Reticulate evolution and rapid development of reproductive barriers upon secondary contact pose challenges for species delineation in a forest fungus

Dabao Sun Lu[1], David Peris[1,2], Jørn Henrik Sønstebø[3], Timothy Y. James[4], Loren H. Rieseberg[5], Sundy Maurice[1], Håvard Kauserud[1], Mark Ravinet[6], Inger Skrede[1]

[1]Department of Biosciences, University of Oslo, Oslo, Norway

[2]Department of Biotechnology, Institute of Agrochemistry and Food Biotechnology (IATA), CSIS, Paterna (Valencia), Spain

[3]Department of Natural Sciences and Environmental Health, University of South-Eastern Norway, Bø, Norway

[4]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, USA

[5]Department of Botany and Biodiversity Research Centre, The University of British Columbia, Vancouver, Canada

[6]School of Life Sciences, University of Nottingham, Nottingham, UK

Abstract

Reproductive barriers within morphospecies of the mushroom-forming fungi tend to be stronger between sympatric lineages, leading to speculation on whether they are being reinforced by selection against hybrids. In this study we use population genomic analyses together with *in vitro* crosses of a global sample of the wood decay fungus *Trichaptum abietinum* to investigate reproductive barriers within this morphotaxon and the processes that have shaped them. Our phylogeographic analyses show that the morphospecies is delimited into six major lineages, one in Asia, two in Europe, and three in North America. The two lineages present in Europe are interfertile and admixed, whereas our crosses show that the North American lineages are reproductively isolated. In Asia a more complex pattern appears, with partial intersterility between multiple sub-lineages that likely originated independently and more recently than the reproductive barriers in North America. We found pre-mating barriers in *T. abietinum* to be moderately correlated with genomic divergence, whereas fitness of the mated hybrids showed a strong positive correlation with increasing genomic divergence. Genome wide association analyses identified candidate genes with

programmed cell death annotations, which are known to be involved in intersterility in distantly related fungi. Our demographic modelling and phylogenetic network analyses suggest that reproductive barriers in *Trichaptum abietinum* could have been reinforced upon secondary contact between lineages that diverged in allopatry during the Pleistocene glacial cycles. Our combination of experimental and genomic approaches demonstrate how *T. abietinum* is a tractable system for studying speciation mechanims.

Introduction

It is estimated that kingdom fungi contains between 1.5 and 7 million species[1-3], yet our knowledge about the mechanisms that have generated this diversity is disproportionally small in comparison to what we know about animals and plants. Unique life cycles, small genomes, and a relatively high vagility may set speciation mechanisms in fungi apart from those of other eukaryotes[4-6]. The development of reproductive barriers is central to the speciation process as they restrict geneflow between populations and thereby fortify them as independently evolving lineages[7]. Work on fungal model organisms[8,9], and pathogens of plants[10] and humans[11], have found that reproductive barriers in fungi can be caused by genetic incompatibilities[12,13], chromosomal rearrangements[14,15] and ecological specialization[16,17]. Another mechanism that can create strong reproductive isolation is reinforcement. Here natural selection works directly against maladaptive hybridization by favoring pre-mating barriers[18], and consequently these are expected to be stronger between populations in sympatry than allopatry[19]. In compiling information on reproductive barriers in fungi, Le Gac and Giraud[20] found pre-mating barriers to be significantly stronger for species pairs in sympatry than allopatry in subclass Agaricomycetes[21], suggesting a role for reinforcement in shaping the reproductive barriers in this group, which contains the mushroom forming fungi.

Many agaricomycete wood decay fungi comprise species complexes with wide distributions and intersterility groups (ISGs) in sympatry[22-27]. In addition, they are easy to culture and cross in the laboratory, which makes them suitable for investigating fungal speciation and reinforcement. In this study we focus on *Trichaptum abietinum* (Pers. ex J.F. Gmel.) Ryvarden, where previous crossing experiments have revealed two sympatric ISGs in North America, and a third European ISG which can mate with both of the North American ISGs[23,28]. *Trichaptum abietinum* is a common decay fungus on dead wood of tree species of *Pinaceae* across boreal and temperate regions. Molecular studies of rDNA[29-31] and genome data[32] have revealed multiple divergent lineages within *T. abietinum*, and two of these have been connected to the ISGs previously detected in North America[31].

In this study we use population genomic analyses in conjunction with lab experiments on a global collection of *T. abietinum* to learn more about the processes that have generated the high species diversity in fungi. First, we apply a phylogeographic framework to assess how many lineages are present globally. We predict that the genetic structure reflects survival in different refugia during the last glacial maximum, but to a lesser extent than in animals and plants due to a high vagility. We then test and compare pre- and post-mating barriers between lineages using crossing and growth experiments, where stronger pre-mating barriers between sympatric lineages can be interpreted as support for reinforcement, although we recognize that other evolutionary processes can produce a similar pattern[19]. We search for the genetic basis of intersterility between lineages with genome scans and genome wide association studies, and expect barrier loci involved in reinforcement selection to show a signature of positive selection, similar to what has been shown in other systems[33-36]. Secondary contact is generally a precondition for reinforcement[18], and we ask if the lineages with pre-mating reproductive barriers show evidence of this.

Results and Discussion

Phylogeography detects six major lineages within the *Trichaptum abietinun* morphotaxon

To uncover the global diversity and genetic structure within the *T. abietinum* morphotaxon, we obtained fruit body samples covering most of its distribution (suppl.table1). The genomes of 358 haploid (monokaryotic) cultures from these samples were sequenced to an average depth of 24x on an Illumina platform and mapped to a reference genome of *T. abietinum*[32] to obtain a SNP dataset for population genomic analyses. We used three different methods to assess the population structure: i) Principal component analysis (PCA) ii) maximum likelihood (ML) phylogeny and iii) Admixture analysis, which in addition to population structure can detect admixture, i.e. when samples contain ancestry from multiple lineages. Altogether, these analyses support the division of *T. abietinum* into six major lineages. Our PCA showed five distinct lineages: three in North America, one in Europe and Siberia, and one in East Asia (Fig.1a). North America C consisted of only two samples from higher elevation sites in Western North America, and is likely underrepresented in our study as these areas were not sampled extensively. The same five lineages were detected in the ML phylogeny (Fig.1b), whereas Admixture[37] split the European samples into two lineages; one consisting of samples from southern Europe and another containing samples from Northern and Eastern Europe extending into Siberia, hereafter referred to as the European and Eurasian lineages, respectively (Fig.1c,1d). A number of samples from Europe and East Asia appeared admixed, and we will discuss these results below (see sections on reproductive barriers and secondary contact). North America C appeared as an admixture of North America A and the European lineage, which may be an artefact caused by the low sample size[38]. In comparing the ITS sequences of our samples to those of Seierstad et al.[31], we determined that North America A and B correspond to the previously detected ISGs of Macrae[23]. While Seierstad et al.[31] inferred a widespread Circumboreal lineage in *T. abietinum* covering both Europe, Asia and North America using rDNA markers, our genome data from a similar spread of populations shows that this group includes three differentiated and geographically restricted lineages; Eurasian, East Asian, and North America B, all with Fst $\geq$ 0.2 (suppl.table2).

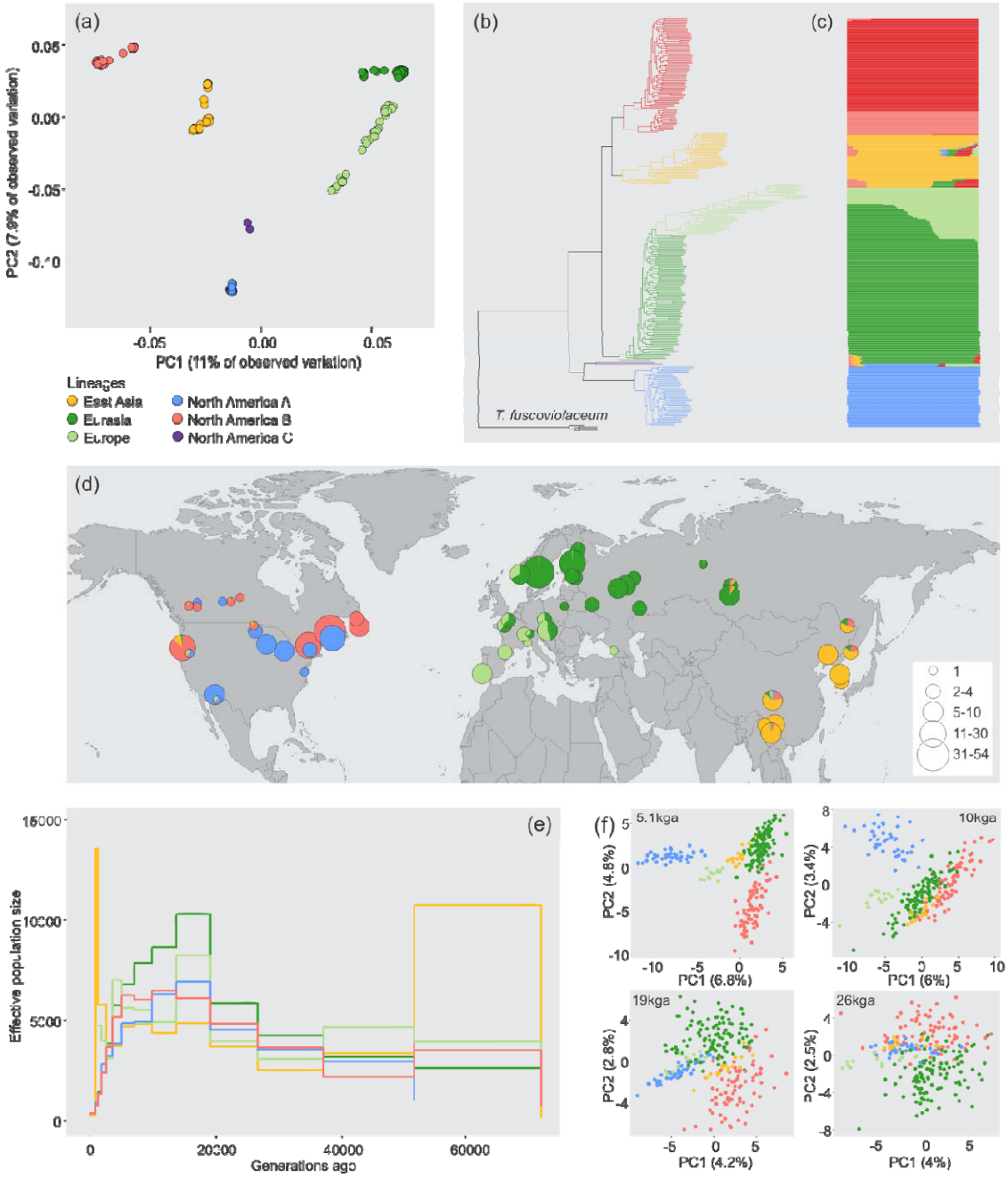

Figure 1. Phylogeographic patterns of *Trichaptum abietinum* based on population genomics data **a**) Principal component analysis (PCA) showing five distinct clusters. A sixth distinct European lineage detected in other analyses is colored in light green. **b)** Maximum likelihood (ML) phylogeny including the sister species *T. fuscoviolaceum* as outgroup showing the same 5 major lineages as in the PCA, with 100% support values for all major lineages **c)** Admixture analysis of K = 6, which had the lowest cross validation error. The samples are ordered in the same order as the ML phylogeny, and identify an additional cluster of North America B from western North America. **d)** Populations plotted with their average admixture proportions on a world map. The size of the circle indicates the number of samples from the population. If two lineages are found at the same site there are two circles i.e. North America A and B in eastern North America and North America C with North

America A and B in western US. **e)** Coalescence analysis showing estimated effective population size through time for selected populations of 5 of the lineages (excluding North America C). **f)** PCA of the pairwise coalescence rate for the same samples as in e) at selected time points. Color denote lineage the sample have been ascribed to.

We performed coalescence analyses in Relate[39] to infer historic relations between the major lineages and their changes in effective population size (Ne). All lineages showed a strong increase in Ne about 20,000 generations ago (Fig.1e). Since the generation time of *T. abietinum* is unknown, the number of generations cannot readily be converted to years. *Trichaptum abietinum* is a primary colonizer that rapidly fruits on recently dead wood substrates, and we note that the rise in Ne coincides with the end of the last glacial maximum 20,000 years ago if we assume a generation time of one year. The lineages did not coalesce on the graph (Fig.1e), possibly indicating limitations of this method for the small genome size of 50 Mb for *T. abietinum*[40]. PCAs of the pairwise coalescence rates between all individual samples back in time show that the Eurasian, East Asian and North America B lineages cluster together at several time periods with similar coalescence rates indicating shared ancestry. We interpret this as high levels of gene flow between these lineages, and it is possible that the rise in Ne 20,000 generations ago is due to increased gene flow between all lineages. At any event, the negative Tajima´s D values for Eurasia indicate that recent population expansions have occurred here (suppl.table3). Interpreting this in light of the greater similarity in pairwise coalescence with East Asia and North America B (Fig.1f), we suggest that the Eurasian lineage has an eastern origin. Similarly, we suggest that North America B originated in Beringia from East Asian ancestors, and migrated to North America over the land bridge, but also spread westward to East Asia and Europe, where it mixed back in with existing lineages (Fig.1f). Southeastern US was a refugium for boreal vegetation when the Laurentide ice sheet covered most of northern North America[41]. We hypothesize that North America A expanded from this refugium considering how they were more frequent in our samples from eastern and southern parts of the continent and that the nucleotide diversity is higher in the south (suppl.table3). Similarly, North America C may have originated south of the Cordilleran ice sheet in western North America. However, further sampling near the historic refugia is needed to verify these hypotheses. Considering how Asia remained mostly ice free during the ice ages[42], more continuous geneflow between populations could have kept the lineages here from diverging to the same extent as on the other continents. The biogeographic patterns we observed of higher levels of genetic diversity

in the south and a greater number lineages in North America, has been found in a number of other Agaricomycetes[22,43-45]. The abovementioned patterns are also known trends for plants and animals[46], suggesting that these fungi are structured by similar biogeographic processes and could have followed the same postglacial migration routes, but more fine grained studies are needed here.

Crossing experiments reveal multiple reproductive barriers

Interfertility in Agaricomycetes has been scored by the ability of two monokaryotic mycelia to form clamp connections upon pairing, as the presence of these structures indicate a dikaryotic mycelium which results from a successful mating between monokaryons[23,47,48] (Fig.2a). To assess mating barriers within the *T. abietinum* morphospecies, two to five selected monokaryotic isolates from each of the six major lineages were crossed in all combinations (Fig.2b). In correspondence with previous findings[23,28], the majority of our crosses between North America A and B produced no clamp connections. The exceptions were four parings between allopatric North America A and B isolates that resulted in sparse production of clamps. All sympatric inter-lineage pairings were completely intersterile, consistent with reinforcement, where stronger pre-mating barriers are expected in sympatry. Our crosses revealed additional ISGs: North America C was completely intersterile with North America A. Surprisingly, the two samples of North America C were not interfertile with each other, and may represent reproductively isolated populations. Moreover, North America A showed almost complete intersterility with the samples from Yunnan and Sichuan in the East Asian lineage. The isolates from Yunnan were also largely intersterile with all the other major lineages. Due to the complex pattern in East Asia, we also crossed 12 isolates from East Asia in all combinations (Fig.2c). This revealed that the Yunnan population is intersterile with the population from Heilongjiang and partially intersterile with the Sichuan population. Interestingly, both of these populations appear to be admixed (Fig1d), alluding to a role for hybridization in causing intersterility here. In any event, these barriers must be recent given the low divergence between the populations, suggesting that pre-mating barriers can evolve rapidly. The Eurasian and European lineage were almost entirely interfertile with each other, whereas the other crosses between lineages showed a patchwork of mixed intersterility lacking clear patterns, indicating that intersterility is governed by several loci, similar to what has been inferred in the wood decay fungus *Heterobasidion*[49].

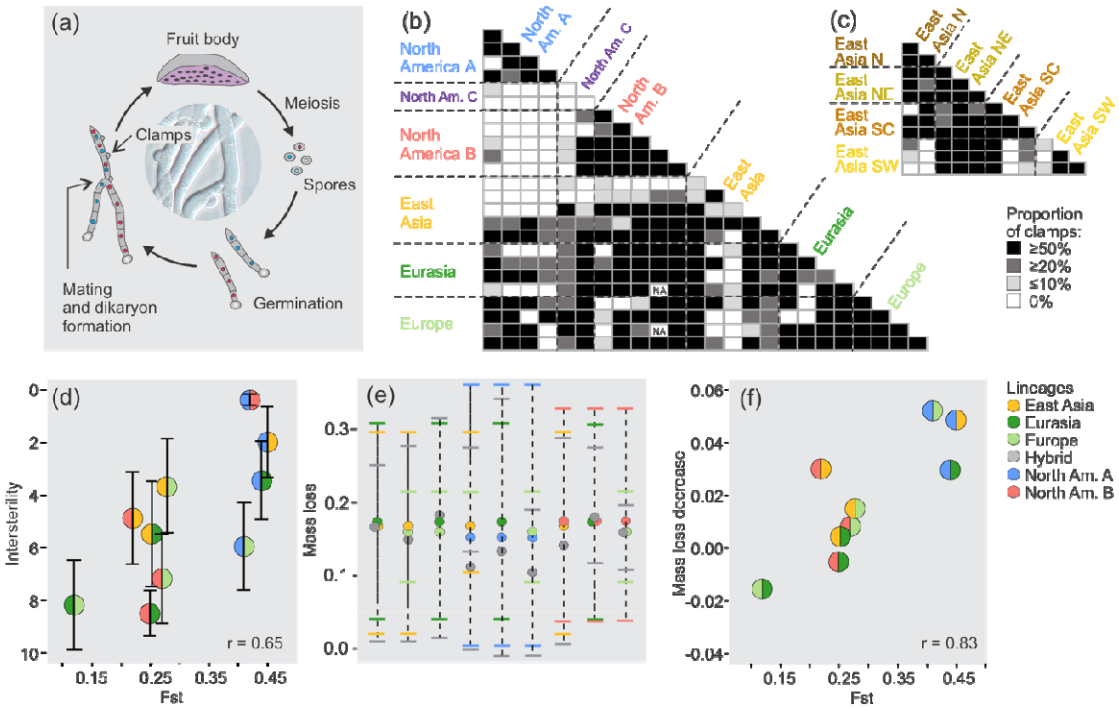

Figure 2. **a)** Overview of typical life cycle for basidiomycetes, including *Trichaptum abietinum*. Basidiospores produced by meiosis in the fruit body germinate into haploid mycelia known as monokaryons. Interfertile monokaryons can mate and form a dikaryotic mycelium, and this is the dominant stage of the life cycle. Dikaryons can be recognized by clamp structures which extend between cell junctions, and their presence or absence in crosses have been used to score intersterility/interfertility between populations. **b)** Crosses of selected monokaryons from all the major lineages. Crosses were scored according to the average proportion of clamps. For each cross we examined 10 cell junctions for clamps. The majority of crosses were set up in triplicates, thus the % is based on 30 possible clamp occurrences. **c)** Crosses between sublineages in East Asia scored the same way as in b) **d)** Intersterility measured as mean number of clamps for crosses between major lineages plotted against genetic distance between major lineages as measured in average Fst between lineages (r^2 = 0.65, p = 0.04). Correlation calculated using Pearsons correlation test. Bars delimit 95% confidence interval. **e)** Proportion of mass loss of decomposed pine wood blocks by hybrid dikaryons as compared to each parental lineage. **f)** Mean of differences in decomposition in hybrid dikaryons as compared to the average decomposition of each parent lineage plotted against genetic distance between major lineages as measured in average Fst between lineages. (r = 0.83, p = 0.005) Correlation calculated using Pearsons correlation test.

We observed a moderate correlation of 0.65 (p = 0.04) when relating Fst values between the major lineages to their degree of reproductive isolation (Fig.2d). When including the incompatibility patterns between the 12 East Asian isolates, the correlation dropped to 0.51 (p = 0.04). The general increase of intersterility with genetic divergence we observe fits with speciation in allopatry where reproductive barriers result from genetic incompatibilities,

which has been proposed for several Agaricomycetes[27,50,51]. However, sufficient build-up of genetic incompatibilities to cause intersterility may take considerable time[52], and we note that in contrast to the overall trend, North America A and the Yunnan population in the East Asian lineage were able to mate with isolates that are equally or more divergent than the isolates they cannot mate with in sympatry or parapatry. Similar patterns have been observed in several basidiomycete species complexes[51,53,54], which suggest these pre-mating barriers are caused by more than the clock-like accumulation of BDMI and may involve reinforcement selection.

Mating tests between monokaryotic cultures only assess prezygotic plasmogamy and do not account for the actual fitness of the dikaryon and its ability to produce a fruit body and complete the life cycle (Fig2a). As a simple assessment of their fitness, we measured the wood decay capacity of hybrid dikaryons between five of the major lineages, and compared it to that of the dikaryons synthesized from within the lineages. In seven out of nine cases, the hybrid dikaryons decomposed wood slower than dikaryons from parent populations (Fig.2e, suppl.fig_dikaryon_fitness), and the hybrids between East Asia and North America A showed a significantly lower decay rate than either parental population. Overall, we observed a strong correlation ($r = 0.83$, $p = 0.005$) between Fst values of parent populations and the reduction in wood decay of their hybrid dikaryons (Fig.2f). Our measurements of decay under lab conditions may not reflect how well dikaryons perform in nature, but we note that the correlation is stronger between wood decay ability and Fst than between intersterility and Fst. Lower hybrid fitness has been documented in other basidiomycete fungi[55,56], and makes a case for reinforcement to avoid investment in unfit hybrids. In the same vein, the increased fitness between hybrids of the Eurasian and European lineage could explain the lack of pre-mating barriers here.

To search for genetic regions associated with intersterility we applied a combination of genome wide association studies and genome scans. We ran a genome wide association study (GWAS) on intersterility with North America A where all samples were included, and intersterility was inferred based on lineage for the samples which we had not included in our crossing experiments. In addition, we conducted a multivariate analysis of correlated phenotypes with SCOPA[57] where the 25 samples in our crossing experiment and their intersterility with all 6 major lineages were analysed together. Assuming that intersterility in basidiomycetes is caused by heterogenic incompatibility, we searched for genomic regions

that are highly differentiated between lineages that cannot mate, but similar in the lineages that can mate, as quantified by ΔFst. Functional annotations of these analyses revealed 407, 1580, and 1357 genes that contained at least one PFAM domain and overlapped with significant GWAS SNPs, SCOPA SNPs, or ΔFst windows, respectively. (Suppl.table4, Suppl.table5, Suppl.table6). The main functions of these lists of genes are difficult to evaluate, and there is a confounding effect of population structure since intersterility is highly correlated with lineage (Suppl.figGWAS). However, there are a few clues on genes to search for; a recent study of the fungus *Podospora* by Ament‐Velásquez et al[58] found genes for somatic incompatibility that regulated cell death to also be causing intersterility between populations. In our dataset, two of the 62 SNPs with 0 p-value in the SCOPA analysis, the 12th most significant GWAS SNP, and 5 of the genes within ΔFst outlier windows were annotated as metacaspases. Metacaspases regulate cell death[59] and are known to function in somatic incompatibility in the fungal phylum Ascomycota[60]. In total, 7 out of 9 genes putatively annotated as metacaspases turned up in one of these three analyses. The metacaspase on scaffold 6 came up in all three analyses, and was identified as ΔFst outlier among populations of the East Asian lineage, suggesting they may play a role in the partial reproductive barriers here. The genes underlying somatic incompatibility are not characterized in agaricomycetes[61], but programmed cell death have been predicted to function in somatic compatibility here[62], and we speculate that this machinery may have been co-opted for intersterility in *T. abietinum*. A test for positive selection in Relate on the increase in frequency of variants in the inferred genealogies did not detect any significant signal for positive selection within the metacaspase gene annotations in North America A, North America B, or the East Asian sublineages, which are the lineages we predict to have experienced reinforcement selection.

Relating reproductive barriers to secondary contact

We carried out additional analyses to assess if the pre-mating barriers in *T. abietinum* could be linked to secondary contact, and by extension, to reinforcement. To investigate secondary contact, we used Treemix[63] to construct phylogenies that can account for admixture of lineages. The suggested 99.8 threshold of explained variation[63] was reached in a phylogenetic network with five admixture events that included admixtures between both the reproductively isolated North America A and B, and the partially reproductively isolated Sichuan and Yunnan populations in East Asia (Fig.3a). The admixture was estimated as weak between the

two North American lineages, suggestive of introgression, but relatively strong in East Asia, consistent with the creation of a new lineage through hybridization. In addition, Treemix identified introgression into the East Asian lineages from an unsampled lineage related to the outgroup *T. fuscoviolaceum*. Kinneberg et al.[64] found support for ancient introgression between *T. abietinum* and *T. fuscoviolaceum*, and it is conceivable that introgressed genes from the *T. fuscoviolaceum* clade have spread into North America A and B through populations in East Asia.

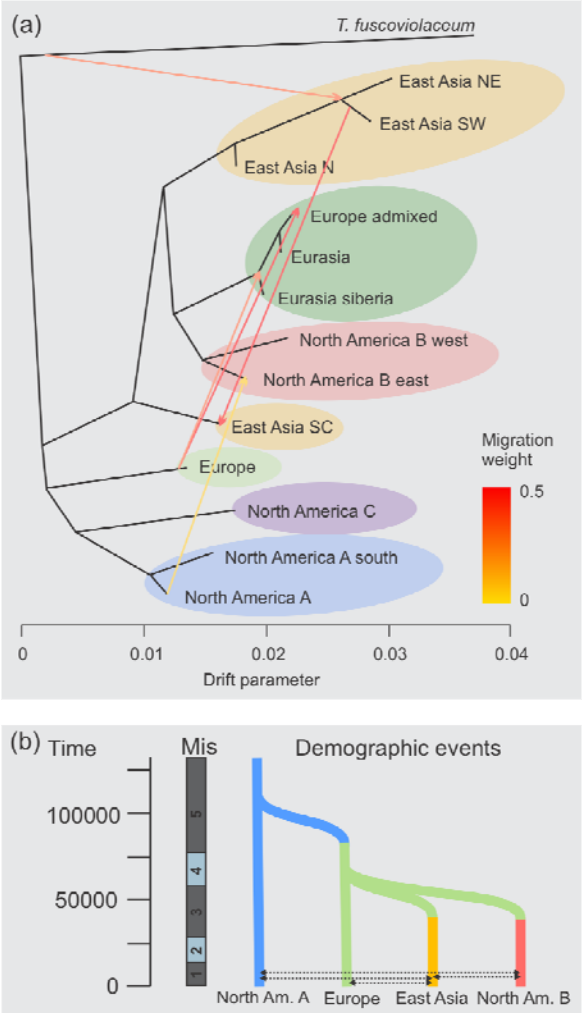

Figure 3. a) Phylogenetic network in Treemix with 5 admixture events, indicated by arrows between lineages and colored by migration weight. b) Coalescence and geneflow events in the best of Fastsimcoal2 model plotted with time, assuming a generation time of 1 year, and major isotopic stages (Mis). Mis captures fluctuations in earth temperature, and cold periods are shown in blue, where Mis2 corresponds to the last glacial maximum.

We applied demographic modelling with Fastsimcoal2[65] to explicitly test for secondary contact between the reproductively isolated North America A and B, and to estimate divergence time between the major lineages (table1). All models with gene flow were significantly better than the model without gene flow. Among the models with gene flow, those including secondary contact between North America A and B had a significantly better likelihood than comparable models with the same number of parameters (table1, suppl.fig.likelihood_distributions). The best model included four migration events, and here the coalescence time between North America A and the other lineages was estimated to 113,000 generations ago. Coalescence time generally precedes divergence time[40], and if we assume a generation time of 1 year, the divergence between North America A and the European lineage happened around the end of the previous interglacial[66]. In the same model, North America B and East Asia can be inferred to have diverged from Europe sometime after their coalescence 60,000 generations ago, after the onset of the last ice age but around a minor transition when the global climate experienced some warming (fig.3b). The best model included geneflow between both East Asia, Europe, and North America A and B, and it is possible that the more recent coalescence of these groups indicated by Relate captured the geneflow that occurred after the main split. It is conceivable that the reproductive barriers between sub-lineages in East Asia and North America A also resulted from reinforcement selection, but they may also be a by-product from geneflow with North America B.

Given the evidence of geneflow between lineages, we conducted a sliding window PCA with Lostruct[67] to investigate if specific genomic regions stood out with a signal of introgression. The analysis indicated a different pattern from the genome average on the second half of Scaffold 5 (Fig4a,4b). A PCA plot of this region revealed that the North America A, B, and C lineages to be more similar here when compared to the PCA of the whole genome (Fig.4c). Inspection of our Fst genome scans showed a similar pattern, where the second half of scaffold 5 had lower divergence between North America A and B than any other Fst comparisons including these lineages (Fig.4d). This could indicate an inversion that has spread from North America A to B. Inversions are predicted to enhance reinforcement and rapid development of reproductive isolation as they tend to suppress the recombination rate[68], and we note that two of the metacaspases from the ΔFst outlier windows fall within this region. If the introgression on scaffold 5 was part of the secondary contact that led to the reinforcement of pre-mating barriers, the more recent clustering of coalescence rates on this

region of scaffold 5 2000-3000 generations ago (Fig.4e) may indicate when the reinforcement selection happened.

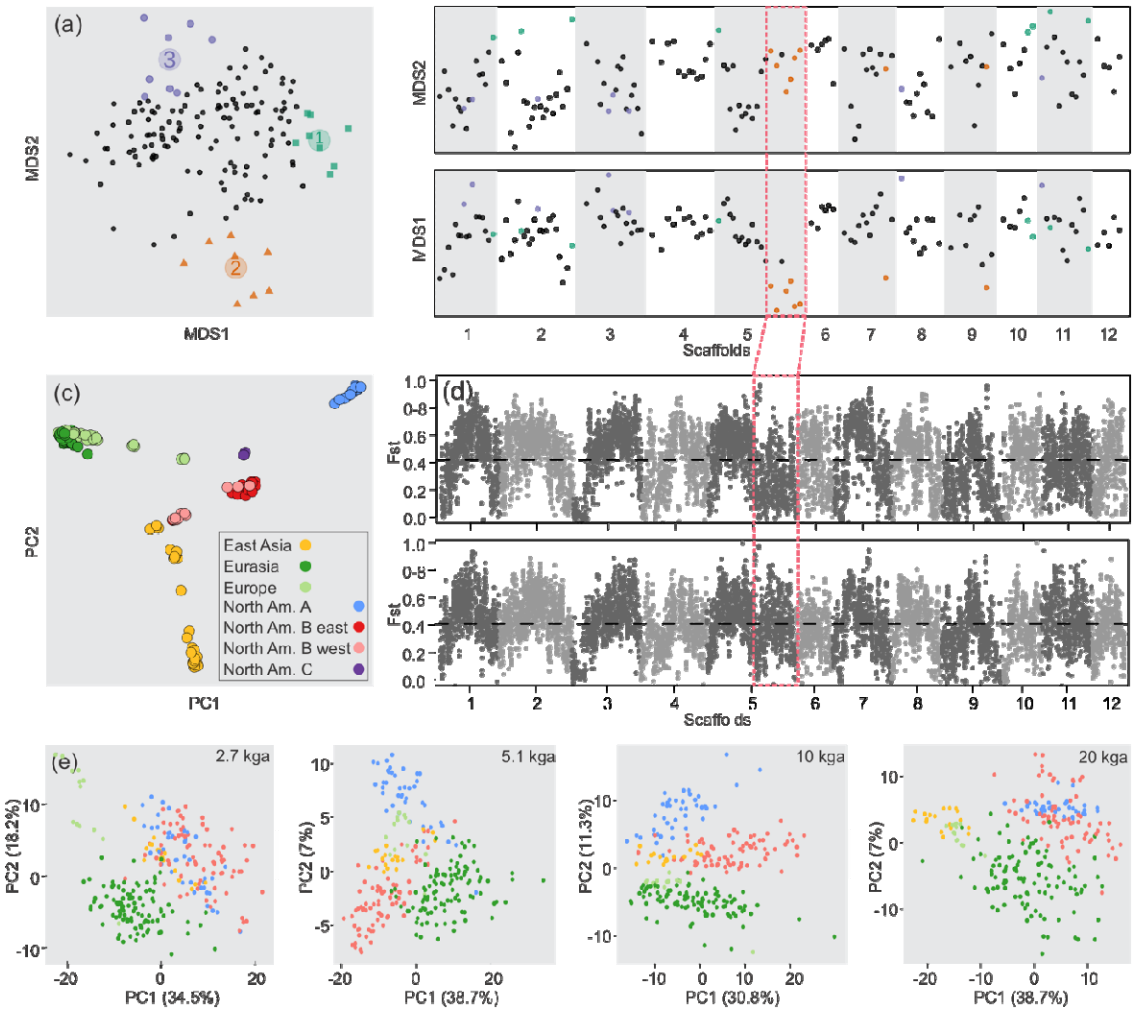

Figure 4. Scaffold 5 in the genome of *Trichaptum abietinum* show a distinct pattern **a)** Local PCA using Lostruct indicate three corners of genomic windows which deviate from the genomic average. **b)** The placement of the Lostruct outlier (corner) windows on MDS1 and MDS2 along the genome. **c)** A PCA of the SNPs in genomic windows of corner 2 show a different structure than the PCA of the genome average in Figure 1a). **d)** Fst genome scan of North America A and North America B compared to Fst of North America A and Eurasia. The second half of Scaffold5 which contains a high proportion genomic windows in corner 2 from the Lostruct analyses is boxed in. **e)** PCA of the coalescent rates of the second half of scaffold5 for selected time points.

Concluding remarks

In this study we show how the morphospecies of a common forest fungus contain multiple differentiated and reproductively isolated lineages. These lineages have not evolved independently of each other, on the contrary, introgression and geneflow appear to be

common. This reticulate pattern of evolution can be interpreted in light of the Pleistocene glacial cycles where previously isolated populations came into secondary contact during the interglacials. Several of the pre-mating barriers in *T. abietinum* fit a scenario of reinforcement, and we interpret the signals of secondary contact between reproductively isolated lineages and decrease of hybrid fitness with increased divergence, as additional support of this. Mating with the wrong lineage can be costly, nonetheless, we find large introgressed regions that have been retained, and we speculate that there could also have been certain adaptive advantages associated with inter-lineage matings. Future experiments on the ability of different hybrid dikaryons to produce fruit bodies with viable spores would allow an assessment of the relative strength of reproductive barriers at all stages of the life cycle and provide a more complete context to evaluate the case we make for reinforcement. Moreover, a controlled crossing scheme and sequencing of spore families from hybrid fruit bodies should be able to pinpoint the genomic basis of the premating barriers in *T. abietinum* and confirm the role of controlled cell death here.

Multiple reproductive barriers across different divergence levels make *T. abietinum* an exceptional study system to test outstanding questions in speciation genomics such as whether particular classes of genes or genetic elements have been involved repeatedly in independent speciation events. However, these qualities also make species concepts difficult to apply. It has been advised that biological species concepts in fungi are most useful in cases of sympatry[5], whereas the phylogenetic species concept will be more discriminating in allopatry since significant phylogenetic divergence can precede the development of reproductive isolation in fungi[51,69,70]. In our case, North America A can be recognized as a distinct species based on both biological and phylogenetic species recognition criteria. Yet, North America A could, in theory, transfer their genetic material to intersterile lineages in North America and Asia through interfertile lineages in Europe. As such, *T. abietinum* resemble a ring species with North America A and B as intersterile end populations. Furthermore, following the criterion of reproductive isolation, the East Asian Yunnan sub-lineage should also be recognized as a distinct biological species. But consolidating a separate Yunnan species with the phylogenetic species concept means splitting up and recognizing both the other East Asian sublineages, and also North America B and Eurasia, as separate species in order to avoid paraphyletic species. It appears that fungal morphospecies, when analysed at a global scale, contain a multitude of lineages with nested levels of divergence and intersterility that are hard to fit within a rigid species concept.

Materials and Methods

Field sampling and culture isolation

To understand the global diversity and structure of the *T. abietinum* morphotaxon we collected fruit body samples in the field or had them sent in by collaborators between 2017 and 2020, where our aim was to cover as much of the known distribution range as possible. Individual collections were made from separate logs when possible, or in the cases where they were sampled from the same log, collections had to be at least 2 m apart to count as separate samples. Whenever possible, we included five samples from each locality. Monospore cultures were isolated by rewetting the dry fruit bodies in sterile H_2O over night before they were attached to a petri dish lid with silicone grease from Merck Millipore (Merck KGaA, Darmstadt, Germany) over a 3% malt extract agar plate containing 25 mg/L streptomycin, 100 mg/L ampicillin, 10 mg/L tetracycline and 1 mg/L benomyl to inhibit the growth of contaminants. The lids with fruit bodies were set to shoot spores for 3 hours to 2 days, and subsequently replaced with sterile lids. These plates were set to germination at 19 °C in the dark for 3-10 days, and germinated spores were excised and placed on new 3% malt extract agar plates with antibiotics and benomyl. After 2 weeks, slides of the mycelia were examined under the microscope to confirm the absence of clamp connections, indicating that they were monokaryotic cultures. To confirm that the isolated cultures were indeed *Trichaptum*, PCR of a small piece of culture tissue was performed using primers ITS1 and ITS4[71] with Phire Plant Direct PCR Kit (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's recommendations: 10 ul Phire Plant PCR Buffer, 4.8 ul H_20, 4 ul of 5 mM primer, 0.2 ul of Phire hot start II DNA polymerase, and 1 ul DNA containing 1 mg mycelium diluted in 20 ul dilution buffer. Amplicons were Sanger Sequenced at GATC Biotech/Eurofins Genomics (Eurofins Scientific, Luxembourg City, Luxembourg) using the ITS4 primer, after enzymatic clean up with Exo-SAP-IT PCR Product Cleanup Reagent (Thermo Fisher Scientific, Waltham, MA, USA) applying a 10x dilution of the Exo-SAP-IT but otherwise following the manufacturer´s instructions. Sanger sequences were blasted to the Unite[72] and NCBI databases[73] to confirm species identity, and aligned in Geneious Prime 2019.0.4 (https://www.geneious.com/) to an ITS2 alignment from Seierstad et al.[31] to ascribe

tentative lineage grouping. All work on the fungal cultures were carried out under a ESCO class II biological safety cabinet (ESCO lifesciences Gmbh, Friedberg, Germany) to avoid contamination.

DNA extraction and Sequencing

To prepare DNA for Illumina sequencing, monokaryotic cultures were grown for 2-3 weeks on sterilised nitex nylon (Sefar AG, Heiden, Switzerland) placed on top of 3% malt extract agar plates. Mycelial tissue was scraped off from plates and homogenized in 2 ml Lysing Matrix E tubes (MP Biomedicals, Santa Ana, CA, USA) on a FastPrep-24 (MP Biomedicals, Santa Ana, CA, USA) for 2 x 20 s at 4.5 m/s^2. DNA was extracted using the E.Z.N.A HP Fungal DNA kit (Omega Bio-Tek, Norcross, GA, USA) supplemented with 30 ul RNaseA (Qiagen, Hilden, Germany) prior to homogenization. The quantity and quality of DNA extractions were assessed with the Invitrogen Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) using the Invitrogen dsDNA BR assay kit (Thermo Fisher Scientific, Waltham, MA, USA), NanoDrop Microvolume Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and on a 1% agarose gel with GelRed (Biotium, Fremont, CA USA). Illumina sequencing was performed on either Hiseq 4000 or Novaseq at the Norwegian Sequencing Centre with library preparation as described in Peris et al.[32]. Briefly, 1 µg of genomic DNA was sheared to a target size of 300-400 bp with a Covaris E220 system (Covaris Inc., Woburn, MA, USA), cleaned on a small volume Mosquito liquid handler (TTP labtech) with a 1:1 ratio of Kapa Pure beads (Roche, Basel, Switzerland). Genomic libraries were generated using Kapa Hyper library prep kit (Roche, Basel, Switzerland) and Illumina UD 96 index kit (Illumina, San Diego, CA, USA) to add barcodes. Samples were pair-end sequenced at a read length of 2 x 150 bp.

Read processing, SNP-calling and filtering

Bioinformatic analyses were performed on the SAGA supercomputer as part of the Sigma2 the National Infrastructure for High Performance Computing and Data Storage in Norway. Illumina reads were examined with FastQC 0.11.8 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and visualized with MultiQC[74]. Samples with less than 1 million reads were excluded from further analyses. Reads were trimmed with TrimGalore 0.6.5 (https://github.com/FelixKrueger/TrimGalore) at q30 and mapped with BWA 0.7.17 -mem[75] to the PacBio assembly of the isolate TA-1010-6-M1 of *T. abietinum* from Peris et al.[32], which is from the North America B lineage and consists of 12

nuclear scaffolds. Mapped reads were converted from sam to bam file format and sorted using SAMtools 1.9[76] and duplicates were marked with the MarkDuplicates module in GATK 4.1.4.0. SNPs were called with GATK4´s pipeline of joint variant discovery[77] by running Haplotypecaller in GVCF mode with ploidy = 1 followed by GenomicsDBImport and GenotypeGVCFs to produce the GVCF-file containing the joint variants in the data. The GVCF-file was quality filtered using BCFtools 1.10.2[76] following an approach outlined in Barth et al.[78] setting the thresholds to exclude sites that have: FS > 40, MQRankSum < −5.0 || > 5.0, ReadPosRankSum < −4, QD < 2, MQ > 40, SnpGap = 10, number of indels = 0, number of alleles = 2, INFO/DP > meanDP/2, missing data <20% and no monomorphic SNPs. The efficiency of filtering was visually controlled in R 4.2.2, and samples with > 50% of sites missing were removed from the VCF-file. A separate GATK4 joint variant discovery pipeline was run on the mapped reads in diploid mode to assess the VCF-file with VCFtools 0.1.16 –het[79]. In the output of VCFtools –het, 10 samples with F values < 0.8 were removed from the VCF-file as they likely represent dikaryons due to the high levels of heterozygosity. The resulting VCF-file contained 353 samples *T. abietinum* and 5 samples of the sister clade *T. fuscoviolaceum,* and was used as basis for further filtering for the different population genomic analyses. In several of the analyses, sites with missing data were removed from the VCF-file using the command bcftools view -g ^miss in BCFtools. To obtain subsets of the data, we used vcf-subset in VCFtools.

Analyses of Population structure

We used Principal Component Analysis (PCA), Admixture and a Maximum Likelihood (ML) phylogeny to assess the population structure in our SNP data. For the PCA and Admixture analyses the 5 *T. fuscoviolaceum* samples were removed from the VCF-file prior to filtering at minor allele frequency (MAF) > 0.05 and to only retain biallelic SNPs in BCFtools. The VCF-file was then LD pruned in PLINK 1.9[80] with r2 threshold of 0.1 calculated within 50 bp windows with a step size of 10 bp to retain 46470 out of 776952 SNPs. The PCA was performed in Eigensoft 7.2.1[81]. Admixture[37] was run with cross validation for 2-10 k clusters. The R package scatterpie (https://github.com/GuangchuangYu/scatterpie) was used to plot the admixture proportions onto a world map. For the maximum likelihood phylogeny, our SNP dataset including 5 samples of *T. fuscoviolaceum* was converted to the phylip format with vcf2phylip.py (https://github.com/edgardomortiz/vcf2phylip). IQtree[82] was run on the phylip file with the options -T AUTO -m GTR+ASC -alrt 1000 -B 1000 using only biallelic sites. The ASC flag accounts for the lack of invariant sites in the SNP data by providing

ascertainment bias correction. The phylogeny from IQtree was visualised in Figtree (http://tree.bio.ed.ac.uk/software/figtree/). We used the main lineages found in the PCA, Admixture and ML phylogeny to ascribe samples to lineages for a Fst genome scan analysis. Here, a VCF-file without *T. fuscoviolaceum* and filtered in BCFtools to contain only biallelic SNPs was converted to a geno file with parseVCF.py. Sliding window genome wide statistics of Fst were calculated from the geno file popgenWindows.py using a window size of 10 kb and step size of 5 kb. Both scripts are available at https://github.com/simonhmartin/genomics_general. The admixed samples in central Europe and the North America C were not included in the analysis due to high admixture proportions and low sample size, respectively. A separate Fst genome scan was run at finer scale where the North America A, North America B, and East Asia was also split up into sublineages. Tajima´s D and pi were calculated in the R package PopGenome[83].

Analyses of Recombination

Coalescence analysis in Relate requires estimates of the recombination rate, and we used LDhelmet[84] to obtain estimates of the recombination rate along each of the 12 nuclear scaffolds. We performed this analysis on 53 North America B samples belonging to the same lineage and population (New Brunswick) as the reference strain TA-1010-6-M1. A VCF-file containing the 53 samples filtered to include only biallelic SNPs and with all missing sites removed with BCFtools was split into individual scaffolds with vcftools --chr. Each VCF-file was then converted to the fasta input file for LDhelmet with python scripts. LDhelmet was run as recommended in the manual with the steps: find_confs -w 50, ldhelmet pade -t 0.01 -x 11, ldhelmet rjmcmc -p output.pade -w 50 -b 50.0 and using a linkage table generated with LDhat [85], and finally extracting the mean output with ldhelmet post_to_text. The linkage table was generated in LDhat complete with the parameters -n 53 -rhomax 100 -n_pts 101 -theta 0.01, and converted to the LDhelmet format using ldhelmet convert_table --input_table. We obtained our estimate of theta with the R package Pegas[86] using scaffold 12 of the same VCF-file of the 53 North America B samples. Scaffold 12 is the smallest scaffold and was chosen to make the computation manageable.

Coalescence analysis

Coalescence analyses were performed with Relate[39] to investigate population histories of the major lineages. The VCF-file without *T. fuscoviolaceum* and filtered to contain only biallelic SNPs and no missing sites in BCFtools were split up into VCF-files for each individual

scaffold with VCFtools --chr. The input haps files for Relate were generated from these VCF files using PLINK/2.00 --export haps[80]. The resulting haps files were then converted to reflect the haploid nature of our sequence data by excising only every other column for each sample. RelateFileFormats -mode FlipHapsUsingAncestor was used to infer ancestral allele state using a consensus sequence of *T. fuscoviolaceum.* The consensus sequence was generated with vcf-consensus in VCFtools using the PacBio reference assembly of *T. abietinum* and extracting the 5 samples of *T. fuscoviolaceum* from the general VCF-file. Relate was then run in --mode All setting the mutation rate to 1e-7 as studies on another rapidly fruiting wood decay fungus found a mutation rate within this order of magnitude[87]. The effective population size (Ne) was set to 10,000, and the recombination map for the Relate analysis was obtained by multiplying the population recombination rate from LDhelmet with 100/(4*Ne), The resulting anc and mut files of each scaffold from the Relate analysis were used to re-estimate coalescence rates and effective population sizes with the Relate script EstimatePopulationSize.sh. Here we combined all scaffolds and specified --num-iter 1 to avoid bias, --threshold 0 to include all sites, and --years_per_gen 1 to specify a generation time of 1 year. EstimatePopulationSize.sh was run separately to obtain the average within coalescence rate of each population that was plotted Fig.1e and used to estimate changes in effective population size for each generation, and also run with all samples without any population specification to obtain coalescence rates between all samples. Principal components of these coalescence rates between all samples were computed with prcomp() in R with normalization of the columns prior to plotting in R. A test for positive selection was calculated on the inferred genealogy for each major lineage separately based on the DetectSelection.sh in Relate on the re-estimated anc and mut files. The North America C lineage was excluded from the analysis due to low sample size (n=2).

Mating experiments

In order to test if monokaryotic isolates could mate and form dikaryotic mycelia, paired isolates were placed 4 cm apart on 3% malt extract agar plates. When the two mycelia had grown out to encounter each other, after 2-3 weeks, a small piece from the contact zone was excised and re-plated on a new 3% malt extract agar plate. Microscope slides were made of this new culture after one week of growth, and examined under a Nikon Eclipse 50i microscope (Nikon Instruments Europe BV, Amsterdam Netherlands) at 400x magnification to assess the frequency of clamps. Each slide was scored for the proportion of clamps present on the first 10 observed septae (cell junctions). If none of the first 10 septae contained clamps,

different areas of the slide were also searched, and if clamps were then detected a note of this was made. The majority of crosses were set up in triplicates, and mating success was evaluated as average proportion septa with clamps across replicates. In the classification of matings used for figure 2, matings were rated as positive if on average 5/10 or more of the septae contained clamps, intermediate if 2/10-4/10 contained clamps, marginal if 1/10 or less contained clamps and in cases were none of the first 10 examined septae contained clamps, but clamps were found in other parts of the slide, and negative if no clamps were found at all. Images of the slides were acquired with a Zeiss Axioplan-2 imaging with Axiocam HRc microscope camera (Zeiss, Oberkochen Germany). The ability to mate between groups were correlated to their average genome wide Fst distance with a Pearson´s correlation test.

Dikaryon fitness experiment

As an assessment of the fitness of dikaryons, 2 cm x 2 cm x 0.5 cm woodblocks of *Pinus sylvestris* placed on top of petri dishes with water agar and inoculated with dikaryons. Decay was measured after 3 months at 19 °C by calculating percentage mass loss in the woodblock. For the experiment we used four hybrid dikaryons between the North America A, North America B, Eurasian, European, and East Asian lineages, and eight dikaryons from within each of these lineages. For the within lineage dikaryons, both parents were from the same or nearby population. Each dikaryon was set up in four replicates. The average difference of mass loss between the parental lineages and the hybrid in each case were correlated to Fst distances with a Pearson´s correlation test. T-tests in R were used to determine if the wood decay ability of the hybrid dikaryons were significantly different than dikaryons of each parent lineage.

Genome Wide Association Studies

We carried out a genome wide association study (GWAS) to associate intersterility with North America A to our SNP data. The phenotype in our GWAS was set as binary according to intersterility with the North America A lineage. For crosses where no clamps or only a few clamps were formed (≤ 1/10) phenotype was set as 0, and for the rest phenotype was set as 1. Intersterility for the samples where we did not have crossing information were inferred based on their lineage, so that it was 0 for North America B, North America C, East Asia SW and East Asia SC. Phenotype was set as 1 for the samples belonging to the Eurasia, Europe, East Asia NE and East Asia N lineages. Map and raw files were produced using PLINK2.00 --recode on VCF file without *T. fuscoviolaceum* filtered for missing sites in BCFtools. The

GWAS was run with the amm_gwas script in R available at https://github.com/arthurkorte/GWAS/blob/master/gwas.r. To account for population structure a VanRaden.2 kinship matrix was calculated with the gMatrix package available at https://github.com/filippob/introduction_to_gwas/blob/master/software/gMatrix_0.2.tar.gz.
To leverage our crossing data and account for intersterility with more than 1 lineage, a subset of the 25 samples for which we had exact crossing information was given to SCOPA[57] for a multivariate analysis to detect correlated phenotypes. A gen file was generated with PLINK2.00 from a VCF-file of the 25 samples with the --export oxford setting as input for the SCOPA analysis. As phenotype, we used intersterility with 12 of the 25 samples to reduce the computational time. A given sample were set to be able to mate with itself for the analysis, as this is governed by the homogenic incompatibility system of the mat-A and mat-B loci that control mating within an ISG, whereas we were interested in detecting the genes that cause intersterility between ISGs.

ΔFst

ΔFst was calculated for 10 kb Fst genome scan windows of intersterile lineages as contrasted to interfertile lineages. A high positive ΔFst thus indicates a region that is highly differentiated between lineages that cannot mate, but similar in the lineages that can mate. We selected our comparisons where the interfertile and intersterile comparisons would involve the same lineage, and where the intersterile and interfertile lineage compared to it would be as closely related to each other as possible. We identified the 95% percentile outlier windows towards the positive tail end of the empirical distribution of all ΔFst windows for each of the 17 comparisons.

Gene function

We used the annotation of our reference genome from[32] as basis of gene function searches. Metacaspases and the pfam domains PF05729 and PF00400, which are associated with het-genes in the sister phylum Ascomycota, were pulled out from the gff file containing the annotations. Tblastx searches for het gene homologs were performed using het gene sequences in *Podospora anserina* and *Neurospora crassa* downloaded from the uniprot database[88] as described in[61]. An enrichment analysis of GO terms within the outlier ΔFst windows, GWAS snps, and SCOPA snps, was conducted in R with Rtracklayer available at https://bioconductor.org/packages/release/bioc/html/rtracklayer.html and topGO available at https://bioconductor.org/packages/release/bioc/html/topGO.html following the setup on:

https://ucdavis-bioinformatics-training.github.io/2018-June-RNA-Seq-Workshop/friday/enrichment.html. For the p-value, a Benjamini-Hochberg corrected threshold of p = 0.05 was used. Similarly, enrichment analyses of PFAM domains falling within genes with significant GWAS SNPs, SCOPA SNPs, and within ΔFst windows were carried out in R. In the enrichment analyses, we considered only the subset of pfams that occurred in our region of interest. The test was set up as described in[89] with Fisher's exact test, and Benjamini-Hochberg corrected threshold of p = 0.05.

Treemix analysis

Treemix infers a maximum likelihood tree including an optional number of m migration edges between branches and thereby can account for hybridization and introgression between lineages. The linkage pruned VCF file from our PCA and Admixture analysis was converted to a strat file using plink2treemix.py (https://github.com/thomnelson/tools/blob/master/plink2treemix.py) and vcf2tremix (https://github.com/speciationgenomics/scripts) for Treemix[63]. Treemix analysis was run with -m edges of 0-10 and a block size of -k 500. Results were evaluated with the R package OptM [90].

Demographic modelling

For the demographic modelling, a subsampled VCF-file filtered for missing sites of 13 samples each from the North America A, North America B, European and Yunnan population of the East Asian lineages covering a similar geographic range were chosen. The model of coalescence order between the lineages we used in these analyses were guided by the ML phylogeny, coalescence analysis, and Treemix, which all indicate North America A as the first lineage to diverge. We included only the most realistic scenarios of geneflow, namely across the Bering strait between North American groups and East Asia, between Eurasia and East Asia, and within the North American groups. Gene flow between North America and Europe, i.e. across the Atlantic, was not considered, with the exception of a simple model with only one event of gene flow. Site frequency spectrum files were generated from the VCF-file with easySFS -a -f (https://github.com/isaacovercast/easySFS). For each demographic model, Fastsimcoal 2.6[65] was initiated with 50 independent runs with -m -n 100000 -L 50 -s 0 -M, which specifies 50 cycles of 100,000 coalescent simulations per cycle. The mutation rate in the tpl files was set to 1e-7, the same as in our coalescence analysis in Relate. The run with the highest likelihood and its estimated parameters for each model was

extracted with scripts available on https://github.com/speciationgenomics/scripts. A likelihood distribution for each model was obtained by running fsc26 -n 1000000 -m -q with the best parameter values for the given model with 100 iterations and collecting the likelihood values for each run using a script available at https://github.com/speciationgenomics/scripts. The likelihood values were then used to compute AIC-distributions of the models, and both were plotted in R. An initial comparison of models with 1, 2 or 3 splits between the lineages identified a model with 2 splits as the model with best support. This model was used as basis for further testing of models with gene flow.

Local PCA and Coalesence Analysis of Scaffold 5

Our local PCA was run with Lostruct[67] using scripts available at https://github.com/petrelharp/local_pca/blob/master/templated/README.md. The analysis was run for 5 kb windows on a 0.05 MAF-filtered and biallelic VCF-file. The second half of scaffold 5 was extracted using VCFtools --from-bp 3213128 --to-bp 6426255, and analysed separately in Relate in the same manner as the other analyses in Relate described under the section on coalescence analysis.

Data visualization

If not otherwise mentioned, data was visualized in R with ggplot2[91], and final graphic design adjustments were made in CorelDRAW (Corel corporation, Ottawa, Canada).

Aknowledgements

This work was supported by the Research Council of Norway (RCN) grant No. RCN 274337 and the University of Oslo and UiO:Life Science internationalization support. David Peris is supported by the RCN grant No. RCN 324253, and the Generalitat Valenciana plan GenT grant No. CIDEGENT/2021/039. We thank all the sample collectors who made this study possible. The computations were performed on resources provided by UNINETT Sigma2 — the National Infrastructure for High Performance Computing and Data Storage in Norway.

References

1 Blackwell, M. The Fungi: 1, 2, 3 ... 5.1 million species? *American Journal of Botany* **98**, 426-438, doi:https://doi.org/10.3732/ajb.1000298 (2011).

2 Hawksworth, D. L. & Lucking, R. Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiol Spectrum* **5**, doi:https://doi.org/10.1128/microbiolspec.FUNK-0052-2016 (2017).

3 Baldrian, P., Větrovský, T., Lepinay, C. & Kohout, P. High-throughput sequencing view on the magnitude of global fungal diversity. *Fungal Diversity*, doi:https://doi.org/10.1007/s13225-021-00472-y (2021).

4 Kohn, L. M. Mechanisms of fungal speciation. *Annual review of phytopathology* **43**, 279-308, doi:https://doi.org/10.1146/annurev.phyto.43.040204.135958 (2005).

5 Giraud, T., Refregier, G., Le Gac, M., de Vienne, D. M. & Hood, M. E. Speciation in fungi. *Fungal genetics and biology* **45**, 791-802, doi:https://doi.org/10.1016/j.fgb.2008.02.001 (2008).

6 Chou, J. Y., Hsu, P. C. & Leu, J. Y. Enforcement of Postzygotic Species Boundaries in the Fungal Kingdom. *Microbiol Mol Biol Rev* **86**, e0009822, doi:10.1128/mmbr.00098-22 (2022).

7 Dobzhansky, T. Speciation as a Stage in Evolutionary Divergence. *The American Naturalist* **74**, 312-321 (1940).

8 Gladieux, P. *et al.* Neurospora from Natural Populations: Population Genomics Insights into the Life History of a Model Microbial Eukaryote. *Methods Mol Biol* **2090**, 313-336, doi:https://doi.org/10.1007/978-1-0716-0199-0_13 (2020).

9 Louis, E. J. Population genomics and speciation in yeasts. *Fungal Biology Reviews* **25**, 136-142, doi:https://doi.org/10.1016/j.fbr.2011.06.001 (2011).

10 Stukenbrock, E. H. Evolution, selection and isolation: a genomic view of speciation in fungal plant pathogens. *New Phytologist* **199**, 895-907, doi:https://doi.org/10.1111/nph.12374 (2013).

11 Vogan, A. A. & Xu, J. Evidence for genetic incompatibilities associated with post-zygotic reproductive isolation in the human fungal pathogen Cryptococcus neoformans. *Genome* **57**, 335-344, doi:10.1139/gen-2014-0077 (2014).

12 Dettman, J. R., Anderson, J. B. & Kohn, L. M. Divergent adaptation promotes reproductive isolation among experimental populations of the filamentous fungus Neurospora. *BMC Evolutionary Biology* **8**, 35, doi:10.1186/1471-2148-8-35 (2008).

13 Anderson, J. B. *et al.* Determinants of divergent adaptation and Dobzhansky-Muller interaction in experimental yeast populations. *Current biology : CB* **20**, 1383-1388, doi:https://doi.org/10.1016/j.cub.2010.06.022 (2010).

14 Charron, G., Leducq, J.-B. & Landry, C. R. Chromosomal variation segregates within incipient species and correlates with reproductive isolation. *Molecular ecology* **23**, 4362-4372, doi:https://doi.org/10.1111/mec.12864 (2014).

15 Yadav, V., Sun, S., Coelho, M. A. & Heitman, J. Centromere scission drives chromosome shuffling and reproductive isolation. *Proceedings of the National Academy of Sciences* **117**, 7917-7928, doi:10.1073/pnas.1918659117 (2020).

16 Stukenbrock, E. H. *et al.* The making of a new pathogen: insights from comparative population genomics of the domesticated wheat pathogen Mycosphaerella

graminicola and its wild sister species. *Genome Research* **21**, 2157-2166, doi:https://doi.org/10.1101/gr.118851.110 (2011).
17 Zaffarano, P. L., McDonald, B. A. & Linde, C. C. Rapid speciation following recent host shifts in the plant pathogenic fungus Rhynchosporium. *Evolution* **62**, 1418-1436, doi:https://doi.org/10.1111/j.1558-5646.2008.00390.x (2008).
18 Butlin, R. Speciation by reinforcement. *Trends in Ecology & Evolution* **2**, 8-13, doi:https://doi.org/10.1016/0169-5347(87)90193-5 (1987).
19 Coyne, J. A. & Orr, H. A. *Speciation*. Vol. 37 (Sinauer Associates Sunderland, MA, 2004).
20 Gac, M. L. & Giraud, T. Existence of a pattern of reproductive character displacement in Homobasidiomycota but not in Ascomycota. *Journal of Evolutionary Biology* **21**, 761-772, doi:https://doi.org/10.1111/j.1420-9101.2008.01511.x (2008).
21 Cannon, P. *et al.* in *State of the World's Fungi 2018* Ch. 1. Definition and diversity, 4–11 (Royal Botanic Gardens, 2018).
22 Anderson, J. B., Korhonen, K. & Ullrich, R. C. Relationships between European and North American biological species of Armillaria mellea. *Experimental Mycology* **4**, 78-86, doi:https://doi.org/10.1016/0147-5975(80)90053-5 (1980).
23 Macrae, R. Pairing incompatibility and other distinctions among Hirschiporus (Polyporus) abietinus, H. fusco-violaceus, and H. laricinus. *Canadian Journal of Botany* **45**, 1371-1398, doi:https://doi.org/10.1139/b67-144 (1967).
24 Stenlid, J. & Karlsson, J.-O. Partial intersterility in Heterobasidion annosum. *Mycological Research* **95**, 1153-1159, doi:https://doi.org/10.1016/S0953-7562(09)80004-X (1991).
25 Mounce, I. & Macrae, R. Interfertility phenomena in Fomes pinicola. *Canadian Journal of Research* **16c**, 354-376, doi:https://doi.org/10.1139/cjr38c-036 (1938).
26 Hallenberg, N. Speciation and distribution in Corticiaceae (Basidiomycetes). *Plant Systematics and Evolution* **177**, 93-110, doi:https://doi.org/10.1007/BF00937830 (1991).
27 Vilgalys, R. & Sun, B. L. Ancient and recent patterns of geographic speciation in the oyster mushroom Pleurotus revealed by phylogenetic analysis of ribosomal DNA sequences. *Proceedings of the National Academy of Sciences* **91**, 4599-4603, doi:https://doi.org/10.1073/pnas.91.10.4599 (1994).
28 Magasi, L. P. Incompatibility factors in Polyporus abietinus, their numbers and distribution. *Mem. of the New York Botanical Garden* **28**, 163-173 (1976).
29 Ko, K. S. & Jung, H. S. Three nonorthologous ITS1 types are present in a polypore fungus Trichaptum abietinum. *Molecular Phylogenetics and Evolution* **23**, 112-122, doi:https://doi.org/10.1016/s1055-7903(02)00009-x (2002).
30 Kauserud, H. & Schumacher, T. Regional and local population structure of the pioneer wood-decay fungus Trichaptum abietinum. *Mycologia* **95**, 416-425, doi:https://doi.org/10.1080/15572536.2004.11833086 (2003).
31 Seierstad, K. S. *et al.* Contrasting genetic structuring in the closely related basidiomycetes Trichaptum abietinum and Trichaptum fuscoviolaceum (Hymenochaetales). *Fungal Biology* **125**, 269-275, doi:https://doi.org/10.1016/j.funbio.2020.11.001 (2021).
32 Peris, D. *et al.* Large-scale fungal strain sequencing unravels the molecular diversity in mating loci maintained by long-term balancing selection. *PLOS Genetics* **18**, e1010097, doi:https://doi.org/10.1371/journal.pgen.1010097 (2022).

33 Ortiz-Barrientos, D., Counterman, B. A. & Noor, M. A. F. The Genetics of Speciation by Reinforcement. *PLOS Biology* **2**, e416, doi:https://doi.org/10.1371/journal.pbio.0020416 (2004).
34 Sæther, S. A. *et al.* Sex Chromosome-Linked Species Recognition and Evolution of Reproductive Isolation in Flycatchers. **318**, 95-97, doi:https://doi.org/10.1126/science.1141506 (2007).
35 Smadja, C. M. *et al.* Seeking signatures of reinforcement at the genetic level: a hitchhiking mapping and candidate gene approach in the house mouse. *Molecular ecology* **24**, 4222-4237, doi:https://doi.org/10.1111/mec.13301 (2015).
36 Hopkins, R., Guerrero, Rafael F., Rausher, Mark D. & Kirkpatrick, M. Strong Reinforcing Selection in a Texas Wildflower. *Current Biology* **24**, 1995-1999, doi:https://doi.org/10.1016/j.cub.2014.07.027 (2014).
37 Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655-1664, doi:10.1101/gr.094052.109 (2009).
38 Lawson, D. J., van Dorp, L. & Falush, D. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications* **9**, 3258, doi:https://doi.org/10.1038/s41467-018-05257-7 (2018).
39 Speidel, L., Forest, M., Shi, S. & Myers, S. R. A method for genome-wide genealogy estimation for thousands of samples. *Nature Genetics* **51**, 1321-1329, doi:https://doi.org/10.1038/s41588-019-0484-x (2019).
40 Rosenberg, N. A. & Feldman, M. Chapter 9. The relationship between coalescence times and population divergence times. *Modern Developments in Theoretical Population Genetics*, 130-164 (2002).
41 Ray, N. & Adams, J. M. A GIS-based Vegetation Map of the World at the Last Glacial Maximum (25,000-15,000 BP). *Internet Archaeology* **11**, doi:https://doi.org/10.11141/ia.11.2 (2001).
42 Batchelor, C. L. *et al.* The configuration of Northern Hemisphere ice sheets through the Quaternary. *Nature Communications* **10**, 3713, doi:https://doi.org/10.1038/s41467-019-11601-2 (2019).
43 Haight, J. E. *et al.* Fomitopsis mounceae and F. schrenkii—two new species from North America in the F. pinicola complex. *Mycologia* **111**, 339-357, doi:https://doi.org/10.1080/00275514.2018.1564449 (2019).
44 Kauserud, H., Svegården, I. B., Decock, C. & Hallenberg, N. Hybridization among cryptic species of the cellar fungus Coniophora puteana (Basidiomycota). *Molecular ecology* **16**, 389-399, doi:https://doi.org/10.1111/j.1365-294X.2006.03129.x (2007).
45 Geml, J., Tulloss, R. E., Laursen, G. A., Sazanova, N. A. & Taylor, D. L. Evidence for strong inter- and intracontinental phylogeographic structure in Amanita muscaria, a wind-dispersed ectomycorrhizal basidiomycete. *Molecular Phylogenetic and Evolution* **48**, 694-701, doi:https://doi.org/10.1016/j.ympev.2008.04.029 (2008).
46 Hewitt, G. The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907, doi:https://doi.org/10.1038/35016000 (2000).
47 Petersen, R. H. & Hughes, K. W. Species and Speciation in Mushrooms: Development of a species concept poses difficulties. *BioScience* **49**, 440-452, doi:https://doi.org/10.2307/1313552 (1999).
48 Pirozynski, K., Rayner, A., Brasier, C. & Moore, D. Evolutionary Biology of the Fungi. *Bioscience* **38**, doi:https://doi.org/10.2307/1310884 (1988).

49 Chase, T. E. & Ullrich, R. C. Five Genes Determining Intersterility in Heterobasidion annosum. *Mycologia* **82**, 73-81, doi:https://doi.org/10.2307/3759965 (1990).
50 Vilgalys, R. Speciation and Species Concepts in the Collybia dryophila Complex. *Mycologia* **83**, 758-773, doi:https://doi.org/10.2307/3760433 (1991).
51 Menolli, N., Jr. *et al.* Global phylogeny of the Shiitake mushroom and related Lentinula species uncovers novel diversity and suggests an origin in the Neotropics. *Molecular Phylogenetic and Evolution* **173**, 107494, doi:https://doi.org/10.1016/j.ympev.2022.107494 (2022).
52 Orr, H. A. Dobzhansky, Bateson, and the genetics of speciation. *Genetics* **144**, 1331-1335 (1996).
53 Aanen, D. K., Kuyper, T. W., Mes, T. H. & Hoekstra, R. F. The evolution of reproductive isolation in the ectomycorrhizal Hebeloma crustuliniforme aggregate (Basidiomycetes) in northwestern Europe: a phylogenetic approach. *Evolution* **54**, 1192-1206, doi:https://doi.org/10.1111/j.0014-3820.2000.tb00554.x (2000).
54 Moncalvo, J.-M. & Buchanan, P. K. Molecular evidence for long distance dispersal across the Southern Hemisphere in the Ganoderma applanatum-australe species complex (Basidiomycota). *Mycological Research* **112**, 425-436, doi:https://doi.org/10.1016/j.mycres.2007.12.001 (2008).
55 Garbelotto, M., Gonthier, P. & Nicolotti, G. Ecological constraints limit the fitness of fungal hybrids in the Heterobasidion annosum species complex. *Appl Environ Microbiol* **73**, 6106-6111, doi:https://doi.org/10.1128/aem.01135-07 (2007).
56 Le Gac, M., Hood, M. E. & Giraud, T. Evolution of reproductive isolation within a parasitic fungal species complex. *Evolution* **61**, 1781-1787, doi:10.1111/j.1558-5646.2007.00144.x (2007).
57 Mägi, R. *et al.* SCOPA and META-SCOPA: software for the analysis and aggregation of genome-wide association studies of multiple correlated phenotypes. *BMC Bioinformatics* **18**, 25, doi:https://doi.org/10.1186/s12859-016-1437-3 (2017).
58 Ament-Velásquez, S. L. *et al.* Allorecognition genes drive reproductive isolation in Podospora anserina. *Nature Ecology & Evolution* **6**, 910-923, doi:https://doi.org/10.1038/s41559-022-01734-x (2022).
59 Tsiatsiani, L. *et al.* Metacaspases. *Cell Death & Differentiation* **18**, 1279-1288, doi:10.1038/cdd.2011.66 (2011).
60 Rico-Ramírez, A. M., Gonçalves, A. P. & Louise Glass, N. Fungal cell death: The beginning of the end. *Fungal Genetics and Biology* **159**, 103671, doi:https://doi.org/10.1016/j.fgb.2022.103671 (2022).
61 Van der Nest, M. A. *et al.* Distribution and evolution of het gene homologs in the basidiomycota. *Fungal Genetics and Biology* **64**, 45-57, doi:https://doi.org/10.1016/j.fgb.2013.12.007 (2014).
62 Auxier, B. *et al.* Cytoplasmic Mixing, Not Nuclear Coexistence, Can Explain Somatic Incompatibility in Basidiomycetes. *Microorganisms* **9**, doi:https://doi.org/10.3390/microorganisms9061248 (2021).
63 Pickrell, J. K. & Pritchard, J. K. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLOS Genetics* **8**, e1002967, doi:https://doi.org/10.1371/journal.pgen.1002967 (2012).
64 Kinneberg, V. B., Lü, D. S., Peris, D., Ravinet, M. & Skrede, I. Introgression between highly divergent fungal sister species. *bioRxiv*, 2022.2008.2026.505392, doi:10.1101/2022.08.26.505392 (2022).

65 Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C. & Foll, M. Robust Demographic Inference from Genomic and SNP Data. *PLOS Genetics* **9**, e1003905, doi:https://doi.org/10.1371/journal.pgen.1003905 (2013).
66 Salonen, J. S. *et al.* Abrupt high-latitude climate events and decoupled seasonal trends during the Eemian. *Nature Communications* **9**, 2851, doi:https://doi.org/10.1038/s41467-018-05314-1 (2018).
67 Li, H. & Ralph, P. Local PCA Shows How the Effect of Population Structure Differs Along the Genome. *Genetics* **211**, 289-304, doi:https://doi.org/10.1534/genetics.118.301747 (2018).
68 Garner, A. G., Goulet, B. E., Farnitano, M. C., Molina-Henao, Y. F. & Hopkins, R. Genomic Signatures of Reinforcement. *Genes* **9**, doi:https://doi.org/10.3390/genes9040191 (2018).
69 Dettman, J. R., Jacobson, D. J., Turner, E., Pringle, A. & Taylor, J. W. Reproductive Isolation and Phylogenetic Divergence in Neurospora: Comparing Methods of Species Recognition in a Model Eukaryote. *Evolution* **57**, 2721-2741, doi:https://doi.org/10.1111/j.0014-3820.2003.tb01515.x (2003).
70 Vilgalys, R. J. & Johnson, J. L. Extensive genetic divergence associated with speciation in filamentous fungi. *Proceedings of the National Academy of Sciences* **84**, 2355-2358, doi:https://doi.org/10.1073/pnas.84.8.2355 (1987).
71 White, T. J., Bruns, T., Lee, S. & Taylor, J. Analysis of phylogenetic relationships by amplification and direct sequencing of ribosomal RNA genes. *A guide to Methods and Applications* (1990).
72 Nilsson, R. H. *et al.* The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Research* **47**, D259-D264, doi:https://doi.org/10.1093/nar/gky1022 (2019).
73 Sayers, E. W. *et al.* Database resources of the national center for biotechnology information. *Nucleic Acids Res* **50**, D20-d26, doi:https://doi.org/10.1093/nar/gkab1112 (2022).
74 Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047-3048, doi:https://doi.org/10.1093/bioinformatics/btw354 (2016).
75 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* **25**, 1754-1760, doi:https://doi.org/10.1093/bioinformatics/btp324 (2009).
76 Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, doi:https://doi.org/10.1093/gigascience/giab008 (2021).
77 Poplin, R. *et al.* Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*, 201178, doi:https://doi.org/10.1101/201178 (2018).
78 Barth, J. M. I. *et al.* Disentangling structural genomic and behavioural barriers in a sea of connectivity. *Molecular ecology* **28**, 1394-1411, doi:https://doi.org/10.1111/mec.15010 (2019).
79 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158, doi:https://doi.org/10.1093/bioinformatics/btr330 (2011).
80 Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, doi:https://doi.org/10.1186/s13742-015-0047-8 (2015).

81 Patterson, N., Price, A. L. & Reich, D. Population Structure and Eigenanalysis. *PLOS Genetics* **2**, e190, doi:https://doi.org/10.1371/journal.pgen.0020190 (2006).
82 Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**, 1530-1534, doi:https://doi.org/10.1093/molbev/msaa015 (2020).
83 Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E. & Lercher, M. J. PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R. *Molecular Biology and Evolution* **31**, 1929-1936, doi:10.1093/molbev/msu136 (2014).
84 Chan, A. H., Jenkins, P. A. & Song, Y. S. Genome-Wide Fine-Scale Recombination Rate Variation in Drosophila melanogaster. *PLOS Genetics* **8**, e1003090, doi:https://doi.org/10.1371/journal.pgen.1003090 (2012).
85 McVean, G. & Auton, A. LDhat 2.1: A package for the population genetic analysis of recombination. *Department of Statistics, Oxford, OX1 3TG, UK* (2007).
86 Paradis, E. pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics* **26**, 419-420, doi:10.1093/bioinformatics/btp696 (2010).
87 Bezmenova, A. V. *et al.* Rapid Accumulation of Mutations in Growing Mycelia of a Hypervariable Fungus Schizophyllum commune. *Molecular Biology and Evolution* **37**, 2279-2286, doi:10.1093/molbev/msaa083 (2020).
88 The UniProt, C. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research* **51**, D523-D531, doi:https://doi.org/10.1093/nar/gkac1052 (2023).
89 Geller, A. M. *et al.* The extracellular contractile injection system is enriched in environmental microbes and associates with numerous toxins. *Nature Communications* **12**, 3743, doi:https://doi.org/10.1038/s41467-021-23777-7 (2021).
90 Fitak, R. R. OptM: estimating the optimal number of migration edges on population trees using Treemix. *Biology Methods and Protocols* **6**, doi:https://doi.org/10.1093/biomethods/bpab017 (2021).
91 Wickham, H. *ggplot2: Elegant Graphics for Data Analysis.* (Springer-Verlag 2016).