UNIVERSITY OF SOUTHAMPTON

FACULTY OF LIFE SCIENCES

School of Biological Sciences

Harnessing the Genetic Diversity of Watercess (*Rorippa nasturtium-aquaticum*) for Improved Morphology and Anti-cancer Benefits: Underpinning Data for Molecular Breeding.

by

Adrienne Payne, BSc.

Thesis for the degree of Doctor of Philosophy

August 2011

UNIVERSITY OF SOUTHAMPTON

Abstract

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCESCENTRE FOR BIOLOGICAL SCIENCES

Doctor of Philosophy

HARNESSING THE GENETIC DIVERSITY OF WATERCESS (*RORIPPA NASTURTIUM-AQUATICUM*) FOR IMPROVED MORPHOLOGY AND ANTI-CANCER BENEFITS: UNDERPINNING DATA FOR MOLECULAR BREEDING

By Adrienne Payne BSc

Watercress (*Rorippa nasturtium-aquaticum*) is a member of the *Brassicaceae* family. Increasing interest in healthy diets has turned the focus to watercress since it is known to contain one of the highest concentrations of the beneficial antioxidant phenethyl isothiocyanate (PEITC). This has exciting possibilities to breed a higher quality crop with increased concentrations of this key antioxidant providing a meachanism to aid in the prevention of cancer. Dwarfism is another desirable characteristic for many agricultural crops. A crop with a reduced stem length produces a stable increased yield and is easier to harvest as well as package. Dwarf mutants have been extensively characterized in many plant species and endogenous phytochrome Gibberellin (GA) is one of the several genes associated with the dwarf phenotype. The three candidate genes of particular interest are: GA2ox, GA3ox and GA20ox. GA20ox and GA3ox are involved in GA biosynthesis whilst GA2ox is involved in GA catabolism.

The aim of my research was to establish a collection of watercress from around the world and breed watercress that not only has a reduced stem length but is also nutritionally beneficial therefore breed an 'ideal' watercress cultivar. Initial screening of the germplasm shows significant variability in stem length (ranging from 7-17cm), stem diameter (ranging from 1.58-3.18mm) and antioxidant concentrations (ranging from 80-140 mmol Fe^{2+} equivalent per gram fresh weight). The antioxidant concentrations were assessed using an adapted antioxidant assay Ferric Reducing Ability of Plasma (FRAP) which revealed a ranking order for the watercress lines. Two lines with high FRAP values, Wx_0033 (143.84 mmol Fe^{2+} equivalent per gram fresh weight) and Wx_0011 (126.44 mmol Fe^{2+} equivalent per gram fresh weight), and one line with a low FRAP value, Wx_0038 (87.49 mmol Fe^{2+} equivalent per gram fresh weight), and also a control line Wx_0001 (130.66 mmol Fe^{2+} equivalent per gram fresh weight) was taken forward for further glucosinolate and isothiocyanate analysis.The lines did indeed vary in the concentration of glucosinolate and isothiocyanate. Wx_0001 had both the highest concentration of phenethyl glucosinolate, 19.35 µmoles per gram weight and phenethyl isothiocyanate, 0.20mg/ml phenethyl isothiocyanate.

The genetic diversity of the watercress lines held within the collection were assessed using the Amplified Fragment Length Polymorphism (AFLP) technique revealing a higher level of variation within (76%) each line than between (24%). Variation in gene expression was analysed using microarrays and verified using real time PCR. From the various data collected and recorded from the watercress germplasm collection this has lead to a breeding programme successfully being initiated at the University of Southampton.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

DECLARATION OF AUTHORSHIP

I, ADRIENNE PAYNE

declare that the thesis entitled

'HARNESSING THE GENETIC DIVERSITY OF WATERCRESS (*RORIPPA NASTURTIUM-AQUATICUM*) FOR IMPROVED ANTI-CANCER AND MORPHOLOGOCAL BENEFITS UNDERPINNING DATA FOR MOLECULAR BREEDING'.

and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- 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;
- where I have consulted the published work of others, this is always clearly attributed;
- 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;
- I have acknowledged all main sources of help;
- 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;
- none of this work has been published before submission

Signed: ……..

Date:……

ACKNOWLEDGEMENTS

I would like to thank Professor Gail Taylor and Dr Graham Clarkson for all their ongoing advice. A special thank you to the Plants and Environment laboratory members for their help and support and also a thank you to Dr Steve Rothwell. I would like to thank HRI Warwick for contribution of seeds to the germplasm collection and my funding bodies University of Southampton and Vitacress Salads Ltd.

I appreciate the collaboration with Breeze Cavell at Southampton General Hospital for bioassay. Also the collaboration with Melanie Dixon in the Environmental Engineering, University of Southampton and Breeze Cavell discussing PEITC detection. A special thanks to Dr John Langley in the Chemistry department, University of Southampton, for testing samples in the GC-MS apparatus and advice given.

A very big thank you to Dr John Rossiter at Imperial College London for allowing me to work in his laboratory and taking the time to show me High Performance Liquid Chromatography (HPLC) and Gas Chromatography Mass Spectrometry (GC/MS) with the end result data for the concentrations of glucosinolate and isothiocynate.

I would like to thank the National Arabidopsis Stock Centre (NASC) for hybridisation of the watercress DNA to the X species chip and RNA to the *Arabidopsis* ATH1 genechip.

I would like to thank Matthew Nelson for his assistance in recording the morphological variation of watercress grown in different growth media. Also like to thank Suzanne Milner for assistance with initial germplasm assessments.

I would also like to thank Dr Jennifer DeWoody, Suzanne Milner, Angela Mills and Stacey Travers for assistance with the grinding of the different watercress lines and Mike Cotton for maintenance of the glasshouse and growth-rooms. I would like to thank Bill Summers, B&W Florida, for advice given for watercress crossing.

A special thank you to the final year undergradraduates; Katie Finn for her work on the RAPDs protocol and Alice Mazzer for her work on the comparison of the FRAP and ORAC assays.

I am grateful for the help provided by Dr Mark Totten at the field site and also with planting out the crosses and assistance at Imperial College London for glucosinolate and isothiocyanate measurement.

I am grateful to Dr Maud Viger, Hazel Smith, Suzanne Milner, Dr Jennifer DeWoody and Dr Mark Totten for proof reading. I am extremely grateful to the help and assistance provided by Suzanne Milner with AFLP analysis.

I would like to thank members of the office; Dr Maud Viger, Yunan Lin, Suzanne Milner and Dr Mark Totten for support and advice throughout.

Final thanks to my Mum, Dad and my brother, Brett, for all their encouragement and support through my further studies and believing in me.

ABBREVIATIONS

3-MSP = 3-methylsufinylpropyl

4-MSB = 4-methylsufinylbutyl

ABI = Applied Biosystems

ABTS = 2,2'azino-bis(3-ethylbenzthiazone-6-sulphonic acid)

ACT2 = Actin 2

AFLP = Amplified Fragment Length Polymorphism

AITC = allyl-isothiocyanate

Ala = Alanine

ANOVA= Analysis of variance

APx = Ascorbate peroxidase

AtGID1 = *Arabidopsis* gibberellin gene

ATH1 = *Arabidopsis* genome array

BITC = benzyl isothiocyanate

BR = Brassinosteroid

BSA = Bovine Serum Albumin

CAT = Catalase

CDF = Chip Description File

cDNA = Complementary Deoxyribonucleic Acid

CEL = Channel Expression Levels

CHISAM = Chloroform/Isoamyl alcohol

CO_2 = Carbon dioxide

CTAB = Hexadecyltrimethylammoniumbromid

CWR = Crop Wild Relative

DEAE – ion exchange = DiEthylAminoEthyl

DAF = Deoxyribonucleic Acid Amplification Fingerprinting

DCM = Dichloromethane

dCTP = deoxycytidine-5'-triphosphate

DELLA = amino acid sequence

DGAT = diacylglycerol

DIM = di(3'-indolyl)methane

DMEM = Dulbecco's Modified Eagle's Medium

DMF = Dimethylformamide

DNA = Deoxyribonucleic acid

dNTP = deoxyribonucleotide triphosphate

DPPH = 2,2, diphenyl-1-picryhydrazyl

DPPH – H = reduced DPPH

*d*wf2 = dwarf 2 mutant

EcoRI = Esherichia coli Restriction I

EGFR = Epidermal Growth Factor Receptor

E_{GOI} = Efficiency of Gene of Interest

EMS = Ethyl methane sulphonate

E_{REF} = Efficiency of Reference Gene

EST = Express Sequence Tag

FAM = Blue fluorescent probe

Fe^{2+} = Iron ion

FRAP = Ferric Reducing Ability of Plasma / Ferric Reducing Antioxidant Power

FW = Fresh Weight

g = grams

gDNA = genomic Deoxyribonucleic Acid

GA = Gibberellic Acid

GAI = Giberrellin Insensitive

GAPDH = Glyceraldehyde 3-phosphate dehydrogenase

GC/MS = Gas Chromatography Mass Spectrometry

GENALEX = Genetic analysis in Excel

GPx = Glutathione perioxidase

GRAS = repressor of gibberellin signalling (GAI, RGA, SCR)

GSL-ELONG – side chain elongation locus

H_2O = Water

H_2O_2 = Hydrogen Peroxide

HAT = Hydrogen Atom Transfer

HCL = Hydrochloric Acid

HEX = Green fluorescent probe

HIF = Hypoxia Inducible Factor

HPLC = High Performance Liquid Chromatography

HRI = Horticultural Research International

I3C = Indole-3-Carbinol

IC50 = Inhibitory Concentration 50

IL = Introgression Lines

Ile = Isoleucine

ITCs = Isothiocyanates

KNOX = KNOTTED 1-LIKE HOMEOBOX gene

Leu = Leucine

M = Molar

MAPK = Mitogen-Activated Protein Kinase

MCF-7 = Michigan Cancer Foundation-7 breast cancer cell line

Met = Methionine

MDA-MB-231 = breast cancer cell line

ml = millilitre

mm = millimeter

MM = Mismatch

mMol/mmol = milli-molar

MMP = Matrix Metalloproteinase

mRNA = Messenger Ribonucleic Acid

MseI = Micrococus species

MT = inner salt

NASC = National Arabidopsis Stock Centre

NCBI = National Centre for Biotechnology Information

nm = nanometer

NNK = 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone

NO = Nitric Oxide

OD = Optical Density

ORAC = Oxygen Radical Absorbance Capacity

OsGID1 = Rice gibberellin gene

OSH1 = *Oryza sativa* homeobox 1

OTC = Over The Counter

PBS = Phosphate Buffer Saline

PCA = Principle Co-ordinate Analysis

PCR = Polymerase Chain Reaction

PEITC = Phenethyl Isothiocyanate

PG = Prostaglandins

Phe = Phenylanine

PM = Perfect Match

QC = Quality Control

qRT-PCR = quantitative PCR

QTL = Quantitative Trait Loci

RAPD-PCR = Random Amplified Polymorphic Deoxyribonucleic Acid Polymerase Chain Reaction

RFLP = Restriction Fragment Length Polymorphism

RI = Recombinant Inbred

RGA = Reppressor of Giberrellic Acid 1-3

RMA = Robust Multichip Average

RNA = Ribonucleic Acid

rpm = revolutions per minute

RPMI = Roswell Park Memorial Institute 1640 medium

SAM = Shoot Apical Meristem

SCF = Skp1-cullin-F-box

SE = Standard Error

SET = Single Electron Transport

SFN = Sulforaphane

SH = sulfhydryl

SHI = SHORT INTERNODES

SLR1 = Slender Rice 1

SLN1 = SLENDER

SNE = SNEEZY

SNP = Single Nucleotide Polymorphism

SOD = Super Oxide Dismutase

SPY = *SPINDLY*

SSR = Simple Sequence Repeat

SYBR green = green qRT-PCR fluorescent dye

T-RFLP = Terminal Restriction Fragment Length Polymorphism

TE buffer = Tris-HCL and EDTA

TEAC = Trolox Equivalent Antioxidant Capacity

TPTZ = (2,4,6-tri(2-pyridyl))-1,3,5-triazine

TRAP = Total Radical Absorbance Potential

Trp = Tryptophan

Tyr = Tyrosine

U = Unit

UPGMA = Unweighted Pair Group Method with Arithmetic mean

Val = Valine

WYSV = Watercress Yellow Spot Virus

ZR = ZYMO Research

CHAPTER ONE: GENERAL INTRODUCTION

1.1 Introduction

Watercress is a popular vegetable accompaniment and salad leaf to many dishes and recent sales have increased due to the finding that watercress contains one of the highest concentrations of a precursor for the beneficial antioxidant phenethyl isothiocyanate (PEITC). This recent finding has exciting possibilities to breed a higher quality crop to meet consumers' demands and the possibility to enhance the concentration of this key antioxidant. The long term over-arching aim of this research is to breed for watercress that is not only aesthetically pleasing to the consumer but also nutritionally beneficial, therefore breed an 'ideal watercress' cultivar. The objectives of my PhD were to provide under pinning information of phenotypic (morphological and biochemical) and genetic variation.

1.1.1 Watercress: Classification, Cultivars and Botanical Description

Watercress (*Rorippa nasturtium-aquaticum*, formerly known as *Nasturtium officinale*) (Figure 1.1) belongs to the Brassicaceae (Cruciferae) family (Palaniswamy and McAvoy. 2001)

Figure 1.1: Watercress (*Rorippa nasturtium-aquaticum*)
Perennial, glabrous, dark green to purple. Stems decumbent below, rooting at the nodes, branched. Rosette absent. Stem leaves all similar or the upper with narrower lobes, petiolate, pinnate with a terminal orbicular to lanceolate, cordate to cuneate at base, abtuse at apex. Inflorescence crowded. Petals about twice as long as sepals. Stamens 6; anthers yellow. Stigma entire to ± emarginate. Seeds numerous, ovoid, brown. Flowering May to October (Rich *et al*. 1991).

There are two species of watercress which have been cultivated – green watercress and brown watercress. In Germany and France only green watercress was cultivated whilst in England during the 19th century both green and brown watercress were cultivated on a large scale. Green watercress was grown predominately in the summer months and the more winter-hardy brown watercress was grown throughout the year but especially in the colder winter months. Brown watercress was most popular with consumers; however it had a higher production cost because of the additional skill and labour required to maintain and vegatively propagate it (Sheridan,1996). Brown watercress soon was replaced in commercial cultivation by green watercress because of the ease of propagation by seed and the lower susceptibility to the fungal crook root disease caused by *Spongospora subterranea*sp. *Nasturtii*. Green watercress now appears to be the only species currently cultivated and consumed around the world (Palaniswamy and McAvoy. 2001) A French stock of green watercress was in common use and characterized by its pale to mid green colour, thin stems and capable of vigorous tillering which made it very productive particularly in the winter months (Stevens, 1983). A late flowering stock of green watercress selected at the National Vegetable Research Station was also in limited use at the same time in which the French watercress was in production. This stock was almost inactive in winter but grew rapidly in spring and produced good crops which were of a similar pale green colour to the French stock. Not only was it late to flower therefore extending the flowering season beyond the natural flowering time of the main winter stock in May but it was also limited in seed production (Stevens, 1983).

However, public preference switched to dark green watercress and hence the arrival of dark green watercress originating from USA which is characterized by its round, dark, glossy leaves and free flowering ability. Since its introduction in the late 1950s growers have carried out selection to improve leaf shape and colour, there are now many similar strains of this stock in existence (Stevens,1983) The three species of watercress are similar in appearance and exhibit morphological plasticity (particularly in leaf shape) (Table 1) (Rich *et al.*, 1991).

Table 1: Three main species of watercress (Rich *et al*., 1991)

Species	Description
Rorippa nasturtium-aquaticum (watercress)	Perennial, glabrous, dark green to purple. Rooting at nodes and branched, rosette absent. Flowering May to October. Common plant of ditches, streams, fens, springs, marshes, wet meadows, ponds and canals. Variable in leaf and size and shape. Interestingly dwarf and giant forms are occasionally found but are thought to be habitat modifications which are not maintained when cultivated.
Rorippa microphylla	Very similar to *R. nasturtium-aquaticum* distinguished by long, thin fruits and finely sculptured seeds. Found in similar places to *R. nasturtium-aquaticum* but less frequent overall.
Rorippa x sterilis (brown, hybrid watercress)	Very similar to *R.nasturtium-aquaticum* and *R.microphyllum* however inflorescences often elongating 30cm or more and the fruits are aborted, dwarfed or deformed. It is the commenest crucifer inter-specific hybrid and is widespread throughout Britain and Ireland – found in similar places to parents and sometimes in the absence of either one or both.

The taxonomic classification of watercress has been complicated and watercress has only recently been renamed in the genus *Rorripa*. The genus *Rorippa* consists of 10 species restricted to diverse habitats in Europe. It is believed that *Rorippa* originated from a cross between ancient ancestors of a *Rorippa* species (n = 16) with a primitive *Cardamine* (n =16). The two species of watercress identified; *Rorippa nasturtium-aquaticum* a diploid (2n = 32) the cultivated variety and *Rorippa microphylla* an tetraploid (2n = 64) have been known to hybridise to produce the sterile triploid (2n = 48) known as *Rorippa sterilis* (Rose, 2001) (Figure 1.2)

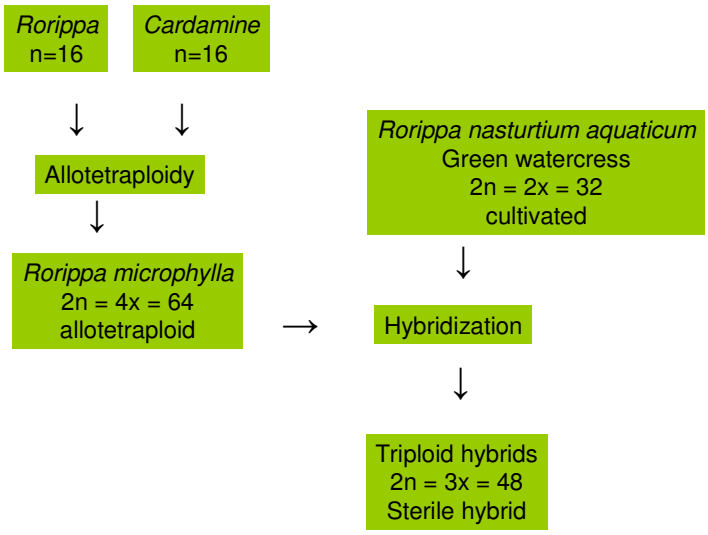


Figure 1.2: Proposed evolutionary pathway for the genus *Rorippa* (Simmonds, 1976)

1.1.2. Cultivation of watercress

Watercress is a perennial plant which thrives in alkaline water hence the dominant growth of watercress in the UK in the chalky south of Britain. The abundant supply of uncontaminated alkaline water with a temperature at source of 10°C or 11°C concentrates watercress production in Hampshire, Dorset and Wiltshire (Stevens, 1983).

Watercress flowers under long day conditions with the dominant growing season May through to September. The shape, size, construction and management of commercial watercress beds tend to vary between regions and growers. The bottom of a bed is a 10cm layer of a mixture of soil and fine washed gravel which is laid on top of the natural substratum and levelled. The bulk of the water which flows through the bed enters from boreholes and is maintained at a constant temperature of 10°C (Crisp, 1970). The beds can be stocked with either young transplants or seed. If seed are germinated on the substratum of the bed than the seedlings develop two leaves before any water is fed through the bed and if young plants are transferred to the bed then they are thinned once or twice before the watercress reaches a suitable stage of development and the first harvest taken (Crisp, 1970).

The watercress is harvested when it is approximately 25cm high, after harvesting a process termed 'patting or rolling' is carried out in order to ensure firm rooting in a bed that has been recently harvested. About 10-14 days after harvesting a fertilizer can be applied so approximately 3-6 weeks after this (during the summer months) the cress is ready for another harvest (Crisp, 1970). Once harvested the watercress is packaged and then distributed to supermarkets to be sold. Pre-packed watercress is predominantly sold as it sustains less physical damage as bunching can cause considerable damage, can be easily cleaned, simple to keep cool, contains no waste and lives in a micro-climate which is protected to some extent from ambient adverse conditions (Stevens, 1983). With bunches insects and other 'foreign bodies' may become trapped in the bunch, the bunch is difficult to keep cool and store and finally bunched watercress contains a high degree of waste (Stevens, 1983). With the current consumer market and convenience lifestyle it is more feasible to sell pre-packed watercress.

The current method for watercress propagation is to propagate the seeds in trays filled with compost and then to lay the trays on a bed of pallets which promote air circulation and reduce disease in a poly tunnel. The seedlings are irrigated and then transferred to a sloping watercress bed covered in fine gravel. The watercress is then grown in continuously flowing borehole water up till the point of harvest. During the

growth of the watercress it is rolled to help create uniform crop sizes. When the crop is ready to harvest it is harvested and sent for washing and packing (Vitacress method of cultivation).

A number of commercial strains of watercress have been isolated and selected for frost resistance, ability to maintain vegetative growth during summer when watercress tends to flower and resistance/tolerance to turnip mosaic virus. However little selective and systematic breeding has yet taken place and no standard commercial cultivars developed. One strain which has been commercially cultivated is Sylvasprings, developed in England and then grown in the United States but showed a lot of genetic diversity so was further selected to obtain a homogenous crop stand in the commercial watercress beds (Palaniswamy and McAvoy, 2001). Due to the susceptibility of watercress to various diseases many selections that have been undertaken have focused on breeding watercress with greater resistance to diseases such as yellow spot virus and crook root disease, a Japanese line was significantly less susceptible to crook root and WYSV and could be used to breed disease-resistant plants in the UK (Walsh, 1992). The ability to breed for desired traits i.e. nutritional and morphological would therefore set a precedent for watercress.

What would the watercress 'ideotype' consist of? Ideally the watercress would have a short (reduced) stem length (short internodes), with dark green, glossy, round leaves of medium size and produced throughout the year (Stevens, 1983) and leafier; however leaf size is restricted by mechanical harvesting. Late flowering is an important trait and the ability to set seed is another key asset (Stevens, 1983) Other important traits would be that watercress was a quick establishing vigorous plant which would allow for a rapid turnover of the crop in the summer and in the winter the crop would grow close to the water with vigorous tillering and produce a high density crop (Stevens, 1983). The crop also needs to withstand frost and other adverse weather conditions (Stevens, 1983). Ultimately the watercress should be a homozygous true breeding line with uniform nutrition and establishment thereby resulting in an even crop (Gilby, 1988). As well as improving the morphological

traits, higher concentrations of the beneficial antioxidant phenethyl isothiocyanate (PEITC) is also desired therefore improving biochemical traits. Finally during cultivation it is desirable to have a reduced requirement for fertilisation utilising phosphates.

There are three methods of basic selection that have been or could be applied by watercress growers; these involve rogueing, mass selection and progeny testing. Rogueing is the most basic system whereby undesirable plants are removed from the existing stock during its different phases of growth through the year (Steven, 1983). Mass selection is more labour intensive and expensive but has many benefits. Instead of removing the worst plants around 100 of the best plants with desirable characteristics are selected and planted in cultivated soil. The seed produced is grown in watercress beds the following year and 100 more plants are further selected (Stevens, 1983). The most complex method of selection is progeny testing which involves the cross pollination of selected plants from which seed is then collected from each individual plant. The seed is then grown and the progeny assessed and the genetic potential of each individual is seen. After determining the better progenies, the best plants from them are seeded together and the seed harvested separately again. This operation is then repeated in successive generations (Stevens, 1983). Any selection that is carried out in artificial conditions will need to be assessed in watercress beds after which they can be planted in cultivated soil. The production of watercress seed in isolated field conditions is required to prevent cross pollination between the newly selected stock and any existing stocks in watercress beds (Stevens, 1983).

The two ‘ideal’ traits I will focus on are the appearance (short stem length with dark and round leaves) and the beneficial nutrients (PEITC). Both traits were ranked highest when discussing the watercress ideotype with Vitacress Ltd. These traits were deemed favourable in order to entice the consumer to buy with the desired appearance and enhance nutrition by the manipulation of beneficial compounds.

1.2 Watercress – a 'functional food'?

The belief in the medicinal power of foods has been widely accepted in philosophy for generations (Milner, 1999). The term 'functional food' is described as the benefits which accompany ingesting foods that go beyond those accounted for merely by the nutriture provided (Milner, 1999). Although the term given to functional food is not a legal term, the concept of functional foods is gaining consumer acceptance (Milner, 1999) and a superfood is simply a term used to describe food which has a high phytonutrient content and therefore may result in health benefits. The Oxford English dictionary describes superfoods as 'food considered especially nutritious or otherwise beneficial to health and well-being'.

A food is considered as a functional food if 'it is consumed as part of a normal diet and demonstrated to have physiological benefits and/or reduce chronic disease beyond basic nutritional function i.e. bioactive compounds' (personal communication Oxford Brookes First Functional Food Conference, 2010). Watercress is packed with a wide range of natural, bioactive plant compounds (phytochemicals) for which there is increasing evidence for beneficial effects on human health. The leafy green salads (lettuce, rocket, spinach and watercress) contain many beneficial nutrients (Table 2).

Table 2: Nutritional composition of 100g of fresh lettuce, rocket, spinach and watercress (USDA, National Database for Standard Reference, 1984)

Nutrient	Lettuce	Rocket	Spinach	Watercress
Water (g)	94.98	91.71	91.4	95.11
Energy (kcal)	15	25	23	11
Protein (g)	1.36	2.58	2.86	2.3
Fat (g)	0.15	0.66	0.39	0.1
Carbohydrate (g)	2.87	3.65	3.63	1.29
Fibre (g)	1.3	1.6	2.2	0.5
Calcium, Ca (mg)	36	160	99	120
Iron, Fe (mg)	0.86	1.46	2.71	0.2
Phosphorus, P (mg)	29	52	49	60
Magnesium, Mg (mg)	13	47	79	21
Potassium, K (mg)	194	369	558	330
Sodium, Na (mg)	28	27	79	41
Vit C, ascorbic acid (mg)	9.2	15	28.1	43
Thiamin (mg)	0.07	0.044	0.078	0.09
Riboflavin (mg)	0.08	0.086	0.189	0.12
Niacin (mg)	0.375	0.305	0.724	0.2
Panothenic acid (mg)	0.134	0.437	0.065	0.31
Vitamin B6 (mg)	0.09	0.073	0.195	0.129

1.2.1 Watercress and Glucosinolates

Watercress is a rich source of phenethyl isothiocyante (PEITC), a glucosinolate derivative which is a key contributor to the distinctive peppery flavour and has been shown in a number of *in vitro* and animal studies to have a range of beneficial anti-cancer effects Watercress contains a plentiful supply of lutein and zeanthin carotenoids as well as considerable amount of vitamins A, C and K (Czapski, 2009) Figure 1.3 demonstrates the wide range of natural bioactive plant compounds in watercress which have beneficial effects on human health ranging from antioxidants (betacarotene, lutein and quercetin) which may help protect the cells of the body from damage from reactive free radicals to a group of phytochemicals called isothiocyanates which have anti-cancer potential (Costain, 2007). I want to focus on the glucosinolates and their conversion to isothiocyanates due to their potential anti-cancer properties.

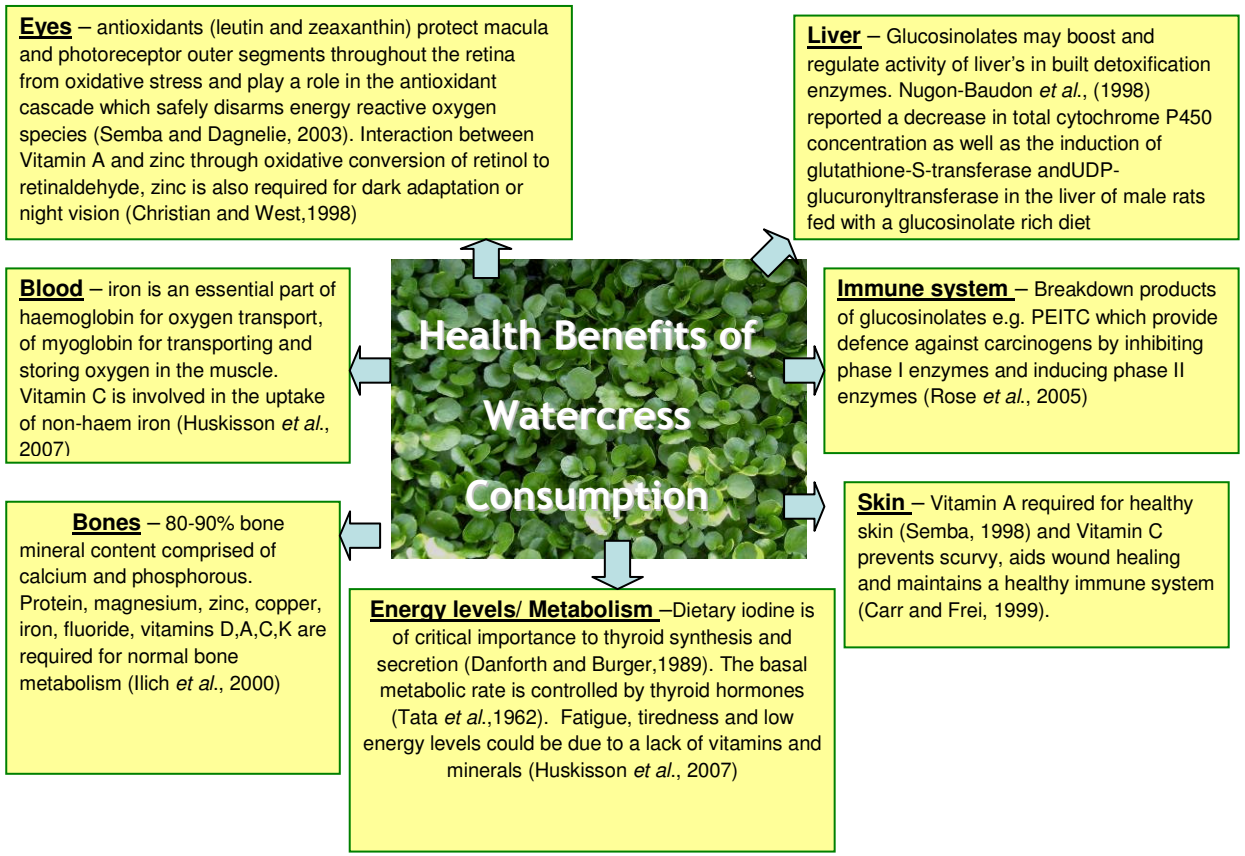


Figure 1.3: The range of beneficial vitamins and minerals contained within watercress (adapted from Costain, 2007).

Cruciferous vegetables are rich sources of glucosinolates; watercress contains a high concentration of glucosinolates per gram weight. (Gill *et al.*, 2007). Watercress contains 17.98µmol/g dry weight 2-Phenethyl-GLS (gluconasturtiin) (Gill *et al*., 2007) whilst broccoli contains 0.4µmol/g^{-1} dry mass, Brussels sprouts 0.5µmol/g^{-1} dry mass, cabbage 0.3µmol/g^{-1} dry mass, cauliflower 0.4µmol/g^{-1} dry mass and kale 4µmol/g^{-1} dry mass (Kushad *et al*., 1999).

Glucosinolates are a class of sulphur-and nitrogen containing glycosides derived from amino acids which serve a protective function in many plants. Upon damage to a cruciferous vegetable, myrosinase is released which results in the degradation of glucosinolates and the formation of isothiocyanates. The glucosinolates can be broken down into several forms. The first step of the breakdown is catalyzed by the enzymes termed myrosinases. The unstable aglycone produced after the removal of the glucose moiety rearranges into isothiocyanates, nitriles or thiocyanates through a spontaneous process which does not require enzymatic activity (Ribnicky *et al*. 2001). It is the isothiocyanates which are of interest, particularly phenethyl isothiocyanate (Figure 1.4), due to its reported ability to inhibit cancer development. Work has investigated the genetic and environmental factors which lead to glucosinolate accumulation in plants and the role of the glucosinolates and their derivatives in plant-herbivore interactions. However there is currently an understanding of how glucosinoates and their products could contribute to a reduction in the risk of carcinogenesis and heart disease (Tarka and Mithen, 2009).

A broccoli has been produced with an enhanced concentration of isothiocyanate using introgression of 3 segments of *Brassica villosa* (a wild relative of broccoli) genome into broccoli (Mithen *et al.* 2003). Broccoli has recently been shown to significantly prolong the survival time of fruit flies and it is the free-radical scavenging activity via the up-regulation of endogenous antioxidant enzymes at both the transcriptional and translational level which prolong the mean survival of the fruit flies (Yuk Man Li , 2008). The phenolics and glucosinolates of broccoli (*Brassica oleracea* var. botrytis L.) contribute significantly to this antioxidant activity of broccoli (Yuk Man Li , 2008). The precise modification of the pathway

which leads to enhanced glucosinolate accumulation is yet to be resolved but it is believed that these genomic segments contain allelic forms of methylalkylmalate synthase genes (Traka and Mithen, 2009). The methylalkylmalate synthase genes are associated with determining the extent of the chain elongation of methione-derived glucosinolates and also are associated with QTLs determining total amount glucosinolates in *Brassica* and *Arabidopsis* (Traka and Mithen, 2009).

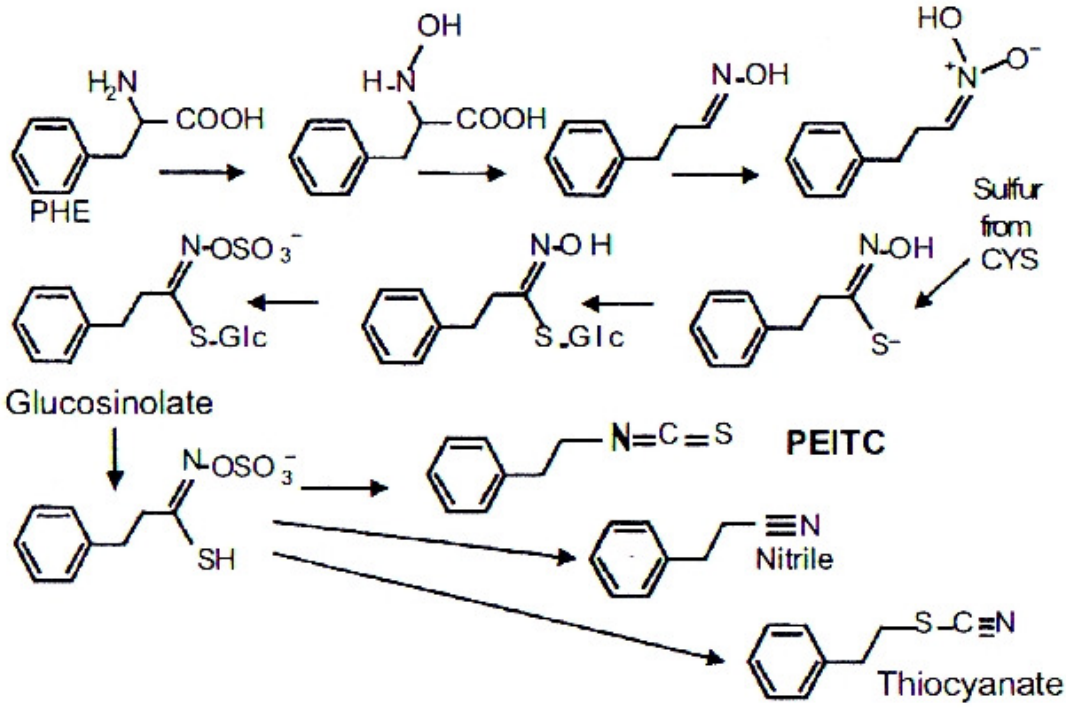


Figure 1.4: Biosynthetic pathway of PEITC (Ribnicky *et al*., 2001)

1.2.2 Watercress and Phenethyl Isothiocyanate

PEITC can inhibit cancer development, as shown in many studies from laboratory and cell cultures through to animal models and human trials, by:

- preventing carcinogen activation by inhibiting phase I enzymes such as cytochrome P450s thereby stopping a potential carcinogen becoming a carcinogen (Rose *et al.*, 2000)
- increasing the ability of cells to resist attack from carcinogens by increasing the activity of detoxifying/antioxidant enzymes known as phase II enzymes e.g. quinone reductase, glutathione-S-transferases and UDPglucuronosyltransferase (Rose *et al*., 2000)
- inhibiting cell cycle progression which inhibits the uncontrolled growth of cancer cells (Chiao *et al.*, 2004)
- inducing apoptosis which is the death of damage or cancer cells (Satyan *et al.*, 2006a)
- Anti-angiogenesis – prevents formation of new blood vessels to the tumour (Kenfield *et al.*, 2007)

Carcinogenesis consists of initiation, promotion and progression thus providing many opportunities for intervention at each stage by phytonutrients (D'Ambrosio, 2007). The list of studies above provides evidence that PEITC intervenes at many stages.

Numerous papers have been published which support the anti-cancer role of PEITC. A study carried out by Rose *et al*. (2005) looked at how broccoli and watercress suppress matrix metalloproteinase-9 activity and invasiveness of human MDA-MB-231 breast cancer cells. The matrix metalloproteinases (MMP) represent a family of proteolytic enzymes that can degrade extracellular matrix (ECM); matrix metalloproteinase 2 (MMP-2) and matrix metalloproteinase 9 (MMP-9) are known to be essential. Elevated levels of these are found in cancerous tissues and tumor cells. This has raised considerable interest into identifying natural and synthetic

compounds that can inhibit MMP activities (Rose *et al.*, 2005) Previous epidemiological evidence has shown that a diet rich in cruciferous vegetables can reduce the incidence of breast cancer in females (Ambrosone *et al.*, 2004), rodent models demonstrated that ITCs (isothiocyanates) can induce phase II detoxification enzymes, induce cell cycle arrest and promote apoptosis in breast cancer cells in vitro (Jackson and Singletary, 2004; Rose *et al.*, 2005).

Rose *et al* (2005) showed that extracts of watercress can inhibit MMP-9 activities as well as inhibiting the invasive potential of the human breast cancer cell line MDA-MB-231. Watercress has anti-invasive and anti-metalloproteinase activities and its phytochemical constituents, the isothiocyantes, can be termed a new class of invasion inhibitors. This paper provides further promising insight into how cruciferous vegetable may contribute to human chemoprevention (Rose *et al.*, 2005) mainly by inhibiting proteolytic enzymes which cause extracellular matrix destruction.

An earlier paper published by Rose *et al*. (2000) highlighted how watercress is an exceptionally rich source of PEITC which inhibits phase I enzymes and induces phase II enzyme but also highlighted the fact that ITCs derived from methylsulfinylalkyl glucosinolates may be more important phase II enzymes inducers than PEITC reporting a 10 to 25-fold greater potency. This paper raised an important finding that watercress may have exceptionally good anticarcinogenic potential due to the fact that it combines a potent inhibitor of phase I enzymes (PEITC) with at least three inducers of phase II enzymes (Rose *et al.*, 2000).

The most recent finding which was published by Gill *et al.* (2007) reported that watercress supplementation in the diet reduces lymphocyte DNA damage and alters blood antioxidant status in healthy adults. Gill *et al.* (2007) report results from in vitro studies suggest that watercress extract can protect cells against DNA damage levels induced by genotoxic agents (affect integrity of cells genetic material, potentially mutagenic or carcinogenic). There was a decrease in DNA damage in lymphocytes in response to watercress consumption in vivo. It appears that

watercress consumption can be linked to a reduced risk of cancer from decreased damage to DNA and possible modulation of antioxidant status by increasing carotenoid concentrations (Gill *et al.*, 2007).

The first study to report an effect of vegetable consumption on metabolism of a lung carcinogen in humans was published by Hecht *et al.* (1995). The paper reported that PEITC inhibited the carcinogenicity of tobacco-specific lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) (a potent pulmonary carcinogen in rodents, also believed to be one of causes of lung cancers in smokers) by inhibiting its metabolic activation. Results from the study confirmed this and supported the hypothesis that PEITC inhibits oxidative metabolism of NNK in humans as seen in rodents and support further development of PEITC as a chemopreventive agent against lung cancer (Hecht *et al.*, 1995).

Chung *et al.* (1992) had previously published a paper showing that PEITC inhibited the lung tumorgenesis induced by a potent tobacco specific carcinogenic nitrosamine in animals, tobacco-specific lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK, nicotine derived nitrosamine) treated rats fed a diet containing PEITC developed significantly fewer lung tumors than (NNK)-treated rats fed a control diet, watercress was used as a dietary source of PEITC (Chung *et al.*, 1992).

Wattenberg (1992) and Siglin *et al.* (1995) reported that PEITC inhibited cancers in rats and mice that are caused by several tobacco specific carcinogens and Meyer *et al.* (1995) inferred that this was due to PEITC acting as both a blocking agent and an inhibitor of tumor initiation via inhibition of cytochrome P450 enzymes and by induction of phase II enzymes (glutathione S-transferase) (Palaniswamy *et al.*, 2003).

In 1992 human uptake of PEITC was quantified by Chung *et al*. The aim of the research was to demonstrate that since watercress is rich in gluconasturtiin (the glucosinolate precursor of PEITC), that PEITC is indeed released in the human body

upon ingestion of watercress (Chung *et al.*, 1992). A urinary metabolite was found, N-acetylcysteine, which is a conjugate of PEITC. The results from the experiment indicate that N-acetylcysteine conjugate of PEITC may be a useful marker for quantitating human exposure to this anticarcinogen (Chung *et al.*, 1992). When cooking vegetables the myrosinase enzyme is inactivated and the level of glucosinolates converted to isothiocycanates is reduced , the conversion still occurs but is considerably less than after ingesting uncooked vegetables (Getahun and Chung, 1999). It is known that the microflora in the intestinal tract in humans is responsible for the hydrolysis of glucosinolates to ITCs (Getahun and Chung, 1999).

Experimental tests in the laboratory using models of cellular activity demonstrate that watercress isothiocynates also suppress the production of pro-flammatory compounds e.g. nitric oxide (NO) and prostaglandins (PG) and these are associated with chronic inflammation and cancer (Czapski, 2009). Further works continues investigating key points in the progression of cancer in which PEITC can intervene including looking at the effect of PEITC on the expression of α and β- tubulins in prostate cancer cells (Yin *et al*., 2009) and induction of autophagic cell death (Powolny *et al*., 2011).

The emergence of 'watercress as nature's richest source of PEITC' provides promising and exciting potential for watercress to be a preventative measure against cancer. Therefore, in terms of humans, PEITC is beneficial, however PEITC can act as deterrent for many invertebrates.

1.3 Watercress and the glucosinolate-myrsoinase system

Watercress possesses the glucosinolate-myrosinase system. This system acts as a chemical defence for terrestrial crucifers. When watercress is damaged the myrosinase mediated hydrolysis of phenethyl glucosinolate to phenethyl isothiocyanate which is toxic is initiated (Newman *et al.*, 1992).

Newman *et al.* (1990) investigated the potential role of defensive compounds in the avoidance of watercress by the amphipod, *Gammarus pseudolimnaeus*. *G. pseudolimnaeus* consumed around five times more yellowed watercress compared to the green watercress which had higher nitrogen content. However results indicated that despite having higher nitrogen content the green watercress also contained seven times more phenethyl isotiocyanate compared to the yellowed watercress. These findings indicate that live watercress is chemically defended to avoid consumption and that the glucosinolate-myrosinase system which has been recognized as the principle deterrent system of terrestrial crucifers is also possessed by watercress which may contribute to defence against herbivory by aqauatic crustaceans (Newman *et al.*, 1990).

Further work carried out by Newman *et al*. (1999) involved a series of choice tests which were carried out on an amphipod, *Gammarus pseudolimnaeus*; limnephilid caddisflies, *Hesperophylax designatus* and *Limnephilus* species and physid snail, *Physella* species. All were given a choice of yellowed senescent watercress or fresh green watercress, the yellowed senescent watercress was preferred over the fresh green watercress which had a higher nitrogen content (Newman *et al.*, 1992). However, upon heating the watercress which deactivates the myrosinase, the species preference switched to the heated green watercress. Therefore deactivation of the myrosinase enzyme, hence isothiocyanate production resulted in a switch of feeding preference indicating that the presence of isothiocyanate does indeed act as a deterrent.

Canola crops have been shown to inhibit soil-borne pathogens in following crops. This effect is believed to be a result of the release of low molecular S-containing compounds i.e. isothiocyanates during microbial degradation of the crop residues (Rumberger and Marschner, 2003). The toxicity of the isothiocyanates is due to their irreversible reaction with sulphydryl groups, amine groups and disulphide bonds of proteins which can lead to enzyme inactivation. Glucosinolates and myrosinase (enzyme which hydrolyzes glucosinolates to isothiocyantes, nitriles and ionic thiocyanate) are separated spatially in intact cells, upon disruption myrosinase

converts glucosinolates to isothiocyanates. The formation of biocidal metabolites from glucosinolates is used in 'biofumigation', a suppression of soil-bourne pests and pathogens after incorporating green manure or harvest residues of canola or other high glucosinolate brassicas (Newman *et al.*, 1992). PEITC can therefore influence the microbial community in the rhizosphere of canola roots and despite the fact that PEITC is rapidly degraded in the soil it can still be a dominant selective factor for the bacterial and eukaryotic rhizosphere community (Rumberger and Marschner, 2003).

This may be one important factor to take into consideration whilst attempting to manipulate the concentration of PEITC – could this have further detrimental effects on the surrounding invertebrate population? However could the increased presence of PEITC in the crop act as a natural pest deterrent therefore reducing the amount of pesticide required and also a higher quality crop with less microbial contamination.

1.4 Factors affecting the concentration of PEITC in Watercress

Numerous environmental factors affect the concentration of PEITC in watercress leaves. Among these are stage of harvest, temperature, photoperiod, short term increase in photosynthetic flux and the proportion of sulphur (Palaniswamy *et al*. 1997). The concentration of PEITC and 2-phen(yl)ethyl glucosinolate, which is a flavour component, is optimal when the plant has approximately 12 to 15 internodes (3-4 weeks after transplanting or 6-7 weeks after sowing) (Palaniswamy and McAvoy, 1997). PEITC concentration is also affected by the photosynthetic period and a short increase in the photosynthetic flux. Palaniswamy *et al.* (1997) findings indicate that PEITC concentration is influenced by a higher photosynthetic flux before harvest but only for plants grown under a short (8 hour) photoperiod. Plants grown under long days have a higher gluconasturtiin concentration when compared to plants grown under short days by 30-40%. The concentration of gluconasturtiin was also increased by 25-40% by growing watercress plants under red light when compared to watercress grown under far red light. Further support is provided by at

least 25% higher concentration of gluconasturtiin in watercress with brief exposure to red light at the end of the main photoperiod compared to an end of day far red light exposure (Palaniswamy *et al*., 1997). Palaniswamy and McAvoy (2001) state that 'one week of bright sunlight whilst growing watercress in open fields can yield more flavourful and healthier produce than if harvested after a period of cloudy days'.

As well as photoperiod influencing PEITC concentrations, temperature can also influence the PEITC concentration. By growing watercress at 15 or 10°C there is as at least a 50% higher concentration of gluconasturtiin compared to watercress grown at 20 or 25°C (Engelen-Eigles *et al*., 2006). In order to optimize the gluconasturtiin concentration and therefore the level of PEITC, watercress needs to be grown at low temperatures, under long days and exposed to red light.

Watercress fertilized with a higher proportion of sulphur yielded a more flavourful crop; leaves of watercress grown in hydroponics with a higher ratio of sulphur to nitrogen also contained a higher concentration of PEITC (Palaniswamy and McAvoy. 2001). Similarily, the phytochemical compostion of broccoli varies with the cultivar, growing conditions and seasons thereby effecting the overall antioxidant activity (Yuk Man Li , 2008).

1.5 Quercetin and Indole-3-carbinol overview

Two other potential anticancer compounds are Quercetin and Indole-3-carbinol (I3C). These are two interesting compounds to take into consideration alongside PEITC due to their beneficial synergistic effects, clearly the plant matrix is complex and individual compounds are difficult to be extrpolated for individual benefits. Breast cancer has become the most frequently diagnosed cancer among women in the United States and is the second most frequent cause of cancer death. In around 30% of breast cancer patients Her-2/*neu* overexpression is directly linked to deregulated activation of intracellular mitogenic signaling, leading to aggressive

tumour behaviour. Also an increase in Her-2/*neu* expression has been found to enhance malignant phenotypes of cancer cells including those with metastatic potential (Jeong *et al.*, 2008; Kenfield *et al.*, 2007). Her-2/*neu* (ErbB2) is a transmembrane tyrosine kinase and acts as a co-receptor for the other EGFR family members. Her-2/*neu* is a member of the subclass I of the receptor kinase (RTK) superfamily which comprises of four members: Her-1/epidermal growth factor receptor (EGFR)/ErbB1, Her-2/*neu*/ErbB2, Her-3/ErbB3 and Her-4/ErbB4. High expression of Her-2/*neu* is associated with a poor prognosis in breast cancer (Jeong *et al.*, 2008). This has led to Her-2/*neu* becoming an important therapeutic target in breast cancer. A study by Jeong *et al*. (2008) revealed that quercetin decreased the level of Her-2/*neu* protein in a time- and dose-dependent manners and also inhibited the downstream survival PI3K-Akt signaling pathway in Her-2/*neu*-overexpressing breast cancer SK-Br3 cells. Quercetin also induced polyubiquitination of the Her-2/*neu*.

Studies in animal models have shown that down-regulation of Her-2/*neu* can suppress tumour growth and dissemination. The first approved monoclonal antibody treatment for breast cancer cells was Trastuzumab (Herceptin). This treatment resulted in significant improvement in patient survival when used in combination with chemotherapy in patients with metastatic Her-2/*neu*. Although trastuzumab is an effective therapeutic agent it was shown to increase the risk of cardiac dysfunction. An alternative therapeutic agent may be quercetin (3,5,7,3',4'-pentahydroxyflavone), a natural product found in many fruits and vegetables (Jeong *et al.*, 2008). Quercetin is a flavonoid present in many vegetables and fruits and is a potent antioxidant which has oxygen radical scavenging properties and inhibits xanthine oxidase band lipid peroxidase in vitro. Quercetin also has anti-tumor, anti-inflammatory, anti-allergic and anti-viral activities and has been extensively studied as a chemoprevention agent in cancer models (Jeong *et al.*, 2008).

Glucobbrassican which is a major component of cruciferous vegetable is a 3-indolylmethyl glucosinolate which is readily hydrolyzed in the acidic conditions of the gut to give Indole-3-carbinol (I3C). Indole-3-carbinol (I3C) conjugates are

phytochemicals expressed in brassica vegetables and demonstrate anticancer activities. I3C in combination with its metabolite di(3'-indolyl)methane (DIM) induce overlapping and unique responses in multiple cancer cell lines and tumors which include growth inhibition, apoptosis and antiangiogenic activities (Jeong *et al.*, 2008; Safe *et al.*, 2008).

It is well researched that cruciferous vegetables contain a wide variety of phytochemicals but research has been based on the isothiocyanates and indole-3-carbinol (I3C) as the major chemopreventive and chemotherapeutic phytochemicals associated with the anticancer activities of cruciferous vegetables. The interactions between the phytochemicals may prove significant and whilst each can have beneficial effects alone the combined effects may be greater. Indeed the synergistic combination of phenethyl isothiocycanate and sulforaphane could be more effective than used alone in preventing inflammation and the associated diseases such as cancer as inflammation is a critical component of tumour progression (Cheung *et al.*, 2009).

1.5.1 Indole-3-carbinol (I3C)

The most extensively studied bioactive food component within Brassicas is I3C which arises from an indolyl-methyl glucosinolate (glucobrassican (Safe *et al.*, 2008)) when Brassicas are crushed or cooked (Kim and Milner, 2005) but also the acidic conditions in the gut cause the conversion of glucobrassican to I3C (Safe *et al.*, 2008). I3C is then converted to the biologically active dimer 3,3'-diindolylmethane (DIM) within the intestinal tract. Changes in cell cycle progression, apoptosis, carcinogen bioactivation and DNA repair account for the anticancer properties of I3C/DIM (Kim and Milner, 2005). I3C/DIM modulates several nuclear transcription factors and I3C exerts anticancer by modulating the expression of several genes that play a key role in many cancer processes (Kim and Milner, 2005). It is multiple aspects of cancer cell cycle regulation and survival that is targeted by I3C and DIM (Weng *et al.*, 2008). The ability of I3C to alter various cancer processes with physiologically achievable concentrations makes it a very

interesting dietary phytochemical (Kim and Milner, 2005) to further research. Safe *et al* 2008 reported that the anti cancer effects are probably due to a combination of I3C and DIM and is difficult to dissect out those responses which are due to I3C alone (Safe *et al.*, 2008). Figure 1.5 below shows the major pathways in cancer cells which I3C and DIM act on: inhibition of cell proliferation and cell cycle progression and induction of cell death or decreased cell survival (Safe *et al.*, 2008). I3C can also been seen to modulate the expression of several genes and proteins which lead to the inhibition of cell growth (Safe *et al.*, 2008)

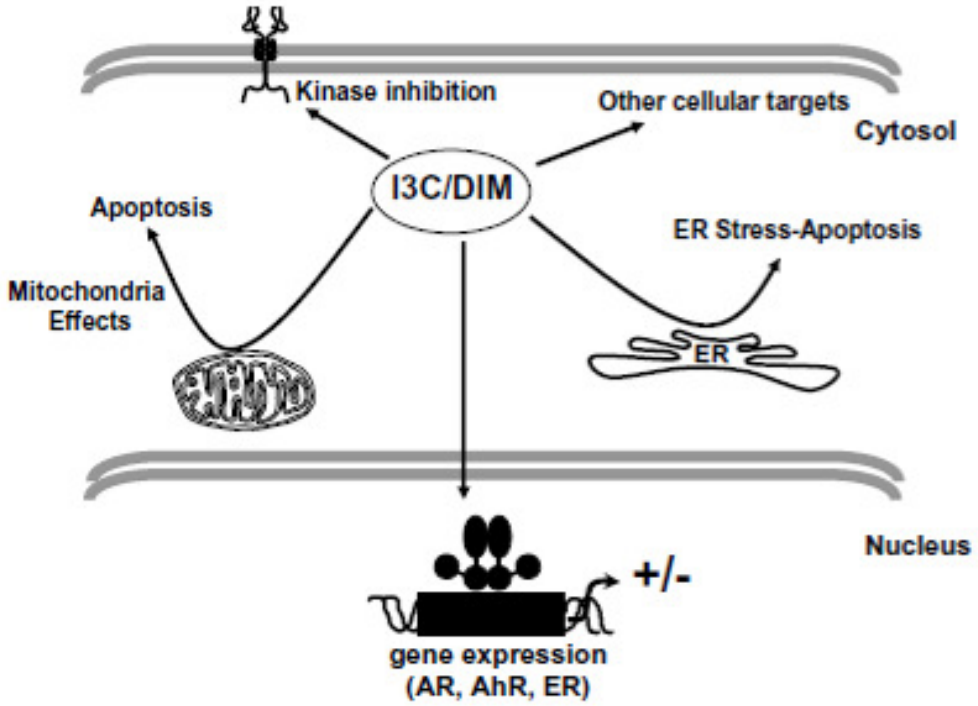


Figure 1.5: Multiple responses and genes in cancer cell lines which are modulated by I3C/DIM (Safe *et al.*, 2008)

The ability of I3C to alter various cancer processes with physiologically achievable concentrations makes it a very interesting and promising dietary phytochemical (Kim and Milner, 2005).

1.5.2 Quercetin

Quercetin (3,3',4',5,7-pentahydroxyflavone) is a plant flavonoid and occurs wide range of fruits and vegetables (Csokay *et al.*, 1997), it is the most common flavonol in the diet (Erlund, 2004). Quercetin inhibits the growth of malignant cells by arresting them in late G1 phase of the cell cycle and also blocks signal transduction pathways by inhibiting protein tyrosine kinase, P1- and PIP-kinases (Csokay *et al.*, 1997). Angiogenesis is characterized by the formation of new vessels from a pre-existing microvascular network and quercetin shows significant antiangiogenic activity. It has also been shown to suppress tumor growth in vitro and in vivo and scavenges free radicals by oxidation (Tan *et al.*, 2003). Quercetin has been shown to decrease the level of the Her/*neu* protein in a time and dose dependent manner (Jeong *et al.*, 2008). Her-2/*neu* (ErbB2) is a transmembrane tyrosine kinase which acts as a co-receptor for other EGFR family members (Jeong *et al.*, 2008). Her2/*neu* overexpression is observed in 30% of all breast cancer patients and is associated with a poor prognosis (Jeong *et al.*, 2008). By using quercetin new therapeutic strategies can be initiated in order to fight breast cancer.

Plants are by far the most abundant and cost-effective renewable resource which are uniquely adapted to complex biochemical synthesis (Raskin *et al.*, 2002). The greatest impact of plant-derived drugs was in the anti-tumour area whereby taxol, vinblastine, vincristine and cheptothecin have dramatically improved effectiveness of chemotherapy against some of the deadliest cancers (Raskin *et al.*, 2002). There are around 250 000 living plant species which contain a much greater diversity of bioactive compounds than any chemical library made by humans (Raskin *et al.*, 2002) and offer a superior source of molecular diversity and novel molecular chemotypes (Raskin *et al.*, 2002). It is interesting to see how these bioactive compounds function independently as well as in synergy. It has long been suggested that for a healthy diet and to reduce the risk of chronic diseases that 5 to 10 servings of a wide variety of fruits are consumed in order to reduce the risk of chronic diseases (Liu, 2004).

In the UK chronic disease is responsible for 75% and therefore we should not underestimate the importance of research investigating and understanding the mechanisms of specific phytochemicals (Hancock *et al*., 2007). Breeding targets have shifted from crop protection via increasing pest and disease resistance towards quality traits i.e. maximization of phytochemicals (Hancock *et al*., 2007). More investment into the various health benefits of certain crops will allow these benefits to be substantiated and clearly identify breeding targets (Hancock *et al*., 2007).

1.6 Dwarf Watercress and The Role of Gibberellins

Increasing the levels of beneficial antioxidants provides one aspect to the watercress breeding programme however another key trait to investigate is the aesthetics, namely a reduction in the stem length. Working alongside Vitacress Ltd. one of the traits inferred in the watercress ideotype was a reduction in stem length. After personal communication with the Vitacress market research team one popular request from consumers was a reduction in the stem i.e. a dwarf phenotype.

'*Our multiple retailer customers believe that consumer preference is for smaller leafed, thinner stemmed watercress. Producing to this specification requires relatively high density crops which are prone to etiolation, especially under poor light levels. The result is the desired reduction in leaf size and stem thickness, but offset by too "stalky" a sprig, with undesirably long internodes and petioles. To try to counter this negative aspect we need to find a strain that has a propensity to a shorter stem length and more compact growth*'

Dr. Steve Rothwell (Group Production and Technical Director Vitacress Ltd.)

(personal communication)

Not only would watercress with a dwarf phenotype fill a 'gap in the market' but also it is believed that a higher concentration of antioxidants is present in the leaves

(personal finding, chapter three). Therefore a dwarf, leafy watercress plant could indeed be a valuable assest to aid in increasing the nutrition of the general public.

Dwarfism is a desirable characteristic for many of the agricultural crops. It may also prove to be a desirable characteristic for watercress; by producing a plant with a reduced stem length it will be more aesthetically pleasing to the consumer and easier for packaging and harvesting and for a variety of ready meal uses. Evans stated that 'dwarfing is one of the most valuable traits in crop breeding because it results in plants that are more resistant to damage by wind and rain (lodging resistant) and have stable increased yields' (Evans, 1993). This proved particularly useful in grain crops such as wheat and rice and the green revolution was fuelled by the introduction of the semi-dwarf trait into cereal crop cultivars (Sakamoto *et al.*, 2003). This may also prove to be of use in vegetable crops such as watercress as well as the grain crops. Dwarfism may in fact be a favourable trait giving the crop increased hardiness to many envirmonmental fluctuations and stresses. By gaining information regarding dwarfism from a range of crops we can apply and adapt this information to the understudied watercress crop.

Dwarf mutants have been extensively characterised in many plant species and endogenous phytochrome giberellin is one of several factors associated with the dwarf phenotype (Sakamoto *et al.*, 2003). There are around 136 naturally occurring gibberellins which have similar chemical structure but few have intrinsic biological activity. There are many different gibberellins, e.g. GA_1, GA_2 and GA_3 (also known as gibberellic acid which is usually the principle component). Gibberellins are plant hormones which regulate growth and influence many developmental processes and are best known for the promotion of stem elongation and application of GA stimulates a dramatic stem elongation in dwarf mutants of peas, maize and many rosette plants (Taiz and Zeiger, 2006). The development and environmental response of plants is a result of the periodic and spatial regulation and the diversity of GA perception and subsequent signal transduction to successive transmitters (Nakajima *et al.*, 2006). The three major commercial uses of gibberellins, mainly GA_3 are (Jackson and Singletary, 2004; Rose *et al.*, 2005; Taiz and Zeiger, 2006):1. Promote growth of fruit crops, 2. Stimulate barley malting process in beer-brewing industry and 3. Increase sugar yield. However in many crop plants a reduction in plant height

is desirable and therefore inhibitors of gibberellin biosynthesis are useful, an example of an inhibitor sprayed in Europe is Cycocel (Taiz and Zeiger, 2006) and chlormequat.

1.6.1 Gibberellin Biosynthesis

There are three main stages in gibberellin biosynthesis (Figure 1.6)

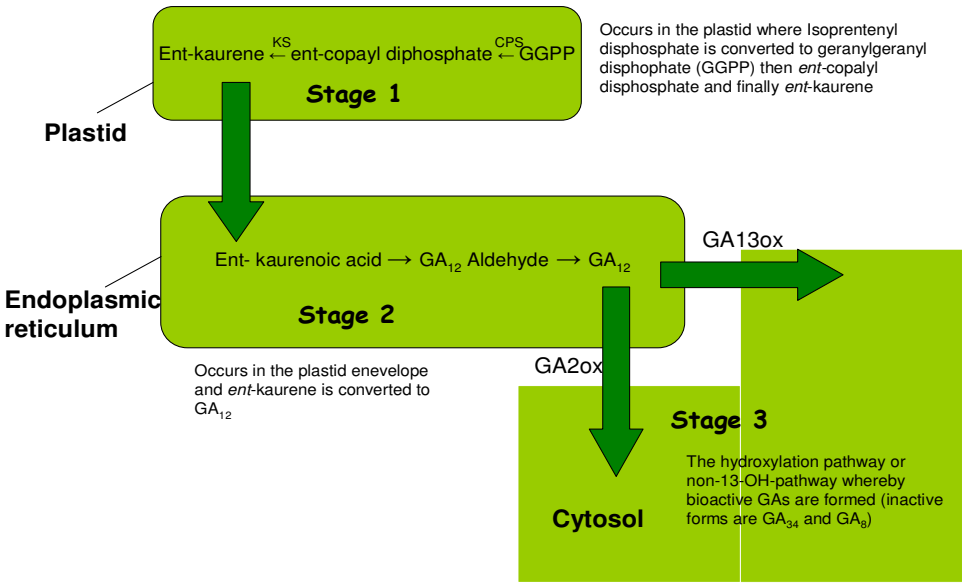


Figure 1.6: Gibberellin biosynthesis (adapted from Taiz and Zeiger, 2006)

Plants which are impaired in GA biosynthesis have generally small and dark leaves; reduced stem length and some are defective in seed germination and floral development and are delayed in flowering time. On the other hand, an increase in GA concentrations or GA signaling results in a tall and spindly phenotype (Fleet and Sun, 2005).

1.6.2 Genes controlling GA metabolism and catabolism

Hormone homeostasis depends on both synthesis and deactivation. Therefore to prevent excess stem elongation GA biosynthesis should be depressed and the catabolism of GA stimulated. This is achieved by down regulation or inhibition of the GA20ox and GA3ox genes which encode the last enzymes for the formation of bioactive GA (Table 3). Alternatively, overexpression of GA2ox results in GA_1 deactivation (Table 3) (Taiz and Zeiger, 2006).

Table 3: The role of GA20ox, GA3ox and GA2ox in GA biosynthesis and catabolism (Taiz and Zeiger, 2006)

Enzyme	Role	Encoded by	Mutation	Mutant Phenotype
GA 20-oxidase (GA20ox)	Catalyzes steps between GA12 and GA9 and between GA53 and GA20	GA5 gene	Ga5	Semi-dwarf
GA 3-oxidase (GA3ox)	Converts GA9 into bioactive GA4 and GA20 into bioactive GA1	GA4 gene	Ga4	Semi-dwarf
GA 2-oxidase (GA2ox)	Deactivates bioactive GAs			No known mutant phenotype
		Ga1, GA2 and GA3	Ga1, ga2 and ga3	Extreme dwarfs

Mutations which result in a dwarf phenotype occur due to the GA response being prevented from taking place or the levels of endogenous GAs are severely reduced i.e. GA-deficient dwarfs (Taiz and Zeiger, 2006). The identification of bioactive GAs and characterization of key enzymes in GA biosynthesis and catabolism will allow researchers to alter the concentrations within plants of bioactive GA thus

affecting plant height (Taiz and Zeiger, 2006). Mutants with dwarf phenotypes have been isolated and analyzed to identify the genes responsible for the phenotypes (Nakajima *et al.*, 2006).

Plant hormone biosynthesis can be modified by the introduction of bacterial genes. Recently biosynthetic pathways have been elucidated for most hormone classes and genes for many of the enzymes have been cloned thereby creating new opportunities to manipulate hormone content in order to study mode of action and regulation of their biosynthesis (Taiz and Zeiger, 2006). The technology also provides means to introduce agriculturally useful traits into crops (Taiz and Zeiger, 2006).

1.6.3 GA 20-oxidase and GA 3-oxidase

Several genes encoding GA biosynthetic enzymes are now available making the genetic engineering of GA content possible. GA 20-oxidase catalyses several consecutive steps in the GA biosynthetic pathway and act at important branch points in the GA biosynthetic pathway making them key enzymes in the pathway (Taiz and Zeiger, 2006).

When GA 20-oxidase is overexpressed in transgenic *Arabidopsis* the resultant phenotype has longer hypocotyls, early flowering, increased stem elongation and reduced seed dormancy which is associated with an increase in bioactive GA_4 (the main bioactive GA). It appears that growth throughout the life cycle of *Arabidopsis* is limited by GA content, particularly GA 20-oxidase which is encoded by at least three genes. Plants with antisense At20ox1 have short hypocotyls and reduced stem elongation, a semi-dwarf phenotype in which the amount of GA_4 in the rosette is reduced to approximately one third of that in the control. Antisense At20ox2 grew normally in long days and slightly shorter during short days when compared to control plants due to a reduction in floral internode length. Finally antisense At20ox3 had reduced plant hypocotyls length early in seed development but grew normally thereafter (Hedden and Philips, 2000).

Alternative approaches to reducing GA content in plants can be achieved using enhanced catabolism or sequestration of bioactive GAs or their precursors. A GA 20-oxidase in pumpkin seeds mainly produces inactive tricarboxylic acid GAs (this being an exception to other 20-oxidases) and when overexpressed in *Arabidopsis* GA_4 is reduced by up to 80% (Hedden and Philips, 2000). However the plants are only slightly dwarfed. GA_4 responsiveness in the stem tissues appears to be less affected than total GA measures in the whole plant. Overexpression of GA 20-oxidase successfully reduced GA content and stem length in *Solanum dulcamara* (Hedden and Philips, 2000). Hydroxylation of GAs at C-2β causes irreversible inactivation. When an enzyme, GA 2-oxidase, which catalyzes the reaction, was isolated from runner beans and overexpressed in *Arabidopsis* the resultant phenotype was an extreme dwarf (Hedden and Philips, 2000). Expression of a single-chain antibody in tobacco against GA_{24}/GA_{19} provides a non-enzymatic approach to produce a dwarf plant (Shimada *et al.*, 1999).

The later stages in GA bisoynthesis consist of a complex homeostatic mechanism and attempts to manipulate GA status may be hampered by the ability of this homeostatic system to compensate for changes in bioactive GA levels (Hedden and Phillips, 2000b). Itoh *et al.* (2002) explored the possibility of modifying the height of rice plants by suppressing the height controlling gene, D18. Gibberellin (GA) 3 ß-hydroxylase catalyzes the final step in which active GA are synthesized in plants and the rice 3 ß-hydroxylase is encoded by OsGA3ox1 and OsGA3ox2. A semi-dwarf rice plant was produced by producing transgenic rice with antisense of D18 cDNA. OsGA3ox2 is identical to the D18 gene and catalyzes the conversion of GA20 to GA1 and may be a key step for regulating the level of active GA(Itoh *et al.*, 2002). The resultant dwarf had reduced levels of D18 mRNA but a higher expression level of OsGA20ox1 which encodes GA 20-oxidase catalyzing step from GA53 to GA20 in the GA biosynthetic pathway (Itoh *et al.*, 2002). In conclusion it appears that successful manipulation of height can be achieved in rice by manipulation of GA 3 ß-hydroxylase expression through decrease in active GA level (Itoh *et al.*, 2002).

Overexpression of the rice homeobox gene *OSH1* led to altered morphology and hormone levels in transgenic tobacco (*Nicotiana tabacum* L.) and GA_1 was among the hormones which were seriously reduced. It appears that overexpression of OSH1 causes a reduction in the levels of GA_1 by suppressing GA 20-oxidase expression, only treatment with GA 20-oxidase caused a promotion in stem elongation in the transgenic tobacco (Hedden and Phillips, 2000b; Kusaba *et al.*, 1998). The level of GA in rice was modified by the overproduction of the GA catabolic enzyme, GA 2-oxidase. When *OsGA20x1*, the gene encoding GA 2-oxidase, is constitutively overexpressed the transgenic rice showed severe dwarfism and failed to set grain due to the involvement of GA in shoot development and reproductive development. However by expressing *OsGA2ox1* in shoots under the control of the promoter of a GA biosynthesis gene, *OsGA3ox2* (*D18*) produced a plant that was semi dwarf and had normal flowering and grain development. This indicates that it is feasible to genetically improve cereal crops via modulation of GA catabolism and bioactive GA content (Hedden and Phillips, 2000b; Sakamoto *et al.*, 2003). Cole *et al.* (1999) looked at the modification of gibberellin production and plant development in *Arabidopsis* by sense and antisense expression if gibberellin 20-oxidase genes. By overexpressing any of the cDNAs (*AtGA20ox1*, *AtGA20ox2* or *AtGA20ox3*) in transgenic *Arabidopsis* expressing sense or antisense copies of each of the GA 20-oxidase cDNAs gave rise to seedlings with elongated hypotocotyls and were 25% taller at maturity (Cole *et al*. 1999). The vegetative rosettes showed a two- to threefold increase in the level of GA_4, indicating GA 20-oxidase normally limits bioactive GA levels. Plants expressing antisense copies of *AtGA20ox1* had short hypocotyls and reduced rates of stem elongation, also there were reduced levels of GA_4 in the rosettes and shoot tips (Cole *et al*., 1999). This paper further demonstrates that the GA levels in plants can be manipulated affecting plant growth and development which can therefore be modified by manipulation of GA 20-oxidase expression in transgenic plants (Cole *et al*., 1999).

1.6.4 GA 2-oxidase

In rice constitutively expressing the gene encoding GA 2-oxidase, OsGA2ox1 results in transgenic rice with severe dwarfism but also a failure to set grain. Ectopic

expression of OsGA2ox1 at the site of bioactive GA synthesis in shoots under the control of a GA biosynthesis gene, OsGA3ox3 (D18) resulted in a semi dwarf phenotype that is normal in flower and grain development (Sakamoto *et al.*, 2003). The stability and inheritance of these traits leads to the possibility of improving cereal crops by modulation of GA catabolism and bioactive GA content (Sakamoto *et al.*, 2003).

In tobacco, the tobacco protein REPRESSION OF SHOOT GROWTH (RSG), a bZIP transcription factor, binds and activates the transcription of the GA biosynthetic gene *Kaurene Oxidase*. Overproduction of the dominant negative mutant form of RS9 confers a GA-deficient dwarf phenotype and prevents upregulation of GA20ox expression, suggesting that RSG also plays a key role in the feedback regulation of GA20ox (Fleet and Sun, 2005; Ishida *et al.*, 2004).

A dwarf transgenic hybrid poplar (*Populus tremula* x *Populus alba*) has been identified and the dwarf phenotype is the result of a hyperactivated gene encoding GA 2-oxidase (GA2ox) (Buscov *et al.*, 2003). The chromosomal locations of several height reducing genes (which encode proteins in GA signalling) in bread wheat were determined. It is then proposed that the reduced height (*Rht*) genes *Rht-B1b* and *Rht-D1b* are replaced with alternative dwarfing genes for bread wheat improvement (Ellis *et al.*, 2005).

Transgenic approaches have also been employed to introduce extreme dwarfing into wheat via transformation of plants with antisense constructs of GA20ox gene and GA3ox genes or by overexpressing GA2ox gene which is responsible for GA_1 deactivation. However the majority of these are transgenic approaches which consumers are cautious of at the current time hence cannot be employed. Perhaps in time such approaches may be employed. The enzymes in gibberellin biosynthesis and those on the gibberellin signaling pathway have been identified, however gibberellin receptors remain unknown and the proposed receptors may be redundant in their functions (Nakajima *et al.*, 2006).

Work continues to advance in the research area of gibberellins with particular interest in gibberellins 2-oxidase. In *Phaseolus coccineus* PcGA2ox1 (a GA 2-oxidase gene) inactivated bioactive gibberellins by 2ß-hydroxylation thereby reducing the availability of the bioactive gibberellins. By introducing PcGA2ox1 into Solanaceae the transgenic plants produced showed a range of dwarf phenotypes associated with a reduction in the bioactive gibberellins. This approach of transgenics to produce a dwarf stature appears to offer a promising option compared to the use of chemical growth retardants (Dijkstra *et al*., 2008). Members of the rice GA2ox family are differentially regulated and can act in concert or individually to control the GA levels during flowering, tillering and seed germination. Improvements in plant structure via semi-dwarfism, increased root systems and higher tiller numbers can be induced by overexpression of wild-type or modified C_{20} GA2oxs (Fang-Lo *et al*., 2008). The GA2ox7 gene encodes a C20-GA deactivation enzyme and loss of function of GA2ox7 suppressed the dwarf phenotype indicating that GA deficiency was due to overexpression of GA2ox7 (Magome *et al*., 2008). Indeed using gibberellins to manipulate plant architecture may prove to be a promising area as chemical growth regulators to modify GA biosynthesis increase production costs, manpower and environmental risks (Bhattacharya *et al.,* 2010).

1.7 DELLA proteins – role in the GA response

DELLA proteins act as pivotal regulators in the GA cascade, in which they are negative regulators (Fleet and Sun, 2005) in monocots and dicots. DELLA proteins are normally localized in the nuclei to suppress GA signalling and are ubiquitinated in presence of GA and digested by the complex of Skp1-cullin-F-box (SCF) family and proteasomes (McGinnis *et al.*, 2003), it is the disappearance of the DELLA protein which releases the suppression on GA signalling therefore allowing growth (Nakajima *et al.*, 2006).

A mutation in the DELLA domains in wheat and rice led to a reduced height. Plant growth is stimulated via the destruction of the DELLA proteins and gibberellin

signaling uses the ubiquitin-proteasome pathway (ubiquitin-26s proteasome pathway) to control expression through protein degradation (Figure 1.7).

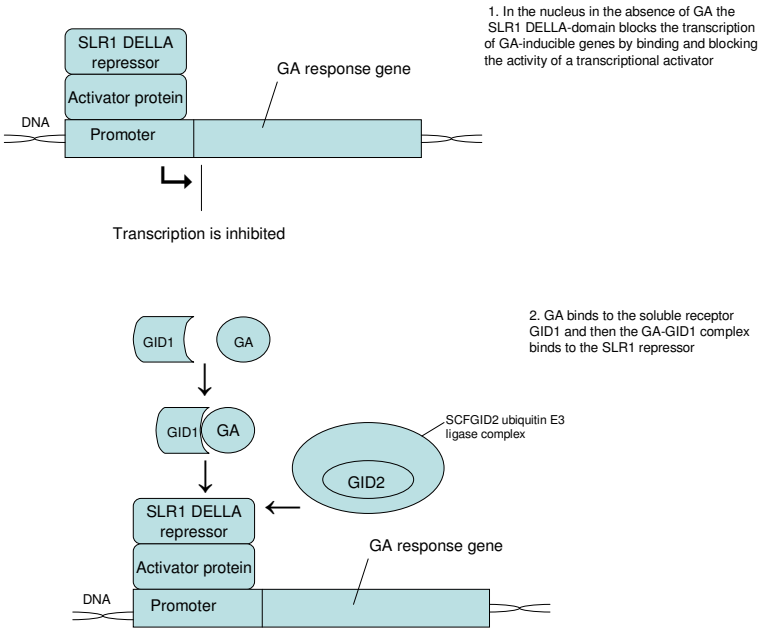


Figure 1.7: Model for GA binding to its receptor and subsequent activation of gene expression in rice (adapted from Taiz and Zeiger, 2006)

Dwarf mutants of rice, wheat, maize and *Arabidopsis* (Table 4) have been identified by their inability to degrade wild type DELLA domain proteins in the presence of bioactive GAs. Therefore for plant growth to be initiated the DELLA protein needs to be degraded and anything which prevents degradation results in a dwarf phenotype (Taiz and Zeiger, 2006). Mutations of or within the DELLA domain of DELLA repressors have been described for species including *Arabidopsis thaliana*, *Triticum aestivum* (wheat), *Zea mays* (maize) and *Hordeum vulgare* (barley) (Willige *et al.*, 2007).

Table 4: Genes responsible for dwarf phenotype

Gene	Function	Trait affected	Species	References
Rht1	Negative regulator of GA response	Reduced height GA insensitive dwarf plants	*Triticum aestivum* (Wheat)	Willege *et al.* 2007
Dwarf8	Negative regulator of GA response	GA insensitive dwarf plants	*Zea mays* (Maize)	Peng *et al.* 1999
Slender1	Negative regulator	Tall slender phenotype, constitutively expressed (overexpression)	*Oryza sativa* (Rice)	Chandler *et al*. 2002
GA – insensitive dwarf 2 (gid2) GID2	Positive regulators of GA	GA insensitive dwarfs	*Arabidopsis Thaliana*	Willige *et al.* 2007
sleepy1 (sly1) SLY1	Positive regulators of GA	GA insensitive dwarfs	*Arabidopsis Thaliana*	Willege *et al.* 2007

1.7.1 Mutations resulting in a dwarf or slender phenotype

There are three types of GA response mutants (Figure 1.8), two of which result in a dwarf phenotype (A and C). The first involves loss of a positive regulator i.e. the F-box proteins which target DELLA degradation and the second mutation to produce a dwarf phenotype is the result of a negative regulator (GA repressor protein) being constitutively expressed. The loss of a negative regulator, DELLA domain repressor proteins results in a tall slender phenotype (D) (Taiz and Zeiger, 2006). Normal growth occurs by degradation of the 26s proteasome (B).

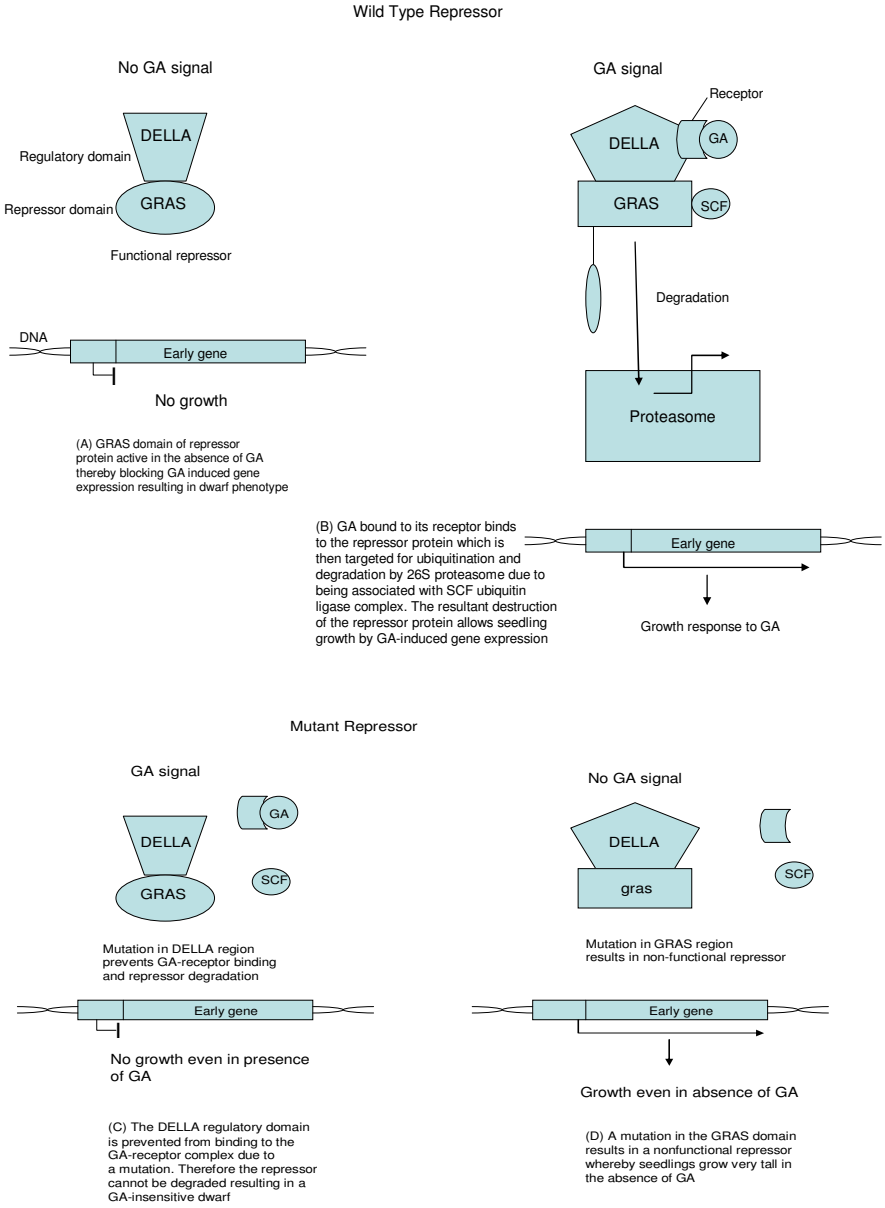


Figure 1.8: GA signalling for growth involves interactions between the GA receptor, an ubiquitin ligase SCF complex and a DELLA-domain repressor protein. The GA repressor (GAI/RGA/SLR1/SLN1) protein contains 2 domains: the regulatory domain (DELLA) and the repressor domain (GRAS). A mutation can result in a dwarf or slender phenotype (adapted from Taiz and Zeiger, 2006)

1.7.2 Interactions between Gibberellins and DELLA proteins

Most work carried out on GA-deficient dwarf mutants have focused on defects in GA biosynthetic enzymes, recent focus has changed to genes which encode GA-catabolising enzymes i.e. GA1-oxidases. When these genes were overexpressed in *Arabidopsis* and poplar the resultant phenotype had decreased height and internode length (Busov *et al.*, 2006; Schomburg *et al.*, 2003). In *Arabidopsis* overexpression of SNEEZY (SNE) which is a homolog of the F-box protein SLY1 rescues the stem elongation of *sly1* mutants through the degradation of DELLA proteins (Strader *et al.*, 2004).

Gibberellin receptors in *Arabidopsis* have been characterized as AtGID1a, AtGID1b and AtGID1c and their recombinant proteins have a high affinity to GA_4 compared to the other GAs (Nakajima *et al.*, 2006). *Arabidopsis* has five DELLA proteins (RGA, GAI, RGL1, RGL2 and RGL3) but no GA receptor mutant has been identified implying the possible redundancy of GA receptors in this plant (Nakajima *et al.*, 2006). GAI and RGA are coupled together to reduce stem elongation growth (Hussain and Peng, 2003). A yeast two hybrid system showed an in vivo interaction between AtGID1 and the *Arabidopsis* DELLA proteins (AtDELLAs) which are negative regulators of GA signaling in the presence of GA_4 (Nakajima *et al.*, 2006). Expression of each AtGID1 clone in the rice gid1-1 mutant rescued the GA insensitive dwarf phenotype (Nakajima *et al.*, 2006). It has been proposed that a protein encoded by rice (*Oryza sativa*), GA insensitive Dwarf1 (OsGID1) could be characterized as a GA receptor ((Ueguchi-Tanaka *et al.*, 2005). The OsGID1 gene encodes a protein which possesses GA-binding activity and its mutation results in a severe dwarf phenotype which does not respond to GA in stem elongation. The mutated gid1 gene has lost GA-binding activity which is essential for the interaction between OsGID1 and GA for normal plant growth. (Ueguchi-Tanaka *et al.*, 2005). OsGID1 interacts with SLR1, a rice DELLA protein of GRAS family and negative regulator of GA signaling (Ueguchi-Tanaka *et al.*, 2005)

A substitution of a conserved amino acid in the C-terminal domain of a DELLA protein causes a dwarf mutant in *Brassica rapa* (Brrga1-d). The mutated amino acid causes dwarfism by preventing an interaction with a protein component required for its degradation. This mutation may be potentially useful in crop species as it confers a nondeleterious dwarf phenotype when transferred to *Arabidopsis* and oilseed rape (Muangprom *et al.*, 2005).

Research continues into the role of DELLAs, in rice the SLR1 DELLA protein needs to be degraded for shoot elongation and seed germination to be initiated (Ueguchi-Tanaka *et al*., 2008). Also the role of *SLEEPY* and *SNEEZY* in *Arabidopsis* signaling was investigated and concluded that the model SLY1 regulates DELLA interaction through interaction with the DELLA-GA-GID1 complex (Arrizumi *et al*., 2011). Promising research continues in the field of manipulating plant height through gibberellins and DELLA proteins.

Gibberellins also play a role in cell elongation in hypocotyls of light grown plants, as well as stems, and ethylene, auxin and brassinolide have synergistic effects with GA promoting cell elongation and hypocotyl length in *Arabidopsis* (Fleet and Sun, 2005; Saibo *et al.*, 2003; Symons and Reid, 2003; Tanaka *et al.*, 2003). Plants adaptively optimise growth and development in ways appropriate to the environmental conditions by sensing specific characteristics such as the surrounding light environment. These characteristics include light quality, intensity and duration of the exposure. It is the light trigger which inhibits seedling hypocotyl growth through the activation of phytochromes and additional photoreceptors. It is evident that light inhibition of *Arabidopsis* hypocotyl growth is in part dependent on DELLAs (family of nuclear growth-restraining proteins that mediate effect of phytochrome gibberellin (GA) on growth) (Achard *et al.*, 2007) . This is supported by the findings that the light inhibition of growth is reduced in DELLA-deficient hypocotyls, light activation of phytochromes promotes the accumulation of DELLAs, GFP tagged DELLA (GFP-RGA) accumulates in the elongating cells of light grown but not dark grown hypocotyls and the transfer of seedlings from light to dark (or vice versa) leads to rapid changes in hypocotyl GFP-RGA accumulation which is paralleled by rapid

alterations in abundance in hypocotyl or transcripts encoding enzymes of GA metabolism (Achard *et al.*, 2007).

Light quality signals enable plants to determine potential neighbour competitors and respond with shade avoidance responses (Djakovic-Petrovic *et al.*, 2007). Amongst the shade avoidance responses, and of particular interest, is increased shoot elongation which enables the plant to have a competitive advantage over its neighbours due to enhanced light capture (Djakovic-Petrovic *et al.*, 2007). Phytochrome B photoreceptors which sense red:far red ratio and the hormone gibberellin interact to coordinate the shade-avoidance responses. The growth suppressing DELLA proteins are targets for GA signaling and are proposed to integrate signals from other hormones (Djakovic-Petrovic *et al.*, 2007). DELLA abundance is regulated during growth responses to neighbours in dense *Arabidopsis* stands and occurs in a red:far-red dependent manner in petioles. With DELLA knockout mutants, observed DELLA breakdown was not sufficient to induce shade-avoidance in petioles (Djakovic-Petrovic *et al.*, 2007). Also, a lack of DELLA breakdown in the *gai* mutant which carries a deletion in the DELLA motif making the GAI protein unrecognizable for GA signaling prevents elongation (Djakovic-Petrovic *et al.*, 2007). Therefore, DELLA proteins, which act as gibberellin signal transduction components play an important role in detecting canopy light signals; DELLA and GA in part regulate the shade-avoidance response (Djakovic-Petrovic *et al.*, 2007).

Photomorphogenic (light-grown) seedlings have shorter hypocotyls compared to dark-grown seedlings and the cotyledons are green and fully expanded. A decrease in the levels of GA result in a partially photomorphogenic phenotype and an increase in the expression of light induced genes, which can be rescued by the loss-of-function *rga* and *gai* mutations which indicate that RGA and GAI repress GA's inhibition of photomorphogenesis in etiolated seedlings (Achard *et al.*, 2003; Alabadi *et al.*, 2004; Fleet and Sun, 2005; Saibo *et al.*, 2003). It appears that red light perception induces photomorphogenesis during de-etiolation and acts partially through GA. Red light treatment of seedlings reduces GA sensitivity in many species

and blue light may also influence GA biosynthesis in a manner that promotes de-etiolation (Fleet and Sun, 2005; Folta *et al.*, 2003; Saibo *et al.*, 2003; Symons and Reid, 2003).

The shoot apical meristem (SAM) is also maintained via exclusion of GA from the interior cells. This is maintained via the suppression of GA biosynthesis by KNOX homeodomain transcription factors which are encoded by KNOTTED 1-LIKE HOMEOBOX (KNOX) genes (Sakamoto *et al.*, 2001).

1.7.3 Possible uses of gibberellins to manipulate height

There is still much to be revealed regarding the role of GA in plant growth and morphogenesis. Further investigation into molecular mechanisms and possible hormone cross-talk will hopefully reveal how GA and other hormone pathways interact to control plant morphogenesis in response to a variety of developmental and environmental cues (Fleet and Sun, 2005; Folta *et al.*, 2003) and the potential to manipulate these processes.

Gibberellins appear to play important roles in the growth of plants and therefore of particular interest are GA2-oxidase, GA3-oxidase and GA20-oxidase. The latter half of this chapter has outlined a large area of research in which *Arabidopsis* is one of the plants which the gibberellin response has been researched and manipulation of gibberellin status has affected the phenotype. As mentioned earlier both *Arabidopsis* and watercress belong to the Brassicaceae family and information gained regarding the effects of gibberellins in *Arabidopsis* may be adapted to possible responses in watercress. It appears that the key targets in order to produce dwarf watercress are the gibberellins, thereby expression of the genes relating to gibberellin biosynthesis (GA3-oxidase or GA20-oxidase) or catabolism (GA2-oxidase) required attention. Also DELLA proteins have been reported to be negative regulators of the GA response and would provide another line of investigation in order to achieve dwarf watercress. A beneficial effect from a variation in gibberellins may prove to be a

dwarf phenotype, whilst in terms of the ultimate goal in the breeding programme a deleterious effect would be a spindly tall phenotype. These are two characteristic phenotypes resulting from a low level of gibberellin biosysnthesis either by an overexpression in GA2-oxidase or antisense GA20-0xidase and GA3-oxidase. A high level of gibberellin could be the result of antisense GA2-oxidase and over-expression of GA20-oxidase or GA3-oxidase.

1.8 Aims and Objectives

The over-arching aim of this research is to provide pre-breeding information on watercress to underpin future breeding efforts for morphology and PEITC. The objectives of the following research is to identify any potential genetic variation held within a watercress germplasm collection with lines sourced from around the world which would aid in establishing a superior cultivar.

1. Establish a worldwide watercress germplasm collection at the University of Southampton.
2. Identify any morphological variation present in the collection with different watercress lines established in a controlled environment and the field.
3. Identify any variation present in the antioxidant power of the different watercress lines established in a controlled environment and in the field and identify differences in individual glucosinolates and isothiocyanates e.g. PEITC concentration.
4. Determine the level of genetic diversity of the University of Southampton watercress germplasm collection.
5. Identify genes of interest to be taken forward to investigate any potential variation in gene expression between lines of particular interest.
6. Initiate preliminary crosses moving towards breeding.

1.9 Thesis overview

The main objective of this study is to explore any underlying variation in an understudied popular crop, watercress, which could be taken forward to initiate a watercress breeding programme. By establishing a global germplasm collection (Chapter two) any phenotypic, biochemical or genetic variation could be investigated. Chapter two focuses on any variation present in the phenotype – namely stem length, stem diameter and leaf to stem ratio. This allowed information on lines with favourable traits i.e. a reduced stem length to be taken forward into the watercress breeding programme (Chapter four). Also by looking at the phenotype any intriguing lines can be selected to investigate the underlying genotype which leads to the resulting phenotype (Chapter five).

The variation in the phenotype observed in Chapter two led to further investigation into any variation in the level of antioxidants (Chapter three). Chapter three initially inspected any variation in the level of antioxidants by use of a simple antioxidant assay, FRAP, which would allow lines to be taken forward for further analysis into glucosinolate and isothiocyanate concentrations. Glucosinolates were selected due to their effect on taste i.e. bitterness as well as the well documented antioxidant benefits. The glucosinolate and isothiocycanate obtained would allow any relationship between glucosinolate and isothiocyanate concentration to be identified and lines with high isothiocyanate concentrations to be selected for the breeding programme. Chapter three also focuses on possible agronomic methods to improve the levels of antioxidants in particular; section of plant, repeated harvest and shelf life. Therefore as well as examining any natural variation within the watercress collection, variation due to agronomy was investigated in this chapter for manipulation to increase the levels of antioxidants. Data extrapolated from this chapter would enable lines to be taken forward into the breeding programme (Chapter four) and finally exploring any possible genetic variation for differences in antioxidant levels (Chapter five).

Chapter four utilises Amplified Fragment Length Polymorphism to deduce any genetic diversity between the watercress lines present in the collection and finally Chapter five uses microarray analysis to explore whether variation in gene expression could explain differences observed in Chapters two and three. Chapter four contributes to this research by exploring any natural diversity present in the collection which may be manipulated for an enhanced crop whilst Chapter five allows the expression of individual genes to be explored for the first time in watercress.

CHAPTER TWO: MORPHOLOGICAL VARIATION WITHIN THE GERMPLASM COLLECTION

2.1 Abstract

Plant height not only affects plant architecture (three dimensional organization of the part of a plant that is above ground (Yang and Hwa, 2008) but is also an important agronomic trait which can be manipulated to improve crop yield (Wang and Li, 2006). The green revolution saw the introduction of many dwarf varieties due to the large number of benefits obtained from dwarfism. Much work has focused on isolating and characterising genes which are involved in the formation of plant architecture, of particular interest are the genes controlling stem elongation. The conservation of these genes is of agronomical importance with respect to improving crop yields (Wang and Li, 2006). Genes encoding gibberellic acid (GA) and brassinosteroid (BR) biosynthetic pathways have been identified giving wide acceptance that GA and BR are two major factors which determine plant height (Wang and Li, 2006).

In order to efficiently manage germplasm collections and breeding programmes the genetic diversity between individuals, populations and gene pools is essential (Yang *et al.*, 2006). Diversity within the collection may lead to breeding a new line with desired characteristics. Differences between the stem length, stem diameter and number of leaves were apparent within the watercress germplasm collection which allowed lines to be selected for future crossing. Of particular interest was the ability to manipulate the stem length. Significant variation was evident in the stem length and stem diameter between the different lines when grown in a controlled environment, ranging from an average of 7cm-17cm and 1.58mm-3.18mm respectively. The results presented here illustrate significant variation in stem length which allowed two lines to be selected for crossing in the breeding programme.

2.2 Introduction

Dwarfism is a desirable trait for many agricultural crops. A crop with a reduced stem length produces a stable increased yield and is easier to harvest as well as package. Dwarf mutants have been extensively characterised in many plant species and endogenous phytohormone, Gibberellin (GA) is associated with the dwarf phenotype. The three candidate genes of particular interest are: GA2ox, GA3ox and GA20ox. GA20ox and GA3ox are involved in GA biosynthesis whilst GA2ox is involved in GA catabolism.

Understanding the mechanisms by which plant height is controlled is not only important in the physiology of the plant but can also play a major role in breeding programmes and agriculture (Ishimaru *et al.*, 2004). Niklas and Enquist (2001) reported a strong relationship between plant height and biomass production in a number of plant species, these were representive species of unicellular algae, aquatic ferns and a wide selection of dicot, monocot and conifer species (Niklas and Enquist, 2001).

Arabidopsis thaliana provided a means whereby the green revolution genes could be identified and cloned. *A. thaliana* is a close relative of the Brassica species and comparative genome analysis between these two can be used to transfer information and resources from this important model organism to agriculturally important crop species (Muangprom and Osborn, 2004). The coding regions in *A. thaliana* and Brassica are highly conserved with Cavell *et al*. (1998) reporting a 85% sequence similarity (Muangprom and Osborn, 2004) The gene order is conserved between species for large segments of chromosomes (Muangprom and Osborn, 2004) This strengthens the possibility of identifying candidate genes of agronomic importance across the Brassica species. As of February 2010 a Brassica array became available at the Nottingham Arabidopsis Stock Centre (NASC). On a Brassica EST microarray made available from John Innes centre, 77% of the protein hits came from *Arabidopsis thaliana*.

Gibberellins play important roles in many different physiological processes such as growth and development including: seed germination, shoot/stem elongation and flower development (Ikeda *et al.*, 2001). Any changes in the tissue sensitivity or GA concentration affect these processes (Ikeda *et al.*, 2001). Mutants which are impaired in GA biosynthesis or response have small, dark green leaves and also a reduced stem length. They are also defective in seed germination and floral development and have a delayed flowering time (Fleet and Sun, 2005). On the other hand, plants which have increased GA have a tall and spindly phenotype (Fleet and Sun, 2005). The response mutants are split into two different phenotypic characters in which the slender mutants have constitutive activation of their GA response whilst the dwarf mutants are deficient in GA detection or signal transduction (Ikeda *et al.*, 2001). This provides two possible routes for the manipulation of plant height.

The DELLA domain is important for GA detection in the GA signal transduction pathway (Ikeda *et al.*, 2001). GA signalling operates as a de- repressible system that is moderated by DELLA domain proteins which act as transcriptional regulators thereby repressing GA responses (Fleet and Sun, 2005). DELLA proteins are highly conserved among different species including *Arabidopsis*, *Brassica*, rice and wheat (Fleet and Sun, 2005) Overexpression of the GA catabolising enzymes e.g. GA2 oxidase results in a decreased height and internode length which has been demonstrated in many species such as *Arabidopsis* and Poplar. This highlights the importance that GA plays in regulating plant height (Fleet and Sun, 2005) A single recessive mutation causes the recessive *slender* rice which is considered to be a constitutive GA response phenotype as the mutant phenotype was twofold taller than the wild type and wild type shoots treated with GA3 had a similar phenotype to the *slender* rice (Ikeda *et al.*, 2001). In *Brassica rapa* *d*wf2 is insensitive to exogenous GA3 for both plant height and flowering time, again inferring that it is not a mutation in the gibberellin biosynthesis pathway (Muangprom and Osborn, 2004). The opposite of the dwarf phenotype is the elongated phenotype. In *Brassica rapa* a single gene mutant (elongated internode [*ein*/*ein*] which results in accelerated shoot elongation. When this mutant is compared to *Brassica rapa* without the mutation an increased internode elongation is apparent (Rood *et al.*, 1990).

The aim was to establish any morphological variation between the watercress lines when grown firstly in a controlled environment and then in the field. The controlled environment would allow a base line for variation between the lines (i.e. genetic variation to become apparent) and by expanding to the field the impact of environmental factors is taken into consideration.

A preliminary experiment was conducted before the lines were planted in order to assess the growth media in which to propagate the watercress lines. Watercress is usually cultivated in slow flowing chalk streams however for research purposes this is not feasible and therefore an ulterior method of cultivation with a standardized growth media and pot size was required. Habegger (1989) investigated how different pot sizes and different substrates affected the growth of watercress. The experiment revealed that the optimum substrate used was a mixture of TKS 1 and humosoil whilst smaller sized pots were more efficient in production (Habegger, 1989). The watercress propagated in the smaller pots had thinner stems and higher shares of leaves and consequently lower nitrate content in contrast to the watercress grown in larger pots which had thicker stems and higher nitrate levels which result in a poor quality plant (Habegger, 1989). Soil structure affects plant growth in many ways, not only does it affect the ability of the roots to grow and supply water and nutrients to leaves but if adverse it can also induce hormonal signals to be sent which slow the growth of the shoot even if the plant is able to take up suitable water and nutrients (Passioura, 1991). The growth and distribution of plants are strongly influenced by environmental factors such as temperature and the availability of nutrients and light (Going *et al.*, 2008). Therefore a constant environment is an important factor to take into consideration when conducting investigations into any variation between different lines.

The aim of this Chapter was to explore the visible phenotypic differences within the watercress germplasm collection with respect to stem length, stem diameter and number of leaves. This would provide information on the lines with favourable traits i.e. a reduced stem length to progress into the watercress breeding programme (Chapter four). Also by examining the phenotype it allows the lines to be chosen to

later investigate the underlying genetics which lead to the resulting phenotype (Chapter five).

2.2.1 Aims of chapter

1. Establish watercress lines in a controlled environment and field.

2. Assess the stem length, stem diameter and leaf to stem ratio to deduce any phenotypic variation between the watercress lines and report any variation deduced.

3. Select watercress lines with a low stem length for the watercress breeding programme

2.3 Materials and Methods

2.3.1 Cultivation of watercress in different growth media

Seeds were cultivated in six different growth media based on a 1:1 ratio;

1) Levingtons (pH 5.3-5.7, mg/litre added; N-72, P-40 and K-120) and vermiculite

2) Levingtons (pH 5.3-5.7, mg/litre added; N-72, P-40 and K-120) and silver sand

3) Levingtons (pH 5.3-5.7, mg/litre added; N-72, P-40 and K-120)

4) Levingtons (pH 5.3-5.7, mg/litre added; N-72, P-40 and K-120) , vermiculite and John Innes no2 (pH 6.5, contains; ground limestone, hoof & horn, superphosphate and potassium sulphate)

5) Levingtons (pH 5.3-5.7, mg/litre added; N-72, P-40 and K-120) and coarse sand

6) Vitacress peat (pH 6, mg/litre added; N-144, P-154 and K-264)

Vitacress stock seeds were sown in 7.5 cm X 7.5cm pots containing six different soil media with 10 replicates per growth media. The pots were randomised and

maintained with an inch of water and monitored daily. The watercress was cultivated under standard glasshouse conditions: 20°C day/15°C night, 12hr day, 06:00-18:00

2.3.2 Pots versus trays

Twenty Vitacress stock seeds were sown in separate trays in the six different growth media listed above. Twenty seeds were sown per tray and the watercress was cultivated under standard glasshouse conditions: 20°C day/15°C night, 12hr day, 06:00-18:00.

2.3.3 Plant material

Seeds were collected from commercial local growers and donated from HRI, Warwick. The seed provided from Warwick HRI was sourced from around the world (Table 5). Half of the seed for each watercress line was stored at -20°C whilst the remaining half was stored at room temperature for assessment. Forty three different watercress lines are present in the collection including watercress seed sourced from Brazil, Italy, France, Denmark and Great Britain.

Table 5: Watercress lines within germplasm collection

Watercress Line	Details
Wx_0001	Vitacress, Great Britain
Wx_0002	Vitacress – French, France
Wx_0003	Maxwell Cress, Alresford, Great Britain
Wx_0004	Maxwell Cress, Alresford, Dark watercress, Great Britain
Wx_0005	Peter Mills, Great Britain
Wx_0006	Peter Mills, Great Britain
Wx_0007	Cricklade nurseries, Great Britain
Wx_0008	John Watts, Great Britain
Wx_0009	John Watts, Great Britain
Wx_0010	John Watts, Great Britain
Wx_0011	John Watts, Great Britain
Wx_0012	John Watts, Great Britain
Wx_0013	John Watts, Great Britain
Wx_0014	John Watts, Great Britain
Wx_0015	John Watts, 1983, Great Britain
Wx_0016	John Watts, Great Britain
Wx_0017	Steve Rothwell, Dark green, Great Britain
Wx_0018	Steve Rothwell, Great Britain
Wx_0019	T.W.Jesty & Partners (Sylvasprings), round leaf and dark green in mid 60s, Great Britain
Wx_0020	De Corato Sementi, Andria, Italy
Wx_0021	Zorz Sementi Via P. Barozzi, Padova, Italy
Wx_0022	Harris seeds, New York, improved broad leaved, United States of America
Wx_0023	Atlee Burpee Warminster, PA, United States of America
Wx_0024	Ben Loyd Bedwyn Stone Museum, Marlborough Wiltshire, Great Britain
Wx_0025	Agroceres Brazil, Agriao - Fohlha Larga, Brazil
Wx_0026	Agroceres Brazil, Agriao D'Agua Fohla Larga, Brazil
Wx_0027	Agroceres Brazil, Agriao D'Agua Fohla Larga, Brazil
Wx_0028	Royal Sluis, large leaves, Holland
Wx_0029	Daehnfeldt, Denmark, ordinary watercress, Denmark
Wx_0030	Daehnfeldt, Denmark, broadleaved watercress, Denmark
Wx_0031	Vilmorin, France, Cresson De Fontaine A Large Feuille, France
Wx_0032	Carl Sperling FRG, Brunnenkresse, Germany
Wx_0033	Bliss New Zealand strain, New Zealand
Wx_0034	5 S Adas line, Unknown
Wx_0035	J Jesty's SRG, Great Britain
Wx_0036	WC4, Unknown
Wx_0037	WC115, Unknown
Wx_0038	WC998, Unknown
Wx_0039	Hairspring watercress, Great Britain
Wx_0040	Vitacress (lot.22625) American land cress, Great Britain
Wx_0041	Vitacress (lot.22637) Variegated American land cress, Great Britain
Wx_0042	Vitacress (lot.23494) American land cress, Great Britain
Wx_0043	Thompson and Morgon, Great Britain
Wx_0044	Mr Fothergills, Great Britain
Wx_0045	Johnsons, Great Britain
Wx_0048	Cook's Garden, Great Britain

2.3.4 Cultivation of lines

Seeds were initially germinated in Petri dishes with lids but not sealed and two layers of moistened filter paper and then individually transferred to Vitacress peat in 7.5cm pots. The pots were then randomised and placed in large trays which were maintained with an inch of water and checked on a daily basis. The watercress lines were raised in a controlled growth room environment, 25°C day/22°C night, 16 hour day, 06:00-22:00. After 7 weeks of growth three different morphological measurements (stem length, diameter and number of leaves) were recorded for each plant. This would quantify the genetic variation in the collection and identify parent material for future crossing experiments. Not all lines germinated (maybe due to seed age, storage or numerous other factors) and therefore not all of the watercress lines have phenotypic (Chapter two, current chapter) and biochemical (Chapter three) measurements taken.

2.3.4.1 Cultivation of the watercress lines in the field, Spetisbury (Dorset)

Lines were initially germinated in 53 X 32 cm (32 squares by 18 squares) sowing trays kept in a polytunnel located in Spetisbury, Dorset (Figure 2.1) with continual misting and then transferred to a watercress bed which is gravel lined with free flowing spring water. The bed was divided into sections to ensure no mixing of the different lines occurred (Figure 2.2).

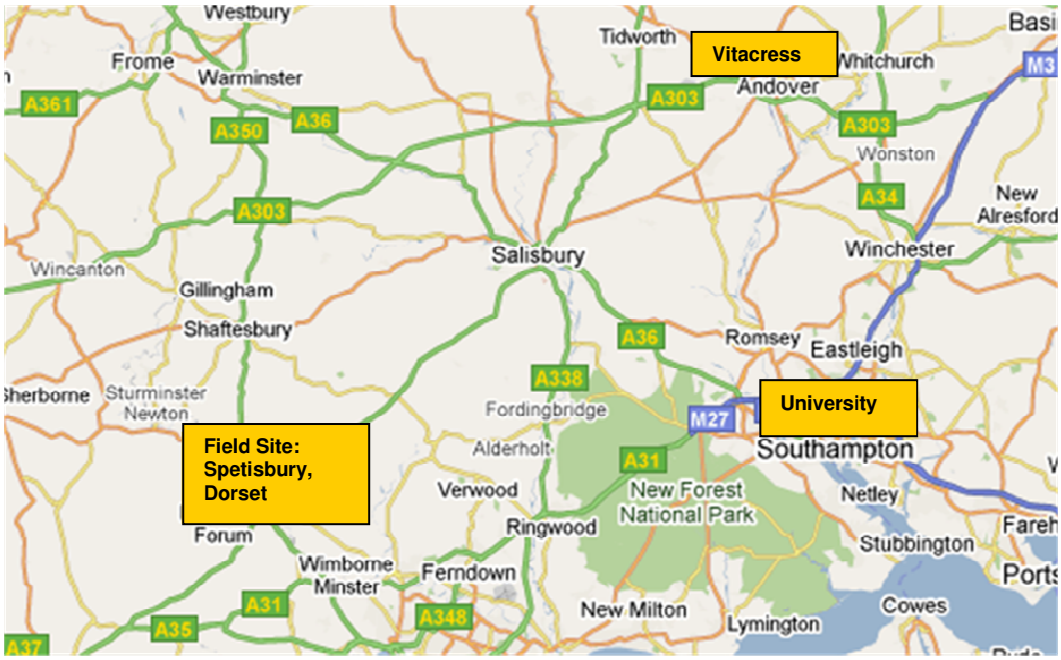


Figure 2.1: Location of Spetisbury field site, main Vitacress site and University of Southampton. The field site, Spetisbury, is where the lines were cultivated.

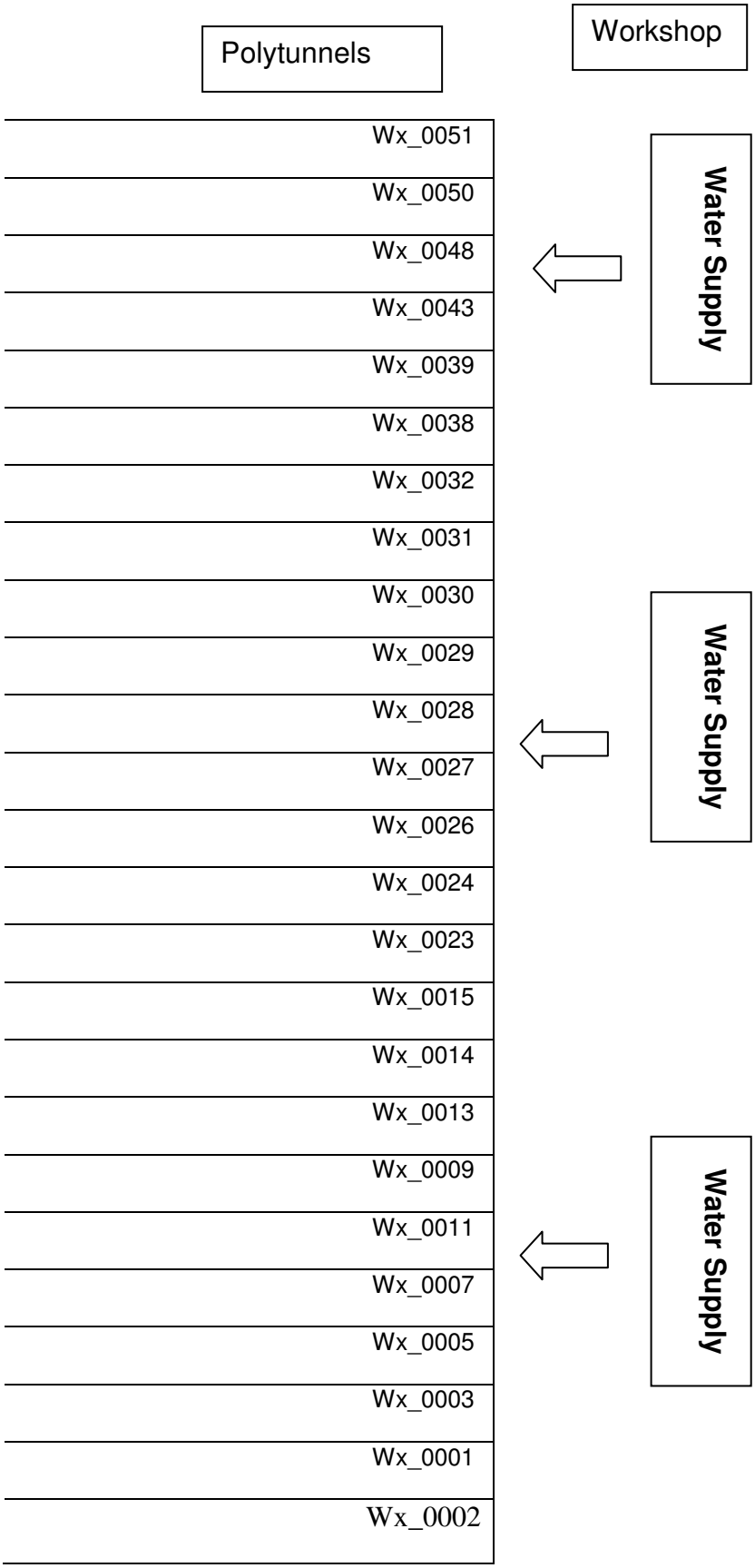


Figure 2.2: Watercress bed plan for cultivation of lines in the field

2.3.4.2 Stem length, diameter and number of leaves

The length of the longest stem on each plant was recorded in cm and the stem diameter at 5cm up the stem was recorded for each plant after seven weeks. The numbers of fully expanded leaves (leaves larger than 2cm by 2cm) were recorded. Five biological replicates were used for each watercress line both in the controlled environment and field; lines grown in the controlled environment were randomised (using Minitab 14 for random design).

2.4 Results

2.4.1 Growth media

A significant difference was observed between the stem length of the Vitacress stock in different growth media ($F_{5,54} = 2.94$, $p<0.05$). The longest stem length was recorded from watercress grown in Levingtons and silver sand with a mean stem length of 19.7cm whilst the mean stem length in vitacress peat was 17cm (Figure 2.3a). The thickness of the stem diameter was significantly different between the watercress in different growth media ($F_{5,54} = 2.51$, $p<0.05$). The thickest stem diameter was recorded from the watercress grown in the vitacress peat and the shortest obtained from levingtons and silver sand (2.44mm and 1.77mm were observed respectively) (Figure 2.3b). Interestingly these results correlate as the thickest mean stem diameter was observed with the longest mean stem length. The vitacress peat also produced the watercress with the highest mean number of leaves (16 per plant) (Figure 2.3c).

2.4.2 Pots versus trays

The watercress was grown in trays with different growth media to assess a more efficicient protocol. Data was statistically analyzed using one way ANOVA. The

Vitacress stock differed significantly in stem length ($F_{5, 113} = 6.87$, $p<0.05$), stem diameter ($F_{5, 115} = 10.26$, $p<0.05$) and number of leaves ($F_{5, 113} = 4.32$, $p<0.05$).

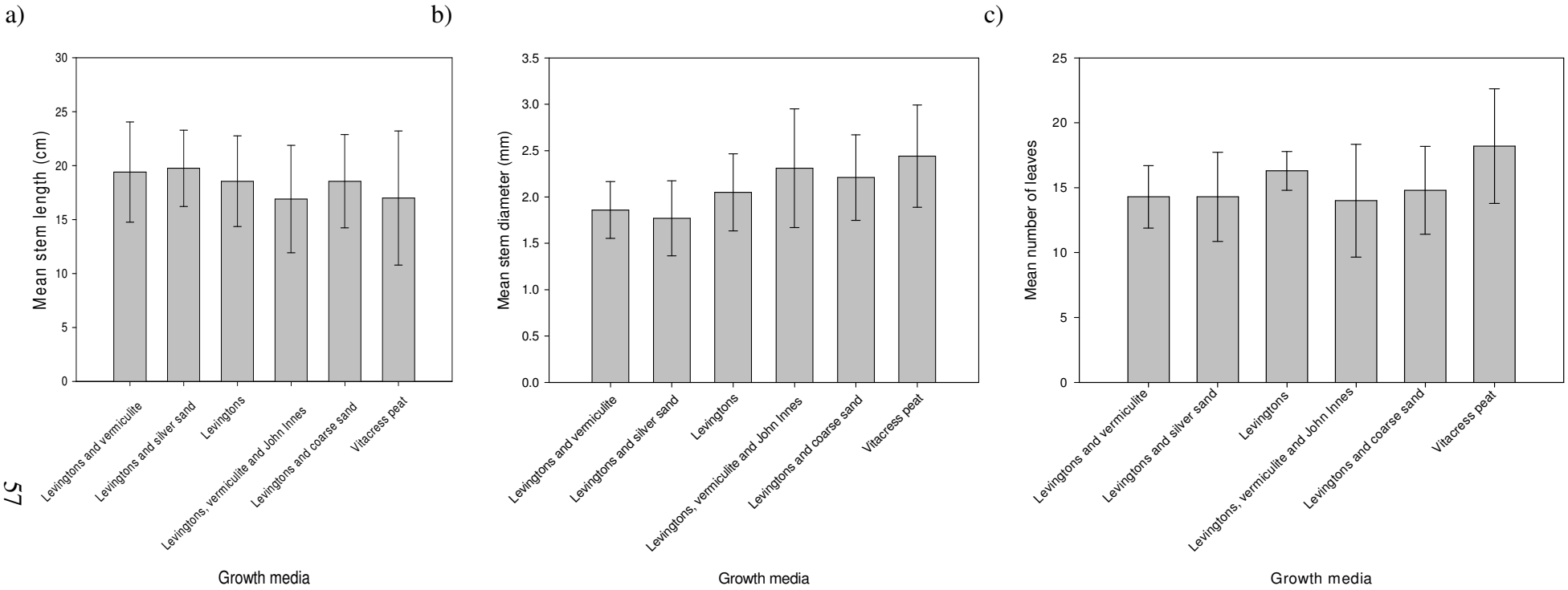


Figure 2.3: Growth trials comparing soils; Mean stem length (±SE) ($F_{5,54}$ = 2.94, p<0.05) (a) stem diameter (±SE) ($F_{5,54}$ = 2.51, p<0.05) (b) and number of leaves (±SE) (c) of Wx_0001 grown in six different growth media.

2.4.3 Stem length

The stem length in the controlled environment significantly differed between the different watercress lines. A line from the USA (Wx_0023) had the lowest mean stem length (7cm), while the highest mean stem length was recorded from a John Watts line (Wx_0011, 26.6cm) (Figure 2.4a). When grown in the field the lowest mean stem length was Wx_0005 (9.5cm) whilst the highest mean stem length was line Wx_0026 (17.1cm) (Figure 2.4b). Not all the watercress lines germinated in the field which caused only a few lines to be compared. In the field Wx_0023 still remains one of the lowest and Wx_0001 remains consistently in the middle in both the controlled environment and field. Also Wx_0002 remains at the higher end of stem lengths in both the controlled environment and field.

When the controlled environment grown watercress are compared to the field grown watercress there is a distinct difference (Figure 2.4c). The lines grown in a controlled environment have a higher stem length when compared to the field grown lines. When plotted in a frequency diagram the stem lengths of the watercress lines grown in the field cluster around 10-12cm and also the 13-14cm grouping whilst the lines grown in the controlled environment cluster at 15-20cm. It is evident in the field the watercress lines develop a longer mean stem length.

There was a significant difference between the watercress lines ($F_{11,96} = 9.05$, $p<0.001$). The condition i.e. controlled environment and field significantly affected the stem length ($F_{1, 96} = 28.16$, $p<0.001$) and there is a significant interaction between the watercress line and condition ($F_{11, 96} = 7.232$, $p<0.001$).

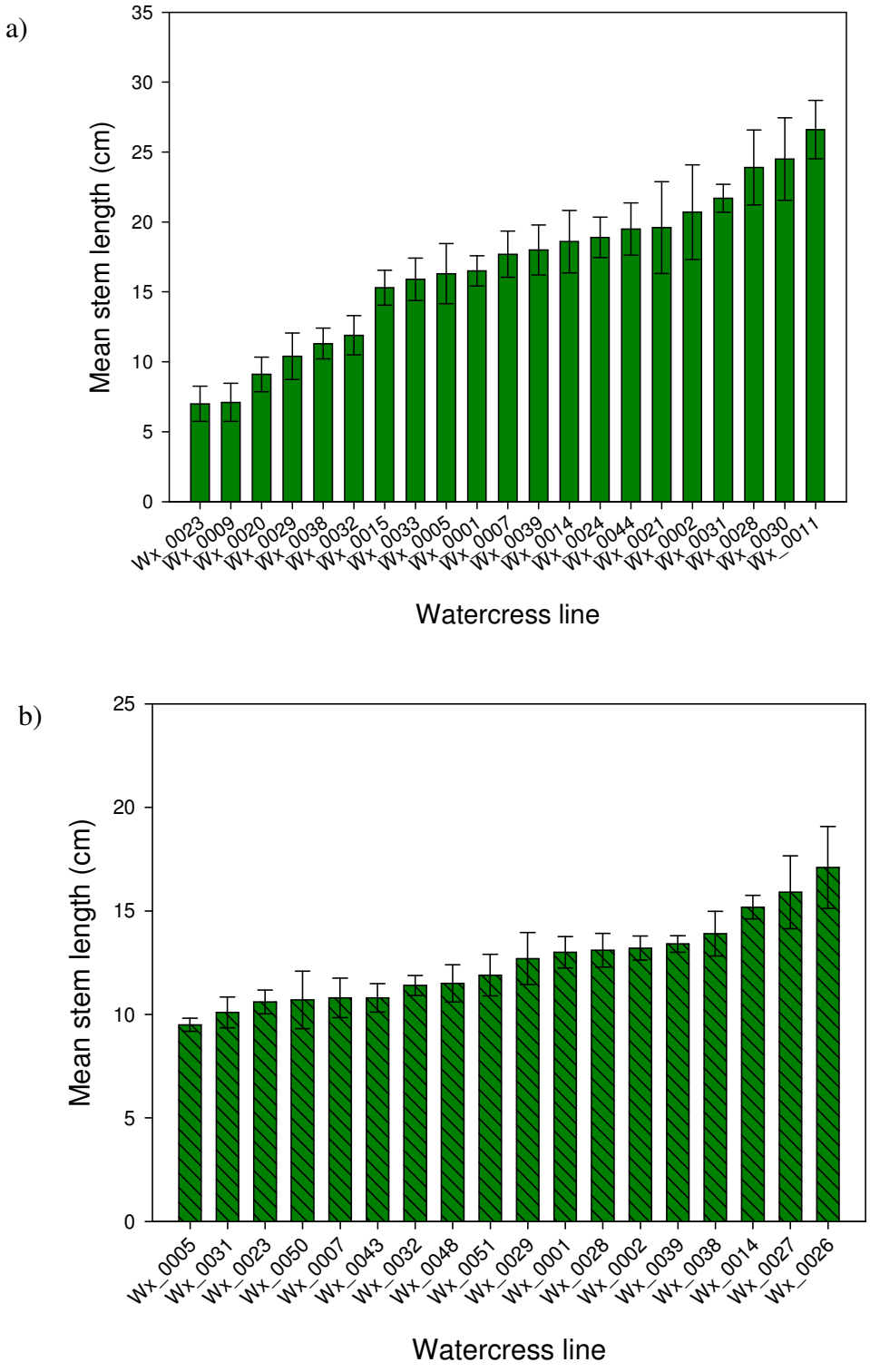


Figure 2.4: Variation in the mean stem length (±SE) of watercress lines (a) grown in a controlled environment and (b) in the field-Spetisbury, Dorset (c) Comparison of the variation in the mean stem length of watercress lines grown in a controlled environment and in the field – Spetisbury, Dorset. There was a significant difference between the watercress lines ($F_{11,96} = 9.05$, $p<0.001$) and the condition i.e. controlled environment and field significantly affected the stem length ($F_{1, 96} = 28.16$, $p<0.001$). Standard error bars present on figure. NB. Fewer lines successfully germinated in the field compared to the controlled environment.

c)

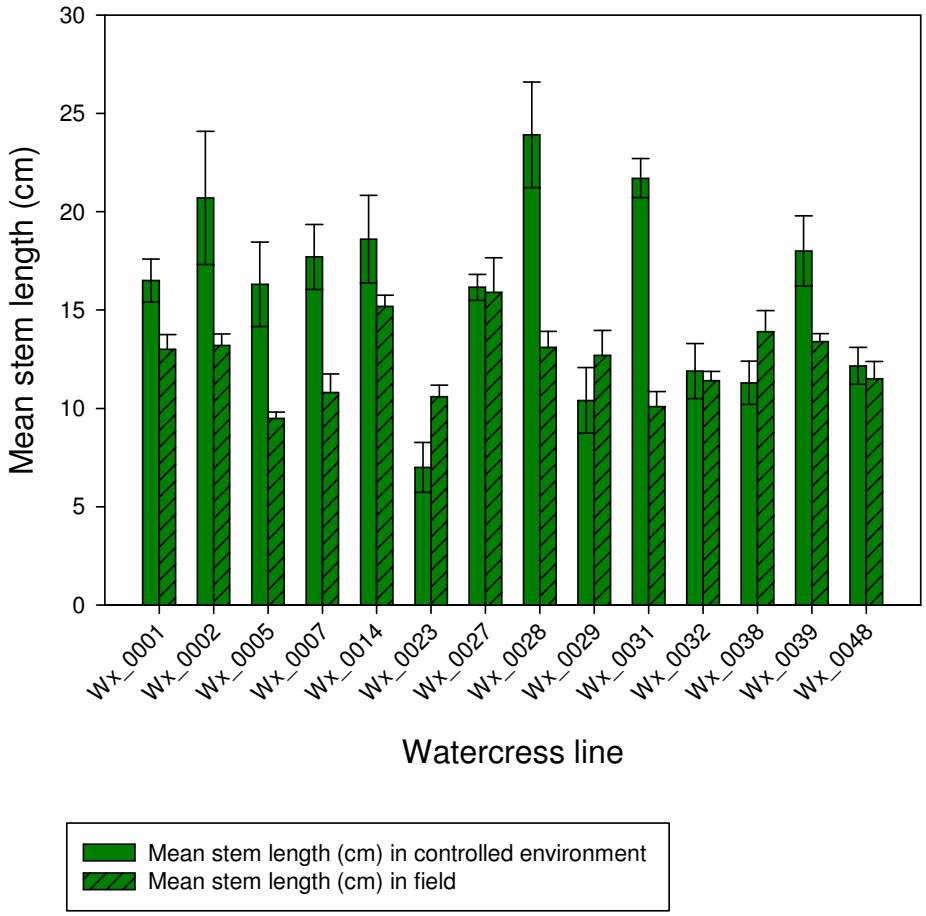


Figure 2.4 Continued

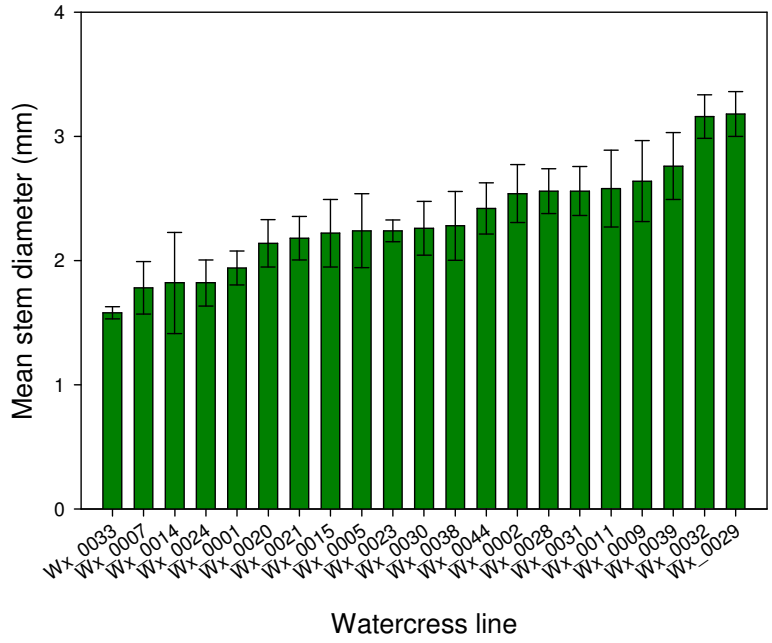
2.4.4 Stem diameter

Stem diameter also significantly differed across the watercress lines in the controlled environment. The smallest stem diameter was observed in a Bliss New Zealand line (Wx_0033, 1.58mm) and the highest in a line from Denmark (Wx_0029, 3.18mm) (Figure 2.5a). When grown in the field Wx_0028 had the lowest mean stem diameter (1.86mm) whilst Wx_0026 the highest (3.4mm) (Figure 2.5b). It is interesting to observe that both in the controlled environment and field Wx_0001 remains to have one of the lowest stem diameters. However Wx_0007 has the second lowest stem diameter in the controlled environment yet has the second highest stem diameter in the field.

The majority of the lines when grown in a controlled environment have a higher mean stem diameter compared to when the lines are grown in the field (Figure 2.5c). When data is presented as a frequency diagram the watercress lines have the highest frequency in 2.0mm to 2.2mm whilst those lines grown in the controlled environment have the highest frequency 2.2-2.6mm.

The stem diameter between the watercress lines significantly differed ($F_{11, 96} = 2.35$, $p<0.05$) but there was no significant difference on stem diameter as a result of condition ($F_{1, 96} = 0.02$). There was a significant interaction between the watercress line and condition ($F_{11, 96} = 4.33$, $p<0.001$).

a)



b)

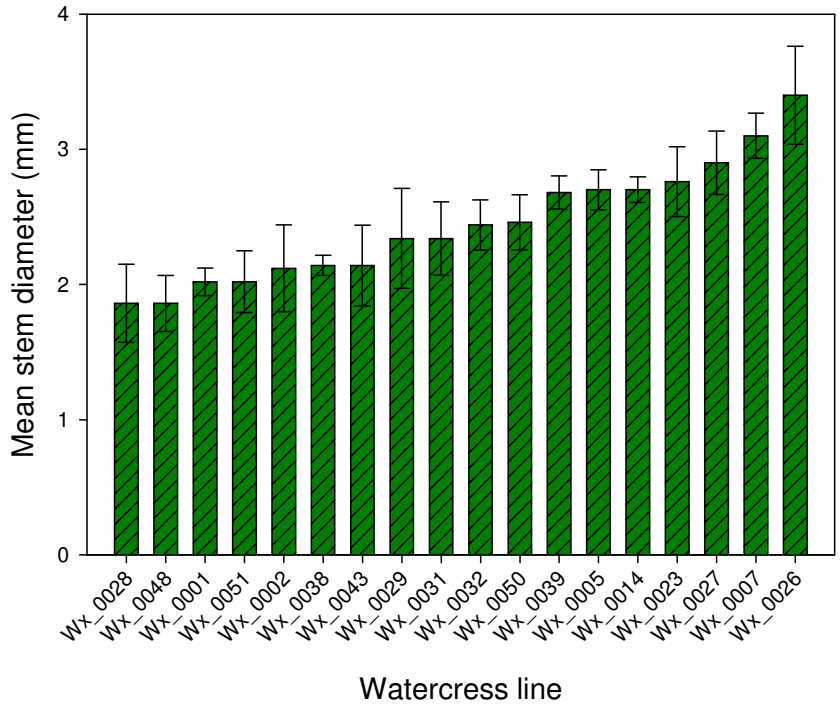


Figure 2.5: Variation in the mean stem diameter (±SE) of watercress lines (a) grown in a controlled environment and (b) in the field – Spetisbury, Dorset (c) Comparison of the variation in the mean stem diameter of watercress lines grown in a controlled environment and in the field – Spetisbury, Dorset. The stem diameter between the watercress lines significantly differed ($F_{11, 96} = 2.35$, $p<0.05$) but there was no significant difference on stem diameter as a result of condition ($F_{1, 96} = 0.02$). NB. Fewer lines successfully germinated in the field compared to the controlled environment.

c)

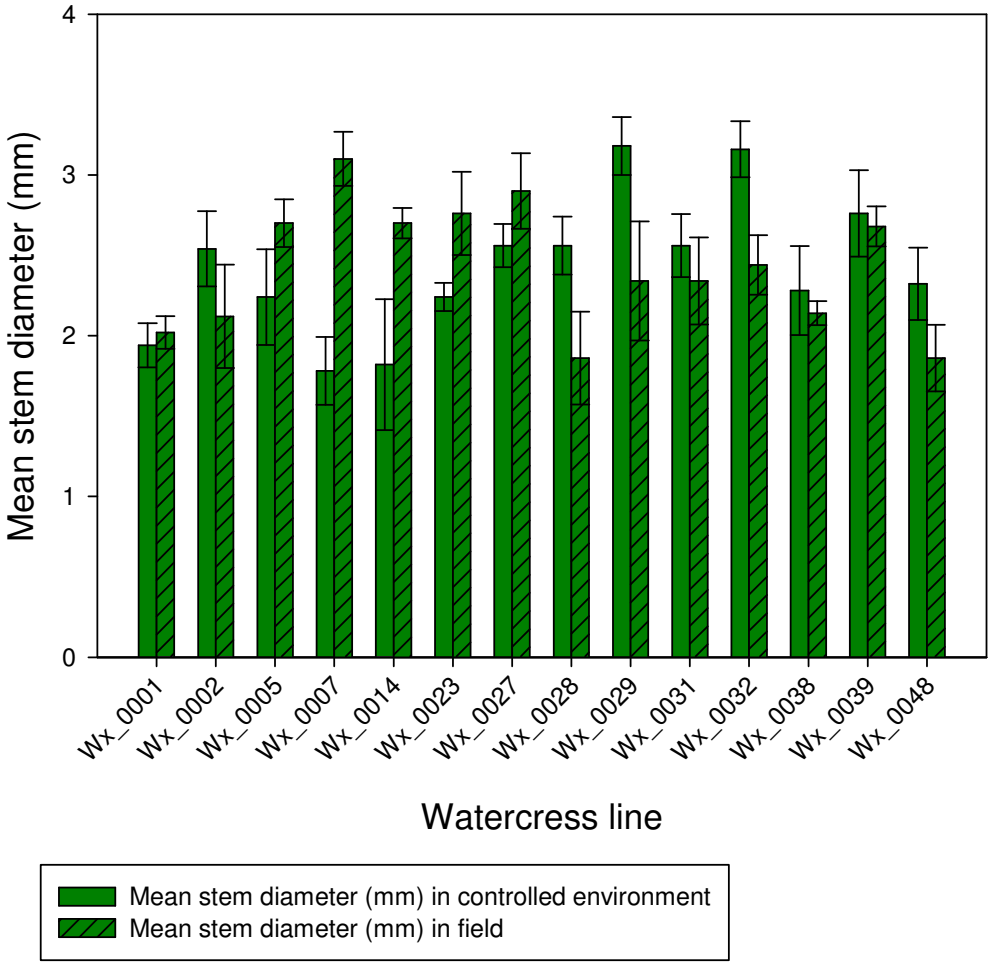


Figure 2.5 Continued

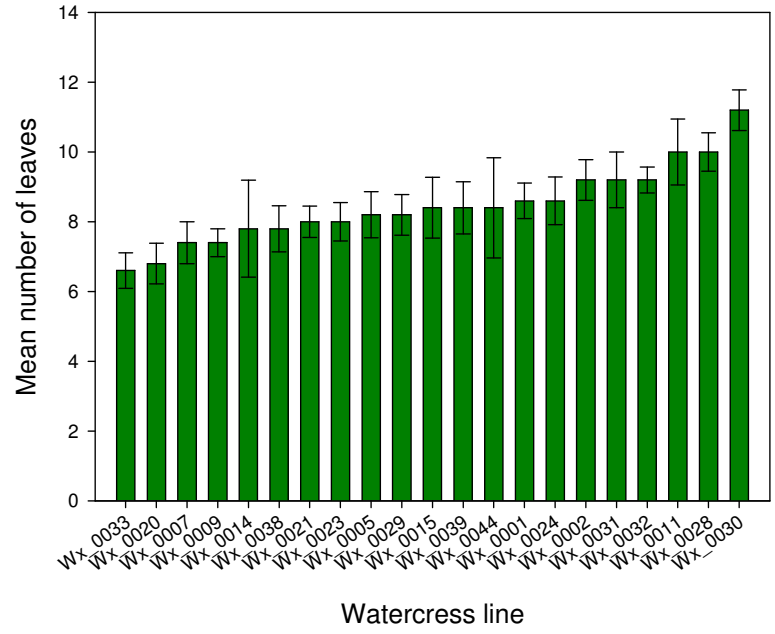
2.4.5 Number of leaves

The number of leaves significantly differed between the different lines. The highest mean number of leaves was present on Wx_0030 whilst the Bliss New Zealand line (Wx_0033) had the lowest mean number of leaves (Figure 2.6a). The Bliss New Zealand line may have had the lowest number of leaves due to its dwarf appearance therefore had many leaves smaller than the stated 2cm X 2cm. In the field Wx_0032 had the lowest mean number of leaves with a total of 6 whilst Wx_0014 had the highest (Figure 2.6b). However in the controlled environment Wx_0032 had one of the highest mean number of leaves.

Again the majority of the lines had a higher mean number of leaves when grown in the controlled environment compared to when grown in the field (Figure 2.6c). The highest frequency of leaves in the field was 7 whilst in the controlled environment the highest frequency was 8-9 leaves.

There was no significant difference in the number of leaves between the controlled environment and field ($F_{11, 96} = 1.67$, $p>0.05$), but there was a significant condition effect ($F_{1, 96} = 17.39$, $p<0.001$). There was a significant interaction between the watercress line and condition ($F_{11, 96} = 2.35$, $p<0.05$).

a)



b)

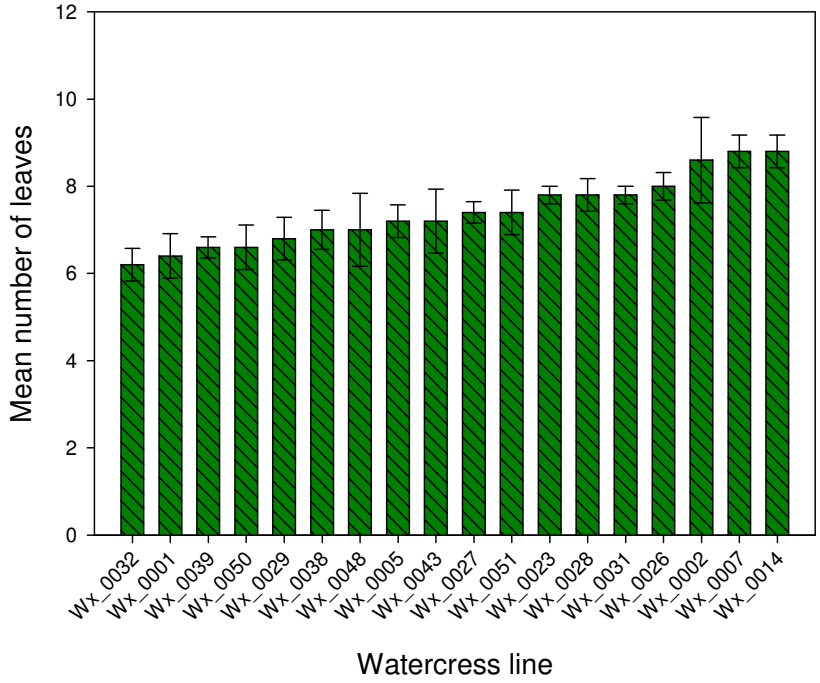


Figure 2.6: Variation in the mean number of leaves (±SE) of watercress lines; grown (a) in a controlled environment and (b) in the field – Spetisbury, Dorset (c) comparison of the number of leaves grown in both conditions. There was no significant difference in the number of leaves between the controlled environment and field ($F_{11, 96} = 1.67$, $p>0.05$), but there was a significant condition effect ($F_{1, 96} = 17.39$, $p<0.001$). NB. Fewer lines successfully germinated in the field compared to the controlled environment.

c)

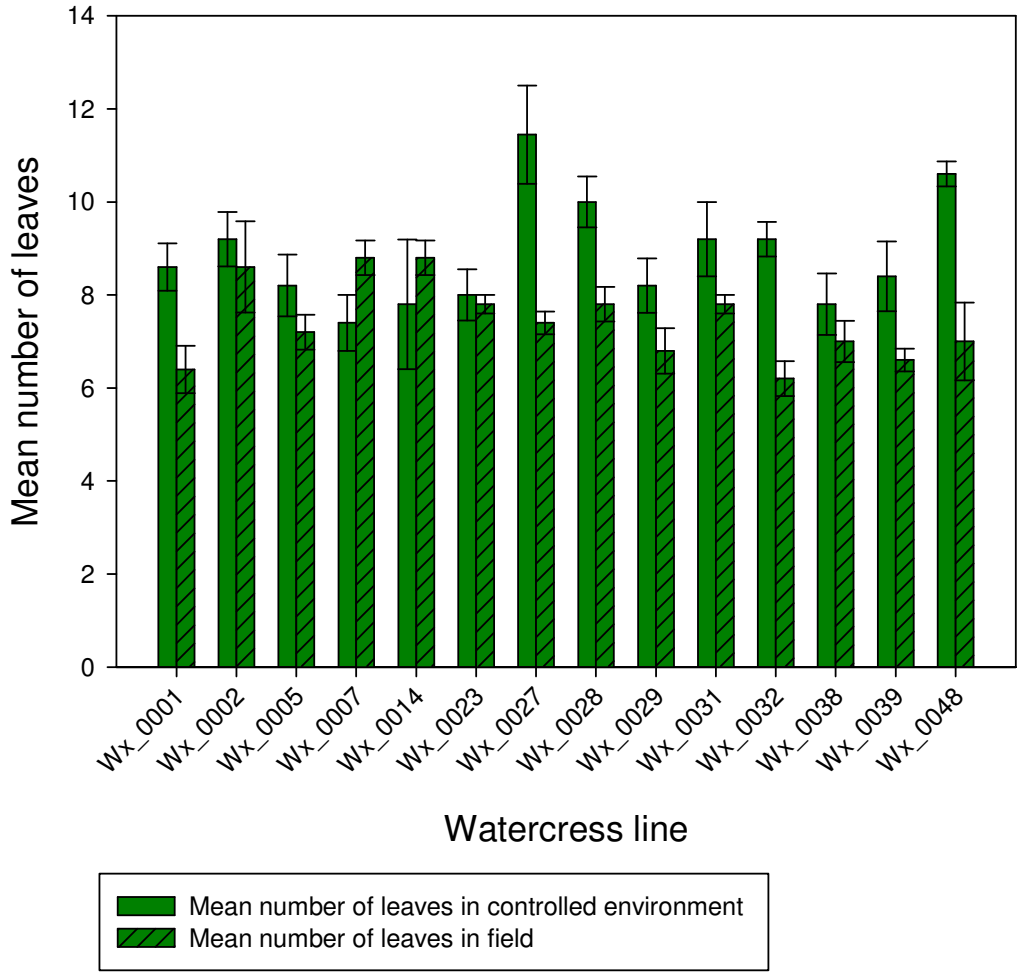


Figure 2.6 Continued

The leaf number index was calculated looking at the leaf number per unit stem (Figure 2.7). Overall it appears that the leaf number index is highest when grown in the field indicating that in the field there are a higher number of leaves per unit stem.

There was a significant difference in the leaf number index between the controlled environment and field ($F_{11, 96} = 7.15$, $p<0.001$), there was no significant condition effect ($F_{1, 96} = 0.41$, $p>0.05$). There was a significant interaction between the watercress line and condition ($F_{11, 96} = 7.58$, $p<0.001$).

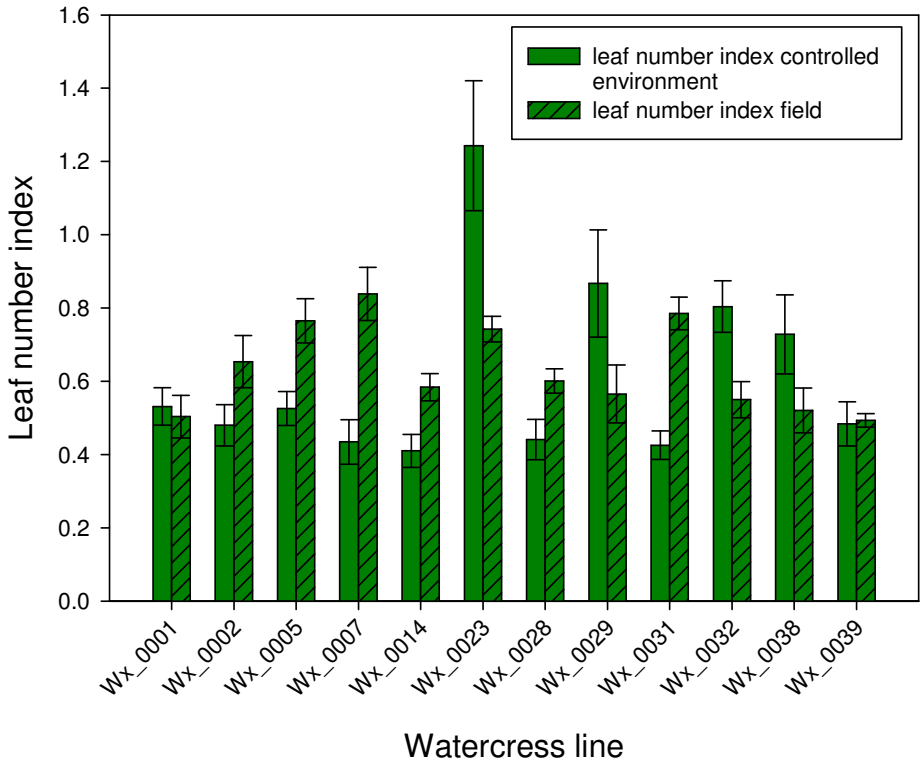


Figure 2.7: Comparison of leaf number index (±SE) in controlled environment and field –Spetisbury, Dorset. There was a significant difference in the leaf number index between the controlled environment and field ($F_{11, 96} = 7.15$, $p<0.001$), there was no significant condition effect ($F_{1, 96} = 0.41$, $p>0.05$). NB. Fewer lines successfully germinated in the field compared to the controlled environment.

2.4.6 Correlations between selected traits

The possible correlations between stem length, stem diameter and number of leaves were investigated for both the lines grown in the controlled environment and field. However it is evident that none of the traits correlate either in the controlled environment or field (Figue 2.8). This reinforces the need to look at each trait individually.

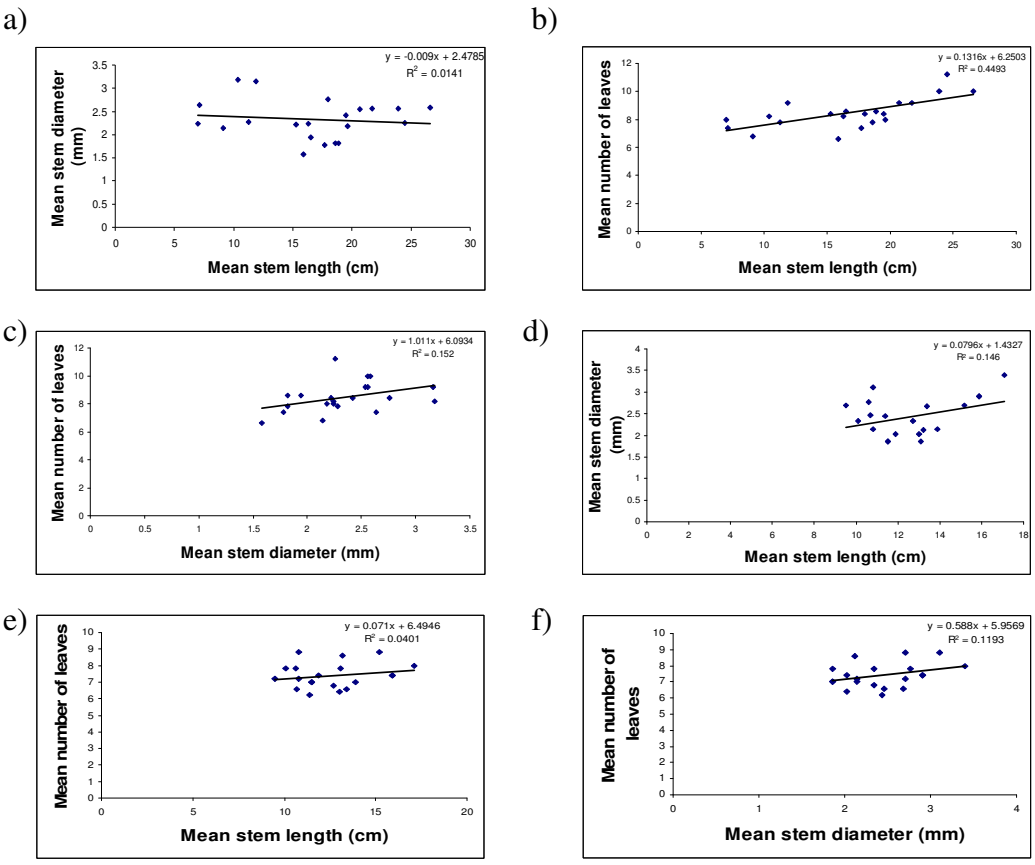


Figure 2.8 a-c) correlations between traits of interest in watercress lines grown in a controlled environment d-f) correlation between traits of interest in watercress lines grown in the field

2.5 Discussion

Results of growth media tests indicate that the environment can cause variation in the growth of watercress (Wx_0001). The watercress' were clearly visually different in the different growth media. In growth medias in which the plants appeared stressed the resultant watercress plants are stunted (Figure 2.4a) and darker in colour. However in the Vitacress peat the resultant watercress plant appeared a lighter green and a much fuller growth. In the vitacress peat the stem length is roughly equal to the average stem length obtained by propagating the watercress in the other growth media but the watercress in the vitacress peat on average tends to have a higher mean stem diameter and number of leaves. The darker green indicates the possible presence of anthocyanins which are a stressed induced response. From these results it was determined to establish the watercress lines in each trial in Vitacress peat for optimal growth and in addition to this it would be the Vitacress peat which the watercress lines would be grown in if taken forward for any commercial production.

One response to stress is the production of anthocyanins. Anthocyanins belong to the diverse class of phenolic compounds collectively named flavonoids, there are believed to be more than 400 anthocyanins found in nature (Kong *et al.*, 2003). The most significant function of anthocyanins is the colour that results from the presence of the anthocyanins in the plants or the plant products in which they occur (Kong *et al.*, 2003). They have long been regarded as an extravagant waste of a plant's genetic resources but may actually be in some instances critical for plant survival (Gould, 2004). The synthesis and vacuolar sequestration of anthocyanin molecules represent a considerable metabolic investment for plant cells (Gould, 2004). It is not likely that these red pigments are simply a default product of a saturated flavonoid biosynthetic pathway as timing of production of anthocyanins is very tightly controlled (Gould, 2004). Anthocyanins can act as antioxidants, phytoalexins or as antibacterial agents (Kong *et al.*, 2003). Anthocyanins have been implicated in tolerance to stressors as diverse as drought, UV-B and heavy metals as well as resistance to herbivores and pathogens (Gould, 2004).

Photooxidative injury may be mitigated by anthocyanins in the leaves by efficiently scavenging free radicals and reactive oxygen species (Gould, 2004). Therefore by growing watercress in a stressed environment the levels of antioxidants may be increased. This implies that the concentration of PEITC may be increased not only through breeding but by manipulating the agronomy too i.e. by causing stress. It has already been identified that watercress needs to be grown at low temperatures, under long days and exposed to red light in order to optimise the concentration of gluconasturtiin, the glucosinolate precursor of PEITC, also high concentration of sulphur (higher ratio of sulphur to nitrogen) can increase PEITC concentration (Engelen-Eigles *et al.*, 2006).

Growing the watercress lines under controlled environments is very important as the environment can affect the size and habit of the plants. Going *et al*. (2008) demonstrated that watercress displays considerable morphological plasticity and acclimates to low light conditions primarily by increasing leaf area and canopy surface area, which supports the growth of the watercress lines in a controlled environment when under experimental conditions. The root biomass and total biomass of seedling grown and vegetative fragment grown watercress decreased with decreasing light levels and hence increased with increasing light levels (Going *et al*., 2008). Morphological adaptations to different light levels are evident by a diffuse canopy and thinner and larger leaves in low light levels which is the result of spreading thylakoids (Enriquez and Sand-Jensen, 2003). Going *et al*, (2008) considered the decrease in canopy and leaf area of watercress in high light treatments as a mechanism to reduce the intensity of light reaching the pigments thereby reducing the risk of damage to photosynthetic apparatus. Equally the increase in canopy and leaf area in low light treatments may be a mechanism to facilitate an increase in the amount of light captured by the pigments and to increase the rate of photosynthesis (Going *et al*, 2008). Going *et al*, (2008) suggested that the morphological plasticity that watercress displays under varied light conditions may confer an adaptive advantage to this plant when growing in heterogeneous light environment common in streams. Interestingly watercress may be one of the few aquatic angiosperms capable of persisting in heavily shaded streams as well as streams with open canopies (Going *et al*, 2008).

Some of the morphological differences observed in the glasshouse experiment in which the watercress was grown in different media may also be a result of varying light intensity in the glasshouse. The watercress acclimates to an increase or decrease in light levels primarily through morphological changes in canopy and leaf area but also through physiological mechanisms including changes in chlorophyll concentration or shifts in the ratio of photosynthetic pigments (Going *et al*, 2008). However we can conclude that the differences observed in the growth room are the result of genetic differences between the different lines.

All the watercress lines in the germplasm collection were grown under standard conditions (light intensity, temperature, pot size, peat) as results from the glasshouse has revealed that the environment can have an affect on the growth morphology of plants. By removing any potential fluctuations, differences between the lines are due to genetic variation and not the environment.

There is significant variation between the watercress lines regarding morphological traits (stem length, stem diameter and number of leaves). This variation is clear in the phenotype but the genotype needs to be investigated in order to determine genetic diversity. One method which can be used to assess genetic diversity within the collection is Amplified Fragment Length Polymorphism (AFLP). Also microarray data will enable the expression of key genes to be examined.

Crop domestication began many years ago with early agriculturalists selecting from wild germplasm for material which was better adapted to human use and cultivation (Frary and Doganlar, 2003). Hence it is often the wild relatives of many commercial cultivars which harbour the highest genetic diversity and therefore selective breeding could combine these favourable traits into an ‘ideal’ cultivar. (Hajjar and Hodgkin, 2007) reported a steady increase in the rate of cultivars containing genes from Crop Wild Relatives (CWR). The CWR provide a broad pool of potential genetic resources which may be of benefit to agriculture (Hajjar and Hodgkin, 2007). Domesticated species differ from their wild ancestors in morphological and

physiological characteristics (Frary and Doganlar, 2003). Assessing the stem length, stem diameter, number of leaves and leaf:stem ratio gave the first set of data in which the variation within the watercress germplasm collection could be established and determined. It would reveal whether a line was present within the collection which may harbour traits i.e a reduced stem length to benefit the current standard industrial cultivar.

Tanksley and McCouch (1997) emphasised the importance of continued sampling of wild germplasm to discover new genes and uses. Within the watercress germplasm collection are lines sourced from different countries which may aid in increasing the gene pool of watercress. Genes from wild relatives have been used for the following improvements: pest and/or disease resistance or abiotic stress tolerance, drought and salt tolerance, increasing yield and improving quality traits of crops (Hajjar and Hodgkin, 2007). Although wild relatives have low yields and are of poor eating quality they are ancient sources of genetic material which form the basic building blocks from which all modern varieties are constructed (McCouch, 2004). An increase in the use of wild relatives may be aided by advances in genomics. DNA markers and sequencing has helped in isolating beneficial genes and selecting traits which are difficult to detect when based solely on phenotype (Hajjar and Hodgkin, 2007). Genetic diversity is obviously essential for evolution in nature but it is also important for the improvement of plant breeding (Harlan, 1975). The last century has seen very active breeding programmes which have in turn led to increases in both the quality and quantity of crops produced (Gepts, 2002)

The job of the plant breeder is to create an improved variety (McCouch, 2004). Hence the aim to produce an improved watercress line by establishing the collection with different lines. It is widely recognised that the wild relatives and early landrace varieties contain an essential pool of genetic variation which will aid in the future improvement of plants (McCouch, 2004). The natural biodiversity present in the form of wild species or old varieties represent genotypes which were selected over evolutionary time for adaptation to natural environments (Zamir, 2008). Thereby by sourcing watercress from a wide range of environments we hope to unlock some

genetic diversity which may aid in a breeding programme. A reduced stem length is considered very important. The variation in stem length is of particular interest for plant breeders given the number of benefits that a dwarf plant provides. Compared to their wild relatives crop plants tend to be shorter and have a more compact growth (Frary and Doganlar, 2003). In agriculture and horticulture the chemical manipulation of gibberellin (GA) status has been widely practiced, in particular GA-regulated developmental processes such as stem elongation (Hedden and Phillips, 2000). Gibberellins are tetracyclic diterpenoid growth factors which are essential for the regulation of stem elongation and other plant developmental processes (Hooley, 1994). It is particularly interesting thereby to focus on gibberellins and their resulting effects on plant height. Future breeding programmes may indeed rely upon the effects of genes (genotype) rather than looking at the final phenotype. It is important therefore to establish which genes contribute to a phenotype.

In rice the total number of elongated internodes and the length of each internode determine height, with a rice plant usually having 4-6 elongated internodes (Yang and Hwa, 2008). Mutant studies have shown that the height of rice plants is controlled in two different ways: genes controlling internode elongation are either expressed in all internodes which allows coordinated elongation of every internode or the genes act in only one or two internodes (Yang and Hwa, 2008). Genes encoding gibberellic and brassinosteroid biosynthetic or signalling pathways were found to act in all elongated internodes (Yang and Hwa, 2008). Wheat that responded abnormally to gibberellin played a major role in the green revolution during the 1960s and 1970s (Peng *et al.*, 1999). The new variety of wheat was shorter, had an increased grain yield and a higher resistance to damage caused by wind and rain (Peng *et al.*, 1999). The reduced response to gibberellin is the result of altered function mutant alleles of the *Rht-1* height regulating genes of wheat (Peng *et al.*, 1999). Mutant alleles that reduce plant height and reduce responses to gibberellin are characteristic of maize d8 and the *Arabidopsis* gai allele (Peng *et al.*, 1999). The possibilities of using *GAI* orthologues to increase yield in a wide range of crop species has proven to be a strong option given the success of transgenic rice containing a mutant *GAI* allele and the resultant dwarfism (Peng *et al.*, 1999).

Two agriculturally important crop plants (rice and wheat) develop a dwarf phenotype as a result of mutations in the gibberellic acid or brassinosteroid biosynthetic or signalling pathways. It is important to dissect these pathways and *Arabidopsis* offers a chance to extrapolate this information. Moreover the genus Brassica is in the same taxonomic family as *Arabidopsis thaliana* and such a close relationship could mean that the genus Brassica will be among the first beneficiaries of the *Arabidopsis* sequence (Lan and Paterson, 2001) hence watrecress may be among one of the first crops which benefit from the model plant *Arabidopsis*. Therefore important agriculture crops such as rice and wheat may uncover varible phenotypes which can then be further explored using *Arabidopsis* and adapted into other crops.

A reduction in gibberellins (GA) responsiveness is conferred by the *Arabidopsis* gai mutant allele (Peng *et al.*, 1997). GAI (wild type) and gai (mutant) differ by a deletion of a 17 amino acid segment from within the amino-terminal, gai is a mutant repressor which is resistant to the effects of GA (Peng *et al.*, 1997). Lucas *et al*. 2008 reported that the central role of the *Arabidopsis thaliana* nuclear transcription factor PIF4 (PHYTOCHROME INTERACTING FACTOR) in the positive control of genes mediating cell elongation (Lucas *et al.*, 2008).

Suppression of GA responses are caused by the gene *SHORT INTERNODES* (*SHI*) when constitutively activated in *Arabidopsis* (Fridborg *et al.*, 2001). SHI expressed in young organs (e.g. shoot apices and root tips) predicts a suppressor GA response to be active in order to prevent premature growth or development (Fridborg *et al.*, 2001). The *SPY* (*SPINDLY*) protein is thought to act as a negative regulator of GA signalling which is achieved through GlcNAc. Mutations in *SPY* cause an elongated growth typical for the response to exogenous application of GA (Fridborg *et al.*, 2001).

In a rapid cycling line of *Brassica rapa* a single gene mutant identified as elongated internode [*ein*/*ein*] caused accelerated shoot elongation. The application of traizole plant growth retardant, paclobutrazol, inhibited shoot elongation thereby restoring a

normal phenotype. Upon application of gibberellin A3 (GA3) the normal genotype is converted back to the *ein* phenotype (Rood *et al.*, 1990). Also discovered in *Brassica rapa* is the dwarf gene *dwf2* which was insensitive to application of GA3. However it was insensitive with regard to both plant height and flowering time indicating it may not be a mutation in the gibberellins biosynthesis pathway (Muangprom and Osborn, 2004). *Brassica napus* is often prone to lodging which then leads to yield loss and difficulty to harvest therefore dwarf genes may be of particular use in oil seed rape (Muangprom and Osborn, 2004).

The molecular basis of brassinosteroid (BR) biosynthesis and function in controlling plant height has been revealed from studies using *Arabidopsis*. Brassinosteroids are plant steroid hormones which regulate a variety of plant growth and developmental processes (Li, 2005). Inhibition of *BRI1, BAK1, BIN2, BES* and *BZR1* which are involved BR signalling resulted in dwarfism. These results suggested that BR signalling plays an essential role in cell elongation thereby affecting plant height (Wang and Li, 2006). An orthologue of BR1 was discovered in rice and it appears that dicotyledonous and monocotyledonous plants probably share conserved molecular events to modify height through the BR biosynthetic and signalling pathways (Wang and Li, 2006).

DELLA proteins function as key repressors of GA-responsive growth by inhibiting GA regulated expression (de Lucas *et al.*, 2008). When accumulated in the nucleus these repressors are rapidly degraded in response to GA. In *Arabidopsis* RGA (encoded by *repressor of gal-3*) and GAI (encoded by *GA insensitive*) are the main repressors controlling hypocotyl growth and stem elongation. With a mutation in the DELLA domain the proteins are resistant to degradation and the resultant dwarf phenotype is GA-insensitive (de Lucas *et al.*, 2008).

The watercress lines with a shorter stem length may be the result of lower GA biosynthesis and higher catabolism of GA. It is the GA20ox and GA3ox genes which need to be down regulated or inhibited (Hedden and Phillips, 2000b) and GA2ox

overexpressed in order to cause GA1 deactivation (Sakamoto *et al.*, 2003). The variation in plant height could also be a result of the ability to cause the destruction of the DELLA proteins as plant growth is stimulated via the destruction of the DELLA proteins and gibberellin signalling uses the ubiquitin-proteasome pathway (ubiquitin-26s proteasome pathway) to control expression through degradation. Further research may reveal the underlying mechanisms for the variation in height in the watercress lines. By creating a list of all the genes which result in a plant with a reduced stem length we can explore the expression levels in the various watercress lines.

As described above there are numerous mechanisms which could result in a crop plant being classified as 'dwarf'. Data collected in this chapter has confirmed that there is indeed a watercress line which has a stem length which is shorter in comparison with the standard line, Wx_0001. Therefore the research has uncovered a line which could aid in the improvement of the current standard cultivar, Wx_0001 with respect to the targeted trait of a reduced stem length. Hence an important step has been taken in the watercress breeding programme in which specific traits (stem length, stem diameter, number of leaves and leaf:stem ratio) have been recorded for future reference in breeding.

2.4.7. Conclusions

1. Watercress lines were established in the controlled environment and field, Spetisbury Dorset.
2. There was a significant difference in the stem length, stem diameter and leaf number index between the different watercress lines. However no significant difference in the number of leaves.
3. Varitation in morphological traits has been reported and recorded for the University of Southampton watercress germplasm collection.
4. Watercress lines were selected and taken forward for the watercress breeding programme (Chapter four).

CHAPTER THREE: BIOCHEMICAL VARIATION WITHIN THE GERMPLASM COLLECTION AND EVIDENCE OF A CORRELATION BETWEEN DIFFERING ANTIOXIDANT ASSAYS

3.1 Abstract

Watercress (*Rorippa nasturtium-aquaticum*) is a popular vegetable accompaniment to many dishes. New research results suggest that watercress contains one of the highest concentrations of a precursor for the beneficial antioxidant, phenethyl isothiocyanate (Gill *et al.*, 2007). Watercress is also perceived as a healthy food which may have a preventative role in several cancers (Gill *et al.*, 2007). This recent finding has the exciting possibility of breeding a higher quality crop with enhanced anti-cancer activity. Here we report the biochemical variation in the watercress germplasm collection and the possibility of crossing different watercress lines in order to increase the concentration of antioxidants. A significant difference was reported in Ferric Reducing Antioxidant Power (FRAP) values between the different watercress lines and this FRAP value dramatically increases when the lines are grown in the field. The values increase from an average of 80-140 mmol Fe^{2+} equivalent per gram fresh weight in the controlled environment to 200-700 mmol Fe^{2+} equivalent per gram fresh weight when grown in the field. The variation in the level of antioxidants according to different pre and post-harvest processes is also reported.

Watercress (*Rorippa nasturtium-aquaticum*) along with rocket (*Eruca sativa*) are members of the *Brassicaceae* family whilst spinach (*Spinacia oleracea*) is a member of the *Amaranthaceae*. These commercial crops are reported to have high concentrations of antioxidants therefore possibly contributing to disease prevention following consumption. Salad leaves have become very popular after reported benefits of consuming. The antioxidant content potential of these species' was assessed using two comparable techniques. The techniques were compared in order to determine consistency between the assays. Antioxidants were assessed by the Ferric Reducing Antioxidant Power (FRAP) assay and the Oxygen Radical Absorbance Capacity (ORAC) assay. Differences between conventional and organic practices were also investigated.

Members of the *Brassicaceae* family are well known for their distinct flavour e.g. Brussels sprouts, horseradish, broccoli and watercress. Indeed both watercress and rocket used in this study are well known for their peppery taste. These flavours are the result of a group of compounds known as glucosinolates. It is the glucosinolates which contribute to the numerous health benefits obtained from consuming these plant products. Many health benefits have been reported but much focus has been placed on the anti-cancer properties of the *Brassicaceae*, therefore investigation of the variation within germplasm collections with regards to the level of glucosinolates is particularly important. By manipulating the level of glucosinolates it may be possible to enhance the anti-cancer properties, of great importance is phenethyl isothiocyanate. Phenethyl isothiocyante is produced by the action of myrosinase on glucosinolate and it is the phenthyl isothiocyanate which can intervene at many stages of cancer thereby reducing progression of the cancer. Hence in this study high performance liquid chromatography was employed to deduce the concentrations of glucosinolates whilst gas chromatography mass spectrometry used to deduce the concentration of isothiocycanates.

Watercress had the highest ferric reducing antioxidant power (562 mmol Fe^{2+} equivalent per gram fresh weight), followed by spinach (275 mmol Fe^{2+} equivalent per gram fresh weight) and rocket (169 mmol Fe^{2+} equivalent per gram fresh weight) respectively. Similarly watercress also had the highest overall ability to scavenge free radicals (59.59%), followed by spinach (27.95%) and rocket (9.81%) respectively. There was little difference between the FRAP and ORAC values for conventional and organic salad. There is variation in salad crops antioxidant potential with the highest values recorded consistently for watercress and that FRAP and ORAC are comparable for measuring antioxidants in these salad crops with similar ranking for each salad crop studied. Differences in the concentration of phenethyl glucosinolate and phenyethyl isothiocyanate were observed between different watercress lines when analysed by high performance liquid chromatography and gas chromatography mass spectrometry. The lowest concentration of phenethyl glucosinolate was observed in a green watercress cultivar (2.74 µmoles per gram dry weight) and the highest in the standard line, Wx_0001 (19.35 µmoles per gram dry weight). The lowest concentration of phenethyl

isothiocyanate was observed in Wx_0038 (0.09 mg/ml) and the highest in Wx_0001 (0.20 mg/ml). The results obtained offer exciting opportunities to use the natural variation of this beneficial compound in watercress to produce a nutritionally enhanced watercress cultivar.

3.2 Introduction

Recently much attention is focused on the nutritional attributes and benefits of food with particular focus on *Brassicas*. Consuming a constant supply of phytochemical containing plants aids defensive mechanisms in humans and reduces the risk of chronic diseases provided that the consumed plant material has a high nutritional content (Chu *et al.*, 2002; Liu R.H, 2002). Prevention is a more effective strategy than treatment of chronic diseases (Chu *et al.*, 2002; Liu R.H, 2002). Green leafy vegetables have the highest antioxidant capacity followed by fruits and root crops (Lako *et al.*, 2007) and the frequent consumption of fruits and vegetables is associated with a lowered risk of cancer, heart disease, hypertension and stroke (Lako *et al.*, 2007).

The many benefits resulting from consumption of fruits and vegetables is the presence of various forms of phytochemicals and antioxidants present in the foods such as carotenoids and polyphenol compounds including flavonoids and anthocyanins (Lako *et al.*, 2007). It is the 'new' health benefits of these foods which may encourage assessment and determination of potential rich sources of antioxidant compounds which could improve cultivar development, post-harvest storage and production practices (Lako *et al.*, 2007). This suggests the possibility of improved public health through the diet (Jeffery *et al.*, 2003).

Brassicas contain isothiocyanates (ITC) which are synthesised and stored as glucosinolates but are released when plant tissues are damaged. This conversion is catalysed by the enzyme myrosinase (thioglucoside glucohydrolase) which is

normally present but separated from the glucosinolates until damage occurs (Zhang, 2004). However, as well as isothiocyanates, non-ITC products including thiocyanates, nitriles, epithionitriles, indoles and oxazolidine are also formed (Zhang, 2004).

Only a small number of ITCs are commonly consumed by humans. Cruciferous vegetables are the principle dietary source of ITCs but crucifers consumed by humans are limited. Crucifers which are particularly rich in certain ITCs include mustard and horseradish - allyl-ITC (AITC) (Nielsen and Rios, 2000), watercress – phenethyl-ITC (PEITC) (Rose *et al.*, 2000), rocket – erucin (Barillari *et al.*, 2005) and broccoli – sulforaphane (SF) (Kushad *et al.*, 1999).

The most important known biological activity of ITCs is their ability to inhibit cancer development (Zhang, 2004). Natural isothiocyanates such as AITC, benzyl ITC (BITC), PEITC and SF and a large number of synthetic analogues are effective inhibitors of chemically induced tumors in one or more organ sites of rodents including the bladder, colon, esophagus, mammary glands, pancreas and stomach (Zhang, 2004).

Watercress is the richest source of the glucosinolate gluconasturtiin which is hydrolyzed by myrosinase to phenethyl isothiocyanate (PEITC) (Palaniswamy *et al.*, 2003). Cancers caused by tobacco-specific carcinogens (e.g. 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, N-nitrosomethyl benzylaamine, benzo[α]pyrene, and N-nitrosobenzyl-methlamine) (Siglin *et al.*, 1995; Wattenberg, 1992) were inhibited in rats and mice (Siglin *et al.*, 1995; Wattenberg, 1992) by PEITC.

PEITC can also act as a blocking agent and inhibit tumour initiation by inhibition of cytochrome P450 enzymes and by induction of phase II enzymes e.g. glutathione S-transeferase (Meyer D J *et al.*, 1995; Palaniswamy *et al.*, 2003) . PEITC has been demonstrated to be an effective inhibitor of hypoxia inducible factor (HIF) which is

a transcription factor that plays an important role in expression of pro-angiogenic factors (Wang *et al.*, 2009). Angiogenesis is the formation of blood vessels which is an essential part of tumor formation. Much focus has been placed on the anti-cancer action of PEITC; a new line of watercress with an enhanced concentration of PEITC would provide a great opportunity to increase the nutritional benefits of watercress. Many factors can cause variation in the levels of antioxidants of *Brassica* (Podsedek, 2007) and these include variety, maturity at harvest, growing condition, soil state and condition of post-harvest storage (Podsedek, 2007).

The major health benefits from plants, as mentioned above, are thought to be due to the presence of antioxidants and their proven ability to scavenge radicals and inhibit chain reaction or break the chain propagation (Podsedek, 2007). A free radical is a free atom, molecule or ion which contains one or more unpaired electrons (Aruoma *et al.*, 1993). Free radicals have been implicated to play an important role in a number of biological processes and can react and damage DNA causing many forms of cancer (Traka and Mithen, 2009). The term 'antioxidant' covers a wide range of different molecules and a common feature is their ability to readily donate electrons/protons whilst remaining stable themselves, thereby acting as reducing agents (Aruoma, 1998) and minimizing damage caused by free radicals. Therefore increased uptake of dietary antioxidants may help in the maintenance of antioxidant status and normal physiological function (Ou *et al.*, 2001). Antioxidants are chemically diverse and important phytonutrients but there is currently no total antioxidant as a nutritional index available for food labelling due to the lack of quantitation methods (Ou *et al.*, 2002). Vitamins C and E, carotenoids, flavonoids and thiol (SH) compounds are the antioxidants commonly found in vegetables, it is the chemical diversity of these antioxidants which makes it difficult to separate and quantify the individual antioxidants from the vegetable matrix (Ou *et al.*, 2002).

Antioxidants which are present in plants and act as a defence system within the plant can deactivate radicals by two major mechanisms: Hydrogen Atom Transfer (HAT) and Single Electron Transfer (SET) with the end result the same regardless of mechanism but the kinetics and potential for side reactions differ (Prior *et al.*, 2005).

Many antioxidant protocols have been devised to determine the antioxidant status of foods. Of particular interest are the Ferric Reducing Ability of Plasma/Ferric Reducing Antioxidant Power (FRAP) and Oxygen Radical Absorbance Capacity (ORAC), in which FRAP utilizes the SET reaction whilst ORAC utilizes the HAT reaction (Prior *et al.*, 2005). Other antioxidant protocols include Trolox Equivalent Antioxidant Capacity (TEAC) and Total Radical Absorbance Potentials (TRAP).

In the SET based methods (FRAP and TEAC) the antioxidants are oxidized by oxidants such as Fe (III) or $ABTS^{+\cdot}$. A single electron is transferred from the antioxidant molecule to the oxidant which deactivates free radicals (equation 3.1)

$$M(n) + A\text{-}H \rightarrow M\ (n\text{-}1) + A\text{-}H+ \quad (M = \text{metal ion}) \qquad \textbf{(eqn 3.1)}$$

The change in absorbance of either antioxidant or oxidant is measured by an ultraviolet-visible spectrometer and the absorbance value is used as the quantitation for the reducing capability of the antioxidant (Ou *et al.*, 2002). In the HAT based methods (ORAC and TRAP) (equation 3.2 and 3.3) a radical initiator is used to generate a peroxyl radical $ROO^{\cdot}$. The $ROO^{\cdot}$ takes a hydrogen atom from the antioxidant and therefore a reaction between the $ROO^{\cdot}$ and target molecule is retarded or initiated (equation 3.2). The result from the reaction is the formation of a stable antioxidant radical from when the oxygen radical abstracts a hydrogen from the antioxidant (Ou *et al.*, 2002).

$$ROO^{\cdot} + A - H \rightarrow ROOH + A^{\cdot} \qquad \textbf{(eqn 3.2)}$$

$$ROO^{\cdot} + FL - H \rightarrow ROOH + FL^{\cdot} \quad (FL = \text{fluorescence}) \qquad \textbf{(eqn 3.3)}$$

(Ou *et al.*, 2002).

FRAP and ORAC are based on two different chemical reactions but it is important to confirm whether these two assays can give comparable values. The HAT and SET reactions (equation 3.1, 3.2 and 3.3) have the net same result but achieved via different chemical reactions. Out *et al.* (2002) reported that the ORAC assay reflected the peroxyl radical scavenging activity whilst FRAP assay only estimates Fe (III) reducing activity. Schlesier (2002) investigated a wide variety of antioxidant assays and recommended that at least two assay are used due to the differences between the test systems.

Spinach has been shown to have an exceptionally rich high phenolic content and high total antioxidant activity (Hecht *et al.*, 1995; Ismail *et al.*, 2004). The antioxidant and anticarcinogenic properties of spinach have been demonstrated (Gil, I *et al.*, 1999), (Nyska *et al.*, 2001), (Ambrosone *et al.*, 2004), (Ismail *et al.*, 2004). Rocket has a long history for medicinal use and is a source of carotenoids (lutein, beta-carotene), vitamin C, folate, flavonoids, glucosinolates (glucoerucin) and fibre (Barillari *et al.*, 2005), (Hedges and Lister, 2005)). Rocket seeds contain the glucosinolate glucoerucin which is hydrolysed to the isothiocyanate erucin (Barillari *et al.*, 2005). Erucin directly affects antioxidant activity by decomposing hydroperoxides and hydrogen peroxide which are known catalysts which trigger the cancer cascade (Hedges and Lister, 2005). Erucin also indirectly affects antioxidant activity by inducing phase II enzymes like sulforphane (Barillari *et al.*, 2005).

Therefore assays (FRAP and ORAC) have been developed in this chapter to determine the antioxidant potential of not just watercress but also spinach and rocket. The FRAP assay was specifically chosen to determine a ranking order for the watercress collection in order to determine lines in which to focus on due to their high total antioxidant power. The FRAP assay acted as a starting point on a previously un-studied collection and once these lines with a high total antioxidant power had been identified research could concentrate on specific antioxidant compounds i.e. for this research the glucosinolates and isothiocynates. The methods applied to determine specific antioxidant concentrations were gas chromatography mass spectrometry (GC-MS) and High Performance Liquid Chromatography

(HPLC). Glucosinolates are of particular interest due to not only their contribution to anti-oxidant benefits but also their contribution to taste. Other potentially important plant phytochemicals include carotenoids, phenolics compounds, folates, vitamins (C, E and K) and finally selenium and iodine.

The glucosinolates, as a consequence of their presence in certain *Brassicaceae* vegetables (cabbage, cauliflower, broccoli) and condiments (mustard, horseradish, wasabi), have been of interest to human society for many years (Halkier and Gershenzon, 2006). The glucosinolates are mainly found in the order Capparales and this includes the agriculturally important plant crops of the *Brassicaceae* family and *Arabidopsis,* the model plant (Wittstock and Halkier, 2002). The *Brassicaceae* contains more than 350 genera and 3000 species (Fahey *et al.*, 2001), the foods and beverages derived from plants are chemically complex and the protective effects could arise from many components or mixtures of components present hence difficult to isolate the beneficial compound (Halliwell, 2007).

Glucosinolates are nitrogen- and sulphur-containing, anionic natural products which upon hydrolysis by the enzyme myrosinase (endogenous thioglucosidase) produce several different products; isothiocyanates, thiocyanates and nitriles (Halkier and Gershenzon, 2006; Wittstock and Halkier, 2002). The distinct taste and flavour of broccoli, cauliflower and cabbage are due to the isothiocyanate hydrolysis product of glucosinolates (Halkier and Gershenzon, 2006).

The extensive use of *Arabidopsis* as a model plant has increased the understanding of glucosinolate biosynthesis (Wittstock and Halkier, 2002). Glucosinolate research has progressed rapidly resulting in the near completion of the elucidation of the core biosynthetic pathway, identification of the first regulators of the pathway and metabolic engineering of specific glucosinolate profiles (Halkier and Gershenzon, 2006). In an attempt to reduce damage from fungal and insect pests breeders have sought to manipulate the glucosinolate concentrations in rapeseed foliage (Mithen and Campos, 1996) as discussed in Halkier and Gershenzon (2006). Brassica

cultivars are finding increased use for 'biofumigation' (Halkier and Gershenzon, 2006). The past decade has seen the glucosinolates identified as potent cancer prevention agents (Halkier and Gershenzon, 2006).

The glucosinolates can be classified by their precursor amino acid and the types of modification to the R group. Aliphatic glucosinolates are derived from alanine (Ala), leucine (Leu), isoleucine (Ile), methionine (Met) or valine (Val) whilst aromatic glucosinolates are derived from phenylanine (Phe) or tyrosine (Tyr) and finally indole glucosinolates are derived from tryptophan (Trp) (Halkier and Gershenzon, 2006). Phenethyl isothiocyanate (PEITC) is an aromatic isothiocyanate (Satyan *et al.*, 2006). Significant differences have been noted among organs in both glucosinolate concentration and composition in *Arabidopsis* (Brown *et al.*, 2003). It is the seeds which have the highest concentration of glucosinolates, over 60µmol per gram dry weight when mature and up to 80µmol per gram dry weight during germination (Brown *et al.*, 2003). *Rorippa* produces a diverse range of glucosinolates (Daxenbichler *et al.*, 1991). For example Daxenbichler *et al.* (1991) reported that present among *Rorippa dubia*, *Rorippa globosa*, *Rorripa hilariana* and *Rorippa indica* were the glucosinolates Siabarin (7-Methylsulphinylheptyl), Hirsutin (8-Methylsulphinyloctyl), Arabin (9-Methylsulpinylnonyl), Capparin (Allyl (2-propenyl)) and Nasturtiin(2-Phenylethyl).

De Quiros *et al*. (2000) carried out QTL mapping of glucosinolates in a RI population de-rived from an F1 hybrid between *Arabidopsis thaliana* ecotypes Columbia and Landsberg *erecta*. A single major QTL coincident with the GSL-ELONG locus was identified which regulates side chain elongation. This is important as the major class of glucosinolates have side chains which are derived from methionine, these then vary in the length of side chain and modifications to structure (de Quiros *et al.*, 2000). When the tissue is disrupted the methionine-derived glucosionlates are of considerable importance as the products produced include isothiocyanates, nitriles, epithiocyanates and thiocyanates (Bones and Rossiter, 1996; de Quiros *et al.*, 2000).

The main glucosinolate present in watercress is phenethyl isothiocyanate (PEITC) and therefore provides a convenient model system (Svanem *et al.*, 1997). A minimum of around 12mg of PEITC is released when 56.8g watercress is consumed (Chung *et al.*, 1992; Hecht *et al.*, 1995; Hecht, 1999). Two separate protocols have been used to assess the total glucosinolate and isothiocycanate content of watercress. High Performance Liquid Chromatography (HPLC) is a sensitive technique which gives the possibility to determine a great number of glucosinolates and does not cause enzymatic degradations or any drastic deviations (Buchner, 1987). Gas Chromatography Mass Spectrometry (GC/MS) is a sensitive procedure for measuring individual isothiocyanates. Numerous isothiocyanates are common in plant extracts hence the chromatographs can become complicated and GC-MS has undoubtedly played an important role with regards to this (Fenwick *et al.*, 1982).

HPLC works using a column packed with a solid granular stationary phase surrounded by a mobile liquid phase. The analyte emerges in a distinct band from the column. The components of HPLC consist of a stationary phase, a column, a mobile phase and delivery system, an injector system, a detector and chart recorder and finally a fraction collector (Wilson and Walker, 2005). On the otherhand GC/MS works by exploiting differences in the partition coeffecients between the stationary phase and the mobile gas phase of the volatised analytes (Wilson and Walker, 2005). This procedure is confined to analytes that are volatile but thermally stable and these analytes are carried through the column by a mobile gas phase. The GC/MS technique is believed to be very powerful when coupled to mass spectrometry (Wilson and Walker, 2005).

The aim of this Chapter was to investigate whether there was any variation in the concentration of antioxidants between the different watercress lines in the germplasm collection. Any resultant variation will lead to selection of individual plants for crossing with the aim of producing an individual line with a higher level of antioxidants. Also within this Chapter any correlation between the FRAP and ORAC assays on organic and conventional samples of rocket, spinach and watercress was investigated. Finally this chapter assessed the concentration of glucosinolates (in

particular phenethyl glucosinolate) and isothiocyanates (in particular phenethyl isothiocyanate) in the watercress lines ranked highest by FRAP using High Performance Liquid Chromatography and Gas Chromatography Mass Spectrometry.

3.2.1 Aims

1. Determine the difference in antioxidant potential (using FRAP) between the watercress lines held in the watercress collection when grown in both the controlled environment and field.

2. Assess agronomic impact on antioxidant power in terms of stage of harvest, repeated harvest, shelf life and section of watercress plant.

3. Determine correlation between two antioxidant assays, FRAP and ORAC.

5. Determine the concentration of phenethyl glucosinolate and phenethyl isothiocyanate within selected watercress lines (Wx_0001, Wx_0011, Wx_0033 and Wx_0038).

3.3 Materials and Methods

3.3.1 Plant material

As stated in 2.3.3

3.3.2 Cultivation of seed lines for biochemical analysis

Lines cultivated as stated in 2.3.4. After 7 weeks of growth the plants were snap frozen for biochemical analysis. Five biological replicates were used for each line.

3.3.3 Field Analysis – Spetisbury, Dorset

As stated in 2.3.4.1. After 7 weeks of growth the plants were snap frozen for biochemical analysis. Five biological replicates were used for each line.

3.3.4 Collection of rocket, spinach and watercress

Rocket, spinach and watercress were randomly selected from commercial supermarkets and immediately snap frozen upon return to the University of Southampton. Eight biological replicates were used for each sample (collected by Alice Mazzer).

3.3.5 Examination of antioxidants in different parts of the watercress

Watercress was collected from the field i.e. ready to harvest, young transplanted and red watercress. The young transplanted, red watercress and half of the ready to

harvest watercress were snap frozen. The remainder of the ready to harvest watercress had the leaves separated from the stems and then the leaves and stems snap frozen separately. Three biological replicates were used for each sample.

3.3.6 Shelf life

Watercress was stored in the cold room at 4°C to replicate the fridge environment once sold to the consumer and samples were taken after 0, 1, 2, 3, 4, 5 and 6 days and snap frozen for FRAP analysis. Three biological replicates were used for each sample.

3.3.7 Repeated harvest

Four lines were randomly chosen and grown in pots in Vitacress peat and then harvested after 7 weeks and snap frozen for FRAP analysis. The plants were then left to re-grow and harvested again after 4 weeks for the 2nd harvest. The lines were then left to re-grow again for another 4 weeks and harvested and finally left for the final 4 weeks and harvested, simulating 4 consecutive harvests. Twelve biological replicates were used for each line.

3.3.8 Harvesting and storing prior to analysis

Plant material from the controlled environment was collected in the morning after 7 weeks of growth and snap frozen in liquid nitrogen and stored at -80°C before carrying out FRAP and the bioassay.

3.3.9 Sap extraction

Plant material was ground in liquid nitrogen using a mortar and pestle. Since it is important to ensure that the material does not defrost the material was kept on dry ice. QIAshredders were labelled and weighed before being filled with ground material from each of the lines and the weight of the fresh material was recorded. The material was then spun at maximum speed (13,000rpm) on a bench top centrifuge before the extracted sap was pipetted into a fresh eppendorf and the final weight recorded.

3.3.10 Assessing antioxidant potential

A protocol termed FRAP known as the Ferric Reducing Ability of Plasma or Ferric Reducing Antioxidant Power was used (Benzie and Strain, 1996). Snap frozen watercress was ground using a pestle and mortar in liquid nitrogen. In order to extract the entire sap contents, QIAshredders were filled with ground material and weighed and then spun at 13,000rpm for 5 minutes. The sap was then pipetted into a new tube and weighed. 10µl of the sap was carried forward for FRAP analysis. The following FRAP reagents were required: acetate buffer (2.04g sodium acetate + 40ml H_2O + 6.4ml acetic acid), TPTZ/HCl (8mg TPTZ + 2.56ml 40mmol HCl) and ferric chloride hexahydrate (0.135g/25ml H_2O). For the assay 25ml acetate buffer + 2.5ml TPTZ/HCl + 2.5ml $FeCl_3.6H_2O$ were combined forming the FRAP reagent. Into each well 5µl sample of the sap and 5µl sterile water was added and 300µl FRAP reagent. Iron standards ranging from 0.25mmol-8mmol were made using Iron Sulphate Heptahydrate. 10µl of each iron standard was pipetted into empty wells and 10µl sterile H_2O was pipetted into a well.

The optical density (OD) was measured after no more than 5 minutes at absorbance 620nm using a plate reading spectrophotometer (Anthos Labtec Instruments). 5 replicates were used for each line.

3.3.11 Bioassay – application of watercress extracts on breast cancer cell lines MCF7

(protocol provided by Breeze Cavell)

MCF-7 cells were planted at a density of 1000 cells per well in 50µl Dulbecco's Modified Eagle's Medium (DMEM) (supplemented with 10% FCS, 50U/ml Penicillin, 50U/ml Steptomycin and 2mM glutamine). The outside cells were filled with 100 µl sterile Phosphate buffer saline (PBS) which prevents the cells from drying out. The cells were then left to adhere in an incubator overnight at 37°C with a humidified atmosphere containing 10% CO_2 . The following day the cells were treated with 3 different watercress lines (Wx_0001, Wx_0011 and Wx_0038) and left for 6 days in incubators. Each 96 well plate also contained untreated cells in triplicate, where no compound was added just 50 µl DMEM and 0.05 µM staurosporine as a positive control, 0.01 µM solution was made up in DMEM and 50 µl added to the wells. Staurosporine is a protein kinase C inhibitor.

After 6 days, media from each of the wells was removed and replaced with 100 µl of a solution made up by mixing 100 µl RPMI (Roswell Park Memorial Institute 1640 Medium) per well with 5 µl CellTiter 96 AQ_{ueous} One Solution Reagent (Promega, UK) per well. The 96 well plates are returned to 5% CO_2 incubator (Galaxy S, RSBiotech) for 1 hour 30 minutes to allow for the colorimetric reaction to occur. Changes in absorbance were measured by a 96 well plate reader at 490nm (Varioskan Flash, Thermo Electron Corporation). An average value obtained from the blanks was subtracted from all other values to allow for any background absorbance of the media.

Metabolically active cells are able to convert the tetrazolium compound 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(-4-sulfophenyl-2H-tetrazolium), inner salt (MTS) that is contained in the CellTiter 96 AQ_{ueous} One Solution Reagent (Promega, UK) to the coloured product formazan. The amount of formazan, as measured by the absorbance at 490nm is said to be directly proportional to the number of living cells.

3.3.12 Chemical reagents

Chemicals used were of analytical grade: Iron (II) sulphate heptahydrate and Iron (III) chloride hexahydrate were obtained from Acros Organics (part of Fisher Scientific), TPTZ (2,4,6-tripyridyl-s-triazine) was purchased from Fluka, Trolox (6-hydroxy-2,5,7,8-tetramethyl-chroman-2-carboxylic acid) from Aldrich Chemistry, DPPH (2,2-diphenyl-1-picrylhydrazyl) from Sigma Life Science (part of Sigma Aldrich), sodium acetate and trihydrate and methanol were purchased from Sigma Alsrich and ethanol from Fisher Scientific.

3.3.13 Samples

Random bags of conventional and organically grown rocket, spinach and watercress were selected from supermarkets and the contents ground under liquid nitrogen and stored at -80°C. Samples were taken straight from the supermarket to the University of Southampton laboratory and snap frozen immediately.

3.3.14 Ferric Reducing Antioxidant Power (FRAP)

As described in 3.3.10

3.3.15 Oxygen Radical Absorbance Capacity (ORAC)

The procedure was devised by Zullo and Ciafardini, (2008) and was followed with reference to that of Kim *et al.* (2005) and Miliauskas *et al.* (2004), moderate adaptations were made. The DPPH peroxyl radical is dark purple in colour in solution with an absorption maximum at 517nm. Quenching of the radical by addition of antioxidants causes a decrease at 517nm due to decolourization as DPPH is reduced to DPPH-H (Friaa and Brault, 2006). The assay therefore measures the change in absorbance when DPPH is mixed with the sample. A 0.3mM solution of

the free radical was prepared and then kept in the dark. The assay technique was applied to samples in sets of eight. Sap samples were again diluted in eppendorfs and then DPPH was added and mixed thoroughly for 30 seconds and then incubated for 10 minutes in the dark at room temperature. After incubation reaction mixtures were rapidly transferred to cuvettes and absorbances measured at 517nm, the spectrophotometer was first blanked with a solution of water and 80% methanol. The percentage of radicals scavenged was calculated by using equation 3.4 (Miliauskas *et al.*, 2004).

DPPH radicals scavenged (%) = (Acontrol–Asample)/Acontrol)x100 **(eqn 3.4)**

Acontrol = absorbance of control

Asample = absorbance of sample

3.3.16 Determination of total glucosinolates by High Performance Liquid Chromatography (HPLC) (in collaboration with Dr John Rossiter of Imperial College of London)

Four leaves from watercress lines Wx_0001, Wx_0011, Wx_0033 and Wx_0038 (three replicates for each line) were taken from 7 week old plants grown in the controlled environment using optimal watercress growing conditions. The material was freeze dried for 24 hours and then weighed to 5 decimal places. After weighing, 1ml methanol is added to the centrifuge tube with a bashing bead. Samples were then put on a tissue lyser for 2 minutes at 30 megahertz. The liquid phase was removed and centrifuged at 12,000 rpm for 1.5minutes. The liquid phase was again removed and stored at -20°C. The samples were then placed in a Techne sample concentrator

at 40°C with a steady flow of nitrogen, the samples stayed in the sample concentrator until all the methanol had evaporated off. 16 columns (DEAE-ion exchange A25) were labelled and an extra two columns labelled for standards. 1ml sephadex was added to the columns. When drained the columns were washed twice with 1ml de-ionised water. The samples were resuspended in 1ml de-ionised water and 100µl barium acetate is added to each sample and 2mg/ml benzylglucosinolate. The columns were then washed twice with 1ml deionised water followed by 0.5ml acetate buffer and 75µl sulphatase. The columns were then left for 12 hours. The following day the columns were washed with separate 500 µl deionised water twice which was collected in an eppendorf and then transferred to a glass vial. The samples were then loaded onto the HPLC for analysis.

Calculation for concentration of glucosinolate is given in equation 3.5.

Glucosinolate concentration =	(area of glucosinolate/area of standard) X (amount of standard (µg/MW)) X (1000/mg of sample) X (response factor, 1)	**(eqn 3.5)**

3.3.17 Determination of isothiocyanates by Gas Chromatography Mass Spectrometry (GC-MS) (in collaboration with Dr John Rossiter of Imperial College of London)

A small section of the stem including leaves approximately 7 cm were taken from 7 week old watercress (Wx_0001, Wx_0011, Wx_0033 and Wx_0038, 3 replicates of each) grown in a controlled environment using optimal watercress growing conditions. Some material from each line was weighed and then a bashing bead added to the eppendorf and eppendorfs placed on a tissue lyser for 2 minutes at 30 megahertz. Then spun at 12,000rpm for 2 minutes and 1ml Dimethylformamide (DMF) added and spun for a further 2 minutes. The dichloromethane liquid phase is transferred to a fresh eppendorf and the volume taken recorded. 10µl dodecanol standard was added to the samples and then the samples spun briefly. Around 1ml of the sample was transferred to an eppendorf containing magnesium sulphate and then spun for 2 minutes at 12, 000rpm and then was transferred to a GC/ MS vial. The vials were then loaded on the Gas Chromatography Mass Spectrometry apparatus. The isothiocyanate concentration was then calculated (equation 3.6).

Calculation for concentration of isothiocycanate;

(area of PEITC/area of sample) X 0.1 **(eqn 3.6)**

3.3.18 Statistical analysis

Statistical analysis was carried out using the software Minitab 15 English. All confidence limits were set at 95%. With normalised data one way ANOVA with 95% confidence intervals were used and when data did not follow a normal distribution General linear model was employed.

3.4 Results

There appears to be significant ($F_{11,71} = 8.26$, $p<0.001$) variation between the watercress lines when grown in a controlled environment. Line Wx_0039 had the lowest mean mmol Fe^{2+} equivalent per gram fresh weight whilst Wx_0033 had the highest mean mmol Fe^{2+} equivalent per gram fresh weight. Wx_0001 was fourth highest with regards to the concentration of antioxidants (Figure 3.1a).

The growth of the watercress in the controlled environment enabled selection of lines to be taken forward for analysis on breast cancer cell lines i.e. high and low antioxidant and control line. Lines Wx_0001, Wx_0011 and Wx_0038 were selected for screening. Upon application to MCF7 breast cancer cells Wx_0011 had the lowest mean IC50 value whilst Wx_0001 the highest. The IC50 value is the required amount of the watercress extract needed to cause 50% cell death of the breast cancer cells. Wx_0011 had the highest concentration of antioxidants as recorded by FRAP and also the lowest IC50 value. Wx_0001 also had one of the highest concentrations of antioxidants recorded by FRAP and Wx_0038 one of the lowest concentrations of antioxidants but their mean IC5O values do not reflect this. Wx_0001 had a higher

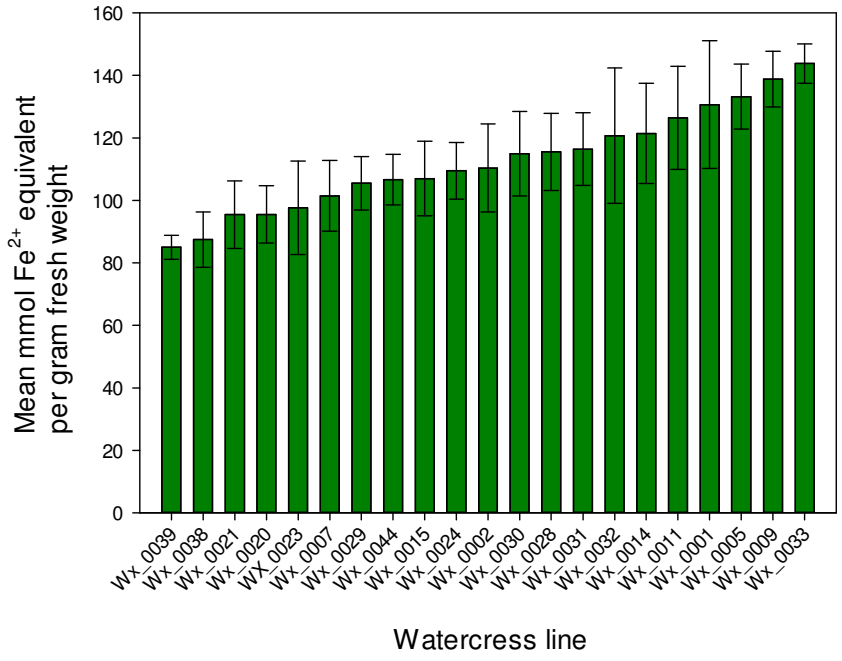
mean IC50 value than Wx_0038 thereby more Wx_0001 extract is required to kill 50% of the breast cancer cells, according to FRAP we would expect this to be line Wx_0038 (Figure 3.2) . There was a significant difference ($F_{2, 11} = 7.03$, $p<0.05$) in IC50 values between the three watercress lines.

There is variation when the lines are grown in the field with Wx_0038 remaining one of the lowest. However when in the field even though Wx_0038 still had the lowest concentrations of antioxidants, overall the concentrations were higher reaching 263 mmol Fe^{2+} equivalent per gram FW (Figure 3.1b). The highest level of antioxidants was present in Wx_0028 with 752 mmol Fe^{2+} equivalent per gram FW.

There is a dramatic increase in antioxidants between controlled environment grown and field grown watercress (Figure 3.1c). Wx_0001 increases from 130 mmol Fe^{2+} equivalent per gram fresh weight in a controlled environment to 529 mmol Fe^{2+} equivalent per gram fresh weight in the field.

There was a significant difference in the antioxidant potential between the watercress lines ($F_{11, 71} = 8.26$, $p<0.001$) and a significant difference on antioxidant potential as a result of condition ($F_{1, 71} = 966.60$, $p<0.001$). There was a significant interaction between the watercress line and condition ($F_{11, 71} = 7.28$, $p<0.001$).

a)



b)

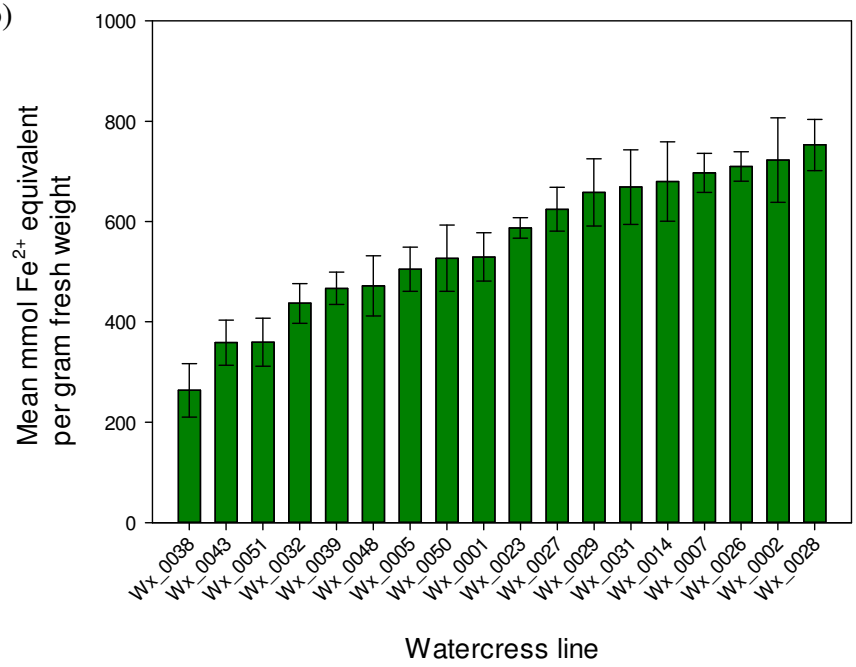


Figure 3.1: Mean mmol Fe^{2+} equivalent per gram FW (±SE) in watercress lines grown in a controlled environment, field and a comparison of both conditions;

(a) antioxidant power in a controlled environment (b) lines grown in the field – Spetisbury, Dorset and (c) antioxidant power of watercress lines grown in a controlled environment compared to those grown in the field. There was a significant difference in the antioxidant potential between the watercress lines ($F_{11, 71} = 8.26$, $p<0.001$) and a significant difference on antioxidant potential as a result of condition ($F_{1, 71} = 966.60$, $p<0.001$). NB. Fewer lines successfully germinated in the field compared to the controlled environment.

c)

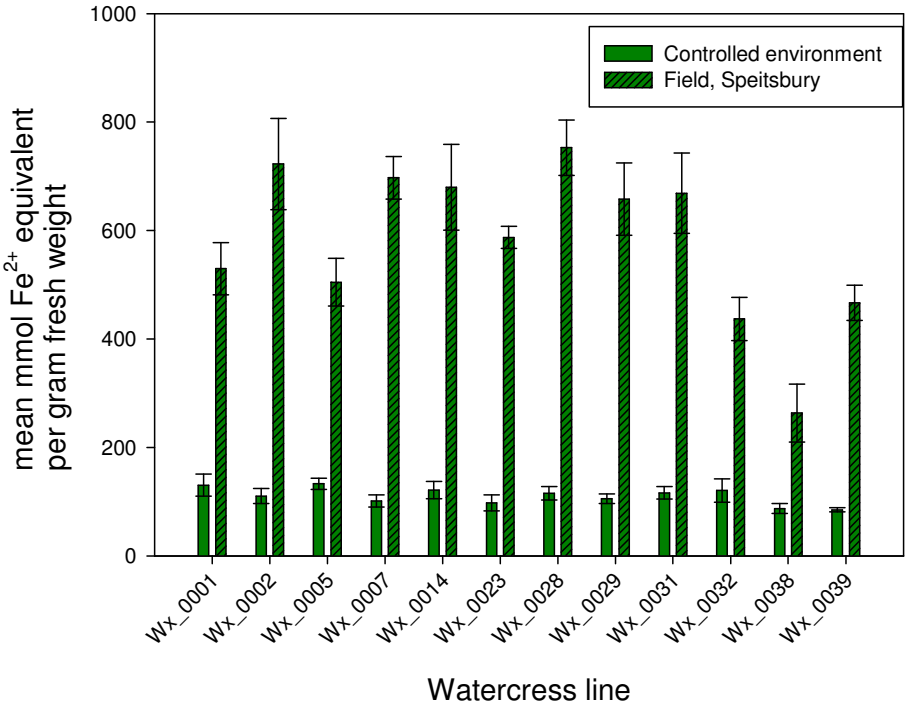


Figure 3.1 continued

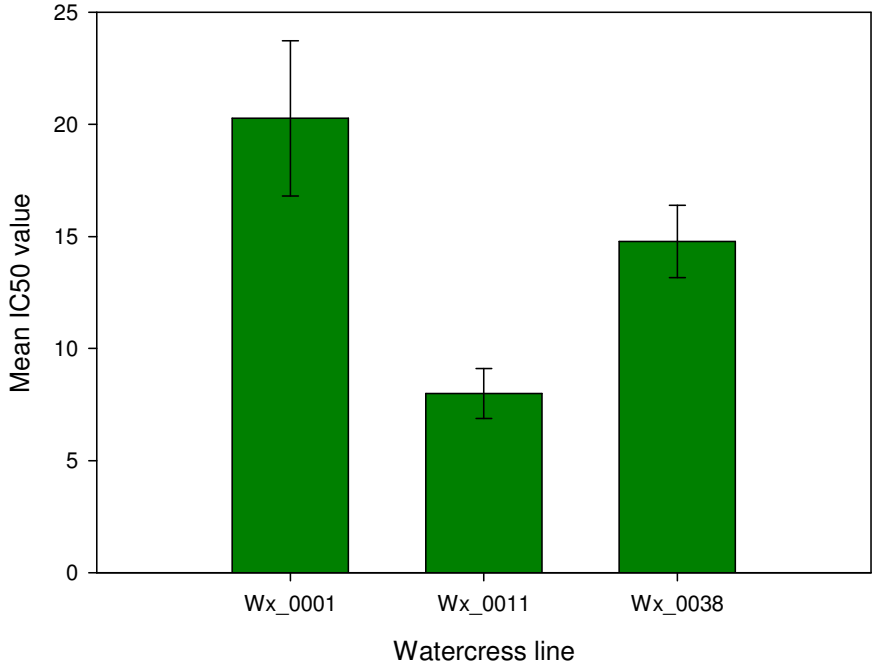


Figure 3.2: Mean IC50 value (±SE) of three different watercress lines extract on the MCF7 breast cancer cell line ($F_{2, 11} = 7.03$, $p<0.05$).

Watercress contains a higher level of antioxidants when compared to spinach and rocket. Watercress contains 562mmol Fe^{2+} equivalent per gram fresh weight followed by spinach 275mmol Fe^{2+}equivalent per gram fresh weigh and finally rocket 169mmol Fe^{2+} per gram fresh weight (Figure 3.3a). There was little variation between the level of antioxidants in organic and conventional produce, with organic having a higher level of antioxidants in rocket and watercress whilst in spinach the conventional spinach has the higher concentration of antioxidants.

There was a significant difference ($F_{2, 42} = 231.26$, $p<0.001$) between the different salad samples (rocket, spinach and watercress) but no significant difference between organic and conventional ($F_{1, 42} =0.49$, $p>0.05$). There was a significant interaction ($F_{2, 42} = 5.23$, $p<0.05$) between the salad sample and condition (i.e. organic and conventional).

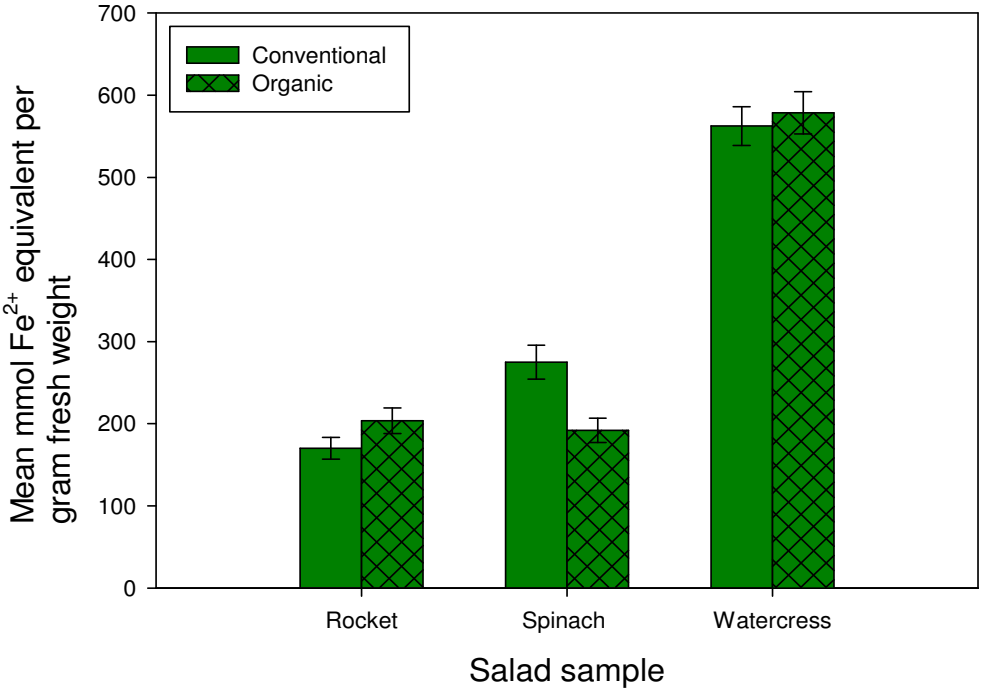
One 'ideal' trait would be a reduced stem length however since another crucial trait is high antioxidants it is important to determine which part of the plant has the highest concentration of antioxidants before aiming to reduce the stem length. It is evident from FRAP that the leaves have the highest FRAP reading, 929mmol Fe^{2+} equivalent per gram fresh weight whilst the stems the lowest, 166mmol Fe^{2+} equivalent per gram fresh weight. It also appears that the young transplanted watercress have the second highest level of antioxidants, 657mmol Fe^{2+} equivalent followed by the red watercress, 631mmol Fe^{2+} equivalent per gram fresh weight and finally whole ready to harvest watercress 604mmol Fe^{2+} equivalent per gram fresh weight (Figure 3.4b). There is significant variation ($F_{4, 10} = 22.50$, $p<0.001$) throughout the watercress plant.

Production methods are very important during salad production in order to minimize loss in quality of the salad and also antioxidants. Differences between the concentrations of antioxidants after each harvest and also the effect of storage were investigated. In two out of the three lines tested the highest concentration of antioxidants were present in the watercress after four harvests 292mmol Fe^{2+}

equivalent per gram fresh weight and 396mmol Fe^{2+} equivalent per gram fresh weight but there appears to be little variation (Figure 3.3b). The level of antioxidants increased in watercress stored at 4°C after 6 days. At day 0 the FRAP reading is 659mmol Fe^{2+} equivalent per gram fresh weight whilst after 6 days this increases to 915mmol Fe^{2+} equivalent per gram fresh weight. Between days 1 and 5 the concentrations of antioxidants increase and decrease, the largest decrease at day 4 when the concentration decreases to 557mmol Fe^{2+} equivalent per gram fresh weight (Figure 3.4a).

Harvesting does not significantly ($F_{1,8} = 1.479$, $p>0.05$) alter the concentration of antioxidants however in two of the three lines assessed the concentration of antioxidants appeared to be highest in the fourth harvest which may be due to the repeated cutting and resulting stress from harvest (Figure 3.3b).

a)



b)

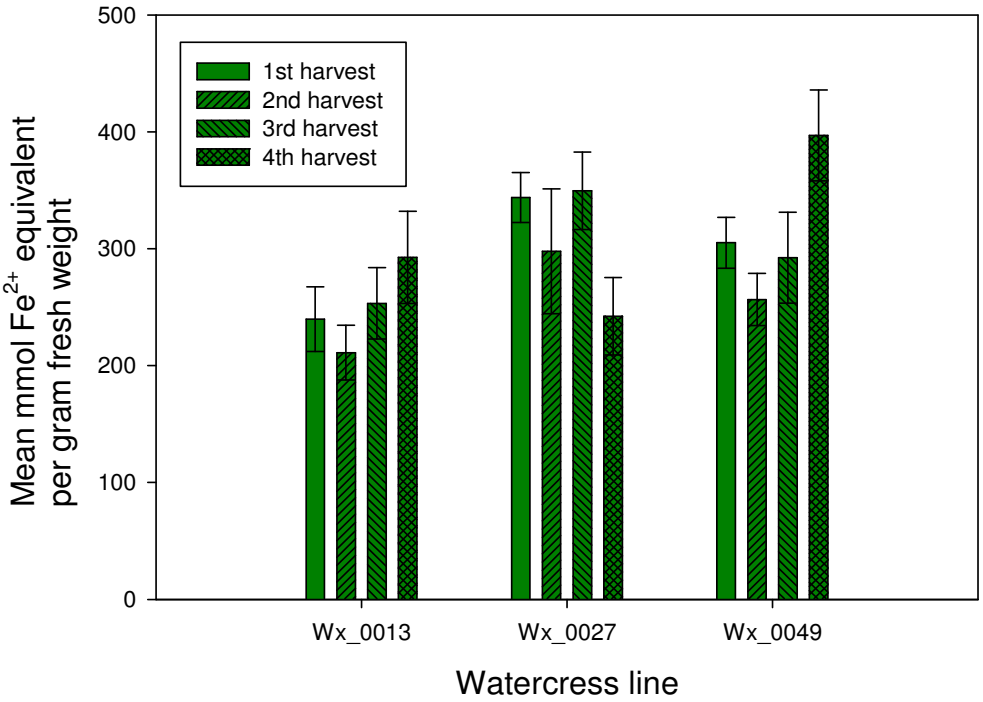
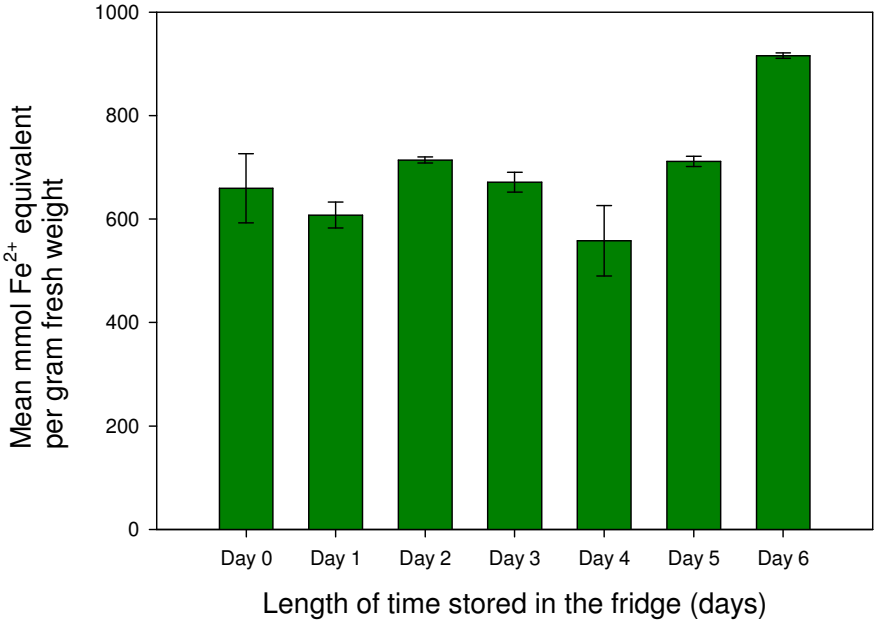


Figure 3.3: Mean mmol Fe^{2+} equivalent per gram FW (±SE) in three different leafy salads (rocket, spinach and watercress) and repeated harvest of three watercress lines (a) antioxidant power of three leafy salads (data collected by Alice Mazzer) (Significant difference ($F_{2, 42}$ = 231.26, p<0.001) between the different salad samples (rocket, spinach and watercress) but no significant difference between organic and conventional ($F_{1, 42}$ =0.49, p>0.05) and (b) antioxidant power after four consecutive harvests in three different watercress lines ($F_{1,8}$ = 1.479, p>0.05).

a)



b)

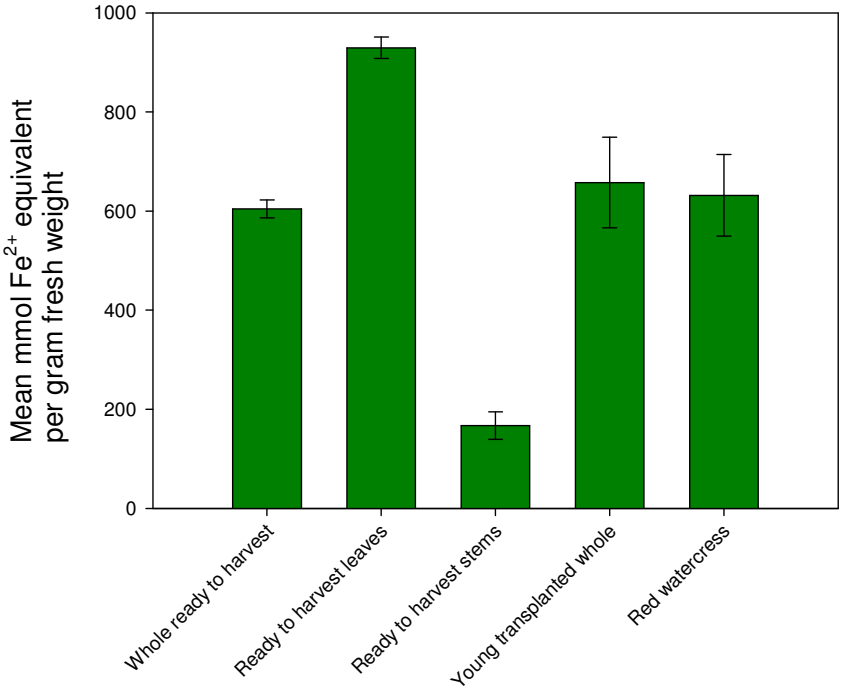


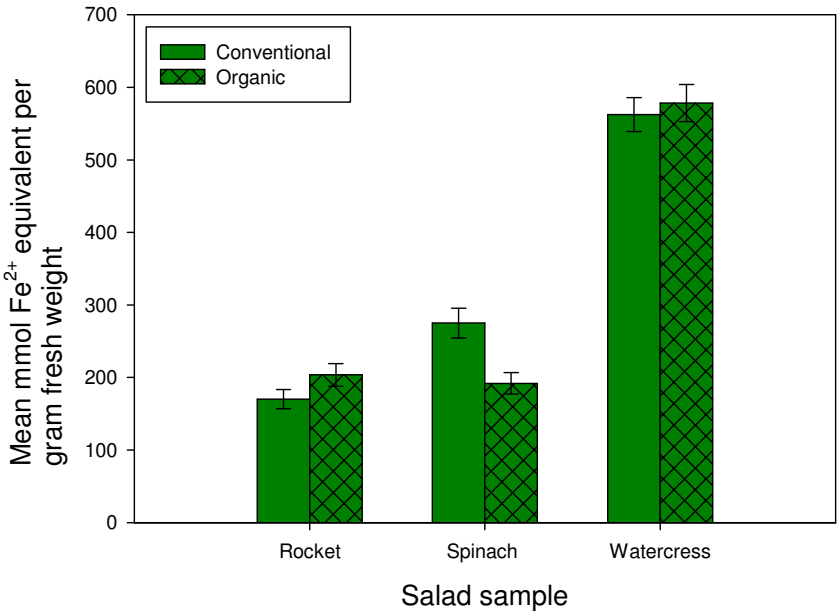
Figure 3.4: Mean mmol Fe^{2+} equivalent per gram FW (±SE) in watercress stored at 3°C over a period of 6 days and varying concentrations throughout the watercress plant

(a) Antioxidant power after storage at 3°C and (b) antioxidant power in different sections of a watercress plant ($F_{4, 10} = 22.50$, $p<0.001$).

The ferric-tripyridyltriazine (Fe^{III}-TPTZ) complex was reduced most strongly to the ferrous (Fe^{II}) by watercress, 562mmol Fe^{2+} equivalent per gram FW and organic watercress, 578mmol Fe^{2+} equivalent per gram FW (Figure 3.5a). Spinach had a lower antioxidant activity 275mmol Fe^{2+} equivalent per gram FW and 19mmol Fe^{2+} equivalent per gram FW organic spinach compared to watercress and rocket was lower than spinach, with the exception of the organic rocket (169mmol Fe^{2+} equivalent per gram FW and 203mmol Fe^{2+} equivalent per gram FW organic rocket). The different salad samples differed significantly ($F_{2, 42} = 231.26$, $p<0.001$) in their antioxidant activity however there was no significant difference ($F_{1, 42} = 0.49$, $p>0.5$) between the cultivation method. The method of cultivation significantly ($F_{2, 42} = 5.23$, $p<0.05$) differed between each salad type.

The highest percentage of radicals scavenged was observed in the watercress samples (59.59% and 61.00%) followed by spinach (27.95% and 20.18%) and then rocket (9.81% and 13.68%) which reflects the results obtained from FRAP (Figure 3.5b). The salad samples significantly differed in their ability to scavenge radicals ($F_{2, 42} = 187.2$, $p<0.05$) but there was no significant difference between the methods of cultivation ($F_{1, 42} = 0.07$, $p>0.05$). The method of cultivation did not have the same effect on each salad type and differed significantly ($F_{2, 42} = 7.29$, $p<0.05$). It appears that the different antioxidant assays rank the salad samples in the same order watercress > spinach > rocket.

a)



b)

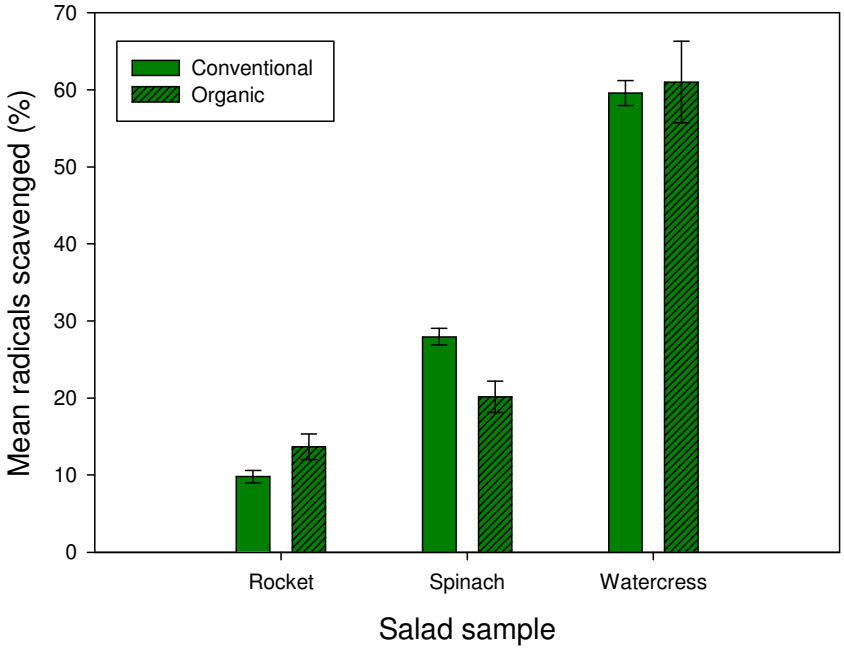


Figure 3.5: FRAP and ORAC assays on rocket, spinach and watercress;

(a) Mean mmol Fe^{2+} equivalent per gram fresh weight (±SE) in different salad samples and a comparison between organic and conventional produce (data collected by Alice Mazzer) (different salad samples differed significantly ($F_{2, 42} = 231.26$, $p<0.001$) in their antioxidant activity however there was no significant difference ($F_{1, 42} = 0.49$, $p>0.5$) between the cultivation method) (b) Mean percentage radicals scavenged (±SE) in different salad samples and a comparison of organic versus conventional produce (data collected by Alice Mazzer) (salad samples significantly differed in their ability to scavenge radicals ($F_{2, 42} = 187.2$, $p<0.05$) but there was no significant difference between the methods of cultivation ($F_{1, 42} = 0.07$, $p>0.05$).

There was a clear correlation ($r^2 = 0.9808$) between the percentage of radicals scavenged in the simplified ORAC assay with the formation of Fe^{II}-TPTZ in the FRAP assay (Figure 3.6). As the percentage of radicals increased so did the formation of Fe^{II} – TPTZ supporting the idea that both methods reflect total antioxidant activity (Ou *et al*., 2002). It was clear that watercress ranked highest for antioxidant content followed by spinach and finally rocket with both FRAP and ORAC supporting these findings. The correlation ($r^2 = 0.9808$) between the assays is evident as the percentage of radicals scavenged in the ORAC assay increases as does the formation of the Fe^{2+} complex.

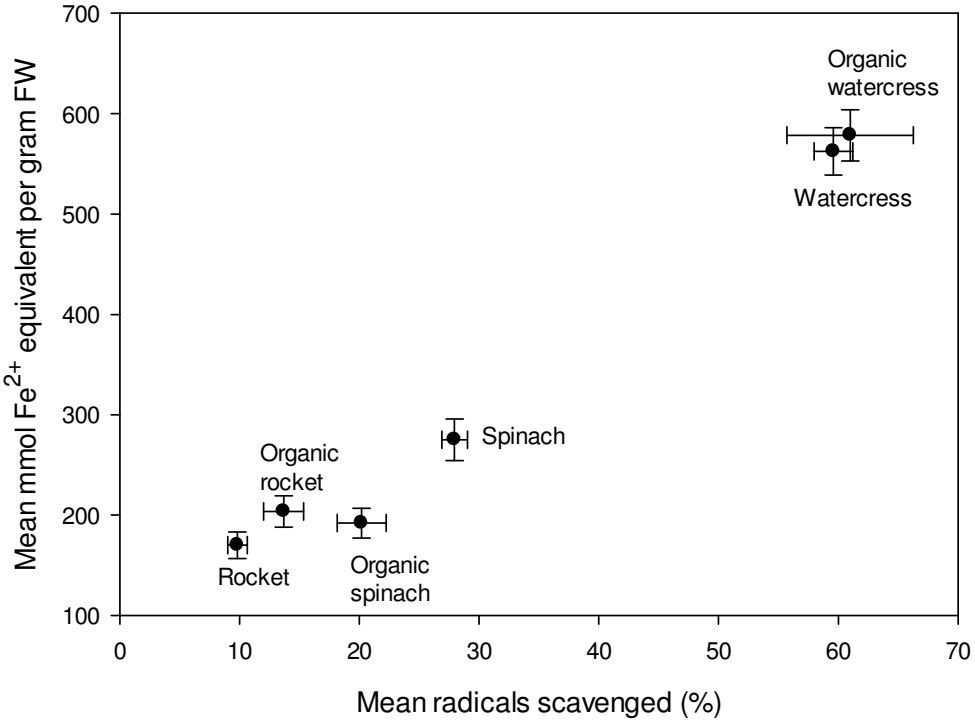


Figure 3.6: Correlation between the FRAP (±SE) and ORAC (±SE) assay;

Three different salad samples (rocket, spinach and watercress) and two different production methods (conventional and organic) were used in the antioxidant assays (data collected by Alice Mazzer).

Overall antioxidant power was assessed using FRAP and HPLC was employed to record the concentration of glucosinolate. The concentration of phenethyl glucosinolate varied from 2.74 µmoles per gram dry weight of watercress up to

19.35 µmoles per gram dry weight of watercress (Figure 3.7). The lowest concentration of phenethyl glucosinolate was observed in a green watercress cultivar and the highest concentration in the standard Vitacress line, Wx_0001. However there were not three biological replicates for Wx_0001 due to loss of sample during maceration of the tissue. The highest then is Wx_0011, 10.44 µmoles per gram dry weight of watercress. A difference between the highest and lowest watercress lines of 7.71 µmoles per gram dry weight of watercress was observed. The results are of a similar pattern to the results obtained from FRAP whereby Wx_0001 and Wx_0011 have the highest level of phenethyl glucosionlate, however Wx_0033 had one of the highest FRAP recordings yet is one of the lowest with regards to phenethyl glucosinolate and Wx_0038 has a higher concentration of phenethyl glucosinolate yet had a much lower FRAP reading compared to Wx_0033. There is no significant difference ($F_{5, 10} = 2.04$, $p>0.05$) in the concentration of phenethyl glucosinolate between the different watercress lines.

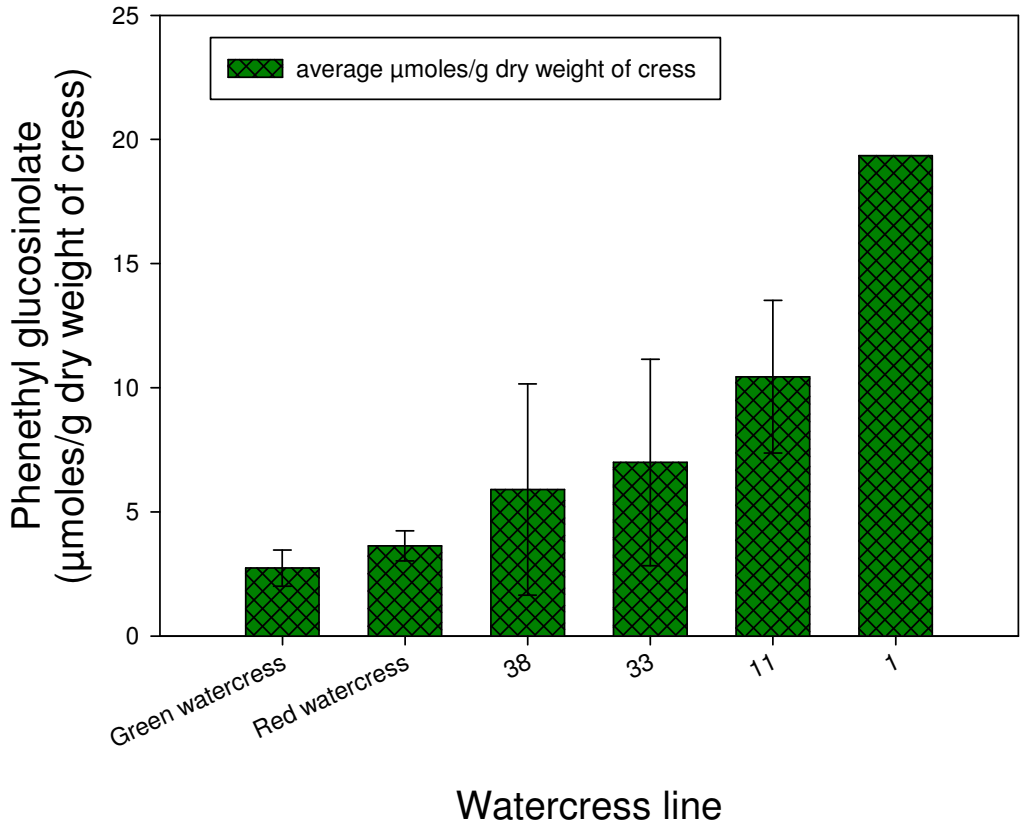


Figure 3.7: Phenethyl glucosinolate concentrations (µmoles/g dry weight of cress) (±SE) in selected watercress lines (Wx_0001 = standard line, Wx_0011 = high antioxidant line, Wx_0033 = high antioxidant line, Wx_0038 = low antioxidant line, green and red watercress = two commercial lines) ($F_{4, 14} = 2.77$, $p>0.05$)

GC/MS analysis was then used to deduce the concentration of PEITC in the samples. The concentrations of PEITC range from 0.09mg/ml in Wx_0038 to 0.20mg/ml in Wx_0001, an increase of 0.11mg/ml (Figure 3.8). Wx_0033 had 0.16mg/ml PEITC whilst Wx_0011 0.09mg/ml. The ranking order obtained for PEITC matches that of FRAP with the exception being that Wx_0033 had the highest FRAP reading followed by Wx_0001, however it appears that Wx_0001 has the higher concentration of PEITC. There was no significant difference ($F_{3,7} = 0.90$, $p>0.05$) in the concentration of PEITC between the different watercress lines.

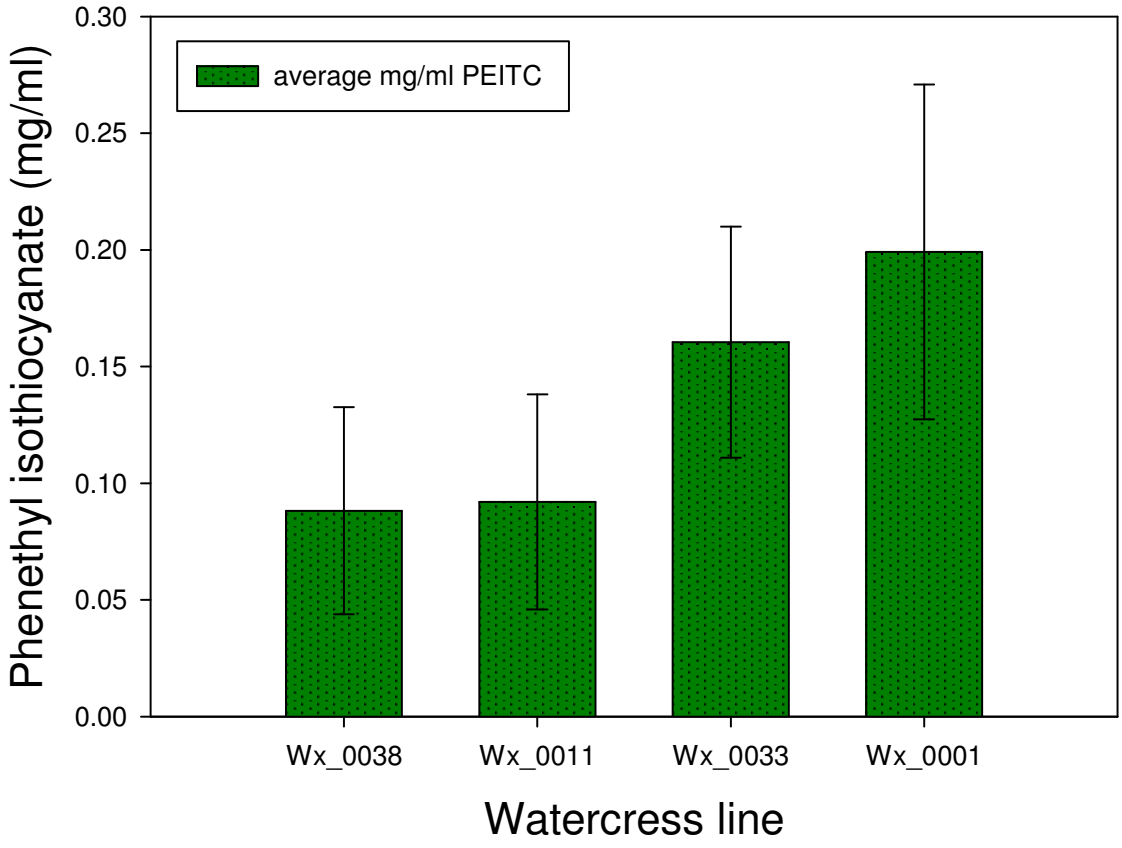


Figure 3.8: Concentration of phenethyl isothiocyanate (mg/ml PEITC) (±SE) in selected watercress lines (Wx_0001 = standard line, Wx_0011 = high antioxidant line, Wx_0033 = high antioxidant line, Wx_0038 = low antioxidant line, green and red watercress = two commercial lines) ($F_{3, 10} = 0.90$, $p>0.05$).

It can clearly be seen that although Wx_0033 has a low concentration of the phenethyl glucosinolate it has one of the highest concetrations of PEITC and therefore a very efficient conversion (Figure 3.9). Surprisingly Wx_0038 also has a high concentration of PEITC, with a higher conversion that Wx_0001 and Wx_0011.

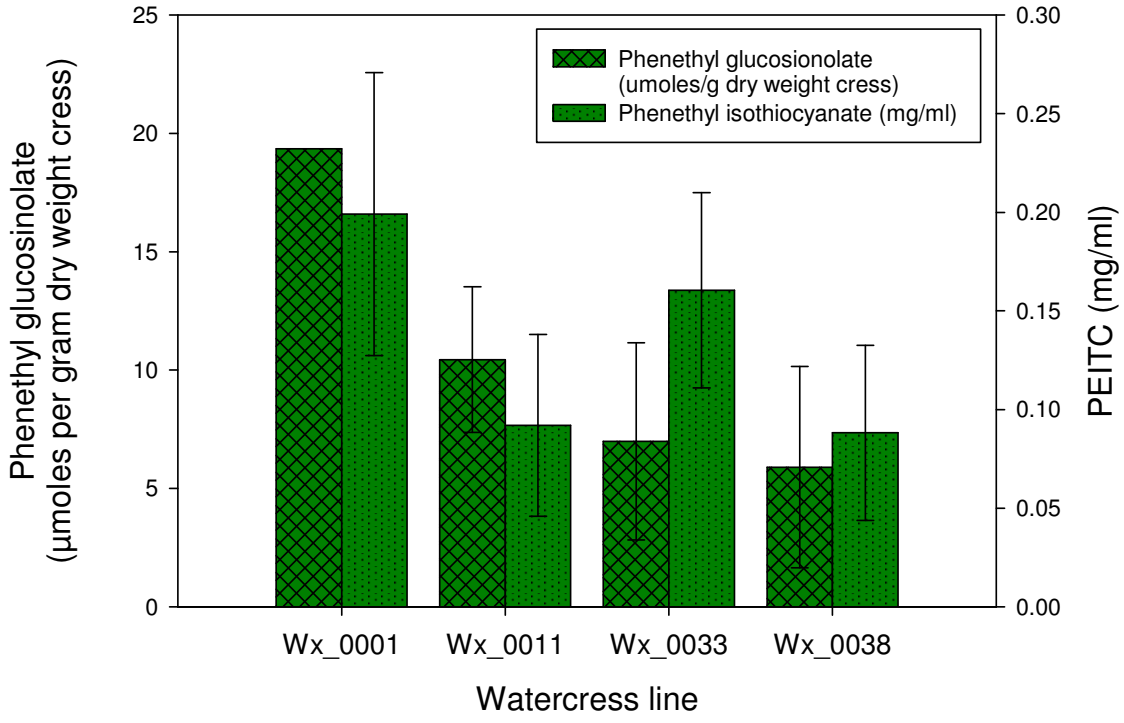


Figure 3.9: Comparison of the concentrations of phenethyl glucosinolate (µmoles per gram dry weight cress) (±SE) and phenethyl isothiocyanate (mg/ml) (±SE) in the selected watercress lines (Wx_0001 = standard line, Wx_0011 = high antioxidant line, Wx_0033 = high antioxidant line, Wx_0038 = low antioxidant line, green and red watercress = two commercial lines).

3.5 Discussion

There is a clear variation in the concentration of antioxidants between the watercress lines held in the germplasm collection and the concentration of antioxidants is exemplified in the field. This exciting variation in the concentration of antioxidants lends well to the breeding programme as the opportunity to breed watercress with a higher concentration of antioxidants compared to the industrial standard (Wx_0001) is feasible. Indeed broccoli varieties have been observed to vary in glucosinolate content which may be related to differences in genotype or may include effects of the environment, changing with the farm soil, season or harvest (Jeffery *et al.*, 2003). The fraction of glucoraphanin, which is the major alkyl glucosinolate in broccoli and is converted to sulforaphane varies with variety (Jeffery *et al.*, 2003). Sulforaphane is biologically active and the finding that it varies with variety suggests that the pathway which directs the formation of sulforaphane may be under genetic regulation (Jeffery *et al.*, 2003). Goffman *et al*. 2002 investigated the genetic variation for tocopherol content in 87 winter rapeseed genotypes. All four derivatives (alpha-, beta-, gamma- and delta tocopherol) were investigated and there were significant genotypic and environmental effects (Goffman and Becker, 2002). The study concluded that the range of genetic variation within the germplasm collection provided enough genetic variation for a possible classical breeding programme aimed at tocopherol content (Goffman and Becker, 2002). Vitamin C content has been shown to vary considerably among *Brassica* and their subspecies (Podsedek, 2007).

Howard *et al*. 2002 reported an effect of genotype on phenolic metabolism and antioxidant capacity in spinach (Howard *et al.*, 2002). This allows breeders to select for increased phenolics, individual and total flavonoids and antioxidant capacity which can then be taken forward for breeding with commercial cultivars (Howard *et al.*, 2002). Therefore the variation evident in the watercress collection may be manipulated for further benefits in breeding with respect to antioxidants as demonstrated with spinach by Howard *et al*. 2002. It is also possible to report variation, as like Jeffery *et al.* 2003, between different watercress lines. Another

possibility highlighted by this paper was the possibility to grow the spinach in different seasons or stress conditions in order to increase the concentration of antioxidants (Howard *et al.*, 2002), this chapter has also explored the effect of agronomy on the concentration of antioxidants.

By discovering the watercress lines with the highest concentration of antioxidants they can be used in crossings with the aim of enhancing the level of antioxidants. Broccoli has recently been bred with a higher concentration of antioxidants in which the genes of interest were found in the wild relative. The watercress germplasm collection has been sourced from around the world hence may contain some genes which result in high antioxidants locked away within the collection. The segments of the wild broccoli genome, *Brassica villosa*, were introgressed into the cultivated broccoli in order to enhance the level of 3-methylsufinylpropyl (3-MSP) and 4-methylsufinylbutyl (4-MSB) which are phase II detoxification enzymes (Mithen *et al.*, 2003). The cultivated broccoli lines have between 3 and 10µmol g^{-1} dry weight of methysulphinalkyl glucosinolates in their florets whereas wild species have between 50 and 100µmol g^{-1} of glucosinolates. However a precise comparison of wild and cultivated accessions is not possible due to the differences in floret morphology and the absolute concentration of glucosinolates in all materials depends on growth conditions (Mithen *et al.*, 2003).

The data presented in this chapter clearly demonstrates an increase in antioxidant concentration. An explanation for the evident increase in antioxidants when the watercress lines are grown in the field could be due to the surrounding environmental fluctuations. Plants encounter many environmental stresses including drought, high salinity, temperature extremes, hypoxia, mineral nutrient deficiency, metal toxicity, pollutants and increased UV-B radiation (Smirnoff, 1998). It is recognised that most of the cultivated species do indeed have wild relatives which can exhibit excellent tolerance to the various abiotic stresses encountered (Bohnert *et al.*, 1995). As a response to stress the concentration of antioxidants may increase in the plant.

One way in which a plant defends itself against herbivory is by the production of volatiles. By releasing volatiles a plant could reduce the number of herbivores by 90% (Kessler and Baldwin, 2001) and it is widely accepted that at low concentrations polyphenolics can defend against herbivores (Lill and Marquis, 2001). It is reasonable then to believe that organic crops which undergo a greater degree of pest attack when compared to highly fertilised conventionally grown crops may develop greater indolyl glucosinolate concentration (Jeffery *et al.*, 2003). Herbivory has been shown to increase glucosinolate content in radish (Agrawal *et al.*, 2010). In wild radish there was little variation in glucosinolate content however upon damage these concentrations then increase (Agrawal *et al.*, 2010). It was the indolyl, non-indolyl and total glucosinolates which increased (Agrawal *et al.*, 2010).

Results do not demonstrate any significant difference between the concentrations of antioxidants in organic and conventionally grown samples of rocket, spinach and watercress.It may be hypothesised that an organically cultivated crop may have a higher concentration of antioxidants compared to a conventionally grown crop due to not receiving protective pesticides/insecticides as antioxidants are coupled to stress responses. However in the data presented here this is not true for the analysis which was performed and an increase in antioxidants is not observed.

However we see an increase in the concentration of antioxidants when the lines are transferred from the controlled environment to the field. PEITC has been reported to be affected by photosynthetic period (Palaniswamy *et al.*, 1997) and also temperature (Engelen-Eigles *et al.*, 2006) which fluctuates more in the field compared to the set conditions of the controlled environment.

Another source of fluctutation in antioxidants could be due to changes post-harvest. There have been many reports of post-harvest handling and storage resulting in loss of glucosinolates and/or the hydrolysis products (Jeffery *et al.*, 2003; Song and Thornalley, 2007). Although not in the case of watercress, it is the post harvest peeling and grating operations which cause many cells to rupture therefore bringing

into contact enzymes and their substrates (Ahvenainen, 1996). With time the quality of commercial salads has been observed to decrease with respect to visual and flavour defects (López-Gálvez *et al.*, 1997).

Post-harvest change can also occur through cooking. Reports of loss of glucosinolates and/or hydrolysis products following various cooking methods (Getahun and Chung, 1999; Song and Thornalley, 2007; Vallejo *et al.*, 2002). After cooking spinach half of the flavonoid content was dissolved in the cooking water whilst the other half remained in the cooked tissue (Gil *et al.*, 1999). It is therefore important to determine methods for harvesting, handling, storage and cooking that provide minimal loss of bioactivity (Jeffery *et al.*, 2003).

In addition to the variation between different watercress lines and possible effects due to agronomy it must be taken into consideration how antioxidants may vary throughout the plant. Palaniswamy *et al*. (2003) highlighted the importance of locating the organ on the plant where glucosinolates are most concentrated. If the organ where the glucosinolates are most concentrated is located then it would be possible to extract these glucosinolates for further analysis (Palaniswamy *et al.*, 2003). This would then lead to exciting possibilities of dietary supplements (Palaniswamy *et al.*, 2003).

Kim and Ishii. (2006) investigated the variation of glucosinolates in seeds, leaves and roots of rocket salad. The highest total glucosinolate content was observed in the seeds and the lowest in the leaves (Kim and Ishii, 2006). Palaniswamy *et al*. (2003) reported a 150% increase in the level of PEITC produced after 40 days compared to that present in young transplanted watercress leaves. Another study carried out by Palaniswamy *et al.* (1995) in which cultural conditions were examined with the intention to optimize chemopreventative agents in watercress, suggested that the leaf is the major site of synthesis and storage of PEITC (Palaniswamy *et al.*, 1995). This study also reveals that the leaves have the highest concentration of antioxidants as

determined by FRAP. Therefore by reducing the stem length, as desired in Chapter two, this should not impact on the concentration of antioxidants.

Watercress plant tissue biomass was investigated under different nutrition regimes, the carotenoid and glucosinolate concentration increased with fertility management; tissue pigment concentrations responded to changes with nitrogen treatment whilst gluconasturiin responded to changes in sulphur fertility (Kopsell *et al.*, 2007). Therefore different nutrient regimes could be manipulated in order to increase the antioxidant concentration.

Watercress is a popular salad accompliment as is spinach and rocket and different vegetables have a range of antioxidant activity which is likely determined by a range of antioxidant components e.g. alpha-tocopherol, beta-carotene, vitamin C, selenium or phenolic compounds (Ismail *et al.*, 2004). The new leaf species (e.g. rocket, watercress and mizuna) are now receiving attention for their antioxidant activity replacing broccoli, cabbage and kale which were more commonly studied (Martinez-Sánchez *et al.*, 2007). It is deemed impractical to quantify all of the compounds in plants which demonstrate antioxidant activity therefore different assays have been developed to quantify the total antioxidant capacity of plant extracts (Howard *et al.*, 2002). The assays (FRAP and ORAC) shows that the antioxidant status of different salad samples (i.e. rocket, spinach and watercress) varies and therefore an intake of a variety of salads could help boost the antioxidant status of individuals. Through selected crop breeding we could achieve crops with an enhanced antioxidant status (nutritional content) thereby increasing the bioavailability of important nutritional components to the consumer.

A study by (Lako *et al.*, 2007) demonstrated that green leafy vegetables have the highest antioxidant capacity followed by fruits and root crops. We can clearly see that watercress has a higher antioxidant status when compared to both rocket and spinach. Out of the three salad samples tested rocket had the overall lowest antioxidant activity. However for the purpose of this study the antioxidant value

recorded was based solely on what was available to the consumer at that current time and did not therefore account in differences in cultural practice, country of origin etc. These results revealed 'on the shelf antioxidant value' and what a consumer would obtain on that day from each of these commonly consumed salads. Martinez-Sanchez *et al*. 2008 reported higher antioxidant activities using the assays FRAP, ABTS and DPPH for watercress when compared to both wild rocket and salad rocket. Ou *et al*. 2002 reported spinach to have one of the highest FRAP and ORAC values when compared to white cabbage, carrots, snap beans, cauliflower, white onions, purple onions, broccoli, tomatoes, beets, peas, red peppers and green peppers. However a comparison to watercress or rocket was not recorded. Hedges and Lister (2005) describe watercress as a 'nutritional heavyweight' which is 'rich in both core nutrients and phytochemicals'. Vitamin C can act as a protective antioxidant and may protect against certain cancers, rocket was recorded as having 15 mg Vitamin C whilst spinach 16 mg and watercress 75mg per 100g fresh raw vegetable (Hecht and Lister, 2005). Watercress also had a higher Vitamin C content than radish, lettuce, cucumber and celery (Hecht and Lister, 2005). Holversen *et al.* (2002) and Dragland *et al.* (2003) report a 1000-fold difference when recording total antioxidants in dietary plants and herbs.

A number of methods have been devised in which to measure the dietary antioxidants in food. The methods focus on different mechanisms of the antioxidant defence system i.e. scavenging of oxygen and hydroxyl radicals, reduction of lipid peroxyl radicals, inhibition of lipid peroxidation or chelation of metal ions (Pulido *et al*. 2000). These methods determine the ability of antioxidant to scavenge free radicals in the medium such as ORAC (Cao *et al.*, 1995) and also TEAC and TRAP assays. The DPPH method measures the scavenging ability of stable radical species by antioxidants whilst the FRAP devised by Benzie and Strain (1996) was based on the reducing ability of plasma as a measure of its antioxidant ability and has been adapted for food samples (Cao *et al.*, 1995), (Pulido *et al.*, 2000). The results presented suggest that FRAP and ORAC have comparable results as both rank the samples tested in the same order of watercress>spinach>rocket.

Halvorsen *et al* 2002 reported numerous benefits of the FRAP assay. The FRAP assay is the only assay that directly measures antioxidants or reductants in a sample as the other assays are more indirect measuring inhibition of free radicals (Halvorsen *et al.*, 2002). The values expressed from the FRAP assay express the corresponding concentration of electron donating antioxidants with the reduction of the ferric ion (Fe^{3+}) to the ferrous ion (Fe^{2+}) (Halvorsen *et al.*, 2002). However FRAP does not react with thiols as they are generally below that of the Fe^{3+}/Fe^{2+} half reaction. Only a limited amount of the glutathione in plants are absorbed by humans with almost no other antioxidant thiols present in dietary plants, indicating FRAP is a suitable assessment for the total antioxidants in plants for human consumption (Halvorsen *et al.*, 2002).

Ou *et al.* (2002) reported, in contrast to the results reported here, that FRAP and ORAC values did not correlate well. The antioxidant activity of 927 freeze-dried vegetable samples were recorded and the FRAP and ORAC values were dependent on species and also the geographical origin and harvest time (Ou *et al.*, 2002). The variability between these methods was concluded to be due to the chemistry behind the assays and that the ORAC method was chemically more relevant to chain-breaking antioxidant activity (Ou *et al.*, 2002). Drawbacks of FRAP were also reported such as inference, reaction kinetics and quantitation methods (Ou *et al.*, 2002).

Martinez-Sánchez *et al.* (2007) reported a high correlation between ABTS, FRAP and DPPH values. With regards to the FRAP and ABTS assay watercress had the highest mg $100g^{-1}$ antioxidant reading and salad rocket the lowest (Martinez-Sánchez *et al.*, 2007) which reinforces our findings that watercress had the highest reducing power (i.e. Fe^{3+}-TPTZ to Fe^{2+}-TPTZ) and radical scavenging ability whilst rocket the lowest.

The two assays used, FRAP and ORAC, do indeed differ by oxidising reagent (Schlesier *et al.*, 2002). The simplified ORAC assay uses radicals whilst FRAP

employs metal ions for oxidation (Schlesier *et al.*, 2002). Therefore the simplified ORAC assay acts by radical reduction using preformed radicals hence a decrease in absorbance whilst the FRAP assay measures the increased absorbance of ferrous ions (Schlesier *et al.*, 2002). However in these salad crops we do see comparative results and Scheliser *et al.* (2002) investigated many different antioxidant assays which gave an idea of the efficacy of plant secondary plant products but strongly recommended that at least two methods are used depending on antioxidant potential expected and the origin of substance (Schlesier *et al.*, 2002).

Glucosinolates (ß-thioglucoside-N-hydroxysulfates) are the precursors of isothiocyanates and are present in 16 families of dicotyledenous angiosperms which include a large number of edible species (Fahey *et al.*, 2001). At least 120 glucosinolates were identified in these plants however only small numbers of such compounds are present in closely related taxonomic groups, it is reported that at least 500 species of non-cruciferous dicotyledonous angiosperms contain one or more of the glucosinolates (Fahey *et al.*, 2001). It is interesting to examine any variation in glucosinolate content between cultivars for an enhanced breeding programme hence a cultivar with high beneficial antioxidant content. The content of glucosinolates has been demonstrated to be highly variable (Fahey *et al.*, 2001; Farnham *et al.*, 2000; Kushad *et al.*, 1999) giving plenty of scope for a potential breeding programme.

Early work on the identification of glucosinolates involved paper or thin-layer chromatography (Danielak and Borkowski, 1969; Fahey *et al.*, 2001). Work then progressed to steam distillation and titration of volatile isothiocyanates; UV spectroscopy of oxazolidinethiones; gas chromatography of volatile isothiocyanates; gas chromatography/UV spectroscopy; UV spectroscopy of thiourea derivatives of isothiocyanates; gas chromatography of trimethylsilyl derivatives of glucosionlates (Fahey *et al.*, 2001). The glucosinolate content of 297 wild plant species was examined by gas chromatography (GC) and then confirmed with mass spectrometry (MS) (Daxenbichler *et al.*, 1991). High Performance Liquid Chromatography (HPLC) is currently the most widely used method (Fahey *et al.*, 2001). Zang (2004) described an assay whereby isothiocyanates react with 1,2-benzenedithiol to form a

cyclic thicarbonyl reaction product 1,3-benzodithiole-2-thione, a cyclocondensation reaction. This then allows individual isothiocycanates to be identified (Zhang, 2004).

The results obtained from HPLC and GC-MS analysis show a clear trend in variation albeit non-significant between some of the lines (Wx_0001, Wx_0011, Wx_0033 and Wx_0038). Interestingly Wx_0001 had the highest antioxidant power and is shown here to have the highest concentration of phenethyl glucosinolate and phenethyl isothiocyanate. It appears that although Wx_0033 has one of the lowest concentrations of phenethyl glucosinolate it has a high concentration of phenethyl isothiocyanate suggesting a very efficient conversion from the glucosinolate to isothiocyanate. The main products of glucosinolate degradation on the autolysis of seeds or leaves were isothiocyanates, no thiocyanate (Gil and MacLeod, 1980). The primary compound from the leaves of the watercress was 2-phenethyl isothiocyanate which had 80% relative abundance (Gil and MacLeod, 1980). Wx_0011 had the second highest concentration of phenethyl glucosinolate and the third highest concentration of phenethyl isothiocyanate. FRAP measures the overall antioxidant power and was used purely to provide a ranking of the watercress with regards to antioxidant properties. This then allowed the lines to be selected for GC/MS and HPLC; highest (Wx_0033), control (Wx_0001) and lowest (Wx_0038). We can clearly see that Wx_0038 did indeed have one of the lowest phenethyl glucosinolate concentrations and the lowest phenethyl isothiocyanate concentration.

Until recently very little research was placed into the beneficial compounds present in watercress. Early work began in 1983 using GC/MS to examine volatiles from the epicuticular wax of watercress. Among the compounds detected in the wax was 2-phenethyl isothiocyanate (Spence and Tucknott, 1983). It was reported that the epicuticular wax traps and concentrates the volatile compounds released and it is these compounds which contribute to the aroma given off when the plant deteriorates (Spence and Tucknott, 1983). We used HPLC to detect total glucosinolates and GC/MS to measure the level of phenethyl isothiocyanate.

Optimal release of phenethyl isothiocyanate is dependent on time, temperature and pH (Ribnicky *et al.*, 2001). Ribnicky *et al*. (2001) investigated the release of PEITC from different cruciferous vegetables and reported Upland cress to release the greatest amount (186µg/g FW) followed by watercress (75µg/g FW) interestingly broccoli was reported as 25µg/g FW, the results were deduced by GC/MS. The release of PEITC from fresh watercress can vary in the source material, the maximal release of PEITC from fresh watercress ranged from 0.1mg/g FW to 3.4mg/g FW depending on the growing conditions or produce quality (Ribnicky *et al.*, 2001). Palaniswamy *et al*. (1997) reinforced this by demonstrating that the concentrations of PEITC ranged from 3-6.7mg/g DW depending on temperature and photoperiod (Ribnicky *et al.*, 2001).

Palaniswamy *et al*. (1997) concluded that watercress grown under 12hours light had more PEITC than plants under 8 hours light. Also that when grown in a temperature of 25°C the PEITC content was higher than those at 15°C (Palaniswamy *et al.*, 1997). Interestingly Palaniswamy *et al*. (2003) reported ontogenic variations of PEITC concentrations in watercress leaves, the highest concentration was reported in the top of the mature leaves between 40 and 60 days after transplanting (Palaniswamy *et al.*, 2003).

The total glucosinolate concentration, along with other plant secondary metabolites, varies due to environmental fluctuations such as photoperiod, light quality and temperature (Engelen-Eigles *et al.*, 2006). Also the glucosinolate profile can change throughout the life cycle of cruciferous plants (Svanem *et al.*, 1997). Watercress is not only grown commercially in streams but also in hydroponics in glasshouses (Engelen-Eigles *et al.*, 2006). By investigating all of the potential sources of variation in the level of glucosinolates it will maybe possible to integrate this into watercress production.

There are a few differences in the results reported by Palaniswamy *et al*. (1997) and Engelen-Eigles *et al.* (2006). They varied in their opinion of which photoperiod and

temperature optimised gluconasturtiin concentration however both were reporting using watercress of different ages. Palaniswamy *et al*. (1997) initiated the experiment 21 days after planting and then harvested 33 days after treatment initiation hence around 7 weeks whilst Engelen-Eigles *et al.* (2006) used younger plants which were harvested at 21 and 28 days, 3-4 weeks old. Engelen-Eigles *et al.* (2006) also argue to measure the naturally occurring glucosinolate compound gluconasturtiin instead of PEITC. Despite any differences Engelen-Eigles *et al.* (2006) reported that the photoperiod, temperature and light quality of the light source can be changed in order to optimise the quality and health benefits from watercress.

There is variation in the level of glucosinolates between plant organs (i.e. roots, leaves, stems and seeds) (Fahey *et al.*, 2001). For example seeds or young sprouts of broccoli contain 70-100µmol total glucosinolates per gram fresh weight whilst late vegetative to reproductive stage plants of the same cultivar may only contain 1-4µmol of total glucosinolates per gram fresh weight (Fahey *et al.*, 2001). Plant age also determines the quantitative and qualitative glucosinolate composition of plants (Fahey *et al.*, 2001). Fahey (2001) outlines many environmental factors which have a significant effect on the level of specific glucosinolates in growing plants. Further work to examine the level of glucosinolates and isothiocyanates would be to investigate the effect of temperature and pH on production of these compounds.

In conclusion to this chapter it is evident that the concentrations of antioxidants vary between the different watercress lines and that some post-harvest processes can affect the concentrations of antioxidants present in the crop. This gives plenty of scope in terms of possible ways not only to breed but also to manipulate processes throughout the production line. It highlights the presence of variation already present within the crop which can be utilised in a breeding programme to produce an enhanced cultivar with beneficial nutrients. Although not significant there also appears to be a trend in the concentration of phenethyl glucosinolate and phenethyl isothiocyanate. With an increase in the replication of each line this trend may become significant hence a possible route to increase the antioxidant benefits

obtained by consuming watercress by enhancing the concentration of glucosinolates or isothiocycanates.

The outcome of this chapter does indeed suggest variation in the antioxidant capacity of the lines held within the collection and now this needs to be narrowed down to specific beneficial compounds in order to proceed with a breeding programme. This has allowed the first steps to be taken for selecting lines which naturally have a high concentration of antioxidants. It also provides an insight into how agronomic processes and post-harvest changes can also impact on the concentration of antioxidants, all of which provides important information for future progression of this crop.

3.5.1 Conclusion

1. There was a significant variation in the antioxidant power between the different watercress lines and a large increase in FRAP values when the watercress lines are grown in the field.

2. There was a significant difference in the IC50 values inferring there was a difference in the ability of the three watercress extracts to kill breast cancer cells.

3. There was a significant difference in FRAP and ORAC values obtained between watercress, spinach and rocket, with watercress having a significantly higher FRAP and ORAC value. However there was no significant difference between conventionally and organically grown crops.

4. A correlation between the two antioxidant assays, FRAP and ORAC, was observed.

5. A significant difference was recorded for different sections of the watercress with leaves having a significantly higher FRAP value.

6. The shelf life of watercress was very variable over six days and repeated harvests did not significantly affect the FRAP values.

7. There was no significant difference in the concentration of phenethyl glucosinolate and phenethyl isothiocyanate between selected watercress lines however a trend in concentrations of this glucosinolate and isothiocyanate were observed.

CHAPTER FOUR: INVESTIGATING THE GENETIC DIVERSITY OF WATERCRESS USING AMPLIFIED FRAGMENT LENGTH POLYMORPHISM AND THE WATERCRESS BREEDING PROGRAMME

4.1 Abstract

The main outcome of plant breeding is to produce a cultivar with the desired phenotype or nutritional composition for both grower and consumer benefits. However in order to establish a successful breeding programme adequate genetic diversity must be held within a collection. This genetic diversity can be assessed using a technique called Amplified Fragment Length Polymorphism (AFLP).

The genetic diversity of the watercress lines held in the germplasm collection was assessed using AFLPs as promising morphological and biochemical variation was evident. Eight primer pairs were used, three labelled with FAM and five with HEX, the FAM and HEX are the fluorescent labels on one of the two primer pairs which allow it to be visualised. Five of the eight primer pairs gave informative amplification for analysis. The AFLPs showed there was a low genetic variation present in the germplasm collection, which is clearly shown in the clustering in the PCA analysis. The lines used in the watercress breeding programme were selected using data from the morphological and biochemical study.

Amplified Fragment Length Polymorphism provides a tool by which informative decisions can be made regarding plant breeding. Recent advances in plant breeding focuses on the underlying genetics rather than selecting from favourable phenotypes. Hence this chapter links the watercress breeding programme initiated from phenotypic and biochemical results with an AFLP protocol exploring the underlying genetic diversity.

4.2 Introduction

At the turn of the 20th century Darwin and Mendel established a scientific basis for plant breeding, Darwin gave the scientific principles of hybridisation and selection whilst Mendel defined the fundamental association between genotype and phenotype (Moose and Mumm, 2008). The aim of plant breeding is to produce an improved cultivar which is suited to the needs of both the farmers and consumers i.e. improved yields and nutritional qualities (Moose and Mumm, 2008). Plant breeding consists of three simple steps; populations or germplasm collections which consist of useful genetic variation assembled and then individuals with superior phenotypes are identified and improved cultivars are developed from these individuals (Moose and Mumm, 2008). The last 30 years have seen the development of genetically engineered plants where genes are introduced by methods other than the usual sexual means (Gepts, 2002). Plant physiologists study the relationship between the crop performance (phenotype) and the environment, however with the many changing global processes (e.g. climate change) resulting in multiple challenges to crop performance changing the crop genome (genotype) through plant breeding and molecular biology maybe the better option (Edmeades *et al.*, 2004). Genomics which studies the genomes of organisms offers many promising tools which could enhance genetic gains and stabilise crop production (Edmeades *et al.*, 2004). By establishing a germplasm collection containing watercress lines sourced from around the world the aim was to build a library of valuable genetic information which can be explored further using different genetic approaches, for example Amplified Fragment Length Polymorphism (AFLP).

As watercress (*Rorippa nasturtium-aquaticum*) has not yet been sequenced there is no genetic map and so important information yielded from the model species *Arabidopsis* must be utilised. Watercress is reported to have a complex genome which allows the hybridisation with closely related species feasible (Sheridan *et al.*, 2001). Sheridan *et al.*, (2001) investigated the feasibility of hybridisation with other species and revealed that the taxa under investigation were placed in three groups: watercresses and *Barbarea verna*, the non-watercress *Rorippa* species and finally

Cardamine, Brassica and *Lepidium*. There appears to be potential for breeding programmes to be initiated and to reassess *R. palustris* and *R. sylvestris* which have previously been chosen in watercress breeding programmes (Sheridan *et al.*, 2001) due to being in the *Rorippa* species.

In order to PCR segments of the watercress genome, primers that recognise non-specific nucleotide sequences are used and then examined on an agarose gel. The process of Random Amplified Polymorphic DNA (RAPD) is relatively simple. It is based on the amplification of random DNA segments with single primers of arbitrary nucleotide sequence and therefore requires no preliminary sequence knowledge (Williams *et al.*, 1990). Sheridan *et al*. (2001) used RAPD-PCR to investigate the genetic diversity within a watercress collection and record any potential diversity that could be used in breeding programs. In the early 1990s Keygene (Blears *et al.*, 1998) developed a technique known as Amplified Fragment Length Polymorphism (AFLP) that has been adopted to identify genetic variation and has been suggested to have a higher discriminatory power in comparison with RAPDs (Virk *et al.*, 1995). The RAPD assay requires pure intact DNA whilst the AFLP has a digest and ligation step before the selective amplification whilst these steps are obsolete with respect to RAPDs. The two enzymes involved in the digest step include an enzyme with an average cutting frequency (EcoRI) and a second with a higher cutting frequency (MseI) (Savelkoul *et al.*, 1999) (Figure 4.1). By using two different enzymes the number of fragments generated for amplification and to produce a fingerprint can be manipulated to produce a fingerprint of the desired complexity (Blears *et al.*, 1998). Also when the fragments are cut with two different enzymes a large number of fingerprints can be produced using the combinations of only a few primers (Blears *et al.*, 1998).

1. DNA digested with restriction endonucleases

5’----------GAATTC----------TTAA----------3’

3’----------CTTAAG----------AATT----------5’

↓ + EcoR1, Mse1

AATTC----------T

G----------AAT

2. Double-stranded DNA adapters are ligated to the ends of the DNA fragments to generate the template for DNA amplification

primer + 1 5’----------A

AATTCN----------NTTA 5’TA Mse1 adapter

TTAA5’ TTAAGN----------NAAT

EcoR1 C----------5’

adapter

The sequence of the adapters and the adjacent restriction site serve as primers binding sites for subsequent amplification of the restriction fragments by PCR

3. Selective nucleotides extending into the restriction fragments are added to the 3’ ends of the PCR primers so that only a subset of restriction fragments are recognized

primer + 3 5’----------AAC

AATTCA----------GTTA

TTAACT----------CAAT

AAC----------5’

Selective amplification with primers +3

4. Only restriction fragments in which the nucleotides flanking the restriction site match the selective nucleotides will be amplified

AATTCAAC----------TTGTTA

TTAAGTTG----------AACAAT

Amplified fragments are analyzed by denaturing polyacrylamide gel electrophoresis to generate a fingerprint

Figure 4.1: AFLP Protocol

(http://tools.invitrogen.com/content/sfs/manuals/aflpi_man.pdf, March 2008)

As mentioned AFLP is not dependent on prior sequence knowledge and double stranded nucleotide adapters complementary to the sticky ends of the restriction site are ligated to the restriction fragments (Blears *et al.*, 1998). The sequence of the adapters and adjacent restriction half site serve as primer binding sites for subsequent PCR amplification (Blears *et al.*, 1998). Upon selective amplification the AFLP primers contain three types of DNA sequence: 5' region complementary to adapter, restriction site sequence and 3'selective nucleotides. Two AFLP primers are used: one primer is complementary to the adapter and adjacent rare-cutter restriction site sequence (EcoRI) whilst the other complementary to adapter and frequent cutter recognition site (MseI) (Blears *et al.*, 1998). Preamplification reduces the amplification products which will be used in the selective amplification. By carrying out two separate steps of amplification the efficiency and reproducibility is increased.

The higher discriminatory power of AFLP allows any variation within the germplasm collection to be highlighted. The AFLP assay has been used extensively in many plants including potato (Van Eck *et al.*, 1995), barley (Becker *et al.*, 1995), rice (Mackill *et al.*, 1996) and *Arabidopsis* (Miyashita *et al.*, 1999). In addition the AFLP assay has detected a large number of polymorphism more efficiently than RFLP (Restriction Fragment Length Polymorphism), RAPD (Random Amplified Fragment Length Polymorphism), microsatellite and SSR (Simple Sequence Repeat Polymorphism) (Miyashita *et al.*, 1999).

In plant genetics AFLP can be applied in plant molecular genetics, breeding, variety identification, germplasm management, indirect selection of agronomically important traits and backcross breeding (Savelkoul *et al.*, 1999). The AFLPs generated by complete restriction endonuclease digestion of total genomic DNA followed by selective PCR amplification and electrophoresis of a subset of the fragments result in a unique, reproducible fingerprint (or profile) for each individual (Meudt and Clarke, 2007) and therefore allows us to examine any potential genetic diversity within the established watercress collection.

As well as determining the level of genetic diversity within the watercress collection, this chapter also sort to establish a breeding programme in which a mapping population could be initiated. By creating a mapping population phenotypic effects can thus be mapped to specific genomic regions (Buerkle and Lexer. 2008). In order for genetic mapping to occur recombination is required to break up the combinations of the parental genotypes (Buerkle and Lexer. 2008). Therefore to construct a linkage map a segregating plant population is required i.e. the population is derived from sexual reproduction. The parents selected to produce the mapping population would have been chosen due to differing in one or more traits of interest (Collard *et al*. 2005), e.g. a mapping population of lettuce was created by crossing a dwarf crisphead lettuce with a butterhead lettuce cultivar (Waycott *et al*. 1999). Therefore the watercress breeding programme was attacked from two different approaches; 1. Two lines which demonstrated two extremes (i.e. a high and low antioxidant power) were selected to establish a mapping population 2. High antioxidant line crossed with the standard Vitacress line to initiate the production of an enhanced cultivar. Single seed descent allowed rapid progression in generations for both of these breeding aims as it allows homozygosity to be reached with minimal selection due to only sampling one seed per individual hence 'single seed descent' (Barkley and Chumley. 2011).

The aim of this Chapter was to explore any sequence variation between the different watercress lines held in the germplasm collection using the Amplified Fragment Length Polymorphism procedure. Following quantification of phenotypic and biochemical variation in Chapters two and three, AFLP would allow us to explore any underlying genetic diversity. Also this chapter initiated the watercress breeding programme to create a mapping population and the data collected in Chapters two and three allowed individual lines to be selected for an advanced watercress cultivar.

4.2.1 Aims

1. Chapters two and three identified variation in both the phenotype and antioxidant capabilities, lines from these chapters were selected for further breeding by single seed descent.

2. Identify any genetic diversity held within the watercress lines in the germplasm collection using Amplified Fragment Length Polymorphism.

4.3 Materials and Methods

4.3.1 Watercress crossing

Plants with the desired traits were grown for 7 weeks and bagged whilst growing to avoid out crossing. It is important to select the best flowers in terms of suitable size and low level of damage to use in the cross (Figure 4.2a). The lines chosen for the crosses were selected from data obtained in Chapters two and three. Wx_0001 as the standard line, Wx_0011 as the highest antioxidant, Wx_0023 as the shortest stem length, Wx_0029 as highest stem diameter and Wx_0038 as the lowest antioxidant, all based on controlled environment trials (Figure 4.2b). To reduce the possibility of self fertilisation, flowers from the female parents must be used before the anthers start to shed the pollen onto the stigma and the male parent should be an open flower which is visibly shedding pollen. Crossing was carried out after 2pm when the plant is most receptive. Suitable flowers were identified on the female plant and were emasculated after 2pm on the first day and tagged (personal communication with Bill Summers).

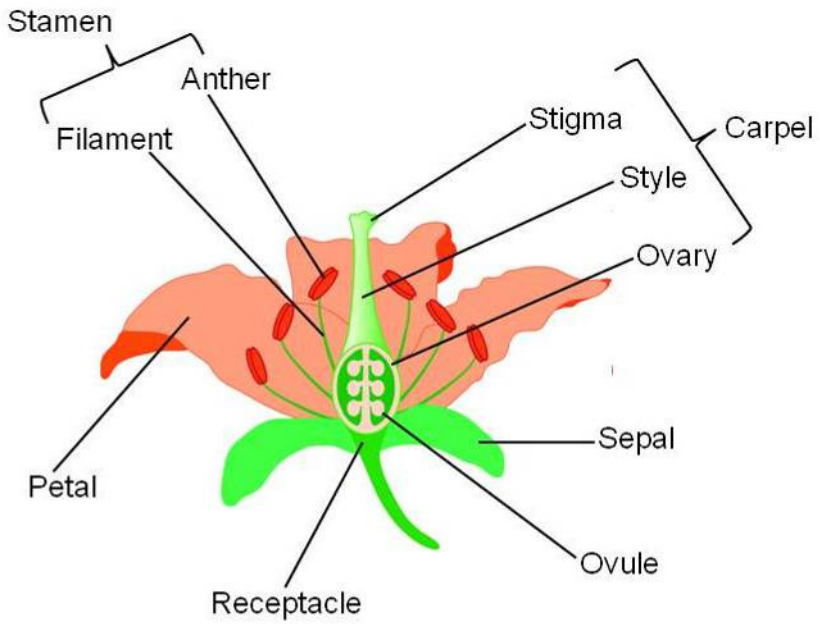
Emasculation was carried out using scissors and jeweller's forceps to remove the flowers which were just above and below (~5cm distance between each flower) the flower to be crossed. Any seed pods (siliques) which were also present were removed. For each flower which was to be crossed on the female parent the sepals,

petals and anthers were removed and the carpel was left intact. It is important that all the anthers were removed and no pollen was shed onto the carpel. The stem should have three to five exposed carpels. An open flower was removed from the male parent and squeezed near the base with forceps. The convex surfaces of the anthers were brushed against the exposed carpel on the female plant. A magnifying glass was used to confirm the pollen on the stigma. The crosses were labelled with tags which allowed the seed to be identified as it matured. Around 30 days post pollination the seed was ready to collect (adapted from Weigel and Glazebrook, 2002).

4.3.2 Single seed descent

The plants which had been crossed were grown until seeds were produced; the seed was collected from each plant. A single seed from each cross were then planted and grown until the production of seed. Again the seed was collected and another single seed progressed to the next generation, this continued until the F9 generation. Two specific crosses were chosen to progress with. The cross of Wx_0001 with Wx_0011 was progressed as this cross was conducted in order to enhance the level of antioxidants in Wx_0001. When Wx_0011 and Wx_0038 were crossed this was carried out in order to produce a mapping population for elucidation of quantitative trait loci (QTL).

a)



b)

♀		♂
Wx_0001		Wx_0038
Wx_0001		Wx_0029
Wx_0001		Wx_0023
Wx_0001		Wx_0011
Wx_0038		Wx_0001
Wx_0029		Wx_0001
Wx_0023		Wx_0001
Wx_0011		Wx_0001
Wx_0038		Wx_0011

Key; Wx_0001 standard line, Wx_0011 highest mean stem length and second highest FRAP, Wx_0023 lowest mean stem length, Wx_0029 highest mean stem diameter, Wx_0030 highest mean number leaves, Wx_0033 lowest mean stem diameter and highest FRAP, Wx_0038 lowest FRAP

Figure 4.2a) flower anatomy b) outline of lines crossed in the breeding programme (NB Figure 5.2a adapted from http://kvhs.nbed.nb.ca/gallant/biology/biology.html)

4.3.3 DNA extraction

Please note the DNA extraction protocol was carried out on tissue harvested from the watercress plants which the morphological (Chapter two) and biochemical (Chapter three) traits were recorded. Three biological replicates for each line were used.

CTAB extraction: The leaf material was ground in liquid nitrogen using a mortar and pestle and a small amount (~200mg) was transferred into an eppendorf. 900µl of CTAB which was prewarmed to 65°C was added to each tube, inverted and incubated at 65ºC for 45 minutes. The eppendorfs were inverted half way through the incubation period. Into each eppendorf 900µl chloroform/isoamyl alcohol (Chisam) was added and the tubes were shaken until well mixed and centrifuged at maximum speed (13,200rpm) for 10 minutes in Eppendorf centrifuge 5417R. 500µlof the aqueous phase was transferred to a fresh eppendorf and 50µl 3M sodium acetate and 333µlcold isopropanol added to the transferred aqueous phase. This was mixed well and incubated at -20°C for 15 minutes and then centrifuged for 10 minutes at maximum speed. The liquid phase was poured out and the pellet (DNA) kept. The tube was turned upside down and gently blotted on tissue. This was then centrifuged for 15 seconds and any remaining liquid pipetted out. The pellet was then washed with 500µl cold 70% ethanol and then centrifuged for 10 minutes at maximum speed. The liquid phase was then pipetted out and centrifuged for an additional 15 seconds to remove any remaining liquid. The pellet was then left to dry at room temperature with the lids open for 20 minutes. 5µl TE solution and 1µl RNase was added to each tube and then left incubating at room temperature overnight. The next day 200µl TE solution, 100µl 3M sodium acetate and 1ml cold absolute ethanol was added to each eppendorf and mixed well and then was centrifuged at maximum speed for 10 minutes. The liquid phase was discarded by pouring out and the pellet could be seen as a small smear on the wall of the tube. Again this was centrifuged for 15 seconds and any remaining liquid pipetted away. The eppendorfs were left to dry with the lids open for 20 minutes at room temperature and then the pellet was resuspended in 50µl water. The Nanodrop (1000

spectrophotometer, Thermo Scientific) gave an estimate of the concentration of DNA. The extracted DNA was stored at -20°C.

4.3.4 Amplified Fragment Length Polymorphism

The AFLP protocol consists of three steps: digestion and ligation of adaptor sequences, pre-selective amplification and selective amplification. The AFLP protocol was carried out using a purchased kit from Invitrogen – AFLP Core Reagent kit (catalogue #10482016).

1) Digest and ligation of adapters;

The following mix was used for the digest: 5µl 5X reaction buffer, 2µl EcoRI/MseI. 7µl was added to 2µl DNA (100-1000ng/µl) and 16µl sterile water. This was incubated at 37°C for 2 hours then 15 minutes at 70°C to inactivate the enzymes (EcoRI and MseI). The following components are then added to the digested DNA: 24µl adapter ligation solution and 1µl T4 DNA ligase. These are mixed at room temperature and then incubated at 20°C for 2 hours.

2) Pre-selective amplification;

After incubation at 20°C a 1:10 dilution was performed taking 10µl of the reaction mixture and adding 90µl of TE buffer. This was taken forward for pre-selective amplification: 10µl biomix, 0.4µl EcoPRE (10µM), 0.4µl MsePRE (10µM), (primer details in Table 6) 7.2µl sterile water and 2µl DNA mix from step 1.
These were mixed gently and the following PCR performed: 95°C for 15 minutes, 29 cycles of 94°C for 30 seconds, 56°C for 1 minute and 72°C for 1 minute followed by 72°C for 5 minutes and finally 4°C for 20 minutes.

3) Selective amplification;

To PCR tubes 10µl biomix, 0.4µl primer (10µM), 0.4µl fluorescently labelled primer (10µM) (primer details given in Table 6), 4.2µl sterile water and 5µl DNA from step 2. These were gently mixed and the following PCR performed: 94°C for 30 seconds then 13 cycles of 65°C which was lowered by 0.7°C each cycle and 72°C for 1 minute then 94°C for 30 seconds followed by 22 cycles of 56°C for 30 seconds and 72°C for 1 minute with a final hold at 4°C for 20 minutes.

5µl was examined on a 1% agarose electrophoresis gel and then 1µl sent to GeneService, Nottingham for analysis.

Table 6: Primers used for pre-selective and selective amplification in AFLP

Primer	Sequence (5' to 3')	Label
EcoPRE	GAC TGC GTA CCA ATT C	
MsePRE	GAT GAG TCC TGA GTA AC	
AACATe	GACTGCGTACCAATTCAA	FAM
AACATm	GATGAGTCCTGAGTAACCAT	None
AACTTe	GACTGCGTACCAATTCAA	HEX
AACTTm	GATGAGTCCTGAGTAACCTT	None
CACAGe	GACTGCGTACCAATTCCA	HEX
CACAGm	GATGAGTCCTGAGTAACCAG	None
CCCACe	GACTGCGTACCAATTCCC	HEX
CCCACm	GATGAGTCCTGAGTAACCAC	None
CGCACe	GACTGCGTACCAATTCCG	HEX
CGCACm	GATGAGTCCTGAGTAACCAC	None
CTCAGe	GACTGCGTACCAATTCCT	FAM
CTCAGm	GATGAGTCCTGAGTAACCAG	None
TACATe	GACTGCGTACCAATTCTA	HEX
TACATm	GATGAGTCCTGAGTAACCAT	None
TCCCTe	GACTGCGTACCAATTCTC	FAM
TCCCTm	GATGAGTCCTGAGTAACCCT	None

4.3.5 ABI Peak Scanner

Data received from GeneService was uploaded into ABI PeakScanner software v1.0 and the peaks for each sample recorded by one person for consistent classification of a peak. The peaks were then scored and analysed in Microsoft Excel using GenALex, 254 peaks were scored in total.

4.3.6 GENALEX calculations

The Analysis of Molecular Variance (AMOVA) is a relatively recent statistical procedure which allows the hierarchical partitioning of genetic variation among populations and regions and the estimation of F statistics (level of heterozygosity) (Flanagan, 2005). Wright's F-statistics (Wright, 1920), first used in 1920 by Wright in Principles of livestock breeding, are widely used to characterise population genetic structure. These statistics allow the partition of genetic diversity within and among populations. The AMOVA framework allows for statistical testing by random permutations (Flanagan, 2005) .

Calculations for the GenALex programme followed the methods of Excoffier *et al.*, (1992), Huff *et al.*, (1993), Peakall *et al.*, (1995) and Michalakis and Excoffier, (1996).

The following outputs were given from the analysis:

Pie Graph (AM) – outputs a pie chart illustrating the distribution of variance

Freq.Dist. (FDAM) – outputs the frequency distribution of permuted PhiPT/Fst/Rst values vs the observed value

Permute Values (PVAM) – outputs the AMOVA values from each permutation

Step by Step (SbySAM) – shows the step by step calculation of AMOVA

4.4 Results

4.4.1 Current breeding status

Analysis of morphological and biochemical variation presented in earlier Chapters (Chapters two and three) revealed variation between the watercress lines. This allowed lines with favourable traits (i.e. reduced stem length and high antioxidant power) to be selected for crossing. It can be seen from Figure 4.3 that the extremes

for each trait that were chosen for the breeding programme. This also shows where the standard Vitacress line, Wx_0001, falls on the frequency graph in comparison to the extremes.

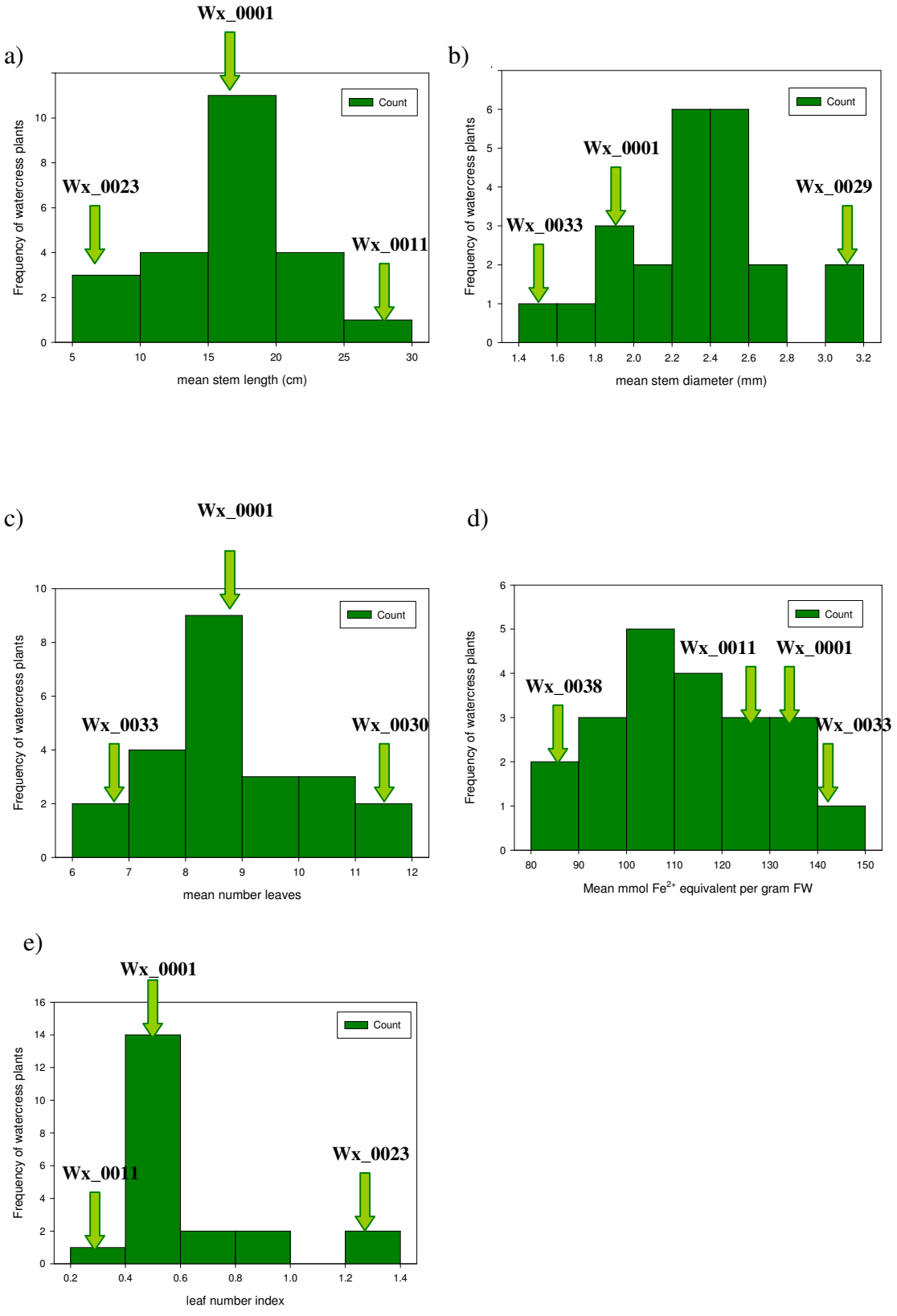


Figure 4.3: Frequency diagrams of each trait of interest (a) mean stem length (b) mean stem diameter (c) mean number of leaves (d) mean mmol Fe^{2+} equivalent per gram fresh weight (e) leaf number index

Two crosses were selected for further breeding, one cross to progress through single seed descent whilst the other cross would investigate the antioxidant power before further progression with single seed descent (Figure 4.4).

The cross of Wx_0038 (lowest antioxidant) with Wx_0011 (highest antioxidant) was taken straight through for single seed descent thereby rapidly progressing generations for a fixed phenotype. Measurements were carried out on the resultant progeny of the cross of Wx_0001 (standard line) and Wx_0011 (highest antioxidant) to examine any morphological differences and tissue frozen for antioxidant analysis. The breeding programme is currently in progression to reach the fixation of the genotype in the ninth generation.

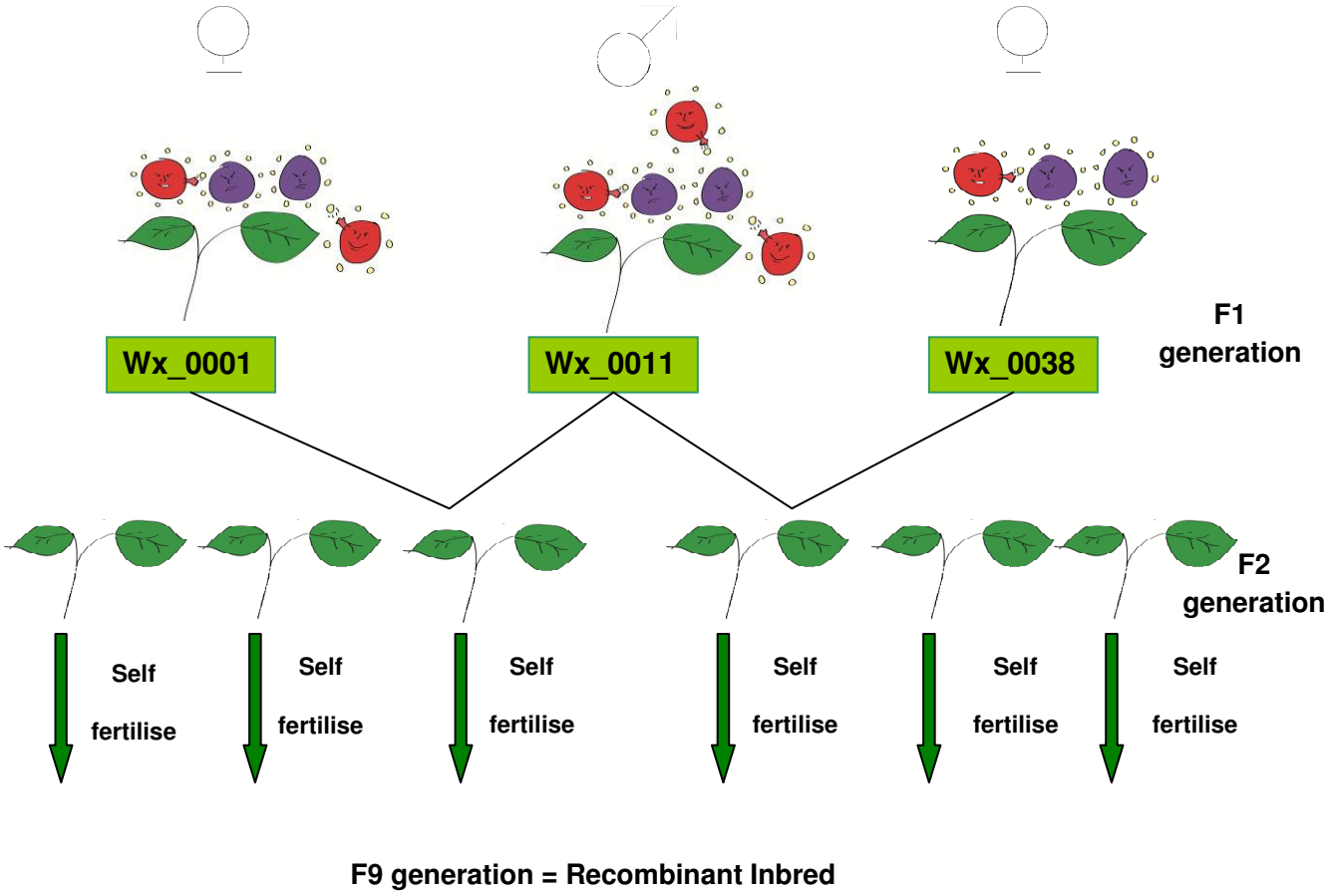


Figure 4.4: Breeding programme incorporating Wx_0001 (standard line), Wx_0011 (high antioxidant line) and Wx_0038 (low antioxidant line).

N.B Wx_0001 was crossed with Wx_0011 in order to produce a line with a high level of antioxidants whilst Wx_0011 and Wx_0038 were crossed to produce a mapping population for QTL elucidation.

One line in which further breeding may not be required is Wx_0033, which not only has an unusual dwarf phenotype (Figure 4.5) but also a high concentration of antioxidants. However it may also be used to be crossed into the standard commercial line, Wx_0001.

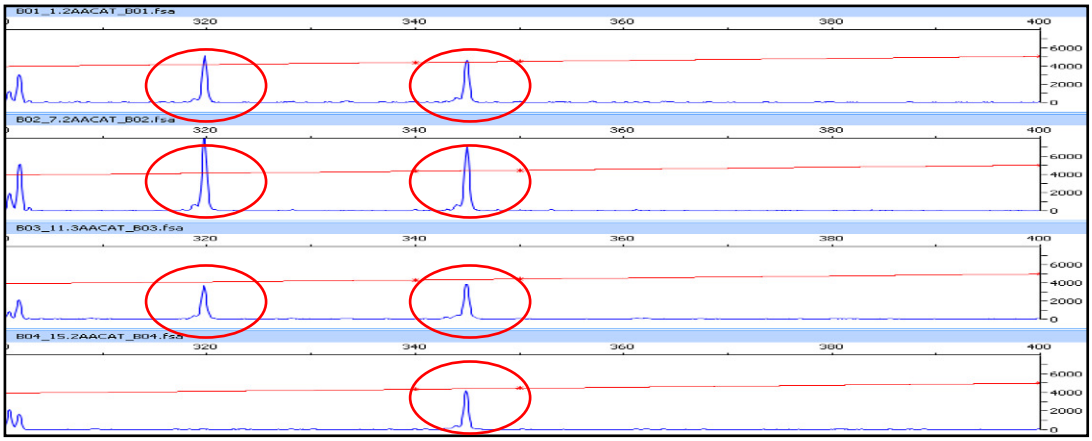
Figure 4.5: An example of the dwarf phenotype of Wx_0033 in comparison to Wx_0001.

4.4.2 AFLP output

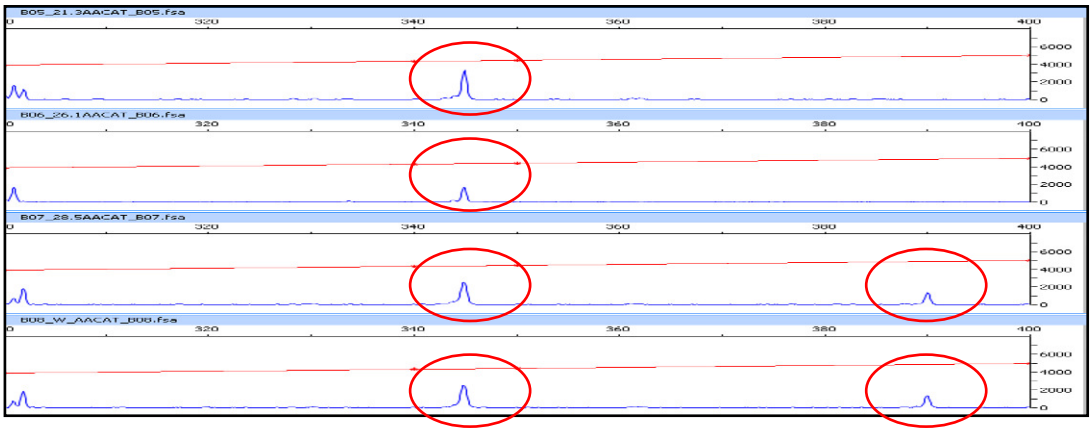
The data output received from ABI Nottingham was initially screened through ABI PeakScanner v1.0. Peaks were recorded for each of the samples sent for analysis. Evidence of variation with difference in peaks produced between the different watercress lines was evident whilst scoring peaks in ABI peak scanner. Some samples e.g. Wx_0009 had an additional peak when compared to Wx_0011 using the FAM primer pair AACAT (Figure 4.6). Interestingly there is evidence that an extra peak is produced when comparing Wx_0013 and Wx_0011, with Wx_0013 having the additional peak using the FAM primer pair set AACAT (Figure 4.6). These differences are also evident using HEX primer pairs, with again Wx_0013 having an additional peak when compared to Wx_0011 (Figure 4.6). The peaks were transferred into excel and then the differences statistically analysed using GenAlEx.

Figure 4.6 a and b cover the same section and clearly demonstrate a peak which is present in Wx_0009 is not present in Wx_0011 or Wx_0013. It is these differences which GenAlEx uses for the genetic diversity calculations.

a)



b)



c)

Figure 4.6: Peak scanner output a) AACAT (FAM) with samples of lines Wx_0009 (top three boxes) and Wx_0011 (last box) b) AACAT (FAM) with samples of lines Wx_0011 (first two boxes) and Wx_0013 (second two boxes) c) AACTT (HEX) with samples of lines Wx_0011 (first two boxes) and Wx_0013 (second two boxes)

The lines which were taken forward for antioxidant analysis in earlier Chapters were analysed via GenALex, Wx_0001, Wx_0011, Wx_0033 and Wx_0038 (Table 7). However contrary to biochemical and morphological results it still appears that there was a higher level of variation within each of the lines compared to between each of the different lines, apart from primer pair CTCAG which again revealed a greater proportion of variation between the watercress lines compared to within each line.

Table 7: Genetic diversity output for both selected lines and all lines held in the germplasm collection

Primer Pair	Selected lines within/among populations (%)	All lines within/among populations (%)
AACAT	79/21	72/28
AACTT	100/0	78/22
CTCAG	34/66	34/66
TACAT	80/20	72/28
TCCCT	88/12	100/0

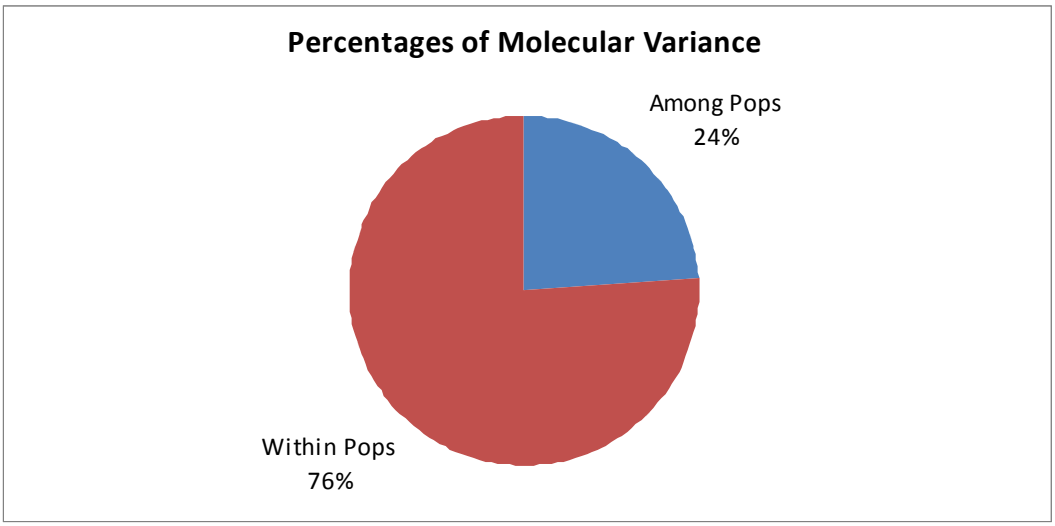
All primer pairs used in this study were combined in one final analysis to identify any potential genetic diversity held within the collection. This allowed an overall Principle Coordinate Analysis and dendogram to be produced (Figure 4.7a and b). It was concluded that the highest level of variation was present within each line rather than between the watercress lines. 76% of the genetic variation lies within the lines whilst only 24% among the lines clearly demonstrating a low genetic diversity between the lines (Figure 4.7a).

Principle Coordinate Analysis is a variant of Principle Component Analysis which is a multivariate test which weights the variables to maximise the differences between individuals (Dytham, 2003). The weights of all the available variants are taken into consideration to provide the maximum discrimination between individuals (Dytham, 2003). The populations are clearly scattered throughout the graph with no same line

represented by population within the same area hence reinforcing the result that most of the genetic variation is within each line rather than between.

All the fragments produced by the five different primer sets were combined in order to also produce an out tree which would allow visualisation of the relatedness (Figure 4.7c). Overall the tree reinforces the findings from individual analysis indicating there is more variation within each line than between. The biological replicates of each line are dispersed throughout the tree and tend not to cluster together.

a)



b)

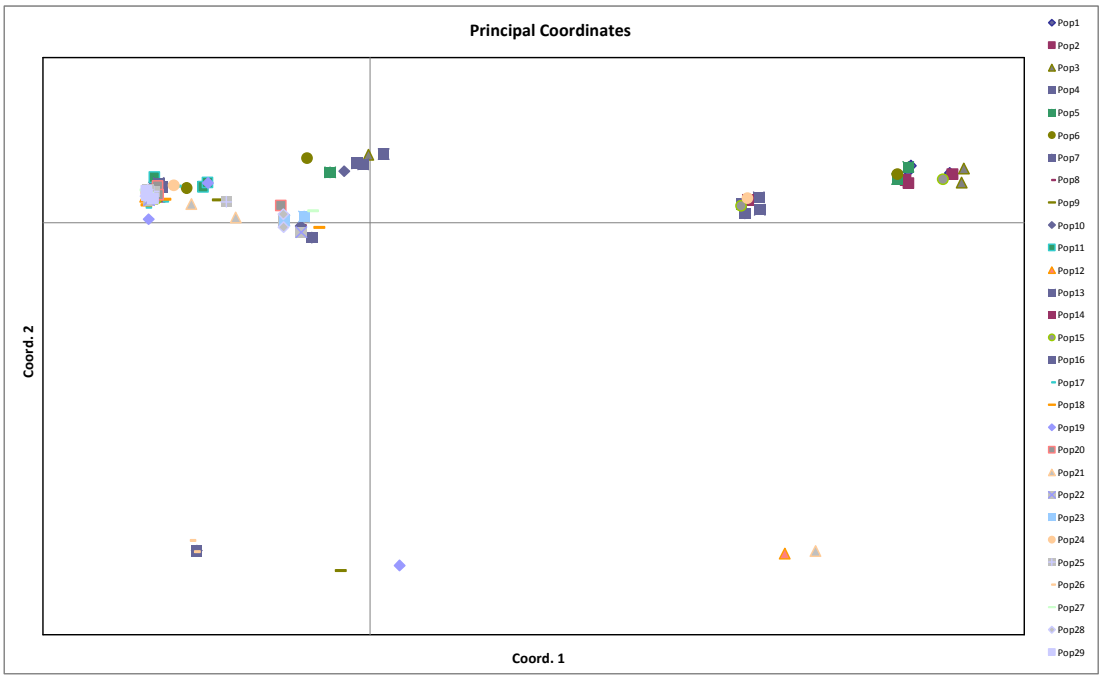


Figure 4.7: (a) percentage of molecular variance apportioned between and within watercress lines, (b) Principle Coordinate Analysis of the watercress lines and (c) Watercress lines displayed in out tree

c)

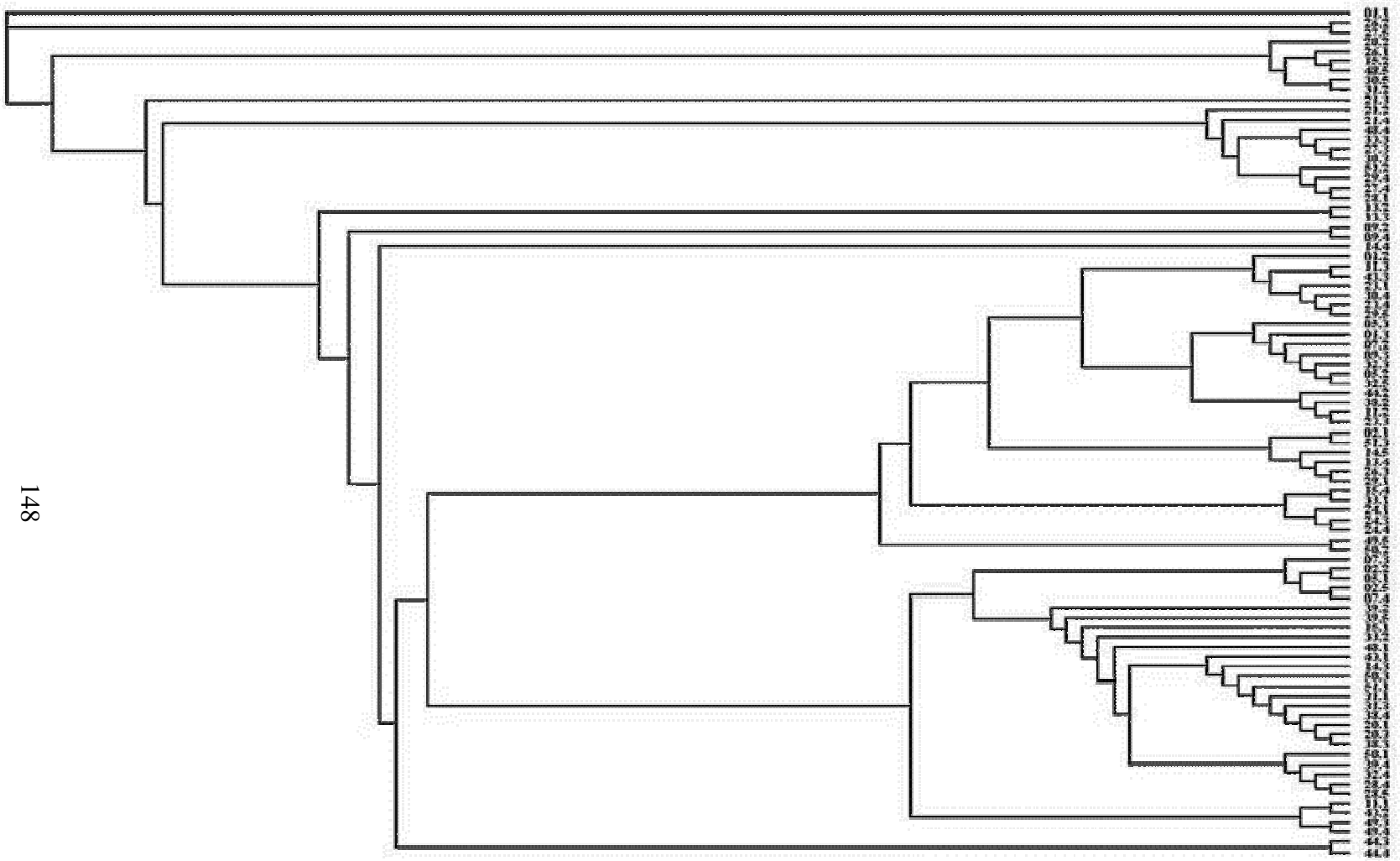


Figure 4.7: (c) Watercress lines displayed in out tree

4.5 Discussion

In order to determine the degree of genetic diversity within the watercress germplasm collection it was necessary to produce a genetic fingerprint for each line. A genetic fingerprint is a unique pattern which can be visualised after PCR amplification. AFLP is a technique for fingerprinting DNAs of any origin or complexity (Blears *et al.*, 1998). In plant molecular genetics the AFLP technique can be applied to assess phylogeny and diversity. AFLP patterns can also provide information on the geographic origins of genotypes and gene pools of plants (Savelkoul *et al.*, 1999). The AFLP technology has aided marker assisted breeding, variety identification (AFLP allows identification of a contaminating variety), germplasm management (by providing AFLP profiles), indirect selection of agronomically important traits and finally backcross breeding (Savelkoul *et al.*, 1999). The focus of this research used AFLP for germplasm management i.e. creating profiles of breeding lines which would then allow predictions of agronomic performance based on relative genetic distance (Savelkoul *et al.*, 1999). Previously RAPDs have been used to assess the genetic diversity but this study progresses the assessment of the gentic diversity of watercress using AFLPs.

Here at the University of Southampton the only currently known watercress breeding programme in the United Kingdom has been established. Final results are not presented as the breeding programme is still being progressed through the generations until the phenotype is fixed in the ninth generation. For many years plant breeders have attempted to enhance yields by selecting individual traits (Rasmusson, 1987). The aim of a large proportion of plant breeding is to improve yields and therefore productivity but another key objective is to develop food staple crops which have seeds which are fortified with essential minerals and vitamins (Bouis, 2007), which would aid developing countries. When varieties are crossed for traits of interest certain characteristics such as taste, cooking qualities and appearance must be maintained. One issue faced with plant breeding for enhanced nutrition is that it is believed that nutrient-enhanced crops will be lower yielding and the additional breeding objective with higher nutrient content will thereby slow down the

production of varieties which have higher yields (Bouis, 2007). A lower dependence on fungicide can be achieved by an enhanced microefficiency due to a higher resistance to root disease (Bouis, 2007). One interesting aspect would have been to have examined the microbial activity on watercress, would the level of microbial activity be lower on lines which had a higher level of PEITC and FRAP values? If this was confirmed then breeding watercress with a higher level of PEITC would not only provide nutritional benefits but also environmental benefits i.e. a reduction in the levels of pesticides and fungicides used.

One of the main objectives to modern plant breeding is to restore some of the diversity which was lost during the domestication for improved agricultural yields (Gur and Zamir, 2004). It is very important to utilise the advantageous traits from wild relatives in agricultural crops (Gur and Zamir, 2004). Thereby by examining the phenotypic and biochemical variation (Chapters two and three) an assessment of the current variation in watercress was assessed. The goals are changing in plant breeding from selection of phenotypes to selection of genes either directly or indirectly (Koornneef and Stam, 2001). Plant breeders are trying to optimise the use of genetic variation in nature by bringing together the alleles in one genotype which maximize yield, resistance to stress etc. (Koornneef and Stam, 2001). However it is of great importance to know how genes function together as they do not function as single entities (Koornneef and Stam, 2001). It is also important to assess the genetic variation within and outside the species to identify the genes which will correct the limitations which limit the optimal functioning of plants (Koornneef and Stam, 2001).

Molecular marker technologies are becoming very important tools for genomic and genetic studies, breeding and biodiversity research (Yang *et al.*, 2006). Another example of a crop which, like watercress, has not been extensively bred is Pigeonpea, a legume crop which is a protein rich food grown predominately in Asia, Africa, Latin America and Australia. The genetic progress on this crop is slow due to small investment in breeding and research towards this crop. The genetic improvement in this crop is also impeded due to limited breeding activity. A quick

and efficient technology for genotyping and precise germplasm diversity assessment is therefore of great importance for the development of this crop (Yang *et al.*, 2006). Hence with the watercress breeding programme established to date we have combined the assessment of morphological and biochemical traits (Chapters two and three) alongside the assessment of genetic diversity (Chapter four). This has provided essential data for progressing breeding in this popular crop.

Plant breeding will remain a vital component of an effective agricultural system and has contributed to the increased crop productivity by creating new genotypes (i.e. varieties) with superior adaptations to the needs of society as well as the resources of the production systems and the demands of nature in the target environments (Lee, 1998). However it needs to be taken into consideration that with all these many benefits there is evidence of a decreased genetic diversity within the elite gene pool and an increased genetic uniformity in crop production (Lee, 1998). This is however not apparent within the watercress collection; this chapter has revealed a high level of genetic diversity which is apportioned to within line variation and explains the standard error bars which are evident on the traits measure in Chapters two and three indicating a large degree of heterogeneity.

Therefore by creating a watercress germplasm collection a repository of genetic diversity is maintained in which certain traits can be selected then bred. The breeding programme currently established is the first step towards creating a mapping population and achieving the desired phenotype (short stem length) and desired concentration of antioxidants.

From the data presented the level of genetic diversity in the watercress germplasm collection, despite the evident morphological and biochemical variation, currently appears to be low between the watercress lines but a large amount of variation is apportioned to within line variation. By establishing a germplasm collection of watercress lines sourced from around the world it is hypothesised that there would be some potential to have isolated a large amount of genetic diversity which could then

be used in a breeding programme. Morphological and biochemical measurements revealed some promising variation between the different watercress lines within the collection which enabled lines to be selected for a breeding programme with the ultimate goal of a crop with an enhanced level of antioxidants. The AFLPs have supported the error bars which are present on the morphological and biochemical traits recorded (Chapters two and three) due to a high level of variation being apportioned to within line and not between.. However this may aid in the breeding programme for the watercress germplasm collection as there is still variation within each line which can then be taken advantage of in the breeding programme. It is apparent from the high within line variation that the watercress population has not yet reached a genetic bottleneck. Li *et al*. 2004 stated that genetic bottlenecks are 'stochastic events that limit genetic variation in a population and result in founding populations that can lead to genetic drift'. For many years crops have been cultivated for a desired property by breeders. However, domestication has resulted in allelic variations of genes gradually being lost and only being recovered by collecting wild accessions (Tanksley and McCouch. 1997). It appears that our data from AFLP analysis of the watercress collection revealed that the watercress present within the collection has not yet reached a genetic bottleneck like most other commercial crops and therefore still has a high level of allelic variations of genes. We can therefore move from selection of phenotype to a closer examination of phenotype i.e. switch from selecting parents by phenotype to screening for the presence of potentially beneficial genes (Tanksley and McCouch. 1997). This is very promising for moving the watercress breeding programme forward as it appears there may still be ample of genetic variation available within this collection. The program used to assess the AFLP output was GenALex 6.4 which allows population genetic data analysis of codominant, haploid and binary genetic data and accommodates the full range of genetic markers available such as allozymes, microsatellites, single nucleotide polymorphisms (SNPs), Amplified Fragment Polymorphisms (AFLPs) and other multilocus markers as well as DNA sequencers (Peakall and Smouse, 2006).

When Principle Coordinate Analysis (PCA) was carried out on the watercress lines held within the current germplasm collection the clustering is distinctive whereby replicates of the each watercress line are clearly spread out and not clustered

together. This demonstrated low genetic diversity between the lines held in the collection and a large variation in genetic diversity within each line. One primer pair did in fact show a higher level of diversity between the lines compared to within the same line suggesting that further primer pairs may reveal further differences in the collection.

The AFLP procedure allowed genetic polymorphisms to be identified by the presence or absence of DNA following the restriction and amplification of the genomic DNA of interest (Blears *et al.*, 1998). The number of these amplified fragments detected on the gel is determined by the complexity of the genomic DNA, the choice of enzymes (in this case EcoRI and MseI) and the number and type of selective nucleotides in the PCR primers (Blears *et al.*, 1998). The genomic DNA is required to be of high purity in order to ensure complete digestion by the restriction endonucleases. This then provides templates for the preamplification utilising primers having single or no selective nucleotides which are then diluted and used in the second amplification process (Blears *et al.*, 1998). RAPDs use arbitrary primers whilst AFLP primers are specific to the adaptor and restriction site sequence. The RAPD technique may be easier to perform than AFLP but is very sensitive to reaction conditions, template DNA concentration and purity whilst AFLP utilises stringent annealing conditions thereby better reproducibility (Blears *et al.*, 1998). The watercress DNA was put through the RAPDs protocol and demonstrated the lack of reproducibility of RAPDs hence a shift to the AFLP procedure. Due to the AFLP technique being able to detect polymorphisms in many loci of the genome simultaneously it is possible to study DNA variation within the entire genome of *A. thaliana* (Miyashita *et al.*, 1999). The AFLP technique will allow fingerprints to be produced and monomorphic bands (generally conserved between individuals of the same type) and polymorphic bands (specific to individuals) to be visualised (Sheridan *et al.*, 2001).

One important consideration to take into account with the AFLP protocol is the choice of fluorescent dyes used. It is evident however that the clarity of results differs depending on the dye used. Pandey *et al.* (2007) deduced that the variation in

T-RFLP (Terminal Restriction Fragment Length Polymorphism) profiles could be attributed to the use of different dyes for labelling primers. The fluorescent dyes, FAM and HEX, are not dissimilar in physical characteristics e.g. absorbance maxima, emission maxima and chemical structure but they are very different in their molecular weight (Pandey *et al.*, 2007). HEX is a hexa-chloro derivative of FAM and has six additional chloride atoms making it a heavier tag, therefore a heavier tag compared to FAM (Pandey *et al.*, 2007). Fewer but clearer peaks were produced with the HEX labelled primers.

Earlier work had reported the limited work on watercress breeding. Claxton *et al.* (1998) highlighted the limited breeding programmes present in the watercress industry hence a lack of cultivars which reinforces the need to evaluate alternative techniques which will allow the generation of variation within the crop. The ultimate goal with the work carried out by Claxton *et al.* (1998) was to breed cultivars with a complete or partial resistance to crook root disease. One possible way to enhance the resistance of watercress to *Spongospora subterranean* f. Sp. *nasturtii* is by looking at the somaclonal variation present in the watercress. Somaclonal variation is the genetic and phenotypic variation among clonally propagated plants of a single donor plant (Kaeppler *et al.*, 2000). It was concluded that somaclonal variation in watercress could be a potential source of heritable genetic variation which could aid in the improvement of the crop (Claxton *et al.*, 1998). Earlier in 1995 investigations were carried out on somaclonal variation, watercress was screened for resistance to crook root disease and revealed significant variation between controls and some of the somaclones (Arnold *et al.*, 1995).

Sheridan *et al*. (2001) used Random Amplified Polymorphic DNA Polymerase Chain Reaction (RAPD-PCR) to analyse the genetic variation in a watercress germplasm collection. In the study 12 watercress populations were used and sweet potato was brought into the study to represent a distant relative whilst Chinese cabbage as a close relative (Sheridan *et al.*, 2001). The Principle Co-ordinate Analysis (PCA) used in the work carried out by Sheridan *et al*. (2001) showed very little variation, similar to the results presented in this Chapter, between the

watercress populations, the Japanese watercress being placed close to the commercial watercresses whilst Chinese cabbage and sweet potato being very different (Sheridan *et al.*, 2001). The Simple-matching and Jaccard distances showed similar results for the watercress populations. The watercress populations ranged from 0.06 and 0.35 units whilst Chinese cabbage was placed at 0.4 and 0.81 distant units and sweet potato between 0.6 and 0.81. Finally when put through UPGMA and neighbour-joining clustering programs the commercial wateresses are placed close to each other (Sheridan *et al.*, 2001). Overall the data presents very little diversity despite the presence of a watercress line from Japan being held within the collection. Concerns have arisen with the lack of genetic diversity within germplasm collection and the option that transgenics may need to be employed in order to produce an improved cultivar. Indeed Grusak (2002) highlighted that conventional breeding can achieve the goal of improved phytochemical concentration in plants. However the genetic diversity available within existing germplasm collections can limit the improvement that can be made through traditional techniques alone (Grusak, 2002). In order to increase the genetic diversity further, collection of watercress from extreme environments may contribute unique genes to the collection. Wild ancestors and primitive crop varieties may tend to produce poorer yields and have poor eating qualities but they are the sources of genetic variation upon which modern varieties are constructed (McCouch, 2004).

However RAPD-PCR carried out on intraspecies revealed many polymorphic bands and allowed the distinct formation of three groups: watercresses and *Barbarea verna*, non-watercress *Rorippa* species and *Cardamine*, *Brassica* and *Lepidium* species (Sheridan *et al.*, 2001). The study proposed the option for hybridisation of *Barbarea verna* into the breeding programme as it was placed closer to watercress than *Cardamine* or the non-watercress *Rorippa* species (Sheridan *et al.*, 2001).

This Chapter has allowed the initial insight into any possible genetic diversity within the watercress germplasm collection. Further AFLP investigation with more primer sets may reveal polymorphisms between the watercress lines; one primer pair has already shown a higher level of diversity between the watercress lines. However the

large variation apportioned within line provides exciting possibilities in this new breeding programme. The watercress breeding programme has progressed using data presented in earlier Chapters and the extremes for each trait of interest (i.e. higher antioxidants and shorter stem length). The crossings are being progressed through single seed descent. The single seed descent method permits the plant breeder to advance generations whilst maintaining the genetic diversity (Johnson and Bernard, 1962). Single seed descent avoids selection in the early generations and waits until a high degree of both homogeneity and homozygosity has been reached. This allows the genotypic differences between the lines to be more pronounced and stable (Van Oeveren and Stam, 1992). Less time is required to obtain homozygous lines as the process of genotype evaluation and selection only begins after reaching the lines in homozygosis (Riva-Souza *et al.*, 2009). The method of single seed descent consists of selfing all plants of the present generation and using a single seed from each plant in order to obtain the next generation (Jansen and Jansen, 1990). Knott and Kumar (1975) also used the single seed descent procedure in wheat breeding whereby a single seed was taken from each F2 plant to produce the F3 generation, the process was repeated for the F4. The F5 lines were then produced from the seed of the F4 plants (Knott and Kumar, 1975). A few more generations are needed to progress through before the result of the crossing are beneficial i.e. a higher level of antioxidants. This is due to an increased likelihood of homozygosity in the plant with progressive generations and therefore a higher stability of the expressed trait or phenotype of interest. Consequently if the cross was successful by the eighth generation the end product would an enhanced antioxidant watercress crop.

4.5.1 Conclusion

1. Watercress breeding programme established, currently in third generation. Seeds are stored ready for further progression into the next generation.

2. There is more genetic variation apportioned within each line than between the watercress lines therefore it is evident that a genetic bottleneck has not yet been reached in this crop which provides a positive scope to breeding. This

broad genetic base is evident in Chapters two and three with the large standard error bars present on the traits observed.

CHAPTER FIVE: EXAMINING THE POTENTIAL VARIATION IN GENE EXPRESSION BETWEEN THREE SELECTED WATERCRESS LINES; Wx_0001, Wx_0011 and Wx_0033

5.1 Abstract

Variation in phenotype may be attributed to differences in gene expression. Therefore examining gene expression may reveal underlying genetic differences between the watercress lines held within the collection. Gene expression can be visualised using microarrays which look at whole genome profiles and validated with real time PCR. Of particular interest were genes relating to the control of plant height and antioxidant concentration. Differences in gene expression of two antioxidant genes, At1g65060 which encodes an isoform of 4-coumarate:CoA ligase (4CL) which is involved in the last step of the general phenylpropanoid pathway and At4g35000 which encodes a microsomal ascorbate peroxidise APX3 which scavenge hydrogen peroxide in plant cells, were recorded for watercress lines Wx_0001, Wx_0011 and Wx_0033. Real time PCR confirmed the microarray gene expression trend and an overall upregulation of the gene encoding 4CL (At1g65060) was recorded for Wx_0001 in comparison to Wx_0011 and Wx_0033. We conclude that there is variation in the expression of genes encoding an ascorbate peroxidase and one encoding 4CL and finally real time PCR can be used to validate microarray results and is a useful tool for quantifying the expression levels of genes of interest.

5.2 Introduction

Watercress has not yet been fully DNA sequenced and therefore few of the genomic resources available in other plant species are available to use here. Also, in contrast to the much researched *Arabidopsis*, the transcriptomes of most agriculturally important plant species are less extensively studied as extensive sequence information and arrays are required to aid research (Hammond *et al.*, 2005a). Biological systems operate in a similar way and it is hypothesised that the genes which participate in these systems are conserved across organisms (Lu *et al.*, 2009). Both watercress and *Arabidopsis* are in the Brassicaceae family thereby knowledge about the model plant *Arabidopsis* could be used to gain a basic understanding of watercress.

Chismar *et al*. (2002) reported the limitation in microarrays for a variety of organisms i.e. the lack of availability and one possible solution is to use microarrays designed for one species to analyse RNA of a closely related species (Chismar *et al.*, 2002). Assuming that there is a conservation of gene sequences between species, a large amount of data can be extrapolated (Chismar *et al.*, 2002). Many papers have reported using arrays designed for a model species to investigate the transcriptome of their interest e.g. Hammond *et al.*, (2005a) investigated *Brassica oleracea* on the ATH1 genechip whereby 98.8% of the available probe sets were retained, *Thlaspi caerulescens* and *Thlaspi arvense* were hybridised to the *Arabidopsis* ATH1 genechip (Hammond *et al.*, 2006), banana (*Musca* species) was hybridised to both the ATH1 and rice genechip (Davey *et al.*, 2009) and finally *Cardamine kokaiensis* hybridised to the ATH1 genechip (Morinaga *et al.*, 2008).

Probe level effects are caused by the insufficient hybridisation of certain transcripts from the target species to their appropriate GeneChip probes designed for the model species. These probe-level effects will substantially impact on the probe-set expression estimates (Hammond *et al.*, 2005). The weight carried by each probe set is inversely related to its distance from the median value of all probes within the

probe set. Therefore probes not generating signals because of sequence polymorphisms with the transcript from the target organism can reduce the quality of information available (Hammond *et al.*, 2005). Some researchers have adopted an RNA-based probe-selection system to counter this problem but a genomic DNA (gDNA)-based probe selection strategy can be used to optimise cross-species transcriptomics (Hammond *et al.*, 2005). The gDNA from the species of interest is labelled and hybridised to the GeneChip array and the perfect-match probes which hybridise efficiently to the gDNA are selected for subsequent use in interpreting GeneChip arrays challenged with RNA from the species of interest (Hammond *et al.*, 2005). This was the approach that was taken here.

Pearl script generates custom probe mask files and then is used to select probe pairs in which the (Perfect Match) PM probe has a gDNA hybridisation greater than a user-defined threshold (Hammond *et al.*, 2005). This script then identifies probe sets retaining one or more PM probes with a gDNA hybridisation intensity greater than the threshold (Hammond *et al.*, 2005). The Chip Description File (CDF) provides the template for the generation of a signal for the probe set when the watercress transcriptome (RNA Channel Expression Level (CEL) files) are analysed (Hammond *et al.*, 2005). The custom CDF files allow information to be extracted from the RNA CEL files for only those probe pairs whose PM probe has a gDNA hybridisation intensity above the imposed threshold (Hammond *et al.*, 2005).

Therefore a new cross species GeneChip has been designed and can be used for species in which they were not designed, in this case we hybridised watercress on an *Arabidopsis* chip to deduce suitability for gene expression analysis. A species close to the target species, i.e. *Arabidopsis* was chosen thereby minimising sequence mis-match effect (Lu *et al.*, 2009).

The *Arabidopsis* ATH-1 chip which contains 22,810 *Arabidopsis* probe sets allows the evaluation of almost all the genes in the *Arabidopsis* genome (Misson *et al.* (2005), has been used to examine the transcriptome of *Brassicacea* species in

(Davey *et al.*, 2009). This high-density oligonucleotide microarray allows the transcriptomics of non-model plants to be probed for example *Arabidopsis halteri*, *Thlaspi caerulescens*, *Thelungiella halophila*, *Brassica oleracea* and *Brassica napus* (Davey *et al.*, 2009).

Arabidopsis ATH-1 GeneChip is the best characterised containing over 22,500 probe sets representing around 24,000 genes. The array is based on information from the international *Arabidopsis s*equencing project completed in December 2000 (Davey *et al.*, 2009). The National Arabidopsis Stock Centre (NASC) provides an array service with a choice of arrays and new GeneChips being added, including a Brassica Exon 1.0ST GeneChip array. The Affymetrix array design consists of multiple probes for each gene and also the presence of mismatch (MM) and perfect match (PM) probes (Bagnaresi *et al.*, 2008; Moore *et al.*, 2005) reported the importance of cross-species genome comparisons for gene expression as well as a tool for analysis of evolution and crop domestication.

The coding regions in *Arabidopsis thaliana* and across the Brassicaceae are highly conserved with Cavell *et al.*, (1998) reporting an 85% sequence similarity (Muangprom and Osborn, 2004). Also on a Brassica Express Sequence Tag (EST) microarray made available from John Innes centre, 77% of the protein hits came from *A. thaliana*. This gave us confidence to proceed with using *Arabidopsis* to progress with watercress knowledge.

Once the microarray data is processed the GeneChip array expression data is verified using real time PCR (qRT–PCR). Valasek and Repa (2005) reported that real time PCR has been accepted as a robust and widely used method for biological investigation as it allows very small amounts of specific nucleic acid sequences to be detected and quantified. Real time PCR is a very important research tool with respect to acting as a rapid and accurate assessment of gene changes as a result of physiology, pathophysiology or development (Valasek and Repa, 2005). The DyNAmo Flash SYBR Green qPCR kit used in real time PCR is based on a hot start

version of a modified *Thermus brockianus* DNA polymerase and the SYBR Green I fluorescent dye (Finnzymes instruction manual). Specific nucleic acid sequences in a sample are distinguished and measured in real time PCR (Valasek and Repa, 2005). Quantification is enabled by recording how quickly the fluorescent signal reaches a threshold level which correlates with the amount of original target sequence (Valasek and Repa, 2005). The complementary DNA (cDNA) from each sample can be used as raw material for real time PCR utilising its precision and sensitivity to determine changes in gene expression (i.e. RNA levels) (Valasek and Repa, 2005). It is the precision and sensitivity of real time PCR which allow subtle changes in the gene expression to be detected (Valasek and Repa, 2005).

The key strategy is to link a change in fluorescence to amplification of DNA thus SYBR green I binds to the minor groove of double stranded DNA (dsDNA) emitting 1000 fold greater fluorescence compared to when in free solution. The greater the amount of dsDNA present then the greater the amount of DNA binding and fluorescent signal from SYBR green I (Valasek and Repa, 2005) therefore, allowing any differences between the lines to be quantified.

The aim of this Chapter was to hybridise watercress RNA to the *Arabidopsis* ATH1 genechip which would allow visual inspection of any variation in gene expression between the different watercress lines. However due to watercress being relatively understudied crop preliminary investigations need to be carried out before deciding which platform was used. Using the X species platform at the National Arabidopsis Stock Centre (NASC) Nottingham, it was confirmed that the hybridisation of watercress DNA was efficient with *Arabidopsis* and therefore allowed progression to the hybridisation of watercress RNA to the *Arabidopsis* ATH1 genechip. The initial X species hybridisation allowed optimisation of the cross-species transcriptomics (Hammond *et al.* 2006). Microarray analysis allowed the deduction of genes of interest i.e. genes which may contribute to antioxidant defence or dwarfism. The gene expression from microarray analysis was then validated using real time PCR. The watercress lines chosen in this part of the study for the comparison in gene expression were lines of particular interest. Wx_0001 was the standard Vitacress line

whilst Wx_0033 demonstrated a dwarf phenotype and a high antioxidant power and Wx_0011 demonstrated a high antioxidant power. The unusual phenotype presented by Wx_0033 in comparison to Wx_0001 and Wx_0011 may be explained by closer inspection of gene expression. Therefore two genes in relation to dwarfism were chosen; one which encoded a gibberellin 2-oxidase and the other encoded a protein with gibberellin 3-beta-hydroxylase activity. A higher level of expression in Wx_0033 compared to Wx_0001 and Wx_0011 is expected for the gene encoding gibberellin 2-oxidase due to its function in causing gibberellin deactivation thereby resulting in dwarfism and a lower level of expression of the gene encoding a protein with giberrellin 3-beta-hydroxylase due to its function in the final steps in gibberellin synthesis.

The levels of gene expression of the genes encoding ascorbate peroxidise and the final stages in the phenylpropanoid pathway are expected to be relatively high in each of the lines. This Chapter will explore how variation in gene expression may contribute to differences between these watercress lines.

5.2.1 Aims

1. Examine the global gene expression of watercress using the *Arabidopsis* ATH1 gene chip.

2. Confirm microarray results using Real Time PCR with genes of interest.

3. Identify potential genes or pathways for furture manipulation with regards to breeding.

5.3 Materials and Methods

5.3.1 Lines selected for analysis

From the many watercress lines contained within the collection three particular lines were selected for variation in gene expression;

Wx_0001 = standard Vitacress line

Wx_0011 = high antioxidant power

Wx_0033 = dwarf phenotype and high antioxidant power

Whole plants (excluding roots) were harvested after the morphological measurements were recorded in Chapter two (2.3.4) and immediately snap frozen in liquid nitrogen. Three biological replicates for each line were used.

5.3.2 DNA extraction

CTAB extraction: The leaf material was ground in liquid nitrogen using a mortar and pestle and a small amount (~200mg) was transferred into an eppendorf. 900µl of CTAB which was prewarmed to 65°C was added to each tube, inverted and incubated at 65ºC for 45 minutes. The eppendorfs were inverted half way through the incubation period. Into each eppendorf 900 µl chloroform/isoamyl alcohol (CHISAM) was added and the tubes were shaken until well mixed and centrifuged at maximum speed (13,200rpm) for 10 minutes in Eppendorf centrifuge 5417R. 500µl of the aqueous phase was transferred to a fresh eppendorf and 50µl 3M sodium acetate and 333µl cold isopropanol added to the transferred aqueous phase. This is mixed well and incubated at -20°C for 15 minutes and then centrifuged for 10 minutes at maximum speed. The liquid phase was poured out and the pellet (DNA) kept. The tube was turned upside down and gently blotted on tissue. This was then centrifuged for 15 seconds and any remaining liquid pipetted out. The pellet was then washed with 500µl cold 70% ethanol and then centrifuged for 10 minutes at maximum speed. The liquid phase was then pipetted out and centrifuged for an additional 15 seconds to remove any remaining liquid. The pellet was then left to dry

at room temperature with the lids open for 20 minutes. 50µl TE solution and 1µl RNase was added to each tube and then left incubating at room temperature overnight. The next day 200µl TE solution, 100µl 3M sodium acetate and 1ml cold absolute ethanol was added to each eppendorf and mixed well and then was centrifuged at maximum speed for 10 minutes. The liquid phase was discarded by pouring out and the pellet could be seen as a small smear on the wall of the tube. Again this was centrifuged for 15 seconds and any remaining liquid pipetted away. The eppendorfs were left to dry with the lids open for 20 minutes at room temperature and then the pellet was resuspended in 50µl water. The nanodrop (1000 spectrophotometer, Thermo Scientific) gave an estimate of the concentration of DNA.

5.3.3 Hybridisation to Xspecies chip

(protocol provided by NASC)

DNA from Wx_0001 was sent to the National Arabidopsis Stock Centre (NASC) for hybridisation to the Xspecies ATH1 genome array. At NASC random primers (octamers) were annealed to the denatured DNA template and were then extended by Klenow fragment in the presence of biotin-14-dCTP. This produced sensitive biotinylated-DNA probes for use in the nonradioactive detection of DNA and RNA. The use of BioPrime DNA labelling system can result in considerable net DNA synthesis to occur and a 10-40 fold amplification of the probe. NASC then used a BioPrime Labelling protocol which is modified from Invitrogen. 500ng DNA was dissolved and diluted in 7µl dilute buffer in a microcentrifuge tube. Then on ice 20µl of 2.5X Random Primer Solution was added. This was then denatured for 5 min at 100°C and was immediately cooled on ice. The following was added whilst on ice; 5µl 10X dNTP mixture, 17µl distilled water (to a total volume of 49µl), this when then mixed briefly and 1µl Klenow fragment added and mixed thoroughly but gently. After these additions the mix was centrifuged (13,200rpm) for 15-30 seconds and incubated at 37°C for 60 minute. Finally 5µl of stop buffer was added. The yield was then tested using 1µlin the nanodrop spectrophotometer. Step 2 was the production of the hybridisation cocktail whereby the following was pipetted into an eppendorf; 3µl BSA (50mg/mL), 2.7µl Herring sperm, 5µl B2 control, 15µl 20x

Euk Control, 150µl 2x Hyb buffer and 70.3µl DEPC water. 54µl of the sample solution was added to give a total volume of 300µl.

In the final step the chip was wet with 100µl water and 100µl 2x buffer and placed in a 45°C oven for 10 minutes at 60rpm. The specimen was incubated at 99°C for 5 minutes and then 45°C for 5 minutes centrifuged at 1500g for 5 minutes and the wetting solution removed. The chip was then filled with 200µl of specimen and hybridised overnight at 45°C for 16 hours at 60rpm.

5.3.4 Xspecies chip analysis – creation of probe mask (.cdf) files

Oligonucleotide GeneChip probes with low/no hybridisation to watercress transcripts were removed using the gDNA probe masking strategy devised by Hammond *et al.*, (2005a) and subsequently used by Davey *et al.*, (2009). Probe pairs from the gDNA CEL files were selected for subsequent watercress transcriptome analysis using a parser written in Perl (www.perl.com) to generate custom probe mask files. The Perl script selects probe pairs in which the PM probe has a gDNA hybridisation intensity greater than a user-defined threshold. This script then identifies probe sets retaining one or more PM probes with a gDNA hybridisation intensity greater than the imposed threshold. Selected probe sets are collated in a custom Chip Description File (CDF). This CDF file provides the template for the generation of a signal for the probe set when analysing the watercress transcriptome. The custom CDF file allows information to be extracted from the RNA CEL files for only those probe pairs whose PM probe has a gDNA hybridisation intensity above the imposed threshold. Both the PM and MM were retained in the CDF file (Hammond *et al.*, 2005a). The Pearl script enables user-specific gDNA hybridisation intensity thresholds for probe mask file generation to be set. Optimal gDNA hybridisation intensity threshold for interpreting transcriptomic data could be determined systematically and empirically (Hammond *et al.*, 2005a). Forty three probe mask files were generated using gDNA hybridisation intensity thresholds from 30 to 16000 and the ATH1 GeneChip array CDF file represents a universal CDF file with no probe pairs excluded (Hammond *et al.*, 2005a). The probe masks were

generated from the ATH1 GeneChip array CDF file which is universal for both *Arabidopsis* and watercress and from which no probe pairs were excluded. The probe masks excluded the maximum number of MM probes whilst retaining the highest PM probes ~90% which represents the biggest difference on the composed graph. The custom CDF file was then loaded onto the GeneSpring computer for analysis of the RNA files.

5.3.5 RNA extraction

The finely ground frozen plant sample was transferred into a Zymo Research (ZR) bashing bead lysis tube and 800µl RNA lysis buffer was added to the sample. The ZR bashingbead lysis tube was then centrifuged at 12000xg for 1 minute. 400µl of the supernatant was transferred into a Zymo-spin IIIC column in a collection tube and centrifuged at 8000xg for 30 seconds and the flow through saved. 0.8 volume ethanol (95-100%) was then added to the flow-through in the collection tube and mixed well (i.e. 320µl ethanol added to 400µl flow through). The mixture was then transferred to a Zymo-spin IIC column in a collection tube and centrifuged at 12000xg for 30 seconds. The flow-through was discarded and 400µl RNA prep buffer was added to the column and centrifuged at 12000xg for 1 minute. The flow-through was discarded and the Zymo-spin IIC column placed back in the collection tube. 800µl RNA wash buffer was added to the column and centrifuged at 12000xg for 30 seconds. The flow-through was discarded and then the Zymo-spin IIC column placed back in the collection tube. This wash step was repeated with 400µl RNA wash buffer and centrifuged at 12000xg for 2 minutes to ensure complete removal of the wash buffer. The Zymo-spin IIC column is carefully removed from the collection tube and placed into a DNase/RNase free tube and 15µl DNase/RNase free water added directly to the column and left to stand for 1 minute. After the minute the tube was then centrifuged at 1000xg for 30 seconds to elute RNA from the column. This was then immediately stored at -80°C (method taken from ZYMO research, 2009).

5.3.6 Microarray analysis and Hierarchial clustering

Microarray analysis was performed according to the Agilent GeneSpring GX version 7.3 analysis workbench manual and Mapman 3.1 instruction manual. RMA normalisation was used for inputting data and statistical analysis was performed with ANOVA using a 1.5 fold change and P value cutoff of 0.1.

Multi-Experiment Viewer (TMeV) was used to perform hierarchial clustering, using the standard default Pearsons and a distance threshold of 0.28. Only genes with a significance of 2-fold after microarray analysis were uploaded into the program.

5.3.7 Real time PCR

5.3.7.1 cDNA synthesis

200ng of RNA from each watercress line was used. 1µl oligo dt was added to each sample and then nuclease free water added to give a total volume of 5µl. The mix was then incubated at 70°C for 5 minutes and then chilled on ice for 5 minutes. To each sample 5.4µl ImPromII buffer, 3.6µl $MgCl_2$, 1.0µl dNTPs and 1µl ImPromII reverse transcriptase was added to each. The samples were then incubated at 25°C for 5 minutes (annealing), 42°C for 1 hour (extension) and 70°C for 15 minutes (inactivate reverse transcriptase). This gave cDNA for each sample ready for real time PCR analysis.

5.3.7.2 Primer design

Genes were selected from microarray analysis and also from MapMan pathway analysis. The chosen full length genomic gene sequences of At1g65060 (encodes an isoform of 4-coumarate:CoA ligase (4CL)), At4g35000 (encodes a microsomal ascorbate peroxidase APX3), At1g78440 (encodes a gibberellin 2-oxidase) and

At1g80340 (encodes a protein with gibberellin 3-beta-hydroxylase activity) were uploaded to the primer designing tool, NCBI primer-BLAST. The product size range was set to 100-250 base pairs and the annealing temperature (T_m) set to 58°C. The optimum GC primer percentage was set to 50% and the GC percentage range at 20-80%. The option to span an exon junction was selected. The reference genes chosen were Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and Actin (ACT2) which were also run through NCBI primer blast.

5.3.7.3 cDNA test dilutions

The concentration of cDNA to use throughout the real time PCR needed to be established. This was carried out by taking 3µl of cDNA from each sample and producing a pool of cDNA with a total volume of 2 µl. 54 µl of sterile de-ionised water was then added to the cDNA pool and this was then used for serial dilutions giving a series of: 1, 1/2, 1/4, 1/8 and 1/16. To each well 5 µl cDNA was added, 0.6µl forward primer, 0.6µl reverse primer, 3.8µl sterile water and 10µl SYBR green 2 X master mix (Finnzymes). The plate was then spun for ~ 2 minutes at 13,200rpm and loaded into the real time PCR (Biorad DNA engine, Chromo 4 thermocycler) for analysis. The cDNA dilutions were tested on all genes chosen.

5.3.7.4 Real Time PCR of lines Wx_0001, Wx_0011 and Wx_0033

Four plates containing two technical replicates and three biological replicates and the control genes were plated. The following genes were chosen for analysis; At1g65060 (encodes an isoform of 4-coumarate:CoA ligase (4CL) which is involved in the last step of the general phenylpropanoid pathway), At4g35000 (encodes a microsomal ascorbate peroxidise APX3), At1g78440 (encodes a gibberellin 2-oxidase), At1g80340 (encodes a protein with gibberellin 3-beta-hydroxylase activity), ACT2 (control gene) and GAPDH (control gene). A 1/4 dilution of the cDNA for each sample was made and then 5 µl pipetted into the correct well. The primer mix for each primer set was then made (each reaction had 0.6 µl F primer, 0.6 µl R primer

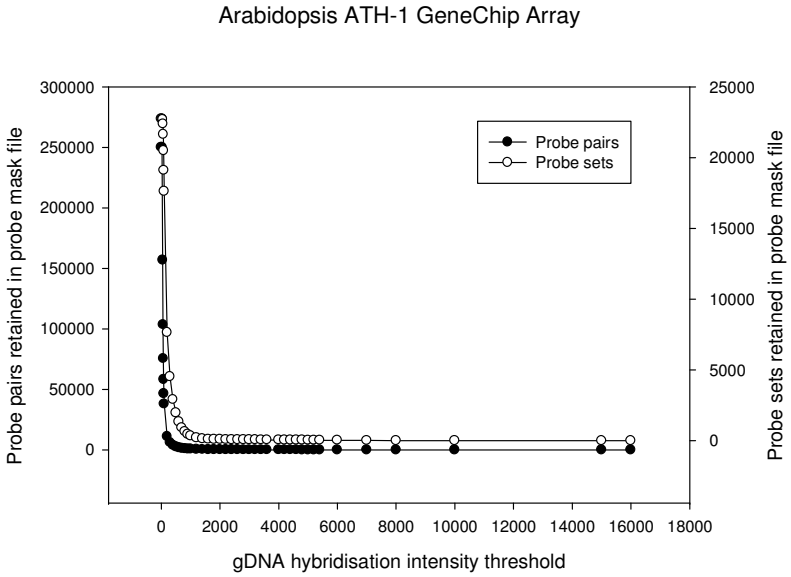
and 3.8 µl sterile de-ionised water). Just before the mix was pipetted into each well 2X SYBR master mix was added (each reaction had 10 µl) due to the light sensitivity of the SYBR green. The plate was placed in the Chromo 4 real time PCR machine (Biorad) and the following program used:

1) 95°C for 10 minutes
2) 95°C for 10 seconds
3) 60°C for 15 seconds
4) 72°C for 15 seconds
5) 75°C for 1 second
6) Plate read
7) Repeat from the second step 39 more times
8) 72°C for 5 minutes
9) Melting curve from 55°C to 95°C, read every 0.2°C, hold 1 second
10) 75°C for 5 minutes
11) End

5.4 Results

The gDNA probe-masking strategy of Hammond *et al.*, (2005a) was used to identify and eliminate oligonucleotide GeneChip probes with low or no hybridisation to watercress transcripts (Davey *et al.*, 2009). The gDNA cell intensity file (.cel) file allows perfect probe pairs showing a high hybridisation signal to watercress gDNA to be selected and probe masks to be created to exclude non-hybridising probe-pairs within that probe-set. The resultant hybridisation to the *Arabidopsis* ATH-1 gene chip was very high with a minimum intensity being set at 16000 (Figure 5.1a) in order to obtain no probe-sets or probe-pairs being retained, which was comparable to Hammond *et al*. 2005a. In order to determine the threshold intensity only the data from 0-2000 was plotted so the greatest distance between the probe-set and probe-pair could be identified. It appears that the threshold intensity was 80 (Figure 5.1b).

a)



b)

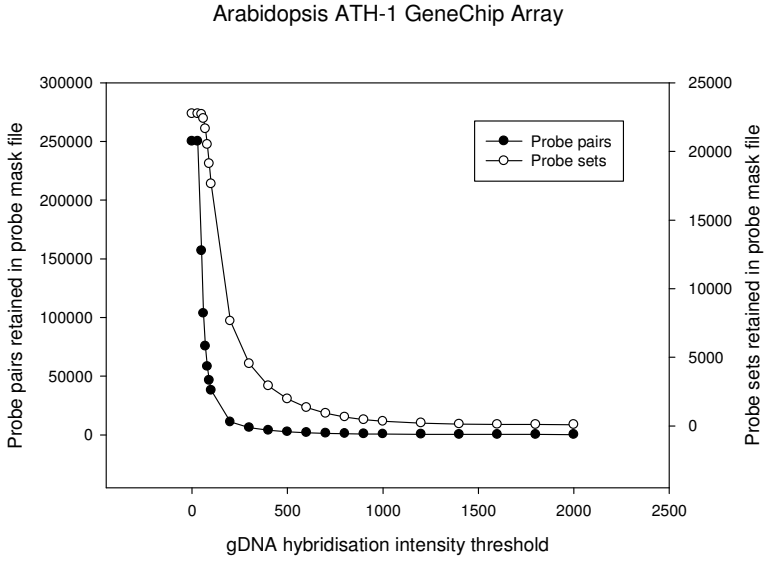


Figure 5.1a and b: *Arabidopsis thaliana* perfect match (PM) probes and probe sets from *Arabidopsis* ATH1 GeneChip array used to study transcriptomes of *Rorippa nasturtium aqauticum*. The genomic DNA hybridization intensity thresholds were used to generate the probe mask files following the hybridization of labelled gDNA to GeneChip arrays. (a) Overall differences between probe pairs and probe sets retained whilst (b) reduced scale to determine threshold level (i.e. biggest difference between probe pairs and probe sets)

After selection of perfect-match probes which hybridise efficiently to the gDNA the perfect match probes are used for interpreting the GeneChip array challenged with RNA from *Rorippa nasturtium aquaticum*. The data is loaded and analysed using the Agilent Genespring Gx Analysis Workbench version 7.3. Initial examination of gene expression was examined and the average expression for each line taken (Figure 5.2). Each continuous line in the line graph corresponds to a single gene's normalised expression (y-axis). In GeneSpring GX v7.3 the microarray data is normalised to centre values around 1.0 therefore any values greater than 1 represent up-regulated genes whilst values below 1 represent down-regulated genes.

There appears to be some quite dramatic differences in gene expression between the three lines being investigated. Some of the largest changes in gene expression were At1g52690 which is a late embryogenesis abundant protein which is up-regulated in Wx_0011 compared to Wx_0001 and Wx_0033. There is a dramatic decrease in expression of At4g35770 in Wx_0033 in comparison with Wx_0001 and Wx_0011 which is a senescence-associated gene. At4g25100, Fe-superoxide dismutase is greatly up-regulated compared to Wx_0001 and Wx_0011, this gene is also up-regulated in Wx_0011 compared to Wx_0001.

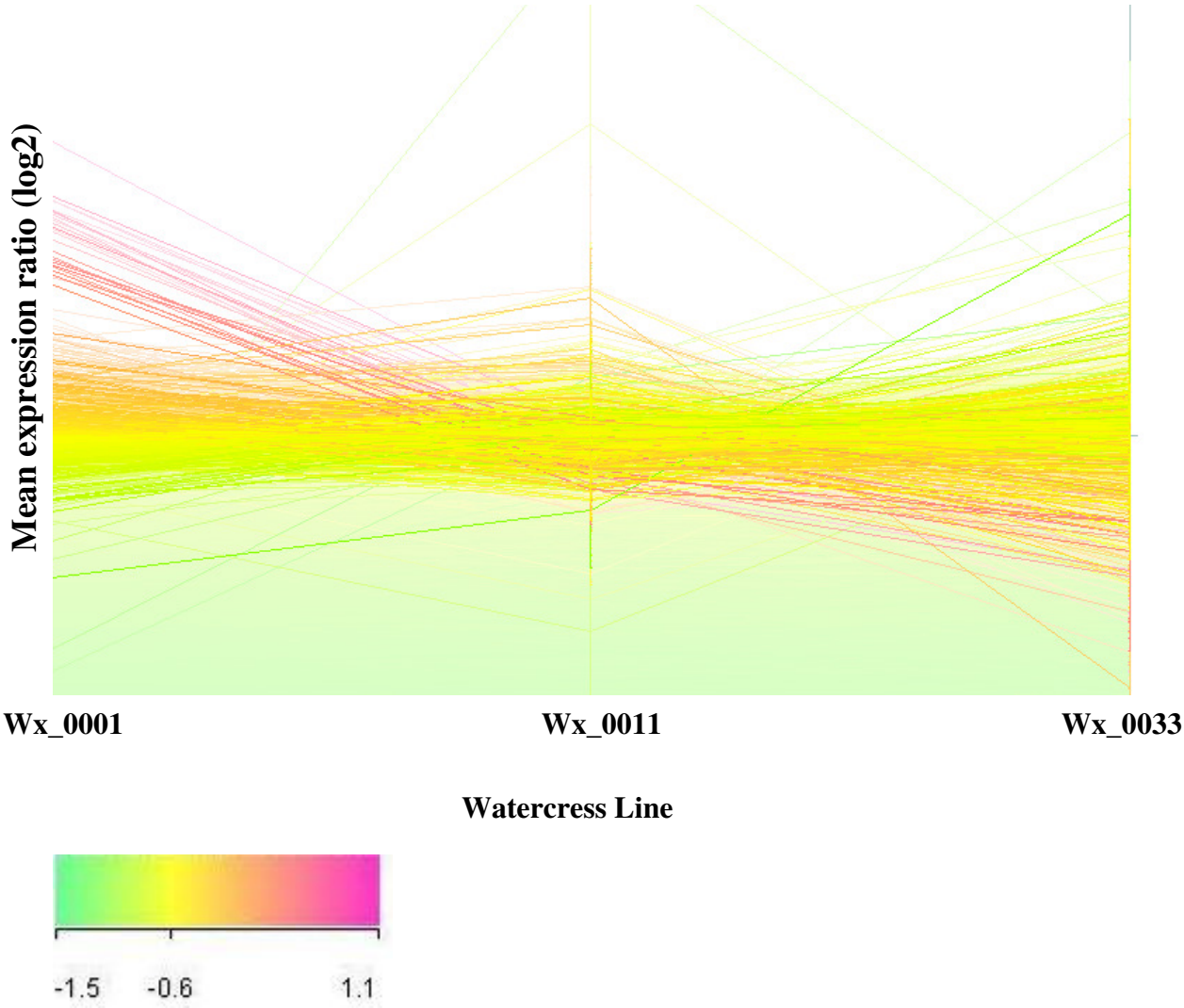
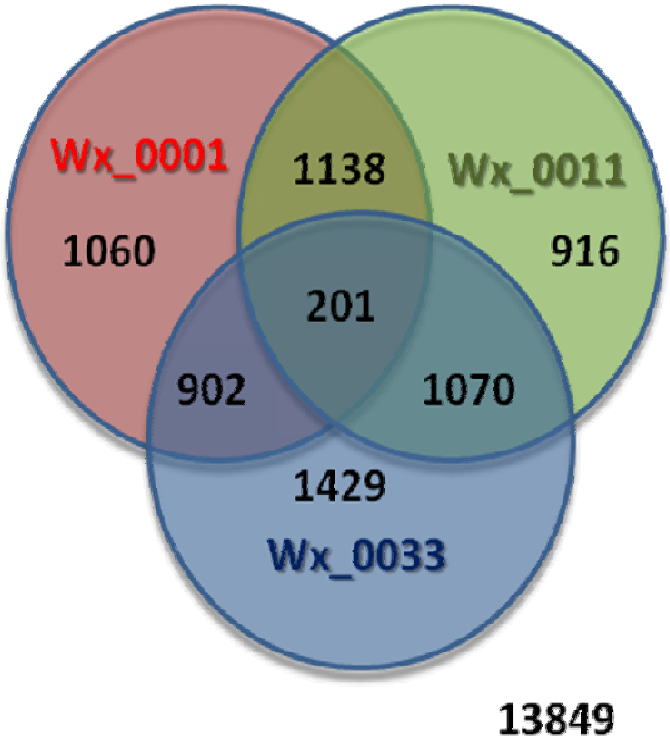


Figure 5.2: Average gene expression for lines Wx_0001 (1), Wx_0011 (11) and Wx_0033 (33). Colour scheme based on expression level of Wx_0001, expression bar indicating that green represents a down-regulation whilst pink at up-regulation of genes when RMA normalisation to the median is performed.

Venn diagrams were created in order to investigate possible relations between the three lines (Figure 5.3). Initially the up-regulation of genes were filtered on expression 1-10 (10 being the maximum in all lines) and examined. The down-regulation of genes were filtered on expression 0.1-1 (0.1 being the minimum in all lines). The examinations of up-regulated and down-regulated genes were carried out on lines Wx_0001, Wx_0011 and Wx_0033. Wx_0001 and Wx_0011 had an up-regulation of 1138 genes in common whilst only 902 genes were up-regulated in common between Wx_0001 and Wx_0033 and Wx_0011 and Wx_0033 had 1070 genes in common. Wx_0033 had the highest number of genes up-regulated 1429 whilst Wx_0011 had the lowest 916. There were 201 genes up regulated which were in common for all 3 lines under investigation. However there were 13849 genes in which there was a down or no change in regulation. These genes were not present in the any of the watercress line gene lists or in any combination of the lines.

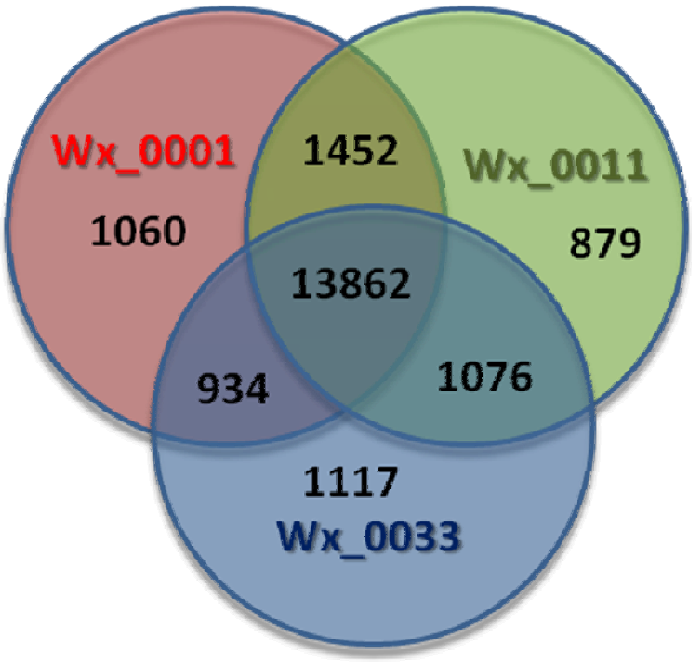
Again for down-regulation Wx_0001 and Wx_0011 have the highest number of genes in common down-regulated, 1452 genes whilst Wx_0001 and Wx_0033 has the lowest number of genes in common down-regulated, 934 genes. Wx_0033 has the overall highest number of down-regulated genes with a total of 1117, whilst Wx_0001 has a total of 1060 genes down-regulated and Wx_0011 879 genes down-regulated. There are a total of 13862 genes in common for all 3 lines down-regulated whilst 185 genes remain either down regulated or no change in expression.

a)



13849

b)



185

Figure 5.3a and b: Venn diagrams of all genes in the array that are a) all up-regulated genes and b) all down-regulated genes in each line; Wx_0001, Wx_0011 and Wx_0033

A list of genes which are significantly changed between the watercress lines was composed using ANOVA using a P value of 0.1 (Table 8). Once the list of significant genes was assembled then gene lists were viewed as volcano plots. A volcano plot shows the –Log (P value) against the Log (fold change). This shows two important measures of differential expression in one plot allowing a decision to be made on which genes to focus on.

Table 8: Significantly differentially expressed genes using ANOVA and a P value of 0.1. Genes highlighted in yellow chosen and taken forward for real time PCR whilst genes highlighted in green are further genes of interest. Function annotation from TAIR website.

Line 1 and 11	Significance (P value)	Function
261354_at (At1g79690)	0.0125 (↑ Wx_0001)	nudix hydrolase homolog 3 (NUDT3)
249037_at (At5g11680)	0.0401 (↑ Wx_0001)	FASCICLIN-like arabinogalactan protein 13 precursor (FLA13)
250332_at (At5g11680)	0.0365 (↑ Wx_0011)	Molecular function unknown
249134_at (At5g43150)	0.0332 (↑ Wx_0011)	Molecular function unknown
245358_at (At4g15900)	0.00926	Encodes a nuclear WD40 protein that is imported into the nucleus.
248029_at (At5g55700)	0.0411	BAM4 may play a regulatory role in starch breakdown.
259340_at (At3g03870)	0.00949	Unknown protein
263600_at (At2g16390)	0.0118	Putative chromatin remodeling protein.
256676_at (At3g52180)	0.0337	Encodes a plant-specific protein phosphatase.
253935_at (At4g26870)	0.00439	Class II aminoacyl-tRNA and biotin synthetases superfamily protein.

Line 1 and 33	Significance (P value)	Function
248311_at (At5g52570)	0.0115	Converts β-carotene to zeaxanthin via cryptoxanthin.
247232_at (At5g64940)	0.0314	Encodes a member of ATH subfamily of ATP-binding cassette (ABC) proteins.
266719_at (At2g46830)	0.0262	Encodes a transcriptional repressor.
256751_at (At3g27170)	0.0245	Member of Anion channel protein family.
262526_at (At1g17050)	0.0463	Encodes a protein with solanesyl diphosphate synthase activity.
258181_at (At3g21670)	0.0488	Encodes a protein with solanesyl diphosphate synthase activity.
249677_at (At5g35970)	0.0385	P-loop containing nucleoside triphosphate hydrolases superfamily protein
254050_s_at (At4g25670)	0.0158	Unknown protein.
250007_at (At5g18670)	0.029	Putative beta-amylase BMY3 (BMY3)
260914_at (At1g02640)	0.00291	Encodes a protein similar to a beta-xylosidase located in the extracellular matrix.
249037_at (At5g44130)	0.0401	FASCICLIN-like arabinogalactan protein 13 precursor (FLA13).
249862_at (At5g22920)	0.0291	CHY-type/CTCHY-type/RING-type Zinc finger protein; FUNCTIONS IN: zinc ion binding;

244992_s_at (no annotation available)	0.0301	
261834_at (At1g10640)	0.0255	Pectin lyase-like superfamily protein.
263845_at (At2g37040)	0.0158	Encodes PAL1, a phenylalanine ammonia-lyase.
253174_at (At4g35090)	0.0275	Encodes a peroxisomal catalase, highly expressed in bolts and leaves.
250665_at (At5g06980)	0.0324	Unknown protein.
250257_at (At5g13770)	0.0334	Pentatricopeptide repeat (PPR-like) superfamily protein.
261354_at (At1g79690)	0.0125	nudix hydrolase homolog 3 (NUDT3).
267035_at (At2g38400)	0.0374	alanine:glyoxylate aminotransferase 2 homolog (AGT3) mRNA.
249798_at (At5g23730)	0.0448	Encodes REPRESSOR OF UV-B PHOTOMORPHOGENESIS 2 (RUP2).
247586_at (At5g60660)	0.0306	A member of the plasma membrane intrinsic protein subfamily PIP2.
245895_at (At5g09230)	0.0102	Encodes SRT2, a member of the SIR2 (sirtuin) family HDAC (histone deacetylase).
254496_at (At4g20070)	0.0394	The gene encoding Arabidopsis thaliana Allantoate Amidohydrolase (AtAAH).
267012_at (At2g39220)	0.0174	PATATIN-like protein 6 (PLP6).
264956_at (At1g76990)	0.0316	ACT domain repeat 3 (ACR3).
249888_s_at (At5g22480)	0.018	ZPR1 zinc-finger domain protein.
263985_at (At2g42750)	0.0499	DNAJ heat shock N-terminal domain-containing protein.
263761_at (At2g21330)	0.0205	Fructose-bisphosphate aldolase 1 (FBA1).
259922_at (At1g72770)	0.0348	Mutant has ABA hypersensitive inhibition of seed germination.
265067_at (At1g03850)	0.0192	Glutaredoxin family protein.
246216_at (At4g36380)	0.0145	Encodes a cytochrome P-450 gene.
247921_at (At5g57660)	0.0327	CONSTANS-like 5 (COL5).
259106_at (At3g05490)	0.0443	Member of a diversely expressed predicted peptide family.
255692_at (At4g00400)	0.0111	Involved in cutin assembly. Is functionally redundant with GPAT4.
257745_at (At3g29240)	0.0355	Protein of unknown function.
249932_at (At5g22390)	0.0218	Protein of unknown function.
245362_at (At4g17460)	0.05	Encodes homeobox protein HAT1.
251356_at (At3g61060)	0.0139	Phloem protein 2-A13 (PP2-A13); CONTAINS InterPro DOMAIN/s: F-box domain, cyclin-like.
255690_at (At4g00360)	0.0183	Encodes a member of the CYP86A subfamily of cytochrome p450 genes.
261907_at (At1g65060)	0.0424 (↑ Wx_0001)	encodes an isoform of 4-coumarate:CoA ligase (4CL),
263151_at (At1g54120)	0.0438	Unknown protein.
254315_at ?	0.0371	
259501_at (At1g15750)	0.0373	Encodes a protein with several WD40 repeats at the C-terminus.
246523_at (At5g15850)	0.0142	Homologous to the flowering-time gene CONSTANS.

259955_s_at (At1g75080)	0.0198	Encodes a positive regulator of the brassinosteroid (BR) signalling pathway.biosynthesis.
250517_at (At5g08260)	0.0412	Serine carboxypeptidase-like 35 (scpl35).
251015_at (At5g02480)	0.0418	HSP20-like chaperones superfamily protein.
260489_at (At1g51610)	0.0303	Cation efflux family protein.
260283_at (At1g80480)	0.0328	Plastid transcriptionally active 17 (PTAC17).
267239_at (At2g02510)	0.00774	NADH dehydrogenase (ubiquinone)s.
247910_at (At5g57410)	0.00729	Afadin/alpha-actinin-binding protein.
265352_at (At2g16600)	0.0278	Encodes cytosolic cyclophilin ROC3.
261767_s_at (At1g15500)	0.0194	ATNTT2.
263600_at (At2g16390)	0.0118	Putative chromatin remodeling protein.
258184_at (At3g21510)	0.0343	Encodes AHP1.
254099_at (At4g25130)	0.0118	Encodes a chloroplast-localized methionine sulfoxide reductase.
258223_at (At3g15840)	0.0263	Encodes a chloroplast-targeted protein localized in the stroma.
261412_at (At1g07890)	0.0159 (↑ Wx_0033)	Encodes a cytosolic ascorbate peroxidase APX1.
264364_at (At1g03330)	0.0233	Small nuclear ribonucleoprotein family protein.
246651_at (At5g35170)	0.031	Adenylate kinase family protein.
264439_at (At1g27450)	0.00639	Denosine phosphoribosyl transferase(E.C:2.4.2.7).
264811_at (At1g08640)	0.0338	Chloroplast J-like domain 1 (CJD1).
257756_at (At3g18680)	0.0294	Amino acid kinase family protein.
246268_at (At1g31800)	0.00648 (↑ Wx_0033)	Encodes a protein with β-ring carotenoid hydroxylase activity.
267153_at (At2g30860)	0.00439 (↑ Wx_0033)	Encodes glutathione transferase belonging to the phi class of GSTs.
265357_at (At2g16740)	0.00405	Ubiquitin-conjugating enzyme 29 (UBC29).
266207_at (At2g27680)	0.0467 (↑ Wx_0033)	NAD(P)-linked oxidoreductase superfamily protein.
253981_at (At4g26670)	0.01	Mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein.
256451_s_at (At1g75170)	0.0238	Sec14p-like phosphatidylinositol transfer family protein.
248128_at (At5g54770)	0.0131	Encodes a thiamine biosynthetic gene.
256676_at (At3g52180)	0.0337	Encodes a plant-specific protein phosphatase.
265966_at (At2g37220)	0.0188	Encodes a chloroplast RNA binding protein.
250339_at (At5g11670)	0.00874	The malic enzyme (EC 1.1.1.40).
246004_at (At5g20630)	0.02	Encodes a germin-like protein.
254098_at (At4g25100)	0.0256 (↑ Wx_0033)	Fe-superoxide dismutase.
263548_at (At2g21660)	0.0418	Encodes a small glycine-rich RNA binding protein.
266673_at (At2g29630)	0.0327	Encodes a protein involved in thiamin biosynthesis.

Line 11 and 33	Significance (P value)	Function
253358_at (At4g32940)	0.0484	Encodes a vacuolar processing enzyme.
253174_at (At4g35090)	0.0275 (↑ Wx_0011)	Encodes a peroxisomal catalase.
256751_at (At3g27170)	0.0245	
249862_at (At5g22920)	0.0291	CHY-type/CTCHY-type/RING-type Zinc finger protein.
245362_at (At4g17460)	0.05	Encodes homeobox protein HAT1.
246304_at (At3g51840)	0.0455 (↑ Wx_0011)	Encodes a short-chain acyl-CoA oxidase.
263985_at (At2g42750)	0.0499	DNAJ heat shock N-terminal domain-containing protein.
259319_at (At3g01090)	0.032	?
265067_at (At1g03850)	0.0192 (↑ Wx_0011)	Glutaredoxin family protein.
260914_at (At1g02640)	0.00291 (↑ Wx_0011)	Encodes a protein similar to a beta-xylosidase located in the extracellular matrix.
254050_s_at (At4g25670)	0.0158	Unknown protein.
253223_at (At35000)	0.0483 (↑ Wx_0011)	Encodes a microsomal ascorbate peroxidase APX3.
250007_at (At5g18670)	0.029	Putative beta-amylase BMY3 (BMY3).
266781_at (At2g28940)	0.0493	Protein kinase superfamily protein.
259340_at (At3g03870)	0.00949	Unknown protein.
250032_at (At5g18170)	0.0349	Encodes the 43 kDa alpha-subunit of the glutamate dehydrogenase.
259922_at (At1g72770)	0.0348	Mutant has ABA hypersensitive inhibition of seed germination.
247586_at (At5g60660)	0.0306	A member of the plasma membrane intrinsic protein subfamily PIP2.
245431_at (At4g17080)	0.0221	Histone H3 K4-specific methyltransferase SET7/9 family protein.
267035_at (At2g38400)	0.0374	Alanine:glyoxylate aminotransferase 2 homolog (AGT3) mRNA.
259293_at (At3g11580)	0.0452	AP2/B3-like transcriptional factor family protein.
263845_at (At2g37040)	0.0158 (↑ Wx_0011)	Encodes PAL1, a phenylalanine ammonia-lyase.
267012_at (At2g39220)	0.0174	PATATIN-like protein 6.
250257_at (At5g13770)	0.0334	Pentatricopeptide repeat (PPR-like) superfamily protein.
247921_at (At5g57660)	0.0327	CONSTANS-like 5 (COL5).
246523_at (At5g15850)	0.0142	Homologous to the flowering-time gene CONSTANS.
257745_at (At3g29240)	0.0355	Protein of unknown function.
264788_at (At2g17880)	0.00218	Chaperone DnaJ-domain superfamily protein.
245895_at (At5g09230)	0.0102	Encodes SRT2, a member of the SIR2 (sirtuin) family HDAC (histone deacetylase).
266106_at (At2g45170)	0.0494	Involved in autophagy. Under nutrient starvation the protein localizes to autophagosomes.
255692_at (At4g00400)	0.0111	Involved in cutin assembly. Is functionally redundant with GPAT4.

248622_at (At5g49360)	0.0304	Encodes a bifunctional {beta}-D-xylosidase/{alpha}-L-arabinofuranosidase.
248764_at (At5g47640)	0.028	Nuclear factor Y, subunit B2" (NF-YB2).
258939_at (At3g10020)	0.00431	Unknown protein.
247910_at (At5g57410)	0.00729	Afadin/alpha-actinin-binding protein.
252669_at (At3g44100)	0.0111	MD-2-related lipid recognition domain-containing protein.
249318_at (At5g40870)	0.00105	Encodes a protein that appears to possess both uridine kinase and uracil phosphoribosyltransferase activities.
253112_at (At4g35970)	0.0266 (↑ Wx_0033)	Encodes a microsomal ascorbate peroxidase APX5.
250739_at (At5g05740)	0.0474	S2P-like putative metalloprotease.
250367_s_at (At5g11200)	0.0115	DEAD/DEAH box RNA helicase family protein.
262202_at (At2g01110)	0.0378	Mutant is Albino and pale green; Chloroplast Protein Translocation (tatC).
253981_at (At4g26670)	0.01	Mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein.
267153_at (At2g30860)	0.00439 (↑ Wx_0033)	Encodes glutathione transferase belonging to the phi class of GSTs.
250339_at (At5g11670)	0.00874	The malic enzyme (EC 1.1.1.40).phosphate pathway.
254098_at (At4g25100)	0.0256 (↑ Wx_0033)	Fe-superoxide dismutase.

The list of significant genes generated from inputting the data into Genespring were then uploaded to the Multi Experiment Viewer (TMeV) (Saeed *et al*. 2006) to investigate clutering of these genes. From this program it was evident that Wx_0001 and Wx_0011 were more closely related in respect to gene expression patterns than with Wx_0033 (Figure 5.4). This could explain the dwarf phenotype (Chapter four) and high antioxidant power (Chapter three) of Wx_0033 compared to Wx_0001 and Wx_0011. This may also explain the 24% among variation presented in Chapter four from the AFLP data due to the differing gene expression of Wx_0033. Seven gene clusters were revealed but only two of these seven had significant GO term categories i.e. enough similarity to classify the cluster with overall functions. In order to reveal the function of the genes in each of the clusters the data was uploaded to AgriGo (Du *et al*. 2010). Clusters 1 and 2 had significant GO term categories in which the genes in cluster 1 are mainly associated with cellular process, cellular component organisation, cellular component biogenesis, multicellular organismal process, developmental process and response to stimulus whilst those in 2 associated with metabolic process, cellular process and response to stimulus. Clusters 3-7 did not have a single GO term which defined the genes in each cluster therefore although they have similar expression patterens the functions vary.

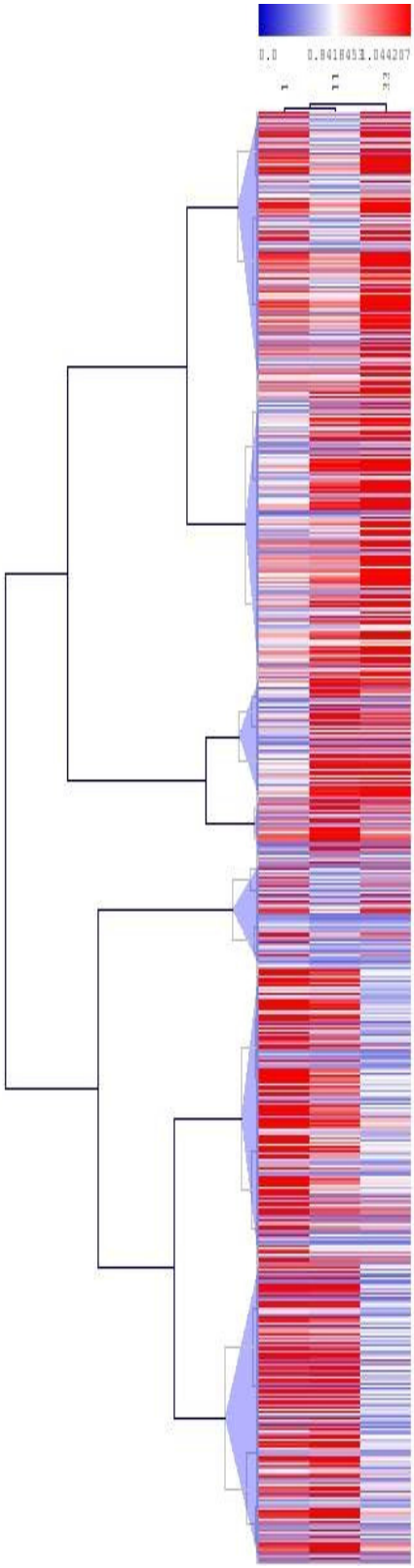


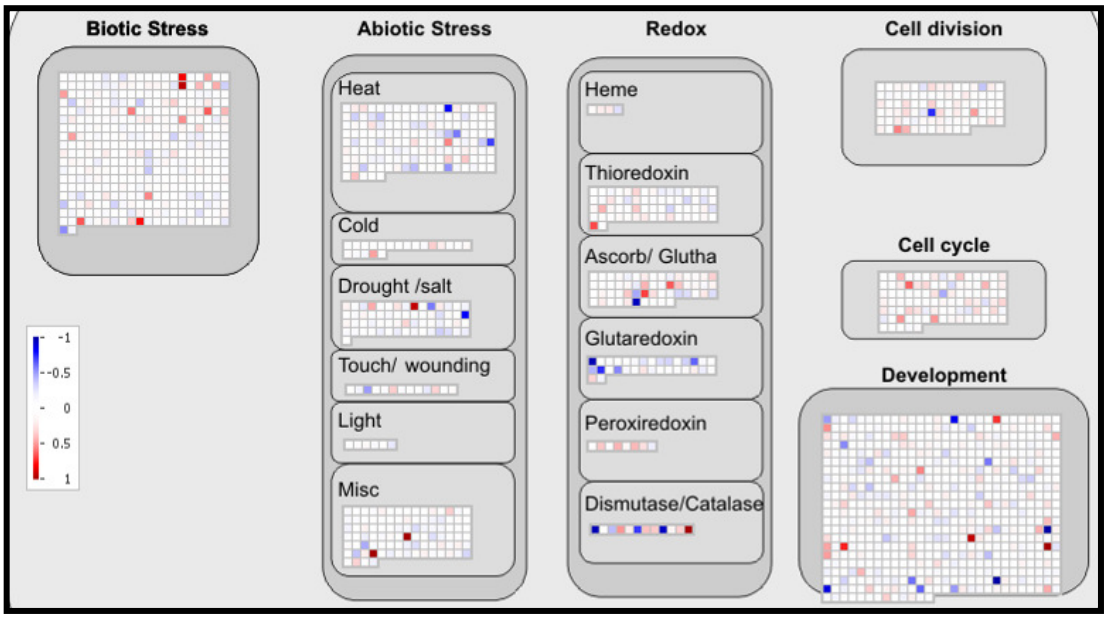
Figure 5.4: Hierarchical clustering of changes in transcript abundance in three different watercress lines (Wx_0001, Wx_0011 and Wx_0033) which was performed using TMeV software from normalised expression ratio data. Lilac triangles represent seven different clusters, colour scale indicates signal normalised ratios (red = upregulated genes, blue = downregulated genes). Only genes with ≥ 2 fold expression change and ANOVA significance p≤ 0.05 are shown.

As well as loading the microarray data into GeneSpring v7.3.1, the data was also loaded into MapMan. MapMan is a user driven tool which displays large data sets and displays gene expression experiments on diagrams of metabolic pathways or other processes: Wx_0033 was compared to Wx_0001, Wx_0011 with Wx_0001 and finally Wx_0033 with Wx_0011. The MapMan program allowed the expression of genes either up or down regulated to be visualised on pathways including large enzyme families, secondary metabolism, phenylpropanoids, regulation overview and cellular response overview. This program is of great importance as it can allow individual genes to be identified and therefore specific pathways can be focused on. The use of blue and red for expression on squares which represent individual genes means patterns in expression can be instantly recognised, blue representing a down-regulation and red an up-regulation.

From MapMan it can be seen that dismutase/catalase genes are upregulated in Wx_0033 compared to Wx_0001. This is evident from the bin which corresponds with dismutase/catalase genes having a few boxes (which relate to single genes) red in colour (i.e. up-expressed) (Figure 5.5).

The genes in this bin are mainly regarding the redox – dismutases and catalases. Some of the genes in this bin which are up-regulated include genes with Fe-superoxide dismutase activity, genes which catalyse the breakdown of hydrogen peroxide into water and oxygen and genes which catalyse the reduction of hydrogen peroxide using heme group as a cofactor therefore protecting cells from toxicity by hydrogen peroxide.The functions of superoxide dismutase (SOD), catalase (CAT), glutathione peroxidase (GPx) and ascorbate peroxidase (APx) are to remove reactive oxygen species in plant cells and therefore protect plants from damage caused by reactive oxygen species. In plants APx and CAT are predominant in detoxification of H_2O_2 while GPx serves additional important roles in the removal of lipid and alkyl peroxides, the up-regulation of glucosinolates is also notable.

Another bin which is up-regulated in Wx_0033 compared to Wx_0001 is the bin referring to secondary metabolism. One gene in this bin belongs to cytochrome P450 and is involved in trytophan metabolism which converts Trp to indo-3-acetaldoxime (IAOx), a precursor to IAA and indole glucosinolates. Another gene up-regulated in Wx_0033 compared to Wx_0001 is a gene which encodes an oxime-metabolising enzyme in the biosynthetic pathway of glucosinolates. Also a gene which encodes a glycosyl hydrolase that localises to peroxisomes and acts as a component of an inducible preinvasion resistance mechanism is up-regulated in this bin in Wx_0033 compared to Wx_0001,



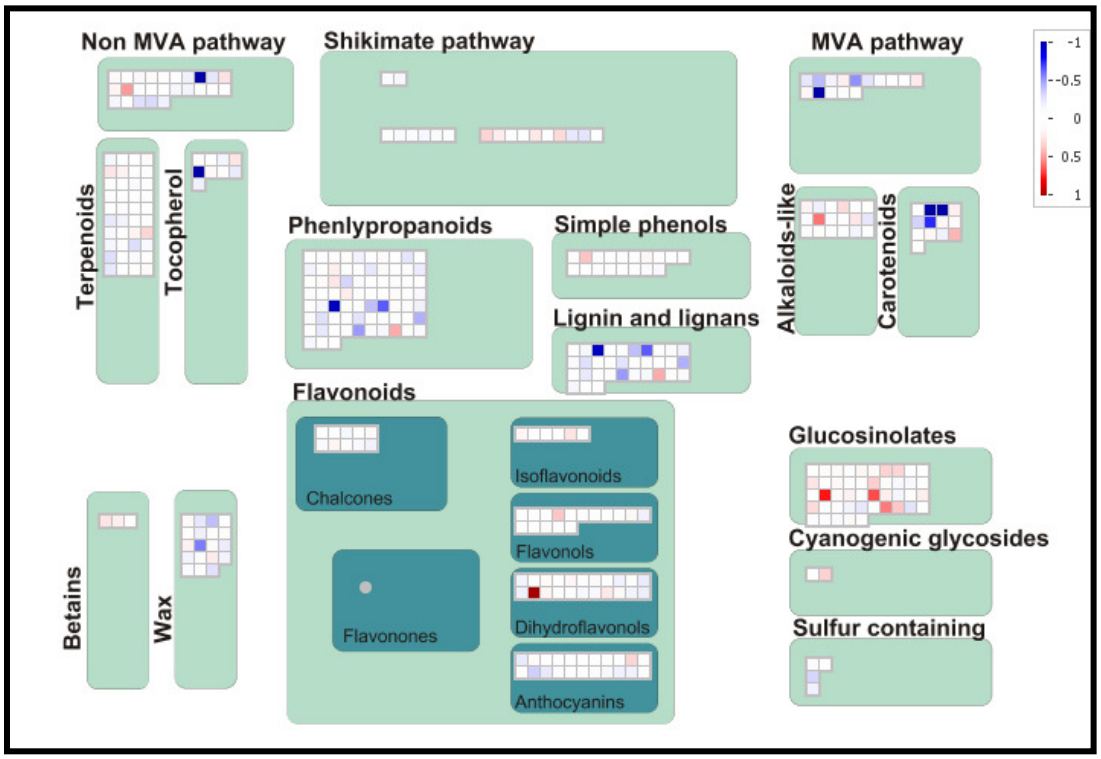
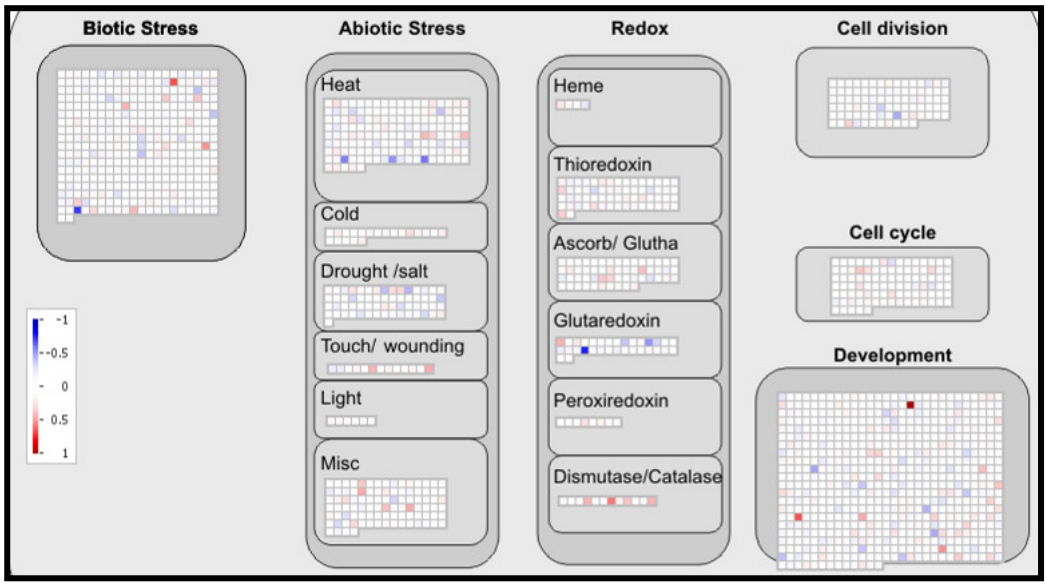


Figure 5.5: Comparison of gene regulation of Wx_0033 on Wx_0001. Ratio of gene expression of Wx_0033 on Wx_0001 in which the red boxes represent an up-regulation of a specific gene in Wx_0033 compared to Wx_0001 whilst blue down-regulation. It is clear that the glucosinolates are up-regulated in Wx_0033 compared to Wx_0001 which supports FRAP findings in Chapter three.

When comparing the regulation of Wx_0011 on Wx_0001 we can clearly see that the bin containing dismutase/catalase genes contains red boxes and are therefore upregulated in Wx_0011. Some of the genes up-regulated in this bin include a gene which catalyses the reduction of hydrogen peroxide using a heme group as a cofactor, a gene which catalyses the breakdown of hydrogen peroxide into water and oxygen, a gene which encodes a peroxisomal catalase and a genewhich has Fe-superoxide dismutase activity. Also there is an up-regulation of the genes in the bin corresponding to glucosinolates (secondary metabolism) in Wx_0011 (Figure 5.6). This includes genes which encode an oxime-metabolising enzyme in the biosynthetic pathway of glucosinolates, a gene which belongs to the flavin monooxygenase S-oxygenase that catalyses the conversion of methythioalkyl glucosinolates to methylsulfinylalkyl glucosinolates and a gene which belongs to cytochrome P450 and is involved in trytophan metabolism. It converts Trp to indo-3-acetaldoxime (IAOx), a precursor to IAA and indole glucosinolates.



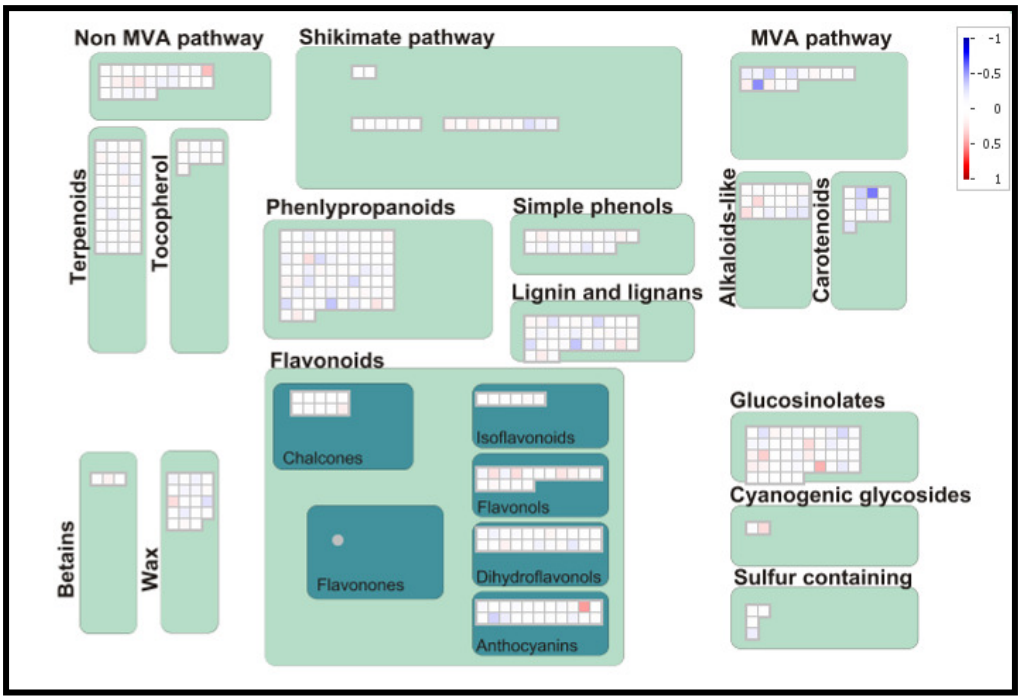


Figure 5.6: Comparison of the regulation of genes of Wx_0011 on Wx_0001. Ratio of gene expression of Wx_0011 on Wx_0001 in which the red boxes represent an up-regulation of a specific gene in Wx_0011 compared to Wx_0001 whilst blue down-regulation. Again although not such an intense or high number of red boxes (which represent individual genes) it appears that there is a higher expression of glucosinolates in Wx_0011 compared to Wx_0001, reflecting FRAP results in Chapter three.

When comparing Wx_0033 to Wx_0011 the bin corresponding to glucosinolates (secondary metabolism) appear to show an up-regulation in Wx_0033 when the regulation of these genes were compared to Wx_0011. Genes which are up-regulated in Wx_0033 compared to Wx_0011 in this bin include a gene which belongs to cytochrome P450 and is involved in trytophan metabolism and converts Trp to indo-3-acetaldoxime (IAOx) which is a precursor to IAA and indole glucosinolates, a gene which encodes an oxime-metabolising enzyme in the biosynthetic pathway of glucosinolates, a gene which is one of the three genes encoding the enzyme 3-isopropylmalate dehydrogenase involved in leucine biosynthesis and a gene which encodes a cytochrome P450 enzyme that catalyses the initial conversion of aldoximes to thiohydroximates in the synthesis of glucosinolates not derived from trytophan. Also we can see that the bin showing the final steps in the phenylpropanoid pathway are upregulated in Wx_0033 when compared to the regulation of these genes in Wx_0001 (Figure 5.7). A gene up-regulated in Wx_0033 compared to Wx_0011 is a gene which encodes a catalytically active cinnamyl alcohol dehydrogenase which uses p-coumaryl aldehyde as a preferred substrate. It can also use caffeyl, coniferyl and d-hydroxyconiferyl aldehydes as substrates.

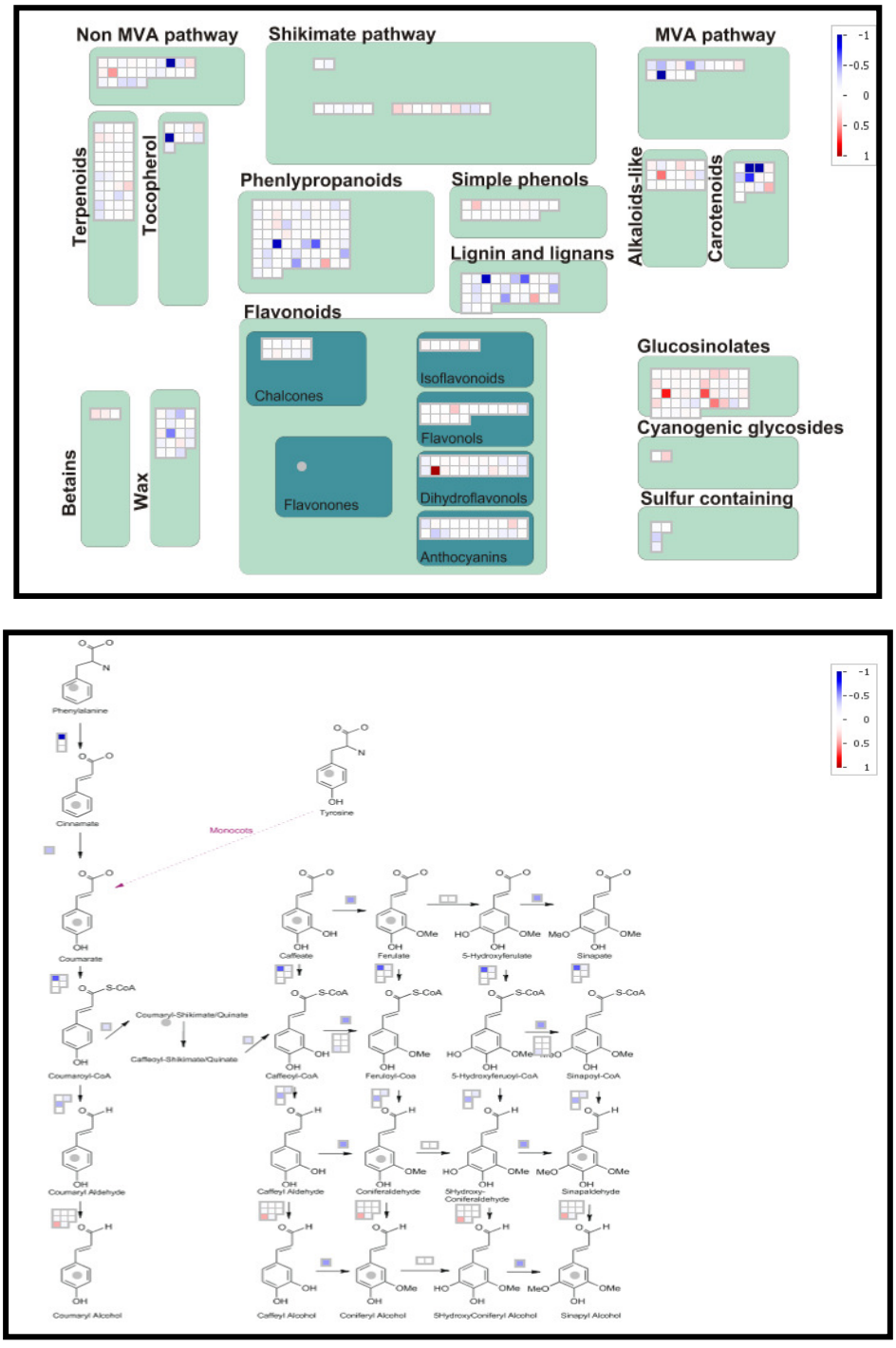


Figure 5.7: Comparison of the regulation of genes of Wx_0033 on Wx_0011. Ratio of gene expression of Wx_0033 on Wx_0011 in which the red boxes represent an up-regulation of a specific gene in Wx_0033 compared to Wx_0011 whilst blue down-regulation.

Microarray analysis gave whole genome profiling which allows the expression pattern of individual genes to be explored whilst MapMan allowed visualisation of specific pathways for gene expression. Using output from both microarray analysis and MapMan four genes were selected for real time PCR analysis which would allow confirmation of gene expression. Genes chosen from microarray analysis include a gene encoding for phenylpropanoids, At1g65060 and a gene encoding ascorbate peroxidase which scavenges hydrogen peroxide, At4g35000. The genes chosen from MapMan were one which encodes a protein with 3 beta-hydroxylase activity and this catalyses the hydroxylation of both GA9 and GA20, At1g80340 (downregulated in Wx_0033 compared to Wx_0011). The second gene chosen was a gene encoding a gibberellin 2-oxidase that acts on C19 gibberellins, At1g78440 (up-regulated in Wx_0033 compared to Wx_0011). Two housekeeping genes were chosen to act as control genes actin (ACT) and glyceraldehyde-3-phosphate dehygrogenase (GAPDH) which were selected from work by Czechowski *et al*. (2005). Real time PCR was being used purely as a confirmation of the microarray data. With the selection of a few genes to examine expression, real time PCR would enable the acceptance of the expression data provided by the ATH1 microarray.

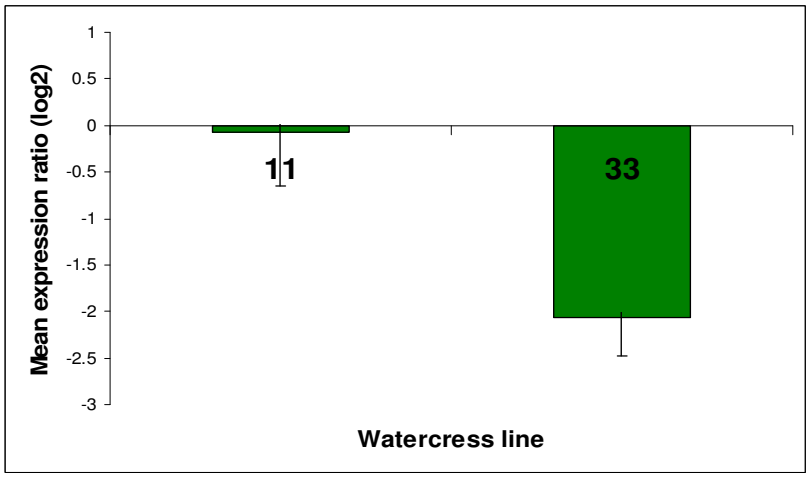
Initial investigation into the efficiency of the primers gave promising results at the different concentrations of cDNA (1,1/2,1/4,1/8 and 1/16). However upon closer inspection it was apparent that the genes At1g78440 and At1g80340 failed, leaving At1g65060 and At4g35000 with the reference gene GAPDH. The results were obtained from a pool of cDNA differing in concentrations; however for further analysis in qRT–PCR the cDNA was at 200ng for each sample.

A cDNA dilution of 1/4 was taken forward for real time PCR analysis using the reference gene GAPDH. In order to calculate the expression ratio the mathematical model proposed by Pfaffl (2001) was followed. This calculation involves calculating the efficiency of each primer and the cycle thresholds recorded and incorporating these into equation 5.1.

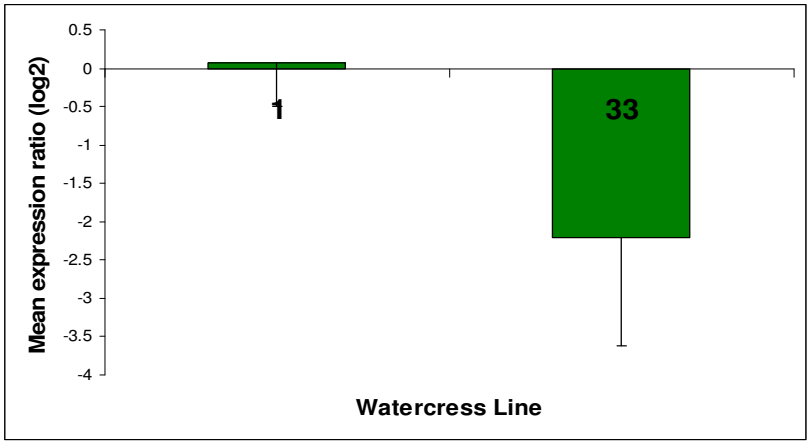
Expression ratio of the target gene normalised to the reference gene and relative to the control = $(E_{GOI})C(t)control-C(t)treatment/(E_{REF})C(t)control-C(t)treatment$ **(eqn 5.1)**

The qRT-PCR results for the gene At1g65060 are in complete agreement with the microarray analysis. When each line is used as a control we can see the results mirror the microarray results. Using Wx_0001 as the control line Wx_0011 is down-regulated but Wx_0033 is down-regulated to a greater extent (Figure 5.8a). When Wx_0011 is used as the control line, Wx_0001 is up-regulated in comparison and Wx_0033 is again down-regulated (Figure 5.8b). Using Wx_0033 as control an up-regulation for both Wx_0001 and Wx_0011 (Figure 5.8c). The error bar also reflects the level of variation within each biological sample, a very small error bar for Wx_0033 due to minimal variation between the biological replicates (Figure 5.9).

a)



b)



c)

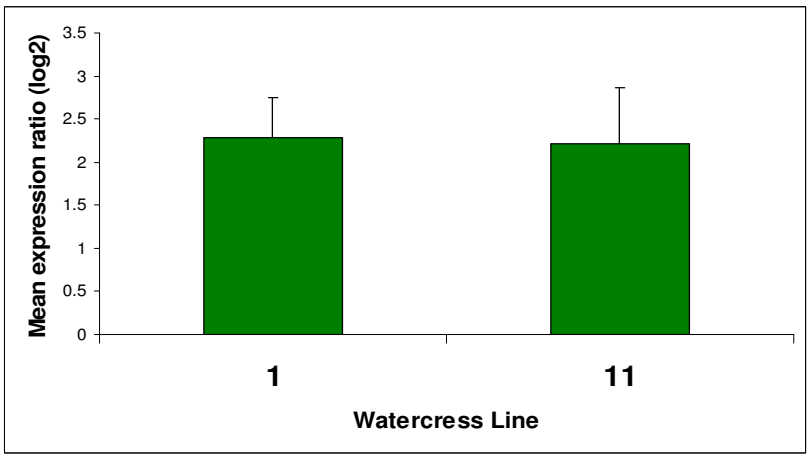


Figure 5.8: Real time PCR expression ratio (±SE) of a gene encoding 4-CL (At1g65060) involved in the final steps of the phenylpropanoid pathway in Wx_0001, Wx_0011 and Wx_0033, (a) gene expression ratio of Wx_0011 and Wx_0033 compared to Wx_0001 (b) gene expression ratio of Wx_0001 and Wx_0033 compared to Wx_0011 (c) gene expression ratio of Wx_0001 and Wx_0011 compared to Wx_0033.

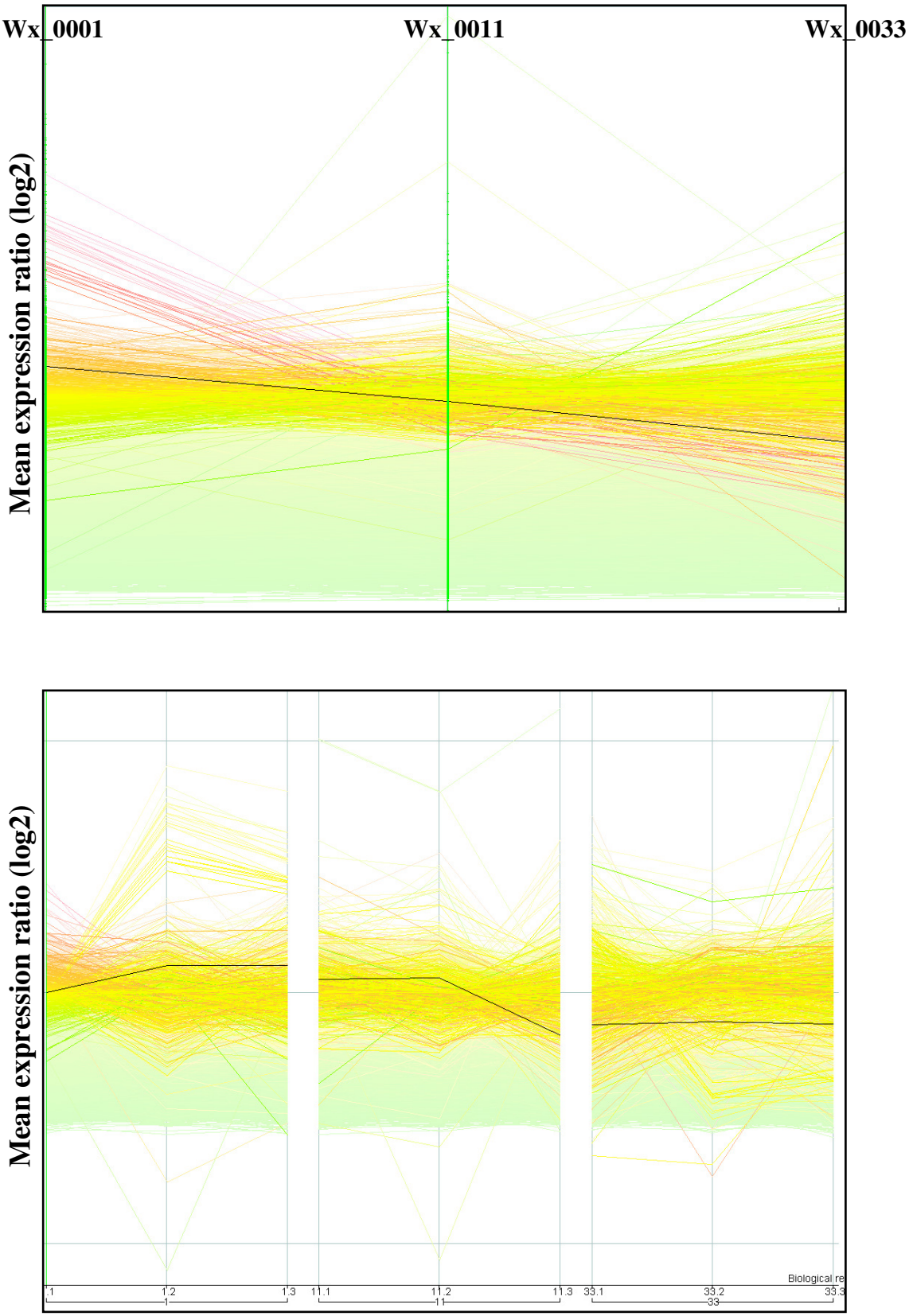
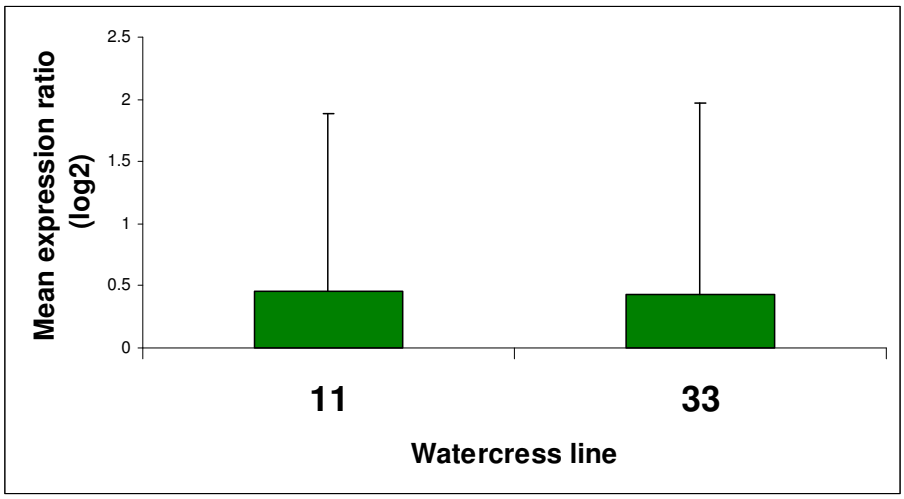


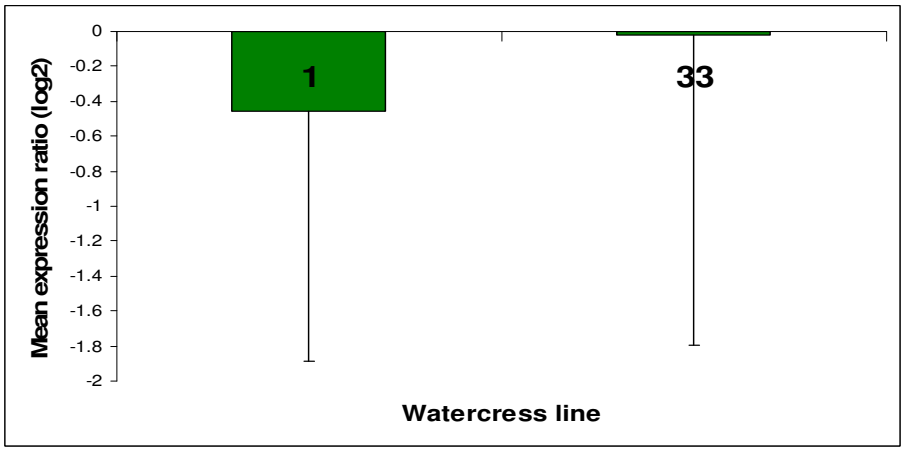
Figure 5.9: Normalised expression of Wx_0001, Wx_0011 and Wx_0033 to *Arabidopsis* ATH1 genechip.

The qRT- PCR results for the gene At4g35000 are not in complete agreement with the microarray analysis, the outlying line is Wx_0033. Wx_0001 as a control we see the up-regulation of both Wx_0011 and Wx_0033 (Figure 5.10a) however in comparison to Wx_0001 with respect to microarray analysis Wx_0033 is down-regulated. With Wx_0011 as the control there was a larger down regulation in Wx_0001 compared to Wx_0033 (Figure 5.10b) which the microarray results also demonstrate (Figure 5.11). Finally using Wx_0033 as the control line Wx_0001 was down-regulated and Wx_0011 is up-regulated (Figure 5.10c) which are again in agreement with the micoarray results (Figure 5.11). The error bars also reinforce the large amount of variation present between the biological replicates (Figure 5.11).

a)



b)



c)

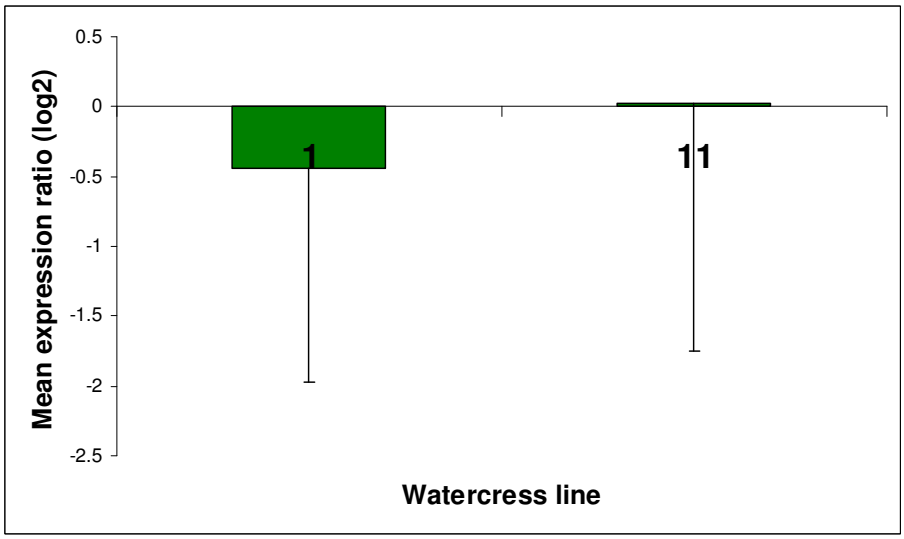


Figure 5.10: Real time PCR expression ratio (±SE) of a gene encoding an ascorbate peroxidase (At4g35000) in Wx_0001, Wx_0011 and Wx_0033, (a) gene expression ratio of Wx_0011 and Wx_0033 compared to Wx_0001 (b) gene expression ratio of Wx_0001 and Wx_0033 compared to Wx_0011 (c) gene expression ratio of Wx_0001 and Wx_0011 compared to Wx_0033.

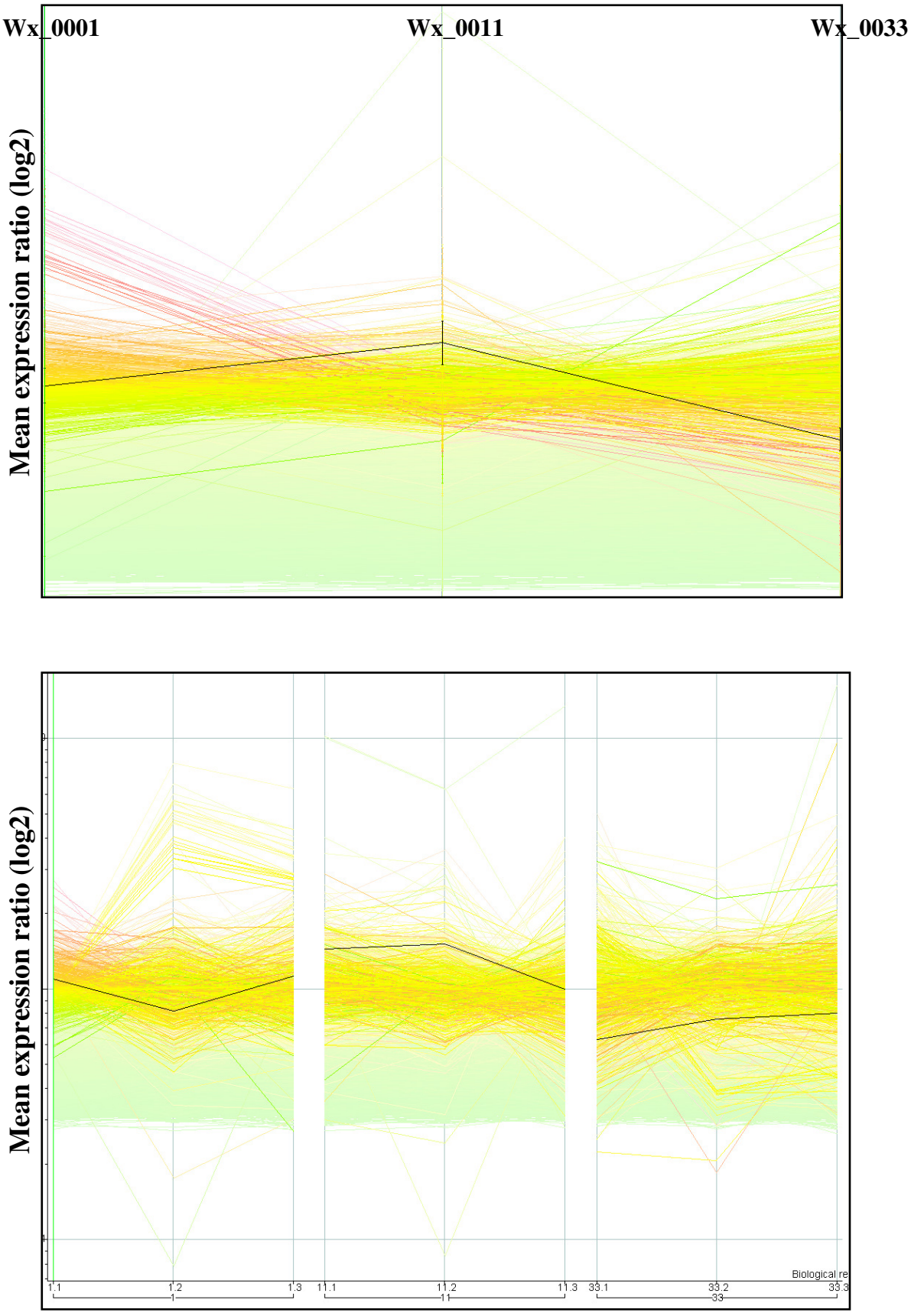
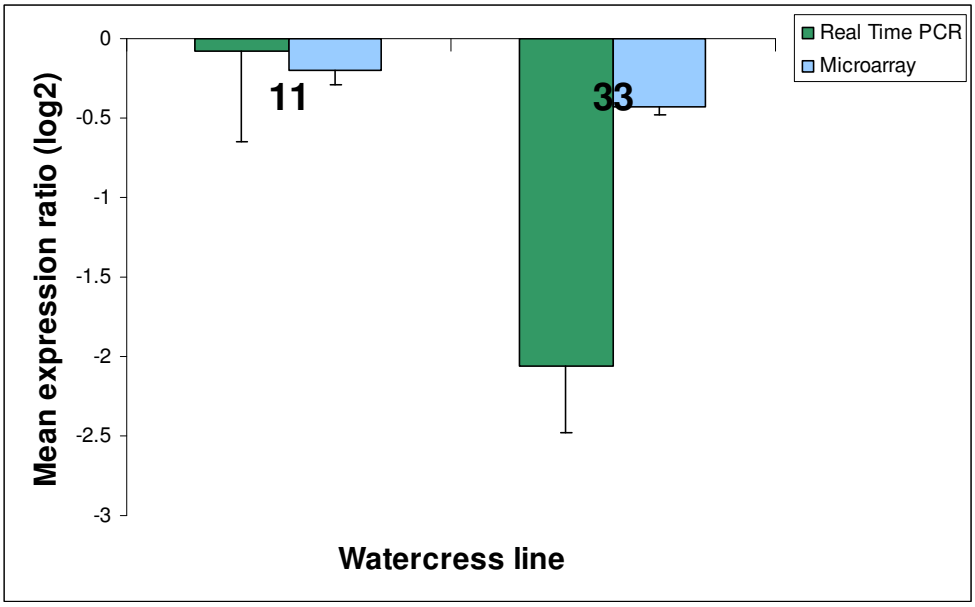


Figure 5.11: Normalised expression of Wx_0001, Wx_0011 and Wx_0033 compared to *Arabidopsis* ATH1 genechip.

It appears that the real time PCR shows a greater down-regulation of genes compared to the microarray analysis; however both show a consistency in down-regulation (Figure 5.12a and b).

(a)



(b)

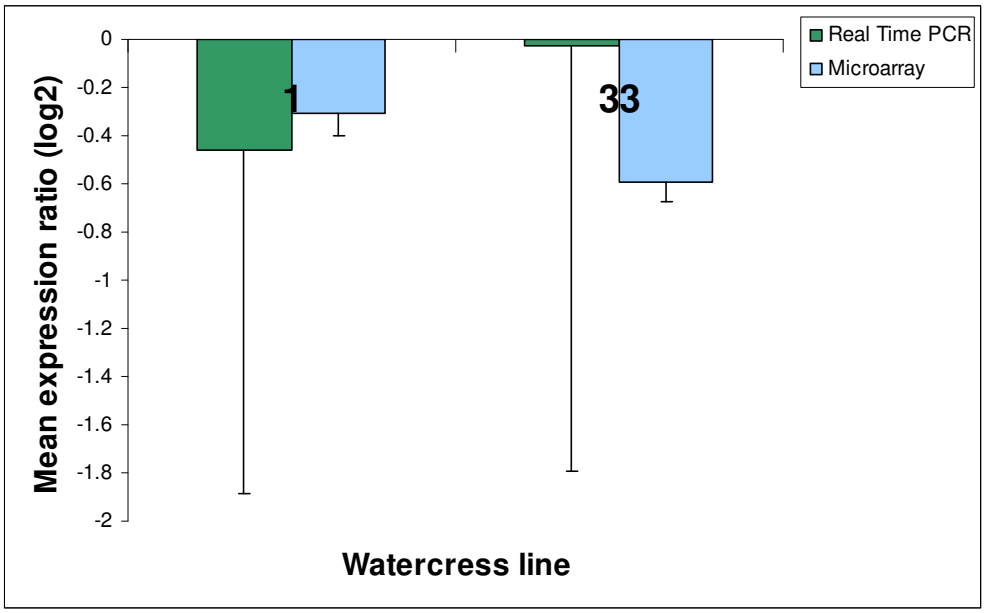


Figure 5.12: Comparison of gene expression between real time PCR analysis and microarray analysis (±SE) (a) gene expression of At1g65060 (4-CL) in Wx_0011 and Wx_0033 (b) gene expression of At4g35000 (ascorbate peroxidase) in Wx_0011 and Wx_0033.

From the data presented here microarray and qRT-PCR data clearly correlate but also that the correlation is not always apparent. There was a significant difference ($F_{1,16}$ =7.87, p<0.05) between Wx_0011 and Wx_0033 with regards to the level of expression of At1g65060 using ANOVA. The microarray results also reported a significant difference between Wx_0011 and Wx_0033 with a fold difference of 1.5 and a p value cut off at 0.1. No significant difference ($F_{1,16}$ =0.04, p>0.05) was observed between Wx_0001 and Wx_0033 with regards to the expression of At4g35000 however the microarray data reported a significant difference with a fold value of 1.5 between Wx_0001 and Wx_0033.

Although not a perfect correlation between real time PCR and microarray, the real time PCR results obtained which act as a quality control are strong enough to support the microarray data. This allows the global gene expression of watercress to be investigated further and provides a powerful tool for future watercress breeding improvement.

5.5 Discussion

This Chapter demonstrated a variation in gene expression between three selected watercress lines with respect to two genes one of which encodes an ascorbate peroxidase defending against hydrogen peroxide whilst the other encodes an isoform of 4-coumarate:CoA ligase (4CL) involved in the final steps of the phenylpropanoid pathway. The highest level of expression of the gene encoding 4CL was Wx_0001 followed by Wx_0011 and finally Wx_0033.Wx_0011 had the highest level of gene expression for the gene encoding ascorbate peroxidase followed by Wx_0001 and Wx_0033. Despite this, in general within each biological replicate there was a large degree of variation and even though overall there was a high level of confirmation with real time PCR there were some small discrepancies.

Initial investigations into the natural variation within the germplasm collection alluded to promising differences in stem length and diameter and also antioxidant power between the lines. Morphological and biochemical differences were therefore recorded, however any possible differences needed to be investigated. Microarray analysis and qRT-PCR were employed to reveal any underlying differences in gene expression. The reliability of real time PCR has not been questioned to date and is considered a rapid, specific and very sensitive way to detect a given nucleic acid target (Gachon *et al.*, 2004). Real time PCR was used to precisely distinguish and measure specific nucleic acid sequences in a sample, it can do so even if there a small quantity present within the sample (Valasek and Repa, 2005). During real time PCR the amplification process is monitored using fluorescent technology, it is how quickly the fluorescent signal reaches threshold intensity which enables quantification of the original target sequence (Valasek and Repa, 2005). Real time PCR utilises conventional PCR but is coupled to state of the art fluorescent chemistries and instrumentation to become what is known as real time PCR (Valasek and Repa, 2005). The DNA polymerases used in real time PCR can only use DNA as their template and therefore RNA cannot be amplified in the same way hence the enzyme reverse transcriptase is employed to produce complementary DNA (cDNA) which is used in the reaction (Valasek and Repa, 2005). The cDNA can be used to

determine changes in gene expression (i.e. RNA levels) (Valasek and Repa, 2005). SYBR green is used as the fluorescent probe, this binds to the minor groove which causes a 1,000 fold greater fluorescence compared to when it is in free solution (Valasek and Repa, 2005) thus allowing the quantification of any differences in gene expression. Genes which are not expected to change between the different experimental conditions are termed housekeeping genes and examples include cylophilin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ribosomal protein 36B4, ß-actin, 18S rRNA and transferring receptor (Valasek and Repa, 2005). In this study GAPDH and actin (ACT2) were chosen as the internal standards, it appeared that GAPDH had a more stable expression compared to ACT2.

Real time PCR is often used for the confirmation of results obtained from gene expression studies using microarray analysis (Gachon *et al.*, 2004). Arguably the reliability of microarray studies can be questioned due to the high number of multigene families, cross hybridisation between cDNA representatives of members of genes on cDNA based can therefore lead to the conclusion being interpreted from the result (Gachon *et al.*, 2004). One huge advantage over the questioned microarray is the fact that thousands of genes can be analysed in one step whilst real time is limited to fewer genes (Gachon *et al.*, 2004). Real time PCR can be used either to confirm microarray analysis (Gachon *et al.*, 2004; Klok *et al.*, 2002) or used to analyse the expression pattern of candidate genes further for fine tuned kinetics (Gachon *et al.*, 2004; Goda *et al.*, 2002). Goda *et al*. 2002 used both microarray and real time PCR to analyse the brassinosteroid-regulated genes in *Arabidopsis*.

In silico analysis can also be used for the analysis of microarray data whereby the array results are compared to information available in the literature and public databases (Chuaqui *et al.*, 2002). As mentioned above real time PCR is the most commonly used method for validating the gene expression from the whole genome profiling of microarray analysis (Morey *et al.*, 2006). However no standard definition of validation currently exists and correlation of real time PCR and microarray data are seldom presented in the literature and frequently non-agreeing data are rarely explained (Morey *et al.*, 2006). Variability in both biological and

technical procedures can have a great impact on both microarray and real time PCR data (Morey *et al.*, 2006). There are many pitfalls for both microarray and real time PCR (Morey *et al.*, 2006) but data normalisation fundamentally differs between the two. Microarray analysis requires global normalisation whilst real time PCR utilizes the expression of one or more reference genes against which all other gene expression is calibrated (Morey *et al.*, 2006). Morey *et al.*(2006) carried out a survey of the literature which revealed a wide range of correlations between microarray and real time PCR data and that the statistical analyses are rarely presented. It was concluded that the correlation between the two methods was affected by the direction of regulation and qPCR ct but not spot intensity, use of composite array data or the use of frozen tissue (Morey *et al.*, 2006). So it appears that there are discrepancies between microarray and real time PCR results and the two do not correlate all the time.

From the results presented in this Chapter a gene encoding 4CL (At1g65060) confirms microarray analysis whilst the other encoding an ascorbate perioxidase (At4g35000) confirms to a certain extent however Wx_0033 not generally following the microarray data. In nature the majority of phenolic compounds are produced via the phenylpropanoid pathway (Yu and Jez, 2008). These molecules produced can go on and have profound effects on plant growth and development (Yu and Jez, 2008). Phenylpropanoids are natural products which are derived from the amino acid L-phenylalanine via deamination by L-phenylalanine ammonia-lyase (PAL), more complicated phenylpropanoids are formed by condensation of phenylpropane unit with which the unit is derived from acetate (Dixon *et al.*, 2002). The phenylpropanoids represent the largest pool of secondary metabolites (Yu and Jez, 2008). The majority of the 'polyphenols' consumed in our diet are made up of phenylpropanoids. The phenylpropanoids can function as anitoxidants due to the multiple hydroxyl groups and unsaturated double bonds which react with radicals and oxidative ions in cells. Also some of the phenylpropanoids can act as phytoestrogens (isoflavones and coumestrols) or chemopreventive anti-cancer agents (reservatol) (Yu and Jez, 2008).

One of the genes of interest chosen for real time PCR was At1g65060 which encodes an isoform of 4-coumarate: CoA ligase (4CL) which is involved in the last step of the phenylpropanoid pathway.

The principle starting molecule in the phenylpropanoid pathway is phenylalanine and the deamination of phenylalanine is catalysed by phenyalalanine ammonia lyase (PAL) into trans-cinnamic acid. The aromatic acid is then oxidised by cinnamate 4-hydroxylase (C4H) and then a cytochrome P450 enzyme yields p-coumaric acid.4-coumaroyl CoA ligase (4CL) attaches a CoA molecule to *p*-coumaric acid generating 4-coumaryl CoA (Yu and Jez, 2008). The 4-coumaroyl CoA provides an active intermediate in multiple branches of the general phenylpropanoid pathway (Dixon *et al.*, 1996). The enzymes of the pathway; PAL, C4H and 4CL are highly conserved among plant species because they are important for the normal growth and development (Yu and Jez, 2008).

In Wx_0011 and Wx_0033 the gene encoding 4CL (At1g65060) was down-regulated in comparison to Wx_0001. At1g65060 is up-regulated in both Wx_0001 and Wx_0011 compared to Wx_0033. This reflects the FRAP and glucosinolate/isothiocyanate content of Wx_0001, Wx_0011 and Wx_0033 reported in earlier Chapter with Wx_0001 having one of the highest FRAP readings thus one of the highest overall antioxidant power, highest HPLC which indicates a high concentration of phenethyl glucosinolate and highest GC/MS reflecting a high concentration of phenethyl isothiocyanate.

Many metabolic processes produce active oxygen species (AOS), to add to this biotic and abiotic stress can give rise to excess concentrations of AOS (Shigeoka *et al.*, 2002). Ascorbate peroxidise (APX) exists as an isoenzyme and plays an important role in the metabolism of hydrogen peroxide (H_2O_2) (Shigeoka *et al.*, 2002). Ascorbate peroxidase can be seen to be distributed in at least four distinct cellular compartments (Shigeoka *et al.*, 2002). Hydrogen peroxide can function as a secondary messenger hence regulates expression of some antioxidative enzymes in

plant cells. The balance between the detrimental and beneficial roles of hydrogen peroxide can be determined by the local concentration of hydrogen peroxide with regards to the function of the organelle and the state of cellular scavengers which include asorbate peroxidase, catalase, ascorbate and glutathione (Shigeoka *et al.*, 2002). Therefore hydrogen peroxide can function as a normal metabolite in both plants and animals without being particularly cytotoxic however if left unchecked it can give rise to harmful free radicals from which the cells must be protected (Halliwell, 1974).

Ascorbate peroxidase (APX) utilises ascorbate (AsA) as its specific electron donor to reduce H_2O_2 to water (Shigeoka *et al.*, 2002). Ascorbate peroxidase has a high specificity for ascorbate as the electron donor (Shigeoka *et al.*, 2002). In plants ascorbate peroxidase in combination with the ascobate-glutathione cycle can function together to prevent the level of H_2O_2 building up to toxic levels (Shigeoka *et al.*, 2002). Photosynthetic organisms have developed various sophisticated AOS-scavenging systems which include the APX isoenzymes (Shigeoka *et al.*, 2002). It is vital for the plant to protect against free radicals e.g. superoxide (O_2^-), hydroxyl radicals, hydrogen peroxide and singlet oxygen (Noctor and Foyer, 1998). The antioxidants and antioxidant enzymes function to interrupt the cascades of uncontrolled oxidation and Table 9 shows the few classes of antioxidant enzymes which have been well characterised (Noctor and Foyer, 1998)

Table 9: Classes of antioxidant enzymes which have been characterized (Noctor and Foyer, 1998)

Enzyme	Abbreviation
Superoxide dismutase	SOD
Ascorbate peroxidase	APX
Monodehydroascorbate reductase	MDHAR
Dehydroascorbate reductase	DHAR
Glutathione reductase	GR
Catalase	CAT
Glutathione peroxidase	GPX
Guaiacol-type peroxidises	
Glutathion S-transferase	GST

At4g35000 encodes a microsomal ascorbate peroxidase, APX3. Other APX isoenzymes include stromal APX, thylakoid membrane bound APX in chloroplast and cytosolic APX (Shigeoka *et al.*, 2002). At4g35000 is upregulated in both Wx_0011 and Wx_0033 compared to Wx_0001. Wx_0011 is also upregulated when compared to Wx_0033. Wx_0001 is down-regulated when compared to both Wx_0011 and Wx_0033. This indicates that both Wx_0011 and Wx_0033 have an efficient in built antioxidant defence system to protect against the various free radicals encountered day to day which reflect the FRAP values recorded in Chapter three.

Unfortunately any differences in expression of At1g78440 and At1g80340 could not be visualised due to the real time PCR failing. At1g78440 encodes a gibberellins 2-oxidase which acts on C19 gibberellins therefore stops the production of gibberellins and At1g80340 encodes a protein with gibberellin 3 beta-hydroxylase activity which has been shown to catalyse the hydroxylation of both GA9 and GA20 therefore involved in the last steps of gibberellin production. These were of particular interest to see the expression due to the presence of the dwarf line Wx_0033 within the collection. Gibberellin 2 oxidase cause GA_1 deactivation therefore we would expect a higher level of expression in the dwarf phenotype. Gibberellin 3 beta hydroxylase catalases the final step in which active GA is synthesised in plants. Itoh *et al.* (2002) explored the possibility of modifying the height of rice plants by suppressing height controlling genes. A semi dwarf rice was produced by producing transgenic rice

with antisense of D18 cDNA (height controlling gene), it was concluded that successful manipulation of height could be achieved by manipulation of GA 3 beta hydroxylase expression through decrease in active GA level (Itoh *et al.*, 2002). We would expect a lower level of expression of gibberellins 3 beta hydroxylase in Wx_0033 to result in the dwarf phenotype.

In conclusion there was an intriguing variation in the level of gene expression between the three different watercress lines (Wx_0001, Wx_0011 and Wx_0033). There is still plenty of scope for further genes to be analysed which have been highlighted in the table produced from the volcano plot of significant genes and in particular genes related to the control of dwarfism. It would also be interesting to grow the lines in the field and examine how the gene expression changes to a fluctuating environment. It can be clearly seen from this Chapter that in Wx_0001 the final step in the phenylpropanoid pathway is up-regulated which is also reinforced in the gene expression analysis from the microarrays and this may contribute to its overall high antioxidant power. This Chapter has demonstrated the ability to detect any differences in the level of gene expression between the three of the watercress lines and this provides a valuable tool to further investigate the lines held within the germplasm collection.

5.5.1 Conclusion

1. Watercress RNA was successfully hybridised to the *Arabidopsis* ATH1 GeneChip.
2. Global gene expression of genes within watercress was observed and confirmed using Real Time PCR.
3. Genes of interest can now be identified and used in future breeding programmes.

CHAPTER SIX: FINAL DISCUSSION

6.1 Project overview

Watercress undoubtedly has a large number of nutritional benefits yet very little breeding has taken place to utilise these advantageous traits. Despite strict regulations in the use of plant products, medicines still contain phytochemicals which are estimated to be valued at US$22,608 million in 1997 and expected to reach a value of US$30,688.5 million in 2002, prescription products and over-the counter (OTC) herbal remedies comprise of 50% of the market (Business communications company Inc (1998) plant-derived drugs: products, technologies and applications study rb-121**,** cited in (Raskin *et al.*, 2002)).

At the dawn of the 21st century 11% of the 252 drugs classified as basic and essential by the World Health Organization were of flowering plant origin (Rates, 2001). In 1897 Friedrich Bayer & Co introduced synthetic acetyl salicylic acid (aspirin), aspirin being the safer synthetic analogue of salicylic acid which is the active ingredient in willow bark (Raskin *et al.*, 2002), this is one example where a pill provides a better option than the natural product. The twentieth century was predominately the synthetic-chemistry-dominated pharmaceutical industry, which replaced natural extracts with synthetic molecules that had no connection to natural products, hence the loss of 'an apple a day keeps the doctor away' (Raskin *et al.*, 2002). When the complex extract matrices which are derived from plant origin display activity in a biological screen there is a need to isolate and characterise these active principle(s) (Cordell, 2000). Natural products were viewed as templates for structure optimisation programs to make perfect new drugs (Raskin *et al.*, 2002) as it can often take up to 6 months to isolate and naturally characterise a natural plant product form a plant extract (Raskin *et al.*, 2002). It is estimated that a chemist can synthesise 200-300 compounds per year or possibly isolate and characterise 100-150 natural products per year (Cordell, 2000). Sadly of the 25,000-500,000 plant species only a small percentage have been investigated phytochemically and an even smaller percentage studied in terms of pharmalogical properties (Rates, 2001).

It is the plant based foods which can offer a diverse mixture of chemical constituents, some of which are essential for human life but also some which can promote good health (Grusak, 2002). Dietary supplements do not contain the balanced combination of phytochemicals found in fruits and vegetables, so by consuming foods originating from plants many diverse types of phytochemicals in various quantities can be consumed which is unlikely to result in toxicity (Liu, 2004). The additive and synergistic effects of phytochemicals can partially explain why no single antioxidant can replace the overall combination of natural phytochemicals in many fruit and vegetables (Liu, 2004). It reinforces the concept that the benefits obtained from fruit and vegetables cannot simply be mimicked by pills or tablets (Liu, 2004) and therefore research and improvement of cultivars is essential.

Although conventional breeding is the preferred route there is concern that the level of diversity within germplasm collections is not sufficient, this will then limit the extent of phytochemical improvement (Grusak, 2002). Transgenic methods are under consideration to enable effectiveness and significant increases in phytochemical content (Grusak, 2002), however public approval is yet to be gained and until that day any improvement in commonly consumed plant products needs to be achieved by conventional breeding hence the establishment of the watercress germplasm collection.

The first step in the watercress breeding programme was to establish a watercress germplasm collection which could potentially represent a wide diversity both genetically and in natural environmental conditions i.e. source watercress seeds from around the world. The next step would involve a simple screening of the collection investigating any potential phenotypic differences and an overall ranking order for differences in antioxidant power. Upon observing variation in morphological and biochemical traits the underlying genetic variation needs to be identified and genetic diversity examined. This would give the underlying fundamentals to a breeding programme.

6.2 Morphological variation within the watercress germplasm collection

In Chapter two the University of Southampton watercress germplasm collection was established and the different watercress lines with successful germination were grown in a controlled environment where temperature and day length could be controlled and any phenotypic differences would be resultant of differences in the genotype and not environmentally determined. The three morphological traits chosen for investigation in Chapter two were stem length, stem diameter and number of leaves and the unusual dwarf phenotype of Wx_0033 was observed and in Chapter three the high antioxidant power of Wx_0033 was noted. The morphological growth trial was extended to the field, Spetisbury Dorset, examining the growth of the watercress lines in watercress beds. The differences in growth could then be compared between the two conditions. Overall the mean stem length is higher in the watercress lines grown in the controlled environment. It may be due to environmental stresses (e.g. wind, temperature) that the watercress remains close to the water hence has a lower stem length and could explain the high concentration of antioxidants present in the lines when grown in the field. The stem diameter is more variable between the controlled environment and field with some lines having a higher stem diameter in the controlled environment whilst others in the field. There is no overall pattern of whether the stem diameter is higher in the controlled environment or field and the number of leaves is very similar whereby there is a mix of whether the controlled environment or field have the higher number of leaves. Also by extending the observation of morphological traits into the field it was possible to see if the growth in the field reflected that of the controlled environment. It was important to examine the growth of the watercress in the field as the watercress is traditionally grown in watercress beds. This chapter offered initial investigation into the phenotype of the watercress lines held within the collection which leads onto the investigation of the biochemistry with respect to antioxidants demonstrated in the following chapter.

6.3 Antioxidant variation within the watercress germplasm collection

Chapter two addressed the phenotypic variation present in the germplasm collection which left the biochemical variation to be addressed. A simple way to enhance people's nutritional diet is by modifying what is consumed day to day. Without venturing into transgenic methods any potential natural variation can be taken advantage of by selective breeding. A simple assay, FRAP, used to measure antioxidant power was used to give a ranking order (Wx_0038 the lowest whilst Wx_0033 highest) for the antioxidant power of the different watercress lines in Chapter three. FRAP was used to give an overall estimate and to indicate any significant variation between the lines, it could not detect individual antioxidants. The watercress lines were grown in a controlled environment, controlling for light and temperature, and a significant variation in overall antioxidant power between the lines was recorded. Antioxidant power of the lines established in the field, Spetisbury, was also recorded. A dramatic increase in antioxidant power was recorded for the watercress lines which were grown in the watercress bed which added to the reduced stem length of the lines in the field observed in Chapter two. This increase may have been due to the fluctuating environmental conditions and the presence of herbivores hence a more stressful environment. Therefore it is evident that the many environmental fluctuations experienced by the watercress lines in the field in Spetisbury lead to a reduced stem length (Chapter two) and increased concentration of antioxidants (Chapter three) which are two favourable traits of which my research is based upon.

It is important when demonstrating the antioxidant power of foods to have more than one assay to support the findings, in Chapter three FRAP and ORAC were used to demonstrate the antioxidant power of three different salad crops (watercress, spinach and rocket). A clear correlation (r^2=0.9808) was seen in the FRAP and ORAC values recorded for each of the salad crops and undoubtedly watercress had both the highest FRAP and ORAC values.

6.4 Agronomic improvement to the crop

As well as exploring the natural variation, which was demonstrated in Chapters two and three, the possible ways to improve the crop agronomically needed to be explored. This involved looking at the harvest number, different sections of the watercress plant, shelf life and also a comparison of watercress to spinach and rocket. No significant difference was seen in the concentration of antioxidants between the 1st, 2nd, 3rd and 4th harvest. One would believe that due to the repeated harvesting this would impose stress on the plant hence an increase in the concentration of antioxidants to protect against the external stress however our results indicated that the plant did not respond to the repeated harvest with four week intervals. One clear aim in plant breeding is to produce a crop with a reduced stem length however, if it is our intention to produce a plant with a reduced stem length but a high concentration of antioxidants it is important to determine the concentration of antioxidants proportioned to the stem. Results indicated that the stems do in fact have the lowest concentration of antioxidants and the leaves the highest therefore a dwarf leafy watercress plant should contain a sufficiently high concentration of antioxidants. A higher concentration of antioxidants were recorded in purchased watercress after six days storage in the fridge at 4℃ compared to the first day in which the watercress was stored. The impact of UV lights in the supermarkets is another possible route to explore in terms of affecting the concentration of antioxidants. Finally a comparison of the antioxidant power to both spinach and rocket needed to be carried out in order to demonstrate the high antioxidant concentration compared to two other popular salad accompaniments. A clear higher concentration of antioxidants was recorded in watercress 562mmol Fe^{2+} equivalent per gram fresh weight compared to 275mmol Fe^{2+} equivalent per gram fresh weight and 169mmo Fe^{2+} equivalent per gram fresh weight, spinach and rocket respectively.

It is often hypothesised that organic crops contain a higher concentration of antioxidants due to being more vulnerable to pest attack as a result of surviving without the use of pesticides/fungicides or any form of chemical. However upon testing no obvious difference in the concentration of antioxidants was recorded between conventionally and organically grown watercress, spinach and rocket.

The use of more than one antioxidant assay to validate results on the concentration of antioxidants contained in the food product is required. The FRAP assay measured the formation of an iron complex due to the presence of antioxidants and this assay correlated with the ORAC assay which measured the disappearance of a free radical due to the presence of antioxidants. The FRAP assay produced a strong blue colour with a high concentration of antioxidants whilst the ORAC showed a reduction in the purple colour with a high concentration of antioxidants.

6.5 The deduction of glucosinolate and isothiocyante concentrations

The antioxidant assays used in Chapter three can however only measure the overall antioxidant power and not deduce individual components therefore two new techniques were developed in Chapter three. High Performance Liquid Chromatography (HPLC) was used to give a measurement of the overall glucosinolate concentration in Wx_0001 (standard line), Wx_0011 (high antioxidant line), Wx_0033 (high antioxidant line) and Wx_0038 (low antioxidant line). Gas chromatography/Mass spectrometry can deduce individual isothiocyanate concentration and was used to measure the concentration of PEITC in the lines selected. Wx_0033 had one of the lowest concentrations of the precursory glucosionlate, phenethyl glucosionolate yet one of the highest concentrations of the isothiocyanate, PEITC implying a very efficient conversion system. The standard line, Wx_0001 had the highest concentration of PEITC and based unfortunately on only one biological replicate the highest concentration of phenethyl glucosionlate and Wx_0038 was consistently low in the glucosinolate prescursor and the isothiocyanate.

These results are promising indicating that there is variability in the antioxidant power of the individual lines thus enabling manipulation for an enhanced cultivar. It is particularly exciting to deduce the concentration of variation in PEITC, currently a popular plant derived product in the anti-cancer research area. Indeed, the greatest impact of plant derived drugs is in the anti-tumour area with the use of taxol, vinblastine, vincristine and camptothecin which have dramatically improved the

effectiveness of chemotherapy against some of the deadliest cancers (Raskin *et al.*, 2002). The bioassay results presented in Chapter three reflect variation in the ability of a few selected lines to kill breast cancer cells. Could watercress potentially be harvested as a rich source of PEITC and as another invaluable tool against cancer? Could PEITC be added to the list of positive plant derived products which assist in chemotherapy? The last twenty years have revealed isothiocyanates as chemopreventative agents. Isothiocyanates also exhibit antitumour activity and target multiple pathways including apoptosis, the MAPK pathway, oxidative stress and the cell cycle machinery (Wu *et al*., 2009). Sulforaphane (SFN), phenethyl isothiocyanate (PEITC) and benzyl isothiocyanate (BITC) are among the isothiocyanates demonstrated to be highly effective at preventing or reducing the risk of cancer induced by carcinogenesis (Wu *et al*., 2009). Chapter three outlined the variation in the concentration of PEITC in the watercress lines which is a promising outcome of the watercress research carried out.

6.6 Genetic diversity within the watercress germplasm collection

Chapters two and three provided new data for the morphological and biochemical of the watercress lines held within the germplasm collection. However the underlying genetics still remained to be explored (Chapters four and five). The essence of variation undoubtedly relies upon genetic diversity; therefore Amplified Fragment Length Polymorphism (AFLP) were employed in Chapter four to deduce the level of genetic diversity held within the watercress germplasm collection as Chapters two and three inferred differences in morphology and chemical composition. Eight primer pairs were used but only five gave informative amplification. The AFLP revealed that the genetic diversity was mainly proportioned to within each line which aids in the explanantion of the variation with their standard error bars presented in Chapters two and three . One primer pair did infer that the genetic diversity lie between the different watercress lines rather than within each line. A greater number of primer pairs are required to quantify any definitive genetic diversity within the collection. However with the current technology the next step instead of AFLPs

would be the use of 454 sequencing where the whole genome can be rapidly sequenced (as mentioned below in future areas of research).

6.7 Gene expression within the watercress germplasm collection

Differences apparent in phenotype demonstrated in Chapter two and biochemical composition demonstrated in Chapter three could be due to underlying differences in gene expression. Chapter four delved into the genetic diversity of the watercress collection whilst Chapter five focused on individual gene expression levels. Microarray studies are used for whole genome profiling whilst real time PCR can be used to validate these results by selecting genes of interest and comparing them to control genes, this was the focus of Chapter five. Of particular interest was the expression of genes relating to antioxidants hence changes in concentrations of antioxidants and also GA oxidases which are involved in the control of plant height. Much work has been carried out manipulating plant height with success due to discovery of the GA oxidases. Real time PCR supported expression levels reported from microarray analysis; therefore further genes of interest can be extrapolated from the microarray analysis and investigated further. Ideally the dwarf phenotype could be explained by the GA oxidase genes and genes of interest identified for the resulting high antioxidant power. Chapter five antioxidant gene expression confirmed the antioxidant assay results, unfortunately the genes relating to GA oxidases did not amplify successfully but the results obtained from the successful gene amplication provided sufficient data for confirmation of the microarray.

6.8 Contributions to watercress breeding through this project

This project initially established an overall sense of the degree of variation held within this watercress germplasm collection by assessing morphological and biochemical (Chapters two and three) traits. It was deemed there was visible variation in the collection therefore the level of the genetic involvement was investigated by looking at the genetic diversity (Chapter four) and differences in

gene expression (Chapter Five). The data obtained in Chapters two and three determined specific lines to be selected for global gene expression in Chapter five.

It was essential to determine the level of genetic diversity present in the collection as this allowed variation to be partitioned to either 'within' or 'between' line variation. The scores analysed from the AFLPs resulted in a larger proportion of varition being assigned to 'within' line variation which is evident by the large degree of standard error present on the morphological and biochemical traits assessed. This gives a positive scope to future breeding in this crop as the wide genetic diversity has not of yet been severly narrowed offering many possibilities for breeding. In order to find an explanation for the variation observed a global gene expression analysis was performed (Chapter five) using the ATH1 genechip. Genes of interest which were believed to be responsible for the variaiton in antioxidant status and stem height (gibberellins) were investigated in order to determine an up-regulation or down-regulation in specific lines (Wx_0001, Wx_0011 and Wx_0033) which thereby may explain some of the observed variation. This research has initaited steps in the molecular breeding of watercress and has provided some fundamental underlying information. Information on morphological and biochemical traits has been given for future selection and a breeding programme initiated which can be taken forward. Global gene expression has been established by hybridisation of watercress RNA to the *Arabidopsis* ATH1 genechip providing a vital tool to watercress research. It is also evident from this research that there is currently a wide genetic diversity of material available in which breeding programmes can take adavantage of.

6.9 Future areas of research

This research has quantified phenethyl glucosinolate and PEITC present in four lines (Wx_0001, Wx_0011, Wx_0033 and Wx_0038). Ideally this should be expanded to the remaining lines held in the collection to examine any further variation. Also the GC/MS assay can be further manipulated via temperature and pH in order to ensure maximum release of PEITC and optimisation of the protocol. In terms of morphological research this can be progressed by investigating root morphology and

any differences between the roots in the watercress lines which could be of use in research connected to phosphate removal. Potentially the larger the root networks the higher amount of phosphate which can be extracted from the ground. This could aid in soil toxicity problems and processes such as phytoremediation. The collection does contain variation between the lines but in order to emphasise this variation watercress needs to be collected from extreme environments and added to this promising collection. As stated it is the wild relatives in which beneficial genes can be obtained.

Crosses have been developed throughout the research but are still continuing to reach the ninth generation. Two lines of crosses have been developed one which has the aim of producing a line with an enhanced concentration of antioxidants whilst the other was based on two extremes (in terms of antioxidant values) in order to produce a mapping population for elucidation of QTL through the development of a molecular genetic map. Upon reaching the ninth generation the line produced for enhanced antioxidants will need to be screened through FRAP comparing to the standard line and parents to confirm if the initial cross was successful. Detailed genetic linkage maps provide an essential tool for selection, identification and organisation of plant genomes (Landry *et al.*, 1987). Construction of a linkage map requires a segregating plant population (i.e. a population derived from sexual reproduction) (Collard *et al*., 2005). The parents selected for this mapping population will differ for one or more traits of interest (Collard *et al*., 2005), in this research the parents in their antioxidant power i.e. Wx_0011 high antioxidant power whilst Wx_0038 low antioxidant power. Two parent plants chosen are crossed which gives rise to an F_1 generation which self pollinate to give rise to an F_2 generation. Inbreeding from the F_2 plants allows the construction of recombinant inbred (RI) lines which consist of a series of homozygous lines each containing a unique combination of chromosomal segments from the original parents (Collard *et al*., 2005), however usually six to eight generations are required. After the breeding population is established DNA markers need to be identified which reveal differences between the parents, termed polymorphic markers (Collard *et al*., 2005). Upon identifying polymorphic markers the whole population, including the parents, need to be screened. The final step involves coding data for each individual of a population and conducting linkage analysis using computer programs (Collard *et al*.

2005). Therefore initial steps have already been taken for construction of a linkage map for watercress.

A paper recently published looked at a genomics based strategy for quantitative trait breeding in tomato. The new genome-assisted breeding approach relies on the identification of genetic mechanisms controlling the trait of interest through functional genomics tools (Barone *et al*. 2009). With this approach, polymorphisms in major gene sequences which are responsible for variability in the expression of the trait under study are exploited for tracking favourable allele combinations in breeding programmes using high throughput genomic technologies (Barone *et al*. 2009). In the tomato breeding programme the introgression lines (ILs) were grown and then screened for fruit antioxidant content for QTL identification and then microarray analysis carried out on differentially performing ILs. Real time PCR is then used for candidate gene validation and functional validation is performed through mutagenesis or transformation. Finally molecular markers are used for assisted breeding (Barone *et al*. 2009).

With the advent of rapidly progressing genetics it is reasonable to consider sequencing of the watercress genome in order to advance further research with this crop. Since the first introduction of next generation sequencing in 2005 these technologies have had a huge impact on research (Morozova and Marra, 2008). Further unprecedented opportunities have been provided for high throughput sequencing technologies from Applied Biosystems, Roche and Illumina (Morozova and Marra, 2008). DNA sequencing is in a rapid period of change, a new generation of instruments which utilise primed synthesis in flow cells to obtain sequence of millions of different DNA templates (Table 10) (Holt and Jones, 2008). There is a continuous improvement in this field by replacement of these new generation of instruments with even more powerful technology as the demand for DNA sequences continues (Holt and Jones, 2008).

Table 10: DNA sequencing technologies (table adapted from Holt and Jones, 2008)

Platform	Method	Instrument configuration	Throughput statistic	Data per run (Gbp)
454 GS-FLX	Pyrosequencing	Single picotiter plate, partitionable into 8 lanes	238bp read	0.1
Illumina	Four-colour SBS with reversible terminators	Single flow cell, partitionable into 8 lanes	35-bp read	1.3
ABI SOLiD	Oligonucleotide ligation with two-base, four colour encoding	Independently controlled dual-flow cells, each partitionable into 8 lanes	35-bp reads, mapped to reference sequence allowing up to three mismatches	4
Helicos HeliScope	Single-colour SBS with virtual terminators	Single 25-lane flow cell	30-bp read	7.5

Using next generation sequencing technologies gives the possibility to resequence entire plant genomes more efficiently and in greater depth than ever before (Varshney *et al*., 2009). The detection and exploitation of genetic variation always has been an integral part of plant breeding. The development of improved cultivars can be accelerated by improving the efficiency of predicting phenotypes from genotypes (Varshney *et al*., 2009). Next generation technologies are expected to advance crop genetics and breeding which would result in crop improvement in the large scale development of molecular markers, association mapping, wide crosses and alien introgression, transcript profiling and population genetics (Varshney *et al*., 2009). It is predicted that there will be an exponential increase in the use of next generation sequencing technologies, and one important take home message is that it will not only be the major crop species which benefit but also the so-called orphan crops e.g. watercress. This will then have a profound positive effect on crop breeding for the future (Varshney *et al*., 2009). The next generation sequencing techniques have super-seeded techniques such as AFLPs, although AFLPs provide an excellent starting point. Therefore a huge step in watercress research and an ultimate goal would be to take adavantage of these rapidly progressing techniques and sequence watercress. This would rapidly progress the watercress research.

The advances in sequencing have allowed for large amounts of data to be collected in a fast and cost-effective way which appears to be very useful for re-sequencing studies of complex organisms or de novo genomic sequencing of organisms with

small genomes (Geraldes *et al*., 2011). Re-sequencing was carried out on the developing poplar xylem transcriptome to allow extensive documentation of the levels and patterns of nucleotide variability which would allow a catalogue to be produced to facilitate tree breeding programme (Geraldes *et al*., 2011). The result of this study was the empirical support for the existence of previously undiscovered genes and adds to the 1,241,214 SNPs discovered when *Populus trichocarpa* Nisqually-1 was sequenced (Geraldes *et al*., 2011).

6.10 Overall summary

This new project has established a leading global germplasm collection at the University of Southampton. This collecion been screened for morphological (Chapter two) and biochemical (Chapter three) traits – in particular a reduced stem length and high antioxidant potential and PEITC content. Large differences were observed in these traits for seven week old plants and individuals were identified for further analysis.

Field trials were conducted alongside a controlled environment for morphological and biochemical analysis. HPLC and GC/MS assays were developed for glucosinolate and isothiocyanate detection (Chapter three) and the use of a bioassay using breast cancer cell lines to determine anticancer potential. Crosses were developed in order to produce a molecular map for the elucidation of QTL and to advance watercress breeding for an enhanced cultivar.Gene expression using an AFFYMETRIX chip was confirmed with real time PCR therefore providing information on the global gene expression of watercress (Chapter five). The genetic diversity within the collection was assessed using AFLP (Chapter four),

This research indicates that to achieve a watercress cultivar with a reduced stem length and high concentration of antioxidants the cultivar needs to be grown in the field; stem length in Chapter two was lower in the field and the concentrations of antioxidants presented in Chapter three were higher in the field. Ultimately by growing the cultivar in the field with many external fluctuations can lead to an

increase in the plant defensive compounds glucosinolates, which in turn lead to an increase in the beneficial isothiocynates. Our results suggest that significant potential exists for future improvement and selection in this crop, potentially with many exciting possibilities for combating cancer with natural plant products.

This project has advanced watercress breeding not only by elucidating morphological variation but also antioxidant values. The nutritional aspect of this crop was investigated by looking at specific antioxidants (glucosinolates and isothiocycantes) and therefore contributed to a more in depth investigation into the antioxidant values. This project has also recorded the genetic diversity of watercress revealing that there is more variation within the lines than between suggesting a genetic bottleneck has not yet been reached in this crop thereby potential for enhancement in this crop using the material available. Finally a global gene expression of this crop is available for investigation of variation into gene expression between the lines which can be of use as a tool for nurtrional and phenotypic research; it has already shown up-regulation of secondary metabolism in some lines.

The work presented in this thesis presents a valuable source of information and data regarding the progression of all aspects of the genetics in watercress breeding. It provides an important base for future research in this field to build on, in particular molecular breeding.

REFERENCE LIST

Achard P, Liao L, Jiang C, Desnos T, Bartlett J, Fu X, Harberd NP. **2007.** DELLAs Contribute to Plant Photomorphogenesis. *Plant Physiology* **143:** 1163-1172.

Achard P, Vriezen WH, Van Der Straeten D, Harberd NP. **2003.** Ethylene Regulates Arabidopsis Development via the Modulation of DELLA Protein Growth Repressor Function. *THE PLANT CELL* **15:** 2816-2825.

Agrawal AA, Conner JK, Johnson MTJ, Wallsgrove R. **2010.** Ecological genetics of an induced plant defence against herbivores: additive genetic variance and costs of phenotypic plasticity. *Evolution: International Journal of Organic Evolution* **56:** 2206-2213.

Ahvenainen R. **1996.** New approaches in improving the shelf life of minimally processed fruit and vegetables. *Trends in Food Science & Technology* **7:** 179-187.

Alabadi D, Gil J, Blazquez MA, Garcia-Martinez JL. **2004.** Gibberellins Repress Photomorphogenesis in Darkness. *Plant Physiology* **134:** 1050-1057.

Ambrosone CB, McCann SE, Freudenheim JL, Marshall JR, Zhang Y, Shields PG. **2004.** Breast Cancer Risk in Premenopausal Women Is Inversely Associated with Consumption of Broccoli, a Source of Isothiocyanates, but Is Not Modified by GST Genotype. *Journal of Nutrition* **134:** 1134-1138.

Ariizumi T, Paulraj KL, Steber CM. 2011. The role of two F-box proteins, SLEEPY1 and SNEEZY in *Arabidopsis* gibberellin signalling. *Plant Physiology* **155:** 765-775.

Arnold D, Flegmann A, Clarkson J. **1995.** Somaclonal variation in watercress for resistance to crook root disease. *Plant Cell Reports* **14:** 241-244.

Aruoma OI, Murcia AM, Butler J, Halliwell B. **1993.** *Journal of Agricultural and Food Chemistry* **41:** 1880.

Aruoma O. **1998.** Free radicals, oxidative stress, and antioxidants in human health and disease. *Journal of the American Oil Chemists' Society* **75:** 199-212.

Bagnaresi P, Moschella A, Beretta O, Vitulli F, Ranalli P, Perata P. **2008.** Heterologous microarray experiments allow the identification of the early events associated with potato tuber cold sweetening. *BMC Genomics* **9:** 176.

Barillari J, Canistro D, Paolini M, Ferroni F, Pedulli GF, Iori R, Valgimigli L. **2005.** Direct Antioxidant Activity of Purified Glucoerucin, the Dietary Secondary Metabolite Contained in Rocket (Eruca sativa Mill.) Seeds and Sprouts. *Journal of Agricultural and Food Chemistry* **53:** 2475-2482.

Barkley A, Chumley FG. 2011. A Doubled Haploid Laboratory for Kansas Wheat Breeding: An Economic Analysis of Biotechnology Adoption. *heartlandplants.com*

Barone A, Di Matteo A, Carputo D, Frusciante L. 2009. High throughput genomics enhances tomato breeding efficiency. *Current Genomics* **10:** 1-9.

Becker J, Vos P, Kuiper M, Salamini F, Heun M. **1995.** Combined mapping of AFLP and RFLP markers in barley. *Molecular and General Genetics* **249:** 65-73.

Benzie IFF, Strain JJ. **1996.** The Ferric Reducing Ability of Plasma (FRAP) as a Measure of "Antioxidant Power": The FRAP Assay. *Analytical Biochemistry* **239:** 70-76.

Bhettacharya A, Kourmpetli S, Davey MR. 2010. Practical Applications of Manipulating Plant Architecture by regulating gibberellin metabolism. *Journal of Plant Growth Regulation* **29:** 249-256.

Blears MJ, De Grandis SA, Lee H, Trevors JT. **1998.** Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. *Journal of Industrial Microbiology and Biotechnology* **21:** 99-114.

Bohnert HJ, Nelson DE, Jensen RG. **1995.** Adaptations to Environmental Stresses. *THE PLANT CELL* **7:** 1099-1111.

Bones AM, Rossiter JT. **1996.** The myrosinase-glucosinolate system, its organisation and biochemistry. *Physiologia Plantarum* **97:** 194-208.

Bouis HE. **2007.** Enrichment of food staples through plant breeding: a new strategy for fighting micronutrient malnutrition. *Nutrition* **16:** 701-704.

Brown PD, Tokuhisa JG, Reichelt M, Gershenzon J. **2003.** Variation of glucosinolate accumulation among different organs and developmental stages of Arabidopsis thaliana. *Phytochemistry* **62:** 471-481.

Buchner R. **1987.** Glucosinolates in rapeseeds: analytical aspects. *World Crops: Production, Utilization, Description.Martinus Nijhoff Publishers, Kluwer Academic Publishers, Dordrecht, The Netherlands* 50-58.

Buerkle AC and Lexer C. 2008. Admixture as the basis for genetic mapping. *Trends in Ecology and Evolution* **23:** 686-694.

Busov V, Meilan R, Pearce D, Rood S, Ma C, Tschaplinski T, Strauss S. **2006.** Transgenic modification of gai or rgl1 causes dwarfing and alters gibberellins, root growth, and metabolite profiles in Populus. *Planta* **224:** 288-299.

Cao G, Verdon CP, Wu AHB, Wang H, Prior RL. **1995.** Automated assay of Oxygen Radical Absorbance Capacity with the COBAS FARA II. *Clin.Chem.* **41:** 1738.

Carr AC, Frei B. 1999. Toward a new recommended dietary allowance for vitamin C based on antioxidant and health effects in humans. *American Journal of Clinical Nutrition* **69:** 1086-1107.

Cavell AC, Lydiate DJ, Parkin IAP, Dean C, Trick M. **1998.** Collinearity between a 30-centimorgan segment of *Arbidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome. *Genome* **41:** 62-69.

Cheung KL, Ookhor T, Kong A. 2009. Synergistic effect of combination of phenethyl isothiothicyanate and sulforaphane or curcumin and sulforaphane in the inhibition of inflamation. *Pharmaceutical Research* **26:** 224-231.

Chiao JW, Wu H, Ramaswamy G, Conaway CC, Chung FL, Wang L, Liu D. **2004.** Ingestion of an isothiocyanate metabolite from cruciferous vegetables inhibits growth of human prostate cancer cell xenografts by apoptosis and cell cycle arrest. *Carcinogenesis* **25:** 1403-1408.

Chismar JD, Mondala T, Fox HS, Roberts E, Langford D, Masliah E, Salomon DR, Head SR. **2002.** Analysis of results varaibilty from high-density oligonucleotide arrays comparing same species and cross species hybridisations. *Biotechniques* **33:** 516-524.

Christian P and West KP Jr. 1998. Interactions between zinc and vitamin A: an update. *American Journal of Clinical Nutrition* **68:** 435s-441s.

Chu YF, Sun J, Wu X, Liu RH. **2002.** Antioxidant and Antiproliferative Activities of Common Vegetables. *Journal of Agricultural and Food Chemistry* **50:** 6910-6916.

Chuaqui RF, Bonner RF, Best CJM, Gillespie JW, Flaig MJ, Hewitt SM, Phillips JL, Krizman DB, Tangrea MA, Ahram M, Linehan M, Knezevic V, Emmert-Buck MR. **2002.** Post-analysis follow-up and validation of microarray experiments. *Nature Genetics* **32:** 509-514.

Chung FL, Morse MA, Eklind KI, Lewis J. **1992.** Quantitation of Human Uptake of the Anticarcinogen Phenethyl Isothiocyanate After A Watercress Meal. *Cancer Epidemiology Biomarkers & Prevention* **1:** 383-388.

Claxton J, Arnold D, Clarkson J, Blakesley D. **1998.** The regeneration and screening of watercress somaclones for resistance to Spongospora subterranea f. sp. nasturtii and measurement of somaclonal variation. *Plant Cell, Tissue and Organ Culture* **52:** 155-164.

Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* **142:** 169-196.

Cordell GA. **2000.** Biodiversity and drug discoveryá-- a symbiotic relationship. *Phytochemistry* **55:** 463-480.

Costain L. 2007. Watercress the original superfood. *Watercress Report*. and 'Superfoods for a super you nourish body, mind and soul'. *Booklet*.

Crisp DT. **1970.** Input and Output of Minerals for a Small Watercress Bed Fed by Chalk Water. *The Journal of Applied Ecology* **7:** 117-140.

Csokay B, Prajda N, Weber G, Olah E. **1997.** Molecular mechanisms in the antiproliferative action of quercetin. *Life Sciences* **60:** 2157-2163.

Czapski J. 2009. Cancer preventing properties of cruciferous vegetables. *Vegetable Crops Research Bulletin* **70:** 5-18.

D'Ambrosio SM. **2007.** Phytonutrients: A more natural approach toward cancer prevention. *Seminars in Cancer Biology* **17:** 345-346.

Danforth E Jr, Burger AG. 1989. The impact of nutrition on thyroid hormone physiology and action. *Annual Review of Nutrition* **9:** 201-227.

Danielak R, Borkowski B. **1969.** Biologically active compounds in seeds of crucifers Part III. Chromatographical search for glucosynolates. *Dissertions in Pharmacy and Pharmacology* **21:** 563-575.

Davey M, Graham N, Vanholme B, Swennen R, May S, Keulemans J. **2009.** Heterologous oligonucleotide microarrays for transcriptomics in a non-model species; a proof-of-concept study of drought stress in Musa. *BMC Genomics* **10:** 436.

Daxenbichler ME, Spencer GF, Carlson DG, Rose GB, Brinker AM, Powell RG. **1991.** Glucosinolate composition of seeds from 297 species of wild plants. *Phytochemistry* **30:** 2623-2638.

de Lucas M, Daviere JM, Rodriguez-Falcon M, Pontin M, Iglesias-Pedraz JM, Lorrain S, Fankhauser C, Blazquez MA, Titarenko E, Prat S. **2008.** A molecular framework for light and gibberellin control of cell elongation. *Nature* **451:** 480-U11.

de Quiros HC, Magrath R, McCallum D, Kroymann J, Scnabelrauch D, Mitchell-Olds T, Mithen R. **2000.** -Keto acid elongation and glucosinolate biosynthesis in Arabidopsis thaliana. *TAG Theoretical and Applied Genetics* **101:** 429-437.

Dijkstra C, Adams E, Bhattacharya A, Page AF, Anthony P, Kourmpetti S, Power JB, Lowe KC, Thomas SG, Hedden P, Phillips AL, Davey MR. 2008. Over-expression of a gibberellin 2-oxidase gene from *Phaseolus coccineus* L.enhances gibberellin inactivation and induces dwarfism in *Solanum* species. *Plant Cell Reports* **27:** 463-470.

Dixon RA, Achnine L, Kota P, Liu CJ, Reddy MSS, Wang L. 2002. The phenylpropanoid pathway and plant defenceΓÇöa genomics perspective. *Molecular Plant Pathology* **3:** 371-390.

Dixon RA, Lamb CJ, Masoud S, Sewalt VJH, Paiva NL. 1996. Metabolic engineering: prospects for crop improvement through the genetic manipulation of phenylpropanoid biosynthesis and defense responses -- a review. *Gene* **179:** 61-71.

Djakovic-Petrovic T, de Wit M, Voesenek LACJ, Pierik R. 2007. DELLA protein function in growth responses to canopy signals. *Plant Journal* **51:** 117-126.

Du Z, Zhou X, Ling Y, Zhang Z, Su Z. 2010. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research* **38:** W64-W70.

Dytham, C. 1999. Choosing and using statistics: A Biologists Guide. *Blackwell Science, USA*.

Edmeades GO, McMaster GS, White JW, Campos H. 2004. Genomics and the physiologist: bridging the gap between genes and crop response. *Field Crops Research* **90:** 5-18.

Ellis MH, Rebetzke GJ, Azanza F, Richards RA, Spielmeyer W. 2005. Molecular mapping of gibberellin-responsive dwarfing genes in bread wheat. *Theoretical and Applied Genetics* **111:** 423-430.

Engelen-Eigles G, Holden G, Cohen JD, Gardner G. 2006. The effect of temperature, photoperiod, and light quality on gluconasturtiin concentration in watercress (Nasturtium officinale R. Br.). *Journal of Agricultural and Food Chemistry* **54:** 328-334.

Erlund I. 2004. Review of the flavonoids quercetin, hesperetin, and naringenin. Dietary sources, bioactivities, bioavailability, and epidemiology. *Nutrition Research* **24:** 851-874.

Excoffier L, Smouse PE, Quattro JM. **1992.** Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics* **131:** 479-491.

Fahey JW, Zalcmann AT, Talalay P. **2001.** The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* **56:** 5-51.

Farnham MW, Stephenson KK, Fahey JW. **2000.** Capacity of Broccoli to Induce a Mammalian Chemoprotective Enzyme Varies among Inbred Lines. *Journal of the American Society for Horticultural Science* **125:** 482-488.

Fenwick GR, Heaney RK, Mullin WJ, VanEtten CH. **1982.** Glucosinolates and their breakdown products in food and food plants. *CRC Critical Reviews in Food Science and Nutrition* **18:** 123-201.

Flanagan NS. 2005. A Guide to GenAlEx, Genetic Analysis in Excel.

Fleet CM, Sun Tp. **2005.** A DELLAcate balance: the role of gibberellin in plant morphogenesis. *Current Opinion in Plant Biology* **8:** 77-85.

Folta KM, Pontin MA, Karlin-Neumann G, Bottini R, Spalding EP. **2003.** Genomic and physiological studies of early cryptochrome 1 action demonstrate roles for auxin and gibberellin in the control of hypocotyl growth by blue light. *The Plant Journal* **36:** 203-214.

Frary A, Doganlar S. **2003.** Comparative genetics of crop plant domestication amd evolution. *Turk J Agric For* **27:** 59-69.

Fridborg I, Kuusk S, Robertson M, Sundberg E. **2001.** The Arabidopsis Protein SHI Represses Gibberellin Responses in Arabidopsis and Barley. *Plant Physiology* **127:** 937-948.

Gachon C, Mingam Ac, Charrier B+. **2004.** Real-time PCR: what relevance to plant studies? *Journal of Experimental Botany* **55:** 1445-1454.

Gepts P. **2002.** A Comparison between Crop Domestication, Classical Plant Breeding, and Genetic Engineering. *Crop Science* **42:** 1780-1790.

Geraldes A, Pang J, Thiessen N, Cezard T, Moore R, Zhao Y, Tam A, Wang S, Friedmann M, Birol I, Jones SJM, Cronk QCB, Douglas, CJ. 2011. SNP discovery in black cottonwood (*Populus trichocarpa*) by population transcriptome resequencing. *Molecular Ecology Resources* **11:** 81-92.

Getahun SM, Chung FL. **1999.** Conversion of glucosinolates to isothiocyanates in humans after ingestion of cooked watercress. *Cancer Epidemiology Biomarkers & Prevention* **8:** 447-451.

Gil M, I, Ferreres F, Tomas-Barberan FA. **1999.** Effect of postharvest storage and processing on the antioxidant constituents (flavonoids and vitamin C) of fresh-cut spinach. *Journal of Agricultural and Food Chemistry* **47:** 2213.

Gil V, MacLeod AJ. **1980.** Degradation of glucosinolates of Nasturtium officinale seeds. *Phytochemistry* **19:** 1657-1660.

Gill CIR, Haldar S, Boyd LA, Bennett R, Whiteford J, Butler M, Pearson JR, Bradbury I, Rowland IR. **2007.** Watercress supplementation in diet reduces lymphocyte DNA damage and alters blood antioxidant status in healthy adults. *American Journal of Clinical Nutrition* **85:** 504-510.

Goda H, Shimada Y, Asami T, Fujioka S, Yoshida S. **2002.** Microarray Analysis of Brassinosteroid-Regulated Genes in Arabidopsis. *Plant Physiology* **130:** 1319-1334.

Goffman F, Becker H. **2002.** Genetic variation of tocopherol content in a germplasm collection of Brassica napus L. *Euphytica* **125:** 189-196.

Going B, Simpson J, Even T. **2008.** The influence of light on the growth of watercress (Nasturtium officinale R. Br.). *Hydrobiologia* **607:** 75-85.

Grusak MA. **2002.** Phytochemicals in plants: genomics-assisted plant improvement for nutritional and health benefits. *Current Opinion in Biotechnology* **13:** 508-511.

Gur A, Zamir D. **2004.** Unused natural variation can lift yield barriers in plant breeding. *PLo S Biol* **2:** 1610-1615.

Habegger RKMaFD. **1989.** A cultivation method for Nasturtium offinale (watercress) groen in a greenhouse. *Acta Horticulturae* **242:** 291-295.

Hajjar R, Hodgkin T. **2007.** The use of wild relatives in crop improvement: a survey of developments over the last 20-áyears. *Euphytica* **156:** 1-13.

Halkier BA, Gershenzon J. **2006.** BIOLOGY AND BIOCHEMISTRY OF GLUCOSINOLATES. *Annual Review of Plant Biology* **57:** 303-333.

Halliwell B. **1974.** Superoxide Dismutase, Catalase and Glutathione Peroxidase: Solutions to the Problems of Living with Oxygen. *New Phytologist* **73:** 1075-1086.

Halliwell B. **2007.** Dietary polyphenols: Good, bad, or indifferent for your health? *Cardiovascular Research* **73:** 341-347.

Halvorsen B, Holte K, Myhrstad MCW, Barikmo I, Hvattum E, Remberg SF, Wold A, Haffner K, Baugerod H, Andersen LF, Moskaug J, Jacobs DR, Blomhoff R. **2002.** A systematic screening of total antioxidants in dietary plants. *Journal of Nutrition* **132:** 461.

Hammond J, Broadley M, Craigon D, Higgins J, Emmerson Z, Townsend H, White P, May S. **2005.** Using genomic DNA-based probe-

selection to improve the sensitivity of high-density oligonucleotide arrays when applied to heterologous species. *Plant Methods* **1:** 10.

Hammond JP, Bowen HC, White PJ, Mills V, Pyke KA, Baker AJM, Whiting SN, May ST, Broadley MR. **2006.** A comparison of the Thlaspi caerulescens and Thlaspi arvense shoot transcriptomes. *New Phytologist* **170:** 239-260.

Hancock RD, McDougall GJ, Stewart D. 2007. Berry fruit as 'superfood'. Hope or hype. *Biologist* **54:** 73-79.

Harlan JR. **1975.** Our Vanishing Genetic Resources. *Science* **188:** 618-621.

Hecht SS, Chung FL, Richie JP, Akerkar SA, Borukhova A, Skowronski L, Carmella SG. **1995.** Effects of Watercress Consumption on Metabolism of A Tobacco-Specific Lung Carcinogen in Smokers. *Cancer Epidemiology Biomarkers & Prevention* **4:** 877-884.

Hecht SS. **1999.** Chemoprevention of Cancer by Isothiocyanates, Modifiers of Carcinogen Metabolism. *Journal of Nutrition* **129:** 768.

Hedden P, Phillips AL. **2000a.** Gibberellin metabolism: new insights revealed by the genes. *Trends in Plant Science* **5:** 523-530.

Hedden P, Phillips AL. **2000b.** Manipulation of hormone biosynthetic genes in transgenic plants. *Current Opinion in Biotechnology* **11:** 130-137.

Hedges LJ, Lister CE. 2005. Nutritional attributes of salad vegetables. 1-35.

Holt RA, Jones SJM. 2008. The new paradigm of flow cell sequencing. *Genome Research* **18:** 839-846.

Hooley R. **1994.** Gibberellins: perception, transduction and responses. *Plant Molecular Biology* **26:** 1529-1555.

Howard LR, Pandjaitan N, Morelock T, Gil MI. **2002.** Antioxidant Capacity and Phenolic Content of Spinach As Affected by Genetics and

Growing Season. *Journal of Agricultural and Food Chemistry* **50:** 5891-5896.

Huff DR, Peakall R, Smouse PE. **1993.** RAPD variation within and among natural populations of outcrossing buffalograss (Buchloe dactyloides (Nutt.) Engelm). *TAG Theoretical and Applied Genetics* **86:** 927-934.

Huskisson E, Maggini S, Ruf M. 2007. The influence of micronutrients on cognitive function and performance. *Journal of International Medical Research* **35:** 1-19.

Hussain A, Peng J. **2003.** DELLA Proteins and GA Signalling in Arabidopsis. *Journal of Plant Growth Regulation* **22:** 134-140.

Ikeda A, Ueguchi-Tanaka M, Sonoda Y, Kitano H, Koshioka M, Futsuhara Y, Matsuoka M, Yamaguchi J. **2001.** slender Rice, a Constitutive Gibberellin Response Mutant, Is Caused by a Null Mutation of the SLR1 Gene, an Ortholog of the Height-Regulating Gene GAI/RGA/RHT/D8. *THE PLANT CELL* **13:** 999-1010.

Ilich JZ, Kerstetter JE. 2000. Nutrition in bone health revisited: a story beyond calcium. *Journal of the American College of Nutrition* **19:** 201-227.

Ishida S, Fukazawa J, Yuasa T, Takahashi Y. **2004.** Involvement of 14-3-3 Signaling Protein Binding in the Functional Regulation of the Transcriptional Activator REPRESSION OF SHOOT GROWTH by Gibberellins. *THE PLANT CELL* **16:** 2641-2651.

Ishimaru K, Ono K, Kashiwagi T. **2004.** Identification of a new gene controlling plant height in rice using the candidate-gene strategy. *Planta* **218:** 388-395.

Ismail A, Marjan ZM, Foong CW. **2004.** Total antioxidant activity and phenolic content in selected vegetables. *Food Chemistry* **87:** 581-586.

Itoh H, Ueguchi-Tanaka M, Sakamoto T, Kayano T, Tanaka H, Ashikari M, Matsuoka M. **2002.** Modification of Rice Plant Height by Suppressing the Height-Controlling Gene, D18, in Rice subtitle_in_Japanese. *Breeding Science* **52:** 215-218.

Jackson SJT, Singletary KW. **2004.** Sulforaphane: a naturally occurring mammary carcinoma mitotic inhibitor, which disrupts tubulin polymerization. *Carcinogenesis* **25:** 219-227.

Jansen RC, Jansen J. **1990.** On the selection for specific genes by single seed descent. *Euphytica* **51:** 131-140.

Jeffery EH, Brown AF, Kurilich AC, Keck AS, Matusheski N, Klein BP, Juvik JA. **2003.** Variation in content of bioactive components in broccoli. *Journal of Food Composition and Analysis* **16:** 323-330.

Jeong J-H, An JY, Kwon YT, Li LY, Lee Y J. **2008.** Quercetin-induced ubiquitination and down regulation of Her-2/neu. *Journal of Cellular Biochemistry* **105:** 585-595.

Johnson HW, Bernard RL. **1962.** Soybean genetics and breeding. *Advances in agronomy* **14:** 149-221.

Kaeppler SM, Kaeppler HF, Rhee Y. **2000.** Epigenetic aspects of somaclonal variation in plants. *Plant Molecular Biology* **43:** 179-188.

Kenfield SA, Chang ST, Chan JM. **2007.** Diet and Lifestyle Interventions in Active Surveillance Patients with Favourable-Risk Prostate Cancer. *Current Treatment Options in Oncology* **8:** 173-196.

Kessler A, Baldwin IT. **2001.** Defensive Function of Herbivore-Induced Plant Volatile Emissions in Nature. *Science* **291:** 2141-2144.

Kim S, Ishii G. **2006.** Glucosinolate profiles in the seeds, leaves and roots of rocket salad (*Eruca sativa* Mill.) and anti-oxidative activities of intact plant powder and purified 4-methoxyglucobrassicin. *Soil Science and Plant Nutrition* **52:** 394-400.

Kim YS, Milner JA. **2005.** Targets for indole-3-carbinol in cancer prevention. *The Journal of Nutritional Biochemistry* **16:** 65-73.

Klok EJ, Wilson IW, Wilson D, Chapman SC, Ewing RM, Somerville SC, Peacock WJ, Dolferus R, Dennis ES. **2002.** Expression Profile Analysis of the Low-Oxygen Response in Arabidopsis Root Cultures. *The Plant Cell Online* **14:** 2481-2494.

Knott DR, Kumar J. **1975.** Comparison of Early Generation Yield Testing and a Single Seed Descent Procedure in Wheat Breeding. *Crop Science* **15:** 295-299.

Kong JM, Chia LS, Goh NK, Chia TF, Brouillard R. **2003.** Analysis and biological activities of anthocyanins. *Phytochemistry* **64:** 923-933.

Koornneef M, Stam P. **2001.** Changing Paradigms in Plant Breeding. *Plant Physiology* **125:** 156-159.

Kopsell DA, Barickman TC, Sams CE, McElroy JS. **2007.** Influence of Nitrogen and Sulfur on Biomass Production and Carotenoid and Glucosinolate Concentrations in Watercress (Nasturtium officinale R. Br.). *Journal of Agricultural and Food Chemistry* **55:** 10628-10634.

Kusaba S, Fukumoto M, Honda C, Yamaguchi I, Sakamoto T, Kano-Murakami Y. **1998.** Decreased GA1 Content Caused by the Overexpression of OSH1 Is Accompanied by Suppression of GA 20-Oxidase Gene Expression. *Plant Physiology* **117:** 1179-1184.

Kushad MM, Brown AF, Kurilich AC, Juvik JA, Klein BP, Wallig MA, Jeffery EH. **1999.** Variation of Glucosinolates in Vegetable Crops of Brassica oleracea. *Journal of Agricultural and Food Chemistry* **47:** 1541-1548.

Lopez-Galvez G, Peiser G, Nie X, Cantwell M. **1997.** Quality changes in packaged salad products during storage. *Zeitschrift Fur Lebensmitteluntersuchung und -Forschung A* **205:** 64-72.

Lako J, Trenerry VC, Wahlqvist M, Wattanapenpaiboon N, Sotheeswaran S, Premier R. **2007.** Phytochemical flavonols, carotenoids and the antioxidant properties of a wide selection of Fijian fruit, vegetables and other readily available foods. *Food Chemistry* **101:** 1727-1741.

Lan TH, Paterson AH. **2001.** Comparative mapping of QTLs determining the plant size of Brassica oleracea. *TAG Theoretical and Applied Genetics* **103:** 383-397.

Landry BS, Kesseli RV, Farrara B, Michelmore RW. **1987.** A genetic map of lettuce (Lactuca sativa L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. *Genetics* **116:** 331.

Lee M. **1998.** Genome projects and gene pools: New germplasm for plant breeding? *Proceedings of the National Academy of Sciences of the United States of America* **95:** 2001-2004.

Li H and Roossinck MJ. 2004. Genetic Bottlenecks Reduce Population Variation in an Experimental RNA virus population. *Journal of Virology* **78:** 10582-10587.

Li J. **2005.** Brassinosteroid signaling: from receptor kinases to transcription factors. *Current Opinion in Plant Biology* **8:** 526-531.

Lill JT, Marquis RJ. **2001.** The effects of leaf quality on herbivore performance and attack from natural enemies. *Oecologia* **126:** 418-428.

Liu R.H. **2002.** Supplement quick fix fails to deliver. *Food Technol.Int.* **1:** 71-72.

Liu RH. **2004.** Potential Synergy of Phytochemicals in Cancer Prevention: Mechanism of Action. *Journal of Nutrition* **134:** 3479S-3485.

Lo S, Yang S, Chen K, Hsing Y, Zeevaart J, Chen L, Yu S. 2008. A novel class of gibberellin 2-oxidases control semidwarfism, tillering and root development in rice. *Plant Cell* **20:** 2603-2618.

Lu Y, Huggins P, Bar-Joseph Z. **2009.** Cross species analysis of microarray expression data. *Bioinformatics* **25:** 1476-1483.

Lucas M, Daviere J-M, Rodriguez-Falcon M, Pontin M, Iglesias-Pedraz JM, Lorrain S, Frankhauser C, Blazquez MA, Titarenko E, Prat S. **2008.** A molecular framework for light and gibberellin control of cell elongation. *Nature* **451:** 480-486.

Mackill DJ, Zhang Z, Redona ED, Colowit PM. **1996.** Level of polymorphism and genetic mapping of AFLP markers in rice. *Genome* **39:** 969-977.

Magome H, Yamguchi S, Hanada A, Kamiya Y, Oda K. 2008. The DFF1 transcriptional activator upregulates expression of a gibberellin-deactivating gene, *GA2ox7*, under high salinity stress in *Arabidopsis. The Plant Journal* **56:** 613-626.

Martinez-Sanchez A, Llorach R, Gil MI, Ferreres F. **2007.** Identification of New Flavonoid Glycosides and Flavonoid Profiles To Characterize Rocket Leafy Salads (Eruca vesicaria and Diplotaxis tenuifolia). *Journal of Agricultural and Food Chemistry* **55:** 1356-1363.

McCouch S. **2004.** Diversifying Selection in Plant Breeding. *PLoS Biol* **2:** e347.

McGinnis KM, Thomas SG, Soule JD, Strader LC, Zale JM, Sun Tp, Steber CM. **2003.** The Arabidopsis SLEEPY1 Gene Encodes a Putative F-Box Subunit of an SCF E3 Ubiquitin Ligase. *THE Plant Cell* **15:** 1120-1130.

Meudt HM, Clarke AC. 2007. Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends in Plant Science* **12:** 106-117.

Meyer D J, Crease D J, Ketterer B. **1995.** Forward and reverse catalysis and product sequestration by human glutathione S-transferases in the reaction of GSH with dietary aralkyl isothiocyanates. *Biochemistry Journal* **306:** 565-569.

Michalakis Y, Excoffier L. **1996.** A Generic Estimation of Population Subdivision Using Distances Between Alleles With Special Reference for Microsatellite Loci. *Genetics* **142:** 1061-1064.

Miliauskas G, Venskutonis PR, van Beek TA. **2004.** Screening of radical scavenging activity of some medicinal and aromatic plant extracts. *Food Chemistry* **85:** 231-237.

Milner JA. **1999.** Functional Foods and Health Promotion. *Journal of Nutrition* **129:** 1395.

Misson J, Raghothama KG, Jain A, Jouhet J, Block MA, Bligny R, Ortet P, Creff A, Somerville S, Rolland N, Doumas P, Nacry P, Herrerra-Estrella L, Nussaume L and Thibaud M-C. 2005. A genome-wide transcriptional analysis using *Arabidopsis thaliana* Affymetrix gene chips determined plant responses to phosphate deprivation. *Proceedings of the National Academy of Sciences of theUnited States of America* **102:** 11934-11939.

Mithen, Mithen R, Faulkner, Faulkner K, Magrath, Magrath R, Rose, Rose P, Williamson, Williamson G, Marquez, Marquez J. **2003.** Development of isothiocyanate-enriched broccoli, and its enhanced ability to induce phase 2 detoxification enzymes in mammalian cells. *TAG Theoretical and Applied Genetics* **106:** 727-734.

Mithen R, Campos H. **1996.** Genetic variation of aliphatic glucosinolates in Arabidopsis thaliana and prospects for map based gene cloning. *Entomologia Experimentalis et Applicata* **80:** 202-205.

Miyashita NT, Kawabe A, Innan H. **1999.** DNA Variation in the Wild Plant Arabidopsis thaliana Revealed by Amplified Fragment Length Polymorphism Analysis. *Genetics* **152:** 1723-1731.

Moore S, Payton P, Wright M, Tanksley S, Giovannoni J. **2005.** Utilization of tomato microarrays for comparative gene expression analysis in the Solanaceae. *Journal of Experimental Botany* **56:** 2885-2895.

Moose SP, Mumm RH. **2008.** Molecular Plant Breeding as the Foundation for 21st Century Crop Improvement. *Plant Physiology* **147:** 969-977.

Morey J, Ryan J, Van Dolah F. **2006.** Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biological Procedures Online* **8:** 175-193.

Morinaga SI, Nagano AJ, Miyazaki S, Kubo M, Demura T, Fukuda H, Sakai S, Hasebe M. **2008.** Ecogenomics of Cleistogamous and Chasmogamous Flowering: Genome-Wide Gene Expression Patterns from Cross-Species Microarray Analysis in Cardamine kokaiensis (Brassicaceae). *Journal of Ecology* **96:** 1086-1097.

Morozova O, Marra MA. 2008. Applications of next-generation sequencing technologies in functional genomics. *Genomics* **92:** 255-264.

Muangprom A, Osborn TC. **2004.** Characterization of a dwarf gene in Brassica rapa, including the identification of a candidate gene. *TAG Theoretical and Applied Genetics* **108:** 1378-1384.

Muangprom A, Thomas SG, Sun TP, Osborn TC. **2005.** A novel dwarfing mutation in a green revolution gene from Brassica rapa. *Plant Physiology* **137:** 931-938.

Nakajima M, Shimada A, Takashi Y, Kim YC, Park SH, Ueguchi-Tanaka M, Suzuki H, Katoh E, Iuchi S, Kobayashi M, Maeda T, Matsuoka M, Yamaguchi I. **2006.** Identification and characterization of Arabidopsis gibberellin receptors. *Plant Journal* **46:** 880-889.

Newman RM, Hanscom Z, Kerfoot WC. **1992.** The watercress glucosinolate-myrosinase system: a feeding deterrent to caddisflies, snails and amphipods. *Oecologia* **92:** 1-7.

Newman RM, Kerfoot WC, Hanscom Z. **1990.** Watercress and amphipods Potential chemical defense in a spring stream macrophyte. *Journal of Chemical Ecology* **16:** 245-259.

Nielsen PV, Rios R. **2000.** Inhibition of fungal growth on bread by volatile components from spices and herbs, and the possible application in active packaging, with special emphasis on mustard essential oil. *International Journal of Food Microbiology* **60:** 219-229.

Niklas KJ, Enquist BJ. **2001.** Invariant scaling relationships for interspecific plant biomass production rates and body size. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 2922-2927.

Noctor G, Foyer CH. **1998.** ASCORBATE AND GLUTATHIONE: Keeping Active Oxygen Under Control. *Annual Review of Plant Physiology and Plant Molecular Biology* **49:** 249-279.

Nugon-Baudon L, Robat S, Flinas J-P, Lory S, Beaune PH. 1998. Effects of the bacterial status of rats on the changes in some liver cytochrome P450 (EC1.14.14.1) apoproteins consequent to a glucosinolate-rich diet. *British Journal of Nutrition* **80:** 231-234.

Nyska A, Lomnitski L, Spalding J, Dunson DB, Goldsworthy TL, Grossman S, Bergman M, Boorman G. **2001.** Topical and oral administration of the natural water-soluble antioxidant from spinach reduces the multiplicity of papillomas in the Tg.AC mouse model. *Toxicology Letters* **122:** 33-44.

Ou B, Hampsch-Woodill M, Prior RL. **2001.** Development and Validation of an Improved Oxygen Radical Absorbance Capacity Assay Using Fluorescein as the Fluorescent Probe. *Journal of Agricultural and Food Chemistry* **49:** 4619-4626.

Ou B, Huang D, Hampsch-Woodill M, Flanagan JA, Deemer EK. **2002.** Analysis of antioxidant activities of common vegetables employing oxygen radical absorbance capacity (ORAC) and ferric reducing antioxidant power (FRAP) assays: a comparative study. *Journal of Agricultural and Food Chemistry* **50:** 3122.

Palaniswamy U, McAvoy R, Bible B. **1997.** Supplemental light before harvest increases phenethyl isothiocyanate in watercress under 8-hour photoperiod. *HortScience* **32:** 222-223.

Palaniswamy U, McAvoy RJ, Bible B, Singha S, Hill D. **1995.** Effect of different nitrogen-sulfur ratios on 2-phenyethyl isothiocyanate (PEITC) levels in watercress. *Horticultural Science* **30:** 789.

Palaniswamy UR, Mcavoy RJ, Bible BB, Stuart JD. **2003.** Ontogenic variations of ascorbic acid and phenethyl isothiocyanate concentrations in watercress (Nasturtium officinale R.Br.) leaves. *Journal of Agricultural and Food Chemistry* **51:** 5504-5509.

Pandey J, Ganesan K, Jain RK. **2007.** Variations in T-RFLP profiles with differing chemistries of fluorescent dyes used for labeling the PCR primers. *Journal of Microbiological Methods* **68:** 633-638.

Passioura JB. **1991.** Soil structure and plant growth. *Australian Journal of soil research* **29:** 717-728.

Peakall R, Smouse PE, Huff DR. **1995.** Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss Buchloe dactyloides. *Molecular Ecology* **4:** 135-148.

Peakall R, Smouse PE. **2006.** GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6:** 288-295.

Peng J, Carol P, Richards DE, King KE, Cowling RJ, Murphy GP, Harberd NP. **1997.** The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes and Development* **11:** 3194-3205.

Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD, Harberd NP. **1999.** Green revolution

genes encode mutant gibberellin response modulators. *Nature* **400:** 256-261.

Pfaffl MW. **2001.** A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29:** e45.

Podsedek A. **2007.** Natural antioxidants and antioxidant capacity of Brassica vegetables: A review. *LWT - Food Science and Technology* **40:** 1-11.

Powolny AA, Bommareddy A, Hahm E, Normolle DP, Beumer JH, Nelson JB, Singh SV. 2011. Chemopreventative potential of the cruciferous vegetable constituent phenethyl isothiocyante in a mouse model of prostate cancer. *Journal of the National Cancer Institute* **103:** 571-584.

Prior RL, Wu X, Schaich K. **2005.** Standardized Methods for the Determination of Antioxidant Capacity and Phenolics in Foods and Dietary Supplements. *Journal of Agricultural and Food Chemistry* **53:** 4290-4302.

Pulido R, Bravo L, Saura-Calixto F. **2000.** Antioxidant Activity of Dietary Polyphenols As Determined by a Modified Ferric Reducing/Antioxidant Power Assay. *Journal of Agricultural and Food Chemistry* **48:** 3396-3402.

Raskin I, Ribnicky DM, Komarnytsky S, Ilic N, Poulev A, Borisjuk N, Brinker A, Moreno DA, Ripoll C, Yakoby N, O'Neal JM, Cornwell T, Pastor I, Fridlender B. **2002.** Plants and human health in the twenty-first century. *Trends in Biotechnology* **20:** 522-531.

Rasmusson DC. **1987.** An Evaluation of Ideotype Breeding. *Crop Science* **27:** 1140-1146.

Rates SMK. **2001.** Plants as source of drugs. *Toxicon* **39:** 603-613.

Ribnicky DM, Poulev A, Henry E, Raskin I. 2001. Seed of Barbarea verna as a rich source of phenethyl iosthiocyanate to provide natural protection from environmental and dietary toxins. *Journal of Nutraceuticals, Functional and Medical Foods* **3:** 43-65.

Rich TCG, Dalby C, Botanical Society of the British Isles. 1991. *Crucifers of Great Britain and Ireland*. London: Botanical Society of the British Isles.

Riva-Souza EMRR, Sudre CP, Pereira MG, Bento C, Matta F. 2009. Genetic parameters and selection for resistance to bacterial spot in recombinant F6 lines of Capsicum annuum. *Crop Breeding and Applied Biotechnology* **9:** 108-115.

Rood SB, Williams PH, Pearce D, MUROFUSHI N, Mander LN, Pharis RP. 1990. A Mutant Gene That Increases Gibberellin Production in Brassica. *Plant Physiology* **93:** 1168-1174.

Rose PC. 2001. Anticarcinogenic compounds in watercress. Thesis, University of East Anglia.

Rose P, Faulkner K, Williamson G, Mithen R. 2000. 7-methylsulfinylheptyl and 8-methylsulfinyloctyl isothiocyanates from watercress are potent inducers of phase II enzymes. *Carcinogenesis* **21:** 1983-1988.

Rose P, Huang Q, Ong CN, Whiteman M. 2005. Broccoli and watercress suppress matrix metalloproteinase-9 activity and invasiveness of human MDA-MB-231 breast cancer cells. *Toxicology and Applied Pharmacology* **209:** 105-113.

Rumberger A, Marschner P. 2003. 2-Phenylethylisothiocyanate concentration and microbial community composition in the rhizosphere of canola. *Soil Biology and Biochemistry* **35:** 445-452.

Saeed L, Bhagabati WK, Braisted JC, Liang W, Sharov V, Howe EA, Thiagarajan M, White JA, Quackenbush J. 2006. TM4 Microarray Software Suite. *Methods in Enzymolgy* **411:** 134-193

Safe S, Papineni S, Chintharlapalli S. **2008.** Cancer chemotherapy with indole-3-carbinol, bis(3'-indolyl)methane and synthetic analogs. *Cancer Letters* **269:** 326-338.

Saibo NJM, Vriezen WH, Beemster GTS, Van Der Straeten D. **2003.** Growth and stomata development of Arabidopsis hypocotyls are controlled by gibberellins and modulated by ethylene and auxins. *The Plant Journal* **33:** 989-1000.

Sakamoto T, Kamiya N, Ueguchi-Tanaka M, Iwahori S, Matsuoka M. **2001.** KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. *Genes and Development* **15:** 581-590.

Sakamoto T, Morinaka Y, Ishiyama K, Kobayashi M, Itoh H, Kayano T, Iwahori S, Matsuoka M, Tanaka H. **2003.** Genetic manipulation of gibberellin metabolism in transgenic rice. *Nat Biotech* **21:** 909-913.

Satyan KS, Swamy N, Dizon DS, Singh R, Granai CO, Brard L. **2006a.** Phenethyl isothiocyanate (PEITC) inhibits growth of ovarian cancer cells by inducing apoptosis: Role of caspase and MAPK activation. *Gynecol.Oncol.* **103:** 261.

Satyan KS, Swamy N, Dizon DS, Singh R, Granai CO, Brard L. **2006b.** Phenethyl isothiocyanate (PEITC) inhibits growth of ovarian cancer cells by inducing apoptosis: Role of caspase and MAPK activation. *Gynecologic Oncology* **103:** 261-270.

Savelkoul PHM, Aarts HJM, de Haas J, Dijkshoorn L, Duim B, Otsen M, Rademaker JLW, Schouls L, Lenstra JA. **1999.** Amplified-Fragment Length Polymorphism Analysis: the State of an Art. *Journal of Clinical Microbiology* **37:** 3083-3091.

Schlesier K, Harwat M, hm V, Bitsch R. **2002.** Assessment of Antioxidant Activity by Using Different In Vitro Methods. *Free Radical Research* **36:** 177-187.

Schomburg FM, Bizzell CM, Lee DJ, Zeevaart JAD, Amasino RM. **2003.** Overexpression of a Novel Class of Gibberellin 2-Oxidases Decreases Gibberellin Levels and Creates Dwarf Plants. *The Plant Ce* **15:** 151-163.

Semba RD. 1998. The role of vitamin A and related retinoids in immune function. *Nutrition Reviews* **56:** 538-548.

Semba RD and Dagneli, G. 2003. Are leutin and zeaxanthin conditionally essential for eye health? *Medical Hypotheses* **61:** 465-472

Sheridan GEC, Claxton JR, Clarkson JM, Blakesley D. **2001.** Genetic diversity within commercial populations of watercress (Rorippa nasturtium-aquaticum), and between allied Brassicaceae inferred from RAPD-PCR. *Euphytica* **122:** 319-325.

Shigeoka S, Ishikawa T, Tamoi M, Miyagawa Y, Takeda T, Yabuta Y, Yoshimura K. **2002.** Regulation and function of ascorbate peroxidase isoenzymes. *Journal of Experimental Botany* **53:** 1305-1319.

Shimada N, Suzuki Y, Nakajima M, Conrad U, Murofushi N, Yamaguchi I. **1999.** Expression of a Functional Single-chain Antibody against GA24/19 in Transgenic Tobacco. *Bioscience, Biotechnology, and Biochemistry* **63:** 779-783.

Siglin JC, Brach DH, Stoner GD. **1995.** Effects of dietary phenethyl isothiocyanate, ellagic acid, sulindac and calcium on the induction and progression of N-nitrosomethylbenzylamine-induced esophageal carcinogenesis in rats. *Carcinogenesis* **16:** 1101-1106.

Simmonds NW. **1976.** *Evolution of crop plants*. Longman, London and New York.

Smirnoff N. **1998.** Plant resistance to environmental stress. *Current Opinion in Biotechnology* **9:** 214-219.

Song L, Thornalley PJ. **2007.** Effect of storage, processing and cooking on glucosinolate content of Brassica vegetables. *Food and Chemical Toxicology* **45:** 216-224.

Spence RM, Tucknott OG. **1983.** Volatiles from the epicuticular wax of watercress (*Rorippa nasturtium-aquaticum*). *Phytochemistry* **22:** 2521-2523.

Strader LC, Ritchie S, Soule JD, McGinnis KM, Steber CM. **2004.** Recessive-interfering mutations in the gibberellin signaling gene SLEEPY1 are rescued by overexpression of its homologue, SNEEZY. *Proceedings of the National Academy of Sciences* **101:** 12771-12776.

Stevens CP. 1983. *Watercress Production of the Cultivated Crops.* ADAS/MAFF Reference book no.136 Grower Books, London.

Svanem PJ, Bones AM, Rossiter JT. **1997.** Metabolism of [[alpha]-14C]-Desulphophenethylglucosinolate in Nasturtium officinale. *Phytochemistry* **44:** 1251-1255.

Symons G, Reid J. **2003.** Interactions Between Light and Plant Hormones During De-etiolation. *Journal of Plant Growth Regulation* **22:** 3-14.

Taiz L, Zeiger E. **2006.** *Plant physiology*. Sunderland, Mass: Sinauer Associates.

Tan Wf, Lin Lp, Li Mh, Zhang YX, Tong Yg, Xiao D, Ding J. **2003.** Quercetin, a dietary-derived flavonoid, possesses antiangiogenic potential. *European Journal of Pharmacology* **459:** 255-262.

Tanaka K, Nakamura Y, Asami T, Yoshida S, Matsuo T, Okamoto S. **2003.** Physiological Roles of Brassinosteroids in Early Growth of Arabidopsis: Brassinosteroids Have a Synergistic Relationship with Gibberellin as well as Auxin in Light-Grown Hypocotyl Elongation. *Journal of Plant Growth Regulation* **22:** 259-271.

Tanksley SD, McCouch SR. **1997.** Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild. *Science* **277:** 1063-1066.

Tata JR, Ernster L, Lindberg O. 1962. Control of basal metabolic rate by thyroid hormones and cellular function. *Nature* **193:** 1058-1060.

Traka M, Mithen R. **2009.** Glucosinolates, isothiocyanates and human health. *Phytochemistry Reviews* **8:** 269-282.

Ueguchi-Tanaka M, Ashikari M, Nakajima M, Itoh H, Katoh E, Kobayashi M, Chow TY, Hsing YIC, Kitano H, Yamaguchi I, Matsuoka M. **2005.** Gibberellin Insensitive Dwarf1 Encodes A Soluble Receptor for Gibberellin. *Nature* **437:** 693-698.

Ueguchi-Tanaka M, Hirano K, Hasegawa Y, Kitano H, Matsuoka M. 2008. Release of the repressive activity of rice DELLA protein SLR1 by gibberellin does not require SLR1 degradation in the *gid2* mutant. *The Plant Cell* **20:** 2437-2446.

Valasek MA, Repa JJ. **2005.** The power of real-time PCR. *Advances in Physiology Education* **29:** 151-159.

Vallejo F, Thomas-Barberan FA, Garcia-Viguera C. **2002.** Glucosinolates and vitamin C content in edible parts of broccoli florets after domestic cooking. *Eur Food Res Technol* **215:** 310-316.

Van Eck HJ, Van der Voort JR, Draaistra J, Van Zandvoort P, Van Enckevort E, Segers B, Peleman J, Jacobsen E, Helder J, Bakker J. **1995.** The inheritance and chromosomal localization of AFLP markers in a non-inbred potato offspring. *Molecular Breeding* **1:** 397-410.

Van Oeveren AJ, Stam P. **1992.** Comparative simulation studies on the effects of selection for quantitative traits in autogamous crops: early selection versus single seed descent. *Heredity* **69:** 342-351.

Varshney RK, Nayak SN, May GD, Jackson SA. 2009. Next generation sequencing technologies and their implications for crop genetics and breeding. *Trends in Biotechnology* **27:** 522-530.

Virk PS, Ford-Lloyd BV, Jackson MT, Newbury HJ. **1995.** Use of RAPD for the study of diversity within plant germplasm collections. *Heredity* **74:** 170-179.

Walsh JA. 1992. Resistant Watercress. Grower August 6th, 18-21.

Wang XH, Cavell BE, Syed Alwi SS, Packham G. **2009.** Inhibition of hypoxia inducible factor by phenethyl isothiocyanate. *Biochemical Pharmacology* **78:** 261-272.

Wang Y, Li J. **2006.** Genes controlling plant architecture. *Current Opinion in Biotechnology* **17:** 123-129.

Wattenberg LW. **1992.** Inhibition of carciongenesis by minor dietary constituents. *Cancer Research* **52:** 2085S-2091S.

Watcott W, Fort SB, Ryder EJ, Michelmore RW. 1999. Mapping morphological genes relative to molecular markers in lettuce (*Lactuca sativa* L.) *Heredity* **82:** 245-251.

Weigel D and Glazebrook J. 2002. *Arabidopsis A Laboratory Manual*. Cold Spring Harbour Laboratory Press, United States of America.

Weir BS and Cockerham CC 1984. Estimating F-statistics for the analysis of population structure**.** *Evolution* **38**: 1358-1370

Weng JR, Tsai CH, Kulp SK, Chen CS. **2008.** Indole-3-carbinol as a chemopreventive and anti-cancer agent. *Cancer Letters* **262:** 153-163.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. **1990.** DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18:** 6531-6535.

Willige BC, Ghosh S, Nill C, Zourelidou M, Dohmann EMN, Maier A, Schwechheimer C. 2007. The DELLA Domain of GA INSENSITIVE Mediates the Interaction with the GA INSENSITIVE DWARF1A Gibberellin Receptor of Arabidopsis. *The Plant Cell* **19:** 1209-1220.

Wilson K and Walker J. 2005. *Principles and Techniques of Biochemistry and Molecular Biology sixth edition*. Cambridge University Press, Hong Kong.

Wittstock U, Halkier BA. 2002. Glucosinolate research in the Arabidopsis era. *Trends in Plant Science* **7:** 263-270.

Wright S. 1920. Principles of Livestock Breeding. *US Department of Ariculture* Bulletin number 905.

Wu X, Zhou Q and Xu K. 2009. Are isothiocyantes potential anti-cancer drugs? *Acta Pharmacologica Sinica* **30:** 501-512.

Yang S, Pang W, Ash G, Harper J, Carling J, Wenzl P, Huttner E, Zong X, Kilian A. 2006. Low level of genetic diversity in cultivated Pigeonpea compared to its wild relatives is revealed by diversity arrays technology. *TAG Theoretical and Applied Genetics* **113:** 585-595.

Yang XC, Hwa CM. 2008. Genetic modification of plant architecture and variety improvement in rice. *Heredity* **101:** 396-404.

Yin P, Kawamura T, He M, Vanaja DK, Young CYF. 2009. Phenethyl isothiocynate induces cell cycle arrest and reduction of α and β-tubulin isotopes in human prostate cancer cells. *Cell Biology International* **33:** 57-64.

Yu O, Jez JM. 2008. Nature's assembly line: biosynthesis of simple phenylpropanoids and polyketides. *The Plant Journal* **54:** 750-762.

Yuk Man Li Hyecyhadzyc. 2008. Broccoli (*Brassica oleracea* var. botrytis L.) improves the survival and up-regulates endogenous antioxidant enzymes in *Drosophila melanogaster* challenged with reactive oxygen species. *Journal of the Science of Food and Agriculture* **88:** 499-506..

Zamir D. **2008.** Plant breeders go back to nature. *Nat Genet* **40:** 269-270.

Zhang Y. **2004.** Cancer-preventive isothiocyanates: measurement of human exposure and mechanism of action. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **555:** 173-190.