

A plant pathology perspective of fungal genome sequencing

Janneke Aylward¹, Emma T. Steenkamp², Léanne L. Dreyer¹, Francois Roets³, Brenda D. Wingfield⁴, and Michael J. Wingfield²

¹Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa; corresponding author e-mail: janneke@sun.ac.za

²Department of Microbiology and Plant Pathology, University of Pretoria, Pretoria 0002, South Africa

³Department of Conservation Ecology and Entomology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

⁴Department of Genetics, University of Pretoria, Pretoria 0002, South Africa

Abstract: The majority of plant pathogens are fungi and many of these adversely affect food security. This mini-review aims to provide an analysis of the plant pathogenic fungi for which genome sequences are publically available, to assess their general genome characteristics, and to consider how genomics has impacted plant pathology. A list of sequenced fungal species was assembled, the taxonomy of all species verified, and the potential reason for sequencing each of the species considered. The genomes of 1090 fungal species are currently (October 2016) in the public domain and this number is rapidly rising. Pathogenic species comprised the largest category (35.5 %) and, amongst these, plant pathogens are predominant. Of the 191 plant pathogenic fungal species with available genomes, 61.3 % cause diseases on food crops, more than half of which are staple crops. The genomes of plant pathogens are slightly larger than those of other fungal species sequenced to date and they contain fewer coding sequences in relation to their genome size. Both of these factors can be attributed to the expansion of repeat elements. Sequenced genomes of plant pathogens provide blueprints from which potential virulence factors were identified and from which genes associated with different pathogenic strategies could be predicted. Genome sequences have also made it possible to evaluate adaptability of pathogen genomes and genomic regions that experience selection pressures. Some genomic patterns, however, remain poorly understood and plant pathogen genomes alone are not sufficient to unravel complex pathogen-host interactions. Genomes, therefore, cannot replace experimental studies that can be complex and tedious. Ultimately, the most promising application lies in using fungal plant pathogen genomics to inform disease management and risk assessment strategies. This will ultimately minimize the risks of future disease outbreaks and assist in preparation for emerging pathogen outbreaks.

Key words:

genome size
pathogen evolution
pathogen lifestyle
plant pathology

Article info: Submitted: 26 October 2016; Accepted: 19 January 2017; Published: 9 February 2017.

INTRODUCTION

Sequencing of fungal genomes is being driven by various groups of scientists having different interests and needs from genomic data. Mycologists desire genome data to understand how fungi live and evolve, while industries require information on how to improve metabolic pathways or how to find new sources of natural products. The medical and plant pathology sectors need this information to understand diseases, improve diagnoses, understand how they function, and ultimately prevent or at least manage disease outbreaks (Kelman 1985). By 2007, the genomes of 42 eukaryotes were available (Cornell *et al.* 2007) and by 2008 the number of fungal genomes exceeded 90 (Park *et al.* 2008). Today, more than 3000 fungi are in completed or ongoing genome projects and the genomes of more than 900 fungal species have been released. The substantial and growing investment in determining genome sequences reflects the positive impact that this field is having on research. Our question here is what the impact has been for plant pathogenic fungi.

This mini-review aims to summarise the number of available fungal plant pathogen genomes, determine their general characteristics, and consider the impact that the availability of these genomes is having on the study of plant pathology. In order to determine which fungal plant pathogens have been sequenced, we surveyed fungal species (including *Microsporidia*, but excluding *Oomycota*) listed in 11 online genome repositories (Table 1), including MycoCosm (Grigoriev *et al.* 2012, 2013), NCBI Genome (www.ncbi.nlm.nih.gov/genome), the Broad Institute (www.broadinstitute.org), and the universal cataloguing database, Genomes OnLine Database (GOLD; Reddy *et al.* 2014). Fungal species that were found in more than one database were clustered and the current classification of all species was verified up to ordinal level using MycoBank (Robert *et al.* 2013) and Index Fungorum (www.IndexFungorum.org). The most recent scientific literature was consulted where the two online reference databases were not in agreement. Synonymous names associated with each species were noted by consulting MycoBank. We used this non-redundant list to accurately determine the number of fungal species with

© 2017 International Mycological Association

You are free to share - to copy, distribute and transmit the work, under the following conditions:

Attribution: You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
Non-commercial: You may not use this work for commercial purposes.
No derivative works: You may not alter, transform, or build upon this work.

For any reuse or distribution, you must make clear to others the license terms of this work, which can be found at <http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode>. Any of the above conditions can be waived if you get permission from the copyright holder. Nothing in this license impairs or restricts the author's moral rights.

Table 1. Fungal genome resources used to populate the genome species list.

Name	Abbreviation	Number of fungal genome projects	URL
<i>Aspergillus</i> Genome Database	AspGD	4	www.aspgd.org
<i>Candida</i> Genome Database	CGD	4	www.candidagenome.org
EnsemblFungi		53	fungi.ensembl.org
Fungal Genome Resource		4	gene.genetics.uga.edu
Genomes OnLine Database	GOLD	3362	genomesonline.org
JGI Genome Portal: MycoCosm		>1200	genome.jgi.doe.gov/fungi
NCBI Genome		>1000	www.ncbi.nlm.nih.gov/genome
PomBase		1	www.pombase.org
<i>Saccharomyces</i> Genome Database	SGD	>50	www.yeastgenome.org
The Broad Institute		>100	www.broadinstitute.org/science/projects/fungal-genome-initiative
The Institute of Bioinformatics and Systems Biology	IBIS	20	www.helmholtz-muenchen.de/en/ibis/institute/groups/fungal-microbial-genomics/resources/index.html
University of Kentucky		29	www.endophyte.uky.edu/

available genome sequences and, specifically, the extent to which fungal plant pathogens have been sequenced.

BEYOND THE 1000 MARK

The lower cost of genome sequencing, due to high-throughput technologies, has encouraged large scale genome initiatives. These include the 5000 Insect Genome Project (i5k; Robinson *et al.* 2011), Genome 10K (Genome 10K, Community of Scientists 2009) and 1000 Plants (www.onekp.com). These projects aim to sample species diversity by sequencing, respectively, whole genomes of insects and vertebrates and the transcriptomes of plant species. Similarly, fungal genome sequencing programmes such as the Fungal Genome Initiative (Fungal Research Community 2002, The Fungal Genome Initiative Steering Committee 2003), the Fungal Genomics Program (Grigoriev *et al.* 2011, Martin *et al.* 2011), and its extension, the 1000 Fungal Genomes (1KFG) Project (Spatafora 2011), have contributed significantly to the number of fungal genomes currently available and continue to do so.

The prevalence of fungi as study organisms is evident when considering the on-going and completed genome sequencing projects. A catalogue of genome projects, GOLD (Reddy *et al.* 2014), began in 1997 with six genome entries (Bernal *et al.* 2001) and in October 2016 included 7422 eukaryote whole genome sequencing projects, of which 3515 (47.4 %) are fungal. Although the respective genome databases (Table 1) list numerous completed and on-going fungal projects, many entries do not represent different species. Of the 1459 completed fungal genome projects in GOLD, slightly more than half (*ca.* 775) are different species, whilst the remainder comprise additional strains of already-sequenced species. To illustrate the extent and prevalence of sequenced genomes in the fungal kingdom, we mapped species with publically available genomes onto ordinal consensus trees (Fig. 1A–C).

The most recent of the fungal genome sequencing initiatives, the five-year international collaborative 1KFG Project, aims to sequence and annotate two species from each of the more than 500 known fungal families (Spatafora 2011). In three years, there has been a shift from obtaining representative genomes for all the fungal phyla (Buckley 2008) to targeting genome sequencing at the family level. By October 2016, the genomes of 1090 different fungal species were publically available (Supplementary File 1). Of this number, the 1KFG Project has released approximately 60 % of the fungal species genomes available.

Although the target of 1000 sequenced fungal species has been reached, the goal of two genomes from each family is a bigger task than the number 1000 (Fig. 1). The goal of having the genome sequences for two representatives have only been achieved in 85 families in the *Ascomycota*, 66 in the *Basidiomycota*, and 11 in the remainder of the fungi. Not surprisingly, some economically and medically important families (e.g. *Aspergillaceae*, *Clavicipitaceae*, *Mucoraceae*, *Mycosphaerellaceae*, *Saccharomycetaceae*, *Tremella-ceae*, *Ustilaginaceae*) have many more than two representatives. Additionally, taxonomic revision and species descriptions continue to generate new fungal families and orders. In the almost ten years since the publication of the Hibbett *et al.* (2007) consensus tree, more than 50 fungal orders have been described, somewhat increasing the workload of the 1KFG Project. Additionally, less than 10 % of the conservative estimate of 1.5 million total fungal species are known (Hawksworth 2012) and new species descriptions continuously emerge. Therefore, the combined goals of sampling fungal biodiversity and sequencing the genomes of representative species are a continuous process.

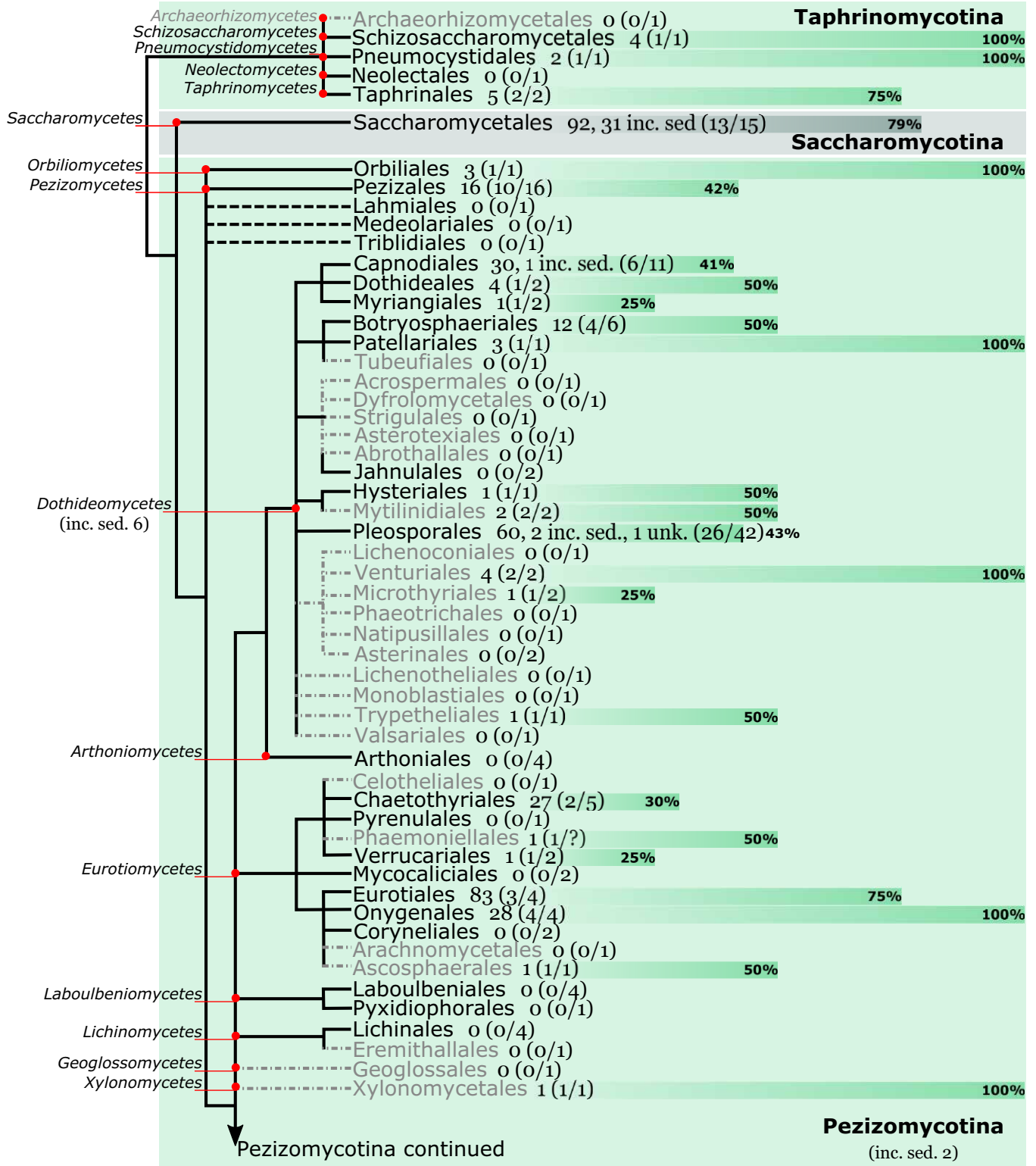


Fig. 1. Ordinal consensus trees depicting the taxonomic (subphylum, class and order) distribution of publicly available genomes for the Ascomycota (A), Basidiomycota (B) and early-diverging fungi (C). The number of sequenced genomes from each order is indicated after the order name. Where sequenced species are not classified into a family or have not been described, these are indicated as *incertae sedis* (inc. sed.) or unknown (unk.), respectively. The number of families with sequenced representatives out of the total number of described families is indicated in brackets. For each order, horizontal bars show the current progress of sequencing two genomes per family, indicated according to the scale bar below the figure. *Dikarya* consensus trees are according to Hibbett *et al.* (2007), while the classification of Spatafora *et al.* (2016) was included in the tree of early-diverging fungi. Orders described after Hibbett *et al.* (2007) have been added in grey (see Supplementary File 3 for references). The figures do not include unclassified fungi that have not been sequenced.

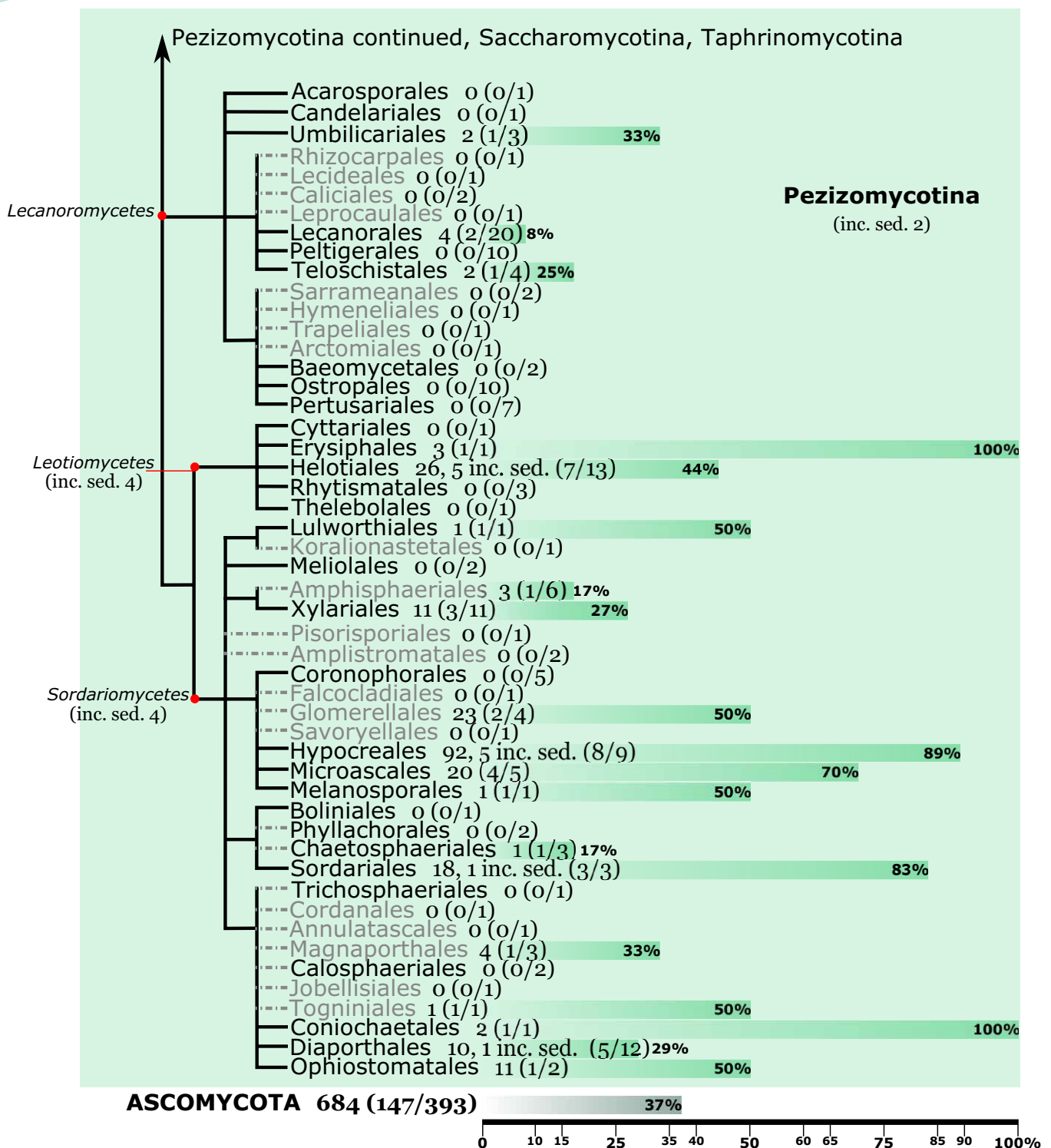


Fig. 1A. (Continued).

ARE PATHOGENS PREFERENTIALLY SEQUENCED?

More than 90 % of known fungal species reside in the subkingdom *Dikarya* (Kirk et al. 2008) comprised of the two largest phyla, *Ascomycota* and *Basidiomycota*. The large number of ascomycete and basidiomycete species for which genome sequences have been determined (Table 2) is, therefore, not an over-emphasis of these common phyla,

but rather reflects the size and diversity of the *Dikarya* (Fig. 2). In the majority of cases, the proportion of sequenced species in the phyla of early-diverging fungi is congruent with the known species, suggesting that genome projects have not neglected them (Fig. 2). *Mucoromycota* has a larger proportion of sequenced species than known species due to the sequencing of several *Mucoraceae* species that cause human mucoromycosis. One phylum (*Olpidiomycotina*) and one subphylum (*Zoopagomycotina*) of early-diverging fungi,

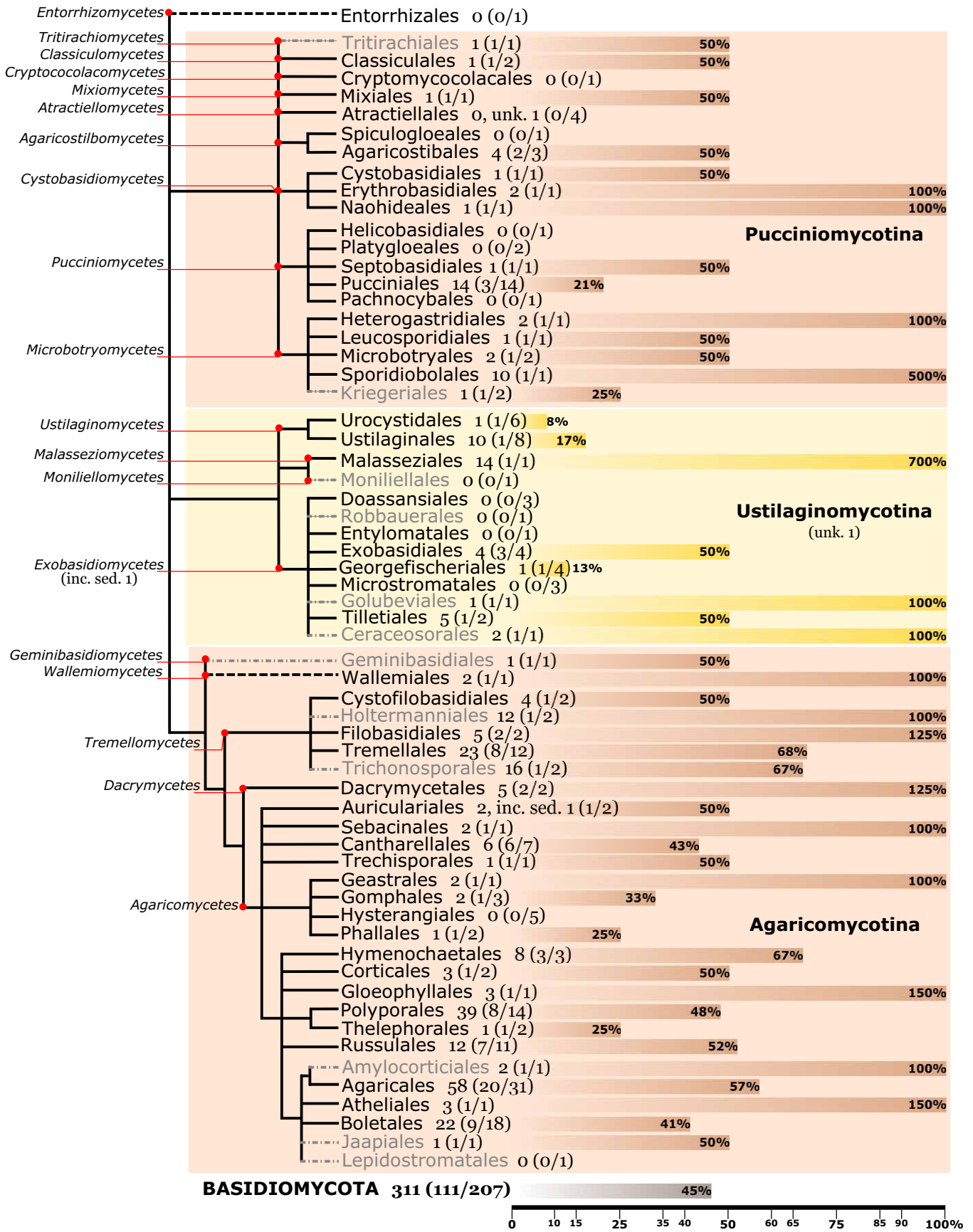


Fig. 1B. (Continued).

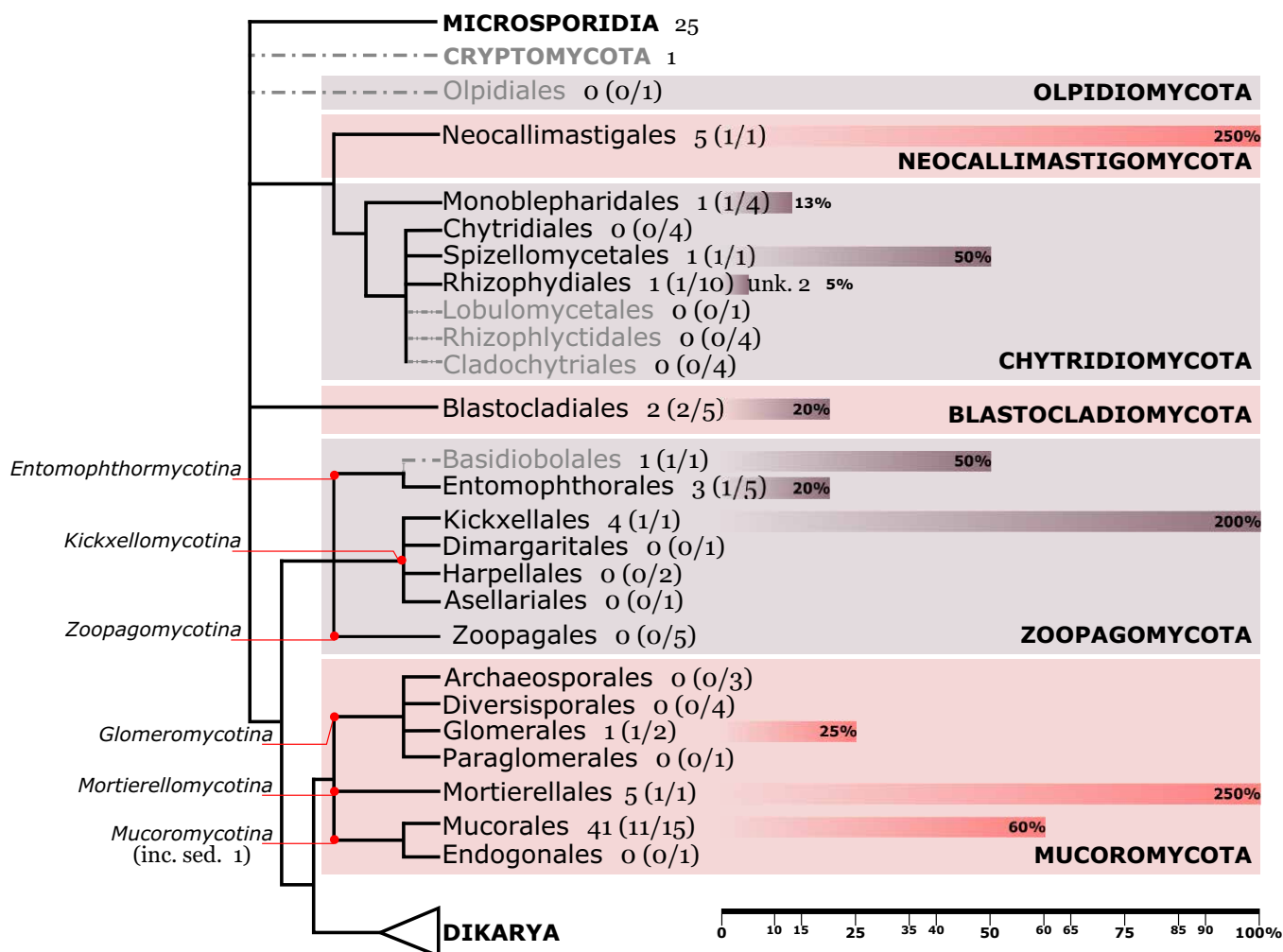


Fig. 1C. (Continued).

however, do not have any sequenced representatives and do not have “targeted” or “in progress” projects listed on GOLD. *Olpidiomycota* was only described recently (Doweld 2013) and its members appear to be poorly known. The lack of *Zoopagomycotina* sequences can likely be ascribed to very few available pure cultures of this predominantly parasitic group of fungi (Spatafora et al. 2016).

Within the subphyla of *Ascomycota* and *Basidio-mycota*, the proportion of sequenced species also largely corresponds to the number of known species (Fig. 3), with the major exception being *Saccharomycotina* (budding yeasts; Fig. 3A). The emphasis placed on this subphylum is even more pronounced when considering not only the number of species, but also the number of strains that have been sequenced. Although *Pezizomycotina* (filamentous fungi) is by far the most species-rich subphylum on the genome list (547 spp.), most sequenced strains are in *Saccharomycotina* (416), maintaining its previous status as the most sequenced subphylum in the kingdom (Cuomo & Birren 2010). Other than members of *Saccharomycotina*, seven highly sequenced (≥ 10 strains) species are listed in GOLD (Table 3). One is of industrial importance (*Rhodotorula toruloides*), whereas the remainder influence food security (*Aspergillus flavus*, *Fusarium oxysporum*, and *Magnaporthe oryzae*) or human health (*Cryptococcus gattii*, *Coccidioides posadasii*, and *Trichophyton rubrum*).

A 2008 report by the American Academy of Microbiology (Buckley 2008) stated that fungal genome sequencing is “heavily” skewed in the favour of human pathogens. At that time whole-genome sequencing of eukaryotes, especially fungi, was in its infancy and the statement was based on “100-150 fungal representatives”. The initial high cost of genome sequencing would have favoured fungi of medical importance, but as the number of sequenced fungi grew and the cost decreased, this pattern was bound to change. We assessed whether pathogens are highly sequenced by consulting recent scientific literature (where available) on each fungal species on the genome list and categorising them according to their significance and reason for being sequenced. The largest (41.4 %) category consisted of pathogenic fungi and fungi of medical importance (Fig. 4), of which plant pathogens were the most prevalent group (49.4 %). Currently, 191 plant pathogenic species have publically available genomes (Supplementary File 2) and all belong to *Dikarya*. Of these, 117 are pathogens of at least one food crop and 43 affect gymnosperms, the majority of which are commercially important (Table 4). The 117 food crop species include pathogens of cereals, fruit, vegetables, and legumes. At least 60 of these species are responsible for diseases on 10 of the 15 global staple food crops (FAO 1995).

Table 2. Number of fungal species from each phylum and subphylum with at least one available genome.

Phylum and subphylum	Sequenced species	Known species ^a
ASCOMYCOTA	684	> 64 000
<i>Pezizomycotina</i>	547	
<i>Saccharomycotina</i>	123	
<i>Taphrinomycotina</i>	11	
<i>Incertae sedis</i>	3	
BASIDIOMYCOTA	311	> 31 000
<i>Agaricomycotina</i>	227	
<i>Pucciniomycotina</i>	43	
<i>Ustilaginomycotina</i>	41	
BLASTOCLADIOMYCOTA	2	> 175
CHYTRIDIOMYCOTA	5	> 700
CRYPTOMYCOTA	1	?
MICROSPORIDIA	25	> 1 300
MUCOROMYCOTA	47	
<i>Glomeromycotina</i>	1	> 165
<i>Mortierellomycotina</i>	5	
<i>Mucoromycotina</i>	41	> 325
NEOCALLIMASTIGOMYCOTA	5	> 20
ZOOPAGOMYCOTA	8	
<i>Entomophthoromycotina</i>	4	> 275
<i>Kickxellomycotina</i>	4	> 260
UNKNOWN	2	
Total sequenced	1090	

^a According to Kirk *et al.* (2008).

Table 3. Fungal species on the Genomes OnLine Database (Bernal *et al.* 2001) with 10 or more completed whole-genome sequencing projects.

Species	Strains	Phylum	Subphylum
<i>Saccharomyces cerevisiae</i>	166	Ascomycota	Saccharomycotina
<i>Magnaporthe oryzae</i>	48	Ascomycota	Pezizomycotina
<i>Candida albicans</i>	35	Ascomycota	Saccharomycotina
<i>Komagataella pastoris</i>	32	Ascomycota	Saccharomycotina
<i>Saccharomyces kudriavzevii</i>	20	Ascomycota	Saccharomycotina
<i>Cryptococcus gattii</i>	18	Basidiomycota	Agaricomycotina
<i>Fusarium oxysporum</i>	17	Ascomycota	Pezizomycotina
<i>Trichophyton rubrum</i>	12	Ascomycota	Pezizomycotina
<i>Aspergillus flavus</i>	10	Ascomycota	Pezizomycotina
<i>Coccidioides posadasii</i>	10	Ascomycota	Pezizomycotina
<i>Rhodotorula toruloides</i>	10	Basidiomycota	Pucciniomycotina
<i>Saccharomyces pastorianus</i>	10	Ascomycota	Saccharomycotina

Clearly, genome sequencing projects have placed an emphasis on plant pathogenic species, specifically those affecting food security or commercial forestry. In general, fungal pathogens are highly represented in the genome list. Furthermore, as the number of available genomes has increased, plant pathogens have replaced human pathogens as the predominantly sequenced category of species. Since most plant pathogens are fungi (Carris *et al.* 2012), the emphasis placed on fungal genome sequencing may (at least partly) be attributed to food security. For example, *M. oryzae*, a plant pathogen for which numerous strains have been sequenced (Table 3), is predicted to increase its distribution range and impact due to increased temperature and carbon dioxide levels (Gautam *et al.* 2013). The sheer number of plant species and their associated disease-causing fungi makes this change in the focus of genome sequencing understandable. Sequencing a large number of plant pathogens that affect a range of plant species is, after all, less of a bias than sequencing many pathogenic species associated with a single species (humans).

GENOME SIZE AND GENE NUMBERS IN PLANT PATHOGENIC FUNGAL GENOMES

As far as we are aware, this review includes the first comprehensive list of plant pathogenic fungal genomes that have been sequenced to date. We, therefore, briefly present an overview of the genome characteristics of these species in comparison to other sequenced fungal species. We specifically looked at genome size and the numbers of genes encoded, because previous studies have revealed a link between plant pathogenicity and genome size and gene content (Duplessis *et al.* 2011, Ohm *et al.* 2012, Spanu *et al.* 2010).

The 1090 fungal species with publically available genome sequences have haploid genome sizes ranging between two and 336 million base pairs (Mbp; Fig. 5A). The majority of these genomes fall within the 30–40 Mbp range (average = 37.2, median = 33.6), consistent with the size distribution of the 1940 entries in the Fungal Genome Size Database (Kullman *et al.* 2005). The genome sizes of sequenced plant pathogens are only slightly, but significantly, larger compared to this “norm”. This difference was most apparent in the pathogenic ascomycetes for which Mann-Whitney U tests indicated the highest level of significance (Median = 38.0; U = 48131, P < 0.01). The average genome size of plant pathogenic basidiomycetes (57.3 Mb) was much larger than that of the plant pathogenic ascomycetes (39.4 Mb) and the remainder of the fungal genomes (34.8 Mb), owing to several pathogenic pucciniomycete (rust) species with genomes larger than 100 Mb.

Somewhat larger genome sizes in plant pathogens are congruent with the hypothesis that they often contain more repeated elements than other species (discussed below) (Castanera *et al.* 2016, Ma *et al.* 2010). Sequenced plant pathogens also have larger genomes than human, animal and opportunistic fungal pathogens (Fig. 5B). Although sequencing has thus far sampled the genome size distribution of the majority of the fungal kingdom, species with excessively large genome sizes have been omitted. This is not necessarily

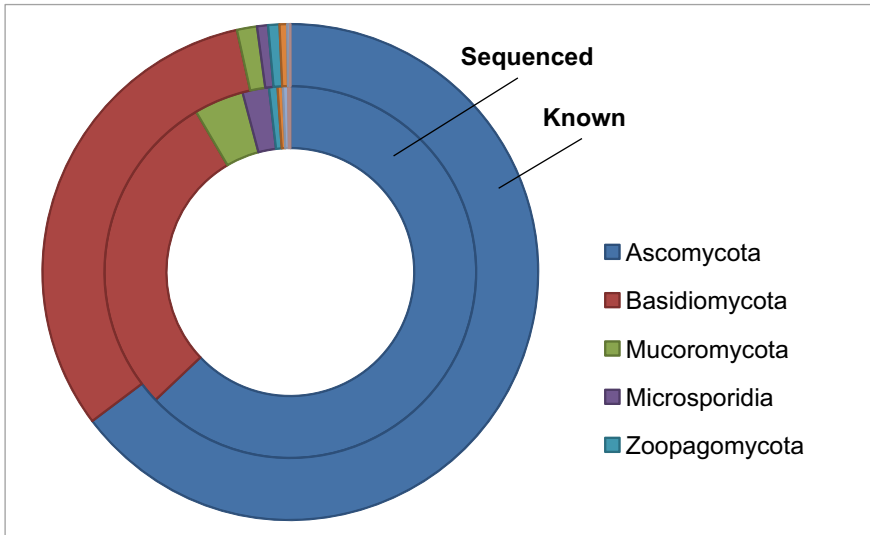


Fig. 2. Comparison between the proportion of known and sequenced fungal species in the major fungal taxonomic groups. The number of known species were obtained from Kirk *et al.* (2008).

only due to the higher cost of sequencing large genomes, but probably also the complexity of obtaining sufficient biomaterial from a single obligately parasitic individual cultured on a live host (Barnes & Szabo 2008). Since the majority of species in *Pucciniomycotina* reside in the order *Pucciniales* of obligate plant pathogens (Kirk *et al.* 2008), the latter may also explain why the proportion of sequenced species in this group is slightly less than the known species (Fig. 3B). Therefore, genome-sequencing efforts so far, most likely underestimate the maximum size of plant pathogen genomes.

Considering the number of predicted open reading frames (ORFs), 714 of the available genomes have publically

accessible gene annotations. The sequenced fungi have, on average, $11\,256 \pm 3\,873$ total predicted ORFs at a density of 351.8 ± 104.0 ORFs per Mbp (Fig. 5C). In comparison to the other genomes, plant pathogenic fungi do not differ in the number of predicted ORFs, but they do have significantly fewer ORFs when accounting for genome size ($U = 35647$; $P < 0.01$). This trend was also observed in the animal pathogenic fungi (including entomopathogens; $U = 8847$; $P < 0.01$). Previous whole-genome studies suggest that the number of coding genes does not necessarily increase with genome size, since transposable elements and repetitive sequences proliferate in large genomes (Kidwell 2002).

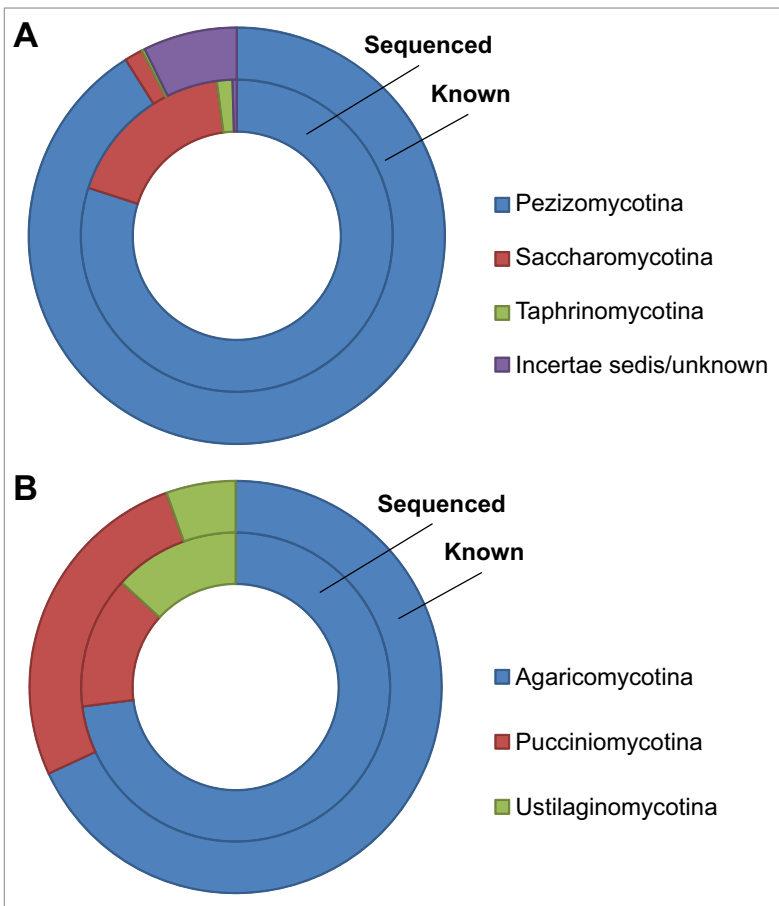


Fig. 3. Comparison between the proportion of known and sequenced fungal species in the subphyla of *Ascomycota* (A) and *Basidiomycota* (B). The number of known species were obtained from Kirk *et al.* (2008).

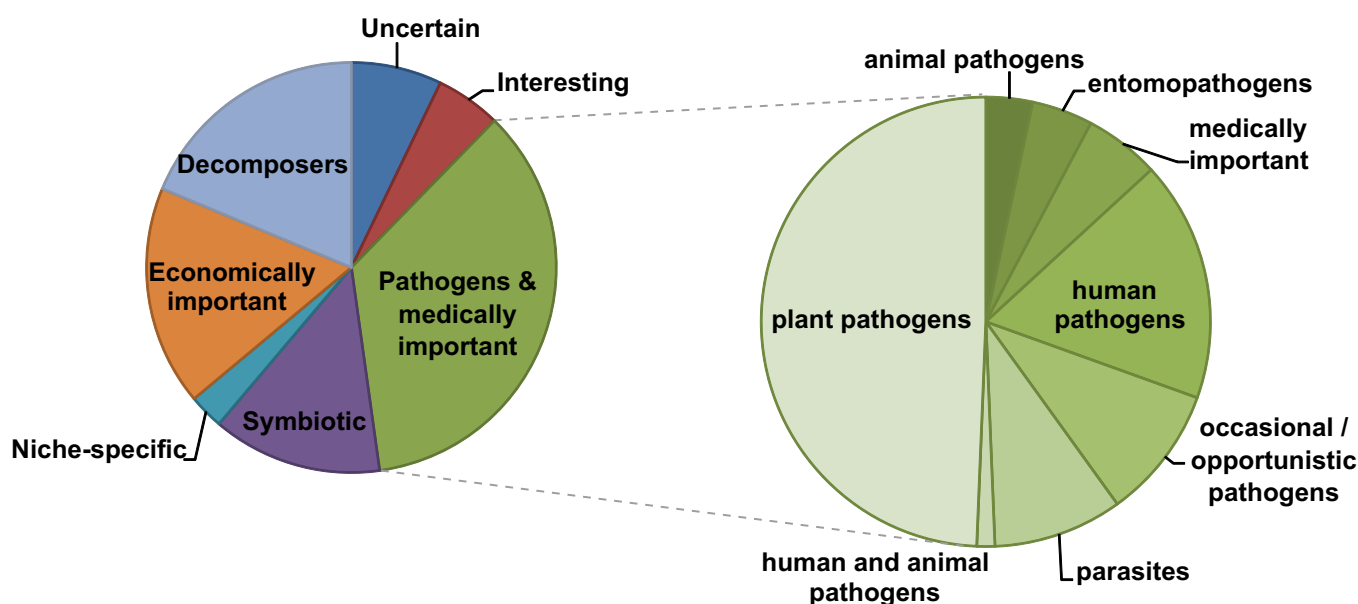


Fig. 4. Categories of significance identified in the 1090 sequenced fungal species. Pathogens comprise the largest category within which plant pathogens are predominant. “Medically important” species represent fungi that are not directly pathogenic, but cause food or environmental contamination. “Interesting” species are studied for their development or metabolism. “Niche-specific” refers to species occupying abiotic niches, whereas “symbiotic” species are associated with other organisms. “Economically important” species have a use in the economy, (e.g. culinary, biocontrol or pharmaceutical industries). Most of the parasites belong to *Microsporidia*.

Table 4. Categories of plants affected by the 191 sequenced fungal plant pathogens.

Plant pathogen categories	Genomes available	%
Cash Crop Pathogens	8	4.2
Food Crop Pathogens	117	61.3
Grains	48	25.2
Fruit	37	19.4
Vegetables	10	5.2
Legumes	11	5.8
Multiple crop types	11	5.8
Gymnosperm Pathogens	43	22.5
Other ^a	23	12.0
TOTAL	171	100

^a Non-gymnosperms not cultivated for food.

The lower number of ORFs/Mb in the genome of plant pathogenic Ascomycetes is, therefore, consistent with their larger genome size possibly being due to repetitive elements. Additionally, some pathogens have lost genes redundant in their lifecycles (Spanu *et al.* 2010), which may also decrease their ORFs/Mb. This trend could, however, not be detected in the genomes of human and opportunistic pathogens (Fig. 5D).

IMPACT OF GENOMES ON PLANT PATHOLOGY

Ever since the advent of plant pathology, researchers have been interested in the biology of plant pathogens and how this can be translated into means for disease control. In this

regard, genome data are not used in isolation, but provide context for observational and experimental data, thereby accelerating the pace of traditional research methods. For emerging pathogens, a disease may be known, but the mechanisms relating to infection biology and virulence are not necessarily understood. In these cases, a genome can provide the first glimpse of the potential effectors and toxins that are present (e.g. Ellwood *et al.* 2010). In some plant pathogens, genomes have resulted in a shift of conventional paradigms. Here a classic example is the discovery of entire horizontally transferrable chromosomes related to pathogenicity in *Fusarium oxysporum* f. sp. *lycopersici* (Ma *et al.* 2010). The primary impacts of genome sequences on plant pathology have been a better understanding of the pathogenicity, life-style and genome evolution of pathogens. Furthermore, genomes are also resources from which genetic tools can be used to mine information.

Pathogenicity and life-style

Secreted and cell surface proteins mediate the interaction between pathogen and host and are often the first to be characterised from plant pathogen genomes. Genome sequences enable *in silico* predictions of secreted virulence proteins (effectors), bypassing traditional enzyme assays or chromatography/spectrometry techniques that were ineffective at detecting less abundant effectors. For the corn smut fungus, *Mycosarcosoma maydis* (syn. *Ustilago maydis*), previous experimental studies were not able to identify the virulence factors that were eventually highlighted by interrogating the genome sequence (Kamper *et al.* 2006). The proteins identified *in silico* could then be used to experimentally determine the function of specific effectors in the infection process (Liu *et al.* 2015). This genome-based effector identification

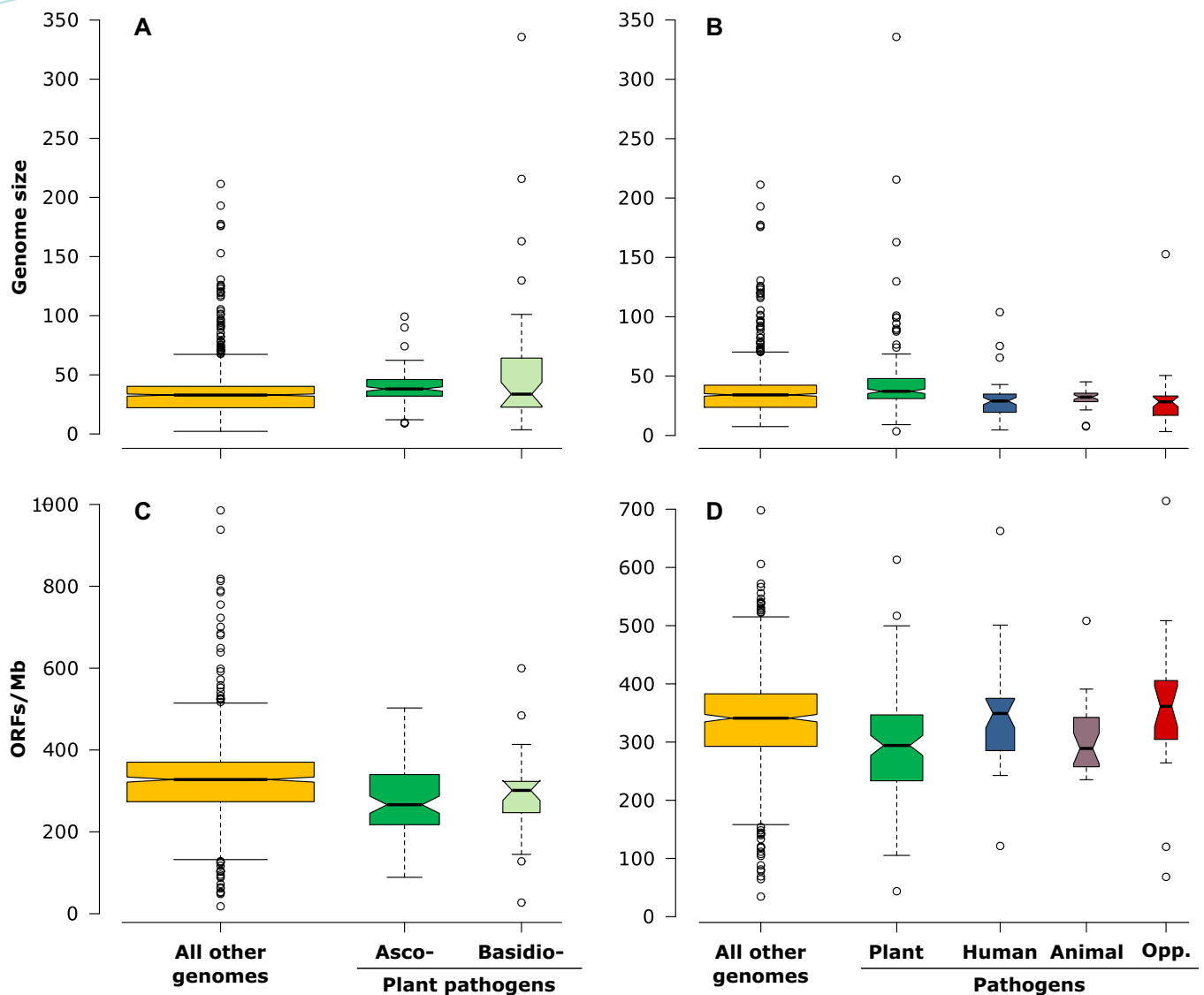


Fig. 5. Genome size (**A** and **B**) and number of open reading frames (ORFs) per million base pairs (Mbp; **C** and **D**) in the plant pathogen genomes compared to the remainder of the genome list and other pathogens. Boxplots were drawn in BoxplotR (Spitzer *et al.* 2014) using the Tukey whisker extent. Width of the boxes is proportional to the square root of the sample size; notches show the 95 % confidence interval of the median. Opp. = opportunistic pathogens. In B and D, animal pathogens include entomopathogenic fungi and *Microsporidia* are excluded from “other” genomes, since they have small genomes with many ORFs/Mb.

and subsequent phenotype determination has contributed significantly to the online Pathogen Host Interactions Database (Urban *et al.* 2015). Similarly, genomic prediction of fungal product biosynthesis genes have been well established (Keller *et al.* 2005) and gene deletion systems could subsequently be used to determine the phenotypes that they confer (Lee *et al.* 2005). Gene inventories of plant pathogens have also revealed proteins not previously known to be involved in pathogenicity, for example the high diversity of membrane transporters in the *F. oxysporum* and *Pyrenochaeta lycopersici* genomes strongly implicates them in the pathogenicity of these fungi (Aragona *et al.* 2014).

Intuitively, cell wall degrading enzymes can be expected to be important in plant pathogenesis and the presence of these enzymes in plant pathogens was well established before the genomic era (Jones *et al.* 1972, Schneider & Collmer 2010). Genome sequences have confirmed that most plant pathogens encode an array of cell wall degrading

enzymes, specifically pectinolytic enzymes in dicot pathogens (Klosterman *et al.* 2011, Olson *et al.* 2012). The diversity of cell wall degrading enzymes in a genome appears to increase with host range, as exemplified by the massive number of carbohydrate degrading enzymes in *Macrophomina phaseolina* that infects over 500 plant species (Islam *et al.* 2012). Exceptions to this perceived norm typically occur in specialised modes of pathogenesis. For example, cell wall degrading enzymes are absent from the genome of the anther smut fungus *Microbotryum lychnidis-dioicae* (Perlin *et al.* 2015). Rather than attacking plant cells, this pathogen has an array of enzymes to influence host development, enabling fungal spores to be substituted for pollen. In contrast, gene inventories suggest that necrotrophic pathogens induce apoptosis in host cells rather than breaking down their cell walls (McDonald *et al.* 2015). Metabolism-related enzymes in fungal genomes, therefore, have great potential to predict infection strategies and lifestyle.

Beyond the analysis of single genome sequences, comparing the genomes of ecologically different strains and species has substantial value. For example, analysis of the *Rhizoctonia solani* AG2-2IIIB genome would have revealed only an abundance of cell wall degrading enzymes. However, comparisons with less aggressive *R. solani* strains revealed that its virulence can be linked to a significant expansion of polysaccharide lyase enzymes (Wibberg *et al.* 2016). Similarly, comparisons of resistant and non-resistant *Penicillium digitatum* strains has enabled identification of mutations conferring tolerance to antifungal compounds (Marcet-Houben *et al.* 2012).

Comparisons between the genomes of 18 dothideo-mycete species has suggested that the number of effectors encoded by these fungi are linked to the pathogenic lifestyle (Ohm *et al.* 2012). The greatest number of effectors was identified from necrotrophic pathogens, whereas hemibiotrophs have apparently reduced their effector arsenal to evade plant defences before they switch to necrotrophy (Ohm *et al.* 2012). Furthermore, multiple genome comparisons have been used to highlight specific genes under diversifying selection, revealing that evolutionary pressure on plant pathogen effector proteins drive their adaptation (Schirawski *et al.* 2010, Stukenbrock *et al.* 2011). Comparison between plant and fungal genomes have also become essential tools to tease apart pathogen and plant RNA sequences when analysing *in planta* transcript data (McDonald *et al.* 2015). The value of multiple genome comparisons has prompted projects such as the Fungal Genome Initiative and the 1KFP to focus not on sequencing single species, but groups of species useful in a comparative context (Grigoriev *et al.* 2013, The Fungal Genome Initiative Steering Committee 2004).

The evolution of different fungal lifestyles is a fascinating topic considered by many comparative genomics studies. Some plant pathogenic fungi with different lifestyles have surprisingly similar gene contents (De Wit *et al.* 2012), yet unique genes mediate their host interactions. The large proportion of unique secreted effector proteins and host-specific hydrolytic enzymes in plant pathogenic fungi implies that host association drives their adaptation and, therefore, evolution (De Wit *et al.* 2012, Duplessis *et al.* 2011, O'Connell *et al.* 2012, Spanu *et al.* 2010). The effect of host association is further emphasised by the diversification of both effectors and hydrolytic enzymes in broad host range pathogens such as *Colletotrichum higginsianum* (O'Connell *et al.* 2012). In contrast, host association cannot explain the loss of primary metabolism genes that led to obligate biotrophy in powdery mildew fungi (Spanu *et al.* 2010). Similarly, selective pressures that mediate the evolution of a hemibiotrophic strategy, where the pathogen transitions between a biotrophic and necrotrophic lifestyle, are poorly understood.

Genome evolution of plant pathogens

The arms race between pathogen and host (Stahl & Bishop 2000) makes pathogen adaptability, or evolutionary potential, particularly interesting (McDonald & Linde 2002). Reproduction and gene diversity are two of the factors that influence evolutionary potential (McDonald & Linde 2002) and these can be estimated from genome sequences. For example, an analysis of the mating type genes that govern

sexual reproduction can provide insights into the mating strategy of a fungus. Heterothallic ascomycete fungi are identified by the occurrence of a single mating type in a genome (Kronstad & Staben 1997), whereas homothallic fungi contain both mating types, either in the same genome or in a dikaryotic cellular state (Wilson *et al.* 2015). Many fungi propagate only vegetatively or sexual reproduction is difficult to observe. In such cases, genomic analyses have been able to reveal the presence of mating type genes (e.g. Bihon *et al.* 2014, Marcet-Houben *et al.* 2012), suggesting that these species could have a cryptic sexual cycle (Bihon *et al.* 2014). The mating type sequence information can subsequently be used to determine the distribution of different mating types in a population (Aylward *et al.* 2016, Haasbroek *et al.* 2014). In contrast, genomes can also reveal the importance of mitotic recombination for generating new allelic combinations. For example, a whole genome survey concluded that mating type genes are completely absent from the tomato pathogen *P. lycopersici* and that it has an expansion of gene modules associated with heterokaryon incompatibility (Aragona *et al.* 2014).

Adaptability to a changing environment or a resistant host can also be mediated by genome plasticity. Transposons are repetitive elements in DNA known to contribute to genome plasticity and evolution (Wöstemeyer & Kreibich 2002). The expansion of repeats in many plant pathogen genomes points to their role in diversification and adaptation (Raffaele & Kamoun 2012, Spanu *et al.* 2010, Thon *et al.* 2006) and has been directly implicated in the pathogenicity of the wheat necrotroph *Pyrenophora tritici-repentis* (Manning *et al.* 2013). Surveys of transposons across plant pathogen genomes have revealed differences in their number and activity between the essential core and dispensable supernumerary chromosomes (Ohm *et al.* 2012, Vanheule *et al.* 2016). In *Fusarium poae*, repeat expansion in the core chromosomes is contained, while the non-essential supernumerary chromosomes have many active transposons that invade the core chromosomes (Vanheule *et al.* 2016). The supernumerary chromosomes also provide opportunity for duplication and diversification of core genes, thereby facilitating adaptation. An example of such diversification and adaptation in the post-harvest spoilage fungus *Penicillium digitatum* is the association of DNA transposons and ABC transporters in drug resistant strains (Sun *et al.* 2013b).

Horizontal gene transfer (HGT) may add novel ecological capabilities to the genomes of recipient species. Although not historically considered relevant to eukaryotic evolution, genome level investigations have revealed multiple HGT events in fungi, often from other kingdoms (e.g. Marcet-Houben & Gabaldón 2010, Sun *et al.* 2013a). Such phylogenomic studies have reported, amongst others, horizontal acquisition of genes that mediate pathogenicity (Friesen *et al.* 2006, Kroken *et al.* 2003, Slot & Rokas 2011, Thynne *et al.* 2015), tolerance to host defences (Marcet-Houben & Gabaldón 2010, Sun *et al.* 2013a), and nutrient uptake and metabolism (Soanes & Richards 2014, Sun *et al.* 2013a). Moreover, a comparative genomics study found evidence of HGT at chromosome level in *F. oxysporum* f. sp. *lycopersici*, as entire pathogenicity-related chromosomes could be transferred between strains (Ma *et al.* 2010).

Acquiring new ecological capabilities through HGT has previously played a causal role in the emergence of new pathogens and will likely do so again in future (Friesen *et al.* 2006, Soanes & Richards 2014, Thynne *et al.* 2015).

Resources for genetic tools

Other than facilitating whole-genome related studies, genome sequences have become ideal resources for mining genetic tools. Previously, species-specific population genetic tools such as microsatellites had to be developed painstakingly by cloning and genome walking (Barnes *et al.* 2001, Burgess *et al.* 2001). Now, any genome sequence enables rapid identification of such genetic markers (e.g. Haasbroek *et al.* 2014). This holds true for diagnostic markers: genome regions that unambiguously and rapidly identify a pathogen and/or differentiate between pathogens can be designed by inspecting whole genome sequences. Although development of such markers in fungi is lagging behind viral and bacterial pathogens, some examples have recently become available. A pathotype specific marker has been developed from the genome of *M. oryzae* f.sp. *tritricum* (Pieck *et al.* 2016) and comparative genomics has detected diagnostic regions in two *Colonectria* species (Malapi-Wight *et al.* 2016) and in *Pseudoperonospora cubensis* (Withers *et al.* 2016). Continued application of fungal genomes to generate identification tools is bound to increase the efficiency of quarantine procedures (McTaggart *et al.* 2016).

CHALLENGES

Although the availability of fungal genomes has dramatically increased our knowledge and understanding of infection processes and genome evolution, there remains much to learn. For example, the regulatory elements in most genomes remain poorly annotated and require complex experimental methodologies for accurate identification (e.g. Shen *et al.* 2012). In a recent review, Schatz (2015) commented that sequencing human genomes has been one of the greatest accomplishments of the past two decades but “one of the greatest pursuits for the next twenty years will be trying to understand what it all means”. The same can be said for fungal genomes. The information that can be gleaned from a genome sequence is bound to increase as our understanding of these sequences grows.

Genome sequences should not be seen to provide “silver bullets”, although they are often sold this way. They provide the blueprint of potential cellular activities, but are not sufficient to unravel the complexity of pathogen-host interactions. For example, in *Fusarium oxysporum*, the cell wall degrading enzymes secreted during infection of tomato displayed a clear succession (Jones *et al.* 1972), an ecologically relevant process that could not be deduced from a gene inventory. In combination with transcriptome data, however, genomic data has revealed how pathogens tolerate host defences (DiGiustini *et al.* 2011) and how hosts can resist pathogen infection (Zhu *et al.* 2012). Experimental work, both *in vitro* and *in plantae*, will remain essential components in studying fungal plant pathogens.

The end goal of studying any host-pathogen relationship is clearly to inform disease management and control. Thus far, identifying specific molecular targets has had little impact

on developing new antifungal inhibitors (Odds 2005) and integrative management strategies must, therefore, be a priority. Ultimately, the elucidated effector proteins, host targets, and the overall insights gained into the biology of pathogens must inform disease management strategies (Maloy 2005). It is also crucial that they inform risk assessment protocols governing biosecurity (McTaggart *et al.* 2016). It is, therefore, essential that the ecological significance of genome patterns is studied to ensure that this knowledge can be extrapolated to emerging pathogen threats.

As revealed by comparative genomics, deciphering plant pathogen evolution is in many cases dependent on being able to do comparisons with species having other lifestyles. A large scale example of this is the revised classification of species previously known as *Zygomycetes* (Spatafora *et al.* 2016); an endeavour possible because of the availability of multiple genome sequences for this group. In this regard, filling in the gaps in the list of sequenced species is crucial to our understanding of relationships and pathogenesis. The challenge is, therefore, to continue sequencing apparently uninteresting or unimportant taxonomic groups along with the economically important in order to ultimately gain a holistic view.

CONCLUSIONS

The activities of independent research groups and several fungal sequencing initiatives (Fungal Research Community 2002, Grigoriev *et al.* 2011, Martin *et al.* 2011, Spatafora 2011, The Fungal Genome Initiative Steering Committee 2003), have resulted in the number of publically available fungal genomes growing exponentially since 1996 when the first genome was sequenced (Goffeau *et al.* 1996). The taxonomic distribution of sequenced fungal genomes is currently roughly congruent with the number of species known from each phylum and subphylum. This is an important and impressive achievement in the goal of sampling biodiversity and representing the phylogenetic groups of the fungal kingdom (Fungal Research Community 2002, The Fungal Genome Initiative Steering Committee 2003). Many of the genomes have been sequenced to sample environmental and ecological diversity. However, investment continues to be primarily focused on projects that have direct human importance. The emphasis on genomes of plant pathogenic fungi has specifically increased subsequent to the Buckley (2008) overview of sequenced fungal species.

The genomes of more than 1 000 fungal species are already publically available and this number is growing steadily. Fungal genomics has enabled rapid characterization of plant pathogen genomes and revealed features that allow better understanding of the biology of these species. It has also made it possible to rapidly develop tools to study pathogen biology and genetics. In a field where delayed action has profound consequences for livelihoods and food security, genome sequences provide us with essential tools to prepare for the emergence of new plant pathogens and future disease outbreaks. In this regard, the medical example provided by Bill Gates (Gates 2015) that the application of available technologies could significantly have reduced the impact of the recent Ebola epidemic also holds for plant

pathology. Particularly in the era of genomics, we have significant tools to deal with the plant disease arms race and we must apply them more actively and aggressively.

ACKNOWLEDGEMENTS

We thank the National Research Foundation (NRF) and the NRF-Department of Science and Technology (DST) Centre of Excellence in Tree Health Biotechnology (CTHB) for financial support.

REFERENCES

- Aragona M, Minio A, Ferrarini A, Valente MT, Bagnaresi P, *et al.* (2014) *De novo* genome assembly of the soil-borne fungus and tomato pathogen *Pyrenochaeta lycopersici*. *BMC Genomics* **15**: 1–12.
- Aylward J, Steenkamp ET, Dreyer LL, Roets F, Wingfield MJ, Wingfield BD (2016) Genetic basis for high population diversity in *Protea*-associated *Knoxdaviesia*. *Fungal Genetics and Biology* **96**: 47–57.
- Barnes C, Szabo L (2008) A rapid method for detecting and quantifying bacterial DNA in rust fungal DNA samples. *Phytopathology* **98**: 115–119.
- Barnes I, Gaur A, Burgess T, Roux J, Wingfield BD, Wingfield MJ (2001) Microsatellite markers reflect intra-specific relationships between isolates of the vascular wilt pathogen *Ceratocystis fimbriata*. *Molecular Plant Pathology* **2**: 319–325.
- Bernal A, Ear U, Kyrpides N (2001) Genomes OnLine Database (GOLD): a monitor of genome projects world-wide. *Nucleic Acids Research* **29**: 126–127.
- Bihon W, Wingfield MJ, Slippers B, Duong TA, Wingfield BD (2014) *MAT* gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. *Fungal Genetics and Biology* **62**: 55–61.
- Buckley M (2008) *The Fungal Kingdom: diverse and essential roles in Earth's ecosystem*. Washington, DCL American Academy of Microbiology.
- Burgess T, Wingfield MJ, Wingfield BD (2001) Simple sequence repeat markers distinguish among morphotypes of *Sphaeropsis sapinea*. *Applied and Environmental Microbiology* **67**: 354–362.
- Carris L, Little C, Stiles C (2012) *Introduction to Fungi*. The Plant Health Instructor. doi: 10.1094/PHI-I-2012-0426-01.
- Castanera R, López-Varas L, Borgognone A, LaButti K, Lapidus A, *et al.* (2016) Transposable elements *versus* the fungal genome: impact on whole-genome architecture and transcriptional profiles. *PLOS Genetics* **12**: e1006108.
- Cornell MJ, Alam I, Soanes DM, Wong HM, Hedeler C, *et al.* (2007) Comparative genome analysis across a kingdom of eukaryotic organisms: specialization and diversification in the Fungi. *Genome Research* **17**: 1809–1822.
- Cuomo CA, Birren BW (2010) The Fungal Genome Initiative and lessons learned from genome sequencing. In: *Methods in Enzymology*: 833–855.
- De Wit PJ, Van Der Burgt A, Ökmen B, Stergiopoulos I, Abd-ElSalam KA, *et al.* (2012) The genomes of the fungal plant pathogens *Cladosporium fulvum* and *Dothistroma septosporum* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. *PLOS Genetics* **8**: e1003088.
- DiGuistini S, Wang Y, Liao NY, Taylor G, Tanguay P, *et al.* (2011) Genome and transcriptome analyses of the mountain pine beetle–fungal symbiont *Grosmannia clavigera*, a lodgepole pine pathogen. *Proceedings of the National Academy of Sciences, USA* **108**: 2504–2509.
- Doweld AB (2013) Nomenclatural novelties. *Index Fungorum* **42**: 1–2.
- Duplessis S, Cuomo CA, Lin Y–C, Aerts A, Tisserant E, *et al.* (2011) Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proceedings of the National Academy of Sciences, USA* **108**: 9166–9171.
- Ellwood SR, Liu Z, Syme RA, Lai Z, Hane JK, *et al.* (2010) A first genome assembly of the barley fungal pathogen *Pyrenophora teres* f. *teres*. *Genome Biology* **11**: 1–14.
- FAO (1995) *Dimensions of Need: an atlas of food and agriculture*. Rome: FAO.
- Friesen TL, Stukenbrock EH, Liu Z, Meinhardt S, Ling H, *et al.* (2006) Emergence of a new disease as a result of interspecific virulence gene transfer. *Nature Genetics* **38**: 953–956.
- Fungal Research Community (2002) *Fungal Genome Initiative (white paper)*.
- Gates B (2015) The next epidemic - Lessons from Ebola. *New England Journal of Medicine* **372**: 1381–1384.
- Gautam H, Bhardwaj M, Kumar R (2013) Climate change and its impact on plant diseases. *Current Science* **105**: 1685–1691.
- Genome 10K Community of Scientists (2009) Genome 10K: a proposal to obtain whole-genome sequence for 10 000 vertebrate species. *Journal of Heredity* **100**: 659–674.
- Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, *et al.* (1996) Life with 6000 Genes. *Science* **274**: 546–567.
- Grigoriev IV, Cullen D, Goodwin SB, Hibbett D, Jeffries TW, *et al.* (2011) Fueling the future with fungal genomics. *Mycology* **2**: 192–209.
- Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, *et al.* (2012) The Genome Portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Research* **40**: D26–D32.
- Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, *et al.* (2013) MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Research*: 10.1093/nar/gkt1183.
- Haasbroek MP, Craven M, Barnes I, Crampton BG (2014) Microsatellite and mating type primers for the maize and sorghum pathogen, *Exserohilum turcicum*. *Australasian Plant Pathology* **43**: 577–581.
- Hawksworth DL (2012) Global species numbers of fungi: are tropical studies and molecular approaches contributing to a more robust estimate? *Biodiversity and Conservation* **21**: 2425–2433.
- Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, *et al.* (2007) A higher-level phylogenetic classification of the *Fungi*. *Mycological Research* **111**: 509–547.
- Islam MS, Haque MS, Islam MM, Emdad EM, Halim A, *et al.* (2012) Tools to kill: genome of one of the most destructive plant pathogenic fungi *Macrophomina phaseolina*. *BMC Genomics* **13**: 1–16.
- Jones TM, Anderson AJ, Albersheim P (1972) Host-pathogen interactions IV. Studies on the polysaccharide-degrading enzymes secreted by *Fusarium oxysporum* f. sp. *lycopersici*. *Physiological Plant Pathology* **2**: 153–166.
- Kamper J, Kahmann R, Bolker M, Ma L–J, Brefort T, *et al.* (2006) Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* **444**: 97–101.

- Keller NP, Turner G, Bennett JW (2005) Fungal secondary metabolism - from biochemistry to genomics. *Nature Reviews Microbiology* **3**: 937–947.
- Kelman A (1985) Plant pathology at the crossroads. *Annual Review of Phytopathology* **23**: 1–12.
- Kidwell MG (2002) Transposable elements and the evolution of genome size in eukaryotes. *Genetica* **115**: 49–63.
- Kirk PM, Cannon PF, Minter DW, Stalpers JA (2008) *Ainsworth & Bisby's Dictionary of the Fungi*. 10th edn. Wallingford: CAB International.
- Klosterman SJ, Subbarao KV, Kang S, Veronese P, Gold SE, et al. (2011) Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. *PLOS Pathogens* **7**: e1002137.
- Kroken S, Glass NL, Taylor JW, Yoder O, Turgeon BG (2003) Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proceedings of the National Academy of Sciences, USA* **100**: 15670–15675.
- Kronstad JW, Staben C (1997) Mating type in filamentous fungi. *Annual Review of Genetics* **31**: 245–276.
- Kullman B, Tamm H, Kullman K (2005) *Fungal Genome Size Database*. <http://www.zbi.ee/fungal-genomesize>.
- Lee B-N, Kroken S, Chou DYT, Robbertse B, Yoder OC, Turgeon BG (2005) Functional analysis of all nonribosomal peptide synthetases in *Cochliobolus heterostrophus* reveals a factor, NPS6, involved in virulence and resistance to oxidative stress. *Eukaryotic Cell* **4**: 545–555.
- Liu J, Yuan Y, Wu Z, Li N, Chen Y, et al. (2015) A novel sterol regulatory element-binding protein gene *sreA* identified in *Penicillium digitatum* is required for prochloraz resistance, full virulence and *erg11* (*cyp51*) regulation *PLOS One* **10**: e0117115.
- Ma L-J, Van Der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, et al. (2010) Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* **464**: 367–373.
- Malapi-Wight M, Demers JE, Veltri D, Marra RE, Crouch JA (2016) LAMP Detection assays for Boxwood Blight pathogens: a comparative genomics approach. *Scientific Reports* **6**: 2640.
- Maloy OC (2005) Plant disease management. *The Plant Health Instructor* **10**: DOI: 10.1094/PHI-I-2005-0202-1001
- Manning VA, Pandelova I, Dhillon B, Wilhelm LJ, Goodwin SB, et al. (2013) Comparative genomics of a plant-pathogenic fungus, *Pyrenophora tritici-repentis*, reveals transduplication and the impact of repeat elements on pathogenicity and population divergence. *Genes|Genomes|Genetics* **3**: 41–63.
- Marcet-Houben M, Gabaldón T (2010) Acquisition of prokaryotic genes by fungal genomes. *Trends in Genetics* **26**: 5–8.
- Marcet-Houben M, Ballester A-R, de la Fuente B, Harries E, Marcos JF, et al. (2012) Genome sequence of the necrotrophic fungus *Penicillium digitatum*, the main postharvest pathogen of citrus. *BMC Genomics* **13**: 1–18.
- Martin F, Cullen D, Hibbett D, Pisabarro A, Spatafora JW, Baker SE, Grigoriev IV (2011) Sequencing the fungal tree of life. *New Phytologist* **190**: 818–821.
- McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**: 349–379.
- McDonald MC, McDonald BA, Solomon PS (2015) Recent advances in the *Zymoseptoria tritici*-wheat interaction: insights from pathogenomics. *Frontiers in Plant Science* **6**.
- McTaggart AR, Van der Nest MA, Steenkamp ET, Roux J, Slippers B, et al. (2016) Fungal genomics challenges the dogma of name-based biosecurity. *PLOS Pathogens* **12**: e1005475.
- O'Connell RJ, Thon MR, Hacquard S, Amyotte SG, Kleemann J, et al. (2012) Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nature Genetics* **44**: 1060–1065.
- Odds FC (2005) Genomics, molecular targets and the discovery of antifungal drugs. *Revista Iberoamericana de Micología* **22**: 229–237.
- Ohm RA, Feau N, Henrissat B, Schoch CL, Horwitz BA, et al. (2012) Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen *Dothideomycetes* fungi. *PLOS Pathogens* **8**: e1003037.
- Olson Å, Aerts A, Asiegbu F, Belbahri L, Bouzid O, et al. (2012) Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. *New Phytologist* **194**: 1001–1013.
- Park J, Park B, Jung K, Jang S, Yu K, et al. (2008) CFGP: a web-based, comparative fungal genomics platform. *Nucleic Acids Research* **36**: D562–D571.
- Perlin MH, Amselem J, Fontanillas E, Toh SS, Chen Z, et al. (2015) Sex and parasites: genomic and transcriptomic analysis of *Microbotryum lychnidis-dioicae*, the biotrophic and plant-castrating anther smut fungus. *BMC Genomics* **16**: 1–24.
- Pieck ML, Ruck A, Farman M, Peterson GL, Stack JP, et al. (2016) Genomics-based marker discovery and diagnostic assay development for Wheat Blast. *Plant Disease*: <http://dx.doi.org/10.1094/PDIS-04-16-0500-RE>.
- Raffaele S, Kamoun S (2012) Genome evolution in filamentous plant pathogens: why bigger can be better. *Nature Reviews Microbiology* **10**: 417–430.
- Reddy T, Thomas AD, Stamatis D, Bertsch J, Isbandi M, et al. (2014) The Genomes OnLine Database (GOLD) v. 5: a metadata management system based on a four level (meta) genome project classification. *Nucleic Acids Research*: gku950.
- Robert V, Vu D, Amor ABH, Van de Wiele N, Brouwer C, J et al. (2013) MycoBank gearing up for new horizons. *IMA Fungus* **4**: 371–379.
- Robinson GE, Hackett KJ, Purcell–Miramontes M, Brown SJ, Evans JD, et al. (2011) Creating a buzz about insect genomes. *Science* **331**: 1386.
- Schatz MC (2015) *The Next 20 years of Genome Research*. doi: 10.1101/020289.
- Schirawski J, Mannhaupt G, Münch K, Brefort T, Schipper K, et al. (2010) Pathogenicity determinants in smut fungi revealed by genome comparison. *Science* **330**: 1546–1548.
- Schneider DJ, Collmer A (2010) Studying plant–pathogen interactions in the genomics era: beyond molecular Koch's postulates to systems biology. *Annual Review of Phytopathology* **48**: 457–479.
- Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, et al. (2012) A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**: 116–120.
- Slot JC, Rokas A (2011) Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. *Current Biology* **21**: 134–139.
- Soanes D, Richards TA (2014) Horizontal gene transfer in eukaryotic plant pathogens. *Annual Review of Phytopathology* **52**: 583–614.
- Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, S et al.

- (2010) Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330**: 1543–1546.
- Spatafora J (2011) 1000 fungal genomes to be sequenced. *IMA Fungus* **2**: (41).
- Spatafora JW, Chang Y, Benny GL, Lazarus K, Smith ME, *et al.* (2016) A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* **108**: 1028–1046.
- Spitzer M, Wildenhain J, Rappsilber J, Tyers M (2014) BoxPlotR: a web tool for generation of box plots. *Nature Methods* **11**: 121–122.
- Stahl EA, Bishop JG (2000) Plant–pathogen arms races at the molecular level. *Current Opinion in Plant Biology* **3**: 299–304.
- Stukenbrock EH, Bataillon T, Dutheil JY, Hansen TT, Li R, Z *et al.* (2011) The making of a new pathogen: Insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species. *Genome Research* **21**: 2157–2166.
- Sun B-F, Xiao J-H, He S, Liu L, Murphy RW, Huang D-W (2013a) Multiple interkingdom horizontal gene transfers in *Pyrenophora* and closely related species and their contributions to phytopathogenic lifestyles. *PLOS One* **8**: e60029.
- Sun X, Ruan R, Lin L, Zhu C, Zhang T, *et al.* (2013b) Genomewide investigation into DNA elements and ABC transporters involved in imazalil resistance in *Penicillium digitatum*. *FEMS Microbiology Letters* **348**: 11–18.
- The Fungal Genome Initiative Steering Committee (2003) *A White Paper for Fungal Comparative Genomics*. Cambridge MA: Whitehead Institute; http://www.broadinstitute.org/annotation/fungi/fgi/FGI_02_whitepaper_2003.pdf.
- The Fungal Genome Initiative Steering Committee (2004) *A White Paper for Fungal Genomics*. Cambridge, MA: The Broad Institute of Harvard and MIT.
- Thon MR, Pan H, Diener S, Papalas J, Taro A, *et al.* (2006) The role of transposable element clusters in genome evolution and loss of synteny in the rice blast fungus *Magnaporthe oryzae*. *Genome Biology* **7**: 1–9.
- Thynne E, McDonald MC, Solomon PS (2015) Phytopathogen emergence in the genomics era. *Trends in Plant Science* **20**: 246–255.
- Urban M, Pant R, Raghunath A, Irvine AG, Pedro H, Hammond-Kosack KE (2015) The Pathogen–Host Interactions database (PHI-base): additions and future developments. *Nucleic Acids Research* **43**: D645–D655.
- Vanheule A, Audenaert K, Warris S, Van de Geest H, Schijlen E, *et al.* (2016) Living apart together: crosstalk between the core and supernumerary genomes in a fungal plant pathogen. *BMC Genomics* **17**: 1–18.
- Wibberg D, Andersson L, Tzelepis G, Rupp O, Blom J, *et al.* (2016) Genome analysis of the sugar beet pathogen *Rhizoctonia solani* AG2–2IIIB revealed high numbers in secreted proteins and cell wall degrading enzymes. *BMC Genomics* **17**: 1–12.
- Wilson AM, Wilken PM, Van der Nest MA, Steenkamp ET, Wingfield MJ, Wingfield BD (2015) Homothallism: an umbrella term for describing diverse sexual behaviours. *IMA Fungus* **6**: 207.
- Withers S, Gongora-Castillo E, Gent D, Thomas A, Ojiambo PS, Quesada-Ocampo LM (2016) Using next-generation sequencing to develop molecular diagnostics for *Pseudoperonospora cubensis*, the Cucurbit Downy Mildew pathogen. *Phytopathology* **106**: 1105–1116.
- Wöstemeyer J, Kreibich A (2002) Repetitive DNA elements in fungi (*Mycota*): impact on genomic architecture and evolution. *Current Genetics* **41**: 189–198.
- Zhu S, Cao Y-Z, Jiang C, Tan B-Y, Wang Z, *et al.* (2012) Sequencing the genome of *Marssonina brunnea* reveals fungus–poplar co-evolution. *BMC Genomics* **13**: 382.