Genome sequencing provides insight into the reproductive biology, nutritional mode and ploidy of the fern pathogen *Mixia osmundae*

Merje Toome[1], Robin A. Ohm[2], Robert W. Riley[2], Timothy Y. James[3], Katherine L. Lazarus[3], Bernard Henrissat[4], Sebastian Albu[5], Alexander Boyd[6], Julianna Chow[2], Alicia Clum[2], Gregory Heller[5], Anna Lipzen[2], Matt Nolan[2], Laura Sandor[2], Natasha Zvenigorodsky[2], Igor V. Grigoriev[2], Joseph W. Spatafora[6] and M. Catherine Aime[1]

[1]Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907, USA; [2]US Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA; [3]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA; [4]Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille University, CNRS UMR 7257, 13288 Marseille, France; [5]Department of Plant Pathology and Crop Physiology, Louisiana State University Agricultural Center, Baton Rouge, LA 70803, USA; [6]Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331, USA

Author for correspondence:
M. Catherine Aime
Tel: +1 765 496 7853
Email: maime@purdue.edu

Received: *13 June 2013*
Accepted: *19 November 2013*

New Phytologist (2014) **202**: 554–564
doi: 10.1111/nph.12653

Key words: biotrophic fungi, CAZy, fungal genomics, Mixiomycetes, *Osmunda* spp., sporogenesis.

Summary

- *Mixia osmundae* (Basidiomycota, Pucciniomycotina) represents a monotypic class containing an unusual fern pathogen with incompletely understood biology. We sequenced and analyzed the genome of *M. osmundae*, focusing on genes that may provide some insight into its mode of pathogenicity and reproductive biology.
- *Mixia osmundae* has the smallest plant pathogenic basidiomycete genome sequenced to date, at 13.6 Mb, with very few repeats, high gene density, and relatively few significant gene family gains.
- The genome shows that the yeast state of *M. osmundae* is haploid and the lack of segregation of mating genes suggests that the spores produced on *Osmunda* spp. fronds are probably asexual. However, our finding of a complete complement of mating and meiosis genes suggests the capacity to undergo sexual reproduction. Analyses of carbohydrate active enzymes suggest that this fungus is a biotroph with the ability to break down several plant cell wall components.
- Analyses of publicly available sequence data show that other *Mixia* members may exist on other plant hosts and with a broader distribution than previously known.

Introduction

Mixia osmundae (Nishida) C.L. Kramer (Pucciniomycotina, Mixiomycetes) is a rarely encountered basidiomycete fungus that has only been isolated from the fronds of ferns belonging to the genus *Osmunda* L. (Pteridophyta, Polypodiopsida). The only records of *M. osmundae* are from Japan and Taiwan on *Osmunda japonica* Thunb. (Nishida *et al.*, 1995; Sugiyama & Katumoto, 2008) and from the United States (Georgia and Michigan) on *Osmunda cinnamomea* L. (Mix, 1947; Kramer, 1958). Mixiomycetes is one of only three monotypic classes described in Dikarya (i.e. Basidiomycota and Ascomycota) and the only one in Pucciniomycotina (Aime *et al.*, 2006; Bauer *et al.*, 2006).

Most studies of *M. osmundae* have been confined to the symptomatic phase on *Osmunda* spp. Infection is recognizable by the small, yellow or brown lesions produced on the fern fronds that develop a powdery white layer of spores (Mix, 1947). *Mixia osmundae* is reported to grow within the plant cell walls, initially forming sparsely septate hyphae that later become coenocytic, containing numerous small nuclei (Mix, 1947; Kramer, 1958). Spores are exogenously and simultaneously produced on the surface of a large sac-like cell that forms from an enlarged swelling within the host epidermal cell wall (Nishida *et al.*, 1995). This kind of sporogenous structure and the manner of spore formation is unique among fungi. Owing to its sac-like appearance, the sporogenous cell was originally interpreted as an ascus and therefore *M. osmundae* was first described in the Ascomycota as *Taphrina osmundae* Nishida (Nishida, 1911). More than 80 yr later, detailed morphological and molecular studies showed that *M. osmundae* holds a rather isolated position within the Basidiomycota, with the simple-septate species now placed together within Pucciniomycotina (Nishida *et al.*, 1995; Sugiyama, 1998; Sjamsuridzal *et al.*, 2002; Aime *et al.*, 2006).

The remaining aspects of the biology and life cycle of *M. osmundae* are not well known, except that the observed spores are monokaryotic (Nishida *et al.*, 1995). It is unknown whether these spores are haploid or diploid and, if haploid, whether they are produced as mitospores or meiospores. Since *M. osmundae* is

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust

very rarely collected and the sporogenous cells have only been observed when growing on the host, spore germination and plant infection have never been observed *in vivo*. When grown on artificial media the fungus reproduces by budding, forming creamy yeast-like colonies (Nishida *et al.*, 1995, 2011b). As there are several groups of fungi in Pucciniomycotina that produce basidiospores that can multiply by budding in culture (Swann *et al.*, 2001; Morrow & Fraser, 2009), it has been speculated that if the spores of *M. osmundae* are meiospores, then the sporogenous cell must represent a novel basidium type (Sugiyama, 1998).

Although the number of available fungal genomes is increasing rapidly, the number of publicly available sequenced genomes within Pucciniomycotina is still small (Grigoriev *et al.*, 2011). These include five species of rust fungi (Pucciniales, Pucciniomycetes) (*Puccinia graminis* f. sp. *tritici* Eriks. & Henning, *Puccinia striiformis* f. *tritici* Erikss., *Puccinia triticina* Erikss., *Melampsora larici-populina* Kleb. and *Cronartium quercuum* f. sp. *fusiforme* Burds. & G.A. Snow) and five species of Microbotryomycetes (*Sporobolomyces roseus* Kluyver & C.B. Niel, *Rhodotorula graminis* Di Menna, *Rhodotorula glutinis* (Frensen.) F.C. Harrison, *Rhodosporidium toruloides* I. Banno (Sporidiobolales) and *Microbotryum violaceum* (Pers.) G. Deml & Oberw. (Microbotryales)) (Cantu *et al.*, 2011; Duplessis *et al.*, 2011; Kumar *et al.*, 2012; www.jgi.doe.gov/fungi; www.broadinstitute.com; www.ncbi.nlm.nih.gov). Thus, of the nine classes and 20 orders described to date within Pucciniomycotina (Schell *et al.*, 2011; Toome *et al.*, 2013; Aime *et al.*, 2014), complete annotated genomes are only available for two classes and three orders; some incomplete and unannotated genomic data are available for seven additional Pucciniomycotina species – six additional rust fungi and a draft genome for *M. osmundae* (Nishida *et al.*, 2011a; www.ncbi.nlm.nih.gov).

Some early researchers considered *M. osmundae* to represent the closest extant taxon descended from the common ancestor of Dikarya (Savile, 1955; Kramer, 1987). While modern DNA sequencing studies have been unable to resolve the branching order of the earliest basidiomycetes, they do show *M. osmundae* as a rather isolated lineage within Pucciniomycotina (Aime *et al.*, 2006), which in turn has been supported as the earliest branching lineage of Basidiomycota (Padamsee *et al.*, 2012). In this study we generated genome data for *M. osmundae* to provide clues to various aspects of its biology, including nutritional mode and reproductive biology, and examined publicly available environmental sequence data to determine whether *Mixia* is possibly more widespread than previously suspected.

Material and Methods

Nucleic acid extraction

A culture of strain IAM14324 was used for the genome sequencing of *Mixia osmundae* (Nishida) C.L. Kramer. This single-spore isolate originated from a collection of *O. japonica* made in May 1993 in Sezawa (Haibara-gun, Nakakawane-cho, Shizuoka Prefecture), Japan (Nishida *et al.*, 1995), and was obtained from the Japan Collection of Microorganisms (Tsukuba, Ibaraki Prefecture, Japan). The cells were grown on minimal media for 2 wk, harvested and ground in liquid nitrogen. DNA was extracted following a modified CTAB protocol described in Padamsee *et al.* (2012). For RNA extraction, a 5-d-old culture grown on Difco™ potato dextrose agar (PDA; Becton, Dickinson and Co., Sparks, MD, USA) was used. RNA was isolated using a RiboPure™ Yeast Kit (Ambion), followed by DNase treatment with TURBO DNase (Ambion), all following the manufacturer's recommendations.

Genome sequencing and annotation

For production of the *M. osmundae* genome, one unamplified whole-genome shotgun (WGS) and one 4 kb Long Mate Pair (LMP) library were generated. Quantified sample libraries were clustered to flow cells and sequenced on an Illumina HiSeq 2000 platform (Illumina, San Diego, CA, USA). For transcriptome sequencing, poly-A RNA was isolated from total RNA and fragmented for first-strand cDNA synthesis. Digested cDNA was amplified by PCR with Illumina Truseq primers and sequenced on the Illumina HiSeq 2000 platform, generating 100 bp paired-end reads.

The genome was annotated using the JGI annotation Pipeline (Grigoriev *et al.*, 2006), which combines several gene prediction and annotation methods, and integrates the annotated genome into the web-based fungal resource MycoCosm (Grigoriev *et al.*, 2012) for comparative genomics. Genome assembly and annotations can be interactively accessed through the JGI fungal genome portal MycoCosm (Grigoriev *et al.*, 2012) at http://jgi.doe.gov/Mosmundae and have also been deposited to DDBJ/EMBL/GenBank under accession no. AYOQ00000000. The version described in this paper is version AYOQ01000000.

Translated protein sequences were assigned to carbohydrate-active enzyme (CAZy) families using the CAZy pipeline (Cantarel *et al.*, 2009) as described in Floudas *et al.* (2012). Multigene families were identified using the JGI clustering pipeline and the core proteome was determined by identifying multigene families that contained at least one member in each of the studied organisms. To assess whether the assembly represents a haploid or diploid genome, we searched for syntenic blocks of genes that are present twice in the genome.

For a more detailed description of the sequencing, assembly and annotation, see the Supporting Information, Methods S1.

Phylogenetic and CAFE analysis

Seventy-one previously determined orthologous protein sequences were obtained from 23 fungal taxa (listed in Table S1) using a modified version of the Hal pipeline (Robbertse *et al.*, 2011). Protein sequences were aligned using MUSCLE (Edgar, 2004) and poorly aligned regions were removed with GBlocks (Castresana, 2000). A superalignment was analyzed using RAxML with a GAMMAPROT model of evolution and branch support was assessed using 100 RAxML bootstraps (Stamatakis, 2006). The resulting tree from the phylogenetic analysis (Fig. 1) was used in CAFE analysis.

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust

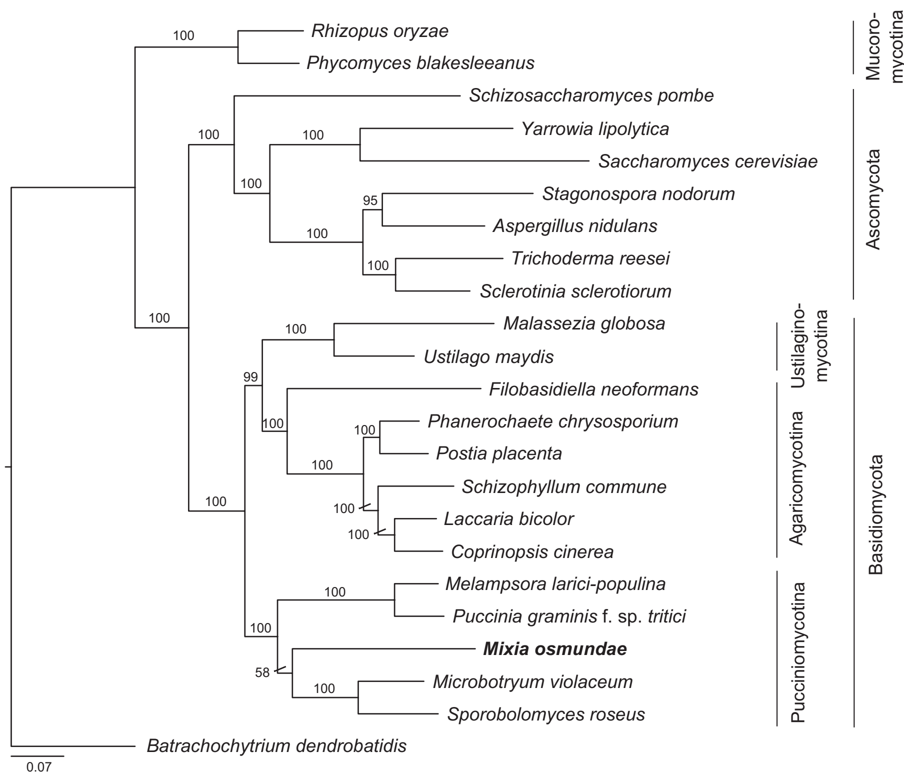

Fig. 1 Consensus tree from RAxML analyses of 71 orthologous proteins from representative fungi with sequenced genomes, including *Mixia osmundae* (in bold). Numbers above branches indicate RAxML bootstrap support values. The chytrid *Batrachochytrium dendrobatidis* was selected as the outgroup.

The CAFE program (De Bie *et al.*, 2006) was used to analyze gain and loss of 3119 Pfams (Punta *et al.*, 2012) found in *M. osmundae* and 22 fungi for comparison (Table S1). To identify families significantly deviating from a random birth/death model of gene evolution, as well as to infer significant branch-specific gains and losses (Viterbi *P*-values), we used a significance threshold of 0.05. Additional information is available in Methods S1.

Putative effector and avirulence genes

Homologs of 51 known fungal avirulence genes (Table S2) were identified by a blastp analysis, using an E-value cutoff of 1E–3. Small secreted proteins (SSPs) were defined as proteins that are smaller than 300 amino acids, have a secretion signal as determined by SignalP 3.0 (Bendtsen *et al.*, 2004), and no transmembrane domain (TMM) as determined by TMHMM 2.0 (Krogh *et al.*, 2001). However, one transmembrane domain is allowed when present in the N-terminal 40 amino acids, as this often corresponds to the secretion signal.

Mating loci and meiosis genes

Only four live cultures of *M. osmundae* are known to exist, all derived from single spore isolates: JCM22182 (= IAM14324), JCM22183 (= IAM14325), JCM22814 (= IAM14326) and JCM22200 (= IAM14511). All four cultures were obtained from the Japan Collection of Microorganisms, maintained on PDA, and DNA was extracted using a MasterPure™ Yeast DNA Purification Kit (Epicentre, Madison, WI, USA). Genes for SXI2, STE3 and MoPh1 were sequenced from each isolate and deposited in GenBank (KF672709–KF672711). See Methods S1 for more details on culture origins and methods.

Protein sequences for the genes involved in meiosis were obtained from the *Saccharomyces* Genome Database (Cherry *et al.*, 2012). The orthologous genes in *M. osmundae* were determined using the BLAST algorithm in MycoCosm (Grigoriev *et al.*, 2012). Detailed search parameters are described in Methods S1.

Screening public sequence databases for Mixiomycetes

A BLASTn search on the NCBI GenBank database (blast.ncbi.nlm.nih.gov) was conducted to locate any sequences sharing high identity with those from *M. osmundae*. Searches were carried out with the entire internal transcribed spacer (ITS) region as well as with the ITS1, 5.8S and ITS2 regions separately. Any sequences sharing at least 95% identity with the *M. osmundae* sequence were selected and aligned using the Muscle algorithm in MEGA 5.10 (Kumar *et al.*, 2008). Additionally, the ITS region was amplified and sequenced from JCM22183, JCM22814 and JCM22200 with primers ITS1F (Gardes & Bruns, 1993) and ITS4 (White *et al.*, 1990), following the protocol described in Toome *et al.* (2013). The ITS sequences are deposited in GenBank (KF559352–KF559354). Maximum likelihood analyses were conducted in RAxML-HPC2 7.2.8 (Stamatakis, 2006; Stamatakis *et al.*, 2008) via the CIPRES Science Gateway (Miller *et al.*, 2010) using the –k option for bootstrap analysis. A selection of the most closely related taxa was made for the final analysis and tree construction.

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust

Results

Genome

The genome of *M. osmundae* (IAM14324) is estimated to be 13.63 Mb in size (150 × sequence read coverage). The sequences are assembled into 204 contigs and 156 scaffolds. The longest scaffold is 2.3 Mb; the next five are all longer than 1 Mb and the first 10 are all longer than 0.5 Mb, together representing *c.* 85% of the genome. There are only 48 scaffold gaps in the entire assembly and a Core Eukaryotic Genes Mapping Approach (CEGMA) analysis (Parra *et al.*, 2007) indicated that the assembly and annotation are 98.3% complete (450 out of 458 CEGMA genes are present). Annotation of the genome resulted in a set of 6903 gene models, 98.7% of which were supported by transcriptome data (96.5% of the models were supported by transcripts over at least 75% of their length, and 82.9% were supported by transcripts over 100% of their length). All the genes and SSPs reported herein were expressed by the fungus growing in pure culture on defined media.

Median gene length in *M. osmundae* is 1528 bp, which is longer than reported in any sequenced Pucciniomycetes to date, but shorter than in the sequenced Microbotryomycetes. Protein length is also medial compared with known genomes from the other two Pucciniomycotina classes. Gene density is 506 genes Mb^{-1} for *M. osmundae* and between 165 and 346 genes Mb^{-1} for the other Pucciniomycotina genomes. This high density of genes is also evidenced by the extremely small number of repeats (0.15% of assembly; 399 repeat-covered regions, 20 909 bp in total) and very short intergenic distance (143 bp in median) within the genome. In addition, although 92% of all genes have at least one intron, these are much shorter than in other Pucciniomycotina (54 bp in median compared with 82 bp average median for the others). Exons in *M. osmundae* are, on average, considerably longer (184 bp) than those found in other Pucciniomycotina genomes (141–158 bp). The mean numbers of introns and exons per gene were three and four, respectively, which is the same found in other Pucciniomycetes genomes, but fewer than in Microbotryomycetes. The guanine-cytosine (GC) content of *M. osmundae* (55.5%) is the highest known in Pucciniomycotina, which is a reflection of the lack of AT-rich repetitive elements that are almost nonexistent in this genome. Detailed genome statistics for all of the published and annotated Pucciniomycotina genomes are presented in Table S3.

The total coding sequence length of the genome is 11.85 Mb and over 70% of the genes have a Swissprot match as well as a functional annotation. The largest protein groups are associated with carbohydrate metabolism, followed by amino acid and lipid metabolism. Gene ontology analysis indicates that most (> 50%) of the gene models have a molecular function (GO:0003674), more than a quarter are associated with biological processes (GO:0008150), and *c.* 20% are predicted to be associated with cellular components (GO:0005575).

The gene models that are associated with different metabolic pathways are proportional to the total number of genes when comparing *M. osmundae* Kyoto Encyclopedia of Genes and Genomes (KEGG) annotations with other Pucciniomycotina genomes. The only categories where *M. osmundae* seems to have proportionally fewer genes are in carbohydrate metabolism, lipid metabolism and glycan biosynthesis and metabolism (Fig. S1). Comparison of the gene models in EuKaryotic Orthologous Groups (KOG) classification showed that *M. osmundae* has fewer loci responsible for extracellular structures, defense mechanisms and signal transduction mechanisms (Fig. S2).

An analysis of conservation of the predicted *M. osmundae* proteins revealed that 2703 out of the 6903 genes were conserved in all analyzed Pucciniomycotina and Ustilaginomycotina, constituting the core proteome. The number of proteins in the core proteome was similar across the analyzed fungi (Fig. S3). An enrichment analysis of GO terms in the core proteome showed that a wide variety of functions were overrepresented (Table S4), probably representing, in the main, the housekeeping genes. SSPs were underrepresented in the core proteome for these fungi.

Carbohydrate-active enzymes

Analyses of carbohydrate-active enzymes detected 78 enzymes belonging to 28 glycoside hydrolase (GH) families, 109 enzymes from 32 glycosyl transferase (GT) families, four enzymes from two polysaccharide lyase (PL) families and 14 enzymes from four carbohydrate esterase (CE) families. Additionally, five carbohydrate-binding module (CBM) families were detected with a total copy number of eight (Table S5).

We detected 24 putative GH enzymes and two putative PL enzymes that are involved in plant biomass degradation (based on Van den Brink & de Vries, 2011; for a list of these enzymes see Table S6), which is fewer than were found in Pucciniales, but more than were found in the saprobic yeast *S. roseus. M. osmundae* has a high number of cellulases and xyloglucanases (GH12, GH45). These enzymes are not widely distributed in Pucciniomycotina; thus far GH45 has been found only in *M. osmundae* and GH12 in *M. osmundae* and rust fungi. The genome also contains three copies of GH26 (β-mannanase), which is more copies than have been found in most other basidiomycetes, indicating that *M. osmundae* could be especially efficient at breaking down mannan in the cell walls of plants. Additionally, *M. osmundae* has one copy of GH93, an enzyme that cleaves α-linked arabinose residues in hemicellulose. This enzyme family was not found in any of the other studied Pucciniomycotina or Ustilaginomycotina genomes.

No evidence was found of genes encoding GH3, GH6, GH7 or GH61 enzymes, which are necessary for cellulose digestion and for converting cellobiose to glucose. Additionally there were no genes encoding for GH10 or GH11 enzymes, implying that *M. osmundae* is not able to break down xylan either. Therefore, although *M. osmundae* seems to possess a set of enzymes that can be used to break down cellulose, it lacks the enzyme sets necessary for depolymerizing it to simple sugars. The genome does encode the GH32 enzyme, however, suggesting the ability to use sucrose and fructose as energy sources.

There are only three PL families represented in Pucciniomycotina in general. PL14 (alginate lyase) was the most evenly

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust

distributed family, which was present in all studied Pucciniomycotina genomes, except for *P. graminis* f. sp. *tritici*; the PL1 (pectate lyase) family was represented by four copies in the rust fungi, but none in *M. osmundae*; another pectate lyase family, PL3, was represented by two copies in *M. osmundae* but was absent from all other Pucciniomycotina species. No copies of GH88 or GH105 – the enzymes that use the products of PL families as a substrate – were found in any of the examined Pucciniomycotina genomes.

Finally, there are several enzyme families present in the *M. osmundae* genome that have no precise known function. The most outstanding GT family was GT31, which is represented by 12 copies in *M. osmundae* and by only three to five in all of the other examined species. Among CE families, CE12 was represented by five copies in *M. osmundae* and was not found in any of the other representatives of Pucciniomycotina. As the function of these families is largely unexplored, their role in the biology of *M. osmundae* remains to be determined.

Phylogenetic position

Analysis of a selection of protein-coding regions confirmed that *M. osmundae* belongs to Pucciniomycotina as shown by previous multilocus studies (Aime *et al.*, 2006). However, the position of *M. osmundae* within Pucciniomycotina cannot be satisfactorily resolved at this time because of limited taxon sampling. Our analyses included genomes from only three of the nine classes of Pucciniomycotina; in these *M. osmundae* is resolved as a sister group to Microbotryomycetes rather than Pucciniomycetes (Fig. 1).

Gene family evolution

The CAFE program (De Bie *et al.*, 2006), which uses a stochastic model of gene birth and death given a phylogenetic tree and a table of gene copy number for each organism in the tree, was used to estimate the gains and losses in gene families (defined by assigning genes to Pfam categories). The gene families with statistically significant losses in *M. osmundae* (16 Pfams) noticeably outnumbered the ones with gains (eight Pfams), which is another correlate of a smaller genome of *M. osmundae*. Further analysis of *M. osmundae* compared with its common node with Microbotryomycetes showed that the gene families with the largest gains were aspartate protease, short-chain hydrogenases, ankyrin repeats and cytochrome P450. Considerable gene loss was detected for sugar transporters and helicase conserved C-terminal domains. As noted in the CAZy analysis section, *M. osmundae* lacks the GH3 family, and CAFE analysis found that this is most likely the result of a recent gene loss. *M. osmundae* also has notable losses in Cu oxidase and ubiquitin families (Fig. 2b, C; Table S7).

Few losses or gains were found to have occurred between the last common ancestor of Pucciniomycotina and Mixiomycetes + Microbotryomycetes (Fig. 2, B). This probably indicates that most extant members of these groups are similar to their common ancestor in this respect. On the other hand, when comparing the Pucciniomycetes and the common ancestor of Pucciniomycotina, there were drastic changes detected in many gene families (Table S7), as was also reported by Duplessis *et al.* (2011). In general, the common ancestor of Pucciniomycotina exhibits few differences from the common ancestor to all Basidiomycota (Fig. 2, A). Greater gene family gain was detected only for ankyrin repeats and cytochrome P450; gene losses were significant for zinc fingers, fungal transporters and ABC transporters. More detailed CAFE results are presented in Table S7.

Pathogenicity-related regions

Out of a selection of 51 known fungal avirulence genes, only three were found in *M. osmundae* (Table S2). The genome encodes 530 secreted proteins (7.7% of total proteins), of which 284 were categorized as SSPs (53% of all secreted proteins). The SSPs had a higher average cysteine percentage than other secreted proteins and only 8% of the SSPs had a Pfam domain. In comparison to other Pucciniomycotina and Ustilaginomycotina genomes, *M. osmundae* has fewer SSPs than rust fungi but more than *S. roseus*, *Ustilago maydis* or *Malassezia globosa*. The cysteine content of the SSPs of *M. osmundae* was more similar to rust fungi than to other compared organisms (Table S8).

A conservation analysis revealed that predicted secreted proteins and SSPs are relatively poorly conserved in related fungi (Fig. S4). The majority of predicted secreted proteins and SSPs were unique to *M. osmundae*, which contrasts with the conservation of the total proteome. Sixteen SSPs were found in at least one other species belonging to the Pucciniomycotina (*S. roseus*, *P. graminis* and *M. larici-populina*), and 23 were also found in at least one of the Ustilaginomycotina (*U. maydis* and *M. globosa*). Only 10 SSPs were part of the core proteome, which means that they are conserved in all six studied Basidiomycota species (Table S9). No annotation terms were statistically significantly overrepresented in the SSPs of *M. osmundae* (data not shown).

Mating type (MAT) loci and meiosis genes

Homologs of the STE3-like pheromone receptors and homeodomain proteins, known to be encoded by MAT loci in Basidiomycota (Kües *et al.*, 2011), were identified from the sequenced *M. osmundae* genome. The first MAT locus (on scaffold 1) contains two adjacent divergently transcribed homeodomain (HD) genes of the HD1 and HD2 types, showing a similar arrangement as observed in other Basidiomycota (Kües *et al.*, 2011). The second MAT locus (on scaffold 7) contains a single putative pheromone receptor (MoSTE3) and two pheromone precursor genes (MoPh1 and MoPh2; Table S10). The genes MoPh1 and MoPh2 are both predicted to produce three nearly identical mature pheromones that are farnesylated (Fig. S5). The comparisons of the MAT loci by DNA sequencing detected no polymorphism among the four available *M. osmundae* single spore isolates.

Several orthologs of genes shown to be involved in meiosis in other fungi were identified in the *M. osmundae* genome. Altogether, 87 out of 127 searched meiosis-related genes were

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust

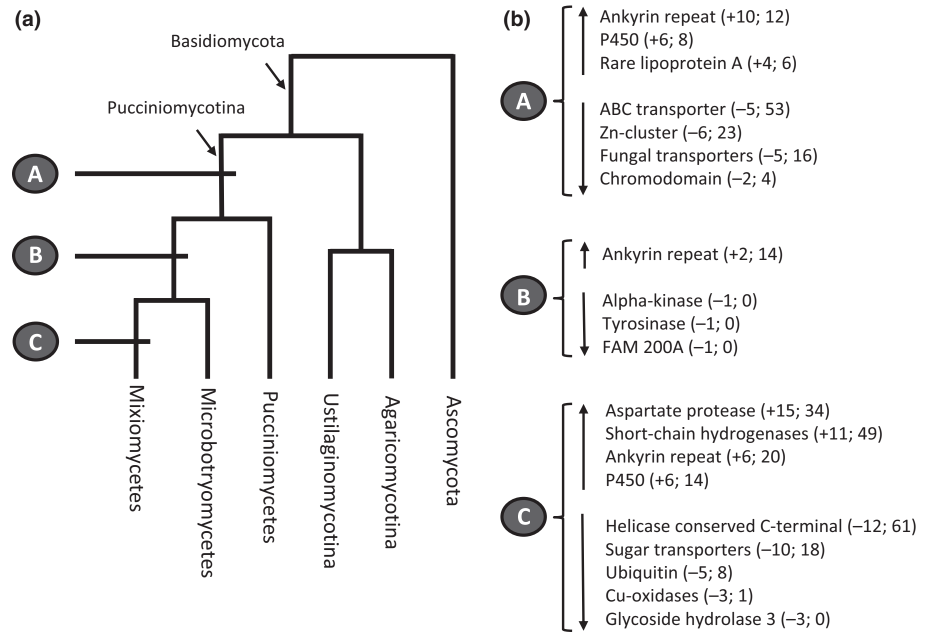

Fig. 2 Gene family gains and losses in Pucciniomycotina. Groups examined in the CAFE analysis are indicated in (a), and significant changes in gene family numbers are listed in (b). A, the common ancestor of all Pucciniomycotina compared with Basidiomycota; B, the common ancestor of Microbotryomycetes + Mixiomycetes compared with the common ancestor of Pucciniomycotina; C, *Mixia osmundae* compared with the common ancestor of Mixiomycetes + Microbotryomycetes. At positions A and B, all gene family gains and losses are shown, whereas at position C, only select changes are presented. In (b), gene families are ordered by the number of genes gained (up arrow) or lost (down arrow) along that branch; the first number in brackets represents the copy number of lost or gained genes and the second number shows the total copy number of this family in the depicted position. More detailed results of the CAFE analysis are presented in Table S7.

detected in *M. osmundae* (see Table S11). Of 31 previously determined core meiosis genes (Halary *et al.*, 2011), 29 orthologs were found in *M. osmundae*, missing only Hop2 and Mnd1. Two genes, Ndt80 and Ime1, determined to have a key role in the meiosis of *Saccharomyces cerevisiae* (Vershon & Pierce, 2000; and references therein) were not detected in *M. osmundae* (Table S11). However, we also failed to detect either of these two genes within any of the other sequenced Pucciniomycotina genomes.

Ploidy of the genome

Searches for syntenic gene blocks at the whole genome level failed to detect any duplicate regions. Additionally, as previously discussed, no heterozygosity or polymorphism was found between any of the putative mating gene homologs (SXI2, STE3 and MoPh1) sequenced from each of the four existent isolates of *M. osmundae*.

Environmental sequences of Mixiomycetes

A BLASTn search on GenBank uncovered five ITS sequences generated by environmental studies that are closely related to *M. osmundae*. Two of these were obtained from the leaves of a bamboo plant (*Yushania exilis* T.P. Yi) in China (Zhang *et al.*, 1997). The other three originated from an examination of phylloplane fungal communities of beech trees (*Fagus sylvatica* L.) in France (Cordier *et al.*, 2012). While the sequences from bamboo leaves contained the full ITS1, 5.8S and ITS2 regions, those from beech leaves covered just the ITS1 region. Maximum likelihood analysis showed that these five sequences and the ITS sequences of all four *M. osmundae* strains form a strongly supported clade (Fig. 3).

Further comparison of the alignments suggests that the environmental sequences represent species that, in all likelihood, are congeneric and possibly conspecific with *M. osmundae*. There were small differences in the ITS regions of the four studied *M. osmundae* isolates, all resulting from single nucleotide polymorphic sites. The five sequences from GenBank were all obtained via cloning (L.G. Clark, pers. comm.; Cordier *et al.*, 2012) and thus were not expected to be heterozygous. All nine analyzed sequences shared 97–99% sequence identity (Fig. S6).

Discussion

Mixia osmundae has a small and very compact genome at 13.6 Mb. To date it is the most completely assembled as well as the smallest known genome in the Pucciniomycotina. The only

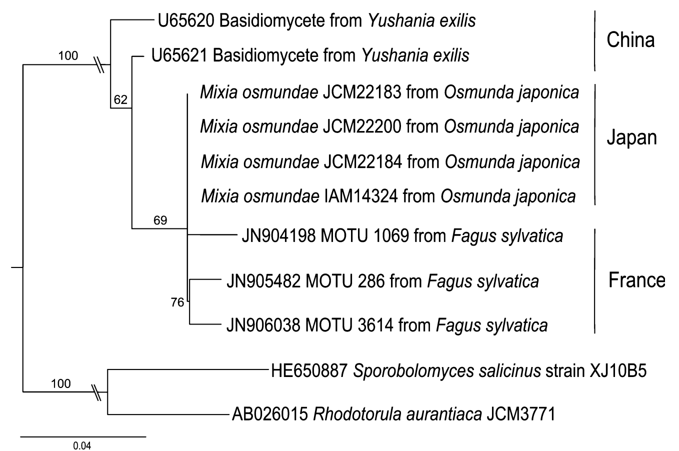

Fig. 3 The maximum likelihood tree of internal transcribed spacer (ITS) sequences of *Mixia osmundae* and five unidentified environmental sequences located by BLASTn searches. Geographic origin is indicated on the right. Numbers above branches indicate RAxML bootstrap support values. *Sporobolomyces salicinus* and *Rhodotorula aurantiaca* (Erythrobasidiomycetes) were used as outgroups.

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust

known Basidiomycota species with smaller genomes are the lipophilic yeast *M. globosa* (8.9 Mb; Xu *et al.*, 2007) and the xerotolerant mold *Wallemia sebi* (9.82 Mb; Padamsee *et al.*, 2012). It has been generally suggested that the genomes of filamentous plant pathogenic fungi are shaped by repeat-driven expansions and that pathogens with larger genomes, such as rust fungi and powdery mildews (at over 100 Mb), have an advantage in adapting in the course of the coevolution with plants (Schmidt & Panstruga, 2011; Kemen & Jones, 2012; Raffaele & Kamoun, 2012; Spanu, 2012). Nevertheless, large genomes are not necessary for pathogenesis, as there are also sequenced plant parasites with relatively small genomes. For example, smut fungi (Ustilaginomycetes, Basidiomycota) have compact genomes averaging at *c.* 20 Mb in size (Schirawski *et al.*, 2010). With the smallest known genome for a basidiomycete plant pathogen, *M. osmundae* could facilitate further studies in identifying genes necessary for a fungus to grow *in planta*. The only known fungal plant pathogens with smaller genomes are the ascomycetes *Eremothecium gossypii* (S.F. Ashby & W. Nowell) Kurtzman and *Taphrina deformans* (Berk.) Tul., which have genome sizes estimated at 9.2 and 13.3 Mb, respectively (Dietrich *et al.*, 2004; Cissé *et al.*, 2013).

The analysis of core genes determined that only 39.2% of all *M. osmundae* genes are shared with other sequenced Pucciniomycotina and Ustilaginomycotina species, indicating that reduction in the genome size is probably the result of the reduction in noncoding parts of the genome, not of loss of protein-encoding genes that are not part of the core proteome. Our analyses also indicate that while approx. half of the genes in the *S. roseus* genome are core genes, this is true for only 15–20% of the genes in rust fungal genomes (Fig. S3). The fact that there are almost no repetitive elements and the genome is very compact might hint at a conservative evolution of *Mixia* and may also explain why it is so rarely found. It is posited that an organism with a small and repeat-poor genome is not as flexible at adapting to new hosts as are organisms with genomes containing excessive repetitive elements (Raffaele & Kamoun, 2012).

Our data suggest that *M. osmundae* is a biotroph, but not an obligate one, as it has a yeast state that can be maintained on artificial media. As a plant-pathogenic fungus, *M. osmundae* is expected to have a set of genes that code for enzymes participating in plant cell wall degradation and the CAZy analysis detected several enzymes for breaking down cellulose as well as other cell wall components (Table S6). Furthermore, *M. osmundae* has a higher copy number of β-mannanases (GH26) compared with other sequenced Basidiomycota species. Recent studies of fern cell walls have determined a new mannan-rich cell wall type for this group (Silva *et al.*, 2011), and the high GH26 copy number in *M. osmundae* may have enhanced its capacity to infect ferns. In fact, the ability to break down cell wall components may be essential for the *in planta* growth of *M. osmundae*, as previous microscopy studies have shown that the fungus grows inside the outer epidermal plant cell wall (Mix, 1947; Kramer, 1958). Additionally, the CAZy analysis suggests that *M. osmundae* is able to use sucrose as a source of energy, which is common in biotrophs as sucrose is found in living plants (Solomon *et al.*, 2003). In general, when compared with the other Pucciniomycotina species, the *M. osmundae* CAZy profile shared more similarities with biotrophic rust fungi than with the saprobic Pucciniomycotina yeast *S. roseus* (Table S5).

When comparing the information obtained from the genome with available morphological and physiological data, there is clear concordance. For example, we did not detect any genes involved in pseudohyphal growth (GO:007124) and, accordingly, the yeast stage of *M. osmundae* is not known to form pseudohyphae. Nor did we detect genes for any enzymes involved in rhamnose (GH78), L-arabinose (GH3, 43, 51, 54), melibiose or raffinose (GH36) degradation, which is in agreement with assimilation studies that show that *M. osmundae* cannot utilize any of these carbon constituents of cellulose, hemicellulose or pectin (Nishida *et al.*, 2011b).

An analysis of secreted proteins was conducted to detect potential effector genes in *M. osmundae*. In total, over 235 putative SSPs were unique to *M. osmundae* and only 49 were conserved in other examined genomes of Pucciniomycotina and Ustilaginomycotina. This large number of unique SSPs shows that *M. osmundae* has a very specific set of potential effectors, and these may be involved in the interaction with its fern host. On the other hand, the set of conserved SSPs could represent a core set of genes needed for plant infection (Table S9).

The CAFE analysis of gene family gain and loss suggests relative genomic stability in *M. osmundae* compared with its ancestors, in contrast to the drastic changes seen in the pathogenic rust fungi (Duplessis *et al.*, 2011). The comparison of gene family gains and losses from the common ancestor of the Pucciniomycotina with those of *M. osmundae* revealed an increase in the copy number of cytochrome P450 throughout the evolution of *Mixia*. A high copy number of cytochrome P450 is characteristic of necrotrophic fungi (Islam *et al.*, 2012); however, despite the constant gain of that gene family in *M. osmundae* and its ancestors, the total copy numbers of the genes in the family still remain lower than what is known for necrotrophic fungi or most Basidiomycota species in general (Olson *et al.*, 2012).

The mating gene loci in *M. osmundae* show similarity to the presumably homologous loci in other Basidiomycota based on neighboring genes (i.e. synteny; Fig. 4). In particular, the pheromone precursor genes are found in proximity to the STE20 gene as observed in other Pucciniomycotina (Coelho *et al.*, 2008, 2011). This association of pheromone genes with the STE20 p21-activated kinase is consistent with other Pucciniomycotina, but is different from that observed in Ustilaginomycotina and Agaricomycotina, in which the pheromones are only found adjacent to the pheromone receptors (Riquelme *et al.*, 2005; Gioti *et al.*, 2013). Other genes found in conserved gene order near STE3 among *M. osmundae*, *R. graminis*, and *S. roseus* are ribosomal proteins (RPS19, RibL18ae), ribonucleoprotein splicing factor (LSm7), and karyopherin (importin) beta (KAP95). These data suggest that the pheromone encoding MAT locus is large and includes multiple proteins, because recombination between STE3 and the pheromone genes of different mating types would create a potentially self-compatible combination of receptor and pheromone.

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust

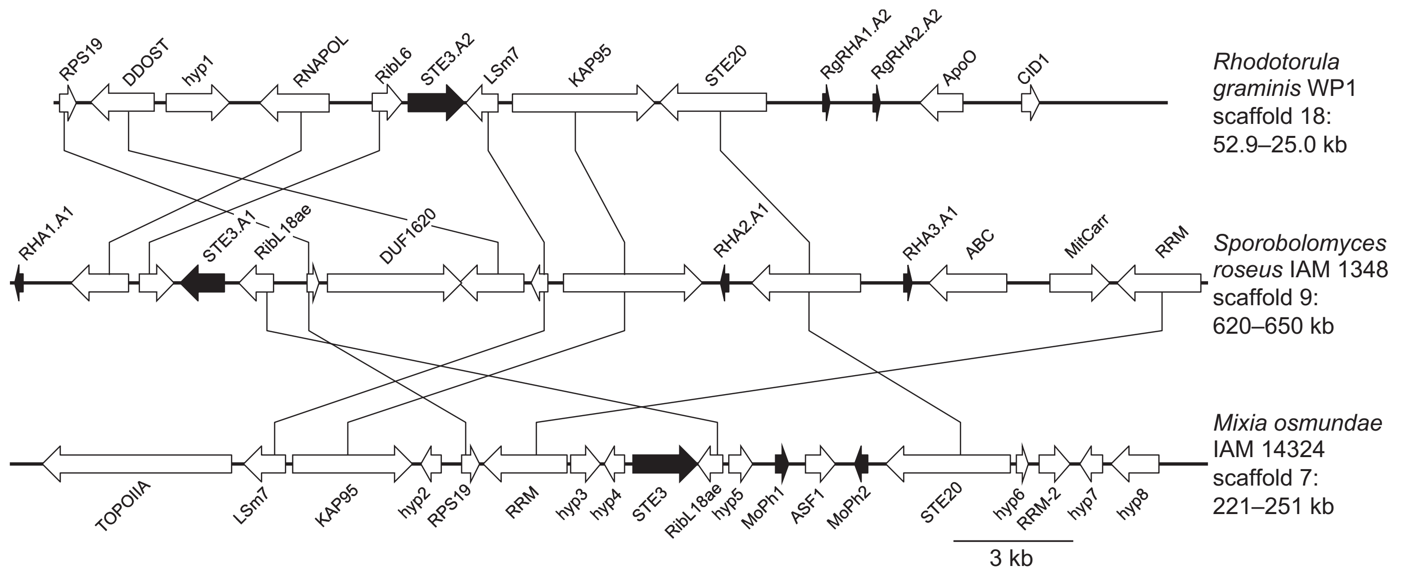

Fig. 4 Genes neighboring the pheromone receptor STE3 show synteny among *Mixia osmundae* and two representative Microbotryomycetes. Arrows indicate the direction of transcription of genes. Arrows in black indicate putative mating type genes (pheromones and pheromone receptors). Gene names are based on those in Coelho *et al.* (2011). hyp1–8 are hypothetical predicted proteins of uncertain function. Homologous genes are connected by vertical lines.

As the two MAT loci (SXI2 and STE3) are located on different scaffolds and are distant from each other, it could be that they are not linked and *M. osmundae* is tetrapolar. Alternatively, the fungus could be bipolar with an HD region as the sole determinant of mating type, as observed in multiple bipolar Agaricomycetes (James *et al.*, 2006). Additional possibilities are that the fungus is completely asexual or homothallic with no polymorphism of MAT genes, or finally that the two MAT regions may be physically linked, but only incompletely, as in the pseudobipolar mating system of Microbotryomycetes yeasts in Pucciniomycotina (Coelho *et al.*, 2010) and *Malassezia* in Ustilaginomycotina (Gioti *et al.*, 2013). If the latter is the case, then the two regions are minimally 1.4 Mb apart, which, though distant, is nonetheless similar to the minimum distance between the two loci in the pseudobipolar *Sporidiobolus salmonicolor* Fell & Tallman (Coelho *et al.*, 2010). The genomic arrangement of MAT homologs thus does not allow us to determine whether *M. osmundae* is bipolar, pseudobipolar or tetrapolar.

The lack of polymorphism between the mating regions of the four studied *M. osmundae* strains has two possible explanations. First, the spores from which the isolated strains originate are mitotic, and as the samples were collected from two locations and 2 yr, the populations of *M. osmundae* sampled in Japan could be clonal. Secondly, *M. osmundae* potentially reproduces sexually, but the spores sampled were without variation at the mating loci because it is homothallic or because only one mating type was isolated by chance. By studying two mating genes in spatially separated and presumably unlinked loci, the probability that these isolates were identical only by chance is low. The probability that the spores are sexually produced from a heterozygous dikaryon and that MAT alleles at SXI2 and STE3 are identical among four meiotic progeny is 1/64, that is, the odds of sampling the same genotype at two segregating loci four times. This calculation assumes that *M. osmundae* is tetrapolar and that both SXI and STE3 are MAT loci whose polymorphism is required for dikaryon formation. If *M. osmundae* was actually pseudobipolar or bipolar, this figure would reduce to 1/8. Nonetheless, 68% of all studied meiosis-related genes and over 93% of all core meiosis genes were found in the *M. osmundae* genome, which, in addition to the finding of a complete set of mating genes, makes it highly probable that *M. osmundae* is competent to undergo sexual reproduction. Although several of the genes that are considered to be essential in the meiosis pathway of *S. cerevisiae* and other eukaryotes (e.g. Ndt80, Ime1, Sum1, Xrs2) were not found in *M. osmundae* (Table S11; Fig. S7), they also were not detected within other Pucciniomycotina and Ustilaginomycotina species that we searched. Moreover, Halary *et al.* (2011) also failed to detect Hop2 and Mnd1 in some Ascomycota and Basidiomycota species that are nevertheless able to reproduce sexually. Therefore, it is probable that several fungi (especially Basidiomycota species) have modified gene sets (compared with *S. cerevisiae*) for regulating meiosis. While there is information available about the meiosis of some basidiomycetes (Burns *et al.*, 2010), the information on Pucciniomycotina meiosis genes is scarce and future studies may reveal new meiosis-specific genes for this subphylum.

As discussed earlier, the chances that the isolates studied in this analysis were produced via sexual reproduction are very small and we infer that the sequenced strain is haploid because of the lack of heterozygosity in the mating loci and the absence of duplicated gene blocks. This would mean that the much-discussed sporogenous structure does not represent a new basidium type. Nevertheless, the sporogenous cell of *M. osmundae* remains unique among Basidiomycota both for its morphology and its mode of spore production. Adding to the uniqueness of the sporogenous cells of *M. osmundae* is the fact that no other species of Pucciniomycotina is known to produce massive quantities of haploid asexual spores

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust

exogenously from a single cell, and the production of mitotic haploid spores is rare overall in this subphylum. Rust fungi produce quantities of haploid asexual spores within spermagonia, but spermagonia may also function as receptive structures for dikaryotization (Swann *et al.*, 2001; Cummins & Hiratsuka, 2003), which is most probably not true for the sporogenous cell of *M. osmundae*. Species of Tritirachiomycetes may also produce haploid conidia, but these differ from *M. osmundae* in that they are produced on hyphal conidiophores (Schell *et al.*, 2011).

Our sequence analysis identified five sequences generated in two independent environmental studies (Zhang *et al.*, 1997; Cordier *et al.*, 2012) that share close identity to *M. osmundae* and most probably represent congeners. Considering the great number of various Pucciniomycotina species that have been detected through high-throughput environmental sequencing projects (e.g. see the discussion in Aime *et al.*, 2014), five sequences in the last 20 yr is not a large number. However, the fact that *M. osmundae* is known from Japan, Taiwan, and the US, and indirectly via environmental sequences from France and China (Mix, 1947; Nishida *et al.*, 1995; Zhang *et al.*, 1997; Sugiyama & Katumoto, 2008; Cordier *et al.*, 2012), shows a wide but infrequently encountered distribution. In both of the studies where *Mixia* sequences were produced, analyzed leaves were picked from the plants and no surface sterilization was performed before DNA extraction and amplification (Zhang *et al.*, 1997; Cordier *et al.*, 2012). Therefore, it is unknown if these isolates were associated with the host plants as endophytes or as epiphytes or simply as inocula from an *Osmunda* fern nearby.

Conclusions

Mixia osmundae was first described 100 yr ago; however, despite efforts from many research groups, major questions about its biology remain. Our genomic analyses have revealed that *M. osmundae* has the smallest genome size to date for a basidiomycete plant pathogen. It is most likely a biotrophic fungus, but lacks significant gene family gains that have been associated with other successful phytopathogens. The majority of the 530 SSPs produced by *M. osmundae* are specific to it and may be involved in the infection of *Osmunda* ferns, whereas 40 SSPs were found to be shared with other Pucciniomycotina species, hinting at a core set of proteins necessary for general infection of plants by basidiomycete phytopathogens. We have determined that the spores produced on fern fronds are haploid and most probably produced via mitosis. Therefore, the sporogenous cells appear to be products of a unique type of asexual reproduction and not a novel basidium type. We also found a near-complete set of *MAT* and core meiosis genes, indicating that *M. osmundae* should be competent to reproduce sexually, although no evidence exists as to what that sexual state may look like. The lack of some meiosis genes (Ndt80 and Ime1) in *M. osmundae* or other analyzed Pucciniomycotina that are essential for meiosis in *S. cerevisiae* probably indicates the evolution of an altered meiosis gene set within this group. Finally, this study found evidence of other species or isolates of *Mixia* known only from environmental sequences of plant leaves, emphasizing the ongoing importance of collecting living isolates in addition to sequence data from ecological and environmental studies.

Acknowledgements

Strains JCM 22200, 22183 and 22184 were provided by the Japan Collection of Microorganisms, RIKEN BRC which is participating in the National BioResource Project of the MEXT, Japan. Dr Lynn G. Clark from Iowa State University is thanked for providing additional information about the environmental sequences obtained from bamboo, and Dr Mahajabeen Padamsee assisted with sample preparation. We are thankful to Dr Michael Perlin and Dr Kenneth Wolfe for granting permission to use the unpublished genome data of *M. violaceum* and *S. roseus*, respectively. We also thank Dr Louis Corrochano for allowing us to use the unpublished genome of *P. blakesleeanus* in phylogenetic analyses. Dr Steve Goodwin is acknowledged for discussions and comments on an earlier version of the manuscript. We thank Kanehisa Labs for allowing the use of the meiosis pathway map. This study was supported by the National Science Foundation Assembling the Fungal Tree of Life project, NSF DEB-0732968. The work conducted by the US Department of Energy Joint Genome Institute is supported by the Office of Science of the US Department of Energy under contract no. DE-AC02-05CH11231.

References

Aime MC, Matheny PB, Henk DA, Frieders EM, Nilsson RH, Piepenbring M, McLaughlin DJ, Szabo LJ, Begerow D, Sampaio JP *et al.* 2006. An overview of the higher-level classification of Pucciniomycotina based on combined analyses of nuclear large and small subunit rDNA sequences. *Mycologia* **98**: 895–905.

Aime MC, Toome M, McLaughlin DJ. 2014. Pucciniomycotina. In: McLaughlin DJ, Spatafora JW, eds. *The Mycota, systematics and evolution, 2nd edn, vol. 7A.* Berlin, Germany: Springer-Verlag, in press.

Bauer R, Begerow D, Sampaio JP, Weiss M, Oberwinkler F. 2006. The simple-septate basidiomycetes: a synopsis. *Mycological Progress* **5**: 41–66.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* **340**: 783–795.

Burns C, Stajich JE, Rechtsteiner A, Casselton L, Hanlon SE, Wilke SK, Savytskyy OP, Gathman AC, Lilly WW, Lieb JD *et al.* 2010. Analysis of the basidiomycete *Coprinopsis cinerea* reveals conservation of the core meiotic expression program over half a billion years of evolution. *PLoS Genetics* **6**: e1001135.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research* **37**: D233–D238.

Cantu D, Govindarajulu M, Kozik A, Wang M, Chen X, Kojima KK, Jurka J, Michelmore RW, Dubcovsky J. 2011. Next generation sequencing provides rapid access to the genome of *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat stripe rust. *PLoS ONE* **6**: e24230.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**: 540–552.

Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR *et al.* 2012. *Saccharomyces* Genome Database: the genomics resource of budding yeast. *Nucleic Acids Research* **40**: D700–D705.

Cissé OH, Almeida JMGCF, Fonseca A, Kumar AA, Salojärvi J, Overmyer K, Hauser PM, Pagni M. 2013. Genome sequencing of the plant pathogen Taphrina deformans, the causal agent of peach leaf curl. *mBio* **4**: e00055–13.

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust

Coelho MA, Gonçalves P, Sampaio JP. 2011. Evidence for maintenance of sex determinants but not of sexual stages in red yeasts, a group of early diverged basidiomycetes. *BMC Evolutionary Biology* **11**: 249.

Coelho MA, Rosa A, Rodrigues N, Fonseca Á, Gonçalves P. 2008. Identification of mating type genes in the bipolar basidiomycetous yeast *Rhodosporidium toruloides*: first insight into the MAT locus structure of the Sporidiobolales. *Eukaryotic Cell* **7**: 1053–1061.

Coelho MA, Sampaio JP, Gonçalves P. 2010. A deviation from the bipolar-tetrapolar mating paradigm in an early diverged basidiomycete. *PLoS Genetics* **6**: e1001052.

Cordier T, Robin C, Capdevielle X, Fabreguettes O, Desprez-Loustau M-L, Vacher C. 2012. The comparison of phyllosphere fungal assemblages of European beech (*Fagus sylvatica*) varies significantly along an elevation gradient. *New Phytologist* **196**: 510–519.

Cummins GB, Hiratsuka Y. 2003. *Illustrated genera of rust fungi, 3rd edn.* St Paul, MN, USA: APS Press, 14–29.

De Bie T, Cristianini N, Demuth JP, Hahn MW. 2006. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**: 1269–1271.

Dietrich FS, Voegeli S, Brachat S, Lerch A, Gates K, Steiner S, Mohr C, Pöhlmann R, Luedi P, Choi S *et al.* 2004. The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science* **304**: 304–307.

Duplessis S, Cuomo CA, Lin YC, Aerts A, Tisserant E, Veneault-Fourrey C, Joly DL, Hacquard S, Amselem J, Cantarel BL *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.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792–1797.

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martinez AT, Otillar R, Spatafora JW, Yadav JS *et al.* 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* **336**: 1715–1719.

Gardes M, Bruns TD. 1993. ITS primers with enhanced specificity for basidiomycetes: application to the identification of mycorrhiza and rusts. *Molecular Ecology* **2**: 113–118.

Gioti A, Nystedt B, Li W, Xu J, Andersson A, Averette AF, Münch K, Wang X, Kappauf C, Kingsbury JM *et al.* 2013. Genomic insights into the atopic eczema-associated skin commensal yeast *Malassezia sympodialis*. *mBio* **4**: e00572–12.

Grigoriev IV, Cullen D, Goodwin SB, Hibbett D, Jeffries TW, Kubicek CP, Kuske C, Magnuson JK, Martin F, Spatafora JW *et al.* 2011. Fueling the future with fungal genomics. *Mycology* **2**: 192–209.

Grigoriev IV, Martines DA, Salamov AA. 2006. Fungal genomic annotation. In: Aurora DK, Berka RM, Singh GB, eds. *Applied mycology and biotechnology, (Bioinformatics), vol.* **6**. Amsterdam, the Netherlands: Elsevier Press, 123–142.

Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, Kuo A, Minovitsky S, Nikitin R, Ohm RA *et al.* 2012. The genome portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Research* **40**: D26–D32.

Halary S, Malik S-B, Lildhar L, Slamovits CH, Hijri M, Corradi N. 2011. Conserved meiotic machinery in *Glomus* spp., a putatively ancient asexual fungal lineage. *Genome Biology and Evolution* **3**: 950–958.

Islam MS, Haque MS, Islam MM, Emdad EM, Halim A, Hossen QMM, Hossain MZ, Ahmed B, Rahim S, Rahman MS. 2012. Tools to kill: genome of one of the most destructive pathogenic fungi *Macrophomina phaseolina*. *BMC Genomics* **12**: 493.

James TY, Srivilai P, Kües U, Vilgalys R. 2006. Evolution of the bipolar mating system of the mushroom *Coprinellus disseminatus* from its tetrapolar ancestors involves loss of mating-type-specific pheromone receptor function. *Genetics* **172**: 1877–1891.

Kemen E, Jones JDG. 2012. Obligate biotroph parasitism: can we link genomes to lifestyles? *Trends in Plant Science* **17**: 448–457.

Kramer CL. 1958. A new genus in the Protomycetaceae. *Mycologia* **50**: 916–926.

Kramer CL. 1987. The Taphrinales. *Studies in Mycology* **30**: 151–166.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* **305**: 567–580.

Kumar S, Kushwaha H, Bachhawat AK, Raghava GPS, Ganesan K. 2012. Genome sequence of the oleaginous red yeast *Rhodosporidium toruloides* MTCC457. *Eukaryotic Cell* **11**: 1083–1084.

Kumar S, Nei M, Dudley J, Tamura K. 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics* **9**: 299–306.

Kües U, James TY, Heitman J. 2011. Mating type in Basidiomycetes: unipolar, bipolar, and tetrapolar patterns of sexuality. In: Pöggeler S, Wöstemeyer J, eds. *The Mycota, evolution of fungi and fungal-like organisms, vol. 14.* Berlin, Germany: Springer-Verlag, 97–160.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Gateway Computing Environments Workshop New Orleans, LA*, 1–8.

Mix AJ. 1947. *Taphrina osmundae* Nishida and *Taphrina higginsii* sp. nov. *Mycologia* **39**: 71–76.

Morrow CA, Fraser JA. 2009. Sexual reproduction and dimorphism in the pathogenic basidiomycetes. *FEMS Yeast Research* **9**: 161–177.

Nishida H, Ando K, Ando Y, Hirata A, Sugiyama J. 1995. *Mixia osmundae*: transfer from the Ascomycota to the Basidiomycota based on evidence from molecules and morphology. *Canadian Journal of Botany* **73**(Suppl. 1): S660–S666.

Nishida H, Nagatsuka Y, Sugiyama J. 2011a. Draft genome sequencing of the enigmatic basidiomycete *Mixia osmundae*. *Journal of General and Applied Microbiology* **57**: 63–67.

Nishida H, Robert V, Sugiyama J. 2011b. *Mixia* C.L. Kramer emend. H. Nishida, K. Ando, Y. Ando, Hirata & Sugiyama (1995). In: Kurtzman CP, Fell JW, Boekhout T, eds. *The Yeasts, a taxonomic study, vol. 3.* London, UK: Elsevier, 1499–1502.

Nishida T. 1911. A contribution to the monograph of the parasitic Exoascaceae of Japan. In: Takahashi Y, ed. *Miyabe-Festschrift.* Rokumeikwan: Tokyo, Japan, 157–212.

Olson Å, Aerts A, Asiegbu F, Belbahri L, Bouzid O, Broberg A, Canbäck B, Coutinho PM, Cullen D, Dalman K *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.

Padamsee M, Kumar TKA, Riley R, Binder M, Boyd A, Calvo AM, Furukawa K, Hesse C, Hohmann S, James TY *et al.* 2012. The genome of the xerotolerant mold *Wallemia sebi* reveals adaptations to osmotic stress and suggests cryptic sexual reproduction. *Fungal Genetics and Biology* **49**: 217–226.

Parra G, Bradnam K, Korf I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **23**: 1061–1067.

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J *et al.* 2012. The Pfam protein families database. *Nucleic Acids Research* **40**: D290–D301.

Raffaele S, Kamoun S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. *Nature Reviews* **10**: 417–430.

Riquelme M, Challen MP, Casselton LA, Brown AJ. 2005. The origin of multiple B mating specificities in *Coprinus cinereus*. *Genetics* **170**: 1105–1119.

Robbertse B, Yoder RJ, Boyd A, Reeves J, Spatafora JW. 2011. Hal: an automated pipeline for phylogenetic analyses of genomic data. *PLoS Currents* **3**: RRN1213.

Savile DBO. 1955. A phylogeny of the Basidiomycetes. *Canadian Journal of Botany* **33**: 60–104.

Schell WA, Lee AG, Aime MC. 2011. A new lineage in Pucciniomycotina: class Tritirachiomycetes, order Tritirachiales, family Tritirachiaceae. *Mycologia* **103**: 1331–1340.

Schirawski J, Mannhaupt G, Münch K, Brefort T, Schipper K, Doehlemann G, Di Stasio M, Rössel N, Mendoza-Mendoza A, Pester D *et al.* 2010. Pathogenicity determinants in smut fungi revealed by genome comparison. *Science* **330**: 1546–1548.

Schmidt SM, Panstruga R. 2011. Pathogenomics of fungal plant parasites: what have we learnt about pathogenesis? *Current Opinion in Plant Biology* **14**: 392–399.

Silva GB, Ionashiro M, Carrara TB, Crivellari AC, Tiné MAS, Prado J, Carpita NC, Buckeridge MS. 2011. Cell wall polysaccharides from fern leaves:

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust

evidence for a mannan-rich type III cell wall in *Adiantum raddianum*. *Phytochemistry* 71: 2352–2360.

Sjamsuridzal W, Nishida H, Yokota A. 2002. Phylogenetic position of *Mixia osmundae* inferred from 28S rDNA comparison. *Journal of General and Applied Microbiology* 48: 121–123.

Solomon PS, Tan K-C, Oliver RP. 2003. The nutrient supply of pathogenic fungi; a fertile field for study. *Molecular Plant Pathology* 4: 203–210.

Spanu PD. 2012. The genomics of obligate (and nonobligate) biotrophs. *Annual Review of Phytopathology* **50**: 91–109.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690.

Stamatakis A, Hoover P, Rougemont J. 2008. A fast bootstrapping algorithm for the RAxML web-servers. *Systematic Biology* 57: 758–771.

Sugiyama J. 1998. Relatedness, phylogeny, and evolution of the fungi. *Mycoscience* 39: 487–511.

Sugiyama J, Katumoto K. 2008. Identity of the plasmodial slime mold *Phytoceratiomyxa osmundae* and the lectotypification of *Taphrina osmundae*, the basionym of *Mixia osmundae*. *Mycoscience* 49: 192–198.

Swann EC, Frieders EM, McLaughlin DJ. 2001. Urediniomycetes. In: McLaughlin DJ, McLaughlin EG, Lemke PA, eds. *The Mycota, systematics and evolution, vol. 7B*. Berlin, Germany: Springer-Verlag, 37–56.

Toome M, Roberson RW, Aime MC. 2013. *Meredithblackwellia eburnea* sp. et gen. nov., Kriegeriaceae fam. nov. and Kriegeriales ord. nov. – toward resolving higher-level classification in Microbotryomycetes. *Mycologia* **105**: 486–495.

Van den Brink J, de Vries RP. 2011. Fungal enzyme sets for plant polysaccharide degradation. *Applied Microbiology and Biotechnology* **91**: 1477–1492.

Vershon AK, Pierce M. 2000. Transcriptional regulation of meiosis in yeast. *Current Opinion in Cell Biology* **12**: 334–339.

White TJ, Bruns TD, Lee S, Taylor JW. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds. *PCR protocols: a guide to methods and applications*. San Diego, CA, USA: Academic Press, 315–322.

Xu J, Saunders CW, Hu P, Grant RA, Boekhout T, Kuramae EE, Kronstad JW, DeAngelis YM, Reeder NL, Johnstone KR *et al.* 2007. Dandruff-associated *Malassezia* genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens. *Proceedings of the National Academy of Sciences, USA* **104**: 18730–18735.

Zhang W, Wendel JF, Clark LG. 1997. Bamboozled again! Inadvertent isolation of fungal rDNA sequences from bamboos (Poaceae: Bambusoideae). *Molecular Phylogenetics and Evolution* **8**: 205–217.

Supporting Information

Additional supporting information may be found in the online version of this article.

Methods S1 Additional information about genome sequencing and phylogenetic, CAFE and mating gene analysis.

Fig. S1 A comparison of KEGG metabolic pathway gene models in Pucciniomycotina.

Fig. S2 A comparison of gene models in Pucciniomycotina based on KOG classification.

Fig. S3 Core and noncore proteome of the published Pucciniomycotina and Ustilaginomycotina genomes.

Fig. S4 Conservation of *Mixia osmundae* predicted proteins in related fungi.

Fig. S5 *Mixia osmundae* pheromone precursors.

Fig. S6 Comparison of the ITS region of *Mixia osmundae* isolates and closest GenBank matches.

Fig. S7 Meiosis pathway in *Saccharomyces cerevisiae* and orthologous genes in *Mixia osmundae*.

Table S1 Fungal genomes used for phylogenetic studies.

Table S2 A list of avirulence genes used for AVR gene analysis in *M. osmundae*

Table S3 Summary statistics for published genomes in Pucciniomycotina

Table S4 Enrichment of functional annotation terms in the core proteome

Table S5 Comparative analysis of the number of putative genes encoding for carbohydrate-active enzymes

Table S6 Plant biomass degradation-related carbohydrate-active enzyme families present in *Mixia osmundae*

Table S7 Gene family evolution data from CAFE analysis presenting all statistically significant gains and losses

Table S8 Comparison of SSPs, secreted proteins and all proteins in Pucciniomycotina and Ustilaginomycotina

Table S9 The list of predicted small secreted proteins of *M. osmundae* and comparative genomes

Table S10 The name and location of the mating genes determined in the genome of *Mixia osmundae*

Table S11 The presence of meiosis-related genes in *Mixia osmundae*

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

No claim to original US Government works
New Phytologist © 2013 New Phytologist Trust