesise cDNA, which was used with the designed primers to clone the *Cel-gphr* CDSs (Figure 5.1). The *Cel-gphr* CDSs were successfully amplified as shown by the predicted product sizes (Figure 5.1). The CDSs of *Cel-gphr-1* and *Cel-gphr-2* ± stop codons were individually cloned into the pENTR/D entry vector to generate pENTR/D *Cel-gphr-1* ± stop codon and pENTR/D *Cel-gphr-2* ± stop codon respectively. The entry vectors were transformed into One Shot® TOP10 chemically competent *E. coli* cells and the plasmid DNA was extracted from a few of the colonies. To confirm that the genes had inserted into the vector in the correct orientation, the entry vectors were analysed by restriction analysis. The restriction enzyme analyses and gel electrophoresis results indicated that the pENTR/D vectors contained *Cel-gphr-1* or *Cel-gphr-2* ± stop codons correctly inserted (Table 2.15). Figure 5.2 shows an example of the predicted digest products for pENTR/D *Cel-gphr-1*.

The entry vectors were sequenced in both forward and reverse directions using primers M13 F (-20) and M13 R outlined in section 2.13.3 (Table 2.5). The sequence analysis confirmed that the *Cel-gphr* sequences corresponded to the CDSs as predicted by WormBase (*Cel-gphr-1* = Y75B8A.16, WBGene00013551; *Cel-gphr-2* = C11H1.2, WBGene00007528) and were inserted into the pENTR/D entry vector correctly. Table 5.2 displays the entry vectors generated containing *Cel-gphr* ± stop codons produced in this study.

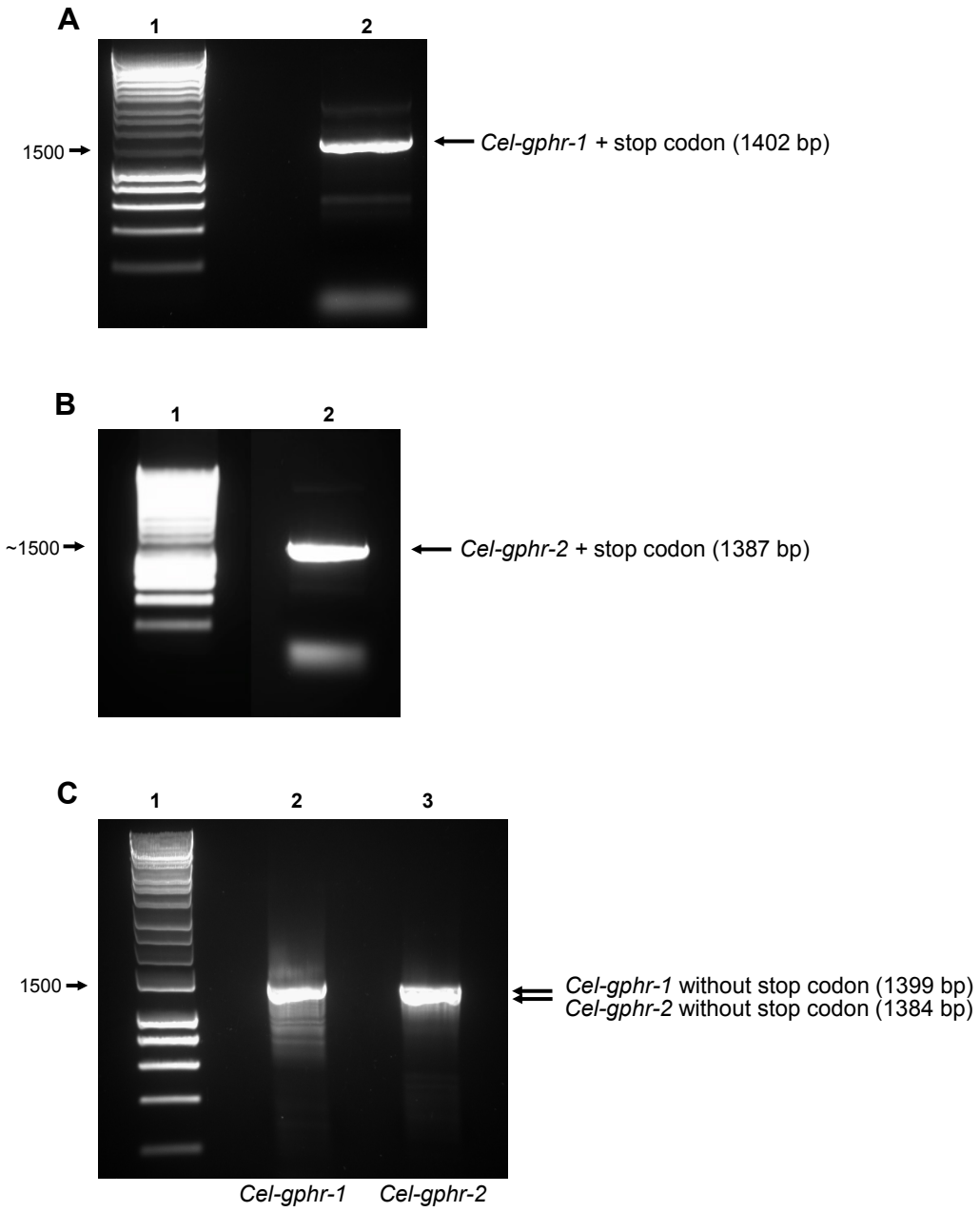


Figure 5.1 *Cel-gphr* products amplified from *C. elegans* cDNA

Cel-gphr gene products ± stop codons amplified by PCR from *C. elegans* cDNA. Predicted product sizes are indicated on right and predicted molecular marker sized indicated on left (lane 1). **A**. *Cel-gphr-1* with stop codon amplified using primers CeGTG1 F and CeGTG1 R (lane 2). **B**. *Cel-gphr-2* with stop codon amplified using primers CeGTG2 F and CeGTG2 R (lane 2). **C**. *Cel-gphr* products amplified without stop codon. CDSs were amplified using primers CeGTG1 F and CeGTG1 NO STOP R for *Cel-gphr-1* (lane 2), CeGTG2 F and CeGTG2 NO STOP R primers for *Cel-gphr-2* (lane 3). Primer sequences outlined in Table 2.10.

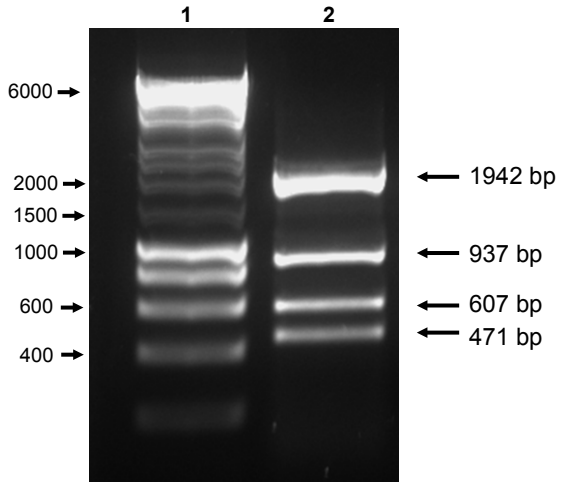


Figure 5.2 Diagnostic restriction analysis of pENTR/D *Cel-gphr-1* entry vector

Analysis of *Cel-gphr-1* cloned into TOPO pENTR/D Entry Vector ™ using restriction enzymes *Dra*I and *Pvu*II. Predicted product sizes are indicated on right and predicted molecular marker sized indicated on left (lane 1).

Table 5.2 The Gateway-compatible entry vectors constructed

Plasmid	**Description**
pENTR/D *Cel-gphr-1* with stop codon	CDS of *Cel-gphr-1*
pENTR/D *Cel-gphr-1* without stop codon	CDS of *Cel-gphr-1* that can be C-terminally tagged
pENTR/D *Cel-gphr-2* with stop codon	CDS of *Cel-gphr-2*
pENTR/D *Cel-gphr-2* without stop codon	CDS of *Cel-gphr-2* that can be C-terminally tagged

Prior to LR recombination reaction, the pENTR/D *Cel-gphr-1* and pENTR/D *Cel-gphr-2* entry vectors were subjected to an overnight restriction digestion with *Sfa*I and *Pvu*I enzymes respectively; this was to linearise the entry vectors by cutting within the kanamycin resistance gene to aid insertion into and selection of the destination vectors. The *Cel-gphr* CDSs were inserted into the destination vectors pMDC32 and pMDC83 via the LR recombination reaction using the linearised entry vectors followed by transformation into *E. coli* DH5α cells. To check that *Cel-gphr-1* and *Cel-gphr-2* CDSs were inserted into pMDC32 and pMDC83 correctly, diagnostic restriction analyses and gel electrophoresis was performed on plasmid DNA preparations extracted from individual colonies containing successful transformants. The results indicate that both *Cel-gphr-1* and *Cel-gphr-2* with stop codons were inserted into pMDC32 and *Cel-gphr-1* and *Cel-gphr-2* without stop codons were inserted into pMDC83 correctly, as shown by the predicted digest product sizes (Table 2.16.) Figure 5.3 shows an example of the predicted digest products for pMDC32 *Cel-gphrs* with stop codons.

The expression vectors were sequenced in both forward and reverse directions using primers pMDC35S and MDCnosR as outlined in section 2.13.4 (Table 2.5). The sequence analysis confirmed that both the *Cel-gphr* sequences corresponded to the CDSs provided by WormBase (*Cel-gphr-1* = Y75B8A.16, WBGene00013551; *Cel-gphr-2* = C11H1.2, WBGene00007528) and were inserted into the pMDC32 and pMDC83 destination vectors correctly. Table 5.3 displays the Arabidopsis expression vectors generated in this study.

5.2.2 Expressing *C. elegans gphr* genes in Arabidopsis plants

The expression vectors were transformed into *Agrobacterium* for subsequent Arabidopsis transformation. A few colonies were selected and analysed using colony PCR to identify successful transformants. The primer sets used contained sequences that would anneal to the full-length of the *Cel-gphr* genes. Gel electrophoresis results indicate that the predicted products produced from colony PCR corresponded to *Cel-gphr-1* or *Cel- gphr-2*. Figure 5.4 shows an example of colony PCR and the predicted products for pMDC32 *Cel-gphr-1* (Figure 5.4 A) and pMDC32 *Cel-gphr-2* (Figure 5.4 B) in GV3850. *Agrobacterium* transformants containing pMDC32 *Cel-gphr-1*, pMDC32 *Cel-gphr-2*, pMDC83 *Cel-gphr-1* and pMDC83 *Cel-gphr-2* were used to transform Arabidopsis WT and *gtg1 gtg2* following the scheme outlined in

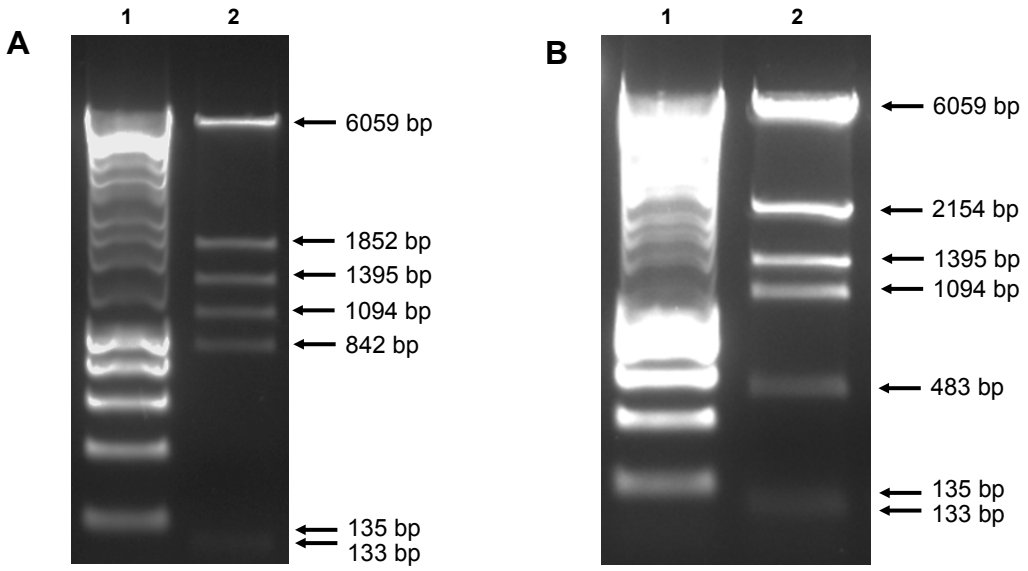


Figure 5.3 Diagnostic restriction analysis of pMDC32 *Cel-gphr* expression vectors

Analysis of *Cel-gphr-1* and *Cel-gphr-2* genes in plant destination vector pMDC32. Predicted product sizes are shown on the figure. Lane 1: molecular markers (see Figure 5.2 for predicted molecular sizes). **A**. Digestion of pMDC32 *Cel-gphr-1* with enzymes *Eco*RV and *Xho*I (lane 2). **B**. Digestion of pMDC32 *Cel-gphr-2* with enzymes *Eco*RV and *Xho*I (lane 2).

Table 5.3 The expression vectors generated for expressing *Cel-gphr* genes in plants (Arabidopsis) in this study

Plasmid	Construct
pMDC32 *Cel-gphr-1*	*P35S::Cel-gphr-1*
pMDC32 *Cel-gphr-2*	*P35S::Cel-gphr-2*
pMDC83 *Cel-gphr-1*	*P35S::Cel-gphr-1::gfp*
pMDC83 *Cel-gphr-2*	*P35S::Cel-gphr-2::gfp*

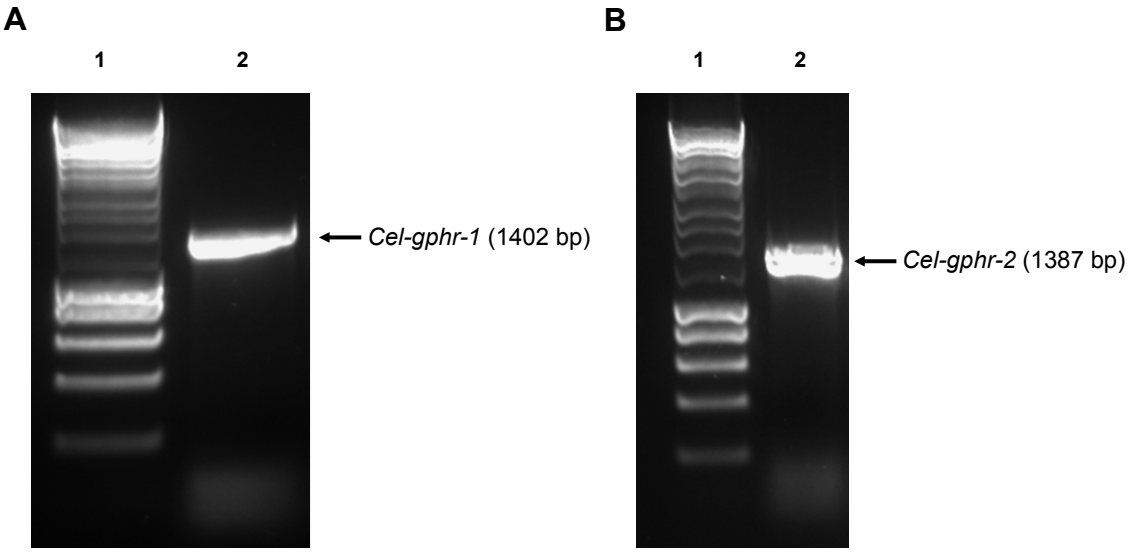


Figure 5.4 *Agrobacterium* transformed with pMDC32 *Cel-gphr* genes

Colony PCR and gel electrophoresis to show the successful transformation of *Agrobacterium* GV3850 with the Arabidopsis expression vectors pMDC32 *Cel-gphr-1* (**A**) and pMDC32 *Cel-gphr-2* (**B**). Predicted product sizes shown on figure. Lane 1: molecular markers (see Figure 5.2 for predicted molecular sizes). **A**. GV3850 pMDC32 *Cel-gphr-1* showing the predicted 1402 bp product (lane 2) using primers CeGTG1 F and CeGTG1 R. **B**. GV3850 pMDC32 *Cel-gphr-2* showing the predicted 1402 bp product (lane 2) using primers CeGTG2 F and CeGTG2 R (primer sequences outlined in Table 2.10).

section 2.3.3 (Figure 2.1). Table 5.4 shows the transgenic lines generated for this chapter. Three to four final independent lines that were 100 % resistant to hygromycin were identified for each plasmid.

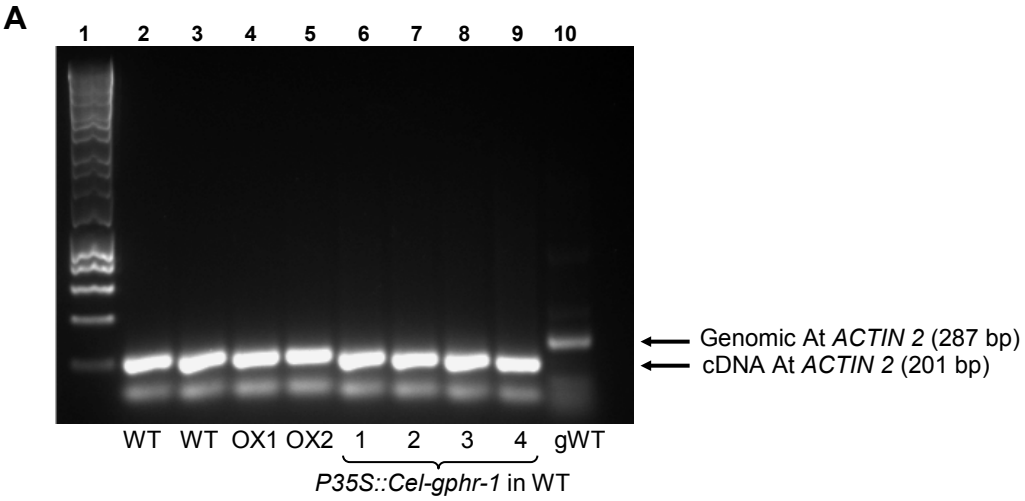
RNA was extracted and cDNA was synthesised to confirm the RNA transcript expression of the *Cel-gphr* products in the T_3 plants. Figure 5.5 and Figure 5.6 are example PCR results showing WT (Figure 5.5) and *gtg1 gtg2* mutant plants (Figure 5.6) transformed with pMDC32 *Cel-gphr-1* and pMDC32 *Cel-gphr-2*. To check that the RNA was good quality and there was no genomic DNA contamination in the samples of RNA extracted, the primers amplifying At *ACTIN2* (*ACTIN 2* = AT3G18780) were used. All plant lines express At *ACTIN 2* at the RNA level (lanes 2 – 9) as shown by the amplification of predicted 201 bp products (Figure 5.5 A). Furthermore, none of these samples (Figure 5.5 A lanes 2 – 9) were contaminated by genomic DNA as shown by the absence of the predicted product size of 287 bp for genomic At *ACTIN 2* (Figure 5.5 A lane 10). All transformed and untransformed WT plants expressed full-length At *GTG1* and At *GTG2* transcripts, as indicated by the predicted 1407 and 1404 bp products respectively (Figure 5.5 B and C). The four independent WT lines transformed with pMDC32 *Cel-gphr-1* (lanes 6 – 9) display the full-length predicted product of 1402 bp (Figure 5.5 D). The four independent WT lines transformed with pMDC32 *Cel-gphr-2* (lanes 4 – 7) display the full-length predicted product of 1387 bp (Figure 5.5 E). WT expressed At *GTG1* as shown by the predicted 327 bp fragment of At *GTG1* (lane 1), while all transformed and untransformed *gtg1 gtg2* double mutant plants do not express the At *GTG1* transcript (Figure 5.6 B lanes 2 – 6). Four independent *gtg1 gtg2* lines transformed with pMDC32 *Cel-gphr-1* or pMDC32 *Cel-gphr2* display the full-length predicted 1402 bp and 1387 bp products for *Cel-gphr-1* and *Cel-gphr-2*, respectively (Figure 5.6 C and D). Three independent *gtg1 gtg2* lines transformed with pMDC32 *Cel-gphr-1::gfp* display *Cel-gphr-1* expression (Figure 5.6 E; predicted band 1402 bp). There appears to be two products produced of very similar size when using primers CeGTG1 F and CeGTG1 R to amplify the full-length *Cel-gphr-1* in the *gtg1 gtg2* double mutant plants (Figure 5.6 C). This could be due to a post-transcriptional splicing event that may have occurred. These genotyping analyses by PCR were conducted using a high number of cycles to check the presence of the At *GTG* genes.

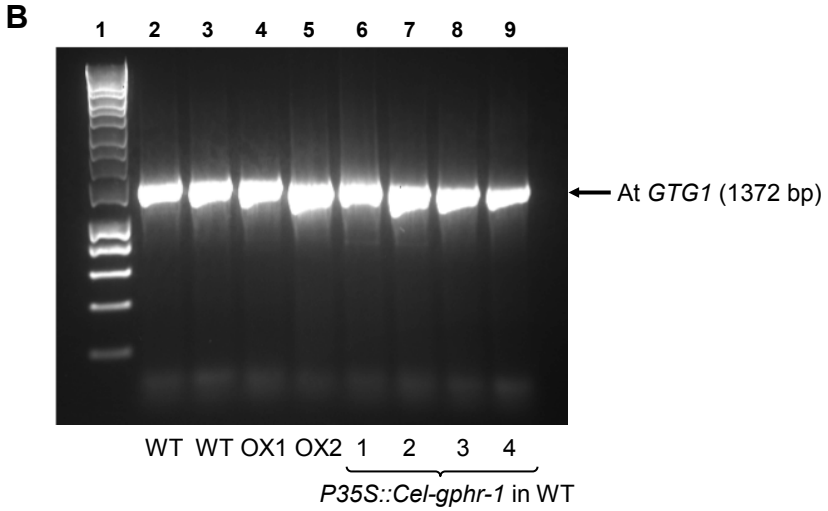
Table 5.4 The transgenic Arabidopsis lines created for expressing *Cel-gphr* genes

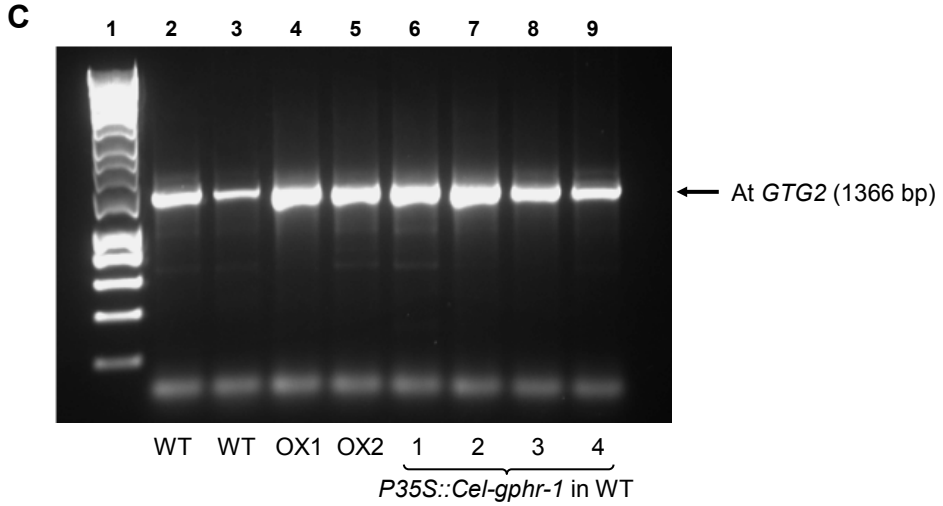
Each plasmid was transformed into WT and *gtg1 gtg2* plants. The construct specifies the promoter the gene is expressed under and whether there are any fusions e.g. to a GFP reporter.

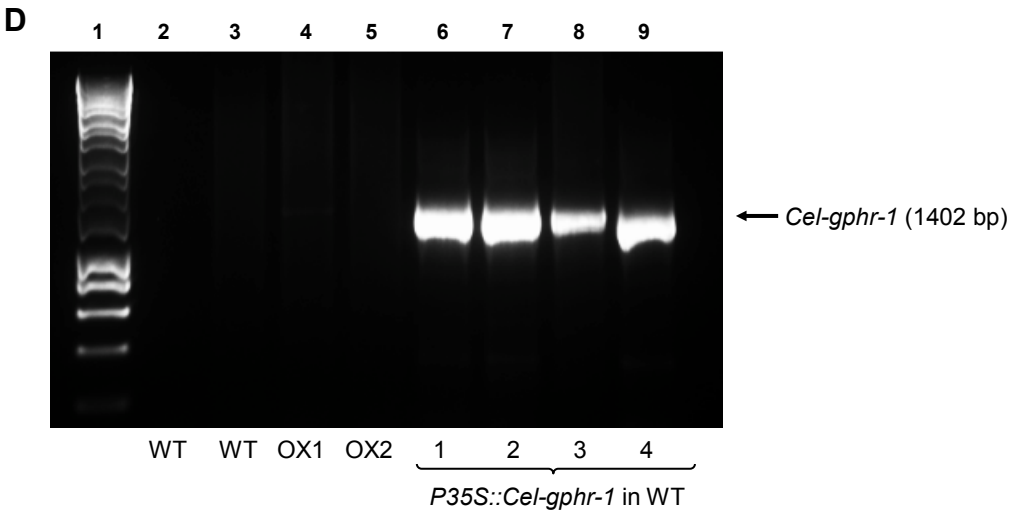
Plasmid	Construct	Transformed into	# of independent lines
pMDC32 *Cel-gphr-1*	*P35S::Cel-gphr-1*	WT	4
pMDC32 *Cel-gphr-1*	*P35S::Cel-gphr-1*	*gtg1 gtg2*	4
pMDC32 *Cel-gphr-2*	*P35S::Cel-gphr-2*	WT	4
pMDC32 *Cel-gphr-2*	*P35S::Cel-gphr-2*	*gtg1 gtg2*	3
pMDC83 *Cel-gphr-1*	*P35S::Cel-gphr-1::gfp*	WT	*
pMDC83 *Cel-gphr-1*	*P35S::Cel-gphr-1::gfp*	*gtg1 gtg2*	3
pMDC83 *Cel-gphr-2*	*P35S::Cel-gphr-2::gfp*	WT	*
pMDC83 *Cel-gphr-2*	*P35S::Cel-gphr-2::gfp*	*gtg1 gtg2*	3

* indicates transgenic plants still undergoing isolation process.


A
1 2 3 4 5 6 7 8 9 10
Genomic At *ACTIN 2* (287 bp)
cDNA At *ACTIN 2* (201 bp)
WT WT OX1 OX2 1 2 3 4 gWT
P35S::Cel-gphr-1 in WT


B
1 2 3 4 5 6 7 8 9
At *GTG1* (1372 bp)
WT WT OX1 OX2 1 2 3 4
P35S::Cel-gphr-1 in WT


C
1 2 3 4 5 6 7 8 9
At *GTG2* (1366 bp)
WT WT OX1 OX2 1 2 3 4
P35S::Cel-gphr-1 in WT



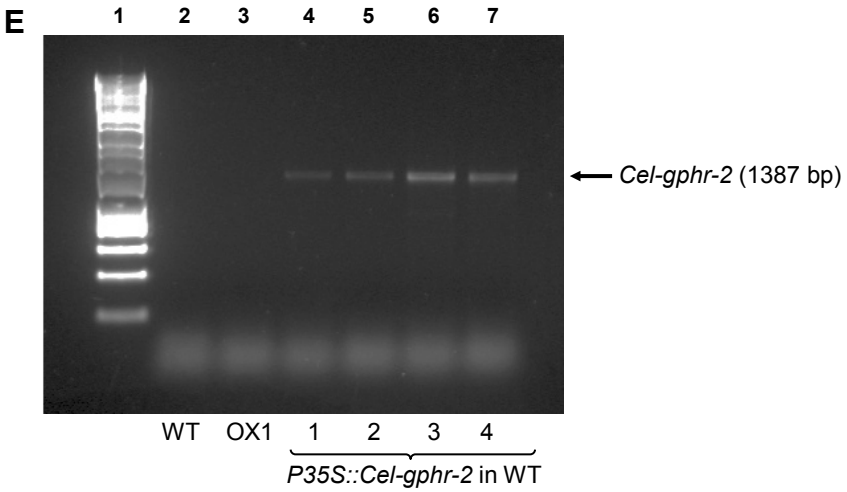
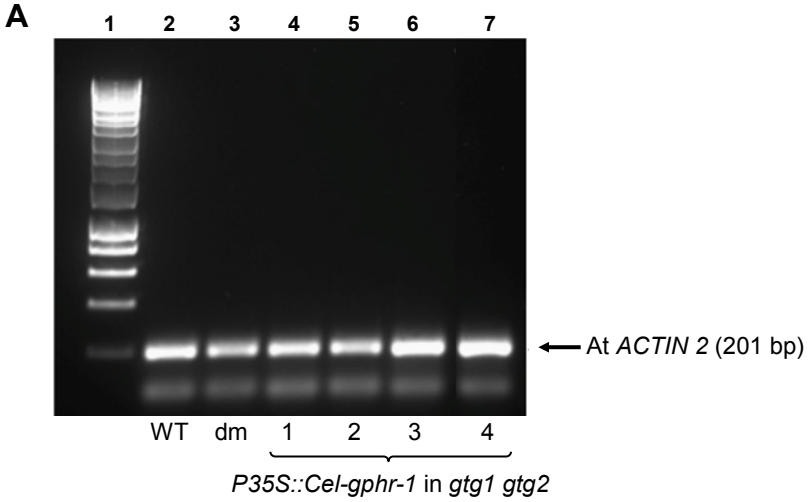
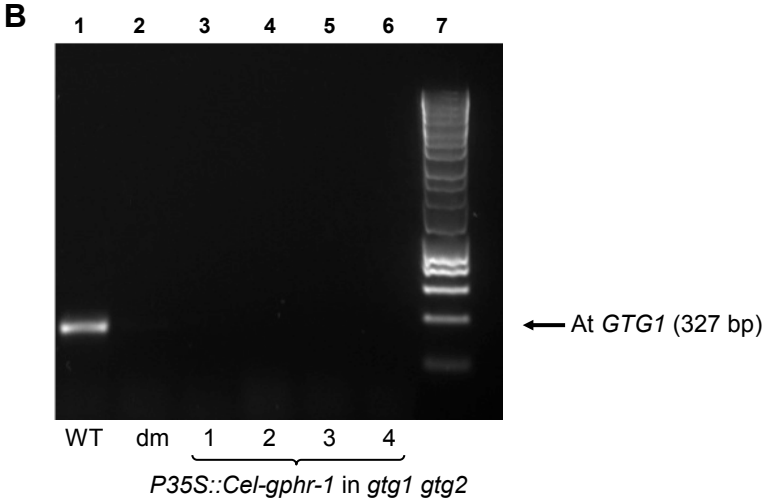
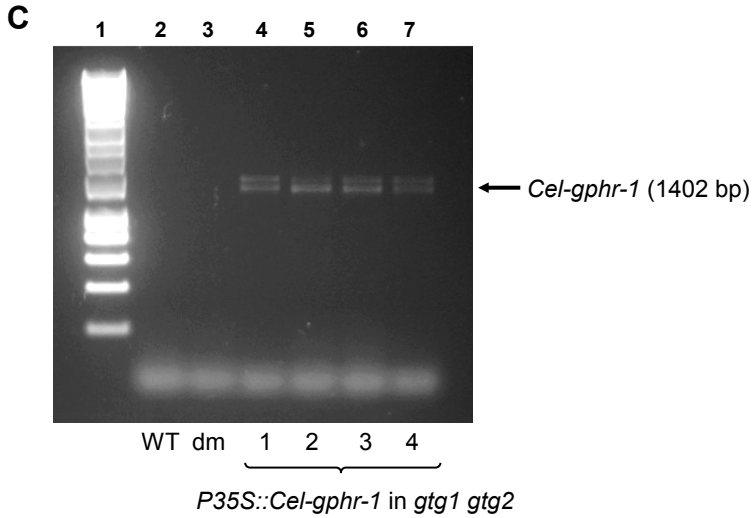


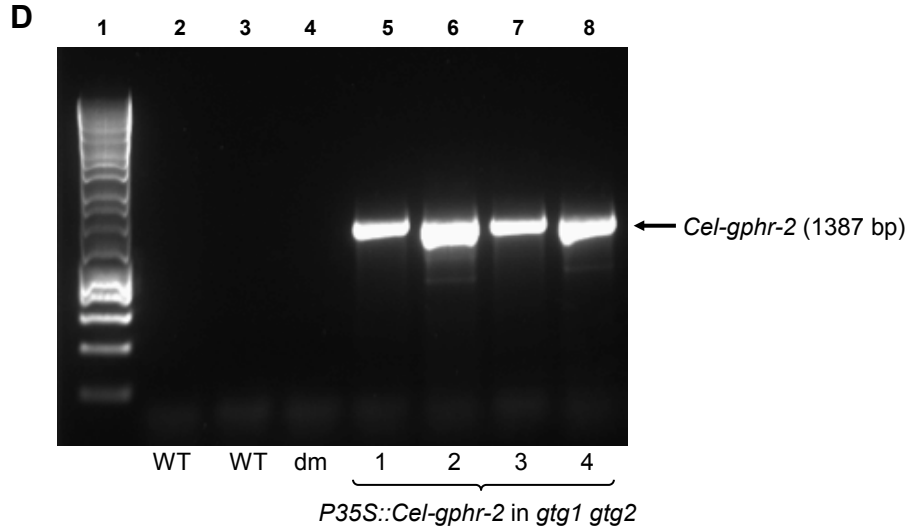
Figure 5.5 Expression of *Cel-gphr* in Arabidopsis WT T_3 lines

PCR and gel electrophoresis using cDNA synthesised from RNA to show the genotypes of WT lines transformed with pMDC32 *Cel-gphr-1* (**A** – **D**) and pMDC32 *Cel-gphr-2* (**E**). Predicted products shown on figure. Lane 1: molecular markers (see Figure 5.1 and Figure 5.2 for predicted molecular sizes), OX = WT **o**vere**x**pressing At *GTGs* (**A** – **D**: At GTG1 and **E**: At GTG2). **A**. Actin2 F and Actin2 R primers amplify a fragment of At *ACTIN 2*. These plant lines show a consistent level of At *ACTIN 2* RNA transcript expression (201 bp predicted product). There is no genomic DNA contamination as shown by the lack of the predicted 287 bp product (lane 10) if genomic DNA was present (gWT = **g**enomic DNA from **WT**). **B**. Arabidopsis *GTG1* expression in all plant lines using primers AtGTG1 F and AtGTG1 R. **C**. Arabidopsis *GTG2* expression in all plant lines using primers AtGTG2 1F and AtGTG2 R. **D**. *Cel-gphr-1* expression in the four WT lines transformed with pMDC32 *Cel-gphr-1* using primers CeGTG1 F and CeGTG2 R (lanes 6 – 9). **E**. *Cel-gphr-2* expression in the four WT lines transformed with pMDC32 *Cel-gphr-2* using primers CeGTG2 F and CeGTG2 R (lanes 4 – 7). Primer sequences outlined in Table 2.10.


A
1 2 3 4 5 6 7
At *ACTIN 2* (201 bp)
WT dm 1 2 3 4
P35S::Cel-gphr-1 in *gtg1 gtg2*


B
1 2 3 4 5 6 7
At *GTG1* (327 bp)
WT dm 1 2 3 4
P35S::Cel-gphr-1 in *gtg1 gtg2*


C
1 2 3 4 5 6 7
Cel-gphr-1 (1402 bp)
WT dm 1 2 3 4
P35S::Cel-gphr-1 in *gtg1 gtg2*



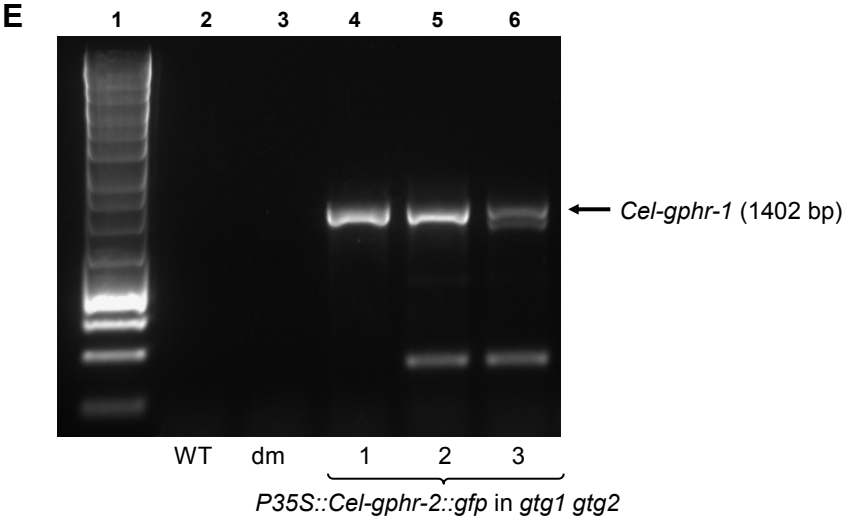


Figure 5.6 Expression of *Cel-gphr* in Arabidopsis *gtg1 gtg2* T_3 lines

PCR and gel electrophoresis using cDNA synthesised from RNA to show the genotypes of *gtg1 gtg2* **d**ouble **m**utant (dm) lines (1 – 4) transformed with pMDC32 *Cel-gphr-1* (**A** – **C**) or pMDC32 *Cel-gphr-2* (**D**). Predicted products shown on figure (see Figure 5.1 and Figure 5.2 for predicted molecular sizes for the molecular markers). **A**. Actin2 F and Actin2 R primers amplify a fragment of At *ACTIN 2*, which shows consistent level of expression in all plant lines. Lane 1: molecular markers. **B**. Arabidopsis GTG1 expression is present only in the WT plant line using primers AtGTG1.ex12 qF1 and AtGTG1.ex12 qR1 (lane 1). Lane 7: molecular markers. **C**. *Cel-gphr-1* expression in four *gtg1 gtg2* mutant lines transformed with pMDC32 *Cel-gphr-1* using primers CeGTG1 F and CeGTG1 R (lanes 4 – 7). Lane 1: molecular markers. **D**. *Cel-gphr-2* expression in four *gtg1 gtg2* mutant lines transformed with pMDC32 *Cel-gphr-2* using primers CeGTG2 F and CeGTG2 R (lanes 5 – 8). Lane 1: molecular markers. **E**. *Cel-gphr-1::gfp* expression in three *gtg1 gtg2* mutant lines transformed with pMDC32 *Cel-gphr-1::gfp* using primers CeGTG1 F and CeGTG1 R (lanes 4 – 6). Primer sequences outlined in Table 2.10.

5.2.3 *C. elegans* GPHR-1 functionally complements Arabidopsis *gtg1 gtg2* mutant phenotypes

Arabidopsis *gtg1 gtg2* double mutants exhibit a range of phenotypes including reduced root and hypocotyl growth, reduced fresh weight and distended epidermal cells in the hypocotyl (Jaffé et al., 2012). The approach to investigate conservation of function was to test whether these phenotypes could be rescued by expressing *Cel-gphr*.

5.2.3.1 The *C. elegans* GPHR-1 rescues the root growth and fresh weight defect in the Arabidopsis *gtg1 gtg2* mutant

Time-course measurements show that root growth in Arabidopsis *gtg1 gtg2* mutants is inhibited compared to WT in the absence of sucrose (Jaffé et al., 2012). The expression of either At GTG1 or At GTG2 in the *gtg1 gtg2* mutant can rescue this defect (Jaffé et al., 2012). To investigate whether the GTG/GPHR family of membrane proteins have a conserved function, *Cel-gphr* genes were constitutively expressed in the *gtg1 gtg2* double mutant. Measurements of the root lengths were conducted at 7, 10, 12 and 14 d to compare the growth of WT, *gtg1 gtg2* and *gtg1 gtg2* expressing *P35S::Cel-gphr-1* (Figure 5.7). Root growth increases from day 7 to 14 for all plant lines; however, root growth rates are different between WT and *gtg1 gtg2* expressing *P35S::Cel-gphr-1* compared to the *gtg1 gtg2* mutant. Over the course of 14 d, there is little root growth in the *gtg1 gtg2* mutant compared to the WT and *gtg1 gtg2* expressing *P35S::Cel-gphr-1* (Figure 5.7). Time-course experiments were carried out for all root assays as in Figure 5.7; however in some experiments at day 12 and day 14, some roots had reached the bottom of the plate so day 10 may be the most valid measurement. In addition, on day 14 of the root length time-course measurements, the fresh weight data was also recorded. For this reason, all root growth images and graphs are shown at 10 d, as in Figure 5.8 A and C, while fresh weight data is shown at 14 d, as in Figure 5.8 B and D. The root growth defect seen in *gtg1 gtg2* can be rescued by both *Cel-gphr-1* and *Cel-gphr-1::gfp* expression (Figure 5.7 and Figure 5.8).

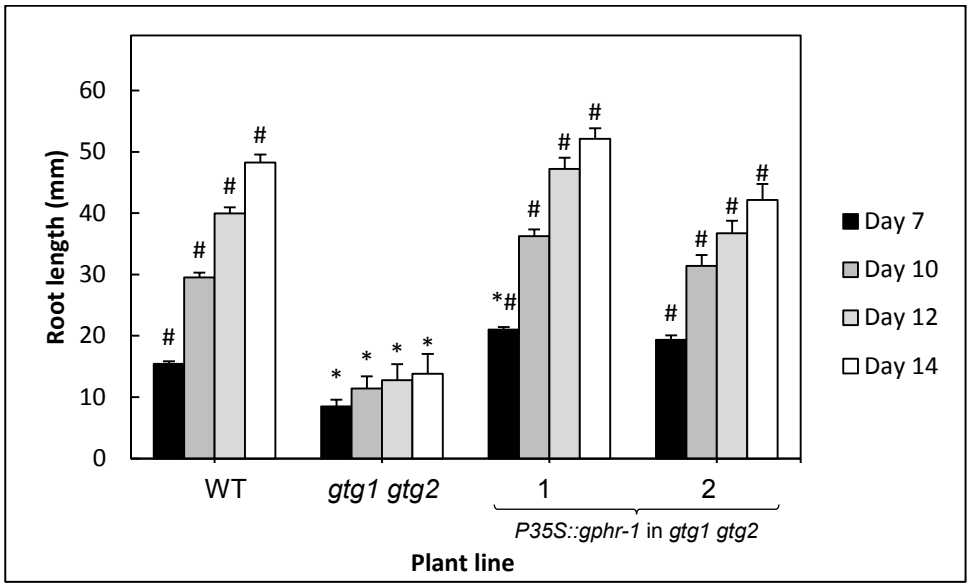


Figure 5.7 Comparing root growth of Arabidopsis *gtg1 gtg2* mutants transformed with pMDC32 *Cel-gphr-1*

Time-course measurements of WT, *gtg1 gtg2* and *gtg1 gtg2* expressing *P35S::Cel-gphr-1* (lines 1 and 2) showing root growth at 7, 10, 12 and 14 d on 0.5 MS containing 0 % sucrose. Data shows the mean root length calculated for six plates (+S.E.) with four seedlings per plant line per plate, each plate containing four plant lines. Results represent one of four independent experiments using four independent *P35S::Cel-gphr-1* expressing lines (lines 1 and 2 are shown). At each time point, there is a significant difference in root length between plant lines as determined by one-way ANOVAs. Day 7: ($F_{3, 20} = 22.65$, $p < 0.0001$), day 10: ($F_{3, 20} = 15.11$, $p < 0.0001$), day 12: ($F_{3, 20} = 11.00$, $p < 0.0001$) and day 14: ($F_{3, 20} = 12.30$, $p < 0.0001$). For each time-point (i.e. compared to the same day), means significantly different to WT are denoted with a * and means significantly different to the *gtg1 gtg2* mutant are denoted with a #, determined by Tukey *post-hoc* tests.

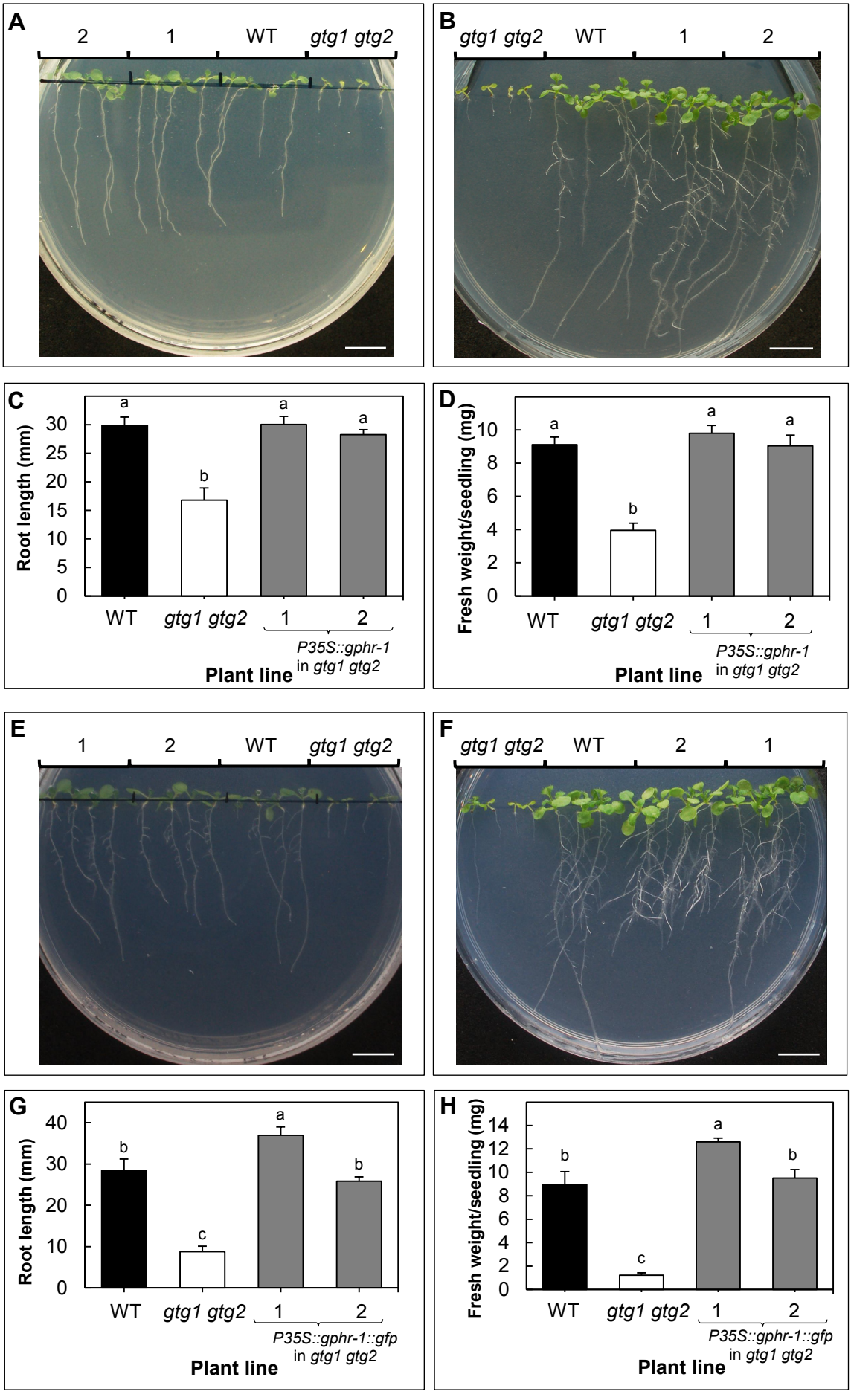

A
2
1
WT
gtg1 gtg2
B
gtg1 gtg2
WT
1
2
C
Root length (mm)
0
5
10
15
20
25
30
a
b
a
a
WT
gtg1 gtg2
1
2
P35S::gphr-1
in gtg1 gtg2
Plant line
D
Fresh weight/seedling (mg)
0
2
4
6
8
10
a
b
a
a
WT
gtg1 gtg2
1
2
P35S::gphr-1
in gtg1 gtg2
Plant line
E
1
2
WT
gtg1 gtg2
F
gtg1 gtg2
WT
2
1
G
Root length (mm)
0
10
20
30
40
b
c
a
b
WT
gtg1 gtg2
1
2
P35S::gphr-1::gfp
in gtg1 gtg2
Plant line
H
Fresh weight/seedling (mg)
0
2
4
6
8
10
12
14
b
c
a
b
WT
gtg1 gtg2
1
2
P35S::gphr-1::gfp
in gtg1 gtg2
Plant line

Figure 5.8 *Cel*-GPHR-1 functionally complements the root length and fresh weight defects of Arabidopsis *gtg1 gtg2*

Root growth and fresh weight per seedling of *gtg1 gtg2* expressing *Cel*-GPHR-1 on 0.5 MS containing 0 % sucrose. Representative images of seedlings at 10 d (**A** and **E**) and 14 d (**B** and **F**) (1 and 2 are *P35S::Cel-gphr-1* or *P35S::Cel-gphr-1::gfp* in *gtg1 gtg2* for **A** and **B** or **E** and **F** respectively). Data shows the mean root length (**C** and **G**) and fresh weight per seedling (**D** and **H**) calculated for six plates (+S.E.) with four seedlings per plant line per plate, each plate containing four plant lines. Results are representative of four independent experiments using four independent *P35S::Cel-gphr-1* or from one experiment using three independent *P35S::Cel-gphr-1::gfp* expressing lines where lines 1 and 2 are shown for both. **C** and **G**. Mean root length at 10 d (+S.E.). There is a significant difference in root lengths between plant lines **C**: ($F_{3, 20}$ = 10.28, $p < 0.0001$), **G**: ($F_{3, 20}$ = 37.18, $p < 0.0001$); one-way ANOVAs. Means not sharing a letter are significantly different, Tukey *post-hoc* test. **D** and **H**. Mean fresh weight per seedling at 14 d (+S.E.). There is a significant difference in fresh weight per seedling between plant lines **D**: ($F_{3, 20}$ = 28.20, $p < 0.0001$), **H**: ($F_{3, 20}$ = 19.55, $p < 0.0001$); one-way ANOVAs. Means not share a letter are significantly different, Tukey *post-hoc* test.

The fresh weight measurements parallel the root growth defect. Fresh weight is reduced in the *gtg1 gtg2* compared to WT; this defect was rescued by *P35S::Cel-gphr-1* expression (Figure 5.8). Figure 5.8 shows the root and fresh weight rescue that was seen in four independent experiments using four independent *gtg1 gtg2* lines expressing *P35S::Cel-gphr-1* (lines 1 and 2 shown here). In addition, three independent *gtg1 gtg2* lines expressing *P35S::Cel-gphr-1::gfp* also show root and fresh weight rescue, indicating that the GFP does not affect functionality. These results signify that *Cel*-GPHR-1 functionally complements the loss of both At *GTG* genes.

On sucrose, the root growth and fresh weight of *gtg1 gtg2* mutants is similar to WT (Jaffé et al., 2012). To test the effect sucrose has on the *gtg1 gtg2* mutants expressing *P35S::Cel-gphr-1*, measurements of root growth and fresh weight data were taken. Sucrose rescues both the root growth and fresh weight defects in *gtg1 gtg2*, and *Cel-gphr-1* expressing lines are not significantly different to *gtg1 gtg2* and WT, under these conditions (Figure 5.9).

5.2.3.2 *Cel*-GPHR-1 rescues hypocotyl defects of the Arabidopsis *gtg1 gtg2* mutant

The Arabidopsis *gtg1 gtg2* mutants exhibit hypocotyl defects affecting their length and epidermal cell morphology (Jaffé et al., 2012). Under low white light conditions, *gtg1 gtg2* mutants have shorter hypocotyls and form large, distended epidermal cells on the hypocotyl surface; this defect is exacerbated in the presence of sucrose and both defects can be rescued by expressing either At GTG1 or At GTG2 (Jaffé et al., 2012). The expression of *P35S::Cel-gphr-1* in *gtg1 gtg2* was also able to rescue these hypocotyl defects (Figure 5.10). *Cel*-GPHR-1 can therefore functionally complement the loss of both At *GTG* genes and restore normal hypocotyl and cellular growth to the *gtg1 gtg2* mutant (Figure 5.10).

5.2.4 *Cel*-GPHR-2 does not appear to functionally complement the Arabidopsis *gtg1 gtg2* mutant phenotypes

Using the lines isolated, *Cel*-GPHR-2 did not rescue *gtg1 gtg2* mutant phenotypes. Measurements comparing the root growth of WT, *gtg1 gtg2* and *gtg1 gtg2* lines expressing *P35S::Cel-gphr-2* or *P35S::Cel-gphr-2::gfp* were conducted in the absence of sucrose. Results indicate that in three independent lines for each, the *Cel-gphr-2* or

Cel-gphr-2::gfp did not restore normal root growth to the *gtg1 gtg2* mutant (Figure 5.11). Furthermore the fresh weight defect of *gtg1 gtg2* was also not restored by either *P35S::Cel-gphr-2* or *P35S::Cel-gphr-2::gfp* expression (Figure 5.11 B and D). Since *Cel*-GPHR-2 could not elicit an effect on *gtg1 gtg2* root growth or fresh weight in the absence of sucrose, no root assays were conducted on sucrose.

P35S::Cel-gphr-2 was also not able to restore normal hypocotyl and cellular growth in the *gtg1 gtg2* mutant in the presence of sucrose (Figure 5.12). Since hypocotyl defects are exacerbated in sucrose, no rescue experiments were conducted in the absence of sucrose. In this case, the lack of phenotype rescue in *gtg1 gtg2* could be because of a silenced or low level of *Cel-gphr-2* RNA transcript, however PCR indicated a good amplification, although this was not carried out quantitatively. Since there is a high level of *Cel-gphr-2* expression in *gtg1 gtg2* as shown by PCR, it could be there could be an expression issue at the protein level. Furthermore, there appears to be a splicing event that occurs for *Cel-gphr-1* in the *gtg1 gtg2* mutant (Figure 5.6 C) but not in WT (Figure 5.5 D), which is not apparent for *Cel-gphr-2* (Figure 5.5 E and Figure 5.6 D).

5.2.5 Expression of *Cel-gphr* genes has no effect on Arabidopsis WT

To determine whether expression of *Cel-gphr* had an effect in the presence of At GTGs, the *Cel-gphr* genes were expressed in WT and phenotypic analysis was carried out. The root growth was similar to or exceeded that of WT when the WT lines expressed *P35S::Cel-gphr-1* or overexpressed *P35S:At GTG1*. At 10 d, the root growth of At GTG1 overexpression lines and WT lines expressing *P35S::Cel-gphr-1* was comparable to or exceeded that of WT (Figure 5.13). Furthermore, there was no significant difference in the fresh weights between these plant lines (Figure 5.13). These results are consistent in three independent experiments across four WT lines expressing *P35S::Cel-gphr-1* (lines 1 – 4) and two *P35S:At GTG1* overexpressing lines (OX1 and OX2). WT lines expressing *P35S::Cel-gphr-2* had no significant effect on either root growth or fresh weight in the absence of sucrose (Figure 5.14), and this was consistent across three independent experiments. In addition hypocotyl growth was unaffected by the expression of *Cel-gphrs* (Figure 5.15). The results in combination suggest that there is negligible effect of expressing *P35S::Cel-gphr* in WT and any differences could be as a result of variations within WT seed. The expression of *P35S::Cel-gphr* in WT did not

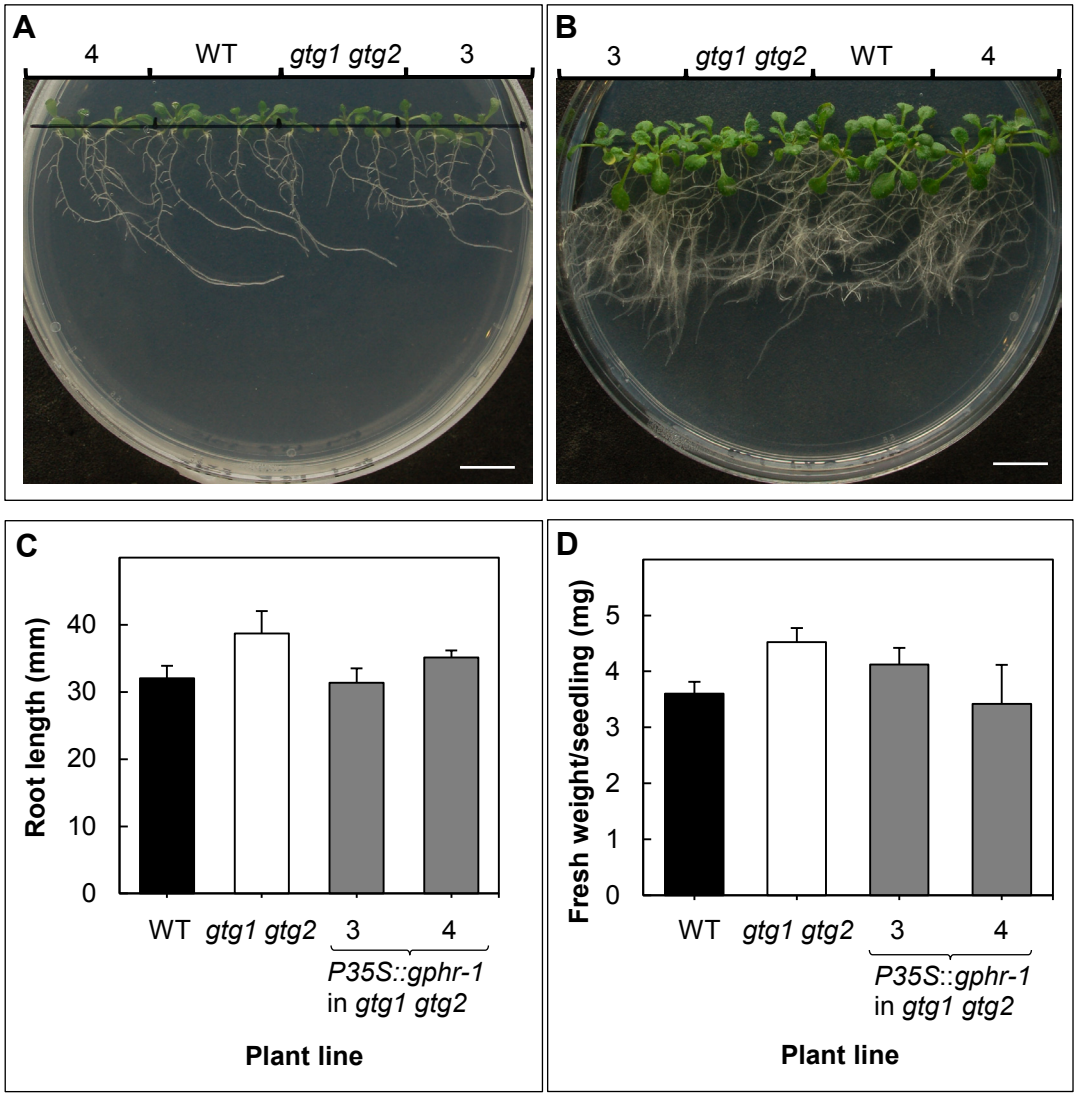


Figure 5.9 Sucrose rescues the root growth defect of the Arabidopsis *gtg1 gtg2* mutant and expression of *C. elegans* GPHR-1 has no marked effect on the response

Root growth and fresh weight of Arabidopsis *gtg1 gtg2* mutant expressing *Cel*-GPHR-1 on 0.5 MS containing 1 % sucrose. Representative images of seedlings at 10 d (**A**) and 14 d (**B**) (3 and 4 are *P35S::Cel-gphr-1* in *gtg1 gtg2*). Data shows the mean root length (**C**) and fresh weight per seedling (**D**) calculated for six plates (+S.E.) with four seedlings per plant line per plate, each plate containing four plant lines. Results represent one of four independent experiments using four independent *P35S::Cel-gphr-1* expressing lines (lines 3 and 4 shown). **C**. Mean root lengths at 10 d (+S.E.). There is no significant difference in root length between plant lines; one-way ANOVA ($F_{3, 20} = 2.23$, $p < 0.117$). **D**. Mean fresh weight at 14 d (+S.E.). There is no significant difference in fresh weight per seedling between plant lines; one-way ANOVA ($F_{3, 20} = 1.46$, $p < 0.256$).

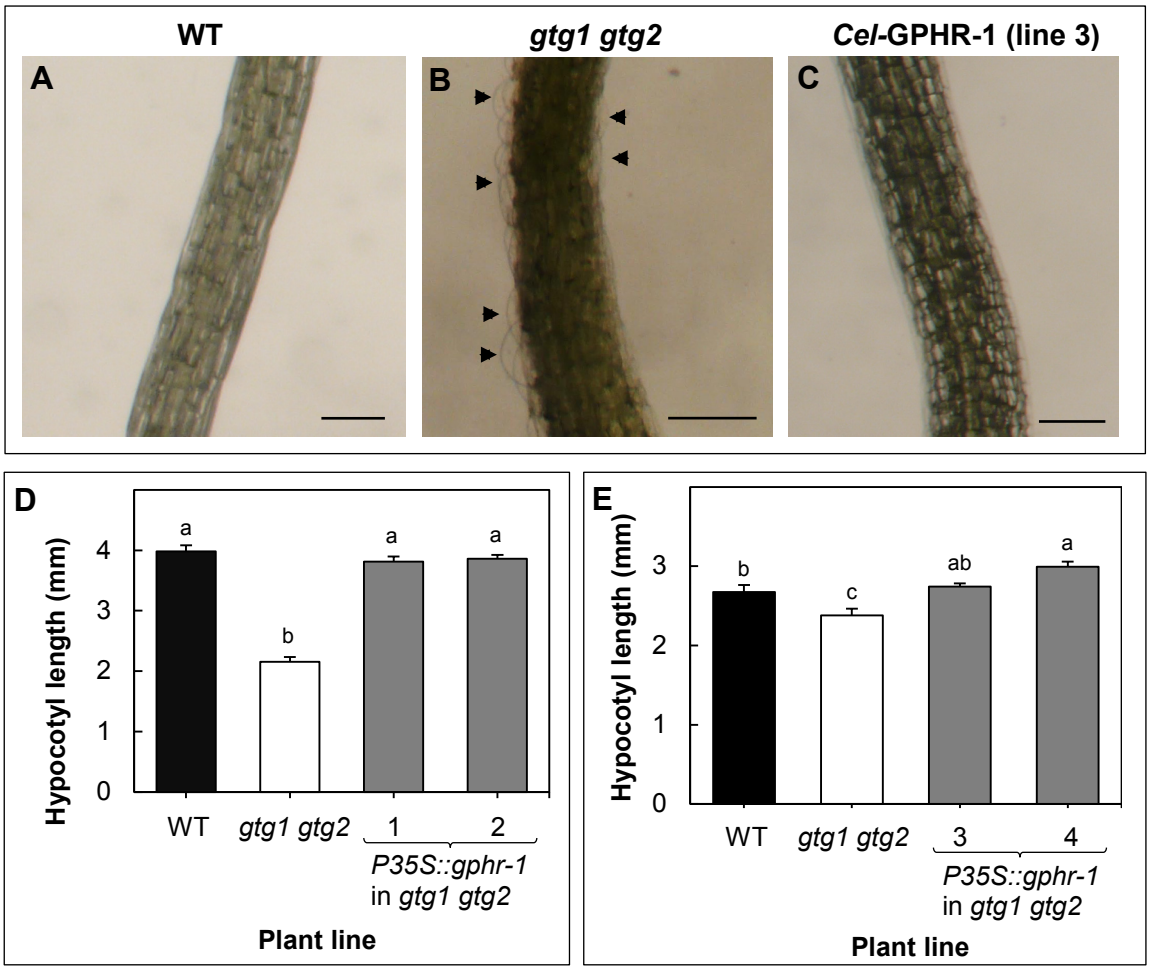


Figure 5.10 *Cel*-GPHR-1 functionally complements hypocotyl defects of the Arabidopsis *gtg1 gtg2* mutant

Hypocotyl growth of Arabidopsis *gtg1 gtg2* mutant expressing *Cel*-GPHR-1 on 0.5 MS containing 1 % (**A** – **D**) and 0 % (**E**) sucrose at 5 d. Short hypocotyl defect of *gtg1 gtg2* is exacerbated in 1 % sucrose (**D**) compared to 0 % sucrose (**E**). Data shows the mean hypocotyl lengths per seedling calculated for six plates (+S.E.) with 15 seedlings per plant line per plate, each plate containing four plant lines. Distended cell growth in *gtg1 gtg2* is rescued by *Cel*-GPHR-1 expression. Light microscopy showing hypocotyls of WT (**A**), *gtg1 gtg2* (**B**) and *gtg1 gtg2* expressing *Cel*-GPHR-1 (line 3) (**C**). Scale bars: 200 µm, arrowheads indicate ballooning cells. Mean hypocotyl length per seedling (+S.E.) in 1 % sucrose (**D**) and 0 % sucrose (**E**). There is a significant difference in hypocotyl length per seedling between plant lines; one-way ANOVAs: (**D**) 1 % sucrose ($F_{3, 20} = 11.57$, $p < 0.0001$) and (**E**) 0 % sucrose ($F_{3, 20} = 12.19$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* tests. Results represent one of four (**D**) and one of three (**E**) independent experiments using four independent *P35S::Cel-gphr-1* expressing lines where lines 1 and 2 (**D**) and 3 and 4 (**E**) are shown.

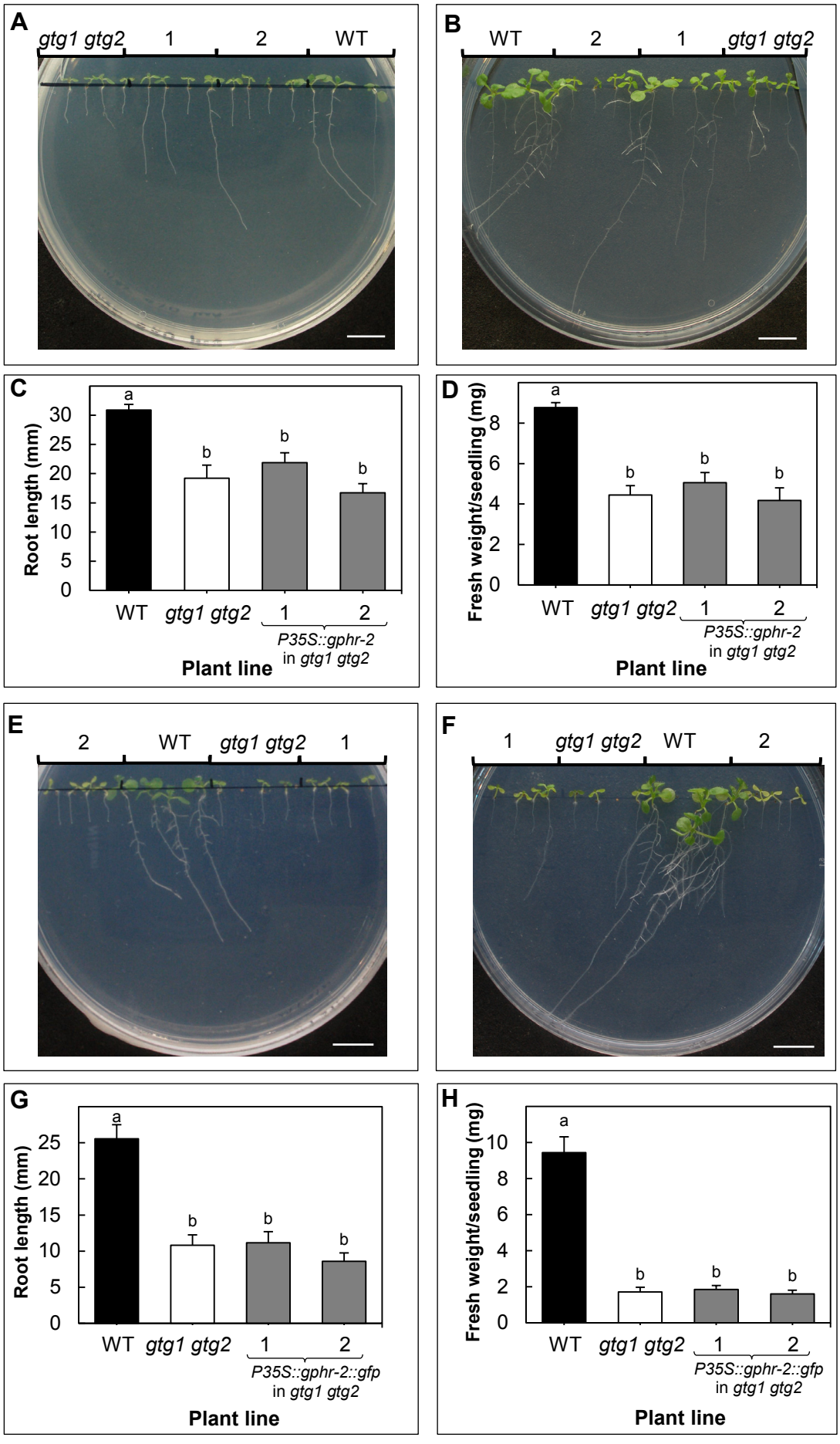

A
gtg1 gtg2
1
2
WT
B
WT
2
1
gtg1 gtg2
C
Root length (mm)
0
5
10
15
20
25
30
a
b
b
b
WT
gtg1 gtg2
1
2
P35S::gphr-2
in gtg1 gtg2
Plant line
D
Fresh weight/seedling (mg)
0
2
4
6
8
a
b
b
b
WT
gtg1 gtg2
1
2
P35S::gphr-2
in gtg1 gtg2
Plant line
E
2
WT
gtg1 gtg2
1
F
1
gtg1 gtg2
WT
2
G
Root length (mm)
0
5
10
15
20
25
a
b
b
b
WT
gtg1 gtg2
1
2
P35S::gphr-2::gfp
in gtg1 gtg2
Plant line
H
Fresh weight/seedling (mg)
0
2
4
6
8
10
a
b
b
b
WT
gtg1 gtg2
1
2
P35S::gphr-2::gfp
in gtg1 gtg2
Plant line

Figure 5.11 *Cel*-GPHR-2 does not rescue root growth and fresh weight defect in the Arabidopsis *gtg1 gtg2* mutant

Root growth and fresh weight per seedling of *gtg1 gtg2* expressing *Cel*-GPHR-2 on 0.5 MS containing 0 % sucrose. Representative images of seedlings at 10 d (**A** and **E**) and 14 d (**B** and **F**) (1 and 2 are *P35S::Cel-gphr-2* for **A** and **B** or *P35S::Cel-gphr-2::gfp* in *gtg1 gtg2* for **E** and **F**). Data shows the mean root length (**C** and **G**) and fresh weight per seedling (**D** and **H**) calculated for six (**C** and **D**) or eight (**G** and **H**) plates (+S.E.) with four seedlings per plant line per plate, each plate containing four plant lines. Results represent one of three experiments using three independent *P35S::Cel-gphr-2* or three independent *P35S::Cel-gphr-2::gfp* expressing lines where lines 1 and 2 are shown for both. **C** and **G**. Mean root length at 10 d (+S.E.). There is a significant difference in root lengths between plant lines; one-way ANOVAs **C**: ($F_{3, 20} = 13.88$, $p < 0.0001$), **G**: ($F_{3, 28} = 12.80$, $p < 0.0001$). **D** and **H**. Mean fresh weight per seedling at 14 d (+S.E.). There is a significant difference in fresh weight per seedling between plant lines **D**: ($F_{3, 20} = 13.63$, $p < 0.0001$), **H**: ($F_{3, 28} = 13.22$, $p < 0.0001$); one-way ANOVAs. Means not share a letter are significantly different, Tukey *post-hoc* test.

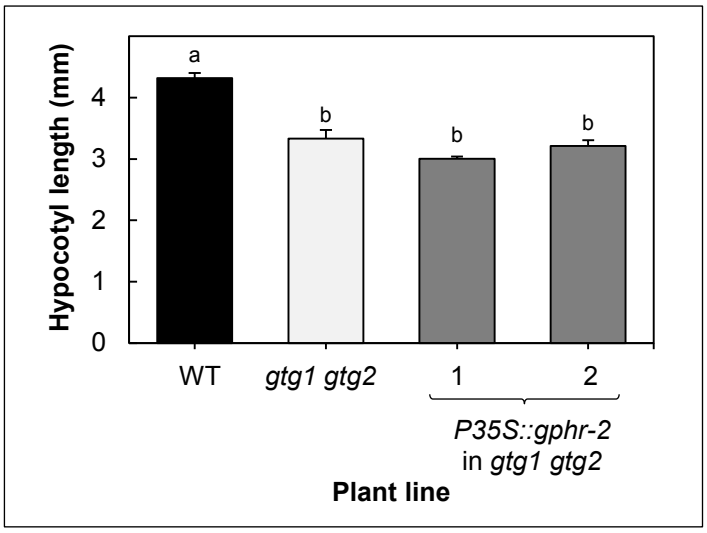


Figure 5.12 *Cel*-GPHR-2 does not functionally complement the hypocotyl defect of the Arabidopsis *gtg1 gtg2* mutant

Hypocotyl growth of Arabidopsis *gtg1 gtg2* expressing *Cel*-GPHR-2 on 0.5 MS containing 1 % sucrose at 5 d. The mean hypocotyl length per seedling (+S.E.) on 1 % sucrose. There is a significant difference in hypocotyl length per seedling between plant lines; one-way ANOVA ($F_{3, 20} = 11.69$, $p < 0.0001$). Means not sharing a letter are significantly different; Tukey *post-hoc* test. Data shows the mean hypocotyl length per seedling calculated for six plates (+S.E.) with 15 seedlings per plant line per plate, each plate containing four plant lines; results from one experiment using four independent *P35S::Cel-gphr-2* expressing lines (1 and 2 shown).

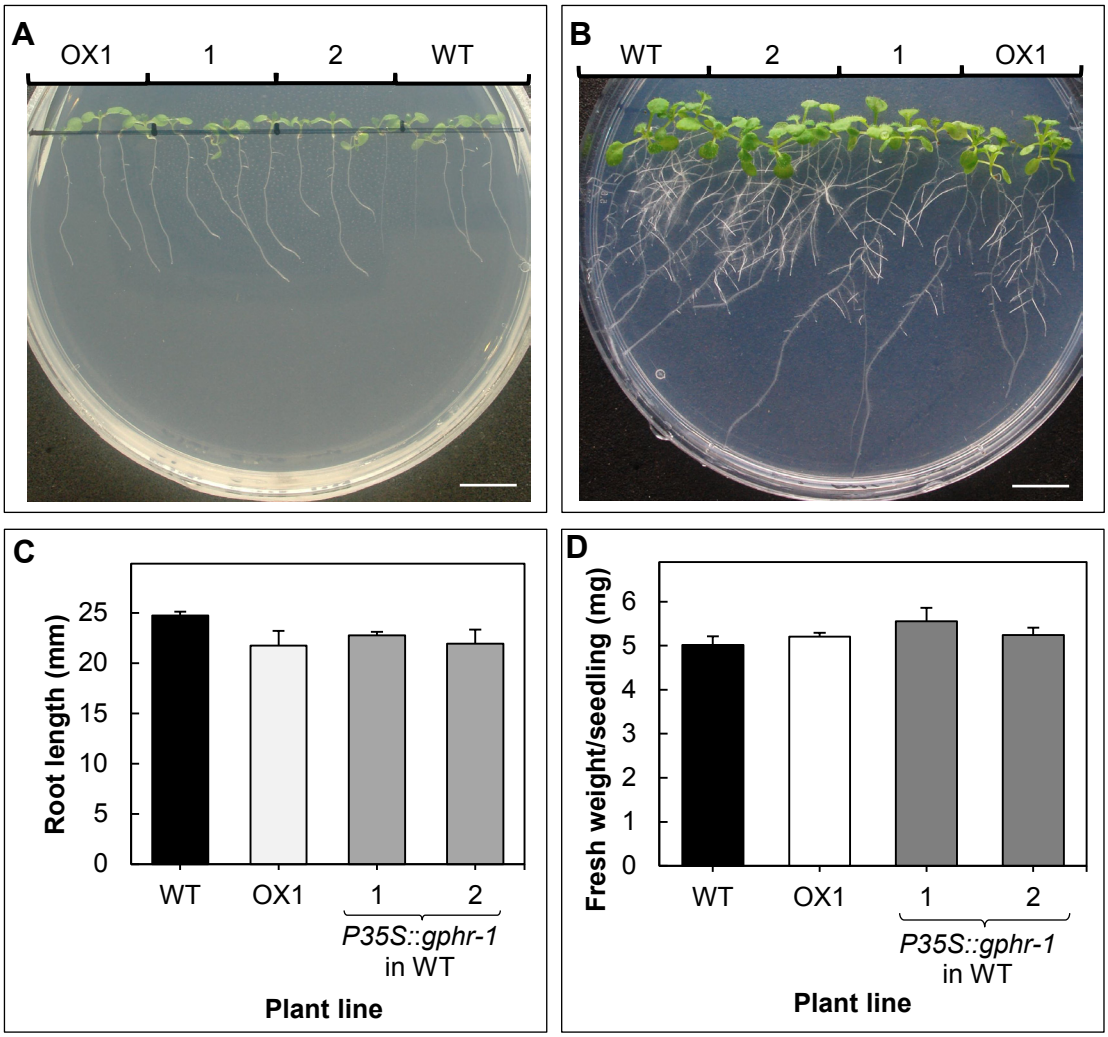


Figure 5.13 Overexpression of At GTG1 and expression of *Cel*-GPHR1 had minimal effect on root length

Root growth of WT expressing *Cel*-GPHR-1 or At GTG1 on 0.5 MS containing 0 % sucrose at 10 d. Data shows the mean root length calculated for six plates (+S.E.) with four seedlings per plant line per plate, each plate containing four plant lines. Results represent one of three independent experiments using four independent *P35S::Cel-gphr-1* expressing lines and 2 *P35S:At GTG1* overexpressing lines. OX = overexpression of At *GTG1* in WT; OX1 and OX2 represent two independent lines isolated previously (Jaffé et al., 2012). Representative images of seedlings at 10 d (**A**) and 14 d (**B**) (OX1 is WT expressing At GTG1, 1 and 2 are WT expressing *P35S::Cel-gphr-1*). **C**. Mean root length at 10 d (+S.E.) using lines 1 and 2. There is no significant difference in root length between plant lines; one-way ANOVA ($F_{3, 20} = 1.70$, $p = 0.199$). **D**. Mean fresh weight per seedling at 14 d (+S.E.). There is no significant difference in fresh weight per seedling between plant lines; one-way ANOVA ($F_{3, 20} = 1.17$, $p = 0.347$).

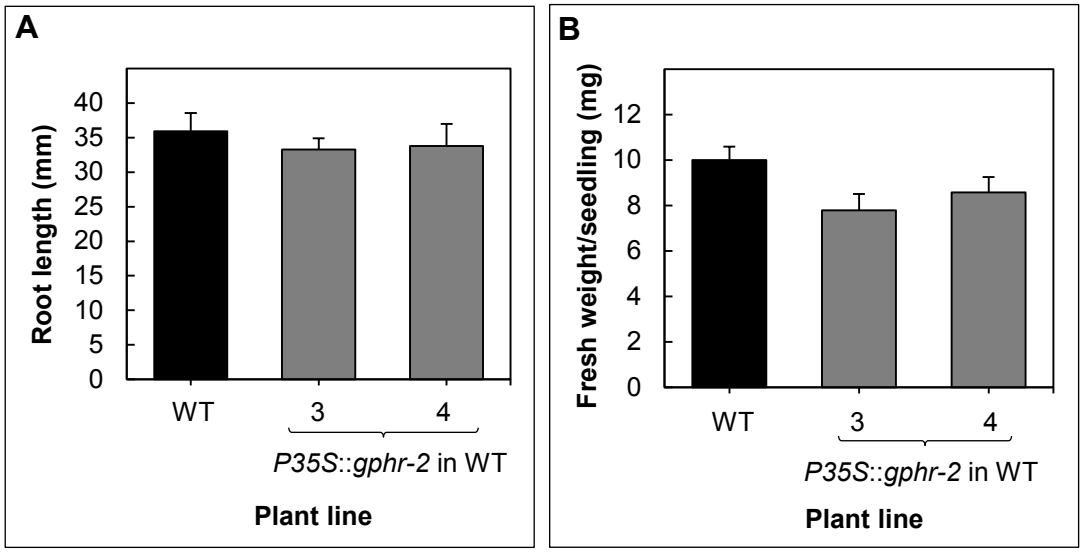


Figure 5.14 *Cel*-GPHR-2 has no effect on root growth and fresh weight of Arabidopsis WT

Root growth and fresh weight per seedling of Arabidopsis WT expressing *Cel*-GPHR-2 on 0.5 MS containing 0 % sucrose. Data shows the mean root length (**A**) and fresh weight per seedling (**B**) calculated for six plates (+S.E.) with four seedlings per plant line per plate, each plate containing four plant lines. Results represent one experiment using four independent *P35S::Cel-gphr-2* expressing lines (lines 3 and 4 shown). **A**. Mean root length at 10 d (+S.E.). There is a significant difference in root length between plant lines; one-way ANOVA ($F_{2, 15} = 0.31$, $p < 0.736$). **B**. Mean fresh weight per seedling at 14 d (+S.E.). There is a significant difference in fresh weight per seedling between plant lines; one-way ANOVA ($F_{2, 15} = 3.42$, $p < 0.060$).

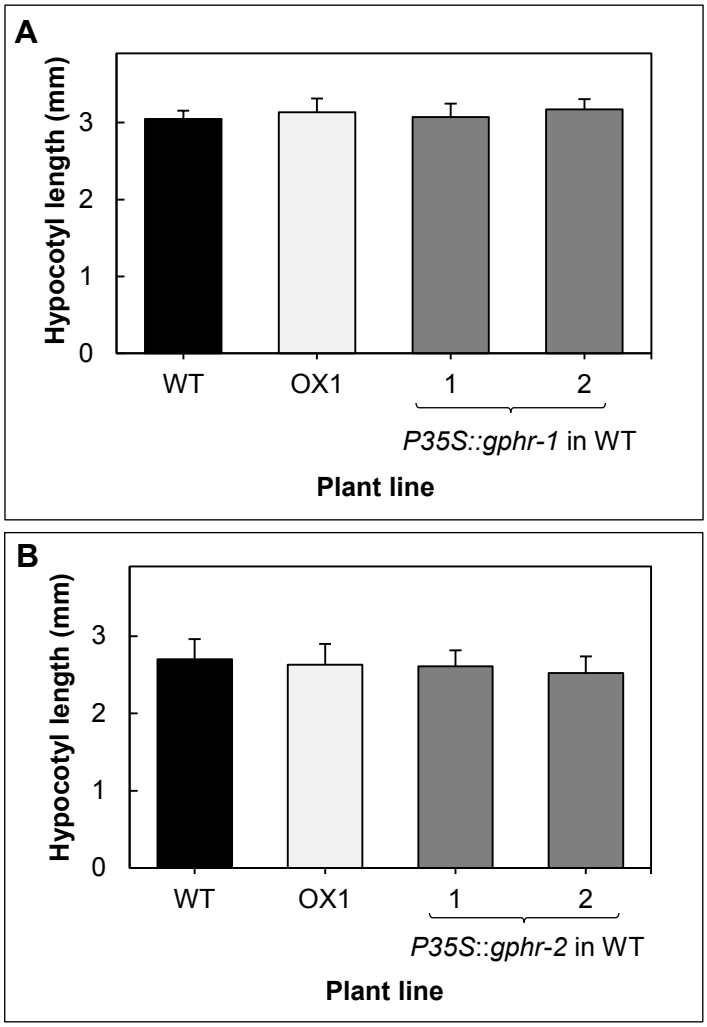


Figure 5.15 *Cel*-GPHR had no effect on Arabidopsis WT hypocotyl lengths

Hypocotyl growth of Arabidopsis WT expressing *Cel*-GPHR-1 or *Cel*-GPHR-*2* on 0.5 MS containing 1 % sucrose at 5 d. **A**. Mean hypocotyl length per seedling (+S.E.) of WT lines expressing *Cel*-GPHR-1 and At GTG1 in 1 % sucrose, calculated for six plates (+S.E.) with 15 seedlings per plant line per plate, each plate containing four plant lines. There is no significant difference in hypocotyl length per seedling between plant lines; one-way ANOVA ($F_{3, 20} = 0.14$, $p = 0.936$). Results represent one of three independent experiments using four independent *P35S::Cel-gphr-1* expressing lines (lines 1 and 2 shown) and two independent *P35S:At GTG1* overexpressing isolated previously (Jaffé et al., 2012); OX1 shown. **B**. Mean hypocotyl length per seedling (+S.E.) of WT lines expressing *Cel*-GPHR-2 and At GTG2 (+S.E.) in 1 % sucrose, calculated for six plates (+S.E.) with 15 seedlings per plant line per plate, each plate containing four plant lines. There is no significant difference in hypocotyl length per seedling between plant lines; one-way ANOVA ($F_{3, 20} = 0.09$, $p = 0.965$). Results from one experiment using four independent *P35S::Cel-gphr-2* expressing lines (lines 1 and 2 shown) and two independent *P35S:At GTG2* overexpressing lines isolated previously (Jaffé et al., 2012); OX1 shown.

consistently cause a difference in root growth and fresh weight (Figure 5.13 and Figure 5.14), or hypocotyl growth (Figure 5.15). These results suggest there is no additional function with WT *GTG/GPHR* genes.

5.2.6 The localisation of the *Cel*-GPHRs in Arabidopsis

Live-cell confocal imaging was carried out in the root tips of segregating (T_2) *gtg1 gtg2* mutant lines expressing *P35S::Cel-gphr-1::gfp* or *P35S::Cel-gphr-2::gfp*. Preliminary results show that both *Cel*-GPHRs localise intracellularly to punctate structures (Figure 5.16 and Figure 5.17). The expression of both *Cel*-GPHR-1 and *Cel*-GPHR-2 appears quite disperse throughout the cells (Figure 5.16 and Figure 5.17). Figure 5.16 and Figure 5.17 represent the localisation data obtained from one independent line for each *Cel*-GPHR-1 and *Cel*-GPHR-2. Further lines for testing are therefore necessary, and a more detailed localisation study carried out in order to fully characterise the localisation of the *Cel*-GPHRs in Arabidopsis.

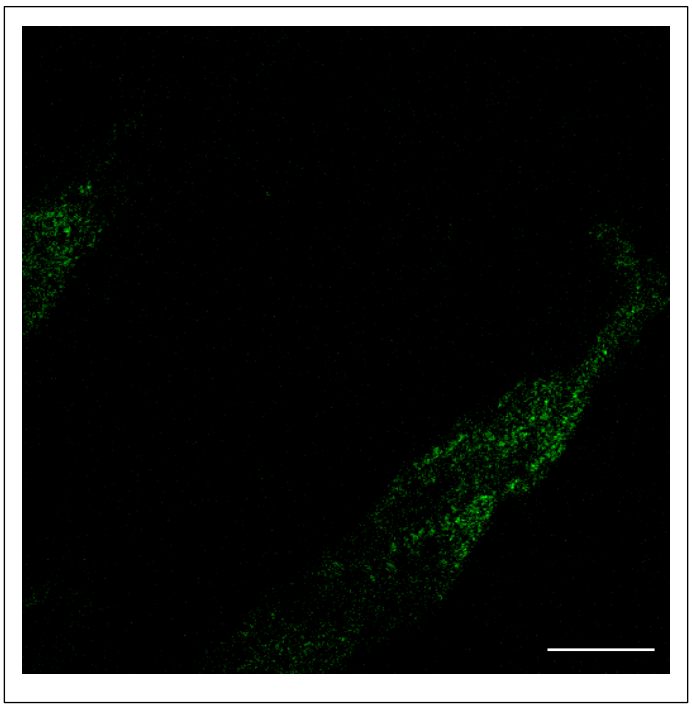


Figure 5.16 *Cel-gphr-1::gfp* localises intracellularly in *gtg1 gtg2*

Confocal microscopy to show *Cel-gphr-1::gfp* expression in the root tip. *Cel*-GPHR-1 localises intracellularly to punctate structures. Image is representative of two seedlings from one segregating (T_2) line expressing *Cel-gphr-1::gfp* in *gtg1 gtg2*. Scale bar: 10 μm.

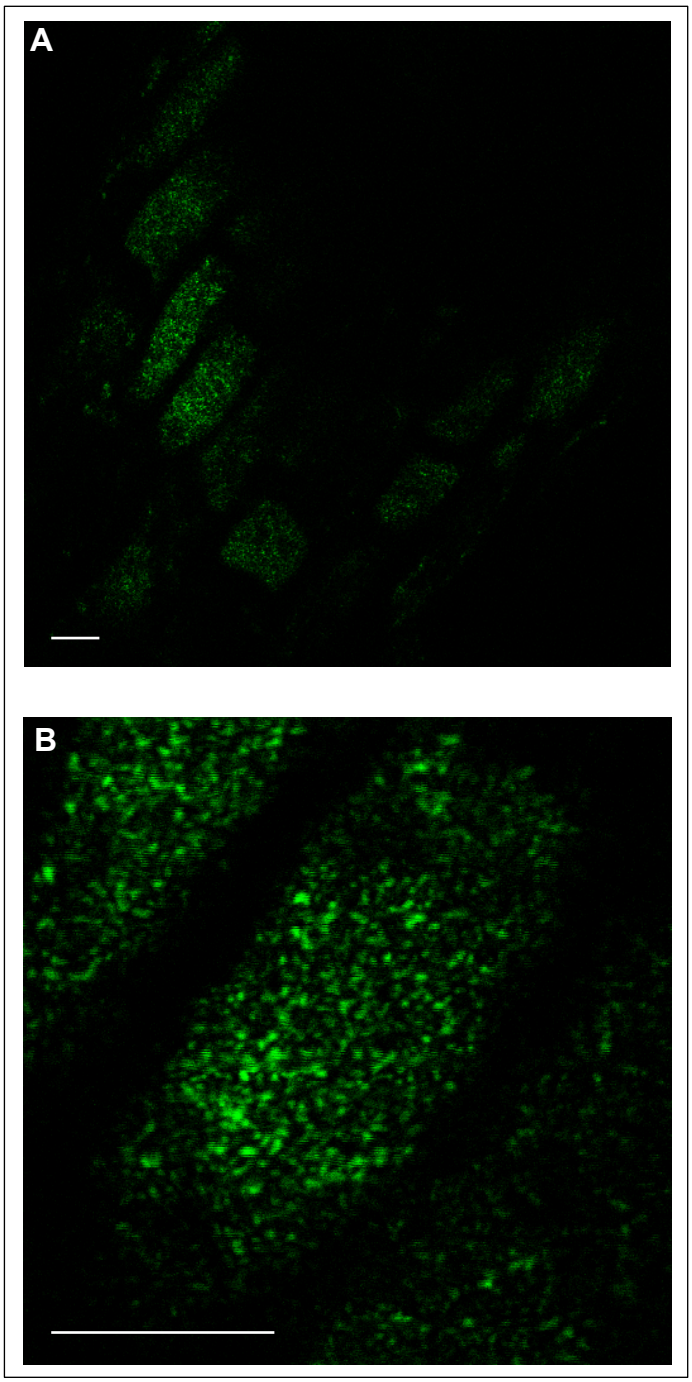


Figure 5.17 *Cel-gphr-2::gfp* localises intracellularly in *gtg1 gtg2*

Confocal microscopy to show *Cel-gphr-2::gfp* expression in the root tip. *Cel*-GPHR-2 localises intracellularly to punctate structures. Image is representative of three seedlings from one segregating (T_2) line expressing *Cel-gphr-2::gfp* in *gtg1 gtg2*. Images show one section of the root tip, scale bars: 10 µm. **A**. *Cel-gphr-2::gfp* expression in a few cells. **B**. Magnified area of (**A**), centred on one cell.

Table 5.5 Percentage identities and similarities between the conserved GTG/GPHR regions using Arabidopsis and *C. elegans* sequences

The DUF3735 and ABA-GPCR domains and intracellular region were predicted by the Pfam database (Bateman et al., 2004). These regions were compared using the ClustalW algorithm (Thompson et al., 2002a, Li, 2003), at the amino acid level. Percentage identities shown in black, percentage similaries shown in light grey.

Predicted region	**Sequence**	**At GTG2**	***Cel*-GPHR-1**	***Cel*-GPHR-2**
DUF3735	At GTG1	90 97	47 74	52 78
	At GTG2		47 76	52 83
	Cel-GPHR-1			76 85
Intracellular segment between TM5 and TM6	At GTG1	79 86	22 38	19 38
	At GTG2		25 43	23 45
	Cel-GPHR-1			52 69
ABA-GPCR	At GTG1	96 98	43 69	43 69
	At GTG2		43 69	44 68
	Cel-GPHR-1			78 89

5.3 Discussion

5.3.1 Generation of transgenic Arabidopsis expressing *C. elegans* GPHRs

The CDSs of the *Cel-gphr* genes (Figure 5.1) were inserted into the TOPO entry vector, pENTR/D (Table 5.2). These entry vectors were used in LR recombination reactions to insert the cloned genes of interest into different destination vectors for subsequent transformation. In this chapter, the CDSs of the *Cel-gphr* genes with and without stop codons were inserted into the Arabidopsis destination vectors pMDC32 and pMDC83 respectively. The *P35S* drives constitutive expression of the *Cel-gphr* genes in pMDC32 and pMDC83, and the latter tags the *Cel*-GPHR at the C-terminus with GFP. These entry vectors (Table 5.2) were also used for inserting the *Cel-gphr* CDSs into destination vectors for different organisms (see sections 7.2.4 and 7.3.3). Restriction digestion was used to confirm correct insertion of the CDSs into entry and destination vectors and sequencing results confirmed that the CDSs corresponded to those supplied by WormBase (Figure 5.2 and Figure 5.3). The expression vectors were transformed into *Agrobacterium* for Arabidopsis transformation and colony PCR results indicated that the plasmids of interest were successfully transformed into *Agrobacterium* (Figure 5.4). The expression vectors in Table 5.3 were transformed into Arabidopsis WT and *gtg1 gtg2*, and for each, several independent T_3 lines were isolated using the method illustrated in Figure 2.1 (Table 5.4). RNA was extracted from the independent T_3 lines for cDNA synthesis. Figure 5.5 and Figure 5.6 are examples of genotyping by PCR. The results showed that *Cel-gphr* transcripts were being expressed in the T_3 lines isolated.

5.3.2 *Cel*-GPHR-1 functionally complements the Arabidopsis *gtg1 gtg2* mutant

Arabidopsis *gtg1 gtg2* mutants exhibit a number of growth and developmental defects. The loss of both At *GTG* genes results in reduced root and hypocotyl growth as well as distended cell structure seen in the hypocotyl epidermal cells (Jaffé et al., 2012). The expression of either At *GTG1* or At *GTG2* can restore normal root, hypocotyl and cellular growth implying that the phenotypes were characteristic of *gtg1 gtg2* mutants (Jaffé et al., 2012). To investigate whether the functions between orthologues are conserved, functional complementation assays can be carried out. Heterologous expression of orthologue proteins in plants has been used to confirm gene functionality

(section 4.1.1). Transgenic Arabidopsis were generated to investigate whether GTG/GPHR proteins from another organism were able to functionally complement the loss of both At GTGs, as this would suggest they have similar roles. *Cel*-GPHR-1 was able to restore normal root, fresh weight, hypocotyl and cellular growth to the *gtg1 gtg2* mutant using four independent *Cel-gphr-1* expressing lines (Figure 5.8 and Figure 5.10). This indicates that the *Cel*-GPHR-1 protein with 39 % protein identity to the At GTGs is able to function in a similar way to both At GTG1 and At GTG2. The root growth and fresh weight of the *gtg1 gtg2* mutant can be restored to normal on sucrose and *Cel*-GPHR-1 expression had no additive growth effect (Figure 5.9).

The At GTGs and *Cel*-GPHRs were expressed under constitutive *P35S* to explore the effects of overexpressing the GTG/GPHRs in plants. The expression of either At GTG1, *Cel*-GPHR-1 or *Cel*-GPHR-2 in WT Arabidopsis produced similar root, hypocotyl and fresh weight growth to that of the WT (Figure 5.13). These results suggest that the At GTGs and *Cel*-GPHR-1 functionally complement the loss of both At GTGs rather than independently stimulate growth and development, especially since the *P35S* is can be more highly active in transgenic plants compared to native promoters. *P35S* has been used to overexpress proteins in WT to observe effects. For example, AGB1 is the Arabidopsis Gβ subunit of G proteins and has a role in K^+ ion channel regulation. AGB1 overexpression lines maintained a WT phenotype; ABA inhibition of K_{in} current was unaltered in AGB1-overexpression lines (Fan et al., 2008).

Overexpression of proteins can be used for a variety of functional studies. Overexpression can reveal that a correct level of gene expression is needed for the required response. For example, overexpression of the Arabidopsis glutamate receptor, At GluR2 caused reduced efficiency of calcium utilisation in transgenic plants (Kim et al., 2001). Alternatively, the overexpression of a protein can further promote a response indicating its role in a specific pathway. For example, a method has been developed to directly follow cytosolic Ca^{2+} concentrations using *Aequorin* luminescence in transgenic plants (Knight et al., 1991). An increase in sucrose-induced luminescence was produced by overexpressing a predicted voltage-gated Ca^{2+} channel, At TPC1 (Furuichi et al., 2001a, Furuichi et al., 2001b). This result implied that the putative At TPC1 channel was indeed a Ca^{2+} permeable channel. The expression of the Arabidopsis or *C. elegans* GTG/GPHRs does not have an additive growth effect suggesting that the root growth

could have been saturated and could not be further stimulated by combined expression of Arabidopsis and *C. elegans* GTG/GPHRs.

Roots appear to skew to the left in Figure 5.8, Figure 5.9, Figure 5.13 and Figure 5.14 (when looking at the roots from an open plate such as in Figure 5.8 B). Plants are capable of various forms of movement and frequently plant organs move in response to environmental cues such as gravity, light, touch and heat (Bell and Charlwood, 1980). In roots, for instance, the columella cells in the root cap perceive gravity and the elongation zone steers the growth response towards this stimulus (Blancaflor and Masson, 2003). When grown vertically on agar, Arabidopsis primary roots show a tendency to diverge from a stringent gravitational direction (Simmons et al., 1995); this is described as root skewing, which is different to root waving (Oliva and Dunand, 2007). This phenomena occurs in a variety of inclined angles in Arabidopsis roots of the Ws ecotype (Simmons et al., 1995). In fact, this occurence is not limited to this ecotype and other Arabidopsis accessions appear to skew different amounts. For example, the Ws and Landsberg ecotypes exhibit a larger deviation than the Col ecotype (Rutherford and Masson, 1996). Since Ws plant lines were used, this explains the root skewing seen in the root assays (Figure 5.8, Figure 5.9, Figure 5.13 and Figure 5.14). In addition, sucrose was shown to enhance root skewing compared when absent from growth media (Murashige and Skoog, 1962). This is in accordance with the results where root skewing was increased in the presence of sucrose (Figure 5.9). Furthermore, both transgenic and WT lines exhibited root skewing so these events were due to the ecotype (Ws), nutrient content and vertical growth.

5.3.3 *Cel*-GPHR-2 does not appear to functionally complement the Arabidopsis *gtg1 gtg2* mutant

Neither *Cel-gphr-2* nor *Cel-gphr-2::gfp* expression appear to restore any of the *gtg1 gtg2* mutant phenotypes indicating it cannot functionally complement the At GTGs (Figure 5.11 – Figure 5.12). This is unexpected since *Cel-gphr-1* and *Cel-gphr-1::gfp* was able to rescue *gtg1 gtg2* phenotypes and both *Cel*-GPHRs share the same percentage identity (39 %) to both At GTG1 and At GTG2 at the amino acid level; however, the *Cel*-GPHRs share only 69 % identity to one another in sequence (Table 3.2). Furthermore, the conserved regions/domains identified by *in silico* analysis (section 3.2.3) were compared between both Arabidopsis and *C. elegans* GTG/GPHR

sequences (Table 5.5). These results indicate that both *Cel*-GPHR1 and *Cel*-GPHR-2 share similar identity and similarity to both At GTGs. Gene silencing can occur transcriptionally and post-transcriptionally. The mechanism of the latter and RNA interference is discussed in section 6.1.3.2. However, since the four *gtg1 gtg2* expressing *P35S::Cel-gphr-2* lines show products for full-length *Cel-gphr-2* transcipt expression (Figure 5.6 D), neither transcriptional or post-transcriptional gene silencing is likely to be occurring. Furthermore, for *Cel-gphr-1* there appeared to be two products near the predicted size for *Cel-gphr-1* (Figure 5.6), which could have resulted from a splicing event in *gtg1 gtg2*. The presence of two products was not seen for *Cel-gphr-2* in *gtg1 gtg2* (Figure 5.6 D), which could be a reason for the lack of functional complementation however it is not clear what these two products are this stage. Furthermore, translation and downstream events could have been disrupted leading to a lack of *Cel*-GPHR-2 protein production. Therefore, testing the production of *Cel*-GPHR-2 protein would be required by either immunostaining or fluorescent microscopy using the T_3 lines expressing the pMDC83 *Cel-gphr-2* construct. Expression of *Cel-gphr-2* is shown at the T_2 level (Figure 5.17) indicating there is expression in *gtg1 gtg2*. In contrast to *Cel*-GPHR-1*, Cel*-GPHR-2 does not rescue the Arabidopsis double mutant phenotypes suggesting the *Cel*-GPHR homologues may have alternative roles (however refer to CHAPTER 6 where the role of *Cel*-GPHRs are explored), unlike the Arabidopsis GTGs which appear to function redundantly.

5.3.4 The localisation the *Cel*-GPHRs in Arabidopsis

The native localisation of the *Cel*-GPHRs is discussed in section 6.3.5. Here we show the localisation of the *Cel*-GPHRs in Arabidopsis. Figure 5.16 and Figure 5.17 show that the *Cel*-GPHRs localise to punctate structures. However, without colocalisation markers, it is difficult to conclude to which organelles these proteins are localised. The expression of the *Cel*-GPHRs appears to be intracellular given the size of the areas hypothesised to be cells (Figure 5.16 and Figure 5.17). However again, without a stain such as FM4-64 to label the plasma membrane it is difficult to conclude which regions correspond to single cells. However, since the *Cel*-GPHR-1 is able to functionally complement *gtg1 gtg2*; it would be prudent to therefore anticipate that *Cel*-GPHR-1 localises to the same site as At GTGs. Since the At GTGs are shown to be Golgi and ER localised (Jaffé et al., 2012), it is predicted that the *Cel*-GPHRs also localise to these organelles but further work is required on these lines.

CHAPTER 6

Isolation and characterisation of the *C. elegans gphr* single and double mutants

6.1 Introduction

6.1.1 *C. elegans* as a model organism for analysing *gphr* function

C. elegans is a commonly used animal model organism (introduced in section 1.5.4). The GTG/GPHR proteins are conserved in eukaryotes and their function is yet to be fully characterised (Maeda et al., 2008, Pandey et al., 2009, Jaffé et al., 2012). Double *gtg1 gtg2* KO models have been created in Arabidopsis (Pandey et al., 2009, Jaffé et al., 2012), a CHO animal cell line (Maeda et al., 2008), and very recently, *Drosophila* (Charroux and Royet, 2014). A skin-specific KO of the *GPHR* in mice has also been created (Tarutani et al., 2012). Comparing KO models from two different kingdoms may allow us to relate any parallel functions of the GTG/GPHRs. Most eukaryotic organisms possess one copy of the *GTG/GPHR* gene, such as *Drosophila*; however, there are some organisms, like Arabidopsis and *C. elegans*, that have two *GTG/GPHR* genes. So far, phenotypic analyses of the single *GTG* mutants in Arabidopsis have shown no marked differences when compared to the WT. However knocking out both *GTGs* in Arabidopsis caused a range of abnormal phenotypes indicating Arabidopsis may function without either one (*GTG1* or *GTG2*) of these genes but not both, suggesting gene functional redundancy (Pandey et al., 2009, Jaffé et al., 2012). In light of this, it would be innovative to explore the functions of the GPHRs in *C. elegans*, not only in single mutants but also by generating a *C. elegans* double KO model. Would the *C. elegans* double *gphr-1 gphr-2* mutant phenotype(s) reveal any functions of these proteins? In addition, gene functional redundancy can be explored as well as dominant-recessive relationships. Only the *Cel-gphr* genes will be discussed in this chapter, therefore both *Cel-gphr-1* and *Cel-gphr-2* will be referred to as *gphr-1* and *gphr-2*, without the species-specific "*Cel-*" prefix.

6.1.2 *C. elegans* genetics

C. elegans possess five pairs of autosomes (I, II, III, IV and V) and a pair of sex chromosomes (XX) (Consortium, 1998). There are two *C. elegans* sexes: a self-fertilising hermaphrodite (matched pair of sex chromosomes, XX) (Figure 6.1 B) and a male (only one sex chromosome, XO) (Figure 6.1 A) (Altun, 2009). The tail of *C. elegans* can be used to differentiate the two *C. elegans* sexes; the males have a distinctive tail shape unlike the tapered tail of the hermaphrodite (Figure 6.1). Reproduction by self-fertilisation produces progeny that are genetically identical i.e. clones (Altun, 2009). The male sex of *C. elegans* arises infrequently (0.01 – 0.02 %) as a result of spontaneous, X chromosome non-disjunction (Ward and Carrel, 1979, Hodgkin and Doniach, 1997). Males are able to mate with hermaphrodites by cross-fertilisation (crossing), allowing genetic exchange between strains thus mediating isolation and maintenance of mutant strains (Altun, 2009). X chromosome non-disjunction can be induced by heatshock (Wood, 1988) and the male line can be maintained by subsequent crossing.

6.1.3 Reverse genetics to explore *gphr* gene function in *C. elegans*

In reverse genetics, functional study of a gene involves observation of phenotypes associated with a gene sequence, and is an important approach which complements forward genetics (Ahringer, 2006). Two methods employed for *C. elegans* reverse genetics are deletion mutants and RNA interference (RNAi).

6.1.3.1 *C. elegans* deletion mutants

C. elegans deletion mutant strains can be generated using ultraviolet/4,5´,8-trimethylpsoralen (UV/TMP) (Yandell et al., 1994). TMP causes single-base changes and deletions under UV exposure (Cimino et al., 1985) by producing interstrand crosslinks and monoadducts (Piette et al., 1985, Sladek et al., 1989). The mutagenic spectrum of TMP was first studied in *C. elegans* in which 21 *unc-22* mutants and two *pal-1* deletion mutants were generated using this reagent (Yandell et al., 1994). Today, UV/TMP is a potent tool in the *C. elegans* KO consortium laboratories (Barstead and Moerman, 2006).

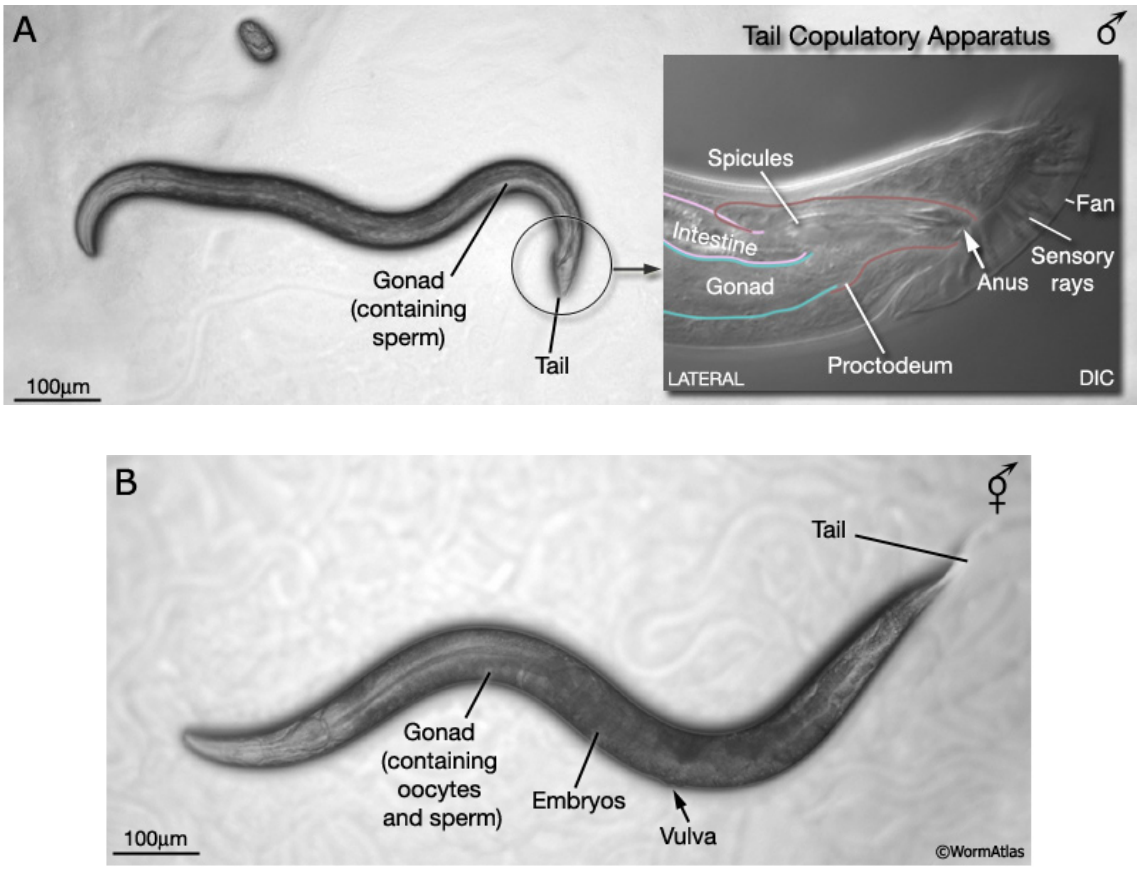


Figure 6.1 Light microscopy images showing male and hermaphrodite anatomy

Images of the two *C. elegans* sexes. **A**. Male *C. elegans*. Inset shows a magnified view of the tail region by Nomarski DIC. **B**. Hermaphrodite *C. elegans*. Figure taken from WormAtlas (Altun, 2009).

Deletion mutants can be searched for and identified using *C. elegans* mutant databases: *Caenorhabditis* Genetics Center (CGC), funded by NIH Office of Research Infrastructure Programs (P40 OD010440; http://www.cgc.cbs.umn.edu/), and the National Bioresource Project (NBP) for the Experimental Animal "Nematode *C. elegans*" (http://www.shigen.nig.ac.jp/c.elegans/). Deletion mutants from these consortiums are isolated using UV/TMP and so background mutations may have occurred during mutagenesis (Piette et al., 1985, Sladek et al., 1989, Yandell et al., 1994). It is therefore highly recommended to backcross any mutant strains with the WT strain, N2, before conducting experiments.

Deletion mutants are widely used in genetic and biochemical analyses to determine the molecular mechanisms underlying biological processes. Characterising mutant phenotypes contributes to understanding how the gene works in *C. elegans*. The egg-laying behaviour, for instance, can be used to examine a variety of fundamental processes in neuronal cell biology and signal transduction such as G protein signalling (Segalat et al., 1995, Brundage et al., 1996). Egg-laying has a well-characterised anatomy (White et al., 1986) involving 16 vulval and uterine muscle cells which are innervated by a pair of serotonergic motor neurons, the hermaphrodite-specific neurons (HSNs) (Desai et al., 1988). The *egl-10* gene, for instance, is one of many identified where egg-laying behaviour in the mutant is severely reduced or abolished (Koelle and Horvitz, 1996). Phenotypic analysis of egg-laying behaviour has consequently been employed to elucidate a number of neuronal pathways in *C. elegans*. With phenotypic information, mutants can also be isolated which enhance or suppress a particular phenotype. Using the egg-laying behaviour as an example, mutations in the *goa-1* gene increase the rate of egg-laying as well as promote a higher frequency of locomotory body bends and other behaviours (Lochrie et al., 1991, Mendel et al., 1995, Segalat et al., 1995). In this way, a whole network of genes has been revealed linked to the same genetic pathway (Patterson and Padgett, 2000).

Sometimes phenotypes are not apparent unless mutations eliminate the activity of two or more loci (Thomas, 1993). Functional redundancy can occur when gene duplication resulting in two or more closely related genes where any one gene can carry out a particular function. For example, the two genes, *end-1* and *end-3*, encoding for GATA-transcription factor homologues, demonstrate gene redundancy. Loss of one of these genes alone causes no phenotype but loss of both however, leads to arrested

embryos lacking an endoderm (Zhu et al., 1997). In addition, transgenic *C. elegans* can be generated for phenotypic rescue experiments. For example, WT or mutant DNA can be transformed into *C. elegans* to test whether they are functional *in vivo* in the nematode; the capacity to rescue phenotypes indicate functional complementation. For all these reasons, *C. elegans* deletion mutants are very useful biological tools for giving insight into gene function.

Using the deletion mutant databases CGC and NBP, *C. elegans gphr* mutant strains can be searched for to confirm their availability. If these mutants are available, they must be backcrossed to remove any background mutations that may have occurred during mutagenesis. Phenotypic analyses and generation of a double mutant whereby both *gphr-1* and *gphr-2* are knocked out in the genome will be carried out. Thus, the double mutant can be studied to determine whether the phenotype reveals anything about the function of the GPHR proteins in *C. elegans*.

6.1.3.2 Gene silencing by RNAi

Sometimes, deletion mutants are not available or there may be complications such as viability. Another approach that could be used to study the functions of the GTG/GPHR family of proteins is RNAi. In 1998 the discovery that double-stranded RNA (dsRNA) could be used to produce interference led to the subsequent employment of RNAi for the manipulation of gene expression in *C. elegans* (Fire et al., 1998). RNAi is an intracellular process by which dsRNA molecules are used to target messenger RNAs (mRNAs) for degradation leading to gene silencing; a process which occurs post-transcriptionally (Fire et al., 1998, Timmons and Fire, 1998). RNA duplexes approximately 20 bps in length are able to suppress gene expression in animals (Elbashir et al., 2001) and in plants (Waterhouse et al., 1998).

The phenomenon is initiated by the cleavage of endogenous dsRNA by an RNAase III nuclease, Dicer (Bernstein et al., 2001). In *C. elegans* RDE-4 detects and binds to exogenous dsRNA and delivers it to Dicer (Parker et al., 2006). Dicer processes long dsRNA by cleaving the duplexes into short interfering RNA (siRNA) approximately 25 bps in length (Bernstein et al., 2001). In some systems, for instance in plants such as Arabidopsis (Axtell et al., 2006), nematodes such as *C. elegans* (Pak and Fire, 2007, Sijen et al., 2007), and fungi such as *Neurospora crassa* (Cogoni and Macino, 1999), an RNA-dependent RNA polymerase (RdRP) also functions to amplify

the primary siRNA and synthesise a population of secondary siRNAs (Figure 6.2) (Baulcombe, 2007). The (secondary) siRNAs are incorporated into a large protein, RNA-induced silencing complex (RISC). RISC contains the Argonaute2 protein which uses one strand of the siRNA duplex to identify complementary sequences of messenger RNAs (mRNAs). Once complementary base-pairing occurs, Argonaute2 cleaves the phosphodiester backbone at a specific site determined from the 5′ end of the single-stranded siRNA (Rand et al., 2005, Sen and Blau, 2005). The cleavage event is essential for removing the anti-guide strand from Argonaute2 to activate RISC leading to the degradation of the mRNA, and in turn resulting in gene silencing by inhibiting mRNA translation (Hammond et al., 2000, Parrish et al., 2000, Rand et al., 2005). Therefore, RNAi by dsRNA acts to degrade mRNA prior to translation, i.e. dsRNA exerts its effect at the post-transcriptional level (Montgomery et al., 1998).

Using RNAi to silence the *gphr* genes in *C. elegans* would be a useful method to confirm double mutant phenotypes. For example, would we see similar phenotypes if both *gphr* genes were knocked out in *C. elegans* compared to if they were knocked down? Phenocopying *gphr-1 gphr-2* double deletion mutants using RNAi would indicate that these observations are characteristic of the loss of both GPHR proteins.

6.1.4 Localisation of the GPHR proteins in *C. elegans*

Studying protein localisation and the current status of what is known regarding this information on the GTG/GPHRs is introduced in section 1.4.2. GUS straining was used to show the At *GTG* genes are widely expressed (discussed in section 1.4.1). To date, the tissue localisation of these proteins is yet to be explored experimentally in an animal system, although is predicted to be ubiquitously expressed in *Drosophila* according to FlyAtlas (http://flyatlas.org/) (Charroux and Royet, 2014) and in humans according to SymAtlast (http://symatlas.gnf.org/SymAtlas/) (Maeda et al., 2008). GFP can be placed under the control of either *gphr* promoters to demonstrate the tissue localisation of the GPHRs in *C. elegans*. Will *Cel-gphr* genes be widely expressed in most tissues like the At *GTG* genes?

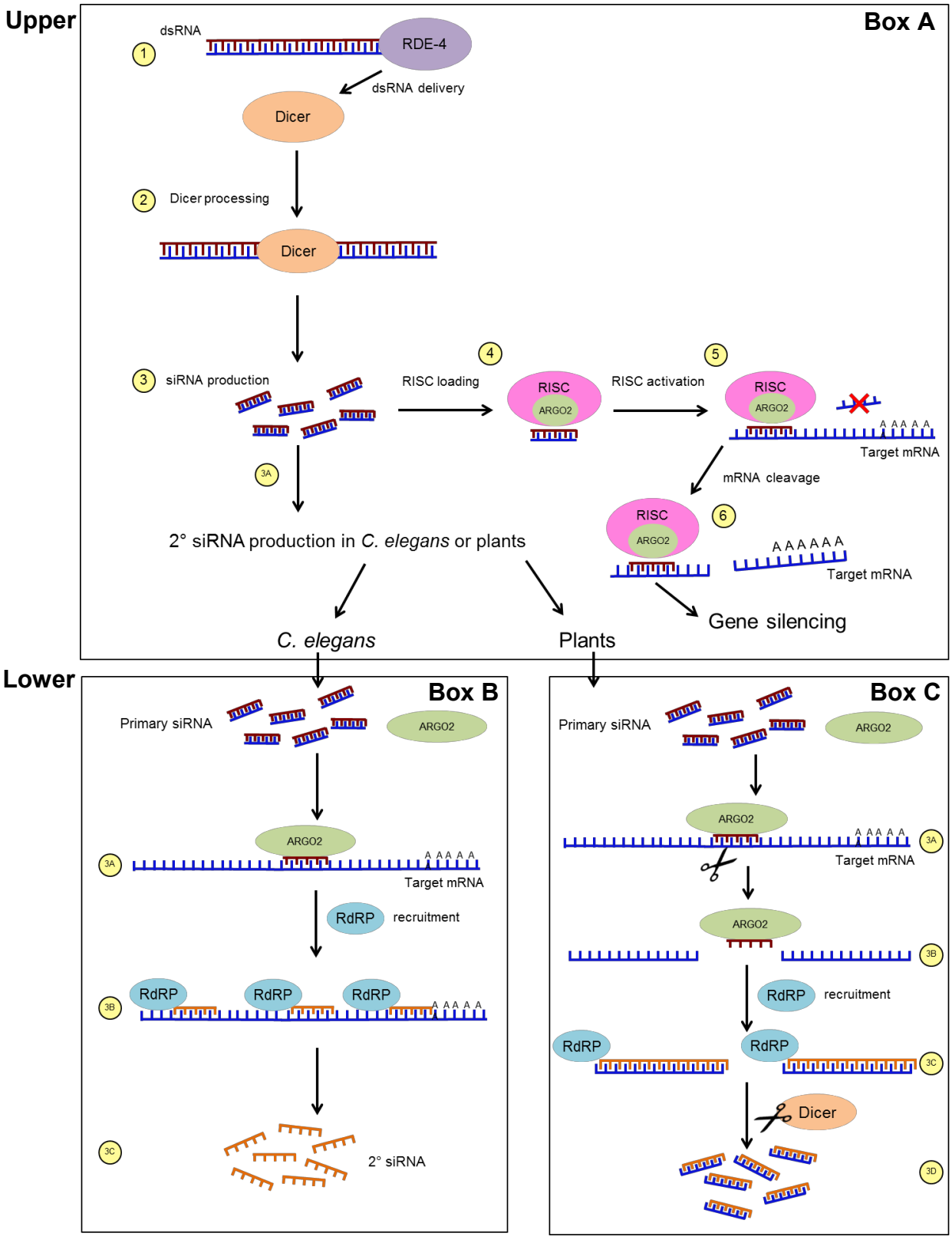


Figure 6.2 Schematic of the RNAi mode of action

Upper: **Box A**. Diagram showing the processes involved in gene silencing by RNAi. In the cytoplasm, RDE-4 delivers dsRNA to Dicer **(1)**. Dicer enzyme processes long dsRNA **(2)** by cleaving the RNA into 20-25 nucleotide siRNAs **(3)**. In plants and nematodes, these primary (1°)

siRNAs serve as templates for secondary (2°) siRNA production **(3A)**. siRNAs are loaded into a large RNA-independent Silencing Complex (RISC) **(4)**. Argonaute2 (ARGO2) protein in RISC uses guide strain (red) of the siRNA to find complementary sequences along mRNA. Once this occurs the anti-guide strand (blue) is cleaved, activating the RISC complex **(5)** resulting in the mRNA cleavage **(6)**, in turn leading to gene silencing.

Lower: Secondary short interfering RNA (2° siRNA) production in plants and nematodes (figure adapted from Baulcombe, (2007)). After the production of 1° siRNAs shown in **Upper Box A step 3**, these 1° siRNAs act as targets for RNA-dependent RNA polymerase (RdRP)-mediated transcription of 2° siRNAs.
Box B. In *C. elegans* (lower left), ARGO2 associates with a 1° siRNA and targets a long single-stranded RNA (ssRNA) **(3A)** leading to the recruitment of RdRP **(3B)**. The RdRP synthesises 22-23 nucleotide 2° siRNAs **(3C)**.
Box C. In plants (lower right), RdRP recruitment is optimal when the long ssRNA has two targets for 1° siRNA (only one shown) **(3A)**. Following mRNA cleavage **(3B)**, RdRP is recruited and converts the targeted ssRNA into a long double-stranded RNA **(3C)**. The long dsRNA is then processed by Dicer to generate a pool of 2° siRNAs **(3D)**.

The subcellular localisation of the GPHR proteins in *C. elegans* is unknown. The GTG/GPHRs have already been described to reside either intracellularly, specifically in the Golgi (Maeda et al., 2008, Jaffé et al., 2012, Charroux and Royet, 2014) and ER (Jaffé et al., 2012, Charroux and Royet, 2014), or at the plasma membrane (Pandey et al., 2009). Since the localisation of proteins is often linked to their function, determining the subcellular localisation of the GTG/GPHRs would be critical for either supporting or rejecting the proposed functions for these proteins (section 1.4.2). On the other hand, the subcellular localisation may even suggest an alternative role for these proteins.

Typically, a known marker is employed to compare the localisation of the protein of interest in colocalisation analyses (Bolte and Cordelières, 2006). Colocalisation evaluates the 3D overlap between different fluorescent reporters of discrete emission wavelengths by analysing whether the targets are located in the same area of the cell or very near to one another (Bolte and Cordelières, 2006). There are two phenomena that umbrella the term colocalisation: co-occurrence and correlation. Co-occurrence is the presence of two fluorophores in the same pixel, and correlation states the relationship between the fluorophores statistically to show biological interaction (Adler et al., 2008). There are a number of methods that can be used to statistically analyse correlation between a pair of molecules including the Pearson's Correlation Coefficient (PCC) (Manders et al., 1992), the colocalisation coefficients (Manders et al., 1993) and the Costes algorithm (Costes et al., 2004).

To investigate whether these proteins are localised at the plasma membrane or intracellularly it would be sensible to begin the study using known markers like those for Golgi and ER. In *C. elegans*, a regularly used Golgi marker is the α-mannosidase-II (AMAN-2) protein (Rolls et al., 2002, Dejima et al., 2009, Broekhuis et al., 2013). AMAN-2 removes mannose residues from *N*-linked oligosaccharides to mature glycoproteins in the Golgi (Zhang et al., 2003). AMAN-2 is broadly expressed in larvae and adults with its strongest expression in the gut wall, pharynx and grinder, hypodermal cells and neurons (Figure 6.3; pharynx region is labelled on figure) (McKay et al., 2003, Paschinger et al., 2006, Dupuy et al., 2007). This would serve as a useful and easily distinguished marker when determining positive transformants. The construct *Paman-2::aman-2::mCherry* used in this study, has AMAN-2 C-terminally tagged with fluorescent protein mCherry, expressed under its own promoter, *Paman-2*.

sp12 encodes a signal peptidase of the ER membrane and is commonly used as a marker to study ER membrane proteins (Rolls et al., 2002). Loss of SP12 activity causes defects in pharyngeal morphology as shown by RNAi (Ferrier et al., 2011). The construct *Peft-4::mCherry::sp12* used in this study, has an mCherry N-terminally tagged SP12 under a constitutive promoter, *Peft-4*. *eef-1A.2* (*eft-4*) encodes a translation elongation factor 1-α and its promoter drives ubiquitous expression (McKay et al., 2003, Murata et al., 2012). *Peft-4* has been used to drive ubiquitous expression of genes for rescue (Murata et al., 2012) and localisation experiments (Dejima et al., 2009, Murata et al., 2012). *eft-4* also encodes an alternatively spliced coding sequence identical to *eft-3*; another ubiquitously expressed translation elongation factor in *C. elegans* (Gönczy et al., 2000, Mitrovich and Anderson, 2000, Maeda et al., 2001, McKay et al., 2003). *Peft-3* has been inserted into a Gateway-compatible *C. elegans* destination vector (section 6.2.4) for constitutive expression (Table 6.1). Both constructs (*Peft-4::mCherry::sp12* and *Paman-2::aman-2::mCherry*) were used as colocalisation markers to demonstrate ER and Golgi localisation of the HUT-1 protein in *C. elegans* (Dejima et al., 2009).

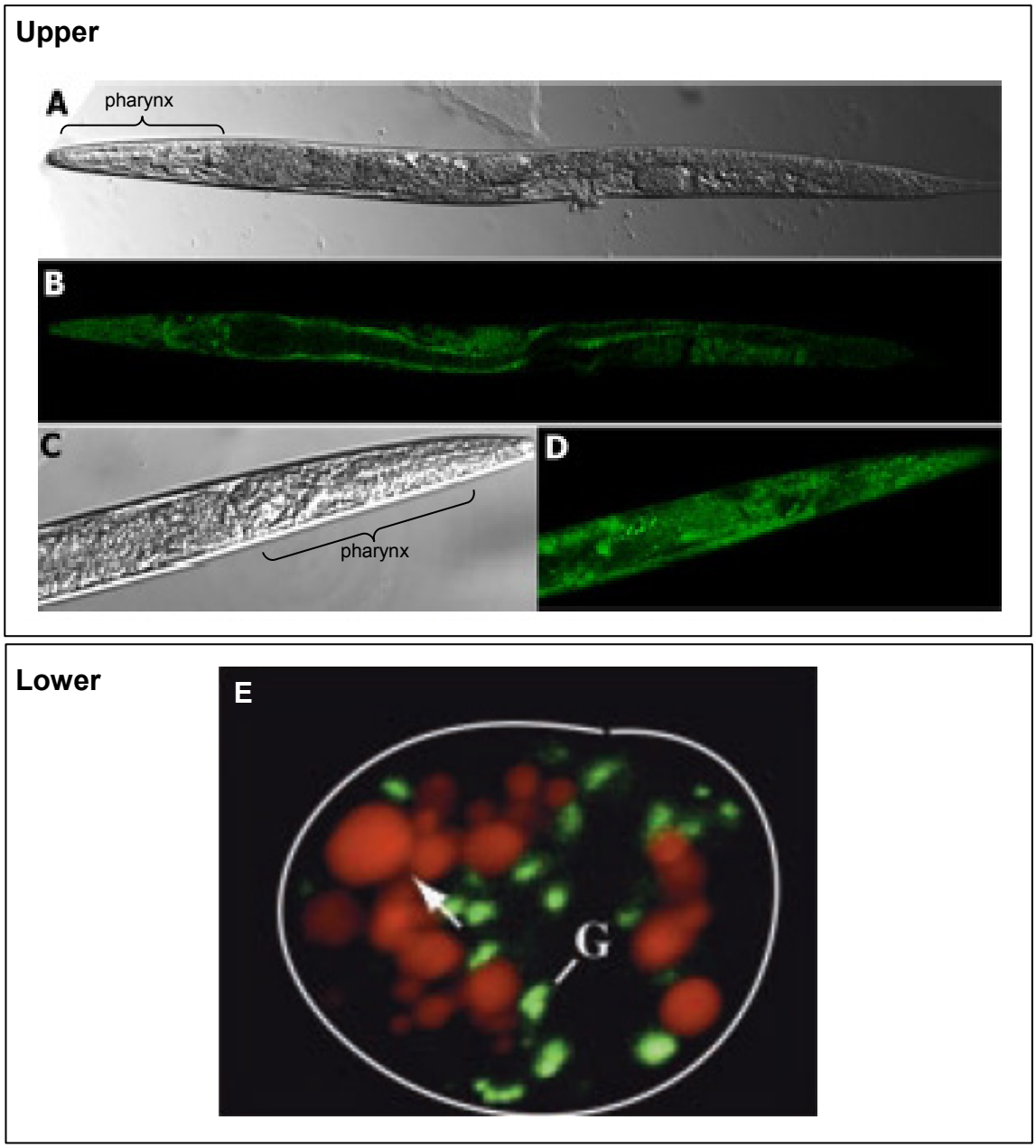


Figure 6.3 *aman-2* promoter-driven green fluorescent protein expression

Upper: A. The tissue localisation of AMAN-2 as shown by GFP expressed under the *aman-2* promoter. Relatively ubiquitous expression as shown by confocal transmission (**A**) and fluorescence (**B**) microscopy of an *Paman-2::gfp* transgenic *C. elegans* adult worm. The head region of an *Paman-2::gfp* transgenic L3 worm is shown at a higher magnification using confocal transmission (**C**) and fluorescence (**D**). Figure taken from Paschinger et al. (2006). **Lower**: **E**. The intracellular Golgi localisation of AMAN-2. *C. elegans* coelomocytes (outlined in white) expressing *aman-2::gfp* (green and punctate foci labelled with G). White arrow indicates large vacuole and endosomal comparments such as lysosomes labelled with BSA-Rhodamine (red). Scale bar: 5 μm. Figure taken from Treusch et al. (2004).

Table 6.1 Destination vectors for expression in *C. elegans*

Gateway-compatible vectors constructed for expression in *C. elegans* (Kuroyanagi et al., 2010).

Vector name	Description
pDEST-PL	Promoterless
pDEST-aex-3p	Pan-neuronal
pDEST-che-2p	Amphid sensory neurons
pDEST-dpy-7p	Hypodermal cells
pDEST-eat-4p	Glutamatergic neurons
pDEST-eft-3p*	Ubiquitous
pDEST-elt-2p	Intestine
pDEST-F25B3.3p	Pan-neuronal
pDEST-ges-1p	Intestine
pDEST-gon-2p	Intestine
pDEST-gst-42p	Intestine
pDEST-gtl-1p	Intestine
pDEST-hsp16-2p	Heat shock inducible
pDEST-hsp16-41p	Heat shock inducible
pDEST-mec-7p	Touch receptor neurons
pDEST-myo-2p*	Pharyngeal muscles
pDEST-myo-3p	Body wall muscles, vulval muscles

* indicates vectors used in this thesis supplied by Prof. H. Kuroyanagi (Tokyo Medical and Dental University, Japan).

6.1.4.1 Using the *C. elegans* pharynx to study GPHR localisation

Expression vectors will need to be constructed to study GPHR localisation in *C. elegans*. There are several destination vectors compatible with the Gateway system (Kuroyanagi et al., 2010), two of which were obtained for this study (Table 6.1). pDEST-myo-3p from this set of vectors (Table 6.1), for example, has been used to explore the expression and function of ASD-2B RNA-binding protein in the body wall muscles (Ohno et al., 2008). Highlighted with an asterix in Table 6.1 are the destination vectors containing the pharyngeal muscle (*myo-2*) and ubiquitous (*eft-3*) promoters. These destination vectors will be used to generate constructs where the expression of the *gphr* genes is driven under these promoters. *myo-2* encodes a muscle-type specific myosin heavy chain isoform (Dibb et al., 1989), expressed solely in the pharyngeal muscles (Jantsch-Plunger and Fire, 1994, Gaudet and Mango, 2002). The Golgi (*Paman-2::aman-2::mCherry*) and ER (*Peft-4::mCherry::sp12*) constructs (obtained from Prof. Yuichi Iino, The University of Tokyo, Japan) also target expression in the pharynx, therefore *Pmyo-2* would be useful for colocalisation studies.

The pharynx (introduced in section 1.5.4.2) is the neuromuscular, tubular pump of approximately 100 µm by 20 µm, which is isolated from the rest of the body by a basement membrane (Avery and You, 2012). Since there are only 20 muscle cells in the pharynx (Figure 1.12) and the structure is more easily distinguishable than distinct neuron sets for example, it would serve as a suitable organ to study the cellular localisation of the *C. elegans* GPHR proteins.

6.1.5 Aims

In order to study the role of the GPHR proteins in *C. elegans*, this project aims to:

1. Use the databases to identify the *gphr-1* and *gphr-2* single mutants in *C. elegans*.
2. Confirm the single mutants, backcross with WT *C. elegans* and sequence final deletion mutants.
3. Generate and characterise a whole animal, double *gphr* KO model, including phenotypic analyses of both single and double mutants; this will also address the question of functional redundancy.

4. Further investigate gene function using RNAi to try to knockdown the *gphr* genes in *C. elegans*.
5. Examine the tissue and cellular localisation of GPHR proteins in *C. elegans*; expression constructs will be made and transgenic *C. elegans* will be used to reveal this.

Overall, these approaches will help provide an insight into the biological role GPHRs play in *C. elegans*.

6.2 Results

6.2.1 Identification of single mutant *C. elegans* strains

C. elegans mutant databases CGC and NBP (section 6.1.3.1) were used to search for and identify deletion mutant strains for *gphr-1* and *gphr-2*. The *C. elegans* strains RB1390 and FXO4228 are deletion mutants for the genes *gphr-1* and *gphr-2* respectively. CGC and NBP provided the RB1390 and FXO4228 strains, respectively. The strains obtained from these databases and used in this thesis are listed in Table 2.2. WormBase (http://www.wormbase.org) is an online resource for the genetics of *C. elegans* and related nematodes. WormBase can aid *in silico* studies such as gene sequence analysis. The *gphr-1* and *gphr-2* genes are located on chromosomes III and X, respectively. The RB1390 strain of *C. elegans* contains the *gphr-1(ok1579)* mutant allele where there is a 1223 bp deletion in *gphr-1* (Figure 6.4 A and B). The FXO4228 strain of *C. elegans* contains the *gphr-2(tm4228)* mutant allele where there is a 413 bp deletion with a TGAAT 5 bp insertion in *gphr-2* (Figure 6.4 C and D).

6.2.2 Backcrossing the *C. elegans* single *gphr* mutant strains

The *C. elegans* single *gphr* mutants detailed above were genotyped to check for homozygosity before backcrossing. Both RB1390 and FXO4228 strains corresponded to the mutant alleles *gphr-1(ok1579)* and *gphr-2(tm4228)* respectively (Figure 6.5). The mutants were isolated using UV/TMP and therefore background mutations may have occurred during mutagenesis (section 6.1.3.1). For this reason a backcrossing scheme was developed (Figure 6.6), to remove any background mutations. Hermaphrodites reproduce through self-fertilisation (i.e. cloning). Therefore, to ensure an X-linked mutation is backcrossed, the X chromosome must be received from a male parent rather than from a hermaphrodite; this ensures recombination in the X chromosome will occur. As mentioned, the *gphr-2* gene is located on chromosome X; this is therefore accounted for in the backcross scheme developed in this project (Figure 6.6).

The *C. elegans* single mutant strains RB1390 and FXO4228 were allowed to grow and reproduce prior to extracting genomic DNA from each population. PCR using primer sets 3F and 1R for *gphr-1* and 1F and 1R for *gphr-2* (Figure 6.4 A and C) amplifies products with predicted molecular weights of 2567 and 1422 bp, for the WT alleles *gphr-1(+)* and *gphr-2(+)*, respectively. Predicted products of 1344 and 1014 bp

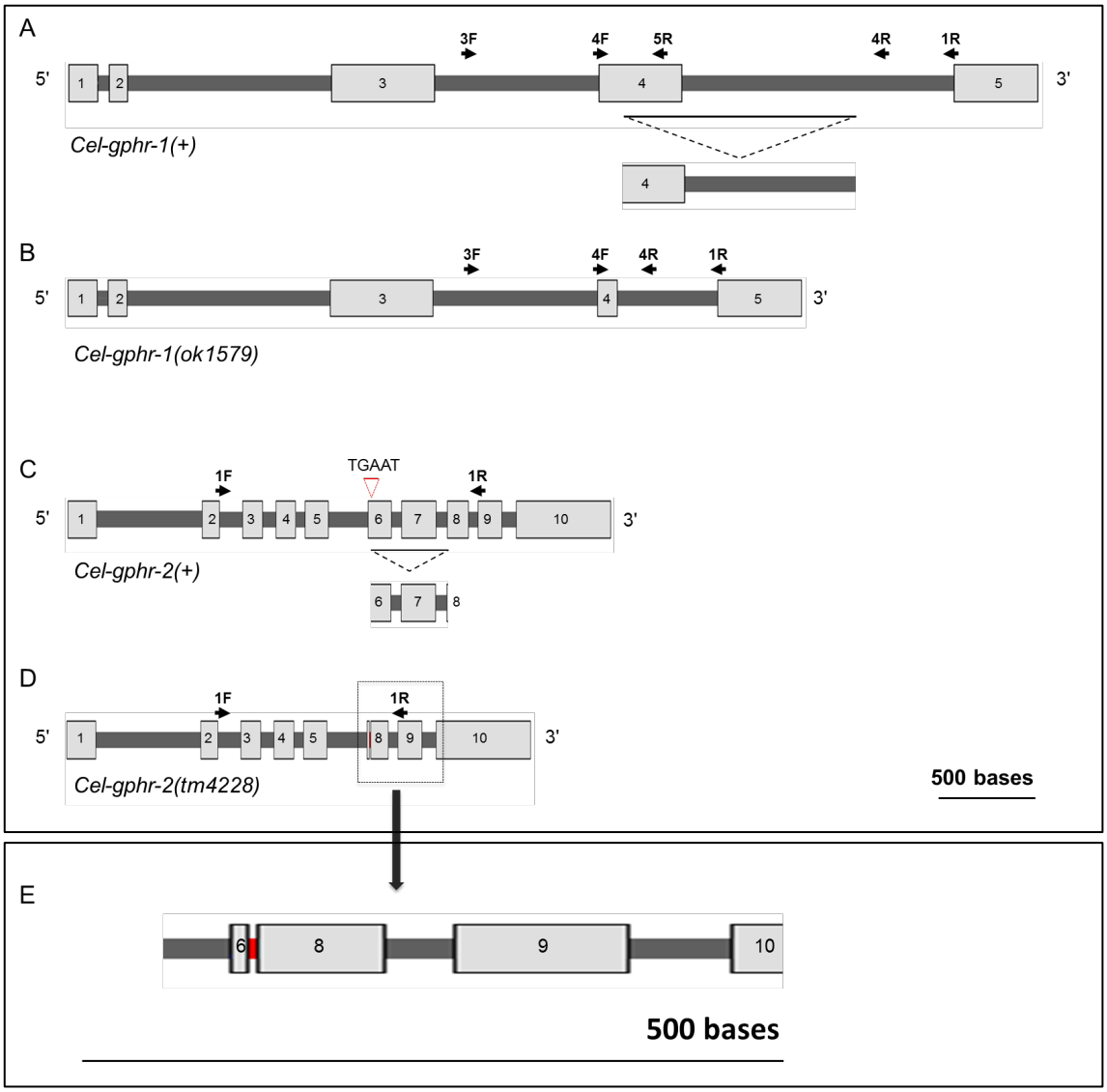


Figure 6.4 Genomic structures of the *gphr* genes and their mutant alleles

The *gphr* genomic structures display the exons and introns of the genes and mutant alleles drawn to scale (scale bars: 500 bases). **A**. *Cel-gphr-1(+)* gene structure with deletion site shown. **B**. *Cel-gphr-1(ok1579)* mutant allele. **C**. *Cel-gphr-2(+)* gene structure with deletion site shown. **D**. *Cel-gphr-2(tm4228)* mutant allele. **E**. Magnified section of *Cel-gphr-2(tm4228)* (inset in **D**). Exons are numbered and shown in light shade grey; introns are shown in dark shade grey. Symbols above the structures indicate primers used for genotyping (Table 2.3); arrows indicate direction of primers. Deleted portions of each gene are indicated by triangles shown in black underneath the WT allele. Insertion in *Cel-gphr-2* is shown by a red triangle above the WT allele in **C** and a red band in **E**.

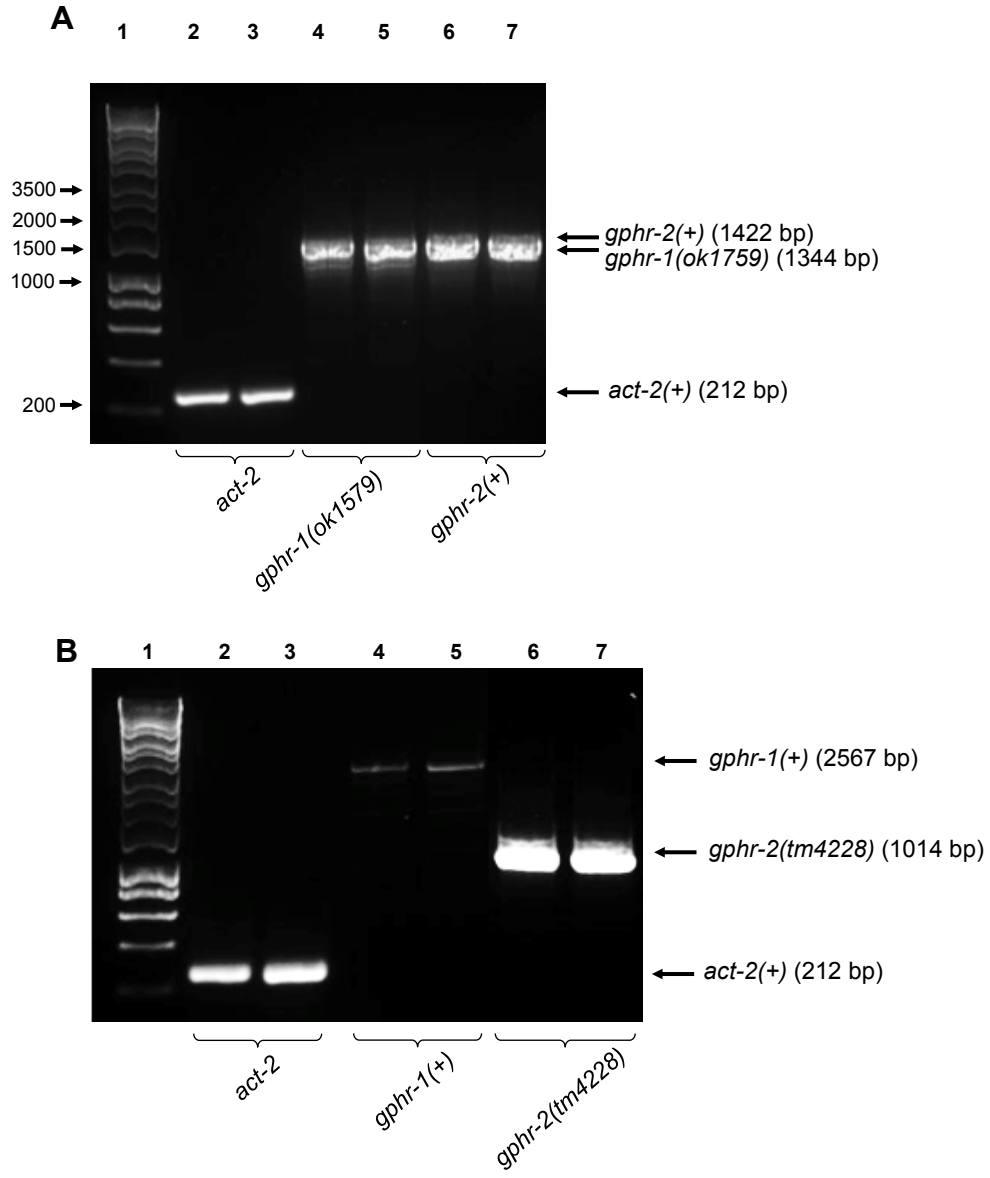


Figure 6.5 Genotyping RB1390 and FXO4228 single mutant strains

PCR and gel electrophoresis using genomic DNA from a whole population of worms from the RB1390 and FXO4228 strains. Predicted product sizes and lanes are shown on figure. Lanes 1: molecular markers **A**. RB1390 strain displaying genomic *actin-2* using primer set CeAct-2 F and CeAct-2 R (lanes 2 and 3), *gphr-2(+)* (using primer set 1F and 1R in Figure 6.4 C) (lanes 4 and 5) and *gphr-1(ok1579)* (using primer set 3F and 1R in Figure 6.4 B) (lanes 6 and 7). **B**. FXO4228 strain displaying genomic *actin-2* (lanes 2 and 3), *gphr-1(+)* (using primer set 3F and 1R from Figure 6.4 A) (lanes 4 and 5) and *gphr-2(tm4228)* (using primer sets 1F and 1R from Figure 6.4 D and primer sequences outlined in Table 2.10) (lanes 6 and 7).

are amplified using these same primer sets (Figure 6.4 B and D) for the mutant alleles *gphr-1(ok1579)* and *gphr-2(tm4228)*. Figure 6.5 A shows that the RB1390 strain contains *gphr-1(ok1579)* and *gphr-2(+)*. In Figure 6.5 A, the WT allele (*gphr-2(+)*) for *gphr-2* is amplified (predicted product size 1422 bp shown in Figure 6.4 C). However, a 1344 bp product is amplified for *gphr-1*, indicative of the mutant allele *gphr-1(ok1579)* (Figure 6.4 B) with no WT product (predicted 2567 bp shown in Figure 6.4 A). The FXO4228 strain contains *gphr-1(+)* and *gphr-2(tm4228)*. In Figure 6.5 the WT allele (*gphr-1(+)*) for *gphr-1* is amplified (predicted product size of 2567 bp shown in Figure 6.4 A). However, a 1014 bp product is amplified for *gphr-2*, indicative of the mutant allele *gphr-2(tm4228)* (Figure 6.4 D) with no WT product (predicted 1422 bp shown in Figure 6.4 C). Both strains are homozygous for their gene mutations as no predicted WT products were produced (Figure 6.5).

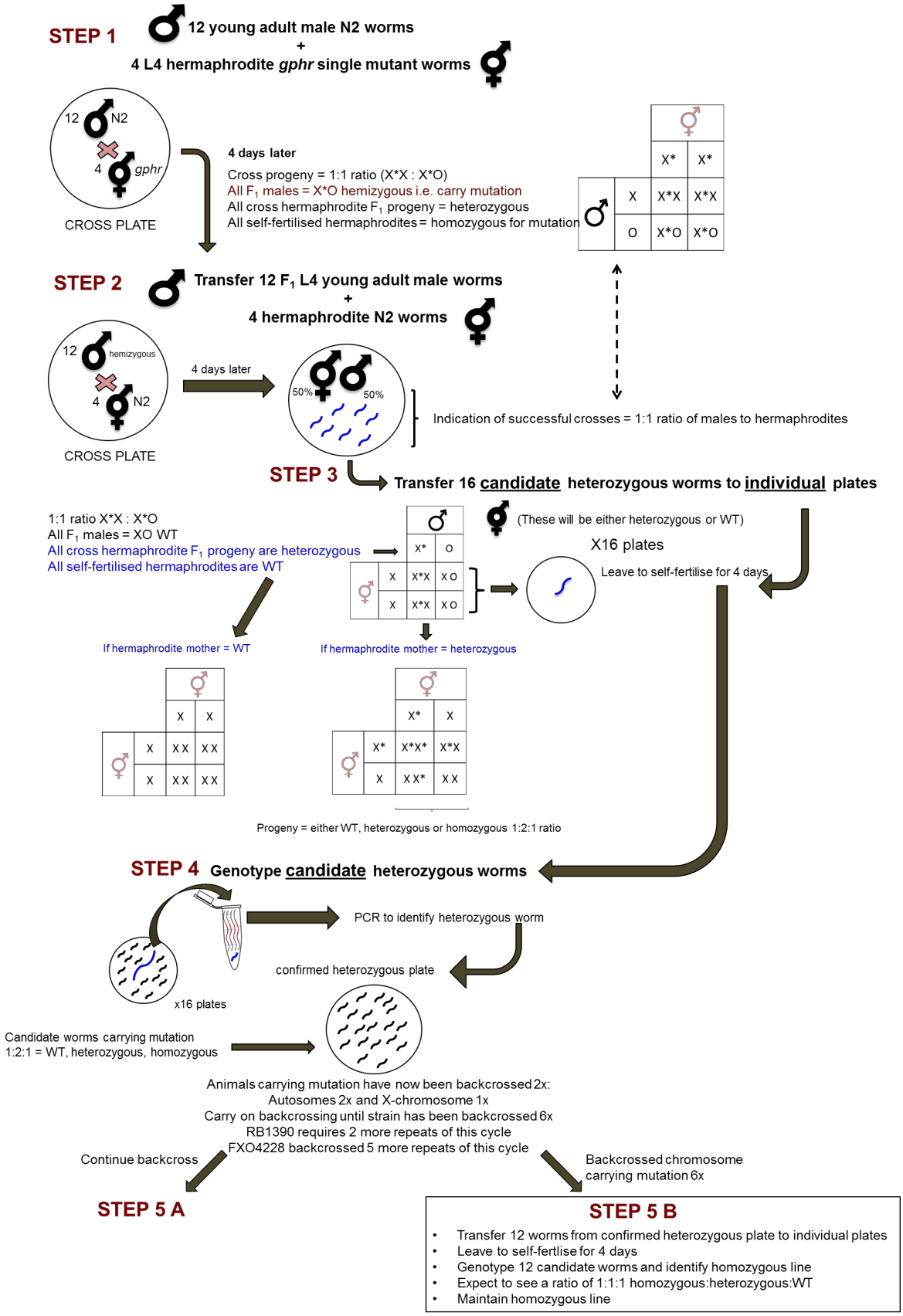
Primers flanking the deletion sites for both *gphr-1* and *gphr-2* (3F and 1R for *gphr-1* and 1F and 1R for *gphr-2* shown in Figure 6.4*)* were used to analyse the sequence in order to assess the deleted portions in both genes (Figure 6.4). Sequencing data confirmed that the *gphr-1(ok1579)* mutant has a 1223 bp deletion of which 326 bp is within the coding region of the 3′ end of the coding sequence (Figure 6.4 B). The deletion begins from the nucleotide 2868 to 4090 inclusive in the genomic sequence (Figure 6.4 B). The sequencing data for the *gphr-2(tm4228)* mutant revealed that there is a 413 bp deletion with a 5 bp insertion of the sequence 'TGAAT' (Figure 6.4 D). The deletion begins from nucleotide 1584 and extends to 1996 inclusive in the genomic sequence, removing introns 6 and 7, exon 7, the majority of exon 6 and 5′ end of exon 8 (Figure 6.4 D).

To omit any background mutations that may have ocurred during mutagenesis, each strain was backcrossed six times with the WT *C. elegans* strain, N2 (Figure 6.6). Male N2 worms were isolated by heatshock, for use in backcrossing hermaphrodite worms from both single mutant strains (Figure 6.6; step 1). After backcrossing, about half the population will be males and the other half will be hermaphrodites; this ratio is a good indication that the cross has occurred (Figure 6.1 step 1). First generation cross progeny (F_1) are heterozygous, i.e. *gphr-1(+/-)* for RB1390 and *gphr-2(+/-)* for FXO4228; all male progeny are hemizygous (Figure 6.6; step 1). Male and hermaphrodites can be distinguished by their tail region (Figure 6.1). After identifying male cross progeny, these were crossed with hermaphrodite N2 worms as all males

contain one copy of the mutation (Figure 6.6; step 2). The resulting F_1 from such a cross will produce either: homozygous, *gphr-1(-/-)* (RB1390) or *gphr-2 (-/-)* (FXO4228); heterozygous, *gphr-1(+/-)* (RB1390) or *gphr-2 (+/-)* (FXO4228); or WT *gphr-1(+/+)* (RB1390) or *gphr-2 (+/+)* (FXO4228) progeny (according to a 1: 2: 1 ratio) (Figure 6.6 step 2). These progeny were left to self-fertilise on individual plates for 4 d before genotyping (Figure 6.6; step 3). Genotyping was conducted using genomic DNA extracted from single worms followed by PCR (Figure 6.6; step 4), which was necessary as there was not a phenotype marking the presence of mutant alleles. 4F and 4R primers were used for genotyping RB1390 (Figure 6.4 A and B), and primers 1F and 1R were used for genotyping FXO4228 (Figure 6.4 C and D). If the F_1 worm is WT (resulting from self-fertilisation), all its progeny (F_2) would be WT; if the F_1 worm is heterozygous (resulting from crossing), *gphr-1(+/-)* (RB1390) or *gphr-2 (+/-)* (FXO4228), F_2 would be either WT, heterozygous or homozygous (Figure 6.6; step 3).

After identifying a heterozygote, 24 of its progeny (hermaphrodites) were transferred to individual plates and each crossed with six male N2 worms (Figure 6.6; step 5). These 24 cross plates were left for 4 d before each candidate worm was genotyped to identify a homozygote (Figure 6.6; step 6). The male cross progeny from a confirmed homozygote were used to continue the backcrossing procedure (Figure 6.6; step 6 and step 2). This cycle of crosses was repeated until six backcrosses were carried out (Figure 6.6). Since the *gphr-2* gene is located on chromosome X, the backcross cycle was repeated such that the autosomes were backcrossed 12 times and the X-chromosome six times.

After the cycles of crossing, to check that these were successfully backcrossed and homozygous for their respective mutation, the whole population genomic DNA was extracted for genotyping. For example, lane 4 of Figure 6.7 A, shows a RB1390 hermaphrodite population of worms that contains two copies of the *gphr-1(ok1579)* mutant allele i.e. a homozygous population of *gphr-1(-/-)*. The homozygosity was also confirmed for the FXO4228 strain that underwent six rounds of X-chromosome crossing as shown in lane 3 of Figure 6.7 B. This lane shows that this population of worms contain two copies of the *gphr-2(tm4228)* mutant allele as only one band of predicted size 1014 bp was observed (Figure 6.7 B). The RB1390 strain will henceforth be denoted RB1390x6 and FXO4228 as FXO4228x6 to indicate these had been backcrossed six times each.


STEP 1
12 young adult male N2 worms
+
4 L4 hermaphrodite gphr single mutant worms
12 N2
4 gphr
CROSS PLATE
4 days later
Cross progeny = 1:1 ratio (X*X : X*O)
All F1 males = X*O hemizygous i.e. carry mutation
All cross hermaphrodite F1 progeny = heterozygous
All self-fertilised hermaphrodites = homozygous for mutation
X* X*
X X*X X*X
O X*O X*O
STEP 2
Transfer 12 F1 L4 young adult male worms
+
4 hermaphrodite N2 worms
12 hemizygous
4 N2
CROSS PLATE
4 days later
50% 50%
Indication of successful crosses = 1:1 ratio of males to hermaphrodites
STEP 3
Transfer 16 candidate heterozygous worms to individual plates
1:1 ratio X*X : X*O
All F1 males = XO WT
All cross hermaphrodite F1 progeny are heterozygous
All self-fertilised hermaphrodites are WT
X* O
X X*X X O
X X*X X O
(These will be either heterozygous or WT)
X16 plates
Leave to self-fertilise for 4 days
If hermaphrodite mother = WT
X X
X X X X X
X X X X X
If hermaphrodite mother = heterozygous
X* X
X* X*X* X*X
X X X* X X
Progeny = either WT, heterozygous or homozygous 1:2:1 ratio
STEP 4 Genotype candidate heterozygous worms
PCR to identify heterozygous worm
confirmed heterozygous plate
x16 plates
Candidate worms carrying mutation
1:2:1 = WT, heterozygous, homozygous
Animals carrying mutation have now been backcrossed 2x:
Autosomes 2x and X-chromosome 1x
Carry on backcrossing until strain has been backcrossed 6x
RB1390 requires 2 more repeats of this cycle
FXO4228 backcrossed 5 more repeats of this cycle
Continue backcross
Backcrossed chromosome carrying mutation 6x
STEP 5 A
STEP 5 B
Transfer 12 worms from confirmed heterozygous plate to individual plates
Leave to self-fertlise for 4 days
Genotype 12 candidate worms and identify homozygous line
Expect to see a ratio of 1:1:1 homozygous:heterozygous:WT
Maintain homozygous line

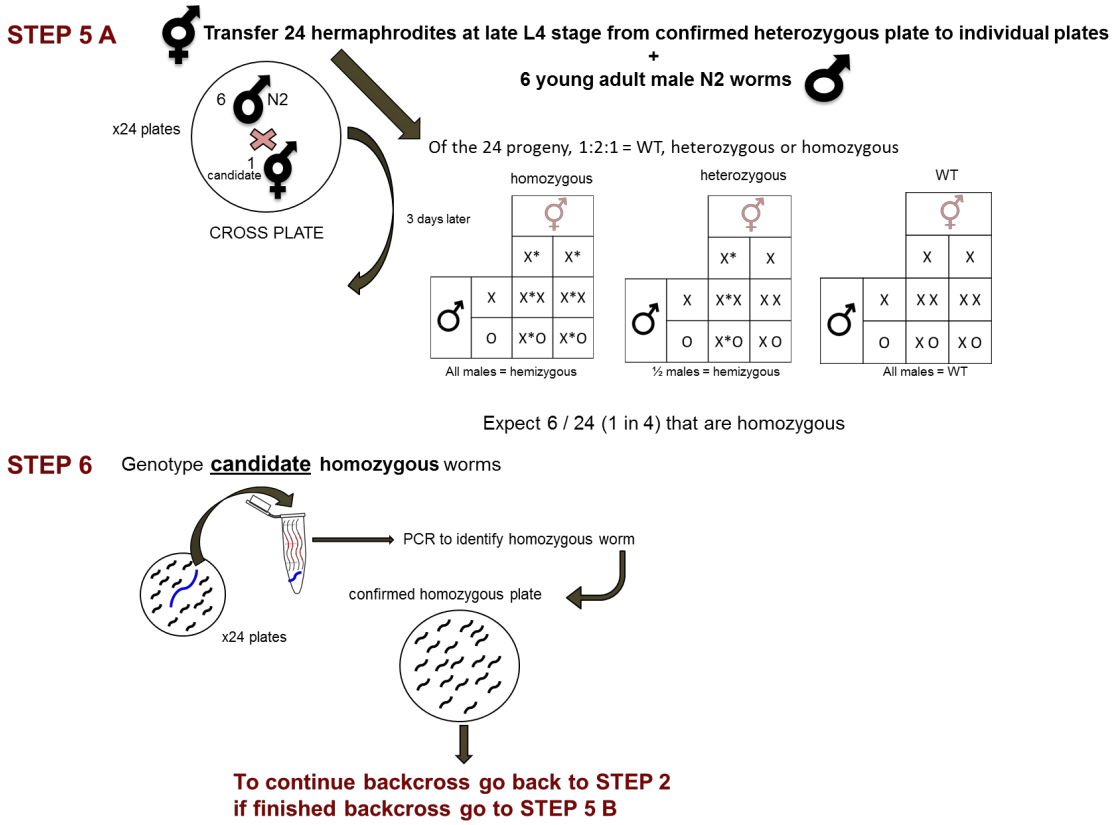


Figure 6.6 Method developed for backcrossing *gphr* single mutants

Procedure used for backcrossing RB1390 and FXO4228 strains. *gphr-1(ok1579)* allele, located on chromosome III, so RB1390 autosomes and X chromosome were backcrossed six and three times, respectively. *gphr-2(tm4228)* allele, located on chromosome X, so FXO4228 autosomes and X chromosome were backcrossed 12 and six times, respectively. Young adult male N2 worms were crossed with single *gphr* mutant L4-stage hermaphrodites (**step 1**). Male cross progeny were crossed with hermaphrodite N2 nematodes (**step 2**). F_1 from such a cross will produce e.g. homozygous, *gphr-1(-/-)*; heterozygous, *gphr-1(+/-)*; or WT *gphr-1(+/+)* (RB1390) progeny (according to a 1: 2: 1 ratio) (**step 2**). These F_1 candidates were left to self-fertilise on individual plates for 4 d (**step 3**) before individual genotyping (**step 4**). If F_1 candidate was WT, all its progeny (F_2) will be WT; if the F_1 candidate was heterozygous (resulted from crossing), *gphr-1(+/-)* (RB1390) or *gphr-2 (+/-)* (FXO4228), all F_2 will be either WT, heterozygous or homozygous (**step 3**). After identifying a heterozygote, 24 of its progeny (hermaphrodites) were transferred to individual plates and each crossed with six male N2 worms (**step 5 A**). These 24 cross plates were left for 4 d before each candidate hermaphrodite was genotyped to identify a homozygote (**step 6**). The male cross progeny from a confirmed homozygote were used to continue the backcrossing procedure (**step 6** and **step 2**). This cycle of crosses was repeated until six backcrosses were carried out and a homozygous line was maintained (**step 5 B**). The equivalent process was carried out for backcrossing the FXO4228, *gphr-2(tm4228)* strain.

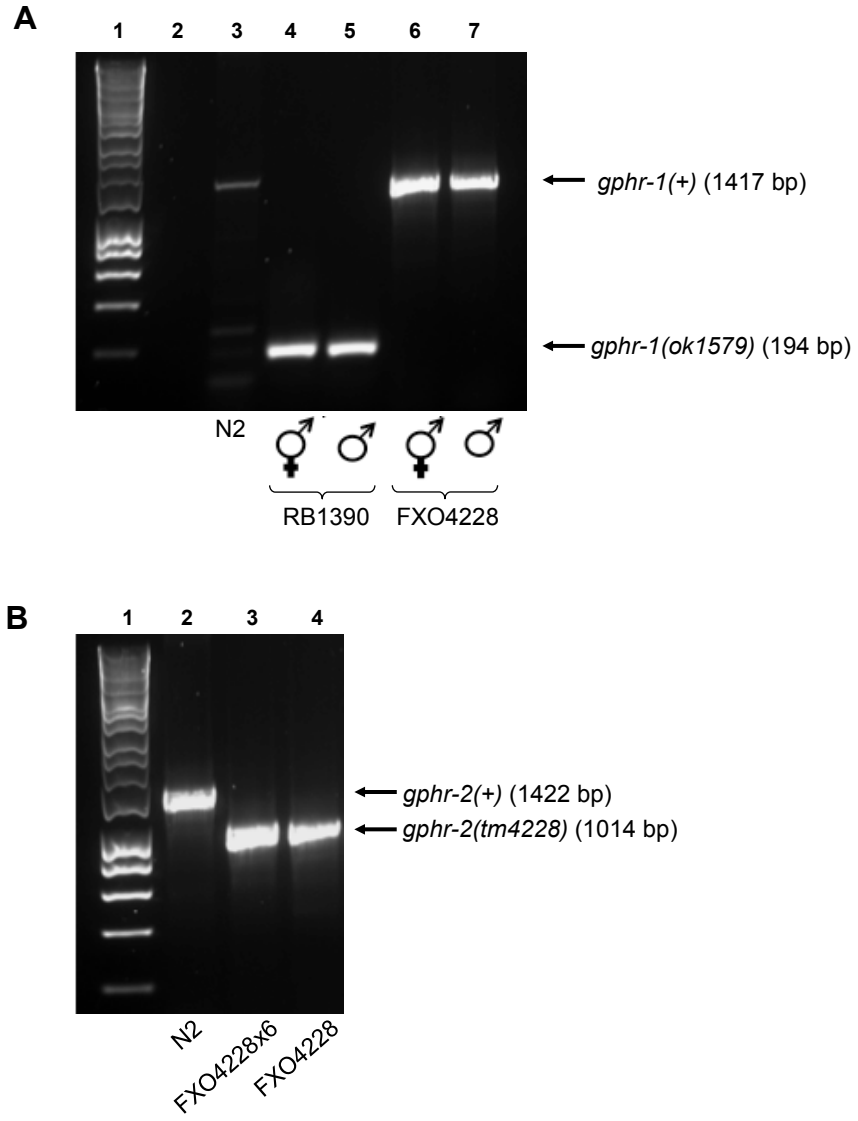


Figure 6.7 Genotyping backcrossed RB1390x6 and FXO4228x6 single *gphr* mutant strains

PCR and gel electrophoresis using genomic DNA from whole populations of worms from the N2, RB1390x6, male RB1390x6 and FXO4228x6 strains. Predicted product sizes and lanes are labelled on figure (see Figure 6.5 A for predicted molecular marker sizes). **A**. N2 strain displays *gphr-1(+)* (lane 3); RB1390x6 hermaphrodite (lane 4) and male (lane 5) strains display *gphr-1(ok1579)* (predicted product size of 194 bp). The primer set used was 4F and 4R (Figure 6.4 A and B) **B**. N2 strain displays *gphr-2(+)* (lane 2), FXO4228 (lane 4) and backcrossed FXO4228x6 (lane 3) strain displays *gphr-2(tm4228)*. The primer set used was 1F and 1R (Figure 6.4 D) and primer sequences outlined in Table 2.10.

6.2.3 Generation and characterisation of the *C. elegans* double mutant strain

Backcrossed single mutant strains RB1390x6 and FXO4228x6 in the *gphr-1* and *gphr-2* genes respectively were obtained in section 6.2.2. To generate a *gphr-1(ok1579) gphr-2(tm4228)* double mutant strain, a crossing procedure was developed (Figure 6.8). Figure 6.8 shows that a *gphr-1(-/-) gphr-2(-/-)* double mutant can be obtained from either a *gphr-1(+/-) gphr-2(+/-)*, *gphr-1(+/-) gphr-2(-/-)* or a *gphr-1(-/-) gphr-2(+/-)* mutant. The likelihood of obtaining a *gphr-1(-/-) gphr-2(-/-)* from a *gphr-1(+/-) gphr-2(+/-)* is 1 in 16, however from a *gphr-1(+/-) gphr-2(-/-)* or a *gphr-1(-/-) gphr-2(+/-)* mutant this chance is 1 in 4. Table 6.2 shows the number of independent experiments that was conducted to generate and characterise a double mutant and as well as the genotypes (*gphr-1(+/-) gphr-2(+/-)* or *gphr-1(+/-) gphr-2(-/-)*) that these double mutants originated from.

Male RB1390x6 worms were isolated by heatshock. Genomic DNA extraction from the resulting population of males was genotyped to confirm that these nematodes were homozygous *gphr-1(ok1579)* mutants (Figure 6.7 A). The male RB1390x6 strain was crossed with hermaphrodite FXO4228x6 (Figure 6.8; step 1). Cross progeny are heterozygous in both *gphr-1(+/-)* and *gphr-2(+/-)*. To identify heterozygous cross progeny, ten candidate worms were transferred to individual plates and left to self-fertilise for 4 d (Figure 6.8; step 2). Subsequently, these ten candidate worms were genotyped to identify a *gphr-1(+/-) gphr-2(+/-)* mutant (Figure 6.8; step 3).

For analysing the *gphr-1* genotype: *gphr-1(+/+)* nematodes (i.e. WT) contain two copies of the WT allele and show only show one product (predicted size of 329 bp using primers 4F and 5R, and 1417 bp using primers 4F and 4R, Figure 6.4 A); homozygous animals for the mutation *gphr-1(-/-)* will contain two copies of the mutant *ok1579* allele and so will only show a predicted product of 194 bp (using primers 4F and 4R, Figure 6.4 B) and no product if using primers 4F and 5R, Figure 6.4 B; heterozygous animals are indicated by having both the predicted, upper 1417 bp (WT allele) and predicted, lower 194 bp (mutant allele) products amplified. After confirming a *gphr-1(+/-) gphr-2(+/-)* mutant (such as in Figure 6.9), its progeny was used to identify a *gphr-1(-/-) gphr-2(-/-)* mutant. To do this, candidate worms from the plate were transferred to individual plates and left to self-fertilise for 4 d (Figure 6.8; step 4); it was expected to obtain one *gphr-1(-/-) gphr-2(-/-)* mutant in 16 worms. Potential

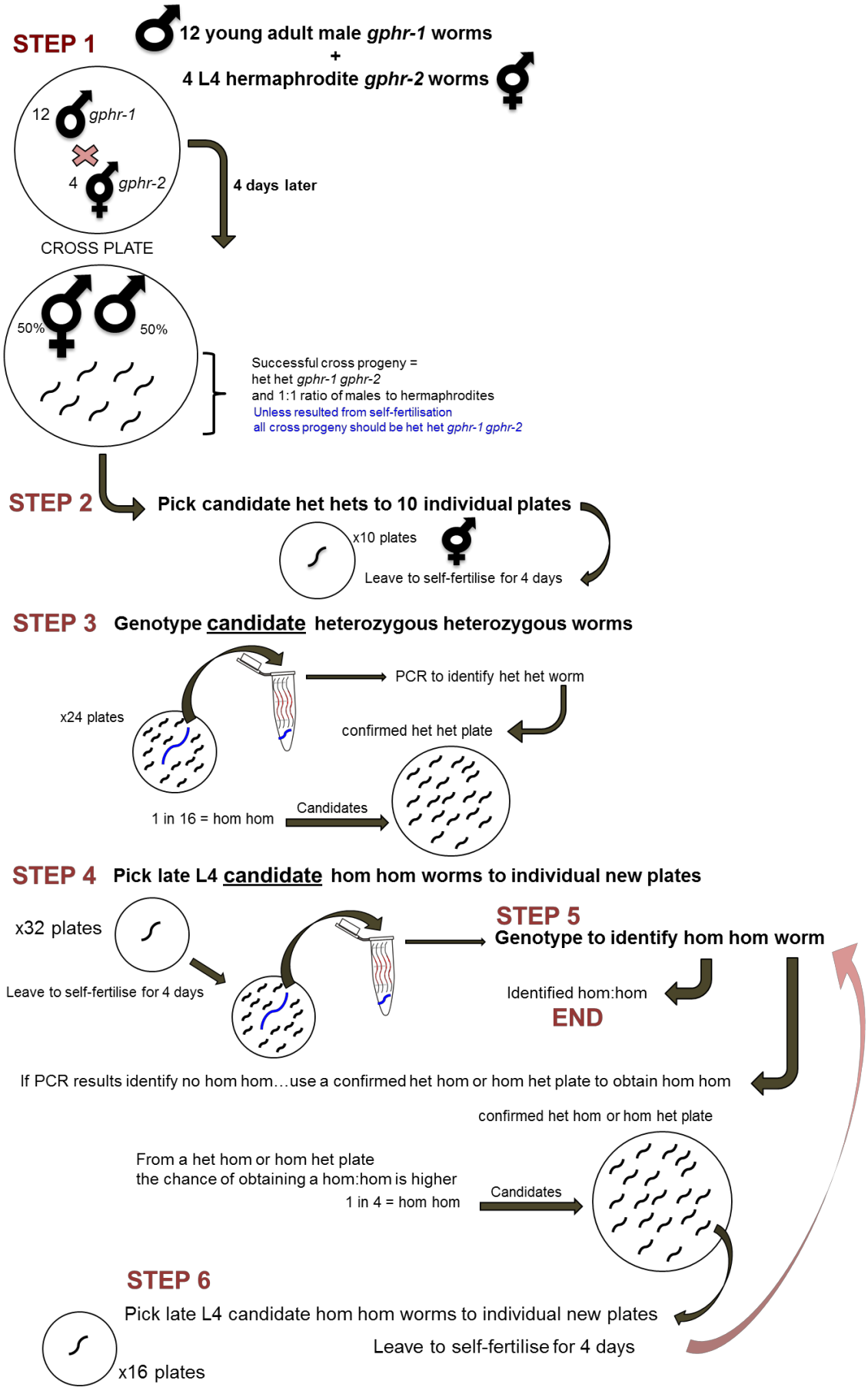

STEP 1
12 young adult male gphr-1 worms
+
4 L4 hermaphrodite gphr-2 worms
12 gphr-1
4 gphr-2
4 days later
CROSS PLATE
50%
50%
Successful cross progeny =
het het gphr-1 gphr-2
and 1:1 ratio of males to hermaphrodites
Unless resulted from self-fertilisation
all cross progeny should be het het gphr-1 gphr-2
STEP 2
Pick candidate het hets to 10 individual plates
x10 plates
Leave to self-fertilise for 4 days
STEP 3
Genotype candidate heterozygous heterozygous worms
x24 plates
PCR to identify het het worm
confirmed het het plate
1 in 16 = hom hom
Candidates
STEP 4
Pick late L4 candidate hom hom worms to individual new plates
x32 plates
Leave to self-fertilise for 4 days
STEP 5
Genotype to identify hom hom worm
Identified hom:hom
END
If PCR results identify no hom hom…use a confirmed het hom or hom het plate to obtain hom hom
confirmed het hom or hom het plate
From a het hom or hom het plate
the chance of obtaining a hom:hom is higher
1 in 4 = hom hom
Candidates
STEP 6
Pick late L4 candidate hom hom worms to individual new plates
Leave to self-fertilise for 4 days
x16 plates

Figure 6.8 Crossing method developed to generate double *gphr* mutant

Male RB1390x6 was crossed with hermaphrodite FXO4228x6. Cross progeny are heterozygous in both *gphr-1(+/-)* and *gphr-2(+/-)* (**step 1**). After reproducing for 4 d, ten candidate worms were transferred to individual plates and left to self-fertilise for 4 d (**step 2**). After self-fertilising for a 4 d, these ten candidate worms were genotyped to identify a *gphr-1(+/-) gphr-2(+/-)* (**step 3**). This plate was used to transfer potential *gphr-1(-/-) gphr-2(-/-)* worms to individual plates and left to self-fertilise for 4 d (**step 4**); it was expected to obtain one *gphr-1(-/-) gphr-2(-/-)* in every 16 worms. Potential homozygous animals in both genes were genotyped to identify a *gphr-1(-/-) gphr-2(-/-)* (**step 5**). If PCR results indicated no *gphr-1(-/-) gphr-2(-/-)* out of the candidates, the isolation process was continued using individual candidate worms from a confirmed *gphr-1(+/-) gphr-2(-/-)* or a *gphr-1(-/-) gphr-2(+/-)* plate (**step 6**). These were left to self-fertilise for 4 d before genotyping (back to **step 5**) until a *gphr-1(-/-) gphr-2(-/-)* was obtained.

Table 6.2 The number of double mutants obtained for each replicate experiment

The experiments conducted to isolate a double *gphr-1(ok1579) gphr-2(tm4228)* mutant. The number of double mutants expected and actual numbers obtained, and the genotype (*gphr-1(+/-) gphr-2(+/-)* or *gphr-1(+/-) gphr-2(-/-)*) it resulted from are shown. Het, heterozygous; hom, homozygous; the independent experiment number where decimal numbers indicate the isolation process continuing from a *gphr-1(+/-)gphr-2(-/-)* mutant, for example experiment 1.1 and 1.2 are from one experiment but 1.2 is continuing from a heterozygous homozygous, *gphr-1(+/-) gphr-2(-/-)* from experiment 1.1. The experiment aim was either to isolate the double mutant or to simultaneously isolate and phenotype the double mutant. For all double homozygous *gphr* mutants obtained, few eggs were laid and only a small proportion hatched of which none developed into adults. In all cases, the F_0 *gphr-1(-/-) gphr-2(-/-)* identified grew to adult stage with little locomotion, as shown by lack of tracks on food (Figure 6.11), but all resulting progeny (F_1) that hatched arrested at L1 larval stage of development (Figure 6.12).

Experiment number	**Experiment aim**	**Genotype of progenitor**	**Number of double mutants**	
			Expected	**Actual**
1.1	Isolation	Het het	1/16	0/16
1.2	Isolation	Het hom	1/4	1/6
1.3	Isolation	Het hom	1/4	3/16
2.1	Isolation	Het het	1/16	0/16
2.2	Isolation and phenotyping	Het hom	1/4	3/13
3	Isolation and phenotyping	Het het	1/16	2/20

homozygous animals mutated in both genes were genotyped to identify a *gphr-1(-/-) gphr-2(-/-)* mutant. If PCR results indicated no *gphr-1(-/-) gphr-2(-/-)* mutant out of the candidates, the self-fertilisation process was continued using individual candidate worms from a confirmed *gphr-1(+/-) gphr-2(-/-)* or a *gphr-1(-/-) gphr-2(+/-)* plate (Figure 6.8 step 5). Figure 6.10 shows an example screen of potential *gphr-1(-/-)* and *gphr-2(-/-)* double mutant candidates while carrying out the cross scheme (Figure 6.8). For example, in lane 6 of Figure 6.10, it shows no product (Figure 6.10 A) and a predicted product of 194 bp (Figure 6.10 B) corresponding to a homozygous *gphr-1(-/-)* mutant. For analysing the *gphr-2* genotype: *gphr-2(+/+)* nematodes (i.e. WT) containing two copies of the WT allele show only show one product (predicted size of 1422 bp using primers1F and 1R, Figure 6.4 C); homozygous animals for the mutation *gphr-2(-/-)* will contain two copies of the mutant *tm4228* allele and so will only show a predicted product of 1014 bp (using primers 1F and 1R, Figure 6.4 D); heterozygous animals are indicated by having both the predicted, upper 1422 bp (WT allele) and predicted, lower 1014 bp (mutant allele) products amplified (Figure 6.10 C). For example, in lane 6 of Figure 6.10 C, it shows the predicted product of 1014 bp, therefore for candidate "a" the genotype is *gphr-1(-/-) gphr-2(-/-)*. The expected ratio and an example of those actually obtained for the various genotypes descended from a *gphr-1(+/-) gphr-2(+/-)* or a *gphr-1(+/-)* and *gphr-2(-/-)* can be found in Table 6.3 and Table 6.4, respectively.

It was noticed that some candidate *gphr-1(-/-) gphr-2(-/-)* worms showed little movement during the isolation process (Figure 6.11); these were suspected to be double *gphr-1(-/-) gphr-2(-/-)* mutants. Indeed, these suspected candidates were confirmed as *gphr-1(-/-) gphr-2(-/-)*. After identifying *gphr-1(-/-) gphr-2(-/-)* mutants by genotyping the progeny using PCR, the eggs laid by the progenitor were left to hatch in an attempt to maintain a *gphr-1(ok1579) gphr-2(tm4228)* line. However, on two occasions (Table 6.2 experiments 1.2 and 1.3), only a few eggs hatched. Of these progeny that did hatch, it was observed that the nematodes exhibited L1 larval arrest. It was also observed that the larva had limited mobility throughout their lifetime before dying after over a week of remaining at the larval stage of development (Figure 6.12).

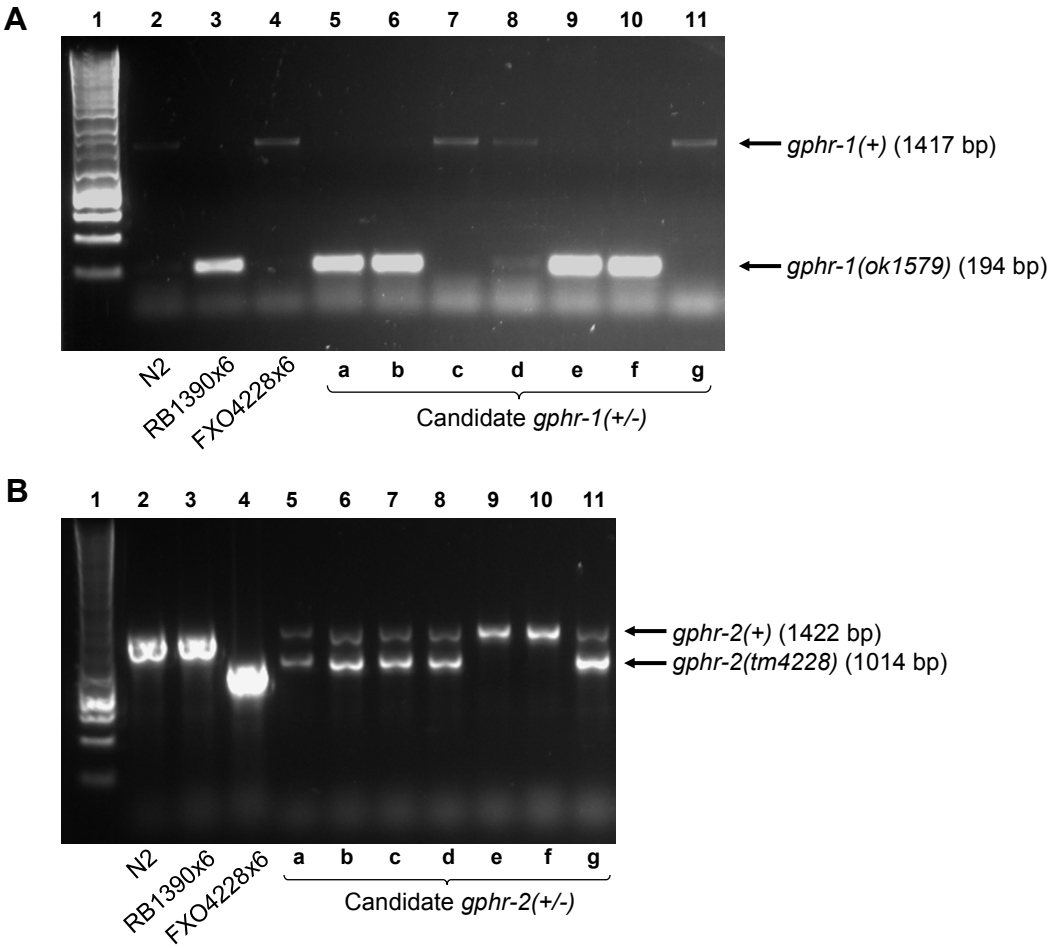


Figure 6.9 Genotyping single worms for screening heterozygous double *gphr* mutants

PCR and gel electrophoresis to show the genotypes of candidate double heterozygous (*gphr-1(+/-) gphr-2(+/-)*) worms. Predicted band sizes and lanes are shown on figure. Lane 1: molecular markers (see Figure 6.5 A for predicted molecular marker sizes). **A**. Primers 4F and 4R (Figure 6.4 A and B) amplify *gphr-1* with seven candidate worms (a – g) displaying WT *gphr-1*(+) (lane 2) and/or mutant *gphr-1(ok1579)* (lane 3) products. **B**. Primers 1F and 1R (Figure 6.4 C and D) to amplify *gphr-2* with seven candidate worms (a – g) displaying WT *gphr-2(+)* (lane 2) and/or mutant *gphr-2(tm4228)* (lane 4) products.

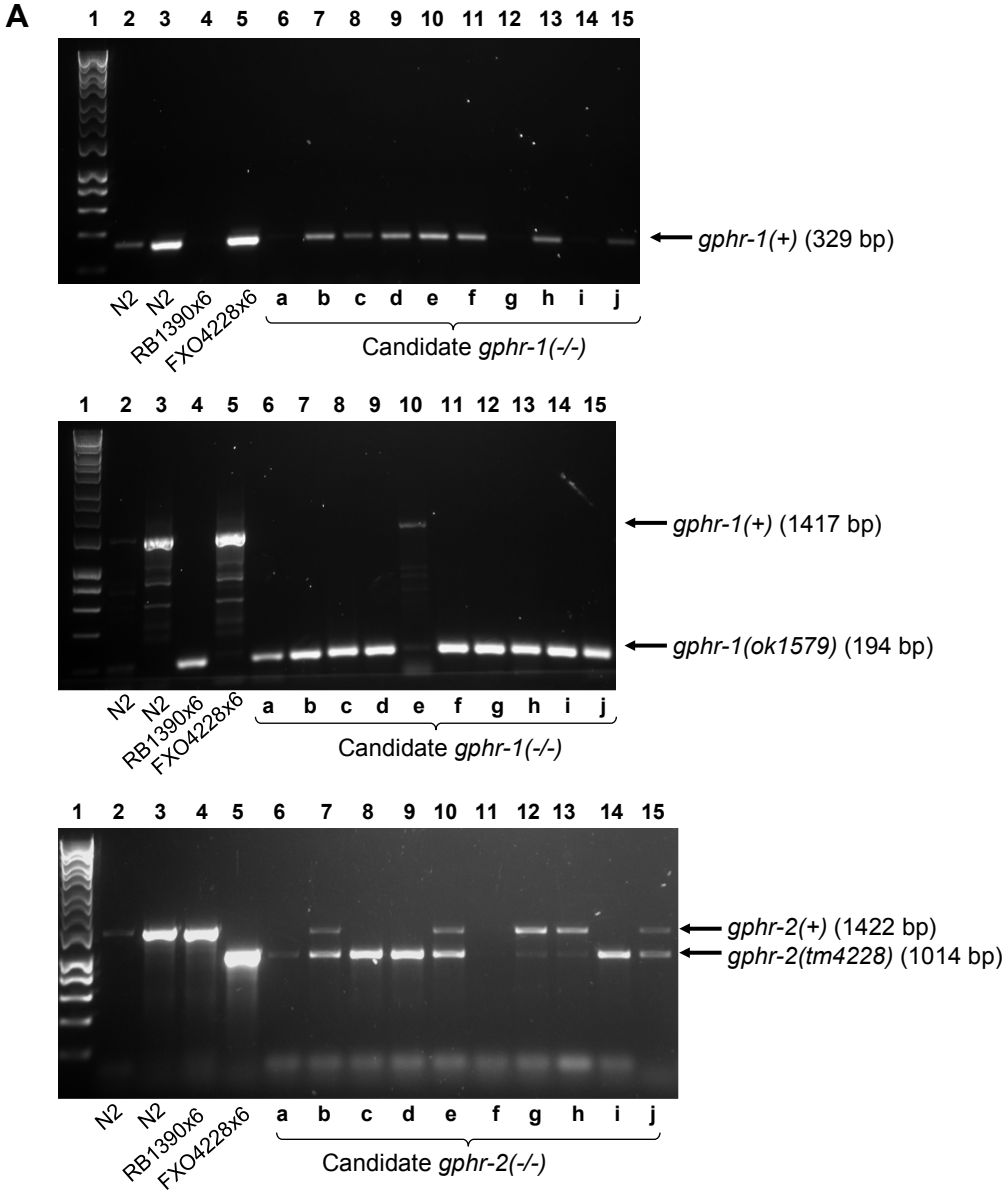


Figure 6.10 Genotyping single worms for screening homozygous double *gphr* mutants

PCR and gel electrophoresis showing the genotypes of candidate *gphr-1(-/-) gphr-2(-/-)* double mutant worms. Predicted band sizes and lanes are shown (see Figure 6.5 A for predicted molecular marker sizes). **A**. Nested primers 4F and 5R (Figure 6.4 A and B) amplify *gphr-1* with ten candidates (a – j) displaying WT *gphr-1(+)* or no product indicative of mutant *gphr-1(ok1579)*. **B**. Primers 4F and 4R (Figure 6.4 A and B) amplify *gphr-1* with ten candidate worms (a – j) displaying WT *gphr-1(+)* (lanes 2 and 3) and/or mutant *gphr-1(ok1579)* (lane 4) products. **C**. Primers 1F and 1R (Figure 6.4 C and D) to amplify *gphr-2* with ten candidate worms (a – j) displaying WT *gphr-2(+)* (lanes 2 and 3) and/or mutant *gphr-2(tm4228)* (lane 5) products.

Table 6.3 Expected and actual numbers of the various genotypes from self-fertilising a *gphr-1(+/-) gphr-2 (+/-)* mutant

This data is taken from experiment #3 of Table 6.2. Expected and actual numbers obtained were not significantly different; χ^2 (8, N = 1) = 0.23, p = 1.000

Genotype	Expected number	Actual number
gphr-1(+) gphr-2(+)	1/16	0/20
gphr-1(-/-) gphr-2(-/-)	1/16	2/20
gphr-1(-/-) gphr-2(+/-)	2/16	2/20
gphr-1(+/-) gphr-2(-/-)	2/16	3/20
gphr-1(+/-)gphr-2(+/-)	4/16	7/20
gphr-1(+) gphr-2(+/-)	2/16	2/20
gphr-1(+/-) gphr-2(+)	2/16	2/20
gphr-1(+) gphr-2(-/-)	1/16	2/20
gphr-1(-/-) gphr-2(+)	1/16	0/20

Table 6.4 Expected and actual numbers of the various genotypes from self-fertilising a *gphr-1(+/-) gphr-2 (-/-)* mutant

This data is taken from experiment #2.2 of Table 6.2. Expected and actual numbers obtained were not significantly different; χ^2 (2, N = 1) = 0.07, p = 0.968

Genotype	Expected number	Actual number
gphr-1(-/-) gphr-2(-/-)	1/4	3/13
gphr-1(+/-) gphr-2(-/-)	1/2	8/13
gphr-1(+) gphr-2(-/-)	1/4	2/13

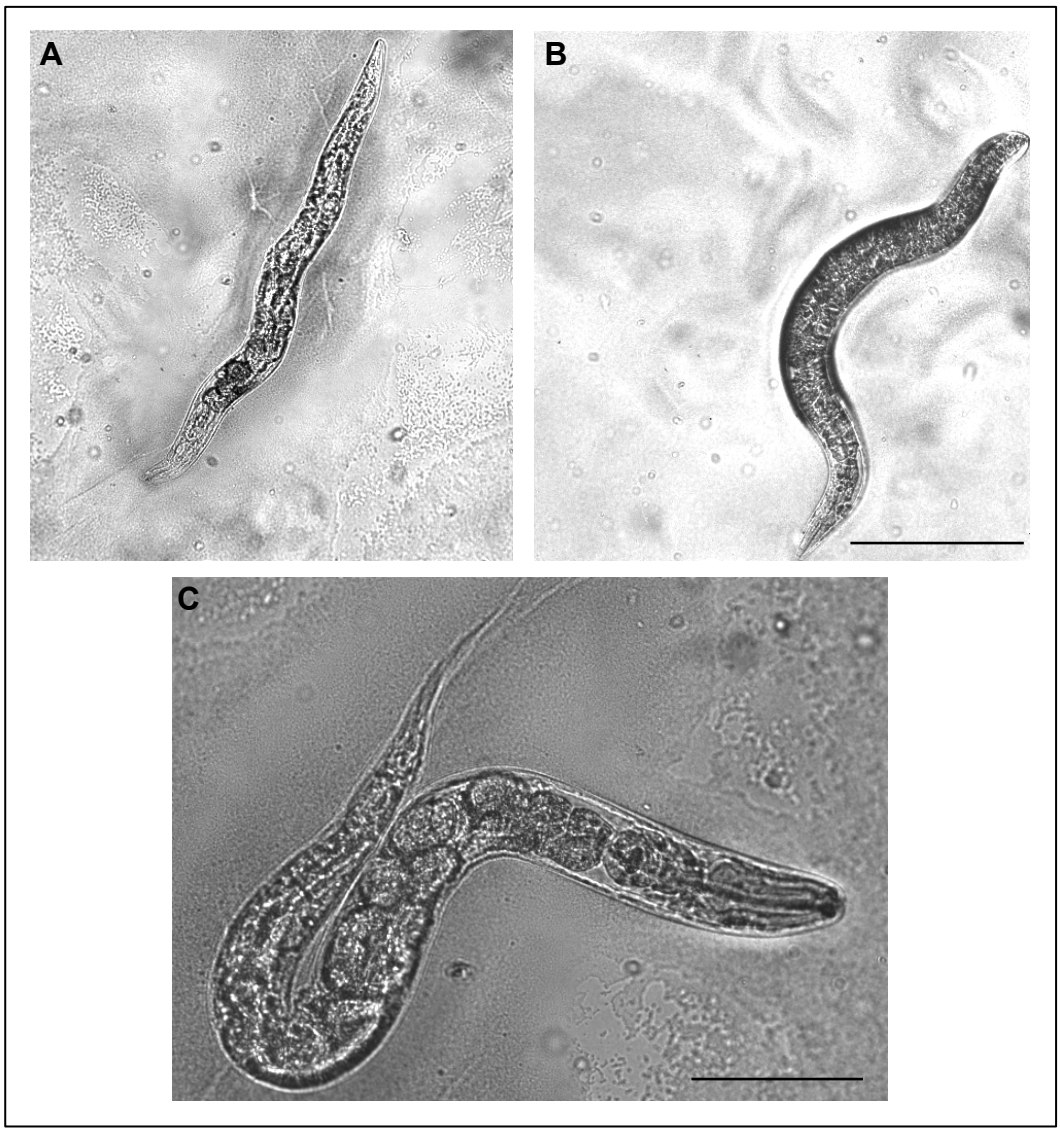


Figure 6.11 F_0 *gphr* double mutants display little locomotion

A – C. Images to show three individual adult (L4+1) F_0 *gphr-1(ok1579) gphr-2(tm4228)* double mutants obtained from a *gphr-1(+/-) gphr-2(-/-)* from experiment 2.2 (Table 6.2). These double mutants exhibited little movement throughout their life time, as demonstrated by the few track marks surrounding them. Images are representative of nine double mutants obtained from four independent attempts of isolating a double mutant (Table 6.2). Scale bars: 100 μm.

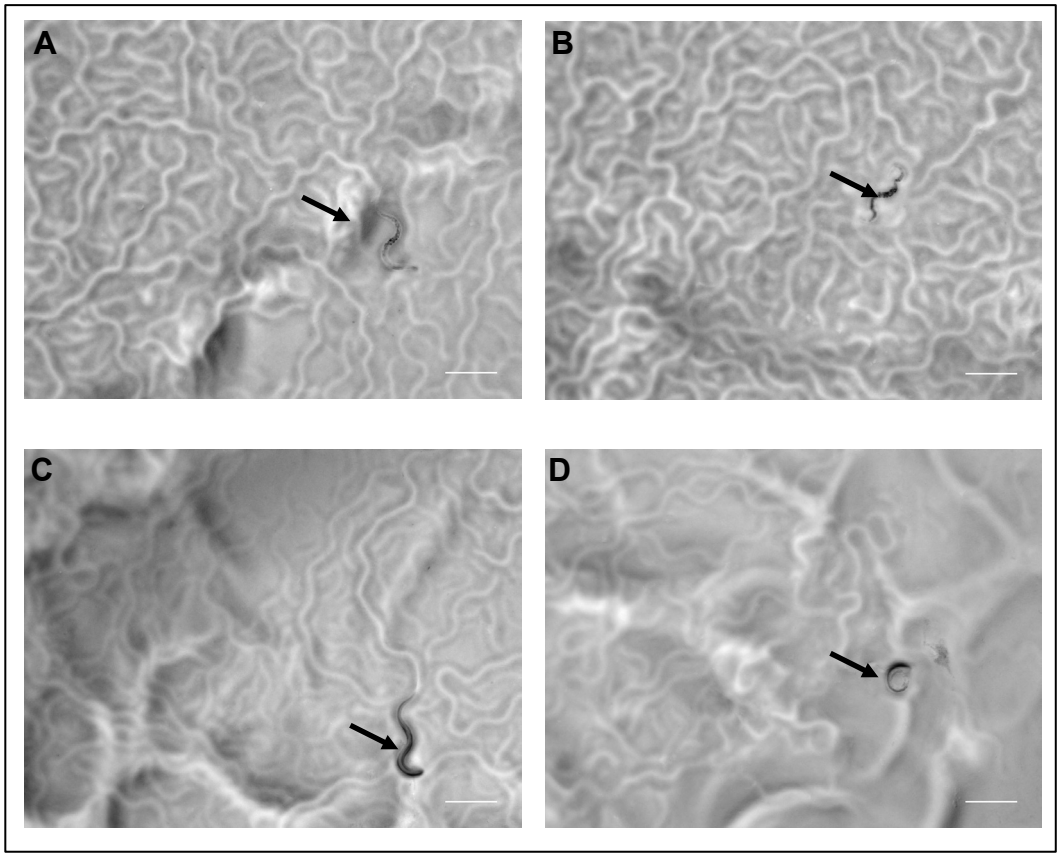


Figure 6.12 F_1 *gphr* double mutants exhibit larval arrest

A – D. Images to show the several *gphr-1(ok1579) gphr-2(tm4228)* double mutants that hatched but arrested at L1 stage. Images were taken one week after hatching, at which stage these should have grown into adults (Figure 1.9 and Figure 1.11). Black arrows point to double mutant larva in each panel; **A – D** are separate worms. Images are representative of larvae obtained from experiment 1.3 (Table 6.2), but seen in all independent attempts to isolate a double mutant. Scale bars: 100 μm.

Because the *gphr-1(ok1579) gphr-2(tm4228)* double mutant strain did not progress through the normal life cycle (Figure 1.11) it could not be easily maintained. Therefore to assess accurately the phenotypes displayed by the double mutant it was necessary to assess their characteristics during the isolation procedure. An isolation and characterisation process was developed in order to accurately genotype and phenotype the *gphr-1(ok1579) gphr-2(tm4228)* double mutant strain. During this simultaneous isolating and phenotyping process, the final genotype of the candidate double *gphr* mutants was unknown until after the characterisation i.e. this process was carried out blind. This is because worm genotyping involves lysing the whole worm and this step must therefore follow after phenotypic characteristics have been measured. A protocol was developed in order to facilitate this and an overview of this is given in Figure 6.13.

6.2.3.1 Defective egg laying, development and hatching phenotype observed *gphr* double mutants

As described in section 6.2.3 it was observed that very few eggs were laid by the *gphr-1(ok1579) gphr-2(tm4228)* double mutant generated on several occasions from independent crosses. In order to test whether there is a significant difference between the number of eggs laid by a WT and *gphr-1(ok1579) gphr-2(tm4228)* double mutant, the scheme described in Figure 6.13 was used. The egg-laying assays were conducted in the presence of abundant food because the egg-laying rate is significantly higher on abundant food compared to the absence of food (Trent, 1982). The total number of eggs laid by a *gphr-1(-/-) gphr-2(-/-)* mutant over a duration of 4 d was significantly less than that of a WT and also of either single *gphr-1(+/-) gphr-2(-/-)* or *gphr-1(-/-) gphr-2(+/-)* mutants (Figure 6.14 A). The total number of eggs laid by the two backcrossed (section 6.2.2) single mutants *gphr-1(ok1579)* (RB1390x6) and *gphr-2(tm4228)* (FXO4228x6) was also assayed, but no significant difference was seen between each or from the WT (Figure 6.14 B). Furthermore, it was observed that the F_0 double mutants exhibited little movement (mentioned in section 6.2.3); therefore images were taken of candidates suspected to be *gphr-1(-/-) gphr-2(-/-)* mutants through lack of movement and reduced eggs laid (Figure 6.11). These were later confirmed by PCR to be *gphr-1(-/-) gphr-2(-/-)* mutants.

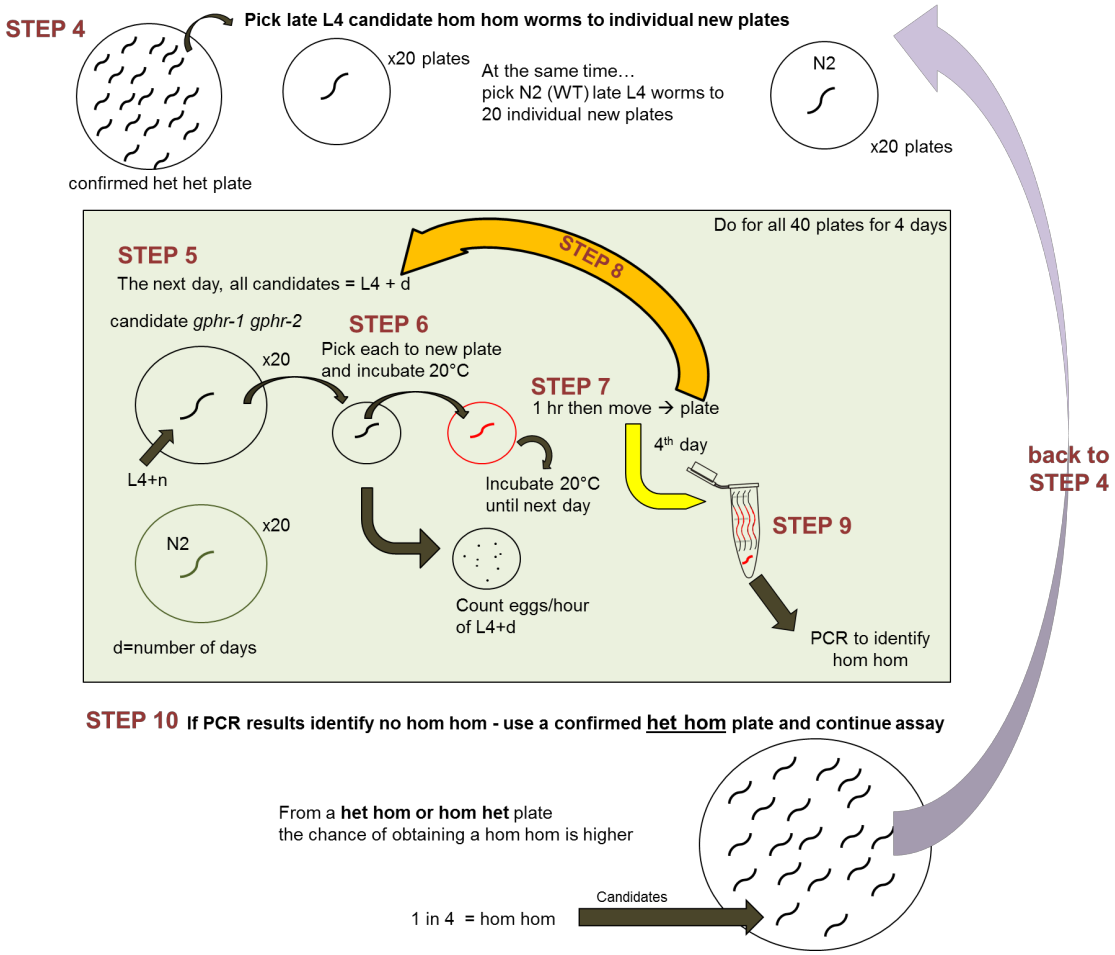


Figure 6.13 Method developed to generate and phenotype the *gphr* double mutant

Method follows steps 1 – 3 from Figure 6.8. 20 N2 (WT control) worms and 20 late L4-staged candidate *gphr-1(-/-) gphr-2(-/-)* worms from an identified *gphr-1(+/-) gphr-2(+/-)* plate (from Figure 6.8 **step 3**) were transferred to individual plates (**step 4**). It was expected to obtain one *gphr-1(-/-) gphr-2(-/-)* mutant in every 16 worms would be obtained. 24 h later (**step 5**), the 1-d old adults were transferred to individual, new plates and incubated for 1 h at 20 ºC (**step 6**) before transferring to another new plate (red) and incubated overnight (24 h) at 20 ºC (**step 7**). The next day (**step 8**) this plate transfer cycle was repeated (**steps 6 – 8**) until 4 d of assay was completed. On the 4th day, the potential homozygous animals in both genes were genotyped to identify a *gphr-1(-/-) gphr-2(-/-)* worm (**step 9**). If no *gphr-1(-/-) gphr-2(-/-)* candidates were identified, the egg laying assay may be continued by using a plate where the candidate worm was a confirmed *gphr-1(+/-) gphr-2(-/-)* or a *gphr-1(-/-) gphr-2(+/-)* worm (**step 10**). The eggs and progeny are quantified for all plates.

A significantly lower number of eggs were laid by the *gphr-1(ok1579) gphr-2(tm4228)* double mutant (see section 6.2.3). Of eggs that were laid, a significantly smaller percentage hatched compared to the WT (Figure 6.15). This was also observed previously although not quantified (see section 6.2.3). In light of the small proportion of eggs hatched (described in section 6.2.3), it was speculated that this could be due to a defect such as abnormal egg development or egg hatching. After following egg development across 2 d, it was discovered that many of the eggs that did not hatch arrested at the 30-cell stage of egg development (Figure 6.16).

6.2.3.2 L1 arrest phenotype observed in *gphr* double mutants

Of the small proportion of *gphr* double mutant eggs that hatched (section 6.2.3), none reached the adult stage of their life cycle (Figure 6.17), even after 10 d. This L1 arrest of *gphr-1(ok1579) gphr-2(tm4228)* double mutants was also observed in all the previous experiments although this phenotype was not quantified during the original isolation process (Table 6.2). Another principal finding was that on all occasions of isolating a double mutant, only the F_1 progeny of the *gphr-1(ok1579) gphr-2(tm4228)* double mutant exhibited larval arrest (Table 6.2). All F_0 *gphr-1(ok1579) gphr-2(tm4228)* double mutants were able to undergo a normal *C. elegans* life cycle whereby the normal molt cycles took place (Figure 1.11). This implies there is maternal rescue in which *gphr-1(+/-) gphr-2(+/-)* or *gphr-1(+/-) gphr-2(-/-)* worms were able to produce viable and fertile *gphr-1(ok1579) gphr-2(tm4228)* double mutants but F_1 worms arising from a *gphr-1(-/-) gphr-2(-/-)* double mutant worm were not.

6.2.3.3 Defective pharyngeal pumping phenotype in *gphr* double mutants

The feeding behaviour was examined in the larvae to see if this was affected compared to the WT. Measuring the pharyngeal pumping rate of the larvae monitored this. The pharyngeal pumping rate was measured in L1 larvae (Figure 6.18 A). The results demonstrate that the *gphr-1(ok1579) gphr-2(tm4228)* L1 larvae pumping rate was significantly lower than that of WT larvae (Figure 6.18 A). The pumping rate of adult single *gphr-1(ok1579)* and *gphr-2(tm4228)* mutants did not show marked differences between genotypes (Figure 6.18 B).

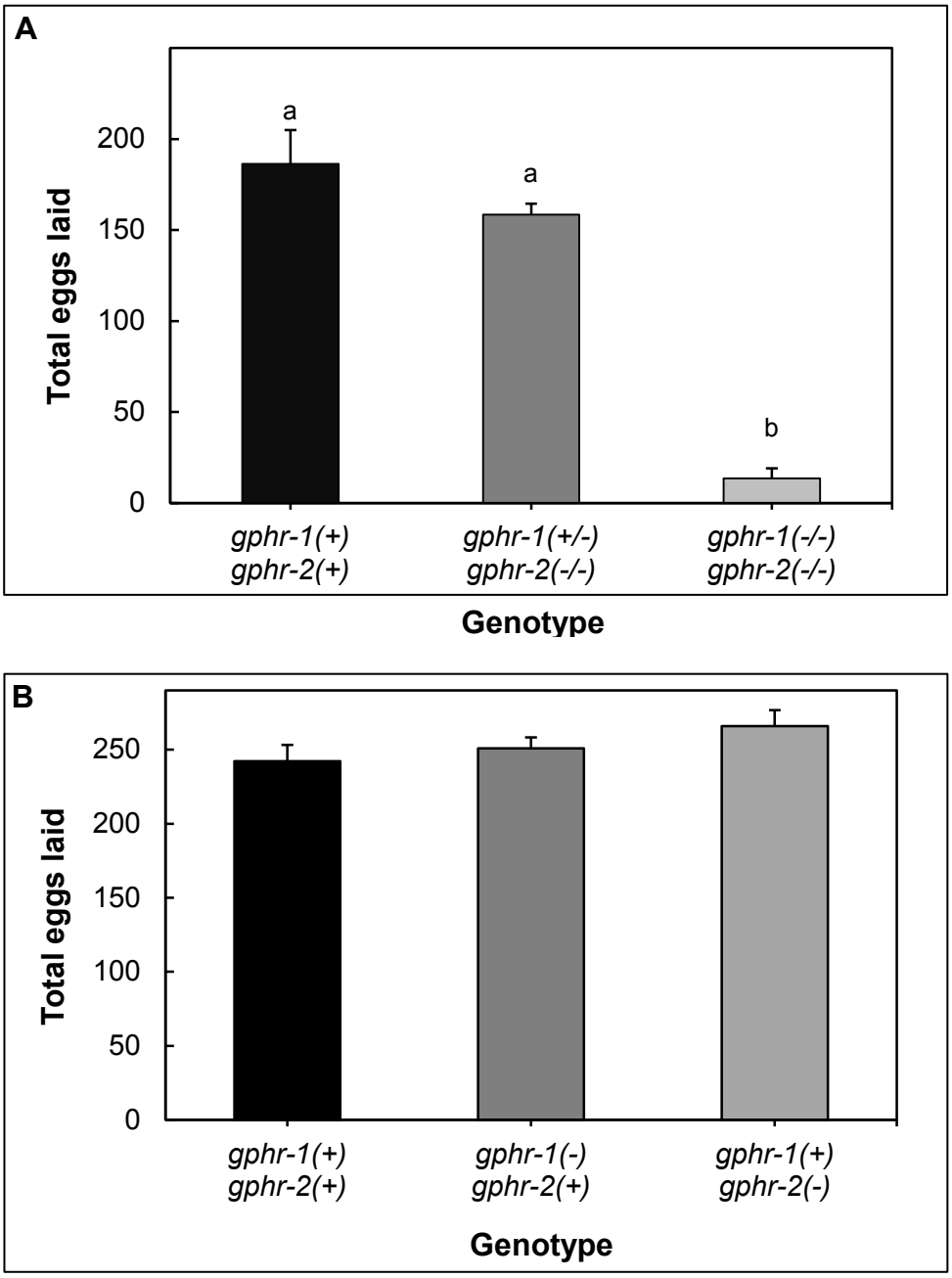


Figure 6.14 Egg laying defect in *gphr* double mutants

A. Mean total number of eggs laid over 4 d (+ S.E.). The number of worms for each genotype is 14 WT, seven *gphr-1(+/-) gphr-2(-/-)*, three *gphr-1(-/-) gphr-2(-/-)*. There is a significant difference between different genotypes and the total mean number of eggs laid; one-way ANOVA ($F_{2,21} = 6.88$, $p = 0.005$). Means not sharing a letter are significant different; Tukey *post-hoc* test. Data represents one of two independent experiments. **B**. Mean number of eggs laid over 4 d by single mutants (+S.E.); number of worms used for each strain was 20. There was no significant difference between the mean eggs laid by WT, or either single *gphr-1* or *gphr-2* mutants; one-way ANOVA ($F_{2,9} = 1.49$, $p = 0.276$).

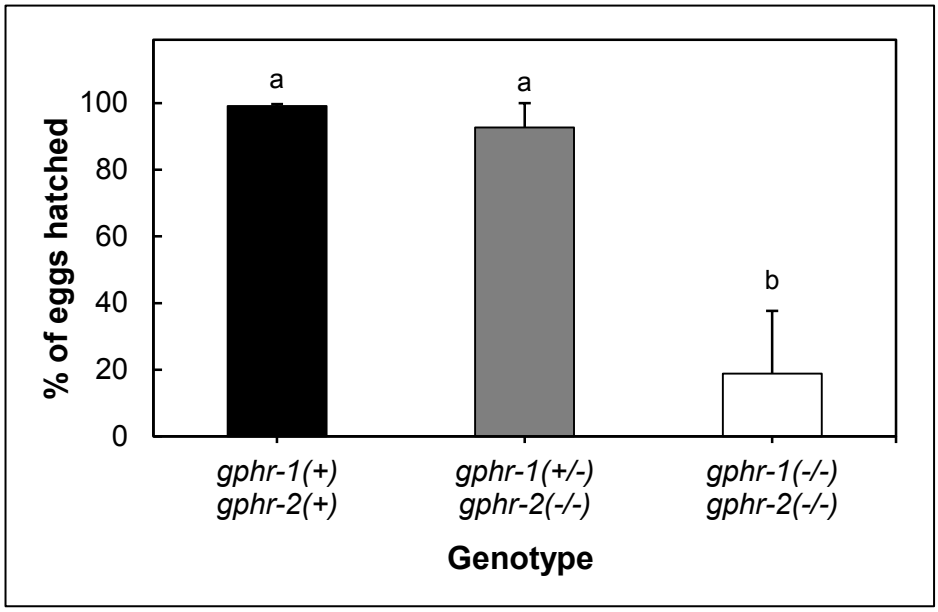


Figure 6.15 Egg hatching is reduced in the *gphr* double mutants

Percentage of hatched eggs laid by adults (+S.E.). The number of worms for each genotype is 14 WT, seven *gphr-1(+/-) gphr-2(-/-)*, three *gphr-1(-/-) gphr-2(-/-)*. There is a significant difference between the genotypes for the percentage of hatched eggs; Kruskal-Wallis test (H = 10.96, d.f. = 2, $p = 0.004$) where the mean ranks were 14.9 (WT), 12.1 (*gphr-1(+/-) gphr-2(-/-)* and 2 (*gphr-1(-/-) gphr-2(-/-)*). A Mann-Whitney test was used as *post-hoc* testing between each pair of groups and means not sharing a letter are significant, $p < 0.004$ (adjusted for ties). Data represents one of two independent experiments.

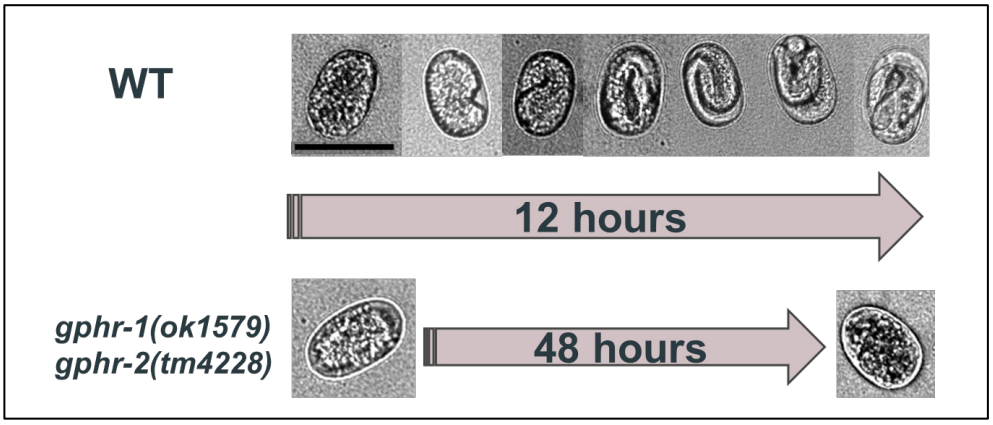


Figure 6.16 Images showing typical egg development of WTs and *gphr* double mutant

Egg development was followed from 1 h after being laid until larval eclosion. Larval eclosion was seen for WT but not for the *gphr-1(ok1579) gphr-2(tm4228)* double mutant. Scale bar indicates 50 µm. Images representative of one experiment, observations in seven eggs for *gphr-1(ok1579) gphr-2(tm4228)*.

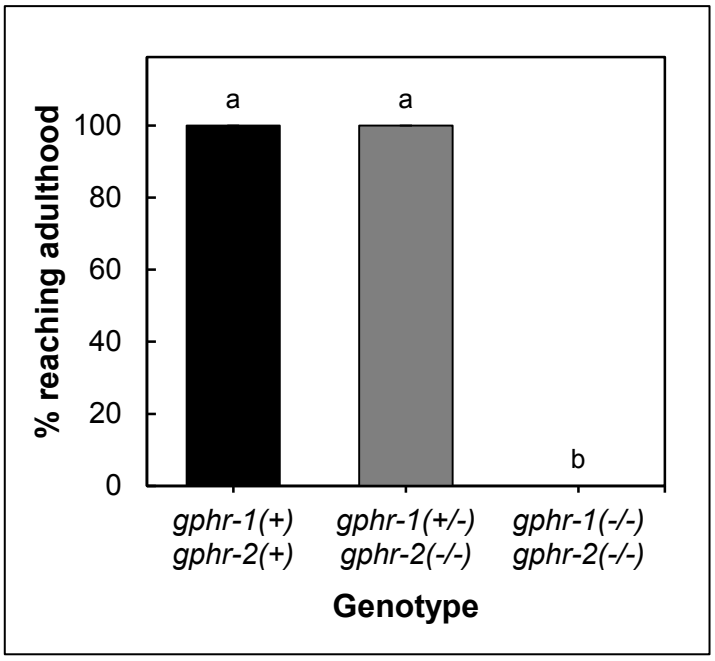


Figure 6.17 All *gphr* double mutants exhibit a larval arrest phenotype

Mean percentage of worms reaching adulthood was scored. The number of worms for each genotype is 14 WT, seven *gphr-1(+/-) gphr-2(-/-)*, three *gphr-1(-/-) gphr-2(-/-)*.There is a significant difference between the genotypes for the percentage of hatched eggs reaching larvae; Kruskal-Wallis test (H = 23, d.f. = 2, $p = 0.001$) where the mean ranks were 14 (WT), 14 (*gphr-1(+/-) gphr-2(-/-)* and two (*gphr-1(-/-) gphr-2(-/-)*). Mann-Whitney tests were used as *post-hoc* testing between each pair of groups and means not sharing a letter are significantly different, $p < 0.0001$ (adjusted for ties). Data represents one of two independent experiments but seen in all *gphr-1(-/-) gphr-2(-/-)* generated (Table 6.2).

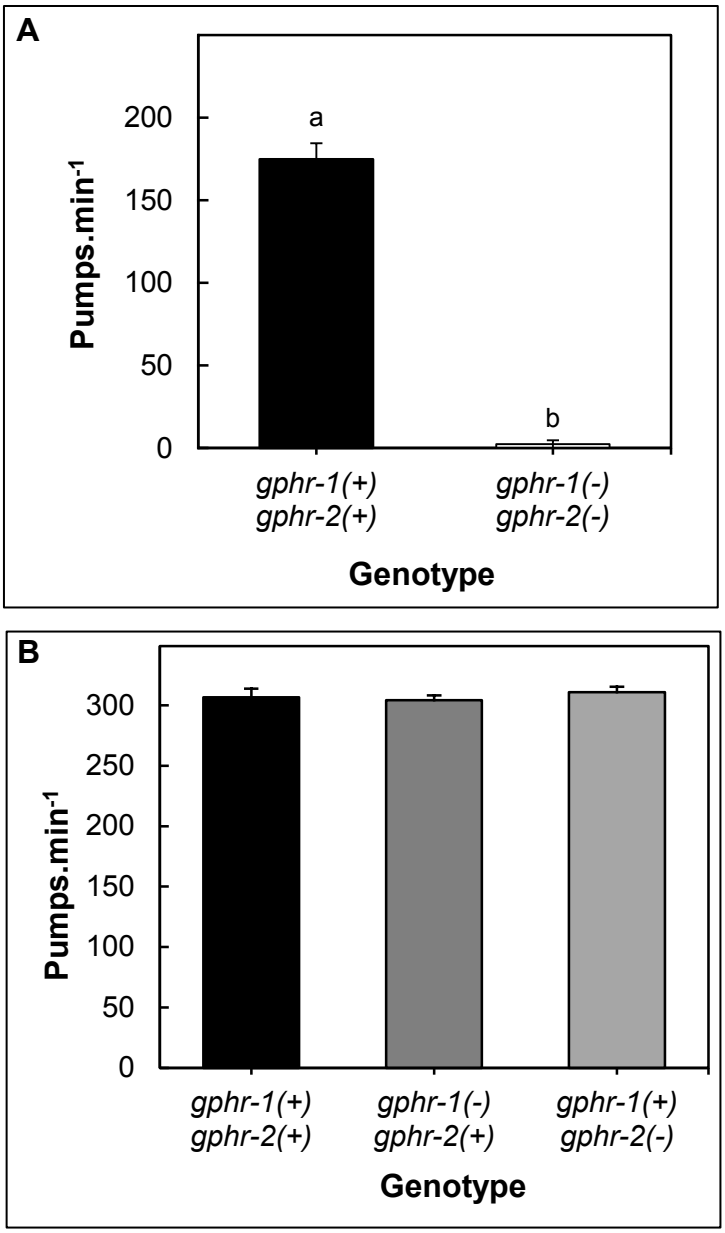


Figure 6.18 Pharyngeal pumping is reduced in *gphr* double mutants

A. Mean pharyngeal pumps.min^{-1} of L1 larvae (+ S.E.). There was a significant difference between the pumps.min^{-1} of WT (median = 175.00 pumps.min^{-1}, n = 7) and double *gphr-1 gphr-2* mutants (median = 0.001 pumps.min^{-1} , n = 6); Mann-Whitney test, $W = 70.0$, $p = 0.0026$ (adjusted for ties). Means not sharing a letter are significantly different, $p < 0.002$. **B**. Mean pharyngeal pumps.min^{-1} by 1-d old adult worms (+S.E.). There was no significant difference between the pumps.min^{-1} of WT, single *gphr-1* or *gphr-2* mutants; one-way ANOVA ($F_{2,9} = 2.59$, $p = 0.129$).

6.2.3.4 Identification of *C. elegans* RNAi feeder strains

To date, there are two *C. elegans* RNAi feeder collection libraries: the ORFeome-based RNAi library (Rual et al., 2004) and the RNAi feeding library constructed using genomic fragments (Kamath et al., 2003). Two RNAi feeding vectors were identified using both collections with both targeting *gphr-2*: pL4440 C11H1.2 ORF has the whole gene ORF cloned into the pL4440 feeding vector (Rual et al., 2004) and pL4440 C11H1.2 contains a genomic fragment (Kamath et al., 2003). Both were obtained and sequenced. pL4440 C11H1.2 contained the correct sequence, an 1887 bp region nearing the 3′ end of the gene. The ORF of *gphr-2* was not cloned into pL4440 as indicated by the lack of alignment of the sequencing data received to the CDS of the gene (with two independent repeats, both of which show a good sequencing signal). Therefore, only pL4440 C11H1.2 was used for the proceeding RNAi assays.

6.2.3.5 RNAi knockdown to generate a *gphr* double mutant

The RB1390x6 (*gphr-1(ok1579)*) strain was treated using: an RNAi feeder, pL4440 C11H1.2 which was induced using isopropyl β-D-1-thiogalactopyranoside (IPTG), and for the controls: the empty bacterial strain (HT115(DE3)), the induced empty vector (pL4440) and the RNAi feeder uninduced (pL4440 C11H1.2). The results indicate that there is no significant difference between the total number of eggs laid when fed with the empty vector and the pL4440 C11H1.2 vector either induced or uninduced (Figure 6.19 A). But there is a significant difference between the total number of eggs laid when fed the negative control, HT115(DE3) bacteria (Figure 6.19 A).

In comparison to the empty vector and uninduced vector controls, there was no significant decrease in the number of eggs laid when fed with induced pL4440 C11H1.2 compared to the empty vector and uninduced pL4440 C11H1.2 vector. However, there was no significant difference between the eggs laid when fed with HT115(DE3) bacteria and both the empty vector and uninduced pL4440 C11H1.2 vector (Figure 6.19 A). This result suggests that the pL4440 vector itself significantly increases the number of eggs laid compared to when it is absent. This could explain why there is no significant difference between the number of eggs laid by the worms when treated with just the HT115(DE3) bacteria and the induced pL4440 C11H1.2 group suggesting that there is a counteractive response where pL4440 increases egg-laying and knockdown of

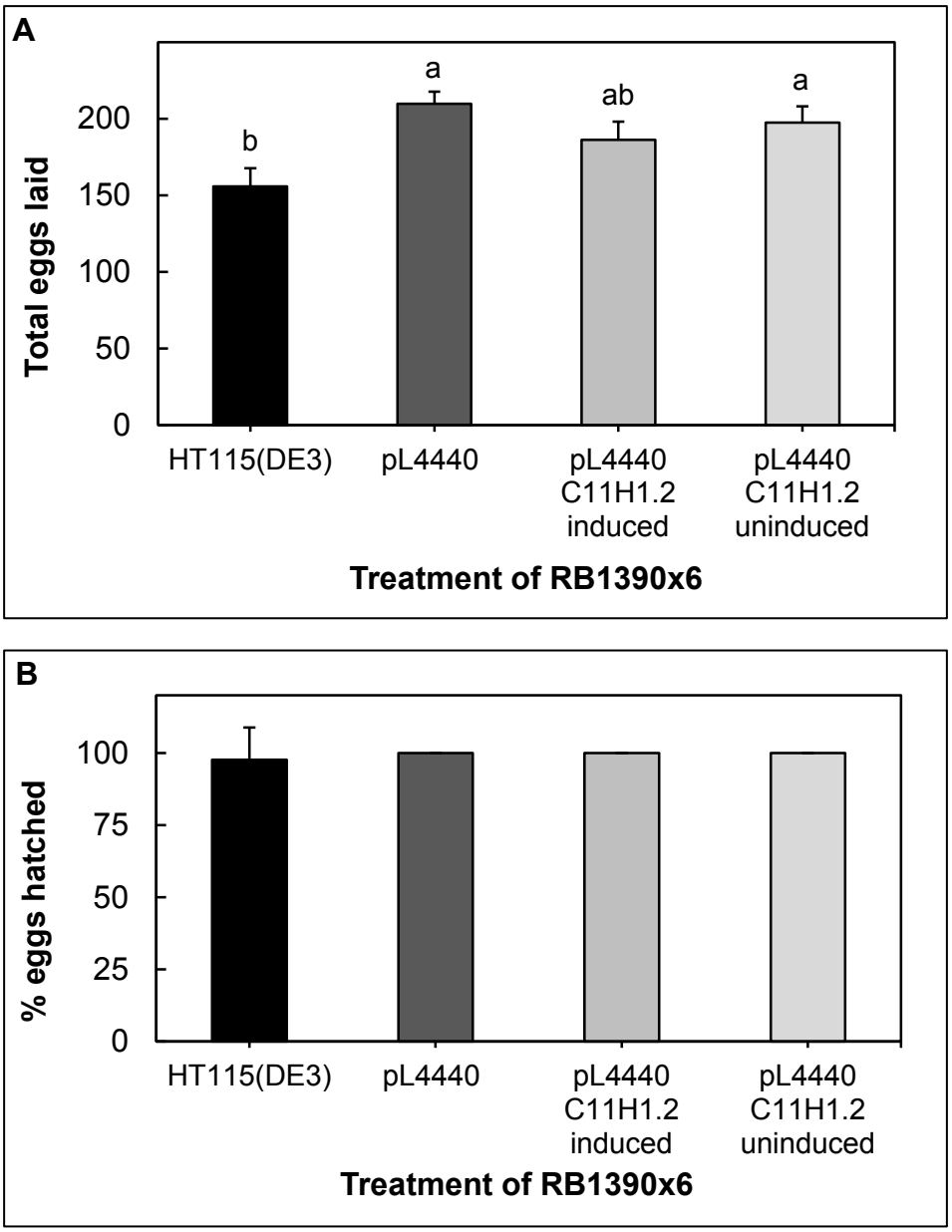


Figure 6.19 The RNAi feeding vector does not reduce number of eggs laid or eggs hatched

A. Mean total number of eggs laid by RB1390x6 (*gphr-1(ok1579)* single mutant strain on different feeding treatments (+S.E.). There is a significant difference between the feeding treatments and the total number of eggs laid; one-way ANOVA ($F_{3,34} = 4.46$, $p = 0.01$). Means not sharing a letter are significantly different; Tukey *post-hoc* test. **B.** Mean percentage of eggs hatched by RB1390x6 (+S.E.). There is no significant difference between the treatment and percentage of hatched eggs; Kruskal-Wallis test (H = 3.33, d.f. = 3, $p = 0.343$) where the mean ranks were 18.3 (HTT115(DE3)), 20.5 (pL4440, pL4440 C11H1.2 induced and pL4440 C11H1.2 uninduced).

gphr-2 decreases egg-laying such that the overall effect is as shown in Figure 6.19 A. Aside from an egg-laying defective phenotype seen in *gphr-1(ok1579) gphr-2(tm4228)* double mutants, an egg-hatching defective phenotype was also observed. Therefore, the percentage of eggs that hatched was also scored to see if a *gphr-1(ok1579)* KO *gphr-2(tm4228)* knockdown mutant would exhibit the same effect when both genes are knocked out as seen in section 6.2.3.1. The results show that the different treatments do not affect hatching behaviour (Figure 6.19 B). These results indicate that feeding the RNAi vector pL4440 C11H1.2 to RB1390x6 (*gphr-1(ok1579)* mutant strain) does not phenocopy the effects seen by knocking out both *gphr* genes in *C. elegans* described in section 6.2.3. It is important to note that RNAi feeding vectors are not always efficient in gene silencing. Therefore, when conducting these experiments, the pLT61 RNAi feeding vector was also used alongside other feeding treatments to check RNAi efficiency. pLT61 is designed to specifically knockdown *unc-22* causing worms to twitch (Timmons et al., 2001), and this anticipated behaviour was observed in all worms when carrying out this control.

6.2.4 Localisation of the GPHR proteins in *C. elegans*

The spatial and subcellular localisation of the *Cel*-GPHRs is yet to be determined. *C. elegans* expression vectors were made for transgenic *C. elegans* generation to investigate the localisation of the *Cel*-GPHRs. The Gateway system (section 5.2.1) was employed to generate the *C. elegans* expression vectors. The entry vectors pENTR/D *Cel-gphr-1* and pENTR/D *Cel-gphr-2* both without stop codons were made previously (section 5.2.1). No Gateway-compatible vectors were available with appropriate promoters and C-terminal fluorescent tagging. For example, pDEST-EGFP is a promoterless vector but allows for C-terminal enhanced GFP (EGFP) tagging for expression in *C. elegans*, while there are other destination vectors that drive tissue-specific expression but offers no C-terminal reporter. The pDEST-EGFP Gateway-compatible destination vector was kindly supplied by Prof. Y. Iino (University of Tokyo, Japan) for this project. Therefore, a specific procedure was designed to generate suitable constructs whereby *gphr* CDSs were tagged with EGFP under specific promoters (Figure 6.20 C).

The *gphr* CDSs were inserted into pDEST-EGFP via LR recombination. This was then used as a template to clone the *C. elegans gphrs* tagged with EGFP into TOPO

entry vector pENTR/D (Figure 6.20 A). Primers were designed to clone the *gphr* CDS tagged with EGFP. The predicted 2216 bp *gphr-1::egfp* (Figure 6.21 A, lane 2) and predicted 2231 bp *gphr-2::egfp* (Figure 6.21 A, lane 3) products were amplified and cloned into the pENTR/D entry vector. Restriction enzyme analysis confirmed that these genes were cloned in the correct orientation as the expected digest sizes 1526 and 3226 bp in lane 3 for pENTR/D *gphr-1::egfp* and, 1852 and 2885 bp in lane 4 for pENTR/D *gphr-2::egfp* were obtained (Figure 6.22 A and B).

These *gphr::egfp* sequences were inserted into *C. elegans* destination vectors pDEST-myo-2-p and pDEST-eft-3-p (Table 6.1) via LR recombination. Restriction analysis (Table 2.16) showed that both *gphr-1* and *gphr-2* tagged with EGFP were inserted correctly into the two vectors pDEST-myo-2-p and pDEST-eft-3-p (as above) generating the following constructs: *Peft-3::gphr-1::egfp*, *Peft-3::gphr-2::egfp*, *Pmyo-2::gphr-1::egfp* and *Pmyo-2::gphr-2::egfp*. These constructs can be used to explore the subcellular localisation of the *Cel*-GPHRs.

To explore the tissue localisation of the *Cel*-GPHRs, promoter-only reporter constructs (*Pgphr-1::egfp* and *Pgphr-2::egfp*) were generated using the Gateway system. *C. elegans* N2 genomic DNA was used to clone genomic fragments containing sequences that were predicted to be promoter regions for the *gphrs* (Figure 6.21 B). The predicted promoter products produced for *gphr-1* and *gphr-2* were 603 and 1190 bp respectively (Figure 6.21 B, lanes 2 and 3 respectively). These genomic sequences comprise the entire intergenic regions, 5′ UTRs, first exon and intron (Figure 6.20 A and B) and were cloned into pDEST-EGFP. Restriction enzyme digestion was used to confirm the entry and expression vectors generated (as above; Table 2.15 and Table 2.16).

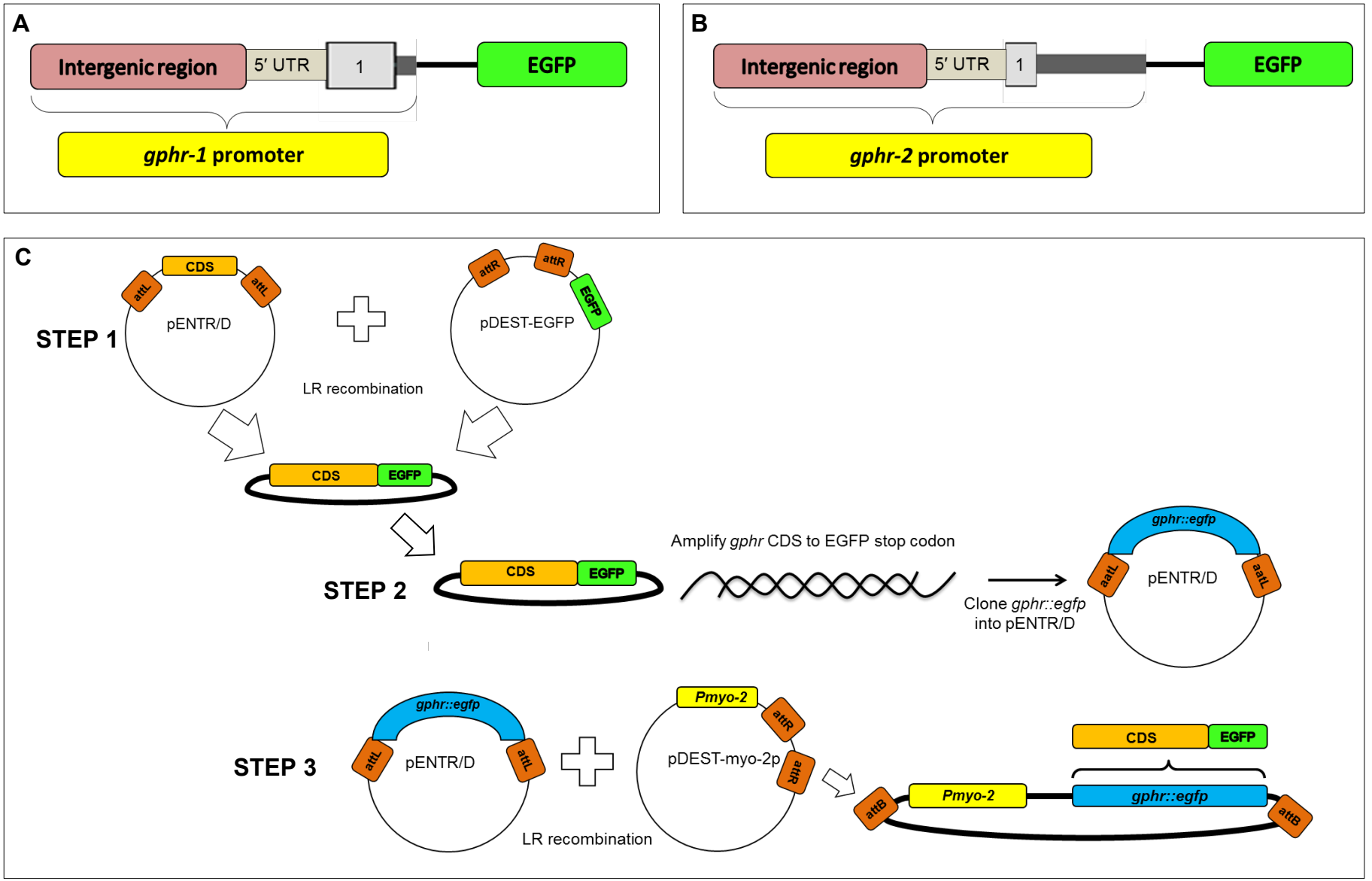
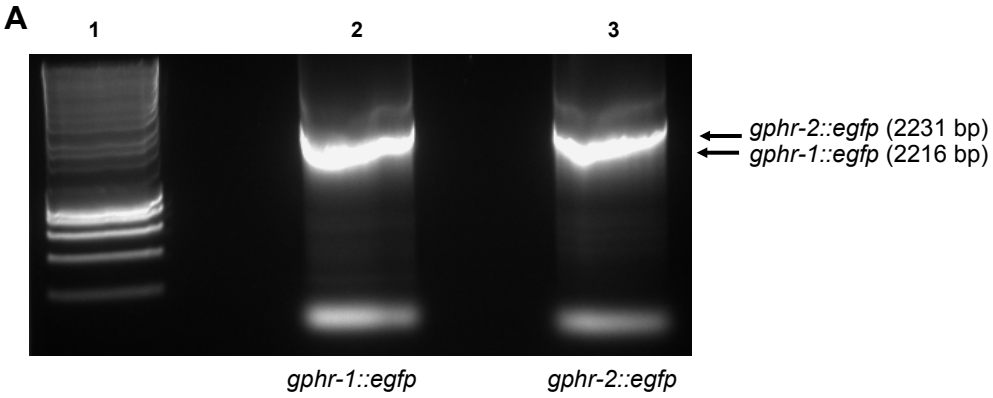
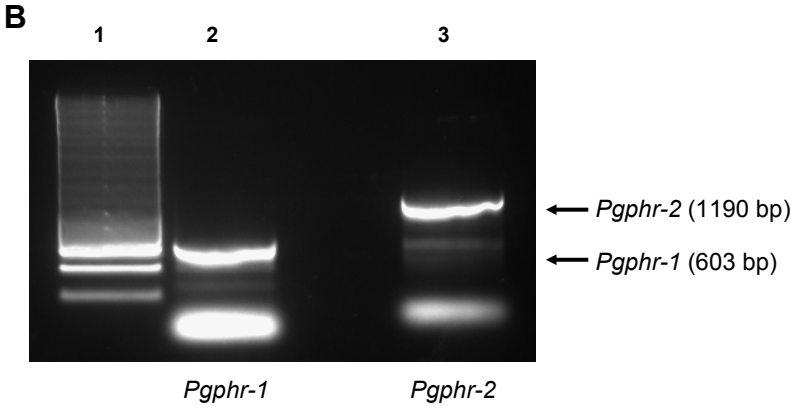
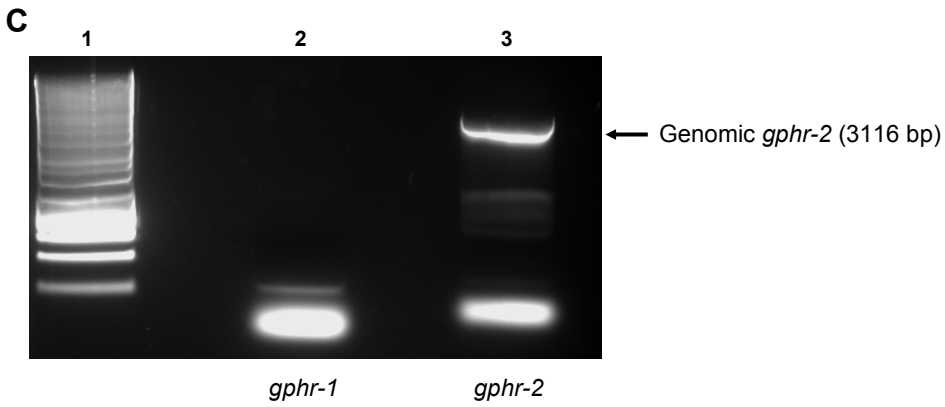

A
Intergenic region
5′ UTR
1
EGFP
gphr-1 promoter
B
Intergenic region
5′ UTR
1
EGFP
gphr-2 promoter
C
STEP 1
CDS
attL
attL
pENTR/D
attR
attR
EGFP
pDEST-EGFP
LR recombination
CDS
EGFP
STEP 2
CDS
EGFP
Amplify gphr CDS to EGFP stop codon
gphr::egfp
aatL
aatL
pENTR/D
Clone gphr::egfp
into pENTR/D
STEP 3
gphr::egfp
attL
attL
pENTR/D
Pmyo-2
attR
attR
pDEST-myo-2p
LR recombination
CDS
EGFP
attB
Pmyo-2
gphr::egfp
attB

Figure 6.20 Expression constructs used for *C. elegans* expression

A. Schematic of promoter-reporter construct for *gphr-1*. **B**. Schematic of promoter-reporter construct for *gphr-2*. **C**. Process developed for cloning *gphr* coding sequences (CDSs) tagged with EGFP into *C. elegans* expression vectors. Transfer CDSs of *gphr* genes (without stop codon) into pDEST-EGFP destination vector using LR recombination reaction (**STEP 1**). The resulting *gphr::egfp* constructs were used as a template for amplifying *gphr::egfp* to clone into pENTR/D entry vector (**STEP 2**). Used LR recombination reaction to transfer *gphr::egfp* sequences into *C. elegans* expression vectors containing appropriate promoters; *Pmyo-2* is shown here (**STEP 3**).


A
1
2
3
gphr-2::egfp (2231 bp)
gphr-1::egfp (2216 bp)
gphr-1::egfp
gphr-2::egfp


B
1
2
3
Pgphr-2 (1190 bp)
Pgphr-1 (603 bp)
Pgphr-1
Pgphr-2


C
1
2
3
Genomic gphr-2 (3116 bp)
gphr-1
gphr-2

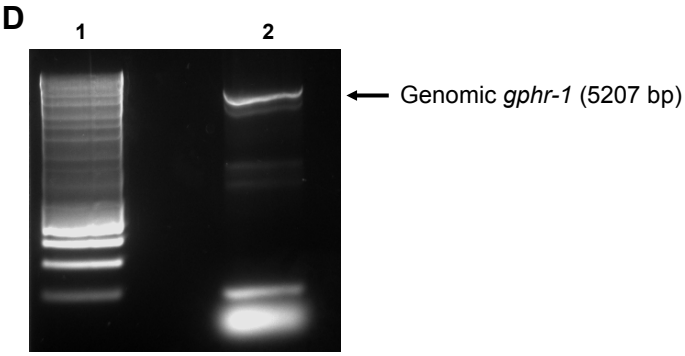

D
1
2
Genomic gphr-1 (5207 bp)

Figure 6.21 *gphr* gene products cloned into pENTR/D entry vector

PCR and gel electrophoresis to show the amplified *gphr* gene products that were cloned into pENTR/D entry vectors. Predicted product sizes and lanes are shown on the figure (see Figure 6.5 A for predicted molecular marker sizes). **A**. CDS of the *gphrs* tagged with EGFP. Lanes: 1 = ladder, 2 = *gphr-1* tagged with EGFP, 3 = *gphr-2* tagged with EGFP. **B**. The *gphr* promoters. Lanes: 1 = ladder, 2 = promoter *of gphr-1*, 3 = promoter of *gphr-2.* **C**. Whole genomic gene products of the *gphrs*. Lanes: 1 = ladder, 2 = *gphr-1*, 3 = *gphr-2*. **D**. Whole genomic gene product of *gphr-1*. Lanes: 1 = ladder, 2 = *gphr-1*.

The whole genomic sequences of the *gphrs* were also amplified to clone into *C. elegans* destination vectors to explore the endogenous localisation of these proteins. The predicted gene size for *gphr-1* is 5207 bp and did not amplify at the first attempt (Figure 6.21 C, lane 2) but a product was produced after another attempt, indicating that the primers, DNA template and parameters were correct (Figure 6.21 D, lane 2). The predicted gene size for *gphr-2* is 3116 bp and the expected product size for this sequence was produced (Figure 6.21 C lane 3). *gphr-2* was successfully cloned into pDEST-EGFP destination vector to generate the construct *Pgphr-2::gphr-2::egfp* (Figure 6.22 C). Cloning of *gphr-1* into pENTR/D entry vector failed despite several attempts, possibly due to the gene size (5207 bp).

The *Peft-3::gphr-1::egfp*, *Peft-3::gphr-2::egfp*, *Pmyo-2::gphr-1::egfp*, *Pmyo-2::gphr-2::egfp* constructs were sequenced and results confirm that both *gphr* sequences corresponded to the CDSs provided by WormBase (*gphr-1* = Y75B8A.16, WBGene00013551; *gphr-2* = C11H1.2, WBGene00007528). Furthermore, the sequencing data showed that EGFP could be read in the same frame, confirming that the genomic sequence and CDS and were successfully tagged with EGFP at the C-terminus.

Each construct was coinjected with ER marker construct, *Peft-4::mCherry::sp12* and Golgi marker construct, *Paman-2::aman-2::mCherry* independently into WT, N2 worms. The markers form tandem arrays with the construct of interest, therefore can be used for positive selection for worms carrying the extrachromosomal array. Positive transformants will exhibit mosaic expression since the extrachromosomal arrays are not stably integrated into the genome. Several transgenic *C. elegans* strains were generated (Table 6.5), while unsuccessfully transformed constructs are shown in Table 6.6. Positive transformants were selected and the strains successfully generated are shown in Table 6.5. Confocal microscopy was used to visualise the *C. elegans* GPHR expression, but first it was important to confirm that the green fluorescence emission was characteristic of EGFP, and not autofluorescence.

To do this, lambda scans were carried out on transformed and untransformed worms (Figure 6.23). The *Pmyo-2::gphr-1::egfp* construct targets expression exclusively to pharyngeal muscle cells (Dibb et al., 1989), therefore green fluorescence should not be observed in additional regions of the worm. However, some worms exhibited green fluorescence in the gut; this was found to be autofluorescence,

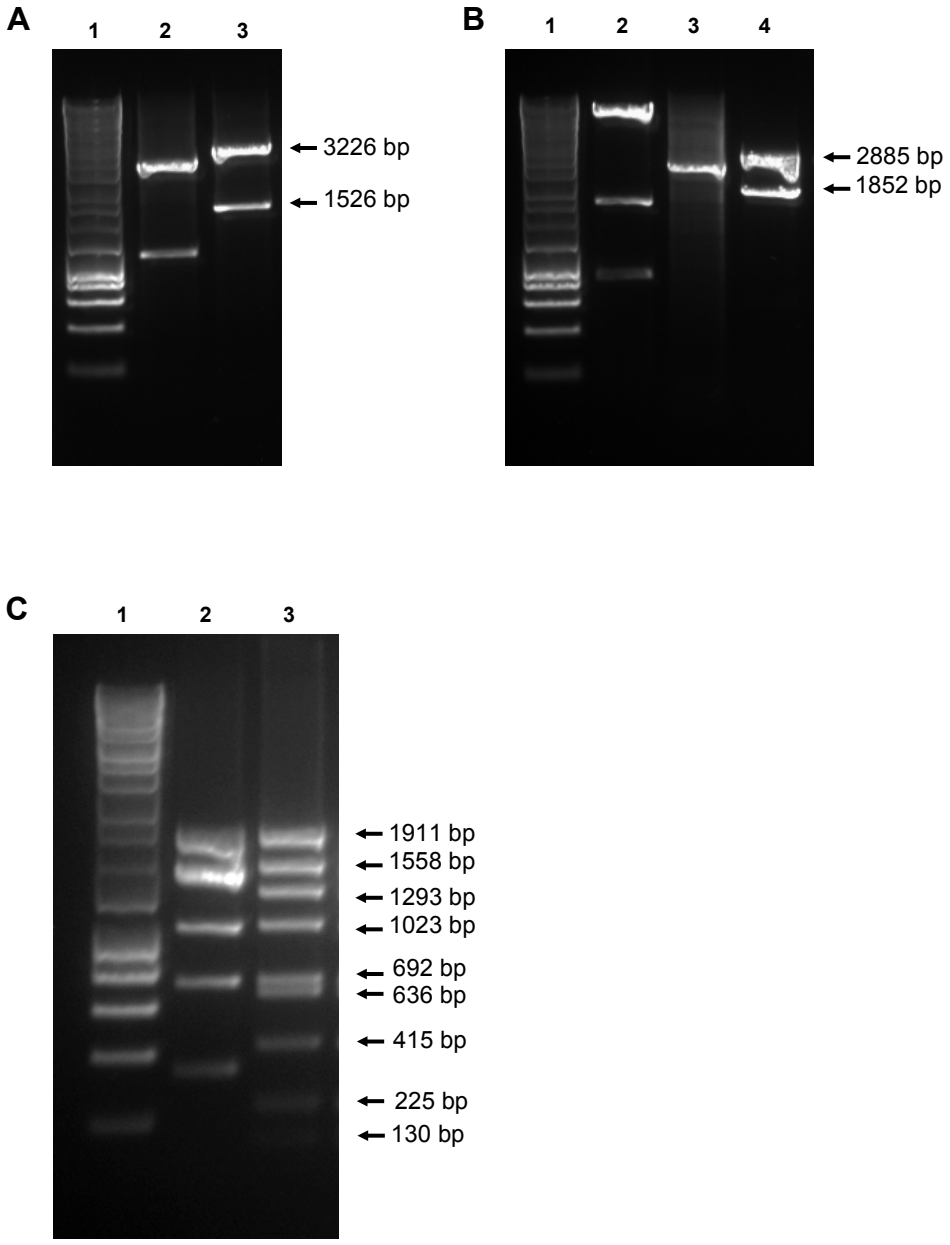


Figure 6.22 Analysis of *Cel-gphrs* cloned into entry and destination vectors

A and **B**. Analysis of *gphr-1::egfp* and *gphr-2::egfp* cloned into pENTR/D entry vector using restriction enzymes *Eco*RV and *Xho*I. Predicted product sizes and lanes are labelled on figures (see Figure 6.5 A for predicted molecular marker sizes). Lane 1: show molecular markers. **A**. pENTR/*D gphr-1::egfp* analysis. *gphr-1::egfp* cloned into pENTR/D successfully (lane 3) and unsuccessfully (lane 2). **B**. pENTR/D *gphr-2::egfp* analysis. *gphr-2::egfp* cloned into pENTR/D (lane 4) with negative, pENTR/D (lane 3) and positive, pAG426GAL-ccdB-EGFP (lane 2) controls displayed (predicted band sizes not shown). **C**. Analysis of *Pgphr-2::gphr-2::egfp* using *Eco*RV and *Xho*I. Whole genomic sequence of *gphr-2* gene cloned into pDEST-EGFP (lane 3) with negative control, pDEST-EGFP (lane 2 – predicted product sizes not shown).

Table 6.5 Transgenic *C. elegans* lines produced for studying GPHR localisation

Extrachromosomal arrays in the N2 strain. An explanation of *C. elegans* transgenic line nomenclature is described in section 2.1.

Independent transgenic line	Aim	Coinjection marker
N2*Ex*1[*Pmyo-2::gphr-1::egfp*]	Subcellular localisation	*Peft-4::mCherry::sp12*
N2*Ex*2[*Pmyo-2::gphr-1::egfp*]	Subcellular localisation	*Peft-4::mCherry::sp12*
N2*Ex*1[*Pmyo-2:: gphr-2::egfp*]	Subcellular localisation	*Peft-4::mCherry::sp12*
N2*Ex*2[*Pmyo-2:: gphr-2::egfp*]	Subcellular localisation	*Peft-4::mCherry::sp12*
N2*Ex*3[*Pmyo-2:: gphr-2::egfp*]	Subcellular localisation	*Peft-4::mCherry::sp12*
N2*Ex*1[*Pgphr-2::gphr-2::egfp*]	Tissue and subcellular localisation	*Paman-2::aman-2::mCherry*
N2*Ex*2[*Pgphr-2::gphr-2::egfp*]	Tissue and subcellular localisation	*Paman-2::aman-2::mCherry*
N2*Ex*1[*Pgphr-1::egfp*]	Tissue localisation	*Peft-4::mCherry::sp12*
N2*Ex*2[*Pgphr-1::egfp*]	Tissue localisation	*Peft-4::mCherry::sp12*
N2*Ex*3[*Pgphr-1::egfp*]	Tissue localisation	*Peft-4::mCherry::sp12*
N2*Ex*4[*Pgphr-1::egfp*]	Tissue localisation	*Peft-4::mCherry::sp12*
N2*Ex*1[*Pgphr-2::egfp*]	Tissue localisation	*Peft-4::mCherry::sp12*
N2*Ex*2[*Pgphr-2::egfp*]	Tissue localisation	*Peft-4::mCherry::sp12*
N2*Ex*3[*Pgphr-2::egfp*]	Tissue localisation	*Peft-4::mCherry::sp12*
N2*Ex*4[*Pgphr-2::egfp*]	Tissue localisation	*Peft-4::mCherry::sp12*

Table 6.6 Attempted but unsuccessful transformations into N2 *C. elegans*

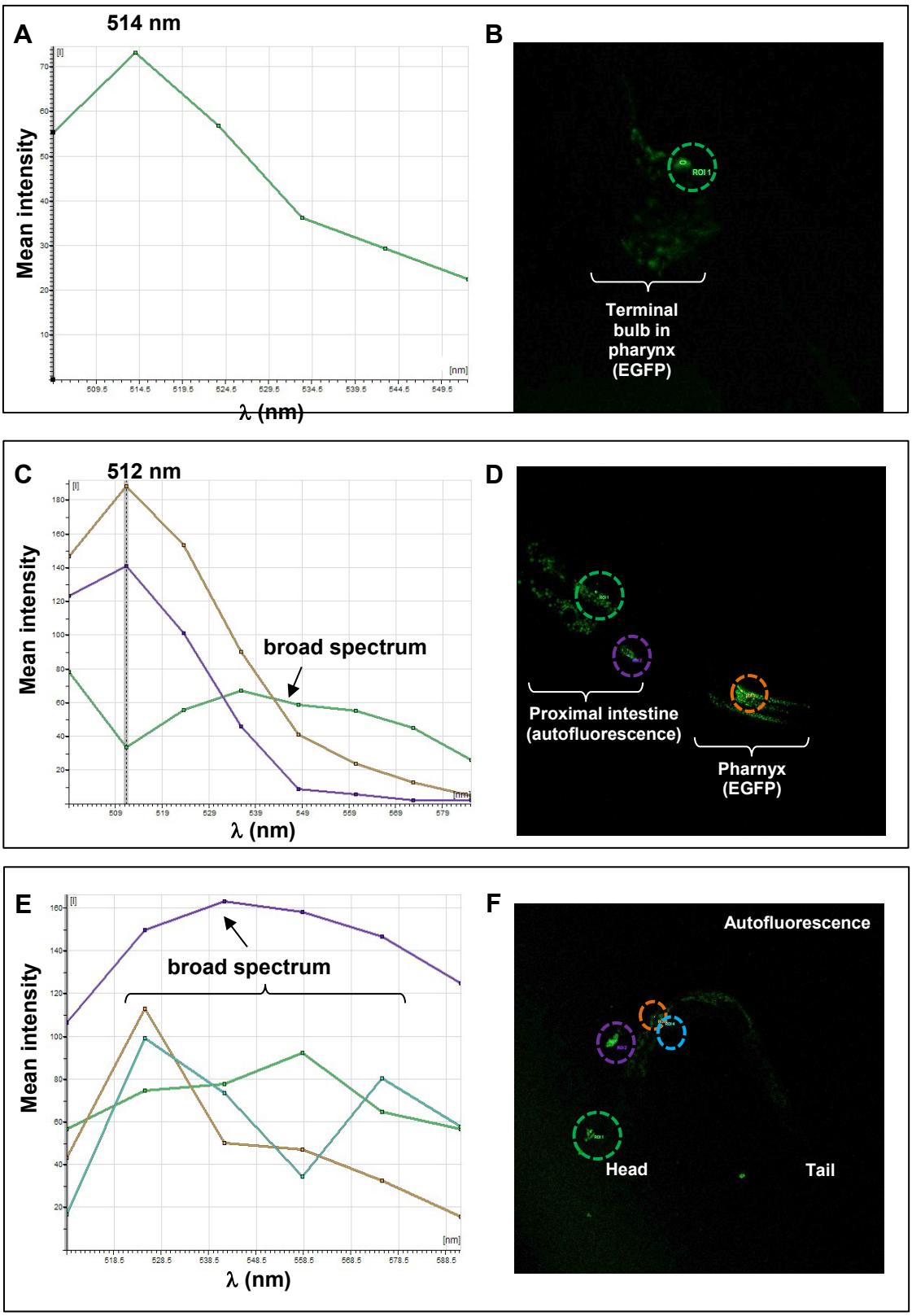
Construct	Coinjection marker	Aim
Pmyo-2::gphr-1::egfp	*Paman-2::aman-2::mCherry*	Subcellular localisation
Pmyo-2::gphr-2::egfp	*Paman-2::aman-2::mCherry*	Subcellular localisation
Peft-3::gphr-1::egfp	*Paman-2::aman-2::mCherry*	Subcellular localisation
Peft-3::gphr-1::egfp	*Peft-4::mCherry::sp12*	Subcellular localisation
Peft-3::gphr-2::egfp	*Paman-2::aman-2::mCherry*	Subcellular localisation
Peft-3::gphr-2::egfp	*Paman-2::aman-2::mCherry*	Subcellular localisation
Pgphr-2::gphr-2::egfp	*Peft-4::mCherry::sp12*	Tissue and subcellular localisation

and was observed commonly in untransformed N2 worms (Figure 6.23). Lambda scans profile a fluorescent sample across a designated range of wavelengths to determine spectral properties of fluorescence. Yellow arrowheads on Figure 6.23 – Figure 6.25 indicate green autofluorescence in gut granules. Similar emission profiles were observed in Figure 6.25 confirming that these areas are indeed autofluorescence, especially since *Pmyo-2* targets expression in pharyngeal muscle cells only (Dibb et al., 1989).

Many fluorescent images are shown in this section as extrachromosmal arrays can exhibit mosaic expression. Therefore, these images show an overview of the localisation results. *Pmyo-2::gphr-1::egfp* expression was targeted to the pharyngeal muscle cells as expected and appeared to be expressed predominantly in the corpus and terminal bulb pharyngeal muscle cells, with weak expression in the isthmus (Figure 6.24 and Figure 6.25). This is consistent in two independent, extrachromosomal lines. Under the *eft-4* promoter, *mCherry::sp12* is as expected, ubiquitously active with strong expression in the digestive and nervous system, pharynx and body wall muscle regions of the adult worm (Figure 6.24). Figure 6.25 shows that *gphr-1* does not colocalise with the ER marker SP12, signifying that *gphr-1* is not an ER membrane protein in pharyngeal muscle cells.

The subcellular and tissue localisation of endogenous GPHR-2 remains inconclusive. The construct (*Pgphr-2::gphr-2::egfp*) designed for investigating the native localisation of GPHR-2 was coinjected with the Golgi marker (*Paman-2::aman-2::mCherry*). The Golgi marker is indicative of positive transformation thus the worms were successfully transformed (Figure 6.26). However, no green fluorescence of the EGFP emission profile was detected and this fluorescence was indicative of autofluorescence (Figure 6.23). This could be as a result of a weak promoter for *gphr-2*.

To investigate tissue localisation of the GPHR proteins, the constructs *Pgphr-1::egfp* and *Pgphr-2::egfp* were coinjected with ER marker (*Peft-4::mCherry::sp12*) as they are ubiquitously expressed. mCherry fluorescence of the ER marker was consistently expressed in 4 independent, extrachromosomal lines for each construct, but no green fluorescence of the EGFP emission profile was detected and rather it was autofluorescence that was observed, analogous to that of the *Pgphr-2::gphr-2::egfp* lines (Figure 6.23 I and J).


A
514 nm
Mean intensity
λ (nm)
B
Terminal bulb in pharynx (EGFP)
C
512 nm
Mean intensity
broad spectrum
λ (nm)
D
Proximal intestine (autofluorescence)
Pharnyx (EGFP)
E
broad spectrum
Mean intensity
λ (nm)
F
Autofluorescence
Head
Tail

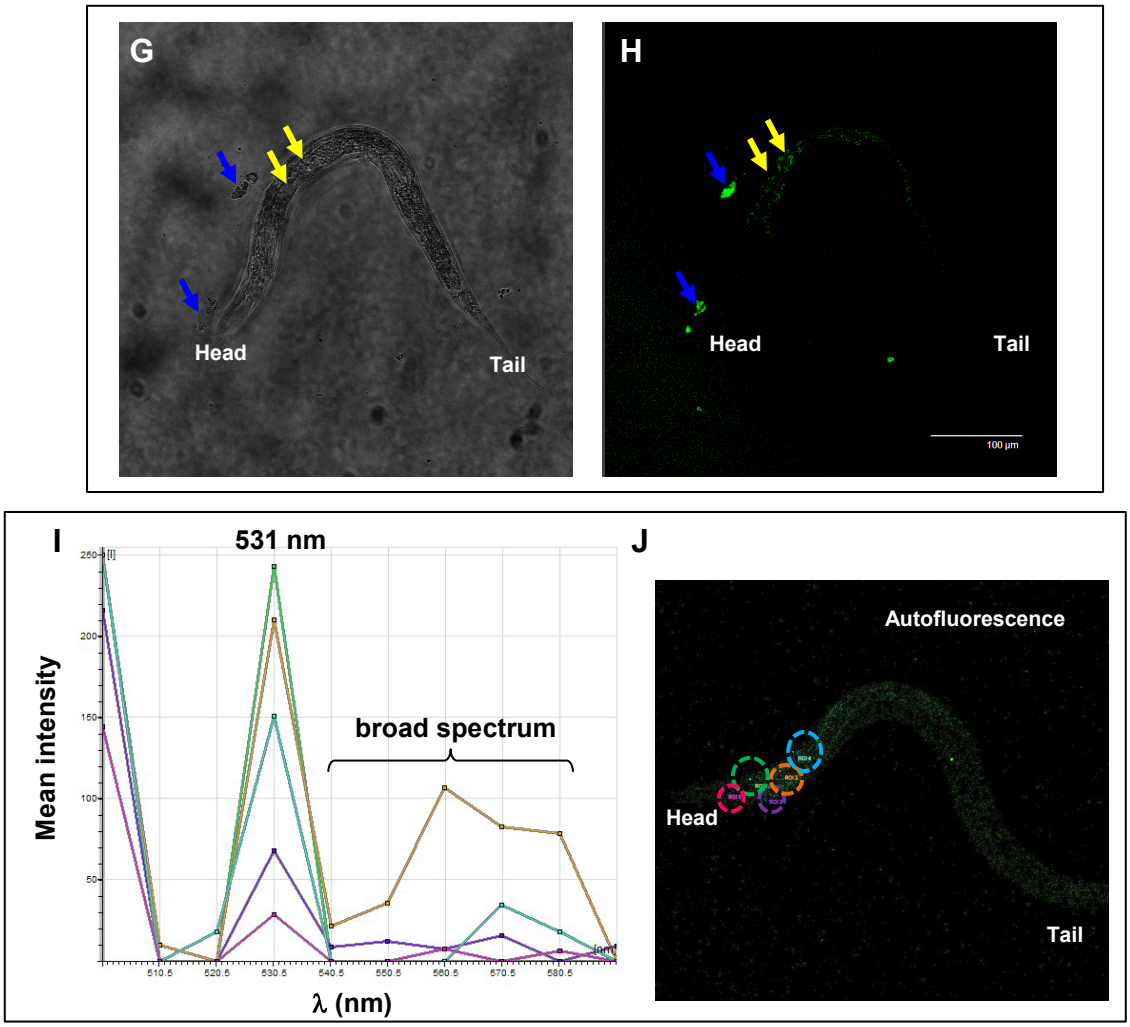
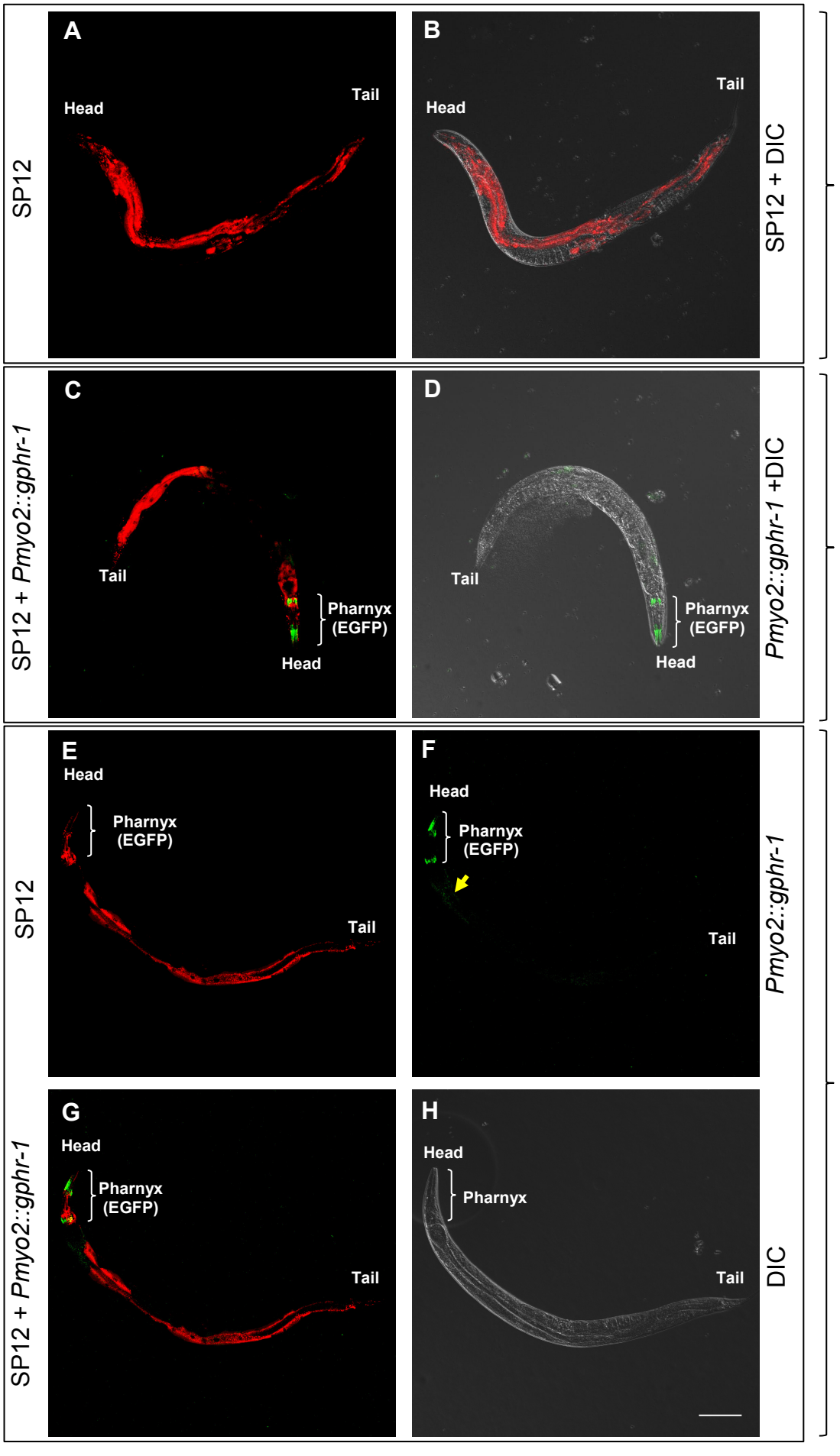


Figure 6.23 Testing for autofluorescence in *C. elegans*

Lambda scans across wavelengths 500 – 590 nm for analysing the emission spectra against mean intensitities of green fluorescence detected in the pharynx and proximal intestine of transformed with *Pmyo-2::Cel-gphr-1::egfp* (**A** – **D**), untransformed (**E** – **H**) and transformed with *Pgphr-2::gphr-2::egfp* (**I** – **J**) N2 worms. Coloured eclipses (labelled region of interests (ROIs) on **B**, **D**, **F** and **J)** correspond to the equivalent emission lines in graphs (**A**, **C**, **E** and **I**, respectively). ROIs analysed are very small; so are highlighted using coloured dotted circles matching original ROIs and emission spectra lines of corresponding graphs. EGFP emission peaks at 511 – 513 nm (**A** and **C**) in the pharynx (*Pmyo-2::Cel-gphr-1::egfp* expression), while broader emission profiles (**E** and **I**) or peak at a different wavelength (531 nm for **I**) indicate green autofluorescence, particularly in the intestine (**D**, **F**, **H** and **J**). **G – H**. Brightfield and fluorescent images showing autofluorescent profiles without ROI labels. Yellow (inside worm) and blue (debris outside worm) arrows indicate green autofluorescence, identified by lambda analysis (**E** and **F)**. Pharynx labelled where there is *gphr-1::egfp* expression in the muscle cells (transformed with *Pmyo-2::Cel-gphr-1::egfp*). **I – J**. Emission profile (**I**) and fluorescent (**J**) image showing autofluorescence and no *Pgphr-2::gphr-2::egfp*.


A
Head
Tail
B
Head
Tail
SP12
SP12 + DIC
C
Tail
Pharnyx
(EGFP)
Head
D
Tail
Pharnyx
(EGFP)
Head
SP12 + Pmyo2::gphr-1
Pmyo2::gphr-1 +DIC
E
Head
Pharnyx
(EGFP)
Tail
F
Head
Pharnyx
(EGFP)
Tail
SP12
Pmyo2::gphr-1
G
Head
Pharnyx
(EGFP)
Tail
H
Head
Pharnyx
Tail
SP12 + Pmyo2::gphr-1
DIC

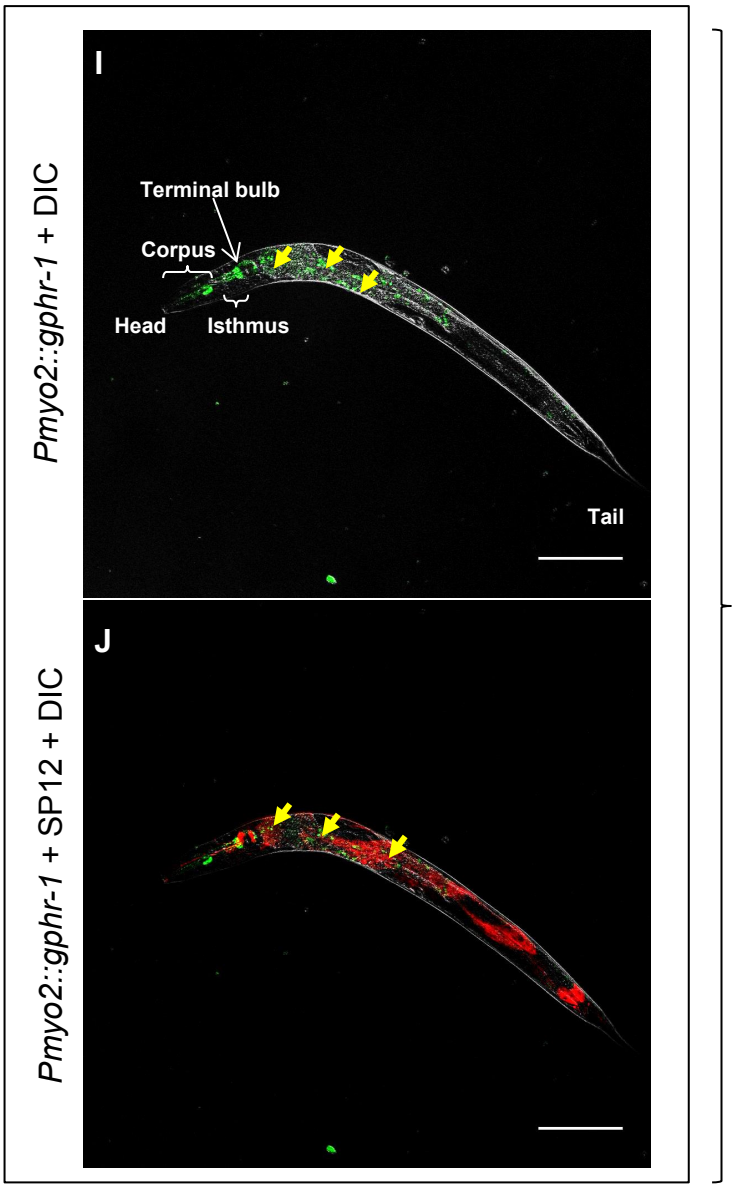
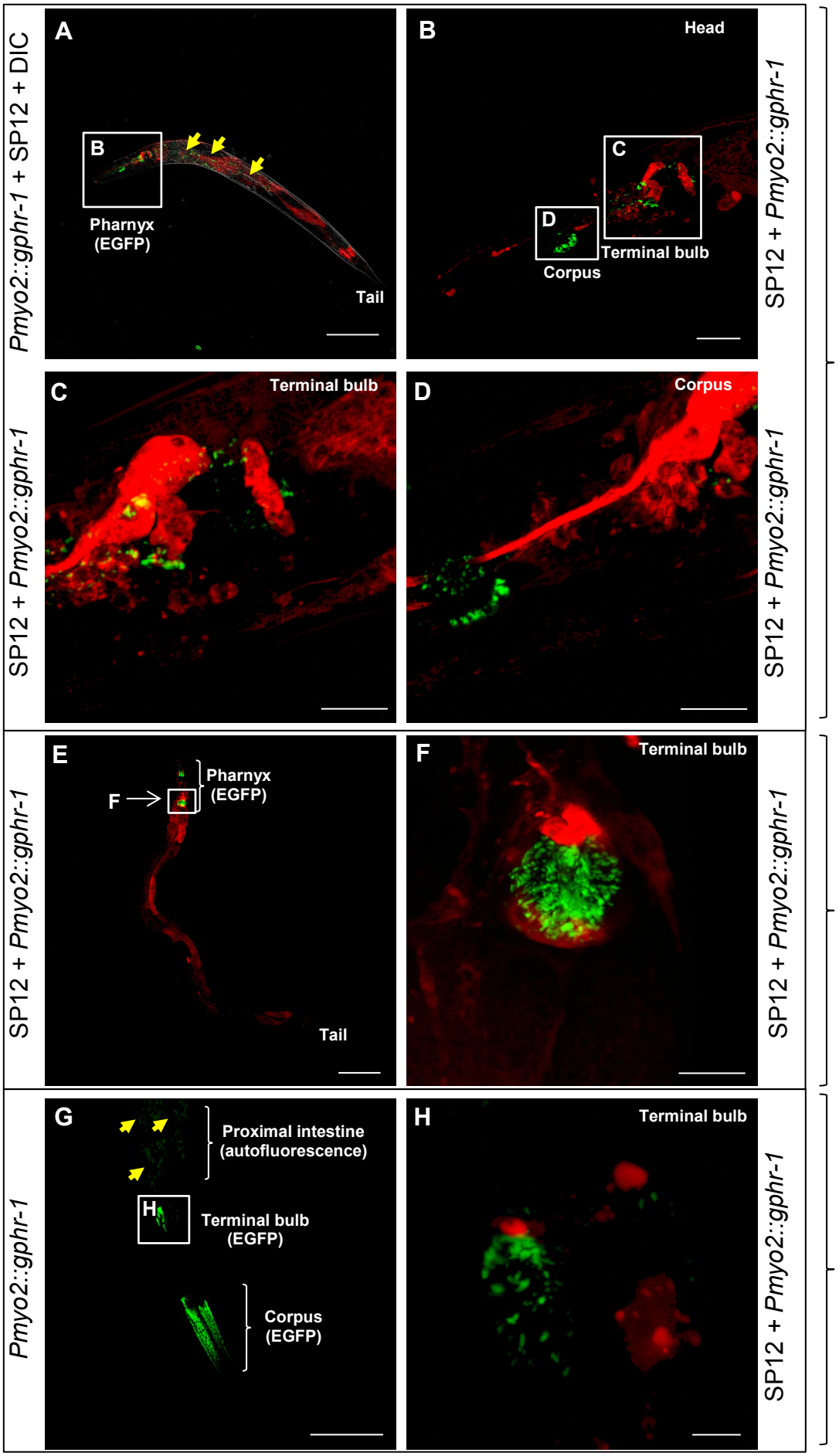
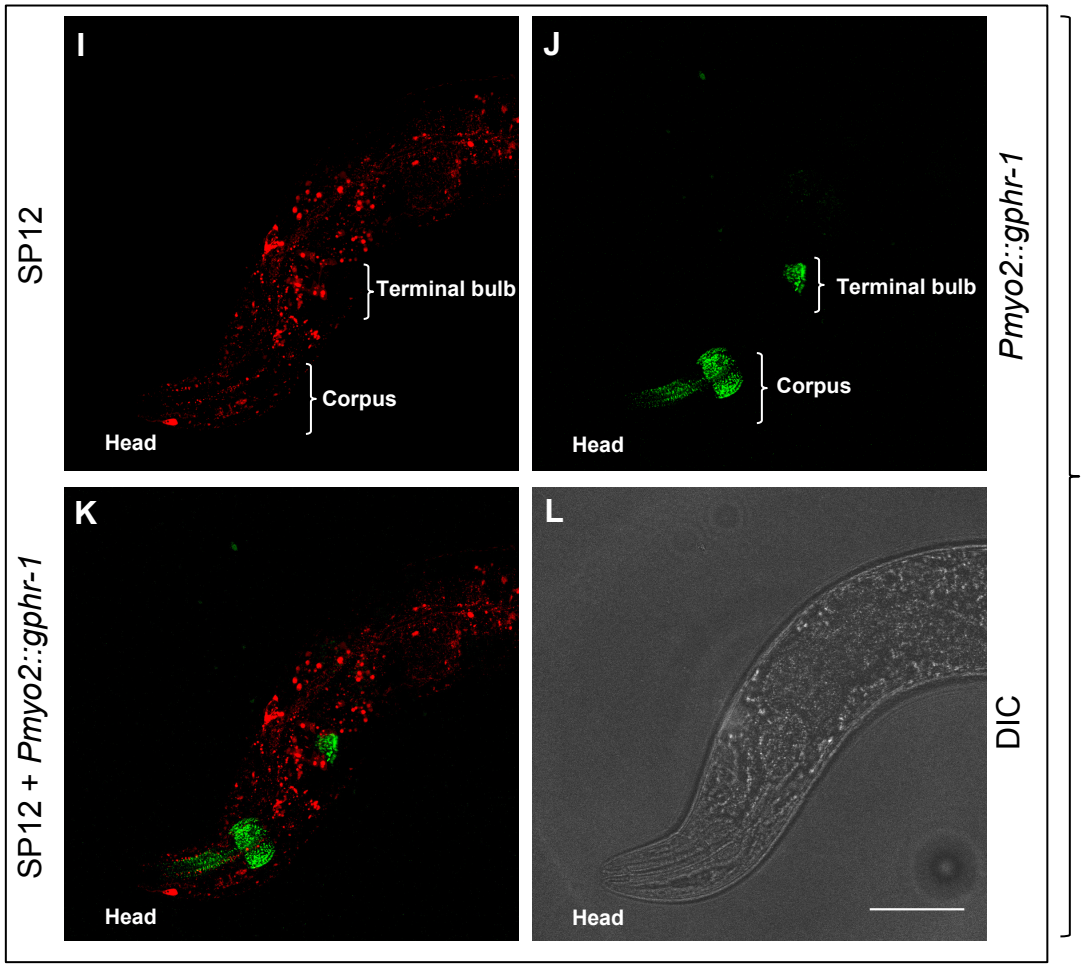


Figure 6.24 The expression of *Pmyo-2::gphr-1::egfp* in *C. elegans*

Transgenic N2 adult worms displaying extrachromosomal arrays of *Pmyo-2::gphr-1::egfp* with marker *Peft-4::mCherry::sp12*. Images from the same worm are grouped by a bracket on the right of the panel. **A** and **B**. ER marker, *mCherry::sp12* (red) under a ubiquitous promoter, *Peft-4*; showing expression in the pharynx, head and tail region, nervous and digestive system, indicated by fluorescent (**A**) and DIC overlaid image (**B**). **C** – **D**. Another worm showing *gphr-1::egfp* (green) expression under the pharyngeal muscle promoter, *Pmyo-2*; merged with fluorescent image of *mCherry::sp12* expression . **E** – **H**. Fluorescent images of *mCherry::sp12* expression (**E**), *gphr-1::egfp* expression (**F**), merged *mCherry::sp12* and *gphr-1::egfp* expression (**G**) and DIC (**H**) of the same worm. Images are representative of two stable extrachromosomal lines of more than 15 worms. **I** – **J**. Another worm showing *gphr-1::egfp* expression in the terminal bulb, isthmus and corpus of the pharynx (labelled). All confocal fluorescent images are maximum projections. Scale bar: 100 μm.


A
B
Pharnyx
(EGFP)
Tail
B
Head
C
D
Terminal bulb
Corpus
C
Terminal bulb
D
Corpus
E
F
Pharnyx
(EGFP)
Tail
F
Terminal bulb
G
Proximal intestine
(autofluorescence)
H
Terminal bulb
(EGFP)
Corpus
(EGFP)
H
Terminal bulb
Pmyo2::gphr-1 + SP12 + DIC
SP12 + Pmyo2::gphr-1
SP12 + Pmyo2::gphr-1
SP12 + Pmyo2::gphr-1
SP12 + Pmyo2::gphr-1
SP12 + Pmyo2::gphr-1
Pmyo2::gphr-1
SP12 + Pmyo2::gphr-1


I
Terminal bulb
Corpus
Head
J
Terminal bulb
Corpus
Head
SP12
Pmyo2::gphr-1
K
Head
L
Head
SP12 + Pmyo2::gphr-1
DIC

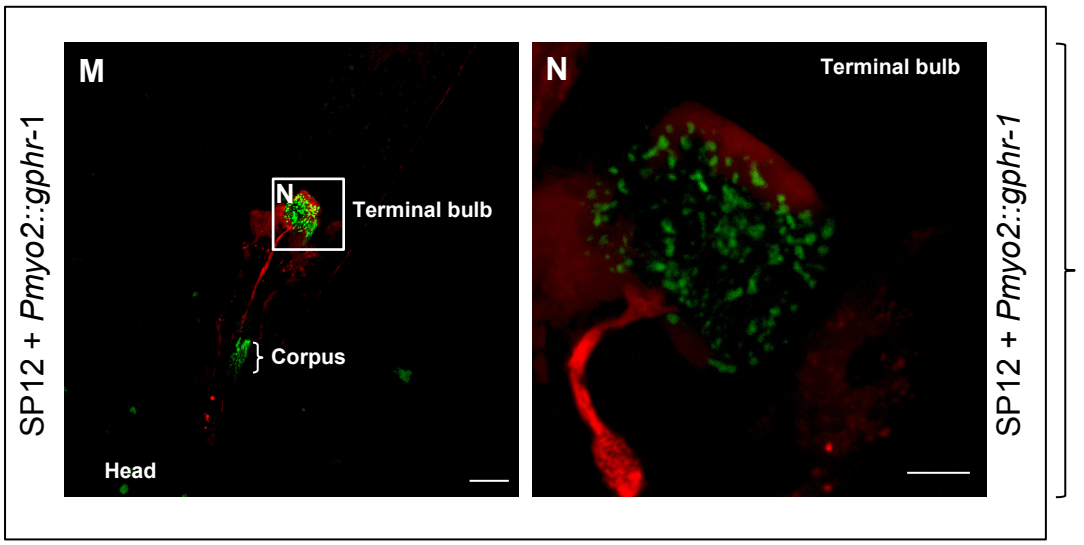

M
N
Terminal bulb
Corpus
Head
N
Terminal bulb
SP12 + Pmyo2::gphr-1
SP12 + Pmyo2::gphr-1

Figure 6.25 The subcellular localisation of *Pmyo-2::gphr-1::egfp* in *C. elegans*

Transgenic N2 adult worms displaying extrachromosomal arrays of *Pmyo-2::gphr-1::egfp* with ER marker *Peft-4::mCherry::sp12*. All images are maximum projections. Merged fluorescent images display *mCherry::sp12* (red) and *gphr-1::egfp* (green) expression. Yellow arrows = green autofluorescence from worm. Images from the same worm are grouped by a bracket on the right of the panel. **A**. Adult worm showing *gphr-1::egfp* expression in the corpus and terminal bulb regions of the pharynx, with constitutive *mCherry::sp12* expression. Higher magnification of pharynx area indicated by white inset and shown in **B**. Scale bar: 100 μm. **B**. Higher magnified pharynx region indicated in **A** showing *mCherry::sp12* and *gphr-1::egfp* expression. Scale bar: 20 μm. Insets indicated by **C** and **D** show further magnifications of terminal bulb (**C**) and corpus **(D)**, respectively. Scale bars: 10 μm. **E**. Another adult worm showing *gphr-1::egfp* expression in the corpus and terminal bulb regions of pharynx. Inset indicated by **F** shows a higher magnification of terminal bulb. Scale bar: 100 μm. **F**. Higher magnified terminal bulb indicated in **E** displaying *gphr-1::egfp* and *mCherry::sp12* expression do not colocalise. Scale bar 10 μm. **G**. Another adult worm showing *gphr-1::egfp* and *mCherry::sp12* expression in the pharynx. Scale bar: 20 μm. Inset indicated by **H** shows a higher magnification of terminal bulb. Scale bar: 5 μm. **H**. Higher magnified terminal bulb indicated in **G** displaying *gphr-1::egfp* and *mCherry::sp12* expression do not colocalise. **I**. Another worm pharynx showing ubiquitous *mCherry::sp12* in the head region. **J**. *gphr-1::egfp* expression in the pharynx corpus and terminal bulb regions of worm **I**. **K**. Merged fluorescent images (**I** and **J**) displaying no colocalisation of *mCherry::sp12* and *gphr-1::egfp*. **L**. DIC of worm head shown in **I**, **J**, **K**.; scale bars: 50 μm. **M**. Another worm showing *gphr-1::egfp* in the corpus and terminal bulb and *mCherry::sp12* expression. Inset indicated by **N**, shows higher magnification of terminal bulb. Scale bar: 50 μm. **N**. Higher magnified terminal bulb indicated in **M** showing *mCherry::sp12* and *gphr-1::egfp* expression do not colocalise. Scale bar: 5 μm.

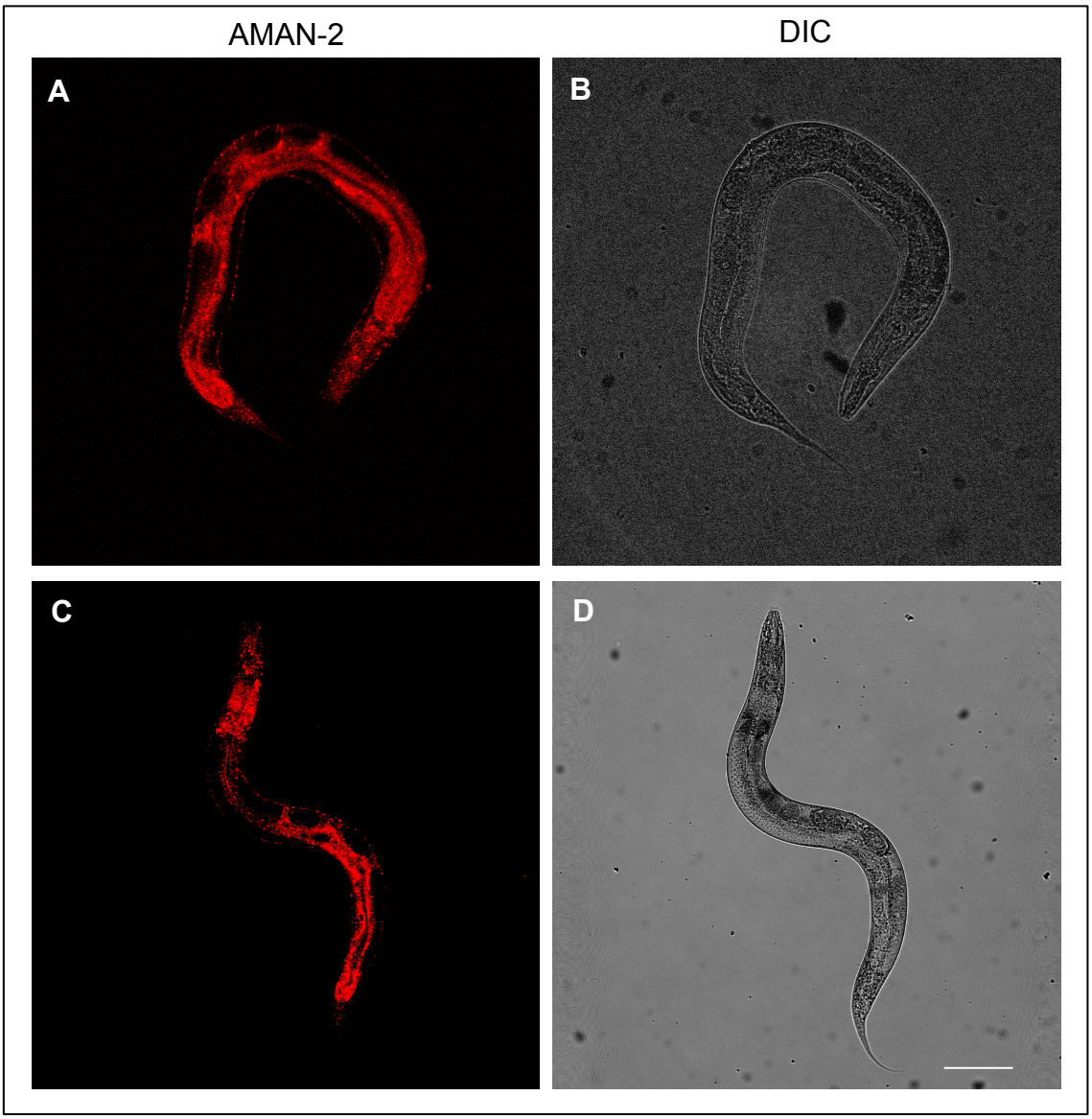


Figure 6.26 AMAN-2 is widely expressed in *C. elegans* adults

Transgenic N2 worms expressing *Paman-2::aman-2::mCherry*. AMAN-2 Golgi protein is relatively ubiquitous with strong expression in the nervous and digestive system, pharynx and hypodermal regions of the worm. All confocal fluorescent images are maximum projections. Scale bar: 100 μm.

Results so far show that GPHR-1 is localised intracellularly. *gphr-1::egfp* in *C. elegans* does not colocalise with the *mCherry::sp12* suggesting that it is not an ER membrane protein. However, we cannot confirm that it is Golgi localised without coexpression with a Golgi marker. Unfortunately no extrachromosomal lines containing both the *Pmyo-2::gphr::egfp* constructs coexpressing the Golgi marker *Paman-2::aman-2::mCherry* were obtained (Table 6.6). The *gphr-1::egfp* expression under the *myo-2* promoter forms punctate foci (Figure 6.25), which is characteristic of Golgi localisation (Figure 6.26 and Figure 6.27 A and B). Furthermore, the expression profiles for Golgi (Figure 6.27 A and B) and ER (Figure 6.27 C and D) membrane proteins are quite distinct from one another. Golgi localisation appears more punctate while ER may appear slightly reticular (Figure 6.27).

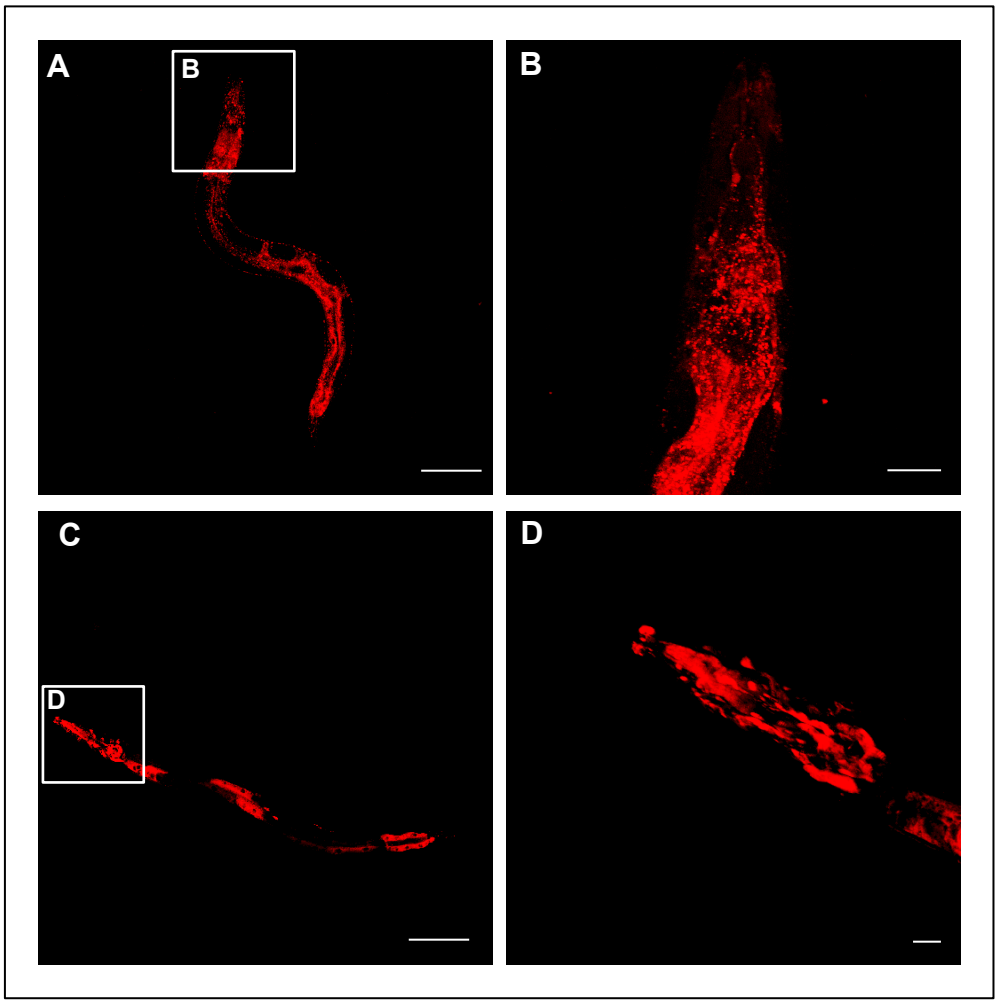


Figure 6.27 AMAN-2 and SP12 expression profiles are distinctive in the pharynx

Transgenic adult N2 worms expressing *Paman-2::aman-2::mCherry* (**A** and **B**) and *Peft-4::mCherry::sp12* (**C** and **D**). All fluorescent images are maximum projections of confocal Z-stacks. **A**. Whole worm fluorescent image showing AMAN-2 Golgi localised protein expression. White inset indicated by **B** shows a higher magnification of the head region. Scale bar: 100 µm. **B**. Higher magnification of head region showing pharynx of the region indicated in **A**. Scale bar: 20 µm. **C**. Whole worm fluorescent image showing SP12 ER localised protein expression. White inset indicated by **D** shows a higher magnification of the head region. Scale bar: 100 µm. **D**. Higher magnification of head region showing pharynx of the region indicated in **C**. Scale bar: 10 µm.

6.3 Discussion

6.3.1 Generating double *gphr* mutants reveal characteristic phenotypes

C. elegans deletion mutants are valuable tools in reverse genetics facilitating insights into gene function (Ahringer, 2006). The RB1390 and FXO4228 strains were obtained from CGC and NBP, and genotyped to confirm the *gphr-1(ok1579)* and *gphr-2(tm4228)* mutant alleles via genomic DNA extraction and PCR (Figure 6.5). Sequencing data confirmed not only the deletion sizes (1223 bp and 413 bp + 5 bp insertion for *gphr-1(ok1579)* and *gphr-2(tm4228)* respectively), but also the flanking sequences of the deletion and insertion sites, as illustrated by Figure 6.4. This verified both the RB1390 mutant allele of *gphr-1* and FXO4228 mutant allele of *gphr-2*.

Deletion mutants generated using this method may contain other background mutations during the mutagenesis process (Piette et al., 1985, Sladek et al., 1989, Yandell et al., 1994). For this reason, both RB1390 and FXO4228 were backcrossed six times with WT *C. elegans* strain, N2. Six cycles of backcross with N2 is sufficient to recover 98.4375 % of WT genes from the recurrent parent (N2) since each backcross will increase the percentage of the genetic material derived from the recurrent parent, N2 $\left(\frac{2^n-1}{2^n}\right)$; n = number of backcross generations (Plant & Soil Sciences eLibrary, 2014). The FXO4228 strain contains the mutant allele *gphr-2(tm4228)*, which is an X-linked mutation. A backcross method was developed in order to ensure the X chromosome was also backcrossed six times (Figure 6.6).

Both single mutant strains were backcrossed six times. Backcrosses were confirmed by single worm lysis followed by PCR and gel electrophoresis to eventually identify single *gphr-1(-/-)* and *gphr-2(-/-)* mutants (Figure 6.7). This method of identifying mutant alleles was necessary since there was not an obvious phenotype to distinguish WT and single mutants. Candidate single mutants were propagated into whole nematode populations to maintain the backcrossed strains. Genomic DNA was extracted from these single mutants to confirm whole population genotypes. Figure 6.7 confirms that both of these strains were successfully backcrossed six times and denoted as RB1390x6 and FXO4228x6 to indicate the number of backcrosses. A male RB1390x6 strain was generated and genotyped to confirm homozygosity (*gphr-1(-/-)*) of the population (Figure 6.7 A). To generate a *gphr-1(ok1579) gphr-2(tm4228)* double

mutant, the male *gphr-1(ok1579)* RB1390x6 strain was crossed with the hermaphrodite *gphr-2(tm4228)* FXO4228x6 strain (Figure 6.8). A *gphr-1(ok1579) gphr-2(tm4228)* double mutant homozygous for both mutations was successfully generated but failed in propagating a line. While knocking out both *gphr* genes was not lethal (F_0), the progeny (F_1) were not able to progress further than the L1 larval developmental stage (Figure 1.11). The results observed indicate that the *gphr-1(ok1579) gphr-2(tm4228)* mutant exhibited larval arrest. This is obviously problematic in maintenance of a *gphr-1(ok1579) gphr-2(tm4228)* double mutant line to study. For this reason, a method (Figure 6.13) had to be developed for quantitative phenotyping during the generation of a *gphr-1(ok1579) gphr-2(tm4228)* mutant and thus limited the phenotypic analyses that could be achieved. However, several clear phenotypes were observed in the double mutant (see below).

In future studies, genetic balances could be used to maintain this lethal phenotype of the *gphr-1(ok1579) gphr-2(tm4228)* mutant. Genetic balancers are a method to maintain lethal phenotypes in *C. elegans*. These are genetic constructs or chromosomal rearrangements allowing lethal or sterile homozygous mutations to be stably maintained in heterozygotes (Edgley et al., 2006). Balancer chromosomes have three key attributes: they suppress recombination with their homologues, carry dominant markers, and they negatively affect reproductive fitness when carried homozygously. In a balanced heterozygote, the WT allele for the gene of interest is carried by a balancer chromosome. For example, GFP balancers have been created such that the GFP is carried on the WT allele, allowing for negative selection i.e. lack of GFP of homozygous animals carrying mutation of interest. For example, *mel-28* encodes a maternal-effect lethal gene (Fernandez and Piano, 2006, Galy et al., 2006), like the *gphr* genes (discussed in section 6.3.4). A GFP-balanced *mel-28* strain was generated whereby the *mel-28* allele and an *unc-22* allele were carried on a WT chromosome, which were balanced by the *qC1* balancer bearing the *qIs26* insertion. In this way a population of *mel-28* homozygotes was generated, identified by the uncoordinated phenotype (Fernandez et al., 2012). Since phenotypic characterisation was limited due to the larval arrest phenotype of the double *gphr* mutants, generating a balanced *gphr* strain could be used for advancing chacterisation.

6.3.2 Double *gphr* mutants are defective in egg-laying, hatching, embryogenesis, pharyngeal pumping and exhibit L1 diapause

Egg-laying behaviour in *C. elegans* is a well-characterised process (White et al., 1986) and used extensively in phenotypic analyses, which has led to the discovery of various fundamental processes in neuronal cell biology and signal transduction pathways (Segalat et al., 1995, Brundage et al., 1996). An egg-laying event is defined as the opening of the vulva to release one or two eggs. These events occur in a specific temporal pattern of short bursts every 20 min but only lasting a few minutes (Waggoner et al., 1998). Results indicate that the number of eggs laid by the *gphr-1(ok1579) gphr-2(tm4228)* mutant is significantly less than both the *gphr-1(-/-) gphr-2(+)* and *gphr-1(-/-) gphr-2(+/-)* (Figure 6.14 A). Since egg-laying episodes occur relatively frequently, there had been sufficient time elapsed (4-d period) to warrant the occurrence of egg-laying events and therefore to conclude that the double mutants essentially lay fewer eggs than WT. Furthermore, WT animals typically lay up to 300 eggs in their lifetime (Schafer, 2005). However, the double *gphr* mutants laid at most 23 eggs in their life time and on most occasions, none.

Defective egg-laying could be due to a number of reasons. The egg-laying neuronal mechanism is well-characterised and it could be that the phenotype observed in the *gphr* double mutant is due to a breakdown in these principal molecular mechanisms. Neuropeptides encoded by the *flp-1* gene play an important role in egg-laying (Waggoner et al., 1998). Proprotein convertases are important for the activation of neuropeptides as they process neuropeptide precursor molecules; PCR/EGL-3 has a role in processing FMRF amide-like peptide (FLP) precursors and neuropeptide-like protein (NLP) precursors in *C. elegans* (Husson et al., 2006). *egl-3/kpc-2* encodes one of four proprotein convertases in *C. elegans* (Thacker and Rose, 2000). Loss of *egl-3/kpc-2* causes defective egg-laying as well as defective pharyngeal pumping, mechanosensation and locomotion (Kass et al., 2001, Jacob and Kaplan, 2003). Neuropeptide activation begins in the ER with the cleavage of the signal peptide and continues in the Golgi (Strand, 1999). Thus without normal function of the ER/Golgi network, neuropeptides cannot be made and can therefore lead to egg-laying defects (Kass et al., 2001, Jacob and Kaplan, 2003). Therefore, the egg-laying defective phenotype seen in the double *gphr* mutants could be explained by a breakdown in the

fundamental elements required for functional egg-laying machinery, such as dysfunctional neuropeptides (Figure 6.28).

Only a small proportion of eggs that were laid by the *gphr-1(ok1579) gphr-2(tm4228)* mutants hatched (Figure 6.15). This indicates that there could be an egg developmental defect, or a defect in eclosion; this is the process of L1 larva exciting the egg, the last step of hatching (Herndon, 2009). Figure 1.10 shows the progressive stages from fertilisation to larva hatching and is introduced in section 1.5.4.1. Failure to undergo these stages at any point will prevent larval development, and in turn hatching. For example, the *rhr-1* gene is essential for normal embryonic development as shown by embryonic arrest via RNAi *rhr-1* knockdown (Ji et al., 2006). Very few progeny managed to escape arrest at the comma stage of embryo development (Figure 1.10) but ultimately any that did escape were arrested at the pretzel stage (Figure 1.10) therefore *rhr-1* is important for embryonic development (Ji et al., 2006).

To explore the possibility of a defect in embryogenesis as a reason for decreased egg hatching, egg development was followed. In most cases, many of the eggs remained at the 30-cell stage (Figure 6.16); wherein successive stages were not attained, as in Figure 1.5. Multinucleated embryos may arise from a cytokinesis defect (Wang et al., 2005), but this cannot be determined from the images taken (Figure 6.16). Nevertheless this certainly would be an interesting concept since the terminal phase of cytokinesis of early *C. elegans* embryos requires protein glycosylation, which occurs in the Golgi. Therefore, further study into embryogenesis of the double *gphr* mutants would be interesting to ascertain whether these are defective in cytokinesis; and thus exhibit impaired protein glycosylation, resulting in the egg development defective phentoypes. This would coincide with the theme that these GPHRs are Golgi membrane proteins that act as anion channels to regulate pH (Maeda et al., 2008). There was a low percentage (up to approximately 20 %) that escaped this embryonic arrest, surpassed gastrulation and hatched (Figure 6.15). There were a few observed to reach the three-fold stage but were unable to hatch. Therefore, there must be an essential process in egg development that is affected by the loss of both *gphr* genes causing defective egg development; perhaps specifically at the 30-cell and three-fold stage. However, at this stage there can be no solid conclusion that the *gphr* genes are either directly associated with embryonic development or fertility.

A small proportion of eggs laid by the double mutant hatched but none of the resulting larvae reached adulthood and instead arrested at the L1 stage and eventually died (Figure 6.17). This finding coincides with the larval arrest phenotype observed in the two *Drosophila GPHR* KO mutants (Charroux and Royet, 2014). Post-embryonic development initiates when the larva feeds after hatching. In the presence of food, cell division resumes allowing the post-embryonic development to be activated 3 h after hatching (Ambros, 2000). Post-embryonic development involves the L1 larva progressing through four larval stages (L1 – L4) before reaching adulthood (Figure 1.11). However, if the embryos hatch in the absence of food the development into adulthood is arrested until food becomes available; L1 larva can survive up to two weeks without feeding (Johnson et al., 1984). Normal molting and development resumes in arrested L1 larva when food becomes available (Slack and Ruvkun, 1997). However, these double mutant larvae were hatched in the presence of food. This prompted testing whether the double mutants were feeding.

C. elegans feeding involves taking up bacteria from the surroundings and transporting it to the intestine and this is achieved through the action of the pharynx. The pharynx is a pump and feeding behaviour that can be measured by counting the number of forward-backward movements of the grinder corresponding to the contraction-relaxation cycle of the pumping motion of the pharynx (Raizen et al., 2012). Indeed, it was found that *gphr-1(ok1579) gphr-2(tm4228)* double mutants had a defect in pharyngeal pumping (Figure 6.18 A); this could be a contributing factor for larval arrest of the double mutant. Typically, WT adult *C. elegans* show a pumping rate of around 250 pumps.min^{-1} (Walker et al., 2002) (Figure 6.18), and the number is reduced in the absence of food. Interestingly, a lower rate of pharyngeal pumping was observed for the N2 L1 larva (~170 pumps.min^{-1}); however, the *gphr-1(ok1579) gphr-2(tm4228)* double mutant pumping rate was less than 10 pumps.min^{-1} on food which is abnormal for typical *C. elegans* feeding behaviour. The motor neuron, MC, initiates pharyngeal muscle contraction and its regulation is dependent on food (Raizen et al., 1995) and serotonin (Riddle et al., 1997, Hobson et al., 2006).

However, the double mutant embryos were hatched in the presence of food and therefore there could be a problem with the mechanosensation, food sensing, and/or endogenous serotonin sensing and/or signalling that may be linked to L1 arrest. Having defects in both egg-laying and pharyngeal pumping simultaneously is not uncommon;

for example it is seen in mutants with reduced serotonin and dopamine levels (Sulston et al., 1975, Avery and Horvitz, 1990). Serotonin and dopamine biosynthesis pathways occur in the cytoplasm (Golan et al., 2011). If these are Golgi-localised proteins it is unlikely that a defect in the synthesis of these neurotransmitters is the cause for decreased egg-laying and pharyngeal pumping. However, there are some mutants that are defective in pharyngeal pumping but do not arrest at the L1 stage. For example, *tph-1(mg280)* mutants are defective in serotonin synthesis resulting in decreased pharyngeal pumping and egg-laying phenotypes but do not exhibit L1 arrest (Sze et al., 2000).

The lack of pharyngeal pumping may be a contributing factor along with other causes for larval arrest. Further supporting this hypothesis, the *Drosophila GPHR* mutant phenotypes were not related to impaired feeding (Charroux and Royet, 2014). Since the *gphr* double mutants exhibit larval arrest in addition to defects in egg-laying and pharyngeal pumping, it is likely that there is a problem with mechanosensation and locomotion since it was observed that the F_0 *gphr* double mutant and F_1 L1 arrested larvae showed reduced movement. This could be analysed by video tracking methods. For example, automatic tracking was used to determine movement features and phenotypes for several mutants such as *eat-4* (Geng et al., 2004). Multiple worms as well as *C. elegans* of different sizes and conformations, such as coiled worms can also be tracked (Roussel et al., 2007). Therefore, future study could explore the locomotion behaviour of these double *gphr* mutants as a possible defect as a contributing factor for L1 arrest. Alternatively, the reduced locomotion could be because the worms are not feeding and lack the energy to do so.

L1 arrest, also known as L1 diapause, is characterised by a failure to initiate post-embryonic development in the absence of food (Johnson et al., 1984), or at a high temperature (Morley and Morimoto, 2004). Since the *gphr* double mutants were hatched in the presence of food and incubated at normal temperature, there could have been a problem with their food and/or temperature sensing/signalling machinery. The insulin/IGF (insulin-like growth factor) signalling pathway involving the DAF-16/FOXO and DAF-2/IGF receptor, is not only central to *C. elegans* dauer formation (Figure 1.10) (Kimura et al., 1997), but also to L1 arrest (Baugh and Sternberg, 2006). For example, the *daf-2(e979)* mutant allele results in embryonic L1 arrest (Vowels and Thomas, 1992, Gems et al., 1998) through impaired insulin/IGF signalling rather than an inability to feed (Baugh and Sternberg, 2006). Furthermore, *asna-1* encodes a

conserved ATPase required for insulin-like peptide secretion and larval development. A mutation in this gene leads to constitutive L1 arrest phenotype (Kao et al., 2007), as seen in the *daf-2* mutants (Baugh and Sternberg, 2006).

Golgi dysfunction has been reported to be linked to larval development by influencing insulin-like peptide secretion, as seen in an RNAi screen predicting genes that interact with ASNA-1 (Billing et al. 2012). For example, the knockdown of Golgi-localised v-SNARE, *ykt-6*, led to Golgi dysfunction that caused defective insulin secretion and ultimately resulted in a larval arrest phenotype (Maekawa et al., 2009, Billing et al., 2012). A defect in egg-laying in the first affected (F_0) generation, followed by L1 arrest of its progeny (F_1), is typical. For example, when endogenous monomethyl branched-chain fatty acids (mmBCFAs) are not produced through *elo-5(RNAi)* inhibition, it causes developmental defects in the intestine and gonads as well as defects in egg-laying in the F_0 and L1 arrest in the F_1 (Kniazeva et al., 2004). The pharyngeal pumping was examined for the F_1 progeny that arrested at L1 stage, but it would be interesting to test the pharyngeal pumping of the F_0 *gphr-1(-/-) gphr-2(-/-)* to see whether the action of the pharynx is affected in this generation. If pharyngeal pumping is limited in the F_0 *gphr-1(-/-) gphr-2(-/-)* mutants it would reveal that the reduced pumping rate seen in the L1s does not contribute to the arrest and could result from the breakdown of, for example, the insulin/IGF signalling pathways (Figure 6.28). However, in the *Drosophila GPHR* mutants, the growth and development defects were shown to be independent of the insulin signalling pathway (Charroux and Royet, 2014). A summary of all the defects observed in the *gphr-1(ok1579) gphr-2(tm4228)* mutant is depicted in Figure 6.28, with all the associated causal defects addressed in this section with many linked to the secretory pathway.

6.3.3 RNAi knockdown does not phenocopy *gphr-1(ok1579) gphr-2(tm4228)*

Without the ability to maintain a *gphr-1(ok1579) gphr-2(tm4228)* mutant line, the phenotyping process is limited. RNAi was therefore employed to further characterise these *gphrs* and/or to phenocopy the double *gphr-1(ok1579) gphr-2(tm4228)* mutants. The results indicate that feeding the *gphr-1(ok1579)* deletion mutant with an RNAi construct designed to knockdown *gphr-2*, did not have an effect on hatching or cause any developmental arrest in embryogenesis or larvae as seen in the double *gphr-1(ok1579) gphr-2(tm4228)* deletion mutants (Figure 6.19). However, there was a

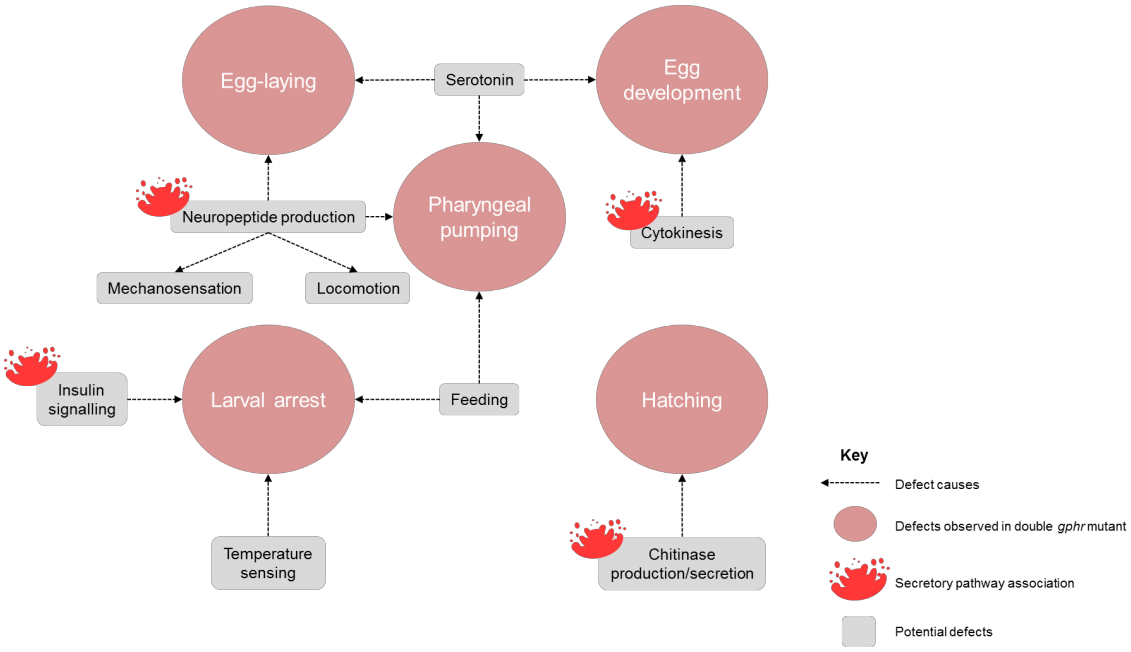


Figure 6.28 Double *gphr* mutant are defective in growth, development and fertility

Schematic showing the defects (pink) observed by the double *gphr-1(ok1579) gphr-2(tm4228)* mutant, with the potential defects (grey) that could cause these phenotypes. For example, a defect in neuropeptide production, which occurs in the secretory pathway causes egg-laying and pharyngeal pumping defects, as well as impaired mechanosensation and locomotion.

significant decrease in eggs laid by the *gphr-1* knockdown *gphr-2(tm4228)* mutants (Figure 6.19). *gphr-1(+/-) gphr-2(-/-)* do not exhibit hatching or developmental arrest defects, but it was observed that there was a slight decrease in eggs laid but it was not found to be significantly different (Figure 6.14). However, this would need to be repeated as these results were only from one experiment. In addition, the feeding treatments were carried out alongside a known working feeding vector to demonstrate that the RNAi method was working under our conditions. pLT61 was also fed to N2 worms to check RNAi efficiency (section 2.5.4). The results (section 2.5.4) indicated that twitching of N2 worms was observed when fed pLT61 as expected from previous studies (Fire et al., 1998, Timmons and Fire, 1998, Timmons et al., 2001).

If *gphr-2* was knocked down by the feeding vector in RB1390, these results suggest that *C. elegans* are able to function with only one GPHR or with low levels of expression of one GPHR. RNAi was not able to phenocopy the *gphr-1(ok1579) gphr-2(tm4228)* mutant line but this could be due to specificity of the RNAi feeding vector. The fragment size cloned into pL4440 was a 1887 bp, as determined from sequencing results. However, it is more common to use a fragment of approximately 800 bp, as seen in RNAi experiments to induce *unc-22* silencing (Timmons et al., 2001). Further RNAi studies could be conducted by creating a new RNAi feeding construct in which a shorter, exon-rich fragment could be inserted into feeding vector (pL4440). Since a phenotype was not observed, it would be important to confirm whether there *gphr-2* knockdown by qPCR and/or Western Blotting before any firm conclusions can be made.

6.3.4 The *gphr* genes are essential and display gene functional redundancy

F_0 *gphr* double mutants are able to develop into adults (Figure 1.11) but F_1 progeny exhibit larval L1 arrest (Figure 6.17). This result indicates that *gphr* genes are essential in *C. elegans*. Essential genes are those that are necessary for growth to a fertile adult and these are defined by three types of mutations: zygotic lethal mutations (lethals), maternal-effect lethal mutations (maternal-effect lethals) and sterile mutations (steriles) (Kemphues, 2005). Lethal mutations of individuals homozygous for the mutation inhibit development into an adult. Maternal-effect lethals homozygous for the mutation stop the hermaphrodite's progeny developing into adults. Maternal-effect genes must be expressed in the mother in order for embryonic development. For example, *zyg-11* is a maternal-effect gene required for early embryogenesis (Carter et

al., 1990). Sterile mutations could arise from defects in germline development, somatic gonad development, oogenesis, spermatogenesis, ovulation or fertilisation (Kemphues, 2005).

By these classifications, the *gphr* genes are maternal-effect essential genes since the F_0 *gphr-1(-/-) gphr-2(-/-)* hermaphrodites that arose from either *gphr-1(+/-) gphr-2(-/-)* or *gphr-1(+/-) gphr-2(+/-)* mutants were able to develop into adults but the progeny (F_1) arrested at the L1 stage (Figure 6.17). As described above, the embryonic development of *gphr* double mutants arrested at two different stages (30-cell and three-fold stage). This is not uncommon for maternal-effect genes, for example *clk-2* embryos arrest at various stages in development (Bénard et al., 2001).

In contrast, the single *gphr-1(ok1579)* and *gphr-2(tm4228)* mutants do not exhibit the phenotypes of the double *gphr* mutants. There are no significant differences between the eggs laid (Figure 6.14 B) and pharyngeal pumping (Figure 6.18 B) of single mutants and WT. No observations were made that indicated that the single mutants had a defect in embryogenesis, hatching or development into adulthood. The absence of such phenotypes observed in the single mutants suggests these two genes are functionally redundant; that it requires knocking out both *gphr-1* and *gphr-2* genes to disclose a phenotype. Functional redundancy can result from gene duplication events in which one gene can carry out a function. For example it requires knocking out of both *end-1* and *end-3* to cause an arrested embryo phenotype (Zhu et al., 1997). Similarly, both Arabidopsis *GTGs* need to be knocked out to elicit marked phenotypes (Pandey et al., 2009, Jaffé et al., 2012). The deletion mutant results suggest that a complete knockout of both *gphr* genes in *C. elegans* is required to elicit severe defects in growth, development and fertility. The *gphr* alleles are recessive since the heterozygosity in the double *gphr* mutants were indistinguishable from WT. In diploid organisms, like *C. elegans*, a "dominant" allele of a gene is expressed over the "recessive" allele in the phenotype. In this case, the interaction between the two *gphr* genes is considered to be recessive as well as functionally redundant; both mutant alleles of *gphrs* need to be present, *ok1579* and *tm4228* to observe the phenotypes.

6.3.5 The localisation of the GPHRs in *C. elegans*

The localisation of proteins is inherently linked to their function, therefore determining *Cel*-GPHR location was a key aim in this project. A variety of fluorescent

images are shown (section 6.2.4) because these originate from extrachromosmal arrays, and are therefore subjective to mosaic expression. To grasp an extensive view on the localisation of these proteins, it was therefore necessary to show a number of images. Those shown are consistent with two independent lines and many worms. To explore the localisation of the *C. elegans* GPHRs, ER and Golgi membrane markers were used, since there are several reports supporting that the GTG/GPHRs are Golgi- and ER-localised in both animals and plants (Maeda et al., 2008, Jaffé et al., 2012, Charroux and Royet, 2014). The ER marker is under a constitutive promoter, *Peft-4* (McKay et al., 2003, Murata et al., 2012) and results displayed expected expression profiles (Figure 6.24). The Golgi marker is broadly expressed under its native promoter (Figure 6.26), which again was as expected (McKay et al., 2003, Paschinger et al., 2006, Dupuy et al., 2007).

Green autofluorescence was observed in some worms (Figure 6.23 – Figure 6.25), and hence it was important to distinguish this from EGFP through lambda scans of the emission spectra (Figure 6.23). Autofluorescence is common in adult worms due to the autofluorescent lipofuscin granules in the intestines of *C. elegans* (Clokey and Jacobson, 1986). If the transformations using the *gphr::egfp* constructs under *Peft-4* had worked, they could have proved quite difficult to analyse if the GPHRs were expressed in the intestines, as the autofluorescence would distort localisation analysis. However, it was important to generate the ubiquitous constructs as these may be used in the future for rescue and localisation experiments to show the functional, subcellular localisation of the proteins, as in the elucidation of the function and localisation of PIGA-1 (Murata et al., 2012).

Pmyo-2 drives gene expression in pharyngeal muscle cells (Dibb et al., 1989). The fluorescent images indicate that *gphr-1::egfp* under *Pmyo-2* was expressed in the pharynx corpus, terminal bulb (Figure 6.24 and Figure 6.25), and although weakly, also in the isthmus (Figure 6.24 I and J). It is quite difficult to decipher whether *gphr-1::egfp* is expressed in all 20 pharyngeal muscle cells. But if not, it could be as a result of mosaic expression due to extrachromosomal arrays, and could explain the reason for the weak expression in the isthmus. In line with Maeda et al. (2008), Jaffé et al. (2012) and Charroux and Royet (2014) *Cel*-GPHR-1 appears to be subcellularly localised. This is in contrast to the plasma-membrane localisation reported by Pandey et al. (2009). There are only 20 muscle cells within the pharynx constituting eight pharyngeal muscle (pm1

– 8) sections (Figure 1.12). The terminal bulb contains pm6, pm7 and pm8, comprising seven cells altogether (Figure 1.12). Figure 6.25 indicates that there is strong expression in this terminal bulb region and at higher magnification it can be seen that *gphr-1::egfp* expression is noticeably intracellular and not at the plasma membrane. It is difficult to distinguish between distinct cells/pm sections due to tri-radiate symmetry (Figure 1.12) without membrane staining, but since there are only a few cells that comprise the terminal bulb, these maximum projections of the Z-stacks would have scanned through whole cells.

Colocalisation described both co-occurrence and correlation (section 6.1.4). The results show that the two fluorophores mCherry (reporter for the ER) and EGFP (reporter for *myo-2::gphr-1*) show no co-occurrence; i.e. do not occur in the same pixel. For this reason, there was no requirement to analyse the correlation (Adler et al., 2008) so statistical analysis of colocalisation was not conducted. GPHR-1 appears to be intracellularly expressed, the results indicate that the GPHR-1 does not colocalise with SP12, an ER membrane protein, and therefore suggesting that GPHR-1 does not reside in the ER membrane (Figure 6.25). However we were not able to conduct a colocalisation experiment using the Golgi marker, and it remains unknown whether GPHRs are localised in the Golgi membrane or not. Looking at the Golgi and ER marker protein localisation profiles it is obvious they are quite distinct from one another (Figure 6.27). The GPHR-1 localisation results are comparable to that of the Golgi membrane protein AMAN-2 (Figure 6.25 and Figure 6.27). Again, although it does not prove they are Golgi membrane-localised the pattern is more consistent with this.

The tissue localisation remains inconclusive since the promoter-only reporters gave no detectable green fluorescence. This could be as a result of incorrect promoter sequence cloned in since although these were checked using restriction enzyme analysis, they were not sequenced. However this is unlikely since subsequent sequencing of *Pgphr-2::gphr-2::egfp* (from two repeats), made from the same genomic template, was found to be correct according to the genomic sequence provided by WormBase. Or it could be as a result of a weak promoter, which could explain why the *Pgphr-2::gphr-2::egfp* construct also produced no detectable EGFP fluorescence.

The *gphr-1(ok1579) gphr-2(tm4228)* mutants exhibit phenotypes similar to those observed in the Arabidopsis (Jaffé et al., 2012) and *Drosophila* (Charroux and Royet, 2014) *GTG/GPHR* KO models. The localisation data indicates that like

mammalian GPHRs, the *Cel*-GPHR-1 is localised intracellularly and this could be at the Golgi. Therefore, the GTG/GPHRs may have a conserved role in eukaryotic cells amd are involved in key biological processes required for growth, development and fertility.

CHAPTER 7

Heterologous expression of *GTG/GPHR* genes in *S. cerevisiae*

7.1 Introduction

7.1.1 Use of *S. cerevisiae* deletion mutants and heterologous expression of the GTG/GPHRs

S. cerevisiae is a commonly used model organism (introduced in section 1.5.5) for studying molecular and cellular biology including genetic functional analysis (Qi et al., 2013) or protein-protein interactions (Botstein et al., 1997, Kharenko et al., 2013). Mutants in other model organisms have also been discussed previously for Arabidopsis (section 4.1.1) and *C. elegans* (6.1.3.1). Likewise, *S. cerevisiae* mutants are extensively used in genetic, biochemical, complementation and functional assays in order to provide insight into biological processes. Characterising phenotypes of deletion mutants can be used to understand how genes work in *S. cerevisiae*. The database EUROpean *Saccharomyces Cerevisiae* ARchive for Functional Analysis (EUROSCARF) is a resource for *S. cerevisiae* WT and deletion strains (http://web.uni-frankfurt.de/fb15/mikro/euroscarf/) (Baker Brachmann et al., 1998). Typically, strains may have deletions in particular enzymes to serve as selection markers. *URA3* encodes for orotidine 5′-phosphate decarboxylase (ODCase) required in pyrimidine ribonucleotide biosynthesis in *S. cerevisiae* (Lacroute, 1968, Umezu et al., 1971). The deletion of *URA3* (*ura3*-Δ) for instance, is useful for positive selection of plasmid transformations since they cannot grow without *URA3* expression or uracil supplementation (Jones, 1991). ODCase converts 5′-fluoroorotic acid (5-FOA) into the toxic compound 5-fluorouracil, therefore 5-FOA can also be used for negative selection (Boeke et al., 1984).

Drop tests are typically used in phenotypic assays to show growth differences; the strains are spotted onto solid media in a dilution series and images are taken to monitor growth over time, as shown in Qin et al. (2007) and Chui et al. (2012). Synthetic complete (SC) medium contains all the necessary nutrients required for growth (Bergman, 2001, 2006). *S. cerevisiae* can grow on a number of different carbon sources

including glucose and galactose (Granot and Snyder, 1993). Different media can be used to reveal growth phenotypes; for example, altering the carbon or nitrogen source, limiting specific nutrients such as iron (Hechenberger et al., 1996), zinc (Gitan et al., 1998, North et al., 2012) or potassium (Sentenac et al., 1992). SC medium without uracil (denoted SC-ura from herein) can be used to select for plasmid transformation if the plasmid contains a *URA3* marker.

There are databases of large-scale screens which identify phenotypes associated with deletion mutants. The information is collected and available on the online resource *Saccharomyces* Genome Database (SGD: yeastgenome.org) (Cherry et al., 2012). These screens are useful starting points for discovering phenotypes. Deletion mutants can also be used in functional complementation assays to reveal molecular mechanisms underlying biological processes of other organisms. *S. cerevisiae* mutant phenotypes can be restored to WT through either the supplement of an ingredient or expression of a gene. Heterologous expression in *S. cerevisiae* is the expression of genes from different species in *S. cerevisiae*, and is often used in complementation assays to address gene functionality. This can be particularly useful when trying to test the functionality of a protein in for example, Arabidopsis. As the Arabidopsis life cycle is much longer than that of *S. cerevisiae* and requires more time to generate Arabidopsis transgenic lines. Today *S. cerevisiae* is a valuable functional expression system for bacterial, animal and plant membrane proteins (Frommer and Ninnemann, 1995).

S. cerevisiae has been exploited for studying membrane transporters including ion channels. For instance, the CLC family of chloride channels exists in plants, animals and fungi. Deleting the *CLC* gene in *S. cerevisiae* causes a defect when grown on iron-deficient media. The Golgi-localised Sc CLC counteracts vesicular acidification essential in iron-deficient conditions for copper cofactor acquisition of a multicopper oxidase, Fet3. Fet3 is required for cation uptake and glutathione homeostasis under such stress conditions (Braun et al., 2010). However, the expression of At *CLC-d* in the Sc *clc*-Δ mutant was able to restore normal growth in iron-deficient conditions. This functional complementation implied that both Sc *CLC* and At *CLC-d* play a similar intracellular role and that At *CLC-d* may act as a chloride channel since it was able to functionally substitute for the loss of *CLC* in *S. cerevisiae* (Hechenberger et al., 1996). There are many examples of functional complementation and conserved biological processes; however function is not always conserved. For example, Tlg2/syntaxin16 is a

TGN protein in *S. cerevisiae* and mammalian cells, and its orthologue in Arabidopsis, SYP43, was used to show that the dynamic aspects of the TGN in plant cells are different to those of animal and *S. cerevisiae* cells (Uemura et al., 2014).

The GTG/GPHRs have been studied in both plant (Pandey et al., 2009, Jaffé et al., 2012, Alvarez et al., 2013) and animal (Maeda et al., 2008, Tarutani et al., 2012, Charroux and Royet, 2014) kingdoms. Whole organism KO mutants of the *GTG/GPHRs* result in several defects in Arabidopsis (Pandey et al., 2009, Jaffé et al., 2012), *C. elegans* (CHAPTER 6) and *Drosophila* (Charroux and Royet, 2014). In addition, a number of phenotypes were also observed in a skin-specific *GPHR* KO model in mice (Tarutani et al., 2012) and a CHO cell line (Maeda et al., 2008). The interaction of recombinant At GTG1 with ABA in *S. cerevisiae* has been investigated (Kharenko et al., 2013), but the endogenous protein, Sc Gtgp (see section 2.1 on nomenclature) is yet to be characterised. The phenotypes observed in whole organism or tissue-specific KO models (Arabidopsis, *C. elegans*, *Mus musculus*, and *Drosophila*) were associated with defects in growth, development and fertility (CHAPTER 4 and CHAPTER 1CHAPTER 6) (Pandey et al., 2009, Jaffé et al., 2012, Tarutani et al., 2012, Alvarez et al., 2013, Charroux and Royet, 2014). Studying GTG/GPHRs from the fungal kingdom is therefore important for determining the function of this family of proteins in eukaryotes.

Characterising a *GTG* KO model in *S. cerevisiae* would be useful for comparing these proteins across three different kingdoms. For example, plant and animal *GTG/GPHR* KO models are defective in growth, so would a slow reproduction rate be a trait for the loss of Gtgp in *S. cerevisiae*? At GTG1, At GTG2 (Jaffé et al., 2012) and *Cel*-GPHR-1 can rescue root and hypocotyl defects of Arabidopsis *gtg1 gtg2* double mutants (CHAPTER 5). While the Arabidopsis and *C. elegans* GTG/GPHRs share ~39 % identity to one another, *S. cerevisiae* Gtgp is only 13 % identical to the orthologues found in Arabidopsis and *C. elegans*. If there are mutant phenotypes in the *S. cerevisiae* would these Arabidopsis and *C. elegans* GTG/GPHRs be able to complement them? The heterologous expression of the *GTG/GPHR* genes in *S. cerevisiae* complementation assays would therefore confirm that there is a conserved function for the GTG/GPHR proteins in eukaryotes shown by organisms from three kingdoms. GTG/GPHRs from different organisms are discussed throughout this chapter, refer to section 2.1 for the species-specific GTG/GPHR symbols including species prefixes.

7.1.2 Using *S. cerevisiae* for protein purification

The structure of proteins can be determined using different techniques such as X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. Producing large quantities of protein and its subsequent purification are two major steps in structural biology. Eukaryotic membrane proteins can be difficult to purify due to challenges encountered with production, solubilisation and purification of appropriate amounts to form crystals diffracting at a high resolution (Clark et al., 2010). For this reason, the number of membrane protein structures characterised is relatively low, as evident from the few membrane protein structures available in the Membrane Proteins of Known 3D Structure database (http://blanco.biomol.uci.edu/mpstruc/listAll/list). The expression and subsequent purification of proteins can be accomplished in *S. cerevisiae* (Rigaut et al., 1999, Scharff-Poulsen and Pedersen, 2013) and bacteria such as *E. coli* (Guan and Dixon, 1991).

The PAP1500-GFP expression system in *S. cerevisiae* combines a high-copy expression system (Pedersen et al., 1996) with membrane protein-GFP fusion technologies (Drew et al., 2008) to produce high amounts of endogenous proteins of satisfactory fluorescence-detection size-exclusion chromatography (FSEC) profiles (Scharff-Poulsen and Pedersen, 2013). In the event that the Arabidopsis and *C. elegans* GTG/GPHRs need to be purified for structural analysis, the proteins can be expressed in *S. cerevisiae*. pYTV is a Gateway-compatible *S. cerevisiae* expression vector with C-terminal tags to facilitate purification of the fusion protein (Figure 2.3 B): 3XFLAG (a triple FLAG epitope sequence), 6XHis (polyhistidine-tag of 6 hisitine residues), a 3C cleavage (a cysteine cleavage site for tag removal e.g. 6XHis) and 2XIgG binding protein (for tandem affinity purification) (Gong et al., 2004). The Arabidopsis and *C. elegans* GTG/GPHRs can be inserted into pYTV for protein purification and possibly structural analysis to further characterise these proteins.

7.1.3 The localisation of the GTG/GPHRs

GFP (introduced in section 1.4.2) is a fundamental reporter used in both localisation and expression studies in *S. cerevisiae*. Fusion proteins expressing a GFP reporter can be used to follow subcellular localisation via live-cell fluorescence microscopy in *S. cerevisiae*. The GFP profiles are distinct for different *S. cerevisiae* organelles (Figure 7.1); this is useful in the absence of a colocalisation marker.

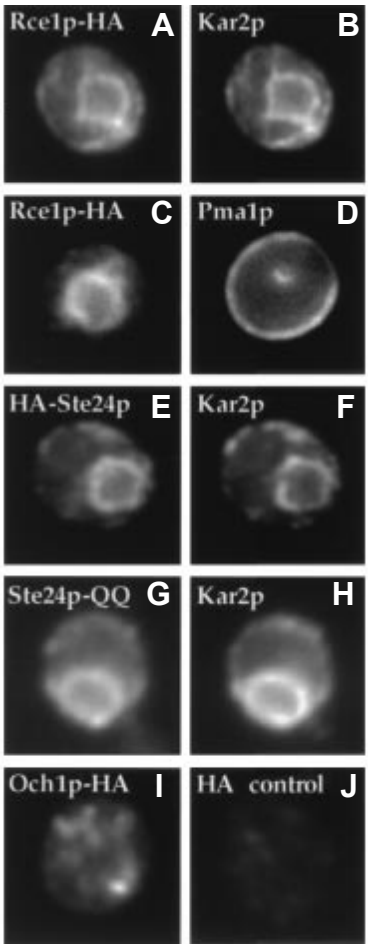


Figure 7.1 The GFP-fusion protein fluorescence profiles are distinct in different *S. cerevisiae* organelles

Immunofluorescence localisation of proteins from different organelles in *S. cerevisiae*. **A – C**. ER localisation. **D**. Plasma-membrane **E – H**. ER localisation and **I**. Golgi localisation. **J**. Control. Figure taken from Schmidt et al. (1998).

Heterologous expression of plant genes in *S. cerevisiae* has confirmed the localisation of a number of membrane proteins. For example, the At CLC-d protein mentioned above, showed a similar staining pattern to the Sc CLC-GFP fusion protein (Hechenberger et al., 1996). The Sc CLC protein has been shown to localise to late- or post-Golgi vesicles (Gaxiola et al., 1998). Together the localisation and functional complementation results helped identify the At CLC-d as an intracellular chloride channel (Hechenberger et al., 1996). Often, plant membrane proteins are expressed in their native species as well as in *S. cerevisiae* to confirm their localisation. *KEA3* encoding a K^+/H^+ antiporter in Arabidopsis, is Golgi-localised when expressed in either *S. cerevisiae* or Arabidopsis (Zheng et al., 2013). However, proteins do not always localise similarly in plants and *S. cerevisiae*. For example, functional heterologous expression of plant plasma membrane H^+-ATPase in *S. cerevisiae* exhibited ER localisation rather than plasma membrane (Villalba et al., 1992).

The Arabidopsis and *C. elegans* GTG/GPHRs can be expressed in *S. cerevisiae* to further study the localisation of these proteins. The Gateway-compatible *S. cerevisiae* destination vector, pAG426GAL-ccdB-EGFP can be used to study the localisation of the Arabidopsis and *C. elegans* GTG/GPHRs (Alberti et al., 2007). It contains a C-terminal tag EGFP under the control of the *GAL1* promoter and selection is via *URA3* (Figure 2.3 A) (Alberti et al., 2007). There is some controversy regarding the subcellular localisation of the GTG/GPHRs (section 1.4.2). Will the expression of the Arabidopsis and *C. elegans* GTG/GPHRs in *S. cerevisiae* parallel any of the described localisations? For all intended studies, *S. cerevisiae* would therefore be a very useful system for characterising the GTG/GPHRs.

7.1.4 Aims

To study the GTG/GPHRs using *S. cerevisiae* as a model organism there are a number of aims:

1. Use the EUROSCARF database to identify an *S. cerevisiae gtg*-Δ deletion mutant.
2. Characterise the *gtg*-Δ mutant using a number of different growth conditions to investigate Gtgp function in *S. cerevisiae*.
3. Generate *S. cerevisiae* expression vectors for heterologous expression of the Arabidopsis and *C. elegans* GTG/GPHRs to:

a. Carry out functional complementation studies with these orthologues in the Sc *gtg*-Δ mutant.
b. Address the localisation of the Arabidopsis and *C. elegans* GTG/GPHRs in *S. cerevisiae*.
c. Develop a system for future purification of the Arabidopsis and *C. elegans* GTG/GPHRs

This will help further investigate conservation of function and the roles the GTG/GPHRs play in *S. cerevisiae*.

7.2 Results

7.2.1 Identification and characterisation of the Sc *gtg*-Δ mutant

In *S. cerevisiae* there appears to be only one *GTG* gene, according to the SGD (yeastgenome.org) and when searching for eukaryotic orthologues using the BLAST algorithm (section 2.9). The EUROSCARF database was used to identify the *S. cerevisiae GTG* deletion (*gtg*-Δ*)* mutant strain, YHR078W. The genotype of this strain is BY4741; *MAT a*; *his3*-Δ1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*:: *kanMX4*. This indicates that it is in the background of the reference WT strain, BY4741 (*MAT a*; *his3*-Δ1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0) and "*gtg*::*kanMX4*" denotes that the complete Sc *GTG* ORF has been replaced by the *kanMX4* gene; this is a hybrid selection marker conferring dominant selection on geneticin and kanamycin (Steiner and Philippsen, 1994, Wach et al., 1994). As shown above, the genotype of WT strain, BY4741 is *MAT a*; *his3*-Δ1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0. This indicates it has a number of deletions including the whole ORF deletion of *URA3* i.e. uracil must be supplemented in the media for both these strains to grow. "*MAT a*" signifies that the haploid cells of the strain produce "a-factor"; a mating pheromone (Betz et al., 1981). Other deletions include a partial ORF deletion of *HIS3*, imidazolegylcerol-phosphate dehydrogenase used in histidine biosynthesis (Alifano et al., 1996); as well as complete ORF deletions of *LEU2*, β-isopropylmalate dehydrogenase in the leucine biosynthesis pathway (Kohlhaw, 1987); and *MET15*, *O*-acetyl homoserine-*O*-acetyl serine sulfhydrylase used for methionine and cysteine biosynthesis (Brzywczy and Paszewski, 1993, Yamagata et al., 1994). The YHR078W strain name describes the location of Sc *GTG*: **Y**east chromosome **VIII** on the **R**ight arm at position **078** encoded by the "**W**atson" strand.

Genomic DNA was extracted from three independent colonies of both BY4741 (candidates a – c) and YHR078W (candidates d – f) strains and these were confirmed to be WT and deletion mutant (*gtg*-Δ), respectively using PCR (Figure 7.2 A). The primers ScGTG 1F and ScGTG 1R amplify the entire Sc *GTG* (Figure 7.3). The presence of the predicted product size (1663 bp) indicates WT Sc GTG^{+} allele (lanes 2 – 4), while the absence of this predicted band indicates the *gtg*-Δ (Figure 7.2 A; lanes 5 – 7). A fragment of *ACT1* (actin) was amplified using the genomic DNA extracted from BY4741 (candidates a – c) and YHR078W (candidates d – f); a predicted product of 92 bp shows the presence of actin (Figure 7.2 B). The results indicate that all six samples

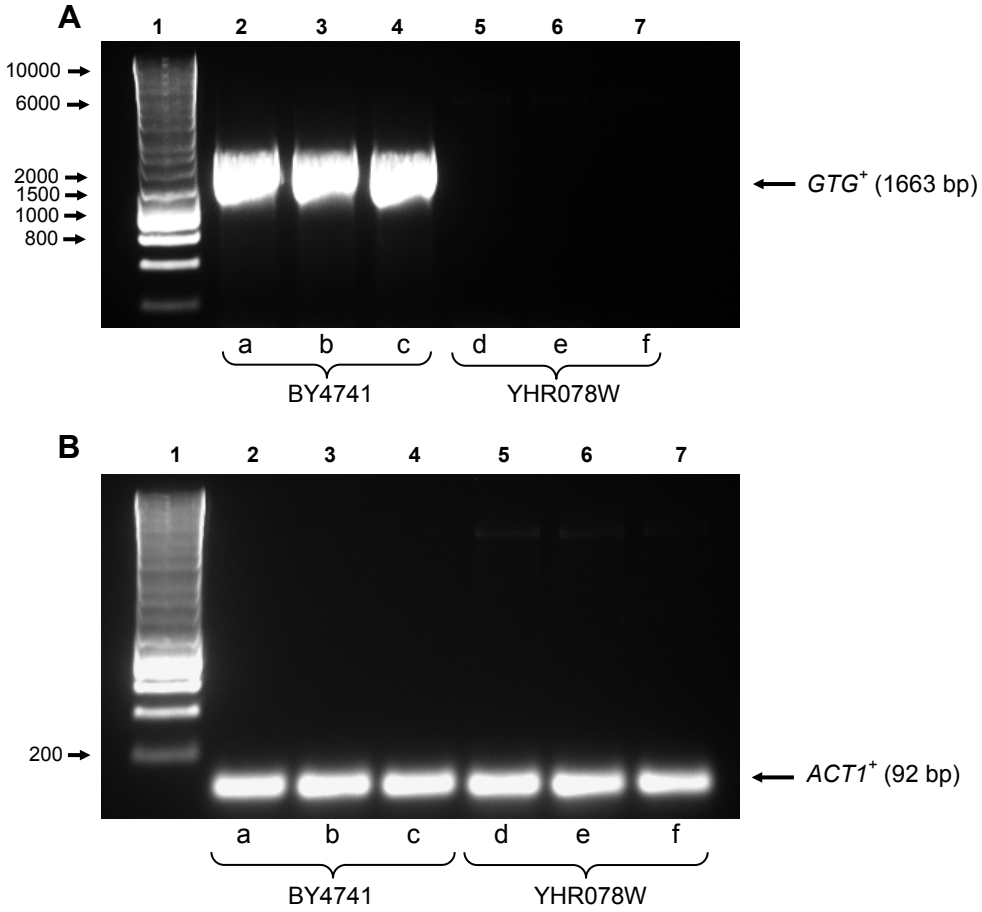


Figure 7.2 The YHR078W strain is the *S. cerevisiae gtg*-Δ mutant

PCR and gel electrophoresis to show the genotype of BY4741 (lanes 2 – 4) and YHR078W (lanes 5 – 7) using genomic DNA isolated from three independent colonies (a – f). Lane 1: molecular markers. **A**. Amplification of the Sc *GTG* gene using primers ScGTG 1F and ScGTG 1R (Figure 7.3). The predicted product (1663 bp) is shown. BY4741 contains WT Sc GTG^+ (GTG^+) and YHR078W contains mutant Sc *gtg*-Δ. **B**. Amplification of *ACT1* gene using primers Sc ACT1 F and Sc ACT1 R. The predicted product (92 bp) is shown ($ACT1^+$).

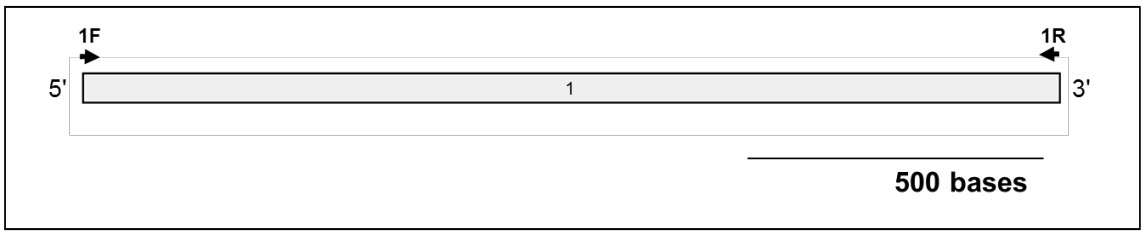


Figure 7.3 Genomic structure of the *S. cerevisiae GTG*

The genomic structure of the *S. cerevisiae GTG* gene comprising a single exon. Primers ScGTG 1F (1F) and ScGTG 1R (1R) are shown above the genomic structure (primer sequences outlined in Table 2.10). Scale bar: 500 bases.

of genomic DNA from BY4741 and YHR078W strains contain approximately equal levels of actin (Figure 7.2 B). The genomic structure shows that the Sc *GTG* gene contains only one exon and no introns (Figure 7.3). The coding and genomic DNA sequences provided by the SGD (systematic name: YHR078W) show that they are 100 % identical as analysed by pairwise alignment using the Needlemann-Wunsch alignment algorithm (http://www.ebi.ac.uk/Tools/psa/).

The SGD reveals a number of phenotypes associated with Sc *GTG* identifed by large-scale screens (Cherry et al., 2012). YHR078W is classed as a viable mutant (Giaever et al., 2002), which is useful as this mutant strain can be maintained. Stress resistance refers to the level of resistance to stress conditions and is related to cellular processes (Cherry et al., 2012). Sc *gtg*-Δ has been associated with decreased stress resistance under zinc-deficiency conditions. North et al. (2012) used zinc-limiting media (LZM) prepared as described in Gitan et al. (1988). The LZM contained 1 mM EDTA and 20 mM citrate to buffer metal availability (Gitan et al., 1998, North et al., 2012). The SGD described Sc *GTG* as a "high osmolarity-regulated protein of unknown function" (Cherry et al., 2012). In a large-scale screen, O'Rourke et al. (1998) identified genes related to the high osmolarity glycerol (HOG) mitogen-activated protein kinase pathway. Sc *GTG* was identified to be regulated by high osmolarity, which was induced by the use of sorbitol and potassium chloride (KCl) (O'Rourke and Herskowitz, 1998). These large-scale screens served as a basis for developing assays for revealing phenotypes of YHR078W differing to the reference (WT) strain.

7.2.2 Expressing the Arabidopsis and *C. elegans* GTG/GPHRs in *S. cerevisiae*

The Gateway system (as described in section 5.2.1) was employed for generating the *S. cerevisiae* expression vectors. The entry vectors pENTR/D *Cel-gphr-1* and pENTR/D *Cel-gphr-2* both without stop codons were constructed as in section 5.2.1. pENTR/D At *GTG1* and pENTR/D At *GTG2* both ± stop codons plasmid stocks were supplied for this project (Dr. L. Williams). The *Cel-gphr-1* and *Cel-gphr-2* without stop codons, At *GTG1* (± stop codon) and At *GTG2* without stop codon CDSs were inserted into both pAG426GAL-ccdB-EGFP and pYTV via LR recombination. CDSs inserted into either pAG426GAL-ccdB-EGFP or pYTV without stop codons have C-terminal tags. For example, pAG426GAL-ccdB-EGFP *Cel-gphr-1* without stop codon would be tagged with EGFP (see Figure 2.3 for vector maps). Restriction enzyme analysis

confirmed that these genes were cloned in the correct orientation as the predicted digest sizes were produced. For example, the predicted digest products for pYTV-*Cel-gphr-1* are 3760, 3328 and 654 bp (Figure 7.4 A; lane 2). Table 7.1 shows the constructs generated using this system. Restriction analyses were carried out as diagnostic checks for all *S. cerevisiae* expression vectors generated (Table 7.1) and their predicted product sizes are listed in Table 2.16. The expression vectors were sequenced in both forward and reverse directions using primers SK and EGFP-NRev (Table 2.10). The sequence results confirmed that the At *GTG* and *Cel-gphr* sequences corresponded to the CDSs provided by TAIR (At *GTG1* = AT1G64990, TAIR Accession # Locus:2010796; At *GTG2* = AT4G27630, TAIR Accession # Locus: 2137702) and WormBase (*Cel-gphr-1* = Y75B8A.16, WBGene00013551 and *Cel-gphr-2* = C11H1.2, WBGene00007528), and were inserted into pAG426GAL-ccdB-EGFP and pYTV vectors correctly.

The expression vectors were transformed into both BY4741 and YHR078W strains for functional analysis (section 7.2.3) and localisation studies (section 7.2.4). They were also transformed into *zrc1*-Δ *cot1*-Δ double mutants (Table 7.2). Transformed strains were grown on SC-ura media to select for positive transformants. Each colony that grew on selective media would have resulted from the transformation of a single *S. cerevisiae* cell containing the plasmid of interest i.e. independent transformant; Table 7.2 displays the transgenic *S. cerevisiae* lines generated in this way. Colony PCR was performed on candidate *S. cerevisiae* colonies to confirm that they contained a transformed plasmid. Figure 7.4 B shows an example of colony PCR for identifying pYTV *Cel-gphr-1* transformed into BY4741. The amplification of a predicted 1399 bp product (using CeGTG1 F and CeGTG1 NO STOP R primers) indicated the presence of the pYTV-*Cel-gphr-1* in BY4741 (Figure 7.4 B; lane 2). Untransformed *S. cerevisiae* colonies would not have grown on Sc-ura media and hence could not be used as a control. However a negative control using untransformed strain DNA could have been used.

7.2.3 Characterisation of the *S. cerevisiae gtg*-Δ mutant

S. cerevisiae drop tests were used to observe if there were any growth differences between the BY4741 (containing WT Sc *GTG*) and YHR078W (Sc *gtg*-Δ) strains under different conditions. The BY4741 and YHR078W strains were grown on solid SC

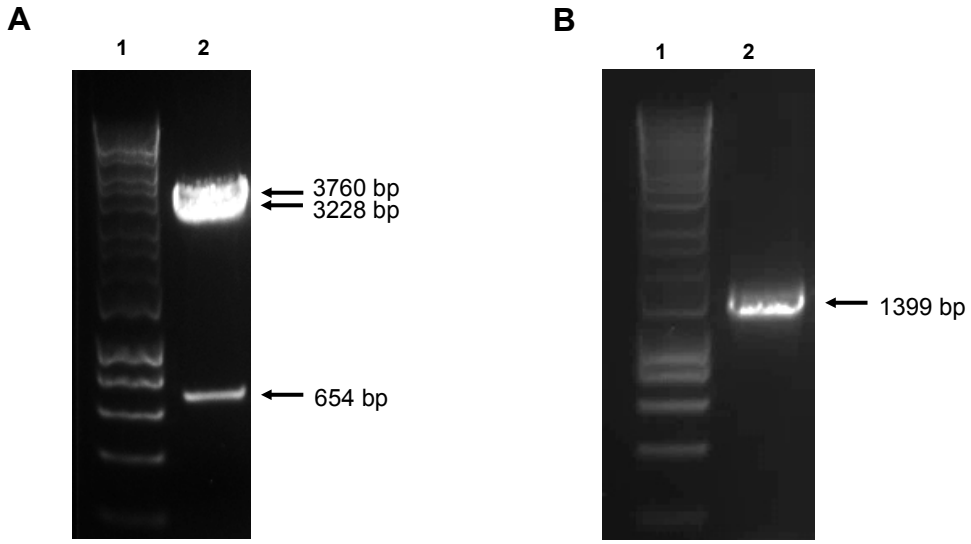


Figure 7.4 Diagnostic analyses for pYTV-*Cel-gphr-1* transformed into *S. cerevisiae*

Restriction analysis showing *Cel-gphr-1* inserted correctly into pYTV vector (**A**), followed by successful transformation into *S. cerevisiae* (**B**). **A**. Restriction enzyme analysis of pYTV-*Cel-gphr-1* using *Eco*RV and *Xho*I. The predicted product sizes for *Cel-gphr-1* inserted into the vector are labelled on figure (lane 2: 3760, 3328 and 654 bp). **B**. Colony PCR to show presence of the transformed plasmid, pYTV-*Cel-gphr-1* in WT (BY4741) strain (lane 2). Amplification of the predicted product for *Cel-gphr-1* (1399 bp), amplified using primers CeGTG1 F and CeGTG1 NO STOP R. Lane 1: molecular markers (see Figure 7.2 for predicted molecular marker sizes).

Table 7.1 *S. cerevisiae* expression constructs generated using the Gateway system

Nomenclature used is specific for each organism (see section 2.1).

Expression construct	Details
pAG426GAL-*Cel-gphr-1::egfp*	*Cel-gphr-1* C-terminally tagged with EGFP under the *GAL1* promoter
pAG426GAL-*Cel-gphr-2::egfp*	*Cel-gphr-2* C-terminally tagged with EGFP under the *GAL1* promoter
pAG426GAL-At *GTG1:EGFP*	At *GTG1* C-terminally tagged with EGFP under the *GAL1* promoter
pAG426GAL-At *GTG2:EGFP*	At *GTG2* C-terminally tagged with EGFP under the *GAL1* promoter
pAG426GAL-At *GTG1*	At *GTG1* under the *GAL1* promoter – can be used as a control for pAG426GAL-At *GTG1:EGFP*
pYTV-*Cel-gphr-1*	*Cel-gphr-1* C-terminally tagged with 3XFLAG, 6XHis, 3C cleavage and 2XIgG under the *GAL1* promoter
pYTV-*Cel-gphr-2*	*Cel-gphr-2* C-terminally tagged with 3XFLAG, 6XHis, 3C cleavage and 2XIgG under the *GAL1* promoter
pYTV-At *GTG1*	At *GTG1* C-terminally tagged with 3XFLAG, 6XHis, 3C cleavage and 2XIgG under the *GAL1* promoter
pYTV-At *GTG2*	At *GTG2* C-terminally tagged with 3XFLAG, 6XHis, 3C cleavage and 2XIgG under the *GAL1* promoter

Table 7.2 Transgenic *S. cerevisiae* strains generated

S. cerevisiae strains transformed with expression constructs (Table 7.1). Refer to section 2.1 for nomenclature. # = number of independent transformants.

#	Strain	Expression construct	Genotype
1	BY4741	pAG426GAL-*Cel-gphr-1::egfp*	[pAG426GAL-*Cel-gphr-1::egfp*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0
4	YHR078W		[pAG426GAL-*Cel-gphr-1::egfp*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*::*kanMX4*
4	*zrc1*-Δ *cot1*-Δ		[pAG426GAL-*Cel-gphr-1::egfp*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *zrc1*::natMX *cot1*::*kanMX4*
4	BY4741	pAG426GAL-*Cel-gphr-2::egfp*	[pAG426GAL-*Cel-gphr-2::egfp*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0
3	YHR078W		[pAG426GAL-*Cel-gphr-2::egfp*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*::*kanMX4*
4	*zrc1*-Δ *cot1*-Δ		[pAG426GAL-*Cel-gphr-2::egfp*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *zrc1*::natMX *cot1*::*kanMX4*
1	BY4741	pAG426GAL-At *GTG1:EGFP*	[pAG426GAL-At *GTG1:EGFP*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0
1	YHR078W		[pAG426GAL-At *GTG1:EGFP*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*::*kanMX4*
1	BY4741	pAG426GAL-At *GTG2:EGFP*	[pAG426GAL-At *GTG2:EGFP*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0
1	YHR078W		[pAG426GAL-At *GTG2:EGFP*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*::*kanMX4*
1	BY4741	pAG426GAL-At *GTG1*	[pAG426GAL-At *GTG1*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0
1	YHR078W		[pAG426GAL-At *GTG1*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*::*kanMX4*
4	BY4741	pYTV-*Cel-gphr-1*	[pYTV-*Cel-gphr-1*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0
4	YHR078W		[pYTV-*Cel-gphr-1*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*::*kanMX4*
3	BY4741	pYTV-*Cel-gphr-2*	[pYTV-*Cel-gphr-2*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0
2	YHR078W		[pYTV-*Cel-gphr-2*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*::*kanMX4*
4	*zrc1*-Δ *cot1*-Δ		[pYTV-*Cel-gphr-2*] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *zrc1*::natMX *cot1*::*kanMX4*
4	BY4741	pAG426GAL-ccdB-EGFP	[pAG426GAL-ccdB-EGFP] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0
4	YHR078W		[pAG426GAL-ccdB-EGFP] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*::*kanMX4*
4	*zrc1*-Δ *cot1*-Δ		[pAG426GAL-ccdB-EGFP] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *zrc1*::natMX *cot1*::*kanMX4*
4	BY4741	pYTV	[pYTV] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0
4	YHR078W		[pYTV] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*::*kanMX4*
4	Δ*zrc1*Δ *cot1*		[pYTV] *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *zrc1*::natMX *cot1*::*kanMX4*

media containing glucose to see if there was any difference in growth. The results do not display any difference between BY4741 and YHR780w strains (control panel of Figure 7.5). Therefore, different growth conditions were successively tested to try and reveal a phenotype for the YHR078W strain.

Sc *GTG* is defined as a high osmolarity-regulated gene of unknown function following a large-scale screen (O'Rourke and Herskowitz, 1998). This formed the basis for using sorbitol and KCl to create high osmolarity conditions for comparing the growth of BY4741 and YHR078W. No growth difference was observed between BY4741 and YHR078W when grown on SC media containing either sorbitol or KCl (Figure 7.5). Similarly, SC media with a varying pH ranging from 4.8 – 6.5 showed no growth differences between BY4741 and YHR078W (Figure 7.6). YHR078W exhibited decreased stress resistance under zinc-deficient conditions (North et al., 2012). In this large-scale screen, zinc was limited using 1 mM EDTA and 20 mM citrate (Gitan et al., 1998). For this reason, EDTA treatment was tested as it chelates a range of divalent cations. BY4741 and YHR078W were grown on solid SC media containing glucose with 0.5 or 1 mM EDTA. Results show that there was no difference in growth between BY4741 and YHR078W in 0.5 mM EDTA; however, YHR078W exhibited a slight decrease in growth in 1 mM EDTA (Figure 7.7). But as the experiment was only carried out once, this would need to be verified.

A growth assay was set up to test the heterologous expression of At *GTG1* and At *GTG1:EGFP*. BY4741 and YHR078W strains were grown alongside BY4741 and YHR078W expressing At *GTG1*, At *GTG1:EGFP* or the empty vector (pAG426GAL-ccdB-EGFP). These strains were grown under the same conditions as the experiment shown in Figure 7.7, but galactose was used as the carbon source instead of glucose to induce the expression of At *GTG1* and At *GTG1:EGFP* (Figure 7.8). No growth was seen in either BY4741 or YHR078W since no uracil was supplied (Figure 7.8). In 0 mM EDTA, no differences in growth was seen between untransformed BY4741 and YHR078W strains; or BY4741 and YHR078W transformed with the empty vector, pAG426GAL-At *GTG1*, or pAG426GAL-At *GTG1:EGFP* (Figure 7.8). Even at 3 d, growth of all strains was severely reduced in 0.5 mM EDTA and even more so in 1 mM EDTA (Figure 7.8). Therefore, the experiment was repeated using a range of lower EDTA concentrations. Figure 7.9 shows the comparison of BY4741 and YHR078W expressing At *GTG1*, At *GTG1:EGFP* or the empty vector grown together on solid

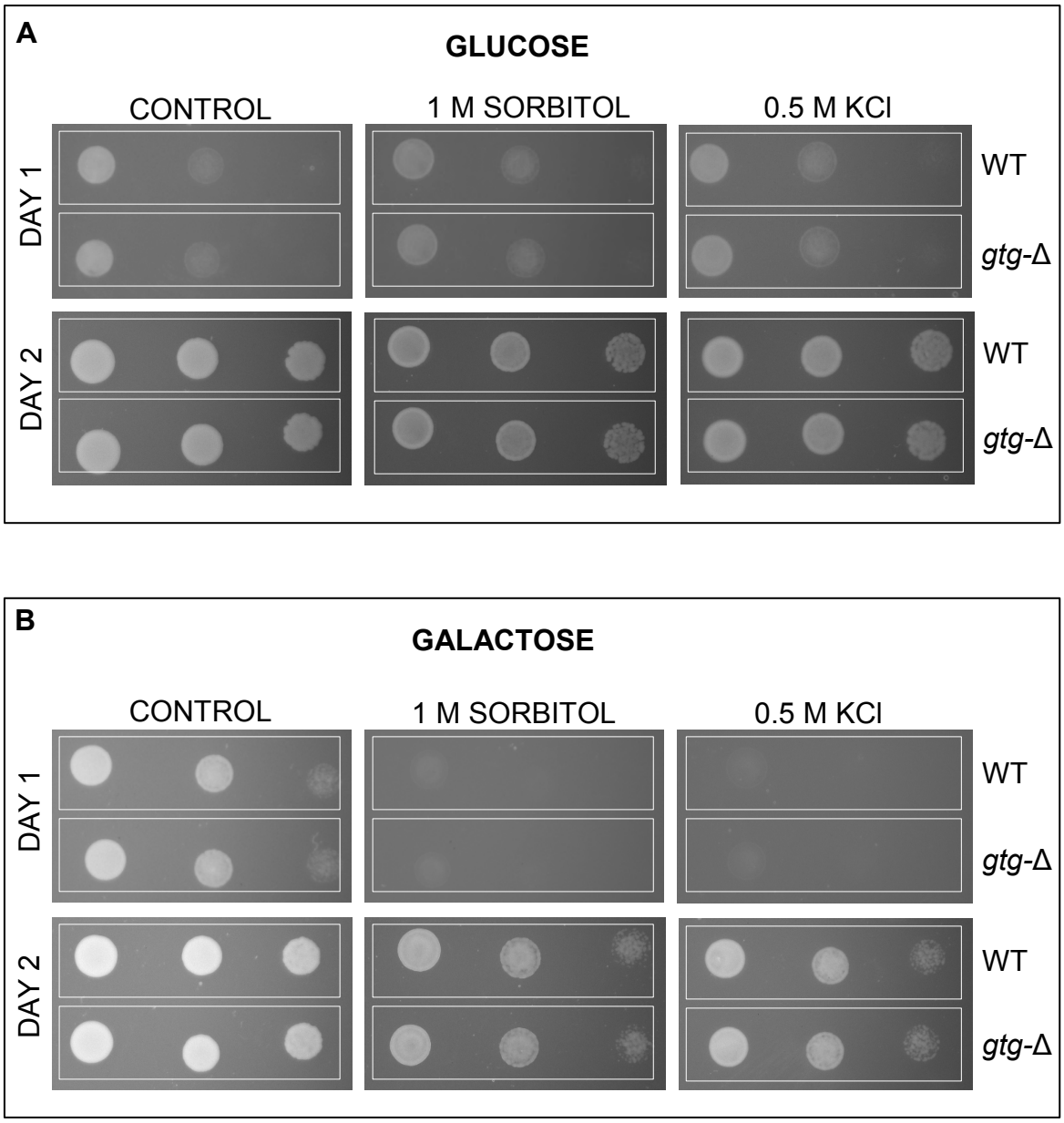


Figure 7.5 Osmotic stress has no effect on the growth of the *gtg*-Δ mutant compared to WT

Images of BY4741 (WT) and YHR078W (*gtg*-Δ) strains grown under osmotic stress caused by 1 M sorbitol or 0.5 M potassium chloride (KCl). Dilution series of WT and mutant *gtg*-Δ strains spotted onto either (**A**) 2 % glucose or (**B**) 2 % galactose SC media. Number of days grown and treatment are labelled on figure. Dilution series from left to right: 1, 1/10, 1/100. Three spots for each dilution were used per strain; WT and mutant *gtg*-Δ were grown on the same plate. Images are from one experiment using three replicate plates.

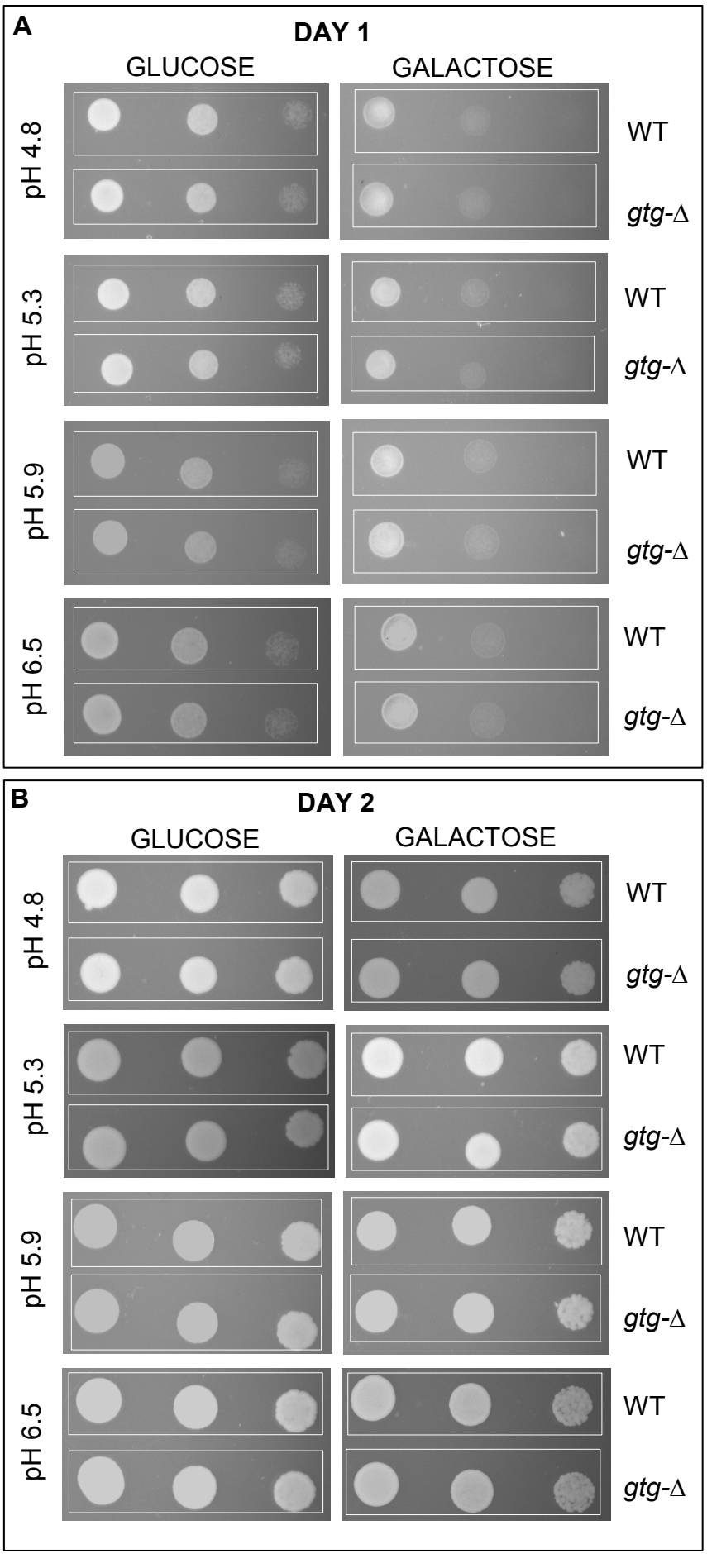

A
DAY 1
GLUCOSE
GALACTOSE
pH 4.8
pH 5.3
pH 5.9
pH 6.5
WT
gtg-Δ
B
DAY 2
GLUCOSE
GALACTOSE
pH 4.8
pH 5.3
pH 5.9
pH 6.5
WT
gtg-Δ

Figure 7.6 Altering pH has no effect on the growth of the *gtg*-Δ mutant compared to WT

Images to show there is no difference between the BY4741 (WT) and YHR078W (*gtg*-Δ) strains under a range of pH conditions after 1 d (**A**) or 2 d (**B**) of growth. Dilution series of WT and *gtg*-Δ mutant strains spotted onto SC media containing either 2 % galactose or 2 % glucose. Dilution series from left to right: 1, 1/10, 1/100. Three spots for each dilution were used per strain; WT and *gtg*-Δ mutant were grown on each plate. Images are from one experiment using two replicate plates.

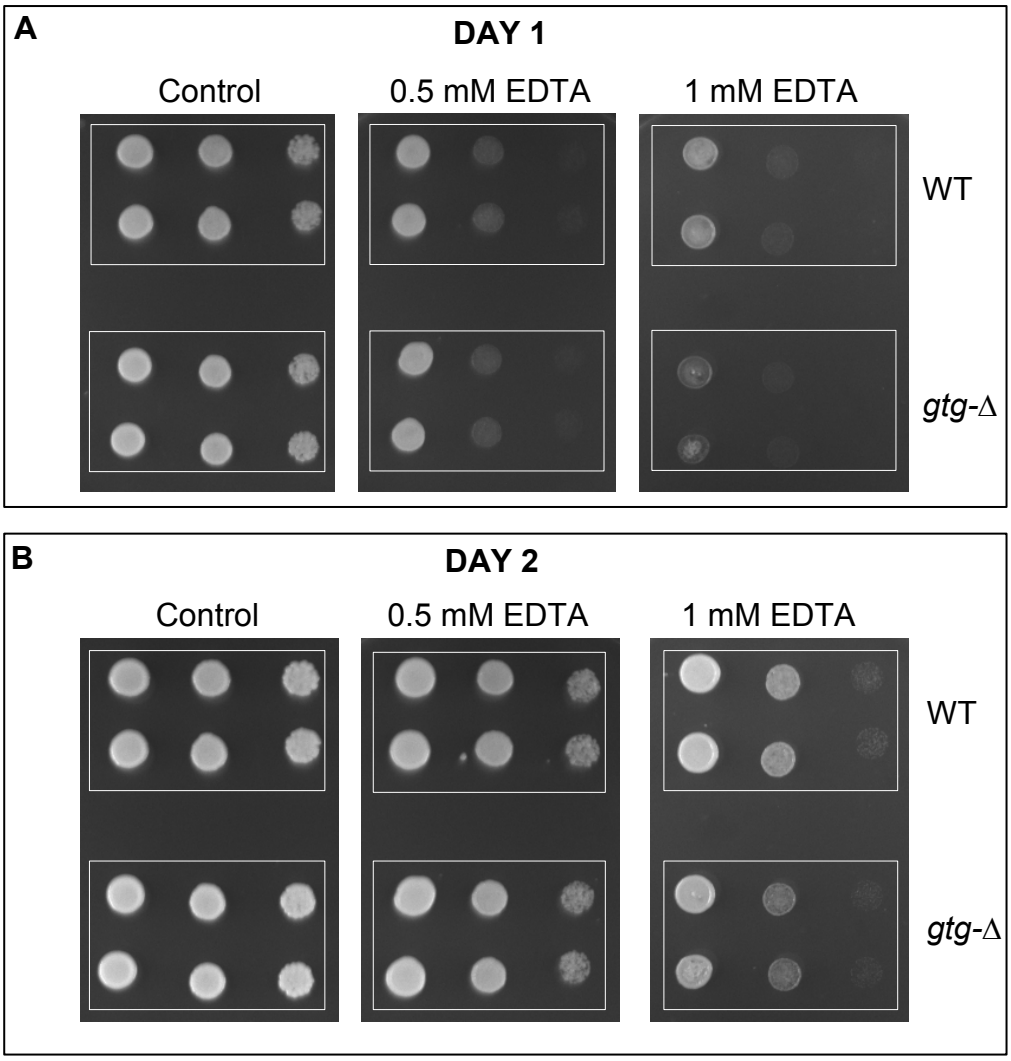


Figure 7.7 *gtg*-Δ mutant growth compared to WT is very slightly reduced with EDTA treatment on glucose

Images to show BY4741 (WT) and YHR078W (*gtg*-Δ) strain growth after 1 d (**A**) and 2 d (**B**) in SC media (control) supplemented with 0.5 mM EDTA and 1 mM EDTA. Dilution series of the WT and *gtg*-Δ mutant strains spotted onto SC 2% glucose media containing 0 (control), 0.5 or 1 mM EDTA. 1 mM EDTA slightly reduced *gtg*-Δ mutant growth compared to WT. Dilution series from left to right: 1, 1/10, 1/100. Three spots for each dilution were used per strain, with two strains per plate. Images are from one experiment using three replicate plates for each EDTA concentration.

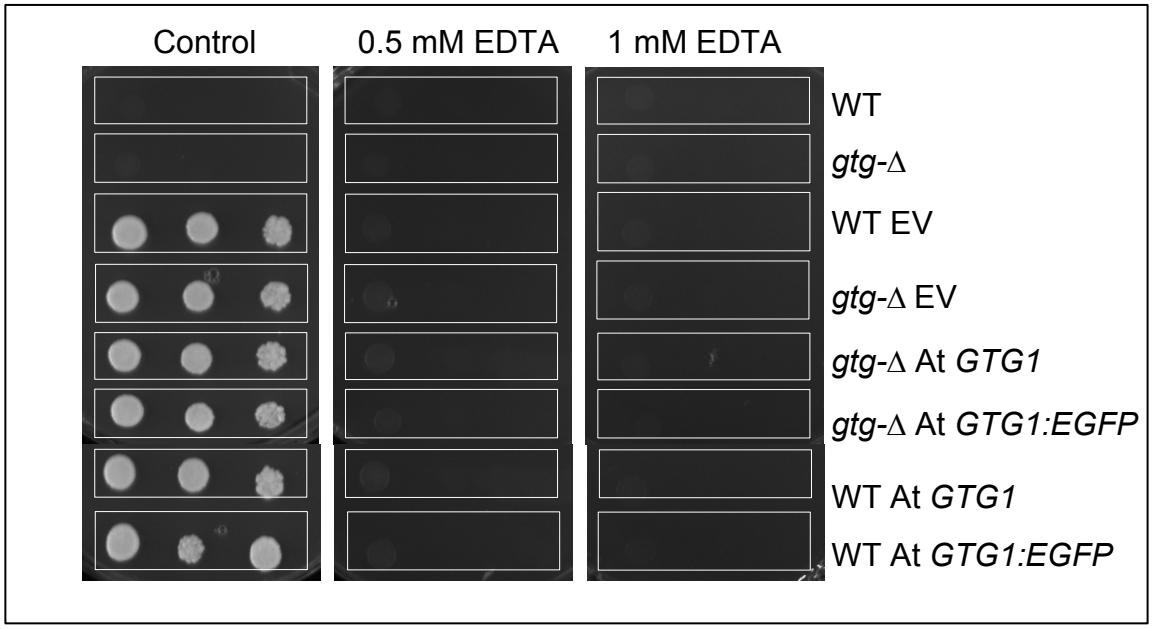


Figure 7.8 *gtg*-Δ mutant growth with EDTA treatment on galactose

BY4741 (WT) and YHR078W (*gtg*-Δ) strains grown on SC-ura media (control) containing 0.5 and 1 mM EDTA at 3 d. Dilution series of WT and *gtg*-Δ mutant strains transformed with the empty vector (EV; pAG426GAL-ccdB-EGFP), pAG426GAL-At *GTG1* (At *GTG1*) or pAG426GAL-At *GTG1:EGFP* (At *GTG1:EGFP*) spotted onto SC-ura 2 % galactose media containing 0, 0.5 or 1 mM EDTA. Growth for 3 d indicates that 0.5 and 1 mM EDTA inhibited the growth of all strains. Dilution series from left to right: 1, 1/10, 1/100. Images are from one experiment with three plates used for each EDTA concentration.

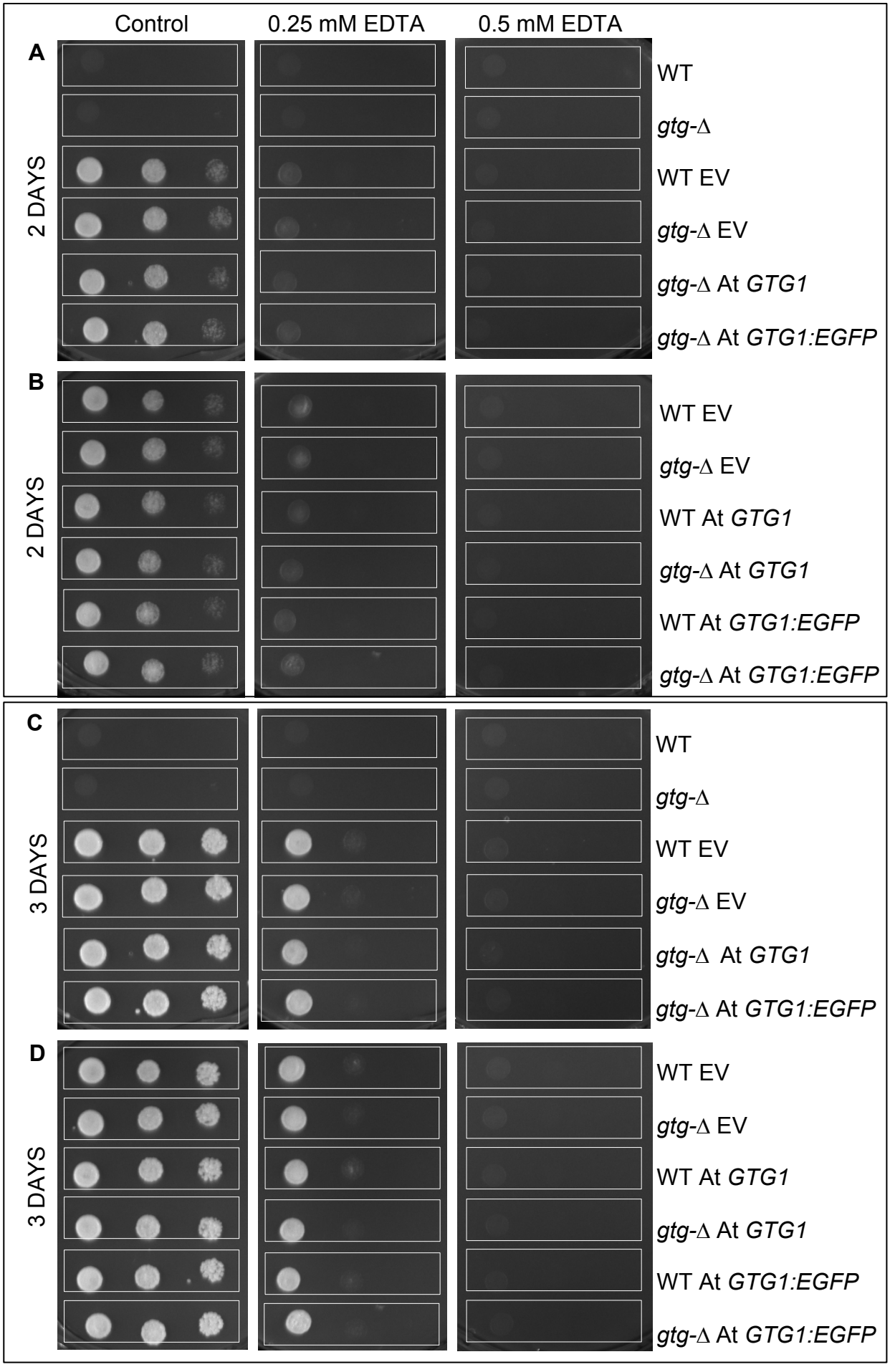

Control
0.25 mM EDTA
0.5 mM EDTA
A
2 DAYS
WT
gtg-Δ
WT EV
gtg-Δ EV
gtg-Δ At GTG1
gtg-Δ At GTG1:EGFP
B
2 DAYS
WT EV
gtg-Δ EV
WT At GTG1
gtg-Δ At GTG1
WT At GTG1:EGFP
gtg-Δ At GTG1:EGFP
C
3 DAYS
WT
gtg-Δ
WT EV
gtg-Δ EV
gtg-Δ At GTG1
gtg-Δ At GTG1:EGFP
D
3 DAYS
WT EV
gtg-Δ EV
WT At GTG1
gtg-Δ At GTG1
WT At GTG1:EGFP
gtg-Δ At GTG1:EGFP

Figure 7.9 The effect of EDTA on growth of *gtg*-Δ mutant on galactose

Images of BY4741 (WT) and YHR078W (*gtg*-Δ) strains transformed with pAG426GAL-At *GTG1* or pAG426GAL-At *GTG1:EGFP* grown on 0, 0.25 mM and 0.5 mM EDTA. **A – D** display a combination of dilution series of WT and *gtg*-Δ mutant strains transformed with empty vector (EV; pAG426GAL-ccdB-EGFP), pAG426GAL-At *GTG1* (At *GTG1*) or pAG426GAL-At *GTG1:EGFP* (At *GTG1:EGFP*) spotted onto SC-ura 2 % galactose media containing 0, 0.25 or 0.5 mM EDTA. **A – D** show different combinations of transgenic WT and *gtg*-Δ strains in the same experiment to highlight that growth spots are similar across different plates. **A** and **C** are the same plate after 2 and 3 d growth respectively, showing the WT and *gtg*-Δ strains, WT and *gtg*-Δ transformed with the empty vector, and the *gtg*-Δ transformed with At *GTG1* or At *GTG1:EGFP*. **B** and **D** are the same plate after 2 and 3 d growth respectively showing WT and *gtg*-Δ transformed with the empty vector, WT and *gtg*-Δ transformed with At *GTG1* or At *GTG1:EGFP*. Dilution series from left to right: 1, 1/10, 1/100. Images are from one experiment and representative of six plates for each EDTA concentration.

SC-ura media containing 0, 0.25 or 0.5 mM EDTA. These images are from the same experiment but different combinations of strains were grown alongside to highlight the similarities or differences of growth compared to different controls (empty vector or untransformed) (Figure 7.9). Similar to previous results (Figure 7.8), there was no growth observed in BY4741 and YHR078W without uracil supplied in the media (Figure 7.9). At 2 d, the 0.25 mM EDTA treatment resulted in very little growth for both BY4741 and YHR078W expressing At *GTG1*, At *GTG1:EGFP* or the empty vector (Figure 7.9). At 3 d, these strains had grown more than at 2 d, but still there was no difference observed between any of the strains (Figure 7.9). In the presence of 0.5 mM EDTA, negligible growth can be seen in any of the strains, which was consistent with the results shown in Figure 7.8. These results suggest that no robust phenotypes specific to the loss of Sc *GTG* or At GTG1 expression can be revealed by using these particular pH, osmotic stress or EDTA treatments.

7.2.4 Localisation of Arabidopsis and *C. elegans* GTG/GPHRs in *S. cerevisiae*

Confocal microscopy was used to visualise the localisation of the Arabidopsis and *C. elegans* GTG/GPHRs in *S. cerevisiae*. Figure 7.10 (I and J) shows that as expected, EGFP expression alone is distributed throughout the cytosol (Alberti et al., 2007). The Arabidopsis and *C. elegans* GTG/GPHR localise to punctate structures (Figure 7.10 and Figure 7.11), which is characteristic of Golgi localisation as well as additional intracellular structures that resemble ER localisation (Figure 7.11 and Figure 7.12). Figure 7.10 – Figure 7.12 are representative images of the localisation results for At GTG1 from 15 cells, At GTG2 from four cells, *Cel*-GPHR-1 from 15 cells and *Cel*-GPHR-2 from eight cells; all from a single BY4741 transformant for each construct (Table 7.2). More than one image for each construct is shown in this section as certain images appeared more punctate (possibly Golgi localisation), while others were more reminiscent of ER localisation. Furthermore no the images indicate no evidence for plasma membrane localization in any of the cells observed.

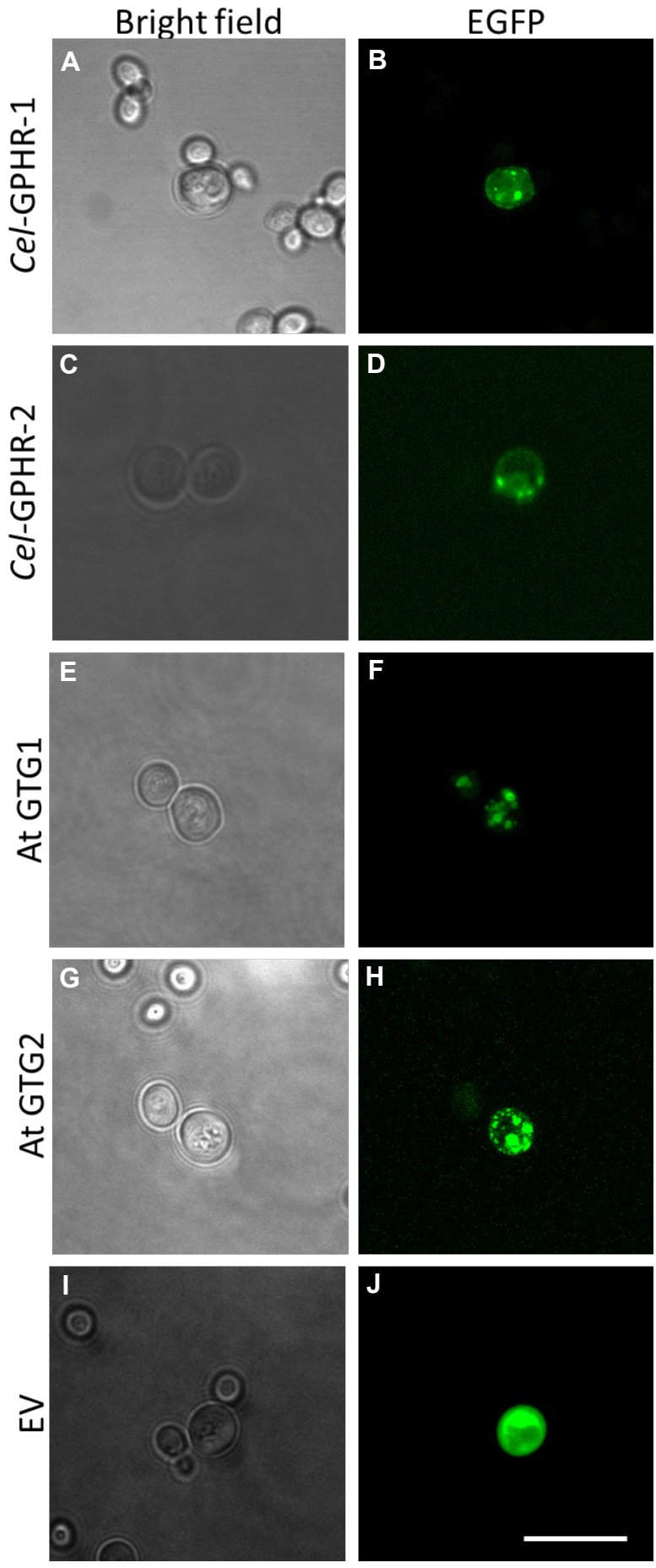

Bright field
EGFP
A
B
Cel-GPHR-1
C
D
Cel-GPHR-2
E
F
At GTG1
G
H
At GTG2
I
J
EV

Figure 7.10 Arabidopsis and *C. elegans* GTG/GPHRs localise to punctate structures

Bright field and fluorescent (maximum projections) images of *S. cerevisiae* BY4741 (WT) expressing GTG/GPHRs with C-terminal EGFP tag under the *GAL1* promoter. **A** and **B**. *Cel*-GPHR-1. **C** and **D**. *Cel*-GPHR-2. **E** and **F**. At GTG1. **G** and **H**. At GTG2. **I** and **J**. Empty vector, pAG426GAL-ccdB-EGFP. These are representative images of the punctate foci seen for the Arabidopsis and *C. elegans* GTG/GPHRs in one transformant for each At *GTG1:EGFP*, At *GTG2:EGFP*, *Cel-gphr-1::egfp*, *Cel-gphr-2::egfp*. Scale bar: 10 μm.

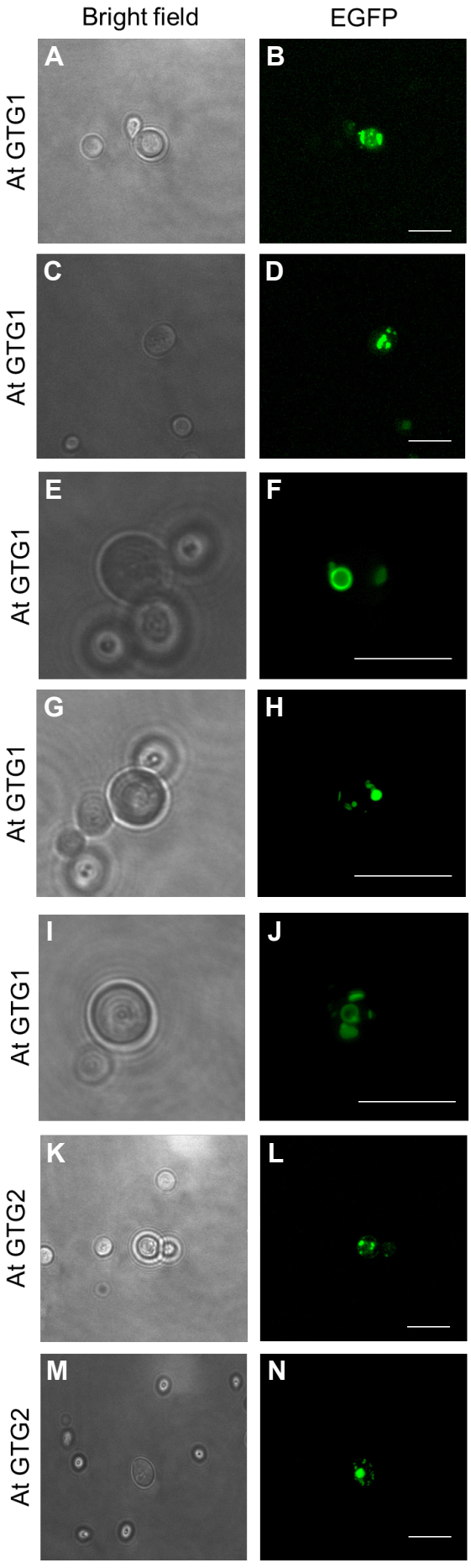

Bright field
EGFP
A
B
At GTG1
C
D
At GTG1
E
F
At GTG1
G
H
At GTG1
I
J
At GTG1
K
L
At GTG2
M
N
At GTG2

Figure 7.11 The Arabidopsis GTGs localise intracellularly

Bright field and fluorescent (maximum projection) images of *S. cerevisiae* expressing GTG/GPHRs with C-terminal EGFP tag under the *GAL1* promoter. BY4741 expressing At *GTG1:EGFP* (**A – J**) and At *GTG2:EGFP* (**K – N**) fusion proteins. **A** – **D** show the punctate foci of At *GTG1:EGFP* expression in *S. cerevisiae* cells. **E – J**. At *GTG1:EGFP* expression that is characteristic of ER localisation. **K – N**. At *GTG2:EGFP* fusion protein resembling punctate foci (**L**) and intracellular structures which could resemble ER localisation (**N**). Scale bar: 10 μm. These are representative images from different cells using the same *S. cerevisiae* transformants as in Figure 7.10.

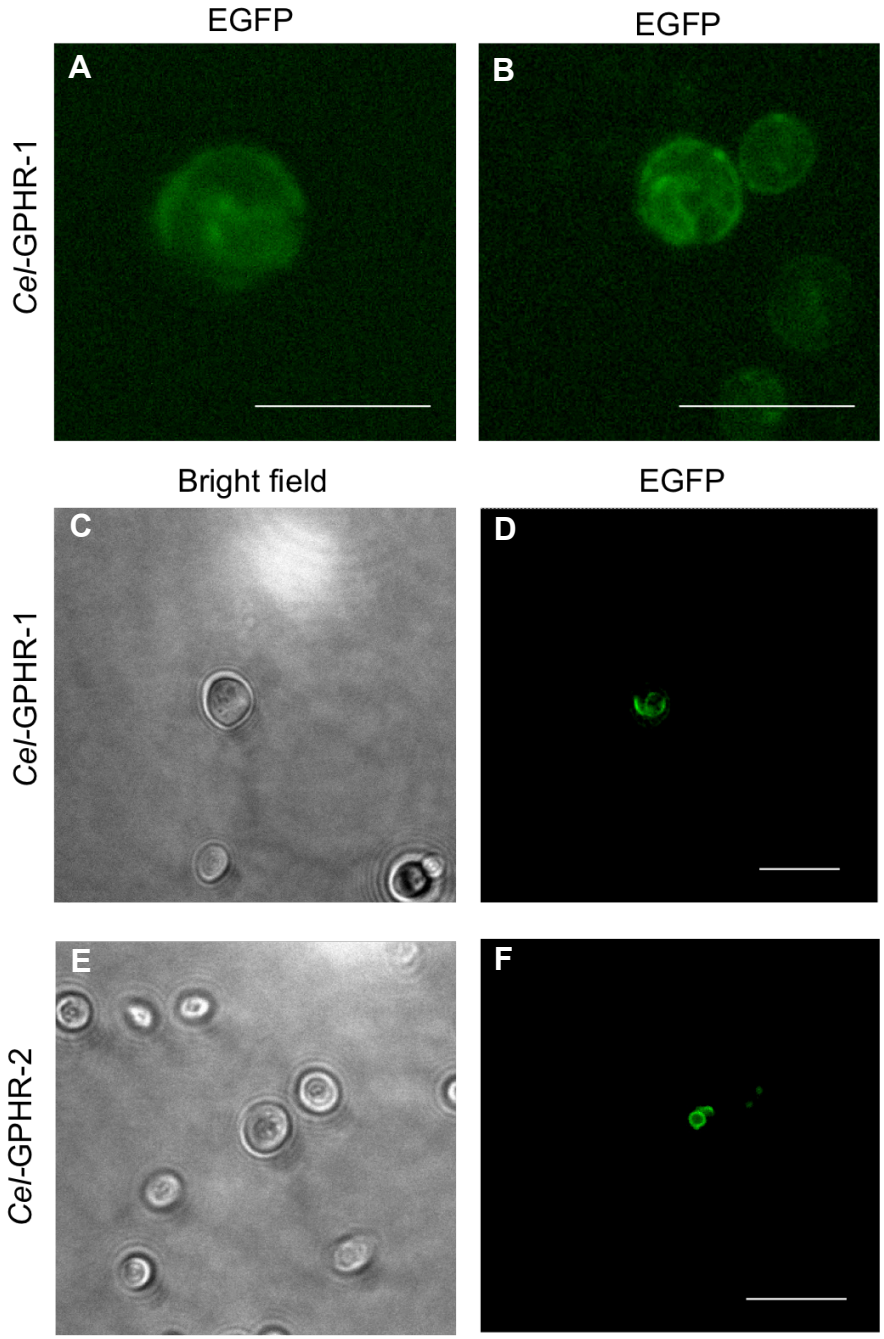


Figure 7.12 *C. elegans* GPHR:EGFP fusion proteins show ER localisation-pattern in some cells

Bright field and fluorescent (maximum projections) images of *S. cerevisiae* expressing GTG/GPHRs with C-terminal EGFP tag under the *GAL1* promoter. BY4741 expressing *Cel-gphr-1::egfp* (**A – D**; *Cel*-GPHR-1) and *Cel-gphr-2::egfp* (**E – F**; *Cel*-GPHR-2) fusion proteins. **A – B**. High magnification fluorescent images to show the localisation of *Cel-gphr-1::egfp* that resembles ER localisation using two different cells from one transformant. **C – F**. Localisation of *Cel-gphr-2::egfp* that resembles ER localisation. Scale bars: 5 µm. These are representative images from different cells using the same *S. cerevisiae* transformants as in Figure 7.10.

7.3 Discussion

The function and localisation of the GTG/GPHRs has been explored in Arabidopsis (Pandey et al., 2009, Jaffé et al., 2012), a CHO cell line (Maeda et al., 2008), mouse (Tarutani et al., 2012), *C. elegans* (CHAPTER 1) and *Drosophila* (Charroux and Royet, 2014). The ABA-At GTG1 interaction has been explored in a fungal system (Kharenko et al., 2013); but functional complementation or localisation of the GTG/GPHRs are yet to be investigated in the fungal kingdom.

There is only one *GTG* gene present in the *S. cerevisiae* genome, like humans, mice and *Drosophila*. The EUROSCARF database was used to identify the *gtg*-Δ deletion mutant strain, YHR078W. The library that YHR078W originated from provided its genotype: *MAT a*; *his3*-Δ 1; *leu2*-Δ 0; *met15*-Δ 0; *ura3*-Δ 0; *gtg*::*kanMX4*, (section 7.2.1). This indicates that Sc *GTG* was replaced by *kanMX4*. Primers were designed to amplify the whole CDS of Sc *GTG*. PCR and gel electrophoresis was used to confirm that YHR078W was a deletion mutant, as shown by the absence of the predicted 1663 bp product (Figure 7.2 A). The genomic DNA isolated from BY4741 contained the Sc *GTG* indicating that it is the WT strain, as shown by the presence of the predicted 1663 bp product (Figure 7.2 A). The *ACT1* gene (predicted 92 bp product) was amplified using the same samples of genomic DNA as in Figure 7.2 A. This showed that the absence of the predicted 1663 bp product for Sc *GTG* in YHR078W (Figure 7.2 A) was due to the absence of the *GTG* gene and not that there was a problem with the genomic DNA isolated (Figure 7.2 B). The Sc *GTG* only contains one exon and no introns in its genomic sequence (Figure 7.3) and the CDS and genomic sequence are identical. The Sc *GTG* genomic structure contrasts to that of both the Arabidopsis and *C. elegans GTG/GPHRs* as these contain many exons and introns (Figure 3.1). The lack of introns is typical for the *S. cerevisiae* genome (Stajich et al., 2007); some fungal species including *S. cerevisiae* (Hirschman et al., 2006) have low intron densities in their genome while others such as *Neurospora crassa* have much higher intron densities (Galagan et al., 2003).

To determine whether the Arabidopsis and *C. elegans* GTG/GPHRs are able to complement any defects of the *S. cerevisiae gtg*-Δ mutant, expression vectors were constructed containing the CDSs of *Cel-gphr* and At *GTG* genes (section 7.2.2). These were also used to study the localisation of these proteins in *S. cerevisiae* (section 7.2.4).

However, to investigate the conservation of function using *S. cerevisiae*, phenotypic characterisation of the Sc *gtg*-Δ mutant was first conducted.

7.3.1 EDTA treatment slightly reduced growth of the *gtg*-Δ mutant

The *gtg*-Δ mutant was subjected to various growth conditions to try to identify a phenotype. In a large-scale survey, YHR078W was associated with having a decreased stress resistance to zinc deficiency using EDTA and citrate (North et al., 2012). This served as the basis for a phenotypic assay. EDTA is a chelating agent owing to its hexadentate structure and therefore is able to sequester divalent cations such as zinc, iron, calcium and magnesium. BY4741 and YHR078W were grown in media containing EDTA to see if cation sequestration would have an effect on their growth. The results indicate that in the presence of glucose, 1 mM EDTA very slightly reduced the growth of YHR078W compared to BY4741, however this was only carried out once (Figure 7.7). In the future, this could be examined more closely by conducting a growth curve assay whereby the absorbance over time is followed. The EDTA effect was very small, so other growth conditions were tested in order to further characterise Sc Gtgp.

There was no difference in growth detected between the BY4741 and YHR078W strains when grown in SC media, using either glucose (Figure 7.5) or galactose (Figure 7.7) as the carbon source. However, both BY4741 and YHR078W strains grew better on SC glucose media (Figure 7.5) than on SC galactose media (Figure 7.6), since the preferred carbon source of *S. cerevisiae* is glucose (Timson, 2007). Sc *GTG* was classified as a high osmolarity-regulated gene in a large-scale screen (O'Rourke and Herskowitz, 1998); for this reason BY4741 and YHR078W were grown under osmotic stress conditions, which can be induced using sorbitol or KCl (Schüller et al., 1994, O'Rourke and Herskowitz, 1998, Giaever et al., 2002, O'Rourke and Herskowitz, 2004). Both BY4741 and YHR078W strains exhibited reduced growth on either sorbitol or KCl treatment compared to SC media, but there was no difference observed between the two strains (Figure 7.5). Deletion mutants of *GPD1, PBS2* and *HOG1* displayed significantly reduced fitness under high osmolarity conditions (Giaever et al., 2002). These genes play important roles in the high osmolarity glycerol (HOP) signal transduction cascade of osmotic homeostasis (Posas et al., 1996, Giaever et al., 2002). Since YHR078W exhibited a decreased growth similar to BY4741 on solid media (Figure 7.5) it suggests that osmotic stress reduced growth for both strains, but

YHR078W had no defect in sensing changes in the conditions tested. Perhaps different high molarity media could be used to try and identify a mutant phenotype.

The BY4741 and YHR078W strains were also grown under a range of pHs, from 4.8 to 6.5, which produced negligible differences in growth, and no differences between the two strains (Figure 7.6). This suggests that these pH conditions tested did not have an effect on the growth of either BY4741 or YHR078W. However the media was only adjusted to these pHs using potassium hydroxide or hydrochloric acid, and not buffered using a buffering agent such as HEPES, MOPS or MES (Kitamura and Itoh, 1987). Furthermore, in a large-scale screen, YHR078W was identified to have an increased resistance to (S)-lactic acid in YPD media (equivalent to SC media) at pH 2.6 (Suzuki et al., 2013). This result suggests that there could be an issue with osmolarity sensing or regulation since it cannot respond normally to the presence of (S)-lactic acid/acidic conditions. Therefore it cannot be concluded that pH does not have an effect on YHR078W with the conditions used and further testing would be required as there is still scope for pH having an effect on YHR078W. To test the effect of pH more accurately, a larger range of different pH conditions obtained with buffering agents could be used.

7.3.2 Heterologous expression of the Arabidopsis and *C. elegans* GTG/GPHRs

EDTA caused a slightly reduced growth in YHR078W compared to BY4741 (Figure 7.7). Arabidopsis or *C. elegans GTG/GPHRs* were cloned into *S. cerevisiae* vectors for phenotypic and localisation studies. First, *S. cerevisiae* expression constructs were generated (section 7.2.2). The constructs in Table 7.1 were made using the Gateway system and confirmed using restriction analysis (Figure 7.4 A) and sequenced before transforming into BY4741, YHR078W and *zrc1*-Δ *cot1*-Δ (Table 7.2). To check whether these were successfully transformed into *S. cerevisiae*, colony PCR was carried out (Figure 7.4 B). Table 7.2 shows the strains created for complementation, localisation and potential purification studies.

As an initial test, a growth assay was set up comparing only the expression of At *GTG1* in BY4741 and YHR078W. This was to test whether At *GTG1* expression had any effect on YHR078W growth in 1 mM EDTA. BY4741 and YHR078W strains expressing At GTG1 were grown on solid SC-ura galactose media containing 0, 0.5 and 1 mM EDTA. However despite growth for 3 d, the growth of all strains was severely

reduced (Figure 7.8). The presence of glucose represses the expression of *GAL* genes (Timson, 2007). Therefore glucose cannot be used for heterologous gene expression under the *GAL1* promoter as the presence of glucose will only repress their expression. For this reason, the EDTA concentration was reduced such that in the next experiment SC-ura galactose media containing lower EDTA concentrations of 0.25 and 0.5 mM was used. In the presence of 0.25 mM EDTA, the strains that contained the transformed plasmid grew since they contain the selectable marker *URA3*. However, at 0.25 mM EDTA, there was no difference seen in growth between the BY4741 and YHR078W strains expressing the empty vector (Figure 7.9), suggesting that the use of galactose as the carbon source cannot elicit the slight reduced growth seen in Figure 7.7. Therefore, no conclusion concerning functional complementation can be made. This posed some difficulty in examining the effects of heterologous expression of the Arabidopsis and *C. elegans* GTG/GPHRs in *S. cerevisiae* since their expression is driven under the *GAL1* promoter which is repressed by glucose (Lohr et al., 1995). One possible solution would be to try and use raffinose, which does not repress or induce the *GAL1* promoter (Kundu et al., 2007, Menendez et al., 2007), and so results in a basal level of expression (Thompson et al., 2003). If raffinose is able to elicit a difference in growth of the YHR078W strain, functional complementation assays could be conducted using the *S. cerevisiae* expression vectors generated (section 7.2.2).

7.3.3 The Arabidopsis and *C. elegans* GTG/GPHRs localise intracellularly

The Arabidopsis, hamster and *Drosophila* GTG/GPHRs are reported to be localised at either the plasma membrane (Pandey et al., 2009), Golgi (Maeda et al., 2008, Jaffé et al., 2012, Charroux and Royet, 2014) and ER (Jaffé et al., 2012, Charroux and Royet, 2014). The *C. elegans* GPHRs did not appear to localise to the ER, but could be Golgi localised in *C. elegans* (section 6.3.5). In line with the reports so far (Maeda et al., 2008, Jaffé et al., 2012, Charroux and Royet, 2014), the Arabidopsis and *C. elegans* GTG/GPHRs localise to punctate foci and structures that resemble Golgi and ER localisation respectively, when expressed in *S. cerevisiae* (Figure 7.10 – Figure 7.12). Despite the lack of a Golgi colocalisation marker, the punctate foci seen in Figure 7.10 and Figure 7.11 are characteristic of Golgi localisation (Figure 7.1), as shown by the Golgi-localised proteins Anp1p (Wooding and Pelham, 1998), Bet3p (Du and Novick, 2001), Och1p (Schmidt et al., 1998) and Emp47p (Restrepo-Hartwig and Ahlquist, 1999). Golgi localisation in *S. cerevisiae* is very distinct to ER localisation or any other

organelles (Figure 7.1), as seen for ER-localised proteins Dpm1p (Restrepo-Hartwig and Ahlquist, 1999), BiP (Wooding and Pelham, 1998) and Kar2p (Schmidt et al., 1998). Additionally, the Arabidopsis and *C. elegans* GTG/GPHRs display fluorescence profiles that resemble ER protein localisation (Figure 7.11 and Figure 7.12). Therefore results suggest that the Arabidopsis and *C. elegans* GTG/GPHRs are localised intracellularly and this could be within the Golgi and/or ER network, which is in accordance with Maeda et al. (2008), Jaffé et al. (2012) and Charroux and Royet (2014), but not Pandey et al. (2009). There are some proteins that localise to more than one organelle, which could be the case for the GTG/GPHRs. For instance, the Svp26 membrane protein is both ER- and Golgi-localised (Noda et al., 2014). However further localisation studies using Golgi and ER markers would need to be conducted to fully conclude that these heterologous proteins are localised to the Golgi and/or ER. Particularly since Golgi-destined proteins passage through the ER in order to reach the Golgi and overexpression of the Golgi proteins could result in proteins aggregating in the ER lumen (Pfeffer and Rothman, 1987).

7.3.4 Further characterisation of Sc Gtgp is required

To further investigate the role of Sc Gtgp, extensive characterisation of YHR078W would need to be completed. Competitive relative fitness is the ability of a mutant strain to grow under conditions of direct competition with a reference e.g. WT strain (Cherry et al., 2012). For example, mutant and reference strains are grown individually, in rich media such as SC media for 24 h, subsequently diluted and grown in new media for a further 4 h. The testing and reference strains are then mixed at a 1:1 ratio, diluted again into fresh media such that a final cell concentration is 5×10^3 cells/mL. These then compete for growth over a specified time-period, for example 21 h represents about 12 generations of growth. The ratios of the two competitors are subsequently quantified at the intial and final time points, for example using a fluorescence-activated cell sorter (Chang and Leu, 2011).

Large-scale surveys have revealed that YHR078W is associated with increased competitive fitness when grown in glycerol (YPG) (Qian et al., 2012), ethanol (YPE) (Qian et al., 2012) or minimal medium (media that contains the minimum nutrients required for growth) (Breslow et al., 2008). But competitive fitness decreased when grown in SC media (Qian et al., 2012). The experiments done so far have not shown any

difference in growth between BY4741 and YHR078W and are classed as "vegetative growth assays" since they are grown individually and not grown in direct competition with one another as explained above (Cherry et al., 2012). Therefore competitive fitness experiments could be set up whereby YHR078W would need to grow under conditions of direct competition with BY4741 (Cherry et al., 2012).

Other stress assays conducted in large-scale screens included decreased resistance of YHR078W to (E)-1,1′-azobis(N,N-dimethylformamide) and aluminum (Tun et al., 2013). Stress resistance is related to cellular processes (Cherry et al., 2012) and since YHR078W stress resistance has been identified to change under different conditions, it indicates that Sc Gtgp may play an important role in a fundamental, cellular process, which is malfunctioning, leading to decreased stress resistance. Although Sc *GTG* was shown to be regulated by high osmolarity (O'Rourke and Herskowitz, 2004) under sorbitol or KCl stress, no growth differences were observed using these conditions in this study (Figure 7.5). YHR078W has been identified in a screen to exhibit abnormal protein transport as Kar2p retention is disrupted and instead Kar2p, which is an ER protein, is secreted (Čopič et al., 2009). This suggests Sc Gtgp is important in the secretory pathway. In the future, characterising the localisation and expression profiles of the Sc Gtgp would further our insight into the role of this protein. The expression vectors reporting localisation could also be used for expression analysis. The amount of GFP can also be quantified in live cells through flow cytometry and compared to a reference strain; this allows for expression level quantification (Niedenthal et al., 1996).

In addition the pAG426-*Cel-gphr-1::egfp*, pAG426-*Cel-gphr-2::egfp* and pYTV-*Cel-gphr-2* were successfully transformed into the *zrc1*-Δ *cot1*-Δ double mutant (Table 7.2). The *S. cerevisiae zrc1*-Δ *cot1*-Δ double mutant is hypersensitive to zinc. This defect was rescued by heterologous expression of *O. sativa* MTP1, a zinc transporter which is localised to the vacuole when expressed in *S. cerevisiae* (Menguer et al., 2013). Since YHR078W was identified to be sensitive to low zinc conditions (North et al., 2012); the effect of expressing the *Cel*-GPHRs in this double mutant could be explored.

The Arabidopsis and *C. elegans* GTG/GPHRs have also been inserted into pYTV (Table 7.1), which can be used for expression, protein purification and subsequent structural studies such as X-ray crystallography. Structural characterisation of the GTG/GPHRs would confirm the membrane topology of these proteins and reveal any

structural features important for its function. Determining their structure will provide an insight into its mechanism of action and whether it resembles any pore-forming or receptor features that could provide evidence for channel, GPCR or G protein activity. These would all be interesting studies to conduct for further characterising the GTG/GPHR family of membrane proteins.

7.3.5 Other proteins identified through large-scale screens share similar phenotypes to YHR078W

There are many *S. cerevisiae* proteins that share similar phenotypes observed through large-scale screens with YHR078W. Some of these are Golgi or ER localised proteins, for instance Per1p is an ER protein required for GPI-phospholipase A2 activity (Fujita et al., 2006). Hoc1p is an α-1,6-mannosyltransferase; a subunit of a Golgi localised complex involved in cell wall mannan biosynthesis (Hashimoto and Yoda, 1997, Neiman et al., 1997, Jungmann et al., 1999). *AGL3* encodes for dolichol-P-Man-dependent α(1-3)-mannosyltransferase. Agl3p is an ER-transmembrane protein involved in *N*-glycosylation of proteins. It catalyses the addition of the sixth mannose moiety to the dolichol-linked oligosaccharide in the ER lumen (Aebi et al., 1996, Sharma Chandra et al., 2001). However, not all of these proteins identified in the surveys are Golgi or ER proteins. Cex1p is a component of the nuclear aminoacylation-dependent tRNA export pathway. Unlike the other proteins mentioned, this protein is cytoplasmic and interacts with nuclear pore component, Nup116p (McGuire and Mangroo, 2007, McGuire and Mangroo, 2012). *CDH1* is a cell-cycle regulated gene encoding an activator of anaphase-promoting complex/cyclosome (APC/C) (Visintin et al., 1997). Therefore these genes identified from large-scale screens that exhibit similar phenotypes to YHR078W suggest that there is a possibility that Sc Gtgp is involved in intracellular biological processes, which could involve the secretory pathway or cytoplasm.

7.3.6 Further characterisation of the GTG/GPHRs is required in *S. cerevisiae*

The results so far indicate that more phenotypic analysis is required for YHR078W. For example, using a wider pH range using buffering agents and amending the high osmolarity media to try and identify a phenotype. Certainly the large-scale screens provide a wider scope for the assays that could be conducted such as competitive fitness assays using different media including SC, minimal, glycerol and

ethanol media and growth in zinc-limiting conditions. The Arabidopsis and *C. elegans* GTG/GPHRs could localise to the same organelles as they appear to display similar intracellular, fluorescence profiles (Figure 7.10 – Figure 7.12). Undoubtedly, further localisation studies are needed to confirm the subcellular localisation of these proteins. These could include using known subcellular markers to examine any colocalisation and this would confirm whether the Arabidopsis and *C. elegans* GTG/GPHRs localise to the same organelles. Furthermore, this could be accomplished by coexpressing the Arabidopsis and *C. elegans* GTG/GPHRs with different fluorescent reporter tags in the same cell. Membrane fractionation could be used to separate subcellular membranes from different organelles for immunofluorescent microscopy, which could utilise the pYTV constructs (Table 7.2). Single-molecule fluorescence spectroscopy can be used to track GFP-tagged GTG/GPHRs (Weiss, 1999, Lippincott-Schwartz et al., 2001). The kinetic movement could reveal characteristic dynamics specific to certain organelles. All further studies would help determine the function of these proteins and whether they support any of the those reported so far.

CHAPTER 8

General Discussion

8.1.1 Plant and animal GTG/GPHRs have a fundamental function affecting fertility, growth and development

The GTG/GPHRs are a novel set of membrane proteins that have been studied in several organisms (Maeda et al., 2008, Pandey et al., 2009, Jaffé et al., 2012, Tarutani et al., 2012, Alvarez et al., 2013, Kharenko et al., 2013, Charroux and Royet, 2014). As mentioned in section 1.2, there are several roles proposed for these proteins, especially in plants (Pandey et al., 2009, Jaffé et al., 2012, Alvarez et al., 2013). In Arabidopsis, distinct single mutants were isolated and displayed no apparent differences to WT (Pandey et al., 2009, Jaffé et al., 2012). These single mutants were used to generate different double At *GTG* null mutants (Table 4.1) and their characterisation has presented various roles for these proteins (Pandey et al., 2009, Jaffé et al., 2012, Alvarez et al., 2013). The Arabidopsis *gtg1 gtg2* (*gtg1-2 gtg2-2* and *gtg1-3 gtg2-3*) mutants isolated by Jaffé et al. (2012) show that these are Golgi and ER proteins required for light-dependent growth and fertility in Arabidopsis. On the other hand, the *gtg1-1 gtg2-1* mutant isolated by Pandey et al. (2009) suggests that these proteins are ABA receptors, acting at the plasma membrane through a mechanism engaging hallmarks of both GPCRs and G proteins. The *gtg1-1 gtg2-1* mutant was hyposensitive to ABA in root and hypocotyl inhibition assays as well as flowering earlier than WT (Pandey et al., 2009). Additional phytohormone assays suggest that the At GTGs are involved in plant hormone cross-talk (Alvarez et al., 2013). Interestingly, the *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* mutants did not exhibit ABA-hyposensitivity (Jaffé et al., 2012) exhibited by *gtg1-1 gtg2-1* (Pandey et al., 2009). However, other interesting phenotypes were observed including shorter roots and hypocotyls, reduced fresh weight as well as distended hypocotyl cells. These phenotypes can be rescued by the expression of either At GTG1 or At GTG2 indicating that these defects are caused by the loss of both At *GTG* genes and that these genes function redundantly (Jaffé et al., 2012).

These reports suggest that these proteins may be implicated in diverse cellular pathways, hence further characterisation is needed to verify their function. In this study, the three distinct Arabidopsis *gtg1 gtg2* double mutants were compared under the same

conditions to confirm the phenotypes previously described and to extend the analysis. Results indicated that the defects exhibited by *gtg1-2 gtg2-2* and *gtg1-3 gtg2-3* are also observed in *gtg1-1 gtg2-1*; these include a smaller rosette phenotype as well as reduced root, hypocotyl and fresh weight growth, and reduced seed number and silique length (section 4.2.2). Seed germination assays conducted in our laboratory indicates that under our conditions, the three *gtg1 gtg2* double mutants do not exhibit ABA hyposensitivity (K. Peaston, M. J. Terry and L. E. Williams, unpublished results). Furthermore, BL hyposensitivity was observed by the *gtg1-1 gtg2-1* mutant in root and hypocotyl assays conducted in the dark (Alvarez et al., 2013). Under low light conditions used in this project, the three *gtg1 gtg2* mutants did not exhibit BL-hyposensitivty (section 4.2.5). Consistent with Jaffé et al. (2012), these results reinforce a role for the At GTGs in growth, development and fertility, but cannot confirm that there is a definite role in phytohormone signalling. Certainly, to investigate their role in hormone signalling further testing would be required since experiments carried out so far are under various conditions (Pandey et al., 2009, Jaffé et al., 2012, Alvarez et al., 2013).

While the GTG/GPHRs have been studied in plants (Pandey et al., 2009, Jaffé et al., 2012, Alvarez et al., 2013), until very recently there were no reports regarding the use of a whole animal *GTG/GPHR* KO model. For this reason, one of the aims of this project was to isolate the first animal *GPHR* KO model. However, *Drosophila GPHR* KO mutants (Dm $GPHR^{P91L}$ and Dm $GPHR^{LL03674}$) have recently been isolated and indicate that the GTG/GPHRs are involved in systemic growth of flies (Charroux and Royet, 2014). Here, a whole organism double *gphr* KO model in the animal *C. elegans* has been characterised indicating comparable defects to the Dm *GPHR* null mutants. Similar to Arabidopsis single mutants (Pandey et al., 2009, Jaffé et al., 2012), the single *gphr* mutants obtained did not exhibit any obvious differences compared to the WT. However, a double *C. elegans gphr* mutant was generated and in doing so a number of phenotypes were revealed. Like the Arabidopsis *gtg1 gtg2* mutants, the *C. elegans* double *gphr* mutant exhibited defects in growth, development and fertility; these manifested as abnormal egg-laying, egg development, hatching, pharyngeal pumping, as well as L1 larval arrest (section 6.3.2). Similar to the *C. elegans* double *gphr* mutant, many of the Dm *GPHR* KO mutants exhibited larval arrest ultimately leading to their death (Charroux and Royet, 2014). However, unlike the *Drosophila GPHR* KO mutants,

no *C. elegans* double *gphr* mutant escaped this and developed into an adult. The Dm *GPHR* KO mutants that escaped larval arrest exhibited dwarfism (Figure 1.2 C). This was characterised by reduced weight and smaller wings, as well as significantly smaller wing cell size and reduced number of wing cells. Clearly, removing the *GPHR* gene in animals causes significant defects in growth and development.

The in-depth study of both plant and animal *GTG/GPHR* KO models shows that these proteins are associated with growth, development and fertility and it is likely that this could be due to a conserved cellular function. Confirming the nature of this is therefore central to understanding their role in eukaryotes and should be a priority in future research.

8.1.2 The GTG/GPHRs are found in all eukaryotes and have a conserved structure

In silico analysis indicates that the GTG/GPHR family of membrane proteins are found exclusively in eukaryotes and are predicted to have a conserved protein architecture (sections 3.2.2 and 3.2.3). So far, the GTG/GPHRs of Arabidopsis (Pandey et al., 2009, Jaffé et al., 2012, Alvarez et al., 2013), a CHO cell line (Maeda et al., 2008), human (Maeda et al., 2008), mouse (Tarutani et al., 2012), *O. sativa* (CHAPTER 4), *C. elegans* (CHAPTER 6), *S. cerevisiae* (CHAPTER 7) and *Drosophila* (Charroux and Royet, 2014) have been studied. *In silico* analyses were conducted on these GTG/GPHRs showing that these proteins are predicted to contain nine TM regions (consensus prediction) and a protein architecture containing DUF3735 and ABA-GPCR domains (section 3.2.3). However since these are only predictions, experimental data would be needed to verify this. Membrane topology can be elucidated using various biochemical approaches. The regions of protein that lie on opposing sides of the membrane are accessible to different reagents. SDM can be employed to introduce targets at specific positions to determine membrane protein topology experimentally. These tags could include cysteine residues, *N*-glycosylation sites or antibody epitopes. Scanning for *N*-glycosylation sites is a popular method for determining topology of eukaryotic proteins, such as the human Band 3 erythrocyte anion channel (Popov et al., 1997) and the *E. coli* ArsB protein that is the membrane subunit of an anion-translocation ATPase (Wu et al., 1992). Cysteine scanning is also used for topographical testing; for example, membrane topology was determined for both Anion

Exchanger 1 (SLC4A1) (Barneaud-Rocca et al., 2013) and the first TM segment of cystic fibrosis transmembrane conductance regulator (CFTR) (Gao et al., 2013). Furthermore this also revealed roles of the first TM domain in gating and selectivity for CFTR (Gao et al., 2013). Reporters can also be attached such as the enzymes alkaline phosphatase (PhoA) or β-galactosidase (LacZ), which are only active in either the periplasm or cytoplasm, respectively (van Geest and Lolkema, 2000). The enzyme activities of PhoA and LacZ can be determined using either XP (5-bromo-4-chloro-3-indolylphosphate) or X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) respectively; in spectroscopic enzyme assays or by immunoblotting (van Geest and Lolkema, 2000). The GTG/GPHRs have been described as GPCR-type G proteins but typically with 7 – 11 TM segments predicted (section 3.2.4), which is not the classic 7TM GPCR structure (Ovchinnikov, 1982, Hargrave et al., 1983, Nathans and Hogness, 1983, Dixon et al., 1986, Kubo et al., 1986). Determining the GTG/GPHR protein topology would therefore be important to clarify the number of membrane-spanning segments and whether this coincides with the 7TM structure of GPCRs.

There are also several conserved residues identified through multiple sequence alignments; these are found in TM3, DUF3735 and the ABA-GPCR domain (Table 3.3). Interestingly, the conserved proline residue found in the predicted TM3 of the GTG/GPHRs has recently been shown to be important for function. When this proline was mutated to a leucine, GPHR function in *Drosophila* was abolished resulting in systemic growth defects (Charroux and Royet, 2014). In this project, the conserved glycine residue of the LS**G** motif in DUF3735 (Table 3.3) was also shown to be functionally relevant. The G166L mutation in At GTG1 was adequate to disrupt its function as shown by its inability to restore Arabidopsis *gtg1 gtg2* double mutant phenotypes (section 4.2.8). Therefore it is possible that the other identified conserved residues, such as the tyrosine in the ABA-GPCR domain, are integral to GTG/GPHR function. The additional residues not yet examined can be targeted in future mutagenesis studies to explore structure/function relationships; this would help identify functionally important residues, especially when the structure of these proteins is yet to be characterised.

8.1.3 Do the GTG/GPHRs have a conserved function?

With a highly conserved sequence and predicted architecture throughout eukaryotes, it was important to determine whether there is a conserved function for this family of proteins. For example, the voltage-gated anion channel (VDAC) family, comprising one α-helix and 13 β-strands, are conserved in eukaryotes (Colombini, 1979), where they play an essential role in apoptosis (programmed cell death) pathways (Tsujimoto and Shimizu, 2000).

Since the Arabidopsis *gtg1 gtg2* mutants have well-characterised phenotypes that can be restored by expressing At GTG1 or At GTG2, this formed the basis for exploring conservation of function. Functional complementation assays were set up to identify whether heterologous expression of GTG/GPHRs from another organism could rescue the Arabidopsis double mutant defects. The root, hypocotyl and fresh weight defects were all rescued by the expression of *Cel*-GPHR-1 (Figure 5.8 and Figure 5.10), suggesting there is conservation of function across kingdoms.

Interestingly, *Cel*-GPHR-2 does not rescue any of the *gtg1 gtg2* seedling growth defects observed, unlike *Cel*-GPHR-1 or either of the At GTGs (discussed above). However, as mentioned in section 5.3.3, protein expression data would be required to corroborate that the lack of rescue by *Cel*-GPHR-2 is not due to low/no expression of the protein. However, preliminary data of a T_2 plant line expressing GFP-tagged *Cel*-GPHR-2 indicate that in there is expression of *Cel*-GPHR-2 (Figure 5.17). However this would need to be confirmed using T_3 plant lines as these were used for phenotypic assays. Sequence analysis shows that *Cel*-GPHR-2 is no less different overall to At GTG1 or At GTG2 than its homologue *Cel*-GPHR-1 at the amino acid level (Table 3.2). Additional sequence analysis was also conducted comparing the conserved predicted intracellular region, DUF3735 and ABA-GPCR domains (Table 5.5), which are considered to be important for function. These data indicate that both *Cel*-GPHR-1 and *Cel*-GPHR-2 share similar identity and similarity to both At GTGs. For instance the ABA-GPCR domain of both *Cel*-GPHR-1 and *Cel*-GPHR-2 are 43 % identical and 69 % similar to At GTG1. The DUF3735 domain of *Cel*-GPHR-2 is 52 % identical and 78 % similar to At GTG1, however surprisingly this is slightly lower when comparing *Cel*-GPHR-1 and At GTG1 (47 % identity and 74 % similarity). On the other hand, the intracellular portion of *Cel*-GPHR-2 has slightly lower identity and similarity to At GTG1 (19 % and 38 % respectively) than *Cel*-GPHR-1, which shares 22 % identity and

38 % similarity to At GTG1 in comparison. Furthermore, the large intracellular region, DUF3735 and ABA-GPCR domains are more similar in the At GTG than the *Cel*-GPHR protein sequences (Table 5.5). For instance, At GTG1 and At GTG2 share 90 % identity and 97 % similarity in the DUF3735 domain. While *Cel-GPHR-1* and *Cel*-GPHR-2 only share 76 % identity and 85 % similarity in this domain. Additional sequence analysis indicates that there are residues shared by At GTG1, At GTG2 and *Cel*-GPHR-1 but not *Cel*-GPHR-2 (Figure 8.1). For example there are a few arginine residues (positively charged) that are found in At GTG1 (e.g. R103 and R108), At GTG2 (e.g. R103 and R08) and *Cel*-GPHR-1 (e.g. R105 and R109) but in *Cel*-GPHR-2 these are glutamine (e.g. Q100 and Q106), which is a polar and uncharged amino acid (Figure 8.1). These differences in residues could be responsible for the similar function of At GTGs and *Cel*-GPHR-1. Therefore the *C. elegans* GPHR homologues could play different roles in the cell where *Cel*-GPHR-1 shares a similar function to the orthologue At GTGs. However, the protein levels of *Cel*-GPHR-2 were not measured and could be the reason for the lack of rescue. There are some genes that are more susceptible to gene silencing when expressed in different organsims due to for example difference in codon usage. Transgene expression can be improved by applying design rules such as removing rare codons and RNA instability sequences. This method has been tested on the silencing susceptible beetle *luc* and bacterial *ims* genes expressed in sugarcane. By using these design rules, a higher level of expression was obtained (Jackson et al., 2014). If after future tests indicate that *Cel*-GPHR-2 protein levels were low in Arabidopsis *gtg1 gtg2* mutants, further testing could involve following these design rules for expressing *Cel-gphr-2* in Arabidopsis.

The seedling growth defects of *gtg1 gtg2* mutants were rescued by either At GTG1, At GTG2 or *Cel*-GPHR-1 expression under the 35S promoter (sections 5.2.3 and 5.2.5). However, no rescue of the short silique phenotype was observed, although this was not measured directly. As mentioned in section 4.3.2, this could be due to the fact that *P35S* does not drive expression in the siliques (Ge et al., 2008), which could explain the lack of rescue. In addition the short silique phenotype was hypothesised to be due to reduced pollen tube extension (Jaffé et al., 2012). Microarray data indicated that the At *GTG* genes are expressed in the siliques during their development but are not highly expressed there once siliques reach maturity, as well as exhibiting low expression levels in pollen (Figure 3.16 and Figure 3.18). To investigate this

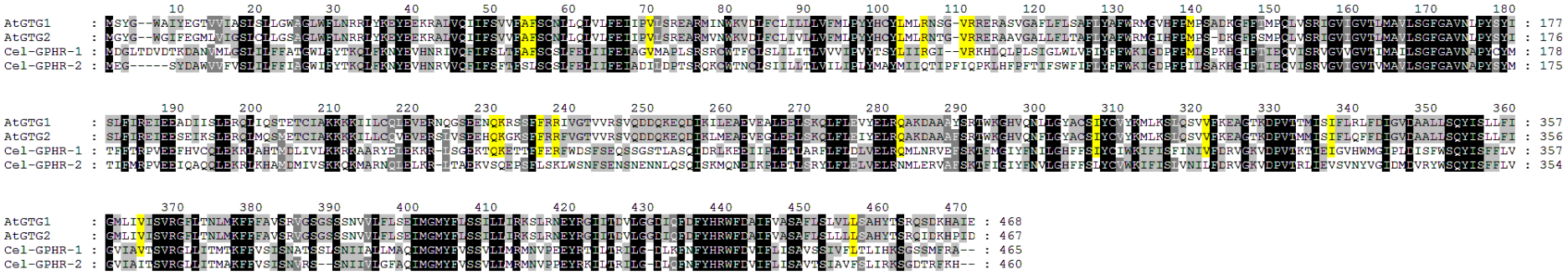


Figure 8.1 Protein sequence alignment comparing *Cel*-GPHR-2 to At GTG1, At GTG2 and *Cel*-GPHR-1

Multiple alignment comparing *Cel*-GPHR-2 to the At GTGs and *Cel*-GPHR-1, with those residues found in At GTG1, At GTG2 and *Cel*-GPHR-1 but not in *Cel*-GPHR-2 are highlighted in yellow. Sequences were aligned using the ClustalW algorithm (Thompson et al., 1994) and presented using GeneDoc (Nicholas and Nicholas Jr, 1997). For the sequences aligned here: black = conserved residues, dark grey = conserved in 3 of 4 organisms, light grey = conserved in 2 of 4 organisms, yellow = conserved in At GTG1, At GTG2, *Cel*-GPHR-1.

experimentally, the Arabidopsis transgenic lines expressing *PAt GTG1:At GTG1:GFP* or *PAt GTG2:At GTG2:GFP* could be used for following *GTG* expression in silique development and pollen grains. These data could be used to explain why there is a short silique phenotype in Arabidopsis *gtg1 gtg2* mutants.

The GTG/GPHRs have been shown to regulate Golgi and ER organelles in two independent animal studies (Maeda et al., 2008, Charroux and Royet, 2014). Maeda et al. (2008) showed that the loss of the hamster *GPHR* in a CHO cell line caused Golgi disorganisation and impairment of Golgi acidification. Similarly in *Drosophila*, the loss of *GPHR* also caused ER and Golgi disorganisation (Charroux and Royet, 2014). Furthermore, in these *Drosophila GPHR* KO models, several growth defects were observed and were restored by expressing either *Drosophila*, human or hamster GPHR proteins, indicating they have a similar role (Charroux and Royet, 2014). In this project, expression of the animal *Cel*-GPHR-1 can restore defective plant Arabidopsis phenotypes implying that plant GTG/GPHRs have the same function as the *C. elegans* orthologue, although *Cel*-GPHR-2 may not share this function. Together these data strongly suggest that there is a conserved function for the GTG/GPHRs.

8.1.4 The GTG/GPHRs are localised to the Golgi/ER

One of the main priorities was to corroborate the localisation of the GTG/GPHRs since there are different accounts. The GTG/GPHRs have been reported to be localised at the plasma membrane in Arabidopsis (Pandey et al., 2009), or at the Golgi in a CHO cell line (Maeda et al., 2008), Arabidopsis (Jaffé et al., 2012) and *Drosophila* (Charroux and Royet, 2014), as well as at the ER in Arabidopsis (Jaffé et al., 2012) and *Drosophila* (Charroux and Royet, 2014). Therefore, the localisation of this family of proteins is yet to be clarified. Here the localisation of both the Arabidopsis and *C. elegans* GTG/GPHRs was investigated using three model organisms.

Localisation results signifies that under its own promoter At GTG2 localised to punctate intracellular structures in Arabidopsis; this protein was also shown to be functional (Figure 4.15 and Figure 4.16). Since At GTGs function redundantly and At GTG1 was shown to be localised to the Golgi and ER network (Jaffé et al., 2012), it is likely that At GTG2 is also Golgi/ER localised. *Cel*-GPHR-1 functionally complemented the Arabidopsis *gtg1 gtg2* mutant and similar to At GTG2, also localised to punctate structures resembling Golgi (Figure 5.16). *Cel*-GPHR-2 also localised to

punctate structures in the Arabidopsis double mutants (Figure 5.17), however neither *Cel-gphr-2* nor *Cel-gphr-2::gfp* expression rescued *gtg1 gtg2* defects. Further supporting that Arabidopsis and *C. elegans* GTG/GPHRs are Golgi and/or ER localised proteins, both localised to punctate intracellular structures that resemble Golgi and ER localisation in *S. cerevisiae* (Figure 7.10 – Figure 7.12). In *C. elegans*, *Cel*-GPHR-1 also localised intracellularly when expressed in the pharynx; the punctate foci observed is reminiscent of Golgi localisation (Figure 6.24 and Figure 6.25). A summary of all the localisation data obtained in this study is given in Table 8.1. All the results support that the GTG/GPHRs in animals and plants are Golgi and ER localised proteins, which is in accordance with the previous reports of the hamster (Maeda et al., 2008), Arabidopsis (Jaffé et al., 2012) and *Drosophila* (Charroux and Royet, 2014) GTG/GPHR proteins, but not with the report of Pandey et al. (2009). Moreover, the localisation data supports that these proteins have a fundamental intracellular role, but what is this mechanism?

Table 8.1 Summary of the localisation results obtained in this study

Sequence	Expressed in organism	Localisation
At GTG1	Arabidopsis	Punctate structures – possibly Golgi and ER
At GTG1	*S. cerevisiae*	Punctate structures – possibly Golgi and ER
At GTG2	Arabidopsis	Punctate structures – possibly Golgi and ER
At GTG2	*S. cerevisiae*	Punctate structures – possibly Golgi and ER
Cel-GPHR-1	Arabidopsis	Punctate structures – possibly Golgi and ER
Cel-GPHR-1	*C. elegans*	Punctate structures – possibly Golgi
Cel-GPHR-1	*S. cerevisiae*	Punctate structures – possibly Golgi and ER
Cel-GPHR-2	Arabidopsis	Punctate structures – possibly Golgi and ER
Cel-GPHR-2	*S. cerevisiae*	Punctate structures – possibly Golgi and ER

8.1.5 Are GTG/GPHRs likely to be the proposed Golgi pH regulator?

There are different roles proposed for the GTG/GPHRs: Golgi pH regulators with anion channel activity (Maeda et al., 2008) or plasma membrane ABA receptors (Jaffé et al., 2012) (Figure 8.2). In this project, many of the defects observed by various *GTG/GPHR* KO models can be explained by Golgi- or secretory pathway-related defects. Moreover, the localisation results (section 8.1.4) and other reports of Golgi and ER localisation of the GTG/GPHRs further validates the hypothesis that the GTG/GPHRs are Golgi-localised proteins (Maeda et al., 2008, Jaffé et al., 2012, Charroux and Royet, 2014). Therefore, the data presented here supports the concept that these proteins are Golgi pH regulators rather than the alternative ABA receptor role; the reasons for this will now be discussed.

Many of the *gtg1 gtg2* mutant phenotypes observed are related to growth deficiency. There are instances where Golgi dysfunction results in reduced growth in Arabidopsis. For example, the *det3* mutant lacking V-ATPase activity in trans-Golgi had reduced hypocotyl growth (Brüx et al., 2008). However, the *det3* seedlings displayed a constitutive photomorphogenic phenotype in the dark, which was not observed in the *gtg1 gtg2* mutants. Furthermore microarray data indicates that *det3* mutants exhibited downregulation of both At *GTG* genes (Figure 3.19). If the At GTGs are the alleged counterion channels, this decline in At *GTG* expression could be to compensate for the loss of V-ATPase activity in the *det3* mutant.

Golgi impairment could result in altered cell wall synthesis and the subsequent hypocotyl growth defects observed. Without the tension and support of a proper cell wall, cellular development may be compromised, leading to growth defects and deformed, ballooning cells. Cells walls comprise carbohydrates such as cellulose, hemicelluloses including xyloglucan, and pectins. Many of these cell wall components are synthesised in the Golgi, such as xyloglucan and pectin (Zhang and Staehelin, 1992). Breakdown of pH homeostasis leads to Golgi dysfunction and could therefore result in malsynthesis of cell wall constituents. *PROCUSTE1* (*PRC1*) encodes a cellulose synthase that is vital for cell elongation in roots and dark-grown hypocotyls (Fagard et al., 2000). It functions to assemble the expanding cell wall as shown by mutant hypocotyls associated with deformed hypocotyl cells (Desnos et al., 1996). In the Arabidopsis *pcr1-1* mutant, roots and hypocotyls exhibited an irregular hypocotyl surface and cross sections displayed irregular cellular structures characterised by

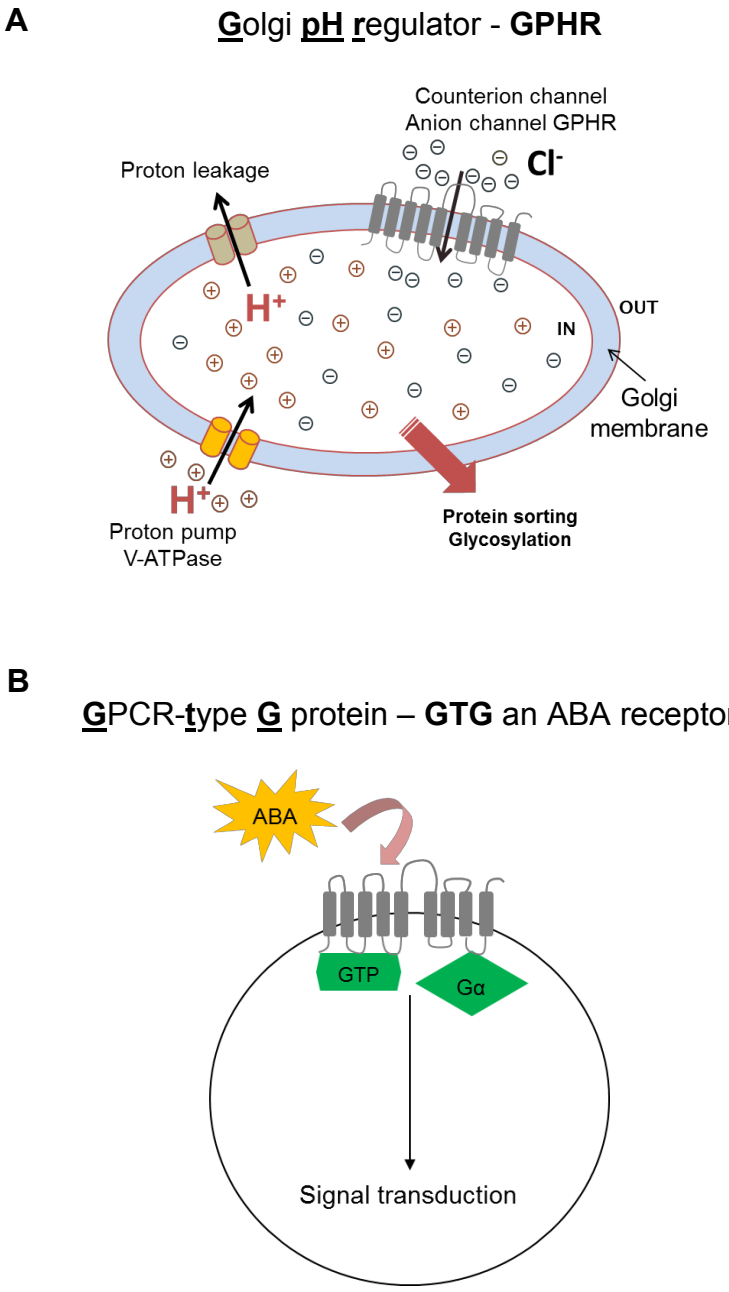


Figure 8.2 The proposed roles for the GTG/GPHR family of membrane proteins

Diagrams to illustrate the potential roles GTG/GPHRs may play in eukaryotes. **A**. Golgi pH regulator at the Golgi membrane. To regulate luminal pH, the GPHR acts as an anion channel to counteract the proton influx. **B**. GPCR-type G protein at the plasma membrane. The GTG acts as an ABA receptor resembling hallmarks of both GPCRs (G protein binding; Gα subunit shown) and G proteins (GTPase activity; GTP shown).

swollen and compressed cells (Desnos et al., 1996); this is similar to the hypocotyl cross sections of the Arabidopsis *gtg1 gtg2* mutants showing a comparable phenotype (Jaffé et al., 2012). This sets a precedent for the *gtg1 gtg2* mutants having defects in Golgi function leading to cell wall defects such as synthesis. This irregular cellular phenotype of the *prc1-1* mutant was also observed in the roots (Desnos et al., 1996). However, the Arabidopsis *gtg1-3 gtg2-3* mutant roots did not appear to exhibit a deformed cellular structure. However, this defect could be more obvious in the *gtg1-2 gtg2-2* mutant since many of the double mutant phenotypes are exacerbated in *gtg1-2 gtg2-2* compared to *gtg1-3 gtg2-3* (Jaffé et al., 2012). Microarray data also identified many genes affected that were involved in cell wall synthesis and modification, such as xyloglucan endotransglycosylase-6 (Jaffé et al., 2012). Growth defects may arise from inadequate cell division, such as that seen in *scarecrow* mutants which exhibited short root lengths due to disruption in asymmetric cell division (Di Laurenzio et al., 1996) or cell elongation, as seen in *prc1-1* mutants (Desnos et al., 1996). In the future this could be investigated in the roots and hypocotyls of the *gtg1 gtg2* mutants.

BRs are involved in both cell expansion and division (Clouse, 1996, Li et al., 1996, Azpiroz et al., 1998, Cheon et al., 2010). Since the application of BL rescued the hypocotyl length defect observed in *gtg1 gtg2* mutants, this supports the hypothesis that the root and hypocotyl growth defects are likely to due to a loss of cell expansion or cell division. There are Golgi processes such as *S*-acylation, which are implicated in cell expansion, division, vascular patterning as well as fertility in Arabidopsis. This was shown by *S. cerevisiae* studies and an Arabidopsis mutant of the *S*-acyl transferase PAT family, At *PAT10*. The At *pat10* mutants exhibited similar phenotypes to the Arabidopsis *gtg1 gtg2* mutants including reduced rosette sizes and smaller siliques; although these abnormal defects were stronger in *pat10* compared to *gtg1 gtg2* mutants (Qi et al., 2013). Therefore cell expansion/cell division defects could explain the Arabidopsis *gtg1 gtg2* mutant phenotypes including reduced root, hypocotyl, fresh weight and rosette growth as well as those related to fertility.

The defects in fertility including reduced seed yield, seed number per silique and silique length are thought to be caused by reduced pollen tube extension in the *gtg1 gtg2* mutants (Jaffé et al., 2012). This defect could be related to ER defects resulting from pH and cation imbalance as seen in Arabidopsis *chx21 chx23* mutants. These double mutants exhibited defects in pollen tube extension caused by dysfunctional ER; correct

pH and cation balance in the ER is involved in shifting the axis of polarity to grow in the direction of the ovule (Lu et al., 2011). In addition, Arabidopsis *VHA-A* encodes for the catalytic subunit of V-ATPase. The At *vha-a* mutant exhibits Golgi disorganisation and results in complete male and partial female gametophytic lethality due to pollen maturation dysfunction (Dettmer et al., 2005). Therefore ER and Golgi pH regulation is integral for normal pollen tube extension, which could explain the fertility defects observed in the *gtg1 gtg2* mutants. Therefore many of the Arabidopsis *gtg1 gtg2* phenotypes can be explained by Golgi/ER malfunctioning and this is summarised in Figure 8.3.

The plant GTG/GPHRs have so far only been studied in the dicot Arabidopsis. Here we initiated analysis of monocot GTG/GPHRs using *O. sativa*. A homozygous Os *gtg* mutant has been isolated which can be used for future studies to see if there are any parallel functions between monocot and dicot GTG/GPHRs. Preliminary observations show that the Os *gtg* mutant resembled similar defects such as reduced size (Figure 4.25) and perhaps a reduced seed yield compared to WT. However, this would need to be conducted with more plants as observations were made only using the one homozygous plant isolated. It is possible that similar defects to the Arabidopsis *gtg1 gtg2* mutants will be observed in the Os *gtg* mutant, since it has been shown that there is a conserved role for these proteins. Obviously, much research will need to be conducted in order to draw any firm conclusions regarding monocot GTG/GPHRs and their functions, and the new Os *gtg* mutant model will provide the possibility for this.

In *C. elegans*, the double *gphr* mutant phenotypes indicate that the GPHRs play a fundamental role in the cell. A defect in egg-laying behaviour (Figure 6.14) may be due to compromised egg-laying machinery, such as neuropeptide synthesis dysfunction (Kass et al., 2001, Jacob and Kaplan, 2003). Interestingly, a voltage-gated anion channel, CLC-2, encoded by *clh-3* modulates HSN motor neuron excitability to regulate egg-laying in *C. elegans* (Branicky et al., 2014). Therefore anion channel activity is central to *C. elegans* egg-laying, suggesting that GTG/GPHRs may be implicated in similar pathways such as anion channel activity. Defects in embryonic development (Figure 6.16) may be explained by failure of the terminal phase of cytokinesis, which is due to Golgi malfunction (Wang et al., 2005). Problems with hatching (Figure 6.15) could be inability to produce and secrete chitinase, an enzyme used to digest the egg shell in order for eclosion to occur (Zhang et al., 2005). The double *gphr* mutants

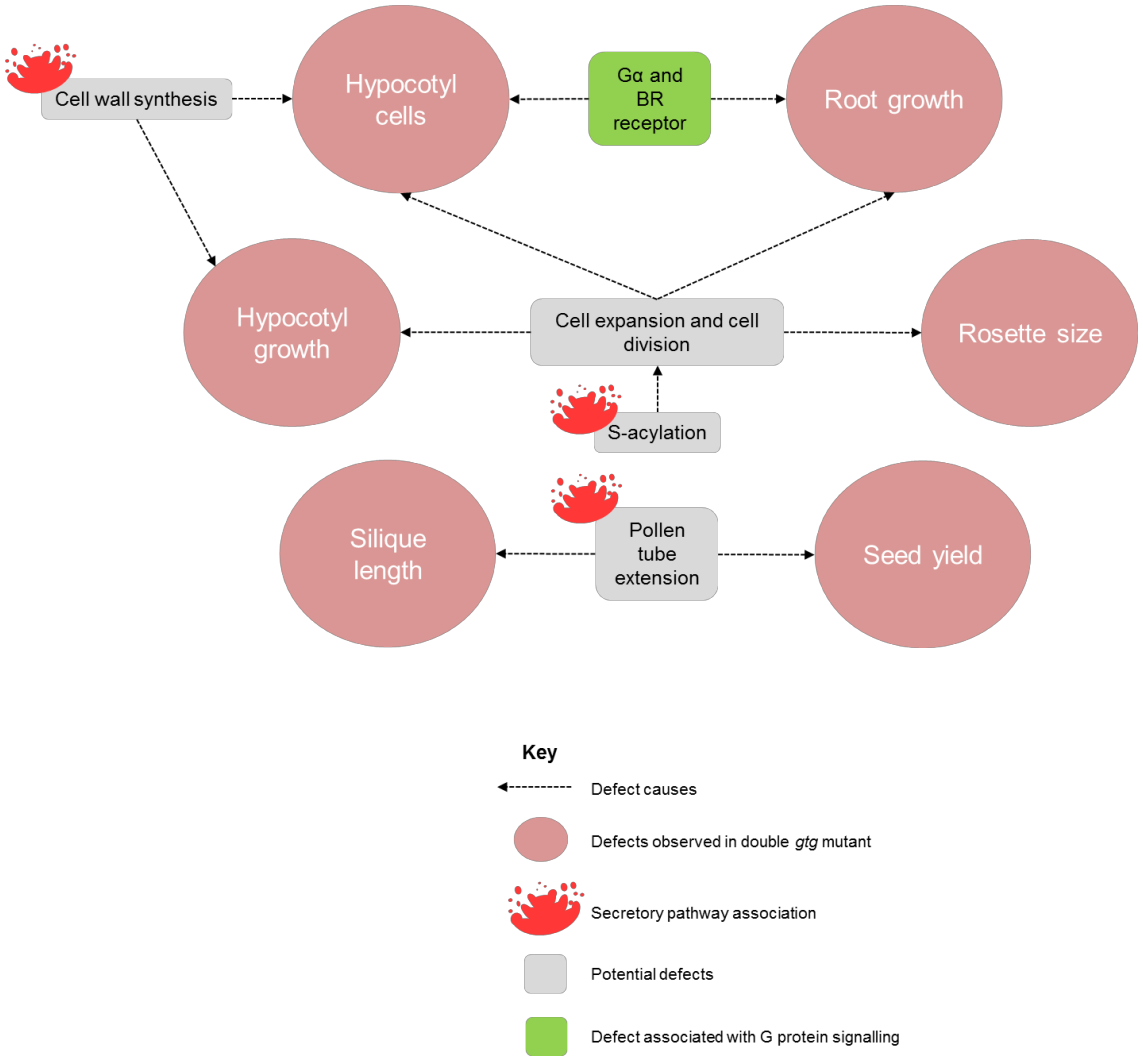


Figure 8.3 Arabidopsis *gtg1 gtg2* mutants exhibit growth, development and fertility defects

Schematic showing the defects (pink) observed by the Arabidopsis double *gtg1 gtg2* mutants, with the potential defects (grey) that could cause these phenotypes. For example, a loss of *S*-acylation causes cell expansion and cell division to malfunction leading to defects in root and hypocotyl growth as well as smaller rosettes and deformed cell growth in hypocotyls and roots.

exhibit L1 arrest (Figure 6.17) which was hypothesised to be due to a lack of feeding and consistent with this, these double mutants showed little to no pharyngeal pumping (Figure 6.18), which is indicative of a reduced feeding behaviour. However, there are some cases where L1 arrest is not attributed solely to a lack of feeding and rather is due to a defect in the insulin/IGF signalling pathway, as seen in *daf-2* mutants (Baugh and Sternberg, 2006).Therefore, there may be a defect in the insulin/IGF signalling pathway in the double *gphr* mutants, and the lack of pharyngeal pumping may contribute to L1 arrest. However, the reduced larval growth rate and partial developmental arrest exhibited by the *Drosophila GPHR* mutants were not attributed to either a feeding defect or insulin/IGF signalling. This was shown by the unaffected expression levels of insulin/IGF signalling transcriptional targets and similar phalloidin staining of wing imaginal discs to WT (Charroux and Royet, 2014). Inability to feed was excluded as a reason for the larval arrest and slow growth through analysing metabolic markers and quantifying mouth hook contractions (Charroux and Royet, 2014). The double *gphr* mutants exhibit abnormal growth, development and fertility phenotypes, which are similar to those systemic growth defects observed in the *Drosophila GPHR* KO models (Charroux and Royet, 2014). Therefore this data is in accordance with those indicating the importance of GTG/GPHR in growth and development.

In *S. cerevisiae*, further characterisation of the Sc Gtgp is required (section 7.3.4). This could include testing a wider pH range using buffering agents and amending the high osmolarity media to try and identify a phenotype for the *gtg*-Δ mutant strain. There are also large-scale screens conducted suggesting the *gtg*-Δ mutant strain exhibits various phenotypes under specific conditions. These provide a wider scope for the assays that could be conducted such as competitive fitness assays using different media including SC, minimal, glycerol and ethanol media and growth in zinc-limiting conditions (North et al., 2012). One of the phenotypes identified from these large-scale screens conducted suggest the *gtg*-Δ mutant strain has abnormal protein transport (Čopič et al., 2009). Furthermore under certain conditions there are also many Golgi and ER protein mutants also sharing similar phenotypes (section 7.3.5). For instance, *agl3*-Δ and *gtg*-Δ mutants both share defects in protein transport and altered stress ressitance (Aebi et al., 1996, Sharma Chandra et al., 2001). These large-scale screens could provide the basis for further investigation of the role of Sc Gtgp and discovery of any phenotypes the *gtg*-Δ mutant may have.

So far, there is more data supporting that GTG/GPHRs are the proposed Golgi pH regulators (Figure 8.4). However these proteins could still be linked to a mechanism that involves GPCRs and/or G proteins. Interestingly, loss-of-function mutations in the Arabidopsis Gα subunit, GPA1 (*gpa1-1* and *gpa1-4*), were crossed with BR receptor mutant, *bri1-5*, and BR-deficient mutant, *det2-1* respectively to generate *gpa1-1 bri1-5* and *gpa1-4 det2-1* double mutants. These exhibited enhanced development defects, displaying shorter roots and hypocotyls as well as less lateral roots (Gao et al., 2008), which are very similar phenotypes to the *gtg1 gtg2* double mutants (Figure 1.3). These results suggest there may be crosstalk between the G protein- and BR-mediated pathways. There is also scope that G protein signalling is responsible for the phenotypes observed in *C. elegans* double *gphr* mutants. In *C. elegans* the *egl-30* gene encodes for the $G_q\alpha$ subunit, knocking out this gene causes egg-laying, locomotion and viability phenotypes (Brundage et al., 1996) similar to the phenotyoes observed by the *C. elegans* double *gphr* mutant generated. Therefore there may still be a role for the GTG/GPHRs as a GPCR-type G proteins in both plants and animals. However, the human GPHR did not exhibit the same GTG binding or GTPase activity as the At GTGs under the same conditions (Pandey et al., 2009), so it would be important to clarify whether other animal GTG/GPHRs show a similar result.

In a recent study, recombinant At GTG1 interacted with ABA as shown by *in vitro* assays and in intact *S. cerevisiae* cells expressing recombinant At GTG1 (Kharenko et al., 2013). ABA and its receptors have recently been found to exist not only plants but also animals (Puce et al., 2004, Bruzzone et al., 2007, Bruzzone et al., 2008, Magnone et al., 2009, Scarfì et al., 2009, Sturla et al., 2009, Li et al., 2011, Sturla et al., 2011). ABA in animals has been shown to be important for immune responses through a PTX receptor/G protein complex involving a cascade of downstream signalling events including activation of cADPR and a rise in intracellular calcium concentration (Bruzzone et al., 2007, Bruzzone et al., 2008, Scarfì et al., 2009). The implications of ABA signalling in animals suggest that ABA functions as a stress hormone in both animals and plants with perhaps a conserved mechanism (Scarfì et al., 2009, Tossi et al., 2012).

In plants there is strong evidence supporting the PYR/PYL/RCAR proteins as ABA receptors (Ma et al., 2009, Park et al., 2009b, Santiago et al., 2009, Yin et al., 2009, Nishimura et al., 2010, Wang et al., 2013), although various receptors could exist

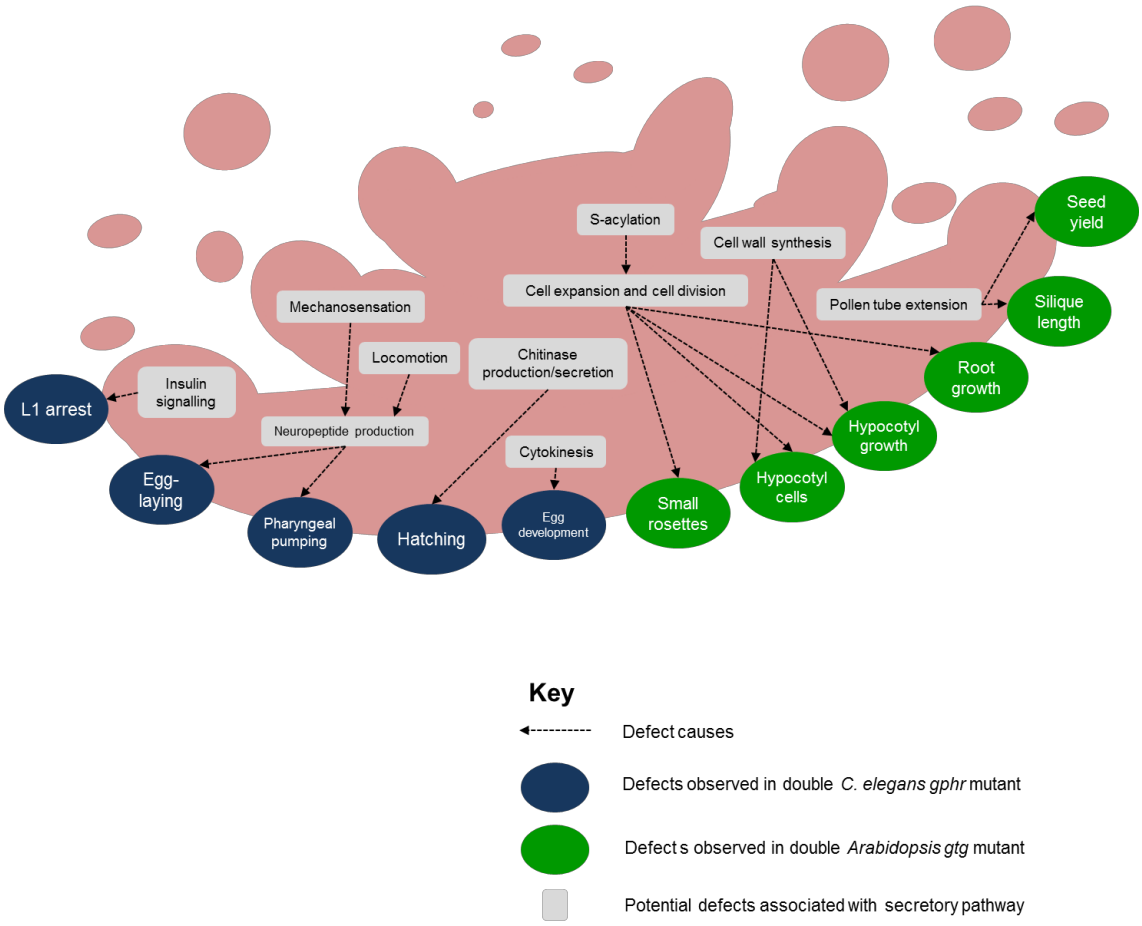


Figure 8.4 Arabidopsis and *C. elegans* double *gtg/gphr* mutant defects are associated with defects in the secretory pathway

Schematic showing the *C. elegans* double *gphr-1 gphr-2* mutant (blue) and Arabidopsis double *gtg1 gtg2* mutant (green) phenotypes observed in this study, with the potential defects associated with the secretory pathway (grey). For example, a defect in cell wall synthesis, which is associated with secretory pathway dysfunction, can cause the cell wall-related phenotype observed in the Arabidopsis *gtg1 gtg2* mutant.

for carrying out the myriad of ABA responses. In contrast, there is inconsistency in the literature and the work here to support the At GTGs as ABA receptors (Christmann and Grill, 2009, Jaffé et al., 2012). However, it was discussed that GTG/GPHRs could be involved in GPCR/G protein signalling (above). There is also data suggesting that ABA functions through a series of signal transduction events likely to involve GPCRs and G proteins (Bruzzone et al., 2007, Bruzzone et al., 2008, Alvarez et al., 2011, Li et al., 2011). Despite the compelling evidence that suggests these proteins have an alternative role to phytohormone signalling, it cannot be ruled out that the GTG/GPHRs could be GPCR-type G proteins involved in ABA signalling. The Arabidopsis GTGs were first proposed as ABA receptors due to the double *gtg* KO mutants exhibiting ABA-hyposensitivity (Pandey et al., 2009) but these experiments are not supported by subsequent analyses (Jaffé et al., 2012). One way to investigate the role of GTG/GPHRs in animal ABA signalling could be to use human or murine *GPHR* mutant cell lines. Since mammalian cells such as granulocytes are stimulated by ABA with measurable responses (Bruzzone et al., 2007), the *GPHR* mutant could be tested to observe whether similar responses occur. An ABA hyposensitive animal cell line would therefore advocate the investigation of GTG/GPHRs and ABA signalling in animals. Since ABA signalling is suggested to be conserved in plants and animals, the data shown here would still corrorborate a conserved function for these proteins if future data further supports GTG/GPHRs as protein modulators by ABA. For example these proteins could be Golgi pH regulators with anion channel activity that are regulated by ABA and the G proteins.

The role for the GTG/GPHRs is conserved between animal and plant kingdoms, but is this function an anion channel as reported by Maeda et al. (2008)? Consequently, future studies could therefore address what this conserved function is. The constructs generated in this project (Table 4.3, Table 5.2 and Table 5.3) have been used to produce expression vectors containing GFP-His tagged Arabidopsis and *C. elegans* GTG/GPHRs for expression in bacteria, human epithelial kidney (HEK) cells and baculovirus-insect cells. The aim will be to use HEK cells for anion channel activity measurements and bacterial or baculovirus cells for purification (Dr. L. Williams, personal communication). The vacuolar expression system is also another methods for testing channel activity. This system was used to show channel activity for the rat lysosomal chloride-proton exchanger, CLC-7. CLC-7-EGFP and expressed in

Arabidopsis mesophyll protoplasts showing expression at the vacuole membrane and exhibited Cl^- antiport activity (Costa et al., 2012). Demonstrating channel activity would confirm the proposed function of the GTG/GPHR proteins as an anion channel in the endomembrane system.

The pH of the Golgi and ER could be assessed to demonstrate that the loss of GTG/GPHR(s) results in dysfunctional acidification of endosomal organelles. This can be accomplished by using pH-sensitive fluorescent sensors, pHluorins, which can be targeted to endosomal organelles such as the ER and Golgi (Benčina, 2013). For example, pH-sensitive sensors have been used to demonstrate that pH homeostasis in endomembranes of Arabidopsis and tobacco cells is essential for normal function and that this was dependent on V-ATPase acidification as well as H^+ efflux mediated by at least one set of exporters such as the endosomal NHX-type antiporters (Martinière et al., 2013). Therefore, the GTG/GPHRs could act as a counterion channel necessary for ion flux facilitating endosomal acidification.

In conclusion, this project has provided a detailed sequence analysis of the GTG/GPHRs using computational methods; this has emphasised that these sequences are highly conserved, from protein structure and domain architecture to having completely conserved residues, possibly integral for structure/function. The GTG/GPHRs are not only conserved at the sequence level but have also been shown to have a conserved function through the use of cross-kingdom heterologous expression. The Arabidopsis *gtg1 gtg2* double mutant phenotypes can be rescued by *Cel*-GPHR-1 expression. In addition, the *Drosophila GPHR* KO mutants exhibit phenotypes that can be rescued by either *Drosophila*, human or hamster GPHR protein expression. Mutant *GPHR* CHO and *Drosophila* cells have altered Golgi pH and human GPHR exhibited anion channel activity. Therefore it is most likely that all genes, including the plant *GTG* genes, encode anion transporters that regulate Golgi pH. Additional phenotypic analysis have been carried out on the Arabidopsis *gtg1 gtg2* mutants showing that distinct mutants behave similarly, including that they are not BL-hyposensitive. Research of the GTG/GPHRs has also been initiated in the moncot species *O. sativa*, as well as a fungal system using *S. cerevisiae*. A broad study of the localisation of both Arabidopsis and *C. elegans* GTG/GPHR proteins has demonstrated that they are intracellularly localised and this is likely to be at the Golgi and ER using three model organisms. All data obtained in this thesis indicates that there is a conserved function

for the GTG/GPHR family of membrane proteins and the results further supports the proposed fundamental role as a Golgi pH regulator.

References

Adam, G. & Marquardt, V. (1986). Brassinosteroids. *Phytochemistry,* **25,** 1787-1799.

Adler, J., Pagakis, S. N. & Parmryd, I. (2008). Replicate-based noise corrected correlation for accurate measurements of colocalization. *Journal of Microscopy,* **230,** 121-133.

Aebi, M. (2013). *N*-linked protein glycosylation in the ER. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* **1833,** 2430-2437.

Aebi, M., Gassenhuber, J., Domdey, H. & Heesen, S. (1996). Cloning and characterization of the *ALG3* gene of *Saccharomyces cerevisiae*. *Glycobiology,* **6,** 439-444.

AGI, T. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature,* **408,** 796-815.

Ahringer, J. 2006. *Reverse genetics* [Online]. Available: http://www.wormbook.org.

Alberti, S., Gitler, A. D. & Lindquist, S. (2007). A suite of Gateway® cloning vectors for high-throughput genetic analysis in *Saccharomyces cerevisiae*. *Yeast,* **24,** 913-919.

Alifano, P., Fani, R., Liò, P., Lazcano, A., Bazzicalupo, M., Carlomagno, M. S. & Bruni, C. B. (1996). Histidine biosynthetic pathway and genes: structure, regulation, and evolution. *Microbiological reviews,* **60,** 44.

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P. & Cheuk, R. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science,* **301,** 653-657.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology,* **215,** 403-410.

Altun, Z. F. a. H., D. H. 2006. Introduction to *C. elegans* anatomy. *WormAtlas.*

Altun, Z. F. a. H., D.H. 2009. Introduction to *C. elegans*. *In:* Herndon, L. A. (ed.) *WormAtlas.*

Alvarez, S., Hicks, L. M. & Pandey, S. (2011). ABA-dependent and -independent G protein signaling in *Arabidopsis* roots revealed through an iTRAQ proteomics approach. *Journal of Proteome Research,* **10,** 3107-3122.

Alvarez, S., Roy Choudhury, S., Hicks, L. M. & Pandey, S. (2013). Quantitative proteomics-based analysis supports a significant role of GTG proteins in regulation of ABA response in *Arabidopsis* roots. *Journal of Proteome Research,* **12,** 1487-1501.

Ambros, V. (2000). Control of developmental timing in *Caenorhabditis elegans*. *Current Opinion in Genetics & Development,* **10,** 428-433.

Arai, M., Mitsuke, H., Ikeda, M., Xia, J.-X., Kikuchi, T., Satake, M. & Shimizu, T. (2004). ConPred II: a consensus prediction method for obtaining transmembrane topology models with high reliability. *Nucleic acids research,* **32,** W390-W393.

Asami, T., Min, Y. K., Nagata, N., Yamagishi, K., Takatsuto, S., Fujioka, S., Murofushi, N., Yamaguchi, I. & Yoshida, S. (2000). Characterization of brassinazole, a triazole-type brassinosteroid biosynthesis inhibitor. *Plant Physiology,* **123,** 93-100.

Avery, L. & Horvitz, H. R. (1990). Effects of starvation and neuroactive drugs on feeding in *Caenorhabditis elegans*. *Journal of Experimental Zoology,* **253,** 263-270.

Avery, L. & Shtonda, B. B. (2003). Food transport in the C. elegans pharynx. *Journal of Experimental Biology,* **206,** 2441-2457.

Avery, L. & You, Y.-J. 2012. C. elegans feeding. *In:* Community, T. C. E. R. (ed.) *WormBook.* WormBook.

Axelsson, M. A., Karlsson, N. G., Steel, D. M., Ouwendijk, J., Nilsson, T. & Hansson, G. C. (2001). Neutralization of pH in the Golgi apparatus causes redistribution of glycosyltransferases and changes in the *O*-glycosylation of mucins. *Glycobiology,* **11,** 633-644.

Axtell, M. J., Jan, C., Rajagopalan, R. & Bartel, D. P. (2006). A two-hit trigger for siRNA biogenesis in plants. *Cell,* **127,** 565-577.

Azpiroz, R., Wu, Y., LoCascio, J. C. & Feldmann, K. A. (1998). An Arabidopsis brassinosteroid-dependent mutant is blocked in cell elongation. *The Plant Cell Online,* **10,** 219-230.

B., W. W. 1988b. *The nematode C. elegans*, Cold Spring Harbor Laboratory Press, New York.

Baker Brachmann, C., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P. & Boeke, J. D. (1998). Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast,* **14,** 115-132.

Bamps, S. & Hope, I. A. (2008). Large-scale gene expression pattern analysis, *in situ*, in *Caenorhabditis elegans*. *Briefings in Functional Genomics & Proteomics,* **7,** 175-183.

Barak, L. S., Tiberi, M., Freedman, N. J., Kwatra, M. M., Lefkowitz, R. J. & Caron, M. G. (1994). A highly conserved tyrosine residue in G protein-coupled receptors is required for agonist-mediated β_2-adrenergic receptor sequestration. *Journal of Biological Chemistry,* **269,** 2790-2795.

Barneaud-Rocca, D., Etchebest, C. & Guizouarn, H. (2013). Structural model of the Anion Exchanger 1 (SLC4A1) and identification of transmembrane segments forming the transport site. *Journal of Biological Chemistry,* **288,** 26372-26384.

Barstead, R. & Moerman, D. 2006. *C. elegans* deletion mutant screening. *In:* Strange, K. (ed.) *C. elegans*. Humana Press.

Bassaganya-Riera, J., Skoneczka, J., Kingston, D., Krishnan, A., Misyak, S., Guri, A., Pereira, A., Carter, A., Minorsky, P. & Tumarkin, R. (2010). Mechanisms of action and medicinal applications of abscisic acid. *Current Medicinal Chemistry,* **17,** 467-478.

Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L. L., Studholme, D. J., Yeats, C. & Eddy, S. R. (2004). The Pfam protein families database. *Nucleic Acids Research,* **32,** D138-D141.

Baugh, L. R. & Sternberg, P. W. (2006). DAF-16/FOXO regulates transcription of *cki-1/Cip/Kip* and repression of *lin-4* during *C. elegans* L1 Arrest. *Current Biology,* **16,** 780-785.

Baulcombe, D. C. (2007). Amplified silencing. *Science,* **315,** 199-200.

Bell, A. & Charlwood, B. V. 1980. *Encyclopedia of plant physiology new series*, Springer-Verlag.

Bénard, C., McCright, B., Zhang, Y., Felkai, S., Lakowski, B. & Hekimi, S. (2001). The *C. elegans* maternal-effect gene *clk-2* is essential for embryonic development, encodes a protein homologous to yeast Tel2p and affects telomere length. *Development,* **128,** 4045-4055.

Benčina, M. (2013). Illumination of the spatial order of intracellular pH by genetically encoded pH-sensitive sensors. *Sensors,* **13,** 16736-16758.

Benfey, P. N., Ren, L. & Chua, N.-H. (1990). Tissue-specific expression from CaMV 35S enhancer subdomains in early stages of plant development. *The EMBO journal,* **9,** 1677.

Bergman, L. W. 2001. Growth and maintenance of yeast. *Two-Hybrid Systems*. Springer.

Bernsel, A., Viklund, H., Hennerdal, A. & Elofsson, A. (2009). TOPCONS: consensus prediction of membrane protein topology. *Nucleic acids research*, gkp363.

Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature,* **409,** 363-366.

Betts, M. J. & Russell, R. B. (2003). Amino acid properties and consequences of substitutions. *Bioinformatics for geneticists,* **317,** 289.

Betz, R., Manney, T. R. & Duntze, W. (1981). Hormonal control of gametogenesis in the yeast *Saccharomyces cerevisiae*. *Gamete Research,* **4,** 571-584.

Billing, O., Natarajan, B., Mohammed, A., Naredi, P. & Kao, G. (2012). A directed RNAi screen based on larval growth arrest reveals new modifiers of *C. elegans* insulin signaling. *PLoS ONE,* **7,** e34507.

Bird, A. F. a. B., J. 1991. *The structure of nematodes,* S. Diego, CA, Academic Press.

Bisht, N. C., Jez, J. M. & Pandey, S. (2011). An elaborate heterotrimeric G-protein family from soybean expands the diversity of plant G-protein networks. *New Phytologist,* **190,** 35-48.

Bissoli, G., Niñoles, R., Fresquet, S., Palombieri, S., Bueso, E., Rubio, L., García-Sánchez, M. J., Fernández, J. A., Mulet, J. M. & Serrano, R. (2012). Peptidyl-prolyl *cis-trans* isomerase ROF2 modulates intracellular pH homeostasis in *Arabidopsis*. *The Plant Journal,* **70,** 704-716.

Blancaflor, E. B. & Masson, P. H. (2003). Plant gravitropism. Unraveling the ups and downs of a complex process. *Plant Physiology,* **133,** 1677-1690.

Boeke, J. D., La Croute, F. & Fink, G. R. (1984). A positive selection for mutants lacking orotidine-5′ -phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Molecular and General Genetics MGG,* **197,** 345-346.

Boekhout, T. & Robert, V. 2003. *Yeasts in food*, Elsevier.

Bolte, S. & Cordelières, F. P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *Journal of microscopy,* **224,** 213-232.

Bork, P., Sander, C. & Valencia, A. (1993). Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. *Protein Science,* **2,** 31-40.

Botstein, D., Chervitz, S. A. & Cherry, J. M. (1997). Yeast as a model organism. *Science (New York, NY),* **277,** 1259-1260.

Boudart, G., Jamet, E., Rossignol, M., Lafitte, C., Borderies, G., Jauneau, A., Esquerré-Tugayé, M. T. & Pont-Lezica, R. (2005). Cell wall proteins in apoplastic fluids of *Arabidopsis thaliana* rosettes: identification by mass spectrometry and bioinformatics. *Proteomics,* **5,** 212-221.

Bracken, C., Iakoucheva, L. M., Romero, P. R. & Dunker, A. K. (2004). Combining prediction, computation and experiment for the characterization of protein disorder. *Current opinion in structural biology,* **14,** 570-576.

Branicky, R., Miyazaki, H., Strange, K. & Schafer, W. R. (2014). The voltage-gated anion channels encoded by *clh-3* regulate egg laying in *C. elegans* by modulating motor neuron excitability. *The Journal of Neuroscience,* **34,** 764-775.

Braun, N. A., Morgan, B., Dick, T. P. & Schwappach, B. (2010). The yeast CLC protein counteracts vesicular acidification during iron starvation. *Journal of Cell Science,* **123,** 2342-2350.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics,* **77,** 71-94.

Breslow, D. K., Cameron, D. M., Collins, S. R., Schuldiner, M., Stewart-Ornstein, J., Newman, H. W., Braun, S., Madhani, H. D., Krogan, N. J. & Weissman, J. S. (2008). A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nature methods,* **5,** 711-718.

Breton, G., Danyluk, J., Charron, J.-B. t. F. & Sarhan, F. (2003). Expression profiling and bioinformatic analyses of a novel stress-regulated multispanning transmembrane protein family from cereals and Arabidopsis. *Plant Physiology,* **132,** 64-74.

Broekhuis, J. R., Rademakers, S., Burghoorn, J. & Jansen, G. (2013). SQL-1, homologue of the Golgi protein GMAP210, modulates intraflagellar transport in C. elegans. *Journal of Cell Science,* **126,** 1785-1795.

Bromberg, Y. & Rost, B. (2007). SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic acids research,* **35,** 3823-3835.

Bromberg, Y., Yachdav, G. & Rost, B. (2008). SNAP predicts effect of mutations on protein function. *Bioinformatics,* **24,** 2397-2398.

Brown, D. M., Zeef, L. A., Ellis, J., Goodacre, R. & Turner, S. R. (2005). Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *The Plant Cell Online,* **17,** 2281-2295.

Brundage, L., Avery, L., Katz, A., Kim, U.-J., Mendel, J. E., Sternberg, P. W. & Simon, M. I. (1996). Mutations in a *C. elegans* $G_q\alpha$ gene disrupt movement, egg laying, and viability. *Neuron,* **16,** 999-1009.

Brüx, A., Liu, T.-Y., Krebs, M., Stierhof, Y.-D., Lohmann, J. U., Miersch, O., Wasternack, C. & Schumacher, K. (2008). Reduced V-ATPase activity in the trans-Golgi network causes oxylipin-dependent hypocotyl growth inhibition in Arabidopsis. *The Plant Cell Online,* **20,** 1088-1100.

Bruzzone, S., Basile, G., Mannino, E., Sturla, L., Magnone, M., Grozio, A., Salis, A., Fresia, C., Vigliarolo, T., Guida, L., De Flora, A., Tossi, V., Cassia, R., Lamattina, L. & Zocchi, E. (2012). Autocrine abscisic acid mediates the UV-B-induced inflammatory response in human granulocytes and keratinocytes. *Journal of Cellular Physiology,* **227,** 2502-2510.

Bruzzone, S., Bodrato, N., Usai, C., Guida, L., Moreschi, I., Nano, R., Antonioli, B., Fruscione, F., Magnone, M., Scarfì, S., De Flora, A. & Zocchi, E. (2008). Abscisic acid Is an endogenous stimulator of insulin release from human pancreatic islets with cyclic ADP ribose as second messenger. *Journal of Biological Chemistry,* **283,** 32188-32197.

Bruzzone, S., Moreschi, I., Usai, C., Guida, L., Damonte, G., Salis, A., Scarfi, S., Millo, E., De Flora, A. & Zocchi, E. (2007). Abscisic acid is an endogenous cytokine in human granulocytes with cyclic ADP-ribose as second messenger. *Proceedings of the National Academy of Sciences,* **104,** 5759-5764.

Brzywczy, J. & Paszewski, A. (1993). Role of *O*-acetylhomoserine sulfhydrylase in sulfur amino acid synthesis in various yeasts. *Yeast,* **9,** 1335-1342.

Byerly, L., Cassada, R. C. & Russell, R. L. (1976). The life cycle of the nematode *Caenorhabditis elegans*: I. Wild-type growth and reproduction. *Developmental Biology,* **51,** 23-33.

Cao, B., Porollo, A., Adamczak, R., Jarrell, M. & Meller, J. (2006). Enhanced recognition of protein transmembrane domains with prediction-based structural profiles. *Bioinformatics,* **22,** 303-309.

Carter, P. W., Roos, J. M. & Kemphues, K. J. (1990). Molecular analysis of *zyg-11*, a maternal-effect gene required for early embryogenesis of *Caenorhabditis elegans*. *Molecular and General Genetics MGG,* **221,** 72-80.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution,* **17,** 540-552.

Cavalier-Smith, T. (1998). A revised six-kingdom system of life. *Biological Reviews,* **73,** 203-266.

Chakravorty, D., Trusov, Y., Zhang, W., Acharya, B. R., Sheahan, M. B., McCurdy, D. W., Assmann, S. M. & Botella, J. R. (2011). An atypical heterotrimeric G-protein γ-subunit is involved in guard cell K^+ channel regulation and morphological development in *Arabidopsis thaliana*. *The Plant Journal,* **67,** 840-851.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. & Prasher, D. (1994). Green fluorescent protein as a marker for gene expression. *Science,* **263,** 802-805.

Chang, S.-L. & Leu, J.-Y. (2011). A tradeoff drives the evolution of reduced metal resistance in natural populations of yeast. *PLoS Genetics,* **7,** e1002034.

Chapman, R. E. & Munro, S. (1994). Retrieval of TGN proteins from the cell surface requires endosomal acidification. *The EMBO journal,* **13,** 2305.

Charroux, B. & Royet, J. (2014). Mutations in the Drosophila ortholog of the vertebrate Golgi pH regulator (GPHR) protein disturb endoplasmic reticulum and Golgi organization and affect systemic growth. *Biology open*, BIO20137187.

Chen, J.-G., Willard, F. S., Huang, J., Liang, J., Chasse, S. A., Jones, A. M. & Siderovski, D. P. (2003). A seven-transmembrane RGS protein that modulates plant cell proliferation. *Science,* **301,** 1728-1731.

Chen, Y., Ji, F., Xie, H., Liang, J. & Zhang, J. (2006). The regulator of G-protein signaling proteins involved in sugar and abscisic acid signaling in Arabidopsis seed germination. *Plant Physiology,* **140,** 302-310.

Cheon, J., Park, S.-Y., Schulz, B. & Choe, S. (2010). *Arabidopsis* brassinosteroid biosynthetic mutant *dwarf7-1* exhibits slower rates of cell division and shoot induction. *BMC plant biology,* **10,** 270.

Cherfils, J. & Zeghouf, M. 2013. *Regulation of small GTPases by GEFs, GAPs, and GDIs*.

Cherry, J. M., Hong, E. L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E. T., Christie, K. R., Costanzo, M. C., Dwight, S. S., Engel, S. R., Fisk, D. G., Hirschman, J. E., Hitz, B. C., Karra, K., Krieger, C. J., Miyasato, S. R., Nash, R. S., Park, J., Skrzypek, M. S., Simison, M., Weng, S. & Wong, E. D. (2012). Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Research,* **40,** D700-D705.

Chervitz, S. A., Hester, E. T., Ball, C. A., Dolinski, K., Dwight, S. S., Harris, M. A., Juvik, G., Malekian, A., Roberts, S. & Roe, T. (1999). Using the Saccharomyces Genome Database (SGD) for analysis of protein similarities and structure. *Nucleic acids research,* **27,** 74-78.

Chilton, M.-D., Drummond, M. H., Merlo, D. J., Sciaky, D., Montoya, A. L., Gordon, M. P. & Nester, E. W. (1977). Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. *Cell,* **11,** 263-271.

Choe, S., Dilkes, B. P., Fujioka, S., Takatsuto, S., Sakurai, A. & Feldmann, K. A. (1998). The *DWF4* gene of *Arabidopsis* encodes a cytochrome P450 that mediates multiple 22α-hydroxylation steps in brassinosteroid biosynthesis. *The Plant Cell Online,* **10,** 231-243.

Choe, S., Dilkes, B. P., Gregory, B. D., Ross, A. S., Yuan, H., Noguchi, T., Fujioka, S., Takatsuto, S., Tanaka, A. & Yoshida, S. (1999). The *Arabidopsis dwarf1* mutant is defective in the conversion of 24-methylenecholesterol to campesterol in brassinosteroid biosynthesis. *Plant physiology,* **119,** 897-908.

Christmann, A. & Grill, E. (2009). Are GTGs ABA's biggest fans? *Cell,* **136,** 21-23.

Christmann, A., Moes, D., Himmelbach, A., Yang, Y., Tang, Y. & Grill, E. (2006). Integration of abscisic acid signalling into plant responses. *Plant Biology,* **8,** 314-325.

Chudakov, D. M., Matz, M. V., Lukyanov, S. & Lukyanov, K. A. (2010). Fluorescent proteins and their applications in imaging living cells and tissues. *Physiological Reviews,* **90,** 1103-1163.

Cimino, G. D., Gamper, H. B., Isaacs, S. T. & Hearst, J. E. (1985). Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Annual review of biochemistry,* **54,** 1151-1193.

Clark, K. M., Fedoriw, N., Robinson, K., Connelly, S. M., Randles, J., Malkowski, M. G., DeTitta, G. T. & Dumont, M. E. (2010). Purification of transmembrane proteins from *Saccharomyces cerevisiae* for X-ray crystallography. *Protein expression and purification,* **71,** 207-223.

Claros, M. G. & von Heijne, G. (1994). TopPred II: an improved software for membrane protein structure predictions. *Computer applications in the biosciences: CABIOS,* **10,** 685-686.

Clokey, G. V. & Jacobson, L. A. (1986). The autofluorescent "lipofuscin granules" in the intestinal cells of *Caenorhabditis elegans* are secondary lysosomes. *Mechanisms of Ageing and Development,* **35,** 79-94.

Clough, S. J. & Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal,* **16,** 735-743.

Clouse, S. D. (1996). Molecular genetic studies confirm the role of brassinosteroids in plant growth and development. *The Plant Journal,* **10,** 1-8.

Clouse, S. D., Langford, M. & McMorris, T. C. (1996). A brassinosteroid-insensitive mutant in *Arabidopsis thaliana* exhibits multiple defects in growth and development. *Plant Physiology,* **111,** 671-678.

Cogoni, C. & Macino, G. (1999). Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature,* **399,** 166-169.

Colombini, M. (1979). A candidate for the permeability pathway of the outer mitochondrial membrane.

Consortium, T. C. e. S. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science,* **282,** 2012-2018.

Čopič, A., Dorrington, M., Pagant, S., Barry, J., Lee, M. C. S., Singh, I., Hartman, J. L. & Miller, E. A. (2009). Genomewide analysis reveals novel pathways affecting endoplasmic reticulum homeostasis, protein modification and quality control. *Genetics,* **182,** 757-769.

Costa, A., Gutla, P. V. K., Boccaccio, A., Scholz-Starke, J., Festa, M., Basso, B., Zanardi, I., Pusch, M., Schiavo, F. L., Gambale, F. & Carpaneto, A. (2012). The *Arabidopsis* central vacuole as an expression system for intracellular transporters: functional characterization of the Cl^-/H^+ exchanger CLC-7. *The Journal of Physiology,* **590,** 3421-3430.

Costes, S. V., Daelemans, D., Cho, E. H., Dobbin, Z., Pavlakis, G. & Lockett, S. (2004). Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysical Journal,* **86,** 3993-4003.

Cox, C. M. & Swain, S. M. (2006). Localised and non-localised promotion of fruit development by seeds in Arabidopsis. *Functional plant biology,* **33,** 1-8.

Curtis, M. D. & Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant physiology,* **133,** 462-469.

De Craene, J.-O., Courte, F., Rinaldi, B., Fitterer, C., Herranz, M. C., Schmitt-Keichinger, C., Ritzenthaler, C. & Friant, S. (2014). Study of the plant COPII vesicle coat subunits by functional complementation of yeast *Saccharomyces cerevisiae* mutants. *PLoS ONE,* **9,** e90072.

De Lean, A., Stadel, J. & Lefkowitz, R. (1980). A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled β-adrenergic receptor. *Journal of Biological Chemistry,* **255,** 7108-7117.

Dejima, K., Murata, D., Mizuguchi, S., Nomura, K. H., Gengyo-Ando, K., Mitani, S., Kamiyama, S., Nishihara, S. & Nomura, K. (2009). The ortholog of human solute carrier family 35 member B1 (UDP-galactose transporter-related protein 1) is involved in maintenance of ER homeostasis and essential for larval development in *Caenorhabditis elegans*. *The FASEB Journal,* **23,** 2215-2225.

Dejima, K., Murata, D., Mizuguchi, S., Nomura, K. H., Izumikawa, T., Kitagawa, H., Gengyo-Ando, K., Yoshina, S., Ichimiya, T., Nishihara, S., Mitani, S. & Nomura, K. (2010). Two Golgi-resident 3′ -phosphoadenosine 5′ -phosphosulfate transporters play distinct roles in heparan sulfate modifications and embryonic and larval development in *Caenorhabditis elegans*. *Journal of Biological Chemistry,* **285,** 24717-24728.

Demaegd, D., Foulquier, F., Colinet, A.-S., Gremillon, L., Legrand, D., Mariot, P., Peiter, E., Van Schaftingen, E., Matthijs, G. & Morsomme, P. (2013). Newly characterized Golgi-localized family of proteins is involved in calcium and pH homeostasis in yeast and human cells. *Proceedings of the National Academy of Sciences,* **110,** 6859-6864.

Dennison, K. L., Robertson, W. R., Lewis, B. D., Hirsch, R. E., Sussman, M. R. & Spalding, E. P. (2001). Functions of AKT1 and AKT2 potassium channels determined by studies of single and double mutants of Arabidopsis. *Plant physiology,* **127,** 1012-1019.

Desai, C., Garriga, G., McLntire, S. L. & Horvitz, H. R. (1988). A genetic pathway for the development of the *Caenorhabditis elegans* HSN motor neurons. *Nature,* **336,** 638-646.

Desnos, T., Orbovic, V., Bellini, C., Kronenberger, J., Caboche, M., Traas, J. & Hofte, H. (1996). *Procuste1* mutants identify two distinct genetic pathways controlling hypocotyl cell elongation, respectively in dark- and light-grown *Arabidopsis* seedlings. *Development,* **122,** 683-693.

Dettmer, J., Schubert, D., Calvo-Weimar, O., Stierhof, Y.-D., Schmidt, R. & Schumacher, K. (2005). Essential role of the V-ATPase in male gametophyte development. *The Plant Journal,* **41,** 117-124.

Di Laurenzio, L., Wysocka-Diller, J., Malamy, J. E., Pysh, L., Helariutta, Y., Freshour, G., Hahn, M. G., Feldmann, K. A. & Benfey, P. N. (1996). The *SCARECROW* gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root. *Cell,* **86,** 423-433.

Dibb, N. J., Maruyama, I. N., Krause, M. & Karn, J. (1989). Sequence analysis of the complete *Caenorhabditis elegans* myosin heavy chain gene family. *Journal of Molecular Biology,* **205,** 603-613.

Diévart, A. & Clark, S. E. (2004). LRR-containing receptors regulating plant development and defense. *Development,* **131,** 251-261.

Dixon, R. A. F., Kobilka, B. K., Strader, D. J., Benovic, J. L., Dohlman, H. G., Frielle, T., Bolanowski, M. A., Bennett, C. D., Rands, E., Diehl, R. E., Mumford, R. A., Slater, E. E., Sigal, I. S., Caron, M. G., Lefkowitz, R. J. & Strader, C. D. (1986). Cloning of the gene and cDNA for mammalian β-adrenergic receptor and homology with rhodopsin. *Nature,* **321,** 75-79.

Dogovski, C., Pi, J. & Pittard, A. J. (2003). Putative interhelical interactions within the PheP protein revealed by second-site suppressor analysis. *J. Bacteriol.,* **185,** 6225-6232.

Doncaster, C. C. (1962). Nematode feeding mechanisms. 1. Observations on rhabditis and pelodera. *Nematologica,* **8,** 313-320.

dos Santos de Lima e Souza, D., de Souza Junior, J. D. A., Grossi-de-Sá, M., Rocha, T. L., Fragoso, R. d. R., de Deus Barbosa, A. E. A., de Oliveira, G. R., Nakasu, E. Y. T., de Sousa, B. A., Pires, N. F., de Alencar Dusi, D. M., Carneiro, R. M. D. G., Romano, E., de Almeida-Engler, J., Engler, G., Martins-de-Sá, C. & Grossi-de-Sá, M. F. (2011). Ectopic expression of a *Meloidogyne incognita* dorsal gland protein in tobacco accelerates the formation of the nematode feeding site. *Plant Science,* **180,** 276-282.

Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T. & MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science,* **280,** 69-77.

Draper, J., Mur, L. A., Jenkins, G., Ghosh-Biswas, G. C., Bablak, P., Hasterok, R. & Routledge, A. P. (2001). *Brachypodium distachyon*. A new model system for functional genomics in grasses. *Plant physiology,* **127,** 1539-1555.

Drew, D., Newstead, S., Sonoda, Y., Kim, H., von Heijne, G. & Iwata, S. (2008). GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in *Saccharomyces cerevisiae*. *Nature protocols,* **3,** 784-798.

Du, L.-L. & Novick, P. (2001). Yeast Rab GTPase-activating protein Gyp1p localizes to the Golgi apparatus and is a negative regulator of Ypt1p. *Molecular Biology of the Cell,* **12,** 1215-1226.

Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M. & Hipps, K. W. (2001). Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling,* **19,** 26-59.

Dupuy, D., Bertin, N., Hidalgo, C. A., Venkatesan, K., Tu, D., Lee, D., Rosenberg, J., Svrzikapa, N., Blanc, A., Carnec, A., Carvunis, A.-R., Pulak, R., Shingles, J., Reece-Hoyes, J., Hunt-Newbury, R., Viveiros, R., Mohler, W. A., Tasan, M., Roth, F. P., Le Peuch, C., Hope, I. A., Johnsen, R., Moerman, D. G., Barabasi, A.-L., Baillie, D. & Vidal, M. (2007). Genome-scale analysis of *in vivo* spatiotemporal promoter activity in *Caenorhabditis elegans*. *Nature Biotechnology,* **25,** 663-668.

Dyson, H. J. & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nature reviews Molecular cell biology,* **6,** 197-208.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research,* **32,** 1792-1797.

Edgley, M. L., Baillie, D. L., Riddle, D. L. & Rose, A. M. (2006). Genetic balancers.

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. & Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature,* **411,** 494-498.

eLibrary, P. S. S. 2014. *Genetic basis for the backcross method* [Online]. Available: http://passel.unl.edu/pages/informationmodule.php?idinformationmodule=959723462&topicorder=3&maxto=7 2014].

Elortza, F., Nühse, T. S., Foster, L. J., Stensballe, A., Peck, S. C. & Jensen, O. N. (2003). Proteomic analysis of glycosylphosphatidylinositol-anchored membrane proteins. *Molecular & Cellular Proteomics,* **2,** 1261-1270.

Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nature protocols,* **2,** 953-971.

Estévez, R. & Jentsch, T. (2002). CLC chloride channels: correlating structure with function. *Current Opinion in Structural Biology,* **12,** 531-539.

Fagard, M., Desnos, T., Desprez, T., Goubet, F., Refregier, G., Mouille, G., McCann, M., Rayon, C., Vernhettes, S. & Höfte, H. (2000). *PROCUSTE1* encodes a cellulose synthase required for normal cell elongation specifically in roots and dark-grown hypocotyls of *Arabidopsis*. *The Plant Cell Online,* **12,** 2409-2423.

Fan, L.-M., Zhang, W., Chen, J.-G., Taylor, J. P., Jones, A. M. & Assmann, S. M. (2008). Abscisic acid regulation of guard-cell K^+ and anion channels in Gβ- and RGS-deficient *Arabidopsis* lines. *Proceedings of the National Academy of Sciences,* **105,** 8476-8481.

Fernandez, A. G., Bargmann, B. O., Mis, E. K., Edgley, M. L., Birnbaum, K. D. & Piano, F. (2012). High-throughput fluorescence-based isolation of live *C. elegans* larvae. *Nature protocols,* **7,** 1502-1510.

Fernandez, A. G. & Piano, F. (2006). MEL-28 is downstream of the Ran cycle and is required for nuclear-envelope function and chromatin maintenance. *Current biology,* **16,** 1757-1763.

Ferrier, A., Charron, A., Sadozai, Y., Switaj, L., Szutenbach, A. & Smith, P. A. (2011). Multiple phenotypes resulting from a mutagenesis screen for pharynx muscle mutations in *Caenorhabditis elegans*. *PLoS ONE,* **6,** e26594.

Finch-Savage, W. E. & Leubner-Metzger, G. (2006). Seed dormancy and the control of germination. *New Phytologist,* **171,** 501-523.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature,* **391,** 806-811.

Fontana, L., Partridge, L. & Longo, V. D. (2010). Extending healthy life span - from yeast to humans. *Science,* **328,** 321-326.

Foster, L. J., Zeemann, P. A., Li, C., Mann, M., Jensen, O. N. & Kassem, M. (2005). Differential expression profiling of membrane proteins by quantitative proteomics in a human mesenchymal stem cell line undergoing osteoblast differentiation. *Stem Cells,* **23,** 1367-1377.

Franck, A., Guilley, H., Jonard, G., Richards, K. & Hirth, L. (1980). Nucleotide sequence of cauliflower mosaic virus DNA. *Cell,* **21,** 285-294.

Franks, C. J., Holden-Dye, L., Bull, K., Luedtke, S. & Walker, R. J. (2006). Anatomy, physiology and pharmacology of *Caenorhabditis elegans* pharynx: a model to define gene function in a simple neural system. *Invert Neurosci,* **6,** 105-122.

Frommer, W. B. & Ninnemann, O. (1995). Heterologous expression of genes in bacterial, fungal, animal, and plant cells. *Annual Review of Plant Physiology and Plant Molecular Biology,* **46,** 419-444.

Fujita, M., Umemura, M., Yoko-o, T. & Jigami, Y. (2006). *PER1* is required for GPI-phospholipase A2 activity and involved in lipid remodeling of GPI-anchored proteins. *Molecular Biology of the Cell,* **17,** 5253-5264.

Fung, B. K.-K. & Stryer, L. (1980). Photolyzed rhodopsin catalyzes the exchange of GTP for bound GDP in retinal rod outer segments. *Proceedings of the National Academy of Sciences of the United States of America,* **77,** 2500-2504.

Furuichi, T., Cunningham, K. W. & Muto, S. (2001a). A putative two pore channel AtTPC1 mediates Ca^{2+} flux in *Arabidopsis* leaf cells. *Plant and Cell Physiology,* **42,** 900-905.

Furuichi, T., Mori, I. C., Takahashi, K. & Muto, S. (2001b). Sugar-induced increase in cytosolic Ca^{2+} in *Arabidopsis thaliana* whole plants. *Plant and Cell Physiology,* **42,** 1149-1155.

Galagan, J. E., Calvo, S. E., Borkovich, K. A., Selker, E. U., Read, N. D., Jaffe, D., FitzHugh, W., Ma, L.-J., Smirnov, S., Purcell, S., Rehman, B., Elkins, T., Engels,

R., Wang, S., Nielsen, C. B., Butler, J., Endrizzi, M., Qui, D., Ianakiev, P., Bell-Pedersen, D., Nelson, M. A., Werner-Washburne, M., Selitrennikoff, C. P., Kinsey, J. A., Braun, E. L., Zelter, A., Schulte, U., Kothe, G. O., Jedd, G., Mewes, W., Staben, C., Marcotte, E., Greenberg, D., Roy, A., Foley, K., Naylor, J., Stange-Thomann, N., Barrett, R., Gnerre, S., Kamal, M., Kamvysselis, M., Mauceli, E., Bielke, C., Rudd, S., Frishman, D., Krystofova, S., Rasmussen, C., Metzenberg, R. L., Perkins, D. D., Kroken, S., Cogoni, C., Macino, G., Catcheside, D., Li, W., Pratt, R. J., Osmani, S. A., DeSouza, C. P. C., Glass, L., Orbach, M. J., Berglund, J. A., Voelker, R., Yarden, O., Plamann, M., Seiler, S., Dunlap, J., Radford, A., Aramayo, R., Natvig, D. O., Alex, L. A., Mannhaupt, G., Ebbole, D. J., Freitag, M., Paulsen, I., Sachs, M. S., Lander, E. S., Nusbaum, C. & Birren, B. (2003). The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature,* **422,** 859-868.

Galy, V., Askjaer, P., Franz, C., López-Iglesias, C. & Mattaj, I. W. (2006). MEL-28, a novel nuclear-envelope and kinetochore protein sssential for zygotic nuclear-envelope assembly in *C. elegans Current biology,* **16,** 1748-1756.

Gao, X., Bai, Y. & Hwang, T.-C. (2013). Cysteine ccanning of CFTR's first transmembrane segment reveals its plausible roles in gating and permeation. *Biophysical Journal,* **104,** 786-797.

Gao, Y., Wang, S., Asami, T. & Chen, J.-G. (2008). Loss-of-function mutations in the Arabidopsis heterotrimeric G-protein α subunit enhance the developmental defects of brassinosteroid signaling and biosynthesis mutants. *Plant and cell physiology,* **49,** 1013-1024.

Gao, Y., Zeng, Q., Guo, J., Cheng, J., Ellis, B. E. & Chen, J.-G. (2007). Genetic characterization reveals no role for the reported ABA receptor, GCR2, in ABA control of seed germination and early seedling development in Arabidopsis. *The Plant Journal,* **52,** 1001-1013.

Garfinkel, D. J. & Nester, E. W. (1980). Agrobacterium tumefaciens mutants affected in crown gall tumorigenesis and octopine catabolism. *Journal of bacteriology,* **144,** 732-743.

Gaudet, J. & Mango, S. E. (2002). Regulation of organogenesis by the *Caenorhabditis elegans* FoxA Protein PHA-4. *Science,* **295,** 821-825.

Gaxiola, R. A., Yuan, D. S., Klausner, R. D. & Fink, G. R. (1998). The yeast CLC chloride channel functions in cation homeostasis. *Proceedings of the National Academy of Sciences,* **95,** 4046-4050.

Ge, X., Wang, H. & Cao, K. (2008). Transformation by T-DNA integration causes highly sterile phenotype independent of transgenes in *Arabidopsis thaliana*. *Plant cell reports,* **27,** 1341-1348.

Geelen, D., Lurin, C., Bouchez, D., Frachisse, J. M., Lelièvre, F., Courtial, B., Barbier-Brygoo, H. & Maurel, C. (2000). Disruption of putative anion channel gene AtCLC-a in Arabidopsis suggests a role in the regulation of nitrate content. *The Plant Journal,* **21,** 259-267.

Gems, D., Sutton, A. J., Sundermeyer, M. L., Albert, P. S., King, K. V., Edgley, M. L., Larsen, P. L. & Riddle, D. L. (1998). Two pleiotropic classes of *daf-2* mutation affect larval arrest, adult behavior, reproduction and longevity in *Caenorhabditis elegans*. *Genetics,* **150,** 129-155.

Geng, W., Cosman, P., Berry, C. C., Feng, Z. & Schafer, W. R. (2004). Automatic tracking, feature extraction and classification of C. elegans phenotypes. *Biomedical Engineering, IEEE Transactions on,* **51,** 1811-1820.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., Arkin, A. P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K.-D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Guldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kotter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L.,

Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C.-y., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W. & Johnston, M. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature,* **418,** 387-391.

Gitan, R. S., Luo, H., Rodgers, J., Broderius, M. & Eide, D. (1998). Zinc-induced Inactivation of the Yeast ZRT1 Zinc Transporter Occurs through Endocytosis and Vacuolar Degradation. *Journal of Biological Chemistry,* **273,** 28617-28624.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H. & Oliver, S. G. (1996). Life with 6000 genes. *Science,* **274,** 546-567.

Golan, D. E., Tashjian, A. H. & Armstrong, E. J. 2011. *Principles of pharmacology: the pathophysiologic basis of drug therapy*, Lippincott Williams & Wilkins.

Goldberg, T., Hamp, T. & Rost, B. (2012). LocTree2 predicts localization for all domains of life. *Bioinformatics,* **28,** i458-i465.

Gönczy, P., Echeverri, C., Oegema, K., Coulson, A., Jones, S. J., Copley, R. R., Duperon, J., Oegema, J., Brehm, M. & Cassin, E. (2000). Functional genomic analysis of cell division in C. elegans using RNAi of genes on chromosome III. *Nature,* **408,** 331-336.

Gong, W., Shen, Y.-P., Ma, L.-G., Pan, Y., Du, Y.-L., Wang, D.-H., Yang, J.-Y., Hu, L.-D., Liu, X.-F., Dong, C.-X., Ma, L., Chen, Y.-H., Yang, X.-Y., Gao, Y., Zhu, D., Tan, X., Mu, J.-Y., Zhang, D.-B., Liu, Y.-L., Dinesh-Kumar, S. P., Li, Y., Wang, X.-P., Gu, H.-Y., Qu, L.-J., Bai, S.-N., Lu, Y.-T., Li, J.-Y., Zhao, J.-D., Zuo, J., Huang, H., Deng, X. W. & Zhu, Y.-X. (2004). Genome-wide ORFeome cloning and analysis of *Arabidopsis* transcription factor genes. *Plant Physiology,* **135,** 773-782.

Gough, J. (2005). Convergent evolution of domain architectures (is rare). *Bioinformatics,* **21,** 1464-1471.

Granot, D. & Snyder, M. (1993). Carbon source induces growth of stationary phase yeast cells, independent of carbon source metabolism. *Yeast,* **9,** 465-479.

Granseth, E., Viklund, H. & Elofsson, A. (2006). ZPRED: predicting the distance to the membrane center for residues in α-helical membrane proteins. *Bioinformatics,* **22,** e191-e196.

Grove, M. D., Spencer, G. F., Rohwedder, W. K., Mandava, N., Worley, J. F., Warthen, J. D., Steffens, G. L., Flippen-Anderson, J. L. & Cook, J. C. (1979). Brassinolide, a plant growth-promoting steroid isolated from *Brassica napus* pollen.

Guan, K. & Dixon, J. E. (1991). Eukaryotic proteins expressed in *Escherichia coli:* an improved thrombin cleavage and purification procedure of fusion proteins with glutathione *S*-transferase. *Analytical biochemistry,* **192,** 262-267.

Guarente, L. & Kenyon, C. (2000). Genetic pathways that regulate ageing in model organisms. *Nature,* **408,** 255-262.

Guo, J., Zeng, Q., Emami, M., Ellis, B. E. & Chen, J.-G. (2008). The *GCR2* gene family is not required for ABA control of seed germination and early seedling development in *Arabidopsis*. *PLoS ONE,* **3,** e2982.

Guo, W., Zuo, Z., Cheng, X., Sun, J., Li, H., Li, L. & Qiu, J.-L. (2014). The chloride channel family gene *CLCd* negatively regulates pathogen-associated molecular pattern (PAMP)-triggered immunity in *Arabidopsis*. *Journal of Experimental Botany,* **65,** 1205-1215.

Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature,* **404,** 293-296.

Hanks, S. K., Quinn, A. M. & Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science,* **241,** 42-52.

Hargrave, P. A., McDowell, J. H., Curtis, D., Wang, J., Juszczak, E., Fong, S.-L., Mohana Rao, J. K. & Argos, P. (1983). The structure of bovine rhodopsin. *Biophysics of structure and mechanism,* **9,** 235-244.

Hashimoto, H. & Yoda, K. (1997). Novel membrane protein complexes for protein glycosylation in the yeast Golgi apparatus. *Biochemical and Biophysical Research Communications,* **241,** 682-686.

Hechenberger, M., Schwappach, B., Fischer, W. N., Frommer, W. B., Jentsch, T. J. & Steinmeyer, K. (1996). A family of putative chloride channels from *Arabidopsis* and functional complementation of a yeast strain with a *CLC* gene disruption. *Journal of Biological Chemistry,* **271,** 33632-33638.

Hecht, M., Bromberg, Y. & Rost, B. (2013). News from the protein mutability landscape. *Journal of molecular biology,* **425,** 3937-3948.

Heginbotham, L., Lu, Z., Abramson, T. & MacKinnon, R. (1994). Mutations in the K^+ channel signature sequence. *Biophysical Journal,* **66,** 1061-1067.

Herndon, L. A., Lints, R. and Hall, D.H. 2009. Glossary E. *WormAtlas.*

Hietakangas, V. & Cohen, S. M. (2009). Regulation of tissue growth through nutrient sensing. *Annual Review of Genetics,* **43,** 389-410.

Hirokawa, T., Boon-Chieng, S. & Mitaku, S. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics,* **14,** 378-379.

Hirschman, J. E., Balakrishnan, R., Christie, K. R., Costanzo, M. C., Dwight, S. S., Engel, S. R., Fisk, D. G., Hong, E. L., Livstone, M. S., Nash, R., Park, J., Oughtred, R., Skrzypek, M., Starr, B., Theesfeld, C. L., Williams, J., Andrada, R., Binkley, G., Dong, Q., Lane, C., Miyasato, S., Sethuraman, A., Schroeder, M., Thanawala, M. K., Weng, S., Dolinski, K., Botstein, D. & Cherry, J. M. (2006). Genome Snapshot: a new resource at the Saccharomyces Genome Database (SGD) presenting an overview of the *Saccharomyces cerevisiae* genome. *Nucleic Acids Research,* **34,** D442-D445.

Hobson, R. J., Hapiak, V. M., Xiao, H., Buehrer, K. L., Komuniecki, P. R. & Komuniecki, R. W. (2006). SER-7, a *Caenorhabditis elegans* 5-HT_7-like receptor, is essential for the 5-HT stimulation of pharyngeal pumping and egg laying. *Genetics,* **172,** 159-169.

Hodgkin, J. & Doniach, T. (1997). Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics,* **146,** 149-164.

Hoekema, A., Hirsch, P., Hooykaas, P. & Schilperoort, R. (1983). A binary plant vector strategy based on separation of *vir*-and T-region of the *Agrobacterium tumefaciens* Ti-plasmid.

Hofmann, K. (1993). TMbase-A database of membrane spanning protein segments. *Biol. Chem. Hoppe-Seyler,* **374,** 166.

Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J. & Nakai, K. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Research,* **35,** W585-W587.

Horton, P., Park, K.-J., Obayashi, T. & Nakai, K. Protein subcellular localisation prediction with WoLF PSORT. APBC, 2006. Citeseer, 39-48.

Horvitz, H. R., Brenner, S., Hodgkin, J. & Herman, R. K. (1979). A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. *Molecular and General Genetics MGG,* **175,** 129-133.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W. & Zimmermann, P. (2008). Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Advances in bioinformatics,* **2008**.

Huala, E., Dickerman, A. W., Garcia-Hernandez, M., Weems, D., Reiser, L., LaFond, F., Hanley, D., Kiphart, D., Zhuang, M. & Huang, W. (2001). The Arabidopsis Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. *Nucleic acids research,* **29,** 102-105.

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P. S., Pagni, M. & Sigrist, C. J. (2006). The PROSITE database. *Nucleic acids research,* **34,** D227-D230.

Huo, N., Vogel, J. P., Lazo, G. R., You, F. M., Ma, Y., McMahon, S., Dvorak, J., Anderson, O. D., Luo, M.-C. & Gu, Y. Q. (2009). Structural characterization of *Brachypodium*

genome and its syntenic relationship with rice and wheat. *Plant molecular biology,* **70,** 47-61.

Huse, J. T., Liu, K., Pijak, D. S., Carlin, D., Lee, V. M.-Y. & Doms, R. W. (2002). β-secretase processing in the trans-Golgi network preferentially generates truncated amyloid species that accumulate in Alzheimer's disease brain. *Journal of Biological Chemistry,* **277,** 16278-16284.

Husson, S. J., Clynen, E., Baggerman, G., Janssen, T. & Schoofs, L. (2006). Defective processing of neuropeptide precursors in *Caenorhabditis elegans* lacking proprotein convertase 2 (KPC-2/EGL-3): mutant analysis by mass spectrometry. *Journal of Neurochemistry,* **98,** 1999-2012.

Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradović, Z. & Dunker, A. K. (2002). Intrinsic disorder in cell-signaling and cancer-associated proteins. *Journal of molecular biology,* **323,** 573-584.

Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z. & Dunker, A. K. (2004). The importance of intrinsic disorder for protein phosphorylation. *Nucleic acids research,* **32,** 1037-1049.

Inoue, T., Higuchi, M., Hashimoto, Y., Seki, M., Kobayashi, M., Kato, T., Tabata, S., Shinozaki, K. & Kakimoto, T. (2001). Identification of CRE1 as a cytokinin receptor from *Arabidopsis*. *Nature,* **409,** 1060-1063.

Inouye, S. & Tsuji, F. I. (1994). *Aequorea* green fluorescent protein: expression of the gene and fluorescence characteristics of the recombinant protein. *FEBS letters,* **341,** 277-280.

Irizarry, S. N., Kutluay, E., Drews, G., Hart, S. J. & Heginbotham, L. (2002). Opening the KcsA K^+ channel: tryptophan scanning and complementation analysis lead to mutants with altered gating. *Biochemistry,* **41,** 13653-13662.

Iyer, R., Iverson, T. M., Accardi, A. & Miller, C. (2002). A biological role for prokaryotic ClC chloride channels. *Nature,* **419,** 715-718.

Jackson, M. A., Sternes, P. R., Mudge, S. R., Graham, M. W. & Birch, R. G. (2014). Design rules for efficient transgene expression in plants. *Plant biotechnology journal*.

Jacob, T. C. & Kaplan, J. M. (2003). The EGL-21 carboxypeptidase E facilitates acetylcholine release at *Caenorhabditis elegans* neuromuscular junctions. *The Journal of Neuroscience,* **23,** 2122-2130.

Jaffé, F. W., Freschet, G.-E. C., Valdes, B. M., Runions, J., Terry, M. J. & Williams, L. E. (2012). G protein–coupled receptor-type G proteins are required for light-dependent seedling growth and fertility in *Arabidopsis*. *The Plant Cell Online,* **24,** 3649-3668.

Jain, M., Nijhawan, A., Arora, R., Agarwal, P., Ray, S., Sharma, P., Kapoor, S., Tyagi, A. K. & Khurana, J. P. (2007). F-box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. *Plant Physiology,* **143,** 1467-1483.

Jantsch-Plunger, V. & Fire, A. (1994). Combinatorial structure of a body muscle-specific transcriptional enhancer in *Caenorhabditis elegans*. *Journal of Biological Chemistry,* **269,** 27021-27028.

Javadpour, M. M., Eilers, M., Groesbeek, M. & Smith, S. O. (1999). Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association. *Biophysical journal,* **77,** 1609-1618.

Jentsch, T. J., Friedrich, T., Schriever, A. & Yamada, H. (1999). The CLC chloride channel family. *Pflügers Archiv,* **437,** 783-795.

Ji, Q., Hashmi, S., Liu, Z., Zhang, J., Chen, Y. & Huang, C.-H. (2006). CeRh1 (*rhr-1*) is a dominant Rhesus gene essential for embryonic development and hypodermal function in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences,* **103,** 5881-5886.

Johnson, T. E., Mitchell, D. H., Kline, S., Kemal, R. & Foy, J. (1984). Arresting development arrests aging in the nematode *Caenorhabditis elegans*. *Mechanisms of Ageing and Development,* **28,** 23-40.

Jones, A. M. & Assmann, S. M. (2004). Plants: the latest model system for G-protein research. *EMBO reports,* **5,** 572-578.

Jones, D., Taylor, W. & Thornton, J. (1994). A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry,* **33,** 3038-3049.

Jones, M. (1991). Orotidylate decarboxylase of yeast and man. *Current topics in cellular regulation,* **33,** 331-342.

Jungmann, J., Rayner, J. C. & Munro, S. (1999). The *Saccharomyces cerevisiae* protein Mnn10p/Bed1p is a subunit of a Golgi mannosyltransferase complex. *Journal of Biological Chemistry,* **274,** 6579-6585.

Kaeberlein, M., McVey, M. & Guarente, L. (1999). The *SIR2/3/4* complex and *SIR2* alone promote longevity in *Saccharomyces cerevisiae* by two different mechanisms. *Genes & Development,* **13,** 2570-2580.

Kaletta, T. & Hengartner, M. O. (2006). Finding function in novel targets: C. elegans as a model organism. *Nat Rev Drug Discov,* **5,** 387-399.

Käll, L., Krogh, A. & Sonnhammer, E. L. (2004). A combined transmembrane topology and signal peptide prediction method. *Journal of molecular biology,* **338,** 1027-1036.

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P. & Ahringer, J. (2003). Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature,* **421,** 231-237.

Kao, G., Nordenson, C., Still, M., Rönnlund, A., Tuck, S. & Naredi, P. (2007). ASNA-1 positively regulates insulin secretion in *C. elegans* and mammalian cells. *Cell,* **128,** 577-587.

Karathia, H., Vilaprinyo, E., Sorribas, A. & Alves, R. (2011). *Saccharomyces cerevisiae* as a model organism: a comparative study. *PloS one,* **6,** e16015.

Karimi, M., Inzé, D. & Depicker, A. (2002). GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation. *Trends in plant science,* **7,** 193-195.

Kass, J., Jacob, T. C., Kim, P. & Kaplan, J. M. (2001). The EGL-3 proprotein convertase regulates mechanosensory responses of *Caenorhabditis elegans*. *The Journal of Neuroscience,* **21,** 9265-9272.

Kawamoto, T., Noshiro, M., Shen, M., Nakamasu, K., Hashimoto, K., Kawashima-Ohya, Y., Gotoh, O. & Kato, Y. (1998). Structural and phylogenetic analyses of RGD-CAP/βig-h3, a fasciclin-like adhesion protein expressed in chick chondrocytes. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression,* **1395,** 288-292.

Kay, R., Chan, A., Daly, M. & Mcpherson, J. (1987). Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science,* **236,** 1299-1302.

Kellokumpu, S., Sormunen, R. & Kellokumpu, I. (2002). Abnormal glycosylation and altered Golgi structure in colorectal cancer: dependence on intra-Golgi pH. *FEBS letters,* **516,** 217-224.

Kemphues, K. 2005. Essential genes. *In:* Community, T. C. E. R. (ed.) *WormBook.* WormBook.

Kharenko, O. A., Choudhary, P. & Loewen, M. C. (2013). Abscisic acid binds to recombinant *Arabidopsis thaliana* G-protein coupled receptor-type G-protein 1 in *Sacaromycese cerevisiae* and *in vitro*. *Plant Physiology and Biochemistry,* **68,** 32-36.

Kim, S. A., Kwak, J., Jae, S.-K., Wang, M.-H. & Nam, H. (2001). Overexpression of the *AtGluR2* gene encoding an *Arabidopsis* homolog of mammalian glutamate receptors impairs calcium utilization and sensitivity to ionic stress in transgenic plants. *Plant and Cell Physiology,* **42,** 74-84.

Kimura, K. D., Tissenbaum, H. A., Liu, Y. & Ruvkun, G. (1997). *daf-2*, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. *Science,* **277,** 942-946.

Kitamura, Y. & Itoh, T. (1987). Reaction volume of protonic ionization for buffering agents. Prediction of pressure dependence of pH and pOH. *Journal of Solution Chemistry,* **16,** 715-725.

Klingler, J. P., Batelli, G. & Zhu, J.-K. (2010). ABA receptors: the START of a new paradigm in phytohormone signalling. *Journal of Experimental Botany*.

Kniazeva, M., Crawford, Q. T., Seiber, M., Wang, C.-Y. & Han, M. (2004). Monomethyl branched-chain fatty acids play an essential role in *Caenorhabditis elegans* development. *PLoS Biology,* **2,** e257.

Knight, M. R., Campbell, A. K., Smith, S. M. & Trewavas, A. J. (1991). Transgenic plant aequorin reports the effects of touch and cold-shock and elicitors on cytoplasmic calcium.

Kobae, Y., Uemura, T., Sato, M. H., Ohnishi, M., Mimura, T., Nakagawa, T. & Maeshima, M. (2004). Zinc transporter of *Arabidopsis thaliana* AtMTP1 is localized to vacuolar membranes and implicated in zinc homeostasis. *Plant and Cell Physiology,* **45,** 1749-1758.

Koelle, M. R. & Horvitz, H. R. (1996). EGL-10 regulates G protein signaling in the *C. elegans* nervous system and shares a conserved domain with many mammalian proteins. *Cell,* **84,** 115-125.

Kohlhaw, G. B. (1987). β-isopropylmalate dehydrogenase from yeast. *Methods in enzymology,* **166,** 429-435.

Kosinski, R. & Zaremba, M. (2007). Dynamics of the model of the *Caenorhabditis elegans* neural network. *Acta Physica Polonica B,* **38,** 2201.

Krogh, A., Larsson, B., Von Heijne, G. & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology,* **305,** 567-580.

Krysan, P. J., Young, J. C. & Sussman, M. R. (1999). T-DNA as an insertional mutagen in *Arabidopsis*. *The Plant Cell,* **11,** 2283-2290.

Krysan, P. J., Young, J. C., TAx, F. & Sussman, M. R. (1996). Identification of transferred DNA insertions within Arabidopsis genes involved in signal transduction and ion transport. *Proceedings of the National Academy of Sciences,* **93,** 8145-8150.

Kubo, T., Fukuda, K., Mikami, A., Maeda, A., Takahashi, H., Mishina, M., Haga, T., Haga, K., Ichiyama, A., Kangawa, K., Kojima, M., Matsuo, H., Hirose, T. & Numa, S. (1986). Cloning, sequencing and expression of complementary DNA encoding the muscarinic acetylcholine receptor. *Nature,* **323,** 411-416.

Kundu, S., Horn, P. J. & Peterson, C. L. (2007). SWI/SNF is required for transcriptional memory at the yeast *GAL* gene cluster. *Genes & Development,* **21,** 997-1004.

Kuroyanagi, H., Ohno, G., Sakane, H., Maruoka, H. & Hagiwara, M. (2010). Visualization and genetic analysis of alternative splicing regulation in vivo using fluorescence reporters in transgenic *Caenorhabditis elegans*. *Nat. Protocols,* **5,** 1495-1517.

Kurtzman, C. P. (2005). Molecular taxonomy of the yeasts. *Yeast,* **10,** 1727-1740.

Lacroute, F. (1968). Regulation of pyrimidine biosynthesis in *Saccharomyces cerevisiae*. *Journal of bacteriology,* **95,** 824-832.

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L. & Garcia-Hernandez, M. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic acids research,* **40,** D1202-D1210.

Larre, C., Penninck, S., Bouchet, B., Lollier, V., Tranquet, O., Denery-Papini, S., Guillon, F. & Rogniaux, H. (2010). Brachypodium distachyon grain: identification and subcellular localization of storage proteins. *Journal of experimental botany,* **61,** 1771-1783.

Le, B. H., Cheng, C., Bui, A. Q., Wagmaister, J. A., Henry, K. F., Pelletier, J., Kwong, L., Belmonte, M., Kirkbride, R., Horvath, S., Drews, G. N., Fischer, R. L., Okamuro, J. K., Harada, J. J. & Goldberg, R. B. (2010). Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proceedings of the National Academy of Sciences,* **107,** 8063-8070.

Lemmon, M., Flanagan, J. M., Hunt, J. F., Adair, B. D., Bormann, B., Dempsey, C. E. & Engelman, D. (1992). Glycophorin A dimerization is driven by specific interactions between transmembrane α-helices. *Journal of Biological Chemistry,* **267,** 7683-7689.

Lemmon, M. A. & Engelman, D. M. (1994). Specificity and promiscuity in membrane helix interactions. *Quarterly reviews of biophysics,* **27,** 157-218.

Lemmon, M. A., Treutlein, H. R., Adams, P. D., Brünger, A. T. & Engelman, D. M. (1994). A dimerization motif for transmembrane α–helices. *Nature Structural & Molecular Biology,* **1,** 157-163.

Leung, M. C. K., Williams, P. L., Benedetto, A., Au, C., Helmcke, K. J., Aschner, M. & Meyer, J. N. (2008). *Caenorhabditis elegans*: an emerging model in biomedical and environmental toxicology. *Toxicological Sciences,* **106,** 5-28.

Lewis, K. C., Selzer, T., Shahar, C., Udi, Y., Tworowski, D. & Sagi, I. (2008). Inhibition of pectin methyl esterase activity by green tea catechins. *Phytochemistry,* **69,** 2586-2592.

Li, H.-H., Hao, R.-L., Wu, S.-S., Guo, P.-C., Chen, C.-J., Pan, L.-P. & Ni, H. (2011). Occurrence, function and potential medicinal applications of the phytohormone abscisic acid in animals and humans. *Biochemical Pharmacology,* **82,** 701-712.

Li, J., Nagpal, P., Vitart, V., McMorris, T. C. & Chory, J. (1996). A role for brassinosteroids in light-dependent development of Arabidopsis. *Science,* **272,** 398-401.

Li, J., Nam, K. H., Vafeados, D. & Chory, J. (2001). *BIN2*, a new brassinosteroid-insensitive locus in *Arabidopsis*. *Plant Physiology,* **127,** 14-22.

Li, K.-B. (2003). ClustalW-MPI: ClustalW analysis using distributed and parallel computing. *Bioinformatics,* **19,** 1585-1586.

Li, S.-C. & Deber, C. M. (1992). Glycine and β-branched residues support and modulate peptide helicity in membrane environments. *FEBS letters,* **311,** 217-220.

Li, X., Romero, P., Rani, M., Dunker, A. K. & Obradovic, Z. (1999). Predicting protein disorder for N-, C-, and internal regions. *Genome Informatics Series***,** 30-40.

Libert, F., Vassart, G. & Parmentier, M. (1991). Current developments in G-protein-coupled receptors. *Current opinion in cell biology,* **3,** 218-223.

Lippincott-Schwartz, J., Snapp, E. & Kenworthy, A. (2001). Studying protein dynamics in living cells. *Nature Reviews Molecular Cell Biology,* **2,** 444-456.

Liu, Y., Engelman, D. M. & Gerstein, M. (2002). Genomic analysis of membrane protein families: abundance and conserved motifs. *Genome Biol,* **3,** 1-0054.0012.

Lochrie, M. A., Mendel, J. E., Sternberg, P. W. & Simon, M. I. (1991). Homologous and unique G protein α-subunits in the nematode *Caenorhabditis elegans*. *Cell Regulation,* **2,** 135-154.

Lohr, D., Venkov, P. & Zlatanova, J. (1995). Transcriptional regulation in the yeast *GAL* gene family: a complex genetic network. *The FASEB Journal,* **9,** 777-787.

Lu, Y., Chanroj, S., Zulkifli, L., Johnson, M. A., Uozumi, N., Cheung, A. & Sze, H. (2011). Pollen tubes lacking a pair of K^+ transporters fail to target ovules in *Arabidopsis*. *The Plant Cell Online,* **23,** 81-93.

Lyu, P. C., Sherman, J. C., Chen, A. & Kallenbach, N. R. (1991). α-helix stabilization by natural and unnatural amino acids with alkyl side chains. *Proceedings of the National Academy of Sciences,* **88,** 5317-5320.

Ma, Y., Szostkiewicz, I., Korte, A., Moes, D., Yang, Y., Christmann, A. & Grill, E. (2009). Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science,* **324,** 1064-1068.

Maeda, I., Kohara, Y., Yamamoto, M. & Sugimoto, A. (2001). Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. *Current Biology,* **11,** 171-176.

Maeda, Y., Ide, T., Koike, M., Uchiyama, Y. & Kinoshita, T. (2008). GPHR is a novel anion channel critical for acidification and functions of the Golgi apparatus. *Nat Cell Biol,* **10,** 1135-1145.

Maekawa, M., Inoue, T., Kobuna, H., Nishimura, T., Gengyo-Ando, K., Mitani, S. & Arai, H. (2009). Functional analysis of GS28, an intra-Golgi SNARE, in *Caenorhabditis elegans*. *Genes to Cells,* **14,** 1003-1013.

Magnone, M., Bruzzone, S., Guida, L., Damonte, G., Millo, E., Scarfì, S., Usai, C., Sturla, L., Palombo, D., De Flora, A. & Zocchi, E. (2009). Abscisic acid released by human monocytes activates monocytes and vascular smooth muscle cell responses involved in atherogenesis. *Journal of Biological Chemistry,* **284,** 17808-17818.

Mandava, N. B. (1988). Plant growth-promoting brassinosteroids. *Annual Review of Plant Physiology and Plant Molecular Biology,* **39,** 23-52.

Manders, E. M., Stap, J., Brakenhoff, G. J., van Driel, R. & Aten, J. A. (1992). Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy. *Journal of Cell Science,* **103,** 857-862.

Manders, E. M. M., Verbeek, F. J. & Aten, J. A. (1993). Measurement of co-localization of objects in dual-colour confocal images. *Journal of Microscopy,* **169,** 375-382.

Mango, S. E. 2007. The C. elegans pharynx: a model for organogenesis. *In:* Community, T. C. E. R. (ed.) *WormBook.* WormBook.

Maple, J. & Møller, S. G. 2007. Mutagenesis in Arabidopsis. *Circadian Rhythms.* Springer.

Marmagne, A., Vinauger-Douard, M., Monachello, D., de Longevialle, A. F., Charon, C., Allot, M., Rappaport, F., Wollman, F.-A., Barbier-Brygoo, H. & Ephritikhine, G. (2007). Two members of the Arabidopsis CLC (chloride channel) family, AtCLCe and AtCLCf, are associated with thylakoid and Golgi membranes, respectively. *Journal of Experimental Botany,* **58,** 3385-3393.

Marshansky, V., Rubinstein, J. L. & Grüber, G. (2014). Eukaryotic V-ATPase: Novel structural findings and functional insights. *Biochimica et Biophysica Acta (BBA)-Bioenergetics,* **1837,** 857-879.

Martinière, A., Bassil, E., Jublanc, E., Alcon, C., Reguera, M., Sentenac, H., Blumwald, E. & Paris, N. (2013). *In vivo* intracellular pH measurements in tobacco and *Arabidopsis* reveal an unexpected pH gradient in the endomembrane system. *The Plant Cell*

Mäser, P., Thomine, S., Schroeder, J. I., Ward, J. M., Hirschi, K., Sze, H., Talke, I. N., Amtmann, A., Maathuis, F. J. & Sanders, D. (2001). Phylogenetic relationships within cation transporter families of Arabidopsis. *Plant Physiology,* **126,** 1646-1667.

McGuire, A. T. & Mangroo, D. (2007). Cex1p is a novel cytoplasmic component of the *Saccharomyces cerevisiae* nuclear tRNA export machinery. *The EMBO journal,* **26,** 288-300.

McGuire, A. T. & Mangroo, D. (2012). Cex1p facilitates Rna1p-mediated dissociation of the Los1p-tRNA-Gsp1p-GTP export complex. *Traffic,* **13,** 234-256.

McKay, S. J., Johnsen, R., Khattra, J., Asano, J., Baillie, D. L., Chan, S., Dube, N., Fang, L., Goszczynski, B., Ha, E., Halfnight, E., Hollebakken, R., Huang, P., Hung, K., Jensen, V., Jones, S. J. M., Kai, H., Li, D., Mah, A., Marra, M., Mcghee, J., Newbury, R., Pouzyrev, A., Riddle, D. L., Sonnhammer, E., Tian, H., Tu, D., Tyson, J. R., Vatcher, G., Warner, A., Wong, K., Zhao, Z. & Moerman, D. G. (2003). Gene expression profiling of cells, tissues, and developmental stages of the nematode *C. elegans*. *Cold Spring Harbor Symposia on Quantitative Biology,* **68,** 159-170.

Mendel, J., Korswagen, H., Liu, K., Hajdu-Cronin, Y., Simon, M., Plasterk, R. & Sternberg, P. (1995). Participation of the protein Go in multiple aspects of behavior in C. elegans. *Science,* **267,** 1652-1655.

Menendez, D., Inga, A., Jordan, J. J. & Resnick, M. A. (2007). Changing the p53 master regulatory network: *ELEMENT*ary, my dear Mr Watson. *Oncogene,* **26,** 2191-2201.

Menguer, P. K., Farthing, E., Peaston, K. A., Ricachenevsky, F. K., Fett, J. P. & Williams, L. E. (2013). Functional analysis of the rice vacuolar zinc transporter OsMTP1. *Journal of Experimental Botany,* **64,** 2871-2883.

Michielse, C. B., Hooykaas, P. J., van den Hondel, C. A. & Ram, A. F. (2008). *Agrobacterium*-mediated transformation of the filamentous fungus *Aspergillus awamori*. *Nature protocols,* **3,** 1671-1678.

Mills, R. F., Doherty, M. L., López-Marqués, R. L., Weimar, T., Dupree, P., Palmgren, M. G., Pittman, J. K. & Williams, L. E. (2008). ECA3, a Golgi-localized P_{2A}-type ATPase, plays a crucial role in manganese nutrition in *Arabidopsis*. *Plant physiology,* **146,** 116-128.

Mitaku, S. & Hirokawa, T. (1999). Physicochemical factors for discriminating between soluble and membrane proteins: hydrophobicity of helical segments and protein length. *Protein engineering,* **12,** 953-957.

Mitrovich, Q. M. & Anderson, P. (2000). Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in C. elegans. *Genes & Development,* **14,** 2173-2184.

Montanini, B., Blaudez, D., Jeandroz, S., Sanders, D. & Chalot, M. (2007). Phylogenetic and functional analysis of the Cation Diffusion Facilitator (CDF) family: improved signature and prediction of substrate specificity. *BMC genomics,* **8,** 107.

Montgomery, M. K., Xu, S. & Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences,* **95,** 15502-15507.

Morley, J. F. & Morimoto, R. I. (2004). Regulation of longevity in *Caenorhabditis elegans* by heat shock factor and molecular chaperones. *Molecular Biology of the Cell,* **15,** 657-664.

Morris, A. J. & Malbon, C. C. (1999). Physiological regulation of G Protein-linked signaling. *Physiological Reviews,* **79,** 1373-1430.

Mount, D. W. (2004). Sequence and genome analysis. *Bioinformatics: Cold Spring Harbour Laboratory Press: Cold Spring Harbour,* **2**.

Müller, A. H. & Hansson, M. (2009). The Barley Magnesium Chelatase 150-kD subunit is not an abscisic acid receptor. *Plant Physiology,* **150,** 157-166.

Murashige, T. & Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia plantarum,* **15,** 473-497.

Murata, D., Nomura, K. H., Dejima, K., Mizuguchi, S., Kawasaki, N., Matsuishi-Nakajima, Y., Ito, S., Gengyo-Ando, K., Kage-Nakadai, E. & Mitani, S. (2012). GPI-anchor synthesis is indispensable for the germline development of the nematode *Caenorhabditis elegans*. *Molecular Biology of the Cell,* **23,** 982-995.

Nakai, K. & Kanehisa, M. (1991). Expert system for predicting protein localization sites in gram-negative bacteria. *Proteins: Structure, Function, and Bioinformatics,* **11,** 95-110.

Nathans, J. & Hogness, D. S. (1983). Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. *Cell,* **34,** 807-814.

Neiman, A. M., Mhaiskar, V., Manus, V., Galibert, F. & Dean, N. (1997). *Saccharomyces cerevisiae HOC1*, a suppressor of *pkc1*, encodes a putative glycosyltransferase. *Genetics,* **145,** 637-645.

Nelson, D. L., Lehninger, A. L. & Cox, M. M. 2008. *Lehninger principles of biochemistry*, Macmillan.

Nicholas, K. B. & Nicholas Jr, H. B. (1997). GeneDoc: a tool for editing and annotating multiple sequence alignments. *Distributed by the author,* **14**.

Niedenthal, R. K., Riles, L., Johnston, M. & Hegemann, J. H. (1996). Green fluorescent protein as a marker for gene expression and subcellular localization in budding yeast. *Yeast,* **12,** 773-786.

Niñoles, R., Rubio, L., García-Sánchez, M. J., Fernández, J. A., Bueso, E., Alejandro, S. & Serrano, R. (2013). A dominant-negative form of Arabidopsis AP-3 β-adaptin improves intracellular pH homeostasis. *The Plant Journal,* **74,** 557-568.

Nishi, T. & Forgac, M. (2002). The vacuolar (H^+)-ATPases — nature's most versatile proton pumps. *Nature Reviews Molecular Cell Biology,* **3,** 94-103.

Nishimura, N., Sarkeshik, A., Nito, K., Park, S.-Y., Wang, A., Carvalho, P. C., Lee, S., Caddell, D. F., Cutler, S. R., Chory, J., Yates, J. R. & Schroeder, J. I. (2010). PYR/PYL/RCAR family members are major *in vivo* ABI1 protein phosphatase 2C-interacting proteins in *Arabidopsis*. *The Plant Journal,* **61,** 290-299.

Noda, Y., Hara, T., Ishii, M. & Yoda, K. (2014). Distinct adaptor proteins assist exit of Kre2-family proteins from the yeast ER. *Biology Open,* **3,** 209-224.

Nordeen, M. H., Jones, S. M., Howell, K. E. & Caldwell, J. H. (2000). GOLAC: an endogenous anion channel of the Golgi complex. *Biophysical Journal,* **78,** 2918-2928.

North, M., Steffen, J., Loguinov, A. V., Zimmerman, G. R., Vulpe, C. D. & Eide, D. J. (2012). Genome-wide functional profiling identifies genes and processes important for zinc-limited growth of *Saccharomyces cerevisiae*. *PLoS Genetics,* **8,** e1002699.

Novick, P., Ferro, S. & Schekman, R. (1981). Order of events in the yeast secretory pathway. *Cell,* **25,** 461-469.

Novick, P., Field, C. & Schekman, R. (1980). Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. *Cell,* **21,** 205-215.
O'Rourke, S. M. & Herskowitz, I. (2004). Unique and redundant roles for HOG MAPK pathway components as revealed by whole-genome expression analysis. *Molecular Biology of the Cell,* **15,** 532-542.
O'Rourke, S. M. & Herskowitz, I. (1998). The Hog1 MAPK prevents cross talk between the HOG and pheromone response MAPK pathways in *Saccharomyces cerevisiae*. *Genes & Development,* **12,** 2874-2886.
Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., Brown, C. J. & Dunker, A. K. (2003). Predicting intrinsic disorder from amino acid sequence. *Proteins: Structure, Function, and Bioinformatics,* **53,** 566-572.
Odell, J. T., Nagy, F. & Chua, N.-H. (1985). Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter.
Offermanns, S. (2003). G-proteins as transducers in transmembrane signalling. *Progress in Biophysics and Molecular Biology,* **83,** 101-130.
Ohgaki, R., van Ijzendoorn, S. C. D., Matsushita, M., Hoekstra, D. & Kanazawa, H. (2010). Organellar Na^+/H^+ exchangers: novel players in organelle pH regulation and their emerging functions. *Biochemistry,* **50,** 443-450.
Ohno, G., Hagiwara, M. & Kuroyanagi, H. (2008). STAR family RNA-binding protein ASD-2 regulates developmental switching of mutually exclusive alternative splicing *in vivo*. *Genes & Development,* **22,** 360-374.
Oliva, M. & Dunand, C. (2007). Waving and skewing: how gravity and the surface of growth media affect root development in Arabidopsis. *New Phytologist,* **176,** 37-43.
Omasits, U., Ahrens, C. H., Müller, S. & Wollscheid, B. (2013). Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics*, btt607.
Ooms, G., Hooykaas, P. J., Van Veen, R. J., Van Beelen, P., Regensburg-Tuïnk, T. J. & Schilperoort, R. A. (1982). Octopine Ti-plasmid deletion mutants of *Agrobacterium tumefaciens* with emphasis on the right side of the T-region. *Plasmid,* **7,** 15-29.
Ooms, G., Klapwijk, P., Poulis, J. & Schilperoort, R. (1980). Characterization of Tn904 insertions in octopine Ti plasmid mutants of *Agrobacterium tumefaciens*. *Journal of bacteriology,* **144,** 82-91.
Østergaard, L. & Yanofsky, M. F. (2004). Establishing gene function by mutagenesis in *Arabidopsis thaliana*. *The Plant Journal,* **39,** 682-696.
Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., Thibaud-Nissen, F., Malek, R. L., Lee, Y., Zheng, L., Orvis, J., Haas, B., Wortman, J. & Buell, C. R. (2007). The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Research,* **35,** D883-D887.
Ovchinnikov, Y. A. (1982). Rhodopsin and bacteriorhodopsin: structure—function relationships. *FEBS Letters,* **148,** 179-191.
Pak, J. & Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science,* **315,** 241-244.
Palmgren, M. G. & Nissen, P. (2011). P-Type ATPases. *Annual Review of Biophysics,* **40,** 243-266.
Pandey, S. & Assmann, S. M. (2004). The *Arabidopsis* putative G protein-coupled receptor GCR1 interacts with the G Protein α-ubunit GPA1 and regulates abscisic acid signaling. *The Plant Cell,* **16,** 1616-1632.
Pandey, S., Nelson, D. C. & Assmann, S. M. (2009). Two novel GPCR-type G proteins are abscisic acid receptors in *Arabidopsis*. *Cell,* **136,** 136-148.
Parinov, S. & Sundaresan, V. (2000). Functional genomics in *Arabidopsis*: large-scale insertional mutagenesis complements the genome sequencing project. *Current Opinion in Biotechnology,* **11,** 157-161.
Park, M.-Y., Chung, M.-S., Koh, H.-S., Lee, D. J., Ahn, S.-J. & Kim, C. S. (2009a). Isolation and functional characterization of the *Arabidopsis salt-tolerance 32* (*AtSAT32*)

gene associated with salt tolerance and ABA signaling. *Physiologia Plantarum,* **135,** 426-435.

Park, S.-Y., Fung, P., Nishimura, N., Jensen, D. R., Fujii, H., Zhao, Y., Lumba, S., Santiago, J., Rodrigues, A., Chow, T.-f. F., Alfred, S. E., Bonetta, D., Finkelstein, R., Provart, N. J., Desveaux, D., Rodriguez, P. L., McCourt, P., Zhu, J.-K., Schroeder, J. I., Volkman, B. F. & Cutler, S. R. (2009b). Abscisic Acid Inhibits Type 2C Protein Phosphatases via the PYR/PYL Family of START Proteins. *Science,* **324,** 1068-1071.

Parker, G. S., Eckert, D. M. & Bass, B. L. (2006). RDE-4 preferentially binds long dsRNA and its dimerization is necessary for cleavage of dsRNA to siRNA. *RNA,* **12,** 807-818.

Paroutis, P., Touret, N. & Grinstein, S. (2004). The pH of the secretory pathway: measurement, determinants, and regulation. *Physiology,* **19,** 207-215.

Parrish, S., Fleenor, J., Xu, S., Mello, C. & Fire, A. (2000). Functional anatomy of a dsRNA trigger: differential requirement for the two trigger strands in RNA interference. *Molecular Cell,* **6,** 1077-1087.

Paschinger, K., Hackl, M., Gutternigg, M., Kretschmer-Lubich, D., Stemmer, U., Jantsch, V., Lochnit, G. & Wilson, I. B. H. (2006). A deletion in the Golgi α-mannosidase II gene of *Caenorhabditis elegans* results in unexpected non-wild-type *N*-glycan structures. *Journal of Biological Chemistry,* **281,** 28265-28277.

Patel, N., Hamamouch, N., Li, C., Hewezi, T., Hussey, R. S., Baum, T. J., Mitchum, M. G. & Davis, E. L. (2010). A nematode effector protein similar to annexins in host plants. *Journal of Experimental Botany,* **61,** 235-248.

Patterson, G. I. & Padgett, R. W. (2000). TGFβ-related pathways: roles in *Caenorhabditis elegans* development. *Trends in Genetics,* **16,** 27-33.

Pedersen, P. A., Rasmussen, J. H. & Jørgensen, P. L. (1996). Expression in high yield of pig $\alpha 1\beta 1$ Na, K-ATPase and inactive mutants D369N and D807N in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry,* **271,** 2514-2522.

Pedersen, S. E. & Ross, E. M. (1982). Functional reconstitution of β-adrenergic receptors and the stimulatory GTP-binding protein of adenylate cyclase. *Proceedings of the National Academy of Sciences of the United States of America,* **79,** 7228-7232.

Persson, B. 2000. Bioinformatics in protein analysis. *Proteomics in Functional Genomics*. Springer.

Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods,* **8,** 785-786.

Pevzner, P. 2000. *Computational molecular biology: an algorithmic approach*, MIT press.

Pfeffer, S. R. & Rothman, J. E. (1987). Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. *Annual review of biochemistry,* **56,** 829-852.

Piette, J., Decuyper-Debergh, D. & Gamper, H. (1985). Mutagenesis of the *lac* promoter region in M13 mp10 phage DNA by 4′ -hydroxymethyl-4,5′ ,8-t rimethylpsoralen. *Proceedings of the National Academy of Sciences of the United States of America,* **82,** 7355-7359.

Piper, P. W. (2006). Long-lived yeast as a model for ageing research. *Yeast,* **23,** 215-226.

Popov, M., Tam, L. Y., Li, J. & Reithmeier, R. A. (1997). Mapping the ends of transmembrane segments in a polytopic membrane protein scanning *N*-glycosylation mutagenesis of extracytosolic loops in the anion exchanger, band 3. *Journal of Biological Chemistry,* **272,** 18325-18332.

Posas, F., Wurgler-Murphy, S. M., Maeda, T., Witten, E. A., Thai, T. C. & Saito, H. (1996). Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1–YPD1–SSK1 "two-component" osmosensor. *Cell,* **86,** 865-875.

Project, I. R. G. S. (2005). The map-based sequence of the rice genome. *Nature,* **436,** 793-800.

Protocols, C. S. H. (2006). Complete Minimal (CM) or Synthetic Complete (SC) and Drop-out Media. *Cold Spring Harbor Protocols,* **2006,** pdb.rec8190.

Puce, S., Basile, G., Bavestrello, G., Bruzzone, S., Cerrano, C., Giovine, M., Arillo, A. & Zocchi, E. (2004). Abscisic acid signaling through cyclic ADP-ribose in hydroid regeneration. *Journal of Biological Chemistry,* **279,** 39783-39788.

Puri, S., Bachert, C., Fimmel, C. J. & Linstedt, A. D. (2002). Cycling of early Golgi proteins via the cell surface and endosomes upon lumenal pH disruption. *Traffic,* **3,** 641-653.

Qi, B., Doughty, J. & Hooley, R. (2013). A Golgi and tonoplast localized S-acyl transferase is involved in cell expansion, cell division, vascular patterning and fertility in Arabidopsis. *New Phytologist,* **200,** 444-456.

Qian, W., Ma, D., Xiao, C., Wang, Z. & Zhang, J. (2012). The Genomic Landscape and Evolutionary Resolution of Antagonistic Pleiotropy in Yeast. *Cell Reports,* **2,** 1399-1410.

Raizen, D., Song, B.-m., Trojanowski, N. & You, Y.-J. 2012. Methods for measuring pharyngeal behaviors. *In:* Community, T. C. E. R. (ed.) *WormBook.* WormBook.

Raizen, D. M., Lee, R. Y. & Avery, L. (1995). Interacting genes required for pharyngeal excitation by motor neuron MC in *Caenorhabditis elegans*. *Genetics,* **141,** 1365-1382.

Rand, T. A., Petersen, S., Du, F. & Wang, X. (2005). Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell,* **123,** 621-629.

Rasmussen, S. G. F., DeVree, B. T., Zou, Y., Kruse, A. C., Chung, K. Y., Kobilka, T. S., Thian, F. S., Chae, P. S., Pardon, E., Calinski, D., Mathiesen, J. M., Shah, S. T. A., Lyons, J. A., Caffrey, M., Gellman, S. H., Steyaert, J., Skiniotis, G., Weis, W. I., Sunahara, R. K. & Kobilka, B. K. (2011). Crystal structure of the β_2 adrenergic receptor-Gs protein complex. *Nature,* **477,** 549-555.

Raz, V., Bergervoet, J. H. & Koornneef, M. (2001). Sequential steps for developmental arrest in Arabidopsis seeds. *Development,* **128,** 243-252.

Razem, F. A., El-Kereamy, A., Abrams, S. R. & Hill, R. D. (2006). The RNA-binding protein FCA is an abscisic acid receptor. *Nature,* **439,** 290-294.

Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. & Jackson, R. B. 2011. *Campbell biology*, Pearson Boston.

Restrepo-Hartwig, M. & Ahlquist, P. (1999). Brome Mosaic Virus RNA replication proteins 1a and 2a colocalize and 1a independently localizes on the yeast endoplasmic reticulum. *Journal of Virology,* **73,** 10303-10309.

Reumann, S. (2004). Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses. *Plant physiology,* **135,** 783-800.

Rhee, S. Y., Beavis, W., Berardini, T. Z., Chen, G., Dixon, D., Doyle, A., Garcia-Hernandez, M., Huala, E., Lander, G. & Montoya, M. (2003). The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic acids research,* **31,** 224-228.

Riddle, D. L., Blumenthal, T., Meyer, B. J. & Priess, J. R. 1997. *Feeding and Defecation - C. elegans II*, Cold Spring Harbor Laboratory Press.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. & Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nature biotechnology,* **17,** 1030-1032.

Risk, J. M., Day, C. L. & Macknight, R. C. (2009). Reevaluation of abscisic acid-binding assays shows that G-protein-coupled receptor 2 does not bind abscisic acid. *Plant Physiology,* **150,** 6-11.

Rivinoja, A., Hassinen, A., Kokkonen, N., Kauppila, A. & Kellokumpu, S. (2009). Elevated Golgi pH impairs terminal *N*-glycosylation by inducing mislocalization of Golgi glycosyltransferases. *Journal of cellular physiology,* **220,** 144-154.

Rivinoja, A., Kokkonen, N., Kellokumpu, I. & Kellokumpu, S. (2006). Elevated Golgi pH in breast and colorectal cancer cells correlates with the expression of oncofetal carbohydrate T antigen. *Journal of cellular physiology,* **208,** 167-174.

Rivinoja, A., Pujol, F. M., Hassinen, A. & Kellokumpu, S. (2011). Golgi pH, its regulation and roles in human disease. *Annals of Medicine,* **0,** 1-13.

Rolls, M. M., Hall, D. H., Victor, M., Stelzer, E. H. K. & Rapoport, T. A. (2002). Targeting of rough endoplasmic reticulum membrane proteins and ribosomes in invertebrate neurons. *Molecular Biology of the Cell,* **13,** 1778-1791.

Romero, P., Obradovic, Z., Kissinger, C., Villafranca, J. & Dunker, A. Identifying disordered regions in proteins from amino acid sequence. Neural Networks, 1997., International Conference on, 1997. IEEE, 90-95.

Rosenbaum, D. M., Rasmussen, S. G. & Kobilka, B. K. (2009). The structure and function of G-protein-coupled receptors. *Nature,* **459,** 356-363.

Ross, E. M. & Gilman, A. G. (1980). Biochemical properties of hormone-sensitive adenylate cyclase. *Annual Review of Biochemistry,* **49,** 533-564.

Rost, B. (1996). [31] PHD: Predicting one-dimensional protein structure by profile-based neural networks. *Methods in enzymology,* **266,** 525-539.

Roussel, N., Morton, C. A., Finger, F. P. & Roysam, B. (2007). A computational model for C. elegans locomotory behavior: application to multiworm tracking. *Biomedical Engineering, IEEE Transactions on,* **54,** 1786-1797.

Rual, J.-F., Ceron, J., Koreth, J., Hao, T., Nicot, A.-S., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S. H., Hill, D. E., van den Heuvel, S. & Vidal, M. (2004). Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based rnai library. *Genome Research,* **14,** 2162-2168.

Russ, W. P. & Engelman, D. M. (2000). The GxxxG motif: a framework for transmembrane helix-helix association. *Journal of molecular biology,* **296,** 911-919.

Rutherford, R. & Masson, P. H. (1996). *Arabidopsis thaliana sku* mutant seedlings show exaggerated surface-dependent alteration in root growth vector. *Plant physiology,* **111,** 987-998.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution,* **4,** 406-425.

Santiago, J., Dupeux, F., Round, A., Antoni, R., Park, S.-Y., Jamin, M., Cutler, S. R., Rodriguez, P. L. & Marquez, J. A. (2009). The abscisic acid receptor PYR1 in complex with abscisic acid. *Nature,* **462,** 665-668.

Santner, A. & Estelle, M. (2009). Recent advances and emerging trends in plant hormone signalling. *Nature,* **459,** 1071-1078.

Scarfì, S., Fresia, C., Ferraris, C., Bruzzone, S., Fruscione, F., Usai, C., Benvenuto, F., Magnone, M., Podestà, M., Sturla, L., Guida, L., Albanesi, E., Damonte, G., Salis, A., De Flora, A. & Zocchi, E. (2009). The plant hormone abscisic acid stimulates the proliferation of human hemopoietic progenitors through the second messenger cyclic adp-ribose. *Stem cells,* **27,** 2469-2477.

Schafer, W. R. 2005. Egg-laying. *In:* Community, T. C. E. R. (ed.) *WormBook.* WormBook.

Scharff-Poulsen, P. & Pedersen, P. (2013). *Saccharomyces cerevisiae*–based platform for rapid production and evaluation of eukaryotic nutrient.

Scheel, O., Zdebik, A. A., Lourdel, S. & Jentsch, T. J. (2005). Voltage-dependent electrogenic chloride/proton exchange by endosomal CLC proteins. *Nature,* **436,** 424-427.

Schlessinger, A., Liu, J. & Rost, B. (2007a). Natively unstructured loops differ from other loops. *PLoS computational biology,* **3,** e140.

Schlessinger, A., Punta, M. & Rost, B. (2007b). Natively unstructured regions in proteins identified from contact predictions. *Bioinformatics,* **23,** 2376-2384.

Schlessinger, A., Punta, M., Yachdav, G., Kajan, L. & Rost, B. (2009). Improved disorder prediction by combination of orthogonal approaches. *PLoS One,* **4,** e4433.

Schlessinger, A., Yachdav, G. & Rost, B. (2006). PROFbval: predict flexible and rigid residues in proteins. *Bioinformatics,* **22,** 891-893.

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D. & Lohmann, J. U. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nature genetics,* **37,** 501-506.

Schmidt, W. K., Tam, A., Fujimura-Kamada, K. & Michaelis, S. (1998). Endoplasmic reticulum membrane localization of Rce1p and Ste24p, yeast proteases involved in carboxyl-terminal CAAX protein processing and amino-terminal a-factor cleavage. *Proceedings of the National Academy of Sciences,* **95,** 11175-11180.

Schriever, A. M., Friedrich, T., Pusch, M. & Jentsch, T. J. (1999). CLC chloride channels in *Caenorhabditis elegans*. *Journal of Biological Chemistry,* **274,** 34238-34244.

Schröder, F., Lisso, J., Lange, P. & Müssig, C. (2009). The extracellular EXO protein mediates cell expansion in Arabidopsis leaves. *BMC plant biology,* **9,** 20.

Schüller, C., Brewster, J., Alexander, M., Gustin, M. & Ruis, H. (1994). The HOG pathway controls osmotic regulation of transcription via the stress response element (STRE) of the *Saccharomyces cerevisiae* CTT1 gene. *The EMBO journal,* **13,** 4382.

Schulze-Gahmen, U., De Bondt, H. L. & Kim, S.-H. (1996). High-resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: bound waters and natural ligand as guides for inhibitor design. *Journal of medicinal chemistry,* **39,** 4540-4546.

Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., Frommer, W. B., Flügge, U.-I. & Kunze, R. (2003). ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. *Plant Physiology,* **131,** 16-26.

Schwarz, E. M., Antoshechkin, I., Bastiani, C., Bieri, T., Blasiar, D., Canaran, P., Chan, J., Chen, N., Chen, W. J., Davis, P., Fiedler, T. J., Girard, L., Harris, T. W., Kenny, E. E., Kishore, R., Lawson, D., Lee, R., Müller, H.-M., Nakamura, C., Ozersky, P., Petcherski, A., Rogers, A., Spooner, W., Tuli, M. A., Van Auken, K., Wang, D., Durbin, R., Spieth, J., Stein, L. D. & Sternberg, P. W. (2006). WormBase: better software, richer content. *Nucleic Acids Research,* **34,** D475-D478.

Schwikowski, B., Uetz, P. & Fields, S. (2000). A network of protein-protein interactions in yeast. *Nat Biotech,* **18,** 1257-1261.

Segalat, L., Elkes, D. & Kaplan, J. (1995). Modulation of serotonin-controlled behaviors by Go in *Caenorhabditis elegans*. *Science,* **267,** 1648-1651.

Sen, G. L. & Blau, H. M. (2005). Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nature cell biology,* **7,** 633-636.

Sentenac, H., Bonneaud, N., Minet, M., Lacroute, F., Salmon, J.-M., Gaymard, F. & Grignon, C. (1992). Cloning and expression in yeast of a plant potassium ion transport system. *Science,* **256,** 663-665.

Seymour, M. K., Wright, K. A. & Doncaster, C. C. (1983). The action of the anterior feeding apparatus of *Caenorhabditis elegans* (Nematoda: *Rhabditida*). *Journal of Zoology,* **201,** 527-539.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E. & Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnology,* **22,** 1567-1572.

Sharma Chandra, B., Knauer, R. & Lehle, L. 2001. Biosynthesis of lipid-linked oligosaccharides in yeast: the *ALG3* gene encodes the dol-P-Man:Man5GlcNAc2-PP-Dol Mannosyltransferase. *Biological Chemistry.*

Shaw, S. L. (2006). Imaging the live plant cell. *The Plant Journal,* **45,** 573-598.

Shen, J., Zeng, Y., Zhuang, X., Sun, L., Yao, X., Pimpl, P. & Jiang, L. (2013). Organelle pH in the Arabidopsis Endomembrane System. *Molecular Plant,* **6,** 1419-1437.

Shen, Y., Xie, J., Liu, R.-d., Ni, X.-f., Wang, X.-h., Li, Z.-x. & Zhang, M. (2014). Genomic analysis and expression investigation of caleosin gene family in Arabidopsis. *Biochemical and biophysical research communications,* **448,** 365-371.

Shewry, P. (2009). Wheat. *Journal of experimental Botany,* **60,** 1537-1553.

Shima, J. E., Komori, T., Taylor, T. R., Stryke, D., Kawamoto, M., Johns, S. J., Carlson, E. J., Ferrin, T. E. & Giacomini, K. M. (2010). Genetic variants of human organic anion transporter 4 demonstrate altered transport of endogenous substrates. *American Journal of Physiology-Renal Physiology,* **299,** F767-F775.

Shirley, B. W., Hanley, S. & Goodman, H. M. (1992). Effects of ionizing radiation on a plant genome: analysis of two Arabidopsis transparent testa mutations. *The Plant Cell Online,* **4,** 333-347.

Sigrist, C. J. A., de Castro, E., Cerutti, L., Cuche, B. A., Hulo, N., Bridge, A., Bougueleret, L. & Xenarios, I. (2012). New and continuing developments at PROSITE. *Nucleic Acids Research*.

Sijen, T., Steiner, F. A., Thijssen, K. L. & Plasterk, R. H. A. (2007). Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science,* **315,** 244-247.

Simmons, C., Söll, D. & Migliaccio, F. (1995). Circumnutation and gravitropism cause root waving in *Arabidopsis thaliana*. *Journal of Experimental Botany,* **46,** 143-150.

Simpson, A. G. & Roger, A. J. (2004). The real 'kingdoms' of eukaryotes. *Current Biology,* **14,** R693-R696.

Slack, F. & Ruvkun, G. (1997). Temporal pattern formation by heterochronic genes. *Annual Review of Genetics,* **31,** 611-634.

Sladek, F. M., Melian, A. & Howard-Flanders, P. (1989). Incision by UvrABC exonuclease is a step in the path to mutagenesis by psoralen crosslinks in *Escherichia coli*. *Proceedings of the National Academy of Sciences,* **86,** 3982-3986.

Snape, J. & Pánková, K. (2007). *Triticum aestivum* (wheat). *Encyclopedia of Life Sciences*.

Snider, C., Jayasinghe, S., Hristova, K. & White, S. H. (2009). MPEx: a tool for exploring membrane proteins. *Protein Science,* **18,** 2624-2628.

Sonnhammer, E. L., Von Heijne, G. & Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. Ismb, 1998. 175-182.

Stajich, J. E., Dietrich, F. S. & Roy, S. W. (2007). Comparative genomic analysis of fungal genomes reveals intron-rich ancestors. *Genome Biol,* **8,** R223.

Steber, C. M. & McCourt, P. (2001). A role for brassinosteroids in germination in Arabidopsis. *Plant Physiology,* **125,** 763-769.

Steiner, S. & Philippsen, P. (1994). Sequence and promoter analysis of the highly expressed *TEF* gene of the filamentous fungus *Ashbya gossypii*. *Molecular and General Genetics* **242,** 263-271.

Stenger, J. E. & Gregory, S. G. (2006). Genomics and bioinformatics. *Genetic Analysis of Complex Diseases, Second Edition***,** 423-454.

Strand, F. L. 1999. *Neuropeptides: regulators of physiological processes*, MIT press.

Sturla, L., Fresia, C., Guida, L., Bruzzone, S., Scarfì, S., Usai, C., Fruscione, F., Magnone, M., Millo, E., Basile, G., Grozio, A., Jacchetti, E., Allegretti, M., De Flora, A. & Zocchi, E. (2009). LANCL2 is necessary for abscisic acid binding and signaling in human granulocytes and in rat insulinoma cells. *Journal of Biological Chemistry,* **284,** 28045-28057.

Sturla, L., Fresia, C., Guida, L., Grozio, A., Vigliarolo, T., Mannino, E., Millo, E., Bagnasco, L., Bruzzone, S., De Flora, A. & Zocchi, E. (2011). Binding of abscisic acid to human LANCL2. *Biochemical and Biophysical Research Communications,* **415,** 390-395.

Sulston, J., Dew, M. & Brenner, S. (1975). Dopaminergic neurons in the nematode *Caenorhabditis elegans*. *The Journal of Comparative Neurology,* **163,** 215-226.

Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Developmental Biology,* **100,** 64-119.

Sussman, M. R., Amasino, R. M., Young, J. C., Krysan, P. J. & Austin-Phillips, S. (2000). The *Arabidopsis* Knockout Facility at the University of Wisconsin–Madison. *Plant Physiology,* **124,** 1465-1467.

Suzuki, T., Sakamoto, T., Sugiyama, M., Ishida, N., Kambe, H., Obata, S., Kaneko, Y., Takahashi, H. & Harashima, S. (2013). Disruption of multiple genes whose deletion causes lactic-acid resistance improves lactic-acid resistance and productivity in *Saccharomyces cerevisiae*. *Journal of Bioscience and Bioengineering,* **115,** 467-474.

Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T. Z., Garcia-Hernandez, M., Foerster, H., Li, D., Meyer, T., Muller, R. & Ploetz, L. (2008). The Arabidopsis Information Resource (TAIR): gene structure and function annotation. *Nucleic acids research,* **36,** D1009-D1014.

Sze, J. Y., Victor, M., Loer, C., Shi, Y. & Ruvkun, G. (2000). Food and metabolic signalling defects in a *Caenorhabditis elegans* serotonin-synthesis mutant. *Nature,* **403,** 560-564.

Szekeres, M., Németh, K., Koncz-Kálmán, Z., Mathur, J., Kauschmann, A., Altmann, T., Rédei, G. P., Nagy, F., Schell, J. & Koncz, C. (1996). Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450, controlling cell elongation and de-etiolation in Arabidopsis. *Cell,* **85,** 171-182.

Takeda, S., Kadowaki, S., Haga, T., Takaesu, H. & Mitaku, S. (2002). Identification of G protein-coupled receptor genes from the human genome sequence. *FEBS letters,* **520,** 97-101.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution,* **30,** 2725-2729.

Tartakoff, A., Vassalli, P. & Detraz, M. (1978). Comparative studies of intracellular transport of secretory proteins. *The Journal of cell biology,* **79,** 694-707.

Tarutani, M., Nakajima, K., Uchida, Y., Takaishi, M., Goto-Inoue, N., Ikawa, M., Setou, M., Kinoshita, T., Elias, P. M., Sano, S. & Maeda, Y. (2012). GPHR-dependent functions of the Golgi apparatus are essential for the formation of lamellar granules and the skin barrier. *Journal of Investigative Dermatology,* **132,** 2019-2025.

Temple, B. R. S. & Jones, A. M. (2007). The plant heterotrimeric G-protein complex. *Annual Review of Plant Biology,* **58,** 249-266.

Thacker, C. & Rose, A. M. (2000). A look at the *Caenorhabditis elegans* Kex2/Subtilisin-like proprotein convertase family. *BioEssays,* **22,** 545-553.

Thomas, J. H. (1993). Thinking about genetic redundancy. *Trends in Genetics,* **9,** 395-399.

Thompson, J. & Gibson, T. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics*.

Thompson, J., Higgins, D. G. & Gibson, T. J. (1994). ClustalW. *Nucleic Acids Res,* **22,** 4673-4680.

Thompson, J. D., Gibson, T. & Higgins, D. G. (2002a). Multiple sequence alignment using ClustalW and ClustalX. *Current protocols in bioinformatics***,** 2.3. 1-2.3. 22.

Thompson, J. S., Snow, M. L., Giles, S., McPherson, L. E. & Grunstein, M. (2003). Identification of a functional domain within the essential core of histone H3 that is required for telomeric and HM silencing in *Saccharomyces cerevisiae*. *Genetics,* **163,** 447-452.

Thompson, R. J., Nordeen, M. H., Howell, K. E. & Caldwell, J. H. (2002b). A large-conductance anion channel of the Golgi complex. *Biophysical Journal,* **83,** 278-289.

Timmons, L., Court, D. L. & Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene,* **263,** 103-112.

Timmons, L. & Fire, A. (1998). Specific interference by ingested dsRNA. *Nature,* **395,** 854-854.

Timson, D. J. (2007). Galactose metabolism in *Saccharomyces cerevisiae*. *Dynamic Biochemistry, Process Biotechnology and Molecular Biology,* **1,** 63-73.

Tossi, V., Cassia, R., Bruzzone, S., Zocchi, E. & Lamattina, L. (2012). ABA says NO to UV-B: a universal response? *Trends in Plant Science,* **17,** 510-517.

Treusch, S., Knuth, S., Slaugenhaupt, S. A., Goldin, E., Grant, B. D. & Fares, H. (2004). *Caenorhabditis elegans* functional orthologue of human protein h-mucolipin-1 is required for lysosome biogenesis. *Proceedings of the National Academy of Sciences of the United States of America,* **101,** 4483-4488.

Tsujimoto, Y. & Shimizu, S. (2000). VDAC regulation by the Bcl-2 family of proteins. *Cell death and differentiation,* **7,** 1174-1181.

Tsuzuki, T., Takahashi, K., Inoue, S.-i., Okigaki, Y., Tomiyama, M., Hossain, M., Shimazaki, K.-i., Murata, Y. & Kinoshita, T. (2011). Mg-chelatase H subunit affects ABA signaling in stomatal guard cells, but is not an ABA receptor in *Arabidopsis thaliana*. *Journal of Plant Research,* **124,** 527-538.

Tun, N. M., O'Doherty, P. J., Perrone, G. G., Bailey, T. D., Kersaitis, C. & Wu, M. J. (2013). Disulfide stress-induced aluminium toxicity: molecular insights through genome-wide screening of *Saccharomyces cerevisiae*. *Metallomics,* **5,** 1068-1075.

Uemura, T., Suda, Y., Ueda, T. & Nakano, A. (2014). Dynamic behavior of the trans-Golgi network in root tissues of *Arabidopsis* revealed by super-resolution live imaging. *Plant and Cell Physiology,* **55,** 694-703.

Umezawa, T., Nakashima, K., Miyakawa, T., Kuromori, T., Tanokura, M., Shinozaki, K. & Yamaguchi-Shinozaki, K. (2010). Molecular basis of the core regulatory network in ABA Responses: sensing, signaling and transport. *Plant and Cell Physiology,* **51,** 1821-1839.

Umezu, K., Amaya, T., Yoshimoto, A. & Tomita, K. (1971). Purification and properties of orotidine-5′-phosphate pyrophosphorylase and orotidine-5′-phosphate decarboxylase from bakers' yeast. *Journal of Biochemistry,* **70,** 249-262.

Vahisalu, T., Kollist, H., Wang, Y.-F., Nishimura, N., Chan, W.-Y., Valerio, G., Lamminmäki, A., Brosché, M., Moldau, H. & Desikan, R. (2008). SLAC1 is required for plant guard cell S-type anion channel function in stomatal signalling. *Nature,* **452,** 487-491.

van Geest, M. & Lolkema, J. S. (2000). Membrane topology and insertion of membrane proteins: search for topogenic signals. *Microbiology and Molecular Biology Reviews,* **64,** 13-33.

van Zanten, M., Koini, M. A., Geyer, R., Liu, Y., Brambilla, V., Bartels, D., Koornneef, M., Fransz, P. & Soppe, W. J. J. (2011). Seed maturation in *Arabidopsis thaliana* is characterized by nuclear size reduction and increased chromatin condensation. *Proceedings of the National Academy of Sciences*.

Villalba, J., Palmgren, M., Berberian, G., Ferguson, C. & Serrano, R. (1992). Functional expression of plant plasma membrane H^+-ATPase in yeast endoplasmic reticulum. *Journal of Biological Chemistry,* **267,** 12341-12349.

Visintin, R., Prinz, S. & Amon, A. (1997). *CDC20* and *CDH1*: a family of substrate-specific activators of APC-dependent proteolysis. *Science,* **278,** 460-463.

Vivian-Smith, A. & Koltunow, A. M. (1999). Genetic analysis of growth-regulator-induced parthenocarpy in *Arabidopsis*. *Plant Physiology,* **121,** 437-452.

von Ehrenstein G., a. S. E. 1980. *Nematodes as biological models*, Academic Press, New York

Von Heijne, G. (1992). Membrane protein structure prediction: hydrophobicity analysis and the positive-inside rule. *Journal of molecular biology,* **225,** 487-494.

Vowels, J. J. & Thomas, J. H. (1992). Genetic analysis of chemosensory control of dauer formation in *Caenorhabditis elegans*. *Genetics,* **130,** 105-123.

Wach, A., Brachat, A., Pöhlmann, R. & Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast,* **10,** 1793-1808.

Waggoner, L. E., Zhou, G. T., Schafer, R. W. & Schafer, W. R. (1998). Control of alternative behavioral states by serotonin in *Caenorhabditis elegans*. *Neuron,* **21,** 203-214.

Wain, H. M., Bruford, E. A., Lovering, R. C., Lush, M. J., Wright, M. W. & Povey, S. (2002). Guidelines for human gene nomenclature. *Genomics,* **79,** 464-470.

Walker, D. S., Gower, N. J., Ly, S., Bradley, G. L. & Baylis, H. A. (2002). Regulated disruption of inositol 1,4,5-trisphosphate signaling in *Caenorhabditis elegans* reveals new functions in feeding and embryogenesis. *Mol Biol Cell,* **13,** 1329-1337.

Wang, H., Spang, A., Sullivan, M. A., Hryhorenko, J. & Hagen, F. K. (2005). The terminal phase of cytokinesis in the *Caenorhabditis elegans* early embryo requires protein glycosylation. *Molecular Biology of the Cell,* **16,** 4202-4213.

Wang, Y., Chen, Z.-H., Zhang, B., Hills, A. & Blatt, M. R. (2013). PYR/PYL/RCAR abscisic acid receptors Regulate K^+ and Cl^- channels through reactive oxygen species-mediated activation of Ca^{2+} channels at the plasma membrane of intact *Arabidopsis* guard cells. *Plant Physiology,* **163,** 566-577.

Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F. & Jones, D. T. (2004a). The DISOPRED server for the prediction of protein disorder. *Bioinformatics,* **20,** 2138-2139.

Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. & Jones, D. T. (2004b). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *Journal of molecular biology,* **337,** 635-645.

Ward, J. M. (2001). Identification of novel families of membrane proteins from the model plant *Arabidopsis thaliana*. *Bioinformatics,* **17,** 560-563.

Ward, S. & Carrel, J. S. (1979). Fertilization and sperm competition in the nematode *Caenorhabditis elegans*. *Developmental Biology,* **73,** 304-321.

Waterhouse, P. M., Graham, M. W. & Wang, M.-B. (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proceedings of the National Academy of Sciences,* **95,** 13959-13964.

Webb, A. A. R., Larman, M. G., Montgomery, L. T., Taylor, J. E. & Hetherington, A. M. (2001). The role of calcium in ABA-induced gene expression and stomatal movements. *The Plant Journal,* **26,** 351-362.

Wege, S., Jossier, M., Filleur, S., Thomine, S., Barbier-Brygoo, H., Gambale, F. & De Angeli, A. (2010). The proline 160 in the selectivity filter of the *Arabidopsis* NO_3^-/H^+ exchanger AtCLCa is essential for nitrate accumulation in planta. *The Plant Journal,* **63,** 861-869.

Weiner, J. J., Peterson, F. C., Volkman, B. F. & Cutler, S. R. (2010). Structural and functional insights into core ABA signaling. *Current Opinion in Plant Biology,* **13,** 495-502.

Weiss, S. (1999). Fluorescence spectroscopy of single biomolecules. *Science,* **283,** 1676-1683.

Weisz, O. A. (2003). Organelle acidification and disease. *Traffic,* **4,** 57-64.

Werck-Reichhart, D. & Feyereisen, R. (2000). Cytochromes P450: a success story. *Genome Biol,* **1,** 3003.3001-3003.3009.

White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences,* **314,** 1-340.

Whittaker, R. H. (1969). New concepts of kingdoms of organisms. *Science,* **163,** 150-160.

Willard, F. S., Kimple, A. J., Johnston, C. A. & Siderovski, D. P. (2005). A direct fluorescence-based assay for RGS domain GTPase accelerating activity. *Analytical biochemistry,* **340,** 341-351.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V. & Provart, N. J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PloS one,* **2,** e718.

Wojas, S., Clemens, S., Hennig, J., Skłodowska, A., Kopera, E., Schat, H., Bal, W. & Antosiewicz, D. M. (2008). Overexpression of phytochelatin synthase in tobacco: distinctive effects of *AtPCS1* and *CePCS* genes on plant response to cadmium. *Journal of Experimental Botany,* **59,** 2205-2219.

Wojas, S., Clemens, S., SkLodowska, A. & Maria Antosiewicz, D. (2010). Arsenic response of *AtPCS1*- and *CePCS*-expressing plants - effects of external As(V) concentration on As-accumulation pattern and NPT metabolism. *Journal of Plant Physiology,* **167,** 169-175.

Wolf, S., Mravec, J., Greiner, S., Mouille, G. & Höfte, H. (2012). Plant cell wall homeostasis is mediated by brassinosteroid feedback signaling. *Current Biology,* **22,** 1732-1737.

Wood, W. B. 1987. *The nematode Caenorhabditis elegans*, Cold Spring Harbour Laboratory.

Wood, W. B. (1988). The nematode *Caenorhabditis elegans*.

Wooding, S. & Pelham, H. R. B. (1998). The dynamics of Golgi protein traffic visualized in living yeast cells. *Molecular Biology of the Cell,* **9,** 2667-2680.

Wu, F.-Q., Xin, Q., Cao, Z., Liu, Z.-Q., Du, S.-Y., Mei, C., Zhao, C.-X., Wang, X.-F., Shang, Y., Jiang, T., Zhang, X.-F., Yan, L., Zhao, R., Cui, Z.-N., Liu, R., Sun, H.-L., Yang, X.-L., Su, Z. & Zhang, D.-P. (2009). The Magnesium-Chelatase H subunit binds abscisic acid and functions in abscisic acid signaling: new evidence in *Arabidopsis*. *Plant Physiology,* **150,** 1940-1954.

Wu, J., Tisa, L. S. & Rosen, B. P. (1992). Membrane topology of the ArsB protein, the membrane subunit of an anion-translocating ATPase. *Journal of Biological Chemistry,* **267,** 12570-12576.

Xiong, X., Xu, D., Yang, Z., Huang, H. & Cui, X. (2013). A single amino-acid substitution at lysine 40 of an *Arabidopsis thaliana* α-tubulin causes extensive cell proliferation and expansion defects. *Journal of Integrative Plant Biology,* **55,** 209-220.

Yachdav, G., Kloppmann, E., Kajan, L., Hecht, M., Goldberg, T., Hamp, T., Hönigschmid, P., Schafferhans, A., Roos, M. & Bernhofer, M. (2014). PredictProtein—an open resource for online prediction of protein structural and functional features. *Nucleic acids research***,** gku366.

Yamagata, S., Isaji, M., Nakamura, K., Fujisaki, S., Doi, K., Bawden, S. & D'Andrea, R. (1994). Overexpression of the *Saccharomyces cerevisiae MET17/MET25* gene in *Escherichia coli* and comparative characterization of the product with *O*-acetylserine· *O*-acetylhomoserine sulfhydrylase of the yeast. *Applied Microbiology and Biotechnology,* **42,** 92-99.

Yandell, M. D., Edgar, L. G. & Wood, W. B. (1994). Trimethylpsoralen induces small deletion mutations in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences,* **91,** 1381-1385.

Yin, P., Fan, H., Hao, Q., Yuan, X., Wu, D., Pang, Y., Yan, C., Li, W., Wang, J. & Yan, N. (2009). Structural insights into the mechanism of abscisic acid signaling by PYL proteins. *Nat Struct Mol Biol,* **16,** 1230-1236.

Yuste, R. (2005). Fluorescence microscopy today. *Nat Meth,* **2,** 902-904.

Zeiger, L. T. a. E. 2006. *Plant Physiology*, Sinaurer Associates, Inc.

Zeyl, C. (2000). Budding yeast as a model organism for population genetics. *Yeast,* **16,** 773-784.

Zhang, G. F. & Staehelin, L. A. (1992). Functional Compartmentation of the Golgi Apparatus of Plant Cells Immunocytochemical Analysis of High-Pressure Frozen-and Freeze-Substituted Sycamore Maple Suspension Culture Cells. *Plant physiology,* **99,** 1070-1083.

Zhang, W., Cao, P., Chen, S., SPENCE, A., Zhu, S., Staudacher, E. & Schachter, H. (2003). Synthesis of paucimannose *N*-glycans by *Caenorhabditis elegans* requires prior actions of UDP-*N*-acetyl-D-glucosamine: *α*-3-D-mannoside *β*1, 2-*N*-acetylglucosaminyltransferase I, *α*3, 6-mannosidase II and a specific membrane-bound *β*-*N*-acetylglucosaminidase. *Biochemal Journal,* **372,** 53-64.

Zhang, Y., Foster, J. M., Nelson, L. S., Ma, D. & Carlow, C. K. S. (2005). The chitin synthase genes *chs-1* and *chs-2* are essential for *C. elegans* development and responsible for chitin deposition in the eggshell and pharynx, respectively. *Developmental Biology,* **285,** 330-339.

Zheng, S., Pan, T., Fan, L. & Qiu, Q.-S. (2013). A novel *AtKEA* gene family, homolog of bacterial K^+ antiporters, plays potential roles in K^+ homeostasis and osmotic adjustment in *Arabidopsis*. *PLoS ONE,* **8,** e81463.

Zhu, J.-K. (2002). Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology,* **53,** 247-273.

Zhu, J., Hill, R. J., Heid, P. J., Fukuyama, M., Sugimoto, A., Priess, J. R. & Rothman, J. H. (1997). *end-1* encodes an apparent GATA factor that specifies the endoderm precursor in *Caenorhabditis elegans* embryos. *Genes & Development,* **11,** 2883-2896.

Zuckerkandl, E. & Pauling, L. (1965). Evolutionary divergence and convergence in proteins. *Evolving genes and proteins,* **97,** 97-166.

Appendix 1

***GTG/GPHR* gene and accession numbers used for phylogenetic analysis**

Species	Gene name	Accession numbers
Acromyrmex echinatior	G5I_00539	EGI70744.1
Acyrthosiphon pisum	GPHR-like	XP_003246954.1
Aedes aegypti	AaeL_AAEL012912	XP_001663047.1
Ailuropoda melanoleuca	GPR89-like	XP_002924599.1
Amblyomma maculatum	G3MPD1_9ACAR	G3MPD1
Amphimedon queenslandica	GPHR-like	XP_003388791.1
Anolis carolinensis	LOC100561234	XP_003219134.1
Anopheles gambiae str. PEST	AGAP011252-PA	XP_309392.4
Apis mellifera	LOC100576864	XP_623554.3
Arabidopsis thaliana	GTG1	NP_176679.2
Arabidopsis thaliana	GTG2	NP_194493.2
Arthrobotrys oligospora	AOL_s00043g768	EGX51749.1
Ascaris suum	GPHR	ADY44017.1
Ashbya gossypii	AGR366Wp	NP_987032.1
Atta cephalotes	H9HUV8_ATTCE	H9HUV8
Batrachochytrium dendrobatidis	BATDEDRAFT_11317	XP_006678841.1
Bos taurus	GPR89	DAA31613.1
Botryotinia fuckeliana	BofuT4_P052760.1	CCD44927.1
Branchiostoma floridae	BRAFLDRAFT_114456	XP_002596694.1
Brachypodium distachyon	GTG1-like	XP_003580421.1
Caenorhabditis brenneri	CAEBREN_26068	EGT45845.1
Caenorhabditis brenneri	CAEBREN_28311	EGT50374.1
Caenorhabditis briggsae	CBG23048	XP_002643118.1
Caenorhabditis briggsae	CBG07249	XP_002645634.1
Caenorhabditis elegans	Y75B8A.16	NP_499588
Caenorhabditis elegans	C11H1.2	NP_510384.2
Caenorhabditis remanei	CRE_24463	XP_003105005.1
Caenorhabditis remanei	CRE_20820	XP_003099973.1
Caligus rogercresseyi	GPR89	ACO11331.1
Callithrix jacchus	GPR89	XP_002759868.1
Camponotus floridanus	GPR89A	EFN72882.1
Candida albicans	Ca019.8068	XP_715923
Candida dubliniensis	Hypothetical protein	XP_002417168.1
Candida tropicalis	Hypothetical protein	XP_002550215.1
Candida glabrata	Hypothetical protein	XP_448174
Candida tenuis	CANTEDRAFT_95811	XP_006689571.1
Canis lupus familiaris	similar to GPR89	XP_849576.1
Capsaspora owczarzaki	GPR89	EFW47716.1
Chaetomium thermophilum	CTHI_0026900	XP_006693148.1
Chlamydomonas reinhardtii	CHLREDRAFT_118641	XP_001695894.1
Chlorella variabilis	CHLNCDRAFT_144997	XP_005848513.1

Clavispora lusitaniae	CLUG_03374	XP_002616133.1
Clonorchis sinensis	CLF_104919	GAA50695.1
Coccomyxa subellipsoidea	COCSUDRAFT_36939	XP_005646724.1
Collectotrichum higginsianum	CH063_07716	CCF36066.1
Colletotrichum graminicola	GLRG_00274	EFQ25130
Cricetulus griseus	GPHR	B2ZXD5.1
Cryptococcus gattii	CGB_D2690W	XP_003193437.1
Cryptococcus neoformans	CNL04370	XP_567988
Culex quinquefasciatus	GPR89A	XP_001869215.1
Danio rerio	GPHR	XP_002663360.1
Dictyostelium discoideum	DDB_G0283855	XP_638846.1
Dictyostelium purpureum	DICPUDRAFT_9193	XP_003290531.1
Drosophila ananassae	GF12626	XP_001958928.1
Drosophila erecta	GG22371	XP_001975543.1
Drosophila grimshawi	GH20796	XP_001985978.1
Drosophila melanogaster	CG8090	NP_611016.2
Drosophila mojavensis	GI20881	XP_002006150.1
Drosophila persimilis	GL11475	XP_002016222.1
Drosophila pseudoobscura	GA20816	XP_001360643.2
Drosophila simulans	GD25634	XP_002081593.1
Drosophila virilis	GJ20616	XP_002049692.1
Drosophila willistoni	GK19445	XP_002062917.1
Drosophila yakuba	GE14878	XP_002086938.1
Drosophila yakuba	GE12259	XP_002091403.1
Entamoeba dispar	EDI_324210	XP_001735348.1
Entamoeba histolytica	EHI_196900	XP_650054.2
Equus caballus	LOC100065612	XP_001499135.1
Eremothecium cymbalariae	Ecym_8167	XP_003648270.1
Exophilia dermatitidis	HMPREF1120_07012	EHY59012.1
Fusarium oxysporum	FOXB_09030	EGU80473.1
Gallus gallus	GPHR	NP_001025962.1
Gasterosteus aculeatus	G3NHC7_GASAC	G3NHC7
Gibberella zeae	FG02573.1	XP_382749.1
Glomerella graminicola	GLRG_00274	EFQ25130.1
Glycine max	GTG-like	XP_003537827.1
Grosmannia clavigera	CMQ_2343	EFX02294.1
Harpegnathos saltator	GPR89A	EFN84497.1
Heterocephalus glaber	GPHR-like	XP_004890518.1
Homo sapiens	GPR89B	NM_016334
Hypocrea virens	TRIVIDRAFT_160103	EHK17618.1
Kazachstania africana	KAFT_0D04750	XP_003957257.1
Kluyveromyces lactis	Hypothetical protein	XP_453651.1
Laccaria bicolour	GPR89	XP_001880078.1
Lachancea thermotolerans	KLTH0E03894p	XP_002553651.1
Latimeria chalumnae	GPHR-like	XP_006011657.1
Leptosphaeria maculans	KLLA0D13200p	CBX90990.1

Loa loa	LOAG_02394	XP_003137980.1
Lodderomyces elongisporus	LELG_01754	XP_001526926.1
Loxodonta Africana	G3T923_LOXAF	G3T923
Ichthyophthirius multifiliis	IMG5_000860	XP_004040162.1
Macaca fascicularis	EGM_01022	EHH50229.1
Macaca mulatta	GPR89B	NP_001244365.1
Magnaporthe oryzae	MGG_05269	XP_359508.1
Medicago truncatula	MTR_5g005440	XP_003610649
Meleagris gallopavo	GPHR-like	XP_003202886.1
Metarhizium acridum	MAC_09753	EFY84201.1
Metarhizium robertsii	MAA_08899	XP_007825088.1
Meyerozyma guilliermondii	PGUG_05521	EDK41423.2
Micromonas pusilla	MICPUCDRAFT_35156	XP_003061350.1
Micromonas sp. RCC299	MICPUN_109606	XP_002506636.1
Mixia osmundae	EQ5_05822	GAA99132.1
Monodelphis domestica	GPHR-like	XP_001368967.1
Monosiga brevicollis	33399	XP_001747836.1
Mus musculus	GPR89	NP_080505
Myceliophthora thermophila	MYCTH_2304629	XP_003663142.1
Mycosphaerella graminicola	GPCR	EGP88931.1
Myotis luciugus	G1P120_MYOLU	G1P120
Nasonia vitripennis	LOC100117322	XP_001601598.1
Naumovozyma castellii	NCAS_0H01730	XP_003677831.1
Naumovozyma dairenensis	NDAI_0C01900	XP_003669093.1
Nectria haematococca	NECHADRAFT_104801	XP_003045249.1
Nematostella vectensis	v1g192049	XP_001626503.1
Neospora caninum	GA20816	XP_003881652.1
Neurospora crassa	NCU00005	XP_956992.1
Neurospora tetrasperma	NEUTE1DRAFT_81495	EGO57694.1
Nomascus leucogenys	GPHR	XP_003268129.1
Oikopleura dioica	GSOID_T00014877001	CBY22955.1
Ornithorhynchus anatinus	LOC100091874	XP_001513405.1
Oryctolagus cuniculus (rabbit)	GPR89B-like	XP_002715696.1
Oryza glaberrima	I1PPE4_ORYGL	I1PPE4
Oryza sativa Japonica Group	OSJNBa0083N12.15	CAD41818.2
Oryzias latipes	LOC101175278	XP_004082021.1
Ostreococcus lucimarinus	OSTLU_89082	ABO99511.1
Ostreococcus tauri	Ot15g00790	XP_003083223.1
Otolemur garnettii	GPHR-like	XP_003793838
Pan troglodytes	GPHRA	NP_001267457.1
Paramecium tetrurelia	GSPATT00029617001	XP_001426384.1
Phaeosphaeria nodorum	SNOG_08079	XP_001798406.1
Pediculus humanus corporis	Phum_PHUM247980	XP_002426313.1
Phytophthora infestans	PITG_07946	XP_002904152.1
Phytophthora ramorum	H3GVB3_PHYRM	H3GVB3
Picea glauca [a]	-	BT114622

Pichia pastoris	PAS_chr3_0427	XP_002492656.1
Pichia sorbitophila	Piso0_004286	XP_004197050.1
Pongo abelii	GPR89-like	XP_002810373.1
Populus trichocarpa	POPTR_0013s07530g	XP_002319243.1
Pyrenophoras teres	PTT_07610	XP_003297272.1
Pyrenophora tritici-repentis	PTRG_08650	XP_001938982.1
Phytophthora sojae	PHYSODRAFT_294857	EGZ29884.1
Rattus norvegicus	GPR89B	NP_001132958.1
Ricinus communis	RCOM_1328090	XP_002520085.1
Saccharomyces cerevisiae	YHR078W	NP_011945.1
Salmo salar	GPHR	B5X1G3.1
Salpingoeca ATCC 50818	PST_06740	XP_004992137.1
Sarcophilus harrisii	G3X243_SARHA	G3X243
Scheffersomyces stipitis	YHN8	XP_001387864.2
Schizophyllum commune	SCHODRAFT_60212	XP_003028479.1
Sclerotinia scleotiorum	SS1g_00556	XP_001598467
Scylla paramamosain	GPR89	ADG60269.1
Selaginella moellendorffii	SELMODRAFT_235594	XP_002988748.1
Solenopsis invicta	SINV_02252	EFZ15867.1
Sordaria macrospora k-hell	SMAC_00097	XP_003351556.1
Sorghum bicolour	SORBIDRAFT_06g027460	XP_002448458.1
Sporisorium reilianum	sr13158	CBQ72450
Sus scrofa	F1SDD4_PIG	F1SDD4
Takifugu rubripes	LOC101076333	XP_003961758.1
Toxoplasma gondii	TGME49_086490	XP_002369275.1
Taeniopygia guttata	GPR89B	XP_002194601.1
Tetrahymena thermophila	TTHERM_00190920	XP_001016669.2
Tetrapisispora phaffii	TPHA_0H01460	XP_003686786.1
Thielavia terrestris	THITE_2119146	XP_003655453.1
Torulaspora delbrueckii	TDEL_0E02630	XP_003681717.1
Tribolium castaneum	GLEAN_10568	XP_967110.1
Trichinella spiralis	Tsp_11055	XP_003381955.1
Trichoderma atroviride	TRIATDRAFT_93659	EHK48643.1
Trichoderma reesei	TRIREDRAFT_107503	EGR48806.1
Trichoderma virens	TRIVIDRAFT_160103	EHK17618.1
Trichoplax adhaerens	TRIADDRAFT_23402	XP_002112186.1
Triticum aestivum [b]	-	-
Triticum urartu	TRIUR3_24949	EMS53699.1
Trypanosoma congolense	TCIL3000_8_1320	CCC91916
Trypanosoma cruzi	Tc00.1047053508973.20	XP_811705.1
Trypanosoma vivax	TVY486_0800970	CCC49489.1
Tuber melanosporum	GSTUM_00004510001	XP_002835982.1
Ustilago maydis	UM06270.1	XP_762417.1
Verticillium dahliae	VDAG_10279	EGY20719.1
Vitis vinifera	VIT_16s0039g01460	XP_002270494.1
Volvox carteri f. nagariensis	VOLCADRAFT_70166	XP_002958986.1

Xenopus (Silurana) tropicalis	GPR89B	XP_002940512.1
Xenopus laevis	GPHR	NP_001086744.1
Yarrowia lipolytica	YALI0B16478p	XP_500975
Zea mays	LOC100272676	NP_001140605.1
Zygosaccharomyces rouxii	ZYRO0D10538p	XP_002496893.1

[a] Protein prediction from mRNA accession.

[b] Obtained from contigs (see section 3.2.1)