UNIVERSITY OF SOUTHAMPTON

THE GENETICS AND GENOMICS OF DROUGHT RESPONSE IN POPULUS

In one volume

Nathaniel Robert Street

Doctor of Philosophy

School of Biological Sciences

September 2005

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES

BIOLOGICAL SCIENCES

Doctor of Philosophy

THE GENETICS AND GENOMICS OF DROUGHT RESPONSE IN POPULUS

Nathaniel Robert Street

The genetic nature of tree adaptation to drought stress was examined by utilising variation in the drought response of an F_2 mapping population from a cross between Populus trichocarpa (93-968) and P. deltoides Bart (ILL-129) known to be highly divergent for a vast range of phenotypic traits.

Phenotyping, QTL analysis and microarray experiments were combined to demonstrate that "genetical genomics" can be used to provide information on adaptation at the species level. The grandparents and F_2 population were subjected to soil drying and contrasting responses to drought across genotypes including for leaf coloration, expansion, and abscission were observed and QTL for these traits were identified. A subset of extreme genotypes exhibiting extreme sensitivity and insensitivity to drought on the basis of abscission were defined and microarray experiments were conducted on these genotypes and the grandparents. The different groups induced a different set of genes; 215 and 125 genes differed in their expression response in control and drought respectively, suggesting species adaptation at the gene expression level. Genes preferentially expressed in drought resistant genotypes overlapped with genes expressed in dormant tissues whereas genes involved in meristem function had a lower expression. Co-location of differentially expressed genes with drought specific and drought responsive QTLs was identified and these represent candidate genes contributing to the variation in drought response.

Contents

List of Figures

List of Tables

Acknowledgements

I thank Professor Gail Taylor for securing the studentship funding for my PhD and for her supervision and guidance. I thank the National Research Environment Council (NERC) for awarding and funding that studentship (NERC ref: NER/S/A/2001/106361). I would also like to thank Gail Taylor for the opportunities given to me to be involved in a range of European collaborative work. Involvement in such collaborations has been stimulating, informative, and has provided me with access to expertise that have helped ensure the success of this work.

I would like to especially thank Caroline Dixon for her continued help with all aspects of setting up, maintaining, and sampling experiments.

I thank James Tucker, Anne Rae, Mathieu Pinnel, Caroline Dixon, and Mike Cotton for their assistance in the collection of cutting material.

I thank Anne Rae, Mathieu Pinnel, Caroline Dixon, and Mike Cotton for their assistance in watering and filling pots with soil for the QTL mapping experiment. I additionally thank Penny Tricker, James Tucker, and Maricela Rodríguez-Acosta for their assistance in filling pots with soil. I thank Harriet Trewin, Nicole Harris, Laura Graham, Sarah Elhag, Sera Bowden, Dan Stapleton, Charlotte Freer-Smith, Penny Tricker, Anne Rae, Mike Cotton, Caroline Dixon and especially Matthew Tallis and James Tucker for their assistance in sampling the QTL mapping experiment. Their patience and assistance was truly appreciated.

I thank James Tucker for his assistance on numerous occasions with the collection of experimental samples and data. I am also indebted to him for the continual debate and discussion he provided on all matters relating to my PhD.

I thank Anne Rae for her guidance and instruction with QTL mapping.

The data on stomatal conductance and photosynthetic rate presented in Chapter 2 were collected by Maricela Rodríguez-Acosta. The chlorophyll data presented in Chapters 4 and 5 were collected by Sarah Elhag. All data were analysed by myself.

I thank Stefan Jansson for the opportunity to run the microarray experiment described in Chapter 2. I am especially grateful to Oskar Skogström for his training in the methods of microarray hybridisation and am eternally grateful to Andreas Sjödin for his extensive training and discussion on the statistical treatment of microarray data and for his invaluable introduction to R. The RT-PCR presented in Chapter 3 was carried out by Oskar Skogström using primers supplied by myself.

List of abbreviations

ABA	Abscisic acid
ABRE	ABA response element
AFLP	Amplified fragment length polymorphism
ANOVA	Analysis of variance
bp	base pair
CAPS	Cleaved-amplified polymorphic sequence
cDNA	complementary DNA
CDPK	Calcium dependant protein kinase
CIM	Composite interval mapping
cM	centi-Morgans
CRAN	Comprehensive R Archive Network
DAP	Days after planting
DAD	Days after drought [initation]
DMF	Dimethylformaide
DNA	Deoxyribonucleic acid
DRE	Dehydration response element
eQTL	expression QTL
EST	Expressed sequence tag
glm	general linear model
HMM	Hidden Markov model
Kb	Kilo-bases
LEA	Late embryogenesis abundant
LOD	Likelihood-of-odds
LPI	Leaf Plastochron Index
LS	Least squares
MAPK	Mitogen activated protein kinase
MAS	Marker-assisted selection
MCMC	Markov chain Monte Carlo
MIM	Multiple interval mapping
MIP	Major intrinsic protein
ML	Maximum Likelihood
mRNA	messenger RNA
NCED	9-*cis*-epoxycarotenoid dioxygenase
PA	Phosphatidic acid
PCA	Principle components analysis. (Also photosynthetic carbon oxidation)
PCR	Polymerase chain reaction. (Also Photosynthetic carbon reduction)
PLC	Phospholipase C
PLS	Partial least squares
PMT	Photomultiplier tube

PP2C	Protein phosphatase 2C
PPLD	Phospholipase D
qRT-PCR	Quantitative reverse-transcription PCR (Also quantitative real-time PCR)
qRR-PCR	Quantitative reverse-transcription real-time PCR
QTL	Quantitative trait locus/loci
RAPD	Randomly amplified polymorphic DNA
RFLP	Random fragment length polymorphism
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RT-PCR	Reverse transcription PCR (Also real-time PCR)
SAM	Shoot apical meristem
SCAR	Sequenced characterised amplified region
SNP	Single-nucleotide polymorphism
SSR	Short sequence repeat (also referred to as microsatellite)
STS	Sequence tagged site
TIFF	Tagged Image File Format (Also commonly TIF)
XTH	xyloglucan endotransglycosylase

Gene names are shown in *CAPITAL* and abbreviations in *CAPS*

Chapter 1

General Introduction

1 General Introduction

The field of biological science is undergoing a fundamental reform; it is maturing from a descriptive to a predictive science. This transition was initiated by two milestone achievements now familiar with all. The first was Darwin's theory of natural selection that allowed an understanding of the forces driving evolutionary change. The second was the determination of the structure of DNA, which enabled scientists to read the base-pair code instruction set for all organisms. As an ever increasing number of organisms have their genomes sequenced, our ability to read and interpret that base-pair code language improves. More recently a revolution has occurred in biology brought about through the development of a host of new technologies that allow simultaneous examination of mRNA, protein, and metabolite levels and through vast improvements in the ability to obtain and analyse sequence data. These technologies have led to the rapid development of the 'genomics' and 'bioinformatics' era and methods and have led to a paradigm shift in the scale of information available about the state of cells and organisms at various biological levels.

This improved knowledge allows one to progress from describing what is happening within a system to predicting how the system will respond based on knowledge of the instruction set used to construct the system. Chaves, Maroco and Pereira (2003) mention the genotype-phenotype gap that currently exists in our understanding of many biological responses. It is this gap that aims to be filled by the move from descriptive to predictive.

1.1 Project overview

My PhD aims to identify the genetic control of drought-stress response in *Populus* and to contribute to the understanding of the genetic architecture controlling this response within the genus *Populus* by identifying the deviations in instruction set (i.e. the base-pair sequence of the genes involved in drought response) that are responsible for

determining the difference in response between two *Populus* species with divergent drought stress responses. A mapping pedigree of F_2 individuals produced from repeated crossing of F_1 hybrids of the two species has been produced and will be used for investigation of the segregation of response within the F_2 genotypes. The parental species (genotypes) are highly divergent in many phenotypic traits and were selected from contrasting climatic regions in the USA.

In order to attain a predictive ability, much empirical work is needed to ascertain the mechanisms responsible for the differences between the species. This first requires the characterisation of those differences. Once the nature of the differences is understood, the aim is to discover the genetic architectural differences underlying the observed divergence of phenotypes, which itself requires a detailed understanding of the response of each species in order to identify where differences exist.

"In order to understand the essential features of a system or process it is necessary to make measurements" (Eriksson *et al.* 1999). Characterisation of the differences requires firstly descriptive, physiological and morphological investigation and examination. Once these broad scale differences have been characterised, experiments can be designed to determine the genetic control of that trait within each species and to explain how the differences between species are obtained (such approaches are discussed in Taiz and Zeiger, 1998).

My principle interests are to explain the differences in leaf morphology and abiotic stress response between the two parental species and within their offspring. An ultimate aim of such research is the ability to predict the outcome of gene combinations in F_2 progeny initially, and in breeding applications thereafter.

One can readily observe a range of leaf forms, be they grasses on a lawn, leaves on a tree, or ivy advancing across a house. The diversity of leaf shape and form is both very apparent and intriguing. Leaves are essentially structures for the capture of photons and gas exchange as well as housing a battery of sensors for monitoring the external environment. The size and shape of leaves can be considered to be the 'best fit' option

for the environment in which the plant evolved. As an aside, it is important to add to this model-view a consideration of the life-history of a species; a species can only evolve such a best-fit option through modifications to its starting instruction set, which forms a continuous progression back to the instruction set of the common ancestor (Futuyama 1998). One realises that these starting limitations provide plenty of scope for modification when considering such divergent examples as that of the cacti, which lack any leaves and achieve the functions mentioned above by use of the stem and needles, through to *Paulownia spp.* trees, which have leaves spanning up to 60 cm and forming a defined canopy structure.

Understanding abiotic stress is of both academic and commercial importance. Drought stress is one of the most important limitations of crop production in the world (Bray, 1997) and loss of harvest biomass is reduced by abiotic stresses to a far greater extent then it is by biotic stresses (Bray, 2000). According to the latest IPCC report (IPCC 2001), drought frequencies are expected to increase globally over the next 50 years and there will be increasing shifts in rainfall patterns, In the UK it is estimated that winters are set to become 20% wetter on average and summers 20% driers in the south (Broadmeadow 2002). The study of abiotic stress also offers the potential to answer many other biological questions as abiotic stress response incorporates complex signalling pathways that bring about a host of changes to the physiology, growth and development of a plant.

The use of model species

Biological organisms represent highly complex systems that exist in a multitude of forms, these forms being grouped together at various levels represented by the phylogenetic tree. It would be impossible to study the details of each organism individually yet it is obviously desirable to understand as much about them as possible. In order to achieve this aim, certain representative organisms are chosen to serve as model species. The most well known is probably the fruit fly, *Drosophila melongaster*, which is used as a model for animal embryogenesis and pattern formation (Gilbert 1997). The first plant species to be globally adopted as a model for

genetic research was *Arabidopsis thaliana*, and accordingly this was the first plant species to have its genome fully sequenced (Arabidopsis Genome Initiative, 2000). To date, little research has exploited *Arabidopsis* as a model for evolutionary, and particularly ecological, studies although this is a trend that appears to be reversing. *Arabidopsis* offers excellent opportunities for studying the dynamics of a population experiencing environmental change (Mitchell-Olds 1996, Koornneef *et al.* 1997). It is also suitable for studying other processes of ecological importance such as competition, resource allocation, and population ecology (Callahan *et al.* 2005, Chen *et al.* 2005)

Populus as a model

As with any model, the exclusive use of *A. thaliana* to model plant development and responses would be too gross a simplification. Traits and adaptive methods present in many species are not represented in *Arabidopsis* (Brunner *et al.* 2004, Taylor 2002, Wullschleger *et al.* 2002, Bradshaw *et al.* 2000). In particular, *Arabidopsis* is an annual species (a very rapid one at that) and is therefore unsuitable for studying the adaptive mechanisms, ecology and evolution of perennial, long lived species. Such perennial species undergo a number of processes that are of great interest including winter dormancy and the production of dormant meristems that can either remain dormant or initiate growth to form branches. Whether a dormant meristem becomes active is dependent on developmental stage of the plant, abiotic environmental triggers such as stress, nutrient availability and light availability, and biotic factors such as competition and herbivory. Winter dormancy and the processes of autumn leaf senescence, bud set, bud burst, and woody tissue formation are also not modelled by *Arabidopsis*. Similarly, long-term adaptation of individuals to a changing environment can not be studied in an annual species. It has therefore proven necessary to select a model for such perennial species. It is the genus *Populus*, and in particular the species *P. trichocarpa,* that has emerged as the accepted model for perennial species and accordingly, this was the first perennial species to have its genome fully sequenced (http://www.ornl.gov/sci/ipgc/ or http://genome.jgi-psf.org/Poptr1/Poptr1.home.html). This is an achievement that positively impacts upon the work of my PhD, making many previously unachievable tasks possible.

Populus species exhibit the most rapid growth of any perennial hardwood species, have a relatively small genome compared to other tree species, are reasonably closely related to *Arabidopsis*, is already of commercial importance, can be readily transformed, clonally propagated, and has a global range: The trembling aspen species (*P. tremula, P. tremuloides)* of Sweden and Canada grow well into the arctic circle and exhibit extreme freezing tolerance, rapid, synchronised spring bud flush with a short growing season followed by rapid, synchronised autumn senescence; the middle eastern *P. euphratica* grows in near desert-like conditions with extreme tolerance of

both water deficit and high salinity; the temperate species of north America exhibit some of the most rapid growth of any poplar species and form many hybrid zones as well as providing clines in a range of environmental conditions in addition to acting as keystone species along many riverbed ecosystems of the USA; the *P. lasiocarpa* species of China has strikingly large leaves and many individuals are monoecious in contrast to the dioecious form of most poplar species.

The genus *Populus* is a member of the Salicaceae family, which includes the willow (*Salix*) genera. The taxonomic classification of poplar has received much debate and study. At present, 29 species are recognised and these are arranged into six sections (Tacamahaca, Leucoides, Aigeros, Populus, Abosa, and Turanga; Eckenwalder 1996). More detail on the taxonomic placement of the Salicaceae is given in Eckenwalder (1996).

The demand for improved, task-specific tree species

Historically it was the case that trees existed in excess of demand and it is only recently that industries making use of tree products have begun to consider the environmental impacts and long term sustainability of their industrial process. For example, the pulp and paper industry is demanding ever improved varieties of trees that allow optimisation of the pulping process. This move comes hand-in-hand with demands on the pulping industry to minimise its environmental pollution and impact, which could be principally achieved through modification of the wood structure of trees (Li *et al.* 2003a). Trees are also gaining popularity as a readily achievable source of renewable, carbon-neutral energy through the chipping and burning of harvested biomass or, alternatively, the production of ethanol as a fuel source from fermented biomass (Dinus *et al.* 2001). Although the technology to establish this industry exists, the commercial success will depend on the provision of new high-yielding genotypes and genotypes that can achieve high biomass yield in a range of environments. Tree species, and in particular *Populus spp.* are also being utilised for phytoremediation; particularly the detoxification of landfill sites and sites contaminated with heavy metals. The efficiency of the genotypes currently used for this purpose could be increased significantly through genetic manipulation or directed breeding to increase heavy metal tolerance, and transportation and uptake. Such industrial use of biological technologies is highly desirable as it is environmentally friendly, can increase local biodiversity, and is sustainable.

Conclusions

The work of my PhD will make use of both pre-established and emergent technologies and methodologies and will combining these together to form a holistic (megavariate) approach to exploring the genetic control and architecture of drought stress response and leaf development, a trait that is highly sensitive to drought stress. An overview of the biological themes to be investigated follows this overview. There then follows an introduction to the principle methods of investigation that I will be making use of. The section on QTL analysis and mapping draws the two methods in together as well as

introducing complementary analysis methods that will be made use of. Additional detail about the application of the methods, both individually and in combination, is given within each of the experimental chapters.

1.2 Drought

Plants need to maintain a specific set of internal conditions in order to grow and function. Any deviation in the external environment from the optimal will place a plant under abiotic stress, be it drought stress, light stress, temperature stress etc. As soon as a stress condition is entered, growth and function become sub-optimal, leading to a decrease in fitness. Natural selection dictates that fitness must be maintained at all times in order for an organism to survive and, as a result, a myriad of stress responses have evolved to enable survival. Trees, unlike annual species, must be able to adapt to both short term acute stress periods as well as environmental perturbations occurring over longer periods of time.

Plants have evolved many physiological and biochemical responses to drought stress (Lichtenthaler 1996, Lichtenthaler 1998). These include stomatal closure (Wilkinson and Davies 2002), alteration of carbon partitioning (Borghetti *et al.* 1998), altered plant growth (Gowing *et al.* 1990, Zhu 2001, Dubos *et al.* 2003), alterations in the partitioning and ratio of sugars, sugar alcohols, and starch (Volaire and Thomas 1995), alterations to the photosynthetic apparatus (Loggini *et al.* 1999), and alterations to the composition of the lipid membrane (Zuniga *et al.* 1990, Maldonadoa 1997).

Drought stress also results in re-modelling of the transcriptome (Bray 1997, Shinozaki and YamaguchiShinozaki 1997, Seki *et al.* 2001a, Bray 2004, 2002, Kreps *et al.* 2002, Seki *et al.* 2002a,b, Oono *et al.* 2003). Genes responding to drought stress can be broken down into those that function directly in protecting against the stress and those that regulate gene expression and signal transduction (Bray 1997, Shinozaki and Yamaguchi-Shinozaki 1997, Thomashow 1999, Hasegawa *et al.* 2000).

As noted by Zhu (2002), an observed stress response can be considered either as adaptive or pathological. Zhu suggests defining adaptive responses into three groups: those that are involved in a) homeostasis, b) detoxification, and c) growth control. When observing the response of a plant to stress it is of key importance to ask oneself

"is this response adaptive or merely pathological?" and to avoid the trap of 'designing' an adaptive explanation for observed responses.

1.2.1 Physiological responses to drought

All plants have set up, in a 'ready and waiting' state, a host of sensory mechanisms to monitor their external environment. Many of these sensory mechanisms remain unknown and uncharacterised and are hard to espy as they respond rapidly and transiently to an external stimulus. It is their downstream effects that are typically observed and investigated. In the case of drought, the first and primary adaptive mechanism a plant initiates in response to detecting a decrease in water availability is to reduce water use and control the rate of water loss. These two adjustments are principally achieved through a reduction in transpiration via stomatal closure (Luan 2002; Wilkinson and Davies 2002) and a reduction in growth rate (Zhu 2001), which reduces the metabolic demand for water. The two processes also form a complex feedback loop whereby stomatal closure impacts on growth.

Stomatal response

If the water supply to any system is reduced, the most obvious first response upon detecting this reduction is to reduce the amount of water being lost through the functioning of the system. This holds true for plants exposed to drought stress. One of the earliest observed responses to drought is the partial or complete closure of stomata (Bray 1997, Luan 2002, Wilkinson and Davies 2002). Stomata regulate the uptake of CO_2 for photosynthesis and the loss of water via transpiration. The aperture of the stomatal pore is under the influence of a wide range of environmental signals via the activation of guard-cell signalling pathways that alter guard cell turgor pressure (Blatt 2000, Schroeder *et al.* 2001b). One of the best understood of these signalling pathways is that induced by ABA, which accumulates in leaves in response to soil drying / water deficit and brings about a the closure of stomata.

Under normal conditions, stomata open and close in a light-dark cycle as this maximises efficiency in the balance between CO_2 supply for photosynthesis, leaf

temperature control, and water loss through transpiration. During the early part of an ideal day (high light, low cloud cover), stomata are open to achieve an optimal flux of CO_2 into the intracellular airspaces of the mesophyll. When light harvesting ceases at night CO_2 is no longer required and stomata close to conserve water (Luan 2002). Midday stomatal closure is often observed as water loss is maximal during this period.

This normal, optimal operating state will, however, require adjustment upon exposure to drought (or other stress). In conditions of limited water supply, survival becomes more important than growth and development. The 'ready and waiting' drought sensors will be activated and this activation will switch on a 'program' that operates the closure of stomata. In the same way, many other programs will be activated, each aiming to adjust physiological, biochemical, and developmental processes in an optimal manner to cope with water deficit.

A stomatal pore is formed between two guard cells. The turgor pressure of these guard cells is finely adjusted to allow control of the size of that pore. When turgor pressure is increased and the cells swell, the pore diameter increases concomitantly allowing a greater flux of CO_2 and greater water loss through the transpiration stream. The turgor pressure of guard cells is controlled principally by adjusting the cytosolic concentrations of K^+, Cl, Ca^{2+}, and malate. Reviews of ion channel movements in stomata can be found in Blatt and Grabov (1997), MacRobbie (1997), and Luan (2002).

ABA induced stomatal closure

Stomatal closure in response to ABA is achieved through changes in guard cell turgor pressure, membrane trafficking, cytoskeletal reorganisation, changes in sugar and organic acid metabolism, changes in pH, the production of numerous secondary messengers, and changes in gene expression (Hetherington 2001). Empirical experiments established that ABA is the key phyto-hormone involved in the control of stomatal opening (Little and Eidt 1968, Mittelheuser and van Steveninck 1969) and that bulk-leaf ABA is increased under drought stress (Wright 1969, Wright and Hiron

1969). Additionally, ABA insensitive mutants (*abi*) tend to wilt readily (Koorneef *et al.* 1984, Roelfsema and Prins 1995). It is now known that other stresses also induce an increase in ABA. For example Kreps *et al.* (2002) shows an increase in ABA concentration in response to cold, osmotic and salt stress and ABA has been shown to play an essential role in the establishment of winter dormancy and winter cold / freezing tolerance (Lang *et al.* 1994, Mantyla *et al.* 1995). Exogenous application of ABA results in the rapid closure of stomata and also prevent subsequent opening (Willmer and Fricker 1996). ABA is known to accumulate in roots in response to soil drying (Wilkinson and Davies 2002) and this ABA is transported via the transpiration stream to the leaves. ABA therefore acts as a long-range chemical signal for stomatal aperture control in response to soil water status.

An effect of ABA has been observed on three ion channels. ABA enhances the activity of outward K^+ channels (Blatt and Armstrong 1993) and slow (S)-type anion channels (Grabov *et al.* 1997, Pei *et al.* 1997), and decreases the activity of inward K^+ channels (Blatt and Armstrong 1993; Lemtiri-Chlieh and MacRobbie 1994, Schwartz *et al.* 1994). Felle *et al.* (2000) also observed an increase in Cl^- that was initiated a short time after application of ABA. Roelfsema *et al.* (2004) show that ABA initiates a transient depolarisation of the guard cell membrane that brought about stomatal closure through the transient activation of S and R type anion channels. Kwak *et al.* (2002) identified a guard cell expressed type 2A protein phosphatase (PP2C) encoding gene, *RCN1*, important in controlling the guard cell ABA signal response. A mutation in the *RCN1* gene results in ABA insensitivity impairs ABA-induced stomatal closure and activation of S type anion channels. *RCN1* was shown to act upstream of Ca^{2+} signalling as all downstream functions responded as in the wild-type. Hunt *et al.* (2003) have also shown a role for phospholipase C (PLC) in the control of stomatal aperture by ABA. Transgenic tobacco plants having reduced PLC levels were only partially able to regulate stomatal aperture in response to applied ABA. A negative regular of the ABA response has also been identified by Leonhardt *et al.* (2004) in an elaborate experiment that involved monitoring the expression profile of the guard cell transcriptome, identification of a PP2C gene that was highly up-regulated by ABA and the subsequent use of a disruption mutant to confirm a role of

the gene encoded by the microarray cDNA. The mutant form of the gene, *AtP2c-HA*, results in ABA hypersensitivity that resulted in more rapid closure of stomata than in the wild-type control. Another gene, *OST1*, was identified by Assmann (2003) that encodes a serine-threonine protein kinase. A mutation in the gene renders stomata insensitive to ABA resulting in the stomata remaining open. *OST1* was shown be involved in reactive oxygen species (ROS) signalling and application of exogenous ROS restores the wild-type phenotype. Interestingly, the wild-type phenotype is also restored by application of Ca^{2+} suggesting that the ROS signal is involved in the increase in Ca^{2+} required for stomatal closure. ROS play an important role in stomatal closure and in many other signalling pathways, providing information about changes in the external environment (Chen and Gaille 2004). ROS signalling has filled in a missing gap in the diurnal pattern of signalling controlling guard cell opening / closure and its oscillations account for the often-observed mid-day closure of stomata (Assmann 2003, Assmann and Wang 2001). H_2O_2 is the ROS involved in guard cell signalling and the levels of H_2O_2 are controlled by the ascorbic acid redox state (Chen and Gaille 2004).

ABA-independent stomatal closure

In addition to the ABA-induced closure of stomata, a second, ABA-independent response is also recognised. This occurs in response to hyperosmotic shocks that lead to rapid water loss (Raschke 1975). This osmotic shock leads to rapid closure of stomata (Zeiger *et al.* 1987, Assmann 1993) in an ABA-independent manner (Assmann *et al.* 2000). This response is typically triggered due to changes in the humidity of the air surrounding a leaf surface. It has been suggested that stretch activated channels (SAC) that are independent of membrane polarity could be involved in this response (Luan 2002). Such SACs have been identified in the plasma membrane of guard cells by Cosgrove and Hedrich (1991). Liu and Luan (1998) have also shown that voltage-gated K^+ channels are involved in this response and that they are regulated by the osmogradient across the plasma membrane. In the same paper by Liu and Luan (1998), the researchers also show that this ion channel activity is closely associated with a change in the actin organisation of the cell in response to cell volume

changes induced by osmotic stress. The organisation of actin filaments in guard cells in relation to stomatal aperture has previously been reported by Kim *et al.* (1995) and Hwang *et al.* (1997).

Leaf expansion

An inhibition of leaf area expansion is one of the earliest symptoms of drought stress (Hsiao 1973). This expansion inhibition can result from reduced turgor pressure, required to achieve an increase in cytosolic volume and therefore cell area, reduced cell division, or reduced expansion of developing cell walls through a decrease in cell wall extensibility (Lu and Neumann 1998).

Cell wall extensibility is controlled by both the activity of cell wall modifying proteins such as expansins (McQueen-Mason *et al.* 1992), xyloglucan endotransglycosylase (XTH; Miller and Fry 2001, Rose *et al.* 2002), and glucanases (Li *et al.* 2003b). Many reports have identified down-regulation of genes involved in cell wall modification, degradation, and synthesis and these genes represent likely mechanisms through which cellular expansion is reduced in response to water deficit (Kreps *et al.* 2002, Seki *et al.* 2002a,b, Kawaguchi *et al.* 2004).

Photosynthesis and oxidative stress

Drought-induced stomatal closure results in a decrease in CO_2 flux into the leaf and therefore a reduction in intra-cellular CO_2 concentration, which impacts upon the photosynthetic apparatus. A decrease in the availability of cellular H_2O also impacts on the functioning of the photosynthetic biochemistry. In drought conditions, excess electrons are produced by photo-system II (PSII) and these must be removed in order to avoid and prevent damage to the chloroplast being caused through the production of reactive oxygen species and other free radicals. This is principally achieved through an increase in the oxygenation of 1,5-bisphophate carboxylase/oxygenase (Rubisco). Numerous studies have shown that the decrease in intracellular CO_2 concentration of drought-stressed leaves results in a decrease in CO_2 assimilation rates that is not reversible upon return the normal concentrations of CO_2 (Lawlor and Cornic 2002, Cornic and Fresneau 2002, Tang *et*

al. 2002). This suggests that drought stress causes changes in mesophyll metabolism. The change in mesophyll metabolism becomes increasingly important as water stress increases and may lead to significant reduction in photosynthetic ability during chronic drought periods (Gimenez *et al.* 1992, Tezara *et al.* 1999).

Although debate continues as to the most significant changes in metabolism induced by drought (Parry *et al.* 2002), a reduction in the activity and quantity of Rubisco has been added to the list of universally observed drought responses (Chaves *et al.* 2003; Tezara *et al.* 1999, Gimenez *et al.* 1992) and vast changes can be seen in the expression of genes encoding the sub-units of Rubisco. The amount of Rubisco in leaves is controlled by the balance between synthesis and degradation. The Rubisco holoenzyme remains relatively stable under drought stress (Webber *et al.* 1994) yet decreases in Rubisco content are observed in response to drought in many species. In tomato (Bartholomew *et al.* 1991), *Arabidopsis* (Williams *et al.* 1994), and rice (Vu *et al.* 1999) rapid decreases were observed in the small subunit of Rubisco, which may suggest a decrease in the rate of synthesis.

A shift from photosynthetic carbon reduction (PCR) to carbon oxidation (PCA) i.e. photorespiration, is essential in drought conditions in order to protect and maintain the photosynthetic machinery, particularly that of PSII that would otherwise release dangerous amount of free electrons into the stroma (Cornic and Fresneau 2002, Medrano *et al.* 2002, Noctor *et al.* 2002; Munne-Bosch and Penuelas 2003). A shift towards photorespiration causes an increase in the production of the oxidant H_2O_2 (Noctor *et al.* 2002). This is an interesting consequence considering the role that H_2O_2 plays in signalling drought stress (discussed below) and may explain the presence of an active, yet limited, photosynthetic system in guard cells. This increase in H_2O_2 alongside increases in antioxidants (ascorbate and glutathione) to compensate form part of a stress-detection / signalling pathway that serves to detect the oxidative load on the plant cell (Noctor *et al.* 2000). Mild increases in oxidative load will trigger acclimatory mechanisms, while increasing load eventually tips to balance over to senescence or (in an un-checked increase in oxidative load), cellular necrosis (Noctor *et al.* 2002). This increase in active oxygen species (AOS) and ROS production and

the resulting increase in antioxidant production can result in an increased level of salicylic acid concentration and this may explain the often-observed overlap between the pathogen-attack response and drought (and other stress) response and explain the up-regulation of pathogenesis-related protein gene expression that is observed as part of the drought-response mechanism.

1.2.2 Biochemical adaptation

As well as changes in plant morphology and physiology, adaptive responses in cellular biochemistry are initiated in response to drought. These mainly serve to maintain cellular homeostasis. Biochemical changes in the photosynthesis system and AOS / ROS production were discussed in the above section on the photosynthesis response to drought stress. A number of other changes occur that serve to maintain cellular integrity and turgor pressure.

Osmoprotectants

Osmoprotectants such as proline, glycinebetaine, mannitol, pinitol, and sugars have been shown to be important in cellular stress tolerance (Bray 1997) and manipulation of genes encoding key enzymes of the osmolyte synthesis or degradation pathways has confirmed this (Kavi-Kishor *et al.* 1995, Tarczynski *et al.* 1993; Thomas *et al.* 1995; Nanjo *et al.* 1999a, b). The raffinose family oligosaccharides raffinose and galactinol have also been shown to accumulate and play a key role in the maintenance of osmotic pressure during drought stress (Taji *et al.* 2002). Verslues and Bray (2004) have recently reported the first two genes, *LWR1* and *LWR2*, that are regulated in response to changes in cellular water potential and have shown that the genes play a role in controlling the accumulation of proline and other solutes as well as influencing the ABA response mechanism.

Cuticle

A role for the cuticle in drought response has also been shown recently by Aharoni *et al.* (2004). They identified that the SHINE clade of AP2 domain transcription factors activates wax biosynthesis resulting in alterations in the cuticle properties that increased the drought tolerance of *Arabidopsis* plants. They showed that plants over-expressing the *SHN* gene were more drought tolerant and they postulate that this was due to a decrease in stomatal density and a decrease in evaporative water loss through the cuticle layer. This is an area of research that has received little attention to date, although alteration in the appearance of leaves is often noted during drought periods and many of these differences are the result of altered reflectance from the leaf surface, which is the result of changes to the wax layer covering the leaf.

Aquaporins

The discovery of aquaporins caused a paradigm shift in the understanding of plant water relations (Lian *et al.* 2004). Aquaporins facilitate and regulate the passive exchange of water across cell membranes (De Groot and Grubmuller 2001, Sui *et al.* 2001). They belong to a highly conserved membrane protein family, MIP (major intrinsic protein: Maurel 1997, Tyerman *et al.* 1999; Tyerman *et al.* 2002). Aquaporins are abundant in both the plasma membrane and vacuolar membrane of plants. The expression of two aquaporin genes in particular has been reported to undergo changes in abundance in response to drought, *RWC3* and *RD28* (Lian *et al* 2004). Aharon *et al.* (2003) have also shown that over-expression of the *Arabidopsis PIP1b* aquaporins in tobacco resulted in a decrease in drought tolerance, with over-expressors wilting more rapidly in response to water stress.

1.2.3 ABA and its role in the drought response

ABA is the most important phytohormone involved in the drought response. It forms a long-distance signalling link between the root system and the above-ground plant tissues and is responsible for the initiation of a multitude of signalling pathways that

bring about increased drought tolerance through the initial of physiological, metabolic, transcriptional, and developmental adaptation mechanisms.

ABA biosynthesis

Mutant analysis has identified a number of genes involved in the biosynthesis of ABA. These mutants have a reduced size even under control conditions, although this may be the result of minor stress influences that remain even in controlled conditions (Zhu 2001). Under drought stress, these mutants rapidly wilt and die. ABA plays two essential roles in drought tolerance. The first and most immediate is in the control of stomatal opening. The other role of ABA is in the induction of gene expression to increase cellular dehydration tolerance.

ABA accumulation during drought stress is the result of both an increase in synthesis and a reduction in the rate of degradation. Several ABA synthesis genes have now been cloned and characterised and the biosynthesis pathway is diagrammatically represented in Figure 1.2.1.

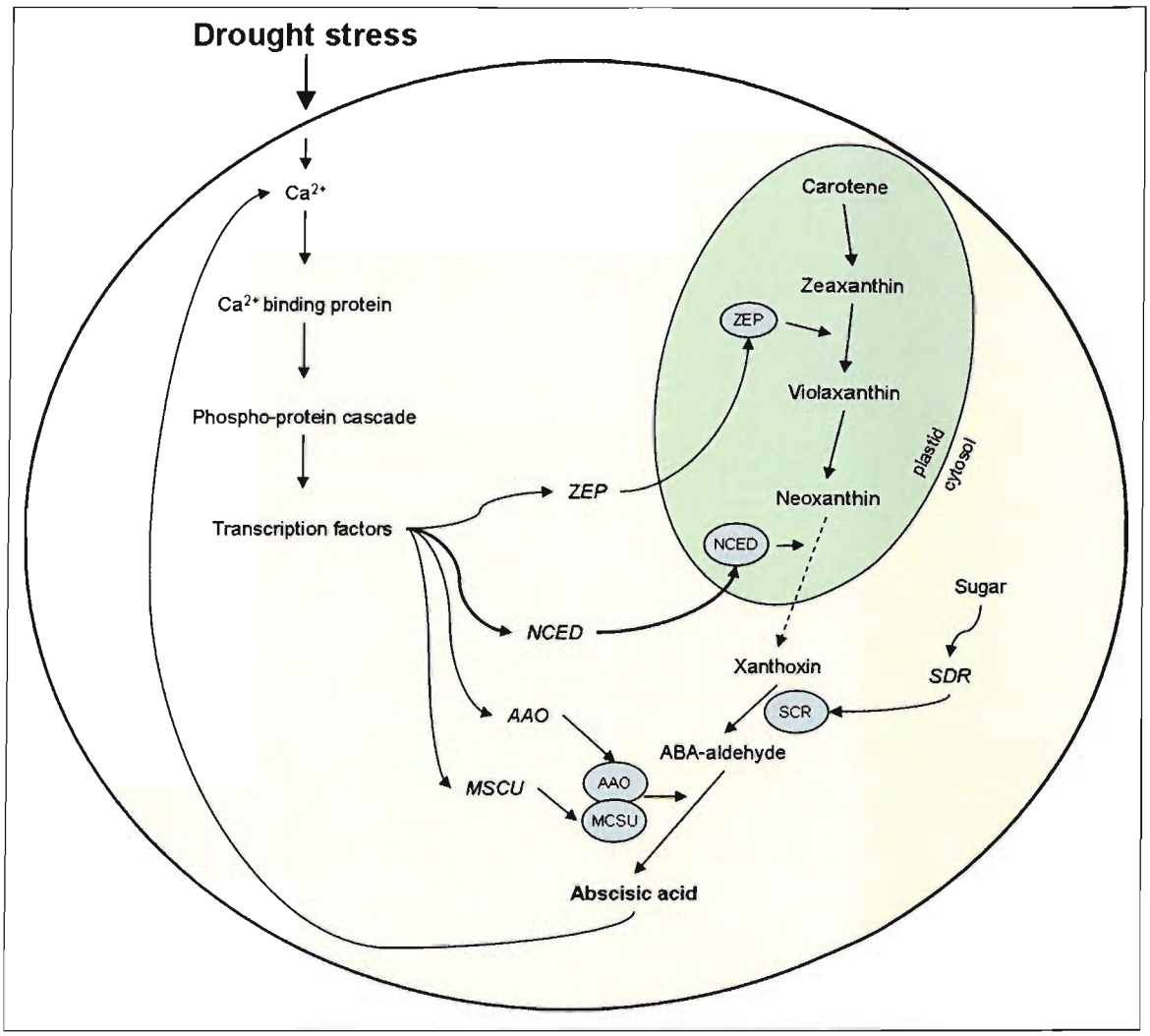


FIGURE 1.2.1 Diagrammatic overview of the ABA biosynthesis pathway and the effect of drought stress signalling on the pathway. Gene names are italicised. The lines linking NCED are in bold to indicate the strong up-regulation of this gene that occurs in response to drought. **NCED**, 9-*cis*-epoxycarotenoid dioxygenase; **AAO**, ABA-Aldehyde Oxidase; **MCSU**, MOCO Sulfurase; **ZEP**, Zeaxanthin Epoxidase.

ABA biosynthesis branches from the carotenoid biosynthesis pathway. Many of the genes involved in the biosynthesis are members of multi-gene families. Where family members serve identical roles, this may represent a means of stress and/or organ specific control of ABA biosynthesis, as well as their being induced by differing environmental stimuli.

Zeaxanthin epoxidase, which converts zeaxanthin to violaxanthin, a two-step epoxidation, was originally identified from the *aba2* mutant of *Nicotiana plumbaginifolia* (Bray 2002).

9-*cis*-epoxycarotenoid dioxygenase (NCED) is the enzyme that cleaves the carotenoid, 9-*cis*-neoxanthin, to result in the first C15 precursor of ABA, xanthoxin. Both putative ABA precursors 9-*cis*-violaxanthin and 9'-*cis*-neoxanthin are cleaved to xanthoxin by VP14. VP14 specifically cleaves xanthophylls in the 9-*cis* configuration (Schwartz *et al.* 1997).

There is an increase in the mRNA transcribed from NCED genes in several species in response to water deficit. *Vp14* gene expression is up-regulated in response to water deficit (Tan *et al.* 1997). Induction of NCED genes by water deficit stress has also been observed in tomato (Burbidge *et al.* 1999) *Arabidopsis* (Neill *et al.* 1998) and avocado (Chernys and Zeevaart 2000).

The final step in the pathway is catalysed by abscisic aldehyde oxidase (AAO). Four genes have been cloned in *Arabidopsis* (Seo *et al.* 2000); yet there are also additional genes in the *Arabidopsis* genome with significant homology to *AAO1-4*.

Biochemical studies have identified that the rate-limiting stage of the synthesis pathway is the reaction catalysed by NCED (Koornneef *et al.* 1998). It has been shown that *ABA1*, *ABA2*, and *ABA3* are up-regulated by drought but there is thus far no evidence of this up-regulation affecting protein levels (Zhu 2002).

ABA perception

In order to function as a chemical signal, the concentration of ABA in the near proximity of the guard cell must be detected. Detection requires a receptor in order to function. Despite many attempts to identify and characterise an ABA receptor, no candidates have yet been identified (Wilkinson and Davies 2002). There is, in fact, no

clear indication of whether the ABA receptor is located intra- or extra-cellularly, or both (Allan *et al.* 1994, Anderson *et al.* 1994, Schwartz *et al.* 1994).

ABA targets

ABA induces the rapid closure of stomata through alterations in the turgor pressure of the guard cells around the stomatal pore. It is known that K^+, Cl^-, and sugars are the chemicals utilised for the control of turgor pressure (Wilkinson and Davies 2002, Taiz and Zeiger 1998, Willmer and Fricker 1996).

The first observed stage of stomatal closure is a transient membrane depolarisation (Willmer and Fricker 1996, Taiz 1998). Schroeder and Hagiwara (1990) showed that this rapid depolarisation is associated with a rapid increase in the concentration of cytosolic Ca^{2+}. This increase in Ca^{2+} is the result of both Ca^{2+} influx and release from internal compartments. (Gilroy *al.* 1990) showed that the controlled release of caged Ca^{2+} was sufficient to trigger stomatal closure in the absence of any ABA signal. ABA also causes an alkalinisation of the cytosol of approximate 0.3 pH units (Irving *et al.* 1992, Blatt and Armstrong 1993) showed that this change in pH activates K^+ efflux channels.

Grabov *et al.* (1997) and Pei *et al.* (1997) showed that ABA also activates S-type anion channels in guard cells. The opening of these anion channels allows Cl^- and $malate^{2-}$ to leave the cytosol down their electrochemical gradient. This loss of Cl^- and $malate^{2-}$results in a long-term depolarisation of the cell membrane that in turn enables the sustained opening of K^+ efflux channels. Membrane depolarisation is also achieved via a Ca^{2+} inhibition of membrane H^+ ATPase (Kinoshita *et al.* 1995).

ABA also inhibits the light-induced opening of stomata (thus overriding the normal stomatal state). Stomatal opening is inhibited via the inactivation of K^+ influx channels, which are normally opened when the plasma membrane is hyperpolarised (by the H^+ pump).

Transmitting the ABA signal

Upon detection of ABA by the hypothesised ABA receptor protein, a signal must be transmitted from the receptor site to the various downstream components of stomatal closure. Ca^{2+}, IP_3, and H^+ are all now known to play roles as secondary messengers mediating the effects of ABA on stomata.

Ca^{2+} concentration can be raised in two ways; Ca^{2+} can be released from intracellular stores or can move into the cell from the external environment. Hamilton *et al.* (2000) and Pei *et al.* (2000) have both presented evidence for the presence of ABA-regulated Ca^{2+} channels in the plasma membrane through the use of patch-clamp techniques. In the case of Pei et al, it was shown that H_2O_2 is required for the activation of the Ca^{2+} that they identified, suggesting an additional role for H_2O_2 in the control of stomatal closure. Following observation by Gilroy *et al.* (1990) and Lee *et al.* (1996) that IP_3 triggers Ca^{2+} elevation in guard cells, other secondary messengers have subsequently been shown to play a role in Ca^{2+} elevation in guard cells. These include IP_6, sphingosine-1-phosphate, and cyclic ADP-ribose (Leckie, McAinsh *et al.* 1998, Lemtiri-Chlieh *et al.* 2000, Ng *et al.* 2001 respectively). It seems likely that the presence of these multiple signalling pathways represents a means of achieving stomatal closure from a range of environmental stimuli, each with a specific signalling pathway.

Cohen (1989) first proposed a role for protein phosphatases in the ABA signal pathway. Such a role comes as no surprise considering that protein phosphatases play a role in nearly all biological signalling systems. Such a role was established in work by Luan *et al.* (1993), Li *et al.* (1994), Thiel and Blatt (1994), and Schmidt *et al.* (1995), mainly through studies using kinase and phosphatase inhibitors. Biochemical studies conducted by Li and Assmann (1996) and Mori and Muto (1997) both identified an ABA-activated protein kinase (AAPK) from the guard cells of *Vicia faba*. Subsequently, (Li *et al.* 2000) have cloned this kinase and shown that it is involved in the regulation of the slow release anion channel by ABA. It was also shown that this regulation is independent of Ca^{2+}. As yet the counter phosphates for the AAPK has not been identified (Luan 2002).

Mutation analysis has identified two isoforms of protein phosphatase 2C (PP2C) that are involved in ABA signalling in guard cells. These two phosphatases were identified as ABA-insensitive plants and are named ABI1 and ABI2 (Leung *et al.* 1994; Meyer *et al.* 1994; Leung *et al.* 1997). Leung *et al.* (1994) and Meyer *et al.* (1994) have shown that the *ABI1* gene encodes a serine/threonine phosphatase. *ABI1* and another PP2C encoding gene from the same clade, *PP2CA*, (Saez *et al.* 2004) are induced by water deficit, low temperature, salinity, and ABA. Both are proposed to be negative regulators of the ABA response (Tahtiharju and Palva 2001). Pei *et al.* (1997) examined the stomatal response of *abi* mutant plants and found that the stomata of these plants remain open upon application of ABA. It was found that this was due to a lack of response by the slow anion channels. Armstrong *et al.* (1995) conducted an experiment using transgenic tobacco expressing a mutant *abi1* gene. In these transgenic plants, inward K^+ channels but not anion channels were affected suggesting divergent functions for the ABI gene between species. Blatt and Grabov (1997) have shown that the insensitivity of the K^+ channels is, in fact, due to a lack of response to the change in pH induced by ABA.

1.2.4 Drought-induced changes in gene expression

Bray (1993), Shinozaki and Yamaguchi-Shinozaki (1997), Zhu *et al.* (1998), Shinozaki and Yamaguchi-Shinozaki (2000), and Ramanjulu and Bartels (2002) offer reviews of genes reported to undergo expressional changes via ABA regulation. Bray (2004) reviews three major experiments using microarrays to study water-deficit stress response in the model plant *Arabidopsis thaliana* Columbia, identifying expression responses common across all three experiments. Posas *et al.* (2000) used a *Saccharomyces cerevisiae* microarray to examine global response to salt stress and found that roughly eight percent of the entire genome underwent an alteration in expression as a result of osmotic stress. Some of these genes would have been salt response specific but many investigations have found vast overlap between salt and

drought stress responses (Ishitani *et al.* 1997, Kasuga *et al.* 1999, Liu *et al.* 2000, Zhu 2001, Salekdeh *et al.* 2002, Xiong and Zhu 2002, Xiong *et al.* 2002, Zhu 2002).

The majority of genes involved in drought response appear to be involved in damage limitation and repair (Zhu 2001). These genes include osmolyte biosynthesis genes, late embryogenesis abundant (LEA)/dehydrins, detoxification enzymes, chaperones, proteases, and ubiquitination related enzymes.

As with many other gene expression responses, drought-induced expression can be broken down into early-response genes and delayed response genes. The early-response genes are induced within minutes of the drought stress being initiated. Their induction does not require the synthesis of new proteins. Early-response gene expression is very often transient and most encode transcription factors that activate the downstream expression of the delayed-response genes. Delayed response genes constitute the majority of genes undergoing expression changes in response to drought. They are usually induced within hours of the drought stress being initiated and their expression is typically sustained through the course of the drought stress and beyond.

ABA-dependant gene expression

It has been found that the ABA-response element (ABRE) is the functional element for ABA-regulated gene expression in maize, barley, rice, tobacco and *Arabidopsis*. The consensus sequence of the ABRE is RYACGTGGYR, where R is a purine containing base and Y is a base containing a pyrimidine. It is defined by the core element ACGT, which was first identified in the *Em* gene from wheat (Bray 1997).

The ABRE has been shown to be sufficient for ABA-regulated gene expression although in some genes association with a coupling element is required (Shen and Ho 1995).

A small family of ABRE-binding bZIP proteins has been isolated using yeast-one hybrid screens (Choi *et al.* 2000, Uno *et al.* 2000). These have been termed ABFs

(ABRE-binding factors). Choi *et al.* (2000) show that there expression is induced by ABA, cold, and drought stress. ABFs have been shown to play vital roles in activating stress response mechanisms (Kang *et al.* 2002). Kang *et al.* (2002) and Kim *et al.* (2004) showed that over-expression of ABF3 results in increased drought tolerance via a promotion of stomatal closure and the expression of genes involved in cellular dehydration tolerance. ABF3 over-expression additionally increased tolerance of cold, heat, and oxidative stress, showing the ABA is involved in multiple stress response and confers multiple stress tolerance.

ABA-independent gene expression

In addition to the ABA-dependent changes in gene expression, there is also a major transcription regulatory pathway that control ABA-independent gene expression in response to low temperature, drought and high salinity stress (Stockinger *et al.* 1997, Liu *et al.* 1998, Haake *et al.* 2002). The ABA-independent dehydration response is based on the dehydration-responsive element (DRE), which is the *cis*-acting element, and the *trans*-acting DNA-binding protein DREB1, have been shown to be involved in the regulation of drought, low temperature and high salinity-induced genes who's regulation is ABA-independent (Yamaguchi-Shinozaki and Shinozaki 1994, Jaglo-Ottosen *et al.* 1998, Kasuga *et al.* 1999, Gilmour *et al.* 2000). These two components are also often referred to as the CRT (C repeat) and the CBF (CRT binding factor).

Many DREs are ABA-independent for their initial induction but sustained induction requires the presence of ABA (Bray 1997). Transcription factors regulating the expressing of ABA-independent stress induced gene expression changes have been identified (Liu *et al.* 1998, Stockinger *et al.* 1997), and over expression of these induces the expression of cold/drought-regulated genes under control conditions and results in increased drought (and cold/freezing) tolerance (Jaglo-Ottosen *et al.* 1998, Kasuga *et al.* 1999). Three DREB1 genes have been identified in the *Arabidopsis* genome, arranged in a tandem array in an 8.7 Kb region (Gilmour *et al.* 1998, Medina *et al.* 1999).

1.2.5 Secondary messengers

A generic signal transduction pathway starts with signal perception, followed by the generation of secondary messengers. Secondary messengers can modulate intracellular Ca^{2+} concentration, often initiating a protein phosphorylation cascade that finally targets proteins directly involved in cellular protection, or transcription factors controlling specific sets of stress-regulated genes. Cold, drought and salinity have been shown to induce transient Ca^{2+} influx into the cell cytoplasm (Sanders *et al.* 1999, Knight 2000).Many recent studies have focused on unravelling the signalling mechanisms initiated during water deficit response (reviewed by Xiong *et al.* 2002, Shinozaki *et al.* 2003).

Several two-component histidine kinases have been identified in *Arabidopsis* (Urao *et al.* 2000), although none have yet been shown to be involved in stress signalling. It has been suggested that AtHK1 could be involved in osmotic stress signalling (Urao *et al.* 1999). In yeast and animals, mitogen-activated protein kinase (MAPK) pathways are responsible for the production of compatible osmolytes and antioxidants. These MAPK pathways are activated by receptors/sensors such as protein tyrosine kinases, G-protein-coupled receptors, and two component histidine kinases.

Osmotic stress was found to rapidly increase PIP_2 levels in cultured *Arabidopsis* cells (Pical *et al.* 1999, DeWald *et al.* 2001). Drought or salt stress also up-regulates the mRNA levels of certain PI-PLC isoforms (Hirayama *et al.* 1995, Kopka *et al.* 1998). This could increase the cleavage of PIP2 to IP3 and diacylglycerol, which are both secondary messengers. IP3 is known to cause a release of Ca^{2+} from intracellular stores (Sanders *et al.* 1999, Schroeder *et al.* 2001a). ABA is known to cause an increase in the levels of IP3.

In a genetic screen using the firefly luciferase reporter under the control of the stress-responsive *RD29A* promoter (Xiong and Zhu 2001, Ishitani, Xiong *et al.* 1997) isolated an *Arabidopsis* mutant *fiery1* (*fry1*) that exhibited an enhanced induction of stress-responsive genes under cold, drought, salt, and ABA treatments. Positional

cloning of the *FRY1* gene revealed that it encodes a bi-functional enzyme with both 3'(2'),5'-bisphosphate nucleotidase and Ins1Pase activities. *FRY1* is identical to the previously described *SAL1* gene that was isolated by its ability to confer increased salt tolerance when expressed in yeast cells (Quintero *et al.* 1996).

Accumulating evidence suggests that phospholipase D (PLD) is also involved in the transduction of stress signals. PLD hydrolyses phospholipids to generate phosphatidic acid (PA), another secondary messenger in animal cells that can activate PI-PLC and protein kinase C (English 1996). In guard cell protoplasts PLD activity mediates ABA-induced stomatal closure (Jacob *et al.* 1999). PLD appears to be activated by osmotic stress through a G-protein (Frank *et al.* 2000).

Drought, cold and salt stress all induce the accumulation of ROS such as superoxide, hydrogen peroxide, and hydroxyl radicals.

It is suggested that Ca^{2+}-dependent protein kinases (CDPKs) and the SOS3 family of Ca^{2+} sensors are major components in coupling this universal inorganic signal to specific protein phosphorylation cascades.
The *Arabidopsis* genome encodes at least 34 putative CDPKs (Harmon *et al.* 2001). A number of studies have shown that they may be involved in abiotic stress signalling (Urao *et al.* 1994, Tahtiharju *et al.* 1997, Hwang *et al.* 2000).

It is not yet clear how CDPK stress signal transduction is integrated with other signalling modules. Recent studies with neural cells suggest that calmodulin perceives local Ca^{2+} and activates a MAPK pathway to regulate target gene expression (Dolmetsch *et al.* 2001), although the connecting point between Ca^{2+}-calmodulin and the MAPK pathway remains unknown. Patharkar and Cushman (2000) identified a CDPK-interacting protein (CSP1) from a yeast-two-hybrid screen. CSP1 is a two-component pseudo-response regulator protein that could serve as a transcriptional activator, suggesting a potential role for CDPK in directly shuttling information to the nucleus to activate gene expression.

1.2.6 Conclusions

Much work has been carried out in recent years to uncover the multiple layers of signalling and gene expression changes involved in the drought / water deficit response and tolerance. Many parts of this signalling network have now been filled in and the increasing number of microarray results allows an increasingly comprehensive picture of gene expression changes induced by drought stress. However, some key gaps remain such as explaining the unexplained steps in the ABA biosynthesis pathway, identifying ABA receptors and identifying the signalling mechanisms that initiate down-stream signal transduction activation. New methods may need to be invented to capture these events.

Many of the genes detailed in the above discussion serve as ideal candidate genes for the investigation of drought response in other species and within QTL mapping populations. Comparison of cross-species responses in gene expression also offers unique opportunity to identify conserved and divergent regulatory mechanisms in drought response and this may play a role in explaining the genetic architectural differences that underlie species differing adaptation mechanisms to drought stress. This requires extensive use of various bioinformatics tools and is something that is only possible due to the sequencing efforts undertaken on rice, poplar and *Arabidopsis*.

1.3 Leaf development

1.3.1 Introduction

Leaves are, in principal, flat three-dimensional structures for the optimised collection of energy stored in photons of a specific wavelength, for the exchange of gasses between the external environment and the internal air spaces of the leaf, and for the perception of the external environment. The form taken by a leaf is the result of natural selection. Leaf shape also, then, leads to questions such as 'why is this leaf shape optimal to this particular evolutionary niche?'. In addition, flowers, and the floral organs, can be considered as modified leaves. In this way, an understanding of leaf development is essential to a full understanding of all seed plants.

As with any organ, leaves are an ordered collection of cells with those cells being arranged into specific tissue types that are located in specific regions and planes of the leaf. It follows from this obvious statement that a leaf, as with any organ, is the result of an ordered series of cell divisions, cell expansion and cell differentiation. The balance between cell division and cell expansion will determine the final dimension of the leaf (Kutschera 1992). It is commonly recognised that for this ordered arrangement to be achieved in any organ of any organism, some form of 'awareness' by an individual cell of it's position within the organ is required. Two principle means of achieving this 'awareness' are utilised. One is the defining of the three dimensions of the organ's structure and the other is cell-cell communication.

Until the last decade of the 20th century, work on dicotyledonous leaf development remained sparse. This was principally due to the complexity of that development and a lack of research methods suitable for its study. Unlike the view of organ development held by classical anatomy, in which the developmental process can be divided into clearly defined zones, dicotyledonous leaves undergo cell division and expansion at the same time and in the same region of the leaf.

It was the adoption and use of *Arabidopsis thaliana* as a model for the study of dicotyledonous plants, alongside the development of suitable genetic research techniques that provided the means to begin investigating and explaining the process of leaf development.

1.3.2 Early leaf development

Leaf primordia initiation

A leaf begins its development with the initiation of a leaf primordium within the SAM. In *Arabidopsis*, genes such as *WUS, CLV1, CLV2, CLV3, KAPP,* and *STM* determine whether a cell will remain as a stem cell in the SAM or whether it will proceed along the developmental pathway for organ formation (Fletcher *et al.* 1999, Jeong *et al.* 1999, Trotochaud *et al.* 1999, Willmann 2000, Becraft 2002, Brand *et al.* 2002, Gallois *et al.* 2002, Searle *et al.* 2003).

Once cells have been committed to proceed along a developmental pathway, *KNOX* genes then play a major role in regulating the function of the SAM. The first *KNOX* gene identified was *knotted1* (*kn1*) from maize, which encodes a homeobox gene (Volbrecht *et al.* 1991) that is expressed in all SAM cells except those that will progress to differentiate into lateral organs (Smith *et al.* 1992). Kerstetter *et al.* (1994) performed a phylogenetic study of *KNOX* genes and found that they sub-divide into two families; Class I and class II *KNOX* genes.

Six *KNOX* homologues have been identified in *Arabidopsis* (Lincoln *et al.* 1994, Granger *et al.* 1996) and they are referred to as *KNAT* genes. Ectopic expression of a class I *KNOX* genes of *Arabidopsis*, *KNAT1*, results in the formation of lobes and ectopic meristems in the leaf blades (Lincoln *et al.* 1994, Chuck at el, 1996). Over-expression of *STM*, another *KNAT* gene, results in the formation of highly disorganised shoot apex, with clusters of small, undeveloped leaf primordia.

Byrne *et al.* (2000) reported the discovery of a gene shown to play a role in controlling the pattern of expression of *KNAT* genes in *Arabidopsis*. The loss-of-function mutant, *as1*, has crinkled, asymmetric leaves, abnormally arranged leaf veins, and the presence of multiple mid veins. The *ASYMETRIC LEVAES* gene is a homologue of the *PHAN* gene of snapdragon, which was previously reported by Waites *et al.* (1998). *PHAN* was shown to encode a MYB transcription factor and the gene's product was suggested to suppress the expression of *KNOX* genes with this acting as a signal for cells to begin proceeding along a developmental pathway. Ori *et al.* (2000) have identified further genes involved in the regulation of *KNOX* genes. They confirmed the expression of *KNOX* genes in the leaves of *as1* and *as2* mutants and additionally identified the *SERRATA* gene as being involved in the control of expression of *KNOX* genes.

Establishment of polarity

The establishment of both proximo-distal and dorsiventral polarity in leaves is an absolutely essential step in the development of leaf shape and as such, it has received much attention in recent years (see Hareven *et al.* 1996, Bohmert *et al.* 1998, Waites *et al.* 1998, Sawa *et al.* 1999, Siegfried *et al.* 1999, Kerstetter *et al.* 2001, and McConnell *et al.* 2001).

Adaxialisation

The abaxial side of a leaf is characterised by the differentiation of the spongy layer and the dorsiventral arrangement of xylem and phloem elements. McConnell and Barton (1998) reported a mutation, *phb-1d*, that results in the adaxialisation of leaves. The strongly adaxialised leaves of *phb-1d* plants have rod-like leaves at their most severe to trumpet shaped leaves in less severe individuals. The *PHB* gene encodes a homeodomain-leucine zipper (HD-ZIP). In a later paper, McConnell *et al.* (2001) identified another mutant, *phv*, with a similar phenotype. The *PHV* gene also encodes an HD-ZIP. A third gene of this gene family, *REVOLUTA/INTERFASCICULAR FIBERLESS* (Ratcliffe *et al.* 2000), also leads to a similar phenotype in mutants (Talbert *et al.* 1995).

A mutation in the well known patterning gene *emb30* (*gnom*) was reported to result in cotyledons with an abnormal dorsiventrality (Shevell *et al.* 1994). Cotyledons of the mutant plants had a palisade layer on their abaxial side and spongy layers on the adaxial side. EMB30 is similar to a yeast gene that encodes a Sec7p-like protein. Sec7p has been shown to be involved in the transport of regulatory ligands in yeast (Shevell *et al.* 1994) and this has led to speculation that EMB30 might be responsible for the uneven distribution of a ligand proposed to activate the PHB gene (McConnell and Barton, 1998).

Additionally, PNH [*pinhead/zwille*] and AGO [*argonaute*] (Lynn *et al.* 1999), have been shown to have similar patterns of expression to PHB and are therefore proposed to play a role in the establishment of bilateral symmetry.

Abaxialisation

The *YABBY* genes have been identified as being involved in the adaxialisation of lateral organs (Siegfried *et al.* 1999). Ectopic expression of *YABBY3* caused the differentiation of abaxial cell types in leaves with no differentiation of adaxial cell types (Siegfried *et al.* 1999). Sawa *et al.* (1999) showed that *YAB* genes are expressed in the abaxial regions of lateral organs in wild type *Arabidopsis*.

Kerstetter *et al.* (2001) showed that a second gene family, *KAN*, is also involved in the regulation of abaxialisation of lateral organs. The mutant phenotype has abaxial leaf tissue converted to an adaxial state. Kerstetter *et al.* used a KAN-GUS fusion protein to visualise the wild type expression of the *KAN* gene and found it to be a mirror image of the expression of *PNH*.

Marginal meristems

The traditionally held view of how a rod shaped protrusion became a flat leaf primordia was that this resulted from activity of the meristems. The meristem is

established along the boundary of the abaxial and adaxial planes. However, it has long been thought that cell division in the marginal or sub-marginal regions of the leaf blade are required for the two-dimensional growth of leaves (Avery 1933).

In 1999, Donnelly *et al.* published a detailed examination of cell cycles in leaves of *Arabidopsis*. They used GUS under the control of the promoter region of a cyclin1 gene (*cyc1At*) as a marker for cell division activity. This cyclin is expressed at the G2/M phase of mitosis. Donnelly *et al.* found that the marginal meristems was active during the earliest stages of the development of the leaf primordia but that it ceased to be activate in subsequent development stages. An important finding of this paper was that leaf area development is dependant on the tissue specific control of the cell cycle. Donnelly *et al.* therefore suggest that the plate meristems might serve a function in the morphogenesis of leaves.

1.3.3 Leaf expansion

Genetic control of leaf expansion

The expansion of the leaf blade (lamina) is the result of both cell division and cell expansion. In monocotyledonous leaves, these processes can be isolated into clearly defined regions of the leaf (see Fiorani *et al.* 2000). In dicotyledonous leaves, the division and expansion of cells occurs without any defined spatial or temporal boundaries (Taylor *et al.* 2003) making this a much harder system to model.

Probably the first report of an *Arabidopsis* leaf mutant phenotype was made by Rédei in 1962 (cited in Tsukaya 2003). He described the *angustifolia* (*an*) mutant, which has narrow and thick leaves compared to the wild type. This phenotype was originally used as a visible marker for chromosome mapping, as were many early reported mutantations. It has subsequently been shown that the *angustifolia* mutation is specific to lateral organs and that the phenotype exhibited is the result of a defect in the elongation of cells in the transverse (width) leaf plane. No effect of cell number is seen. The mutant phenotype was exhibited in all cell types examined (Tsukaya *et al.*

1994, Tsuge *et al.* 1996). Tsuge *et al.* (1996) reported that parenchymatous cells are particularly strongly affected. Mutant type parenchymatous cells expand considerably less in the leaf width plane and have enhanced expansion in the leaf depth plane, as compared to the wild type. They therefore speculate that the *AN* gene is an important regulator of the polar elongation of leaf cells in the leaf-width direction specifically.

With the knowledge of the existence of a gene regulating cell expansion in the transverse plane, it was predicted that such a gene would also exist for regulating expansion in the longitudinal (leaf-length) plane. Tsuge *et al.* (1996) also reported a mutant with a polar defect in cell elongation in the longitudinal plane of lateral organs and was named *rotundifolia3-1* (*rot3-1*). Again, no change in cell number was observed as compared to the wild type. The double mutant of *rot3-1* and *an* shows an additive phenotype, suggesting that the two regulators act independently of each other (Tsuge *et al.* 1996).

Hanson *et al.* (2001) also reported that the over-expression of an HD-ZIP transcription factor, *ATHB13*, resulted in a similar phenotype in the cotyledons of the over-expresser to that of *an* mutants but only when over expressing mutants were grown on high concentrations of sugars. Unlike *an* mutants, it was only the epidermal cells of the *ATHB13* over-expresser that exhibited abnormal polar elongation. The possible role and function of *ATHB13* remains unknown.

The *ROT3* gene has been shown to encode the Cytochrome P450 CYP90C1 (Kim *et al.* 1998). This cytochrome has regions homologous to those of steroid hydroxylases. The majority of cytochromes so far characterised encode enzymes involved in the biosynthesis of brassinosteroids (Kim *et al.* 1998). Kim *et al.* (1999) examined the expression of the *ROT3* gene and found that does not exhibit organ specificity. However, over-expression of the *ROT3* gene resulted in plants with longer leaves than the wild type. No change was observed in the leaf width. No alterations were observed in organs other than leaves and flowers and it therefore appears that ROT3 regulates cell elongation specifically in the leaf length direction.

Mutations affecting only the number of cells along a leaf axis have also been identified. The *an3* mutant has narrow leaf blades or normal length and the *compact rosette* (*cro*) *4-1* mutant has short and narrow leaves (Tsukaya 2002). Both these mutants have cells of normal size with only the number of cells being affected. This is evidence for the regulation of both cell division and cell elongation along specific leaf axis being under genetic control and with both processes contributing towards leaf expansion.

The above mentioned combination of cell division and cell expansion contributing to leaf area expansion has also been characterised in more mature leaves of various *Populus* species in the lab of Gail Taylor. Workers in the lab have examined both the spatial and temporal patterns of leaf area expansion and have examined epidermal imprints from these leaves, measuring epidermal cell area both spatially and temporally (Taylor *et al.* 2003). In *Populus euramericana*, it was found that spatial patterns of cell production (a surrogate for cell division) can be defined within the leaf and that the rate of cell production varies over time. Work carried out on the parents of a *Populus* mapping pedigree (see Bradshaw *et al.* 1994) showed that leaf area is achieved differently in different species (Ferris *et al.* 2002). Ferris *et al.* (2002) additionally showed that QTL can be mapped for cell area. Cell area is a normally distributed trait within the F_2 progeny of the mapping population (Robinson 2004), showing that this is a quantitative, segregating trait and that the picture is not a simple one requiring just a few genes of major influence to determine leaf area.

Cortical microtubules and the cell wall

As is also the case in animals, the polarised expansion of plant cells is controlled by the organisation and orientation of cortical microtubules (Cyr 1994, Shibaoka 1994). The possible influence of cortical microtubules on leaf shape was shown by Burk *et al.* (2001). They examined a mutant, *fra2*, that has a defect in the gene for a katanin-like protein. In mutants, the establishment of cortical microtubules is delayed after cytokinesis and this results in alterations in the shape of organs, in which cells are

shorter but wider than in the wild type. Leaves of *fra2* plants have an altered leaf width: length ratio.

Conclusions

The genes involved in the control of leaf expansion and leaf shape are being discovered at a rapid pace, principally through the screening of mutants with altered leaf form. This has led to some major discoveries over the past decade with some of the key genes having been identified. However, the function of these genes in plant growth in response to stress has not been examined. It is know that abiotic stress, particularly drought stress, causes alteration in leaf structure, shape, and most importantly size and rate of expansion. It will be fascinating to see if the pattern formation genes that have been identified in controlling leaf development have their expression altered in response to drought stress and whether this is the mechanism through which altered leaf form is achieved. Studying such gene expression is very difficult, requiring the isolation of leaf primordia from the SAM. However, another approach is possible through the use of QTL analysis (which is discussed below). In this approach, the influence of any genes influencing leaf form, expansion etc would be identified as QTL and it may be possible to identify the genes underlying the QTL. Poplar provides as excellent model for the study of this process as it is highly drought sensitive and a mapping population is available (Family 331, see below) that shows segregation of drought response mechanisms meaning that QTL detection for these traits will be possible. Applying stress to effect changes in leaf characteristics as an excellent way of studying leaf development. Alterations in gene expression can be observed through the use of microarrays (see below) and candidate genes can be examined for a role in leaf development.

Explanation of leaf development is of immense commercial importance as leaves are the site of carbon assimilation and therefore will have major impacts on plant growth and development. This has been shown in the commercially important species, poplar (Robinson 2004, Rae *et al.* 2004). Therefore, identifying candidate genes involved in

the control of leaf development through perturbations introduced by the application of abiotic stress may play an important role in identifying targets for improvement of the biomass yield of this and other species.

1.4 Microarray Technology

1.4.1 Introduction

The rapid development of sequencing and computer technology has provided the means to sequence and begin the annotation of many important model organisms including *Haemophilus influenzae* (small, non-motile Gram-negative bacterium), *Saccharomyces cerevisiae* (yeast), *Caenorhabditis elegans* (soil nematode), *Vibrio cholerae* (bacteria causing cholera), *Drosophila melanogaster* (fruit fly), more than 30 microbial genomes, *Arabidopsis thaliana, Mus musculus* (mouse) and *Homo sapiens*. Work on the sequencing of the Poplar genome has now been completed and the annotation is currently being finalised.

Sequence information alone is of little use, revealing only that there are *x* number of genes and that they are at location x within the genome. To turn this research into something of use, it is essential to determine the function of all genes. This information will then provide a list of all the genes from which an organism is constructed. From there the challenge turns to understanding the control of those genes and how that control results in the organism under study, its development, and response to the environment. This requires investigation into: time and place of RNA expression during development; sub-cellular localisation and inter-molecular interaction of protein products; physiological response; primary DNA sequence in coding and regulatory regions; polymorphic variation within a species or subgroup; and the subsequent arrangement of that data into regulons (a regulon is defined as a group of genes under similar expressional control).

The sheer magnitude of the problem means that traditional gene-by-gene approaches are not sufficient and new methods and technologies have been invented to provide a high-throughput means of analysis. These high throughput gene analysis methods are becoming more and more important in many areas of biological and biomedical

research as well as in clinical diagnostics (particularly in the field of cancer diagnosis), providing biologists with a more global view of biological processes at the genome level.

Microarrays allow the parallel screening of gene expression levels for (potentially) all genes from an organism either at a particular time or in response to a treatment. This ability has changed the way in which research is being carried out. Whereas a northern blot, or other equivalent techniques, are highly directed investigative methods requiring often substantial prior knowledge, microarrays allow purely explorative investigation to be carried out (commonly referred to as global expression analysis). This has brought with it many problems, many of which remain unresolved at present. However, this technology is now firmly established as a valuable research tool and its use will soon be as routine as the now ubiquitous northern blot.

The invention of the microarray technique was an inevitable progression from the northern blot, which is now a well established and accepted laboratory tool. All that was required to make the progression was the application of knowledge and technologies developed in other industrial areas (such as the production of microchips) that had developed means of accurately directing the application of small amounts of material onto a substrate. The development of microarrays has mirrored that of the micro-processor; they have rapidly become smaller, contain an higher density of spots, are cheaper to produce and are more efficient. Although the development of the technique was relatively simple, its implications are far reaching and have caused a major wave through nearly all fields of biological research.

There are a number of variations on the microarray principle. Here, the focus is on the use of cDNA arrays as this is relevant to the research of this thesis. The other principle array technology is that of Affymetrix. A review of array technologies can be found in Schena *et al.* (1995), Kehoe (1999), and Deyholos and Galbraith (2001). Although microarrays are typically used to produce an mRNA profile, the technology can be adapted to investigate genomic DNA (therefore becoming genomics as opposed to transcriptomics). In such an approach, genomic DNA is restricted using a combination

of restriction enzymes. Those restriction fragments are then cloned and stored in vectors and are subsequently spotted on to an array. These arrays are then screen against restricted genomic DNA. For an example of this application see Nouzova *et al.* (2001).

As is now the accepted standard, the immobilised cDNA is regerred to as the 'probe' and the labelled cDNA as the 'target' (Phimister 1999). A diagrammatic overview of the microarray process of provided in Figure 1.4.1.

1.4.2 Production of cDNA Microarrays

Probe Production

In principle, a microarray is a spotted array of unique Expression Sequence Tags (ESTs) that have be immobilised onto a substrate. The production of the required EST libraries will not be considered here but an introduction to their production can be found in Brown (2002). The selection of ESTs to be used requires careful planning when producing an array (Churchill 2002). If producing an array for general screening and exploratory investigation, the ideal is to have as wide a coverage of the genome as is possible with no bias being introduced as to the type of EST selected. If, however, the research is to be more directed (e.g. how does drug A affect tissue types B-E) then careful selection of the relevant ESTs will yield the most productive results. Alternatively all members of a super genes family could be spotted onto an array to enable expression profiling under various treatments. As with all biological experiments, the set up of the experiment will also affect its analysis. The methods used to statistically analyse data should always be kept in mind when designing all aspects of an experiment.

Once an EST collection has been produced and the relevant ESTs selected from it, these are then amplified and quality control checked to ensure the presence of a single amplicon per probe as well as to estimate concentration.

Array Spotting

An array is then produced by the robotic deposition of the probes in a regular pattern onto a suitable substrate. For low density arrays (<36000 elements, Deyholos and Galbraith 2001), nylon filters can be used. At higher densities, the deformability of nylon filters prevents printing at higher densities. Instead, a solid substrate must be used. An array is considered to be high density when the elements have a centre - centre spacing of < 0.5 mm (Schena *et al.* 1995). Typically, glass slides of similar dimensions to those of a typical microscope slide are used as the solid substrate. These slides are coated with a hydrophobic layer to increase probe adhesion and maintain element integrity. Elements printed on glass slides can have a spacing of as little as 60 μm using current technologies. Glass is a particularly suitable substrate due to its low fluorescence and porosity. The latter minimises hybridisation volumes while the former is an important factor in the later stage of image acquisition.

Typically, a spotter contains a head that houses a number of 'pins'. These pins are usually slotted in order to collect probes through capillary action. Each slotted pin holds a large enough volume to deposit approximately 400 elements. In most spotters, a number of glass slides are arranged on a vibration-free surface. The elements are then deposited onto the slides in a set pattern and in a set order. This can introduce batch and slide effects. For example the last slide 'printed' often has less element material deposited onto it. When the pin is empty, it is either refilled if it does not have the capacity to spot all required elements per probe in one run, or is washed and filled with the next probe. After all probes have been deposited, the slides are dried and washed to remove unbound material. The most commonly used spotters now have heads with 48 pins, allowing them to spot 100 slides in 17 hours. Any dust or vibration will lead to variation in the size and shape of the spotted elements and eliminating these influences is of vital importance. Other environmental factors such as temperature and humidity also greatly affect spot quality and must be tightly controlled.

A number of variations exist on the above outlined spotting process. The most appealing of them deposits probes onto the array substrate using the same technology

that is utilised in an inkjet printer to deposit drops of ink onto a sheet of paper. This technology is rapidly developing and should eventually lead to production of easy to use 'array printers' requiring little specialist knowledge to operate them. It is now also possible to purchase ready-made oligonucleotide collections suitable for array spotting that containing (near) full genome coverage of most organisms for which a genome sequence is available (Qiagen is the principle supplier of such oligonucleotide sets).

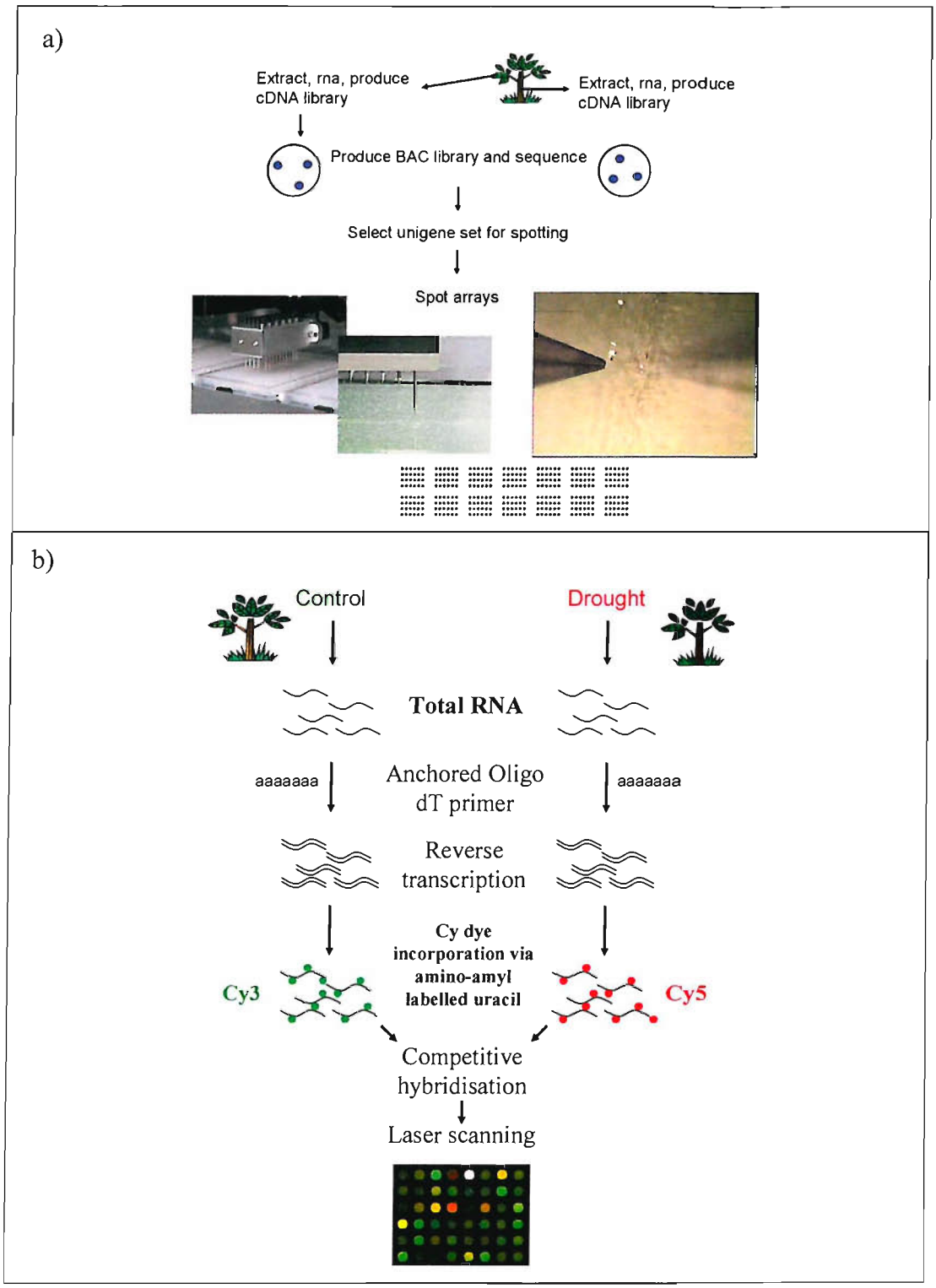


FIGURE 1.4.1 Overview of the microarray process **(a)** construction of a cDNA library and microarray, **(b)** process of hybridising and scanning cDNA microarrays (modified from talk by James Tucker, Southampton University, Uk).

1.4.3 Target Production

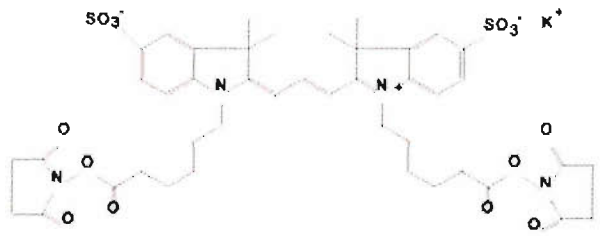
For the production of the labelled targets, RNA is extracted using a suitable protocol. For the extraction of Poplar RNA, a modified version of Chang *et al.* (1993) is typically used (e.g. Andersson *et al.* 2004, Moreau *et al.* 2005). This protocol extracts total RNA from plant tissues with high phenolic and secondary compound content. Other protocols commonly used extract poly(A)+ mRNA through the used of beads coated in oligo dT primer. The method of dye incorporation varies depending on the method used. RNA amplification can be used where only small amounts of RNA are available.

After RNA extraction, a reverse transcription is performed in order to obtain a cDNA copy of the RNA population this cDNA will be an antisense copy of the RNA). Either during this reverse transcription stage (direct incorporation) or as a later stage (indirect incorporation), a fluorescent dye is incorporated as a label. The most commonly used dyes are the cyanine dyes Cy3 and Cy5 (Amersham Pharmacia Biotech, Uppsala, Sweden). The structure of the Cy3 and Cy5 dyes can be seen in Figure 1.4.2.

Typically, for a cDNA array, two targets are hybridised simultaneously (competitive hybridisation), each one being labelled with a different dye flore. In the sample in which a gene is more highly expressed, more labelled cDNA will be present and therefore will be more highly represented on the array.

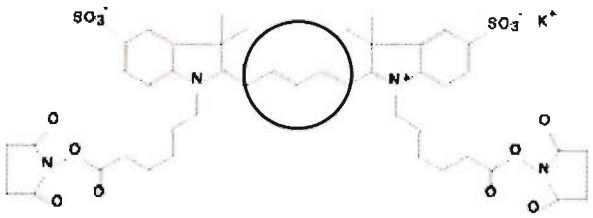
Hybridisation can either be performed manually, under a coverslip, or in an automated hybridisation station, in which the slide is placed into a sealed chamber and the labelled targets added to the chamber. Hybridisation typically is carried out over a 14 hour period. After this period, the slides are washed to remove any unbound target material. Various factors can be adapted during the pre-hybridisation, hybridisation, and washing stages to alter the stringency of binding in order to maximise the efficiency and specificity of the assay. Conditions must be optimised as too low a stringency can lead to a high levels of cross-hybridisation, therefore rendering results

inconclusive; too high a stringency will lead to low intensities, making quantification of assay problematic.



CY3™ REACTIVE DYE

ABSORPTION – 552NM
EMISSION - 568NM
EXTINCTION - 130000



CY5™ REACTIVE DYE

ABSORPTION - 650NM
EMISSION - 667NM
EXTINCTION - 200000

FIGURE 1.4.2 Chemical structure and physical properties of the Cy3 and Cy5 scribe dyes. Cy5 can be seen to contain one extra C bond (indicated above), which accounts for its slightly different chemical and binding qualities. adapted from Amersham Cy dye literature.

1.4.4 Scanning the Array

After hybridisation the amount of dye is quantified. A diagrammatic representation can be seen in Figure 1.4.1(b). The first stage of this process is to scan the array using a laser scanner. Laser scanners are particularly useful as they have a very high resolution and can accurately excite the fluorescent dyes used. A laser is focused at each point on the array with the excitation wavelength of each dye. The amount of light emitted at the emission wavelength of the dye (the dyes are referred to as channels) is then recorded. Scanning is either a one pass or two-pass process depending on the scanner being used. The information from the scanner is transferred to a computer, converted into pixel-intensity data ans stored in a RAW data format (typically a 16-bit greyscale TIFF [Tagged Image File Format] image). The size of the image file obtained is dependant on the resolution at which the array has been scanned: This is determined by the density of probes on the array; the higher the density, the higher the resolution required.

Common settings that can be adjusted when scanning are the laser power and the photomultiplier tube (PMT) voltage (gain). These two settings affect the amount of signal returned from each spot. At this point settings are adjusted to return an even signal from both dye channels at a scanning intensity that balances the signal: noise ratio, number of observable spots and number of saturated spots (those returning a signal beyond the range of the scanner).

1.4.5 Image Analysis

After obtaining a scanned image of an array, the spot elements must be identified. This is largely an automated process that is performed by image analysis software, although the degree to which human intervention is required varies between different software packages.

Typically, a grid is applied to the array that defines the layout and spacing of the spots with this grid serving as the starting point for spot location. Once the spots have been identified, pixels must be segmented into those that actually constitute the spot and those that constitute background (i.e. not a spot). A number of alternative methods exist for both spot location and pixel intensity quantification. The oldest method plots a histogram of pixel intensity (Chen *et al.* 1997a). A mask is chosen around each spot and a histogram is plotted of the intensity of pixels within the mask. A cut-off value is then defined for background signal level. All pixels above that level are classified as foreground and pixels below that intensity are classified as background and excluded from further analysis. The software package QuantArray (QuantArray Analysis Software, Perkin Elmer Life Sceince, Boston, USA) utilises this method. This method is computationally simple but can lead to over-estimation of foreground pixel intensity and under-estimation of background pixel intensity, and foreground pixels are not necessarily connected.

Other methods have been designed to locate spots based on the assumption that spots are composed of connected groups of foreground pixels. This is implemented by fitting a circle of constant diameter to all spots. The diameter of this spot is then allowed to be resized by a specified amount above or below the set value to allow for the inherent imperfections of the spotting process. GenePix (GenePix Pro, Axon Instrument Inc, California, USA) is an example of a software package utilising this method.

Two additional methods that do not assume circularity of spots are the watershed method (Beucher and Meyer 1993) and that of seeded region growing (Adams and Bischof 1994). Both methods require starting information about pixels or seeds. Adjoining pixels are added to the spot until adjacent pixels appear to be progressively less intense. The software package Spot (Adams and Bischof 1994) utilises the seeded region growing approach.

After locating spots, foreground intensity is calculated. This is usually estimated as the median pixel intensity for the spot with the assumption that this is then directly proportional to the number of bound dye molecules (and thus the number of cDNA molecules). The median is used in preference to the mean pixel intensity due to the resilience of the median to outlier values.

There are various methods of calculating background intensities. One method considers pixels that are outside of the spot but within a boundary box. Another method, used by the QuantArray software package, considers an area between two concentric circles outside of the spot mask. This method has advantages in that it does not use the pixels immediately around the spot boundary. These can be problematic with certain feature finding methods (particularly the seeded region and watershed methods). Another method, used by the GenePix software package, is to examine the 'valley' regions of the array, which are the background regions furthest from the spots. Yang *et al.* (2002) carried out a comparison of the various methods for segmentation (spot identification) and background calculation and observed that the background calculation method had the greatest effect on the calculated log-ratios (see later) of intensities. They found that the software package SPOT provided the mostoptimal results. This package uses another method of background calculation called morphological opening, the details of which can be found in Yang *et al.* (2002).

After spot location has been performed, human visual inspection is still currently required to check for misidentified spots and artefacts such as salts and dust that have not been washed from the array. Salts This visual inspection is open to bias and the results differ between different researchers. Due to this, effort is being put into improving the current automated methods to make them reliable to a degree where human intervention will not be required.

One further quality control check that can be applied to spots is a mean: median filter. A 'high quality' spot will have a similar mean and median and any spots that do not can be flagged as bad and excluded from analysis. In addition, spots with saturated pixels can be excluded, as these spots yield unreliable estimates of intensity. It has

been estimated that any spot containing more than 10% saturated pixels will have a skewed mean (pers. comm. Andreas Sjödin). The easiest method for carrying out these additional filtering stages is to write a script in a statistical language such as R, which is an open-source implementation of the language S (see later).

After background intensities have been obtained, the following calculation is performed: $R = Rf - Rb$ and $G = Gf - Gb$, where R stands for adjusted red value, Rf for Red foreground and Rb for red background (G defines the green channel values). Until recently, it was common practise to subtract background with R and G values used for further analysis. However, as background is often over estimated, background subtraction increases variance error, and many arrays now have low and even background levels, opinion is shifting towards using Rf and Gf as the values for further analysis, providing arrays pass basic quality control checks (Wu 2002a).

After these stages, data is typically output in the form of a tab delimited text file containing, at minimum, the intensity data for each spot in each channel along with a spot identifier.

1.4.6 Visual Inspection

In addition to the above data collection processes, it is advisable to perform a number of visual inspections, both of the array image itself and of the data colleted. A number of diagnostic plots are now available to give an overview of the raw data and to allow appropriate selection of a normalisation method (see later).

The first inspection that should always be carried out is a simple (yet essential) visual inspection of the array images collected. This is usually best done by creating a false-coloured image representation from the scanned TIF files, as greyscale images are hard to examine visually. The Cy3 channel is typically coloured green and the Cy5 channel red with the green / red value of the spot representing spot intensity. These two images are then combined to create an overlay image in which spots having equal

intensities in both channels are coloured yellow, spots with a higher Cy3 intensity are green and those with higher Cy5 intensities are red. These images offer easy identification of hybridisation artefacts (e.g. salt spots and dust), spot uniformity, evenness of hybridisation (to be assumed if your array layout is random) and background levels. A 'quick glance' assessment of the number of differentially expressed spots is also obtained. If most spots are not yellow (on a global gene array), this serves as an indication that a problem exists.

Until Dudoit *et al.* (2000), the most common graphical display of numerical data had been a scatter plot of log$R$ verses log$G$ values (Figure 1.4.3a). This plot is simple to produce, however, the very strong correlation that exists between the two channel intensity values invariably masks those features of the plot that are of interest (one is interested in the deviation of points from the diagonal). Dudoit *et al.* proposed an alternative plot that has been rotated through 45° and the axes have been re-scaled to allow easier identification of fluctuation from the horizontal (Figure 1.4.3b). This plot is know as the MA plot where $M = \log_2 R/G$ i.e. ratio and $A = ((\log_2 R) + (\log_2 G)/2)$ i.e. sum intensity. M is plotted on the vertical axis and A on the horizontal. This plot makes it easier to identify non-linear relationships in the data. This plot is also referred to as an RI plot, for ratio intensity.

Two other plots are also commonly examined. The first is the box plot. Box plots are typically plotted for each print tip group within and array and for each array across a set of replicates. These offer a very good indication of consistency and variance between the units. The other common plot is the spatial plot (Figure 1.4.4). A spatial plot of background and foreground intensities can be constructed in order to identify problematic areas of an array. Various styles of plot are available within the R software package.

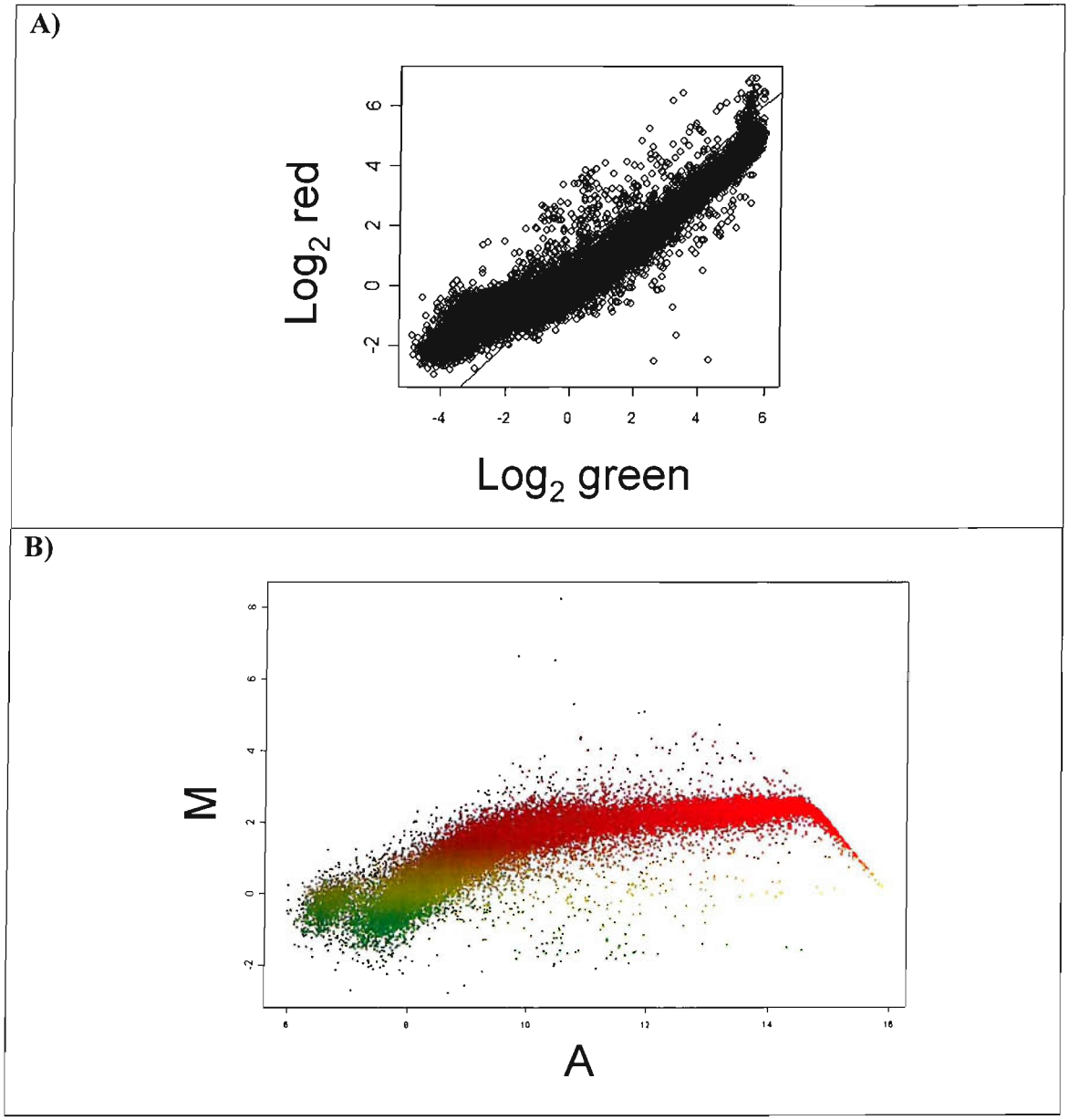


FIGURE 1.4.3 (A) $\log_2$ red intensity (cy5) plotted against $\log_2$ green intensity (cy3), **(b)** the same data represented as an ma plot where m = $\log_2$ rg and a = (($\log_2 r$) + ($\log_2 g$)/2). data are from a self x self test array. the same RNA sample from a mature *P. trichocarpa* leaf was used to label both dye channels. the hybridisation was conducted at Umeå University, Sweden using the pop1 array (see http://www.upscbase.db.umu.se/ for details of array construction).

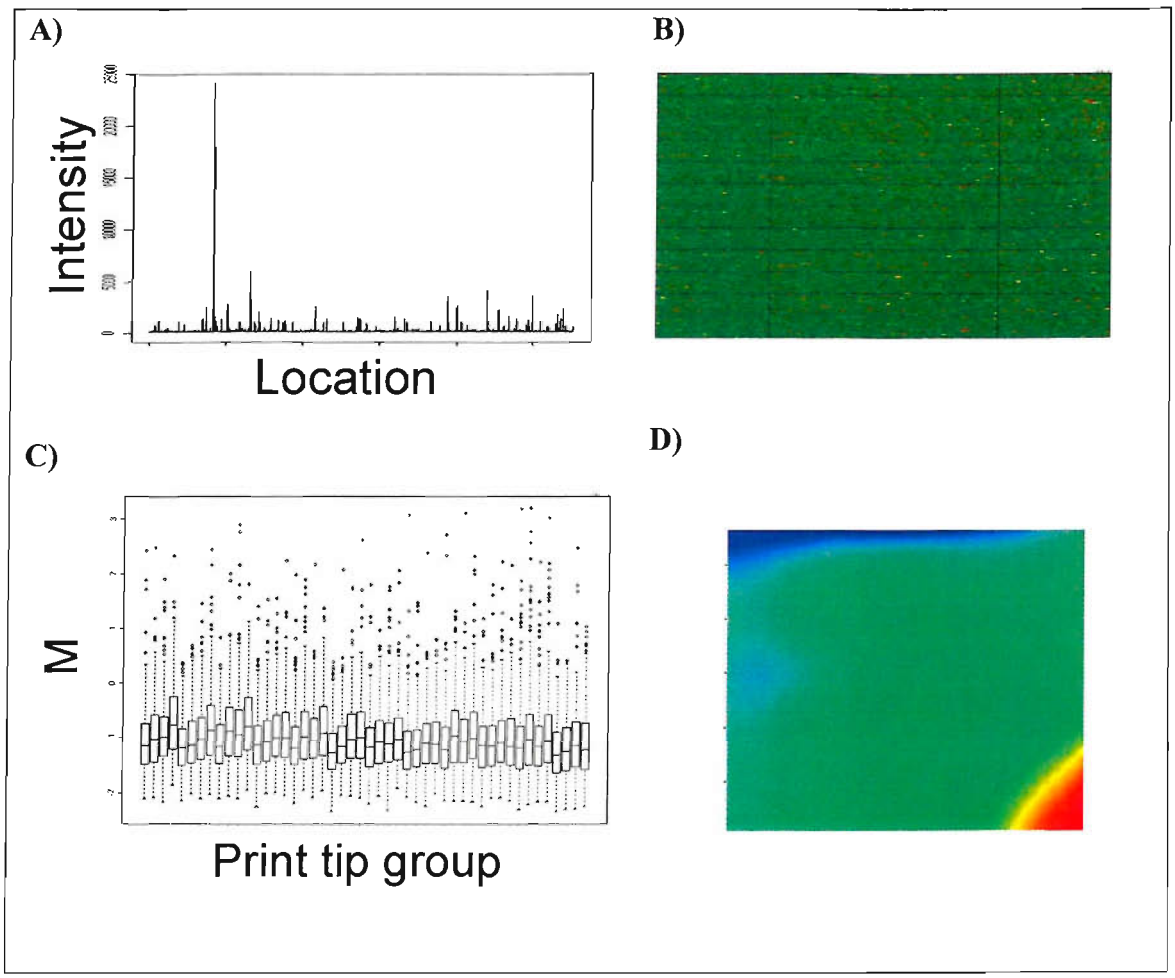


FIGURE 1.4.4 Example quality control plots **(a)** background intensity in the red (cy5) channel. the y axis represents location along the array surface, **(b)** false-colour representation of the foreground expression ratio calculated as m values where $m = \log_2 rg$, **(c)** box plot of foreground m values for each print tip group, **(d)** false-colour representation of the difference in intensity value of the red and green foreground channels, calculated as r-g.

Data are from a self x self test array. the same RNA sample from a mature *P. trichocarpa* leaf was used to label both dye channels. the hybridisation was conducted at Umeå University, Sweden using the pop1 array (see http://www.upscbase.db.umu.se/ for details of array construction).

1.4.7 Normalisation

There are a vast number of stages at which bias can be introduced during the microarray method. These can be from the spotting of the array or any of the subsequent wet (RNA extraction, cDNA synthesis, dye coupling, hybridisation) or dry (scanning, spot detection, background correction, intensity estimation) stages of the process. A bias introduced at any of these stages will interfere with the detection of biological differences between samples (i.e. they are all sources of noise).

The most universal bias is due to differences in the labelling efficiencies and scanning properties of the two dyes used for probe labelling. This is often compounded by the fact that different scanner settings are used for the two channels. Other common biases result from spatial variations across the array, uneven hybridisation, and differences between slides. Any change in stringency will also introduce biases.

A good method for the assessment of whether array data will require normalisation is to perform a self x self test, where one sample of RNA is split into two aliquots with one aliquot labelled with Cy3 and one with Cy5. It is expected that no differential expression will be observed and that a perfect correlation will exist between the two dyes channels. However, there will frequently be a clearly visible dye bias, with the red intensities being lower than the green intensities. Clearly this is not a biological result and would lead to the introduction of erroneous conclusions being introduced when assessing differential expression of samples from different sources. As well as a dye bias, it is often observed that certain areas of an array have higher intensities than other and that differences exist between arrays. Another clearly observed pattern is an intensity dependence (Figure 1.4.5).

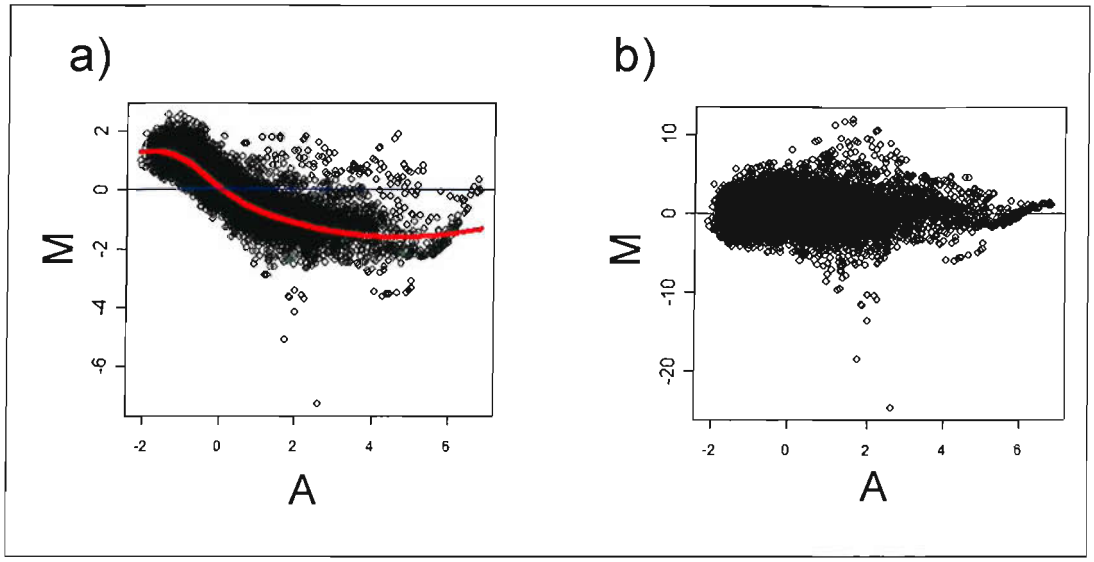


FIGURE 1.4.5 (A) MA plot showing strong intensity dependence. red line represents the loess (locally weighted linear regression) fitted curve, **(b)** the same data after the loess normalisation has been applied.

$M = \log_2 rg$, $A = ((\log_2 r)+(\log_2 g)/2)$.

Normalisation is usually carried out first within an array and then between arrays. The 'traditional', most basic normalisation technique assumes that any red-green bias is constant across the log-scale of intensities and across the array. Intensity values are then corrected by subtracting a constant c to give normalised values: $M = M - c$. The global constant is typically the median spot intensity, although many alternatives exist. Kerr *et al.* (2000) and Wolfinger *et al.* (2001) have proposed the use of ANOVA (Analysis of Variance) models for normalisation. The end result of this is very similar to the subtraction of a global constant. Global normalisation methods are still the most widely used despite the fact that there is extensive evidence of spatial and intensity dependant biases.

In order to deal with these spatial and intensity dependencies, a method that allows c to vary between spots in an intensity-dependent manner is required. If $c(A)$ is the height of the curve at each value of A, M values are normalised by subtracting this curve: $M = M - c(A)$. The curve is estimated using a robust scatterplot smoother, for

example locally weighted regression (LOESS or LOWESS) (Yang *et al.* 2002). Other intensity approaches have also been proposed and are explained in Cui *et al.* (2002). At present, LOWESS transformation is the most widely accepted and used non-linear transformation method although opinion is constantly shifting in the field of microarray analysis and examination of the most up to date literature will be required to assess recent advances and current opinion.

In order to account for differences across an array, it is also possible to use a different curve for different regions of the array. This is usually implemented based on print-tip groups.

When comparing multiple arrays, a cross-array normalisation is usually required due to differences in starting RNA quality, labelling differences or scanner setting alterations between arrays. This is achieved using a simple scaling of the M values from the series of arrays so that each array has the same median absolute deviation.

The above methods of normalisation utilise all of the spots on the array. An alternative is to base normalisation on a subset of spot that have been included specifically for the purpose of normalisation. The first proposed approach to achieving this was to use housekeeper genes. However, these often show a sample bias and were extremely hard to identify for plant species, although recent work is progressing the availability of plant control genes (Czechowski *et al.* 2005). Such housekeeping genes do not indicate any intensity-dependant bias as they are generally highly expressed at all times. An alternative approach is to use a titration series of spots that are replicated across the array. The set of Lucidea controls used on many microarrays are an example of such a titration series. These controls are artificial genes that generate pre-determined signal intensities that do not change across samples or experiments and are species independent (although work carried out at Umeå suggest some degree of cross-hybridisation with Poplar material). Suitable normalisation methods utilising such titration gradients have been developed. Essentially, normalisation is based just on those spots rather than on all spots. See Hill *et al.* (2001) for example. If exploratory

plots suggest strong differences between array regions then a sub-array (local) loess normalisation can additionally be applied.

1.4.8 Data Analysis

The below discussion presumes that a global expression array is being analysed as opposed to an array spotted with pre-selected genes of interest.

There are a number of approaches to the analysis of microarray data. The choice is determined by the question being asked. There are three principle approaches; that of clustering, examination of differentially expressed genes and expression profiling (guilt-by-association functional classification). These can usually be broken down into applications. Clustering methods are most often associated with time-course experiments, differential expression with treatment-response type experiments and expression profiling is often used to identify cancer cell , tissue type or disease type.

Transformation

After collection of the raw red and green values and any background subtraction has been carried out, a preliminary stage to data analysis is the transformation of the raw data values. For practical reasons, this step is usually carried out before the plotting of quality control checks as these plots typically use transformed values (as in the MA plot). Data transformation is used to achieve favourable statistical properties such as additivity and stabilisation of variances (Sokal and Rohlf 1995). The most commonly applied transformation is a logarithmic one, typically to base 2 for array data. Recently it has been questioned whether the logarithmic transformation is the most suitable for microarray data. Other proposed transformations are the shift-log transformation, the lin-log (linear-logarithmic), variance stabilising transformations towards which current emphasis is falling (VSN, Rocke and Durbin 2001) and the arcsine transformation. A review of these and analysis of their performance can be found in Cui *et al.* (2002) and briefly in Wu *et al.* (2002).

1.4.9 Differential Expression

The identification of differentially expressed genes is the priciple use to which microarrays are put. However, it is a challenging one. Firstly a suitable statistical test must be chosen that can indicate the evidence for differential expression. After indication values have been calculated, a critical value must be selected to allow grouping of those genes that can be considered significantly differentially expressed. This typically yields a list containing in the region of 100 genes of interest and these must then be scrutinised. In many experiments using microarrays, the purpose is to identify a set of candidate genes for further study (post-array evaluation and confirmation). The aim is usually to identify the key 10-20 genes for follow-up experiments. Selecting these from the 100 or so identified as being significantly differentially expressed is a crucial step of the data analysis and can be extremely time intensive.

Fold change and standard deviation

By far the simplest method for differential expression identification is to consider any gene with greater than an arbitrarily defined (typically 2X) fold change to be differentially expressed. This is easily achieved by plotting an MA plot and drawing horizontal lines at +1 and -1 (on a $\log_2$ scale) and subsequently creating a list of all genes above (up-regulated) and below (down-regulated) those lines. This is the basic method used by the Genespring software (Genespring, Silicon Genetics, Redwood City, USA) where a Log / Log scatterplot of raw and control intensities is used and a line draw along the diagonal corresponding to the specified fold change level.

An alternative to the fold change is to calculate the standard deviation of the distribution of $\log_2$ ratio values and selecting genes that have a ratio of less than or more than a selected number of standard deviations from the mean (Quackenbush 2002). Again, these cut-off values are easily visualised as parallel horizontal lines on an MA plot. This method is essentially the equivalent of performing a z test on the data (Quackenbush 2002). However, due to differences in the variability of data across the

intensity range, this approach is also inadequate as it stands a large chance of misidentifying genes as being differentially expressed at low intensities and of missing interesting genes at higher intensities.

The above approach has no logical grounding or justification and has been shown to be inadequate as a means of detecting differential expression. Such a method includes no estimation of whether the detected differential expression is a true result or a statistical fluctuation (Baggerly *et al.* 2001). This is currently the most contentious area of microarray data analysis with no clear consensus having emerged to date, although, reassuringly, more recent publications are making ever-improved use of statistically robust methods for differential gene identification (for example, Rensink *et al.* 2005).

The *t* test

The classical *t* test is a simple statistical means of selecting differentially expressed genes (Cui and Churchill 2003). However, Cui and Churchill point out the limitations of the *t* test. To overcome some of these limitations, various modified versions of the *t* test have been proposed. The SAM (Storey and Tibshirani 2000) *t* test (*S*-test) aims to stabilise the gene-specific variance by adding a small, positive constant to the denominator of the gene-specific *t* test (Tusher 2001).

Another now commonly used statistic is the *B* statistic proposed in Lonnstedt and Speed (2002) and implemented in the LIMMA package of R (Smyth 2004). This test allows for gene specific variances and also combines information across genes using the Bayes method. Pan (2002) discusses additional modifications to the *t* test.

A commonly used approach is to combine the methods mentioned so far into a single-combined script that outputs a plot with those spots that meet either one (fold change), two (fold and *t* test), or three (fold, *t* test, and *B* test) criteria identified and later output as a list of genes. Genes meeting all three criteria are considered to be robustly-

differentially-expressed. The term robust is used in the appropriate manner as applied to and defined for statistical analysis and as defined in Zar (1999).

ANOVA Based Analysis

A more statistically 'sound' method of analysis involves the use of a mixed- or fixed-model ANOVA test as proposed by (Kerr 2000), and discussed further in Kerr and Churchill (2001). Similar models have also been proposed by Pritchard (2001) and Wolfinger *et al.* (2001). The ANOVA model is applied to transformed data (typically $\log_2$ transformed) of raw intensities. Such a model has particular strengths in identifying sources of variation in the data that are attributable to factors other than differential expression. In this way, the ANOVA model is being used to normalise the data, as was mentioned above. Another advantage of this method is that data are not converted into ratios, which involves a loss of information. F-test values are computed for dye-effects, spot-effects, array-effects and biological-effects.

Wu *et al.* (2002) describe a freely available package called MAANOVA, which is a collection of functions implementing ANOVA-based analysis of gene expression data from two colour cDNA microarray experiments. This package (a package referring to a collection of functions) has been made available for use in Matlab (MathWork Corp) or in the open-source software package R (Ihaka and Gentleman 1996). As R now contains a wealth of tools for the analysis of microarray data, this platform is particularly suitable. The R package YASMA (an extension of the SMA [statistics for microarrays] package) also offers functions for ANOVA-based analysis although only with fixed effect models. An extensive introduction to ANOVA based analysis and explanation of fixed and mixed model tests can be found in (Cui and Churchill 2003) and (Wu *et al.* 2002).

Controlling the Error Rate

When carrying out any statistical test, there exists a risk of making either a type I or type II error. In the case of microarray analysis, a type I error occurs when a gene is

selected as being differentially expressed when, in fact, it is not. A type II error is the opposite: a gene that is differentially expressed fails to be identified.

Microarrays typically contain many thousands of spots and this requires the performing of an equal number of statistical tests in the case of differential gene identification. In such a situation, the accumulation of type I errors may result in a large number of false positives. A number of approaches are available to control the associated error rate.

Historically and classically, the most common approach is to control the family-wise error rate (Sokal and Rolf 1995). This is the probability of making one or more type I errors. The most commonly applied family-wise error rate correction is the Bonfferoni correction (Zar 1999). This correction requires significance at the nominal level divided by the number of tests. Cui and Churchill (2003) give details of other correction methods, including permutation analysis. The criteria for meeting the family-wise error rate are very stringent and result in a large reduction in statistical power. As the aim of differential expression analysis is to produce a list of genes for later analysis by the investigator, a certain number of false positives may be considered acceptable.

With this in mind, Reiner *et al.* (2001) proposed controlling the false discovery rate (FDR), which is the proportion of false positives among all of the rejected null hypotheses. Storey (2002) has further adapted this idea and proposes controlling the positive false discovery rate. The false discovery rate is computed after a list of differentially expressed genes has been produced and will estimate the proportion of those genes that will represent false positives. It is less stringent than the family wise error rate corrections and therefore retains more of the power of the original statistical test. The software package SAM and it's newer replacement PAM (Storey and Tibshirani 2000) are freely available packages for either R or Microsoft Excel with false discovery rate calculation and visualistion (in the case of the R version) abilities.

1.4.10 Cluster Analysis

Beyond differential expression analysis, another powerful use of microarrays is the organisation of genes into biologically relevant groups. As mentioned previously, this is done for three main reasons:

Functionally related genes can be expected to be co-expressed. Identifying co-expressed genes through guilt-by-association, can allow the classification of previously un-classified genes, co-expressed genes can identify regulatory control mechanisms, and gene expression varies between cell type and cellular state. Cells with similar patterns of regulation and expression can be identified as a cell type or as a cell under a known stress. In medical fields this is particularly important as a patient with an unknown condition can be screened and the condition identified by the pattern of gene expression.

Clustering is the process of separating a set of objects into subsets based on their similarity (Gilbert *et al.* 2000). The aim is usually to identify clusters that contain objects that are very similar to each other (low intra-cluster variability) while being distant from members of other clusters (high inter-cluster variability). There are a number of statistical tools available to perform cluster analysis. Commonly used clustering methods include hierarchical clustering, self Organising Maps, *k*-means clustering, and principal components analysis (also referred to as single value decomposition). Sturn and Tibrishani (2000) offers an excellent introduction to these methods as well as comparison between them.

1.4.11 Experimental Design

As with any biological experiment, experimental design is the most important step of the experimental process: A poorly designed experiment will limit the conclusion that can be drawn from the data or, even more seriously, result in incorrect conclusions being drawn and stated from the data. Experimental design can be flawed in many different ways. Microarrays previously suffered from their associations with biochemical and molecular backgrounds. These disciplines are unaccustomed to uncertainties inherent in many biological experiments and the subsequent requirement for statistical analysis; researches in biochemistry and molecular biology are more familiar with assays providing with yes / no answers of a nature that microarrays can not supply. Historically in biology, the people who know about statistics are the ecologists, who are only just beginning to wake up to the potential of microarrays and to comment on the design of microarrays experiments.

One of the most classical errors to make when planning an experiment is that of pseudoreplication. This is a topic covered in detail in Morrison and Morris (2000) and Hurlbert (1984) from an ecology perspective and touched upon in Churchill (2002). All of the issues discussed in Morrison and Morris (2000) and Hurlbert (1984) are equally applicable to the design of array experiments, or indeed any biological experiment involving inference.

Conducting microarray experiments requires some special considerations. For example, to compare two treatments, should one hybridise both against a control and compare the results from different slides to each other or should the two treatment samples be hybridised against each other directly on the same array? A brief introduction to these issues can be found in Churchill (2002), Quackenbush (2001) and Kerr *et al.* (2000).

Technical replication v's Biological replication

As mentioned above, this is an issue often mis-handled by researchers or one that is given no consideration at all. These two fundamental units of replication serve different functions and must be handled differently when analysing data. Technical replication at appropriate levels are essential in microarray analysis to allow assessment of the reliability of the technique and to identify biases. If technical replicates are not considered, the number of confounding influences becomes too great for biological conclusions to be drawn reliably. However, technical replication provides no data from which to draw biological conclusions and using them as biological replicates is fundamentally and absolutely inappropriate. Biological replicates are essential to show that the observed result is not peculiar to just one individual but that it is a reproducible result that is representative of the system under consideration. The ideal design will maximise biological replication while providing adequate levels of technical replication to provide a sound basis for placing faith in the experimental method being used.

Assuring adequate replication in microarray experiments can be problematic due to the high costs involved. The number of replicates required depends on a number of factors. If the investigation is a 'look-see' screen then fewer replicates may be acceptable as further work will confirm or deny results and conclusions drawn. However, if the aim is to provide evidence for a statement such as 'Cyclin D is up-regulated as a result of treatment *x*' based purely on array data, then far greater replication will be required.

It is due to these reasons that journals have placed increasing restrictions on the publication of microarray data. Many only allow publication of microarray data alongside confirmatory studies, such as northern blots or RT-PCR (reverse transcription PCR).

Conclusions

The above section contains the key references and introductory references for each topic under discussion. A far more comprehensive set of references can be downloaded and searched as an Endnote (or other format) database file from http://popyomics.biol.soton.ac.uk/~nat/ref.html.

It is clear that microarrays are an invaluable addition to the field of biological research. They offer the potential to answer questions at a number of different levels through the different ways in which microarray data can be interrogated. For example, questions can be asked about signalling pathways through the use of clustering methods, genes involved in the response to stress can be used by identifying differential expression between control and treated plants, and questions about genetic architecture can be approached through the hybridising of related species together on an array and examining both differentially expressed genes and emergent gene networks (this will relate to *cis*- and *trans*-regulatory differences respectively).

The development of microarray technology has stimulated a new field of statistical research in biology and one that will benefit many related areas of biological research. It also serves to more formally cement an important link between statisticians, computer scientists, and biologists, forcing each to learn the skills and knowledge of the other to forge new understanding of complex data.

1.5 QTL mapping

1.5.1 General introduction

The vast majority of observable traits in higher organisms can readily be seen to be quantitative. A stand of trees will consist of trees of varying heights, weight of a group of undergraduate students will vary across a range and the cold-tolerance of bees within a hive will vary. If data are collected on each individual from these groups and those data plotted as a histogram, it will be seen that the data form a normal distribution with the majority of individuals having a phenotype near the centre of the distribution. There will also be a spread of values around the mean (which in the case of a normal distribution is the centre of the distribution) with the shape of the distribution histogram resembling the classical-normal bell-shaped curve. The existence of such quantitatively varying traits can not be explained and modelled in the context of classical-Mendelian genetics, where by the population can be divided into a subset of qualitative groups with the frequency of those groups describing the Mendelian inheritance controlling the trait (e.g. the famous smooth or wrinkled pea-pod example). This posed a challenge to breeders and researchers, academically and practically respectively, as most traits of agronomic and commercial interest are quantitative. Quantitative traits are additionally influenced by environmental factors (for example beer consumption will affect weight of undergraduates, nutrient availability will affect tree height), unlike their Mendelian counter-parts that are always either smooth or wrinkled (for example).

There is now a long history behind the study of genetic factors underlying qualitative traits. It was Sax (1923) who first detailed the concept of association between a continuously variable phenotypic trait and a discrete trait (marker). This was later added to by Thoday (1961), who developed the first statistical approach mapping quantitative traits.

Early studies were limited in their ability to accurately identify the genomic loci controlling the quantitative trait of interest. The first studies relied on the use of observable discrete phenotypic traits (such as eye colour, seed coat pattern) that were identified as being associated with variation in the continuous trait of interest. These studies did not locate the discrete trait within the genome; they merely showed association between the discrete trait and the quantitative one. In order to form a link to the genome position, a number of markers were required that could be scored in the progeny of a cross between two parents with contrasting phenotypes. Due to the presence of linkage disequilibrium, markers will display varying degrees of linkage to each other (this being dependent on the genetic distance between them), and to the loci contributing to the continuous trait of interest. The markers can be positioned in relation to each other, and the genetic distances (measured in centi-Morgans, cM) between them be determined, by construction of a linkage disequilibrium map (Lynch and Walsh, 1998). A process of quantitative trait loci (QTL) mapping can then be carried out where-by a marker that shows association to a quantitative trait is identified and its position on the linkage map noted.

During the past half-century there has been an explosion in the availability of markers for use in QTL mapping. Gel electrophoresis provided the first increase in marker density, allowing the identification of isozymes that could be scored in the offspring of a mapping population. This was followed by the introduction of genomic based markers, which saw an explosion in the number of markers that could be identified. This increase in the number of available markers has led to ever-improved QTL detection resolution. These markers are detailed below.

Early studies on annual crops plant species suggested that quantitative traits were under the control of a small number of genes with large effect (major genes) that controlled a substantial proportion of the phenotypic trait variation (Paterson *et al.* 1988, Stuber *et al* 1992). This is often referred to as the oligogenic model quantitative variation. Otto and Jones (2000) describe this distribution of gene effects as a negative exponential distribution. A sparsity of markers meant that the numerous genes with small effect could not be studied individually in these studies; their influence was

inferred collectively via resemblance between relatives (Lynch and Walsh 1998). Strauss *et al.* (1992) suggested that forest-tree species would differ to their annual relatives due to the large population size and long generation time of trees, which would require a greater ability to adapt to changing environments over long time periods. They suggested that forest-tree species would follow a polygenic model where a large number of genes, each with a small effect, would control quantitative traits and that this would severely limit the success of QTL mapping. However, work by Groover *et al.* (1994), Bradshaw and Stettler (1995), and Grattapaglia *et al.* (1995) showed that QTL were readily identifiable in a range of tree species with the proportion of explained phenotypic variation due to the QTL varying from low (~1%) to very high (~98%).

1.5.2 DNA markers

Saturated marker densities began to be available first with the introduction of restriction fragment length polymorphisms (RFLPs, Botstein *et al.* 1980). Since then, a number of additional DNA-based marker technologies have evolved. The most recent make use of the polymerase chain reaction (PCR) in addition to high through-put sequencing and, more recently, EST data and microarrays.

The various marker technologies differ in a number of factors that determine their information content. These parameters include their mode of inheritance (dominant or co-dominant) and the polymorphic information content (Botstein *et al.* 1980). Some marker technologies, such as amplified fragment length polymorphisms (AFLPs) are also covered by international patents, meaning that the cost of royalty payments may become a consideration in the choice of marker technology.

Randomly amplified polymorphic DNA (RAPD) markers

RAPDs are a PCR based marker technology where short random primers are used to amplify anonymous genomic segments. They therefore represent an approach that is simple to use and set up. However, the results obtained are dependant on the PCR

reaction setup and conditions, making results difficult to reproduce. RAPDs have a dominant mode of inheritance, which limits their information content although Plomion *et al.* (1996) detail how this limitation can be largely overcome by saturating the genetic map with dominant markers in *trans*. This approach marks both homologs of the allele and therefore the markers become similar to a co-dominant marker. RAPDs are also limited due to the low frequency of identified polymorphism.

Amplified fragment length polymorphism (AFLP) markers

The AFLP marker system is a propriety technology (licence held by Keygene, Wageningen, Netherlands) developed by Vos *et al.* (1995). Generating AFLPs involves initial digestion of DNA with two restriction enzymes followed by ligation of double stranded adaptors to the ends of the restricted fragments. The fragments are then amplified by PCR using primers that anneal to the adaptors. However, a number of selective nucleotides (1-4) are added to the primers. These extend beyond the adaptors into the genomic restrictions fragments. This results in amplification of a subset of the total pool of restricted fragments. Different combinations of selective nucleotides and restriction enzymes can be used. Where there is polymorphism in the restriction site between the two parents of a mapping population, differences will be observed in the amplified products of the PCR reaction. The PCR reaction is run on a sequencing gel and the amplified restriction fragments visualised by radio isotope staining or through the inclusion of a fluorescent dye attached to a nucleotide of the PCR primers. Each restriction enzyme – primer combination will result in the amplification and visualisation of a number of bands and polymorphic bands can then be identified and scored in the progeny of the mapping population. AFLP have a dominant mode of inheritance and therefore suffer from the same information limitations as RAPDs. However, the multiplex nature of the method makes it highly favourable as numerous markers are identified.

Microsatellite or simple sequence repeat (SSR) markers

AFLPs and RAPDs are both limited by the dominant mode of their inheritance. Microsatellites overcome this limitation as they have a co-dominant mode of

inheritance. This makes them particularly suitable for use in out-breeding tree mapping pedigrees such as the one used in this study (see details below). As the level of heterozygosity is high in out-bred trees, SSRs are highly likely to provide fully-informative markers i.e. a marker where all four allelic configurations are distinguishable in the segregating progeny of the mapping pedigree. SSRs are frequently transferable across species, particularly closely related ones. This makes them invaluable in aligning linkage maps derived from differing crosses of closely related species. An example of this is the use of SSR markers to co-align the genetic maps of four mapping pedigrees forming the basis of the POPYOMICS project (www.soton.ac.uk/~popyomic), to which the outcome of this study is contributing. The mapping pedigrees used in the POPYOMICS study will be discussed in more detail below.

SSRs are DNA sequences composed of varying numbers of short tandem repeats (Hamada *et al.* 1992). PCR amplification using primers that anneal to conserved flanking regions is used to detect polymorphisms in repeat number between the parental genotypes and within the progeny of the mapping pedigree being examined. SSR markers have been developed for a number of tree species including *Pinus radiata* D. Don. (Devey *et al.* 1996), *Pinus strobus* L. (Echt *et al.* 1996), *Picea sitchensis* (van der Ven and McNichol 1996), *Pseudotsuga menziesii* (Slavov *et al.* 2004), and eucalyptus (Kirst *et al.* 2004). Attention has largely been focused on species of commercial importance due to the level of investment required and high costs involved in developing SSR markers. Although the information content of each SSR marker is high, the method is resource and labour intensive. Only one SSR can be identified from each primer set and SSR detection is far more efficient from where SSR enriched genomic libraries have been constructed (Karagyozov *et al.* 1993)

In recent years a large and ever growing number of SSRs have been identified in various poplar species including *Populus nigra* (Smulders *et al.* 2001, van der Schoot *et al.* 2000), *Populus tremuloides* (Dayanandan *et al.* 1998), and *Populus trichocarpa* (Tuskan *et al.* 2004, Frewen *et al.* 2000). Tuskan *et al* developed their SSRs from the recently released *Populus trichocarpa* genome sequence and these SSRs therefore

represent a spread of intergenic, exonic and intronic regions making them an invaluable resource as they provide coverage of genomic regions not covered by other marker types. The frequency of location within the different genomic regions (intergenic, exonic, and intronic) varied, as one would expect and as has been detailed by various studies (Morgante *et al.* 2002, Temnykh *et al.* 2001, Toth *et al.* 2000).

Single-nucleotide polymorphisms (SNPs) as genetic markers

SNPs (Kruglyak 1997, Kwok *et al.* 1996) represent single base-pair alterations, insertions, or deletions in the coding sequence. Their frequency of occurance varies throughout the genome, with higher frequencies occurring within intergenic regions of DNA. SNPs represent a highly abundant pool of genetic variation and the application of this variation as genetic markers is only recently becoming popular. The use of SNPs has been increasing in recent years due to the ever increasing amounts of sequence information becoming available, improvements in SNP detection and the development of high-throughput genotyping methods (Gut 2001, Shi 2001, Syvanen 2001). Interest in the use of SNPs for association studies for the genetic dissection of complex traits in humans and other out-bred species has also bolstered the popularity of SNPs as genetic markers (Risch and Merikangas 1996).

SNPs can be readily identified where ESTs have been derived from a mixture of genotypes or closely related species and where the EST sequence information is available in a database. A range of bioinformatics tools are now available for the *in silico* identification of SNPs from sequence databases (Barker *et al.* 2003, Schmid *et al.* 2003). An extensive EST resource, comprising >100,000 EST sequences (these representing approximately 15,-20,000 contig sequences, pers. comm. Stefan Janson, UPSC, Umeå, Sweden) collected from a range of tissue libraries has been made publicly available (www.poppel.fysbot.umn.edu) in addition to other rapidly developing EST collections that are being collated as part of a European-funded project (www.picme.org), which contains all ESTs represented in the French INRA EST collection (http://mycor.nancy.inra.fr/PoplarDB/), and that of AspenDB (http://aspendb.mtu.edu/), an EST collection made for the investigation of regulation

of phenolic glycoside metabolism in *Populus*. Other extensive *Populus* EST collections are also expected to be publicly available shortly (pers. comm. Rolf Sera, Genome Canada, University British Columbia, Canada, http://www.treenomix.com/). SNPs can also be identified from the *Populus* genome where heterozygous alleles have been identified.

Identified SNPs must be confirmed through PCR amplification and sequencing. Forest tree species pose a particular problem here as most individuals are highly heterozygous. This necessitates the ability to discriminate between the two haplotypes for any heterozygous nucleotides and sequencing errors. Bioinformatics tools have been developed to aid in this discrimination (Marth *et al.* 1999, Nickerson *et al.* 1997) and high quality sequencing also aids the discrimination between the two possibilities.

Other marker systems

I have limited the above discussion to the most widely used marker systems in forest tree research. A number of other markers systems have also been developed. As mentioned briefly, RFLP were once common markers but have largely been replaced by AFLPs due to the multiplex nature of the AFLP system. Another marker system that is similar to AFLPs is cleaved amplified polymorphic sequences (CAPS). CAPS differ to AFLP in that the DNA fragments are first amplified and subsequently cleaved with restriction enzymes (Iwata *et al.* 2001). Sequence-tagged-sites (STS) markers are polymorphisms detected by locus-specific amplification using primers that produce present/absent marker phenotypes (Perry and Bousquet 1998, Tsumura *et al.* 1997). Sequence characterised amplified region (SCAR) markers are based on sequenced AFLP or RAPD markers (Gosselin *et al.* 2002, Paran and Michelmore 1993). More recently, EST sequences have allowed the analysis of polymorphisms based on SNP markers in transcribed regions (Temesgen *et al.* 2001).

1.5.3 QTL mapping methods

Alongside the increase in marker density and marker technologies has been a progression and expansion in the methods available to map QTL. QTL mapping is a multi-step process. Firstly, two individuals that are polymorphic for the trait of interest must be identified. A set of offspring is then produced by the crossing of these two parents (some different crossing strategies are detailed below). Next, a set of polymorphic markers must be identified in the parents and then these genotype of those markers scored within the mapping population. Marker-genotype data is then used to construct a linkage (or genetic) map. A phenotypic score of the trait of interest is then recorded from each of the genotypes in the mapping population (preferably with replication at the genotype level). QTL mapping is then the process in linking together these sources of data with the aim of identifying markers that are linked to the control of the phenotypic trait that has been scored (or measured). The outcome of QTL mapping is then identification of a region(s) of the genome that is involved in the genetic control of the trait under investigation. This information can then be used to aid breeding designs in what has been termed the process of marker assisted selection (MAS). MAS has not been used extensively to date to provide improved genotypes or increased speed of breeding, but some projects are nearing completion (Pers. Comm. Steve Hanley, Rothemstead Research, UK). This is possibly due to the problem that most quantitative traits are under the control of few to many QTL and that these QTL are dependant on the genetic background in which they exist. Many QTL involve epistatic interactions and these must be considered if MAS is to be successful.

QTL mapping, in principle, examines the association of markers with a phenotypic trait. QTL analysis is necessary since the effect of individual QTL is too small to be tracked and identified by segregation analysis in pedigrees. QTL mapping therefore makes use if markers that are in linkage disequilibrium and that do display Mendelian inheritance. If a marker is linked to the trait of interest then a differences in mean values of the phenotypic trait will be observed among individuals with different genotypes at that marker locus. The success of QTL mapping is dependent on the

marker density of the genetic map (the higher the density, the greater the number of loci that will be represented) and the size of the mapping population being used (the greater the size, the greater the number of meioses that will be represented). The combination of these two factors will determine the resolution of the QTL mapping (i.e. the cM distance to which the QTL can be isolated to lie within).

Regression analysis

Regression analysis is the most basic form of QTL mapping available. It is computationally simple and can be achieved in most statistical software available. Regression analysis makes use of a general linear model (glm) to regress the phenotypic value onto the marker genotype. A marker that is linked the a locus controlling the phenotypic trait will generate a stronger regression than one that is less weakly linked or not linked at all. The strength of the regression between markers and the QTL will decrease as the genetic distance between them increases, and association will break down completely when markers lie on another chromosome (in the simple case of a single gene / QTL model in any case). In this way, the most likely position of the QTL can be indicated. Kearsey and Hyne (1994) have extended this approach to deal with multiple-QTL models (but only multiple QTL lying on the same chromosome).

Interval mapping

Interval mapping was popularised by Lander and Botstein (1989) primarily through the provision of a set of software tools to carry out their method of maximum likelihood (ML) interval mapping. Interval mapping is similar to basic regression analysis except that the regression is carried out across a marker interval. ML interval mapping calculates the likelihood of the odds (LOD) score, which is the $\log_{10}$ of the likelihood of the odds ratio. LOD is the probability of a QTL being located within a marker interval measured against the probability of the result occurring by chance. In this way interval mapping localises a QTL between two genetic markers (referred to as flanking markers) rather than locating it to a single locus, as does simple regression analysis. ML interval mapping requires the use of custom written software and is

computationally intensive (particularly when extended to multi-QTL models). An alternative is the use of least-squares (LS) multiple regression as proposed by Haley and Knott (1992). They discuss the relative merits of the two approaches. LS regression is computation far more simple and consistently yields similar results to ML methods.

Composite interval mapping

Composite interval mapping (CIM) is an extension of ML interval mapping whereby non-linked markers are included as cofactors in a multiple regression analysis (Jansen and Stam 1994, Zeng 1994). This reduces the influence of multiple or linked QTL and reduces the bias in estimates of QTL detection that can be introduced by an inappropriate single-QTL model (Schork *et al.* 1993). CIM has been extended to multiple traits as well as QTL by trait and environment interactions (Jiang and Zeng 1995). The outcome of CIM will be dependent on the markers used as cofactors in the regression model and this can introduce selection bias. However, it provides a more powerful means of QTL detection than standard interval mapping and many software packages are now available that make use of CIM.

Multiple interval mapping

The multiple interval mapping (MIM) extension of interval mapping again includes additional markers as cofactors in a multiple regression analysis (Kao *et al.* 1999). In the case of MIM, multiple models are tested under a stable model emerges that identifies the main and interaction effects for multiple QTL models. Again, this method will be influenced by the selection of cofactor markers and the possible influence of cofactor selection must be considered by the investigator. This is a problem familiar to users of stepwise regression where the order of factors entered into a multiple regression model must be determined by the investigator (Miller 1990)

Declaring a QTL as present: the multiple testing problem

All of the above methods calculate a critical cut-off score above which a QTL is determined to be significant. This score is calculated based on the model parameters and consideration of the distribution character of the data. All of the above methods calculate multiple tests and this increases the chance of making a type I error. This chance increases with marker density. Where more than 20 markers are being tested, this chance becomes as QTL may often be declared as significant where actually none is present. The chance of a false-positive is highest around regions with a major QTL and this introduces the presence of what has been termed 'ghost qtl'. These problems can be decreased by accounting for the level of multiple testing and adjusting the required significance threshold accordingly. This is most often achieved by bootstrapping or permutation analysis of the data set whereby data are re-shuffled a number of times (bootstrapping) or continuously reshuffled until convergence (permutation) and the test statistic re-calculated. These re-shuffled datasets are essentially random and the significance threshold will be adjusted until no QTL are detected within the shuffled data (as these will be due to chance). The outcome of this process is that only QTL of major effect are able to be detected so many QTL of smaller effect will be missed. This is not a problem for a trait where heritability is high and a small number of QTL have large effect of the trait of interest.

Advancing methodologies for QTL mapping

Within the past four to five years a new approach, based on Bayesian statistics, has been emerging for the detection and analysis of QTL. Bayesian statistics differs from classical statistical approaches in that it makes no underlying assumptions about the distribution of data. Instead, probability distributions are computed from each dataset and this information is used for the underlying parameters in a model. Bayesian analysis is preferable as it can fully take into account all unknowns involved in QTL mapping; principally number, location, effect, and the genotypes of markers and QTL. Most work to date has concentrated on biallelic models from simple mapping crosses. However, a paper by Yi and Xu (2001) offered an insight into the potential and flexibility of this method and its application to very complex mating designs. A review

of articles on Bayesian mapping of QTL was provided by Hoeschele *et al.* (1997). Since then, work has progressed to consider multiple models with different numbers of QTL (Stephens and Fisch 1998, Sillanpaa and Arjas 1999, 1998, Heath 1997, Satagopan *et al.* 1996). The methods proposed in these papers are all based on the reversible jump methodology of Green (1995) for constructing a sampler that jumps between models of different dimensions and all make use of the Markov chain Monte Carlo algorithm (MCMC). Markov chains are now used extensively for a number of bioinformatics applications including efficient BLAST searches and gene prediction post genome sequencing. The optimal model is selected when convergence is achieved with the model dimensions yielding the number of QTL identified. The main drawback of Bayesian modeling is that it is incredibly computationally intensive and model convergence is, at present, slow. As yet, no full Bayesian-based software has been developed for use but the power and flexibility of the method will ensure its continued development. A number of interim approaches are currently available that make use of a mixture of the Bayesian and classical approach, for example Ball (2001) proposes a such an approach. An example of software making use of a mixed approach is the R/qtl package developed at the Churchill lab, which makes use of Hidden Markov chain models (HMM). This package is available for the R statistical language and can cope with complex pedigrees with four segregating alleles. It can be downloaded or installed internally within the R environment via CRAN (Comprehensive R Archive Network).

1.5.3 Mapping pedigree designs in forest tree species

Due to the high genetic load and inbreeding depression in forest trees, the production of commonly used mapping pedigrees was not possible in trees. For instance, inbred lines can not be constructed due to high genetic load. Forest trees are highly heterozygous and this complicates the mapping of QTL considerably as there can be, and often are, four segregating alleles present at each locus. Researches of forest trees therefore had to apply novel breeding designs in order to facilitate QTL mapping as well as adapting existing QTL mapping methods and software.

Two-way pseudo-testcross

The two-way pseudo-testcross was proposed when it was realized that genetic linkage maps could be constructed for each parent individually and inbred line models then used for QTL mapping where dominant markers were used (Grattapaglia and Sederoff 1994). When this strategy is used, three inherited marker types are identified: (a) markers inherited from the pollen parent and having a segregation ratio of 1:1; (b) markers inherited from the seed parent and having a segregation ratio of 1:1; and (c) markers inherited from both parents and having a segregation ration of 3:1. The two-way pseudo-testcross has been used extensively in conjunction with dominant mode inheritance markers (Cervera *et al.* 2001, Wu *et al.* 2000). Use of markers that segregate in both parents, as well as use of co-dominant markers such as SSRs, allows alignment of the two parental maps and therefore QTL mapped on the two maps. The two-way pseudo-testcross is usually used to construct intra-specific full-sib families or in F_1 inter-specific families. The POP1, POP3, and POP4 populations of the POPYOMICS project are all examples of this type of population.

Double pseudo-testcross

In this design, an F_1 hybrid is back-crossed to a parental species in order to produce an F_2 pedigree. However, unlike in a standard backcross, a substitution parent of the same

species is used to perform the backcross in order to overcome inbreeding depression (Myburg *et al.* 2003). The resulting F_2 allows for high resolution QTL mapping.

Inter-specific F_2

Use of divergent, highly polymorphic parental species where hybridisation possible allows for the construction of F_2 populations, such as the one used in this study. In the example of this study, two *Populus* species from different sections of the genus were hybridised and an F_2 population constructed from a cross of two of the resulting F_1 offspring. This is depicted in Figure 1.3.1. This family forms the POP1 pedigree of the POPYOMICS population and is called Family 331 (Bradshaw and Stettler 1993). F_2 crosses provide excellent QTL mapping due to the large number of segregating alleles represented. In the case of Family 331, all morphological, physiological, and biochemical traits studied to date have displayed segregation within the F_2 population making this pedigree an excellent model for studying environmental adaptation and the genetic architecture of traits at all levels of the organism. The F_2 genetic map is a combined parental map and can contain a mix of semi- and fully- informative markers. The advent of new QTL mapping methods means that all available marker information can be taken into account when using the genetic map for QTL mapping.

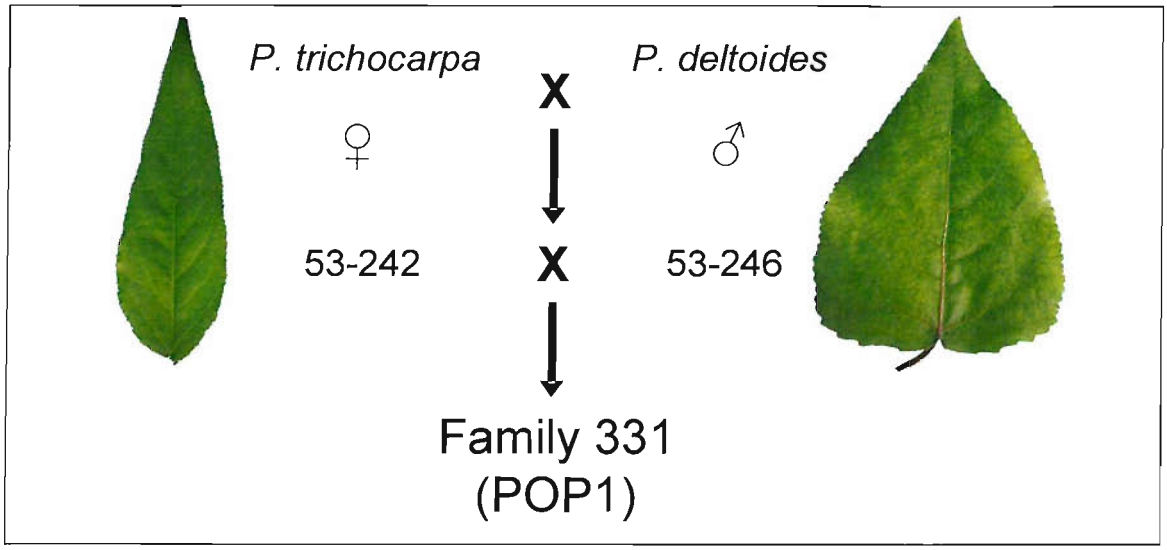


FIGURE 1.3.1 Overview of the production and structure of family 331. the *Populus trichocarpa* (93-968) female was crossed with the *P. deltoides* bart (ill-129) male to produce an F_1 population (Family 53). two F_1 genotypes (53-246 male and 53-242 female) were selected and crossed to produce an F_2 population (family 331, referred to as pop1 populations in the popyomics project, www.soton.ac.uk/~popyomic).

Chapter 2

Characterising patterns of leaf development and drought response in *P. deltoides* and *P. trichocarpa*

2 Characterising patterns of leaf development and drought response in *P. deltoides* and *P. trichocarpa*

2.1 OVERVIEW

Leaves are of fundamental importance to the growth and functioning of plants as they are the site of carbohydrate synthesis and the principal site of environmental sensing. Two species of the model genus *Populus* were selected as having highly divergent leaf size, shape and structure and were grown in order to characterise the mechanisms of leaf area expansion.

P. deltoides was found to have large, deltoid shaped leaves that are constructed principally through the process of cell division, as shown previously by Van Volkenburgh (1999). In contrast *P. trichocarpa* leaves were found to be smaller, lanceolate shaped with area expansion being achieved principally through the process of cell expansion, resulting in a considerably larger mature epidermal cell area than that of *P. deltoides*.

It was shown that the two species respond differentially to the application of a chronic and acute drought stress with *P. deltoides* exhibiting a rapid and significant reduction in leaf area expansion rate and mature leaf area in response to drought. In contrast, *P. trichocarpa* showed no significant leaf area or expansion rate drought response.
These two species therefore serve as an ideal model system with which to study the process of leaf area development. Drought stress results in perturbation of the normal developmental pattern in a species-specific manner.

2.2 INTRODUCTION

Leaf expansion results from a combination of cell division and expansion. In monocotyledonous leaves, these processes can be isolated into clearly defined regions of the leaf (see Fiorani *et al.* 2000). In dicotyledonous leaves, the division and expansion of cells occurs without any defined spatial or temporal boundaries making this a much harder system to model (Taylor *et al.* 2003). Much progress has been made in improving understanding of leaf development, principally through work carried out using the model species *Arabidopsis*. For example, genes controlling the width (Tsukaya 2003), length (Tsuge *et al.* 1996), and cell number (Tsukaya 2002) within expanding leaves have been characterised and mapped, as have many pattern formation genes functioning within the shoot apical meristem (see Fleming 2005 for a recent review of the process of leaf development).

The processes of cell division and cell expansion contributing to leaf area expansion have been characterised in leaves of various *Populus* species in the lab of Gail Taylor. Workers in the lab have examined both the spatial and temporal patterns of leaf area expansion and have examined epidermal imprints from these leaves, measuring epidermal cell area both spatially and temporally (Taylor *et al.* 2003). In *P. euramericana*, it was found that spatial patterns of cell production (a surrogate for cell division) can be defined within the leaf and that the rate of cell production varies over time. Work carried out on the grandparents of a *Populus* mapping pedigree (see Bradshaw *et al.* 1994) showed that leaf area is achieved by different means in different species (Ferris *et al.* 2002). Ferris *et al.* (2002) additionally showed that QTL can be mapped for cell area. Cell area is a normally distributed trait within the F_2 progeny of the mapping population (Robinson 2003), showing that this is a quantitative, segregating trait and that the picture is not a simple one requiring just a few genes of major influence to determine leaf area.

Two species of poplar used extensively for scientific studies, and forming the grandparents of the *Populus* mapping progeny discussed above, have highly divergent leaf shape/morphology, size and anatomy. *P. deltoides* leaves are large, deltoid in

shape having long, flattened petioles and are amphistomatic. In contrast, leaves of *P. trichocarpa* have narrow, lanceolate shaped with rounded, short petioles and are monostomatic (Critchfield 1960). These two species of the model *Populus* genus may therefore represent a model system for comparative studies to examine the control of leaf development processes. The two species were grown in order to attain a detailed characterisation of the mechanisms of leaf area expansion. Additionally, two drought regimes of differing severity were imposed in order to examine the degree to which drought stress causes alterations and perturbations in the normal developmental pattern of leaf expansion in the two species.

2.3 Materials and Methods

Growth Conditions

30 uniform cuttings (30cm) of *Populus deltoides* (clone ILL-129) and *Populus trichocarpa* (clone 93-968) were obtained from a field site under coppice management (for details see Rae *et al.* 2004). Cuttings were taken in the winter of 2001 and stored at 4 °C. Cuttings were soaked in cold water at room temperature for 24 hours before planting. Cuttings were planted in plastic cylinders (Hepworth drainage pipe) of diameter 13 cm and length 75 cm. The base of each cylinder was sealed with a fine mesh. 10 KG of John Innes No. 3 (see http://www.jic.bbsrc.ac.uk/corporate/Media_and_Public/compost.html) compost was added to each 'pot' and lightly compacted and watered thoroughly before insertion of the cuttings. Cuttings were inserted so that only one bud was above the soil line. Any buds from below the soil line that flushed were removed. Cuttings were planted on 7th May 2002. Cuttings were then grown in the growth chambers at Southampton University, UK. Initially the top of the pots was sealed with perforated cling film to ensure an humid environment. The cling film was removed each day to allow watering. Cling film was removed before it restricted the new growth.

Growth rooms were set to a 16 hour photoperiod. Photosyntheically Active Radiation (PAR) was 160 μmol m^{-2} s^{-1} ±10. Day temp was 22 °C and night temp at 15 °C. Humidity was approximately 50%. All cuttings were watered to field capacity for a period of 20 days. On 27th May 2002 the drought treatments were initiated. Chronic drought treated plants received 50 ml water every third day. Water was completely withheld from acute treated plants. Control plants were watered to field capacity throughout. All plants were watered with distilled water. Plants were given a balanced NPK treatment to ensure that no nitrogen stress was introduced (Osmocote, Scotts Europe B.V., Netherlands [N 14: P 13: K 13]). Due to space limitations the two species were placed separately in two growth chambers. A weekly rotation between chambers was carried out to negate any chamber effect. Rows and columns were rotated weekly to negate any spatial chamber effect.

Physiological Measurements

Leaf length was calculated using a flexible ruler constructed from 1mm graph paper. For the purpose of Leaf Extension Rate (LER) measurements, leaves were labelled by tying a piece of fine cotton loosely around the petiole. Leaf Plastochron Index (LPI) measurements were calculated as in Erickson and Michelini (1957). Weekly measurements of height were obtained using a standard extendable measuring tape accurate to ±0.5 cm. Measurements of leaf area were taken at weekly intervals. Leaf area was calculated by obtaining a digital image using a Nikon CoolPix 5000 camera (Nikon, UK). Images were then imported into Metamorph (Version5, Universal Imaging Corporation, Downingtown, Philadelphia, USA). Leaf area, leaf length and width were obtained. Leaves n+1, n-1, n-3, n-5, n-7 (as defined by Erickson and Michelini 1957) were removed at the base of the petiole and a $1cm^2$ area of each veinal area on the right hand side of the leaf (looking towards the tip) from the abaxial and adaxial surface was covered in clear nail varnish (Natural Clear, Boots Ltd, UK). The nail varnish was air dried at room temperature for 10 mins and then removed from the leaf by placing a strip of 'sellotape' over the patch and peeling it off of the leaf surface. In this way an impression of the epidermis was/can be created. The strip of sellotape was then placed on a glass microscope slide and a digital image of the epidermal imprint captured. All images were captured at a magnification of x400. Images were then imported into Metamorph and the cell area of 10 cells obtained for each image.

2.4 RESULTS

Contrasting patterns of leaf development and leaf expansion

Leaf development in the two species of *Populus* was examined to form a developmental and physiological framework for subsequent studies. The two species have strikingly different leaf morphologies: *P. deltoides* has deltoid-shaped leaves (Van Volkenburgh 1999) that can be sub-divided into six inter-veinal segments on either side of the mid-vein in contrast to the lanceolate-shaped leaves of *P. trichocarpa*, which sub-divide into nine inter-veinal leaf segments (Figure 2.1b). Figure 2.1a shows the development of leaf area in relation to Leaf Plastochron Index (LPI, see Taylor *et al.* 2003 for discussion).

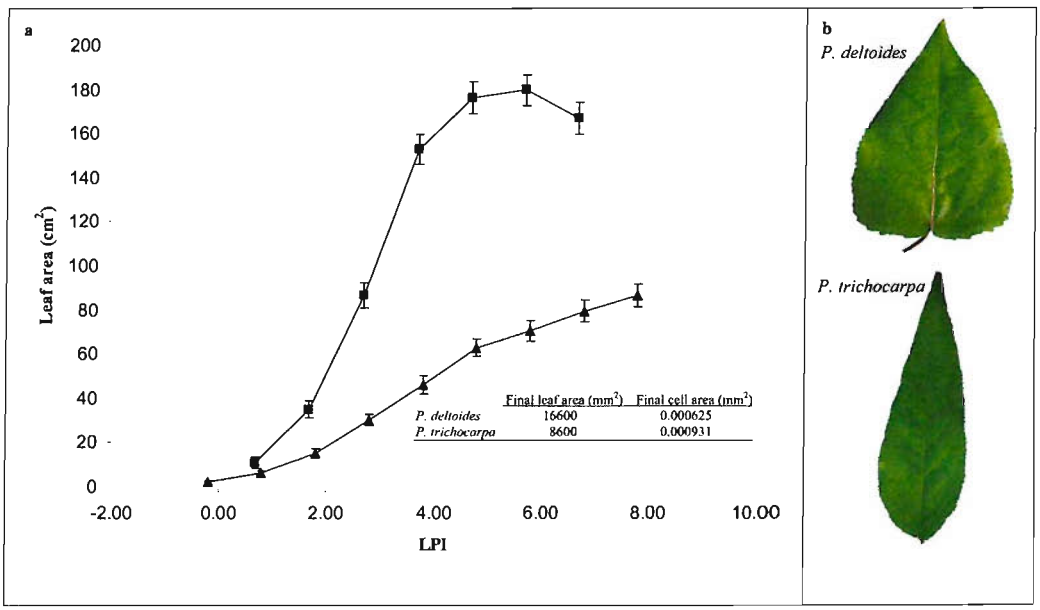


Figure 2.1 Leaf area expansion in *P. deltoides* (■) and *P. trichocarpa* (▲) **(a)** in relation to LPI, **(b)** leaf images of the two species. Each point represents the mean ± s.d. where n=10

P. deltoides attained a mature leaf area of almost double that of *P. trichocarpa* and the temporal pattern of leaf expansion also differed between the two species, with a rapid expansion of leaf area in *P. deltoides* between LPI 2-5 contrasting the steady-rate expansion of leaf area in *P. trichocarpa* (Figure 2.1a). These data are re-expressed as

area expansion rates in Figure 2.2, where it can be seen that *P. deltoides* leaves show a distinct, rapid expansion phase in contrast to the much slower and steady-rate expansion of *P. trichocarpa* leaves.

Clear morphological differences were also seen at the level of the epidermal cell; *P. trichocarpa* cells were significantly larger than those of *P. deltoides* (Table 2.1). This difference in mature epidermal cell area reflects the different mechanisms through which leaves of the two species expand. Table 2.1 gives cell production rates for the two species, showing that *P. deltoides* increased leaf area principally through the process of cell division in contrast to *P. trichocarpa*, which underwent significant expansion of epidermal cell area.

Table 2.1 Cell area, cell production rate, leaf area, and epidermal cell number in control and acute drought conditions. Values are calculated from LPI 5 leaves. Measurements were taken 14 DAP at mid-day. Asterisks indicate outcome of two-way ANOVA test: ns not significant, * $p < 0.05$, ** $p < 0.01$. Letter superscripts indicate the results of a post-hoc Tukey test.

	Cell area (mm^2 x10^6)		Cell production rate (cells day^{-1} x10^4)		Leaf area (mm^2 x10^2)		Epidermal cell number x10^6	
	P. deltoides	*P. trichocarpa*	*P. deltoides*	*P. trichocarpa*	*P. deltoides*	*P. trichocarpa*	*P. deltoides*	*P. trichocarpa*
Control	458^a	716^b	22.09^a	4.93^b	84^a	30^b	18.34^a	4.19^b
Drought	395^a	432^c	13.70^c	6.76^c	54^c	26^b	13.67^c	6.02^d
Treatment	**		**		**		**	
Species	**		***		***		***	
Species × Treatment	*		ns		*		ns	

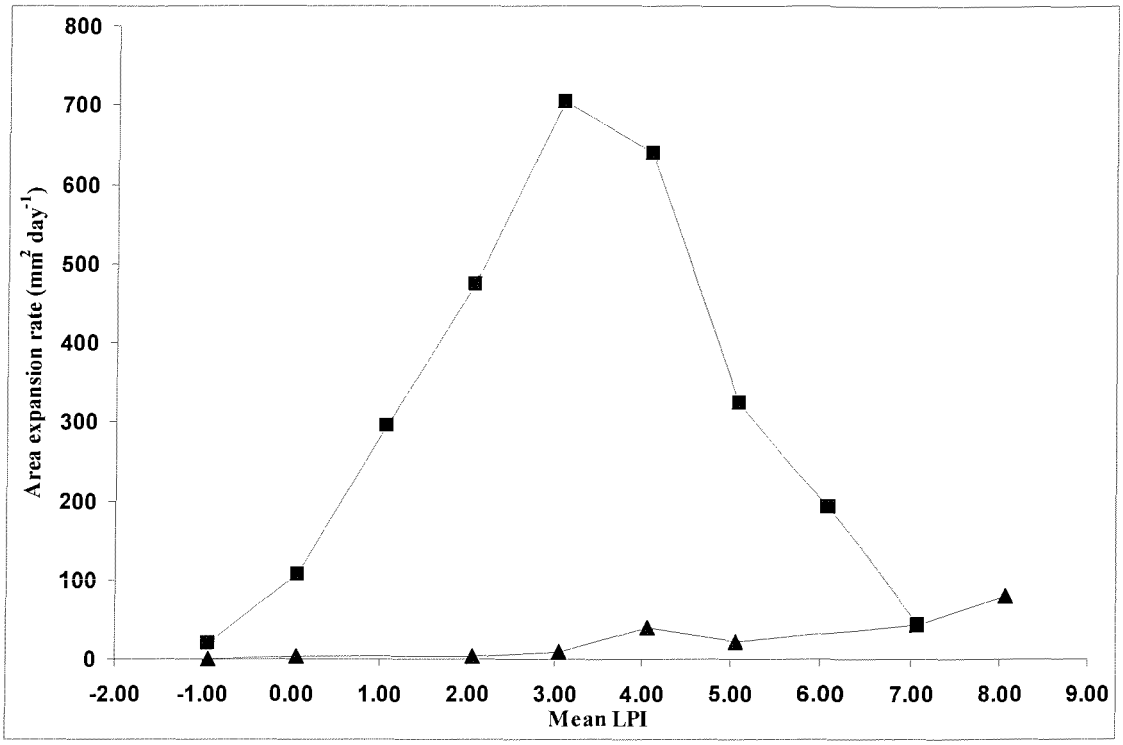


Figure 2.2 Mean leaf area expansion rate for *P. deltoides* (■) and *P. trichocarpa* (▲) plotted against mean LPI, each value represents the mean of 10 individuals.

Leaf expansion was further subdivided into veinal areas within each species, and expansion rates for each inter-veinal area at each LPI were calculated. Figure 2.3 shows the base-tip profile. A rapid expansion of *P. deltoides* leaves in the basal leaf segments between LPI 1-3 can be seen. In contrast, *P. trichocarpa* leaves exhibit a more uniform rate of basal-petal expansion, with sections 1-5 having the highest expansion rates as can be seen in Figure 2.4. Note the different scales in Figure 2.3 and Figure 2.4.

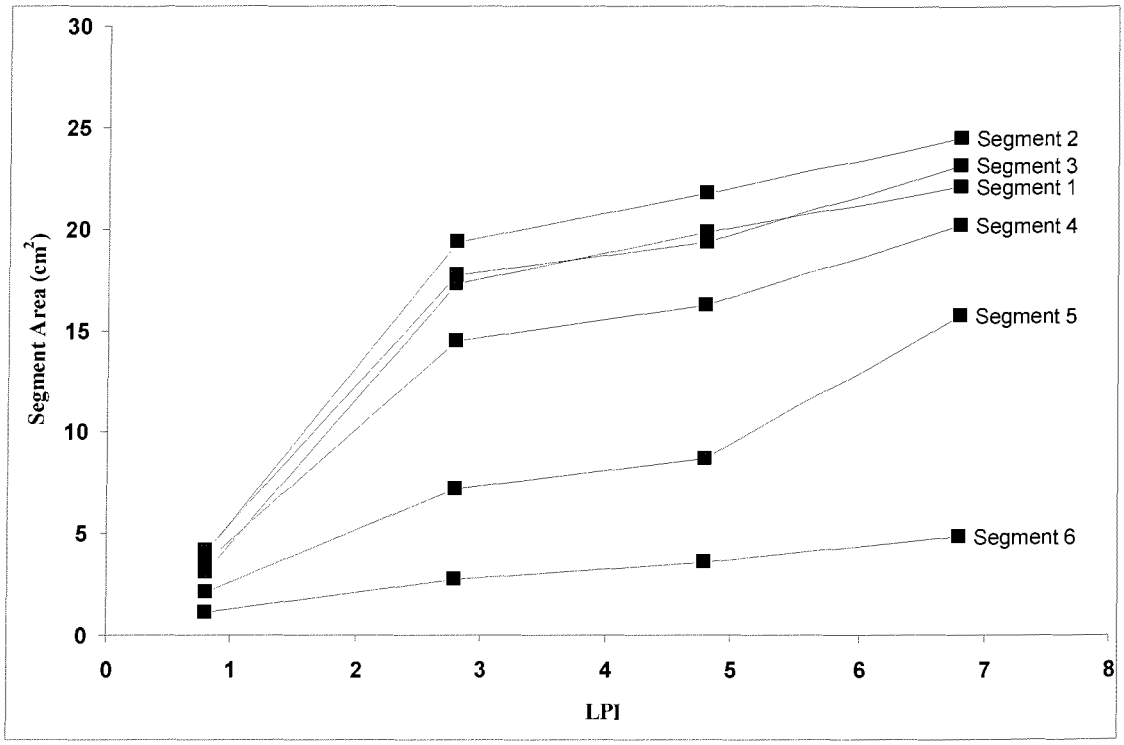


Figure 2.3 Leaf inter-veinal area plotted against LPI for *P. deltoides* leaves. Segments are numbered sequentially from the base to the tip of the leaf. Each value represents the mean of 10 individuals.

Figure 2.5 shows a false-coloured spatial representation of expansion rate for *P. trichocarpa* and *P. deltoides* grown under control conditions. Images for both species are plotted on the same colour scale. The difference in expansion rate ($p < 0.001$) can clearly be seen as can the early flush of expansion in segments 1-3 of *P. deltoides* leaves. It is this early expansion in the lower leaf segments that creates the characteristic shapes of the *P. deltoides* leaf.

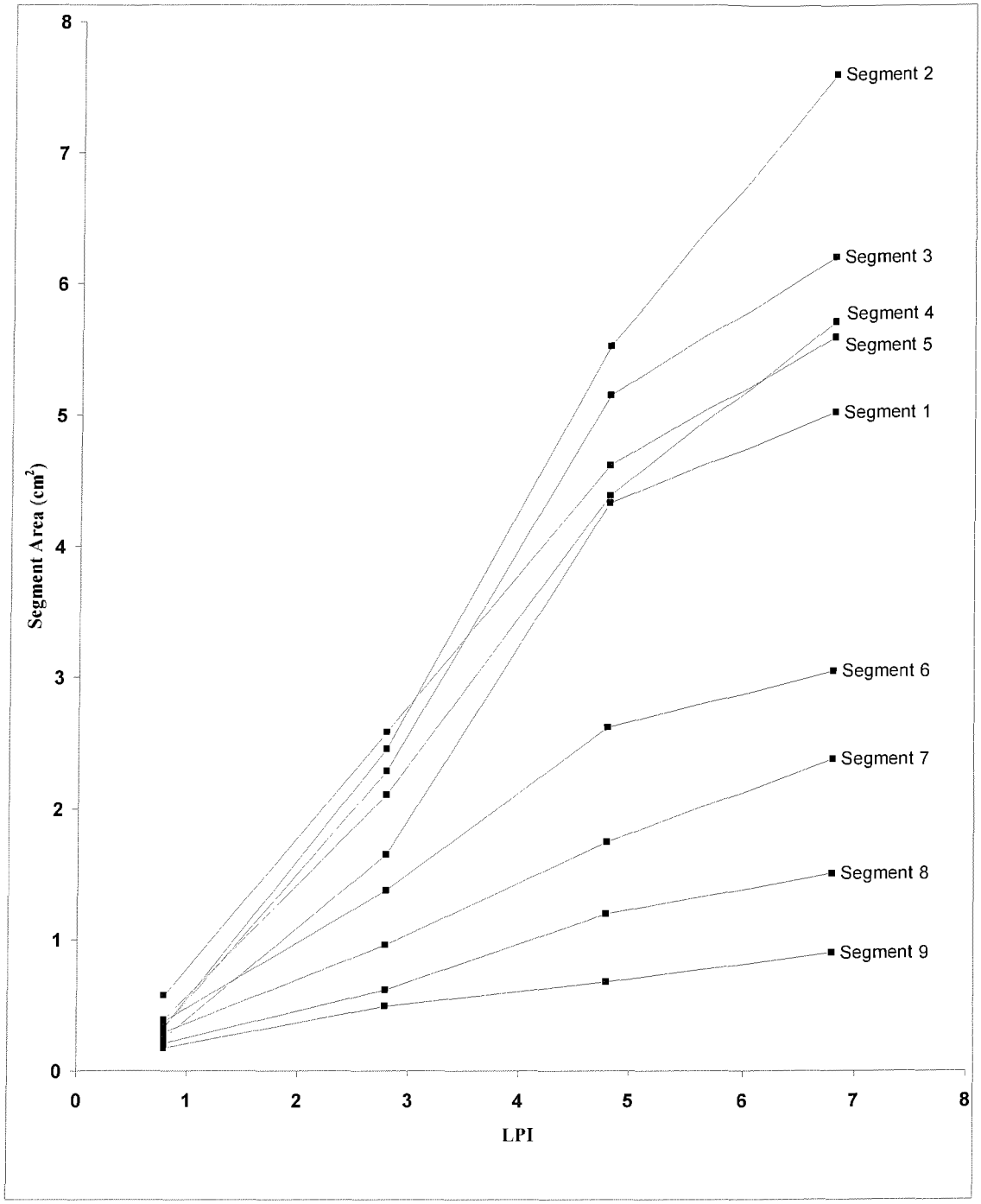


Figure 2.4 Leaf inter-veinal area plotted against LPI for *P. trichocarpa* leaves. Leaf segments are numbered sequentially from the base to the tip of the leaf. Each value represents the mean of 10 individuals.

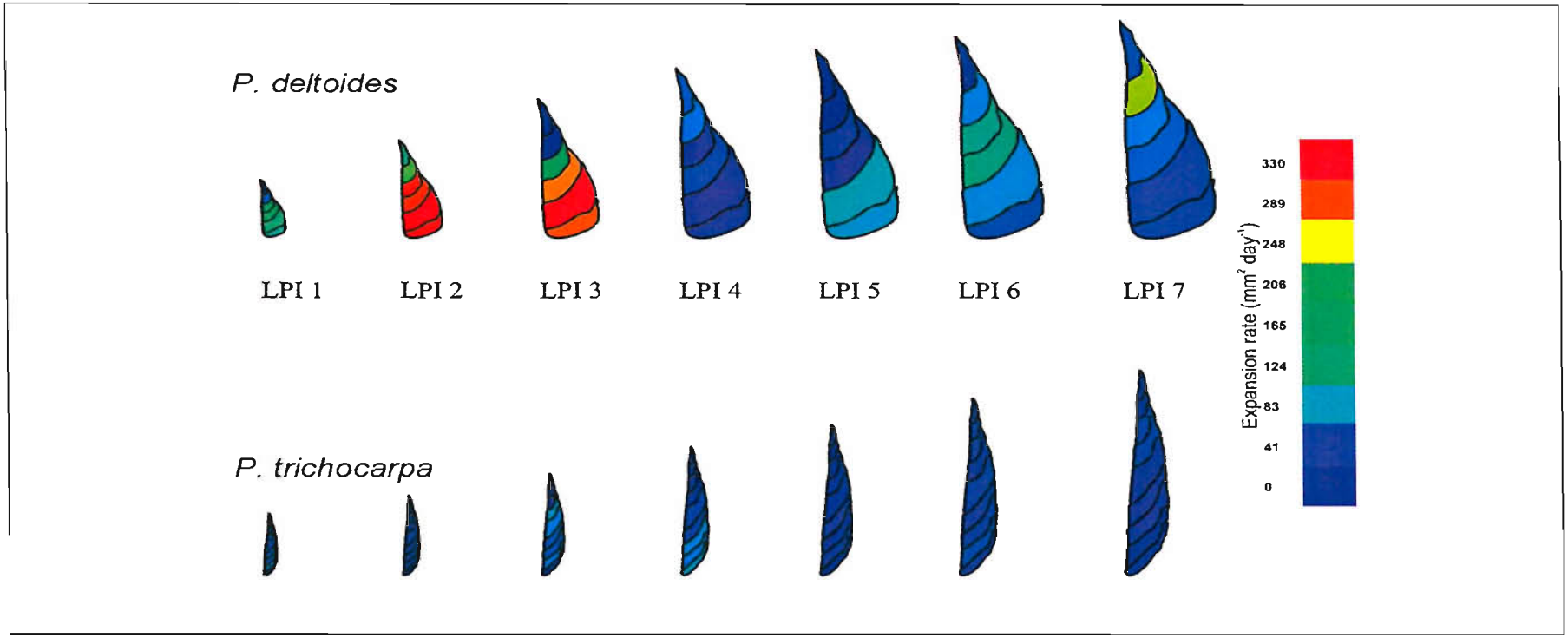


Figure 2.5 False colour spatial representation of leaf segment absolute expansion rate (mm^{-2} day^{-1}) for *P. deltoides* and *P. trichocarpa*. A template leaf was created for each species by drawing around the veinal sections of a representative mature leaf in Photoshop. Each veinal area represents a separate part of the image and each was coloured according to the expansion rate within that vienal section. Colour was calculated on an RGB scale and by fitting a regression line between the minimum and maximum values across both species so that both appear on the same scale. Each LPI was then re-scaled based on the mean leaf area at each LPI measured using the mature leaf area to represent 100%.

Contrasting patterns of cell size and production

Figure 2.6 shows the mean cell area for each leaf segment of *P. deltoides* and *P. trichocarpa* leaves. As is common in dicotyledonous leaves, no clear base-tip pattern can be seen, suggesting that cell division and expansion occur throughout the leaf. Figure 2.6 clearly shows the contrasting mechanism of leaf growth for the two species: *P. trichocarpa* exhibits a significant change in cell size ($p<0.001$) between LPI 3 and 8; *P. deltoides* shows no significant change in cell area between LPI 3 and 8. This suggests that *P. deltoides* expands leaf area principally through the process of cell division, with the epidermis being constructed from a large number of smaller cells. *P. trichocarpa* leaves expand through cell expansion as well as cell division, the result being a leaf with fewer, larger cells per unit area.

The spatial pattern of cell production rate within leaf segments for *P. trichocarpa* and *P. deltoides* is represented in Figure 2.7. The pattern of cell division in *P. deltoides* clearly matched the pattern of leaf expansion seen in Figure 2.5. *P. trichocarpa* shows a much lower rate of cell production (which can be considered a surrogate measure of cell division), which corresponds to the higher rate of cell expansion seen in Figure 2.6.

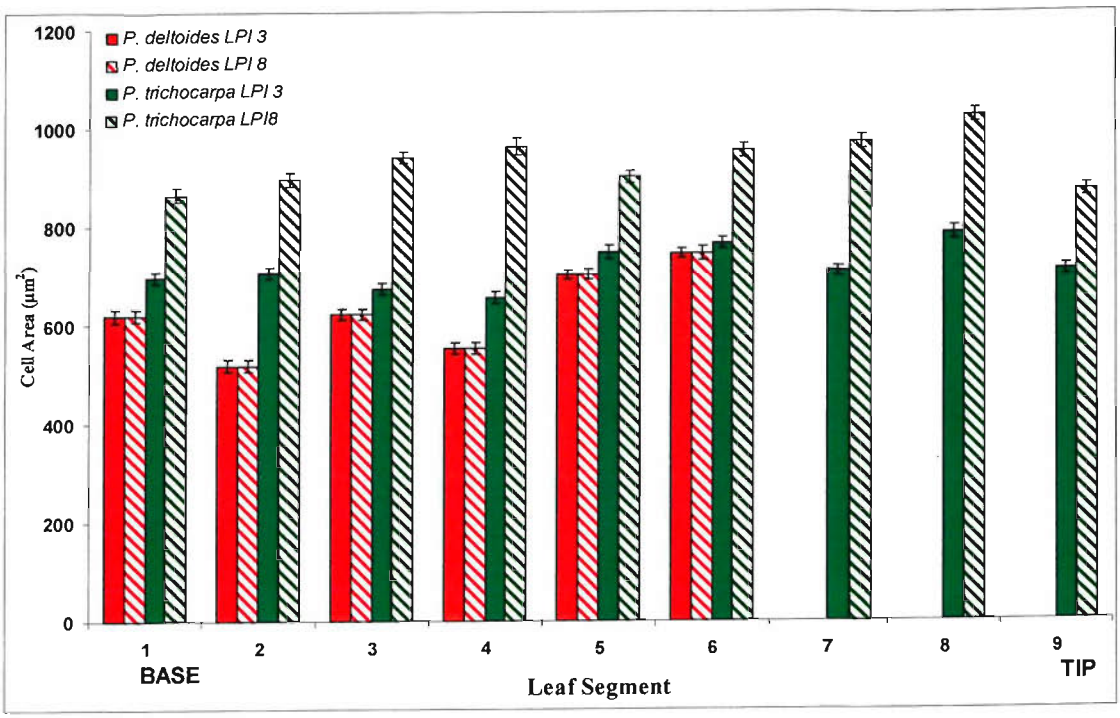


Figure 2.6 Mean inter-veinal cell area for *P. deltoides* (red) and *P. trichocarpa* (green) in LPI 3 (solid bars) and LPI 8 leaves (hatched bars). Leaf segments (inter-veinal areas) are numbered from base to tip. Values are mean ± s.d. where n=10.

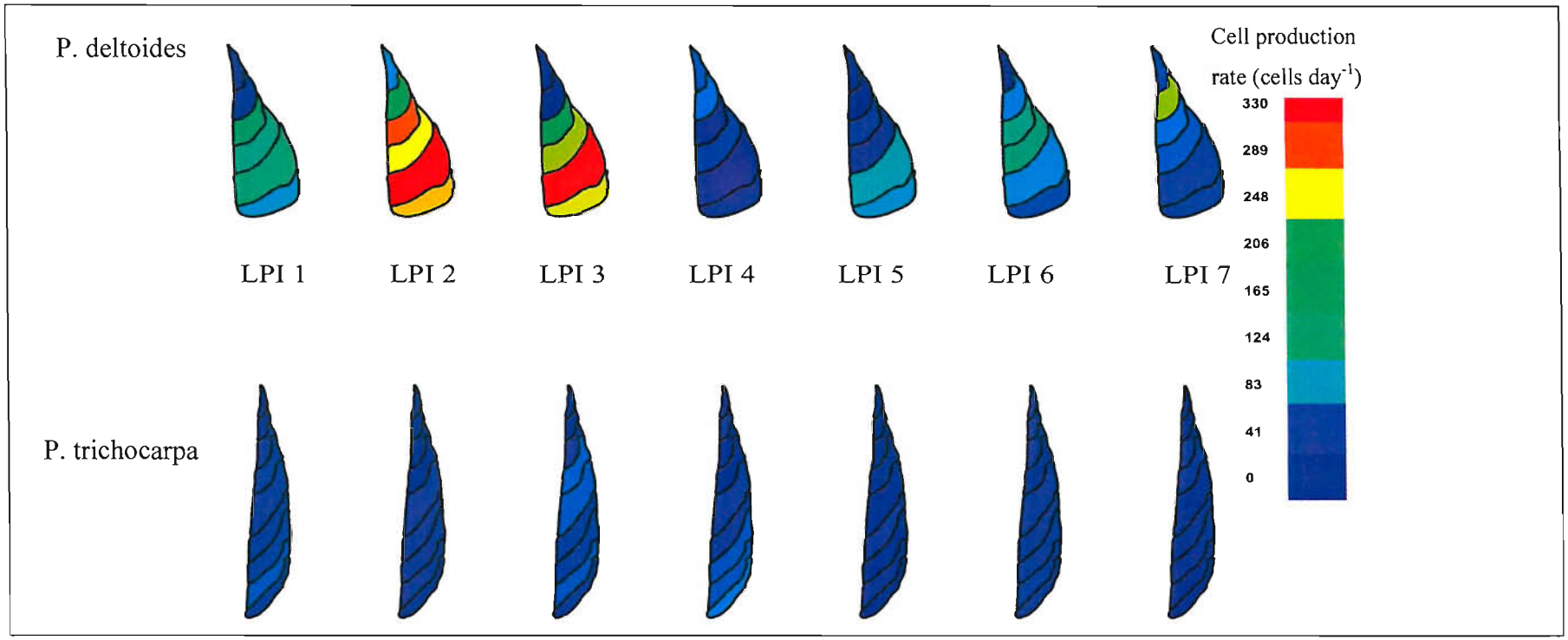


Figure 2.7 False colour spatial representation of leaf segment cell production rate (cells day^{-1}) for *P. deltoides* and *P. trichocarpa*. A template leaf was created for each species by drawing around the veinal sections of a representative mature leaf in Photoshop. Each veinal area represents a separate part of the image and each was coloured according to the expansion rate within that vienal section. Colour was calculated on an RGB scale and by fitting a regression line between the minimum and maximum values across both species so that both appear on the same scale.

Contrasting leaf development and function in response to drought

It has been shown that *P. deltoides* and *P. trichocarpa* exhibit different survival adaptations to drought stress. Figure 2.8 shows LER for the two species under both chronic and acute drought stress. *P. trichocarpa* shows little adaptation to chronic drought stress but does show a significant reduction in LER in response to acute drought stress ($p = 0.001$). However, whether this is an adaptation response or simply a limit on growth due to lack of water to maintain cell turgor can not be established from the present study. *P. deltoides* can be seen to show a gradual adaptation to both chronic and acute drought stress. On day 53, LER was significantly lower ($p=0.01$) in chronic and acute ($p=0.05$) stressed leaves. By day 60 the response can be seen even more clearly with chronic stressed leaves showing a sharp decline in LER ($p <= 0.001$). Acute stressed leaves were reduced in size, possibly due to an inability to maintain turgor pressure.

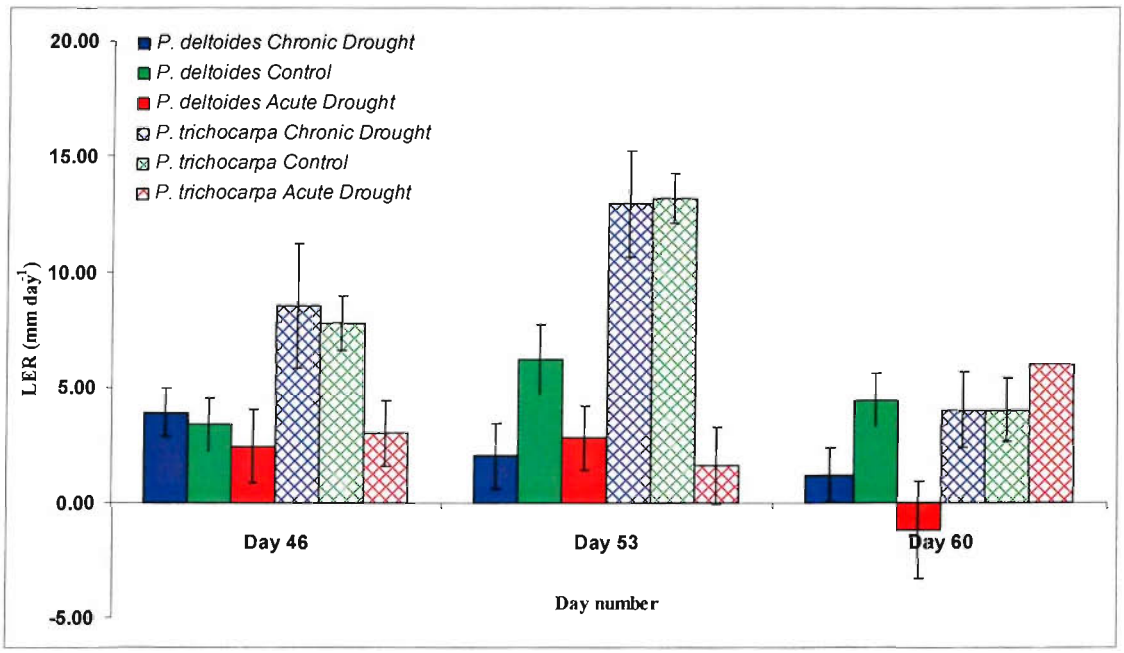


Figure 2.8 Leaf Extension Rate of *P. deltoides* (solid bars) and *P. trichocarpa* (hatched bars) Ln+1 leaves at a series of time points during exposure to chronic (blue) and acute (red) drought stress. Values shown are ± s.d where n=10 for control plants and n=5 for drought plants.

Figure 2.9 shows leaf area for *P. deltoides* and *P. trichocarpa* under chronic and acute drought stress. As can clearly be seen, *P. trichocarpa* exhibits no reduction in leaf area in response to either drought stress. *P. deltoides* shows a contrasting adaptation with significant reductions in leaf area in response to both chronic ($p=0.05$) and acute ($p=0.001$) drought stress.

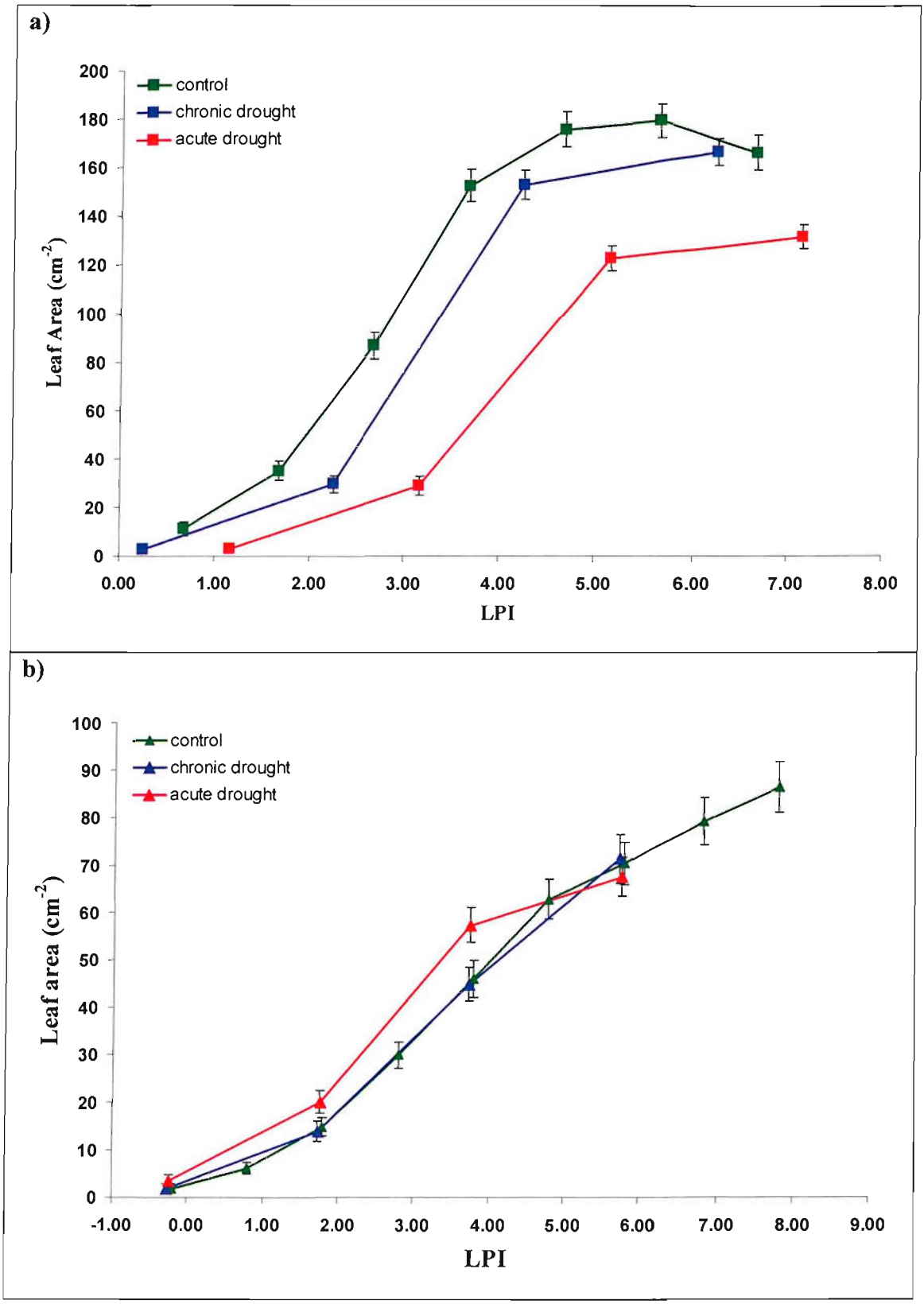


Figure 2.9 Mean leaf area plotted against mean LPI for leaves of **(a)** *P. deltoides,* **(b)** *P.trichocarpa.* in control (green), chronic (blue) and acute (red) drought. Values are shown ± s.d where n=10 for control and n=5 for drought treatments.

2.5 DISCUSSION

The developmental pattern of leaf expansion in *P. deltoides* and *P. trichocarpa* was found to differ dramatically (Figure 2.1). *P. deltoides* produced leaves that were approximately twice as large as those of *P. trichocarpa*. Detailed analysis of the spatial patterns of expansion within defined veinal segments of leaves in the two species showed that expansion rates varied across the leaf surface (Figures 2.3-2.5). For leaves of both species, mature leaf shape could be explained by the spatial pattern of epidermal cell division, as indicated by cell production rate calculations (Figure 2.7). Cell size did not differ across the lamina, which is in agreement with the findings of Taylor *et al.* (2003). However, it can clearly be seen that the rate of cell production differs dramatically between the two species, with cell production rates in *P. trichocarpa* being insufficient to account for leaf area expansion in this species. Figure 2.6 shows that in *P. trichocarpa* the process of cell expansion is also of importance in the process of leaf area expansion. These two contrasting mechanisms for achieving leaf area expansion are reflected in the calculation of cell number within the mature leaf epidermis (Table 2.1). The two species therefore represent an ideal model system for investigating the spatial control of cell division as well as the control of the process of cell expansion.

The work of Robinson (2003) showed that the control of leaf area, mature epidermal cell area, and the processes of leaf area expansion and cell division and expansion are under the control of many genomic loci (QTL) within a QTL mapping pedigree produced from a cross between the *P. trichocarpa* and *P. deltoides* species used for this investigation. Comparative studies within these two species may therefore offer a means of identifying genes controlling or influencing leaf development and examination of whether identified genes are located within the genomic regions identified by Robinson (2003).

The expansion of the leaf blade (lamina) is the result of both cell division and cell expansion. In monocotyledonous leaves, these processes can be isolated into clearly defined regions of the leaf (see Fiorani *et al.* 2000). In dicotyledonous leaves, the division and expansion of cells occurs without any defined spatial or temporal boundaries (Taylor *et al.* 2003).

Many genes are now known to function in the control of leaf area development including those affecting rates of cell division, cell expansion and those that influence rates of expansion in the various dimensions of the leaf (Fleming 2005). The first report of an *Arabidopsis* leaf mutant phenotype was made by Rédei in 1962 (cited in Tsukaya 2003). He described the *angustifolia* (*an*) mutant, which has narrow and thick leaves compared to the wild type. This phenotype was originally used as a visible marker for chromosome mapping, as were many early-reported mutantations. It has subsequently been shown that the *angustifolia* mutation is specific to lateral organs and that the phenotype exhibited is the result of a defect in the elongation of cells in the transverse (width) leaf plane. No effect on cell number is seen. The mutant phenotype was exhibited in all cell types examined (Tsukaya *et al.* 1994, Tsuge *et al.* 1996). Tsuge *et al.* (1996) reported that parenchymatous cells are particularly strongly affected. Mutant type parenchymatous cells expand considerably less in the leaf width plane and have enhanced expansion in the leaf depth plane, as compared to the wild type. They therefore speculate that the *AN* gene is an important regulator of the polar elongation of leaf cells in the leaf-width direction specifically. Tsuge *et al.* (1996) also reported a mutant with a polar defect in cell elongation in the longitudinal plane of lateral organs and it was named *rotundifolia3-1* (*rot3-1*). Again, no change in cell number was observed as compared to the wild type. The double mutant of *rot3-1* and *an* shows an additive phenotype, suggesting that the two regulators act independently of each other (Tsuge *et al.* 1996). Hanson *et al.* (2001) reported that the over-expression of an HD-ZIP transcription factor, *ATHB13*, resulted in a similar phenotype in the cotyledons of the over-expresser to that of *an* mutants but only when over expressing mutants were grown on high concentrations of sugars. Unlike *an* mutants, it was only the epidermal cells of the *ATHB13* over-expresser that exhibited abnormal polar elongation. The possible role and function of *ATHB13* remains unknown. The *ROT3* gene has been shown to encode the Cytochrome P450 CYP90C1 (Kim *et al.* 1998), which has regions homologous to those of steroid hydroxylases. Kim *et al.* (1999) examined the expression of the *ROT3* gene and found that it did not exhibit organ specificity. However, over-expression of the *ROT3* gene resulted in plants with longer leaves than the wild type but with leaves of wild-type width. These findings suggest that response to the gene product may be organ specific and that there exists an, as yet, unidentified interaction.

Mutations affecting only the number of cells along a leaf axis have also been identified. The *an3* mutant has narrow leaf blades of normal length and the *compact rosette* (*cro*) *4-1* mutant has short and narrow leaves (Tsukaya 2002). Both mutants have cells of normal size, with only the number of cells being affected. This is evidence for the regulation of both cell division and cell elongation along a specific leaf axis being under genetic control and with both processes contributing towards leaf expansion.

In *Populus euramericana*, it was found that spatial patterns of cell production (which was argued as a surrogate for cell division) can be defined across the surface of the leaf epidermis and that the rate of cell production varies over time (Taylor *et al.* 2003). Work carried out on the parents of a *Populus* mapping pedigree (see Bradshaw *et al.* 1994) has shown that leaf area is achieved differently in different species (Ferris *et al.* 2002, Robinson 2003). Ferris *et al.* (2002) and Robinson (2003) additionally showed that QTL can be mapped for cell area. Cell area is a normally distributed trait within the F_2 progeny of the mapping population (Robinson 2004), indicating that it is under quantitative control of many genes, and that the control of these genes is segregating within the population.

It has long been known that drought stress causes a rapid reduction in the rate of leaf area expansion in many species, although the exact mechanisms through which this is achieved are not yet understood (Zhu 2001, Wilkinson and Davies 2002, Sibole *et al.* 2003). Reduced leaf growth is a response of particular importance in *Populus* as leaf size and canopy longevity are closely linked to productivity in this genus (Rae *et al.* 2004). Perturbations to the normal developmental pattern, such as those induced by drought stress, offer potential for the identification of genes involved in the control of leaf development, and which may be difficult to identify during the normal developmental pattern of plants grown under control conditions.

P. deltoides and *P. trichocarpa* exhibited divergent responses to drought stress at the level of leaf area development. *P. deltoides* leaves showed a significant reduction in leaf area at all LPIs and a reduction in mature leaf area in response to both chronic and acute drought stress, with the reduction being significantly greater in the acute stress

(Figure 2.9a). In contrast *P. trichocarpa* leaves showed no leaf area response to either drought stress. However, it can be observed in Figure 2.9b, that in the acute drought stress no measurements were possible beyond LPI 6 for *P. trichocarpa* as more mature leaves formed rapidly spreading necrotic lesions followed by leaf shedding, making measurements impossible.

The results shown in Table 2.1 and Figure 2.8 show that the observed responses to drought stress in both species were the outcome of altered LER and cell expansion / division control. Although no significant changes in leaf area or LER were observed in *P. trichocarpa*, changes were seen in the epidermal cell area and cell production rate in response to drought: cells became smaller and production rate was increased. This may represent a re-modelling of the leaf development process in response to drought. In contrast, *P. deltoides* showed reductions in cell area, production rate and LER with the combination of these accounting for the significant reductions in leaf area observed. These two contrasting responses may represent broader-scale adaptive stress mechanisms employed by the two species.

It is somewhat surprising that still so little is known about the fundamental process of leaf development. Although the first in-roads towards an explanation have been made through the isolation and characterisation of mutations such as *AN* and *ROT*, much more remains to be explained. There may be many reasons for this apparent immaturity of understanding: Studying leaf development is an experimentally intricate problem; many of the genes of importance will only be expressed in the apical meristem and a narrow zone below it. Extracting sufficient RNA from exact locations within a meristem is technically demanding and requires exact selection of leaf stages if comparative methods such as microarrays, RT-PCR, or subtractive library screening methods (among others) are to be used. Indeed the problem of selecting leaves of the same developmental age poses a huge problem to any researcher as it is very hard to achieve a collection of plants that mature at exactly the same rate, even when they represent clonal material. This problem means that changes in gene expression resulting from the developmental process and maturation of leaf function and structure will overwhelm the likely slight and subtle expression levels and changes that will account for differences in leaf size and shape. This is the principle reason that advance in the field has thus far relied upon mutant identification and characterisation. It is in

light of these challenges that examining leaf development in response to an environmental perturbation such as drought stress might prove useful. In such a case, a clear difference may be observed between the control and treated conditions. This approach has it own inherent problem: does the stress-response mechanism for altering leaf development function in the normal developmental process? This is a question that can only be answered through functional characterisation of any genes identified. Another comparative method that may prove fruitful and is yet to be exploited is the comparison of closely related species with highly divergent leaf forms and the isolation of genes differing in their developmental profiles. The two species examined in this chapter have proven to serve as an ideal model for such an approach, as have other species being used by poplar researchers in other groups. Questions remain about how to employ them for the isolation of developmental genes: Can point-in-time comparisons of leaves presumed to be at the same developmental stage be used or must the full developmental profile of genes within each species be examined at the dynamics of expression be compared for the two species? It seems likely that the latter will prove necessary and that examination of leaf stages as close to emergence from the meristem will prove the most informative. In the case of *P. deltoids* and *P. trichocarpa*, correlation between mapped QTL regions and any identified genes can also be examined.

It has been shown that two *Populus* species, *P. deltoides* and *P. trichocarpa,* expand leaves through contrasting mechanisms and that this process is controlled spatially. The two species show divergent leaf area responses to drought stress with *P. deltoides* significantly reducing leaf area and expansion rates and *P. trichocarpa* maintaining control-condition rates of leaf area expansion and mature leaf area. These species therefore represent a suitable model system for the study of leaf area development and its control.

Chapter 3

Characterising drought response of *Populus*

3 Characterising drought response of *Populus*

3.1 OVERVIEW

An understanding of drought stress response is essential to the future genetic improvement of *Populus* for the production of biomass as a source of carbon-neutral energy and to ensure the availability of viable genotypes for timbre production in a future environment. We have linked drought-induced changes in leaf development and function to the transcriptional responses in two contrasting species of poplar, *P. deltoides* (ILL-129) and *P. trichocarpa* (93-968), selected from relatively dry and wet environments in central Illinois and western Washington respectively. At the physiological level, distinctly contrasting responses were observed for both leaf development and gaseous exchange parameters: *P. deltoides* displayed rapid stomatal closure and reduced mature leaf area in contrast to *P. trichocarpa*, which showed minimal stomatal closure and no reduction in leaf area. Examination of global transcript profiles within functional gene categories, examined here using a *Populus* cDNA array, mirrors these physiological responses. To identify differentially expressed genes, Partial Least Squares Discriminant Analysis (PLS-DA) was utilised. This efficiently identified 622 ESTs that were differentially expressed in response to drought regardless of species. Major expression changes were observed in cell rescue and defence pathways, cellular structure and organisation, and photosynthesis.

3.2 INTRODUCTION

Plants have evolved a host of response mechanisms to allow survival under less than optimum conditions, since the likelihood is that they will, at some point during their development, be exposed to biotic or abiotic stress. This is particularly true of long-lived species such as trees. Drought stress in particular is the primary cause of crop loss worldwide (Boyer 1982), reducing yields by as much as 50 % (Bray *et al.* 2000), and along with temperature, determines the global distribution of major vegetation biomes (Graham, Graham and Wilcox 2003). According to the latest IPCC report (IPCC 2001), drought frequencies are expected to increase globally over the next 50 years and there will be increasing shifts in rainfall patterns.

Populus is now firmly established as a model organism (Brunner *et al.* 2004, Taylor 2002, Wullschleger *et al.* 2002, Bradshaw *et al.* 2000) and has been added to the list of organisms for which the entire genome has been sequenced (http://genome.jgi-psf.org/Poptr1/Poptr1.home.html). *Populus* as a model offers the opportunity to study stress response in a perennial, deciduous tree, that is also being grown extensively as a commercial biomass crop for the production of carbon-neutral energy (Tuskan and Walsh 2001), often in environments subjected to drought stress (Amlin and Rood 2003) and where yield may be impaired (Lindroth and Båth, 1999, Pinon and Valadon 1997). It is known that drought tolerance varies considerably between genotypes of *Populus*, both inter- and intra-specifically (Marron *et al.* 2002, Gebre *et al.* 1998, Tschaplinski *et al.* 1998, Robison and Raffa 1998, Harvey and van den Driessche 1997, Ridolfi *et* al. 1996, Cochard *et al.* 1996, Gebre and Kuhns 1993) suggesting that the genus provides a good model from which to investigate the genetic architecture and adaptive responses to this and other stresses. Despite extensive physiological and morphological descriptions of *Populus* response to drought, to date, little work has been undertaken to explain differences at the level of the gene.

Plant drought stress responses are typified by a rapid closure of stomata (Marron *et al.* 2002, Wilkinson and Davies 2002, Cochard *et al.* 1996, Cornic 1994) and reduction in leaf growth (Sibole *et al.* 2003, Wilkinson and Davies 2002). Reduced leaf growth is a response of particular importance in *Populus as* leaf size and canopy longevity are

closely linked to productivity in this genus (Rae *et al.* 2004). The next observed response to drought stress is often osmotic adjustment and, later, cellular protective mechanisms, alterations to the plant-cell wall (Niinemets 2001) and lipid-membrane structure (Spiteller 2003). Rapid stomatal closure is of vital importance in *Populus* to protect against xylem cavitation (embolism) as much as to reduce water loss through the transpiration stream (Harvey and van den Driessche 1997, Blake *et al.* 1996, Cochard *et al.* 1996). An inability to close stomata leading to leaf and branch shedding as a consequence of fatal xylem cavitation may be an adaptive mechanism, reducing the demand on the transpiration stream (Rood *et al.* 2000, Chen *et al.* 1997a, Cochard *et al.* 1996). Stomatal closure additionally results in a decrease in intracellular carbon (C_i), causing an increase in the rate of photorespiration and in the production of active oxygen species within the chloroplasts (Lawlor and Cornic 2002, Noctor *et al.* 2002, Foyer *et al.* 1997, Smirnoff 1993). In order to maintain cellular function under such stress conditions, proteins involved in damage repair, detoxification, and correct protein formation may be synthesised *de novo*, as may enzymes that detoxify reactive oxygen species (Lawlor and Cornic 2002, Noctor *et al.* 2002, Bray 1997).

Microarrays have become an essential laboratory tool and have been used to examine the response to osmotic and drought stress (Kawasaki *et al. 2004,* 2001, Kreps *et al.* 2002, Ozturk *et al.* 2002, Seki *et al.* 2002 a,b, 2001 a,b, Bohnert *et al.* 2001, Yale and Bohnert 2001, Posas *et al.* 2000, Rep *et al.* 2000), but research in plants has focused on the model, *Arabidopsis thaliana.* In *Populus*, a cDNA microarray has been developed that consists of 13526 ESTs. This resource therefore represents a wide coverage of the entire *Populus* genome, as discussed fully by Andersson *et al.* (2004). Such global cDNA arrays allow transcript-profiling of essentially the entire genome of a large number of organisms (Deyholos and Galbraith 2001, Richmond and Sommerville 2000, Schuchardt *et al.* 2000). Although of tremendous importance to plant biology, the herbaceous model Arabidopsis may be limited in providing results relevant to long-lived species such as trees that produce secondary meristems, woody tissue, and undergo periods of dormancy.

In this study, two genotypes of *Populus,* a male *P. deltoides* (ILL-129) selected from central Illinois (39 °N), and a female *P. trichocarpa* (93-968) selected from western Washington (48 °N), were exposed to drought stress. These two genotypes have

previously been used to produce an F_2 mapping pedigree (Family 331; Bradshaw and Stettler 1993) that is currently being utilised as part of a biomass-yield improvement study (Rae *et al.* 2004) and for which a large number of QTL for leaf, stem, and biomass traits have already been identified (Ferris *et al.* 2002, Wu, Bradshaw and Stettler 1998, 1997, Bradshaw and Stettler 1995). It is also known that the two species may exhibit contrasting responses to drought stress, with *P. trichocarpa* found predominantly in wet, riparian habitats in the north and west of the USA, whilst *P. deltoides* may tolerate much drier environments across the central and eastern parts of the American continent, where it is native. Here we present the physiological and transcriptional responses of the two genotypes to drought stress.

3.3 MATERIALS AND METHODS

Plant material and growth conditions

Plant material was grown as in Materials and Methods, Chapter 2.

Physiological measurements

Leaf Plastochron Index (LPI) was calculated as in Erickson and Michelini (1957) and shown to be applicable to poplar by Larson and Isebrands (1971). Detailed discussion can be found in Taylor *et al.* (2003). The time per LPI was calculated by examining the growth of individual leaves. Leaf area of each LPI was recorded 61 DAP. A digital image of each leaf and a scale bar was obtained using a Nikon CoolPix 5000 camera (Nikon, UK) and leaf area calculated from the image using MetaMorph (Version 5, Universal Imaging Corporation, Philadelphia, USA).

Physiological gas exchange data for stomatal conductance and rate of photosynthesis were collected using a LiCor 6400 IRGA (LiCor, USA). Measurements were taken on leaves LPI 0, LPI 5, and LPI 9 one day and nine days after the initiation of drought (52 and 60 DAP). IRGA chamber parameters were: PAR 300 μmol m^{-2} s^{-1}, CO_2 400 μmol mol^{-1} and constant fixed flow of 400 μmol s^{-1}. Measurements were made on attached leaves at mid-day.

Microarray experimental design and analysis

In total, 12 successful microarrays were hybridised. Initially a loop design was used but slides comparing the two species directly failed. Therefore the 12 arrays used represent three independent biological replicates from each species and each treatment (total of 12 independent biological sample with n=3 at each level) with direct comparisons made between treatment within each species. Species comparisons were made indirectly by comparing across arrays. A technical replicate of each slide was made with dyes in the reverse orientation. Cui and Churchill (2003) discuss the proper statistical handling of such a design.

The microarrays used are detailed in Andersson *et al.* (2004). A DNA microarray was developed containing 13,490 elements spotted in duplicate. These elements were

selected from 36,354 ESTs obtained from seven cDNA libraries. All EST sequences were annotated and functionally classified.

Material for microarray analysis

65 DAP (14 days of drought treatment), leaf LPI 0 from each tree was flash frozen in liquid N for subsequent RNA extraction (Doyle and Doyle 1987: modified by Chang *et al.* 1993) and microarray analysis. RNA was prepared as in Chang *et al.* (1993) with the following modifications. No spermadine was used in the extraction buffer and 2.67% ß-mercaptoethanol was used. An additional extraction step was performed after precipitation with 2.5 M LiCl. RNA concentrations were determined spectrophotometrically (GeneQuant, Amersham-Pharmacia Biotech, Uppsala, Sweden) and RNA quality was assessed by running a 1% (w/v) agarose gel.

cDNA synthesis

50 µg of total RNA suspended in 9 µl DEPC-H_2O was denatured with 1 µl Oligo(dT)-anchor (Cybergene AB) at 70 °C for 5 min and then cooled on ice. mRNA was then reverse-transcribed with 6 µl 5x RT buffer, 0.6 µl 50x dNTP mix (25 mM dA-, dC-, dGTP, 20mM aa-dUTP, 5 mM dTTP), 3 µl 10 mM DTT , 1 µl RNAse inhibitor (30 U, Invitrogen) and 1.5 µl Superscript II (300 U, Invitrogen), and 7.9 µl DEPC-H_2O by incubating at 42 °C for 3 h. The reaction was stopped with 10 µl 0.5 M EDTA, RNA degraded by adding 10 µl 1M NaOH and incubating at 65 °C for 15 min. The reaction was then neutralised by addition of 50 µl 1M Hepes (pH 7.0).

cDNA was purified using Qiaquick columns (Qiagen) according to the included instruction with the exception that the wash buffer was replaced by a Phosphate Wash Buffer (pH 8.0 - 5mM KPO4, http://pga.tigr.org/sop/M004_1a.pdf). cDNA was eluted from the membrane twice with 30 µl ddH_2O with a one min incubation for each elution. Samples were then dried in a Speedvac (Savant, DNA SpeedVac) at 40 °C for 60 mins.

Indirect Cy3/5-Dye coupling

Dyes were coupled by first re-suspending the Cy3/Cy5 (Amersham Biosciences, Uppsala, Sweden) in 120 µl 0.1 M $NaHCO_3$ (pH 9.0). 15 µl of Cy3/Cy5 was then

added to the dry cDNA. Dyes were coupled in the dark for 2.5 - 3 h at room temperature.

Cy3 labelled-cDNA was purified using a Qiaquick column as described above with the exception of an extra washing step and that labelled cDNA was eluted with 41 µl Phosphate Elution Buffer (pH 8.5 - 4mM KPO4, http://pga.tigr.org/sop/M004_1a.pdf) that was incubated on the membrane for one min. The Cy5 labelled target was then eluted into the same eppendorf in the same manner.

Hybridisation

Hybridisation was performed in an ASP (Automated Slide Processor, Lucidea ASP Hybridisation Station, Amersham-Pharmacia Biotech, Uppsala, Sweden). Pre-hybridisation buffer was 50 % Formamide, 5x SSC, 2.5x Denhart's solution. The hybridisation solution contained the labelled cDNA, 25% Formamide, 5x SSC, 0.22% SDS, 1µl tRNA, and 0.42µg Oligo-dA(80mer). Wash buffer 1 was 0.8x SSC, 0.03 % SDS. Wash buffer 2 was 0.2x SSC. Wash buffer 3 was 0.05x SSC, 2 mM KPO_4. 100% Isopropanol was used to clean slides after washing. The ASP used a custom-washing script, treating the slides one by one, giving the slides an identical wash between hybridisation chambers, with the small hybridisation time side effect.

Scanning

Arrays were scanned at 5 µm resolution, using a Scanarray 4000 Microarray Analysis System scanner (GSI Luminomics). Scanner settings were PMT 80% for Cy5, 85% for Cy3 and laser power of 90-99% depending on signal strength.

Image analysis

Spot data were extracted using GenePix (Version 4.1 Pro, Axon Instruments Inc, California, USA). Settings for the spot diameter resize feature were set to <75% and >150%, and CPI (Composite Pixel Intensity) was set to 300.

Data output from GenePix were examined for quality control purposes. A set of custom scripts written in the statistical language R (http://www.r-project.org) were used to perform these checks. The Bioconductor R-packages (www.bioconductor.org) were used for the import of Genepix data. Mean foreground and median background

intensities were used for background subtraction. Additionally, data were filtered based on A-value ($(\log_2 R + \log_2 G)/2$) to remove spots with low intensities in both channels. Log_2 A-value threshold was set to 8.0, this corresponding to a raw intensity of 256. This was used as an alternative to flagging spots with a low intensity in either channel, which can remove useful data.

GeneSpring analysis

Before loading data into GeneSpring, data were LOWESS normalised and the dye-swap and non dye-swap values for each biological replicate were subsequently combined to give a mean value. This data preparation stage was achieved using scripts written in R. These values were then loaded into GeneSpring (Version 6.2, Silicon Genetics, USA) for subsequent analysis giving a biological replication of three. GeneSpring automatically produces a mean of replicate features on an array and these means are used in subsequent statistical analysis. A divide by control channel normalisation was applied to provide ratio of expression values.

PLS-DA analysis

Normalised data were imported into SIMCA P (Version 10, Umetrics, Sweden). Data were scaled using mean centring. For investigating the effects of stress, the arrays were divided into two sets; control and stress. PCA and PLS-DA analysis were then performed. For PLS-DA, a VIP score of 2 was used as the threshold for significant expression (Nguyen and Rocke 2002).

***In silico* comparisons with PopulusDB**

To obtain digital expression profiles of the 100 genes assigned the highest VIP score, the library distribution of all clusters corresponding to these uni-gene clones was exported from PopulusDB. For each of the 100 genes, clone frequency in a particular library was determined and the resulting tables (one for up-regulated and one for down-regulated genes) were analysed according to Ewing *et al.* (1999) and clustered correlation maps were generated. To calculate ROF (relative occurrence factors), the similarities between transcriptomes, occurrence in a particular library (yes or no) was scored for each of the genes. The relative frequency of clusters of the up-regulated list that was found in each of the libraries was calculated, for example 18 of the 74 (24 %)

most up-regulated clusters/genes were found in the young leaf (C) library and 19 of the 26 most down-regulated (73 %). ROF was calculated as the ratio of these relative frequencies, i.e. 24/73.

Real-time quantitative PCR confirmation of array results

qRR-PCR (Real-Time Reverse-Transcription quantitative Polymerase Chain Reaction) confirmation was conducted using the same RNA as for the arrays. SYBR Green (SYBR Green Supermix, BioRad, USA) was used to detect the relative, quantitative amounts of RNA on a BioRad iCycler RT-PCR machine (Bio-Rad, USA). For each biological replicate, four technical replications were performed. cDNA synthesis was preformed using Superscript II (Invitrogen), and cDNA concentrations were measured using a NanoDrop spectrophotometer (Nanodrop Technologies, USA). PCR reactions were not optimised but instead the PCR efficiency was calculated using LinRegPCR (Ramakers 2003) and the relative values calculated according to the formula presented by Pfaffl (2001). Graphs and calculations were performed in R using custom scripts.

3.4 RESULTS

Contrasting leaf development and function in response to drought

Trees grown under drought stress manifested contrasting physiological responses: *P. deltoides* showed a significant reduction in leaf area at all LPIs (Figure 3.1), with the greatest reduction in leaf area in young leaves; *P. trichocarpa* showed no significant alteration in leaf area in response to drought. Similar patterns of response were observed for photosynthesis (A, Figure 3.2a), and stomatal conductance (g_s, Figure 3.2b). *P. deltoides* exhibited a rapid and significant closure of stomata in response to drought and a concomitant reduction in the rate of photosynthesis. *P. trichocarpa* exhibited a markedly smaller stomatal response in all except the youngest leaf examined, with photosynthetic rates mirroring this pattern of response. Both species showed a reduction in mature cell area in response to drought, although this reduction was only significant in *P. trichocarpa*, as revealed by *post-hoc* testing of the data (Table 3.1). Cell production rates were significantly altered by drought stress, with the two species showing opposite responses. Cell production rate was reduced in *P. deltoides*. In contrast, *P. trichocarpa* cell production rate was increased in drought combined with a significant decrease in mature cell area, with no significant reduction in leaf area. These data are reflected in the calculation of epidermal cell number given for the two species (Table 3.1). Observed damage at the leaf level also differed between the two species. Figure 3.3 shows the observed pattern of damage / response to stress for the two species. Clearly it can be seen that *P. deltoides* leaves underwent controlled senescence over a period of five days in response to drought. In contrast, *P. trichocarpa* leaves developed necrotic lesions that rapidly spread across the leaf, typically from tip to base. These lesions spread to cover the entire leaf within one day. Leaves did not dehydrate as in senescence and remained attached to the tree for a number of days during which they gradually dehydrated before being shed.

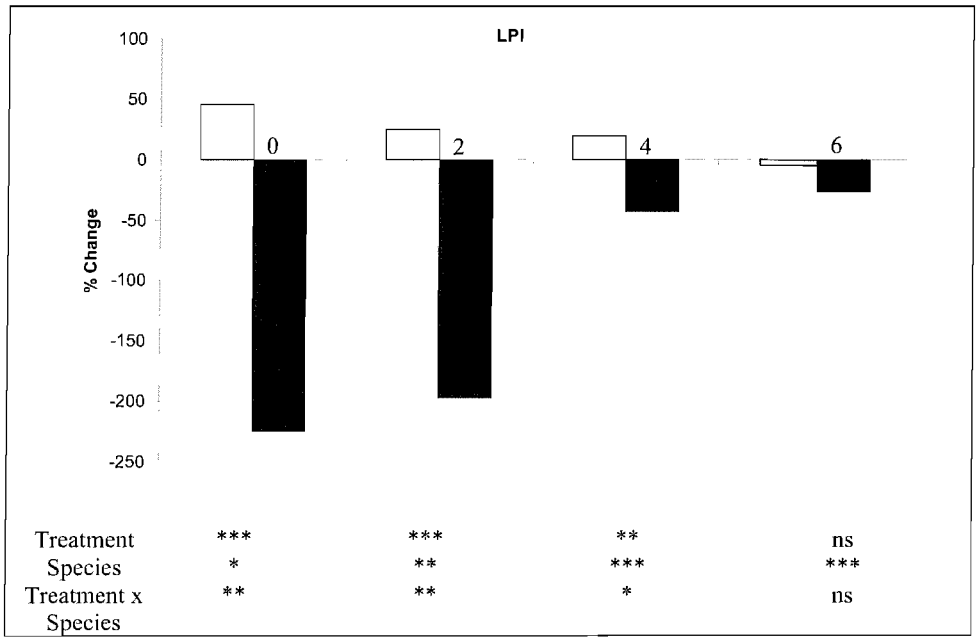
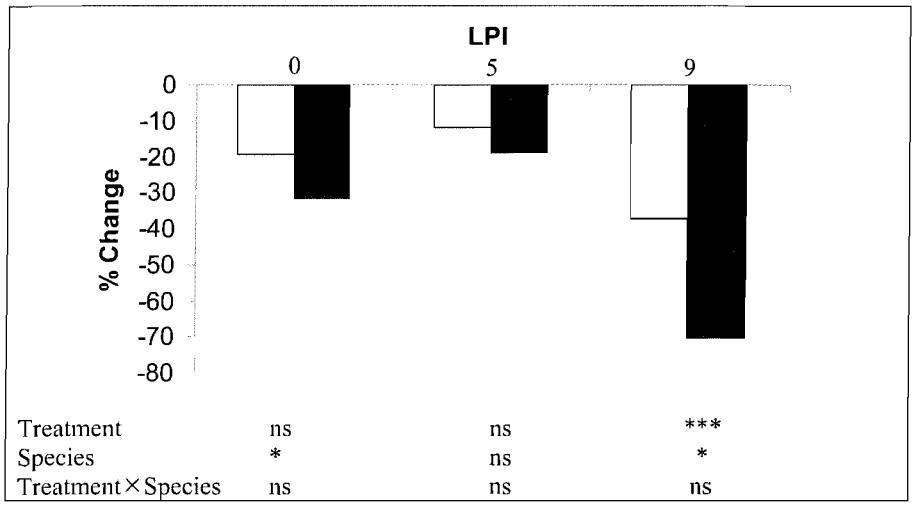


Figure 3.1 Change in leaf area (%) in response to drought at contrasting LPI for *P. deltoides* (■) and *P. trichocarpa* (□). Results of a two-way ANOVA test are shown for each LPI. ns not significant, * $p<0.05$, ** $p<0.01$, *** $p<0.01$. Percentage change was calculated as ([drought-control]/[control])*100 and indicates the percentage difference between the control and treated conditions. Data was collected after 14 days of drought treatment and all measurements were taken at mid-day.

Table 3.1 Cell area, cell production rate, leaf area, and epidermal cell number. Values are calculated from LPI 5 leaves. Data was collected after 14 days of drought treatment and all measurements were taken at mid-day. Asterisks indicate outcome of two-way ANOVA test: ns not significant, * $p < 0.05$, ** $p < 0.01$. Letter superscripts indicate the results of a post-hoc Tukey test.

	Cell area (mm^2 x10^6)		Cell production rate (cells day^{-1} x10^4)		Leaf area (mm^2 x10^2)		Epidermal cell number x10^6	
	P. deltoides	*P. trichocarpa*	*P. deltoides*	*P. trichocarpa*	*P. deltoides*	*P. trichocarpa*	*P. deltoides*	*P. trichocarpa*
Control	458^a	716^b	22.09^a	4.93^b	84^a	30^b	18.34^a	4.19^b
Drought	395^a	432^c	13.70^c	6.76^c	54^c	26^b	13.67^c	6.02^d
Treatment	**		**		**		**	
Species	**		***		***		***	
Species × Treatment	*		ns		*		ns	

(a)



(b)

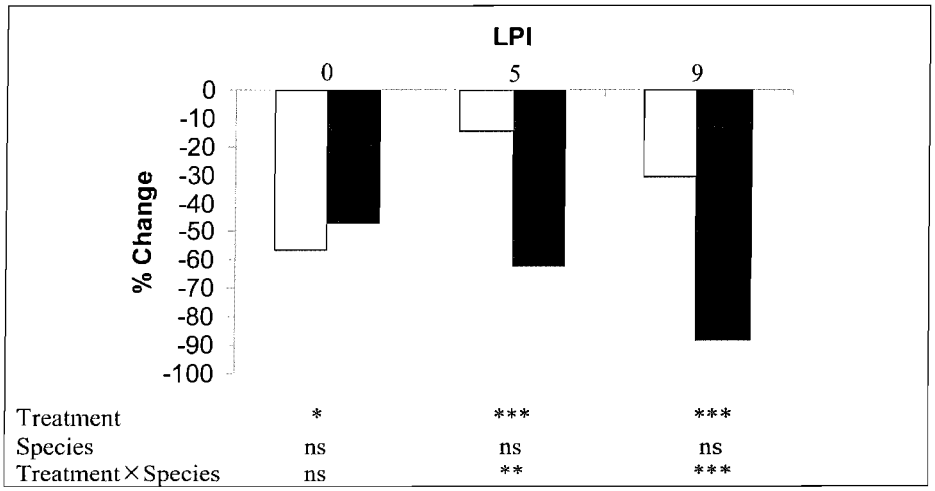


Figure 3.2 (a) Change in rate of photosynthesis (%) in response to drought for *P. deltoides* (■) and P. *trichocarpa* (□) under drought. Results of a two-way ANOVA test are indicated below, **(b)** Percentage change in stomatal conductance for *P. deltoides* (■) and *P. trichocarpa* (□) under drought. The results of a two-way ANOVA test are indicated below. ns not significant, * $p<0.05$, ** $p<0.01$, *** $p<0.01$. Percentage change was calculated as ([drought-control]/[control])*100 and indicates the percentage difference between the control and treated conditions. Data was collected after 14 days of drought treatment and all measurements were taken at mid-day.

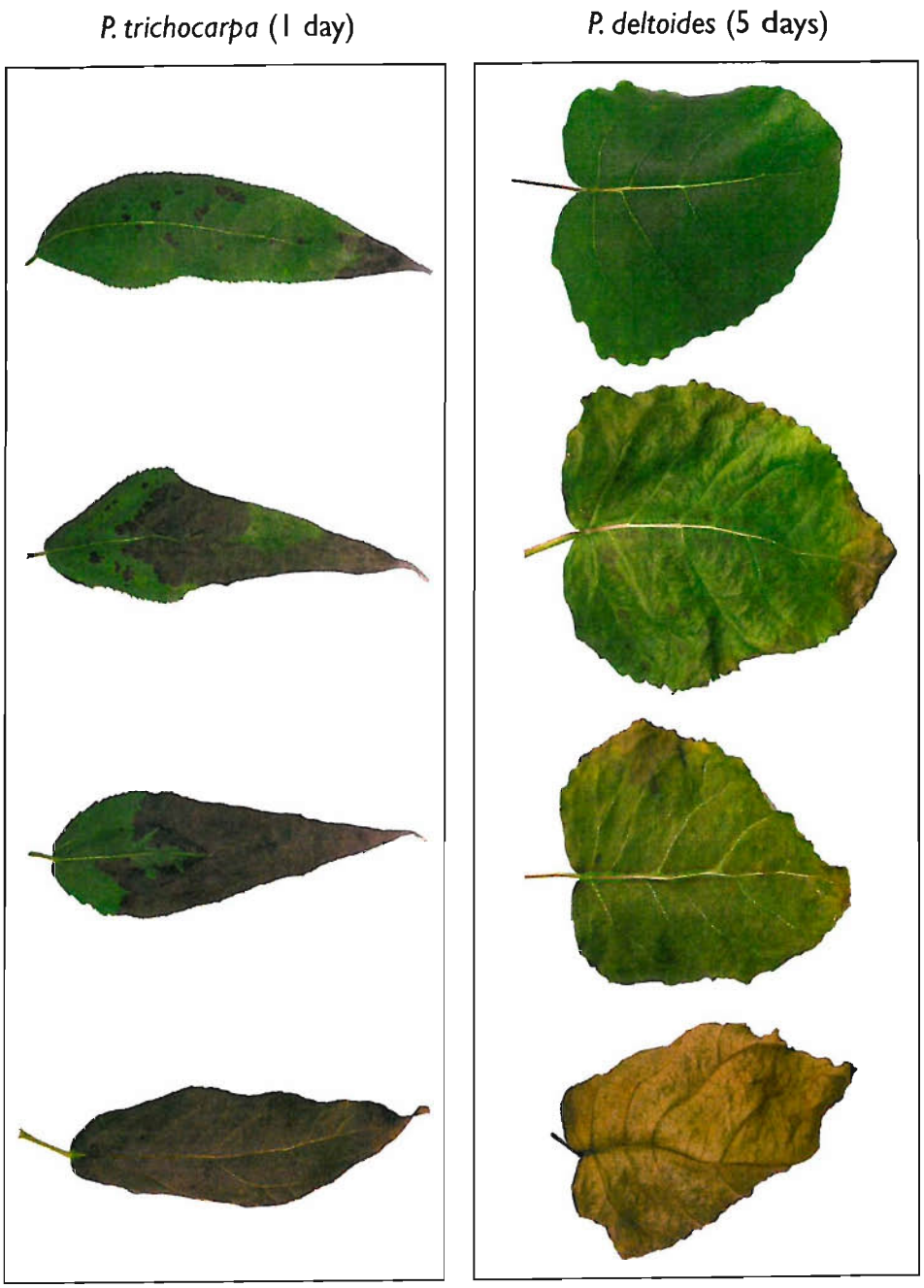


Figure 3.3 The progression of leaf death / senescence in *P. deltoides* and *P. trichocarpa* in response to drought stress. The time taken to progress through the stages after the first appearance of any visible symptoms is indicated in parenthesis for each species and the degree of visible damage at various time points is shown. In both species, symptoms had appeared in semi-mature – mature leaves after 14 days of drought stress.

Common transcriptional response to drought

Clearly, the two species responded to drought stress by contrasting developmental and physiological responses, and on-going studies of the F_2 population derived from these parents suggest that these responses are heritable (see Chapter 4). Common responses in the two parents are shown in Figure 3.4.

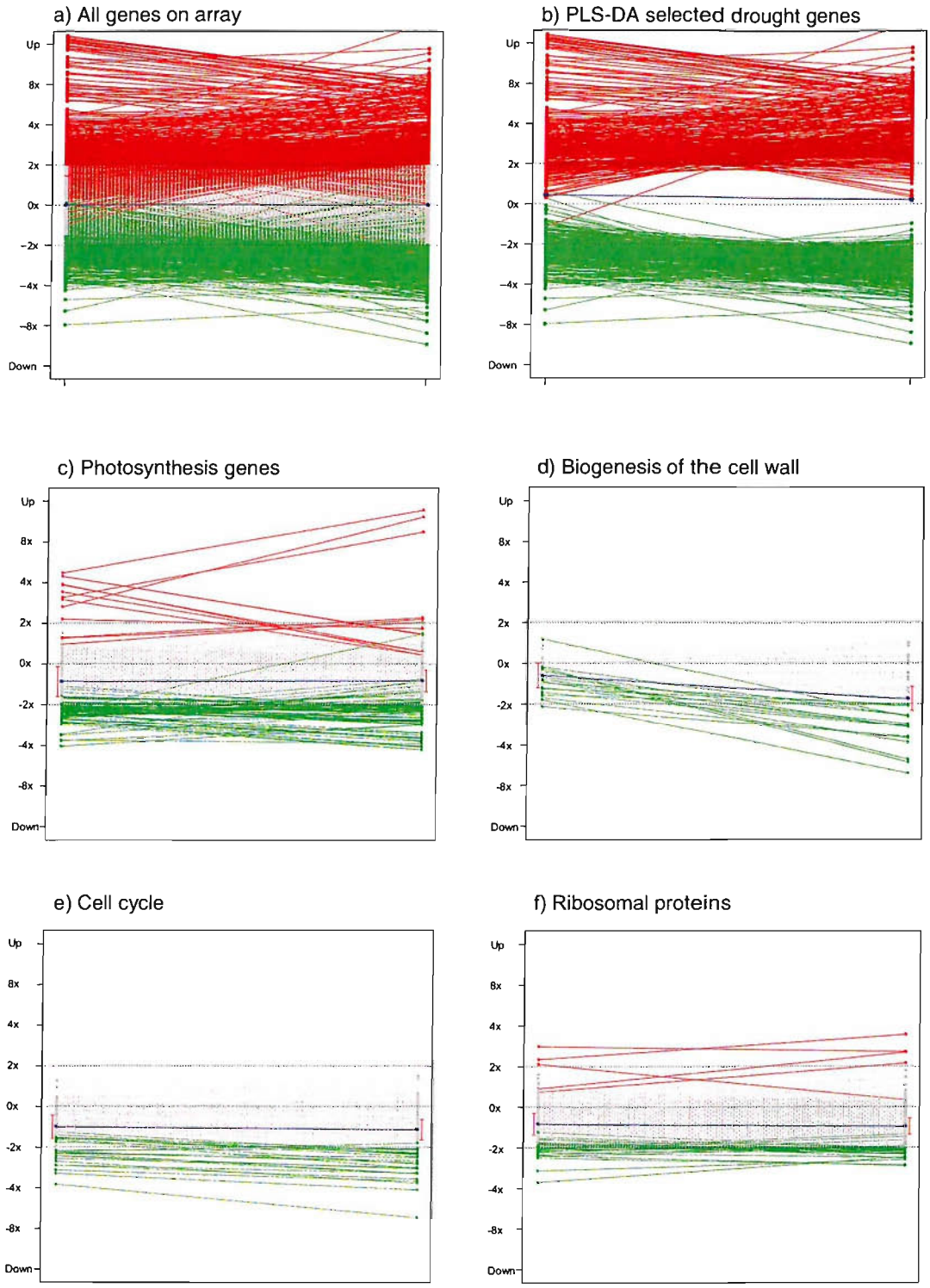


Figure 3.4 Legend on following page

Figure 3.4 The ratio of expression for *P. deltoides* is plotted on the left and *P. trichocarpa* on the right of each plot. Each gene is connected by a line. Lines in green represent genes that are 2-fold or more down-regulated in response to drought in at least one species; genes in red are up-regulated by 2-fold or more; genes in grey have less than a 2-fold change in expression: The mean expression for all genes is represented by the blue line ± s.d. **(a)** The expression ratio for all features on the array is shown, **(b)** the expression ratio of genes identified as differentially expressed using the PLS-DA method, **(c)** the expression ratio for genes in the 'Photosynthesis' UPSC-MIPS category, **(d)** the expression ratio for genes in the 'Cell wall' UPSC-MIPS category, **(e)** the expression ratio for genes in the 'Cell cycle' UPSC-MIPS category, **(f)** the expression ratio for genes in the 'Ribosomal proteins' UPSC-MIPS category.

A global analysis of transcript changes in the two species during drought stress was performed using cDNA microarrays, with three biological replicates. To ensure that microarray analysis results were a true reflection of mRNA level changes in the leaves examined, qRR-PCR (quantitative real-time reverse transcription-PCR) was performed to confirm the array results. Five genes were selected from the microarray analysis that were significantly up-regulated, and five that were significantly down-regulated, and their expression in both species was analysed. Figure 3.5 shows the qRR-PCR results along-side the array results.

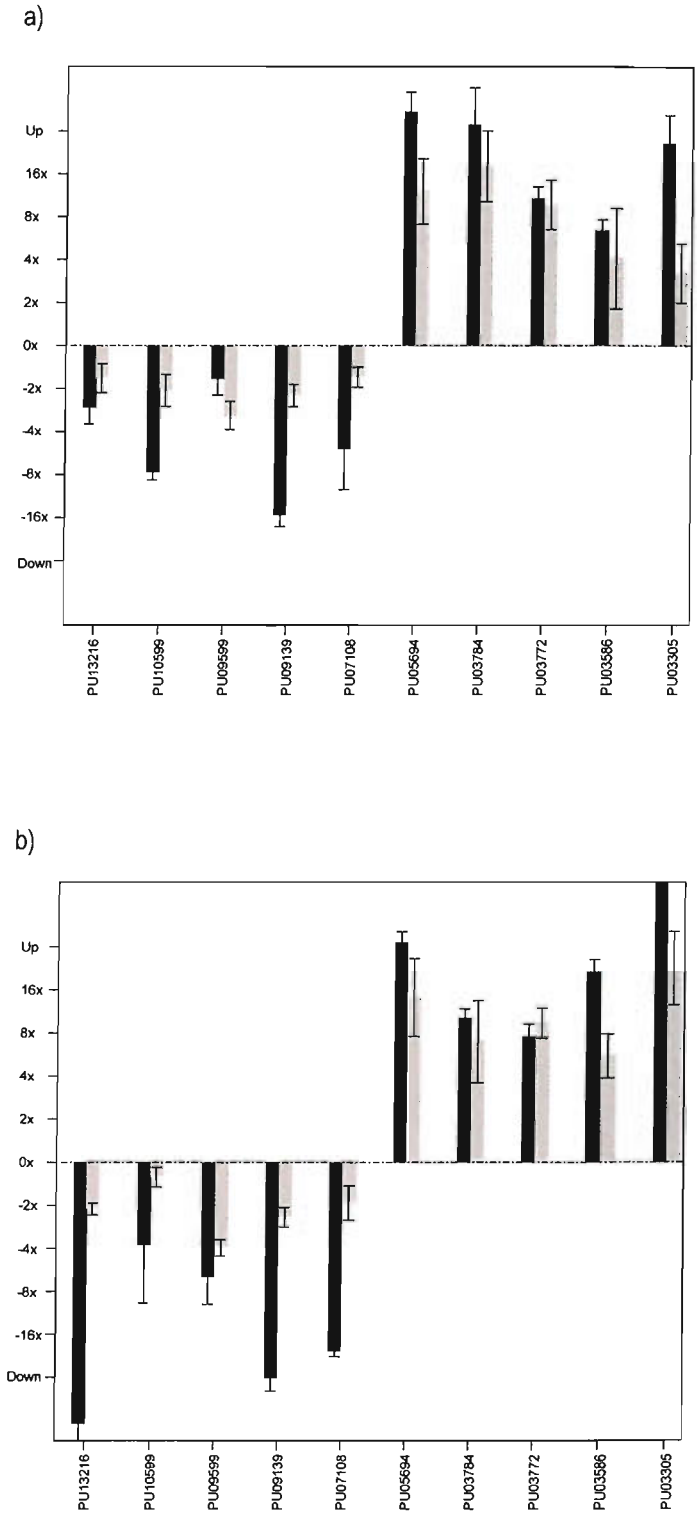


Figure 3.5 The mean expression ratio is shown for 5 up- and 5 down-regulated genes ± s.d. in *Populus deltoides* **(a)** and *P. trichocarpa* **(b)**. Black bars represent the ratios calculated from RT-PCR and grey bars from microarray data. PU number refer to identifiers in the PopulusDB database. Samples for gene expression analysis were taken at mid-day after 14 days of drought.

In general, the ratios between mRNA levels in control and drought samples were higher in the qRR-PCR assay than in the microarray assay, but in all cases the trend was the same; genes that were selected as significantly up- or down-regulated from the microarray data showed the same pattern in the qRR-PCR assay. The fact that microarrays underestimate differences as compared with qRR-PCR is not uncommon and needs to be considered if absolute expression levels are to be determined, but is not of importance for the conclusions presented here.

As in other species studied before, drought stress resulted in a profound remodelling of the transcriptome in *Populus*. Although many genes did not change their expression levels, many others were up-or down-regulated more than 8-fold in response to drought treatment (Figure 3.4a). Figure 3.4a depicts the expression of all genes present on the array: Not surprisingly, the mean expression line lies along the 1:1 ratio. Although some genes with differential expression between *P. deltoides* and *P. trichocarpa* can be observed, the majority of genes have similar expression ratios in the two species. To identify genes that were differentially expressed in response to drought, the multivariate statistical method PLS-DA (Partial Least Squares Discriminant Analysis) was used. In Figure 3.4b, only the 622 that were identified as being differentially expressed in response to drought in both species are shown and it can be seen that the method was highly effective at selecting only genes that were differentially expressed.

To obtain information about the overall changes to the transcriptome that were induced, the 622 differentially expressed genes in response to drought were grouped into UPSC-MIPS (Bhalerao *et al.* 2003) categories (Table 3.2).

Table 3.2 Functional classification of genes identified as being differentially expressed in drought (regardless of species).

Functional process	Up-regulated genes	Down-regulated genes
Metabolism		
amino-acid metabolism	3	0
nitrogen and sulphur metabolism	2	5
nucleotide metabolism	1	0
C-compound and carbohydrate metabolism	6	6
lipid fatty-acid and isoprenoid metabolism	0	3
Energy		
electron transport and membrane-associated energy conservation	1	0
fermentation	2	0
metabolism of energy reserves (glycogen trehalose)	0	2
glyoxylate cycle	1	0
photosynthesis	7	50
Cell Growth Cell Division and DNA Synthesis		
DNA synthesis and replication	0	2
cell cycle control and mitosis	0	1
growth regulators	1	1
Transcription		
mRNA transcription	12	3
Protein Synthesis		
ribosomal proteins	0	5
translation (initiation elongation and termination)	0	1
Protein Destination		
protein destination	0	1
protein folding and stabilization	3	2
storage proteins	10	0
protein targeting sorting and translocation	1	0
proteolysis	13	2
other protein-destination activities	1	0
Transport Facilitation		
ion channels	1	0
ion transporters	2	1
sugar and carbohydrate transporters	0	1
lipid transporters	2	0
Intracellular Transport		
vacuolar transport	0	1
Cellular Biogenesis		
biogenesis of cell wall (cell envelope)	0	11
Cellular Communication / Signal Transduction		
cellular communication / signal transduction	0	2
intracellular communication	8	7
Cell Rescue Defence Cell Death and Ageing		
stress response	6	7

Functional process	Up-regulated genes	Down-regulated genes
disease virulence and defence	2	8
detoxification	12	3
ageing	1	0
Cellular Organization		
cellular organization	1	0
organization of cell wall	0	20
organization of plasma membrane	2	5
organization of cytoskeleton	0	10
organization of chromosome structure	4	6
chloroplast organization	0	1
Development		
development of eukaryotes	0	1
Unclassified Proteins	102	77
Not classified	97	73

Excluding genes in the not-classified and unclassified categories, the largest expression response was in the down-regulation of genes associated with photosynthesis. Genes involved in cellular organisation also underwent significant down-regulation, particularly those involved in the organisation of the cell wall and cytoskeleton, which include expansins, tubulin alpha chain genes, XETs, extensins, and arabinogalactan (Figure 3.4d). Large up-regulation was seen in the category of cell rescue, defence, death, and ageing, particularly in genes involved in detoxification. The category of protein destination, particularly storage proteins and proteolysis, also contained a substantial number of up-regulated ESTs. The full list of these 622 genes is available as supplementary data (Table S1). Genes of particular interest within this list include genes involved in ABA biosynthesis (protein phosphatase 2C, 9-*cis*-epoxycarotenoid dioxygenase, and lipoxygenase), ABA signalling and ABA binding factors (ABF1, ABF3), and ABA response (WCOR413, RD22, stellar K^+ outward rectifying channel), all of which were up-regulated. Many genes observed to be up-regulated in response to drought by previous studies (Seki 2002 a,b, Bray 2004) were also identified here, including LEA genes (Lea 4, 5-D, 14-A), osmoprotectants, chaperones and other stress-related genes (HSPs, glutathione-s-transferase, LTCOR 11), transcription factors (bzips, zinc finger family, and myb family), alcohol dehydrogenase, and an aquaporin, among others. A number of candidate genes for leaf expansion control including RGA1 (response to gibberellic acid 1) and GAI1 (Gibberellic acid insensitive 1), Phantastica, and a number of cell-cycle genes (cyclin

A1, A2, mitotic check point protein, retinoblastoma) were down-regulated (See Flemming 2005 for a recent review on leaf development).

The expression profiles of all UPSC-MIPS categories (containing all genes, not only the differentially expressed) was examined. Graphs for all categories, in the same format as Figure 3.4, are available for download from the UPSC-BASE website. Many genes involved in photosynthesis were down-regulated in response to drought. Figure 3.4c shows the expression of all genes in the photosynthesis category; the blue line represents the mean expression for this category (almost 1.5-fold down-regulation). However, a small number of genes in this category were up-regulated, including ESTs encoding early light-inducible proteins (ELIPS, which are know to be stress induced; Bhaleao *et al.* 2003, Kreps *et al.* 2002), beta-amylase, sucrose synthase, sucrose-phosphate synthase, catalase 2, pyruvate kinase and the photosystem II 10 kDa polypeptide PsbR. These up-regulated ESTs were not induced to the same extent in the two species (Figure 2.5c). Those encoding beta-amylase, sucrose synthase, and ELIPs were more highly induced in *P. deltoides* while those encoding sucrose-phosphate synthase, pyruvate kinase and PsbR were more highly induced in *P. trichocarpa*.

Other UPSC-MIPS categories that contained many down-regulated genes were 'cell cycle' (Figure 3.4d) and 'cell wall' (Figure 3.4e). Another category with a clear downward trend was 'ribosomal proteins' (Figure 3.4f), indicating that overall protein synthesis was decreased in the drought-stressed leaves. Similar graphs for all UPSC-MIPS categories and sub-categories are provided as supplementary material (available from the UPSC-BASE website). Clearly, the categories identified within the list of differentially expressed genes obtained by the PLS analysis that contained large numbers of genes up- or down-regulated (Table 3.2) provides a good overview of the general transcriptome changes that occurred during drought stress: The leaves were, by large, reprogrammed from growth and photosynthesis to stress response.

By focusing on functional categories, important genes could be missed, for example those that were regulated in the opposite direction to the majority of genes in a category or those that, due to lack of knowledge of their function, have not yet been classified. It is therefore necessary to look directly at the expression changes of individual genes. A statistically robust way of analysing transcriptome differences is to

apply the multivariate statistical method of Partial Least-squares discriminant analysis (PLS-DA), which calculates a VIP (Variable Importance) score for each gene. The higher the VIP score, the more important a gene is in explaining the separation of the unit variables (here, control and drought) and therefore, the more likely it is to be influenced by drought. We performed this analysis on the dataset and Table 3.3 lists the 100 uni-genes with the highest assigned VIP score, either positive (up-regulation in response to drought) or negative (down-regulation). Notably, 74 of these genes were up-regulated in response to drought whereas only 26 were down-regulated. This is likely to be a reflection of the overall transcriptome changes occurring during drought. Since the photosynthesis category by far is the highest expressed category in leaf tissues (about 35 % of ESTs in young *Populus* leaves, Bhalorao *et al.* 2003), the approximately 1.5-fold mean down-regulation of this category makes a large overall contribution to the drought-induced transcriptome changes. As a consequence, genes that were up-regulated during drought were more likely to get a high VIP score. Within the up-regulated genes, a range of functions were observed including many genes involved in stress or pathogen response, organisation of the cell wall and cytoskeleton, mobilisation of reserves, transcription factors, hormonal response and signalling, in addition to those that had no informative annotation (expressed protein). Some of these showed very large changes in expression levels with $\log_2$ ratios exceeding 5 (32-fold induction). Among the most down-regulated genes, several photosynthesis-related genes were, not surprisingly, found as well as genes identified from the functional class analysis presented above.

Table 3.3 Top 100 differentially expressed uni-genes in drought determined by VIP (Variable Importance) score. Where multiple ESTs representing the same gene were present, the replicate with the highest VIP score in both species has been selected. PU identifiers relate to the PopulusDB database.

PU number	Description	Log_2 ratio *P. deltoides*	Log_2 ratio *P. trichocarpa*
PU03140	expressed protein	5.840564	3.913188
PU03578	expressed protein	5.785237	4.104617
PU03243	cellulase, putative	5.598332	3.586209
PU03393	splicing factor like protein	5.549764	3.772832
PU03233	expressed protein	5.544281	4.014845
PU03283	microtubule-associated protein EB1-like protein	5.541052	3.627803
PU03117	Major storage protein. Populus canadensis (Carolina poplar).	5.529708	3.833205
PU03669	protease inhibitor II	5.240601	4.355533
PU03457	expressed protein	4.989634	3.591804
PU03696	Probable coproporphyrinogen oxidase. Neurospora crassa	4.921542	3.591527
PU03543	expressed protein	4.829781	3.68578
PU03380	Isocitrate lyase (EC 4.1.3.1). Ricinus communis (Castor bean)	4.792002	3.470903
PU05694	Late embryogenesis abundant protein Lea5-D. Gossypium hirsutum (Upland cotton)	4.765503	4.688682
PU03772	ATP synthase subunit 8. Ostrinia nubilalis (European corn borer)	4.683751	4.351391
PU03333	expressed protein	4.667735	3.070424
PU03640	RGA1, giberellin repsonse modulation protein	4.6331	3.200594
PU03355	RNA-binding protein mec-8, putative	4.461964	3.743573
PU03138	expressed protein	4.439798	3.194071
PU03099	expressed protein	4.202411	3.381248
PU03741	ubiquitin fusion-degradation protein	4.176422	3.161345
PU03435	expressed protein	3.972959	3.528362
PU09474	expressed protein	3.884947	2.648653

PU number	Description	Log_2 ratio *P. deltoides*	Log_2 ratio *P. trichocarpa*
PU10012	expressed protein	3.866638	2.639249
PU03721	expressed protein	3.790681	2.578839
PU03197	expressed protein	3.776635	2.571973
PU03827	Thaumatin-like protein PR-5b precursor (Fragment). Cicer arietinum (Chickpea) (Garbanzo)	3.752298	1.645333
PU04004	IPP transferase - like protein	3.715061	3.191935
PU09267	expressed protein	3.685141	2.503934
PU03835	expressed protein	3.624672	3.209231
PU03636	RNA-binding protein	3.591912	3.147724
PU06720	plastidial w-3 fatty acid desaturase. Picea abies (Norway spruce) (Picea excelsa).FAD3	3.540047	3.658006
PU03098	Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)	3.432237	2.976725
PU12175	HEN1	3.31664	3.052551
PU03501	expressed protein	3.134053	2.822001
PU13018	Jasmonic acid 2. Lycopersicon esculentum (Tomato)	3.071722	3.047464
PU09211	homeobox-leucine zipper protein ATHB-12	3.033776	3.198226
PU05185	MADS-box protein. Cucumis sativus (Cucumber).ERAF17.	3.022584	2.455127
PU03549	Protein kinase	2.976379	2.558339
PU03555	expressed protein	2.975822	2.452988
PU03208	CONSTANS B-box zinc finger family protein	2.974595	2.593555
PU03201	RING zinc finger ankyrin protein	2.965532	2.229438
PU03302	dnaJ protein homolog atj3	2.894374	2.557773
PU03562	expressed protein	2.881769	3.565079
PU01636	Histone H1C. Nicotiana tabacum (Common tobacco)	2.791939	2.193174
PU03861	expressed protein	2.786669	3.838039
PU03270	AP2 domain protein RAP2.3	2.750677	4.117268
PU03552	Meiosis protein mei2-related	2.697478	3.699562
PU09521	Hypothetical protein PH0688. Pyrococcus horikoshii.PH0688.	2.650567	2.17646
PU03347	Spo0B-associated GTP-binding protein. Bacillus subtilis.OBG	2.6452	2.119967

PU number	Description	Log_2 ratio *P. deltoides*	Log_2 ratio *P. trichocarpa*
PU03691	LTCOR11. Lavatera thuringiaca.LTCOR11.	2.627676	2.65747
PU11293	Cytochrome oxidase subunit II (EC 1.9.3.1).Diadasia rinconis.	2.58361	2.765657
PU07842	EIN3-like protein. Cucumis melo (Muskmelon).CMEIL2.	2.560656	2.63226
PU03556	expressed protein	2.517905	2.894293
PU11307	expressed protein	2.476864	3.588652
PU03456	expressed protein	2.411376	2.823915
PU03305	Seed maturation protein LEA 4. Glycine canescens	2.410458	5.784839
PU11243	CG11759 protein. Drosophila melanogaster (Fruit fly)	2.340974	3.614088
PU11301	NADH dehydrogenase subunit 2. Taenia crassiceps	2.320555	3.51716
PU10362	NAC domain protein, putative	2.316195	2.703296
PU03667	expressed protein	2.312229	3.420828
PU03665	calmodulin-binding protein	2.294455	3.688038
PU03309	Alcohol dehydrogenase 2 (EC 1.1.1.1). Vitis vinifera (Grape)	2.2807	4.378843
PU03101	expressed protein	2.279823	2.774695
PU03219	pore protein homolog	2.275005	2.84056
PU05410	Metallothionein-like protein type 3. Carica papaya (Papaya).	2.233789	3.472763
PU11237	CHP-rich zinc finger protein, putative	2.170257	3.306125
PU12209	senescence-associated protein 12	2.070198	3.514139
PU09928	expressed protein	2.067619	3.993822
PU10175	expressed protein	2.029495	3.687748
PU04020	AAA-metalloprotease FtsH. Pisum sativum (Garden pea)	1.880546	3.325543
PU13174	expressed protein	1.808456	4.182124
PU11789	Metallothionein-like protein. Quercus suber (Cork oak)	1.688919	3.559488
PU10629	expressed protein	1.433377	3.922222
PU12789	Extracellular insoluble cystatin. Daucus carota (Carrot)	1.363146	3.562526

PU number	Description	Log_2 ratio *P. deltoides*	Log_2 ratio *P. trichocarpa*
PU03044	expressed protein	-1.16484	-3.82825
PU09319	GAB3. Homo sapiens (Human)	-1.37502	-3.6079
PU00341	beta-glucosidase like protein	-1.66659	-3.2853
PU12657	seed imbibition protein, putative	-1.90656	-2.8159
PU09091	Carbonic anhydrase. Populus tremula x Populus tremuloides.	-1.96745	-2.84175
PU06397	fasciclin-like arabinogalactan-protein (FLA10)	-1.98123	-2.7555
PU03656	Class IV endochitinase (EC 3.2.1.14). Vitis vinifera (Grape)	-2.07749	-3.93825
PU09075	A588R protein. Paramecium bursaria chlorella virus 1 (PBCV-1)	-2.22047	-2.80228
PU02003	Tubulin alpha chain. Prunus dulcis (Almond) (Prunus amygdalus)	-2.30841	-3.34927
PU09376	Fructose-1,6-bisphosphatase (EC 3.1.3.11). Solanum tuberosum (Potato).	-2.33246	-2.58716
PU08961	thylakoid lumenal 17.4 kD pentapeptide repeat family protein, chloroplast precursor	-2.35542	-2.94516
PU04746	thaumatin-like protein	-2.44545	-2.58643
PU08554	expressed protein	-2.46269	-2.80565
PU09599	Retinoblastoma-related protein 1. Populus tremula x Populus tremuloides.RB1	-2.46368	-2.86544
PU08710	Chlorophyll A/B binding protein precursor. Gossypium hirsutum (Upland cotton).	-2.4993	-2.40406
PU09073	Similar to pericentriolar material 1. Homo sapiens (Human).	-2.52022	-3.10511
PU08934	expressed protein	-2.52638	-2.23208
PU08975	hsp 70-like protein	-2.53282	-2.1928
PU12424	Hypothetical 39.5 kDa protein. Caenorhabditis elegans.T26C11.2.	-2.5403	-3.0711
PU09305	GDSL-motif lipase	-2.57275	-2.25359
PU13488	Ribulose bisphosphate carboxylase small chain, chloroplast precursor(EC 4.1.1.39)	-2.57814	-2.90189
PU09273	Auxin-binding protein ABP20 precursor. Prunus persica (Peach)	-2.61168	-2.27273
PU08514	Chlorophyll A-B binding protein 3C, chloroplast precursor. Lycopersicon esculentum (Tomato).	-2.7416	-2.14384
PU08410	expressed protein	-2.85098	-1.92926
PU08532	expressed protein	-3.39921	-3.34937
PU08650	Auxin-binding protein ABP19a precursor. Prunus persica (Peach)	-3.56151	-3.17343

The EST database PopulusDB (http://www.populus.db.umu.se/) makes rapid digital expression profiling of many *Populus* genes possible in order to gain information on their expression levels in the different tissues and treatment libraries used to construct the microarrays used here (Sterky *et al.* 2004). The library distributions of ESTs (with ESTs typically corresponding to genes) from different clusters was examined for the list of 100 genes with the highest VIP score (Table 3.3). When presented as a clustered correlation map (according to Ewing *et al.* 1999), it is evident that the most highly up-regulated genes (Fig 3.6a) have a very different distribution pattern to the most highly down-regulated genes(Fig 3.6b).

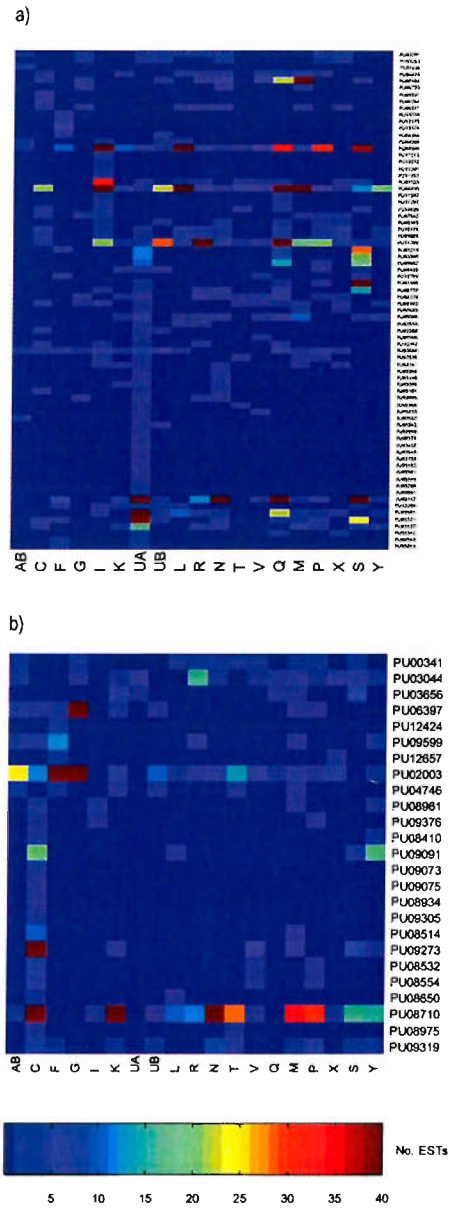


Figure 3.6 The frequency of ESTs location within each cDNA library used to construct the microarray is shown on a false colour scale for **(a)** ESTs up-regulated in response to drought, **(b)** ESTs down-regulated in response to drought.

Library codes are: **R**, Root; **S**, Imbibed seeds, **T**, Shoot meristem; **M**, Female catkins; **C**, Young leaf; **L**, Cold stressed leaves; **P**, Petioles; **K**, Apical shoot; **N**, Bark; **I**, Senescing leaf; **Y**, Virus/fungus-infected leaves; **UA**, Dormant cambium; **G**, Tension wood; **UB**, Active cambium; **AB**, cambial zone; **F**, Flower buds; **V**, Male catkins; **Q**, Dormant bud; **X**, Wood cell-death.

The up-regulated regulon(s) were abundantly expressed in the dormant cambium library (UA) as well as the dormant buds (Q) and senescing leaf (I) libraries. By contrast, the genes in these regulons show low representation in cambial tissues (cambial zone, AB, and tension wood, G), apical meristems (T) and, interestingly, young leaves (C). The genes in the down-regulated regulon(s) were almost entirely expressed in young leaves (C). Such data can additionally be expressed numerically as the comparison of the fraction of up-regulated/down-regulated genes found in each library (the Relative occurrence factor, ROF). If this ratio is 1, genes found in a library have an equal chance of being in the up- or down-regulated dataset; ratios over 1 indicate that genes were more likely to be found in the up-regulated dataset. If the ROF for the different libraries are compared (Table 3.4), the similarity between the transcriptome of drought stressed leaves and those expressed in dormant tissues is evident. In contrast, the regulons that were down-regulated in drought stress show highest similarity to young leaf and cambial-zone transcriptomes. Drought-induced genes often overlap with those induced by low temperatures. For the genes with the highest VIP scores in this experiment this was not evident, as these genes were not abundant in the cold stress leaf library (L).

Table 3.4. Similarities between up-and down-regulated genes during drought and the transcriptome of different *Populus* tissues/treatments. ROF; relative occurrence factor (see Material and Methods). A higher value indicates greater similarity between ESTs present in a library and ESTs identified as being differentially expressed in response to drought.

Library	Material	ROF
UA	Dormant cambium	9.14
Q	Dormant buds	3.05
I	Senescing leaves	2.81
N	Bark	1.76
X	Wood cell death	1.49
R	Roots	1.34
UB	Active cambium	1.19
L	Cold stressed leaves	1.05
S	Imbibed seeds	0.90
G	Tension wood	0.88
P	Petioles	0.78
F	Flower buds	0.65
V	Male catkins	0.65
K	Apical shoot	0.57
M	Female catkins	0.53
T	Apical meristem	0.53
C	Young leaves	0.33
AB	Cambial zone	0.30

3.5 DISCUSSION

This study provides the first genome-wide investigation into the genetic architecture of drought-induced transcriptome changes in two closely-related perennial tree species. Trees are long-lived and likely to be adapted to changing environmental conditions, but investigation of this at the level of the gene has, in the past, been limited by the availability of genomic resources. Sequencing of the *Populus* genome and access to a poplar array (Andersson *et al.* 2004) re-affirms *Populus* as the model tree where such investigations are possible. Here, our approach has been to link transcript profiles to the contrasting physiological responses observed in two *Populus* species.

Contrasting physiological mechanisms of response to drought in *P. deltoides* and *P. trichocarpa.*

P. deltoides and *P. trichocarpa* manifested contrasting physiological responses, perhaps reflecting the adaptation of the species to the drought regimes of their natural ranges. In leaves of *P. deltoides,* senescence was initiated (Figure 3.3), stomata closed rapidly in response to drought with a concomitant reduction in the rate of photosynthesis (Figure 3.2b) and leaf area was significantly reduced at all LPIs tested (Figure 3.1). In contrast, leaves of *P. trichocarpa* formed necrotic lesions in semi-mature to mature leaves, followed by leaf shedding with smaller reductions in stomatal conductance and photosynthetic activity, and no reduction in leaf area. At the onset of leaf necrosis, lesions spread rapidly across the leaf (within hours) in contrast to *P. deltoides*, which took days to develop chlorotic leaves (Figure 3.3). It has been argued that leaf shedding is an adaptive drought response, particularly in perennial species, designed to reduce load on the transpiration stream (Rood *et al.* 2000), typically resulting from fatal xylem cavitation in the leaf mid-ribs (Hukin *et al.* 2005) as opposed to the formation of an abscission zone. It has been shown that *P. deltoides* is actually more prone to xylem cavitation, with the later onset of cavitation resulting from the ability of the species to rapidly close stomata (Rood *et al.* 2000). Cochard *et al.* (1996) have shown that a hybrid of *P. trichocarpa* and *P. koreana* x *P. trichocarpa* cv. 'Peace' is unable to close stomata in response to drought or ABA application and Hukin *et al.* (2005) have recently shown that this clone readily develops mid-rib

cavitations in response to drought. The data presented in Figure 3.2(b) reveal a similar, limited stomatal response for *P. trichocarpa* in response to drought and it is therefore likely that this limited ability to close stomata may lead to rapid xylem cavitation resulting in leaf shedding. Whether this is an evolved, adaptive trait (hydraulic segmentation through cavitation; Hukin *et al.* 2005) or is merely an unavoidable consequence of the xylem structure of *Populus* species is not clear. The results shown in Figure 3.2(b) additionally indicate that the lack of stomatal response may be dependant on the developmental age of the leaf, which is consistent with the findings of Ridolfi *et al.* (1996) and Hukin *et al.* (2005). Liu *et al.* (2001a) showed that increases in leaf ABA concentration are greater in young leaves of *Salix dasyclados* Wimm. than in mature leaves and that a potential source of this ABA is the drying root zone (Liu *et al.* 2001b). It is possible that the same pattern occurs in *Populus spp.*, offering a potential explanation for the age-dependent response to drought in *P. trichocarpa*. Possible causes of the reduced stomatal response include an inability to detect or respond to ABA, a break in one of the signalling pathways that initiates stomatal closure, or the existence of a mechanical restraint preventing stomatal closure. Potential candidates include the K^+ channel gene KPT1 (Langer *et al.*, 2004), which is important in conferring rapid stomatal closure in *Populus* species in response to ABA, OST1 (Assmann 2003) and ABI1 and ABI2 (Lueng, Merlot and Giraudat 1997).

In addition to the differences in leaf shape, the epidermal cellular construction of the leaves also contrasts between the species. This is important since the epidermal cell layer is considered to exert a controlling influence on leaf development (Kutschera 1992). The epidermis of *P. deltoides* leaves is constructed from many small cells compared to the smaller number of significantly larger cells in *P. trichocarpa*, a heritable difference for which QTL have already been identified in the mapping population derived from these two species (Ferris *et al.* 2002). These data are reflected in the values for cell production rate and epidermal cell number given in Table 3.1, with *P. deltoides* requiring a significantly higher cell production rate than that of *P. trichocarpa* in order to achieve its higher mature leaf area. It is interesting to speculate as to whether the control of these contrasting modes of development may be identified at the level of gene expression from the microarray analysis undertaken here, and whether differences in cell cycle and cell expansion-related proteins are apparent for

the species both in control and drought conditions. In response to drought, the leaf growth responses of the two species contrasted markedly: *P. deltoides* reduced leaf area at all leaf ages (LPIs) examined in contrast to *P. trichocarpa*, which showed no significant response in leaf area to drought, although epidermal cell area was significantly affected in this species (Table 3.1).

Drought-induced remodelling of the transcriptome

Regardless of whether the main physiological differences between the two species is the result of their contrasting stomatal responses to drought, both species induced highly similar sets of genes and switched off a highly similar set of genes. Presented here are the results of a profound change in the transcriptome of the two species in response to drought stress. Considerable changes were seen in the expression of genes in the UPSC-MIPS category organisation and biogenesis of the cell wall and cytoskeleton. These expression changes suggest that mechanical alterations were made to the composition, structure, and extensibility of the cell wall and cytoskeleton in order to protect cells against damage resulting from cellular dehydration. The expression of developmentally controlled cell-wall extension was reduced in response to drought (XETs and expansins) while expression of genes associated with hydrolytic-degradation activity (pectinases and cellulases) was increased. This bears similarity to the results reported in Bray (2004). Photosynthetic genes were largely down-regulated (with the exception of some stress-induced genes) and the regulons that were most sensitive to drought-induction were those that are typically associated with dormancy (as evidenced by similarities to the transcriptomes of dormant cambium and dormant buds). Even though drought typically induces leaf senescence, the genes observed to be induced were more similar to dormant tissues than to leaves undergoing autumn senescence, and did not significantly overlap with genes expressed in cold-stressed leaves. These profound changes in the leaf transcriptome during drought stress are likely to be induced in order to protect the leaf from the acute stress imposed, with these protective mechanisms more closely resembling the protection of over-wintering tissues rather than those induced in leaves exposed to acute cold stressed.

Multivariate statistics is a useful tool to delineate complex transcriptional interactions

PLS-DA has shown itself to be a viable and useful alternative over other methods used to analyse array data (Nguyen and Rocke 2002). Examination of the PCA score and loading plots for such data offers insight into how experimental groups and genes cluster together, both in isolation and in relation to each other. It has been shown that this method functions efficiently in the identification of differentially expressed genes. When examining the data separated on a control-drought basis, 622 genes were identified as being differentially expressed in response to drought, with approximately equal numbers being up- and down-regulated.

In conclusion, contrasting physiological and transcriptional responses to drought stress in two species of *Populus* have been identifies. Genes involved in cellular developmental, control and expansion have been identified in addition to stress response genes. It can be concluded therefore that the physiological response to drought for these two *Populus* species differs, providing a good model system to study genetic mechanisms.

Chapter 4

Identifying genomic regions controlling the drought response of *Populus*

4 Identifying genomic regions controlling the drought response of *Populus*

4.1 OVERVIEW

The use of populations of related individuals to map quantitative trait loci (QTL) represents a powerful means of identifying genomic regions involved in the control of polymorphic traits. The mapping of stress response traits represents a particularly effective method of identifying genomic regions containing *trans*-acting regulatory elements controlling the stress response.

An F_2 mapping population from a cross between *Populus trichocarpa* Torr. & Gray (93-968) and *P. deltoides* Bart (ILL-129) known to be highly divergent for a vast range of phenotypic traits was subjected to drought stress. The *P. deltoides* and *P. trichocarpa* grandparents were selected from relatively dry and wet regions of the USA respectively and it was therefore hypothesised that natural variation in response to drought would segregate in the F_2. This can be used as a mechanism to gain insight into the genetic nature of tree adaptation to drought stress. The F_2 population was subjected to soil drying and contrasting responses to drought across genotypes including for leaf coloration, expansion, and abscission were observed. A number of drought specific and drought response QTL were identified and these provide a basis for developing an understanding of the genetic architecture of the *Populus* drought response.

4.2 INTRODUCTION

The study of traits under the influence of many genes (polygenic control / epistatic control) is one of the most challenging tasks facing geneticists and biologists. While some questions can be answered through the application of reverse genetics (anti-sense lines, knock-outs, over-expressers etc), such single gene examinations yield only limited data on the interactive effects that exist in the control of many traits. QTL analysis forms an important part of the reverse genetics approach and plays an essential role in the analysis of polymorphic traits as it identifies the genomic regions contributing to their control.

At the whole-plant level, drought stress response involves changes in morphology, physiology, biochemistry, and gene expression. Traits at any of these levels can be measured and subsequently mapped to identify QTL. Co-location of QTL for multiple traits within a single genomic region may represent the mapping of a *trans*-acting element, although such interpretation must be cautious in the case of traits where strong correlations are expected (for example biomass and height).

The drought stress response of plants is highly complex, involving many different signalling pathways and protection / response mechanisms. It is known that there are at least two key response elements that are sensitive to drought stress: the ABRE controlling ABA-dependent response mechanisms and the DREB controlling ABA-independent mechanisms (Shinozaki and Yamaguchi-Shinozaki 2000). The transcriptions factors DREB1 and DREB2 interact with a *cis*-acting element (DRE) and results in the inductions of stress tolerance genes (Liu *et al.* 1998, Stockinger *et al.* 1997). ABA is the principle plant hormone involved in the drought response and is rapidly accumulated in response to drought stress (Loci *et al.* 2001). ABA promotes the rapid closure of stomata (Leung and Giraudat 1998) and the induction of expression of stress-related genes (Lang *et al.* 1994, Yamaguchi-Shinozaki and Shinozaki 2004, 1994). It is thought that ABA is synthesised from carotenoids (Zeevaart and Creelman 1988) and three principle genes involved in the biosynthesis of ABA have been isolated: Zeaxanthin epoxidase (ZEP; Marin *et al.* 1996) catalyses the epoxidation of zeaxanthin to produce epoxycarotenoid; *9-cis*-epoxycarotenoid

(NCED; Schwartz *et al.* 1997) catalyses the cleavage reaction of epoxycarotenoid to form xanthonin; and abscisic aldehyde (AAO; Seo *et al.* 2000) catalyses the final step of ABA biosynthesis, converting ABA aldehyde to ABA.

Limitations in water supply induced by soil drying result in the closure of stomata, which restricts the flux of CO_2 into the cell, therefore reducing fixation rates. This results in the accumulation of excess light energy entering cells and a requirement for the induction of mechanisms to dissipate the excess light energy in order to protect against photo-oxidative destruction (Schwanz and Polle 2001). This can be achieved through a lowering of photochemical efficiency via the action of the xanthophyll cycle (Demmig-Adams *et al.* 1996) or through maintenance of electron flux through alternative pathways such as the Mehler-peroxidase reaction (Asada 1999, Foyer, Noctor and MorotGaudry 1997). Both mechanisms for dissipation result in the formation of reactive oxygen species (ROS) such as O_2^- and H_2O_2 and the release of transition metals such as Cu and Fe, which increases the risk of oxidative stress through Fenton-type reactions (Moran *et al.* 1994). Antioxidative systems play a vital role in protecting against these induced increases in ROS production. Both enzymatic defences such as superoxide dismutases (SODs) and catalases and peroxidases (Noctor and Foyer 1998), and metabolites such as ascorbate, glutathione and tocopherol contribute to the control of ROS levels in drought-stressed plants (Mittler *et al.* 2004). In order for these mechanisms of ROS scavenging to remain effective their components must be maintained in a reduced state. For example the active form of antioxidants is achieved through the reduction of ascorbate free radicals by the activity of monodehydroascorbate radical reductase and through the consumption of NAD(P)H (Borracino *et al.* 1989) or through the action of the ascorbate-glutathione cycle (Foyer and Halliwell 1976). It has been shown that drought stress results in the induction of antioxidative systems in a number of herbaceous plant species (Smirnoff 1993). For example, a drought tolerant maize strain significantly increased antioxidants in response to drought stress, in contrast to a susceptible strain, which maintained lower levels of antioxidants (Pastori and Trippi 1992), and it has recently been shown that the principle difference between a resistant and a susceptible clone of aspen at the ASPENFACE (http://aspenface.mtu.edu/) facility is a mutation in the copper/zinc superoxide dismutase gene (Pers. Comm. D. Karnosky, MTU, Michigan, USA).

Previous work presented in Chapters 1 and 2 established that the grandparents of the mapping pedigree Family 331 have highly divergent drought response mechanisms and it can therefore be expected that the control of these traits will segregate within the population of F_2 individuals. Examination of drought response traits within the F_2 genotypes provides the opportunity to study the control of phenotypic traits within both control and drought conditions and the response of those traits to drought. In order to identify genomic regions involved in the control of the *Populus* drought response it is essential to measure a suitable set of trait data. The results reported in Chapters 1 and 2 were used to inform a choice of traits that showed divergent responses between the grandparental *P. trichocarpa* and *P. deltoides* species. These included abscission of leaves, leaf chlorosis and necrosis, chlorophyll content, which additionally allows the quantification of carotenoid content, and leaf expansion. It has previously been suggested that chlorophyll content can serve as a good indicator of drought tolerance (This *et al.* 2000).

4.3 Materials and Methods

Growth Conditions

Hardwood cuttings of 167 genotypes of Family 331 were obtained from a field site in the UK (see Rae *et al.* 2004 for details). Cuttings were taken in December 2003 and stored at 4 °C until use and were soaked in cold water at room temperature for 24 hours before planting into 25 l plastic pots filled with lightly compacted JI3 compost (JI3; http://www.jic.bbsrc.ac.uk\corporate/Media_andPublic/compost1.html). Cuttings were planted with one bud above the soil line.

Plant material was grown in a greenhouse (n=2 control and drought) and polytunnel (n=1 control and drought) facility owned by, and near to, the University of Southampton (UK). Before planting, pots were arranged into columns consisting of four rows of pots in the greenhouse and two rows of pots in the polytunnel. Columns had a separation of .75 m. A photographic record of the greenhouse facility is given in Figure 4.1. In the greenhouse, the left-hand two rows of pots within each column were assigned control and the right-hand two drought. In the polytunnel, the two columns in the left side were assigned to control and the two columns in the right side assigned to drought. Neither site was affected by shading from surrounding obstructions so within-site variation in light, temperature, airflow, and humidity was minimal. Average day temp in the greenhouse was 26 °C and in the polytunnel was 28.5 °C. Night temp was 22 °C in the greenhouse and 21 °C in the polytunnel. Day length was uncontrolled. Average humidity in the greenhouse was 60 % and in the polytunnel was 66 %.

Cuttings were pre-assigned random positions within the experiment based on position number of the pots, which were numbered sequentially across the greenhouse and polytunnel. However, it was ensured that two replicates of each genotype would occur in each treatment in the greenhouse and one rep in each treatment in the polytunnel. Each pot was labelled with genotype, position, and treatment information.

Cuttings were planted on 05/06/2004. During establishment, all cuttings were watered daily by overhead sprinklers. Once established, watering continued by use of a hosepipe with each pot being watered daily. The drought treatment was initiated 131

days after planting (DAP). Water was withheld from drought trees for a period of seven days. Drought trees were then given 0.5 l of water and soil drying continued. Soil percentage moisture content was recorded 17 days after drought (DAD) using a Delta-T ML2x ThetaProbe connected to an HH2 Moisture Meter (Delta-T Devices, Cambridge, UK). Average percentage moisture content was 27% and 15% for control and drought treatments respectively. An ANOVA test for treatment and genotype was used to confirm that no genotype or treatment*genotype effects were present. This tests for, and eliminates, the possibility that pots containing small trees dry at a different rate to those containing larger trees.

Physiological data collection

Data previously collected for this population has shown that Plastochron Index (and therefore LPI) differs significantly between genotypes and is therefore an unsuitable method for sampling leaves of equivalent developmental age. We therefore use the leaf that would form LPI 0 as a reference leaf. A 12 mm leaf disc from leaf *l*4 was sampled into DMF (dimethylformamide) for subsequent analysis of chlorophyll content and carotenoid content.

At 18 DAD all plants were scored for leaf number, the number of abscised leaves, the number of leaves with necrotic lesions, and the number of yellow(ing) leaves. Tree height was also scored to the nearest cm. On 9 DAD the petiole of leaf *l*0 was tagged with coloured cotton thread and a digital image containing a scale bar was acquired (Canon EOS 300D, Canon UK Ltd., Surrey, UK). Leaf area, leaf length (along the mid-rib) and leaf width (at widest point at right angle to leaf mid-rid) for both days were then measured from the digital images using ImageJ (http://rsb.info.nih.gov/ij/). Expansion and extension rates and length : width were then calculated.

Figure 4.1 Photographic record of planting arrangement and tree growth in the greenhouse facility. Photographs were taken one day before planting, 14 days after planting, and 60 days after planting.

Physiological data analysis

All data analysis and manipulation was performed in the statistical language R (http://www.r-project.org) using the nlme and nortest packages. Data were first filtered to remove all genotypes that had less than two replicates present in both conditions. For the traits abscission, necrosis, and chlorosis the percentage of leaves that had abscised etc was calculated ([trait value/number of leaves]*100). Percentage effect of drought was calculated for each trait using the formula ([{control-drought}/drought]*100). Normality was then tested for all traits and data transformed using box-cox normalisation where required. Homogeneity of variance was also tested using a Bartlett test. For normally distributed data, a two-way ANOVA test was conducted with genotype set as a random factor. Normality of the residuals was additionally tested using an Anderson-Darling test (all traits returned a non-significant result). Non-normal data was tested using Kruskall-Wallis test. Genotype means were calculated for all traits and exported for QTL analysis. Chlorophyll a, b, total chlorophyll, and carotenoid content were calculated. Measurements of chlorophyll content were made by measuring the absorption of DMF extracts (diluted 1:4) at 647,

664, and 480 nm using a spectrophotometer (Hitachi U20000) as detailed in Wellburn (1994) using the calculations stated for the 1-4 nm range. Subsequently chlorophyll a: b and carotenoid: total chlorophyll were calculated.

QTL mapping

QTL were mapped using the freely available web-based program QTLExpress (Seaton *et al.* 2002). The out-breeding module of the program was used (pers. comm. Anne Rae, University of Southampton). Permutation testing implemented in QTLExpress was used to establish the critical F value for declaring a QTL present (1000 permutation, see Churchill and Deorge 1994). QTL confidence intervals were calculated using an F two drop-off (the cM distance taken for the F value to drop by two from the highest F value location). The genetic linkage map was produced by Tuskan *et al.* (pers. comm.) and consists of 91 SSR markers genotyped on 350 individuals and 92 fully informative AFLPs genotyped on 165 individuals. SSR primer sequences (http://www.ornl.gov/sci/ipgc/ssr_resource.htm) were located on the genome sequence to align the genetic and physical maps and to provide correct orientation of linkage groups (i.e. 3' to 5').

4.4 RESULTS

Physiological drought traits segregate within the population

The percentage effect of drought for the traits total chlorophyll, relative leaf expansion rate, percentage of leaves with visible necrosis and abscission, leaf length extension rate, carotenoid content, carotenoid: total chlorophyll, and absolute leaf area expansion are shown in Figure 4.2 and Table 4.1 gives trait values for the control and drought conditions for the F_2 populations for all traits. The frequency distributions of a subset of these traits are additionally shown in Figure 4.3.

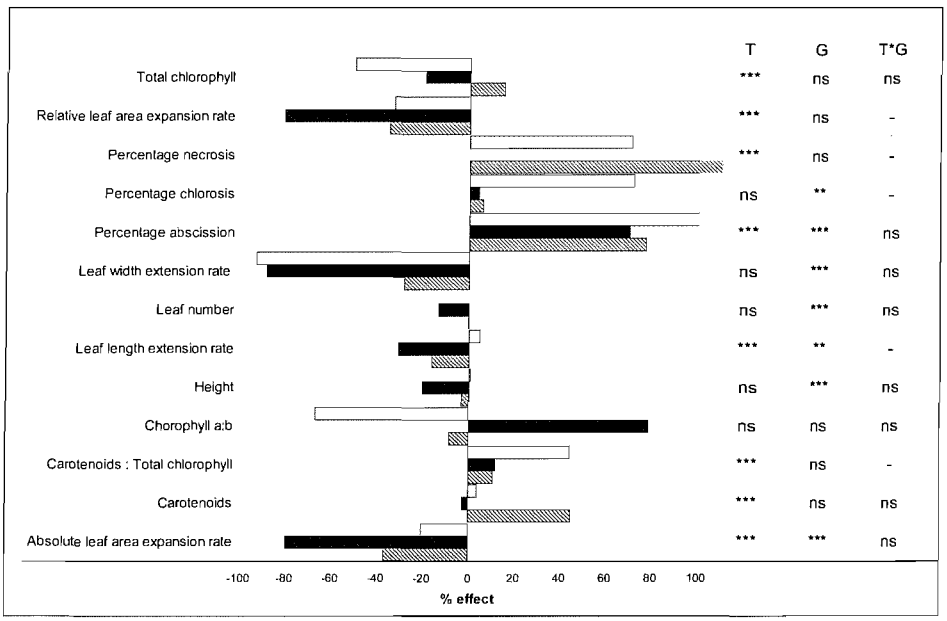


Figure 4.2 Physiological overview of the response to drought of the grandparental *P. deltoides* (■), *P. trichocarpa* (□), and F_2 (▒) mapping population. Values shown are percentage effect. The results of an ANOVA are shown for each trait; * $p<0.05$, ** $p<0.01$, *** $p<0.01$. ANOVA codes are: T - treatment; G - genotype; T*G - treatment by genotype interaction. Percentage effect was calculated as ([drought-control]/[control])*100.

Table 4.1 Summary of physiological trait data recorded in the F_2 population in control and drought conditions. Asterisks indicate outcome of two-way ANOVA test: ns not significant, * $p < 0.05$, ** $p < 0.01$. Letter superscripts indicate the results of a post-hoc Tukey test. +Non-normally distributed trait tested using Kruskal-Wallis test. Percentage effect was calculated as ([drought-control]/[control])*100.

	Trt	Genotype	Trt x Genotype	F2 mean control	F2 mean drought	% effect
Height	0.074	***	ns	90.1	87.1	-3.3
Abscission	***	***	ns	19.7	34.9	77.5
Chlorosis+	ns	ns	-	2.2	6	5.98
Necrosis+	***	ns	-	0.6	5.5	799.88
Soil % water	***	ns	ns	26.4	16.4	-38
Number of leaves	ns	***	ns	22.6	22.5	-0.26
Chlorophyll a	***	ns	ns	14.3	18.2	27.6
Chlorophyll b+	***	ns	-	6.7	6.2	-8.1
a:b+	ns	ns	-	4.1	4.2	2.8
Total chlorophyll	***	ns	ns	21.2	24.4	15.2
Carotenoids	***	ns	ns	2.1	3.1	44.7
Carotenoids : Total chlorophyll	***	ns	ns	0.1	0.1	10.6
Leaf area 1	ns	***	ns	15.24	14.8	-2.89
Leaf area 2	***	***	ns	99.55	64.82	-34.89
Leaf length 1	ns	*	**	6.09	5.76	-5.38
Leaf length 2	***	***	ns	13.99	11.52	-17.65
Leaf width 1	ns	*	ns	4.7	4.97	5.64
Leaf width 2	***	***	ns	11.68	9.26	-20.7
Length: Width 1+	**	ns	-	1.62	1.48	-8.61
Length: Width 2	*	***	ns	1.33	1.41	6.06
Relative leaf area expansion rate+	***	ns	-	6.39	3.75	-41.31
Absolute leaf area expansion rate	***	**	ns	84.31	50.02	-40.67
Leaf length extension rate	***	**	ns	7.9	5.76	-27.09
Leaf width extension rate	***	ns	ns	6.98	4.29	-38.46

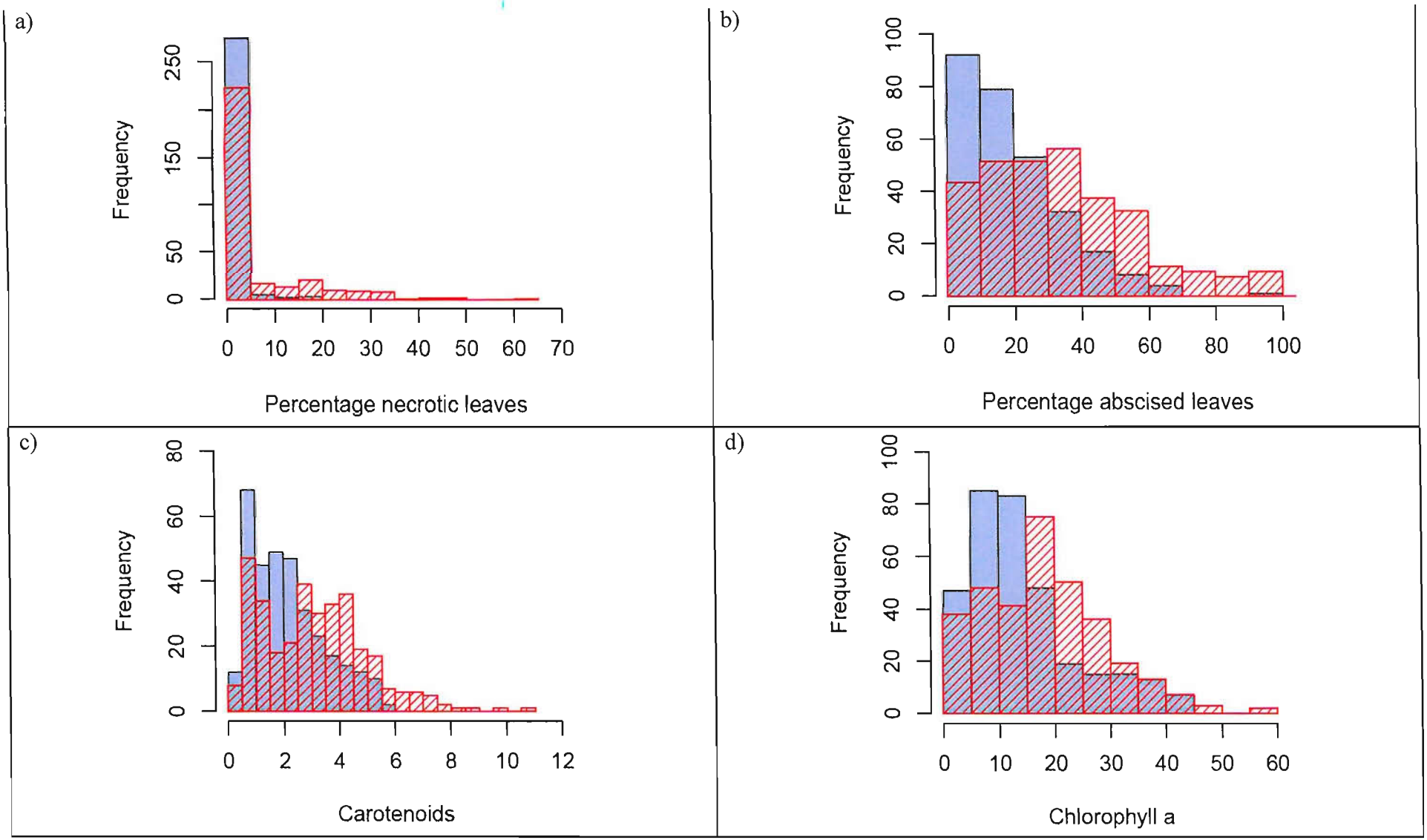


Figure 4.3 Frequency distribution of physiological trait data in control (blue bars) and drought (red hatched bars) conditions. **(a)** Percentage leaves with visible necrosis, **(b)** Percentage leaves abscised, **(c)** Carotenoid content (mg ml^{-1}), **(d)** Chlorophyll a content (mg ml^{-1}). Data are shown in their raw form without normalisation and so are in actual units. All data were normalised prior to statistical analysis.

QTL for all trait data were mapped using the out-breeding module of the freely available web-based program QTLExpress (Seaton *et al.* 2002). The genetic linkage map was produced by Tuskan *et al.* (Pers. comm.) and consists of 91 SSR markers genotyped on 350 individuals and 92 fully informative AFLPs genotyped on 165 individuals. QTL were mapped from the physiological data in control conditions, drought, and for percentage effect response to drought (these being termed 'response' QTL). In total 25 QTL were mapped in control conditions, 44 in drought, and 30 for percentage effect. Information on all QTL including confidence intervals, maternal and paternal effects, percentage variance explained, and statistical information can be found in Table 4.2 and QTL locations are visualised in Figures 4.3-4.7.

Table 4.2 Summary of QTL mapped in the F2 population in control and drought conditions. Trait abbreviations are: **Width 1** – Leaf width 9 DAD (days after drought – see Materials and Methods); **Width 2** – Leaf width 16 DAD; **W. ext.** – Leaf width extension rate; **Length 1** – Leaf length 9 DAD; **Length 2** – Leaf length 16 DAD; **L. ext.** – Leaf length extension rate; **L:W 1** – Leaf length : width 9 DAD; **L:W 2** – Leaf length : width 16 DAD; **Area 1** – Leaf area 9 DAD; **Area 2** – Leaf area 16 DAD; **Abs. exp** – Leaf area absolute expansion rate; **Rel. exp** – Leaf area relative expansion rate; **Ab** – Percentage abscised leaves; **Cl** – Percentage chlorotic leaves; **Nec** – Percentage necrotic leaves; **Leaf** – Number of leaves; **Ht** – Height; **a** – Chlorophyll a content; **b** – Chlorophyll b content; **Total** – Total chlorophyll content; **a:b** – Chlorophyll a:b; **Car** – Carotenoid content; **Car:cl** – Carotenoids : Total chlorophyll

* **C** – control, **D** – drought, **R** – response

Trait	C/D/R*	LG	Position (cM)	Confidence Interval (cM)	F	p	% variance explained
Total	R	I	0	-	6.43	0.02	10
Total	C	I	0	-	5.81	0.014	9
Ht	C	I	11	7-21	6.41	0.029	10
L.ext.	D	I	11	5-21	11.21	0.03	10.9
a	C	I	16	-30	5.44	0.047	10
Car	C	I	19	5-35	5.45	0.038	8
Leaf	D	I	68	41-96	5.33	0.052	10
a:b	D	I	97	78-117	5.62	0.009	9
Cl	D	I	166	157-172	5.91	0.021	11
Width 2	C	II	38	19-60	10.04	0.038	9.8
Leaf	R	III	0	-25	5.18	0.046	7.9
Ht	C	III	2	-22	7.29	0.004	11.4
Leaf	C	III	9	-27	9.35	0.001	14.6
Length 1	D	III	23	7-42	13.34	0.007	12.9
L:W2	C	III	52	26-68	10.22	0.041	10
Length 2	C	III	52	28-61	12.95	0.007	12.5
L.ext.	C	III	53	-65	10.11	0.031	9.9
Nec	C	IV	5	-35	4.99	0.026	9.4
Ht	C	IV	15	-50	4.86	0.034	9
a	R	IV	49	33-76	4.83	0.052	8
Car	R	IV	62	37-82	4.68	0.051	5
Car	D	IV	64	31-	4.44	0.052	7
a	R	IX	0	-3	4.7	0.018	8
Car	R	IX	0	-3	5.52	0.009	9
Car	D	IX	0	-4	3.98	0.03	7
Total	R	IX	0	-4	3.69	0.03	6
Width 2	C	IX	2	-	11.31	0.003	11
Abs.exp.	C	IX	4	-	13.6	0.037	13.1
Area 2	C	IX	5	-	13.47	0	13
Area 1	C	IX	7	1-	8.2	0.024	8.1
Area 1	D	IX	7	1-	7.63	0.043	7.6
Ht	R	IX	7	-	4.27	0.019	6.3

Trait	C/D/R*	LG	Position (cM)	Confidence Interval (cM)	F	p	% variance explained
Leaf	R	IX	7	0-	4.81	0.024	7.2
Length 2	C	V	43	17-62	9.24	0.04	9.1
L:W 2	C	V	44	24-63	9.56	0.042	9.4
L.ext.	R	V	51	30-71	8.55	0.048	8.4
L:W 2	R	V	54	40-63	18.75	0	17.6
Length 2	R	V	59	37-77	9.3	0.041	9.1
Ht	D	V(I)	0	-8	4.94	0.035	7.4
Ht	R	V(I)	2	-18	5.19	0.029	7.9
Ab	D	V(I)	57	37-70	5.63	0.016	11
Nec	D	V(II)	13	-	3.66	0.036	7
Area 2	R	VI	5	1-20	12.27	0.02	11.9
L.ext.	R	VI	10	2-24	12.52	0.018	12.1
Rel.exp.	R	VI	10	3-18	17.5	0.002	16.5
Abs.exp.	R	VI	11	2-20	14.17	0.008	13.6
L:W 2	D	VI	11	-26	13.82	0.008	13.3
Ab	R	VI	34	22-41	4.95	0.052	9
Length 1	D	VI	62	51-70	14.55	0.004	13.9
CI	D	VI	64	55-74	5.28	0.044	9.8
Abs.exp.	D	VI	102	88-120	10.46	0.039	10.2
Area 2	D	VI	102	85-119	10.55	0.037	10.3
Abs exp.	D	VII	13	-21	10.2	0.026	10
Area 2	D	VII	15	2-21	10.3	0.033	10.1
L.ext.	D	VII	15	1-24	10.49	0.011	10.2
Length 2	D	VII	16	9-22	16.63	0	15.8
Rel.exp.	R	VII	16	6-55	9.59	0.024	9.4
L:W 2	D	VII	37	18-	11.85	0.008	11.5
CI	C	VIII	0	-	3.88	0.031	6.8
Nec	C	VIII	0	-	3.88	0.034	7
Length 1	R	VIII	2	-13	10.43	0.008	10.2
Area 1	C	VIII	8	-20	9.86	0.008	9.7
Total	R	VIII	18	7-	7.78	0.001	12
a	R	VIII	19	7-	7.02	0.003	11
Car	R	VIII	25	11-	7.8	0.003	12
Total	D	VIII	25	6-	5.84	0.01	9
a	D	VIII	27	11-	6.3	0.006	10
Car	D	VIII	27	16-	6.7	0.005	11
Car	D	X	20	2-41	4.47	0.046	7
Abs.exp.	C	X	46	12-58	9.18	0.037	9
b	C	XI	14	-26	5.2	0.027	2
b	R	XI	29	-	4.46	0.05	7
W.ext	D	XI	44	33-	10.12	0.027	9.9
Ht	D	XI	55	11-	5.29	0.026	8
Leaf	D	XI	55	35-	4.78	0.035	9
L.ext.	D	XII	20	1-	8.53	0.038	8.4
b	D	XIV	22	10-	4.64	0.025	8
Ht	D	XIV	22	7-	4.6	0.02	6.9
Ab	D	XIV	23	10-	4.65	0.027	9
Ht	C	XIV	28	1-	4.33	0.033	6.4
Leaf	R	XIX	12	-	3.81	0.051	5.4
Area 1	D	XIX	13	-25	10.95	0.01	10.7
Nec	R	XIX	14	-26	5.56	0.007	10

Trait	C/D/R*	LG	Position (cM)	Confidence Interval (cM)	F	p	% variance explained
a:b	R	XIX	17	-	4.69	0.008	8
Length 1	D	XIX	22	13-	13.39	0.002	12.9
a:b	D	XIX	27	-	3.34	0.024	2
Ab	D	XIX	29	7-	5.13	0.017	10
Cl	D	XV	0	-15	4.47	0.043	8
L:W 1	D	XV	0	-17	10.15	0.021	9.9
Length 1	D	XV	0	-28	8.61	0.047	8.5
Car:cl	R	XVII	16	2-32	4.77	0.032	8
Width 1	D	XVII	36	29-59	9.2	0.033	9
Length 1	D	XVIII	11	-28	9.94	0.024	9.7
Width 1	C	XVIII	23	-30	13	0.011	12.5
Car	R	XVIII	36	10-47	5.17	0.033	8
Car	D	XVIII	37	26-49	5.95	0.014	10
Ab	D	XVIII	43	33-57	7.02	0.005	13
a:b	R	XVIII	47	25-62	4.74	0.013	8
a:b	D	XVIII	51	36-69	4.66	0.001	8

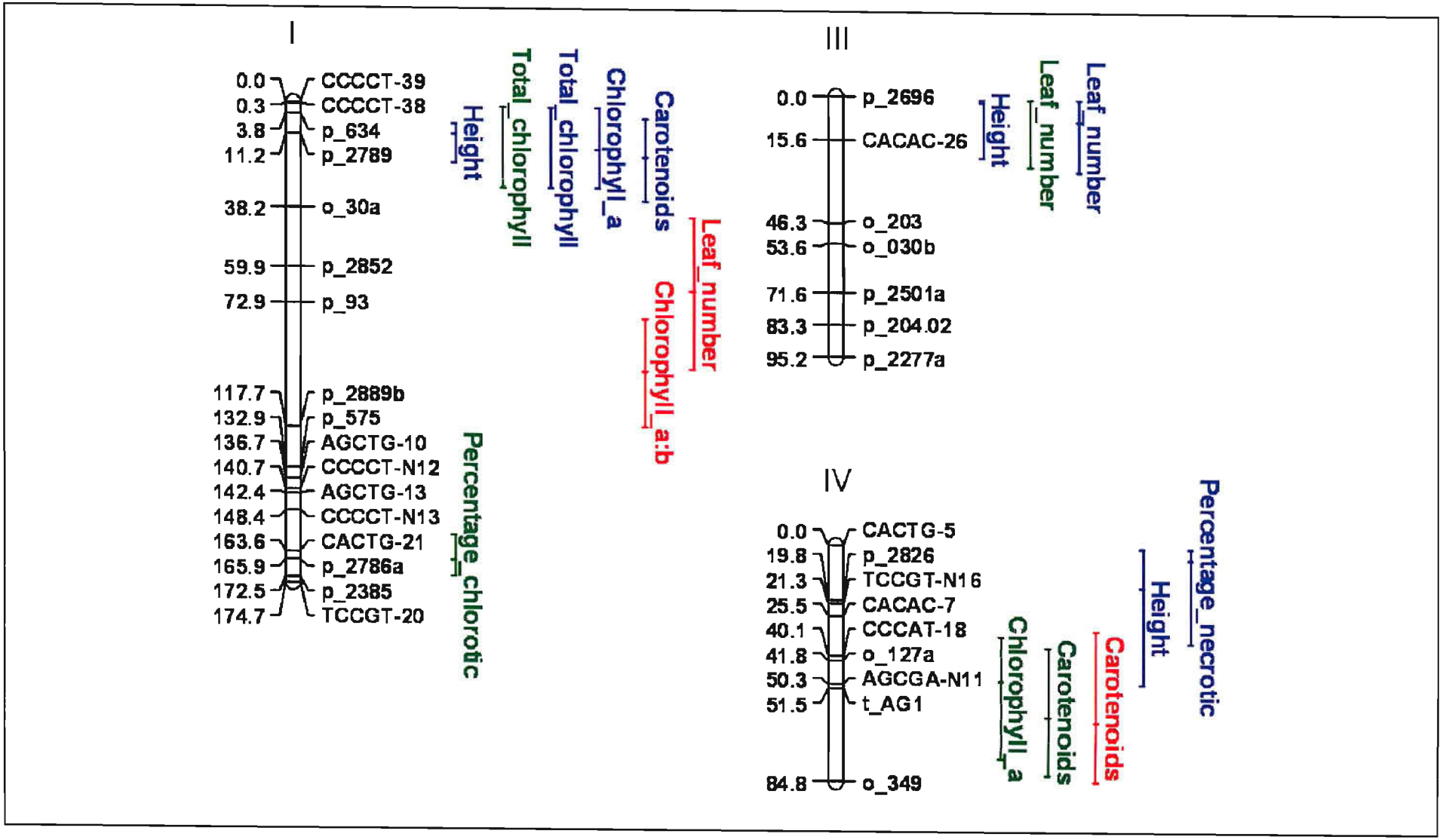


Figure 4.4 QTL locations for trait data recorded in control (blue), drought (red), and response to drought (green). QTL positions are shown ± Confidence Interval as determined by bootstrap analysis. Marker names are shown to the right of the linkage group and cM distances to the left.

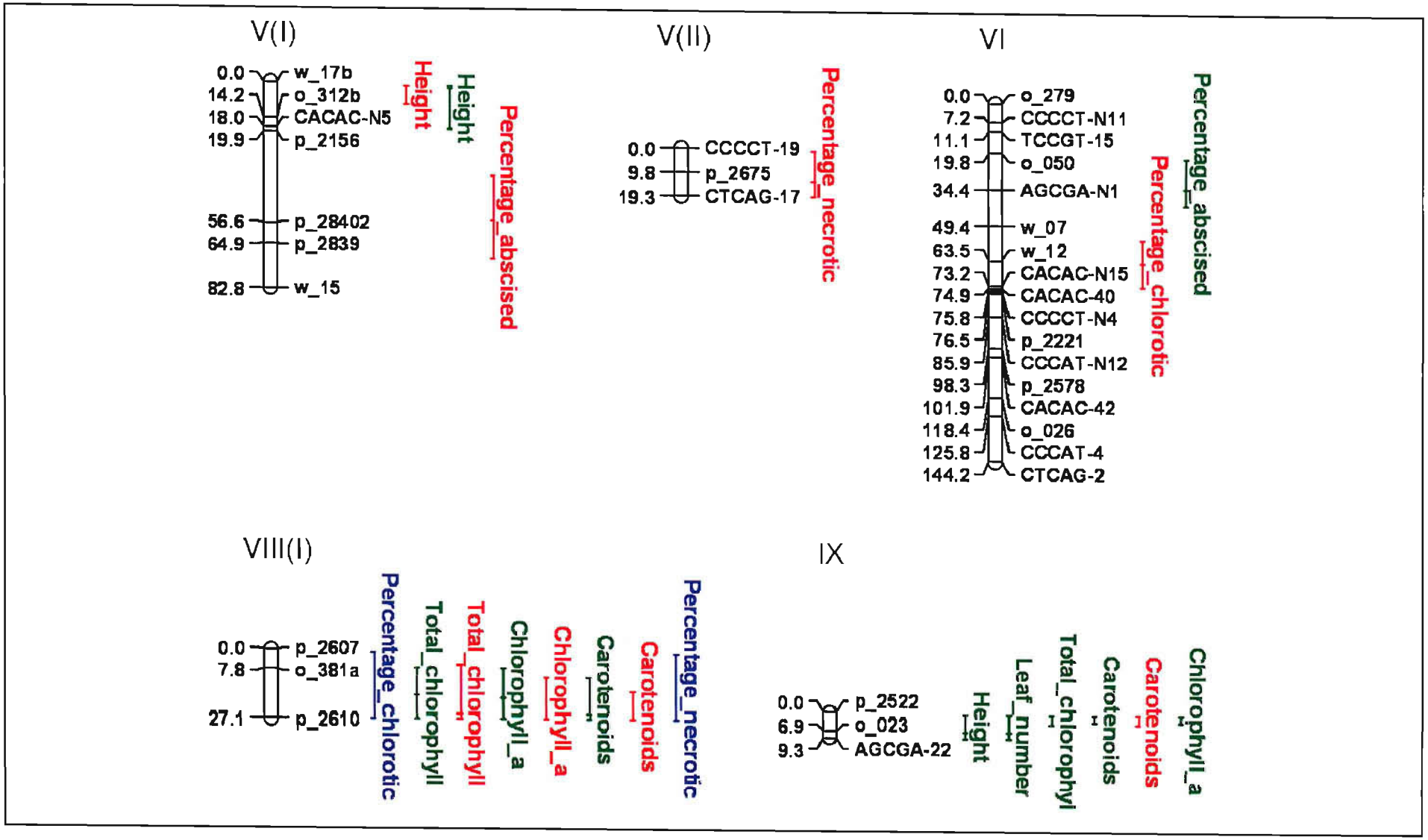


Figure 4.5 QTL locations for trait data recorded in control (blue), drought (red), and response to drought (green). QTL positions are shown ± Confidence Interval as determined by bootstrap analysis. Marker names are shown to the right of the linkage group and cM distances to the left.

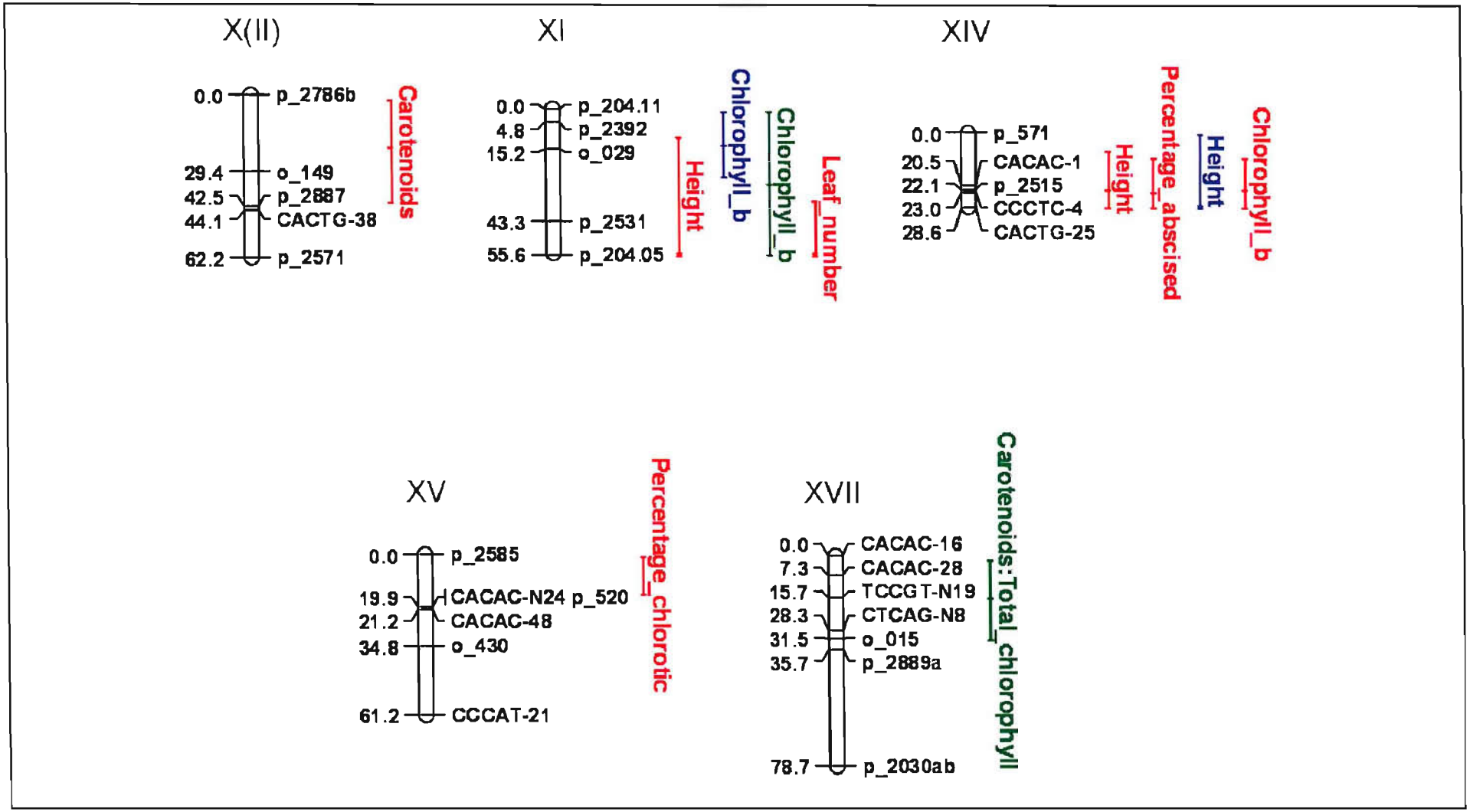


Figure 4.6 QTL locations for trait data recorded in control (blue), drought (red), and response to drought (green). QTL positions are shown ± Confidence Interval as determined by bootstrap analysis. Marker names are shown to the right of the linkage group and cM distances to the left.

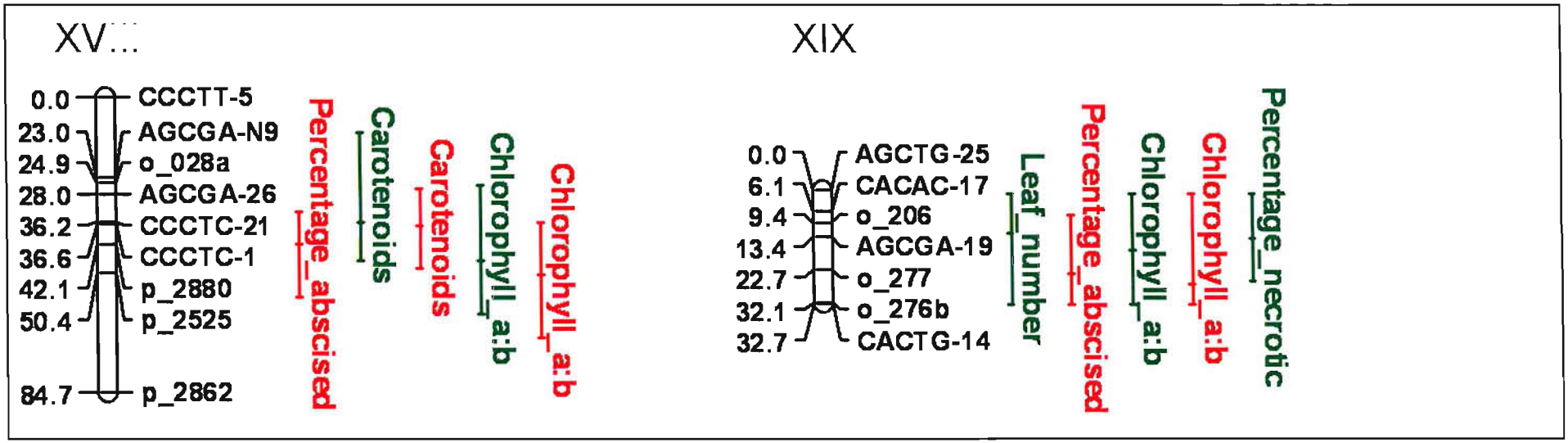


Figure 4.7 QTL locations for trait data recorded in control (blue), drought (red), and response to drought (green). QTL positions are shown ± Confidence Interval as determined by bootstrap analysis. Marker names are shown to the right of the linkage group and cM distances to the left.

4.5 DISCUSSION

Marked transgressive segregation (segregation in both directions that exceeds either parental value e.g. range of leaf areas extends below and above that of either parental genotype) in drought response was observed for all traits recorded in the F_2 population. The traits total chlorophyll, relative leaf expansion rate, percentage of leaves with visible necrosis and abscission, leaf length extension rate, carotenoid content, carotenoid: total chlorophyll, and absolute leaf area expansion were all significantly altered by drought (Figure 2a, Table S1). Significant genotype effects were seen for percentage chlorotic and abscised leaves, leaf length and width extension rate, leaf number, height, and absolute leaf area expansion rate. The percentage of abscised leaves, leaf area expansion, and chlorophyll and carotenoid content were the most highly affected traits in response to drought. Clearly drought stress initiates a remodelling of growth and development patterns and pigment composition and content of genotypes within the F_2 population. The control of this remodelling is segregating within the population resulting in a large diversity in the scale and nature of the induced drought response mechanism.

The marked segregation of drought response mechanisms within the population shows it to be a suitable model system for elucidating the genetic control of drought response. QTL were mapped to identify the genomic regions involved in the control of traits examined and in the control of their response to drought. Six QTL were common to control and drought and 54 were specific to the drought and/or response to drought . It is possible that where a QTL maps only in drought (with no supporting evidence for the presence of a QTL in the control condition) and/or as a response QTL, that this represents the control of a stress-specific response mechanism. Co-location of QTL specific to the drought treatment or response were identified on LGs IV, VI, VII, VIII, IX, XI, XVIII, and XIX (Figures 4.3-4.6). These all represent regions of the genome that function in controlling drought response with effects on multiple drought response/tolerance related traits. It is possible that such a region contains a *trans*-acting element that is affecting the expression of a number of genes functioning in the control of these traits. However, it is also possible that a *cis*-acting element of a structural gene has been located and that this element affects the expression on a gene affecting only

one of the co-locating traits. In such a case a change in the value of one trait is directly affecting another trait. For example, it is possible that the scale and nature of a change in pigment composition (i.e. differential changes in carotenoid content) affects the severity of leaf abscission (see LG XVIII in Figure 4.6). Such a correlation could exist through an increased ability to control ROS levels or though an altered level of ABA biosynthesis, both of which could result from changes in carotenoid content (Iuchi *et al.* 2001, Demmig-Adams 1996). Trait-specific drought/response QTL were also identified for some traits, for example a QTL for height on LG V(I) that explains approximately 8% of the trait variance in both conditions, and these possibly represent the locations of genes controlling the drought/stress response in a trait-specific manner.

Little work has concentrated on the genetic improvement of tree species for drought tolerance, although the requirement to do so will increase as our climate continues to change (Broadmeadow, Ray and Samuel 2005, Broadmeadow 2002). Poplar is currently already grown in regions subjected to drought stress (Amlin and Rood 2003) and where yield may be impaired (Pinon and Valadon 1997, Lindroth and Båth 1999). As the UK, and the wider world, intensifies efforts to increase electricity and fuel production from renewable means, the availability of genotypes of tree that can maintain high yield, desirable mechanical and compositional properties and can tolerate biotic and abiotic stress will become essential. Poplar is particularly suitable for these uses (Tuskan 2001) and the now extensive genetics and genomics base of information and research tools available for the genus will make it the logical choice for future development and improvement programmes (Brunner *et al.* 2004), through either genetic modification or through marker assisted selection (MAS). It is known that drought tolerance varies considerably between genotypes of *Populus*, both inter- and intra-specifically (Gebre *et al.* 1993, 1998, Cochard, Ridolfi and Dreyer 1996, Harvey and Driessche 1997, Robison and Raffa 1998, Tschaplinski *et al.* 1998, Marron *et al.* 2002) suggesting that the genus provides a good model from which to investigate the genetic architecture and adaptive responses to this and other stresses.

QTL mapping provides an invaluable starting point for identifying the genomic regions important to such programmes. Such approaches have been undertaken principally in

crop species, where the emphasis is on maintaining seed yield and quality under stress conditions (Sanguineti *et al.* 1999, Tuinstra, Ejeta and Goldsbrough 1998). Such studies have concentrated on traits known to be tightly associated with drought tolerance such as rooting characteristics (Li *et al.* 2005, Mu *et al.* 2003, Zheng *et al.* 2003, Price *et al.* 2002, Venuprasad *et al.* 2002), ABA content (Landi *et al.* 2005, Liu *et al.* 2005, Sanguineti *et al.* 1999), and flowering time (Xiao *et al.* 2004), where early flowering and seed-set represent a drought escape mechanism (Lanceras *et al.* 2004, Xiao *et al.* 2004). Some recent attempts have been made to integrate studies carried out on different crop species, principally rice, maize, and wheat (Sawkins *et* al. 2004, Salvi *et al.* 2003) with a view to identifying QTL that are important in determining drought tolerance not just between species but across genera. Such an approach will also soon be possible in poplar through comparative mapping comparison of studies in *Salix* (Ronnberg-Wastljung 2005) and poplar (Pers. Comm. Tuskan J, ORNL, Oak Ridge, USA). A degree of linkage between crop studies and studies in poplar can also be achieved through comparison of syntenic regions of the two genomes (Stirling *et al.* 2003). The challenges facing the tree breeder differ considerably to those of the crop breeder principally due to the perennial nature of trees: fitness, yield and quality must be maintained over long periods of time until the crop is ready for harvest (typically 15 years for poplar trees grown as single stem trees). For trees grown as short rotation coppice (SRC), the rootstock must be capable of producing adequate yields over multiple rotation cycles.

It has been shown that the F_2 *Populus* mapping pedigree, Family 331, shows marked segregation for a number of drought-sensitive traits. In particular, percentage leaf abscission and carotenoid content were significantly affected by drought stress. Numerous QTL were mapped in common to both control and drought conditions in additional to , and more importantly, in common to only drought and response to drought conditions. These latter QTL represent areas of the *Populus* genome that are sensitive to drought stress and are involved in the control of stress-response mechanisms. Segregation of response mechanisms within the population may result from the emergence of the divergent stress response mechanisms, perhaps inherited from the grandparental species. These results provide a firm foundation for subsequent studies aimed at elucidating the genetic basis of drought response in *Populus*.

Chapter 5

Linking the genetics and genomics of the *Populus* drought response

5 Linking the genetics and genomics of the *Populus* drought response

5.1 Rationale

This chapter draws together data presented in previous chapters in order to form bridges between the different approaches taken to examine the response of *Populus* to drought stress. To this end, some data presented in previous chapters are re-presented here, but in different form and in context of the approach taken within this chapter.

In order to ensure the success of the approach, the microarray data first presented in Chapter 3 were re-analysed using a newly available analysis method and database resource (UPSC-BASE, http://www.upscbase.db.umu.se/), the results of which are presented in this chapter. Although the conclusions drawn from the new analysis method are broadly similar to the original, the new method highlights some differences between the grandparental species that were not previously apparent. In addition, an alternative functional classification system to that presented in chapter 3 (UPSC-MIPS) became available, due principally to the release of the *Populus* genome sequence and parallel improvements in the bioinformatics resource base available. This new classification system (GO) has been used to re-examine the functional annotation of genes involved in the grandparental drought response at the transcriptional level.

The work presented in this chapter has been prepared as a paper that has been submitted to Plant Journal to appear in a special edition that will be published in parallel to the publication of the *Populus* genome sequence. This chapter appears in the format and language-use of that submission. Figures referred to as supplementary in this chapter are available on the CD included with this thesis or from http://popyomics.biol.soton.ac.uk/~nat/thesis.html. The 'index.html' file on the CD lists the supplementary tables as links to html pages containing the gene lists or to pdf version of figures.

5.2 OVERVIEW

The genetic nature of tree adaptation to drought stress was examined by utilising variation in the drought response of an F_2 mapping population from a cross between *Populus trichocarpa* (93-968) and *P. deltoides* Bart (ILL-129) known to be highly divergent for a vast range of phenotypic traits.

Phenotyping, QTL analysis and microarray experiments were combined to demonstrate that "genetical genomics" can be used to provide information on adaptation at the species level. The grandparents and F_2 population were subjected to soil drying and contrasting responses to drought across genotypes including for leaf coloration, expansion, and abscission were observed and QTL for these traits were identified. A subset of extreme genotypes exhibiting extreme sensitivity and insensitivity to drought on the basis of abscission were defined and microarray experiments were conducted on these genotypes and the grandparents. The different groups induced a different set of genes; 215 and 125 genes differed in their expression response in control and drought respectively, suggesting species adaptation at the gene expression level. Genes preferentially expressed in drought resistant genotypes overlapped with genes expressed in dormant tissues whereas genes involved in meristem function had a lower expression. Co-location of differentially expressed genes with drought specific and drought responsive QTLs was identified and these represent candidate genes contributing to the variation in drought response.

5.3 INTRODUCTION

Plants have evolved a host of response mechanisms to enable survival under stress conditions. Drought stress is the primary cause of crop loss worldwide (Boyer 1982), reducing yields by as much as 50 % (Bray *et al.* 2000), and along with temperature, determines the global distribution of major vegetation biomes (Graham, Graham and Wilcox 2003). According to the latest IPCC report (IPCC 2001), drought frequencies are expected to increase globally over the next 50 years and there will be increasing shifts in patterns of rainfall.

Populus is now firmly established as a model organism (Brunner *et al. 2004,* Taylor 2002, Wullschleger *et al.* 2002, Bradshaw *et al.* 2000) and has been added to the list of organisms for which the entire genome has been sequenced (http://genome.jgi-psf.org/Poptr1/Poptr1.home.html). *Populus* as a model offers the opportunity to study stress response in a perennial, deciduous tree, that is also being grown extensively as a commercial biomass crop for the production of carbon-neutral energy (Tuskan and Walsh 2001), often in environments subjected to drought stress (Amlin and Rood 2003) and where yield may be impaired (Lindroth and Båth, 1999, Pinon and Valadon 1997). It is known that drought tolerance varies considerably between genotypes of *Populus*, both inter- and intra-specifically (Marron *et al.* 2002, Gebre *et al.* 1998, 1993, Robison and Raffa, 1998, Tschaplinski *et al.* 1998, Harvey and Driessche 1997, Cochard, Ridolfi and Dreyer 1996) suggesting that the genus provides a good model from which to investigate the genetic architecture and adaptive responses to this and other stresses. Despite extensive physiological and morphological descriptions of *Populus* response to drought, to date, little work has been undertaken to explain differences at the level of the gene.

Plant responses to drought are typified by a rapid closure of stomata (Hukin *et al.* 2005, Marron *et al.* 2002, Wilkinson and Davies 2002, Cochard *et al.* 1996, Cornic 1994) and reduction in leaf growth (Sibole *et al.* 2003, Wilkinson and Davies 2002). Reduced leaf growth is a response of particular importance in *Populus* as leaf size and canopy longevity are closely linked to productivity in this genus (Rae *et al.* 2004). Osmotic

adjustment, cellular protective mechanisms, alterations to the plant-cell wall (Niinemets 2001) and lipid-membrane structure (Spiteller 2003) all occur as the drought stress progresses. In order to maintain cellular function, proteins involved in damage repair, detoxification, and correct protein formation may be synthesised *de novo*, as may enzymes that detoxify reactive oxygen species (Lawlor 2002, Noctor *et al.* 2002, Bray 1997).

At the molecular level, adaptation and speciation are the result of mutation in protein coding regions or in regulatory regions (Kirst *et al.* 2005, 2004) and, at least in the case of speciation, genome rearrangements. It is of interest to see if differences in drought response, which is a highly adaptive trait, can be used to gain insight into the genetic differences between closely related species. It is now possible to form a bridge between the methods of QTL mapping, the 'candidate gene' approach, and transcriptomics (Weigel and Nordborg 2005, Borevitz *et al.* 2003). Genome-wide poplar arrays can be linked to the physical sequence, providing the genomic location of each EST: QTL mapping aims to identify regions of a genome involved in the control of a trait; it therefore follows that a list of all ESTs on a microarray, and existing within a QTL region, can be examined for differential expression. The expression levels of these genes can then be quantified within a mapping population and the data used to map QTL, with such QTL being termed expression QTL (eQTL, Kirst 2005). When mapping eQTLs, there are two possible outcomes: either a *cis*- or *trans*- regulatory element will be mapped. If the polymorphism lies in a *cis*-regulatory region, then the mapped QTL should co-locate to the gene of interest and is likely to represent a structural gene; if the polymorphism lies in a *trans*-regulatory region, the QTL will locate to the transcription regulator(s). This approach also has potential for identifying transcription regulators acting on a regulon. In this case, the expression of a set of genes would co-locate to a single QTL, which would locate the transcription factor regulating the expression of that regulon. This approach was proposed by Jansen and Nap (2001) but, to date, has not been fully exploited. It was recently shown to be viable in a study of lignin synthesis in *Eucalyptus* (Kirst *et al.* 2004, 2005), cell wall composition of *Zea mays* (Hazen *et al.* 2003), and drought response of *Oryza* (Hazen *et al.* 2004) and other work has examined the co-location of candidate genes with QTL or regions of introgression (Baxter *et al.*

2005, Silva *et al.* 2005). The approach has also been applied to asthma susceptibility in mice (Karp *et al.* 2000), and for ovariole number in *Drosophila* (Wayne and Mcintyre 2002). Limitations to this approach exist - principally that it assumes expression changes serve as the underlying causal mechanism of phenotypic divergence - however it seems reasonable to assume that this will be true in at least some cases. These limitations aside, the combination of these approaches offers new opportunities to investigate causal mechanisms underlying adaptive traits and speciation.

In this study we have used the grandparents and the F_2 mapping pedigree that was initially described by Bradshaw and Stettler (1993), and for which many QTL for traits of adaptive significance are already available. Variation in response to drought in this population was assessed and this was complemented by analysis of the transcriptional response of both the grandparental species and the F_2 genotypes to drought, using a sub-set of extreme genotypes. We hypothesise that this pedigree will be a valuable tool to understand tree adaptation to drought since the grandparental genotypes were selected from a relatively wet (*P. trichocarpa*) and dry (*P. deltoides*) part of the USA respectively.

5.4 Materials and methods

Plant material and growth conditions:

Grandparental response to drought

30 uniform cuttings each of *P. deltoides* (clone ILL-129) and *P. trichocarpa* (clone 93-968) were obtained from a field site in the UK (see Rae *et al.* 2004 for details) Following storage at 4 °C, cuttings were soaked in cold water at room temperature for 24 h prior to planting and following this were planted into rigid plastic tubes (length 75 cm, diameter 13 cm) with one bud above the soil line. The tubes were filled with 10 Kg John Innes No. 3 compost (John Innes Manufacturers Association, Harrogate, UK).

Cuttings were grown in controlled growth rooms at University of Southampton, UK. Day and night temp was 22 °C and 16 °C respectively, PAR 160 μmol $m^{-2}s^{-1}$, day length was 16 h. VPD was uncontrolled but showed limited fluctuation around 50% RH. All cuttings were watered to field capacity for 49 days. The drought treatment was initiated 50 DAP (Days after planting). Water was completely withheld form stressed trees (n=5). Control plants (n=10) were watered to field capacity. Trees were given a balanced NPK fertiliser (Osmocote, Scotts Europe B.V., Netherlands [N 14: P 13: K 13]) throughout the experimental treatment.

Natural variation in population response to drought

Hardwood cuttings of 167 genotypes of Family 331 (referred to as POP1 in the POPYOMICS project, www.soton.ac.uk/~popyomic/) were obtained from a field site in the UK (see Rae *et al*, 2004 for details). Cuttings were stored at 4 °C until used, then soaked in water at room temperature for 24 hours before planting into 25 l plastic pots filled with lightly compacted JI3 compost. Cuttings were planted with one bud above the soil line. Plant material was grown in a greenhouse facility near to the University of Southampton (UK). Before planting, pots were arranged into blocks consisting of four columns of pots running the length of the greenhouse, with 0.75 m gaps between blocks. The two left-hand and two right-hand columns of pots within each block were assigned control and drought respectively. The site was not affected by shading so within-site variation in light, temperature, airflow, and humidity was

minimal. Average day temp was 26 °C and night temp was 22 °C. Day length and VPD were uncontrolled. Cuttings were pre-assigned to random positions within the experiment and each pot was labelled with genotype, position, and treatment. Cuttings were planted on 05/06/2004. During establishment, all cuttings were watered daily by overhead sprinklers. Once established, daily watering continued by hosepipe. The drought treatment was initiated 131 days after planting (DAP). Water was withheld from drought trees for a period of seven days. Drought trees were then given 0.5 l of water and soil drying continued for a further seven days. Percentage soil moisture (% SM) was recorded 17 days after drought (DAD) using a Delta-T ML2x ThetaProbe connected to an HH2 Moisture Meter (Delta-T Devices, Cambridge, UK). Average percentage moisture content was 27% and 15% for control and drought treatments respectively. An ANOVA test for treatment and genotype was used to confirm that no genotype or treatment*genotype effects were present for % SM. Control trees were watered to field capacity throughout. Trees were sampled 17 DAD.

Physiological measurements:

Grandparental response to drought

Leaf area expansion

Leaf Plastochron Index (LPI) was calculated as in Erickson and Michelini (1957) and shown to be applicable to poplar by Larson and Isebrands (1971). Detailed discussion can be found in Taylor *et al.* (2003). The time per LPI was calculated by examining the growth of individual leaves. Leaf area of each LPI was recorded 61 DAP. A digital image of each leaf and a scale bar was obtained using a Nikon CoolPix 5000 camera (Nikon, UK) and leaf area calculated from the image using MetaMorph (Version 5, Universal Imaging Corporation, Philadelphia, USA).

Photosynthesis and transpiration

Physiological gas exchange data for stomatal conductance and rate of photosynthesis were collected using a LiCor 6400 IRGA (LiCor, Nebraska, USA). Measurements were taken on leaves LPI 0, LPI 5, and LPI 9 one day and nine days after the initiation of drought (52 and 60 DAP). IRGA chamber parameters were: PAR 300 μmol m^{-2} s^{-1},

CO_2 400 μmol mol^{-1} and constant fixed flow of 400 μmol s^{-1}. Measurements were made on attached leaves at mid-day.

Natural variation in population response to drought

On 17 DAD, leaf *l6* was flash frozen in liquid N for subsequent analysis of transcript levels. Leaf *l*0 was defined as the first fully unfurled leaf. Data previously collected for this population has shown that Plastochron Index (and therefore LPI) differs significantly between genotypes and is therefore an unsuitable method for sampling leaves of equivalent developmental age. We therefore use the leaf that would form LPI 0 as a reference leaf. A 12 mm leaf disc from leaf *l*4 was sampled into DMF (Dimethylformamide) for subsequent analysis of chlorophyll content and carotenoid content. At 18 DAD all plants were scored for leaf number, the number of abscised leaves, the number of leaves with necrotic lesions, and the number of chlorotic leaves. Tree height was also scored to the nearest cm. On 9 DAD the petiole of leaf *l*0 was tagged with coloured cotton thread and a digital image containing a scale bar was acquired (Canon EOS 300D, Canon UK Ltd., Surrey, UK). Leaf area, leaf length (along the mid-rib) and leaf width (at widest point at right angle to leaf mid-rid) for both days were then measured from the digital images using ImageJ (http://rsb.info.nih.gov/ij/). Expansion and extension rates and length : width were then calculated.

Physiological data analysis

All data analysis and manipulation was performed in the statistical language R (http://www.r-project.org) using the nlme and nortest packages. Data were first filtered to remove all genotypes that had less than two replicates present in both conditions. For the traits abscission, necrosis, and chlorosis the percentage of leaves that had abscised etc was calculated ([trait value/number of leaves]*100). Percentage effect of drought was calculated for each trait using the formula ([{control-drought}/drought]*100). Normality was then tested for all traits and data transformed using box-cox normalisation where required. Homogeneity of variance was also tested using a Bartlett test. For normally distributed data, a two-way ANOVA test was conducted with genotype set as a random factor. Normality of the residuals was additionally tested using an Anderson-Darling test (all traits returned a non-significant

result). Non-normal data was tested using Kruskall-Wallis test. Genotype means were calculated for all traits and exported for QTL analysis. Chlorophyll a, b, total chlorophyll, and carotenoid content were calculated. Measurements of chlorophyll content were made by measuring the absorption of DMF extracts (diluted 1:4) at 647, 664, and 480 nm using a spectrophotometer (Hitachi U20000) as detailed in Wellburn (1994) using the calculations stated for the 1-4 nm range. Subsequently chlorophyll a: b and carotenoid: total chlorophyll were calculated.

A set of extreme sensitive and insensitive genotypes were defined on the basis of percentage response to drought for the percentage of abscised leaves. This trait was selected as a late-onset indicator of drought tolerance. The five highest and lowest ranked genotypes were selected for subsequent transcriptome analysis.

QTL mapping

QTL were mapped using the freely available web-based program QTLExpress (Seaton *et al.* 2002). The out-breeding module of the program was used (pers. comm. Anne Rae, University of Southampton). Permutation testing implemented in QTLExpress was used to establish the critical F value for declaring a QTL present (1000 permutation, see Churchill and Deorge 1994). QTL confidence intervals were calculated using an F two drop-off (the cM distance taken for the F value to drop by two from the highest F value location). The genetic linkage map was produced by Tuskan *et al.* (pers. comm.) and consists of 91 SSR markers genotyped on 350 individuals and 92 fully informative AFLPs genotyped on 165 individuals. SSR primer sequences (http://www.ornl.gov/sci/ipgc/ssr_resource.htm) were located on the genome sequence to align the genetic and physical maps and to provide correct orientation of linkage groups (i.e. 3' to 5').

Microarray experimental design:

Grandparental response to drought

65 DAP (14 days of drought treatment), leaf LPI 0 from each tree was flash frozen in liquid N for subsequent RNA extraction (Doyle and Doyle 1987: modified by Chang *et al.* 1993) and microarray analysis. RNA was prepared as described above and in Chang *et al.* (1993) with the following modifications. No spermadine was used in the extraction buffer and 2.67% ß-mercaptoethanol was used. An additional extraction

step was performed after precipitation with 2.5 M LiCl. RNA concentrations were determined spectrophotometrically (GeneQuant, Amersham-Pharmacia Biotech, Uppsala, Sweden) and RNA quality was assessed by running a 1% (w/v) agarose gel. Further details of the array procedure can be found below in the 'common array methods' section.

In total, 12 successful microarrays were hybridised representing three independent biological replicates from each species and each treatment with direct comparisons made between treatment and control within each species. Species comparisons were made indirectly by comparing across arrays. A technical replicate of each slide was made with dyes in the reverse orientation, as discussed in Cui and Churchill (2003). The microarrays used are detailed in Andersson *et al.* (2004) and are referred to as POP1 arrays on the Umeå Plant Science Centre (UPSC) PopulusDB website (http://www.populus.db.umu.se/). The POP1 array is a cDNA microarray developed containing 13,490 elements spotted in duplicate. These elements were selected from 36,354 ESTs obtained from seven cDNA libraries. All EST sequences were annotated and functionally classified and full information is available from the PopulusDB website (http://www.populus.db.umu.se/).

Extreme-genotype microarray experimental design

Leaf *l*6 from each replicate of the extreme genotypes was used for RNA extraction. RNA was extracted from leaf material as detailed in the grandparental experiment below. However, RNA concentration was measured using a Nanodrop spectrophotometer (Nanodrop Technologies, USA) and quality was checked using a Bioanalyzer (Agilent 2100 Bioanalyzer, Agilent Technologies, Waldbronn, Germany). A pooled analysis strategy was then used for microarray investigation. Four RNA pools were made: High-abscission control (Hc, n=11); high-abscission drought (Hd, n=12); low-abscission control (Lc, n=9); and low-abscission drought (Ld, n=11). The number of replicates within each pool reflects the number of surviving plants but n=>2 per genotype in all conditions. The amount of RNA used to construct the pools from each genotype was normalised to ensure that each genotype had equal representation within the pool. A loop design was then used to make all possible comparisons. The loop was replicated four times with half of the replicates in one dye orientation (Cy5/Cy3) and four in the opposite dye orientation (Cy3/Cy5). The POP2 array was

used for this experiment and contains 24,735 probes representing >100,000 ESTs from 18 tissues (described by Sterky *et al.* 2004). Full details are available from the PopulusDB website.

Microarray methods:

cDNA synthesis

50 µg of total RNA suspended in 9 µl DEPC-H_2O was denatured with 1 µl Oligo(dT)-anchor (Cybergene AB) at 70 °C for 5 min and then cooled on ice. mRNA was then reverse-transcribed with 6 µl 5x RT buffer, 0.6 µl 50x dNTP mix (25 mM dA-, dC-, dGTP, 20mM aa-dUTP, 5 mM dTTP), 3 µl 10 mM DTT , 1 µl RNAse inhibitor (30 U, Invitrogen) and 1.5 µl Superscript II (300 U, Invitrogen), and 7.9 µl DEPC-H_2O by incubating at 42 °C for 3 h. The reaction was stopped with 10 µl 0.5 M EDTA, RNA degraded by adding 10 µl 1M NaOH and incubating at 65 °C for 15 min The reaction was then neutralised by addition of 50 µl 1M Hepes (pH 7.0).

cDNA was purified using Qiaquick columns (Qiagen) according to the included instruction with the exception that the wash buffer was replaced by a Phosphate Wash Buffer (pH 8.0 - 5mM KPO4, http://pga.tigr.org/sop/M004_1a.pdf). cDNA was eluted twice with 30 µl ddH_2O with a one min incubation for each elution. Samples were dried in a Speedvac (Savant, DNA SpeedVac) at 40 °C for 60 mins.

Indirect Cy3/5-Dye coupling

Dyes were coupled by first re-suspending the Cy3/Cy5 (Amersham Biosciences, Uppsala, Sweden) in 120 µl 0.1 M $NaHCO_3$ (pH 9.0). 15 µl of Cy3/Cy5 was then added to the dry cDNA. Dyes were coupled in the dark for 2.5 - 3 h at room temperature. Cy3 labelled-cDNA was purified using a Qiaquick column as described above with the exception of an extra washing step and that labelled cDNA was eluted with 41 µl Phosphate Elution Buffer (pH 8.5 - 4mM KPO4, http://pga.tigr.org/sop/M004_1a.pdf) that was incubated on the membrane for one min. The Cy5 labelled target was then eluted into the same eppendorf in the same manner.

Hybridisation

Hybridisation was performed in an ASP (Automated Slide Processor, Lucidea ASP Hybridisation Station, Amersham-Pharmacia Biotech, Uppsala, Sweden). Pre-

hybridisation buffer was 50 % Formamide, 5x SSC, 2.5x Denhart's solution. The hybridisation solution contained the labelled cDNA, 25% Formamide, 5x SSC, 0.22% SDS, 1µl tRNA, and 0.42 µg Oligo-dA(80mer). Wash buffer 1 was 0.8x SSC, 0.03 % SDS. Wash buffer 2 was 0.2x SSC. Wash buffer 3 was 0.05x SSC, 2 mM KPO_4. 100% Isopropanol was used to clean slides after washing. The ASP used a custom-washing script, treating the slides one by one, giving the slides an identical wash between hybridisation chambers, with the small hybridisation time side effect.

Scanning and Image Analysis

Arrays were scanned at 5 µm resolution, using a Scanarray 4000 Microarray Analysis System scanner (PerkinElmer, Boston, USA). Scanner settings were PMT 80% for Cy5, 85% for Cy3 and laser power of 90-99% depending on signal strength for the acute experiment. For the extreme experiment, arrays were scanned at four settings of increasing laser power and PMT (Laser power – 60, 80, 100, 100 %; PMT – 70, 70, 70, 80 %) at 10 µm resolution, using a ScanarrayLite Microarray Analysis System scanner (PerkinElmer). Regression analysis was then applied to the scans using a UPSC-BASE plug-in. This produces a unified data file from all four scans with the effect of increasing the dynamic range of intensities for which spot intensity data can be extracted (Dudley *et al.* 2002). Spot data was extracted using GenePix (Version 4.1 and 5.0) for the grandparental and population experiments respectively (Axon Instruments Inc, California, USA). Settings for the spot diameter resize feature were set to <75% and >150%, and CPI (Composite Pixel Intensity) was set to 300.

Data analysis

Data output from GenePix were imported into UPSC-BASE, where they are publicly available (http://www.upscbase.db.umu.se/). All data were examined for quality control purposes using plug-ins integral to the UPSC-BASE analysis pipeline (see Sjodin *et al.* submitted). Background subtraction was achieved by removing the local median background intensity from the spot foreground median intensity. A print-tip LOWESS normalisation was then applied and spots flagged as bad had a negative weighting factor of 0.1 applied to them. Data were subsequently filtered based on A-value $((\log_2 R + \log_2 G)/2)$ to remove spots with low intensities in both channels. Log_2 A-value threshold was set to 8.0, this corresponding to a raw intensity of 256. B-statistics implemented in the LIMMA package for R (Smyth 2004;

http://bioinf.wehi.edu.au/limma/) and made available as a UPSC-BASE plug-in were then used to select genes with a high probability of differential expression in response to drought for the grandparental experiment and for all comparisons between the population extreme groups. A B-statistic value of 0 indicates a 50:50 probability of differential expression and as the B value increases, so too does the probability that the gene is differentially expressed. In this way, the B value can be considered as a confidence measure of differential expression. We used an arbitrary cut-off of 5 to yield lists of genes with a high probability of differential expression. FDR adjusted p values are also given. Grandparental species differences were tested by exporting normalised data from UPSC-BASE and importing this into GeneSpring (V 7.2, Agilent Technologies, Redwood City, CA) where a 1-way ANOVA was performed both with and with out Bonferroni correction. These genelists were then imported into UPSC-BASE.

***In silico* comparisons with PopulusDB**

To obtain digital expression profiles of genes with B-statistic scores above 5 for the grandparental and extreme arrays,, the library distribution of all clusters corresponding to represented uni-genes was exported from PopulusDB using a UPSC-BASE plug-in. For each of the 100 genes, clone frequency in a particular library was determined and the resulting tables (one for up-regulated and one for down-regulated genes) were analysed according to Ewing *et al.* (1999) and clustered correlation maps were generated. To calculate ROF (relative occurrence factors), the similarities between transcriptomes, occurrence in a particular library (yes or no) was scored for each of the genes. The relative frequency of clusters of the up-regulated list that was found in each of the libraries was calculated, for example 18 of the 74 (24 %) most up-regulated clusters/genes were found in the young leaf (C) library and 19 of the 26 most down-regulated (73 %). ROF was calculated as the ratio of these relative frequencies, i.e. 24/73.

All microarray data collected from the two experiments is available for download from the UPSC-BASE website. Experiments grandparents and population are numbers 0009 and 036 respectively.

5.5 Results and discussion

Contrasting physiological mechanisms of drought response in *P. deltoides* and *P. trichocarpa*

We tested the hypothesis that the grandparental species would exhibit contrasting drought responses by exposing them to an acute drought and making detailed physiological and gene expression assessments.. *P. deltoides* and *P. trichocarpa* manifested contrasting physiological responses, perhaps reflecting the adaptation of the species to the drought regimes of their natural ranges. In leaves of *P. deltoides,* senescence was initiated (Figure 5.1a), stomata closed rapidly in response to drought with a concomitant reduction in the rate of photosynthesis (Figure 5.1b) and leaf area was significantly reduced at all LPIs tested (Figure 5.1c).

In contrast, leaves of *P. trichocarpa* formed necrotic lesions in semi-mature to mature leaves, followed by leaf shedding with smaller reductions in stomatal conductance and photosynthetic activity, and no reduction in leaf area. At the onset of leaf necrosis, lesions spread rapidly across the leaf (within hours) in contrast to *P. deltoides*, which took days to develop chlorotic leaves (Figure 5.1). It has been argued that leaf shedding is an adaptive drought response, particularly in perennial species, designed to reduce load on the transpiration stream (Rood *et al.* 2000), typically resulting from fatal xylem cavitation in the leaf mid-ribs (Hukin *et al.* 2005) as opposed to the formation of an abscission zone. Perhaps an alternative evolutionary-adaptation argument for such an adaptation could be that xylem cavitation is advantageous in drought stress as it results in branch shedding: branch shedding is of particular importance to poplar as a mechanism of vegetative propagation, with shed branches moving down-stream and eventually growing into a vegetative clone; such young, vegetative growth requires less water for growth than a mature tree. It has been shown that *P. deltoides* is actually more prone to xylem cavitation, with the later onset of cavitation resulting from the ability of the species to rapidly close stomata (Rood *et al.* 2000). Cochard *et al.* (1996) have shown that a hybrid of *P. trichocarpa* and *P. koreana* x *P. trichocarpa* cv. 'Peace' is unable to close stomata in response to drought

or ABA application and Hukin *et al.* (2005) have recently shown that this clone readily develops mid-rib cavitations in response to drought.

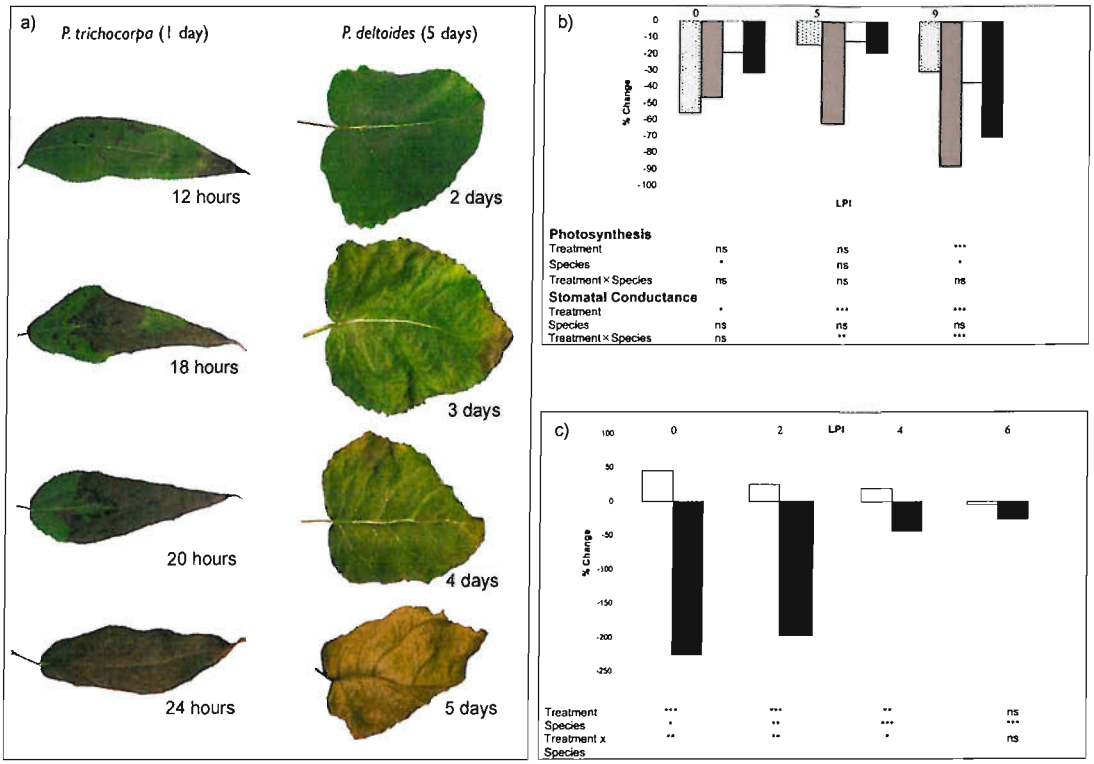


Figure 5.1 Physiological response to drought stress in *Populus trichocarpa* and *P. deltoides* **(a)** Progression of leaf senescence / damage in response to drought for *P. trichocarpa* and *P. deltoides*. The time taken to progress through the stages after the first appearance of any visible symptoms is indicated in parenthesis for each species and the degree of visible damage at various time points is shown. In both species, symptoms had appeared in semi-mature – mature leaves after 14 days of drought stress. **(b)** Percentage effect on the rate of photosynthesis in response to drought for *P. deltoides* (■) and *P. trichocarpa* (□) and percentage effect on stomatal conductance for *P. deltoides* (■) and *P. trichocarpa* (□). **(c)** Percentage effect on leaf area (%) in response to drought at contrasting LPIs for *P. deltoides* (■) and *P. trichocarpa* (□). Results of a two-way ANOVA test are shown for each LPI.

ns not significant, * p>0.05, ** p>0.01, *** p>0.01. In all cases, percentage effect was calculated as ([drought-control]/[control])*100.

The data presented in Figure 5.1(b) reveal a similar, limited stomatal response for *P. trichocarpa* in response to drought and it is therefore likely that this limited ability to close stomata may lead to rapid xylem cavitation resulting in leaf shedding. Whether this is an evolved, adaptive trait (hydraulic segmentation through cavitation; Hukin *et al.* 2005) or is merely an unavoidable consequence of the xylem structure of *Populus* species is not clear. The results shown in Figure 5.1(b) additionally indicate that the lack of stomatal response may be dependant on the developmental age of the leaf, which is consistent with the findings of Ridolfi *et al.* (1996) and Hukin *et al.* (2005). We can conclude therefore that the physiological response to drought for these two *Populus* species differs, providing a good model system to study genetic mechanisms.

Physiological drought traits segregate within the population

Marked transgressive segregation in drought response was observed for all traits recorded in the F_2 population. The traits total chlorophyll, relative leaf expansion rate, percentage of leaves with visible necrosis and abscission, leaf length extension rate, carotenoid content, carotenoid: total chlorophyll, and absolute leaf area expansion were all significantly altered by drought (Figure 5.2a, Table 5.1).

Table 5.1 Summary of extreme genotype trait response in control and drought treatments and in response to drought. Asterisks indicate outcome of two-way ANOVA test: ns not significant, * $p < 0.05$, ** $p < 0.01$. Letter superscripts indicate the results of a post-hoc Tukey test. +Non-normally distributed trait tested using Kruskal-Wallis test. Genotype(extreme) indicates nesting of genotype within extreme group. Percentage effect was calculated as ([drought-control]/[control])*100.

					Control mean		Drought mean		% effect	
	Trt	Extreme	Genotype(extreme)	Trt * Extreme	High	Low	High	Low	High	Low
Height	ns	ns	*	ns	83.25	102.15	87.23	84.64	4.78	-17.14
Abscission	***	ns	**	***	2.67	14.22	36.06	16.24	1249.84	14.22
Chlorosis	ns	ns	ns	-	6.75	3.79	4.15	1.85	-38.42	-51.28
Necrosis	*	ns	ns	-	0	0.4	2.1	5.14	-	1170.11
Chlorophyll a	ns	ns	ns	ns	11.17	16.94	17.08	16.36	52.91	-3.42
Chlorophyll b	ns	ns	ns	-	2.95	13.22	5.79	6.19	95.93	-53.13
Total chlorophyll	ns	ns	ns	ns	14.15	30.19	22.9	22.59	61.9	-25.18
Chorophyll a:b	ns	ns	ns	ns	3.81	4.76	3.79	3.81	-0.65	-20
Carotenoids	*	ns	ns	ns	1.84	2.04	3.01	3.14	63.18	53.7
Carotenoids : Total chlorophyll	**	ns	ns	*	0.12	0.09	0.14	0.16	12.68	63.73
Leaf area 1	ns	*	ns	ns	17.26	12.82	18.08	11.42	4.75	-10.92
Leaf area 2	*	ns	ns	ns	102	107.6	84.23	59.31	-17.42	-44.88
Leaf length 1	*	*	ns	ns	7.21	6.25	6.14	4.53	-14.84	-27.52
Leaf length 2	ns	ns	*	n	13.64	15.07	14.57	12.71	6.82	-15.66
Leaf width 1	**	ns	ns	ns	4.79	3.56	6.18	5.08	29.02	42.7
Leaf width 2	*	ns	ns	ns	12.79	11	9.71	7.49	-24.08	-31.91
Length: Width 1	**	ns	ns	ns	1.84	2.04	1.29	1.18	-29.89	-42.16
Length: Width 2	ns	ns	ns	ns	1.2	1.57	1.64	1.78	36.67	13.38
Relative leaf area expansion rate	**	*	ns	ns	5.21	7.38	3.9	4.3	-25.14	-41.73
Absolute leaf area expansion rate	**	ns	ns	ns	84.7	94.8	66.15	47.89	-21.9	-49.48
Leaf length extension rate	ns	ns	ns	ns	6.43	8.82	8.43	8.2	31.1	-7.03
Leaf width extension rate	**	ns	ns	ns	8	7.44	3.53	2.42	-55.88	-67.47

Significant genotype effects were seen for percentage chlorotic and abscised leaves, leaf length and width extension rate, leaf number, height, and absolute leaf area expansion rate. The percentage of abscised leaves, leaf area expansion, and chlorophyll and carotenoid content were the most highly affected traits in response to drought. These apparent differences prompted us to map the genomic loci (QTL) that were responsible for the variation.

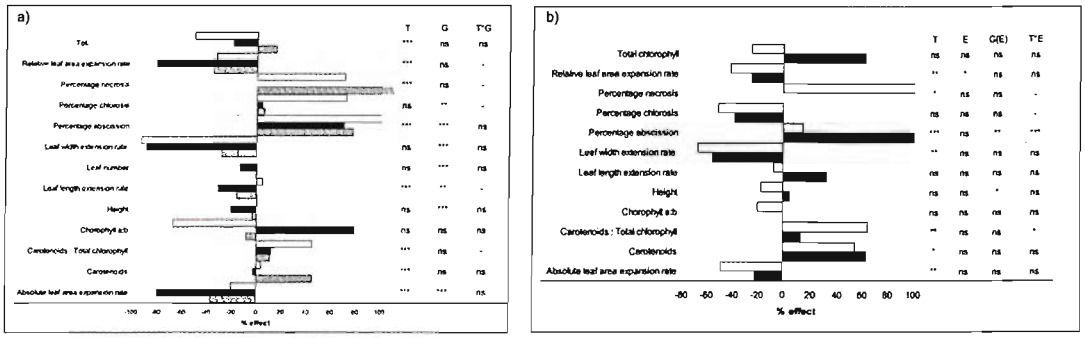


Figure 5.2 (a) Physiological overview of the response to drought of the F_2 *Populus* mapping family **(b)** Physiological overview of the extreme F_2 genotype groups for the abscission response to drought. Values shown are percentage effect. The results of an ANOVA are shown for each trait; * $p<0.05$, ** $p<0.01$, *** $p<0.01$. ANOVA codes are: T - treatment; G - genotype; T*G - treatment by genotype interaction; E - extreme group; G(E) - genotype nested within extreme group. Percentage effect was calculated as ([drought-control]/[control])*100.

QTL were mapped using the out-breeding module of the freely available web-based program QTLExpress (Seaton *et al.* 2002). The genetic linkage map was produced by Tuskan *et al.* (pers. comm.) and consists of 91 SSR markers genotyped on 350 individuals and 92 fully informative AFLPs genotyped on 165 individuals. QTL were mapped from the physiological data in control conditions, drought, and for percentage effect response to drought (these being termed 'response' QTL). In total 25 QTL were mapped in control conditions, 44 in drought, and 30 for percentage effect. Information on all QTL including confidence intervals, maternal and paternal effects, percentage variance explained, and statistical information can be found in Chapter 4, Table 4.2. Six QTL were common to control and drought and 54 were specific to the drought and/or response to drought . It is possible that where a QTL maps only in drought

(with no supporting evidence for the presence of a QTL in the control condition) and/or as a response QTL that this represents the control of a stress-specific response mechanism.

Co-location of QTL specific to the drought treatment or response were identified on LGs IV, VI, VII, VIII, IX, XI, XVIII, and XIX (a subset are shown in Figure 5.3). These all represent regions of the genome that function in controlling drought response with effects on multiple drought response/tolerance related traits. Trait-specific drought/response QTL were also identified for some traits, for example a QTL for height on LG V(I) that explains approximately 8% of the trait variance in both conditions, and these possibly represent the locations of genes controlling the drought/stress response in a trait-specific manner.

Having established that there was a clearly identifiable genetic basis to the segregation of drought response with the population, we examined the drought response within the grandparents and the F_2 genotypes at the transcriptional level to identify differentially expressed genes involved in drought response/tolerance. The use of high throughput mRNA quantifications represents a powerful method by which the differentially expressed gene(s) mapping within a QTL region may be identified. Such an approach may prove to be a useful, and relatively simple, means of breaching the QTL-gene barrier (Hazen and Kay 2003).

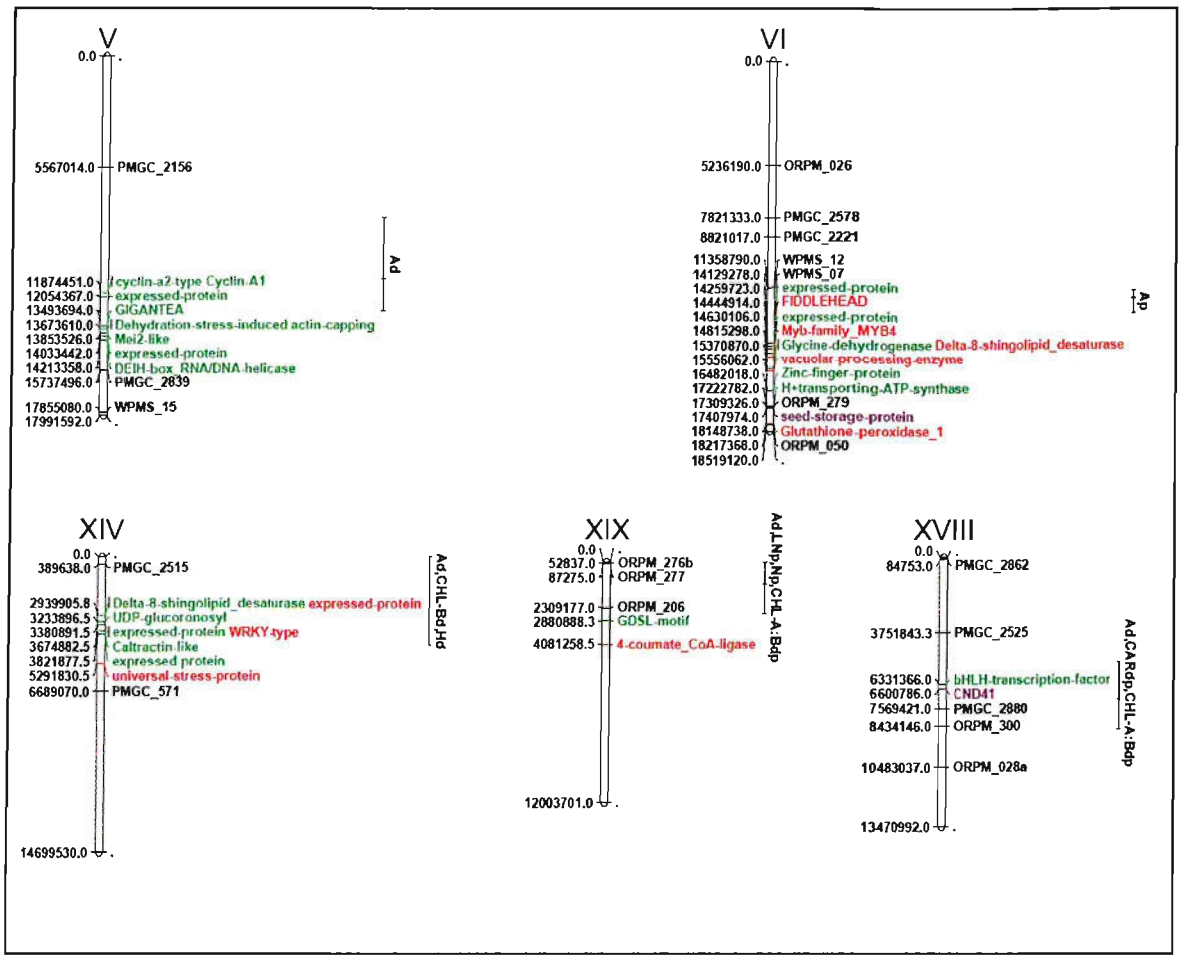


Figure 5.3 The co-location of abscission QTL with genes differentially expressed in response to drought in the grandparental species, *Populus deltoides* and *P. trichocarpa* (red), extreme F_2 genotype groups for the abscission response to drought (green), and both (purple). Each extreme group is composed of the highest or lowest five genotypes for the trait percentage abscission in response to drought. Where QTL for other traits co-locate to abscission QTL, these are shown. Trait names are in CAPS followed by a lowercase letter indicating treatment conditions. Traits are A – abscission, N – Necrosis, CHL-b – chlorophyll b content (µg ml-1), CHL-a:b – chlorophyll a:b, CAR – carotenoid content (µg ml-1), H – height. Conditions are d – drought, p – percentage effect calculated as ([control-drought]/[control])*100. Chromosomes are represented by the open bars with the chromosome number labelled. The values to the left of the bar represent calculated base-pair positions. Genes names and SSR identifiers are shown to the right of the bar. The position of genes and SSRs along a chromosome and of SSRs and QTL along LGs were calculated on a ratio scale and then converted to base-pair positions. Only SSR markers are shown. Confidence Intervals were defined using an F 2 drop off.

Drought-induced remodelling of the grandparental transcriptomes

We exposed the grandparental species to an acute drought stress treatment, sampled leaf material (l_0 – see Experimental procedures) from control and drought stressed plants (after 14 days of drought), prepared RNA from the leaves and analysed the samples on the POP1 microarrays using the experimental design shown in Figure 5.4(a).

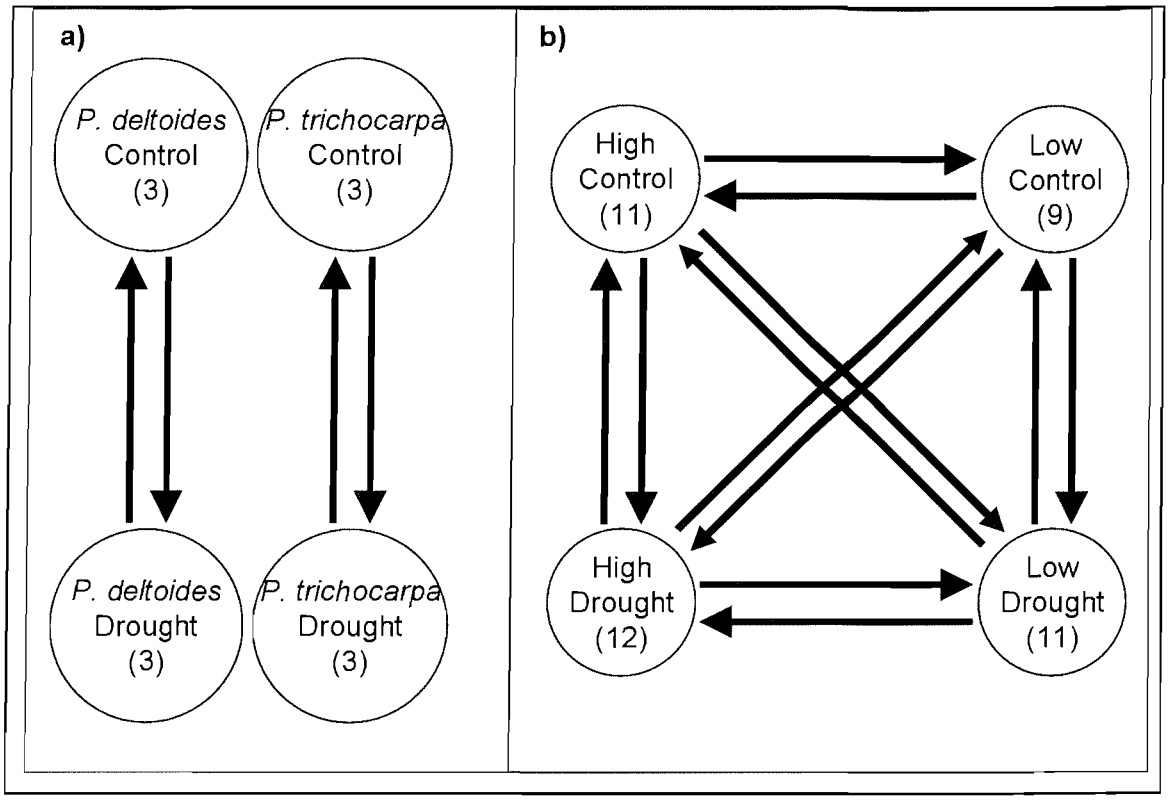


Figure 5.4 Experimental design of microarray experiments. Arrows indicate dye orientations. Replicate arrays conducted with cDNA synthesised from the same RNA extraction but with dyes in the reverse orientation were performed for all comparisons made. (a) Microarray experimental design to examine the transcriptional drought response of *Populus deltoides* and *P. trichocarpa*. The number of biological replicate arrays is indicated in brackets. (b) Microarray experimental design to examine and compare the transcriptional response of extreme F_2 genotype groups for the abscission response in control and drought stress conditions. Each circle represents a pool of cDNA with pools representing the five highest or lowest abscission drought response genotypes sampled in either control or drought conditions. The number of independent biological replicates constituting the pool represented by each circle is shown in brackets.

The results showed that regardless of whether the main physiological differences between the two grandparental species are the result of their contrasting stomatal responses to drought or not, in both species a highly similar set of changes in gene expression was induced in response to drought. As in other species, drought stress

resulted in a profound remodelling of the transcriptome with many genes having more than an 8-fold change in expression (Table S1). To gain a holistic view of changes in gene expression, the functional role of genes that were commonly differentially expressed in both species was examined. Functional classification of genes in PopulusDB (Sterky *et al.* 2004, Moreau *et al.* 2005) is based on Gene Ontology (GO, www.geneontology.org) assignment of mapped gene models from the poplar genome sequence. GO terms are of three types (Molecular function, Biological process and Cellular component). Here we show only data on Biological Process categories as we felt these were the most intuitive for informing biological interpretation. The dataset classified according to Molecular function or Cellular component is available on request. A full list of the 302 differentially expressed genes (229 up-regulated, 73 down-regulated) is available in Table S1 and per-species lists are available in Tables S2 (*P. deltoides*) and S3 (*P. trichocarpa*). The most obvious result was that many genes involved in photosynthesis were down-regulated in response to drought, a commonly observed response to drought stress (Kreps *et al.* 2002, Seki *et al.* 2002). Figure 5.5 shows the GO category hierarchy for down-regulated genes with significant categories shown in green. Non-significant child categories are drawn where necessary to complete the hierarchy. A clear down-regulation of the photosynthetic apparatus can be seen, with this decrease in photosynthetic ability feeding into the observed decrease in metabolism. Although the clear trend within the photosynthesis category was that of down-regulation (the category mean was 1.5 fold down), a small number of photosynthetic genes were up-regulated, including those encoding early light-inducible proteins (ELIPs, which are know to be stress induced; Kreps *et al.* 2002, Bhaleao *et al.* 2003), beta-amylase, sucrose synthase, sucrose-phosphate synthase, catalase 2, pyruvate kinase and the photosystem II 10 kDa polypeptide PsbR. These up-regulated genes were not induced to the same extent in the two species. Those encoding beta-amylase, sucrose synthase, and ELIPs were more highly induced in *P. deltoides* while those encoding sucrose-phosphate synthase, pyruvate kinase and PsbR were more highly induced in *P. trichocarpa*.

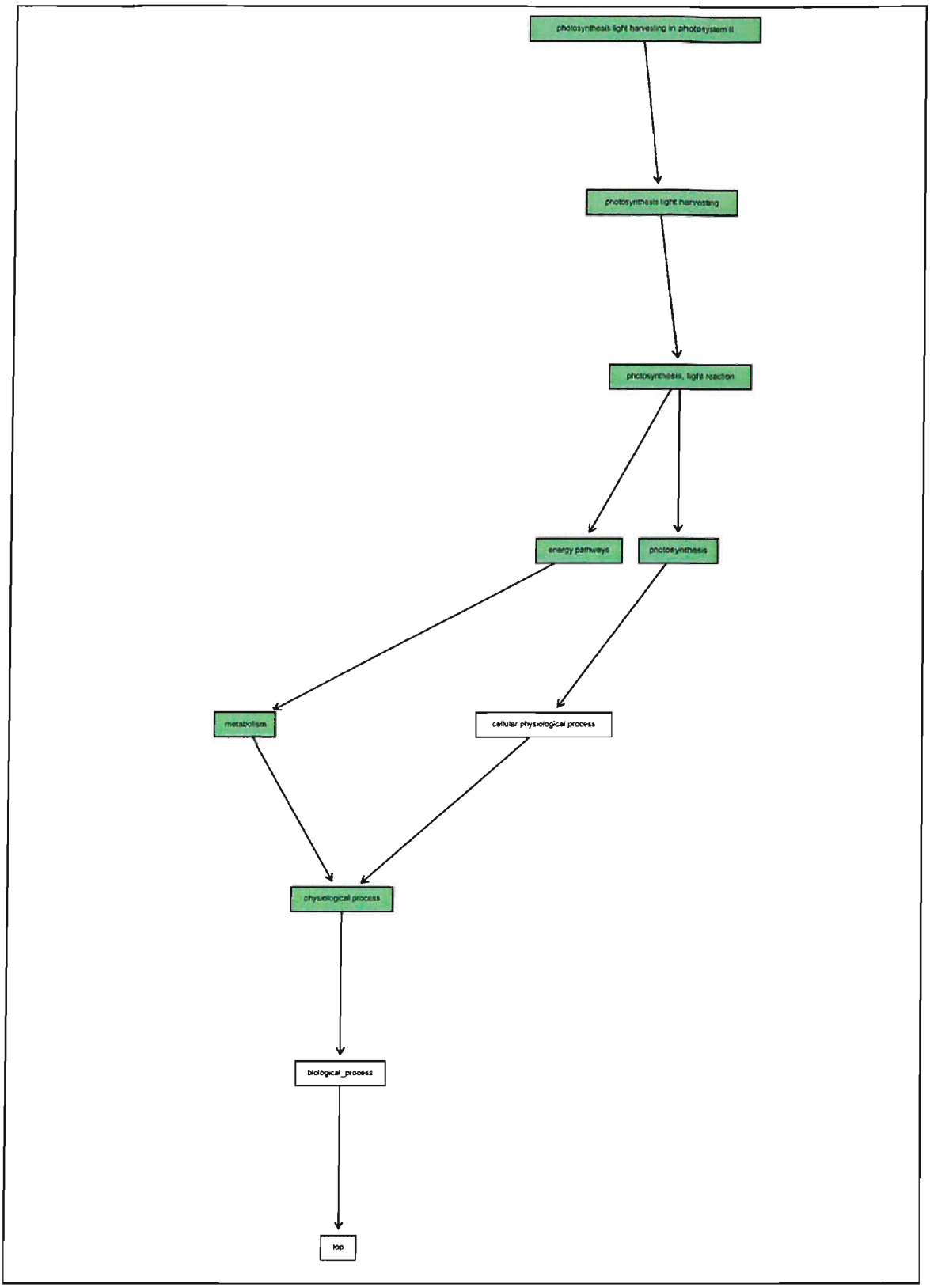


Figure 5.5 Hierarchical view of Gene Ontology categories significantly over-represented by down-regulated genes common to both *Populus deltoides* and *P. trichocarpa* in response to drought. Significant categories are shown in green.

Other GO categories that contained many down-regulated genes were 'biosynthesis' and the child categories 'macromolecule biosynthesis' and 'protein synthesis', indicating that overall protein synthesis was decreased in the drought-stressed leaves. Large up-regulation was seen in the categories 'response to water' and 'response to water deprivation', as well as the more general 'response to stress' category. Up-regulation was also observed in the 'protein ubiquitination' category, suggesting that as well as the down-regulation of protein biosynthesis, there is also an increase in protein degradation / turnover. Categories involved in wax biosynthesis and cuticle biosynthesis were also up-regulated, suggesting that bio-mechanical adaptation was occurring at the level of leaf structure. Alterations to the cuticle and wax layers are often observed in response to drought stress in order to minimise latent water loss through the epidermis (Aharoni *et al.* 2004).

In addition to looking at classes of genes that were regulated in response to drought, individual items from the list of differentially expressed genes (Table S1) were investigated. Additional genes of interest showing expression changes include an ABA binding factors (AREB1), a negative modulator of ABA activity (CBL-9), a ubiquitin-conjugating enzyme, an early response to drought protein (ERD7), and ethylene-insensitive 3 (EIN3), which was recently shown to ABA-induced closure of stomata in response to drought (Tanaka *et al.* 2005), all of which were up-regulated in drought. Many genes observed to be up-regulated in response to drought by previous studies (Bray *et al.* 2004, Seki *et al.* 2002) were also identified here, including LEA genes (Lea 4, 5-D, 14-A), osmoprotectants, chaperones and other stress-related genes (HSPs, glutathione-s-transferase, LTCOR 11), transcription factors (bzips, zinc finger family, and myb family), and alcohol dehydrogenase among others. A number of candidate genes for leaf expansion control including RGA2 (response to gibberellic acid 2) and GAI1 (Gibberellic acid insensitive 1), Phantastica, and a number of cell-cycle genes (cyclin A1, A2, mitotic check point protein, retinoblastoma) were down-regulated during drought stress (See Fleming 2005 for a recent review on leaf development).

In this way, we were able to identify genes in regulons that were differentially expressed during drought stress in both species. Could we get further insight into the other expression characteristics of these regulons? The EST database PopulusDB makes possible rapid digital expression profiling of many *Populus* genes to gain

information on their expression levels in different tissues and after different treatments (Sterky *et al.* 2004). We investigated the library distributions of ESTs from different clusters, corresponding to genes, for the list of 100 genes with most significant differential expression in both species (highest VIP score determined by PLS-DA analysis). When presented as a clustered correlation map according to Ewing *et al.* (1999), it became evident that the most up-regulated genes (Figure 5.6a) have a very different expression pattern than the most down-regulated (Figure 5.6b). This can be expressed numerically as the comparison of the fraction of up-regulated/down-regulated genes found in each library, a Relative Occurrence Factor (ROF, Table 5.2), described in Experimental procedures. If this ratio is 1, the genes found in a library have an equal chance of being in the up- and down-regulated dataset, and ratios over 1 indicate that they are more likely to be found in the up-regulated dataset. The up-regulated regulon(s) are abundantly expressed especially in the dormant cambium (UA), but also in the dormant buds (Q) and senescing leaves (I). By contrast, the genes in these regulons are almost not expressed in cambial tissues (cambial zone, AB, and tension wood, G), in apical meristems (T) and not even in young leaves (C). The genes in the down-regulated regulon(s) are almost all expressed in young leaves (C). Notably, the regulons induced during the stress conditions employed here did not overlap with the regulons induced by cold stress (L).

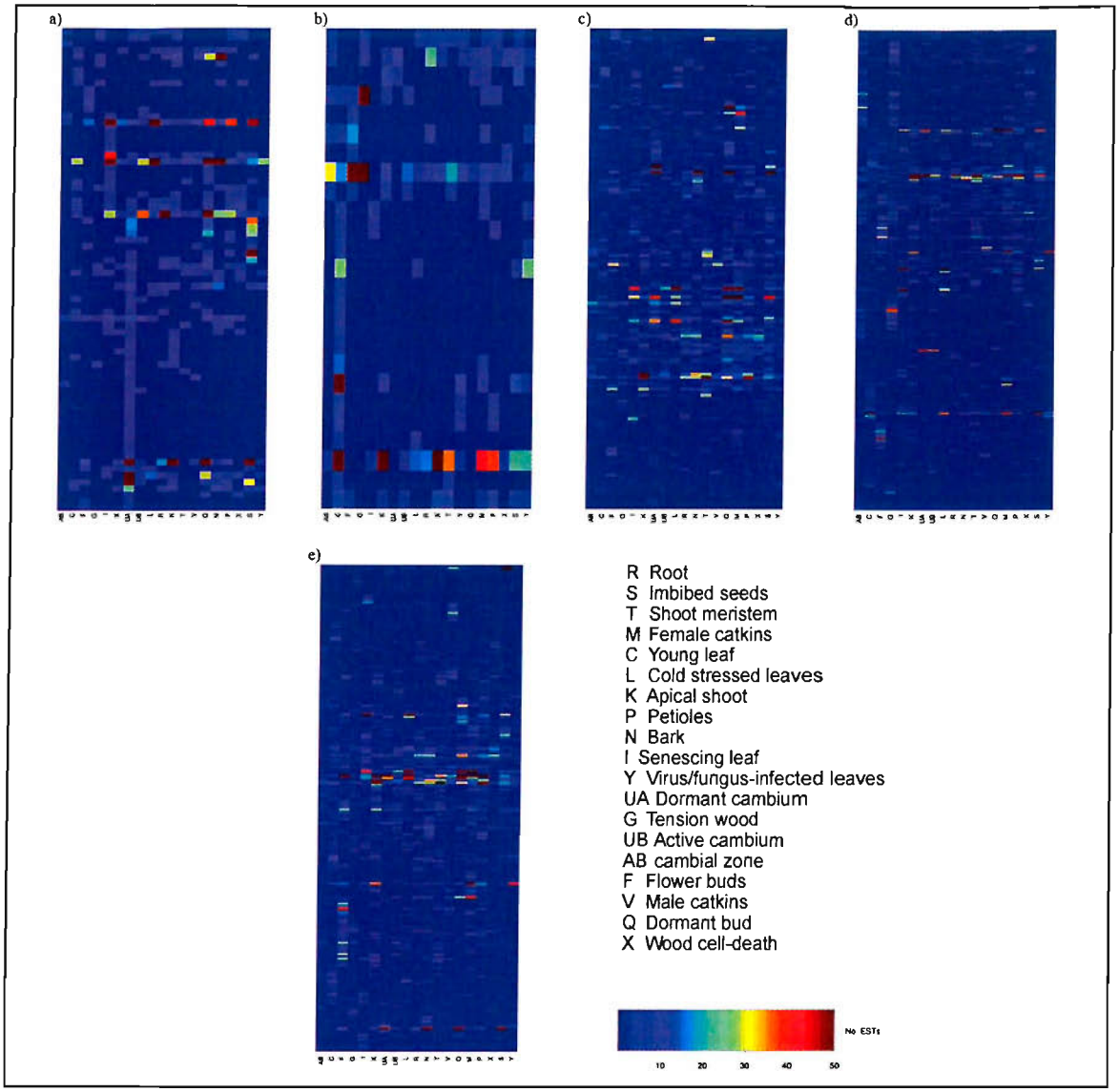


Figure 5.6 The frequency of EST located within each cDNA library in PopulusDB for **(a)** ESTs commonly up-regulated in response to drought in *Populus deltoides* and *P. trichocarpa*, **(b)** ESTs commonly down-regulated in response to drought in *P. deltoides* and *P. trichocarpa*, **(c)** ESTs up-regulated in response to drought in low abscission extreme genotypes, **(d)** ESTs down-regulated in response to drought in low abscission extreme genotypes, **(e)** ESTs that are differentially expressed between high and low extreme genotypes in control and drought conditions.

Table 5.2 Similarities between up-and down-regulated genes during drought and the transcriptome of different *Populus* tissues/treatments. ROF; relative occurrence factor (see material and methods)

Library	Material	ROF
UA	Dormant cambium	9.14
Q	Dormant buds	3.05
I	Senescing leaves	2.81
N	Bark	1.76
X	Wood cell death	1.49
R	Roots	1.34
UB	Active cambium	1.19
L	Cold stressed leaves	1.05
S	Imbibed seeds	0.90
G	Tension wood	0.88
P	Petioles	0.78
F	Flower buds	0.65
V	Male catkins	0.65
K	Apical shoot	0.57
M	Female catkins	0.53
T	Apical meristem	0.53
C	Young leaves	0.33
AB	Cambial zone	0.30

When the digital expression profiles of genes with most significant differential expression between the species after drought stress were analysed separately, additional differences became apparent (data not shown, but available in UPSC-BASE). The regulons must induced in *P. trichocarpa* correlated best with those expressed in senescing leaves, and those down-regulated were typically expressed in the shoot meristem and apical shoot libraries, consistent with the finding that *P. trichocarpa* was more severely affected by the drought, resulting in an rapid decrease in cell division and instead cell death (data not shown, but see Chapter 2).

Taken together, the gross transcriptome responses upon drought stress in *Populus* seems to be biologically adequate and was by large similar to changes reported in earlier studies in other plants.

Species-specific transcriptional response to drought

Rather than simply gaining an overview of the transcriptional response to drought of *Populus* in general, we wanted to identify genes that may contribute to the genetic architectural differences accounting for the contrasting adaptive drought mechanisms of *P. deltoides* and *P. trichocarpa*. This is an analysis that requires a more stringent statistical treatment of the data and an ANOVA was conducted to determine genes differentially expressed between the two species based on ratio-of-expression values. 569 genes that were differentially expressed between the species were identified (Table S4) with 59 of these remaining significant after applying a Bonferroni correction (Table 5.3).

Table 5.3 ESTs that are differentially expressed between *Populus deltoides* and *P. trichocarpa* in response to drought as identified by one-way ANOVA after Bonferroni correction. PU numbers represent gene identifiers for the PopouluDB database (http://www.populus.db.umu.se/). Values for *P. trichocarpa* and *P. deltoides* are M values ($\log_2$ R/G).

PU	*P. deltoides*	*P. trichocarpa*	Decription
PU13206	*6.475598*	*2.198361*	inositol-3-phosphate synthase, putative / myo-inositol-1-phosphate synthase, putative / MI-1-P synthase, putative very strong similarity to SP|Q38862 Myo-inositol-1-phosphate synthase isozyme 2 (EC 5.5.1.4) (MI-1-P synthase 2) (IPS 2) {Arabidopsis thaliana}; identical to SP|Q9LX12| Probable inositol-3-phosphate synthase isozyme 3 (EC 5.5.1.4) (Myo- inositol-1-phosphate synthase 3) (MI-1-P synthase 3) (IPS 3) {Arabidopsis thaliana}; contains Pfam profile PF01658: Myo-inositol-1-phosphate synthase
PU07271	*4.100153*	*1.825729*	calcium-transporting ATPase, plasma membrane-type, putative / Ca2+-ATPase, putative (ACA10) identical to SP|Q9SZR1 Potential calcium-transporting ATPase 10, plasma membrane-type (EC 3.6.3.8) (Ca(2+)-ATPase isoform 10) {Arabidopsis thaliana}; similar to SP|Q9LF79 Calcium-transporting ATPase 8, plasma membrane-type (EC 3.6.3.8) (Ca(2+)-ATPase isoform 8) {Arabidopsis thaliana}
PU10380	*2.601606*	*8.789288*	metallothionein protein, putative (MT2A) identical to Swiss-Prot:P25860 metallothionein-like protein 2A (MT-2A) (MT-K) (MT-1G) [Arabidopsis thaliana]
PU04599	*2.418198*	*7.885675*	Hypothetical 7.2 kDa protein
PU00164	*2.256046*	*1.147243*	vacuolar calcium-binding protein-related
PU09111	*2.208862*	*5.768294*	ATP synthase delta chain, chloroplast, putative / H(+)-transporting two-sector ATPase, delta (OSCP) subunit, putative similar to SP|P32980 ATP synthase delta chain, chloroplast precursor (EC 3.6.3.14) {Nicotiana tabacum}; contains Pfam profile PF00213: ATP synthase F1, delta subunit
PU10934	*2.172932*	*6.569877*	ABC transporter family protein
PU10578	*1.96325*	*7.085015*	ABC transporter family protein
PU05080	*1.796904*	*3.205011*	kinesin light chain-related low similarity to kinesin light chain [Plectonema boryanum] GI:2645229; contains Pfam profile PF00515 TPR Domain
PU05949	*1.771464*	*3.079621*	dihydroorotate dehydrogenase family protein / dihydroorotate oxidase family protein low similarity to SP|Q12882 Dihydropyrimidine dehydrogenase [NADP+] precursor (EC 1.3.1.2) (DPD) (DHPDHase) (Dihydrouracil dehydrogenase) (Dihydrothymine dehydrogenase) {Homo sapiens}; contains Pfam profile PF01180: Dihydroorotate dehydrogenase
PU04104	*1.76984*	*9.142346*	1,4-alpha-glucan branching enzyme / starch branching enzyme class II (SBE2-1) nearly identical to starch branching enzyme class II [Arabidopsis thaliana] GI:619939
PU09561	*1.711779*	*3.286455*	oxidoreductase, 2OG-Fe(II) oxygenase family protein similar to IDS3 [Hordeum vulgare][GI:4514655], leucoanthocyanidin dioxygenase [SP|P51091][Malus domestica]; contains PF03171 2OG-Fe(II) oxygenase superfamily domain
PU00987	*1.660026*	*3.753532*	expressed protein contains Pfam profile PF05129: Putative zinc binding domain (DUF701)

PU	*P. deltoides*	*P. trichocarpa*	Decription
PU12988	*1.580198*	*2.957931*	dihydroorotate dehydrogenase family protein / dihydroorotate oxidase family protein low similarity to SP|Q12882 Dihydropyrimidine dehydrogenase [NADP+] precursor (EC 1.3.1.2) (DPD) (DHPDHase) (Dihydrouracil dehydrogenase) (Dihydrothymine dehydrogenase) {Homo sapiens}; contains Pfam profile PF01180: Dihydroorotate dehydrogenase
PU09412	*1.535174*	*0.327163*	RWP-RK domain-containing protein similar to nodule inception protein [Lotus japonicus] GI:6448579; contains Pfam profile: PF02042 RWP-RK domain
PU02845	*1.497441*	*2.74512*	ABC transporter family protein similar to ABC transporter homolog PnATH GI:7573600 from [Populus nigra]
PU04708	*1.438145*	*0.440084*	invertase/pectin methylesterase inhibitor family protein / DC 1.2 homolog (FL5-2I22) similar to SP|Q42534 Pectinesterase 2 precursor (EC 3.1.1.11) (Pectin methylesterase 2) (PE 2) {Arabidopsis thaliana}; contains Pfam profile PF04043: Plant invertase/pectin methylesterase inhibitor; FL5-2I22 mRNA for DC 1.2 homolog, partial cds GI:11127598
PU12103	*1.419192*	*2.464271*	ATP-dependent Clp protease proteolytic subunit (ClpR1) (nClpP5) identical to nClpP5 GB:BAA82069 GI:5360595 from [Arabidopsis thaliana]; identical to cDNA nClpP5 (nuclear encoded ClpP5) GI:5360594
PU07050	*1.411856*	*0.310094*	expressed protein
PU10024	*1.350112*	*0.465823*	leucine-rich repeat family protein contains leucine rich-repeat (LRR) domains Pfam:PF00560, INTERPRO:IPR001611; similar to Hcr2-0B [Lycopersicon esculentum] gi|3894387|gb|AAC78593
PU12091	*1.337708*	*5.679871*	starch synthase, putative similar to starch synthase SP:Q42857 from [Ipomoea batatas]
PU04424	*1.336257*	*0.168624*	GCN5-related N-acetyltransferase (GNAT) family protein contains Pfam profile PF00583: acetyltransferase, GNAT family
PU13521	*1.335028*	*5.671338*	starch synthase, putative similar to starch synthase SP:Q42857 from [Ipomoea batatas]
PU02560	*1.27536*	*0.209652*	O-methyltransferase family 2 protein similar to caffeic acid O-methyltransferase, Pinus taeda, gb:U39301
PU13236	*1.256461*	*2.771864*	AP2 domain-containing transcription factor, putative contains similarity to AP2 domain transcription factor
PU12538	*1.223985*	*0.278769*	NAD-dependent epimerase/dehydratase family protein similar to sugar epimerase BlmG from Streptomyces verticillus GI:9937230; contains Pfam profile PF01370 NAD dependent epimerase/dehydratase family
PU04760	*1.1844*	*2.537971*	sucrose-phosphatase 1 (SPP1) identical to sucrose-phosphatase (SPP1) [Arabidopsis thaliana] GI:11127757
PU11992	*1.167603*	*2.389933*	expressed protein contains Pfam profile PF05129: Putative zinc binding domain (DUF701)
PU04858	*1.162962*	*3.141314*	pseudo-response regulator 7 (APRR7) identical to pseudo-response regulator 7 GI:10281004 from [Arabidopsis thaliana]
PU12564	*1.143505*	*2.291262*	NADP-dependent oxidoreductase, putative similar to probable NADP-dependent oxidoreductase (zeta-crystallin homolog) P1 [SP|Q39172][gi:886428] and P2 [SP|Q39173][gi:886430], Arabidopsis thaliana; similar to allyl alcohol dehydrogenase GI:9758497 from [Arabidopsis thaliana]
PU00351	*1.116586*	*3.863891*	polyubiquitin (UBQ14) identical to GI:166795; similar to N. sylvestris hexameric polyubiquitin, GenBank accession number M74101
PU07550	*1.105191*	*0.245513*	Beta-galactosidase, putative / lactase, putative similar to beta-galactosidase precursor SP:P48980 from [Lycopersicon esculentum]
PU11800	*1.077015*	*3.639903*	expressed protein contains Pfam profile: PF04601 protein of unknown function (DUF569
PU04754	*1.067533*	*0.394291*	expressed protein contains Pfam profile PF04784: Protein of unknown function, DUF547; expression supported by MPSS

PU	*P. deltoides*	*P. trichocarpa*	Decription
PU04882	*1.066044*	*2.698525*	starch synthase, putative similar to starch synthase SP:Q42857 from [Ipomoea batatas]
PU10422	*1.060048*	*2.589327*	1-aminocyclopropane-1-carboxylate oxidase
PU10214	*1.031735*	*2.233607*	1-aminocyclopropane-1-carboxylate oxidase
PU05957	*1.025986*	*2.597359*	CBL-interacting protein kinase 25 (CIPK25) identical to CBL-interacting protein kinase 25 [Arabidopsis thaliana] gi|17646697|gb|AAL41008
PU12828	*0.953839*	*3.718451*	expressed protein
PU11545	*0.944425*	*3.078609*	expressed protein
PU01231	*0.93581*	*0.331329*	expansin, putative (EXP8) similar to expansin 2 GI:7025493 from [Zinnia elegans]; alpha-expansin gene family, PMID:11641069
PU09275	*0.924274*	*0.493374*	monodehydroascorbate reductase, putative similar to cytosolic monodehydroascorbate reductase GB:BAA77214 [Oryza sativa]
PU03733	*0.922522*	*0.308485*	expressed protein
PU10560	*0.918033*	*2.316057*	expressed protein
PU07219	*0.917932*	*3.605562*	auxin-responsive protein, putative similar to auxin-induced protein AIR12 (GI:11357190) [Arabidopsis thaliana]; similar to stromal cell derived factor receptor 2 (GI:20381292) [Mus musculus]
PU09965	*0.859618*	*0.124811*	expansin-related similar to blight-associated protein p12 precursor [Citrus jambhiri] gi|4102727|gb|AAD03398; similar to beta-expansin [Oryza sativa] gi|8118428|gb|AAF72986; expansin-related gene, PMID:11641069, www.bio.psu.edu/expansins
PU08846	*0.847*	*0.342712*	expressed protein
PU08140	*0.835628*	*2.696852*	auxin-responsive protein, putative similar to auxin-induced protein AIR12 (GI:11357190) [Arabidopsis thaliana]; similar to stromal cell derived factor receptor 2 (GI:20381292) [Mus musculus]
PU09410	*0.817462*	*0.122454*	expansin-related similar to blight-associated protein p12 precursor [Citrus jambhiri] gi|4102727|gb|AAD03398; similar to beta-expansin [Oryza sativa] gi|8118428|gb|AAF72986; expansin-related gene, PMID:11641069, www.bio.psu.edu/expansins
PU03608	*0.810901*	*2.2006*	phosphorylase family protein contains Pfam PF01048: Phosphorylase family
PU08629	*0.810684*	*3.166807*	expressed protein
PU01216	*0.808822*	*0.224556*	Zinc finger (C3HC4-type RING finger) family protein contains Pfam profile: PF00097 zinc finger, C3HC4 type (RING finger)
PU12638	*0.797644*	*0.369748*	DEAD box RNA helicase, putative similar to RNA helicase [Rattus norvegicus] GI:897915; contains Pfam profiles PF00270: DEAD/DEAH box helicase, PF00271: Helicase conserved C-terminal domain
PU01731	*0.741206*	*0.276708*	hypothetical protein
PU08382	*0.719738*	*0.208747*	expansin, putative (EXP8) similar to expansin 2 GI:7025493 from [Zinnia elegans]; alpha-expansin gene family, PMID:11641069
PU03127	*0.708841*	*2.755436*	glutaredoxin, putative similar to glutaredoxin [Ricinus communis] SWISS-PROT:P55143
PU03615	*0.658281*	*2.462611*	expressed protein
PU05730	*0.62813*	*2.633208*	expressed protein
PU12382	*0.496574*	*0.946178*	3'(2'),5'-bisphosphate nucleotidase / inositol polyphosphate 1-phosphatase / FIERY1 protein (FRY1) (SAL1)

This list identified genes in regulons that were differently regulated in the two species - regulons that may be important for the different drought responses in the two species. Among the genes more highly expressed during drought in *P. trichocarpa* were those encoding blight-associated protein P12, a putative disease resistance protein, a ripening-related protein, an expansin (with an apparent leaf-specific expression pattern, deduced from the library distribution of EST clones), a methyl transferase, and an epimerase/dehydratase. Genes with an informative annotation that were more highly induced in *P. deltoides* included those encoding enzymes in starch metabolism, a granule-bound glycogen/starch synthase, a 1-4-alpha-glucan branching enzyme (starch branching enzyme), and proteases (ClpR1 and ubiquitin). In addition, a 1-aminocyclopropane-1-carboxylate oxidase (ACC oxidase) was induced in *P. deltoides* to a considerably greater extent than in *P. trichocarpa*. There are many ACC oxidase genes expressed in *Populus* and data-mining of PopulusDB showed that these genes often had drastically different expression patterns; this particular ACC oxidase gene (PU10422, PU10214 Table 5.3), appears to be almost exclusively expressed in senescing leaves.

Although the commonality in the transcriptional responses of the two species to drought was large, the data presented here indicate that there were a number of genes that differed in expression in response to drought between the two *Populus* species, suggesting that at least a part of their contrasting physiological responses to drought may be genetically controlled through differences in gene regulation. To provide further insight into this, we performed a drought stress experiment using the F_2 offspring of the grandparent clones.

Transcriptional separation of drought response within the population

In order to assess the extent to which gene expression may account for the separation of F_2 genotypes in their drought response, ideally the transcriptome of all individuals in the population should be analysed, during drought stress, using DNA microarrays , similar to the approach Kirst *et al.* (2005) have taken to examine wood-forming tissues. Since this would represent an enormous effort, we tried to construct an experimental design that allowed us to reduce the number of necessary hybridisations.

We identified two extreme groups of genotypes based on the trait 'percentage effect response to drought of the percentage of abscised leaves' and pooled the labelled cDNA obtained from individual replicates leaves from genotypes within each group (Figure 5.4b). Our aim was to investigate whether large-scale differences could be identified in the transcriptional response of the two extreme groups and we felt that a pooled loop design provided the most efficient strategy for testing this, although we are aware of the limitations of such a pooling strategy. This approach can be considered similar to that of performing a bulk segregant analysis (Borevitz *et al.* 2003). Abscission was selected as it is a late onset indicator of drought adaptation response and shows a significant treatment and genotype effect within the F_2 genotypes, with drought-specific QTL being mapped that account for a total of 43% of trait variance. Figure 5.2a shows the percentage effect response to drought for the physiological traits recorded and Figure 5.2b shows the separation of the extreme genotype groups for those traits. Although percentage abscission was the most significant separator of the extreme groups, relative leaf area expansion rate and carotenoid : total chlorophyll content also showed significant differences (Table 5.1). We did not expect all genes with differential response to drought between the grandparents to have a differential expression between the extreme genotype groups as only a sub-fraction of these genes would be expected to segregate in the F_2 population. As the POP2 array had subsequently been developed, containing almost twice as many elements, we were able to assay differential expression of many more genes in this experiment.

RNA was extracted from individual replicates of each genotype and a pool of genotypes within each extreme group was then created for transcriptional analysis. Analysis of this dataset revealed a striking difference in the transcriptional response of the two extreme groups with only 65 genes being commonly differentially expressed in response to drought (55 of these being up-regulated and 10 down-regulated). This proved that our experimental design using a pooling strategy was efficient enough to detect segregating patterns of gene expression, allowing us to gain some insights into the genomic structure of the transcriptional drought response and showing that a major factor distinguishing the two phenotypic extreme groups was differences in gene expression.

In the high abscission group there were 128 genes with a high probability of differential expression in response to drought (Table S5), and in the low abscission group 386 (Table S6). The loop design employed also allowed data to be interrogated from a high-low extreme group perspective in both control and drought conditions. In drought there were 48 genes more highly expressed in the high abscission extreme group and 77 genes with lower-expression (Table S7). In this comparison, a gene with a higher expression level in the high abscission extreme group than in the low will have a positive $\log_{(2)}$ ratio value, and a gene with higher expression in the low abscission extreme group will have a negative $\log_{(2)}$ ratio. In control conditions 96 genes were more highly expressed and 119 showed lower expression (Table S8). These represent constitutive differences in gene expression across the population that could contribute to drought susceptibility and tolerance, although it is possible that some of these genes (or their *trans*-acting regulatory factors) could be genetically linked to genes controlling the drought response.

The functional role of genes based on GO categories (as described above) was examined. Table 5.4 gives GO categories that were significantly over-represented with differentially expressed genes in either the high abscission or low abscission extremes in control and drought conditions. These data allow two questions to be asked: 1) Are there classes of genes that differ in their constitutive expression and could account for differences in drought response? 2) Are there drought responsive gene classes that differ in their degree or direction of expression change that could account for the differential drought response?

In control conditions, high abscission extremes had higher expression levels of genes involved in hormonal response and signalling, including ABA mediated signalling and response, jasmonic acid and ethylene mediated signalling, and categories involved in biotic stress and pathogenesis responses. The expression differences for ABA signalling, biotic stress, and pathogenesis responses remained significant in drought conditions. It is interesting that ethylene signalling was more highly expressed in the high abscission extremes as ethylene plays a key role in controlling the senescence and abscission of leaves, both in response to stress and in autumnal senescence (Buchanan-Wollaston *et al.* 2005, Andersson *et al.* 2004, Lim *et al.* 2003). It is also of note that a number of biotic stress and pathogenesis response categories are significantly more

highly expressed in high abscission lines as this may suggest that these genotypes initiate a response that has a greater overlap with the HR-type response of plants to biotic attack. Indeed, these genotypes regulate genes involved in ROI removal such as glutathione peroxidase and superoxide dismutase, to a lesser extent than low abscission response genotypes, which is associated with a pathogenesis-type response (Mittler 2002) and the control of ROS scavenging is critical in determining the observed response to a stress condition (Mittler *et al.* 2004).

Table 5.4 Gene Ontology (GO) Biological Process categories that were significantly over-represented by genes with expression differences between high and low abscission extreme genotypes in control and drought conditions.

(5.4 i) Categories that have higher expression in high abscission genotypes in control conditions

GO ID	p-value	Description
GO:0009607	0.003893	response to biotic stimulus
GO:0006952	0.006005	defense response
GO:0009755	0.008658	hormone-mediated signalling
GO:0009737	0.013095	response to abscisic acid stimulus
GO:0007165	0.013749	signal transduction
GO:0009725	0.014563	response to hormone stimulus
GO:0009871	0.016874	jasmonic acid and ethylene-dependent systemic resistance, ethylene mediated signaling pathway
GO:0046173	0.016874	polyol biosynthesis
GO:0006114	0.016874	glycerol biosynthesis
GO:0050832	0.025209	defense response to fungi
GO:0015803	0.025209	branched-chain aliphatic amino acid transport
GO:0009814	0.027542	defense response to pathogen, incompatible interaction
GO:0007154	0.028728	cell communication
GO:0009873	0.033476	ethylene mediated signaling pathway
GO:0006662	0.033476	glycerol ether metabolism
GO:0006641	0.033476	triacylglycerol metabolism
GO:0006638	0.033476	neutral lipid metabolism
GO:0006639	0.033476	acylglycerol metabolism
GO:0006071	0.033476	glycerol metabolism
GO:0046486	0.033476	glycerolipid metabolism
GO:0042829	0.034527	defense response to pathogen
GO:0006457	0.036371	protein folding
GO:0006729	0.041675	tetrahydrobiopterin biosynthesis
GO:0009827	0.041675	cell wall modification (sensu Magnoliophyta)
GO:0009816	0.041675	defense response to pathogenic bacteria, incompatible interaction
GO:0042830	0.041675	defense response to pathogenic bacteria
GO:0046146	0.041675	tetrahydrobiopterin metabolism
GO:0009269	0.049808	response to dessication
GO:0009738	0.049808	abscisic acid mediated signalling
GO:0009620	0.049808	response to fungi
GO:0042742	0.049808	defense response to bacteria
GO:0046165	0.049808	alcohol biosynthesis

(5.4 ii) Categories that have higher expression in low abscission genotypes in control conditions

GO ID	p-value	Description
GO:0009699	2.34E-07	phenylpropanoid biosynthesis
GO:0009812	8.59E-07	flavonoid metabolism
GO:0009813	8.59E-07	flavonoid biosynthesis
GO:0042398	1.98E-06	amino acid derivative biosynthesis
GO:0009698	1.98E-06	phenylpropanoid metabolism
GO:0019748	2.58E-06	secondary metabolism
GO:0019438	1.83E-05	aromatic compound biosynthesis
GO:0006575	2.13E-05	amino acid derivative metabolism
GO:0006725	0.000538	aromatic compound metabolism
GO:0019725	0.004902	cell homeostasis
GO:0006873	0.004902	cell ion homeostasis
GO:0050801	0.004902	ion homeostasis
GO:0042592	0.004902	homeostasis
GO:0006519	0.006082	amino acid and derivative metabolism
GO:0030002	0.006354	anion homeostasis
GO:0009714	0.006354	chalcone metabolism
GO:0009715	0.006354	chalcone biosynthesis
GO:0009718	0.006354	anthocyanin biosynthesis
GO:0046688	0.006354	response to copper ion
GO:0030319	0.006354	di-, tri-valent inorganic anion homeostasis
GO:0042180	0.006354	ketone metabolism
GO:0042181	0.006354	ketone biosynthesis
GO:0016098	0.006354	monoterpenoid metabolism
GO:0016099	0.006354	monoterpenoid biosynthesis
GO:0046283	0.006354	anthocyanin metabolism
GO:0030643	0.006354	phosphate ion homeostasis
GO:0009443	0.012669	pyridoxal 5'-phosphate salvage
GO:0008614	0.018947	pyridoxine metabolism
GO:0009828	0.018947	cell wall loosening (sensu Magnoliophyta)
GO:0009831	0.018947	cell wall modification during cell expansion (sensu Magnoliophyta)
GO:0010051	0.018947	vascular tissue pattern formation (sensu Tracheophyta)
GO:0010035	0.018947	response to inorganic substance
GO:0010038	0.018947	response to metal ion
GO:0042816	0.018947	vitamin B6 metabolism
GO:0042547	0.025187	cell wall modification during cell expansion
GO:0009664	0.028255	cell wall organization and biogenesis (sensu Magnoliophyta)
GO:0009827	0.031389	cell wall modification (sensu Magnoliophyta)
GO:0046916	0.031389	transition metal ion homeostasis
GO:0045229	0.043399	external encapsulating structure organization and biogenesis
GO:0007047	0.043399	cell wall organization and biogenesis
GO:0007389	0.043682	pattern specification

(5.4 iii) Categories that have higher expression in high abscission genotypes in drought conditions

GO ID	p-value	Description
GO:0007165	0.000927	signal transduction
GO:0007154	0.002129	cell communication
GO:0009607	0.00383	response to biotic stimulus
GO:0010075	0.004236	regulation of meristem size
GO:0010073	0.008455	meristem maintenance
GO:0006468	0.008985	protein amino acid phosphorylation
GO:0006952	0.014719	defense response
GO:0009934	0.016845	regulation of meristem organization
GO:0016310	0.018705	phosphorylation
GO:0009816	0.021015	defense response to pathogenic bacteria, incompatible interaction
GO:0042830	0.021015	defense response to pathogenic bacteria
GO:0006793	0.0237	phosphorus metabolism
GO:0006796	0.0237	phosphate metabolism
GO:0009738	0.025169	abscisic acid mediated signaling
GO:0040008	0.025169	regulation of growth
GO:0042742	0.025169	defense response to bacteria
GO:0009832	0.033429	cell wall biosynthesis (sensu Magnoliophyta)
GO:0009933	0.033429	meristem organization
GO:0009618	0.037535	response to pathogenic bacteria
GO:0009617	0.045698	response to bacteria

(5.4 iv) Categories that have higher expression in low abscission genotypes in drought conditions

GO ID	p-value	Description
GO:0019725	0.001483	cell homeostasis
GO:0006873	0.001483	cell ion homeostasis
GO:0050801	0.001483	ion homeostasis
GO:0042592	0.001483	homeostasis
GO:0030002	0.00353	anion homeostasis
GO:0019740	0.00353	nitrogen utilization
GO:0006808	0.00353	regulation of nitrogen utilization
GO:0030319	0.00353	di-, tri-valent inorganic anion homeostasis
GO:0030643	0.00353	phosphate ion homeostasis
GO:0009937	0.007048	regulation of gibberellic acid mediated signaling
GO:0009938	0.007048	negative regulation of gibberellic acid mediated signaling
GO:0009740	0.010556	gibberellic acid mediated signaling
GO:0006878	0.010556	copper ion homeostasis
GO:0009739	0.014052	response to gibberellic acid stimulus
GO:0046916	0.017537	transition metal ion homeostasis
GO:0030005	0.021011	di-, tri-valent inorganic cation homeostasis
GO:0006875	0.031367	metal ion homeostasis
GO:0009968	0.031367	negative regulation of signal transduction
GO:0006334	0.031367	nucleosome assembly
GO:0006801	0.034797	superoxide metabolism
GO:0009966	0.034797	regulation of signal transduction
GO:0019430	0.034797	removal of superoxide radicals
GO:0006333	0.038216	chromatin assembly or disassembly
GO:0006869	0.045022	lipid transport

Even more striking differences were revealed when examining genes that were more highly expressed in the low abscission extremes. In control conditions many categories involved in secondary metabolite synthesis including phenylpropanoid biosynthesis, flavonoid biosynthesis, anthocyanin biosynthesis, chalcone biosynthesis, and monoterpenoid biosynthesis showed significantly higher expression in the low abscission extremes. There were also categories involved in cell wall modification, organisation and pattern definition. The apparent emphasis on homeostasis and cellular protection may indicate that resources are being utilised to maintain the cellular integrity and biochemical functionality of leaves. For example, many of these categories could result in an enhanced ability to tolerate increases in reactive oxygen species and other drought-induced biochemical stresses. In drought conditions many categories involved in cellular homeostasis maintenance were represented, as were categories for gibberellic acid response and signalling, suggesting a functional role for this hormone in the drought response of trees that do not readily abscise leaves when exposed to drought stress.

We then investigated the expression patterns of the regulons that were differentially expressed, either in drought or in control conditions, in drought-sensitive vs. drought-resistant clones in the population. As for the common responses to drought, we studied the digital expression profiles in PopulusDB for the gene lists identified above. Within the high abscission extremes there was no strong pattern of library distribution, perhaps as a result of the small number of differentially expressed genes (data not shown). More obvious patterns were seen for the low abscission extremes with the greatest number of up-regulated genes being located within the dormant bud and dormant cambium tissue libraries (Figure 5.6c), perhaps reflecting the adaptive response of these genotypes. A combined list of genes differing in their expression between the high and low extremes in drought and control conditions is shown in Figure 5.6(e) where the greatest number of ESTs were found in the young leaf, apical and shoot meristem libraries as well as the senescing leaf, flower bud and petiole libraries.

It is interesting to attempt to assign adaptive interpretation for the segregation of gene expression observed between the extreme genotypes to their abscission response. For example, the higher expression of meristem genes in high abscission extremes could

correlate to our observations that these genotypes maintain healthy and actively expanding leaves in a narrow zone below the apical meristem. In contrast, the low abscission genotypes initiated a response with similarity to that of dormancy. This may suggest differing adaptive mechanisms for tolerating drought, with low abscission extremes becoming dormant to avoid drought stress, and high abscission extremes adapting their physiology in order to tolerate a stress period; this possibly represents transgressive segregation of grandparental gene expression control and reflects differences in the nature of drought stress that the two species are exposed to in their natural ranges.

Integrating transcriptional and QTL data

We wanted to examine the degree to which differentially expressed genes co-locate to genomic regions identified by QTL analysis. If co-location occurs, this could be due to differences in *cis*-acting elements (promoter sequences) in the differentially expressed genes. This increases the probability that a gene is involved in the control of drought response, although it does not prove a causal link since the expression may be regulated in *trans* by changes in e.g. a transcription factor regulating the drought response. Weigel and Nordborg (2005) discuss the evidence required for forming a causal link between gene and phenotype.

Both the grandparental and extreme genotype datasets provide candidates for explaining segregation of drought response at the transcriptional level. As discussed above, there were considerable differences between the two sets of candidate genes, with a greater number of co-locating genes being identified from the extreme genotype comparison. Perhaps examination of genotypes at the population distribution extremes for the trait of interest provides a more informative set of candidates than those derived from examination of the parental species.

Figure 5.3 shows the location of a selection of QTL. Using the sequence of SSR marker primers as a link between the physical and genetic maps, we estimated the extent to which the genetic map provides coverage of the physical sequence, as indicated by the alignment of LGs and chromosomes. The positional ratios of genes along a chromosome and of QTL along LGs were calculated as a means to examine

co-location. This resulted in 13 genes that differ in their expression between extreme groups and that co-locate to genomic regions identified by QTL mapping.

In the high abscission extreme group, five genes co-locating with QTLs were more highly expressed than in the low abscission extreme group, with two of these genes being expressed in control conditions, two in drought conditions, and one in both conditions. In drought there was a WRKY-type DNA binding protein of unknown function on LG XIV (PU09430), and the FIDDLEHEAD gene on LG VI (PU23162), which is involved in the synthesis of long-chain lipids localised to the cuticle. FIDDLEHEAD is similar to genes encoding proteins related to β-ketoacyl-CoA synthases and chalcone synthases (Pruitt *et al.* 2000). In control there was an expressed protein of unknown function on LG XIV (PU06181) and Glutathione peroxidase 1 on LG VI (PU29386), which is involved in the oxidative stress response. The nucleoid DNA binding protein CND41 located on LG XVIII (PU09068, PU09712, PU09672) was more highly expressed in the high abscission extreme group in both conditions and additionally is present in the list of genes that were differentially expressed between the grandparental species in response to drought. This DNA binding protease has been shown to function in the degradation of denatured Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase) and antisense tobacco cells with reduced expression of CD41 exhibited delayed senescence (Kato *et al*, 2004). CND41 has also been shown to affect gibberellin content with a resultant dwarfism phenotype in tobacco (Nakano *et al.* 2003). Other gibberellin signalling and biosynthesis genes are differentially expressed to differing degrees between the high and low extreme groups (See Table 5.4 and S7-8). Other phenotypic effects identified in the tobacco antisense line (Kato *et al.* 2004) were mirrored in the drought response segregation observed in the extreme abscission genotypes examined.

In the low abscission extreme group, seven genes were more highly expressed; six of these were expressed in control conditions and one in drought. In drought there was a universal stress protein (USP) family protein similar to ER6 on LG XIV (PU08728). In control, on LG VI there was a myb family transcription factor MYB4 that is involved in the control of cinnemate-4-hydroxylase (PU21938; Hemm *et al.* 2001, Jin *et al.* 2000), a delta-8-sphingolipid desaturase of unknown function but with potential function in signal transduction (PU09788), a cysteine proteinase that is expressed in

vegetative tissues during cell death and in response to stress (PU11221; Kinoshita *et al*, 1999). On LG XIX there was a GDSL-motif lipase/hydrolase family protein similar to family II lipase EXL3 (PU27165, PU08538, PU09305) that has no ascribed function but was identified in a screen of genes regulated by APETALA3 and PISTILLATA (Zik and Irish 2003), and 4-coumarate:CoA ligase 3/4-coumaroyl-CoA synthase 3 (4CL3, PU12912), which encodes an isoform of 4-coumate:CoA ligase (4CL) and that is involved in the last step of the general phenylpropanoid pathway (Ehlting *et al.* 1999).

The co-location of genes differentially expressed in response to drought between the grandparental species was also examined. One gene, CND41 as discussed above, was common to the grandparental and extreme experiments. An additional 20 genes are located within the QTL regions discussed above including a seed storage protein, a calcium binding protein, delta-8-sphingolipid desaturase, and GIGANTEA, which was represented by four ESTs.

Conclusions

We have shown that *P. deltoides* and *P. trichocarpa* have contrasting responses to drought, and have used a genetic and genomic approach to study the genetics of this difference. We have shown that the divergent drought response of two *Populus* species exhibits transgressive segregation within an F_2 population and that this results in the emergence of highly contrasting adaptive drought responses. We observed dramatic segregation within the population in the abscission response to drought and the comparison of the transcriptional response of a set of high and low abscission genotypes revealed a striking and perhaps surprising degree of separation.

Although the common transcriptional responses to drought stress provides information of genes and regulons induced by drought stress (i.e. genes typically expressed by dormant tissues), this type of study does not necessarily provide good candidates for genes responsible for natural variation in this trait, since it is quite likely that the different alleles of the genes regulating the drought response may not have particularly contrasting transcript levels. Instead, we believe that the "genetical genomics" approach may be needed to understand the drought response at the level of genetic variation. A number of genes with differential levels of expression response to drought

between the two extreme groups co-located to genomic regions identified by QTL analysis. These genes may provide clues as to the genetic mechanisms through which species adaptation to drought has been achieved. A similar study in rice identified large-scale divergence of the transcriptional response for genotypes with divergent osmotic adjustment responses to drought stress (Hazen *et al.* 2004) and this may suggest that gene expression control plays a key role in the mechanisms of species divergence.

We believe that the results of most general significance from our studies are that: 1) microarrays can be used to identify genes differentially expressed in two closely related species exposed to stress; 2) functional classification and digital expression profiling of the differentially expressed genes can give insights into the different strategies employed by the two species and; 3) a transcriptomics experimental design similar to a bulk segregant analysis can be used to gain further insight into the species-specific responses and provide candidate genes for further testing. For instance, a gene homologous to CND41 that plays a role in the control of senescence in tobacco leaves co-locates to a QTL on LG XVIII that explains 13% of the variation in the abscission response to drought and co-located with QTL for carotenoid content and chlorophyll a:b mapped in both treatment and response conditions. Further studies may tell whether this, or other, genes are responsible for the differences in the architecture of the drought response in these two, and other, *Populus* species.

Chapter 6

General Discussion

6 General discussion

Overview of results

This thesis has examined the genetic nature of *Populus* adaptation to drought stress by utilising variation in the drought response of an F_2 mapping population (Bradshaw and Stettler 1993) from a cross between *Populus trichocarpa* (93-968) and *P. deltoides* Bart (ILL-129) known to be highly divergent for a vast range of phenotypic traits (Rae *et al.* 2004).

Phenotyping, QTL analysis and microarray methods were combined to demonstrate that "genetical genomics" (Janson and Nap 2001) can be used to provide information on adaptation at the species level. The grandparents and F_2 population were subjected to soil drying and contrasting responses to drought across genotypes including for leaf colouration, expansion, and abscission were observed and QTL for these traits were identified. A subset of extreme genotypes exhibiting extreme sensitivity and insensitivity to drought on the basis of leaf abscission were identified and microarray experiments were conducted on these extreme genotype groups and the grandparental species. The extreme genotype groups induced markedly different sets of genes in response to drought: 215 and 125 genes differed in their expression response in control and drought respectively. This result, presented in Chapter 5, suggests that species adaptation may result from changes at the gene expression level. Genes preferentially expressed in drought resistant genotypes overlapped with genes expressed in dormant tissues, whereas genes involved in meristem function had a lower expression. Co-location of differentially expressed genes with drought specific and drought responsive QTLs was identified and these represent candidate genes contributing to the variation in drought response.

The work presented in this thesis provides the first genome-wide examination of the drought response of *Populus*, and indeed of any perennial, woody species. The results presented in Chapter 3 show that there is a large degree of similarity in the remodelling of the transcriptome induced by drought stress in two divergent *Populus* species from different sections of the *Populus* genus. Comparison of the transcriptional changes reveals a high degree of similarity to changes induced in other

model species (Seki *et al.* 2002 a,b, Hazen *et al.* 2005). A number of interesting changes in gene expression were observed including changes in genes involved in the ethylene biosynthesis and signalling pathways. It was also observed that changes in the expression of genes involved in the biosynthesis of, and response to, gibberellic acid appeared to be important constituents of the *Populus* drought response.

It is known that drought tolerance varies considerably between genotypes of *Populus*, both inter- and intra-specifically (Cochard, Ridolfi & Dreyer, 1996; Gebre *et al.,* 1993, 1998; Harvey & Driessche, 1997; Marron *et al.,* 2002; Robison & Raffa, 1998; Tschaplinski *et al.,* 1998). Examination of the two species used throughout this thesis revealed that, at the morphological and physiological level, *P. deltoides* and *P. trichocarpa* have divergent drought responses: *P. deltoides* exhibited a 'classic' drought avoidance response with rapid reductions in stomatal conductance and leaf area expansion and the induction of leaf senescence; in contrast, *P. trichocarpa* showed little response except in young leaves, and mature leaves developed rapidly-spreading necrotic lesions, ultimately resulting in leaf shedding in response to drought. Differences in stomatal response of the two species mirrored those reported by others (Hukin *et al.* 2005, Ridolfi *et al.* 1996). The results shown in Chapter 1 show that these two species employ contrasting mechanisms to expand leaf area and that these mechanisms are differentially affected by exposure to drought stress. Chapter 4 showed that there was also clear segregation for the control of leaf area expansion and leaf shape within the F_2 population making this an ideal genetic model with which to elucidate the constituent components of the control of leaf area development in future research.

Populus as an emerging model system

Populus research has undergone a transformation in the past year: the release of the genome sequence has provided numerous new opportunities to answer questions that were previously time-consuming, or intractable. For example, when the data presented in Chapter 3 were first analysed, no genome sequence was available and it was a task of considerable magnitude attempting to ascertain the cluster structure of ESTs in order to provide a uni-gene set for analysis. Subsequently, the PopulusDB database was updated to include reference to the genome sequence; this provided a uni-gene

EST set as well as new links to bioinformatics systems, such as the Gene Ontology (GO) hierarchy system used for the analysis and interpretation of data in Chapter 5. These developments, and improvements in microarray analysis methods, prompted the re-analysis of the grandparental microarray experiment, the results of which re-analysis are presented in Chapter 5.

The availability of the genome sequence required application of the information provided within the context of the work presented in this thesis. Two of the most important implications arising were the ability to locate genes within chromosomes, and the discovery that *Populus* contains two highly-homologous copies of the majority of genes due to a genome duplication event (pers. comm. Stefan Janson, UPSC, Umeå, Sweden; Gerry Tuskan, ONRL, Oak Ridge, USA). This second fact will be of critical importance when examining and interpreting the genetic architecture of an examined trait or trait response. The genome sequencing project made use of existing genetic maps to anchor sequence scaffolds to chromosomes and this has made forming a link between the genetic map used in this thesis and the genome sequence immediately possible. This ability has been exploited in Chapter 5 to align and orient linkage groups from the genetic map to the genome sequence and to subsequently examine the overlap of QTL within linkage groups to gene locations within chromosomes. The work presented here is the first to make use of this approach in *Populus* and, as such, serves as a proof-of-concept as well as providing valuable insights into the genetic architecture of the *Populus* drought response.

Genetic improvement of *Populus*

Understanding of plant development and response to abiotic and biotic stress has been significantly advanced by the application of genetics approaches such as QTL mapping and, more importantly through the use of microarray technology. Combinations of these approaches allows the selection of specific genes for genetic improvement programmes through the use of marker assisted selection or through transformation systems. These advances provide clear economic and social benefits and implications. However, they also advance the ability to model the response of systems to abiotic stress, biotic attack, and to altered expression of genes at specific times during a developmental pathway. This knowledge serves as a vital addition to

the holistic, systems biology level understanding of plants and, ultimately the emergence of plant biology as a fully predictive science. The addition of *Populus* as a model system (Brunner *et al.* 2004, Taylor 2002, Wullschleger *et al.* 2002, Bradshaw *et al.* 2000) is important as it provides the opportunity to model responses not observed in *Arabidopsis*, such as winter dormancy, the development of secondary meristems and woody tissue development. Poplar also exists as a keystone species in many ecosystems in which it grows and therefore represents a bridge from genomic modelling to that of ecosystems interaction modelling. The ability to select genes of adaptive importance for coping with environmental change will also significantly advance the ability to breed, or design, trees to supply the developing needs of existing and emerging industries, such as that of pulp production and renewable energy production.

The work presented in this thesis not only advances understanding of the genetic architecture of divergent drought response mechanisms employed by *Populus* species, it additionally provides a set of candidate genes for informing the design of a desired response to abiotic stress. The divergent response mechanisms employed by the grandparental species of Family 331 showed marked transgressive segregation within the F_2 population, with the emergence of two distinctly divergent abiotic stress response mechanisms: that of 'escape through dormancy' and that of 'tolerance through adaptation of development'. Previous work has shown that the control of ROS is a key determinant of the stress response initiated (Mittler *et al.* 2004) and the results presented in Chapter 5 suggest that ROS control is an important aspect of the divergent *Populus* drought response mechanisms observed in the extreme genotype groups examined. This suggests that trees with a desired response, designed to match the nature of the stress to which the tree will be exposed (i.e. magnitude, duration), could be produced through manipulation of genes involved in control of ROS responses.

The work presented in this thesis serves to provide a grounding for directing species improvement projects: It has examined the end-result of two divergent stress response mechanisms and the segregation of aspects of those responses. However, numerous challenges remain before this information can be advanced to the point of directed genome improvement; the results of the microarray experiments presented in Chapters

3 and 5 have shown that a very large number of genes are involved in controlling the divergent response mechanisms observed and it would be impractical to selectively alter the expression of such a large number of genes. Instead, it will be essential to use this information to identify transcription factors that are controlling these induced changes in gene transcription – therefore requiring only selective expression changes in that smaller number of transcription factors that are acting as 'master control' switches. The QTL mapping results presented in Chapter 4 may serve the purpose of locating some of these transcription factors. The challenge would now be to fine-map these regions to identify those transcription factors. This can be achieved through the mapping of eQTL for a subset of genes shown to have divergent drought responses within the F_2 population in the hope that some will identify co-locating QTL; such a result would likely indicate the identification of a commonly-acting transcription factor. Where such a QTL additionally co-locates to a trait QTL there would be sufficient evidence to warrant functional characterisation of that gene through genetic manipulation in the form of under- or over-expression. Many previous attempts have been made to manipulate stress response without working backwards from an understanding of the factors controlling differing response mechanisms and these attempts have commonly failed to achieve their goals. Taking a combined forward and reverse genetics approach may prove more fruitful.

Marker assisted selection as a means of gene introduction

Marker assisted selection (MAS) is an attractive method for the genetic improvement of species as it avoids the social problems associated with transformation and genetic modification and successful introduction of genomic regions can be tested for during the seedling stage; this is clearly advantageous compared to the alternative of growing plants to maturity before phenotype data can be collected for analysis (Bradshaw 1998). Despite the clear potential of MAS and its successful demonstration in agricultural crops (Toojinda *et al.* 1998), there have so far been no demonstrations in trees (Taylor 2002). The fact that most agronomically important traits are quantitative increases the challenge of selection: it is relatively easy to incorporate a single gene into a recipient region of the genome; however, quantitative traits, controlled by

several genes, prove more difficult. MAS has always been a proposed use for the results of QTL location (Ribaut and Hoisington 1998).

Genetic improvement through transformation

Transformation is a desirable tool as genes can be introduced into a genome without influencing the expression of a large number of pre-existing genes. In contrast, traditional breeding results in an uncontrolled re-shuffle of the genome, and is restricted to genes already existent in the genotype (Bradshaw and Strauss 2001). Gene transfer technology has dramatically reduced the time required to improve genotypes (Rishi *et al.* 2001) and can be achieved either through the introduction of novel genes from other organisms or through the manipulation of existing gene expression patterns. Candidate genes identified from studies in crops species can therefore be introduced into transformed poplars, and the effects of the introduction examined. Poplar can be transformed either through direct gene transfer or by use of *Agrobacterium*-mediated gene transfer. There have now been a number of successful demonstrations of genetic improvement of poplar, including stress tolerance (Jing *et al.* 2004), insect and herbicide resistance, phytoremediation ability, and lowering of lignin content (Rishi *et al.* 2001). Transgenic poplars have been produced with reduced lignin concentrations and increased cellulose production, and have been tested in field trials, with transformants out-performing the controls (Dinus 2000). Two genes with independent effects have been co-transferred into *Populus tremuloides* through the use of an *Agrobacterium* vector, resulting in the simultaneous reduction of lignin content by up to 52 % and increased cellulose content of up to 30 % (Li *et al.* 2003a). Public concern exists that transgenic trees may result in contamination of the natural gene-pool of wild tree species. Such concerns can be alleviated through the incorporation of sterility into transgenic trees. Sterile, triploid hybrids have been found in the Family 331 pedigree and are not uncommon in *P. trichocarpa* x *P. deltoides* hybrids. Such sterile hybrids could be used as a starting point for genetic modification, since transgenes would not be released into the environment in pollen or ovules (Bradshaw and Stettler 1993), although characterisation of these triploid genotypes is currently limited. It is also possible that poplars modified for accelerated flowering could accelerate the process of pedigree development and MAS (Brunner *et al,* 2002; Rishi *et al.* 2001).

Approaches to understand genetic architecture

Understanding of the suitability of approaches such as that presented in Chapter 5 will become important in the coming years as researchers increasingly make use of linkage disequilibrium mapping, which requires an informed set of candidate genes to be available: The co-location of differentially expressed genes within QTL regions provides strong support for the involvement of those genes in the trait response being examined, although it does not prove a causal link (Weigel and Nordberg 2005). As such, the method developed in Chapter 5 can be used as a screening tool to identify candidate genes for subsequent fine-mapping within a linkage disequilibrium population. Two such populations are currently available: one is a collection of native European *Populus nigra*, and the second a collection of *Populus trichocarpa*, which is held at UBC, Canada by Quentin Cronk (pers. comm. University British Columbia, Vancouver, Canada). An alternative approach is that presented by Kirst *et al.* (2004, 2005), where the expression of genes within all genotypes of a mapping population is examined using microarrays. This approach was proposed by Jansen and Nap (2001) but, to date, has not been fully exploited. It was recently shown to be viable in a study of lignin synthesis in *Eucalyptus* (Kirst *et al.,* 2004, 2005), and variations on the approach have been used to examine the control of cell wall composition of *Zea mays* (Hazen and Kay, 2003), and drought response of *Oryza* (Hazen *et al.,* 2004). Other work has examined the co-location of candidate genes with QTL or regions of introgression (Baxter *et al.,* 2005; Silva *et al.*, 2005). The approach has also been applied to asthma susceptibility in mice (Karp *et al.,* 2000), and for ovariole number in *Drosophila* (Wayne and Mcintyre, 2002). While the approach of mapping expression for all genotypes of a population has clear advantages and provides extremely informative insights into the control of a trait or response, it is technically challenging to analyse and is prohibitively expensive when adequate replication is employed: such an approach would require the use of many hundreds of microarrays for the population used in this thesis. It is also questionable whether the approach provides adequate mapping resolution to be considered ultimately worthwhile. The combined approach proposed in this thesis followed by subsequent fine-mapping studies provides a more affordable and approachable alternative.

Gene expression as a mechanism of speciation

The analysis presented in Chapter 5 shows that, although the similarity in the transcriptional drought response of *P. deltoides* and *P. trichocarpa* was large, differences between the species could be identified at the level of gene transcription, suggesting that at least a part of their contrasting physiological responses to drought may be genetically controlled through differences in gene regulation. These differences were shown to be segregating within the F_2 population produced form a cross between the two species. Chapters 4 and 5 examine and present the variation in response to drought in this population and the analysis of the transcriptional response of both the grandparental species and the F_2 genotypes to drought, using a sub-set of extreme genotypes. These datasets provide candidate genes for explaining the genetic architecture of the drought response of *Populus*, and perhaps other, species.

Although the common transcriptional responses to drought stress provides information on genes and regulons induced by drought stress (i.e. genes typically expressed by dormant tissues), this type of study does not necessarily provide good candidates for genes responsible for natural variation in this trait, since it is quite likely that the different alleles of the genes regulating the drought response may not have particularly contrasting transcript levels. Instead, the "genetical genomcs" approach may be needed to understand the drought response at the level of genetic variation. Chapter 5 showed that a number of genes with differential levels of expression response to drought between the two extreme groups identified for leaf abscission in response to drought co-located to genomic regions identified by QTL analysis. These genes may provide clues as to the genetic mechanisms through which species adaptation to drought has been achieved. A similar study in rice identified large-scale divergence of the transcriptional response for genotypes with divergent osmotic adjustment responses to drought stress (Hazen *et al.* 2004) and this may suggest that gene expression control plays a key role in the mechanisms of species divergence.

In summary, this project has characterised the drought response of contrasting *Populus* species at both the physiological and transcriptional levels. The genetic architecture of these responses has been examined within a population produced from a cross between

the two species using both transcriptional and genetic approaches. Distinct patterns of gene expression were found to exist between the extremes of the population distribution for drought response, with a number of these co-locating to regions of the genome identified as functioning in the control of drought response. The identified contrasting responses may represent defined drought survival strategies and the information obtained on drought response at the gene expression and genetic levels provides valuable insight for designing and breeding trees that are more able to cope with future climate conditions.

7 Publications and presentations arising

7.1 Publications arising from the work of this thesis

The Genetics and Genomics of drought response in *Populus* (2005). **Street NR,** Skogström O, Sjödin A, Tucker J, Rodríguez-Acosta M, Nilsson P, Jansson S, Taylor G. *Submitted: The Plant Journal*

Transcriptional and Physiological Analysis Following Acute and Chronic Drought Stress in Two Populus Species from a Three Generation Mapping Pedigree (2003) Street NR, Rodriguez-Acosta M, Skogström O, Sjödin A, Jansson S, Nilsson P, Bradshaw HD, and Taylor G. *Tree Biotechnology Conference*.

7.2 Presentations given on the work of this thesis

Investigating drought stress using the parents of Family 331. May 2003. *Oral presentation* – POPYOMICS project meeting, Italy.

How good is your array? May 2003. *Poster presentation* – Postgraduate symposium, Southampton University.

Transcriptional response to drought. Feb 2004. *Oral presentation* – POPYOMICS project meeting, Belgium.

Tree genes in drought. May 2004. *Oral presentation* – Postgraduate symposium, University of Southampton. *Prize for best talk awarded.*

Response of POP1 to drought stress. Sept 2004. *Oral presentation* – POPYOMICS project meeting, France.

Abiotic stress: Linking microarrays, QTL and ELPs. Oct 2004. *Poster presentation* – 12th New Phytologist Symposium: Functional genomics of environmental adaptation in *Populus*. Gatlinburg, Tennessee, USA.

Response of POP1 to drought and ozone stress: QTL mapping and candidate genes. Feb 2005. *Oral presentation* – POPYOMICS project meeting, Sweden.

7.3 Other publications arising during the writing of this thesis

QTL for growth and development in elevated carbon dioxide in two model plant genera: a novel approach for understanding plant adaptation to climate change? (2005). Rae AM, Graham LE, Street NR, Hughes J, Hanley ME, Tucker J, Taylor G. *Submiited Global Change Biology*

The genetic basis of plant adaptation to elevated CO_2 (2002) G Taylor, PJ Tricker, L Graham, **NR Street**, AM Rae, MJ Tallis, S Jansson. *Oral presentation* – International Botanical Congress, Vienna, Austria.

Identifying G x E interactions influencing growth characteristics and QTL discovery for an F_2 population of hybrid *Populus* (*P. deltoides* x *P. trichocarpa*) grown at three contrasting sites (2005) Rae AM, Pinel MPC, Bastien C, Sabatti M, Street NR, Tucker J, Dixon C, Taylor G. *In preparation* - submission to *TAG*

The transcriptome of *Populus* in elevated CO_2 (2005). Taylor G, **Street NR**, Tricker P.J, Sjödin A, Graham L, Skogström O, Calfapietra C, Scarascia-Mugnozza G, Jansson S. *New Phytologist* 167: 143-154

Morphological and physiological traits influencing biomass productivity in short-rotation coppice poplar (2004) Rae AM, Robinson KM, Street NR, Taylor G. *Canadian Journal of Forest Research* 34(7): 1488-1498

The Potential of Genomics and Genetics in Free Air Carbon Dioxide Enrichment Experiments. (2005) Gail Taylor, Penny J. Tricker, Laura E Graham, Matthew J Tallis, Anne M Rae, Harriet Trewin, **Nathaniel R. Street**. *Book chapter*

All protocols developed are available from my website: http://populus.biol.soton.ac.uk/~nat

8 References

Adams, R., and Bischof, L. (1994) Seeded Region Growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16**: 641-647.

Aharon, R., Shahak, Y., Wininger, S., Bendov, R., Kapulnik, Y., and Galili, G. (2003) Overexpression of a plasma membrane aquaporin in transgenic tobacco improves plant vigor under favorable growth conditions but not under drought or salt stress. *Plant Cell* **15**: 439-447.

Aharoni, A., Dixit, S., Jetter, R., Thoenes, E., van Arkel, G., and Pereira, A. (2004) The SHINE Clade of AP2 Domain Transcription Factors Activates Wax Biosynthesis, Alters Cuticle Properties, and Confers Drought Tolerance when Overexpressed in Arabidopsis. *Plant Cell* **16**: 2463.

Allan, A.C., Fricker, M.D., Ward, J.L., Beale, M.H., and Trewavas, A.J. (1994) Two Transduction Pathways Mediate Rapid Effects of Abscisic Acid in Commelina Guard Cells. *Plant Cell* **6**: 1319-1328.
Amlin, N.M., and Rood, S.B. (2003) Drought stress and recovery of riparian cottonwoods due to water table alteration along Willow Creek, Alberta. *Trees - Stucture and Function* **17**: 351-358.

Anderson, B.E., Ward, J.M., and Schroeder, J.I. (1994) Evidence for an Extracellular Reception Site for Abscisic-Acid in Commelina Guard-Cells. *Plant Physiology.* **104**: 1177-1183.

Andersson, A., Keskitalo, J., Sjodin, A., Bhalerao, R., Sterky, F., Wissel, K., Tandre, K., Aspeborg, H., Moyle, R., Ohmiya, Y., Bhalerao, R., Brunner, A., Gustafsson, P., Karlsson, J., Lundeberg, J., Nilsson, O., Sandberg, G., Strauss, S., Sundberg, B., Uhlen, M., Jansson, S., and Nilsson, P. (2004) A transcriptional timetable of autumn senescence. *Genome Biology* 5: R24.

Arabidopsis Genome Initiative. (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* **408**: 796-815.

Armstrong, F., Leung, J., Grabov, A., Brearley, J., Giraudat, J., and Blatt, M.R. (1995) Sensitivity to Abscisic-Acid of Guard-Cell K+ Channels Is Suppressed by Abi1-1, a Mutant Arabidopsis Gene Encoding a Putative Protein Phosphatase. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **92**: 9520-9524.

Assmann, S.M. (2003) OPEN STOMATA1 opens the door to ABA signaling in Arabidopsis guard cells. *Trends Plant Science.* **8**: 151-153.

Assmann, S.M., and Wang, X.Q. (2001) From milliseconds to millions of years: Guard cells and environmental responses. *Current Opinions Plant Biology* **4**: 421-428.

Avery, G.S. (1933) Structure and development of the tobacco leaf. *American Journal of Botany* **20**: 565-593.

Baggerly, K.A., Coombes, K.R., Hess, K.R., Stivers, D.N., Abruzzo, L.V., and Zhang, W. (2001) Identifying differentially expressed genes in cDNA microarray experiments. *Journal of Computational Biology* **8**: 639-659.

Ball, R.D. (2001) Bayesian Methods for Quantitative Trait Loci Mapping Based on Model Selection: Approximate Analysis Using the Bayesian Information Criterion. *Genetics* **159**: 1351-1364.

Barker, G., Batley, J., O'Sullivan, H., Edwards, K.J., and Edwards, D. (2003) Redundancy based detection of sequence polymorphisms in expressed tag data using autoSNP. *Bioinformatics* **19**: 421-422.

Bartholomew, D.M., Bartley, G.E., and Scolnik, P.A. (1991) Abscisic-Acid Control of Rbcs and Cab Transcription in Tomato Leaves. *Plant Physiology.* **96**: 291-296.

Baxter, C.J., Sabar, M., Quick, W.P., and Sweetlove, L.J. (2005) Comparison of changes in fruit gene expression in tomato introgression lines provides evidence of genome-wide transcriptional changes and reveals links to mapped QTLs and described traits. *Journal of Experimental Botany* **56**: 1591-1604.

Becraft, P.W. (2002) Receptor kinase signaling in plant development. *Annual Review Cell Developmental Biology* **18**: 163-192.

Beucher, S., and Meyer, F. (1993) The morphological approach to segmentation: the watershed transformation. Mathematical morphology in image processing. *Optical Engineering* **34**: 433-481.

Bhalerao, R., Keskitalo, J., Sterky, F., Erlandsson, R., Bjorkbacka, H., Birve, S.J., Karlsson, J., Gardestrom, P., Gustafsson, P., Lundeberg, J., and Jansson, S. (2003) Gene Expression in Autumn Leaves. *Plant Physiology* 131: 430-442.

Blake, T.J., Sperry, J.S., Tschaplinsky, T.J., and Wang, S.S. (1996). Water relations. NRC Research Press, National Research Council of Canada, Ottawa.

Blatt, M.R. (2000) Cellular signalling and volume control in stomatal movements in plants. *Annual Review Cell Developmental Biology* **16**: 221-241.

Blatt, M.R., and Armstrong, F. (1993) K+ Channels of Stomatal Guard-Cells - Abscisic-Acid-Evoked Control of the Outward Rectifier Mediated by Cytoplasmic Ph. *Planta* **191**: 330-341.

Blatt, M.R., and Grabov, A. (1997) Signalling gates in abscisic acid-mediated control of guard cell ion channels. *Physiologia Plantarum* **100**: 481-490.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. (1998) AGO1 defines a novel locus of Arabidopsis controlling leaf development. *Embo Journal* **17**: 170-180.

Bohnert, H.J., Ayoubi, P., Borchert, C., Bressan, R.A., Burnap, R.L., Cushman, J.C., Cushman, M.A., Deyholos, M., Fischer, R., and Galbraith, D.W. (2001) A genomics approach towards salt stress tolerance. *Plant Physiology and Biochemistry* 39: 295-311.

Borevitz, J.O., Liang, D., Plouffe, D., Chang, H.-S., Zhu, T., Weigel, D., Berry, C.C., Winzeler, E., and Chory, J. (2003) Large-Scale Identification of Single-Feature Polymorphisms in Complex Genomes. *Genome Research* 13: 513-523.

Borghetti, M., Cinnirella, S., Magnani, F., and Saracino, A. (1998) Impact of long-term drought on xylem embolism and growth in Pinus halepensis Mill. *Trees-Structure and Function* **12**: 187-195.

Borraccino, G., Dipierro, S., and Arrigoni, O. (1989) Interaction Of Ascorbate Free-Radical Reductase With Sulfhydryl-Reagents. *Phytochemistry* **28**: 715-717.

Botstein, D., White, R.L., Skolnick, M., and Davis, R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**: 314-331.

Boyer, J.S. (1982) Plant Productivity and Environment. *Science* **218**: 443-448.

Bradshaw, H.D. (1998). Case history in genetics of long-lived plants: Molecular approaches to domestication of a fast-gowing forest tree: *Poplus. In* Molecular dissection of complex traits., *A.H. Paterson*, ed. Boca Raton, CRC Press

Bradshaw, H.D., and Stettler, R.F. (1993) Molecular-Genetics of Growth and Development in Populus.1. Triploidy in Hybrid Poplars. *Theoretical and Applied Genetics* **86**: 301-307.

Bradshaw, H.D., and Stettler, R.F. (1995) Molecular-Genetics of Growth and Development in Populus.4. Mapping Qtls with Large Effects on Growth, Form, and Phenology Traits in a Forest Tree. *Genetics* **139**: 963-973.

Bradshaw, H.D., and Strauss, S.H. (2001). Breeding strategies for the domestication of poplar. *In* Poplar culture in North America, *D.I. Dickman, J.G. Isebrands, J.E. Eckenwalder, and J. Richardson*, eds. NRC Research Press, Ottawa, Canada.

Bradshaw, H.D., Ceulemans, R., Davis, J., and Stettler, R. (2000) Emerging model systems in plant biology: Poplar (Populus) as a model forest tree. *J. Plant Growth Regulation* **19**: 306-313.

Bradshaw, H.D., Villar, M., Watson, B.D., Otto, K.G., Stewart, S., and Stettler, R.F. (1994) Molecular-Genetics of Growth and Development in Populus.3. A Genetic-Linkage Map of a Hybrid Poplar Composed of Rflp, Sts, and Rapd Markers. *Theoretical and Applied Genetics* **89**: 167-178.

Brand, U., Grunewald, M., Hobe, M., and Simon, R. (2002) Regulation of CLV3 expression by two homeobox genes in Arabidopsis. *Plant Physiology* **129**: 565-575.

Bray, E.A. (1997) Plant responses to water deficit. *Trends in Plant Science* **2**: 48-54.

Bray, E.A. (2002) Abscisic acid regulation of gene expression during water- deficit stress in the era of the Arabidopsis genome. *Plant Cell Environment* **25**: 153-161.

Bray, E.A. (2004) Genes commonly regulated by water-deficit stress in Arabidopsis thaliana. *Journal Experimental Botany* **55**: 2331-2341.

Bray, E.A., Bailey-Serres, J., and Weretilnyk, E. (2000). Responses to abiotic stress.*In* Responses to abiotic stress., *W. Gruissen, B. Buchanan, and R. Jones*, eds. American Society of Plant physiologists, Rockville, USA. pp: 1158-1249.

Broadmeadow, M. (2002). Climate Change: Impact on UK Forests. Crown Copyright, Norwich, UK.
Broadmeadow, M.S.J., Ray, D., and Samuel, C.J.A. (2005) Climate change and the future for broadleaved tree species in Britain. *Forestry* **78**: 145-161.

Brown, T.A. (2002). Genomes 2. Bios Scientific Publishers Ltd, Oxford, UK.

Brunner, A.M. (2002). Controlling maturation and flowering for forest tree domestication.*In* Transgenic plants: current innovations and future trends., *C.N.J. Stewart*, ed. Horizon Scientific Press, Wymondham, USA.

Brunner, A.M., Busov, V.B., and Strauss, S.H. (2004) Poplar genome sequence: functional genomics in an ecologically dominant plant species. *Trends in Plant Science* **9**: 49-56.

Buchanan-Wollaston, V., Page, T., Harrison, E., Breeze, E., Lim, P.O., Nam, H.G., Lin, J.-F., Wu, S.-H., Swidzinski, J., Ishizaki, K., and Leaver, C.J. (2005) Comparative transcriptome analysis reveals significant differences in gene expression and signalling pathways between developmental and dark/starvation-induced senescence in Arabidopsis. *The Plant Journal* 42: 567-585.

Burbidge, A., Grieve, T.M., Jackson, A., Thompson, A., McCarty, D.R., and Taylor, I.B. (1999) Characterization of the ABA-deficient tomato mutant notabilis and its relationship with maize Vp14. *Plant Journal* **17**: 427-431.

Burk, D.H., Liu, B., Zhong, R.Q., Morrison, W.H., and Ye, Z.H. (2001) A katanin-like protein regulates normal cell wall biosynthesis and cell elongation. *Plant Cell* **13**: 807-827.

Byrne, M.E., Barley, R., Curtis, M., Arroyo, J.M., Dunham, M., Hudson, A., and Martienssen, R.A. (2000) Asymmetric leaves1 mediates leaf patterning and stem cell function in Arabidopsis. *Nature* **408**: 967-971.

Callahan, H.S., Dhanoolal, N., and Ungerer, M.C. (2005) Plasticity genes and plasticity costs: a new approach using an Arabidopsis recombinant inbred population. *New Phytologist* **166**: 129-140.

Cervera, M.T., Storme, V., Ivens, B., Gusmao, J., Liu, B.H., Hostyn, V., Van Slycken, J., Van Montagu, M., and Boerjan, W. (2001) Dense genetic linkage maps of three Populus species (Populus deltoides, P-nigra and P-trichocarpa) based on AFLP and microsatellite markers. *Genetics* **158**: 787-809.

Chang, S., Puryear, J., and Cairney, J. (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* **11**: 113-116.

Chaves, MM., Maroco, J.P., and Pereira, J.S. (2003). Understanding plant responses to drought – from genes to the whole plant. *Functional Plant Biology* **30**: 239-264.

Chen, S.L., Wang, S.S., Altman, A., and Huttermann, A. (1997a) Genotypic variation in drought tolerance of poplar in relation to abscisic acid. *Tree Physiology* **17**: 797-803.

Chen, W., Chang, S., Hudson, M., Kwan, W.-K., Li, J., Estes, B., Knoll, D., Shi, L., and Zhu, T. (2005) Contribution of transcriptional regulation to natural variations in Arabidopsis. *Genome Biology* **6**: R32.

Chen, Y., Dougherty, E.R., and Bittner, M.L. (1997b) Ratio-Based Decisions and the Quantitative Analysis of cDNA Microarray Images. *Journal of Biomedical Optics* **2**: 364-374.

Chen, Z., and Gallie, D.R. (2004) The Ascorbic Acid Redox State Controls Guard Cell Signaling and Stomatal Movement. *Plant Cell* **16**: 1143-1162.

Chernys, J.T., and Zeevaart, J.A.D. (2000) Characterization of the 9-cis-epoxycarotenoid dioxygenase gene family and the regulation of abscisic acid biosynthesis in avocado. *Plant Physiology* **124**: 343-353.

Choi, D.W., and Close, T.J. (2000) A newly identified barley gene, Dhn12 encoding a YSK2 DHN, is located on chromosome 6H and has embryo-specific expression. *Theoretical and Applied Genetics* **100**: 1274-1278.

Choi, H.-i., Hong, J.-h., Ha, J.-o., Kang, J.-y., and Kim, S.Y. (2000) ABFs, a Family of ABA-responsive Element Binding Factors. *Journal of Biological Chemistry* **275**: 1723-1730.

Chuck, G., Lincoln, C., and Hake, S. (1996) KNAT1 induces lobed leaves with ectopic meristems when overexpressed in Arabidopsis. *Plant Cell* **8**: 1277-1289.

Churchill, G.A. (2002) Fundamentals of experimental design for cDNA microarrays. *Nature Genetics* **32 Suppl 2**: 490-495.

Churchill, G.A., and Doerge, R.W. (1994) Empirical Threshold Values for Quantitative Trait Mapping. *Genetics* **138**: 963-971.

Cochard, H., Ridolfi, M., Dreyer, E. (1996) Responses to Water Stress in an ABA-Unresponsive Hybrid Poplar (Populus koreana x trichocarpa cv. Peace). II. Hydraulic Properties and Xylem Embolism. *New Phytologist* **134**: 455-461.

Cohen, P. (1989) The Structure and Regulation of Protein Phosphatases. *Annual Review Biochemistry* **58**: 453-508.

Cornic, G. (1994). Drought stress and high light effects on leaf photosynthesis. Bios Scientific Publishers, Oxford, UK.

Cornic, G., and Fresneau, C. (2002) Photosynthetic carbon reduction and carbon oxidation cycles are the main electron sinks for photosystem II activity during a mild drought. *Annals of Botany* **89**: 887-894.

Cosgrove, D.J., and Hedrich, R. (1991) Stretch-Activated Chloride, Potassium, and Calcium Channels Coexisting in Plasma-Membranes of Guard-Cells of Vicia-Faba L. *Planta* **186**: 143-153.

Critchfield, W.B. (1960) Leaf dimorphism in *Populus trichocarpa*. *American Journal of Botany* **47**: 699-711.

Cui, X., and Churchill, G. (2003) Statistical tests for differential expression in cDNA microarray experiments. *Genome Biology* **4**: 210.

Cui, X., M. Kathleen Kerr, and Churchill, G.A. (2002) Data Transformations for cDNA Microarray Data. *Lab Report Paper* http://web.ntpu.edu.tw/~hwangyt/93fall/studygp93/paper1.pdf

Cyr, R.J. (1994) Microtubules in Plant Morphogenesis - Role of the Cortical Array. *Annual Review of Cell Biology* **10**: 153-180.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.-R. (2005) Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis. *Plant Physiology* **139**: 5-17.

Dayanandan, S., Rajora, O.P., and Bawa, K.S. (1998) Isolation and characterization of microsatellites in trembling aspen (*Populus tremuloides*). *Theoretical and Applied Genetics* **96**: 950-956.

de Groot, B.L., and Grubmuller, H. (2001) Water permeation across biological membranes: Mechanism and dynamics of aquaporin-1 and GlpF. *Science* **294**: 2353-2357.

Demmig-Adams, B., and Adams Iii, W.W. (1996) The role of xanthophyll cycle carotenoids in the protection of photosynthesis. *Trends in Plant Science* **1**: 21.

Devey, M.E., Bell, J.C., Smith, D.N., Neale, D.B., and Moran, G.F. (1996) A genetic linkage map for *Pinus radiata* based on RFLP, RAPD and microsatellite markers. *Theoretical and Applied Genetics* **92**: 673-679.

DeWald, D.B., Torabinejad, J., Jones, C.A., Shope, J.C., Cangelosi, A.R., Thompson, J.E., Prestwich, G.D., and Hama, H. (2001) Rapid accumulation of phosphatidylinositol 4,5-bisphosphate and inositol 1,4,5-trisphosphate correlates with calcium mobilization in salt-stressed Arabidopsis. *Plant Physiology* **126**: 759-769.

Deyholos, M.K., and Galbraith, D.W. (2001) High-density microarrays for gene expression analysis. *Cytometry* **43**: 229-238.

Dinus, R.J., Payne, P., Sewell, N.M., Chiang, V.L., and Tuskan, G.A. (2001) Genetic modification of short rotation popular wood: Properties for ethanol fuel and fiber productions. *Critical Reviews in Plant Science* **20**: 51-69.

Dolmetsch, R.E., Pajvani, U., Fife, K., Spotts, J.M., and Greenberg, M.E. (2001) Signaling to the nucleus by an L-type calcium channel - Calmodulin complex through the MAP kinase pathway. *Science* **294**: 333-339.

Donnelly, P.M., Bonetta, D., Tsukaya, H., Dengler, R.E., and Dengler, N.G. (1999) Cell cycling and cell enlargement in developing leaves of Arabidopsis. *Developmental Biology* **215**: 407-419.

Doyle, J.J., and Doyle, J.L. (1987) DNA isolation from leaves. *Phytochemical Bulletin* **19**: 11-15.

Dubos, C., Le Provost, G., Pot, D., Salin, F., Lalane, C., Madur, D., Frigerio, J.M., and Plomion, C. (2003) Identification and characterization of water-stress-responsive genes in hydroponically grown maritime pine (Pinus pinaster) seedlings. *Tree Physiology* **23**: 169-179.

Dudley, A.M., Aach, J., Steffen, M.A., and Church, G.M. (2002) Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA* **99**: 7554-7559.

Dudoit, S., Yang, Y.H., Callow, M.J., and Speed, T.P. (2002) Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistical Sinica* **12**: 111-139.

Echt, C.S., May-Marquardt, P., Hseih, M., and Zahorchak, R. (1996) Characterization of microsatellite markers in eastern white pine. *Genomes* **39**: 1102-1108.

Eckenwalder, J.E. (1996). Systematics and evolution of *Populus*. *In* Biology of *Populus*., *R.F. Stettler, H.D.J. Bradshaw, P.E. Heilman, and T.M. Hinckley*, eds. NRC Research Press, Ottawa, Canada pp: 7-30.

Ehlting, J., Buttner, D., Wang, Q., Douglas, C.J., Somssich, I.E., and Kombrink, E. (1999) Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. *The Plant Journal* **19**: 9-20.

English, D. (1996) Phosphatidic acid: A lipid messenger involved in intracellular and extracellular signalling. *Cellular Signalling* **8**: 341-347.

Erickson, R.O., and Michelini, F.J. (1957) The Plastochron Index. *American Journal of Botany* **44**: 297-305.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., and Wold, S. (1999). Introduction to Multi- and Megavariate Data Analysis using Projection Methods (PCA & PLS). Umetrics, Umea, Sweden.

Ewing, R.M., Kahla, A.B., Poirot, O., Lopez, F., Audic, S., and Claverie, J.M. (1999) Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Research* **9**: 950-959.

Felle, H.H., Hanstein, S., Steinmeyer, R., and Hedrich, R. (2000) Dynamics of ionic activities in the apoplast of the sub- stomatal cavity of intact Vicia faba leaves during stomatal closure evoked by ABA and darkness. *Plant Journal* **24**: 297-304.

Ferris, R., Long, L., Bunn, S.M., Robinson, K.M., Bradshaw, H.D., Rae, A.M., and Taylor, G. (2002) Leaf stomatal and epidermal cell development: identification of putative quantitative trait loci in relation to elevated carbon dioxide concentration in poplar. *Tree Physiology* **22**: 633-640.

Fiorani, F., Beemster, G.T.S., Bultynck, L., and Lambers, H. (2000) Can Meristematic Activity Determine Variation in Leaf Size and Elongation Rate among Four Poa Species? A Kinematic Study. *Plant Physiology* **124**: 845-856.

Fleming, A.J. (2005) The control of leaf development. *New Phytologist* **166**: 9-20.

Fletcher, L.C., Brand, U., Running, M.P., Simon, R., and Meyerowitz, E.M. (1999) Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot meristems. *Science* **283**: 1911-1914.

Flexas, J., and Medrano, H. (2002) Drought-inhibition of photosynthesis in C-3 plants: Stomatal and non-stomatal limitations revisited. *Annals Botany* **89**: 183-189.

Foyer, C., Noctor, G., and MorotGaudry, J.F. (1997) Oxygen: Friend or foe for plants. *Biofutur* 27-29.

Foyer, C.H., and Halliwell, B. (1977) Presence Of Glutathione And Glutathione Reductase In Chloroplasts - Proposed Role In Ascorbic-Acid Metabolism. *Planta* **133**: 21-25.

Frank, W., Munnik, T., Kerkmann, K., Salamini, F., and Bartels, D. (2000) Water deficit triggers phospholipase D activity in the resurrection plant Craterostigma plantagineum. *Plant Cell* **12**: 111-123.

Frewen, B.E., Chen, T.H.H., Howe, G.T., Davis, J., Rohde, A., Boerjan, W., and Bradshaw, H.D. (2000) Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. *Genetics* **154**: 837-845.

Gallois, J.L., Woodward, C., Reddy, G.V., and Sablowski, R. (2002) Combined SHOOT MERISTEMLESS and WUSCHEL trigger ectopic organogenesis in Arabidopsis. *Development* **129**: 3207-3217.

Gebre, G.M., and Kuhns, M.R. (1993) Effects of Water-Stress Preconditioning on Gas-Exchange and Water Relations of Populus-Deltoides Clones. *Canadian Journal Forestry Research* **23**: 1291-1297.

Gebre, G.M., Tschaplinski, T.J., Tuskan, G.A., and Todd, D.E. (1998) Clonal and seasonal differences in leaf osmotic potential and organic solutes of five hybrid poplar clones grown under field conditions. *Tree Physiology* **18**: 645-652.

Gentleman, R., Carey, V., Bates, D., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J., and Zhang, J. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* 5: R80.

Gilbert, D.R., Schroeder, M., and van Helden, J. (2000) Interactive visualization and exploration of relationships between biological objects. *Trends in Biotechnology* **18**: 487-494.

Gilbert, S.F. (1997). Developmental Biology. Sinauer Associates Inc., Massachusetts, USA.

Gilmour, S.J., Sebolt, A.M., Salazar, M.P., Everard, J.D., and Thomashow, M.F. (2000) Overexpression of the Arabidopsis CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiology* **124**: 1854-1865.

Gilmour, S.J., Zarka, D.G., Stockinger, E.J., Salazar, M.P., Houghton, J.M., and Thomashow, M.F. (1998) Low temperature regulation of the Arabidopsis CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *Plant Journal* **16**: 433-442.

Gilroy, S., Read, N.D., and Trewavas, A.J. (1990) Elevation of Cytoplasmic Calcium by Caged Calcium or Caged Inositol Trisphosphate Initiates Stomatal Closure. *Nature* **346**: 769-771.

Gimenez, C., Mitchell, V.J., and Lawlor, D.W. (1992) Regulation of photsynthesic rate of 2 sunflower hybrids under water stress. *Plant Physiology* **98**: 516-524.

Gosselin, I., Zhou, Y., Bousquet, J., and Isabel, N. (2002) Megagametophyte-derived linkage maps of white spruce (*Picea glauca*) based on RAPD, SCAR and ESTP markers. *Theoretical and Applied Genetics* **104**: 987-997.

Gowing, D.J., Davies, W.J., and Jones, H.G. (1990) A positive root-sourced signal as an indicator of soil drying in apple, *Malus* x *domestica* Borkh. *Journal Experimental Botany* **41**: 1535-1540.

Grabov, A., Leung, J., Giraudat, J., and Blatt, M.R. (1997) Alteration of anion channel kinetics in wild-type and abi1-1 transgenic Nicotiana benthamiana guard cells by abscisic acid. *Plant Journal* **12**: 203-213.

Graham, L.E., Graham, J.M., and Wilcox, L.W. (2002). Plant Biology. Prentice Hall, UK.

Granger, C.L., Callos, J.D., and Medford, J.I. (1996) Isolation of an Arabidopsis homologue of the maize homeobox Knotted-1 gene. *Plant Molecular Biology* **31**: 373-378.

Grattapaglia, D., Bertolucci, F.L., and Sederoff, R.R. (1995) Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross strategy and RAPD markers. *Theoretical and Applied Genetics* **90**: 933-947.

Green, P.J. (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model dtermination. *Biometrika* **82**: 711-732.

Groover, A., Devey, M., Fiddler, T., Lee, J., Megraw, R., Mitchelolds, T., Sherman, B., Vujcic, S., Williams, C., and Neale, D. (1994) Identification of Quantitative Trait Loci Influencing Wood Specific-Gravity in an Outbred Pedigree of Loblolly-Pine. *Genetics* **138**: 1293-1300.

Gut, I.G. (2001) Automation in genotyping of single nucleotide polymorphisms. *Human Mutations* **17**: 475-492.

Haake, V., Cook, D., Riechmann, J.L., Pineda, O., Thomashow, M.F., and Zhang, J.Z. (2002) Transcription Factor CBF4 Is a Regulator of Drought Adaptation in Arabidopsis. *Plant Physiology* **130**: 639-648.

Haley, C.S., and Knott, S.A. (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**: 315-324.

Hamada, H., Petrino, G.M., and Kakunaga, T. (1982) A novel repeated element with Z-DNA-forming potential is widely found in evolutionary diverse eukaryotic genomes. *Proceedings of the National Academy of Sciences of the United States of America* **79**: 6465-6469.

Hamilton, D.W.A., Hills, A., Kohler, B., and Blatt, M.R. (2000) Ca2+ channels at the plasma membrane of stomatal guard cells are activated by hyperpolarization and abscisic acid. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 4967-4972.

Hanson, J., Johannesson, H., and Engstrom, P. (2001) Sugar-dependent alterations in cotyledon and leaf development in transgenic plants expressing the HDZhdip gene ATHB13. *Plant Molecular Biology* **45**: 247-262.

Hareven, D., Gutfinger, T., Parnis, A., Eshed, Y., and Lifschitz, E. (1996) The making of a compound leaf: genetic manipulation of leaf architecture in tomato. *Cell* **84**: 735-744.

Harmon, A.C., Gribskov, M., Gubrium, E., and Harper, J.F. (2001) The CDPK superfamily of protein kinases. *New Phytologist* **151**: 175-183.

Harvey, H.P., and vandenDriessche, R. (1997) Nutrition, xylem cavitation and drought resistance in hybrid poplar. *Tree Physiology* **17**: 647-654.

Hasegawa, P.M., Bressan, R.A., Zhu, J.K., and Bohnert, H.J. (2000) Plant cellular and molecular responses to high salinity. *Annual Review of Plant Physiology and Plant Molecular Biology* **51**: 463-499.

Hazen, S., Pathan, M., Sanchez, A., Baxter, I., Dunn, M., Estes, B., Chang, H., Zhu, T., Kreps, J., and Nguyen, H. (2004) Expression profiling of rice segregating for drought tolerance QTLs using a rice genome array. *Functional and Integrative Genomics* **5**: 104-116

Hazen, S.P., and Kay, S.A. (2003) Gene arrays are not just for measuring gene expression. *Trends in Plant Science* **8**: 413-416.

Hazen, S.P., Hawley, R.M., Davis, G.L., Henrissat, B., and Walton, J.D. (2003) Quantitative Trait Loci and Comparative Genomics of Cereal Cell Wall Composition. *Plant Physiology* **132**: 263-271.

Heath, S.C. (1997) Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *American Journal of Human Genetics* **61**: 748-760.

Hemm, M.R., Herrmann, K.M., and Chapple, C. (2001) AtMYB4: a transcription factor general in the battle against UV. *Trends in Plant Science* **6**: 135.

Hetherington, A.M. (2001) Guard cell signalling. *Cell* **107**: 711-714.

Hill, A.A., Brown, E.L., Whitley, M.Z., Tucker-Kellogg, G., Hunter, C.P., and Slonim, D.K. (2001) Evaluation of normalization procedures for oligonucleotide array data based on spiked cRNA controls. *Genome Biology* **2**: RESEARCH0055.

Hirayama, T., Ohto, C., Mizoguchi, T., and Shinozaki, K. (1995) A Gene Encoding a Phosphatidylinositol-Specific Phospholipase-C Is Induced by Dehydration and Salt Stress in

Arabidopsis- Thaliana. *Proceedings of the National Academy of Sciences of the United States of America* **92**: 3903-3907.

Hoeschele, I., Uimari, P., Grignola, F.E., Zhang, Q., and Gage, K.M. (1997) Advances in statistical methods to map quantitative trait loci in outbred populations. *Genetics* **147**: 1445-1457.

Hsiao, T.C. (1973) Plant responses to water stress. *Annual Review of Plant Physiology* **24**: 519-570.

Hughes, T.R., Marton, M.J., Jones, A.R., Roberts, C.J., Stoughton, R., Armour, C.D., Bennett, H.A., Coffey, E., Dai, H., and He, Y.D. (2000) Functional Discovery via a Compendium of Expression Profiles. *Cell* 102: 109-126.

Hukin, D., Cochard, H., Dreyer, E., Thiec, D.L., and Bogeat-Triboulot, M.B. (2005) Cavitation vulnerability in roots and shoots: does Populus euphratica Oliv., a poplar from arid areas of Central Asia, differ from other poplar species? *Journal of Experimental Botany* **56**: 2003-2010

Hunt, L., Mills, L.N., Pical, C., Leckie, C.P., Aitken, F.L., Kopka, J., Mueller-Roeber, B., McAinsh, M.R., Hetherington, A.M., and Gray, J.E. (2003) Phospholipase C is required for the control of stomatal aperture by ABA. *Plant Journal* 34: 47-55.

Hurlbert, S.H. (1984) Pseudoreplication and the Design of Ecological Field Experiments. *Ecological Monographs* **54**: 187-211.

Hwang, I., Sze, H., and Harper, J.F. (2000) A calcium-dependent protein kinase can inhibit a calmodulin- stimulated Ca2+ pump (ACA2) located in the endoplasmic reticulum of Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 6224-6229.

Hwang, J.U., Suh, S., Yi, H.J., Kim, J., and Lee, Y. (1997) Actin filaments modulate both stomatal opening and inward K+- channel activities in guard cells of Vicia faba L. *Plant Physiology* **115**: 335-342.

Ihaka, R., and Gentleman, R. (1996) R: A Language for Data Analysis and Graphics. *Journal Computational and Graphical Statistics* **5**: 299-314.

IPCC. (2001). Climate change 2001: the scientific basis. Summary for policymakers and technical summary. Cambridge University Press, Cambridge, UK.

Irving, H.R., Gehring, C.A., and Parish, R.W. (1992) Changes in Cytosolic Ph and Calcium of Guard-Cells Precede Stomatal Movements. *Proceedings of the National Academy of Sciences of the United States of America* **89**: 1790-1794.

Ishitani, M., Xiong, L.M., Stevenson, B., and Zhu, J.K. (1997) Genetic analysis of osmotic and cold stress signal transduction in Arabidopsis: Interactions and convergence of abscisic acid- dependent and abscisic acid-independent pathways. *Plant Cell* **9**: 1935-1949.

Iuchi, S., Kobayashi, M., Taji, T., Naramoto, M., Seki, M., Kato, T., Tabata, S., Kakubari, Y., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2001) Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis. *The Plant Journal* 27: 325-333.

Iwata, H., Ujino-Ihara, T., Yoshimura, K., Nagasaka, K., Mukai, Y., and Tsumura, Y. (2001) Cleaved amplified polymorphic sequence markers in sugi, Cryptomeria japonica D. Don, and their locations on a linkage map. *Theoretical and Applied Genetics* **103**: 881-895.

Jacob, T., Ritchie, S., Assmann, S.M., and Gilroy, S. (1999) Abscisic acid signal transduction in guard cells is mediated by phospholipase D activity. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 12192-12197.

Jaglo-Ottosen, K.R., Gilmour, S.J., Zarka, D.G., Schabenberger, O., and Thomashow, M.F. (1998) Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science* **280**: 104-106.

Jansen, R.C., and Stam, P. (1994) High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136**: 1447-1455.

Jansen, R.C., and Nap, J.P. (2001) Genetical genomics: the added value from segregation. *Trends in Genetics* **17**: 388-391.

Jeong, S., Trotochaud, A.E., and Clark, S.E. (1999) The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell* **11**: 1925-1933.

Jiang, C., and Zieng, Z.-B. (1995) Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* **140**: 1111-1127.

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B., and Martin, C. (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *Embo Journal.* **19**: 6150-6161.

Jing, Z.P., Gallardo, F., Pascual, M.B., Sampalo, R., Romero, J., de Navarra, A.T., and Canovas, F.M. (2004) Improved growth in a field trial of transgenic hybrid poplar overexpressing glutamine synthetase. *New Phytologist* **164**: 137-145.

Kang, J.Y., Choi, H.I., Im, M.Y., and Kim, S.Y. (2002) Arabidopsis basic leucine zipper proteins that mediate stress- responsive abscisic acid signaling. *Plant Cell* **14**: 343-357.

Kao, C.-H., Zeng, Z.-B., and Teasdale, R. (1999) Multiple interval mapping for quantitative trait loci. *Genetics* **152**: 1203-1216.

Karagyozov, L., and Kalcheva, I.D.C., V.M. (1993) Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. *Nucleic Acids Research* **21**: 3911-3912.

Karp, C.L., Grupe, A., Schadt, E., Ewart, S.L., Keane-Moore, M., Cuomo, P.J., Kohl, J., Wahl, L., Kuperman, D., Germer, S., Aud, D., Peltz, G., and Wills-Karp, M. (2000) Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma. *Nature Immunology* 1: 221-226.

Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nature Biotechnology* **17**: 287-291.

Kato, Y., Murakami, S., Yamamoto, Y., Chatani, H., Kondo, Y., Nakano, T., Yokota, A., and Sato, F. (2004) The DNA-binding protease, CND41, and the degradation of ribulose-1,5-bisphosphate carboxylase/oxygenase in senescent leaves of tobacco. *Planta* **220**: 97-104.

Kawaguchi, R., Girke, T., Bray, E.A., and Bailey-Serres, J. (2004) Differential mRNA translation contributes to gene regulation under non-stress and dehydration stress conditions in *Arabidopsis thaliana. The Plant Journal* **38**: 823-839.

Kawasaki, S., Borchert, C., Deyholos, M., Wang, H., Brazille, S., Kawai, K., Galbraith, D., and Bohnert, H.J. (2001) Gene expression profiles during the initial phase of salt stress in rice. *Plant Cell* 13: 889-905.

Kearsey, M.J., and Hyne, V. (1994) QTL analysis: a simple 'marker regression' approach. *Theoretical and Applied Genetics* **89**: 698-702.

Kehoe, D.M., Villand, P., and Somerville, S. (1999) DNA microarrays for studies of higher plants and other photosynthetic organisms. *Trends Plant Science* **4**: 38-41.

Kerr, M.K., and Churchill, G.A. (2001) Statistical design and the analysis of gene expression microarray data. *Genetics Research* 77: 123-128.

Kerr, M.K., Martin, M., and Churchill, G.A. (2000) Analysis of variance for gene expression microarray data. *Journal Computational Biology* 7: 819-837.

Kerstetter, R., Vollbrecht, E., Lowe, B., Veit, B., Yamaguchi, J., and Hake, S. (1994) Sequence-Analysis and Expression Patterns Divide the Maize Knotted1-Like Homeobox Genes into 2 Classes. *Plant Cell* **6**: 1877-1887.

Kerstetter, R.A., Bollman, K., Taylor, R.A., Bomblies, K., and Poethig, R.S. (2001) KANADI regulates organ polarity in Arabidopsis. *Nature* 411: 706-709.

Kim, G.-T., Tsukaya, H., and Uchimiya, H. (1998) The ROTUNDIFOLIA3 gene of Arabidopsis thaliana encodes a new member of the cytochrome P-450 family that is required for the regulated polar elongation of leaf cells. *Genes Dev.* **12**: 2381-2391.

Kim, G.T., Tsukaya, H., Saito, Y., and Uchimiya, H. (1999) Changes in the shapes of leaves and flowers upon overexpression of cytochrome P450 in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 9433-9437.

Kim, J.-B., Kang, J.-Y., and Kim, S.Y. (2004) Over-expression of a transcription factor regulating ABA-responsive gene expression confers multiple stress tolerance. *Plant Biology Journal* **2**: 459-466.

Kinoshita, T., Nishimura, M., and Shimazaki, K.I. (1995) Cytosolic Concentration of Ca2+ Regulates the Plasma-Membrane H+-Atpase in Guard-Cells of Fava-Bean. *Plant Cell* **7**: 1333-1342.

Kinoshita, T., Yamada, K., Hiraiwa, N., Kondo, M., Nishimura, M., and Hara-Nishimura, I. (1999) Vacuolar processing enzyme is up-regulated in the lytic vacuoles of vegetative tissues during senescence and under various stressed conditions. *The Plant Journal* **19**: 43-53.

Kirst, M., Basten, C.J., Myburg, A.A., Zeng, Z.-B., and Sederoff, R.R. (2005) Genetic Architecture of Transcript-Level Variation in Differentiating Xylem of a Eucalyptus Hybrid. *Genetics* **169**: 2295-2303.

Kirst, M., Myburg, A.A., De Leon, J.P.G., Kirst, M.E., Scott, J., and Sederoff, R. (2004) Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of eucalyptus. *Plant Physiology* **135**: 2368-2378.

Kishor, P.B.K., Hong, Z.L., Miao, G.H., Hu, C.A.A., and Verma, D.P.S. (1995) Overexpression of Delta-Pyrroline-5-Carboxylate Synthetase Increases Proline Production and Confers Osmotolerance in Transgenic Plants. *Plant Physiology* **108**: 1387-1394.

Knight, H. (2000). Calcium signaling during abiotic stress in plants. *In* International Review of Cytology - a Survey of Cell Biology, Vol 195, pp: 269-324.

Koorneef, M., Reuling, G., and Karssen, C.M. (1984) The isolation and characterisation of abscisic acid-insensitive mutants of *Arabidopsis thaliana*. *Physiolical Plantarum* **61**: 377-383.

Koornneef, M., Alonso-Blanco, C., and Peeters, A.J.M. (1997) Genetic approaches in plant physiology. *New Phytologist* **137**: 1-8

Koornneef, M., Leon-Kloosterziel, K.M., Schwartz, S.H., and Zeevaart, J.A.D. (1998) The genetic and molecular dissection of abscisic acid biosynthesis and signal transduction in Arabidopsis. *Plant Physiology and Biochemistry* **36**: 83-89.

Kopka, J., Pical, C., Gray, J.E., and Muller-Rober, B. (1998) Molecular and enzymatic characterization of three phosphoinositide-specific phospholipase C isoforms from potato. *Plant Physiology* **116**: 239-250.

Kreps, J.A., Wu, Y., Chang, H.-S., Zhu, T., Wang, X., and Harper, J.F. (2002) Transcriptome Changes for Arabidopsis in Response to Salt, Osmotic, and Cold Stress. *Plant Physiology* **130**: 2129-2141.

Kruglyak, L. (1997) The use of a genetic map of biallelic markers in linkage studies. *Nature Genetics* **17**: 21-24.

Kutschera, U. (1992) The Role of the Epidermis in the Control of Elongation Growth in Stems and Coleoptiles. *Botanica Acta* **105**: 246-252.

Kwak, J.M., Moon, J.-H., Murata, Y., Kuchitsu, K., Leonhardt, N., DeLong, A., and Schroeder, J.I. (2002) Disruption of a Guard Cell-Expressed Protein Phosphatase 2A Regulatory Subunit, RCN1, Confers Abscisic Acid Insensitivity in Arabidopsis. *Plant Cell* **14**: 2849-2861.

Kwok, P.Y., Deng, Q., Zakeri, S.L., Taylor, S.L., and Nickerson, D.A. (1996) Increasing information content of STS-based genome maps: Identifying polymorphisms in mapped STSs. *Genomics* **31**: 123-126.

Lanceras, J.C., Pantuwan, G., Jongdee, B., and Toojinda, T. (2004) Quantitative trait loci associated with drought tolerance at reproductive stage in rice. *Plant Physiology* **135**: 384-399.

Landi, P., Sanguineti, M.C., Salvi, S., Giuliani, S., Bellotti, M., Maccaferri, M., Conti, S., and Tuberosa, R. (2005) Validation and characterization of a major QTL affecting leaf ABA concentration in maize. *Molecular Breeding* **15**: 291-303.

Lang, V., Mantyla, E., Welin, B., Sundberg, B., and Palva, E.T. (1994) Alterations in Water Status, Endogenous Abscisic-Acid Content, and Expression of Rab18 Gene During the Development of Freezing Tolerance in Arabidopsis-Thaliana. *Plant Physiology* **104**: 1341-1349.

Langer, K., Levchenko, V., Fromm, J., Geiger, D., Steinmeyer, R., Lautner, S., Ache, P., and Hedrich, R. (2004) The poplar K+ channel KPT1 is associated with K+ uptake during stomatal opening and bud development. *Plant Journal* **37**: 828-838.

Larson, P.R., and Isebrands, J.G. (1971) The Plastochron Index as Applied to Developmental Studies of Cottonwood. *Canadian Journal of Forest Research* **1**: 1-11.

Lawlor, D.W., and Cornic, G. (2002) Photosynthetic carbon assimilation and associated metabolism in relation to water deficits in higher plants. *Plant Cell Environment* **25**: 275-294.

Leckie, C.P., McAinsh, M.R., Allen, G.J., Sanders, D., and Hetherington, A.M. (1998) Abscisic acid-induced stomatal closure mediated by cyclic ADP- ribose. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 15837-15842.

Lemtirichlieh, F., and Macrobbie, E.A.C. (1994) Role of Calcium in the Modulation of Vicia Guard-Cell Potassium Channels by Abscisic-Acid - a Patch-Clamp Study. *Journal of. Membrane Biology* **137**: 99-107.

Leonhardt, N., Kwak, J.M., Robert, N., Waner, D., Leonhardt, G., and Schroeder, J.I. (2004) Microarray expression analyses of Arabidopsis guard cells and isolation of a recessive abscisic acid hypersensitive protein phosphatase 2C mutant. *Plant Cell* **16**: 596-615.

Leung, J., and Giraudat, J. (1998) Abscisic acid signal transduction. *Annual Review Plant Physiology and Plant Molecular Biology* **49**: 199-222.

Leung, J., Merlot, S., and Giraudat, J. (1997) The Arabidopsis ABSCISIC ACID-INSENSITIVE2 (ABI2) and ABI1 genes encode homologous protein phosphatases 2C involved in abscisic acid signal transduction. *Plant Cell* **9**: 759-771.

Leung, J., Bouvierdurand, M., Morris, P.C., Guerrier, D., Chefdor, F., and Giraudat, J. (1994) Arabidopsis Aba Response Gene Abi1 - Features of a Calcium- Modulated Protein Phosphatase. *Science* **264**: 1448-1452.

Li, J.X., and Assmann, S.M. (1996) An abscisic acid-activated and calcium-independent protein kinase from guard cells of fava bean. *Plant Cell* **8**: 2359-2368.

Li, J.X., Wang, X.Q., Watson, M.B., and Assmann, S.M. (2000) Regulation of abscisic acid-induced stomatal closure and anion channels by guard cell AAPK kinase. *Science* **287**: 300-303.

Li, L., Zhou, Y., Cheng, X., Sun, J., Marita, J.M., Ralph, J., and Chiang, V.L. (2003a) Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **100**: 4939-4944.

Li, W., Luan, S., Schreiber, S.L., and Assmann, S.M. (1994) Cyclic AMP Stimulates K+ Channel Activity in Mesophyll Cells of Vicia faba L. *Plant Physiology* **106**: 957-961.

Li, Y., Jones, L., and McQueen-Mason, S. (2003b) Expansins and cell growth. *Current Opinions in Plant Biology* **6**: 603-610.

Li, Z.C., Mu, P., Li, C.P., Zhang, H.L., Li, Z.K., Gao, Y.M., and Wang, X.K. (2005) QTL mapping of root traits in a doubled haploid population from a cross between upland and lowland japonica rice in three environments. *Theoretical And Applied Genetics* **110**: 1244-1252.

Lian, H.-L., Yu, X., Ye, Q., Ding, X.-S., Kitagawa, Y., Kwak, S.-S., Su, W.-A., and Tang, Z.-C. (2004) The Role of Aquaporin RWC3 in Drought Avoidance in Rice. *Plant Cell Physiology* **45**: 481-489.

Lichtenthaler, H.K. (1996) Vegetation stress: An introduction to the stress concept in plants. *Journal of Plant Physiology* **148**: 4-14.

Lichtenthaler, H.K. (1998). The stress concept in plants: An introduction. *Annals of the New York Academy of Science* **851**: 187-198.

Lim, P.O., Woo, H.R., and Nam, H.G. (2003) Molecular genetics of leaf senescence in Arabidopsis. *Trends in Plant Science* **8**: 272.

Lincoln, C., Long, J., Yamaguchi, J., Serikawa, K., and Hake, S. (1994) A Knotted1-Like Homeobox Gene in Arabidopsis Is Expressed in the Vegetative Meristem and Dramatically Alters Leaf Morphology When Overexpressed in Transgenic Plants. *Plant Cell* **6**: 1859-1876.

Lindroth, A., and Bath, A. (1999) Assessment of regional willow coppice yield in Sweden on basis of water availability. *Forestry and Ecological Management* **121**: 57-65.

Little, C.H.A., and Eidt, D.C. (1968) Effects of abscisic acid on bud break and transpiration in woody species. *Nature* **220**: 498-499.

Liu, H.Y., Zou, G.H., Liu, G.L., Hu, S.P., Li, M.S., Yu, X.Q., Mei, H.W., and Luo, L.J. (2005) Correlation analysis and QTL identification for canopy temperature, leaf water potential and spikelet fertility in rice under contrasting moisture regimes. *China Science Bulletin* **50**: 317-326.

Liu, L., McDonald, A.J.S., Stadenberg, I., and Davies, W.J. (2001a) Abscisic acid in leaves and roots of willow: significance for stomatal conductance. *Tree Physiology* **21**: 759-764.

Liu, L., McDonald, A.J.S., Stadenberg, I., and Davies, W.J. (2001b) Stomatal and leaf growth responses to partial drying of root tips in willow. *Tree Physiology* **21**: 765-770.

Liu, Q., Zhao, N.M., Yamaguch-Shinozaki, K., and Shinozaki, K. (2000) Regulatory role of DREB transcription factors in plant drought, salt and cold tolerance. *China Science Bulletin* **45**: 970-975.

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. *Plant Cell* **10**: 1391-1406.

Loggini, B., Scartazza, A., Brugnoli, E., and Navari-Izzo, F. (1999) Antioxidative defense system, pigment composition, and photosynthetic efficiency in two wheat cultivars subjected to drought. *Plant Physiology* **119**: 1091-1099.

Lonnstedt, I., and Speed, T. (2002) Replicated microarray data. *Statistical Sinica* **12**: 31-46.

Lu, Z., and Neumann, P.M. (1998) Water-stressed maize, barley, and rice seedlings show species diversity in mechanisms of leaf growth inhibition. *Journal of Experimental Botany* **49**: 1945-1952.

Luan, S. (2002) Signalling drought in guard cells. *Plant Cell Environment* **25**: 229-237.

Luan, S., Li, W.W., Rusnak, F., Assmann, S.M., and Schreiber, S.L. (1993) Immunosuppressants Implicate Protein Phosphatase Regulation of K+ Channels in Guard-Cells. *Proceedings of the National Academy of Sciences of the United States of America* **90**: 2202-2206.

Lynch, M., and Walsh, B. (1998). Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA, USA.

Lynn, K., Fernandez, A., Aida, M., Sedbrook, J., Tasaka, M., Masson, P., and Barton, M.K. (1999) The PINHEAD/ZWILLE gene acts pleiotropically in Arabidopsis development and has overlapping functions with the ARGONAUTE1 gene. *Development* **126**: 469-481.

MacRobbie, E.A.C. (1997) Signalling in guard cells and regulation of ion channel activity. *Journal of Experimental Botany* **48**: 515-528.

Maldonado, C.A., Zuñiga, G.E., J., C.L., and Alberdi, M. (1997) Effect of water stress on frost resistance of oat leaves. *Environmental and Experimental Botany* **38**: 99-107.

Mantyla, E., Lang, V., and Palva, E.T. (1995) Role of Abscisic-Acid in Drought-Induced Freezing Tolerance, Cold-Acclimation, and Accumulation of Lt178 and Rab18 Proteins in Arabidopsis-Thaliana. *Plant Physiology* **107**: 141-148.

Marin, E., Nussaume, L., Quesada, A., Gonneau, M., Sotta, B., Hugueney, P., Frey, A., and MarionPoll, A. (1996) Molecular identification of zeaxanthin epoxidase of Nicotiana plumbaginifolia, a gene involved in abscisic acid biosynthesis and corresponding to the ABA locus of Arabidopsis thaliana. *Embo Journal* **15**: 2331-2342.

Marron, N., Delay, D., Petit, J.M., Dreyer, E., Kahlem, G., Delmotte, F.M., and Brignolas, F. (2002) Physiological traits of two Populus x euramericana clones, Luisa Avanzo and Dorskamp, during a water stress and re- watering cycle. *Tree Physiology* **22**: 849-858.

Marth, G.T., Korf, I., Yandell, M.D., Yeh, R.T., Gu, Z.J., Zakeri, H., Stitziel, N.O., Hillier, L., Kwok, P.Y., and Gish, W.R. (1999) A general approach to single-nucleotide polymorphism discovery. *Nature Genetics* 23: 452-456.

Maurel, C. (1997) Aquaporins and water permeability of plant membranes. *Annual Review of Plant Physiology and Plant Molecular Biology* **48**: 399-429.

McConnell, J.R., and Barton, M.K. (1998) Leaf polarity and meristem formation in Arabidopsis. *Development* **125**: 2935-2942.

McConnell, J.R., Emery, J., Eshed, Y., Bao, N., Bowman, J., and Barton, M.K. (2001) Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. *Nature* **411**: 709-713.

McQueen-Mason, S., Durachko, D.M., and Cosgrove, D.J. (1992) Two endogenous proteins that induce cell wall extension in plants. *The Plant Cell* **11**: 1425-1433.

Medina, J., Bargues, M., Terol, J., Perez-Alonso, M., and Salinas, J. (1999) The Arabidopsis CBF gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration. *Plant Physiology* **119**: 463-469.

Medrano, H., Escalona, J.M., Bota, J., Gulias, J., and Flexas, J. (2002) Regulation of Photosynthesis of C3 Plants in Response to Progressive Drought: Stomatal Conductance as a Reference Parameter. *Annals of Botany* **89**: 895-905.

Meyer, K., Leube, M.P., and Grill, E. (1994) A Protein Phosphatase 2c Involved in ABA Signal-Transduction in Arabidopsis-Thaliana. *Science* **264**: 1452-1455.

Miller, A.J. (1990). Subset Selection in Regression. Chapman and Hall, London, UK.

Miller, J.G., and Fry, S.C. (2001) Characteristics of xyloglucan after attack by hydroxyl radicals. *Carbohydrate Research* **332**: 389-403.

Mitchell-Olds, T. (1996) Genetic Constraints on Life-History Evolution: Quantitative-Trait Loci Influencing Growth and Flowering in Arabidopsis thaliana. *Evolution* **50**: 140-145.

Mittelheuser, C.J., and van Steveninck, R.F.M. (1969) Stomatal closure and inhibition of transpiration induced by (RS)-abscisic acid. *Nature* **221**: 281-282.

Mittler, R. (2002) Oxidative stress, antioxidants and stress tolerance. *Trends In Plant Science* **7**: 405-410.

Mittler, R., Vanderauwera, S., Gollery, M., and Van Breusegem, F. (2004) Reactive oxygen gene network of plants. *Trends In Plant Science* **9**: 490-498.

Moran, J.F., Becana, M., Iturbeormaetxe, I., Frechilla, S., Klucas, R.V., and Apariciotejo, P. (1994) Drought Induces Oxidative Stress In Pea-Plants. *Planta* 194: 346-352.

Moreau, C., Aksenov, N., Lorenzo, M., Segerman, B., Funk, C., Nilsson, P., Jansson, S., and Tuominen, H. (2005) A genomic approach to investigate developmental cell death in woody tissues of Populus trees. *Genome Biology* **6**: R34.

Morgante, M., Hanafey, M.K., and Powell, W. (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics* **30**: 194-200.

Mori, I.C., and Muto, S. (1997) Abscisic acid activates a 48-kilodalton protein kinase in guard cell protoplasts. *Plant Physiology* **113**: 833-839.

Morrison, D.A., and Morris, E.C. (2000) Pseudoreplication in experimental designs for the manipulation of seed germination treatments. *Austral Ecology* **25**: 292.

Mu, P., Li, Z.C., Li, C.P., Zhang, H.L., Wu, C.M., Li, C., and Wang, X.K. (2003) QTL mapping of the root traits and their correlation analysis with drought resistance using DH lines from paddy and upland rice cross. *China Science Bulletin* **48**: 2718-2724.

Munne-Bosche, S., and Penulelas, J. (2003) Photo- and antioxidative protection, and a role for salicyclic acid during drought and recovery in field-grown *Phillyrea angustifolia* plants. *Planta* **217**: 758-766.

Myburg, A.A., Griffin, A.R., Sederoff, R.R., and Whetten, R.W. (2003) Comparative genetic linkage maps of *Eucalyptus grandis, Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. *Theoretical and Applied Genetics* **107**: 1028-1048.

Nakano, T., Nagata, N., Kimura, T., Sekimoto, M., Kawaide, H., Murakami, S., Kaneko, Y., Matsushima, H., Kamiya, Y., Sato, F., and Yoshida, S. (2003) CND41, a chloroplast nucleoid protein that regulates plastid development, causes reduced gibberellin content and dwarfism in tobacco. *Physiologia Plantarum.* 117: 130-136.

Nanjo, T., Kobayashi, M., Yoshiba, Y., Kakubari, Y., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1999a) Antisense suppression of proline degradation improves tolerance to freezing and salinity in Arabidopsis thaliana. *FEBS Letters* **461**: 205-210.

Nanjo, T., Kobayashi, M., Yoshiba, Y., Sanada, Y., Wada, K., Tsukaya, H., Kakubari, Y., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1999b) Biological functions of proline in morphogenesis and osmotolerance revealed in antisense transgenic Arabidopsis thaliana. *Plant Journal* 18: 185-193.

Neill, S.J., Burnett, E.C., Desikan, R., and Hancock, J.T. (1998) Cloning of a wilt-responsive cDNA from an Arabidopsis thaliana suspension culture cDNA library that encodes a putative 9-cis- epoxy-carotenoid dioxygenase. *Journal of Experimental Botany* **49**: 1893-1894.

Ng, C.K.Y., Carr, K., McAinsh, M.R., Powell, B., and Hetherington, A.M. (2001) Drought-induced guard cell signal transduction involves sphingosine-1-phosphate. *Nature* **410**: 596-599.

Nguyen, D.V., and Rocke, D.M. (2001). Partial Least Squares Proportional Hazard Regression for Application to DNA Microarray Data. *Bioinformatics* **18**: 1625-1632

Nguyen, D.V., and Rocke, D.M. (2002) Tumor classification by partial least squares using microarray gene expression data. *Bioinformatics* **18**: 39-50.

Nickerson, D.A., Tobe, V.O., and Taylor, S.L. (1997) PolyPhred: Automating the detection and genotyping of single nucleotide substitutions using flourescence-based resequencing. *Nucleic Acids Research* **25**: 2745-2751.

Niinemets, U. (2001) Global-scale climatic controls of leaf dry mass per area, density, and thickness in trees and shrubs. *Ecology* **82**: 453-469.

Noctor, G., and Foyer, C.H. (1998) Ascorbate and glutathione: Keeping active oxygen under control. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**: 249-279.

Noctor, G., Veljovic-Jovanovic, S., Driscoll, S., Novitskaya, L., and Foyer, C.H. (2002) Drought and oxidative load in the leaves of C-3 plants: a predominant role for photorespiration? *Annals Botany* **89**: 841-850.

Nouzova, M., Neumann, P., Navratilova, A., Galbraith, D.W., and Macas, J. (2001) Microarray-based survey of repetitive genomic sequences in Vicia spp. *Plant Molecular Biology* **45**: 229-244.

Ori, N., Eshed, Y., Chuck, G., Bowman, J., and Hake, S. (2000) Mechanisms that control knox gene expression in the Arabidopsis shoot. *Development* **127**: 5523-5532.

Otto, S.P., and Jones, C.D. (2000) Detecting the undetected: estimating the total number of loci underlying a quantitative trait. *Genetics* **156**: 2093-2107.

Ozturk, Z.N., Talame, V., Deyholos, M., Michalowski, C.B., Galbraith, D.W., Gozukirmizi, N., Tuberosa, R., and Bohnert, H.J. (2002) Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. *Plant Molecular Biology* **48**: 551-573.

Pan, W. (2002) A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments. *Bioinformatics* **18**: 546-554.

Paran, I., and Michelmore, R.W. (1993) Development of reliable PCR-based markers linked to downy mildew resistence genes in lettuce. *Theoretical and Applied Genetics* **127**: 181-197.

Parry, M.A.J., Andralojc, P.J., Khan, S., Lea, P.J., and Keys, A.J. (2002) Rubisco Activity: Effects of Drought Stress. *Annals Botany* **89**: 833-839.

Pastori, G.M., and Trippi, V.S. (1992) Oxidative Stress Induces High-Rate Of Glutathione-Reductase Synthesis In A Drought-Resistant Maize Strain. *Plant And Cell Physiology* **33**: 957-961.

Paterson, A.H., Lander, E.S., Hewitt, J.D., Peterson, S., Lincoln, S.E., and Tanksley, S.D. (1988) Resolution of Quantitative Traits into Mendelian Factors by Using a Complete Linkage Map of Restriction Fragment Length Polymorphisms. *Nature* **335**: 721-726.

Patharkar, O.R., Cushman, J C. (2000) A stress-induced calcium-dependent protein kinase from *Mesembryanthemum crystallinum* phosphorylates a two-component pseudo-response regulator. *Plant Journal* **24**: 679-691.

Pei, Z.M., Kuchitsu, K., Ward, J.M., Schwarz, M., and Schroeder, J.I. (1997) Differential abscisic acid regulation of guard cell slow anion channels in Arabidopsis wild-type and abi1 and abi2 mutants. *Plant Cell* **9**: 409-423.

Perry, D.J., and Bousquet, J. (1998) Sequence-tagged-site (STS) markers of arbitrary genes: Development, characterization and analysis of linkage in black spruce. *Genetics* **149**: 1089-1098.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real- time RT-PCR. *Nucleic Acids Research* **29**: art. no.-e45.

Phimister, B. (1999) Going global. *Nature Genetics* **21**: 1-1.

Pical, C., Westergren, T., Dove, S.K., Larsson, C., and Sommarin, M. (1999) Salinity and Hyperosmotic Stress Induce Rapid Increases in Phosphatidylinositol 4,5-Bisphosphate, Diacylglycerol Pyrophosphate, and Phosphatidylcholine in *Arabidopsis thaliana* Cells. *Journal of Biological Chemistry* **274**: 38232-38240.

Pinon, J., and Valadon, A. (1997) Behaviour towards major parasites of poplar cultivars registered in European Union. *Annals of Forest Science* **54**: 19-38.

Plomion, C., Liu, B.H., and Omalley, D.M. (1996) Genetic analysis using trans-dominant linked markers in an F2 family. *Theoretical and Applied Genetics* **93**: 1083-1089.

Posas, F., Chambers, J.R., Heyman, J.A., Hoeffler, J.P., de Nadal, E., and Arino, J. (2000) The transcriptional response of yeast to saline stress. *Journal of Biological Chemistry* **275**: 17249-17255.

Price, A.H., Cairns, J.E., Horton, P., Jones, H.G., and Griffiths, H. (2002) Linking drought-resistance mechanisms to drought avoidance in upland rice using a QTL approach: progress and new opportunities to integrate stomatal and mesophyll responses. *Journal of Experimental Botany* **53**: 989-1004.

Pritchard, C.C., Hsu, L., Delrow, J., and Nelson, P.S. (2001) Project normal: defining normal variance in mouse gene expression. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **98**: 13266-13271.

Pruitt, R.E., Vielle-Calzada, J.-P., Ploense, S.E., Grossniklaus, U., and Lolle, S.J. (2000) FIDDLEHEAD, a gene required to suppress epidermal cell interactions in Arabidopsis, encodes a putative lipid biosynthetic enzyme. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **97**: 1311-1316.

Quackenbush, J. (2002) Microarray data normalization and transformation. *Nature Genetics* **32 Suppl 2**: 496-501.

Quintero, F.J., Garciadeblas, B., and RodriguezNavarro, A. (1996) The SAL1 gene of Arabidopsis, encoding an enzyme with 3'(2'),5'-bisphosphate nucleotidase and inositol polyphosphate 1-phosphatase activities, increases salt tolerance in yeast. *Plant Cell* **8**: 529-537.

R Development Core Team (2004). R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org

Rae, A.M., Robinson, K.M., Street, N.R., and Taylor, G. (2004) Morphological and physiological traits influencing biomass productivity in short-rotation coppice poplar. *Canadian Journal of Forest Research* **34**: 1488-1498.

Ramakers, C., Ruijter, J.M., Deprez, R.H.L., and Moorman, A.F.M. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters* **339**: 62-66.

Ramanjulu, S., and Bartels, D. (2002) Drought- and desiccation-induced modulation of gene expression in plants. *Plant Cell Environment* **25**: 141-151.

Raschke, K. (1975) Stomatal action. *Annual Review of Plant Physiology* **26**: 309-340.

Ratcliffe, O.J., Riechmann, J.L., and Zhang, J.Z. (2000) INTERFASCICULAR FIBERLESS1 is the same gene as REVOLUTA. *Plant Cell* **12**: 315-317.

Redei, G.P. (1962) Single locus heterosis. *Z. Vererbungs.* **93**: 164-170.

Reiner, A., Yekutieli, D., and Benjamini, Y. (2001) Identifying Differentially Expressed Genes Using False Discovery Rate Controlling Procedures. *Bioinformatics* **19**: 368-375

Rensink, W., Iobst, S., Hart, A., Stegalkina, S., Liu, J., and Buell, C. (2005) Gene expression profiling of potato responses to cold, heat, and salt stress. *Functional and Integrative Genomics* **5**: 201.

Rep, M., Krantz, M., Thevelein, J.M., and Hohmann, S. (2000) The transcriptional response of Saccharomyces cerevisiae to osmotic shock - Hot1p and Msn2p/Msn4p are required for the induction of subsets of high osmolarity glycerol pathway- dependent genes. *Journal of Biological Chemistry* **275**: 8290-8300.

Ribaut, J.M., and Hoisington, D. (1998) Marker-assisted selection: new tools and strategies. *Trends In Plant Science* **3**: 236-239.

Richmond, T., and Somerville, S. (2000) Chasing the dream: plant EST microarrays. *Current Opinions in Plant Biology* **3**: 108-116.

Ridolfi, M., Fauveau, M.L., Label, P., Garrec, J.P., and Dreyer, E. (1996) Water stress in an ABA-unresponsive hybrid poplar (Populus koveana x tvichocavpa cv Peace) - Response. *New Phytologist* **134**: 445-454.

Risch, N., and Merikangas, K. (1996) The future of genetic studies of complex human diseases. *Science* **273**: 1516-1517.

Rishi, A.S., Nelson, N.D., and Goyal, A. (2001) Improvement of *Populus* through genetic engineering. *Indian Journal of Plant Physiology* **6**: 119-126.

Robinson, K.M. (2004). Biomass determinants and their use as yield predictors in salicaceae. *PhD Thesis*, School of Biological Sciences, University of Southampton pp: 299.

Robison, D.J., and Raffa, K.F. (1998) Productivity, drought tolerance and pest status of hybrid Populus: Tree improvement and silvicultural implications. *Biomass and Bioenergy.* **14**: 1-20.

Rocke, D.M., and Durbin, B. (2001) A model for measurement error for gene expression arrays. *Journal of Computational Biology* **8**: 557-569.

Roelfsema, M.R.G., and Prins, H.B.A. (1995) Effect of abscisic acid on stomatal opening in isolated epidermal strips of *abi* mutants of *Arabidopsis thaliana. Physiologia Plantarum* **95**: 373-378.

Roelfsema, M.R.G., Levchenko, V., and Hedrich, R. (2004) ABA depolarizes guard cells in intact plants, through a transient activation of R- and S-type anion channels. *Plant Journal* **37**: 578-588.

Ronnberg-Wastljung, A.C., Glynn, C., and Weih, M. (2005) QTL analyses of drought tolerance and growth for a Salix dasyclados x Salix viminalis hybrid in contrasting water regimes. *Theoretical And Applied Genetics* **110**: 537-549.

Rood, S.B., Patino, S., Coombs, K., and Tyree, M.T. (2000) Branch sacrifice: cavitation-associated drought adaptation of riparian cottonwoods. *Trees - Stucture and Function* **14**: 248-257.

Rose, J.K.C., Braam, J., Fry, S.C., and Nishitani, K. (2002) The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. *Plant Cell Physiology* **143**: 1421-1435.

Saez, A., Apostolova, N., Gonalez-Guzman, M., Gonalez-Garcia, M.P., Nicolas, C., Lorenzo, O., and Rodriguez, P.L. (2004) Gain-of-function and loss-of-function phenotypes of the protein phosphatase 2C HAB1 reveal its role as a negative regulator of abscisic acid signalling. *The Plant Journal* **37**: 354-369.

Salekdeh, G.H., Siopongco, J., Wade, L.J., Ghareyazie, B., and Bennett, J. (2002) A proteomic approach to analyzing drought- and salt- responsiveness in rice. *Field Crop Research* **76**: 199-219.

Salvi, S., Talame, V., Sanguineti, M.C., Landi, P., Giuliani, S., Bellotti, M., Sponza, G., Maccaferri, M., Stefanelli, S., Rotondo, F., and Tuberosa, R. (2003) Genomic approaches to dissect the genetic basis of drought tolerance in maize and barley. *Journal of Experimental Botany* 54: 21-22.

Sanders, D., Brownlee, C., and Harper, J.F. (1999) Communicating with calcium. *Plant Cell* **11**: 691-706.

Sanguineti, M.C., Tuberosa, R., Landi, P., Salvi, S., Maccaferri, M., Casarini, E., and Conti, S. (1999) QTL analysis of drought related traits and grain yield in relation to genetic variation for leaf abscisic acid concentration in field-grown maize. *Journal of Experimental Botany* **50**: 1289-1297.

Satagopan, J.M., Yandell, B.S., Newton, M.A., and Osborn, T.G. (1996) A Bayesian approach to detect quantitative trait loci using Markov chain Monte Carlo. *Genetics* **144**: 805-816.

Sawa, S., Watanabe, K., Goto, K., Kanaya, E., Morita, E.H., and Okada, K. (1999) FILAMENTOUS FLOWER, a meristem and organ identity gene of Arabidopsis, encodes a protein with a zinc finger and HMG- related domains. *Genes and Development* **13**: 1079-1088.

Sawkins, M.C., Farmer, A.D., Hoisington, D., Sullivan, J., Tolopko, A., Jiang, Z., and Ribaut, J.M. (2004) Comparative Map and Trait Viewer (CMTV): an integrated bioinformatic tool to construct consensus maps and compare QTL and functional genomics data across genomes and experiments. *Plant Molecular Biology* **56**: 465-480.

Sax, K. (1923) The association of size difference with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* **8**: 552-560.

Schadt, E.E., Monks, S.A., Drake, T.A., Lusis, A.J., Che, N., Colinayo, V., Ruff, T.G., Milligan, S.B., Lamb, J.R., Cavet, G., Linsley, P.S., Mao, M., Stoughton, R.B., and Friend, S.H. (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422: 297-302.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**: 467-470.

Schmid, K.J., Sorensen, T.R., Stracke, R., Torjek, O., Altmann, T., Mitchell-Olds, T., and Weisshaar, B. (2003) Large-scale identification and analysis of genome-wide single- nucleotide polymorphisms for mapping in Arabidopsis thaliana. *Genome Research* **13**: 1250-1257.

Schmidt, C., Schelle, I., Liao, Y.J., and Schroeder, J.I. (1995) Strong Regulation of Slow Anion Channels and Abscisic-Acid Signaling in Guard-Cells by Phosphorylation and Dephosphorylation Events. *Proceedings of the National Academy of Sciences of the United States of America* **92**: 9535-9539.

Schork, N., Boehnke, M., and Terwilliger, J. (1993) Two-trait-locus linkage analysis: A powerful strategy for mapping complex traits. *American Journal of Human Genetics* **53**: 1127-1136.

Schroeder, J.I., and Hagiwara, S. (1990) Repetitive Increases in Cytosolic Ca-2+ of Guard-Cells by Abscisic-Acid Activation of Nonselective Ca-2+ Permeable Channels. *Proceedings of the National Academy of Sciences of the United States of America* **87**: 9305-9309.

Schroeder, J.I., Kwak, J.M., and Allen, G.J. (2001a) Guard cell abscisic acid signalling and engineering drought hardiness in plants. *Nature* **410**: 327-330.

Schroeder, J.I., Allen, G.J., Hugouvieux, V., Kwak, J.M., and Waner, D. (2001b) Guard cell signal transduction. *Annual Review of Plant Physiology and Plant Molecular Biology* **52**: 627-658.

Schuchhardt, J., Beule, D., Malik, A., Wolski, E., Eickhoff, H., Lehrach, H., and Herzel, H. (2000) Normalization strategies for cDNA microarrays. *Nucleic Acids Research* 28: E47.

Schwanz, P., and Polle, A. (2001) Differential stress responses of antioxidative systems to drought in pendunculate oak (Quercus robur) and maritime pine (Pinus pinaster) grown under high CO2 concentrations. *Journal of Experimental Botany* **52**: 133-143.

Schwartz, A., Wu, W.H., Tucker, E.B., and Assmann, S.M. (1994) Inhibition of Inward K+ Channels and Stomatal Response by Abscisic-Acid - an Intracellular Locus of Phytohormone Action. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 4019-4023.

Schwartz, S.H., Tan, B.C., Gage, D.A., Zeevaart, J.A.D., and McCarty, D.R. (1997) VP14 of maize catalyzes the carotenoid cleavage reaction of abscisic acid biosynthesis. *Plant Physiology* **114**: 798-798.

Searle, I.R., Men, A.E., Laniya, T.S., Buzas, D.M., Iturbe-Ormaetxe, I., Carroll, B.J., and Gresshoff, P.M. (2003) Long-distance signaling in nodulation directed by a CLAVATA1- like receptor kinase. *Science* 299: 109-112.

Seaton, G., Haley, C.S., Knott, S.A., Kearsey, M., and Visscher, P.M. (2002) QTL Express: mapping quantitative trait loci in of simple and complex pedigrees. *Bioinformatics* **18**: 339-340.

Seki, M., Narusaka, M., Yamaguchi-Shinozaki, K., Carninci, P., Kawai, J., Hayashizaki, Y., and Shinozaki, K. (2001a) Arabidopsis encyclopedia using full-length cDNAs and its application. *Plant Physiology and Biochemistry* **39**: 211-220.

Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y., and Shinozaki, K. (2001b) Monitoring the Expression Pattern of 1300 Arabidopsis Genes under Drought and Cold Stresses by Using a Full-Length cDNA Microarray. *Plant Cell* **13**: 61-72.

Seki, M., Ishida, J., Narusaka, M., Fujita, M., Nanjo, T., Umezawa, T., Kamiya, A., Nakajima, M., Enju, A., Sakurai, T., Satou, M., Akiyama, K., Yamaguchi-Shinozaki, K., Carninci, P., Kawai, J., Hayashizaki, Y., and Shinozaki, K. (2002a) Monitoring the expression pattern of around 7,000 Arabidopsis genes under ABA treatments using a full-length cDNA microarray. *Functional and Integrative Genomics* 2: 282-291.

Seki, M., Narusaka, M., Ishida, J., Nanjo, T., Fujita, M., Oono, Y., Kamiya, A., Nakajima, M., Enju, A., Sakurai, T., Satou, M., Akiyama, K., Taji, T., Yamaguchi-Shinozaki, K., Carninci, P., Kawai, J., Hayashizaki, Y., and Shinozaki, K. (2002b) Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *Plant Journal* 31: 279-292.

Seo, M., Peeters, A.J.M., Koiwai, H., Oritani, T., Marion-Poll, A., Zeevaart, J.A.D., Koornneef, M., Kamiya, Y., and Koshiba, T. (2000) The Arabidopsis aldehyde oxidase 3 (AA03) gene product catalyzes the final step in abscisic acid biosynthesis in leaves. *Proceedings of the National Academy of Sciences of the United States of America* 97: 12908-12913.

Shen, Q.X., and Ho, T.H.D. (1995) Functional Dissection of an Abscisic-Acid (Aba)-Inducible Gene Reveals 2 Independent Aba-Responsive Complexes Each Containing a G-Box and a Novel Cis-Acting Element. *Plant Cell* 7: 295-307.

Shevell, D.E., Leu, W.M., Gillmor, C.S., Xia, G.X., Feldmann, K.A., and Chua, N.H. (1994) Emb30 Is Essential for Normal-Cell Division, Cell Expansion, and Cell-Adhesion in Arabidopsis and Encodes a Protein That Has Similarity to Sec7. *Cell* **77**: 1051-1062.

Shi, Z.X., Chen, X.M., Line, R.F., Wellings, C.R. (2001) Development of resistance gene analog polymorphism markers for the *Yr9* gene resistance to wheat stripe rust. *Genome* **44**:509–516.

Shibaoka, H. (1994) Plant Hormone-Induced Changes in the Orientation of Cortical Microtubules - Alterations in the Cross-Linking between Microtubules and the Plasma-Membrane. *Annual Review of Plant Physiology and Plant Molecular Biology* **45**: 527-544.

Shinozaki, K., and YamaguchiShinozaki, K. (1997) Gene expression and signal transduction in water-stress response. *Plant Physiology* **115**: 327-334.

Shinozaki, K., and Yamaguchi-Shinozaki, K. (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Current Opinions in Plant Biology* **3**: 217-223.

Shinozaki, K., Yamaguchi-Shinozaki, K., and Seki, M. (2003) Regulatory network of gene expression in the drought and cold stress responses. *Current Opinions in Plant Biology* **6**: 410-417.

Sibole, J.V., Cabot, C., Poschenrieder, C., and Barcelo, J. (2003) Efficient leaf ion partitioning, an overriding condition for abscisic acid-controlled stomatal and leaf growth responses to NaCl salinization in two legumes. *Journal of Experimental Botany* **54**: 2111-2119.

Siegfried, K.R., Eshed, Y., Baum, S.F., Otsuga, D., Drews, G.N., and Bowman, J.L. (1999) Members of the YABBY gene family specify abaxial cell fate in Arabidopsis. *Development* **126**: 4117-4128.

Sillanpaa, M.J., and Arjas, E. (1998) Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. *Genetics* **148**: 1373-1388

Sillanpaa, M.J., and Arjas, E. (1999) Bayesian mapping of multiple quantitative trait loci from incomplete outbred offspring data. *Geneticcs* **151**: 1605-1619.

Silva, C., Garcia-Mas, J., Sanchez, A.M., Arus, P., and Oliveira, M.M. (2005) Looking into flowering time in almond (*Prunus dulcis* (Mill) D. A. Webb): the candidate gene approach. *Theoretical and Applied Genetics* **110**: 959.

Slavov, G.T., Howe, G.T., Yakovlev, I., Edwards, K.J., Krutovskii, K.V., Tuskan, G.A., Carlson, J.E., Strauss, S.H., and Adams, W.T. (2004) Highly variable SSR markers in Douglas-fir: Mendelian inheritance and map locations. *Theoretical and Applied Genetics* In press:

Smirnoff, N. (1993) Tansley Review.52. The Role of Active Oxygen in the Response of Plants to Water-Deficit and Desiccation. *New Phytologist* **125**: 27-58.

Smith, L.G., Greene, B., Veit, B., and Hake, S. (1992) A Dominant Mutation in the Maize Homeobox Gene, Knotted-1, Causes Its Ectopic Expression in Leaf-Cells with Altered Fates. *Development* **116**: 21

Smulders, M.J.M., van der Schoot, J., Arens, P., and Vosman, B. (2001) Trinucleotide repeat microsatellites for black poplar (*Populus nigra* L.). *Molecular Ecology Notes* **1**: 188-190.

Smyth, G.K. (2004) Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Statistical Applications in Genetics and Molecular Biology* **3**: Article 3.

Sokal, R.R., and Rohlf, J.F. (1995). Biometry. W.H. Freeman and Company, New York, USA.

Spiteller, G. (2003) The relationship between changes in the cell wall, lipid peroxidation, proliferation, senescence and cell death. *Physiologia Plantarum* **119**: 5-18.

Stephens, D.A., and Fisch, R.D. (1998) Bayesian analysis of quantitative trait locus data using reversible jump Markov chain Monte Carlo. *Biometrics* **54**: 1334-1347.

Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P., Brunner, A.M., Charbonnel-Campaa, L., Lindvall, J.J., Tandre, K., Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalerao, R.P., Nilsson, O., Sandberg, G., Karlsson, J., Lundeberg, J., and Jansson, S. (2004) A Populus EST resource for plant functional genomics. *Proceedings of the National Academy of Sciences of the United States of America* 101: 13951-13956.

Stirling, B., Yang, Z.K., Gunter, L.E., Tuskan, G.A., and Bradshaw, H.D. (2003) Comparative sequence analysis between orthologous regions of the Arabidopsis and Populus genomes reveals substantial synteny and microcollinearity. *Canadian Journal of Forest Research* **33**: 2245-2251.

Stockinger, E.J., Gilmour, S.J., and Thomashow, M.F. (1997) Arabidopsis thaliana CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 1035-1040.

Storey, J.D. (2002) A Direct Approach to Flase Discovery Rate. *Journal of the Royal Statistical Society, Series B* **64**: 479-798.

Storey, J.D., and Tibshirani, R. (2000). SAM Thresholding and False Discovery Rates for Detecting Differential Gene Expression in DNA Microarrays.*In* The Analysis of Gene Expression Data: Methods and Software. Springer, New York, USA.

Strauss, S.H., Lande, R., and Namkoong, G. (1992) Limitations of molecular-marker-aided selection in forest tree breeding. *Canadian Journal of Forest Research* **22**: 1050-1061.

Stuber, C.W., Lincoln, S.E., Wolff, D.W., Helentjaris, T., and Lander, E.S. (1992) Identification of genetic factors contributing to heterosis in a hybrid from 2 elite maize inbred lines using molecular markers. *Genetics* **132**: 823-839.

Sturn, A. (2000). Cluster Analysis for Large Scale Gene Expression Studies. Graz University of Technology and The Institute for Genomic Research, Rockville, USA.

Sturn, A., Quackenbush, J., and Trajanoski, Z. (2002) Genesis: cluster analysis of microarray data. *Bioinformatics* **18**: 207-208.

Sui, H.X., Han, B.G., Lee, J.K., Walian, P., and Jap, B.K. (2001) Structural basis of water-specific transport through the AQP1 water channel. *Nature* **414**: 872-878.

Syvanen, A.C. (2001) Accessing genetic variation: Genotyping single nucleotide polymorphisms. *Nature Reviews Genetics* **2**: 930-942.

Tahtiharju, S., Sangwan, V., Monroy, A.F., Dhindsa, R.S., and Borg, M. (1997) The induction of kin genes in cold-acclimating Arabidopsis thaliana. Evidence of a role for calcium. *Planta* **203**: 442-447.

Tähtiharju, S., and Palva, T. (2001) Antisense inhibition of protein phosphatase 2C accelerates cold acclimation in *Arabidopsis thaliana. The Plant Journal* **26**: 461-470.

Taiz, L., and Zeiger, E. (1998). Plant Physiology. Sinauer Associates Inc., Sunderland, MA, USA.

Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2002) Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana. *The Plant Journal* **29**: 417-426.

Talbert, P.B., Adler, H.T., Parks, D.W., and Comai, L. (1995) The Revoluta Gene Is Necessary for Apical Meristem Development and for Limiting Cell Divisions in the Leaves and Stems of Arabidopsis-Thaliana. *Development* **121**: 2723-2735.

Tan, B.C., Schwartz, S.H., Zeevaart, J.A.D., and McCarty, D.R. (1997) Genetic control of abscisic acid biosynthesis in maize. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 12235-12240.

Tanaka, Y., Sano, T., Tamaoki, M., Nakajima, N., Kondo, N., and Hasezawa, S. (2005) Ethylene Inhibits Abscisic Acid-Induced Stomatal Closure in Arabidopsis. *Plant Physiology* **138**: 2337-2343.

Tang, A.-C., Kawamitsu, Y., Kanechi, M., and Boyer, J.S. (2002) Photosynthetic oxygen evolution at low water potential in leaf discs lacking an epidermis. *Ann Bot* **89**: 861-870.

Taylor, G. (2002) Populus: Arabidopsis for Forestry. Do We Need a Model Tree? *Annals Botany* **90**: 681-689.

Taylor, G., Tricker, P.J., Zhang, F.Z., Alston, V.J., Miglietta, F., and Kuzminsky, E. (2003) Spatial and Temporal Effects of Free-Air CO2 Enrichment (POPFACE) on Leaf Growth, Cell Expansion, and Cell Production in a Closed Canopy of Poplar. *Plant Physiology* **131**: 177-185.

Temesgen, B., Brown, G.R., Harry, D.E., Kinlaw, C.S., Sewell, M.M., and Neale, D.B. (2001) Genetic mapping of expressed sequence tag polymorphism (ESTP) markers in loblolly pine (Pinus taeda L.). *Theoretical and Applied Genetics* **102**: 664-675.

Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S., and McCouch, S. (2000) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon association and genetic marker potential. *Genome Research* **11**: 1441-1452.

Tezara, W., Mitchell, V.J., Driscoll, S.D., and Lawlor, D.W. (1999) Water stress inhibits plant photosynthesis by decreasing coupling factor and ATP. *Nature* **401**: 914-917.

Thiel, G., and Blatt, M.R. (1994) Phosphatase Antagonist Okadaic Acid Inhibits Steady-State K+ Currents in Guard-Cells of Vicia-Faba. *The Plant Journal* **5**: 727-733.

This, B., D.C. Souyris, I. Teulat, B. (2000) QTL study of chlorophyll content as a genetic parameter of drought tolerance in barley. *Barely Genetics Newsletter* **30**: 20-24.

Thoday, J.M. (1961) Location of poly genes. *Nature* **191**: 368-370

Thomashow, M.F. (1999) Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. *Annual Review of Plant Physiology and Plant Molecular Biology* **50**: 571-599.

Toth, G., Gaspari, Z., and Jurke, J. (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research* **10**: 967-981.

Trotochaud, A.E., Hao, T., Wu, G., Yang, Z.B., and Clark, S.E. (1999) The CLAVATA1 receptor-like kinase requires CLAVATA3 for its assembly into a signaling complex that includes KAPP and a Rho- related protein. *The Plant Cell* **11**: 393-405.

Tschaplinski, T.J., Tuskan, G.A., Gebre, G.M., and Todd, D.E. (1998) Drought resistance of two hybrid Populus clones grown in a large-scale plantation. *Tree Physiology* **18**: 653-658.

Tsuge, T., Tsukaya, H., and Uchimiya, H. (1996) Two independent and polarized processes of cell elongation regulate leaf blade expansion in Arabidopsis thaliana (L) Heynh. *Development* **122**: 1589-1600.

Tsukaya, H. (2002). Interpretation of mutants in leaf morphology: Genetic evidence for a compensatory system in leaf morphogenesis that provides a new link between cell and organismal theories. *International Review of Cytology* **217**: 1-39.

Tsukaya, H. (2003) Organ shape and size: a lesson from studies of leaf morphogenesis. *Current Opinion in Plant Biology* **6**: 57-62.

Tsukaya, H., Tsuge, T., and Uchimiya, H. (1994) The cotyledon: a superior system for studies of leaf development. *Planta* **195**: 309-312.

Tsumura, Y., Suyama, Y., Yoshimura, K., Shirato, N., and Mukai, Y. (1997) Sequence-tagged-sites (STSs) of cDNA clones in *Cryptomeria japonica* and their evaluation as molecular markers in conifers. *Theoretical and Applied Genetics* **94**: 764-772.

Tuinstra, M.R., Ejeta, G., and Goldsbrough, P. (1998) Evaluation of near-isogenic Sorghum lines contrasting for QTL markers associated with drought tolerance. *Crop Science* **38**: 835-842.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 5116-5121.

Tuskan, G.A., and Walsh, M.E. (2001) Short-rotation woody crop systems, atmospheric carbon dioxide and carbon management: A US case study. *The Forestry Chronicle* **77**: 259-264.

Tuskan, G.A., Gunter, L.E., Yang, Z.M.K., Yin, T.M., Sewell, M.M., and DiFazio, S.P. (2004) Characterization of microsatellites revealed by genomic sequencing of Populus trichocarpa. *Canadian Journal of Forest Research* **34**: 85-93.

Tyerman, S.D., Niemietz, C.M., and Bramley, H. (2002) Plant aquaporins: multifunctional water and solute channels with expanding roles. *Plant, Cell and Environment* **25**: 173-194.

Tyerman, S.D., Bohnert, H.J., Maurel, C., Steudle, E., and Smith, J.A.C. (1999) Plant aquaporins: their molecular biology, biophysics and significance for plant water relations. *Journal of Experimental Botany* **50**: 1055-1071.

Uno, Y., Furihata, T., Abe, H., Yoshida, R., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2000) Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 11632-11637.

Urao, T., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2000) Two-component systems in plant signal transduction. *Trends in Plant Science* **5**: 67-74.

Urao, T., Katagiri, T., Mizoguchi, T., Yamaguchishinozaki, K., Hayashida, N., and Shinozaki, K. (1994) 2 Genes That Encode Ca2+-Dependent Protein-Kinases Are Induced by Drought and High-Salt Stresses in Arabidopsis-Thaliana. *Molecular Genetics and Genomics* **244**: 331-340.

Urao, T., Yakubov, B., Satoh, R., Yamaguchi-Shinozaki, K., Seki, M., Hirayama, T., and Shinozaki, K. (1999) A transmembrane hybrid-type histidine kinase in arabidopsis functions as an osmosensor. *The Plant Cell* **11**: 1743-1754.

van der Schoot, J., Pospiskova, M., Vosman, B., and Smulders, M.J.M. (2000) Development and characterization of microsatellite markers in black poplar (*Populus nigra* L.). *Theoretical and Applied Genetics* **101**: 317-322.

van der Ven, W.T.G., and McNichol, R.J. (1996) Microsatellite as DNA marker in Sitka spruce. *Theoretical and Applied Genetics* **93**: 613-617.

Van Volkenburgh, E. (1999) Leaf expansion - an integrating plant behaviour. *Plant, Cell and Environment* **22**: 1463-1473.

Venuprasad, R., Shashidhar, H.E., Hittalmani, S., and Hemamalini, G.S. (2002) Tagging quantitative trait loci associated with grain yield and root morphological traits in rice (Oryza sativa L.) under contrasting moisture regimes. *Euphytica* **128**: 293-300.

Verslues, P.E., and Bray, E.A. (2004) LWR1 and LWR2 Are Required for Osmoregulation and Osmotic Adjustment in Arabidopsis. *Plant Physiology* **136**: 2831-2842.

Villar, M., Lefevre, F., Bradshaw, H.D., and duCros, E.T. (1996) Molecular genetics of rust resistance in poplars (Melampsora larici-populina Kleb Populus sp) by bulked segregant analysis in a 2x2 factorial mating design. *Genetics* **143**: 531-536.

Volaire, F., and Thomas, H. (1995) Effects of Drought on Water Relations, Mineral Uptake, Water-Soluble Carbohydrate Accumulation and Survival of 2 Contrasting Populations of Cocksfoot (Dactylis-Glomerata L). *Annals of Botany* **75**: 513-524.

Vollbrecht, E., Veit, B., Sinha, N., and Hake, S. (1991) The Developmental Gene Knotted-1 Is a Member of a Maize Homeobox Gene Family. *Nature* **350**: 241-243.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. (1995) Aflp - a New Technique for DNA-Fingerprinting. *Nucleic Acids Research* 23: 4407-4414.

Vu, J.C.V., Baker, J.T., Pennanen, A.H., Allen, L.H., Bowes, G., and Boote, K.J. (1998) Elevated CO2 and water deficit effects on photosynthesis, ribulose bisphosphate carboxylase-oxygenase, and carbohydrate metabolism in rice. *Physiologia Plantarum* **103**: 327-339.

Waites, R., Selvadurai, H.R.N., Oliver, I.R., and Hudson, A. (1998) The PHANTASTICA gene encodes a MYB transcription factor involved in growth and dorsoventrality of lateral organs in Antirrhinum. *Cell* **93**: 779-789.

Wayne, M.L., and McIntyre, L.M. (2002) Combining mapping and arraying: An approach to candidate gene identification. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 14903-14906.

Webber, A.N., Nie, G.Y., and Long, S.P. (1994) Acclimation of Photosynthetic Proteins to Rising Atmospheric Co2. *Photosynthesis Research* **39**: 413-425.

Weigel, D., and Nordborg, M. (2005) Natural Variation in Arabidopsis. How Do We Find the Causal Genes? *Plant Physiology* **138**: 567-568.

Wellburn, A.R. (1994) The spectral determination of chlorophyll-A and chlorophhyll-B, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *Journal of Plant Physiology* **144**: 307-313.

Wilkinson, S., and Davies, W.J. (2002) ABA-based chemical signalling: the co-ordination of responses to stress in plants. *Plant, Cell and Environment* **25**: 195-210.

Williams, J., Bulman, M.P., and Neill, S.J. (1994) Wilt-Induced ABA Biosynthesis, Gene-Expression and Down- Regulation of Rbcs Messenger-RNA Levels in Arabidopsis-Thaliana. *Physiologia Plantarum* **91**: 177-182.

Willmann, M.R. (2000) CLV1 and CLV3: negative regulators of SAM stem cell accumulation. *Trends in Plant Science* **5**: 416-416.

Willmer, C., and Fricker, M. (1996). Stomata. Chapman and Hall, London, UK.

Wolfinger, R.D., Gibson, G., Wolfinger, E.D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C., and Paules, R.S. (2001) Assessing gene significance from cDNA microarray expression data via mixed models. *Journal of Computational Biololgy* **8**: 625-637.

Wright, S.T.C. (1969) An increase in the 'inhibitor-B' content of detached wheat leaves following a period of wilting. *Planta* **86**: 10-20.

Wright, S.T.C., and Hiron, R.W.P. (1969) (+)-Abscisic acid, the growth inhibitor induced in wheat leaves by a period of wilting. *Nature* **224**: 719-720.

Wu, H., Kerr, K.M., Cui, X., and Churchill, G.A. (2002) MAANOVA: A Software Package for the Analysis of Spotted cDNA Microarray Experiments. *Technical Report*: http://www.jax.org/staff/churchill/labsite/pubs/Wu_maanova.pdf

Wu, R., Bradshaw, H.D., and Stettler, R.F. (1997) Molecular genetics of growth and development in Populus (Salicaceae).5. Mapping quantitative trait loci affecting leaf variation. *American Journal of Botany* **84**: 143-153.

Wu, R., Bradshaw, H.D., and Stettler, R.F. (1998) Developmental quantitative genetics of growth in Populus. *Theoretical and Applied Genetics* **97**: 1110-1119.

Wu, R.L., Han, Y.F., Hu, J.J., Fang, J.J., Li, L., Li, M.L., and Zeng, Z.B. (2000) An integrated genetic map of Populus deltoides based on amplified fragment length polymorphisms. *Theoretical and Applied Genetics* **100**: 1249-1256.

Wu, T.D. (2002) Large-scale analysis of gene expression profiles. *Briefings in Bioinformatics* **3**: 7-17.

Wullschleger, S.D., Jansson, S., and Taylor, G. (2002) Genomics and Forest Biology: *Populus* Emerges as the Perennial Favorite. *The Plant Cell* **14**: 2651-2655.

Xiao, Y.N., Li, X.H., Zhang, S.H., Wang, X.D., Li, M.S., and Zheng, Y.L. (2004) Identification of quantitative trait loci (QTLs) for flowering time using SSR marker in maize under water stress. *Korean Journal of Genetics* **26**: 405-413.

Xiong, L., and Zhu, J.K. (2002) Molecular and genetic aspects of plant responses to osmotic stress. *Plant, Cell and Environment* **25**: 131-139.

Xiong, L.M., and Zhu, J.K. (2001) Abiotic stress signal transduction in plants: Molecular and genetic perspectives. *Physiologia Plantarum* **112**: 152-166.

Xiong, L.M., Schumaker, K.S., and Zhu, J.K. (2002) Cell signaling during cold, drought, and salt stress. *The Plant Cell* **14**: S165-S183.

Yalcin, B., Willis-Owen, S.A.G., Fullerton, J., Meesaq, A., Deacon, R.M., Rawlins, J.N.P., Copley, R.R., Morris, A.P., Flint, J., and Mott, R. (2004) Genetic dissection of a behavioral quantitative trait locus shows that *Rgs2* modulates anxiety in mice. *Nature Genetics* 36: 1197

Yale, J., and Bohnert, H.J. (2001) Transcript expression in Saccharomyces cerevisiae at high salinity. *Journal of Biological Chemistry* **276**: 15996-16007.

Yamaguchi-Shinozaki, K., and Shinozaki, K. (1994) A Novel Cis-Acting Element in an Arabidopsis Gene Is Involved in Responsiveness to Drought, Low-Temperature, or High-Salt Stress. *The Plant Cell* **6**: 251-264.

Yamaguchi-Shinozaki, K., and Shinozaki, K. (2004) Signal transduction pathways for the activation of ABA- responsive gene expression in drought-stress responsse. *Plant and Cell Physiology* **45**: S24-S24.

Yang, Y.H., Buckley, M.J., Dudoit, S., and Speed, T.P. (2002) Comparison of methods for image analysis on cDNA microarray data. *Journal of Computational and Graphical Statistics* **11**: 108-136.

Yi, N., and Xu, S. (2001) Bayesian Mapping of Quantitative Trait Loci Under Complicated Mating Designs. *Genetics* **157**: 1759-1771.

Yoshida, S. (2003) Molecular regulation of leaf senescence. *Current Opinion in Plant Biology* **6**: 79.

Zar, J.H. (1999). Biostatistical Analysis. Prentice Hall, New Jersey, USA.

Zeevaart, J.A.D., and Creelman, R.A. (1988) Metabolism and Physiology of Abscisic Acid. *Annual Review of Plant Physiology and Plant Molecular Biology* **39**: 439-473.

Zeiger, E., Farquhar, G.D., and Cowan, I.R. (1987). Stomatal function. Stanford University Press, Stanford, USA.

Zeng, Z.-B. (1994) Precision mapping of quantitative trait loci. *Genetics* **136**: 1347

Zheng, B.S., Yang, L., Zhang, W.P., Mao, C.Z., Wu, Y.R., Yi, K.K., Liu, F.Y., and Wu, P. (2003) Mapping QTLs and candidate genes for rice root traits under different water-supply conditions and comparative analysis across three populations. *Theoretical and Applied Genetics* **107**: 1505-1515.

Zhu, J.K. (2001) Plant salt tolerance. *Trends in Plant Science* **6**: 66-71.

Zhu, J.K. (2002) Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology* **53**: 247-273.

Zhu, J.K., Liu, J.P., and Xiong, L.M. (1998) Genetic analysis of salt tolerance in Arabidopsis: Evidence for a critical role of potassium nutrition. *The Plant Cell* **10**: 1181-1191.

Zhu, T., Budworth, P., Han, B., Brown, D., Chang, H.-S., Zou, G., and Wang, X. (2001) Toward elucidating the global gene expression patterns of developing Arabidopsis: Parallel analysis of 8300 genesby a high-density oligonucleotide probe array. *Plant Physiology and Biochemistry* **39**: 221-242.

Zik, M., and Irish, V.F. (2003) Global Identification of Target Genes Regulated by APETALA3 and PISTILLATA Floral Homeotic Gene Action. *The Plant Cell* **15**: 207-222.

Zuniga, G.E., Fernandez, J., Cristi, R., Alberdi, M., and Corcuera, L.J. (1990) Lipid Changes in Barley Seedlings Subjected to Water and Cold Stress. *Phytochemistry* **29**: 3087-3090.