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°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.

138 **RESULTS**

139

140 **Genome structure and content**

141 We produced a 31.6 Mb *P. noxius* genome reference assembly from a Japanese
142 KPN91 isolate combining both Pacbio and Illumina technologies. The nuclear
143 genome of *P. noxius* is assembled into 12 scaffolds with six assembled into
144 chromosomes from telomere to telomere **(Figure 2)**, while the mitochondrial genome
145 is assembled into a single sequence of 163.4 kb. To date chromosome-level
146 assemblies are available only for a few basidiomycetous species[22-24] including *P.*
147 *noxius* (Supplementary Table 1). For a comprehensive understanding of genome
148 evolution amongst members of the hymenochaetoid clade, we also sequenced and
149 assembled three additional species: *P. sulphurascens, P. lamaensis* and *P. pini*, as
150 well as two more isolates (A42 and 718-S1) of *P. noxius*. The three assemblies of *P.*
151 *noxius* have N50s of 2.4-2.7 Mb, whilst the other three genome assemblies comprise
152 30.7-53.3 Mb with N50 of 570 kb-2.7 Mb. Together, the genome assembly quality of
153 these Hymenochaetaceae are among the most contiguous and complete of the
154 sequenced basidiomycetes.

155

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

171 *mediterranea* show an expansion of their genome sizes to 53-63 Mb and are also
172 grouped together.
173
174 **Analysis of transposable elements and evidence of repeat-induced point mutation**
175 **(RIP)**
176 A combination of *de novo* and reference repeat detection classified a total of 53-297
177 repeat families (Supplementary Table 2). The transposable element (TE) content
178 makes up 1.4-27.5% of the assemblies in the Hymenochaetales order (Supplementary
179 Table 3) and is associated with their genome size and taxonomic relationship. The
180 Class I transposons make up the majority of repetitive elements in these fungal
181 genomes (Supplementary Table 3). Manual annotation of the 16 most abundant full-
182 length transposon sequences revealed that these TEs belong to the *Gypsy* and *Copia*
183 superfamilies of the LTR retrotransposons, and the *Tad1* clade from the LINE
184 retrotransposons (Supplementary Figure 1). The *Gypsy* elements are 5,295 to 7,629 bp
185 in length and contain an open reading frame encoding 1,523 amino acid residues,
186 which contain conserved domains in the order of protease, reverse transcriptase,
187 RNase H, integrase and chromodomain (Supplementary Figure 1). Chromodomain-
188 containing *Gypsy* or chromoviruses are widely distributed in eukaryotes and present
189 in many basidiomycetes[27]. In the case of *P. pini,* 588 full-length copies of two
190 repeats from the *Gypsy* family (Supplementary Figure 1) make up 10% of the total
191 assembly. In contrast, a clear exception is *P. sulphurascens* where a *Tad1* class LINE
192 dominates the TE content of the genome and a *Copia* class of LTR-retrotransposon is
193 found to be more abundant than *Gypsy*.
194
195 Copies of TEs are not distributed uniformly across the genome: the six complete
196 chromosomes in the *P. noxius* assembly show clear signatures of Basiodiomycete
197 centromeres (located at the middle of chromosome and comprised solely of TEs) and
198 enriched at subtelomeres; TEs are distributed as distinct 80.4-118.4 kb blocks (**Figure**
199 **2**), which are typical in fungi[22]. These clusters also show a much lower average GC
200 content suggesting a hallmark of fungal repeat-induced point mutation*[28]* (RIP)
201 favouring T or A changes (Supplementary Figure 2). In Basidiomycota such
202 mutational patterns were previously only observed in the Pucciniomycotina
203 subphylum[29]. However, homologs of known genes responsible for RIP are present
204 in all four Hymenochaetales species (Supplementary Table 4). Analysis of

205 dinucleotide transitions[30] has shown significantly higher CpN→TpN mutations, but
206 not vice versa (Supplementary Figure 3), suggesting that RIP have successfully
207 limited TE expansions especially in the compact genomes of *P. noxius* and *P.*
208 *lamaensis*. Interestingly, each species has a different preferred mode of dinucleotide
209 changes: CpA→TpA mutations clearly dominate in *P. noxius* compared to CpT→TpT
210 and CpG→TpG in both *P. pini* and *P. sulphurascens.*
211
212
213 **Chromosome biology**
214 Like other complete chromosomes of basidiomycetes (Supplementary Table 1), the *P.*
215 *noxius* genome also presents little variability in gene density along its scaffolds with
216 the exception of TE clusters (Supplementary Figure 4). The number of known
217 karyotypes in Basidiomycota ranges from 11 to 14 chromosome pairs (Supplementary
218 Table 1), which suggests a possible ancestor with similar chromosome numbers. In an
219 attempt to characterize chromosome architecture and evolution amongst the
220 Agaricomycotina sub-division, we constructed a linkage network of seven selected
221 species based on single-copy orthologs between species pairs. We identified 13
222 distinct linkage groups (LG) providing strong evidence that chromosome macro-
223 synteny is largely conserved since the common ancestor of Agaricomycetes (**Figure**
224 **3a**; Supplementary Figure 5). Such a relationship is extended to Dacrymycetes where
225 multiple scaffolds can be predominantly assigned to different linkage groups, but it is
226 no longer apparent when compared to Tremellomycetes. Certain scaffolds are found
227 to connect different linkage groups, implying inter-chromosomal rearrangements. For
228 example, *P. noxius* scaffold1 is strongly clustered in LG11 but also shows linkage to
229 LG10, implying a translocation from scaffold5 (**Figure 3**). Within each linkage group,
230 gene collinearity is no longer apparent, suggesting high levels of intra-chromosomal
231 rearrangements which have been observed in different fungal groups[31] (**Figure 3b**).
232
233 Like most species, the most common telomere repeating unit in basidiomycetes is
234 TTAGGG[32], although other combinations such as T(n)AG(n)[33] or
235 TTTAGGGG[34] have been described elsewhere. A systematic survey of the selected
236 published genomes reveals the predominance of TTAGGG repeat across
237 basidiomycetes, including *P. pini* and *P. sulphurascens* (Supplementary Table 5). The

238 karyotype of *P. noxius* has not been reported, but together with *P. lamaensis* we found
239 10 scaffolds that end with overrepresented 13.6-66.9 copies of a TTAGGTG heptamer,
240 suggesting a novel telomeric repeat compared to other basidiomycetes[32].
241 Interestingly, the *Dacryopinax primogenitus* assembly contains 31 scaffold ends with
242 TTAGGG telomeric repeats. Of those, 11 scaffolds have telomeric repeats at both
243 ends revealing extensive fission of at least 21 chromosomes. **Figure 3b** shows one
244 example of six scaffolds with such repeats present on the ends in *D. primogenitus*
245 compared to two scaffolds in both *P. noxius* and *Coprinopsis cinerea.*
246
247 **Gene families associated with Hymenochaetales and *Phellinus noxius***
248 We sought to identify genes and protein domains specific to Hymenochaetales by
249 determining when a new gene family arose and if the family has expanded or
250 contracted. In total, 7,125-11,659 proteins in the Hymenochaetales order are clustered
251 together in 5,184 families. Acquisition of gene families was mainly found at the tips
252 of the phylogeny (531-8,055 families) suggesting each species has a repertoire of
253 specific genes. The seven Hymenochaetaceae species have a total of 62 enriched
254 protein domains compared to other basidiomycetes (Supplementary Figure 6). Gain of
255 domains such as fungal specific transcription factors (Fungal_trans; 53.8 *vs* 41.6
256 copies) and peroxidase associated protein domains (Peroxidase_ext; 16.7 *vs* 3.9
257 copies) in these Hymenochaetaceae species reflect their white rot nature. Conversely,
258 579 gene families were inferred to have been lost in the common ancestor of *P.*
259 *lamaensis* and *P. noxius*. Analysis of gene ontology (GO) term enrichment on these
260 otherwise preserved gene families in *P. sulphurascens* and *P. pini* revealed a common
261 theme of intracellular signal transduction (Supplementary Table 6). Several signal
262 transduction pathways may have become more specialized in the ancestor of *P. noxius*
263 and *P. lamaensis*. For example, multiple copies of NmrA (nitrogen metabolite
264 repression) were lost in these two species, which act as regulators in nitrogen-
265 dependent control of secondary metabolism[35].
266
267 We sought to identify novel and expanded gene families in *P. noxius* despite its
268 compact genome (Supplementary Figure 7). In total, 74% (7,313) of 9,833 predicted
269 genes from *P. noxius* have orthologs from at least one of the other basidiomycetes,
270 suggesting most of the basidiomycetous core genes are conserved. Consistent with the
271 fact that *P. noxius* is a fast grower[5], genes encoding 1,3-beta-glucan synthase

272 responsible for the formation of beta-glucan components of the fungal cell wall[36]
273 are expanded (14 compared to an average of 1.8 copies). Notably, counts of WD40
274 protein domains are highest in *P. noxius* and *P. lamaensis* despite their small genome
275 size (Supplementary Table 7a). This domain is important in protein-protein
276 interactions of cellular networks and is usually associated with additional domains
277 with other functional activities[37]. Interrogating this expansion revealed the
278 association of WD40 domains with the AAA and NACHT domains (Supplementary
279 Table 7b), both of which are NTPase domains and such combinations are commonly
280 found in nucleotide-binding-oligomerization-domain like receptors (NLRs). This gene
281 family is highly diverse in fungi and was found to be central to the process of
282 programmed cell death and implicated in fungal vegetative incompatibility and
283 general nonself recognition [38]. A maximum likelihood phylogeny of the NACHT
284 domain proteins shows different expansions of NLR subfamilies in fungi
285 (Supplementary Figure 8). In particular, the C2-NACHT-WD40(n) subfamily has
286 only been found exclusively in a few basidiomycetes[39] and is present in the highest
287 copy number in *P. noxius* and *P. lamaensis*. The strikingly high abundances of WD40
288 domains in *P. noxius* and *P. lamaensis* are a result of tandem expansions near
289 chromosome ends (**Figure 2**). Other expansions include UDP-
290 glucuronosyltransferases, which catalyse conjugation reactions in the
291 biotransformation of xenobiotics released by its host environment [40].
292
293
294 ***P. noxius* displays a bipolar heterothallic mating system**
295 Mate recognition of sexual reproduction in Basidiomycota is known to be controlled
296 by two unlinked loci, named as *A* and *B* locus. Different from the divergently
297 transcribed *HD1-HD2* observed in most species in Agaricomycetes[41], a conserved
298 head-to-tail orientation of *HD1-HD2* in *HD* pair 1 was found in Hymenochaetales
299 **(Figure 4)**. Alignment of 100-kb sequences upstream and downstream of *A* locus in *P.*
300 *noxius* isolates KPN91, 718-S1 and A42 revealed that *A* locus is highly polymorphic
301 (*HD* pairs in particular) despite well-conserved flanking regions (Supplementary
302 Figure 9). For *B* locus, only one *STE3* encoding seven transmembrane helices was
303 identified; four pheromone genes were identified but not physically linked to the
304 highly monomorphic *STE3* (Supplementary Figure 10, Supplementary Table 8,
305 Supplementary Information), which is characteristic of a bipolar mating system.

306 To further confirm this observation, allele diversity was analyzed by resequencing *A*
307 and *B* locus from 10 single-basidiospore isolates originating from a single basidiocarp
308 (Supplementary Table 9). Sequencing of *STE3* revealed two highly similar alleles (b1,
309 b2), with 99.5% amino acid identity. The only variant (244V/A) is considered a semi-
310 conservative mutation (valine and alanine are nonpolar aliphatic acids) and may have
311 minor or no effect on protein function. Previous studies have shown that although
312 STE3 is not involved in mating type determination in bipolar fungi, variations can
313 still be observed[42-44]. Primer walking of *A* locus revealed two distinct haplotypes
314 (Supplementary Figure 11): the a1 and a2 alleles of *HD* pair I shared an overall 78%
315 nucleotide identity; the two alleles of *HD* pair II were highly diverse and the HD1
316 domain in a2 allele contains a 1-bp and a 9-bp deletions and has become a
317 pseudogene. The loss of HD1 domain was also found in the *HD* pair I of 718-S1
318 (Figure 4), suggesting that at least one of the HD1 motifs in *A* locus is dispensable for
319 a functional HD1-HD2 heterodimer in *P. noxius*. The presence or absence of specific
320 HD domains reflects phylogenetic characteristics and has been commonly observed in
321 fungi[45]. Clear segregation of two distinct alleles at *A* locus (Supplementary Figure
322 11) suggests a heterothallic bipolar reproductive mode in *P. noxius*.
323
324 Distinct from all the other fungi, *P. noxius* has an exceptionally highly expanded *A*
325 locus across a ~60-kb region **(Figure 4)**. There are two pairs of HD1-HD2 gene
326 located ~50-kb away from each other. The characteristic large interval between *HD*
327 pairs has only been reported in *Schizophyllum commune* (~55 kb) and *Flammulina*
328 *velutipes* (~70 kb), in which the genomic separation likely emerged through
329 inversions or transpositions of gene clusters surrounding *HD*[46]. In the case of *P.*
330 *noxius,* the large interval is a result of accumulation of genes transposed from
331 different parts of the genome (**Figure 4**; Supplementary Figure 12).
332
333
334 **Distribution of CAZymes and secondary metabolic gene clusters**
335 Carbohydrate-active enzymes (CAZymes) are a diverse set of enzymes involved in
336 the synthesis and degradation of carbohydrates[47]. Efficient degradation of lignin, a
337 tough biopolymer present in woody plants, by white rot fungi such as members of
338 Hymenochaetales requires an arsenal of ligninolytic enzymes[48]. Consistent with
339 this, CAZymes spanning eight families of lignin-degrading enzymes (AA1-AA8)

340 which include laccases, peroxidases, oxidases, and reductases, and a lytic
341 polysaccharide mono-oxygenase family (LPMO; AA9) were found in all
342 Hymenochaetales species (Supplementary Table 10). Specialization includes the
343 identification of copper-dependent LPMO AA11 and oxidoreductase from AA12
344 family only in *P. lamaensis* and *P. sulphurascens*, respectively. The most abundant
345 and diverse among the CAZymes annotated in these four species were glycoside
346 hydrolases (GHs). GHs act on diverse carbohydrates, including cellulose (e.g. GH1,
347 GH3, GH5, GH6, GH7 and GH12), hemicellulose (e.g. GH5, CE1, GH3 and GH27),
348 pectin (e.g. GH28, GH43, GH78, CE8, CE12, PL1 and PL3), starch (GH13, GH15
349 and GH31), and chitin (e.g. GH18, GH20, GT2 and CBM5) (Supplementary Table
350 10). Despite its compact genome, *P. noxius* contains a comprehensive repertoire of
351 carbohydrate-active enzymes; a total of 4.23% (416 proteins) of its proteome were
352 identified as CAZymes which is more than those in the other Hymenochaetales
353 (Supplementary Figure 13). Taken together, the diversity of the CAZymes encoded in
354 the *P. noxius* genome is consistent with its ability to infect a wide range of hosts.
355
356 The Basidiomycota phylum is known for the production of huge and diverse sets of
357 secondary metabolites (SM) such as terpenoid compounds which have roles in
358 virulence, survival, and host specificity[49]. Investigations on SM gene clusters
359 revealed that nonribosomal peptide synthase (NRPS), type 1 polyketide synthase
360 (T1PKS), and terpene synthase (TS) clusters were common among the proteomes of
361 Hymenochaetales. For *P. noxius*, a total of 18 putative clusters were found, which
362 included one lantipeptide, one NRPS, two T1PKS, and 11 TS clusters. These SM
363 clusters contained 4-37 annotated biosynthesis genes (Supplementary Table 11),
364 distributed evenly in eight of the 12 scaffolds of *P. noxius* (**Figure 2**). Clusters of
365 lantipeptides were found in only three out of 19 compared fungi, including *P. noxius*
366 and its closest relative *P. lamaensis.* Lantipeptides are primarily known for their
367 antimicrobial activities, although some have other roles such as in fungal
368 morphegenesis[19].
369
370
371
372
373

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

408 upregulated in mycelial mats. This is consistent with the fact that *P. noxius* as a white
409 rot has the capacity for continuously degrading plant materials including lignin and
410 responding to xenobiotics elicited by plant defence response[53] (Supplementary
411 Table 12). Import pathways such as Major Facilitator Superfamily (MFS-1)
412 transporters was *in silico* predicted to have a role in transporting intermediate
413 decomposition products during wood decay [54]. This is also true for *P. noxius,*
414 where 45% (51) out of 113 members of the MFS-1 transporter were upregulated in at
415 least one of the field conditions. Only 10% of the *P. noxius* secretome were predicted
416 to be putative effectors[55], which are significantly upregulated in *in planta*
417 conditions (Supplementary Table 13). The upregulated gene list also contains
418 hydrophobins (e.g. cerato-platanin), which contributed to virulence of various
419 fungi[56] and were upregulated in previous inoculation experiments of *P.*
420 *sulphurascens*[3].
421
422 The creamy and brown mycelial mat on infected trees are physically in close
423 proximity and reflects relatively early and middle stages of pathogenesis rather than
424 the late stage of vigorous decomposition. Their global transcriptome profiles were
425 more similar to each other when compared with the other two conditions. Enrichment
426 analysis of differentially expressed genes shows different strategies used by *P. noxius*
427 (Supplementary Table 14): the creamy mycelium shows enrichment of upregulated
428 genes involved in DNA replication such as ATPases and DNA topoisomerase,
429 suggesting *P. noxius* is actively progressing outwards along the tree bark
430 (Supplementary Figure 15). In contrast, members of P450 and CAZymes are
431 significantly upregulated in the dark mycelial mat, suggesting an even higher
432 expression of these genes during the middle stage of colonization.
433
434
435 **Sequence analysis of 60 *P. noxius* strains**
436 To further understand the regional dissemination of *P. noxius* in Taiwan and Japanese
437 offshore islands, we sequenced the genomes of 60 isolates (Figure 6a) originating
438 from diseased trees in urban cities as well as trails in natural parks. This collection
439 was sequenced to a median depth of coverage of 35X, and additional mate pair
440 libraries were sequenced in two isolates (718-S1 and A42) which were used to
441 characterize synteny relationship amongst strains. An average of 96% mapping rate

442 was achieved after aligning reads from each strain to the KPN91 reference genome.
443 The descriptions and statistics of the strains are provided in Supplementary Table 15.
444
445
446 **Stable structural variation and frequencies of heterokaryons in *P. noxius***
447 The detection of frequency of heterokaryons and polyploids of *P. noxius* in nature
448 remain challenging as its arthospores contain 1-5 nuclei (based on quantification of
449 145 arthospores from A42 in this study) and multiple allelic fragments in SSRs (2-4
450 alleles per locus) are commonly found in populations [4, 6]. All the strains were
451 isolated from either a single arthospore or fungal mat, and thus allowed for the
452 analysis of variation in genome structure. We found no deviation in coverage across
453 every scaffold in all strains (Supplementary Figure 16), suggesting a stable number of
454 chromosome copies in *P. noxius.* To distinguish the frequency of mono- or
455 heterokaryons in nature, heterozygous SNPs and minor allele frequency (MAF)
456 distribution was inferred in each strain (Supplementary Figure 17). We identified four
457 groups (A to D) that clearly differ in heterozygosity and MAF. The group with the
458 lowest heterozygosity of averaging 0.2 % showed a random MAF distribution. In
459 particular, strains 718-S1 and A42 each isolated from single basidiospores fall in this
460 group suggesting they are monokaryotic in nature. Conversely, the other three groups
461 have a much higher heterozygosity, with the highest group of 1.6% heterozygosity
462 displaying a peak around 50% in MAF distribution. Peaks around 27%-33% are also
463 observed in two of the groups, and this is not due to multiple nuclei as the
464 distributions of the number of nuclei per cell in seven randomly selected isolates are
465 the same (Supplementary Figure 18). Groups B-D may refer to different compositions
466 of two or more genetically distinct nuclei. The number of monokaryons present in
467 nature is 40% and can be found across all regions surveyed in this study.
468
469
470 **Diversification of *P. noxius* across Pacific Ocean islands**
471 To examine the population structure of *P. noxius* in this area of Pacific Islands, we
472 performed principal-component analysis (Figure 6b; Supplementary Figure 19) on the
473 SNP variants from individual phased haploid genomes. Similar to previous report
474 based on SSRs on highly diversified populations between Japanese offshore
475 islands[4], the major principal component divided Ogasawara islands samples from

476 the rest of the samples, which are located geographically 1,210 km apart. The second
477 and third principal components defined a tight cluster of Taiwanese isolates with
478 considerable overlaps between cities despite being 170 km apart. Most isolates from
479 Ryukyu islands can be differentiated from those from Taiwan. Interestingly, we
480 identified three isolates from Taiwan that are distinctive from both the Taiwan-
481 Ryukyu and Ogasawara island clusters, suggesting a possibility of more populations
482 within the Taiwan island.
483
484 We constructed a maximum likelihood phylogeny based on 1,837,281 high
485 confidence SNPs of *P. noxius* (Figure 6c). Concordant with the PCA, the strains from
486 Ogasawara islands form a distinctive lineage. We also observed that the majority of
487 Taiwan and Ryukyu islands strains are grouped together forming another major
488 lineage, suggesting the possibility of gene flow in this region despite physical
489 separations by the sea. We inferred the population structure in 144,426 unlinked
490 bialleic sites using the Bayesian model-based clustering approach[57]. Consistent
491 with the phylogeny, fastSTRUCTURE[57] also identified two Taiwan-Ryukyu and
492 Ogasawara lineages circulating in this region (Figure 6c, Supplementary Figure 20).
493
494
495 **High nucleotide diversities in *P. noxius* populations**
496 The *de novo* assemblies of three strains of *P. noxius* have on average 97% nucleotide
497 identity and are largely co-linear to each other with apparent genome translocations
498 between Taiwan/Ryukyu (A42 and KPN91) and the Ogasawara (718-S1) isolate
499 (Supplementary Figure 21). We quantified sequence diversity using θs and θπ and
500 categorised them into whole genome average, four-fold synonymous sites and
501 replacement sites (Table 1). An large number of SNPs are found in the genome of *P.*
502 *noxius*, averaging one SNP identified in every 20-59 bp. Nucleotide diversity at
503 synonymous sites θπ-syn is 15.8 and 19.2 per kilobase in Ogasawara island and
504 Taiwan-Ryukyu populations, respectively. This is ~5 fold higher when compared to
505 the majority of species[58] and is likely an underestimate of true diversity as indels
506 were not considered. Extremely high diversity has been reported in natural
507 populations of *Schizophyllum commune*[59]. The Tajima's D is strongly negative in
508 the Taiwan and Ryukyu islands lineage, suggesting an excess of low frequency alleles
509 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.

611 **FIGURES AND TABLES**

612

613 **Figure 1. Life stages of *Phellinus noxius* and comparative genomic analysis of**
614 **Hymenochaetales species.** (**a**) The mycelial mat with young creamy leading front
615 and aged brown section. (**b**) In advanced stage of decay, the hyphae form a network
616 of brown zone lines permeating the soft and white wood tissue. (Lower left and lower
617 right) (**c and d**) Basidiocarps are perennial and can be resupinate (**c**) or grow into a
618 sessile bracket-like conk with a broad basal attachment (**d**). The distinctive greyish-
619 brown surface is the hymenial layer with irregularly polygonal pores, containing four-
620 spored basidia, ellipsoid and hyaline basidiospores, but no hymenial setae. (**e**) The
621 phylogeny of four *Phellinus* species with 15 other species of Basidiomycota based on
622 a concatenated alignment of single-copy orthologous genes. All nodes have 100 out of
623 100 bootstrap replicates. The numbers of gained ("+") and lost ("-") gene families
624 along each branch of the phylogeny is annotated.

625

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

629

630 **Figure 3. Linkage group (LG) network of Basidiomycota.** (**a**) The linkage groups
631 were identified by linking single-copy orthologs of scaffolds between species pairs.
632 All scaffolds included in this plot are larger than 500kb. Each directional edge points
633 toward the reference chromosome. Edges were weighted by the number of single-
634 copy orthologs, and an edge was filtered if it has a weight smaller than 20 or less than
635 40% of the sum of all weights out of a node. (**b**) Cross-mapping of single-copy
636 orthologs in LG10 and 11. Scaffold names are shown at the upper-right side of each
637 sequence, with detectable telomeric regions labelled as 'T' at the upper-left side.

638

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

644

645 **Figure 5. WD40 domain containing genes upregulated in the transcriptome of**
646 ***Phellinus noxius* at different conditions.** RNA-seq counts of these genes (in log of
647 FPM, fragments per million) for *P. noxius* cultured in PDA, as mycelial mat and
648 fruiting body on infected trees are presented as different colours in a heatmap sorted
649 according to their physical positions on the assembly. The results of differentially
650 expressed genes from pairwise DEseq2 comparisons (padj < 0.01) are shown in
651 different colours each corresponding to a specific life stage of *P. noxius*. Different
652 combinations of protein domains are categorised based on their initial domain from
653 the N termini. Genes with combination that is shared no more than four other genes
654 are categorised into 'Others' (3 of ANAPC4, 1 of BOP1NT, 1 of CAF1C_H4-bd, 5 of
655 Pkinase, 1 of PRP4, 1 of Striatin, 1 of Tup_N).
656
657 **Figure 6. The population genomics of 60 *Phellinus noxius* isolates.** (**a**) Map of
658 Taiwan and offshore islands of Japan showing origins of the 60 sampled *P. noxius*
659 isolates. Ogasawara islands were conveniently drawn below Ryukyu island so this
660 does not represent their actual location. (**b**) PCA plot of 60 *P. noxius* isolates using
661 genome-wide variation data sampled from 13 islands by the first three eigenvectors.
662 (**c**) Top: Phylogenetic tree with 100 bootstrap using SNPs computed from alignment
663 to KPN91 reference. Point separating the Taiwan-Ryukyu and Ogasawara island
664 isolates was used as root. Nodes with >90% bootstrap were labelled with circles.
665 Bottom: fastSTRUCTURE analysis of the linkage independent pruned set of variation
666 data. A model with two ancestral components (K=2) had the highest likelihood to
667 explain the variation of genome-wide structure on the 60 isolates. Also see
668 Supplementary Figure 20 for different K.
669
670 **Table 1.** Polymorphism in the two regional lineages of *Phellinus noxius.*
671
672
673
674
675
676
677
678

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

713 against the KAPA Library Quantification Kit (KAPA Biosystems) for Illumina
714 platform. The transcriptome libraries were pooled at equal molar concentrations, and
715 PE2*151nt multiplexed sequencing was conducted on HiSeq 2500 sequencer.
716
717 **Nuclear quantification**
718 Nuclei in the growing hyphal tips were stained following the procedure described by
719 Chung et al. (2015). The mycelium on the slide was mounted in 20 µl of a DAPI (4',6-
720 diamidino-2-phenylindole) solution (10 µg/ml in ddH_2O) for an hour, destained in
721 ddH_2O for 30 min, then observed under a OLYMPUS BX41 microscope (Shinjuku-
722 ku, Tokyo, Japan) equipped with filter cube U-MWU2 (BP 330–385 nm, LP 420 nm).
723 Images were captured by using a Canon (Ohta-ku, Tokyo, Japan) digital camera EOS
724 700D. One hundred hyphal cells per strain were counted.
725
726 **Genome assembly**
727 Genome assembly of different species was carried out with Falcon[66] (ver 0.5.0) and
728 were improved using Quiver[67] and finisherSC[68]. For assembly of individual
729 strains of *P. noxius*, Illumina paired end reads were trimmed with Trimmomatic[69]
730 (ver 0.32; options LEADING:30 TRAILING:30 SLIDINGWINDOW:4:30
731 MINLEN:50) and subsequently assembled using SPAdes[70] (ver 3.7.1). Multiple
732 mate-pair reads were available for three strains of *P. noxius* (KPN91, A42 and 718-S1)
733 and they were assembled using ALLPATH-LG[71] assembler (ver 49688) and
734 improved using Pilon[72]. The *P. noxius* assembly was further merged with
735 metassembler[73], miss-assemblies were identified using REAPR[74] and manually
736 corrected.
737
738 **Gene predictions and functional annotation**
739 For *P. noxius,* the gene predictor Augustus[75] (v 3.2.1) was trained on a gene
740 training set of complete core genes from CEGMA[76] and subsequently used for
741 manual curation of ~1000 genes. Annotation was then run by providing introns as
742 evidence from RNA-seq data. For *P. lamaensis*, *P. sulphurascens* and *P. pini*, genes
743 were predicted using Braker1[77] pipeline that automatically use RNA-seq mappings
744 as evidence hints and retraining of GeneMark-ES[78] and Augustus. Gene product
745 description was assigned using blast2go[79] (ver 4.0.7) and GO term assignment were
746 provided by ARGOT2.5[80]. The web server dbCAN[81] (HMMs 5.0, last accessed

747 September 5 2016) was used to predict CAZymes from the protein sequences of all
748 species, while AntiSMASH[82] (version 3.0) was used to predict secondary
749 metabolite gene clusters. For dbCAN results, only hits with <= 1 x 10e-5 e-value, and
750 >= 30% HMM coverage were considered, while overlapping domains were resolved
751 by choosing hits with the smallest *P* value.
752
753 **Repeat analysis and repeat induced mutation analysis**
754 Consensus sequences of repetitive elements in the assembly of the four
755 Hymenochaetales species were identified based on UCLUST[83] clustering of 80%
756 sequence identity from two repeat detection programs: RepeatModeler
757 (www.repeatmasker.org) and TransposonPSI (http://transposonpsi.sourceforge.net/).
758 Annotations of these sequences were based on output of both programs and BLAST
759 results from the National Center for Biotechnology Information (NCBI) nonredundant
760 database. RepeatMasker (v3.2.8, http://www. repeatmasker.org) was used to calculate
761 the distribution of each repeat and its abundance. Consensus sequences of the most
762 abundant copy were manually curated based on pfam[84] (ver 30.0) domain
763 prediction from the longest open reading frame. Whenever a LTR retrotransposon
764 was identified, long terminal repeats (LTR) and associated features were predicted
765 using LTRharvest[85]. Repeat induced mutation (RIP) analysis was carried out using
766 RIPCAL[30]
767
768 **Comparative genomics analysis**
769 Gene families were determined by OrthoFinder[86] (ver 1.0.6). Then, MAFFT[87]
770 (ver 7.271) was then used to align sequences in each of 1,127 single-copy orthogroups.
771 Alignments results with less than 10% alignment gaps were concatenated, and the
772 outcome was taken to compute a maximum likelihood phylogeny using RAxML[88]
773 with 100 bootstrap replicates. Gene family gain and loss in different positions along
774 the global phylogeny leading to *P. noxius* were inferred using dollop[89]. Pfam and
775 GO enrichments were carried out on these gene families using TopGO[90] (ver
776 2.10.0). Sequences to be included in the phylogenetic tree for NACHT domain
777 containing genes were selected on the basis of the presence of pfam domain PF05729.
778
779
780

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

815 parameters were 1 cycle of 95°C for 3 min, 30 cycles of 95°C for 30 s, 54°C for 30 s,
816 and 72°C for 60–270 s (according to different product sizes, ~2 kb/min), and a final
817 extension step of 72°C for 10 min. The PCR products were sequenced by Genomics
818 Biotechnology Inc. (Taipei, Taiwan). DNA trace data were visualized using 4Peaks
819 (http://nucleobytes. com/4peaks) and assembled using DNA Sequence Assembler in
820 DNA Baser (http://www. dnabaser.com).
821
822 **Population genomics**
823 Paired end reads of 60 *P. noxius* strains were aligned to the KPN91 reference using
824 Smalt v5.7 (www.sanger.ac.uk/resources/software/smalt/). Removal of PCR
825 duplicates and bam file sorting were implemented with Picard
826 (http://broadinstitute.github.io/picard/) and samtools[101] (ver 1.3-20-gd49c73b). The
827 first round of variant identification was implemented in Varscan[102] (v.2.4.0) and
828 degree of heterokaryon was inferred in each strain based on allele frequency and total
829 number of heterozygous SNPs called. The final list of SNPs was ascertained by
830 combining evidences from samtools[101] and FreeBayes[103] (ver1.0.2-16-
831 gd466dde). A maximum likelihood phylogeny of the SNPs segregating in these 60
832 strains was produced using FastTree[104]. plink[105] (v1.9) was used to subset
833 biallelic SNPs without linkage (filtering options: --maf 0.05 --indep-pairwise 50 5
834 0.2), which were clustered using fastSTRUCTURE[57] to determine number of
835 populations. Strains were phased using samtools[101] and one haplotype was chosen
836 at random. Consensus sequences were generated from each strain using bcftools[106]
837 and population genetics analyses were conducted using Variscan[107].
838
839
840
841
842
843
844
845
846
847
848

849 **DECLARATIONS**

850 **Ethics approval, consent to participate and consent for publication:**

851 Not applicable in this study.

852

853 **Availability of data and material:**

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

858

859 **Competing interests:**

860 The authors declare that they have no competing interests.

861

862 **Funding:**

863 I.J.T. and J.T.L were funded by MOST (105-2628-B-001-002-MY3 and 105-2313-B-
864 001 -003). C.L.C. was funded by Office of General Affairs, National Taiwan
865 University.

866

867 **Author's contributions:**

868 Strain cultivation and preparation: H.H.L., J.T.L, H.M.K, M.A., T.H., Y.O., N.S. and
869 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.
870 Strain sequencing and assembly: C.L.C., H.H.L., C.Y.C., M.J.L., T.K. and I.J.T.
871 Annotation and manual curation: H.M.K., J.T.L. and I.J.T. Comparative genomics
872 analysis: T.H.K, D.L., M.B.R., H.M.K. and I.J.T. Population genomics analysis:
873 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.
874 RNA-seq analysis: J.T.L and I.J.T. Wrote the manuscript: C.L.C, J.T.L, H.H.L, T.K
875 and I.J.T. Conceived and directed the project: C.L.C., T.K. and I.J.T.

876

877 **Acknowledgements:**

878 We thank Yen-Ping Hsueh (IMB, Academia Sinica) and Robert Waterhouse
879 (University of Lausanne) for commenting the manuscript. We thank Chian-mei Wei,
880 Kuan-chun Chen (High Throughput Genomics Core at Biodiversity Research Center,
881 Academia Sinica) for sequencing. We thank Tun-Tschu Chang and Yu-Ching Huang
882 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.

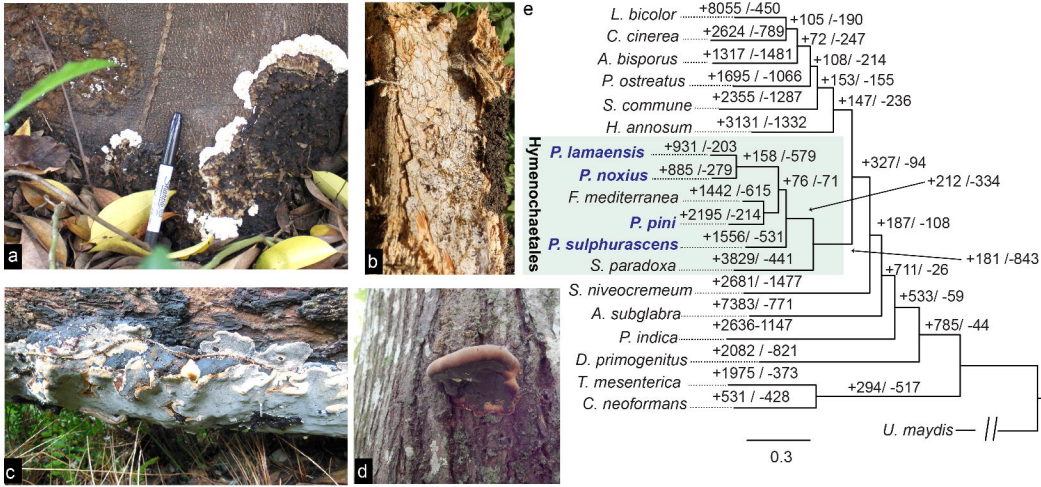
982 32. Lucía R, Gúmer P, Raúl C, Francisco S, G. PA: *Basidiomycetes Telomeres – A*
983 *Bioinformatics Approach.* InTech; 2011.
984 33. Foulongne-Oriol M, Murat C, Castanera R, Ramirez L, Sonnenberg AS:
985 **Genome-wide survey of repetitive DNA elements in the button mushroom**
986 **Agaricus bisporus.** *Fungal Genet Biol* 2013, **55:**6-21.
987 34. Edman JC: **Isolation of telomere like sequences from Cryptococcus**
988 **neoformans and their use in high-efficiency transformation.** *Mol Cell Biol*
989 1992, **12:**2777-2783.
990 35. Tudzynski B: **Nitrogen regulation of fungal secondary metabolism in fungi.**
991 *Front Microbiol* 2014, **5:**656.
992 36. Douglas CM: **Fungal beta(1,3)-D-glucan synthesis.** *Med Mycol* 2001, **39**
993 **Suppl 1:**55-66.
994 37. Leipe DD, Koonin EV, Aravind L: **STAND, a class of P-loop NTPases**
995 **including animal and plant regulators of programmed cell death:**
996 **multiple, complex domain architectures, unusual phyletic patterns, and**
997 **evolution by horizontal gene transfer.** *J Mol Biol* 2004, **343:**1-28.
998 38. Bidard F, Clave C, Saupe SJ: **The transcriptional response to nonself in the**
999 **fungus Podospora anserina.** *G3 (Bethesda)* 2013, **3:**1015-1030.
1000 39. Van der Nest MA, Olson A, Lind M, Velez H, Dalman K, Brandstrom Durling
1001 M, Karlsson M, Stenlid J: **Distribution and evolution of het gene homologs**
1002 **in the basidiomycota.** *Fungal Genet Biol* 2014, **64:**45-57.
1003 40. Jancova P, Anzenbacher P, Anzenbacherova E: **Phase II drug metabolizing**
1004 **enzymes.** *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* 2010,
1005 **154:**103-116.
1006 41. James TY, Sun S, Li W, Heitman J, Kuo HC, Lee YH, Asiegbu FO, Olson A:
1007 **Polyporales genomes reveal the genetic architecture underlying tetrapolar**
1008 **and bipolar mating systems.** *Mycologia* 2013, **105:**1374-1390.
1009 42. James TY, Lee M, van Diepen LT: **A single mating-type locus composed of**
1010 **homeodomain genes promotes nuclear migration and heterokaryosis in**
1011 **the white-rot fungus Phanerochaete chrysosporium.** *Eukaryot Cell* 2011,
1012 **10:**249-261.
1013 43. James TY, Srivilai P, Kues U, Vilgalys R: **Evolution of the bipolar mating**
1014 **system of the mushroom Coprinellus disseminatus from its tetrapolar**
1015 **ancestors involves loss of mating-type-specific pheromone receptor**
1016 **function.** *Genetics* 2006, **172:**1877-1891.
1017 44. Niculita-Hirzel H, Labbe J, Kohler A, le Tacon F, Martin F, Sanders IR, Kues
1018 U: **Gene organization of the mating type regions in the ectomycorrhizal**
1019 **fungus Laccaria bicolor reveals distinct evolution between the two mating**
1020 **type loci.** *New Phytol* 2008, **180:**329-342.
1021 45. Kües U: **From two to many: multiple mating types in Basidiomycetes.**
1022 *Fungal Biology Reviews* 2015, **29:**126-166.
1023 46. van Peer AF, Park S-Y, Shin P-G, Jang K-Y, Yoo Y-B, Park Y-J, Lee B-M,
1024 Sung G-H, James TY, Kong W-S: **Comparative genomics of the mating-**
1025 **type loci of the mushroom *Flammulina velutipes* reveals widespread**
1026 **synteny and recent inversions.** *PLOS ONE* 2011, **6:**e22249.
1027 47. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B: **The**
1028 **carbohydrate-active enzymes database (CAZy) in 2013.** *Nucleic Acids Res*
1029 2014, **42:**D490-495.
1030 48. Dashtban M, Schraft H, Syed TA, Qin W: **Fungal biodegradation and**
1031 **enzymatic modification of lignin.** *Int J Biochem Mol Biol* 2010, **1:**36-50.

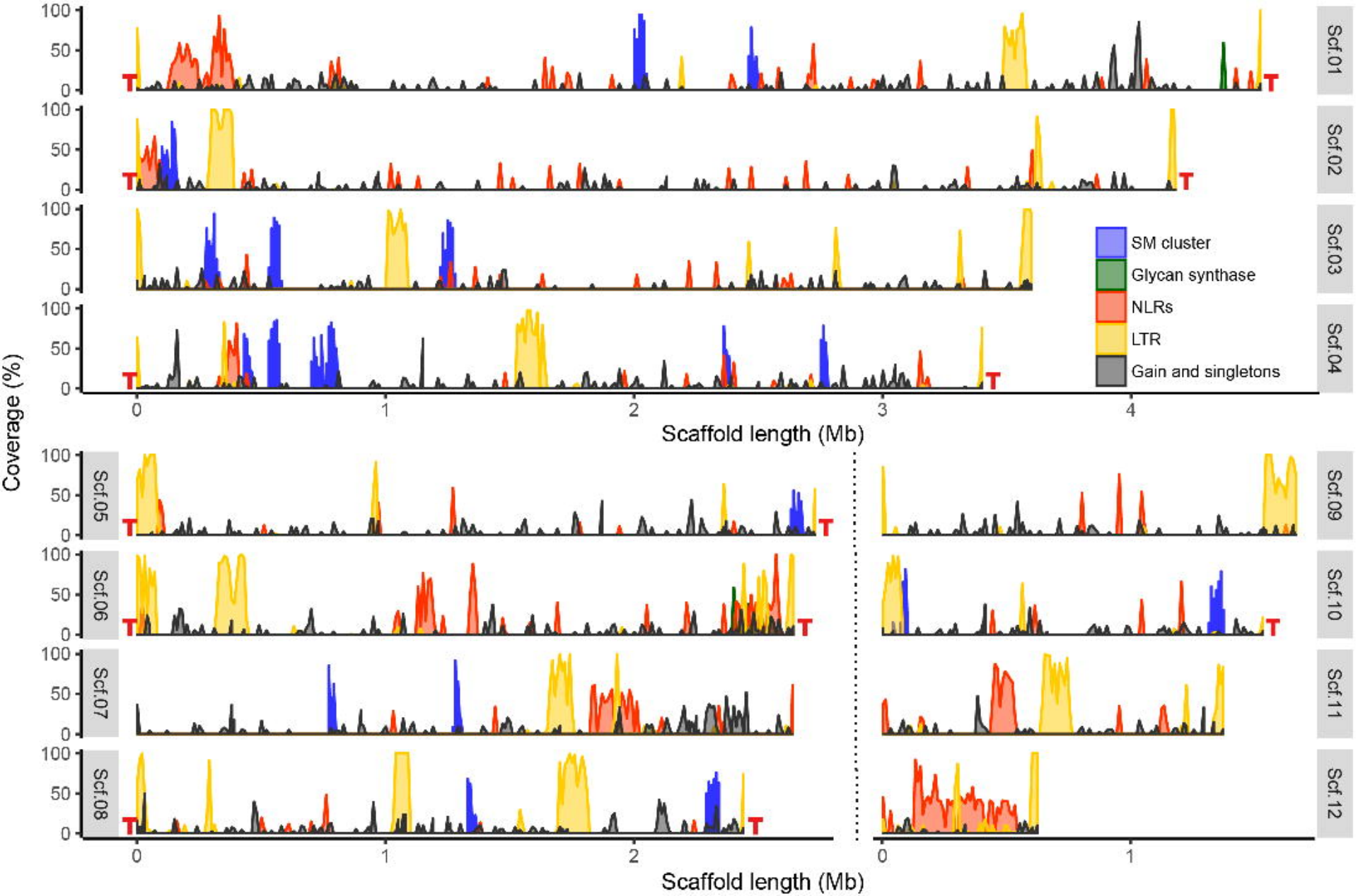
1032 49. Osbourn A: **Secondary metabolic gene clusters: evolutionary toolkits for**
1033 **chemical innovation.** *Trends Genet* 2010, **26:**449-457.
1034 50. Paoletti M, Saupe SJ: **Fungal incompatibility: evolutionary origin in**
1035 **pathogen defense?** *Bioessays* 2009, **31:**1201-1210.
1036 51. Poggeler S, Kuck U: **A WD40 repeat protein regulates fungal cell**
1037 **differentiation and can be replaced functionally by the mammalian**
1038 **homologue striatin.** *Eukaryot Cell* 2004, **3:**232-240.
1039 52. He Q, Cheng P, He Q, Liu Y: **The COP9 signalosome regulates the**
1040 **Neurospora circadian clock by controlling the stability of the SCFFWD-1**
1041 **complex.** *Genes Dev* 2005, **19:**1518-1531.
1042 53. Lah L, Podobnik B, Novak M, Korosec B, Berne S, Vogelsang M, Krasevec N,
1043 Zupanec N, Stojan J, Bohlmann J, Komel R: **The versatility of the fungal**
1044 **cytochrome P450 monooxygenase system is instrumental in xenobiotic**
1045 **detoxification.** *Mol Microbiol* 2011, **81:**1374-1389.
1046 54. Nagy LG, Riley R, Bergmann PJ, Krizsan K, Martin FM, Grigoriev IV, Cullen
1047 D, Hibbett DS: **Genetic Bases of Fungal White Rot Wood Decay Predicted**
1048 **by Phylogenomic Analysis of Correlated Gene-Phenotype Evolution.** *Mol*
1049 *Biol Evol* 2017, **34:**35-44.
1050 55. Sperschneider J, Gardiner DM, Dodds PN, Tini F, Covarelli L, Singh KB,
1051 Manners JM, Taylor JM: **EffectorP: predicting fungal effector proteins**
1052 **from secretomes using machine learning.** *New Phytol* 2016, **210:**743-761.
1053 56. Guzman-Guzman P, Aleman-Duarte MI, Delaye L, Herrera-Estrella A,
1054 Olmedo-Monfil V: **Identification of effector-like proteins in Trichoderma**
1055 **spp. and role of a hydrophobin in the plant-fungus interaction and**
1056 **mycoparasitism.** *BMC Genet* 2017, **18:**16.
1057 57. Raj A, Stephens M, Pritchard JK: **fastSTRUCTURE: Variational Inference**
1058 **of Population Structure in Large SNP Data Sets.** *Genetics* 2014, **197:**573-
1059 589.
1060 58. Leffler EM, Bullaughey K, Matute DR, Meyer WK, Segurel L, Venkat A,
1061 Andolfatto P, Przeworski M: **Revisiting an old riddle: what determines**
1062 **genetic diversity levels within species?** *PLoS Biol* 2012, **10:**e1001388.
1063 59. Baranova MA, Logacheva MD, Penin AA, Seplyarskiy VB, Safonova YY,
1064 Naumenko SA, Klepikova AV, Gerasimov ES, Bazykin GA, James TY,
1065 Kondrashov AS: **Extraordinary Genetic Diversity in a Wood Decay**
1066 **Mushroom.** *Mol Biol Evol* 2015, **32:**2775-2783.
1067 60. Swain E, Stukey J, McDonough V, Germann M, Liu Y, Sturley SL, Nickels
1068 JT, Jr.: **Yeast cells lacking the ARV1 gene harbor defects in sphingolipid**
1069 **metabolism. Complementation by human ARV1.** *J Biol Chem* 2002,
1070 **277:**36152-36160.
1071 61. Olson A, Aerts A, Asiegbu F, Belbahri L, Bouzid O, Broberg A, Canback B,
1072 Coutinho PM, Cullen D, Dalman K, et al: **Insight into trade-off between**
1073 **wood decay and parasitism from the genome of a fungal forest pathogen.**
1074 *New Phytol* 2012, **194:**1001-1013.
1075 62. Miyauchi S, Navarro D, Grisel S, Chevret D, Berrin JG, Rosso MN: **The**
1076 **integrative omics of white-rot fungus Pycnoporus coccineus reveals co-**
1077 **regulated CAZymes for orchestrated lignocellulose breakdown.** *PLoS One*
1078 2017, **12:**e0175528.
1079 63. Wendland J: *Growth, Differentiation and Sexuality.* 3 edn: Springer; 2016.

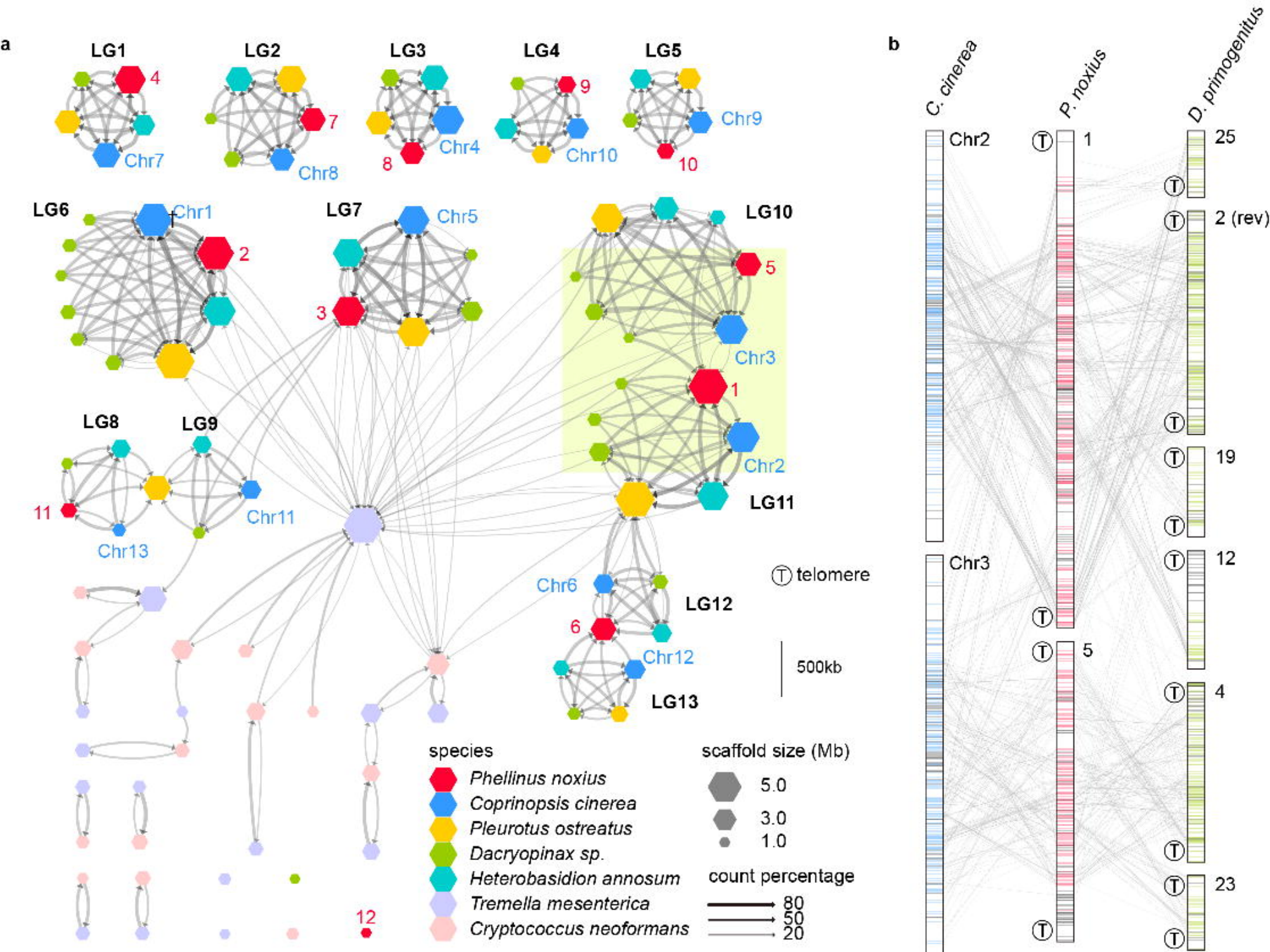
1080 64. Hattori T, Abe Y, Usugi T: **Distribution of clones of *Phellinus noxius* in a**
1081 **windbreak on Ishigaki Island.** *European Journal of Forest Pathology* 1996,
1082 **26:**69-80.
1083 65. Klopfenstein NB, Pitman EWI, Hanna JW, Cannon PG, Stewart JE, Sahashi N,
1084 Ota Y, Hattori T, Akiba M, Shuey L, et al: **A preliminary bioclimatic**
1085 **approach to predicting potential distribution of Phellinus noxious and**
1086 **geographical areas at risk from invasion.** In *Proceedings of the 63rd Annual*
1087 *Western International Forest Disease Work Conference*. 2016: 137-139.
1088 66. Chin C-S, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A,
1089 Dunn C, O'Malley R, Figueroa-Balderas R, Morales-Cruz A, et al: **Phased**
1090 **diploid genome assembly with single-molecule real-time sequencing.**
1091 *Nature Methods* 2016, **13:**1050-1054.
1092 67. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A,
1093 Copeland A, Huddleston J, Eichler EE, et al: **Nonhybrid, finished microbial**
1094 **genome assemblies from long-read SMRT sequencing data.** *Nature*
1095 *Methods* 2013, **10:**563-569.
1096 68. Lam KK, LaButti K, Khalak A, Tse D: **FinisherSC: a repeat-aware tool for**
1097 **upgrading de novo assembly using long reads.** *Bioinformatics* 2015,
1098 **31:**3207-3209.
1099 69. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for**
1100 **Illumina sequence data.** *Bioinformatics* 2014, **30:**2114-2120.
1101 70. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS,
1102 Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al: **SPAdes: A New**
1103 **Genome Assembly Algorithm and Its Applications to Single-Cell**
1104 **Sequencing.** *Journal of Computational Biology* 2012, **19:**455-477.
1105 71. Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES,
1106 Nusbaum C, Jaffe DB: **ALLPATHS: de novo assembly of whole-genome**
1107 **shotgun microreads.** *Genome Res* 2008, **18:**810-820.
1108 72. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo
1109 CA, Zeng Q, Wortman J, Young SK, Earl AM: **Pilon: An Integrated Tool**
1110 **for Comprehensive Microbial Variant Detection and Genome Assembly**
1111 **Improvement.** *PLoS ONE* 2014, **9:**e112963.
1112 73. Wences AH, Schatz MC: **Metassembler: merging and optimizing de novo**
1113 **genome assemblies.** *Genome Biology* 2015, **16**.
1114 74. Hunt M, Kikuchi T, Sanders M, Newbold C, Berriman M, Otto TD: **REAPR:**
1115 **a universal tool for genome assembly evaluation.** *Genome biology* 2013,
1116 **14:**R47.
1117 75. Stanke M, Tzvetkova A, Morgenstern B: **AUGUSTUS at EGASP: using**
1118 **EST, protein and genomic alignments for improved gene prediction in the**
1119 **human genome.** *Genome biology* 2006, **7 Suppl 1:**S11.11-18.
1120 76. Parra G, Bradnam K, Korf I: **CEGMA: a pipeline to accurately annotate**
1121 **core genes in eukaryotic genomes.** *Bioinformatics (Oxford, England)* 2007,
1122 **23:**1061-1067.
1123 77. Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M: **BRAKER1:**
1124 **Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET**
1125 **and AUGUSTUS.** *Bioinformatics* 2016, **32:**767-769.
1126 78. Borodovsky M, Lomsadze A: **Eukaryotic gene prediction using**
1127 **GeneMark.hmm-E and GeneMark-ES.** *Curr Protoc Bioinformatics* 2011,
1128 **Chapter 4:**Unit 4 6 1-10.

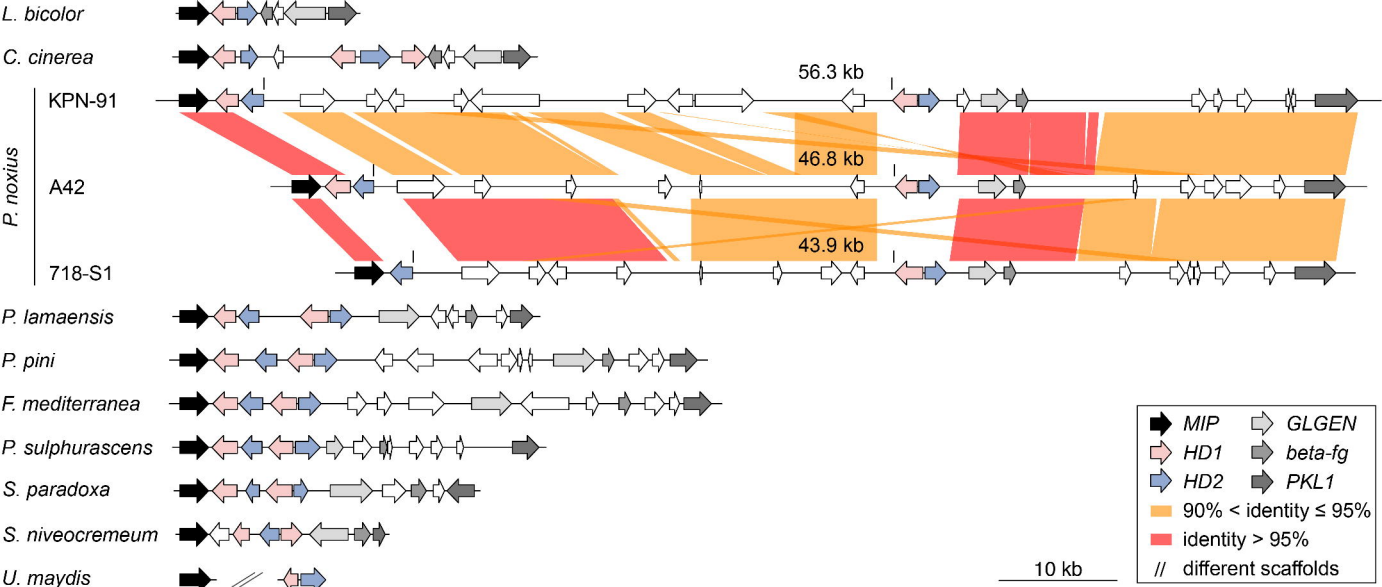
1129 79. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M: **Blast2GO:**
1130 **a universal tool for annotation, visualization and analysis in functional**
1131 **genomics research.** *Bioinformatics* 2005, **21:**3674-3676.
1132 80. Lavezzo E, Falda M, Fontana P, Bianco L, Toppo S: **Enhancing protein**
1133 **function prediction with taxonomic constraints--The Argot2.5 web server.**
1134 *Methods* 2016, **93:**15-23.
1135 81. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y: **dbCAN: a web resource for**
1136 **automated carbohydrate-active enzyme annotation.** *Nucleic Acids Res*
1137 2012, **40:**W445-451.
1138 82. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY,
1139 Fischbach MA, Müller R, Wohlleben W, et al: **antiSMASH 3.0—a**
1140 **comprehensive resource for the genome mining of biosynthetic gene**
1141 **clusters.** *Nucleic Acids Research* 2015, **43:**W237-W243.
1142 83. Edgar RC: **Search and clustering orders of magnitude faster than BLAST.**
1143 *Bioinformatics* 2010, **26:**2460-2461.
1144 84. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter
1145 SC, Punta M, Qureshi M, Sangrador-Vegas A, et al: **The Pfam protein**
1146 **families database: towards a more sustainable future.** *Nucleic Acids*
1147 *Research* 2016, **44:**D279-D285.
1148 85. Ellinghaus D, Kurtz S, Willhoeft U: **LTRharvest, an efficient and flexible**
1149 **software for de novo detection of LTR retrotransposons.** *BMC*
1150 *Bioinformatics* 2008, **9:**18.
1151 86. Emms DM, Kelly S: **OrthoFinder: solving fundamental biases in whole**
1152 **genome comparisons dramatically improves orthogroup inference**
1153 **accuracy.** *Genome Biol* 2015, **16:**157.
1154 87. Katoh K, Standley DM: **MAFFT: iterative refinement and additional**
1155 **methods.** *Methods Mol Biol* 2014, **1079:**131-146.
1156 88. Stamatakis A: **RAxML-VI-HPC: maximum likelihood-based phylogenetic**
1157 **analyses with thousands of taxa and mixed models.** *Bioinformatics (Oxford,*
1158 *England)* 2006, **22:**2688-2690.
1159 89. Felsenstein J: **PHYLIP (Phylogeny Inference Package) version 3.6.**
1160 *Distributed by the author Department of Genome Sciences, University of*
1161 *Washington, Seattle* 2005.
1162 90. A A, J R: **topGO: Enrichment Analysis for Gene Ontology.** *R package*
1163 *version 2260* 2016.
1164 91. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P,
1165 Chaisson M, Gingeras TR: **STAR: ultrafast universal RNA-seq aligner.**
1166 *Bioinformatics* 2012, **29:**15-21.
1167 92. Liao Y, Smyth GK, Shi W: **featureCounts: an efficient general purpose**
1168 **program for assigning sequence reads to genomic features.** *Bioinformatics*
1169 2014, **30:**923-930.
1170 93. Love MI, Huber W, Anders S: **Moderated estimation of fold change and**
1171 **dispersion for RNA-seq data with DESeq2.** *Genome Biology* 2014, **15**.
1172 94. Young MD, Wakefield MJ, Smyth GK, Oshlack A: **Gene ontology analysis**
1173 **for RNA-seq: accounting for selection bias.** *Genome Biology* 2010, **11:**R14.
1174 95. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H,
1175 Maslen J, Mitchell A, Nuka G, et al: **InterProScan 5: genome-scale protein**
1176 **function classification.** *Bioinformatics* 2014, **30:**1236-1240.

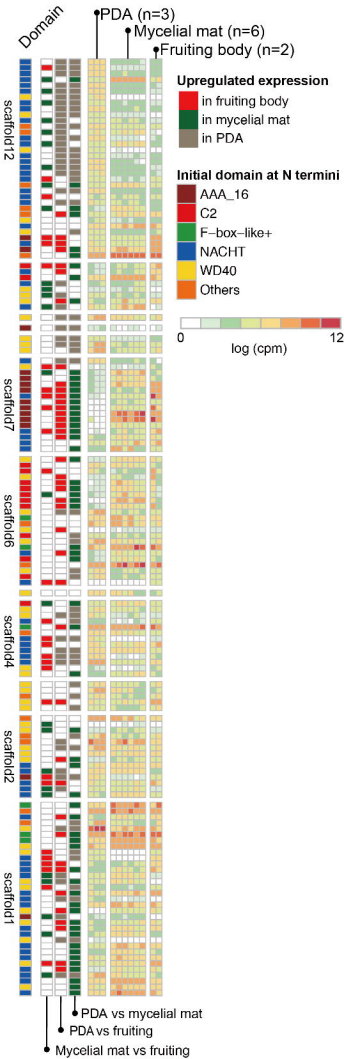
1177 96. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R,
1178 Salamov A, Zhao X, Korzeniewski F, et al: **MycoCosm portal: gearing up**
1179 **for 1000 fungal genomes.** *Nucleic Acids Research* 2014, **42:**D699-D704.
1180 97. Guy L, Roat Kultima J, Andersson SGE: **genoPlotR: 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

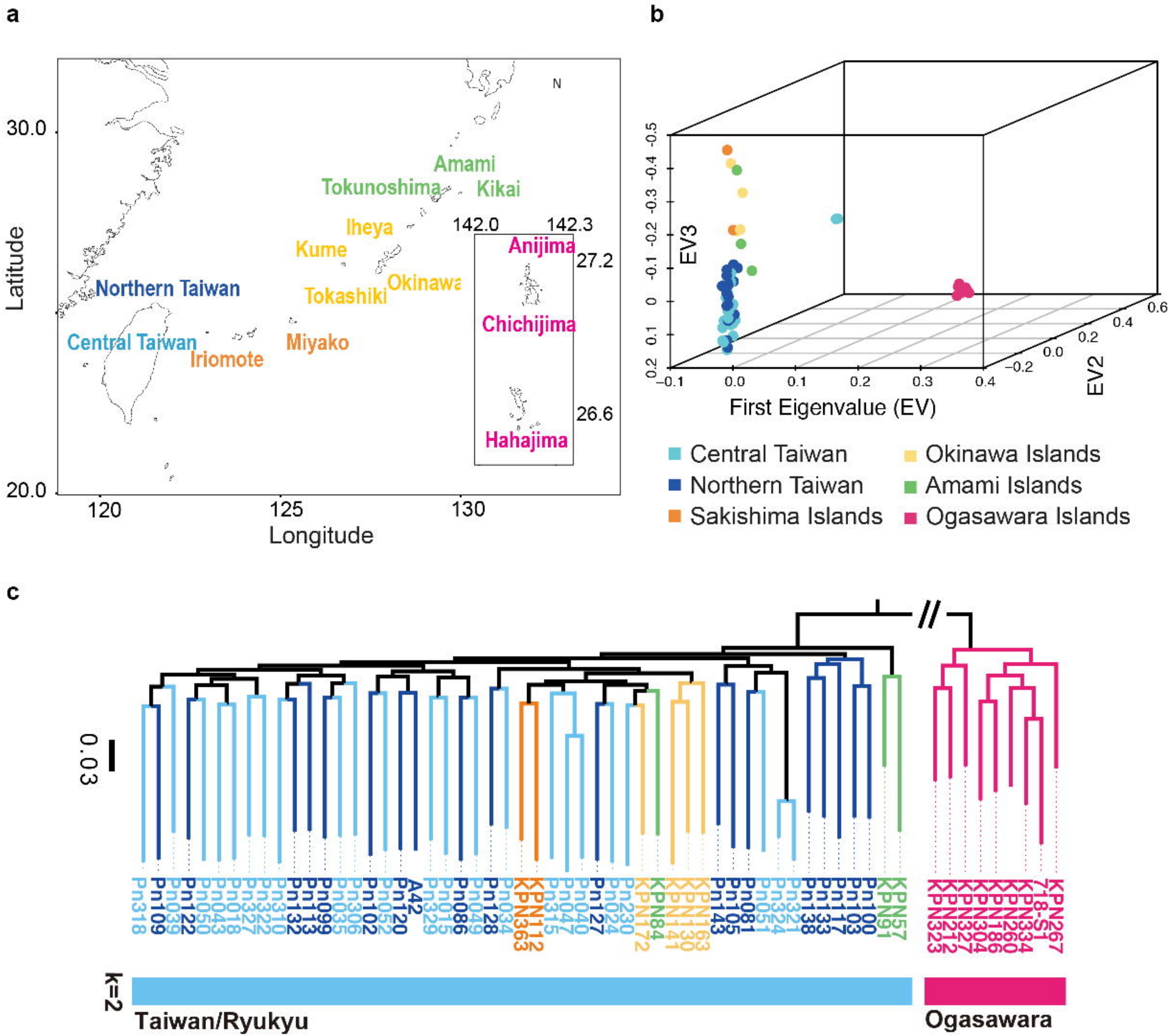

a
b
c
d
e
Hymenochaetales
L. bicolor
C. cinerea
A. bisporus
P. ostreatus
S. commune
H. annosum
P. lamaensis
P. noxius
F. mediterranea
P. pini
P. sulphurascens
S. paradoxa
S. niveocremeum
A. subglabra
P. indica
D. primogenitus
T. mesenterica
C. neoformans
U. maydis
+8055 /-450
+105 /-190
+2624 /-789
+72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


Coverage (%)
Scf.01
Scf.02
Scf.03
Scf.04
Scf.05
Scf.06
Scf.07
Scf.08
Scf.09
Scf.10
Scf.11
Scf.12
SM cluster
Glycan synthase
NLRs
LTR
Gain and singletons
Scaffold length (Mb)
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
10 kb
MIP
HD1
HD2
GLGEN
beta-fg
PKL1
90% < identity ≤ 95%
identity > 95%
// different scaffolds


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