Analysis of plant and fungal transcripts from resistant and susceptible phenotypes of *Leptospermum scoparium* challenged by *Austropuccinia psidii*

Rebekah A. Frampton[1*], Louise S. Shuey[2*], Charles C. David[1], Georgia M. Pringle[1**], Falk Kalamorz[1], Geoff S. Pegg[2], David Chagné[3], Grant R. Smith[1***]

[1] The New Zealand Institute for Plant and Food Research Limited, Lincoln 7608, New Zealand

[2] The Queensland Department of Agriculture and Fisheries, Brisbane, Queensland 4001, Australia

[3] The New Zealand Institute for Plant and Food Research Limited, Palmerston North 4410, New Zealand

* Rebekah A. Frampton and Louise S. Shuey contributed equally to this study

** Current Address: Te Kaunihera o Tāmaki Makaurau, Auckland Council, Auckland 1010, New Zealand

*** Corresponding author Grant R. Smith; grant.smith@plantandfood.co.nz

Abstract

Austropuccinia psidii is the causal pathogen of myrtle rust disease of Myrtaceae. To gain understanding of the initial infection process, gene expression in germinating *Austropuccinia psidii* urediniospores and in *Leptospermum scoparium* inoculated leaves were investigated via analyses of RNAseq samples taken 24 and 48 hours post inoculation (hpi). Principal component analyses of transformed transcript count data revealed differential gene expression between the uninoculated *L. scoparium* control plants that correlated with the three plant leaf resistance phenotypes (immunity, hypersensitive response and susceptibility). Gene expression in the immune resistant plants did not significantly change in response to fungal inoculation, while susceptible plants showed differential expression of genes in response to fungal challenge. A putative disease resistance gene, jg24539.t1, was identified in the *L. scoparium* hypersensitive response phenotype family. Expression of this gene may be associated with the phenotype and could be important for further understanding the plant hypersensitive response to *A. psidii* challenge. Differential expression of pathogen genes was found between samples taken 24 and 48 hpi, but there were no significant differences in pathogen gene expression that were associated with the three different plant leaf resistance phenotypes. There was a significant decrease in the abundance of fungal transcripts encoding three putative effectors and a putative carbohydrate-active enzyme between 24 and 48 hpi, suggesting that the encoded proteins are important during the initial phase of infection. These transcripts, or their translated proteins, may be potential targets to impede the early phases of fungal infection by this wide-host range obligate biotrophic basidiomycete.

Introduction

Austropuccinia psidii (G. Winter) Beenken *comb. nov.* (Beenken 2017) is the causal agent of myrtle rust disease on a wide range of Myrtaceae around the world [480 species in 69 genera (Soewarto et al. 2019a)]. This obligate, biotrophic basidiomycete was first described by Winter (1884) as *Puccinia psidii* G. Winter on the leaves of *Psidium guajava* L. (= *Psidium pomiferum*) in Brasil, and has a number of synonyms including *Bullaria psidii, Dicaeoma psidiia* and *Uredo rangelii* (Carnegie and Giblin 2014). *Austropuccinia psidii* is an invasive fungal pathogen that is now widely dispersed and found throughout the American continent and Caribbean, Hawaii, Asia, Australia, New Caledonia, Aotearoa (New Zealand) and South Africa (Carnegie and Pegg 2018). Stewart et al. (2018) identified nine distinct genetic clusters (C1–C9) of *A. psidii* based on multilocus genotype (MLG) analysis and host range. The pandemic strain, comprising the C1 and C3 clusters, is now causing the functional extinction of Myrtaceae from the sub-tropical east coast of Australia (Fensham et al. 2020) and presents a generational extinction threat to a range of Australian east coast tropical species (Fensham and Radford-Smith 2021).

Myrtle rust, resulting from infection by the pandemic strain (du Plessis et al. 2019), was first found in Aotearoa on *Metrosideros kermadecensis* (Kahikā Rangitāhua, Kermadec pōhutukawa) growing on Rangitāhua (Raoul Island), the largest of the Kermadec Islands, in April 2017. A month later the pathogen was found in Kerikeri in Te Ika-a-Māui (North Island) on *M. excelsa* (pōhutukawa; Ho et al. 2019). The pathogen is now widely established throughout most of Te Ika-a-Māui, has also been found on the west coast of Te Waipounamu (South Island) and as far south as Ōtautahi (Christchurch) on the east coast of Te Waipounamu (Anonymous 2021).

Currently, there are 18 recognised endemic and indigenous Myrtaceae in six genera in Aotearoa. These species are considered taonga (treasure) by Māori, who have a role in providing guardianship over these species. The reduction from the 27 species noted in Smith et al. (2020) is a consequence of the findings by Heenan et al. (2023) from a single nucleotide polymorphism (SNP) analysis of *Kunzea* that resulted in recognition of a single species, *K. ericoides*, in Aotearoa and the findings of a SNP analysis of *Leptospermum scoparium* (mānuka) by Chagné et al. (2023) which provided 'little support for taxonomic revision and subdividing *L. scoparium* into segregate species' within Aotearoa, negating the recent taxonomic determinations of *Leptospermum repo* (de Lange and Schmid 2021) and *L. hoipolloi* (Schmid et al. 2023). These 18 endemic and indigenous species are susceptible to infection by the pandemic strain of *A. psidii* (Smith et al. 2020). The pandemic strain of *A. psidii* has now caused the localised extinction of *Lophomyrtus bullata* (ramarama) in the east cape region of Te Ika-a-Māui (Gifford 2021). Additionally, some New Zealand Myrtaceae are also susceptible to infection by the

South African strain (Roux et al. 2016) of *A. psidii* (Soewarto et al. 2021). Resistance to this pathogen is present in some species, in particular mānuka (*L. scoparium*) and kānuka (formerly *Kunzea robusta* and *K. linearis,* now *K. ericoides*), however the percentage of resistant plants in the provenances tested is relatively low (Smith et al. 2020). Additionally, both the leaf and stem of *L. scoparium* and *K. ericoides* are infected by the pathogen: analysis of the phenotype distribution of leaf/stem infection in plants from a range of sibling families suggested that the leaf and stem resistances are the result of independent disease resistance mechanisms (Smith et al. 2020).

Obligate biotrophic pathogens specifically interact with, and manipulate, their hosts to obtain resources for reproduction without triggering host defence responses. Flor (1942) first described the genetics of these interactions in the *Melampsora lini – Linum usitatissimum* pathosystem leading to the gene-for-gene hypothesis (Flor 1955) that defined the genetics of the specificity of the pathogen-host interaction. The extensive host range of the pandemic strain of *A. psidii* presents a significant challenge to understand the basis of pathogenicity and the mechanisms that this obligate biotrophic fungus utilises to infect and reproduce on its wide range of hosts (Soewarto et al. 2019a). Loci that are associated with resistance to *A. psidii* have been identified in different Myrtaceae. The first locus reported, *Puccinia psidii* resistance gene 1 (Ppr1) (Junghans et al. 2003b), has been successfully used in the Brasilian *Eucalyptus grandis* breeding program, although a new race of *A. psidii* has been reported to have now broken that resistance (Almeida et al. 2021). Four additional independent resistance loci were found in *Eucalyptus globulus*: Ppr2 and Ppr3 were associated with disease symptom expression, whilst Ppr4 and Ppr5 were associated with hypersensitive resistance (Butler et al. 2016). The presence of these resistance loci, whilst beneficial, is intriguing as the pathogen does not share co-evolutionary history with most of its current hosts.

Results from a range of investigations to understand the molecular/biochemical basis of host plant resistance to *A. psidii* have been reported including metabolic analyses (Moffitt et al. 2022, Sekiya et al. 2021) and transcriptomic analyses of *Arillastrum gummiferum*, *Syzygium longifolium* and *Tristaniopsis glauca* (Soewarto et al. 2019b), *E. grandis* (Santos et al. 2020, Swanepoel et al. 2021, Swanepoel et al. 2023), *Melaleuca quinquenervia* (Hsieh et al. 2018) and *Syzygium luehmannii* (Tobias et al. 2018). Several putative host resistance mechanisms were identified in these studies, including over-expression of receptor-like kinases, nucleotide-binding site leucine-rich repeat proteins (Hsieh et al. 2018, Tobias et al. 2018), protein kinase leucine rich receptors (Santos et al. 2020), endochitinases (Soewarto et al. 2019b) and brassinosteroid mediated signalling genes (Swanepoel et al. 2021). These studies have largely focused on the plant response to infection. Two fungal transcripts, a homolog of a fungal cellulase (P07982) and a homolog of a 'fungal plant-induced rust protein' (O00057), were noted in samples taken five dpi (Hsieh et al. 2018). A recent study identified 890 differentially

expressed *A. psidii* genes in an *E. grandis* infection model but only one gene was significantly differentially expressed at 1 dpi (Swanepoel et al. 2023). No genes were identified at 12 hours post inoculation (hpi) or 2 days post inoculation (dpi). Additionally, in the above studies only the Swanepoel et al. (2021) and (2023) investigations used plants for which the myrtle rust disease resistance rating had been pre-determined. In those studies, only immune resistant and highly susceptible plants were used, and hypersensitive resistant plants were not included.

Tobias et al. (2016) proposed that the most likely mechanism for the resistance levels found in Australian Myrtaceae was 'a common Myrtaceae effector hub' that 'on modification, triggers host recognition and response' as the 'proportions of resistant plants are problematic to explain without a co-evolved selective pressure'. In this study we sought evidence to test this hypothesis using Aotearoa providence *L. scoparium* to gain an initial understanding of the molecular basis of host resistance following challenge by the pandemic strain of *A. psidii*. This study also sought to understand the basis of fungal pathogenicity on these *L. scoparium* plants for which the resistance phenotype was already known, in particular did the pathogen gene expression change after inoculation onto immune resistant, hypersensitive resistant and susceptible *L. scoparium* plants?

Materials and Methods

Leptospermum scoparium **plants and experimental design.** Plants were selected from four seed families sourced from plants from the East Cape region of Te Ika-a-Māui that had been previously assessed for leaf and stem resistance as described in Smith et al. (2020). These shade-house potted plants were cut back to remove infected material and regrown for 8 weeks as the pathogen is not systemic, as per Swanepoel et al. (2021). All selected plants were stem-infection resistant (S1). In total 24 plants from four seed families (F01, F02, F03, F04) were selected that had been leaf-resistance phenotyped as per Smith et al (2020) (L1, immune resistant; L2, hypersensitive resistant; L5, highly susceptible). Twelve were L1S1 (leaf immune resistant, stem resistant) plants, four L2S1 (leaf hypersensitive resistant, stem resistant) plants and eight L5S1 (leaf highly susceptible, stem resistant) susceptible plants (**Table 1**). Half of the plants were randomly assigned as controls, the other half were assigned to the inoculated group, with plants from each of the four families in both the inoculated and control treatments.

Inoculation. After regrowth the control plants were sprayed with the inoculation solution (two drops of Tween® 20 per 100 mL of sterile distilled water). The inoculated plants were sprayed with inoculation solution containing 1 x 10^5 *A. psidii* urediniospores per mL. The inoculated seedlings were then covered with plastic sheeting and hot tap water (60°C) was applied to the lower plastic sheet creating a sealed environment to maintain humidity and leaf wetness. The covered plants were then placed into a controlled-environment chamber in darkness at 18°C and 80% relative humidity. After 24h, the plastic coverings were removed and seedlings were transferred into a shade house and watered for 10 min twice daily as described by Smith et al. (2020).

RNA preparation and sequencing. Each plant was sampled 24 and 48 hpi by removing 8 to 12 leaves from directly below the stem apical meristem which were immediately snap frozen in liquid nitrogen and then ground in liquid nitrogen with a mortar and pestle prior to total RNA extraction using the Norgen Plant/Fungi Total RNA Purification Kit, including the optional chloroform extraction step. The RNA preparation quality was analysed by Australian Genome Research Facility (AGRF, www.agrf.org.au) Melbourne, prior to single lane Illumina NovaSeq 6000 S4 sequencing with 150 bp PE reads. Image analysis was performed in real time by the NovaSeq Control Software (NCS) v1.6.0 and Real Time Analysis (RTA) v3.4.4, running on the instrument computer. The Illumina bcl2fastq 2.20.0.422 pipeline was used to generate the sequence data for 150 bp paired end reads, (Illumina, San Diego, CA, USA).

Sequencing data quality checking, clean-up and mapping to the reference genome. Ribosomal RNA sequences were removed from the raw RNAseq data using SortMeRNA (version 2.1b; Kopylova et al.

2012). Adaptors, low-quality sequences, and homo-polymers were removed using Trimmomatic (version 0.36; Bolger et al. 2014). FastQC (version 0.11.7; Andrews 2010) and MultiQC (version 1.7; Ewels et al. 2021) were used throughout the data processing steps to ascertain the integrity of the data. Picard tools (version 2.9.4; Broad Institute https://broadinstitute.github.io/picard/) was used to add metadata to the read sets and the STAR aligner (version 2.6.1d; Dobin et al. 2013) was used to align sequence reads to the *L. scoparium* genome (Thrimawithana et al. 2019; permission obtained 29/03/2023 from Kaitiaki Māori via the Aotearoa Genomic Data Repository, https://data.agdr.org.nz/) and to the *A. psidii* genome (Tobias et al. 2021). Count data for the expressed genes were obtained using HTSeq count (version 0.9.1; Anders et al. 2014).

Differential gene expression analysis. DESeq2 (Love et al. 2014) was used for the differential expression analysis. R (version 4.0) was used for the analysis, and the results were visualised using ggplot2 (version 3.3.5; Wickham 2016). Treatment and plant family were combined into one factor referred to as combo. The DESeq2 design for the *L. scoparium* data was: design =~ time + combo + time:combo. This formula modelled the difference between the two time points, the difference between the treatments and plant families (represented by the combo factor), and any treatment-plant family differences over time. The *A. psidii* design was simpler and only the difference at the two time points was modelled (design =~ time). The false discovery rate/alpha was set to 0.1. All reported p-values are the adjusted p-values as generated by the DESeq2 analysis. Principal component analysis (PCA) was performed on the 100 genes with the smallest adjusted p-value (i.e., with the most significant changes in gene expression) after variance stabilising transformation (vst), and hierarchical clustering was performed on the data from all genes following vst.

BLAST. The predicted *L. scoparium* (Thrimawithana et al. 2019) and *A. psidii* (Tobias et al. 2021) genes were compared against sequences in the SwissProt database ("UniProt: a worldwide hub of protein knowledge", 2018), using the BLAST algorithm through the BLAST+ command line application (Camacho et al. 2009) to identify sequences with putative functions that share sequence similarity to the genes.

Pathway enrichment analysis. The predicted protein sequences from the *L. scoparium* and *A. psidii* genomes were submitted to the Kyoto Encyclopaedia of Genes and Genomes (KEGG) annotation server to obtain KEGG ontologies (Kanehisa 2019, Kanehisa and Goto 2000, Kanehisa et al. 2019). The pathway analysis was done using gage (version 2.37.0; Luo et al. 2009) and the results visualised using pathview (version 1.28.0; Luo and Brouwer 2013).

Co-expression analysis. Co-expression analysis was undertaken by creating correlation matrices, using DESeq2 results and the R package qgraph (Epskamp et al. 2012). Data were visualised using correlation graphs for selected gene groups of interest.

Carbohydrate-active enzyme identification. dbCAN3 (Zhang et al. 2018) was used to identify carbohydrate-active enzymes (CAZymes) in the *A. psidii* genome. Additional CAZymes were identified through manual analysis of gene sequences in the genome.

Results

Sequencing

RNA sequencing libraries. In total, 48 leaf samples were taken over the two time points and RNA extracted. One F02_L1S1 sample RNA extraction failed, and two RNA library creations failed (F02_L1S1 and F04_L5S1) resulting in 45 sequence libraries over the four families and two time points (**Table 1**). 353 giga base pairs (Gbp) of data containing 1.167 billion reads was created (average number of reads per sample was 25.90 million (range 22.84 to 31.24) and the average data yield per sample was 7.83 Gbp (range 6.90 to 9.43; **Supplementary Table 1**).

***Leptospermum scoparium* differential gene expression**

Changes in gene expression are influenced by *L. scoparium* family, time, and treatment. PCA and hierarchical clustering methods were used to understand the underlying substructure of the *L. scoparium* gene expression data. These analyses suggested that the differential gene expression patterns were strongly linked to sampling time, plant family, and treatment. These factors were subsequently included in the final DESeq2 design. The PCA plot from the final DESeq2 analysis design is shown in **Figure 1A-D**. The sample points are coloured by **(A)** family, **(B)** sampling time (hpi), **(C)** treatment, and **(D)** phenotype. The first principal component accounts for 26% of the variation and separates the samples by family while the second principal component accounts for 12% of the variation and sampling time and treatment appear to drive separation along this component. Phenotype does not provide clear separation with a mixed cluster of L1S1 and L2S2 samples in the centre of the plot.

Differences between uninoculated *L. scoparium* families. There are differences in gene expression between the different *L. scoparium* families even without *A. psidii* inoculation (control group). The number of significantly differentially expressed genes for each family for different comparisons are in **Table 2**. The up- and down-regulated genes for each family compared to F04_L5S1 are listed in **Supplementary File 1**. Nine genes were identified as differentially expressed in more than one family, a Venn diagram illustrating this is shown in **Figure 1E** and the nine genes differentially expressed by F01_L1/2S1, F02_L1S1, and F03_L2S1 are listed in **Table 3**. No metabolic pathways were identified as having significant changes in gene expression using gage pathway analysis.

***Leptospermum scoparium* families respond to inoculation with *A. psidii*.** The four *L. scoparium* families had different responses at both 24 and 48 hpi following inoculation with *A. psidii*. The up- and down-regulated genes for each family comparing the inoculated plants to the control plants at 24 hpi are listed in **Supplementary File 2**. The up- and down-regulated genes for each family comparing the

inoculated plants to the control plants at 48 hpi are listed in **Supplementary File 3**. Only a small number of genes (3-14) were differentially expressed at 24 hpi for all families while 1343 genes were up-regulated and 373 down-regulated by F04_L5S1 (leaf susceptible, stem resistant) at 48 hpi (**Table 2**). F01_L1/2S1 (leaf immune and hypersensitive response resistant, stem resistant) had a single up-regulated gene, F02_L1S1 (leaf immune resistant, stem resistant) had nine up-regulated and four down-regulated and no differentially expressed genes were identified for F03_L2S1 (leaf hypersensitive response resistant, stem resistant) at 48 hpi. The one up-regulated gene (jg20441.t1) from F01_L1/2S1 is also up-regulated in both the F02_L1S1 and F04_L5S1 families. The protein sequence from this gene has a top blastp match to a BTB/POZ domain-containing protein (AT2G30600) from *Arabidopsis thaliana*. There was no other overlap in differentially expressed genes at either the 24 or 48 hpi timepoints. The gene jg24539.t1 is significantly up-regulated in F03_L2S1 plants inoculated with *A. psidii* compared to control plants at 24 hpi (**Supplementary File 2**). The blastp result suggests this gene contains a leucine-rich repeat and has similarity to disease resistance protein RUN1-like isoform X1 (XP_018717429.1) and X2 (XP_018717437.1) from *E. grandis* and other disease resistance proteins found in a range of *Syzygium* species.

Austropuccinia psidii differential gene expression

Gene expression changes over time. Hierarchical clustering and PCA were used to assess the similarities between samples and to identify determinants that differentiated them. Both the heat map and PCA plot in **Figures 2A** and **2B** show sampling time to be a major differentiating factor. The 24 hpi samples are highly similar and tightly clustered, while the 48 hpi samples are less tightly clustered but are separate from the 24 hpi samples. Both putative carbohydrate-active enzymes (CAZymes) and putative effectors were present in the differential gene expression analysis of the top ten fungal genes ranked by adjusted p-value after analysis of differential gene expression between the two sampling time points (24 hpi and 48hpi) in **Figure 2C**.

Host plant phenotype or family have limited impact on fungal gene expression. The *A. psidii* gene expression data were analysed by comparing the 24 hpi to the 48 hpi sampling point. The expression profiles of both putative effectors (**Figure 2D**) and putative carbohydrate-active enzymes (CAZymes) (**Figure 2E**) change between 24 and 48 hpi. Both the effector and the CAZyme heatmaps appear to independent of the both plant family and resistance phenotype, suggesting that the changes in the pathogen expression profiles in the first 48 hours are not in response to any external criteria (e.g. a plant hyper-sensitive response). When the host plant family was included as a factor in the analysis, only 134 *A. psidii* genes were identified as significantly differentially expressed after inoculation to

plants in the F02_L1S1 and F04_L5S1 families (**Table 4**). No significant differentially expressed fungal genes were identified after inoculation to plants in the F03_L2S1 and F04_L5S1 families. When sampling time was the only factor used in the analysis, 151 up-regulated and 332 down-regulated fungal genes were identified. The DESeq2 design using only sampling time as a factor was subsequently chosen as the best representation of the data and used from here-on.

Changes in *A. psidii* gene expression at 24 hpi compared to 48 hpi. Transcripts for 16 predicted effectors were more abundant at 24 hpi compared to 48 hpi (**Figure 2D**). Twelve predicted CAZymes were more abundant at 24 hpi compared to 48 hpi. Additionally, there were three predicted CAZymes with the opposite expression pattern (APSI_P010.11473.t1, APSI_P021.13488.t1, and APSI_P013.4183.t1; **Figure 2E**). The 10 genes with the greatest up-regulation and the 10 with the greatest down-regulation at 48 hpi compared to 24 hpi are listed in **Table 5**. The full gene list is in **Supplementary File 4**. Investigation of gene pathways using gage (Luo et al. 2009) identified several up-regulated pathways relating to increased fungal growth at 48 hpi compared to 24 hpi. This included the ribosome, biosynthesis of secondary metabolites, biosynthesis of amino acids, and oxidative phosphorylation pathways. Only the ribosome pathway was significantly up-regulated (p-value 1.48×10^{-21}). No significantly down-regulated pathways were identified. A correlation analysis to identify patterns in gene analysis was performed on the 10 genes with the smallest adjusted p-value (**Figure 2C**). APSI_P010.11473.t1, a putative patulin synthase and a CAZyme, expression was negatively correlated with two other CAZymes and three effectors. The three effectors' expression levels were all positively correlated with each other.

Discussion

***Leptospermum scoparium* response to infection.** This is the first report of the differential transcriptomic response of *L. scoparium* (mānuka) plants to challenge by germinating urediniospores of *A. psidii*. Similar to previous studies (Swanepoel et al. 2021, Swanepoel et al. 2023), the resistance phenotypes of the plants in this study were previously characterised. In contrast to those studies, where only immune resistant, and highly susceptible plants were selected for transcriptomic analysis (ratings 1 and 5 respectively on the modified Junghans scale (Junghans et al. 2003a)) hypersensitive resistant plants (rating 2 on the modified Junghans scale (Smith et al. 2020)) were also included in this study. Santos et al. (2020) used clones of two well-studied plant *E. grandis* genotypes (resistant CLR385 and susceptible CLR220), whilst in other myrtle rust gene expression studies the resistance phenotype was determined during the study itself (Hsieh et al. 2018, Tobias et al. 2018) or observed at sample collection (Soewarto et al. 2019b). Additionally, in this study and that of Santos et al. (2020) the genetic relationship of the plants (seed family siblings or clones) was known (**Table 1**). For this pathogen-host study, knowing the resistance phenotype of the plants was important as *L. scoparium* has two tissue-specific resistances, leaf and stem, that are hypothesised to be genetically independent based on disease resistance phenotype distributions within, and between, seed families (Smith et al. 2020). To minimise the complexity of the experimental design and subsequent transcript analyses, only stem-infection resistant (S1) *L. scoparium* plants were selected for this study. Stem infection resistance is hypothesised to be an immune resistance, as no equivalent to L2 (leaf hypersensitive resistance) was observed during the original phenotyping (Table 2 in Smith et al. (2020)).

Six Myrtaceae have been assessed in previous myrtle rust transcriptomic studies: *E. grandis, A. gummiferum, S. longifolium*, *T. glauca, M. quinquenervia* and *S. luehmannii* (Hsieh et al. 2018, Santos et al. 2020, Soewarto et al. 2019b, Swanepoel et al. 2021, Swanepoel et al. 2023, Tobias et al. 2018). This study is the first to use *L. scoparium* and the first myrtle rust transcript study to investigate the response to fungal challenge by sibling plants from known seed families. Whilst the *L. scoparium* siblings are not biological replicates, as previously noted by Smith et al. (2020), the original seed collected from the mother plants is considered to have been open pollinated as Myrtaceae have late-acting self-incompatibility (Gibbs 2014), although self-compatibility exceptions have been noted in some Myrtaceae (Schmidt-Adam et al. 1999). Thus, the *L. scoparium* plants grown from seed from a seed family and used in this study are believed to have a common maternal genetic base with additional genetics via open pollination from one or more paternal plants.

Analysis of the top 100 *L. scoparium* significantly expressed genes confirmed that plant family was the factor with the strongest influence on the data structure (**Figure 1**). Differences between the families

were wide ranging with only nine genes common between the resistant families (F01_L1/2S1, F02_L1S1, and F03_L2S1) and the susceptible family (F04_L5S1; **Figure 1E** and **Table 3**). This finding is consistent with the results found using SNP-based methods (Chagné et al. 2023, Koot et al. 2022) where geographic populations were evident in the genomic data from *L. scoparium* plants sampled across New Zealand.

The low number of differentially expressed genes identified for F01_L1/2S1 and F03_L2S1 are likely due to the small number of samples in each group. However, it is interesting that these families cluster together in the PCA plot (**Figure 1A-D**) despite the mix of phenotypes. It is possible that the F01_L1/2S1 plants with a L1S1 phenotype also carry L2S1 immune resistance genes like their L2S1 siblings and members of the F03_L2S1 family. This L2 hypersensitive response would not be necessary or measurable if the plant also has genes for the L1 immune resistant phenotype. One gene (jg20441.t1) was up-regulated in the F01_L1/2S1, F02_L1S1, and F04_L5S1 families at 24 hpi when comparing inoculated to control plants. The predicted protein from this gene contains a BTB/POZ domain. These domains are key factors enabling protein-protein interactions and proteins with BTB/POZ domains have a wide variety of roles in cells, including transcription repression, protein ubiquination and degradation, and cytoskeleton regulation (Stogios et al. 2005). Further study of this gene and its protein will be required to understand if it has an important role in the response of *L. scoparium* to *A. psidii* infection. When inoculated with *A. psidii*, only a few genes showed a change in expression level in the immune resistant F02_L1S1 plants (**Table 2**). This suggests that the basis of this resistance is preformed and is not reliant on plant recognition of infection and response. This limited change in gene expression in leaf immune resistant plants was similar to that observed in *E. grandis* by Santos et al. (2020). The susceptible F04_L5S1 plants had numerous differentially expressed genes suggesting the plant had sensed the presence of the pathogen but ultimately the response was not successful. This contrasts with the findings of Tobias et al. (2018) and Swanepoel et al. (2021) where a greater number of differentially expressed genes were observed in the resistant plants of *S. luehmannii* and *E. grandis*, compared to the susceptible plants. The different responses between these experiments could be due to the different plant species or to the limited sampling points within each experiment. Further research investigating gene expression changes over more time points may reveal plants with different disease phenotypes responding differentially to infection in this non-coevolved pathosystem.

This study included plants with a hypersensitive resistance phenotype (rating 2 on the modified Junghans scale (Smith et al. 2020)). While some differentially expressed genes were found at 24 hpi when inoculated plants were compared to control plants, no differentially expressed genes were identified at 48 hpi. Despite the small sample size in these experiments, we were able to identify one

potential disease resistance gene that may be associated with the hypersensitive response in these plants. The gene, jg24539.t1, was identified as a potential disease resistance gene linked to the L2 hypersensitive response phenotype of the F03_L2S1 family. Similar genes have been annotated in the genomes of *E. grandis* and other *Sygyzium* species. These genes have similarity to the grape 'resistance to *Uncinula necator*' (*MrRUN1*) gene, that confers resistance to the fungal pathogen *U. necator* (synonym *Erysiphe necator*), the causal agent of powdery mildew (Feechan et al. 2013). Future studies using more plants with this phenotype could enable the full hypersensitive resistance mechanism to be elucidated. However, the low abundance of hypersensitive plants in *L. scoparium* populations will make elucidation challenging as only 2.3% of the *L. scoparium* families screened by Smith et al. (2020) contained plants with the L2S1 resistance phenotype.

No transcripts associated with other forms of disease resistance were found in these analyses. For example, other studies had found transcripts associated with glutathione S-transferases (Hsieh et al. 2018) and changes to brassinosteroid signalling (Swanepoel et al. 2021). Tobias et al. (2016) suggested that the range of responses by the different plant species challenged by *A. psidii* may reflect the lack of co-evolution between the plant hosts and this pathogen and that the most likely mechanism for the resistance levels found in Australian Myrtaceae was 'a common Myrtaceae effector hub' that 'on modification, triggers host recognition and response' as the 'proportions of resistant plants are problematic to explain without a co-evolved selective pressure'. The results from this study that each resistant *L. scoparium* family (F01, F02, F03) has a unique gene expression profile, with only nine transcripts shared by all families (**Figure 1E**), provides little support for the hypothesis of 'a common Myrtaceae effector hub' as there is significant diversity of genetic responses in this single species to fungal challenge. A much larger study with many more families and species would be required to validate or refute this hypothesis.

***Austropuccinia psidii* gene expression.** As noted previously most of the previous *A. psidii*-host studies either did not note, or reported only a limited number of, fungal transcripts in their analysis. The *E. grandis* study by Swanepoel et. al (2023) reported 890 *A. psidii* transcripts, including 43 candidate effector protein genes at 0.5, 1, 2 and 5 day sample time points, with most reads mapping to the susceptible host samples taken 5 days after inoculation. The top 10 differentially expressed genes between 24 and 48 hpi in this study were not present in the top 100 most highly expressed *A. psidii* genes (based on fragments per kilobase of transcript per million reads mapped (FPKM)), in the susceptible or resistant host infections in the *E. grandis* infection study of Swanepoel et al (2023). These two studies used different plant species, but did have two aligned sampling time points (24 h/1 day and 48 h/2 days). The different results from this study and that of Swanepoel et al (2023) reflect one of the challenges of investigating this multi-host pathosystem. The top three differentially

expressed fungal transcripts in this study encoded putative CAZymes, suggesting a key role for these proteins in enabling the pathogen to breech the plant cell wall. APSI_P001.7032 has two predicted transcripts (t1 and t2): both are predicted to be differentially expressed but further work is required to confirm this result. Blastp identified a cellulase domain in APSI_P001.7032 (pfam00150; glycosyl hydrolase family 5). Proteins with similarity to APSI_P001.7032 are conserved across a range of fungi, including *Puccinia* spp., other Basidiomycota as well as several Ascomycota species. The protein encoded by APSI_P010.11473 is also conserved across a wide range of fungi. The blastp search against the SwissProt database had a match to patulin synthase while a general blastp search detected a Rossmann-fold NAD(P)H/NAD(P)(+) binding (NADB) domain and a GMC oxidoreductase domain (pfam00732/pfam05199). This gene family is predicted to be involved in the degradation of lignocellulose (Sützl et al 2018). These putative functional domains in these CAZymes suggests they may have key roles in the initial phases of infection of the plant host.

Over 1,200 putative effector genomic sequences have been identified in the *A. psidii* genome and mapped to the two haploid genomes: 617 to haplotype 1 and 616 to haplotype 2 (Edwards et al. 2022). Effectors are relativity small proteins '...that facilitate pathogen entry into the host interior, suppress plant immune perception, and alter host physiology for pathogen benefit…' (Toruño et al. 2016). The presence of three predicted *A. psidii* effectors in the top ten differentially expressed pathogen genes across all the inoculated plants samples suggests a key role for these proteins in successfully establishing infection. Interestingly, the top ten differentially expressed genes found in this study are not present in the FPKM ranked supplementary tables S2: The top 100 most highly expressed *A. psidii* genes in the resistant interaction, nor S3: The top 100 most highly expressed *A. psidii* genes in the susceptible interaction, supplemental files of the findings of Swanepoel et al. (2023) at https://apsjournals.apsnet.org/doi/suppl/10.1094/PHYTO-07-22-0257-R/suppl_file/PHYTO-07-22-0257-R.sm2.xlsx . BLAST searches of the three effectors sequences did not identify any conserved domains therefore classic biochemical and interaction studies will be required to understand their role in the infection process. Genes with sequence similarity to APSI_P001.5292.t1 and APSI_P014.1260.t1 are present in other *Puccinia* species suggesting a conserved function. APSI_P005.10948.t1 appears to be unique to *A. psidii*.

These early expressed fungal genes found in this study are potentially critical to facilitate pathogen entry into the plant and manipulation of the plant cells for access to nutrition and for suppression of host detection and response systems. Further research to investigate this initial infection time period (up to 24 hpi) in detail is required to understand the function of these early expressed pathogen genes so that novel management techniques that directly target this pathogen can be developed.

Acknowledgements

Funding: We gratefully acknowledge the support provided by the Ministry of Business, Innovation and Employment, New Zealand via the Beyond Myrtle Rust Research Program (Endeavour C09X1806)

We acknowledge Rangitāne O Manawatū, Genomics Aotearoa, and the New Zealand eScience Infrastructure for facilitating access to the *L. scoparium* genome data set. We thank mana whenua and landowners for kindly allowing scientists and seed collectors access to their lands to receive plant samples and undertake this research.

References

Almeida, R. F., Machado, P. S., Damacena, M. B., Santos, S. A., Guimarães, L. M. S., Klopfenstein, N. B., and Alfenas, A. C. 2021. A new, highly aggressive race of *Austropuccinia psidii* infects a widely planted, myrtle rust-resistant, eucalypt genotype in Brazil. For. Pathol. 51:e12679.

Anders, S., Pyl, P. T., and Huber, W. 2014. HTSeq – A Python framework to work with high-throughput sequencing data. bioRxiv:002824.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Anonymous. 2021. Myrtle rust find reported in Christchurch. https://www.landcareresearch.co.nz/discover-our-research/biodiversity-biosecurity/ecosystem-resilience/beyond-myrtle-rust/news/myrtle-rust-find-reported-in-christchurch/

Beenken, L. 2017. *Austropuccinia*: a new genus name for the myrtle rust *Puccinia psidii* placed within the redefined family Sphaerophragmiaceae (Pucciniales). Phytotaxa 297:53.

Bolger, A. M., Lohse, M., and Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Butler, J. B., Freeman, J. S., Vaillancourt, R. E., Potts, B. M., Glen, M., Lee, D. J., and Pegg, G. S. 2016. Evidence for different QTL underlying the immune and hypersensitive responses of *Eucalyptus globulus* to the rust pathogen *Puccinia psidii*. Tree Genet. Genom. 12:1-13.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421-421.

Carnegie, A. J., and Giblin, F. R. 2014. *Austropuccinia psidii* (myrtle rust). in: CABI. https://www.cabidigitallibrary.org/doi/10.1079/cabicompendium.45846

Carnegie, A. J., and Pegg, G. S. 2018. Lessons from the incursion of myrtle rust in Australia. Annu. Rev. Phytopathol. 56:457-478.

Chagné, D., Montanari, S., Kirk, C., Mitchell, C., Heenan, P., and Koot, E. 2023. Single nucleotide polymorphism analysis in *Leptospermum scoparium* (Myrtaceae) supports two highly differentiated endemic species in Aotearoa New Zealand and Australia. Tree Genet. Genom. 19:31.

de Lange, P. J., and Schmid, L. M. H. 2021. *Leptospermum repo* (Myrtaceae), a new species from northern Aotearoa/New Zealand peat bog habitats, segregated from *Leptospermum scoparium* s. l. Ukrainian Botanical Journal 78:247-265.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21.

du Plessis, E., Granados, G. M., Barnes, I., Ho, W. H., Alexander, B. J. R., Roux, J., and McTaggart, A. R. 2019. The pandemic strain of *Austropuccinia psidii* causes myrtle rust in New Zealand and Singapore. Australas. Plant Pathol. 48:253-256.

Edwards, R. J., Dong, C., Park, R. F., Tobias, P. A. 2022. A phased chromosome-level genome and full mitochondrial sequence for the dikaryotic myrtle rust pathogen, *Austropuccinia psidii*. bioRxiv preprint doi: https://doi.org/10.1101/2022.04.22.489119; version posted April 22, 2022.

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. 2012. qgraph: Network visualizations of relationships in psychometric data. J. Statistical Software 48:1-18.

Ewels, P., Lundin, S., and Käller, M. 2021. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047-3048.

Feechan, A., Anderson, C., Torregrosa, L., Jermakow, A., Mestre, P., Wiedemann-Merdinoglu, S., Merdinoglu, D., Walker, A.R., Cadle-Davidson, L., Reisch, B., Aubourg, S., Bentahar, N., Shrestha, B., Bouquet, A., Adam-Blondon, A.-F., Thomas, M.R. and Dry, I.B. 2013. Genetic dissection of a TIR-NB-LRR locus from the wild North American grapevine species *Muscadinia rotundifolia* identifies paralogous genes conferring resistance to major fungal and oomycete pathogens in cultivated grapevine. Plant J, 76: 661-674.

Fensham, R. J., and Radford-Smith, J. 2021. Unprecedented extinction of tree species by fungal disease. Biol. Conserv. 261:109276.

Fensham, R. J., Carnegie, A. J., Laffineur, B., Makinson, R. O., Pegg, G. S., and Wills, J. 2020. Imminent extinction of Australian Myrtaceae by fungal disease. Trends Ecol. Evol. 35:554-557.

Flor, H. H. 1942. Inheritance of pathogenicity in a cross between physiologic races 22 and 24 of *Melampsora lini*. Phytopathology 32:653-669.

Flor, H. H. 1955. Host-parasite interaction in flax rust - its genetics and other implications. Phytopathology 45:680-685.

Gibbs, P. E. 2014. Late-acting self-incompatibility – the pariah breeding system in flowering plants. New Phytol. 203:717-734.

Gifford, A. 2021. Ramarama gone from Raukūmara as myrtle rust spreads. in: Waatea News. https://waateanews.com/2021/11/05/ramarama-gone-from-raukumara-as-myrtle-rust-spreads/

Heenan, P. B., McGlone, M. S., Mitchell, C. M., McCarthy, J. K., and Houliston, G. J. 2023. Genotypic variation, phylogeography, unified species concept, and the 'grey zone' of taxonomic uncertainty in kānuka: recognition of *Kunzea ericoides* (A.Rich.) Joy Thomps. sens. lat. (Myrtaceae). N. Z. J. Bot.:1-30.

Ho, W. H., Baskarathevan, J., Griffin, R. L., Quinn, B. D., Alexander, B. J. R., Havell, D., Ward, N. A., and Pathan, A. K. 2019. First report of myrtle rust caused by *Austropuccinia psidii* on *Metrosideros kermadecensis* on Raoul Island and on *M. excelsa* in Kerikeri, New Zealand. Plant Dis. 103:2128-2128.

Hsieh, J. F., Chuah, A., Patel, H. R., Sandhu, K. S., Foley, W. J., and Külheim, C. 2018. Transcriptome profiling of *Melaleuca quinquenervia* challenged by myrtle rust reveals differences in defense responses among resistant individuals. Phytopathology 108:495-509.

Junghans, D. T., Alfenas, A. C., and Maffia, L. A. 2003a. Escala de notas para quantificação da ferrugem em Eucalyptus. Fitopatol Bras 28:184-188.

Junghans, D. T., Alfenas, A. C., Brommonschenkel, S. H., Oda, S., Mello, E. J., and Grattapaglia, D. 2003b. Resistance to rust (*Puccinia psidii* Winter) in *Eucalyptus*: mode of inheritance and mapping of a major gene with RAPD markers. Theor. Appl. Genet. 108:175-180.

Kanehisa, M. 2019. Toward understanding the origin and evolution of cellular organisms. Protein Sci. 28:1947-1951.

Kanehisa, M., and Goto, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28:27-30.

Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., and Tanabe, M. 2019. New approach for understanding genome variations in KEGG. Nucleic Acids Res. 47:D590-D595.

Koot, E., Arnst, E., Taane, M., Goldsmith, K., Thrimawithana, A., Reihana, K., González-Martínez, S. C., Goldsmith, V., Houliston, G., and Chagné, D. 2022. Genome-wide patterns of genetic diversity, population structure and demographic history in mānuka (*Leptospermum scoparium*) growing on indigenous Māori land. Horticulture Res. 9:uhab012.

Kopylova, E., Noé, L., and Touzet, H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28:3211-3217.

Lorrain, C., Gonçalves dos Santos, K. C., Germain, H., Hecker, A., and Duplessis, S. 2019. Advances in understanding obligate biotrophy in rust fungi. New Phytol. 222:1190-1206.

Love, M. I., Huber, W., and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550-550.

Luo, W., and Brouwer, C. 2013. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics 29:1830-1831.

Luo, W., Friedman, M. S., Shedden, K., Hankenson, K. D., and Woolf, P. J. 2009. GAGE: generally applicable gene set enrichment for pathway analysis. BMC Bioinformatics 10:161.

Moffitt, M. C., Wong-Bajracharya, J., Shuey, L. S., Park, R. F., Pegg, G. S., and Plett, J. M. 2022. Both constitutive and infection-responsive secondary metabolites linked to resistance against *Austropuccinia psidii* (myrtle rust) in *Melaleuca quinquenervia*. Microorganisms 10:383.

Roux, J., Granados, G. M., Shuey, L., Barnes, I., Wingfield, M. J., and McTaggart, A. R. 2016. A unique genotype of the rust pathogen, *Puccinia psidii*, on Myrtaceae in South Africa. Australas. Plant Pathol. 45:645-652.

Santos, S. A., Vidigal, P. M. P., Guimarães, L. M. S., Mafia, R. G., Templeton, M. D., and Alfenas, A. C. 2020. Transcriptome analysis of *Eucalyptus grandis* genotypes reveals constitutive overexpression of genes related to rust (*Austropuccinia psidii*) resistance. Plant Mol. Biol. 104:339-357.

Schmid, L. M. H., de Lange, P. J., and Marshall, A. J. 2023. *Leptospermum hoipolloi* (Myrtaceae), a new species from Aotearoa/New Zealand, segregated from *Leptospermum scoparium* s. l. Ukrainian Botanical Journal 80:173-198.

Schmidt-Adam, G., Gould, K. S., and Murray, B. G. 1999. Floral biology and breeding system of pohutukawa (*Metrosideros excelsa*, Myrtaceae). N.Z. J. Bot. 37:687-702.

Sekiya, A., Marques, F. G., Leite, T. F., Cataldi, T. R., de Moraes, F. E., Pinheiro, A. L. M., Labate, M. T. V., and Labate, C. A. 2021. Network analysis combining proteomics and metabolomics reveals new insights into early responses of *Eucalyptus grandis* during rust infection. Front. Plant Sci. 11.

Smith, G. R., Ganley, B. J., Chagné, D., Nadarajan, J., Pathirana, R. N., Ryan, J., Arnst, E. A., Sutherland, R., Soewarto, J., Houliston, G., Marsh, A. T., Koot, E., Carnegie, A. J., Menzies, T., Lee, D. J., Shuey, L. S., and Pegg, G. S. 2020. Resistance of New Zealand provenance *Leptospermum scoparium*, *Kunzea robusta*, *Kunzea linearis*, and *Metrosideros excelsa* to *Austropuccinia psidii*. Plant Dis. 104:1771-1780.

Soewarto, J., Giblin, F., and Carnegie, A. J. 2019a. *Austropuccinia psidii* (myrtle rust) global host list. https://www.anpc.asn.au/wpcontent/uploads/2019/09/MyrtleRust_GlobalHostList_final_4.xlsx

Soewarto, J., Hamelin, C., Bocs, S., Mournet, P., Vignes, H., Berger, A., Armero, A., Martin, G., Dereeper, A., Sarah, G., Carriconde, F., and Maggia, L. 2019b. Transcriptome data from three endemic Myrtaceae species from New Caledonia displaying contrasting responses to myrtle rust (*Austropuccinia psidii*). Data in Brief 22:794-811.

Soewarto, J., Somchit, C., du Plessis, E., Barnes, I., Granados, G. M., Wingfield, M. J., Shuey, L., Bartlett, M., Fraser, S., Scott, P., Miller, E., Waipara, N., Sutherland, R., and Ganley, B. 2021. Susceptibility of native New Zealand Myrtaceae to the South African strain of *Austropuccinia psidii*: A biosecurity threat. Plant Pathol. 70:667-675.

Stewart, J. E., Ross-Davis, A. L., Graça, R. N., Alfenas, A. C., Peever, T. L., Hanna, J. W., Uchida, J. Y., Hauff, R. D., Kadooka, C. Y., Kim, M.-S., Cannon, P. G., Namba, S., Simeto, S., Pérez, C. A., Rayamajhi, M. B., Lodge, D. J., Arguedas, M., Medel-Ortiz, R., López-Ramirez, M. A., Tennant, P., Glen, M., Machado, P. S., McTaggart, A. R., Carnegie, A. J., and Klopfenstein, N. B. 2018. Genetic diversity of the myrtle rust pathogen (*Austropuccinia psidii*) in the Americas and Hawaii: Global implications for invasive threat assessments. For. Pathol. 48:e12378.

Stogios, P. J., Downs, G. S., Jauhal, J. J. Nandra, S. K., and Privé, G. G. 2005. Sequence and structural analysis of BTB domain proteins. Genome Biol 6, R82.

Sützl, L., Laurent, C.V.F.P., Abrera, A.T.,Schütz, G., Ludwig, R., and Haltrich, D. 2018. Multiplicity of enzymatic functions in the CAZy AA3 family. Appl Microbiol Biotechnol 102, 2477–2492.

Swanepoel, S., Visser, E. A., Shuey, L., and Naidoo, S. 2023. The *in planta* gene expression of *Austropuccinia psidii* in resistant and susceptible *Eucalyptus grandis*. Phytopathology. 113:1066-1076.

Swanepoel, S., Oates, C. N., Shuey, L. S., Pegg, G. S., and Naidoo, S. 2021. Transcriptome analysis of *Eucalyptus grandis* implicates brassinosteroid signaling in defense against myrtle rust (*Austropuccinia psidii*). Front. For. Global Change 4.

Thrimawithana, A. H., Jones, D., Hilario, E., Grierson, E., Ngo, H. M., Liachko, I., Sullivan, S., Bilton, T. P., Jacobs, J. M. E., Bicknell, R., David, C., Deng, C., Nieuwenhuizen, N., Lopez-Girona, E., Tobias, P. A., Morgan, E., Perry, N. B., Lewis, D. H., Crowhurst, R., Davies, K. M., Chagné, D., and Schwinn, K. E. 2019. A whole genome assembly of *Leptospermum scoparium* (Myrtaceae) for mānuka research. N.Z. J. Crop Hortic. Sci. 47:233-260.

Tobias, P. A., Guest, D. I., Külheim, C., and Park, R. F. 2018. *De novo* transcriptome study identifies candidate genes involved in resistance to *Austropuccinia psidii* (myrtle rust) in *Syzygium luehmannii* (riberry). Phytopathology 108:627-640.

Tobias, P. A., Guest, D. I., Külheim, C., Hsieh, J.-F., and Park, R. F. 2016. A curious case of resistance to a new encounter pathogen: myrtle rust in Australia. Mol. Plant Pathol. 17:783-788.

Tobias, P. A., Schwessinger, B., Deng, C. H., Wu, C., Dong, C., Sperschneider, J., Jones, A., Luo, Z., Zhang, P., Sandhu, K., Smith, G. R., Tibbits, J., Chagné, D., and Park, R. F. 2021. *Austropuccinia psidii*, causing myrtle rust, has a gigabase-sized genome shaped by transposable elements. G3 Genes|Genomes|Genetics 11:jkaa015.

Toruño, T.Y., Stergiopoulos I., Coaker G. 2016. Plant-pathogen effectors: cellular probes interfering with plant defenses in spatial and temporal manners. Annu. Rev. Phytopathol. 54:419–441.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Winter, G. 1884. Repertorium. Rabenhorstii fungi europaei et extraeuropaei. Cent XXXI et XXXII. Hedwigia 23:164-172.

Xia, C., Qiu, A., Wang, M., Liu, T., Chen, W., and Chen, X. 2022. Current status and future perspectives of genomics research in the rust fungi. Int J Mol Sci. 23(17):9629.

Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., Busk, P. K., Xu, Y., and Yin, Y. 2018. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 46:W95-W101.

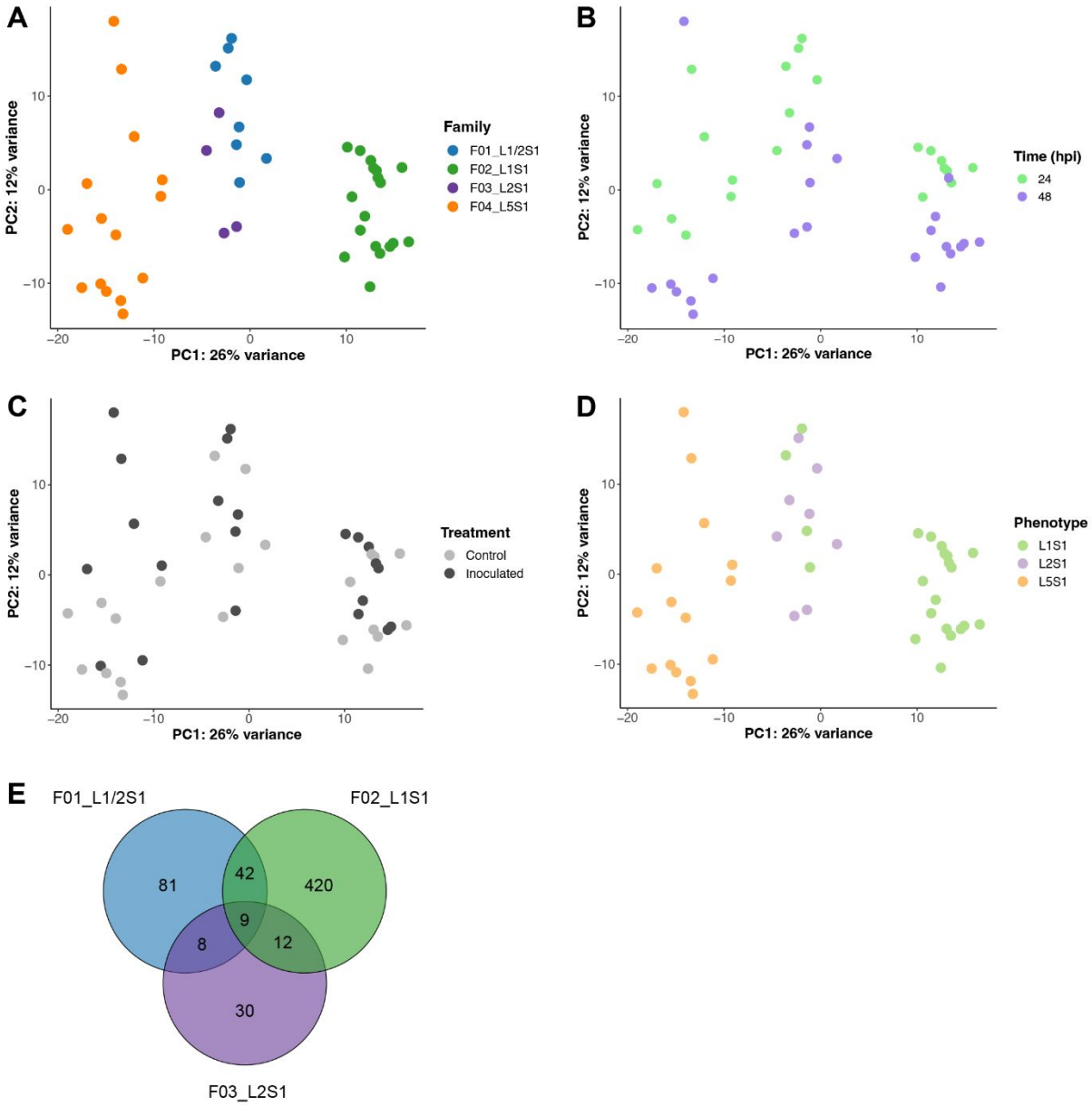

Figure 1. Differentially expressed *Leptospermum scoparium* genes are linked to family, sampling time, and treatment. Principal component analysis (PCA) plots were used to understand *L. scoparium* sample similarity and underlying substructures in the data. The PCA plot was generated using variance-stabilising transformation of gene expression count data from the 100 genes with the most significant changes in gene expression when comparing samples from 48 and 24 hours post inoculation (hpi). The sample points are coloured based on **(A)** family **(B)** sampling time (hpi), **(C)** treatment (inoculated or control), and **(D)** phenotype (L1, immune leaf resistance; L2 hypersensitive leaf resistance; L5 leaf susceptible, S1 stem resistance). The samples separate based on family along the first principal component axis and sampling time along the second principal component axis. Treatment also appears to contribute to the variance in the second principal component. The families are F01, F02, F03, F04. The phenotypes are: L1S1 leaf immune resistance, stem resistant; L2S1 leaf hypersensitive resistant, stem resistant and L5S1 leaf highly susceptible, stem resistant. Family F01 contains both L1S1 and L2S1 plants. **(E)** Venn diagram showing the number of common differentially expressed

genes across families with resistance to *Austropuccinia psidii* when compared to the susceptible family (F04_L5S1). Each family has a unique gene expression profile. Nine transcripts are shared by all families. These differences in gene expression were observed in the control plants that were not inoculated with *A. psidii*.

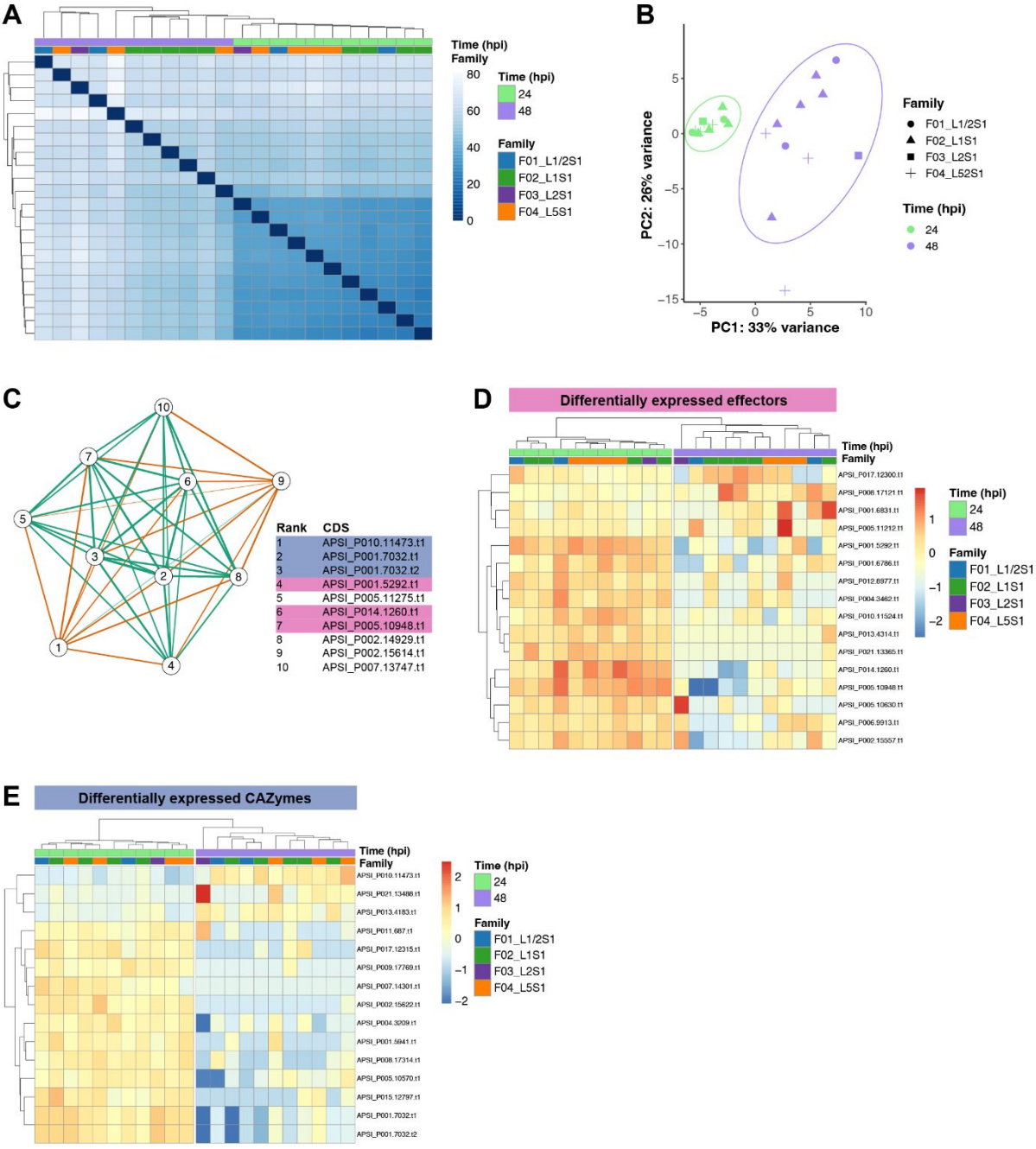

Figure 2. Identification of differentially expressed *Austropuccinia psidii* genes. (A) Hierarchical clustering and heatmap of *A. psidii* gene expression data. Analysis of *A. psidii* sample similarity and underlying substructures in the data was performed using variance-stabilising transformation of gene expression count data. The colour gradient corresponds to the correlation of gene expression for sample pairs. The samples cluster together based on the sampling time. **(B)** Principal component analysis (PCA) plot generated using variance-stabilising transformation of gene expression count data from the 100 *A. psidii* genes with the most significant changes in gene expression when comparing samples from 48 and 24 hours post inoculation (hpi). The sample points are coloured based on sampling time (hpi) and the shapes depict the different families. The samples cluster based on sampling time. **(C)** Correlation network diagram of the top 10 *A. psidii* genes ranked by adjusted p-

value following differential gene expression analysis comparing 24 and 48 hpi. No cut-off was applied to the fold change. Green edges represent positive correlations, orange edges represent negative correlations, and the edge width represents the strength of the correlation. Genes highlighted blue in the legend are carbohydrate-active enzymes (CAZymes) and genes highlighted pink are effectors (as identified in Tobias et al. (2021)). The expression profiles of **(D)** effectors and **(E)** carbohydrate-active enzymes (CAZymes) change over time. The variance-stabilised count data for effector and CAZyme genes with significant changes in gene expression at 48 hpi compared to 24 hpi was used in a hierarchical cluster analysis that is depicted in the heatmaps. The family (F01, F02, F03, F04) phenotypes are: L1S1 leaf immune resistant, stem resistant; L2S1 leaf hypersensitive resistant, stem resistant and L5S1 leaf highly susceptible, stem resistant.

Table 1. The assignment of *Leptospermum scoparium* seed family plants to treatments and successful RNAseq libraries based on pre-determined leaf infection resistance phenotype. Seed families, F01, F02, F03, F04. Leaf resistance phenotype L1, immune resistant; L2, hypersensitive resistant; L5, highly susceptible. Stem resistance phenotype S1, resistant.

Phenotype	**Treatment**	**Family**	**Number of Plants**	**Number of RNAseq Libraries (24 hpi)**	**Number of RNAseq Libraries (48 hpi)**
Immune resistant (L1S1)	Inoculated	F01_L1/L2S1	1	1	1
		F02_L1S1	5	4	5
	Control	F01_L1/L2S1	1	1	1
		F02_L1S1	5	4	5
Hypersensitive resistant (L2S1)	Inoculated	F03_L2S1	1	1	1
		F01_L1/L2S1	1	1	1
	Control	F03_L2S1	1	1	1
		F01_L1/L2S1	1	1	1
Susceptible (L5S1)	Inoculated	F04_L5S1	4	4	3
	Control	F04_L5S1	4	4	4
Total			24	22	23

Table 2. Number of significantly up- and down-regulated genes in *Leptospermum scoparium*. Results for several comparisons were extracted from the DESeq2 analysis. Significant genes were defined as genes with a ≥ 2-fold change in gene expression with an adjusted p-value < 0.1. Seed families, F01, F02, F03, F04. Leaf resistance phenotype L1, immune resistant; L2, hypersensitive resistant; L5, highly susceptible. Stem resistant phenotype S1, resistant. The F01 family contained both L1 and L2 plants, whilst the F04 family contained only L5 plants.

Comparison	Supplementary File	Family	Number of up-regulated genes	Number of down-regulated genes
Uninoculated plants compared to F04_L5S1 (susceptible) at 24 h	1	F01_L1/2S1	71	69
		F02_L1S1	222	261
		F03_L2S1	23	36
Inoculated versus uninoculated plants at 24 h	2	F01_L1/2S1	6	4
		F02_L1S1	3	5
		F03_L2S1	14	6
		F04_L5S1	6	6
Inoculated versus uninoculated plants at 48 h	3	F01_L1/2S1	1	0
		F02_L1S1	9	4
		F03_L2S1	0	0
		F04_L5S1	1343	372

Table 3. *Leptospermum scoparium* genes that are differentially expressed by F01_L1/2S1, F02_L1S1, and F03_L2S1 compared to F04_L5S1 at 24 hours post inoculation in uninoculated plants. Seed families, F01, F02, F03, F04. Leaf resistance phenotypes: L1, immune resistant; L2, hypersensitive resistant; L5, highly susceptible. Stem resistant phenotype S1, resistant. The F01 family contained both L1 and L2 plants, whilst the F04 family only contained L5 plants. A putative gene function is listed based on the top blastp match using a SwissProt database.

CDS	Top blastp match from SwissProt database	E-value
jg27885.t1	SNF2 domain-containing protein CLASSY 2	0.0
jg6587.t1	Potassium transporter 10	0.29
jg7658.t1	Aspartic proteinase CDR1	$2.17x10^{-102}$
jg8213.t1	Alginate lyase	$9.95x10^{-178}$
jg11269.t1	Putative clathrin assembly protein At5g35200	0.0
jg3844.t1	Coiled-coil domain-containing protein 124	$4.28x10^{-19}$
jg23001.t1	Probable pectinesterase 8	$4.48x10^{-165}$
jg5220.t1	Probable protein phosphatase 2C 4	1.9
jg15853.t1	E3 ubiquitin-protein ligase RFI2	$2.01x10^{-61}$

Table 4. Number of significantly up- and down-regulated genes in *Austropuccinia psidii*. One analysis included host plant family and time as factors while a second analysis only included time as a factor. Significant genes were defined as genes with a ≥ 2-fold change in gene expression at 48 hours post inoculation (hpi) compared to 24 hpi with an adjusted p-value < 0.1.

Analysis (48 hpi compared to 24 hpi)	**Samples**	**Number of up-regulated genes**	**Number of down-regulated genes**
With family and time as factors	F01_L1/2S1	0	0
	F02_L1S1	21	38
	F03_L2S1	0	0
	L04_L5S1	53	22
With time as a factor	All samples	151	332

Table 5. The 10 *Austropuccinia psidii* genes with the greatest up- and down-regulation in expression at 48 hours post inoculation (hpi) compared to 24 hpi. A putative gene function is listed based on the top blastp match using a SwissProt database.

CDS	Top blastp match from SwissProt database	Log_2 Fold Change	Adjusted p-value
Up-regulated			
APSI_P020.4955.t1	4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate synthase	7.01	1.7E-05
APSI_P005.11227.t1	Trigger factor	6.46	1.8E-04
APSI_P004.3809.t1	Sodium/glucose cotransporter 4	4.91	1.2E-02
APSI_P021.13488.t1	Levanase	4.82	7.5E-04
APSI_P003.2293.t1	Endoribonuclease dcr-1	4.73	2.3E-03
APSI_P007.14480.t1	Protein adenylyltransferase SelO	4.66	3.3E-02
APSI_P007.14353.t1	Glutamate racemase	4.39	4.1E-02
APSI_P017.12307.t1	40S ribosomal protein S0	4.17	1.6E-05
APSI_P012.8820.t1	Calcium-transporting ATPase 1, plasma membrane-type	4.16	1.1E-05
APSI_P020.4940.t1	RNA-binding post-transcriptional regulator cip2	4.08	5.3E-03
Down-regulated			
APSI_P007.14121.t1	Zinc finger CCCH domain-containing protein 4	-3.70	1.5E-05
APSI_P009.17743.t1	UPF0307 protein AHA_3937	-3.56	9.3E-05
APSI_P002.15622.t1	Mannosyl-oligosaccharide alpha-1,2-mannosidase	-3.44	8.9E-05
APSI_P013.4407.t1	Proline—tRNA ligase	-3.12	1.3E-06
APSI_P002.15849.t1	Probable serine/threonine-protein kinase PBL19	-3.09	5.1E-04
APSI_P004.2910.t1	Putative pre-16S rRNA nuclease	-3.08	4.8E-04
APSI_P021.13365.t1	Ribonuclease PH*	-3.05	1.0E-03
APSI_P002.15017.t1	Chorion transcription factor Cf2	-3.05	1.6E-05
APSI_P015.13082.t1	Protein-L-isoaspartate O-methyltransferase	-2.97	5.7E-05
APSI_P007.14301.t1	Pectinesterase	-2.90	1.8E-02

***** identified as an effector by Tobias et al. (2021).

Supplementary Data

Supplementary Table 1: Summary of raw sequencing data from RNA sequencing of *Leptospermum scoparium* inoculated with *Austropuccinia psidii.*

RNA library ID (family_phenotype-sampling time (hours)-rep)	Phenotype*	Number of sequencing reads	Data yield (giga base pairs)
F01_L1/2S1-24-1	L2S1	24,844,699	7.50
F01_L1/2S1-24-2	L1S1	31,236,353	9.43
F01_L1/2S1-24-3	L2S1	25,349,719	7.66
F01_L1/2S1-24-4	L1S1	27,770,460	8.39
F01_L1/2S1-48-1	L2S1	22,844,016	6.90
F01_L1/2S1-48-2	L1S1	24,675,087	7.45
F01_L1/2S1-48-3	L2S1	28,968,214	8.75
F01_L1/2S1-48-4	L1S1	26,778,752	8.09
F02_L1S1-24-1	L1S1	25,571,216	7.72
F02_L1S1-24-2	L1S1	24,151,296	7.29
F02_L1S1-24-3	L1S1	24,127,316	7.29
F02_L1S1-24-4	L1S1	24,125,363	7.29
F02_L1S1-24-5	L1S1	24,774,761	7.48
F02_L1S1-24-6	L1S1	24,582,202	7.42
F02_L1S1-24-7	L1S1	25,082,451	7.57
F02_L1S1-24-8	L1S1	25,133,201	7.59
F02_L1S1-48-1	L1S1	23,946,707	7.23
F02_L1S1-48-2	L1S1	26,840,251	8.11
F02_L1S1-48-3	L1S1	24,283,215	7.33
F02_L1S1-48-4	L1S1	30,322,405	9.16
F02_L1S1-48-5	L1S1	23,637,777	7.14
F02_L1S1-48-6	L1S1	26,651,415	8.05
F02_L1S1-48-7	L1S1	23,650,804	7.14
F02_L1S1-48-8	L1S1	26,720,209	8.07
F02_L1S1-48-9	L1S1	24,202,557	7.31
F02_L1S1-48-10	L1S1	26,672,290	8.06
F03_L2S1-24-1	L2S1	25,872,989	7.81
F03_L2S1-24-2	L1S1	28,707,603	8.67
F03_L2S1-48-1	L2S1	28,683,809	8.66
F03_L2S1-48-2	L2S1	26,431,364	7.98
F04_L5S1-24-1	L5S1	27,893,400	8.42
F04_L5S1-24-2	L5S1	24,989,932	7.55
F04_L5S1-24-3	L5S1	22,929,891	6.92
F04_L5S1-24-4	L5S1	29,622,616	8.95
F04_L5S1-24-5	L5S1	24,921,960	7.53
F04_L5S1-24-6	L5S1	26,480,051	8.00
F04_L5S1-24-7	L5S1	25,795,846	7.79
F04_L5S1-24-8	L5S1	25,551,598	7.72

F04_L5S1-48-1	L5S1	27,337,290	8.26
F04_L5S1-48-2	L5S1	28,405,721	8.58
F04_L5S1-48-3	L5S1	25,301,121	7.64
F04_L5S1-48-4	L5S1	25,155,384	7.60
F04_L5S1-48-5	L5S1	24,518,878	7.40
F04_L5S1-48-6	L5S1	25,979,209	7.85
F04_L5S1-48-7	L5S1	25,700,671	7.76
Total		1,167,222,069	353

Note: F02_L1S1-24-9, F04_L5S1-48-8 library creation failed, F02_L1S1-24-10 RNA preparation failed.

*****The phenotypes are: L1S1 leaf immune resistant, stem resistant; L2S1 leaf hypersensitive resistant, stem resistant and L5S1 leaf highly susceptible, stem resistant.