The genomic landscape of tree rot in *Phellinus noxius* and its Hymenochaetales members

Chia-Lin Chung*+[1,2], Jiaye T Lee[3,4,5], Mitsuteru Akiba[6], Hsin-Han Lee[1], Tzu-Hao Kuo[3], Dang Liu[3,7], Huei-Mien Ke[3], Toshiro Yokoi[6], Marylette B Roa[3,8], Meiyeh J Lu[3], Ya-Yun Chang[1], Pao-Jen Ann[9], Jyh-Nong Tsai[9], Chien-Yu Chen[10], Shean-Shong Tzean[1], Yuko Ota[6,11], Tsutomu Hattori[6], Norio Sahashi[6], Ruey-Fen Liou[1,2], Taisei Kikuchi*[12] and Isheng J Tsai*+[3,4,5,7]

+ Contributed equally

* Corresponding authors: clchung@ntu.edu.tw, taisei_kikuchi@med.miyazaki-u.ac.jp and ijtsai@gate.sinica.edu.tw

[1]Department of Plant Pathology and Microbiology, National Taiwan University, Taiwan

[2]Master Program for Plant Medicine, National Taiwan University, Taiwan

[3]Biodiversity Research Center, Academia Sinica, Taipei, Taiwan

[4]Biodiversity Program, Taiwan International Graduate Program, Academia Sinica and National Taiwan Normal University

[5]Department of Life Science, National Taiwan Normal University

[6]Department of Forest Microbiology, Forestry and Forest Products Research Institute, Tsukuba, Japan

[7]Genome and Systems Biology Degree Program, National Taiwan University and Academia Sinica, Taipei, Taiwan

[8]Philippine Genome Center, University of the Philippines, Diliman, Quezon City, Philippines 1101

[9]Plant Pathology Division, Taiwan Agricultural Research Institute

[10]Department of Bio-industrial Mechatronics Engineering, National Taiwan University, Taiwan

[11]College of Bioresource Sciences, Nihon University, Fujisawa, Japan

[12]Division of Parasitology, Faculty of Medicine, University of Miyazaki, Miyazaki, Japan

ABSTRACT

Background

The order Hymenochaetales of filamentous fungi contains some of the most aggressive wood decayers that have caused major forest incidents as early as 1971. It is also phylogenetically placed between the well-studied Agaricales and Ustillaginales orders, making Hymenochaetales species interesting for investigating the evolutionary transition between the orders of Basidiomycota.

Results

We report four whole-genomes from this group including the tree rot pathogen *Phellinus noxius*, the wood decomposing fungus *Phellinus lamaensis*, the laminated root rot pathogen *Phellinus sulphurascens*, and the trunk rot pathogen *Porodaedalea pini*. While expansion of transposable elements constitutes the main difference between these genomes ranging from 31- to 53-megabases, their karyotype is conserved with distant fungi. We found many genes reflecting their ability as white-rot fungi, and transcriptome data from *in planta* colonization of *P. noxius* revealed members of carbohydrate active enzymes, P450s, and nucleotide-binding-oligomerization-domain like receptors that are upregulated at different infection phases. Population genomics analyses of 60 *P. noxius* isolates across the Asia Pacific region revealed extraordinary genetic diversity and mono-/heterokaryotic nature, and that the isolates can be divided into two genetically separated lineages with contrasting evolutionary scenarios.

Conclusions

These genomes shed light on the basic biology and population genetics of fungal tree pathogens. The combined genome and transcriptome analysis of *P. noxius* can serve as a starting point for understanding the survival strategies of this devastating pathogen, from which new control methods could be developed.

Keywords

Phellinus – Comparative genomics – Population genomics – Fungal biology – Tree pathogens

INTRODUCTION

Under most circumstances, fungi coexist with trees or act as saprotrophs responsible for carbon and nitrogen cycling in forest systems. However, some fungi are also one of the dominant groups of pathogens causing diseases in trees. There has been an emergence of tree disease outbreaks in different parts of the world such as the ash dieback[1], Dutch elm disease[2], laminated root rot caused by *Phellinus sulphurascens*[3], and brown root rot caused by *Phellinus noxius*[4-6]. Factors contributing to this phenomenon include climate change[7] and human activity[8]. If interventions are not implemented early and effective, pathogen infections can kill millions of trees and the spread can become very difficult to stop[9].

Hymenochaetales is dominated by wood decay fungi and belongs to Agaricomycetes of Basidiomycota. Most of which are saprotrophic but some also exhibit pathogenic lifestyles that have been recorded in major forest incidents as early as 1971 in different parts of the world[10, 11]. In particular, *P. noxius* has a very wide host range, spanning more than 200 broadleaved and coniferous tree species (at least 59 families) including many agricultural, ornamental, and forest trees such as longan, litchi, camphor, banyan, and pine[12, 13]. Inoculation assays showed that only seven out of 101 tested tree cultivars (92 species) exhibited high resistance[14]. Moreover, despite the differences in virulence, *P. noxius* isolates exhibited no host specificity[14-16].

The life cycle of *P. noxius*[13] is thought to be similar to other important root-rotting basidiomycetes, such as *Armillaria mellea* and *Heterobasidion annosum*. The new infection can start from previously infected plants/stumps or colonized wood debris, from which the mycelium of *P. noxius* grows to infect the lateral and tap roots of the host tree[13]. An invasion to the cortex and lignified xylem is usually observed[17], sometimes accompanied by gradual expansion of the mycelial mat to the basal stem (**Figure 1a**). Diseased trees with decayed roots may then show symptoms of foliar chlorosis, thinning leaves, defoliation, and eventually decline within a few months to several years. The damaged and fragile roots (**Figure 1b**) also make the trees easily toppled over by strong winds and heavy rains. Basidiocarps are occasionally formed on trunks of infected trees (**Figure 1c-d**). The sexual reproduction system of *P. noxius* has remained unclarified, partly due to the lack of clamp connections for diagnosing compatibility[13].

Brown root rot disease caused by *P. noxius* is widespread in tropical and subtropical areas in Southeast and East Asia, Oceania, Africa, Central America and the Caribbean. The geographical distribution appears to be related to the range of growth temperature range of *P. noxius*: it grows over 10-12^{o}C to 36 ^{o}C, with optimum growth at 30 ^{o}C. In the past 20 years, brown root rot disease has become a serious threat to a variety of perennial fruit trees, ornamental and landscape trees, and shade trees in Taiwan[6] and in Ryukyu and Ogasawara Islands of Japan[4]. In Australia, it occurs in the natural and commercial plantation forests and orchards along the east coast, and has killed many trees in the Greater Metropolitan area of Brisbane Queensland[18]. In West Africa and China, brown root rot was reported as the most devastating root disease attacking the rubber plantations[16]. Trees in urban areas and plantation forestry in Hong Kong and Macao have also been seriously affected[19, 20]. Recent population genetics studies based on simple sequence repeat (SSR) markers suggested that *P. noxius* may spread over short distances via root-to-root contact of the hosts, and the genetically variable basidiospores are likely involved in long-distance dispersal and the establishment of unique clones in new disease foci[4, 6].

Here we report the genome sequences of four species from Hymenochaetales: white-rot fungus *Phellinus noxius*, *Phellinus lamaensis* (wood decomposing fungus that causes a white pocket rot only on dead heartwood trees), *Phellinus sulphurascens* (syn. *Coniferiporia sulphurascens*[21]; pathogen responsible for laminated root rot in Douglas-fir/ true fir), and *Porodaedalea pini* (syn. *Phellinus pini;* trunk pathogen of conifers). To investigate the genomic basis of tree pathogenicity in these fungi, we compared: i) these genomes against other representative species of Basidiomycota; ii) the transcriptomes of basidiocarps and mycelial mat from infected live trees and in culture; and iii) whole genome sequencing of 60 *P. noxius* isolates from regions endemic to brown root rot disease. We identified several features that are conserved throughout basidiomycetes in the genome of *P. noxius* as well as genes upregulated during infection. Population genomics analyses of *P. noxius* isolates revealed its bipolar reproductive mode and two major genetic lineages with different evolutionary scenarios. Together, this study provides insights into an important tree filamentous pathogen by applying large-scale comparative and population genomics approaches.

RESULTS

Genome structure and content

We produced a 31.6 Mb *P. noxius* genome reference assembly from a Japanese KPN91 isolate combining both Pacbio and Illumina technologies. The nuclear genome of *P. noxius* is assembled into 12 scaffolds with six assembled into chromosomes from telomere to telomere **(Figure 2)**, while the mitochondrial genome is assembled into a single sequence of 163.4 kb. To date chromosome-level assemblies are available only for a few basidiomycetous species[22-24] including *P. noxius* (Supplementary Table 1). For a comprehensive understanding of genome evolution amongst members of the hymenochaetoid clade, we also sequenced and assembled three additional species: *P. sulphurascens, P. lamaensis* and *P. pini*, as well as two more isolates (A42 and 718-S1) of *P. noxius*. The three assemblies of *P. noxius* have N50s of 2.4-2.7 Mb, whilst the other three genome assemblies comprise 30.7-53.3 Mb with N50 of 570 kb-2.7 Mb. Together, the genome assembly quality of these Hymenochaetaceae are among the most contiguous and complete of the sequenced basidiomycetes.

A total of 9,833-18,103 genes were predicted in each species, which are 82-94% complete as assessed with the BUSCO[25] Basidiomycota dataset (Supplementary Table 1). To compare these predicted proteins to those of other basidiomycetes to explore chromosome and gene family evolutionary dynamics, we selected the proteomes of fifteen species from the 1000 Fungal Genomes Project that are highly finished or for use as outgroups (Supplementary Table 1). The Hymenochaetales species have median intergenic and intron lengths of 507-634 bp and 59-60 bp, respectively, which are comparable with those observed in genomes of other basidiomycetes. The maximum likelihood phylogeny based on 1,127 single-copy orthologs placed these species with *Fomitiporia mediterranea* and *Schizopora paradoxa,* two other species of the Hymenochaetaceae group with strong bootstrap support (**Figure 1e**). This phylogenetic relationship is consistent with previous findings[26], and species with similar genome sizes and pathogenic habit are grouped together. For instance, *Phellinus noxius* and *P. lamaensis* with compact genome sizes of ~31 Mb are grouped together, while the trunk rot pathogens *P. pini* and *F.*

mediterranea show an expansion of their genome sizes to 53-63 Mb and are also grouped together.

Analysis of transposable elements and evidence of repeat-induced point mutation (RIP)

A combination of *de novo* and reference repeat detection classified a total of 53-297 repeat families (Supplementary Table 2). The transposable element (TE) content makes up 1.4-27.5% of the assemblies in the Hymenochaetales order (Supplementary Table 3) and is associated with their genome size and taxonomic relationship. The Class I transposons make up the majority of repetitive elements in these fungal genomes (Supplementary Table 3). Manual annotation of the 16 most abundant full-length transposon sequences revealed that these TEs belong to the *Gypsy* and *Copia* superfamilies of the LTR retrotransposons, and the *Tad1* clade from the LINE retrotransposons (Supplementary Figure 1). The *Gypsy* elements are 5,295 to 7,629 bp in length and contain an open reading frame encoding 1,523 amino acid residues, which contain conserved domains in the order of protease, reverse transcriptase, RNase H, integrase and chromodomain (Supplementary Figure 1). Chromodomain-containing *Gypsy* or chromoviruses are widely distributed in eukaryotes and present in many basidiomycetes[27]. In the case of *P. pini,* 588 full-length copies of two repeats from the *Gypsy* family (Supplementary Figure 1) make up 10% of the total assembly. In contrast, a clear exception is *P. sulphurascens* where a *Tad1* class LINE dominates the TE content of the genome and a *Copia* class of LTR-retrotransposon is found to be more abundant than *Gypsy*.

Copies of TEs are not distributed uniformly across the genome: the six complete chromosomes in the *P. noxius* assembly show clear signatures of Basiodiomycete centromeres (located at the middle of chromosome and comprised solely of TEs) and enriched at subtelomeres; TEs are distributed as distinct 80.4-118.4 kb blocks (**Figure 2**), which are typical in fungi[22]. These clusters also show a much lower average GC content suggesting a hallmark of fungal repeat-induced point mutation*[28]* (RIP) favouring T or A changes (Supplementary Figure 2). In Basidiomycota such mutational patterns were previously only observed in the Pucciniomycotina subphylum[29]. However, homologs of known genes responsible for RIP are present in all four Hymenochaetales species (Supplementary Table 4). Analysis of

dinucleotide transitions[30] has shown significantly higher CpN→TpN mutations, but not vice versa (Supplementary Figure 3), suggesting that RIP have successfully limited TE expansions especially in the compact genomes of *P. noxius* and *P. lamaensis*. Interestingly, each species has a different preferred mode of dinucleotide changes: CpA→TpA mutations clearly dominate in *P. noxius* compared to CpT→TpT and CpG→TpG in both *P. pini* and *P. sulphurascens.*

Chromosome biology

Like other complete chromosomes of basidiomycetes (Supplementary Table 1), the *P. noxius* genome also presents little variability in gene density along its scaffolds with the exception of TE clusters (Supplementary Figure 4). The number of known karyotypes in Basidiomycota ranges from 11 to 14 chromosome pairs (Supplementary Table 1), which suggests a possible ancestor with similar chromosome numbers. In an attempt to characterize chromosome architecture and evolution amongst the Agaricomycotina sub-division, we constructed a linkage network of seven selected species based on single-copy orthologs between species pairs. We identified 13 distinct linkage groups (LG) providing strong evidence that chromosome macro-synteny is largely conserved since the common ancestor of Agaricomycetes (**Figure 3a**; Supplementary Figure 5). Such a relationship is extended to Dacrymycetes where multiple scaffolds can be predominantly assigned to different linkage groups, but it is no longer apparent when compared to Tremellomycetes. Certain scaffolds are found to connect different linkage groups, implying inter-chromosomal rearrangements. For example, *P. noxius* scaffold1 is strongly clustered in LG11 but also shows linkage to LG10, implying a translocation from scaffold5 (**Figure 3**). Within each linkage group, gene collinearity is no longer apparent, suggesting high levels of intra-chromosomal rearrangements which have been observed in different fungal groups[31] (**Figure 3b**).

Like most species, the most common telomere repeating unit in basidiomycetes is TTAGGG[32], although other combinations such as T(n)AG(n)[33] or TTTAGGGG[34] have been described elsewhere. A systematic survey of the selected published genomes reveals the predominance of TTAGGG repeat across basidiomycetes, including *P. pini* and *P. sulphurascens* (Supplementary Table 5). The

karyotype of *P. noxius* has not been reported, but together with *P. lamaensis* we found 10 scaffolds that end with overrepresented 13.6-66.9 copies of a TTAGGTG heptamer, suggesting a novel telomeric repeat compared to other basidiomycetes[32]. Interestingly, the *Dacryopinax primogenitus* assembly contains 31 scaffold ends with TTAGGG telomeric repeats. Of those, 11 scaffolds have telomeric repeats at both ends revealing extensive fission of at least 21 chromosomes. **Figure 3b** shows one example of six scaffolds with such repeats present on the ends in *D. primogenitus* compared to two scaffolds in both *P. noxius* and *Coprinopsis cinerea.*

Gene families associated with Hymenochaetales and *Phellinus noxius*

We sought to identify genes and protein domains specific to Hymenochaetales by determining when a new gene family arose and if the family has expanded or contracted. In total, 7,125-11,659 proteins in the Hymenochaetales order are clustered together in 5,184 families. Acquisition of gene families was mainly found at the tips of the phylogeny (531-8,055 families) suggesting each species has a repertoire of specific genes. The seven Hymenochaetaceae species have a total of 62 enriched protein domains compared to other basidiomycetes (Supplementary Figure 6). Gain of domains such as fungal specific transcription factors (Fungal_trans; 53.8 *vs* 41.6 copies) and peroxidase associated protein domains (Peroxidase_ext; 16.7 *vs* 3.9 copies) in these Hymenochaetaceae species reflect their white rot nature. Conversely, 579 gene families were inferred to have been lost in the common ancestor of *P. lamaensis* and *P. noxius*. Analysis of gene ontology (GO) term enrichment on these otherwise preserved gene families in *P. sulphurascens* and *P. pini* revealed a common theme of intracellular signal transduction (Supplementary Table 6). Several signal transduction pathways may have become more specialized in the ancestor of *P. noxius* and *P. lamaensis*. For example, multiple copies of NmrA (nitrogen metabolite repression) were lost in these two species, which act as regulators in nitrogen-dependent control of secondary metabolism[35].

We sought to identify novel and expanded gene families in *P. noxius* despite its compact genome (Supplementary Figure 7). In total, 74% (7,313) of 9,833 predicted genes from *P. noxius* have orthologs from at least one of the other basidiomycetes, suggesting most of the basidiomycetous core genes are conserved. Consistent with the fact that *P. noxius* is a fast grower[5], genes encoding 1,3-beta-glucan synthase

responsible for the formation of beta-glucan components of the fungal cell wall[36] are expanded (14 compared to an average of 1.8 copies). Notably, counts of WD40 protein domains are highest in *P. noxius* and *P. lamaensis* despite their small genome size (Supplementary Table 7a). This domain is important in protein-protein interactions of cellular networks and is usually associated with additional domains with other functional activities[37]. Interrogating this expansion revealed the association of WD40 domains with the AAA and NACHT domains (Supplementary Table 7b), both of which are NTPase domains and such combinations are commonly found in nucleotide-binding-oligomerization-domain like receptors (NLRs). This gene family is highly diverse in fungi and was found to be central to the process of programmed cell death and implicated in fungal vegetative incompatibility and general nonself recognition [38]. A maximum likelihood phylogeny of the NACHT domain proteins shows different expansions of NLR subfamilies in fungi (Supplementary Figure 8). In particular, the C2-NACHT-WD40(n) subfamily has only been found exclusively in a few basidiomycetes[39] and is present in the highest copy number in *P. noxius* and *P. lamaensis*. The strikingly high abundances of WD40 domains in *P. noxius* and *P. lamaensis* are a result of tandem expansions near chromosome ends (**Figure 2**). Other expansions include UDP-glucuronosyltransferases, which catalyse conjugation reactions in the biotransformation of xenobiotics released by its host environment [40].

P. noxius **displays a bipolar heterothallic mating system**

Mate recognition of sexual reproduction in Basidiomycota is known to be controlled by two unlinked loci, named as *A* and *B* locus. Different from the divergently transcribed *HD1-HD2* observed in most species in Agaricomycetes[41], a conserved head-to-tail orientation of *HD1-HD2* in *HD* pair 1 was found in Hymenochaetales **(Figure 4)**. Alignment of 100-kb sequences upstream and downstream of *A* locus in *P. noxius* isolates KPN91, 718-S1 and A42 revealed that *A* locus is highly polymorphic (*HD* pairs in particular) despite well-conserved flanking regions (Supplementary Figure 9). For *B* locus, only one *STE3* encoding seven transmembrane helices was identified; four pheromone genes were identified but not physically linked to the highly monomorphic *STE3* (Supplementary Figure 10, Supplementary Table 8, Supplementary Information), which is characteristic of a bipolar mating system.

To further confirm this observation, allele diversity was analyzed by resequencing *A* and *B* locus from 10 single-basidiospore isolates originating from a single basidiocarp (Supplementary Table 9). Sequencing of *STE3* revealed two highly similar alleles (b1, b2), with 99.5% amino acid identity. The only variant (244V/A) is considered a semi-conservative mutation (valine and alanine are nonpolar aliphatic acids) and may have minor or no effect on protein function. Previous studies have shown that although STE3 is not involved in mating type determination in bipolar fungi, variations can still be observed[42-44]. Primer walking of *A* locus revealed two distinct haplotypes (Supplementary Figure 11): the a1 and a2 alleles of *HD* pair I shared an overall 78% nucleotide identity; the two alleles of *HD* pair II were highly diverse and the HD1 domain in a2 allele contains a 1-bp and a 9-bp deletions and has become a pseudogene. The loss of HD1 domain was also found in the *HD* pair I of 718-S1 (Figure 4), suggesting that at least one of the HD1 motifs in *A* locus is dispensable for a functional HD1-HD2 heterodimer in *P. noxius*. The presence or absence of specific HD domains reflects phylogenetic characteristics and has been commonly observed in fungi[45]. Clear segregation of two distinct alleles at *A* locus (Supplementary Figure 11) suggests a heterothallic bipolar reproductive mode in *P. noxius*.

Distinct from all the other fungi, *P. noxius* has an exceptionally highly expanded *A* locus across a ~60-kb region **(Figure 4)**. There are two pairs of HD1-HD2 gene located ~50-kb away from each other. The characteristic large interval between *HD* pairs has only been reported in *Schizophyllum commune* (~55 kb) and *Flammulina velutipes* (~70 kb), in which the genomic separation likely emerged through inversions or transpositions of gene clusters surrounding *HD*[46]. In the case of *P. noxius,* the large interval is a result of accumulation of genes transposed from different parts of the genome (**Figure 4**; Supplementary Figure 12).

Distribution of CAZymes and secondary metabolic gene clusters

Carbohydrate-active enzymes (CAZymes) are a diverse set of enzymes involved in the synthesis and degradation of carbohydrates[47]. Efficient degradation of lignin, a tough biopolymer present in woody plants, by white rot fungi such as members of Hymenochaetales requires an arsenal of ligninolytic enzymes[48]. Consistent with this, CAZymes spanning eight families of lignin-degrading enzymes (AA1-AA8)

which include laccases, peroxidases, oxidases, and reductases, and a lytic polysaccharide mono-oxygenase family (LPMO; AA9) were found in all Hymenochaetales species (Supplementary Table 10). Specialization includes the identification of copper-dependent LPMO AA11 and oxidoreductase from AA12 family only in *P. lamaensis* and *P. sulphurascens*, respectively. The most abundant and diverse among the CAZymes annotated in these four species were glycoside hydrolases (GHs). GHs act on diverse carbohydrates, including cellulose (e.g. GH1, GH3, GH5, GH6, GH7 and GH12), hemicellulose (e.g. GH5, CE1, GH3 and GH27), pectin (e.g. GH28, GH43, GH78, CE8, CE12, PL1 and PL3), starch (GH13, GH15 and GH31), and chitin (e.g. GH18, GH20, GT2 and CBM5) (Supplementary Table 10). Despite its compact genome, *P. noxius* contains a comprehensive repertoire of carbohydrate-active enzymes; a total of 4.23% (416 proteins) of its proteome were identified as CAZymes which is more than those in the other Hymenochaetales (Supplementary Figure 13). Taken together, the diversity of the CAZymes encoded in the *P. noxius* genome is consistent with its ability to infect a wide range of hosts.

The Basidiomycota phylum is known for the production of huge and diverse sets of secondary metabolites (SM) such as terpenoid compounds which have roles in virulence, survival, and host specificity[49]. Investigations on SM gene clusters revealed that nonribosomal peptide synthase (NRPS), type 1 polyketide synthase (T1PKS), and terpene synthase (TS) clusters were common among the proteomes of Hymenochaetales. For *P. noxius*, a total of 18 putative clusters were found, which included one lantipeptide, one NRPS, two T1PKS, and 11 TS clusters. These SM clusters contained 4-37 annotated biosynthesis genes (Supplementary Table 11), distributed evenly in eight of the 12 scaffolds of *P. noxius* (**Figure 2**). Clusters of lantipeptides were found in only three out of 19 compared fungi, including *P. noxius* and its closest relative *P. lamaensis*. Lantipeptides are primarily known for their antimicrobial activities, although some have other roles such as in fungal morphegenesis[19].

374 **Transcription profiles of *P. noxius* during infection of trees**

375 Upon infection of a tree, which usually starts from the root, the mycelial mat of *P.*
376 *noxius* can grow and expand on the surface of the basal stems. Occasionally,
377 basidiocarps were found in the diseased tree. To investigate genes and gene families
378 that are likely to contribute to *in planta* colonization and the stress response of *P.*
379 *noxius*, we characterized transcriptomes of mycelial mats between brown and creamy
380 sections from two adjacent Chinese elm (*Ulmus parvifolia*) trees, fruiting body, and
381 mycelia grown in potato dextrose agar (PDA) medium (2-3 biological replicates for
382 each condition; Supplementary Figure 14a). Samples from naturally infected trees
383 were used because artificial inoculation of *P. noxius* has remained challenging.
384 Limited success has been achieved by wound inoculation of seedlings on the basal
385 stem[14], but the infection site and symptom development did not fully reflect the
386 phenomenon in nature. Global transcriptome profiles of *P. noxius* are grouped
387 together according to conditions (Supplementary Figure 14b), which show a precise
388 regulatory difference of *P. noxius* at different stages. In particular, 151 (58%)
389 members of the most expanded gene family, NACHT- or WD40-encoding proteins,
390 were found to be differentially expressed in at least one pairwise comparison of
391 transcriptome profiles from different conditions (**Figure 5**). Of those, 10-35 genes
392 were found upregulated solely in one condition, suggesting a major role of this gene
393 family across multiple conditions. Other than the involvement in the incompatibility
394 reaction[50], WD40-associated proteins have been found to be central to many
395 developmental processes in ascomycetes, such as fruiting body formation in *Sordaria*
396 *macrospora*[51] and the regulation of circadian clock in *Neurospora crassa*[52].
397 Interestingly, we observed that the WD40 domain containing genes with similar
398 transcriptional profile and domain combinations are physically clustered together,
399 suggesting retained function on some of the paralogous genes since tandem
400 duplication.
401
402 When compared with transcriptome profiles of *P. noxius* grown in PDA, Cytochrome
403 P450, and CAZymes gene families dominated the list of genes upregulated in field
404 samples (mycelial mat and fruiting body). For instance, despite lacking lignin-
405 degrading peroxidases, *P. noxius* encodes 5, 2, and 11 copies of laccase, glyoxal
406 oxidase, and manganese peroxidase genes with predicted signal peptides, respectively.
407 Of these members of Auxiliary Activity (AA) family, 67% (12) are significantly

upregulated in mycelial mats. This is consistent with the fact that *P. noxius* as a white rot has the capacity for continuously degrading plant materials including lignin and responding to xenobiotics elicited by plant defence response[53] (Supplementary Table 12). Import pathways such as Major Facilitator Superfamily (MFS-1) transporters was *in silico* predicted to have a role in transporting intermediate decomposition products during wood decay [54]. This is also true for *P. noxius,* where 45% (51) out of 113 members of the MFS-1 transporter were upregulated in at least one of the field conditions. Only 10% of the *P. noxius* secretome were predicted to be putative effectors[55], which are significantly upregulated in *in planta* conditions (Supplementary Table 13). The upregulated gene list also contains hydrophobins (e.g. cerato-platanin), which contributed to virulence of various fungi[56] and were upregulated in previous inoculation experiments of *P. sulphurascens*[3].

The creamy and brown mycelial mat on infected trees are physically in close proximity and reflects relatively early and middle stages of pathogenesis rather than the late stage of vigorous decomposition. Their global transcriptome profiles were more similar to each other when compared with the other two conditions. Enrichment analysis of differentially expressed genes shows different strategies used by *P. noxius* (Supplementary Table 14): the creamy mycelium shows enrichment of upregulated genes involved in DNA replication such as ATPases and DNA topoisomerase, suggesting *P. noxius* is actively progressing outwards along the tree bark (Supplementary Figure 15). In contrast, members of P450 and CAZymes are significantly upregulated in the dark mycelial mat, suggesting an even higher expression of these genes during the middle stage of colonization.

Sequence analysis of 60 *P. noxius* strains

To further understand the regional dissemination of *P. noxius* in Taiwan and Japanese offshore islands, we sequenced the genomes of 60 isolates (Figure 6a) originating from diseased trees in urban cities as well as trails in natural parks. This collection was sequenced to a median depth of coverage of 35X, and additional mate pair libraries were sequenced in two isolates (718-S1 and A42) which were used to characterize synteny relationship amongst strains. An average of 96% mapping rate

was achieved after aligning reads from each strain to the KPN91 reference genome. The descriptions and statistics of the strains are provided in Supplementary Table 15.

Stable structural variation and frequencies of heterokaryons in *P. noxius*

The detection of frequency of heterokaryons and polyploids of *P. noxius* in nature remain challenging as its arthospores contain 1-5 nuclei (based on quantification of 145 arthospores from A42 in this study) and multiple allelic fragments in SSRs (2-4 alleles per locus) are commonly found in populations [4, 6]. All the strains were isolated from either a single arthospore or fungal mat, and thus allowed for the analysis of variation in genome structure. We found no deviation in coverage across every scaffold in all strains (Supplementary Figure 16), suggesting a stable number of chromosome copies in *P. noxius*. To distinguish the frequency of mono- or heterokaryons in nature, heterozygous SNPs and minor allele frequency (MAF) distribution was inferred in each strain (Supplementary Figure 17). We identified four groups (A to D) that clearly differ in heterozygosity and MAF. The group with the lowest heterozygosity of averaging 0.2 % showed a random MAF distribution. In particular, strains 718-S1 and A42 each isolated from single basidiospores fall in this group suggesting they are monokaryotic in nature. Conversely, the other three groups have a much higher heterozygosity, with the highest group of 1.6% heterozygosity displaying a peak around 50% in MAF distribution. Peaks around 27%-33% are also observed in two of the groups, and this is not due to multiple nuclei as the distributions of the number of nuclei per cell in seven randomly selected isolates are the same (Supplementary Figure 18). Groups B-D may refer to different compositions of two or more genetically distinct nuclei. The number of monokaryons present in nature is 40% and can be found across all regions surveyed in this study.

Diversification of *P. noxius* across Pacific Ocean islands

To examine the population structure of *P. noxius* in this area of Pacific Islands, we performed principal-component analysis (Figure 6b; Supplementary Figure 19) on the SNP variants from individual phased haploid genomes. Similar to previous report based on SSRs on highly diversified populations between Japanese offshore islands[4], the major principal component divided Ogasawara islands samples from

the rest of the samples, which are located geographically 1,210 km apart. The second and third principal components defined a tight cluster of Taiwanese isolates with considerable overlaps between cities despite being 170 km apart. Most isolates from Ryukyu islands can be differentiated from those from Taiwan. Interestingly, we identified three isolates from Taiwan that are distinctive from both the Taiwan-Ryukyu and Ogasawara island clusters, suggesting a possibility of more populations within the Taiwan island.

We constructed a maximum likelihood phylogeny based on 1,837,281 high confidence SNPs of *P. noxius* (Figure 6c). Concordant with the PCA, the strains from Ogasawara islands form a distinctive lineage. We also observed that the majority of Taiwan and Ryukyu islands strains are grouped together forming another major lineage, suggesting the possibility of gene flow in this region despite physical separations by the sea. We inferred the population structure in 144,426 unlinked bialleic sites using the Bayesian model-based clustering approach[57]. Consistent with the phylogeny, fastSTRUCTURE[57] also identified two Taiwan-Ryukyu and Ogasawara lineages circulating in this region (Figure 6c, Supplementary Figure 20).

High nucleotide diversities in *P. noxius* populations

The *de novo* assemblies of three strains of *P. noxius* have on average 97% nucleotide identity and are largely co-linear to each other with apparent genome translocations between Taiwan/Ryukyu (A42 and KPN91) and the Ogasawara (718-S1) isolate (Supplementary Figure 21). We quantified sequence diversity using θs and θπ and categorised them into whole genome average, four-fold synonymous sites and replacement sites (Table 1). An large number of SNPs are found in the genome of *P. noxius*, averaging one SNP identified in every 20-59 bp. Nucleotide diversity at synonymous sites θπ-syn is 15.8 and 19.2 per kilobase in Ogasawara island and Taiwan-Ryukyu populations, respectively. This is ~5 fold higher when compared to the majority of species[58] and is likely an underestimate of true diversity as indels were not considered. Extremely high diversity has been reported in natural populations of *Schizophyllum commune*[59]. The Tajima's D is strongly negative in the Taiwan and Ryukyu islands lineage, suggesting an excess of low frequency alleles present in the population possibly as a result of high mutation rate[59] or population

510 expansion. In the Ogasawara population, nucleotide diversity is reduced compared to
511 the Taiwan/Ryukyu lineage and Tajima's D is positive implying a reduced effective
512 population size; *P. noxius* may have been introduced recently in these islands. By *de*
513 *novo* assembly and annotation of these isolates, we were able to identify a total of 25
514 differential content of Pfam domains between these two lineages ($P \leq 0.05$; Wilcoxon
515 rank-sum test; Supplementary Figure 22). Of them, Arv1 can be found in 37/51
516 strains in Taiwan-Ryukyu lineage which may be related to the traffic of sterol[60].
517 Together our data suggest that the two *P. noxius* lineages may have derived from
518 genetically distinct gene pools and have undergone divergent evolutionary scenarios,
519 possibly as a result of different time of introduction, and different environments
520 and/or human interference in Taiwan-Ryukyu *vs* Ogasawara areas.
521
522
523 **DISCUSSION**
524
525 Here we report four high-quality genome sequence of Hymenochaetales species that
526 are global tree pathogens of particular importance. Orthologous relationships with
527 other complete genomes of basidiomycetes have confirmed conservation of
528 karyotypes with few fusion or breaks in Agaricomycetes. Our study has shown how a
529 pathogenic wood decay fungus regulates global transcription at different stages of
530 colonization on tree host in nature. Such orchestrated responses have been
531 demonstrated in a few other white rot fungi, including *Pycnoporus coccineus* during
532 cultivation in three types of lignocellulosic substrates (i.e. cereal straw, softwood, and
533 hardwood)[55], and *Heterobasidion irregular* during growth in wood substrates, in
534 the cambial zone of pine, and as fruiting bodies[22]. Although the initial phase for *P.*
535 *noxius* to invade the root of a living tree remains to be elucidated, our combined
536 genome and transcriptome analysis reiterate the importance of CAZymes and P450
537 families in host adaptation, the involvement of MFS-1 in transporting the intermediate
538 decomposition products, and the roles of abundant WD40-encoding proteins in
539 modulating cell differentiation and responses to environmental stimuli. The results
540 can serve as a starting point for understanding the survival strategies of *P. noxius*,
541 from which new control methods could be developed.
542

543 Population genomics analyses of *P. noxius* suggest that it is a hypervariable species.
544 Our investigations into mating type loci and genome-wide heterozygosity further
545 indicated that the genetic hyperdiversity can be attributed to a bipolar heterothallic
546 reproductive system and heterokaryotic nature, though gene flow and/or high
547 mutation rate may also play some roles. The exceptional large separation between *HD*
548 pairs would allow higher probability of recombination between the physically distant
549 *HD* genes[41, 46], thus resulting in progeny with more diverse mating types which
550 are ready to mate. Notably, both monokaryotic and heterokaryotic state of *P. noxius*
551 mycelia are prevalent in the nature, and some isolates likely contained more than two
552 genetically different nuclei[6]. In addition, some isolates were able to produce
553 basidiocarps by themselves when cultured on sawdust medium (P.-J. Ann, pers.
554 comm.), suggesting dikaryotization or homokaryotic fruiting occurred spontaneously
555 or in response to certain conditions[63]. It would be interesting to further investigate
556 the complex regulatory mechanisms underlying anastomosis, karyogamy, and meiotic
557 division during vegetative growth and basidiocarp formation. Transcriptional
558 diversity among genetically variable individuals is also warrant further exploration.
559
560 How *P. noxius* is spread in regions of Asia has been the focus of a few studies[4, 6,
561 64]. Field observations support the root-to-root spread as a major mode of
562 transmission. Human activity such as planting of infected seedlings may have
563 promoted the movement of the strains between the islands of Taiwan and the Ryukyu
564 region, which provide an artificial environment for population to increase. The
565 within-population hyperdiversity and sporadic pattern of new disease foci also
566 support the involvement of basidiospores in *P. noxius* dissemination. Abundant
567 basidiospores produced from perennial fruiting bodies may have the ability to travel
568 long distance (e.g. the basidiospores of *Heterobasidion annosum* and *Peniophora*
569 *aurantiaca* were captured 50–500 km and ~1000 km apart from the inoculum source).
570 Following the dispersal, highly variable individuals of *P. noxius* would have good
571 chance of successful colonization. A preliminary model based on 19 bioclimatic
572 variables of known locations of ~100 *P. noxius* isolates from south eastern Asia,
573 Australia, and Pacific islands predicted that extensive global regions are at risk, which
574 includes a big part of the South American continent[65].
575
576

CONCLUSION

The Hymenochaetales is phylogenetically placed between the better-studied Agaricales and Ustillaginales orders, making the reference assembly of *P. noxius* KPN91 an attractive genome to study the evolutionary transition between these orders of Basidiomycota. It should be of continuous interest to confirm this observation when more reference genome assemblies become available. Genetic hyperdiversity of *P. noxius* suggests that the pathogen may have greater adaptability to different environments and stresses. Much more attention has to be paid on how global warming and increasing frequency of extreme weather events would enhance disease severity and allow the colonization of a fast grower like *P. noxius* in the warm temperate zones.

FIGURES AND TABLES

Figure 1. Life stages of *Phellinus noxius* and comparative genomic analysis of Hymenochaetales species. (**a**) The mycelial mat with young creamy leading front and aged brown section. (**b**) In advanced stage of decay, the hyphae form a network of brown zone lines permeating the soft and white wood tissue. (Lower left and lower right) (**c and d**) Basidiocarps are perennial and can be resupinate (**c**) or grow into a sessile bracket-like conk with a broad basal attachment (**d**). The distinctive greyish-brown surface is the hymenial layer with irregularly polygonal pores, containing four-spored basidia, ellipsoid and hyaline basidiospores, but no hymenial setae. (**e**) The phylogeny of four *Phellinus* species with 15 other species of Basidiomycota based on a concatenated alignment of single-copy orthologous genes. All nodes have 100 out of 100 bootstrap replicates. The numbers of gained ("+") and lost ("-") gene families along each branch of the phylogeny is annotated.

Figure 2. The twelve assembled scaffolds of *Phellinus noxius* with telomeres (denoted as T). The plot shows the distribution and density of LTR retrotransposons and various gene families of interest denoted in different colours.

Figure 3. Linkage group (LG) network of Basidiomycota. (**a**) The linkage groups were identified by linking single-copy orthologs of scaffolds between species pairs. All scaffolds included in this plot are larger than 500kb. Each directional edge points toward the reference chromosome. Edges were weighted by the number of single-copy orthologs, and an edge was filtered if it has a weight smaller than 20 or less than 40% of the sum of all weights out of a node. (**b**) Cross-mapping of single-copy orthologs in LG10 and 11. Scaffold names are shown at the upper-right side of each sequence, with detectable telomeric regions labelled as 'T' at the upper-left side.

Figure 4. Synteny around the *A* mating locus. The analysis included *Phellinus noxius, Porodaedalea pini, Phellinus sulphurascens, Phellinus lamaensis*, two other species from Hymenochaetales species (*Fomitiporia mediterranea, Schizopora paradoxa*), two species from Agaricales (*Laccaria bicolar, Coprinopsis cinerea*), *Sistotremastrum niveocremeum*, and *Ustilago maydis*.

Figure 5. WD40 domain containing genes upregulated in the transcriptome of *Phellinus noxius* at different conditions. RNA-seq counts of these genes (in log of FPM, fragments per million) for *P. noxius* cultured in PDA, as mycelial mat and fruiting body on infected trees are presented as different colours in a heatmap sorted according to their physical positions on the assembly. The results of differentially expressed genes from pairwise DEseq2 comparisons (padj < 0.01) are shown in different colours each corresponding to a specific life stage of *P. noxius*. Different combinations of protein domains are categorised based on their initial domain from the N termini. Genes with combination that is shared no more than four other genes are categorised into 'Others' (3 of ANAPC4, 1 of BOP1NT, 1 of CAF1C_H4-bd, 5 of Pkinase, 1 of PRP4, 1 of Striatin, 1 of Tup_N).

Figure 6. The population genomics of 60 *Phellinus noxius* isolates. (**a**) Map of Taiwan and offshore islands of Japan showing origins of the 60 sampled *P. noxius* isolates. Ogasawara islands were conveniently drawn below Ryukyu island so this does not represent their actual location. (**b**) PCA plot of 60 *P. noxius* isolates using genome-wide variation data sampled from 13 islands by the first three eigenvectors. (**c**) Top: Phylogenetic tree with 100 bootstrap using SNPs computed from alignment to KPN91 reference. Point separating the Taiwan-Ryukyu and Ogasawara island isolates was used as root. Nodes with >90% bootstrap were labelled with circles. Bottom: fastSTRUCTURE analysis of the linkage independent pruned set of variation data. A model with two ancestral components (K=2) had the highest likelihood to explain the variation of genome-wide structure on the 60 isolates. Also see Supplementary Figure 20 for different K.

Table 1. Polymorphism in the two regional lineages of *Phellinus noxius*.

679 **METHODS**

680

681 **Strain preparation and sequencing**

682 Genome sequencing of three *Phellinus* species and *P. pini* (Supplementary
683 Information) was performed using both Pacific Biosciences (*P. noxius*, *P. lamaensis,*
684 *P. sulphurascens* and *P. pini*) and Illumina (*P. noxius*) platforms. DNA was isolated
685 using the CTAB method[6]. At least 15 µg DNA was used for a 20 kb library prep
686 according to the manufacturer's instructions. Sequencing was performed on a Pacific
687 Biosciences RS II system using 8 SMRT cell per run, P6C4 chemistry and 360 min
688 movie time. A total of 5-7 SMRT cells were run per species yielding a raw depth of
689 coverage of 173-266X. Illumina shotgun genomic libraries were sheared by Covaris
690 and genomic libraries prepared using TruSeq DNA LT Sample Prep Kit (Illumina).
691 Input of 1 µg sheared DNA was used for end-repair, A-tailing, adaptor ligation, and
692 gel size selection. Size range at 600-700bps range was selected form gel and
693 amplified by 5 cycles of PCR. In addition, Illumina mate-pair libraries were generated
694 using 8 µg of genomic DNA with gel size selection of tagmented DNA at 2-4kb, 4-
695 7kb, 7-10kb, and 10-20kb, and amplified by 10~15 cycles of PCR. These libraries
696 were normalized by KAPA Library Quantification Kit (KAPA Biosystems), and
697 pooled equally for PE2*300 sequencing on MiSeq V2 sequencer.

698

699 To aid annotation for each of the species in this study, 7- to 21-day old mycelia from
700 PDA cultures were used for RNA-seq. For field samples, *Phellinus noxius* mycelial
701 mat on creamy and dark region of two Chinese elm trees (*Ulmus parvifolia*) adjacent
702 to each other in Academia Sinica, Taiwan were scraped down using alcohol wiped
703 tweezers. The fruiting body of *Phellinus noxius* was found on *Machilus kobu* trees,
704 Ogasawara island, Japan and was removed aseptically. On diseased trees, dried (and
705 possibly dead) mycelial mats are most commonly observed. Fresh mycelial mat and
706 fruiting body only occasionally occur and we were not able to collect both samples
707 from the same tree. RNA was isolated with Direct-zol (Zymoresearch) according to
708 manufacturer's instructions. RNA-seq libraries were constructed using the Illumina
709 TruSeq Stranded mRNA HT Sample Prep Kit with the dual index barcoded adaptors.
710 Input of 3 µg of total RNA was used for each sample for two rounds of oligo-dT bead
711 enrichment, and the ligated cDNA were amplified by 10 cycles of PCR. The Stranded
712 mRNA libraries were quantified by Qubit and molar concentrations normalized

against the KAPA Library Quantification Kit (KAPA Biosystems) for Illumina platform. The transcriptome libraries were pooled at equal molar concentrations, and PE2*151nt multiplexed sequencing was conducted on HiSeq 2500 sequencer.

Nuclear quantification

Nuclei in the growing hyphal tips were stained following the procedure described by Chung et al. (2015). The mycelium on the slide was mounted in 20 µl of a DAPI (4',6-diamidino-2-phenylindole) solution (10 µg/ml in ddH_2O) for an hour, destained in ddH_2O for 30 min, then observed under a OLYMPUS BX41 microscope (Shinjuku-ku, Tokyo, Japan) equipped with filter cube U-MWU2 (BP 330–385 nm, LP 420 nm). Images were captured by using a Canon (Ohta-ku, Tokyo, Japan) digital camera EOS 700D. One hundred hyphal cells per strain were counted.

Genome assembly

Genome assembly of different species was carried out with Falcon[66] (ver 0.5.0) and were improved using Quiver[67] and finisherSC[68]. For assembly of individual strains of *P. noxius*, Illumina paired end reads were trimmed with Trimmomatic[69] (ver 0.32; options LEADING:30 TRAILING:30 SLIDINGWINDOW:4:30 MINLEN:50) and subsequently assembled using SPAdes[70] (ver 3.7.1). Multiple mate-pair reads were available for three strains of *P. noxius* (KPN91, A42 and 718-S1) and they were assembled using ALLPATH-LG[71] assembler (ver 49688) and improved using Pilon[72]. The *P. noxius* assembly was further merged with metassembler[73], miss-assemblies were identified using REAPR[74] and manually corrected.

Gene predictions and functional annotation

For *P. noxius,* the gene predictor Augustus[75] (v 3.2.1) was trained on a gene training set of complete core genes from CEGMA[76] and subsequently used for manual curation of ~1000 genes. Annotation was then run by providing introns as evidence from RNA-seq data. For *P. lamaensis*, *P. sulphurascens* and *P. pini*, genes were predicted using Braker1[77] pipeline that automatically use RNA-seq mappings as evidence hints and retraining of GeneMark-ES[78] and Augustus. Gene product description was assigned using blast2go[79] (ver 4.0.7) and GO term assignment were provided by ARGOT2.5[80]. The web server dbCAN[81] (HMMs 5.0, last accessed

September 5 2016) was used to predict CAZymes from the protein sequences of all species, while AntiSMASH[82] (version 3.0) was used to predict secondary metabolite gene clusters. For dbCAN results, only hits with <= 1 x 10e-5 e-value, and >= 30% HMM coverage were considered, while overlapping domains were resolved by choosing hits with the smallest *P* value.

Repeat analysis and repeat induced mutation analysis

Consensus sequences of repetitive elements in the assembly of the four Hymenochaetales species were identified based on UCLUST[83] clustering of 80% sequence identity from two repeat detection programs: RepeatModeler (www.repeatmasker.org) and TransposonPSI (http://transposonpsi.sourceforge.net/). Annotations of these sequences were based on output of both programs and BLAST results from the National Center for Biotechnology Information (NCBI) nonredundant database. RepeatMasker (v3.2.8, http://www. repeatmasker.org) was used to calculate the distribution of each repeat and its abundance. Consensus sequences of the most abundant copy were manually curated based on pfam[84] (ver 30.0) domain prediction from the longest open reading frame. Whenever a LTR retrotransposon was identified, long terminal repeats (LTR) and associated features were predicted using LTRharvest[85]. Repeat induced mutation (RIP) analysis was carried out using RIPCAL[30]

Comparative genomics analysis

Gene families were determined by OrthoFinder[86] (ver 1.0.6). Then, MAFFT[87] (ver 7.271) was then used to align sequences in each of 1,127 single-copy orthogroups. Alignments results with less than 10% alignment gaps were concatenated, and the outcome was taken to compute a maximum likelihood phylogeny using RAxML[88] with 100 bootstrap replicates. Gene family gain and loss in different positions along the global phylogeny leading to *P. noxius* were inferred using dollop[89]. Pfam and GO enrichments were carried out on these gene families using TopGO[90] (ver 2.10.0). Sequences to be included in the phylogenetic tree for NACHT domain containing genes were selected on the basis of the presence of pfam domain PF05729.

781 **RNA-seq analysis of *P. noxius* at different conditions**

782 For differential gene expression analysis, paired-end Illumina RNA-seq data were
783 mapped against *P. noxius* genome using STAR[91] and number of read counts per
784 gene were determined with featureCounts[92] (ver 1.5.0-p1). DESeq2[93] was used to
785 determine the reliability of our replicates and determine differentially expressed genes
786 between conditions. Pairwise comparison of each transcript is identified as
787 differentially expressed if adjusted p-value is lower than 0.01 and the absolute value
788 of log2(fold change) is greater than 1. GO term enrichment of up- and down-regulated
789 genes between conditions was carried out with GOseq[94].
790
791 **MAT locus**
792 Homologs of the conserved genes in mating locus *A* and *B* were annotated in *P.*
793 *noxius*, *P. pini*, *P. sulphurascens* and *P. lamaensis* (e-value < 10^{-5}) then subjected to
794 InterProScan 5[95] (ver 5.20-59.0) and Pfam[84] (ver 30.0) for protein signature
795 prediction. Syntenic alignment of *A* locus among *Phellinus* spp. and other species
796 (sequences/annotations retrieved from Joint Genome Institute MycoCosm[96]) was
797 plotted using genoPlotR package[97] in R. The sequences of *A* locus in KPN91, 718-
798 S1 and A42 were further compared by MUMmer[98] (ver 3.20) and PipMaker[99]
799 (http://pipmaker.bx.psu.edu/cgi-bin/pipmaker?basic). For identification of candidate
800 pheromone genes, all the potential open reading frames were filtered for small
801 peptides with C-terminal CaaX motif, then searched for pheromone homologies
802 against Pfam-A and scanned for farnesylation signal by PrePs[100]
803 (http://mendel.imp.ac.at/PrePS/).
804
805 The sequences of *HD* and *STE3* genes were analyzed for 10 single-basidiospore
806 isolates originating from a basidiocarp by the dideoxy termination method (primers
807 listed in Supplementary Table 16). Long-range PCR followed by primer walking was
808 performed to sequence the highly diverse regions containing *HD1-HD2* in *A* locus.
809 The outer primers were designed manually based on the alignment of all the isolates;
810 the inner primers were developed step by step according to previous sequencing
811 results. The PCR reaction was performed in 30-µl reaction mixture containing 50 to
812 100 ng genomic DNA, 0.2 mM dNTP, 1X Ex Taq buffer [proprietary, containing
813 20mM Mg^{2+}] (Takara Bio Inc., Japan), 0.67 µM forward and reverse primers, and
814 TaKaRa Ex Taq® DNA polymerase (Takara Bio Inc., Japan). The thermal cycling

parameters were 1 cycle of 95°C for 3 min, 30 cycles of 95°C for 30 s, 54°C for 30 s, and 72°C for 60–270 s (according to different product sizes, ~2 kb/min), and a final extension step of 72°C for 10 min. The PCR products were sequenced by Genomics Biotechnology Inc. (Taipei, Taiwan). DNA trace data were visualized using 4Peaks (http://nucleobytes. com/4peaks) and assembled using DNA Sequence Assembler in DNA Baser (http://www. dnabaser.com).

Population genomics

Paired end reads of 60 *P. noxius* strains were aligned to the KPN91 reference using Smalt v5.7 (www.sanger.ac.uk/resources/software/smalt/). Removal of PCR duplicates and bam file sorting were implemented with Picard (http://broadinstitute.github.io/picard/) and samtools[101] (ver 1.3-20-gd49c73b). The first round of variant identification was implemented in Varscan[102] (v.2.4.0) and degree of heterokaryon was inferred in each strain based on allele frequency and total number of heterozygous SNPs called. The final list of SNPs was ascertained by combining evidences from samtools[101] and FreeBayes[103] (ver1.0.2-16-gd466dde). A maximum likelihood phylogeny of the SNPs segregating in these 60 strains was produced using FastTree[104]. plink[105] (v1.9) was used to subset biallelic SNPs without linkage (filtering options: --maf 0.05 --indep-pairwise 50 5 0.2), which were clustered using fastSTRUCTURE[57] to determine number of populations. Strains were phased using samtools[101] and one haplotype was chosen at random. Consensus sequences were generated from each strain using bcftools[106] and population genetics analyses were conducted using Variscan[107].

DECLARATIONS

Ethics approval, consent to participate and consent for publication:

Not applicable in this study.

Availability of data and material:

Genome assemblies and annotations are available from NCBI under whole genome shotgun (WGS) ID: NBII00000000 (*P. noxius*), NBBA00000000 (*P. sulphurascens*), NBAY00000000 (*P. pini*) and NBAZ00000000 (*P. lamaensis*). Bioproject and Biosample ID of raw data are available in Supplementary Table 15, 17 and 18.

Competing interests:

The authors declare that they have no competing interests.

Funding:

I.J.T. and J.T.L were funded by MOST (105-2628-B-001-002-MY3 and 105-2313-B-001 -003). C.L.C. was funded by Office of General Affairs, National Taiwan University.

Author's contributions:

Strain cultivation and preparation: H.H.L., J.T.L, H.M.K, M.A., T.H., Y.O., N.S. and T.K. Strain provider: C.L.C., R.F.L., S.S.T., P.J.A., J.N.T., M.A., T.H., Y.O., N.S. Strain sequencing and assembly: C.L.C., H.H.L., C.Y.C., M.J.L., T.K. and I.J.T. Annotation and manual curation: H.M.K., J.T.L. and I.J.T. Comparative genomics analysis: T.H.K, D.L., M.B.R., H.M.K. and I.J.T. Population genomics analysis: H.M.K., J.T.L. and I.J.T. Mating locus analysis: C.L.C., H.H.L., Y.Y.C., T.H.K., I.J.T. RNA-seq analysis: J.T.L and I.J.T. Wrote the manuscript: C.L.C, J.T.L, H.H.L, T.K and I.J.T. Conceived and directed the project: C.L.C., T.K. and I.J.T.

Acknowledgements:

We thank Yen-Ping Hsueh (IMB, Academia Sinica) and Robert Waterhouse (University of Lausanne) for commenting the manuscript. We thank Chian-mei Wei, Kuan-chun Chen (High Throughput Genomics Core at Biodiversity Research Center, Academia Sinica) for sequencing. We thank Tun-Tschu Chang and Yu-Ching Huang for strain collection and consultation.

REFERENCES

1. Gross A, Holdenrieder O, Pautasso M, Queloz V, Sieber TN: **Hymenoscyphus pseudoalbidus, the causal agent of European ash dieback.** *Mol Plant Pathol* 2014, **15:**5-21.
2. Potter C, Harwood T, Knight J, Tomlinson I: **Learning from history, predicting the future: the UK Dutch elm disease outbreak in relation to contemporary tree disease threats.** *Philos Trans R Soc Lond B Biol Sci* 2011, **366:**1966-1974.
3. Williams HL, Sturrock RN, Islam MA, Hammett C, Ekramoddoullah AK, Leal I: **Gene expression profiling of candidate virulence factors in the laminated root rot pathogen Phellinus sulphurascens.** *BMC Genomics* 2014, **15:**603.
4. Akiba M, Ota Y, Tsai IJ, Hattori T, Sahashi N, Kikuchi T: **Genetic Differentiation and Spatial Structure of Phellinus noxius, the Causal Agent of Brown Root Rot of Woody Plants in Japan.** *Plos One* 2015, **10:**e0141792.
5. Ann P, Chang T, Ko W: **Phellinus noxius Brown Root Rot of Fruit and Ornamental Trees in Taiwan.** *Plant disease* 2002, **86**.
6. Chung C-L, Huang S-Y, Huang Y-C, Tzean S-S, Ann P-J, Tsai J-N, Yang C-C, Lee H-H, Huang T-W, Huang H-Y, et al: **The Genetic Structure of Phellinus noxius and Dissemination Pattern of Brown Root Rot Disease in Taiwan.** *Plos One* 2015, **10:**e0139445.
7. Goberville E, Hautekeete NC, Kirby RR, Piquot Y, Luczak C, Beaugrand G: **Climate change and the ash dieback crisis.** *Sci Rep* 2016, **6:**35303.
8. Fisher MC, Henk DA, Briggs CJ, Brownstein JS, Madoff LC, McCraw SL, Gurr SJ: **Emerging fungal threats to animal, plant and ecosystem health.** *Nature* 2012, **484:**186-194.
9. Cunniffe NJ, Cobb RC, Meentemeyer RK, Rizzo DM, Gilligan CA: **Modeling when, where, and how to manage a forest epidemic, motivated by sudden oak death in California.** *Proc Natl Acad Sci U S A* 2016, **113:**5640-5645.
10. H HG: *Diseases of forest and shade trees of the United States.* U.S. Dept. Of Agriculture Forest Service Handbook Number 386; 1971.
11. Sahashi N, Akiba M, Ishihara M, Ota Y, Kanzaki N: **Brown root rot of trees caused by Phellinus noxius in the Ryukyu Islands, subtropical areas of Japan.** *Forest Pathology* 2012, **42:**353-361.
12. Sahashi N, Akiba M, Takemoto S, Yokoi T, Ota Y, Kanzaki N: **Phellinus noxius causes brown root rot on four important conifer species in Japan.** *European Journal of Plant Pathology* 2014, **140:**869--873.
13. Ann P-J, Chang T-T, Ko W-H: **Phellinus noxius brown root rot of fruit and ornamental trees in Taiwan.** *Plant Disease* 2002, **86:**820--826.
14. Ann PJ, Lee HL, Huang TC: **Brown Root Rot of 10 Species of Fruit Trees Caused by Phellinus noxius in Taiwan.** *Plant Disease* 1999, **83:**746-750.
15. Sahashi N, Akiba M, Ishihara M, Miyazaki K, Kanzaki N: **Cross Inoculation Tests with Phellinus noxius Isolates from Nine Different Host Plants in the Ryukyu Islands, Southwestern Japan.** *Plant Disease* 2010, **94:**358--360.
16. Nandris D, Nicole M, Geiger JP: **Variation in virulence among Rigidoporus lignosus and Phellinus noxius isolates from West Africa1.** *European Journal of Forest Pathology* 1987, **17:**271-281.

17. Chang T-T: **Decline of some forest trees associated with brown root rot caused by Phellinus noxius.** *Plant Pathology Bulletin* 1992, **1:**90-95.
18. Schwarze F, Jauss F, Spencer C, Hallam C, Schubert M: **Evaluation of an antagonistic Trichoderma strain for reducing the rate of wood decomposition by the white rot fungus Phellinus noxius.** *Biological Control* 2012, **61:**160--168.
19. Wu J, Peng SL, Zhao HB, Tang MH, Li FR, Chen BM: **Selection of species resistant to the wood rot fungus *Phellinus noxius*.** *European Journal of Plant Pathology* 2011, **130:**463-467.
20. Huang H, Sun L, Bi K, Zhong G, Hu M: **The effect of phenazine-1-carboxylic acid on the morphological, physiological, and molecular characteristics of *Phellinus noxius*.** *Molecules* 2016, **21:**613.
21. Zhou L-W, Vlasák J, Dai Y-C: **Taxonomy and phylogeny of *Phellinidium* (Hymenochaetales, Basidiomycota): A redefinition and the segregation of *Coniferiporia* gen. nov. for forest pathogens.** *Fungal Biology* 2016, **120:**988-1001.
22. Stajich JE, Wilke SK, Ahren D, Au CH, Birren BW, Borodovsky M, Burns C, Canback B, Casselton LA, Cheng CK, et al: **Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom Coprinopsis cinerea (Coprinus cinereus).** *Proceedings of the National Academy of Sciences* 2010, **107:**11889-11894.
23. Alfaro M, Castanera R, Lavín JL, Grigoriev IV, Oguiza JA, Ramírez L, Pisabarro AG: **Comparative and transcriptional analysis of the predicted secretome in the lignocellulose-degrading basidiomycete fungus Pleurotus ostreatus.** *Environmental Microbiology* 2016, **18:**4710-4726.
24. Foulongne-Oriol M, Rocha de Brito M, Cabannes D, Clément A, Spataro C, Moinard M, Dias ES, Callac P, Savoie J-M: **The Genetic Linkage Map of the Medicinal Mushroom Agaricus subrufescens Reveals Highly Conserved Macrosynteny with the Congeneric Species Agaricus bisporus.** *G3: Genes|Genomes|Genetics* 2016, **6:**1217-1226.
25. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM: **BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs.** *Bioinformatics* 2015, **31:**3210-3212.
26. Larsson K-H, Parmasto E, Fischer M, Langer E, Nakasone KK, Redhead Sa: **Hymenochaetales: a molecular phylogeny for the hymenochaetoid clade.** *Mycologia* 2006, **98:**926-936.
27. Castanera R, Borgognone A, Pisabarro AG, Ramirez L: **Biology, dynamics, and applications of transposable elements in basidiomycete fungi.** *Appl Microbiol Biotechnol* 2017, **101:**1337-1350.
28. Selker EU: **Premeiotic instability of repeated sequences in Neurospora crassa.** *Annu Rev Genet* 1990, **24:**579-613.
29. Horns F, Petit E, Yockteng R, Hood ME: **Patterns of Repeat-Induced Point Mutation in Transposable Elements of Basidiomycete Fungi.** *Genome Biology and Evolution* 2012, **4:**240-247.
30. Hane JK, Oliver RP: **RIPCAL: a tool for alignment-based analysis of repeat-induced point mutations in fungal genomic sequences.** *BMC Bioinformatics* 2008, **9:**478.
31. Hane JK, Rouxel T, Howlett BJ, Kema GH, Goodwin SB, Oliver RP: **A novel mode of chromosomal evolution peculiar to filamentous Ascomycete fungi.** *Genome Biol* 2011, **12:**R45.

32. Lucía R, Gúmer P, Raúl C, Francisco S, G. PA: *Basidiomycetes Telomeres – A Bioinformatics Approach.* InTech; 2011.
33. Foulongne-Oriol M, Murat C, Castanera R, Ramirez L, Sonnenberg AS: **Genome-wide survey of repetitive DNA elements in the button mushroom Agaricus bisporus.** *Fungal Genet Biol* 2013, **55:**6-21.
34. Edman JC: **Isolation of telomere like sequences from Cryptococcus neoformans and their use in high-efficiency transformation.** *Mol Cell Biol* 1992, **12:**2777-2783.
35. Tudzynski B: **Nitrogen regulation of fungal secondary metabolism in fungi.** *Front Microbiol* 2014, **5:**656.
36. Douglas CM: **Fungal beta(1,3)-D-glucan synthesis.** *Med Mycol* 2001, **39 Suppl 1:**55-66.
37. Leipe DD, Koonin EV, Aravind L: **STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer.** *J Mol Biol* 2004, **343:**1-28.
38. Bidard F, Clave C, Saupe SJ: **The transcriptional response to nonself in the fungus Podospora anserina.** *G3 (Bethesda)* 2013, **3:**1015-1030.
39. Van der Nest MA, Olson A, Lind M, Velez H, Dalman K, Brandstrom Durling M, Karlsson M, Stenlid J: **Distribution and evolution of het gene homologs in the basidiomycota.** *Fungal Genet Biol* 2014, **64:**45-57.
40. Jancova P, Anzenbacher P, Anzenbacherova E: **Phase II drug metabolizing enzymes.** *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* 2010, **154:**103-116.
41. James TY, Sun S, Li W, Heitman J, Kuo HC, Lee YH, Asiegbu FO, Olson A: **Polyporales genomes reveal the genetic architecture underlying tetrapolar and bipolar mating systems.** *Mycologia* 2013, **105:**1374-1390.
42. James TY, Lee M, van Diepen LT: **A single mating-type locus composed of homeodomain genes promotes nuclear migration and heterokaryosis in the white-rot fungus Phanerochaete chrysosporium.** *Eukaryot Cell* 2011, **10:**249-261.
43. James TY, Srivilai P, Kues U, Vilgalys R: **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* 2006, **172:**1877-1891.
44. Niculita-Hirzel H, Labbe J, Kohler A, le Tacon F, Martin F, Sanders IR, Kues U: **Gene organization of the mating type regions in the ectomycorrhizal fungus Laccaria bicolor reveals distinct evolution between the two mating type loci.** *New Phytol* 2008, **180:**329-342.
45. Kües U: **From two to many: multiple mating types in Basidiomycetes.** *Fungal Biology Reviews* 2015, **29:**126-166.
46. van Peer AF, Park S-Y, Shin P-G, Jang K-Y, Yoo Y-B, Park Y-J, Lee B-M, Sung G-H, James TY, Kong W-S: **Comparative genomics of the mating-type loci of the mushroom *Flammulina velutipes* reveals widespread synteny and recent inversions.** *PLOS ONE* 2011, **6:**e22249.
47. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B: **The carbohydrate-active enzymes database (CAZy) in 2013.** *Nucleic Acids Res* 2014, **42:**D490-495.
48. Dashtban M, Schraft H, Syed TA, Qin W: **Fungal biodegradation and enzymatic modification of lignin.** *Int J Biochem Mol Biol* 2010, **1:**36-50.

49. Osbourn A: **Secondary metabolic gene clusters: evolutionary toolkits for chemical innovation.** *Trends Genet* 2010, **26:**449-457.
50. Paoletti M, Saupe SJ: **Fungal incompatibility: evolutionary origin in pathogen defense?** *Bioessays* 2009, **31:**1201-1210.
51. Poggeler S, Kuck U: **A WD40 repeat protein regulates fungal cell differentiation and can be replaced functionally by the mammalian homologue striatin.** *Eukaryot Cell* 2004, **3:**232-240.
52. He Q, Cheng P, He Q, Liu Y: **The COP9 signalosome regulates the Neurospora circadian clock by controlling the stability of the SCFFWD-1 complex.** *Genes Dev* 2005, **19:**1518-1531.
53. Lah L, Podobnik B, Novak M, Korosec B, Berne S, Vogelsang M, Krasevec N, Zupanec N, Stojan J, Bohlmann J, Komel R: **The versatility of the fungal cytochrome P450 monooxygenase system is instrumental in xenobiotic detoxification.** *Mol Microbiol* 2011, **81:**1374-1389.
54. Nagy LG, Riley R, Bergmann PJ, Krizsan K, Martin FM, Grigoriev IV, Cullen D, Hibbett DS: **Genetic Bases of Fungal White Rot Wood Decay Predicted by Phylogenomic Analysis of Correlated Gene-Phenotype Evolution.** *Mol Biol Evol* 2017, **34:**35-44.
55. Sperschneider J, Gardiner DM, Dodds PN, Tini F, Covarelli L, Singh KB, Manners JM, Taylor JM: **EffectorP: predicting fungal effector proteins from secretomes using machine learning.** *New Phytol* 2016, **210:**743-761.
56. Guzman-Guzman P, Aleman-Duarte MI, Delaye L, Herrera-Estrella A, Olmedo-Monfil V: **Identification of effector-like proteins in Trichoderma spp. and role of a hydrophobin in the plant-fungus interaction and mycoparasitism.** *BMC Genet* 2017, **18:**16.
57. Raj A, Stephens M, Pritchard JK: **fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets.** *Genetics* 2014, **197:**573-589.
58. Leffler EM, Bullaughey K, Matute DR, Meyer WK, Segurel L, Venkat A, Andolfatto P, Przeworski M: **Revisiting an old riddle: what determines genetic diversity levels within species?** *PLoS Biol* 2012, **10:**e1001388.
59. Baranova MA, Logacheva MD, Penin AA, Seplyarskiy VB, Safonova YY, Naumenko SA, Klepikova AV, Gerasimov ES, Bazykin GA, James TY, Kondrashov AS: **Extraordinary Genetic Diversity in a Wood Decay Mushroom.** *Mol Biol Evol* 2015, **32:**2775-2783.
60. Swain E, Stukey J, McDonough V, Germann M, Liu Y, Sturley SL, Nickels JT, Jr.: **Yeast cells lacking the ARV1 gene harbor defects in sphingolipid metabolism. Complementation by human ARV1.** *J Biol Chem* 2002, **277:**36152-36160.
61. Olson A, Aerts A, Asiegbu F, Belbahri L, Bouzid O, Broberg A, Canback B, Coutinho PM, Cullen D, Dalman K, et al: **Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen.** *New Phytol* 2012, **194:**1001-1013.
62. Miyauchi S, Navarro D, Grisel S, Chevret D, Berrin JG, Rosso MN: **The integrative omics of white-rot fungus Pycnoporus coccineus reveals co-regulated CAZymes for orchestrated lignocellulose breakdown.** *PLoS One* 2017, **12:**e0175528.
63. Wendland J: *Growth, Differentiation and Sexuality.* 3 edn: Springer; 2016.

64. Hattori T, Abe Y, Usugi T: **Distribution of clones of *Phellinus noxius* in a windbreak on Ishigaki Island.** *European Journal of Forest Pathology* 1996, **26:**69-80.
65. Klopfenstein NB, Pitman EWI, Hanna JW, Cannon PG, Stewart JE, Sahashi N, Ota Y, Hattori T, Akiba M, Shuey L, et al: **A preliminary bioclimatic approach to predicting potential distribution of Phellinus noxious and geographical areas at risk from invasion.** In *Proceedings of the 63rd Annual Western International Forest Disease Work Conference*. 2016: 137-139.
66. Chin C-S, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, Dunn C, O'Malley R, Figueroa-Balderas R, Morales-Cruz A, et al: **Phased diploid genome assembly with single-molecule real-time sequencing.** *Nature Methods* 2016, **13:**1050-1054.
67. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, et al: **Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data.** *Nature Methods* 2013, **10:**563-569.
68. Lam KK, LaButti K, Khalak A, Tse D: **FinisherSC: a repeat-aware tool for upgrading de novo assembly using long reads.** *Bioinformatics* 2015, **31:**3207-3209.
69. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence data.** *Bioinformatics* 2014, **30:**2114-2120.
70. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al: **SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing.** *Journal of Computational Biology* 2012, **19:**455-477.
71. Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, Nusbaum C, Jaffe DB: **ALLPATHS: de novo assembly of whole-genome shotgun microreads.** *Genome Res* 2008, **18:**810-820.
72. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM: **Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement.** *PLoS ONE* 2014, **9:**e112963.
73. Wences AH, Schatz MC: **Metassembler: merging and optimizing de novo genome assemblies.** *Genome Biology* 2015, **16**.
74. Hunt M, Kikuchi T, Sanders M, Newbold C, Berriman M, Otto TD: **REAPR: a universal tool for genome assembly evaluation.** *Genome biology* 2013, **14:**R47.
75. Stanke M, Tzvetkova A, Morgenstern B: **AUGUSTUS at EGASP: using EST, protein and genomic alignments for improved gene prediction in the human genome.** *Genome biology* 2006, **7 Suppl 1:**S11.11-18.
76. Parra G, Bradnam K, Korf I: **CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes.** *Bioinformatics (Oxford, England)* 2007, **23:**1061-1067.
77. Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M: **BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS.** *Bioinformatics* 2016, **32:**767-769.
78. Borodovsky M, Lomsadze A: **Eukaryotic gene prediction using GeneMark.hmm-E and GeneMark-ES.** *Curr Protoc Bioinformatics* 2011, **Chapter 4:**Unit 4 6 1-10.

79. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics* 2005, **21:**3674-3676.
80. Lavezzo E, Falda M, Fontana P, Bianco L, Toppo S: **Enhancing protein function prediction with taxonomic constraints--The Argot2.5 web server.** *Methods* 2016, **93:**15-23.
81. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y: **dbCAN: a web resource for automated carbohydrate-active enzyme annotation.** *Nucleic Acids Res* 2012, **40:**W445-451.
82. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, et al: **antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters.** *Nucleic Acids Research* 2015, **43:**W237-W243.
83. Edgar RC: **Search and clustering orders of magnitude faster than BLAST.** *Bioinformatics* 2010, **26:**2460-2461.
84. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, et al: **The Pfam protein families database: towards a more sustainable future.** *Nucleic Acids Research* 2016, **44:**D279-D285.
85. Ellinghaus D, Kurtz S, Willhoeft U: **LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons.** *BMC Bioinformatics* 2008, **9:**18.
86. Emms DM, Kelly S: **OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy.** *Genome Biol* 2015, **16:**157.
87. Katoh K, Standley DM: **MAFFT: iterative refinement and additional methods.** *Methods Mol Biol* 2014, **1079:**131-146.
88. Stamatakis A: **RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models.** *Bioinformatics (Oxford, England)* 2006, **22:**2688-2690.
89. Felsenstein J: **PHYLIP (Phylogeny Inference Package) version 3.6.** *Distributed by the author Department of Genome Sciences, University of Washington, Seattle* 2005.
90. A A, J R: **topGO: Enrichment Analysis for Gene Ontology.** *R package version 2260* 2016.
91. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics* 2012, **29:**15-21.
92. Liao Y, Smyth GK, Shi W: **featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.** *Bioinformatics* 2014, **30:**923-930.
93. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.** *Genome Biology* 2014, **15**.
94. Young MD, Wakefield MJ, Smyth GK, Oshlack A: **Gene ontology analysis for RNA-seq: accounting for selection bias.** *Genome Biology* 2010, **11:**R14.
95. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al: **InterProScan 5: genome-scale protein function classification.** *Bioinformatics* 2014, **30:**1236-1240.

1177 96. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R,
1178 Salamov A, Zhao X, Korzeniewski F, et al: **MycoCosm portal: gearing up**
1179 **for 1000 fungal genomes.** *Nucleic Acids Research* 2014, **42:**D699-D704.
1180 97. Guy L, Roat Kultima J, Andersson SGE: **genoPlotR: comparative gene and**
1181 **genome visualization in R.** *Bioinformatics* 2010, **26:**2334-2335.
1182 98. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C,
1183 Salzberg SL: **Versatile and open software for comparing large genomes.**
1184 *Genome Biology* 2004, **5:**R12.
1185 99. Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R,
1186 Hardison R, Miller W: **PipMaker--a web server for aligning two genomic**
1187 **DNA sequences.** *Genome research* 2000, **10:**577--586.
1188 100. Maurer-Stroh S, Eisenhaber F: **Refinement and prediction of protein**
1189 **prenylation motifs.** *Genome Biology* 2005, **6:**R55.
1190 101. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G,
1191 Abecasis G, Durbin R: **The Sequence Alignment/Map format and**
1192 **SAMtools.** *Bioinformatics (Oxford, England)* 2009, **25:**2078-2079.
1193 102. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA,
1194 Mardis ER, Ding L, Wilson RK: **VarScan 2: Somatic mutation and copy**
1195 **number alteration discovery in cancer by exome sequencing.** *Genome*
1196 *Research* 2012, **22:**568-576.
1197 103. E G, G M: **Haplotype-based variant detection from short-read sequencing.**
1198 *arXiv* 2012, **1207:**3907.
1199 104. Price MN, Dehal PS, Arkin AP: **FastTree: Computing Large Minimum**
1200 **Evolution Trees with Profiles instead of a Distance Matrix.** *Molecular*
1201 *Biology and Evolution* 2009, **26:**1641-1650.
1202 105. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ: **Second-**
1203 **generation PLINK: rising to the challenge of larger and richer datasets.**
1204 *Gigascience* 2015, **4:**7.
1205 106. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA,
1206 Handsaker RE, Lunter G, Marth GT, Sherry ST, et al: **The variant call**
1207 **format and VCFtools.** *Bioinformatics* 2011, **27:**2156-2158.
1208 107. Vilella AJ, Blanco-Garcia A, Hutter S, Rozas J: **VariScan: Analysis of**
1209 **evolutionary patterns from large-scale DNA sequence polymorphism data.**
1210 *Bioinformatics* 2005, **21:**2791-2793.
1211

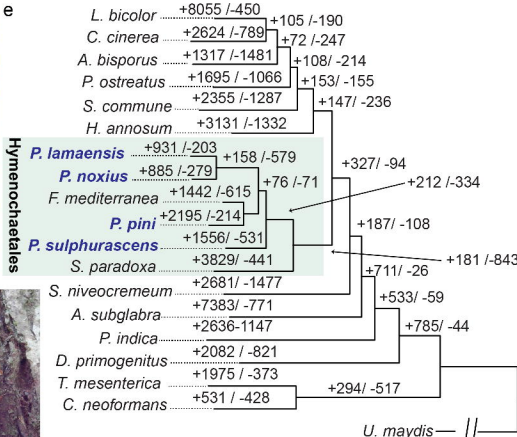

a

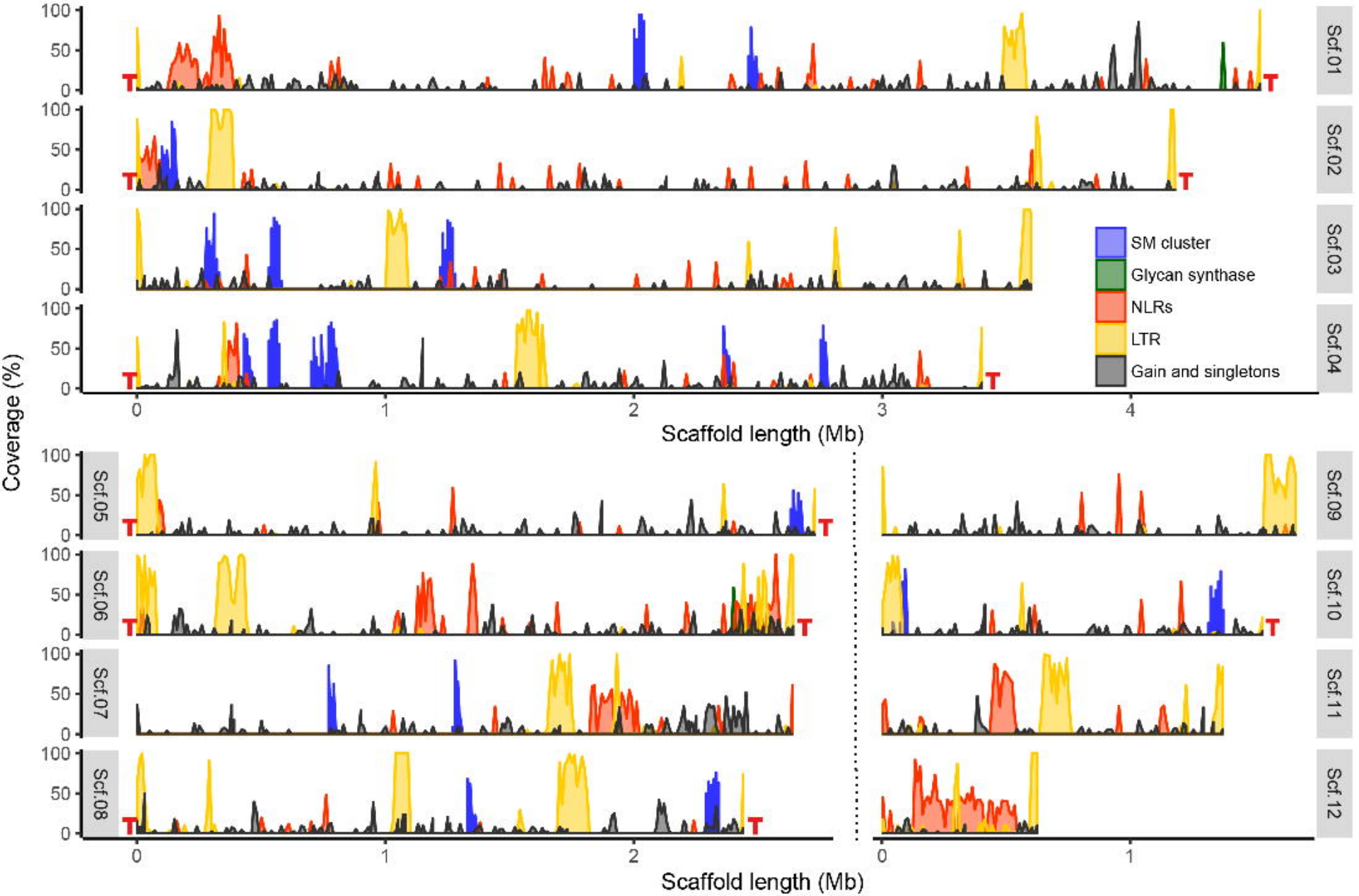

b

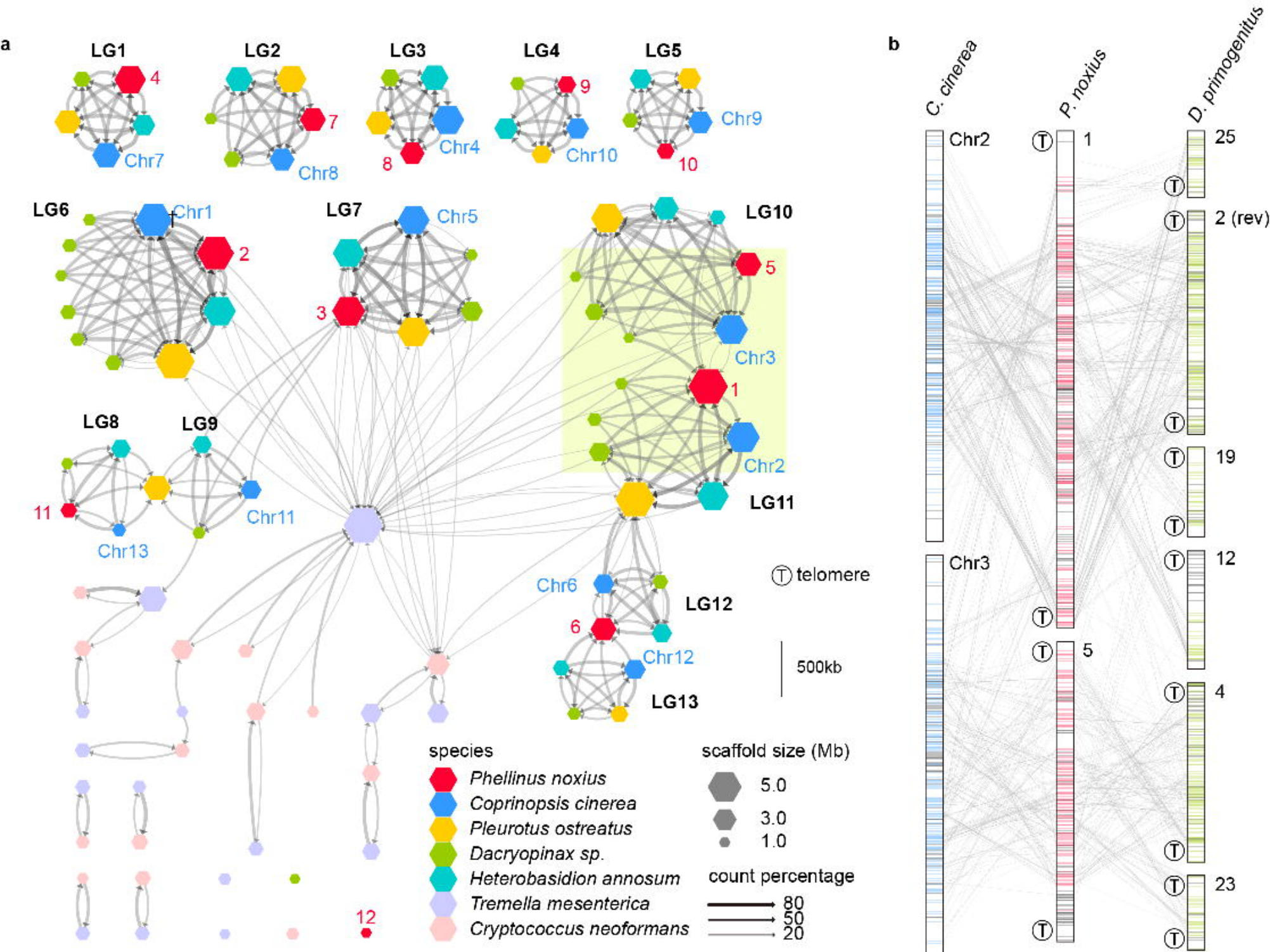

c

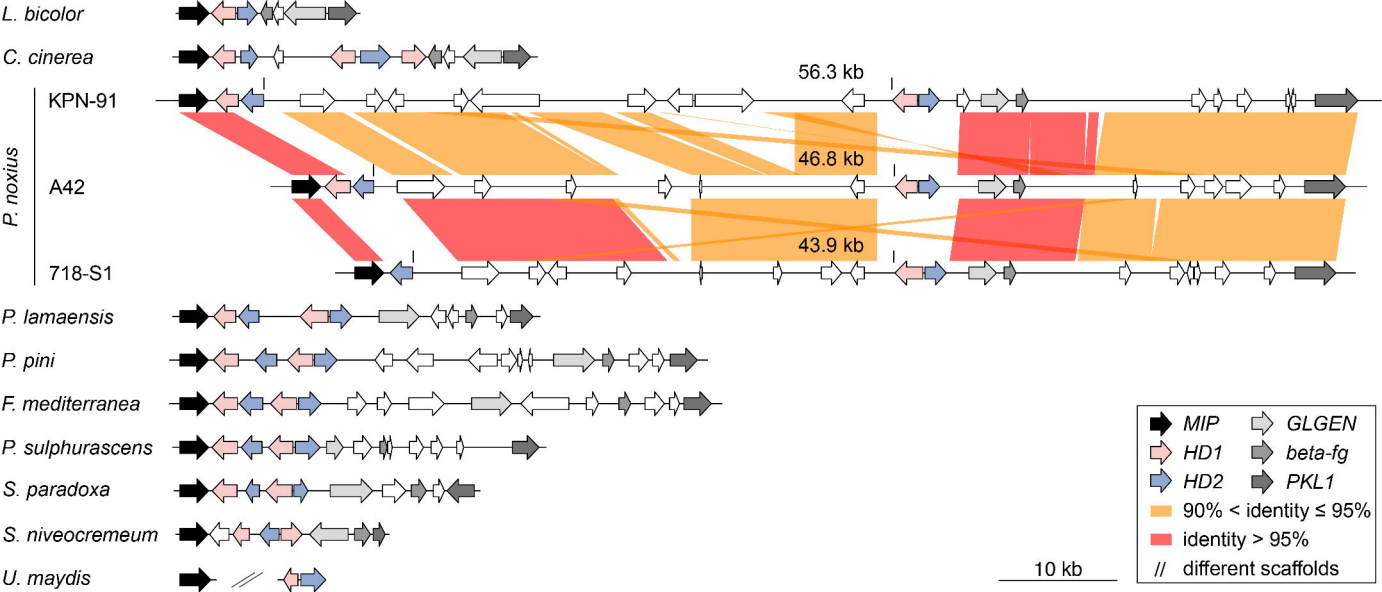

d

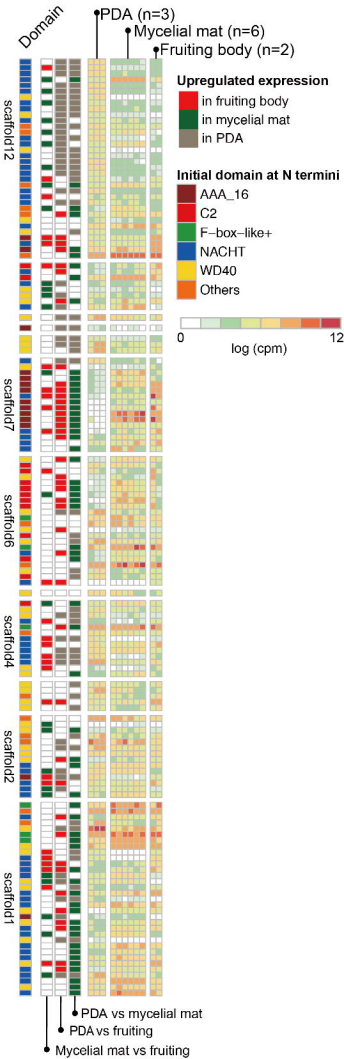
e

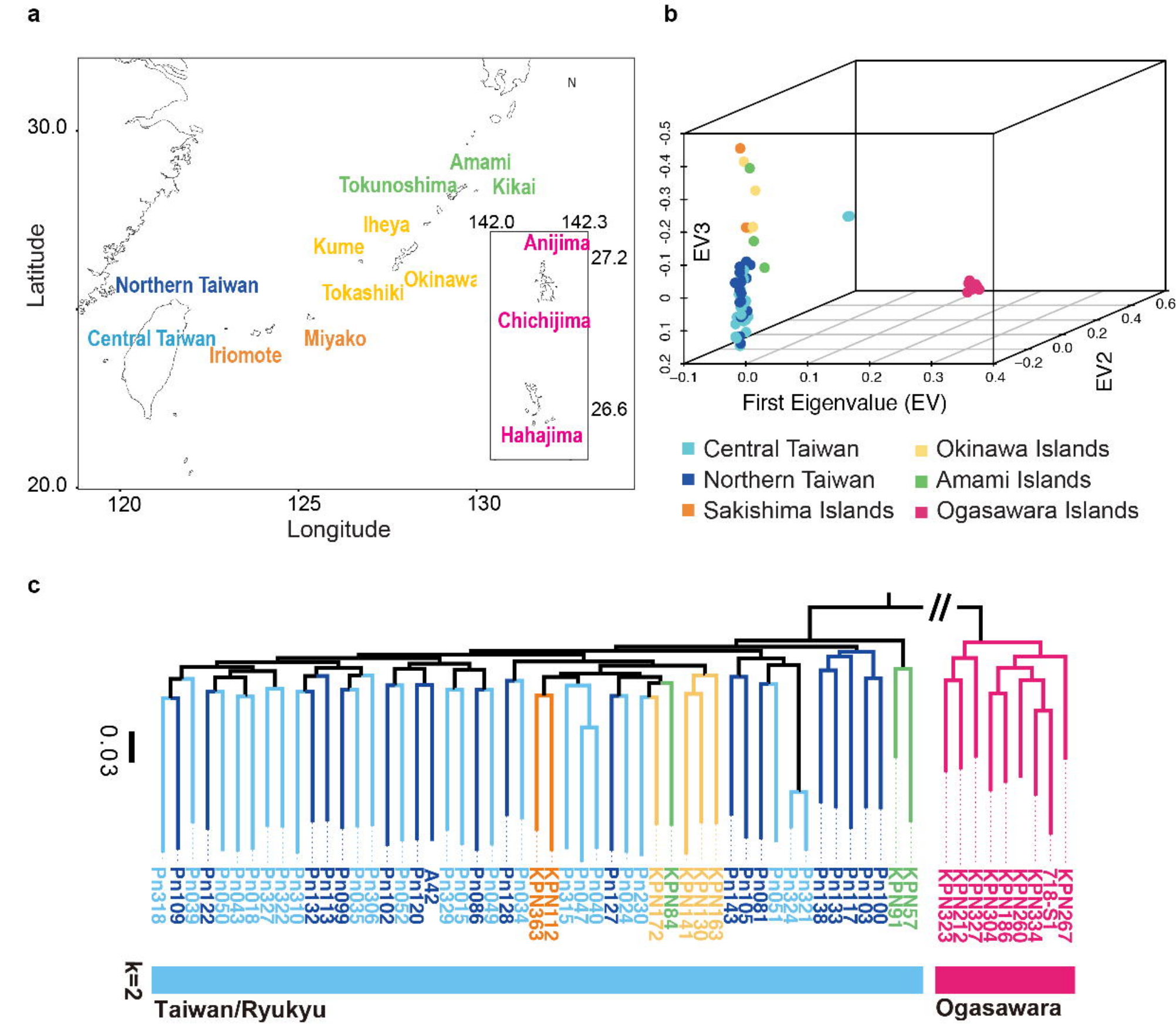

L. bicolor
+8055 /-450
+105 /-190
C. cinerea
+2624 /-789
+72 /-247
A. bisporus
+1317 /-1481
+108/ -214
P. ostreatus
+1695 / -1066
+153/ -155
S. commune
+2355 /-1287
+147/ -236
H. annosum
+3131 /-1332
Hymenochaetales
P. lamaensis
+931 /-203
+158 /-579
P. noxius
+885 /-279
+76 /-71
F. mediterranea
+1442 /-615
P. pini
+2195 /-214
P. sulphurascens
+1556/ -531
S. paradoxa
+3829/ -441
+327/ -94
+212 /-334
+187/ -108
+181 /-843
+711/ -26
S. niveocremeum
+2681/ -1477
A. subglabra
+7383/ -771
+533/ -59
P. indica
+2636-1147
+785/ -44
D. primogenitus
+2082 / -821
T. mesenterica
+1975 / -373
C. neoformans
+531 / -428
+294/ -517
U. maydis
0.3


Coverage (%)
100
50
0
T
Scf.01
Scf.02
Scf.03
Scf.04
SM cluster
Glycan synthase
NLRs
LTR
Gain and singletons
0
1
2
3
4
Scaffold length (Mb)
Scf.05
Scf.06
Scf.07
Scf.08
Scf.09
Scf.10
Scf.11
Scf.12
Scaffold length (Mb)


a
LG1
LG2
LG3
LG4
LG5
LG6
LG7
LG8
LG9
LG10
LG11
LG12
LG13
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
Chr13
telomere
500kb
species
Phellinus noxius
Coprinopsis cinerea
Pleurotus ostreatus
Dacryopinax sp.
Heterobasidion annosum
Tremella mesenterica
Cryptococcus neoformans
scaffold size (Mb)
5.0
3.0
1.0
count percentage
80
50
20
b
C. cinerea
P. noxius
D. primogenitus
Chr2
Chr3
1
5
25
2 (rev)
19
12
4
23


L. bicolor
C. cinerea
P. noxius
KPN-91
A42
718-S1
56.3 kb
46.8 kb
43.9 kb
P. lamaensis
P. pini
F. mediterranea
P. sulphurascens
S. paradoxa
S. niveocremeum
U. maydis
MIP
HD1
HD2
GLGEN
beta-fg
PKL1
90% < identity ≤ 95%
identity > 95%
// different scaffolds
10 kb


Domain
PDA (n=3)
Mycelial mat (n=6)
Fruiting body (n=2)
Upregulated expression
in fruiting body
in mycelial mat
in PDA
Initial domain at N termini
AAA_16
C2
F−box−like+
NACHT
WD40
Others
0
12
log (cpm)
scaffold12
scaffold7
scaffold6
scaffold4
scaffold2
scaffold1
PDA vs mycelial mat
PDA vs fruiting
Mycelial mat vs fruiting


a
b
c
Latitude
Longitude
30.0
20.0
120
125
130
142.0
142.3
27.2
26.6
N
Amami
Tokunoshima
Kikai
Iheya
Kume
Anijima
Okinawa
Tokashiki
Northern Taiwan
Chichijima
Central Taiwan
Miyako
Iriomote
Hahajima
EV3
EV2
First Eigenvalue (EV)
Central Taiwan
Northern Taiwan
Sakishima Islands
Okinawa Islands
Amami Islands
Ogasawara Islands
0.03
k=2
Taiwan/Ryukyu
Ogasawara

Population	Analyzed Sites	Segregating sites (S)	Singletons	θπ (x1,000)	θs (x1,000)	Tajima's D
Whole genome						
Ogasawara (n=9)	30,028,158	506,231	174,924	6.6	6.3	0.3
TaiwanRyukyu (n=51)	31,376,691	1,538,598	545,324	8.0	10.9	-0.9
TaiwanRyukyu subset (n=9)[1]	30,640,276	749,681	408,021	8.1	8.9	-0.5
Synonymous Sites						
Ogasawara (n=9)	4,651,751	182,879	59,464	15.8	14.7	0.4
TaiwanRyukyu (n=51)	4,758,211	513,717	161,353	19.2	23.6	-0.7
TaiwanRyukyu subset (n=9)[1]	4,704,161	269,060	136,780	19.3	20.5	-0.4
Replacement Sites						
Ogasawara (n=9)	9,794,069	71,497	25,981	2.9	2.8	0.2
TaiwanRyukyu (n=51)	10,018,114	230,689	93,385	3.6	5.3	-1.2
TaiwanRyukyu subset (n=9)[1]	9,904,211	103,907	60,313	3.6	4.0	-0.6

[1] A random of nine isolates were chosen to repeat the analysis.