UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES

School of Biological Sciences

USING MICROARRAYS TO ELUCIDATE THE GENETIC BASIS OF WOOD DENSITY IN SITKA SPRUCE AND POPLAR

by
Nicole Harris

Thesis for the degree of Doctor of Philosophy

June 2008

Correction Sheet

Location	Page 17, line 4
Before change	
Change	Addition of line: Thanks to the BBSRC (Biotechnology and Biological Sciences Research Council) for providing the funding for this project.
After change	

Location	Page 19, line 25
Before change	Pine xtension-like protein
Change	Addition of 'e' to 'extension'
After change	Pine extension-like protein

Chapter One Corrections

Location	Page 37
Change	Addition of 'sinapyl alcohol' to the diagram

Location	Page 53, line 26
Before change	The primary cell wall is membrane-like and rich in pectins
Change	Removal of 'membrane-like and'
After change	The primary cell wall is rich in pectins

Location	Page 60, line 25
Before change	
Change	Removal of '(i.e. produced during the first twelve to fifteen years after planting)'
After change	

Location	Page 61, line 5
Before change	Two contracting species…
Change	Change of the word 'contracting'
After change	Two contrasting species…

Chapter Two Corrections

Location	Page 65, line 4
Before change	
Change	Addition of the line: 'The modified protocol was developed by Nathaniel Street and James Tucker (pers.comm.) using poplar leaves.
After change	

Location	Page 66, line 11
Before change	Hybridisation was performed….
Change	Addition of the words 'for all arrays'
After change	Hybridisation for all arrays was performed…

Location	Page 66, line 17
Before change	
Change	Addition of the line 'A full description of the method used is provided in section 4.2.4.'
After change	

Location	Page 66, line 18
Before change	The slides were scanned…
Change	Addition of 'for the first Sitka experiment'
After change	The slides for the first Sitka experiment were scanned…

Location	Page 66, line 19
Before change	
Change	Addition of the line: The slides for experiments two and three were scanned using a Genetix scanner on default settings.
After change	

Chapter Three Corrections

Location	Page 69, line 27
Before change	The work described in this thesis involved a collaboration between…
Change	Removal of 'a'
After change	The work described in this thesis involved collaboration between…

Location	Page 70
Before change	
Change	Addition of line: Basic density measurements were collected using the water displacement method (see Lee, 1997).
After change	

Location	Page 73
Before change	
Change	Addition of the line: The Pilodyn measurements were collected by Forest Research at the same time that the wood cores were collected.
After change	

Location	Page 75
Before change	The low density clones to the left of the graph have higher pilodyn measurements.
Change	Addition of 'tend to', and 'although the correlation is weak (see figure 3.2).'
After change	The low density clones to the left of the graph tend to have higher pilodyn measurements, although the correlation is weak (see figure 3.2).

Location	Page 84
Before change	Average density and diameter measurements for all clones, high density clones and low density clones.
Change	Change to figure legend
After change	Average density and diameter measurements for all clones from experiment 35. Average wood density is also shown for the selected high density and low density clones.

Chapter Four Corrections

Location	Page 94, line 3
Before change	
Change	Removal of the paragraph beginning 'Wood quality can be determined by…'
After change	

Location	Page 94, line 21
Before change	As the aim of this project is to identify genes involved with wood formation…

Change	Change the word ‘formation’ for ‘density’
After change	As the aim of this project is to identify genes involved with wood density…

Location	Page 95, line 18
Before change	Fluorescently labelled mRNA is hybridised….
Change	Change ‘mRNA’ for ‘cDNA’
After change	Fluorescently labelled cDNA is hybridised….

Location	Page 100, line 4
Before change	Cambial samples were collected during May 2004 from 150 trees.
Change	
After change	Wood-forming tissue samples were collected during May 2004 from 150 trees (see Chapter 2 for methodology).

Location	Page 100, line 6
Before change	From this group of samples 60 samples were selected for microarray analysis.
Change	
After change	Of these 150 trees, only 10 out of the 15 replicates per clone had physiological data collected, so 100 samples were available for analysis. From this group, 60 samples were selected for microarray analysis.

Location	Page 101, line 4
Before change	
Change	Addition of the line ‘For this experiment, samples were only selected from 9 of the 10 clones available, due to a lack of mRNA after extraction from the low density clone C20145.’
After change	

Location	Page 102, line 2
Before change	The ten replicates for each clone were paired up to produce ten pairs of samples…
Change	
After change	The ten replicates for clone C20133 were paired up with the ten replicates from clone C20166 to produce ten pairs of samples with contrasting wood densities (to be hybridised to ten arrays).

Location	Page 104, line 13
Before change	After the heat denature was completed…
Change	Change the word ‘denature’ for ‘denaturation’
After change	After the heat denaturation was completed…

Location	Page 105
Before change	
Change	Addition of LNA to the list of abbreviations
After change	

Location	Page 114, line 13
Before change	
Change	Addition of the line '(i.e. there were no genes that increased in expression level in correlation with wood density).'
After change	

Location	Page 136
Before change	
Change	Addition of figure 4.13
After change	

Location	Page 143
Before change	Many of the genes…
Change	Use an actual percentage
After change	44% of the genes…

Chapter five corrections

Location	Page 145
Before change	It will then be possible to compare the gene lists….
Change	
After change	This will be compared with the gene lists……

Location	Page 149
Before change	The genotypes were then pooled into four groups for hybridisation to the four Affymetrix arrays available.
Change	
After change	RNA was extracted from each replicate, and the genotypes were then pooled into four groups for hybridisation to the four Affymetrix arrays available. Each pool contained RNA from two genotypes, with two replicates representing each genotype.

Chapter Six Correction

Location	Page 172
Before change	
Change	Addition of 'The type of approach suggested above may be more suitable for research, in terms of assigning definitive functions to some of our genes of interest. It is possible to assign putative functions based on homology to other sequences, but altering the genotype to assess changes in the phenotype is a useful method for validating putative functions.'
After change	

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES

SCHOOL OF BIOLOGICAL SCIENCES

Doctor of Philosophy

Using Microarrays to Elucidate the Genetic Basis of Wood Density in Sitka Spruce and Poplar

By Nicole Louise Harris

As the global population continues to increase, so will the demand for timber (and other raw materials) for building, construction, and also for the pulping industry. The high demand for wood and the increasing human population mean that natural forests are being lost and degraded. A potential solution to this problem is to improve the productivity of our plantation forests to relieve the pressure on natural forests in terms of sustainable wood production. This project is the first to use newly available microarray technology to study differential gene expression in cambial tissue of high versus low wood density field grown samples from two contrasting species, Sitka spruce (gymnosperms) and poplar (angiosperms). Genes up-regulated in high-density Sitka spruce and poplar samples had functions in cell formation and expansion, with down-regulated genes having functions in lignin biosynthesis, stress-response and defence. Plantation trees could be screened at a young age to assess their expression of candidate genes to speed up the breeding and selection process.

List of contents

List of figures

List of tables

DECLARATION OF AUTHORSHIP

I, ……NICOLE HARRIS……, [please print name]

declare that the thesis entitled [enter title]

USING MICROARRAYS TO ELUCIDATE THE GENETIC BASIS OF WOOD DENSITY IN SITKA SPRUCE AND POPLAR

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

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission, **or** [delete as appropriate] parts of this work have been published as: [please list references]

Signed: ……..
Date:…….

Acknowledgements

I would like to thank my supervisors Gail Taylor and Joan Cottrell for help, advice and patience with this project. Thanks to the BBSRC (Biotechnology and Biological Sciences Research Council) for providing the funding for this project. I would also like to acknowledge the Forestry Commission and the SCRI for the original density, lignin and cellulose data for the experiment 35 trees, and the Forestry Commission for the diameter and pilodyn data. Thanks to Harriet Trewin for help with collecting samples, Dave Watterson at the Forestry Commission for organising health and safety checks, Steve Lee at the Forestry Commission for permission to use the experiment 37 trees, and help with all aspects of this thesis. Thanks to James Duff and Rob Sykes for help and advice with collecting wood cores; Steven Ralph, Jörg Bohlmann and Sharon Janscik at UBC for providing and helping with the spruce microarrays; Rick White at UBC for help with microarray analysis; Nat Street for help and advice with the microarray analysis; Nat Street and James Tucker for help with RNA extraction; and NASC for the poplar arrays. I would particularly like to thank Caroline Dixon for help on various field trips and for transportation. I would finally like to thank all the members of the Plant and Environment Lab for their helpful discussion and advice.

List of abbreviations

2-ME	2-mercaptoethanol
4CL	Hydroxycinnamate:CoA ligase
ANOVA	Analysis of variance
C	Centigrade
C3H	*P*-coumarate 3-hydroxylase
C4H	Cinnamic acid 4-hydroxylase
CAD	Hydroxycinnamyl alcohol dehydrogenase
CCoAOMT	Caffeoyl-CoA O-methyltransferase
CCR	Hydroxycinnamoyl-CoA reductase
cDNA	Complementary DNA
cDNA-AFLP	cDNA amplified fragment length polymorphism
CHISAM	Chloroform : Isoamyl alcohol
cm	Centimetre
CoA	Coenzyme A
COMT	Caffeic acid O-methyltransferase
CQT	Hydroxycinnamoyl CoA:quinate hydroxycinnamoyltransferase
CST	Hydroxycinnamoyl CoA:shikimate hydroxycinnamoyltransferase
CTAB	Hexadecyltrimethylammonium bromide
DDRT-PCR	Differential display reverse transcription PCR
DNA	Deoxyribonucleic acid
EDTA	Ethylenediamine tetraacetic acid
eQTLs	Expression quantitative trait loci
EST	Expressed sequence tag
EtOH	Ethanol
F5H	Ferulate 5-hydroxylase
FSD	Family-specific domain
g	Gram
G	Guaiacyl
GAs	Gibberellins
GAX	Glucuronoarabinoxylans
h	Hours
H	*P*-hydroxyphenyl
HCl	Hydrogen chloride
IAA	Indole-3-acetic acid
iAFLP	Introduced amplified fragment length polymorphism

JGI	Joint Genome Institute
K	Thousand
Kg	Kilogram
Km	Kilometre
L	Litre
LiCl	Lithium Chloride
LOWESS	Locally weighted scatterplot smoothing
LNA	Locked nucleic acid
m	Metre
M	Molar
Mbp	Million base pairs
MFA	Microfibril angle
MFs	Microfibrils
ml	Millilitre
µl	Microlitre
mm	Millimetre
µm	Micrometre
MLG	Mixed-linkage glucans
M_q H_2O	Milli-Q water
mRNA	Messenger RNA
MW	Molecular weight
N	North
NaCl	Sodium chloride
NaOH	Sodium hydroxide
PAL	Phenylalanine ammonia-lyase
PCR	Polymerase chain reaction
PELP	Pine extension-like protein
PCD	Programmed cell death
PMT	Photomultiplier tube
PVP	Polyvinyl pyrrolidone
QCI	Queen Charlotte Islands
QTL	Quantitative trait locus
r	Correlation
r_a	Genotypic correlation
RC4D	Restriction fragment-coupled differential display
r_g	Genetic correlation
RNA	Ribonucleic acid

RPM	Revolutions per minute
S	Syringyl
S1	First layer of secondary cell wall
S2	Second layer of secondary cell wall
S3	Third layer of secondary cell wall
SAD	Sinapyl alcohol dehydrogenase
SAGE	Serial analysis of gene expression
SBH	Sequencing by hybridisation
SDS	Sodium dodecyl sulfate
SH	Subtractive hybridisations
SSC	Saline sodium citrate
SSTE	Defined in chapter 2
TE	TE (Tris + EDTA) buffer
TIFF	Tagged image file format
TRIS	2-amino-2-hydroxymethyl-1,3-propanediol
Trp	Tryptophan
XyG	Xyloglucans

1 General Introduction

1.1 Project aims

The aim of this project is to determine the link between wood density and gene expression in Sitka spruce and Poplar.

1.2 Project overview

The current estimate for the total world human population is 6.6 billion, with a population increase of approximately 75 million people per year (U.S. Census Bureau). Figure 1.1 shows estimated world population growth until the year 2050. Along with this massive population increase, in some areas the rate of consumption of raw materials has been growing at an even faster rate than the population (Bowyer *et al*., 2003). According to the Food and Agriculture Organisation (FAO), the world lost an estimated 14.6 million hectares of natural forests per year between the years 1990 and 2000 (FAO Forest Resources Assessment 2000).

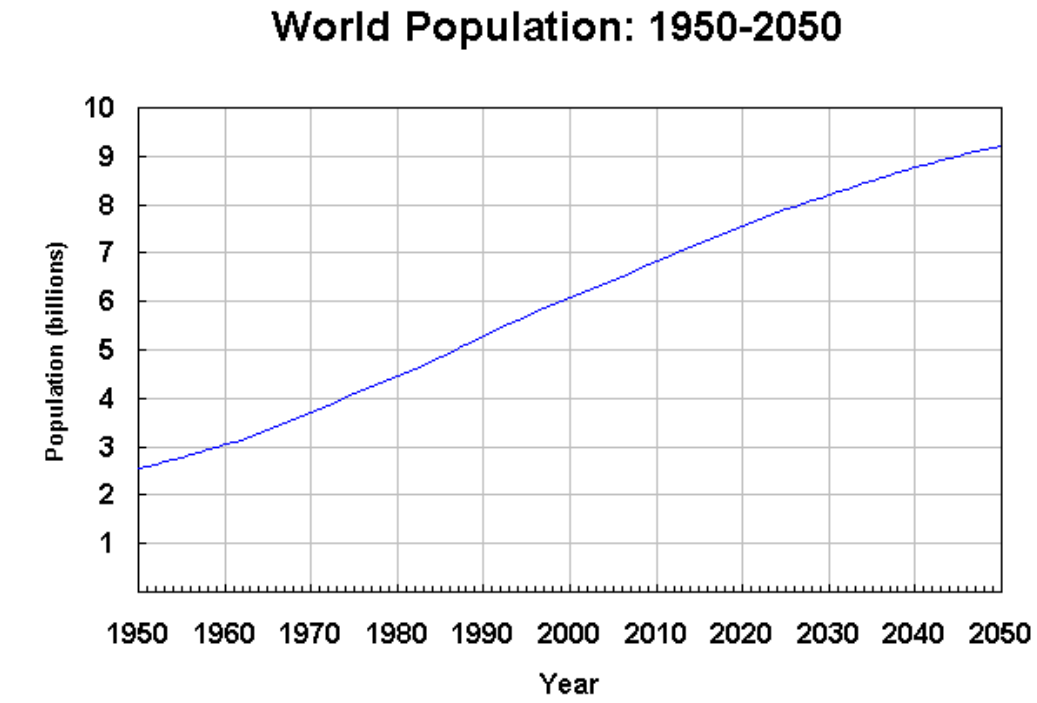


Figure 1.1: Predicted increase in world population from 1950-2050

The demand for tree biomass is predicted to increase in the future – both to act as a sink for atmospheric carbon dioxide, and to provide a renewable alternative to fossil fuels (Campbell *et al*., 2003). As the global population continues to increase, so will the demand for timber (and other raw materials) for building, construction, and also for the pulping industry. Projections indicate that the demand for wood will increase until the year 2030 (Whetten and Sederoff, 1991). The high demand for wood and the increasing human population mean that natural forests are being lost and degraded. Deforestation is responsible for approximately 20% of carbon emissions annually (Hall, 2008); and this has to be reduced. The area available for forestry is also decreasing; as the rising population has led to an increase in urbanisation and conversion of land for agricultural purposes. Plantation forests must produce a higher yield of timber than natural forests, with shorter rotation times, to meet the demand (Fenning and Gershenzon, 2002).

Breeding programmes are usually aimed at improving wood quality and density, including growth rate and the shape of the wood (Tzfira *et al*., 1998), but these are slow to implement. Tree improvement is hindered by the time needed to reach sexual maturity and the age at which performance can be assessed, as many traits can only be evaluated at maturity. For this reason, consideration must be given to biotechnological techniques which may offer the opportunity to accelerate current rates of improvement of our forest trees. Any procedure that improves the accuracy, speed and age at which evaluation of genetic value can be assessed and reduces the generation time would be of considerable benefit (Verhaegen *et al*., 1997). Improvement of yield of wood produced and improvement of the wood properties (e.g. strength) would be ideal targets.

1.3 General introduction to wood

Why is wood necessary to the tree?
Wood is necessary for several reasons. It provides mechanical support to the tree. It allows long distance conductance of water and nutrients from the roots to the leaves. It ensures the constant life of the tree through the cyclic activity of the cambium, producing new xylem and phloem each year.

Why is wood useful to man?
Wood pulp is used to produce paper and sawn logs produce timber for construction purposes. Wood is also used as a source of energy.

Why study wood formation?
Wood formation is a good model system for the analysis of differentiation in higher plants, and also programmed cell death. Its development is controlled by a wide variety of factors both endogenous and exogenous. From the analysis of wood formation it may be possible to identify genes involved in the direct modification of wood properties.

Why improve wood quality?
Timber being used for the construction industry will be allocated a strength grade. Producing Sitka spruce with stronger wood will increase its strength grading and also its commercial value. By improving wood quality and increasing the productivity of commercial plantations, we can protect the natural forest ecosystems from pressure caused by the increase in demand for wood.

1.4 General introduction to Sitka spruce

Conifers are one of the most genetically variable groups of forest tree species (Leibenguth and Shoghi, 1998), due to their wide geographical distributions. Sitka spruce is one of about 40 species from the genus *Picea,* from the order *Pinaceae*. It is a commercially important species (Lee *et al*., 2002) from the north temperate coastal rain forests of western North America. It can grow to a height of 80 m or more, and may attain diameters of over 5 m (Gymnosperm Database, http://www.conifers.org/pi/pic/sitchensis.htm). Timber derived from Sitka spruce is used for boat, aeroplane, ship and house construction, pallets, board manufacture, musical instruments and paper.

David Douglas originally introduced Sitka spruce into Great Britain in the early 1800s when it was known as *Pinus menziesii*; it was later described as *Pinus sitchensis* by botanist Bongard, and was then recognised as a member of the genus *Picea* by Carrière. It is now known as *Picea sitchensis* (Bong.) Carr. Common names include airplane spruce, coast spruce, Menzies spruce, silver spruce, tideland spruce, western spruce, and yellow spruce.

The natural range of Sitka spruce extends for 2,900 km from 39º N in northern California to 61º N latitude in south-central Alaska. It distribution is restricted to areas of maritime climate due to its requirements for abundant moisture, mild winters and cool summers.

1.5 General Introduction to Poplar

Populus L. is a genus of approximately 30 species of poplars, aspens and cottonwoods in the family Salicaceae. These species are native to temperate and boreal parts of the northern hemisphere (Stettler *et al*., 1996). The species are all large, single trunked trees that spread clonally from their vigorous root systems using sucker shoots (Stettler *et al*., 1996)

Poplar wood has multiple uses including timber, pulp (paper) and as a fuel source. The genus *Populus* has been divided into six sections: *Populus* (aspens and white poplar), Aigeiros (black poplars or cottonwoods), Tacamahaca (balsam poplars), Leucoides (necklace poplars), Turanga (subtropical poplars) and Abaso (Mexican poplars). Poplars are a dioecious species, and they flower before leaf emergence in the spring (Stettler *et al*., 1996).

1.6 Wood types

Woods can be classified into two types: softwoods and hardwoods. The trees which produce both types of wood belong to the division Spermatophytes (seed plants); however they are then separated into different subdivisions. Hardwoods are angiosperms, meaning they produce their seeds within ovaries, and softwoods are gymnosperms, their seeds (embryos) are not protected by a covering layer. Softwoods, such as Sitka spruce, have needle-like leaves (for reduced evapotranspiration during the summer) and mostly remain green throughout the year. These trees are referred to as conifers due to the cones inside which their seeds are produced.

Softwoods and hardwoods also vary in their manner of wood formation. They differ in the type and arrangement of cells produced (Bowyer *et al*., 2003).

Hardwoods – dicotyledons (angiosperms)	Softwoods – conifers (gymnosperms)
Seeds produced within ovaries	Seeds not protected by a covering layer
Broad leaves	Needle-like leaves
Deciduous	Evergreen (except *Larix*)
Many cell types in the wood	Few cell types in the wood
Wood contains vessel elements	Wood consists mainly of tracheids and paranchyma

Table 1.1: Differences between hardwoods and softwoods

1.6.1 Softwoods

Approximately 95% of a softwood tree is composed of tracheids (see figure 1.2) whose function is to conduct water and also to provide support to the tree (Lichtenegger *et al*., 1999). The tracheid cell is closed at each end and is perforated with pits, which occur as matched pairs with adjacent cells.

Ray cells, grouped together to form rays, enable horizontal movement of substances across the tree (Bowyer *et al*., 2003). Rays can be composed of either ray parenchyma cells, ray tracheids or a mixture of the two types. Longitudinal tracheids are oriented parallel to the stem axis, and rays run horizontally through the tree stem (figure 1.3). Some softwood species also contain a few other types of cells, including longitudinal parenchyma and epithelial cells (which secrete resin).

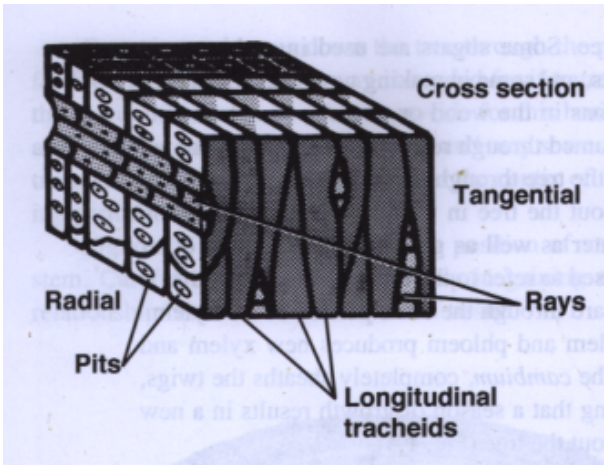


Figure 1.2: representation of a block of softwood (taken from Bowyer et al., 2003).

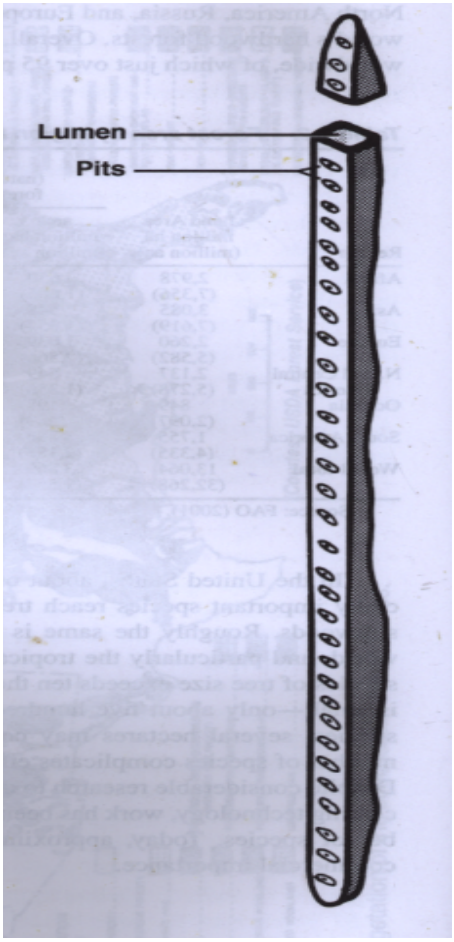


Figure 1.3: representation of a softwood longitudinal tracheid (taken from Bowyer et al., 2003).

Other types of cells may also be present in certain

species. Parenchyma cells are involved in the storage and movement of food substances, and make up a minor proportion of the volume of some softwood species. Ray tracheids have thick cells wall and are similar in shape to longitudinal tracheids. The ray parenchyma cells can have either thin or thick walls.

Sitka spruce (and other conifers such as larch and Douglas-fir) contain resin canals; an intercellular space surrounded by parenchyma cells (Bowyer *et al*., 2003). Parenchyma cells release resin into the canal. The resin is thought to have a function in wound healing and in conferring resistance to insect attack. The specialised parenchyma cells that secrete the resin are referred to as epithelial cells (Bowyer et al., 2003).

1.6.2 Hardwoods

Hardwoods are made up of four major types of cells; vessel elements, longitudinal parenchyma, ray parenchyma and fibre tracheids (Bowyer *et al*., 2003). Vessel elements are much shorter and wider than softwood tracheids, and they link end to end to form vessels for conduction of water. Fibre tracheids are the cells of hardwood xylem. They are shorter than softwood longitudinal tracheids, and they tend to be rounded in cross section. As the most of the water conduction occurs through the vessels, the main job of the fibre tracheids is to provide support. Strength in hardwoods is directly proportional to the percentage of thick-walled fibre tracheids (Bowyer *et al*., 2003).

Parenchyma cells occur as either longitudinal cells or ray cells, both types function as storage units. A few hardwood species contain gum canals (similar to resin canals in softwoods) which are surrounded by short epithelial parenchyma cells.

1.7 Wood

A detailed knowledge of the processes involved in xylogenesis (wood formation) is required to identify the factors that affect wood quality in Sitka spruce. The physical properties of wood are determined by the genes expressed during xylogenesis (Kirst *et al*., 2004); as well as environmental factors such as site quality, wind, and slope (Macdonald and Hubert, 2002).

1.7.1 Wood formation

Wood formation (xylogenesis) is the successive addition of secondary xylem formed by the division and subsequent differentiation of the cambial cells. This process is made up of five steps: cell division, cell enlargement, cell wall thickening, lignification and programmed cell death (Yang *et al*., 2004). The step of cell wall thickening includes the biosynthesis and deposition of cellulose, hemicelluloses, and cell wall proteins (Le Provost *et al*., 2003). Lignification occurs after cell division and expansion as it restricts the expansion of the cell wall (Patzlaff *et al*., 2003). The formation of wood is controlled by many factors including photoperiod, temperature, and phytohormones (Plomion *et al*., 2001). Wood properties (physical and chemical) are derived from wood cell morphology and composition (Allona *et al*., 1998).

1.7.1.1 Cell division and expansion

Wood cells are produced by the vascular cambium, a thin layer of tissue that is found between the phloem and the xylem (i.e. the wood and the bark). The cambium produces new xylem and phloem tissue. The cambium is initially produced by the apical meristem; this is the region of cell division at the top of the main stem.

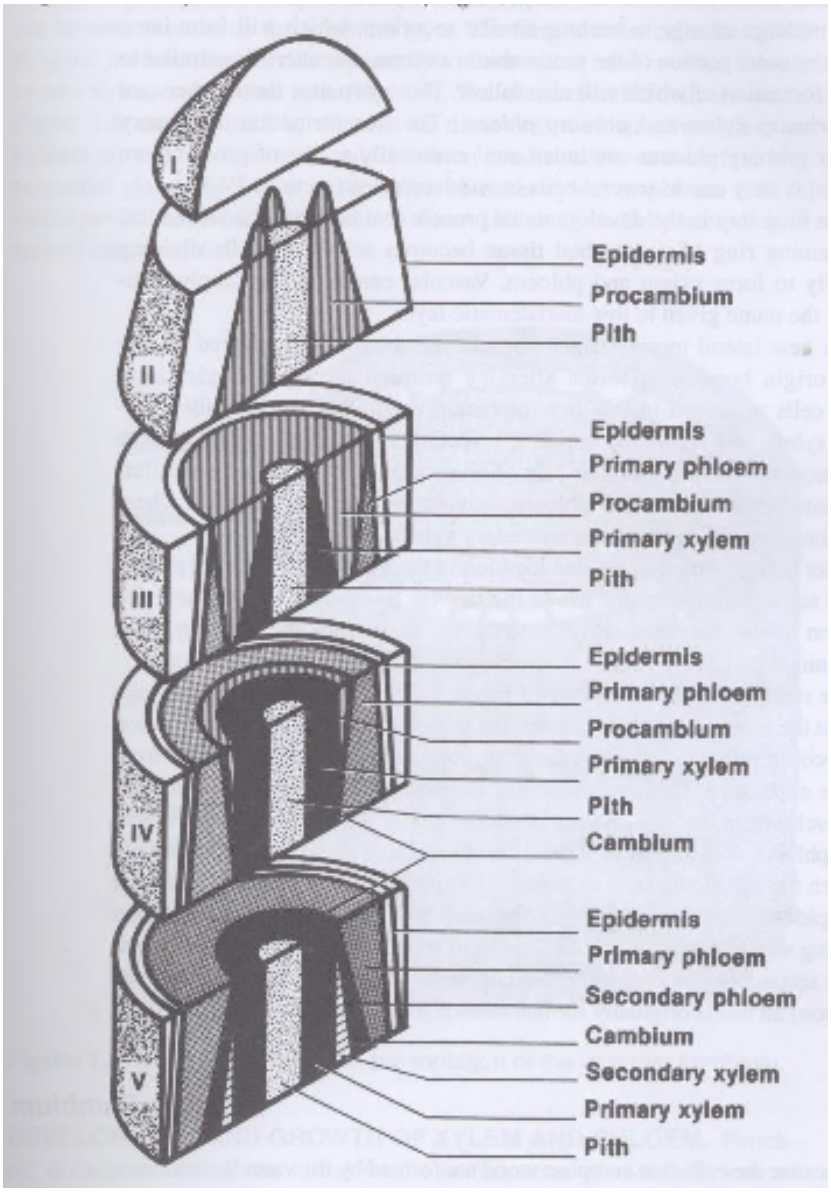


Figure 1.4: a developing stem (Bowyer et al., 2003) showing the many stages of stem growth (wood formation, starting with the formation of the primary xylem and phloem, leading through to a dividing meristematic region that produces secondary phloem and secondary xylem).

Figure 1.4 shows the development of various layers of tissue starting with the apical meristem. Cells divide and elongate resulting in growth, and after this cell differentiation occurs. The first layer to be formed is the epidermis, a single layer of cells that have waxy outer walls to prevent water loss. The second layer is the

procambium, which will eventually form the vascular cambium. In the very centre of the stem the pith is formed.

When the procambium reaches a maximum size it starts to form layers of primary xylem and primary phloem (Bowyer *et al*., 2003). The procambium cells change into these two tissues leaving a very small ring of procambium tissue. This ring of tissue then becomes active, producing both xylem and phloem. This layer of active tissue is referred to as the vascular cambium.

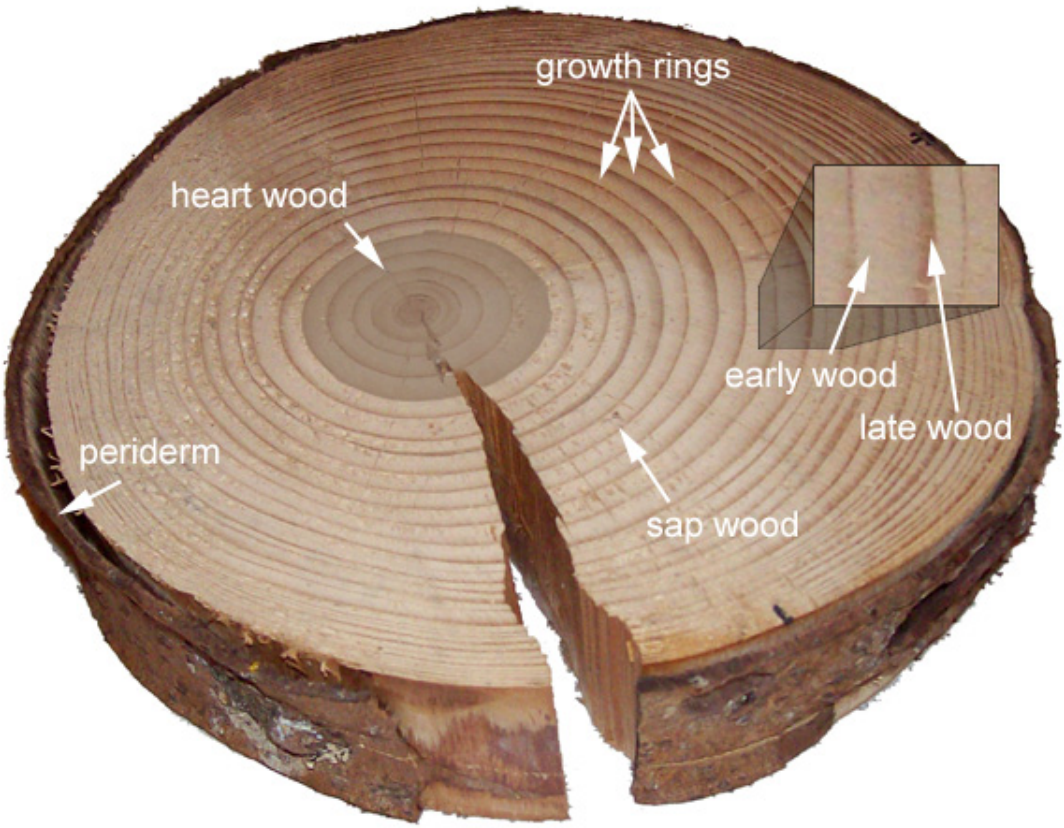


Figure 1.5: the structure of wood

Figure 1.5 shows the individual layers of tissue that are formed during xylogenesis. The new ring of xylem added during each growth season is referred to as a growth ring. Each growth ring can be split into two distinct phases, the

early wood and the late wood. Early wood is formed at the start of the growing season when conditions are most favourable. Late wood is formed when growth slows down towards the end of the growing seasons. Latewood cells are smaller in diameter than earlywood cells, and also have thicker walls and smaller lumens (Bowyer *et al*., 2003).

The vascular cambium contains several types of cells, including juvenile cells, called initials, and the phloem and xylem mother cells, which are produced by the initials (Plomion *et al*., 2001). There are two types of initials; fusiform initials and ray initials. When a fusiform initial divides it forms two cells. One of these will remain an active cambial cell, and the other will become either a xylem or phloem mother cell. These mother cells expand in diameter and length, and deposition of lignin after cell-wall thickening completes the process of maturation (Bowyer *et al*., 2003). Ray initials divide to produce either phloem or xylem ray cells or new ray initials.

Vascular cambium becomes dormant when the temperature drops, reactivating in spring in response to various signals. During dormancy, cell production is suspended, and response to growth-promoting signals is inhibited (Schrader *et al*., 2004b).

1.7.1.2 Cell wall thickening

The xylem and phloem mother cells divide and expand to produce mature xylem and phloem cells. The rate of expansion is determined by xyloglucan endotransglycosylases, endoglucanases, expansins, pectin methyl esterases and pectinases (Plomion *et al*., 2001). Secondary cell wall formation occurs after cell expansion has finished, as the process of lignification prevents further cell growth. During secondary cell wall formation the biosynthesis of four compounds occurs: polysaccharides (cellulose and hemicelluloses), lignins, cell wall proteins and other compounds (tannins, pectins, cell wall proteins, etc.).

Cellulose is the major structural element of plant cell walls (Williamson *et al*., 2002), and it constitutes approximately 40-45% of the dry weight of wood (Bowyer *et al*., 2003). It is a long-chain carbohydrate, $(C_6H_{10}O_5)_n$ that is formed from a sugar nucleotide (constituent of RNA or DNA). Glucose molecules ($C_6H_{12}O_6$) are joined together through the removal of a molecule of H_2O to form a cellulose chain. Figure 1.6 shows the chemical linkage of two glucose molecules into a cellulose chain. The units are connected through β linkages, which cannot be digested by humans (Bowyer *et al*., 2003).

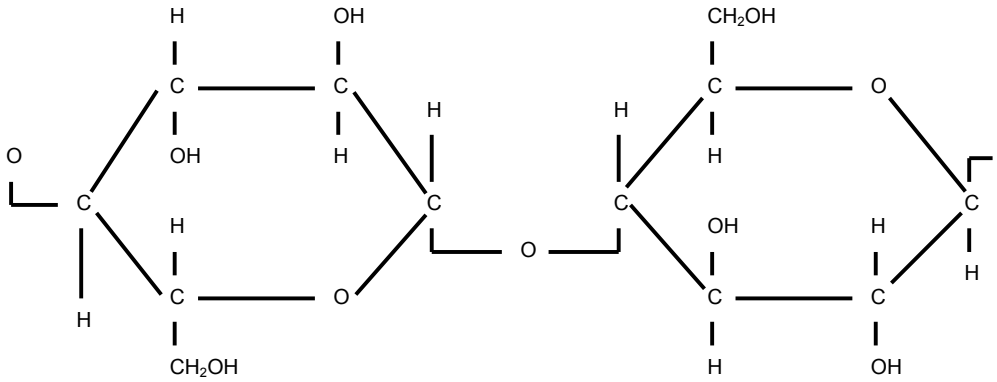


Figure 1.6: The removal of an H_2O molecule to link two glucose molecules into a cellulose chain (adapted from Bowyer et al., 2003).

Cellulose chains form from anhydroglucose (glucose attached to a mononucleotide) during cell wall thickening. The chains lengthen and aggregate into crystalline networks which merge to form microfibrils (Bowyer *et al*., 2003). It is thought that 36 chains form a microfibril (Doblin *et al*., 2002). These microfibrils are the main structural element of most plant cell walls (Williamson *et al*., 2002). They provide strength and flexibility to plant tissues (Peng *et al*., 2002), and serve as a scaffold for other wall components to bind to (Lerouxel *et al*., 2006).

Other sugars apart from glucose are produced during photosynthesis, including galactose and mannose (six-carbon sugars), and xylose and arabinose (five-

carbon sugars). Lower molecular weight (than cellulose) polysaccharides called hemicelluloses are synthesised from these sugars (Bowyer *et al*., 2003). The chains formed are shorter in length than cellulose chains, and tend to be branched polymers. Hemicelluloses can be classified into four groups: xyloglucans (XyG), glucomannans, glucuronoarabinoxylans (GAX) and mixed-linkage glucans (MLG); the proportion of sugars which constitutes them varies, and all hemicelluloses except MLG contain sidechains (Lerouxel *et al*., 2006). Celluose microfibrils link with hemicelluoses to produce a cross-linked matrix (Lerouxel *et al*., 2006).

Pectin is a class of macromolecule (rich in galacturonic acid) found within the cellulose-hemicellulose matrix (Willats *et al*., 2001), and also in the middle lamella between primary cell walls. It is rarely found in secondary cell walls. Pectin functions as an adhesive between cells, controls cell wall porosity and allows extension of the cellulosic matrix (Willats *et al*., 2001).

1.7.1.3 Lignification and Programmed Cell Death

Lignin is found in plant cells that have secondary cell walls, such as tracheids and vessel elements (Rogers and Campbell, 2004). It is deposited in the walls of cells after division and expansion. It provides rigidity and strength to the plant (and is therefore important for determining wood properties), and makes the cell walls impermeable to water (Whetten and Sederoff, 1995). Lignin is also found between plant cells; where its function is to bind adjacent cells together. It is possible that lignin also has a role in plant defence by adding to a wood's toxicity (Bowyer *et al*., 2003); it has also been shown that lignification can be induced through wounding or disease (Whetten and Sederoff, 1995).

Lignin is a chemical compound derived from the amino acid phenylalanine. It is formed by the removal of water from sugars, creating aromatic compounds. Different lignins are created through the polymerisation of three monolignols; p-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol (see figure 1.7). The three

monolignols give rise to p-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) type lignin respectively (Rogers and Campbell, 2004).

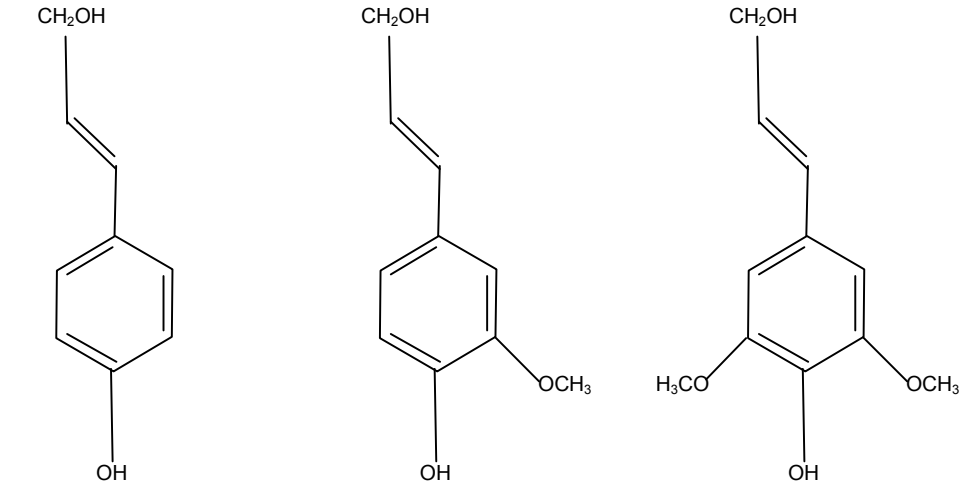


Figure 1.7: Three monolignols that are polymerised to form lignin.

Three biosynthetic pathways are required to produce these monolignols; shikimic acid metabolism (which produces the amino acid phenylalanine), phenylpropanoid metabolism (synthesises hydroxycinnamoyl-CoA-thioesters used in a two-step reduction), and finally monolignol biosynthesis by phenyl substitution (Rogers and Campbell, 2004).

Phenylalanine
PAL
Cinnamic acid
C4H
p-Coumaric acid
4CL
CST
C3T
p-Coumaroyl-CoA
CCR
p-Coumaraldehyde
CAD
p-Coumaryl alcohol (H)
Coumaroyl-Shikimate
Coumaroyl-Quinate
C3'H
Caffeoyl-Shikimate
Caffeoyl-Quinate
CST
C3T
Caffeoyl-CoA
COMT
CCoAOMT
Feruloyl-CoA
CCR
Coniferaldehyde
F5H/C4d5H
5-Hydroxyconiferaldehyde
COMT/4d0MT
Sinapaldehyde
CAD
Coniferyl alcohol (G)
SAD
Sinapyl alcohol (S)
UDP-glucose glucosyltransferase
B-glucosidase
Peroxidase/Laccase
LIGNINS

Figure 1.8: pathway depicting lignin biosynthesis (adapted from Rogers and Campbell, 2004). The enzymes listed are: phenylalanine ammonia-lyase (PAL), cinnamic acid 4-hydroxylase (C4H), hydroxycinnamate: CoA ligase (4CL), hydroxycinnamoyl-CoA reductase (CCR), hydroxycinnamyl alcohol dehydrogenase (CAD), hydroxycinnamoyl CoA:quinate hydroxycinnamoyltransferase (CQT), hydroxycinnamoyl CoA:shikimate hydroxylcinnamoyltransferase (CST), p-coumarate 3-hydroxylase (C3H), caffeoyl-CoA O-methyltransferase (CCoAOMT), coniferyl aldehyde 5-hydroxylase/ferulate 5-hydroxylase (F5H), caffeic acid O-methyltransferase (COMT) and sinapyl alcohol dehydrogenase (SAD).

H-type lignins (derived from *p*-coumaryl alcohol) are found in grasses and the xylem cells of reaction wood. G-type lignins (derived from coniferyl alcohol) are found in gymnosperms and angiosperms. S-type lignins (derived from sinapyl alcohol) are only found in angiosperms (Rogers and Campbell, 2004).

After ligninfication, the xylem elements undergo programmed cell death (Plomion *et al*., 2001), which is controlled by auxins and cytokinins. The cellular contents are destroyed and the vacuoles collapse. The secondary cell wall remains intact.

1.7.2 Wood quality – commercially important traits

1.7.2.1 Wood density and growth rate

Wood is a complex composite material that can contain varying amounts of water absorbed within the fibre. Density is the weight of a substance per unit volume, in this case the weight of the dry wood per volume measurement (kg/m^3). A tree with a high density has less air within the wood than a tree with a low density (Lee, 1997). Basic density varies greatly within and between species, being strongly influenced by geographic location, site fertility, age and genotype. It can also be influenced by silviculture (Macdonald and Hubert, 2002).

Wood density is widely regarded as an important indicator of general wood quality. High wood density has been shown to be strongly linked to superior

strength, stiffness, hardness and working properties of sawn timber, as well as wood pulp yield and paper-making quality.
Wood density varies between juvenile and mature wood. Juvenile wood is 'the secondary xylem at the centre of a tree formed throughout the life of the tree' (Bowyer *et al*., 2003). Juvenile wood is considered to extend from the pith outward to between the 12th or 18th annual ring (Thompson, 1992), and is less dense than the mature wood that is produced later on in the life of the tree. Improving the density of juvenile wood is one factor that can be addressed in tree improvement programmes. Wood density also varies within a growth season; earlywood (produced at the start of the growth season) has a lower density than latewood.

Breeding for increased wood density may incur significant penalties through negative associations with other traits. Lee *et al*. (2002) found the correlation between early wood density measurements (annual ring group 6-9) and later diameter measurements in Sitka spruce to be approximately r_A = - 0.70, indicating that it will be difficult to improve density and diameter concurrently. Genetic parameters for wood, fibre, stem quality and growth traits have been estimated in 11-year-old birch (*Betula pendula* Roth) and the negative correlation between density and diameter (r_G) was calculated to be - 0.53 (Stener and Hedenberg, 2003). This decrease in density is caused by the early wood width within a ring increasing without a corresponding increase in late wood (Savill and Sanders, 1983). Despite this negative genotypic correlation between diameter and wood density, only a small reduction in wood density was observed when diameter was used as the selected trait (Stener and Hedenberg, 2003), indicating that it may be possible to select for one of these traits without incurring a negative impact on the other.

The effects of selecting for both growth and wood quality on phenology and growth of seedlings have been assessed in lodgepole pine (Wang *et al*., 2003). Four subpopulations were assessed: 1) fast height growth and high wood

density; 2) slow height growth and high density; 3) fast height growth and low density; and 4) slow height growth and low density. The population undergoing selection for fast growth and high density exhibited higher xylem conductivity than the slower growing populations, and increased resistance to xylem cavitation. Wang *et al*. (2003) suggest that it should be possible to select families that combine traits for fast growth, high density and drought tolerance.

1.7.2.2 Microfibril angle

Microfibril angle (MFA) is the angle between parallel cellulose chains (in the tracheid cell wall) and the fibre axis (Vainio *et al*., 2002), and is one characteristic that contributes to the strength of wood. The MFA of a tree is measured in the S2 (second) layer of the secondary cell wall as the angle between cellulose fibrils and the longitudinal cell axis (Lichtenegger *et al*., 1999), using X-ray scattering and optical microscopy (Sarén *et al*., 2004). The S2 layer forms the main portion of the cell wall (Wang *et al*., 2001). MFA plays a significant function in the mechanical properties of wood, i.e. adaptation to wind and gravitational forces (Sarén *et al*., 2004). Microfibril angle (in Norway spruce) gets smaller the further from the pith the measurement is taken, and becomes constant in mature wood (Sarén *et al*., 2004). Therefore, it is possible to obtain several different MFA measurements within an individual depending upon the location in the tree, as juvenile measurements of MFA will vary from mature wood measurements. Measurements in coniferous species gave a MFA range of 37º to 55º for juvenile wood and between 7º and 20º for mature wood (Wang *et al*., 2001). Trees with a small MFA are more prone to wind damage, and the stem can break more easily (Vainio *et al*., 2002). However, it has been suggested that having a small MFA is the most efficient way for a tree to carry its weight (Vainio *et al*., 2002).

Microfibril angle variation has been examined between provenances of Irish grown Sitka spruce, and also the relationship between microfibril angle and wood strength and stiffness has been studied (Treacy *et al*., 2000). The results showed that microfibril angle is strongly correlated with wood stiffness, and Treacy *et a*l.

(2000) suggest that in cases where the objective of improvement programmes is to provide stronger timber, then trees should be selected on the basis of microfibril angle rather than wood density.

1.7.2.3 Spiral Grain

Spiral grain refers to the orientation of the tracheids in a tree stem (Hannrup *et al*., 2002), and it is formed during the process of cell maturation within the vascular cambium (Schulgasser and Witztum, 2007). The grain orientation can be to either the left or the right. Juvenile wood tends to have a higher grain angle than mature wood. In two trials with 19-year old Norway spruce, Hannrup *et al*. (2002) showed that grain angle to the left increases to a maximum at annual ring 4, followed by a decrease in grain angle.

Spiral grain has been correlated with twist in dried timber in Sitka spruce (Hansen and Roulund, 1998). Deviation of the tracheid orientation from the vertical axis has been linked to warping and a reduction in strength (Hannrup *et al*., 2002). It is therefore extremely disadvantageous to the production of sawn timber. Grain angle is affected by both environmental and genetic factors (Forsberg and Warensjö, 2001). Severe twist in boards leads to their rejection or downgrading at the sawmill. The level of twist exhibited by the wood can be linked to its distance from the pith. When comparing two types of wood with the same grain angle, wood formed closer to the pith will twist more than wood formed further away (Hannrup *et al*., 2002).

Fortunately, spiral grain shows low levels of correlations with other traits, and no or little genotype x environment interaction. It has been suggested that spiral grain in juvenile wood can be reduced by 0.5-1º (mean spiral grain in the two trials being 3.8 and 4.1º) while concurrently improving stem straightness and diameter in Sitka spruce (Hansen and Roulund, 1998).

1.7.2.4 Earlywood / latewood

Cambium is dormant during the winter. It becomes active again in springtime, when cell division resumes. Conditions for growth are most favourable at the start of the growing season, so growth through cell division and expansion is rapid. The cells formed at this time tend to have large lumens and thinner cell walls (Plomion *et al*., 2001). This wood is referred to as earlywood (or springwood).

Conditions for growth become less favourable towards the end of the growing season (due to drought, less daylight, etc.), so cambial activity slows down. The cells produced at this time have smaller lumens and thicker cell walls. This wood is referred to as latewood (or summerwood). An example of the cells produced during the growth season is shown in figure 1.9. Latewood has a higher density than earlywood due to the smaller cells and the thicker cell walls. It has been suggested that water deficit induces latewood formation, and the slower rate of cell division results in more resources being available for cell wall synthesis (Bowyer *et al*., 2003).

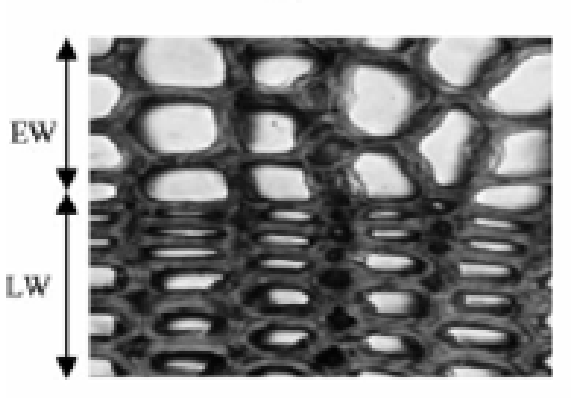


Figure 1.9: early and latewood cells (taken from Plomion et al., 2001)

1.7.2.5 Juvenile / maturewood

Juvenile wood is formed during a period of fast growth at the beginning of the life of a tree. This type of wood is said to occur for the first 12-18 annual rings produced by the tree from the pith (Thompson, 1992), though other papers suggest shorter time spans, such as 12-15 annual rings (Lee, 1997). After this

period of growth, the tree will switch from producing juvenile wood and start to produce maturewood.

1.7.2.6 Reaction wood

Reaction wood is formed in response to environmental influences, i.e. when a tree leans to one side. In gymnosperms reaction wood is referred to as compression wood; and it is formed on the underside of a leaning tree (the side that is subjected to compression stress).

The properties of compression wood are very different to those of 'normal' wood. The tracheids in compression wood are shorter, cellulose content is lower and lignin content is increased (Bowyer *et al*., 2003). In terms of desirability for timber, compression wood has a higher level of longitudinal shrinkage compared to normal wood, causing the wood to bend and warp. When comparing compression wood and normal wood of the same density, normal wood is stronger, though compression wood generally has a higher density.

Tension wood is formed on the upper (or tension) side of a leaning stem/branch. It is higher in cellulose content that normal wood, and is 5-10% greater in density (Bowyer *et al*., 2003), although it is not as strong.

Other factors that are commercially important and affect wood quality include knots, tracheid length, tree/log size, stem straightness and stem taper. Stem straightness and form have been selected since the start of the GB Sitka spruce breeding programme (Lee, 1999a). Selection for tracheid length (in *Pinus sylvestris*) has given rise to a 3% increase in juvenile and mature wood (Hannrup *et al*., 2000). Fibre size also has an effect on growth traits. There is also evidence in this species of a positive correlation between tree height and fibre size, so improving growth could produce longer and thicker fibres (Ericsson and Fries, 2004). In Sitka spruce these factors can be affected by silvicultural methods including rotation length, spacing, respacing and thinning, nursing mixtures,

pruning, weed suppression and fertiliser use (Macdonald and Hubert, 2002). Timber quality has been defined in terms of stability, strength, straightness and stiffness (Hannrup *et al*., 2004), so traits affecting these properties will be of most interest to tree breeders.

All the above factors have an effect on wood quality. Timber quality can be measured in terms of strength. Strength is categorised using the European Standard EN 14081, and can be graded either visually or by machine. This grading takes into account the species, geographical origin, dimensional requirements and requirements for use (BS EN 14081). Measurements obtained during grading include bending strength, tension strength, compression strength, shear strength, modulus of elasticity and density. Strength reducing characteristics include knots, slope of grain, density, rate of growth and warp (linked to moisture content).

As there are so many different factors that determine wood quality, it is important to address those that will have the most effect on the technological properties of the timber (Brazier, 1967). Breeding for improved wood characteristics would be a sensible choice as the Sitka spruce plantations in Britain 'contain a high proportion of vigorous, healthy and well-formed trees (Brazier, 1967). There is little point in selecting for improved physical characteristics when these characteristics are satisfactory across much of a plantation.

1.8 Breeding

1.8.1 Traditional breeding methods

Plant domestication started approximately 10,000 years ago, however, tree improvement (also termed domestication) through the use of conventional breeding is a much more recent discipline (Campbell *et al*., 2003), and it focuses on the most commercially important species.

The GB Sitka spruce breeding programme commenced in 1963. The major goal of the breeding programme is to develop a range of populations suited to various site-types, with improved wood traits beneficial for timber production (Lee, 1997). The programme started with the selection of 'plus trees' growing in Britain, individuals that are superior in height, diameter and form. These trees were thought to originate from the Queen Charlotte Islands (QCI), British Columbia (Lee, 1999a). The collection of trees selected is called the 'base population'. Wind-pollinated seed were collected from all the plus trees, and the progeny were assessed for height, diameter and form (tested in progeny trials at three contrasting sites). The progeny from a single mother tree are called a 'half-sib family', as the seed only have the female parent in common. Assessment of the progeny involves comparing the families with trees grown from a standard QCI seedlot (in comparative trials relative to a QCI control).

Parents (maternal) are allocated a breeding value based on the performance of their progeny. The progeny must be assessed to ensure that the superior traits exhibited by the mothers are heritable. The plus trees re-selected after progeny assessment constitute the breeding population.

Grafts of the selected plus trees (based on tested heritabilities) can be planted to produce a tested clonal seed orchard, which can be managed to produce seed commercially. Different traits can be emphasised in different orchards. Plus trees selected to produce specific orchards are called the production population.

Information about the breeding programme in Sitka spruce is available from the Forest Research website http://www.forestresearch.gov.uk/website/forestresearch.nsf/ByUnique/INFD-5WNJBW.

Gains can only be made in a tree improvement programme if there is sufficient genetic variability that can be exploited (i.e. if it is heritable). There are three levels of variability that can be exploited: species level variability; provenance level; and individual tree level (Fletcher, 1992). Costa E Silva *et al*. (1998) estimated clonal correlations for various wood properties in Sitka spruce, and found that phenotypic variation was only low for lignin content, so potential genetic gain may be lower for that trait than for height and diameter selection. Problems with Sitka spruce include: 1) poor mechanical properties, 2) large and frequent knots, 3) warping of the wood, and 4) variability in wood properties (Rozenberg and Cahalan, 1997).

Progress in tree improvement is hampered for several reasons that are related to the biology of the tree species. Most of the commercially important species are obligate outcrossers (selfed crossed result in very low seed numbers), which restricts the use of selfing and backcrossing as aids to breeding. Traits that are economically important to the forestry industry can only be assessed for the identification of a superior individual when a tree is relatively mature (Tzfira *et al*., 1998). The ability to predict age-age genetic correlations can bring forward selection ages in tree improvement programmes (Gwaze *et al*., 2000). Juvenile/mature correlations in Sitka spruce of up to r = 0.87 have been estimated from 6-year height for 27-year volume, 6-year height and 15-year diameter, and 10-year height with 22-year diameter (Lee, 1992). Gill (1987) showed correlations between three-year height in Sitka spruce and later measures of growth, and suggested that withholding progeny tests to later years will give little further advantage.

The benefits from using genetically improved planting stock, which have been available from the early 1990s, have been quantified (Lee, 1999b). Gains in height and stem form vary from 8% and 2% respectively at ten years of age, increasing to 15% for stem form at fifteen years of age (Lee, 1999b). Gains in diameter at fifteen years of age are estimated to be 20% (with no drop in wood density).

Breeding programmes have been set up in all the areas where Sitka spruce is a commercially important species (OECD, 2002). Despite the fact that Sitka spruce is native to British Columbia, its use as a plantation species is limited due to risk of attack by the white pine weevil (*Pissodes strobi*), so their breeding program emphasises resistance to weevil attack. Breeding programmes are also being carried out in Denmark and northern France, where Sitka spruce is an important plantation species.

The use of age-age genetic correlations means that selection ages can be decreased, but even earlier methods of identifying superior individuals would be beneficial for accelerating the improvement process. It may be possible to use genetic markers to identify superior genotypes in much younger material than would be possible with traditional breeding approaches.

1.8.2 Genetic methods (genes, genomics, genetics)

Screening populations for superior genotypes can take many years when using traditional breeding methods due to the reasons outlined above. The exploitation of genetic diversity in breeding programmes can be used as a mechanism for early selection of elite genotypes (Boerjan, 2005). Using a functional genomics approach to identify candidate genes offers promise to speed up the process of selection.

In order to use the candidate gene approach in the process of selection within a population, it is necessary to draw up a list of desirable (candidate) genes that

are thought to affect the important traits of interest (in this case, wood density). The use of microarray experiments to determine candidate genes have been widely reported. Paux *et al*. (2004) studied wood formation in *Eucalyptus* by comparing gene expression in leaf and xylem material on an array. This study identified genes that are preferentially expressed in xylem; including those involved in cell wall biosynthesis and remodelling (Paux *et al*., 2004).

Hertzberg *et al*. (2001) sampled 30µm thick sections from the wood-forming zone in poplar to identify genes involved in wood formation. RNA from different developmental stages of xylogenesis were hybridised to an array to identify differentially expressed genes (Hertzberg *et al*., 2001). Genes encoding lignin and cellulose biosynthetic enzymes were found to be under strict developmental stage-specific transcriptional regulation (Hertzberg *et al*., 2001).

The table below (table 1.2) lists several studies that have used microarray technology to study xylogenesis.

Author	**Species**	**Tissue type**	**Results**
Hertzberg *et al*., 2001	Poplar	Xylem	Identified potential regulators of xylogenesis
Kirst *et al*., 2004	Eucalyptus	Xylem	Lignin negatively correlated with growth
Paux *et al*., 2004	Eucalyptus	Xylem and leaves	Identified genes up-regulated in xylem compared to leaves
Schrader et al., 2004a	Poplar	Cambium	Identified potential regulators of cambial stem cell identity
Schrader et al., 2004b	Poplar	Cambium	Identified genes that showed differential expression between active and dormant samples
Yang *et al*., 2004	*Pinus taeda*	Xylem and needles	Identified genes up-regulated in xylem compared to needles

Table 1.2: Studies using microarray technology to identify genes with a role in wood formation

Prior to the widespread use of microarrays, other techniques were used to study wood formation. Allona *et al*. (1998) created a cDNA library from the xylem of loblolly pine undergoing gravitational stress (*Pinus taeda*). RNA was extracted from the compression wood xylem to enable library formation, which was followed by sequencing (Allona *et al*., 1998). Of the 1,097 sequences obtained, 55% showed similarity to previously described sequences, and approximately 10% of these are thought to be involved with cell wall formation (Allona *et al*., 1998).

In a more recent study, a combination of microarray technology and QTL (quantitative trait locus) mapping has been used, which is referred to as genetical genomics (Boerjan, 2005). QTL mapping allows for the detection of chromosomal regions that affect the trait (or traits) of interest (Pot *et al*., 2006). Kirst *et al*. (2004) measured transcript abundance for 2,608 genes (thought to be involved in wood formation) in Eucalyptus, in association with their diameter measurements. The faster-growing trees showed a reduction in transcript levels for genes involved in lignin biosynthesis (when compared to the slower-growing progeny), and an increase in transcript levels for cell-wall associated genes (Kirst *et al*., 2004). eQTLs (QTLs combined with gene expression data) were identified for all of the previously identified lignin genes that correlated with growth (Kirst *et al*., 2004).

These studies illustrate rapid ways of identifying candidate genes, for which a population could then be screened to identify individuals with high levels of expression of specific genes. This would be a much quicker method of selection than assessing traits when the trees are mature.

1.9 Targets for genetic manipulation

Crop domestication can focus on a wide range of traits, so it would be beneficial to manipulate those traits that could lead to largest improvement, such as wood yield or wood density (Campbell *et al*., 2003). Genetic manipulation can be used to bypass the long generation times that hinder traditional tree breeding (Boerjan, 2005). Much of the effort into genetic engineering has been directed towards modifying lignin (Boerjan, 2005).

1.9.1 Lignin

The function of lignin within a tree (to provide rigidity and strength) makes it an important target for genetic modification. The control of lignin biosynthesis is a key step in wood formation, as the formation of a lignified secondary cell wall prevents further cell expansion.

Figure 1.8 depicts the lignin biosynthesis pathway. At the start of the pathway, ammonia is removed from phenylalanine to produce cinnamate (Whetten and Sederoff, 1995). The reaction is catalysed by the enzyme phenylalanine ammonia-lyase (PAL). Manipulation of the genes encoding this enzyme to control lignin biosynthesis would be inappropriate, as PAL activity is necessary for producing a range of phenylpropanoid compounds (Rogers and Campbell, 2004).

4CL is the third enzyme in the pathway (figure 1.8), and its activity generates the phenolic precursors for lignin and flavonoid biosynthesis (Hu *et al*., 1999). The expression of different 4CL enzymes may control the abundance of the lignin precursors, the cinnamyl alcohols (Whetten and Sederoff, 1995). Hu *et al*. (1999) produced transgenic aspen with a down-regulated 4CL gene. This resulted in up to a 45% decrease in lignin, which was compensated for by a 15% increase in cellulose (Hu *et al*., 1999). The down-regulation of 4CL expression was not associated with any negative side-effects on growth (Hu *et al*., 1999), as leaf, root, and stem growth were increased. The purpose of reducing lignin content in

wood would be to make pulping for the paper industry easier. Table 1.3 lists other studies using genetic modification of genes involved in lignin biosynthesis.

Author	Species	Gene	Result
Hu *et al*., 1999	Poplar	4CL, down-regulated	Decrease in lignin content, increase in cellulose content
Li *et al*., 2003	Poplar	4CL and CAld5H	S/G ligninratio increase accelerated cell maturation
Meyermans *et al*., 2000	Poplar	CCoAOMT, down-regulated	Decrease in lignin content, increase in S/G lignin ratio
Pilate *et al*., 2002	Poplar	CAD and COMT, down-regulated	Improved pulping characteristics
Piquemal *et al*., 1998	Tobacco	CCR, down-regulated	Increase in S/G lignin ratio
Patzlaff *et al*., 2003	*Pinus taeda*	*Pt*MYB, over-expression	Increase in lignin deposition

Table 1.3: Studies of the genetic modification of the lignin

Another step in the lignin biosynthesis pathway that has received attention is the reduction of cinnamaldehydes to cinnamyl alcohols, the precursors of lignin. This step, the last in the pathway, is catalysed by cinnamyl alcohol dehydrogenase (CAD). Various attempts have been made to manipulate CAD activity. MacKay *et al*. (1997) discovered a mutant loblolly pine in which the expression of the gene encoding CAD was lower than normal, and the recessive allele causing this phenotype maps to the same region as the CAD locus (MacKay *et al*., 1997). This reduction led to the modification of the lignin composition (through increased levels of coniferaldehyde) but only slight decrease in lignin content (MacKay *et al*., 1997).

The manipulation of lignin content and structure affects cellulose content (Li *et al*., 2003), the amount of lignin deposition (Patzlaff *et al*., 2003), as well as other growth properties, and may be an important factor in determining wood density.

1.9.2 Auxin

Auxins are one of the major five classes of plant hormones. The other classes are cytokinins, gibberellins, abscisic acid and ethylene. Studies assessing the affect of phytohormones on growth have focused on auxins and gibberellins. Auxins are plant hormones involved in development; they are required for plant growth. Auxin is involved in many plant processes, including the growth of roots and shoots, responses to light and gravity, and cell differentiation. Indole-3-acetic acid (IAA) was the first auxin to be isolated from plants, and is recognised as the key auxin (Woodward and Bartel, 2005).

There are two ways in which IAA is produced by plants; Trp-dependent biosynthesis and Trp-independent biosynthesis. Several pathways may be available for each type of synthesis; none of these pathways are fully understood. A table of genes thought to be involved in IAA biosynthesis from *Arabidopsis* is listed by Woodward and Bartel (2005).

Uggla *et al*. (1996) reported the existence of a concentration gradient of the endogenous auxin, indole-3-acetic acid, across the vascular cambium in *Pinus sylvestris* trees. This auxin gradient could have a role in positional signalling, with the amount of IAA present at a particular site affecting the development of the cambial cells. This auxin gradient was confirmed in a study by Moyle *et al*. (2002), who showed that the expression of IAA genes varies along a developmental gradient of cambial cells in hybrid aspen.

1.9.3 Gibberellins

Gibberellins (GAs) are plant hormones that stimulate growth in shoots and leaves. They are also involved in flowering (the breaking of bud dormancy), and seed germination. More than 100 different gibberellins have been identified.

GAs are synthesised through the isoprenoid pathway from mevalonic acid (Eriksson *et al*., 2000).The levels of GA are regulated by gibberellin biosynthesis

genes, one example of which is GA 20-oxidase. It has also been suggested that end-product repression is involved with the gene regulation (Eriksson *et al*, 2000).

GA induces stem growth through enhanced cell division and cell elongation. Cell expansion can only occur after loosening of the cell wall (to allow for increased water uptake). GA activity may be limited to young cells due to the orientation of their microfibrils. Auxin has the ability to reorient the microfibril deposition to a transverse direction, so this hormone can promote elongation in cells that have already stopped growing (Kende and Zeevaart, 1997).

Eriksson *et al.* (2000) have reported improved growth rate and biomass in hybrid aspen due to the over expression of a GA 20-oxidase. This gene catalyses the conversion of the C_{20} gibberellins to the precursors (GA_9/GA_{20}) of the active gibberellins (GA_4/GA_1). The trees that over expressed this gene showed faster growth, larger leaves, an increased number of longer xylem fibres, and increased biomass (Eriksson *et al*., 2000).

1.9.4 Cell wall

The cell wall is formed during cell differentiation. Its role is to maintain cell shape and to provide structural support to the cell. The cell wall prevents cell expansion when water enters the cell, resulting in an increase in turgor pressure. The middle lamella is the first layer of the cell wall to be formed after cell division has occurred. It is located between the cells, and its role is to ensure the adhesion of a cell with the cells adjacent to it (Plomion *et al*., 2001). The primary cell wall then starts to develop, and this occurs at the beginning of cell differentiation. The primary cell wall is rich in pectins (Bowyer *et al*., 2003). The cell wall (both primary and secondary) consists of several layers of cellulose microfibrils (MFs). The microfibrils are held together by hydrogen bonds, and they are coated with hemicelluloses whose function is to link the MFs into a network. The last layer to form is the secondary cell wall, this starts to develop when the maturing cell has

finished expanding. The secondary cell wall is composed of three layers; the S_1, S_2 and S_3 (see figure 1.10). The organisation of the cellulose microfibrils is different in each layer. The S2 is the thickest of the three layers, and is considered to be the most important in terms of mechanical strength (Plomion *et al*., 2001).

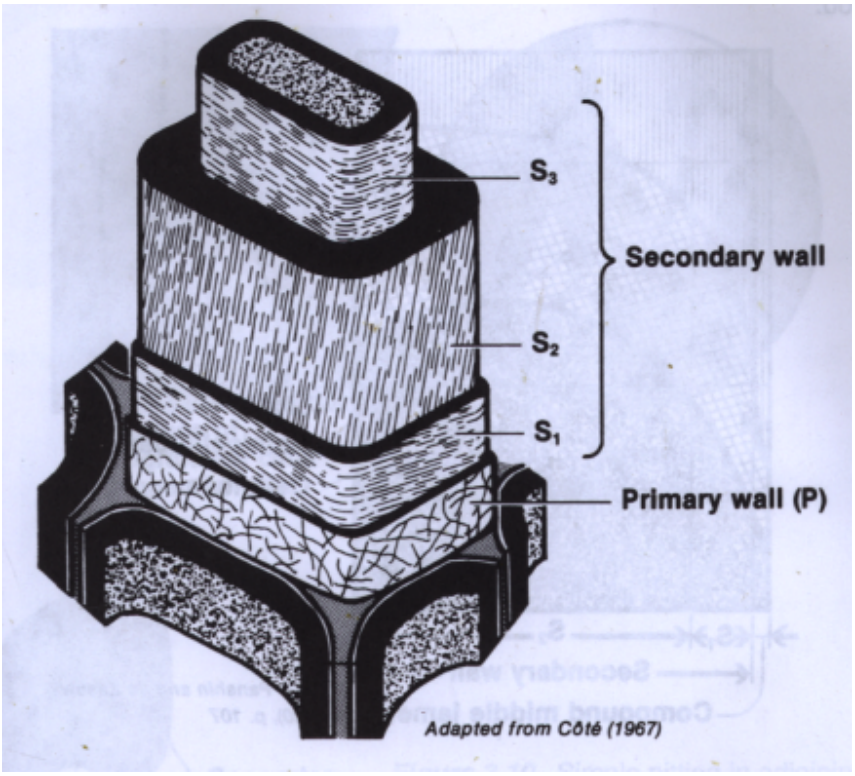


Figure 1.10: Representation of a mature cell wall (Bowyer et al., 2003).

Plant growth depends upon cell growth, which is the combined result of cell elongation and cell division. Plant cells usually increase in volume by 10- to 1000-fold through cell wall expansion and vacuolation (Cosgrove, 1997). For cell elongation to occur, the cell wall must 'loosen', causing a decrease in turgor pressure, so that more water can be taken up by the cell. For cell wall extension to occur, a minimum level of turgor pressure (or 'stress') is required. The rate of extension is dependant upon the amount by which the turgor pressure exceeds this minimum level of stress (Cosgrove, 1997). This relationship between growth and turgor is referred to as wall extensibility (Cosgrove, 1997).

It may be possible to screen plants for high expression levels of genes involved in cell wall synthesis. Bao *et al*. (1992) identified a pine extension-like protein (PELP) from differentiating xylem and mature wood of loblolly pine (*Pinus taeda*). The results show that PELP is present in the secondary cell walls of earlywood, but is not as prevalent in latewood (Bao *et al*., 1992).

Aspeborg *et al*. (2005) identified genes involved in cell wall biosynthesis using microarray expression profiling in hybrid aspen. Twenty five glycosyltransferases were identified, of which four were cellulose synthase genes (Aspeborg *et al*., 2005).

Paux *et al*. (2004) identified differentially expressed genes in eucalyptus by comparing secondary xylem and leaves on a microarray. Sixteen genes were classified with the function of cell wall biogenesis, including a cellulose synthase, xyloglucan endotransglycosylase, caffeoyl Co-A O-methyltransferase and an arabinogalactan protein (Paux *et al*., 2004).

Rather than identifying genes for the purpose of genetic modification, genomic techniques can be used to determine candidate genes involved in different aspects of wood formation, for example cell wall synthesis, lignin and cellulose synthesis, cell growth and division, initiation of dormancy, etc. A population can then be screened for these desirable genes, which can be used as selection criteria for a breeding programme.

1.10 Methods of Analysing Gene Expression

There are many different methods available for analysing gene expression, including polymerase chain reaction (PCR), serial analysis of gene expression (SAGE), cDNA-amplified fragment length polymorphism analysis (cDNA-AFLP), restriction fragment-coupled differential display (RC4D), differential display reverse transcription PCR (DDRT-PCR) and screening of cDNA libraries. Oligonucleotide and DNA-fragment based microarrays are also used to analyse gene expression.

Differential display reverse transcription PCR (DDRT-PCR) was developed by Liang and Pardee (1992) to determine transcript patterns. Restriction fragment length polymorphism-coupled domain directed differential display (RC4D) was developed by Fischer *et al* (1995), and is a combination of DDRT-PCR and amplified fragment length polymorphism (AFLP). This method is used to detect differentially expressed genes from individual families. Bachem *et al* (1996) developed the cDNA-AFLP technique, where cDNA restriction fragments are amplified using PCR. Serial analysis of gene expression (SAGE) is used to detect sequence abundance (Velculescu *et al*., 1995). iAFLP was described by Kawamoto *et al*. (1999), and is used to measure the concentration of known transcripts.

1.10.1 Serial Analysis of Gene Expression (SAGE)

SAGE is used to determine the relative levels of messenger RNA (mRNA) molecules within a population (Pleasance *et al*., 2003). Short specific tags are produced for each mRNA, producing a SAGE library that is representative of the samples under investigation (Fizames *et al*., 2004). The SAGE protocol assumes that a tag is able to identify a unique transcript. The abundance of a transcript in the mRNA population is measured by observing how many times the tag sequence is present. The use of SAGE does not require expensive equipment (Fizames *et al*., 2004), and the throughput is comparable to that of microarray technology (Kuhn, 2001). The original SAGE protocol requires large amounts of

mRNA (Peters *et al*., 1999), which can be a limitation depending on the species and tissue being studied. However, modifications to the original protocol have been made so that SAGE can be made using very small amounts of mRNA (Peters *et al*., 1999; Datson *et al*., 1999). The major disadvantage of SAGE is the use of EST clusters or available cDNA sequences to assign tags to genes, so that many of the transcripts cannot be identified. This means that many of the tags in the SAGE libraries will not be assigned to a gene and are left unidentified. Only 23.1% of the tags produced for rice could be identified from EST and cDNA sequences in databases (Matsumura *et al*., 1999). For *Arabidopsis* leaves, 27.9% of the SAGE tags were matched to EST or cDNA databases (Jung *et al*., 2003) and for Arabidopsis pollen, 45% (Lee and Lee, 2003). Large collections of EST or cDNA sequences are required to enable the identification of the transcripts, so the use of this method to identify differentially expressed genes in spruce could be difficult.

1.10.2 Differential display reverse transcription PCR (DDRT-PCR)

This technique was developed to identify differentially expressed mRNA species, and then to clone the corresponding cDNAs (Kuhn, 2001). mRNA is reverse transcribed, and the cDNA is then amplified (using PCR). The products are labelled during the reaction and these can then be separated on a DNA sequencing gel and visualised autoradiographically (Kuhn, 2001). The PCR fragments that are produced are of varying lengths.

This technique is very simple to use and has been applied to many different species. Unfortunately, this technique also generates an unacceptable number of false-positive results, suggesting that genes are differentially expressed when in fact they are not. Another method for verification of the results is required, so a large amount of time, labour and mRNA is necessary. It has been suggested that cDNA-AFLP is a more acceptable method than DDRT-PCR (Habu *et al*., 1997).

1.10.3 RFLP-coupled domain-directed differential display (RC4D)

This protocol is based on reverse transcriptase PCR, and is used to detect differentially expressed genes. Fischer *et al*. (1995) suggested that it is not necessary to detect all differences in gene expression, and that sometimes it may be possible to detect differences only within a certain gene family. RC4D takes advantage of family-specific domains (FSDs, used to define membership of different gene families). Instead of arbitrary decameric primers, longer primers are directed at the FSD, increasing priming efficiency and reproducibility (Fischer *et al*., 1995).

The technique of RC4D combines cDNA-AFLP with DDRT-PCR (Kuhn, 2001). cDNAs that are members of the same gene family will be selectively amplified because of the use of primers directed against an FSD (Kuhn, 2001). The products of amplification will all be of a similar length, so RFLP is introduced to reduce the size of the cDNAs so that they can be separated on a gel. RC4D is more reproducible than DDRT-PCR because of the use of the longer primers. One drawback occurs if the gene has a restriction site within the FSD, as the primer binding site will be lost during restriction.

1.10.4 Introduced amplified fragment length polymorphism (iAFLP)

iAFLP is used to measure the concentrations of known transcripts in numerous different probes (Kuhn, 2001), and the method is based on PCR. cDNAs from each probe are restricted, then selectively amplified with a primer specific to the gene of interest (Kuhn, 2001).

iAFLP is comparable to microarray technology in terms of throughput and accuracy (Kawamoto et al., 1999; Kuhn, 2001). Sequence information is required to use this technique as a primer binding site is required (Kuhn, 2001).

1.10.5 cDNA-AFLP

This technique involves three steps. The first is the restriction of cDNA followed by the ligation of oligonucleotide adapters. Secondly, restriction fragments are selectively amplified using PCR primers, and thirdly, the amplified fragments are analysed on a gel.

The occurrence of mismatches is rare unless transcript abundance is very high, so this procedure produces very few false-positives. One drawback is the need for appropriate restriction sites (Kuhn, 2001), so it is necessary to repeat the procedure with various enzymes to ensure that every cDNA has been visualised.

1.10.6 Screening of cDNA libraries

This is the oldest method used to isolate differentially expressed cDNAs. For this procedure to be accurate and efficient, the mRNA must comprise of more than 0.05% of the total mRNA in one cell and less than 0.01% in the other (Kuhn, 2001). Subtractive hybridisation (SH) can be used to increase the concentration of specific sequences in a cDNA probe, which improves the sensitivity of the reaction.

The construction of a representative library is very laborious and renders the whole process as time inefficient in terms of time and effort, especially if looking for rare mRNAs. Also, the technique generates a large percentage of false-positives.

If a particular study requires quantitative measurements of gene expression, and no sequence information is available for the species of interest (as is the case with *Picea sitchensis*), then cDNA microarray technology offers an appropriate option.

1.11 Problems

1.11.1 Lack of a 'model' conifer species

Researchers working on conifer species do not have the same resources available to them as those working on poplar trees. The entire genome in poplar has been sequenced, and this confers on it the status of a 'model' species. Poplar has a relatively small genome (450-550 Mbp), and a large number of genetic maps are available (Taylor, 2002). It may therefore be beneficial to do some comparison work between Sitka spruce and poplar to make use of these resources.

It is possible to study secondary growth in *Arabidopsis thaliana*, as it forms vascular cambium and secondary xylem (Nieminen *et al*., 2004). However, the small physical size of this species makes it difficult to extract sufficient amounts of woody tissue for microarray analysis (Mijnsbrugge *et al*., 2000), so it may not be the ideal choice for genetic studies. Using a true woody system for the analysis of wood formation will cover any developmental processes that are unique to trees (Taylor, 2002).

1.11.2 Wood quality of 'fast-growth' wood species

Producing trees that grow faster so that they can be harvested at an earlier age has its own problems. Wood produced during the first twelve to fifteen years (approximately) of a tree's life is referred to as juvenile wood (Lee, 1997). The quality of juvenile wood tends to be much lower (in terms of strength) than that of mature wood. Plantation species that are grown quickly and harvested earlier will have a much higher proportion of juvenile wood to mature wood, leading to a lower quality end product if used for timber.

1.12 Aims

The aim of this PhD is to study gene expression in the wood-forming tissues of Sitka spruce and poplar that show variation in terms of their wood density. This is with the objective of identifying the key genes which are involved in determining the complex trait of high wood density. A clonal trial of Sitka spruce will be used for gene expression analysis, and a *P. trichocarpa* x *P. deltoides* F_2 family for the cross-species comparison. These contrasting species were used to determine whether the same compliment of genes are up-regulated in high wood density individuals in different species.

In order to select an appropriate sample of trees from which to base our experiments, it will be necessary to collect measurements to characterise the differences in physiology. Once this selection has been made, experiments can be designed to determine how wood density is genetically controlled. This will be achieved through the use of microarray technology to study differential gene expression.

2 Materials and Methods

The details for the general methods and materials used in this thesis are contained in this chapter. Methods which only relate to individual chapters will be found in a materials and methods section within the specific chapter.

2.1 Physiological measurements

2.1.1 Indirect assessment of wood density - Pilodyn device

The Pilodyn provided an indirect, fast and non-destructive assessment of wood density. It was originally developed in Switzerland to determine the extent of soft rot in telephone poles (Hansen, 2000). When triggered, the spring-loaded Pilodyn drives a steel pin into the tree with a fixed force each time, and the depth of penetration is inversely proportional to the density of the wood. The Pilodyn only provides an estimate of relative wood density, not a measurement of actual wood density. The pin penetrates to a maximum of 40 mm, so the measurement provided is an estimate of the density of the outer rings of the tree only.

2.1.2 Xylem scrapes – Sitka spruce

It was important that a suitable protocol be developed for extracting cambium/xylem from mature spruce trees, as the RNA had to be kept as clean as possible for future microarray work. Tissue extracted from a xylem scrape must be removed as swiftly as possible and snap-frozen immediately in order to limit RNA degradation, and to ensure that it is suitable for further analysis. To achieve this, samples were snap frozen in liquid nitrogen and then transported in a cryoshipper.

A fresh pair of latex gloves was worn for the removal of each sample to prevent cross-contamination. The bark was cut away at breast height (1.3 m) from an approximately sized 5 cm (vertical) x 2 cm (horizontal) section of tree using a sharp knife. The knife was then cleaned with ethanol and dried with clean tissue

after every sample to prevent contamination. Using a fresh razor blade for each sample, the exposed xylem was excised (figure 2.1), wrapped in aluminium foil, sealed with an adhesive label and snap-frozen in liquid nitrogen. On return to the laboratory the samples were then stored in a deep freeze at -80° C.

Figure 2.1: Removal of xylem tissue

Samples were collected at a height of 1.3 m for every tree, as all the trees are approximately the same height. This height coincides with the location of where the pilodyn measurement was taken.

The trees were sampled at three time points during the growing season. The first sampling took place from the 4th-7th of May 2004. 300 samples were collected during this period, 2 samples from each of the 150 trees selected. The second

sampling took place from the 13th-15th of July 2004 and the third sampling from the 7th-9th of September 2004. For each of these samplings, 150 samples were collected, with one sample from each tree. For each sampling date, samples were taken from the same compass orientation on each tree.

2.1.3 Xylem scrapes – Poplar

A similar technique to the cambial scrapes taken in spruce was used for collecting poplar cambium. A fresh pair of gloves was worn for every sample to prevent contamination. The top 50 cm of the leader stem was removed using secateurs, and the entire layer of thin bark was then peeled off. Using a fresh razor blade for each sample, the xylem was scraped off, wrapped in aluminium foil and frozen in liquid nitrogen.

2.2 Molecular biology methods

2.2.1 RNA extraction

The RNA extraction protocol used was a modified version of Chang *et al*. (1993). The modified protocol was developed by Nathaniel Street and James Tucker (pers.comm.) using poplar leaves. The CTAB extraction buffer was warmed in 50 ml falcon tubes at 65º C. The ground xylem tissue was added to the falcon tube along with 400 µl of 2-ME and vortexed. The suspension was incubated for 5-10 min at 65º C. 15ml of CHISAM was then added.

The tubes were centrifuged for 20 minutes at room temperature in a swing-out rotor at 4500 rpm. The upper phase was transferred to a new 50 ml falcon tube with 15ml of CHISAM and re-centrifuged. The upper phase (approx 12 ml) was transferred to a JA-20 tube (Oakridge) with 0.25 volume of lithium chloride, and left overnight on ice to precipitate in a coldroom at 4º C.

The tubes were centrifuged at 10,000 rpm at 4º C for 30 minutes in a Beckman JA-20 rotor. Meanwhile, the SSTE was warmed to 60º C. The supernatant was removed, and the pellet re-suspended in 700 µl SSTE. This was then transferred to a 2 ml Eppendorf tube and incubated for a few minutes at 60º C. 700 µl of CHISAM was added, and the tubes centrifuged at 10,000 rpm at room temperature for 10 minutes. The upper phase was transferred to a new tube with 700 µl of CHISAM and re-centrifuged. The upper phase (approx 600 µl) was transferred to a new tube with 1.2 ml of 99.8% EtOH and left to precipitate at -70º C for one hour. The tubes were centrifuged at 13,000 rpm for 30 minutes at 4º C. The pellet was washed twice with 1 ml of 70% EtOH by centrifuging at 10,000 rpm for two minutes. The pellet was air dried at room temperature for a few minutes, then re-suspended in 100 µl of RNase free MQ H_2O.

2.2.2 RNA quantification

The RNA was quantified using a NanoDrop® ND-1000 Spectrophotometer, using 1 µl for each sample. The sample was pulled into a column between fibre optic ends, and a pulsed xenon flash lamp provided the light source. 260/230 ratios of 1.8-2.2 were commonly observed. Ratios within this range indicated that pure RNA had been isolated from the samples. Ratios from outside this range indicated that contamination or RNA degradation may be a problem.

A small section of the samples were analysed using the Agilent 2100 bioanalyser 'lab on a chip'. The program produces an electropherogram summary and gel picture for each sample loaded onto the chip.

2.2.3 Microarray hybridisation

Hybridisation for all spruce arrays was performed using the Treenomix Spruce cDNA Microarray Guide Version 2.0 (Ralph and Yueh, 2004), on a 21.8 K spruce microarray. The samples were reverse transcribed with Cy3 and Cy5 dye primers. The probe was left to hybridise for 16 hours in a 60 º C water bath. Slides were then washed, dried, and the dendrimers allowed to hybridise for 3 hours, also in a 60ºC water bath. The slides were then washed and dried ready for scanning. A full description of the method used is provided in section 4.2.4.

The slides for the first Sitka experiment were scanned using a Scan Array Express Packard Bioscience scanner. The laser power is fixed at 90%. A full description of the method is provided in chapter four. The slides for experiments two and three were scanned using a Genetix scanner on default settings.

2.3 Solutions and buffers

2.3.1 Chisam

Chloroform : isoamyl alcohol, 24:1 (192 ml : 8 ml).

2.3.2 CTAB extraction buffer

2% (10 g) CTAB, 2% (10 g) PVP, 100 mM (50ml@1M) Tris-HCl (pH 8.0), 25 mM (25ml@0.5M) EDTA, 2M (58.57 g) NaCl. Add MQ H_2O to make up to 500 ml. Autoclave.

2.3.3 0.5M EDTA

Dissolve 186.1 g $Na_2EDTA\text{-}2H_2O$ in 700 ml MQ H_2O, then use NaOH to adjust the pH to 8.0. Add more MQ H_20 to make the solution up to 1L.

2.3.4 70% EtOH

35 ml 100% EtOH, 15 ml MQ H_2O.

2.3.5 10M LiCl

Dissolve 42.39 g of LiCl in 100 ml of MQ H_2O.

2.3.6 5M NaOH

Dissolve 20 g of NaOH (MW 40.0) in 80 ml of MQ H_2O. Add more MQ H_2O to make the solution up to 100 ml.

2.3.7 10% SDS

Dissolve 100 g of SDS (MW 288.38) in 900ml of MQ H_2O. Adjust to pH 7.2 with HCl. Add more MQ H_2O to make the solution up to 1 L.

2.3.8 3.3M Sodium acetate

Dissolve 449.1 g of sodium acetate·3H_2O (MW 136.08) in 800 ml of MQ H_2O. Adjust to pH 5.2 using glacial acetic acid. Add more MQ H_2O to make the solution up to 1 L.

2.3.9 20x SSC

Dissolve 175.3 g of NaCl (MW 58.44) and 88.2 g of sodium citrate (MW 294.1) in 800 ml of MQ H_2O. Adjust to pH 7.0 with NaOH. Add more MQ H2O to make the solution up to 1 L.

2.3.10 2x SSC / 0.2% SDS

Mix 100 ml of 20x SSC, 20 ml of 10% SDS and 880 ml of MQ H_2O.

2.3.11 SSTE

1M (11.7 g) NaCl, 0.5% (10ml@10%) SDS, 10 mM (2ml@1M) Tris-HCl (pH 8.0), 1 mM (400µl@0.5M) EDTA. Add MQ H_2O to make the solution up to 200 ml.

2.3.12 TE pH 8.0

Mix 10 ml of 1 M Tris (pH 8.0), 2 ml EDTA (pH 8.0) and 988 ml MQ H_2O. Autoclave.

2.3.13 1M Tris-HCl

Dissolve 24.2 g Tris in 150 ml MQ H2O, then use NaOH to adjust the pH to 8.0. Add more MQ H_2O to make the solution up to 200 ml.

3 Phenotypic Analysis of Wood Quality In Sitka spruce

3.1 Introduction

The aim of this project is to identify the genes involved in the determination of wood quality, so the first step was to select an appropriate set of trees to analyse that showed differences in wood quality. Wood quality is an ambiguous term, and can be defined in many ways, taking into account physical and chemical properties, aesthetic properties, ease of harvesting, etc. For the purpose of this project wood quality will be defined as 'the resultant of physical and chemical characteristics possessed by a tree or a part of a tree that enable it to meet the property requirements for different end products' (Mitchell, 1961). The trait selected for further analysis is wood density.

Wood quality is a complex, multi-component trait. Several measurements may therefore be indicators of wood quality. These include growth measurements (height and diameter), density (direct and indirect measurements), and lignin and cellulose content. Analysis of gene expression from trees showing different levels of growth and wood density may help to identify important genes involved in wood formation. The identity and level of expression of the genes involved in the process of wood formation will ultimately determine the physical properties of the wood produced (Kirst *et al*., 2004).

To look at wood quality, first it is necessary to identify trees that show physiological and chemical differences in their wood properties, i.e. differences in density, lignin and cellulose concentration. The work described in this thesis involved collaboration between Southampton University and Forest Research, an agency of the Forestry Commission. Forest Research has a long history of breeding improved Sitka spruce and has been responsible for setting up a wide

range of Sitka trials across Britain. Therefore several experimental populations were available for use through Forest Research. A large collection of Sitka spruce trees were available in Newcastleton (OS grid reference NY506881, latitude = 55.1847, longitude = -2.77892) in the Scottish Borders. The collection consisted of two clonal trials (experiments 35 and 36), and a diallel cross experiment (experiment 37). Experiments 35 and 36 were established from cuttings taken from trees belonging to 6 unrelated full-sib families (8 genotypes per family). There are 50 genotypes (48 experimental and two lots of one control genotype) in each experiment, represented by 15 replicates of each genotype to provide a final total of 750 trees per experiment.

We decided to work on a sub-sample of the clonal trees from experiment 35 for this project for the following reasons. The experiment consisted of a large number of trees with replication of all genotypes across the site in a randomised single plant plot design. This would allow us to determine how variable the wood properties were within and between genotypes. Diameter and indirect density measurements (using the pilodyn, see methods below) had already been collected for all 750 trees by Forest Research, at 12 years of age (Steve Lee, pers.comm). Also, in a previous project working on the same set of trees, direct wood densities, cellulose and lignin measurements had been collected by the Scottish Crop Research Institute (SCRI) (Nuopponen et al., 2006). These measurements were based on 8mm wood cores taken at breast height from 50 clones of 15 years of age during the winter season of 2003/2004. Ten replicates of each clone were collected by Forest Research and processed by the Scottish Crop Research Institute (SCRI). For the current work, these direct density and growth (diameter) measurements of the individuals were examined. Basic density measurements were collected using the water displacement method (see Lee, 1997). Lee (1997) demonstrated a negative relationship between growth rate and wood density. For our study we were interested in studying genotypes which combined average growth rate with extremes of either high or low wood density. In order to select which genotypes to study those with below average growth rate

were discarded and, from the remainder, the five clones at each extreme of the normal distribution for wood density were selected. For our study, we therefore selected 10 replicates each of 5 high and 5 low density genotypes from the 50 genotypes available to provide 100 trees in total. We based our selection criteria on clones rather than individual trees because wood density has been shown to be a highly heritable trait with a large genetic component (Lee *et al*., 2002).

3.2 Results

Wood density is a complex trait and we used the data available from the SCRI experiment to determine the range of wood density in the clonal trial and to explore the relationship between wood density and lignin and cellulose content. These two substances are important constituents of the secondary cell wall and may be key determinants of wood density. The samples taken for the SCRI work consisted of a core taken at breast height through the entire trunk. Values therefore reflect an average of the wood across all the growth rings in a given tree.

Figure 3.1 shows the spread of wood density measurements for each of the 48 experimental genotypes from experiment 35. Each genotype has data for 10 out of its 15 replicates, so there are 480 measured trees in total.

Boxplot of density measurements for each clone from experiment 35

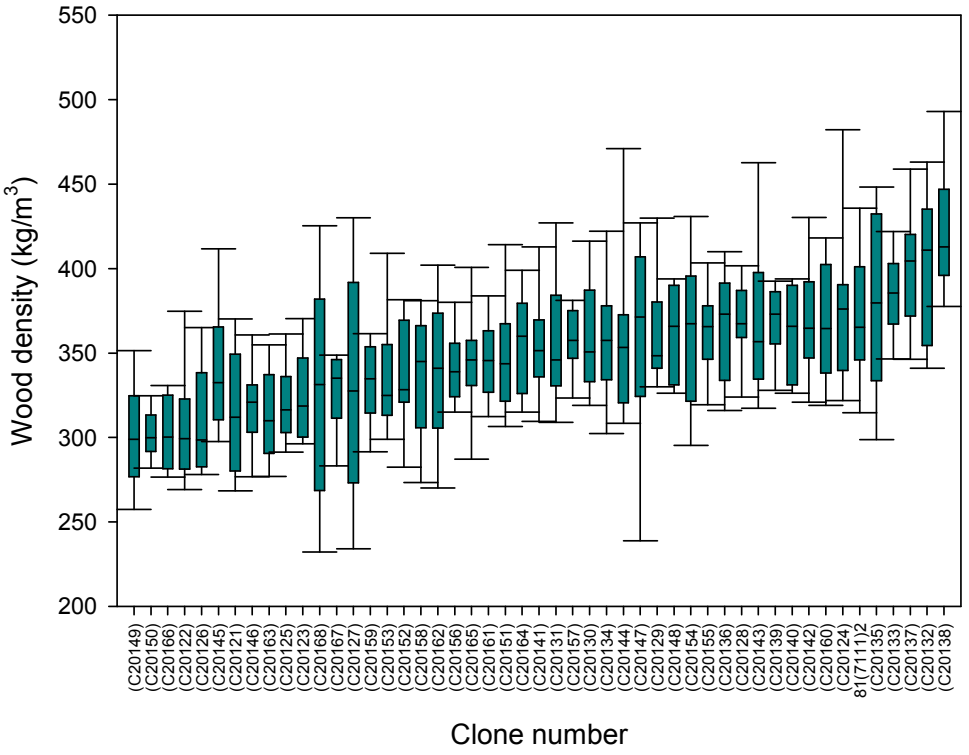


Figure 3.1: density measurements (collected by the SCRI) for all clones from experiment 35 at Newcastleton. Each boxplot shows the median, the 5th and 95th percentile. The bars show the lowest and highest extremes of density for each clone. Each clone has ten measured replicates out of fifteen total replicates.

The spread of wood density measurements is large across all of the clones (figure 3.1). Some of the clones have a very broad spread of wood density measurements (e.g. clones 27 and 68), thereby demonstrating that they are sensitive to their growing environment. As the site conditions were relatively uniform this may reflect the effect of competition with more or less vigorous neighbouring trees. Other clones have a much narrower spread of wood densities (e.g. clones 33 and 50), demonstrating greater stability in terms of wood density. The clone with the largest range of 199 kg/m^3 is C20127. The clone with the smallest range of 44 kg/m^3 is C20150.

The direct method used by SCRI to measure wood density is time consuming and Forest Research routinely use a pilodyn method which is much simpler, faster and can be taken from outside the tree without the need for coring. The Pilodyn measurements were collected by Forest Research at the same time that the wood cores were collected. Pilodyn measurements provide indirect measures of wood density, so these should show a negative correlation with density. Figure 3.2 shows the wood density for each individual graphed against its pilodyn measurement. As the pilodyn measure reflects the depth of penetration of a spring loaded pin, unlike average direct wood density measurements based on specific gravity of a whole wood core, the pilodyn provides an indirect measure of the density of only the bark and outer rings of wood.

Scatterplot for density against pilodyn for all measured trees from experiment 35

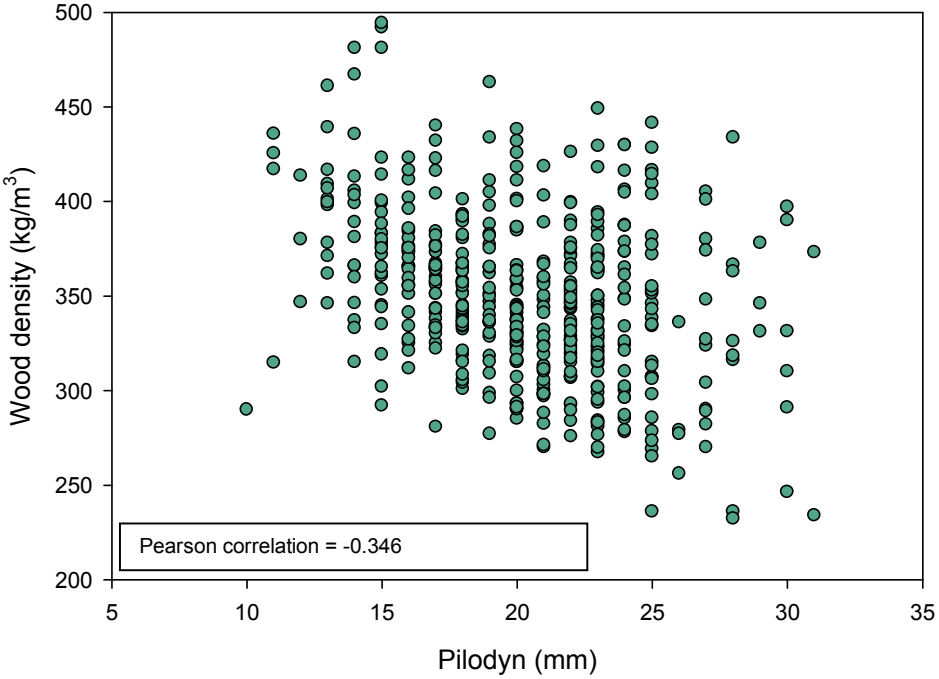


Figure 3.2: density against pilodyn measurements for all clones from experiment 35 at Newcastleton. There is a slight negative correlation (-0.346) between wood density and pilodyn penetration.

Figure 3.2 gives a Pearson correlation between density and pilodyn of -0.346. The correlation value normally achieved between density and pilodyn is approximately -0.6 (Lee, 1997).

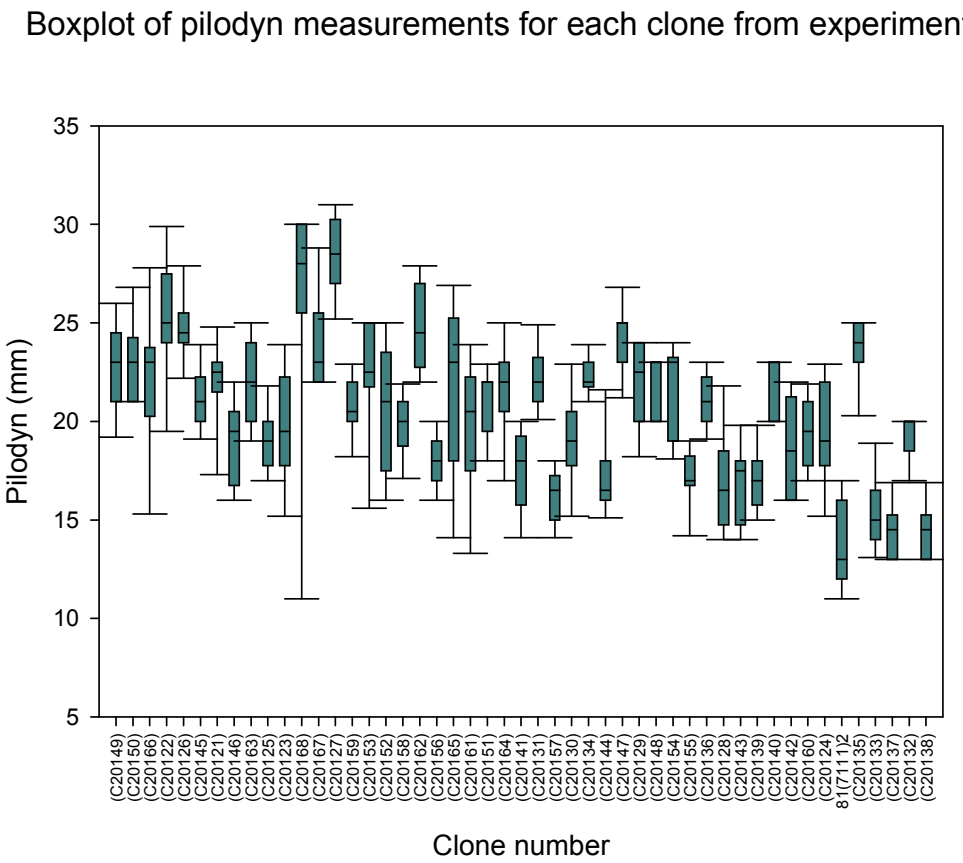


Figure 3.3: pilodyn measurements for all clones from experiment 35 at Newcastleton. Each boxplot shows the median, the 5th and 95th percentile. The bars show the lowest and highest extremes of density for each clone. Each clone has ten measured replicates out of fifteen total replicates.

Figure 3.3 shows the spread of pilodyn measurements for each clone, with the clones plotted in the same order as for figure 3.1 (lowest direct density on the left to highest direct density on the right). The low density clones to the left of the graph tend to have higher pilodyn measurements, although the correlation is weak (see figure 3.2). These pilodyn measurements cannot be used to predict density.

Figure 3.4 graphs the relationship between wood density and pilodyn penetration at an average clonal level.

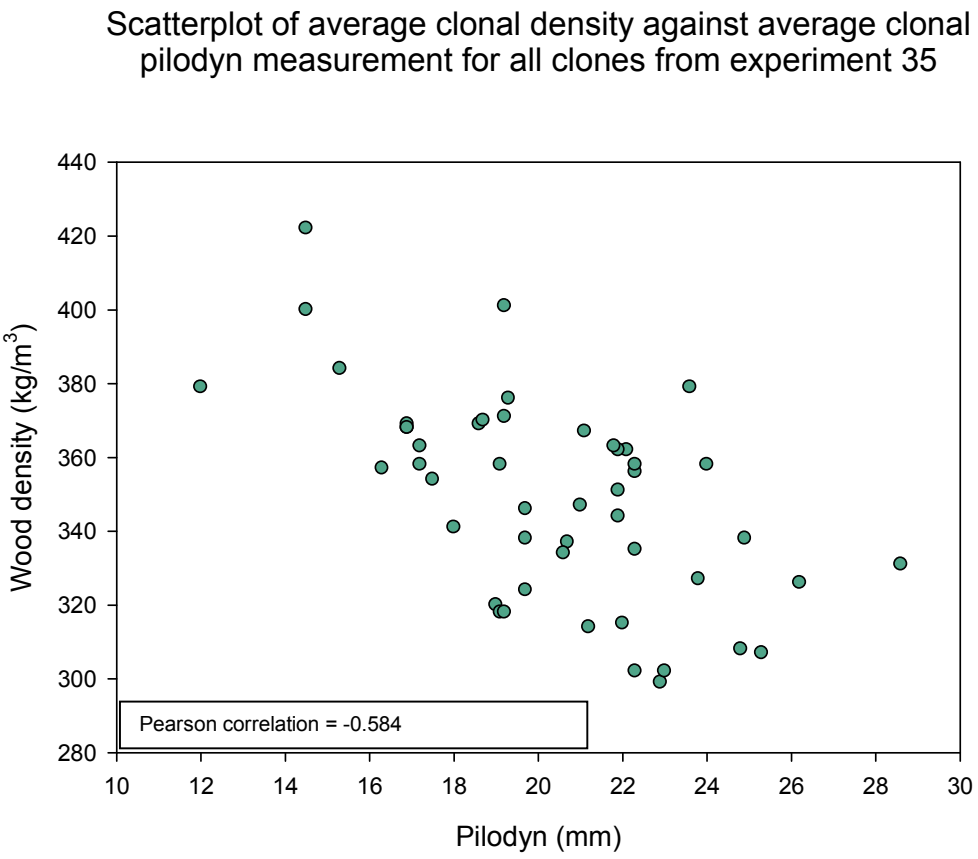


Figure 3.4: average density against average pilodyn measurements for each clone from experiment 35. The clonal mean level correlation between density and pilodyn is -0.584.

There is a stronger correlation (-0.584) between pilodyn penetration and wood density when the measurements are assessed as a clonal mean level (figure 3.4).

Figure 3.5 shows the relationship between density and diameter measurements for the trees from experiment 35.

Scatterplot for density against diameter for all measured trees from experiment 35

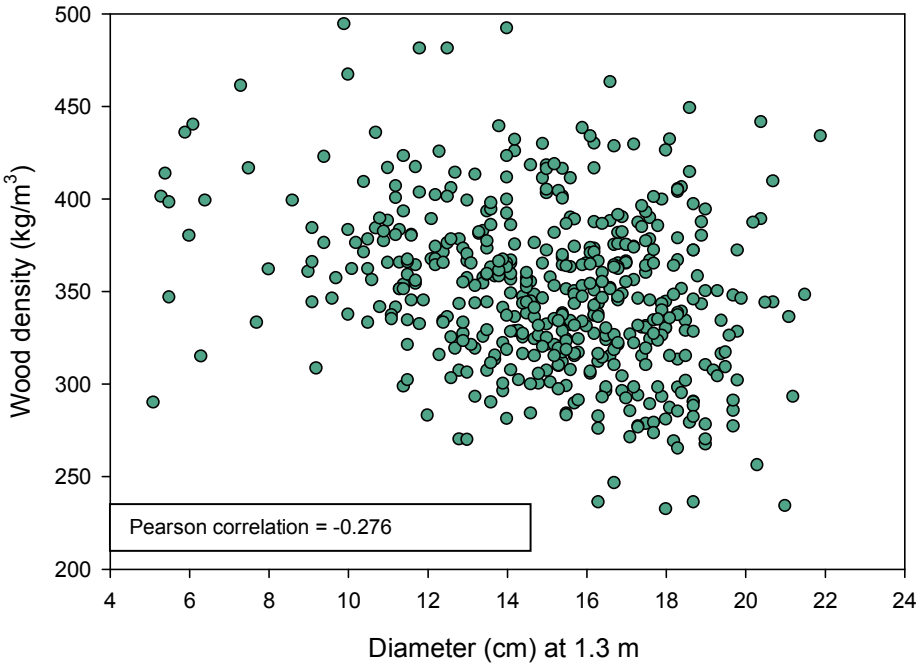


Figure 3.5: Direct density against diameter measurements for all clones from experiment 35 at Newcastleton. There is a slight negative correlation (-0.276) between wood density and tree diameter.

The Pearson correlation coefficient between density and diameter is -0.276. This negative relationship between density and diameter must be taken into consideration when selecting clones for further analysis. It is necessary to select both high and low density clones that have similar growth rates (i.e. measures of diameter), so that differences in density cannot be explained by the level of growth (e.g. trees with high density will have small diameters and trees with low density will have large diameters).

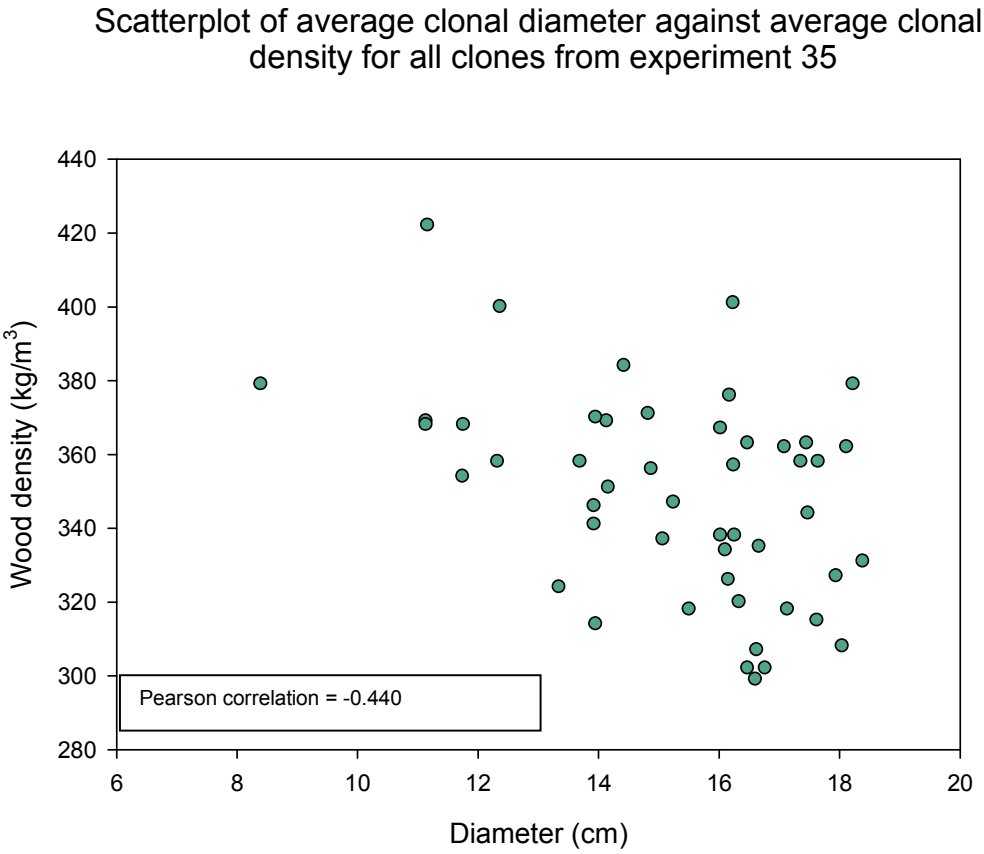


Figure 3.6: Average density against average diameter measurements for each clone from experiment 35. The clonal mean level correlation between density and diameter is -0.440.

There is a stronger correlation (-0.440) between diameter and wood density when the measurements are assessed as a clonal mean level (figure 3.6).

Figure 3.7 graphs the relationship between wood density and lignin content at an individual tree level. Lignin content deceases slightly with an increase in wood density (Pearson correlation = -0.177).

Scatterplot for density against lignin for all measured trees from experiment 35

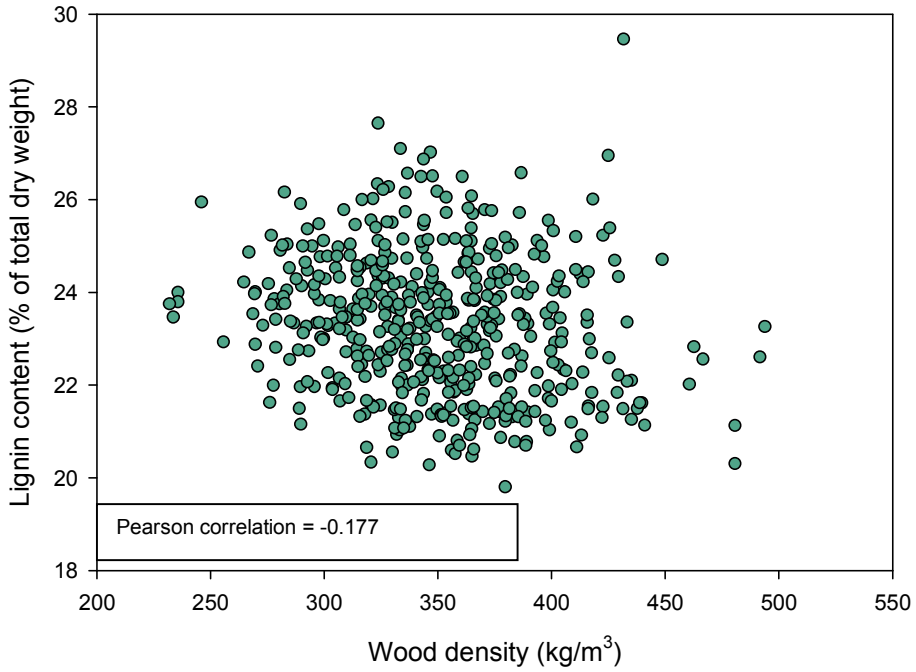


Figure 3.7: Direct density against lignin measurements for all clones from experiment 35 at Newcastleton. There is a slight negative correlation (-0.177) between wood density and lignin content.

Figure 3.8 shows the correlation between wood density and lignin content at a clonal mean level. There is an increase in the negative correlation between lignin content and wood density when the measurements are assessed as a clonal average (figure 3.8).

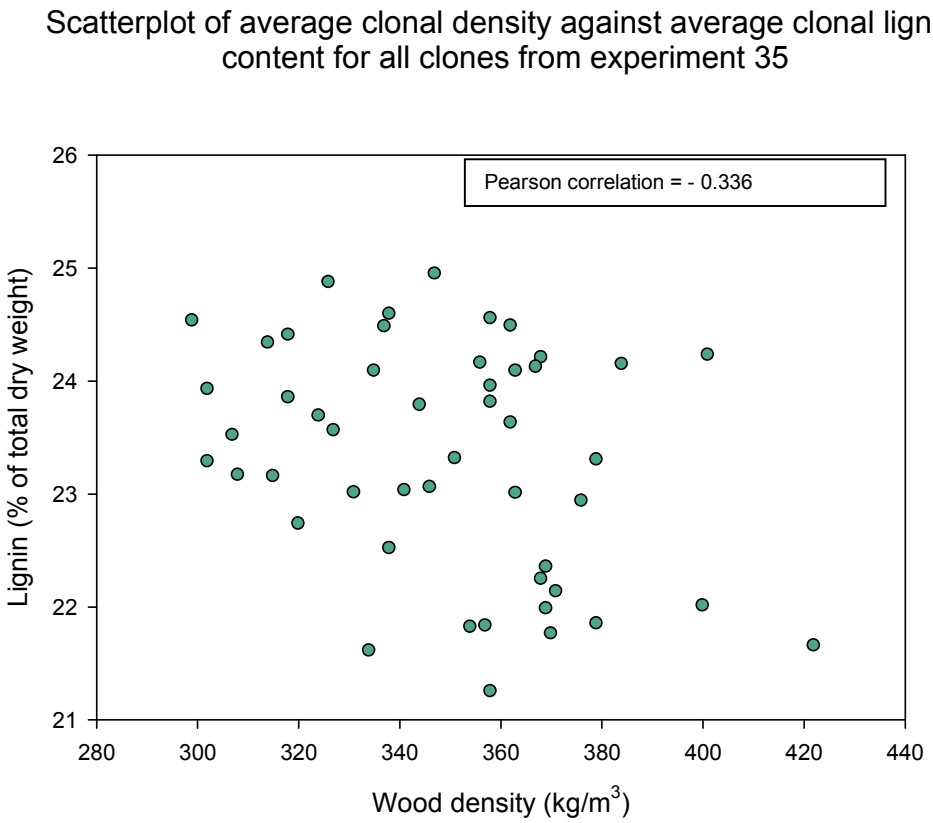


Figure 3.8: average density against average lignin measurements for each clone from experiment 35. The clonal mean level correlation between density and lignin is -0.336.

Scatterplot of density against cellulose for all measured trees from experiment 35

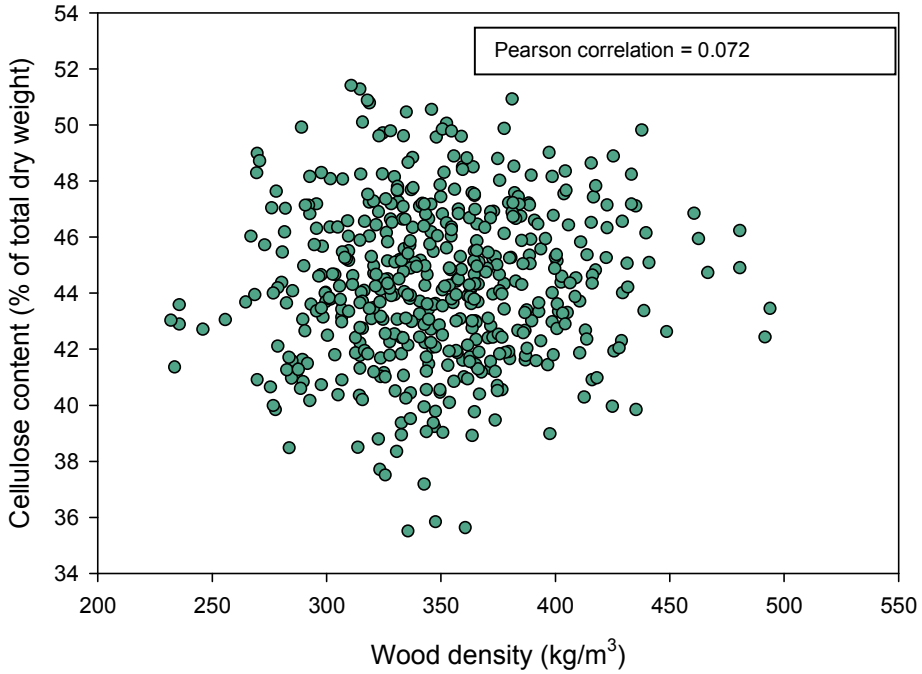


Figure 3.9: density against cellulose measurements for all clones from experiment 35 at Newcastleton. There is a poor correlation (0.072) between wood density and cellulose content.

There is no significant correlation between cellulose content and wood density, either at an individual tree level (figure 3.9) or at a clonal mean level (figure 3.10).

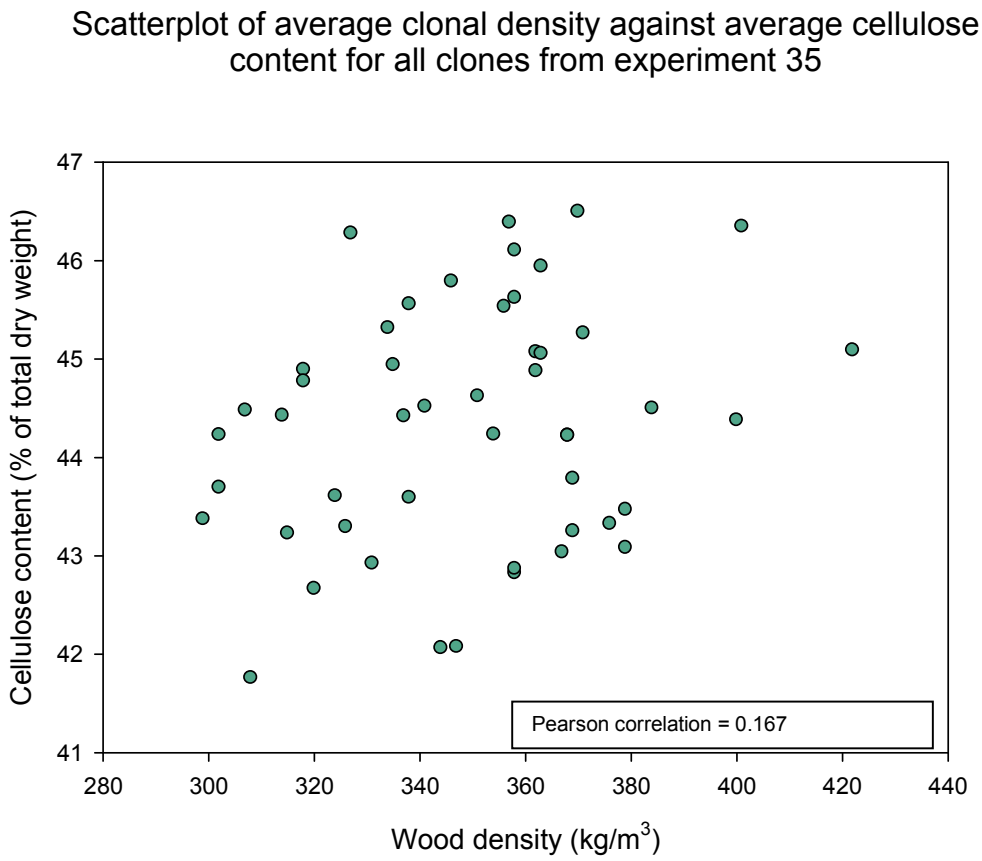


Figure 3.10: average density against average cellulose measurements for each clone from experiment 35. The clonal mean level correlation between density and cellulose is 0.167.

10 out of the 50 clones available were selected for gene expression study. The clones were of either high or low density, and showed average growth rates (indicated by diameter measurements). Lignin, cellulose and pilodyn measurements were not considered when selecting these clones, due to the poor correlation with density.

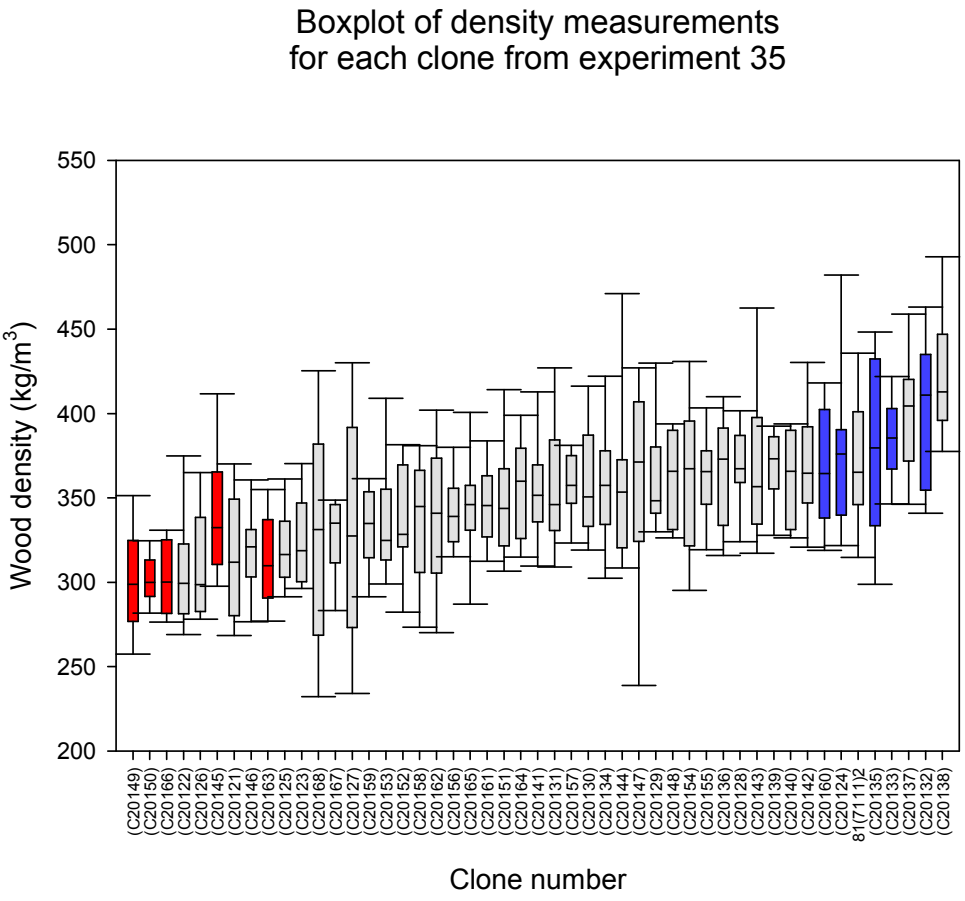


Figure 3.11: Direct density measurements for all clones from experiment 35 at Newcastleton. Each boxplot shows the median, the 5th and 95th percentile. The bars show the lowest and highest extremes of density for each clone. Each clone has ten measured replicates out of fifteen total replicates. The plots highlighted in blue have been selected as high density clones, and the plots highlighted in red selected as low density clones.

Figure 3.11 shows all 48 experimental clones, with the 10 selected clones highlighted. The high density clones are shown in blue, and the low density clones in red.

	Wood density (kg/m^3)	Tree diameter (cm)
All clones	350	15.1
High density clones	382	16.0
Low density clones	307	15.9

Table 3.1: Average density and diameter measurements for all clones from experiment 35. Average wood density is also shown for the selected high density and low density clones. The average diameter (growth rate) is similar for the high density and low density clones.

Table 3.1 shows the average density and diameter measurements for all of the clones combined, the high density clones and the low density clones. Both the high density and the low density clones selected have a similar growth rate (diameter) to all of the clones combined.

The five high density clones selected have an average density of 382kg/m^3. The low density clones selected have an average density of 307kg/m^3. A one-way ANOVA showed these averages to be significantly different ($p < 0.001$).

The high density clones have an average diameter of 16.0cm, and the low density clones have an average diameter of 15.9cm. There is no significant difference between the two averages. The two types of clones (high and low density) have similar growth rates.

Although table 3.1 shows a statistically significant difference between the high and low density clones in terms of density, there is considerable overlap between the low and high density clones when the individual density measurements are taken into account.

Boxplot of density measurements for each selected clone from experiment 35

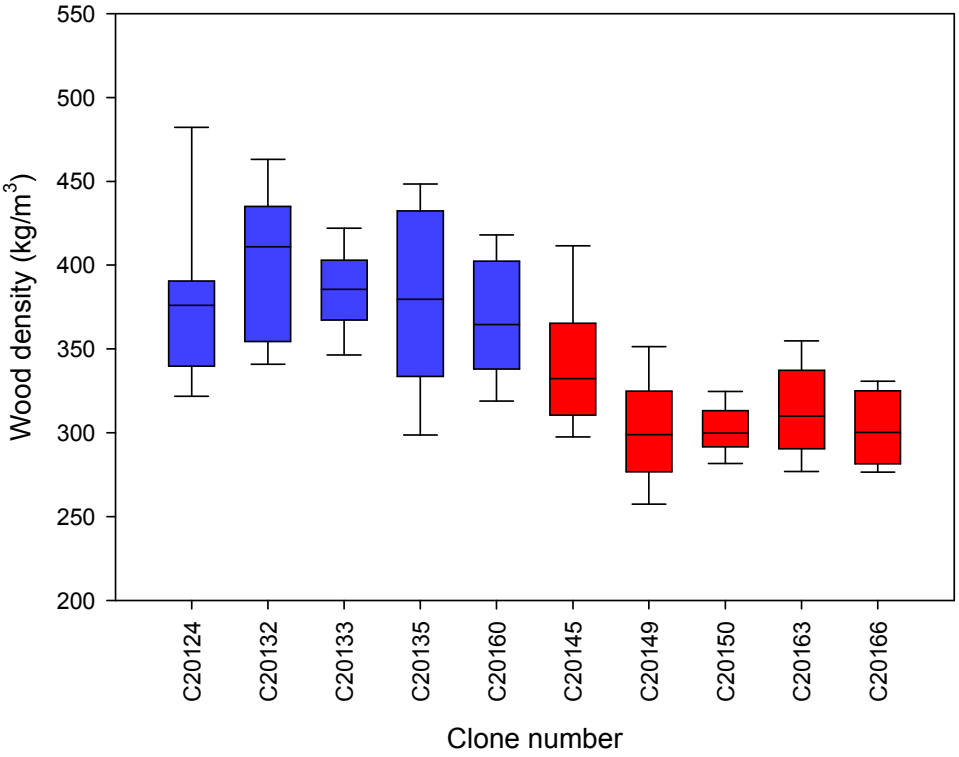


Figure 3.12: Direct density measurements for all selected clones from experiment 35 at Newcastleton. Each boxplot shows the median, the 5th and 95th percentile. The bars show the lowest and highest extremes of density for each clone. Each clone has ten measured replicates out of fifteen total replicates. The plots highlighted in blue have been selected as high density clones, and the plots highlighted in red selected as low density clones.

The boxplot in figure 3.12 shows the density measurements for the 10 selected clones. The blue bars represent the 5 high density clones and the red bars represent the 5 low density clones. All of the high clones show some level of overlap with regard to density measurements with all of the red clones. This overlap between the high and low density clones was unavoidable and demonstrates the problems attached to working with field grown samples. This within clone variability will need to be taken into account when designing the gene expression experiments.

Due to the variation of measurements within a clone, it is important to check whether a tree's position within the site has an effect on its density measurement. A general linear model for the selected clones showed that there was no effect for block, row or column on the density measurements of the clones (p-values of 0.753, 0.977 and 0.570 respectively).

The boxplot in figure 3.13 shows the lignin measurements for each of the 5 high density and 5 low density clones.

Boxplot of lignin measurements for each selected clone from experiment 35

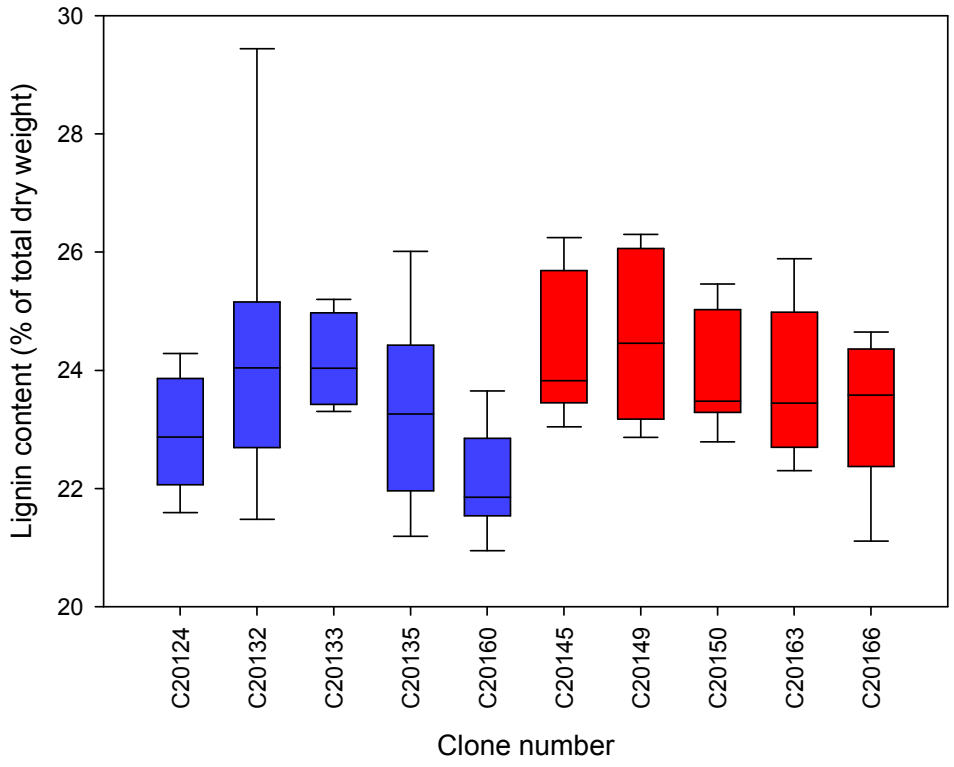


Figure 3.13: lignin measurements for all selected clones from experiment 35 at Newcastleton. Each boxplot shows the median, the 5th and 95th percentile. The bars show the lowest and highest extremes of lignin for each clone. Each clone has ten measured replicates out of fifteen total replicates. The plots highlighted in blue have been selected as high density clones, and the plots highlighted in red selected as low density clones.

The high density clones (in blue) tend to have a lower lignin content than the low density clones (in red). A one-way ANOVA test shows that the lignin measurements for the selected clones are significantly different ($P<0.05$), with the high density clones having a significantly lower level of lignin than the low density clones.

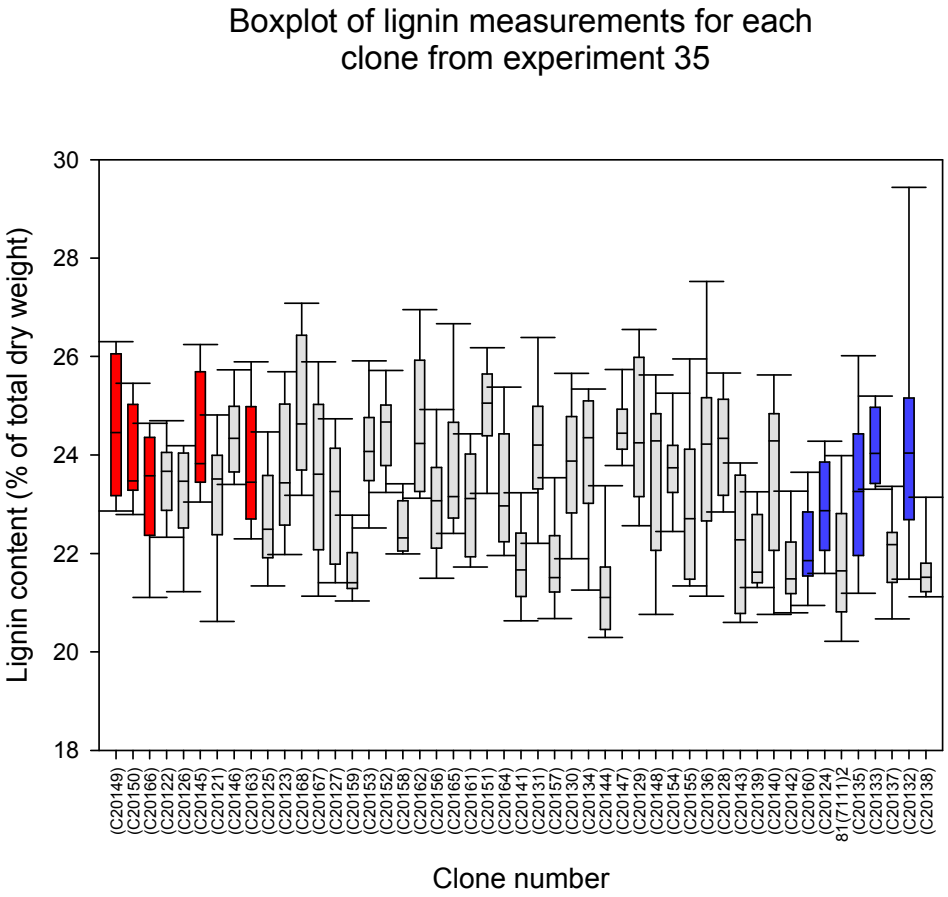


Figure 3.14: lignin measurements for all clones from experiment 35 at Newcastleton. Each boxplot shows the median, the 5th and 95th percentile. The bars show the lowest and highest extremes of lignin for each clone. Each clone has ten measured replicates out of fifteen total replicates. The plots highlighted in blue have been selected as high density clones, and the plots highlighted in red selected as low density clones. The clones are ordered based on their density, with the lowest density clones to the left of the graph, and the highest density to the right.

Figure 3.14 shows the overlap in lignin measurements for the high density (blue) and low density (red) clones, even though the average lignin measurements for each group of clones are significantly different.

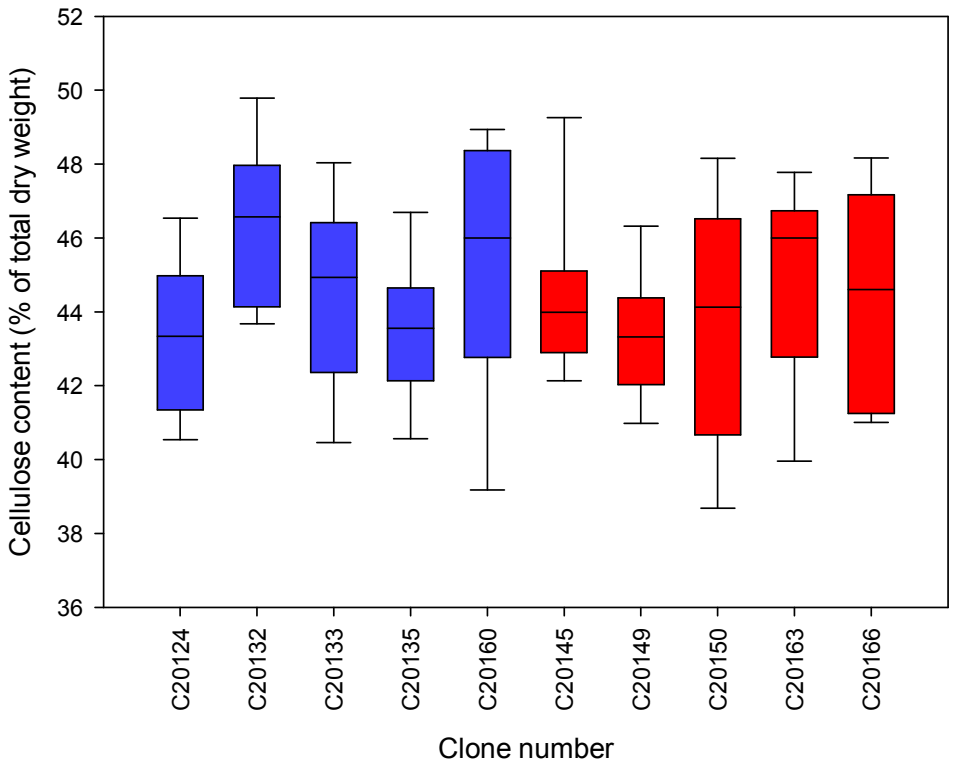


Figure 3.15: cellulose measurements for all selected clones from experiment 35 at Newcastleton. Each boxplot shows the median, the 5th and 95th percentile. The bars show the lowest and highest extremes of cellulose for each clone. Each clone has ten measured replicates out of fifteen total replicates. The plots highlighted in blue have been selected as high density clones, and the plots highlighted in red selected as low density clones.

Figure 3.15 shows the cellulose measurements for each of the 5 high density and 5 low density clones. The overlap of the cellulose measurements between the high and low density clones is much greater than with the lignin measurements. A one-way ANOVA shows no significant difference between the cellulose measurements for the two types of clones.

To summarise, the selected clones are significantly different with regard to their density measurements, though they all have similar growth measurements. The lignin concentration is significantly higher in the low density clones compared with the high density clones. There is no difference in cellulose content between the two types of clones.

3.3 Discussion

The density measurements obtained from the clonal trial (experiment 35) are within the range of those typical of British grown Sitka spruce. The average density of all of the replicates from experiment 35 is 350 kg/m^3, which is the same as the approximate average density measurements obtained from other British Sitka spruce plantations (Lee, 1999a).

Each of the 48 experimental clones has a maximum of ten measured replicates. The wood densities of some of the clones vary across a wide range whereas others exhibit more consistent density across replicates of the same clone (figure 3.1). It has been shown that there was no effect of block, row or column on the density of individual trees, so other factors must be causing these within clone differences. Savill and Sanders (1983) have shown that spacing has little effect on wood density, and attributed 63% of the variation in wood density to differences between trees within spacings, so it is likely to be non-silvicultural factors causing the differences in density and growth. Other studies have shown large variation in terms of growth measurements within a single Sitka spruce plot (Lee *et al*., 1999c); although these are not clonal experiments. A study on wood variation in *Pinus radiata* clones shows a similar range in wood density (300-400 kg/m^3) to this study (Lindström *et al*., 2004). Competition between individuals could be the cause of some of the physiological variation within the clones. One replicate of each clone is planted randomly within a block, so a replicate of a clone within one block will have a different set of neighbours to a replicate of the same clone planted in a different block. The relationship of a tree with its neighbours could have a significant impact on its phenotype, especially now that the trial is at the closed canopy stage. The density values used in this study were calculated from wood cores that covered the total wood accumulated within the lifetime of the trees. Competition from neighbours is less of a factor at a young age, so selection based on these density values should reduce some of the competitive effect, compared to selection based on outer ring density

measurements. This may explain why the correlation between pilodyn and direct wood density measurements was so low. The indirect density measurements from the pilodyn may only be reflecting tissue that has been produced during the closed canopy stage when competition has more of an effect on phenotype.

The negative correlation between wood density and diameter needs to be taken into account when selecting clones for further analysis, as the clones required need to cover a range of densities yet have similar growth measurements. Trees with more vigorous growth increase their earlywood component within an annual ring, but do not show a corresponding increase in the latewood component, leading to a reduction in wood density (Brazier, 1970). The high density and low density clones selected for microarray analysis show no significant difference in terms of growth measurements, so we can be confident that any difference in density between the two groups of clones cannot be explained by variation in growth rates.

Figure 3.10 shows the density range for the selected clones (highlighted in blue for high density and in red for low density) with all of the unselected experimental clones. It would be ideal restrict the choice of clones to those that have a small range of density measurements (and therefore exhibit greater consistency across replicates), for example clones 28, 40, 55 and 57. Unfortunately, these clones were either not high or not low enough in their density measurements, or their growth measurements varied too much from the experimental average.

Pilodyn measurements have previously been used to obtain indirect measures of density for screening in breeding programmes (Taylor *et al*., 2006). The results in this study give a relatively low correlation between density and pilodyn when compared to other studies (Lee, 1997). Due to the way that a pilodyn operates, the measurements obtained only relate to the outer ring of wood formed on the tree, so it does not provide an accurate estimate of whole tree density. As density

is positively correlated with distance from the pith (Taylor, 2006), the pilodyn measurements collected may overestimate whole tree density (Taylor, 2006).

The high and low density clones selected differed significantly in terms of their lignin content, with the low density clones having a higher lignin content than the high density clones (figure 3.13). As with the density measurements, there is a large amount of within clone variation with regards to lignin content (shown in figure 3.15). Kirst *et al*. (2004) also found this result in Eucalyptus, with faster-growing trees containing less lignin (22.4%) than slower-growing trees (24.8%), although there is some overlap of the lignin measurements for the two groups. Kirst *et al*. (2004) suggested that high levels of lignin could reduce carbon availability for growth. The effect of lignin reduction on growth has also been highlighted in transgenic studies; Hu *et al*. (1999) down-regulated 4CL (part of the lignin biosynthetic pathway) and this resulted in a 45% reduction in lignin, and enhanced leaf, root and stem growth.

So why are the clones in this experiment exhibiting such variation in their phenotypes? It may be that some clones are more responsive to their surroundings than others, or that some traits may be under stronger genetic control than other traits (Wu and Stettler, 1998), so are not affected as much by their environment. The effect of environment on a number of traits related to leaf production and branching has been studied in poplar (Wu and Stettler, 1998), with all traits showing a significant difference across sites with the exception of sylleptic branching. Phenotypic plasticity has been defined as 'the differences of a phenotypic trait produced by a genotype in response to changes in the environment' (Ma *et al*., 2008), or 'the ability of a genotype to generate a range of different phenotypes, depending on the environment' (Chambel *et al*., 2005).

Lack of clonal uniformity is a problem in genetic studies, as the decision has to be made whether to select samples on a clonal or an individual basis. For this study, the decision was made to select samples on a clonal basis to reduce the

variation caused by genetic factors between the trees. This approach has been adopted in other studies in spruce (Sarén *et al*., 2004). The results presented showed better correlations between diameter and density, and also lignin and density, when considered at a clonal mean level. As wood density is a highly heritable trait, the clonal approach may give us a greater likelihood of success in identifying genes than using a single ramet per clone approach. Using field trials for genetic studies leads to a large number of variables that need to be accounted for, such as site, competition, exposure, silviculture; and it is more difficult to ensure that all sources of variation are minimised when using field-grown trees. However, the clonal field trial set out on a uniform site offered the best available material to work on. Ideally, a clonal trial of mature trees planted at much wider spacing might have produced less within clone variation by reducing the interaction between neighbouring trees, but such material was not available.

From a collection of 50 clones, 10 clones have now been selected for use in microarray experiments. These 10 clones represent two groups with either high or low density clones, all with average growth rates, so they will be suitable for assessing the genes involved in producing high density wood. These clones have been selected on the basis of their density and growth (diameter) measurements.

Despite the overlap in density measurements between some of the high and low density clones, these were the best choice for gene expression analysis due to the wide range of densities exhibited and the average growth rates. Using these trees, it may be possible to assess both the genetic and the environmental effects on wood formation through microarray analysis, by hybridising samples with different densities to the same array to check for differential gene expression.

4 Sitka Spruce Gene Expression

4.1 Introduction

Analysing differences in gene expression between trees with low density wood and trees with high density wood may enable us to discover key genes that are involved in the control of wood quality. There are many different ways of analysing gene expression (summarised in chapter one), only some of which are appropriate for the use in our study of wood density in Sitka spruce as lack of sequence information for Sitka spruce means that some of the approaches are not a practical option. However, the recent development by Genome Canada of spruce cDNA microarrays offers a new and appropriate tool for high-throughput analysis of gene expression in these species. Southampton University was fortunate to be granted access to these microarrays through a generous collaboration with Kermit Ritland and Joerg Bohlmann of Genome Canada, Vancouver.

4.1.1 Sample Selection

As the aim of this project is to identify genes involved with wood density, we need to measure gene expression in samples from the region of the tree where wood formation occurs. The best place to collect samples is therefore the layer of cambium and xylem located directly underneath the bark of each tree. Gene expression is likely to vary with time of year, so the level of expression of 'wood formation' genes may be different at the beginning compared to the middle of the growth season. It is thus important to ensure that samples applied in a microarray run are all taken at the same stage in the growing season.

4.1.2 Microarrays

The following sections describe the theory and practice involved in developing and using microarrays for gene expression studies.

Recognition (or hybridisation) is the ability of a nucleic acid strand to recognise complementary sequences through the process of base pairing (Lipshutz *et al*., 1999). The use of oligonucleotide arrays was first proposed in the late 1980's, when several groups independently developed the idea of determining sequences using hybridisation to a set of oligonucleotides (Hoheisel, 1997). This technology was developed with the aim of sequencing by hybridisation (SBH). Two types of microarrays have been developed; DNA fragment-based microarrays and oligonucleotide-based microarrays (e.g. Affymetrix GeneChip™). DNA fragment-based arrays are produced by depositing DNA fragments at defined locations onto a substrate, usually a microscope slide. Oligonucleotide-based arrays are made using two techniques: photolithography and solid-phase DNA synthesis.

Fluorescently labelled cDNA is hybridised to the arrays, and the unbound probe is washed away. The fluorescent probe is then excited by light, and the signal emitted is a reflection of the abundance of the corresponding sequence in the original probe (Kehoe *et al*., 1999). Fluorescent tags that vary in their excitation and emission optima can be used to label two different probes (e.g. mRNA extracted from two different individuals of the same species). The fluorescent cyanine dyes Cy3 and Cy5 are most commonly used. The two probes are allowed to hybridise with the same microarray, thereby removing any problems with hybridisation differences between two arrays. After hybridisation, fluorescence measurements are made with a microscope, or a sophisticated scanner, that illuminates each DNA spot and measures fluorescence for each dye separately; these measurements are used to determine the ratio, and in turn the relative abundance, of the sequence of each specific gene in the two mRNA or DNA samples (Brown and Botstein, 1999).

4.1.2.1 Oligonucleotide vs. DNA arrays

It has been shown that there is no difference in sensitivity between oligonucleotide and DNA probes (Kane *et al*., 2000), although it has been suggested that oligonucleotide-based arrays are more sensitive to nucleotide mismatches, because the sequence that hybridises is very short (Kehoe *et al*., 1999).

Oligonucleotide-based microarrays are more consistent from one array to the next than are DNA fragment-based arrays (Kehoe *et al*., 1999). One major disadvantage of oligonucleotide-based arrays is that sequence knowledge is required for their construction, and so their application is restricted to species where the majority of the genome has been sequenced. In contrast, sequence information is not required for DNA-fragment based arrays, so this type of array is useful for assessing gene expression in species such as Sitka spruce, whose genome has not been sequenced. Labelling of probes, hybridisation, signal detection and data analysis are performed in the same way for both DNA fragment-based arrays and oligonucleotide-based arrays.

4.1.2.2 Substrate choice

The arrays that were used in this work were attached to a glass substrate. The substrate on which the DNA fragments are attached can be either nylon or glass, each of which has both advantages and disadvantages. The glass surface can be treated (e.g. with poly-L-lysine, amino silanes or amino-reactive silanes) so that DNA samples can be covalently attached to it. The treatment of the surfaces of the glass slide enhances both the hydrophobicity of the slide, and the adherence of the DNA that has been spotted on to it (Duggan *et al*., 1999). Glass is also more resilient to high temperatures, and is less likely to suffer detachment of the DNA when washed with solutions of high ionic strength (Cheung *et al*., 1999). Glass naturally emits only a low level of fluorescence, so it will contribute less to background ‘noise’. One of the major benefits of using a glass array is that samples can be assessed simultaneously using fluorescent dyes to label the samples, so that they can hybridise to the same array. This type of experiment

cannot be performed using nylon arrays, hybridisations can only be either serial or parallel (Cheung *et al*., 1999). Also, high density arrays can only be prepared on solid substrates, because deformation of nylon filters occurs with high density printing.

4.1.2.3 Preparing the array

DNA fragment-based arrays

The probes are spotted onto a suitable substrate in a regular pattern, usually onto a microscope slide (for high-density arrays). The slides are treated (various different methods are available) to increase adhesion of the probe to the slide, and to produce a uniform spot shape. Modified oligonucleotides can be covalently attached to a mercaptosilane-activated glass slide, which minimises side reactions involving other functional groups (Rogers *et al*., 1999). Once all the material has been spotted, the arrays are washed to remove any unbound material.

Oligonucleotide-based arrays

Light is directed through a photolithographic mask to specific areas on the glass substrate, producing localised areas of photodeprotection (Lipshutz *et al*., 1999). Hydroxyl-protected deoxynucleosides are incubated within the surface, and chemical coupling occurs in the illuminated regions. Light is then directed to different regions of the substrate using a different mask, and the same steps are repeated. The methods behind the production of DNA ‘chips’ are documented by McGall *et al*. (1997).

Target and probe size is an important consideration for both DNA fragment-based and oligonucleotide-based arrays. A common size for probes is generally between 20 to 30 bases, although the use of longer probes has been suggested to increase the specificity of the hybridisation reaction (Graves, 1999). Target size will affect the efficiency of the hybridisation reaction, as larger probes will

take longer to diffuse. It has been shown that even a small sequence hybridising to a large piece of DNA can uniquely identify an important gene (Graves, 1999).

4.1.2.4 Target RNA production and labelling

For the production of the target, RNA is extracted from the species of interest. One problem with array technology is the amount of RNA that is required. If the amount of RNA extracted is not large enough for microarray analysis, the RNA can be amplified, though this may introduce biases in the amplification process. Also, gene expression may change dramatically after disturbance, so the extracted material needs to be frozen immediately (e.g. in liquid nitrogen), and RNA extraction completed swiftly (Watson *et al*., 1998).

Fluorescent dyes are used to label the targets, and these can be incorporated when producing a cDNA copy of the RNA population. The amount of RNA labelled for each sample must be the same to produce accurate results. The cyanine dyes Cy3 and Cy5 vary in their emission and excitation optima and so are commonly used to label two distinct samples.

4.1.2.5 Hybridisation

Hybridisation can be performed in two ways, either using a hybridisation machine, or manually. A hybridisation machine consists of a chamber in to which the array is placed, and the targets are then added. After hybridisation is complete, the array is washed to remove any target that has not bound to the slide. The array can then be analysed to assess gene expression.

4.1.2.6 Signal detection

Laser confocal fluorescence scanning allows us to measure the molecular binding events that have occurred on an array (Lipshutz *et al*., 1999). Each spot is ‘excited’ using a laser with the particular wavelength specific to the dye used. Assuming that two different targets (and therefore two different dyes) have been used, the amount of light emitted at each spot is recorded at the two appropriate

wavelengths. Two images will be produced from this scanning, one for each dye used.

4.1.2.7 Data processing

Spot location from within the images is the first step of data analysis. Sensitivity is determined by the minimum signal that can be detected above background 'noise', and also on the incorporation of the label into the target RNA samples (Richmond and Somerville, 2000). There are several methods available for calculating intensity, and detecting an acceptable level of signal.

Data analysis is considered to be the most underdeveloped part of microarray technology (Richmond and Somerville, 2000). Normalisation of the data is the first step of data analysis, allowing comparison of separate fluorescent channels or separate experiments. There are several methods available for normalisation of the data. One method is identifying a set of genes whose expression levels remain fairly constant under varying conditions, termed 'housekeeping genes'. It is possible to adjust the fluorescent signal emitted by each target based on the expression level of the housekeeping genes. Another method, which is not based upon a set of housekeeping genes, assumes that most genes will stay constant in their level of expression under the conditions they are exposed to, so the mean, mode or median can be used to normalise the data. This method would not work for microarrays that are based on only a small number of genes. Richmond and Somerville (2000) suggest that, as all of the methods provide similar results, the Student-normalisation (based on standard deviation) might be the most appropriate as it is the simplest to use.

When the microarray process has been completed, it should be possible to identify a list of differentially expressed genes that can be analysed further.

4.2 Materials and Methods

4.2.1 Experiment One

4.2.1.1 Samples

Wood-forming tissue samples were collected during May 2004 from 150 trees (see Chapter 2 for methodology). The trees consisted of 15 ramets from each of five high wood density and 5 low wood density clones as described in Chapter 3. Of these 150 trees, only 10 out of the 15 replicates per clone had physiological data collected, so 100 samples were available for analysis. From this group, 60 samples were selected for microarray analysis. The samples were selected on the basis of their wood density measurements and the amount of RNA they yielded for microarray analysis. Each cambial sample that was collected was immediately frozen in liquid nitrogen and maintained constantly under these conditions prior to RNA extraction. RNA extraction was performed using a modified version of Chang *et al*. (1993). 40 µg of total RNA was used from each sample for microarray analysis. Experiment 1 was performed in the laboratories at the University of British Columbia, Vancouver.

4.2.1.2 Experimental Design

The experimental design was based on a statistical model developed by Rick White (Stats Department, UBC). On the advice of staff at UBC where the first experiment was formed, samples were paired randomly to ensure that both array channels contained samples which covered the whole range of wood densities across the 30 available arrays, and that each channel had an even number of samples graded 'high density' and 'low density' (table 4.1).

Clone	Wood density (H/L)	№ of samples
C20124	H	7
C20132	H	6
C20133	H	8
C20135	H	7
C20149	L	6
C20150	L	7

C20160	H	6
C20163	L	7
C20166	L	6

Table 4.1: Samples selected for experiment one. These samples were selected from the five high density clones and the five low density clones. The samples were paired using a statistical model (produced by Rick White, UBC) to ensure consistency across both dye channels.

For this experiment, samples were only selected from 9 of the 10 clones available, due to a lack of mRNA after extraction from the low density clone C20145.

Due to the limitation in the number of microarray slides that were available each of the 56 samples was hybridised only once, so there were no technical replicates or dye swaps included in the experimental design. The inclusion of technical replicates (i.e. dye swaps) in the 30 arrays we had available to use would have halved the number of samples that could be hybridised.

Two of the completed arrays showed problems when scanned, so these had to be repeated using 2 of the arrays from the 30 available. For this reason, the final analysis is based on 56 samples hybridised to 28 arrays.

4.2.2 Experiment Two

For experiments two and three, greater emphasis was placed on pairing samples with contrasting wood densities. Experiment two was completed in the Plant Science lab at Southampton University, using the same type of microarrays used in experiment one. This approach used only 2 clones, each of which had a smaller range of (within clone) density measurements than the other selected clones. This approach was taken to attempt to reduce the effect of environmental factors (such as competition) on wood density, by selecting clones that showed greater consistency across replicates. The protocol used for microarray hybridisation was easily transferred to the Southampton lab.

4.2.2.1 Samples

Two clones with relatively narrow within clone variation were selected from the 10 clones available. Clone C20133 was selected as the high density clone, and C20166 as the low density clone. The ten replicates from each clone were used for this experiment. The samples used in this experiment were all collected in July 2004.

4.2.2.2 Experimental Design

The ten replicates for clone C20133 were paired up with the ten replicates from clone C20166 to produce ten pairs of samples with contrasting wood densities (to be hybridised to ten arrays). The pairing was done in such a way that the highest density sample from clone C20133 was paired with the lowest density sample from C20166, then the second highest from C20133 with the second lowest from C20166, etc. By pairing the samples in this manner, gene expression could still be assessed on a clonal basis, rather than just assessing individual replicates.

	C20133 (high)		**C20166 (low)**		
Pair number	Replicate	Density (kg/m^3)	Replicate	Density (kg/m^3)	Difference in density (kg/m^3)
1	64	423	713	276	147
2	478	413	95	277	136
3	551	399	152	283	116
4	172	394	458	288	106
5	630	389	126	292	97
6	127	381	534	308	73
7	687	375	676	319	56
8	346	373	633	324	49
9	715	349	332	326	23
10	504	346	562	331	15

Table 4.2: Sample pairing for experiment two. Two clones were selected for this experiment, one high density clone (C20133), and one low density clone (C20166).

4.2.3 Experiment Three

The disadvantage of the approach used in Exp 2 was that some of the paired samples applied to the microarrays did not exhibit extremes of high vs. low wood density. In an attempt to improve this, the next experiment was set up such that each pair of samples had one very high and one very low wood density sample.

4.2.3.1 Samples

Samples were selected from four clones, two high density clones (C20132 and C20135) and two low density clones (C20149 and C20163). The two replicates representing each clone were chosen because they exhibited the highest wood density for the high wood density clones and the lowest wood densities for the low wood density clones. In this way the difference in the wood density of each given pairing was as large as possible. Four arrays were completed in total. The samples used in this experiment were collected in July 2004.

4.2.3.2 Experimental Design

Two ramets of each of four clones were selected for experiment three, two high density clones (C20132 and C20135), and two low density clones (C20149 and C20163) (table 4.3).

	C20132 (high)		**C20149 (low)**	
Pair number	Replicate	Density (kg/m^3)	Replicate	Density (kg/m^3)
1	731	462	475	256
2	98	432	327	279
	C20135 (high)		**C20163 (low)**	
3	683	429	154	306
4	649	409	610	309

Table 4.3: Sample pairing for experiment three. The samples were paired together as 'extreme' arrays – the two samples hybridised to each array had a difference in density of at least 100kg/m^3.

4.2.4 Microarray Hybridisation

The method used for completing the spruce arrays was developed by Steven Ralph and Hesther Yueh (2004).

4.2.4.1 Reverse transcriptase reaction

1st strand synthesis reactions were prepared in 200 µl PCR tubes containing 40 µg of total RNA dissolved in 18 µl of nuclease-free water (kit vial #10), and 2 µl of Cy3 or Cy5 primer (kit vial #5). The tubes were incubated at 80° C for 10 minutes and then chilled immediately on ice for at least two minutes. The tubes were centrifuged briefly to collect the contents at the bottom of the tubes.

In the meantime, a reverse transcription master mix (for each reaction) containing 8 µl 5X first strand buffer, 4 µl dNTPs mix (kit vial #3), 4 µl 0.1M DTT, 3 µl superscript II and 1 µl superase-IN RNase inhibitor (kit vial #4), was made up at room temperature. If there were 6 tubes in total (e.g. for three arrays), the volumes added to the master mix would be six times the amount stated (e.g. 48 µl 5X first strand buffer). 20 µl of the master mix was added to each reaction. The contents were gently mixed and then the tubes centrifuged briefly. They were then incubated at 42° C for 2 hours.

After this incubation, 7µl of a 0.5M NaOH/0.05M EDTA solution (made fresh that day with 14µl of 5M NaOH, 14µl 0.5M EDTA pH 8.0 and 112µl nuclease-free H_2O) was added to each reaction. The tubes were then incubated for 15 minutes at 65°C, and then cooled down to 4°C.

After the heat denaturation was completed, 10 µl of 1 M Tris-Cl pH 7.5 was added. The tubes were briefly centrifuged. The Cy3 and Cy5 reactions for a given hybridisation were combined into a single 1.5 ml tube, and 46 µl of TE pH 8.0 was used to rinse each PCR tube before being added to the mix (giving a total volume of 160 µl).

4.2.4.2 Probe precipitation

3 µl of linear acrylamide (kit vial #5), 16 µl of 3.3M Na Acetate pH 5.2 and 400 µl of 100% ethanol was added to each pooled cDNA mixture. The contents were left to precipitate at -80° C for at least one hour.

After the precipitation step, the tubes were centrifuged at 14,000 rpm at 4° C for 50 minutes. The pellet was then washed with 500 µl of 70% ethanol, and spun for an additional 10 minutes. The pellet was then allowed to air dry at room temperature for 15 minutes before being re-suspended in 18 µl of nuclease-free water (kit vial #10) for 10 minutes in a 65° C heat block.

4.2.4.3 Prewash

During the 15 minute centrifugation step, the slides were prewashed. 750 ml of Milli-Q H_2O (covered with cling film to prevent dust contamination) was heated to 95° C on a heat block. Each slide was placed into a coplin jar containing 0.1% SDS and gently mixed by inverting for 20 seconds, then left to sit on a bench for 4 minutes and 40 seconds. The buffer solution was replaced and the process repeated. The slides were then transferred to a coplin jar containing Milli-Q H_2O, gently mixed by inverting for 20 seconds, and then left to sit on a bench for 1 minute and 40 seconds. The water was then replaced and the process repeated.

The slides were then transferred to a slide rack, and the rack placed end down with the printed slides facing upwards into the 95° C Milli-Q H_2O for 3 minutes. The slides were then immediately transferred to 50mL Falcon tubes containing a tissue tightly packed at the bottom. The tubes were centrifuged (without their caps) for 3 minutes at 2000 rpm. N_2 gas was blown gently over the slides to remove any dust, and they were then placed into the hybridisation chambers.

4.2.4.4 1st hybridisation

2x SDS-based hybridisation buffer (kit vial #6) was thawed at room temperature for 15 minutes (wrapped in aluminium foil to protect it from the light). This was then mixed briefly and heated in a water bath at 65° C for 15 minutes. The buffer was then vortexed thoroughly, centrifuged and kept in a heat block at 65° C until added to the master mix.

Salmon sperm DNA was denatured at 95° C for 10 minutes, and then chilled on ice. The master mix was prepared with 22.5 µl of 2x SDS buffer, 4 µl LNA dT

blocker (kit vial #9), and 0.5 µl salmon sperm DNA per reaction. The master mix was kept at 65° C until added to the cDNA.

27 µl of the hybridisation master mix was added to the 18 µl of cDNA probe, and the contents gently vortexed, and then centrifuged. The probes were then denatured in an 80° C heat block for 10 minutes, and then transferred to a 65° C heat block until added to the slides.

N2 gas was blown over the required number of 22x60 mm (1.5 mm thickness) coverslips to remove the dust. 20 µl of Milli-Q H_2O was added to each of the two wells on the hybridisation chamber to minimise probe evaporation. A single cDNA probe was vortexed, centrifuged, and added in a straight line 5 mm inwards from the long edge of the slide. Any bubbles were popped with a fresh pipette tip. The coverslip was added to the slide so that the hybridisation solution covered the entire printed area of the slide. The hybridisation chamber was then closed, and placed into a 60° C water bath to hybridise for 16 hours. This process was repeated with the remaining cDNA probes.

4.2.4.5 Post 1st hybridisation wash

A coplin jar containing 2x SSC/0.2% SDS was placed into a water bath set at 65° C. The hybridisation chambers were removed from the 60° C water bath and disassembled. The slides were placed into a coplin jar containing 2x SSC/0.2% SDS at room temperature and left for 5 minutes to allow the coverslips to slide off. Each slide was transferred to the coplin jar containing 2x SSC/0.2% SDS and placed in a 65° C water bath to incubate without agitation for 15 minutes.

Each slide was then transferred to a coplin jar containing 2x SSC at room temperature. The jar was inverted for 30 seconds, then left to sit for 4minutes and 30 seconds. This is repeated twice to give a total time of 15 minutes.

The slides were then transferred to a jar containing 0.2x SSC at room temperature. The jar was inverted for 30 seconds then left to sit for 4 minutes and

30 seconds. The wash buffer was replaced and the 5 minute wash repeated. The buffer was replaced for a 2nd time and washed for a final 5 minutes.

The slides were transferred to 50 ml Falcon tubes containing tissue tightly packed at the bottom, and were centrifuged (without the cap) for 3 minutes at 2000 rpm. N_2 gas was gently blown over the slides to remove any dust, and they were then placed into the hybridisation chambers.

4.2.4.6 2nd hybridisation

2x SDS-based hybridisation buffer and 3DNA Array 350 capture reagent (kit vial #1) were thawed at room temperature for 15 minutes (wrapped in aluminium foil to protect from the light). They were then mixed briefly and incubated in a 65° C and 55° C water bath, respectively, for 15 minutes. Then were then vortexed, centrifuge and stored in a 65° C heat block.

The master mix was then prepared with 22.5 µl of 2x SDS buffer, 17.5 µl nuclease-free water, and 2.5 µl of each Cy3 and Cy5 dendrimer per reaction. The master mix was denatured at 80° C for 10 minutes, and then transferred to a 65° C heat block until added to the slides.

N_2 gas was blown over the required number of coverslips to remove any dust. 20 µl of Milli-Q H_2O was added to both wells on the hybridisation chamber. A single dendrimer probe was vortexed, centrifuged, and added to each slide in a line 5 mm inwards from the long edge. Bubbles were popped with a fresh pipette tip. A coverslip was added so that the hybridisation solution covered the entire printed area of the slide. The hybridisation chamber was closed, and placed in a 60° C water bath to hybridise for 3 hours. This procedure was repeated with the remaining dendrimer probes.

4.2.4.7 Post 2nd hybridisation wash

The same wash procedure as for the post 1st hybridisation wash was repeated. After drying, the slides were ready for scanning.

4.2.5 Array Scanning

The arrays from experiment one conducted in UBC were scanned using a Scan Array Express Packard Bioscience scanner. Lasers one (633nm) and three (543nm) were set to 90% power by default. The slides were inserted with a gentle downwards force to prevent the printed DNA from scraping against the upper wall of the opening.

Partial scans were conducted for each slide to determine optimal scanning conditions, at a resolution of 10μm. The correct fluorophores were selected (cyanine 3 and cyanine 5), and the photomultiplier tube (PMT) gain adjusted during the scans to increase or decrease the amount of fluorescence captured. Several small scans were completed across the slide to obtain suitable PMT values. A full scan was then completed. The output files are saved as TIFF (tagged image file format) images ready for processing.

The arrays for experiments two and three (performed at Southampton University) were scanned using a Genetix scanner to produce a TIFF file for each channel. A visual inspection from the composite image of the arrays could be made during scanning.

4.2.6 Image processing

Image processing is completed to measure the signal intensity (background and foreground) of each spot on the array. Experiment one arrays were processed using ImaGene software (BioDiscovery, Marina Del Ray, USA). A grid was overlaid on each array to help with spot location, and then the program identified the pixels that make up a spot. The results were saved as text files, ready for quality control evaluation and analysis.

4.2.7 Analysis

4.2.7.1 UBC analysis

Normalisation was completed for channel, array and clone effects, and then an r-script was used to determine the level of correlation between sample density and gene expression.

4.2.7.2 Genespring analysis

The files were normalised using an R-script (written by Nathaniel Street). The normalisation included a print-tip LOWESS, which checks for consistency between print-tip groups (i.e. array subgrids), and a scaled normalisation between arrays. The normalised files were then ready for analysis using GeneSpring (Version 6.3, Silicon Genetics, USA).

4.3 Results

4.3.1 Experiment One

4.3.1.1 Introduction

The array used was a 21.8 K spruce array (Sitka spruce and White spruce) based on a combination of tissues produced by the Forestry Department, the Michael Smith Genome Sciences Centre and the Biotechnology Research Institute in Montreal.

4.3.1.2 RNA – quantity and quality

All the samples used in the experiment were quantified using the NanoDrop® ND-1000 Spectrophotometer. The samples all contained a minimum of 40 µg of RNA, which is the amount required from one sample for hybridisation. 25% of the samples were run on an Agilent 2100 bioanalyser to assess the quality of the extracted RNA. Figure 4.1 shows the electrophoresis file run summary for these samples.

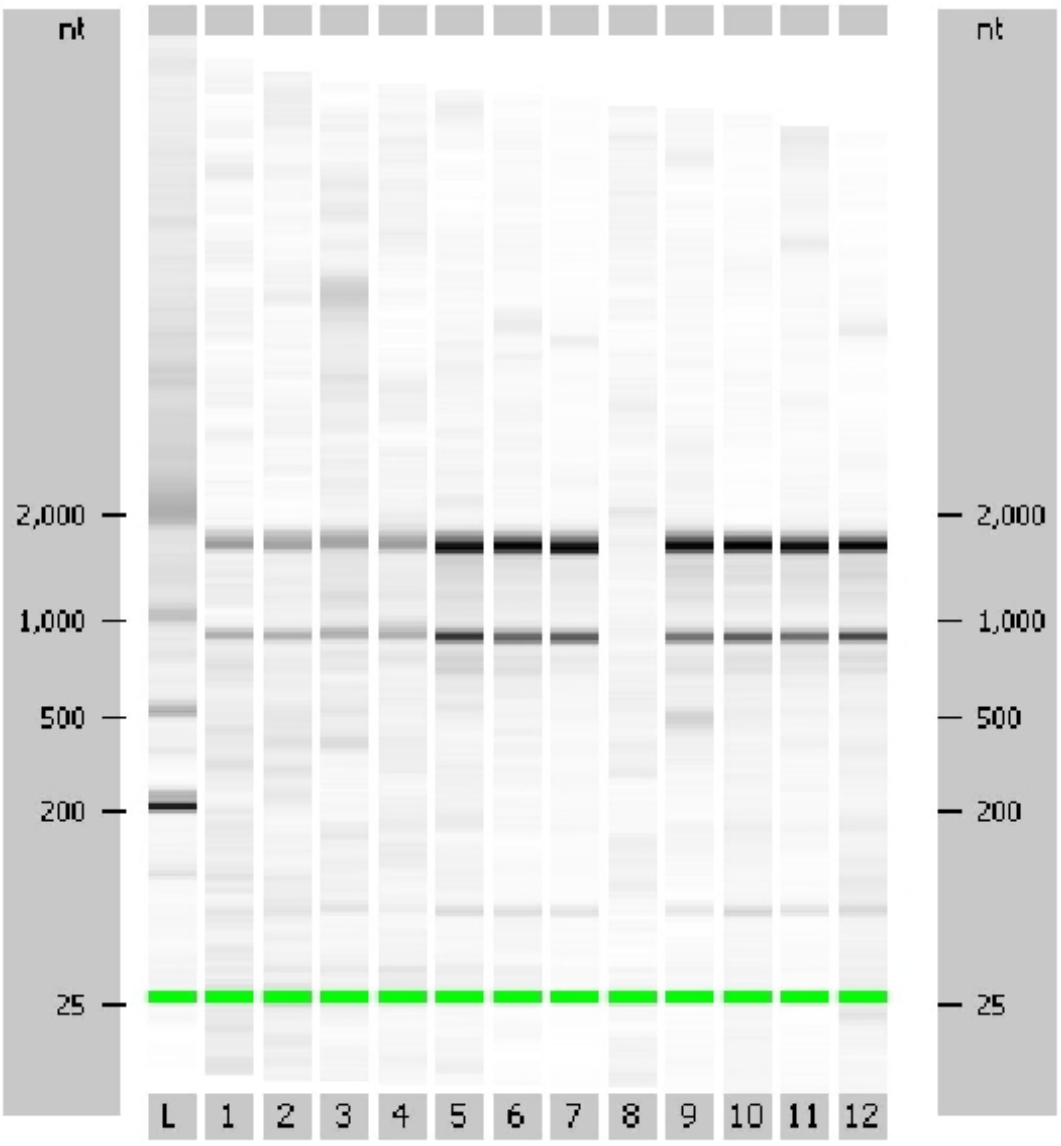


Figure 4.1: electrophoresis file run summary for 12 of the samples used in experiment one. The first lane on the left of the figure contains the RNA 6000 size standard ladder. With the exception of sample 8, the presence of two discrete bands for all samples at molecular weights approx 900 and 1800 represent the 28S and 18S rRNA bands (for eukaryotic samples), indicate that the RNA remains intact following extraction and there is no evidence of degradation. The RNA in lane 8 appears to have degraded, so would not be suitable for analysis.

The concentration of RNA in the samples is estimated by comparison of the intensity of their fluorescence with that of the ladder of known RNA concentration, and the fragments of known size provided by the ladder are used to identify the ribosomal peaks in the samples (figure 4.2).

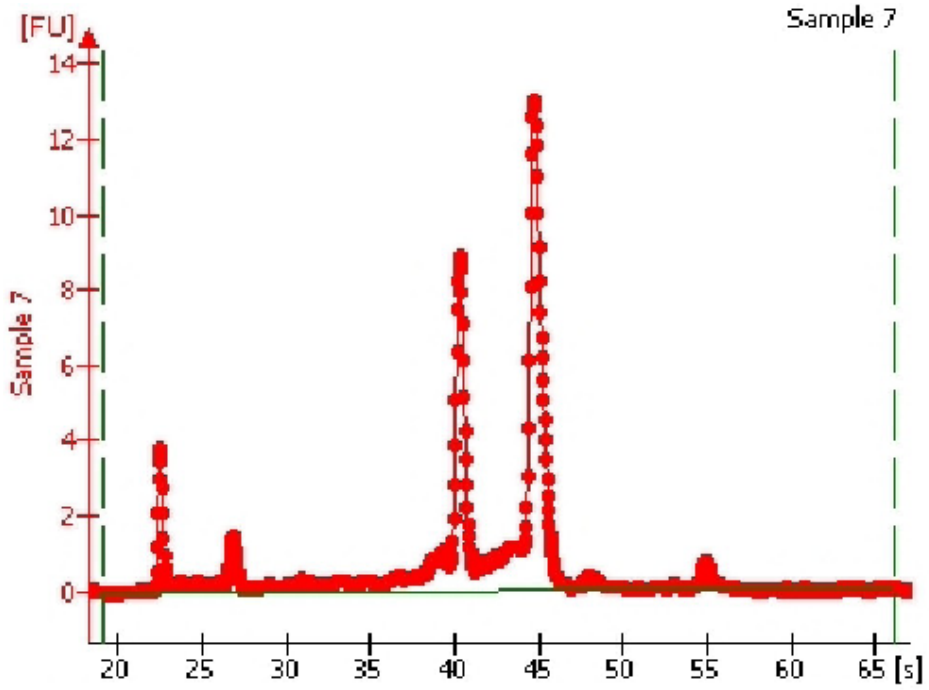


Figure 4.2: Electropherogram from experiment one. A dye was added to each channel on the bioanalyser chip, and the samples were then detected by their resulting fluorescence. The intensity of fluorescence of the sample is plotted on the y-axis, with time (in seconds) on the x-axis. The RNA quantity was estimated by integrating the area under the peaks. The first large peak on the x axis at 40 seconds represents the 18S band, and the second large peak at 45 seconds represents the 28S band. A 2:1 ratio of 28S:18S rRNA indicates intact RNA.

To assess the quality of extracted RNA, a ratio was calculated from the peak areas of the ribosomal bands, 18S/28S for eukaryotic RNA. A ratio of 1.8-2.2 for eukaryotic RNA was considered acceptable. Deviation from this indicates RNA degradation. The left peak from the electropherogram (figure 4.2) represents the 18S band and the right peak the 28S band. The two small peaks to the left represent small RNA (5S and 5.8S peaks), presence of which varies depending

on the extraction protocol used. This RNA should represent 5-10% maximum of the 18S/28S RNA.

4.3.1.3 Microarrays

4.3.1.3.1 UBC analysis

These arrays were first analysed using an R-script produced by Rick White (stats. Department, UBC). Approximately 700 of the spots had a P value of less than 0.05, but these were considered to be differentially expressed by chance, rather than showing a significant result. When a more stringent Q-value (Storey and Tibshirani, 2003) was applied, no spots were considered to be differentially expressed. The Q-value is an adjusted version of the P-value based upon the false discovery rate (FDR). A definition of the Q-value for a particular spot is 'the expected proportion of false positives incurred when calling that feature significant' (Storey and Tibshirani, 2003). From this we can conclude that there is no correlation between the density measurements for each sample and the levels of gene expression shown on the arrays (i.e. there were no genes that increased in expression level in correlation with wood density).

These results were calculated using a statistical model produced by Rick White (personal communication). As this approach adopts a rather stringent approach to detecting up-regulation of gene expression, we also used a different method to look for differences in the gene expression of our samples. The text files for 10 of the arrays were imported into Genespring, for a less stringent form of analysis. These 10 arrays were selected on the basis of the densities of the samples arrayed to them. These selected arrays had samples arrayed with extreme densities (e.g. a difference of at least 100 kg/m^3) between the two samples).

The ten 'extreme' density arrays were analysed using a two-fold up and down cut-off to check for differentially expressed genes.

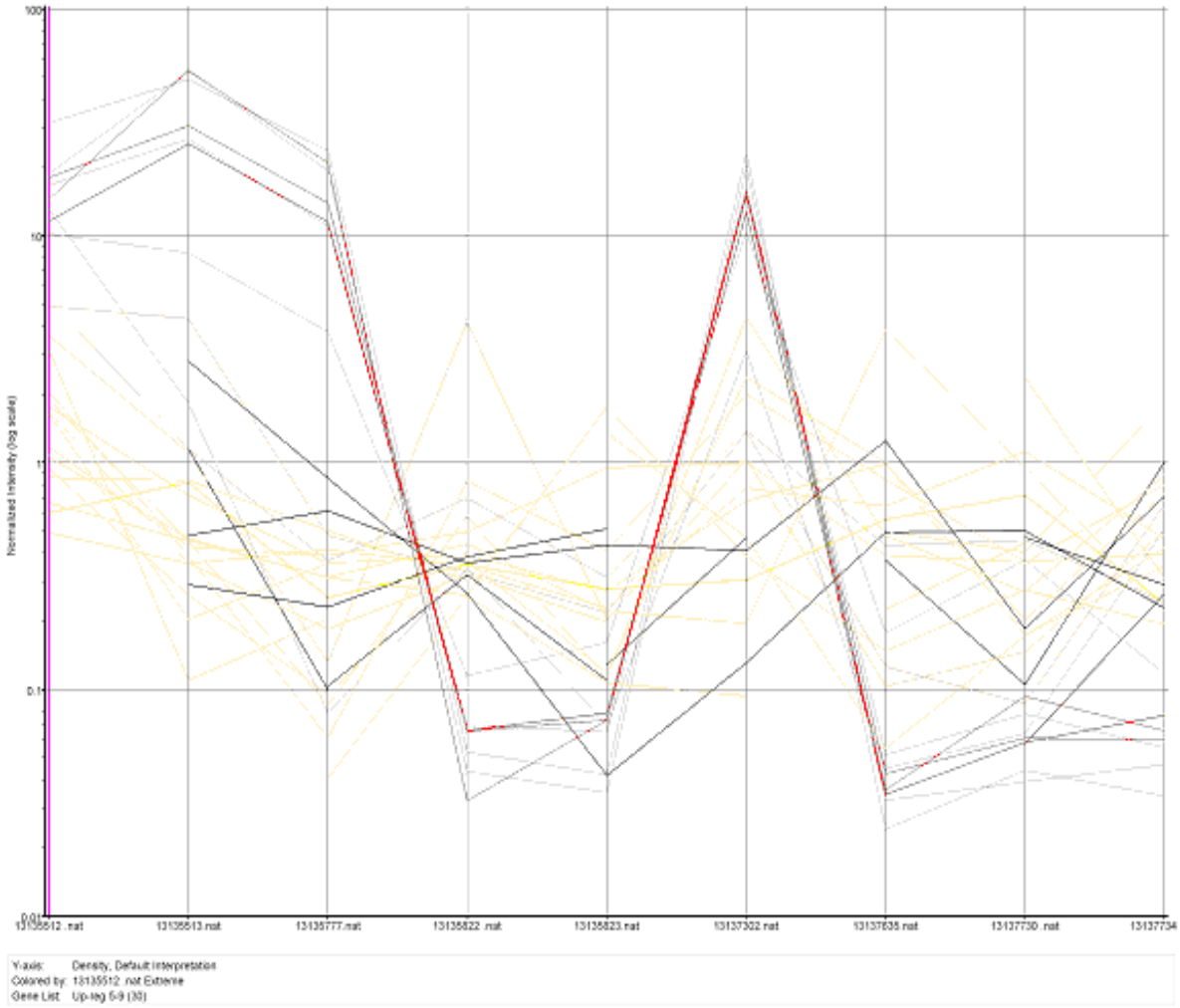


Figure 4.3: Two-fold up-regulated genes across 5 of 9 extreme density arrays from experiment one. The samples hybridised to each array had a minimum difference in density of 100kg/m^3. The x-axis shows the nine arrays that were selected for this analysis, and how the expression of the listed genes changes from one array to the next. The y-axis shows the level of expression, with anything above 2 considered up-regulated, and anything below 0.5 considered down-regulated.

30 genes passed the 2-fold change filter. These genes are listed in table 4.4.

Gene ID	Putative annotations (*Arabidopsis thaliana*)	AGI #	MIPS functional group
WS00828_A07	No significant hit		
WS00925_D02	Aquaglyceroporin	At4g18910	
WS0262_A03	Osmotin-like protein (OSM34)	At4g11650	
WS0094_C05	Disease resistance-responsive family protein	At5g42510	Cell rescue, defence and virulence
WS00930_I18	No significant hit		
WS0092_M09	Tubulin beta-2/beta-3 chain (TUB3)	At5g62700	Cell rescue, defence, virulence; interaction
WS00926_D01	Expressed protein	At5g50730	Unclassified
WS00929_H24	No significant hit		
WS01027_D10	Aminotransferase-related	At5g36160	Metabolism
WS01028_O15	Peroxidase 12 (PER12)	At1g71695	Cell rescue, defence, virulence; interaction
WS0106_J08	No significant hit		
WS0262_P08	Peroxidase 57 (PER57)	At5g17820	Cell rescue, defence, virulence; interaction
WS01011_P05	No significant hit		
WS01012_C04	Isocitrate dehydrogenase	At5g03290	Energy
WS01016_K10	No significant hit		
WS0071_P17	Glycine hydroxymethyltransferase (GHMT)	At5g26780	
WS00722_P03	Receptor protein kinase-related	At3g22060	Interaction with the environment
WS0073_H02	No significant hit		
WS00730_B03	Myrosinase-binding protein	At1g52030	Protein with binding function
MGC1001	No significant hit		
MGC1199	No significant hit		
WS0019_A03	Terpene synthase/cyclase family protein	At5g48110	Subcellular localisation
WS0092_B07	Hydrolase, alpha/beta fold family protein	At1g26360	Protein fate
WS0092_D23	No significant hit		
WS0092_H18	Transcriptional regulator Sir2 family protein	At5g09230	
WS0102_G09	No significant hit		
WS0064_M15	PfkB-type carbohydrate kinase family protein	At5g58730	Metabolism
WS0075_N05	No significant hit		
WS0085_L08	No significant hit		
WS00914_D10	No significant hit		

Table 4.4: List of two-fold up-regulated genes across 5 of 9 extreme density arrays from experiment one. Putative functions were assigned based on a BLASTx search of the TAIR (The Arabidopsis Information Resource) database. Functional classification is provided by MIPS (Munich Information Centre for Protein Sequences).

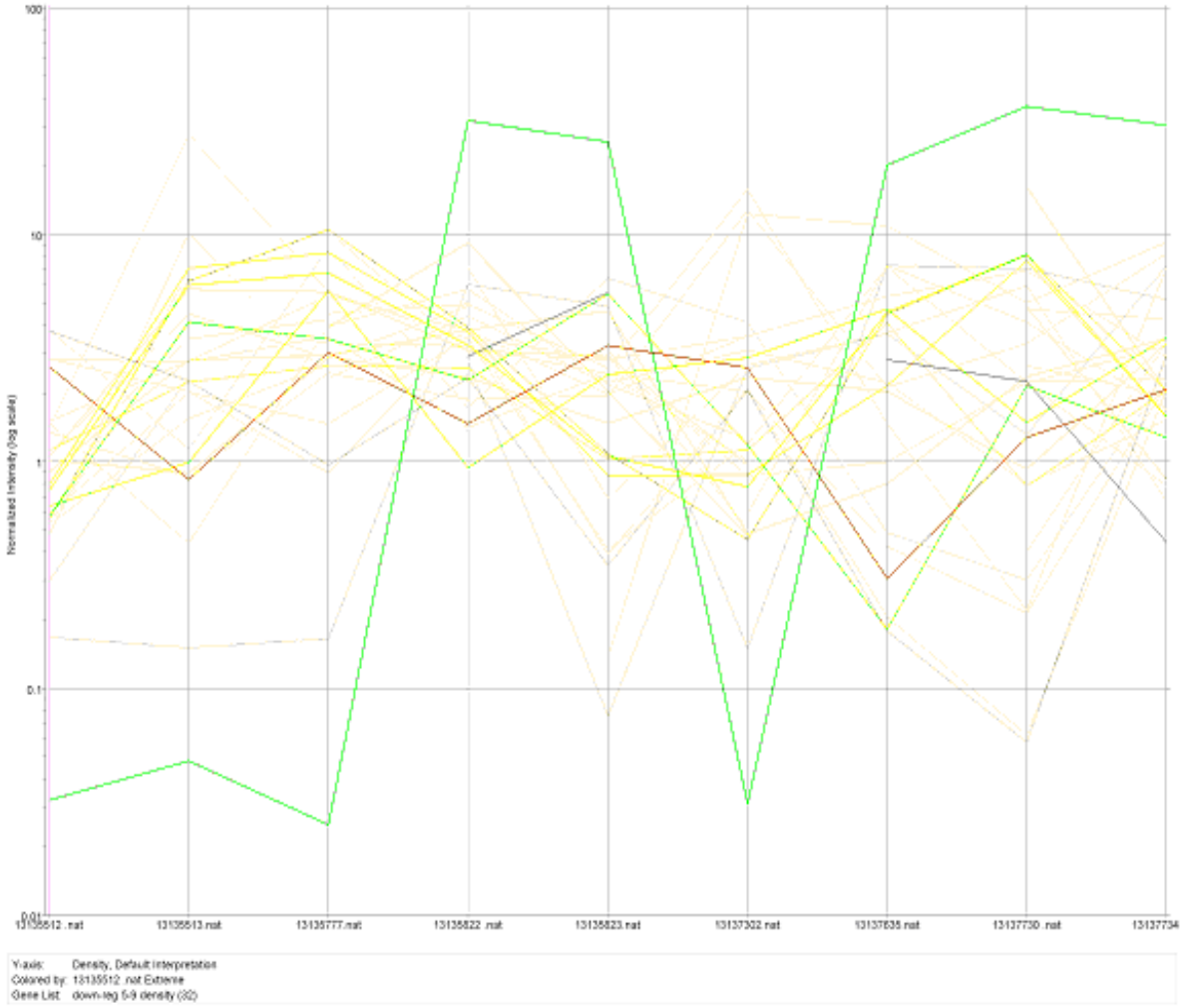


Figure 4.4: Two-fold down-regulated across 5 of 9 extreme density arrays from experiment one. The samples hybridised to each array had a minimum difference in density of 100kg/m^3. The x-axis shows the nine arrays that were selected for this analysis, and how the expression of the listed genes changes from one array to the next. The y-axis shows the level of expression, with anything above 2 considered up-regulated, and anything below 0.5 considered down-regulated.

Gene ID	Putative annotations (*Arabidopsis thaliana*)	AGI #	MIPS functional group
WS0042_M20	Leucine-rich repeat family protein	At1g69545	
WS00110_B19	Expressed protein	At3g56290	Subcellular localisation
WS01033_C04	Transferase family protein	At5g17540	Unclassified
IS0014_E17	Leucine-rich repeat family protein	At5g49760	Metabolism
WS00924_B06	No significant hit		
WS0099_J20	Cytochrome P450 98A3	At2g40890	
WS0097_A24	No significant hit		
WS00941_A05	BURP domain-containing protein	At1g49320	Unclassified
WS00916_N15	Hypothetical protein	At1g27110	Unclassified
WS00937_P14	Seven transmembrane MLO family protein	At2g17430	
WS00938_M21	40S ribosomal protein S17	At5g04800	Protein synthesis
WS01035_F15	No significant hit		
WS0106_J01	Leucine-rich repeat family protein	At1g69545	
WS0107_E10	Allene oxide synthase (AOS)	At5g42650	Metabolism
WS0101_E21	Leucine-rich repeat family protein	At1g69545	
WS01012_N24	No significant hit		
WS01026_B05	No significant hit		
WS0024_D18	Aspartyl protease family protein	At2g03200	Protein fate
WS0058_F16	Isoflavone reductase	At1g75290	Protein with binding function
WS00725_D20	No significant hit		
WS0075_D20	No significant hit		
MGC15287	No significant hit		
WS0022_I11	Lipid transfer protein 3 (LTP3)	At5g59320	Protein with binding function
WS0087_L03	Early nodulin-related	At5g25940	Subcellular localisation
WS00914_P22	No significant hit		
WS00916_C09	No significant hit		
WS0261_E15	Caffeoyl-CoA 3-O-methyltransferase	At4g34050	Metabolism
WS0023_D03	Lipid transfer protein 4	At5g59310	Protein with binding function – lipid binding
WS0078_L10	Leucine-rich repeat family protein	At5g49760	Metabolism
WS00815_K09	No significant hit		
WS00816_F21	Beta-fructosidase	At1g55120	Metabolism

Table 4.5: List of two-fold down-regulated genes across 5 of 9 extreme density arrays from experiment one. Putative functions were assigned based on a BLASTx search of the TAIR (The

Arabidopsis Information Resource) database. Functional classification is provided by MIPS (Munich Information Centre for Protein Sequences).

Five arrays were then selected from the 30 available for analysis based on their lignin measurements. Each array was selected due to maximum difference in lignin concentration of the two samples hybridised. Gene lists were determined using a two-fold expression cut off.

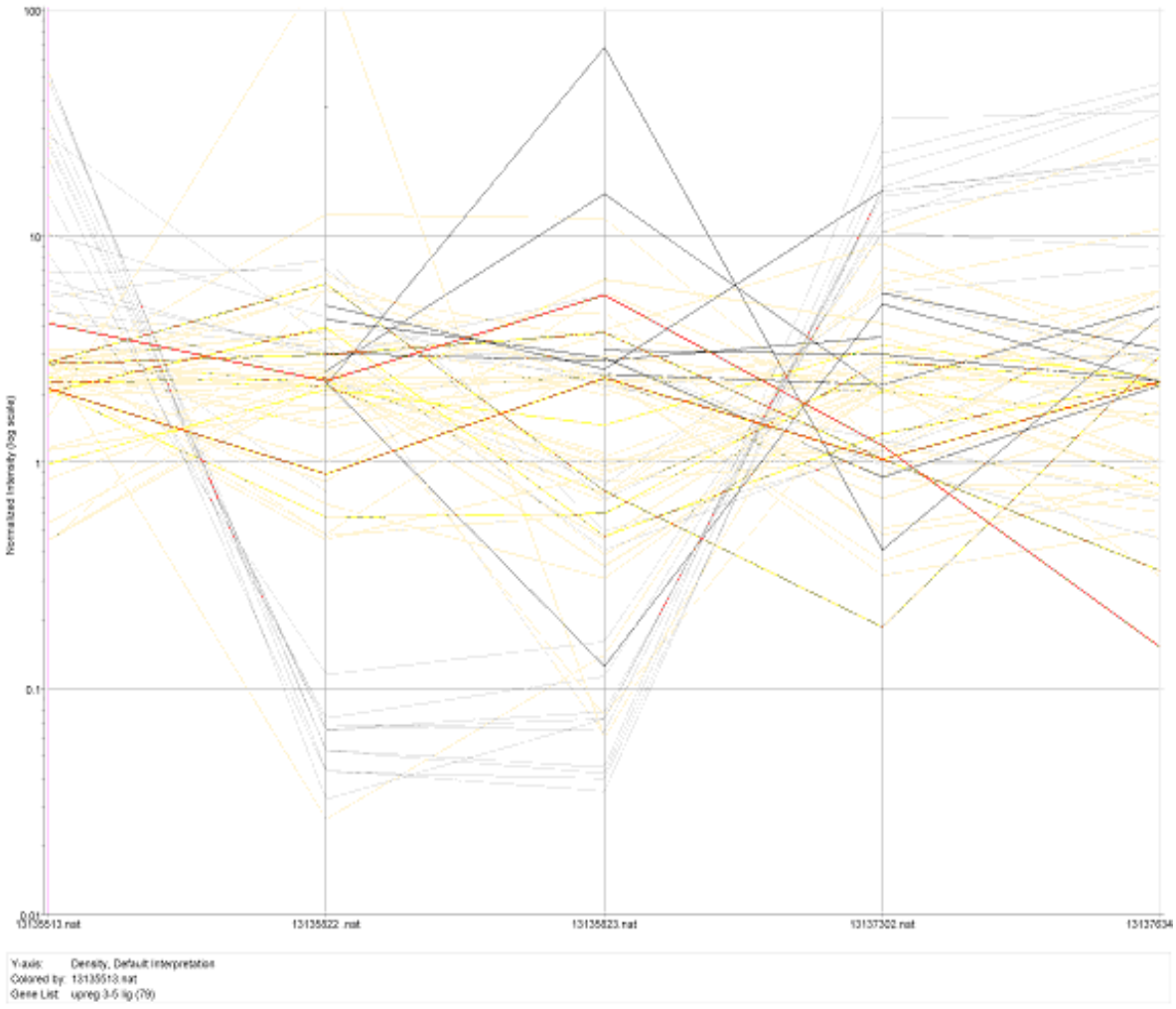


Figure 4.5: Two-fold up-regulated genes across 3 of 5 extreme lignin arrays from experiment one. The x-axis shows the five arrays that were selected for this analysis, and how the expression of the listed genes changes from one array to the next. The y-axis shows the level of expression,

with anything above 2 considered up-regulated, and anything below 0.5 considered down-regulated.

Gene ID	Putative annotations (*Arabidopsis thaliana*)	AGI #	MIPS functional category
WS01035_F15	No significant hit		
WS0101_K05	No significant hit		
WS0089_O02	No significant hit		
WS00926_N10	Ethylene-responsive protein (ER6)	At1g09740	Cell rescue, defense and virulence – stress response
WS00930_I18	No significant hit		
WS00921_P03	No significant hit		
WS00922_B10	No significant hit		
WS00937_P14	Seven transmembrane MLO family protein	At2g17430	
WS00938_M21	40S ribosomal protein S17 (RPS17D)	At5g04800	Protein synthesis
WS00929_E22	No significant hit		
WS0094_L20	No significant hit		
WS00918_N19	Ubiquinol-cytochrome C reductase complex 7.8 kDa protein	At1g15120	Energy
WS00918_O04	No significant hit		
WS00916_C09	No significant hit		
WS00916_N11	Aquaglyceroporin / NOD26-like major intrinsic protein 2 (NLM2)	At4g18910	
WS00918_P09	DEAD/DEAH box helicase	At5g54910	Cell cycle and DNA processing
WS0092_M09	Tubulin beta-2/beta-3 chain (TUB3)	At5g62700	
WS00921_L17	Nodulin MtN3 family protein	At5g53190	Unclassified
WS0092_H02	Ribosomal protein S14 mitochondrial family protein	At2g34520	
WS0092_H18	Transcriptional regulator Sir2 family protein	At5g09230	
WS0094_I03	PQ-loop repeat family protein / transmembrane family protein	At4g36850	Unclassified
WS01027_B13	No significant hit		
WS01033_C04	Transferase family protein	At5g17540	Unclassified
WS01025_H24	No significant hit		
WS01026_H11	Cytochrome P450	At2g26710	
WS0104_G16	UDP-glucoronosyl/UDP-glucosyl transferase family protein	At4g01070	
WS0107_P09	No significant hit		
WS0261_E15	Caffeoyl-CoA 3-O-methyltransferase	At4g34050	Metabolism
WS01040_J24	Transferase family protein	At3g03480	
WS0106_C23	No significant hit		
WS0101_E21	Leucine-rich repeat family protein	At1g69545	

WS01011_P05	No significant hit		
WS00941_A05	BURP domain-containing protein	At1g49320	Unclassified
WS00941_O06	No significant hit		
WS01012_C04	Isocitrate dehydrogenase, putative / NAD+ isocitrate dehydrogenase	At5g03290	Energy
WS01025_C06	No significant hit		
WS01025_G16	ERF domain protein 9 (ERF9)	At5g44210	Cell cycle and DNA processing
WS01017_A12	Disease resistance protein (TIR-NBS-LRR class)	At5g46260	Cell rescue, defense and virulence
WS01024_E15	Serine/threonine protein kinase family protein	At1g66880	Metabolism
WS0263_M18	Cytochrome P450	At3g14690	
WS0042_M20	Leucine-rich repeat family protein	At1g69545	
WS0043_D09	No significant hit		
WS0023_D03	Lipid transfer protein 4 (LTP4)	At5g59310	Cellular transport
WS0042_G13	Leucine-rich repeat family protein	At1g69545	
WS0045_N01	No significant hit		
WS0058_F16	NADPH oxidoreductase	At1g75290	Protein with binding function
WS0058_G14	No significant hit		
WS0048_G14	No significant hit		
WS0048_G23	Osmotin-like protein (OSM34)	At4g11650	Interaction with the environment
WS00715_G11	Expressed protein		
MGC1001	No significant hit		
MGC1199	No significant hit		
IS0013_H08	No significant hit		
IS0014_E17	Leucine-rich repeat family protein	At5g49760	Metabolism
Pa2420	No significant hit		
WS0019_C03	Glyceraldehyde 3-phosphate dehydrogenase A (G3PDH)	At3g26650	Metabolism
WS0022_D09	Dehydration-responsive protein (RD22)	At5g25610	Storage protein; cell rescue, defense and virulence
WS00110_B19	Expressed protein	At3g56290	Subcellular localisation
WS0013_M03	No significant hit		
WS00822_K11	No significant hit		
WS00823_M04	No significant hit		
WS00817_D03	Pentatricopeptide (PPR) repeat-containing protein	At5g46100	Unclassified
WS00819_D04	No significant hit		
WS00827_O16	No significant hit		
WS0084_K04	No significant hit		
WS0084_L08	No significant hit		
WS00828_D14	No significant hit		

WS0084_B07	No significant hit		
WS00914_P22	No significant hit		
WS0073_A07	No significant hit		
WS0073_D24	Cellulose synthase	At5g05170	Metabolism
WS0022_I11	Lipid transfer protein 3 (LTP3)	At5g59320	Protein with binding function
WS00725_D20	No significant hit		
WS0075_D20	No significant hit		
WS00811_L03	No significant hit		
WS00815_K09	No significant hit		
WS0075_K01	Non-symbiotic hemoglobin 1 (HB1) (GLB1)	At2g16060	Protein with binding function; Cell rescue, defense and virulence
WS00811_J15	No significant hit		
WS00816_F21	Beta-fructosidase	At1g55120	Metabolism

Table 4.6: List of two-fold up-regulated genes across 3 of 5 extreme lignin arrays from experiment one. Putative functions are included. Putative functions were assigned based on a BLASTx search of the TAIR (The Arabidopsis Information Resource) database. Functional classification is provided by MIPS (Munich Information Centre for Protein Sequences).

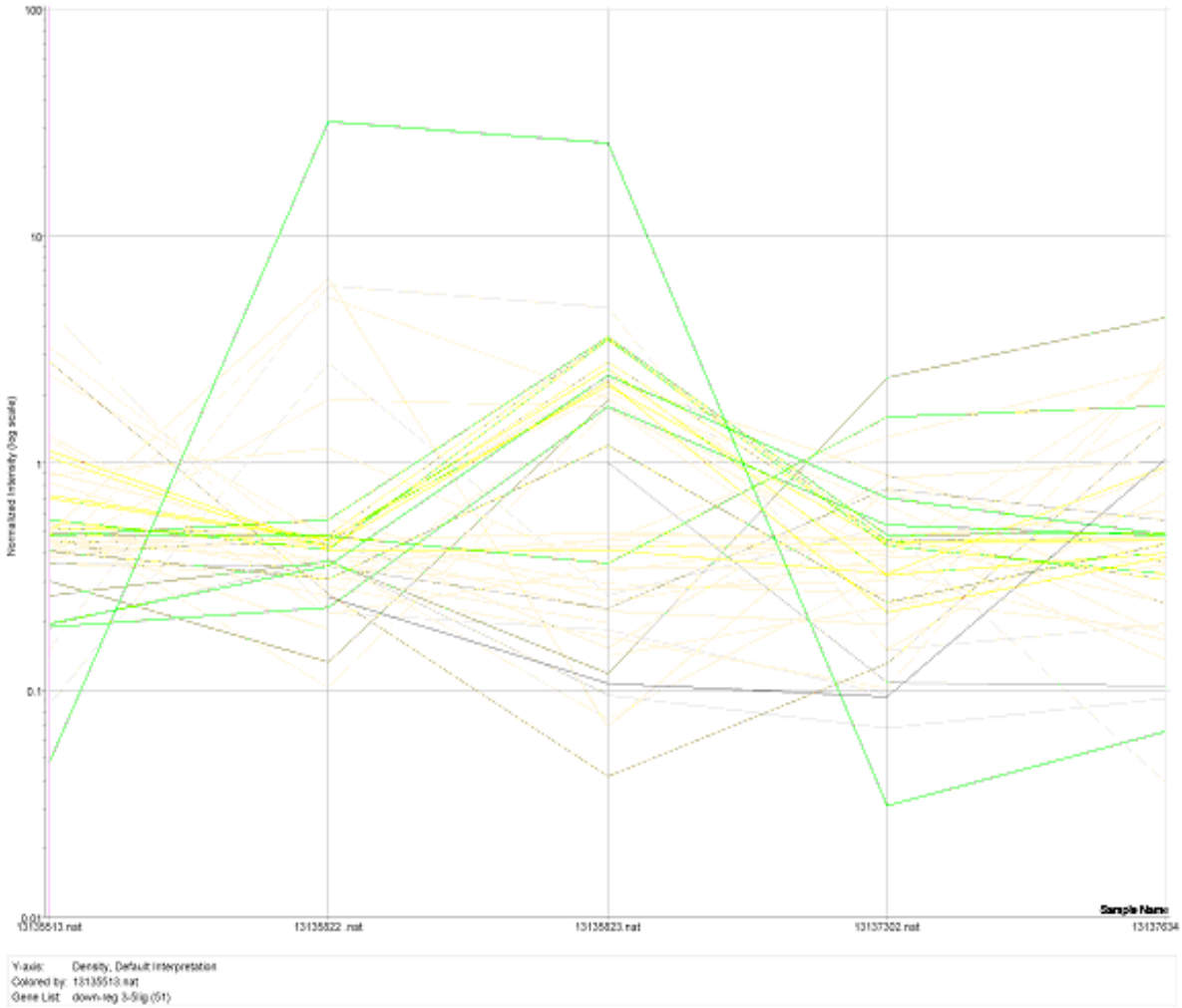


Figure 4.6: Two-fold down-regulated genes across 3 of 5 extreme lignin arrays from experiment one. The x-axis shows the five arrays that were selected for this analysis, and how the expression of the listed genes changes from one array to the next. The y-axis shows the level of expression, with anything above 2 considered up-regulated, and anything below 0.5 considered down-regulated.

Gene ID	**Putative annotations (*Arabidopsis thaliana*)**	**AGI #**	**MIPS functional group**
WS0262_A03	Osmotin-like protein (OSM34)	At4g11650	Interaction with the environment
WS01037_M17	Cytochrome P450 family protein	At3g52970	Protein with binding function; cellular transport
WS00914_D10	No significant hit		
WS00819_G02	No significant hit		
WS00828_A07	No significant hit		
WS0262_N08	Peroxidase 12 (PER12)	At1g71695	Cell rescue,

			defence and virulence
WS00926_D01	Expressed protein	At5g50730	
WS0083_D23	Two-component responsive regulator / response regulator 4 (ARR4)	At1g10470	Transcription
WS00929_F11	Malate oxidoreductase	At1g79750	
WS00927_O09	Quercetin 3-O-methyltransferase 1 (OMT1)	At5g54160	
WS0085_O15	Expressed protein	At2g20830	Unclassified
WS00912_O21	No significant hit		
WS00925_J22	O-methyltransferase family 2 protein	At1g51990	Metabolism (phenylpropanoid)
WS00925_D02	Aquaglyceroporin / NOD26-like major intrinsic protein 2 (NLM2)	At4g18910	Cellular transport
WS0089_N16	Inositol-1(or 4)-monophosphatase	At3g02870	Metabolism
WS02613_D23	No significant hit		
WS0261_A24	Quercetin 3-O-methyltransferase 1	At5g54160	
WS01024_E07	No significant hit		
WS0264_O12	No significant hit		
WS0264_C14	AWPM-19-like membrane family protein	At1g29520	Unclassified
WS0262_P08	Peroxidase 57 (PER57)	At5g17820	Cell rescue, defence and virulence
WS00941_N02	No significant hit		
WS00931_G03	Scarecrow-like transcription factor 3 (SCL3)	At1g50420	Transcription
WS00929_H14	DNA-binding bromodomain-containing protein	At3g27260	
WS01021_F20	No significant hit		
WS01016_H09	Peroxidase 57 (PER57)	At5g17820	Cell rescue, defence and virulence
WS0058_F20	No significant hit		
WS0041_D17	No significant hit		
WS0039_I22	Nodulin MtN21 family protein	At1g75500	Unclassified
WS00730_B03	Myrosinase-binding protein	At1g52030	Protein with binding function
WS0064_M15	pfkB-type carbohydrate kinase family protein	At5g58730	Metabolism
WS0063_N04	Protein kinase family protein	At4g21410	Metabolism
IS0012_N03	Jacalin lectin family protein	At1g60110	Unclassified
GFP	No significant hit		
E.coli fus-A	No significant hit		
WS0035_O02	Pyrophosphate--fructose-6-phosphate 1-phosphotransferase alpha subunit	At1g76550	Metabolism
WS0019_B11	No significant hit		
MGC15287	No significant hit		
WS00823_L11	17.4 kDa class I heat shock	At3g46230	Cell rescue,

	protein		defence and virulence; interaction with the environment
WS00823_J17	MADS-box protein (AGL20)	At2g45660	Transcription
WS00820_A15	No significant hit		
WS01013_F04	Leucine-rich repeat transmembrane protein kinase	At5g61480	Metabolism
WS00828_H13	No significant hit		
WS00826_N05	Expressed protein	At5g20540	
WS0078_I14	No significant hit		
WS0076_G13	AWPM-19-like membrane family protein	At1g29520	Unclassified
WS0075_N05	No significant hit		
WS00814_F24	Peroxidase 57 (PER57)	At5g17820	Cell rescue, defence and virulence
WS00810_J12	AWPM-19-like membrane family protein	At1g29520	Unclassified
WS0079_M05	Integral membrane family protein	At4g15610	Unclassified

Table 4.7: List of two-fold down-regulated genes across 3 of 5 extreme lignin arrays from experiment one. Putative functions were assigned based on a BLASTx search of the TAIR (The Arabidopsis Information Resource) database. Functional classification is provided by MIPS (Munich Information Centre for Protein Sequences).

4.3.2 Experiment two

4.3.2.1 Introduction

The microarray work for experiment two was completed at Southampton using the same 21.8 K spruce array as in experiment one.

4.3.2.2 RNA

All the samples used in the experiment were quantified using the NanoDrop® ND-1000 Spectrophotometer. The samples all contained a minimum of 40 µg of RNA, which is the amount required from one sample for hybridisation. 50% of the samples were run of an Agilent 2100 bioanalyser for quality control. Figure 4.15 shows the electrophoresis file run summary.

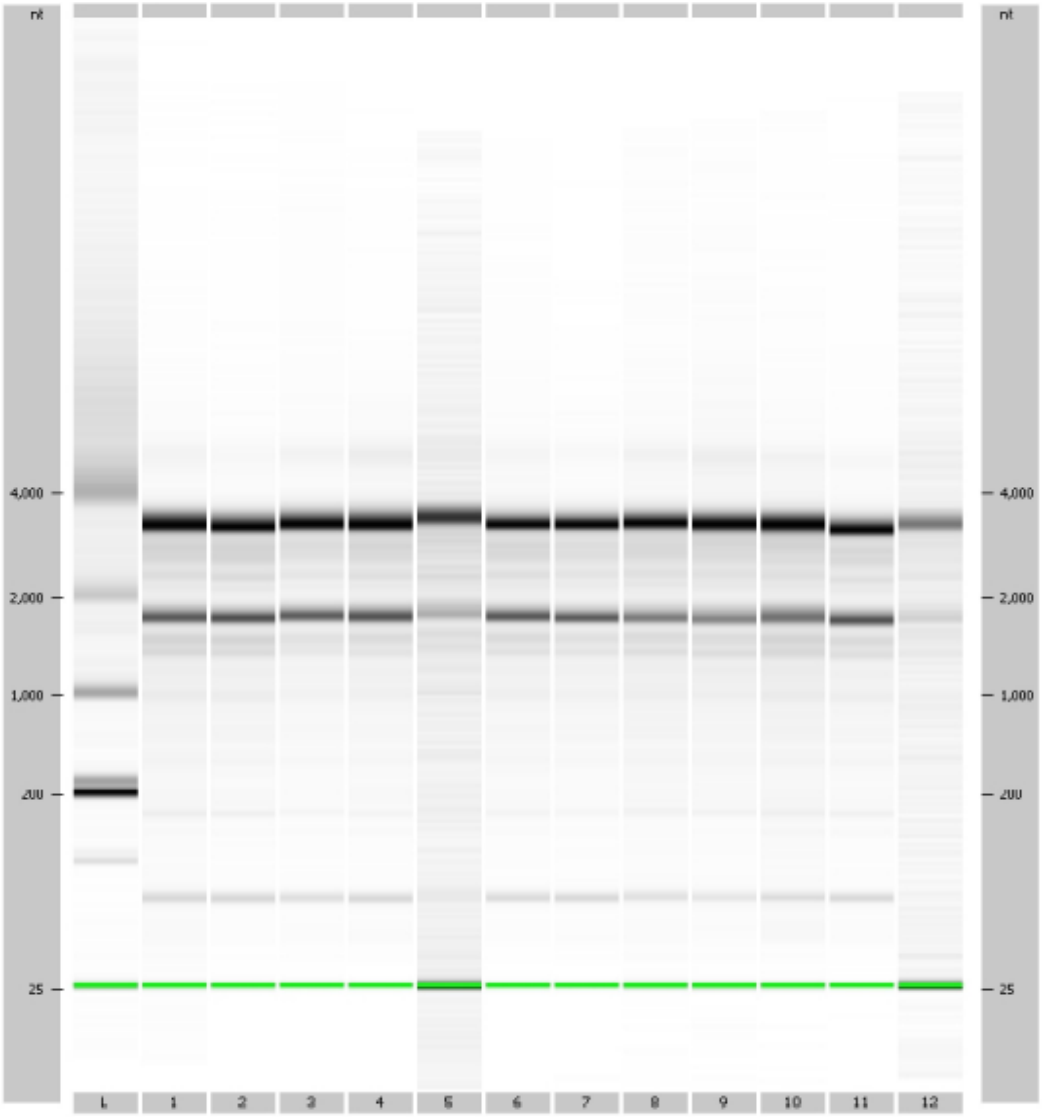


Figure 4.7: electrophoresis file run summary for experiment two. The two bands shown across all samples represent the 28S and 18S rRNA bands (for eukaryotic samples), indicating intact RNA.

In figure 4.7, well L contains the RNA 6000 ladder. The unknown samples are compared to this ladder to determine the RNA concentration and to identify the ribosomal peaks.

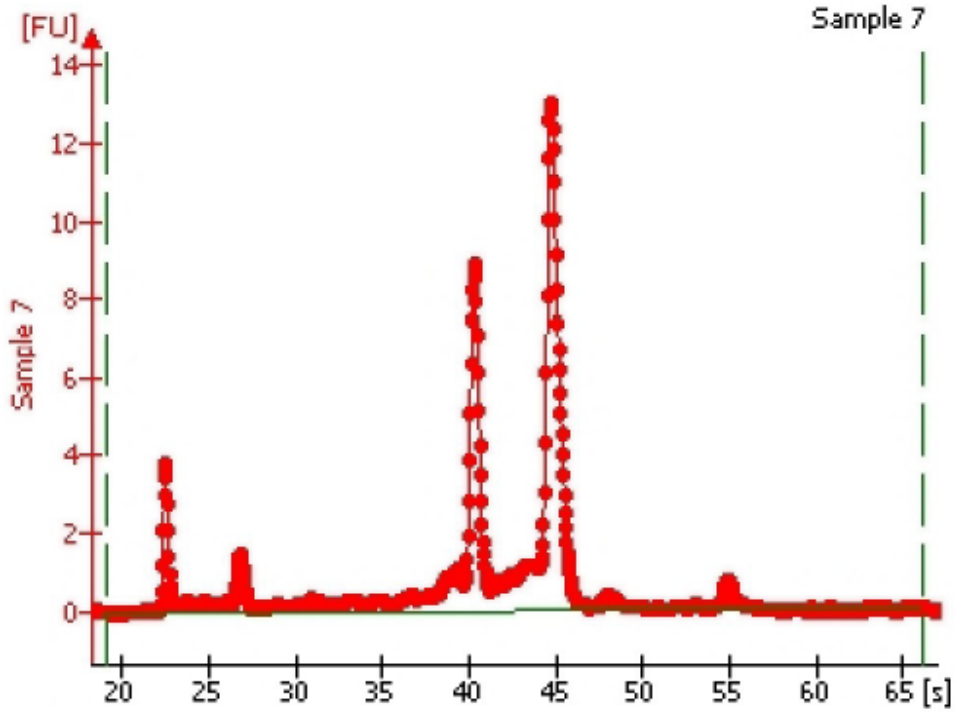


Figure 4.8: electropherogram from experiment two. A dye is added to each channel on the bioanalyser chip, and the samples are then detected by their fluorescence. The fluorescence of the sample is plotted on the y-axis, with time (in seconds) on the x-axis. The RNA quantity is measured by integrating the area under the peaks. The first large peak represents the 18S band, and the second large peak represents the 28S band. A 2:1 ratio of 28S:18S rRNA indicates intact RNA.

4.3.2.3 Microarrays

To compare analysis techniques, these arrays were analysed first using a b-stat calculation, and second using a fold-change calculation to generate gene lists. The b-stat calculation identified genes that have a 95% (or greater) certainty of being differentially expressed.

Gene ID	Putative annotations (*Arabidopsis thaliana*)	AGI#	Putative annotations (*Populus*)	NCBI accession number	MIPS functional group
WS0107_P09	No significant hit		No significant hit		
WS01037_C01	Oligosaccharyl transferase STT3 subunit family protein	At5g19690	No significant hit		Metabolism
WS01035_B16	Cytochrome P450	At4g36220	Cytochrome P450	ACE96926	Protein with binding function
WS01033_P09	No significant hit		No significant hit		
WS01030_M20	No significant hit		No significant hit		
WS01026_A13	Glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein	At3g48950	No significant hit		Metabolism
WS01024_O15	Expressed protein	At1g49170	No significant hit		Subcellular localisation
WS01012_N24	No significant hit		No significant hit		
WS0101_K05	No significant hit		No significant hit		
WS0099_E04	Strictosidine synthase family protein	At3g51440	Unknown	ABK96449	Metabolism
WS00939_C20	Lectin protein kinase	At4g02420	Lectin-like protein kinase	BAA82556	Metabolism
WS00939_B18	3' exoribonuclease family domain 1	At3g07750	Unknown	ABK95169	Transcription
WS00937_D23	Leucine-rich repeat family protein	At1g69545	TIR/NBS/LRR protein	CAC95124	
WS00936_P17	Lectin protein kinase	At4g02420	Lectin-like protein kinase	BAA82556	Metabolism
WS00929_G09	Latex-abundant family protein	At1g02170	Unknown	ABK95057	Unclassified
WS00920_P17	Peroxidase	At1g71695	Peroxidase	CAA66037	Cell rescue, defence and virulence
WS0092_A17	No significant hit		No significant hit		
WS00914_P22	No significant hit		No significant hit		
WS00825_J12	Histone H2A	At1g54690	Unknown	ABK95779	Protein with binding function

WS0058_H13	Histone H2A	At1g54690	Unknown	ABK95779	Protein with binding function
WS0042_M20	Leucine-rich repeat family protein	At1g69545	TIR/NBS/LRR protein	CAC95124	
WS0042_G13	Leucine-rich repeat family protein	At1g69545	TIR/NBS/LRR protein	ABF81427	
WS0018_C09	Transferase family protein	At5g23940	Unknown	ABK94894	Development
WS0011_D08	No significant hit		No significant hit		

Table 4.8: Up-regulated genes from b-stat analysis, experiment two. Putative functions were assigned based on a BLASTx search of the TAIR (The Arabidopsis Information Resource) database, and a BLASTx search of the NCBI (National Centre for Biotechnology Information) for comparison to Poplar sequences. Functional classification is provided by MIPS (Munich Information Centre for Protein Sequences).

Gene ID	Putative annotations (*Arabidopsis thaliana*)	AGI#	Putative annotations (*Populus*)	NCBI accession number	MIPS functional group
WS0269_N01	No significant hit		No significant hit		
WS0269_C15	No significant hit		No significant hit		
WS0261_K03	Epoxide hydrolase	At3g51000	Unknown	ABK96207	Metabolism
WS01030_I10	No significant hit		No significant hit		
WS01027_M13	No significant hit		No significant hit		
WS01024_K17	No significant hit		No significant hit		
WS01018_O06	Mitochondrial transcription termination factor-related	At5g64950	Unknown	ABK95347	Subcellular localisation
WS01012_C09	Senescence-associated family protein	At5g60220	Unknown	ABK95639	Cell fate
WS01011_A04	No significant hit		No significant hit		
WS01010_F11	No significant hit		NBS type disease resistance protein	ABF81447	Cell rescue, defence and virulence
WS01010_E16	No significant hit		No significant hit		
WS0097_O17	No significant hit		NBS type disease resistance protein	ABF81420	Cell rescue, defence and virulence
WS00938_E24	No significant hit		NBS type disease resistance protein	ABF81447	Cell rescue, defence and virulence
WS00931_I10	Double-stranded DNA-binding family protein	At1g29850	Unknown	ABK95623	Protein with binding function
WS00928_F16	Sulfotransferase family protein	At3g45070	Unknown	ABK94985	Metabolism
WS00920_J22	Jacalin lectin family protein	At1g19715	No significant hit		
WS00911_I10	No significant hit		No significant hit		
WS00910_L16	Pectinesterase family protein	At3g60730	Putative pectin methylesterase	CAC01624	Development; biogenesis of cellular components
WS00910_J21	Expressed protein	At3g58110	No significant hit		Subcellular localisation

WS0091_C09	No significant hit		No significant hit		
WS00821_B17	Expressed protein	At1g67790	No significant hit		Unclassified
WS0079_M09	No significant hit		Caffeoyl-CoA O-methyltransferase	AAF44689	Metabolism
WS0078_M14	Expressed protein	At3g44380	No significant hit		Unclassified
WS0075_I17	No significant hit		No significant hit		
WS0075_B04	Disease resistance protein	At4g27190	NBS-LRR type disease resistance protein	ABF81443	Cell rescue, defence and virulence
WS0073_A07	No significant hit		No significant hit		
WS00727_D12	UDP-glucoronosyl/UDP-glucosyl transferase family protein	At2g36790	Unknown	ABK96547	Metabolism
WS00720_G15	Jacalin lectin family protein	At1g19715	No significant hit		
WS0071_C13	UDP-glucoronosyl/UDP-glucosyl transferase family protein	At4g01070	Unknown	ABK96361	Metabolism
WS0063_K02	Jacalin lectin family protein	At1g19715	No significant hit		
WS0051_K21	SEC14 cytosolic factor	At1g01630	Unknown	ABK93168	Cellular transport
WS0046_D22	No significant hit		No significant hit		

Table 4.9: Down-regulated genes from b-stat analysis, experiment two. Putative functions were assigned based on a BLASTx search of the TAIR (The Arabidopsis Information Resource) database, and a BLASTx search of the NCBI (National Centre for Biotechnology Information) for comparison to Poplar sequences. Functional classification is provided by MIPS (Munich Information Centre for Protein Sequences).

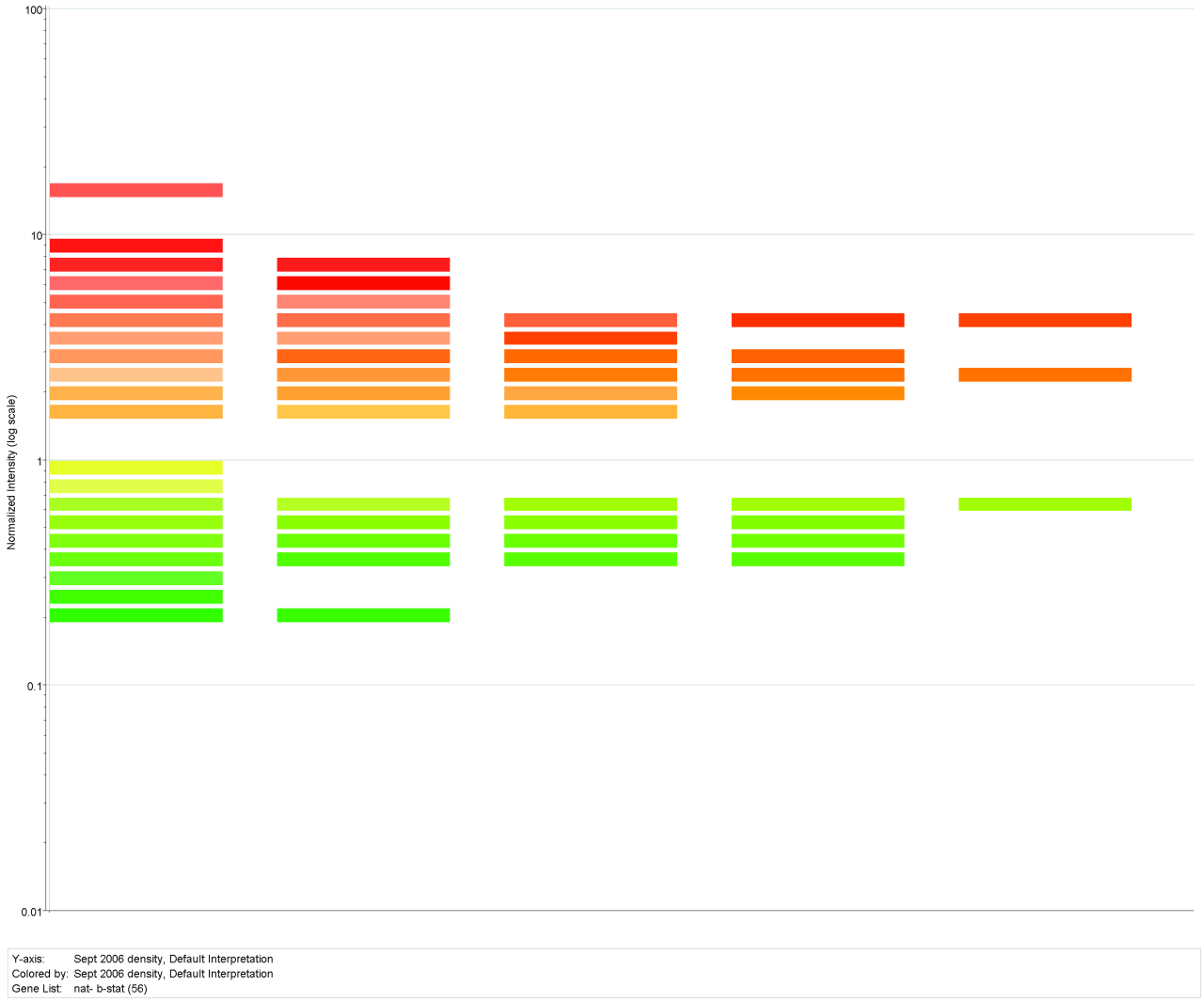


Figure 4.9: Differentially expressed genes from the b-stat analysis, experiment two. The red bars depict the up-regulated genes, and the green bars represent the down-regulated genes. The distance above 1 on the vertical axis shows the degree of up-regulation and the distance below 1 shows the degree of down-regulation of particular, identified genes. The distance along the x-axis is a measure of confidence.

4.3.3 Experiment three

4.3.3.1 Introduction

The microarray work for experiment three was completed at Southampton University using the same 21.8 K spruce array as in experiments one and two.

4.3.3.2 RNA

All the samples used in the experiment were quantified using the NanoDrop® ND-1000 Spectrophotometer. The samples all contained a minimum of 40 µg of RNA, which is the amount required from one sample for hybridisation. 100% of the samples were run of an Agilent 2100 bioanalyser for quality control. Figure 4.20 shows the electrophoresis file run summary.

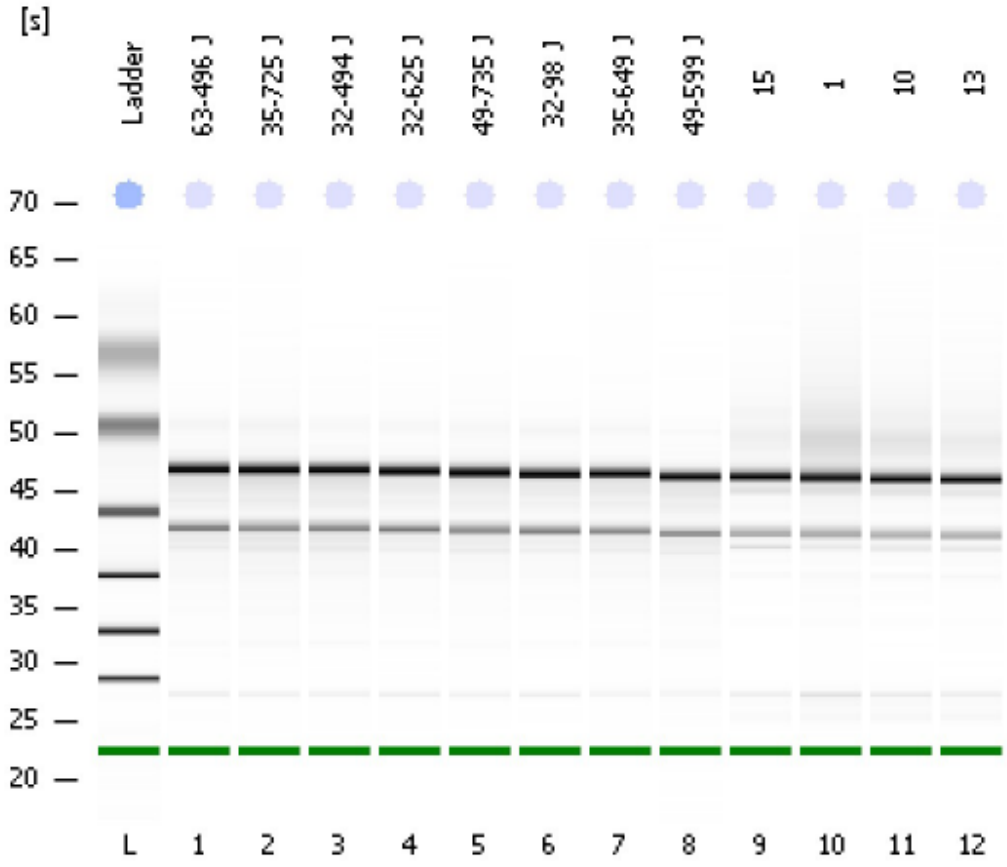


Figure 4.10: Electrophoresis file run summary for experiment three. The two bands shown across all samples represent the 28S and 18S rRNA bands (for eukaryotic samples), indicating intact RNA.

Figure 4.11 below shows the electropherogram for C20163-496 J, one of the samples used in experiment three.

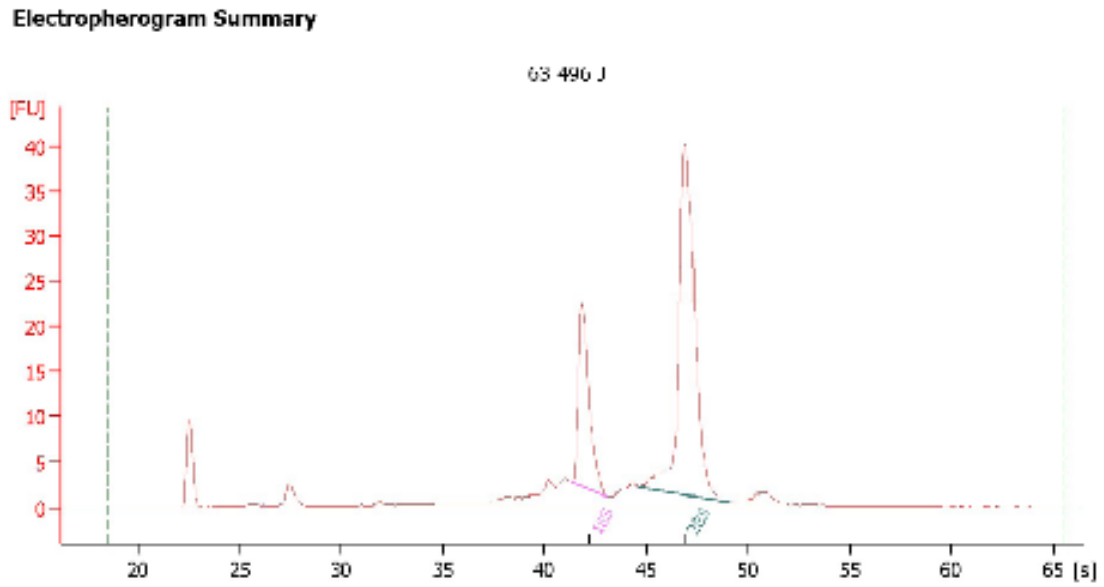


Figure 4.11: electropherogram from experiment three, sample C20163-496 J. The fluorescence of the sample is plotted on the y-axis, with time (in seconds) on the x-axis. The RNA quantity was estimated by integrating the area under the peaks. The first large peak represents the 18S band, and the second large peak represents the 28S band. A 2:1 ratio of 28S:18S rRNA indicates intact RNA.

4.3.3.3 Microarray

Gene lists were identified for this experiment using a 2-fold expression cut off. The genes identified in figure 4.12 were at least 2-fold up-regulated in all four arrays.

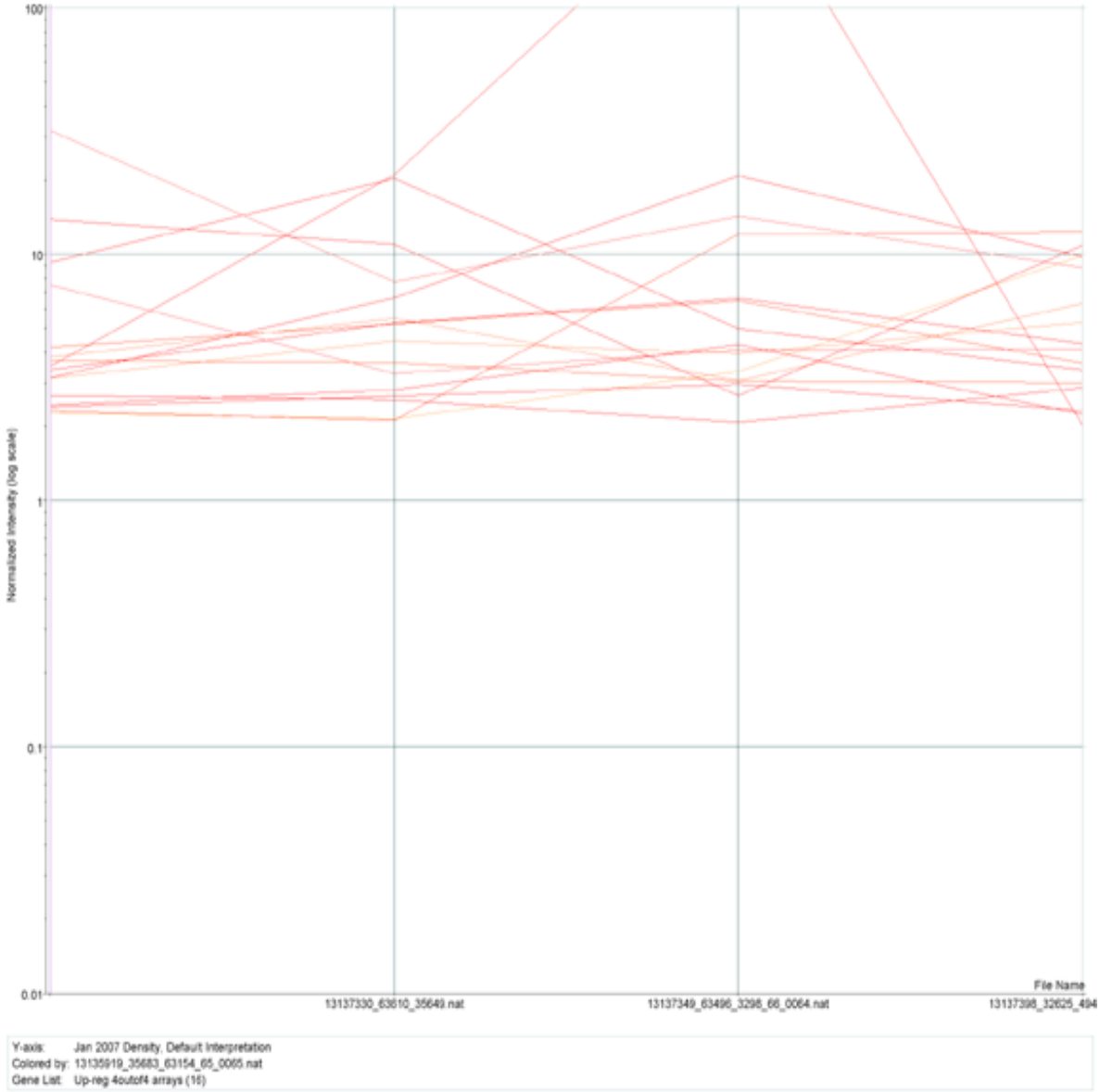


Figure 4.12: Genes 2-fold up-regulated across all four arrays from experiment three. The x-axis shows the four arrays that were selected for this analysis, and how the expression of the listed genes changes from one array to the next. The y-axis shows the level of expression, with anything above 2 considered up-regulated.

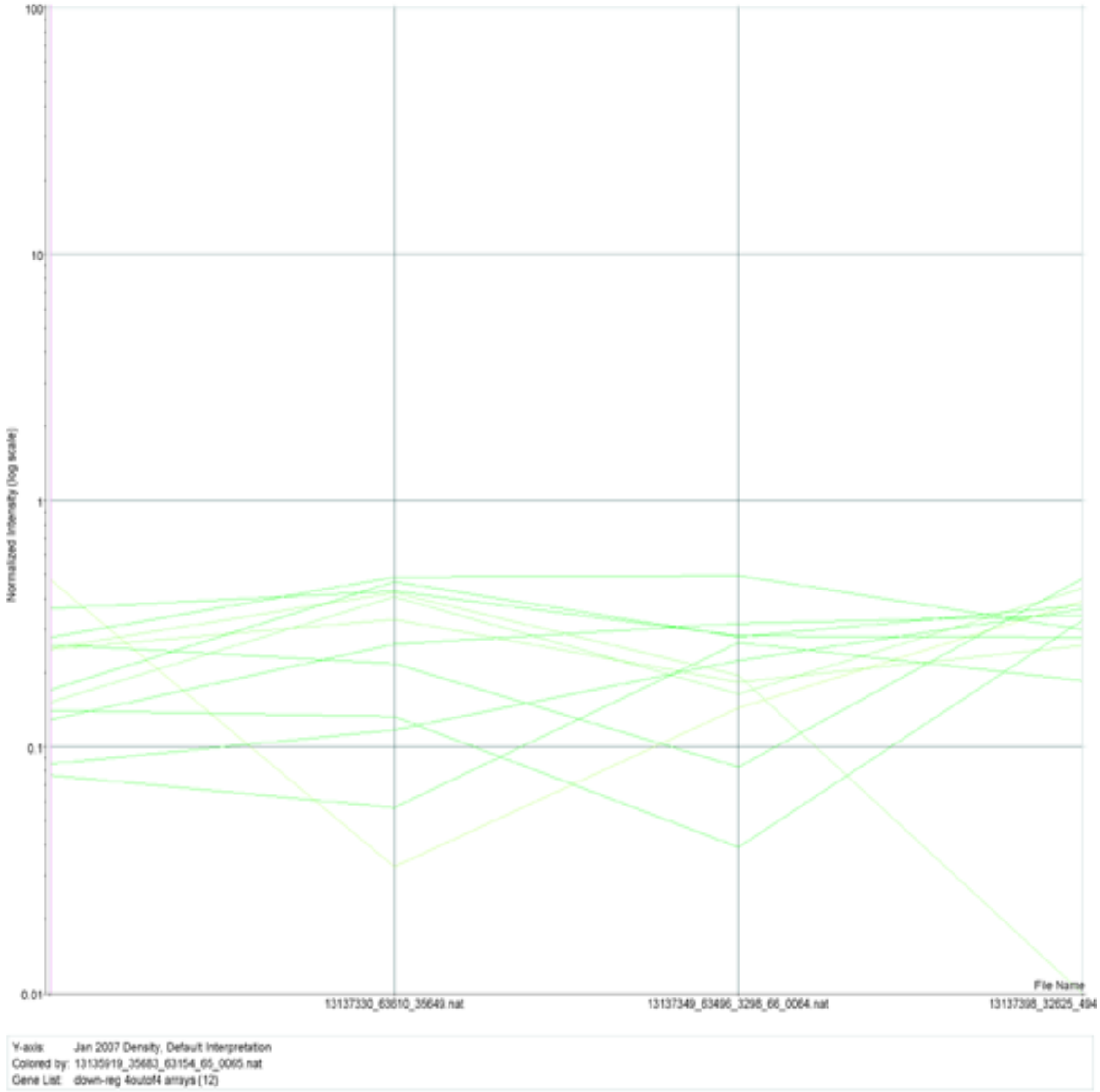


Figure 4.13: Genes 2-fold down-regulated across all four arrays from experiment three. The x-axis shows the four arrays that were selected for this analysis, and how the expression of the listed genes changes from one array to the next. The y-axis shows the level of expression, with anything below 0.5 considered down-regulated.

Gene ID	Putative annotations (*Arabidopsis thaliana*)	AGI#	Putative annotations (*Populus*)	NCBI accession number	MIPS functional group
WS01024_N23	No significant hit		No significant hit		
WS01012_N24	No significant hit		No significant hit		
WS00931_C04	Expansin, putative (EXP8)	At2g40610	Alpha-expansin 2	AAR09169	Cell fate – cell growth
WS00926_E21	Aminoacylase, putative / N-acyl-L-amino-acid amidohydrolase	At4g38220	Unknown	ABK94846	
WS02613_E18	No significant hit		No significant hit		
WS0261_O15	No significant hit		No significant hit		
WS01033_K02	Leucine-rich repeat transmembrane protein kinase	At5g63930	Leucine rich repeat protein	ACE97245	
WS01031_O02	GDSL-motif lipase/hydrolase family protein	At1g54790	No significant hit		Metabolism
WS0054_L12	No significant hit		No significant hit		
WS0013_M05	Amino acid permease, putative (AUX1)	At2g38120	AUX1-like protein	AAF21982	Cellular transport
WS0012_J09	Peroxidase	At5g05340	Peroxidase	AAX53172	Cell rescue, defence and virulence
IS0014_E17	Leucine-rich repeat family protein	At5g49760	Unknown	ABK93750	Metabolism
WS00914_P22	No significant hit		No significant hit		
WS0078_G10	Leucine-rich repeat protein kinase	At2g28970	Unknown	ABK93453	Metabolism
WS00730_N19	No significant hit		No significant hit		
WS0072_J11	No significant hit		No significant hit		

Table 4.10: Up-regulated genes from 2-fold analysis, experiment three. Putative functions were assigned based on a BLASTx search of the TAIR (The Arabidopsis Information Resource) database, and a BLASTx search of the NCBI (National Centre for Biotechnology Information) for comparison to Poplar sequences. Functional classification is provided by MIPS (Munich Information Centre for Protein Sequences).

Gene ID	Putative annotations (*Arabidopsis thaliana*)	AGI #	Putative annotations (*Populus*)	NCBI accession number	MIPS functional group
WS01010_E16	No significant hit		No significant hit		
WS00943_D03	Expressed protein	At5g44450	No significant hit		Unclassified
WS0094_N01	No significant hit		No significant hit		
WS0264_B09	Protease inhibitor/seed storage/lipid transfer protein (LTP)	At3g22620	Unknown	ABK94530	Unclassified
WS0262_P08	Peroxidase 57 (PER57)	At5g17820	Peroxidase	BAE16616	Cell rescue, defence and virulence
WS01010_F11	No significant hit		NBS type disease resistance protein	ABF81447	Cell rescue, defence and virulence
WS00927_O09	Quercetin 3-O-methyltransferase 1	At5g54160	Caffeic acid 3-O-methytransferase	ACF04799	Metabolism
WS00814_E06	No significant hit		No significant hit		
WS00729_E03	No significant hit		No significant hit		
WS00939_A02	No significant hit		No significant hit		
WS00938_E24	No significant hit		NBS type disease resistance protein	ABF81447	Cell rescue, defence and virulence
WS0093_M06	Glycosyl hydrolase family 3 protein	At5g49360	Xylan 1,4-beta-xylosidase	CAJ65922	Metabolism

Table 4.11: Down-regulated genes from 2-fold analysis, experiment three. Putative functions were assigned based on a BLASTx search of the TAIR (The Arabidopsis Information Resource) database, and a BLASTx search of the NCBI (National Centre for Biotechnology Information) for comparison to Poplar sequences. Functional classification is provided by MIPS (Munich Information Centre for Protein Sequences).

4.4 Discussion

Many sources of variation can lead to differences in wood density within a population of trees, and also within a tree itself. This variation can be either genetic or environmental. Environmental variables can be silvicultural (forest management practices), such as spacing, thinning, or pruning; or site-specific, such as weather conditions, slope, etc. (MacDonald and Hubert, 2002).

Variation within a tree can arise from several sources, such as juvenile wood compared to mature wood. Changes in climate within growing season mean that the early wood formed during the first part of the growing season has different characteristics to the late wood formed during the latter part of the growing season (Gion *et al*., 2005). To remove variability with regards to juvenile wood, mature wood, early wood and late wood, samples were collected at the same height and orientation for every tree and each microarray run was based on samples that were all collected at the same time of year.

In this study we have attempted to remove sources of environmental variation to enable us to study the genetic component of wood formation, using a 21.8K spruce microarray. Analysis of expression patterns has revealed genes of interest in each of the Sitka spruce experiments.

4.4.1 Experiment One

No differentially expressed genes were detected using the UBC analysis and this may indicate that this analytical approach was overly stringent. The rationale of this approach was to search for a correlation between gene expression and wood density in the samples applied to the arrays. A second, less stringent approach was therefore also adopted that was based on the fold-change method. Other studies have used fold-change criteria to identify genes of interest, for example Hertzberg *et al*. (2001), used a 1.8 fold cut-off to determine significant expression.

A total of 30 genes were up-regulated in at least five out of the nine arrays that had a large difference (more than 100 kg/m^3) in wood density between the two hybridised samples. Sixteen of these genes (53%) could be assigned a putative function. The up-regulated genes span a range of MIPS functional groups, including defence (such as the peroxidases) and metabolism.

In total, 31 genes were over two-fold down-regulated in five of the same nine high-density arrays. Twenty of these (65%) could be assigned a putative function. One of the down-regulated genes was a putative caffeoyl-CoA 3-0-methyltransferase, which has a role in the lignin biosynthetic pathway (Whetten and Sederoff, 1995).

WS01033_C04, also down-regulated in five of the nine arrays, shows homology to a putative fructokinase-like enzyme up-regulated in tension wood of poplar (Andersson-Gunnerås *et al*., 2006). This up-regulation is thought to lead to increased levels of cellulose being deposited in the cell walls, matched by a decrease in lignin deposition (Andersson-Gunnerås et al., 2006).

Five arrays were then selected based on the lignin content of the samples hybridised for a lignin analysis. 80 genes were up-regulated in three out of the five arrays, and 47 genes were down-regulated. There is only one up-regulated gene from the lignin biosynthetic pathway (caffeoyl-CoA 3-O-methyltransferase; Whetten and Sederoff, 1995). The other up-regulated genes are involved in defense and metabolism.

The down-regulated genes from the five lignin arrays have putative functions involved with defence, metabolism and transcription.

4.4.2 Experiment Two

A total of 24 genes were up-regulated in the experiment 2 arrays. Sixteen of these genes (67%) could be assigned a putative function. The up-regulated

genes span a range of MIPS functional groups, including metabolism, transcription and development.

An up-regulated putative transferase protein (WS0018_C09) shows homology to a transferase that is up-regulated in *Arabidopsis thaliana* flowers when compared to leaves. The up-regulated putative peroxidase (WS00920_P17) shows homology to a peroxidase isolated from poplar xylem (Christensen *et al*., 2001) that is thought to have a function in secondary cell wall metabolism. Syringaldzine undergoes peroxidase-directed oxidation, and this process has been isolated in cells that undergo secondary cell wall formation (Christensen *et al*., 2001).

In total, 32 genes were down-regulated in the experiment 2 arrays. Sixteen of these (50%) could be assigned a putative function. Several of the down-regulated genes have been assigned putative functions in defence, such as the NBS/LRR (nucleotide binding site/leucine rich repeat) resistance genes. These genes showed homology to defence genes identified in white poplar (Zhang *et al*., 2008), that are thought to have a role in resistance to diseases attacking the organs.

4.4.3 Experiment Three

A total of 16 genes were up-regulated in the experiment 3 arrays. Eight of these genes (50%) could be assigned a putative function. The up-regulated genes span a range of MIPS functional groups, including metabolism and cell growth.

The up-regulated putative expansin (WS00931_C04) shows homology to an expansin identified in poplar (Gray-Mitsumune *et al*., 2004), shown to be abundant in the secondary xylem and a member of the α-expansin gene family. Expansins reside in the cell wall and aid in the formation of the cellulose-hemicellulose network found in primary walls (Gray-Mitsumune *et al*., 2004).

Expansins are considered to be important regulators of cell wall properties, by catalysing cell wall extension (Darley *et al*., 2001).

Auxin has previously been identified as a regulator of wood formation in poplar (Moyle *et al*., 2002). WS0013_M05 shows homology with a putative auxin permease, *AUX1* (Swarup *et al*., 2004), whose function is to transport IAA (indole-3-acetic acid). IAA transport through polar and phloem channels is dependent on *AUX1* activity (Swarup *et al*., 2004).

In total, 12 genes were down-regulated in the experiment 3 arrays. Five of these (42%) could be assigned a putative function. One of these down-regulated genes shows homology to a NBS/LRR resistance gene (Zhang *et al*., 2008). Putative NBS/LRR genes were also down-regulated in the experiment 2 arrays.

Genes involved with senescence and disease resistance have been identified in other microarray experiments. Schrader et al. (2004b) studied gene expression differences between active and dormant poplar cambial tissue, showing that senescence proteins and a disease resistance protein were up-regulated in dormancy.

4.5 Conclusion

When focusing on the first Sitka spruce experiment, a considerable amount of overlap is observed to occur between the up-regulated and down-regulated gene lists for both the density analyses and the lignin analyses. This first experiment was based on samples collected in May 2004, which may have been rather early in the year for a tree to be fully expressing the genes involved in wood production, as the trees may have been just emerging from dormancy. In retrospect, it would have been better to have used samples that had been collected later in the growing season. Indeed, an examination of the literature reveals that other studies have used samples collected during the height of the growing season (Schrader *et al*., 2004a). The decision was therefore made to use samples collected during July 2004 for the second and third Sitka experiments.

The clonal aspect of the experimental design for experiment one may have contributed to the inconsistencies in gene expression, as some of the clones had a very wide spread of density and/or lignin measurements. For this reason the two clones that were used for experiment two were selected because they had a much smaller range of density measurements, so showed greater consistency in physiology across all of the replicates.

Some of the inconsistencies in the expression data may be caused by a lack of precision in the sampling technique. Hertzberg *et al*. (2001) used tangential sectioning to cover five positions in developmental sequence, and so they were able to separate samples into meristematic cells, early expansion, late expansion, secondary wall formation and late cell maturation for transcript profiling (Hertzberg et al. 2001). Tangential sectioning has its own associated difficulties, as every tree has a different thickness of cambium, making it difficult to match individual samples together (Schrader *et al*., 2004a). However, unlike Hertzberg *et al*. (2001) whose objective was to examine differences in expression

of genes in the different regions of the developing xylem and cambial tissues, our goal was to search for differences of expression between high and low wood density genotypes. Our approach of using the combined tissues from across the cambial and developing xylem was therefore valid. With the limited availability of microarrays it was not practicable to incorporate the added complexity of tissue type into our experimental design. However, now that we have some idea of which genes may be up-regulated in high wood density types it would be possible to determine which of the tissue types exhibit the greatest difference in expression of these genes in high vs. low wood density samples by using the Herztberg *et al*. (2001) tangential sectioning method.

44% of the genes which showed differential expression had no known function. It has been suggested that sequences with no homology to other ESTs may correspond to rare transcripts making it unlikely that they will have been sequenced in other EST projects to date (Gion *et al*., 2005). In Sterky *et al*. (1998), working on large-scale production of ESTs from the wood-forming tissues of two poplars, 12% of their ESTs showed no similarity to any others in sequencing projects to date, and it was suggested that these ESTs may have a specific role in wood formation. However, these were some of the first large scale ESTs established from wood forming tissues and our knowledge has advanced since this work was done.

The next step in this study was to see if any of the genes generated from experiments two and three would also be identified from a similar experiment in a different species. Poplar was chosen for this next experiment, as its sequenced genome and large numbers of publicly available annotated ESTs makes it an ideal candidate for gene expression studies.

5 Poplar Gene Expression

5.1 Introduction

The aim of this study is to determine the link between gene expression and wood density in poplar. This will be compared with the genes lists generated from Sitka samples with any genes identified in this experiment. Species comparison between Sitka and poplar offers a greater stringency when identifying genes of interest, and this will be the first study to use this technique.

The Sitka experiment and the poplar experiment differ in terms of genotypes, age, silviculture, etc. Sitka samples were collected when the trees were approximately 15-years-old, whereas the poplar trees were coppiced in 2000, so were sampled when producing wood that may have an equivalent density to that of 4-year-old juvenile wood. This may have an effect on any gene expression patterns, as juvenile wood tends to have a lower density than mature wood. It would have been ideal to have an exact comparison with both sets of trees producing wood of the same age.

Many studies into the gene expression of wood formation have been conducted on a range of species including eucalyptus (Kirst *et al*., 2004; Paux *et al*., 2004), pine (Allona et al., 1998; Kirst *et al*., 2003), Arabidopsis (Chaffey *et al*., 2002; Zhao *et al*., 2005) and poplar (Hertzberg *et al*., 2001; Schrader *et al*., 2004a; Sterky *et al*., 1998). The poplar studies have arisen from the large EST resource now available for poplar (Sterky *et al*., 2004).

Poplar is an ideal species to use for genetic studies due to its relatively small, sequenced genome (50 times smaller than pine, ref. International *Populus* Genome Consortium), ease of clonal replication (Sterky *et al*., 2004) and rapid growth (Rae *et al*., 2004).

The objective of the work completed in this chapter was to determine whether there were any consistent differences in the gene expression in the cambial tissues of poplars belonging to family 331 which exhibited contrasting wood density properties. This study therefore uses RNA extracted from the cambium of high and low wood density poplar genotypes applied to Affymetrix microarrays to identify genes involved in the determination of high wood density. Family 331, based at Headley, UK, was used for the purpose of this study.

The Family 331 F_2 population was created from a cross between a female *P. trichocarpa* and a male *P. deltoides*. From the resulting F1 family, two siblings were crossed to form an F2 family of 90 genotypes (in 1988) and 320 genotypes (in 1990) (Rae *et al*., 2008). The pedigree was imported into the UK in 1999, and planting took place at Headley in May 2000 (Rae *et al*., 2004).

5.2 GeneChip® Poplar Genome Array

The GeneChip® Poplar Genome Array, produced by Affymetrix, was chosen for use in this study. Affymetrix use light-directed oligonucleotide synthesis to produce their arrays (Lipshutz *et al*., 1999). GeneChip® arrays are 'single channel' arrays, therefore only one sample is hybridised per chip, and analysis is performed between arrays.

GeneChip® arrays use perfect match (PM) and mismatch (MM) probes to act as specificity controls (Lipshutz *et al*., 1999). The MM probes are an exact match to their partner PM probe except for a change to a single base in the centre of the probe. This probe pairing is used to detect non-specific hybridisation.

The poplar array contains 61,251 probe sets representing 56,055 transcripts, and is based on mRNAs and ESTs from GenBank® for all Populus species up to April 26, 2005. Additional information on the poplar array is available on the Affymetrix website (www.affymetrix.com).

5.3 Materials and Methods

5.3.1 Physiological Measurement

Family 331 was planted at Headley in May 2000. 300 full-sib progeny genotypes were planted in three randomised blocks, one replicate from each genotype per block (Rae et al., 2004). A double row of buffer trees (*P. deltoides* x *Populus nigra* L. 'Gaver') was planted round the outside of the trial, and irrigation of the site occurring during the growing season as necessary (Rae *et al*., 2004). Blocks A and B had last been coppiced in 2002, and block C in 2004. The replicates from blocks A and B were chosen for this experiment because they provided older growing material. Growth measurements were collected during January 2006.

Height measurements were collected for the tallest stem of each replicate (identified visually) using a measuring pole. Stem diameter (to the nearest millimetre) was measured for the tallest stem at a height of 1m using digital callipers. The total number of stems per replicate was counted. All the trees from blocks A and B were then coppiced. A small section (approx. 10cm long) was removed from each main stem at a height of 1.3 m for the purpose of wood density determination. These sections were sent to Forintek, Canada for density determination.

5.3.2 Experimental Design

Of the 300 genotypes from family 331, 84 genotypes were represented in both block A and block B during July 2006 when the cambium samples were collected. The other 216 genotypes were either only represented in one block or no blocks (due to failure to recover from coppicing). On the basis of wood density, six genotypes (three from each extreme of the wood density distribution) were selected for microarray analysis, totalling 12 samples (two replicates of each genotype, one from block A and one from block B). The high density genotypes were selected on the basis that both replicates had an outer ring density that exceeded 450 kg/m^3, and the low density genotypes were selected when both

replicates had an outer ring density lower than 0.360 g/cm^3. The family range for wood density was from 0.28 to 0.54 g/cm^3.

GENOTYPE	BLOCK	HEIGHT (m)	DIAMETER (cm)	HIGH/LOW WOOD DENSITY CATEGORY	DENSITY (g/cm^3)
1082	A	5.32	36.90	HIGH	0.453
1082	B	4.95	35.30	HIGH	0.504
1731	A	5.31	38.1	HIGH	0.474
1731	B	3.52	23.5	HIGH	0.500
1630	A	4.95	35.70	HIGH	0.473
1630	B	5.90	41.20	HIGH	0.497
1650	A	3.16	19.30	LOW	0.295
1650	B	4.59	26.60	LOW	0.318
1809	A	3.81	27.00	LOW	0.348
1809	B	4.14	31.20	LOW	0.360
1866	A	4.49	35.00	LOW	0.350
1866	B	4.08	29.50	LOW	0.334

Table 5.1: Measurements of growth and wood density for the high and low density selected poplar genotypes.

RNA was extracted from each replicate, and the genotypes were then pooled into four groups for hybridisation to the four Affymetrix arrays available. The decision to pool samples was made in order to maximise the number of replicates that could be included into the experimental design. Peng *et al*. (2003) have shown that pooling of RNA is a statistically valid method for microarray experiments. Each pool contained RNA from two genotypes, with two replicates representing each genotype. The genotypes were grouped as follows:
High pool one (HP1) – 1630A, 1630B, 1082A and 1082B
High pool two (HP2) – 1630A, 1630B, 1731A and 1731B

Low pool one (LP1) – 1809A, 1809B, 1866A and 1866B
Low pool two (LP2) – 1650A, 1650B, 1866A and 1866B

5.3.3 RNA extraction and Microarrays

The RNA was extracted, quantified and quality checked as detailed in section 2.2. The selected samples were pooled as described above and sent to the Nottingham Arabidopsis Stock Centre (NASC) for hybridisation to Affymetrix GeneChips. A single sample (or in this case, pooled samples) of RNA was hybridised to a single chip using a single-colour detection system (Do and Choi, 2006).

5.4 Results

5.4.1 Physiological data

The following graphs present the growth data collected for family 331 prior to the 2006 coppicing. It is necessary to assess these correlations to ensure the selection of samples where the density is not being affected by other factors (e.g. block, growth rate, etc.).

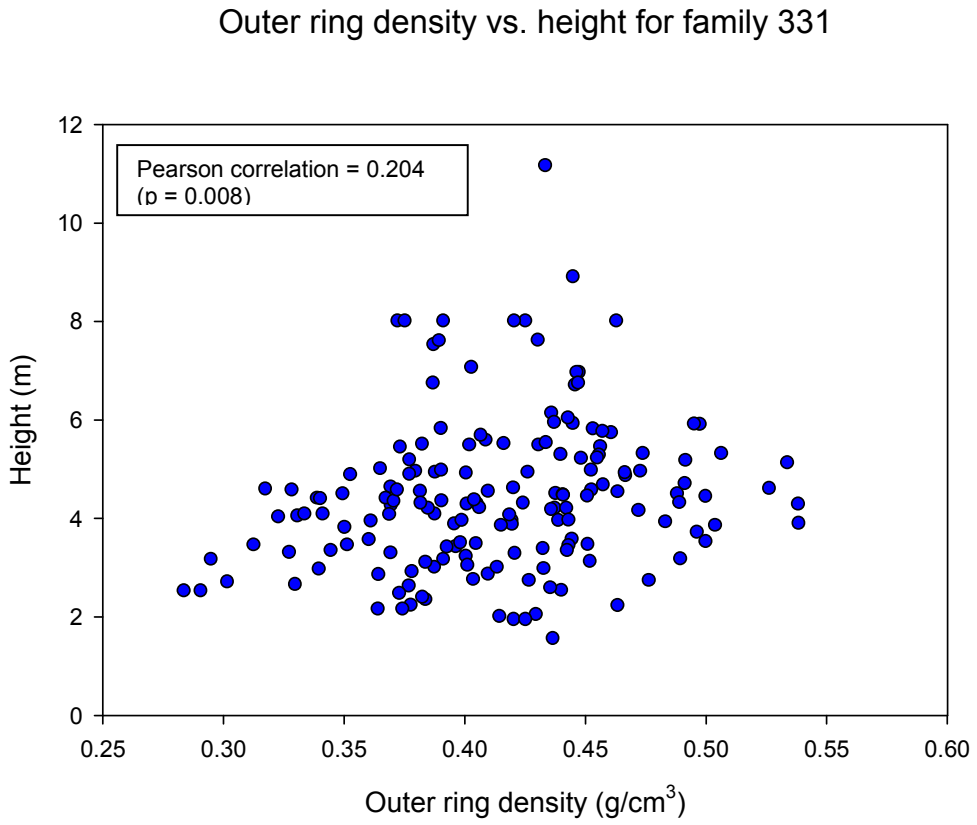


Figure 5.1: Outer ring density measurements plotted against main stem height for each individual in family 331 (blocks A and B).

Figure 5.1 shows that there is little correlation between outer ring density and height (r = 0.204). Therefore, genotypes which re-grow after coppicing do not necessarily produce low density wood.

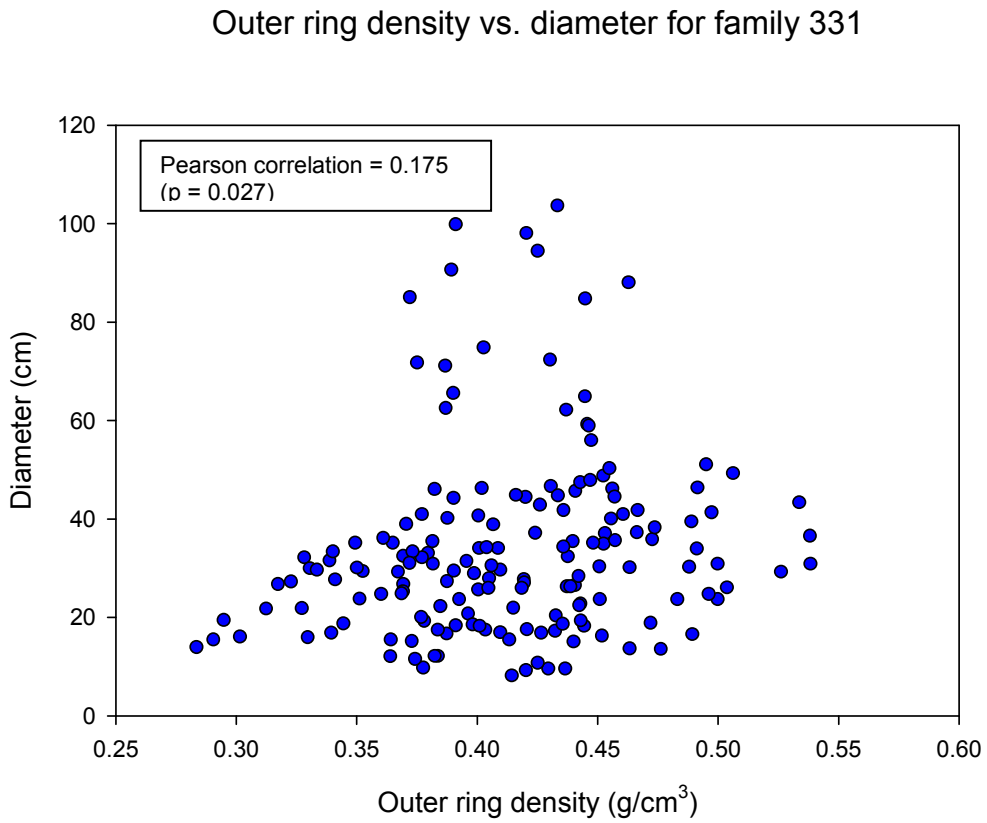


Figure 5.2: Outer ring density measurements plotted against main stem diameter for each individual in family 331 (blocks A and B). There is no correlation between average density and diameter (r = 0.175).

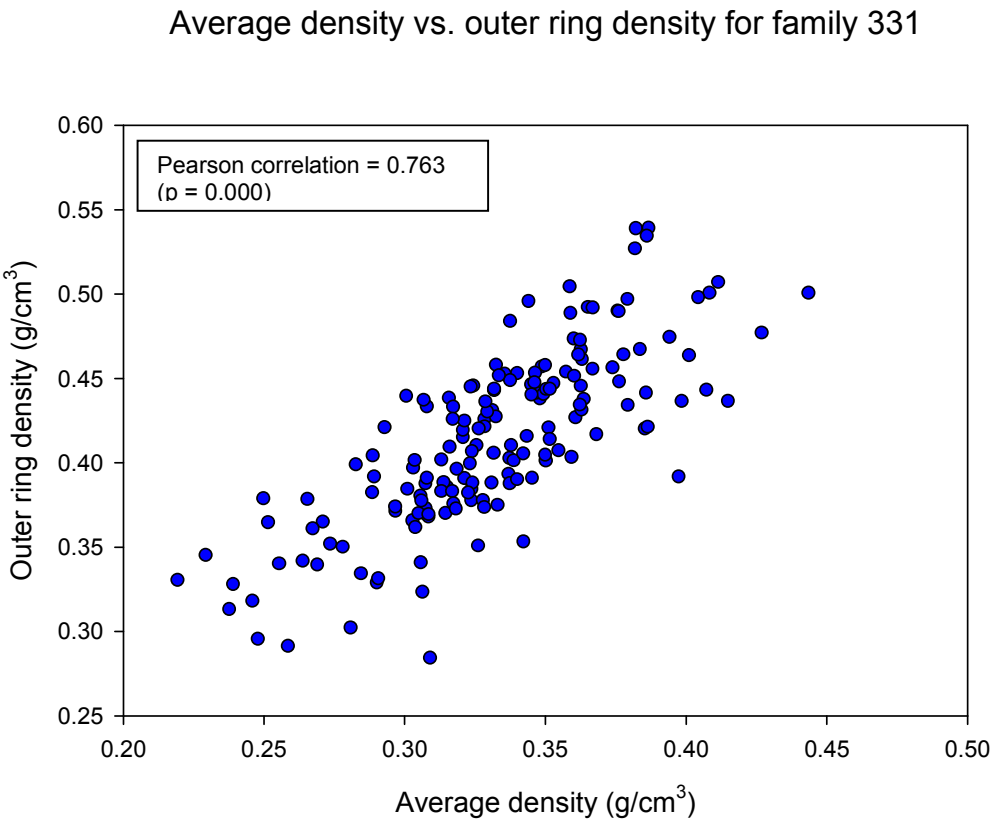


Figure 5.3: Average density plotted against outer ring density for each individual in family 331(blocks A and B). There is a strong positive correlation between the two measurements (r = 0.763).

It was necessary to check that other variables (for example block, position, growth rate, etc.) were not affecting the density measurements for the trees in family 331. A one-way ANOVA for outer ring density (the measurement that was used to select appropriate genotypes) with block as a factor showed no significant effect (p = 0.246). There was also no significant association between block and height (one-way ANOVA, p = 0.370), and between block and diameter (one-way ANOVA, p = 0.784).

Outer ring density vs. diameter for the selected genotypes from family 331

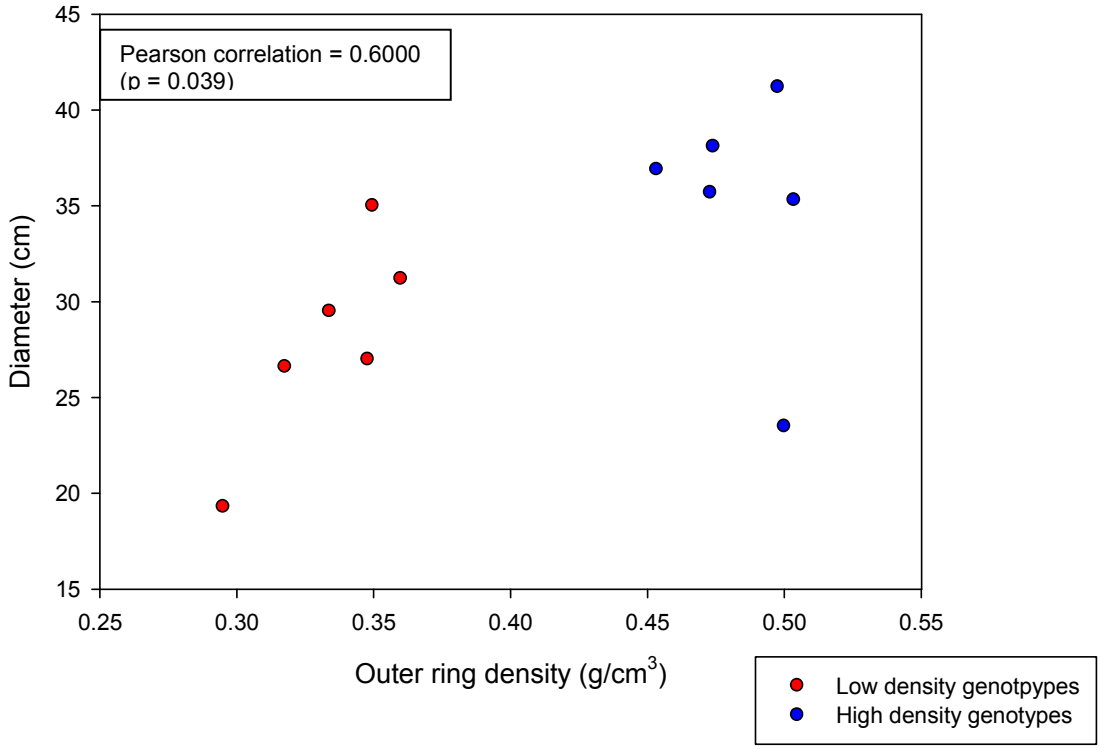


Figure 5.4: Outer ring density plotted against diameter for the selected genotypes in family 331 (blocks A and B). There is a correlation between outer ring density and diameter for the selected genotypes (r = 0.600, p = 0.039).

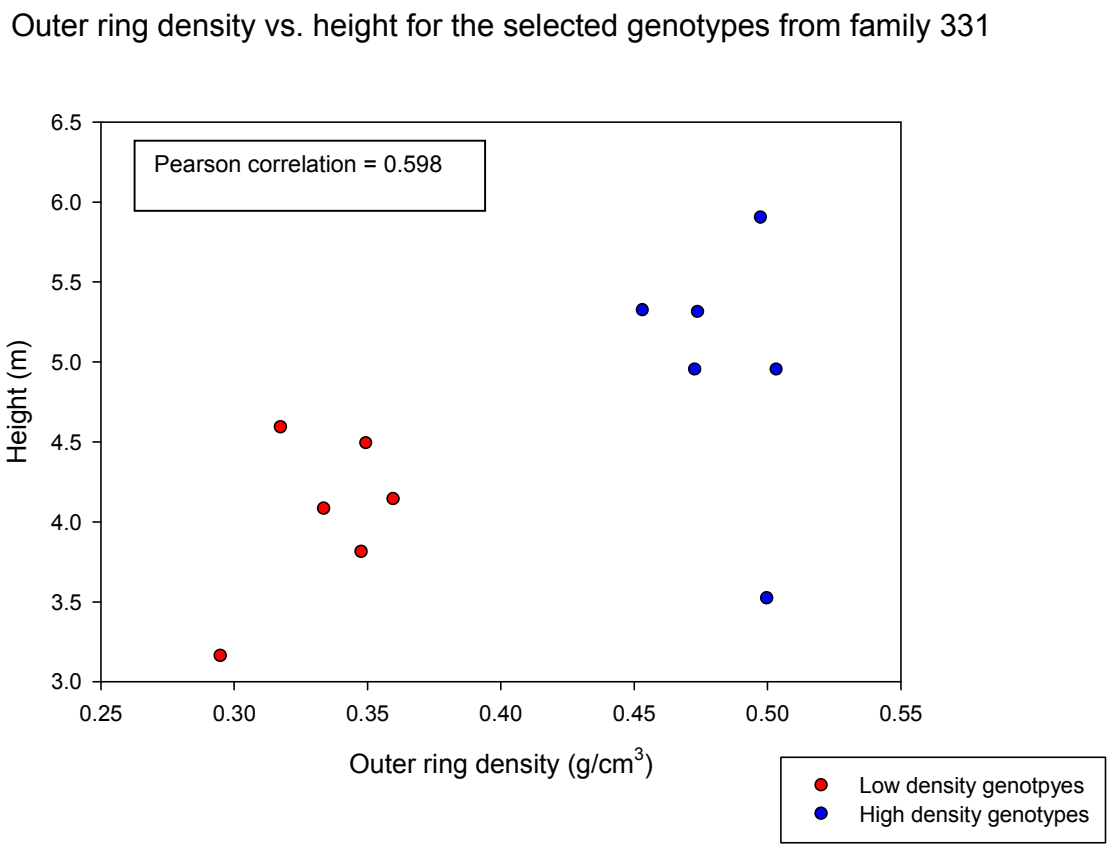


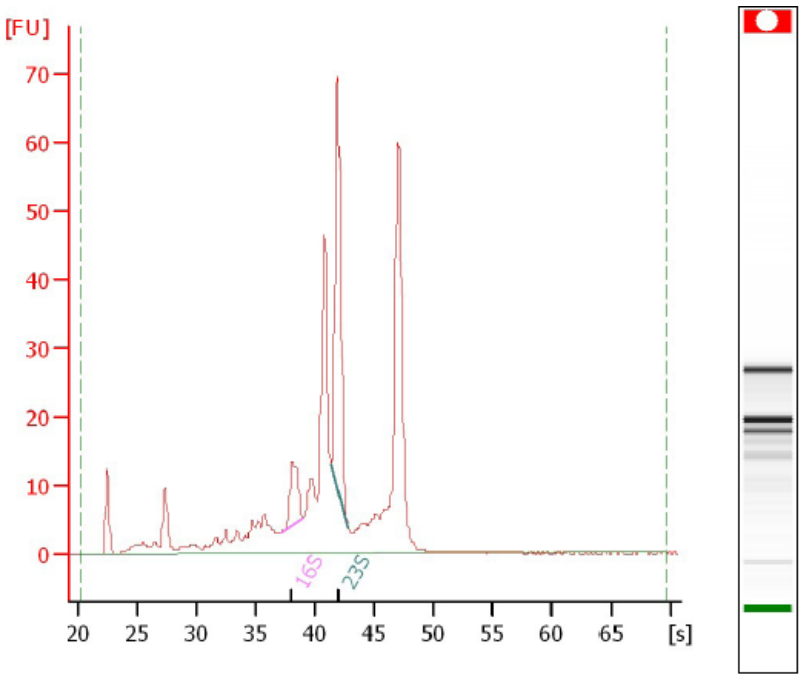
Figure 5.5: Outer ring density plotted against height for the selected genotypes in family 331 (blocks A and B). There is a correlation between outer ring density and height for the selected genotypes (r = 0.598).

Was there a statistically significant difference in density between the two groups of selected genotypes? A one-way ANOVA showed that there was a significant difference in outer ring density between the high and low genotypes. Were these differences in density being explained by difference in growth rate (height and diameter)? A one-way ANOVA showed that there was a significant difference in height measurements between the high and low genotypes ($p = 0.079$), but it was the high density genotypes that had the larger diameters. The same was true with regards to the height of the individuals. A one-way ANOVA showed that the high-density genotypes were significantly taller than the low density genotypes ($p = 0.041$). These genotypes would not have been suitable for this experiment had the low density pool had significantly higher growth rates than the high density pool, as the differences could have accounted for the variation in

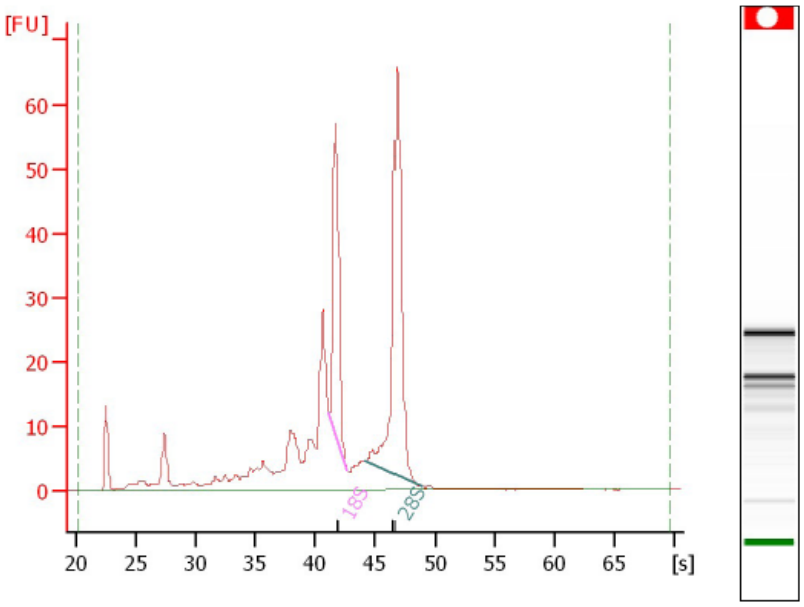
density. In this case, the high density pool had the higher growth rates, so the genotypes selected were suitable for further analysis.

5.4.2 RNA extraction

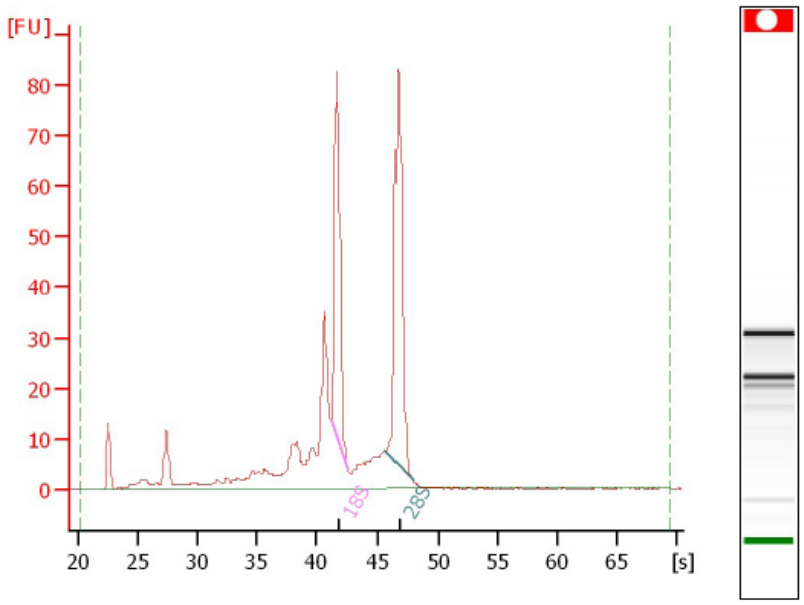
a) Low Pool 1



b) Low Pool 2



c) High Pool 1



d) High Pool 2

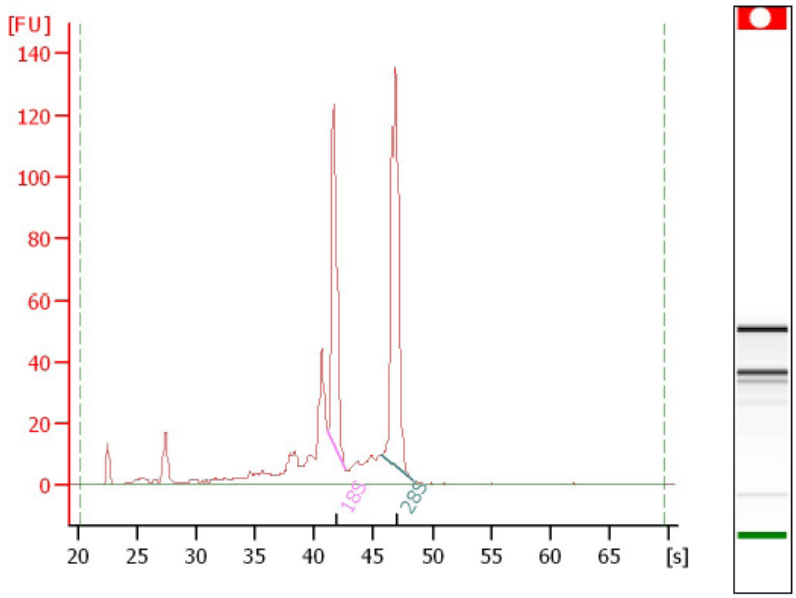


Figure 5.6: Electropherograms from each pool of RNA (A is low density pool 1, B low density pool 2, C high density pool 1 and D high density pool 2). The fluorescence emitted by the sample is plotted on the y-axis, with time (in seconds) on the x-axis. The RNA quantity was estimated by integrating the area under the peaks. The first large peak represents the 18S band, and the second large peak represents the 28S band. A 2:1 ratio of 28S:18S rRNA indicates intact RNA.

5.4.3 Microarray data

Each pool of RNA (HP1, HP2, LP1 and LP2) was hybridised to a separate poplar array (four in total). A robust multi-array average (RMA) correction (Irizarry *et al*., 2003) was applied to the data, which was then normalised by scaling using the median signal intensity (Kreil and Russell, 2005) to allow for overall differences in signal strength. This corrected for differences in hybridisation efficiency across different arrays belonging to the same experiment (Hoffmann *et al*., 2002).

A gene on an Affymetrix GeneChip is represented by 11-20 oligonucleotide probe pairs (Irizarry *et al*., 2003), with each probe pair consisting of a perfect match (PM probe) and a mismatch (MM) probe (used for measuring non-specific binding). Intensity values are recorded as the average difference between the PM probe and the MM probe for a set of probes specific to a gene (Hoffmann *et al*., 2002).

For this experiment, a t-test for differential expression was applied to all the genes (significant to a 5% level) as well as a two-fold cut off for added stringency. The following figures represent the differentially expressed genes identified.

Figure 5.7: Genes up-regulated in the high-density arrays. The left-hand side of the chart shows the expression levels of the 119 up-regulated genes in the high-pool 1 and high-pool 2 arrays. The right-hand side of the chart shows the corresponding down-regulated expression levels of the same genes in the low-pool 1 and low-pool 2 arrays. The y-axis shows the expression level for each gene, with anything above 2 considered up-regulated, and anything below 0.5 considered down regulated.

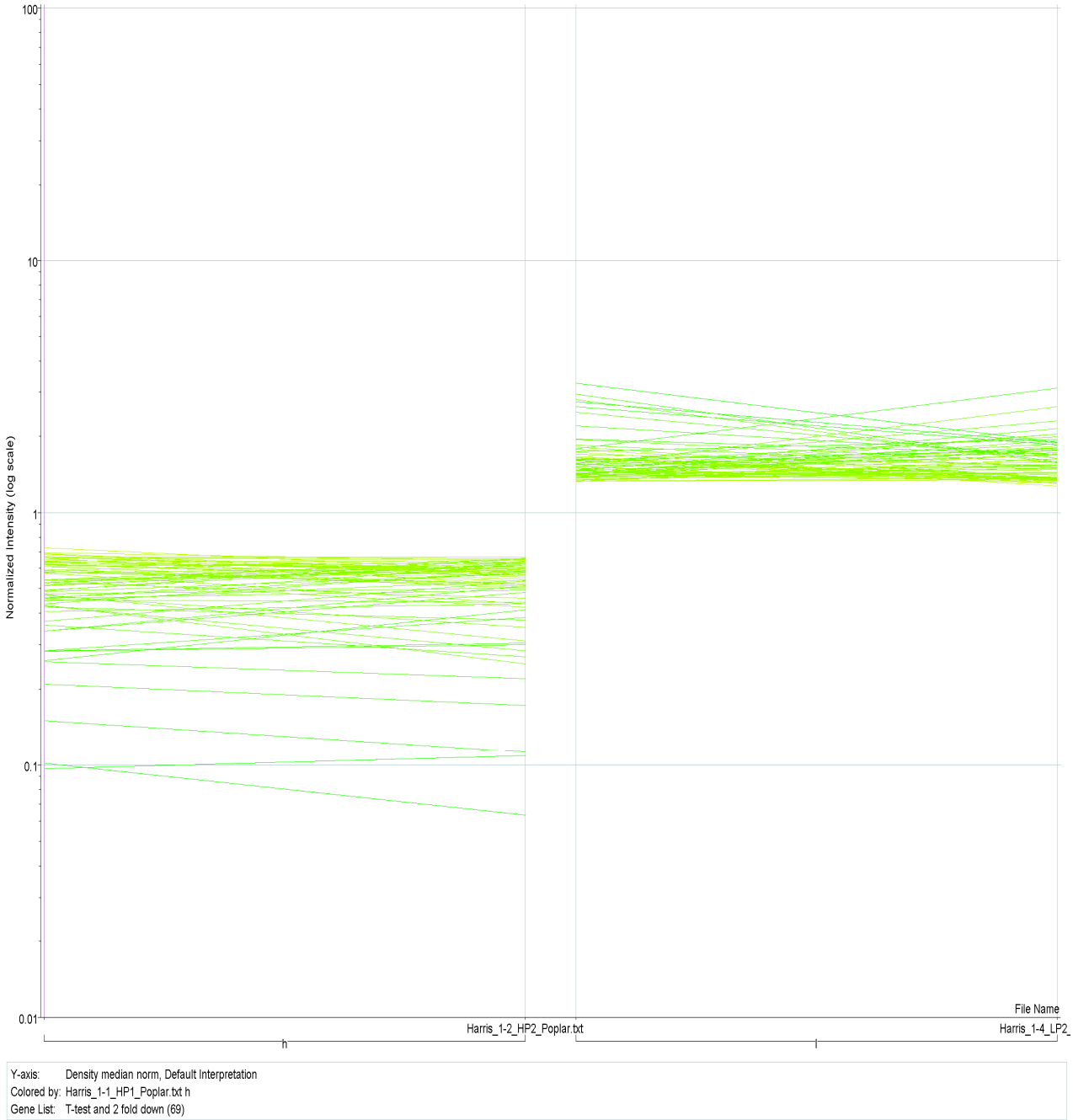


Figure 5.8: Genes down-regulated in the high-density arrays. The left-hand side of the chart shows the expression levels of the 69 down-regulated genes in the high-pool 1 and high-pool 2 arrays. The right-hand side of the chart shows the corresponding up-regulated expression levels of the same genes in the low-pool 1 and low-pool 2 arrays. The y-axis shows the expression level for each gene, with anything above 2 considered up-regulated, and anything below 0.5 considered down regulated.

Figures 5.8 and 5.9 show the gene expression levels for the differentially expressed genes from the poplar arrays. These genes were annotated by comparison with *Arabidopsis thaliana*, JGI best hit and TGI annotation. The tables below list the differentially expressed annotated genes.

Gene ID	**Putative annotation**	**NCBI accession number/AGI #**	**MIPS functional group**
PtpAffx.98646.1.S1_at	Expressed protein		
PtpAffx.206300.1.S1_s_at	Acetyl co-enzyme A carboxylase carboxyltransferase		
PtpAffx.9932.2.A1_a_at	Pectin methylesterase	AK222211	
Ptp.6857.1.S1_s_at	B-like cyclin	AT1G76310	Cell cycle and DNA processing
PtpAffx.31067.1.A1_at	Hypothetical protein	AT1G21510	Unclassified
PtpAffx.217542.1.S1_s_at	Ta-11 related non-LTR retroelement protein		
Ptp.3885.1.S1_s_at	Expressed protein		
Ptp.2060.1.S1_at	Hypothetical protein		
PtpAffx.75907.1.A1_at	Copper-exporting ATPase		
Ptp.2667.1.S1_s_at	Arabinogalactan protein		
PtpAffx.138263.1.S1_at	Zinc finger (C3HC4-type RING finger) family protein	AT5G55970	Protein with binding function
PtpAffx.29726.1.A1_s_at	SNAP25 homologous protein	AAM62553	
PtpAffx.92655.1.A1_at	Kinesin motor-related protein		
Ptp.5620.1.S1_s_at	SEC14 cytosolic factor family protein		
PtpAffx.66534.1.S1_s_at	Expressed protein		
PtpAffx.100235.1.S1_s_at	AUX1-like permease	AT1G77690	Cellular transport

PtpAffx.56737.1.A1_at	WD-40 repeat family protein		
PtpAffx.72161.1.A1_s_at	Reticulon family protein		
PtpAffx.215594.1.S1_at	Unknown protein		
PtpAffx.212682.1.S1_at	Kinesin motor family protein (NACK1)		
Ptp.4899.1.S1_s_at	Myb-related protein	AT1G75250	Transcription
PtpAffx.78267.1.S1_s_at	Seven transmembrane MLO family protein, mildew resistance	AT4G02600	Cell rescue, defence and virulence
PtpAffx.95744.1.A1_at	Ribosomal protein L23A		
PtpAffx.47796.1.S1_s_at	Alkaline alpha galactosidase		

Table 5.2: List of annotated genes (putative functions) up-regulated in the high-density arrays. These genes are also down-regulated in the low-density arrays.

Gene ID	**Putative annotations**	**NCBI accession number/AGI #**	**MIPS functional group**
PtpAffx.120153.1.S1_s_at	Xyloglucan endotransglucosylase	AT5G65730	Metabolism
Ptp.868.1.S1_s_at	Protein phosphatase 2C	AT2G29380	Metabolism; protein fate
Ptp.6069.1.S1_a_at	Auxin-responsive GH3 family protein	At1g23160	Interaction with the environment
PtpAffx.54126.1.S1_s_at	NAM-like protein		
Ptp.386.2.A1_s_at	RNA binding		
PtpAffx.249.58.A1_x_at	Aquaporin 2	AAM63463	
PtpAffx.5968.2.S1_s_at	AT-rich element binding factor 3		
PtpAffx.146058.1.A1_at	Dof zinc finger protein		
Ptp.5268.1.S1_at	Acyl-CoA oxidase		
PtpAffx.4736.1.A1_s_at	Expressed protein		
PtpAffx.30722.1.A1_s_at	Coronatine-insensitive 1		
PtpAffx.127847.1.A1_at	DNA binding		

PtpAffx.51460.1.S1_a_at	Expressed protein		
PtpAffx.54186.2.S1_s_at	Histidine-containing phosphotransfer protein		
PtpAffx.19246.1.S1_at	BEL1-like homeodomain transcription factor		
PtpAffx.249.332.S1_at	Caffeic acid O-methyltransferase	AT4G35160	Metabolism (carbohydrate)
Ptp.7956.2.S1_s_at	Oxidoreductase, 2OG-Fe(II) oxygenase family protein	AT5G54000	Unclassified
PtpAffx.249.195.S1_s_at	40S ribosomal protein S9		
PtpAffx.10247.1.A1_a_at	Calcium-binding protein		
Ptp.4050.17.A1_at	Metallothionein		
Ptp.6093.1.S1_at	Protein kinase family protein	AT5G58350	Metabolism; protein fate
PtpAffx.18850.1.A1_at	Harpin-induced family protein		
PtpAffx.11491.1.S1_a_at	Aquaporin NIP1.2		
Ptp.1490.1.S1_at	Subtilisin-like protease		
PtpAffx.1659.1.A1_s_at	Ribosomal protein L14		
PtpAffx.10032.1.S1_a_at	H+-pyrophosphatase	BAE98517	
PtpAffx.23427.1.S1_s_at	Glutathione S-transferase GST 18	AT2G29420	Metabolism; defense
PtpAffx.10247.1.A1_at	Calcium-binding protein		
Ptp.5242.3.S1_s_at	Chlorophyll a/b binding protein CP24		
PtpAffx.80487.1.A1_at	Gag-pol protein		
PtpAffx.42878.1.A1_at	Scarecrow-like 1	AT1G21450	Transcription
PtpAffx.11491.2.A1_a_at	Aquaglyceroporin		
PtpAffx.754.1.A1_s_at	Oxidoreductase 20G-Fe(II) oxygenase family protein		
PtpAffx.6959.1.S1_at	Glutaredoxin family protein	AT5G11930	Cellular transport; defence
Ptp.3400.1.S1_s_at	Glutamate		

	decarboxylase		
PtpAffx.31202.1.S1_at	Homeobox protein knotted-1-like LET12		

Table 5.3: List of annotated genes (putative functions) down-regulated in the high-density arrays. These genes are also up-regulated in the low-density arrays.

Figure 5.8 lists the genes that were up-regulated in the high-density genotype arrays. Of the 119 identified, only 24 (20.2%) have been successfully annotated.

Ptp.6857.1.S1_s_at shows homology to a putative B-like cyclin in *Arabidopsis thaliana*, CYCB2;4 (Vandepoele *et al*., 2002). These cyclins help regulate progression through the eukaryotic cell cycle (Vandepoele *et al*., 2002).

As with the Sitka experiments, one of the up-regulated genes from the poplar experiment showed homology to an auxin influx carrier, LAX3 (Swarap *et al*., 2008). LAX3 has been shown to regulate the expression of several cell-wall-remodelling enzymes (Swarap *et al*., 2008).

Also in the up-regulated gene list, Ptp.4899.1.S1_s_at shows homology to a putative MYB protein (At1g75250) in Arabidopsis thaliana (Yanhui *et al*., 2006). MYB proteins have roles in defence and developmental processes (Yanhui *et al*., 2006).

Figure 5.9 lists the genes that were down-regulated in the high-density genotype arrays. Of the 69 identified, 36 (52.2%) have been successfully annotated.

A putative caffeic acid O-methyltransferase (PtpAffx.249.332.S1_at) was down-regulated in the high-density poplar arrays. This gene is involved in the lignin biosynthesis pathway (Whetten and Sederoff, 1995). Kirst *et al*. (2004) showed that the expression of genes involved with lignin biosynthesis is negatively correlated with growth in Eucalyptus. This is reflected by the Sitka growth

measurements from Chapter 3, where the high-density selected clones had a significantly lower lignin content than the low-density selected clones.

PtpAffx.120153.1.S1_s_at shows homology with a xyloglucan endotransglucosylase (XET) from *Arabidopsis thaliana* (Yokoyama and Nishitani, 2001). These XETs are involved in cell-wall construction (Yokoyama and Nishitani, 2001).

Also in the down-regulated list, PtpAffx.23427.1.S1_s_at shows homology with a glutathione *S*-transferase (Wagner *et al*., 2002). Glutathione S-transferases have roles in plant defence, specifically detoxification (Wagner et al., 2002).

5.5 Discussion

This experiment has used microarray analysis to identify genes that were differentially expressed in high density and low density samples of coppiced hybrid poplar. By using a two-fold cut off and a t-test for added stringency, 119 up-regulated and 69 down-regulated genes were identified.

Studies in other species such as Eucayptus (Paux *et al*., 2004; Kirst *et al*., 2004) have shown some overlap with the results generated from this study. Genes involved with lignin biosynthesis look to be an important area on which to focus, as a gene from the lignin pathway was down-regulated in high-density arrays. This result matches that of a wood formation study in Eucalyptus (Kirst et al., 2004). Hu *et al*. (1999) showed that repression of lignin biosynthesis (specifically a 4-coumarate: coenzyme A ligase) resulted in reduced lignin content, increased cellulose content and increased stem growth in poplar. It is possible that by decreasing lignin biosynthesis more carbon is available for growth (Kirst *et al*., 2004).

Auxin has previously been shown to be associated with wood formation in a study in Eucalyptus (Paux *et al*., 2004). Its function as a positional signal in pattern formation in plants was determined in a study on *Pinus sylvestris* (Uggla et al., 1996). Auxin is critical for plant growth and the control of developmental processes (Woodward and Bartel, 2005), and its role could be an area of focus for future studies on wood formation.

The up-regulated pectin methylesterase could have an important impact on wood density, as it is currently thought to determine fibre width and length (Siedlecka *et al*., 2008). In Eucalyptus, an increase in basic density from pith to bark is matched by an increase in fibre length (Igartúa *et al*., 2003). Up-regulation of this gene leading to an increase in fibre length could be linked to increased wood density.

In total, 60 out of the 188 genes identified by this study have been annotated. Other studies have also generated lists of differentially expressed genes in relation to wood formation with no sequence similarity to ESTs currently available. These genes should be an area of interest for future studies as they could provide new candidate genes involved in wood formation.

6 General Discussion

6.1 Overview of results

This project is the first to use newly available microarray technology to study differential gene expression in cambial tissue of high versus low wood density field grown samples from two contrasting species, Sitka spruce (gymnosperms) and poplar (angiosperms). Gene expression was studied in these samples using a 21.8K spruce array and an Affymetrix Poplar Genome Array. During the course of the work, a method for the extraction of high quality RNA was successfully optimised for spruce and poplar cambial tissue and this provided good quality RNA which was essential for the gene expression studies. Careful consideration was given to the choice of the most appropriate material to sample for the differential expression studies. Also, with experience, it was possible to improve the design of the experiments to make optimal use of the samples by making the best pairings of samples for application to the microarrays. This analysis succeeded in producing a list of genes that show differential expression between high density and low wood density samples. Although the function of many of these candidate genes remains unknown, it was possible to attribute a likely function to many of them. It was of considerable interest that several of the candidate genes appeared on the lists for both spruce and poplar. This is despite the fact that these are taxonomically very different species and the samples for the two species were grown under very different conditions.

6.1.1 Sitka spruce (Chapter 4)

The first Sitka experiment produced some rather contradictory results which were difficult to reconcile and explain. None of the genes showed consistent levels of expression across all nine of the density arrays, and those genes that did pass the (less-stringent) filter showed large variation in their expression levels. The gene expression levels for experiments two and three were much more

consistent, as some gene showed differential expression across all of the arrays in an experiment.

There are several reasons why this variation may have occurred. The samples used in this experiment were collected in May 2004, a time at which the cambial tissues may still have been in a relatively dormant state. This may not therefore have been an optimal time-point to look for genes involved in wood formation. However, these were the only samples that had been taken when the trip to the Vancouver laboratory had been scheduled. There was therefore no choice of which sampling date should be used to provide material for use in Canada. In the later experiments which were performed in the Southampton laboratory samples were available from later in the growing season. The decision was therefore made to switch to samples collected in July 2004 for the following Sitka spruce experiments, and much more consistent patterns of expression were shown.

The design of experiment one may also have contributed to this variation in gene expression. The design was constructed on a clonal basis rather than an individual replicate basis (i.e. selecting replicates of clones which exhibited extremes of wood density and average diameters). As figure 3.1 shows, some of the clones exhibit a wide range of density measurements, while others show much greater consistency. Some of the clones also demonstrated similar high levels of variability in their lignin and cellulose contents. The choice of clones was based on clonal averages, and this within clone variation was not given sufficient consideration. Although the arrays that were eventually selected for further analysis in experiment one showed extreme differences in density measurements, the variation shown by the diameter and lignin measurements may have led to the inconsistent results. It may have been more appropriate to select individual replicates for this experiment, thereby removing the clonal element of this study and concentrating on pairing high versus low wood density individuals as opposed to clones.

Experiment one showed a gene involved in the lignin biosynthesis pathway (CCoA-OMT) to be down-regulated in high-density samples. This might be expected from the lignin measurements collected for these samples (figure 3.14), where the high-density clones had significantly lower lignin contents than the low-density clones. Kirst *et al*. (2004) have shown lignin genes to be negatively correlated with growth.

Experiments two and three showed greater consistency in their expression patterns than experiment one. Overall trends showed storage proteins, and genes with functions relating to defence and senescence to be down-regulated in high density samples. Schrader *et al*. (2004b) showed stress-response and senescence genes to be highly up-regulated in dormant tissue.

Genes with roles in cell division, expansion and growth were up-regulated in the high density samples from experiments two and three. These include genes such as expansins, and auxin transporters. As wood formation involves the processes of cell division and expansion, the indication that genes involved in these processes are associated with wood density is not surprising.

6.1.2 Poplar (Chapter 5)

The results from the poplar experiment showed similar trends to those of the second and third Sitka experiments. Once again, a gene (CCoA-OMT) involved in lignin biosynthesis was down-regulated in high-density samples, as was a stress response gene. Up-regulated genes included an auxin transporter and a B-like cyclin.

6.1.3 Comparison

Some interesting comparisons can be made between the Sitka spruce results and the poplar results. For example, amino acid permease, (*AUX1*), which is involved in auxin transportation, was shown to be up-regulated in high-density Sitka spruce samples (experiment three) and poplar samples. A role for auxin as a regulator of wood formation has previously been suggested (Moyle *et al*.,

2002). Also, genes with putative functions relating to cell formation and expansion were up-regulated in both high density Sitka spruce and poplar. These include genes such as expansin, glycoside hydrolase and one which codes for the synthesis of an arabinogalactan protein.

A lignin gene (CCoA-OMT was down-regulated in high-density Sitka spruce samples (experiment one) as well as high-density poplar samples. Kirst *et al*. (2004) has suggested that increased lignin content leads to a lack of carbon availability for tree growth, which could account for the relationship seen in our samples between lignin and wood density. Other genes that were observed to be down-regulated in Sitka spruce and poplar included those with functions in stress-response and defence.

This is the first study of wood formation that has compared gene expression in cambial tissues of a field grown gymnosperm and angiosperm. Even though gymnosperms and angiosperms differ in terms of their physiology, biochemistry and anatomy of their wood formation process, the evidence presented in this thesis suggests that some of the fundamental determinants of wood density may be common to members of both taxonomic divisions.

The genes listed in this section are ideal candidates for further analysis, as they show differential expression in high versus low wood density individuals in more than one species and many have previously been identified in other studies as having a role in wood formation.

6.2 Potential Applications

The aim of this study has been to determine a link between gene expression in the wood-forming tissues and wood density in Sitka spruce and poplar. This section considers how a knowledge of the key genes involved in the determination of wood density can be used towards developing approaches that will eventually lead to practical improvements in British forestry.

One possible application could be the genetic transformation route, where multiple copies of a candidate gene can be inserted into a genome to bring about an improvement in a specific desirable trait such as wood density. Poplar has been identified as a model species due to the ease by which it can be genetically transformed (Taylor, 2002), so it would be a better candidate than Sitka spruce for this type of approach. Genetic transformation does however have its limitations, a major one being strong opposition to field-based studies of modified trees. For this reason, a screening approach might be more acceptable.

The type of approach suggested above may be more suitable for research, in terms of assigning definitive functions to some of our genes of interest. It is possible to assign putative functions based on homology to other sequences, but altering the genotype to assess changes in the phenotype is a useful method for validating putative functions.

Plantation trees could be screened at a young age to assess their expression of a certain gene, with those which exhibit over-expression in relation to the rest of the plantation being selected as ‘plus-trees’. An ability to speed up the breeding and selection process in this way could eventually lead to an increase in the productivity of our plantation forests, and thereby reduce the pressure on natural forests in terms of sustainable quality wood production (Sedjo, 1999).

A knowledge of the genes which are the key determinants of wood density also offers the opportunity to search for sequence variation in these genes. Tree breeders are interested in the genetic variation that is present in their material and an investigation of the DNA sequence of these important genes could reveal desirable variants associated with improvements in wood density. For example, a SNP (single nucleotide polymorphisms) in one of these candidate genes may alter one of the amino acids that make up a protein whose function is to act as an enzyme to catalyse a key step in wood formation. If any of these SNPs could be linked to wood density then they could be used as markers for early selection of high wood density genotypes.

6.3 Further Work

The next step in this study should be an attempt to perform a more detailed search for putative functions for the large number of un-annotated genes, by using software such as MAPMAN (Thimm *et al*., 2004). With such a large proportion of each gene lists showing no matches to any existing ESTs, or having no function assigned to them, it seems likely that further annotation would identify more genes of interest expressed in both Sitka spruce and poplar. The expression of the genes could then be studied in a wider range of samples using PCR.

It would be of interest to study gene expression patterns across an entire growing season, either using microarrays to access large numbers of genes, or using a real-time PCR approach to develop a time sequence for the expression patterns of the most important differentially expressed genes identified in this study. Schrader *et al*. (2004b) assessed gene expression differences between active and dormant cambial tissues, but there have been no studies that follow the same individual trees throughout a growing season. Monthly sample collection between May and September would give 5 time points with which to compare expression patterns.

The current study is unique in that it studies gene expression in relatively old, field grown Sitka spruce. The majority of the studies published to date have used much younger trees in their experiments. For example, Paux *et al*. (2004) used 6-year-old eucalyptus trees, Kirst *et al*. (2004) used 20-month-old eucalyptus, and Yang *et al*. (2004) used 12-15 year-old loblolly pine. It remains unknown whether the genes identified by their studies would continue to show the same up-regulated expression in high-density samples from older trees which have made the transition to mature wood production. A long-term trial following a cohort of individuals from juvenile wood production through to mature wood production could answer this question. If genes show consistent expression across the life of

an individual, gene expression could be screened very early in life when trees were at a young age, which could markedly speed up the breeding process.

List of references

Allona, I., Quinn, M., Shoop, E, Swope, K., St. Cyr, S., Carlis, J., Riedl, J., Retzel, E., Campbell, M.M., Sederoff, R. and Whetten, R.W. (1998) Analysis of xylem formation in pine by cDNA sequencing. Proceedings of the National Academy of Sciences USA 95, 9693-9698.

Andersson-Gunnerås, S., Mellerowicz, E.J., Love, J., Segerman, B., Ohmiya, Y., Coutinho, P.M., Nilsson, P., Henrissat, B., Moritz, T. and Sundberg, B. (2006) Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary cell wall biosynthesis. *The Plant Journal* 45, 144-165.

Aspeborg, H., Schrader, J., Coutinho, P.M., Stam, M., Kallas, Å, Djerbi, S., Nilsson, P., Denman, S., Amini, B., Sterky, F., Master, E., Sandberg, G., Mellerowicz, E., Sundberg, B., Henrissat, B. and Teeri, T.T. (2005) Carbohydrate-active enzymes involved in the secondary cell wall biogenesis in hybrid aspen. *Plant Physiology* 137, 983-997.

Bao, W., O'Malley, D.M. and Sederoff, R.R. (1992) Wood contains a cell-wall structural protein. *Proceedings of the National Academy of Sciences* USA 89, 6604-6608.

Bachem, C.W., van der Hoeven, R.S., de Bruijn, S.M., Vreugdenhil, D., Zabeau, M. and Visser, R.G. (1996) Visualisation of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. *Plant Journal* 9, 745-753.

Boerjan, W. (2005) Biotechnology and the domestication of forest trees. *Current Opinion in Biotechnology* 16, 159-166.

Bourquin, V., Nishikubo, N., Abe, H., Brumer, H., Denman, S., Eklund, M., Christiernin, M., Teeri, T.T., Sundberg, B. and Mellerowicz, E.J. (2002) Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls of vascular tissues. *The Plant Cell* 14, 3073-3088.

Bowyer, J.L., Shmulsky, R. and Haygreen, J.G. (2003) Forest Products and Wood Science: An Introduction. Iowa State Press.

Brazier, J.D. (1967) Timber improvement I: A study of the variation in wood characteristics in young Sitka spruce. Forestry 40(2), 117-128.

Brazier, J.D. (1970) Timber improvement II: The effect of vigour on young-growth of Sitka spruce. *Forestry* 43, 135-150.

Brown, P.O. and Botstein, D. (1999) Exploring the new world of the genome with DNA microarrays. *Nature Genetics* 21, 33-37.

Busov, V.B., Johannes, E., Whetten, R.W., Sederoff, R.R., Spiker, S.L., Lanz-Garcia, C. and Goldfarb, B. (2004) An auxin-inducible gene from loblolly pine (*Pinus taeda* L.) is differentially expressed in mature and juvenile-phase shoots and encodes a putative transmembrane protein. *Planta* 218, 916-927.

Campbell, M.M., Brunner, A.M., Jones, H.M., and Strauss, S.H. (2003) Forestry's fertile crescent: the application of biotechnology to forest trees. *Plant Biotechnology Journal* 1, 141 – 154.

Chaffey, N., Cholewa, E., Regan, S. and Sundberg, B. (2002) Secondary xylem formation in *Arabidopsis*: a model for wood formation. *Physiologia Plantarum* 114, 594-600.

Chambel, M.R., Climent, J., Alía, R. and Valladares, F. (2005) Phenotypic plasticity: a useful framework for understanding adaptation in forest species. *Invest Agrar: Sist Recur For* 14(3), 334-344.

Chang, S., Puryear, J. and Cairney, J. (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* 11(2), 113-116.

Cheung, V.G., Morley, M., Aguilar, F., Massimi, A., Kucherlapati, R. and Childs, G. (1999) Making and reading microarrays. *Nature Genetics* 21, 15-19.

Christensen, J.H., Overney, S., Rohde, A., Diaz, W.A., Bauw, G., Simon, P., Van Montagu, M. and Boerjan, W. (2001) The syringaldazine-oxidizing peroxidase PXP 3-4 from poplar xylem: cDNA isolation, characterisation and expression. Plant Molecular Biology 47, 581-593.

Cosgrove, D.J. (1997) Relaxation in a high-stress environment: The molecular bases of extensible cell walls and cell enlargement. *The Plant Cell* 9, 1031-1041.

Costa E Silva, J., Wellendorf, H. and Pereira, H. (1998) Clonal variation in wood quality and growth in young Sitka spruce (*Picea sitchensis* (Bong.) Carr.): estimation of quantitative genetic parameters and index selection for improved pulpwood. *Silvae Genetica* 47(1), 20-33.

Darley, C.P., Forrester, A.M., and McQueen-Mason, S.J. (2001) The molecular basis of plant cell wall extension. *Plant Molecular Biology 47*, 179-195.

Datson, N.A., van der Perk-de Jong, J., van den Berg, M.P., de Kloet, E.R. and Vreugdenhil, E. (1999) MicroSAGE: a modified procedure for serial analysis of gene expression in limited amounts of tissue. *Nucleic Acids Research* 27(5), 1300-1307.

Déjardin, A., Leplé, J., Lesage-Descauses, M., Costa, G. and Pilate, G. (2004) Expressed sequence tags from poplar wood tissues – a comparative analysis from multiple libraries. *Plant Biology* 6, 55-64.

Do, J.H. and Choi, D.K. (2006) Normalisation of microarray data: single-labelled and dual-labelled arrays. *Molecules and Cells* 22(3), 254-261.

Doblin, M.S., Kurek, I., Jacob-Wilik, D. and Delmer, D.P. (2002) Cellulose biosynthesis in plants: from genes to rosettes. *Plant & Cell Physiology* 43(12), 1407-1420.

Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P. and Trent, J.M. (1999) Expression profiling using cDNA microarrays. *Nature Genetics* 21, 10-14.

Eriksson, M.E., Israelsson, M., Olsson, O. and Moritz, T. (2000) Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fibre length. *Nature Biotechnology* 18, 784-788.

Ericsson, T. and Fries, A. (2004) Genetic analysis of fibre size in a full-sib *Pinus sylvestris* L. progeny test. *Scandinavian Journal of Forest Research* 19, 7-13.

Fenning, T.M. and Gershenzon, J. (2002) Where will the wood come from? Plantation forests and the role of biotechnology. *Trends in Biotechnology* 20(7), 291 – 296.

Fischer, A., Saedler, H. and Theissen, G. (1995) Restriction fragment length polymorphism-coupled domain-directed differential display: a highly efficient technique for expression analysis of multigene families. *Proceedings of the National Academy of Sciences of the USA* 92, 5331- 5335.

Fizames, C, Muños, S., Cazettes, C., Nacry, P., Boucherez, J., Gaymard, F., Piquemal, D., Delorme, V., Commes, T., Doumas, P., Cooke, R., Marti, J., Sentenac, H. and Gojon, A. (2004) The Arabidopsis root transcriptome by serial analysis of gene expression. Gene identification using the genome sequence. *Plant Physiology* 134, 67-80.

Fletcher, A.M. (1992) Breeding improved Sitka spruce. In, *Super Sitka for the 90's*, ed. D.A. Rook, Forestry Commission Bulletin 103, 11-24. HMSO, London.

Forsberg, D. and Warensjö, M. (2001) Grain angle variation: a major determinant of twist in sawn *Picea abies* (L.) Karst. *Scandinavian Journal of Forest Research* 16, 269-277.

Gill, J.G.S. (1987) Juvenile-mature correlations and trends in genetic variances in Sitka spruce in Britain. *Silvae Genetica* 36(5-6), 189-194.

Gion, J., Lalanne, C., Le Provost, G., Ferry-Dumazet, H., Paiva, J., Chaumeil, P., Frigerio, J., Brach, J., Barré, A., de Daruvar, A., Claverol, S., Bonneu, M., Sommerer, N., Negroni, L. and Plomion, C. (2005) The proteome of maritime pine wood forming tissue. *Proteomics* 5, 3731-3751.

Graves, D.J. (1999) Powerful tools for genetic analysis come of age. *Trends in Biotechnology* 17, 127-134.

Gray-Mitsumune, M. Mellerowicz, E.J., Abe, H., Schrader, J., Winzéll, A., Sterky, F., Blomqvist, K., McQueen-Mason, S., Teeri, T.T. and Sundberg, B. (2004) Expansins abundant in secondary xylem belong to subgroup A of the α-expansin gene family. *Plant Physiology* 135, 1552-1564.

Gray-Mitsumme, M., Blomquist, K., McQueen-Mason, S., Teeri, T.T., Sundberg, B. and Mellerowicz, E.J. (2008) Ectopic expression of a wood-abundant expansin PttEXPA1 promotes cell expansion in primary and secondary tissues in aspen. *Plant Biotechnology Journal* 6, 62-72.

Gwaze, D.P., Bridgwater, F.E., Byram, T.D., Woolliams, J.A. and Williams, C.G. (2000) Predicting age-age genetic correlations in tree-breeding programs: a case study of *Pinus taeda* L. *Theoretical and Applied Genetics* 100, 199-206.

Habu, Y., Fukada-Tanaka, S., Hisatomi, Y. and Iida, S. (1997) Amplified restriction fragment length polymorphism-based mRNA fingerprinting using a single restriction enzyme that recognises a 4-bp sequence. *Biochemical and Biophysical Research Communications* 234, 516-521.

Hall, A. (2008) Better RED than dead: paying the people for environmental services in Amazonia. *Philosophical Transactions of the Royal Society B* 363, 1925-1932.

Hannrup, B., Ekberg, I. and Persson, A. (2000) Genetic correlations among wood, growth capacity and stem traits in *Pinus sylvestris*. *Scandinavian Journal of Forest Research* 15, 161-170.

Hannrup, B., Grabner, M., Karlsson, B., Müller, U., Rosner, S., Wilhelmsson, L. and Wimmer, R. (2002) Genetic parameters for spiral-grain angle in two 19-year-old clonal Norway spruce trials. *Annals of Forest Science* 59, 551-556.

Hannrup, B., Cahalan, C., Chantre, G., Grabner, M., Karlsson, B., Le Bayon, I., Jones, G.L., Müller, U., Pereira, H., Rodrigues, J.C., Rosner, S., Rozenberg, P., Wilhelmsson, L. and Wimmer, R. (2004) Genetic parameters of growth and wood quality traits in *Picea abies*. *Scandinavian Journal of Forest Research* 19, 14-29.

Hansen, J.K. (1999) Genetic variation of spiral grain in Sitka spruce growing in Demnark. Multiple-trait selection for improved timber quality. PhD thesis. The Royal Veterinary and Agricultural University.

Hansen, J.K. and Roulund, H. (1998) Genetic parameters for spiral grain in two 18-year-old progeny trials with Sitka spruce in Denmark. *Canadian Journal of Forest Research* 28, 920-931.

Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A., Erlandsson, R., Blomqvist, K., Bhalerao, R., Uhlén, M., Teeri, T.T., Lundeberg, J., Sundberg, B., Nilsson, P. and Sandberg, G. (2001) A transcriptional roadmap to wood formation. *Proceedings of the National Academy of Sciences USA* 98(25), 14732-14737.

Hoffmann, R., Seidl, T. and Dugas, M. (2002) Profound effect of normalisation on detection of differentially expressed genes in oligonucleotide microarray data analysis. *Genome Biology* 3(7), 33.

Hoheisel, J.D. (1997) Oligomer-chip technology. *Trends in Biotechnology* 15, 465-469.

Hu, W., Harding, S.A., Lung, J., Popko, J.L., Ralph, J., Stokke, D.D., Tsai, C. and Chiang, V.L. (1999) Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nature Biotechnology* 17, 808-812.

Igartúa, D.V., Monteoliva, S.E., Monterubbianesi, M.G. and Villegas, M.S. (2003) Basic density and fibre length at breast height of *Eucalyptus globulus spp. globulus* for parameter prediction of the whole tree. *IAWA Journal* 24(2), 173-184.

Irizarry, R.A., Bolstad, B.M., Clollin, F., Cope, L.M., Hobbs, B. and Speed, T.P. (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research* 31(4), e15.

Jung, S., Lee, J. and Lee, D. (2003) Use of SAGE technology to reveal changes in gene expression in *Arabidopsis* leaves undergoing cold stress. *Plant Molecular Biology* 52, 553-567.

Kane, M.D., Jatkoe, T.A., Stumpf, C.R., Lu, J., Thomas, J.D. and Madore, S.J. (2000) Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Research* 28(22), 4552-4557.

Kawamoto, S., Ohnishi, T., Kita, H., Chisaka, O. and Okubo, K. (1999) Expression profiling by iAFLP: a PCR-based method for genome-wide gene expression profiling. *Genome Research* 9, 1305-1312.

Kehoe, D.M., Villand, P. and Somerville, S. (1999) DNA microarrays for studies of higher plants and other photosynthetic organisms. *Trends in Plant Science* 4(1), 38-41.

Kende, H. and Zeevaart, J.A.D. (1997) The five "classical" plant hormones. *The Plant Cell* 9, 1197-1210.

Kirst, M., Johnson, A.F., Baucom, C., Ulrich, E., Hubbard, K., Staggs, R., Paule, C., Retzel, E., Whetten, R. and Sederoff, R. (2003) Apparent homology of expressed genes from wood-forming tissues of loblolly pine (*Pinus taeda* L.) with *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* USA 100, 7383-7388.

Kirst, M., Myburg, A.A., De León, 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.

Kreil, D.P. and Russell, R.R. (2005) There is no silver bullet – a guide to low-level data transforms and normalisation methods for microarray data. *Briefings in Bioinformatics* 6(1), 86-97.

Kuhn, E. (2001) From library screening to microarray technology: strategies to determine gene expression profiles and to identify differentially regulated genes in plants. *Annals of Botany* 87, 139-155.

Le Provost, G., Paiva, J., Pot, D., Brach, J and Plomion, C. (2003) Seasonal variation in transcript accumulation in wood-forming tissues of maritime pine (*Pinus pinaster* Ait.) with emphasis on a cell wall glycine-rich protein. *Planta* 217, 820-830.

Lee, J. and Lee, D. (2003) Use of serial analysis of gene expression technology to reveal changes in gene expression in Arabidopsis pollen undergoing cold stress. *Plant Physiology* 132, 517-529.

Lee, S.J. (1992) Likely increases in volume and revenue from planting genetically improved Sitka spruce. In, Super Sitka for the 90's, ed. D.A. Rook, Forestry Commission Bulletin 103, 61-74. HMSO, London.

Lee, S.J (1997) The genetics of growth and wood quality in Sitka spruce estimated using mixed model analysis techniques. PhD thesis. University of Edinburgh.

Lee, S.J. (1999a) Improving the timber quality of Sitka spruce through selection and breeding. Forestry 72(2), 123-133.

Lee, S.J. (1999b) Predicted genetic gains from Sitka spruce production populations. *Information Note FCIN26*, Forestry Commission.

Lee, S.J. (1999c) Comparison of sawlog quantity and quality between Sitka spruce seedlots originating from Washington and Queen Charlotte Islands. *Information Note FCIN25*, Forestry Commission.

Lee, S.J., Woolliams, J., Samuel, C.J.A. and Malcom, D.C. (2002) A study of population variation and inheritance in Sitka spruce. III. Age trends in genetic parameters and optimum selection ages for wood density, and genetic correlations with vigour traits. *Silvae Genetica* 51(4), 143-151.

Leibenguth, F. and Shoghi, F. (1998) Analysis of random amplified polymorphic DNA markers in three conifer species. *Silvae Genetica* 47, 120-126.

Lerouxel, O., Cavalier, D.M., Liepman, A.H. and Keegstra, K. (2006) Biosynthesis of plant cell wall polysaccharides – a complex process. *Current Opinion in Plant Biology* 9, 621-630.

Li, J., Yang, H., Peer, W.A., Richter, G., Blakeslee, J., Bandyopadhyay, A., Titapiwantakun, B., Undurraga, S., Khodakovskaya, M., Richards, E.L., Krizek, B., Murphy, A.S., Gilroy, S. and Gaxiola, R. (2005) *Arabidopsis* H+-PPase AVP1 regulates auxin-mediated organ development. *Science* 310, 121-125.

Li, L., Zhou, Y., Cheng, X., Sun, J., Marita, J.M., Ralph, J. and Chiang, V.L. (2003) Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. *Proceedings of the National Academy of Sciences USA* 100(8), 4939-4944.

Liang, P. and Pardee, A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257, 967-971.

Lichtenegger, H., Reiterer, A., Stanzl-Tschegg, S.E. and Fratzl, P. (1999) Variation of cellulose microfibril angles in aoftwoods and hardwoods – a possible strategy of mechanical optimization. *Journal of Structural Biology* 128, 257-269.

Lindström, H., Harris, P., Sorensson, C.T. and Evans, R. (2004) Stiffness and wood variation of 3-year-old *Pinus radiata* clones. *Wood Science and Technology* 38, 579-597.

Lipshutz, R.J., Fodor, S.P.A., Gingeras, T.R. and Lockhart, D.J. (1999) High density synthetic oligonucleotide arrays. *Nature Genetics* supplement 21, 20-24.

Loopstra, C.A., Puryear, J.D. and No, E.G. (2000) Purification and cloning of an arabinogalactan-protein from xylem of loblolly pine. *Planta* 210, 686-689.

Ma, C., Yu, Q., Berg, A., Drost, D., Novaes, E., Fu, G., Yap, J.S., Tan, A., Kirst, M., Cui, Y. and Wu, R. (2008) A statistical model for testing the pleiotropic control of phenotypic plasticity for a count trait. *Genetics* 179, 627-636.

Macdonald, E. and Hubert, J. (2002) A review of the effects of silviculture on timber quality of Sitka spruce. *Forestry* 75(2), 107-138.

MacKay, J.J., O'Malley, D.M., Presnell, T., Booker, F.L., Campbell, M.M., Whetten, R.W. and Sederoff, R.R. (1997) Inheritance, gene expression, and lignin characterisation in a mutant pine deficient in cinnamyl alcohol dehydrogenase. *Proceedings of the National Academy of Sciences* USA 94, 8255-8260.

Matsumura, H., Nirasawa, S. and Terauchi, R. (1999) Transcript profiling in rice (*Oryza sativa* L.) using serial analysis of gene expression (SAGE). *The Plant Journal* 20(6), 719-726.

McGall, G.H., Barone, A.D., Diggelmann, M., Fodor, S.P.A., Gentalen, E. and Ngo, N. (1997) The efficiency of light-directed synthesis of DNA arrays on glass substrates. *Journal of the American Chemical Society* 119(22), 5081-5090.

Meyermans, H., Morreel, K., Lapierre, C., Pollet, B., De Bruyn, A., Busson, R., Herdewijn, P., Devreese, B., Van Beeumen, J., Marita, J.M., Ralph, J., Chen, C., Burggraeve, B., Van Montagu, M., Messens, E. and Boerjan, W. (2000) Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of CCoAOMT, an enzyme involved in lignin biosynthesis. *The Journal of Biological Chemistry* 275(47), 36899-36909.

Mijnsbrugge, K.V., Meyermans, H., Montagu, M.V., Bauw, G. and Boerjan, W. (2000) Wood formation in poplar: identification, characterisation, and seasonal variation of xylem proteins. *Planta* 210, 589-598.

Mitchell, H.L. (1961) A concept of intrinsic wood quality, and non-destructive methods for determining quality in standing timber. Report No. 2233. Forest Products Laboratory, Madison, Wisconsin.

Moyle, R., Schrader, J., Stenberg, A., Olsson, O., Saxena, S., Sandberg, G. and Bhalerao, R.P. (2002) Environmental and auxin regulation of wood formation

involves members of the Aux/IAA gene family in hybrid aspen. *The Plant Journal* 31(6), 675-685.

Nieminen, K.M., Kauppinen, L. and Helariutta, Y. (2004) A weed for wood? Arabidopsis as a genetic model for xylem development. *Plant Physiology* 135, 653-659.

Nuopponen, M.H., Birch, G.M., Sykes, R.J., Lee, S.J. and Stewart, D. (2006) Estimation of wood density and chemical composition by means of diffuse reflectance mid-infrared Fourier transform (DRIFT-MIR) spectroscopy. *Journal of Agricultural and Food Chemistry* 54, 34-40.

Organisation for Economic Co-operation and Development (2002) Consensus document on the biology of Picea sitchensis (Bong.) Carr. (Sitka spruce).

Patzlaff, A., McInnis, S., Courtenay, A., Surman, C., Newman, L.J., Smith, C., Bevan, M.W., Mansfield, S., Whetten, R.W., Sederoff, R.R. and Campbell, M.M. (2003) Characterisation of a pine MYB that regulates lignification. *The Plant Journal* 36, 743-754.

Patzlaff, A., Newman, L.J., Dubos, C., Whetten, R.W., Smith, C., McInnis, S., Bevan, M.W., Sederoff, R.R. and Campbell, M.M. (2004) Characterisation of PtMYB1, an R2R3-MYB from pine xylem. *Plant Molecular Biology* 53(4), 597-608.

Paux, E., Tamasloukht, M., Ladouce, N., Sivadon, P. and Grima-Pettenati, J. (2004) Identification of genes preferentially expressed during wood formation in *Eucalyptus*. *Plant Molecular Biology* 55, 263-280.

Peng, L., Kawagoe, Y., Hogan, P. and Delmer, D. (2002) Sitosterol-β-glucoside as primer for cellulose synthesis in plants. *Science* 295, 147-150.

Peng, X., Wood, C.L., Blalock, E.M., Chen, K.C., Landfield, P.W. and Stromberg A.J. (2003) Statistical implications for pooling RNA samples for microarray experiments. *BMC Bioinformatics* 4(26).

Peters, D.G., Kassam, A.B., Yonas, H., O'Hare, E.H., Ferrell, R.E. and Brufsky, A.M. (1999) Comprehensive transcript analysis in small quantities of mRNA by SAGE-lite. *Nucleic Acids Research* 27(24), e39.

Pilate, G., Guiney, E., Holt, K., Petit-Conil, M., Lapierre, C., Leplé, J., Pollet, C., Mila, I., Webster, E.A., Marstorp, H.G., Hopkins, D.W., Jouanin, L., Boerjan, W., Schuch, W., Cornu, D. and Halpin, C. (2002) Field and pulping performances of transgenic trees with altered lignification. *Nature Biotechnology* 20, 607-612.

Piquemal, J., Lapierre, C., Myton, K., O'Connell, A., Schuch, W., Grima-Pettenati, J. and Boudet, A. (1998) Down-regulation of CCR induces significant changes of lignin profiles in transgenic tobacco plants. *The Plant Journal* 13(1), 71-83.

Pleasance, E.D., Marra, M.A. and Jones, S.J.M. (2003) Assessment of SAGE in transcript identification. *Genome Research* 13, 1203-1215.

Plomion, C., Leprovost, G. and Stokes, A. (2001) Wood Formation in Trees. *Plant Physiology* 127, 1513-1523.

Pot, D., Rodrigues, J., Rozenberg, P., Chantre, G., Tibbits, J., Cahalan, C., Pichavant, F. and Plomion, C. (2006) QTL and candidate genes for wood properties in maritime pine (*Pinus pinaster* Ait.). *Tree Genetics and Genomes* 2(1), 10-24.

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.

Rae, A.M., Street, N.R., Robinson, K.M., Harris, N. and Taylor, G. (2008) The genetic architecture of biomass in short rotation coppice poplar. BMC Biology, in submission.

Ralph, S. and Yueh, H. (2004) Treenomix spruce cDNA microarray guide version 2.0 – March 7th, 2004.

Richmond, T. and Somerville, S. (2000) Chasing the dream: plant EST microarrays. *Current Opinion in Plant Biology* 3, 108-116.

Rogers. L.A. and Campbell, M.M. (2004) The genetic control of lignin deposition during plant growth and development. *New Phytologist* 164, 17-30.

Rogers, Y., Jiang-Baucom, P., Huang, Z., Bogdanov, V., Anderson, S. and Boyce-Jacino, M.T. (1999) Immobilisation of oligonucleotides onto a glass support via dulsifide bonds: a method for preparation of DNA microarrays. *Analytical Biochemistry* 266(1), 23-30.

Rozenberg, P. and Cahalan, C. (1997) Spruce and wood quality: genetic aspects (a review). *Silvae Genetica* 46(5), 270-279.

Sarén, M-P., Serimaa, R., Andersson, S., Saranpää, P., Keckes, J. and Fratzl, P. (2004) Effect of growth rate on mean microfibril angle and cross-sectional shape of tracheids on Norway spruce. *Trees* 18, 354-362.

Savill, P.S. and Sanders, A.J. (1983) The influence of early respacing on the wood density of Sitka spruce. *Forestry* 56(2), 109-120.

Schrader, J., Nilsson, J., Mellorowicz, E., Berglund, A., Nilsson, P., Hertzberg, M. and Sandberg, G. (2004a) A high-resolution transcript profile across the wood-forming meristem of poplar identifies potential regulators of cambial stem cell identity. *Plant Cell* 16, 2278-2292.

Schrader, J., Moyle, R., Bhalerao, R., Hertzberg, M., Lundeberg, J., Nilsson, P. and Bhalerao, R.P. (2004b) Cambial meristem dormancy in trees involves extensive remodelling of the transcriptome. *The Plant Journal* 40, 173-187.

Schulgasser, K. and Witztum, A. (2007) The mechanism of spiral grain formation in trees. *Wood Science and Technology* 41(2), 133-156.

Siedlecka, A., Wiklund, S., Péronne, M., Micheli, F., Leśniewska, J., Sethson, I., Edlund, U., Richard, L., Sundberg, B. and Mellerowicz, E.J. (2008) Pectin methyl esterase inhibits intrusive and symplastic cell growth in developing wood cells of Populus. *Plant Physiology* 146, 554-565.

Stener, L-G. and Hedenberg, Ö. (2003) Genetic parameters of wood, fibre, stem quality and growth traits in a clone test with *Betula pendula*. *Scandinavian Journal of Forest Research* 18, 103-110.

Sterky, F., Regan, S., Karlsson, J., Hertzberg, M.m Rohde, A., Holmberg, A., Amini, B., Bhalerao, R., Larsson, M., Villarroel, R., Van Montagu, M., Sandberg, G., Olsson, O., Terri, T.T., Boerjan, W., Gustafsson, P., Uhlén, M., Sundberg, B. and Lundeberg, J. (1998) Gene discovery in the wood-forming tissues of poplar: analysis of 5692 expressed sequence tags. *Proceedings of the National Academy of Sciences USA* 95, 13330-13335.

Sterky, F., Bhalerao, R., Unneberg, P., Segerman, B., Nilsson, P., Brunner, A., Charbonnel-Campaa, L., Lindvall, J., Tandre, K., Strauss, S., Sundberg, B., Gustafsson, P., Uhlén, M., Bhalerao, R., 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 USA* 101(38), 13951-13956.

Stettler, R.F., Bradshaw, H.D., Jr,. Heilman, P.E. and Hinckley, T.M. (1996) Biology of *Populus* and its implications for management and conservation. NRC Research Press, Ottawa, Ontario, Canada.

Storey, J.D. and Tibshirani, R. (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences USA* 100(16), 9440-9445.

Swarap, K., Benková, E., Swarap, R., Casimiro, I., Péret, B., Yang, Y., Parry, G., Nielsen, E., De Smet, I., Vanneste, S., Levesque, M.P., Carrier, D., James, N.,

Calvo, V., Ljung, K., Kramer, E., Roberts, R., Graham, N., Marillonnet, S., Patel, K., Jones, J.D.G., Taylor, C.G., Schachtman, D.P., May, S., Sandberg, G., Benfey, P., Friml, J., Kerr, I., Beeckman, T., Laplaze, L. and Bennett, M.J. (2008) The auxin influx carrier LAX3 promotes lateral root emergence. Nature Cell Biology 10(8), 946-960.

Swarap, R., Kargul, J., Marchant, A., Zadik, D., Rahman, A., Mills, R., Yemm, A., May, S., Williams, L., Millner, P., Tsurumi, S., Moore, I., Napier, R., Kerr, I.D. and Bennett, M.J. (2004) Structure-function analysis of the presumptive *Arabidopsis* auxin permease AUX1. The Plant Cell 16, 3069-3083.

Taylor, G. (2002) *Populus*: Arabidopsis for forestry. Do we need a model tree? *Annals of Botany* 90, 681-689.

Taylor, A.J. (2006) Wood density determination in *Picea sitchensis* using computerised tomography: who do density measurements compare with measurements of pilodyn pin penetration? Honours Thesis, University of Wales, Bangor.

Terol, J., Bargues, M. and Pérez-Alonso, M. (2000) ZFWD: a novel subfamily of plant proteins containing a C3H zinc finger and seven WD40 repeats. *Gene* 260, 45-53.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L.A., Rhee, S.Y. and Stitt, M. (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological pathways. *The Plant Jounal* 37, 914-939.

Thompson, D.A. (1992) Growth of Sitka spruce and timber quality. In, *Super Sitka for the 90's*, ed. D.A. Rook, Forestry Commission Bulletin 103, 54-60. HMSO, London.

Treacy, M., Evertsen, J. and Dhubháin, Á.N. (2000) A comparison of mechanical and physical wood properties of a range of Sitka spruce provenances. COFORD.

Tuominen, H., Puech, L., Fink, S. and Sundberg, B. (1997) A radial concentration gradient of indole-3-acetic acid is related to secondary xylem development in hybrid aspen. *Plant Physiology* 115, 577-585.

Tzfira, T., Zuker, A. and Altman, A. (1998) Forest-tree biotechnology, genetic transformation and its application to future forests. *Trends in Biotechnology* 16, 439 – 446.

Uggla, C., Moritz, T., Sandberg, G. and Sundberg, B. (1996) Auxin as a positional signal in pattern formation in plants. *Proceedings of the National Academy of Sciences USA* 93, 9282-9286.

Vandepoele, K., Raes, J., De Veylder, L., Rouźe, P., Rombauts, S. and Inzé, D. (2002) Genome wide analysis of core cell cycle genes in Arabidopsis. *The Plant Cell* 14, 903-916.

Vainio, U., Andersson, S., Serimaa, R., Paakkari, T., Saranpää, P., Treacy, M. and Eversten, J. (2002) Variation of microfibril angle between four provenances of Sitka spruce (Picea sitchensis [Bong.] Carr.). *Plant Biology* 4, 27-33.

Velculescu, V.E., Zhang, L., Vogelstein, B. and Kinzler, K.W. (1995) Serial analysis of gene expression. *Science* 270, 484-487.

Verhaegen, D., Plomion, C., Gion, J.-M., Poitel, M., Costa, P. and Kremer, A. (1997) Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. *Theoretical and Applied Genetics* 95, 597-608.

Wagner, U., Edwards, R., Dixon, D.P. and Mauch, F. (2002) Probing the diversity of the Arabidopsis glutathione S-transferase gene family. *Plant Molecular Biology* 49, 515-532.

Wang, H.H., Drummond, J.G., Reath, S.M., Hunt, K. and Watson, P.A. (2001) An improved fibril angle measurement for wood fibres. *Wood Science and Technology* 34, 493-503.

Wang, T., Aitken, S.N. and Kavanagh, K.L. (2003) Selection for improved growth and wood quality in lodgepole pine: effects on phenology, hydraulic architecture and growth of seedlings. *Trees – Structure and Function* 17, 269-277.

Watson, A., Mazumder, M., Stewart, M. and Balasubramanian, S. (1998) Technology for microarray analysis of gene expression. *Current Opinion in Biotechnology* 9, 609-614.

Whetten, R. and Sederoff, R. (1991) Genetic engineering of wood. *Forest Ecology and Management* 43, 301 – 316.

Whetten, R. and Sederoff, R. (1995) Lignin Biosynthesis. *The Plant Cell* 7, 1001-1013.

Willats, W.G.T., McCartney, L., Mackie, W. and Knox, J.P. (2001) Pectin: cell biology and prospects for functional analysis. *Plant Molecular Biology* 47, 9-27.

Williamson, R.E., Burn, J.E. and Hocart, C.H. (2002) Towards the mechanism of cellulose synthesis. *Trends in Plant Science* 7(10), 461-467.

Woodward, A.W. and Bartel, B. (2005) Auxin: regulation, action and interaction. *Annals of Botany* 95, 707-735.

Wu, R. and Stettler, F. (1998) Quantitative genetics of growth and development in *Populus*. III. Phenotypic plasticity of crown structure and function. *Heredity* 81, 299-310.

Yang, L., Wang, R., Ren, F., Liu, J., Cheng, J. and Lu, Y. (2005) AtGLB1 enhances the tolerance of *Arabidopsis* to hydrogen peroxide stress. *Plant Cell Physiology* 46(8), 1309-1316.

Yang, S., van Zyl, L., No, E. and Loopstra, C.A. (2004) Microarray analysis of genes preferentially expressed in differentiating xylem of loblolly pine (*Pinus taeda*). *Plant Science* 166, 1185-1195.

Chen Yanhui, Yang Xiaoyuan, He Kun, Liu Meihua, Li Jigang, Gao Zhaofeng, Lin Zhiqiang, Zhang Yunfei, Wang Xiaoxiao, Qiu Xiaoming, Shen Yunping, Zhang Li, Deng Xiaohui, Luo Jingchu, Deng Xing-Wang, Chen Zhangliang, Gu Hongya and Qu Li-Jia (2006) The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Molecular Biology* 60, 107-124.

Yokoyama, R. and Nishitani, K. (2001) A comprehensive expression analysis of all members of a gene family encoding cell-wall enzymes allowed us to predict *cis*-regulatory regions involved in cell-wall construction in specific organs of Arabidopsis. *Plant Cell Physiology* 42(10), 1025-1033.

Zhang, Q., Zhang, Z., Lin, S., Zheng, H., Lin, Y., An, X., Li, Y. and Li, H. (2008) Characterisation of resistance gene analogs with a nucleotide binding site isolated from a triploid white poplar. Plant Biology 10, 310-322.

Zhang, Y., Brown, G., Whetten, R., Loopstra, C.A., Neale, D., Kieliszewski, M.J. and Sederoff, R.R. (2003) An Arabinogalactan protein associated with secondary cell wall formation in differentiating xylem of loblolly pine. *Plant Molecular Biology* 52(1), 91-102.

8 Appendicies

8.1 Raw data from Experiment 35

Clone	extract-ives	cellul-ose	lignin	density	Clone	extract-ives	Cellul-ose	lignin	density
81(7111)2-90-1	0.63	45.12	24.04	401	81(7111)2-530-1	0.46	41.60	21.94	290
81(7111)2-97-2	0.70	47.14	23.98	346	81(7111)2-568-1	1.29	39.82	22.08	436
81(7111)2-104	0.59	44.30	21.40	378	81(7111)2-631-1	0.55	42.04	20.26	347
81(7111)2-116-1	0.64	41.77	19.79	380	81(7111)2-654-1	0.49	42.64	20.90	413
81(7111)2-163-1	0.70	43.41	22.68	345	81(7111)2-729-1	0.51	40.32	22.76	315
81(7111)2-165-1	0.66	38.95	23.53	398	81(7111)2-505-1	0.70	44.39	20.55	357
81(7111)2-326-1	0.58	42.95	21.37	364	81(7111)2-584-1	0.39	40.98		360
81(7111)2-345-1	0.72	42.94	21.70	399	81(7111)2-620-1	0.39	43.68	21.30	341
81(7111)2-470-1	0.72	46.12	21.60	440	81(7111)2-743-1	0.43	42.98	20.45	365
81(7111)2-499-1	0.50	47.40	22.98	417					
C20121-71-1	0.57	40.14	24.46	293	C20121-511-1	1.27	40.48	20.54	330
C20121-145-1	0.76	41.28	23.29	372	C20121-586-1	0.78	43.57	22.72	294
C20121-179-1	0.89	40.07	23.77	354	C20121-628-1	0.84	39.20	21.36	348
C20121-318	0.73	47.60	23.84	278	C20121-696-1	0.71	46.00	24.85	267
C20121-455-1	0.74	49.58	23.01	334	C20121-707-1	0.75	44.35	23.73	281
C20122-63-2	0.72	42.19	24.23	315	C20122-517-1	0.63	41.55	22.53	285
C20122-138-2	0.54	43.99	22.99	300	C20122-554-1	0.71	45.13	23.17	310
C20122-190-1	0.50	46.10	23.89	378	C20122-603-1	0.58	45.64	24.75	299
C20122-330-1	0.71	48.96	23.99	270	C20122-660-2	0.57	44.62	23.82	298
C20122-477-1	0.61	43.91	23.52	269	C20122-727-1	0.56	42.68	22.31	346
C20123-87-1	0.65	39.34	23.38	333	C20123-540-1	1.20	42.46	24.27	301
C20123-130-1	1.01	47.15	21.95	296	C20123-566-1	0.74	41.13	22.69	325
C20123-194-1	2.87	43.23	23.49	341	C20123-636-1	0.68	44.59	25.02	312

C20123-320-1	1.00	48.05	22.25	302	C20123-653-1	0.75	44.64	25.09	365
C20123-473-1	0.70	42.37	25.76	371	C20123-746-1	0.82	43.11	23.01	299
C20124-92-1	0.90	41.57	22.99	360	C20124-519-1	0.60	40.67	23.83	375
C20124-148-1	0.61	43.31	21.86	393	C20124-567-1	0.52	44.24	22.13	343
C20124-158-1	0.83	40.53	23.41	377	C20124-641-1	0.65	45.89	23.94	390
C20124-344-1	0.74	44.67	21.56	321	C20124-681-1	0.64	46.61	22.75	328
C20124-463-1	0.64	42.40	22.59	492	C20124-702-1	1.01	43.37	24.32	388
C20125-67-1	0.69	40.17	23.41	316	C20125-512-1	0.75	40.34	24.51	305
C20125-110-1	0.66	42.38	21.28	333	C20125-578-1	0.58	44.06	22.43	317
C20125-195-1	0.63	41.69	21.89	363	C20125-612-1	0.55	41.44	22.55	345
C20125-335-1	0.58	44.63	21.92	304	C20125-657-1	0.43	43.97	22.96	300
C20125-474-1	0.71	43.03	22.28	325	C20125-701-1	0.65	44.94	24.12	290
C20126-77-1	0.74	39.81	21.98	278	C20126-509	0.54	39.00	24.08	351
C20126-102-1	0.54	41.08	23.32	287	C20126-585	0.79	41.68	24.03	284
C20126-173-1	0.57	45.48	22.70	310	C20126-608	0.61	40.81	21.14	290
C20126-311-1	0.45	44.16	22.79	279	C20126-665	0.59	43.66	23.80	366
C20126-452-1	0.72	41.03	24.20	334	C20126-703	0.82	40.87	23.62	307
C20127-84-1	0.78	41.33	23.45	234	C20127-503		48.21	22.06	434
C20127-122-2	0.67	41.77	21.89	304	C20127-594	0.95	37.68	24.60	324
C20127-157-2	0.94	45.21	23.16	373	C20127-639	0.99	42.98	21.40	390
C20127-321-2	0.86	43.55	23.98	236	C20127-690	0.78	44.05	23.36	286
C20127-479-2	4.64	41.41	24.75	397	C20127-747	0.81	43.04	21.45	331
C20128-51-2	0.94	43.98	25.53	399	C20128-516	0.63	40.37	24.08	367
C20128-149-2	0.71	45.47	24.73	365	C20128-579	0.69	44.94	22.84	367
C20128-181-2	0.61	45.28	23.27	374	C20128-613	0.57	43.31	22.90	402
C20128-347-2	0.58	46.41	24.60	322	C20128-691	0.61	45.14	23.45	341
C20128-464-2	0.56	44.39	25.00	383	C20128-723	0.93	42.93	25.68	366
C20129-	0.68	45.06	24.10	339	C20129-	0.68	49.54	22.52	348

52-2					506				
C20129-124-2	0.58	43.91	24.71	330	C20129-592	0.71	46.89	24.17	373
C20129-159-2	0.67	44.19	25.82		C20129-629	0.82	43.98	24.32	430
C20129-303-2	0.66	42.63	26.56	387	C20129-677	0.72	45.09	23.22	365
C20129-489-2	0.68	42.85	26.48	343	C20129-733	0.65	44.63	22.98	344
C20130-73-2	0.50	45.46	24.18	347	C20130-513	0.48	47.66	22.87	337
C20130-114-2	0.49	46.01	24.61	319	C20130-598	0.78	44.79	22.68	418
C20130-187-2	0.73	47.55	23.81	364	C20130-646	0.67	47.55	21.81	383
C20130-328-2	0.61	46.31	23.21	338	C20130-700	0.83	44.25	25.70	354
C20130-472-2	0.74	43.05	23.94	320	C20130-730	0.82	43.58	25.31	401
C20131-96-2	0.72	41.18	25.44	344	C20131-529	0.67	47.23	23.65	332
C20131-109-2	0.66	49.85	24.19	378	C20131-587	0.93	44.94	23.97	343
C20131-177-2	0.75	44.48	24.84	327	C20131-622	0.89	50.03	22.28	353
C20131-339-2	0.62	43.11	24.31	307	C20131-685	0.89	44.36	24.22	403
C20131-498-2	0.65	43.60	26.49	348	C20131-711	0.81	46.53	22.20	430
C20132-98-2	0.55	44.18	29.44	432	C20132-549	0.46	46.57	23.03	391
C20132-128-2	0.69	47.07	24.38	341	C20132-625	0.54	48.86	22.57	426
C20132-188-2	0.55	46.97	25.14	358	C20132-659	0.33	49.79	21.48	438
C20132-325-2	0.58	44.09	24.04	351	C20132-731	0.52	45.91	22.81	463
C20132-494-2	0.57	43.68	25.18	411					
C20133-64-2	0.79	46.30	25.21	423	C20133-504	0.43	46.15	23.72	346
C20133-127-2	0.73	45.24	24.26	381	C20133-551	0.56	46.74	23.30	399
C20133-172-2	0.78	42.26	25.10	394	C20133-630	0.41	48.18	23.31	390
C20133-346-2	0.58	42.39	24.93	373	C20133-687	0.64	44.63	23.81	376
C20133-478-2	0.82	40.26	24.38	413	C20133-715	0.62	42.83	23.46	350
C20134-70-2	4.78	41.90	25.37	426	C20134-546	0.61	44.98	24.40	375
C20134-132-2	0.71	43.04	25.10	345	C20134-552	0.56	42.95	23.92	337
C20134-193-2	0.95	42.05	24.30	348	C20134-614	0.79	42.80	23.52	372

C20134-314-2	1.17	43.66	25.10	300	C20134-661	0.60	44.29	21.23	367
C20134-487-2	0.94	41.00	25.10	324	C20134-705	0.81	41.58	21.51	387
C20135-94-2	0.85	40.43	26.16	350	C20135-550	0.82	43.78	22.41	337
C20135-139-2	0.51	44.10	24.34	324	C20135-559	0.84	41.74	22.61	387
C20135-168-2	0.63	42.59	24.69	449	C20135-649	0.72	44.51	22.01	409
C20135-324-2	0.53	43.33	23.95	296	C20135-683	0.63	42.27	21.82	429
C20135-500-2	0.46	46.87	23.91	372	C20135-725	0.83	45.06	21.12	441
C20136-74-2	0.58	39.49	26.55	337	C20136-547	0.90	47.18	23.42	389
C20136-143-2	0.42	41.15	24.70	368	C20136-573	0.42	44.85	22.84	371
C20136-164-2	0.47	41.74	24.49	315	C20136-609	0.48	46.72	23.98	385
C20136-313-2	0.39	43.14	27.63	324	C20136-695	0.60	42.50	22.13	375
C20136-453-2	0.94	41.83	24.47	411	C20136-738	0.53	41.77	21.02	399
C20137-89-2	0.44	42.33	22.26	414	C20137-536	0.49	43.28	22.32	406
C20137-118-2	0.75	43.40	20.85	378	C20137-580	0.60	43.34	21.60	439
C20137-166-2	0.58	46.65	23.46	372	C20137-615	0.68	44.56	22.42	403
C20137-304-2	0.53	46.82	22.00	461	C20137-699	0.56	42.19	22.10	343
C20137-482-2	0.49	45.42	22.45	371	C20137-742	0.65	45.79	20.65	411
C20138-75-2	0.54	42.39	21.51	377	C20138-515	0.53	44.61	21.53	417
C20138-131-2	0.60	46.63	21.30	383	C20138-583	0.61	47.08	21.24	435
C20138-183-2	0.75	47.12	21.52	423	C20138-611	0.49	45.26	21.64	400
C20138-337-2	0.58	43.84	21.18	409	C20138-656	0.45	44.33	22.29	407
C20138-495-2	0.45	46.20	21.11	481	C20138-712	0.52	43.42	23.24	
C20139-86-2	0.55	49.69	21.55	325	C20139-525	0.72	41.40	21.38	353
C20139-133-2	0.44	48.86	22.04	356	C20139-557	0.58	41.42	22.92	365
C20139-175-2	0.42	42.64	22.75	384	C20139-616	0.64	44.27	21.30	386
C20139-349-2	0.54	43.62	21.41	393	C20139-674	0.80	41.63	21.69	380
C20139-467-2	0.44	42.00	23.29	388	C20139-750	0.56	42.31	21.50	366
C20140-	0.53	45.32	21.20	389	C20140-	0.68	41.26	22.99	367

59-2					502				
C20140-103-2	0.36	42.70	22.47	401	C20140-581	0.55	45.12	22.80	365
C20140-185-2	1.63	43.93	21.15	373	C20140-647	0.67	42.55	23.52	386
C20140-334-2	0.64	45.08	22.21	367	C20140-673	0.59	43.05	22.93	359
C20140-469-2	0.67	41.65	22.42	324	C20140-716	0.67	41.84	21.83	357
C20141-85-2	0.67	44.46	21.86	358	C20141-520	0.66	47.19	21.20	380
C20141-137-2	0.71	39.91	22.16	343	C20141-572	0.43	46.43	20.60	366
C20141-199-2	0.56	46.62	23.24	345	C20141-650	0.61	40.88	21.47	416
C20141-319-2	0.57	42.94	22.13	307	C20141-686	0.60	45.16	20.92	332
C20141-486-2	0.81	44.86	23.18	337	C20141-732	0.56	43.87	21.45	362
C20142-60-2	0.60	47.21	22.71	319	C20142-508	0.55	47.41	20.76	384
C20142-105-2	0.33	42.72	22.01	364	C20142-575	0.53	47.45	21.50	365
C20142-176-2	0.64	41.91	21.22	357	C20142-607	0.44	45.03	21.47	432
C20142-306-2	0.81	48.61	23.33	416	C20142-693	0.74	48.81	21.09	338
C20142-480-2	0.81	47.99	22.07	376	C20142-706	0.40	47.83	21.46	350
C20143-81-2	1.00	42.48	23.72	351	C20143-538	0.55	47.67	20.58	356
C20143-112-2	0.91	41.80	23.36	333	C20143-560	0.93	45.23	21.28	423
C20143-169-2	0.71	44.70	22.54	467	C20143-627	0.67	45.27	20.79	359
C20143-317-2	0.57	42.97	23.55	357	C20143-662	2.89	45.02	20.76	389
C20143-461-2	0.65	43.07	22.02	335	C20143-740	0.78	44.00	23.85	316
C20144-80-2	0.61	44.72	21.41	370	C20144-532	0.73	45.55	21.97	334
C20144-147-2	3.10	46.81	20.50	358	C20144-588	0.39	47.25	20.32	321
C20144-196-2	0.44	46.79	20.89	351	C20144-632	0.55	45.22	21.32	380
C20144-350-2	0.66	50.76	20.64	319	C20144-675	0.75	45.44	21.64	307
C20144-476-2	0.60	44.87	20.29	481	C20144-726	0.52	43.62	23.53	356
C20145-99-2	0.51	42.11	26.13	336	C20145-542	0.70	46.49	23.49	416
C20145-141-2	0.60	44.65	23.80	303	C20145-571	0.82	43.88	26.26	329
C20145-151-2	0.52	42.38	23.01	313	C20145-617	0.52	43.45	25.54	321

C20145-336-2	0.53	44.57	23.32	297	C20145-669	0.38	44.09	23.85	363
C20145-466-2	0.51	43.07	23.55	371	C20145-714	0.58	49.56	24.41	359
C20146-82-2	0.73	45.19	25.76	309	C20146-539	0.74	45.07	25.49	330
C20146-121-2	0.64	51.25	23.38	315	C20146-558	0.53	44.27	24.78	312
C20146-200-2	0.57	48.12	23.88	330	C20146-637	0.60	40.94	24.51	285
C20146-316-2	0.48	45.12	23.68	327	C20146-689	0.59	40.62	24.17	276
C20146-492-2	0.32	44.74	23.60	334	C20146-737	0.82	43.59	24.82	364
C20147-083	0.60	41.69	23.86	344	C20147-535	0.50	44.18	25.80	364
C20147-120-2	0.60	42.98	24.85	364	C20147-556	1.09	41.85	24.46	378
C20147-162-2	1.18	41.88	25.17	380	C20147-606	0.71	43.25	24.33	404
C20147-348-2	0.72	42.86	23.78	236	C20147-651	0.47	42.01	24.67	428
C20147-497-2	0.67	43.65	24.20	265	C20147-749	2.93	44.32	24.42	416
C20148-58-2	0.54	43.56	23.37	347	C20148-528	0.60	47.15	24.29	329
C20148-111-2	0.47	45.54	22.25	394	C20148-600	0.51	47.14	24.47	385
C20148-161-2	0.65	37.48	24.29	326	C20148-618	0.55	46.79	24.95	344
C20148-308-2	0.53	47.12	20.68	389	C20148-658	0.64	46.43	24.80	393
C20148-493-2	0.40	44.47	21.49	332	C20148-708	0.58	45.01	25.70	386
C20149-55-2	0.65	42.23	25.50	328	C20149-514	0.52	46.32	24.76	302
C20149-136-2	0.51	43.75	23.26	302	C20149-599	0.69	43.61	26.14	283
C20149-160-2	0.64	40.88	22.86	270	C20149-634	0.64	43.65	26.32	324
C20149-327-2	0.75	42.08	23.40	279	C20149-663	1.18	41.89	26.03	354
C20149-475-2	2.09	43.02	22.91	256	C20149-735	0.44	46.28	24.15	296
C20150-93-2	0.77	41.83	23.62	313	C20150-523	0.71	46.32	23.32	305
C20150-144-2	0.58	40.70	25.46	298	C20150-590	0.60	43.80	23.30	302
C20150-156-2	2.32	40.57	22.74	289	C20150-605	0.40	44.45	23.92	326
C20150-310-2	0.71	38.47	25.44	314	C20150-684	0.71	48.27	23.33	298
C20150-485-2	0.51	45.43	24.89	281	C20150-720	0.41	47.11	23.25	293
C20151-	0.60	39.75	23.40	348	C20151-	0.52	40.94	25.99	419

57-2					531				
C20151-101-2	0.64	39.44	25.03	374	C20151-561	0.68	43.74	24.96	306
C20151-170-2	0.46	37.16	25.08	343	C20151-623	0.55	43.58	25.53	344
C20151-342-2	0.46	43.39	23.20	307	C20151-655	0.42	47.33	26.20	326
C20151-465-1	0.59	39.74	25.37	365	C20151-721	0.49	45.68	24.72	337
C20152-61-2	0.75	35.48	25.72	336	C20152-545	0.42	46.13	24.33	327
C20152-129-2	0.62	41.16	24.67	321	C20152-593	0.37	46.86	24.94	381
C20152-198-2	0.83	44.94	23.68	321	C20152-619	0.48	46.66	25.09	363
C20152-490-2	0.53	41.80	24.77	376	C20152-688	0.39	46.99	23.89	282
20153-69-1	0.47	41.75	22.51	328	C20152-718	0.86	49.76	23.24	328
20153-150-2	0.34	45.86	22.61	315	C20153-537	0.64	44.21	26.00	322
20153-192-2	0.41	40.91	24.64	362	C20153-569	0.48	47.73	23.77	338
20153-323-2	0.64	46.53	23.94	317	C20153-640	0.71	45.33	24.21	414
20153-462-2	0.65	43.44	24.33	298	C20153-679	0.60	45.58	25.12	353
20154-62-2	0.52	41.55	24.09	392	C20153-728	0.56	48.04	23.77	307
20154-134-2	0.36	43.27	22.39	387	C20154-507	0.35	47.17	23.33	342
20154-184-2	0.40	46.80	25.35	293	C20154-576	0.65	46.40	23.99	406
20154-338-2	0.53	43.68	22.96	316	C20154-601	0.47	46.31	23.49	369
20154-456-2	0.39	47.50	24.42	365	C20154-680	0.47	49.58	24.12	323
20155-79-2	0.49	45.91	24.99	396	C20154-745	0.61	47.15	23.34	434
20155-142-2	0.48	38.89	22.13	364	C20155-533	0.40	44.93	23.83	366
20155-189-2	0.52	45.83	21.52	366	C20155-570	0.47	44.30	26.06	365
20155-301-2	0.58	42.51	22.75	330	C20155-621	0.38	47.51	23.43	404
20155-468-2	0.48	41.53	22.66	372	C20155-668	0.41	50.85	21.36	318
20156-68-2	0.53	38.32	24.16	331	C20155-739	0.46	48.28	21.34	352
20156-146-2	0.55	40.53	21.80	344	C20156-527	0.86	44.81	22.62	359
20156-182-2	0.36	45.80	25.01	327	C20156-555	0.36	47.06	21.46	334
20156-322-2	0.57	43.51	23.53	351	C20156-624	0.60	48.50	22.21	382

20156-471-2	0.44	43.62	23.61	315	C20156-670	0.51	48.21	22.38	315
20157-56-2	0.48	40.49	23.54	375	C20156-741	0.85	44.80	23.52	355
20157-125	0.44	44.02	22.13	334	C20157-518	0.50	48.47	20.68	360
20157-180-2	0.52	49.82	22.14	351	C20157-595	0.45	48.77	21.39	375
20157-340-2	0.42	44.64	21.51	322	C20157-645	0.44	49.75		355
20157-484-2	0.56	43.76	22.58	355	C20157-652	0.30	46.96	21.04	365
20158-053	0.60	40.41	21.98	337	C20157-736	0.60	47.19	21.48	382
20158-119	0.48	43.32	23.44	310	C20158-510	0.32	48.69	22.39	271
20158-155	0.50	43.83	23.16	363	C20158-589	0.52	47.40	22.24	350
20158-312	0.92	43.91	22.14	357	C20158-644	0.55	50.90	22.67	381
20158-459	0.89	44.05	23.04	376	C20158-664	0.58	48.14	22.05	293
20159-76-2	0.65	42.11	22.86	362	C20158-709	0.51	44.91	22.05	340
20159-106-2	0.62	44.21	21.31	352	C20159-522	0.41	46.61	21.02	334
20159-178-2	1.73	40.82	21.53	356	C20159-565	0.46	46.48	21.33	353
20159-341-2	0.47	45.13	22.04	333	C20159-638	0.47	49.89	21.48	289
20159-488-2	0.55	42.71	21.31	316	C20159-698	0.77	46.55	22.01	310
20160-72-2	0.32	38.92	23.72	333	C20159-722	0.41	48.64	21.22	336
20160-123-2	0.29	43.42	21.65	319	C20160-521	0.47	47.14	21.66	343
20160-153-2	0.45	44.85	22.72	400	C20160-574	0.38	48.48	20.91	364
20160-343-2	0.44	41.52	22.81	363	C20160-602	0.32	48.99	21.22	398
20160-481-2	0.67	43.18	22.95	306	C20160-692	0.96	47.80	21.82	418
C20161-88-2	0.78	46.33	21.82	355	C20160-744	1.94	48.32	21.89	405
C20161-117-2	0.58	44.86	23.93	354	C20161-543	0.44	43.91	22.40	336
C20161-197-2	0.50	41.86	24.29	367	C20161-563	0.44	46.87	24.45	323
C20161-302-2	0.55	44.14	23.77	328	C20161-604	0.57	45.85	23.49	386
C20161-491-2	0.74	43.89	22.75	337	C20161-671	0.55	51.38	21.71	311
C20162-66-2	0.68	35.60	26.48	361	C20161-704	0.54	48.79	21.97	362
C20162-	0.66	44.31	23.50	327	C20162-	0.47		23.22	373

113-2					526				
C20162-167-2	0.66	41.17	25.74	374	C20162-582	0.77	50.07	24.52	316
C20162-309-2	0.63	39.34	27.00	347	C20162-626	0.52	47.6383	23.11	405
C20162-454-2	0.60	40.22	25.13	335	C20162-666	0.45	48.26	23.95	270
C20163-54-2	0.54	43.03	25.89	290	C20162-724	0.76	45.69	23.27	273
C20163-115-2	0.47	42.52	24.57	326	C20163-541	0.52	47.10	23.11	291
C20163-154	0.62	44.19	23.74	306	C20163-564	0.53	46.00	22.30	355
C20163-307-2	0.49	44.22	24.76	355	C20163-610	0.47	46.00	23.45	310
C20163-496-2	0.43	39.96	25.21	277	C20163-672	0.47	47.78	22.87	331
C20164-100-2	0.58	38.77	25.38	323	C20163-748	0.58	46.36	22.52	343
C20164-107-2	0.57	40.41	25.02	321	C20164-548	0.61	45.49	23.36	366
C20164-186-2	0.66	41.29	24.45	315	C20164-591	0.45	48.13	21.96	399
C20164-331-2	0.48	43.94	24.41	377	C20164-648	0.50	46.49	22.82	329
C20164-451-2	1.52	42.40	22.97	341	C20164-697	0.66	46.70	22.17	382
C20165-78-2	0.70	39.03	26.85	344	C20164-710	0.70	48.38	22.30	360
C20165-135-2	0.57	38.45	25.02	284	C20165-501	0.57	45.83	22.74	337
C20165-174-2	0.76	42.74	24.55	315	C20165-596	0.53	45.38	22.65	336
C20165-305-2	0.65	35.81	24.45	348	C20165-635	0.57	43.75	22.38	365
C20165-460-2	0.72	40.52	23.24	350	C20165-678	0.61	46.23	23.07	355
C20166-95-2	0.97	43.97	23.71	277	C20165-717	2.06	42.86	22.91	405
C20166-126-2	0.84	41.46	24.63	292	C20166-534	0.67	45.24	23.45	308
C20166-152-2	0.49	41.23	23.74	283	C20166-562	0.74	47.63	21.05	331
C20166-332-2	0.56	40.98	24.65	326	C20166-633	0.58	48.23	23.13	325
C20166-458-2	0.99	41.25	24.27	288	C20166-676	0.71	45.27	22.62	320
C20167-65-2	0.55	45.70	24.98	295	C20166-713	0.71	47.02	21.61	276
C20167-108-2	0.82	42.20	25.12	346	C20167-524	0.44	50.52	22.29	346
C20167-191-2	1.72	41.92	25.98	317	C20167-597	0.71	46.02	22.50	349
C20167-329-2	0.76	46.15	25.00	282	C20167-643	0.77	46.60	21.82	335

C20167-457	0.83	45.72	24.71	346	C20167-694	0.59	47.49	22.17	318
C20168-091-1	6.59	42.62	24.98	291	C20167-734	0.59	50.43	21.05	335
C20168-140	0.49	43.00	23.73	232	C20168-544	0.58	39.93	26.93	425
C20168-171-1	0.76	42.37	27.08	334	C20168-577	0.56	44.26	24.63	363
C20168-315-1	0.68	41.80	23.74	318	C20168-667	1.01	47.65	23.18	331
C20168-483-1	0.64	42.68	25.93	246	C20168-719	0.82	45.34	23.66	401

8.2 Raw data for Poplar experiment (family 331)

position	genotype	height (m)	diam (cm)	shoots	Av. Density	Av.den.outer ring
38	242	7.61	72.2	14	0.331379	0.43059
192	242	8	87.9	25	0.4012929	0.463136
58	246	6.49	50.1	7	0.3229656	0.389858
269	246	6.94	60.2	15		0.413165
220	968	8	94.3	15	0.3286304	0.425382
119	968	8	99.7	32	0.3976098	0.391391
268	1059	5	33.8	6	0.295646	0.350077
267	1059	4.9	36.9	11		0.438541
76	1060	3.45	21.6	4	0.237929	0.312764
10	1060	4.34	38.8	7	0.296979	0.370872
231	1061	6.34	50	9	0.3856654	0.512719
268	1061	4.44	28.6	10		0.37538
217	1062	5.92	64.7	15	0.362839	0.445058
244	1062	4.29	27.8	15	0.3318514	0.405319
190	1063	4.08	27.2	7	0.3311004	0.387773
213	1063	4.46	45.5	12	0.3471634	0.441055
179	1067	4.08	27.5	11	0.2640527	0.341375
123	1067	4.04	29.8	13	0.2908838	0.330907
168	1068	6.7	59.1	20	0.3451612	0.445938
125	1068	5.28	39.9	19	0.3740927	0.455929
296	1075	7.6	84.6	20	0.3724176	0.395586
76	1075	6.88	56.7	23		0.446507
125	1076	4.81	36.1	8	0.3550178	0.507465
283	1076	5.23	48.3	7		0.441831
136	1077	3.42	20.6	9	0.3032973	0.396544
188	1077	2.62	19.9	7	0.3061865	0.377072
265	1078	4.97	48.6	16	0.3402658	0.452589
161	1078	4.31	39.3	14	0.3763198	0.489236
115	1080	4.86	41.6	9	0.3838572	0.466862
106	1080	5.22	50.1	14	0.367	0.455156
278	1082	5.32	36.9	7	0.3434398	0.453351
299	1082	4.95	35.3	9		0.503523

134	1085	3.92	23.5	5	0.337701	0.483449
157	1085	4.7	35.4	3	0.3320112	0.442413
234	1085	3.7	21.1	3	0.3356045	0.425683
206	1086	4.95	32.9	8	0.3058943	0.379911
60	1086	3.43	19	1		0.293635
165	1086	2.75	17.3	9	0.2889177	0.403787
80	1087	5	35	12	0.302988	0.365234
122	1087	4.6	29.1	19	0.382071	0.526484
242	1090	4.28	36.4	18	0.38234	0.538439
113	1090	4.91	40.5	20	0.3389816	0.400848
178	1093	3.54	22.6	4	0.281855	0.349304
71	1093	2.67	19	3		0.347215
232	1095	4.54	46.3	17	0.3638142	0.456282
281	1095	5.18	44.3	15		0.472503
291	1102	4.69	29	7	0.3194815	0.457491
46	1102	4.88	30	7		0.405622
88	1103	4.51	28	6	0.3269927	0.380009
33	1103	5.79	44.8	13		0.466201
180	1104	2.85	15.3	2	0.271207	0.364509
2	1104	4.57	34.8	7	0.3465672	0.452776
83	1106	2.65	15.8	1	0.219561	0.330008
121	1106	2.39	12	3	0.313338	0.3827
253	1108	4.67	38.4	8	0.3207478	0.406425
89	1108	5.03	36.2	7		0.398953
176	1114	6.74	71	12	0.3076609	0.387119
17	1114	3.12	16.1	2	0.3358268	0.452115
241	**1118**	2.67	17.2	4	0.3047133	0.417252
280	1118	4.26	32.1	10		0.430863
97	1120	3.94	36	10	0.304017	0.36131
245	1120	3.88	31.3	25	0.3187403	0.395887
44	1122	6.75	55.7	25	0.4007409	0.427321
291	1122	5.23	37.3	6		0.436197
207	1125	3.95	27.6	4	0.3856234	0.41971
218	1125	4.44	30.7	8	0.4085042	0.500104
272	1126	3.17	16.4	5	0.3759403	0.489578
207	1126	2.23	9.6	4	0.2656725	0.377933
25	1127	3.84	20.6	7	0.3052183	0.367937
97	1127	4.82	33.1	6		0.434366
73	1128	3.71	31.9	10	0.3003158	0.402481
53	1128	5.61	37.4	13		0.425021
8	1131	3.56	24.6	8	0.2676163	0.360514
124	1131	5.5	45.9	21	0.3171304	0.382615
127	1135	3.9	22.9	1	0.294928	0.391924
36	1135	4.16	26.8	2		0.44067
63	1140	4.31	30.6	3	0.3197183	0.403612
278	1140	4.17	26.2	7		0.421484
186	1149	5.45	46	7	0.3489484	0.456412
141	1149	5.68	38.7	7	0.3548618	0.406885

123	1158	3.16	18.2	6	0.289425	0.391344
204	1158	3	15.3	10	0.35177	0.413592
90	1163	2.91	19.1	6	0.2500927	0.378326
212	1163	3.22	25.5	19	0.3503635	0.400759
124	1186	4.2	22.1	9	0.3153595	0.38503
108	1186	5.21	35	12	0.33773	0.448423
245	1582	5.58	33.9	11	0.3162663	0.408912
187	1582	4.53	30	7	0.3619723	0.463585
109	1584	4.39	27.6	9	0.3241388	0.402457
83	1584	4.63	31.4	10		0.377888
85	1586	2.95	14.6	9	0.2498313	0.356059
40	1586	3.42	23.2	9		0.372269
145	1591	6.13	41.6	12	0.4151289	0.436173
236	1591	5.44	33.2	1	0.296959	0.373542
189	1594	4.5	32.2	11	0.3160055	0.437933
216	1594	4.3	37	10	0.321493	0.424383
126	1599	4.88	29.2	6	0.3424944	0.352891
265	1599	4.45	30.2	7	0.3604818	0.450975
147	1609	5.6	46.3	15	0.3735064	0.436188
8	1609	4.31	33.7	8		0.456294
263	1610	11.16	103.5	16	0.3795589	0.433622
180	1610	7.52	62.4	7	0.3375463	0.387297
250	1616	5.42	46.9	7	0.337647	0.419097
276	1616	5.36	52.9	18		0.399326
129	1619	4.02	31.1	11	0.3604223	0.520787
293	1619	4.02	31.2	4		0.404387
221	1630	4.95	35.7	3	0.3602708	0.473007
112	1630	5.9	41.2	7	0.404588	0.49766
164	1632	4.26	32.3	17	0.31472	0.369643
128	1632	4.89	40.8	13	0.328043	0.377367
71	1634	4.46	34.4	12	0.3742894	0.387143
49	1634	4.02	28.1	8		0.381103
122	1637	3.28	17.4	6	0.3286835	0.420949
239	1637	3.57	18.1	8	0.3236678	0.444601
61	1640	2.7	15.9	4	0.2810535	0.30182
182	1640	4.7	33.8	13	0.3670778	0.491455
162	1641	4	28.4	3	0.325615	0.440898
64	1641	3.42	22.3	12		0.389677
75	1641	1.61	7.9	2		0.385716
159	1642	4.61	44.3	15	0.3514142	0.420347
135	1642	4.67	35.5	10	0.3327415	0.457568
52	1645	4.63	26.6	9	0.3071613	0.369616
159	1645	4.97	44.1	1	0.3455322	0.390529
64	1650	3.16	19.3	4	0.2480793	0.295106
15	1650	4.59	26.6	4	0.246168	0.317638
249	1652	4.62	30.1	5	0.3096558	0.384911
6	1652	4.84	31.4	4		0.442886
193	1667	1.94	9.1	2	0.293095	0.420585

107	1667	2.04	9.4	6	0.32961	0.429791
2	1674	6.59	52.4	24	0.3290977	0.379987
77	1674	7.66	70.9	44		0.385724
243	1676	4.28	33.9	4	0.3038445	0.401091
117	1676	4.3	30.7	7	0.2886818	0.381968
5	1678	2	8	12	0.320976	0.414651
200	1678	3.38	17.1	9	0.317522	0.432633
184	1679	3.85	25.9	8	0.3589178	0.503992
246	1679	3.46	23.5	14	0.333725	0.451208
167	1685	6.08	43.7	13	0.3408538	0.464766
74	1685	6.69	63	22		0.435671
65	1686	5.28	34.2	6	0.3364532	0.394353
98	1686	5.33	45.5	5		0.404951
66	1689	6.96	55.8	27	0.3764922	0.447654
225	1689	7.6	90.5	28	0.3401642	0.389617
202	1695	5.03	27.3	14	0.299098	0.347538
79	1695	4.94	29.4	23		0.371707
18	1700	3.34	18.6	14	0.2295685	0.34474
167	1700	3.45	23.6	11	0.273758	0.351532
287	1703	4.77	29.6	3	0.3428152	0.502346
58	1703	3.5	20.3	3		0.414339
81	1709	3.87	26.9	16	0.3265405	0.419828
14	1709	4.93	42.7	19	0.360979	0.426344
228	1719	4.06	25.8	5	0.320968	0.418821
19	1719	2.53	14.9	2	0.349572	0.440289
227	1720	3.06	17.1	3	0.294146	0.381672
44	1720	3.62	23.8	1		0.373068
251	1722	5.82	65.4	10	0.3214897	0.39038
229	1722	5.94	62	13	0.3638203	0.437296
223	1726	5.37	55.9	10	0.2927394	0.376532
66	1726	5.14	56.2	8		0.352728
226	1731	5.31	38.1	17	0.3943432	0.474002
191	1731	3.52	23.5	15	0.443769	0.500143
171	1732	4.57	32	7	0.2904068	0.328563
223	1732	3.1	17.3	14	0.3012783	0.38396
105	1734	5.16	42.4	10	0.3472232	0.440385
48	1734	5.21	40.6	6		0.375318
34	1738	2.16	8.8	5	0.3504335	0.444935
88	1738	3.36	17.6	6		0.399338
277	1768	7.6	84.9	37	0.3078076	0.372427
264	1768	8	71.6	36	0.317609	0.375371
154	1774	5.48	46.1	13	0.3374628	0.402203
206	1774	6.03	47.3	20	0.3504838	0.443127
39	1779	3.3	21.7	10	0.2393113	0.327535
131	1779	3.48	25.8	16	0.3423758	0.404952
153	1784	2.15	11.9	3	0.2517355	0.364285
103	1784	2.22	13.5	9	0.3780173	0.463628
188	1786	4.49	30.1	7	0.3591448	0.488293

249	1786	3.85	21.8	12	0.343676	0.415287
200	1791	5.73	40.8	10	0.363273	0.460892
242	1791	3.81	29.9	11	0.3263936	0.350399
24	1799	5.81	36.9	3	0.357568	0.453361
127	1799	5.29	35.3	8	0.3452784	0.439974
141	1801	6.96	58.8	13	0.3531565	0.44668
198	1801	6.74	47.7	8	0.3463805	0.447196
98	1809	3.81	27	9	0.2808893	0.347901
95	1809	4.14	31.2	11		0.359914
94	1810	1.48	7.4	9	0.254772	0.324087
70	1810	2.23	13	12		0.47289
72	1831	5.17	46.2	7	0.3653704	0.491748
186	1831	5.48	46.5	15	0.3630476	0.430879
114	1835	3.5	18.4	8	0.282919	0.398449
233	1835	1.94	10.6	11	0.317334	0.425487
131	1837	3.41	23.5	10	0.3371928	0.392833
261	1837	1.55	9.4	3	0.3070045	0.436772
293	1843	3.15	22	6	0.3093408	0.347324
271	1843	2.29	10.8	1		0.418658
271	1847	2.18	11.6	5	0.3275765	0.423997
279	1847	2.76	17.8	8		0.421523
216	1851	3.21	21.8	15	0.3299347	0.461082
31	1851	2.16	11.5	8		0.422926
267	1865	4.92	37.1	8	0.362824	0.466607
143	1865	4.21	30.4	7	0.3241685	0.406309
239	1866	4.49	35	13	0.2782515	0.349639
228	1866	4.08	29.5	14	0.2847558	0.333833
215	1867	4.15	18.7	4	0.362581	0.472275
178	1867	3.04	18.1	9	0.3132873	0.401348
201	1878	5.18	32	1	0.323903	0.377386
153	1878	3.95	28.8	8	0.323463	0.399012
53	1879	4.84	38.7	7	0.3182305	0.445779
107	1881	4.93	40	8	0.3140703	0.387966
129	1881	4.57	30.9	17	0.3184608	0.372241
133	1884	5.31	49.1	8	0.4116648	0.506511
171	1884	5.12	43.2	8	0.386249	0.533942
298	1891	2.52	15.3	6	0.2588007	0.290861
217	1891	3.29	25.1	9	0.3052223	0.369511
210	1903	2.34	12	8	0.3240787	0.383987
263	1903	2.52	13.8	12	0.3093155	0.283868
222	1915	4.47	26.4	4	0.3860018	0.440947
262	1915	3.44	19.2	5	0.3320273	0.443317
78	1918	8.9	84.6	20	0.3246001	0.445112
259	1918	8	97.9	37	0.386592	0.420796
84	1923	4.02	27.1	8	0.306617	0.323051
220	1923	4.17	34.2	22	0.398702	0.436073
273	1924	3.89	30.7	7	0.3869452	0.538669
201	1924	3.96	22.6	11	0.3518535	0.443265

40	1925	4.2	26.1	4	0.3483062	0.437588
224	1925	2.73	13.4	7	0.427078	0.476528
187	1926	3	16.5	8	0.3243567	0.387627
257	1926	2.47	15	9	0.3285248	0.373191
172	1930	2.86	16.8	9	0.3380783	0.409884
26	1930	3.71	24.6	11	0.3794078	0.496459
82	1931	4.4	31.4	5	0.2691828	0.339079
175	1931	4.07	24.7	4	0.3087438	0.369
37	1939	5.51	44.7	10	0.3683774	0.416443
205	1939	7.06	74.7	20	0.3595057	0.40297
290	1941	4.54	35.3	15	0.3228416	0.381762
16	1941	2.96	16.7	2	0.2556627	0.339742
146	1944	2.74	14.5	1	0.3024837	0.380027
69	1944	2.91	17.4	2		0.536438
140	1945	4.4	29.1	4	0.3086704	0.367501
235	1945	2.58	18.5	9	0.3288867	0.435715
211	1952	4.54	29.5	8	0.3258006	0.409927
24	1952	4.39	33.2	5	0.3059965	0.340415
35	1958	2.97	20.2	1	0.3081703	0.43293
140	1958	4.37	34.1	12	0.3501922	0.404191
266	1961	3.95	26.1	7	0.300817	0.439068
190	1961	2.15	11.4	7	0.33322	0.374466
75	1962	2.73	16.7	3	0.3327163	0.426886
184	1962	3.34	22.3	29	0.407467	0.442681
209	1964	5.76	44.4	11	0.3502088	0.457361
248	1964	5.53	44.6	20	0.362525	0.433858
120	2842	4.35	29.3	8	0.3081028	0.390583
176	2842	5.91	50.9	18	0.3442908	0.495344
256	2844	3.51	25.5	10	0.3127378	0.386907
104	2847	2.56	12.8	10	0.3716117	0.453366