

# Repeated loss of function at HD mating-type genes and of recombination in anther-smut fungi

Received: 11 April 2024

Accepted: 14 May 2025

Published online: 28 May 2025

Check for updates

Elise A. Lucotte<sup>1</sup> , Paul Jay<sup>1,2</sup>, Quentin Rougemont<sup>1</sup>, Lorelei Boyer<sup>1</sup>, Amandine Cornille<sup>3</sup>, Alodie Snirc<sup>1</sup>, Amandine Labat<sup>1</sup>, Elizabeth Chahine<sup>1</sup>, Marine Duhamel<sup>1</sup>, Alice Namias<sup>1</sup>, Jacob Gendelman<sup>4</sup>, Wen-Juan Ma<sup>4,5</sup>, Roxanne K. Hayes<sup>6</sup>, Shikhi Baruri<sup>6</sup>, Joseph P. Ham<sup>6</sup>, Michael H. Perlin<sup>6</sup>, Michael E. Hood<sup>4,7</sup>, Ricardo C. Rodríguez de la Vega<sup>1,7</sup> & Tatiana Giraud<sup>1,7</sup>

Basidiomycete fungi typically have two mating-type loci controlling mating compatibility, *HD* and *PR*, residing on different chromosomes. Loss-of-function in mating compatibility has been reported at the *PR* genes in a few heterothallic basidiomycetes, but not for the *HD* genes. In *Microbotryum* anther-smut fungi, there have been repeated linkage events between the *HD* and *PR* loci through chromosome fusions, leading to non-recombining regions. Here, we found that two sister *Microbotryum* species parasitizing *Dianthus* plants, *M. superbum* and *M. shykoffianum*, as well as the distantly related *M. scorzonerae*, have their *HD* and *PR* loci on different chromosomes, but with the *PR* chromosome fused with a part of the ancestral *HD* chromosome. In addition, recombination suppression has extended stepwise, generating evolutionary strata. In all three species, the *HD* genes lost their function in mating compatibility, natural diploid strains being often homozygous at the *HD* locus. Strains could be homozygous for a disrupted *HD2* gene, that was hardly expressed during mating. Mating tests confirmed that a single genetic factor controlled mating compatibility and that haploid strains with identical *HD* alleles could mate and produce hyphae. This study shows that a unifactorial mating-type determinism can evolve, repeatedly, from a bifactorial system, by different mechanisms.

Mating systems, controlling the degree of selfing/outcrossing, influence the level of gene flow, genetic load and adaptability in natural populations<sup>1–7</sup>. A wide diversity of mating systems occurs in nature, controlled by mating compatibility rules (e.g., separate sexes or mating types), with frequent evolutionary transitions<sup>5,8–11</sup>. In many animals and dioecious plants, different sexes are determined by sex

chromosomes<sup>12</sup>. In hermaphroditic plants, mating compatibility is often controlled by a self-incompatibility locus, which prevents mating between identical alleles at this locus<sup>13</sup>. In fungi with mating types, gamete compatibility is controlled at the haploid stage, where only cells with different mating types can successfully mate<sup>14,15</sup>. Fungal mating types are determined by one or two loci<sup>16</sup>, corresponding to

<sup>1</sup>Université Paris-Saclay, CNRS, AgroParisTech, Ecologie Société Evolution, Gif sur Yvette, France. <sup>2</sup>Center for GeoGenetics, University of Copenhagen, Copenhagen, Denmark. <sup>3</sup>Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE—Le Moulon, Gif-sur-Yvette, France. <sup>4</sup>Department of Biology, Amherst College, Amherst, MA, USA. <sup>5</sup>Department of Biology, Research group of Ecology, Evolution and Genetics, Vrije Universiteit Brussel, Brussels, Belgium.

<sup>6</sup>Department of Biology, Program on Disease Evolution, University of Louisville, Louisville, KY, USA. <sup>7</sup>These authors jointly supervised this work: Michael E. Hood, Ricardo C. Rodríguez de la Vega, Tatiana Giraud. ✉ e-mail: [elise.lucotte@cnrs.fr](mailto:elise.lucotte@cnrs.fr); [tatiana.giraud@universite-paris-saclay.fr](mailto:tatiana.giraud@universite-paris-saclay.fr)

uni- and bifactorial systems, respectively. There have been multiple transitions in fungi for changes in the number of loci controlling mating types, between one mating type, two mating types or even none, i.e., lack of discrimination for mating<sup>8,15</sup>.

Recombination is typically suppressed at loci determining sexes and mating types, which ensures proper function in these complex phenotypes, for example by linking together genes for lock-and-key functions, thus avoiding self-compatibility. In many systems, recombination suppression has furthermore extended stepwise outward from these mating-compatibility loci in plants, animals and fungi<sup>17–24</sup>, and the reason why is debated<sup>23,25,26</sup>. It has long been accepted that recombination suppression extended on sex chromosomes because it was beneficial to link sexually antagonistic genes (i.e., with an allele being beneficial in only one sex) to the sex-determining locus<sup>27</sup>. However, little evidence could be found in favor of this hypothesis<sup>28</sup>, and it cannot explain the stepwise extension of recombination suppression on fungal mating-type chromosomes, as they lack sexual antagonism and other kinds of antagonistic selection<sup>17,23,29,30</sup>.

Other hypotheses have therefore been developed, in order to account for the ubiquity of stepwise extension of recombination suppression on sex-related chromosomes: the neutral accumulation of sequence divergence proximal to the mating-compatibility locus that would reduce recombination rates<sup>31</sup>, the spread of transposable elements adjacent to the non-recombining mating-compatibility locus, together with their epigenetic marks<sup>32</sup>, deleterious mutation sheltering<sup>17,23,26,33</sup> or the selection of non-recombining fragments carrying fewer deleterious mutations than the population average<sup>26,34</sup>.

Following recombination suppression, Muller's ratchet and Hill-Robertson effect<sup>35,36</sup> lead to genomic degeneration, in terms of gene losses, weaker gene expression, lower frequencies of optimal codons and transposable element accumulation<sup>37–42</sup>. Rearrangements also rapidly accumulate<sup>17,18,43</sup>, so that recombination can be challenging to restore after the non-recombining fragments have become highly degenerated<sup>26</sup>.

Basidiomycete fungi typically have two mating-type loci controlling mating compatibility<sup>16</sup>: (i) the *PR* locus, controlling pre-mating fusion with pheromone and pheromone receptor genes, and (ii) the *HD* locus, controlling post-mating dikaryotic growth, with two homeodomain genes (*HD1* and *HD2*) whose products heterodimerize to form an active transcription factor<sup>16,44</sup>; *HD1* and *HD2* heterodimerize exclusively with non-allelic counterparts to form an active transcription factor. This mechanism prevents mating with other haploid cells carrying identical alleles at the *HD* locus, thereby promoting non-self-interactions. The *PR* and *HD* loci are usually on different chromosomes and may harbor multiple alleles, and each contains several genes linked together<sup>16</sup>. Some basidiomycetes in contrast are unifactorial, i.e., with a single segregating unit controlling mating type, which is caused most often by HD-PR linkage, as in the crop pathogen *Ustilago hordei*<sup>45</sup>, or the human pathogens *Malassezia* spp.<sup>46,47</sup> and *Cryptococcus neoformans*<sup>48</sup>, or more rarely by the loss of mating-compatibility role by the *PR* locus, as in the mushrooms *Coprinellus disseminatus*<sup>49</sup> and *Volvariella volvacea*<sup>50</sup>. A loss of mating-compatibility role by the *HD* locus in heterothallic fungi has only been observed so far in experimental mutants with their two *HD* genes fused<sup>51</sup>.

*Microbotryum* fungi (Basidiomycota) are plant-castrating pathogens, producing their spores in the anthers of diseased plants, and are therefore called anther-smut fungi. Most *Microbotryum* species specialize on one or a few host plants<sup>52–55</sup>. These fungi are model organisms in ecology, disease transmission, host specialization and the evolution of reproductive isolation, host-pathogen costructure, mating system and sex-related chromosome<sup>53,55–64</sup>. Anther-smut fungi mostly undergo selfing via intra-tetrad mating<sup>60,65,66</sup>, in which case carrying a single segregating mating-type locus is advantageous because it maximizes the odds of gamete compatibility<sup>17</sup>. These fungi have in fact repeatedly evolved HD-PR linkage, with at least nine events

of chromosomal rearrangements and recombination suppression between the two mating-type loci<sup>17,18,40,67</sup>. In addition, two non-sister *Microbotryum* species still have their mating-type loci on different chromosomes although each is linked to its centromere, resulting in the same odds of gamete compatibility as HD-PR linkage under intra-tetrad mating<sup>67,68</sup>. Recombination suppression extended stepwise beyond mating-type loci in several lineages, forming evolutionary strata, i.e., fragments displaying decreasing differentiation with increasing distance to mating-type loci along the ancestral gene order<sup>17,18,67</sup>.

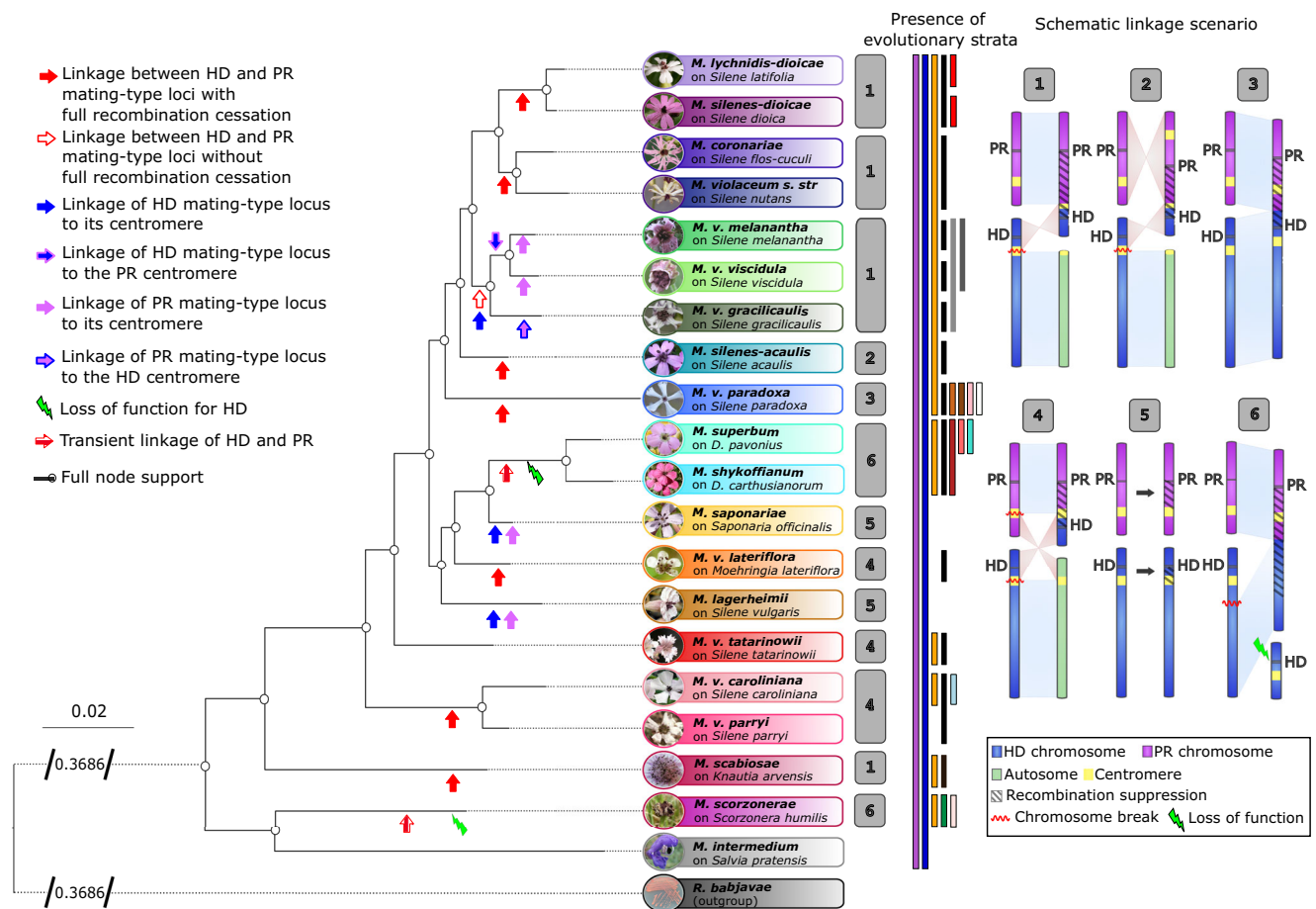
The non-recombining regions on both mating-type chromosomes in *Microbotryum* fungi display signs of genomic degeneration, as neither ever recombines in these regions. For example, polymorphism is common for haploid sporidia of one or the two mating types failing to grow in vitro in some species, likely due to deleterious mutations linked to mating-type loci<sup>60,69–71</sup>. Indeed, with time since recombination suppression, these mating-type chromosomes accumulate transposable elements, non-synonymous mutations, gene losses, rearrangements and non-optimal codons<sup>18,40–43</sup>. Transposable elements accumulate rapidly following recombination suppression, and some specific transposable element families expanded preferentially on *Microbotryum* mating-type chromosomes, such as *Helitrons* and *Copia/Ty3* elements<sup>42,72</sup>. These patterns are especially pronounced in species with linked HD and PR loci and extensive suppression of recombination, in contrast to the genomes retaining the ancestral state of unlinked mating types, that exhibit lesser severity of degenerative processes<sup>18,40–43</sup>.

Surveys of the diversity within *Microbotryum* fungi continue to yield novel insights; preliminary analyses suggested that the species parasitizing *Dianthus* plant species carry unlinked *HD* and *PR* loci, and yet display very high transposable element content compared to other *Microbotryum* species<sup>72</sup>. Three *Microbotryum* species occur in *Dianthus* plants, in sympatry, and they are not distinguishable based on morphology or host plant<sup>54,55,73,74</sup>. Focusing on two of these species, *M. superbum* and *M. shykoffianum*, our aim was to characterize their mating-type chromosomes, elucidating whether they displayed recombination suppression, and if so, to ascertain whether the mating-type loci were linked together or if each was linked to its respective centromere on separate chromosomes. We found in both species a large non-recombining mating-type chromosome resulting from the fusion of the entire ancestral PR chromosome and a part of the ancestral HD chromosome that, surprisingly, did not contain the *HD* locus. Genome comparison suggested that the translocation likely occurred via a transient stage of whole PR and HD chromosome fusion, a few rearrangements and then excision of a part of the HD chromosome arm. Remarkably, homozygosity was observed at the *HD* genes in natural populations, indicating that they were not involved anymore in mating compatibility. The *HD2* gene appeared disrupted in several strains of *M. superbum*, was homozygous in diploids, and showed low expression levels under mating conditions. This novel observation among basidiomycete fungi of loss of role in mating compatibility for *HD* genes was also found in a distant congeneric species, *M. scorzonerae*, with unlinked *PR* and *HD* loci, but homozygosity at *HD* genes and unifactoriality. These results demonstrate the potential for genetic innovation in compatibility systems that directly influence how mating system variation is achieved in nature. In addition, we found multiple evolutionary strata in the three species, i.e., young extensions of recombination suppression on the mating-type chromosomes.

## Results

### The large non-recombining PR chromosome includes a part of the ancestral HD chromosome but not the HD genes

We obtained highly contiguous genome assemblies from one diploid strain from each of two *Microbotryum* species parasitizing *Dianthus* plants: *M. superbum* (aka MvDsp2<sup>54</sup>, strain 1065) and *M. shykoffianum*



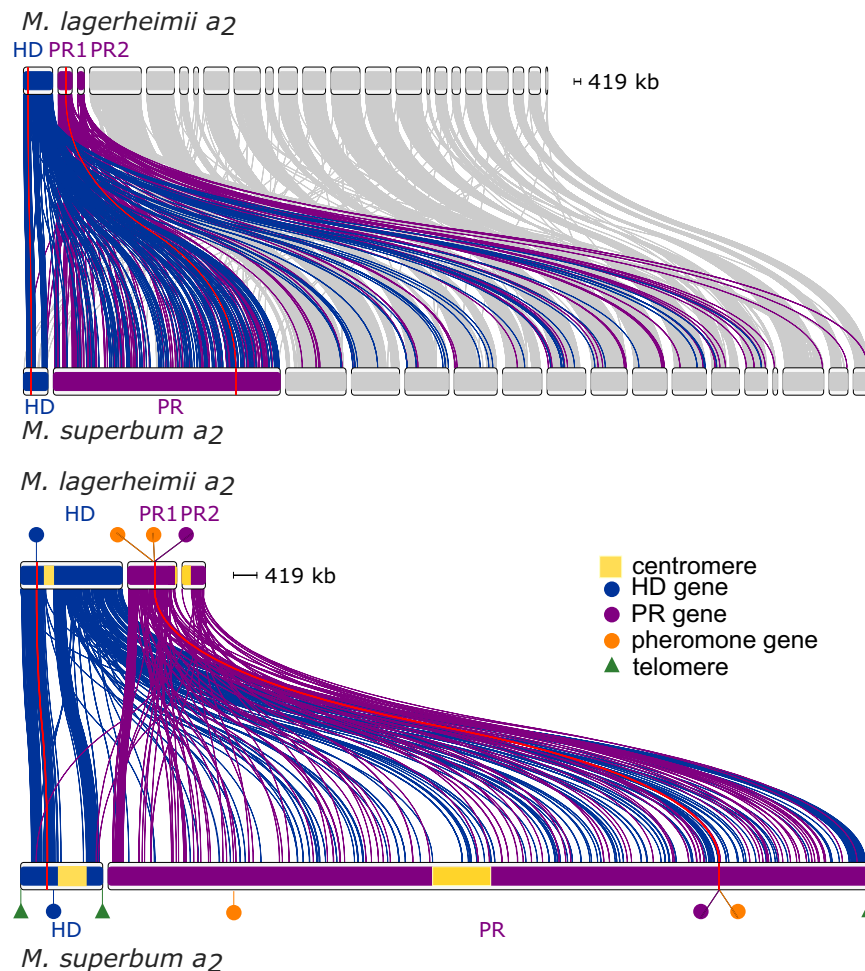
**Fig. 1 | Species tree and chromosomal arrangements of the mating-type chromosomes compared to the ancestral state.** The species tree is illustrated with pictures of diseased plants parasitized by the different *Microbotryum* species. The possible intermediate steps for the scenario 6 are depicted on Fig. 3 and S18. Arrows of different colors on the tree branches indicate mating-type locus linkage, or their linkage to centromeres, or loss of function. Colored bars at the right of the phylogeny indicate the presence of evolutionary strata suppression recombination

beyond mating-type loci. Pictures are from Michael E. Hood and Julian Woodman. The species tree was reconstructed by maximum likelihood on a concatenated alignment of 977 fully conserved single-copy genes. Diagrams based on ref. 67. Source data are provided as a Source Data file. Diseased plants photo credits: *Dianthus pavonius* and *Dianthus carthusianorum* by M. E. Hood, *Scorzonera humilis* by J. Woodman, with permission. Other photos are from Duhamel et al.<sup>42</sup>.

(aka MvDsp1<sup>54</sup>, strain c212) (Tables S1 and S2). The phylogenetic tree based on 2391 single-copy orthologs, representing 5.2 mb of nucleotide sequences, placed them as sister species (Fig. 1). In all four haploid genomes (two haploid genomes per strain, of opposite PR mating types, named  $a_1$  and  $a_2$ ), the PR and HD genes were found on different contigs (Fig. 2 and S1–S3) and with telomeric repeats identified on half of these mating-type contig edges (Figs. S4–S7), showing that the two mating-type loci are located on different chromosomes. We compared these mating-type chromosomes to those of *M. lagerheimii*, inferred to be a good proxy for the ancestral gene order of *Microbotryum* mating-type chromosomes, being highly collinear to those of the distant species *M. intermedium*<sup>17</sup>. We found the PR gene to be located in both *M. superbum* and *M. shykoffianum* on a chromosome corresponding to the fusion of the ancestral PR chromosome and a large part of the ancestral HD chromosome. The remaining part of the ancestral HD chromosome, carrying the HD gene, was found on a separate chromosome (Fig. 2 and S1–S4). This new PR chromosome also carried two pheromone genes (as in other *Microbotryum* species<sup>43,67,75</sup>, Table S3) and was non-recombining across most of its length, as it was widely rearranged between species and between  $a_1$  and  $a_2$  chromosomes in both *M. superbum* and *M. shykoffianum* (Fig. 2 and S1–S5). Only small regions were collinear, situated at both edges of the PR chromosome, corresponding to recombining pseudo-autosomal regions (PAR,

Figs. S1–S5). The centromeres of PR chromosomes were predicted within the non-recombining region for both *M. superbum* and *M. shykoffianum* (Fig. 2 and S1–S5). Telomeres were predicted at both edges of the HD chromosome in the  $a_1$  and  $a_2$  genomes of *M. superbum*, and at both edges of the PR chromosome in the  $a_1$  and  $a_2$  genomes of *M. shykoffianum* (Figs. S1–S5). At least one set of PR or HD chromosomes were thus completely assembled telomere-to-telomere in each species. The PR chromosome was not assembled telomere-to-telomere in all cases, which is frequent in *Microbotryum* fungi with non-recombining mating-type chromosomes because of high transposable element content<sup>18,40,42</sup>. The separate chromosome carrying the HD genes was very small in both species, and collinear between the two mating types as well as between species (Fig. 2 and S1–S5), suggesting ongoing recombination.

In the *M. shykoffianum* genome, the HD chromosome was much shorter in the  $a_2$  assembly than in the  $a_1$  assembly, while still carrying the HD gene (Figs. S4–S5). The mapping of reads from the reference *M. shykoffianum* diploid genome on the  $a_1$  genome assembly showed double coverage in the missing regions of the HD chromosome in the  $a_2$  genome assembly, indicating the collapse of the two homologous chromosomes in the corresponding regions in the  $a_1$  assembly (Fig. S8). The coverage was however not doubled in a small region at the edge of one pseudo-autosomal region, and the corresponding



**Fig. 2 | Synteny and rearrangements between the  $a_2$  genomes of *Microbotryum superbum* and *M. lagerheimii*, with all contigs (top) or only the mating-type chromosomes (bottom).** Contigs are drawn to scale. The *M. lagerheimii* genome represents a proxy for the ancestral state before recombination suppression and with two separate mating-type chromosomes (the PR chromosome is split into two contigs in this genome but they constitute a single chromosome<sup>17</sup>). The HD mating-type chromosome from *M. lagerheimii* is represented in blue and the PR mating-type chromosome in purple, with links to orthologous genes to the *M. superbum*

mating-type chromosomes. The positions of the HD and PR mating-type genes are indicated in red, and by red links between the two genomes, and by blue and purple circles. The pheromone genes are indicated within orange circles. Centromeres are indicated in yellow and telomeres with green triangles. The PR chromosome shows a substantial increase in size and a chaos of rearrangements. The links between *M. lagerheimii* mating-type chromosomes and *M. shykoffianum* autosomes likely represent repeats undetected in our filters, especially as many points to ends of contigs. Source data are provided as a Source Data file.

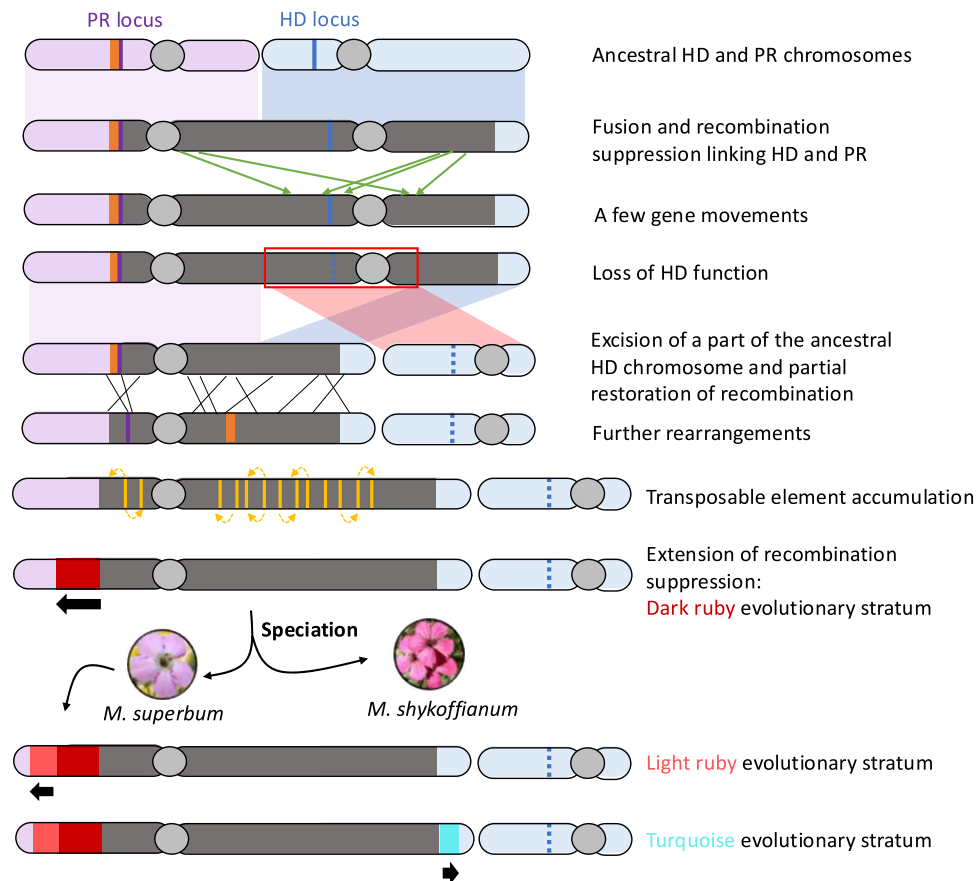
genomic region was not found in the  $a_2$  genome assembly, suggesting that this small region may be genuinely missing in the *M. shykoffianum*  $a_2$  genome (Fig. S8).

In both *M. superbum* and *M. shykoffianum*, two genes present on the PR chromosome in *M. lagerheimii* were found in the recombining region of the HD chromosome, as well as seven genes present on the HD chromosome region that is fused with PR in *M. superbum* and *M. shykoffianum* (genes with green color or green links in Figs. 3, 4, S4–S6 and Table S4). This suggests an ancient fusion between the whole PR and HD chromosomes, followed by very few genomic rearrangements, rapidly succeeded by an excision of a part of the ancestral short HD chromosome arm that would have retained the ancestral gene order (Fig. 3). Such an excision may have occurred in a single step or by chromosome fissions followed by a new fusion. The fission of the current small HD chromosome probably occurred in an ancestor of *M. superbum* and *M. shykoffianum*, as the nine genes found reshuffled between HD and PR chromosomes were located on the same chromosomes in the two species.

Nucleotide polymorphism data further indicated that the PR chromosome was mostly non-recombining and that the HD genes were

not linked to the PR gene in the *Microbotryum* fungi parasitizing *Dianthus* plants. We used SNPs identified across 149 genomes sequenced with the Illumina technology, corresponding to strains collected on different plant individuals, from various *Dianthus* species in western Europe, on which the two *Microbotryum* species studied here occur. We found high levels of LD (linkage disequilibrium) across most of the PR chromosome (mean  $r^2 = 0.68$ ), while LD levels were lower between HD and PR chromosomes (mean  $r^2 = 0.49$ ) and within the HD chromosome (mean  $r^2 = 0.54$ ). These values were higher than those within autosomes (mean  $r^2 = 0.37$ ), but were similar to those between the HD chromosomes and autosomes (mean  $r^2 = 0.47$ ), thus not indicating linkage (Fig. S9).

The high per-gene synonymous divergence ( $d_s$ ) between alleles on the  $a_1$  and  $a_2$  PR mating-type chromosomes (Figs. 4, 5) further confirmed the lack of recombination along most of the PR chromosome, while the HD chromosome appeared recombining. Indeed, because selfing in *Microbotryum* anther-smut fungi yields high homozygosity levels,  $d_s$  values are mostly zero between alleles in recombining regions such as autosomes, PARs of the PR chromosome and in the HD chromosome (Figs. 4, 5). In contrast, the rearranged region of



**Fig. 3 | Inferred scenario for the evolution of the mating-type chromosomes in *Microbotryum superbum* and *M. shykoffianum*.** Pictures of the diseased plants parasitized by the two species are shown. Recombination suppression is indicated by colors corresponding to the different evolutionary strata. The rearrangements

between ancestral HD and PR chromosomes before the HD fragment excision are indicated by green links. See an alternative scenario in Fig. S18. Diseased plants photo credits: *Dianthus pavonius* and *Dianthus carthusianorum* by M. E. Hood.

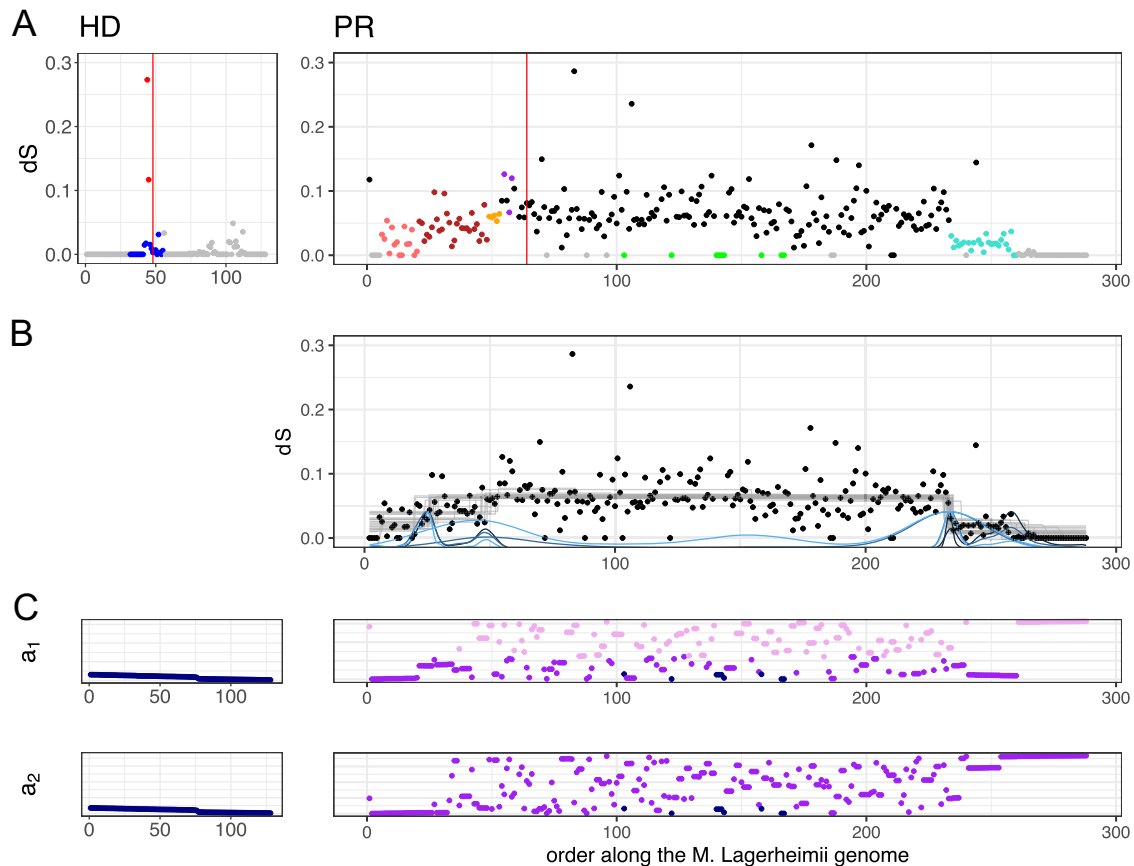
the PR chromosome displayed high  $d_s$  values in both *M. superbum* and *M. shykoffianum* (Figs. 4, 5). Following previous studies on *Microbotryum* fungi, the region initially linking the PR and HD genes was called the black stratum (Figs. 4, 5), although this stratum does not link the PR and HD mating-type loci any more in *M. superbum* and *M. shykoffianum* due to the proposed fission of the HD-containing chromosome arm (Fig. 3). This black stratum evolved before the divergence between *M. superbum* and *M. shykoffianum*, as 73% of the genes showed full trans-specific polymorphism between these two species (Fig. 6): the alleles clustered according to  $a_1$  or  $a_2$  mating type rather than species in gene genealogies, indicating recombination suppression before speciation. We found footprints of previously inferred evolutionary strata (i.e., the blue, purple and orange strata; Figs. 4, 5), that were shown to have evolved around the HD and PR genes ancestrally in the *Microbotryum* clade<sup>17,18,23</sup>.

### Gene disruption, homozygosity and segregation analyses revealed loss-of-function of HD genes in mating-type determination

The alleles of the two tightly-linked HD genes (*HD1* and *HD2*) appeared functional in the two haploid reference genomes of *M. shykoffianum* and in the  $a_1$  genome of *M. superbum*, but not in the  $a_2$  reference genome of *M. superbum*. The *HD2* gene indeed appeared disrupted, being separated into two different gene models in the  $a_2$  genome in *M. superbum*, with an early stop codon and a new start codon (Fig. 7A). This disruption deleted the homeodomain of the protein (Fig. 7A). We analyzed the HD gene expression using a different *M. superbum* strain

(6P/6D, Table S2) than that of the reference genome (1065). The genome assembly of the strain used for gene expression (6P/6D) was homozygous for the disrupted *HD2* allele (Fig. 7A), while it was able to cause normal disease symptoms in plants. In this strain, the two predicted gene models at the *HD2* locus were expressed at much lower levels than the *HD1* gene during mating conditions (Fig. 7B, C), when HD genes should play a role in initiating dikaryotic growth. In contrast, both HD genes are evenly expressed under mating conditions in *M. lychnidis-dioicae* (Fig. 7B, C). Analyses of non-synonymous substitutions based on gene genealogies with all available species (Fig. 1) indicated significant positive selection in the disrupted *HD2* gene, suggesting a possible neofunctionalization or relaxed selection (significant test for selection intensification:  $K = 22.64$ ;  $p = 0.001$ ,  $LR = 11.37$ ).

Mating tests confirmed that HD genes have lost their mating-type determination function in *M. superbum*. We isolated the four meiotic products of 12 full tetrads from a strain collected on *D. pavonius* in the same site as the reference strain (Table S2). *Microbotryum* produces linear tetrads allowing assessment of first versus second division segregation<sup>60,66</sup>. We tested whether each of these haploid cell lineages would mate with tester haploid strains from one of the tetrads that carried  $a_1$  or  $a_2$  PR alleles, respectively, as determined by PCR<sup>77</sup>. The typical situation in heterothallic basidiomycete fungi is that (i) only haploid cells carrying different PR alleles can fuse and (ii) only haploid cells carrying different HD alleles can produce hyphae when mated. In all 12 tetrads but one, two of the four meiotic products conjugated with the  $a_1$  tester and the other two



**Fig. 4 | Differentiation and rearrangements between the mating-type chromosomes in *Microbotryum superbum*.** **A** Per-gene synonymous divergence ( $d_s$ ) plotted along the ancestral gene order (taking as proxy the gene order along the *M. lagerheimii* mating-type chromosomes), the HD chromosome on the left and the PR chromosome on the right; the red vertical lines indicate the positions of the HD and PR genes; the two red points on the HD chromosome are the HD genes; the PR gene is not plotted here as the alleles are too differentiated. The points are colored according to their evolutionary stratum assignment: turquoise, light and dark ruby for the *M. superbum* specific evolutionary strata, black for the evolutionary stratum shared by *M. superbum* and *M. shykoffianum*, blue, orange and purple for the

ancient evolutionary strata shared by most *Microbotryum* species. The pseudo-autosomal regions (PARs) are in gray. The green points correspond to genes that were ancestrally on the PR chromosome but found in the HD chromosomes in *M. superbum* or reciprocally. **B** Change-point analysis identifying changes in mean  $d_s$  levels. Blue and black curves are the density of the posterior distributions of the changepoint locations. Thin gray lines are the average  $d_s$  of the inferred strata. **C** Rearrangements compared to the ancestral gene order illustrated by plotting the gene rank in the current gene order as a function of the gene rank in the ancestral gene order. Source data are provided as a Source Data file.

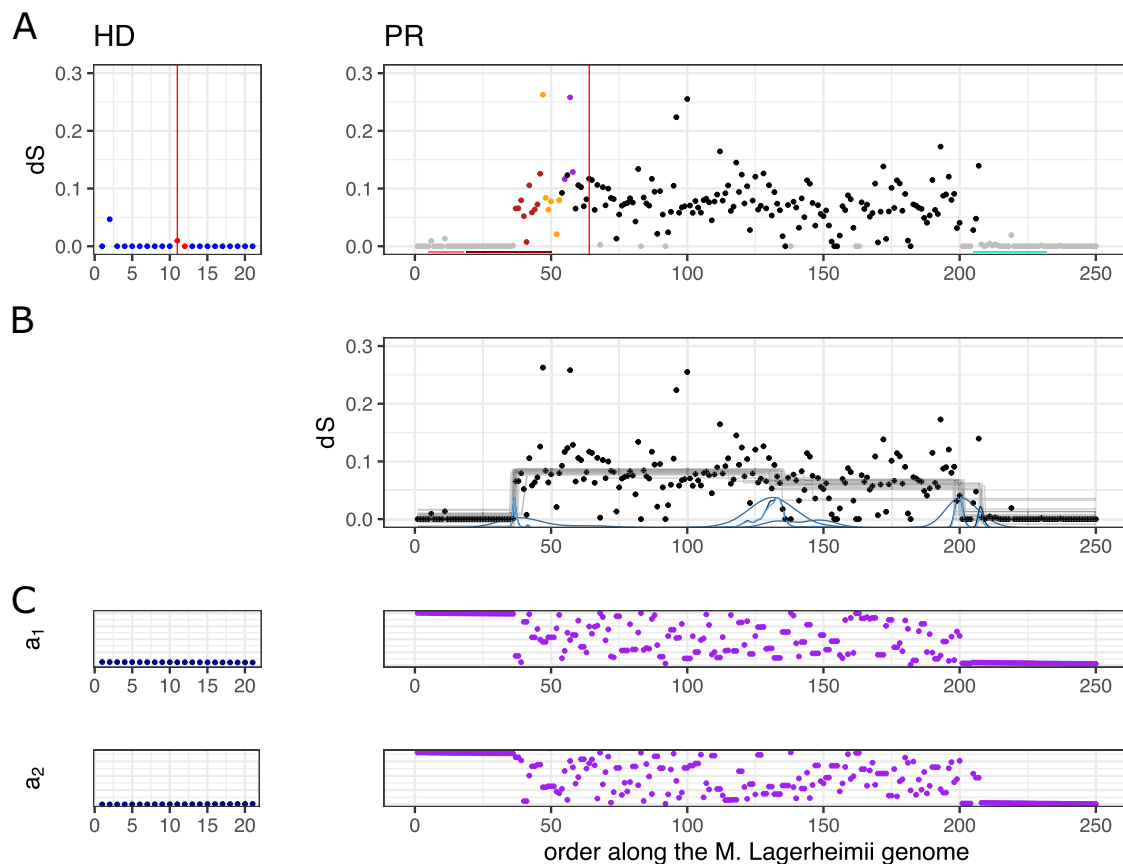
with the  $a_2$  testers, as expected, with PR alleles segregating at the first meiotic division, consistent with linkage of the PR gene to its centromere shown above. Supporting the view that the HD genes no longer have a role in mating-type determination, all meiotic products compatible for conjugation with a given mating tester also produced hyphae (Table S5). If HD genes were still involved in mating compatibility, only half of conjugated pairing would produce hyphae due to independent segregation/assortment of PR and HD mating type loci, as genomic assemblies show that they are unlinked.

Additionally, PCR tests confirmed that these 12 meiotic tetrads (Table S2) resulted from a field-collected diploid parental genotype that contained identical HD alleles, which indicates the lack of mating-type role of HD genes. We indeed designed primers that would specifically amplify only the allele present in the  $a_1$  or  $a_2$  genome of the *M. superbum* 1065 strain (Table S6). The primers for the PR gene segregated as expected in tetrads, with the  $a_1$  and  $a_2$  alleles segregating at the first meiotic division (Fig. S10). In contrast, the HD1 and HD2 genes gave PCR products only with the primer pair designed to amplify the alleles in the  $a_2$  genome of the *M. superbum* 1065 strain, and in all four meiotic products of all tetrads, indicating homozygosity of the diploid parent (Fig. S10). Sanger sequencing of the amplicons confirmed that all meiotic products were identical at the HD genes. Therefore,

heterozygosity at the HD genes was not required in these strains for successful mating and hyphae production. We also inoculated *D. chinensis* in the greenhouse and showed these homozygous pairs could cause disease (5 out of 24 inoculated plants, as typical under inoculations), with completely normal symptoms, i.e., anthers full of dark and smutty spores and ovary reduction. In one tetrad, two meiotic products produced hyphae alone without any tester strains (Fig. S10) and the PCR indicated that these sporidia contained the two PR alleles, likely following abnormal meiotic segregation (i.e., non-disjunction), as previously reported in *Microbotryum* fungi (back then called *Ustilago violacea*)<sup>78</sup>. This sporidial line however did not produce disease upon plant inoculation.

We further pulled out two tetrads from each of six additional field-collected *M. superbum* strains (Table S2). The sequencing of PCR amplicons of HD and PR genes indicated that two strains were homozygous for the HD genes, confirming that mating can occur between cells with identical HD alleles; the other strains were heterozygous. Across tetrads, the PR alleles were not always associated to the same HD alleles, which confirms that the HD and PR loci are unlinked in this species (Table S7).

The HD1 and HD2 coding sequences in *M. shykoffianum* c212 were not disrupted but were 100% identical between the  $a_1$  and  $a_2$  genomes



**Fig. 5 | Differentiation and rearrangements between the mating-type chromosomes in *Microbotryum shykoffianum*.** **A** Per-gene synonymous divergence ( $d_S$ ) plotted along the ancestral gene order (taking as proxy the gene order along the *M. lagerheimii* mating-type chromosomes), the HD chromosome on the left and the PR chromosome on the right. The red vertical lines indicate the positions of the HD and PR genes; the two red points on the HD chromosome are the HD genes; the PR gene cannot be plotted here as the alleles are too differentiated. The points are colored according to their evolutionary stratum assignment: black and dark ruby for the evolutionary strata shared by *M. superbum* and *M. shykoffianum*, and blue,

orange and purple for the ancient evolutionary strata shared by most *Microbotryum* species. The pseudo-autosomal regions (PARs) are in gray. The position of the *M. superbum* specific evolutionary strata are indicated by color bars at the bottom. **B** Change-point analysis identifying changes in mean  $d_S$  levels. Blue and black curves are the density of the posterior distributions of the changepoint locations. Thin gray lines are the average  $d_S$  of the inferred strata. **C** Rearrangements compared to the ancestral gene order illustrated by plotting the gene rank in the current gene order as a function of the gene rank in the ancestral gene order. Source data are provided as a Source Data file.

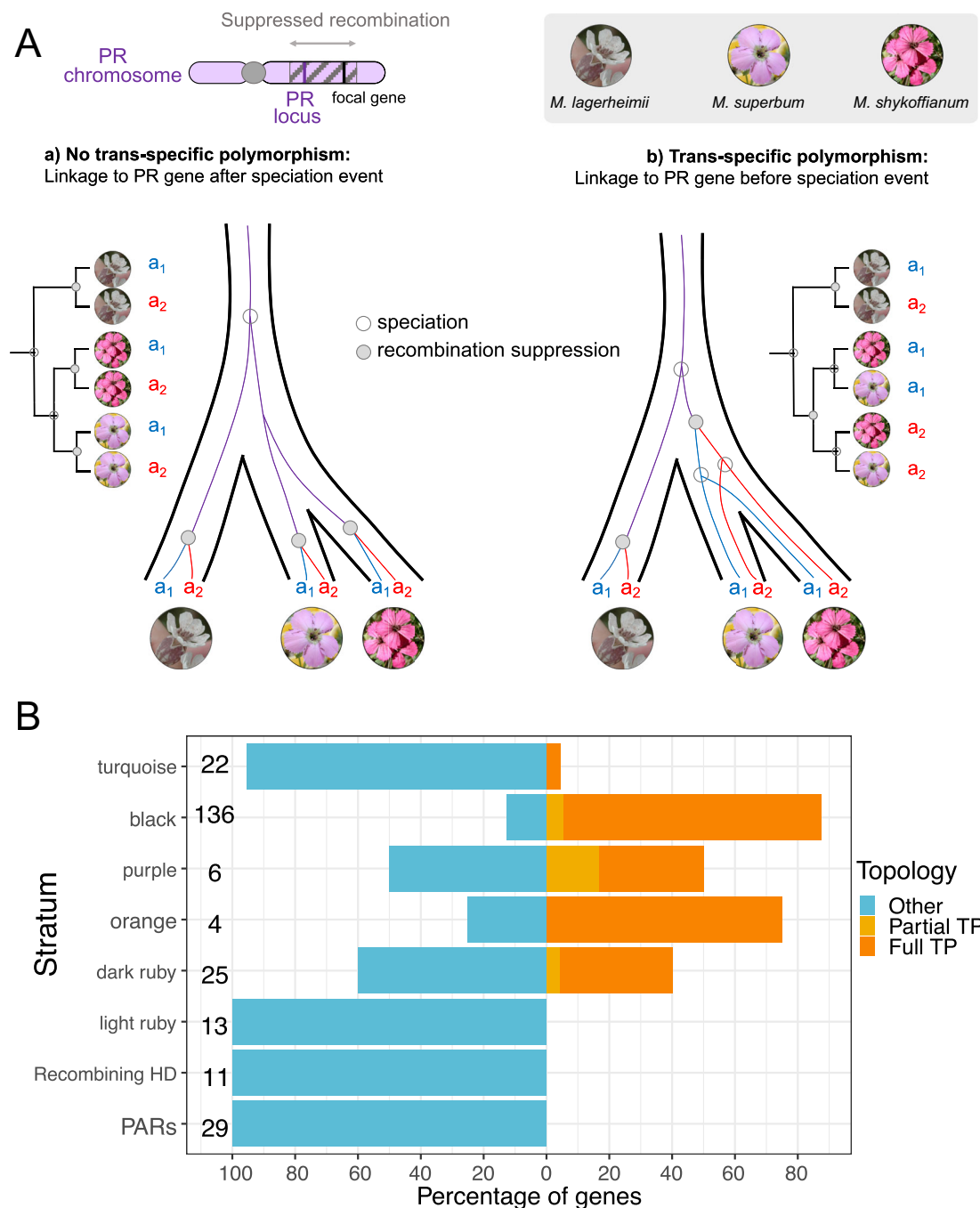
of the reference strain (Fig. 7A), which has been checked in HiFi reads, while the PR alleles showed high differentiation and trans-specific polymorphism as expected. This strain had been sequenced as a diploid collected from a diseased plant, which reinforces the view that the HD genes can be homozygous in *Microbotryum* strains from *Dianthus* species, and that, therefore, HD genes are not involved anymore in mating compatibility.

Analysis of the HD genes in the genomes sequenced with the Illumina technology confirmed that they could be homozygous in natural strains. We could assemble de novo three of the genomes, originating from strains collected on two different *D. pavonius* plants (strains 235 and 267, belonging to *M. superbum*) and a *D. carthusianorum* plant (strain c165, belonging to *M. shykoffianum*), thus allowing to obtain reliably phased alleles. We checked that the genomes carried both PR alleles, i.e., that we did obtain diploid genomes after culturing. In all three strains, we found the HD genes to be completely homozygous.

### Loss of function of HD genes in *M. scorzonerae* and genomic rearrangements

We also identified a loss of HD gene function and homozygosity of field-collected samples in a distant species, *M. scorzonerae* (Fig. 1). Oxford Nanopore sequencing of two haploid genomes isolated from

the same meiotic tetrad and with opposite PR alleles (Table S2), revealed identical sequences for both haploid genomes at HD1 and HD2 genes. The HD2 gene contained an additional intron compared to the functional *M. lychnidis-dioicae* HD2 gene (Fig. 7A) and was shorter due to a deletion in the highly variable N-terminal heterodimerization domain, which is challenging to align across species. The N-terminal heterodimerization domain of HD2 was also shorter in *M. scorzonerae* than in other species. The  $a_1$  and  $a_2$  genomes for *M. scorzonerae* did not assemble as well as for *M. superbum* and *M. shykoffianum* (Table S1). For example, the N50 statistic for *M. scorzonerae* was half of the other assembly N50 statistics. Nevertheless, we identified two contigs in the  $a_1$  *M. scorzonerae* assembly that carried regions corresponding to the ancestral PR and HD chromosomes, and also parts of *M. lagerheimii* autosomes, indicating that rearrangements occurred between the ancestral mating-type chromosomes (Fig. S11A). The contigs carrying parts of the ancestral PR and HD chromosomes were collinear between the  $a_1$  and  $a_2$  *M. scorzonerae* genomes, indicating ongoing recombination, with the exception of a relatively small non-recombining region around the PR locus (Figs. S11B and S12A–C). The HD and PR genes were not in the same contigs. This suggests that there was a transient stage, as in *M. superbum* and *M. shykoffianum*, where the HD and PR genes were linked by chromosomal fusion of ancestral HD and PR chromosomes,



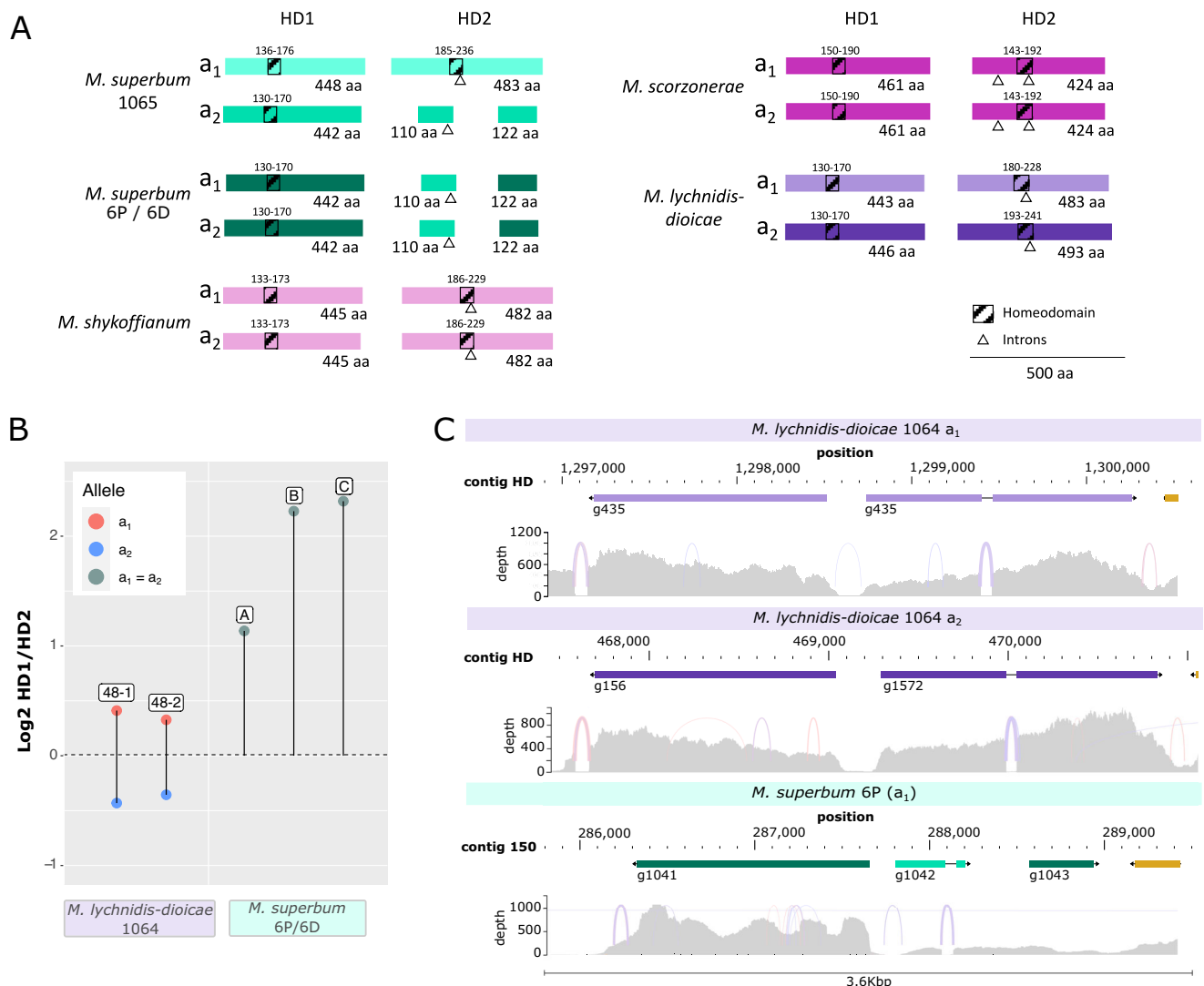
**Fig. 6 | Trans-specific polymorphism (TP) of genes present along the mating-type chromosomes in *Microbotryum superbum*, considering also *M. shykoffianum* and *M. lagerheimii*.** **A** Schematic representation of trans-specific polymorphism in genes completely linked to mating-type loci by suppressed recombination, allowing to assess whether recombination cessation occurred before or after speciation, here between *M. lagerheimii*, *M. superbum* and *M. shykoffianum*. **B** Percentage of genes within each evolutionary stratum with the following genealogy pattern: full trans-specific polymorphism (i.e., across both *M. superbum* and *M. shykoffianum*), partial trans-specific polymorphism (i.e., with only

one allele showing clustering by mating type and the other by species, or with one allele missing, i.e., not informative), and “others” which show no evidence of trans-specific polymorphism. The total number of genes analyzed per stratum is indicated at the left of each barplot. The PAR (pseudo-autosomal region) category corresponds here to the two PARs of the PR chromosome pooled together. Source data are provided as a Source Data file. Diseased plants photo credits: *Dianthus pavonius* and *Dianthus carthusianorum* by M. E. Hood. Other photos are from Duhamel et al.<sup>42</sup>.

as well as autosomes, followed by chromosomal rearrangements and subsequent loss-of-function of *HD* genes in mating-type determination. However, we could not detect telomeres in any of these contigs.

Additionally, we performed mating tests between meiotic products of one tetrad from each of 22 *M. scorzonerae* natural diploid strains (Table S2), which again indicated compatibility segregating at a

single mating-type locus: when cells were compatible for *PR*-determined conjugation, they also produced hyphae in 100% of the cases, despite the probable lack of *HD-PR* linkage, and they could produce hyphae even when bearing the same *HD* alleles (Fig. S10D). This strongly suggests a loss-of-function of *HD* genes in mating-type determination also in *M. scorzonerae*.



**Fig. 7 | Expression and structure of the HD genes in *Microbotryum* fungi.** **A** HD1 and HD2 gene and protein structures and functional domains (N-terminal to C-terminal), in two strains of *M. superbum* (1065 and 6P/6D), and one strain of each of *M. shykoffianum*, *M. scorzonerae* and *M. lychnidis-dioicae*. Protein sequence length is given in number of amino-acids (aa). Intron positions are shown with a black triangle and homeodomain positions by a dashed rectangle. Identical amino-acid sequences are shown with identical colors. **B** Expression analysis of HD genes. Log<sub>2</sub> ratio of the HD1-to-HD2 TPM (transcripts per million of mapped reads) in mated conditions of two *Microbotryum* species (*M. lychnidis-dioicae* and *M. superbum*). Balanced expression generates Log<sub>2</sub> ratios close to zero (dashed line). Data is shown for both alleles of *M. lychnidis-dioicae*. The diploid *M. superbum* strain used

was homozygous at the HD gene. Normalized read counts were averaged for both genes predicted at the locus homologous to HD2 in *M. superbum* 6P. The different replicates are shown for each species (48-1 and 48-2 for *M. lychnidis-dioicae* and A, B and C for *M. superbum*). **C** Genome tracks showing, in each of *M. lychnidis-dioicae* a<sub>1</sub> (top), *M. lychnidis-dioicae* a<sub>2</sub> (middle) and *M. superbum* (bottom), the HD locus at the top and the per-position read mapping depth at the bottom, along genome position, in a representative RNAseq sample per species in mating conditions. Arcs represent splitted RNAseq reads showing perfect splicing of the inferred introns. Note that RNAseq reads were not stranded, so the strand-of-origin cannot be identified. RNAseq read mapping was displayed with jbrowse 2. Source data are provided as a Source Data file.

### Progressive recombination suppression in the mating-type chromosome: evolutionary strata

We found evidence of four young evolutionary strata specific to the mating-type chromosome of *M. superbum* or of *M. shykoffianum*, therefore resulting from recent extensions of recombination cessation. Indeed, when plotting the per-gene synonymous divergence between mating-type chromosomes in *M. superbum* 1065 along the ancestral gene order (taking *M. lagerheimii* gene order as a proxy<sup>18,67</sup>), we found a region with significantly lower *d<sub>s</sub>* values at the edge of the non-recombining region, while *d<sub>s</sub>* values were significantly different from the neighboring pseudo-autosomal region (Fig. 4 and S13). We called this region the “turquoise” evolutionary stratum following previous color names given to young evolutionary strata in *Microbotryum* fungi. This turquoise stratum is currently situated within the non-recombining region of the PR mating-type chromosome (Figs. S4,

S5 and S14), while still remaining much less rearranged than the rest of the non-recombining region, as expected for a young evolutionary stratum. This genomic region had in contrast zero *d<sub>s</sub>* values in *M. shykoffianum* and has remained in the pseudo-autosomal region (Fig. 5 and S4, S5, S15). Additional evidence for the evolution of the turquoise stratum post-dating speciation between *M. shykoffianum* and *M. superbum* is the lack of trans-specific polymorphism (Fig. 6). A single gene displayed full trans-specific polymorphism in the turquoise stratum and it was situated just at the limit with the black stratum, suggesting that it may actually belong to the black stratum. Altogether, these findings show that the turquoise stratum has moved into the non-recombining region in *M. superbum* after its divergence from *M. shykoffianum*, thus constituting a young evolutionary stratum (Fig. 3).

We found footprints of another young evolutionary stratum, in *M. superbum*, on the other side of the PR mating-type chromosome in the

ancestral chromosomal arrangement. The  $d_s$  plot indeed showed high  $d_s$  values in the left edge of the PR chromosome (Fig. 4 and S13), in the region that we called “ruby”.

A changepoint analysis confirmed the inference of young evolutionary strata in *M. superbum*, with the turquoise and ruby strata showing significantly lower mean  $d_s$  value than the black stratum, with discrete changes in mean values (Fig. 4). The changepoint analysis even indicated that the ruby stratum was divided into two distinct sub-strata of different ages, that we called the “dark” and “light” ruby strata, with significant different mean  $d_s$  values and different locations: the dark ruby stratum was rearranged, while the light ruby stratum was still at the edge of the chromosome (Fig. 4, S4 and S13). The black, turquoise, light and dark ruby strata had significantly different  $d_s$  mean values than their nearby strata (Fig. S13).

The dark ruby stratum was only partly present in *M. shykoffianum* (Fig. 5), suggesting either an independent origin in the two lineages or an ancestral stratum that would have partly restored recombination in *M. shykoffianum*. This latter hypothesis is supported by the lack of detected change point in *M. superbum* within the dark ruby stratum, indicating it evolved in a single step (Fig. 4). Moreover, trans-specific polymorphism was found for the genes from the right part of dark ruby stratum that have high  $d_s$  values in the two species (Figs. 4, 5).

In *M. scorzonerae*, genes around the PR locus were highly rearranged between mating-type chromosomes and compared to the *M. lagerheimii* gene order (Figs. S11, S12), indicating ancient recombination suppression, which was further supported by high  $d_s$  values (Figs. S16, S17). This non-recombining region extended farther than the purple and orange strata that are shared among all *Microbotryum* species, indicating an extension of recombination suppression specific to *M. scorzonerae*. We called this extension the emerald stratum, and it encompassed not only a part of the ancestral PR chromosome, but also a small fragment of the ancestral HD chromosome, from the middle of the MC03B *M. lagerheimii* contig: this fragment indeed displays elevated  $d_s$  values in *M. scorzonerae* (Fig. S16) and is currently located near the PR locus because of the inter-chromosomal rearrangements (Figs. S11, S17). In addition, we detected footprints of a younger evolutionary stratum, called quartz, with very low  $d_s$  values (Fig. S17), being collinear between  $a_1$  and  $a_2$  PR mating-type chromosomes but having moved from the PAR to the middle of the non-recombining region in the  $a_2$  PR chromosome (Fig. S11). The non-zero  $d_s$  values extend a bit farther towards the PAR than the translocated fragment, suggesting that the quartz stratum may be larger than this translocated fragment.

We used the divergence time between *M. lychnidis-dioicae* and *M. silenae-dioicae* as a calibration point (0.42 million years<sup>79</sup>) to estimate the age in million years (MY) of the various evolutionary strata and their confidence intervals (CI), based on autosomal single-copy genes. We estimated the following ages: 1.55 MY for the black stratum (CI 1.21–1.92), 1.25 MY for the dark ruby stratum (CI 0.75–2.23), 0.28 MY for the light ruby stratum (CI 0.16–0.45), 1.61 MY for the emerald stratum (CI 1.03–2.33) and 0.06 MY for the quartz stratum (CI 0–0.16).

### Degeneration of the PR mating-type chromosomes

In both *M. superbum* and *M. shykoffianum*, the PR mating-type chromosome shows chaos of rearrangements between the  $a_1$  and  $a_2$  genomes, and in comparison to the ancestral gene order, taking as proxy the *M. lagerheimii* PR chromosome (Fig. 2, S1–S5). As expected in non-recombining regions, the PR chromosome has also accumulated transposable elements (TEs), as seen by the much higher TE content in the  $a_1$  and  $a_2$  PR chromosomes in *M. superbum* (65% and 67% of bp, respectively) compared to its autosomes (mean 35% of bp; Crawford-Howell's  $t$ -test = 2.97,  $df$  = 15,  $p$  = 9.45e-3) (Fig. 8). Some TE families seem to have preferentially expanded (Fig. 8). For example, the PR chromosome displays higher density (percentage of base pairs occupied by TEs) in the following repeated elements compared to

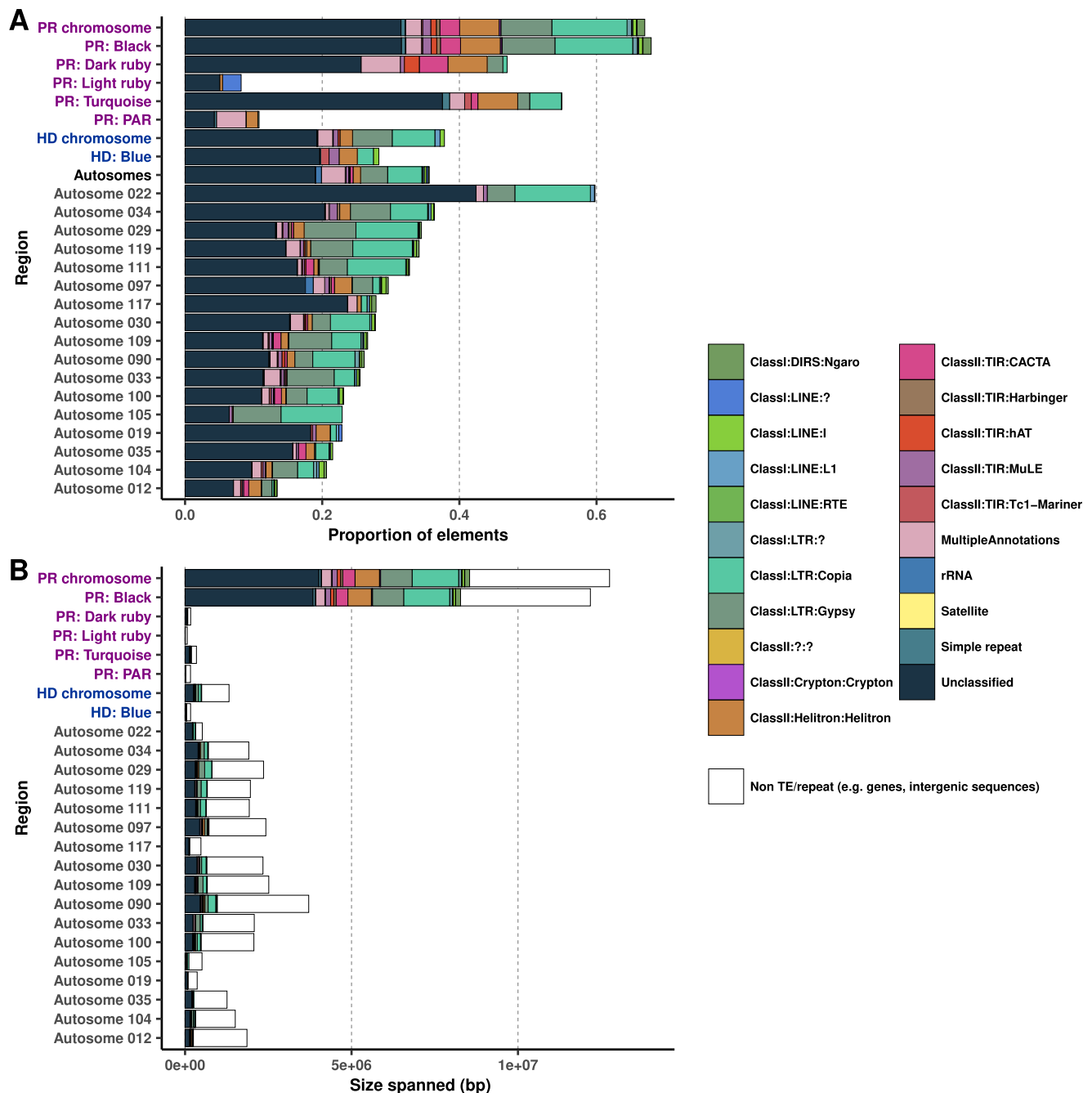
autosomes: MuLe, CACTA elements, hAT, *Helitrons* and *RTE* (Table S8). In contrast, other TE families did not show higher density of TEs in the PR chromosome compared to autosomes: c1-Mariner, LTR Copia and LINE L1 (Table S8). As a result of TE expansion, the PR chromosome has become much larger than its ancestral recombining state (Fig. 8; 12,930,922 vs 942,259 bp). In contrast, the HD chromosome did not have an elevated TE load (mean 36% of bp, Fig. 8), in agreement with the inference of continued recombination in this chromosome. Overall, the genome size of *M. superbum* and *M. shykoffianum* (ca. 45–48 Mb, Table S1) was larger than that of other *Microbotryum* species, typically more around 30 Mb, as *M. scorzonerae*.

### Discussion

A wide variety of breeding systems have been reported across fungi, including universal haploid compatibility, a single mating-type locus or two mating-type loci<sup>8,14</sup>. The role of PR and HD genes in mating-type determination and their location on different chromosomes is ancestral in basidiomycetes, corresponding to a bifactorial system, i.e., with two different mating-type loci, each encompassing multiple genes linked together<sup>16</sup>. Having two distinct loci controlling mating compatibility can favor outcrossing, while having a single mating-type locus is advantageous under selfing: one mating-type locus yields 50% compatibility among gametes of a given diploid individual across tetrads versus only 25% with two mating-type loci<sup>58</sup>. There have been multiple independent transitions across the Basidiomycota from bifactorial to unifactorial systems, i.e., with a single genetic locus determining mating compatibility. Unifactorial compatibility can occur through two main pathways: linkage of PR and HD loci or the loss of function at one of the mating type loci<sup>8</sup>. Linkage between mating-type loci has been observed in *Ustilago hordei*<sup>45</sup>, *Malassezia* spp.<sup>46,47</sup>, *Microbotryum* spp.<sup>18,67</sup>, as well as in other basidiomycetes<sup>8,80–83</sup>.

Loss-of-function at the PR locus has been observed in several basidiomycete fungi, including *Pholiota nameko*<sup>84</sup>, *Coprinnellus disseminatus*<sup>49</sup> and others<sup>8</sup>, but never at the HD locus in natural populations of heterothallic basidiomycete fungi so far. We demonstrated here HD loss-of-function in mating-type determinism by a combination of various lines of evidence: (i) gene disruption in several natural strains and reduced expression level of the HD2 gene, (ii) the possibility of mating between haploid cells with identical HD alleles (and some even homozygous for disrupted HD2 alleles), with the ability of hyphal growth and plant infection for dikaryons homozygous at the HD locus, and (iii) the occurrence of diploid strains in nature that are homozygous at the HD genes. This shows that unifactoriality, i.e., the segregation of only two mating types in progenies, can be convergently acquired through a variety of different genomic pathways in selfing fungi, for which it is beneficial. Furthermore, we found convergent events of HD gene loss-of-function after a transient HD-PR chromosome fusion, in two very distantly related *Microbotryum* fungi. Altogether, this sheds light on the power of selection and on the repeatability of evolution, and shows that similar phenotypes can be achieved repeatedly, via similar or different mechanisms.

In addition, the possibility of HD loss-of-function has important implications for our understanding of the sexual cycle and gene functions in basidiomycete fungi. The two tightly linked genes of the HD locus, HD1 and HD2, are indeed thought to be essential for completing the life cycle, as they heterodimerize to activate the dikaryotic growth and can only do so between different allelic forms at the two genes (using  $b_1$  and  $b_2$  to denote different alleles and where haploid genomes typically carry either HD1  $b_1$  with HD2  $b_1$  or HD1  $b_2$  with HD2  $b_2$ , dikaryotic growth is triggered by HD1  $b_1$  with HD2  $b_2$  or HD1  $b_2$  with HD2  $b_1$ ). Uniting the two compatible variants in a haploid genome could allow dikaryotic growth without the need to mate with a different haplotype, as shown experimentally with a mutant carrying a chimeric homeodomain protein in the mushroom *Coprinus cinereus*<sup>51</sup>. The need for HD heterodimers may be bypassed for dikaryotic growth,



**Fig. 8 | Scaffold size and transposable element (TE) content in the genome of *Microbotryum superbum* a<sub>1</sub>.** The TE content is given as a proportion (A) or absolute content in base pair (bp, B). The relative proportions of transposable element classes in each scaffold are represented by the sizes of the colored bars. The PR chromosome is presented as a whole (“PR chromosome”), and then separated into its different evolutionary strata (blue, black, turquoise, light and dark ruby) and the

pseudo-autosomal regions (PAR). The blue stratum on the HD chromosome is also separated from the rest of the chromosome. Only scaffolds longer than 300 kb are displayed. The average proportion of the various repetitive elements observed on autosomes is shown on panel A (“Autosomes”). Statistical analyses were performed using only scaffolds larger than 300 kb, i.e., using  $n = 24$  scaffolds. Source data are provided as a Source Data file.

as shown in the invasive Californian death cap mushroom, *Amanita phalloides*, that can mate as a haploid homothallic species, despite the two *HD* genes in haploid genomes being unable to form a heterodimer<sup>85</sup>. Another possibility may be as in the homothallic basidiomycete yeast *Cystofilobasidium capitatum*, in which the *HD1/HD2* heterodimer has likely been replaced by a *HD2* homodimer<sup>86</sup>. This hypothesis would explain the positive selection detected on the new *HD* coding sequences. However, in our species, strains could be homozygous for a disrupted *HD2* allele without any homeodomain, while it is the *HD2* homeodomain that binds DNA, and the

heterodimerization domain of *HD2* in *M. scorzonerae* was much shorter than in other species. In addition, the *HD2* gene was found little expressed in mating conditions in *M. superbum*, so it may be that HD genes are not needed any more for dikaryotic growth, or that only a single form of the *HD1* gene can be sufficient, which would represent a unique finding. The *HD1* gene indeed always seems complete and was found expressed at similar levels as in *M. lychnidis-dioicae* with functional *HD* genes, but was found homozygous in several natural isolates in *M. superbum*, *M. shykoffianum* and *M. scorzonerae*. Knocking out the whole genes with CRISPR-Cas9 in future studies would help verify the

molecular function. The non-zero level of expression of the *HD2* gene despite the early stop codon suggests incomplete degradation of transcripts by NMD (nonsense-mediated decay)<sup>87,88</sup>.

Given the loss of the *HD* role in mating-type determination, it was intriguing to discover that the PR chromosome had incorporated a fragment of the ancestral HD chromosome and was non-recombining throughout most of its length in *M. superbum* and *M. shykoffianum*. Indeed, the large non-recombining regions in *Microbotryum* fungi are generally partly due to the linkage of *HD* and *PR* genes, even if further extension of the non-recombining region has repeatedly evolved<sup>18,67</sup>. In the *Microbotryum* species parasitizing *Dianthus* and *Scorzonera* plants analyzed here, a transient stage seems to have occurred, with *HD-PR* linkage by fusion of the whole PR and HD ancestral chromosomes, followed by very few rearrangements, and then excision of a part of the ancestral HD chromosome, likely once *HD* has lost its function. This hypothesis is supported by our finding of genes belonging to the ancestral HD-containing chromosome arm being now located on the PR chromosome in each of *M. superbum* and *M. shykoffianum*. The transient stage with *HD* and *PR* loci linked thanks to chromosomal fusion would be advantageous under selfing, as it increases the odds of compatibility among gametes<sup>17</sup>. After loss of function of the *HD* genes, the linkage of *HD* and *PR* loci would not be beneficial, and chromosome fission could occur. The fission must have occurred rapidly after the fusion, before the HD chromosome fragment that separated lost its ancestral order. An alternative hypothesis could be that the fusion occurred only between the ancestral PR chromosome and a part of the ancestral HD chromosome without an intermediate step of whole-chromosome fusion (Fig. S18), but this fusion would provide no fitness advantage in terms of gamete compatibility; it would seem surprising that a chromosomal fusion occurred between the ancestral PR chromosome and a fragment of the ancestral HD chromosome, but without the *HD* genes. Fusion of chromosomes to the mating-type chromosome has been suggested to possibly occur for deleterious mutation sheltering, just as evolutionary strata; it would be a striking coincidence that this precisely occurred for an HD chromosome arm, but it cannot be excluded. The reshuffling of the “green” genes between the HD and PR chromosomes could have been facilitated by a *PR-HD* fusion, but it could also have happened without it. Reversion to unlinked mating-type chromosomes has been suggested in *Malassezia*-related fungi<sup>47</sup>, but both mating-type loci kept their functions, such that reversion was likely driven by changes in mating systems. Here, the chromosome fission may have followed the loss of function of the *HD* genes, as this loss of function would offset the benefit of *HD-PR* linkage under selfing mating systems.

This study also provided further support for the existence of three young, species-specific evolutionary strata associated with the mating type loci in *Microbotryum* fungi. Indeed, *M. superbum*, *M. shykoffianum* and *M. scorzonerae* displayed high differentiation between mating-type chromosomes and rearrangements within the non-recombining region, but with lower differentiation and lower levels of gene order reshuffling in young evolutionary strata than the rest of the non-recombining region. This confirms that stepwise recombination suppression can evolve in fungi despite the lack of sexual antagonism<sup>17,23</sup>. The evolutionary cause for such stepwise and repeated recombination suppression events may be a selection for a lower load or deleterious-mutation sheltering<sup>17,23,26,33,34</sup>, neutral fixation of inversions or genetic differences impairing recombination<sup>31</sup>. The proximal mechanism of recombination suppression may be the observed movements of fragments from the pseudo-autosomal region into the non-recombining region, forming the young strata, as previously reported for the red stratum in *M. lychnidis-dioicae*<sup>17</sup>. However, such movements can also be a consequence, rather than a cause, of recombination suppression.

As expected, the older part of the non-recombining region on the PR mating-type chromosome shows signs of degeneration, with chaos

of rearrangements and high TE load. Some TE families have preferentially expanded, i.e., *Copia*, Ty3 (*Gypsy*) and *Helitrons*, corresponding to the same families that repeatedly expanded in mating-type chromosomes of other *Microbotryum* species<sup>42</sup>.

Given such accumulated load that results from mutational degeneration, it has been suggested that reversion towards recombination could be selected for under the lower-loaded-sheltering hypothesis, unless dosage compensation evolves<sup>26,34</sup>. Our scenario actually proposes an early reversal of recombination suppression along a part of the ancestral HD chromosome. The chaos of rearrangements on both mating-type chromosomes in the oldest evolutionary strata, however, reinforces the view that rearrangements accumulate following recombination suppression such that it seems challenging to restore recombination when load has accumulated<sup>26</sup>. Recombination restoration would be even more difficult if the proximal mechanism of recombination suppression is the movement of PARs into non-recombining regions, as observed for young evolutionary strata. Indeed, recombination restoration would then require the exact same movement back, with exactly the same breakpoints, which seems highly implausible. Our findings suggest that recombination restoration may occur when rearrangements are not too extensive, as evidenced by (i) the footprints of recombination suppression followed by chromosome separation and recombination restoration in the HD chromosome arm, (ii) the reversion of a part of the dark ruby stratum to a recombining state in *M. shykoffianum*. However, recombination restoration now seems impossible in the non-recombining region of the PR chromosome, given the chaos of rearrangements there in both mating-type chromosomes.

In conclusion, our study brings insights into the evolution of breeding systems and sex-related chromosomes, reporting cases of loss-of-function of *HD* mating-type genes in fungi, and fusion of ancestral mating-type chromosomes without linkage of mating-type genes. Our findings further reinforce the view that recombination suppression can extend progressively in organisms without sexual antagonism.

## Methods

### Study systems

*Microbotryum* fungi, parasitizing *Dianthus* and *Silene* species, castrate plants by producing their spores in the anthers of host plants and aborting the ovaries. Diseased plants are identifiable by dark-colored spore masses in their flowers. Most *Microbotryum* fungi are specialized on a single host species and most host species harbor a single *Microbotryum* species, especially in the *Silene* genus<sup>53–55,89</sup>. One exception is the clade of *Microbotryum* fungi parasitizing wild carnation relatives in the pink family (*Dianthus* genus, Caryophyllaceae), in which several closely-related and sympatric *Microbotryum* species parasitize multiple *Dianthus* plant species, with host range overlapping to some extent<sup>54,55,73,74</sup> (Fig. 1). Three species have been formally described so far in this group<sup>74</sup>, but they are challenging to recognize based on their host plant or morphology. We therefore assigned strains to species based on the published sequences of the gene barcodes used for describing the species (EF1- $\alpha$  and beta-tubulin<sup>54,74</sup>; see below). *Microbotryum scorzonerae* parasitizes flowers of *Scorzonera humilis*, in the Asteraceae family<sup>90</sup>. Unlike other *Microbotryum* species, *M. scorzonerae* does not cause spore formation on anthers alone. It instead causes spore formation within the entire floral head (Fig. 1).

### Genomic analyses

**Strains and DNA extraction for long-read sequencing and species identification.** We extracted high molecular weight DNA for long-read sequencing as previously described<sup>17,18</sup>. For the strains isolated from *Dianthus* hosts, DNA sequencing based on PacBio (Pacific Bioscience) HiFi long-read sequencing was performed at the GenoToul sequencing facility (Toulouse, INRAE, France) in 2021. Size selection and

circularization was performed by the sequencing platform. The sequencing instrument of HiFi reads was a Sequel II. We isolated haploid sporidia of opposite mating types from a single meiosis (tetrad) from the strain 1065, collected in July 2011 on *Dianthus pavonius*, in Italy (coordinates 44.189-7.688, near the Garelli Refugium in the Alps) by Michael Hood, Janis Antonovics and Emme Bruns. Based on preliminary analyses of SNPs, we also identified a strain belonging to another *Microbotryum* species parasitizing *Dianthus* plants: strain C212 (CZ\_D24), collected in 2016 on *Dianthus carthusianorum*, in the Czech Republic, coordinates 49.639759-14.201153 (Bohemia, The Vltava river basin, to the north of Zduchovice) by Klara Koupilova. This later strain had been cultivated on Petri dishes at the haploid stage after spreading diploid teliospores that each underwent meiosis; we therefore sequenced a mixture of haploid sporidia resulting from multiple meioses from a single diploid individual. The mixture is thus equivalent to a diploid genome. Cultivation of sporidia was performed on PDA (potato dextrose agar). The samples were collected in Italy, which is not a party of the Nagoya protocol, and in the Czech Republic, which does not regulate access to its genetic resources in relation to the Nagoya Protocol. Using the best BLAST hits of the EF1- $\alpha$  and  $\beta$ -tubulin genes against the reference sequences<sup>54,74</sup>, we assigned the strain 1065 to *M. superbum* (aka MvDsp2) and the strain C212 to *M. shykoffianum* (aka MvDsp1)<sup>54,74</sup>.

We extracted genomic DNA from the strain used for the RNAseq experiments and sequenced its genome with the ONT Oxford Nanopore technology; this strain was isolated from the same population as *M. superbum* 1065 and was confirmed to belong to the same species using the barcode genes as explained above. For extracting DNA, we slightly modified a previous protocol<sup>91</sup>. The 6P strain was grown on nutrient rich media (yeast extract peptone dextrose, YPD) at room temperature for 4 days, and then in 4 ml of liquid yeast extract peptone dextrose (YPD) overnight at room temperature on a shaker at 200 rpm. Cultures were pelleted by centrifugation at 13,000 rpm for 1 min and the supernatants discarded. The pellet was washed with 1 ml of dH<sub>2</sub>O, pelleted and resuspended in 0.5 ml of lysis buffer (0.5 M NaCl, 0.01 M EDTA at pH 8.0, 0.2 M Tris-Cl at pH 7.5, 1% SDS) and 0.3 g of sterile 0.5 mm glass beads were added and quick spun for 5 s. To this mixture, we added 250  $\mu$ l PCI (25:24:1 v/v phenol: chloroform: Isoamyl alcohol) and vortexed for 4.5 min. The cells were centrifuged at 13,000 rpm for 3 min. Using wide bore tips, the upper phase was collected to a clean tube. Again 0.25 ml PCI was added and centrifuged at 13,000 rpm for 3 min. The upper phase was collected to a clean tube and 1 ml of ice cold 100% ethanol was added and centrifuged for 5 min. Next, the wash step was repeated with 70% ice cold ethanol twice and the supernatant was discarded. The DNA pellet was vacuum-aspirated and dried on for 10 min at room temperature and resuspended in 40  $\mu$ l TE (pH 8.0). Then, the genomic DNA was treated with PureRec RNaseA enzyme (ZymoResearch, 1 mg/1 ml) for 15 min with 2  $\mu$ l at room temperature to remove any RNA contamination. Finally, the purification was done with Genomic DNA clean and concentrator kit from Zymo Research. DNA concentration and purity were assessed using NanoDrop One<sup>c</sup> (Thermo Fisher Scientific, Waltham, MA).

For the *M. scorzonerae* strain, obtained in 2018 from Cefn Cribwr, Wales, by Julian Woodman, sporidia of opposite PR mating types were isolated from a single meiosis (tetrad). Although a party to the Nagoya Protocol, the UK does not regulate access to its genetic resources. From cultures on PDA, extracted haploid genomic DNA was used to generate long-read sequences with MinION (Oxford Nanopore Technologies, FLO-MIN106 SpotON Flow Cell Mk I R9 Version), and short-read Illumina sequencing.

**RNAseq from *M. superbum*.** For optimizing gene prediction and analyzing *HD* gene expression, we sequenced RNA from one of the focal species, *M. superbum*. We used for RNAseq another strain (6P/6D) from the same *D. pavonius* population as the strain 1065 (coordinates

44.191, 7.685, near the Garelli Refugium in the Italian Alps); we isolated a<sub>1</sub> (6P) and a<sub>2</sub> (6D) haploid strains, and used them separately as well as pooled in mating experiments. For extractions from haploid cells grown on PDA, we used the Zymo RNA Extraction Kit DirectZol RNA Miniprep plus, Cat. No. R2072 (Zymo Research, Irvine, California). Prior to RNA extractions, each haploid fungal strain was grown on two PDA Petri dishes at 28 °C for 24 h. Extractions were performed on fungal cells from the two PDA Petri dishes for each haploid sample replicate. Fungal mating was accomplished in the following steps. Haploid *M. superbum* a<sub>1</sub> and a<sub>2</sub> cells were grown on YPD agar for 4 days prior to the mating assay. Haploid cells were then suspended in distilled and autoclaved water before adjusting to 10<sup>9</sup> cells/mL. Suspended fungal cells were combined in equal proportions before being spotted in 50  $\mu$ l spots on water agar Petri dishes. Plates were left to incubate at 13 °C for 48 h. Resulting cells were visualized under a light microscope to confirm the occurrence of mating via the formation of conjugation tubes between cells. Under these conditions we typically observed about 30% of the cells engaged in mating. Extractions using the Zymo RNA extraction kit DirectZol RNA Miniprep plus, Cat. No. R2072, were performed on mated fungal tissue that was aggregated from the water agar petri dishes for each mated sample. Paired-end RNAseq of three biological replicates per condition (mated, haploid a<sub>1</sub> and haploid a<sub>2</sub>) were performed by CD Genomics (Shirley, New York).

**Analysis of *HD* gene expression.** For the gene expression analysis on *HD* genes in *M. superbum*, we used the culture condition in which *HD* genes are expected to be expressed, i.e. the three *M. superbum* replicates under mating conditions at 48 h. The expression of *HD* genes in *M. lychnidis-dioicae* was analyzed as a control of expectations for functional *HD* genes. For *M. lychnidis-dioicae*, we used available data, i.e., two replicates under mating conditions at 48 h<sup>91</sup>. To check whether *HD* genes were expressed, we mapped the paired-end reads with STAR 2.7.11a<sup>92</sup>, explored with samtools<sup>93</sup>, quantified with featureCounts from subread package with -p -t CDS -M -fraction options<sup>94</sup> and displayed the SNP coverage tracks with jbrowse 2<sup>95</sup>. Counts were transformed to transcripts per million (TPM) using the following formula:

$$TPMi = \frac{Counts_i \div Length_i(in\ kb)}{\sum(Counts \div Length(in\ kb)) \div 1e6}$$

To compare the relative expression of *HD* genes, we performed a log<sub>2</sub> fold-change of TPM of *HD1* over TPM of *HD2*. We averaged the TPM of the two genes in the locus homologous to *HD2* in *M. superbum* a<sub>1</sub> (6P).

**Genome assemblies and gene prediction.** HiFi-based assemblies were obtained with the subreads with QV 20 or higher. For *M. superbum* 1065, we sequenced separately two haploid genomes, considered to be of compatible mating types based on their position in the isolated linear tetrad. As mating types typically segregate in the first meiotic division in *Microbotryum*, cells at opposite edges in the linear tetrad are indeed usually of opposite mating types<sup>60,76</sup>. We assembled these two genomes separately with Canu 2.2<sup>96</sup> using the fastq reads files as input, and setting the -pacbio-hifi parameter. Additionally, we sequenced the a<sub>1</sub> haploid strain (strain 6P) used for RNAseq with a R10.4.1 flow cell in a PromethION P24 instrument. ONT reads were basecalled with dorado (ont-doradod-for-promethion v7.1.4) on super-accurate mode. We assembled the reads with flye 2.9.4-b1799<sup>97</sup> using -g 48.7 m -i 3 --nano-raw as input parameters. For the strain collected on *D. carthusianorum*, we sequenced sporidial cultures expected to correspond to mixtures of meiotic products from a diploid individual. Therefore, we used the software Hifiasm 0.16.1-r375<sup>98</sup>, which can separate haplotypes, with default parameters. We performed quality assessment of the resulting assembly by computing N50 and

L50 statistics and by assessing genome completeness with BUSCO v5.4<sup>99</sup> using the basidiomycota\_odb10 database as reference genes.

The two alternative mating types of *M. scorzonerae* were sequenced on Amherst laboratory ONT platform. Reads longer than 1000 bp of raw ONT were kept for genome assemblies. Trimmed reads were assembled with Canu v1.7b with default parameters<sup>96</sup>. The resulting assemblies were polished with Pilon v1.22 (<https://github.com/broadinstitute/pilon/wiki>) with default parameters, using Illumina 150 bp paired-end reads belonging to the two mating types of the same strain. To that end, short-read sequences were checked for quality using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), trimmed with Trimmomatic<sup>100</sup> using default parameters. Then, three rounds of polishing were run iteratively with Pilon, with Illumina short reads aligned to the polished assemblies obtained from the previous round and indexed at each iteration using bwa-mem2<sup>101</sup>. We then assessed genome completeness with BUSCO v5.4<sup>99</sup> using the basidiomycota\_odb10 as reference genes.

Repeated regions and low complexity DNA sequences were identified using RepeatModeler<sup>102</sup>. The genome was then softmasked using RepeatMasker open-4.0.9<sup>103</sup> with the de novo detected repeats as well as custom libraries of candidate transposable elements (TEs) from our previous studies<sup>67</sup> and existing fungal TEs from RepBase<sup>104</sup>. To minimize the rates of false negatives in the gene prediction steps, we excluded from the softmasking TEs that were labeled as “unknown” as their TE status was uncertain. We assessed whether the proportion of base pairs occupied by transposable elements (TEs) from each family differed between the PR chromosome and the autosomes using Crawford and Howell’s *t*-tests (two-sided)<sup>105,106</sup>. This test was the most appropriate here as we compared a single case (the PR chromosome) against a small set of controls (the autosomes). Additionally, we calculated the *z*-score to quantify the deviation of TE content in the PR chromosome relative to the autosomes.

Genes were then predicted on the soft-masked genome using the software Braker version 2.1.6<sup>107</sup> which uses Diamond, ProHint 2.6.0, GeneMark 4 and AUGUSTUS 3.4.0. To help gene prediction, we combined RNAseq data from *M. superbum* (this study), *M. intermedium*<sup>40</sup> and *M. lychnidis-dioicae*<sup>91</sup>, and a set of highly conserved protein sequences previously identified across multiple *Microbotryum* species. This set of proteins contains fully conserved single-copy genes not overlapping with the mating-type chromosomes from 18 published *Microbotryum* genomes, all predicted with the same pipeline<sup>17,18,67,68</sup>. The Illumina RNAseq data from 12 *M. superbum*, together with 5 *M. intermedium*<sup>40</sup> and 17 *M. lychnidis-dioicae*<sup>91</sup> samples previously obtained (Table S9), were processed with fastp<sup>108</sup>, aligned to the reference genome with the software GSNAP and GMAP<sup>109</sup>, combined and sorted with samtools<sup>93</sup> and filtered with Augustus<sup>110</sup> to keep only the unique best match (-uniq option). Following Braker’s documentation, we performed gene predictions separately for RNAseq data and protein databases. For the gene prediction, we ran five rounds of gene prediction and kept the run showing the highest BUSCO score. We eventually combined the protein and RNAseq gene predictions using TSEBRA<sup>111</sup> and assessed the quality of the final gene set with BUSCO using the basidiomycota\_odb10 as reference using a dedicated workflow (<https://github.com/QuentinRougemont/EASystrata>)<sup>112</sup>. De novo detection of transposable elements (TEs) was performed as described previously<sup>42</sup>.

**Orthogroup construction.** We performed orthology reconstruction using a set of 47 published genomes from our previous work<sup>40,67</sup> and including the newly assembled genomes for three species studied here. The *Rhodospiridium babjavae* genome was used as an outgroup<sup>67</sup>. We identified transposable elements and predicted genes in each genome using Braker v2.1.6 with the protein database built as explained above and we used in addition, for the *M. superbum*, *M. intermedium* and *M. lychnidis-dioicae* genomes, RNAseq data from

these species (*n* = 12, 5 and 17 conditions for *M. superbum*, *M. intermedium* and *M. lychnidis-dioicae*, respectively, Table S9). When multiple isoforms were present, we kept the gene prediction with the longest transcript. This extensive dataset was then fed to OrthoFinder<sup>113</sup> with default parameters to reconstruct orthogroups.

**Telomere and centromere prediction.** To detect telomeres, we counted the occurrence of the telomere-specific motif “TTAGGG” and its reverse complement “CCCTAA” within 1 kb windows along the contigs. Peaks occur when telomeres are present as they are constituted by multiple repeats of this motif. Centromeres were predicted by searching for the centromeric repeats previously described in *Microbotryum* fungi<sup>43,67</sup>.

**Identification of mating-type chromosomes and mating-type determining genes.** We identified the contigs belonging to the mating-type chromosomes (Table S10) by (i) detecting the contigs carrying *PR* and *HD* genes using BLASTN 2.6.0<sup>114</sup> (Table S3) with the gene sequences from *M. lagerheimii* as queries<sup>77,115</sup>, (ii) plotting the ortholog links for all contigs of our focal species against the reference genome of the *M. lagerheimii* strain 1253 (*a*<sub>1</sub>)<sup>17</sup> using Rideogram<sup>116</sup>, and (iii) plotting the ortholog links using Rideogram between the *a*<sub>1</sub> and *a*<sub>2</sub> genomes of our focal species. The pheromone genes were identified using BLASTN 2.6.0<sup>114</sup>, with the gene sequences<sup>75</sup> as queries.

We also generated circos plots using circlize<sup>117</sup>. In circos plots, we oriented contigs in the direction involving the fewest inversions and keeping the recombining regions of the mating-type chromosomes, called pseudo-autosomal regions (PARs), at the edges and non-inverted between mating-type chromosomes.

**Analysis of selection on *HD* genes.** For analyzing selection on *HD* genes, we aligned the predicted coding sequences of the *HD2* gene for all the *Microbotryum* species with available genomes (Fig. 1) using the software MACSE v2<sup>118</sup>. The *HD2* gene that was predicted as two separate gene models in the *a*<sub>2</sub> genome of *M. superbum* was manually merged to form one sequence, conserving the alignment with the other species. We used the RELAX method from the hyphy package<sup>119</sup> to detect relaxed or intensified selection on the branch of the *HD2* gene in the *a*<sub>2</sub> genome of *M. superbum*, compared to the *HD2* gene sequence in the *a*<sub>1</sub> and *a*<sub>2</sub> *M. lagerheimii* branches.

**Detection of evolutionary strata.** We used as a proxy of the ancestral state the chromosomal arrangement and gene order of *M. lagerheimii*, as done in previous studies<sup>17,68</sup>. In *M. lagerheimii*, *PR* and *HD* loci are located on different chromosomes<sup>17</sup>. For each species, synonymous divergence (*d*<sub>s</sub>) was computed from the alignment of *a*<sub>1</sub> and *a*<sub>2</sub> allele sequences, obtained with macse<sup>120</sup>, using the yn00 v4.9f program of the PAML package<sup>121</sup> and plotted using the ggplot2 library of R<sup>122</sup>. Pseudo-autosomal, recombining regions were identified as regions (i) with null synonymous divergence on *d*<sub>s</sub> plots, as expected in these highly selfing fungi, (ii) with the same gene order between mating types and (iii) the same gene order as in *M. lagerheimii*, as assessed on circos plots. Non-recombining regions were identified in contrast as regions with non-null synonymous divergence on *d*<sub>s</sub> plots, and being most often rearranged compared to the ancestral gene order in synteny plots. In order to objectively identify the limits of evolutionary strata, we looked for changes in mean *d*<sub>s</sub> along the ancestral gene order along the mating-type chromosomes using a changepoint analysis performed using the R package mcp<sup>123</sup>. This analysis relies on Bayesian regression to infer the location of changes in means of the *d*<sub>s</sub> values.

**Trans-specific polymorphism.** Trans-specific polymorphism (i.e., the clustering of alleles per mating type across species) can be used to study the age of recombination suppression linking genes to mating-

type loci. Indeed, as soon as a gene is fully linked to a mating-type locus, its alternative alleles will remain associated to alternative mating types, even across speciation events. Mutations will therefore accumulate independently in alleles associated with alternative mating types. In a genealogy, alleles will therefore be grouped according to mating type rather than according to species. The node at which the alleles associated with the alternative mating types diverge indicates the time of recombination cessation<sup>17,18,23</sup>.

We performed codon-based alignment with macse v2.05<sup>118</sup> of one-to-one orthologs of genes ancestrally located on mating type chromosomes (416 genes). Gene trees were obtained with IQ-TREE 1.6.1<sup>124</sup> with default parameters. Tree topologies were assessed using a homemade awk command line, available on demand. Gene trees were midpoint rooted with gotree<sup>125</sup> and subtrees containing the *M. shykoffianum* and *M. superbum*  $a_1$  alleles recovered with nw\_clade from Newick utils<sup>126</sup>. If the subtree contains only two tips it means that  $a_1$  alleles were recovered as monophyletic (i.e., displaying trans-specific polymorphism), we repeated the procedure for the  $a_2$  alleles and counted the number of trees showing full (both alleles recovered as monophyletic clades) or partial (one only  $a_1$  or  $a_2$  were recovered as monophyletic) trans-specific polymorphism. Gene trees in which neither  $a_1$  nor  $a_2$  were recovered as monophyletic were collated as “other” topologies, often following the species phylogeny (i.e., alternative alleles from the same species clustered together, *M. superbum* and *M. shykoffianum* being recovered as sister lineages).

A large part of the HD chromosome is collapsed in the *M. shykoffianum* genome assembly, which means that it was not correctly separated between  $a_1$  and  $a_2$  genomes during the assembly. Therefore, we could not run trans-specific polymorphism analyses in this region because the genes could only be predicted in the  $a_1$  genome assembly. However, the assembly collapse means that the  $a_1$  and  $a_2$  alleles were highly similar for the genes in this genomic region. Mapping the reads to the genome confirmed that there were no SNPs in this region. There should therefore be no trans-specific polymorphism in the HD chromosome.

**Dating of evolutionary strata.** For each gene assigned to an evolutionary stratum in either of the three focal species, we identified the orthologous genes in the mating-type chromosomes of four reference species (eight haploid assemblies): *M. lagerheimii* and *M. intermedium*, displaying an ancestral-like gene order<sup>17</sup>, as well as *M. lychnidis-dioicae* and *M. silenes-dioicae*, two species for which the divergence time has been estimated<sup>79</sup>. We filtered out orthologous groups with a copy number different from one in any of the haploid genomes. For the dark ruby stratum, we only kept the genes assigned to the stratum in both *M. shykoffianum* and *M. superbum*. We identified a minimum of 6 (dark ruby stratum) and a maximum of 31 (black stratum) orthologous groups per stratum. We obtained codon-based alignments for each orthologous group with translatorX<sup>127</sup> using muscle<sup>128</sup> as protein aligner. For each stratum, we concatenated the codon-based alignments of its orthologous groups (alignment length range: 9528 - 75651 base-pairs) and reconstructed an ultrametric tree with the least square dating method<sup>129</sup> under the TPM + F + G4 substitution model, as implemented in iqtree2 (version 2.2.6<sup>124</sup>). We used the divergence time between *M. lychnidis-dioicae* and *M. silenes-dioicae* as a calibration point (0.42 million years<sup>79</sup>).

**Illumina sequencing: strains, methods and analysis.** A total of 165 samples were collected on different diseased *Dianthus* plant individuals, from various species (*D. pavonius*, *D. seguieri*, *D. carthusianorum*, *D. hyssopifolius*, *D. albinus*, *D. superbus*, *D. caryophyllus*, *D. deltoideus* and *Stellaria holostea*) in several locations across Europe. For each strain, diploid spores from anthers of a single flower were spread on a Petri dish containing potato dextrose agar (PDA) and ampicillin, and let grow at 23 °C under artificial light. Spores in a given flower are

produced by a single diploid individual<sup>130</sup>. When placed on a nutritive media, the diploid spores undergo meiosis and the resulting haploid sporidia then replicate clonally. There were thousands of haploid sporidia harvested on the Petri dishes representing the numerous meiotic products of a single diploid individual. However, lethal alleles are sometimes found linked to one mating type in some strains, so that some mixtures of recovered sporidia can include only the  $a_1$  or  $a_2$  mating-type chromosome<sup>60,69</sup>. Cells were harvested from the medium and stored at -20 °C until DNA extraction, using the Macherey-Nagel NucleoSpin Soil® kit following the manufacturer's instructions. The quality of the DNA extracted was then assessed by measuring the ratio of 230/260 and 280/260 nm with a NanoDrop 2000 spectrophotometer (Thermo Scientific). A Qubit 2.0 fluorometer was used to measure DNA concentration. Preparation of DNA libraries for sequencing was performed using the Illumina TruSeq Kits®. Paired-end libraries of 2 × 150 bp fragments with an insert size of 300 bp were prepared with Illumina TruSeq Nano DNA Library Prep® kit. Sequencing was performed on a HiSeq 2500 Illumina sequencer, with an average coverage of 10–15X. We sequenced 49 genomes of *Microbotryum* strains from *D. carthusianorum*, 87 from *D. pavonius*, 12 from *D. seguieri*, and 17 from other hosts across western Europe. The samples were collected in Italy, which is not a party of the Nagoya protocol, in the Czech Republic, Germany, Finland, UK and the Netherlands, which do not regulate access to their genetic resources in relation to the Nagoya Protocol, in Switzerland, where the Nagoya laws only require for fundamental research to keep track of information, and in France, where there was an exception for micro-organisms at the time of sampling.

We followed GATK pipeline best practices for read mapping and SNP calling. Illumina reads from the 165 whole genomes were mapped onto the *M. superbum* reference genome. We constructed a “hybrid” reference genome by merging the whole genome *M. superbum* 1065  $a_1$  and the two mating-type chromosomes of the *M. superbum* 1065  $a_2$  genome (tig000000068 and tig000000031, respectively harboring the  $a_2$  PR and HD alleles). We mapped the reads against the hybrid reference genome using BWA-MEM2<sup>131</sup> (<https://github.com/bwa-mem2/bwa-mem2>, t-4 option). For mating-type chromosomes, only reads mapping to a single location were considered for further analyses. We used GATK Haplotype Caller to obtain gVCF from the bam mapping files using the haploid flag (ploidy = 1) for the mating-type chromosomes, and using the diploid flag (ploidy = 2) for the rest of the genome. We used CombineGVCFs to combine gvcf of the autosomes and mating-type chromosomes for a given individual, and to obtain the vcf file. We filtered out highly genetically related (i.e., individuals with pairwise KING-robust kinship estimates among individuals >0.354)<sup>132</sup> with plink version 2.0<sup>133</sup>, and individuals with more than 20% of positions with missing genotype. We computed pairwise linkage disequilibrium between genome-wide biallelic sites with minor allele frequency higher than 0.25 using Plink v1.90b6.26. We used for this analysis only reads mapped to the genome *M. superbum* 1065  $a_1$  and the 149 strains that passed filtering.

For our analysis of homozygosity of HD genes in the genomes sequenced with the Illumina technology, we used the three strains for which we could obtain good enough de novo assemblies and with both PR alleles: strains 235 and 267 on two *D. pavonius* plants in Italy in 2011 by M. E. Hood, Janis Antonovics and Emme Bruns (GPS coordinates: 44.14482, 7.55401 and 44.21239, 7.71804, respectively), and c165 on *D. carthusianorum* in France in 2016 by Henri Maté (GPS coordinates: 47.9683333-7.2533333). We assembled the Illumina reads with spades v3.15.5<sup>134</sup> using the --careful mode. Contigs containing the mating-type genes PR, HD1 and HD2 were identified by whole genome alignment of the assemblies against the mating-type chromosomes of each haploid of the reference strain *M. superbum* 1065. We identified the position of the HD genes by using miniprot 0.12-r237<sup>135</sup> mapping the alternative alleles of HD1 and HD2 (i.e., non disrupted as in the  $a_1$  reference and the

two predicted gene models in the homologous locus of the  $a_2$  reference). All three strain cultures harbored the two alleles for the *PR* gene, but they were found to carry a single *HD* allele, i.e., a homozygous *HD* locus. We manually annotated the *HD* genes of strains 235 and c165.

**Analyses of functional domains in *HD* genes.** Functional domains of the HD proteins of both mating types were predicted in *M. lychnidis-dioicae* 1064, *M. superbum* 1065, 6P, 6D, 235, 267, *M. shykoffianum* c212, c165, and *M. scorzonerae* 1528 by the InterPro server<sup>36</sup>, and compared with previously characterized homeodomain proteins in the Pfam database, as done previously<sup>86</sup>.

### Tetrad isolation and PCR analyses

**Sample and culture description.** We pulled out one tetrad of from additional 22 *M. scorzonerae* strains (Table S2), from the plants *Scorzonera humilis* (strain 1528, collected in 2018 in Cefn Cribwr, Wales, UK 51°32'08"N 3°38'52"W by Julian Woodman), *Tragopogon pratensis* (strain 1535 collected in 2018 in Länsimäentie, Helsinki, Finland, 60°13'54.753"N 25°6'57.1716"E, by Juha Tuomola), *Tragopogon pratensis* (strain 1544, collected in 2018 in De Weere, Netherlands, 52°44'13.33"N 4°59'22.5"E, by Justus and Maureen Houthuesen-Hulshoff), and *Tragopogon porrifolius* (strain 1560 collected in 2018 in De Weere, Netherlands, 52°44'13.33"N 4°59'22.5"E 2018, by Justus and Maureen Houthuesen-Hulshoff). The samples were collected in Finland, the Netherlands and the UK, which do not regulate access to their genetic resources in relation to the Nagoya Protocol. We also pulled out two tetrads from each of six additional *M. superbum* strains (Table S2), collected on *D. pavonius* in 2019 in Garelli (Italy 44.1887, 7.6876). Italy is not a party of the Nagoya protocol.

To obtain pure haploid cultures from *M. scorzonerae* and *M. superbum* samples, teliospores were first germinated by plating on water agar and incubated at room temperature for -12 h. In *Microbotryum* fungi, a linear basidium containing a meiotic tetrad of cells forms during germination<sup>60</sup>. Meiosis occurs in the basidium, yielding four haploid cells. The four haploid basidium cells then undergo mitosis to form haploid sporidia. These sporidia can be separated from the basidium and isolated into separate cultures by micromanipulation<sup>60</sup>. In order to assess segregation patterns of the mating-type loci across multiple meioses, sporidia cells were isolated from one tetrad of each of 22 diploid strains in *M. scorzonerae* and 2 tetrads from each of six diploid *M. superbum* strains collected from *D. pavonius* (Table S2). Each sporidium cell was then moved by microcapillary pipetting to a potato dextrose agar (PDA) plate, and incubated at room temperature. Two of every four cells in each tetrad divided only a few times after isolation in *M. scorzonerae* and then stopped division. This inviability has been seen in other *Microbotryum* species and is called haplo-lethality<sup>60</sup>. Sporidial cultures which grew successfully were maintained continually on PDA and were stored as samples in silica powder at -18 °C.

**PCR analyses.** Primers were designed, using Primer3<sup>37</sup>, that would specifically amplify, for each the *PR*, *HD1* and *HD2* mating-type genes, either of the two alleles that were present in the two haploid reference genomes of *M. superbum*, a priori compatible for mating given their opposite position in the tetrad<sup>60</sup> (Table S6).

For amplifying fragments of *PR* or *HD* genes, DNA templates were prepared by suspending portions of fungal cultures (quantity not exceeding the tip of a sterile toothpick) in 40 µL of ultrapure water. Boiling PCR was performed in a final volume of 30 µL containing 3 µL of fungal suspension, 15 µL of 2X Buffer (containing a mix of dNTPs 5 mM in addition of bovine serum albumin (BSA) 20 mg/ml, W1 detergent 1%, home made 10X Buffer (Ammonium Sulfat 1 M, Tris HCl 1 M, MgCl<sub>2</sub>·6H<sub>2</sub>O 1 M and sterile deionized H<sub>2</sub>O), 1.2 µL of each forward and reverse primers (10 µM), 0.15 µL of Taq DNA Polymerase (5 U/µL, MP Biomedicals) and 9.8 µL of sterile H<sub>2</sub>O. The touchdown PCR

amplification protocol consisted of 5 min at 94 °C for pre-denaturation, 10 cycles of denaturation at 94 °C for 30 s, annealing at 60 °C (with 1 °C decrements from 60 °C to 50 °C at every cycle) for 30 s, and extension at 72 °C for 1 min. This was followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, and extension at 72 °C for 1 min. The reaction was finished with a final extension for 5 min at 72 °C. The PCR products were revealed and confirmed by the GelRed Nucleic Acid Stain 10,000X (Biotium) and photographed with UV transillumination using the VWR Genosmart gel documentation system after loading on 1.5% agarose (Nippon Genetics Europe GmbH) TAE gels. PCR products were sequenced using Sanger technology at Genewiz.

**Mating tests and plant inoculations.** To check for compatibility between cultures in tetrads, mating tests were performed between pairs of cultures from the same meiotic tetrad. Six-day old cultures were suspended in sterile water. Suspensions were then combined in equal amounts and plated in small spots, less than 10 µl each, onto water agar plates with 0.5 ml phytol solution and onto water agar plates with 0.5 ml sterile water. The phytol solution was 0.006% phytol and 0.8% ethanol<sup>60</sup>. Phytol has been shown to induce hyphal growth in *Microbotryum* fungi<sup>38</sup>. Plates were refrigerated at 5 °C. Spots were checked for hyphae using an inverted scope between 2 and 4 days after plating. We used cultures of *M. superbum* sporidia of opposite PR mating-type alleles but identical HD alleles to inoculate seeds of *Dianthus chinensis* as previously described<sup>60</sup>. We grew the plants in the greenhouse and checked for symptoms in anthers. The species *D. chinensis* was used as it is easier to grow under the greenhouse and to make it flower than *D. pavonius* or *D. carthusianorum*, and as the focal *Microbotryum* fungi are found on various *Dianthus* hosts in nature.

**Statistics and reproductibility.** No statistical method was used to predetermine sample size. No data were excluded from the analyses. The experiments were not randomized. The Investigators were not blinded to allocation during experiments and outcome assessment.

**Softwares for figures.** For figures illustrating synteny, we used the package circlize<sup>117</sup> or Rideogram<sup>116</sup> based on alignments performed with mummer<sup>139</sup>. We used ggplot2 for  $d_5$  vs ancestral gene rank scatter plots, extant vs ancestral gene rank scatter plots, log2 expression ratios and bar plots. RNAseq read mapping was visualized with jbrowse 2.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Whole genome shotgun sequencing, RNAseq data and genome assemblies generated for this work are available at GenBank BioProjects PRJNA1169769 (<https://www.ncbi.nlm.nih.gov/bioproject/1169769>), PRJNA1206992 (<https://www.ncbi.nlm.nih.gov/bioproject/1206992>) and PRJNA1206994 (<https://www.ncbi.nlm.nih.gov/bioproject/1206994>). BioSample, Genome and SRA accession numbers are detailed in Table S9 for the RNAseq data, and in Table S11 for whole genome sequences. The RNAseq quantifications, species tree and gene trees are deposited in dryad (<https://doi.org/10.5061/dryad.0k6djhb89>), along with the submitted versions of genome assemblies and the de novo gene annotation of *M. superbum* strain 6P and mating-type chromosomes of *M. lychnidis-dioicae* strain 1064. Source data are provided with this paper.

### Code availability

The pipeline to annotate genomes is available at the following link: <https://github.com/QuentinRougemont/EASYstrata> or <https://doi.org/10.>

5281/zenodo.14917635<sup>40</sup>. Relevant customized codes are available at the following link: <https://github.com/eliselucotte/MicrobotryumOnDianthus> or <https://doi.org/10.5281/zenodo.14918233<sup>41</sup>>.

## References

- Lande, R. & Schemske, D. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* **39**, 24–40 (1985).
- Charlesworth, D. & Charlesworth, B. Inbreeding depression and its evolutionary consequences. *Annu. Rev. Ecol. Evol. Syst.* **18**, 237–268 (1987).
- Charlesworth, D., Morgan, M. & Charlesworth, B. Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. *Evolution* **44**, 1469–1489 (1990).
- Charlesworth, D. Plant sex determination and sex chromosomes. *Heredity* **88**, 94–101 (2002).
- Igic, B., Lande, R. & Kohn, J. Loss of self-incompatibility and its evolutionary consequences. *Int. J. Plant Sci.* **169**, 93–104 (2008).
- Hereford, J. Does selfing or outcrossing promote local adaptation? *Am. J. Bot.* **97**, 298–302 (2010).
- Lande, R. Evolution of phenotypic plasticity in colonizing species. *Mol. Ecol.* **24**, 2038–2045 (2015).
- Nieuwenhuis, B. P. S. et al. Evolution of uni- and bifactorial sexual compatibility systems in fungi. *Heredity* **111**, 445–455 (2013).
- Goldberg, E. & Igić, B. Tempo and mode in plant breeding system evolution. *Evolution* **66**, 3701–3709 (2012).
- Chantha, S., Herman, A., Platts, A., Vekemans, X. & Schoen, D. Secondary evolution of a self-incompatibility locus in the Brassicaceae genus *Leavenworthia*. *PLoS Biol.* **11**, e1001560 (2013).
- Hansch, E., Herron, M., Wiens, J., Nozaki, H. & Michod, R. Repeated evolution and reversibility of self-fertilization in the volvocine green algae. *Evolution* **72**, 386–398 (2018).
- Beukeboom, L. & Perrin, N. *The Evolution of Sex Determination* (Oxford University Press, 2014).
- Vekemans, X., Poux, C., Goubet, P. & Castric, V. The evolution of selfing from outcrossing ancestors in Brassicaceae: what have we learned from variation at the S-locus? *J. Evol. Biol.* **27**, 1372–1385 (2014).
- Billiard, S. et al. Having sex, yes, but with whom? Inferences from fungi on the evolution of anisogamy and mating types. *Biol. Rev.* **86**, 421–442 (2011).
- Billiard, S., López-Villavicencio, M., Hood, M. & Giraud, T. Sex, outcrossing and mating types: unsolved questions in fungi and beyond. *J. Evol. Biol.* **25**, 1020–1038 (2012).
- Coelho, M. A., Bakkeren, G., Sun, S., Hood, M. E. & Giraud, T. Fungal sex: the basidiomycota. *Microbiol. Spectr.* **5**, 1–30 (2017).
- Branco, S. et al. Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. *Proc. Natl. Acad. Sci. USA* **114**, 7067–7072 (2017).
- Branco, S. et al. Multiple convergent supergene evolution events in mating-type chromosomes. *Nat. Commun.* **9**, 2000 (2018).
- Charlesworth, D. The status of supergenes in the 21st century: recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations. *Evol. Appl.* **9**, 74–90 (2016).
- Charlesworth, D. Evolution of recombination rates between sex chromosomes. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160456 (2017).
- Charlesworth, D. When and how do sex-linked regions become sex chromosomes? *Evolution* **75**, 569–581 (2021).
- Bergero, R. & Charlesworth, D. The evolution of restricted recombination in sex chromosomes. *Trends Ecol. Evol.* **24**, 94–102 (2009).
- Hartmann, F. E. et al. Recombination suppression and evolutionary strata around mating-type loci in fungi: documenting patterns and understanding evolutionary and mechanistic causes. *N. Phytol.* **229**, 2470–2491 (2021).
- Vittorelli, N. et al. Stepwise recombination suppression around the mating-type locus in the fungus *Schizothecium tetrasporum* (Ascomycota, Sordariales). *PLoS Genet.* **19**, e1010347 (2022).
- Ironside, J. E. No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *Bioessays* **32**, 718–726 (2010).
- Jay, P., Jeffries, D., Hartmann, F. E., Véber, A. & Giraud, T. Why do sex chromosomes progressively lose recombination? *Trends Genet.* **40**, 564–579 (2024).
- Charlesworth, D., Charlesworth, B. & Marais, G. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95**, 118–128 (2005).
- Wright, A. E., Dean, R., Zimmer, F. & Mank, J. E. How to make a sex chromosome. *Nat. Commun.* **7**, 12087 (2016).
- Bazzicalupo, A. L., Carpentier, F., Otto, S. P. & Giraud, T. Little evidence of antagonistic selection in the evolutionary strata of fungal mating-type chromosomes (*Microbotryum lychnidis-dioicae*). *G3-Genes Genomes Genet.* **9**, 1987–1998 (2019).
- Ma, W.-J., Carpentier, F., Giraud, T. & Hood, M. Differential gene expression between fungal mating types is associated with sequence degeneration. *Genome Biol. Evol.* **12**, 243–258 (2020).
- Jeffries, D. L., Gerchen, J. F., Scharmann, M. & Pannell, J. R. A neutral model for the loss of recombination on sex chromosomes. *Philos. Trans. R. Soc. B Biol. Sci.* **376**, 20200096 (2021).
- Kent, T. V., Uzunovic, J. & Wright, S. I. Coevolution between transposable elements and recombination. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160458 (2017).
- Antonovics, J. & Abrams, J. Y. Intratetrad mating and the evolution of linkage relationships. *Evolution* **58**, 702–709 (2004).
- Lenormand, T. & Roze, D. Y Recombination arrest and degeneration in the absence of sexual dimorphism. *Science* **375**, 663 (2022).
- Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294 (1966).
- Muller, H. J. Some genetic aspects of sex. *Am. Nat.* **66**, 118–138 (1932).
- Charlesworth, B. & Charlesworth, D. The degeneration of Y chromosomes. *Philos. Trans. R. Soc. B-Biol. Sci.* **355**, 1563–1572 (2000).
- Charlesworth, D. The timing of genetic degeneration of sex chromosomes. *Philos. Trans. R. Soc. B* **376**, 20200093 (2021).
- Bachtrog, D. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genet.* **14**, 113–124 (2013).
- Carpentier, F. et al. Tempo of degeneration across independently evolved nonrecombining regions. *Mol. Biol. Evol.* **39**, msac060 (2022).
- Fontanillas, E. et al. Degeneration of the nonrecombining regions in the mating-type chromosomes of the anther-smut fungi. *Mol. Biol. Evol.* **32**, 928–943 (2015).
- Duhamel, M., Hood, M. E., Rodríguez de la Vega, R. C. & Giraud, T. Dynamics of transposable element accumulation in the non-recombining regions of mating-type chromosomes in anther-smut fungi. *Nat. Commun.* **14**, 5692 (2023).
- Badouin, H. et al. Chaos of rearrangements in the mating-type chromosomes of the anther-smut fungus *Microbotryum lychnidis-dioicae*. *Genetics* **200**, 1275–1284 (2015).
- Kües, U., James, T. Y. & Heitman, J. Mating type in basidiomycetes: unipolar, bipolar, and tetrapolar patterns of sexuality. In *Evolution of Fungi and Fungal-Like Organisms* (eds Pöggeler S., Wöstemeyer J.) (Springer, 2011).

45. Bakkeren, G. & Kronstad, J. W. Linkage of mating-type loci distinguishes bipolar from tetrapolar mating in basidiomycetous smut fungi. *Proc. Natl. Acad. Sci. USA* **91**, 7085–7089 (1994).
46. Xu, J. et al. Dandruff-associated *Malassezia* genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens. *Proc. Natl. Acad. Sci. USA* **104**, 18730–18735 (2007).
47. Coelho, M. A., Ianiri, G., David-Palmaa, M. & Heitman, J. Frequent transitions in mating-type locus chromosomal organization in *Malassezia* and early steps in sexual reproduction. *Proc. Natl. Acad. Sci. USA* **120**, e2305094120 (2023).
48. Lengeler, K. B. et al. Mating-type locus of *Cryptococcus neoformans*: a step in the evolution of sex chromosomes. *Eukaryot. Cell* **1**, 704–718 (2002).
49. James, T. Y., Srivilai, P., Kues, U. & Vilgalys, R. Evolution of the bipolar mating system of the mushroom *Coprinellus disseminatus* from its tetrapolar ancestors involves loss of mating-type-specific pheromone receptor function. *Genetics* **172**, 1877–1891 (2006).
50. Chen, B. Z. et al. Fruiting body formation in *Volvariella volvacea* can occur independently of its MAT-a-controlled bipolar mating system, enabling homothallic and heterothallic life cycles. *Genes Genomes Genet.* **6**, 2135–2146 (2016).
51. Kues, U. et al. A chimeric homeodomain protein causes self-incompatibility and constitutive sexual development in the mushroom *Coprinus cinereus*. *EMBO J.* **13**, 4054–4059 (1994).
52. Hood, M. E. et al. Distribution of the anther-smut pathogen on species of the Caryophyllaceae. *N. Phytol.* **187**, 217–229 (2010).
53. Hartmann, F. E. et al. Understanding adaptation, coevolution, host specialization, and mating system in castrating anther-smut fungi by combining population and comparative genomics. *Annu. Rev. Phytopathol.* **57**, 431–457 (2019).
54. Le Gac, M., Hood, M. E., Fournier, E. & Giraud, T. Phylogenetic evidence of host-specific cryptic species in the anther smut fungus. *Evolution* **61**, 15–26 (2007).
55. Refrégier, G. et al. Cophylogeny of the anther smut fungi and their caryophyllaceous hosts: Prevalence of host shifts and importance of delimiting parasite species for inferring cospeciation. *BMC Evol. Biol.* **8**, 100 (2008).
56. Hood, M. E., Antonovics, J. & Koskella, B. Shared forces of sex chromosome evolution in haploids and diploids. *Genetics* **168**, 141–146 (2004).
57. Feurtey, A. et al. Strong phylogeographic co-structure between the anther-smut fungus and its white campion host. *N. Phytol.* **212**, 668–679 (2016).
58. Giraud, T., Yockteng, R., Lopez-Villavicencio, M., Refrégier, G. & Hood, M. E. The mating system of the anther smut fungus, *Microbotryum violaceum*: selfing under heterothallism. *Euk. Cell* **7**, 765–775 (2008).
59. Hartmann, F. E. et al. Congruent population genetic structures and divergence histories in anther-smut fungi and their host plants *Silene italica* and the *Silene nutans* species complex. *Mol. Ecol.* **29**, 1154–1172 (2020).
60. Hood, M. E. & Antonovics, J. Intratetrad mating, heterozygosity, and the maintenance of deleterious alleles in *Microbotryum violaceum* (= *Ustilago violacea*). *Heredity* **85**, 231–241 (2000).
61. Alexander, H. M., Thrall, P. H., Antonovics, J., Jarosz, A. M. & Oudemans, P. V. Population dynamics and genetics of plant disease: a case study of anther-smut disease. *Ecology* **77**, 990–996 (1996).
62. Antonovics, J., Iwasa, Y. & Hassel, M. P. A generalized model of parasitoid, veneral, and vector-based transmission processes. *Am. Nat.* **145**, 661–675 (1995).
63. Hood, M. E., Antonovics, J. & Heishman, H. Karyotypic similarity identifies multiple host-shifts of a pathogenic fungus in natural populations. *Infect. Genet. Evol.* **2**, 167–172 (2003).
64. Hartmann, F. E. et al. Higher gene flow in sex-related chromosomes than in autosomes during fungal divergence. *Mol. Biol. Evol.* **37**, 668–682 (2020).
65. Zakharov, I. A. Intratetrad mating and its genetic and evolutionary consequences. *Genetika* **41**, 508–519 (2005).
66. Zakharov, I. A. Intratetrad mating as the driving force behind the formation of sex chromosomes in fungi. *Trends Genet. Evol.* **6**, <https://doi.org/10.24294/tge.v24296i24291.24252> (2023).
67. Duhamel, M. et al. Onset and stepwise extensions of recombination suppression are common in mating-type chromosomes of *Microbotryum* anther-smut fungi. *J. Evol. Biol.* **35**, 1619–1634 (2022).
68. Carpentier, F. et al. Convergent recombination cessation between mating-type genes and centromeres in selfing anther-smut fungi. *Genome Res.* **29**, 944–953 (2019).
69. Thomas, A., Shykoff, J., Jonot, O. & Giraud, T. Sex-ratio bias in populations of the phytopathogenic fungus *Microbotryum violaceum* from several host species. *Int. J. Plant Sci.* **164**, 641–647 (2003).
70. Kaltz, O. & Shykoff, J. A. Sporidial mating-type ratios of teliospores from natural populations of the anther smut fungus *Microbotryum* (equals *Ustilago*) *violaceum*. *Int. J. Plant Sci.* **158**, 575–584 (1997).
71. Oudemans, P. V. et al. The distribution of mating-type bias in natural populations of the anther-smut *Ustilago violacea* on *Silene alba* in Virginia. *Mycologia* **90**, 372–381 (1998).
72. Horns, F., Petit, E. & Hood, M. E. Massive expansion of Gypsy-Like retrotransposons in *Microbotryum* fungi. *Genome Biol. Evol.* **9**, 363–371 (2017).
73. Petit, E. et al. Co-occurrence and hybridization of anther-smut pathogens specialized on *Dianthus* hosts. *Mol. Ecol.* **26**, 1877–1890 (2017).
74. Denchev, C. M., Giraud, T. & Hood, M. E. Three new species of anthericolous smut fungi on Caryophyllaceae. *Mycol. Balc* **6**, 79–84 (2009).
75. Xu, L., Petit, E. & Hood, M. E. Variation in mate-recognition pheromones of the fungal genus *Microbotryum*. *Heredity* **116**, 44–51 (2016).
76. Hood, M. E. & Antonovics, J. Mating within the meiotic tetrad and the maintenance of genomic heterozygosity. *Genetics* **166**, 1751–1759 (2004).
77. Devier, B., Aguilera, G., Hood, M. & Giraud, T. Ancient trans-specific polymorphism at pheromone receptor genes in basidiomycetes. *Genetics* **181**, 209–223 (2009).
78. Day, A. & Jones, J. The production and characteristics of diploids in *Ustilago violacea*. *Genet. Res. Camb.* **11**, 63–81 (1968).
79. Gladieux, P. et al. Maintenance of fungal pathogen species that are specialized to different hosts: allopatric divergence and introgression through secondary contact. *Mol. Biol. Evol.* **28**, 459–471 (2011).
80. Lee, N., Bakkeren, G., Wong, K., Sherwood, J. E. & Kronstad, J. W. The mating-type and pathogenicity locus of the fungus *Ustilago hordei* spans a 500-kb region. *Proc. Natl. Acad. Sci. USA* **96**, 15026–15031 (1999).
81. Fraser, J. A. et al. Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms. *Plos Biol.* **2**, 2243–2255 (2004).
82. Rabe, F. et al. A complete toolset for the study of *Ustilago bromivora* and *Brachypodium* sp as a fungal-temperate grass pathosystem. *Elife* **5**, e20522 (2016).
83. Sun, S., Coelho, M. A., Heitman, J. & Nowrousian, M. Convergent evolution of linked mating-type loci in basidiomycete fungi. *PLoS Genet.* **15**, e1008365 (2019).
84. Yi, R. et al. Genomic structure of the A mating-type locus in a bipolar basidiomycete, *Pholiota nameko*. *Mycol. Res.* **113**, 240–248 (2009).

85. Wang, Y.-W. et al. Invasive Californian death caps develop mushrooms unisexually and bisexually. *Nat. Commun.* **14**, 6560 (2023).
86. Cabrita, A. et al. Multiple pathways to homothallism in closely related yeast lineages in the Basidiomycota. *mBio* **12**, <https://doi.org/10.1128/mbio.03130-03120> (2021).
87. Chang, Y. F., Imam, J. S. & Wilkinson, M. F. The nonsense-mediated decay RNA surveillance pathway. *Annu Rev. Biochem.* **76**, 51–74 (2007).
88. Causier, B. et al. Conservation of nonsense-mediated mRNA decay complex components throughout eukaryotic evolution. *Sci. Rep.* **7**, 16692 (2017).
89. de Vienne, D. M., Hood, M. E. & Giraud, T. Phylogenetic determinants of potential host shifts in fungal pathogens. *J. Evol. Biol.* **22**, 2532–2541 (2009).
90. Liu, T., Tian, H., He, S. & Guo, L. *Microbotryum scorzonerae* (Microbotryaceae), new to China, on a new host plant. *Mycotaxon* **108**, 245–247 (2009).
91. Perlín, M. H. et al. Sex and parasites: Genomic and transcriptomic analysis of *Microbotryum lychnidis-dioicae*, the biotrophic and plant-castrating anther smut fungus. *BMC Genom.* **16**, 1–24 (2015).
92. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
93. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Giga-Science* **10**, giab008 (2021).
94. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
95. Diesh, C. et al. JBrowse 2: a modular genome browser with views of synteny and structural variation. *Genome Biol.* **24**, 74 (2023).
96. Koren, S. et al. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
97. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
98. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
99. Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A. & Zdobnov, E. M. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).
100. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
101. Walker, B. J. et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**, e112963 (2014).
102. Flynn, J. M. et al. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci. USA* **117**, 9451–9457 (2020).
103. Smit, A., Hubley, R. & Green, P. RepeatMasker Open-4.0. <https://www.repeatmasker.org/>, (2013–2015).
104. Bao, W., Kojima, K. K. & Kohany, O. Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* **6**, 11 (2015).
105. Crawford, J. R. & Howell, D. C. Comparing an individual's test score against norms derived from small samples. *Clin. Neuropsychol.* **12**, 482–486 (1998).
106. Crawford, J. R. & Garthwaite, P. H. Single-case research in neuropsychology: a comparison of five forms of t-test for comparing a case to controls. *Cortex* **48**, 1009–1016 (2012).
107. Bruna, T., Hoff, K. J., Lomsadze, A., Stanke, M. & Borodovsky, M. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genom. Bioinform.* **3**, lqaa108 (2021).
108. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
109. Wu, T. D., Reeder, J., Lawrence, M., Becker, G. & Brauer, M. J. GMAP and GSNAP for genomic sequence alignment: enhancements to speed, accuracy, and functionality. In *Statistical Genomics: Methods and Protocols* (eds Mathé, E., Davis, S) (Springer New York, 2016).
110. Keller, O., Kollmar, M., Stanke, M. & Waack, S. A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics* **27**, 757–763 (2011).
111. Gabriel, L., Hoff, K. J., Bruna, T., Borodovsky, M. & Stanke, M. TSEBRA: transcript selector for BRAKER. *BMC Bioinform.* **22**, 566 (2021).
112. Rougemont, Q. et al. EASYstrata: a new workflow to infer evolutionary strata along sex chromosomes and other supergenes. *bioRxiv*, (2025). 2025.2001.2006.631483.
113. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).
114. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
115. Petit, E. et al. Linkage to the mating-type locus across the genus *Microbotryum*: insights into non-recombining chromosomes. *Evolution* **66**, 3519–3533 (2012).
116. Hao, Z. et al. Rldeogram: drawing SVG graphics to visualize and map genome-wide data on the ideograms. *PeerJ Comput. Sci.* **6**, e251 (2020).
117. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).
118. Ranwez, V., Douzery, E. J. P., Cambon, C., Chantret, N. & Delsuc, F. MACSE v2: Toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. *Mol. Biol. Evol.* **35**, 2582–2584 (2018).
119. Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L. & Scheffler, K. RELAX: detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* **32**, 820–832 (2015).
120. Ranwez, V., Harispe, S., Delsuc, F. & Douzery, E. J. P. MACSE: Multiple alignment of coding sequences accounting for frame-shifts and stop codons. *PLoS ONE* **6**, e22594 (2011).
121. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
122. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis* 2nd edn. (Springer Publishing Company, 2009).
123. Lindeløv, J. K. Mcp: an R package for regression with multiple change points. <https://doi.org/10.31219/osf.io/fzqxv> (2020).
124. Minh, B. Q. et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
125. Lemoine, F. & Gascuel, O. Gtree/Goalign: toolkit and Go API to facilitate the development of phylogenetic workflows. *NAR Genom. Bioinform.* **3**, lqab075 (2021).
126. Junier, T. & Zdobnov, E. M. The Newick utilities: high-throughput phylogenetic tree processing in the Unix shell. *Bioinformatics* **26**, 1669–1670 (2010).
127. Abascal, F., Zardoya, R. & Telford, M. J. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* **38**, W7–W13 (2010).
128. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
129. To, T.-H., Jung, M., Lycett, S. & Gascuel, O. Fast dating using least-squares criteria and algorithms. *Syst. Biol.* **65**, 82–97 (2016).

130. Lopez-Villavicencio, M. et al. Multiple infections by the anther smut pathogen are frequent and involve related strains. *PLoS Pathog.* **3**, e176 (2007).
  131. Vasimuddin, M., Misra, S., Li, H. & Aluru, S. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. *Proc. IEEE International Parallel and Distributed Processing Symposium (IPDPS)* 314–324 (Rio de Janeiro, Brazil, 2019).
  132. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
  133. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
  134. Pribelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes de novo assembler. *Curr. Protoc. Bioinform.* **70**, e102 (2020).
  135. Li, H. Protein-to-genome alignment with minimap2. *Bioinformatics* **39**, btad014 (2023).
  136. Mitchell, A. et al. The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* **43**, D213–D221 (2015).
  137. Untergasser, A. et al. Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).
  138. Castle, A. J. & Day, A. W. Isolation and identification of  $\alpha$ -tocopherol as an inducer of the parasitic phase of *Ustilago violacea*. *Phytopathology* **74**, 1194–1200 (1984).
  139. Kurtz, S. et al. Versatile and open software for comparing large genomes. *Genome Biol.* **5**, R12 (2004).
  140. Rougemont, Q., Lucotte, E. A. & Boyer, L. EASYstrata: a new workflow to infer evolutionary strata along sex chromosomes and other supergenes. *Zenodo*, <https://doi.org/10.5281/zenodo.14917635> (2025).
  141. Lucotte, E. A., Rougemont, Q., Boyer, L. & Jay, P. Repeated loss of function at HD mating-type genes and repeated recombination suppression in anther-smut fungi, *Zenodo*, <https://doi.org/10.5281/zenodo.14918233> (2025).
- JPH sequenced in ONT the genome of the strain used for RNAseq. E.A.L., P.J., Q.R., L.B. and R.C.R.d.I.V. analyzed the PacBio genomes. A.C. acquired the Illumina genomes and performed SNP calling. P.J. and E.A.L. analyzed the Illumina genomes. M.E.H. and J.G. performed tetrad isolation and mating tests. A.L., E.C., M.E.H. and J.G. performed tetrad segregation analyses. M.D. identified the centromeres and TEs. M.H.P. and R.K.H. acquired the expression data. R.C.R.d.I.V. analyzed the RNAseq data and assembled the Illumina genomes. E.A.L., P.J. and R.C.R.d.I.V. produced the figures. A.N. analyzed the structure of HD genes. T.G. and E.A.L. wrote the manuscript draft with contributions by P.J., M.E.H., Q.R., M.H.P. and R.K.H. All authors edited the manuscript. The final version of the manuscript was prepared by E.A.L., R.C.R.d.I.V. and T.G.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-60222-5>.

**Correspondence** and requests for materials should be addressed to Elise A. Lucotte or Tatiana Giraud.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Acknowledgements

This work was supported by the Louis D. Foundation award and Evol-SexChrom ERC advanced grant #832352 to T.G., a National Institute of Health (NIH) grant R15GM119092 to M.E.H., National Science Foundation (NSF) grants #2007449 and #2419465 to M.H.P., and a Human Frontier Science Program (HFSP) fellowship grant number LT0033/2022-L to P.J. We thank Klara Koupilova, Janis Antonovics, Emme Bruns, Henri Maté and Juha Tuomola for the *M. v. Dianthus* strains, Justus & Maureen Houthuesen-Hulshoff, Julian Woodman, Juha Tuomola for the *M. scorzonerae* strains. We thank Julian Woodman for the picture of *M. scorzonerae*.

## Author contributions

T.G., M.E.H., A.C. and R.C.R.d.I.V. conceptualized the study, acquired funding and supervised the study. M.E.H., T.G. and A.C. acquired the *Microbotryum* strains. A.S. extracted DNA for PacBio sequencing. M.E.H. and W.J.M. sequenced and analyzed the *M. scorzonerae* genome. SB and

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025