



# Crop-Zone Weed Mycobiomes of the South-Western Australian Grain Belt

Pippa J. Michael<sup>1</sup>, Darcy Jones<sup>1</sup>, Nicole White<sup>2</sup>, James K. Hane<sup>1\*</sup>, Michael Bunce<sup>2</sup> and Mark Gibberd<sup>1</sup>

<sup>1</sup> Centre for Crop and Disease Management, School of Molecular and Life Sciences, Curtin University, Perth, WA, Australia,

<sup>2</sup> TRENDLab, School of Molecular and Life Sciences, Curtin University, Perth, WA, Australia

## OPEN ACCESS

### Edited by:

Dhanushka Nadeeshan  
Wanasinghe,  
Kunming Institute of Botany, China

### Reviewed by:

Sajeewa S. N.  
Maharachchikumbura,  
University of Electronic Science  
and Technology of China, China  
Mathabatha Evodia Setati,  
Stellenbosch University, South Africa

### \*Correspondence:

James K. Hane  
james.hane@curtin.edu.au

### Specialty section:

This article was submitted to  
Fungi and Their Interactions,  
a section of the journal  
Frontiers in Microbiology

Received: 09 July 2020

Accepted: 28 October 2020

Published: 24 November 2020

### Citation:

Michael PJ, Jones D, White N,  
Hane JK, Bunce M and Gibberd M  
(2020) Crop-Zone Weed Mycobiomes  
of the South-Western Australian Grain  
Belt. *Front. Microbiol.* 11:581592.  
doi: 10.3389/fmicb.2020.581592

In the absence of a primary crop host, secondary plant hosts may act as a reservoir for fungal plant pathogens of agricultural crops. Secondary hosts may potentially harbor heteroecious biotrophs (e.g., the stripe rust fungus *Puccinia striiformis*) or other pathogens with broad host ranges. Agricultural grain production tends toward monoculture or a limited number of crop hosts over large regions, and local weeds are a major source of potential secondary hosts. In this study, the fungal phyllospheres of 12 weed species common in the agricultural regions of Western Australia (WA) were compared through high-throughput DNA sequencing. Amplicons of D2 and ITS were sequenced on an Illumina MiSeq system using previously published primers and BLAST outputs analyzed using MEGAN. A heatmap of cumulative presence-absence for fungal taxa was generated, and variance patterns were investigated using principal components analysis (PCA) and canonical correspondence analysis (CCA). We observed the presence of several major international crop pathogens, including basidiomycete rusts of the *Puccinia* spp., and ascomycete phytopathogens of the *Leptosphaeria* and *Pyrenophora* genera. Unrelated to crop production, several endemic pathogen species including those infecting Eucalyptus trees were also observed, which was consistent with local native flora. We also observed that differences in latitude or climate zones appeared to influence the geographic distributions of plant pathogenic species more than the presence of compatible host species, with the exception of Brassicaceae host family. There was an increased proportion of necrotrophic Ascomycete species in warmer and drier regions of central WA, compared to an increased proportion of biotrophic Basidiomycete species in cooler and wetter regions in southern WA.

**Keywords:** mycobiome, phyllosphere, fungi, plant pathogen, weeds

## INTRODUCTION

In the absence of a primary crop host, secondary plant hosts may act as a reservoir for fungal plant pathogens that cause economically significant crop diseases. In a crop monoculture environment, readily available secondary hosts are commonly weed species (Narayanasamy, 2011), which may serve as important sources of inoculum for agricultural crops that are sown or planted in the following season. Outbreaks of crop diseases caused by fungi are typically monitored based on observations of disease symptoms or fungal spore counts, making early preventative control

measures (i.e., fungicide application) difficult or impossible to implement (Lindahl and Kuske, 2013). Epidemiological studies have highlighted the need for eradication of potential sources of inoculum (i.e., weeds) between growing seasons so as to restrict the incidence and spread of disease to newly planted crops (Narayanasamy, 2011). These studies have focused on a small number of pathogen species infecting major crops; however, little is known of the larger pool of fungal species that co-inhabit the plant phyllosphere (leaf surface and interior). Pathogens constitute a relatively minor proportion of a fungal community—or mycobiome (Sapkota et al., 2015; Donovan et al., 2018). Limited information on the composition of plant mycobiomes is available, with the majority of mycobiome surveys focusing on soil and human body environments (Donovan et al., 2018). Plant pathogens are not commonly featured, despite occasional reports in non-plant mycobiomes (Weyrich et al., 2017; Donovan et al., 2018). A study by Sapkota et al. (2015) found most fungal species to be common to all host cereal crops, with the exception of a few crop-specific pathogens, indicating that the majority of fungi in leaves are not host-specific and that host-specific pathogens live in a “sea” of non-specific fungi. How plant-associated mycobiomes react to inter-species interactions, environmental conditions, host genotype, and agronomic practices is largely unexplored.

Over 99,000 fungal species are documented; however, total diversity is estimated to be between 2.2 and 3.8 million species (Hawksworth and Lücking, 2017). Traditional culture-dependent studies have revealed an immense diversity of fungal communities colonizing the plant phyllosphere, although more recent metagenomic studies indicate that this may be underestimated (Peršoh, 2015). Phyllosphere mycobiomes are highly variable among leaves within an individual plant, with factors such as leaf position, canopy height, and leaf age shown to influence leaf-to-leaf variability (Kinkel, 1997). An overlap between fungi found in the phyllosphere and the air spora has also been reported (Levetin and Dorsey, 2006). A study on beech tree (*Fagus sylvatica*) looking at variability at four different spatial scales (tree, branch, group of leaves, and individual leaf) found the majority of variation occurred at the smallest spatial scale (i.e., between individual leaves), with intra-host variability of phyllosphere fungal populations distinctly greater than inter-host variability (Cordier et al., 2012). However, within a single tree canopy, mycobiome profiles become more similar with decreasing distance, suggesting that these differences may be minimized for a single sapling when it is small and has only a few leaves. When analyzing differences between individual trees, dissimilarity between mycobiomes was linked with genetic rather than geographic distance between trees (Cordier et al., 2012). Several studies have also found that genetic makeup of the plant host at both the species and cultivar/ecotype level was the major factor influencing fungal diversity on plant leaves (Joshee et al., 2009; Bálint et al., 2013; Hunter et al., 2015; Sapkota et al., 2015), with spatial and seasonal factors also having significant but lesser impacts. While Sapkota et al. (2015) found crop genotype was the principal factor in explaining mycobiome diversity in cereal phyllospheres, within each of the individual crops studied (wheat, winter, and spring barley), location also played an important role.

Blixt et al. (2010) found that variation of community composition was greater between fields of diseased wheat (*Triticum aestivum*) than within fields as did Zimmerman and Vitousek (2012) who found that among-site diversity of a single tree species (*Metrosideros polymorpha*) contributed more than within-site diversity to the overall fungal community richness. In addition to spatial, temporal, and genetic factors, phyllosphere mycobiomes are also influenced by management factors, with fungicide use shown to have an impact on the composition of cereal leaf mycobiomes (Karlsson et al., 2014; Sapkota et al., 2015).

In this study, we identified the plant pathogens present in the phyllosphere mycobiomes of commonly found weed species adjacent to cropping fields. We also report differences in weed mycobiome composition in response to host species and spatial/climatic factors. This information on weed host-specific or region-specific association with pathogen species may become an increasingly important factor in developing new methods for crop disease management in the future.

## MATERIALS AND METHODS

### Field Sampling

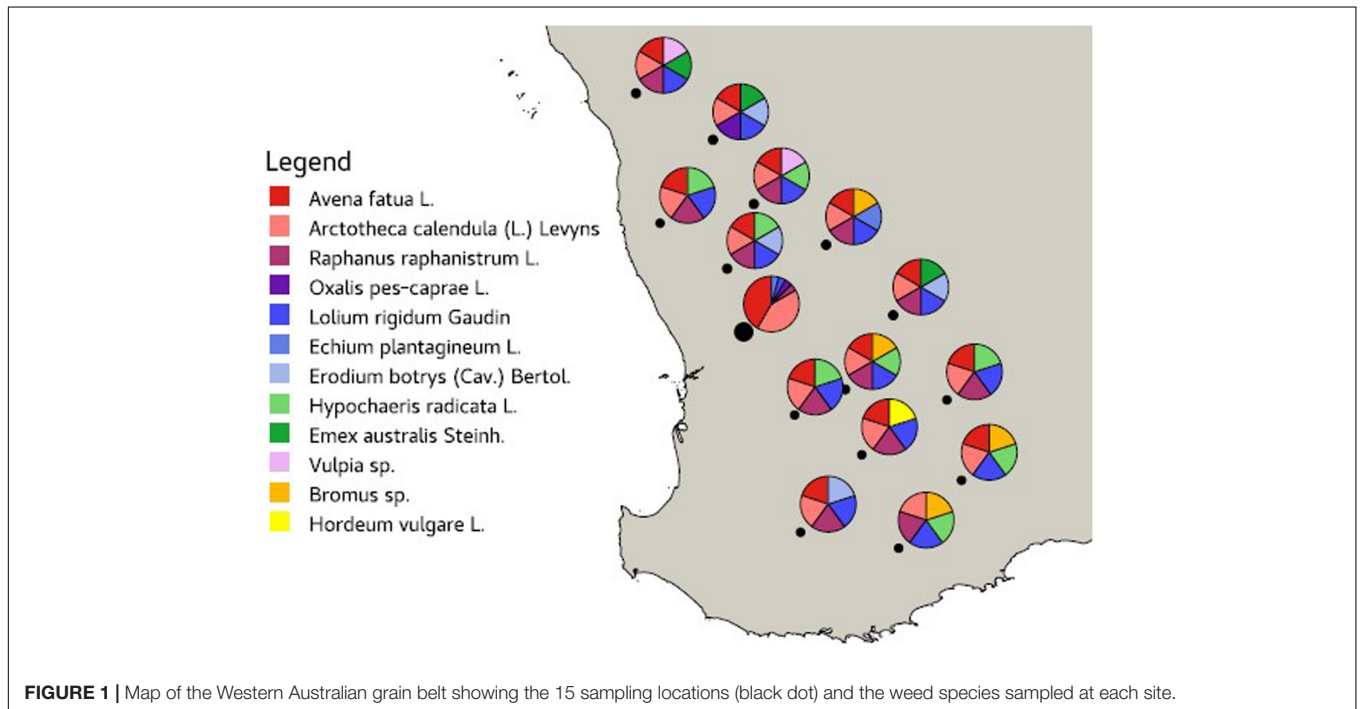
Leaves from 12 common agricultural weed species (Table 1) were sampled from 15 locations across the Western Australian grain belt (Figure 1) over a 5-day period during autumn, 2016. All sites were located on road verges adjacent to cropping paddocks and selected for maximum number of weed species present. Environmental and spatial data were recorded for each site (Table 2), with climatic data from 1950 to 2020. The optimal number of individual plants required per site to be representative of the fungal biota present needed to be determined. Therefore, 10 leaves of two weed species (*Lolium rigidum* and *Raphanus raphanistrum*) were sampled from one site (Yoting) and the preliminary metagenome sequencing results were used to benchmark the appropriate number of leaves for sequencing (Supplementary Figure 1). Thus, for each weed species sampled at each site, one uppermost leaf from 10 individual plants were stored at 4°C during transport to the lab (24–48 h), then at –20°C. Each set of 10 leaves was randomly sub-sampled using a 5-mm disc punch and combined in a 5-ml sample vial until full prior to DNA sequencing.

### DNA Extraction and Quantification

DNA was extracted using a QIAmp Plant Mini Kit (Qiagen, Venlo) with a modified protocol including (1) 600 µl of API digestion buffer; (2) overnight sample digestion at 65°C; and (3) the following day, addition of 195 µl of P3 to digests prior to putting on ice. A total volume of 400 µl was placed in a QIAcube, after which the remaining QIAmp Plant Mini Kit protocol was followed. Genomic DNA (gDNA) extracts were eluted in 100 µl of AE buffer and stored at –20°C, and then quantified and assessed for quality via real-time quantitative polymerase chain reaction (qPCR) at three dilutions (1:1, 1:10, and 1:100). Primer pairs were designed to amplify a highly conserved region of the 26S gene between the D1/D2 domains (Q1\_F 5' GTTGTTTGGGAATGCAGCTC 3' and QB3\_R 5'

**TABLE 1** | Sampling incidence of 12 common weed species from 15 locations within the Western Australian grain belt.

Species	Family	Group	Common name	No. locations sampled
<i>Arctotheca calendula</i> L.	Asteraceae	Dicotyledon	Capeweed	14
<i>Avena fatua</i> L.	Poaceae	Monocotyledon	Wild oats	13
<i>Bromus</i> spp.	Poaceae	Monocotyledon	Brome grass	4
<i>Echium plantagineum</i> L.	Boraginaceae	Dicotyledon	Patterson's curse	2
<i>Emex australis</i> Steinh.	Polygonaceae	Dicotyledon	Doublegee	3
<i>Erodium</i> spp.	Geraniaceae	Dicotyledon	Erodium	4
<i>Hordeum</i> spp.	Poaceae	Monocotyledon	Barley grass	1
<i>Hypochoeris radicata</i> L.	Asteraceae	Dicotyledon	Flatweed	8
<i>Lolium rigidum</i> Gaud	Poaceae	Monocotyledon	Annual ryegrass	14
<i>Oxalis pes-caprae</i> L.	Oxalidaceae	Dicotyledon	Soursob	1
<i>Raphanus raphanistrum</i> L.	Brassicaceae	Dicotyledon	Wild radish	12
<i>Vulpia</i> spp.	Poaceae	Monocotyledon	Silvergrass	2

**FIGURE 1** | Map of the Western Australian grain belt showing the 15 sampling locations (black dot) and the weed species sampled at each site.

AGTGCTTTTCATCTTTCCCTCAC 3'). qPCR was performed in 25- $\mu$ l reactions containing 1  $\times$  PCR Gold Buffer, 2.5 mM MgCl<sub>2</sub>, 0.4 mg/ml BSA, 0.25 mM of each dNTP, 0.4  $\mu$ M of forward and reverse primer, 0.25  $\mu$ l of AmpliTaq Gold, 0.6  $\mu$ l of SYBR Green, and 2  $\mu$ l of gDNA. The qPCR cycling conditions included an initial heat denaturation at 95°C for 5 min, 40 cycles of 95°C for 30 s, 52°C for 30 s, and 72°C for 45 s, and then a final extension at 72°C for 10 min. From the qPCR results, an optimal DNA concentration was selected for DNA sequencing, which was free of inhibition and yielded DNA of sufficient quality, as reported to facilitate reproducible quantitative data by (Murray et al., 2011).

### High-Throughput DNA Sequencing

The D2 and ITS2 amplicons were sequenced on an Illumina MiSeq system utilizing previously published primers that were

modified with a unique 8-bp Multiplex Identifier tag (MID-tag) and MiSeq adaptors for paired-end sequencing. For the D2 domain primers, U1\_F (Putignani et al., 2008) and NL4\_R (Kurtzman and Robnett, 1998) were utilized, and for the ITS2 region, fITS7\_F (Ihrmark et al., 2012) and ITS4\_R (White et al., 1990) were used. Independent MID-tagged qPCR setup for samples and controls were prepared in a physically separate ultra-clean laboratory and were carried out using each primer set in 25- $\mu$ l reactions containing 1  $\times$  PCR Gold Buffer, 2.5 mM MgCl<sub>2</sub>, 0.4 mg/ml BSA, 0.25 mM of each dNTP, 0.4  $\mu$ M of forward and reverse MID-tag primer, 0.25  $\mu$ l of AmpliTaq Gold, 0.6  $\mu$ l of SYBR Green, and 2  $\mu$ l of gDNA. The cycling conditions for qPCR using the U1\_F/NL4\_R (52°C annealing) and fITS7\_F/ITS4\_R (54°C annealing) primer sets were as follows: initial heat denaturation at 95°C for 5 min, followed by 40

**TABLE 2** | Spatial and climatic details (averaged over 1950–2020) for each location.

Location	Latitude (°S)	Longitude (°E)	Annual rainfall (mm)	Annual mean temp (°C)	Annual min temp (°C)	Annual max temp (°C)	No. weeds sampled
Badgingarra	−30.2111	115.4599	521	19.2	12.3	25.6	5
Borden	−34.0256	118.2634	377	15.8	9.8	21.7	5
Buntine	−29.9852	116.5618	319	19.4	12.2	26.5	6
Harrismith	−32.9292	117.8300	356	16.5	9.6	22.8	5
Hyden	−32.2848	118.8303	342	17.1	9.9	24.5	5
Kalannie	−30.4651	117.4111	307	19.2	12.0	26.0	6
Kojonup	−33.8383	117.1103	505	15.5	9.1	21.6	5
Morawa	−29.0817	115.9865	308	20.0	12.8	27.6	6
Mullewa	−28.6862	115.1780	352	19.6	13.0	26.6	6
Newdegate	−33.2284	119.0020	357	16.0	9.5	22.9	5
Nungarin	−31.2907	118.1990	296	18.2	11.4	25.3	6
Pingelly	−32.4618	117.0406	427	16.7	9.9	23.7	5
Toodyay	−31.4870	116.4431	485	18.1	11.2	24.9	3
Walebing	−30.7444	116.2491	432	18.3	11.6	25.1	6
Yoting	−32.1621	117.6333	325	17.4	10.2	24.2	4

Data sourced from <https://www.longpaddock.qld.gov.au/silo/>.

cycles of 95°C for 30 s; 52°C or 54°C for 30 s (annealing step); and 72°C for 45 s followed by final extension at 72°C for 10 min. Multiplex Identifier-tagged PCR amplicons were generated in duplicate for each sample and pooled together to minimize the effects of PCR stochasticity. The pooled amplicons were quantified on a LabChip (with high-sensitivity chip) and then combined to produce a final library of equimolar ratio per sample. The final library was then purified on a PippinPrep with a size selection gate of 300–600 base pair capture following the manufacturer's protocol (PerkinElmer). The purified library was diluted with purified water and re-run on the LabChip to determine the volume required (2 nM) for Illumina MiSeq paired-end sequencing. For each MID-tagged qPCR assay, extraction and PCR controls were included, and if found to contain amplifiable DNA, these reactions were incorporated into the pooled MID-tagged DNA sequencing library. Illumina MiSeq sequencing was performed using a MiSeq Reagent Kit v2 (500 cycles) 250-bp paired-end protocol as per the manufacturer's instruction. Paired-end reads were stitched using Illumina's MiSeq Reporter software.

## DNA Sequence Quality Filtering and Analyses

Stitched sequences were separated from unstitched, which had low overall phred scores and were discarded. MID-tags and gene-specific primers were trimmed from the sequences allowing for no mismatch in length or base composition using QIIME1 (extract\_barcode.py) and de-multiplexed using QIIME2 (Caporaso et al., 2010). De-multiplexed reads were then trimmed and filtered using v1.9 ( $\leq 6$  undetermined bases, quality cutoffs: 5' = 25 and 3' = 22) (Martin, 2011). Reads were then error-corrected using deblur (Amir et al., 2017) within QIIME2 (Caporaso et al., 2010), using a custom database of ITS [combining

UNITE (Kõljalg et al., 2013), ITS2 (Ankenbrand et al., 2015), and NCBI bioproject: PRJNA177353] and LSU [combining RDP (Cole et al., 2013) and NCBI bioproject: PRJNA51803] sequences, but discarding those below 258 bp for ITS and 240 bp for D2 sequences. Similar sequences to these clustered consensus sequences in the NCBI nt and GSS databases were also found using BLAST + v2.2.6 (blastn). The BLAST databases were filtered before searching to exclude sequences with NCBI taxonomic IDs in the subtree below “12908” (unknown and environmental samples) or containing the keyword “Uncultured” in the sequence name.

## Data Analysis

BLAST outputs were analyzed using MEGAN v6.11.2 (Huson et al., 2016) (Weighted LCA, Min Score: 50, Max Expected: 1e-5, Min Percent Identity: 70, Top Percent: 10.0, Min Support Percent: 0.05, Min Support: 1, Percent to cover: 50). Taxonomic lineages for OTUs were found and manipulated using custom scripts<sup>1</sup> based on the NCBI taxonomy database. OTUs with taxonomic assignments were associated with the samples and summarized to generate presence-absence profiles. Presence-absence profiles of fungal taxa were generated for each sample using a cumulative approach, whereby the presence of a low-level taxon also automatically assigned presence to its higher-level taxa. Patterns of variance were investigated using principal components analysis (PCA) using the scikit-learn python package (Pedregosa et al., 2011) as well as by performing canonical correspondence analysis (CCA) using the R package “vegan” (Oksanen et al., 2019). Species richness was also measured using the Shannon–Weaver and Simpson diversity indices in the R package “vegan.” To infer the specific identities of OTUs, OTU sequences were searched with BLAST against the custom database developed for this study (see above). BLAST matches were filtered to contain

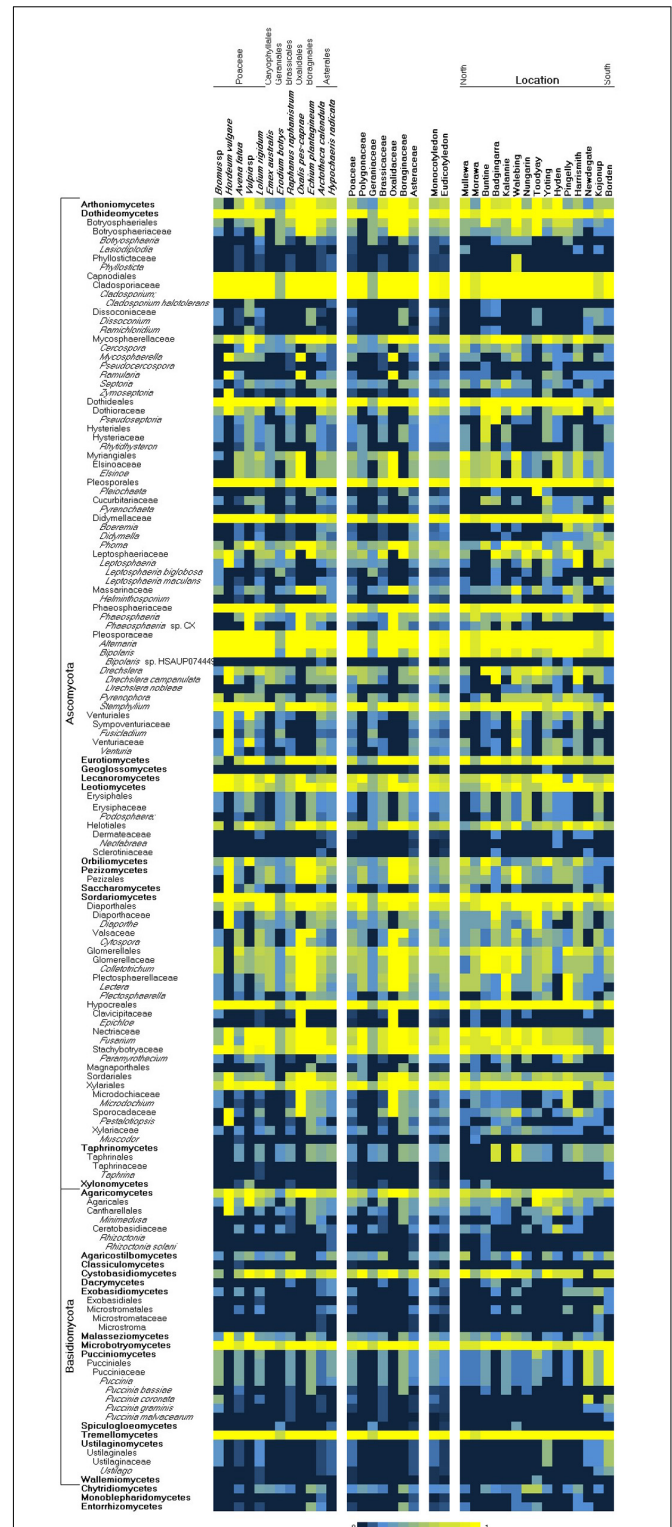
<sup>1</sup><https://github.com/darcyabjones/acc-to-tax>

alignments with greater than 95% identical matches, and covering at least 90% of the OTU sequence.

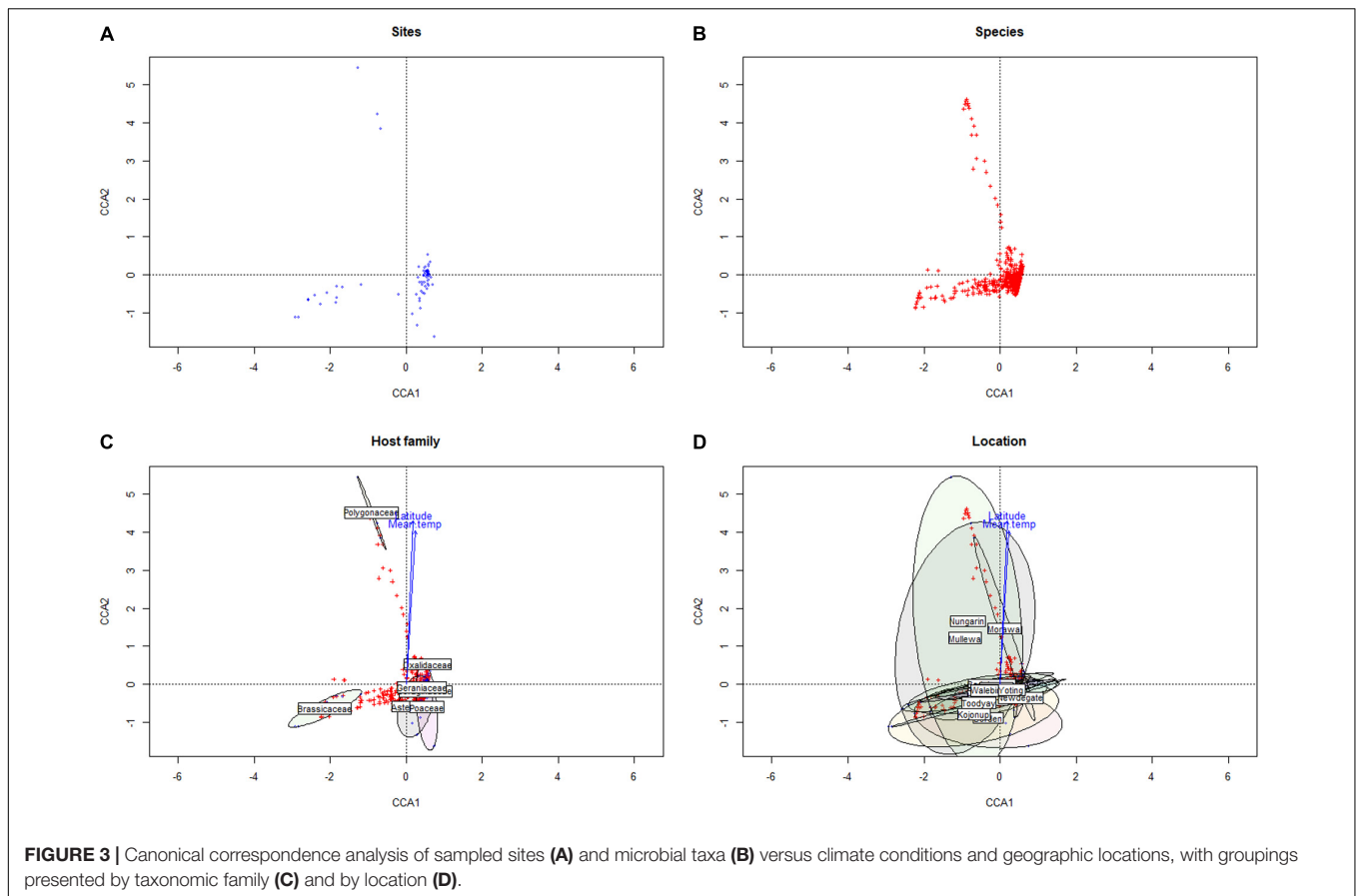
## RESULTS AND DISCUSSION

### Crop-Zone Weeds Are Host to a Wide Range of Plant-Associated Fungi

Overall, 2391 and 1819 OTUs were detected from the sequenced ITS and D2 loci, respectively (Supplementary Table 1). As an indicator of the specificity of identification, 1231 ITS and 1086 D2 OTUs were directly assigned to NCBI taxonomy identifiers at genus level or lower via BLASTN to Genbank (nt and GSS), while 7 and 15 could not be assigned to any taxonomic identifiers (Supplementary Table 2). Most OTUs were assigned to taxa at or below the fungal classes Dothideomycetes, Tremellomycetes, and Sordariomycetes (Supplementary Table 3). Species richness was highest in weeds within the Brassicaceae and Polygonaceae, followed by other host families that were not dissimilar. Among the fungal species taxa detected in weed samples, the heteroecious rusts of the *Puccinia* spp. (Zhao et al., 2016) (*Puccinia graminis*, *Puccinia bassiae*, *Puccinia coronata*, and *Puccinia malvacearum*) were prominently represented across multiple West Australian weed hosts (Supplementary Table 3). These reports of common Australian weeds as alternate hosts for the *Puccinia* rusts are analogous to a growing collection of reports across Europe, the United States, and Asia (Roelfs, 1982; Jin et al., 2010; Wang and Chen, 2013; Zhao et al., 2016). In this study, *P. graminis* was observed on *Bromus*; *P. bassiae* on *Echium* and *Avena*; *P. coronata* on *Bromus* and *Lolium*; and *P. malvacearum* on *Bromus*, *Lolium*, *Avena*, *Echium*, and *Raphanus*. Additional species of agricultural relevance that were identified tended to infect multiple hosts, including the following: *Leptosphaeria maculans*, *Leptosphaeria biglobosa*, *Drechslera campanulata* (syn. *Pyrenophora semeniperda*; ring spot of oat), and *Drechslera nobleae* (syn. *Pyrenophora phaeocomes*, pathogen of rye and pastures). Species endemic to Australia, including several species associated with native eucalypts (*Cladriella eucalypti*, *Fusculina eucalypti*, *Plectosphaera eucalypti*, *Phacidiella eucalypti*, *Phaeococcomyces eucalypti*, and *Saccharata eucalyptorum*) (Supplementary Table 2), were detected, which were consistent with the local environment but unrelated to crop production. Several genera that contain well-known crop pathogen species were also detected, however, their species-level identities were inconclusive. These included *Cladosporium*, *Zymoseptoria*, *Phaeosphaeria*, *Bipolaris*, *Pyrenophora*, *Venturia*, *Colletotrichum*, *Fusarium*, *Taphrina*, and *Ustilago*. We looked for (using covariance analysis of presence-absence variation) consistent association of two or more species across samples, as well as the inverse mutual exclusion of species (Supplementary Table 4), but unlike similar previous studies (Sapkota et al., 2017), we did not observe strong evidence of species complexes or potential bio-control agents across these samples. While culture-independent methods, which are based on the extraction and analysis of nucleic acids, can provide a comprehensive overview of microbial communities, they have drawbacks such as the detection of DNA from dead cells (Müller and Ruppel,



**FIGURE 2** Heatmap of the observed presence of common fungal pathogen genera/species by weed host species and collection location (left/North to right/South). A presence score from 0 to 1 (shown in blue–yellow) indicates the proportion of observed presence of a taxon across all samples of a common species or location, where 0 indicates total absence and 1 indicates ubiquity.



2014) and dependence on well-annotated databases (Blackwell, 2011; Aylward et al., 2017). It has been reported that up to 20% of fungal sequences in major databases such as GenBank may be misidentified (Bridge et al., 2003; Nilsson et al., 2008). However, the internal transcribed spacer regions (ITS1 and ITS2) have been demonstrated to be taxonomic markers for fungi due to their length and discriminative sequence variation (Lindahl and Kuske, 2013). Nevertheless, an issue we encountered with the taxonomic mapping of OTUs was that mappings to the species level were not consistently achieved. Species level (or lower taxa) would be required in order to reliably assign a crop disease to an OTU and infer its potential host range. Species-level mapping was dependent on multiple factors including representation bias across different taxa in sequence databases, variable map-ability of reads across different taxa, and potential for some species to be present but undetectable. In this study, OTU alignments more often supported the genus-level over species-level mappings; hence, this limited our ability to infer fully the disease risks posed by crop-zone weeds. We have presented a taxonomic summary of genera likely to contain plant-pathogenic species in Figure 2, although in most cases, the presence of pathogenic species was inconclusive, as endophytic species may have also been detected.

## Regional Biases

Within the Western Australian regions sampled, there appeared to be an overall trend for a relatively increased presence of

Ascomycota in northern sample sites and a corresponding relative increase in Basidiomycota (particularly biotrophic rusts of the *Puccinia* spp.) in southern sites (Figure 2). Multivariate clustering of samples tended to form groups based on geographic locations rather than by the weed host (Supplementary File 1), suggesting that location and/or climate zone was a major factor influencing the weed phyllosphere. This was also supported by CCA (Figure 3 and Supplementary File 2), which indicated that the phyllosphere composition was influenced by the continuous variables, latitude and mean temperature. Species distributions in the CCA corroborated the initial observation of a necrotrophic bias in hotter and dryer northern regions and a corresponding biotrophic bias in cooler and wetter southern regions (Figure 3 and Supplementary File 2). CCA also indicated clustering of hosts of the Polygonaceae family with a strong association with northern latitudes and increased temperatures, which was due to the three *Emex* (syn. doublegee) samples having been exclusively obtained from sites in the northern wheatbelt (Mullewa, Morawa, and Nungarin). However, a host cluster by Brassicaceae (composed of 13 *R. raphanistrum* samples) along CCA1 could not be explained by latitude or temperature vectors and suggests that unique host-specific pathogens may be present on this weed species. The other host families were represented at all sites and did not exhibit any clustering by CCA.

This approximate north-to-south division may reflect that the northern regions in this study are warmer and drier, which may

suppress saprophytic and/or pathogen-suppressive microbial activity that would break down the stubble containing pathogen inoculum—a common problem with cereal necrotrophs (typically of the Ascomycota). Conversely, the southernmost regions of WA are cooler and have increased rainfall, which likely increases the spread and survivability of heteromecic/obligate biotrophic pathogens (typically of the Basidiomycota). Although crop pathogen species were in the relative minority within the phyllosphere, this appears to be consistent to other studies linking climate zones to the distribution of fungal pathogens (Bebber et al., 2014; Fisher et al., 2018).

## CONCLUSION

This survey of the weed phyllospheres local to cereal crop-growing regions of Western Australia demonstrates the utility of this approach for the monitoring of plant pathogenic species and could be adapted for the purpose of monitoring for pathogen reservoirs and emerging crop disease risks. We report the presence of several important crop pathogen species within the phyllospheres of local weed hosts and observe that host range and climate zone are both important factors in determining their geographic distributions.

## DATA AVAILABILITY STATEMENT

Sequence data was deposited in NCBI BioProject: <https://www.ncbi.nlm.nih.gov/bioproject/672330>.

## AUTHOR CONTRIBUTIONS

PM, JH, and MG contributed to the conception and design of the study. PM conducted the field sampling. NW contributed to DNA extraction and sequencing. DJ and JH contributed to DNA sequence quality filtering and data analysis. PM, JH, DJ, and NW contributed to the manuscript. All authors contributed to the article and approved the submitted version.

## REFERENCES

- Amir, A., McDonald, D., Navas-Molina, J. A., Kopylova, E., Morton, J. T., Xu, Z. Z., et al. (2017). Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems* 2:e0191-16. doi: 10.1128/mSystems.00191-16
- Ankenbrand, M. J., Keller, A., Wolf, M., Schultz, J., and Förster, F. (2015). ITS2 database V: twice as much. *Mol. Biol. Evol.* 32, 3030–3032. doi: 10.1093/molbev/msv174
- Aylward, J., Steenkamp, E. T., Dreyer, L. L., Roets, F., Wingfield, B. D., and Wingfield, M. J. (2017). A plant pathology perspective of fungal genome sequencing. *IMA Fungus* 8, 1–15. doi: 10.5598/imafungus.2017.08.01.01
- Bálint, M., Tiffin, P., Hallström, B., O'hara, R. B., Olson, M. S., Fankhauser, J. D., et al. (2013). Host genotype shapes the foliar fungal microbiome of balsam poplar (*Populus balsamifera*). *PLoS One* 8:e53987. doi: 10.1371/journal.pone.0053987
- Bebber, D. P., Holmes, T., and Gurr, S. J. (2014). The global spread of crop pests and pathogens. *Glob. Ecol. Biogeogr.* 23, 1398–1407. doi: 10.1111/geb.12214

## FUNDING

This study was conducted within the Centre for Crop and Disease Management, a cooperative research center funded by the Grains and Research Development Corporation (GRDC) and Curtin University.

## ACKNOWLEDGMENTS

This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian and Western Australian Governments.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.581592/full#supplementary-material>

**Supplementary Figure 1** | Fungal species accumulation curves across 10 samples of *Lolium rigidum* and *Raphanus raphanistrum* leaves sampled from Yoting.

**Supplementary Table 1** | OTU frequencies for ITS and D2 regions versus host taxonomy and sample locations.

**Supplementary Table 2** | Assignment of taxa to OTUs via BLAST.

**Supplementary Table 3** | Combined ITS and D2 relative frequencies of microbial taxa versus host taxonomy and sample locations. Each microbial taxon detected by either an ITS or D2 OTU sequence is expressed in terms of a number ranging from 0 to 1, where 1 indicates universal occurrence across host taxa or sample locations, and 0 indicates no occurrences.

**Supplementary Table 4** | Co-variance tests for association of taxon pairs across samples. Taxon pairs with high or low co-variance may indicate biologically relevant associations or mutually exclusions across samples in this study.

**Supplementary File 1** | Multivariate analyses of microbial taxa testing for relationships versus host species or sample locations.

**Supplementary File 2** | Canonical correspondence analyses of microbial taxa testing for relationships between host taxa, pathogen taxa, sample location and climate.

- Blackwell, M. (2011). The fungi: 1, 2, 3. 5.1 million species? *Am. J. Bot.* 98, 426–438. doi: 10.3732/ajb.1000298
- Blixt, E., Olson, Å., Lindahl, B., Djurlle, A., and Yuen, J. (2010). Spatiotemporal variation in the fungal community associated with wheat leaves showing symptoms similar to stagonospora nodorum blotch. *Eur. J. Plant Pathol.* 126, 373–386. doi: 10.1007/s10658-009-9542-z
- Bridge, P. D., Roberts, P. J., Spooner, B. M., and Panchal, G. (2003). On the unreliability of published DNA sequences. *New Phytol.* 160, 43–48. doi: 10.1046/j.1469-8137.2003.00861.x
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., Mcgarrell, D. M., Sun, Y., et al. (2013). Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42, D633–D642. doi: 10.1093/nar/gkt1244
- Cordier, T., Robin, C., Capdevielle, X., Desprez-Loustau, M.-L., and Vacher, C. (2012). Spatial variability of phyllosphere fungal assemblages: genetic distance

- predominates over geographic distance in a European beech stand (*Fagus sylvatica*). *Fungal Ecol.* 5, 509–520. doi: 10.1016/j.funeco.2011.12.004
- Donovan, P. D., Gonzalez, G., Higgins, D. G., Butler, G., and Ito, K. (2018). Identification of fungi in shotgun metagenomics datasets. *PLoS One* 13:e0192898. doi: 10.1371/journal.pone.0192898
- Fisher, M. C., Hawkins, N. J., Sanglard, D., and Gurr, S. J. (2018). Worldwide emergence of resistance to antifungal drugs challenges human health and food security. *Science* 360, 739–742. doi: 10.1126/science.aap7999
- Hawksworth, D. L., and Lücking, R. (2017). Fungal diversity revisited: 2.2 to 3.8 million species. *Microbiol. Spectr.* 5, 1–2. doi: 10.1128/microbiolspec.FUNK-0052-2016
- Hunter, P. J., Pink, D. A., and Bending, G. D. (2015). Cultivar-level genotype differences influence diversity and composition of lettuce (*Lactuca* sp.) phyllosphere fungal communities. *Fungal Ecol.* 17, 183–186. doi: 10.1016/j.funeco.2015.05.007
- Huson, D. H., Beier, S., Flade, I., Górská, A., El-Hadidi, M., Mitra, S., et al. (2016). MEGAN community edition - interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* 12:e1004957. doi: 10.1371/journal.pcbi.1004957
- Ihrmark, K., Bödeker, I., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., et al. (2012). New primers to amplify the fungal ITS2 region-evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiol. Ecol.* 82, 666–677. doi: 10.1111/j.1574-6941.2012.01437.x
- Jin, Y., Szabo, L. J., and Carson, M. (2010). Century-old mystery of *Puccinia striiformis* life history solved with the identification of *Berberis* as an alternate host. *Phytopathology* 100, 432–435. doi: 10.1094/Phyto-100-5-0432
- Joshee, S., Paulus, B. C., Park, D., and Johnston, P. R. (2009). Diversity and distribution of fungal foliar endophytes in New Zealand podocarpaceae. *Mycol. Res.* 113, 1003–1015. doi: 10.1016/j.mycres.2009.06.004
- Karlsson, I., Friberg, H., Steinberg, C., and Persson, P. (2014). Fungicide effects on fungal community composition in the wheat phyllosphere. *PLoS One* 9:e111786. doi: 10.1371/journal.pone.0111786
- Kinkel, L. L. (1997). Microbial population dynamics on leaves. *Annu. Rev. Phytopathol.* 35, 327–347. doi: 10.1146/annurev.phyto.35.1.327
- Köljal, U., Nilsson, R. H., Abarenkov, K., Tedersoo, L., Taylor, A. F., Bahram, M., et al. (2013). Towards a unified paradigm for sequence-based identification of fungi. *Mol. Ecol.* 22, 5271–5277. doi: 10.1111/mec.12481
- Kurtzman, C. P., and Robnett, C. J. (1998). Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Van Leeuwenh.* 73, 331–371. doi: 10.1023/A:1001761008817
- Levetin, E., and Dorsey, K. (2006). Contribution of leaf surface fungi to the air spora. *Aerobiologia* 22, 3–12. doi: 10.1007/s10453-005-9012-9
- Lindahl, B. D., and Kuske, C. R. (2013). “Metagenomics for study of fungal ecology,” in *The Ecological Genomics of Fungi*, ed. F. Martin (Hoboken, NJ: Wiley-Blackwell), 279–303. doi: 10.1002/9781118735893
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200
- Müller, T., and Ruppel, S. (2014). Progress in cultivation-independent phyllosphere microbiology. *FEMS Microbiol. Ecol.* 87, 2–17. doi: 10.1111/1574-6941.12198
- Murray, D. C., Bunce, M., Cannell, B. L., Oliver, R., Houston, J., White, N. E., et al. (2011). DNA-based faecal dietary analysis: a comparison of qPCR and high throughput sequencing approaches. *PLoS One* 6:e25776. doi: 10.1371/journal.pone.0025776
- Narayanasamy, P. (2011). “Detection of fungal pathogens in the environment,” in *Microbial Plant Pathogens Detection and Disease Diagnosis*, ed. P. Narayanasamy (Cham: Springer), 201–244. doi: 10.1007/978-90-481-9735-4
- Nilsson, R. H., Kristiansson, E., Ryberg, M., Hallenberg, N., and Larsson, K.-H. (2008). Intraspecific ITS variability in the kingdom Fungi as expressed in the international sequence databases and its implications for molecular species identification. *Evol. Bioinform.* 4, 193–201. doi: 10.4137/EBO.S653
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). *vegan: Community Ecology Package. R Package Version 2.5-6*. Available online at: <https://CRAN.R-project.org/package=vegan> (accessed September 11, 2020).
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Peršoh, D. (2015). Plant-associated fungal communities in the light of meta’omics. *Fungal Divers.* 75, 1–25. doi: 10.1007/s13225-015-0334-9
- Putignani, L., Paglia, M. G., Bordi, E., Nebuloso, E., Pucillo, L. P., and Visca, P. (2008). Identification of clinically relevant yeast species by DNA sequence analysis of the D2 variable region of the 25-28S rRNA gene. *Mycoses* 51, 209–227. doi: 10.1111/j.1439-0507.2007.01472.x
- Roelfs, A. P. (1982). Effects of barberry eradication on stem rust in the United States. *Plant Dis.* 66, 177–181. doi: 10.1094/Pd-66-177
- Sapkota, R., Jørgensen, L. N., and Nicolaisen, M. (2017). Spatiotemporal variation and networks in the mycobiome of the wheat canopy. *Front. Plant Sci.* 8:1357. doi: 10.3389/fpls.2017.01357
- Sapkota, R., Knorr, K., Jørgensen, L. N., O’hanlon, K. A., and Nicolaisen, M. (2015). Host genotype is an important determinant of the cereal phyllosphere mycobiome. *New Phytol.* 207, 1134–1144. doi: 10.1111/nph.13418
- Wang, M. N., and Chen, X. (2013). First report of Oregon grape (*Mahonia aquifolium*) as an alternate host for the wheat stripe rust pathogen (*Puccinia striiformis* f. sp. *tritici*) under artificial inoculation. *Plant Dis.* 97, 839–839. doi: 10.1094/Pdis-09-12-0864-Pdn
- Weyrich, L. S., Duchene, S., Soubrier, J., Arriola, L., Llamas, B., Breen, J., et al. (2017). Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature* 544, 357–361. doi: 10.1038/nature21674
- White, T. J., Bruns, T., Lee, S., and Taylor, J. (1990). “Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics,” in *PCR Protocols: a Guide to Methods and Applications*, eds M. Innis, D. Gelfand, J. Sninsky, and T. J. White (Michigan: Academic Press), 315–322. doi: 10.1016/b978-0-12-372180-8.50042-1
- Zhao, J., Wang, M., Chen, X., and Kang, Z. (2016). Role of alternate hosts in epidemiology and pathogen variation of cereal rusts. *Annu. Rev. Phytopathol.* 54, 207–228. doi: 10.1146/annurev-phyto-080615-095851
- Zimmerman, N. B., and Vitousek, P. M. (2012). Fungal endophyte communities reflect environmental structuring across a Hawaiian landscape. *Proc. Natl. Acad. Sci. U.S.A.* 109, 13022–13027. doi: 10.1073/pnas.1209872109

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Michael, Jones, White, Hane, Bunce and Gibberd. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.