Linking genetics and chemistry to minimise bark stripping in *Pinus radiata*

Judith Ssali Nantongo
BSc (Forestry), MSc (Agroforestry)

A thesis submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy
School of Natural Sciences, University of Tasmania (UTAS)
June 2021

Declarations

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

The thesis may be available for loan, copying and communication in accordance with the Copyright Act 1968

Signature   Date …26.06.2021…

Judith S. Nantongo, Candidate

Statement of co-authorship

The following people and institutions contributed to the work undertaken as part of this thesis:

Candidate: Nantongo Ssali Judith, School of Natural Sciences

Author 1: O'Reilly-Wapstra Julianne, University of Tasmania

Author 2: Potts M. Brad, University of Tasmania

Author 3: Fitzgerald Hugh, University of Tasmania

Author 4: Rodemann Thomas, University of Tasmania

Author 5: Newman Jess, University of Tasmania

Author 6: Davies W. Noel, University of Tasmania

Author 7: Dungey Heidi, SCION Research

Author 8: Klapste Jaroslav, SCION Research

Author 9: Graham Natalie, SCION Research

Author 10: Frickey Tancred, SCION Research

Author 11: Telfer Emily, SCION Research

Author 12: Aurik Don, Timberlands Pacific Pty Ltd

Author 13: Elms Stephen, Hancock Victoria Plantations

Contribution of work by co-authors for each paper:

Paper 1: Located in Chapter 2:

Nantongo JS, Potts BM, Fitzgerald H, Newman J, Elms S, Aurik D, Dungey H, O'Reilly-Wapstra J (2020) Quantitative genetic variation in bark stripping of *Pinus radiata*. *Forests* 11, 1356.

- Conception, design and resource allocation: Author 1, Author 2, Author 7, Author 12 and Author 13
- Data collection from the field: Candidate, Author 3 and Author 5
- Data curation and analysis: Candidate, Author 1 and Author 2
- Writing, review & editing: Candidate, Author 1, Author 2, Author 3, Author 5, Author 7, Author 12 and Author 13
- Provided pedigree data files that were pertinent to the chapter: Author 7
- Management of field trials: Author 12 and Author 13

Paper 2: Located in Chapter 3

Nantongo JS, Potts BM, Davies NW, Fitzgerald H, Rodemann T, O'Reilly-Wapstra J (2021) Variation in constitutive and induced chemistry in the needles, bark and roots of *Pinus radiata.* Submitted to *Trees*

- Experimentation: Candidate
- Sample collection and management; Candidate and Author 3
- Data collection of all laboratory work: Candidate, Author 3 and Author 6
- Data curation and analysis: Candidate, Author 1, Author 2 and Author 6
- Writing, review & editing: Candidate, Author 1, Author 2, Author 3, Author 4, Author 6, Author 12 and Author 13

Paper 3: Located in Chapter 4

Nantongo JS, Potts BM, Davies NW, Elms S, Aurik D, Fitzgerald H, O'Reilly-Wapstra J (2021) Chemical traits that predict susceptibility of *Pinus radiata* to marsupial bark stripping. Submitted to *Journal of Chemical Ecology*

- Experimentation: Candidate, Author 3
- Sample collection and management; Candidate and Author 3
- Data collection of all field work: Candidate and Author 3
- Data curation and analysis: Candidate, Author 1, Author 2 and Author 6
- Writing, review & editing: Candidate, Author 1, Author 2, Author 3, Author 12 and Author 13
- Management of field trials: Author 12 and Author 13

Paper 4: Located in Chapter 5

Nantongo JS, Potts BM, Davies NW, Fitzgerald H, Rodemann T, O'Reilly-Wapstra J (2021) Developing near infrared spectroscopy (NIRS) models for predicting chemistry and responses to stress in *Pinus radiata. Journal of Near Infrared Spectroscopy*, 09670335211006526.

- Experimentation: Candidate
- Sample collection and management; Candidate and Author 3
- Data collection of all laboratory work: Candidate, Author 3, Author 4 and Author 6
- Data curation and analysis: Candidate, Author 1, Author 2, Author 4 and Author 6
- Writing, review & editing: Candidate, Author 1, Author 2, Author 3, Author 4 and Author 6

Paper 5: Located in Chapter 6

Nantongo JS, Potts BM, Davies NW, Rodemann T, Fitzgerald H, O'Reilly-Wapstra J (2021)

Additive genetic variation in *Pinus radiata* near infrared estimated bark chemistry and the chemical traits associated with variation in bark stripping. Submitted to *Heredity*

- Experimentation: Candidate

- Sample collection and management; Candidate and Author 3

- Data collection of all laboratory work: Candidate, Author 3, Author 4, Author 6

- Data curation and analysis: Candidate, Author 1, Author 2, Author 4, Author 6

- Writing, review & editing: Candidate, Author 1, Author 2, Author 3, Author 4, Author 6

Paper 6: Located in Chapter 7

Nantongo JS, Potts BM, Klapste J, Graham N, Dungey H, Fitzgerald H, O'Reilly-Wapstra J (in prep.) Genomic selection for resistance to bark stripping and associated chemical compounds in radiata pine.

- Sample collection and management; Candidate and Author 3

- Data collection of all laboratory work: Candidate, Author 3 and Author 9

- Data curation and analysis: Candidate, Author 1, Author 2, Author 7, Author 8 and Author 9

- Writing, review & editing: Candidate, Author 1, Author 2, Author 3, Author 7, Author 8 and Author 9

- Provided pedigree data files that were pertinent to the chapter: Author 7

Paper 7: Located in Chapter 8

Nantongo JS, Potts BM, Frickey T, Telfer E, Dungey H, Fitzgerald H, O'Reilly-Wapstra J (2021). Constitutive and induced transcriptome analysis of the needles and bark of *Pinus radiata*. Submitted to *BMC genomics*

- Experimentation: Candidate

- Sample collection and management; Candidate and Author 3

- Data collection of all laboratory work: Candidate, Author 3, Author 10 and Author 11

- Data curation and analysis: Candidate, Author 1, Author 2, Author 7, Author 10 and Author 11

- Writing, review & editing: Candidate, Author 1, Author 2, Author 3, Author 7, Author 10 and Author 11

We, the undersigned, endorse the above stated contribution of work undertaken for each of the papers in preparation contributing to this thesis

Judith Ssali Nantongo
Candidate
School of Natural Sciences
Plant Science
University of Tasmania

Julianne O'Reilly-Wapstra
Primary supervisor
Associate Professor
School of Natural Sciences (Plant Science)
University of Tasmania

Greg Jordan
Head, Discipline of Biological Sciences
University of Tasmania

Signed ....

Date………26.06.2021………………..

Signed ……………   …

Date………26.06.2021………………..

Signed

Date …26.06.2021….

Brad M. Potts
Supervisor
Professor, Forest Genetics
School of Natural Sciences (Plant Science)
ARC Training Centre for Forest Value
University of Tasmania

Thomas Rodemann
Supervisor
Deputy Director - Central Science Laboratory
Senior Research Fellow, Vibrational Spectroscopy
University of Tasmania

Signed ……   ………….

Date………26.06.2021……………..

Signed ……   ….

Date…………26.06.2021………..

Statement of ethical conduct

The experimental work for the field trials in this thesis was performed with the approval of the University of Tasmania Animal Ethics Committee (Permit No. A0015577).

Judith Ssali Nantongo
Candidate
School of Natural Sciences
Plant Science
University of Tasmania

Signed ……   .

Date………26.06.2021………………..

Overall abstract

Pinus radiata (radiata pine) is native to California but the main plantation softwood in both Australia and New Zealand where it has been subject to breeding to improve productivity and wood quality. However, in many plantations in Australia trees are subject to bark stripping by mammalian (mainly marsupial) herbivores, which can markedly reduce the genetic gain achieved in breeding programmes. This thesis explores the mechanisms and potential for exploiting natural genetic variation in resistance/susceptibility to minimise bark stripping in *P. radiata*. The study uses field and nursery experiments, integrating results from quantitative genetics, genomics, gene expression, analytical chemistry and rapid phenotyping to understand the genetic basis of variation in bark stripping and associated physical and chemical traits.

To understand the genetic control and the stability of the genetic signal, bark stripping was scored in three *Pinus radiata* family trials (Chapter 2). Quantitative genetic analysis, using pedigree-based mixed linear models (ABLUP) revealed significant additive genetic variation in bark stripping. Non-additive genetic effects were insignificant. While narrow-sense heritability estimates were low, the significant genetic signal was relatively stable across sites. The highest damaged families had approximately two-fold more bark removed than the least browsed families. Selecting the top 20% least susceptible families for planting could potentially reduce bark stripping by up to 22%. In the two older trials, reduced bark stripping was genetically associated with the presence of thick and rough bark while the presence of obstructive branches and needles on the lower stem (stem access) reduced bark stripping in the younger field trial. These important physical traits were also under significant genetic control. A positive additive correlation between prior height and bark stripping in the younger trial suggests that selecting faster growing trees may make *P. radiata* more vulnerable in the early stages of tree growth. However, when accounting for these physical and growth traits as covariates, significant additive genetic variation in bark stripping was still evident suggesting that genetic-based chemical properties of the bark were also important.

To provide a framework to assess if chemical traits mediated variation in bark-stripping, the plant-wide constitutive and induced chemistry was first assessed. In a shade house experiment, induction over a 4-week period was achieved by treating trees with methyl jasmonate (a chemical stressor) and artificial bark stripping (Chapter 3), following which 81 chemical compounds were quantified in the needles, stem and roots. These plant parts had different constitutive chemical profiles, with quantitatively and qualitatively more secondary plant compounds in the bark. After treatment, an

overall upregulation of terpenes and phenolics and a down-regulation of sugars and fatty acids was observed. However, the quantitative and qualitative chemical responses differed between plant parts, treatments and time period over the 4 weeks of the experiment. Stronger responses were observed for primary compared to secondary metabolites suggesting their potential roles in plant stress responses, including bark stripping.

To identify the specific constitutive and induced chemical traits that differentiated families with extreme levels of bark stripping, 21 of the most damaged and 21 of the least damaged families were selected from a fenced area within the younger trial used in Chapter 2. This field experiment examined the constitutive and induced chemistry with 83 compounds quantified (Chapter 4). Of the constitutive chemical traits in the bark, specific sugars, phenolics and terpenes were significantly different between the resistant and susceptible families. The bark sugars - fructose and glucose - and the phenolics - phenyl ethanol and benzene acetic acid - increased in the more susceptible families. The bark sesquiterpenoids - bicyclogermacrene and an unknown sesquiterpenoid alcohol - increased in the less susceptible families. The resistant and susceptible families could not be separated based on induced bark chemistry nor constitutive and induced needle chemistry.

An important aim of this thesis was to generate genetic parameters for the chemical traits and identify those that are genetically correlated with bark stripping. To gain sufficient sample size for the genetic study, Chapter 5 explored the potential of near infra-red spectroscopy (NIRS) in qualitative classification and in quantifying the amounts of compounds identified in the samples from the methyl jasmonate treated, artificially bark stripped and non-treated trees used in Chapter 3. NIRS was successful in qualitatively separating samples from different plant parts as well as separating the treated from non-treated samples. NIRS models with high accuracy were developed for individual sugars, terpenes, phenolics and fatty acids. Highest accuracy models were developed for the sugars - glucose and fructose, suggesting practical application of such models, while models for most secondary compounds were able to give proximate amounts. The NIRS modelling was extended to quantify the chemistry for all samples in Chapter 6 for quantitative genetic analysis.

Quantitative genetic analysis of the NIRS predicted values of 65 compounds in the bark using pedigree-based mixed models (ABLUP) showed significant additive genetic variation for individual chemical traits with low to moderate narrow-sense heritability estimates (Chapter 6). Results further showed strong positive genetic correlation of the sugar – glucose with bark stripping and a strong negative correlation with the unknown sesquiterpenoid alcohol. The results strengthen the findings based on the wet chemistry of the extreme families in Chapter 4. More positive genetic correlations

were detected between bark stripping and fatty acids and an unknown diterpenoid, possibly due to the increased sample size gained from using NIRS prediction. No additive genetic variation in inducibility was detected and non-additive genetic variation in the constitutive chemistry was also not significant. However, most of the heritability estimates were low, implying that response to selection will be slow for these traits. Therefore, the potential improvement in the heritability estimates was tested using genomic models.

To develop the genomic models, SNP genotyping was performed on needles from trees collected in Chapter 6; giving a total of 15,624 SNPs (Chapter 7). Using linear mixed models, the narrow-sense heritability estimates based on genomic models, genomic best linear unbiased prediction (GBLUP) and single-step GBLUP (ssGBLUP) were substantially better for both resistance and chemical traits compared to estimates obtained from the pedigree-based models (ABLUP). For the chemical traits, the average of the univariate GBLUP heritability estimates was1.6- fold higher than the average of the univariate ABLUP heritabilities, suggesting that the SNPs were able to capture additional genetic information. Similarly, the heritability of all the compounds based on the trivariate ssGBLUP was 1.7 – fold higher than the trivariate ABLUP estimates. The predictive ability (PA) of the ssGBLUP was comparable to the trivariate ABLUP model. Similarly, the PA of univariate GBLUP was mostly comparable to the univariate generalised ridge regression (GRR) with a few exceptions. The better performance of the GBLUP over the GRR for most traits suggests that the traits are quantitative in nature, influenced by many genes.

The final chapter examined the expression of genes following methyl jasmonate and bark stripping with the aim of linking the chemical phenotypes in Chapter 3 to the underlying molecular activity, both in the needles and bark samples (Chapter 8). RNA was extracted and sequenced to yield 100bp paired-end sequences and each sample was sequenced to a depth of 20m reads per sample. After aligning the project transcriptome with the available *P. radiata* transcriptome, gene expression analysis showed up- and downregulation of genes associated with primary and secondary metabolism, with differences in transcript expression in the needles and the bark, between treatments and at different times. Consistent with the chemistry results, the genes that were related to secondary metabolism were also mainly up-regulated. Genes related to primary metabolism were more responsive than those related to secondary metabolism by up-regulation or down-regulation. Methyl jasmonate and bark-stripping showed many non-overlapping responses. Whereas maximum expression of the transcripts was observed 7 days after treatment, on the same population stronger

chemical changes were detected 14 and 21 days after treatment, suggesting a time-lag between gene and phenotypic expression.

Overall, the results indicated the potential for selection of less susceptible germplasm for operational plantings as a strategy to reduce the effects of bark stripping in plantation forestry. Selection against bark stripping in *P. radiata* can also be performed indirectly based on physical or chemical traits. Further tests may be required to establish the stability of the less susceptible families when planted as a monoculture.

Acknowledgements

Table of Contents

Chapter 1: General introduction

1.0   Herbivory in conifers

Conifers of the pine family (Pinaceae) comprise a diverse group of economically and ecologically important species found around the world in a diverse range of habitats from the sub-arctic to the tropics (Whitehill *et al.* 2016). In both natural and planted conifer forests, interactions between trees and herbivores are important determinants of forest productivity (Endress *et al.* 2016). However, in most planted forests, needles, bark and root herbivores can be detrimental to plant fitness, reducing growth, survival and reproductive output (Leidinger *et al.* 2019; O'Reilly-Wapstra *et al.* 2002). These detrimental effects translate into significant direct and indirect social-economic losses. In British Columbia for example, infestation by the mountain pine beetle has been projected to cause a loss in GDP of $57 billion by 2050 from reduction of timber sales (Corbett *et al.* 2015). Similarly large economic impacts of mammalian herbivory in conifers have been reported in Australia, Asia, Europe, North America and other parts of the world (Endress *et al.* 2016; Mayle *et al.* 2009; Miller *et al.* 2014; White 2019). Management of these herbivore pests to mitigate the effects of damage is costly. For example, the cost of protecting *Thuja plicata* seedlings against deer browsing during the initial stages of growth in British Columbia has been estimated to reach CAD$ 25 million per annum (Russell 2008). Accordingly, there has been interest in the exploitation of natural plant defences and tolerance mechanisms as a management solution (Telford *et al.* 2014).

1.1   Conifer defence traits

The arms race between herbivores and conifers over millions of years has led to the evolution of conifer defences that underpin natural resistance and tolerance to herbivory (Ehrlich and Raven 1964; Franceschi *et al.* 2005). Natural defences potentially offer environmentally safe, durable and less costly opportunities for forest management against herbivores (Johnson 2011; Russell 2008; Sniezko and Koch 2017). In some conifers, phenotypic and quantitative genetic variation in resistance and tolerance against herbivores has been detected (Moreira *et al.* 2013b; Zas *et al.* 2017; Zas *et al.* 2011). This variation is a prerequisite for breeding and genetic improvement programs (Alfaro *et al.* 2004; Russell 2008). To maximise the exploitation of such variation in forest management, it is critical to understand the mechanisms controlling quantitative variation in resistance and the plant traits underlying this variation (Kliebenstein 2014).

In conifers, a wide spectrum of both physical and chemical traits that act directly or indirectly to reduce herbivore damage or its deleterious effects have been characterised (Franceschi *et al.* 2005; Iason *et*

al. 2011; Moreira *et al.* 2013a; O'Reilly-Wapstra *et al.* 2007). Physical traits that play a role in defence include thick bark, constitutive and traumatic resin ducts and specialised phloem parenchyma cells (Franceschi *et al.* 2005). Chemical traits include secondary metabolites, mainly terpenoids and phenolics, although recently studies suggest that primary metabolites can directly or indirectly contribute to resistance or susceptibility (Gershenzon 1994; Page *et al.* 2013; Tauzin and Giardina 2014; Tiffin 2000). Defences can be constitutive (preformed) or induced (change in levels in response to herbivory). Constitutive defences are thought to be beneficial in environments where herbivory is consistent, while induced defences are thought to have evolved in environments where herbivory varies temporally. Studies, however, show tremendous variation in the expression of these traits between species (Raffa *et al.* 2017), genotypes (Moreira *et al.* 2013b), ontogenic stages (Erbilgin and Colgan 2012), time of expression (Hood and Sala 2015; Schmidt *et al.* 2005) and plant tissues (Moreira *et al.* 2012a). The response to specific defences by different herbivores is also variable, reflecting differences in feeding adaptations (Iason *et al.* 2011; Vourc'h *et al.* 2002b). The feeding pattern of a single herbivore species is often correlated with multiple plant traits, which may include physical traits as well as primary and secondary compounds (Agrawal and Weber 2015; Russell 2008). Identifying the specific chemistry that is correlated to herbivory in such diverse circumstances can be a complex task, necessitating phenotypic screening of large numbers of plants in detail (sometimes in an untargeted manner or without a prior hypothesis on traits), which requires fast and cost-effective analytical tools. In addition, large population sizes that provide sufficient power and resolution are required to connect genotype to phenotype (Singh *et al.* 2019).

1.2   Phenotyping for large-scale assessment of defence traits

Powerful analytical tools for the separation, characterization and quantification of the vast diversity of compounds in plants are available. The most common analytical technique is mass spectrometry coupled to chromatographic techniques (e.g. liquid chromatography–mass spectrometry, capillary electrophoresis–mass spectrometry and gas chromatography–mass spectrometry) (Jorge *et al.* 2016). Although these techniques are efficient, sample preparation or analysis procedures are complicated and time-consuming. Recently, studies have explored the potential application of spectroscopic methods to enable fast, low-cost and large-scale chemotyping for genetic studies. Near infrared spectroscopy (NIRS) for example has been used for measuring plant constituents and assessing the effects of a wide range of biotic and abiotic stressors in plants (Coops and Stone 2005; Radeloff *et al.* 1999). NIRS could possibly be applied to directly model herbivory related stresses in conifer plantations. In tree species including conifers, high accuracy of prediction for herbivore damage has been attained using near infrared spectroscopy (Henery *et al.* 2008; Radeloff *et al.* 1999). NIRS

has also been used to assess phytochemical variability, monitoring changing patterns of constituents and discriminating tree genotypes (O'Reilly-Wapstra *et al.* 2013a; Quentin *et al.* 2017; Stackpole *et al.* 2011). Although such technological advances in large-scale plant phenotyping can accelerate selection of germplasm with desirable defence traits, the prediction of complex traits and their employment in genetic improvement programs requires a detailed understanding of the genetic architecture underpinning variation in the phenotypic traits.

1.3   Factors determining the response to selection of resistance and associated traits

The phenotypic differences between individual organisms are ascribed to underlying genetic and environmental variation or their interaction. In quantitative genetics, the phenotype (P) is expressed as a function of the genotype (G) and environment (E) and their interaction, where P=G+E+G×E (Falconer and Mackay 1996). Hence, the phenotypic variance (σ^2_P) can be written as $\sigma^2_P=\sigma^2_G+\sigma^2_E+\sigma^2_{G\times E}$, where σ^2_G =genotypic variation, σ^2_E= environmental variation and $\sigma^2_{G\times E}$ = variance due to genotype by environment interaction. Analysis of these variance terms is used to evaluate additive and non-additive genetic effects, heritability as well as expected genetic gain. Heritability can be broad-sense (H^2), *i.e.* the proportion of phenotypic variance attributable to genetic causes, while narrow-sense heritability (h^2) is the proportion attributable to additive gene effects (Falconer and Mackay 1996). A high heritability indicates a high correlation between the genetic effects (breeding value) and the phenotype. In contrast, if environmental or genotype by environment (G × E) effects are large relative to the additive component then the heritability is low (Hallauer *et al.* 2010). G × E can take different forms, including plasticity arising from changes in variance among genotypes across environments, and plasticity resulting from genotype rank changes among environments (Aspinwall *et al.* 2015). Genotypes may also show variable linear or non-linear responses to continuous environmental variation (Aspinwall *et al.* 2015). In the context of the breeder's equation, narrow-sense heritability has been used to predict genetic gain (ΔG), measured by the difference between a selected population and its offspring population where, $\Delta G = h^2 \times$ selection differential. If the phenotypic variance is similar, traits with higher heritability should exhibit higher genetic gain and, hence, faster genetic improvement or breeding progress than those with low heritability. Several studies have documented significant additive genetic variation and low to moderate heritability estimates for resistance and traits associated with herbivory in conifers (Moreira *et al.* 2013b; Raffa and Smalley 1995; Wainhouse *et al.* 2009; Zas *et al.* 2017), as well as significant genetic gain (Zas *et al.* 2017). In practice, selection may not be based just on the individual's own performance, but additionally or exclusively on that of its relatives using a selection index or best-linear unbiased prediction (BLUP). BLUP models have especially become popular

approaches in practical breeding value evaluations because they are simple and have low computational demands (Liu *et al.* 2008)

Genetic gain can also be constrained by antagonistic genetic correlations (trade-offs) between traits, where if two traits are positively correlated, selection on one trait can indirectly improve the other trait through indirect selection. Conversely, negative correlations between traits of interest makes simultaneous selection for both categories a challenge. Similarly, if there is a plant cost to defence, a trade-off between allocation to defence and to other functions that increase plant fitness, such as growth, would be expected in the absence of herbivores (Ivey *et al.* 2009). In the majority of cases, plant species produce more than one defensive chemical compound, and trade-offs could exist in allocation to different chemicals. In conifers, mixed evidence for the existence of trade-offs has been presented (Deslauriers *et al.* 2015; Moreira *et al.* 2014; Villari *et al.* 2014), however, trade-offs are not generally expected between multiple defence traits that complement each other or where resources are not limited (Lamara *et al.* 2018; Sampedro *et al.* 2011).

Despite the potential genetic gain there are still very few examples of operational conifer breeding programs against insect and mammalian herbivores (Alfaro *et al.* 2008; Russell 2008; Sniezko and Koch 2017). Traditional pedigree-based methods of breeding involve recurrent cycles of selection, mating, and testing. These methods face diverse challenges, including often low heritability, long breeding cycles, late flowering, variable juvenile-mature correlations and emerging pests and diseases (Gamal El-Dien *et al.* 2016; Goddard 2009; Hayes *et al.* 2009; Iwata *et al.* 2011; Klápště *et al.* 2018; Stejskal *et al.* 2018; Suontama *et al.* 2018). Estimation of additive genetic variance from among-family variation also requires the existence of large phenotyped and pedigreed populations involving numerous parents and their families. More recently, however, the genomic era has provided the opportunity for the use of genomic tools in breeding programs that can overcome some of the challenges above (Crossa *et al.* 2017; Meuwissen *et al.* 2001).

1.4 Application of molecular tools in plant defences

The use of dense genome-wide single nucleotide polymorphism (SNP) markers in genomic prediction of complex phenotypes has become a promising approach in genetic improvement programs. Genomic prediction can increase the genetic progress of breeding programs by increasing predictive accuracy of breeding values, reducing generation intervals or shortening the breeding cycles. Genomic selection was developed especially for quantitative traits to overcome the challenges of traditional pedigree-based methods and marker-assisted selection (Arruda *et al.* 2016). It is fast becoming popular in plant

breeding because of recent advances in high-throughput marker technologies and accompanying reduction in the costs of genotyping. In conifers, one study did not show any improvement in herbivory resistance through genomic selection (Lenz *et al.* 2020), possibly due to the few markers utilised, and evidence for whether genomic prediction models can improve selection accuracy of herbivory resistance or the associated traits in conifers is not clear. However, evidence for potential improvement in genetic prediction for resistance traits is available from pathosystems (Resende *et al.* 2012b). The benefit of genomics has also been clearly illustrated for other economically important traits (Kliebenstein 2014; Li *et al.* 2016; Resende *et al.* 2012b). In loblolly pine (*Pinus taeda*) for example, prediction accuracy improved by 53–112% using genomic compared with pedigree-based selection for growth traits (Resende *et al.* 2012a).

The use of genomics is based on the principle that phenotypes are the result of variation in gene sequence and gene expression and subsequent molecular modifications that vary across cells, tissues, organisms, populations and environments (Ralph *et al.* 2006; Whitehill *et al.* 2016). The extent to which phenotypic variation is modulated by gene expression has been less studied relative to gene sequence modifications (Idaghdour and Awadalla 2013). Gene expression data can be useful for prediction of complex traits (Gao *et al.* 2017; Li *et al.* 2019). Its importance is based on the premise that most markers and the associated variants are not located in protein-coding regions, except for recent studies that generate genetic markers from protein coding regions through exome capture sequencing (Telfer *et al.* 2019). The transcriptome is also an efficient method for gene discovery and is a major reference for annotation of both coding and non-coding genes; it can provide information to study specific pathways for specific physiological states. In understanding the defence responses that occur following herbivory, studies have shown that large-scale transcriptional responses occur when herbivores attack plants (Kovalchuk *et al.* 2015; Reymond *et al.* 2004), which may act as an intermediate phenotype between genomic DNA sequence variation and more complex cellular, organ, or whole-plant phenotypes.

1.5   Consideration of mammalian herbivores

Most evidence on the role of conifer defences in herbivory comes from insect systems; comparatively fewer studies have examined the relationships between conifers and mammalian herbivores. Mammals may show divergent responses to plant chemical traits due to behavioural, morphological, biochemical, and population-level adaptations (Boyle 1999; Raffa 2014). In conifer–insect systems for example, monoterpenes have been shown to exert major defence roles (Raffa and Smalley 1995; Seybold *et al.* 2006), since they are exploited as primary chemical cues and/or precursors for the aggregation

pheromones of insects. This may not be the case for mammalian herbivores that can rely on visual cues (Stutz *et al.* 2017). Several Australian marsupials for example possess the capability to ingest and metabolise a range of dietary terpenes and phenols that would be toxic to many other herbivore species (Boyle 1999; El-Merhibi *et al.* 2007). Also, in contrast to insect species that complete their life cycle in a single host, for example the well-studied conifer bark beetles (Schowalter 2012), mammalian bark browsers that readily adjust foraging behaviour on spatial and temporal criteria may be less affected by chemical defences (Gill 1992; Miller *et al.* 2014). A few conifer studies have shown genetic variation in susceptibility to mammalian herbivory (Dimock *et al.* 1976; Iason *et al.* 2011; Miller *et al.* 2014; Silen *et al.* 1986) and selection based on chemical traits has been demonstrated in one breeding programme (Russell 2008). However, these studies have been focused on needle herbivory, and genetic variation in mammalian bark stripping has not been examined in conifers.

1.6   The bark stripping problem

Bark stripping mammals may exert selective pressures on conifer populations through their impact on growth, reproduction, and survival (Gill *et al.* 2000; Nagaike 2019; Zamora Nasca *et al.* 2018). In *Picea abies* in the Czech Republic, for example, Cukor *et al.* (2019) showed that bark stripping reduced tree diameters by 64% compared to non-stripped trees. Similarly, in *Picea sitchensis*, survival was reduced by up to 13% in severely bark stripped trees in Scotland (Welch and Scott 1998), while 80% of trees died in subalpine coniferous forests in Japan following bark stripping by deer (Iijima and Nagaike 2015). In addition to direct impacts on mortality and reduced growth, bark stripping can cause wood degradation with adverse impacts on the forest industry. The extent of bark stripping in conifers depends on individual tree characteristics such as size, age and species, bark traits, and the presence of lower branches (Gill 1992; Klich 2017), the species of herbivore (Ligot *et al.* 2013), as well as habitat and environmental characteristics (Di Bitetti 2019; Iijima and Nagaike 2015; Kobashikawa and Koike 2016). Various control and management measures have been implemented to mitigate the effects of bark stripping by mammalian herbivores, which include culling of animals, fencing, wrapping of trees, diversionary feeds and different silvicultural practices (Di Bitetti 2019; Kimball *et al.* 2011; Kobashikawa *et al.* 2019; Smith *et al.* 2020; Turek *et al.* 2016). The use of natural conifer resistance to minimise mammalian bark damage has not been explored.

1.7   *Pinus radiata* in Australia

Pinus radiata (D. Don, Pinales: Pinaceae) is native to coastal California (USA) (Offord 1964) but has become the most widely planted forest tree in the southern hemisphere with extensive plantations in New Zealand, Chile, Australia and South Africa. In Australia, it dominates the softwood plantation

estate (ABARES 2018). Plantations are distributed across five states in New South Wales, Victoria, South Australia, Tasmania and Western Australia and contributed up to 16.8 million cubic metres of saw logs between 2015–19 (ABARES 2018). Tree breeding programs aimed at improving growth, form, health and wood traits of *P. radiata* have been ongoing in Australia and New Zealand since the 1950s (Dungey *et al.* 2009; Wu *et al.* 2007). However, the needles, bark and roots of *P. radiata* are attacked by numerous insects, fungi and mammals (Hernandez-Escribano *et al.* 2018; Mead 2013; Miller *et al.* 2014; Reglinski *et al.* 2017), with up to 600 pests and pathogens recorded (Brockerhoff and Bulman 2014). In Australia, genetic improvement of health has focussed on resistance to *Dothistroma* needle blight and insect pests such as Monterey pine aphid, *Essigella californica* (Ivković *et al.* 2010a; Li *et al.* 2018; Sasse *et al.* 2009). However, bark stripping by the brushtail possum (*Trichosurus vulpecula*), the swamp wallaby (*Wallabia bicolor),* Bennett's wallaby (*Macropus rufogriseus* subspecies *rufogriseus*), grey kangaroo (*Macropus giganteus*), common wombat (*Vombatus ursinus)*, rabbit (*Oryctolagus cuniculus)* and the red-bellied pademelon (*Thylogale billardierii*) has compromised the realization of expected genetic gains in some regions (ABARES 2018). Nevertheless, there is little information on the extent to which observed variation in bark stripping is under genetic control and what the underlying plant traits are driving this variation.

In Tasmania, bark stripping by the native Bennett's wallaby (*Macropus rufogriseus* ssp. *rufogriseus*), eastern grey kangaroo (*Macropus giganteus* ssp. *tasmaniensis*) and the red-bellied pademelon (*Thylogale billardierii*) has become the most important pest problem in conifer plantations (ABARES 2018; Miller *et al.* 2014; Page *et al.* 2013). Smith *et al.* (2020) listed factors that potentially predispose Tasmanian *P. radiata* plantations to bark stripping. These include ease of access, topography and weather conditions although there is still limited support for these factors. Page *et al.* (2013) suggested that increases in the amount of sugars in the *P. radiata* bark relative to other forest vegetation may be responsible for bark stripping, which was not supported by Smith *et al.* (2020) who showed that bark stripping occurs even in presence of other forages. Generally, larger reserves of sugars are known to be associated with increased cold tolerance in pine species (Bansal and Germino 2009; Ögren *et al.* 1997) and are a potential food resource for winter bark stripping herbivores. These studies, however, did not examine intraspecific variation in bark stripping. In other *P. radiata* populations, variation in herbivory has been associated with total secondary metabolites (Moreira *et al.* 2013b), but individual compounds contributing to the variation are still unclear. Various bark stripping management strategies, including provision of diversionary feed and managing animal densities, have been suggested (Page *et al.* 2013), but currently fencing and tree guards are the most commonly used practices. Given that the financial costs of managing mammalian herbivores in *P. radiata* plantations is high, there is untapped

potential to examine the use of natural conifer resistance in mitigating the effects of bark stripping, especially given that intraspecific genetic variation in bark stripping that has been noted in a few families (Miller *et al.* 2014).Therefore, there is a need to assess the potential use of natural resistance to minimize bark stripping by the marsupials in *Pinus radiata* in Australia.

1.8   The thesis

This thesis presents results of an extensive research project examining the genetic-based variability of bark stripping of *P. radiata* and the plant traits conferring variation in resistance. The overall aim was to facilitate the selection of less susceptible germplasm for operational deployment as a strategy to reduce the effects of bark stripping in plantation forestry. This study was based in Tasmania, Australia. The first part of the study investigated if variation in bark stripping is under genetic control and examined the importance of genotype by environment interactions based on additive-pedigree based relationships (Chapter 2). Second, the chemical traits influencing bark stripping were identified; first by understanding plant-wide constitutive and induced chemistry in a more controlled nursery environment (Chapter 3), then, by examining the specific traits that differentiate susceptibility categories in a common garden field experiment (Chapter 4). To understand the genetic control of these chemical traits, near infrared spectroscopy (NIRS) models were developed to facilitate chemotyping of large sample sizes of root, bark and needles collected from a nursery trial (Chapter 5). This methodology was then extended to the field by examining the genetic variation in the NIRS predicted chemistry based on additive-pedigree relationships (Chapter 6). Improvement in heritabilities and breeding value predictions were then tested using genomic models, with Chapter 7 presenting the relative predictive ability of genomic versus pedigree-based models. Finally, a transcriptomic study was undertaken to compare the effects of bark stripping and another stressor on gene expression (Chapter 8).

Chapter 2 - Quantitative genetic variation in bark stripping of *Pinus radiata*

2.0 Abstract

Natural variation in plant susceptibility to mammalian bark stripping may provide a strategy for forest managers to mitigate herbivore impacts in *Pinus radiata* plantations. Predicting the ability of the associated traits to respond to selection requires knowledge about the amount of genetic variation for that trait, but also about all environmental factors. This study examined the extent to which variation in mammalian bark stripping in *Pinus radiata* plantations by native marsupials is under additive genetic control, whether the additive genetic effects are stable across sites (G x E) and explained by physical plant traits such as bark characteristics, stem accessibility and growth rate. Bark stripping was assessed at ages of four and five years in two trials that shared most families, comprising 101 and 138 open-pollinated half-sib families. A third younger trial comprised 74 full-sib control-pollinated families was assessed at two and three years after planting. Significant additive genetic variation in bark stripping was demonstrated in all trials, with narrow-sense heritability estimates between 0.06 - 0.14. Within sites, the amount of additive genetic variation detected increased with the level of bark stripping. Across both sites, the genetic signal was stable (no significant G x E). No significant non-additive effect (family effect) on bark stripping was detected in the full-sib family trial, where an estimated genetic gain of up to 22.1% can be achieved by selecting 20% of the less susceptible families. Overall, results indicate that selection for reduced susceptibility is possible, with potential genetic gains for deployment and breeding. However, the physical traits that were genetically correlated, and likely influenced, the amount of bark removed from the trees by the marsupials appeared to depend upon tree age. In the older trials these traits included bark features (presence of rough bark, rough bark height and bark thickness) whereas in the younger trial where rough bark was not developed, it was the presence of obstructive branches or needles on the stem. In the younger trial, a positive genetic correlation between prior height and bark stripping was detected suggesting that faster growing trees exhibit more bark stripping than slower growing trees. However, the presence of unexplained genetic variation after accounting for these physical factors suggests other explanatory plant traits may be involved such as chemical traits.

2.1 Introduction

Herbivores are important determinants of plant productivity in managed and natural plant systems (Schowalter 2016). In the case of managed conifer forests browsing may have significant economic impacts through the deleterious effects on focal tree growth, reproduction and survival (Ward *et al.* 2004). Potential strategies for forest, tree or herbivore management to reduce these impacts are costly and difficult to implement on a large scale (Kimball *et al.* 2011; Nichols *et al.* 2016). Exploitation of natural resistance is a potential strategy for managing browsing damage (Telford *et al.* 2014). In conifers, natural resistance leading to variation in herbivory damage among and within populations, as well as within individuals and tissues has been reported (Miller *et al.* 2014; Zas *et al.* 2014). This variation, if genetically based, is key to the evolution of resistance mechanisms (Ehrlich and Raven 1964; Maron *et al.* 2019) and a prerequisite for genetic improvement in tree breeding and deployment populations (Falconer and Mackay 1996; White *et al.* 2007). The response to selection described by the concept of genetic gain is measured by the difference between a selected population and its offspring population. The narrow-sense heritability of a trait, defined as the proportion of the phenotypic variance accounted for by additive genetic effects, is an important indicator of the extent to which parents can influence the expression of traits in the progeny (Falconer and Mackay 1996). Significant additive genetic variation in resistance to herbivore damage has been predicted in several conifers (O'Reilly-Wapstra *et al.* 2007; Zas *et al.* 2017; Zas *et al.* 2005).

Despite the potential to exploit genetic variation in resistance to herbivores to improve productivity in conifers, other factors that can constrain a genetic increase in resistance need consideration. Genetic gain can, for example, be constrained if the expression of resistance varies according to environmental conditions. Such genotype by environment interaction (G X E) can result in change in variances among genotypes across environments or the relative resistance rankings of plant genotypes (Klápště *et al.* 2020b; O'Reilly-Wapstra *et al.* 2005; Wise and Rausher 2013). The nature of genetic correlations is also important, where if two traits are favourably correlated, selection can simultaneously improve both by indirect selection. Conversely, adverse correlations between traits of interest makes simultaneous selection for the involved traits a challenge. Additionally, the patterns of resistance to herbivory can change as trees develop, which may be related to toughening of plant tissues or increase in secondary metabolites with age (Barton and Koricheva 2010; Lawrence *et al.* 2003; Russell 2008).

Genetic variation in resistance to insect herbivores has been reported in many conifer studies (Moreira *et al.* 2013b; Raffa and Smalley 1995; Wainhouse *et al.* 2009; Zas *et al.* 2017), and some operational breeding programmes are in existence (Alfaro *et al.* 2008). However, there is limited evidence of

genetically-based variation in susceptibility to mammalian herbivory. A few studies have indicated the existence of genetic variation in needle herbivory by mammals (Dimock *et al.* 1976; Duncan *et al.* 2001; Russell 2008; Silen *et al.* 1986) but not for bark damage. Mammalian bark stripping has been reported on various conifers, including *Pinus radiata*, where it impacts the forest industry through its effect on growth, reproduction, and survival as well as its effect on the timber quality resulting from wood decay (Cukor *et al.* 2019; Iijima and Nagaike 2015; Miller *et al.* 2014; Welch and Scott 1998). However, most studies of mammalian bark stripping in conifers have been undertaken at the phenotypic level and no studies at the genetic level were found. Since investment in defences may vary with plant tissues (Moreira *et al.* 2012a) and genetic correlations in resistance to different herbivores may not be universal in conifers (Iason *et al.* 2011; Moreira *et al.* 2013a), the genetic basis of the observed variation in mammalian bark stripping is unclear.

*Pinus radiata (*D. Don, Pinales: Pinaceae) is a softwood species native to California (Eldridge 1979) but is widely planted with over 4 million hectares globally (Mead 2013). In Australia, it is the major softwood plantation species covering approximately 770,000 ha (ABARES 2018). The plantations are distributed across five states - New South Wales, Victoria, South Australia, Tasmania and Western Australia - and contributed up to 16.8 million cubic metres of saw logs between 2015–19 (ABARES 2018). Tree breeding programmes aimed at improving growth, form, health and wood traits of *P. radiata* have been ongoing in Australia and New Zealand since the 1950s (Dungey *et al.* 2009; Wu *et al.* 2007). Genetic improvement of health has focussed on resistance to diseases, especially the *Dothistroma* needle blight and insect pests such as Monterey pine aphid, *Essigella californica* (Ivković *et al.* 2010a; Li *et al.* 2018; Sasse *et al.* 2009). Variation in resistance to insect pests and pathogens has been shown to have strong additive genetic variation (Graham *et al.* 2018; Moreira *et al.* 2013b; Sasse *et al.* 2009; Suontama *et al.* 2019). However, while bark stripping by native marsupials (wallabies and kangaroos) might compromise the realization of expected genetic gains, especially in Tasmania (ABARES 2018), little is known of the genetic basis of the variation in the damage to assess whether plantations can be genetically improved to reduce damage. In Tasmania, *P. radiata* is the only pine species grown commercially, covering 75,900 ha (Downham and Gavran 2018) and producing over 17,013 m^3 of sawlog and 142,547 tonnes of pulpwood (Forestry Tasmania 2017). Bark stripping by the Bennett's wallaby (*Macropus rufogriseus*) that occurs between the ages of 1-6yrs has become the most important pest problem (ABARES 2018; Miller *et al.* 2014; Page *et al.* 2013), affecting up to 40% of the plantations, with up to 80% of trees affected in some plantations (Miller *et al.* 2014). Given that the financial costs of managing mammalian herbivores through fencing and culling in *P. radiata* plantations

are high (DPIPWE 2018), the use of genetic variation in resistance to bark stripping is of considerable interest as a sustainable avenue for minimising the damage of *P. radiata*.

This study aimed to:

1) determine the extent to which bark stripping is under additive genetic control and if this is stable across sites and age;

(2) determine the genetic correlation between the level of bark stripping, growth, stem and bark traits; and

3) estimate possible genetic gains in reducing bark stripping damage from field-based selection of the least damaged families.

2.2 Materials and methods

2.2.1 Genetic trials

The field trials used in this study were established in Tasmania (Figure 2.1) by Timberlands Pacific Pty with seeds from the Radiata Pine Breeding Company. Three trials were studied – Beulah, Payanna and Wilmot (Table 2.1). The progeny planted in Beulah and Payanna consisted of fourth generation breeding populations of half-sib (open-pollinated; OP) families while that of Wilmot was established from third generation full-sib (cross-pollinated; CP) families. The OP families were assumed to be half-sibs (Bannister 1969). The Wilmot CP families (n = 74 families) was derived from 55 parents and 54 unique grandparents. The OP families at Beulah (n = 101 families) and Payanna (n = 138 families) were derived from 101 and 138 mothers and 194 and 195 grand/great grandparents, respectively, highlighting the high pedigree diversity of the populations in the study and that the results were not biased by a few founder ancestors (Table 2.1). The progeny were from parents that were selected for vigour, stem form, wood properties and branch characteristics (Dungey *et al.* 2009; Wu *et al.* 2007). The trials planted at Beulah and Payanna were planted in the same year with 98 common families, plus 3 families which were unique to Beulah and 40 families which were unique to Payanna. G X E was tested based on these two trials. The younger Wilmot trial did not have parents in common with those represented in the other two trials but had some common ancestry deeper in the pedigree. All trials were replicated in a randomized incomplete block design, with families represented as single tree plots within each block. Seedlings were raised in pots and were planted in rows, at a spacing of 3 m by 3 m. Since trials were established in clear-cut coupes formerly dominated by *P. radiata*, the remaining debris was gathered to form wind rows which separated different sets of blocks within each trial.

In Tasmania, the major bark stripping pest of *P. radiata* - the Bennett's wallaby (*Macropus rufogriseus*) (Page *et al.* 2013) varies in population density depending on location. The Beulah and Wilmot field trials

were situated in mid-north Tasmania where the density of the Bennett's wallaby is estimated as approximately 32.0 animals/km^2 (DPIPWE 2019). The Payanna field trial was situated in the north-east where the estimated density of the Bennett's wallaby is approximately 26.9 animals/km^2 (DPIPWE 2019). Apart from operational culling (DPIPWE 2019), the field trials at Beulah and Payanna were freely accessible by the animals. The Wilmot site was, however, fenced for the first two years of growth. After two years, the gates to the trial were opened during winter for approximately 2 months to allow the marsupial herbivores access to 20 of the 26 replicates (6 replicates spread throughout the trial remained independently fenced). The gates were closed to stop browsing when browsing was evident across all 20 replicates. However, after the first bark stripping assessment, the animals accessed the trees for further bark stripping that was also scored.
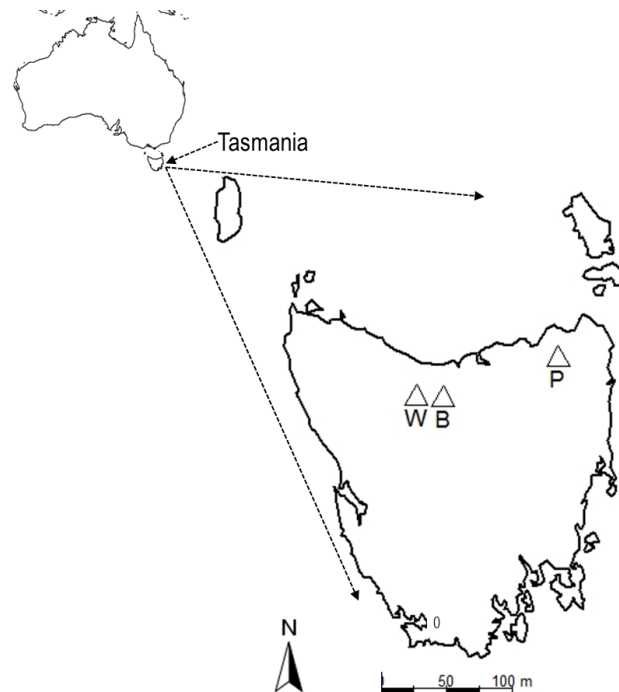
Figure 2.1: A map of Australia showing the location of Tasmania and the location of the three *P. radiata* genetic trials within Tasmania that were studied - P = Payanna, B= Beulah and W=Wilmot.

Table 2.1: Description of the genetic trials used for the study indicating the location and date of establishment. All trials were replicated in a randomized incomplete block design, with families represented as single tree plots within each block. For each trial, the number of replicates, incomplete blocks as well as the families, parents, grandparents and total number of trees are indicated. The total number of trees denotes the trees that were alive at the time of assessment. Selected traits (indicated in Tables 2.2 – 2.4) were measured twice and the age at which the first and second assessments were performed is shown

Genetic trial	Latitude (°)	Longitude (°)	Elevation (m)	Date planted	Replicates	Incomplete blocks	Families	Parents	Grand parents	# trees assessed	Time of first assessment (age years)	Time of second assessment (age years)
Beulah	14.468984	146.41178	336	2011	25	75	101	101	194	2002	2015 (4 yrs)	2016 (5 yrs)
Payanna	41.123591	147.70581	170	2011	20	80	138	138	195	2668	2015 (4 yrs)	2016 (5 yrs)
Wilmot	41.454271	146.10680	580	2015	26	78	74	55	54	1372	2017 (2 yrs)	2018 (3 yrs)

2.2.2   Assessment of bark stripping damage and related traits

Bark stripping damage on the stems of the *P. radiata* trees was recorded visually on individual plants in the field trials. At Beulah and Payanna, the damage was scored on an ordered categorical scale assigning zero (0) to plants with no evidence of bark-stripping, 1 = <20% of the circumference stripped; 2 = 20 - 50% of the circumference stripped; 3 = 50 – 100% of the circumference stripped. At Wilmot, more categories were included i.e., 3 = 50 - 75%, 4 = >75%, 5 = 100% damage (completely ring barked). Except for the scores 0 and 100, the remaining scores were converted to class mid-point values (average of the range values) for final analyses. Other tree traits including tree height, basal diameter (BD), diameter at breast height (DBH = 1.3 m), bark thickness, presence/absence of rough bark, stem access and survival were also assessed (Tables 2.2 – 2.4). Rough bark was assessed by visual inspection of presence or absence of bark fissures (e.g. Figure 2.2), and when present rough bark height was measured as the height from the ground to the top of the rough bark on the side of the stem with the highest rough bark cover. Bark thickness was estimated at DBH with a custom-made bark probe. The trees at the younger Wilmot trial had not yet developed rough bark and were still too small to measure diameter at breast height (DBH) or bark thickness at the time of assessment. Instead, basal diameter and stem access were assessed at 2 years of age. Stem access describes the presence of stem needles and obstructive branches in the first 1 m of the stem which may prevent ready access to the bark. This trait was subjectively assessed in a categorical decile scale, where 0 =all stem covered with needles and branches, 10 = up to 10cm covered, 20 = up to 20cm covered, until 100 where no needles or branches were found within 1 m. Survival (trees alive) at all trials was assessed at the time of the first bark stripping assessment. Some variables were assessed twice (Tables 2.2 – 2.4). The first assessment at Payanna and Beulah was made at 4 years and then at 5 years after planting. At the Wilmot trial the first and second assessments were made at 2 and 3 years after planting, respectively. For bark stripping, recurrent assessments included old damage except where the tree had completely healed with no clear signs of earlier damage. Therefore, the two scores are not independent.

Figure 2.2: Smooth (left) and rough bark (right) of 4-year-old *P. radiata*.

2.2.3 *Spatial analyses*

Presence of spatial heterogeneity in data may cause inaccurate estimation of genetic parameters (Zas *et al.* 2007). To detect small-scale spatial effects within the blocks that were not accounted for by the experimental replicates and incomplete blocks on each site, a spatial term was introduced in the linear mixed models. The general linear mixed model is represented as;

$$y = X\beta + Zu + e, \tag{1}$$

where, y is the vector of phenotypic observations for the traits that were assessed from each site (Tables 2.2 – 2.4). β is a vector of fixed effects and this included the mean and missing values (mv) except where covariates were fitted (Table 2.9). u is the vector of random effects which included replicates, blocks within replicates, tree (additive genetic effect – estimated using the relationship matrix derived from the pedigree file for trial trees and their ancestors) and family (specific combining ability) terms, and e is a vector of random residuals. X and Z correspond to design matrices relating the observations in y to the fixed and random effects in β and u, respectively. The joint distribution of the random terms was assumed to be multivariate normal, with means and (co)variances defined as:

$$\begin{bmatrix} U \\ e \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \right)$$

where 0 is a null matrix, and G and R are (co)variance matrices for effects in u and e, respectively (Costa e Silva *et al.* 2001; Dungey *et al.* 2014; Isik *et al.* 2017). The error term e was then fitted with a two dimensional spatial term (Costa e Silva *et al.* 2001; Dutkowski *et al.* 2006). To set the spatial term, every tree was uniquely identified by a row and column position within each trial, setting the absent, dead, and filler trees to missing values (Costa e Silva *et al.* 2001). The missing values (mv) were included as a fixed factor in the models (Dungey *et al.* 2014). The error term in model 1 was then partitioned into spatially correlated (ξ) and uncorrelated (η) residuals. The spatially correlated error (ξ) was modelled using a first-order separable autoregressive model in the row and column directions (Butler *et al.* 2009; Dungey *et al.* 2013; Suontama *et al.* 2019). However, in addition to the two-dimensional separable first-order autoregressive spatial model, an independent residual (nugget) was also added in the random terms with the following form:

$$\sigma^2 \Sigma = \sigma^2 (\Sigma p_{col} \otimes \Sigma p_{row}) + \psi I_{150} \tag{2}$$

where σ^2 is the spatial variance, $\otimes$ is the Kronecker product, and (p) is a first-order autoregressive correlation matrix with autocorrelation p for columns and rows (Butler *et al.* 2009).

For a visual inspection of the presence of small-scale spatial effects, semivariograms of the residuals from the above spatial models were generated in R v. 3.6.0 (R Core Team 2018) using the sp and gstat packages. The variograms were generated for selected scores of bark stripping and height but

16

the importance of spatial models was tested for all the variables that were assessed (see results). Flat variograms are expected for randomly distributed data. If spatial dependence is present, semivariance will be small at short distances, will increase at intermediate distances, and will reach an asymptote at longer distances (Zas *et al.* 2007). The direction of strongest spatial correlation in each site was used to generate the variogram although there was spatial autocorrelation in all directions (results not shown). Of the five common models (linear, exponential, spherical, Gaussian and Matern), the Matern model was selected as the most suitable suggesting presence of irregular spatial patterns in both height and bark stripping (Haskard 2007).

2.2.4 *Estimation of additive genetic variation and heritability within sites*

Because of spatial effects (see results), spatial models were used to obtain variance components to enable testing of the presence of additive genetic variation for each trait that was assessed (Butler *et al.* 2009). To test whether the additive genetic variation were greater than zero, full models were compared with respective reduced models using one-sided log likelihood ratio tests (LRT) with one degree of freedom in ASReml (Butler *et al.* 2009). Where rough bark height was fitted in the model, all trees without rough bark were fitted as missing values for the trait being modelled. Since the Beulah and Payanna trials comprised only open-pollinated families, the family term was excluded from the models. For the Wilmot trial that had full-sib families, both the additive and family terms were included in the initial models allowing additive genetic variances and specific combining ability to be estimated. The significance of the family term was first tested using one-sided LRT with one degree of freedom in ASReml (Butler *et al.* 2009). However, the family term was excluded in the final models because it was not significant for all traits (see results; Table 2.4). Then, the additive genetic variation was also tested as mentioned above.

For the binary trait i.e. rough bark, a generalized linear mixed model with a binomial distribution and logistic link function was used to estimate variance components (Butler *et al.* 2009). For this model, design (replicates and incomplete blocks) and spatial (autoregressive) terms were fitted as random variables. The importance of the random factors in the binomial models were evaluated using the Akaike information criterion (Burnham and Anderson 2004). The presence of additive genetic variation for survival was not tested due to high survival in all the three trials.

For all models, individual narrow-sense heritability ($\hat{h}^2$) was estimated from univariate spatial models, that included all the design terms as described above as the additive genetic variance divided by the sum of the additive genetic variance $\hat{\sigma}_a^2$ and the error variance $\hat{\sigma}_e^2$ (Butler *et al.* 2009) as below:

17

$$\hat{h}^2 = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_e^2} \qquad (3)$$

For the binary trait, the residual error variance on the underlying logistic scale, which is 3.28987 (Butler *et al.* 2009), was used in the heritability estimations. Estimates of the associated standard error for the estimated heritability were obtained from the average information matrix, using a Taylor series approximation in ASREML (Butler *et al.* 2009).

To test the improvement of the models after fitting the spatial autocorrelation structure to the residuals, the spatial model was compared with the corresponding reduced model that did not fit a spatially correlated residual using a two-tailed likelihood ratio test (LRT) with 3 degrees of freedom (Dungey *et al.* 2013). Significance was considered at the 5% significance level.

2.2.5   Type B genetic correlations

In evaluating G X E interaction, type B genetic correlations are used to assess the relative performance of traits measured in different environments (Burdon 1977). The across-site additive genetic correlation of bark stripping, height, DBH and bark traits was assessed for Payanna and Beulah. The same traits at the two sites were treated as different traits (Yamada 1962). The fixed term was only the mean while the random terms comprised the additive genetic (i.e. tree) term and site-specific replicates and incomplete blocks. The unstructured design allowed covariation between additive genetic variation, but the site-specific terms were considered as independent. Spatial models were not used at this stage to avoid singularities. The additive genetic correlation (r_a) was estimated as:

$$r_a = \frac{cov_a(x,y)}{\sqrt{\sigma_{ax}^2 \cdot \sigma_{ay}^2}} \qquad (4)$$

where $cov_a(x,y)$ is the additive genetic covariance between traits x and y, σ_{ax}^2 is the additive genetic variance components for trait x, and σ_{ay}^2 is the additive genetic variance components for trait y. Standard errors of the genetic correlations were estimated as mentioned above (Butler *et al.* 2009) .

To test whether genetic correlations were significantly different from zero, a full model was compared with the respective reduced model that had the additive covariances fixed to zero using two-sided LRT with one degree of freedom in ASReml. To test whether genetic correlations were equal to one (perfect correlation), a full model was compared with the respective model that had the additive covariances fixed to one using one-sided likelihood ratio tests (LRT) with one degree of freedom in ASReml (Butler *et al.* 2009).

2.2.6 Phenotypic and type A genetic correlations

Within sites (Type A) additive genetic (r_g) and Pearson's phenotypic correlations among traits were estimated directly from non-spatial bivariate models as defined above (equation 4) with mean as the fixed term and the design (replicates and incomplete blocks) as well as the additive genetic component as random terms. The family term was not included for Wilmot estimates since it was not significant (see results). The models allowed covariation at the residual, replicates and incomplete block levels. For binomial models, the bivariate models did not fit a link function (O'Reilly-Wapstra *et al.* 2014). To test whether genetic correlations were significantly different from zero, a full model was compared with the model where additive covariance was fixed to zero using two-sided LRT with one degree of freedom in ASReml (Butler *et al.* 2009). The test that the Pearson's phenotypic correlations were not equal to zero was done in R using the function cor.test.

To test if the measured physical traits fully explained the additive genetic variation in bark stripping, linear models (Equation 1) were re-run with the traits that significantly correlated with bark stripping included as covariates (See results; Tables 2.6 – 2.9). For the selected covariates, the models were modified as follows.

Beulah

- Bark stripping (year 4) = μ + height (year 4) + mv + random + spatial + ε
- Bark stripping (year 5) = μ + bark thickness (year 5) + rough bark height (year 5) + mv + random + spatial + ε

Payanna

- Bark stripping (year 5) = μ + bark thickness (year 5) + rough bark height (year 5) + mv + random + spatial + ε

Wilmot

- Bark stripping (year 2) = μ + stem access (year 2) + mv + random + spatial + ε
- Bark stripping (year 3) = μ + height (year 2) + mv + random + spatial + ε

As previously mentioned, the significance of the additive genetic variation in bark stripping from zero was tested with one-tailed LRT; mv = missing values (that comprised the absent, dead, and filler trees) and ε = independent residual. The random term included the replicates, blocks within replicates and the additive genetic variation.

2.2.7 Estimation of genetic gain

To determine the family rankings for estimation of genetic gain for a single selection event, family best linear unbiased predictors (BLUPs) were estimated for all trials using univariate family models with the spatial term as described above. The additive tree term was not included in the model at this stage, and the differences between families in this case includes additive genetic and specific combining effects. Genetic gain was estimated at the family level since the parents of the families are already available in seed orchards and can be used to generate new seed for selected families for deployment. Family selection is also associated with lower bias than selection within families and strongly correlates with the realized gain than selection within progenies (Viana *et al.* 2009).

The percentage of genetic gain was estimated as the difference between the average BLUPs of the least susceptible families for each trial and the average family BLUPs of the population, divided by the average BLUPs (plus mean score of trait) of the population, multiplied by 100 (Dungey *et al.* 2014; Isik *et al.* 2017). Genetic gain was estimated for selection of a different proportion of families in the different sites. Estimates were based on assessments at year 5 (2016) for Payanna and Beulah and the year 2 assessment (2017) for Wilmot that exhibited the highest heritability estimates.

2.3   Results

2.3.1   *Differences between sites in bark stripping and associated traits*

The percentage of trees exhibiting bark stripping was variable across sites, with 95%, 70% and 52% of the trees experiencing some level of back stripping at Beulah, Payanna and Wilmot, respectively (Figure 2.3). Consistently, the amount of bark removed at Beulah (Table 2.2) was higher than the other two sites (Tables 2. 3 and 2.4). Within site, the mean bark stripping was greater in the second assessment (a year after the first assessment), except for the younger Wilmot trial where slightly lower levels were observed; possibly due to fast wound recovery in young trees. Beulah, which had the highest levels of bark stripping, had a higher proportion of dead or missing trees and the trees were shorter at age 4 years ($\bar{x}$ = 391 ± 89 cm) than the partner trial at Payanna ($\bar{x}$ = 573 ± 94 cm). The DBH of the trees in Beulah was also lower than that of Payanna. However, survival was high in all the sites (Tables 2.2 – 2.4).

The bark traits were different between the older (Beulah and Payanna) and younger (Wilmot) trials. The young Wilmot trees had no rough bark and the bark was too thin to be assessed even at the time of the 2nd assessment at age of 3 years. At 4 years of age, 1% of the trees in Payanna had rough bark compared with 5.3% at Beulah, but by 5 years of age 76% and 42.6% respectively of the trees had rough bark. Of the trees with rough bark at Payanna and Beulah, the mean height of rough bark on

Payanna trees ($\bar{x} = 93.5 \pm 60.4$ cm) was higher than at Beulah ($\bar{x} = 47.9 \pm 64.3$ cm) (Tables 2. 3 and 2.4). These differences in bark development at age 5 years may reflect different growth rates with the faster growing trial at Payanna having a greater rough bark development.
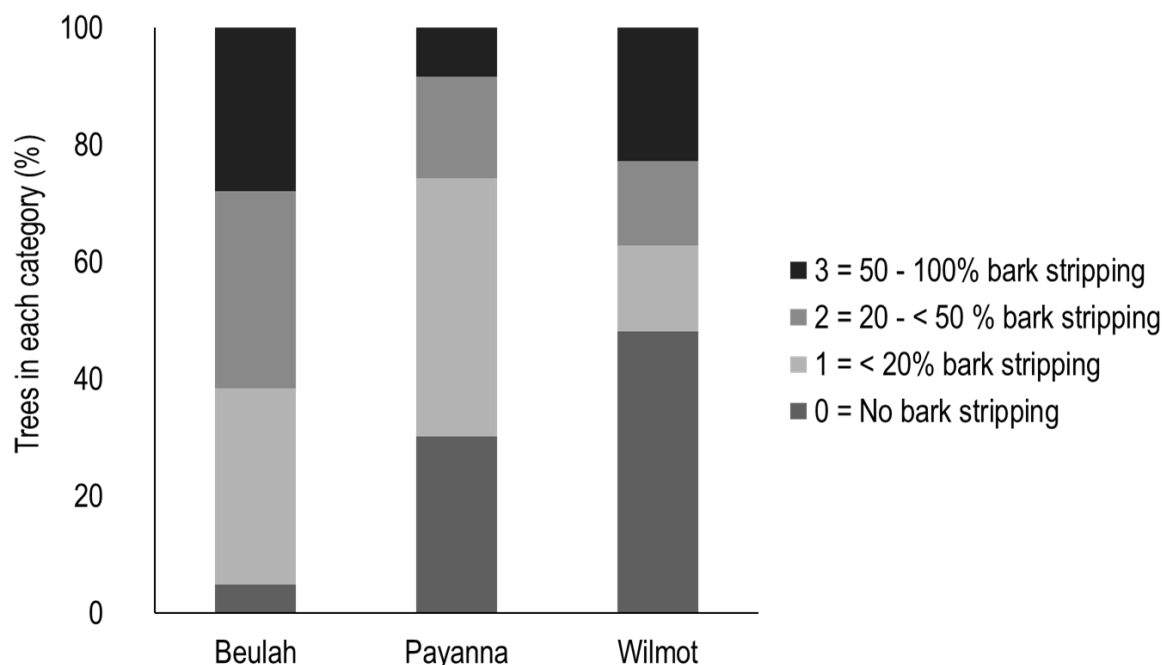
Figure 2.3: The percentage of trees in each bark stripping category at the three field trials – Beulah, Payanna and Wilmot. Bark stripping was recorded as the proportion of the bark that was removed relative to the circumference. 0 indicates no evidence of bark stripping, 1 = trees that had 1 - <20% of the circumference stripped; 2 = 20 - <50%; 3 = 50 - 100%. For this graph the upper 3 categories used at the Wilmot trial were merged to match the categories of the other two trials. The number of trees assessed in the three trials; Beulah (in 2016, age = 5 years), Payanna (in 2016, age = 5 years) and Wilmot (in 2017, age = 2 years) were 2002, 2668 and 1372, respectively.

2.3.2   Spatial effects

Traits assessed in the three field trials were not randomly distributed within the blocks as indicated by the high significance of the spatial term (Tables 2.2 – 2.4). For bark stripping, more damage occurred at the edge of the blocks possibly due to cover from the windrow. For purposes of illustration, Figure 2.4 shows that the scale of the spatial dependence or the point at which the semivariogram tended to be flat at short distances was less than the sizes of the blocks which were on average 15 m in the longest direction.
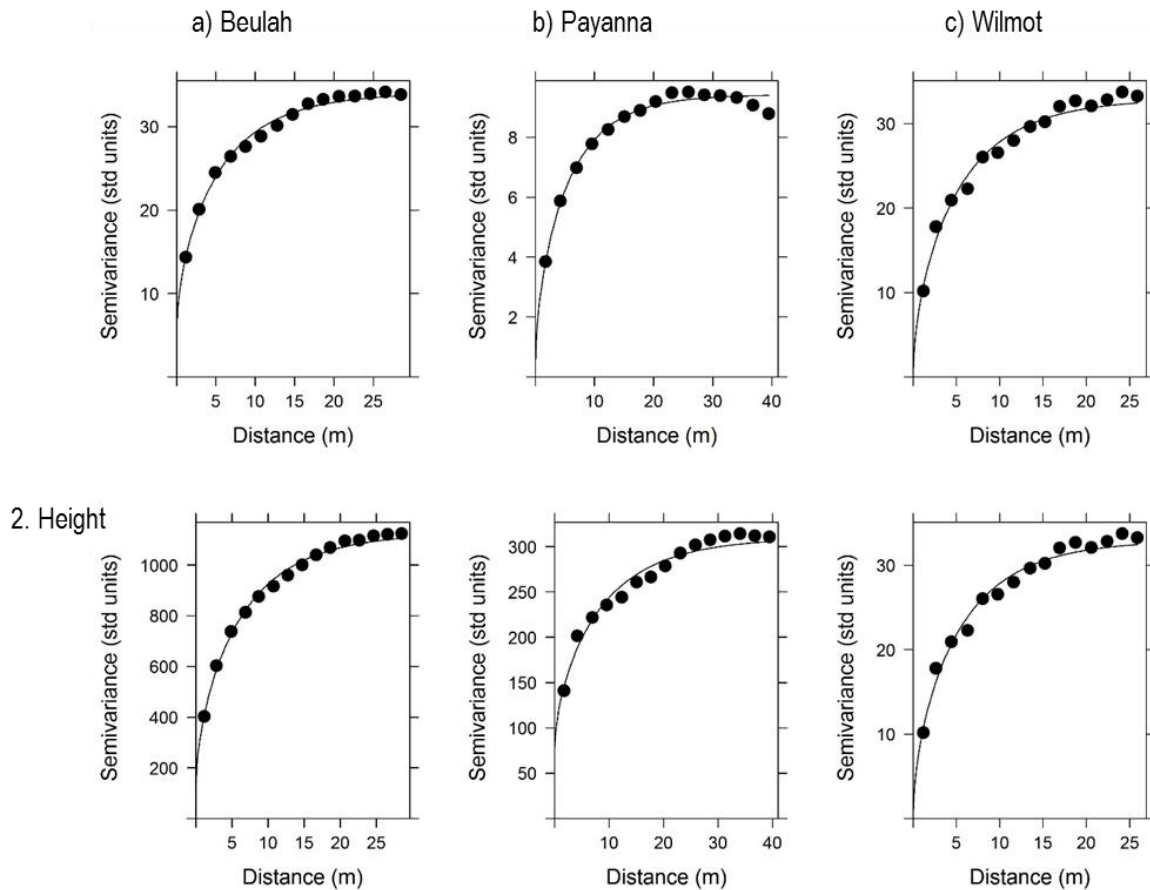
Figure 2.4: Semivariograms of the residuals generated from the spatial models showing small-scale spatial variation of bark stripping (above) and height (below) for a) Beulah, b) Payanna and c) Wilmot after adjusting for random genetic, replicate and incomplete blocks within replicate effects using mixed models. Estimates are based on assessments for year 5 (2016) at Beulah and Payanna and at year 2 (2017) for Wilmot, except for the height variogram for Payanna that used year 4 height since only one height measurement was taken for this trial.

2.3.3 Additive genetic variation for bark stripping

There was significant ($p < 0.001$) additive genetic variation for bark-stripping at all trials for at least one of the yearly measurements, and heritability estimates ranged between 0.06 - 0.14 (Tables 2.2 – 2.4) The least additive genetic variation and heritability estimates were associated with the first bark stripping assessment at Payanna ($p < 0.05$; Table 2.3) and the second assessment in the younger Wilmot trial ($p < 0.05$; Table 2.4). The highest heritability estimate ($h^2 = 0.14 \pm 0.04$) was recorded in 2016 (age 5) at Payanna.

Within site, the time of assessment was important in detecting significant additive genetic variation for bark stripping, and in part this appeared to reflect the intensity of damage. At Payanna, for example, the amount of bark stripping in the first assessment (age 4 years) was lower than in the following year

(means of 8% versus 16%, respectively) (Tables 2.2, 2.3), and exhibited less additive genetic variation and lower heritability suggesting low distinction between preferred and non-preferred families at low bark stripping. However, bark stripping across years was highly additively genetically correlated at Beulah (r_a = 0.78 ± 0.16), Payanna (r_a = 0.91 ± 0.23) and Wilmot (r_a = 0.99 ± 0.27) indicating that family choice by the animals was consistent across years although the scores across years were not independent.

Most other traits assessed showed significant additive genetic variation in one or all trials (Tables 2.2 – 2.4). There was significant additive genetic variation for height in the older Payanna and Beulah trials (p <0.001) as well the younger Wilmot trial (p <0.001). The heritability estimate for rough bark height was tending to 1 in Beulah showing a strong family variation of rough bark. However, heritability reduces as more trees develop rough bark so the distinction between family reduces.

Table 2.2: Statistics for each trait assessed at Beulah. SD = standard deviation, h^2(se)=narrow-sense heritability and its associated standard error (se) estimated with univariate spatial models. The significance of the additive genetic variation and the spatial effect on the model are also presented and were tested with likelihood ratio tests from univariate spatial models. The importance of the binary trait-rough bark (marked *) was tested with the Akaike information criteria (AIC) and + signifies that the parameter tested was important. Genetic variation was not estimated (NE) for survival since it was high for all families

Trait	Year assessed	Age of trees (years)	# of trees assessed	Mean	Minimum	Maximum	SD	h^2 (se)	Significance of additive genetic variation	Significance of spatial model
Bark stripping (%)	2015	4	1957	36.44	0.00	75.00	26.89	0.14±0.05	<0.001	<0.001
Bark stripping (%)	2016	5	2002	40.08	0.00	75.00	26.55	0.12±0.04	<0.001	<0.001
Height (cm)	2015	4	2032	391.18	40.00	707.00	89.38	0.18±0.05	<0.001	<0.001
Height (cm)	2016	5	2002	473.92	1.00	807.00	117.29	0.13±0.05	<0.001	<0.001
DBH (cm)	2016	5	1975	8.67	1.00	14.90	2.46	0.09±0.04	<0.05	<0.001
Bark thickness (mm)	2016	5	1355	6.67	2.50	13.50	1.26	0.10±0.03	<0.01	<0.001
Rough bark*	2016	5	2002	0.42	0.00	1.00	0.49	0.09±0.03	--46.82+	-72.14+
Rough bark height (cm)	2016	5	847	47.88	2.00	420.00	63.21	1.00±0.06	<0.001	0.372
Survival	2016	5	2699	0.72	0.00	1.00	0.45	NE	NE	NE

Table 2.3: Statistics for each trait assessed at Payanna. SD = standard deviation, h^2(se)=narrow-sense heritability and its associated standard error (se). The significance of the additive genetic variation and the spatial effect on the model are also presented and were tested with likelihood ratio tests from univariate spatial models. The importance of the binary trait-rough bark (marked *) was tested with the Akaike information criteria (AIC) and + signifies that the parameter tested was important. Genetic variation was not estimated (NE) for survival since it was high for all families

Trait	Year assessed	Age (years)	Sample size	Mean	Minimum	Maximum	SD	h^2 (se)	Significance of additive genetic variation	Significance of spatial model
Bark stripping (%)	2015	4	2647	8.21	0.00	75.00	15.99	0.07±0.03	<0.05	0.270
Bark stripping (%)	2016	5	2668	16.77	0.00	75.00	21.21	0.14±0.04	<0.001	<0.001
Height (cm)	2015	4	2648	573.92	130.00	827.00	94.06	0.10±0.04	<0.001	<0.001
DBH (cm)	2016	5	1727	7.90	3.00	19.50	6.48	0.03±0.00	0.205	<0.001
Bark thickness (mm)	2016	5	1727	6.01	0.00	16.00	1.78	0.11±0.04	<0.010	<0.001
Rough bark*	2016	5	1694	0.77	0.00	1.00	0.42	0.07±0.04	-29.90[+]	-35.12[+]
Rough bark height (cm)	2016	5	1327	93.54	5.00	350.00	60.26	0.53±0.10	<0.001	<0.050
Survival	2016	5	2760	0.97	0.00	1.00	0.18	NE	NE	NE

Table 2.4: Statistics for each trait assessed at Wilmot. SD = standard deviation, h^2(se)=narrow-sense heritability and its associated standard error (se). The significance of the additive genetic variation, family term (specific combining ability) and the spatial effect on the model are also presented and were tested with likelihood ratio tests from univariate spatial models. The importance of rough bark was tested with the Akaike information criteria (AIC) and + signifies that the parameter tested was important. Genetic variation was not estimated (NE) for survival since it was high for all families

Trait	Year of assessment	Age (years)	Sample size	Mean	Minimum	Maximum	SD	h^2 (se)	Significance of additive genetic variation	Significance of family effect	Significance of spatial model
Bark stripping (%)	2017	2	1372	23.96	0.00	100.00	33.18	0.09±0.03	<0.001	>0.05	<0.001
Bark stripping (%)	2018	3	1269	22.50	0.00	100.00	25.10	0.06±0.03	<0.05	>0.05	<0.001
Bark strip height (cm)	2017	2	706	2.19	1.00	7.00	1.45	0.12±0.04	<0.001	>0.05	<0.001
Height (cm)	2017	2	1372	147.40	10.00	248.00	33.51	0.07±0.04	<0.001	>0.05	<0.001
Height (cm)	2018	3	1275	231.09	30.00	382.00	51.17	0.08±0.04	<0.001	>0.05	<0.001
Height (cm)	2020	5	1230	544.33	40.00	780.00	99.08	0.11±0.04	<0.001	>0.05	<0.001
Basal diameter (cm)	2017	2	140	2.99	1.00	5.30	0.78	0.04±0.05	>0.05	>0.05	<0.001
DBH (cm)	2020	5	1230	103.40	5.00	190.00	42.97	0.03±0.02	<0.01	>0.05	<0.001
Stem access	2017	2	1371	49.49	0.00	100.00	26.26	0.09±0.03	<0.05	>0.05	<0.001
Survival	2017	2	1372	0.86	0.00	1.00	0.40	NE	NE	NE	NE

2.3.4 Genetics x environment interaction

The additive genetic correlations for the same traits assessed in the two partner trials – Beulah and Payanna - are shown in Table 2.5. There was no significant across-site genetic correlation for bark stripping at year 4 ($r_a = 0.23 \pm 0.39$) possibly due to a high randomness of damage at the low damage intensity observed at Payanna. However, with cumulative and new damage at year 5 heritability increased and a significant and high positive across-site additive genetic correlation was observed for bark stripping ($r_a = 0.76 \pm 0.25$). This indicates a high correspondence of family ranks between the two trials at this age. Similarly, high correlations were obtained for year 5 rough bark ($r_a = 0.74 \pm 0.42$) and height at year 4 ($r_a = 0.91 \pm 0.32$).

Table 2.5: Across-site (Type-B) additive genetic correlation (r_a) for each trait assessed in both Beulah and Payanna and the corresponding standard error (se). The chi-square value (χ^2) associated with the likelihood ratio test (LRT) that the p-value of the genetic correlation is different from zero (P[r_a=0]) is also indicated, and was generated with two-tailed LRTs. The significance that the genetic correlation is equal to 1 (P[r_a=1]) was also performed with LRT for correlations that were significantly different from zero. The partner trials at these sites share 98 families and were the same age. The importance of the genetic correlation for rough bark was tested with the Akaike information criteria (AIC) and for this trait + signifies that the parameter tested was important

Trait	r_g	se(r_g)	χ^2 [r_g=0]	P[r_g=0]	P[r_g=1]
Bark stripping (year 4)	0.23	0.39	0.44	>0.05	
Bark stripping (year 5)	0.76	0.25	10.00	<0.01	<0.01
Height (year 4)	0.91	0.32	14.50	<0.001	<0.001
Rough bark (year 5)	0.74	0.42	23.80	-10.72+	-22.53+
Rough bark height (year 5)	0.25	0.21	1.50	>0.05	
Bark thickness (year 5)	0.53	0.85	2.70	>0.05	

2.3.5   *Traits that influenced bark stripping*

The phenotypic (0.34 to 0.71) and genetic (0.78 to 1.00; Type A) correlations between the various measures or ages of bark stripping were generally highly significantly positively correlated within a trial (Tables 2.6 – 2.8). The phenotypic and genetic correlations of bark stripping with growth traits—height and stem diameter—varied from positive to negative (Tables 2.6 – 2.8). In general, bark stripping was significantly negatively correlated with stem diameter at the phenotypic level and while this trend was still evident at the genetic level, these correlations were not significant. The association with height was more variable. At the phenotypic level the correlation was generally negative, but the genetic correlation tended to be negative in the older trials (Tables 2.6 & 2.7). The exception was the year 5 bark stripping at Beulah which was significantly positively correlated with height (r_g = 0.33 ± 0.34), possibly due to the tendency for rough bark height to negatively genetically correlate with height at this age (see below). In the younger Wilmot trial, the phenotypic correlation between bark stripping and height was evident at the onset of bark stripping (age 2 years) but this subsequently became significantly negative consistent with an adverse effect of bark stripping. In contrast, the genetic correlations between bark stripping and height were generally positive indicating that families that were initially faster growing had more bark stripping, although this was only significantly greater than zero in the case of 3year bark stripping and 2-year height.

In the older trials, bark stripping was phenotypically significantly negatively correlated with year 5 bark thickness, the presence of rough bark and the rough bark height, with one exception (Payanna bark stripping at age 4 years versus 5-year rough bark height, possibly because there were very few trees with rough bark at age 4) (Tables 2.6 & 2.7). This negative association was also evident at the genetic level, with significant negative correlations of year 5 bark stripping with bark thickness (r_g = -0.48 ± 0.22), rough bark (r_g = -0.47 ± 0.19) and rough bark height (r_g = -0.37 ± 0.17) but only in Payanna (Table 2.7), suggesting that these traits influence the level of bark stripping at the phenotypic and genetic levels. Compared to Beulah, we noted a rapid increase in the amount of rough bark between 4 and 5 years (77% of the trees in Payanna had rough bark in year 5 compared to 42% at Beulah), emphasizing importance of this trait in minimizing bark stripping but this depends on its extent in the population. In the younger Wilmot trial, rough bark had not developed at the time of bark stripping. However, in this case, both the 2 and 3 year bark stripping and 2 year bark strip height were highly significantly positively correlated with stem access at the phenotypic (0.31 to 0.42) and genetic (0.87 to 1.00) levels (Table 2.8).

The genetic correlation between bark stripping for year 2 and 3 and height or DBH for year 5 in Wilmot were negative suggesting a tendency for bark stripping to reduce performance. A negative genetic correlation was detected between year 5 diameter and bark stripping for year 2 (r_g = -0.84 ± 0.16) and year 3 (r_g = -0.77 ± 0.22). Similarly, a nonsignificant negative correlation between the first height increment (ΔHt2-3) and bark stripping for year 2 (r_g = -0.35 ± 0.27) and year 3 (r_g = -0.26 ± 0.29) was detected, indicating a reduction in performance in the initial years of bark stripping with recovery in the later years. Phenotypic correlations were however mostly lower than genetic correlations suggesting that the use of phenotypic correlations as surrogates for genetic correlations to evaluate traits will result in the underestimation of potential gains from indirect selection.

Table 2.6: Phenotypic (above diagonal) and genetic (below diagonal) correlations of traits assessed at the Beulah trial. The significance of the difference of the genetic correlations from zero was tested using two-tailed likelihood ratio tests, while the test that the Pearson's phenotypic correlations were not equal to zero was performed in R using the function cor.test. Significance levels, ***= $p <0.001$, **= $p <0.01$, *= $p <0.05$. NA not applicable as only assessed on a specific cohort of individuals

	Bark stripping (year 4)	Bark stripping (year 5)	Height (Year 4)	Height (year 5)	Bark thickness (year 5)	DBH (year 5)	Rough bark (Year 5)	Rough bark height (year 5)
Bark stripping (year 4)		0.42 (0.02) ***	-0.26 (0.02)***	-0.29 (0.02)***	-0.18 (0.02)***	-0.26 (0.02)***	-0.08 (0.02)***	-0.17 (0.03)***
Bark stripping (year 5)	0.78 (0.16)***		-0.32 (0.02)***	-0.40 (0.02)***	-0.25 (0.02)***	-0.42 (0.02)***	-0.25 (0.02)***	-0.25 (0.03)***
Height (year 4)	-0.19 (0.25)	-0.09 (0.02)		0.86 (0.01)***	0.49 (0.02)***	0.79 (0.01)***	0.31 (0.02)***	0.13 (0.03)***
Height (year 5)	-0.09 (0.27)	0.33 (0.34)	0.98 (0.04)***		0.50 (0.02)***	0.80 (0.01)***	0.33 (0.02)***	0.10 (0.03)**
Bark thickness (year 5)	-0.24 (0.28)	-0.37 (0.29)	0.24 (0.25)	0.03 (0.33)		0.57 (0.02)***	0.34 (0.02)***	0.34 (0.03)***
DBH (year 5)	-0.26 (0.02)	-0.04 (0.37)	0.78 (0.12)*	0.73 (0.15)	0.40 (0.30)		0.33 (0.02)***	0.17 (0.03)***
Rough bark (year 5)	-0.27 (0.22)	-0.25 (0.22)	0.12 (0.21)	-0.05 (0.24)	0.31 (0.23)	0.21 (0.27)		NA
Rough bark height (year 5)	-0.39 (0.27)	-0.52 (0.26)	0.02 (0.01)	0.02 (0.018)	0.10 (0.02)	0.37 (0.09)*	NA	

Table 2.7: Phenotypic (above diagonal) and genetic (below) correlations of the traits that were assessed in Payanna trial. The significance of the difference of the genetic correlations from zero was tested using two-tailed likelihood ratio tests, while the test that the Pearson's phenotypic correlations were not equal to zero was performed in R using the function cor.test. Significance levels, ***= p <0.001, **= p <0.01, *= p <0.05. NA not applicable as only assessed on subset of individuals which had developed rough bark

	Bark stripping (year 4)	Bark stripping (year 5)	Height (year 4)	Bark thickness (year 5)	DBH (year 5)	Rough bark (year 5)	Rough bark height (year 5)
Bark stripping (year 4)		0.34 (0.02)***	-0.30 (0.02)***	-0.18 (0.02)***	0.01 (0.02)	-0.12 (0.02)***	0.00 (0.02)
Bark stripping (year 5)	0.91 (0.23)**		-0.28 (0.02)***	-0.21 (0.02)***	-0.11 (0.02)***	-0.29 (0.02)***	-0.17 (0.03)***
Height (year 4)	-0.12 (0.34)	-0.20 (0.24)		0.36 (0.02)***	0.22 (0.06)***	0.23 (0.02)***	0.19 (0.03)***
Bark thickness (year 5)	-0.34 (0.51)	-0.48 (0.22)*	0.24 (0.27)		-0.02 (0.02)	0.26 (0.02)***	0.29 (0.03)***
DBH (year 5)	-0.34 (0.33)	-0.40 (0.34)	0.89 (0.27)**	0.56 (0.42)		0.02 (0.02)	0.32 (0.03)***
Rough bark (year 5)	0.05 (0.31)	-0.47 (0.19)*	-0.41 (0.28)	0.52 (0.25)	-0.58 (0.57)		NA
Rough bark height (year 5)	-0.16 (0.25)	-0.37 (0.17)*	0.05 (0.21)	0.70 (0.17)**	0.11 (0.37)	NA	

Table 2.8: Phenotypic (above diagonal) and genetic (below) correlations of the traits that were assessed in Wilmot trial. Significance of genetic correlations was tested as being different from zero using log likelihood tests, while the test that the Pearson's phenotypic correlations were not equal to zero was performed in R using the function cor.test. Significance levels, ***= p<0.001, **= p <0.01, *= p <0.05, ΔHt2-3= Height (year 3) minus Height (year 2), and ΔHt3-5 = Height (year 5) minus Height (year 3).

	Bark stripping (year 2)	Bark stripping (year 3)	Height (year 2)	Height (year 3)	Height (year 5)	diameter at 10 cm (year 2)	DBH (year 5)	Stem access (year 2)	Bark strip height (year 2)	ΔHt2-3	ΔHt3-5
Bark stripping (year 2)		0.71 (0.02)***	040 (0.03)***	-0.24 (0.03)***	0.43 (0.04)***	-0.23 (0.08)**	-0.54 (0.02)***	0.41 (0.02)***	0.89(0.02)***	-0.46 (0.02)***	0.45 (0.04)***
Bark stripping (year 3)	1.00 (0.04)**		-0.82 (0.31)*	-0.24 (0.03)***	0.50 (0.04)***	-0.29 (0.55)	-0.29 (0.03)***	1.00 (0.09)***	0.62 (0.02)***	-0.39 (0.03)***	0.52 (0.04)***
Height (year 2)	0.56 (0.32)	0.00 (0.03)		0.80 (0.02)***	0.53 (0.02)***	0.70 (0.06)***	0.49 (0.02)***	-0.11 (0.03)***	0.10 (0.03)***	0.59 (0.03)***	0.48 (0.03)***
Height (year 3)	0.12 (0.)	0.34 (0.41)	0.79 (0.17)*		0.69 (0.18)***	0.65 (0.07)***	0.81 (0.01)***	0.40 (0.38)	-0.17 (0.03)	0.75 (0.01)***	0.63 (0.03)***
Height (year 5)	0.42 (0.26)	0.38 (0.32)	0.45 (0.28)	0.77 (0.16)**		-0.82 (1.60)	0.44 (0.30)	0.49 (0.27)	0.53 (0.23)	0.49 (0.02)***	0.89 (0.01)***
Diameter at 10 cm (year 2)	-0.05 (0.57)	-0.27 (0.07)*	0.22 (0.55)	-0.20 (0.62)	0.47 (0.05)***		0.72 (0.02)***	-0.20 (0.55)	0.03 (0.56)	-0.12 (0.08)***	0.34 (0.06)***
DBH (year 5)	-0.84 (0.16)**	-0.74 (0.24)*	-0.44 (0.42)	0.49 (0.29)	0.90 (0.01)***	0.75 (0.29)		-0.66 (0.24)	-0.88 (0.15)**	0.62 (0.02)***	0.72 (0.01)***
Stem access (year 2)	0.95 (0.10)***	0.36 (0.03)***	-0.76 (0.25)*	-0.28 (0.03)***	0.50 (0.04)***	-0.24 (0.08)**	-0.42 (0.03)***		0.42 (0.02)***	-0.34 (0.03)***	0.50 (0.04)***
Bark strip height (year 2)	0.97 (0.02)***	0.98 (0.02)***	0.66 (0.28)	0.27 (0.33)	0.34 (0.03)***	-0.13 (0.08)	-0.47 (0.02)***	0.87 (0.13)***		-0.44 (0.03)***	0.35 (0.04)***
ΔHt2-3	-0.35 (0.27)	-0.20 (0.35)	0.38 (0.33)	0.89 (0.10)**	0.65 (0.22)*	-0.15 (0.59)	0.68 (0.23)	-0.18 (0.33)	-0.25 (0.29)		0.53 (0.03)***
ΔHt3-5	0.47 (0.27)	0.38 (0.31)	0.33 (0.34)	0.53 (0.27)	0.94 (0.04)***	-0.99 (0.76)	0.19 (0.37)	0.47 (0.27)	0.55 (0.22)	0.44 (0.29)	

1

To test if the observed additive genetic variation in bark stripping could be fully explained by tree physical traits, linear models were run with traits that significantly correlated with bark stripping (i.e. height, stem access, rough bark and bark thickness) as covariates (Table 2.9). However, additive genetic variation in bark stripping was still significant after accounting for this covariation at the phenotypic level. In the younger Wilmot trial for example, after simultaneously including stem access as a covariate in the model for year 2 bark stripping, the additive genetic variation was still significant (LRT χ^2 = 15.6, p<0.001) but heritability reduced from h^2 = 0.09 ± 0.03 to h^2 = 0.05 ± 0.02. Similarly, in Payanna and Beulah, the additive genetic variation for bark stripping was significant after accounting for covariation with selected traits (Table 2.9). The heritability in Beulah improved after fitting covariates but that of Payanna reduced.

Table 2.9: Genetic estimates derived from models that accounted for covariation. Additive genetic variation (V_a), h^2(se) = narrow-sense heritability and its associated standard error (se) estimated with univariate spatial models. The significance of the additive genetic variation was tested with likelihood ratio tests (LRT). The LRT chi-square value (χ^2) associated with the additive genetic variation is also indicated. mv = missing values, ε = residual. The random term included the replicates, blocks within replicates and the additive genetic variation

Model	LRT χ^2 [V_a>0]	P-value [V_a>0]	h^2 ± se
Beulah			
Bark stripping (year 4) = μ + height (year 4) + mv + random + spatial + ε	18.00	<0.001	0.14 ± 0.05
Bark stripping (year 5) = μ +bark thickness (year 5) + mv + rough bark height (year 5) + random + spatial + ε	20.28	<0.001	0.17 ± 0.05
Payanna			
Bark stripping (year 5) = μ + bark thickness (year 5) + mv + rough bark height (year 5) + random + spatial + ε	21.96	<0.001	0.12 ± 0.04
Wilmot			
Bark stripping (year 2) = μ + stem access (year 2) + mv + random + spatial + ε	15.60	<0.001	0.05 ± 0.02
Bark stripping (year 3) = μ + height (year 2) + mv + random + spatial + ε	11.90	<0.001	0.06 ± 0.02

2.3.6   Estimation of genetic gain

Up to 37% reduction in future bark stripping was predicted based on family selection using the best linear unbiased prediction (BLUPs) (Figure 2.5). The highest predicted genetic gain was established in the younger, full-sib progeny at the Wilmot trial. The half sib-OP families at Payanna and Beulah gave relatively low genetic gain predictions. Selecting 20% of the least susceptible families in each trial resulted in an expected gain in reduction of bark stripping of 7.9%. at Beulah, 3.8% at Payanna and 22.1% at Wilmot.
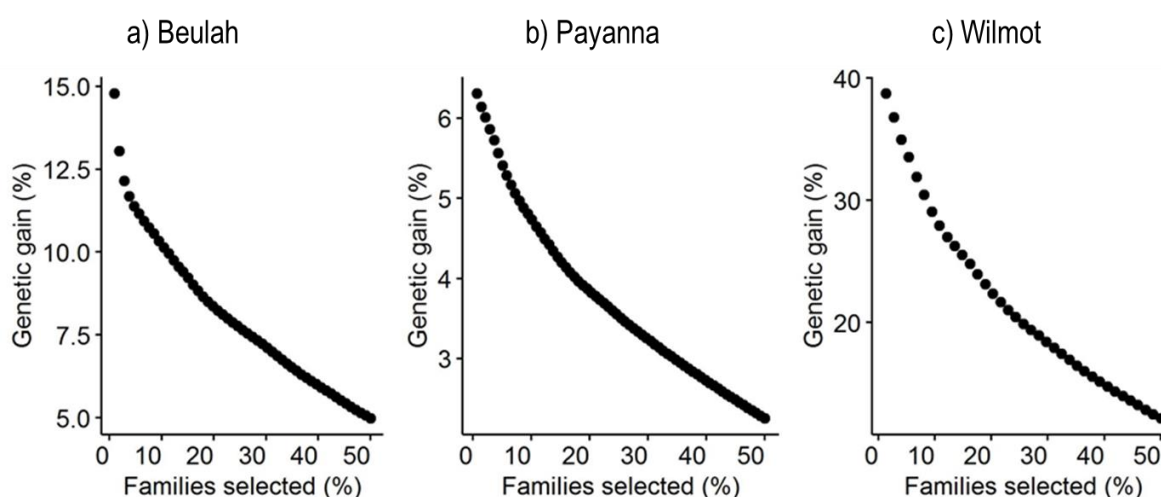
Figure 2.5: Predicted genetic gain (%) in reduction in the average amount of bark stripping as a function of the proportion of families selected in the different sites. Estimates are based on family BLUPs for year 5 (2016) at (a) Beulah and (b) Payanna and at year 2 (2017) for (c) Wilmot.

2.4   Discussion

The current study used several field-based genetic trials established with a large number of families to show the existence of additive genetic variation in susceptibility to bark stripping damage among *P. radiata* families and that when well-expressed appears relatively stable across sites. Depending on plantation age, bark features and the presence of needles or branches covering the stem were an important determinant of the amount of bark removed from the trees by the marsupials. Importantly, these physical traits are under significant additive genetic control and can be enhanced through selection. Up to 22.1% reduction in bark stripping can be achieved by selecting 20% of the less susceptible families although the stability of the less susceptible families, when planted separately needs further testing. Whereas the highest gains were detected in the younger Wilmot trial, lower genetic gains in the older trials could result from the development of other features such as thick bark that reduce the damage and possibly the differentiation of damage between families.

The estimated narrow-sense heritability of bark stripping (0.06 to 0.14) was within the range of what has been reported for the damage of conifer bark from insect herbivores (0.02-0.40) (de la Mata *et al.* 2017; King *et al.* 1997; Moreira *et al.* 2013b; Yanchuk *et al.* 2008; Zas *et al.* 2017; Zas *et al.* 2005). Similarly, it is in the range of the narrow-sense heritabilities reported for resistance to pathogens (h^2 = 0.05 to 0.69) in Australian *P. radiata* plantations (Li *et al.* 2018). Although heritability was also lower than what was observed for *P. radiata* damage by an insect defoliator *Thaumetopoea pityocampa* (h^2= 2.04 ± 1.29) in this case the estimate was inflated possibly by a very small sample size or hidden relatedness of the individuals used (Moreira *et al.* 2013b). Estimates were also markedly lower than what has been reported for mammalian browsing on the needles of Douglas fir (h^2 = 0.73) (Silen *et al.* 1986). Since *P. radiata* has not coevolved with the marsupials and animals have generally had a relatively minor impact on the natural *P. radiata* stands in California, its place of origin, the presence of genetic variation could be a by-product of variation in resistance traits to other native pests in California where *P. radiata* originates (Mead 2013).

Within a particular site, results indicated stability of genetic variation in the different years depicted by the high genetic correlation of the damage between years. Phenotypic correlations were moderate but significant in most cases. This may show that the families damaged in one year are consistently damaged in the subsequent year or it may be an effect of confounding since year 2 scores included old damage. Although the plant traits that correlate with bark stripping are likely to change with age as depicted by the difference in bark traits at 4 and 5 years, the changes are possibly consistent among the families for each year. The stability of the genetic signal between periods of contrasting developmental stages, (for example at 2 years before rough bark development vs 5 years) is however not known. In the different years the genetic estimates are likely to be in part influenced by the amount of bark stripping. For most biotic stresses, differentiation between susceptibility categories becomes clearer with increase in damage intensity since high damage levels reduce the possibility of random escape from damage (Dieters *et al.* 1996; Dungey *et al.* 2009; King and Alfaro 2009; Zas *et al.* 2017). However, cumulative damage estimates are contingent upon survival of the trees and the rate of bark recovery after the initial damage. In the Wilmot site, for example, the cumulative bark stripping was lower possibly due to quicker recovery of the younger trees with low damage and the death of heavily damaged trees (but these factors were not assessed). The threshold damage to enable differentiation of the genotypes is also not known since susceptible individuals may appear resistant if they are located in areas of low bark stripping intensity and conversely the resistance in less susceptible genotypes can be overcome in regions of high intensity damage. Therefore, to correctly identify less

susceptible genotypes, selected families should be tested in areas of high bark stripping intensity. However, extreme browsing intensity may also lead to browsing of resistant families.

Across sites, results showed that when genetic variation is well expressed the bark stripping in Beulah and Payanna, was strongly genetically correlated i.e. there was no evidence of G X E, reflecting the fact that genes influencing bark stripping and their expression is similar in the different environments. This suggests that the relative resistance of different families was not dependent on the environment and that the families that received relatively less bark stripping in Payanna were less bark stripped in Beulah and vice versa. Available studies that have shown strong positive correlations between the trials for damage by insect herbivores in *Pinus radiata* (Sasse *et al.* 2009) and in spruce (Alfaro *et al.* 2008; Mottet *et al.* 2015; Zas *et al.* 2017) even when the pest was present of low density (Alfaro *et al.* 2008). However, in the present case the high genetic correlations signalling low G X E were true for trials of the same age that had similar physical features to deter bark stripping, and how G X E is affected when the families are at different development stages needs to be tested.

This study provides correlative evidence that physical traits play a role in moderating feeding preferences in the field. The role of the physical traits appears to vary depending on the age of the trees. In older trees (>3yrs) the physical traits of the bark (i.e. thick and rough bark) may offer protection against bark stripping. In the younger trees <3yrs, because the bark features are not well developed, the presence of obstructive needles and low branches are a significant explanatory factor. In *Pinus radiata*, bark thickness is positively correlated with age, height and DBH (Murphy and Cown 2015). The negative association of bark stripping with rough bark development in *Pinus radiata* has been previously noted (Miller *et al.* 2014). Studies in other conifers have also indicated the importance of rough and thick bark (Jiang *et al.* 2005; Kuiters *et al.* 2006; Månsson and Jarnemo 2013; Nopp-Mayr *et al.* 2011) and other physical features such as obstructive branches on the stem, in deterring mammalian herbivores (Månsson and Jarnemo 2013). Thick bark may be difficult to detach but has also been associated with high density of resin canals and gritty-textured sclereids (King *et al.* 2011; vanAkker *et al.* 2004). These traits often occur simultaneously but the relative importance of each trait is not known. In contrast, to the negative association observed for mammal damage, damage by insects has been positively correlated with rough bark in conifer species (Ferrenberg and Mitton 2014), which highlights that resistance strategies may be herbivore specific. Rough bark supports wood-boring insects as the fissures are a safe place for oviposition or protection from natural enemies and environmental extremes. It is possible these organisms may be unpalatable to bark stripping animals which could further contribute to reducing susceptibility. The bark features showed substantial

genetic variation and with no genetic constraints to their selection indicated by the tendency for positive genetic correlations among themselves Similarly, the bark features showed no obvious G X E interaction. The low G X E in the bark traits may partly explain the low G X E for bark stripping.

Height also genetically correlated with bark stripping although the trend contrasted in the trials of different ages. In the young population (Wilmot), the herbivores appeared to be attracted to faster growing trees in contrast to the two older plantations where faster growing trees were less damaged. Comparison of relative size of trees damaged by insect and mammalian herbivores in other *P. radiata* populations (Zas *et al.* 2008), other conifer (Mottet *et al.* 2015; Zas *et al.* 2005) and non-conifer (Mayle *et al.* 2009) species suggests that trees most often damaged are those which are larger, dominant and growing most rapidly. Damage to larger trees has been correlated with high sap volume, phloem thickness and sap sweetness (Kenward *et al.* 1996). Also, faster growing trees are known to invest less in defence (Ferrenberg *et al.* 2015). In young trees, bigger trees possibly give more bark per unit stripping than smaller trees and could possibly produce have more sugars. In the study populations however, bigger trees were also associated with more obstructive branches which may have countered the positive effect of height on bark stripping to a certain extent. However, it has also been suggested that fast growing trees should be able to recover more quickly than slow growing trees (Gianoli and Salgado-Luarte 2017) - an aspect of tolerance that needs further research for *P. radiata* bark stripping. For commercial *P. radiata* plantations intended for timber production, tolerance will however be less desirable since bark stripping exposes tissues to fungal attack with subsequent rotting which reduces timber quality (Cukor *et al.* 2019). Bark stripping may also reduce tolerance to freezing that may be associated with subsequent chemical changes (Fedderwitz *et al.* 2020). In contrast, results showed that in older conifer plantations animals prefer smaller trees possibly with less developed bark features, depicted in the tendency for positive correlation between rough bark or bark thickness with height and DBH. Selectivity of smaller trees by bark stripping mammals seems to be the norm in mature conifer plantations (Akashi and Terazawa 2005; Jiang *et al.* 2005; Månsson and Jarnemo 2013). However, the presence of residual variation that could not be explained by physical or growth traits suggests the possible involvement of chemical features in driving differences in bark stripping. This is yet to be tested.

Assuming the family ranking for bark stripping is not expected to change in different environments as this study suggests, the genetic gain for resistance in bark stripping was estimated to be between 4-19% when 20% of the most resistant families are selected for deployment. This is in the genetic gain range that has been estimated for herbivory in conifers (Silen *et al.* 1986). In Douglas-fir, based on

37

selecting 10% of the most resistant families against deer browsing, genetic gain was estimated at 11% (Silen *et al.* 1986). Similarly, against the pine weevil in *Picea abies*, genetic gain varied between 8-50% (Zas *et al.* 2017). However, the genetic gain will be influenced by the proportion of the less susceptible families selected and applies to current field setting. Whether or not the less susceptible families retain their rank when grown separately needs to be tested.

2.5   Conclusion

Variation in bark stripping is under low but significant additive genetic control and when well expressed the genetic signal appears to be relatively stable across different environments. This provides an opportunity for selection for reduced susceptibility with potential genetic gains for deployment and breeding. Based on phenotypic and genetic correlations several physical plant traits were identified as likely contributing to the variation in bark stripping. Initial variation in damage may be affected directly by plant size and accessibility but later, factors like bark thickness and bark texture become important. However, the presence of unexplained genetic variation after accounting for these physical factors suggests other explanatory plant traits may be involved such as chemical traits.

Chapter 3: Variation in constitutive and induced chemistry in the needles, bark and roots of *Pinus radiata*

3.0   Abstract

The capacity to cope with pests and pathogens depends in part on the variation of defences within trees. In conifers, studies have shown extensive variation in allocation of secondary metabolites to above ground plant parts. Plant-wide variation in constitutive and induced chemistry that includes the chemistry of roots has rarely been studied, and few studies have also considered the distribution and role of primary metabolites like the sugars in defence or susceptibility. Here we examined the constitutive and induced variation of primary and secondary metabolites in the needles, bark and roots of *Pinus radiata*. A total of 81 compounds were examined. Results indicate differential constitutive qualitative and quantitative allocation of individual sugars, fatty acids, mono-, sesqui- and di-terpenoids as well as phenolics between the needles, bark and roots. Constitutively, the bark overall had more compounds and a higher amount of most secondary compounds. Within plant parts, the allocation of compounds in the roots differed from that of the needles and bark, for example, glucose dominated in the needles and bark and fructose dominated in the roots. Of the fully identified secondary compounds, monoterpenoids dominated in all the three plant parts but with different qualitative patterns. Following methyl jasmonate and artificial bark stripping treatments, a marked reduction in sugars but weaker changes in secondary compounds were detected in the needles and bark. Responses in the roots were minor but the few that were detected were mostly in response to the bark stripping treatment. Changes in correlations among chemicals within plant parts and between the same compound across the different plant parts were also detected after stress treatments. Overall, results showed that the constitutive composition in the roots differs from that of the bark and needles in *P. radiata* and inducibility is stronger in the primary than secondary metabolites. Differential response of the plant parts to treatments was also detected. This detailed assessment of *P. radiata* chemistry in the needles, bark and roots, before and after stress will potentially facilitate the identification of traits associated with susceptibility or resistance to mammalian bark stripping.

3.1   Introduction

The chemistry of different parts of a plant may not be homogeneous given the different roles they play in plant-environment interactions. In conifers, differences in chemical composition and concentration between plant parts of individual trees of the same species have been well-studied from the perspective of defences (Moreira *et al.* 2012a; Tomlin *et al.* 2000). The main secondary metabolites, the terpenes and phenolics (Franceschi *et al.* 2005), have been implicated in resistance to various pests and pathogens in the needles and bark (Lundborg 2016; Moreira *et al.* 2012a; Reglinski *et al.* 2017). The defences that are constitutive or induced locally at the site of damage or plant/tissue wide (i.e. systemically), can differ among plant parts (Franceschi *et al.* 2005; Iason *et al.* 2011; Moreira *et al.* 2013a; Moreira *et al.* 2012a). The time course of induced defences following herbivore cues or chemical elicitors such as methyl jasmonate has also been shown to differ for different compounds and between plant parts (Lewinsohn *et al.* 1991; Miller *et al.* 2005). Various theories explain the spatial and temporal distribution of chemistry in plants. However, within an individual, the optimal defence hypothesis predicts that defences will be concentrated in parts or tissues that are at higher risk of herbivory and/or tissues that are more valuable (McKey 1974). In conifers, more studies have focussed on above ground parts showing differential allocation of chemistry between needles, bark or xylem (Lewinsohn *et al.* 1991; Miller *et al.* 2005). However, plant-wide variation in constitutive and induced chemistry that includes the chemistry of roots has rarely been studied.

Roots can be equally at risk from pests and pathogens as above ground parts, especially in the juvenile stages (Hernandez-Escribano *et al.* 2018; Moreira *et al.* 2012a; Senior *et al.* 2018) and consequently might also be well defended. Roots can also be involved in above-ground responses to stresses by their direct effects on water and nutrient acquisition, and through correlation, trade-offs or constraints with the bark and/or needles (Huber *et al.* 2005; Moreira *et al.* 2012b). However, few studies have investigated conifer root defences and the relationship between below and above ground defences is poorly understood (Moreira *et al.* 2012b; Poopat 2013). Results from these studies indicate that there is variability in root secondary chemistry, and this can change in response to different treatments. For example, in response to above ground treatment with methyl jasmonate, Poopat (2013) and respectively showed changes in some root monoterpenes in *Pinus edulis* and root physiological traits in *Pinus pinaster*. Similarly, application of methyl jasmonate to the roots of *Pseudotsuga menziesii* caused significant changes in the above ground tissues (Huber *et al.* 2005). In *Pinus radiata*, a link between below and above ground defences has been signalled by the elevated above ground resistance to stem infections following root exposure to *Fusarium circinatum* which causes pitch canker (Swett and Gordon 2017). Root defences are particularly relevant given the

importance of chemical communication at the root-soil interface (Senior *et al.* 2018), the effects of plant-derived metabolites on the soil microbiome assemblage (Senior *et al.* 2018), and the possible influence of this microbiome on plant resistance (Mhlongo *et al.* 2018).

While secondary metabolites are often implicated in plant defence, there is increasing evidence that sugars, such as glucose, fructose and sucrose, play direct and indirect roles in tolerance and resistance in many herbivore-plant systems. Sugars are the primary chemical substrates for structural material in plants other than lignin, but as part of the pool of non-structural carbohydrates (Hartmann and Trumbore 2016) they provide energy for defence responses and may also act as signal molecules (Schwachtje and Baldwin 2008; Tauzin and Giardina 2014). The biosynthesis of terpenes has a metabolic cost (Gershenzon 1994), in which case, a high supply of sugars should lead to increased resistance (Clancy 1992). A positive correlation between primary and secondary chemistry has also been shown in other conifers (Sampedro *et al.* 2011; Villari *et al.* 2014). Studies also indicate that reallocation of sugars in stressed plants is a major consequence of the induced response to herbivory. Reduction of sugars in the stem tissues following application of methyl jasmonate was for example reported in *P. pinaster* (Sampedro *et al.* 2011), which may explain the increase in biomass in fine roots (Moreira *et al.* 2012b). In other trees, evidence of herbivore-induced resource reallocation to roots in response to herbivory, including increased transport of sugars, has been demonstrated (Babst *et al.* 2008). However, the direction of resource movement will depend on the kind of herbivory (Frost and Hunter 2008). Nevertheless, roots are a key sink for sugars (Babst *et al.* 2005), and this sink of non-structural carbohydrates is important for enhancing tolerance (Zhou *et al.* 2015). The importance of sugars in defence may also be linked to their potential as a food source for herbivores, in which case they may attract herbivores in contrast to their positive roles in resistance and tolerance. Although no studies show that sugars alone are sufficient to differentiate resistant and susceptible hosts, there is a clear indication that sugars play a role in host selectivity (Kurek *et al.* 2019; Snyder 1992). Few studies have also considered the distribution of primary metabolites like the sugars in conifers (Cranswick *et al.* 1987).

The conifer *Pinus radiata* (D. Don, Pinales: Pinaceae) is native to California (Axelrod 1988) and is one of the most widely planted commercial timber trees in temperate regions of the world, mainly because of its fast growth rate and wood quality (Burdon *et al.* 2017). However, the needles, bark and roots of *P. radiata* are attacked by numerous insects, fungi and animals (Hernandez-Escribano *et al.* 2018; Mead 2013; Miller *et al.* 2014; Reglinski *et al.* 2017), with up to 600 pests and pathogens recorded (Brockerhoff and Bulman 2014). While the constitutive and induced chemistry of *Pinus radiata* needles

and bark have been reported (Bonello *et al.* 2001; Moreira *et al.* 2012a), there are no comparative studies which include roots. There is also limited understanding of the timing of induction across plant parts. Of particular interest is the response of bark chemistry as bark stripping by mammalian herbivores has become a major problem in many *Pinus* species (Arhipova *et al.* 2015), including in *P. radiata* plantations in Australia (Miller *et al.* 2014; Page *et al.* 2013). The chemical defences of *P. radiata* against bark stripping by mammals have not yet been investigated. In other conifer species, very few studies have documented defences against mammalian herbivores and the role of terpenes and phenolics remains unclear (Bucyanayandi *et al.* 1990; Ilse and Hellgren 2007; Snyder 1992; Zhang and States 1991). In *P. radiata*, a single study has implicated sugars. Page *et al.* (2013) attributed bark stripping to higher sugars in the bark compared to surrounding food sources. However, the interaction of sugars with secondary compounds to explain the observed variation in mammalian bark stripping has not yet been investigated. Here we examined the induced chemical responses to bark stripping in *P. radiata* and compared these responses to the comparatively well-documented responses to methyl jasmonate. We asked: 1) Are there qualitative and quantitative differences in constitutive primary and secondary chemistry between the needles, bark and roots of *Pinus radiata*?, 2) What chemical changes occur in the different plant parts following artificial bark stripping and above ground application of methyl jasmonate?, and 3) do the amounts of compounds correlate between and within plant parts?

3.2   Materials and Methods

3.2.1   Experimental design

In 2015, 6-month-old seedlings of 18 full-sib families (each with at least 4 seedlings) of *Pinus radiata* (D. Don) originating from the Radiata Pine Breeding Company deployment population were obtained from a commercial nursery. Seedlings were transferred into 145 mm x 220 mm pots containing 4L of basic potting mix (composted pine bark 80% by volume, coarse sand 20%, lime 3 kg/m3 and dolomite 3 kg/m3) and raised outside in a fenced area (to protect against animals) at the University of Tasmania, Hobart. At 2 years of age, the 18 families were randomly allocated to three treatment groups (methyl jasmonate-MJ, artificial bark stripping-strip and control) each with 6 families (n=6 seedlings per treatment). Within each family and treatment, the individuals from were again randomly allocated to different sampling times (T0-T4). T0 represents the time immediately before treatment applications. T1, T2, T3 and T4 represents respective sampling times at 7, 14, 21 and 28 days after treatment (MJ and strip) application. While all T0 seedlings, irrespective of group allocation were not treated, all the seedlings allocated to control were not treated all through the experimental period. MJ

was applied to seedlings of 6 families (n=6), strip treatments were applied to another 6 families (n=6), and the remaining 6 families were kept as controls (Figure 1). 25mM MJ (Moreira *et al.* 2013a) was applied by spraying the stem and needles with a fine mist from a hand sprayer until 'just before run-off'. The treated seedlings were sprayed in a well-ventilated area away from untreated seedlings to avoid cross contamination (Moreira *et al.* 2013a). For strip treatment, plants were artificially stripped by removing a 30 cm vertical strip, beginning 2 cm from the ground and covering 50% of the stem circumference, representative of the common upper threshold for browsing observed in natural field conditions. The three treatment groups (control, strip and MJ) were arranged in a randomized block design of 3 blocks in a shade house, each block comprising a treatment plot of two families (Figure 3.1). Seedlings of each family were linearly arranged in a family plot. The treatment plots were separated within each block to minimise any interference which may arise from the MJ and bark strip treatments.

a

Treatments	control	MJ	strip	Sampled before treatment for constitutive chemistry
Time T0	n=6	n=6	n=6	
T1	n=6	n=6	n=6	Sampled 7 days after treatment application
T2	n=6	n=6	n=6	Sampled 14 days after treatment application
T3	n=6	n=6	n=6	Sampled 21 days after treatment application
T4	n=4	n=4	n=5	Sampled 28 days after treatment application

b

Block 1

T0	T1	T2	T3	T4
4	1	3	2	5
4	5	1	2	3

| 2 | 4 | 1 | 3 | |
| 5 | 4 | 3 | 1 | 2 |

| 2 | 3 | 1 | 5 | 4 |
| 1 | 3 | 5 | 4 | 2 |

Block 2

T0	T1	T2	T3	T4
3	2	4	1	
4	1	3	2	5

| 2 | 5 | 1 | 4 | 3 |
| 4 | 5 | 2 | 1 | 3 |

| 2 | 5 | 1 | 3 | 4 |
| 4 | 3 | 1 | 5 | 2 |

Block 3

T0	T1	T2	T3	T4
2	4	3	5	1
4	1	5	2	3

| 2 | 1 | 3 | 4 | 5 |
| 2 | 4 | 3 | 1 | |

| 1 | 2 | 3 | 4 | |
| 2 | 4 | 3 | 1 | |

Figure 3.1: Experimental lay out in the shade house. a) At T0, seedlings were destructively harvested just before treatment applications. At 7(T1), 14(T2), 21(T3) and 28(T4) days after treatment, one seedling from each family (total number of seedlings per sampling time =18, equivalent to the number of families, and n= 6 are the seedlings per treatment) were destructively harvested. **b**) Shows the experimental design of the treatments. Each set of squares is a table where we placed seedlings and represents a replicate block. Each row represents a different family. Each treatment plot comprised two family plots in which seedlings of the same family were arranged linearly in a row and a seedling sampled at random for studying temporal changes in chemistry. Each square represents a seedling and the number represents the order in which they were

sampled. Squares of the same colours received the same treatment i.e. yellow = MJ treated, blue = strip and green = control. The blank squares represent missing individuals.

3.2.2 *Sample processing and phytochemical extractions*

Each family had a minimum of 4 seedlings from which one seedling was randomly sampled weekly from T0-T4 (Figure 3.1). T0 represents the time immediately before treatment applications. T1, T2, T3 and T4 represents respective sampling times at 7, 14, 21 and 28 days after treatment application. For constitutive estimates, one seedling from each family (n=18; 3 treatments x 6 families) was destructively harvested just before application of treatments (T0). Thereafter, one seedling from each family (n=18 per time; 3 treatments x 6 families) were destructively harvested at T1, T2, T3 and T4 to investigate induced changes in chemistry (Figure 3.1). At harvesting the seedling was divided into 3 parts; the upper part with most needles for needle collection, the middle stem for bark collection and the stump in the soil for root harvesting. Because of the small size of the trees (average height = 130 cm), most of the needles, bark and roots on the plant were collected. Needles were cut off the upper stem using scissors. From the middle stem, the bark was carefully peeled off avoiding the wood using a knife. The knife and scissors were cleaned with ethanol after every harvest to avoid cross contamination. For the roots, each stump was carefully removed from the pot and soil gently separated from the roots. The roots were further shaken and rinsed in water to remove any remaining soil and both fine and coarse roots were cut from the stump using garden clippers that were cleaned between samples. Needles, bark and roots from each seedling were kept separate and immediately put in a cooler and transported to a −20°C freezer for storage until chemical extraction.

Chemical extractions (targeting terpenes, phenolics and sugars) were undertaken randomly and separately from each of the three plant parts (needles, bark and roots) from each seedling sampled. Dichloromethane (DCM) and acetone were used to extract polar and semipolar components respectively (Jones *et al.* 2002; Sasidharan *et al.* 2011). For DCM-extracted compounds, frozen material was cut into smaller sizes and 1.5 g of tissue was weighed and extracted in 10 ml of 99.9% dichloromethane (Jones *et al.* 2002). An internal standard (n-heptadecane) was added to dichloromethane at the concentration of 100 ppm. Acetone extracts were made from 50 g of freeze dried, ground material in 10 ml of 95% acetone, and sugars were extracted from 50 g of freeze dried, ground material in 10 ml of hot water (Jones *et al.* 2002). Rutin was used as the internal standard for the acetone extracts and final concentration of 0.2mg/L was added to each sample. Acetone extracts were then reconstituted by mixing 600 µl of the extract with 600 µl of 98:2 acetonitrile/water and then 1ml of each extract was transferred into a vial and stored at −20°C until analysis. An extra sample of

each tissue per individual was weighed, dried in the oven at 110°C for 72 hrs and reweighed to convert samples from wet to dry weights. The DCM extracts that comprised the mono-, sesqui- and diterpenoids and volatile phenolic compounds were then analysed by gas chromatography-mass spectrometry (GC-MS). However, the diterpenoid resin acids in the GC-MS were quantified by the ultra-high-performance liquid chromatography-mass (UHPLC-MS). The acetone extracts that comprised the diterpenoid resin acids and fatty acids were analysed in the defrosted state by the UHPLC-MS. The sugars were also quantified by the UHPLC-MS.

3.2.3   GC-MS analyses

Gas chromatography-mass spectrometry (GC-MS) analyses were carried out on a Varian 3800 GC coupled to a Bruker-300 triple quadrupole mass spectrometer, using helium carrier gas at 1.2 mL/min in constant flow mode. The column was an Agilent DB-5 column (30 m x 0.25 mm internal diameter and 0.25 µm film). Injections of 1 µL were made using a Varian CP-8400 autosampler and a Varian 1177 split/splitless injector in split mode with a 4:1 split ratio. The injector temperature was 250ºC. The column oven was started at 60ºC then ramped to 290ºC at 15ºC/min with a 1-minute hold at the final temperature. The ion source was held at 220ºC, and the transfer line at 290ºC. Electron ionisation mass spectra at 70eV were acquired over the range m/z 35 to 400 over 130 ms, with additional Selected Ion Monitoring (SIM) channels in 4 different time windows, all with 15 ms dwell time per channel. Window 1 from 0 to 8 minutes included m/z 41, 68, 69, 71, 93, 104, 135, 148, 151, 178, window 2 from 8 to 9.7 minutes included m/z 81, 107, 121, 162, window 3 from 9.7 to 14.5 minutes included m/z 69, 91, 109, 137, 177, 229.1, 239.2, 240.2 and window 4 from 14.5 minutes to the end included m/z 81, 109, 134 and 221.1. These ions were chosen based on initial full scan analyses to select compounds to target. Peak areas of relevant characteristic ions were measured using Bruker Workstation MS Data Review version 7.0, then scaled up to the equivalent total ion current based on measurements of the proportion the diagnostic ion was of the full spectrum (based on a good quality spectrum). All adjusted peak areas were finally expressed as ratios of the total ion current for each compound to the internal standard, n-heptadecane. The preliminary identification of compounds was based on the comparison of the retention time and mass spectra with the National Institute of Standards and Technology mass spectra library (NIST 2014). Most compounds gave ions that were structurally characteristic. The DCM components were expressed as milligrams of heptadecane equivalents (HE) per gram of dry weight of the sample (mg HE/g dw).

3.2.4   LC-MS analyses

To analyse the acetone extracts, 12 µL aliquots were injected using a Waters Acquity H-series UHPLC coupled to a Waters Xevo triple quadrupole mass spectrometer operating MassLynx 4.1 software. A Waters Acquity UHPLC BEH C18 column (2.1 x 100 mm x 1.7 micron particles) was used, with 1% acetic acid (Solvent A) and acetonitrile (Solvent B) at a flow rate of 0.35 mL/min and, after an initial hold for 30 seconds at 85% A:15% B, a linear gradient was followed to 45% A:55% B at ten minutes, then a further linear gradient to 5% A:95% B at 15 minutes with a one minute hold at the final value, before re-equilibration to starting conditions for 4 minutes. The mass spectrometer was operated in negative ion electrospray mode. The ion source temperature was 150°C, the desolvation gas was nitrogen at 950 L/hr, and the desolvation temperature was 450°C and needle voltage 2.7 kV. Based on some trial full scan analyses, a range of target ions were included in Selected Ion Monitoring (SIM) mode with dwell time of 27 ms per channel, as well as a full scan from m/z 120 to 1200 over 250 ms. The SIM channels chosen (with cone voltages in brackets) were 277 (20), 285(25), 289 (25), 297 (C19 fatty acid standard)(25), 303(25), 317(30), 319(30), 333 (30), 349 (30), 365 (30), 405(30), 465(35), 481 (35), 495 (40), 561 (30), 575 (30), 577 (30), 579 (30), 609.1 (rutin standard) (40), 709.2 (40), 739.2 (40), 863.2 (45). Data were analysed with MassLynx and TargetLynx software. The LC-MS analytes were expressed as milligrams of rutin equivalents (RE) per gram of dry weight of the sample (mg RE/g dw).

To further characterise the diterpenoid resin acids, selected dichloromethane and acetone extracts were evaporated to dryness. The residue was redissolved in methanol/ chloroform/ HCl (10:1:1) and heated at 80°C for 1 hr. The methylated diterpenoid resin acids were then extracted with hexane/chloroform (4:1) and the resulting extract analysed by GC-MS. However, to verify the retention times for final identification of the resin acids by UHPLC-MS, standards namely; abietic acid, neoabietic acid, dehydroabietic acid, palustric acid, levopimaric acid, pimaric acid and isopimaric acid were purchased from Santa Cruz Biotechnology and analysed by UHPLC-MS.

For sugars, 1 µL aliquots were injected using the instrument described above for the acetone extracts. A Waters Acquity UHPLC BEH Amide column (2.1 x 50 mm x 1.7-micron particles) was used, with 0.4 % ammonia in water (Solvent A) and acetonitrile (Solvent B) at a flow rate of 0.37mL/min, with a gradient from 20% A:80% B to 28% A:72% B at 2 minutes, before immediate re-equilibration to initial conditions for 2.5 minutes. The mass spectrometer was operated in negative ion electrospray mode, and selected ion monitoring was used to detect the specific sugars, which were quantified by external calibration. The ion source temperature was 150°C, the desolvation gas was nitrogen at 1000 L/hr, and the desolvation temperature was 450°C and needle voltage 2.7 kV. The [M-H]- ion was monitored

for monosaccharides and inositol at m/z 179.1 and for sucrose and other disaccharides at m/z 341.1. Cone voltages were 17 V for monosaccharides and 24 V for disaccharides. Standard curves were created over the range 0 to 500 ppm with standards at 0, 10, 25, 50, 100, 250 and 500 ppm. Non-linear (second-order) equations were used for sucrose (typical $R^2= 0.99996$) while linear fits were used for the monosaccharides (typical $R^2 = 0.9992$). Due to potential drift in MS response, the full standard curve was repeated after every tenth sample, and a 250 ppm QC was run in the middle of each set of ten. Under these conditions, fructose eluted at 1.11 minutes, glucose at 1.38 minutes, sucrose at 2.09 minutes and inositol at 2.49 minutes. In a typical large set of analyses, with 18 250 ppm QC samples run, the average values reported for fructose, glucose and sucrose were 249 ppm, 248 ppm and 250 ppm respectively, with relative standard deviations of 2.0%, 1.5% and 1.8% respectively. Accurate mass data for molecular formula assignment of major unknowns were acquired by direct infusion on a Thermo Orbitrap mass spectrometer operating at a resolution of 15,000.

Compounds that were quantified by the GC-MS were summed according to functional groups and expressed in their relevant units. Therefore, in addition to individual compounds, analysis was made of total compound groups i.e. total monoterpenoids, sesquiterpenoids, GC-MS diterpenoids and phenolics. In addition, because LC-MS sugars were quantified in absolute amounts, total sugars were also derived as the sum of individual sugars. The rest of the compounds analysed by the LC-MS i.e. the fatty acids and the LC-MS diterpenoids were not summed, except for one group of resin acids ($C_{20}H_{30}O_2$ resin acids [41]) that eluted together that was considered in the total compound groups. All amounts were expressed relative to the internal standard except the sugars that were measured in absolute amounts. The terpenes and phenolics are broadly categorised under secondary metabolites while the sugars as well as fatty acids are primary compounds. All individual compounds that were identified have been given a unique identifier based on Supplementary Table 10 for ease of location in the tables.

3.2.5   Statistical analyses

All statistical analyses were conducted using the software R (version 3.6.0) (R Core Team 2018). Principal components analysis (PCA), using FactoMinerR version 1.41 (Lê *et al.* 2008) was used to summarise the overall difference between samples from the three plant parts, to understand the constitutive chemistry (T0 samples) contributing to the variation and visualize differences between plant parts. The PCA was based on the correlation matrix among all identified chemicals compounds. Differences between plant parts along each principal component were tested with the Kruskal–Wallis (KW) one-way analysis of variance since the normality assumption did not hold. Comparisons

between plant parts were undertaken for the first 13 principal components (PCs) and significance was set at p-value of 0.05/39 (13 PCs X 3 parts) = 0.001 following Bonferroni adjustment.

To further characterize the differences in the plant parts, arithmetic means and standard errors (se) were calculated for all compounds that were detected. For compounds detected in more than one plant part, differences among plant parts were tested with Kruskal–Wallis (KW) one-way analysis of variance and where significant results were obtained a Dunn's test (Dn) was performed to identify the parts which were significantly different. Bonferroni adjustments (Bf) were made to account for multiple tests for individual compounds. Significant p-values were considered at 0.05/n, where n is the number of statistical tests (McDonald 2009), for example a p-value of 0.05/48 = 0.001 was considered significant for tests between plant parts for individual monoterpenoids, where 48 = 16 monoterpenoids X 3 plant parts.

To further consider relationships between individual constitutive compounds between plant parts, Spearman's rank correlations were used. The tests were performed on all the compounds that were detected in more than one plant part and unadjusted p-values were reported. To have an idea of correlations within plant parts, Spearman's rank correlations were tested for only the compound groups (total mono-, sesqui- and GC-MS diterpenoids, sugars as well as the LC-MS resin acid group)

To detect induced changes, the mean amounts of compounds after treatment (induced; T1-T4) were compared. Induced changes for each time were detected by comparing the MJ and strip treatments to the control for that time (Figure 3.1) (Morris *et al.* 2006). Spearman's rank correlations were also tested for induced chemistry to detect whether there are any shifts in the correlations following treatment. Between plant parts, the correlations between plant parts for all compounds that were detected in more than one plant part were assessed from T1 – T4 (but for only methyl jasmonate induced chemistry for simplicity of presenting the results). Within plant parts, Spearman's rank correlations were tested between total compound groups at T1 (7 days after treatment) only. Absolute p-values are reported for the correlations.

3.3   Results

Across all samples, a total of 81 compounds or compound groups were detected. Full or partial identification was achieved for 52 of these, which included 15 monoterpenoids, 20 diterpenoids, 3 sesquiterpenoids, 7 phenolic compounds, 4 sugars and 3 fatty acids (Table 3.1). The 28 unidentified compounds/groups had molecular weights ranging between 104 – 770, and where group allocation

was possible these included terpenes and sugars (Table 3.1). The major $C_{20}H_{30}O_2$ [41] diterpenoid resin acids (Table 3.1) that had very close retention times by LC - MS were measured as a group. All the 81 compounds were given a unique number for ease of identification in the tables.

3.3.1   Constitutive differences between plant parts
Overall differences

Of the 81 compounds/groups, 62 compounds were detected in the bark, 45 in the needles and 35 in the roots (Table 3.1). While 20 of these compounds were detected in all the three plant parts, 22, 13 and 5 compounds were detected only in the bark, needles and roots respectively (Supplementary Figure 3.1; Table 3.1). The constitutive (T0) data for all plant parts, was reduced to 13 principal components that had eigen values greater than one and these explained 90.3% of the total variation. PC1, PC2 and PC3 explained 37.2%, 20.2% and 7.3% of the chemical variation, respectively. The three plant parts were clearly differentiated in the two-dimensional space defined by the first (PC1) and second (PC2) principal components (Figure 3.2).
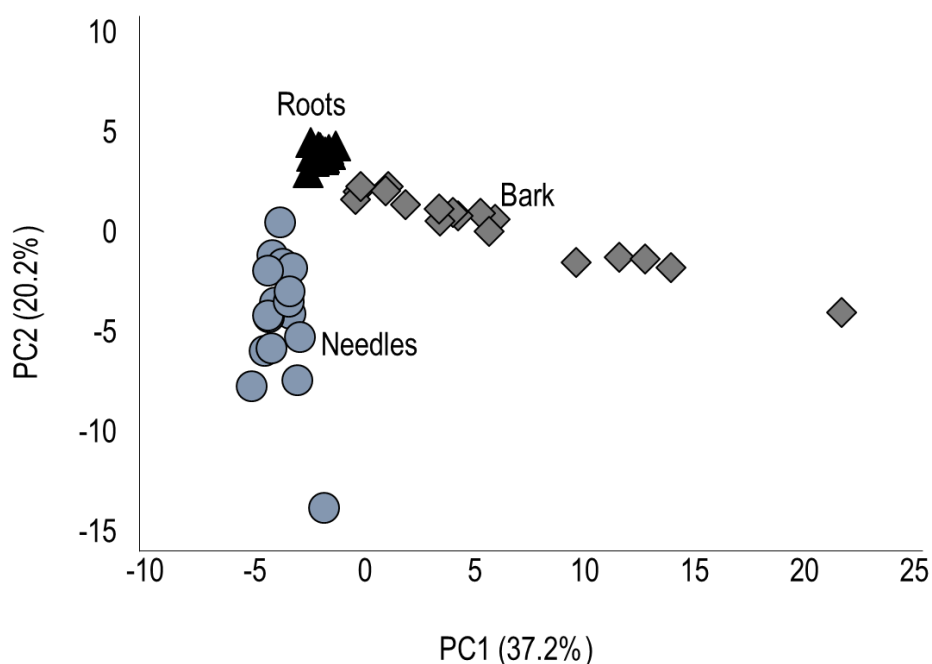
Figure 3.2: PC1 vs PC2 separating the needle, bark and root samples of *Pinus radiata* seedlings based on their constitutive chemistry. The principal component analysis was based on all compounds detected at T0. The three plant parts were clearly differentiated in the two-dimensional space defined by the first (PC1) and second (PC2) principal components. PC1 and PC2 respectively explained 37.2% and 20.2% of the chemical variation. PC1 strongly differentiated the needles and bark, with bark samples having high values relative to the needles ($p < 0.001$; *Dn, Bf*) and the roots ($p < 0.01$; *Dn, Bf*). PC1 did not differentiate between needles and roots. Increasing values along PC1 were mainly associated with unknown diterpene Mol Wt 272 [38], unknown m/z 109 B [34], vanillin [74], unknown m/z 109 A[33] and citronellic acid [7] (top 5 compounds). The root samples were strongly differentiated from the needle samples along PC2 ($p < 0.001$; *Dn, Bf*), with the bark samples intermediate and marginally discrete from the roots ($p < 0.05$; *Dn, Bf*) but not from the needles. PC2 values were higher in the roots than the bark ($p < 0.05$; *Dn, Bf*). Increasing values in the needles along PC2 were mainly associated with palmitic acid[85],, unknown Mol Wt 770 [98], unknown Mol Wt 740 B [97], unknown C20H30O6 A [52], unknown C20H30O2 A [47] (top 5 compounds). The third principal component explained 7.3% of the variance further separated the roots from the bark ($p < 0.05$; *Dn, Bf*) and needles ($p < 0.05$; *Dn, Bf*) but not needles from bark. Roots had higher values on PC3 that was mainly associated with glucose [77], copalol [24], agathadiol [22], unknown m/z 109 A [33] and β-pinene [4] (top 5 compounds). The other principal components did not provide further resolution between the three plant parts. The numbers in parentheses are the identifiers to enable quick location of the compounds in the tables.

Differences between plant parts for total compound groups

Plant parts differed in the amounts of total compound groups. These differences are summarised in Figure 3.3. In summary, while total monoterpenoids were lower in the roots, they were not significantly different among plant parts after Bonferroni (Bf) adjustment. Total GC-MS diterpenoids were lower in the roots than the bark (p <0.01; *Dn, Bf*) and needles (p <0.001; *Dn, Bf*) but were not significantly different between needles and bark. Sesquiterpenoids were not detected in the roots and were not significantly different between the bark and needles. The amount of total phenolics were almost absent in the roots hence were lower than in the bark (p <0.001; *Dn, Bf*) and needles (p <0.01; *Dn, Bf*). Needles and bark also differed in total phenolics (p <0.05; *Dn, Bf*). Total sugars were lower in the roots than in the needles or bark (p <0.001; *Dn, Bf*) with no difference between needles and bark. The LC-MS $C_{20}0H_{30}O_2$ resin acids [41] that eluted as a group were higher in the bark (p <0.001; *Dn, Bf*) than the needles, with the roots intermediate and not significantly different from the bark or the needles. Of the secondary compounds identified by the GC-MS, the total monoterpenoids were higher than diterpenoids, sesquiterpenoids and phenolics in all the three plant parts. The total sesquiterpenoids were the smallest fraction of the total identified GC-MS terpenoids.

Differences between plant parts for dominant individual compounds

Within total compound groups, there were some individual compounds that were dominant, and these varied with plant part (Table 3.1, Figure 3.4). The monoterpenoids were dominated by β/α -pinene [1,4] followed by β-phellandrene [3] in all the three plant parts (Figure 3.4). The GC-MS diterpenoids were dominated by agatholal [23] and agathadiol [22] both in the bark and needles and by methyl levopimarate [27] in the roots. The sesquiterpenoids in the needles were dominated by the unknown sesquiterpenoid alcohol [21] and those in the bark dominated by trans-farnesol [20]. No sesquiterpenoids were detected in the roots. The phenolics were dominated by trans-coniferyl alcohol [73] in the needles and the bark. Ethyl 4-ethoxybenzoate [62] was the only phenolic compound detected in the roots. Glucose [77] dominated the sugars in the needles and bark, but fructose [76] dominated in the roots (Figure 3.4). However, sucrose was not detected in the needles. Of the fatty acids, palmitic acid [85] dominated in the needles and linoleic acid [83] in the needles and the roots. Overall, based on the dominant compounds the needles and the bark had a more similar chemical profile relative to the roots.
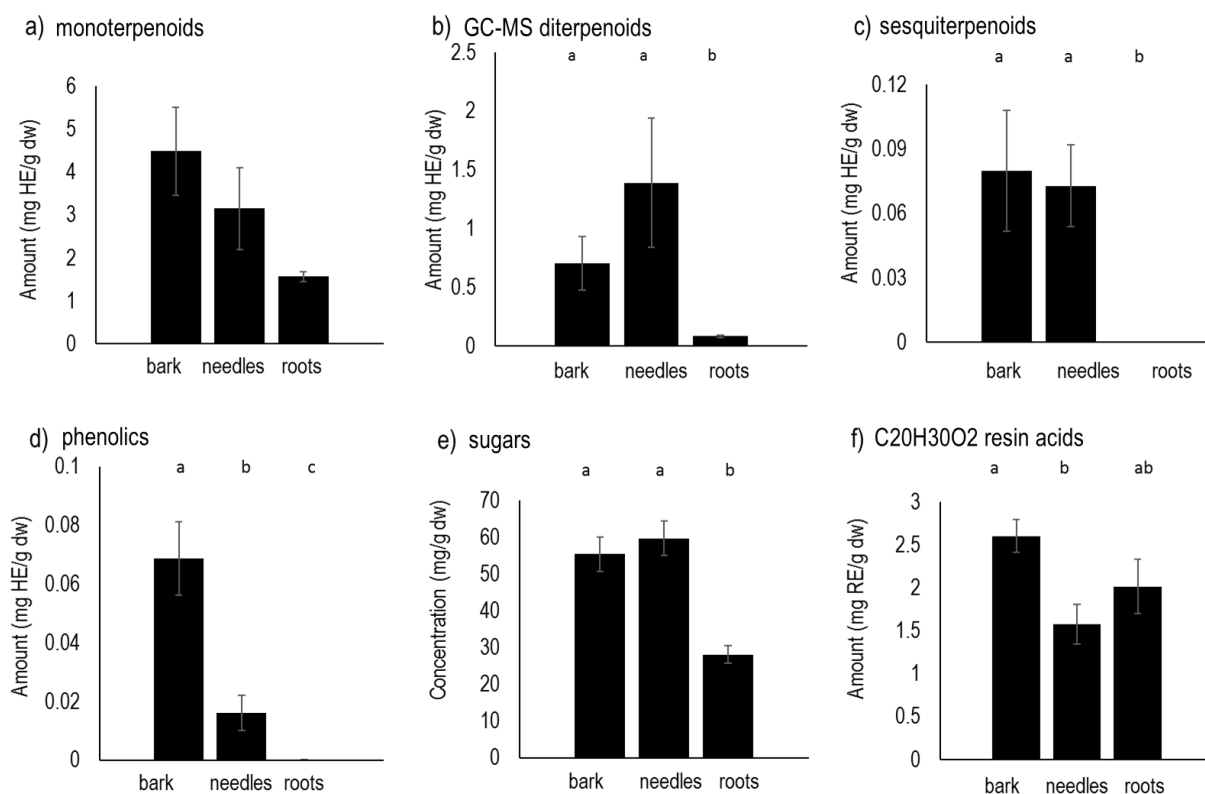
Figure 3.3: Relative amounts (mg/g±SE) of total: a) monoterpenoids, b) GC-MC diterpenoids, c) sesquiterpenoids, d) phenolics, e) sugars and f) C20H30O2 resin acids in the bark, needles and roots of *Pinus radiata* seedlings at T0. No sesquiterpenoids were found in the roots. Different letters adjacent to the plot indicate significant differences (p < 0.05) between plant parts based on Dunn's multiple comparisons test for post hoc evaluations with Bonferroni adjustments. The resin acid group was not derived by summation of individual resin acids quantified by the LC-MS. This group represents a group of LC-MS resin acids that eluted together. The terpenes and phenolics are expressed as mg heptadecane equivalents (HE)/g dw and resin acids are expressed as mg rutin equivalents (RE)/g dw, as no absolute quantitation was carried out on these analytes. The absolute amounts of sugars are quantified in mg/g dw.
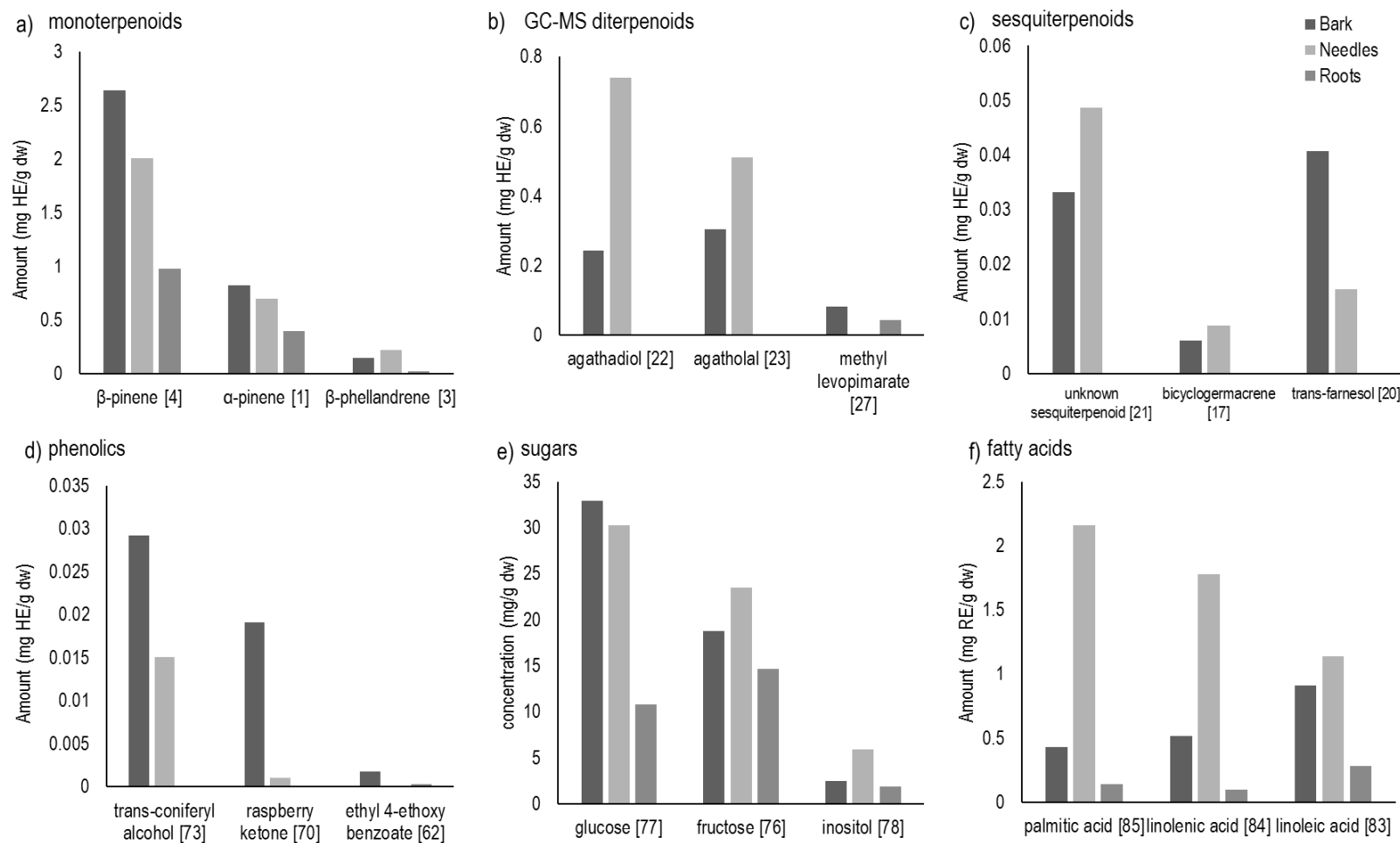
Figure 3.4: Mean constitutive amounts of the dominant individual: a) monoterpenoids, b) diterpenoids, c) sesquiterpenoids, d) phenolics, e) sugars and f) fatty acids at T0, showing differences in the dominant compounds of the bark, needles and roots of *Pinus radiata*. The terpenes and phenolics are expressed as mg heptadecane equivalents (HE)/g dw and fatty acids are expressed as mg rutin equivalents (RE)/g dw, as no absolute quantitation was carried out on these analytes. The absolute amounts of sugars are quantified in mg/g dw.

Table 3.1: Mean and standard error (se) of the relative constitutive (T0) amounts of all the identified secondary and primary metabolites in *Pinus radiata* and the changes that occurred after treatment in the needles [N], bark [B] and roots [R]) at different times. The GC-MS components (monoterpenoids, sesquiterpenoids, GC-MS diterpenoids and phenolics) are expressed as milligrams of heptadecane equivalents (HE) per gram of dry weight of the sample (mg HE/g dw) and the LC-MS analytes (LC-MS diterpenoids and fatty acids) are expressed as milligrams of rutin equivalents (RE) per gram of dry weight of the sample (mg RE/g dw). Sugars are expressed in mg/g dw. Compounds labelled "A", "B", "C" and "D" are isomers. The letters adjacent to the means relate to differences in the means, where different letters indicate that the means are significantly different, otherwise, they are not. Significance was set at $p < 0.05$ based on Kruskal Wallis with Dunn test with Bonferroni correction. Changes after methyl jasmonate [MJ] or strip [S] treatments are also indicated where, ↑ and ↓ signify that the amount of compound increased or reduced respectively after treatment in the bark[B], needles[N] or roots[R]. For the induced changes, the unadjusted p-values based on Kruskal Wallis tests are indicated. Only induced changes associated with a $p < 0.05$ are indicated but the bold values retained their significance after Bonferroni correction. T0 represents the time before treatment applications. T1, T2, T3 and T4 represents 1, 2, 3 and 4 weeks after treatment application, respectively. All compounds were given a unique identifier based on Supplementary Table 10 (after Chapter 9), for ease of identification. The identifiers in this table are not sequential as some compounds indicated in Supplementary Table 10 were not identified in this data set

	Molecules	Mean ± se (T0)			T1	T2	T3	T4
		Bark (n=18)	Needles(n=18)	Roots (n=18)				
	total monoterpenoids	4.488±1.058ᵃ	3.152±0.743ᵃ	1.564±0.369ᵃ		↑B MJ (0.02)	↑B MJ (0.01)	
	total GC-MS diterpenoids	0.705±1.662ᵃ	1.391±0.328ᵃ	0.085±0.020ᵇ	↑R MJ (0.04)	↑B MJ (0.04)	↓R S (0.008)	
	total sesquiterpenoids	0.080±0.019ᵃ	0.073±0.017ᵃ					
	total phenolics	0.069±0.016ᵃ	0.016±0.004ᵇ	0.000±0.000ᶜ		↓N MJ (0.03); ↓N S (0.03)		
	total sugars	55.350±4.689ᵃ	59.668±4.657ᵃ	28.083±2.408ᵇ	↓**B MJ (0.0008)**, S (0.03); ↓**N MJ (0.0007)**	↓**B MJ (0.0004)**; ↓B S (0.02); ↓**N MJ (0.002)**	↓B MJ (0.006); ↓N (0.03)	↓B MJ (0.008)
	Monoterpenoids							
1	α-pinene	0.824±0.204ᵃ	0.699±0.171ᵃ	0.394±0.038ᵃ			↑B MJ (0.003); ↑B S (0.02); ↓R S (0.04)	↓R S (0.01)
2	α -terpineol	0.001±0.000			↑B MJ (0.01)			
3	β-phellandrene	0.147±0.049ᵃ	0.218±0.091ᵃ	0.023±0.006ᵃ		↑B MJ (0.01)		
4	β-pinene	2.640±0.594ᵃ	2.004±0.676ᵃ	0.974±0.078ᵃ		↑B MJ (0.01)	↑B MJ (0.006)	↑N MJ (0.03)
5	camphene	0.014±0.003ᵃ	0.011±0.004ᵃ			↑B MJ (0.02)	↑B MJ (0.002); ↑B S (0.008)	
6	citronellal	0.072±0.021						
7	citronellic acid	0.023±0.006						
8	citronellol	0.144±0.039						
9	γ-terpinene	0.009±0.003						
10	limonene	0.124±0.037ᵃ	0.213±0.068ᵃ	0.011±0.001ᵇ				
11	linalool	0.003±0.001					↑B MJ (0.003)	↑B MJ (0.04)
12	myrtenoic acid			0.002±0.000				
13	sabinene	0.346±0.117ᵃ	0.006±0.002ᵇ	0.105±0.018ᵃ			↑N S (0.02)	
14	terpinene-4-ol	0.011±0.003						
15	terpinolene	0.130±0.047ᵃ		0.055±0.008ᵃ				↓R MJ (0.04)
16	unknown Mol Wt 150	0.011±0.004ᵃ	0.004±0.003ᵇ					

	Sesquiterpenoids							
17	bicyclogermacrene	0.006±0.001[a]	0.009±0.002[b]					
20	trans-farnesol	0.041±0.021[a]	0.015±0.003[a]					
21	unknown sesquiterpenoid alcohol	0.033±0.015[a]	0.049±0.014[a]					
	GC-MS diterpenoids							
22	agathadiol	0.243±0.057[a]	0.739±0.120[a]					
23	agatholal	0.305±0.072[a]	0.510±0.145[a]					↑N$_{MJ (0.01)}$
24	copalol		0.136±0.066					
25	levopimaral	0.023±0.006[a]		0.024±0.003[a]		↑B$_{MJ (0.02)}$; ↑B$_{S (0.03)}$	↓R$_{S (0.02)}$	
26	methyl dehydroabietate	0.023±0.005[a]	0.006±0.001[a]	0.018±0.004[a]			↑B$_{MJ (0.03)}$; ↓R$_{S (0.01)}$	
27	methyl levopimarate	0.082±0.019[a]		0.043±0.010[a]		↑B$_{MJ (0.006)}$ ↑B$_{S (0.01)}$	↑B$_{MJ (0.04)}$; ↑B$_{MJ (0.03)}$; ↓R$_{S (0.02), MJ (0.02)}$	
28	unknown $C_{19}H_{26}$	0.029±0.007					↑B$_{MJ (0.02)}$	
	LC-MS diterpenoids							
29	dehydroabietic acid	4.050±0.287[a]	0.897±0.116[b]	1.727±0.093[b]				↑B$_{MJ (0.04)}$
30	unknown diterpene-1	0.213±0.203						
31	unknown diterpene-2			0.154±0.037			↓R$_{s (0.01)}$	
32	unknown diterpene-3	1.218±0.415[a]	0.000±0.000[b]	0.047±0.011[a]				
33	unknown m/z 109 A	0.056±0.014[a]		0.017±0.001[a]		↑B$_{MJ (0.006)}$; ↑B$_{S (0.01)}$	↑B$_{MJ (0.03)}$; ↑B$_{S (0.03)}$; ↓R$_{S (0.01)}$	
34	unknown m/z 109 B	0.120±0.023[a]	0.016±0.004[b]				↓R$_{s (0.01)}$	
35	unknown m/z 121			0.030±0.011			↓R$_{MJ (0.01)}$	
36	unknown m/z 134	0.486±0.086				↑B$_{MJ (0.005)}$		
37	unknown m/z 239		0.027±0.007[a]	0.001±0.000[a]	↓R$_{S (0.03)}$			
38	unknown Mol Wt 273	0.020±0.005					↑B$_{MJ (0.02)}$	
39	unknown $C_{20}H_{30}O_2$ A			0.602±0.052				
40	unknown $C_{20}H_{30}O_2$ B		0.760±0.106					
41	$C_{20}H_{30}O_2$ resin acids	2.603±0.613[a]	1.576±0.372[b]	2.016±0.475[ab]				↑B$_{S (0.03)}$
42	unknown Mol Wt 304 A		0.003±0.001					↑N$_{MJ (0.03)}$
43	unknown Mol Wt 304 B		0.106±0.016		↓N$_{MJ (0.008)}$	↑N$_{MJ (0.04)}$		
44	unknown Mol Wt 304 C	0.204±0.028						
45	unknown Mol Wt 316	3.494±0.337			↑B$_{S (0.006)}$		↑B$_{MJ (0.009)}$; ↑B$_{S (0.01)}$	↑B$_{S (0.02)}$; ↑B$_{MJ (0.03)}$
46	unknown $C_{20}H_{30}O_3$	3.512±0.365[a]	0.565±0.083[b]	1.829±0.106[a]	↑B$_{S (0.01)}$	↓R$_{S (0.04)}$	↑B$_{MJ (0.009)}$; ↑B$_{S (0.01)}$; ↓R$_{S (0.03)}$	↑B$_{MJ (0.02)}$; ↑B$_{S (0.02)}$
47	unknown $C_{20}H_{32}O_3$ A	1.587±0.203[a]	1.389±0.243[a]	1.172±0.094[a]	↑B$_{S (0.04)}$			↑B$_{S (0.01)}$; ↑N$_{MJ (0.04)}$
48	unknown $C_{20}H_{32}O_3$ B		0.087±0.028				↑N$_{S (0.04)}$	↑N$_{S (0.04)}$
49	unknown $C_{20}H_{32}O_3$ C		1.042±0.169					↑B$_{S (0.01)}$; ↑B$_{MJ (0.02)}$

No.	Compound							
50	unknown C20H30O4	10.994±1.002^a	1.201±0.165^b	1.858±0.100^b	↑B $_{S\ (0.03)}$	↓R $_{MJ\ (0.01)}$; ↓R $_{S\ (0.01)}$;	↑B $_{S\ (0.02)}$; ↑N $_{S\ (0.04)}$; ↓R $_{S\ (0.03)}$	↑B $_{MJ\ (0.02)}$; ↑B $_{S\ (0.02)}$;
51	unknown C20H30O5	2.768±0.287^a		0.076±0.014^b			↑B $_{MJ\ (0.008)}$; ↓R $_{MJ\ (0.04)}$	
52	unknown C20H30O6 A	0.144±0.035					↑B $_{MJ\ (0.02)}$	↑B $_{S\ (0.005)}$
53	unknown C20H30O6 B		0.191±0.051					↑N $_{MJ\ (0.04)}$
54	unknown C20H30O6 C	1.37±0.071			↑B $_{S\ (0.02)}$		↑B $_{MJ\ (0.004)}$, ↑B $_{S\ (0.01)}$	↑B $_{MJ\ (0.01)}$
	phenolics							
56	anethole	0.001±0.000			↑B $_{MJ\ (0.004)}$	↑B $_{MJ\ (0.01)}$	↑B $_{MJ\ (0.003)}$	
62	ethyl 4-ethoxybenzoate	0.002±0.000^a		0.0002±0.000^a		↓N $_{MJ\ (0.03)}$; ↓N $_{S\ (0.03)}$		
65	methyl eugenol	0.001±0.000					↑B $_{MJ\ (0.03)}$	
68	pinosylvin dimethyl ether	0.006±0.003						
70	raspberry ketone	0.019±0.004^a	0.001±0.001^b					
73	trans coniferyl alcohol	0.029±0.006^a	0.015±0.006^a		↓B $_{MJ\ (0.03)}$			
74	vanillin	0.010±0.002			↑B $_{MJ\ (0.04)}$		↑B $_{MJ\ (0.006)}$	
	Sugars							
76	fructose	18.764±1.821^{ab}	23.522±2.054^a	14.658±1.319^b	↓**B** $_{MJ\ (0.001)}$; ↓N $_{MJ\ (0.001)}$	↓**B** $_{MJ\ (0.0004)}$; ↓B $_{S\ (0.02)}$; ↓N $_{MJ\ (0.003)}$	↓B $_{MJ\ (0.006)}$; ↓N $_{MJ\ (0.04)}$	↓R $_{S\ (0.04)}$
77	glucose	32.888±2.887^a	30.250±2.911^a	10.802±0.933^b	↓**B** $_{MJ\ (0.0003)}$, $_{S\ (0.02)}$; ↓**N** $_{MJ\ (0.0005)}$	↓**B** $_{MJ\ (0.0004)}$; ↓B $_{S\ (0.02)}$; ↓**N** $_{MJ\ (0.002)}$	↓**B** $_{MJ\ (0.002)}$	↓B $_{MJ\ (0.005)}$; ↓N $_{MJ\ (0.04)}$
78	inositol	2.466±0.470^a	5.896±0.530^b	1.830±0.275^a	↓N $_{MJ\ (0.02)}$	↓**N** $_{MJ\ (0.002)}$	↓B $_{S\ (0.03)}$; ↓**N** $_{MJ\ (0.002)}$	
79	sucrose	1.230±0.327^a		0.793±0.147^a	↓B $_{MJ\ (0.03)}$, $_{S\ (0.01)}$	↓**B** $_{MJ\ (0.0002)}$; ↓B $_{S\ (0.01)}$	↓B $_{MJ\ (0.007)}$; ↓R $_{MJ\ (0.03)}$	↓B $_{MJ\ (0.008)}$
80	unknown disaccharide A	0.013±0.009^a	0.003±0.003^a	0.169±0.019^b	↑N $_{MJ\ (0.03)}$; ↑R $_{S\ (0.01)}$	↑R $_{S\ (0.04)}$		
81	unknown disaccharide B	0.162±0.025^a	0.109±0.013^a	0.027±0.010^b	↓B $_{MJ\ (0.04)}$; ↓**N** $_{MJ\ (0.001)}$	↓B $_{MJ\ (0.04)}$; ↓**N** $_{MJ\ (0.001)}$	↓**B** $_{MJ\ (0.002)}$; ↓**N** $_{MJ\ (0.002)}$	
82	unknown monosaccharide	1.133±0.224^a	0.278±0.113^b					
	fatty acids							
83	linoleic acid	0.908±0.059^a	1.139±0.092^a	0.285±0.022^b	↓B $_{MJ\ (0.02)}$; ↓N $_{MJ\ (0.008)}$	↓**B** $_{MJ\ (0.004)}$	↓B $_{MJ\ (0.009)}$; ↓B $_{S\ (0.006)}$	
84	linolenic acid	0.514±0.081^a	1.778±0.254^a	0.095±0.008^b	↓**N** $_{MJ\ (0.003)}$		↓B $_{MJ\ (0.04)}$; ↓B $_{S\ (0.01)}$	↑R $_{S\ (0.04)}$
85	palmitic acid	0.427±0.033^a	2.162±0.085^b	0.143±0.021^c	↓**N** $_{MJ\ (0.003)}$			↑R $_{MJ\ (0.007)}$
	Unknowns							
86	unknown m/z 104	0.002±0.001						
87	unknown m/z 111			0.005±0.002				
88	unknown m/z 162	0.027±0.005						
89	unknown m/z 272		0.107±0.008					↑N $_{S\ (0.04)}$
90	unknown Mol Wt 274	0.854±0.073^a		0.142±0.009^b	↑B $_{S\ (0.02)}$	↓R $_{MJ\ (0.02)}$; ↓R $_{S\ (0.02)}$	↓R $_{S\ (0.03)}$; ↑B $_{MJ\ (0.02)}$; ↑B $_{S\ (0.02)}$;	↑B $_{MJ\ (0.03)}$

91	unknown m/z 302		0.816±0.192		↓N S (0.03)			
92	unknown Mol Wt 358		0.098±0.021				↑N S (0.04)	
93	unknown Mol Wt 362		0.096±0.025					
94	unknown Mol Wt 406 A	0.175±0.013[a]		0.266±0.024[a]	↑B S (0.04)			
96	unknown Mol Wt 740 A	0.098±0.017						
97	unknown Mol Wt 740 B		0.188±0.021					↑N S (0.003)
98	unknown Mol Wt 770		0.092±0.01					↑N S (0.01)

1

3.3.2   Constitutive correlations

Between plant parts for total compound groups

To examine the degree to which the amounts of compounds in the needles, bark and roots are correlated, Spearman rank correlations (r_s) of the variation in constitutive chemistry (T0) between plant parts are shown in Table 3.2. The amounts of sesquiterpenoids in the needles and bark positively correlated among individuals ($r_s = 0.55$, $p < 0.05$). Total phenolics also correlated between the needles and the roots ($r_s = 0.48$, $p < 0.05$). No other significant correlation was detected for compound groups between plant parts.

Between plant parts for individual compounds

For individual compounds between plant parts for T0, only 5 significant correlations were detected (Table 3.2). Three of these involved positive correlations between amounts of limonene [10], bicyclogermacrene [17] and unknown sesquiterpenoid alcohol [21] in the needles and the bark ($r_s = 0.57$ to 0.64). The other significant positive correlation was the amount of glucose [77] in the needles and roots ($r_s = 0.54$, $p < 0.01$). The only significant negative correlation detected was linoleic acid [83] levels in the bark and needles ($r_s = -0.53$, $p < 0.05$) (Table 3.2).

Within plants parts for total compound groups

Within plant parts, Spearman rank correlations focussed on total compound groups (Table 3.3). For these groups, high and mostly positive correlations were detected. In the bark, the amount of total monoterpenoids positively correlated with total phenolics ($r_s = 0.92$, $p < 0.001$) and with total GC-MS diterpenoids ($r_s = 0.97$, $p < 0.001$). In the needles, high positive correlation was detected between total monoterpenoids and sesquiterpenoids ($r_s = 0.90$, $p < 0.001$) and the GC-MS diterpenoids ($r_s = 0.83$, $p < 0.001$). In the roots, the only significant correlation detected was between total monoterpenoids and GC-MS diterpenoids ($r_s = 0.62$, $p < 0.01$).

Table 3.2: Spearman rank correlations between the bark [B], needles [N] and roots [R] of individual and total compounds that were identified in more than one plant part. Correlations were estimated before treatment (T0). Induced correlations from T1-T4 were estimated for only methyl jasmonate samples, for simplicity of illustrations. T1, T2, T3 and T4 represents 1, 2, 3 and 4 weeks after treatment application, respectively. The unadjusted p-values that the correlations are different from zero are indicated as * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$. All compounds were given a unique identifier based on Supplementary Table 10 (after Chapter 9), for ease of identification

	Compound/Part	B/N		B/R		N/R		T1	T2	T3	T4
						T0					
	compound groups										
	total monoterpenoids	0.29		-0.08		0.06					
	total GC-MS diterpenoids	0.12		-0.10		0.38					
	total sesquiterpenoids	0.55	*					0.83(B/N)*			
	total phenolics	-0.36		0.06		0.48	*				
	total sugars	0.49		0.43		0.94	**				
	Individual compounds										
1	α-pinene	0.31		-0.14		0.18					
3	β-phellandrene	0.30		0.28		0.14		1.0 (B/R)*** ; 0.9(B/N)*; 0.9(R/N)*	0.9 (R/N)*		0.75 (B/R)** ; 0.56 (B/N)*
4	β-pinene	0.28		-0.03		-0.07					
5	camphene	0.21									
10	limonene	0.64	**	-0.20		-0.18					
13	sabinene	0.37		-0.14		0.05		0.94 (R/N)**			
15	terpinolene			-0.04							
16	unknown Mol Wt 150	-0.08									
17	bicyclogermacrene	0.57	*								
20	trans-farnesol	0.41									
21	unknown sesquiterpenoid alcohol	0.61	**								
22	agathadiol	0.13									

#	Name										
23	agatholal	0.19									
25	levopimaral			0.01							
26	methyl dehydroabietate	0.13		0.16		0.38					
27	methyl levopimarate			0.09							1.00 (B/R)***
29	dehydroabietic acid	0.12		-0.14		0.2					
32	unknown diterpene-3	0.21		0.66	**	-0.07					
33	unknown m/z 109 A			0.25							
34	unknown m/z 109 B	0.15									
37	unknown m/z 239					-0.27					
41	C20H30O2 resin acids	0.11		0.05		0.46		0.94 (B/R)**			0.90(B/R)*
46	unknown C20H30O3	-0.26		0.36		0.13					
47	unknown C20H32O3 A	0.14		0.48		-0.42				0.58 (B/N)**	0.66 (B/N)*
50	unknown C20H30O4	0.09		0.03		0.25					-1.00(N/R)***
51	unknown C20H30O5			-0.06				0.84 (B/R)*		-0.88(B/R)*	
62	ethyl 4-ethoxybenzoate			-0.24				0.83 (B/R)*	-0.88(B/N)*		
70	raspberry ketone	0.23						0.48 (B/N)*			
73	trans-coniferyl alcohol	-0.28									
76	fructose	-0.37		0.08		0.49					-0.90 (N/R)*
77	glucose	0.66		0.43		0.77					
78	inositol	0.09		-0.26		0.54					
79	sucrose			-0.14							
80	unknown disaccharide A	-0.20		-0.13		-0.65					
81	unknown disaccharide B	0.49		-0.15		-0.88	*				0.89 (N/R)***
82	unknown monosaccharide	0.42									
83	linoleic acid	-0.53	*	0.02		0.06			0.98 (B/R)***		

84	linolenic acid	-0.26	-0.25	-0.21				
85	palmitic acid	-0.27	-0.18	0.04		-0.94 (B/R)**		-0.9 (B/N)*
90	unknown Mol Wt 274		-0.09					
93	unknown Mol Wt 406 A		0.07					

Table 3.3: Spearman's rank correlations of the constitutive (T0) amounts of total compound groups within the bark [B], needles [N] and roots [R]. The unadjusted p-values that the correlation is different from zero are indicated as * p <0.05, ** p <0.01 and *** p <0.001. No sesquiterpenoids were detected in the roots

		Total monoterpenoids		Total GC-MS diterpenoids		Total sesquiterpenoids	Total phenolics	Total sugars
Total GC-MS diterpenoids	B	0.97	***					
	N	0.83	***					
	R	0.62	**					
Total sesquiterpenoids	B	0.62	**	0.55	*			
	N	0.90	***	0.77				
	R	NA		NA				
Total phenolics	B	0.92	***	0.89	***	0.55 *		
	N	0.19		0.15		0.25		
	R	-0.07		-0.05		NA		
Total sugars	B	-0.11		-0.11		-0.26	-0.10	
	N	-0.19		-0.01		-0.09	-0.03	
	R	0.30		0.32			0.18	
C20H30O2 resin acids	B	0.32		0.28		0.52 *	0.30	-0.41
	N	0.43		0.46		0.45	0.08	0.17
	R	-0.21		-0.27		NA	-0.10	-0.07

3.3.3 Induced responses

Total compound groups

To understand how the total compound groups responded to treatment, pair-wise Kruskal Wallis tests comparing the MJ and strip treatments to the controls at each point in time revealed minor changes in secondary compounds (Table 3.1). While non -significant changes were detected in the bark and needles for total monoterpenoids and phenolics, no response was detected in the total, sesqui- and diterpenoids at any time of measurements (Table 3.1), for both MJ and strip treated plants. The GC-MS diterpenoids was the only total compound group that non-significantly responded to treatment in the roots at T3 in bark stripped samples (Table 3.1).

Contrary to the lack of significant change in secondary compounds after treatment, the amount of total sugars was significantly reduced in the needles ($p < 0.001$; *Dn, Bf*) and bark ($p < 0.001$; *Dn, Bf*), but not in the roots. The reduction occurred only in the MJ treated plants at T1 and T2. In the strip-treated plants, the reduction in total sugars was noted but was not significant after Bonferroni correction (Figure 3.5)..

The time progression of total compound groups appeared to differ between total compound groups, plant parts and treatments (Figure 3.5), for instance total sugars reduced after MJ treatment in all plant parts at all time points but tended to increase at some points following stripping. Overall, for the secondary compound groups, while an overall increase in amounts was detected over time in the bark for the total monoterpenoids, GC-MS diterpenoids and total phenolics, an overall reduction was observed in the needles and roots (Figure 3.5a and b). However, it was observed that in the case of the GC-MS diterpenoids, a reduction in the needles was followed by almost an equivalent increase of the total compounds in the bark at each time point, while for the monoterpenoids and phenolics reduction at T2 in the needles was followed by an increase in the amounts of these compounds at T3, which could signal translocation of metabolites from the needles to the bark. The LC-MS C20OH3002 resin acids [41] was the only total compound group that notably changed over time by reducing in amount for both MJ and strip-treated plants at T2 (Figure 3.5), but these changes were not significantly different between the treated and control samples. The total sugars were generally reduced after treatment for both MJ and strip induced chemistry. Comparing the treatments, differences were noted in the bark especially for the total monoterpenoids and GC-MS diterpenoids that increased over time in MJ-treated plants (Figure 3.5a) but reduced in the strip-treated plants (Figure 3.5b), compared to the control. Although these changes were non-significant, this suggests that bark stripping may cause different chemical responses relative to other stressors.
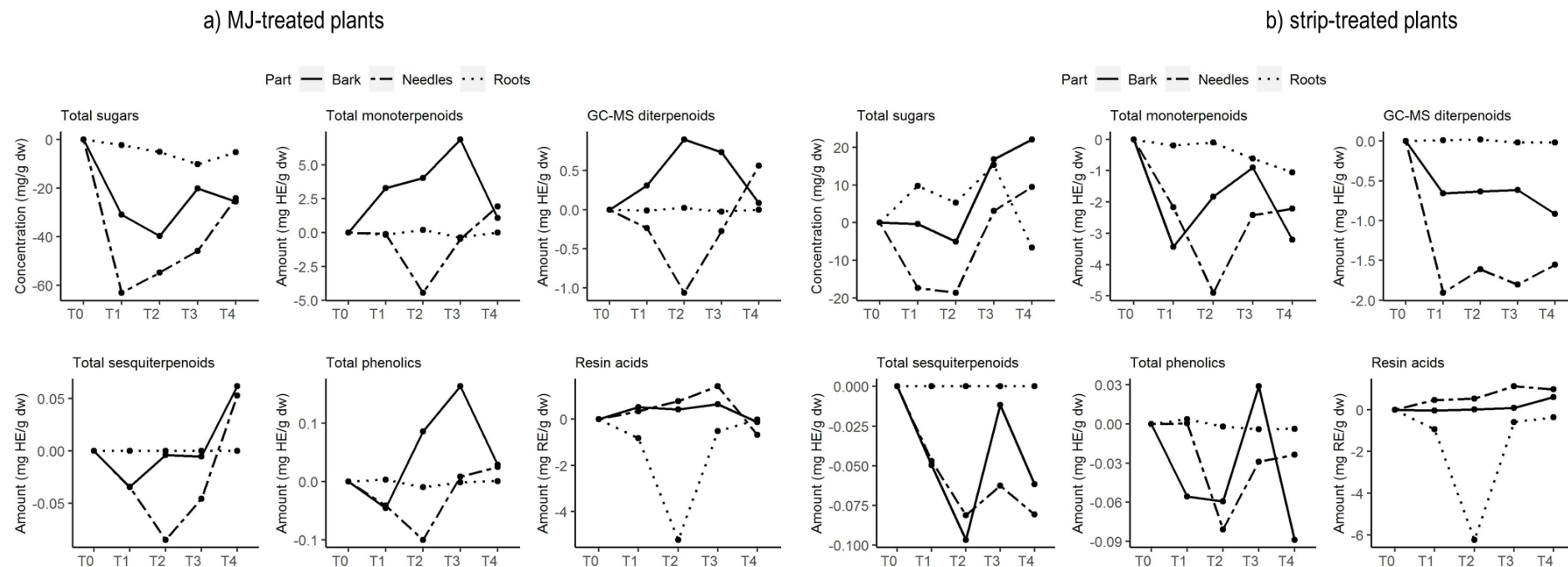
Figure 3.5: Average progressive change in the amounts of total compounds in a) methyl jasmonate (MJ) and b) bark strip-treated plants relative to the controls in the bark, needles and roots of *Pinus radiata* seedlings. At each time point, induced changes were detected by comparing the mean values for the MJ and control treatments (mean of treatment – mean of control) for a specific time. The comparisons were undertaken at each sampling time: T0 - before treatment applications, then, T1, T2, T3 and T4, which are respectively 7, 14, 21 and 28 days after treatment application. The terpenes and phenolics are expressed as mg heptadecane equivalents (HE)/ g dw, while the resin acids are expressed as mg rutin equivalents (RE)/g dw, as no absolute quantitation was carried out on these analytes. The absolute amounts of sugars are quantified in mg/g dw.

Individual compounds

To detect the induced changes at individual compound level, results showed that similar to the non-significant changes in total monoterpenoids and total phenolics, individual monoterpenoids (α-terpineol, α-pinene, camphene, linalool) and phenolics (vanillin and anethole) responded to treatment especially in the MJ treated bark samples at different times but were marginally non-significant after Bonferroni adjustment (Table 3.1). Several unknown diterpenoids: unknown $C_{20}0H_{30}O_6$ C [54], unknown $C_{20}0H_{30}O_3$ [46] and unknown m/z 316 [45] in the bark also reduced with similar magnitude (Table 3.1). However, compounds in the roots responded to bark stripping more than MJ treatment. In the roots, α-pinene[1], methyl levopimarate [27], levopimaral [25], methyl dehydroabietate [26] and several unknowns; the diterpenoids unknown $C_{20}H_{30}O_3$ [46] and unknown $C_{20}H_{30}O_4$ [50], unknown Mol Wt 274 [90], unknown $C_{19}H_{26}$ [28] and unknown m/z 109 A [33] increased in bark stripped but not MJ treated samples but were marginally not significant after Bonferroni adjustment (p <0.1). This may indicate some level of specificity in plant responses to bark stripping relative to other stressors like MJ. Almost all the amounts of secondary compounds that reduced in the roots increased in the bark and/or needles, which may suggest a translocation from the roots to the bark or needles.

Stronger changes were however, detected in individual sugars. In the bark of MJ treated plants, fructose [76] and glucose [77] reduced significantly at T1-T2. Glucose [77] and the unknown disaccharide [81] also reduced significantly at T3. In the bark of strip-treated plants, non-significant reduction in glucose and fructose were noted at T1 and T2. In the needles, fructose [76] and glucose [77] reduced in the MJ treated plants at T1 and T2. Inositol in the needles also reduced at T2 (p <0.01; *Dn, Bf*). The unknown disaccharide [81] also reduced in the needles from T1-T3. The responses of individual sugars in the roots were not evident although sucrose reduced in the roots of MJ treated plants at T3, but the changes were not significant after p-adjustment. The individual fatty acids were more responsive in the needles than in the bark (Table 3.1). However, linoleic acid [83], reduced in the bark of strip-treated samples at T3 and in the roots palmitic acid increased non significantly in the MJ treated plants at T4 (Table 3.1, Figure 3.5).

As observed with the total compound groups, the induced responses and time progression of the individual compounds also varied between individual compounds, plant parts and treatments. For individual compounds that were identified in at least two plant parts, this is illustrated in Supplementary Figure 3.2. Secondary compounds for example generally increased in the bark but reduced in the needles and the roots following MJ treatment (Supplementary Figure 3.2). In contrast, following strip treatment, an overall reduction in secondary compounds was observed in all the plant

parts with a few exceptions for example the diterpenoids- levopimaral [25], methyl dehydroabietate [26] and methyl levopimarate [27] that increased in the bark following both MJ and strip treatments. Also, the sugars returned to normal levels faster in the strip than in the MJ treated plants. Overall, MJ caused greater responses than bark stripping in the bark and needles.

Spearman's rank correlations after treatment
Between plant parts
To detect the influence of treatments on the correlations, Spearman's rank correlations after treatment showed more significant correlations for total compound groups and individual compounds between plant parts. Induced positive correlations were for example detected between the total sesquiterpenoids of the needles and the bark and the resin acids of the bark and the roots (Table 3.2). For individual compounds, the bark and the needles mostly exhibited positive correlations after treatment except for the negative correlation of alpha-pinene and palmitic acid at T4 (Table 3.2)

Within plant parts
Within plants Spearman's rank correlations after treatment were performed only for compound groups at T1. Results showed that more correlations that were not detected at T0 became evident, for example significant negative correlation between total monoterpenoids and total phenolics in the needles. In contrast, some correlations detected at T0 were not significant after treatment, for example the correlation between the monoterpenoids and diterpenoids in the roots was not significant after treatment. Instead, in the roots a positive correlation was detected between the monoterpenoids and the resin acids in both the MJ and strip-induced chemistry. More significant correlations were detected with MJ- induced chemistry than strip-induced chemistry (Tables 3.4 a and b).

Table 3.4: Spearman's rank correlations of compound groups within plant parts at T1 in a) MJ-treated and b) strip-treated plants. P values were set as $p < 0.05$*, $p < 0.01$** and $p < 0.001$***. The values were not Bonferroni adjusted

a) MJ-treated plants

		Total monoterpenoids		Total GC-MS diterpenoids		Total sesquiterpenoids		Total phenolics	Total sugars
Total GC-MS diterpenoids	B	0.94	**						
	N	0.94	**						
	R	-0.14							
Total sesquiterpenoids	B	0.54		0.60					
	N	0.83	*	0.89	*				
	R	NA							
Total phenolics	B	0.43		0.37		-0.26			
	N	-0.83	*	-0.77		-0.54			
	R	0.09		0.31		NA			
Total sugars	B	0.71		0.60		0.09		0.43	
	N	-0.71		-0.60		-0.43		0.89 *	
	R	-0.54		-0.09		NA		-0.66	
C20H30O2 resin acids	B	-0.09		0.03		0.71		-0.37	-0.60
	N	0.54		0.60		0.83		-0.37	-0.14
	R	0.83	*	-0.31		.NA		-0.20	-0.31

b) strip-treated plants

		Total monoterpenoids	Total GC-MS diterpenoids	Total sesquiterpenoids	Total phenolics	Total sugars
Total GC-MS diterpenoids	B	0.09				
	N	0.77				
	R	0.14				
Total sesquiterpenoids	B	0.81	0.32			
	N	0.71	0.94 **			
	R	NA	NA			
Total phenolics	B	-0.31	-0.09	0.14		
	N	0.03	-0.14	-0.31		
	R	0.61	0.49			
Total sugars	B	-0.31	0.77	-0.03	-0.09	
	N	0.60	0.37	0.26	-0.31	
	R	0.27	0.11	.NA	0.54	
C20H30O2 resin acids	B	0.49	0.03	0.09	-0.49	0.71
	N	0.83 *	0.09	0.46	-0.54	-0.14
	R	0.83 *	-0.14	.NA	0.64	-0.26

1

3.4   Discussion

This study showed that: (i) *P. radiata* harbors a diversity of primary and secondary compounds that occur in one or other plant part and that the chemical profile between bark, needles and roots is quite different; (ii) the chemical compounds in the roots responded to above-ground treatment with methyl jasmonate and artificial bark stripping but the progression of the changes over time differed for the chemical compounds, plant parts and treatments; and that (iii) mostly positive correlations between the amounts of compounds occur within and between plant parts. To date, comparatively few studies have examined plant wide variation in chemistry across plant parts and this is the first study to examine the secondary compound composition of roots in *P. radiata* and how it varies to other plant parts. Results have shown that *P. radiata* roots have high numbers and amounts of both primary and secondary metabolites, with unique compounds also being recorded. However, the number of compounds detected in the roots was less than that detected in the bark and the needles. The allocation patterns of the amount of compounds in the roots vs bark and needles also varied depending on the class of chemical compounds, for example, with the sugars, fructose dominated in the roots while glucose dominated in the bark and the needles.

The temporal and spatial variation in the amount of secondary metabolites in a plant is thought to be shaped by organisms that interact with the plant and the nature of the interactions (Franceschi *et al.* 2005; McKey 1979). According to the optimal defence theory, secondary metabolites that act as defences will be concentrated in parts or tissues that are at higher risk of herbivory and/or tissues that are more valuable (McKey 1974). In this present study, the higher number of compounds in the bark may imply that the bark is more protected since it is nutrient rich and sought by many herbivores (Felicijan *et al.* 2015; Franceschi *et al.* 2005). In addition, bark damage may have a stronger impact on plant fitness as it increases the risk of secondary infection to the wood (Franceschi *et al.* 2005; Welch *et al.* 1988) and this can reduce the mechanical, hydraulic, and physiological integrity of a tree. Additionally, severe bark stripping can lead to complete ring-barking of the tree and tree death. Studies also show that *Pinus radiata* can easily compensate for a relatively high level of defoliation (Eyles *et al.* 2011; Lombardero *et al.* 2016), which may explain why the number and amount of secondary compounds was lower in the needles compared to the bark. Similar patterns have been detected in other *P. radiata* populations (Apetrei *et al.* 2011; Lundborg 2016). The lower amounts of mono and di-terpendoids and phenolics in the roots and the absence of sesquiterpenes altogether, implies that this plant part is not attacked by pests and pathogens to the same degree as the above

68

ground parts. However, currently, there are no exhaustive surveys of aboveground versus belowground herbivory in most conifers.

In regards to specific individual secondary compounds, some compounds were unique or were expressed more in one plant part than the other parts. Although the compounds that have been well described in terms of their involvement in defence activities in conifers such as α/ β-pinene, camphene and limonene were found in all three plant parts, several monoterpenoids, diterpenoids and phenolics were below detection levels in the needles and the roots. Of note was the missing sesquiterpenoids in the roots. The role of sesquiterpenoids in pest-conifer relationships is not well established but their potential exploitation as precursors for insect hormones and also as signals to attract insect parasitoids in tri-trophic interactions (Celedon and Bohlmann 2019) is consistent with the reduced exposure of roots to pests. We demonstrated an overlap in presence and quantity of some compounds between plant parts and this may be beneficial where some pathogens and herbivores are non-specific in plant part preferences; for instance, the pitch canker affects both below and above ground parts (Mead 2013). Conversley, different tissues may be attacked by different herbivores and pathogens and plant part specific profiles in compounds may cater to this, for example, bark-eating herbivores rarely overlap with leaf herbivores (Leimu and Koricheva 2006). All identified compounds have been previously reported in conifers (Salem *et al.* 2014; Zhang *et al.* 2016a), and most in *Pinus radiata* (Cool and Zavarin 1992). However, in this work many unidentified diterpenoids were also important in differentiating the needles, bark and roots and require further investigation.

The constitutive amount of total and individual sugars and fatty acids was highest in the needles, in this study, which is consitent with other studies in *P. radiata* (Cranswick *et al.* 1987) and other conifer species (Dobbelstein *et al.* 2019; Piper *et al.* 2017). Interestingly, we did not detect any sucrose in the needles. Sucrose is the end product of photosynthesis and the primary sugar transported in the phloem of most plants and is expected to be higher than gulcose and fructose (Tauzin and Giardina 2014). Its absence may suggest that the needles were not photosynthetically very active at the time of sample collection or that sucrose is inherently lower than the other sugars in *P. radiata*. Cranswick *et al.* (1987) also showed that sucrose was relatively low compared to glucose or fructose in *P. radiata* needles compared to the bark and the roots. Sugars in conifers are very prone to seasonal changes and, therefore, the amount detected in this study could possibly be affected by sampling time (Cranswick *et al.* 1987; Dobbelstein *et al.* 2019). The sampling in this study was done at the onset of winter (May-June), when photosynthetic activity of most plants is reduced. Generally, the distribution of non-structural carbon compounds (NSC); including glucose and fructose between

photosynthesizing needles and non- photosynthesizing plants parts has been viewed as a passive sink–source process resulting from imbalances between carbon supply and demand (Wiley and Helliker 2012). However, the results may also highlight compound-specific, environmental or even seasonal/age related accumulation of primary and secondary metabolites in *Pinus radiata*.

Maintaining metabolite diversity may pose possible conflicts in resource allocation in the plant that manifest as negative correlations between pairs of traits that share a resource (Kant *et al.* 2015; Saeki *et al.* 2014). In the constitutive amounts, the results, however, showed limited trade-offs in resource investment between and within plant parts for individual primary and secondary metabolites, or between secondary compounds, except for one fatty acid - linolenic acid. Similar studies with pine species (Deslauriers *et al.* 2015; Sampedro *et al.* 2011; Villari *et al.* 2014) including *P. radiata* (Moreira *et al.* 2013b) have also not found strong evidence of trade-offs. For constitutive secondary metabolites, it has been suggested that trade-offs are not expected where resources are sufficient (Sampedro *et al.* 2011) and when reduction in herbivory is achieved by multiple defence traits (Pearse *et al.* 2018), for example in the juveniles of *P. radiata* that support diverse defoliating and bark-specialized insects and mammals and root feeding insects and pathogens (Mead 2013; Moreira *et al.* 2012a). Trade-offs are also uncommon for long-lived trees that need to defend themselves against a wide array of herbivores (Iason *et al.* 2011; Snyder 1993). Sometimes, trade-offs are disrupted in crops like *P. radiata* that have undergone breeding and cultivation that reduces pest and pathogen pressures in non-native habitats and subsequent investment in secondary compounds (Kempel *et al.* 2011).

Roots, bark and needles were differentially responsive to MJ and stripping

This is the first study to show induced chemical responses in roots of *Pinus radiata* to above ground stressors. In response to treatment, roots mostly reduced the amounts of secondary metabolites and increased the sugars and fatty acids. Individual compounds in the roots responded more to artificial bark stripping than methyl jasmonate. Although the changes were non-significant after the Bonferroni adjustment, the results suggest that roots are more sensitive to mechanical damage than other kinds of above ground stresses. Overall, differential responses were detected between compound groups, plant parts, and treatments at each time period. Stronger responses were detected in primary compounds compared to secondary compounds, with a very strong overall reduction of sugars especially in the bark and the needles. The consistent reduction of glucose, fructose, sucrose and fatty acids at various times of measurement following the treatments suggests their significant involvement in induced stress responses. The reconfiguration of sugars following herbivory and similar

stress treatments could result from diminished photosynthesis, which would decrease the overall pool of energy reserves as a result of damage to photosynthetic machinery, loss of photosynthetic tissue, and/or disruption of the vasculature affecting water and sugar transport, and/or from a diversion of resources (Gershenzon 1994; Huot *et al.* 2014; Schwachtje and Baldwin 2008). While the reduction in photosynthesis has been observed in *P. radiata* after methyl jasmonate treatment (Gould *et al.* 2008), this has been explained by the following premises:

I. resistance traits are costly and frequently up-regulated after attack - the cost is reflected as trade-offs that manifest as negative correlations among chemical traits or between chemicals and growth, reproduction or storage; For this study, within plant parts, although there was a tendency for induced sugars to negatively correlate with secondary compounds, the correlations were not significant and so there is no evidence that defence is costly. This contrasted with similar studies in conifers that provide evidence that the biosynthesis of terpenes is energetically demanding and induction relies on phloem carbohydrate resources (Goodsman *et al.* 2013; Raffa *et al.* 2017; Roth *et al.* 2018). Instead of a carbohydrate-based resource trade-off, Machado *et al.* (2017) suggested hormonal antagonism to explain the reduction of sugars; However, sugars may also be used as carbon skeletons for replacement of lost tissues (Stein and Granot 2019).

II. resources are translocated to areas inaccessible to herbivores to support the physiological adjustments for subsequent recovery- an aspect of tolerance that has been implicated in other *Pinus* species (Moreira *et al.* 2012b; Sampedro *et al.* 2011). However, this was not strongly supported by this study because the unknown disaccharide increased nonsignificantly in the roots;

III. sugars function as signals in defence pathways, where sugar signals can reportedly be generated either by carbohydrate concentration and relative ratios to other metabolites, such as C:N or by flux through sugar-specific sensors and/or transporters (Eveland and Jackson 2012). Although it was noted that the ratio of total monoterpenoids to total sugar increased 5-fold in MJ treated samples (data not shown), the detailed analysis was out of scope for this chapter;

IV. induced changes in primary metabolism could themselves be defensive, but whether sugars have a direct defensive role needs to be tested in *Pinus radiata*. The direct role of sugars in defence has been demonstrated in *Pseudotsuga menziesii*, where sucrose increased the resistance of the trees to western spruce budworm (*Choristoneura occidentalis)* (Clancy 1992). Similarly, in most fungal pathogen–plant systems, a high level of sugars in plant tissues enhances plant resistance (Morkunas and Ratajczak 2014). However, indirectly, the

reduction in sugars that was observed in this study may reduce forage quality for herbivores or flow into the root-soil interface to attract beneficial soil microbes as an indirect defence response (Schwachtje and Baldwin 2008; Zhou *et al.* 2015).

Few other studies in conifers have directly demonstrated a reduction in sugars following real herbivory or MJ application (Roth *et al.* 2018) while some work has shown a change in the physiological functions of affected/treated plants with implications for sugar levels (Gould *et al.* 2008; Ralph *et al.* 2006; Roth *et al.* 2018). The reduction in sugars contradicts observations with *Pinus pinaster* that did not show any response in bark sugars following methyl jasmonate treatment (Sampedro *et al.* 2011) and in *Fagus sylvatica* bark that showed an increase in glucose and fructose after bark stripping (Saint-Andrieux *et al.* 2009). Sometimes, responses of sugars will depend on the amount of bark removed from the plant, for example girdling experiments show an accumulation in leaves and bark above the girdle and a strong decline in soluble sugar and starch concentrations in organs below the girdle (roots) (Li *et al.* 2003).

The strong reduction of fatty acids in the needles and the bark following treatments in this and similar studies (López-Goldar *et al.* 2018) shows that they may have a role in plant defences, because fatty acids can be precursors to the formation of secondary compounds (Kachroo and Kachroo 2009). Additionally, direct defence properties of specific fatty acids against the fungus *Dothistroma pini* in the needles of *Pinus radiata* have been documented (Franich *et al.* 1983).

For secondary compound groups and individual compounds, relatively weak up- and down-ward responses of individual mono- and diterpenoid compounds in the bark, needles or roots following both bark stripping or methyl jasmonate treatments were detected. The sesquiterpenoids did not respond to either bark stripping or MJ application. This contrasted with results from Moreira *et al.* (2013a) who detected significant changes in individual terpenoids of the bark after 5 days of treatment with the large pine weevil and the pine processionary caterpillar in *P. radiata* and *P. pinaster*. Therefore, differences in induction within and between populations may reflect the timing of assessment and the nature of the treatments to which the plants are subjected. Nevertheless in this study, the patterns of induced changes differed between plant parts. While an overall increase in amounts of compounds was detected over time in the bark, an overall reduction was observed in the needles and roots. The patterns seemed, in part, to reflect reallocation of compounds from the needles to the bark since the reduction in the needles was synchronized with the increment in the amount of compounds in the bark at the same time point or the subsequent time point. It was also noted that most compounds that tended to reduce in the

72

roots mostly increased in the bark. The negative correlations of compounds between plant parts after treatment may also suggest reallocation of the involved compounds. Metabolite reallocation following stress have been suggested in conifers (Hammerbacher *et al.* 2011; Wu *et al.* 2017). Especially for phenolics, studies indicate that induced accumulation can be due to rapid translocation (Hammerbacher *et al.* 2011). For terpenoids, although passive changes have been indicated in the constitutive form, it has been suggested that all stress-induced terpenes are entirely of a *de novo* nature (Wu *et al.* 2017).

The responsive compounds that included particularly the monoterpenoids and phenolics are normally thought to increase defence against stresses (Karban and Myers 1989). The monoterpenoids (e.g. α-pinene, linalool and camphene) that marginally increased following treatment in this study have been extensively studied in insect/plant systems as feeding deterrents, pheromone precursors and synergists, and also as feeding attractants at low concentrations (Erbilgin *et al.* 2017; Miller and Borden 2000). However, the roles of these compounds as induced defences against mammalian herbivores has been poorly studied and volatiles may be important in mammals that use olfactory cues in determining forage quality, such as marsupials (Stutz *et al.* 2017). The roles of phenolics and diterpenoids in defence against major herbivore groups in conifers has not yet been thoroughly investigated (Sunnerheim-Sjöberg and Hämäläinen 1992; Vourch *et al.* 2002; Whitehill and Bohlmann 2019). The importance of phenolics has been often studied in relation to abiotic stress, where they play important role in several physiological processes to improve the tolerance and adaptability of plants under suboptimal conditions (Close and McArthur 2002; Sharma *et al.* 2019). Some unidentified compounds significantly responded to the treatments suggesting that they could have a direct, additive or synergistic defence role, possibly in less studied herbivores.

Induced responses following real or artificial herbivory have been widely reported in conifers. Within *Pinus radiata*, most studies have reported significant increase in total or individual terpenes and phenolics following herbivory (Gould *et al.* 2009; Lundborg *et al.* 2019; Moreira *et al.* 2012a; Zas *et al.* 2014), with a few exceptions (Lombardero *et al.* 2013). The differences in timing of expression of different compounds in the needles and the bark were similar to other observations in *P. radiata* (Reglinski *et al.* 2017) and other *Pinus* species (Raffa and Smalley 1995; Reglinski *et al.* 2017). The expression of compounds at different times is thought to be related to the associated costs of expression (Bonello *et al.* 2006). While bark stripping can elicit a response in *Pinus radiata*, the quantitative responses at each time point were weaker than those from methyl jasmonate treatment. Other studies have also reported similar results in *P. radiata* (Gould *et al.* 2009). In other conifer species, studies

showed limited induction in needles or bark following artificial wounding (Hudgins 2003; Pham et al. 2016).

After treatment, this study detected more significant correlations of compounds within and between plant parts compared to the chemistry before treatment. The negative correlations are usually attributed to resource limitations especially in this case where the traits involved rely on a common pool of resources (Sampedro *et al.* 2011). Whether the observed trade-offs are genetic or can translate into reduced growth remains to be tested. Gould *et al.* (2008) also showed that *in P. radiata* the costs of expressing defences may be short lived and easily compensated for and are expressed only under nutrient deficiency (Sampedro *et al.* 2011). Hence further understanding the genetic basis of these trade-offs should provide a foundation for the development of breeding strategies. However, the induced responses of methyl jasmonate were more visible in the bark than in the needles or roots, which was contrary to the theory of trade-offs between constitutive and induced chemical defences that has been documented in pine trees (Moreira *et al.* 2014; Sampedro *et al.* 2011).

3.5   Conclusion

This study has shown that plant parts harbour unique assemblages of compounds. The allocation of compounds to the roots differs from that of the needles and bark. The results showed strong changes in sugars and fatty acids that has not be documented in most conifers and limited induction of secondary compounds. The importance of sugars and fatty acids in conifer defences need a more comprehensive investigation. This detailed assessment of *P. radiata* chemistry in the needles, bark and roots, before and after stress will potentially facilitate the identification of related defence traits. However, incorporating the effect of genetics and age for target herbivore species will further our understanding of *Pinus radiata* defences, and their potential in the management of browsing herbivores.
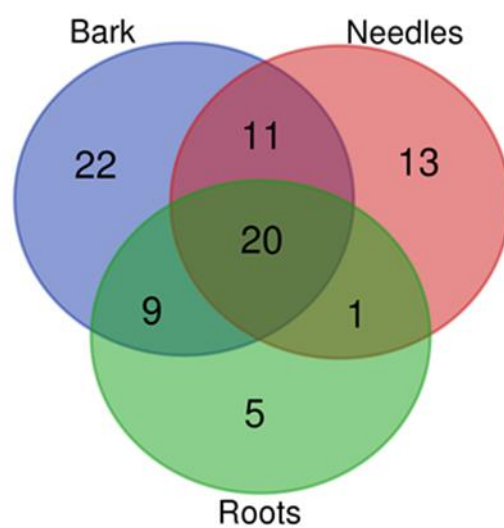
Supplementary Figure 3.1: Number of compounds detected in the bark, needles and roots of *Pinus radiata* at T0 (before treatment).

a) Monoterpenoids
   Monoterpenoids - MJ
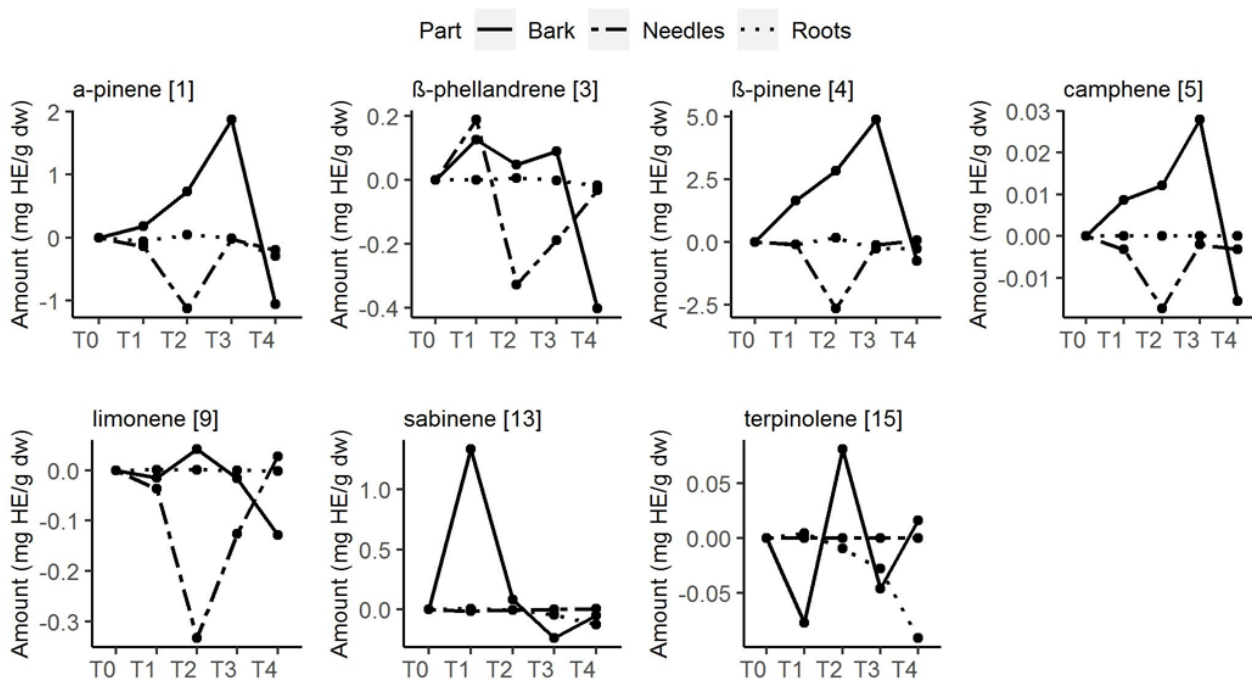
Monoterpenoids - strip
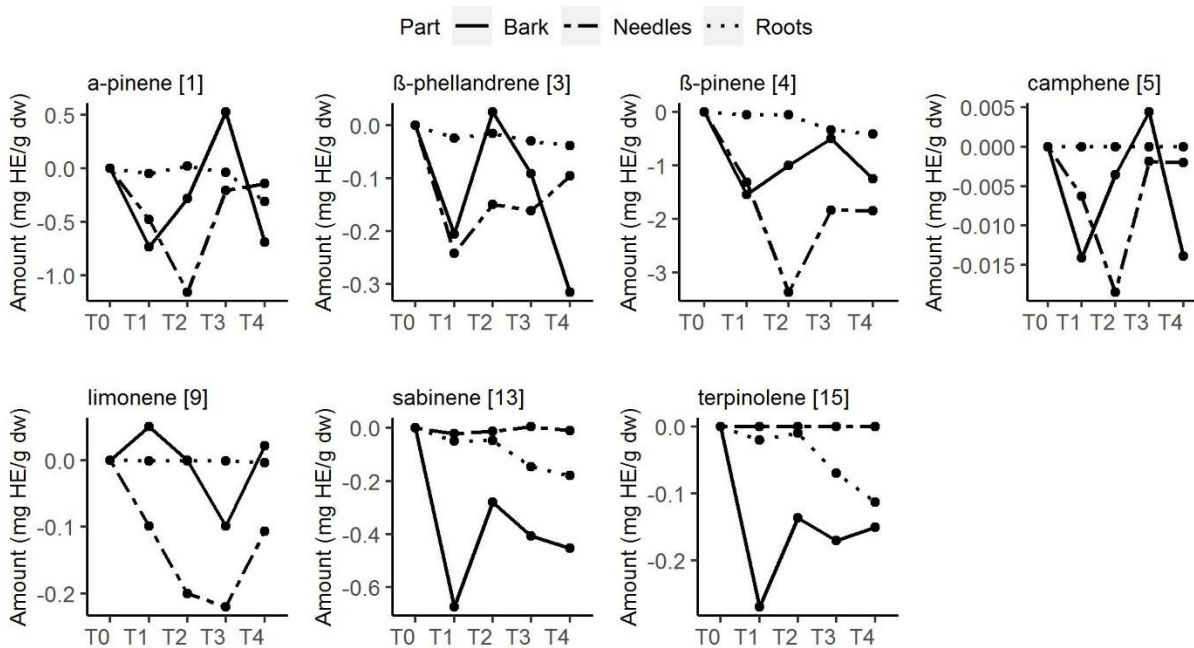
b) Sesquiterpenoids

Sesquiterpenoids - MJ
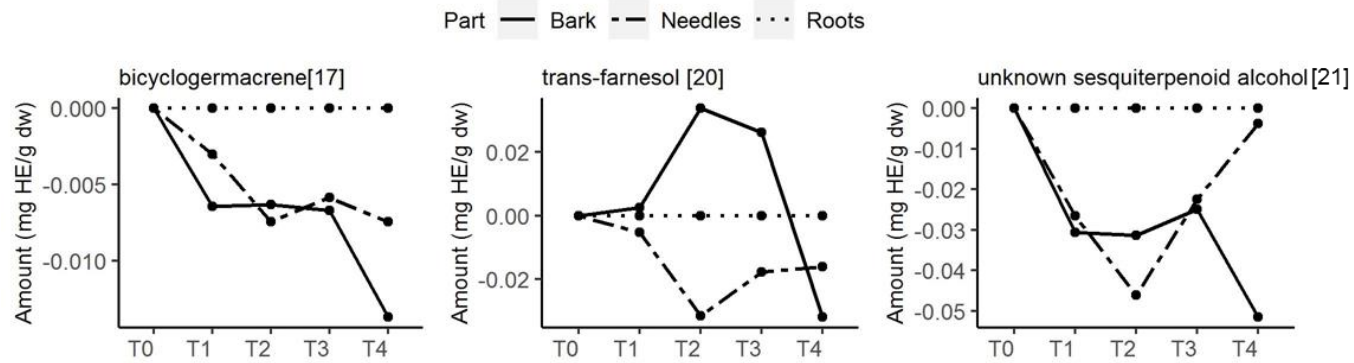
Sesquiterpenoids – strip

c) Diterpenoids

Diterpenoids - MJ

Diterpenoids - strip

Part — Bark — Needles ··· Roots

Part — Bark — Needles ··· Roots

78
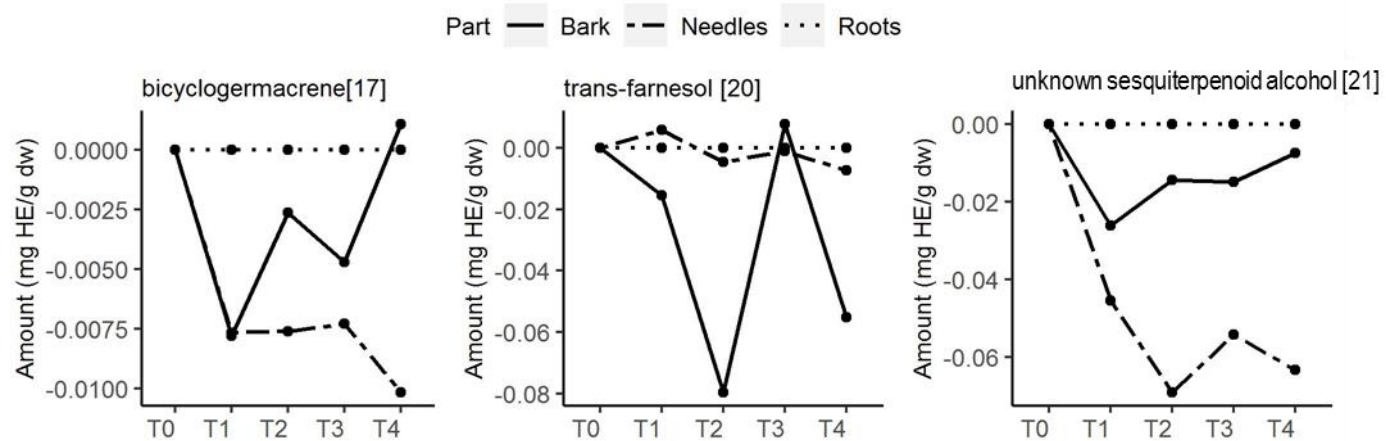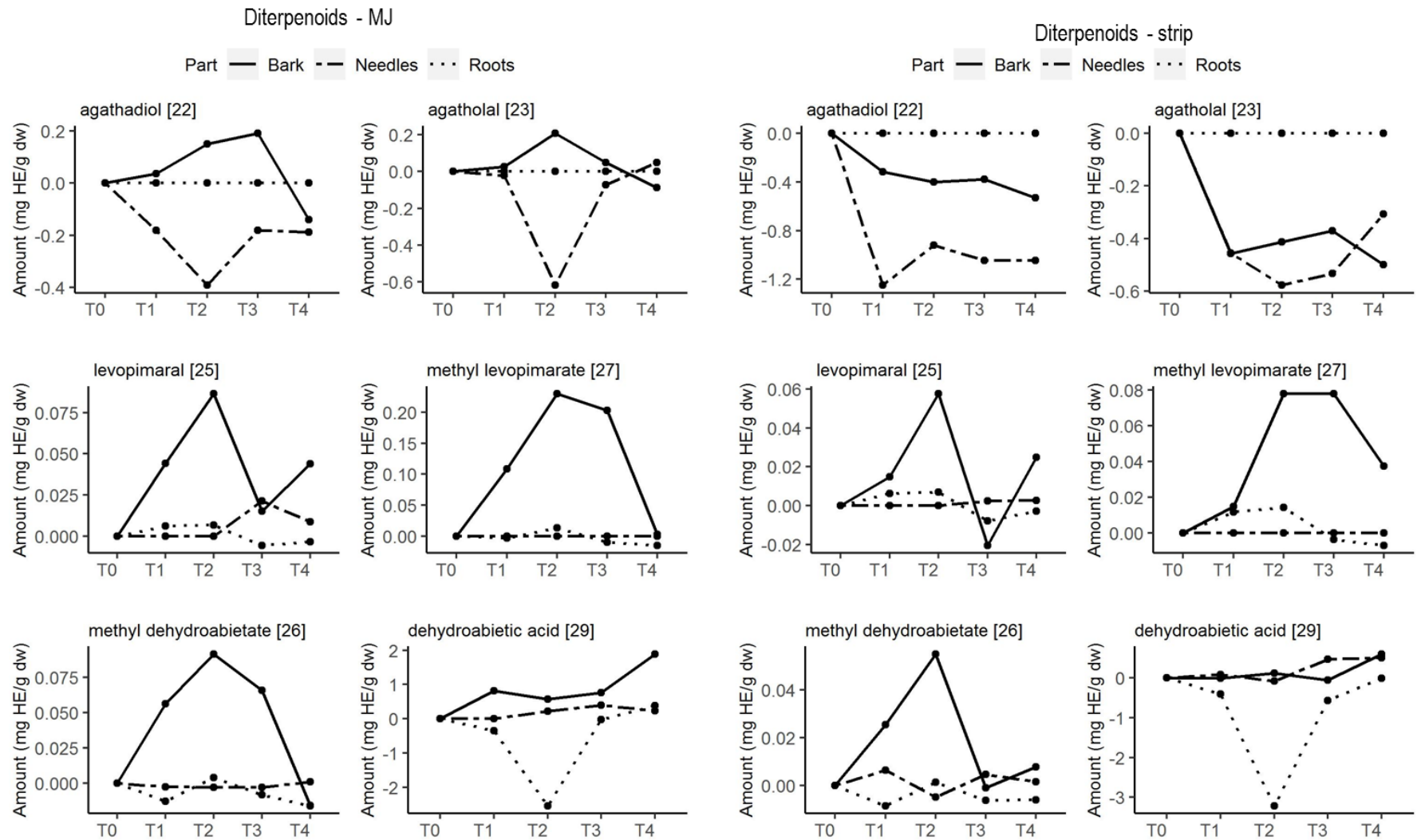
d) Phenolics

Phenolics - MJ
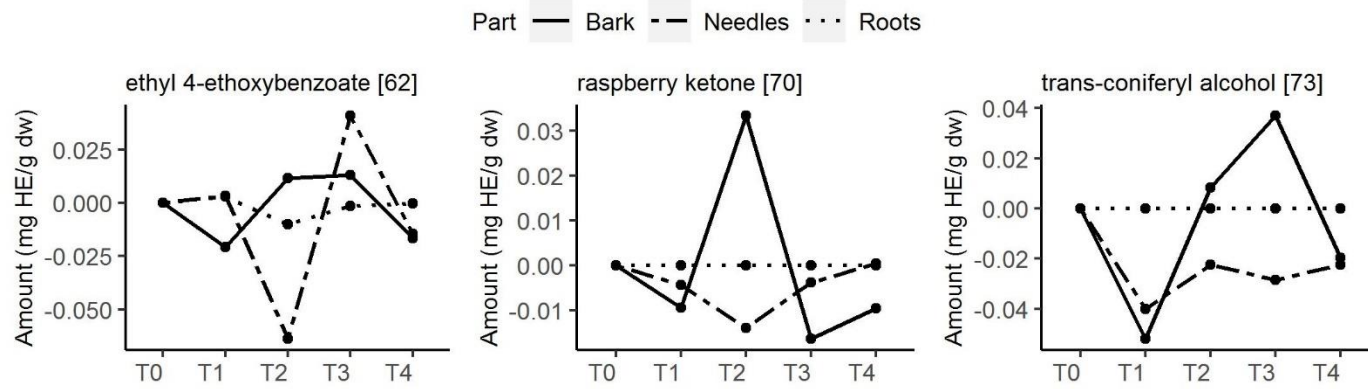
Phenolics – strip
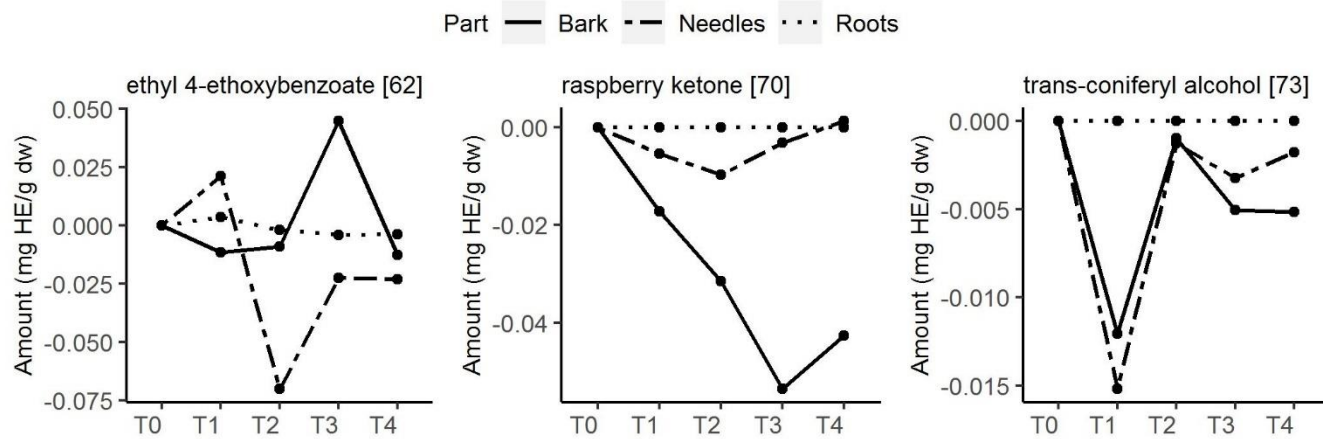
e) Sugars
Sugars - MJ
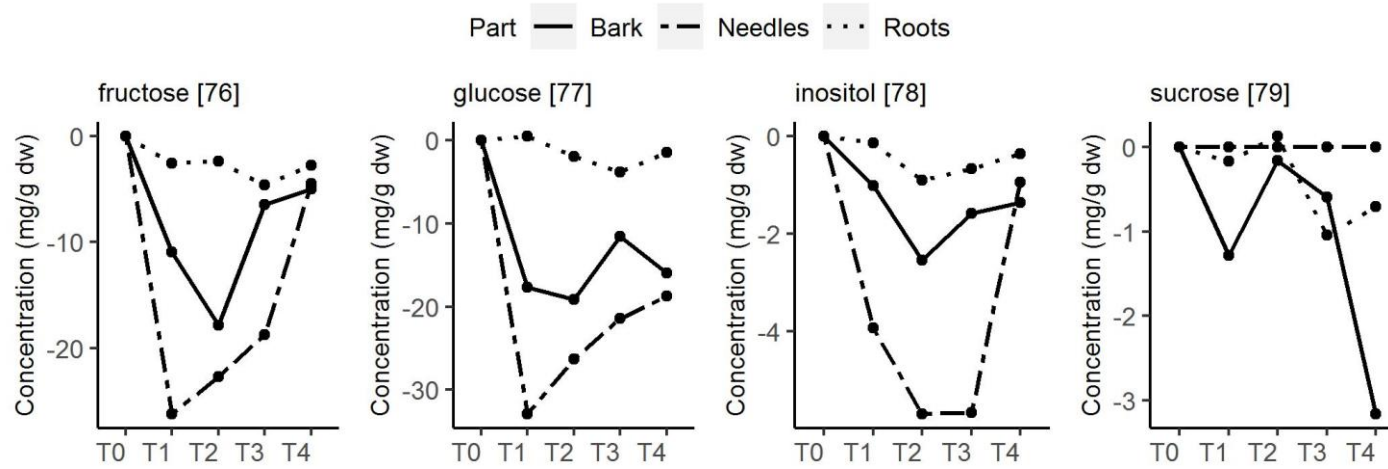
Sugars – strip
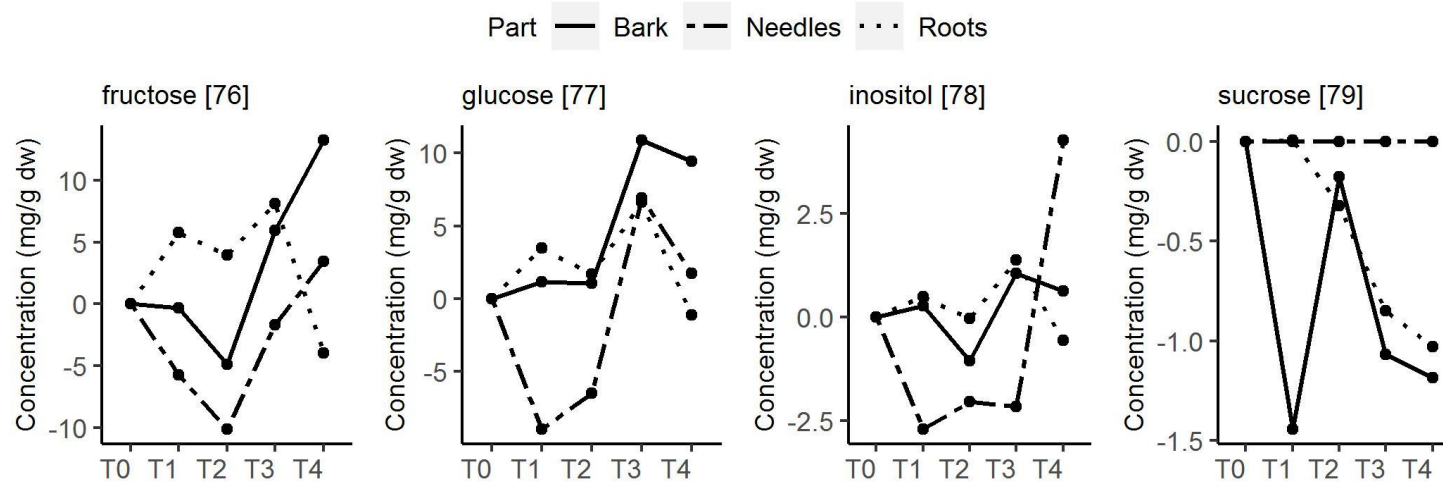
f) Fatty acids

Fatty acids - MJ
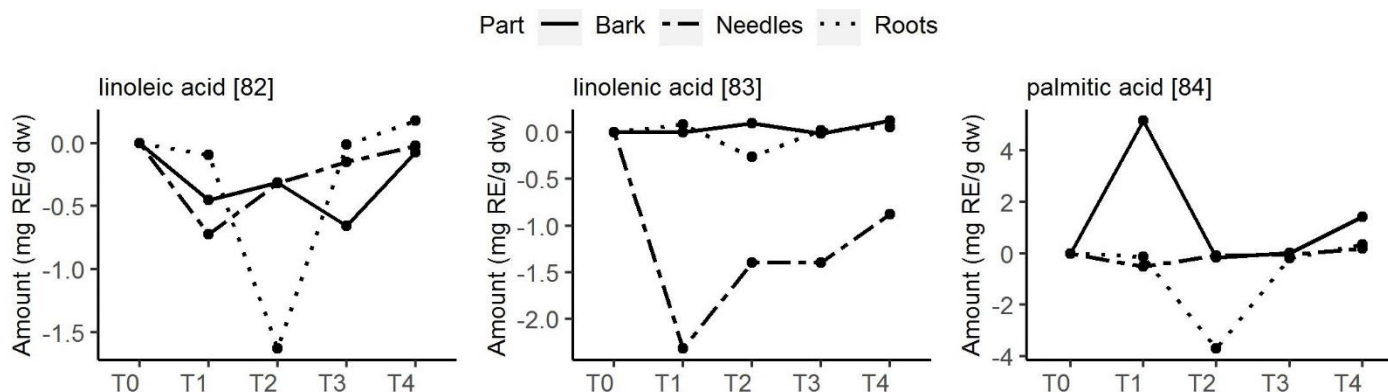
Fatty acids – strip
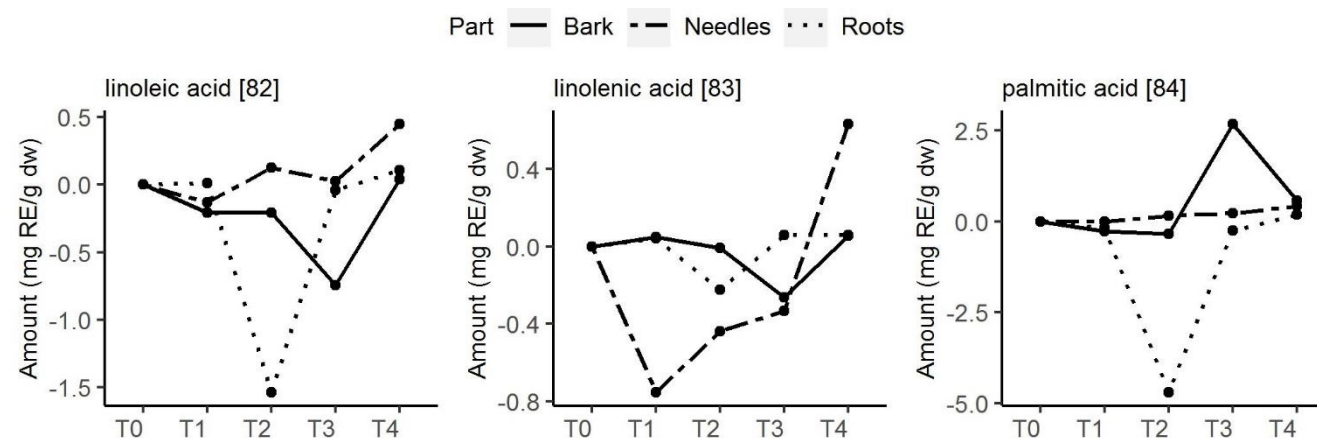
Supplementary Figure 3.2: Time progression and average change in the amount of individual compounds that were identified in at least two plant parts and their response to treatment relative to the control in the bark, needles and roots of *Pinus radiata* seedlings. At each time point, induced changes were detected by comparing the mean values for the treated and control samples (mean of treatment – mean of control) for a specific time. The comparisons were undertaken at each sampling time: T0 - before treatment applications, then, T1, T2, T3 and T4 which are respectively 7, 14, 21 and 28 days after treatment application.

CHAPTER 4: Chemical traits that predict susceptibility of *Pinus radiata* to marsupial bark stripping

4.0 Abstract

Herbivory is determined by a balance between primary metabolites that are the sources of nutrition and secondary metabolites that act as defences. Identifying the compounds that influence herbivory may be a useful tool in the management of forest systems. This study aimed to detect and identify both constitutive and induced compounds that are associated with genetic differences in susceptibility of *Pinus radiata* to bark stripping by marsupials. An untargeted profiling of 83 primary and secondary compounds of the needles and bark samples from seedlings of 21 susceptible and 21 resistant families was undertaken. These were among the most and least damaged families, respectively, identified within a trial of 74 families that were exposed to natural field bark stripping by marsupials. Experimental field plants were protected from bark stripping by marsupials and were subjected to artificial bark stripping treatment to examine induced chemistry compared to unstripped control plants which were also used to examine constitutive chemistry between resistant and susceptible families. Machine learning (random forest), partial least squares plus discriminant analysis (PLS-DA) and principal components analysis with discriminant analysis (PCA-DA) as well as univariate methods were used to identify the most important total compound groups and individual compounds differentiating the resistant and susceptible families. In the bark, the constitutive amount of two sesquiterpenoids - bicylogermacrene and an unknown sesquiterpenoid alcohol - were shown to be of higher levels in the resistant families, whereas the constitutive sugars; fructose and glucose as well individual phenolics were higher in the more susceptible families. The chemistry of the needles was not useful in differentiating the resistant and susceptible families. After artificial bark stripping, the terpenes, sugars and phenolics responded in both the resistant and susceptible families by increasing or reducing their amounts that levelled the differences in the amounts of the compounds between the different resistant and susceptible classes at the constitutive level. Overall, based on the families with extreme values for less and more susceptibility, differences in the amounts of secondary compounds were very subtle and susceptibility may outweigh defence in this non-native tree herbivory system.

4.1   Introduction

Host plant resistance may be influenced by constitutive and induced primary and secondary chemical compounds, as well as morphological, physical and life history characteristics (Carmona *et al.* 2011; Mumm and Hilker 2006). Identifying which plant traits drive differences in herbivore damage is a key challenge in understanding plant-herbivore interactions. In conifers, differences between herbivory resistant and susceptible genotypes have been associated with variation in constitutive physical traits including outer bark, exudation of oleoresin from wounds, resin ducts, specialized phloem parenchyma cells and induced anatomical responses (Franceschi *et al.* 2005; Whitehill *et al.* 2019). Chemical traits also play a role, where constitutive and induced terpenoids and phenolics have been identified as the most important secondary chemical groups in conifers (Keeling and Bohlmann 2006). Their roles as toxins, digestion inhibitors, deterrents, host recognition cues or precursors to physical defence systems in both the bark and needles has been documented for major herbivore groups ranging from insects to mammals (Farentinos *et al.* 1981; O'Reilly-Wapstra *et al.* 2007; Snyder 1992). However, studies show tremendous variation in the expression of these traits between species (Raffa *et al.* 2017), genotypes (Moreira *et al.* 2013a) and tissues (Chapter 3 ; Moreira *et al.* 2012a). The response to specific defences by different herbivores is also variable (Iason *et al.* 2011; Vourc'h *et al.* 2002b).

While the role of secondary compounds in defence of conifers has been a focus of previous studies, the role of primary compounds in directly differentiating susceptibility classes has not been as widely investigated (Clancy 1992; Zou and Cates 1994). Available evidence suggests that overall attractiveness to herbivores is determined by the balance between secondary and primary compounds in plants (Agrawal and Weber 2015; Kimball *et al.* 1998; Kurek *et al.* 2019; Snyder 1992). Sugars can for example be an attractant to herbivores (Felicijan *et al.* 2015; Kurek *et al.* 2019) although it has also been suggested that a high supply of sugars may increase resistance by enhancing the formation of terpenes, which have a high metabolic cost (Gershenzon 1994). The role of primary compounds as precursors to secondary compounds and as sources of energy and stored reserves for compensatory regrowth after damage is relatively well known, as is their signalling role (Goodsman *et al.* 2013; Raffa *et al.* 2017; Roth *et al.* 2018; Schwachtje and Baldwin 2008). Hence, to accurately predict variation in resistance, a detailed characterization of both the constitutive and induced primary and secondary chemical profile of a species is needed.

Despite our general understanding of chemical defences in conifers, most of the reported studies are associated with insects (Phillips and Croteau 1999). Little is known about how specific defence mechanisms act against mammalian herbivores. The impacts of defences may differ between

vertebrate and invertebrate herbivores reflecting differences in physiological adaptations, body sizes, herbivore-associated cues and the amount of tissues removed by the herbivore (Boyle 1999; Raffa 2014). In conifer–insect systems for example, monoterpenes have been shown to exert major defence roles (Raffa and Smalley 1995; Seybold *et al.* 2006) since they are exploited as primary chemical cues and/or precursors for the aggregation pheromones of insects. This may not be the case for mammalian herbivores that mostly rely on visual cues (Stutz *et al.* 2017). Also, in contrast to insect species that complete their life cycle in a single host, for example the well-studied conifer bark beetles (Schowalter 2012), mammalian bark browsers that readily adjust foraging behaviour on spatial and temporal criteria may be less affected by chemical defences (Gill 1992; Miller *et al.* 2014; Rea *et al.* 2014). While a few studies provide some evidence of the importance of conifer terpenes and phenolics on mammalian browsing (Farentinos *et al.* 1981; Snyder 1992), most instances of resistant traits have been circumstantial, with limited experimental evidence.

Pinus radiata is native to California but is a major plantation tree species in Australia that experiences significant bark stripping by kangaroos and wallabies (Miller *et al.* 2014; Page *et al.* 2013). In several Australian *P. radiata* family trials, variation in bark stripping was only partially associated with physical structures, suggesting that chemical traits may also be important (Chapter 2 ; Miller *et al.* 2014). While no specific compounds have been reported to differentiate resistant and susceptible genotypes for most *P. radiata* herbivores, the constitutive and induced chemical compounds in the bark and needles that can potentially be involved in defence have been documented (Chapter 3 ; Lundborg 2016; Moreira *et al.* 2013a; Moreira *et al.* 2012a; Reglinski *et al.* 2017). Page *et al.* (2013) also indicated that seasonal increase in *P. radiata* bark stripping by the marsupials may be driven by higher levels of sugars in the *P. radiata* bark relative to other winter forages, but no genetic-based intraspecific differences within *P. radiata* populations were documented. To my knowledge, whether primary compounds contribute to genetic variation in bark stripping has not been tested in conifer species.

The present study investigated the chemical compounds associated with family differences in the susceptibility of *P. radiata* to bark stripping. This study aimed to determine: 1) the constitutive compounds in bark and the needles that differentiate relatively least damaged (hereafter referred to as resistant) and relatively most damaged (hereafter referred to as susceptible) families; 2) how the bark and needle chemistry of the resistant and susceptible plants respond to artificial bark stripping and 3) whether the resistant and susceptible families can be classified based on the induced bark and needles chemistry.

4.2   Materials and methods

4.2.1   Field trials

The genetic field trial used for this study is the same decribed in Chapter 2. The trial was established at Wilmot in Tasmania, Australia (-41.454271^{0}N, 146.106801^{0}E, 580 m ASL) in 2015. Plant material was sourced from the New Zealand Radiata Pine Breeding Company. This material comprised 74 full-sib families that were planted in the field in an incomplete randomised block design of 26 replicates and 3 incomplete blocks within each replicate that were planted in a square of rows and columns. The families were represented in each replicate as single tree plots and four filler positions were planted with deployment seedlots. The families were from 55 parents and 54 grandparents that were selected for vigour, stem form and branch characteristics. The field trial was fenced to prevent browsing. In 2017 (25 months after planting), the gates were opened during winter for about two months to allow browsers access. To exclude browsers from some of the trees for the induction experiment (Experiment 2, below), six of the 26 replicates were further fenced. These six replicates are hereafter referred to as the 'protected replicates' and were spread randomly throughout the trial. The remaining 20 replicates were freely accessible to browsing for 2 months for characterisation of the variation in mammalian bark stripping between families to determine the relatively resistant and susceptible families. The major herbivore responsible for bark stripping in Tasmanian plantations of *P. radiata* is the Bennett's wallaby (*Macropus rufogriseus* subspecies *rufogriseus*). The density of the Bennett's wallaby within the mid-North of Tasmania where the genetic trial is situtated was estimated at 32.0 animals/km^2 (DPIPWE 2019).

4.2.2   Experiment 1: Characterization of resistant and susceptible families

At 25 months, in the 20 replicates that were accessible to browsers, the amount of bark removed from a tree relative to the stem circumference was scored on a scale of 0 - 5; 0 = no damage, 1 = <25%, 2 = 25 - 50%, 3 = 50 - 75%, 4 = >75%, 5 = 100% damage (completely ring barked). Except for scores = 0 and 100, the rest of the scores were converted to mid-point values for final analyses (Chapter 2). When trees exhibited bark stripping, the length of the strip was also scored on a scale of 1-7. where 1 = ~10 cm, 2 = ~ 20 cm and so on.The height of all trees in all replicates was also measured at the same time (height 1). At ~ 30 months, all 26 replicates were exposed to natural browsing, and ~2 months later, after this second episode of browsing another height measurement was taken (height 2).

To select the families that were more susceptible or less susceptible to mammalian bark stripping, best linear unbiased predictions (BLUPs) of family variation in bark stripping were estimated using linear mixed models described in Chapter 2. The general model is represented below

$$y = X\beta + Zu + e,$$

where y is a vector of the responses (bark stripping), β is a vector of fixed effects (overall mean), X and Z are design matrices associated with the fixed and random effects, u is the vector of random effects (replicates, incomplete blocks within replicates and family), and e is a vector of random residuals. Restricted maximum likelihood (REML) estimates of (co)variance parameters and their standard errors were obtained by using the average information REML algorithm implemented in the ASREML-R version 4.1 (Butler *et al.* 2009).

Based on the family BLUPs, 21 of the least damaged (resistant) and 21 of the most damaged (susceptible) families (42 families in total) were selected from the 74 families represented in the trial for further investigation. Families were chosen to maximise the differences in bark stripping BLUPs, but at the same time maintaining parental diversity in each population so that differences are not biased by a few parents. Therefore, some families with extreme values were not included if the parents were already well-represented. In total, 17 and 22 parents were represented in the selected resistant and susceptible families respectively, indicating a diverse representation of genotypes in each group.

4.2.3   Experiment 2: Chemistry experimental design and phytochemical analysis

Three weeks after the mammalian bark stripping assessment in Experiment 1, an experiment to assess constitutive and induced chemical differences between the susceptible and resistant families was initiated using all the individuals in the six protected replicates (n=393 trees). Half of the individuals sampled were subject to artificial bark stripping (treated trees) and half were untreated and used as controls. To apply the treatment, alternate plants in the six replicates were systematically treated regardless of family. The tree at one corner of each replicate was selected as a control tree, the next one in the column was selected for treatment and this pattern was consistently followed across the six replicates of the trial. The artificial bark stripping treatment (T0) was applied by removing a vertical strip of 15 cm of bark, starting 2 cm above the ground, and covering 30% of the stem circumference (Figure 1). The dimensions were selected based on the most common browsing level observed in Experiment 1. Three weeks after treatment (based on time of the biggest induced response observed in Chapter 3), bark and needle samples were collected from all control and treated trees. This bark sample was of similar size to the initial strip removed. It was taken from a similar height on both control and treated trees, being ~1 cm above the first strip on the treated trees, and was the bark strip used in the chemical analysis (Figure 4.1). All bark samples were collected from the north side of the stem. Needles were also sampled from each tree, ensuring mixture of young and old needles mostly at the top of the crown.

Samples were kept in a cool box until transportation to a -20ºC freezer where they were stored until chemical extraction.

After 3 weeks, a second strip of bark was removed from trees selected for treatment. The difference between the comparable samples from treated and untreated trees gave estimates of induced chemistry

After 3 weeks, a strip of bark was removed from control trees at the same height as treated trees. Constitutive chemistry was quantified from control samples

At T0, a strip of bark was removed from trees selected for treatment

Figure 4.1: Artificial bark stripping treatment and position of bark samples taken from 2-year old *Pinus radiata* in the Wilmot field trial.

Chemical extractions were performed on only the 42 families identified as resistant and susceptible in Experiment 1. For each of the 42 families, two random samples were selected for extraction; one from a control and the other from a treated individual. Chemical extractions (targeting terpenes, phenolics and sugars) were undertaken separately from the needles and bark from each tree according to the methods described in Chapter 3 with the exception that (i) the dichloromethane (DCM) extraction was done in 5mL using 75mg of fresh material, and for the acetone extraction, nonadecanoic acid was used instead of rutin as an internal stardard at a concentration of 0.25mg/L. The sugars were extracted using hot water as previously described (Chapter 3). An extra sample of each plant part from each indivual was weighed, dried in the oven at 110°C for 72 hrs and reweighed to convert samples from wet to dry weights. The percentage water content was also expressed as the difference between the wet and dry weight of the sample divided by the wet weight. The extracts were stored at -18˚C until analysis. The DCM extracts were analysed by gas chromatography-mass spectrometry (GC-MS) while the acetone extracts and the sugars were analysed by ultra-high-performance liquid chromatography-mass spectrometry (UHPLC-MS). The GC-MS and UHPLC-MS conditions, as well as the procedure for identification and quantification of peaks, is detailed in Chapter 3. In addition to individual compounds, the total amounts of monoterpenoids, GC-MS diterpenoids, sesquiterpenoids, phenolics and sugars were derived by summing the respective individual compounds that belonged to each category (referred to as "total compound groups"). All quantified compounds were given a unique identifier (Supplementary Table 10) for ease of identification in the tables.

4.2.4   Statistical analysis

Univariate and/or multivariate statistical methods are routinely used to rank and select the most important chemical compounds. Univariate methods such as analysis of variance consider each variable independently in contrast to multivariate approaches that consider several or all variables simultaneously, evaluating the joint distribution of some or all variables and estimating their relevance to the observed variation (Vinaixa *et al.* 2012). The most widely used multivariate methods for chemical data including the unsupervised principal component analysis and random forest, the supervised partial least squares- discriminant analysis, as well as univariate techniques, have been subject to detailed reviews (Chen *et al.* 2013; Saccenti *et al.* 2014). Random forest is a machine learning technique that uses multiple independent decision trees that are trained independently on a random subset of data to predict another set of independent samples (Breiman 2001). In understanding plant defences, multivariate techniques could be valuable since defence is usually achieved by multiple interdependent traits (Carmona *et al.* 2011). However, compared to univariate techniques, the implementation of multivariate techniques is computationally demanding and often difficult to interpret, and sometimes both techniques may give different results depending on parameterisation (Saccenti *et al.* 2014). Therefore, the implementation of both multivariate and univariate data analysis is strongly recommended to maximize the extraction of relevant information (Saccenti *et al.* 2014).

To identify the most important total compound groups and individual compounds that differentiated the resistant and susceptible families based on constitutive (control samples) and induced (treated samples) bark and needle chemistry, the multivariate techniques - random forest (RF) (Breiman 2001), partial least squares plus linear discriminant analysis (PLS-LDA) (Preda *et al.* 2007) and principal component analysis-discriminant analysis (PCA-DA) (Jombart and Collins 2015) - were used. Random forest analysis was done using the ranger package (Wright and Ziegler 2015) while PLS-DA and PCA-DA were executed using the pls and pca packages, respectively. All analyses were done in R v.3.6.0 (R Core Team 2018). Classification by the three methods was evaluated and compared through a 20 - fold cross-validation using a common standard error. The models were assessed according to the accuracy of classification, where accuracy was defined as the percentage of correct predictions to the total number of input samples. These models were fitted in two phases:

(i)   Estimation of model accuracy: this was done fitting separate models for each subsample of the data pertaining to specific pairwise comparisons. For comparing resistant versus susceptible families four such pair-wise comparisons were undertaken for each sample type: (1) bark-constitutive chemistry, (2) bark-induced chemistry, (3) needle-constitutive chemistry

and (4) needle-induced chemistry. Separate analyses were done using the total compound groups and individual compounds resulting in eight analyses.

(ii)   Variable importance: to select the variables of importance in differentiating the susceptible and resistant families, the chemistry from all subsamples were included as separate variables (individual needles and bark, total compound groups, constitutive and induced samples) in a single model (termed 'combined analysis'). In this case, the same chemical compound and total compound group in each of the four sample types above were treated as separate variables in the model. This was done to allow variable importance (i.e the relative contribution to classification accuracy) to be compared across all subsamples (1- 4 above) and compound type (individual compounds and total compound groups). All measures of importance were scaled to have a maximum value of 100, where 0 means that the variable does not contribute to the classification and 100 signifies the most important variable across all tissues (bark and needles) and treatments (constitutive and induced) and compound type. Five total compound groups (monoterpenoids, GC-MS diterpenoids, sesquiterpenoids, phenolics and sugars) were ranked by the models and are presented, but for the individual compounds only the 10 most important variables in either of the multivatiate methods are presented.

Univariate analysis used the Kruskal–Wallis one-way analysis of variance (KW) in R v. 3.6.0 (R Core Team 2018) to test the differences in the amounts of the selected compounds between the resistant and susceptible families since the normality assumption did not hold for all compounds. No Bonferroni correction was applied. Arithmetic means and standard error (se) for the constitutive amounts of compounds in the bark and needles of the resistant and susceptible families were also calculated in R v. 3.6.0 (R Core Team 2018).

To identify the most important total compound groups and individual compounds that responded to the artificial bark stripping treatment, the three multivariate methods; RF, PLS-DA and PCA-DA were used to identify the variables of importance for the resistant and susceptible families. To estimate model accuracy, separate models for each subsample of the data pertaining to specific pairwise comparisons were fitted. Pair-wise comparisons of the chemistry of control versus treated plants were undertaken for: (1) the bark of resistant families, (2) the bark of susceptible families, (3) the needles of resistant families and (4) needles of susceptible families. Separate analyses were done using the total compound groups and individual compounds resulting in eight analyses. A combined analysis as described above, that included all samples was run so that the importance for all comparisons is

comparable. For the selected compounds, the amounts of compounds quantified in the control and treated samples were compared using Kruskal–Wallis (KW) one-way analysis of variance mentioned above and absolute p-values are reported.

Differences between the resistant and susceptible families were also tested for the relative water content and tree height using Kruskal–Wallis (KW) one-way analysis of variance mentioned above. The height differences between the resistant and susceptible families and associated standard errors were tested separately for the six protected replicates and 20 unprotected replicates using Kruskal–Wallis (KW) one-way analysis of variance. Absolute p-values were reported for differences in water content and height.

4.3   Results

4.3.1   Identifying resistant and susceptible families based on field bark stripping

Bark stripping by the marsupials occurred on 52% of the 1370 trees assessed in the 20 unprotected replicates. Overall, the amount of bark removed ranged from 0 to 100% of the stem circumference and averaged 49% of amount of damage when considering only the bark stripped trees. The average percentage of bark removed from the 21 most susceptible families ($\bar{x}$ = 31.76 ± 1.56 %) was significantly higher than for the 21 most resistant families ($\bar{x}$ = 18.35 ± 1.22 %) (KW, X^2 = 27.30, p < 0.001). On the damaged plants, the length of the strip of bark removed was almost two-fold longer in the susceptible ($\bar{x}$=1.43 ± 0.07 cm) than in the resistant ($\bar{x}$ = 0.81 ± 0.05 cm) families (KW, X^2 = 28.65, p < 0.001).

For the unprotected trees in the 20 replicates, the susceptible families were significantly taller ($\bar{x}$ = 150.70 ± 1.41 cm) than the resistant families ($\bar{x}$ = 144.63 ± 1.49 cm) (KW, X^2 = 4.59, p < 0.05) for the first height measurement taken at the time of browsing assessment, suggesting a positive relationship between bark stripping and height. The height difference was in a similar direction in the six protected replicates, but not significant (susceptible $\bar{x}$ = 168.75 ± 1.32 cm, resistant $\bar{x}$ = 162.13 ± 1.26 cm; KW, X^2 = 2.53, p = 0.11), possibly in part due to the reduced sample size. For the second height measurement, 6 months after exposure to bark stripping, the height of the susceptible and resistant families in the 20 browsed replicates increased by 40% and 39% respectively but did not differ (KW, X^2 = 2.82, p = 0.09) suggesting a reduction in growth rate for the susceptible families following bark stripping. The percentage water content of the bark of resistant ($\bar{x}$ = 66.88 ± 0.67 %) and susceptible families ($\bar{x}$ = 66.58 ± 0.50 %) did not differ (KW, X^2 = 0.23, p = 0.63).

4.3.2 LC-MS and GC-MS analysis of compounds in the bark and needles

Chemical assessment of terpenes, phenolics and sugars in the bark and needles of *Pinus radiata* detected a total of 83 compounds or compound groups. Of these, 14 monoterpenoids, 20 diterpenoids, 18 phenolic compounds, 5 sesquiterpenoids, 5 sugars and 3 fatty acids were fully or partially identified while 18 compounds remained unidentified. Some compounds were found in either the bark or needles and consequently 65 compounds were identified in the bark and 65 compounds in the needles. The means of all the constitutive compounds in the bark and needles of resistant and susceptible families are listed in Supplementary Table 4.1.

4.3.3 Classification of resistant and susceptible families based constitutive bark chemistry

(i)   Model accuracy

Using constitutive bark chemistry (control samples), classification into resistant or susceptible families was associated with moderate accuracy, where 61 - 68% of the individuals were correctly classified based on total compound groups and 62 - 72% of the individuals were correctly classified based on individual compounds (Table 4.1). Partial least squares plus linear discriminant analysis (PLS-LDA) had the highest accuracy for most comparisons compared to principal component analysis-discriminant analysis (PCA-DA) and random forest (RF).

Table 4.1: Accuracy (%) associated with the classification of individuals into resistant and susceptible classes based on constitutive (control) total compound groups and individual compounds in the bark for three multivariate methods. Total compound groups were derived by summing the individual compounds in each category (see methods). The multivariate methods were partial least squares plus linear discriminant analysis (PLS-LDA), random forest (RF) and principal component analysis-discriminant analysis (PCA-DA). Accuracy was defined as the percentage of correct predictions to the total number of samples tested. Identification of resistant and susceptible families based on field bark stripping assessments

	Total compound groups	Individual compounds
PLS-DA	68	72
RF	61	70
PCA-DA	65	62

(ii)   Important variables

Total compound groups

Based on constitutive chemistry (control samples) in the bark, the three multivariate models ranked total phenolics as the most important compound group classifying or differentiating the resistant and susceptible families, followed by sugars (Figure 4.2).

Figure 4.2: Ranking of total compound groups according to their importance in the models in differentiating the resistant and susceptible families based on constitutive bark chemistry. The importance for each compound was ranked according the relative contribution to the accuracy of classifications in a combined analysis, where 0 = variable was not important and 100 = variable was most important. PLS = partial least squares – linear discriminant; RF = random forest; PCA = principal component – discriminant analysis.

The univariate tests of the constitutive levels of total compound groups showed that susceptible families were 2 - fold higher in total constitutive phenolics than the resistant families (KW, $X^2 = 7.04$, $p<0.01$) (Table 4.2, Figure 4.3). No significant differences were detected for the other total compound groups (i.e. sugars, mono-, sesqui-, diterpenoids and $C_{20}H_{30}O_2$ resin acids; Supplementary Table 4.1).
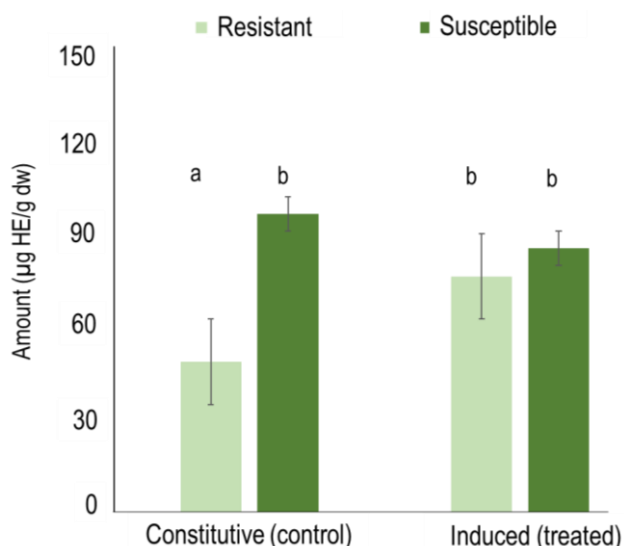
Figure 4.3: Mean ± standard error of the constitutive (control) and induced (treated) amount of total phenolics for the resistant and susceptible families. Total phenolics were derived by summing the individual compounds. The amount of phenolics are expressed as micrograms of heptadecane equivalents (HE) per gram of dry weight of

the sample. The Kruskal-Wallis test indicated that the four means were significantly different ($p < 0.05$) and means with different letters are significantly different at the at the $p < 0.05$ level, based on Dunn's test.

Individual compounds

Using constitutive bark chemistry (control samples), for individual compounds, the three multivariate models identified the following primary (sugars and fatty acids) and secondary compounds (phenolics and terpenoids) as most important in differentiating the resistant and susceptible families (Figure 4.4):

  – two sugars – fructose [76] and glucose [77];

  – one fatty acid - linolenic acid [84];

  – three phenolics - phenyl ethanol [67], trans-ferulic acid [72] and benzene acetic acid [57];

  – a monoterpenoid - citronellal [6];

  – two sesquiterpenoids – bicyclogermacrene [17] and an unknown sesquiterpenoid alcohol [21];

  – two diterpenoids - unknown $C20H32O3$ B [48], unknown $C20H30O6$ D [55]; and

  – a compound of unknown group - unknown m/z 104 [86].

The constitutive levels of glucose, benzene acetic acid and bicyclogermacrene in the bark were ranked as the most important of all variables respectively by the random forest (scoring 100) , partial least squares (scoring 91) and by principal component analysis (scoring 100) (Figure 4.4).
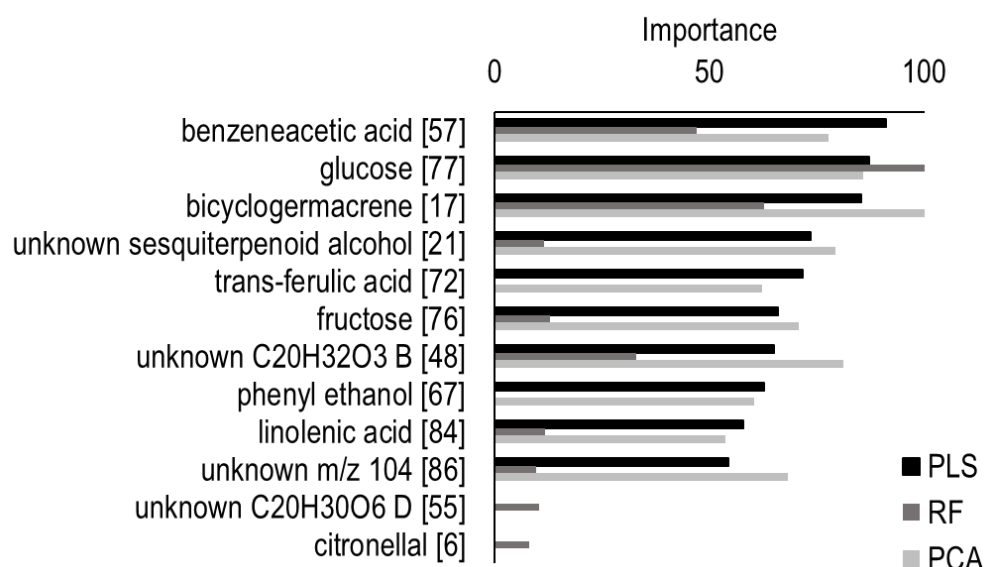
Figure 4.4: The most important compounds selected to differentiate the resistant and susceptible families for the constitutive (control) bark chemistry (i.e. those among top 10 in each analysis are presented). The importance for each compound was ranked according the relative contribution to the accuracy of classifications in the combined analysis, where 0 = variable not important and 100 = variable was most important. PLS = partial least squares –

linear discriminant; RF = random forest; PCA = principal component – discriminant analysis. The numbers in parentheses are identifiers given to each compound for ease of location in the tables.

Univariate tests for the constitutive amounts of the most important individual compounds in the bark showed that the terpene compounds bicyclogermacrene [17] and the unknown sesquiterpenoid alcohol [21] had significantly higher amounts in the resistant families (Figure 4.5, Table 4.2). Both compounds were more than 3-fold higher in the resistant than susceptible families. The amounts of trans-ferulic acid [72], phenyl ethanol, [67] citronellal [6], linolenic acid [84] and the unknown C20H30O6 D [55] did not significantly differ between the resistant and susceptible families, while the amounts of the remaining compounds were higher in the susceptible families. Fructose [76] and glucose [77], for example were 1.2 times higher in the susceptible than in the resistant families (Figure 4.5, Table 4.2). The unknown diterpenoid [48], benzene acetic acid [57] and the unknown compound [86] were respectively 2.9, 2.4 and 2.1 -fold higher in the susceptible than the resistant families.
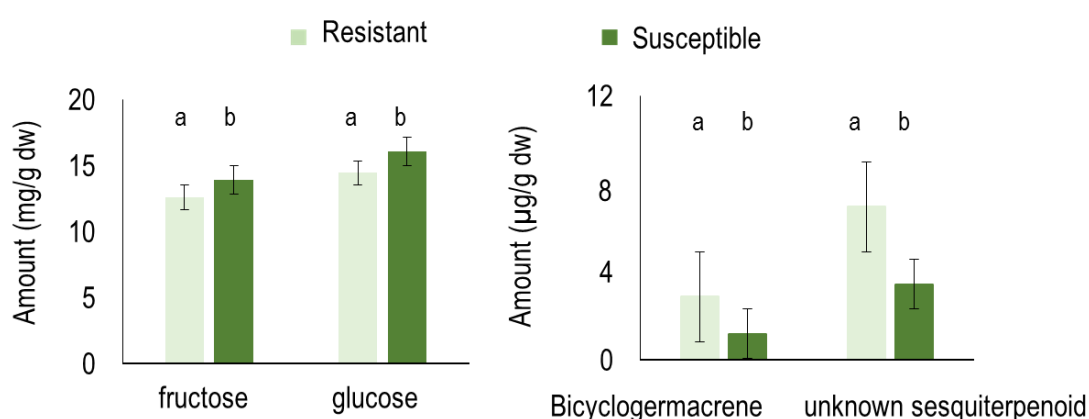
Figure 4.5: Mean ± standard error of constitutive sugars (fructose and glucose) that were higher in the susceptible families and sesquiterpenes (bicyclogermacrene and unknown sesquiterpenoid alcohol) that were higher in resistant families. The sugars are expressed in absolute amounts. The amount of terpenes are expressed as micrograms of heptadecane equivalents (HE) per gram of dry weight of the sample. Means with different letters within a compound are significantly different between resistant and susceptible families at the $p < 0.05$ level the based on Kruskal-Wallis pair-wise comparisons.

Table 4.2: Mean ± standard error (se) of the constitutive amounts of individual compounds and total compound groups in the bark identified to be important in differentiating resistant [R] and susceptible [S] families (Figures 4.2, 4.4). Unadjusted Kruskal-Wallis p-values are indicated. The amount of terpenoids [T] and phenolics [P] are expressed as micrograms of heptadecane equivalents per gram of dry weight of the sample. The amounts of fatty acids [F] and unknown compounds [U] are indicated as micrograms of nonadecanoic acid equivalents per gram of dry weight of the sample. The sugars [S] are expressed in absolute amounts (µg/g dw). The unadjusted p-value of the difference in the amounts between resistant and susceptible families is based on Kruskal Wallis test. Resistant and susceptible families were identified based on field bark stripping assessments. All compounds were given a unique identifier based on Supplementary Table 10 (after Chapter 9), for ease of identification

ID	Compound	Group	Constitutive amount in resistant families ($\overline{x}$ ± se)	Constitutive amount in susceptible families ($\overline{x}$ ± se)	P-value
6	citronellal	T	60.97 ± 10.94	52.42 ± 20.28	0.170
7	citronellic acid	T	22.27 ± 3.75	19.38 ± 2.82	0.753
17	bicyclogermacrene	T	3.73 ± 0.85	1.20 ± 0.40	0.003
21	unknown sesquiterpenoid alcohol	T	9.17 ± 2.26	3.43 ± 1.14	0.019
48	unknown C20H32O3 B	T	468.08 ± 59.41	1368.2 ± 451	0.017
55	unknown C20H30O6 D	T	3739.48±576.56	4712.55±767.28	0.148
57	benzene acetic acid	P	23.76 ± 4.63	57.4 ± 10.61	0.023
67	phenyl ethanol	P	1.88 ± 0.86	8.86 ± 3.57	0.052
72	trans-ferulic acid	P	40.05 ± 5.99	65.7 ± 9.99	0.068
76	fructose	S	13538.19 ± 895.89	16159.82 ±948.77	0.037
77	glucose	S	15197.83 ± 917.81	18736.28 ± 894.95	0.011
84	linolenic acid	F	6808.79 ± 369.06	7849.75 ± 470.88	0.116
86	unknown m/z 104	U	1.28 ± 0.39	2.71 ± 0.78	0.045
	total phenolics	P	48.28±6.37	95.96±13.16	0.008
	total sugars	S	41728.11±2247.20	47802.10±2215.30	0.085

4.3.4   Classification of resistant and susceptible families based constitutive needle chemistry

(i)   Model accuracy

The classification of families into resistant and susceptible based on constitutive needle chemistry (Table 4.3), was associated with less accuracy compared to the models based on bark chemistry

(Table 4.1). The accuracy for the needles never exceeded 59% which indicates that needle chemistry is not an accurate predictor of bark damage.

Table 4.3: Accuracy (%) associated with the classification of individuals into resistant and susceptible classes based on constitutive (control) total compound groups in the needles for three multivariate methods. Total compound groups were derived by summing the individual compounds in each category (see methods). The multivariate methods were partial least squares plus linear discriminant analysis (PLS-LDA), random forest (RF) and principal component analysis-discriminant analysis (PCA-DA). Accuracy was defined as the percentage of correct predictions to the total number of samples tested. Identification of resistant and susceptible families based on field bark stripping assessments

	Total compound groups	Individual compounds
PLS-DA	46	48
RF	38	59
PCA-DA	33	44

(ii)　　Important variables

Total compound groups

Using constitutive needle chemistry**,** the contribution of the total compound groups to the models was very low but the total sugars followed by sesquiterpenoids were the most important in differentiating the resistant and susceptible families with an importance value of <40 (Figure 4.6). The univariate tests showed no significant differences between the resistant and susceptible families in the constitutive amounts of any total compound groups including the sugars (Supplementary Table 4.1).
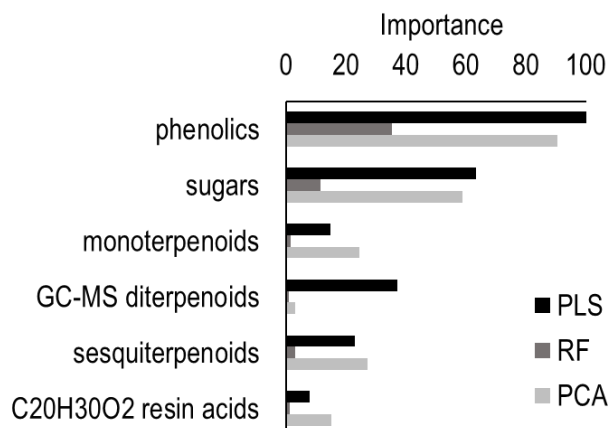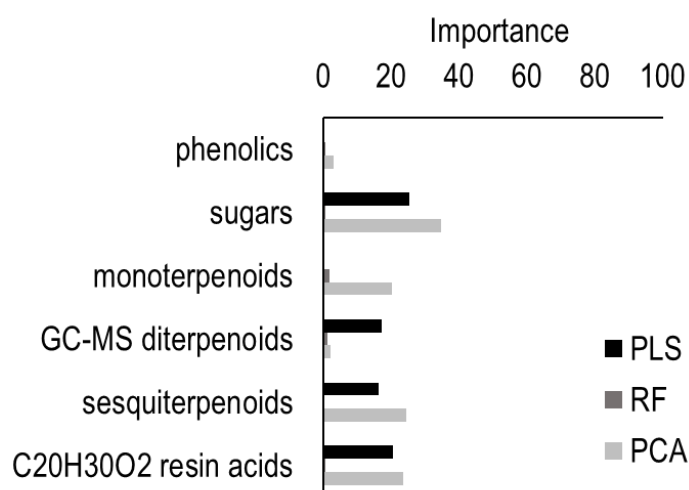

Figure 4.6: Ranking of total compound groups according to their importance in the models in differentiating the resistant and susceptible families based on constitutive needle chemistry. The importance for each compound was ranked according the relative contribution to the accuracy of classifications in a combined analysis, where 0

= variable was not important and 100 = variable was most important. PLS = partial least squares – linear discriminant; RF = random forest; PCA= principal component – discriminant analysis.

Individual compounds

Based on constitutive needle chemistry, the most important individual compounds that differentiate the resistant and susceptible families were topped by compounds of unknown groups [97][86] , an unknown diterpenoids [48] and a sugar – inositol [78] but these never exceeded the importance value of 75 (Supplementary Figure 4.1). There was limited overlap of the important compounds for different models. The important compounds based on constitutive needle chemistry were mostly different from those that were most important in the bark (Supplementary Figure 4.1). Some of the selected compounds were unique to the needles (i.e. sesquiterpenoid-caryophyllene[18], phenolic-chavicol[58], unknown diterpenoid - $C20H30O6$ B[53] and compounds of unknown group- unknown m/z 358[92] and unknown m/z 740 B[97]). This shows that important compounds selected either in the needles or the bark reflect the uniqueness of chemistry in each plant part.

Univariate tests on individual compounds showed no differences between the resistant and susceptible families in constitutive needle chemistry (Supplementary Table 4.1) except for the unknown m/z 104[86]. The constitutive amount of this compound in the needles was two-fold higher in the susceptible than in the resistant families (KW, X^2 = 4.7, p = 0.03) (Supplementary Table 4.1).

4.3.5   Differential induction of resistant and susceptible families to artificial bark stripping

To understand how the resistant and susceptible families responded to treatment, bark and needle chemistry from treated and control samples was compared separately for the resistant and susceptible families.

(i)   Model accuracy

Models for classifying the treated from untreated samples were associated with relatively high accuracy in both the susceptible and resistant families based on bark chemistry (maximum of 77%) but accuracy was lower based on the needle chemistry (maximum of 67%) indicating limited systemic effects of bark stripping (Table 4.4).

Table 4.4: Accuracy (%) associated with the classification of individuals into control or treated based on total groups in the bark and the needles of resistant and susceptible families for three multivariate methods. Total compound groups were derived by summing the individual compounds in each category (see methods). The multivariate methods were partial least squares plus linear discriminant analysis (PLS-LDA), random forest (RF) and principal component analysis-discriminant analysis (PCA-DA). Accuracy was defined as the percentage of correct predictions to the total number of samples tested. Identification of resistant and susceptible families based on field bark stripping assessments

	Resistant		Susceptible	
	Total compounds	Individual	Total compounds	Individual
Bark chemistry				
PLS-DA	70	77	71	62
RF	68	60	71	69
PCA-DA	72	65	73	66
Needle chemistry				
PLS-DA	67	63	44	37
RF	40	48	48	53
PCA-DA	60	41	36	44

(ii)   Important variables

Total compound groups

Multivariate models showed that in the bark, the total phenolics were more important in differentiating treated and control samples in the resistant families followed by monoterpenoids (Figure 4.7a). However, the importance levels never exceed 65 and were rarely consistent in the three multivariate analyses. By contrast, in the susceptible families, the sugars consistently ranked as the most important variables (scoring 100 for all the three models) and the total phenolics were of little importance, highlighting differences in responses to treatment between the resistant and susceptible families (Figure 4.7b). In the needles, the importance of all total compound groups was low, consistent with no/weak responses of needle chemistry to treatment (Figure 4.7c, d). Even then, sugars and phenolics were the most important compound groups in differentiating treated samples.

Univariate comparisons for the total compound groups showed that artificial bark stripping caused a significant increase in total phenolics and a decrease in the total monoterpenoids in the bark of resistant families (Table 4.5). In the bark of susceptible families, bark stripping caused a decrease in

total sugars (Table 4.5) supporting a difference in induced responses between resistant and susceptible families. No significant changes were detected in other total compound groups in the bark or any total compound group in the needles.

Individual compounds

The individual compounds that were important in differentiating treated from control samples in the bark and the needles of resistant and susceptible families are shown in Supplementary Figure 4.2. For both the bark and the needles, the top compounds identified by each model were mostly individual phenolics and sugars suggesting that these were more responsive to treatment. The selected compounds were mostly different in the resistant and susceptible families highlighting differences in responses to treatment.

Univariate tests for individual compounds showed a significant reduction in the amount of fructose and glucose in the bark (p <0.001), but only in the susceptible families resulting in levels comparable to the resistant families (Figure 4.8, Table 4.5). Phenyl ethanol [67] and benzene acetic acid [57] marginally increased in the treated bark of the resistant but not the susceptible families. In the bark, the terpenoid compounds - α-pinene [1], citronellal [6] and citronellol [8] reduced in the resistant families, while the sugar inositol [78] reduced in both resistance classes (Table 4.5). Changes in the needles that denote systemic responses to artificial bark stripping were only marginally significant (p <0.05) and involved an increase in linolenic acid [84] (R) and a decrease in coniferyl alcohol [59] (R and S) and trans-ferulic acid [72] (R) (Table 4.5).
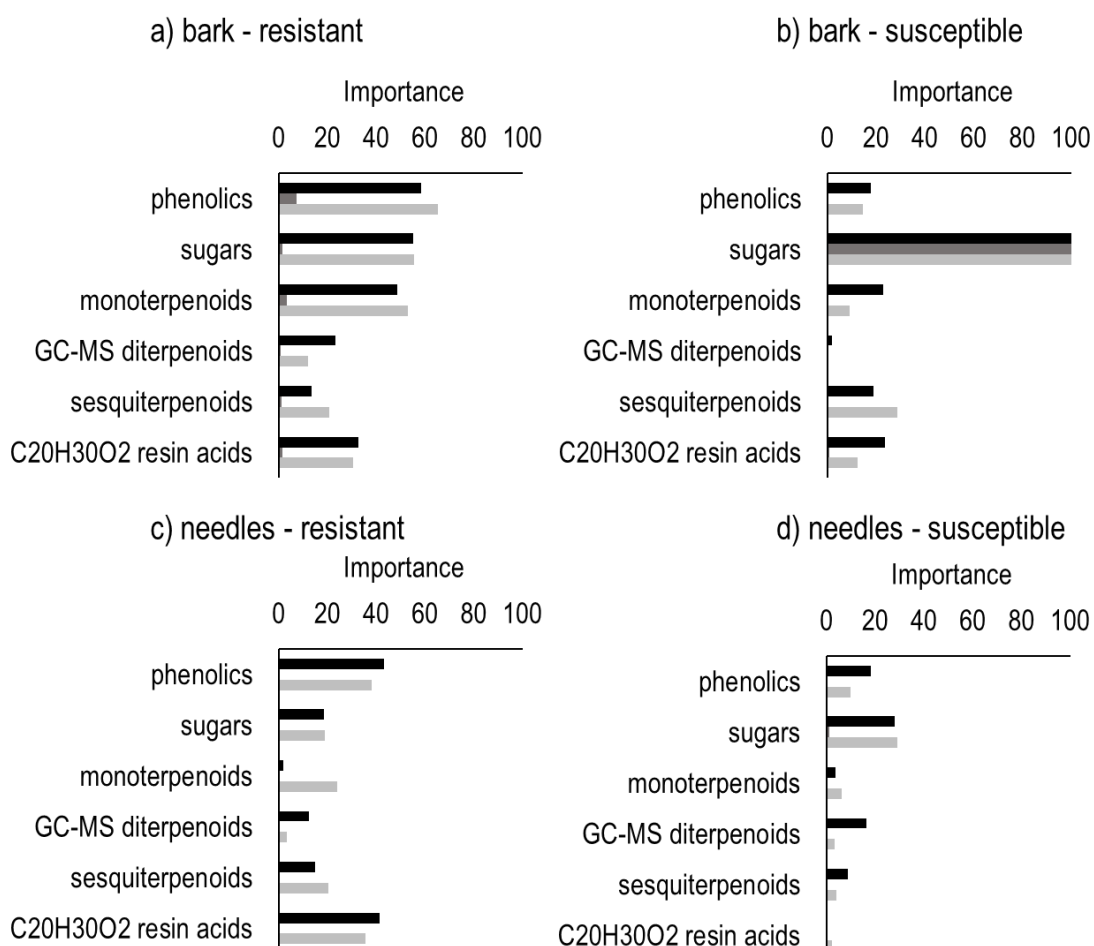
Figure 4.7: Ranking of total compound groups according to their importance in differentiating the control and treated plants in the resistant and susceptible plants in the a) bark and b) needles. The importance for each compound was ranked according the relative contribution to the accuracy of classifications in the combined analysis, where 0 = variable not important and 100 = variable was most important. PLS = partial least squares – linear discriminant; RF = random forest; PCA = principal component– discriminant analysis.
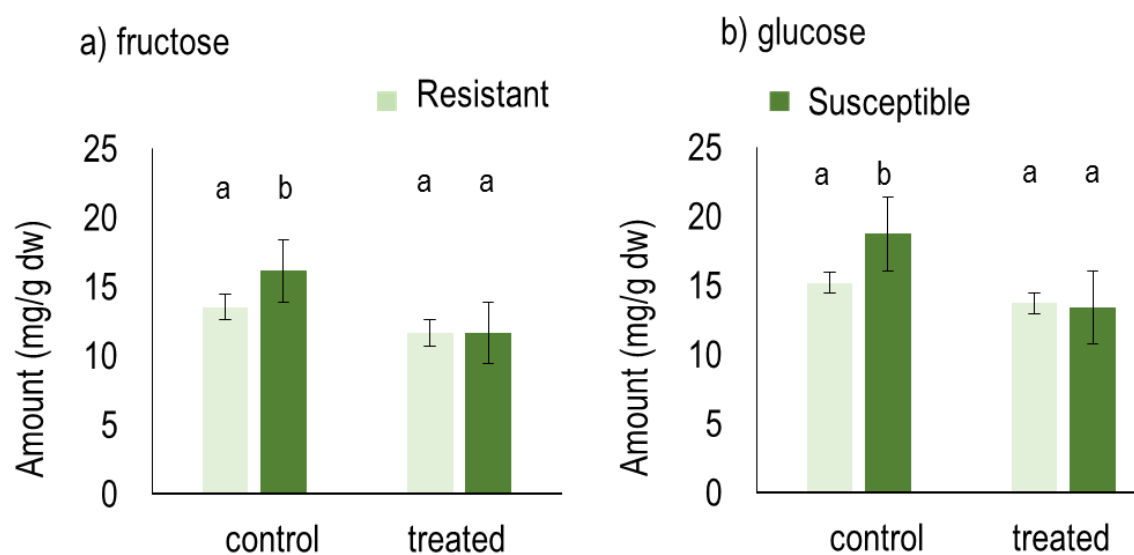
Figure 4.8: Mean ± standard error of constitutive amount of sugars; (a) fructose and (b) glucose that were higher in the susceptible families before treatment (control) but were reduced to levels in the resistant families after treatment. The sugars are expressed in absolute amounts. Different letters on each graph indicate significant differences in mean amounts at the $p < 0.05$ level (unadjusted) based on Kruskal-Wallis tests of the four panels on each graph

Table 4.5: Compounds and total compound groups in the bark [B] and needles [N] that responded to the artificial bark stripping treatment in the resistant [R] and susceptible [S] families. The Kruskal-Wallis (KW) chi-square value (x^2) and unadjusted p-values for the difference between the constitutive and induced amounts of each compound are shown, as well as the direction of change in the bark stripped treatment compared to the control (+ = increase, - = decreased after treatment). Compounds that were selected by the models but did not show any significant changes after treatment based on KW are not listed. Significance was set at $p < 0.05$. Identification of resistant and susceptible families based on field mammalian bark stripping assessments. All compounds were given a unique identifier based on Supplementary Table 10 (after Chapter 9), for ease of identification

Identifier	Compound	KW chi-square	P-value (unadjusted)	Direction	Part	Category
	Terpenoids					
1	α-pinene	5.2	0.021	-	B	R
6	citronellal	7.5	0.006	-	B	R
8	citronellol	4.6	0.031	-	B	R
	total monoterpenoids	4.3	0.038	-	B	R
43	unknown m/z 304 B	4.9	0.028	+	B	S
48	unknown C20H32O3 B	6.7	0.010	-	B	S
	phenolics					
57	benzene acetic acid	4.6	0.031	+	B	R
59	coniferyl alcohol	4.3	0.038	+	B	R
		6.1	0.010	-	N	R
		4.0	0.040	-	N	S
65	methyl eugenol	8.3	0.004	+	B	S
67	phenyl ethanol	4.4	0.035	+	B	R
72	trans-ferulic acid	4.4	0.031	-	N	S
	total phenolics	6.5	0.011	+	B	R
	sugars					
76	fructose	12.3	0.000	-	B	S
77	glucose	14.4	0.000	-	B	S
78	Inositol	4.6	0.031	-	B	R
		5.8	0.016	-	B	S
	total sugars	15.9	0.000	-	B	S
	fatty acids					
84	linolenic acid	5.4	0.020	+	N	R

4.3.6 Classification of resistant and susceptible families based in induced chemistry

(i)   Model accuracy

To test whether the resistant and susceptible families could be classified based on induced bark and needle chemistry (treated samples), comparatively lower accuracy was obtained for models developed with total compound groups and individual compounds (maximum = 58) (Table 4.6) compared to models based on constitutive chemistry (maximum = 72) (Table 4.1). Although the results above have shown responses to treatment of the resistant and susceptible families, these induced changes diminish the chemical differences between resistant and susceptible families and consequently explain the low accuracy associated with the classification.

Table 4.6: Accuracy (%) associated with the classification of individuals into resistant and susceptible classes based on induced (control) total compound groups and individual compounds in the bark and the needles for three multivariate methods. Total compound groups were derived by summing the individual compounds in each category (see methods). The multivariate methods were partial least squares plus linear discriminant analysis (PLS-LDA), random forest (RF) and principal component analysis-discriminant analysis (PCA-DA). Accuracy was defined as the percentage of correct predictions to the total number of samples tested. Identification of resistant and susceptible families based on field mammalian bark stripping assessments

	Induced bark chemistry		Induced needle chemistry	
	Total compound groups	Individual compounds	Total compound groups	Individual compounds
PLS-DA	45	58	51	47
RF	37	50	52	42
PCA-DA	32	34	57	45

(ii)   Important variables

Total compound groups

For induced bark chemistry, total monoterpenoids were the most important in differentiating susceptible and resistant families but their importance was lower compared to the classification based on constitutive chemistry (Figure 4.8). Total phenolics were more important based on the induced needle chemistry. In both cases the importance values did not exceed 63.
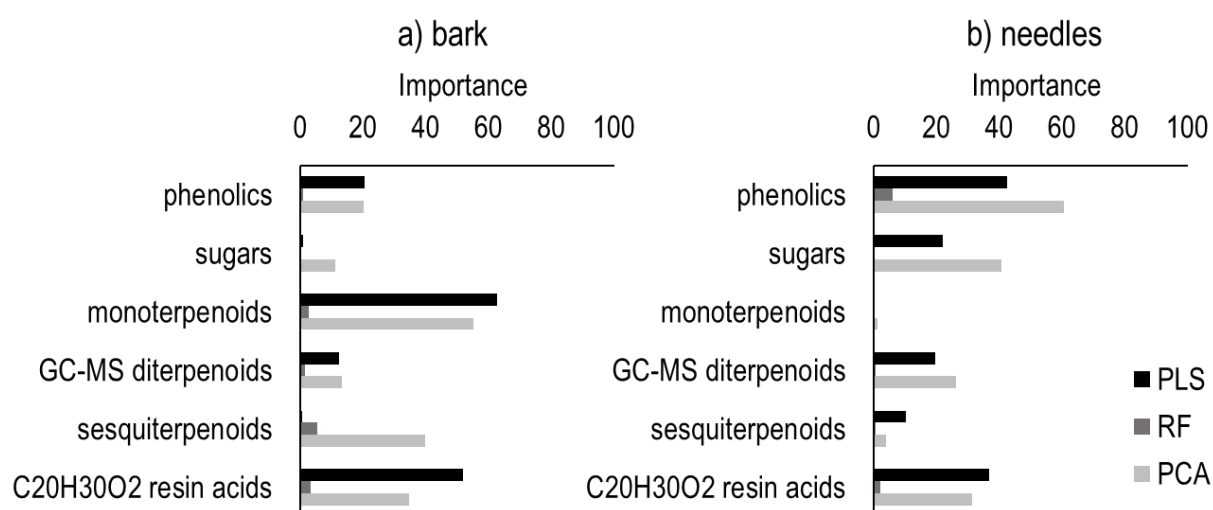
Figure 4.8: Ranking of total compound groups according to their importance in differentiating the resistant and susceptible families based on induced (a) bark and (b) needle chemistry (treated). The importance for each compound was ranked according to the relative contribution to the accuracy of classifications in the combined analysis, where 0 = variable not important and 100 = variable was most important. PLS = partial least squares – linear discriminant; RF = random forest; PCA = principal component– discriminant analysis.

Individual compounds

The important individual compounds in differentiating the resistant and susceptible families based on treated samples (induced bark and needle chemistry) were similarly summarised (Supplementary Figure 4.3). In the bark, the monoterpenoid- β-pinene [4] and the phenolic compound - methyl eugenol [65] were on top of the selected compounds. The important individual compounds based on induced bark chemistry were largely different from those based on the constitutive bark chemistry except for the phenyl ethanol [67], benzene acetic acid [57], unknown $C20H30O6$ D [55] and unknown m/z 104 [86]. Similarly, in the needles, the important compounds were dominated by a sesquiterpenoid-caryophyllene [18]. However various phenolic compounds as well as sugars were important. Only a proportion of compounds [18],[59],[67],[84],[86],[92],[97] identified as important based on induced needle chemistry were represented in the models developed for constitutive chemistry (Supplementary Figure 4.3). The important compounds identified by each of the multivariate analyses were generally non-overlapping indicating randomness of the selections consistent with the low accuracy of the associated models (Table 4.6).

Univariate tests on individual compounds showed no differences between the resistant and susceptible families in the induced chemistry of both the bark and the needles (KW test; data not shown) confirming that resistant and susceptible families are not well differentiated on induced chemistry.

4.4   Discussion

Overall, three major findings emerged from this study. Firstly, constitutive chemistry is more important in classifying families that are resistant and susceptible to bark stripping than induced chemistry. The most important constitutive chemical groups or compounds in differentiating resistant and susceptible families were: (a) the total phenolics, as well as individual phenolic compounds that were higher in susceptible families; (b) the sugars- glucose and fructose were higher in the susceptible families, and (c) the sesquiterpenoids that were higher in resistant families. Secondly, induced bark chemistry was not important for the classification of the resistance classes, but differential induction was detected between the resistant and susceptible families, which may have implications on the attractiveness of the plants after the initial damage. Thirdly, needle chemistry differed little between resistance classes and barely responded to the bark stripping treatment. The differences between the resistant and susceptible family groups in our study reflect genetic-based differences as families are randomly distributed within multiple replicates of a common-garden field trial.

Family variation in bark stripping damage is associated with differences in both the primary and secondary compounds. Individual sugars especially glucose and fructose found in the bark were identified as important predictors of susceptibility to bark stripping. Sugars being linked to herbivory is not surprising since the phloem is nutrient rich and sought by many herbivores. However, evidence of genetic-based intraspecific variation in herbivory being related to sugars has not been demonstrated in many conifers to date. In *P. radiata*, Page *et al.* (2013) attributed bark stripping to higher sugars in the bark compared to surrounding forages, although no intraspecific differences were studied. Other studies have also correlated differences in mammalian damage to differences in the sugars and other nutritional and mineral elements in coniferous (Tamura and Ohara 2005) and non-coniferous trees (Kurek *et al.* 2019; Saint-Andrieux *et al.* 2009) but not at the genetic level. The lack of genetic studies involving sugars could be related to the premise that traits coupled tightly with fitness show low or no genetic variation (Mousseau and Roff 1987) and thus the accumulation of non-structural carbon compounds (NSC); including glucose and fructose, in non-photosynthetic plant parts has been viewed as a passive sink–source process resulting from imbalances between carbon supply and demand (Wiley and Helliker 2012). However, the demonstration of genetic differences in sugar accumulation in *P. radiata* suggests that storage of NSC compounds is an actively regulated process and identifying genes that regulate the sugar allocation process may further enhance the understanding of bark stripping. Overall, since photosynthesis is the primary determinant of crop productivity, determining the genetic variation in physiological mechanisms underlying the variation in growth and storage of photosynthates between the resistant and susceptible families in *P. radiata* may further our

understanding of variation in bark stripping. The results of the present study contrast with studies which suggest that high levels of sugars in plant tissues may directly or indirectly enhance plant resistance (Clancy 1992; Morkunas and Ratajczak 2014; Zou and Cates 1994). In *Pteudotsuga menziesii* for example sucrose provided higher resistance against the budworm (Clancy 1992). Sugars have also been indirectly associated with defence as a source of energy or precursors to secondary compounds (Schwachtje and Baldwin 2008).

Constitutive levels of total and some individual phenolics, especially benzene acetic acid were also predictors of susceptibility, with their amount higher in the susceptible families, consistent with observations in *Pinus sylvestris* (Sunnerheim-Sjöberg and Hämäläinen 1992). Although phenolics are largely considered to be defensive (Franceschi *et al.* 2005), evidence for the association of phenolics with mammalian browsing in conifers is inconsistent. While various studies have shown no effect of phenolics on needle and bark browsing (Hansson *et al.* 1986; Sauvé and Cǒté 2007), a number of studies have indicated that phenolics may increase resistance (Lindroth and Batzli 1984; Radwan 1972; Radwan and Crouch 1978). The association of phenolics with susceptibility in this study points to the possibility that phenolics may have an attractant rather than a defensive effect, especially in low concentrations. The phenolic compounds, benzene acetic acid and phenyl ethanol, that have higher levels in the bark of the susceptible families, have diverse medicinal properties (Bredsdorff *et al.* 2015; Madan *et al.* 2016) and may be beneficial to the marsupials. On bark stripping in *P. radiata*, Smith *et al.* (2020) for example suggested that non-preferred food sources, such as bark in small quantities, form part of a mixed diet, which may have positive effects on nutrition and digestion, as predicted by the nutrient balance hypothesis (Westoby 1978). The association of phenolics with susceptibility could also be related to the herbivores in question. Australian marsupials are reported to possess the capability to ingest and metabolise a range of secondary metabolites such as phenols and terpenes that would be toxic to many other herbivore species (Boyle 1999; El-Merhibi *et al.* 2007). Consistently, the present results provided little support for the role of monoterpenoids and diterpenoids in defence against marsupial bark stripping, despite various studies implicating these groups of compounds as defences against mammalian herbivores for both bark (Bucyanayandi *et al.* 1990; Pederson and Welch 1985; Snyder 1992; Zhang and States 1991) and needles (Iason *et al.* 2011; Vourc'h *et al.* 2002a).

While constitutive levels of individual mono and di-terpenoids were not important in differentiating resistant and susceptible families, specific sesquiterpenoids (bicyclogermacrene and an unknown sesquiterpenoid alcohol) in the bark were shown to have significantly higher levels in resistant families, suggesting that they could be defensive against bark stripping. Bicyclogermacrene has been also

identified in the needles and bark of many conifer species (Amri *et al.* 2013; Lebouvier *et al.* 2013), evidence of a defence role has not been well demonstrated except in *Pinus sylvestris* where it was higher in the branches that were not attacked by the pine processionary moth (Achotegui-Castells *et al.* 2013). While the specific defence roles of sesquiterpenes against herbivores remain largely unknown, their potential exploitation as precursors for insect hormones and as signals to attract insect parasitoids in tri-trophic interactions has been suggested for insect herbivores (Celedon and Bohlmann 2019). Even then, the differences in the amounts of sesquiterpenes between the resistant and susceptible families in this study were marginal suggesting that resistance could result from synergistic or additive interactions between different compounds rather than the effect of individual compounds (Scalerandi *et al.* 2018). Synergy has been demonstrated in *P. pinaster* where, genetic variation in resistance against the pine weevil was largely explained by the multivariate concentration and a blend of secondary metabolites, rather than by bivariate correlations with individual compounds (López-Goldar *et al.* 2018). Similar observations were made in *Spodoptera litura* (tobacco cutworm) where mixtures of trans-anethole plus thymol, citronellal and α-terpineol had an almost ten times stronger effect on the mortality rate than would have been the case with simply an additive effect (Hummelbrunner and Isman 2001). Such interactions may need further investigation in conifers, and describing of the amount of bark eaten by the marsupials as a function of the amount of these secondary metabolites, whether linear or curvilinear as described by Moore and DeGabriel (2012) may also need further study.

Resistant and susceptible families were distinguishable based on constitutive but not the induced bark chemical compounds. This is due to differential responses to treatment, highlighting the differences in preparedness to herbivory between the resistant and susceptible families. While the resistant families are more guarded constitutively, the susceptible families seem to respond strongly to percieved herbivory to attain a chemical profile similar to the resistant families. A key finding from the study is the rapid reduction of bark glucose and fructose levels in treated susceptible plants to levels comparable with the resistant families. If sugars are a major driver of bark stripping, this observation suggests that after the initial damage the likelihood of further stripping may be minimised by the reduced sugar content in the bark. This is similar to observations by Moreira *et al.* (2013b) on the resistance of *P. radiata* to the processionary moth where, damage to plants genetically differed significantly among pine families in the constitutive condition, but this genetic variation disappeared after methyl jasmonate induction (Moreira *et al.* 2013b). Even then, the moth consumed 20% less phloem in the treated plants but with no genetic variation of damage between families (Moreira *et al.* 2013b), but these observations were not associated with sugars. These results, however, contrasted the observation in *P. pinaster* where induced chemistry was more important in predicting differences in susceptibility to the pine

weevil (López-Goldar *et al.* 2018), potentially highlighting species specific herbivory responses. For bark stripping mammals, however, this will depend on how fast (in hours) the sugar reduction occurs and also how long the effects persist. Although earlier results (Chapter 3) indicated that sugar levels remained low for up to three weeks after the initial bark stripping, this time progression especially early reduction of the sugar levels needs to be understood in view of repeated browsing that has been observed in the field. It is also possible that chemical responses depend on the amount of bark removed from the trees such that trees that receive low intitial bark stripping may have lower chemical responses which may cause repeated bark stripping. On the other hand, ring barking may also cause an accumulation in needles and bark above the ring barked area as demonstrated in girdling studies (Li *et al.* 2003), and this may also increase the bark stripping. Therefore, the relationship between amount of bark removed and the chemical responses may need further investigation.

In addition to the changes in the sugars described above, an induction of mainly the phenolics was also evident, with a general reduction in the amounts in the needles and an increase in the amounts of the bark. In the constitutive chemistry, phenolics were higher in the susceptible than the resistant families. After treatment, the phenolics reduced non-significantly in the susceptible families but increased significantly in the resistant families. This opposing response may suggest that phenolics play different physiological roles in stress responses of the resistant and susceptible families. In the resistant families, the increase in phenolics may augment resistance to herbivory (Roitto *et al.* 2009) as an adaptation to the pests and pathogens in the native habitats (Mead 2013). In the susceptible families however, the roles of phenolics may be more related to inducible stress cell division – used as precursors to replace lost tissues; and regulation of photosynthetic activity – aimed to reduce the amount of sugars (Hammerbacher *et al.* 2011). Studies that have examined the defence properties of a larger number of conifer phenolic compounds have also indicated that the role of phenolics is specific, where some are defensive while others have been associated with susceptibility (Danielsson *et al.* 2011; Ganthaler *et al.* 2017)

Defence-growth trade-offs were also implicated by the study. Although the height difference between the resistant and susceptible families was not significant in the families where chemistry was assayed (protected replicates), possibly due to small sample size, we noted an overall reduction in height in resistant families in the unprotected families with larger sample size. The higher amount of sesquiterpenoids in more resistant, slower growing families is consistent with the defence-growth trade-off theory that suggests accumulation of defence compounds in slow growing trees (Ferrenberg *et al.* 2015). Although theories explain the trade-offs in terms of resource allocation (Coley *et al.* 1985), the

patterns of intraspecific variation observed in this study, where preference for constitutive investment in some families and the simultaneous preference for induction in other families in the same environment do not clearly fit the theoretical predictions. In this case, genetics is a strong factor in explaining the intraspecific chemical variation and therefore, examining how genetics interacts with the environment to shape chemical variation is important for further understanding of the evolution of metabolites and their inclusion in *P. radiata* breeding programmes. However, the premise of reduced growth following increased storage of non-structural carbohydrates like glucose (Wiley and Helliker 2012) is inconsistent with the results of this study that showed a positive relationship between amounts of sugars and height, where susceptible families that had more sugars in the bark grew faster than the resistant families.

Multivariate and univariate techniques generally depicted the same patterns in terms of ranking the compounds. Univariate techniques also have extra information such as the amount of the compounds and the direction of the effect. Multivariate techniques were associated with high accuracy where the absolute differences in amounts of compounds were larger. Comparing the multivariate techniques, the premise that non-linear machine learning methods such as random forest (RF) provide superior generalised predictive ability for metabolomics data when compared to linear alternatives, such as partial least squares discriminant analysis (PLS-DA) was not supported. Various studies on plant metabolomics show that random forest has performed poorer than PLS (Lee *et al.* 2018; Mendez *et al.* 2019). However, it has been suggested that non-linear machine learning algorithms may perform better with larger data sets (Gromski *et al.* 2015; Mendez *et al.* 2019). The linear methods could also be better in analyzing highly collinear and noisy data (Gromski *et al.* 2015). Therefore, linear methods like PLS may be recommended where small samples are involved. Specifically, however, random forests can tend to overfit some data distributions consequently reducing performance of the model. Therefore, defining the most optimum model parameters may improve the performance of random forests (Gromski *et al.* 2015).

4.5   Conclusion

Variation in bark stripping is associated with consititutive differences in sesquiterpenoids, sugars and phenolics, and would appear to be an interplay between defence and attractant compounds of the bark. Specific bark sesquiterpenoids are better predictors of constitutive resistance to marsupial damage than other terpenoid groups. Phenolics and particularly sugars predict increased susceptibility both as total compound groups and for specific compounds. Induced chemistry did not directly differentiate between resistant and susceptible families, suggesting the main variation in bark stripping was due to consitutive chemistry of the bark. The main induced response involved the susceptible plants becoming

more like the resistant plants in their chemistry – particularly the sugars. Studies on the mechanisms and function of selected individual compounds, including possible synergies, are required to further our understanding of *Pinus radiata* defences against mammalian herbivores. Understanding the genetic architecture of the traits will enable their selection in breeding programmes.

Supplementary Table 4.1: Constitutive amounts of compounds (± se) in the resistant [R] and susceptible [S] families. The amounts of terpenes and phenolics were expressed as micrograms of heptadecane equivalents per gram of dry weight of the sample. The amounts fatty acids and unknown compounds are indicated as micrograms of nonadecanoic acid equivalents per gram of dry weight of the sample. The sugars are expressed in absolute amounts (µg-1 dw). The amounts in bold were significantly different between susceptible and resistant families based on the Kruskal-Wallis test. The bold values show that amounts were different and significance levels were set at p<0.05. The p-values are indicated in Table 4.1. The absence of a value indicates that the compound was not detected. Ethyl phenol and 4-ethyl guaiacol were detected only in the bark of susceptible families at the constitutive level. All compounds were given a unique identifier based on Supplementary Table 10 (after Chapter 9), for ease of identification. The identifiers in this table are not sequential as some compounds indicated in Supplementary Table 10 (after Chapter 9) were not identified in this data set

		Bark		Needles	
	monoterpenoids	R	S	R	S
1	α-pinene	985.31±131.09	828.12±139.46	27.72±5.67	30.58±5.10
2	α-terpineol	32.4±4.33	32.1±4.78	5.44±0.79	4.84±0.54
3	β-phellandrene	89.85±16.17	98.28±17.04	7.58±1.58	9.31±1.72
4	β-pinene	1939.23±217.40	1797.65±234.75	91.63±23.20	83.7±12.36
5	camphene	8.73±1.06	7.67±1.13	0.36±0.08	0.38±0.05
6	citronellal	60.97±10.94	52.42±20.28		
7	citronellic acid	22.27±3.75	19.38±2.82	0.45±0.07	0.41±0.06
8	citronellol	67.05±10.31	49.41±9.64	0.54±0.12	0.45±0.08
9	γ-terpinene	6.29±1.96	7.91±3.39	0.16±0.04	0.19±0.04
10	limonene	77.58±18.71	58.02±9.76	10.9±3.54	13.93±2.87
11	linalool	8.87±2.05	8.79±2.45	0.08±0.03	0.06±0.03
13	sabinene	179.4±72.71	225.58±102.93	0.25±0.06	0.64±0.40
14	terpinene-4-ol	28.47±7.25	38.64±15.72	0.96±0.12	1.06±0.18

15	terpinolene	94.2±41.67	94.56±42.83	0.28±0.05	0.32±0.06
16	unknown mol Wt 150	3.04±0.61	4.43±1.45	0.06±0.01	0.09±0.03
	sesquiterpenoids				
17	bicyclogermacrene	**3.73±0.85**	**1.2±0.40**	1.31±0.28	1.59±0.28
18	caryophyllene			3.25±0.83	4.73±0.98
19	γ-elemene			0.86±0.19	0.99±0.18
20	trans-farnesol	92.81±70.18	25.00±5.20	4.91±0.61	4.08±0.57
21	unknown sesquiterpenoid alcohol	**9.17±2.26**	**3.43±1.14**	3.13±0.72	3.46±0.67
	GC-MS diterpenoids				
22	agathadiol	452.89±73.71	690.47±195.59	60.71±8.30	76.9±18.07
23	agatholal	309.76±46.34	439.13±104.22	50.7±6.14	53.47±9.40
24	copalol	39.34±7.47	39.72±8.76	13.3±2.40	11.17±2.38
25	levopimaral	13.52±2.54	11.2±1.76		
26	methyl dehydroabietate	12.58±1.87	11.64±1.24	1.28±0.40	1.27±0.20
27	methyl levopimarate			0.8±0.21	0.99±0.31
	LC-MS diterpenoids				
29	dehydroabietic acid	24442.49±2162.37	25673.56±2050.45	11221.69±734.77	11627.06±721.21
32	unknown diterpene-3	242.23±54.01	177.22±44.46		
33	unknown m/z 109 A	17.77±2.32	16.94±2.53	0.96±0.25	1.19±0.37
34	unknown m/z 109 B	21.19±4.98	20.01±3.01	1.38±0.30	1.95±0.46
37	unknown m/z 239	6.78±1.53	6.26±0.76	1.86±0.29	2.1±0.32
38	unknown Mol Wt 272	15.77±8.30	7.68±1.13	1.33±0.19	1.69±0.30
39	unknown C20H30O2 A			3030.39±507.80	3308.57±465.02

#	Compound	Col1	Col2	Col3	Col4
40	unknown C20H30O2 B			8420.81±435.08	8449.47±453.92
41	C20H30O2 resin acids	25396.23±1573.16	24917.33±1290.34	18677.6±1770.91	20015.39±1360.72
42	unknown m/z 304 A	144.11±22.31	117.95±17.00		
43	unknown m/z 304 B			1730.46±117.66	1918.6±148.71
45	unknown m/z 316	14661.66±1655.38	13732.81±1365.94		
46	unknown C20H30O3	27227.17±2944.18	27003.59±2404.41	9456.52±971.45	9719.52±775.34
47	unknown C20H32O3 A	18633.86±1907.07	23443.21±2894.35	22973.88±2201.10	22799.7±2474.18
48	unknown C20H32O3 B	**468.08±59.41**	**1368.2±451**	1263.68±371.03	2791.28±1002.14
49	unknown C20H32O3 C			20267.68±2147.54	18611.66±2203.79
50	unknown C20H30O4	61811.5±5566.7	64458.76±4990.16	39497.05±3365.43	41769.93±3218.07
51	unknown C20H30O5	11957.42±1604.43	12153.43±1334.87		
52	unknown C20H30O6 A	215.77±27.32	210.97±27.81		
53	unknown C20H30O6 B			7000.45±903.77	6881.16±677.54
54	unknown C20H30O6 C	6527.7±804.83	7454.21±898.88		
55	unknown C20H30O6 D	3739.48±576.56	4712.55±767.28		
	phenolics				
56	anethole/estragole	1.79±0.20	1.92±0.29	2.13±0.59	1.67±0.50
57	benzene acetic acid	**23.76±4.63**	**57.4±10.61**	77.36±12.26	75.62±10.71
58	chavicol			6±2.33	5.79±1.32
59	coniferyl alcohol	4.64±0.66	4.54±0.42	1.6±0.23	1.93±0.26
60	eugenol			2.52±0.79	3.5±1.11
61	ethyl phenol		0.53±0.44		
63	4-ethyl guaiacol		3.86±3.36		
64	isoeugenol			4.63±1.08	4.64±1.07

65	methyl eugenol	0.18±0.04	0.17±0.04		
66	p-Menth-1-en-7,8-diol	12.23±0.98	11.36±1.05		
67	phenyl ethanol	1.88±0.86	8.86±3.57	3.37±0.48	4.53±0.96
68	pinosylvin dimethyl ether	2.73±0.66	5.42±2.77		
69	piperitone			1.8±0.41	2.14±0.75
70	raspberry ketone	7.11±1.15	10.87±2.05	3.49±0.53	3.12±0.32
71	thymol	9.09±1.47	8.01±1.7	1.01±0.34	0.72±0.14
72	trans-ferulic acid	40.05±5.99	65.7±9.99	1.47±0.26	1.26±0.15
74	vanillin	4.47±0.31	4.33±0.27	0.65±0.09	0.75±0.12
75	zingerone			9.7±1.23	8.35±0.70
	sugars				
76	fructose	**13538.19 ± 895.89**	**16159.82 ±948.77**	23265.12±1278.92	24232.10±1413.42
77	glucose	**15197.83 ± 917.81**	**18736.28 ± 894.95**	25374.16±1520.41	26730.87±1369.46
78	inositol	12382.38±1215.36	12247.40±980.40	7820.02±301.22	8425.61±334.02
79	sucrose	2.83±0.01	2.06±0.00	90.53±19.98	80.53±16.70
80	unknown disaccharide			276.55±62.58	240.48±57.50
81	unknown monosaccharide	606.87±97.57	656.52±134.88		
	fatty acids				
82	linoleic acid	16356.86±982.02	16809.72±985.13	11255.66±711.71	12481.71±823.26
83	linolenic acid	6808.79±369.06	7849.75±470.88	27692.04±1123.94	29992.07±1410.38
84	palmitic acid	16222.82±823.76	15832.8±764.60	17526.74±615.29	18267.19±723.25
	unknowns				
85	unknown m/z 104	**1.28±0.39**	**2.71±0.78**	**0.27±0.10**	**0.49±0.13**
88	unknown m/z 272			781.77±58.6	742.04±69.59

89	unknown m/z 274	1131.95±124.29	1273.18±121.01		
90	unknown m/z 358			1194.28±156.75	941.52±109.98
91	unknown m/z 362			1134.06±128.18	1022.41±139.97
92	unknown m/z 406 A	523.01±57.35	490.91±46.55		
93	unknown m/z 406 B	5111.61±439.65	5900.64±569.80		
95	unknown m/z 740 B			2109.65±163.91	2755.18±363.27
	compound groups				
	total monoterpenoids	3615.87±390.91	3334.34±515.22	146.39±34.22	145.98±20.08
	total GC-MS diterpenoids	828.08±123.17	1192.17±307.02	126.79±16.39	143.81±29.29
	total sesquiterpenoids	153.14±69.43	105.28±13.24	16.54±2.31	18.04±2.08
	total phenolics	**48.28±6.37**	**95.96±13.16**	115.91±13.35	115.58±12.37
	total sugars	41728.11±2247.20	47802.10±2215.30	56826.38±2771.34	59709.59±2681.51

1
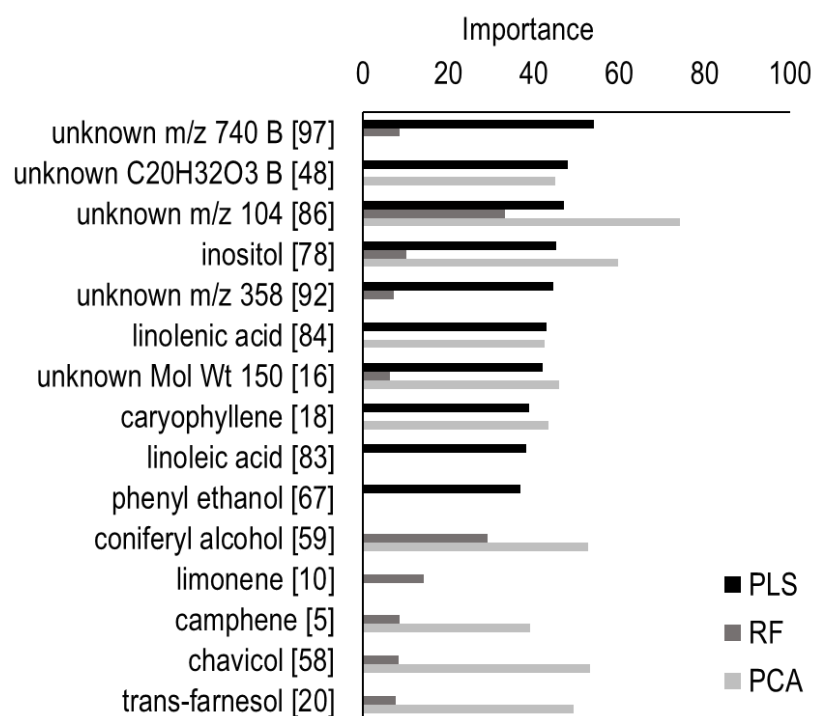
Supplementary Figure 4.1: The most important compounds selected to differentiate the resistant and susceptible families based on constitutive (control) needle chemistry. The importance for each compound was ranked according the relative contribution to the accuracy of classifications in the combined analysis, where 0 = variable not important and 100 = variable was most important. PLS = partial least squares – linear discriminant; RF = random forest; PCA = principal component – discriminant analysis.

a) bark - resistant

Importance

b) bark - susceptible

Importance

Legend: PLS, RF, PCA

c) needles - resistant

Importance

Compound	
unknown C20H30O6 B [53]	
linolenic acid [83]	
coniferyl alcohol [59]	
unknown m/z 362 [93]	
vanillin [74]	
benzene acetic acid [57]	
unknown C20H30O4 [50]	
unknown C20H32O3 A [47]	
unknown C20H32O3 C [49]	
trans-farnesol [20]	
unknown Mol Wt 150 [16]	
phenyl ethanol [67]	
zingerone [75]	
unknown m/z 272 [89]	

d) needles - susceptible

Importance

Compound	
trans-ferulic acid [72]	
coniferyl alcohol [59]	
vanillin [74]	
unknown C20H32O3 B [48]	
unknown m/z 109 A [33]	
unknown m/z 109 B [34]	
methyl levopimarate [27]	
trans-farnesol [20]	
unknown m/z 239 [37]	
unknown C20H32O3 C [49]	
glucose [77]	
unknown C20H30O3 [46]	
phenyl ethanol [67]	
agatholal [23]	
unknown m/z 304 A [42]	
zingerone [75]	
unknown C20H32O3 A [47]	
fructose [76]	
unknown m/z 272 [89]	
chavicol [58]	
linalool [11]	

■ PLS
■ RF
■ PCA

Supplementary Figure 4.2: The most important compounds selected to differentiate the control and treated plants in the bark of resistant (a) and susceptible (b) families and in the needles of resistant (c) and susceptible (d) families resistant and susceptible families. The importance for each compound was ranked according the relative contribution to the accuracy of classifications in the combined analysis, where 0 = variable not important and 100 = variable was most important. PLS = partial least squares – linear discriminant; RF = random forest; PCA = principal component – discriminant analysis.

118
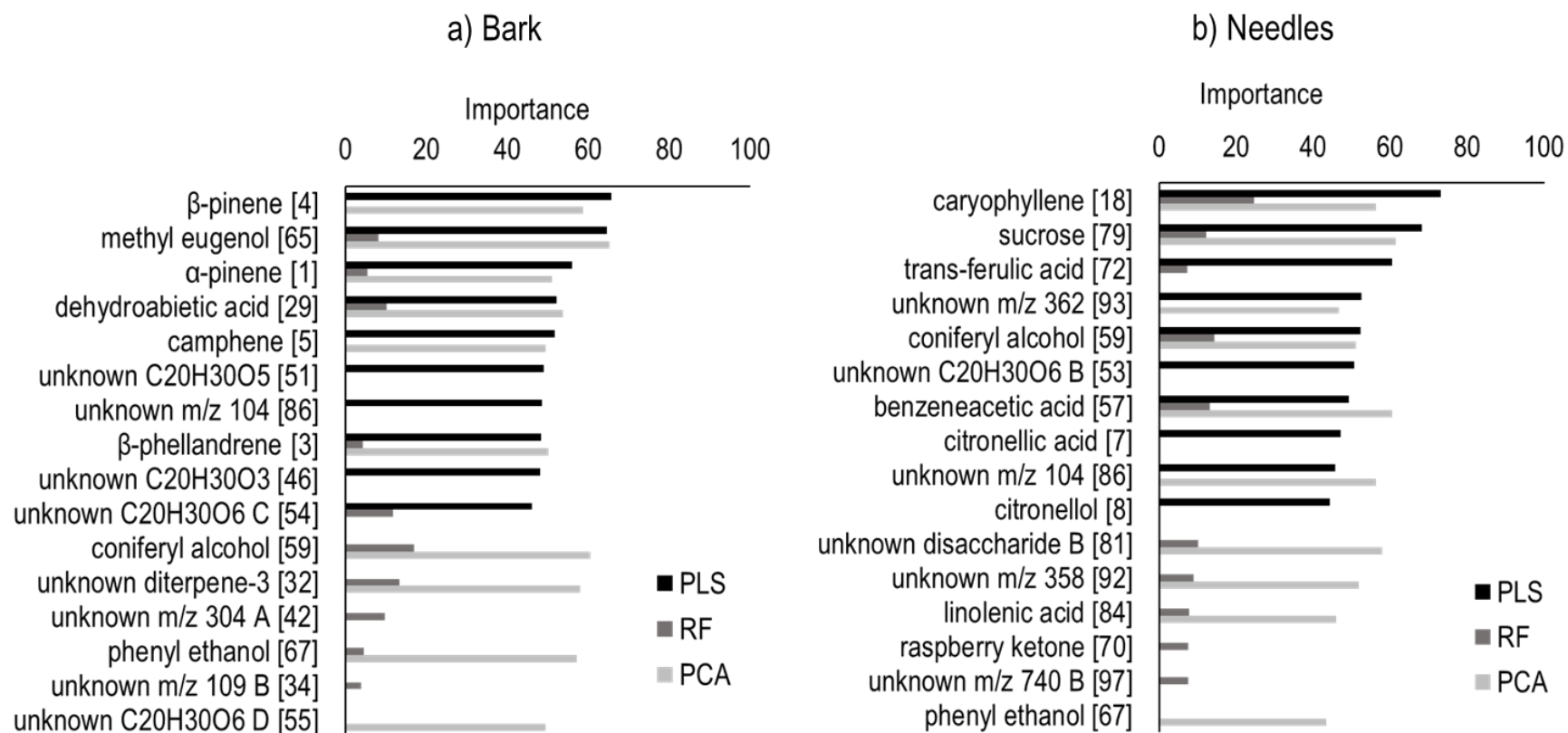
Supplementary Figure 4.3: The most important compounds selected to differentiate the resistant and susceptible families based on induced (treated) a) bark and b) needle chemistry. The importance for each compound was ranked according the relative contribution to the accuracy of classifications in the combined analysis, where 0 = variable not important and 100 = variable was most important. PLS = partial least squares – linear discriminant; RF = random forest; PCA = principal component– discriminant analysis.

CHAPTER 5: Developing near infrared spectroscopy (NIRS) models for predicting chemistry and responses to stress in *Pinus radiata* (D. Don)

5.0　　Abstract

The incorporation of important chemical traits in breeding programmes requires the estimation of genetic parameters especially the levels of additive genetic variation and this requires a large number of samples from pedigreed populations. Conventional wet chemistry procedure for chemotyping are slow and expensive and are not a practical option. In this study, the near infrared spectral properties of the needles, bark and roots before and after exposure to stress treatments were investigated as an alternative approach. The aim was to test the capability of near infrared (NIR) spectroscopy to (i) discriminate samples exposed to artificial bark stripping and methyl jasmonate assessed 7, 14, 21 and 28 days after treatment from untreated samples and (ii) quantitatively predict individual chemical compounds in the three plant parts. Using principal components analysis (PCA) on the spectral data, we were able to differentiate between treated and untreated samples for the individual plant parts. Based on partial least squares-discriminant analysis (PLS-DA) models, the best discrimination of treated from non-treated samples with the smallest root mean square error of prediction was achieved in the fresh needles (RMSECV, 0.24, R^2 = 0.81) and fresh inner bark (R^2 = 0.79, RMSE = 0.25) for methyl jasmonate treated samples after 14 days and 21 days respectively after treatment. For the strip treatment, the smallest error of prediction was achieved 21 days after treatment (R^2 = 0.69, RMSE = 0.30) in the bark. Calibrations developed for individual chemical compounds correlating the NIR spectra with chemistry data using partial least squares (PLS) regression gave models with high coefficient of determination for fructose (R^2 = 0.84, RPD = 1.5, PRL = 0.71, RER = 7.25) and glucose (R^2 = 0.83, RPD = 1.9, PRL = 1.14, RER = 8.50) as well as two unknown diterpenoids (both R^2 = 0.72, 0.75, RER = 10.76, 11.29). Overall, the results indicated the ability to separate stressed and unstressed plants with NIR spectroscopy and to accurately predict glucose, fructose and diterpenoids levels and to provide proximate values for several fatty acids, mono- and diterpenoids and phenolics, which provides an optimistic outlook for the use of NIRS-based models for the larger-scale prediction of the *P. radiata* chemistry needed for quantitative genetic studies.

5.1 Introduction

The use of natural resistance to minimize herbivory is a desirable alternative to existing costly and less effective techniques. To achieve this necessitates genetic selection of resistant genotypes but this requires phenotyping large sample sizes to gain sufficient power and resolution (Mir *et al.* 2019; Singh *et al.* 2019). In conifer-herbivory relationships, resistance has been attributed to physical properties and secondary chemical compounds especially the constitutive and induced terpenes and phenolics (Franceschi *et al.* 2005; Moreira *et al.* 2012a). Recent studies have also indicated the role of primary chemical compounds as influencing herbivory – either as attractants or indirectly contributing to defence (Goodsman *et al.* 2013; Page *et al.* 2013; Raffa *et al.* 2017; Roth *et al.* 2018; Schwachtje and Baldwin 2008). Common methods that have been used to characterize the resistance traits include histological and morphological examination, thin layer chromatography, high-performance liquid chromatography, gas chromatography, liquid chromatography–mass spectrometry and gas chromatography–mass spectrometry (Danielsson *et al.* 2011; Whitehill *et al.* 2016). Most of these methods are time-consuming, labour-intensive and expensive. This has constrained most studies to the phenotypic level, limiting our understanding of defences at the genetic level (Jannink *et al.* 2010). The incorporation of genetically controlled chemical traits in resistance breeding programmes will depend on the availability of genetic variation (Falconer and Mackay 1996) but quicker and cheaper methods of determining genetic variation are required.

Novel high-throughput phenotyping approaches such as near-infrared spectroscopy (NIRS) are fast, non-invasive, low-cost, and environmentally safe analytical methods, that have become well-established for measuring plant physico-chemical constituents. NIRS has also emerged as a potential method for assessing the effects of a wide range of plant biotic and abiotic stressors, including herbivory (Coops and Stone 2005; Radeloff *et al.* 1999). Most stresses influence or are influenced by primary and secondary organic compound composition with functional groups - C-H, N-H, S-H and O-H - that have specific absorbance patterns in the NIR region (4000 -14000 cm^{-1}). These patterns can be qualitatively and quantitatively analysed for specific physiological states of the plants (Araus and Cairns 2014; Couture *et al.* 2016). There are examples in conifers and other tree species where NIRS has been used to predict herbivore damage, assess phytochemical variability, predict terpenoids and phenolics and predict palatability and nutrition levels (Couture *et al.* 2016; O'Reilly-Wapstra *et al.* 2013a; Quentin *et al.* 2017; Radeloff *et al.* 1999; Villamuelas *et al.* 2017). The applicability of NIRS for these uses depends on: 1) the presence of distinct physico-chemical differences between different levels of the condition under investigation (e.g. different susceptibility classes), 2) the ability of NIRS to generate characteristic spectra of these differences, and 3) the availability of calibration data that can

be well correlated with the spectra for quantitative predictions. Given the commercial interest in *Pinus radiata* and the significant impact of herbivore damage in plantations, it would be an advantage to use NIRS as a rapid, high-throughput phenotyping tool in an array of applications, including herbivory resistance breeding and deployment.

In *Pinus radiata*, the constitutive and induced chemistry has been characterised in the needles and bark (Moreira *et al.* 2012a; Reglinski *et al.* 2017) and roots (Chapter 3) using conventional wet chemistry techniques, as well as the physical traits (Chapters 2). Except for Moreira *et al.* (2012a) who demonstrated genetic control for total compounds, the genetic control of the individual traits has not been documented. NIRS could potentially be used to quantify the individual chemical compounds associated with observed differences in herbivory for a larger data set to provide genetic estimates. While spectroscopic techniques to identify damaged *P. radiata* plants has been utilised (Coops and Stone 2005), NIRS has rarely been used to examine the relationship between chemistry and damage by herbivores. However, NIRS has been used successfully in *P. radiata* in the field of wood science where wood properties such as lignin, monosaccharide composition (Fahey *et al.* 2018), physical wood properties (McLean *et al.* 2014; Schimleck *et al.* 2002) and bark extractives (Schimleck and Yazaki 2003) have been successfully predicted. In the present study, the suitability of NIRS as a classification and prediction tool in *P. radiata* was tested, specifically its ability to: 1) differentiate plant parts - needle, bark and root samples - based on their differences in spectra; 2) differentiate samples exposed to stress by simulated bark stripping and methyl jasmonate application from untreated control samples; and 3) quantitatively predict primary and secondary compounds in the needles, bark and roots.

5.2 Materials and methods

5.2.1 Experimental design

The plant material, experimental design and wet chemical analytical methods were described fully in Chapter 3. In brief, seedlings of 18 full-sib families of *Pinus radiata* were raised outside in a fenced area (to protect against animals) at the University of Tasmania, Hobart. At 2 years of age, 6 families were selected for treatment with either 25 mM methyl jasmonate (MJ) spraying or artificial bark stripping. The other 6 families were not treated (control plants). The three experimental groups (control, strip and MJ) were replicated 3 times and arranged in a randomized block design of 3 blocks in a shade house. One individual per family was then destructively harvested just before treatment application (T0), then 7 (T1), 14(T2), 21(T3) and 28(T4) days after treatment application collecting needles, bark and root samples separately. Most of the needles, bark and roots on the plant were collected according to the methods in Chapter 3. NIRS scanning of the fresh needle and bark samples was done immediately

after collecting the samples. After scanning, the sample was divided into two; one subsample for freeze drying and grinding for analysis of dried-ground samples and the other for reference wet chemistry analysis below.

5.2.2   Reference chemistry analysis

Quantitative chemical predictions require calibration against reference data that has been obtained following standard methods for example conventional wet chemistry techniques. To obtain the reference chemistry, we used wet chemistry methods as described in Chapter 3. In summary, extractions from the needles, bark and roots targeting terpenes and phenolics were made using dichloromethane (DCM) and acetone with respectively n-heptadecane and rutin as the standards (Jones *et al.* 2002; Sasidharan *et al.* 2011). Hot water was used to extract sugars. The DCM extracts that mainly comprised the volatile mono-, sesqui- and diterpenes and phenolics were analysed using gas chromatography-mass spectrometry (GC-MS). The acetone extracts containing mainly diterpenoids and fatty acids were analysed using ultra-high-performance liquid chromatography - mass spectrometry (UHPLC-MS) as well as the sugars. The DCM components were expressed as milligrams of heptadecane equivalents per gram of dry weight of the sample (mg HE/g dw) and the acetone analytes were expressed as milligrams of rutin equivalents per gram of dry weight of the sample (mg RE/g dw). The sugars were expressed in absolute amounts (mg/g dw). The mean and standard deviation of the amounts of each compound were calculated in the R software (version 3.6.0) (R Core Team 2018). Some samples were extracted in triplicates for estimation of laboratory error. The amounts of compounds quantified in the different plant parts were presented earlier and these include terpenes, phenolics, sugars, fatty acids and compounds that were not classified into the four classes (unknown compounds) (Chapter 3). All compounds were given an identifier (based on Supplementary Table 10) for ease of location in the tables.

5.2.3   Spectra collection and pre-treatment

Using a Bruker MPA Fourier Transform NIR spectrometer (Bruker, Germany) in the diffuse reflectance mode (12000 to 3800 cm^{-1}), spectral data were collected from 85 fresh and 85 dried-ground needle, 85 dried-ground root, 85 fresh outer bark, 85 fresh inner bark and 85 dried-ground bark samples. No spectral data were collected from fresh roots. Spectral data of fresh samples were obtained immediately after sample collection and again after the samples were freeze-dried and ground. Grinding of the freeze-dried samples was undertaken using a Cyclotec 1093 sample mill (FOSS, Denmark) that was cleaned between samples. The inner and outer sides of the fresh bark were scanned at 5 different points using a fibre optic probe resulting in an average spectrum per surface.

Bark thickness measurements of the samples were taken just before scanning using a Vernier calliper providing an average thickness of ~0.7mm. Fresh needles were cut into ~3cm pieces to allow homogeneity, then measured in a rotating sample cup through a glass bottom (9 cm diameter). The dried-ground bark, needles and roots were measured in 7 ml glass vials. Each spectrum was collected at 8 cm^{-1} and reflectance (R) data was stored as log (1/R). Background measurements were made after every after 2 hours. The OPUS (ver. 7.2; Bruker Optik GmbH, Germany) program was employed for the spectral data collection and to predict the most appropriate spectral pre-treatments. All qualitative and quantitative analyses were performed using the Unscrambler® X software (CAMO software version 10.2, CAMO AS, Trondheim, Norway).

Prior to the quantitative analysis, principal component analysis (PCA) based on a correlation matrix was used on the raw spectral data to eliminate spectral outliers based on Mahalanobis distance (Rousseeuw *et al.* 2006) and to observe any potential clustering of samples. PCA was performed separately for different comparisons of plant parts, treatments and time. To show the separation of plant parts, the PCA plot was generated using dried ground samples. The significance of the clustering along each PCA axis was tested using Kruskal–Wallis (KW) one-way analysis of variance in the R software (version 3.6.1) (R Core Team 2018). Partial least squares (PLS) was used to develop quantitative models using spectra and reference data (obtained from wet chemistry analysis methods or treatment groups). Full (also called leave-one- out) cross-validation methods were applied due to a limited number of samples (Zornoza *et al.* 2008). Cross-validation entails a set of techniques that partition the dataset and repeatedly generate models and test their future predictive power. The partitioning can be performed in different ways. The general format is that of a "leave k-observations-out" analysis. In such an analysis, the entire dataset is typically divided into k smaller observations. Full cross validation/leave-one-subject-out approach repeatedly splits the data but instead of creating k-folds, the dataset is split according to the number of subjects in the dataset, where, one subject is randomly selected for the testing purposes while the other subjects are used for training the model. This procedure is repeated until all the subjects have been used as test dataset (Koul *et al.* 2018). For chemical predictions, factors that were automatically selected by the algorithm were retained in the final models. To predict treated from untreated samples, a maximum of 2 factors was selected to prevent overfitting (Gowen *et al.* 2011). In most cases, spectral data were transformed by pre-treatments before the calibration process to remove spectral differences that are not related to physico-chemical properties of the samples such as physical differences due to sample preparation (Rinnan *et al.* 2009). Scatter-correction methods and spectral derivatives are the most widely used pre-processing techniques (Rinnan *et al.* 2009).

The performance of the PLS model was evaluated according to the: root mean square error of cross-validation (RMSECV), coefficient of determination (R^2) of the plot between the predicted values and the reference values, the predictive to lab error (PRL) and the residual predictive deviation (RPD) - which is defined as the standard deviation of observed values divided by the RMSECV. Based on published criteria (Malley *et al.* 2004; Saeys *et al.* 2005), R^2 values that range between 0.66 - 0.80 indicate approximate quantitative predictions, good prediction for R^2 values between 0.81 - 0.90 and excellent for R^2 > 0.90. Similarly, a residual predictive deviation (RPD) > 3 indicates an excellent prediction, between 2.5 - 3 predictions can be classified as good, 2 - 2.5 makes approximate quantitative predictions and an RPD < 2 is considered insufficient for applications. Other studies indicate that an R^2 of range 0.54 - 0.7 is acceptable and useful for initial screening studies (Malley *et al.* 2004; Quentin *et al.* 2017; Schimleck *et al.* 2003). PRL values ranging from 0.43 to 1.88 were established for amounts approaching laboratory precision but prediction errors within 2 times of the standard wet chemistry precision are sufficient for application (Yang *et al.* 2017). Similarly, the ratio of the range of the original data to RMSE (ratio error range -RER) was estimated. A minimum RER of 6.00 has been suggested as sufficient for detecting differences between classes of samples and for initial screening (Malley *et al.* 2004). In addition to the models developed for each plant part, spectra from different plant parts were aggregated and analysed together to reflect the scenario of developing global models for predicting chemistry using NIRS. Only dried-ground roots, needles and bark samples were used.

5.3   Results and discussion

5.3.1   *Characteristics of NIR spectra of the needles, bark and roots*

The spectral features of *P. radiata* needles, bark and roots were consistent with spectra from other coniferous species (Rautiainen *et al.* 2018; Toscano *et al.* 2017), where the critical peaks obtained are located in NIR region II (8500–5500 cm^{-1}) and region III (5500–4000 cm^{-1}) (Figure 5.1).
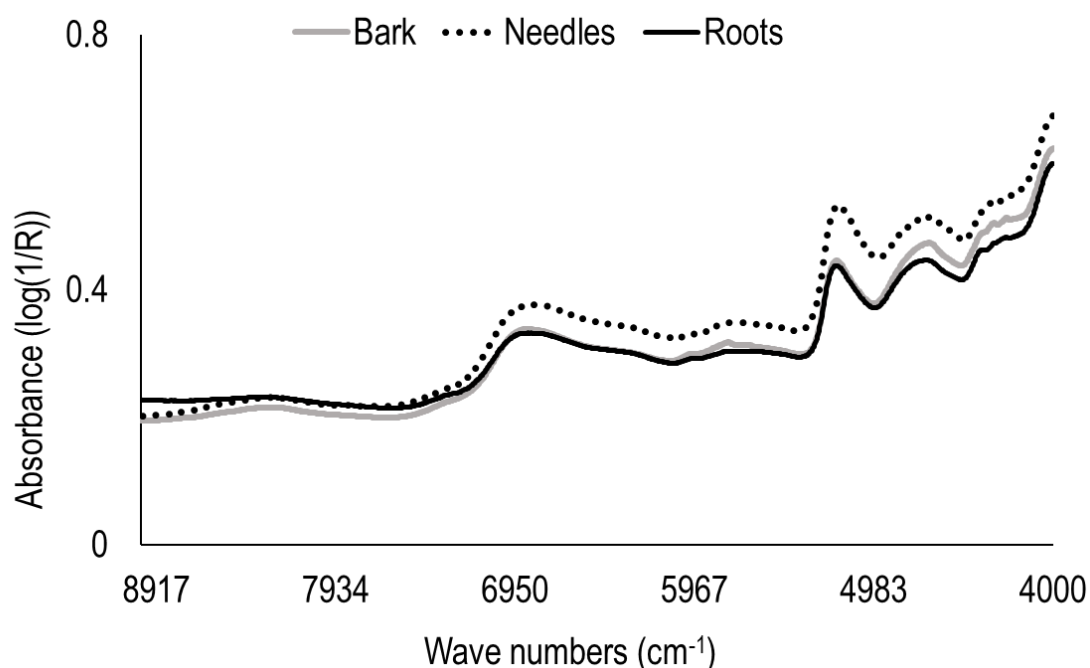
Figure 5.1: Representative averaged raw spectra of the needles, bark and roots of *Pinus radiata* derived from dried-ground samples. Spectra from wet samples (not shown) were very similar.

The derivatised spectra highlighted several spectral features between 4000 to 4300 cm^{-1}, 4700 to 5000 cm^{-1} and 5500 to 6000 cm^{-1} and 7100 to 7262 cm^{-1} (Figure 5.2). These peaks have been assigned to various functional groups that are pertinent to the major compounds detected by wet chemistry. Major peaks between 4400 and 4000 cm^{-1} have been associated with diterpenoids (Invernizzi *et al.* 2018). These peaks are very well resolved because diterpenoids were found to be abundant in the *P. radiata* needles, bark and roots. However, other terpenoids have been assigned to peaks between at 4775 and 5665 cm^{-1} (Invernizzi *et al.* 2018; Ma *et al.* 2019). The dominant peaks at ~5100 cm^{-1}, 4760 cm^{-1} and 4358 cm^{-1} are related to O-H functional groups in water and sugars (Invernizzi *et al.* 2018). Specifically, glucose, sucrose and fructose that were also identified in our samples have been related to peaks between 4398 and 4878 cm^{-1} (Rambla *et al.* 1997) and at 5186 cm^{-1} (Giangiacomo 2006). These peaks are strong as primary compounds are abundant in plants. The spectral region between 6000 and 5500 cm^{-1} is related to a CH_2 and CH_3 of proteins and fatty acids (Beć *et al.* 2018; Grabska *et al.* 2017; Invernizzi *et al.* 2018). Phenolics that have major peaks between ~6000 - 7000 cm^{-1} (Beć *et al.* 2018; Ma *et al.* 2019; Schwanninger *et al.* 2011) were weakly displayed, which is consistent with the wet chemistry results that quantified very few phenolics in all the plant parts. Various terpenes, including mono-, sesqui- and diterpenes have been associated with peaks between 7000 cm^{-1} and 8000 cm^{-1} (Ma *et al.* 2019). Weaker peaks were observed in the region 1 (12500 – 8500 cm^{-1}) (data not shown), which is associated with second and third overtones of functional groups CH_3 and CH_2 in terpenes

126

(Schwanninger *et al.* 2011) and phenolics (Ferrer-Gallego *et al.* 2011). However, because of the overlapping feature of NIR spectra, the regions indicated are proxies and may overlap with peaks from other constituents.
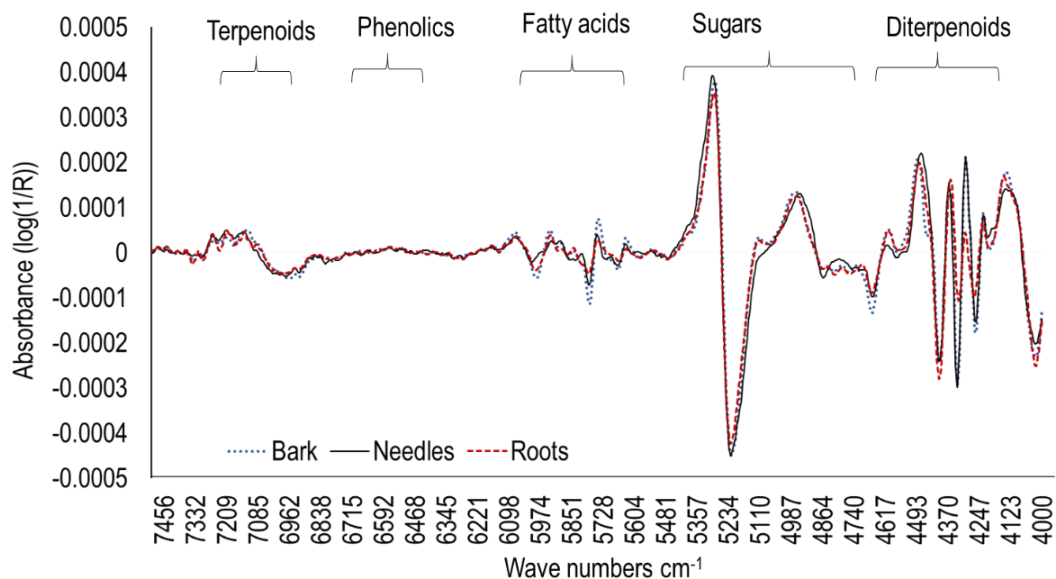
Figure 5.2: Second derivative using 13 smoothing points of the averaged spectra at T0 reflecting the major wave numbers contributing to *Pinus radiata* chemistry in the dried-ground needles, bark and roots.

5.3.2   Qualitative classification of plant parts

To inspect differences in spectra between plant parts, derivatised spectra collected from dried-ground needles, bark and roots of untreated plants at T0 (constitutive profile) were simultaneously analysed. The PCA plot (PC1 vs PC2) of the spectra indicated a spatial separation of the needles, bark and roots (Figure 5.3). Most spectral variance between the plant parts was explained by the first 4 principal components. PC1, PC2, PC3 and PC4 contributed 54.3%, 24.7%, 6% and 3%, respectively. PC1 mainly separated the needles from the bark and the roots but PC2 significantly separated the bark from the roots.

127
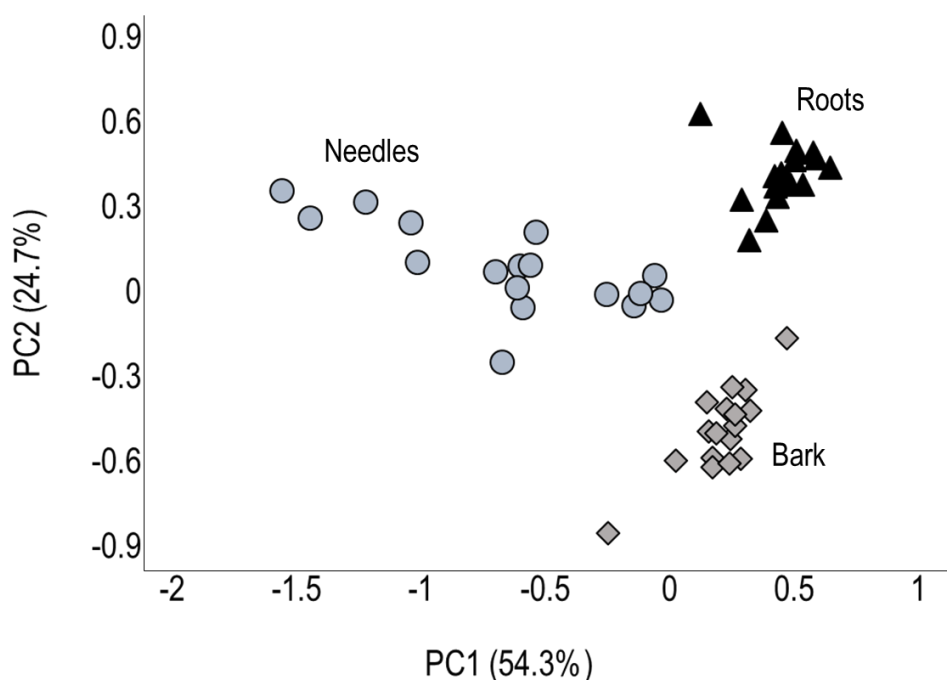
Figure 5.3: PCA plot (PC1 vs PC2) of the derivatised spectra (2nd derivative, 11 smoothing points) collected from dried-ground samples of needles, bark and roots. Only spectral data collected at T0 (before application of treatments) are displayed. The X and Y axes were scaled by multiplying the scores with 10^3.

To understand the chemistry that contributes to differentiating plant parts, averaged spectra at T0 of each part show differences in most wave numbers suggesting major differences in most chemical groups between the needles, bark and roots (Figure 5.2). Because NIRS is influenced by macro-components, the major differences in plant parts could be related to the dominant macro molecules including lignin, cellulose, hemicellulose and non-cellulosic neutral polysaccharides (Tenhaken 2015) rather than the relatively smaller molecules that were targeted by wet chemistry. However, the differences detected in the spectra above are consistent with the trends that were observed from GC-MS and LC-MS data on the same sample set (Chapter 3). Overall, the amount of compounds assigned as diterpenoids was lower in the roots than in the needles and the bark in the wave numbers between 4150 - 4381 cm^{-1} (Figure 5.2) (Invernizzi *et al.* 2018). Yet, the amount of compounds in the wave numbers between 5060 – 5500 cm^{-1} (Figure 5.2; assigned to sugars) were lower in the roots than in the needles and the bark. Similarly, the amount of compounds in several wave numbers (Figure 5.2), for example the wave numbers between 5750 - 5573 cm^{-1} and 5997 cm^{-1} (assigned as diterpenoids) (Invernizzi *et al.* 2018), were higher in the bark.

5.3.3   Spectral differences between treated and control samples

Both PCA and PLS analyses separated the treated from control samples in the inner and outer fresh bark, the dry-ground bark, the fresh and dry needles and the roots suggesting chemical differences

between the control and treated samples. The best prediction contrasting treated and control samples was determined by high coefficient of determination (R^2) and low root mean square error of cross validation (RMSECV) of partial least squares (PLS) models.

For the bark collected from MJ treated plants, the best prediction in the bark was achieved at T2 (14 days after treatment) for the dried-ground bark samples (RMSECV = 0.26, R^2 = 0.78) and at T3 (21 days after treatment) in the fresh inner bark samples (RMSECV = 0.25, R^2 = 0.79) (Figure 5.4A). For the bark collected from the strip treated plants, better predictions were achieved at T3 for both the dried-ground bark (RMSECV = 0.41, R^2 = 0.45) and the fresh inner bark (RMSECV = 0.30, R^2 = 0.69) (Figure 5.4B). In the strip treated samples, however, prediction was associated with a higher error (Figure 5.4B) indicating weaker responses to treatment compared to samples from MJ treatment. Figure 5.5 is a PCA score plot of raw spectra collected from the inner bark at T3 illustrating separation in space of treated and untreated samples, suggesting that the chemistry of the treated samples differs from the control.

For the needles and roots, the best prediction was achieved in the fresh needles (RMSECV = 0.24, R^2 = 0.81) and ground-dried roots (RMSECV = 0.38, R^2 = 0.51) at T2 from MJ treated plants (Figure 5.4C, 5.4D). The trends in the bark differed from those observed for the needles and roots (Figure 5.4C, 5.4D) suggesting differences between plant parts in the temporal response to treatments. On the same populations large changes were detected in primary and secondary compounds after treatment, with differences in time progression of the compounds in the different plant parts (Chapter 3). The strongest quantitative changes were detected 14 and 21 days after treatment application. In other *P. radiata* populations, Reglinski *et al.* (2017) also observed peak expression of induced chemistry between 14 and 21 days in the bark and between 21 and 28 days in the needles.
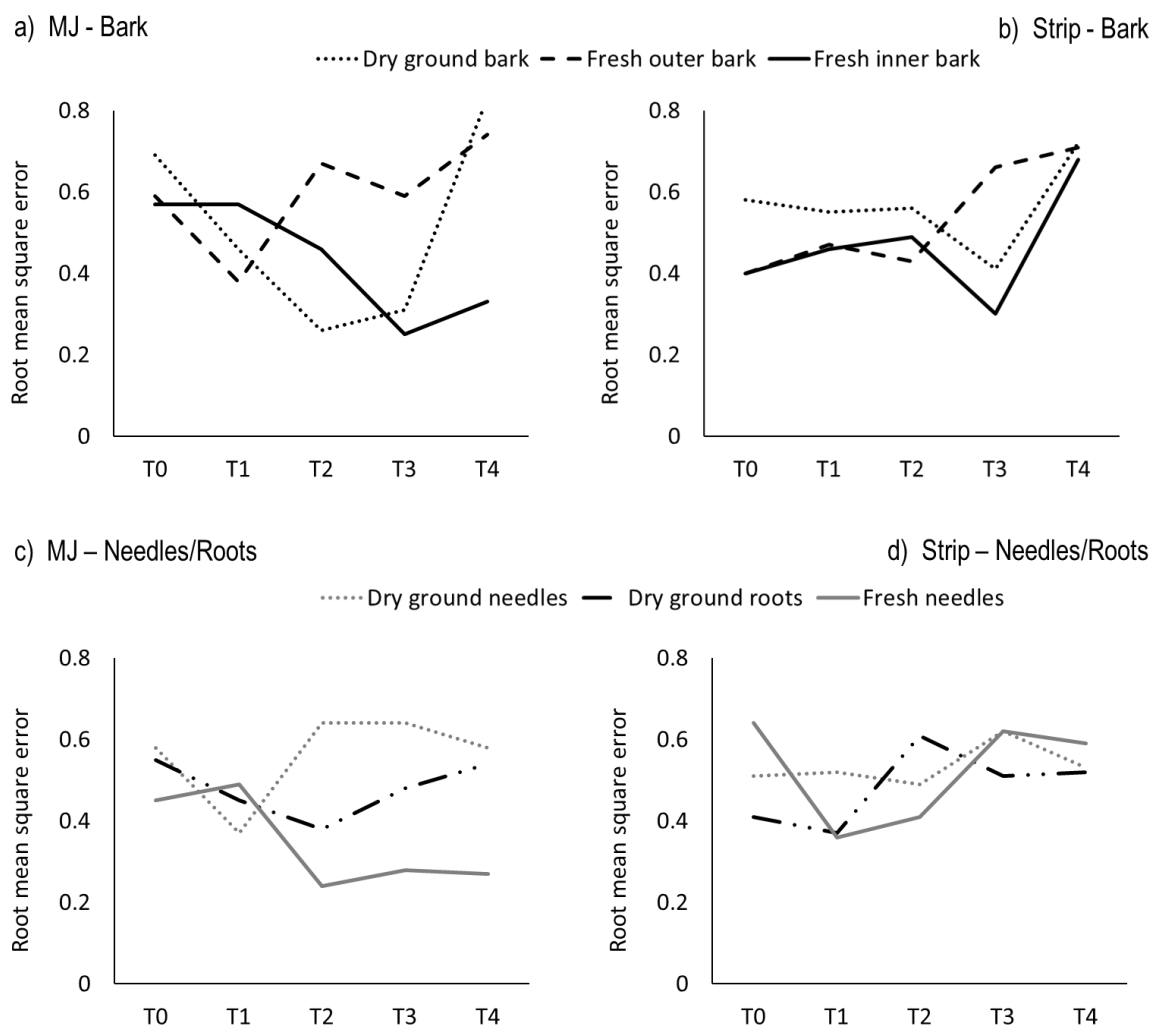
Figure 5.4: Partial least squares regression (PLS) root mean square error (RMSECV) associated with separating control bark, needle or root samples from; A) bark samples from MJ treatment, and B) bark samples from strip treatment, C) needle and root samples from MJ treatment and D) needle and root samples from strip treatment. No spectra were collected from fresh roots. Only two factors were used to avoid over fitting of the models.
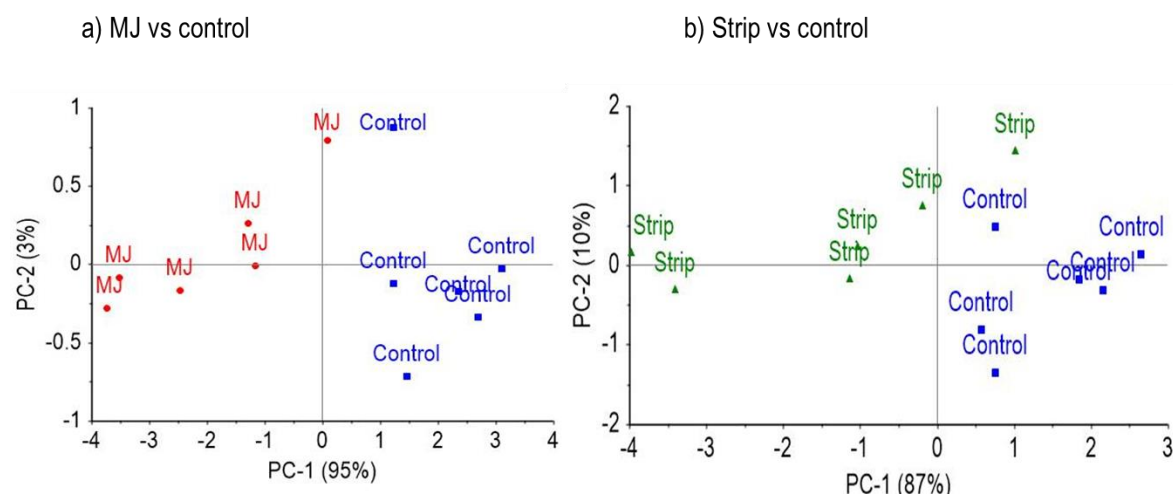
Figure 5.5: PC1 vs PC2 separates (a) methyl jasmonate - MJ and (b) strip treated samples and the controls. Results are based on the spectra collected from the inner bark three weeks (T3) after treatment application.

To identify the chemistry associated with the differences between the two treatments, regression coefficients of derivatised spectra showed that strip and MJ treated samples differed from the control samples especially in wave numbers between 5060 and 5500 cm^{-1} and between 4690 and 4950 cm^{-1} (Figure 5.6) that have been assigned to sugars. This suggests that sugars show more pronounced responses to treatments compared to other compound groups, which is consistent with the observations from wet chemistry. The wet chemistry results showed major induced changes in the sugars after treatment with less pronounced changes in terpenes and phenolics (Chapter 3). Sugars have been shown to drastically reduce in these treated samples (Chapters 3, 4) which may be related to increase consumption of sugars to supply energy, reduced photosynthesis during stress or the use of sugars as precursors for secondary defence compounds (Schwachtje and Baldwin 2008). Other studies with *P. radiata* also indicate qualitative and quantitative changes in the secondary chemistry of the needles, bark and roots following artificial wounding and methyl jasmonate application (Keefover-Ring *et al.* 2016; Lombardero *et al.* 2013; Moreira *et al.* 2013a; Moreira *et al.* 2009). The secondary compounds that respond to stress are thought to have some defensive functions (Schwachtje and Baldwin 2008).

In addition, the spectra variations indicated some treatment-specific responses. Based on the spectra collected from the inner bark, bark stripping caused responses at several wave numbers such as between 4316 cm^{-1} and ~5087 cm^{-1} that weakly responded in methyl jasmonate treated samples (Figure 5. 6). This suggests that the nature of expressed chemistry is dependent on the type of stress to which plants are subjected, consistent with earlier results (Chapter 3). Bark stripping may cause different chemical responses from methyl jasmonate application since it causes direct tissue loss, exposure of the phloem to potential infections and desiccation.
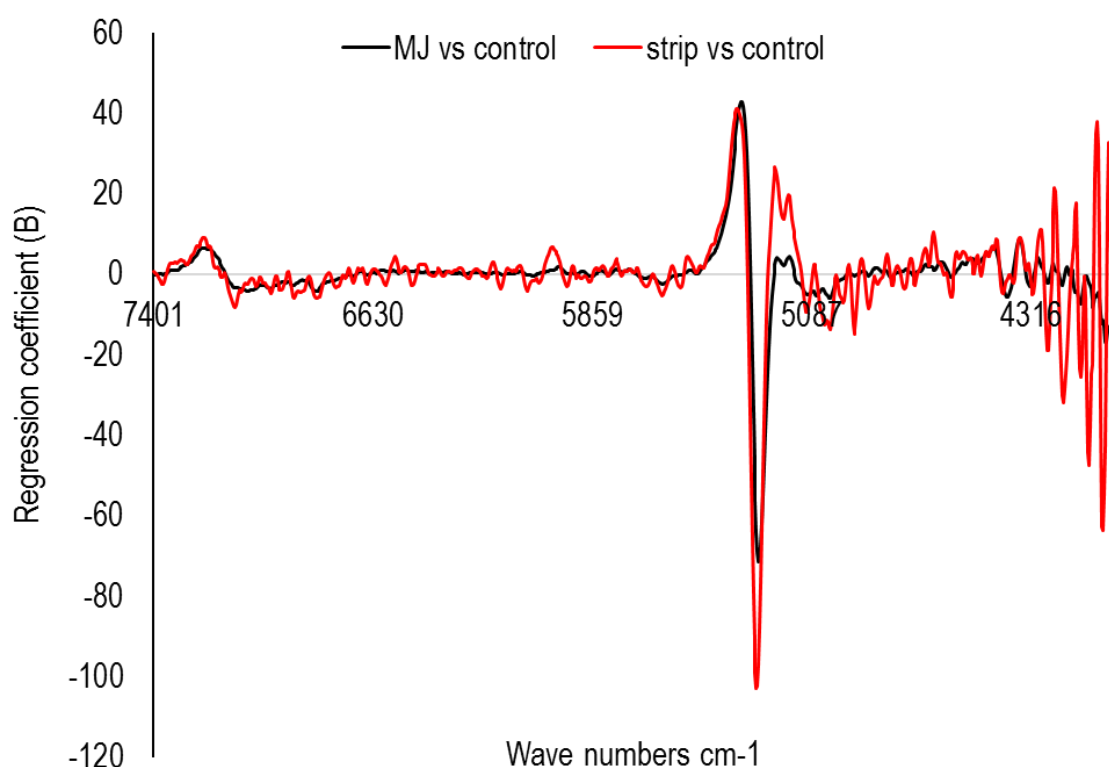
Figure 5.6: Regression coefficients differentiating responses of *P. radiata* bark to methyl jasmonate (MJ) and bark stripping (strip). The model is based on spectral data collected from the fresh inner bark at T3.

5.3.4 Quantitative prediction of individual compounds

NIRS models were developed for all the 81 compounds/groups and total compound groups that are listed in Supplementary Table 5.1. The prediction accuracy varied considerably between individual compounds. Of the primary compounds, glucose [77] (R^2 = 0.83, RPD = 1.9, PRL = 1.25, RER =8.50) and fructose [76] (R^2 = 0.84, RPD = 1.5, PRL = 2.72, RER = 7.25) exhibited the highest predictive accuracy (Figure 5.7, 5.8 A & B). The total sugars also exhibited a very high prediction accuracy (R^2 = 0.89, RPD = 1.1, RER =14.12). Generally, non-structural carbohydrates such as sugars show high prediction potential in tissues of *Pinus* (Acquah *et al.* 2018; Fahey *et al.* 2018) and other tree species (Quentin *et al.* 2017; Ramirez *et al.* 2015; Rubert-Nason *et al.* 2013), possibly because they exist in very high quantities (Cranswick *et al.* 1987). The prediction of glucose in bark was slightly lower than what has been observed in *P. radiata* wood (Fahey *et al.* 2018), which may be related to the relatively higher amounts of sugars stored in the wood of *P. radiata* (Cranswick *et al.* 1987). The wave numbers that offered best prediction for both glucose (6051.6 to 4597 cm^{-1}) and fructose (7501.8 to 4246.5 cm^{-1}) are consistent with the wave numbers established for these sugars (Rambla *et al.* 1997). Other primary compounds including fatty acids also had promising R^2 with higher RER values than sugars (Supplementary Table 5.1). The low RPD values, lower than what has been established in other studies suggests skewness of the data (Malley *et al.* 2004; Quentin *et al.* 2017; Schimleck *et al.* 2003).
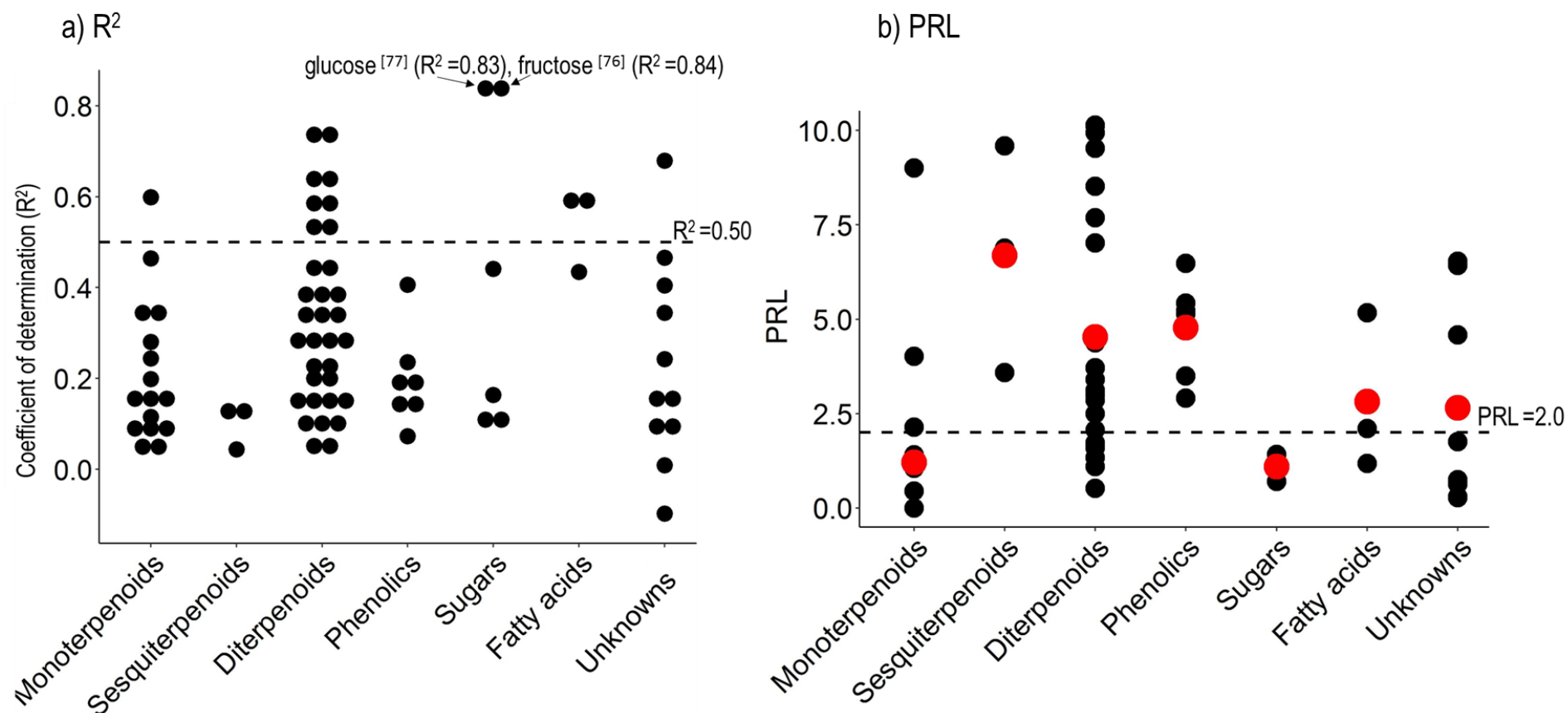
Figure 5.7: The coefficient of determination (R^2) (a) and prediction relative to lab error (PRL) (b) for the NIRS models for the compounds listed in Supplementary Table 5.1. The back dots for each graph indicate individual estimates and the red dots in (b) indicate the average PRL for each compound group. Models with $R^2 > 0.5$ can be confidently used for quantitative predictions as well as models with PRL~ 2.0.

Of the secondary compounds, the highest prediction was achieved for unknown isomeric diterpenoids; unknown $C_{20}H_{32}O_3$ C [49] (R^2 = 0.75, RPD = 2.02, PRL = 3.72, RER = 11.29) and unknown $C_{20}H_{32}O_3$ A [47] (R^2 = 0.72, RPD = 1.9, PRL = 4.73, RER = 10.76) (Supplementary Table 5.1 and 5.8 C & D). However, there are several compounds that had promising R^2 that included limonene[10] (R^2 = 0.60, RPD = 1.6, PRL = 0.45, RER = 24.16), dehydroabietic acid [29] (R^2 = 0.54, RPD = 1.5, PRL = 2.99, RER = 6.4), unknown resin acid groups - unknown $C_{20}H_{30}O_3$ [46] (R^2 = 0.63, RPD = 1.6, PRL = 3.70, RER = 51.35) and several other unknown diterpenoids (R^2 = 0.60 - 0.68, RPD = 1.6 -1.8) (Supplementary Table 5.1). The relatively low R^2 values in the present study for most individual compounds (Supplementary Table 5.1) contrast with other studies that have established very high values of prediction for terpenes, phenolics and other secondary compounds in other plant species (Couture *et al.* 2016; Juliani *et al.* 2006; Schulz *et al.* 2003). For example NIR models have been produced for cinnamon (*Cinnamomum zeylanicum*) and clove (*Syzygium aromaticum*) which have outstanding R^2 > 0.99 for all terpenes and phenolics (Juliani *et al.* 2006; Schulz *et al.* 2003). The amounts of secondary total compounds groups also exhibited low predictive accuracy (Supplementary Table 5.1, Figure 5.10a)

Various factors affect the predictive accuracy of the NIRS models and these include accuracy of the reference data and statistical distribution (Chu *et al.* 2005; Pérez-Marín *et al.* 2012). Most of the amounts of compounds were negatively skewed showing that there was limited representation of samples with higher relative amounts. This caused poor modelling of the amounts of the skewed compounds. The RER and RPD values have especially been indicated to be sensitive to skewed distributions (Malley *et al.* 2004). Figure 5.9 illustrates the statistical distribution of some well and poorly predicted compounds in this study. This suggests that increasing sample size to capture a wider range of values; including more higher values would enhance model performance. Consistently, compounds that were well predicted had a better spread of values in contrast to those that were poorly predicted that were more skewed (Figure 5.9). It has also been suggested that non-normal data may be better modelled with non-linear methods that were not tested in this study (Pérez-Marín *et al.* 2012).
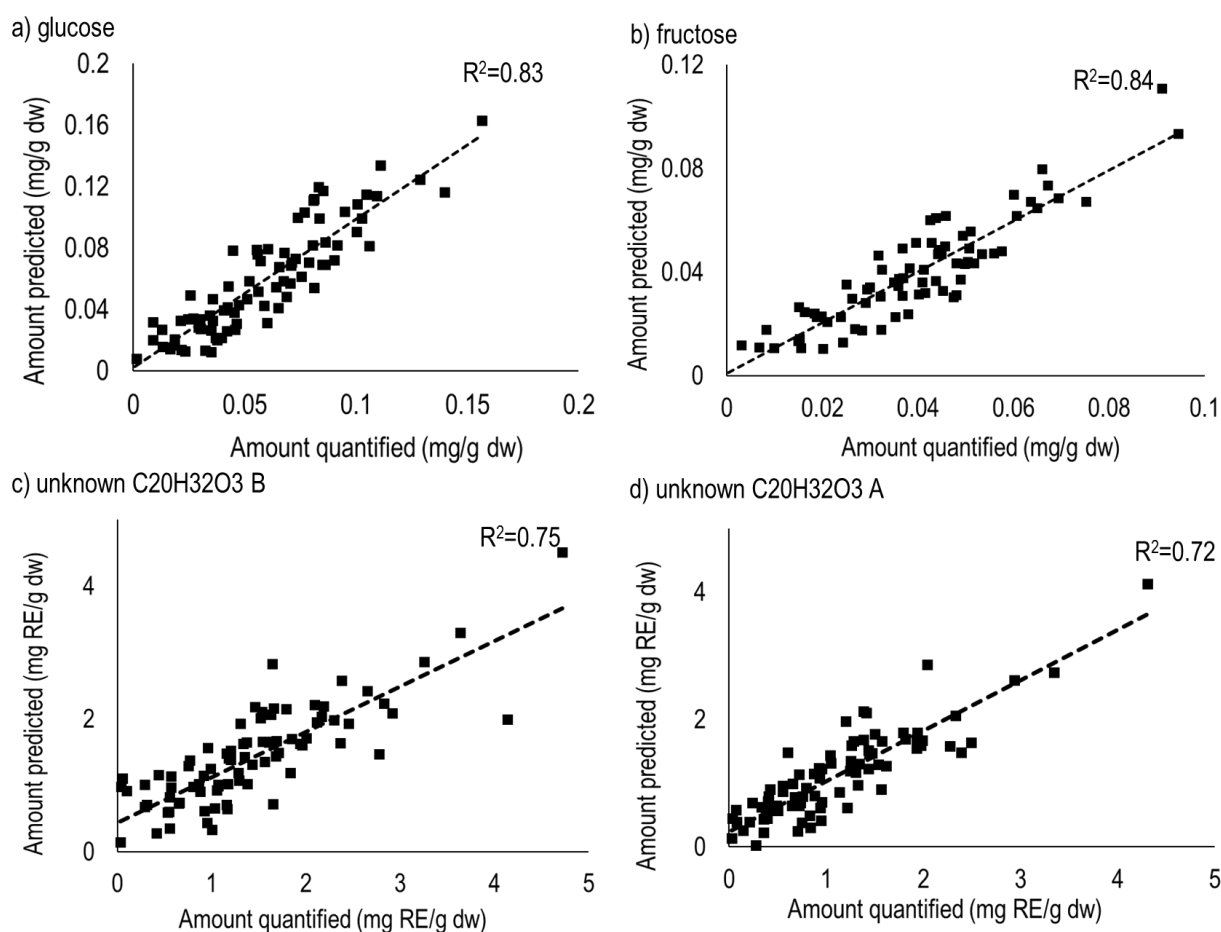
Figure 5.8: Linear relationship between the amount of compounds measured by standard laboratory procedures and that predicted by NIRS of the top two predicted primary compounds: a) glucose, b) fructose (top row) and secondary (bottom row) compounds: c) unknown C20H32O3 B, d) unknown C20H32O3 A. The absolute amounts of sugars are quantified in mg/g dw. The unknown secondary compounds are expressed as mg rutin equivalents (RE)/g dw, as no absolute quantitation was carried out on these analytes.
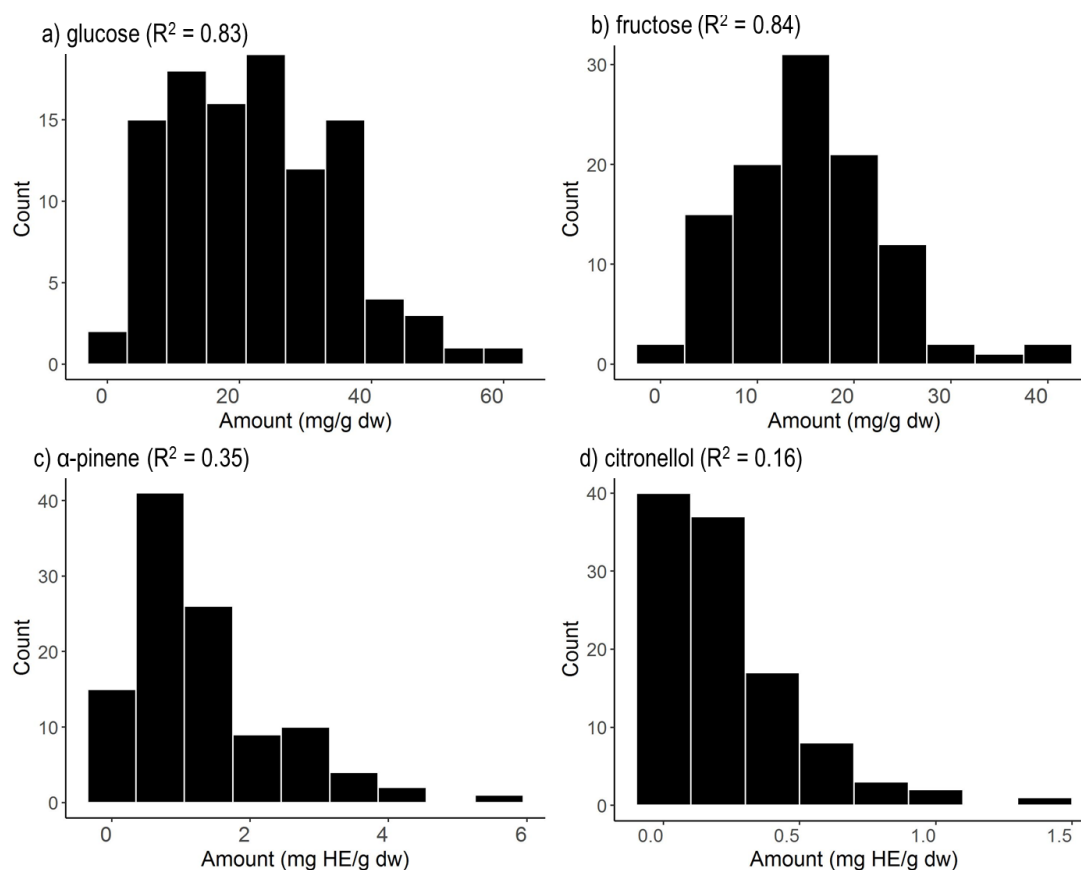
Figure 5.9: Histograms showing the statistical distribution of compounds that had high and low R^2 and respectively a better spread and a strong right skew. The absolute amounts of sugars; (a) glucose and (b) fructose are quantified in mg/g dw. The secondary compounds: (c) α-pinene and (d) citronellol are expressed as mg heptadecane equivalents (HE)/g dw, as no absolute quantitation was carried out on these analytes.

In the present study, predictive accuracy partly depended on the plant part from which the spectra were collected and the nature of the samples. There was, for example, comparatively better prediction of sugars in the dried ground bark than in the fresh bark or needles and roots. In contrast, fatty acids and monoterpenes were better predicted in the dried-ground roots and the diterpenoids were better in the dry-ground needles (data not shown). Figure 5.10b illustrates the variation in predictability of primary compounds in the various *P. radiata* plant parts and the preparation of samples. While the amounts of compounds are known to vary between *P. radiata* needles, bark and roots (Cranswick *et al.* 1987; Moreira *et al.* 2012a) and consequently the model accuracy (Riley and Crider 2000), sample drying has been shown to improve predictions by eliminating the variability in the water content. This improves spectral specificity for constituents such as sugar which depends on O-H groups in NIR interactions (Rubert-Nason *et al.* 2013; Stuart 2004). Our results especially showed that drying and grinding reduced the percentage contribution of PC1 to the total variation and altered the significant spectral ranges for the various analyses. Grinding also produced homogeneous samples that reduced spectral differences due to spatial heterogeneity.
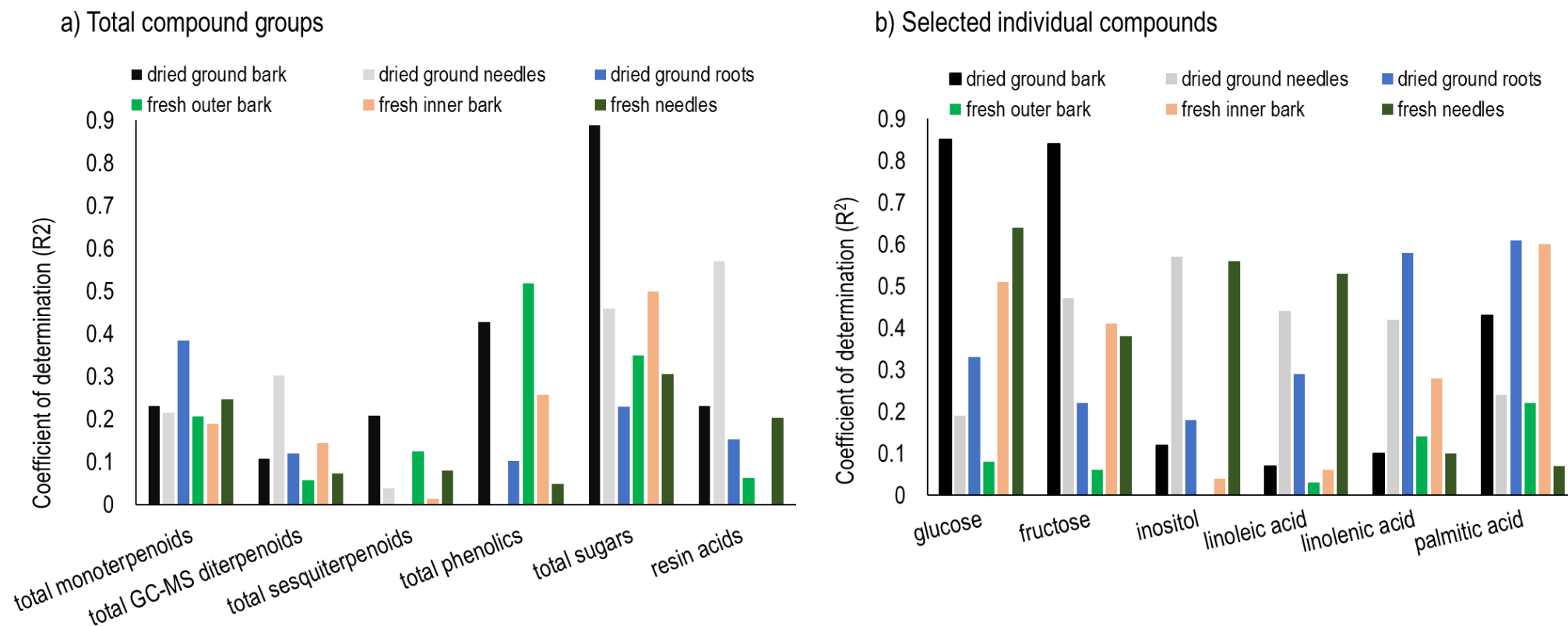
Figure 5.10: Differences in coefficient of determination (R^2) in different plant parts and sample types for a) total compound groups and b) selected primary compounds.

In terms of developing global models for *P. radiata* chemistry, the combined analysis of the spectra from dried-ground roots, needles and bark improved predictability of some compounds like β-pinene [4] (R^2 = 0.61, RPD = 1.61), α-pinene [1] (R^2 = 0.40, RPD = 1.3) and sabinene compared to the part specific models that are presented in Supplementary Table 5.1. While the R^2 of glucose [77] in a global model (R^2= 0.81, RPD = 2.3) was similar to tissue specific models, global model accuracy slightly reduced for fructose [76] (R^2 = 0.70, RPD =1.81), limonene [10] (R^2 = 0.2, RPD = 1.12) and unknown C20H32O3 B [48] (R^2 = 0.68, RPD = 1.75). This suggests that opting to amalgamate sample spectra for a global model will depend on target compounds. Apart from (Ramirez *et al.* 2015) who found optimal calibration for non-structural carbohydrates for samples originating from a broad range of plant species and tissue types, most studies show that heterogenous samples may lead to bias in validation and cross-validation results (Cécillon *et al.* 2009; Rubert-Nason *et al.* 2013).

5.4   Conclusion

This study demonstrates the potential to use *Pinus radiata* NIRS-based models to classify samples from different plant parts and differentiate samples from stressed and unstressed plant parts. NIRS models well-predicted sample glucose, fructose and selected terpenoids as well as provided proximate values for several fatty acids, monoterpenoids, diterpenoids and phenolics. Hence, NIRS models show potential to be useful in all aspects of natural herbivory prevention in other plant species, such as predicting chemistry for a larger number of samples for genetic studies. However, it is expected that the predictive accuracy could be improved if calibrations are developed with a larger sample size, and determining optimal sample sizes could be a subject for further investigation. Assigning wavelengths to secondary metabolites remains an important research goal in NIR spectroscopy.

Supplementary Table 5.1: Results of the cross-validated predictions for various chemical properties using partial least-squares (PLS) regression. For compounds detected in more than one plant part (Chapter 3), models were developed with spectra collected from all plant parts (bark, needles and roots) and state of samples (fresh and dry). However, only the best models (based on R^2) achieved from specific plant parts are shown. The table shows coefficient of determination (R^2) of the plot between the predicted values and the reference values, the root mean square error of cross-validation (RMSECV), the number of factors used in the model (rank), the residual predictive deviation (RPD) - which is defined as the standard deviation of observed values divided by the RMSECV, the predictive to lab error (PRL) and the ratio of the range of the original data to RMSE (ratio error range (RER). The best wave number range for the prediction as well as the pre-processing applied to the models are also indicated. Some compounds were not detected in the triplicates and hence have PRL=NA. The letters: "A', "B" and "C" against compounds denote isomers. Compound identifiers (Id) were given to each compound for ease of location. The identifiers in this table are not sequential as some compounds indicated in Supplementary Table 10 (after Chapter 9) were not identified in this data set

Id	Compounds	R^2	RMSECV	Rank	RPD	PRL	RER	Wave number range	Pre-processing	Plant part
	total compound groups									
	total monoterpenoids	0.38	0.41	6	1.27		6.61	6101.7-5449.9	multiplicative scatter correction (MSC)	roots
	total GC-MS diterpenoids	0.30	1.22	2	1.05		8.36	9403.3-6097.9	2nd derivative	dry needles
	total sesquiterpenoids	0.21	0.11	3	1.08		5.31	9403.3-7498; 6101.7-5449.9	no spectral processing	dry bark
	total phenolics	0.43	0.13	5	1.26		6.14	9403.3-7498; 6101.7-4597.5	2nd derivative	dry bark
	total sugars	0.89	0.02	9	2.68		14.12	7602.098-3802.972	MSC	dry bark
	monoterpenoids									
1	α-pinene	0.35	0.65	13	1.20	4.01	6.55	6102-5450	SLS	dry needles
2	α-terpineol	0.28	0.00	2	1.20	0.34	10.77	8451-7498; 5778-5450	1st derivative	dry bark
3	β-phellandrene	0.34	0.13	4	1.20	1.05	15.59	9403-7749; 5026-4598	2nd derivative	dry needles
4	β-pinene	0.46	0.24	6	1.40	0.45	6.47	6102-5450	MSC	roots
5	camphene	0.24	0.01	4	1.20	3.49	5.83	9401.8-7498.3; 5774.1-5448.2	SLS	fresh needles
6	citronellal	0.15	0.06	1	1.10	5.16	7.39	8450-7497	2nd derivative	outer bark
7	citronellic acid	0.10	0.02	3	1.10	13.96	4.97	6826.8-4246.5	SLS	inner bark

8	citronellol	0.16	0.22	12	1.10	25.72	4.41	9403-7497.9; 5777.7-5449.9	1st derivative	inner bark
9	γ-terpinene	0.04	0.01	1	1.00	NA	5.60	6101.7-5773	2nd derivative	inner bark
10	limonene	0.60	0.00	3	1.60	0.45	24.16	7502-5450;4601-4247	1st derivative + MSC	roots
11	linalool	0.12	0.01	2	1.10	1.21	8.56	8450.6-7497.9; 6101.7-5773.8	1st derivative	inner bark
12	myrtenoic acid	0.07	0.00	2	1.03	0.28	6.94	4601.4-4424	1st derivative + vector normalisation (SNV)	dry roots
13	sabinene	0.08	0.07	1	1.00	1.48	4.33	9403-7424.7; 4601.4-4424	1st derivative + SNV	roots
14	terpinene-4-ol	0.09	0.01	1	1.05	6.13	11.57	6101.7-5449.9	2nd derivative	inner bark
15	terpinolene	0.20	0.12	3	1.10	1.74	6.37	9403-5450	1st derivative + MSC	dry bark
16	unknown Mol Wt 150	0.15	0.02	8	1.08	35.13	6.09	8450.6-7497.9; 6101.7-5773.8	constant offset elimination	inner bark
	sesquiterpenoids									
17	bicyclogermacrene	0.04	0.01	1	1.02	3.59	6.54	5774.1-5448.2; 4601.6-4424.1	MSC	fresh needles
20	trans-farnesol	0.12	0.08	9	1.07	6.88	5.13	9403.2-7497.9; 6101.7-5449.9; 4601,3-4423.9	constant offset elimination	inner bark
21	unknown sesquiterpenoid alcohol	0.14	0.04	2	1.08	9.59	6.39	9401-6098.1	constant offset elimination	fresh needles
	GC-MS diterpenoids									
22	agatholal	0.35	0.38	4	1.24	2.07	6.79	9403-6098; 5026-4598	2nd derivative	dry needles
23	agathadiol	0.19	0.30	5	1.11	1.59	5.90	9403-7749;5026-4598	MSC	outer bark
24	copalol	0.10	0.12	3	1.05	0.52	9.81	8451-7498.3;6100-5448.2; 4601.6-4248.6	MSC	fresh needles
25	levopimaral	0.17	0.01	5	1.09	1.34	5.92	6102-5450	SNV	roots
26	methyl dehydroabietate	0.38	0.04	2	1.27	1.73	8.57	9403.3-6097.9	1st derivative + SLS	dry bark
27	methyl levopimarate	0.38	0.01	6	1.27	1.10	6.39	7502-4247	2nd derivative	roots
28	unknown C19H26	0.65	0.07	12	1.69	NA	10.23	9403.2-7497.9	SLS	inner bark
	LC-MS diterpenoids									
29	dehydroabietic acid	0.54	0.27	4	1.48	2.99	6.40	7501.8-5449.9	2nd derivative	dry needles

30	unknown diterpene-1	0.39	0.33	4	1.28	9.94	7.18	6799.8-6097.9	MSC	dry bark
31	unknown diterpene-2	0.10	0.22	1	1.06	NA	4.69	5453.8-4246.5	2nd derivative	dry roots
32	unknown diterpene-3	0.04	1.46	1	1.02	10.12	4.69	7501.8-6097.8	1st derivative + MSC	inner bark
33	unknown m/z 109 A	0.28	0.08	5	1.18	2.50	6.53	6799.8-6097.9; 4601.4-4246.5	SNV	dry bark
34	unknown m/z 109 B	0.28	0.25	3	1.18	NA	5.89	7501.8-4597.5	2nd derivative	inner bark
35	unknown m/z 121	0.16	0.03	4	1.09	NA	5.96	9403-7498; 4601.4-4246.5	min-max normalisation	dry roots
36	unknown m/z 134	0.11	0.42	1	1.06	NA	16.86	7501.8-6097.8	1st derivative + MSC	inner bark
37	unknown m/z 239	0.22	0.00	6	1.13	NA	5.90	6101.7-5449.6; 4601.4-4246.5	1st derivative + SLS	dry roots
38	unknown Mol Wt 272	0.29	0.05	5	1.19	10.14	6.64	9403-7498; 5777-5450;4424-4247	1st derivative	outer bark
39	unknown C20H30O2 A	0.20	0.20	5	1.13	3.12	13.73	5778-5450; 4601-4247	SLS	dry roots
40	unknown C20H30O2 B	0.32	0.31	12	1.22	7.01	5.12	7502-4598	constant offset elimination	dry needles
41	C20H30O2 resin acids	0.57	0.63	10	1.13	3.12	7.43	5778-5450; 4601-4247	straight line subtraction (SLS)	dry needles
42	unknown m/z 304 A	0.14	5.56	3	1.08	NA	0.00	6100-5774.1; 4426.1-4248.6	constant offset elimination	fresh needles
43	unknown m/z 304 B	0.22	0.07	2	1.13	9.52	4.62	5349.6-4597.5	SNV	dry needles
44	unknown m/z 304 C	0.36	0.14	6	1.25	8.52	10.97	9403.3-7498; 5453.8-4597.5	1st derivative	dry bark
45	unknown m/z 316	0.16	0.94	1	1.09	7.69	8.02	9403-7498; 5778-5459	2nd derivative	outer bark
46	unknown C20H30O3	0.63	0.17	14	1.64	3.70	51.35	6102-5450	SLS	dry needles
47	unknown C20H32O3 A	0.72	0.44	5	1.90	4.37	10.76	7501.8-6097.9	1st derivative + SNV	dry needles
48	unknown C20H32O3 B	0.06	0.05	1	1.03	2.86	23.53	442-4246.5	1st derivative + SLS	dry needles
49	unknown C20H32O3 C	0.75	0.38	6	2.02	3.72	11.29	7501.8-6047.7	1st derivative + MSC	dry needles
50	unknown C20H30O4	0.60	0.42	10	1.58	3.40	6.01	6475.9-5449.9; 4424-4246.5	SLS	dry needles
51	unknown C20H30O5	0.53	0.93	5	1.45	NA	5.62	8451-7498	MSC	outer bark
52	unknown C20H30O6 A	0.28	0.07	10	1.18	NA	7.89	6101.7-5773.8	constant offset elimination	inner bark
53	unknown C20H30O6 B	0.43	0.08	6	1.32	1.67	11.40	7500.2-5448.2; 4601.6-4248.6	SLS	fresh needles
54	unknown C20H30O6 C	0.23	0.77	9	1.14	NA	7.55	9403.2-7497; 5777.7-5449.9; 4601.3-4423.9	constant offset elimination	inner bark
	phenolics									
56	anethole	0.18	0.02	3	1.10	3.50	6.14	9403.2-8450.6; 6101.7-5449.9	SLS	inner bark

62	ethyl 4ethoxy benzoate	0.16	0.01	2	1.09	NA	31.57	8450.6-7497.9; 5777.7-5449.9	2nd derivative	inner bark
65	methyl eugenol	0.13	0	2	1.07	5.14	57.52	9403.2-7497.9; 6101.7-5773.8	1st derivative	inner bark
68	pinosylvin dimethyl ether	0.07	0.01	1	1.04	2.91	20.27	6475.9-5449.9	SNV	dry bark
70	raspberry ketone	0.21	0.05	3	1.12	6.48	5.33	9403.2-7748.6	1st derivative + SLS	inner bark
73	trans-coniferyl alcohol	0.24	0.03	4	1.12	5.22	6.37	9403-8450;4423-4246	SLS	outer bark
74	vanillin	0.41	0.02	3	1.30	5.43	7.36	4601.4-4246.5	no spectral processing	dry bark
	sugars									
76	fructose	0.84	1.50	7	1.50	0.71	7.25	8450,6-7498; 6101.7-5449.9	SLS	dry bark
77	glucose	0.83	1.89	9	1.90	1.14	8.50	8450,6-7498; 6101.7-5449.9	SLS	dry bark
78	inositol	0.16	0.24	1	1.10	1.12	10.53	6101.7-5773.9; 4424-4246.5	SLS	dry bark
79	sucrose	0.44	0.28	6	1.30	1.42	7.69	5778-5450	2nd derivative	dry bark
80	unknown disaccharide A	0.26	0.00	10	1.16	0.00	5.79	9403.3-4246.3	1st derivative + MSC	
81	unknown disaccharide B	0.49	0.00	5	1.39	0.00	6.77	9403.3-8450.6; 4601.4-4246.5	SNV	
82	unknown monosaccharide	22.35	0.00	6	1.13	0.00	6.37	6101.7-5449.9; 4601.4-4246.5	1st derivative + SNV	
	fatty acids									
83	linoleic acid	0.43	0.35	7	1.3	2.11	6.19	7501.8-6097.9	min-max normalization	dry needles
84	linolenic acid	0.57	0.08	11	1.5	1.18	14.75	8450.6-7498;4601.4-4246.5	no spectral processing	roots
85	palmitic acid	0.61	0.72	4	1.6	5.17	21.08	8451-7498	no spectral processing	roots
	unknowns									
86	unknown m/z 104	0.01	0.00	1	0.98	0.63	12.20	5025-4598	MSC	outer bark
87	unknown m/z 111	-0.10	0.01	1	1.00	0.75	4.36	5777.7-5449.9; 4424-4246.5	constant offset elimination	dry roots
88	unknown m/z 162	0.16	0.05	9	1.09	NA	6.98	7501.8-6097.8	constant offset elimination	inner bark
89	unknown m/z 272	0.15	0.03	1	1.09	0.27	4.83	5777.7-5449.9; 4601.4-4424	MSC	dry needles
90	unknown m/z 274	0.47	0.18	5	1.37	NA	11.01	9403.3-7424.7	SLS	dry bark
91	unknown m/z 302	-0.66	0.56	1	1.00	NA	6.98	7500.2-6798.2	min-max normalization	fresh needles
92	unknown m/z 358	0.11	0.06	6	1.09	6.53	6.40	9401.8-8449; 5774.1-5448.2	min-max normalisation	fresh needles
93	unknown m/z 362	0.68	0.05	3	1.77	1.75	9.55	6800-6098	1st derivative +MSC	dry needles

94	unknown m/z 406 A	0.40	0.09	2	1.29	0.31	6.60	7502-4247	no spectral processing	roots
96	unknown m/z 740 A	0.24	0.04	6	1.15	6.42	6.99	9403.3-7498; 6101.7-5349.6	SNV	dry bark
97	unknown m/z 740 B	0.34	0.08	3	1.24	4.58	2.92	6101.7-5349.6	SNV	dry needles
98	unknown m/z 770	0.08	2.44	2	1.04	NA	0.10	6799.8-6097.9; 5025.6-4597.5	2nd derivative	dry needles

1

CHAPTER 6: Additive genetic variation in *Pinus radiata* near infrared estimated bark chemistry and the chemical traits associated with variation in bark stripping

6.0   Summary

Secondary metabolites have been suggested as a major mechanistic link between genetic variation in herbivory levels of *Pinus radiata*. The potential to incorporate these chemical traits into breeding/deployment programmes in part depends on the presence of additive genetic variation for the relevant chemical traits. In this chapter, near-infrared spectroscopy was used to quantify the constitutive and induced levels of 65 compounds in the bark of trees from 74 full-sib families of *P. radiata* growing in a designed field trial, with between 3 and 6 trees per family. The trees sampled for chemistry were protected from browsing and induced levels were obtained by subjecting half of the trees to artificial bark stripping. The effect of this treatment on bark chemistry was assessed along with narrow-sense heritability, the significance of non-additive (family) genetic effects and the additive genetic correlations of compounds with marsupial bark stripping, that was observed in unprotected replicates of the field trial. The results indicated: (i) significant additive genetic variation, with low-moderate narrow-sense heritability estimates for most compounds; (ii) while significant induced effects were detected for 50% of the chemicals, no significant genetic variation in inducibility was detected; and (iii) sugars, fatty acids and a diterpenoid positively correlated with bark stripping while a sesquiterpenoid negatively correlated with mammalian bark stripping. In the absence of browsing, a trade-off with height and with other chemical compounds ($r_g = -0.85 \pm 0.22$, $p < 0.01$) was detected for selecting higher amounts of the sesquiterpenoids. However, overall, results showed that there is potential to incorporate near infrared estimated chemical traits into breeding/deployment programmes. This is possible with the use of near-infrared spectroscopy for large-scale phenotyping and the present study has shown that quantitative genetic analyses of chemical traits produce associations with bark stripping that mostly conform with those obtained using standard wet chemistry procedures.

6.1   Introduction

Defence against herbivory is achieved by constitutive and inducible physical and chemical traits that act directly or indirectly on herbivore feeding (Franceschi *et al.* 2005; Hudgins *et al.* 2004) and understanding their genetic architecture is of interest to the field of evolutionary ecology as well as plant breeding (Johnson 2011). In *Pinus* species and other conifers, physical traits include bark thickness and texture, constitutive and traumatic resin ducts and specialized phloem parenchyma cells (Franceschi *et al.* 2005; Hudgins *et al.* 2004). The chemical traits include secondary metabolites mainly terpenoids and phenolics, where higher amounts are linked to increased resistance to mammalian and insect herbivores in the needles and the bark (Chapter 4 ; Chiu *et al.* 2017; Iason *et al.* 2011; Zhang and States 1991). A few studies have also directly or indirectly associated the amounts of primary compounds with herbivory responses (Chapter 4 ; Gershenzon 1994; Tauzin and Giardina 2014; Tiffin 2000). These chemical and physical traits, that are often present in basal levels in plants, increase or reduce following real or artificial herbivory (Miller *et al.* 2005; Raffa and Smalley 1995; Sampedro *et al.* 2011). In *Pinus* species, both constitutive and induced traits have been shown to be under genetic control (Baradat and Yazdani 1988; Iason *et al.* 2011; Ott *et al.* 2011; Westbrook *et al.* 2015; Zhang *et al.* 2016a) and are potentially amenable to natural and artificial selection. However, for different traits, there is tremendous variation in the extent to which phenotypic selection on parents will impact on progeny, which is determined by the amount of additive genetic variation and hence narrow-sense heritability. Other factors being constant, traits with low heritability will respond more slowly than traits with higher heritability (Falconer and Mackay 1996).

While the presence of sufficient additive genetic variation for traits is an important requirement, current theories on the evolution of plant resistance predict the existence of evolutionary trade-offs (negative genetic correlations) between resistance and fitness traits, or between individual traits that can constrain their selection in breeding programmes (Huot *et al.* 2014). Terpenes and phenolics are carbon-based and their production requires carbon resources, resulting in potential conflicts among the compounds or with other plant functions such as growth (Sampedro *et al.* 2010; Sampedro *et al.* 2011). When there is a genetic-basis in these trade-offs, improving resistance through selection and breeding could negatively impact growth or other defence traits, and vice versa. Mixed evidence for the existence of trade-offs in *Pinus* species has been documented (Sampedro *et al.* 2010; Sampedro *et al.* 2011) but generally trade-offs are not expected in environments that are resource-rich as predicted by the resource availability hypothesis (Coley *et al.* 1985; Sampedro *et al.* 2011) and growth–differentiation balance hypothesis (Lorio 1986). Similarly, where multiple traits are required for effective defence, limited trade-offs will be observed among such traits (Carmona and Fornoni 2013).

In *Pinus radiata*, although genetic variation in herbivory has been documented for both mammals and insect herbivores (Chapter 2 ; Chapter 4 ; Moreira *et al.* 2013b), the associated defence mechanisms are not well established. A few studies have documented the involvement of physical structures such as thick bark, rough bark and obstructive branches on the stem in deterring herbivores (Chapter 2 ; Miller *et al.* 2014). Most recent studies have focussed on chemical defences and have found some relationships between the chemical defences and *P. radiata* herbivores (Moreira *et al.* 2013a; Moreira *et al.* 2013b; Sampedro *et al.* 2011) but the proportion of phenotypic variation explained by additive genetic variation has been estimated for only a few compound groups (Moreira *et al.* 2012a; Moreira *et al.* 2013b). Recently it was identified that sugars in the bark of *P. radiata* may contribute to susceptibility to mammalian herbivores (Chapter 4). There was little evidence of a relationship between needle chemistry and susceptibility to mammalian bark stripping (Chapter 4). Studies also indicate substantial response of *P. radiata* to both real and artificial herbivory and stress elicitors mostly by increasing the terpenes and phenolics and reducing the sugars (Chapter 3 ; Chapter 4 ; Moreira *et al.* 2013a; Reglinski *et al.* 2019). However, there is still limited support for the role of induced chemistry in deterring herbivores or its variation between families (Chapter 4 ; Moreira *et al.* 2013a). The presence of trade-offs between growth and chemistry and between different chemical traits has been demonstrated in *P. radiata* mainly at the phenotypic level (Chapter 3 ; Gould *et al.* 2008; Reglinski *et al.* 2019). However, the existence of a genetic-basis for the trade-offs for individual compounds has not been investigated.

This study examined additive genetic variation of primary and secondary metabolites in the bark of *Pinus radiata* using a field trial of full-sib families. The aims of this study were to; 1) examine the extent to which variation in *P. radiata* bark chemistry is under genetic control, 2) test whether there are genetic differences in the inducibility of bark chemical traits, 3) identify compounds that genetically correlate with mammalian bark stripping, and 4) examine correlations among selected chemical compounds and their correlation with growth. Because needle chemistry does not appear to be important in differentiating families which are resistant and susceptible to mammalian bark stripping (Chapter 4), this study focuses only on compounds quantified in the bark.

6.2   Materials and methods

6.2.1   Genetic trials

The genetic field trial used for this study is decribed in Chapters 2 and 4. It was established at Wilmot in northern Tasmania (-41.454271^{0}N, 146. 106801^{0}E, 580 m ASL), Australia in 2015 using genetic material sourced from the New Zealand Radiata Pine Breeding Company (RPBC). The genetic material

comprised 74 full-sib (cross-pollinated; CP) families generated from 55 unique parents and 54 grandparents which were planted out in the field in a randomised incomplete block design of 26 replicates, three incomplete blocks per replicate and each family represented as a single-tree plot within each replicate. The field trial was fenced to prevent bark stripping by native mammals. The dominant native herbivore on site was the Bennett's wallaby (*Macropus rufogriseus* subspecies *rufogriseus*). The density of the Bennett's wallaby within the area was estimated at 32 animals/km^2 (DPIPWE 2018). In 2017 (when seedlings were 25 months of age), the gates of the trial were opened during winter for about two months to allow access to animals to do bark stripping. Six of the 26 replicates were further protected using internal fencing to totally exclude the herbivores, to allow chemistry to be assayed in the absence of uncontrolled browsing (see chemistry experiment described below). The remaining 20 replicates were freely accessible to browsing and were used to assess the genetic variation in susceptibility to mammalian bark stripping.

6.1.2 Experiment 1: Assessment of mammalian bark stripping

The details of the bark stripping assessment are presented in Chapters 2 and 4. In brief, at 2 years of age, after ~2 months of exposure to mammalian bark stripping, the amount of bark removed by the marsupials was scored in the 20 replicates (n = 1550 plants), on a scale of 0-5; 0 = no damage, 1 = <25%, 2 = 25 - 50%, 3 = 50 - 75%, 4 = >75%, 5 = 100% damage (completely ring barked). At the same time the height of trees from all 26 replicates was measured. By this time, bark stripping had not differentially influenced the height of the different families as depicted from the non-significant tree*protection term (described later in the methods). The average height of the trees at the time of bark stripping assessment in the 20 unprotected replicates was 147.4 ± 0.90 cm and 163.7 ± 1.54 cm for trees in the 6 protected replicates. Browse scores were converted to mid-point values for data analyses, except for scores 0 and 100 (Chapter 2). The bark stripping scores used were spatially adjusted to minimize the spatial effects as detailed in Chapter 2.

6.2.3 Experiment 2: Chemistry experimental design and chemical analysis

Three weeks after the bark stripping assessment was conducted, an experiment was initiated to assess the constitutive and induced chemical differences among all the 74 families using trees in the 6 protected replicates (n = 393 plants). Half of the plants were subject to artificial bark stripping (treated trees) at T0 and half were untreated and used as controls (more details of the sampling are presented in Chapter 4). The treatment was applied by removing a vertical strip of 15 cm of bark, starting 2 cm above the ground, and covering 30% of the stem circumference (Figure 6.1). The dimensions were selected based on the most common browsing level observed in Experiment 1. Three weeks after the

treatment was applied (Chapter 5), bark samples were collected from both the control and the treated trees. This second bark sample was collected from all the trees ~1 cm above the stripped part on the treated trees as illustrated in Figure 6.1. On the control trees, a bark sample of similar size was collected from a similar height as the one from the treated plants (Figure 6.1). All bark samples were collected from the north side of the stem. Samples were kept in a cool box until transportation to the laboratory for near infra-red spectroscopy (NIRS) scanning of fresh samples. After scanning, each sample was divided into 2 parts; one part was stored in a -20⁰C freezer until wet chemical extraction and the other was freeze-dried and ground using a Cyclotec 1093 sample mill (FOSS, Denmark) for NIRS scanning of dried-ground samples.
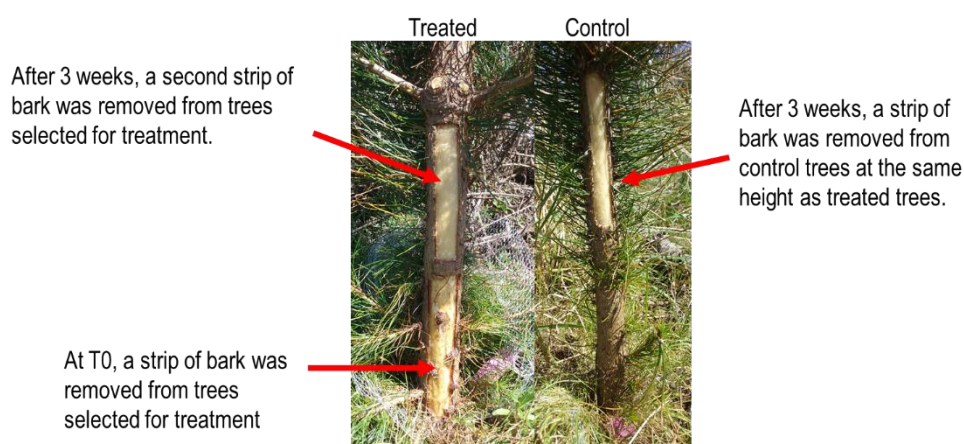
Figure 6.1: Fifteen month-old *Pinus radiata* trees showing the artificial bark stripping treatment (lower left) and how bark was sampled for chemical analysis (upper strip removed) from treated (left) and untreated control (right) plants. The artificial bark stripping treatment was applied by removing a vertical strip of bark 15 cm long, starting 2 cm above the ground and covering 30% of the circumference. After 3 weeks a bark strip for chemical analysis was collected 1cm above the treated area of the treatment tree and at similar height for the control tree.

6.2.4   Near infrared reflectance spectroscopy models and wet chemical analysis

Assessment of chemistry is conventionally performed using wet chemistry procedures. However, the need for large sample sizes for genetic analysis puts a constraint on the use of wet chemistry given the associated cost and labour (Siesler and Ozaki 2002). The ability of NIRS to accurately predict the amount of primary and secondary compounds in *P. radiata* enabled fast, low cost and large scale chemotyping for this genetic study. Robust models for predicting *P. radiata* chemistry in the needles and the bark have been reported in Chapter 5.

To predict the chemistry of *P. radiata* using NIRS, bark samples from all trees in the 6 protected replicates were scanned when fresh and when freeze dried and ground according to the methods in Chapter 5. The fresh bark samples were divided into two; the part closer to the original strip (proximal)

and one further from the strip (distal). For both the proximal and distal samples, spectra were collected from the inner and outer sides. Briefly, the fresh bark samples were scanned using an optic fibre reflective probe at five different points and the spectra averaged. The dry samples were scanned in diffuse reflective mode through the bottom of glass vials. Each spectrum was collected at 8 cm^{-1} using the OPUS (ver. 7.2; Bruker Optik GmbH, Germany) program and reflectance (R) data was stored as log (1/R). All qualitative and quantitative analyses were performed using the Unscrambler® X software (CAMO software version 10.2, CAMO AS, Trondheim, Norway).

Principal component analysis (PCA) was carried out on the spectra of all samples, to enable selection of a subset (~0.3 of the total number of samples) of samples for wet chemical analysis (see below) for model calibration purposes. 150 samples that were representative of the spectral variation present in the entire sample set were selected. One hundred of these were used to develop the models based on cross-validation and 50 samples were used for external validation before predicting the unknown samples (n=243). Partial least squares regression (PLS) models of the spectral data and quantitative chemical data of the 150 calibration samples were developed to predict the amount of quantified chemical constituents in all the other samples. In most cases, spectral data were transformed by pre-treatments before the calibration process (Rinnan *et al.* 2009). The performance of the PLS models was evaluated according to the root mean square error, coefficient of determination (R^2) of the plot between the predicted values and the reference values and the number of factors used in the model. The better model of either the cross-validated or the externally validated model was used to predict the chemistry of the unknown samples. The metrics for the final predictive models for each compound or group of compounds can be found in Supplementary Table 6.1.

Wet chemical extractions that targeted terpenes, phenolics and sugars were performed separately for the bark from each tree using three extraction solvents: dichloromethane (DCM), acetone and hot water, according to the methods documented in Chapter 3. DCM extracts comprised the volatile terpenoids and phenolics and the acetone extracts comprised the diterpenoids and fatty acids. The DCM extracts were analysed by gas chromatography-mass spectrometry (GC-MS) and the acetone extracts and the hot-water extracted sugars were analysed by ultra-high-performance liquid chromatography-mass spectrometry (UHPLC-MS). The preliminary identification of the compounds was based on the comparison of the retention time and mass spectra with the National Institute of Standards and Technology mass spectra library (NIST 2014). However, to verify the retention times for final identification of diterpenoid resin acids by UPLC-MS, standards of abietic acid, neoabietic acid, dehydroabietic acid, palustric acid, levopimaric acid, pimaric acid and isopimaric acid were purchased from Santa Cruz Biotechnology and analysed by UPLC-MS. The components were expressed as

equivalents to the respective internal standard, except for the sugars that were measured in absolute amounts. Some compound peaks that could not be resolved on chromatograms were reported as groups of compounds as shown in Supplementary Table 6.1. A unique number was given to each compound for ease of identification in the tables. Some samples were extracted in triplicates for estimation of lab-error to enable calculation of the NIRS predictive error relative to the lab error mentioned above (PRL). The compounds that were included in the final data analysis were selected based on 2 criteria. First was the ratio of the range of the original data to the RMSE (ratio error range (RER). A minimum RER of 6.00 has been suggested as sufficient for detecting differences between classes of samples and for initial screening (Malley *et al.* 2004). Secondly, among those that did not meet this criteria, further selection was done based on PRL and in this case, a PRL < 3.00 was selected based on suggestions that prediction errors within approximately twice the standard wet chemistry precision are sufficient for application (Yang *et al.* 2017).

6.2.5 *Estimation of additive genetic variation in chemical traits*

For all chemical traits, the presence of additive genetic variation was tested based on variance components generated by fitting univariate linear mixed models in ASReml v4.1 (Gilmour *et al.* 2015). The general linear mixed model is represented below

$$y = X\beta + Zu + e, \tag{1}$$

and in this study, y is a vector of the amounts of individual compounds, β is a vector of fixed effects (i.e. treatment-inducibility) and X and Z are design matrices associated with the fixed and random effects. u, is the vector of random effects and in the full model, the random effects included replicates, blocks within replicates, family (specific combining ability - SCA), tree (additive genetic effect) and the tree x treatment interaction (differential inducibility) that were assumed to be normally distributed. The model term, e is the vector of random residuals. The model was fitted using a three-generation pedigree file.

The significance of the fixed treatment effect (inducibility) was tested using the Wald-F statistics (Gilmour *et al.* 2015). The significance of the random terms was sequentially tested in univariate models using likelihood ratio tests (LRT) starting with family (SCA), differential inducibility (tree x treatment) and then the additive genetic variation. Full models were compared with respective reduced models using one-sided likelihood ratio tests with one degree of freedom (Gilmour *et al.* 2015). Bonferroni's correction was applied to the p-values associated with SCA, inducibility, differential inducibility and additive genetic variation to reduce the chances of obtaining false-positive results (type I errors) when multiple tests are performed (McDonald 2009). The Bonferroni correction was applied within each compound group (monoterpenoids, diterpenoids, sesquiterpenoids, phenolics, sugars, fatty

acids or unknowns) by multiplying the p-value derived from the likelihood ratio tests (LRT) by the total number of compounds tested within each compound group. Significant p-values were considered at 0.05/n, where n is the number of statistical tests (McDonald 2009), for example a p-value of 0.05/13 =0.004 was considered significant for additive genetic variation of individual monoterpenoids, where 13 = the number of monoterpenoids retained in the final dataset (Table 6.1).

From univariate analyses, the individual narrow-sense heritability ($\hat{h}^2$) was estimated as the additive genetic variance divided by the sum of the additive genetic variance $\hat{\sigma}_a^2$ and the error variance as below:

$$\hat{h}^2 = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_e^2} \tag{2}$$

The associated standard errors were estimated through "delta method" using ASReml (Gilmour *et al.* 2015) based on Taylor expansion (Lynch and Walsh 1998). The variance components used for this heritability calculation were derived from the above model (Equation 1) excluding the family and the interaction terms since they were not significant for all compounds after Bonferroni correction (see results).

6.2.7   *Genetic correlations between chemistry and mammalian bark stripping*

To determine the relationship between genetic variation in specific chemical compounds and amount of mammalian bark stripping, genetic correlations were estimated in trivariate models. In each model, height, spatially adjusted bark stripping (Chapter 2) and one chemical compound were fitted as response variables. Height was the only trait that was assessed in all the 26 replicates and therefore acted as the bridging trait between the 20 unprotected replicates where mammalian bark stripping was scored and the 6 protected replicates where chemistry was assessed. The family term was not fitted at this stage. The terms "protected" and treatment were fitted as fixed effects. The fixed term "protected" was fitted for height to distinguish the 20 plots that were not protected and from which bark stripping was estimated from the 6 protected plots from which chemistry was estimated. The treatment term was fitted only to the chemical compound. The design terms (replicates and blocks) and additive genetic term were retained as random terms in the trivariate models. The unstructured variance-covariance structure was fitted for the tree term and a diagonal matrix for the replicates and blocks within replicates.
.

The genetic correlation r_g between two traits measured was estimated as:

$$r_g = \frac{cov_a(x,y)}{\sqrt{\sigma_{ax}^2 \cdot \sigma_{ay}^2}} \tag{3}$$

where $cov_a(x, y)$ is the additive genetic covariance between traits x and y, σ^2_{ax} is the additive genetic variance components for trait x, and σ^2_{ay} is the additive genetic variance components for trait y. The standard error for each genetic correlation was estimated using a Taylor series approximation in ASReml (Gilmour *et al.* 2015). To test if the additive genetic correlation was different from zero, a full trivariate mixed model with all the model terms that allowed covariance between bark stripping and the chemical compound was fitted and was compared with a model where the additive covariance was fixed to zero. A two-tailed LRT with one degree of freedom was used. No adjustment was applied to the p–values of the correlations for compounds that were associated with bark stripping in Chapter 4 as there were clear *a priori* reasons for specifically testing these compounds. However, for interpreting significance of any new correlations, Bonferroni's correction was applied within compound groups as indicated above.

6.2.8   *Genetic correlations among chemical compounds and height*

For the chemical compounds that had significant additive genetic correlations with mammalian bark stripping, the genetic correlation between chemical traits and height were estimated to test for genetic-based trade-offs with growth. This was done using the trivariate models described above. Before this model was fitted and genetic correlation tests undertaken, a random tree*protected term was fitted in the trivariate model (model 3) and its significance tested using a one-tailed LRT. This aimed to test if the height of different families differentially responded to bark stripping by the time of assessment. However, there was no evidence for a significant tree*protected interaction effect on height (results not shown) suggesting that the height measured in the presence of browsing was likely unaffected by the browsing at the time of measurement (Chapter 2). Therefore, LRT for the genetic correlations proceeded with this term out of the model, by comparing a full model that allowed the covariation between height and the chemical compound to that where the covariation was fixed to zero.

Bivariate models were used to test the genetic correlations among all compounds that had a significant genetic correlation with bark stripping. Bivariate models included the treatment as a fixed term that was fitted for both compounds. The tree and design terms were included as random terms. The unstructured variance-covariance structure was considered for all the random terms. Bonferroni's correction to the correlations was not applied at this stage. Pearson's phenotypic correlations were also estimated in ASReml from bivariate models (Gilmour *et al.* 2015) and the test that the phenotypic correlations were different from zero was done using the cor.test function of R v 3.6.1.

6.3 Results

6.3.1 Predictions of chemical traits

Near infrared spectroscopy models were developed for all 65 compounds quantified in the bark by wet chemical analysis (Supplementary Table 6.1). Better calibration models with higher R^2 were mostly developed with the spectra collected from the dry ground bark compared to fresh bark with few exceptions (Supplementary Table 6.1). Therefore, the chemical predictions presented were derived using NIRS models developed with spectra collected from the dry ground bark.

Based on dry ground samples, the predictive accuracy of NIRS models, determined by the RER, PRL and R^2 varied considerably between compounds (Figure 6.2). Of the primary compounds, the sugars, glucose [77] (RER = 11.12, PRL = 1.76, R^2 = 0.79) and fructose [76] (RER =10.55, PRL = 1.63, R^2 = 0.77) showed the highest predictive power (Supplementary Table 6.1). Of the secondary compounds, the highest prediction was achieved for unknown diterpenoids; unknown C20H32O3 C [49] (RER = 12.52, PRL = 4.50, R^2 = 0.83), unknown C20H30O3[46] (RER =14.79, PRL= 3.72, R^2 = 0.83), unknown m/z 316[45] (RER = 11.87, PRL = 2.97, R^2 = 0.72) and unknown C20H30O5[51] (RER = 12.24, PRL = 4.59, R^2 = 0.71) as well as monoterpenoids; α-pinene[1] (RER = 7.63, PRL = 0.81, R^2=0.73) and β-pinene [4] (RER = 10.30, PRL =1.01, R^2 = 0.73) (Figure 6.2). However, there are several compounds that had R^2 > 0.50 and these included the sugar; inositol[78], fatty acids (linoleic acid[83] and linolenic acid[84]) and secondary compounds that included monoterpenes (camphene[5], citronellal[6]), and diterpenoids (agathadiol[22], agatholal[23], copalol[24], levopimaral[25], dehydroabietic acid[29] and several unknown diterpenoids[33] [45] [50] [51] [52] [54]) as well as unknown compounds [90] [94] (Supplementary Table 6.1). The proceeding results however focussed on the 54 compounds that were retained (Table 6.1) after the selection criteria defined in the methods.
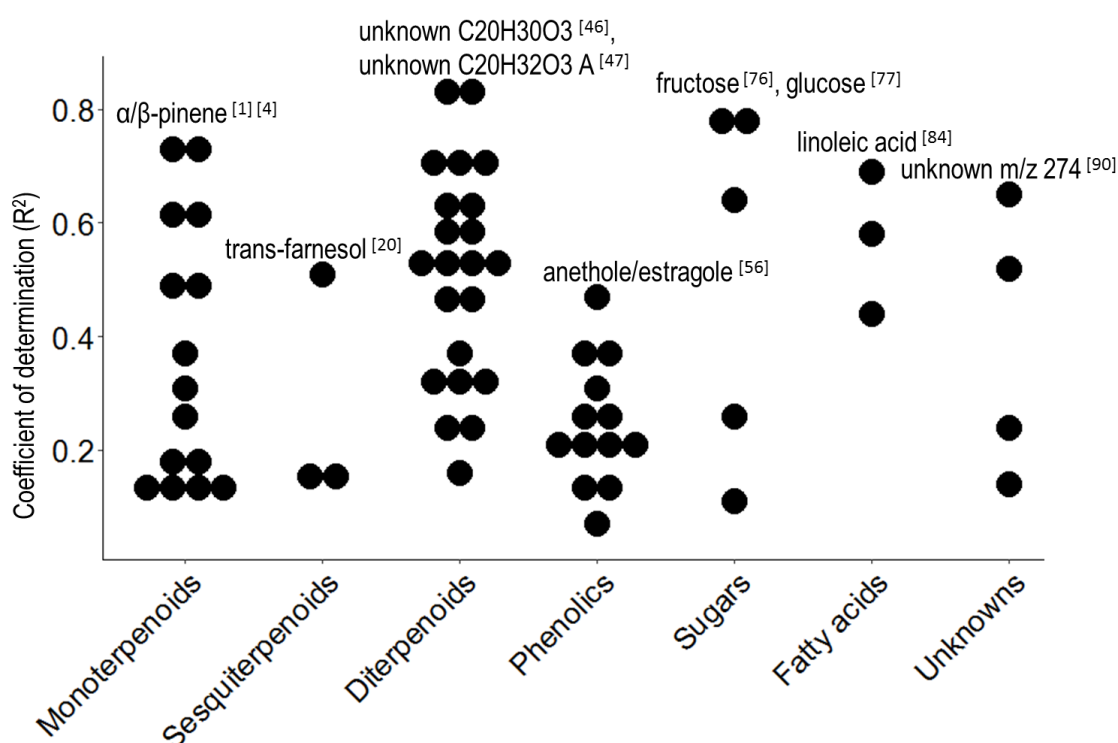
Figure 6.2: Dot plot of the distribution of the coefficient of determination (R^2) for the NIRS PLS models for the 65 chemical compounds identified in the bark. Each dot represents the one R^2 estimate for a specific compound and these have been grouped by major compound groups. The figure also shows the compound that exhibited the highest R^2 estimate in each major compound group.

6.3.2 Inducibility of chemical traits

Only 27 out of 54 compounds (50%) responded to treatment (Table 6.1) by increasing or reducing their amounts, but 39% (n = 21) retained their significance after Bonferroni adjustment. The strongest increment in the amount of compounds was detected for the phenolic compound; trans-ferulic acid [72] (Δ+39.58%, p<0.001). In contrast, the sugars reduced following treatment, where inositol [78], fructose [76] and glucose [77] reduced by 23.64%, 22.19% and 18.51% respectively (Table 6.1). Only 6 out of 54 (11%) compounds, comprising three monoterpenoids [1,5,7], a sesquiterpenoid [20] and two diterpenoids [24, 49] showed significant (p < 0.05) genetic differences in inducibility as indicated by the unadjusted p-values of the tree by treatment interaction term (Table 6.1). However, these interactions were not significant after Bonferroni correction. There is thus little evidence to suggest the presence of genetic variation in chemical inducibility and this term was not included in the genetic models used to estimate heritabilities.

6.3.3 Family (SCA) variation

Based on unadjusted probabilities, 30% of the compounds showed significant (P<0.05) non-additive genetic variation (i.e. family variation) including several monoterpenoids [1,3,4,5,7,9,10], a sesquiterpenoid

[20], diterpenoids [32,33,41,52,54], a phenolic compound [67], a fatty acid [84] and a compound of unknown class [87]. However, the family term, after Bonferroni adjustment, was not significant for any of the compounds (Table 6.1) so the family term was also excluded from the models used to estimate heritability.

6.3.4   *Genetic variation in P. radiata chemistry*

Using univariate models minus the family term and the random interaction, significant, adjusted levels of additive genetic variation were evident for most of the selected chemical compounds, with narrow-sense heritability values ranging between 0.00 - 0.48 and standard error between 0.02 - 0.13 (Figure 6.3, Table 6.1). No evidence of significant additive genetic variation was found for compounds with heritability values lower than 0.09. Only 12 compounds including two monoterpenoids, a sesquiterpenoid, four diterpenoids, three phenolic compounds, a sugar and an unknown compound did not show significant additive genetic variation.

Of the secondary compound groups, considering only compounds with significant additive genetic variation, the heritability of monoterpenoids ($\bar{x}$ = 0.27 ± 0.10) appeared to be consistently higher compared to sesquiterpenoids ($\bar{x}$ = 0.17 ± 0.08), diterpenoids ($\bar{x}$ = 0.25 ± 0.10) and phenolics ($\bar{x}$ = 0.13 ± 0.06). Generally, the heritability for the terpenoids was higher than that of the sugars ($\bar{x}$ = 0.22 ± 0.09) but not of fatty acids ($\bar{x}$ = 0.29 ± 0.10). There was no relationship between the univariate narrow-sense heritability estimate and the NIRS predictive accuracy for the 54 selected compounds as indicated by (i) NIRS coefficient of determination (R^2) and (ii) ratio of NIRS root mean square error (RMSE) relative to the laboratory error – PRL, and the range error ratio - RER (Figure 6.4). The standard error associated with the heritability estimates was in the same range irrespective of the NIRS model accuracy (Table 6.1).
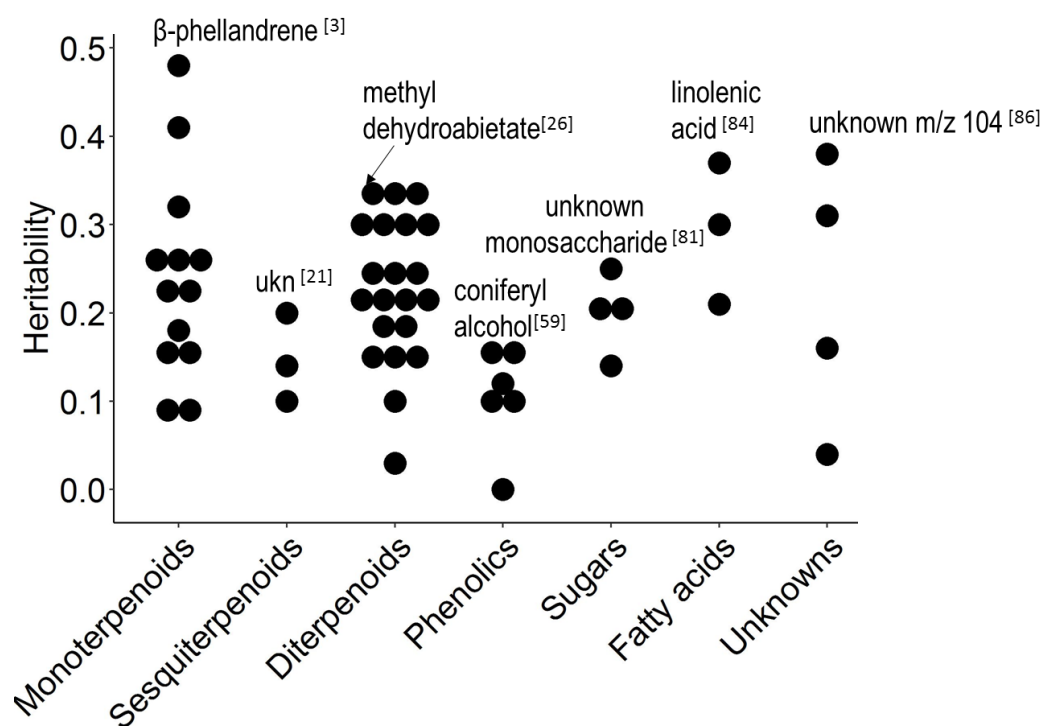
Figure 6.3: Dot plot of the distribution of estimated narrow-sense heritabilities for selected chemical compounds in the bark. 54 chemical compounds that had RER > 6 or PRL < 3 were included in the plot. Each dot represents a heritability estimate. The figure also shows the compound that exhibited the highest heritability estimate in each group, where ukn = unknown sesquiterpenoid alcohol.
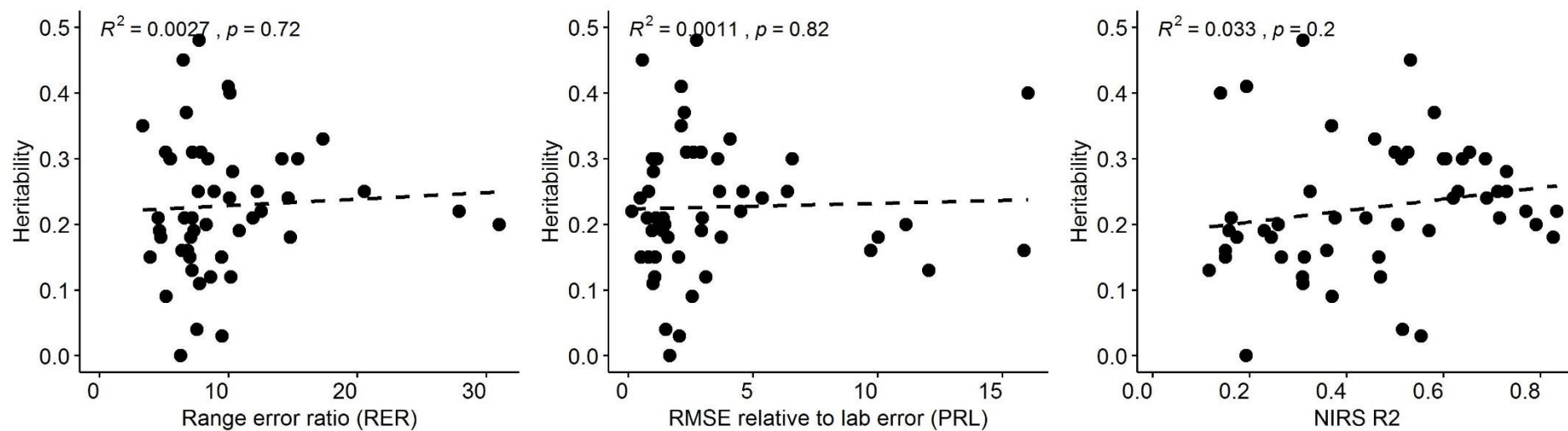
Figure 6.4: The linear relationship between heritability and the range error ratio (RER), ratio of NIRS root mean square error (RMSE) relative to the laboratory error (PRL) and NIRS coefficient of determination (R^2). Only compounds listed in Table 6.1 are included

Table 6.1: Mean and standard deviation (SD) of compounds, the significance of fixed and random terms, narrow-sense heritability (h^2) and genetic correlations from the univariate models fitted for the bark chemical components. For the various terms, the unadjusted probabilities are shown and bolded when significant following Bonferroni adjustment ($p<0.05$). The change that occurred in the different compounds following treatment (inducibility) is shown, where negative values (-) indicate reduction in the treated samples relative to the control and positive values indicate increase in the compounds relative to the control. The unadjusted likelihood ratio test (LRT) p-values associated with the significance of the random terms; differential inducibility (tree x treatment), family (specific combining ability) and tree (additive genetic variation) are also shown. The narrow-sense heritability and standard error (se) and the genetic correlation plus standard error (se) of the individual chemical compound with mammalian bark stripping and height are also included. The unadjusted p-value for the LRT test from zero are indicated for the genetic correlations (se) of the chemical compounds with bark stripping. For the correlation with height, only compounds that significantly correlated with bark stripping after Bonferroni adjustment were tested and unadjusted p- values are indicated in Table 6.2. The GC-MS components (monoterpenoids, sesquiterpenoids, GC-MS diterpenoids and phenolics) are expressed as micrograms of heptadecane equivalents (HE) per gram of dry weight of the sample (µg HE/g dw) and the LC-MS analytes (LC-MS diterpenoids and fatty acids) are expressed as micrograms of nonadecanoic acid equivalents (RE) per gram of dry weight of the sample (µg NE/g dw). Sugars are expressed in µg/g dw. All compounds were given a unique identifier based on Supplementary Table 10 (after Chapter 9), for ease of identification

id	Group		fixed effect (treatment-inducibility)		LRT p-values (adjusted) for random terms				genetic correlation with bark stripping(rg) ± se	P-value	genetic correlation with height (rg) ± se
		mean±SD	% inducibility	Wald p-value	differential inducibility (tree*treatment)	family	additive genetic variation	narrow-sense heritability ± se			
	monoterpenoids										
1	α-pinene	782.61±419.98	-4.26	0.666	0.021	0.040	**<0.001**	0.25±0.10	-0.20±0.29	0.498	0.02±0.33
2	α-terpineol	29.49±18.71	-2.63	0.737	0.063	0.070	0.045	0.08±0.07	-0.17±0.42	1.000	-0.56±0.45
3	β-phellandrene	80.19±39.76	-13.79	**<0.001**	0.258	0.021	**<0.001**	0.48±0.13	0.10±0.24	0.488	-0.02±0.29

158

4	β-pinene	1840.94±996.04	-0.36	0.913	0.103	0.007	**<0.001**	0.27±0.10	-0.01±0.28	1.000	-0.04±0.32
5	camphene	7.81±3.84	10.67	0.036	0.038	0.024	**<0.001**	0.26±0.10	-0.12±0.29	0.671	-0.00±0.33
6	citronellal	42.35±49.75	3.12	0.710	0.054	0.160	**<0.001**	0.18±0.09	-0.14±0.32	0.313	-0.32±0.36
7	citronellic acid	18.72±8.97	2.62	0.608	0.028	0.034	**<0.001**	0.32±0.11	0.08±0.27	1.000	0.03±0.31
8	citronellol	49.01±33.91	-5.68	0.287	0.058	0.054	**<0.001**	0.22±0.09	-0.13±0.29	1.000	-0.31±0.32
9	γ-terpinene	5.87±3.63	5.26	0.430	0.376	0.026	**0.001**	0.16±0.08	0.29±0.28	1.000	0.04±0.35
10	limonene	54.30±18.25	-1.24	0.562	0.189	0.008	**<0.001**	0.41±0.12	0.34±0.24	0.084	0.15±0.29
13	sabinene	149.82±91.47	2.51	0.715	0.344	0.139	**<0.001**	0.15±0.09	0.16±0.31	0.403	-0.01±0.37
15	terpinene-4-ol	26.88±16.59	-1.77	0.831	0.500	0.067	**<0.001**	0.23±0.09	0.32±0.27	0.260	-0.28±0.32
16	unknown Mol Wt 150	3.52±1.85	0.00	0.931	0.212	0.000	0.010	0.10±0.07	0.22±0.34	1.000	-0.02±0.40
	sesquiterpenoids										
17	bicyclogermacrene	1.93±0.75	5.26	0.494	0.231	0.500	0.022	0.10±0.05	0.09±0.30	0.888	-0.14±0.34
20	trans-farnesol	18.16±15.36	-18.27	**0.008**	0.018	0.018	**0.004**	0.14±0.08	0.08±0.35	0.752	-0.37±0.39
21	unknown sesquiterpenoid alcohol	4.37±1.68	9.52	**0.003**	0.444	0.500	**0.001**	0.20±0.09	-0.69±0.22	**0.008**	-0.85±0.22
	GC-MS diterpenoids										
22	agathadiol	550.26±506.94	18.94	0.054	0.123	0.186	0.003	0.21±0.10	-0.03±0.31	0.752	0.00±0.36
23	agatholal	340.57±213.96	8.91	0.160	0.324	0.500	**0.000**	0.22±0.09	-0.12±0.29	0.708	0.18±0.33
24	copalol	34.83±18.34	4.26	0.390	0.028	0.336	**<0.001**	0.29±0.10	-0.09±0.28	0.767	-0.20±0.32
25	levopimaral	13.34±7.42	-5.63	0.196	0.354	0.389	**<0.001**	0.30±0.10	0.08±0.27	0.762	0.04±0.31
26	methyl dehydroabietate	14.04±6.15	-14.01	**<0.001**	0.292	0.058	**<0.001**	0.34±0.11	0.22±0.25	0.298	-0.00±0.30
	LC-MS diterpenoids										
29	dehydroabietic acid	24704.5±5156.17	0.48	0.593	0.202	0.500	0.500	0.03±0.02	-0.27±0.37	0.462	0.39±0.39
32	unknown diterpene-3	184.25±116.46	-21.14	**<0.001**	0.189	0.005	**0.000**	0.15±0.09	0.30±0.29	0.247	0.03±0.35

33	unknown m/z 109 A	17.44±10.98	-6.28	0.204	0.086	0.006	**<0.001**	0.33±0.11	0.07±0.27	0.675	-0.11±0.31
34	unknown m/z 109 B	21.38±6.12	-2.31	0.166	0.500	0.139	**0.001**	0.19±0.09	0.52±0.24	0.030	0.77±0.26
37	unknown m/z 239	6.63±2.33	-5.80	0.021	0.240	0.095	0.107	0.10±0.07	0.35±0.30	0.237	0.70±0.30
38	unknown Mol Wt 272	7.79±2.91	-13.10	**<0.001**	0.254	0.071	**0.002**	0.15±0.07	0.34±0.28	0.237	0.30±0.34
41	C20H30O2 resin acids	26090.98±6003.77	-5.00	0.020	0.075	0.034	0.022	0.15±0.07	-0.35±0.31	0.273	-0.43±0.35
45	unknown m/z 316	13772.93±5570.38	4.12	0.257	0.224	0.259	**0.001**	0.21±0.09	-0.29±0.29	0.337	-0.05±0.34
46	unknown C20H30O3	25954.20±9277.74	-4.85	0.129	0.161	0.054	**0.001**	0.18±0.08	0.15±0.30	0.624	-0.03±0.35
47	unknown C20H32O3 A	692.45±303.45	5.13	0.240	0.011	0.133	**<0.001**	0.21±0.09	-0.04±0.30	0.888	-0.13±0.34
48	unknown C20H32O3 B	20689.06±10286.11	-11.08	**<0.001**	0.090	0.079	**<0.001**	0.25±0.10	0.37±0.24	0.077	0.56±0.25
50	unknown C20H30O4	60605.73±18178.65	-5.51	0.037	0.107	0.095	**0.000**	0.24±0.09	0.05±0.28	0.841	0.07±0.32
51	unknown C20H30O5	11354.00±4595.08	-0.03	0.959	0.115	0.107	**<0.001**	0.25±0.10	0.07±0.28	0.806	0.13±0.32
52	unknown C20H30O6 A	223.74±132.64	10.71	0.095	0.291	0.035	**<0.001**	0.31±0.11	0.18±0.27	0.517	0.01±0.31
54	unknown C20H30O6 C	6297.58±1894.94	2.65	0.412	0.110	0.009	**<0.001**	0.30±0.11	0.32±0.25	0.209	0.16±0.30
55	unknown C20H30O6 D	3583.38±1316.9	-6.60	0.008	0.500	0.078	**<0.001**	0.33±0.11	0.61±0.19	**0.006**	0.45±0.26
	phenolics										
56	anethole/estragole	1.63±0.65	-16.67	**<0.001**	0.145	0.208	**0.005**	0.12±0.06	0.00±0.58	0.752	0.43±1.00
59	coniferyl alcohol	4.63±1.23	6.67	**0.006**	0.274	0.254	0.037	0.16±0.08	0.34±0.27	0.237	0.39±0.33
67	phenyl ethanol	5.13±4.5	-5.66	0.218	0.500	0.036	0.431	0.00±0.02	0.52±0.56	1.000	0.15±0.61
71	thymol	7.92±5.66	-8.24	**<0.001**	0.500	0.154	0.103	0.09±0.08	0.30±0.33	1.000	-0.40±0.37
72	trans-ferulic acid	58.25±27.78	39.58	**<0.001**	0.205	0.500	**0.000**	0.11±0.06	0.48±0.27	0.051	0.32±0.35
74	vanillin	4.56±0.76	6.82	**<0.001**	0.500	0.500	**<0.001**	0.15±0.07	-0.28±0.30	1.000	0.15±0.34
	sugars										
76	fructose	13303.09±3607.02	-22.19	**<0.001**	0.500	0.500	**0.001**	0.21±0.09	0.55±0.23	**0.018**	0.06±0.31

77	glucose	15298.65±4309.32	-18.50	**<0.001**	0.376	0.500	**0.001**	0.20±0.09	0.80±0.20	**0.002**	0.62±0.24
78	inositol	10935.47±3551.37	-23.64	**<0.001**	0.354	0.500	0.060	0.14±0.08	-0.14±0.33	0.671	-0.01±0.38
82	unknown monosaccharide	521.33±236.64	24.12	**<0.001**	0.208	0.084	**0.003**	0.25±0.10	-0.47±0.28	0.066	-0.48±0.32
	fatty acids										
83	linoleic acid	16914.05±3747.38	-9.00	**<0.001**	0.500	0.500	**<0.001**	0.30±0.11	0.68±0.16	**0.001**	0.69±0.22
84	linolenic acid	7689.81±1454.91	1.73	0.430	0.500	0.015	**<0.001**	0.37±0.12	0.65±0.19	**0.004**	0.50±0.26
85	palmitic acid	16656.04±2440.36	4.05	**<0.001**	0.500	0.402	**0.004**	0.21±0.09	0.08±0.27	0.752	0.07±0.31
	unknowns										
86	unknown m/z 104	2.64±1.36	-16.67	**<0.001**	0.336	0.020	**<0.001**	0.38±0.12	0.42±0.22	**0.042**	0.27±0.28
90	unknown m/z 274	1115.47±398.55	-9.04	**0.004**	0.100	0.052	**<0.001**	0.31±0.11	0.03±0.28	0.888	0.05±0.32
94	unknown m/z 406 A	530.63±167.07	8.84	**0.001**	0.500	0.500	0.500	0.04±0.04	0.77±0.56	0.067	0.77±0.66
95	unknown m/z 406 B	5984.92±2308.79	8.03	0.021	0.346	0.500	**<0.001**	0.16±0.08	-0.46±0.27	0.058	-0.68±0.27

6.3.5 Traits genetically associated with mammalian bark stripping

A greater number of positive than negative genetic correlations between compounds and bark stripping were detected suggesting that preference may be a stronger driver for bark stripping than defence (Table 6.1; Figure 6.5). Significant unadjusted positive genetic correlations were detected between bark stripping and the sugars, glucose [77] (r_g = 0.80 ± 0.20, p <0.01) and fructose [76] (r_g = 0.55 ± 0.23, p <0.05; fatty acids- linoleic acid [83] (r_g = 0.68 ± 0.16, p <0.01) and linolenic acid [84] (r_g = 0.65 ± 0.19, p <0.01) as well as one diterpenoid – unknown C20H30O6 D [55] (r_g = 0.61 ± 0.19, p <0.01). The only significant negative genetic correlation was observed between bark stripping and an unknown sesquiterpenoid alcohol [21] (r_g = -0.69, p <0.05, unadjusted) (Figure 6.5, Table 6.1, Table 6.2).
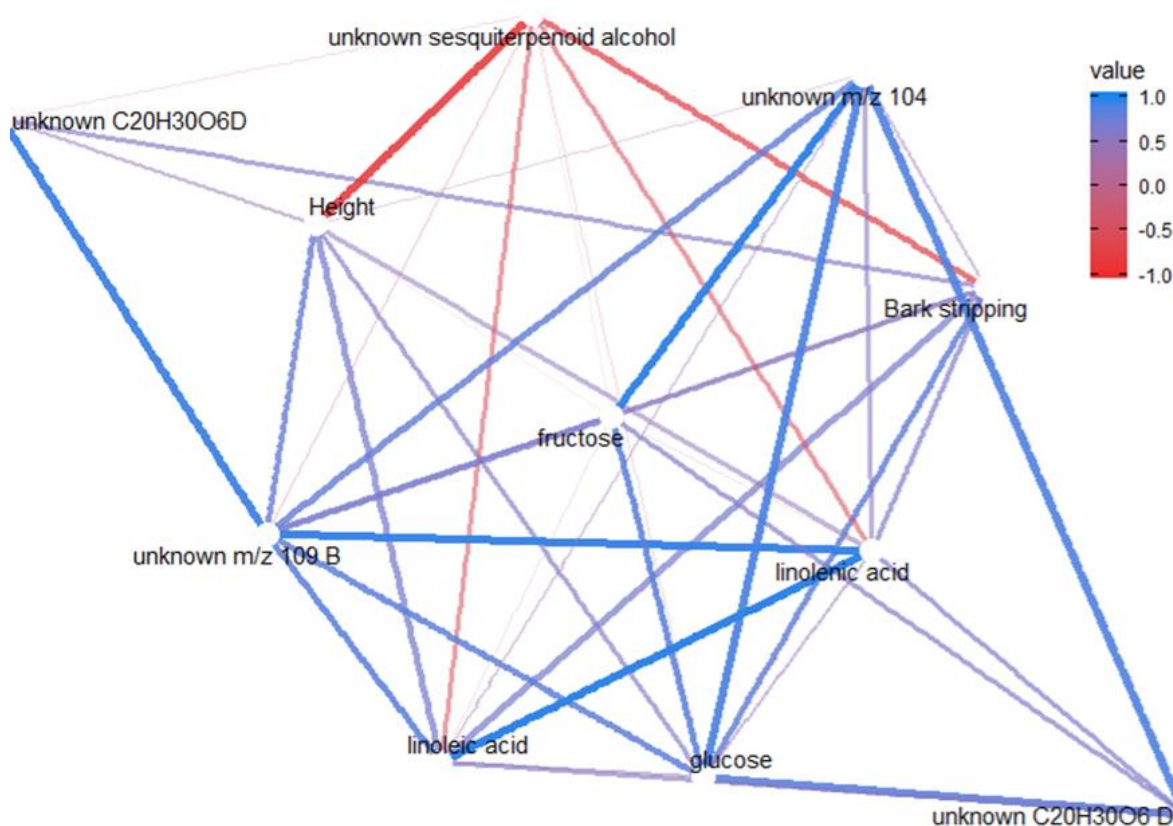
Figure 6.5: A network diagram showing the genetic correlations between different traits. Blue indicates a positive relationship and red indicates a negative relationship. A thicker line indicates a stronger correlation.

6.3.6 Genetic correlations among compounds and with height

In the trivariate models, the genetic correlation between mammalian bark stripping and height was positive, but non-significant (r_g = 0.40 ± 0.29, p =0.11). However, several of the chemical compounds correlated with bark stripping were genetically correlated with height. A significant negative genetic correlation was detected between the unknown sesquiterpene alcohol [21] and height (r_g = -0.85±0.22, $\chi 2$ = 7.1, p <0.01), suggesting that selecting for higher amounts of this compound will reduce growth in

the absence of herbivores (Table 6.2). Similarly, a positive correlation of sugars with height indicates that selecting for reduced sugars may result in reduced early growth. The strong positive genetic correlation between sugars and height shows that fast growing trees store more sugar in the bark than slow growing trees and conversely slow growing trees stored less sugar and more sesquiterpenes.

Among the compounds that significantly correlated with bark stripping, genetic correlations indicated that selecting for higher amounts of the unknown sesquiterpenoid alcohol will slightly reduce the amount of the fatty acid - linoleic acid. There was no evidence for a genetic correlation between the unknown sesquiterpenoid alcohol and glucose (Table 6.2). The positive genetic correlations also indicate that selecting for low sugars will shift the fatty acids and the unknown diterpenoid in the same direction, offering possibilities for multi-trait selection.

Phenotypic correlations between compounds showed similar trends as genetic correlations. However, significant correlations were detected at the phenotypic level that were not detected at genetic level (Table 6.1) suggesting that such correlations possibly pertain to this specific environment. Where genetic correlations were significant, the corresponding phenotypic correlations were smaller, except for the phenotypic correlations of the unknown sesquiterpenoid alcohol with glucose, fructose and the unknown diterpenoid (unknown $C_{20}H_{30}O_6$ D) that were higher than genetic correlations.

Table 6.2: Genetic (below) and phenotypic (above) correlations and the standard error of the compounds that significantly correlated with mammalian bark stripping. The phenotypic correlation between bark stripping with the compounds was not estimated since the chemistry was estimated in separate plots from where bark stripping was estimated (see methods). Unadjusted p-values are indicated

		Bark stripping	Height	unknown sesquiterpenoid alcohol	unknown C20H30O6 D	fructose	glucose	linoleic acid	linolenic acid
	Bark stripping								
	Height	0.40 (0.29)		0.24 (0.04) ***	-0.13 (0.04) **	-0.12 (0.00) *	0.36 (0.08)	-0.09 (0.00)	0.00 (0.00)
21	unknown sesquiterpenoid alcohol	-0.69 (0.22) *	-0.85 (0.22) **		-0.28 (0.12) ***	-0.37 (0.05) ***	-0.44 (0.05) ***	-0.18 (0.19) ***	-0.17 (0.05) **
55	unknown C20H30O6 D	0.61 (0.19) **	0.45 (0.26)	-0.15 (0.26)		0.28 (0.40)	0.41 (0.09) ***	0.48 (0.09) ***	0.35 (0.09) ***
76	fructose	0.55 (0.23) *	0.06 (0.31)	-0.06 (0.27)	0.64 (0.18) *		0.85 (0.03) ***	0.07 (0.05)	0.05 (0.05)
77	glucose	0.80 (0.20) **	0.62 (0.24) **	-0.19 (0.27)	0.85 (0.14) **	0.84(0.09) ***		0.07 (0.15) *	0.18 (0.12) ***
83	linoleic acid	0.68 (0.16) **	0.69 (0.22) ***	-0.51 (019) *	0.59(0.17) *	0.14 (025)	0.44 (0.23)		0.58 (0.06) ***
84	linolenic acid	0.65 (0.19) **	0.50 (0.26) **	-0.52 (0.21)	0.65 (0.16) **	0.12 (026)	0.47 (0.23)	0.97 (0.06) ***	

6.4   Discussion

The results of this study showed that; (i) most of the bark primary and secondary chemical compounds of *Pinus radiata* are under additive genetic control, with non-additive effects of little significance, (ii) there are weak family differences for inducibility, (iii) sugars and fatty acids genetically, positively correlate with mammalian bark stripping while an unknown sesquiterpenoid negatively correlated with bark stripping, (iv) the sesquiterpenoid alcohol negatively correlated with height while the remainder of the compounds positively or did not correlate with height. Genetic differences in the constitutive and induced variation in patterns of secondary and primary metabolites have been noted in earlier studies in *P. radiata* and other pine species (Sampedro *et al.* 2010; Zhang *et al.* 2016a). In *P. radiata*, the presence of additive genetic variation has been previously detected for total terpenes and total phenolics (Moreira *et al.* 2012a; Moreira *et al.* 2013b) as well as a few individual monoterpenoids (Burdon *et al.* 1992a; Burdon *et al.* 1992b). However, with the large number of families tested in this study, this is the first to estimate narrow-sense heritability and genetic correlations for individual primary and secondary compounds in *P. radiata* bark*.*

Within the secondary compound groups, the monoterpenoids appeared to have higher narrow-sense heritability compared to diterpenoids and phenolics. Differences in additive genetic variances of traits may be explained by the importance of these traits to fitness, whereby significant negative correlations exist between heritability and measures of fitness for different traits (Mousseau and Roff 1987). Following this theory, this suggests that diterpenoids and phenolics may be more related to fitness consequences in *P. radiata* in this study. The physiological role of diterpenoids is not well established, but given the diverse direct and indirect roles of phenolics in constitutive or inducible stress responses, particularly in lignin and pigment biosynthesis and in their ability to repel or kill many microorganisms, as well as their roles in cell division, hormonal regulation, photosynthetic activity and nutrient mineralization and direct (Bhattacharya *et al.* 2010; Hammerbacher *et al.* 2011), they may be important for plant fitness. In most conifers, constitutive and induced monoterpenoids have been implicated in defence against pathogens, insect pests as well as some mammalian herbivores (Bucyanayandi *et al.* 1990; Iason *et al.* 2011; O'Reilly-Wapstra *et al.* 2007; Vourch *et al.* 2002) compared to other terpenoids or phenolics. If these results have fitness consequences, we could expect that the heritability of monoterpenoids should be lower. The higher heritability of monoterpenoids in this case could suggest that, survival of *P. radiata* was more related to defence against abiotic than against biotic stresses since naturally the trees grew on nutrient impoverished and often acidic soils (McDonald and Laacke 1990), and hence the higher importance (interpreted by the low heritability) of phenolic compounds.

Differences in heritability can also arise from differences in the levels of environmental or residual variance, nonadditive genetic variance or a combination of all these effects, rather than from different levels of additive genetic variation (Price and Schluter 1991). Although nonadditive genetic variation was not significant for the traits in this study, accounting for both spatial (i.e. between-population variation) and temporal (i.e. cohort variation) heterogeneity in environmental conditions may improve the estimates (Chapter 2). It has also been noted for secondary compounds that the relative amount of additive genetic variation may be related to the relative amount of compounds harboured by the plant, where metabolites that occur in higher amounts have been found to have higher heritability estimates than those with small amounts (Haviola *et al.* 2006). For *P. radiata* bark, the amounts of monoterpenoids often dominate the other terpenoid components (Chapter 3), which may explain their high heritability estimates compared to other secondary metabolites. While this may be linked to the accuracy of quantification using conventional wet chemistry or spectroscopic methods (Chapter 5), there was no link between NIRS accuracy and estimates of genetic parameters in this study. Overall, while heritability estimates for different secondary compounds may be variable between studies, the low to moderate heritability values from different studies (Sampedro *et al.* 2010; Zhang *et al.* 2016a) indicate that the secondary compounds in the bark of *P. radiata* and other conifers may have important fitness consequences and demonstrate sufficient additive variance to be potentially responsive to natural or artificial selection.

The average narrow-sense heritability estimates of the sugars was lower than the terpenoids consistent with the theory that low heritability may result from intensive natural selection for traits that are tightly coupled with fitness (Mousseau and Roff 1987). Sugars are at the core of plant survival and development, as sugars and sugar derivatives are the conduit of the carbon fixed during photosynthesis (Patrick *et al.* 2013). However, sugars are also very sensitive to environmental changes (Bansal and Germino 2009; Cranswick *et al.* 1987), which can impact heritability estimates. Low genetic variation for bark and wood sugars has been detected in other *P. radiata* studies (Cranswick *et al.* 1987; Donaldson *et al.* 1997) and in other conifers species, observations have been mixed. For example while no genetic variation was observed for sugars in the bark of juvenile *Pinus pinaster* (Sampedro *et al.* 2011), strong genetic control of glucose has been observed in the wood of *Pseudotsuga menziesii* (Ukrainetz *et al.* 2008). These counterexamples may also suggest tissue and species-specific differences in additive genetic variation for sugars.

Fatty acids exhibited the highest average heritability compared to other compound groups, which may suggest that they are of relatively low importance for *P. radiata* fitness. However, the negative genetic

correlation of the fatty acid linoleic acid and the unknown sesquiterpenoid alcohol suggests that fatty acids may be partly linked to the formation of sesquiterpenoids. The storage and structural functions, as well as direct defence properties against pathogens of fatty acids have been documented in the needles of *P. radiata* (Franich *et al.* 1983). Actually some of the most studied signalling molecules like Jasmonic acid belong to a group of compounds formed by the oxygenation of fatty acids (Kachroo and Kachroo 2009), emphasizing the role of fatty acids in stress responses. In this Moreira *et al.* (2012a)study, fatty acids were also positively associated with bark stripping. The narrow-sense heritability for fatty acids in the bark has not been reported for most conifers. Overall, the results suggest that selection against primary compounds especially the sugars can be incorporated in the current *P. radiata* breeding/deployment programmes to reduce bark stripping by the marsupials. However, studies have noted that as environments change, heritability may change, and therefore testing the expression of the additive genetic variation in different environments, to get an understanding of the genotype by environment interactions (G x E) will facilitate the incorporation of the traits in different breeding programmes.

Changes in the amounts of primary and secondary compounds were observed in response to artificial bark stripping. This is consistent with previous studies in this thesis and other *P. radiata* studies that show a reduction in the amounts of sugars and an increase in some of the secondary compounds after treatment (Chapter 3 ; Chapter 4 ; Moreira *et al.* 2012a; Sampedro *et al.* 2011). However, genetic differences in inducibility did not appear to be evident for individual compounds in the present study despite an earlier study (Chapter 4) illustrating that the amounts of individual terpenes, phenolics and sugars reduced or increased differentially between susceptible and resistant families. While several compounds showed genetic variation in the induced response to bark stripping, the effect was weak and not significant after statistical correction for multiple testing. In contrast, Moreira *et al.* (2013b) and Sampedro *et al.* (2011) found high genetic variation in inducibility of stem resins in *P. radiata* and diterpenoids in *P. pinaster* respectively. Presence of genetic variation in inducibility suggests that this trait can be selected for. However, overall, the results for the populations studied in this thesis suggest that selection for reduced susceptibility of *P. radiata* to bark stripping is more feasible based on the constitutive than the induced chemistry.

Similar to ealier observations based on the top susceptible and resistant families (Chapter 4), the sugars; fructose and glucose positively correlated with bark stripping, which emphasizes the genetic basis of this association to susceptibility. Of the secondary compounds, an unknown sesquiterpenoid alcohol best correlated with reduced bark stripping and this is consistent with earlier characterisation

that identified the importance of this compound in differentiating the more and less susceptible groups of families (Chapter 4). In contrast, the sesquiterpenoid - bicylogermacrene that was the major compound that differentiated the less and more susceptible families in Chapter 4 did not exhibit significant additive genetic variation and did not genetically correlate with bark stripping. This could possibly be due to non-linear genetic associations with bark stripping. Non-linear genetic relationships, where the range of a trait varies drastically from one extreme to another have been detected in *Arabidopsis thaliana* (Vasseur *et al.* 2019) and this may have affected the genetic parameters for bicyclogermacrene that were estimated based on linear models. Sometimes, bivariate correlations may give more positive results than a model with additional covariates, especially where the extra variables do not strongly correlate with the two major variables of the model (Song *et al.* 2019). In this study, genetic correlations between chemistry and bark stripping were estimated in trivariate models with bark stripping, a chemical compound and height as response variables, and because no bark stripped trees were chemotyped it was not possible to undertake bivariate genetic correlations between bark stripping and chemical compounds. So possibly an experimental design that favors this analysis may give further insights into the correlations. The only terpenoid that significantly correlated was one unknown diterpenoid that genetically, positively correlated with bark stripping. This diterpenoid was also highlighted earlier (Chapter 4) where its amount was higher in the susceptible compared to the resistant families (although this was non-significant). This positive association contrasts with the documented role of diterpenoids in reducing herbivory in conifers (Franceschi *et al.* 2005) and the fitness-related properties gestured by the low heritability estimates of diterpenoids as described above. This observation may be in part due to the capability of the marsupials to ingest and metabolise a range of terpenes that would be toxic to many other herbivore species (Boyle 1999; El-Merhibi *et al.* 2007). It could also be related to the sample size used in the study since accurate estimation of genetic correlation requires large sample sizes. It could also be due to the nature of experimental design that could not allow direct correlation of individual chemistry and bark stripping, which could have introduced more errors in the correlations. Consistently, across this system there was no evidence for the importance of monoterpenoids in determining bark stripping.

There was evidence of defence-growth trade-off for the unknown sesquiterpenoid alcohol implying that this compound will potentially be reduced via indirect responses to selection in *P. radiata* programmes where height is targeted as a key breeding objective. The trade-offs were also implicated earlier (Chapter 4). Similarly, the positive correlation between sugars and height suggests that if bark stripping is mainly driven by the sugars, positive selection for early growth in the presence of bark stripping will increase the vulnerability of the populations to bark stripping. Positive correlations between herbivory of

the bark and height are common in conifers (Lenz *et al.* 2020; Zas *et al.* 2017) and may be explained by fast-growing trees potentially investing less in secondary compounds especially in the presence of resource constraints (Ferrenberg *et al.* 2015; Moreira *et al.* 2015). In this study, genetic correlations suggest that fast-growing trees invested less in secondary compounds and more in sugars. The direction of the correlation in this study was similar to the univariate models presented in Chapter 2 and non-parametric tests of Chapter 4 where a significant positive correlation between height and bark stripping was observed in the same population. However, in the linear models that included three response variables -height, bark stripping and chemical compounds as response variables, bark stripping was not significantly correlated with height. This may imply that height is an associational rather than independent predictor of susceptibility, especially given its positive correlation with bark glucose.

The genetic correlations between individual compounds were mostly positive or non-significant, except for the negative correlation between the unknown sesquiterpenoid alcohol and the fatty acid, linoleic acid. Positive genetic correlations suggest simultaneous investment in multiple traits and is common for traits that interact together to perform a given function, and hence the potential for multi-trait selection. For chemical defences, this is important since defence is potentially achieved by multiple interdependent primary and secondary compounds (Chapter 4). On the other hand, negative genetic correlations indicate trade-offs between investment in the affected traits (Moreira *et al.* 2015). In this study, for instance, to invest in higher amounts of the defence compound (unknown sesquiterpenoid) will come at the expense of fatty acids, in presence of limited resources. This observation may explain the strong reduction in fatty acids detected in Chapter 3, consistent with suggestions that fatty acids can be precursors to the formation of secondary compounds (Kachroo and Kachroo 2009). The lack of significant negative correlations of the secondary compounds with sugars in this and earlier studies (Chapter 3) contradicts models of physiological trade-offs which postulate that non-structural carbohydrates (NSCs), e.g. glucose, are pivotal in defence-growth relationships (Herms and Mattson 1992; Lombardero *et al.* 2000; Moreira *et al.* 2015; Sampedro *et al.* 2011). In this case, fatty acids were more traded for defence functions. Although, a negative correlation between height and the unknown sesquiterpenoid was detected, this growth-defence trade-off were not directly linked to the sugars in this study.

6.5 Conclusion

There is significant additive genetic variation for primary and secondary compounds in the bark of *P. radiata* with low-moderate heritability estimates. While glucose and fatty acids predicted susceptibility, the unknown sesquiterpnoid alcohol was a strong predictor of reduced bark stripping. The unknown sesquiterpenoid alcohol was genetically, negatively correlated with height whereas glucose and fructose as well as the fatty acids genetically, positively correlated with height, suggesting that positive selection for early-age height would shift the chemistry of the plants towards increased susceptibility. However, whether or not these traits affect performance subsequent to browsing needs testing. The use of NIR offers opportunities for large scale chemical phenotyping and has allowed genotyping for chemical traits with results conforming to observations obtained using standard wet chemistry procedures. Nevertheless, there is still a need for calibration improvement for most of the compounds which may be achieved by increasing sample size.

Supplementary Table 6.1: The statistics used to select the best models, the range error ration-RER, prediction relative to the lab error -PRL, and Near infrared spectroscopy coefficient of determination -R2, in the different bark tissue types. Final prediction was based on the models developed with the dried-ground bark. The bold compounds were dropped from the final genetic analyses since their model parameters did not meet the set criteria. Compound identifiers (ID) were given to each compound for ease of location based on Supplementary Table 10 (after Chapter 9)

ID	Compound	dried ground bark			Inner bark proximal			Inner bark distal			Outer bark proximal			Outer bark distal		
		RER	PRL	R^2	RER	PRL	R^2	RER	PRL	R^2	RER	PRL	R^2	RER	PRL	R^2
	monoterpenoids															
1	α-pinene	7.63	0.81	0.73	4.38	1.42	0.14	4.47	1.39	0.18	4.08	1.52	0.00	4.11	1.51	0.02
2	α-terpineol	5.15	2.56	0.37	5.02	2.63	0.09	5.00	2.64	0.08	4.77	2.77	-0.01	4.75	2.78	-0.02
3	β-phellandrene	7.72	2.74	0.31	5.44	3.89	0.00	5.52	3.83	0.02	5.57	3.80	0.04	5.51	3.84	0.02
4	β-pinene	10.30	1.01	0.73	5.35	1.94	0.14	5.66	1.84	0.22	5.15	2.02	0.05	6.51	1.60	0.14
5	camphene	5.50	1.14	0.60	4.96	1.27	0.23	4.43	1.42	0.24	3.93	1.60	0.03	4.65	1.35	0.11
6	citronellal	8.89	3.65	0.63	10.98	2.95	0.21	6.95	4.67	0.13	6.88	4.71	0.10	6.57	4.94	0.03
7	citronellic acid	7.85	2.91	0.50	7.07	3.23	0.25	5.37	4.25	0.10	5.13	4.45	0.01	5.13	4.45	0.01
8	citronellol	4.64	3.78	0.48	4.08	4.29	0.17	4.03	4.34	0.15	3.86	4.53	0.07	3.74	4.68	0.01
9	γ-terpinene	7.07	10.01	0.17	7.07	10.01	-0.01	7.00	10.11	-0.02	7.09	9.98	-0.01	7.07	10.01	-0.02
10	limonene	9.98	2.11	0.19	3.84	5.49	0.01	3.84	5.49	0.01	3.84	5.49	0.00	3.80	5.54	-0.01
11	**linalool**	4.99	6.95	0.14	5.05	6.86	0.04	4.42	7.85	0.24	3.93	8.82	0.02	4.00	8.68	0.06
13	sabinene	6.37	15.85	0.15	6.50	15.52	0.00	6.42	15.71	-0.02	6.71	15.04	0.05	6.40	15.76	-0.03
14	**terpinolene**	5.81	18.94	0.14	6.14	17.93	-0.01	6.14	17.93	-0.01	6.23	17.68	0.00	6.10	18.06	-0.03
15	terpinene-4-ol	8.27	11.12	0.26	7.95	11.57	-0.01	7.88	11.68	-0.03	7.99	11.51	-0.01	7.91	11.62	-0.02
16	unknown Mol Wt 150	7.15	12.05	0.12	6.93	12.43	0.01	6.87	12.55	-0.02	6.95	12.40	0.01	6.88	12.52	-0.01
	sesquiterpenoids															
17	bicyclogermacrene	6.99	0.80	0.15	6.61	0.85	0.01	6.75	0.83	0.05	6.70	0.84	0.03	6.70	0.84	0.03
20	trans-farnesol	74.55	1.25	0.51	12.53	7.43	-0.01	12.64	7.36	-0.01	12.64	7.36	-0.01	12.53	7.43	-0.01
21	unknown sesquiterpenoid alcohol	7.18	1.11	0.16	6.88	1.16	0.00	6.88	1.16	0.00	6.94	1.15	0.02	9.83	0.81	0.08
	GC-MS diterpenoids															
22	agathadiol	10.81	0.94	0.57	7.85	1.30	0.24	7.27	1.40	0.11	9.89	1.03	0.17	6.87	1.49	0.01
23	agatholal	14.60	0.47	0.62	7.94	0.87	0.22	7.49	0.92	0.12	10.60	0.65	0.21	7.05	0.98	0.01
24	copalol	5.46	0.98	0.64	3.81	1.41	0.16	3.94	1.36	0.21	3.55	1.51	0.02	3.49	1.54	0.00
25	levopimaral	5.12	2.34	0.53	4.88	2.46	0.08	5.17	2.32	0.18	4.79	2.50	0.03	4.66	2.57	0.00
26	methyl dehydroabietate	3.35	2.12	0.37	3.28	2.16	0.00	3.28	2.16	0.01	3.25	2.18	0.04	3.35	2.12	0.04
	LC-MS diterpenoids															
29	dehydroabietic acid	9.49	2.05	0.55	5.95	3.28	-0.10	5.91	3.30	-0.11	6.07	3.21	-0.05	5.92	3.30	-0.11

32	unknown diterpene-3	4.63	1.38	0.23	3.78	1.69	0.05	3.71	1.72	0.02	3.74	1.70	0.04	3.68	1.73	0.00
33	unknown m/z 109 A	6.48	0.56	0.53	4.88	0.74	0.03	5.32	0.68	0.18	8.62	0.42	0.29	4.80	0.75	0.00
34	unknown m/z 109 B	7.29	2.93	0.16	6.83	3.13	0.00	6.91	3.09	0.02	6.75	3.17	0.01	6.87	3.11	0.02
37	unknown m/z 239	7.76	0.98	0.31	7.14	1.07	0.11	7.12	1.07	0.10	6.87	1.11	0.03	7.08	1.08	0.09
38	unknown Mol Wt 272	66.17	0.51	0.31	12.28	2.78	0.00	12.36	2.76	0.02	12.36	2.76	0.01	31.45	1.08	0.05
41	C20H30O2 resin acids	9.46	2.01	0.47	8.36	2.27	0.22	7.17	2.65	0.07	6.98	2.72	0.03	6.89	2.75	-0.01
43	**unknown m/z 304 B**	2.00	16.30	0.25	2.64	12.34	0.00	2.64	12.37	0.00	2.65	12.30	0.00	2.64	12.37	0.00
45	unknown m/z 316	11.87	2.97	0.72	5.56	6.33	-0.05	5.59	6.30	-0.04	5.79	6.08	0.04	5.67	6.21	-0.01
46	unknown C20H30O3	14.79	3.72	0.83	6.03	9.13	-0.02	6.08	9.05	0.00	6.34	8.69	0.08	6.13	8.98	0.01
47	unknown C20H32O3 A	12.52	4.50	0.83	7.73	7.30	0.26	5.68	9.92	0.06	7.17	7.86	0.17	5.78	9.75	0.10
48	unknown C20H32O3 B	20.52	6.38	0.33	6.47	20.22	-0.02	6.47	20.22	-0.02	6.47	20.22	-0.01	6.47	20.22	-0.01
50	unknown C20H30O4	10.08	5.37	0.69	5.61	9.63	0.10	5.38	10.05	0.03	5.64	9.59	0.11	5.40	10.01	0.04
51	unknown C20H30O5	12.24	4.59	0.71	5.82	9.64	0.00	5.75	9.76	-0.02	6.25	8.98	0.13	5.87	9.56	0.02
52	unknown C20H30O6 A	14.14	6.58	0.6	13.32	6.98	0.07	8.52	10.91	-0.03	8.68	10.71	0.00	8.57	10.85	-0.01
54	unknown C20H30O6 C	15.36	3.58	0.51	8.09	6.80	0.21	7.18	7.66	0.00	8.07	6.82	0.20	7.24	7.61	0.01
55	unknown C20H30O6 D	17.31	4.08	0.46	9.35	7.56	0.21	8.33	8.49	0.01	9.14	7.73	0.17	8.30	8.52	0.00
	phenolics															
56	anethole/estragole	10.18	3.12	0.47	5.58	5.69	-0.01	6.64	4.78	-0.04	7.03	4.51	0.17	6.28	5.05	0.05
57	**benzene acetic acid**	5.95	3.33	0.25	4.76	4.17	0.10	4.77	4.15	0.11	4.96	4.00	0.17	4.62	4.29	0.05
59	coniferyl alcohol	6.82	9.71	0.36	6.05	10.94	0.10	6.39	10.37	0.19	5.89	11.25	0.10	6.18	10.72	0.13
61	**ethyl phenol**	1.51	NA	0.2	1.51	NA	-0.05	1.51	NA	-0.05	1.31	NA	-0.03	1.51	NA	-0.05
63	**4-ethyl guaiacol**	689.72	0.02	0.23	4.95	3.06	-0.05	4.95	3.06	-0.05	4.99	3.04	-0.03	4.92	3.09	-0.06
65	**methyl eugenol**	0.03	220.59	0.22	4.46	1.52	0.06	4.68	1.45	0.15	4.40	1.54	0.03	4.68	1.45	0.14
66	**p-Menth-1-en-7,8-diol**	2.64	1.60	0.12	4.92	0.86	0.18	4.30	0.98	0.08	0.00	0.00	0.00	4.21	1.01	0.04
67	phenyl ethanol	6.26	1.66	0.19	6.71	1.55	0.00	6.60	1.58	-0.03	6.65	1.56	0.00	6.64	1.57	-0.02
68	**pinosylvin dimethyl ether**	5.81	7.62	0.07	6.27	7.07	0.01	6.27	7.07	0.01	6.27	7.06	0.01	6.24	7.10	0.01
70	**raspberry ketone**	3.52	13.52	0.15	4.42	10.75	0.13	4.14	11.48	0.01	4.46	10.67	0.17	4.12	11.53	0.00
71	thymol	4.52	0.75	0.38	5.78	0.59	0.14	6.02	0.56	0.20	5.50	0.62	0.04	5.67	0.60	0.10
72	trans-ferulic acid	8.59	1.07	0.31	6.58	1.39	0.12	6.41	1.43	0.07	6.76	1.36	0.16	6.52	1.41	0.10
74	vanillin	3.93	1.08	0.27	3.77	1.13	-0.05	4.05	1.05	0.08	3.56	1.20	-0.02	4.47	0.95	0.02
	sugars															
76	fructose	10.55	1.63	0.77	0.06	294.10	0.31	6.53	2.62	0.40	5.56	3.08	0.30	7.36	2.33	0.52
77	glucose	11.12	1.76	0.79	0.07	295.27	0.40	6.65	2.94	0.40	9.23	2.12	0.49	6.95	2.81	0.45
78	inositol	0.01	34.81	0.64	0.00	5668.48	0.05	0.01	48.14	0.31	6.25	0.05	0.01	0.01	41.67	0.46
79	**sucrose**	7062.29	0.47	0.11	70.25	47.22	0.08	6708.29	0.49	0.00	6.10	544.04	0.03	6658.23	0.50	-0.02

81	unknown monosaccharide	7.25	NA	0.26	0.07	NA	0.01	7.16	NA	0.00	7.40	NA	0.06	7.23	NA	0.02
	fatty acids															
82	linoleic acid	8.40	0.97	0.69	4.25	1.91	-0.04	4.20	1.93	-0.06	4.24	1.91	-0.02	4.25	1.91	-0.04
83	linolenic acid	6.72	2.24	0.58	4.36	3.44	-0.05	4.25	3.53	-0.10	4.32	3.48	-0.08	4.32	3.48	-0.08
84	palmitic acid	6.57	1.39	0.44	4.40	2.07	-0.14	4.34	2.10	-0.17	4.38	2.08	-0.12	4.43	2.06	-0.13
	unknowns															
86	unknown m/z 104	10.11	16.02	0.14	9.90	16.36	0.02	9.86	16.42	0.01	10.01	16.17	0.04	9.88	16.39	0.02
89	unknown m/z 274	7.20	2.62	0.65	4.12	4.58	0.00	4.05	4.66	-0.04	4.35	4.34	0.09	8.14	2.32	0.03
92	unknown m/z 406 A	7.54	1.50	0.52	4.75	2.38	-0.12	5.25	2.16	0.09	5.25	2.16	0.11	4.80	2.36	-0.10
93	unknown m/z 406 B	4.74	1.59	0.24	4.12	1.83	-0.01	4.26	1.77	0.05	4.09	1.84	0.01	4.12	1.83	-0.01

Chapter 7: Genomic selection for resistance to bark stripping and associated chemical compounds in radiata pine

7.0   Abstract

The integration of genomic data into genetic evaluations can facilitate the rapid selection of superior genotypes and accelerate the breeding cycle in trees. In this chapter, 420 controlled-pollinated trees from 74 families were genotyped using a 50K axiom SNP chip array. A total of 15,624 high-quality SNPs were used to develop genomic prediction models for bark stripping, height and selected primary and secondary chemical compounds in the bark. Genetic parameters from different genomic prediction methods; univariate best linear unbiased prediction based on a genomic derived additive relationship matrix (GBLUP), trivariate single-step GBLUP which integrated the genomic and pedigree derived additive relationship matrix (ssGBLUP) and the univariate generalised ridge regression (GRR) were compared to equivalent univariate or trivariate pedigree-based predictions (ABLUP). The influence of the statistical distribution of data on the genetic parameters was assessed. Results indicated that the heritability estimates were improved by up to 2-fold with genomic models compared to the equivalent pedigree based ABLUP models. However, the predictive ability of the ssGBLUP was not markedly different from the pedigree-based model (ABLUP) for most traits. Compared with GBLUP, allowing for heterogeneity in marker effects through the use of GRR did not markedly improve predictive ability over GBLUP, arguing that most of the chemical traits are modulated by many genes with small effects. Overall, the traits with low ABLUP heritability benefited more from genomic models compared to the traits with high ABLUP heritability. There was no evidence that data skewness or presence of outliers affected the genomic or pedigree-based genetic estimates.

7.1　Introduction

The implementation of genomic prediction in plants has offered new possibilities for maximising genetic gains for economically important traits (Crossa *et al.* 2017; Meuwissen *et al.* 2001) and can potentially enhance the efficiency of selecting herbivory-resistant phenotypes in conifers. In conifers, breeding for resistance against pests and diseases has mainly relied on conventional phenotype-based methods (Alfaro *et al.* 2004; Carson 1989) and has been facilitated by quantitative genetic studies that investigate the genetic basis of the resistance mechanisms. Although results from these studies mostly indicate that resistance and the associated chemical traits are under genetic control and can respond to selection, often low narrow-sense heritability estimates (Chapter 2 ; Chapter 5 ; Moreira *et al.* 2013b; Zas *et al.* 2017) reduce the precision of breeding value predictions. In addition, the inherently long generation intervals of trees and high phenotyping costs are always a challenge in tree breeding. Therefore, the potential improvement in prediction accuracy of breeding values for traits with low heritability (Gamal El-Dien *et al.* 2016; Goddard 2009; Hayes *et al.* 2009; Iwata *et al.* 2011; Klápště *et al.* 2018; Stejskal *et al.* 2018; Suontama *et al.* 2018), coupled with the predicted reduction in the length of breeding cycles (Klápště *et al.* 2018; Thistlethwaite *et al.* 2017), should be a major motivation for incorporating genomic selection in breeding for resistance in conifers. However, the effectiveness of genomic prediction depends on the improvement in the accuracy of breeding value predictions.

The factors that affect the accuracy of breeding value predictions and, hence, the expected response to genomic-informed selection have been well documented (Desta and Ortiz 2014; Klápště *et al.* 2018; Momen *et al.* 2018; Stejskal *et al.* 2018). A key factor, however, is the choice of statistical methods for genomic estimated breeding value predictions that differ with respect to the assumptions regarding the distribution of marker effects. Methods such as the genomic best linear unbiased prediction (GBLUP), single-step GBLUP (ssGBLUP) and ridge regression best linear unbiased prediction (RR-BLUP) assume that all marker effects follow the same distribution and each explains a very small amount of variance (Legarra *et al.* 2009; Meuwissen *et al.* 2001; Misztal *et al.* 2013). In contrast, linear regularized (penalized) regression models such as generalised ridge regression (GRR), least absolute shrinkage and selection operator (LASSO) and elastic net as well as nonlinear Bayesian methods that include Bayes A/B/C/Cπ/R account for heterogeneity of marker effects (Meuwissen *et al.* 2017). Studies that have evaluated the relative predictive performance of the different approaches mostly indicate that the optimum approach is partly dependent on the genetic architecture and heritability of the trait involved (Hayes and Goddard 2001; Momen *et al.* 2018; Ogutu *et al.* 2012; Ratcliffe *et al.* 2017; Wang *et al.* 2018a). Genetic architecture describes genotype-phenotype

relationships for the loci contributing to phenotypic variation, and includes the number of loci and their genomic location, number of alleles per locus, magnitude of their effects, patterns of pleiotropy, mode of gene action and epigenetic effects (Momen *et al.* 2018). Complex quantitative traits that are controlled by many genes with small effects, such as likely for resistance to herbivory (Kliebenstein 2014; Lenz *et al.* 2020), may be better predicted by methods that do not prioritize individual genetic markers (Desta and Ortiz 2014; Lenz *et al.* 2020). In conifers, however, studies also indicate that some herbivory resistance, as well as associated chemical resistance traits, may be controlled by genes with major effects (Porth *et al.* 2011). For such traits that are controlled by major genes, prediction accuracy can be favourably estimated by models that apply variable selection and differential shrinkage of allelic effects (Gianola *et al.* 2009; Resende *et al.* 2012a). Additionally, for resistance which is mostly scored on qualitative or semiquantitative scales, and chemical data that is skewed (Burdon *et al.* 1992a), regression models that support non-normal data may be more appropriate (Kärkkäinen and Sillanpää 2012). Therefore, for less studied traits, it is important to evaluate a broad range of statistical methods to identify those that can better model the additive genetic variance. Accordingly, we tested the ability of three selected statistical methods - GBLUP, ssGBLUP and GRR - to estimate breeding values for, (i) resistance of radiata pine (*Pinus radiata* D. Don) to marsupial bark stripping and (ii) the bark chemical traits, some of which may impact susceptibility. These methods were selected to represent alternative approaches and assumptions related to marker effects (Meuwissen *et al.* 2017; Wang *et al.* 2018b).

In radiata pine, selection for pest resistance has mostly been based on conventional approaches involving visual selection and trait screening over several pedigreed generations (Carson 1989; Dungey *et al.* 2009). Although there does not appear to be operational breeding programmes focussed on reducing the susceptibility of radiata pine to herbivory, various quantitative genetic studies have indicated the potential for selection against insect and mammalian bark damage and the associated chemical traits in various radiata pine populations (Chapter 2 ; Chapter 5 ; Chapter 6 ; Moreira *et al.* 2013b). However, bark stripping and associated chemical traits, that include terpenes, phenolics and sugars, have low to moderate pedigree-based heritability estimates (Chapters 2, 6) and would possibly benefit from genomic selection. In conifer-herbivore systems, only one study has examined the potential benefits of genomic selection, in this case for white pine weevil resistance in Norway spruce (Lenz *et al.* 2020). To my knowledge, there is no study incorporating genomic selection for chemical compounds related to susceptibility or resistance mechanisms in breeding programmes of conifers. For other economically important traits in radiata pine, studies have indicated potential genetic gain from using genomic selection (Li and Dungey 2018; Whitehill *et al.* 2016) and similar concepts could be

adopted for resistance breeding. Currently, there are marker panels to identify known biotic threats such as *Dothistroma pini* (Li *et al.* 2015b) and other genomic resources in radiata pine (Telfer *et al.* 2019; Telfer *et al.* 2018) that can facilitate detailed genomic dissection of resistance and other traits of interest.

The present study aimed to:

(1) use genomic models to estimate genetic parameters for bark stripping and selected primary and secondary chemical compounds in the bark, including those associated with bark stripping in radiata pine. A key focus was examination of the stability and potential improvement in heritability estimates of the bark chemical compounds and their additive genetic correlation with bark stripping and height; and (2) assess the predictive ability of the implemented models and evaluate the impact of data statistical distribution on genomic prediction.

7.2   Materials and methods

7.2.1   Plant material

The genetic field trial at Wilmot in Tasmania, Australia (-41.454271º N, 146.106801º E, 580 m ASL) described in Chapters 2, 4 and 6 was used for this study. Plant material was sourced from the New Zealand Radiata Pine Breeding Company (RPBC). This comprised 74 full-sib families that were planted in the field experiment in an incomplete randomised block design of 26 replicates with 78 incomplete blocks and single tree plots, with a total number of 1970 of trees. The families represented 55 unique parents and 54 grandparents. 20 replicates (1550 trees) were accessible to browsing for assessment of bark stripping. Six replicates (420 trees) were protected from bark stripping from which samples for the chemical analysis and for genotyping were collected (Supplementary Figure 7.1). The sample size is sufficient to reach the benchmark accuracy of conventional genomic selection (Grattapaglia and Resende 2011). Alternate trees in these replicates had been subject to artificial bark stripping (details in Chapters 2, 4 and 6).

7.2.2   Phenotypic data

a)   Bark stripping

From the 20 unprotected replicates, bark stripping damage by marsupials was recorded on an individual-plant basis at 2 years of age. The damage was scored on a categorical scale assigning zero (0) to non - stripped plants, 1 = <20% of the circumference stripped; 2 =20-50%; 3= 50-75% and 4= 75->100%, 5=100% of the circumference stripped. At the same time, height was assessed in all the 26

replicates. From the 6 protected replicates, needle and bark samples were collected for chemistry and genetic analysis.

b) Chemical analysis using near infrared reflectance spectroscopy

The chemical data used in this chapter was the same as used in Chapter 6. The chemical compounds i.e. terpenes, phenolics, fatty acids, sugars and unknown compounds were predicted by near infrared spectroscopy (NIRS) according to the methods documented in Chapter 5 and the associated wet chemistry methods are documented in Chapter 4. In brief, chemical extractions were done on 150 bark samples and NIRS prediction was done for the rest of the samples. Wet chemical extraction with dichloromethane (DCM) was carried out in 5ml of 0.75mg of fresh material and the acetone extraction was performed in 10 ml of 95% acetone using 50g of freeze-dried ground material. The sugars were extracted from 50 g of freeze dried, ground material in 10 ml of hot water (Jones *et al.* 2002). The DCM extracts were analysed by gas chromatography-mass spectrometry (GC-MS) while the acetone extracts and sugars were analysed by ultra-high-performance liquid chromatography-mass spectrometry (UHPLC-MS). The procedures for the GC-MS and UHPLC-MS are detailed in Chapter 3. For the NIRS analysis, samples were scanned when fresh and when freeze dried and ground according to the methods in Chapter 5. Near infrared reflectance spectroscopy models were developed to predict the amounts of all chemical compounds quantified in the bark as indicated in Chapter 6. The stronger model of either the cross-validated or the externally validated model was used to predict the chemistry of the unknown samples. The amount of each of the 65 compounds listed in Chapter 6 was predicted by NIRS and as a proof of concept only the 25 compounds in the bark predicted with models with $R^2 >$ 0.5 (Chapter 6) were selected for this study. Four of these compounds were shown to positively correlate with bark stripping in Chapter 6. The descriptive statistics and the statistical distribution of all traits considered in the study are shown in Table 7.1.

Table 7.1: Descriptive statistics of the traits used for genomic selection. Only bark compounds with NIRS models $r^2 > 0.5$ were selected for this study (Chapter 6), in addition to bark stripping and height. The monoterpenoids (M), sesquiterpenoids (SS) and GC-MS diterpenoids (DG) compound groups are expressed as milligrams of heptadecane equivalents per gram of dry weight of the sample. The LC-MS diterpenoids (DL), fatty acids (F) and the unknown compounds (U) are expressed as milligrams of nonadecanoic equivalents per gram of dry weight. The sugars (S) are expressed in absolute amounts (Min = minimum, Max = maximum, SD = standard deviation). Each chemical compound was given a unique identifier based on Supplementary Table 10 for ease of location in the tables

Id	Compound	Compound group	Min	Mean	Max	SD	skewness	kurtosis
	Bark stripping		0.00	25.20	100.00	33.20	1.35	0.62
	Height (cm)		77.00	163.70	257.00	30.40	-0.35	0.72
1	α-pinene	M	0.02	3.82	0.78	0.42	2.08	9.56
4	β-pinene	M	-0.38	8.34	1.84	0.99	1.40	6.11
5	camphene	M	0.00	0.03	0.01	0.00	0.88	3.15
8	citronellal	M	-0.06	0.46	0.04	0.05	2.09	12.83
20	trans-farnesol	SS	-0.02	0.10	0.02	0.02	1.00	2.88
22	agathadiol	DG	-0.54	3.89	0.55	0.51	2.00	7.99
23	agatholal	DG	-0.10	1.56	0.34	0.21	1.57	4.81
24	copalol	DG	0.00	0.18	0.03	0.02	2.33	12.09
25	levopimaral	DG	0.00	0.05	0.01	0.01	1.25	3.01
29	dehydroabietic acid	DL	8.53	39.35	24.71	5.18	0.72	-0.12
33	unknown m/z 109 A	DL	-0.01	0.07	0.02	0.01	0.84	1.94
45	unknown m/z 316	DL	2.48	44.63	13.75	5.58	1.20	3.62
46	unknown C20H30O3	DL	7.56	88.00	25.93	9.27	1.65	7.25
47	unknown C20H32O3 A	DL	-0.29	92.88	20.71	10.29	1.86	7.81
50	unknown C20H30O4	DL	22.24	175.53	60.58	18.16	1.53	5.94
51	unknown C20H30O5	DL	3.09	39.25	11.34	4.59	1.58	5.40
52	unknown C20H30O6 A	DL	-0.06	1.16	0.22	0.13	1.59	7.50
54	unknown C20H30O6 C	DL	1.79	15.23	6.29	1.89	0.90	1.77
76	fructose	S	-0.10	2.12	1.33	0.36	-0.30	0.01
77	glucose	S	0.31	3.06	1.53	0.43	0.09	0.30
78	inositol	S	0.17	2.30	1.09	0.36	0.29	0.13
83	linoleic acid	F	7.12	27.00	16.91	3.75	-0.08	-0.21
84	linolenic acid	F	1.11	12.31	7.69	1.46	-0.17	0.67
90	unknown m/z 274	U	0.36	3.72	1.11	0.40	1.68	6.62
94	unknown m/z 406 A	U	0.08	1.02	0.53	0.17	-0.02	-0.27

7.2.3 Genotyping

From the 6 protected replicates, needle samples were collected from all individuals (n=420) and stored at -80^0C before DNA extraction. Total genomic DNA was extracted using a commercial NucleoSpin Plant II kit (Machery-Nagel, Duren, Germany) with modifications (Telfer *et al.* 2013). DNA purity and concentration were evaluated using a NanoDrop 2000 spectrophotometer and quantified using the Agilent 5200 fragment analyser. The samples were genotyped using the 50K axiom SNP chip for radiata pine developed on the Illumina platform (Thermofischer) (Telfer, 2019). Currently this is the densest SNP array for radiata pine, capable of assaying in excess of 80,000 SNPs. 396 individuals were included in the final genotype data with a total of 27,000 SNPs. This genotype data was filtered to include only SNPs with a mean allele frequency (MAF) >0.01 and maximum missing data of 0.4% using the rrBLUP package. This filtering resulted in the retention of 15,624 SNPs for analysis. The genotyping reproducibility rate was high (99.9%), as estimated from 10 samples that were replicated during DNA extraction.

7.2.4 Statistical methods

Three selected statistical methods were compared - genomic best linear unbiased predictor (GBLUP), single-step genomic BLUP (ssGBLUP) and generalised ridge regression (GRR). GBLUP and ssGBLUP do not estimate individual marker effects and the two methods rely on different sample sizes. GBLUP used only the samples that were genotyped and phenotyped (n=396) and these were the plants in the 6 protected replicates for which chemistry data was available (see Chapter 6). The ssGBLUP included all individuals in the Wilmot trial with documented pedigree that had been phenotyped for height (26 replicates, n=1970). Of these, 20 replicates were exposed to bark stripping and 6 replicates were protected (see above). The genomic models were compared to the pedigree based (ABLUP) model involving the same individuals, which is the standard method used for breeding value prediction using the expected relatedness among individuals based on pedigree information.

The ssGBLUP and GBLUP models are similar to the ABLUP models detailed in Chapter 6, except that the average numerator relationship matrix A in the ABLUP is substituted with the realized genomic relationship matrix (G) in GBLUP and with the H-matrix that combines G and A matrices in ssGBLUP (Christensen and Lund 2010).

The G-matrix was computed using the "A.mat" function in the R package rrBLUP (Endelman 2011) from the marker data following (VanRaden 2008):

$$G = \frac{ZZ\prime}{2\sum_i p_i(1-p_i)}$$

(1)

where Z where $Z = M - P$, M is the matrix of genotypes coded 0, 1, and 2 as reference allele homozygote, heterozygote, and alternative allele homozygote, respectively, and P is the vector of doubled frequencies for alternative alleles, p is the frequency of the alternative allele at i^{th} loci.

The single-step GBLUP (ssGBLUP) combines the pedigree relationship matrix A and the genomic relationship matrix G, in one matrix, H and hence simultaneously uses information from genotyped and non-genotyped individuals. The H-matrix is defined by

$$H= \begin{bmatrix} A_{11} + A_{12}A_{22}^{-1}(G_s - A_{22})A_{22}^{-1}A_{21} & A_{12}A_{22}^{-1}(G_s - A_{22}) \\ (G_s - A_{22})A_{22}^{-1}A_{21} & (G_s - A_{22}) \end{bmatrix}$$

(2)

where, A_{11}, represents the relationship matrix for the non-genotyped individuals (20 replicates, n = 1550), A_{12} and A_{21} are relationship matrices between genotyped and non-genotyped individuals (26 replicates, n = 1970) while A_{22} is the pedigree-based relationship matrix for genotyped individuals (6 replicates, n = 420) and G_s is the scaled marker-based matrix, G-matrix for only the genotyped individuals (see scaling below). Moreover, A_{22}^{-1} denotes the inverse of A_{22}. And is denoted as;

$$H^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & G_S^{-1} A_{22}^{-1} \end{bmatrix} + A_{11}^{-1}$$

(3)

Forming the H-matrix above involves 2-major steps; i) the G-matrix is rescaled (G_a) such that the average of its diagonal elements (avg.diag) is equal to the average of the diagonal elements of the A matrix, such that the G and A matrices are compatible. This is also done for the average of the non-diagonal elements (avg.offdiag). Following (Gao *et al.* 2012), this was done by adjustment factors, α and β to all elements of G.

$$G_a = G\beta + \alpha,$$

(4)

where α and β are adjustment factors derived from the following equations:

$$\begin{cases} \text{Avg.diag (G) } \beta + \alpha = \text{Avg.diag } (A_{22}) \\ \text{Avg.offdiag (G) } \beta + \alpha = \text{Avg.offdiag } (A_{22}) \end{cases}$$

(5)

ii) The G matrix is usually not positive semi-definite, which is one of the mixed linear model assumptions, and weighting of the genomic and pedigree-based relationship matrices is required as follows:

$$G_s = G_a(1-w) + A_{22}w$$

(6)

where G_s is a rescaled genomic relationship matrix based on the SNP data, G_a is the adjusted genomic relationship matrix (equation 4) and w is the weighting factor that represents the fraction of total additive variance that is not captured by markers and A_{22} is the additive relationship matrix from the full pedigree. The weight (w) can take any value between 0 and 1, where the model fitted with w=0 is equivalent to the GBLUP model while the model where w=1 is equivalent to ABLUP. For the present study, a low w of 0.05 was selected to give high weighting to the genomic data as 5% pedigree errors were detected during the attempted pedigree reconstruction (Martini *et al.* 2018).

Linear models for estimation of variance components

Variance components based on the univariate pedigree-based relationship matrix (A) have been previously documented (Chapter 6). This chapter presents the results from univariate and trivariate ABLUP for comparison with the genomic models. The variance components for the ABLUP, ssGBLUP or GBLUP were obtained in ASReml v4.1 (Gilmour *et al.* 2015) using a general linear additive genetic model as,

$$y = X\beta + Zu + e, \tag{7}$$

where, *y* is the response variable (height, spatially adjusted bark stripping and a chemical variable); *β*, a vector of fixed effects; u is the vector of random effects which included replicates, incomplete blocks within replicates, tree (additive genetic effect – estimated using either the A, G or H relationship matrices) terms and e is a vector of random residuals. The random family (specific combining ability) and the tree*treatment terms were excluded from the analyses because they were generally non-significant in previous analyses (Chapter 6). X and Z correspond to design matrices relating the observations in y to the fixed and random effects in β and u, respectively. The vector of random additive effect (u) in equation is assumed to follow a normal distribution $u \sim N(0,A(G/H)\sigma^2)$, where σ^2 is the additive genetic variance based on the A, G or H matrices. For single models, the fixed term *β* contained the overall mean and the treatment term that was fitted for the chemical but for multi-trait models another term, "protected" was fitted for height as a fixed term to differentiate the measurements from the 20 replicates that were unprotected – where bark stripping was scored, from those from the 6 protected replicates from which chemistry was estimated. To test whether the additive genetic variation was greater than zero, full models were compared with respective reduced models using a one-tailed likelihood ratio test (Gilmour *et al.* 2015). This was done for only the GBLUP model and for the trivariate models, tests were done for compounds [76,77,83,84] that exhibited significant correlation with bark stripping and height in Chapter 6.

Narrow-sense heritability estimates were derived from univariate ABLUP and GBLUP as well as trivariate ABLUP and trivariate ssGBLUP models (see below for trivariate models). Individual narrow-sense heritability ($\hat{h}^2$) was estimated as the additive genetic variance divided by the sum of the additive genetic variance $\hat{\sigma}_a^2$ and the error variance $\hat{\sigma}_e^2$ as below:

$$\hat{h}^2 = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_e^2} \tag{8}$$

Estimates of the associated standard error for the traits were obtained directly using Taylor series expansion ("delta method") (Gilmour *et al.* 2015). The univariate ABLUP heritability values presented were those from Chapter 6.

Comparisons of the heritability estimated were made between pedigree and genomic based models. The univariate GBLUP estimates were compared to univariate ABLUP, and similarly the trivariate ssGBLUP was compared with the trivariate ABLUP. A two-tailed paired t-test was used to test the difference in the average heritability estimates in R v 3.6.0.

Generalised ridge regression

Generalised ridge regression (GRR) estimates marker effects using linear and penalized parameters based on only the genotyped individuals (n=396). GRR alters the notations of parameters *a* and *Z* in Eq. (1) by allowing variable shrinkage for different markers through the introduction of a diagonal matrix following a two-step process (Shen *et al.* 2013).

In the first step, the predicted breeding values are obtained following the mixed model by the summing of all the marker effects of an individual tree (similar to model 7), where, X is a vector of 1. The terms e (vector for random residuals) and Z (the design matrix for SNP effects) are respectively assumed to follow a normal distribution, i.e. $e \sim N(0, I\sigma_e^2)$ and $u \sim N(0, I\sigma_u^2)$, where I is an identity matrix. The solution for the marker effects is given by the following equation:

$$\hat{u} = (z'z + \lambda I)^{-1} z'y \tag{9}$$

where, $\lambda = {\sigma_e^2}/{\sigma_u^2}$, the ridge penalty parameter is the ratio between the residual and marker variances (Shen *et al.* 2013; Veerman *et al.* 2019). The method assumes equal distribution of marker effects, and therefore all effects are equally shrunk towards zero.

In the second step, a marker-specific shrinkage parameter is imposed on the BLUP for $\hat{u}$ where λI becomes *diag* (λ). The equation now becomes

$$\hat{u} = (z'z + \text{diag}(\lambda))^{-1}z'y$$

where λ is a vector of shrinkage parameters and $\hat{u}$ is the BLUP marker effect (from step 1). Generalized ridge regression was implemented in the "bigRR" package in R (Rönnegård and Shen 2016).

Genetic correlations

To estimate the genetic correlations among traits based on the ssGBLUP, trivariate models were fitted including height, spatially adjusted bark stripping and one chemical compound as response variables. Similar to the trivariate ABLUP models in Chapter 6, the fixed treatment effect was fitted to the chemical compound and the fixed protected effect was fitted to height. The random terms included the additive genetic variation, replicates and incomplete blocks within replicates. The multivariate analyses assumed heterogeneous independent residuals. The genetic correlation (r_g) between two traits measured was estimated as:

$$r_g = \frac{cov_a(x,y)}{\sqrt{\sigma^2_{ax} \cdot \sigma^2_{ay}}} \tag{10}$$

where $cov_a(x, y)$ is the additive genetic covariance between traits x and y, σ^2_{ax} is the additive genetic variance components for trait x, and σ^2_{ay} is the additive genetic variance components for trait y. Standard errors were estimated in ASReml (Gilmour *et al.* 2015). The accuracy of the genetic correlation estimates were compared with those obtained with the trivariate ABLUP (Chapter 6). This was done by using a two-tailed paired t-test to test the difference in the standard errors of the correlation in R v 3.6.0.

Cross-validation scheme for estimating predictive ability

The predictive ability assesses the potential of the models to estimate the breeding value of individuals with yet-to-be observed phenotypes (Momen *et al.* 2018). To test the predictive ability of the 3 genomic and ABLUP methods, a 10-fold cross-validation scheme was implemented in ASReml (Gianola and Schön 2016; Utz *et al.* 2000). Within each model, individuals were randomly subdivided into 10 subsets (i.e., folds) irrespective of family, and a leave one out procedure was repeated 10 times until all individuals had their breeding values predicted. The average estimated breeding value from all iterations was computed for each individual. Predictive ability of phenotypes was then defined as the Pearson correlation between the genomic estimated breeding values predicted in cross-validation and the observed phenotypes (Momen *et al.* 2018).

To detect if data distribution was correlated to predictive ability, the coefficient of skewness (a measure of symmetry) and kurtosis (a measurement about the extremities [i.e. tails] of the distribution of data, that provides an indication of the presence of outliers) were estimated for each compound using functions SKEW and KURT respectively in excel. The expected skewness of a normal distribution is zero (0) while a dataset with no tails has a kurtosis of 3.0. The relationship between predictive ability for each and skewness and kurtosis was assessed using a scatter plot.

7.3   Results

7.3.1   ABLUP - univariate

The additive genetic estimates for the 25 chemical compounds from the univariate pedigree-based model have been previously presented (Chapter 6) and are presented here for comparative purposes (Table 7.2). The narrow-sense heritability values of the 25 chemical compounds ranged between 0.03 - 0.37 and averaged to 0.23 ± 0.09 (Table 7.2). Apart from four compounds - agathadiol [22], dehydroabietic acid [29], inositol [78] and unknown m/z 406 A [94], the chemical compounds showed significant additive genetic variation (Table 7.2). Significant additive genetic variation but low heritability was exhibited for height ($h^2 = 0.04 \pm 0.02$) and the spatially adjusted bark stripping ($h^2 = 0.07 \pm 0.03$) based on univariate models (Table 7.2).

7.3.2   GBLUP - univariate

Based on the single trait GBLUP that used only the relationship matrix derived from genomic data (G-matrix), heritability estimates ranged from $0.04 – 0.60$ (average 0.37 ± 0.10), suggesting that the markers were sufficient in capturing the additive genetic variation despite the small sample size. The average of heritability estimates for the chemical traits (0.37 ± 0.10) was 1.6- fold higher than the average of the univariate ABLUP heritabilities (0.23 ± 0.09; $t_{24} = -9.16$, $p < 0.001$) (Table 7.2). The strongest improvement was exhibited with compounds that had the lowest ABLUP heritability. For example, a 3.5 - fold increase in heritability was detected for height. Also, the heritability of the unknown m/z 406 A[94] ($h^2 = 0.04 \pm 0.05$) increased 2.5-fold with the GBLUP ($h^2 = 0.10 \pm 0.06$) while that of citronellal [8], trans-farnesol[20] and unknown C20H30O3[46] increased 2.2-fold. Apart from the heritability of linoleic acid [83] that slightly reduced, the compounds experienced between 1.1-2.09 - fold increment with the GBLUP. Significant additive genetic variation was also detected for agathadiol [22], dehydroabietic acid [29], inositol [78] and unknown m/z 406 A [94] that were not significant with the ABLUP models. Bark stripping heritability estimates were not derived using GBLUP since bark stripping had missing values for the trees were genotyped.

7.3.3   ssGBLUP - trivariate

Based on the trivariate ssGBLUP which combines pedigree (A) and genomic (G) matrices, low heritability values were estimated for bark stripping (h^2 = 0.12 ± 0.04) and height (h^2 = 0.07 ± 0.03). These estimates were constant irrespective of chemical compound fitted in the model. Compared to the trivariate ABLUP, the heritability for both bark stripping and height improved 1.4 - fold (Table 7.2). The trivariate ssGBLUP heritability estimates for individual bark compounds ranged between 0.13 – 0.77. Relative to the heritability of all the compounds based on the trivariate ABLUP (average = 0.27 ± 0.10), the heritability improved 1.7 – fold with ssGBLUP model (average =0.46 ± 0.10; t_{24} = -8.79, p < 0.001). Similar to GBLUP, the compounds that had high heritability with ABLUP showed the less relative improvement with ssGBLUP (Figure 7.1, Table 7.2). The four compounds that did not show any significant additive genetic variation with the univariate ABLUP displayed significant additive genetic variation with this model (results not shown).

Table 7.2: Narrow-sense heritability (h2) and standard error (se) estimates of bark stripping, height and selected chemical compounds quantified in *P. radiata* bark based on univariate models for ABLUP and GBLUP and trivariate ABLUP and ssGBLUP. The trivariate models included height, spatially adjusted bark stripping and one chemical compound. The predictive ability (PA) defined as the correlation of breeding values from 10-fold cross validation is also indicated. The significance that the additive genetic variation from the univariate ABLUP and GBLUP was greater than zero was tested using the one-tailed likelihood ratio test (Chapter 6), where * = p < 0.05, ** = p < 0.01 and *** p < 0.001. Significance tests were not done for the trivariate models. M = monoterpenoids, S=sesquiterpenoids, DG = GC-MS diterpenoids, GL=LC-MS diterpenoids, S=Sugars, F=Fatty acids, U=Unknown compounds. Each chemical compound is given a unique identifier (Id) based on Supplementary Table 10 (after Chapter 9), for ease of location in the tables

Id	Compound	Group	Narrow-sense heritability (h2) (se)					Predictive ability				
			ABLUP-univariate		GBLUP-univariate	ABLUP-trivariate	ssGBLUP-trivariate	ABLUP-univariate	GBLUP-univariate	GRR-univariate	ABLUP-trivariate	ssGBLUP-trivariate
	Bark stripping		0.07 (0.03)	***		0.09 (0.03)	0.12 (0.04)	0.46			0.11	0.11
	Height		0.04 (0.02)	**	0.14 (0.08)	0.05 (0.02)	0.07 (0.03)	0.37			0.36	0.10
1	α-pinene	M	0.25 (0.10)	***	0.32 (0.10)***	0.25 (0.10)	0.32 (0.11)	0.76	0.22	0.24	0.74	0.82
4	β-pinene	M	0.27 (0.10)	***	0.45 (0.11)***	0.34 (0.12)	0.60 (0.11)	0.74	0.30	0.32	0.74	0.75
5	camphene	M	0.26 (0.10)	***	0.45 (0.11)***	0.29 (0.11)	0.55 (0.11)	0.75	0.33	0.33	0.75	0.78
8	citronellal	M	0.18 (0.09)	***	0.40 (0.12)***	0.19 (0.09)	0.48 (0.11)	0.68	0.35	0.32	0.71	0.67
20	trans-farnesol	SS	0.14 (0.08)	***	0.31 (0.11)***	0.14 (0.08)	0.43 (0.11)	0.66	0.27	0.26	0.60	0.14
22	agathadiol	DG	0.21 (0.10)		0.43 (0.11)***	0.22 (0.10)	0.46 (0.11)	0.74	0.25	0.28	0.74	0.74
23	agatholal	DG	0.22 (0.09)	***	0.33 (0.10)***	0.24 (0.10)	0.38 (0.11)	0.73	0.28	0.19	0.73	0.77
24	copalol	DG	0.29 (0.10)	***	0.51 (0.11)***	0.28 (0.10)	0.51 (0.11)	0.79	0.37	0.37	0.79	0.82
25	levopimaral	DG	0.30 (0.10)	***	0.48 (0.11)***	0.32 (0.11)	0.53 (0.11)	0.79	0.36	0.37	0.79	0.70
29	dehydroabietic acid	DL	0.03 (0.02)		0.04 (0.03)**	0.10 (0.07)	0.18 (0.09)	0.33	-0.13	-0.06	0.30	0.31
33	unknown m/z 109 A	DL	0.33 (0.11)	***	0.60 (0.11)***	0.36 (0.12)	0.77 (0.10)	0.81	0.47	0.43	0.81	0.88

45	unknown m/z 316	DL	0.21 (0.09)	*	0.32 (0.11)***	0.25 (0.10)	0.41 (0.12)	0.71	0.20	0.18	0.70	0.71
46	unknown C20H30O3	DL	0.18 (0.08)	*	0.40 (0.12)***	0.21 (0.10)	0.50 (0.13)	0.69	0.26	0.27	0.69	0.71
47	unknown C20H32O3 A	DL	0.21 (0.09)	***	0.44 (0.11)***	0.21 (0.09)	0.48 (0.11)	0.71	0.32	0.32	0.78	0.74
50	unknown C20H30O4	DL	0.24 (0.09)	***	0.41 (0.11)***	0.28 (0.11)	0.50 (0.12)	0.73	0.27	0.28	0.73	0.54
51	unknown C20H30O5	DL	0.25 (0.10)	***	0.43 (0.11)***	0.29 (0.11)	0.52 (0.12)	0.75	0.26	0.30	0.76	0.64
52	unknown C20H30O6 A	DL	0.31 (0.11)	***	0.45 (0.11)***	0.34 (0.11)	0.57 (0.11)	0.79	0.29	0.30	0.79	0.81
54	unknown C20H30O6 C	DL	0.30 (0.11)	***	0.43 (0.10)***	0.38 (0.12)	0.54 (0.11)	0.77	0.24	0.29	0.81	0.79
76	fructose	S	0.21 (0.09)	**	0.31 (0.10)***	0.30 (0.10)	0.47 (0.10)	0.58	0.29	0.28	0.57	0.46
77	glucose	S	0.20 (0.09)	**	0.37 (0.10)***	0.27 (0.10)	0.49 (0.10)	0.63	0.18	0.29	0.59	0.64
78	inositol	S	0.14 (0.08)		0.29 (0.10)***	0.16 (0.09)	0.32 (0.11)	0.60	0.08	0.07	0.60	0.16
83	linoleic acid	F	0.30 (0.11)	***	0.27 (0.09)***	0.55 (0.13)	0.46 (0.10)	0.70	0.18	0.19	0.69	0.68
84	linolenic acid	F	0.37 (0.12)	***	0.40 (0.10)***	0.45 (0.12)	0.47 (0.10)	0.81	0.29	0.39	0.80	0.80
90	unknown m/z 274	U	0.31 (0.11)	**	0.37 (0.11)***	0.32 (0.11)	0.39 (0.11)	0.80	0.27	0.28	0.81	0.81
94	unknown m/z 406 A	U	0.04 (0.04)		0.10 (0.06)*	0.04 (0.05)	0.13 (0.09)	0.46	0.00	-0.03	0.28	0.33

1

7.3.4   Univariate versus trivariate models

The application of trivariate models improved the genetic estimates of most traits (Table 7.2). With the ABLUP, a 1.3 – fold increase in heritability of both bark stripping and height was detected with trivariate ABLUP compared with the univariate estimates, with no change in the standard error of the estimates (Table 7.2). The heritability estimates of the chemical compounds also improved in trivariate (average $h^2 = 0.27 \pm 0.10$) compared to the univariate ABLUP (average $h^2 = 0.23 \pm 0.09$; $t_{24} = 3.46$, $p < 0.01$). However, the heritability of α-pinene [1], trans-farnesol [20], one unknown diterpenoid [49] and unknown m/z 406 A [94] did not change, while that of copalol [24] slightly reduced in the trivariate ABLUP models indicating that the improvement of genetic estimates from correlated traits in multi-trait models is trait specific. With the ssGBLUP, a 50% reduction in heritability was detected for height, compared to the univariate GBLUP, although the ssGBLUP was associated with lower standard error of measurement suggesting that the ssGBLUP estimate is more accurate. With the ssGBLUP an improvement in the heritability estimates was detected for most of the chemical compounds (average $h^2 = 0.46 \pm 0.10$) relative to GBLUP (average $h^2 = 0.37 \pm 0.10$; $t_{24} = 7.14$, $p < 0.001$), except copalol [24], indicating the relative advantage of additional phenotypic information from correlated traits.

Of the chemical compounds, the largest increment in heritability for both trivariate pedigree and genomic models relative to the univariate ABLUP was observed for compounds that had the lowest heritability estimate in the univariate models (Figure 7.1). The highest improvement was for example detected for dehydroabietic acid [29], where a 3.3 - fold and a 4.5 - fold improvement was detected with trivariate ABLUP and ssGBLUP respectively relative to the univariate models

7.3.5 Genetic correlations

The genetic correlations of chemical compounds with bark stripping or height varied from positive to negative for both trivariate ABLUP and ssGBLUP. Although most correlations retained the signs, shifts from the low negative ABLUP to low positive ssGBLUP genetic correlations were detected (Table 7.3). The magnitude of the ABLUP positive correlations marginally increased in magnitude compared to the genetic correlation from the ssGBLUP (Table 7.3). The ABLUP negative correlations either changed to positive or reduced in magnitude with the ssGBLUP. However, for both the ABLUP positive and negative, correlations, the standard error (se) of most estimates significantly reduced with the ssGBLUP compared to the trivariate ABLUP for both bark stripping (ABLUP average se = 0.29, ssGBLUP average se = 0.26, $t_{24} = 3.52$, $p < 0.01$) and height (ABLUP average se = 0.32, ssGBLUP average se = 0.28, $t_{24} = 5.30$, $p < 0.001$), suggesting more accurate measurements with this model. The correlation

estimates that had the highest ABLUP standard errors more strongly improved in accuracy than those that had low errors of estimation (Figure 7.1, Table 7.3).

The genetic correlations between the sugars – glucose [77] and fructose [76] as well as the fatty acids - linoleic acid [83] and linolenic acid [84] with bark stripping that were significant with the trivariate ABLUP model were still significant with the trivariate ssGBLUP, although the magnitude of the genetic correlations slightly reduced (Table 7.3). The genetic correlations of compounds with height that were significant with the ABLUP models were non- significant with the ssGBLUP except for the genetic correlation between height and linoleic acid that retained its significance (r_g = 0.62 ± 0.24, p < 0.05)

Table 7.3: Additive genetic correlation of different bark chemical compounds with bark stripping and height estimated using the pedigree-based method (ABLUP) and single step GBLUP (ssGBLUP) trivariate models. The multivariate models always included the spatially adjusted bark-stripping scores and height and one of the listed bark chemical compounds as response variables. The significance that the additive genetic correlation (r_g) is different from zero was tested using a two-tailed likelihood ratio test (Chapter 6) and are presented here for comparative purposes. * = $p <0.05$, ** = $p <0.01$ and *** $p <0.001$, se = standard error of the genetic correlation. The GBLUP and GRR models are not included here since they were univariate. Significance tests were done for all ABLUP models and for correlations that were significant, likelihood ratio tests were further done on the ssGBLUP estimates. Each chemical compound is given a unique identifier (Id) based on Supplementary Table 10 (end of Chapter 9), for ease of location in the tables

	Compound	Genetic correlation with bark stripping (r_g) (se)		Genetic correlation with height (r_g) (se)	
		trivariate ABLUP	trivariate ssGBLUP	trivariate ABLUP	trivariate ssGBLUP
	Height	0.42 (0.29)	0.43 (0.27)		
1	α-pinene	-0.20 (0.29)	-0.01 (0.28)	0.02 (0.33)	0.15 (0.30)
4	β-pinene	-0.01 (0.28)	0.14 (0.25)	-0.04 (0.32)	0.15 (0.26)
5	camphene	-0.12 (0.28)	0.01 (0.26)	-0.00 (0.33)	0.07 (0.27)
8	citronellal	-0.14 (0.32)	-0.17 (0.27)	-0.32 (0.36)	-0.16 (0.29)
20	trans-farnesol	0.08 (0.35)	0.25 (0.26)	-0.37 (0.39)	0.01 (0.29)
22	agathadiol	-0.03 (0.31)	0.14 (0.27)	0.00 (0.36)	0.18 (0.29)
23	agatholal	-0.12 (0.29)	0.07 (0.27)	0.18 (0.33)	0.24 (0.29)
24	copalol	-0.09 (0.28)	0.05 (0.26)	-0.20 (0.32)	-0.02 (0.28)
25	levopimaral	0.08 (0.27)	0.16 (0.25)	0.04 (0.31)	0.11 (0.27)
29	dehydroabietic acid	-0.27 (0.37)	0.10 (0.33)	0.39 (0.39)	0.43 (0.32)

33	unknown m/z 109 A	0.07 (0.27)		0.14 (0.24)		-0.11 (0.31)		0.16 (0.25)	
45	unknown m/z 316	-0.29 (0.29)		-0.15 (0.28)		-0.05 (0.34)		0.16 (0.29)	
46	unknown C20H30O3	0.15 (0.30)		0.33 (0.25)		-0.03 (0.35)		0.30 (0.27)	
47	unknown C20H32O3 A	-0.04 (0.30)		0.13 (0.26)		-0.13 (0.34)		0.05 (0.28)	
50	unknown C20H30O4	0.05 (0.28)		0.18 (0.26)		0.07 (0.32)		0.23 (0.27)	
51	unknown C20H30O5	0.07 (0.28)		0.17 (0.26)		0.13 (0.32)		0.29 (0.27)	
52	unknown C20H30O6 A	0.18 (0.27)		0.25 (0.24)		0.01 (0.31)		0.20 (0.26)	
54	unknown C20H30O6 C	0.32 (0.25)		0.37 (0.23)		0.16 (0.30)		0.29 (0.25)	
76	fructose	0.55 (0.23)	*	0.51 (0.22)	*	0.06 (0.31)		0.08 (0.27)	
77	glucose	0.80 (0.20)	**	0.71 (0.19)	**	0.62 (0.24)	**	0.52 (0.23)	
78	inositol	-0.14 (0.33)		-0.04 (0.29)		-0.01 (0.38)		-0.01 (0.32)	
83	linoleic acid	0.68 (0.16)	**	0.65 (0.19)	**	0.69 (0.22)	*	0.61 (0.24)	*
84	linolenic acid	0.65 (0.19)	**	0.62 (0.20)	*	0.50 (0.26)	**	0.47 (0.25)	
90	unknown m/z 274	0.03 (0.28)		0.16 (0.26)		0.05 (0.32)		0.23 (0.28)	
94	unknown m/z 406 A	0.77 (0.56)		0.25 (0.36)		0.77 (0.66)		0.46 (0.39)	
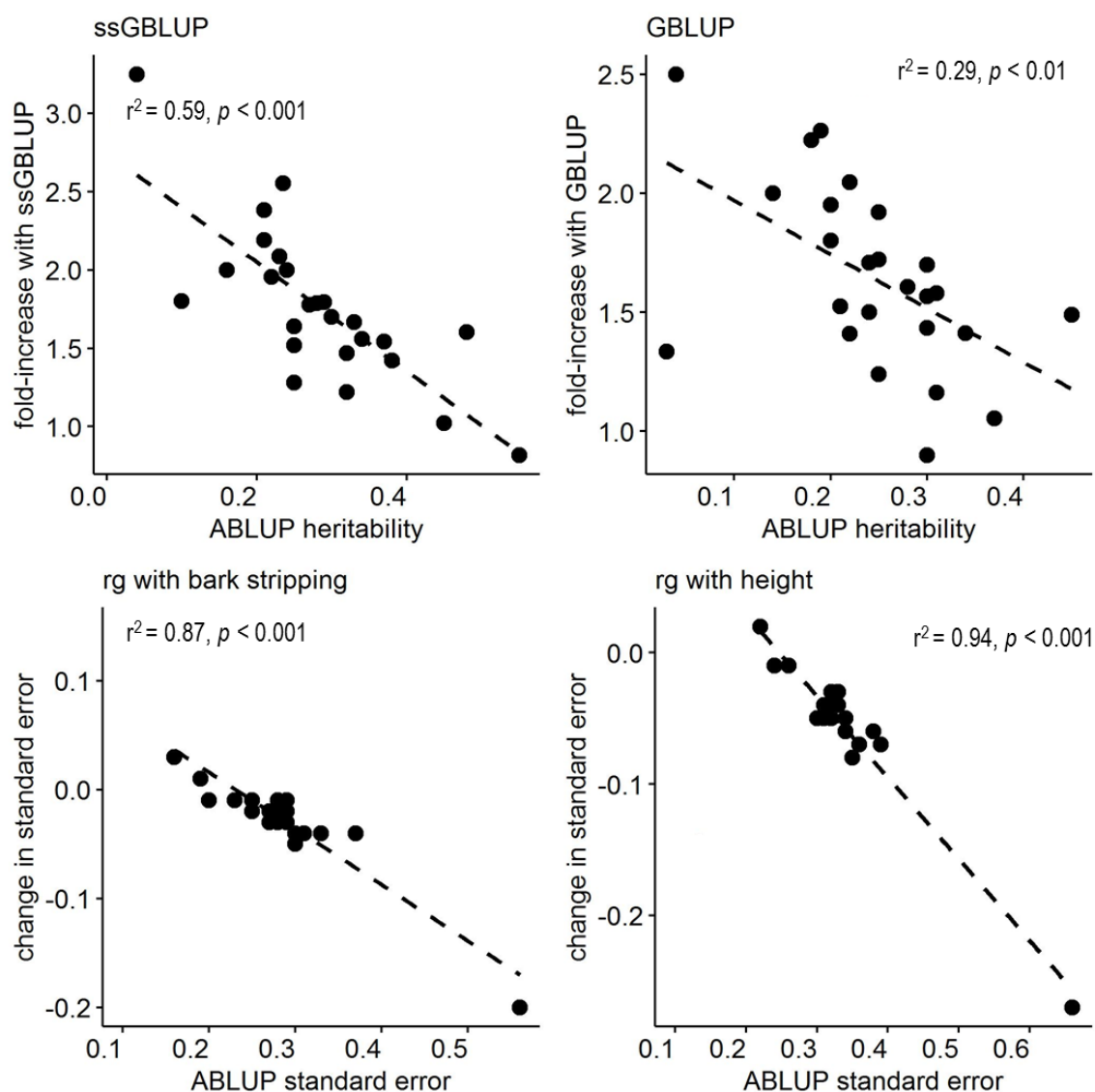
Figure 7.1: Above: The relationship between ABLUP heritability and improvement in heritability with the ssGBLUP and GBLUP models for the 25 chemical compounds in Table 7.2. Univariate GBLUP was compared to univariate ABLUP and the trivariate ssGBLUP was compared to trivariate ABLUP Below: Change in standard error (se) of the genetic correlations (r_g) of the chemical compounds with bark stripping or height. The change was calculated as the difference between the standard error associated with ABLUP estimates and that of ssGBLUP (se ABLUP r_g - se ssGBLUP r_g) estimate for both bark stripping and height.

7.3.5 Predictive ability

The predictive ability (PA) for the height and the chemical phenotypes varied between -0.13 to 0.88 depending on the trait and statistical model (Table 7.2). The PA for bark stripping was not estimated for the GBLUP and GRR since none of the bark stripped trees were genotyped. Overall, genomic models did not improve the PA for bark stripping, height and the chemical compounds. Comparing univariate models, the univariate ABLUP (average PA = 0.70) outperformed the GBLUP (average PA =0.25) and the GRR (average PA = 0.26). The PA for the univariate GBLUP and GRR were comparable (t_{24} = -1.2, p = 0.23). However, a few compounds, for example the sugar - glucose [77], a fatty acid - linolenic acid [84], and a diterpenoid[29] were better predicted by GRR compared to GBLUP suggesting that these compounds may be controlled by genes with major effects. Similarly, no relative advantage was detected for the trivariate ssGBLUP (average PA =0.65) over the trivariate ABLUP (average PA =0.69) models (t_{24} = 0.12, p = 0.12). The PA for height reduced with trivariate ssGBLUP (PA = 0.10) compared to trivariate ABLUP (PA =0.36).

Within genomic models, PA for trivariate ssGBLUP was higher than that of the univariate GBLUP and GRR, indicating the advantage of the additional non-phenotyped trees. However, the PA for trivariate and univariate ABLUP did not differ (t_{24} = -0.97, p = 0.34). The linear relationship between heritability and predictive ability (PA) for the 25 chemical compounds, was high indicating that there was a tendency for compounds with higher heritability estimates to reach higher predictive ability for all models (Figure 7.2).

In trivariate models, higher PA is expected in the presence of high genetic correlation between the traits. However, there was no linear relationship between the predictability of the chemical traits and the genetic correlation of the chemical traits with bark stripping or height (Supplementary Figure 7.2). Instead, the absolute standard error of the correlations showed a negative linear relationship with predictive ability, suggesting that the precision of the estimate impacts PA rather than the magnitude. The genetic correlations that were associated with low ABLUP error of estimating the genetic correlation were better predicted that those that had higher genetic correlation standard errors (Figure 7.3).
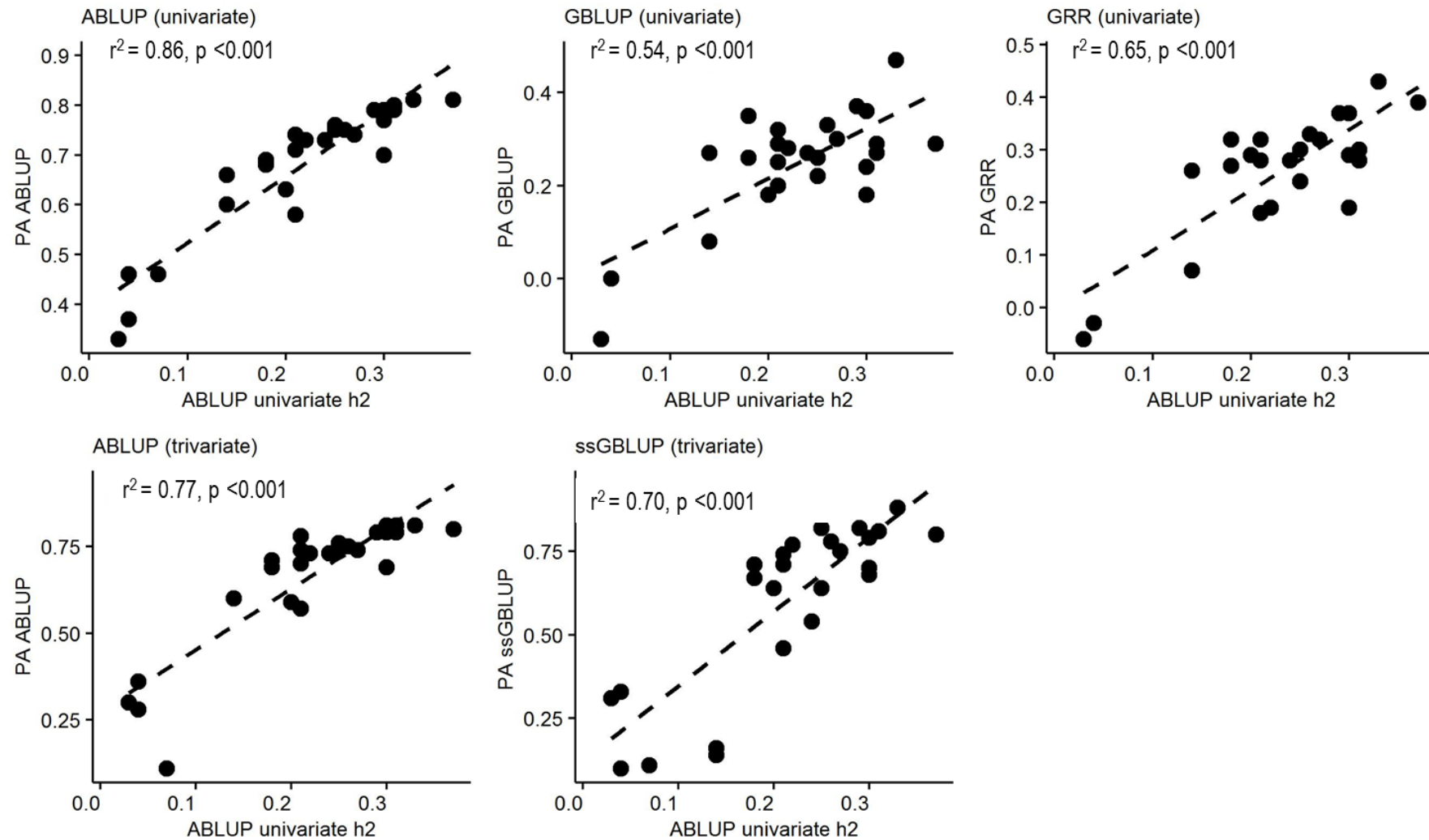
Figure 7.2: The coefficient of determination (r^2) showing the relationship between ABLUP narrow-sense heritability (h^2) and predictive ability of univariate ABLUP, GBLUP and GRR as well as trivariate ABLUP and ssGBLUP, models. The heritability estimates were derived from univariate models.

1 To detect the effect of the distribution of phenotypic data on predictive ability (PA), there was a

2 tendency for positively skewed chemical compounds to have higher heritability estimates for all models

3 (Figure 7.5) suggesting that a higher amount of compound may be better predicted. Similarly, there was

4 a tendency for compounds with expected kurtosis coefficient (~3.0) to have higher heritability estimates

5 for all the models (Figure 7.5) indicating that the presence of outliers in the data affects the prediction of

6 breeding values more than the symmetry for these models.
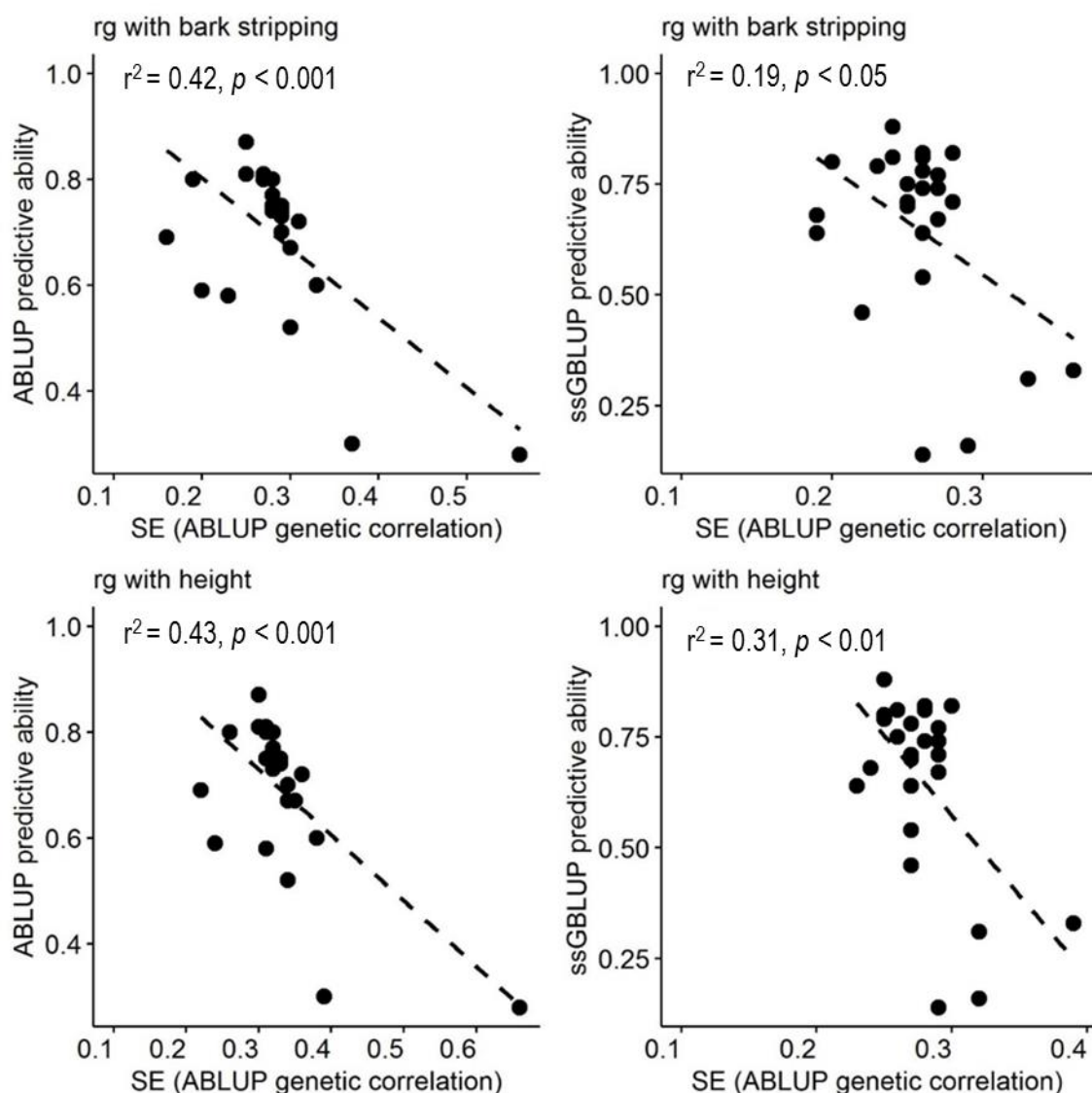
Figure 7.3: The coefficient of determination (r^2) between predictive ability of the ABLUP and ssGBLUP models and the standard error (SE) of the trivariate ABLUP genetic correlation of chemical compounds with bark stripping (above) and height (below). The trivariate models included spatially adjusted bark stripping, height and a chemical compound
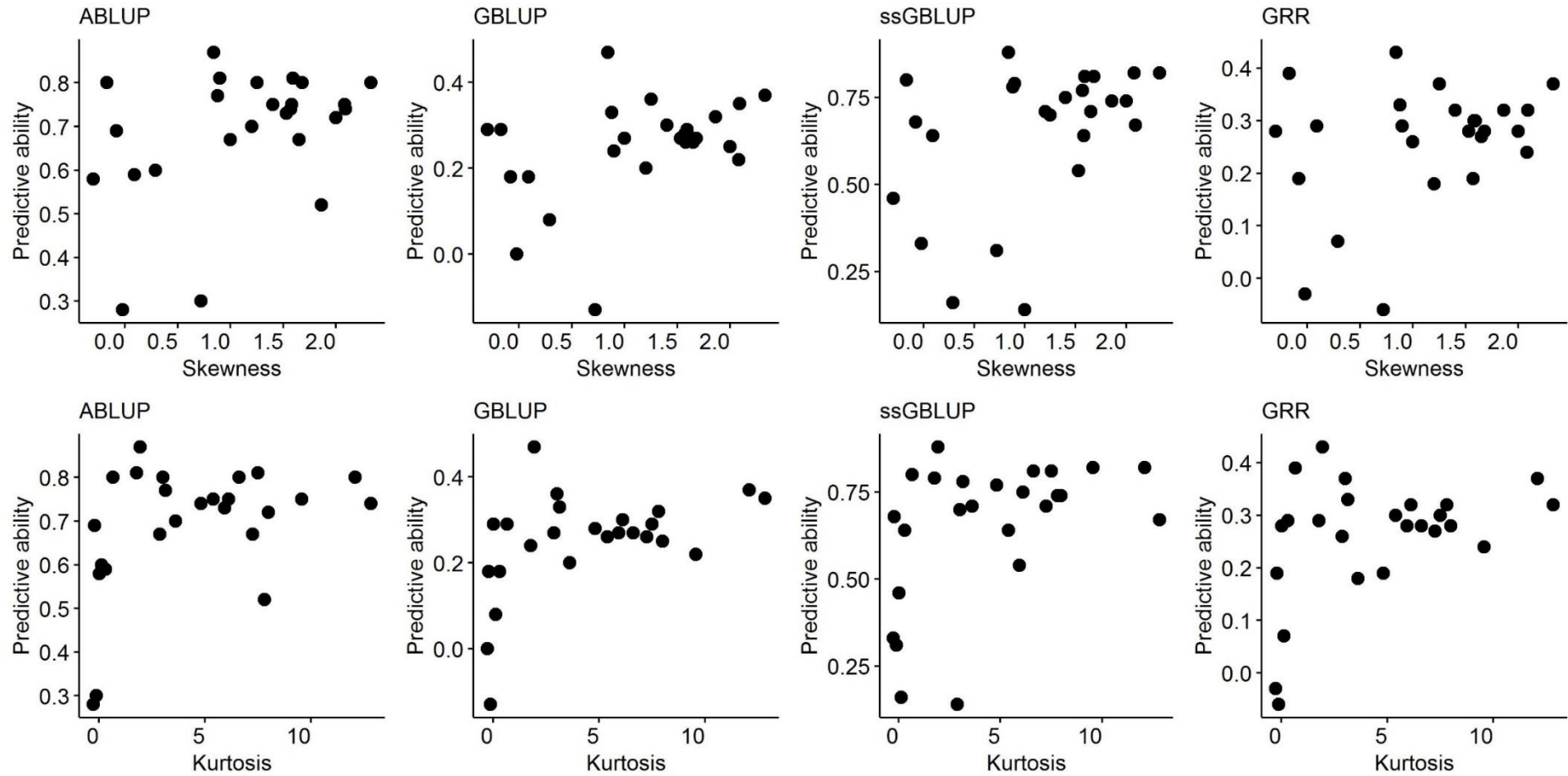
Figure 7.4: The relationship between predictive ability (PA) of the ABLUP, GBLUP, ssGBLUP and GRR and data distribution coefficient; skewness (a measure of symmetry) and kurtosis (a measurement of tails). Data with normal distribution has a skewness of zero and a kurtosis of 3.0, and therefore theoretically highest PA is expected at these points.

7.4   Discussion

In this study, we aimed to compare genomic models and conventional pedigree-based models for resistance and chemical traits and also understand how statistical distribution of data impacts the prediction models. Results showed: (i) significant improvement in heritability estimates for traits with genomic models; (ii) improvement in the accuracy of genetic correlation estimates with genomic models; (iii) comparable ability of the pedigree (ABLUP) and genomic based models to predict future phenotypes (i.e. predictive ability); and (iv) minimal impact of statistical distribution of the data on the genetic estimates.

Various studies have provided evidence that genomic models, using realized relationships based on marker information lead to a substantial increase in the prediction accuracies for various traits in trees compared to those using relationships based on pedigree information (Klápště *et al.* 2020a; Klápště *et al.* 2014; Ratcliffe *et al.* 2017), which is related to their ability to accurately trace pedigree relatedness and the ability to track Mendelian sampling variance (Klápště *et al.* 2018). In forest tree breeding, the single-step GBLUP (ssGBLUP) that combines phenotype, pedigree and genomic information through a single-step evaluation approach has been highlighted as the preferred strategy for evaluation of breeding values since it has not been practical to genotype all trees in the large progeny tests used in most forest tree breeding programmes (Ratcliffe *et al.* 2017). In this study, while direct evaluations for the impact of genomic models were made for chemical compounds and height, it was not possible to directly evaluate the impact of genomic models on bark stripping estimates since the genotyped trees were not scored for bark stripping. Therefore, the improvement detected in the genetic estimates of bark stripping indirectly links to the changes in the accuracy of estimating the genetic estimates of the chemical compound and height that were simultaneously analysed as response variables in trivariate models. A direct assessment of bark stripped, genotyped trees will give a better outlook of the impact of genomic models to the genetic parameters of bark stripping. While there are few comparative studies relating to herbivory in conifers, one recent study assessed the genomic prediction of resistance against weevil herbivory in spruce (Lenz *et al.* 2020). Lenz *et al.* (2020) indicated a two-fold reduction in the GBLUP heritability compared to pedigree-based prediction for the number of weevil attacks, and also the wood property traits assessed in the study, which could be related to the fewer SNPs (Chen *et al.* 2018) that was 4 - fold less than what has been used in this current study. Chen *et al.* (2018) also noted that improvement in genetic estimates increases with the number of families especially in full -sib populations, so possibly the 40 families (35 parents) used by Lenz *et al.* (2020) were not sufficient to capture the additive genetic variation, although the trees per family were greater compared to this

current study that used 75 families with utmost 6 trees per family. This may also imply that the number of families may be more important than the number of trees selected per family. In Norway spruce for example, PA for several wood traits stabilized as the number of trees within-family reached six (Chen *et al.* 2018). Therefore, the optimal parameters to realise the benefits of genomic selection for herbivory may need more research. However, more evidence of the superior performance of genomic models in resistance studies is available from pathosystems in *Pinus radiata* (Klápště *et al.* 2020b) and other conifers (Carpenter *et al.* 2018; González-Camacho *et al.* 2018; Resende *et al.* 2012a).

Comparing heritability estimates for ABLUP and genomic models showed that the application of marker-based models significantly improved the narrow-sense heritability estimates of individual chemical compounds compared to the pedigree-based (ABLUP) method. The chemical traits that did not exhibit significant additive genetic variation with the univariate ABLUP model did with the GBLUP model. The GBLUP performed better than the ABLUP, indicating that markers provided additional information. The ssGBLUP, compared to the trivariate ABLUP further improved the estimates highlighting the importance of additional individuals. However, for some traits like linoleic acid [83], no difference between GBLUP and HBLUP were detected, indicating that in this case the A-matrix did not provide additional genetic information, implying that the benefit of additional non-phenotyped data in the ssGBLUP is trait specific. Overall, the ssGBLUP and GBLUP should in theory perform well for traits that are under quantitative genetic control. Indeed there is evidence for their superior performance for different quantitative traits in conifers such as growth and wood traits (Cappa *et al.* 2019; Gamal El-Dien *et al.* 2018; Ratcliffe *et al.* 2017). The genetic control and genomic selection of most chemical compounds have been less studied, and the genetic architecture is less known. However, based on our results the high predictive ability estimates from the trivariate ssGBLUP model, equivalent to the trivariate ABLUP suggest a strong quantitative genetic control of the amounts of these compound groups. All terpenoids for example, are formed directly by one group of terpene synthase genes that can be modified by various enzyme classes, such as the cytochrome P450 hydroxylases (Bohlmann and Croteau 1999; Celedon and Bohlmann 2019). However, the extent to which the terpene synthases are modified by other genes is not well studied. Results however, indicated that there was a tendency for some compounds such as the fatty acid compound; linolenic acid and the sugar; glucose to be controlled by genes with major effects. Comparing GBLUP and GRR, a few individual compounds these compounds tended to be better predicted by the GRR that assumed heterogeneity of marker effects suggesting that there is a tendency for such compounds to be influenced by genes with major effects. In other plants, the amount of some sugars, for example, have been shown to be multigenic and controlled by many quantitative trait loci (Calenge *et al.* 2006). Although there do not appear to be

many comparable studies in plants, enhanced performance of non-BLUP genomic methods has been observed in livestock for fatty acids that are controlled by a few dominant genes (Verbyla *et al.* 2009). In practice, traits controlled by major genes can easily be incorporated in breeding programmes (Pérez-de-Castro *et al.* 2012).

Similar to the chemical traits, the heritability for height in this study improved with genomic models consistent with the observations in Norway spruce populations (Chen *et al.* 2018). The results of this study however contrasted with observations in white spruce and Norway spruce (Gamal El-Dien *et al.* 2018; Lenz *et al.* 2020), where multi-fold reduction in heritability values for height were documented for GBLUP vs ABLUP models. Likewise, in Douglas-fir, Thistlethwaite *et al.* (2019) did not show any relative improvement in heritability estimates in height with genomic models.

Multi-trait models mostly gave higher heritability estimates compared to univariate for most of the compounds, which is the norm for most traits (Karaman *et al.* 2020; Lenz *et al.* 2020). Multi-trait selection models can improve the accuracy of predictions by taking advantage of the genetic correlations between traits (Calus & Veerkamp, 2011). In our case, the benefits of multi-trait models were realised since chemical traits had fewer phenotypic records and could be better predicted genetically when coupled with other traits that were extensively assessed (Covarrubias-Pazaran *et al.* 2018). The increase in heritability estimates with multi-trait over single-trait models was especially high for traits that had very low heritability values in the single trait models. The magnitude of the genetic correlation from ABLUP relative to ssGBLUP however, did not change, consistent with Lenz *et al.* (2020) for the wood and herbivory traits in spruce. However, evidence for better performance of the ssGBLUP was derived from the reduction in the standard errors of the estimates.

The predictive ability (PA) (i.e. the ability to predict future phenotypes) did not vary markedly between the trivariate ABLUP and the ssGBLUP. This finding is consistent with some studies in conifers (Beaulieu *et al.* 2014; Chen *et al.* 2018) but not others where, potential improvement of PA with genomic models was shown (Gamal El-Dien *et al.* 2016; Goddard 2009; Hayes *et al.* 2009; Iwata *et al.* 2011; Klápště *et al.* 2018; Stejskal *et al.* 2018; Suontama *et al.* 2018). The lack of improvement in PA for the genomic vs ABLUP in this study could be due to various reasons; Firstly, the genotyped reference was not large enough to improve the genomic predictive ability, although it improved the heritability estimates for the current study population. Few comparative studies exist in conifers, but in animal studies for example, PA drastically reduced for ssGBLUP compared with ABLUP when the genotyped reference population was small (Lourenco *et al.* 2015; Song *et al.* 2019). In pigs, the change

in PA was insignificant for ssGBLUP compared with ABLUP when the genotyped reference population size was <500 (Song *et al.* 2019). Secondly, the parameters selected for the construction of the H-matrix; especially the w (the proportion of the genetic variation not captured by the markers) as well as α and β -the scaling factors have a significant impact on the predictive ability of the ssGBLUP. In this study, a low w was selected, which signified that most of the additive genetic variance was captured by the markers, which may not have been the case especially given that the genotyped population was small. The effect of the scaling factor, w, has been assessed in various studies and a range of optimum w for different traits, up to 0.95 have established (Oliveira *et al.* 2019; Song *et al.* 2019). Even then, these studies have indicated that w, α and β can be population and trait specific such that using the same parameters for different traits may lead to inaccuracy of prediction. Therefore, determining the optimal parameters for these traits is worth investigating. Thirdly, since the ABLUP and ssGBLUP were trivariate models, studies have also indicated that predictive ability in multi-trait models is influenced by the strength of the genetic correlation between the traits (Song *et al.* 2019), where a positive correlation is predicted between the magnitude of genetic correlation and predictive ability. In this study, the ssGBLUP genetic correlation did not improve relative to ABLUP, and consequently the PA. Even then, the study showed no linear relationship between predictive ability and genetic correlation that could be related to the change in direction that was observed in some correlations. Instead, predictive ability was strongly negatively correlated with the standard error of the genetic correlation estimates. In addition, for some compounds, for example glucose that seemed to have SNPs with major effects contributing to the genetic estimations, PA could be improved by utilizing the subset of markers with the largest positive effects rather than all markers (Chen *et al.* 2018). Finally, the spread of the data around the mean was of concern in this study since chemical data is often skewed and this distribution has a potential influence on the estimation of the genetic parameters (Kärkkäinen and Sillanpää 2012; Muranty *et al.* 2015). However, skewness (a measure of symmetry) and kurtosis (a measurement about the extremities, i.e. tails, of the distribution of data, that provides an indication of the presence of outliers) had negligible impact on the predictive ability of the pedigree- based or the genomic-based models. The data used in the study did not deviate strongly from normality for most compounds, where data with substantial departure from normality has as an absolute skew value > 2 and kurtosis of > 7 (Kim 2013). Overall, minimal PA was low for traits with low ABLUP consistent with other studies (Arojju *et al.* 2019; Lenz *et al.* 2020).

Comparing the univariate and trivariate models, the PA for trivariate ssGBLUP was higher than that of the univariate GBLUP and GRR. It has been indicated that prediction accuracy for a low-heritability trait could be significantly improved by multivariate genomic selection when a correlated high-heritability trait

was available (Jia and Jannink 2012). In contrast, where correlations are low or non-significant, there will be erroneous information sharing across traits leading to poor performance of multivariate models (Jia and Jannink 2012). In this study because, most of the correlations were not significant, and that PA for univariate ABLUP did not differ from the trivariate ABLUP, the better PA performance of the trivariate ssGBLUP compared to the univariate GBLUP and GRR could be due to a relatively small sample size used in these models. The heritability for height was lower than what has been estimated in other *P. radiata* populations using both ABLUP and genomic models (Li *et al.* 2018), which could be related to the age of the trees. Weng *et al.* (2007) for example, showed significant improvement in the heritability of *P. banksiana* with age. In Chapter 2, the heritability estimates of height in the older trials was also higher than that of the younger trial assessed in this study.

7.5   Conclusion

The results indicate improvement in the genomic models to predict heritability of primary and secondary chemical traits compared to the traditional ABLUP. Chemical traits that did not show evidence of additive genetic variation with the ABLUP model were significant with genomic models. Improvement in accuracy of genetic correlations was detected. Since genomic models, are important for traits with low heritability, this should be a motivation for its employment in forestry where most traits are complex. For predicting future phenotypes, no relative advantage was detected for genomic over pedigree prediction for bark stripping, while some improvements for predicting specific terpenes, sugars and fatty acids was observed.

1. Whole field

26 replicates,
78 incomplete blocks
n=1970,
74 full-sib families
55 and 54 grand
parents

2. Sections

6 replicates
protected,
n=420

20 replicates
exposed to
bark stripping,

n=1550

3. Data collected

Bark chemistry, SNP
genotyping, height

bark stripping scores,
height

Supplementary Figure 7.1: Set up of the genetic trial and the data collected from the different sections. However, the 6 protected replicates were randomly spread in the entire field. Height was the only variable assessed in all the 26 replicates, and in the trivariate models it was the bridging trait between the 20 unprotected and the 6 protected replicates.
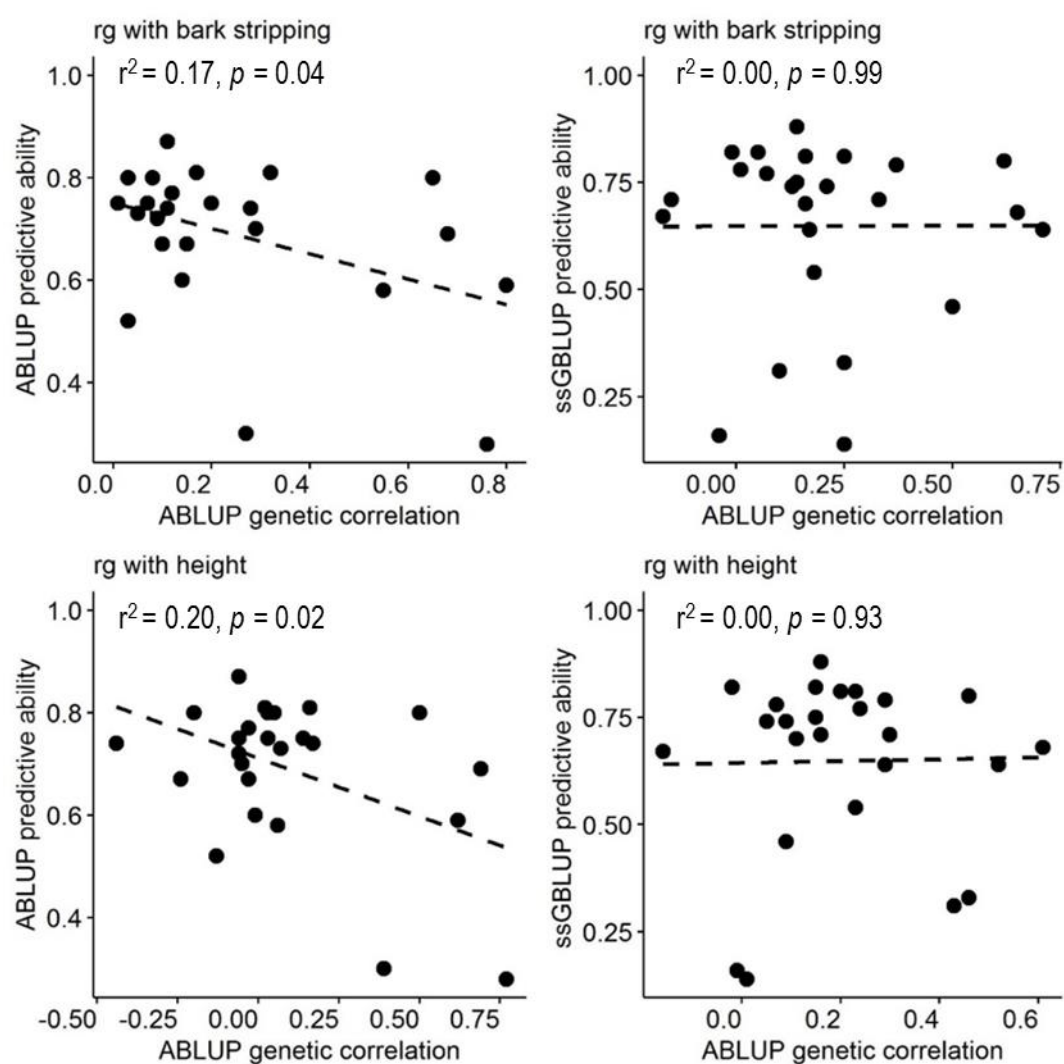
Supplementary Figure 7.2: The coefficient of determination (r^2) between the ABLUP genetic correlation of chemical compounds with bark stripping (above) and height (below). The trivariate models included spatially adjusted bark stripping, height and a chemical compound.

Chapter 8: Constitutive and induced transcriptome analysis of the needles and bark of *Pinus radiata*

8.0 Abstract

Plants are attacked by diverse insect and mammalian herbivores and respond with different physical and chemical defences. The changes in the phenotype usually underly transcriptional changes. Simulated herbivory has been used to study the transcriptional and other early regulation events of these plant responses. In this study, constitutive and induced transcriptional responses to artificial bark stripping were compared in the needles and the bark of *Pinus radiata* to the responses from application of the plant stressor, methyl jasmonate. The time progression of the responses was assessed over a 4-week period. The constitutive transcriptome was dominated by genes related to defence and photosynthesis and did not differ between the needles and the bark. Following bark stripping and methyl jasmonate treatments, there was an up-regulation and down-regulation of genes associated with primary and secondary metabolism both in the needles and the bark. The genes related to primary metabolism were more responsive compared to those associated with secondary metabolism. The up-regulation of genes related to sugar break-down and the repression of photosynthesis related genes was consistent with the strong down-regulation of sugars that was observed in the previous chapter (Chapter 3). While the regulation of genes involved in signalling, photosynthesis, carbohydrate and lipid metabolism, defence and water stress was mimicked well by the treatments, non-overlapping transcripts were detected between the needles and the bark, treatments and different times of assessment. Methyl jasmonate caused more responses in the bark than bark stripping although the peak of expression following both treatments was detected 7 days post treatment application. The effects of bark stripping were only localised, and no systemic changes were detected in the needles. Whether the gene expression changes are heritable and how they differ between resistant and susceptible families identified in previous work (Chapter 4) needs further investigation.

8.1 Introduction

Plants have evolved a variety of constitutive and inducible defences to resist and tolerate herbivory. An assessment of the genetic mechanisms that influence these defences will enhance our understanding of their evolution (Anderson and Mitchell-Olds 2011). Although studies suggest that structural changes in DNA are the major source of genetic variation (Mitchell-Olds *et al.* 2007), the phenotypic outcomes of several traits have also been linked to gene expression (D'Agui *et al.* 2016; Eldar *et al.* 2009; Guo *et al.* 2016; Li *et al.* 2019; Raj *et al.* 2010) but the genes and genetic pathways that underlie most phenotypes are still unknown (Mitchell-Olds *et al.* 2007). To date, most gene expression studies have focussed on identifying transcripts or genes showing differential expression, or pathways associated with a phenotype (case/control) or condition (treated/untreated). In conifers for example, transcript abundance has been examined with respect to biotic and abiotic environmental factors such as herbivory (Lamara *et al.* 2018; Verne *et al.* 2011), pathogens (Kovalchuk *et al.* 2019), artificial wounding (Ralph *et al.* 2006), drought (Behringer *et al.* 2015), light intensity (Ranade *et al.* 2019), seasonal changes (Cronn *et al.* 2017), chemical stressors like methyl jasmonate (Liu *et al.* 2017a), as well as associated phenotypic traits such as resistance and chemical composition (Lamara *et al.* 2018; Verne *et al.* 2011).

In conifer-herbivory studies, most gene expression studies have focused on understanding induced defence responses, with a premise that these may be more important than constitutive defences as they are cost effective and expressed only when required (Kant *et al.* 2015; Moreira *et al.* 2014). Global transcriptome responses have been studied in the needles and the bark, monitoring the expression of a wide range of genes related to the biosynthesis of primary and secondary compounds plus structural components (Du *et al.* 2018; Kānberga-Siliṇa *et al.* 2017; Kolosova 2010; Litvak and Monson 1998; Miller *et al.* 2005; Ralph *et al.* 2006; Zulak and Bohlmann 2010). Most of these genes are expressed at basal levels in plants but some are only expressed in the presence of an appropriate stimulus. Some of the genes significantly respond to herbivory cues, by increasing or reducing their expression locally at the site of the perceived effect or systemically throughout the plant (Byun-McKay *et al.* 2006; Keeling and Bohlmann 2006; Miller *et al.* 2005). Studies also show a high overlap in the genes that are differentially expressed when plants are subjected to different biotic and abiotic stresses (Kovalchuk *et al.* 2015; Reymond *et al.* 2004). However, the genes that show differential expression have been indicated to differ within and between target plant species (Kolosova 2010; Verne *et al.* 2011), between plant tissues (Martin *et al.* 2002; Miller *et al.* 2005), as well as biotic agents (Korth 2003) and applied treatments (Men *et al.* 2013). Intra-specific differences in the timing of transcript expression have also been observed, where plants may respond to injury within hours or days, with short or long lasting

effects (Litvak and Monson 1998; Liu *et al.* 2017a; Martin *et al.* 2002; Miller *et al.* 2005). Plant responses to different classes of herbivores may differ due to differences in herbivore oral secretions or mode of feeding and the amount of plant tissue damage (Agrawal 2000; Korth 2003; Ohse *et al.* 2017). While available conifer studies have documented changes in gene expression in response to insect herbivory (Kovalchuk *et al.* 2015; Ralph *et al.* 2006), there are no studies from the perspective of mammalian herbivory, and none that link changes in gene expression to changing chemistry. Mammalian bark herbivory is fundamentally different from insect herbivory in the mode of feeding (Chapter 2 ; Kant *et al.* 2015) and possibly the oral secretions. This particularly applies to mammalian bark stripping, which is of increasing concern to managers of conifer forests world-wide including *Pinus radiata* plantations in *Australia* (ABARES 2018; Cukor *et al.* 2019; Nagaike).

Pinus radiata is native to California (Eldridge 1979) but is now a major plantation species in Australia, where native marsupials (wallabies and kangaroos) are currently the primary cause of bark stripping. Native mammals strip the bark of the trees during the early stages of growth (Miller *et al.* 2014; Page *et al.* 2013; Smith *et al.* 2020), reducing tree growth rate (Chapters 2 and 4). Chemical profiling in *P. radiata* shows that needles and bark respond differentially to bark stripping and other forms of real and simulated herbivory, by mostly increasing levels of secondary compounds, especially terpenes and phenolics (Lundborg *et al.* 2019; Moreira *et al.* 2012a), and reducing levels of sugars and fatty acids (Chapters 3, 4 and 6). This suggests changes in the expression of underlying genes that subsequently transforms the chemical phenotype. The differences in timing of the induced changes in terpenes, phenolics and sugars (Chapter 3 ; Reglinski *et al.* 2017) suggest corresponding differences in the expression of the underlying genes. The molecular basis of these induced changes have not been characterised, although various studies have documented *P. radiata* transcriptomes with respect to changes in ontogeny and during wood formation (Alvarez *et al.* 2011; Li *et al.* 2013; Telfer *et al.* 2018).

The present study aimed to quantify and compare the transcriptome changes that occur in response to artificial bark stripping of *Pinus radiata* and whole plant stress induced by application of the chemical stressor, methyl jasmonate. The longer-term goal is to identify genes that mediate the previously observed induced chemical responses to simulated bark stripping in *P. radiata*. The specific aims of the study were to: 1) characterise and compare the constitutive transcriptome of *P. radiata* needles and bark; and 2) identify genes which are differentially expressed following artificial bark stripping (aimed at mimicking mammalian bark stripping) and whole plant application of methyl jasmonate and compare the induced responses. The results are discussed in view of the holistic chemistry that has been characterised on the same individuals with the same treatments (Chapter 3).

8.2   Materials and Methods

8.2.1   Experimental design

The experiment and the genetic material used for this transcriptomic study are reported in Chapter 3. In 2015, 6-month-old seedlings from 18 full-sib families of *Pinus radiata* (D. Don) originating from the Radiata Pine Breeding Company deployment population, were obtained from a commercial nursery. Seedlings were transferred into 145 mm x 220 mm pots containing 4L of basic potting mix (composted pine bark 80% by volume, coarse sand 20%, lime 3 kg/m^3 and dolomite 3 kg/m^3) and raised outside in a fenced area (to protect against animal damage) at the University of Tasmania, Hobart. At 2 years of age, methyl jasmonate was applied to seedlings of 6 families, artificial bark stripping treatments were applied to another 6 families, and the remaining 6 families were kept as controls (Chapter 3; Figure 8.1). Methyl jasmonate (MJ) was applied in a 25 mM solution by spraying the whole plant with a fine mist from a hand sprayer until 'just before run-off'. The treated seedlings were sprayed in a well-ventilated area away from untreated seedlings to avoid cross contamination (Moreira *et al.* 2013a). For bark stripping (strip),18 plants were artificially stripped by removing a 30 cm vertical strip, beginning 2 cm from the ground and covering 50% of the stem circumference, which is the average upper threshold of browsing observed in natural field conditions for families (Chapter 4). The treatment groups (control, strip and MJ) were replicated 3 times and arranged in a randomized block design of 3 blocks in a shade house. Four seedlings of each family were grown in a line plot and one was chosen at random for destructive harvesting at each time. The treatment plots were separated within each block to minimise any interference from treatments.
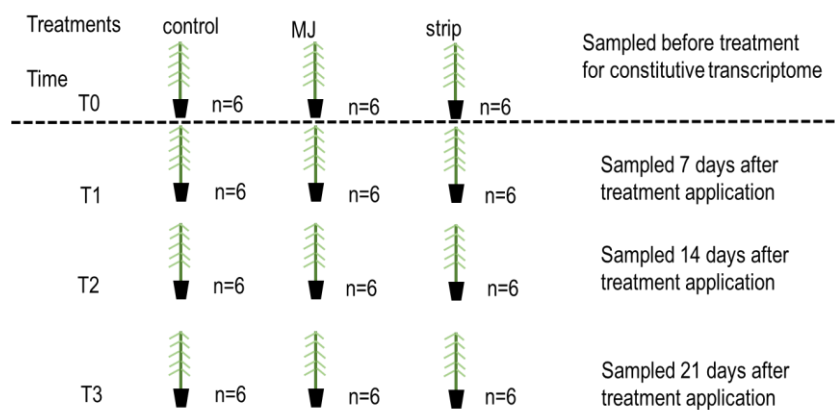
Figure 8.1: The treatments, sample size and pairwise comparisons that were made for each time and for the two treatments - bark stripping (strip) and methyl jasmonate (MJ). The seedlings of each family were grown in a line plot and one was chosen at random for destructive harvesting at each time (T0 to T3). At T0, the sampled seedlings were destructively harvested just before treatment applications. At 7 (T1), 14 (T2) and 21 (T3) days after treatment, one seedling from each family and each treatment was destructively harvested (n=3 blocks x 3 treatments x 2 families = 18).

8.2.2　Sample processing

One seedling from each family was sampled weekly from T0-T3 (Figure 1). T0 represents the time immediately before treatment applications. T1, T2 and T3 represents respective sampling times at 7, 14 and 21 days after treatment application. Seedlings were destructively sampled at each time and randomly selected from each family. Each treatment was equally represented by a random seedling from a total of 6 families at each sampling time. In total 72 needle and 72 bark samples were collected (3 treatments x 3 replicates x 2 families x 4 sample times = 72). Up to 20 young needles were randomly collected per seedling from different parts of the crown. The bark was sampled from different points of the stem carefully avoiding the wood as presented in Chapter 3. Individual samples were kept separate providing 144 samples for sequencing (2 plant parts x 3 treatments x 6 families x 4 sample times). The needles and bark samples were immediately frozen in liquid nitrogen and were stored at −80 °C until RNA extraction. The 6 families sampled from each treatment at each time point were treated as biological replicates. No technical replicates were included. This sampling occurred at the same time when the tissue for the chemistry assays reported in Chapter 3 was sampled.

8.2.3　RNA extraction and sequencing

RNA from all the 144 bark and needle samples was extracted using the Spectrum™ Plant Total RNA kit (Sigma Aldrich, lot # SLBW2113). The RNA extraction was random with respect to part, sampling time, treatment, family and shade house replicate. The quality and quantity of the RNA extracts were assessed with an Agilent 5200 Fragment Analyzer (Palo Alto, California, USA). One sample had poor quality RNA and was excluded from further processing. Using the high-quality RNA samples, 143 separate libraries were prepared with a 6-bp nucleotide bar-coding tag for each library. To construct the library, approximately 1 µg of total RNA was used following the MGIEasy RNA Directional Library Prep Kit (MGI). Paired-end sequencing was performed using the Beijing Genomics Institute (BGI) MGISEQ-2000 sequencer according to the manufacturer's instructions, yielding 100-bp paired-end reads and a total of 20m reads per sample. Tagged cDNA libraries were sequenced in separate lanes. The library for each lane was selected at random. The quality of RNAseq sequences was assessed using FastQC version 0.11.8 (Andrews 2018). Quality trimming and filtering of data was performed using Trimmomatic v 0.39 (Bolger *et al.* 2014). On average, 99.9% of the sequences were retained at phred33 (Ewing and Green 1998). A *de novo* assembly of the pooled transcriptome was attempted using TRINITY v2.9.0 (Grabherr *et al.* 2011) but due to the excessive computation requirements, could not be completed with the available resources in the required timeframe. Accordingly, the filtered reads were aligned to the *Pinus radiata* reference transcriptome that is harboured at the SCION Research Institute in New Zealand (Telfer *et al.* 2018) using SALMON v0.14.1(Patro *et al.* 2017). This reference transcriptome

was assembled from a range of *Pinus radiata* genotypes and tissue types that were collected at different developmental and temporal stages. Most of the samples were from healthy seedlings under normal growth conditions but also included some pathogen infected seedlings (Telfer *et al.* 2018). The reference transcriptome has a total of 279,510 unique transcripts.

8.2.4 Differential transcripts expression analysis

Statistical analysis of differential expression was performed using edgeR v3.24.3 (Robinson *et al.* 2010) package in R (v3.6.0) (R Core Team 2018). Transcripts were first filtered retaining only transcripts with a minimum expression fold change of 2 and with a minimum of 100 copies of a single transcript in at least two groups. To adjust for library sizes and skewed expression of transcripts, the estimated abundance values were normalized using the trimmed mean of M-values normalization method included in edgeR. To select the cut-off false discovery rate (FDR) for the different groups being compared, transcript expression was initially compared between the samples collected from the control plants (n=6), MJ-allocated (n=6) or strip-allocated groups (n=6) at T0 (before treatment) to check the inherent differences between sample groups. The p-values at which no differential expression was detected between these groups was set as the FDR for downstream pairwise comparisons. Respectively, the p-value for detecting differentially expressed transcripts (DET) in the treated needles following both MJ and bark stripping was set at 1.0×10^{-11}. A p-value of 1.0×10^{-18} was set to detect DET in MJ treated bark and 1.0×10^{-10} to detect expression in the bark stripped samples. For each plant part, comparisons were thereafter made between the control (n=6) and methyl jasmonate (MJ, n=6) and the control (n=6) and bark stripping (strip, n=6) treatments at each sampling time (T1, T2, T3) (Figure 8.1). Twelve pairwise comparisons were performed. Venn diagrams were used to find the transcripts that were identified as significantly differentially expressed across different comparisons and were generated using bioinformatics.psb.ugent.be/webtools/Venn/.

Unsupervised cluster analysis was performed to detect dominant, relative expression patterns across the needles and the bark and the treatments. Following Ralph *et al.* (2006), a subset of 500 transcripts with the highest variability and highest expression across the 143 libraries were selected in edgeR for this analysis. Clustering and heat maps were generated using heatmap.2 function from the gplots package in R, with a matrix of Euclidean distances from the log2 counts of normalised transcripts.

8.2.5 Sequence similarity search

For sequence similarity search and functional analysis of differentially expressed transcripts (DETs) the transcripts were blasted against the nucleotide BLAST database using BLASTn

(https://blast.ncbi.nlm.nih.gov/Blast.cgi). BLAST analysis revealed that more *P. radiata* transcripts were most similar to those predicted from genome sequences of *P. taeda* (BLASTn with e- value <0.0001). Other species, mostly *Pinus sylvestris*, *P. monticola*, *Picea stichensis* and *Pseudotsuga menziesii*, showed high similarity with the *P. radiata* transcripts. Annotations of selected transcripts was done by comparing *P. radiata* transcripts to the sequences in the SwissProt annotated genes (Bairoch and Apweiler 2000) using cut-off values ≤ 1. To get clear patterns of the responses, only transcripts associated with genes of known function were included. However, there were many uncharacterised transcripts and proteins of unknown functions.

8.2.6   GO enrichment analysis

To understand differences in the biological process, cellular component and molecular function categories, gene ontology (GO) enrichment analysis was performed on selected transcripts. At T0, enrichment of differentially expressed transcripts aimed to understand the constitutive differences of the GO processes between the transcriptome of the needles and the bark. In addition, the GO enrichment that was performed on selected T1 transcripts aimed to understand differences in the up-regulated and down-regulated transcripts as well as differences in the induced transcriptome of the strip and MJ treated samples. Due to the limited annotation resources available for conifers, gene family annotations were obtained using genomes of 10 species: *Arabidopsis thaliana*, *Citrus sinensis*, *Cucumis sativus*, *Oryza sativa, Populus trichocarpa*, *Prunus persica, Saccharomyces cerevisiae*, *Theobroma cacao, Vitis vinifera* and *Zea mays*. Annotations were obtained via protein analysis through evolutionary relationships (PANTHER) version 14.1(Mi *et al.* 2019). GO analysis was done by comparing the GO terms enriched in the different conditions (time x treatment x part).

8.3   Results

8.3.1   The Pinus radiata reference transcriptome and read mapping

RNA-seq of *P. radiata* generated a total of 2,860 million 100-bp PE reads with approximately 20 million reads from each of the 143 samples. Only 0.8% of the reads were mapped to the reference transcriptome but 87.6% of the reference transcriptome was represented among the study transcripts. However, after the filtration criteria described above, only 6312 unique transcripts (2.6% of the reference transcriptome) were retained as the expression of the other transcripts was too low.

8.3.2   Overall summary of the transcriptome

To visualise the overall relationships between the transcriptome from the different samples, the unsupervised hierarchical clustering of the top 500 variable transcripts in the transcriptome showed that the major differences were concerning the plant parts (top x-axis). Among the most variable transcripts of the needles and the bark (y-axis), two major clusters of transcripts (1 & 2) categorised based on expression patterns by part, time and treatment were identified (Figure 8.2). Within the clusters, we noted genes that were:

    (i)   up-regulated in the needles relative to the bark and generally non- responsive to treatment;

    (ii)   up-regulated in the bark relative to the needles and generally non-responsive to treatment;

    (iii)   up-regulated in either the needles or the bark and responsive to treatment;

    (iv)   not differentially expressed between the needles and the bark but responded to treatment
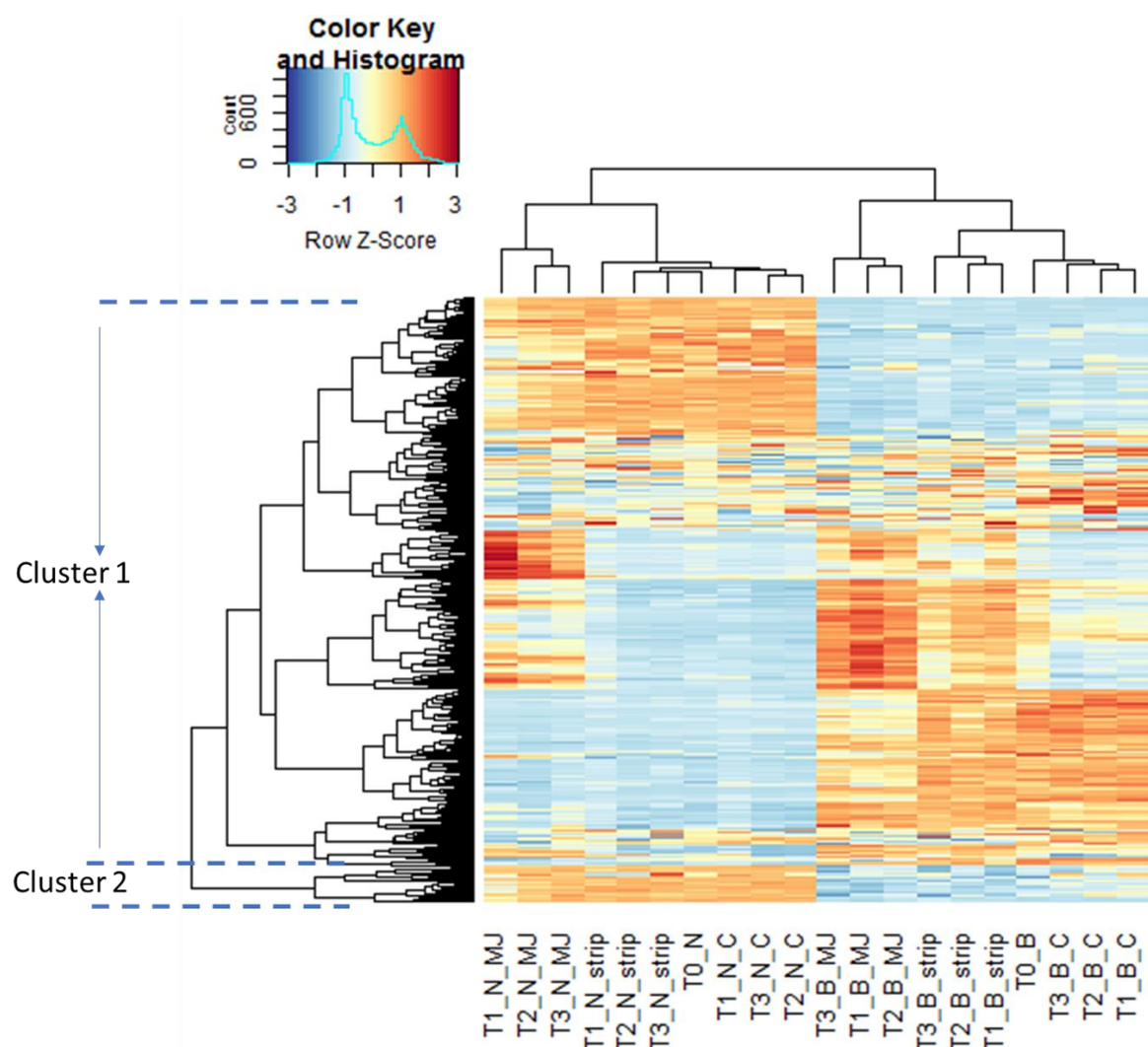
Figure 8.2: Hierarchical cluster analysis of the top 500 most variable transcripts selected by edgeR in the needles (N) and bark (B) treated with methyl jasmonate (MJ) and artificial bark stripping (strip) and control (C), 7 (T1), 14 (T2) and 21 (T3) days after treatment application. Transcripts (rows) and time/part/treatment categories (columns) were clustered using Euclidean distance. The Z-score is calculated by subtracting the score of the individual from the mean of the grand mean of all the individuals and then divided by the standard deviation. The colours; yellow = mean expression, blue= expression below the mean and red = expression above the mean. The categories on the X-axis were reshuffled based on similarity. Clusters 1 and 2 are highlighted as the major nodes on the Y-axis.

8.3.3 Differences in the constitutive needle and bark transcriptome

Almost all the 6312 transcripts analysed were detected in both the needles and bark. Five transcripts were detected in the needles only and 13 in the bark only, and most of these were uncharacterised (Table 8.1). Gene level annotation of the top 10 transcripts expressed in each plant parts are listed in Table 8.2. The type 2 light-harvesting chlorophyll a/b-binding polypeptide (TLH) that harvests photons that are converted to biochemical energy and biomass through photosynthesis, was the most expressed gene in both the needles and the bark and was represented by different copies of transcripts (isoforms). The needles also had other photosynthesis-related genes such as ribulose bisphosphate carboxylase/oxygenase (RuBisCO) and PSI-D1 precursor (Table 8.2) possibly due to its major role in photosynthesis. Genes related to secondary metabolism were also detected among these top 10 genes, suggesting that constitutive defence is important in *P. radiata*. These included dehydrin [2], metallothionein [3], chalcone synthase [4], defensin [5] and pathogenesis-related proteins [8] and were more prominent in the bark than in the needles but their relative expression was not significantly different between the needles and the bark.

Approximately 1478 out of the 6312 transcripts (23.4%) were differentially expressed in the needles or the bark. Of these, 938 were up-regulated in the bark compared to the needles, while 540 transcripts were up-regulated in the needles. The 10 most strongly differentially expressed transcripts in both bark and needles are shown in Table 8.3. Besides the general function-genes and those related with photosynthesis, there was an upregulation of genes related to terpene [B9] and lipids biosynthesis [B7] in the bark and those related to sugars [N4] and phenolics biosynthesis [N1] in the needles. Of note were the genes involved in sugar transport in both the needles [N3] and the bark [B2].

Table 8.1: Transcripts that were unique to each *Pinus radiata* plant part in the constitutive transcriptome as assessed at T0 (sampled before treatment). The SCION transcript code, predicted gene name and predicted functions of the known genes are indicated

P. radiata transcript code	Gene name	Gene function
Transcripts expressed in the needles but not in the bark at T0		
NZPradTrx008090_C01	Unknown	
NZPradTrx102814_C01	Hypothetical protein 0_2136_01	
NZPradTrx114705_C04	PREDICTED: uncharacterized LOC101213828	
NZPradTrx119356_C01	Repetitive proline-rich cell wall protein 2 precursor, putative	Key determinant of many cell wall proteins https://www.uniprot.org/uniprot/Q40375
NZPradTrx138443_C01	Unknown	
Transcripts expressed in the bark but not in the needles at T0		
NZPradTrx105287_C05	Chloroplast ELIP early light-induced protein	Prevents photooxidative stress (Hutin *et al.* 2003)
NZPradTrx068786_C02	Unknown	
NZPradTrx110900_C02	Unknown	
NZPradTrx158724_C01	Unknown	
NZPradTrx111161_C02	Embryo-abundant protein	May act as a cytoplasm protectant during desiccation. https://www.uniprot.org/uniprot/P46520
NZPradTrx032755_C01	Unknown	
NZPradTrx054373_C01	Unknown	
NZPradTrx151188_C01	Unknown	
NZPradTrx007008_C01	Unknown	
NZPradTrx069030_C01	Unknown	
NZPradTrx081218_C01	Unknown	
NZPradTrx154223_C01	PREDICTED: tetrahydrocannabinolic acid synthase-like	Catalyzes the oxidative cyclization of the monoterpene moiety in cannabigerolic acid https://www.uniprot.org/uniprot/Q8GTB6
NZPradTrx189870_C01	Uninformative	

Table 8.2: Top most expressed transcripts in the constitutive transcriptome of the bark and the needles as assessed at T0 (sampled before treatment), indicating the SCION transcript name, gene name and predicted function. Some transcripts were represented by different copies of the transcripts (isoforms- represented by different transcript codes in each row) and the percentages of transcripts that represented by each isoform are indicated. Each isoform has a superscript linking it to its corresponding percentage number of transcripts identified. B[a]= first isoform identified in the bark for the gene, N[a] =first isoform one identified in the needles etc. The transcripts were not differentially expressed between the bark and the needles

ID	*P. radiata* code (or isoforms)	Gene name	predicted gene function	percentages of transcripts (out of 6312)	
				Bark	Needles
1	NZPradTrx107583_C02[Ba, Na] NZPradTrx050124_C01[Bb, Nb] NZPradTrx118940_C01[Bc, Nc] NZPradTrx107583_C01[Nd] NZPradTrx107583_C03[Ne] NZPradTrx050061_C01[Nf]	Light-harvesting chlorophyll a/b-binding polypeptide (Lhcb2) mRNA	Absorb sunlight and transfer the excitation energy to the core complexes of PSII in order to drive photosynthetic electron transport (Liu *et al.* 2013b)	1.46[Ba],0.28[Bb], 0.25[Bc]	1.99[Na], 0.95[Nb], 1.07[Nc], 0.51[Nd], 0.51[Ne], 0.33[Nf],
2	NZPradTrx100458_C02[Ba] NZPradTrx100458_C03[Bb]	Dehydrin 7	Involved in dehydration stress (Stival Sena *et al.* 2018)	1.38[Ba], 0.60[Bb]	
3	NZPradTrx112612_C02[Ba, Na] NZPradTrx085990_C02[Bb] NZPradTrx094970_C01[Nb] NZPradTrx094970_C02[Nc]	Metallothionein 3	Play important roles in metal homeostasis and protection against heavy metal toxicity (Nevrtalova *et al.* 2014)	0.82[Ba],0.29[Bb]	0.58[Nc],1.75[Na], 0.66[Nb]
4	NZPradTrx052720_C01[Ba] NZPradTrx115271_C03[Bb] NZPradTrx078806_C01[Bc, Na] NZPradTrx115271_C02[Bd] NZPradTrx115271_C05[Be]	Chalcone synthase	Plays crucial roles in phenolic biosynthesis (Dixon and Paiva 1995)	0.70[Ba], 0.37[Bb], 0.35[Bc],0.27[Bd], 0.26[Be]	0.30[Na]

#		Protein	Function		
5	NZPradTrx050994_C02 [Ba] NZPradTrx050994_C01 [Bb]	Defensin	Inhibit the growth of a broad range of pathogens, including bacteria, fungi and viruses (Ermakova *et al.* 2016a; Picart *et al.* 2012).	0.61 [Ba], 0.53 [Bb]	
6	NZPradTrx076819_C01	TCTP-like protein	Implicated in important cellular processes, such as cell growth, cell cycle progression, malignant transformation and in the protection of cells against various stress conditions and apoptosis (Bommer and Thiele 2004)	0.42	
7	NZPradTrx062252_C01 [Ba,] NZPradTrx107621_C01 [Bb]	Nonspecific lipid transfer protein	Play important roles in resistance to biotic and abiotic stress. have the ability to bind or transfer various types of hydrophobic molecules in vitro, such as fatty acids, fatty acyl-CoA, phospholipids, glycolipids and cutin monomers (Liu *et al.* 2015a)	0.27 [Ba], 0.26 [Bb]	
8	NZPradTrx116410_C12	Pathogenesis-related protein 10	Show biological activities related to disease resistance (Liu and Ekramoddoullah 2006)	0.26	
9	NZPradTrx077717_C01	LP3-1	Implicated in water-stress https://www.uniprot.org/uniprot/Q5G154	0.24	
10	NZPradTrx100333_C02	ASR protein	Involved in sugar and abscisic acid signalling (Çakir *et al.* 2003)	0.25	0.24
11	NZPradTrx098632_C01	Translation elongation factor-1 alpha	Catalyses the transfer of aminoacylated-tRNAs (Sasikumar *et al.* 2012)		
12	NZPradTrx098233_C03 [Na] NZPradTrx064995_C01 [Nb] NZPradTrx064875_C01 [Nc] NZPradTrx098233_C01 [Nd]	Ribulose bisphosphate carboxylase//oxygenase	Catalyses carboxylation of RuBP in the first step of the Calvin cycle of photosynthesis (Tabita 1999)		1.57 [Na], 0.59 [Nb], 0.53 [Nc], 0.36 [Nd], 0.30 [Ne], 0.22 [Nf]

	NZPradTrx098233_C05 [Ne] NZPradTrx064875_C02 [Nf]			
13	NZPradTrx098207_C02 [Na] NZPradTrx098207_C01 [Nb]	Cysteine proteinase inhibitor CPI-3	Involved in plant development and defence, especially in the regulation of stress responses (Li *et al.* 2015a)	0.77 [Na], 0.27 [Nb]
14	NZPradTrx105813_C01	PREDICTED: probable fructose-bisphosphate aldolase 2, chloroplastic-like	It plays a key role in glycolysis and gluconeogenesis https://www.uniprot.org/uniprot/Q944G9	0.37
15	NZPradTrx111299_C01 [Na] NZPradTrx100425_C01 [Nb]	PREDICTED: oxygen-evolving enhancer protein 1, chloroplastic-like isoform 2	Stabilizes the manganese cluster which is the primary site of water splitting https://www.uniprot.org/uniprot/P23321	0.35 [Na], 0.32 [Nb]
16	NZPradTrx065162_C02	Thiazole biosynthetic enzyme	Thiamine synthesis and DNA damage tolerance (Liu *et al.* 2015b)	0.34
17	NZPradTrx184720_C01	PSI-D1 precursor	PsaD can form complexes with ferredoxin and ferredoxin-oxidoreductase in photosystem I (PS I) reaction centre. https://www.uniprot.org/uniprot/Q9S7H1	0.22

Table 8.3: Top 10 differentially expressed genes at T0 (before treatment). These are genes that were differentially expressed in either plant part without treatment

Part		*P. radiata* code	Predicted gene name	Predicted gene function
Bark	B1	NZPradTrx054097_C01	Homeobox transcription factor KN3	Central regulators of meristem cell identity (Guillet-Claude *et al.* 2004)
	B2	NZPradTrx073079_C03	Transporter, putative	Sugar transport (Weig *et al.* 1994)
	B3	NZPradTrx087709_C01	Homeobox transcription factor KN1	Central regulators of meristem cell identity (Namroud *et al.* 2010)
	B4	NZPradTrx055579_C01	Mini zinc finger 1	Regulates several development aspects, including photomorphogenesis, apical dominance, longevity, flower morphology and fertility, as well as root and stem elongation (https://www.uniprot.org/uniprot/Q9CA51)
	B5	NZPradTrx048496_C01	Plastid phosphate translocator	Involved in the exchange of metabolites and inorganic phosphate between stroma and cytosol (Bockwoldt *et al.* 2019)
	B6	NZPradTrx101882_C01	Auxin-induced protein 5NG4, putative	Transmembrane transporter activity especially during root formation (Busov *et al.* 2004)
	B7	NZPradTrx103825_C01 NZPradTrx103825_C04	PREDICTED: GDSL esterase/lipase At5g03610-like	Lipid catabolic process (https://www.uniprot.org/uniprot/Q9LZS7)
	B8	NZPradTrx184572_C01	G1-like protein	Polymerizes the backbones of non-cellulosic polysaccharides (hemicelluloses) of plant cell wall https://www.uniprot.org/uniprot/Q570S7
	B9	NZPradTrx055645_C01 NZPradTrx096935_C03	PREDICTED: squalene monooxygenase-like	Converts squalene into oxidosqualene, the precursor of all known angiosperm cyclic triterpenoids (Rasbery *et al.* 2007)
	B10	NZPradTrx093053_C01	Ribulose 1,5-bisphosphate carboxylase/oxygenase small subunit	Catalyses carboxylation of RuBP in the first step of the Calvin cycle of photosynthesis (Tabita 1999)
Needles				

N1	NZPradTrx115678_C04 NZPradTrx115678_C05	Anthocyanidin reductase	Involved in the biosynthesis of proanthocyanidins (Zhu *et al.* 2015)
N2	NZPradTrx090889_C01	Cytochrome P450 CYPA2	Oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen https://www.uniprot.org/uniprot/A9F9S4
N3	NZPradTrx114954_C01 NZPradTrx086877_C02	Glucosyltransferase	Transfer of glucose (Chen *et al.* 2016)
N4	NZPradTrx088783_C01	Glucose-1-phosphate adenylyltransferase, putative	Involved in the pathway starch biosynthesis (https://www.uniprot.org/uniprot/Q688T8)
N5	NZPradTrx086324_C01	PREDICTED: LOB domain-containing protein 1-like	Involved in the repression of the homeobox gene BP https://www.uniprot.org/uniprot/Q9FKZ3-1
N6	NZPradTrx065580_C01	Catalase	crucial antioxidant enzymes that mitigates oxidative stress to a considerable extent by destroying cellular hydrogen peroxide to produce water and oxygen (Nandi *et al.* 2019)
N7	NZPradTrx049683_C01	Photosystem II core complex proteins psbY2C chloroplast precursor	Photosystem II (PSII) is a multi-component pigment-protein complex that is responsible for water splitting, oxygen evolution, and plastoquinone reduction (Lu 2016)
N8	NZPradTrx097448_C02	ribonucleoprotein, chloroplast, putative	involved in chloroplast RNA processing (Tillich *et al.* 2009)
N9	NZPradTrx119685_C01	SOUL heme-binding protein	plays an active role in primary plant metabolic pathways as well as in stress signalling (Shanmugabalaji *et al.* 2020)
N10	NZPradTrx184701_C01	chloroplast ribosomal protein S1	involvement in translation initiation via positioning of initiation mRNA–protein complexes (mRNPs), and the potential involvement of these unique domains in the processivity of chloroplast translation (Manuell *et al.* 2007)

To assess the overall differences in the needles and the bark based on the top 100 differentially expressed transcripts in each plant part, GO annotation of the genes that were up-regulated in the needles or bark showed quantitative differences in enrichment of all the molecular but not biological or cellular GO categories. In the molecular terms, an overall difference in enrichment was detected in all GO-terms with the catalytic activity being more enriched in both the needle and bark transcripts (Figure 8.3).
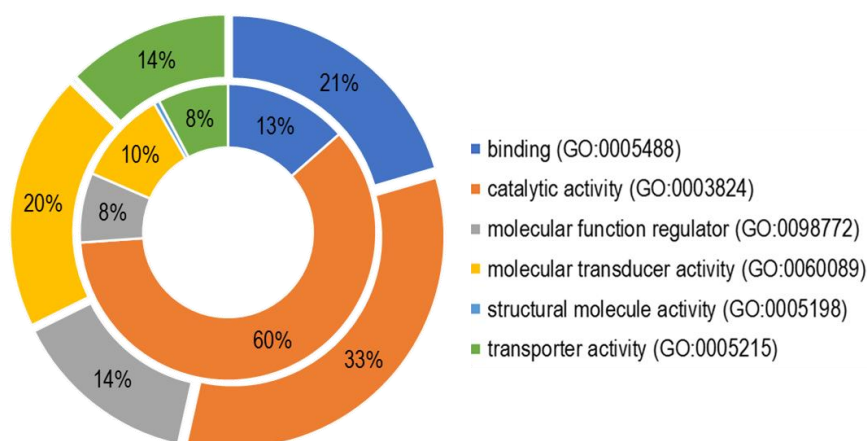
Figure 8.3: The genes that were enriched in the different molecular functions in the needles (inner circle) and the bark (outer circle). The GO enrichment was based on the top 100 genes that were differentially expressed in each plant part.

8.3.4   *Gene expression in the needles and the bark after treatment*

After treatment, considering all time points, more transcripts responded to treatment in the needles than in the bark and more transcripts were up-regulated than down-regulated (Figure 8.4). Most differential expression was detected 7 days (T1) after MJ and strip treatments and decline thereafter, although differential expressed transcripts (DETs) were still evident in both treatments 21 days later (Figure 8.4). Bark stripping did not cause any systemic response in the needles at any time point, suggesting a localised effect. While the main response to MJ treatment was in the needles, DETs in the bark still exceeded the needles following bark stripping. Of the transcripts that were differentially expressed between the bark and needles at T0, only 20% of those expressed in the bark responded and 1% of those expressed on the needles responded following either of the treatments, indicating that the transcripts that did not differ between the needles and the bark at T0 showed stronger responses to treatment. One uncharacterised transcript (NZPradTrx091980_C05) that was not present in the transcriptome of untreated samples was present after treatment. One isoform of ribulose bisphosphate carboxylase preprotein (NZPradTrx098233_C06) that regulates the first step of carbon dioxide fixation during photosynthesis - was present before treatment but was missing in all the samples in the bark and the needles after treatment, including the untreated samples.
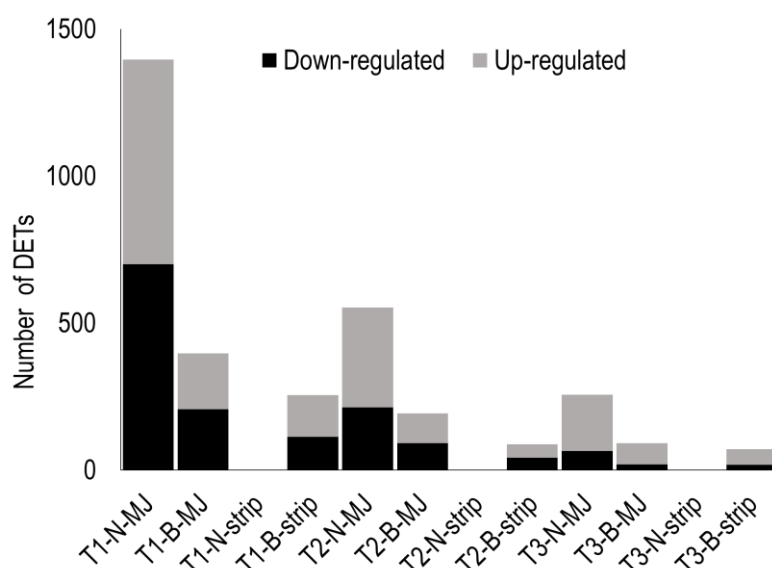
Figure 8.4: Number of differentially expressed transcripts (DETs) in *Pinus radiata* needles (N) and bark (B) following methyl jasmonate (MJ) and bark stripping (strip) treatments quantified 7 (T1), 14 (T2) and 21 (T3) days after treatment. No differential expression was detected in the needles following the strip treatment.

Annotations of the top ten genes that were up-regulated or down-regulated for each condition (time x treatment x part) are presented in Table 8.4. Based on these genes, the following functions were detected, indicating that multiple genes are involved in coordinating plant responses to stress:

1) primary metabolism; involved in synthesis or breakdown and transport of sugars, lipids and aminoacids such as hexokinases, esterase/lipases, glucose-1-phosphate adenylyltransferase and cell wall invertases. The genes involved in photosynthesis and in sugar and fatty acid synthesis were down-regulated. Those involved in breakdown of sugars and fatty acids as well as those involved with amino-acid synthesis were up-regulated.

2) secondary metabolism; involved in biosynthesis of phenolics , alkaloids and terpenes such as chalcone synthases, anthocyanidin reductase , (RS)-norcoclaurine 6-O-methyltransferase-like and lipoxygenase 2. These were mostly up-regulated.

3) digestive inhibitors; these block the normal digestion and absorption of nutrients by vertebrate and invertebrate herbivores (eg chloroplast threonine deaminase 1 precursor) and were up-regulated

4) pathogenesis-related (PR) protein families; includes enzymes that degrade the cell walls of pathogenic fungi for example chitinases, thaumatin proteins, and glucanases. These were also up-regulated.

5) genes involved with physical strengthening of the cell-wall such as proline-rich arabinogalactan protein were up-regulated.

6) transcription factors which are key regulators of the activation or repression of differentially expressed transcripts for example bHLH126-like were up-regulated

7) phytohormones and signalling molecules, including jasmonic, ethylene and abscisic acid signalling such as mitogen activated protein kinase 6, lanC-like protein 2-like, 1-aminocyclopropane-1-carboxylate oxidase 3, abscisic acid (ABA) and leucine-rich repeat-containing protein 40-like were up-regulated

8) general catalysts such as the P450 gene family were down-regulated

9) molecules involved in transcription, such as cleavage and polyadenylation specificity factor subunit 5-like and these were down-regulated

10) molecules involved in broad biotic and abiotic stress responses such as mitogen activated protein kinase and endo-1,3;1,4-beta-D-glucanase-like that were mostly up-regulated

11) broad function genes, for example transporters like lysine histidine transporter 2-like and aquaporin-like protein were up-regulated.

However, it is also recognised that some genes may belong to more than one category

8.3.5 Differences in responses between bark stripping and methyl jasmonate treatments

A subset (15.6%) of the transcriptome was differentially expressed in only one treatment (strip or MJ) (Figure 8.5 top, Table 8.4). Similarly, non-overlapping transcripts were detected at different times in the needles and bark (Figure 8.5 bottom, Table 8.4).

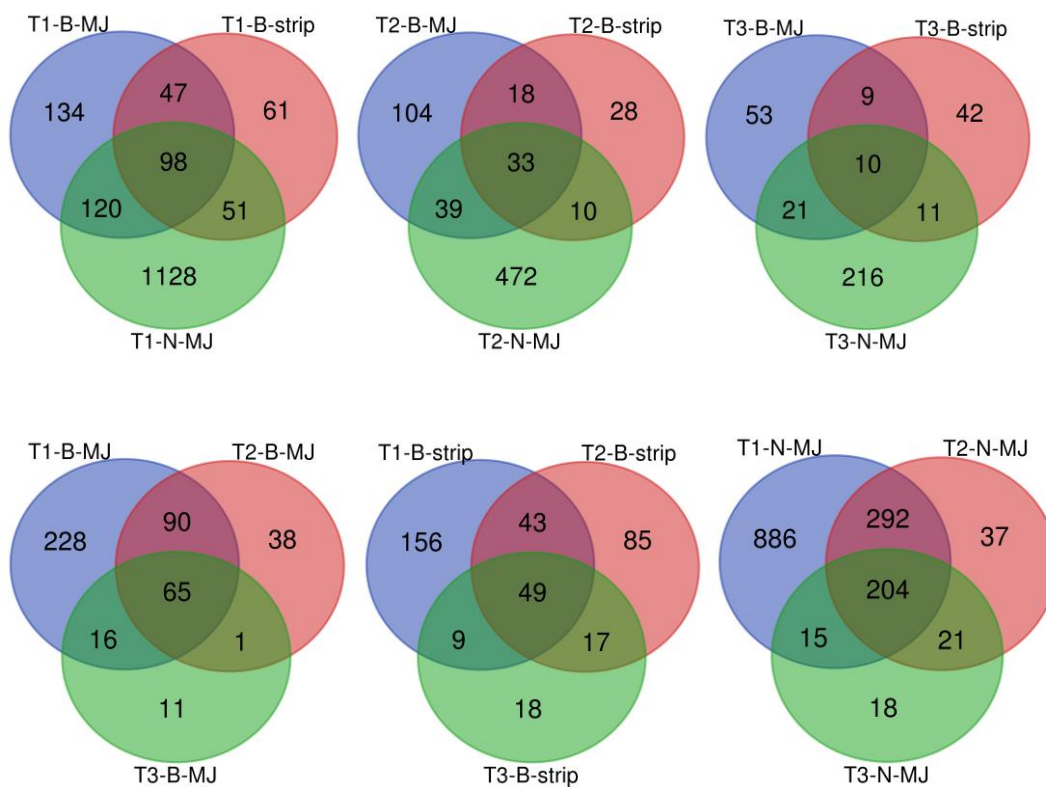**Figure 8.5**: Top: Venn diagrams showing the number of unique transcripts that were differentially expressed for methyl jasmonate (MJ) and bark stripping (strip) at each time; T1, T2 and T3 and for each treatment. Bottom: showing the number of unique transcripts that were differentially expressed across different times within each treatment and part. No differential expression was detected in the needles following bark stripping.

Table 8.4: Top 10 genes differentially expressed in each of the time periods from T1 to T3 in the bark (B) and needles (N) following bark stripping (S) or methyl jasmonate (MJ) treatment of two-year old *Pinus radiata* plants

The SCION transcript code, predicted gene name and predicted functions of the known genes are indicated. Some genes were represented by more than one transcript (isoforms -different *P. radiata* codes that represent one gene in column 1) and multiple copies of an isoform as indicated by the numbers in the parentheses, for example +(2) = two copies of an isoforms relating to the gene were identified, where + = up-regulation, - = down-regulation. The numbers in the parentheses along with the gene names represent the core function of the gene among the 11 broad categories identified in the previous page, for example for the Peptide transporter PTR3-A-like [1], [1] denotes that this gene was associated with primary metabolism (see categories on p.217-218) . However, it is recognised that some genes may fall in more than one category

P. radiata code	Gene name	Function	T1-B-MJ	T1-B-S	T1-N-MJ	T2-B-MJ	T2-B-S	T2-N-MJ	T3-B-MJ	T3-B-S	T3-N-MJ
NZPradTrx081530_C01	Peptide transporter PTR3-A-like [1]	Facilitates amino acid induction (Barnes *et al.* 1998)	+								
NZPradTrx115883_C01	Granule-bound starch synthase, partial [1]	Responsible for amylose synthesis (Miao *et al.* 2014)	-								
NZPradTrx113785_C01	GDP-D-mannose-3',5'-epimerase [1]	Central enzyme of the major ascorbate biosynthesis pathway in higher plants that converts GDP-d-mannose to GDP-l-galactose (Gilbert *et al.* 2009)	+								
NZPradTrx065162_C02	Thiazole biosynthetic enzyme [2]	Thiamine synthesis and DNA damage tolerance (Liu *et al.* 2015b)	-								
NZPradTrx083866_C01	1-aminocyclopropane-1-carboxylate oxidase 3[7]	Production of ethylene, that functions as a mediator of responses to external stimuli, such as wounding (Houben and Van de Poel 2019)	+	+		+	+				
NZPradTrx117447_C01 NZPradTrx117447_C02 NZPradTrx091619_C02	PREDICTED: transcription factor bHLH126-like [6]	Transcription factors play a central role in a number of biological processes, producing, for example, the induction of specific genes in response to particular stimuli as well as controlling the cell type specific or developmentally regulated expression of other genes (Latchman 2008)	+ (2)		+(2)				+(3)	+	
NZPradTrx113021_C04 NZPradTrx117482_C10	Cytochrome P450 CYPC [8]	Key players in plant development and defence (Xu *et al.* 2015)	-(2)			-					
NZPradTrx103647_C02	Oleoyl-acyl carrier protein thioesterase, partial [1]	Plays an essential role in chain termination during de novo fatty acid synthesis https://www.uniprot.org/uniprot/Q42561	-								
NZPradTrx111880_C01 NZPradTrx132560_C01	Cell wall invertase [1]	Mediates reduced export of sucrose or enhanced import of hexoses at the site of infection (Proels and Hückelhoven 2014)	+				+		+ (2)		
NZPradTrx186688_C01 NZPradTrx187077_C01	DNA binding protein, putative [9]	DNA binding proteins serve two principal functions: to organize and compact the chromosomal DNA and to regulate and effect the processes of transcription, DNA replication, and DNA recombination(Travers 2001).				+		-	-		
NZPradTrx065807_C02	PREDICTED: cleavage and polyadenylation specificity factor subunit 5-like [9]	Component of the cleavage factor Im (CFIm) complex that functions as an activator of the pre-mRNA 3'-end cleavage and polyadenylation processing required for the maturation of pre-mRNA into functional mRNAs https://www.uniprot.org/uniprot/Q16630				-					
NZPradTrx095732_C01	Thaumatin-like protein [4]	Involved in local responses of roots to colonization by non-pathogenic plant growth-promoting rhizobacteria (PGPR) fluorescent Pseudomonas spp.(Léon-Kloosterziel *et al.* 2005)	-	-,+ (2)				+	+	+	

ID	Annotation	Function	C1	C2	C3	C4	C5	C6	C7
NZPradTrx064724_C01 NZPradTrx108940_C08 NZPradTrx087317_C02									
NZPradTrx038584_C01	Chloroplast threonine deaminase 1 precursor [3]	Useful in isoleucine (Ile) biosynthesis and impairing digestive processes in the insect gut (Chen et al. 2007)	+	+	+	+		+	+
NZPradTrx111230_C01	Triacylglycerol lipase, putative [1]	Release fatty acids from a number of different substrates (Padham et al. 2007)	-						
NZPradTrx084103_C02	PREDICTED: glutamate--cysteine ligase, chloroplastic-like [4]	Seems to play an important role in controlling the expression of resistance responses like the regulation of salicylic acid (SA) and phytoalexin (camalexin) production. Involved in resistance to fungal and bacterial pathogens. https://www.uniprot.org/uniprot/P46309	+						
NZPradTrx074370_C02, NZPradTrx132647_C01	PREDICTED: lysine histidine transporter 2-like [11]	Amino acid-proton symporter. Transporter with a broad specificity for neutral and acidic amino acids https://www.uniprot.org/uniprot/Q9LRB5	+		+		+		+(2)
NZPradTrx098051_C01	PREDICTED: endo-1,3;1,4-beta-D-glucanase-like [10]	Implicated in responses to stress, wounding, and pathogen infection (Rezzonico et al. 1998)	+						
NZPradTrx053937_C01	2-methyl-6-phytylbenzoquinone methyltranferase [10]	One of the regulators of the composition of tocopherols-class of compounds that function as lipid soluble antioxidants that are extremely potent quenchers of singlet oxygen and free radical species (Shintani et al. 2002)	-						
NZPradTrx119228_C01	4-hydroxyphenyl-pyruvate dioxygenase [1]	Plays an important role in degrading aromatic amino acids (Fritze et al. 2004)		+		+			
NZPradTrx184501_C01	PREDICTED: 50S ribosomal protein L6, chloroplastic-like [9]	Binds directly to 23S ribosomal RNA and is located at the aminoacyl-tRNA binding site of the peptidyltransferase centre. https://www.uniprot.org/uniprot/O23049	-						
NZPradTrx186075_C01	PREDICTED: hexokinase-1-like [1]	Fructose and glucose phosphorylating enzyme https://www.uniprot.org/uniprot/Q42525		-			-		
NZPradTrx105399_C03	PREDICTED: leucine-rich repeat-containing protein 40-like [10]	Plays crucial roles in development and stress responses (Liu et al. 2017b)	-						
NZPradTrx051602_C02	Sodium-bile acid cotransporter, putative [11]	Is involved in photorespiratory metabolism (South et al. 2017)	-						
NZPradTrx082621_C01	Mitogen activated protein kinase 6 [10]	Involved in oxidative stress-mediated signalling cascade (such as ozone) https://www.uniprot.org/uniprot/Q39026		+					
NZPradTrx033779_C01	PREDICTED: pentatricopeptide repeat-containing protein At1g62670, mitochondrial-like [9]	binds one or several organellar transcripts, and influences their expression by altering RNA sequence, turnover, processing, or translation (Barkan and Small 2014)	-						
NZPradTrx184660_C01	PREDICTED: PGR5-like protein 1A, chloroplastic-like [1]	Ferredoxin-plastoquinone reductase involved in cyclic electron flow (CEF) around photosystem I https://www.uniprot.org/uniprot/Q8H112	-						
NZPradTrx097586_C01	Type III chlorophyll a /b-binding protein [1]	functions as a light receptor, it captures and delivers excitation energy to photosystems with which it is closely associated https://www.uniprot.org/uniprot/P27523	-						

ID	Name	Description				
NZPradTrx101698_C02	PrMC3 [2]	Predicted to encode a chalcone-synthase-like protein (Walden *et al.* 1999)	-	-		
NZPradTrx117804_C07	PREDICTED: probable carboxylesterase 2 [1]	Carboxylesterases hydrolyse esters of short-chain fatty acids (Marshall *et al.* 2003)	-			
NZPradTrx100227_C01	PREDICTED: medium-chain-fatty-acid--CoA ligase [1]	Catalyses the esterification, concomitant with transport, of exogenous fatty acids into metabolically active CoA thioesters for subsequent degradation or incorporation into phospholipids https://www.uniprot.org/uniprot/P38135	+			
NZPradTrx081530_C01	PREDICTED: peptide transporter PTR3-A-like [1]	Facilitates amino acid induction (Barnes *et al.* 1998)	+			
NZPradTrx192941_C01	Beta-amylase	The primary function of β-amylase is involvement in starch breakdown in plants (Kaplan and Guy 2004)	+			
NZPradTrx052040_C01	PREDICTED: oleosin 16 kDa-like [10]	May have a structural role to stabilize the lipid body during desiccation of the seed by preventing coalescence of the oil. https://www.uniprot.org/uniprot/Q42980	-			
NZPradTrx108711_C04	PREDICTED: putative UDP-rhamnose:rhamnosyltransferase 1-like [1]	Involved in fatty acid metabolism (van der Sluis and Erasmus 2016)		+		
NZPradTrx112833_C08 NZPradTrx112833_C07	Tify domain containing protein [9]	Found in a variety of plant transcription factors https://pfam.xfam.org/family/PF06200		+	+	+
NZPradTrx071306_C02	PREDICTED: transmembrane ascorbate ferrireductase 1-like [10]	Catalyses ascorbate-dependent trans-membrane ferric-chelate reduction https://www.uniprot.org/uniprot/Q8L856		+		
NZPradTrx051982_C01	PREDICTED: histone H2B.2-like isoform 2 [9]	Histones thereby play a central role in transcription regulation, DNA repair, DNA replication and chromosomal stability https://www.uniprot.org/uniprot/Q5QNW6		-		
NZPradTrx119456_C01	PR10-1.13 [10]	Involved in defence against pathogen infection and other environmental stresses (Liu *et al.* 2005)		+		
NZPradTrx053878_C02 NZPradTrx053878_C01 NZPradTrx053878_C03	Aldehyde dehydrogenase [1]	Involved in plant metabolism and contribute to aldehyde homeostasis to eliminate toxic aldehydes (Zhao *et al.* 2017)			+(3)	+(3)
NZPradTrx087148_C01	PREDICTED: lanC-like protein 2-like [7]	May play a role in abscisic acid (ABA) signalling https://www.uniprot.org/uniprot/F4IEM5			+	
NZPradTrx115807_C06	Hydrolase, putative [10]	Enzyme which catalyses hydrolysis reaction, i.e. the addition of the hydrogen and hydroxyl ions of water to a molecule with its consequent splitting into two or more simpler molecules. https://www.uniprot.org/keywords/KW-0378			+	+
NZPradTrx112951_C03	Embryo-abundant protein [10]	May act as a cytoplasm protectant during desiccation. https://www.uniprot.org/uniprot/P46520			+	

Transcript ID	Annotation	Description				
NZPradTrx097637_C01	PREDICTED: leucoanthocyanidin dioxygenase-like [2]	Involved in anthocyanin and protoanthocyanidin biosynthesis by catalysing the oxidation of leucoanthocyanidins into anthocyanidins https://www.uniprot.org/uniprot/Q96323	+			
NZPradTrx112166_C01	Peroxidase-like protein, partial [10]	Response to oxidative stress https://www.uniprot.org/uniprot/Q24925	+		+	
NZPradTrx082621_C01	Mitogen activated protein kinase 6 [7]	Play key roles in the transduction of environmental and developmental signals through phosphorylation of downstream signalling targets (Jagodzik et al. 2018)	+			
NZPradTrx110107_C07	PREDICTED: transcription factor aborted microspores-like	Required for male fertility and pollen differentiation, especially during the post-meiotic transcriptional regulation of microspore development within the developing anther https://www.uniprot.org/uniprot/Q9ZVX2		+		
NZPradTrx112236_C02	Laccase [2]	Involved in phenolic metabolism and functioning of cell wall (Ranocha et al. 2002)		+		
NZPradTrx089433_C01	Lipoxygenase 2 [2]	Is essential for formation of green leaf volatiles and five-carbon volatiles (Mochizuki et al. 2016)		+		
NZPradTrx109272_C04	Malic enzyme, putative [1]	Catalyse the oxidative decarboxylation of malate to form pyruvate, a reaction important in a number of metabolic pathways (Zhang et al. 2016b)		-	-	
NZPradTrx107808_C01	Putative flavoprotein-containing polyamine oxidase, partial [2]	Involved in drought stress response and flavonoid biosynthesis (Kamada-Nobusada et al. 2008)		+		
NZPradTrx049513_C01 NZPradTrx049513_C02	Putative proline-rich arabinogalactan protein 4 [5]	Contribute to the strengthening of cell walls of quickly growing organs (Hijazi et al. 2014)			+	
NZPradTrx079868_C01	PREDICTED: (RS)-norcoclaurine 6-O-methyltransferase-like [2]	Involved in the biosynthesis of (S)-coclaurine, the common precursor of all benzylisoquinoline alkaloids https://www.uniprot.org/uniprot/Q6WUC1			-	
NZPradTrx054832_C01	Aquaporin-like protein [11]	Main role of aquaporins in plants is transport of water and other small neutral molecules across cellular biological membranes (Kapilan et al. 2018)			+	
NZPradTrx069597_C01	Acetyl-CoA carboxylase BCCP subunit [1]	Catalyses the first committed step of fatty acid synthesis, the carboxylation of acetyl-CoA to malonyl-CoA (Sasaki and Nagano 2004)			-	
NZPradTrx117954_C05	E-alpha-bisabolene synthase [2]	Involved in defensive oleoresin formation in conifers in response to insect attack or other injury. Involved in sesquiterpene (C15) olefins biosynthesis https://www.uniprot.org/uniprot/O81086			+	
NZPradTrx087252_C01	TPA: putative GID1-like gibberellin receptor [7]	Involved in gibberellin signalling (Sun 2011)				+
NZPradTrx074370_C01 NZPradTrx112166_C01	Putative proline transporter [11]	Mediates the amino acid proline and glycine betaine transport https://www.uniprot.org/uniprot/P92961				+(2)
NZPradTrx113904_C06/ NZPradTrx101343_C01	PREDICTED: clavaminate synthase-like protein At3g21360-like [10]	Associated with metal ion binding and oxido-reductase activity https://www.uniprot.org/uniprot/Q9LIG0				+

Annotations of the unique top differentially expressed transcripts in each category shown in the Venn diagram indicated that methyl jasmonate caused differential expression of more genes that are directly involved in the metabolism of sugars, fatty acids and amino acids in the bark and the needles (Table 8.5). The strip induced transcriptome had among the top genes those involved in defence against pathogens, such as chitinases and PR10 and defensins. Bark stripping also caused water-stress responsive genes as well as genes related to replacement of tissues (Table 8.5). The difference in the representation of genes is likely related to the kind of damage incurred by the two stressors. Both stressors caused an expression of genes related to secondary metabolism, including metabolism of monoterpenes (e.g. geranyl diphosphate synthase), phenolics (e.g. laccases) and alkaloids (e.g. phenylalanine ammonia-lyase). Genes associated with lignification of cell walls were also identified for both treatments in the needles and the bark emphasising the role of cell wall physical properties in stress responses. For some genes, the same gene was represented by different isomorphs in the different conditions such as geranyl diphosphate synthase in the B-strip and N-MJ.

Some genes were consistently differentially expressed from T1-T3 (Figure 8.5 bottom). Across all the treatments and parts, this involved only 6 genes that were all up-regulated (Table 8.5). Annotations of these transcripts mostly showed genes related to amino acid synthesis, suggesting that amino acids are integral parts of the plant stress responses. Specific treatments also showed some genes that consistently responded to treatment from T1-T3. In the methyl jasmonate-induced transcriptome of the bark (B-MJ), 6 transcripts were involved and were mostly up-regulated. Annotations of these transcripts showed that the genes were mostly involved in generating energy from various substrates, particularly glucose and fatty acids. In the needles treated with methyl jasmonate (N-MJ), 114 transcripts were consistently expressed from T1-T3. Among the top 10 differentially expressed genes were those directly associated with chemical and physical structures, for example 4hydroxyphenylpyruvate dioxygenase involved in phenolic biosynthesis and the hydroxyproline-rich glycoproteins that are important structural components of the cell wall. The others included growth regulating genes (Table 8.5). However, among the 144 transcripts, defence related genes were also detected, and these included; class I chitinase and 1,8-cineole synthase, protein transparent testa involved in the accumulation of condensed tannins and acyl carrier protein involved in cell death execution or direct antimicrobial activity. These genes were mostly up-regulated. In the strip-induced transcriptome of the bark (B-strip), only the DNA binding protein was up-regulated consistently from T1 to T3.

Table 8.5: Number of transcripts that were uniquely expressed for each condition (time x treatment x plant part) and the five dominant genes in each category. These transcripts were not expressed at any other time or treatment. (+) =up-regulated and (-) =down-regulated

Condition	Unique transcripts	*P. radiata* code	Gene name	Predicted gene function	direction
T1-B-MJ	96	NZPradTrx115883_C02	granule bound starch synthase 1a precursor	Involved in the pathway starch biosynthesis https://www.uniprot.org/uniprot/P0C585	-
		NZPradTrx184661_C01	PREDICTED: putative caffeoyl-CoA O-methyltransferase At1g67980-like	Involved in the reinforcement of the plant cell wall. Also involved in the responding to wounding or pathogen challenge by the increased formation of cell wall-bound ferulic acid polymers https://www.uniprot.org/uniprot/Q9C9W3	-
		NZPradTrx108036_C04	Cytochrome b reductase	Required for the NADH-dependent electron transfer involved in the desaturation and hydroxylation of fatty acids and in the desaturation of sterol precursors https://www.uniprot.org/uniprot/Q9ZNT1	-
		NZPradTrx119186_C01	DEAD-box RNA helicase	Ubiquitous in RNA-mediated processes and function by coupling cycles of ATP binding and hydrolysis to changes in affinity for single-stranded RNA https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3032546/	-
		NZPradTrx060156_C02	PREDICTED: probable rhamnose biosynthetic enzyme 1	Involved with nucleotide-sugar metabolic process https://www.uniprot.org/uniprot/A0A1U7W8H4	+
		NZPradTrx119948_C01	PREDICTED: protein HOTHEAD-like	Required to limit cellular interactions between contacting epidermal cells during floral development (Krolikowski *et al.* 2003)	+
		NZPradTrx119070_C01	PREDICTED: mitochondrial-processing peptidase subunit alpha-like	Substrate recognition and binding subunit of the essential mitochondrial processing protease (MPP), which cleaves the mitochondrial sequence off newly imported precursors proteins. https://www.uniprot.org/uniprot/P29677	+
		NZPradTrx110606_C03 NZPradTrx110606_C04	snakin	Active against fungal and bacterial plant pathogens (Berrocal-Lobo *et al.* 2002)	-
		NZPradTrx094750_C01	PREDICTED: zinc finger CCCH domain-containing protein 20-like	Known to play important roles in RNA processing as RNA-binding proteins in animals (Wang *et al.* 2008)	-
		NZPradTrx119288_C01	PREDICTED: blue copper protein	Redox proteins whose role is to shuttle electrons from an electron donor to an electron acceptor (De Rienzo *et al.* 2000)	-
T1-B-strip	39	NZPradTrx111276_C02	low molecular weight heat-shock protein	Expressed in plants experiencing high-temperature stress (Hernandez and Vierling 1993)	-
		NZPradTrx109179_C02 NZPradTrx077717_C01	LP3-1	Shown to be up-regulated in response to water deficit stress and to also act as transcription factors for genes likely involved in hexose transport (Lecoy and García-Gil 2020)	-
		NZPradTrx112152_C04	PREDICTED: L-type lectin-domain containing receptor kinase IV.1-like	Involved in resistance response to the pathogenic oomycetes, Promotes hydrogen peroxide H_2O_2 production and cell death https://www.uniprot.org/uniprot/Q9LXA5	+
		NZPradTrx082734_C01	Casparian strip domain-like gene	Recruit the lignin polymerisation machinery necessary for the deposition of Casparian strips in the endodermis https://www.ebi.ac.uk/interpro/entry/InterPro/IPR006459/	-

		NZPradTrx105759_C05	Methyl esterase 13	Involved in jasmonic and salicylic acid metabolic process https://www.uniprot.org/uniprot/F4IE65	+
		NZPradTrx042090_C01	Geranyl diphosphate synthase	Catalyses the condensation of dimethylallyl diphosphate and isopentenyl diphosphate to geranyl diphosphate, the key precursor of monoterpene biosynthesis (Burke *et al.* 1999)	+
		NZPradTrx064702_C01	Class II chitinase	Involved in the defence response against pathogen and fungal infection (de A. Gerhardt *et al.* 1997)	-
		NZPradTrx105720_C01	Defensin	Inhibit the growth of a broad range of pathogens, including bacteria, fungi and viruses (Ermakova *et al.* 2016a; Picart *et al.* 2012).	-
		NZPradTrx119059_C01	Annexin p33	Central regulators or effectors of plant growth and stress signalling (Mortimer *et al.* 2008)	-
		NZPradTrx118949_C01	Peroxiredoxin	Guardians against oxidative stress and modulators of peroxide signalling (Perkins *et al.* 2015)	-
T1-N-MJ	751	NZPradTrx110565_C01	UDP-sulfoquinovose synthase	Involved in the biosynthesis of sulfolipids found in thylakoid membranes. Converts UDP-glucose and sulfite to the sulfolipid head group precursor UDP-sulfoquinovose https://www.uniprot.org/uniprot/O48917	-
		NZPradTrx064995_C02	Chloroplast ribulose bisphosphate carboxylase/oxygenase activase alpha1, partial	Catalyses carboxylation of RuBP in the first step of the Calvin cycle of photosynthesis (Tabita 1999)	-
		NZPradTrx088104_C02	RNA polymerase sigma factor rpoD, putative	Sigma factors are initiation factors that promote the attachment of RNA polymerase to specific initiation sites and are then released https://www.uniprot.org/uniprot/P00579	-
		NZPradTrx081803_C01	PREDICTED: mitochondrial carnitine/acylcarnitine carrier-like protein-like	Mediates the transport of acylcarnitines of different length across the mitochondrial inner membrane from the cytosol to the mitochondrial matrix for their oxidation by the mitochondrial fatty acid-oxidation pathway https://www.uniprot.org/uniprot/O43772	-
		NZPradTrx086144_C02	Chloroplast omega-6 fatty acid desaturase	Chloroplast omega-6 fatty acid desaturase introduces the second double bond in the biosynthesis of 16:3 and 18:3 fatty acids, important constituents of plant membranes. It is thought to use ferredoxin as an electron donor and to act on fatty acids esterified to galactolipids, sulfolipids and phosphatidylglycerol https://www.uniprot.org/uniprot/P46312	-
		NZPradTrx065194_C01	Glutamate--ammonia ligase	Key enzyme of ammonium assimilation and recycling in plants where it catalyses the synthesis of glutamine from ammonium and glutamate(Lothier *et al.* 2011)	-
		NZPradTrx077590_C01	PREDICTED: ATP synthase gamma chain, chloroplastic-like	Produces ATP from ADP in the presence of a proton gradient across the membrane. The gamma chain is believed to be important in regulating ATPase activity and the flow of protons through the CF_0 complex https://www.uniprot.org/uniprot/Q01908	-
		NZPradTrx064646_C01	PREDICTED: photosystem I reaction center subunit XI, chloroplastic-like	Involved in photosynthesis https://www.uniprot.org/uniprot/Q41385	-

		NZPradTrx115121_C05	glutathione peroxidase-like protein, partial	Protects cells from phospholipid hydroperoxides and nonphospholipid peroxides during oxidative stress https://www.uniprot.org/uniprot/P36014	+
		NZPradTrx186664_C01	F353614_1 senescence-associated protein SPA15	May be involved in the regulation of leaf senescence https://www.uniprot.org/uniprot/Q65XF2	-
T2-B-MJ	18	NZPradTrx192941_C01	Beta-amylase	The primary function of β-amylase is involvement in starch breakdown in plants (Kaplan and Guy 2004)	+
		NZPradTrx076831_C01	UV-B receptor 1	Involved in response to UV-B (Loyola *et al.* 2016)	+
		NZPradTrx044917_C01	Putative cyclophilin	Involved in various physiological processes including transcriptional regulation, organogenesis, photosynthetic and hormone signalling pathways, stress adaptation and defence responses (Barbosa dos Santos and Park 2019)	-
		NZPradTrx119079_C01	Xyloglucan endotransglucosylase/hydrolase 13	Cleaves and religates xyloglucan polymers, an essential constituent of the primary cell wall, and thereby participates in cell wall construction of growing tissues https://www.uniprot.org/uniprot/Q9FKL8	-
		NZPradTrx037564_C01	PREDICTED: bidirectional sugar transporter SWEET3-like	Mediates both low-affinity uptake and efflux of sugar across the plasma membrane https://www.uniprot.org/uniprot/Q6NQN5	-
		NZPradTrx118938_C01	Glycine-rich RNA-binding protein	Plays a role in RNA transcription or processing during stress. Binds RNAs and DNAs sequence with a preference to single-stranded nucleic acids. https://www.uniprot.org/uniprot/Q03250	-
		NZPradTrx109658_C01	Probable aquaporin	Main role of aquaporins in plants is transport of water and other small neutral molecules across cellular biological membranes (Kapilan *et al.* 2018)	-
		NZPradTrx094541_C02	PREDICTED: methionine gamma-lyase-like	Involved in amino acid catabolism (Ravanel *et al.* 1998)	+
T2-B-strip	12	NZPradTrx119456_C01	PR10-1.13	Involved in defence against pathogen infection and other environmental stresses (Liu *et al.* 2005)	+
		NZPradTrx098320_C05	PREDICTED: LOB domain-containing protein 1-like	Controls the proximal-distal patterning in petals and the adaxial-abaxial determination of leaves. Involved in the repression of the homeobox gene BP https://www.uniprot.org/uniprot/Q9FKZ3-1	+
		NZPradTrx073494_C01	TPA: putative ARF GTPase-activating domain family protein	Have potential roles in cell migration, central to normal physiology in embryogenesis, the inflammatory response and wound healing (Campa and Randazzo 2008)	-
		NZPradTrx103835_C01	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	Involved in the biosynthesis of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), two major building blocks of terpenoid compounds https://www.uniprot.org/uniprot/P62617	+
		NZPradTrx102746_C02	S6 ribosomal protein	Key downstream effector of the target of rapamycin (TOR) kinase pathway that regulates various biological processes, including translation, synthesis of ribosomal proteins, and transcription of rRNA (Kim *et al.* 2014)	+
		NZPradTrx094959_C01	Pathogenesis-related protein 10	Function in a wide range from cell wall rigidification to signal transduction and antimicrobial activity (Liu and Ekramoddoullah 2006)	+

		NZPradTrx096309_C03	Dirigent-like protein	Predominant roles in defence responses, secondary metabolism, and fiber biosynthesis (Li *et al.* 2017)	+
		NZPradTrx105315_C01	PREDICTED: uncharacterized LOC101219508		-
		NZPradTrx077043_C01	FAD/NAD(P)-binding oxidoreductase domain-containing protein	Catalyse a wide variety of redox reactions with many different substrates (Sellés Vidal *et al.* 2018)	-
		NZPradTrx110593_C01	PREDICTED: chaperonin CPN60-2, mitochondrial-like	Implicated in mitochondrial protein import and macromolecular assembly. May facilitate the correct folding of imported proteins. May also prevent misfolding and promote the refolding and proper assembly of unfolded polypeptides generated under stress conditions in the mitochondrial matrix. https://www.uniprot.org/uniprot/Q05046	+
T2-N-MJ	30	NZPradTrx118421_C03	Caffeic acid O-methyltransferase	Catalyses the conversion of caffeic acid to ferulic acid and of 5-hydroxyferulic acid to sinapic acid. The resulting products may subsequently be converted to the corresponding alcohols that are incorporated into lignins https://www.uniprot.org/uniprot/Q06509	+
		NZPradTrx079649_C05 NZPradTrx079649_C03 NZPradTrx079649_C02	Geranyl diphosphate synthase	Catalyses the condensation of dimethylallyl diphosphate and isopentenyl diphosphate to geranyl diphosphate, the key precursor of monoterpene biosynthesis (Burke *et al.* 1999)	+
		NZPradTrx122822_C01	PREDICTED: F-box protein GID2-like	Essential component of the SCF-type E3 ligase complex, SCF(GID2), a complex that positively regulates the gibberellin signalling pathway https://www.uniprot.org/uniprot/Q9STX3	+
		NZPradTrx083848_C01	Chlorophyllase	The first enzyme involved in chlorophyll (Chl) degradation and catalyses the hydrolysis of ester bond to yield chlorophyllide and phytol (Tsuchiya *et al.* 1999)	+
		NZPradTrx103321_C01	Phenylalanine ammonia-lyase	Phenylalanine aminomutase that catalyses the rearrangement of L-phenylalanine to R-beta-phenylalanine. Catalyses the first committed step in the biosynthesis of the side chain of the alkaloid taxol (paclitaxel) https://www.uniprot.org/uniprot/Q68G84	+
		NZPradTrx071573_C01	Starch synthase isoform II	Contribute to the extension of glucan chains in the synthesis of starch (Edwards *et al.* 1999)	+
		NZPradTrx105898_C01	Glutamate-1-semialdehyde 2,1-aminomutase	Essential enzyme in the pathway that leads to the synthesis of chlorophyll and other tetrapyrroles in plants and some bacteria (Tyacke *et al.* 1995)	-
		NZPradTrx182827_C01	PREDICTED: LRR receptor-like serine/threonine-protein kinase FLS2-like	Constitutes the pattern-recognition receptor (PPR) that determines the specific perception of flagellin (flg22), a potent elicitor of the defence response to pathogen-associated molecular patterns (PAMPs) https://www.uniprot.org/uniprot/Q9FL28	+
		NZPradTrx184681_C01	FK506 binding-like protein	Involved in diverse cellular functions including protein folding, cellular signalling, apoptosis and transcription (Tong and Jiang 2016)	+
		NZPradTrx094486_C01	Putative UDP-glucose:flavonoid glucosyltransferase	Enhances the solubility of anthocyanins (Chen *et al.* 2011)	+

T3-B-MJ	4	NZPradTrx083714_C01	PREDICTED: protein GLUTAMINE DUMPER 1-like	Involved in the regulation of the amino acid metabolism. involved in the salicylic acid (SA) pathway and in the geminivirus-host interaction https://www.uniprot.org/uniprot/O81775	+
		NZPradTrx053990_C01	PREDICTED: cytochrome P450 71A1-like	Involved in the metabolism of compounds associated with the development of flavour in the ripening fruit process, possibly by acting as trans-cinnamic acid 4-hydrolase https://www.uniprot.org/uniprot/P24465	+
		NZPradTrx105443_C01	GMP synthase [glutamine-hydrolyzing] subunit A, putative	Involved in the de novo biosynthesis of guanosine nucleotides https://www.brenda-enzymes.org/enzyme.php?ecno=6.3.5.2	+
		NZPradTrx112236_C01	Laccase	Involved in phenolic metabolism and functioning of cell wall (Ranocha *et al.* 2002)	+
T3-B-S	13	NZPradTrx087634_C02	Properoxidase	Involved in lignification, cell elongation, stress defence and seed germination (Shigeto and Tsutsumi 2016)	+
		NZPradTrx103699_C01	Oxidoreductase, 2OG-Fe(II) oxygenase family protein	Involved in defence against pathogens (Van Damme *et al.* 2008)	+

The time progression of individual genes was variable between methyl jasmonate and strip induced genes. For purposes of illustrating the differences in temporal expression of individual genes, Figure 8.6 below shows the time progression of the top 10 genes that were differentially expressed in the bark compared to the needles at T0 (listed in Table 8.2) and how they responded to methyl jasmonate and strip treatments. The genes were up-regulated or down-regulated following both treatments. Although results show stronger quantitative responses at T1 for most up-regulated and down-regulated genes, the direction of responses for some genes differed between methyl jasmonate and strip treatments. Figure 8.6 shows chalcone synthase involved in phenolic metabolism was up-regulated in the strip treatment but down-regulated after methyl jasmonate treatment. Metallothionein was consistently down-regulated in strip-treated transcriptome but was up-regulated at T3 in methyl jasmonate induced transcriptome.
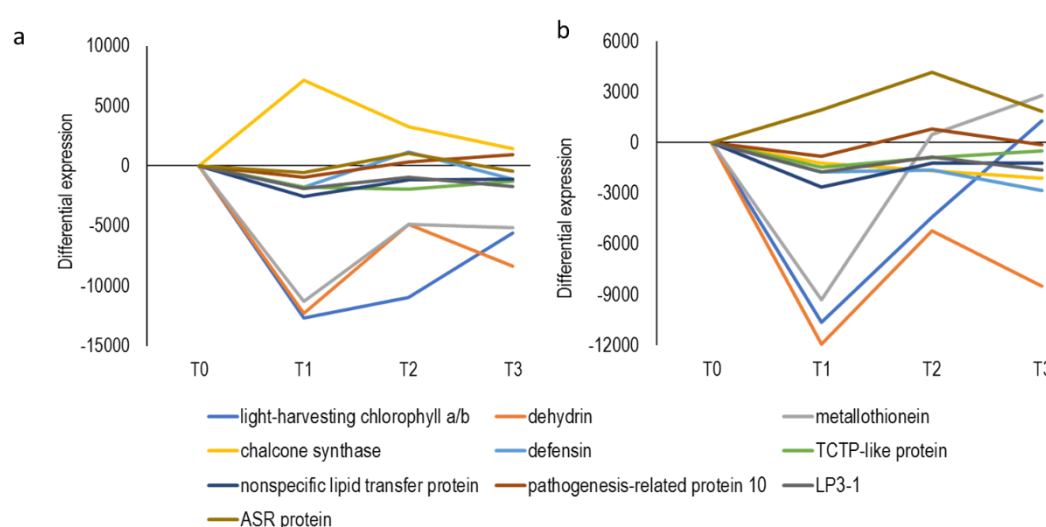
Figure 8.6: Time progression of the top 10 genes that were differentially expressed in the bark relative to the needles at T0 following, a) strip treatment and b) methyl jasmonate treatment. Average change in expression was estimated at each time point by comparing the raw counts for the strip or methyl jasmonate induced transcripts and the transcripts from control treatments (mean of treatment – mean of control) for a specific time and were adjusted according to the basal expression at T0. T0 -before treatment applications, then, T1, T2 and T3 are respectively 7, 14 and 21 days after treatment application.

8.3.7 *Functional annotation and gene ontology (GO) classification of differentially expressed transcripts*

To assess the overall effect of the treatments across different gene families and molecular processes, an enrichment analysis of GO terms was performed for the up-regulated and down-regulated transcripts for each condition (time x treatment x plant part). An overall similarity was detected of the major GO terms for genes that were up- and down-regulated in both the strip and methyl jasmonate treatments.

235

For example, in the GO-molecular function (Figure 8.7), most genes relating to catalytic activity were enriched both in the needles and the bark. However, the enrichment was stronger for the genes that were up-regulated in the methyl jamonate-expressed transcripts, while in the strip-expressed transcripts enrichment was stronger in the down-regulated genes.
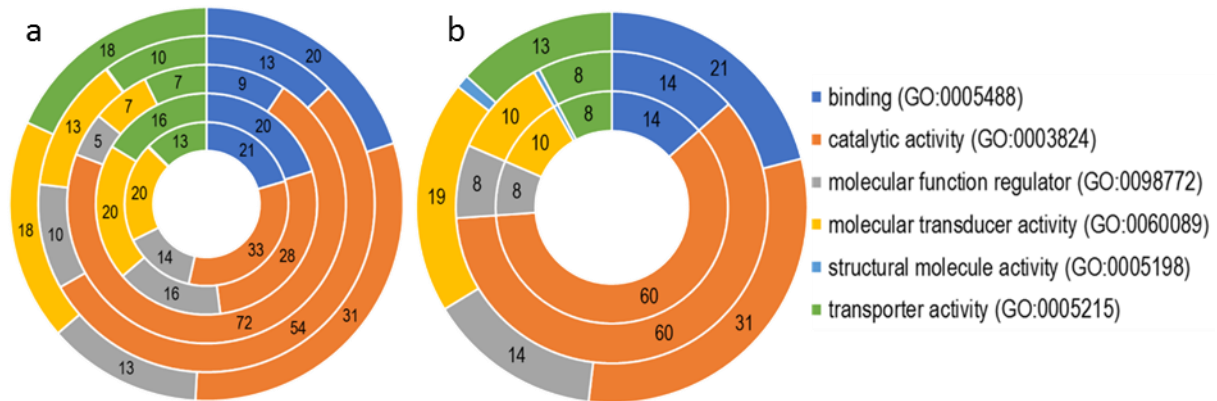

Figure 8.7: GO-molecular categories showing differences in the treatments, and plant parts, as well as differences in enrichment of the different GO terms for the up-regulated and down-regulated transcripts for each condition at T1. The top 100 differentially expressed transcripts were selected in each condition for this analysis. The categories in order from the centre are; a) transcripts in the bark up-regulated relative to the needles at T0, up-regulated in T1-B-strip, down-regulated in T1-B-strip, up-regulated in T1-B-MJ and down-regulated in T1-B-MJ, b) transcripts up-regulated in the needles relative to the bark at T0, up-regulated in T1-N-MJ and down-regulated in T1-N-MJ). No differential expression was detected in the needles of bark stripped trees at any time point.

Comparing GO terms for the constitutive and induced transcriptome, results indicated that some gene functions that were not strongly enriched at T0 with respect to plant parts were enriched after treatment, for example, genes related to response to stimulus (GO:0050896), plasmodesma (GO:0009506) and cell junction (GO:0030054) were strongly enriched in the transcriptome of the bark stripped samples (Figure 8.8).
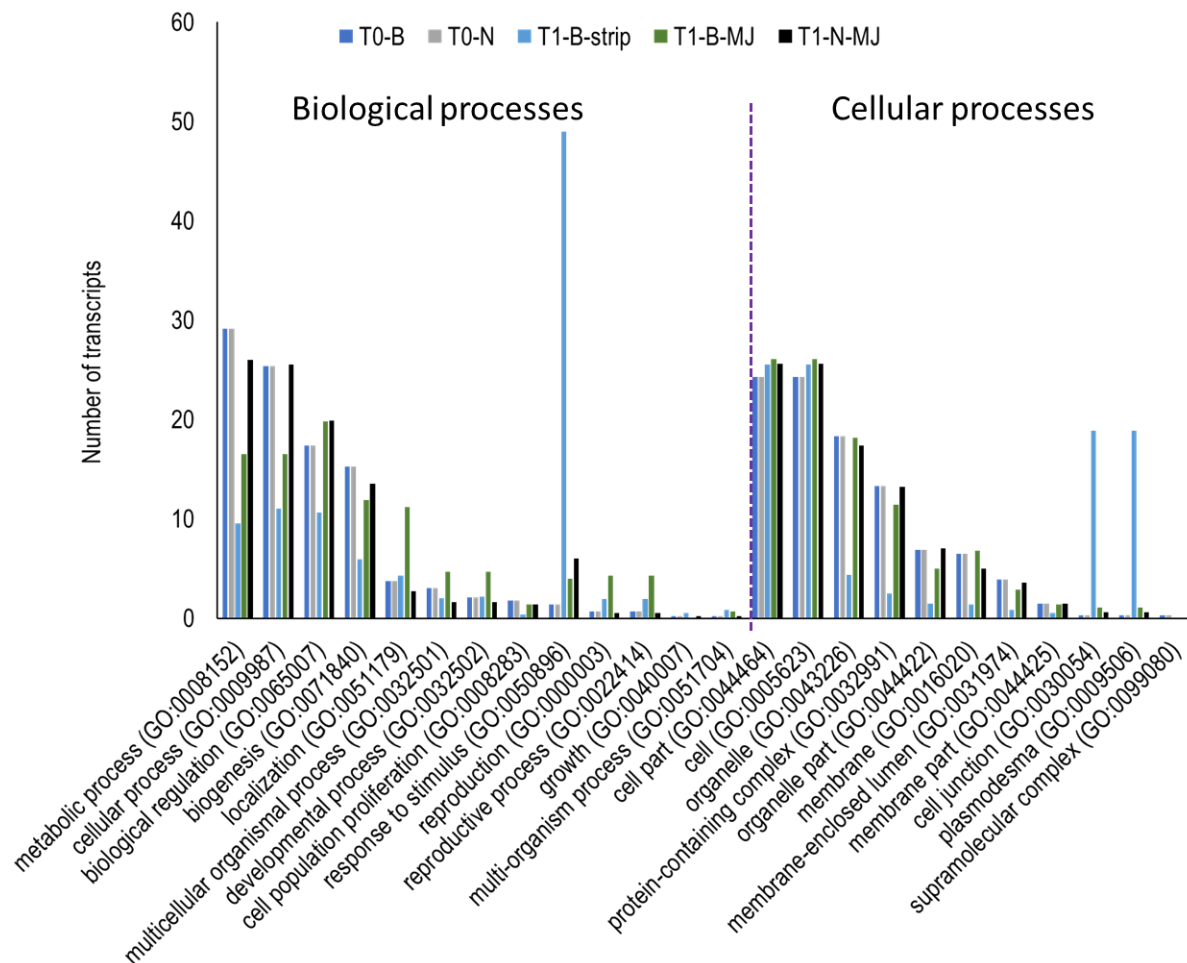
Figure 8.8: Number of transcripts enriched in each biological and cellular categorization of up-regulated and down-regulated genes in *Pinus radiata* needles (N) and bark (B) between plant parts at T0 and after treatment with methyl jasmonate (MJ) or bark stripping (strip) at T1. The categorization is based on gene ontology (GO) annotations of the top 100 differentially expressed transcripts in each category. No differences were detected in the biological and cellular processes of up-regulated or down-regulated genes.

8.4 Discussion

We aimed to understand the differences in the constitutive needle and bark transcriptomes and the changes that occur following bark stripping and how they compare with those of methyl jasmonate that have been documented in several conifer species (Celedon *et al.* 2017; Kānberga-Siliņa *et al.* 2017; Liu *et al.* 2017a; Men *et al.* 2013). While the results are based on a partial transcriptome, comparing the needle and bark transcriptome as assessed prior to treatment (T0) showed that, there were minimal qualitative differences in the transcriptomes. However, after treatment there was strong transcriptional response in both the needles and the bark following both methyl jasmonate and strip treatments, with different and sometimes non-overlapping in responses between plant parts, treatments and at each time of sampling. Differences in responsiveness were also detected between the classes of genes,

where genes related to primary metabolism were more responsive to treatments by up-regulation or down-regulation compared to genes associated with secondary metabolism. This is the first study to illustrate transcriptional responses to bark stripping.

Among the genes that were homogeneously expressed between the bark and the needles were those related to basic life fuctions especially those related to primary and secondary metabolism. Genes for example, ribulose bisphosphate carboxylase/oxygenase (RuBisCO) and a chlorophyll a/b binding protein were dominant both in the transcriptome of the needles and the bark. Similar observations were made in the needles of other *P. radiata* populations (Alvarez *et al.* 2016) and *Pinus monticola* (Liu *et al.* 2013a), although these studies did not analyse how the transcriptomes change with treatment and the observations were limited to one plant part. Genes directly related to secondary metabolism for example chalcone synthases, dehydrins and defensins were among the basal genes, highlighting the importance of constitutive defences in *P. radiata*. Chalcone synthase has been identified in other conifers (Baker and White 1996; Richard *et al.* 2000) and plays crucial roles in phenolic biosynthesis (Dixon and Paiva 1995). Defensins have also been detected in various conifers where they inhibit the growth of a broad range of pathogens, including bacteria, fungi and viruses (Ermakova *et al.* 2016b; Picart *et al.* 2012). Dehydrins that represent a family of genes for drought tolerance have been detected in spruces and in other *Pinaceae* (Stival Sena *et al.* 2018). Metallothioneins that were strongly expressed both in the bark and the needles are important in protection against heavy metal toxicity (Nevrtalova *et al.* 2014) and have been documented mainly in *Pseudotsuga menziesii* (Chatthai *et al.* 1997; Chatthai *et al.* 2004) and could reflect an adaptation to leached, heavy metal enriched soils in the coastal sites of California where *Pinus radiata* originates (Keator 2002). However, while the above genes are expressed at high amounts equally in the bark and needles, some transcripts were up-regulated in the needles or the bark. More up-regulation was detected in the bark, which contrasted with the higher expression of transcripts in the needles than the bark reported in other *P. radiata* populations (Alvarez *et al.* 2016). In both parts up-regulated genes were predominantly related to the synthesis and transfer of macro- and micro-molecules, as well as transcription factors which are the key molecular switches orchestrating the regulation of plant responses to a variety of stresses.

After treatment with methyl jasmonate and bark stripping, there was an up-regulation and down-regulation of several genes involved in both primary and secondary metabolism both in the bark and needles, consistent with other studies that have characterised responses to other stressors (Kānberga-Siliņa *et al.* 2017; Liu *et al.* 2017c). The top genes that were up-regulated or down-regulated in the present study overlap with those observed in similar studies that have characterised the transcriptomes

with contrasting sources of stress in conifers (Celedon *et al.* 2017; Fox *et al.* 2017; Liu *et al.* 2013a; Liu *et al.* 2017c; Ralph *et al.* 2006), suggesting that changes in gene expression following stress are relatively conserved. Among the top expressed genes, results showed a down-regulation of hexokinases, granule-bound starch synthase and sodium-bile acid cotransporter as well as genes related with photosynthesis, suggesting reduction in sugar metabolism in the treated plants. However, cell wall invertase that mediates export of sucrose or enhanced import of hexoses at the site of infection was up-regulated in both methyl jasmonate and strip treated plants. Cell wall invertase (CWI) is an enzyme that cleaves sucrose, the major transport sugar in plants, irreversibly yielding glucose and fructose, which can be taken up by plant cells (Proels and Hückelhoven 2014; Tauzin and Giardina 2014). An increase in CWI should ideally lead to a reduction in sucrose, which is consistent with the drastic reduction in the amounts of sucrose that has been observed following methyl jasmonate and strip treatments in *P. radiata* (Chapters 3 and 6). The up-regulation of CWI would also suggest an increase of glucose and fructose, but this was not the case as a strong reduction in the amounts of glucose and fructose was observed in treated samples (Chapter 3). This suggests that although fructose and glucose may be potentially enhanced by an increased break down of sucrose, their utilisation for energy and carbon skeletons for other organic compounds or for tissue recovery exceeds their production, supporting the concept that defence is costly in terms of energy (Gershenzon 1994). Gould *et al.* (2008) detected a repression of photosynthesis in *P. radiata* as a response to stress that could lead to a reduction of sugars. However, the repression of sugars in the needles of bark stripped trees (Chapter 3) without a consequent change in gene expression shows that reduction of sugars following defence activation does not always correlate with changes in protein profiles, indicating that mechanisms other than transcriptional level changes are involved in reconfiguration of the sugars. Sugars have also been shown to function as signalling molecules, in a manner similar to hormones (Tauzin and Giardina 2014; Trouvelot *et al.* 2014), but their down-regulation contrasts to the up-regulation of other signalling molecules. However, according to Eveland and Jackson (2012) sugar signals are generated either by relative ratios to other metabolites, such as C:N not necessarily carbohydrate concentration.

Additionally, genes related to fatty acid metabolism, for example medium-chain-fatty-acid--CoA ligase and UDP-rhamnose:rhamnosyltransferase, were up-regulated and those related to fatty acid hydrolysis such as carboxylesterase were down-regulated. Observations on the same population showed a reduction in fatty acids following treatment (Chapter 3), suggesting their potential use as precursors to the formation of secondary compounds (Kachroo and Kachroo 2009). Accumulating evidence has suggested lipids and lipid metabolites as important regulators of plant defence (Shah 2005). Genes

related to amino acid synthesis were also among the top expressed genes. Increase in amino acid levels have been detected in plants under stress and is hypothesized to protect plant cells against dehydration (Al-Asbahi *et al.* 2012; Joshi and Jander 2009). Amino acid accumulation has been observed to be strongly related to abscisic acid signalling (Al-Asbahi *et al.* 2012). Molecules related to abscisic acid signalling were also strongly up-regulated.

Genes related directly to secondary metabolism were not detected among the top differentially expressed genes although they are abundant in the constitutive transcriptome consistent with the observations in spruce (Verne *et al.* 2011). However, the relatively weak transcriptional response to treatment of individual genes related to secondary metabolism in this study contrasts with other studies (Liu *et al.* 2017a; Ralph *et al.* 2006) and could possibly be due to the timing of the sampling, which was done 7 days after treatment application. In various studies maximum expression of genes is shown to be attained within 5 days after treatment application (Liu *et al.* 2017a; Ralph *et al.* 2006). On the same population, a weak response of terpenes and phenolics was observed following similar treatments (Chapters 3 and 6), which probably suggests an inherently weak response of secondary compounds and associated genes to stress in *P. radiata*. Defence genes being strongly expressed in the constitutive but not in the induced transcriptome may suggest existence of trade-offs in induced gene expression (Kim *et al.* 2020), analogous to the trade-offs in constitutive versus induced chemical responses that have been detected in *P. radiata* (Moreira *et al.* 2014). Although alkaloids have not been well researched as important defence compounds in conifers, genes related to alkaloid biosynthesis such as RS-norcoclaurine 6-O-methyltransferase were among the top expressed genes but was down-regulated after treatment. There were also many proteins of unknown functions that were up-regulated or down-regulated at various time points, which potentially explains the many unknown chemical compounds that were quantified on the same plants.

Considerable overlap was observed between the methyl jasmonate and the strip induced transcriptome. However, results also indicate that bark stripping can induce transrips that are not induced with methyl jasmonate and *vice versa*. Defence responses for bark stripping may differ from methyl jasmonate since bark stripping causes tissue and water loss at the injured sites, and damaged plants are also easily infected by pathogens through these wounds. In this case both defence and repair responses are required. Hence the dominant genes in the strip-induced transcriptome involved pathogenesis-related (PR) genes and those related to fibre synthesis. The expression of PR genes could also be related to the historical relationship between *P. radiata* and various pathogens (Offord 1964). No systemic transcript responses were observed in the needles to bark stripping. Coupled with

the chemical changes that were observed in the needles following bark stripping on the same population for example the reduction of glucose and fructose at T1 and T2, (Chapter 3), this observation suggests that some chemical stress responses, possibly those involving sugars may not involve on-site gene expression changes and may result from passive reallocation of chemistry within the plant. For other compounds like terpenes, it has been indicated that passive changes normally occur only in the constitutive environment and that stress-induced changes in terpenes are entirely of a *de novo* nature (Wu *et al.* 2017).

Relative to the chemical changes that were observed on the same population, whereas maximum expression of the transcripts was observed at T1 (7 days after treatment), more chemical changes were detected at T2 (14 days) and T3 (21 days) suggesting a time-lag between gene and phenotypic expression. This discrepancy may be associated with trade-offs between gene expression and other cellular resources, including the nutritional quality of the plant (Kim *et al.* 2020).

On functional analysis, one GO-term that was significantly enriched after treatment was *response to stimuli* (Figure 8.8) and consistently, genes related to signalling were among the top expressed genes. For example 1-aminocyclopropane-1-carboxylate oxidase, which is related to production of ethylene; lanC-like protein 2-like for abscissic acid and Tify domain containing protein for jasmonates were strongly responsive. Ethylene is one of the major signalling molecules in plant defences in addition to others, such as jasmonic acid, salicylic acid and abscisic acid (Yang *et al.* 2019). Ethylene can act synergistically or antagonistically with jasmonic acid in the regulation of both stress and developmental responses. The connection between these two signalling pathways has been demonstrated genetically to be the transcription factor for the ethylene response (Lorenzo *et al.* 2003), that was also strongly expressed. This suggests that jasmonates, abscisic acid and ethylene are involved in induced responses of *Pinus radiata* under different stresses. The involvement of jasmonates and ethylene in induced defence responses has been shown in other pine species (Hudgins and Franceschi 2004). In other species, abscisic acid has been shown to be involved in defence responses and has been reported to play a negative role in the regulation of the major photosynthesis gene - type 2 light-harvesting chlorophyll a/b-binding polypeptide (Liu *et al.* 2013b), which was reduced after treatment in this current study.

8.5   Conclusion

There are marked quantitative differences in the needle and bark transcriptome of *Pinus radiata* both in the constitutive and induced states. The transcriptome triggered by bark stripping substantially differed from methyl jamonate triggered responses suggesting that some molecular aspects of bark stripping may differ from other biotic and abiotic responses, which contributes to the understanding of plant molecular responses to diverse stresses. Gene annotation revealed that some of the differentially expressed transcripts have putative functions in plant defence signalling, transcription regulation, biosyntheses of primary and secondary metabolites and other biological processes. The diversity of these genes reflects the complexity of stress responses. The expressed genes provide a basis for further identification of candidate genes that affect bark stripping through variation in their expression levels while the uncharacterized genes that responded to simulated herbivory and methyl jasmonate provide a rich resource for future studies. Gene expression can be used by breeders to exploit phenotype variability among individuals within or between populations. It also remains to be tested whether variations in the transcript levels, particularly the differentially expressed components in reponse to the artificial stress treatments can be linked to the susceptibility classes identified in the field (e.g Chapter 4) and whether they are heritable.

Chapter 9: General discussion

This thesis examined the quantitative and molecular genetic-based variation in bark stripping of *P. radiata* by marsupial herbivores and the mechanistic traits conferring variation in damage. The forestry sector is significantly impacted by damage caused by browsing herbivores (Gill *et al.* 2000; Månsson and Jarnemo 2013; Miller *et al.* 2014; Smith *et al.* 2020; Turek *et al.* 2016) and natural variation in plant susceptibility to herbivory (Alfaro *et al.* 2013; Dimock *et al.* 1976; Russell 2008) may provide a strategy for forest managers to mitigate herbivore impacts in managed plant systems. The results in this thesis show that tree genetics influences the amount of bark removed from trees by marsupial herbivores and that there are potential opportunities to reduce levels of bark stripping by exploiting natural variation in *P. radiata* susceptibility. A major finding of this study is the significant additive genetic variation in bark stripping (Chapter 2) in the current breeding populations of *P. radiata*, and that plant physical and primary and secondary metabolites influence feeding preferences (Chapters 2, 4 and 6), and these can be selected for/against to reduce susceptibility. Predictive genomic approaches have also shown to increase selection accuracy. These will shorten generation intervals. The genomic approaches as well as gene expression may also assist the detection of novel allelic variants and disclose the genomic potential of adaptation to herbivory and other biotic and abiotic stresses.

The physical traits include the bark features (presence of rough bark, rough bark height and bark thickness) that are more important in older plantations, and the presence of obstructive branches or needles on the stem in younger plantations. In both cases, these traits showed significant additive variation. The thick bark generally reduces damage of conifers from herbivores as well as pathogens with few exceptions (Shibata 1998; Tomlin and Borden 1997). The function of bark thickness may be related to the ratio of the outer bark (as the first line of protection) to the inner bark as a source of nutrition. A thicker outer bark may suggest a more passive defence function of the bark, as dead tissue cannot react against threats, but relies solely on its constitution such as texture and chemical composition. On the other hand, a thick inner bark may imply a more dynamic strategy, where the living cells play a role in defence by actively responding to the threat (Morris and Jansen 2016). Therefore, examination of genetic-based variation of the relative proportion of outer and inner bark, may be important to fully understand the importance of bark thickness to the browsing marsupial herbivores. The importance of rough bark in reducing bark stripping by the marsupial herbivores contrasts with the situation for insect herbivores where smooth bark appears to reduce damage of *Pinus* trees based on the slippery' hypothesis (Ferrenberg and Mitton 2014). Such contrasts suggest herbivore specific

responses to physical traits and are important considerations in the presence of multiple herbivores. In Tasmania and other parts of Australia, the phloem-feeding aphid *Essigella californica* represents a potential threat to the productivity of *P. radiata* plantations (Ivković *et al.* 2010b; Sasse *et al.* 2009). *Sirex noctilio* which is still considered the most serious insect problem in radiata pine plantations in Australia (Nahrung *et al.* 2015) appears to prefer rough bark *Pinus* species (Haavik *et al.* 2017). Also, the relationship of the physical traits with other aspects of production such as timber quality due to knots and processing costs due to different bark traits (Murphy and Cown 2015) needs further analysis. In these populations, the physical traits explained only some proportion of the genetic variation in bark stripping, consistent with chemical traits being another potential mechanism of variation.

Secondary metabolites have been suggested as the major mechanistic link between genetic variation in *P. radiata* and levels of herbivory (Lundborg et al. 2019; Moreira et al. 2012). Indeed, the untargeted chemical profiling of experimental plants (Chapter 3) showed that *P. radiata* harbors high chemical diversity within the plant, providing a platform for herbivores to respond. The constitutive and induced chemical profiles differ in the needles, bark and roots (Chapter 3) as may be expected according to the optimal defence theory (McKey 1974). The constitutive and induced chemistry in the needles and the bark of *P. radiata* and other conifers has been previously studied and similar patterns of allocation of secondary compounds have been documented (Lundborg et al. 2019; Moreira et al. 2012). However, this is the first report on plant-wide above and below ground chemistry in *P. radiata,* showing significant levels of secondary compounds in the roots. While root chemistry has been less studied in conifers (Poopat 2013), this topic has received increased attention in the area of plant–root herbivore interactions in non-coniferous systems, as well as plant-soil microbiota interactions (Rasmann and Agrawal 2008), where root chemistry influences root herbivores, soil biota as well as above ground resistance (Swett and Gordon 2017). In *P. radiata* and other conifers, root chemistry may be especially important with respect to protection against root pathogens which may be a threat, especially to seedlings (Reglinski *et al.* 2009). Currently, the root pathogens are managed by chemical agents (Reglinski *et al.* 2009), however, the enhancement of natural resistance to root pathogens may reduce reliance on agrochemicals. In this thesis, the attempts to understand how root chemistry may contribute to susceptibility of *P. radiata* to bark stripping can be inferred from the allocation of compounds, including the trade-offs between below ground and above ground chemistry. For example, the significant positive correlation of total phenolics between the needles and the roots may suggest that the state of phenolic-based defence of the needles can be correlated to the roots. Although the correlations of the amounts of compounds between the roots and the aboveground parts were not prominent, an understanding that roots have distinct qualitative and quantitative, constitutive and

induced chemical profiles relative to the bark and the needles provides a rich information resource for future related studies.

This study included an examination of variation in the primary plant compounds and possible relationships with plant stress and bark-stripping. When examining conifer defence or susceptibility to herbivores, most studies have targeted secondary compounds, and simultaneous examination of primary and secondary compounds is rare (Raffa *et al.* 2017). Primary compounds especially the non-structural carbohydrates (NSC) that include sugars such as glucose and fructose have been studied in *P. radiata* and other conifers predominantly in relation to different abiotic conditions such as seasonal changes or soil properties (Cranswick *et al.* 1987; Tinus *et al.* 2000). This is based on the premise that NSC storage could be used by trees to cope with stress or alternatively that NSC storage competes with growth under stress conditions (Wiley and Helliker 2012). Consistently, the comprehensive examination of stressed plants in this study showed that primary compounds were more responsive to stress treatments than secondary compounds (Chapter 3).

 A similar pattern was also detected at the molecular level where genes related to primary compounds were more strongly expressed following a stressor treatment relative to those related to secondary metabolism. Genes with reduced variance of expression in a selected line may reflect the influence of selective breeding, which is consistent with the artificial selection that has been ongoing in *P. radiata* (Hughes and Buitenhuis 2010). Genes related to primary compounds metabolism were mostly down-regulated while those related to catabolism were up-regulated.

As predicted from the initial experiments (Chapter 3), an important finding in this study was that variation in sugars influenced the degree of bark-stripping by the marsupial herbivores. Characterising the differences between the extreme resistant and susceptible phenotypes indicated that constitutive glucose was more important than the other compounds in differentiating the two groups of plants. Glucose was also strongly genetically correlated to bark stripping. Although in the induced chemistry, the sugars could not separate the resistant and susceptible families, their importance in bark stripping was still signalled by their strong reduction in the susceptible but not in the resistant families. This indicates that after the initial taste by the animals of the susceptible families, their sugar amounts will reduce to the level of the resistant families. So, if sugars strongly determine susceptibility, the attractiveness of the susceptible families will be reduced after the first taste which can minimise multiple damage. Sugars; being important in determining bark stripping is not trivial since nutrition is the primary driver for herbivory and in both *P. radiata* and non-conifers trees, sugars have been correlated with

bark stripping (Kurek *et al.* 2019; Page *et al.* 2013). However, this is the first study to demonstrate the role of sugars in determining genetic differences in mammalian herbivory in conifers. Available studies in non-conifers linking sugars to bark stripping have been undertaken on an ad hoc basis and with limited experimental evidence (Kurek *et al.* 2019; Saint-Andrieux *et al.* 2009). This study has also demonstrated that NSC concentrations are not strictly a passive response to source:sink (Wiley and Helliker 2012) but are under genetic control. The reduction in sugars could be related to the increase in sugar breakdown as a response to stress treatments shown by the up-regulation of genes like the sucrose invertase. It could also be due to repression of photosynthesis that was detected from the down regulation of photosynthesis related genes. Therefore, how gene expression links to the resistant and susceptible categories will enhance the understanding of the molecular mechanisms involved in bark stripping.

In relation to other conifer/herbivore systems, few studies have shown a direct role of sugars in herbivory (Clancy 1992), although their indirect roles in signalling in a manner similar to hormones or as sources of energy have been more widely researched (Deslauriers *et al.* 2015; Tauzin and Giardina 2014; Trouvelot *et al.* 2014). In other studies, sugars may lead to increased resistance, for example, sucrose increased the resistance of *Pseudotsuga menziesii* trees to western spruce budworm (*Choristoneura occidentalis*) (Clancy 1992). Similarly, in most fungal pathogen–plant systems, a high level of sugars in plant tissues enhances plant resistance (Morkunas and Ratajczak 2014). This suggests that different sugars can have contrasting roles in different systems. *P. radiata* is a non-native in Australia and the high amount of sugars could be a defence adaptation to pathogens in California, the native origin where pathogens were more important than mammal herbivory (Mead 2013). If high amounts of sugars promote resistance to pathogens, the selection of trees with less sugars may have direct implications on pathogens like *Dothistroma* needle blight that have been detected in Tasmania and other Australian *P. radiata* plantations (Podger and Wardlaw 1990). The relationship between sugars and pathogens is also important since tissues exposed after bark stripping are prone to pathogens.

Overall, results from this thesis suggest that selecting and potentially breeding for trees that have lower amounts of sugars in the bark may reduce the bark stripping. The potential to reduce sugars through selection was signalled by the significant additive genetic variation with moderate pedigree- based (ABLUP) narrow- sense heritability estimates (h^2 = 0.14 - 0.31). This falls in the range of the traits that are targeted for improvement in *Pinus radiata* e.g. stem straightness and branching (h^2 = 0.23 to 0.55), DBH (h^2 = 0.22 - 0.32) (Gapare *et al.* 2012), height (h^2 = 0.21) (Li *et al.* 2018) as well as disease

resistance (h^2 = 0.28–0.48) (Ismael *et al.* 2020). While the ABLUP estimates were promising, the estimates were significantly improved using genomic best linear unbiased predictions (GBLUP) and single-step GBLUP (ssGBLUP) by up to 2-fold (Chapter 6). This suggests that genomic-based models can increase accuracy of selections. However, sugars appear to accumulate in the bark during winter as an adaptive response for conifers to the cold (Bansal and Germino 2009; Ögren *et al.* 1997), and an understanding of the seasonal changes in sugars between the resistant and susceptible families may provide a further understanding of the basis on the winter bark stripping. In *P. radiata*, Cranswick *et al.* (1987) showed existence of seasonal variation of sugars in the needles, bark and roots but how these contribute to intraspecific variation in bark stripping is not known.

The positive correlation between prior height and bark stripping in the initial stages of growth is however a challenge, where selection for fast growth as a common breeding objective in *P. radiata* may make trees more vulnerable to browsing, especially at early ages. Positive correlation between growth and herbivory is a common trend in herbivore studies (King *et al.* 1997; Mottet *et al.* 2015; Zas *et al.* 2008; Zas *et al.* 2005). However, in this study further analysis of the relationship between height and bark stripping showed that height is not an independent factor but is possibly related to the accumulation of sugars or the reduction of secondary metabolites following fast growth (Ferrenberg *et al.* 2015; Kenward *et al.* 1996). On the other hand, it has also been suggested that fast growing - susceptible trees should be able to recover more quickly from herbivory damage than slower growing-resistant trees (Gianoli and Salgado-Luarte 2017; Wiley and Helliker 2012) - an aspect of tolerance that needs further research for *P. radiata* bark stripping. For commercial *P. radiata* plantations intended for timber production, tolerance will however be less desirable since bark stripping exposes tissues to fungal attack with subsequent rotting which reduces timber quality (Cukor *et al.* 2019). Bark stripping may also reduce tolerance to freezing that may be associated with subsequent chemical changes (Fedderwitz *et al.* 2020).

While sugars became a dominant story in this thesis, the importance of constitutive secondary compounds- phenolics and sesquiterpenoids was also noted. Total and individual phenolics were also important in differentiating the resistant and susceptible families although they increased in the susceptible families and were genetically, positively correlated with bark stripping. Phenolics seem to be more important when considered as a sum of their individual components in contrast to the sugars where individual but not total compounds are more important. Also contrary to the strong reduction observed in the sugars, the phenolics responded to stress treatments by mostly increasing their amounts. However, the changes in the amount of individual phenolics were not as intense as those

observed with the reduction in the sugars. The consistent role of phenolics in defence against mammalian herbivores is still not clear as there is contrasting evidence in the literature (Radwan 1972; Sunnerheim-Sjöberg and Hämäläinen 1992). However, the direct and indirect roles of phenolics in constitutive or inducible stress responses, such as cell division, hormonal regulation, photosynthetic activity and nutrient mineralization (Bandau *et al.* 2015; Hammerbacher *et al.* 2011) have been suggested. It is also possible that phenolics may have an attractant rather than a defensive effect, especially in low concentrations. Of the terpenes, sesquiterpenes, bicyclogermacrene and an unknown sesquiterpenoid alcohol were more important in reducing susceptibility to bark stripping by marsupials in contrast to monoterpenes that have been more implicated in conifer susceptibility to bark insect herbivores (Raffa and Smalley 1995). For other mammalian herbivores which browse on needles such as deer, monoterpenes have been suggested as important deterrents (Elliott and Loudon 1987; Russell 2008; Vourc'h *et al.* 2002a), suggesting that defences differ for different herbivores (Korth 2003; Ohse *et al.* 2017). Australian marsupials have evolved to ingest and metabolise a range of dietary terpenes and phenols that would be toxic to many other herbivore species (Boyle 1999; El-Merhibi *et al.* 2007). This thesis shows that secondary compounds can also potentially be selected for as they exhibited strong additive genetic variation with low-moderate pedigree and genomic based heritability. Previous studies indicate that intraspecific variation for secondary metabolites in most plant species is high and may be subject to selection (Iason *et al.* 2011; Moore *et al.* 2014; O'Reilly-Wapstra *et al.* 2013b).

Near infrared spectroscopy (NIRS) was used to accurately quantify the amount of compounds in the different plant parts, which enabled the chemotyping of a larger number of samples for genetic studies. Although NIR spectroscopy is not a new technology, its use for quantifying secondary compounds in trees has not been very common (Couture *et al.* 2016) being partly due to the scepticism in its ability to quantify compounds that often occur in very small quantities. However, this study has shown that high accuracy of prediction can be attained even for low quantity metabolites. Other studies that have compared the accuracy of NIRS and conventional wet chemistry also show very high correlation between the two methods (Harris *et al.* 2018). Therefore, studies focussing on the application of phenolics, terpenes, and other secondary metabolites, as well as sugars or fatty acids in areas of stress and defence responses, plant physiological, plant extractives among others can be enhanced; taking advantage of the cost effectiveness, rapid analysis, recent improvements in the interpretability of NIR spectra and wide-applicability to a variety of samples. The work in this thesis was done using benchtop spectrometers, however, the flexible instrumentation including hand-held NIR spectrometers enable on-site examination of samples (Yan *et al.* 2019) making the analysis even faster.

In conclusion, based on the extreme less susceptible and more susceptible families, differences in the amounts of secondary compounds were very subtle and more positive than negative genetic correlations of compounds with bark stripping were observed suggesting that the importance of nutrition may outweigh defence in this non-native *P. radiata* system.

Supplementary Table 10: Compounds that were detected in the needles, bark and roots of, in the shade house and the field experiments. ✓ =detected. The identifier is used in the text for ease of identification of the compound. Compounds labelled "A", "B", "C" and "D" are isomers

Identifier		Shade house samples (Chapters 3 and 5)			Field samples (Chapters 4 and 6	
		Bark	Needles	Roots	Bark	Needles
	Monoterpenoids					
1	α-pinene	✓	✓	✓	✓	✓
2	α-terpineol	✓			✓	✓
3	β-phellandrene	✓	✓	✓	✓	✓
4	β-pinene	✓	✓	✓	✓	✓
5	camphene	✓	✓		✓	✓
6	citronellal	✓			✓	
7	citronellic acid	✓			✓	✓
8	citronellol	✓			✓	✓
9	γ-terpinene	✓			✓	✓
10	limonene	✓	✓	✓	✓	✓
11	linalool	✓			✓	✓
12	myrtenoic acid			✓		
13	sabinene	✓	✓	✓	✓	✓
14	terpinene-4-ol	✓			✓	✓
15	terpinolene	✓		✓	✓	✓
16	unknown Mol Wt 150	✓	✓		✓	✓
	Sesquiterpenoids					
17	bicyclogermacrene	✓	✓		✓	✓
18	caryophyllene					✓
19	γ-elemene					✓
20	trans-farnesol	✓	✓		✓	✓
21	unknown sesquiterpenoid alcohol	✓	✓		✓	✓
	GC-MS diterpenoids					
22	agathadiol	✓	✓		✓	✓
23	agatholal	✓	✓		✓	✓
24	copalol		✓		✓	✓
25	levopimaral	✓		✓	✓	
26	methyl dehydroabietate	✓	✓	✓	✓	✓
27	methyl levopimarate	✓		✓		✓
28	unknown C19H26	✓				

LC-MS diterpenoids

#		C1	C2	C3	C4	C5
29	dehydroabietic acid	✓	✓	✓	✓	✓
30	unknown diterpene-1	✓				
31	unknown diterpene-2			✓		
32	unknown diterpene-3	✓	✓	✓	✓	
33	unknown m/z 109 A	✓		✓	✓	✓
34	unknown m/z 109 B	✓	✓		✓	✓
35	unknown m/z 121			✓		
36	unknown m/z 134	✓				
37	unknown m/z 239		✓	✓	✓	✓
38	unknown Mol Wt 272	✓			✓	✓
39	unknown C20H30O2 A			✓		✓
40	unknown C20H30O2 B		✓			✓
41	C20H30O2 resin acids	✓	✓	✓	✓	✓
42	unknown m/z 304 A		✓		✓	
43	unknown m/z 304 B		✓			✓
44	unknown m/z 304 C	✓				
45	unknown m/z 316	✓			✓	
46	unknown C20H30O3	✓	✓	✓	✓	✓
47	unknown C20H32O3 A	✓	✓	✓	✓	✓
48	unknown C20H32O3 B		✓		✓	✓
49	unknown C20H32O3 C		✓			✓
50	unknown C20H30O4	✓	✓	✓	✓	✓
51	unknown C20H30O5	✓		✓	✓	
52	unknown C20H30O6 A	✓			✓	
53	unknown C20H30O6 B		✓			✓
54	unknown C20H30O6 C	✓			✓	
55	unknown C20H30O6 D				✓	

phenolics

#		C1	C2	C3	C4	C5
56	anethole	✓			✓	✓
57	benzene acetic acid				✓	✓
58	chavicol					✓
59	coniferyl alcohol				✓	✓
60	eugenol					✓
61	ethyl phenol				✓	
62	ethyl 4-ethoxybenzoate	✓		✓		
63	4-ethyl guaiacol				✓	
64	isoeugenol					✓

#	compound					
65	methyl eugenol	✓			✓	
66	p-Menth-1-en-7,8-diol				✓	
67	phenyl ethanol				✓	✓
68	pinosylvin dimethyl ether	✓			✓	
69	piperitone					✓
70	raspberry ketone	✓	✓		✓	✓
71	thymol				✓	✓
72	trans-ferulic acid				✓	✓
73	trans-coniferyl alcohol	✓	✓			
74	vanillin	✓			✓	✓
75	zingerone					✓
	sugars					
76	fructose	✓	✓	✓	✓	✓
77	glucose	✓	✓	✓	✓	✓
78	inositol	✓	✓	✓	✓	✓
79	sucrose	✓		✓	✓	✓
80	unknown disaccharide A	✓	✓	✓		
81	unknown disaccharide B	✓	✓	✓		✓
82	unknown monosaccharide	✓	✓		✓	
	fatty acids					
83	linoleic acid	✓	✓	✓	✓	✓
84	linolenic acid	✓	✓	✓	✓	✓
85	palmitic acid	✓	✓	✓	✓	✓
	unknowns					
86	unknown m/z 104	✓			✓	✓
87	unknown m/z 111			✓		
88	unknown m/z 162	✓				
89	unknown m/z 272		✓			✓
90	unknown m/z 274	✓		✓	✓	
91	unknown m/z 302		✓			
92	unknown m/z 358		✓			✓
93	unknown m/z 362		✓			✓
94	unknown m/z 406 A	✓		✓	✓	
95	unknown m/z 406 B				✓	
96	unknown m/z 740 A	✓				
97	unknown m/z 740 B		✓			✓
98	unknown m/z 770		✓			

References

ABARES (2018) Australia's State of the Forests Report Australian Government Department of Agriculture and Water Resources, Canberra.

Achotegui-Castells A, Llusià J, Hódar JA, Peñuelas J (2013) Needle terpene concentrations and emissions of two coexisting subspecies of Scots pine attacked by the pine processionary moth (*Thaumetopoea pityocampa*). *Acta Physiologiae Plantarum* **35**, 3047-3058.

Acquah GE, Via BK, Gallagher T, Billor N, Fasina OO, Eckhardt LG (2018) High throughput screening of elite loblolly pine families for chemical and bioenergy traits with near infrared spectroscopy. *Forests* **9**, 418.

Agrawal AA (2000) Specificity of induced resistance in wild radish: causes and consequences for two specialist and two generalist caterpillars. *Oikos* **89**, 493-500.

Agrawal AA, Weber MG (2015) On the study of plant defence and herbivory using comparative approaches: how important are secondary plant compounds. *Ecology Letters* **18**, 985-991.

Akashi N, Terazawa K (2005) Bark stripping damage to conifer plantations in relation to the abundance of sika deer in Hokkaido, Japan. *Forest Ecology and Management* **208**, 77-83.

Al-Asbahi AA, Al-Maqtari MA, Naji KM (2012) ABA biosynthesis defective mutants reduce some free amino acids accumulation under drought stress in tomato leaves in comparison with *Arabidopsis* plants tissues. *Journal of Stress Physiology & Biochemistry* **8**, 179-192.

Alfaro RI, King JN, Brown RG, Buddingh SM (2008) Screening of Sitka spruce genotypes for resistance to the white pine weevil using artificial infestations. *Forest Ecology and Management* **255**, 1749-1758.

Alfaro RI, King JN, vanAkker L (2013) Delivering Sitka spruce with resistance against white pine weevil in British Columbia, Canada. *The Forestry Chronicle* **89**, 235-245.

Alfaro RI, vanAkker L, Jaquish B, King J (2004) Weevil resistance of progeny derived from putatively resistant and susceptible interior spruce parents. *Forest Ecology and Management* **202**, 369-377.

Alvarez C, Valledor L, Hasbún R, Sanchez-Olate M, Ríos D (2011) Variation in gene expression profile with aging of *Pinus radiata* D. Don. *BMC Proceedings* **5**, P62.

Alvarez C, Valledor L, Sáez P, Hasbún R, Sánchez-Olate M, Cañal MJ, Ríos D (2016) Changes in gene expression in needles and stems of *Pinus radiata* rootstock plants of different ontogenic age. *American Journal of Plant Sciences* **7**, 1205-1216.

Amri I, Hanana M, Gargouri S, Jamoussi B, Hamrouni L (2013) Comparative study of two coniferous species (*Pinus pinaster* Aiton and *Cupressus sempervirens* L. var. *dupreziana* [A. Camus] Silba) essential oils: chemical composition and biological activity. *Chilean Journal of Agricultural Research* **73**, 259-266.

Anderson JT, Mitchell-Olds T (2011) Ecological genetics and genomics of plant defenses: evidence and approaches. *Functional Ecology* **25**, 312-324.

Andrews S (2018) FastQC: a quality control tool for high throughput sequence data. In. ' 0.11.8 edn. (Babraham Bioinformatics, Babraham Institute: Cambridge, United Kingdom)

Apetrei CL, Tuchilus C, Aprotosoaie AC, Oprea A, Malterud KE, Miron A (2011) Chemical, antioxidant and antimicrobial investigations of *Pinus cembra* L. bark and needles. *Molecules* **16**, 7773-7788.

Araus JL, Cairns JE (2014) Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science* **19**, 52-61.

Arhipova N, Jansons A, Zaluma A, Gaitnieks T, Vasaitis R (2015) Bark stripping of *Pinus contorta* caused by moose and deer: wounding patterns, discoloration of wood, and associated fungi. *Canadian Journal of Forest Research* **45**, 1434-1438.

Arojju SK, Cao M, Jahufer MZZ, Barrett BA, Faville MJ (2019) Genomic predictive ability for foliar nutritive traits in perennial ryegrass. *Genes, Genomes, Genetics* **10**, 695-708.

Arruda MP, Lipka AE, Brown PJ, Krill AM, Thurber C, Brown-Guedira G, Dong Y, Foresman BJ, Kolb FL (2016) Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (*Triticum aestivum* L.). *Molecular Breeding* **36**, 84.

Aspinwall MJ, Loik ME, Resco de Dios V, Tjoelker MG, Payton PR, Tissue DT (2015) Utilizing intraspecific variation in phenotypic plasticity to bolster agricultural and forest productivity under climate change. *Plant, Cell & Environment* **38**, 1752-1764.

Axelrod D (1988) Paleoecology of a Late Pleistocene Monterey pine at Laguna Niguel, southern California. *International Journal of Plant Sciences* **149**, 458-464.

Babst BA, Ferrieri RA, Gray DW, Lerdau M, Schlyer DJ, Schueller M, Thorpe MR, Orians CM (2005) Jasmonic acid induces rapid changes in carbon transport and partitioning in *Populus*. *New Phytologist* **167**, 63-72.

Babst BA, Ferrieri RA, Thorpe MR, Orians CM (2008) *Lymantria dispar* herbivory induces rapid changes in carbon transport and partitioning in *Populus nigra*. *Entomologia Experimentalis et Applicata* **128**, 117-125.

Bairoch A, Apweiler R (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research* **28**, 45-48.

Baker SM, White EE (1996) A chalcone synthase/stilbene synthase DNA probe for conifers. *Theoretical and Applied Genetics* **92**, 827-831.

Bandau F, Decker VHG, Gundale MJ, Albrectsen BR (2015) Genotypic tannin levels in *Populus tremula* impact the way nitrogen enrichment affects growth and allocation responses for some traits and not for others. *PLoS One* **10**, e0140971.

Bannister MH (1969) An early progeny trial in *Pinus radiata* 1. Juvenile characters. *New Zealand Journal of Botany* **7**, 316-335.

Bansal S, Germino MJ (2009) Temporal variation of nonstructural carbohydrates in montane conifers: similarities and differences among developmental stages, species and environmental conditions. *Tree Physiology* **29**, 559-568.

Baradat P, Yazdani R (1988) Genetic expression for monoterpenes in clones of *Pinus sylvestris* grown on different sites. *Scandinavian Journal of Forest Research* **3**, 25-36.

Barbosa dos Santos I, Park S-W (2019) Versatility of cyclophilins in plant growth and survival: a case study in *Arabidopsis*. *Biomolecules* **9**, 20.

Barkan A, Small I (2014) Pentatricopeptide repeat proteins in plants. *Annual Review of Plant Biology* **65**, 415-442.

Barnes D, Lai W, Breslav M, Naider F, Becker JM (1998) *PTR3*, a novel gene mediating amino acid-inducible regulation of peptide transport in *Saccharomyces cerevisiae*. *Molecular Microbiology* **29**, 297-310.

Barton KE, Koricheva J (2010) The ontogeny of plant defense and herbivory: characterizing general patterns using meta-analysis. *The American Naturalist* **175**, 481-493.

Beaulieu J, Doerksen T, Clément S, MacKay J, Bousquet J (2014) Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity* **113**, 343-352.

Beć KB, Grabska J, Czarnecki MA (2018) Spectra-structure correlations in NIR region: spectroscopic and anharmonic DFT study of n-hexanol, cyclohexanol and phenol. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **197**, 176-184.

Behringer D, Zimmermann H, Ziegenhagen B, Liepelt S (2015) Differential gene expression reveals candidate genes for drought stress response in *Abies alba* (Pinaceae). *PLoS One* **10**, e0124564.

Berrocal-Lobo M, Segura A, Moreno M, López G, García-Olmedo F, Molina A (2002) Snakin-2, an antimicrobial peptide from potato whose gene is locally induced by wounding and responds to pathogen infection. *Plant physiology* **128**, 951-961.

Bhattacharya A, Sood P, Citovsky V (2010) The roles of plant phenolics in defence and communication during Agrobacterium and Rhizobium infection. *Molecular Plant Pathology* **11**, 705-719.

Bockwoldt M, Heiland I, Fischer K (2019) The evolution of the plastid phosphate translocator family. *Planta* **250**, 245-261.

Bohlmann J, Croteau R (1999) Diversity and variability of terpenoid defences in conifers: molecular genetics, biochemistry and evolution of the terpene synthase gene family in grand fir (*Abies grandis*). In 'Insect–Plant Interactions and Induced Plant Defence. (Eds DJ Chadwick and JA Goode) pp. 132-149. (John Wiley & Sons: Chicester, West Sussex, UK).

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120.

Bommer U-A, Thiele B-J (2004) The translationally controlled tumour protein (TCTP). *The International Journal of Biochemistry & Cell Biology* **36**, 379-85.

Bonello P, Gordon TR, Herms DA, Wood DL, Erbilgin N (2006) Nature and ecological implications of pathogen-induced systemic resistance in conifers: a novel hypothesis. *Physiological and Molecular Plant Pathology* **68**, 95-104.

Bonello P, Gordon TR, Storer AJ (2001) Systemic induced resistance in Monterey pine. *Forest Pathology* **31**, 99-106.

Boyle RR (1999) The metabolic fate of dietary terpenes in folivorous marsupials. University of Tasmania, Hobart

Bredsdorff L, Wedebye EB, Nikolov NG, Hallas-Møller T, Pilegaard K (2015) Raspberry ketone in food supplements – High intake, few toxicity data – A cause for safety concern? *Regulatory Toxicology and Pharmacology* **73**, 196-200.

Breiman L (2001) Random forests. *Machine Learning* **45**, 5-32.

Brockerhoff E, Bulman L (2014) Biosecurity risks to New Zealand's plantation forests and the rationale for pathway risk management. *New Zealand Journal of Forestry* **59**, 3-8.

Bucyanayandi JD, Bergeron J-M, Menard H (1990) Preference of meadow voles (*Microtus pennsylvanicus*) for conifer seedlings: chemical components and nutritional quality of bark of damaged and undamaged trees. *Journal of Chemical Ecology* **16**, 2569-2579.

Burdon RD (1977) Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. *Silvae Genetica* **26**, 168-175.

Burdon RD, Gaskin RE, Low CB, Zabkiewicz JA (1992a) Clonal repeatability of monoterpene composition of cortical oleoresin of *Pinus radiata*. *New Zealand Journal of Forestry Science* **22**, 299-305.

Burdon RD, Libby WJ, Brown AG (2017) 'Domestication of Radiata Pine.' (Springer Nature: Cham, Switzerland).

Burdon RD, Zabkiewicz JA, Andrew IA (1992b) Genetic survey of *Pinus radiata*. 8: Population differences in monoterpene composition of cortical oleoresin. *New Zealand Journal of Forestry Science* **22**, 257-273.

Burke CC, Wildung MR, Croteau R (1999) Geranyl diphosphate synthase: Cloning, expression, and characterization of this prenyltransferase as a heterodimer. *Proceedings of the National Academy of Sciences* **96**, 13062-13067.

Burnham KP, Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research* **33**, 261-304.

Busov VB, Johannes E, Whetten RW, Sederoff RR, Spiker SL, Lanz-Garcia C, 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.

Butler DG, Cullis BR, Gilmour AR, Gogel BJ (2009) ASReml-R reference manual. In. ' Version 3 edn. pp. 149. (Queensland Government Department of Primary Industries and Fisheries: Brisbane, Qld)

Byun-McKay A, Godard K-A, Toudefallah M, Martin DM, Alfaro R, King J, Bohlmann J, Plant AL (2006) Wound-induced terpene synthase gene expression in Sitka spruce that exhibit resistance or susceptibility to attack by the white pine weevil. *Plant Physiology* **140**, 1009-1021.

Çakir B, Agasse A, Gaillard C, Saumonneau A, Delrot S, Atanassova R (2003) A grape ASR protein involved in sugar and abscisic acid signaling. *The Plant Cell* **15**, 2165-2180.

Calenge F, Saliba-Colombani V, Mahieu S, Loudet O, Daniel-Vedele F, Krapp A (2006) Natural variation for carbohydrate content in *Arabidopsis*. Interaction with complex traits dissected by quantitative genetics. *Plant Physiology* **141**, 1630-1643.

Campa F, Randazzo PA (2008) Arf GTPase-activating proteins and their potential role in cell migration and invasion. *Cell Adhesion & Migration* **2**, 258-262.

Cappa EP, de Lima BM, da Silva-Junior OB, Garcia CC, Mansfield SD, Grattapaglia D (2019) Improving genomic prediction of growth and wood traits in *Eucalyptus* using phenotypes from non-genotyped trees by single-step GBLUP. *Plant Science* **284**, 9-15.

Carmona D, Fornoni J (2013) Herbivores can select for mixed defensive strategies in plants. *New Phytologist* **197**, 576-585.

Carmona D, Lajeunesse MJ, Johnson MTJ (2011) Plant traits that predict resistance to herbivores. *Functional Ecology* **25**, 358-367.

Carpenter MA, Goulden DS, Woods CJ, Thomson SJ, Kenel F, Frew TJ, Cooper RD, Timmerman-Vaughan GM (2018) Genomic selection for ascochyta blight resistance in pea. *Frontiers in Plant Science* **9**, 1878.

Carson SD (1989) Selecting *Pinus radiata* for resistance to Dothistroma needle blight. *New Zealand Journal of Forestry Science* **19**, 3-21.

Cécillon L, Barthès BG, Gomez C, Ertlen D, Genot V, Hedde M, Stevens A, Brun JJ (2009) Assessment and monitoring of soil quality using near‐infrared reflectance spectroscopy (NIRS). *European Journal of Soil Science* **60**, 770-784.

Celedon JM, Bohlmann J (2019) Oleoresin defenses in conifers: chemical diversity, terpene synthases and limitations of oleoresin defense under climate change. *New Phytologist* **224**, 1444-1463.

Celedon JM, Yuen MMS, Chiang A, Henderson H, Reid KE, Bohlmann J (2017) Cell-type- and tissue-specific transcriptomes of the white spruce (*Picea glauca*) bark unmask fine-scale spatial patterns of constitutive and induced conifer defense. *The Plant Journal* **92**, 710-726.

Chapter 2 Quantitative genetic variation in bark stripping of *Pinus radiata*. In 'Thesis.').

Chapter 3 Variation in constitutive and induced chemistry in the needles, bark and roots of *Pinus radiata*. In 'Thesis.').

Chapter 4 Chemical traits that predict susceptibility of *P. radiata* to marsupial bark stripping In 'Thesis.').

Chapter 5 Developing near infrared spectroscopy (NIRS) models for predicting chemistry and responses to stress in Pinus radiata (D. Don).

Chapter 6 Additive genetic variation in Pinus radiata bark chemistry and the chemical traits associated with variation in bark stripping.

Chatthai M, Kaukinen KH, Tranbarger TJ, Gupta PK, Misra S (1997) The isolation of a novel metallothionein-related cDNA expressed in somatic and zygotic embryos of Douglas-fir: regulation by ABA, osmoticum, and metal ions. *Plant Molecular Biology* **34**, 243-254.

Chatthai M, Osusky M, Osuska L, Yevtushenko D, Misra S (2004) Functional analysis of a Douglas-fir metallothionein-like gene promoter: transient assays in zygotic and somatic embryos and stable transformation in transgenic tobacco. *Planta* **220**, 118-128.

Chen H, Gonzales-Vigil E, Wilkerson CG, Howe GA (2007) Stability of plant defense proteins in the gut of insect herbivores. *Plant Physiology* **143**, 1954-1967.

Chen T, Cao Y, Zhang Y, Liu J, Bao Y, Wang C, Jia W, Zhao A (2013) Random forest in clinical metabolomics for phenotypic discrimination and biomarker selection. *Evidence-Based Complementary and Alternative Medicine* **2013**, 11.

Chen W-H, Hsu C-Y, Cheng H-Y, Chang H, Chen H-H, Ger M-J (2011) Downregulation of putative UDP-glucose: flavonoid 3-*O*-glucosyltransferase gene alters flower coloring in *Phalaenopsis*. *Plant Cell Reports* **30**, 1007-1017.

Chen Y-Y, Xu Z-S, Xiong A-S (2016) Identification and characterization of DcUSAGT1, a UDP-glucose: sinapic acid glucosyltransferase from purple carrot taproots. *PLoS One* **11**, e0154938.

Chen Z-Q, Baison J, Pan J, Karlsson B, Andersson B, Westin J, García-Gil MR, Wu HX (2018) Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. *BMC Genomics* **19**, 946.

Chiu CC, Keeling CI, Bohlmann J (2017) Toxicity of pine monoterpenes to mountain pine beetle. *Scientific Reports* **7**, 8858.

Christensen OF, Lund MS (2010) Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution* **42**, 2.

Chu X-l, Yuan H-f, Lu W-z (2005) Effects of the accuracy of reference data on NIR prediction results. *Guang Pu Xue Yu Guang Pu Fen Xi* **25**, 886-889.

Clancy KM (1992) The role of sugars in western spruce budworm nutritional ecology. *Ecological Entomology* **17**, 189-197.

Close DC, McArthur C (2002) Rethinking the role of many plant phenolics – protection from photodamage not herbivores? *Oikos* **99**, 166-172.

Coley PD, Bryant JP, Chapin FS, III (1985) Resource availability and plant antiherbivore defense. *Science* **230**, 895-899.

Cool LG, Zavarin E (1992) Terpene variability of mainland *Pinus radiata*. *Biochemical Systematics and Ecology* **20**, 133-144.

Coops NC, Stone C (2005) A comparison of field-based and modelled reflectance spectra from damaged *Pinus radiata* foliage. *Australian Journal of Botany* **53**, 417-429.

Corbett LJ, Withey P, Lantz VA, Ochuodho TO (2015) The economic impact of the mountain pine beetle infestation in British Columbia: provincial estimates from a CGE analysis. *Forestry* **89**, 100-105.

Costa e Silva J, Dutkowski GW, Gilmour AR (2001) Analysis of early tree height in forest genetic trials is enhanced by including a spatially correlated residual. *Canadian Journal of Forest Research* **31**, 1887-1893.

Couture JJ, Singh A, Rubert‑Nason KF, Serbin SP, Lindroth RL, Townsend PA (2016) Spectroscopic determination of ecologically relevant plant secondary metabolites. *Methods in Ecology and Evolution* **7**, 1402-1412.

Covarrubias-Pazaran G, Schlautman B, Diaz-Garcia L, Grygleski E, Polashock J, Johnson-Cicalese J, Vorsa N, Iorizzo M, Zalapa J (2018) Multivariate GBLUP improves accuracy of genomic selection for yield and fruit weight in biparental populations of *Vaccinium macrocarpon* Ait. *Frontiers in Plant Science* **9**.

Cranswick AM, Rook DA, Zabkiewicz JA (1987) Seasonal changes in carbohydrate concentration and composition of different tissue types of *Pinus radiata* trees. *New Zealand Journal of Forestry Science* **17**, 229-45.

Cronn R, Dolan PC, Jogdeo S, Wegrzyn JL, Neale DB, St. Clair JB, Denver DR (2017) Transcription through the eye of a needle: daily and annual cyclic gene expression variation in Douglas-fir needles. *BMC Genomics* **18**, 558.

Crossa J, Pérez-Rodríguez P*, et al.* (2017) Genomic selection in plant breeding: methods, models, and perspectives. *Trends in Plant Science* **22**, 961-975.

Cukor J, Vacek Z, Linda R, Vacek S, Marada P, Šimůnek V, Havránek F (2019) Effects of bark stripping on timber production and structure of Norway spruce forests in relation to climatic factors. *Forests* **10**, 320.

D'Agui H, Fowler W, Lim SL, Enright N, He T (2016) Phenotypic variation and differentiated gene expression of Australian plants in response to declining rainfall. *Royal Society Open Science* **3**, 160637.

Danielsson M, Lundén K*, et al.* (2011) Chemical and transcriptional responses of Norway spruce genotypes with different susceptibility to *Heterobasidion* spp. infection. *BMC Plant Biology* **11**, 154.

de A. Gerhardt LB, Sachetto-Martins G, Contarini MG, Sandroni M, de P. Ferreira R, de Lima VM, Cordeiro MC, de Oliveira DE, Margis-Pinheiro M (1997) *Arabidopsis thaliana* class IV chitinase is early induced during the interaction with *Xanthomonas campestris*. *FEBS Letters* **419**, 69-75.

de la Mata R, Hood S, Sala A (2017) Insect outbreak shifts the direction of selection from fast to slow growth rates in the long-lived conifer *Pinus ponderosa*. *Proceedings of the National Academy of Sciences* **114**, 7391-7396.

De Rienzo F, Gabdoulline RR, Menziani MC, Wade RC (2000) Blue copper proteins: A comparative analysis of their molecular interaction properties. *Protein Science* **9**, 1439-1454.

Deslauriers A, Caron L, Rossi S (2015) Carbon allocation during defoliation: testing a defense-growth trade-off in balsam fir. *Frontiers in Plant Science* **6**.

Desta ZA, Ortiz R (2014) Genomic selection: genome-wide prediction in plant improvement. *Trends in Plant Science* **19**, 592-601.

Di Bitetti MS (2019) Primates bark-stripping trees in forest plantations – A review. *Forest Ecology and Management* **449**, 117482.

Dieters MJ, Hodge GR, White TL (1996) Genetic parameter estimates for resistance to rust (*Cronartium quercuum*) infection from full-sib tests of slash pine (*Pinus elliottii*), modelled as functions of rust incidence. *Silvae Genetica* **45**, 235-242.

Dimock EJ, II, Silen RR, Allen VE (1976) Genetic resistance in Douglas-fir to damage by snowshoe hare and black-tailed deer. *Forest Science* **22**, 106-121.

Dixon RA, Paiva NL (1995) Stress-induced phenylpropanoid metabolism. *The Plant Cell* **7**, 1085-1097.

Dobbelstein E, Fink D, Öner-Sieben S, Czempik L, Lohaus G (2019) Seasonal changes of sucrose transporter expression and sugar partitioning in common European tree species. *Tree Physiology* **39**, 284-299.

Donaldson LA, Croucher M, Uprichard JM (1997) Clonal variation of wood chemistry variables in radiata pine (*Pinus radiata* D. Don.) wood. *Holzforschung* **51**, 537-542.

Downham R, Gavran M (2018) Australian plantation statistics 2018 update. Australian Government Department of Agriculture and Water Resources, Canberra.

DPIPWE (2018) A planning toolkit for managing browsing and grazing losses from wallabies and brushtail possums on farms in Tasmania. In. ' Ed. P Department of Primary Industries, Water and Environment). (Tasmanian Government: Launceston, Tasmania)

DPIPWE (2019) Annual state-wide spotlight surveys, Tasmania 2018/19. Regional summary: priority harvested species. In. ' Ed. DoPIPWa Environment). (Tasmanian Government: Tasmania)

Du M, Ding G, Cai Q (2018) The transcriptomic responses of *Pinus massoniana* to drought stress. *Forests* **9**, 326.

Duncan AJ, Hartley SE, Thurlow M, Young S, Staines BW (2001) Clonal variation in monoterpene concentrations in Sitka spruce (*Picea sitchensis*) saplings and its effect on their susceptibility to browsing damage by red deer (*Cervus elaphus*). *Forest Ecology and Management* **148**, 259-269.

Dungey HS, Brawner JT, Burger F, Carson M, Henson M, Jefferson P, Matheson AC (2009) A new breeding strategy for *Pinus radiata* in New Zealand and New South Wales. *Silvae Genetica* **58**, 28-38.

Dungey HS, Russell JH, Costa e Silva J, Low CB, Miller MA, Fleet KR, Stovold GT (2013) The effectiveness of cloning for the genetic improvement of Mexican white cypress *Cupressus lusitanica* (Mill.). *Tree Genetics & Genomes* **9**, 443-453.

Dungey HS, Williams NM, Low CB, Stovold GT (2014) First evidence of genetic-based tolerance to red needle cast caused by *Phytophthora pluvialis* in radiata pine. *New Zealand Journal of Forestry Science* **44**, 31.

Dutkowski GW, Costa e Silva J, Gilmour AR, Wellendorf H, Aguiar A (2006) Spatial analysis enhances modelling of a wide variety of traits in forest genetic trials. *Canadian Journal of Forest Research* **36**, 1851-1870.

Edwards A, Borthakur A, Bornemann S, Venail J, Denyer K, Waite D, Fulton D, Smith A, Martin C (1999) Specificity of starch synthase isoforms from potato. *European Journal of Biochemistry* **266**, 724-736.

Ehrlich PR, Raven PH (1964) Butterflies and plants: a study in coevolution. *Evolution* **18**, 586-608.

El-Merhibi A, Ngo SNT, Jones BR, Milic NL, Stupans I, McKinnon RA (2007) Molecular insights into xenobiotic disposition in Australian marsupials. *Australasian Journal of Ecotoxicology* **13**, 7-18.

Eldar A, Chary VK, Xenopoulos P, Fontes ME, Losón OC, Dworkin J, Piggot PJ, Elowitz MB (2009) Partial penetrance facilitates developmental evolution in bacteria. *Nature* **460**, 510-514.

Eldridge KG (1979) Seed collections of *Pinus radiata* and *P. muricata* in California. FAO,Report Book No. 9, Canberra.

Elliott S, Loudon A (1987) Effects of monoterpene odors on food selection by red deer calves (*Cervus elaphus*). *Journal of Chemical Ecology* **13**, 1343-1349.

Endelman JB (2011) Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome* **4**, 250-255.

Endress BA, Naylor BJ, Pekin BK, Wisdom MJ (2016) Aboveground and belowground mammalian herbivores regulate the demography of deciduous woody species in conifer forests. *Ecosphere* **7**, e01530.

Erbilgin N, Cale JA, Hussain A, Ishangulyyeva G, Klutsch JG, Najar A, Zhao S (2017) Weathering the storm: how lodgepole pine trees survive mountain pine beetle outbreaks. *Oecologia* **184**, 469-478.

Erbilgin N, Colgan LJ (2012) Differential effects of plant ontogeny and damage type on phloem and foliage monoterpenes in jack pine (*Pinus banksiana*). *Tree Physiology* **32**, 946-957.

Ermakova EA, Faizullin DA*, et al.* (2016a) Structure of Scots pine defensin 1 by spectroscopic methods and computational modeling. *International Journal of Biological Macromolecules* **84**, 142-152.

Ermakova EA, Faizullin DA*, et al.* (2016b) Structure of Scots pine defensin 1 by spectroscopic methods and computational modeling. *International Journal of Biological Macromolecules* **84**, 142-152.

Eveland AL, Jackson DP (2012) Sugars, signalling, and plant development. *Journal of Experimental Botany* **63**, 3367-3377.

Ewing B, Green P (1998) Base-calling of automated sequencer traces using *phred*. II. Error probabilities. *Genome Research* **8**, 186-194.

Eyles A, Smith D, Pinkard EA, Smith I, Corkrey R, Elms S, Beadle C, Mohammed C (2011) Photosynthetic responses of field-grown *Pinus radiata* trees to artificial and aphid-induced defoliation. *Tree Physiology* **31**, 592-603.

Fahey LM, Nieuwoudt MK, Harris PJ (2018) Using near infrared spectroscopy to predict the lignin content and monosaccharide compositions of *Pinus radiata* wood cell walls. *International Journal of Biological Macromolecules* **113**, 507-514.

Falconer DS, Mackay TFC (1996) 'Introduction to Quantitative Genetics.' 4th edn. (Prentice Hall: Harlow, UK).

Farentinos RC, Capretta PJ, Kepner RE, Littlefield VM (1981) Selective herbivory in tassel-eared squirrels: role of monoterpenes in ponderosa pines chosen as feeding trees. *Science* **213**, 1273-1275.

Fedderwitz F, Björklund N, Anngren R, Lindström A, Nordlander G (2020) Can methyl jasmonate treatment of conifer seedlings be used as a tool to stop height growth in nursery forest trees? *New Forests* **51**, 379-394.

Felicijan M, Novak M, Kraševec N, Urbanek Krajnc A (2015) Antioxidant defences of Norway spruce bark against bark beetles and its associated blue-stain fungus. *Agricultura* **12**, 9-18.

Ferrenberg S, Kane JM, Langenhan JM (2015) To grow or defend? Pine seedlings grow less but induce more defences when a key resource is limited. *Tree Physiology* **35**, 107-111.

Ferrenberg S, Mitton JB (2014) Smooth bark surfaces can defend trees against insect attack: resurrecting a 'slippery'hypothesis. *Functional Ecology* **28**, 837-845.

Ferrer-Gallego R, Hernández-Hierro JM, Rivas-Gonzalo JC, Escribano-Bailón MT (2011) Determination of phenolic compounds of grape skins during ripening by NIR spectroscopy. *LWT - Food Science and Technology* **44**, 847-853.

Fox H, Doron-Faigenboim A, Kelly G, Bourstein R, Attia Z, Zhou J, Moshe Y, Moshelion M, David-Schwartz R (2017) Transcriptome analysis of *Pinus halepensis* under drought stress and during recovery. *Tree Physiology* **38**, 423-441.

Franceschi VR, Krokene P, Christiansen E, Krekling T (2005) Anatomical and chemical defenses of conifer bark against bark beetles and other pests. *New Phytologist* **167**, 353-376.

Franich RA, Gadgil PD, Shain L (1983) Fungistaxic effects of *Pinus radiata* needle epicuticular fatty and resin acids on *Dothistroma pini*. *Physiological Plant Pathology* **23**, 183-195.

Fritze IM, Linden L, Freigang J, Auerbach G, Huber R, Steinbacher S (2004) The crystal structures of *Zea mays* and Arabidopsis 4-hydroxyphenylpyruvate dioxygenase. *Plant Physiology* **134**, 1388-1400.

Frost CJ, Hunter MD (2008) Herbivore-induced shifts in carbon and nitrogen allocation in red oak seedlings. *New Phytologist* **178**, 835-845.

Gamal El-Dien O, Ratcliffe B, Klápště J, Porth I, Chen C, El-Kassaby YA (2016) Implementation of the realized genomic relationship matrix to open-pollinated white spruce family testing for disentangling additive from nonadditive genetic effects. *Genes, Genomes, Genetics* **6**, 743-753.

Gamal El-Dien O, Ratcliffe B, Klápště J, Porth I, Chen C, El-Kassaby YA (2018) Multienvironment genomic variance decomposition analysis of open-pollinated Interior spruce (*Picea glauca × engelmannii*). *Molecular Breeding* **38**, 26.

Ganthaler A, Stöggl W, Kranner I, Mayr S (2017) Foliar phenolic compounds in Norway spruce with varying susceptibility to *Chrysomyxa rhododendri:* analyses of seasonal and infection-induced accumulation patterns. *Frontiers in Plant Science* **8**.

Gao H, Christensen OF, Madsen P, Nielsen US, Zhang Y, Lund MS, Su G (2012) Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. *Genetics Selection Evolution* **44**, 8.

Gao N, Martini JWR, Zhang Z, Yuan X, Zhang H, Simianer H, Li J (2017) Incorporating gene annotation into genomic prediction of complex phenotypes. *Genetics* **207**, 489-501.

Gapare WJ, Ivković M, Dutkowski GW, Spencer DJ, Buxton P, Wu HX (2012) Genetic parameters and provenance variation of *Pinus radiata* D. Don. 'Eldridge collection' in Australia 1: growth and form traits. *Tree Genetics & Genomes* **8**, 391-407.

Gershenzon J (1994) Metabolic costs of terpenoid accumulation in higher plants. *Journal of Chemical Ecology* **20**, 1281-1328.

Giangiacomo R (2006) Study of water–sugar interactions at increasing sugar concentration by NIR spectroscopy. *Food Chemistry* **96**, 371-379.

Gianola D, de los Campos G, Hill WG, Manfredi E, Fernando R (2009) Additive genetic variability and the Bayesian alphabet. *Genetics* **183**, 347-363.

Gianola D, Schön C-C (2016) Cross-validation without doing cross-validation in genome-enabled prediction. *Genes, Genomes, Genetics* **6**, 3107-3128.

Gianoli E, Salgado-Luarte C (2017) Tolerance to herbivory and the resource availability hypothesis. *Biology Letters* **13**, 20170120.

Gilbert L, Alhagdow M*, et al.* (2009) GDP-$_D$-mannose 3,5-epimerase (GME) plays a key role at the intersection of ascorbate and non-cellulosic cell-wall biosynthesis in tomato. *The Plant Journal* **60**, 499-508.

Gill R, Webber J, Peace A (2000) The economic implications of deer damage - a review of current evidence. Forest Research Agency, Wrecclesham, Surrey.

Gill RMA (1992) A review of damage by mammals in north temperate forests: 1. Deer. *Forestry* **65**, 145-169.

Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R (2015) 'ASReml user guide release 4.1 structural specification.' 4.1 edn. (VSN International: Hemel Hempstead, UK).

Goddard M (2009) Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica* **136**, 245-257.

González-Camacho JM, Ornella L, Pérez-Rodríguez P, Gianola D, Dreisigacker S, Crossa J (2018) Applications of machine learning methods to genomic selection in breeding wheat for rust resistance. *The Plant Genome* **11**.

Goodsman DW, Lusebrink I, Landhäusser SM, Erbilgin N, Lieffers VJ (2013) Variation in carbon availability, defense chemistry and susceptibility to fungal invasion along the stems of mature trees. *New Phytologist* **197**, 586-594.

Gould N, Reglinski T, Northcott GL, Spiers M, Taylor JT (2009) Physiological and biochemical responses in *Pinus radiata* seedlings associated with methyl jasmonate-induced resistance to *Diplodia pinea*. *Physiological and Molecular Plant Pathology* **74**, 121-128.

Gould N, Reglinski T, Spiers M, Taylor JT (2008) Physiological trade-offs associated with methyl jasmonate-induced resistance in *Pinus radiata*. *Canadian Journal of Forest Research* **38**, 677-684.

Gowen AA, Downey G, Esquerre C, O'Donnell CP (2011) Preventing over-fitting in PLS calibration models of near-infrared (NIR) spectroscopy data using regression coefficients. *Journal of Chemometrics* **25**, 375-381.

Grabherr MG, Haas BJ*, et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**, 644-652.

Grabska J, Ishigaki M, Beć KB, Wójcik MJ, Ozaki Y (2017) Correlations between structure and near-infrared spectra of saturated and unsaturated carboxylic acids. Insight from anharmonic density functional theory calculations. *The Journal of Physical Chemistry A* **121**, 3437-3451.

Graham NJ, Suontama M, Pleasants T, Li Y, Bader MK-F, Klápště J, Dungey HS, Williams NM (2018) Assessing the genetic variation of tolerance to red needle cast in a *Pinus radiata* breeding population. *Tree Genetics & Genomes* **14**, 55.

Grattapaglia D, Resende MD (2011) Genomic selection in forest tree breeding. *Tree Genetics & Genomes* **7**, 241-255.

Gromski PS, Muhamadali H, Ellis DI, Xu Y, Correa E, Turner ML, Goodacre R (2015) A tutorial review: Metabolomics and partial least squares-discriminant analysis–a marriage of convenience or a shotgun wedding. *Analytica chimica acta* **879**, 10-23.

Guillet-Claude C, Isabel N, Pelgas B, Bousquet J (2004) The evolutionary implications of *knox-I* gene duplications in conifers: correlated evidence from phylogeny, gene mapping, and analysis of functional divergence. *Molecular Biology and Evolution* **21**, 2232-2245.

Guo Z, Magwire MM, Basten CJ, Xu Z, Wang D (2016) Evaluation of the utility of gene expression and metabolic information for genomic prediction in maize. *Theoretical and Applied Genetics* **129**, 2413-2427.

Haavik LJ, Dodds KJ, Allison JD (2017) Suitability of eastern pines for oviposition and survival of *Sirex noctilio* F. *PLoS One* **12**, e0174532-e0174532.

Hallauer AR, Carena MJ, Miranda Filho JB (2010) 'Quantitative genetics in maize breeding.' (Springer: New York).

Hammerbacher A, Ralph SG, Bohlmann J, Fenning TM, Gershenzon J, Schmidt A (2011) Biosynthesis of the Major Tetrahydroxystilbenes in Spruce, Astringin and Isorhapontin, Proceeds via Resveratrol and Is Enhanced by Fungal Infection. *Plant Physiology* **157**, 876.

Hansson L, Gref R, Lundren L, Theander O (1986) Susceptibility to vole attacks due to bark phenols and terpenes in *Pinus contorta* provenances introduced into Sweden. *Journal of Chemical Ecology* **12**, 1569-1578.

Harris PA, Nelson S, Carslake HB, Argo CM, Wolf R, Fabri FB, Brolsma KM, van Oostrum MJ, Ellis AD (2018) Comparison of NIRS and wet chemistry methods for the nutritional analysis of haylages for horses. *Journal of Equine Veterinary Science* **71**, 13-20.

Hartmann H, Trumbore S (2016) Understanding the roles of nonstructural carbohydrates in forest trees – from what we can measure to what we want to know. *New Phytologist* **211**, 386-403.

Haskard KA (2007) An anisotropic Matern spatial covariance model: REML estimation and properties. Ph.D. Thesis, University of Adelaide, Adelaide

Haviola S, Saloniemi I, Ossipov V, Haukioja E (2006) Additive genetic variation of secondary and primary metabolites in mountain birch. *Oikos* **112**, 382-391.

Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009) Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science* **92**, 433-443.

Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819-1829.

Henery ML, Henson M, Wallis IR, Stone C, Foley WJ (2008) Predicting crown damage to *Eucalyptus grandis* by *Paropsis atomaria* with direct and indirect measures of leaf composition. *Forest Ecology and Management* **255**, 3642-3651.

Herms DA, Mattson WJ (1992) The dilemma of plants: to grow or defend. *The Quarterly Review of Biology* **67**, 283-335.

Hernandez-Escribano L, Iturritxa E, Aragonés A, Mesanza N, Berbegal M, Raposo R, Elvira-Recuenco M (2018) Root infection of canker pathogens, *Fusarium circinatum* and *Diplodia sapinea,* in asymptomatic trees in *Pinus radiata* and *Pinus pinaster* plantations. *Forests* **9**, 128.

Hernandez LD, Vierling E (1993) Expression of low molecular weight heat-shock proteins under field conditions. *Plant Physiology* **101**, 1209-1216.

Hijazi M, Roujol D, Nguyen-Kim H, del Rocio Cisneros Castillo L, Saland E, Jamet E, Albenne C (2014) Arabinogalactan protein 31 (AGP31), a putative network-forming protein in *Arabidopsis thaliana* cell walls? *Annals of Botany* **114**, 1087-1097.

Hood S, Sala A (2015) Ponderosa pine resin defenses and growth: metrics matter. *Tree Physiology* **35**, 1223-1235.

Houben M, Van de Poel B (2019) 1-aminocyclopropane-1-carboxylic acid oxidase (ACO): the enzyme that makes the plant hormone ethylene. *Frontiers in Plant Science* **10**.

Huber DPW, Philippe RN, Madilao LL, Sturrock RN, Bohlmann J (2005) Changes in anatomy and terpene chemistry in roots of Douglas-fir seedlings following treatment with methyl jasmonate. *Tree Physiology* **25**, 1075-1083.

Hudgins JW, Christiansen E, Franceschi VR (2004) Induction of anatomically based defense responses in stems of diverse conifers by methyl jasmonate: a phylogenetic perspective. *Tree Physiology* **24**, 251-264.

Hudgins JW, Franceschi VR (2004) Methyl jasmonate-induced ethylene production is responsible for conifer phloem defense responses and reprogramming of stem cambial zone for traumatic resin duct formation. *Plant Physiology* **135**, 2134-2149.

Hughes A, Buitenhuis A (2010) Reduced variance of gene expression at numerous loci in a population of chickens selected for high feather pecking. *Poultry Science* **89**, 1858-1869.

Hummelbrunner LA, Isman MB (2001) Acute, sublethal, antifeedant, and synergistic effects of monoterpenoid essential oil compounds on the tobacco cutworm, *Spodoptera litura* (Lep., Noctuidae). *Journal of Agricultural and Food Chemistry* **49**, 715-720.

Huot B, Yao J, Montgomery BL, He SY (2014) Growth–defense tradeoffs in plants: a balancing act to optimize fitness. *Molecular Plant* **7**, 1267-1287.

Hutin C, Nussaume L, Moise N, Moya I, Kloppstech K, Havaux M (2003) Early light-induced proteins protect *Arabidopsis* from photooxidative stress. *Proceedings of the National Academy of Sciences* **100**, 4921-4926.

Iason GR, O'Reilly-Wapstra JM, Brewer MJ, Summers RW, Moore BD (2011) Do multiple herbivores maintain chemical diversity of Scots pine monoterpenes? *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 1337-1345.

Idaghdour Y, Awadalla P (2013) Exploiting gene expression variation to capture gene-environment interactions for disease. *Frontiers in Genetics* **3**.

Iijima H, Nagaike T (2015) Susceptible conditions for debarking by deer in subalpine coniferous forests in central Japan. *Forest Ecosystems* **2**, 33.

Ilse LM, Hellgren EC (2007) Indirect interactions among dendrophages: Porcupines predispose pinyon pines to bark beetle attack. *Forest Ecology and Management* **242**, 217-226.

Invernizzi C, Rovetta T, Licchelli M, Malagodi M (2018) Mid and near-infrared reflection spectral database of natural organic materials in the cultural heritage field. *International Journal of Analytical Chemistry* **2018**.

Isik F, Holland J, Maltecca C (2017) 'Genetic data analysis for plant and animal breeding.' (Springer International: Cham, Switzerland).

Ismael A, Suontama M, Klápště J, Kennedy S, Graham N, Telfer E, Dungey H (2020) Indication of quantitative multiple disease resistance to foliar pathogens in *Pinus radiata* D.Don in New Zealand. *Frontiers in Plant Science* **11**.

Ivey CT, Carr DE, Eubanks MD (2009) Genetic variation and constraints on the evolution of defense against spittlebug (*Philaenus spumarius*) herbivory in *Mimulus guttatus*. *Heredity* **102**, 303-311.

Ivković M, Baltunis B, Gapare W, Sasse J, Dutkowski G, Elms S, Wu H (2010a) Breeding against dothistroma needle blight of radiata pine in Australia. *Canadian Journal of Forest Research* **40**, 1653-1660.

Ivković M, Gapare WJ, Wharton T, Jovanovic T, Elms S, McRae TA, Wu HX (2010b) Risks affecting breeding objectives for radiata pine in Australia. *Australian Forestry* **73**, 265-278.

Iwata H, Hayashi T, Tsumura Y (2011) Prospects for genomic selection in conifer breeding: a simulation study of *Cryptomeria japonica*. *Tree Genetics & Genomes* **7**, 747-758.

Jagodzik P, Tajdel-Zielinska M, Ciesla A, Marczak M, Ludwikow A (2018) Mitogen-activated protein kinase cascades in plant hormone signaling. *Frontiers in Plant Science* **9**.

Jannink J-L, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. *Briefings in Functional Genomics* **9**, 166-177.

Jia Y, Jannink J-L (2012) Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics* **192**, 1513-1522.

Jiang Z, Ueda H, Kitahara M, Imaki H (2005) Bark stripping by sika deer on veitch fir related to stand age, bark nutrition, and season in northern Mount Fuji district, central Japan. *Journal of Forest Research* **10**, 359-365.

Johnson MTJ (2011) Evolutionary ecology of plant defences against herbivores. *Functional Ecology* **25**, 305-311.

Jombart T, Collins C (2015) A tutorial for discriminant analysis of principal components (DAPC) using *adegenet* 2.0.0. In. ' (Imperial College London, MRC Centre for Outbreak Analysis and Modelling: London)

Jones TH, Potts BM, Vaillancourt RE, Davies NW (2002) Genetic resistance of *Eucalyptus globulus* to autumn gum moth defoliation and the role of cuticular waxes. *Canadian Journal of Forest Research* **32**, 1961-1969.

Jorge TF, Rodrigues JA, Caldana C, Schmidt R, van Dongen JT, Thomas‐Oates J, António C (2016) Mass spectrometry‐based plant metabolomics: Metabolite responses to abiotic stress. *Mass Spectrometry Reviews* **35**, 620-649.

Joshi V, Jander G (2009) *Arabidopsis* methionine γ-lyase is regulated according to isoleucine biosynthesis needs but plays a subordinate role to threonine deaminase. *Plant Physiology* **151**, 367-378.

Juliani HR, Kapteyn J, Jones D, Koroch AR, Wang M, Charles D, Simon JE (2006) Application of near‑infrared spectroscopy in quality control and determination of adulteration of African essential oils. *Phytochemical Analysis* **17**, 121-128.

Kachroo A, Kachroo P (2009) Fatty acid–derived signals in plant defense. *Annual Review of Phytopathology* **47**, 153-176.

Kamada-Nobusada T, Hayashi M, Fukazawa M, Sakakibara H, Nishimura M (2008) A putative peroxisomal polyamine oxidase, AtPAO4, is involved in polyamine catabolism in *Arabidopsis thaliana*. *Plant and Cell Physiology* **49**, 1272-1282.

Kānberga-Siliņa K, Rauda E, Šķipars V, Vivian-Smith A, Yakovlev I, Krivmane B, Šņepste I, Ruņģis D (2017) Transcriptomic response to methyl jasmonate treatment of Scots pine (*Pinus sylvestris*) seedlings. *Environmental and Experimental Biology* **15**, 257-274.

Kant MR, Jonckheere W*, et al.* (2015) Mechanisms and ecological consequences of plant defence induction and suppression in herbivore communities. *Annals of Botany* **115**, 1015-1051.

Kapilan R, Vaziri M, Zwiazek JJ (2018) Regulation of aquaporins in plants under stress. *Biological Research* **51**, 4.

Kaplan F, Guy CL (2004) *β*-amylase induction and the protective role of maltose during temperature shock. *Plant Physiology* **135**, 1674-1684.

Karaman E, Lund MS, Su G (2020) Multi-trait single-step genomic prediction accounting for heterogeneous (co)variances over the genome. *Heredity* **124**, 274-287.

Karban R, Myers JH (1989) Induced plant responses to herbivory. *Annual Review of Ecology and Systematics* **20**, 331-348.

Kärkkäinen HP, Sillanpää MJ (2012) Back to basics for Bayesian model building in genomic selection. *Genetics* **191**, 969-987.

Keator G (2002) 'Introduction to Trees of the San Francisco Bay Region.' (University of California Press: Berkeley, CA).

Keefover-Ring K, Trowbridge A, Mason CJ, Raffa KF (2016) Rapid induction of multiple terpenoid groups by ponderosa pine in response to bark beetle-associated fungi. *Journal of Chemical Ecology* **42**, 1-12.

Keeling CI, Bohlmann J (2006) Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. *New Phytologist* **170**, 657-675.

Kempel A, Schädler M, Chrobock T, Fischer M, van Kleunen M (2011) Tradeoffs associated with constitutive and induced plant resistance against herbivory. *Proceedings of the National Academy of Sciences* **108**, 5685-5689.

Kenward RE, Dutton JCF, Parish T, Doyle FIB, Walls SS, Robertson PA (1996) Damage by grey squirrels. I. Bark-stripping correlates and treatment. *Quarterly Journal of Forestry* **90**, 135-142.

Kim H-Y (2013) Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative dentistry & endodontics* **38**, 52-54.

Kim J, Darlington A, Salvador M, Utrilla J, Jiménez JI (2020) Trade-offs between gene expression, growth and phenotypic diversity in microbial populations. *Current Opinion in Biotechnology* **62**, 29-37.

Kim Y-K, Kim S, Shin Y-j, Hur Y-S, Kim W-Y, Lee M-S, Cheon C-I, Verma DPS (2014) Ribosomal protein S6, a target of rapamycin, is involved in the regulation of rRNA genes by possible epigenetic changes in *Arabidopsis*. *Journal of Biological Chemistry* **289**, 3901-3912.

Kimball BA, Nolte DL, Engeman RM, Johnston JJ, Stermitz FR (1998) Chemically mediated foraging preference of black bears (*Ursus americanus*). *Journal of Mammalogy* **79**, 448-456.

Kimball BA, Pfund F, Gourley M, Griffin DL, Russell JH (2011) Silvicultural attempts to induce browse resistance in conifer seedlings. *International Journal of Forestry Research* **2011**, 7.

King JN, Alfaro RI (2009) Developing Sitka spruce populations for resistance to the white pine weevil: summary of research and breeding program. Ministry of Forests and Range, British Columbia, Victoria, BC

King JN, Alfaro RI, Lopez MG, Van Akker L (2011) Resistance of Sitka spruce (*Picea sitchensis* (Bong.) Carr.) to white pine weevil (*Pissodes strobi* Peck): characterizing the bark defence mechanisms of resistant populations. *Forestry* **84**, 83-91.

King JN, Yanchuk AD, Kiss GK, Alfaro RI (1997) Genetic and phenotypic relationships between weevil (*Pissodes strobi*) resistance and height growth in spruce populations of British Columbia. *Canadian Journal of Forest Research* **27**, 732-739.

Klápště J, Dungey HS, Graham NJ, Telfer EJ (2020a) Effect of trait's expression level on single-step genomic evaluation of resistance to Dothistroma needle blight. *BMC Plant Biology* **20**, 205.

Klápště J, Lstibůrek M, El-Kassaby YA (2014) Estimates of genetic parameters and breeding values from western larch open-pollinated families using marker-based relationship. *Tree Genetics & Genomes* **10**, 241-249.

Klápště J, Meason D, Dungey HS, Telfer EJ, Silcock P, Rapley S (2020b) Genotype-by-environment interaction in coast redwood outside natural distribution - search for environmental cues. *BMC Genetics* **21**, 15.

Klápště J, Suontama M, Dungey HS, Telfer EJ, Graham NJ, Low CB, Stovold GT (2018) Effect of hidden relatedness on single-step genetic evaluation in an advanced open-pollinated breeding program. *Journal of Heredity* **109**, 802-810.

Klich D (2017) Selective bark stripping of various tree species by Polish horses in relation to bark detachability. *Forest Ecology and Management* **384**, 65-71.

Kliebenstein DJ (2014) Quantitative genetics and genomics of plant resistance to insects. *Annual Plant Reviews* **47**, 235-262.

Kobashikawa S, Koike S (2016) Spatiotemporal factors affecting bark stripping of conifer trees by Asiatic black bears (*Ursus thibetanus*) in Japan. *Forest Ecology and Management* **380**, 100-106.

Kobashikawa S, Trentin B, Koike S (2019) The benefit of wrapping trees in biodegradable material netting to protect against bark stripping by bears extends to surrounding stands. *Forest Ecology and Management* **437**, 134-138.

Kolosova N (2010) Transcriptome analysis of conifer defense against bark beetle-associated blue-stain fungi and white pine weevil. University of British Columbia, Vancouver, BC, Canada

Korth KL (2003) Profiling the response of plants to herbivorous insects. *Genome Biology* **4**, 221.

Koul A, Becchio C, Cavallo A (2018) Cross-Validation Approaches for Replicability in Psychology. *Frontiers in Psychology* **9**.

Kovalchuk A, Raffaello T, Jaber E, Keriö S, Ghimire R, Lorenz WW, Dean JFD, Holopainen JK, Asiegbu FO (2015) Activation of defence pathways in Scots pine bark after feeding by pine weevil (*Hylobius abietis*). *BMC Genomics* **16**, 352-352.

Kovalchuk A, Zeng Z*, et al.* (2019) Dual RNA-seq analysis provides new insights into interactions between Norway spruce and necrotrophic pathogen *Heterobasidion annosum s.l*. *BMC Plant Biology* **19**, 2.

Krolikowski KA, Victor JL, Wagler TN, Lolle SJ, Pruitt RE (2003) Isolation and characterization of the *Arabidopsis* organ fusion gene *HOTHEAD*. *The Plant Journal* **35**, 501-511.

Kuiters AT, van der Sluijs LAM, Wytema GA (2006) Selective bark-stripping of beech, *Fagus sylvatica*, by free-ranging horses. *Forest Ecology and Management* **222**, 1-8.

Kurek T, Todys J, Pazdrowski W, Szymański M, Łukowski A (2019) Intensity of stripping and sugar content in the bark and the bast of European beech (*Fagus sylvatica*). *Open Life Sciences* **14**, 19-28.

Lamara M, Parent GJ, Giguère I, Beaulieu J, Bousquet J, MacKay JJ (2018) Association genetics of acetophenone defence against spruce budworm in mature white spruce. *BMC Plant Biology* **18**, 231.

Latchman DS (2008) Regulation of transcription factor synthesis. In 'Eukaryotic Transcription Factors. ' Fifth edn. (Ed. DS Latchman) pp. 271-311. (Academic Press: New York).

Lawrence R, Potts BM, Whitham TG (2003) Relative importance of plant ontogeny, host genetic variation, and leaf age for a common herbivore. *Ecology* **84**, 1171-1178.

Lê S, Josse J, Husson F (2008) FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software* **25**, 1-18.

Lebouvier N, Hue T, Hnawia E, Lesaffre L, Menut C, Nour M (2013) Acaricidal activity of essential oils from five endemic conifers of New Caledonia on the cattle tick *Rhipicephalus* (*Boophilus*) *microplus*. *Parasitology Research* **112**, 1379-1384.

Lecoy J, García-Gil MR (2020) Evolutionary analysis of LP3 gene family in conifers: an ASR homolog. In. Vol. This article is a preprint and has not been certified by peer review'. pp. 2020.03.27.011197: bioRxiv)

Lee SY, Mediani A, Maulidiani M, Khatib A, Ismail IS, Zawawi N, Abas F (2018) Comparison of partial least squares and random forests for evaluating relationship between phenolics and bioactivities of Neptunia oleracea. *Journal of the Science of Food and Agriculture* **98**, 240-252.

Legarra A, Aguilar I, Misztal I (2009) A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science* **92**, 4656-4663.

Leidinger J, Seibold S, Weisser WW, Lange M, Schall P, Türke M, Gossner MM (2019) Effects of forest management on herbivorous insects in temperate Europe. *Forest Ecology and Management* **437**, 232-245.

Leimu R, Koricheva J (2006) A meta-analysis of genetic correlations between plant resistances to multiple enemies. *The American Naturalist* **168**, E15-E37.

Lenz PRN, Nadeau S, Mottet MJ, Perron M, Isabel N, Beaulieu J, Bousquet J (2020) Multi‐trait genomic selection for weevil resistance, growth, and wood quality in Norway spruce. *Evolutionary Applications* **13**, 76-94.

Léon-Kloosterziel KM, Verhagen BWM, Keurentjes JJB, Van Pelt JA, Rep M, Van Loon LC, Pieterse CMJ (2005) Colonization of the *Arabidopsis* rhizosphere by fluorescent *Pseudomonas* spp. activates a root-specific, ethylene-responsive *PR-5* gene in the vascular bundle. *Plant Molecular Biology* **57**, 731-748.

Lewinsohn E, Gijzen M, Croteau R (1991) Defense mechanisms of conifers: differences in constitutive and wound-induced monoterpene biosynthesis among species. *Plant Physiology* **96**, 44.

Li C-Y, Weiss D, Goldschmidt EE (2003) Girdling affects carbohydrate-related gene expression in leaves, bark and roots of alternate-bearing citrus trees. *Annals of Botany* **92**, 137-143.

Li N, Zhao M*, et al.* (2017) A novel soybean dirigent gene *GmDIR22* contributes to promotion of lignan biosynthesis and enhances resistance to *Phytophthora sojae*. *Frontiers in Plant Science* **8**, 1185.

Li R, Wang W, Wang W, Li F, Wang Q, Xu Y, Wang S (2015a) Overexpression of a cysteine proteinase inhibitor gene from *Jatropha curcas* confers enhanced tolerance to salinity stress. *Electronic Journal of Biotechnology* **18**, 368-375.

Li X, Yang X, Wu HX (2013) Transcriptome profiling of radiata pine branches reveals new insights into reaction wood formation with implications in plant gravitropism. *BMC Genomics* **14**, 768.

Li Y, Dungey HS (2018) Expected benefit of genomic selection over forward selection in conifer breeding and deployment. *PLoS One* **13**, e0208232-e0208232.

Li Y, Dungey HS, Carson M, Carson S (2018) Genotype by environment interaction for growth and *Dothistroma* resistance and clonal connectivity between environments in radiata pine in New Zealand and Australia. *PLoS One* **13**, e0205402.

Li Y, Telfer E, Wilcox PL (2015b) New Zealand forestry enters the genomics era—applications of genomics in tree breeding. *New Zealand Journal of Forestry* **60**, 23-25.

Li Y, Wilcox P, Telfer E, Graham N, Stanbra L (2016) Association of single nucleotide polymorphisms with form traits in three New Zealand populations of radiata pine in the presence of genotype by environment interactions. *Tree Genetics & Genomes* **12**, 63.

Li Z, Gao N, Martini JWR, Simianer H (2019) Integrating gene expression data into genomic prediction. *Frontiers in Genetics* **10**.

Ligot G, Gheysen T, Lehaire F, Hébert J, Licoppe A, Lejeune P, Brostaux Y (2013) Modeling recent bark stripping by red deer (*Cervus elaphus*) in South Belgium coniferous stands. *Annals of Forest Science* **70**, 309-318.

Lindroth RL, Batzli GO (1984) Plant phenolics as chemical defenses: effects of natural phenolics on survival and growth of prairie voles (*Microtus ochrogaster*). *Journal of Chemical Ecology* **10**, 229-244.

Litvak ME, Monson RK (1998) Patterns of induced and constitutive monoterpene production in conifer needles in relation to insect herbivory. *Oecologia* **114**, 531-540.

Liu F, Zhang X, Lu C, Zeng X, Li Y, Fu D, Wu G (2015a) Non-specific lipid transfer proteins in plants: presenting new advances and an integrated functional analysis. *Journal of Experimental Botany* **66**, 5663-5681.

Liu J-J, Ekramoddoullah AKM (2006) The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses. *Physiological and Molecular Plant Pathology* **68**, 3-13.

Liu J-J, Ekramoddoullah AKM, Piggott N, Zamani A (2005) Molecular cloning of a pathogen/wound-inducible PR10 promoter from *Pinus monticola* and characterization in transgenic *Arabidopsis* plants. *Planta* **221**, 159-169.

Liu J-J, Sturrock RN, Benton R (2013a) Transcriptome analysis of *Pinus monticola* primary needles by RNA-seq provides novel insight into host resistance to *Cronartium ribicola*. *BMC Genomics* **14**, 884.

Liu J-J, Williams H, Li XR, Schoettle AW, Sniezko RA, Murray M, Zamany A, Roke G, Chen H (2017a) Profiling methyl jasmonate-responsive transcriptome for understanding induced systemic resistance in whitebark pine (*Pinus albicaulis*). *Plant Molecular Biology* **95**, 359-374.

Liu P-L, Du L, Huang Y, Gao S-M, Yu M (2017b) Origin and diversification of leucine-rich repeat receptor-like protein kinase (*LRR-RLK*) genes in plants. *BMC Evolutionary Biology* **17**, 47.

Liu Q, Wei Y, Xu L, Hao Y, Chen X, Zhou Z (2017c) Transcriptomic profiling reveals differentially expressed genes associated with pine wood nematode resistance in masson pine (*Pinus massoniana* Lamb.). *Scientific Reports* **7**, 4693.

Liu R, Xu Y-H, *et al.* (2013b) Light-harvesting chlorophyll *a/b*-binding proteins, positively involved in abscisic acid signalling, require a transcription repressor, WRKY40, to balance their function. *Journal of Experimental Botany* **64**, 5443-5456.

Liu X-Q, Rong J-Y, Liu X-Y (2008) Best linear unbiased prediction for linear combinations in general mixed linear models. *Journal of Multivariate Analysis* **99**, 1503-1517.

Liu Y, Zhang W, Xie L, Liu H, Gong G, Zhu B, Hu Y (2015b) *Acthi*, a thiazole biosynthesis enzyme, is essential for thiamine biosynthesis and CPC production in *Acremonium chrysogenum*. *Microbial Cell Factories* **14**, 50.

Lombardero MJ, Ayres MP, Bonello P, Cipollini D, Herms DA (2016) Effects of defoliation and site quality on growth and defenses of *Pinus pinaster* and *P. radiata*. *Forest Ecology and Management* **382**, 39-50.

Lombardero MJ, Ayres MP, Lorio Jr PL, Ruel JJ (2000) Environmental effects on constitutive and inducible resin defences of *Pinus taeda*. *Ecology Letters* **3**, 329-339.

Lombardero MJ, Pereira-Espinel J, Ayres MP (2013) Foliar terpene chemistry of *Pinus pinaster* and *P. radiata* responds differently to methyl jasmonate and feeding by larvae of the pine processionary moth. *Forest Ecology and Management* **310**, 935-943.

López-Goldar X, Villari C, Bonello P, Borg-Karlson A-K, Grivet D, Zas R, Sampedro L (2018) Inducibility of plant secondary metabolites in the stem predicts genetic variation in resistance against a key insect herbivore in maritime pine. *Frontiers in Plant Science* **9**.

Lorenzo O, Piqueras R, Sánchez-Serrano JJ, Solano R (2003) Ethylene Response Factor1 integrates signals from ethylene and jasmonate pathways in plant defense. *The Plant Cell* **15**, 165-178.

Lorio PL, Jr. (1986) Growth-differentiation balance: a basis for understanding southern pine beetle-tree interactions. *Forest Ecology and Management* **14**, 259-273.

Lothier J, Gaufichon L*, et al.* (2011) The cytosolic glutamine synthetase GLN1;2 plays a role in the control of plant growth and ammonium homeostasis in *Arabidopsis* rosettes when nitrate supply is not limiting. *Journal of Experimental Botany* **62**, 1375-1390.

Lourenco D, Tsuruta S*, et al.* (2015) Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. *Journal of animal science* **93**, 2653-2662.

Loyola R, Herrera D*, et al.* (2016) The photomorphogenic factors UV-B RECEPTOR 1, ELONGATED HYPOCOTYL $_5$, and HY$_5$ HOMOLOGUE are part of the UV-B signalling pathway in grapevine and mediate flavonol accumulation in response to the environment. *Journal of Experimental Botany* **67**, 5429-5445.

Lu Y (2016) Identification and roles of photosystem II assembly, stability, and repair factors in *Arabidopsis*. *Frontiers in Plant Science* **7**, 168.

Lundborg L (2016) Effects of methyl jasmonate on chemical defenses of conifer seedlings in relation to feeding by *Hylobius abietis*. Doctoral thesis Thesis, KTH Royal Institute of Technology, Stockholm

Lundborg L, Sampedro L, Borg-Karlson A-K, Zas R (2019) Effects of methyl jasmonate on the concentration of volatile terpenes in tissues of Maritime pine and Monterey pine and its relation to pine weevil feeding. *Trees* **33**, 53-62.

Ma L, Peng Y, Pei Y, Zeng J, Shen H, Cao J, Qiao Y, Wu Z (2019) Systematic discovery about NIR spectral assignment from chemical structural property to natural chemical compounds. *Scientific Reports* **9**, 9503.

Machado RAR, Baldwin IT, Erb M (2017) Herbivory-induced jasmonates constrain plant sugar accumulation and growth by antagonizing gibberellin signaling and not by promoting secondary metabolite production. *New Phytologist* **215**, 803-812.

Madan SS, Wasewar KL, Pandharipande SL (2016) Modeling the adsorption of benzeneacetic acid on CaO_2 nanoparticles using artificial neural network. *Resource-Efficient Technologies* **2**, S53-S62.

Malley DF, Martin PD, Ben-Dor E (2004) Application in analysis of soils. In 'Near-Infrared Spectroscopy in Agriculture. Vol. 44. (Eds CA Roberts, J Workman, Jr. and JB Reeves, III) pp. 729-784. (ASA, CSSA, SSSA Books: Madison, WI).

Månsson J, Jarnemo A (2013) Bark-stripping on Norway spruce by red deer in Sweden: level of damage and relation to tree characteristics. *Scandinavian Journal of Forest Research* **28**, 117-125.

Manuell AL, Quispe J, Mayfield SP (2007) Structure of the chloroplast ribosome: novel domains for translation regulation. *PLoS Biology* **5**, e209.

Maron JL, Agrawal AA, Schemske DW (2019) Plant-herbivore coevolution and plant speciation. *Ecology* **100**, e02704.

Marshall SDG, Putterill JJ, Plummer KM, Newcomb RD (2003) The carboxylesterase gene family from *Arabidopsis thaliana*. *Journal of Molecular Evolution* **57**, 487-500.

Martin D, Tholl D, Gershenzon J, Bohlmann J (2002) Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. *Plant Physiology* **129**, 1003-1018.

Martini JWR, Schrauf MF*, et al.* (2018) The effect of the H^{-1} scaling factors τ and ω on the structure of **H** in the single-step procedure. *Genetics Selection Evolution* **50**, 16-16.

Mayle BA, Proudfoot J, Poole J (2009) Influence of tree size and dominance on incidence of bark stripping by grey squirrels to oak and impact on tree growth. *Forestry* **82**, 431-444.

McDonald JH (2009) 'Handbook of biological statistics.' 2nd edn. (Sparky House Publishing Baltimore).

McDonald PM, Laacke RJ (1990) *Pinus radiata* D. Don, Monterey pine. *Silvics of North America* **1**, 433-441.

McKey D (1974) Adaptive patterns in alkaloid physiology. *The American Naturalist* **108**, 305-320.

McKey D (1979) The distribution of secondary compounds within plants. In 'Herbivores: Their interaction with secondary plant metabolites. ' 1st edn. (Eds GA Rosenthal, DH Jansen and SW Applebaum) pp. 56–133. (Academic Press: New York).

McLean JP, Jin G, Brennan M, Nieuwoudt MK, Harris PJ (2014) Using NIR and ATR-FTIR spectroscopy to rapidly detect compression wood in *Pinus radiata*. *Canadian Journal of Forest Research* **44**, 820-830.

Mead DJ (2013) 'Sustainable management of *Pinus radiata* plantations.' (Food and Agriculture Organization of the United Nations (FAO): Rome).

Men L, Yan S, Liu G (2013) *De novo* characterization of *Larix gmelinii* (Rupr.) Rupr. transcriptome and analysis of its gene expression induced by jasmonates. *BMC Genomics* **14**, 548.

Mendez KM, Reinke SN, Broadhurst DI (2019) A comparative evaluation of the generalised predictive ability of eight machine learning algorithms across ten clinical metabolomics data sets for binary classification. *Metabolomics* **15**, 150.

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819-1829.

Meuwissen THE, Indahl UG, Ødegård J (2017) Variable selection models for genomic selection using whole-genome sequence data and singular value decomposition. *Genetics Selection Evolution* **49**, 94.

Mhlongo MI, Piater LA, Madala NE, Labuschagne N, Dubery IA (2018) The chemistry of plant–microbe interactions in the rhizosphere and the potential for metabolomics to reveal signaling related to defense priming and induced systemic resistance. *Frontiers in Plant Science* **9**.

Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD (2019) PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research* **47**, D419-D426.

Miao H, Sun P, Liu W, Xu B, Jin Z (2014) Identification of genes encoding granule-bound starch synthase involved in amylose metabolism in banana fruit. *PLoS One* **9**, e88077.

Miller A, O'Reilly-Wapstra J, Potts B (2014) Genetic variation in bark stripping among *Pinus radiata*. National Centre for Future Forest Industries, Hobart.

Miller B, Madilao LL, Ralph S, Bohlmann J (2005) Insect-induced conifer defense. White pine weevil and methyl jasmonate induce traumatic resinosis, de novo formed volatile emissions, and accumulation of terpenoid synthase and putative octadecanoid pathway transcripts in Sitka spruce. *Plant Physiology* **137**, 369-382.

Miller DR, Borden JH (2000) Dose-dependent and species-specific responses of pine bark beetles (Coleoptera: Scolytidae) to monoterpenes in association with pheromones. *The Canadian Entomologist* **132**, 183-195.

Mir RR, Reynolds M, Pinto F, Khan MA, Bhat MA (2019) High-throughput phenotyping for crop improvement in the genomics era. *Plant Science* **282**, 60-72.

Misztal I, Aggrey SE, Muir WM (2013) Experiences with a single-step genome evaluation. *Poultry Science* **92**, 2530-2534.

Mitchell-Olds T, Willis JH, Goldstein DB (2007) Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics* **8**, 845-856.

Mochizuki S, Sugimoto K, Koeduka T, Matsui K (2016) Arabidopsis lipoxygenase 2 is essential for formation of green leaf volatiles and five-carbon volatiles. *FEBS Letters* **590**, 1017-1027.

Momen M, Mehrgardi AA, Sheikhi A, Kranis A, Tusell L, Morota G, Rosa GJM, Gianola D (2018) Predictive ability of genome-assisted statistical models under various forms of gene action. *Scientific Reports* **8**, 12309.

Moore BD, Andrew RL, Külheim C, Foley WJ (2014) Explaining intraspecific diversity in plant secondary metabolites in an ecological context. *New Phytologist* **201**, 733-750.

Moore BD, DeGabriel JL (2012) Integrating the effects of PSMs on vertebrate herbivores across spatial and temporal scales. In 'The ecology of plant secondary metabolites: from genes to global processes. Vol. 226. (Eds GR Iason, M Dicke and SE Hartley). (Cambridge University Press: Cambridge).

Moreira X, Lundborg L, Zas R, Carrillo-Gavilán A, Borg-Karlson A-K, Sampedro L (2013a) Inducibility of chemical defences by two chewing insect herbivores in pine trees is specific to targeted plant tissue, particular herbivore and defensive trait. *Phytochemistry* **94**, 113-122.

Moreira X, Mooney KA, Rasmann S, Petry WK, Carrillo-Gavilán A, Zas R, Sampedro L (2014) Trade-offs between constitutive and induced defences drive geographical and climatic clines in pine chemical defences. *Ecology Letters* **17**, 537-546.

Moreira X, Sampedro L, Zas R (2009) Defensive responses of *Pinus pinaster* seedlings to exogenous application of methyl jasmonate: Concentration effect and systemic response. *Environmental and Experimental Botany* **67**, 94-100.

Moreira X, Zas R, Sampedro L (2012a) Differential allocation of constitutive and induced chemical defenses in pine tree juveniles: a test of the optimal defense theory. *PLoS One* **7**, e34006.

Moreira X, Zas R, Sampedro L (2012b) Genetic variation and phenotypic plasticity of nutrient reallocation and increased fine root production as putative tolerance mechanisms inducible by methyl jasmonate in pine trees. *Journal of Ecology* **100**, 810-820.

Moreira X, Zas R, Sampedro L (2013b) Additive genetic variation in resistance traits of an exotic pine species: little evidence for constraints on evolution of resistance against native herbivores. *Heredity* **110**, 449-456.

Moreira X, Zas R, Solla A, Sampedro L (2015) Differentiation of persistent anatomical defensive structures is costly and determined by nutrient availability and genetic growth-defence constraints. *Tree Physiology* **35**, 112-123.

Morkunas I, Ratajczak L (2014) The role of sugar signaling in plant defense responses against fungal pathogens. *Acta Physiologiae Plantarum* **36**, 1607-1619.

Morris H, Jansen S (2016) Bark: its anatomy, function and diversity. In 'International Dendrology Society Yearbook 2016. pp. 51-61. (International Dendrology Society: Kingston, Herefordshire, UK).

Morris WF, Traw MB, Bergelson J (2006) On testing for a tradeoff between constitutive and induced resistance. *Oikos* **112**, 102-110.

Mortimer JC, Laohavisit A, Macpherson N, Webb A, Brownlee C, Battey NH, Davies JM (2008) Annexins: multifunctional components of growth and adaptation. *Journal of Experimental Botany* **59**, 533-544.

Mottet M-J, DeBlois J, Perron M (2015) High genetic variation and moderate to high values for genetic parameters of *Picea abies* resistance to *Pissodes strobi*. *Tree Genetics & Genomes* **11**, 58.

Mousseau TA, Roff DA (1987) Natural selection and the heritability of fitness components. *Heredity* **59**, 181-197.

Mumm R, Hilker M (2006) Direct and indirect chemical defence of pine against folivorous insects. *Trends in Plant Science* **11**, 351-358.

Muranty H, Troggio M*, et al.* (2015) Accuracy and responses of genomic selection on key traits in apple breeding. *Horticulture Research* **2**, 15060.

Murphy G, Cown D (2015) Within-tree, between-tree, and geospatial variation in estimated *Pinus radiata* bark volume and weight in New Zealand. *New Zealand Journal of Forestry Science* **45**, 18.

Nagaike T (2019) Effects of heavy, repeated bark stripping by *Cervus nippon* on survival of *Abies veitchii* in a subalpine coniferous forest in central Japan. *Journal of Forestry Research*.

Nahrung HF, Ramsden M, Griffiths M (2015) Sirex woodwasp range expansion in Australia: performance and parasitism on two commercial pine species. *Forestry* **89**, 310-315.

Namroud M-C, Guillet-Claude C, Mackay J, Isabel N, Bousquet J (2010) Molecular evolution of regulatory genes in spruces from different species and continents: heterogeneous patterns of linkage disequilibrium and selection but correlated recent demographic changes. *Journal of Molecular Evolution* **70**, 371-386.

Nandi A, Yan L-J, Jana CK, Das N (2019) Role of catalase in oxidative stress- and age-associated degenerative diseases. *Oxidative Medicine and Cellular Longevity* **2019**, 9613090.

Nevrtalova E, Baloun J, Hudzieczek V, Cegan R, Vyskot B, Dolezel J, Safar J, Milde D, Hobza R (2014) Expression response of duplicated *metallothionein 3* gene to copper stress in *Silene vulgaris* ecotypes. *Protoplasma* **251**, 1427-1439.

Nichols CP, Drewe JA, Gill R, Goode N, Gregory N (2016) A novel causal mechanism for grey squirrel bark stripping: The Calcium Hypothesis. *Forest Ecology and Management* **367**, 12-20.

NIST (2014) NIST/EPA/NIH Mass Spectral Library. In 'Mass Spectral Search Program. ' (U.S. Government National Institute for Standards and Technology)

Nopp-Mayr U, Reimoser F, Völk F (2011) Predisposition assessment of mountainous forests to bark peeling by red deer (*Cervus elaphus* L.) as a strategy in preventive forest habitat management. *Wildlife Biology in Practice* **7**, 66-89.

O'Reilly-Wapstra JM, Freeman JS, Barbour R, Vaillancourt RE, Potts BM (2013a) Genetic analysis of the near-infrared spectral phenome of a global *Eucalyptus* species. *Tree Genetics & Genomes* **9**, 943-959.

O'Reilly-Wapstra JM, Hamilton M, Gosney B, Whiteley C, Bailey JK, Williams D, Wardlaw T, Vaillancourt RE, Potts BM (2014) Genetic correlations in multi-species plant/herbivore interactions at multiple genetic scales: implications for eco-evolutionary dynamics. In 'Eco-Evolutionary

Dynamics. (Eds J Moya-Laraño, J Rowntree and G Woodward) pp. 267-295. (Academic Press: London).

O'Reilly-Wapstra JM, Iason GR, Thoss V (2007) The role of genetic and chemical variation of *Pinus sylvestris* seedlings in influencing slug herbivory. *Oecologia* **152**, 82-91.

O'Reilly-Wapstra JM, McArthur C, Potts BM (2002) Genetic variation in resistance of *Eucalyptus globulus* to marsupial browsers. *Oecologia* **130**, 289-296.

O'Reilly-Wapstra JM, Miller AM, Hamilton MG, Williams D, Glancy-Dean N, Potts BM (2013b) Chemical variation in a dominant tree species: population divergence, selection and genetic stability across environments. *PLoS One* **8**, e58416.

O'Reilly-Wapstra JM, Potts BM, McArthur C, Davies NW (2005) Effects of nutrient variability on the genetic-based resistance of *Eucalyptus globulus* to a mammalian herbivore and on plant defensive chemistry. *Oecologia* **142**, 597-605.

Offord HR (1964) Diseases of Monterey pine in native stands of California and in plantations of western North America. Pacific Southwest Forest & Range Experiment Station, Berkeley.

Ögren E, Nilsson T, Sundblad LG (1997) Relationship between respiratory depletion of sugars and loss of cold hardiness in coniferous seedlings over‐wintering at raised temperatures: indications of different sensitivities of spruce and pine. *Plant, Cell & Environment* **20**, 247-253.

Ogutu JO, Schulz-Streeck T, Piepho H-P (2012) Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. *BMC Proceedings* **6**, S10.

Ohse B, Hammerbacher A, Seele C, Meldau S, Reichelt M, Ortmann S, Wirth C (2017) Salivary cues: simulated roe deer browsing induces systemic changes in phytohormones and defence chemistry in wild‐grown maple and beech saplings. *Functional Ecology* **31**, 340-349.

Oliveira HR, Lourenco DAL, Masuda Y, Misztal I, Tsuruta S, Jamrozik J, Brito LF, Silva FF, Schenkel FS (2019) Application of single-step genomic evaluation using multiple-trait random regression test-day models in dairy cattle. *Journal of Dairy Science* **102**, 2365-2377.

Ott DS, Yanchuk AD, Huber DPW, Wallin KF (2011) Genetic variation of lodgepole pine, *Pinus contorta* var. *latifolia*, chemical and physical defenses that affect mountain pine beetle, *Dendroctonus ponderosae*, attack and tree mortality. *Journal of Chemical Ecology* **37**, 1002.

Padham AK, Hopkins MT, Wang T-W, McNamara LM, Lo M, Richardson LGL, Smith MD, Taylor CA, Thompson JE (2007) Characterization of a plastid triacylglycerol lipase from Arabidopsis. *Plant Physiology* **143**, 1372-1384.

Page DE, Close D, Beadle CL, Wardlaw TJ, Mohammed CL (2013) Seasonal dynamics in understorey abundance and carbohydrate concentration in relation to browsing and bark stripping of Tasmanian *Pinus radiata* plantations. *Forest Ecology and Management* **296**, 98-107.

Patrick JW, Botha FC, Birch RG (2013) Metabolic engineering of sugars and simple sugar derivatives in plants. *Plant Biotechnology Journal* **11**, 142-156.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* **14**, 417-419.

Pearse IS, Paul R, Ode PJ (2018) Variation in plant defense suppresses herbivore performance. *Current Biology* **28**, 1981-1986.e2.

Pederson JC, Welch BL (1985) Comparison of ponderosa pines as feed and nonfeed trees for Abert squirrels. *Journal of Chemical Ecology* **11**, 149-157.

Pérez-de-Castro AM, Vilanova S, Cañizares J, Pascual L, Blanca JM, Díez MJ, Prohens J, Picó B (2012) Application of genomic tools in plant breeding. *Current genomics* **13**, 179-195.

Pérez-Marín D, Fearn T, Guerrero JE, Garrido-Varo A (2012) Improving NIRS predictions of ingredient composition in compound feedingstuffs using Bayesian non-parametric calibrations. *Chemometrics and Intelligent Laboratory Systems* **110**, 108-112.

Perkins A, Nelson KJ, Parsonage D, Poole LB, Karplus PA (2015) Peroxiredoxins: guardians against oxidative stress and modulators of peroxide signaling. *Trends in Biochemical Sciences* **40**, 435-445.

Phillips MA, Croteau RB (1999) Resin-based defenses in conifers. *Trends in Plant Science* **4**, 184-190.

Picart P, Pirttilä AM, Raventos D, Kristensen H-H, Sahl H-G (2012) Identification of defensin-encoding genes of *Picea glauca*: characterization of PgD5, a conserved spruce defensin with strong antifungal activity. *BMC Plant Biology* **12**, 180.

Piper FI, Fajardo A, Hoch G (2017) Single-provenance mature conifers show higher non-structural carbohydrate storage and reduced growth in a drier location. *Tree Physiology* **37**, 1001-1010.

Podger FD, Wardlaw TJ (1990) Spring needle-cast of *Pinus radiata* in Tasmania: I. Symptoms, distribution, and association with *Cyclaneusma minus*. *New Zealand Journal of Forestry Science* **20**, 184-205.

Poopat P (2013) The ups and downs of conifer defense: linking aboveground herbivory and induced needle and root defenses. Undergraduate Honours Thesis 467 Thesis, University of Colorado, Boulder, CO

Porth I, Hamberger B, White R, Ritland K (2011) Defense mechanisms against herbivory in *Picea*: sequence evolution and expression regulation of gene family members in the phenylpropanoid pathway. *BMC Genomics* **12**, 608.

Preda C, Saporta G, Lévéder C (2007) PLS classification of functional data. *Computational Statistics* **22**, 223-235.

Price T, Schluter D (1991) On the low heritability of life‐history traits. *Evolution* **45**, 853-861.

Proels RK, Hückelhoven R (2014) Cell-wall invertases, key enzymes in the modulation of plant metabolism during defence responses. *Molecular Plant Pathology* **15**, 858-864.

Quentin AG, Rodemann T, Doutreleau M-F, Moreau M, Davies NW, Millard P (2017) Application of near-infrared spectroscopy for estimation of non-structural carbohydrates in foliar samples of *Eucalyptus globulus* Labilladière. *Tree Physiology* **37**, 131-141.

R Core Team (2018) R: A language and environment for statistical computing. In. ' (R Foundation for Statistical Computing: Vienna, Austria)

Radeloff VC, Mladenoff DJ, Boyce MS (1999) Detecting jack pine budworm defoliation using spectral mixture analysis: separating effects from determinants. *Remote Sensing of Environment* **69**, 156-169.

Radwan MA (1972) Differences between Douglas-fir genotypes in relation to browsing preference by black-tailed deer. *Canadian Journal of Forest Research* **2**, 250-255.

Radwan MA, Crouch GL (1978) Selected chemical constituents and deer browsing preference of Douglas fir. *Journal of Chemical Ecology* **4**, 675-683.

Raffa K (2014) Terpenes tell different tales at different scales: glimpses into the chemical ecology of conifer - bark beetle - microbial interactions. *Journal of Chemical Ecology* **40**, 1-20.

Raffa KF, Mason CJ, Bonello P, Cook S, Erbilgin N, Keefover-Ring K, Klutsch JG, Villari C, Townsend PA (2017) Defense syndromes in lodgepole – whitebark pine ecosystems relate to degree of historical exposure to mountain pine beetles. *Plant, Cell & Environment* **40**, 1791-1806.

Raffa KF, Smalley EB (1995) Interaction of pre-attack and induced monoterpene concentrations in host conifer defense against bark beetle-fungal complexes. *Oecologia* **102**, 285-295.

Raj A, Rifkin SA, Andersen E, van Oudenaarden A (2010) Variability in gene expression underlies incomplete penetrance. *Nature* **463**, 913-918.

Ralph SG, Yueh H, *et al.* (2006) Conifer defence against insects: microarray gene expression profiling of Sitka spruce (*Picea sitchensis*) induced by mechanical wounding or feeding by spruce budworms (*Choristoneura occidentalis*) or white pine weevils (*Pissodes strobi*) reveals large-scale changes of the host transcriptome. *Plant, Cell & Environment* **29**, 1545-1570.

Rambla FJ, Garrigues S, de la Guardia M (1997) PLS-NIR determination of total sugar, glucose, fructose and sucrose in aqueous solutions of fruit juices. *Analytica Chimica Acta* **344**, 41-53.

Ramirez JA, Posada JM, Handa IT, Hoch G, Vohland M, Messier C, Reu B (2015) Near‐infrared spectroscopy (NIRS) predicts non‐structural carbohydrate concentrations in different tissue types of a broad range of tree species. *Methods in Ecology and Evolution* **6**, 1018-1025.

Ranade SS, Delhomme N, Garcia-Gil MR (2019) Global gene expression analysis in etiolated and de-etiolated seedlings in conifers. *PLoS One* **14**, e0219272.

Ranocha P, Chabannes M, Chamayou S, Danoun Sd, Jauneau A, Boudet A-M, Goffner D (2002) Laccase down-regulation causes alterations in phenolic metabolism and cell wall structure in poplar. *Plant Physiology* **129**, 145-155.

Rasbery JM, Shan H, LeClair RJ, Norman M, Matsuda SPT, Bartel B (2007) *Arabidopsis thaliana* squalene epoxidase 1 is essential for root and seed development. *Journal of Biological Chemistry* **282**, 17002-17013.

Rasmann S, Agrawal AA (2008) In defense of roots: a research agenda for studying plant resistance to belowground herbivory. *Plant Physiology* **146**, 875-880.

Ratcliffe B, Gamal El-Dien O, Cappa EP, Porth I, Klápště J, Chen C, El-Kassaby YA (2017) Single-step BLUP with varying genotyping effort in open-pollinated *Picea glauca*. *Genes, Genomes, Genetics* **7**, 935-942.

Rautiainen M, Lukeš P, Homolová L, Hovi A, Pisek J, Mõttus M (2018) Spectral properties of coniferous forests: a review of in situ and laboratory measurements. *Remote Sensing* **10**, 207.

Ravanel S, Gakière B, Job D, Douce R (1998) The specific features of methionine biosynthesis and metabolism in plants. *Proceedings of the National Academy of Sciences* **95**, 7805-7812.

Rea RV, Hjeljord O, Härkönen S (2014) Differential selection of North American and Scandinavian conifer browse by northwestern moose (*Alces alces andersoni*) in winter. *Acta Theriologica* **59**, 353-360.

Reglinski T, Spiers TM, Dick MA, Taylor JT, Gardner J (2009) Management of phytophthora root rot in radiata pine seedlings. *Plant Pathology* **58**, 723-730.

Reglinski T, Taylor JT, Northcott GL, Ah Chee A, Spiers M, Wohlers M (2019) Growth environment and seedling age affect constitutive and inducible defence in radiata pine. *Plant Pathology* **68**, 1481-1492.

Reglinski T, Taylor JT, Northcott GL, Ah Chee A, Spiers M, Wohlers M, Hill RA (2017) Biochemical responses associated with induced resistance to *Colletotrichum acutatum* in *Pinus radiata* seedlings treated with methyl jasmonate and *Trichoderma* spp. *Forest Pathology* **47**, e12350.

Resende MFR, Muñoz P, Acosta JJ, Peter GF, Davis JM, Grattapaglia D, Resende MDV, Kirst M (2012a) Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. *New Phytologist* **193**, 617-624.

Resende MFR, Muñoz P*, et al.* (2012b) Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L.). *Genetics* **190**, 1503-1510.

Reymond P, Bodenhausen N, Van Poecke RMP, Krishnamurthy V, Dicke M, Farmer EE (2004) A conserved transcript pattern in response to a specialist and a generalist herbivore. *The Plant Cell* **16**, 3132-3147.

Rezzonico E, Flury N, Meins F, Beffa R (1998) Transcriptional down-regulation by abscisic acid of pathogenesis-related β-1,3-glucanase genes in tobacco cell cultures. *Plant Physiology* **117**, 585-592.

Richard S, Lapointe G, Rutledge RG, Séguin A (2000) Induction of chalcone synthase expression in white spruce by wounding and jasmonate. *Plant and Cell Physiology* **41**, 982-987.

Riley MR, Crider HM (2000) The effect of analyte concentration range on measurement errors obtained by NIR spectroscopy. *Talanta* **52**, 473-484.

Rinnan Å, Berg Fvd, Engelsen SB (2009) Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry* **28**, 1201-1222.

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140.

Roitto M, Rautio P, Markkola A, Julkunen-tiitto R, Varama M, Saravesi K, Tuomi J (2009) Induced accumulation of phenolics and sawfly performance in Scots pine in response to previous defoliation. *Tree Physiology* **29**, 207-216.

Rönnegård L, Shen X (2016) Genomic prediction and estimation of marker interaction effects. In. ' bioRxiv)

Roth M, Hussain A, Cale JA, Erbilgin N (2018) Successful colonization of lodgepole pine trees by mountain pine beetle increased monoterpene production and exhausted carbohydrate reserves. *Journal of Chemical Ecology* **44**, 209-214.

Rousseeuw PJ, Debruyne M, Engelen S, Hubert M (2006) Robustness and outlier detection in chemometrics. *Critical Reviews in Analytical Chemistry* **36**, 221-242.

Rubert-Nason KF, Holeski LM, Couture JJ, Gusse A, Undersander DJ, Lindroth RL (2013) Rapid phytochemical analysis of birch (*Betula*) and poplar (*Populus*) foliage by near-infrared reflectance spectroscopy. *Analytical and Bioanalytical Chemistry* **405**, 1333-1344.

Russell J (2008) Deployment of deer-resistant western redcedar (*Thuja plicata*). In 'National Proceedings: Forest and Conservation Nursery Associations - 2007. (Eds RK Dumroese and LE Riley) pp. 149-153. (U. S. Department of Agriculture, Forest Service, Rocky Mountain Research Station: Fort Collins, CO).

Saccenti E, Hoefsloot HCJ, Smilde AK, Westerhuis JA, Hendriks MMWB (2014) Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics* **10**, 361-374.

Saeki Y, Tuda M, Crowley PH (2014) Allocation tradeoffs and life histories: a conceptual and graphical framework. *Oikos* **123**, 786-793.

Saeys W, Mouazen AM, Ramon H (2005) Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy. *Biosystems Engineering* **91**, 393-402.

Saint-Andrieux C, Bonenfant C, Toïgo C, Basille M, Klein F (2009) Factors affecting beech *Fagus sylvatica* bark stripping by red deer *Cervus elaphus* in a mixed forest. *Wildlife Biology* **15**, 187-197.

Salem MZM, Ali HM, Basalah MO (2014) Essential oils from wood, bark, and needles of *Pinus roxburghii* Sarg. from Alexandria, Egypt: antibacterial and antioxidant activities. *BioResources* **9**, 7454-7466.

Sampedro L, Moreira X, Llusia J, Peñuelas J, Zas R (2010) Genetics, phosphorus availability, and herbivore-derived induction as sources of phenotypic variation of leaf volatile terpenes in a pine species. *Journal of Experimental Botany* **61**, 4437-4447.

Sampedro L, Moreira X, Zas R (2011) Costs of constitutive and herbivore-induced chemical defences in pine trees emerge only under low nutrient availability. *Journal of Ecology* **99**, 818-827.

Sasaki Y, Nagano Y (2004) Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding. *Bioscience, Biotechnology, and Biochemistry* **68**, 1175-1184.

Sasidharan S, Chen Y, Saravanan D, Sundram KM, Latha LY (2011) Extraction, isolation and characterization of bioactive compounds from plants' extracts. *African Journal of Traditional, Complementary, and Alternative Medicines* **8**, 1-10.

Sasikumar AN, Perez WB, Kinzy TG (2012) The many roles of the eukaryotic elongation factor 1 complex. *WIREs RNA* **3**, 543-555.

Sasse J, Elms S, Kube P (2009) Genetic resistance in *Pinus radiata* to defoliation by the pine aphid *Essigella californica*. *Australian Forestry* **72**, 25-31.

Sauvé DG, Cŏté SD (2007) Winter forage selection in white-tailed deer at high density: balsam fir is the best of a bad choice. *Journal of Wildlife Management* **71**, 911-914.

Scalerandi E, Flores GA, Palacio M, Defagó MT, Carpinella MC, Valladares G, Bertoni A, Palacios SM (2018) Understanding synergistic toxicity of terpenes as insecticides: contribution of metabolic detoxification in *Musca domestica*. *Frontiers in Plant Science* **9**, 1579.

Schimleck LR, Doran JC, Rimbawanto A (2003) Near infrared spectroscopy for cost effective screening of foliar oil characteristics in a *Melaleuca cajuputi* breeding population. *Journal of Agricultural and Food Chemistry* **51**, 2433-2437.

Schimleck LR, Evans R, Matheson AC (2002) Estimation of *Pinus radiata* D. Don clear wood properties by near-infrared spectroscopy. *Journal of Wood Science* **48**, 132-137.

Schimleck LR, Yazaki Y (2003) Analysis of *Pinus radiata* D. Don bark by near infrared spectroscopy. *Holzforschung* **57**, 520-526.

Schmidt A, Zeneli G, Hietala AM, Fossdal CG, Krokene P, Christiansen E, Gershenzon J (2005) Induced chemical defenses in conifers: biochemical and molecular approaches to studying their function. In 'Chemical Ecology and Phytochemistry in Forest Ecosystems. Vol. 39. (Ed. JT Romeo) pp. 1-28. (Elsevier: Amsterdam).

Schowalter TD (2012) Ecology and management of bark beetles (Coleoptera: Curculionidae: Scolytinae) in southern pine forests. *Journal of Integrated Pest Management* **3**, A1-A7.

Schowalter TD (2016) 'Insect Ecology: An Ecosystem Approach.' 4th edn. (Academic Press).

Schulz H, Schrader B, Quilitzsch R, Pfeffer S, Krüger H (2003) Rapid classification of basil chemotypes by various vibrational spectroscopy methods. *Journal of Agricultural and Food Chemistry* **51**, 2475-2481.

Schwachtje J, Baldwin IT (2008) Why does herbivore attack reconfigure primary metabolism? *Plant Physiology* **146**, 845-851.

Schwanninger M, Rodrigues JC, Fackler K (2011) A review of band assignments in near infrared spectra of wood and wood components. *Journal of Near Infrared Spectroscopy* **19**, 287-308.

Sellés Vidal L, Kelly CL, Mordaka PM, Heap JT (2018) Review of NAD(P)H-dependent oxidoreductases: Properties, engineering and application. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1866**, 327-347.

Senior JK, Potts BM, O'Reilly‑Wapstra JM, Bissett A, Wooliver RC, Bailey JK, Glen M, Schweitzer JA (2018) Phylogenetic trait conservatism predicts patterns of plant‑soil feedback. *Ecosphere* **9**, e02409.

Seybold SJ, Huber DPW, Lee JC, Graves AD, Bohlmann J (2006) Pine monoterpenes and pine bark beetles: a marriage of convenience for defense and chemical communication. *Phytochemistry Reviews* **5**, 143-178.

Shah J (2005) Lipids, lipases, and lipid-modifying enzymes in plant disease resistance. *Annual Review of Phytopathology* **43**, 229-260.

Shanmugabalaji V, Grimm B, Kessler F (2020) Characterization of a plastoglobule-localized $SOUL_4$ heme-binding protein in *Arabidopsis thaliana*. *Frontiers in Plant Science* **11**.

Sharma A, Shahzad B, Rehman A, Bhardwaj R, Landi M, Zheng B (2019) Response of phenylpropanoid pathway and the role of polyphenols in plants under abiotic stress. *Molecules* **24**, 2452.

Shen X, Alam M, Fikse F, Rönnegård L (2013) A novel generalized ridge regression method for quantitative genetics. *Genetics* **193**, 1255-1268.

Shibata Ei (1998) Effects of Japanese cedar inner bark nutritional quality on development of *Semanotus japonicus* (Coleoptera: Cerambycidae). *Environmental Entomology* **27**, 1431-1436.

Shigeto J, Tsutsumi Y (2016) Diverse functions and reactions of class III peroxidases. *New Phytologist* **209**, 1395-1402.

Shintani DK, Cheng Z, DellaPenna D (2002) The role of 2-methyl-6-phytylbenzoquinone methyltransferase in determining tocopherol composition in *Synechocystis* sp. PCC6803. *FEBS Letters* **511**, 1-5.

Siesler HW, Ozaki Y (2002) 'Near-infrared spectroscopy: principles, instruments, applications.' (Wiley-VCH: Weinheim, Germany).

Silen RR, Randall WK, Mandel NL (1986) Estimates of genetic parameters for deer browsing of Douglas-fir. *Forest Science* **32**, 178-184.

Singh D, Wang X, Kumar U, Gao L, Noor M, Imtiaz M, Singh RP, Poland J (2019) High-throughput phenotyping enabled genetic dissection of crop lodging in wheat. *Frontiers in Plant Science* **10**, 394.

Smith AH, Ratkowsky DA, Wardlaw TJ, Mohammed CL (2020) Ease of access to an alternative food source enables wallabies to strip bark in Tasmanian *Pinus radiata* plantations. *Forests* **11**, 387.

Sniezko RA, Koch J (2017) Breeding trees resistant to insects and diseases: putting theory into application. *Biological Invasions* **19**, 3377-3400.

Snyder MA (1992) Selective herbivory by Abert's squirrel mediated by chemical variability in ponderosa pine. *Ecology* **73**, 1730-1741.

Snyder MA (1993) Interactions between Abert's squirrel and ponderosa pine: the relationship between selective herbivory and host plant fitness. *The American Naturalist* **141**, 866-879.

Song H, Zhang J, Zhang Q, Ding X (2019) Using Different Single-Step Strategies to Improve the Efficiency of Genomic Prediction on Body Measurement Traits in Pig. *Frontiers in genetics* **9**, 730-730.

South PF, Walker BJ, Cavanagh AP, Rolland V, Badger M, Ort DR (2017) Bile acid sodium symporter BASS6 can transport glycolate and is involved in photorespiratory metabolism in *Arabidopsis thaliana*. *The Plant Cell* **29**, 808-823.

Stackpole DJ, Vaillancourt RE, Alves A, Rodrigues J, Potts BM (2011) Genetic variation in the chemical components of *Eucalyptus globulus* wood. *Genes, Genomes, Genetics* **1**, 151-159.

Stein O, Granot D (2019) An Overview of Sucrose Synthases in Plants. *Frontiers in plant science* **10**, 95-95.

Stejskal J, Lstibůrek M, Klápště J, Čepl J, El-Kassaby YA (2018) Effect of genomic prediction on response to selection in forest tree breeding. *Tree Genetics & Genomes* **14**.

Stival Sena J, Giguère I, Rigault P, Bousquet J, MacKay J (2018) Expansion of the dehydrin gene family in the Pinaceae is associated with considerable structural diversity and drought-responsive expression. *Tree Physiology* **38**, 442-456.

Stuart BH (2004) 'Infrared spectroscopy: fundamentals and applications.' (John Wiley & Sons, Ltd: Chicester, West Sussex, UK).

Stutz RS, Croak BM, Proschogo N, Banks PB, McArthur C (2017) Olfactory and visual plant cues as drivers of selective herbivory. *Oikos* **126**, 259-268.

Sun T-p (2011) The molecular mechanism and evolution of the GA–GID1–DELLA signaling module in plants. *Current Biology* **21**, R338-R345.

Sunnerheim-Sjöberg K, Hämäläinen M (1992) Multivariate study of moose browsing in relation to phenol pattern in pine needles. *Journal of Chemical Ecology* **18**, 659-672.

Suontama M, Klápště J, Telfer E, Graham N, Stovold T, Low C, McKinley R, Dungey H (2018) Efficiency of genomic prediction across two *Eucalyptus nitens* seed orchards with different selection histories. *Heredity* **122**, 370-379.

Suontama M, Li Y, Low CB, Dungey HS (2019) Genetic improvement of resistance to cyclaneusma needle cast in *Pinus radiata*. *Canadian Journal of Forest Research* **49**, 128-133.

Swett CL, Gordon TR (2017) Exposure to a pine pathogen enhances growth and disease resistance in *Pinus radiata* seedlings. *Forest Pathology* **47**, e12298.

Tabita FR (1999) Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: A different perspective. *Photosynthesis Research* **60**, 1-28.

Tamura N, Ohara S (2005) Chemical components of hardwood barks stripped by the alien squirrel *Callosciurus erythraeus* in Japan. *Journal of Forest Research* **10**, 429-433.

Tauzin AS, Giardina T (2014) Sucrose and invertases, a part of the plant defense response to the biotic stresses. *Frontiers in Plant Science* **5**, 293-293.

Telfer E, Graham N, Macdonald L, Li Y, Klápště J, Resende Jr M, Neves LG, Dungey H, Wilcox P (2019) A high-density exome capture genotype-by-sequencing panel for forestry breeding in *Pinus radiata*. *PLoS One* **14**.

Telfer E, Graham N, Macdonald L, Sturrock S, Wilcox P, Stanbra L (2018) Approaches to variant discovery for conifer transcriptome sequencing. *PLoS One* **13**, e0205835.

Telfer E, Graham N, Stanbra L, Manley T, Wilcox P (2013) Extraction of high purity genomic DNA from pine for use in a high-throughput genotyping platform. *New Zealand Journal of Forestry Science* **43**, 3.

Telford A, Cavers S, Ennos RA, Cottrell JE (2014) Can we protect forests by harnessing variation in resistance to pests and pathogens? *Forestry* **88**, 3-12.

Tenhaken R (2015) Cell wall remodeling under abiotic stress. *Frontiers in Plant Science* **5**, 771.

Thistlethwaite FR, Ratcliffe B, Klápště J, Porth I, Chen C, Stoehr MU, El-Kassaby YA (2017) Genomic prediction accuracies in space and time for height and wood density of Douglas-fir using exome capture as the genotyping platform. *BMC Genomics* **18**, 930.

Thistlethwaite FR, Ratcliffe B, Klápště J, Porth I, Chen C, Stoehr MU, El-Kassaby YA (2019) Genomic selection of juvenile height across a single-generational gap in Douglas-fir. *Heredity* **122**, 848-863.

Tiffin P (2000) Mechanisms of tolerance to herbivore damage: what do we know? *Evolutionary Ecology* **14**, 523-536.

Tillich M, Hardel SL*, et al.* (2009) Chloroplast ribonucleoprotein CP31A is required for editing and stability of specific chloroplast mRNAs. *Proceedings of the National Academy of Sciences* **106**, 6002-6007.

Tinus RW, Burr KE, Atzmon N, Riov J (2000) Relationship between carbohydrate concentration and root growth potential in coniferous seedlings from three climates during cold hardening and dehardening. *Tree Physiology* **20**, 1097-1104.

Tomlin ES, Antonejevic E, Alfaro RI, Borden JH (2000) Changes in volatile terpene and diterpene resin acid composition of resistant and susceptible white spruce leaders exposed to simulated white pine weevil damage. *Tree Physiology* **20**, 1087-1095.

Tomlin ES, Borden JH (1997) Thin bark and high density of outer resin ducts: interrelated resistance traits in Sitka spruce against the white pine weevil (Coleoptera: Curculionidae). *Journal of Economic Entomology* **90**, 235-239.

Tong M, Jiang Y (2016) FK506-binding proteins and their diverse functions. *Current Molecular Pharmacology* **9**, 48-65.

Toscano G, Rinnan Å, Pizzi A, Mancini M (2017) The use of near-infrared (NIR) spectroscopy and principal component analysis (PCA) to discriminate bark and wood of the most common species of the pellet sector. *Energy & Fuels* **31**, 2814-2821.

Travers A (2001) DNA-binding proteins. In 'Encyclopedia of Genetics. (Eds S Brenner and JH Miller) pp. 541-544. (Academic Press: New York).

Trouvelot S, Héloir M-C*, et al.* (2014) Carbohydrates in plant immunity and plant protection: roles and potential application as foliar sprays. *Frontiers in Plant Science* **5**, 592-592.

Tsuchiya T, Ohta H, Okawa K, Iwamatsu A, Shimada H, Masuda T, Takamiya K-i (1999) Cloning of chlorophyllase, the key enzyme in chlorophyll degradation: Finding of a lipase motif and the induction by methyl jasmonate. *Proceedings of the National Academy of Sciences* **96**, 15362-15367.

Turek K, Kamler J, Procházka L (2016) The impact of thinning type on bark stripping damage intensity caused by red deer (*Cervus elaphus* L.). *Baltic Forestry* **22**, 246-250.

Tyacke RJ, Contestabile R, Grimm B, Harwood JL, John RA (1995) Reactions of glutamate semialdehyde aminotransferase (glutamate-1-semialdehyde 2,1 aminomutase) with vinyl and acetylenic substrate analogues analysed by rapid scanning spectrophotometry. *Biochemical Journal* **309**, 307-313.

Ukrainetz NK, Kang K-Y, Aitken SN, Stoehr M, Mansfield SD (2008) Heritability and phenotypic and genetic correlations of coastal Douglas-fir (*Pseudotsuga menziesii*) wood quality traits. *Canadian Journal of Forest Research* **38**, 1536-1546.

Utz HF, Melchinger AE, Schön CC (2000) Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. *Genetics* **154**, 1839-1849.

Van Damme M, Huibers RP, Elberse J, Van den Ackerveken G (2008) Arabidopsis *DMR6* encodes a putative 2OG-Fe(II) oxygenase that is defense-associated but required for susceptibility to downy mildew. *The Plant Journal* **54**, 785-793.

van der Sluis R, Erasmus E (2016) Xenobiotic/medium chain fatty acid: CoA ligase – a critical review on its role in fatty acid metabolism and the detoxification of benzoic acid and aspirin. *Expert Opinion on Drug Metabolism & Toxicology* **12**, 1169-1179.

vanAkker L, Alfaro RI, Brockley R (2004) Effects of fertilization on resin canal defences and incidence of *Pissodes strobi* attack in interior spruce. *Canadian Journal of Forest Research* **34**, 855-862.

VanRaden PM (2008) Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science* **91**, 4414-4423.

Vasseur F, Fouqueau L, de Vienne D, Nidelet T, Violle C, Weigel D (2019) Nonlinear phenotypic variation uncovers the emergence of heterosis in *Arabidopsis thaliana*. *PLoS Biology* **17**.

Veerman JR, Leday GGR, van de Wiel MA (2019) Estimation of variance components, heritability and the ridge penalty in high-dimensional generalized linear models. *Communications in Statistics - Simulation and Computation*, 1-19.

Verbyla KL, Hayes BJ, Bowman PJ, Goddard ME (2009) Accuracy of genomic selection using stochastic search variable selection in Australian Holstein Friesian dairy cattle. *Genetics Research* **91**, 307-311.

Verne S, Jaquish B, White R, Ritland C, Ritland K (2011) Global transcriptome analysis of constitutive resistance to the white pine weevil in spruce. *Genome Biology and Evolution* **3**, 851-867.

Viana JMS, Faria VR, da Costa e Silva A (2009) Bias in the prediction of genetic gain due to mass and half-sib selection in random mating populations. *Genetics and Molecular Biology* **32**, 497-506.

Villamuelas M, Serrano E*, et al.* (2017) Predicting herbivore faecal nitrogen using a multispecies near-infrared reflectance spectroscopy calibration. *PLoS One* **12**, e0176635-e0176635.

Villari C, Faccoli M, Battisti A, Bonello P, Marini L (2014) Testing phenotypic trade-offs in the chemical defence strategy of Scots pine under growth-limiting field conditions. *Tree Physiology* **34**, 919-930.

Vinaixa M, Samino S, Saez I, Duran J, Guinovart JJ, Yanes O (2012) A guideline to univariate statistical analysis for LC/MS-based untargeted metabolomics-derived data. *Metabolites* **2**, 775-795.

Vourc'h G, De Garine-Wichatitsky M, Labbé A, Rosolowski D, Martin J-L, Fritz H (2002a) Monoterpene effect on feeding choice by deer. *Journal of Chemical Ecology* **28**, 2411-2427.

Vourc'h G, Vila B, Gillon D, Escarré J, Guibal F, Fritz H, Clausen TP, Martin JL (2002b) Disentangling the causes of damage variation by deer browsing on young *Thuja plicata*. *Oikos* **98**, 271-283.

Vourch G, Russell J, Martin J-L (2002) Linking deer browsing and terpene production among genetic identities in *Chamaecyparis nootkatensis* and *Thuja plicata* (Cupressaceae). *Journal of Heredity* **93**, 370-376.

Wainhouse D, Staley JT, Jinks R, Morgan G (2009) Growth and defence in young pine and spruce and the expression of resistance to a stem-feeding weevil. *Oecologia* **158**, 641.

Walden AR, Walter C, Gardner RC (1999) Genes expressed in *Pinus radiata* male cones include homologs to anther-specific and pathogenesis response genes. *Plant Physiology* **121**, 1103-1116.

Wang D, Guo Y, Wu C, Yang G, Li Y, Zheng C (2008) Genome-wide analysis of CCCH zinc finger family in Arabidopsis and rice. *BMC Genomics* **9**, 44.

Wang J, Zhou Z, Zhang Z, Li H, Liu D, Zhang Q, Bradbury PJ, Buckler ES, Zhang Z (2018a) Expanding the BLUP alphabet for genomic prediction adaptable to the genetic architectures of complex traits. *Heredity* **121**, 648-662.

Wang X, Xu Y, Hu Z, Xu C (2018b) Genomic selection methods for crop improvement: Current status and prospects. *The Crop Journal* **6**, 330-340.

Ward AI, White PCL, Smith A, Critchley CH (2004) Modelling the cost of roe deer browsing damage to forestry. *Forest Ecology and Management* **191**, 301-310.

Weig A, Franz J, Sauer N, Komor E (1994) Isolation of a family of cDNA clones from *Ricinus communis* L. with close homology to the hexose carriers. *Journal of Plant Physiology* **143**, 178–183.

Welch D, Scott D (1998) Bark-stripping damage by red deer in a Sitka spruce forest in western Scotland IV. Survival and performance of wounded trees. *Forestry* **71**, 225-235.

Welch D, Staines BW, Scott D, Catt DC (1988) Bark stripping damage by red deer in a Sitka spruce forest in western Scotland II. Wound size and position. *Forestry* **61**, 245-254.

Weng Y, Tosh K, Park Y, Fullarton M (2007) Age-related trends in genetic parameters for jack pine and their implications for early selection. *Silvae Genetica* **56**, 242-252.

Westbrook JW, Walker AR*, et al.* (2015) Discovering candidate genes that regulate resin canal number in *Pinus taeda* stems by integrating genetic analysis across environments, ages, and populations. *New Phytologist* **205**, 627-641.

Westoby M (1978) What are the biological bases of varied diets? *The American Naturalist* **112**, 627-631.

White TCR (2019) The cause of bark stripping of young plantation trees. *Annals of Forest Science* **76**, 105.

White TW, Adams WT, Neale DB (2007) 'Forest Genetics.' (CABI Publishing: Cambridge, MA).

Whitehill JGA, Bohlmann J (2019) A molecular and genomic reference system for conifer defence against insects. *Plant, Cell & Environment* **42**, 2844-2859.

Whitehill JGA, Henderson H, Schuetz M, Skyba O, Yuen MMS, King J, Samuels AL, Mansfield SD, Bohlmann J (2016) Histology and cell wall biochemistry of stone cells in the physical defence of conifers against insects. *Plant, Cell & Environment* **39**, 1646-1661.

Whitehill JGA, Yuen MMS, Henderson H, Madilao L, Kshatriya K, Bryan J, Jaquish B, Bohlmann J (2019) Functions of stone cells and oleoresin terpenes in the conifer defense syndrome. *New Phytologist* **221**, 1503-1517.

Wiley E, Helliker B (2012) A re-evaluation of carbon storage in trees lends greater support for carbon limitation to growth. *New Phytologist* **195**, 285-289.

Wise MJ, Rauscher MD (2013) Evolution of resistance to a multiple‐herbivore community: genetic correlations, diffuse coevolution, and constraints on the plant's response to selection. *Evolution* **67**, 1767-1779.

Wright MN, Ziegler A (2015) ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software* **77**, 1508.04409.

Wu C, Pullinen I, Andres S, Kiendler-Scharr A, Kleist E, Wahner A, Wildt J, Mentel TF (2017) ^{13}C labelling study of constitutive and stress-induced terpenoide missions from Norway spruce and Scots pine. In 'Biogeosciences. ' pp. preprint)

Wu HX, Eldridge KG, Matheson AC, Powell MP, McRae TA (2007) Successful introduction and breeding of radiata pine to Australia. In 'Growing Forest Values. Proceedings of ANZIF 2007 Conference; 3-7 June, 2007; Coffs Harbour, NSW. pp. 506-517. (Institute of Foresters of Australia and New Zealand Institute of Forestry: Canberra).

Xu J, Wang X-y, Guo W-z (2015) The cytochrome P450 superfamily: Key players in plant development and defense. *Journal of Integrative Agriculture* **14**, 1673-1686.

Yamada Y (1962) Genotype by environment interaction and genetic correlation of the same trait under different environments. *The Japanese Journal of Genetics* **37**, 498-509.

Yan H, Xu Y-C, Siesler HW, Han B-X, Zhang G-Z (2019) Hand-Held Near-Infrared Spectroscopy for Authentication of Fengdous and Quantitative Analysis of Mulberry Fruits. *Frontiers in Plant Science* **10**.

Yanchuk AD, Murphy JC, Wallin KF (2008) Evaluation of genetic variation of attack and resistance in lodgepole pine in the early stages of a mountain pine beetle outbreak. *Tree Genetics & Genomes* **4**, 171-180.

Yang J, Duan G, Li C, Liu L, Han G, Zhang Y, Wang C (2019) The crosstalks between jasmonic acid and other plant hormone signaling highlight the involvement of jasmonic acid as a core component in plant response to biotic and abiotic stresses. *Frontiers in Plant Science* **10**, 1349.

Yang Z, Nie G, Pan L, Zhang Y, Huang L, Ma X, Zhang X (2017) Development and validation of near-infrared spectroscopy for the prediction of forage quality parameters in *Lolium multiflorum*. *PeerJ* **5**, e3867.

Zamora Nasca LB, Relva MA, Núñez MA (2018) Ungulates can control tree invasions: experimental evidence from nonnative conifers and sheep herbivory. *Biological Invasions* **20**, 583-591.

Zas R, Björklund N, Nordlander G, Cendán C, Hellqvist C, Sampedro L (2014) Exploiting jasmonate-induced responses for field protection of conifer seedlings against a major forest pest, *Hylobius abietis*. *Forest Ecology and Management* **313**, 212-223.

Zas R, Björklund N, Sampedro L, Hellqvist C, Karlsson B, Jansson S, Nordlander G (2017) Genetic variation in resistance of Norway spruce seedlings to damage by the pine weevil *Hylobius abietis*. *Tree Genetics & Genomes* **13**, 111.

Zas R, Moreira X, Sampedro L (2011) Tolerance and induced resistance in a native and an exotic pine species: relevant traits for invasion ecology. *Journal of Ecology* **99**, 1316-1326.

Zas R, Sampedro L, Moreira X, Martíns P (2008) Effect of fertilization and genetic variation on susceptibility of *Pinus radiata* seedlings to *Hylobius abietis* damage. *Canadian Journal of Forest Research* **38**, 63-72.

Zas R, Sampedro L, Prada E, Fernández-López J (2005) Genetic variation of *Pinus pinaster* Ait. seedlings in susceptibility to the pine weevil *Hylobius abietis* L. *Annals of Forest Science* **62**, 681-688.

Zas R, Solla A, Sampedro L (2007) Variography and kriging allow screening *Pinus pinaster* resistant to *Armillaria ostoyae* in field conditions. *Forestry* **80**, 201-209.

Zhang S, Jiang J, Luan Q (2016a) Genetic and correlation analysis of oleoresin chemical components in slash pine. *Genetics and Molecular Research* **15**, gmr.15038982.

Zhang X, States JS (1991) Selective herbivory of ponderosa pine by Abert squirrels: a re-examination of the role of terpenes. *Biochemical Systematics and Ecology* **19**, 111-115.

Zhang Y, Smallbone LA, diCenzo GC, Morton R, Finan TM (2016b) Loss of malic enzymes leads to metabolic imbalance and altered levels of trehalose and putrescine in the bacterium *Sinorhizobium meliloti*. *BMC Microbiology* **16**, 163.

Zhao J, Missihoun TD, Bartels D (2017) The role of Arabidopsis aldehyde dehydrogenase genes in response to high temperature and stress combinations. *Journal of Experimental Botany* **68**, 4295-4308.

Zhou S, Lou Y-R, Tzin V, Jander G (2015) Alteration of plant primary metabolism in response to insect herbivory. *Plant Physiology* **169**, 1488-1498.

Zhu Y, Wang H, Peng Q, Tang Y, Xia G, Wu J, Xie D-Y (2015) Functional characterization of an anthocyanidin reductase gene from the fibers of upland cotton (*Gossypium hirsutum*). *Planta* **241**, 1075-1089.

Zornoza R, Guerrero C, Mataix-Solera J, Scow KM, Arcenegui V, Mataix-Beneyto J (2008) Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. *Soil Biology and Biochemistry* **40**, 1923-1930.

Zou J, Cates RG (1994) Role of Douglas fir (*Pseudotsuga menziesii*) carbohydrates in resistance to budworm (*Choristoneura occidentalis*). *Journal of Chemical Ecology* **20**, 395-405.

Zulak KG, Bohlmann J (2010) Terpenoid biosynthesis and specialized vascular cells of conifer defense. *Journal of Integrative Plant Biology* **52**, 86-97.