Repeated loss of function at HD mating-type genes and of recombination suppression without mating-type locus linkage in anther-smut fungi

Elise Lucotte[1§], Paul Jay[1], Quentin Rougemont[1], Loreleï Boyer[1], Amandine Cornille[2], Alodie Snirc[1], Amandine Labat[1], Elizabeth Chahine[1], Marine Duhamel[1], Jacob Gendelman[3], Roxanne Kaaren Hayes[4], Michael H. Perlin[4], Michael E. Hood[3]*, Ricardo C. Rodríguez de la Vega[1]*, Tatiana Giraud[1]*[§]

* These authors co-supervised the study

[1] Ecologie Systematique et Evolution, CNRS, Université Paris-Saclay, AgroParisTech, 91198 Gif-sur-Yvette, France

[2] Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE - Le Moulon, 91190 Gif-sur-Yvette, France

[3] Department of Biology, Amherst College, 01002-5000 Amherst, Massachusetts, United State of America.

[4] Department of Biology, Program on Disease Evolution, University of Louisville, Louisville, KY 40292, USA

§ Corresponding authors:

tatiana.giraud@universite-paris-saclay.fr

elucotte@gmail.com

IDEEV, Bâtiment 680, 12 route RD128, 91190 Gif-sur-Yvette, France

phone: +33 1 69 15 56 69 fax: +33 1 69 15 46 97

Keywords: mating-type chromosomes, recombination suppression, fungi, sex chromosomes, mating-type loci, basidiomycota, evolutionary strata, *Dianthus,* sheltering hypothesis, stepwise recombination suppression

ABSTRACT

A wide diversity of mating systems occur in nature, with frequent evolutionary transitions in mating-compatibility mechanisms. Basidiomycete fungi typically have two mating-type loci controlling mating compatibility, HD and PR, usually residing on different chromosomes. In *Microbotryum* anther-smut fungi, there have been repeated events of linkage between the two mating-type loci through chromosome fusions, leading to large non-recombining regions. By generating high-quality genome assemblies, we found that two sister *Microbotryum* species parasitizing *Dianthus* plants, *M. superbum* and *M. shykoffianum*, as well as the distantly related *M. scorzonarae*, have their HD and PR mating-type loci on different chromosomes, but with the PR mating-type chromosome fused with a part of the ancestral HD chromosome. Furthermore,progressive extensions of recombination suppression have generated evolutionary strata. In all three species, rearrangements suggest the existence of a transient stage of HD-PR linkage by whole chromosome fusion, and, unexpectedly, the HD genes lost their function. In *M. superbum*, multiple natural diploid strains were homozygous, and the disrupted HD2 alleles was hardly expressed. Mating tests confirmed that a single genetic factor controlled mating compatibility (i.e. PR) and that haploid strains with identical HD alleles could mate and produce infectious hyphae. The HD genes have therefore lost their function in the control of mating compatibility in these *Microbotryum* species. While the loss of function of PR genes in mating compatibility has been reported in a few basidiomycete fungi, these are the first documented cases for the loss of mating-type determination by HD genes in heterothallic fungi. The control of mating compatibility by a single genetic factor is beneficial under selfing and can thus be achieved repeatedly, through evolutionary convergence in distant lineages, involving different genomic or similar pathways.

INTRODUCTION

Mating systems, corresponding to the degree of selfing/outcrossing, influence gene flow, genetic load and adaptability in natural populations [1,2,3,4,5,6,7]. A wide diversity of mating systems occur in nature, controlled by mating compatibility systems (e.g. separate sexes or mating types), with evolutionary transitions between these systems being frequent [5,8,9,10,11,12]. In many animals and plants, different sexes are determined by sex chromosomes [13]. In plants, mating compatibility is often further controlled by a self-incompatibility locus, which prevents mating between identical alleles at this locus [14]. In fungi with mating types, gamete compatibility is controlled at the haploid stage, where only cells with different mating types can successfully mate [15,16], mating types being determined by one or two loci [17], the systems being then called uni- and bifactorial, respectively. There have been multiple transitions in fungi for changes in the number of loci controlling mating types, between one, two or even none, i.e., lack of discrimination for mating [8,16].

Recombination is typically suppressed at loci determining sexes and mating types, which ensures proper function in these complex phenotypes. Recombination suppression has furthermore often extended stepwise outward from these mating-compatibility loci in plants, animals and fungi [18,19,20,21,22,23,24,25], and the reason why is still debated [24,26]. It has long been accepted that recombination suppression extended on sex chromosomes because if was beneficial to link sexually antagonistic genes (i.e., with an allele beneficial in only one sex) to the sex-determining locus [27]. However, little evidence could be found in favor of this hypothesis [28], 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 [18,24,29].

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 [30], the spread of transposable elements adjacent to the non-recombining mating-compatibility locus, together with their epigenetic marks [31], or the selection of non-recombining fragments carrying fewer deleterious mutations than the population average while having their load sheltered by a permanently heterozygous allele [32-34].

Following recombination suppression, Muller's ratchet and Hill-Robertson effect [35,36] lead to genomic degeneration, in terms of gene losses, weaker gene expression, lower frequencies of optimal codons, and transposable element accumulation [37,38,39,40,41,42]. Rearrangements also rapidly accumulate [18,19,43], so that recombination can be challenging to restore after the non-recombining fragments have become highly degenerated [32,34].

Basidiomycete fungi typically have two mating-type loci controlling mating compatibility [17]: the PR locus, controlling pre-mating fusion with pheromone and pheromone receptor genes, and the HD locus, controlling post-mating dikaryotic growth with two homeodomain genes whose products heterodimerize to form an active transcription factor [44]. These two loci are usually on different chromosomes and may harbor multiple alleles. Some basidiomycetes in contrast carry a single mating-type locus, which is caused most often by HD-PR linkage, as in the crop pathogen *Ustilago hordei* [45] and human pathogens *Malassezia spp* [46,47], or more rarely by the loss of mating-compatibility role by the PR locus, as in the mushrooms *Coprinellus disseminatus* [48] and *Volvariella volvacea* [49]. A loss of mating-compatibility role by the HD locus has only been observed so far in experimental mutants with their two HD genes fused [50].

Microbotryum fungi (Basidiomycota) are plant-castrating pathogens, producing their spores in the anthers of diseased plants and therefore called anther-smut fungi. Most *Microbotryum* species specialized on one or a few host plants [51,52,53,54]. *Microbotryum* fungi are model organisms in as diverse fields as ecology, disease transmission, host specialization and reproductive isolation evolution, host-pathogen costructure, mating system and sex-related chromosome evolution [52,54,55,56,57,58,59,60,61,62,63]. *Microbotryum* fungi mostly undergo selfing via intra-tetrad mating [59,64,65], in which case carrying a single mating-type locus is advantageous because it maximizes the odds of gamete compatibility [18]. *Microbotryum* 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 [18,19,40,66]. 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 selfing [66,67]. 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 [18,19,66].

The non-recombining regions on both mating-type chromosomes in *Microbotryum* fungi display signs of genomic degeneration, as neither ever recombines. For example, polymorphism is common for haploid sporidia of one or the two mating types failing to grow *in vitro* in some *Microbotryum* species, likely due to deleterious mutations linked to mating-type loci [59,68-70]. Indeed, with time since recombination suppression, these mating-type chromosomes accumulate transposable elements, non-synonymous mutations, gene losses, rearrangements and non-optimal codons [19,40-43]. 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 [42,71]. 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 [19,40-43].

Surveys of the diversity within *Microbotryum* fungi continue to yield novel insights, and preliminary analyses suggested that the species parasitizing *Dianthus* plant species carried unlinked HD and PR loci, yet displaying very high transposable element content compared to other *Microbotryum* species [71]. Using two *Microbotryum* species on *Dianthus*, our aim was therefore to characterize the mating-type chromosomes of these species, elucidate whether the mating-type loci display localized recombination suppression, and if so, whether the mating-type loci were linked together or to their respective centromeres. We found in both species a large non-recombining mating-type chromosome issued from the fusion of the entire ancestral PR chromosome and an arm of the ancestral HD chromosome but, surprisingly, not the arm carrying the HD locus. Genome comparison indicated that the translocation occurred via a transient stage of whole chromosome fusion, a few rearrangements and then separation of a part of the HD chromosome arm. Remarkably, homozygosity was observed at the HD genes, indicating that they were not involved anymore in mating-type determination, and one of the HD1 alleles appeared disrupted in *M. superbum*. This novel observation among basidiomycete fungi was also found in a distant congeneric species, *M. scorzonerae*, with unlinked PR and HD loci, but homozygosity and non-functionalization of HD genes, and rearrangements between ancestral HD and PR chromosomes. 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.

Material and 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 [52,53,54,72]. 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 [53,54,73,74] (Figure 1). Three species have been formally described so far in this group [74], 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; [53,74]; see below). *Microbotryum scorzonerae* parasitizes flowers of *Scorzonera humilis*, in the Asteraceae family [75]. Unlike other *Microbotryum*, *M. scorzonerae* does not cause spore formation on anthers alone. It instead causes spore formation within the entire floral head (Figure 1).

<u>Genomic analyses</u>

Strains, DNA extraction for PacBio sequencing and species identification

We extracted DNA as previously described [18,19]. For the *Dianthus* strains, DNA sequencing based on PacBio (Pacific Bioscience) HiFi long-read sequencing was performed at the GenoToul sequencing facility (Toulouse, INRAE, France) in 2021. 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) [76]. 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 [53,74], we assigned the strain 1065 to *M. superbum* (aka MvDsp2) and the strain C212 to *M. shykoffianum* (aka MvDsp1). For the *M. scorzonarae* strain, obtained in 2018 from Cefn Cribwr, Wales, by Julian Woodman, sporidia of opposite 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), and sequence accuracy was improved by also sequencing the genomes with Illumina technologies.

RNAseq from *M. superbum* and *M. intermedium*

We generated RNAseq data for gene prediction and HD gene expression analysis. For optimizing gene prediction, we sequenced RNA from one of the focal species, *M. superbum,* and also from an outgroup to our studied *Microbotryum* species, *M. intermedium*. For *M. intermedium,* we used the strain BM12-12.1 (the same strain as used previously for genome sequencing [18,40]), under two different conditions: 1) haploid cells of a single PR mating type (a_1 or a_2) grown on nutrient medium (PDA) and then mixed in equal proportions of the two mating types before RNA extraction ("H" condition), as well as additional replicates grown separately on nutrient agar, whose transcriptomes were analyzed separately, and 2) mixtures of cells of the two mating types under mating conditions ("M" condition), i.e. on water agar. Total RNA was isolated from haploid *M. intermedium* cells using the Monarch® Total RNA Miniprep Kit (New England Biolabs). For the nutrient medium condition, haploid strains (a_1 or a_2) were streaked onto potato dextrose agar (PDA) and grown for 3-4 days at 22 °C. Cells were scraped, suspended in DNA/RNA protection solution, disrupted mechanically in the presence of acid-washed glass beads (0.5 mm dia), and then total RNA was extracted following the manufacturer's protocol. For the mating condition, haploid a_1 and a_2 cultures were first grown separately in yeast extract peptone dextrose (YPD) broth overnight. Cell density was measured with a spectrophotometer. The concentration of each culture was adjusted to O.D.600 1.0. Equal volumes of the two haploid cultures were mixed and plated on water agar. These

plates were incubated for four days at 14°C, after which wet mounts were prepared from each "mated" plate to verify the presence of conjugation tubes, indicating active mating behavior. Cells were scraped and then total RNA was isolated from mated haploid cells using the Qiagen RNeasy Plant Mini Kit. After total RNA isolation, several quality control measures were taken. Concentration and purity were assessed using a NanoDrop 2000 spectrophotometer (Thermo Scientific); 260/280 and 260/230 ratios>1.8 were considered satisfactory for RNAseq application. Additionally, cDNA was prepared and used as a template for intron-spanning primers in PCR reactions to verify the lack of genomic DNA contamination. Bioanalyzer analysis was completed to detect intact 18S and 23S RNA as a measure of overall RNA quality. After passing all quality control measurements, RNA was sent for sequencing to CD Genomics (Shirley, NY, USA). Three replicates were sequenced per condition and means across the three replicates were analyzed.

For *M. superbum,* we used another strain from the same *D. pavonius* population as the strain 1065 (coordinates 44.191, 7.685, near the Garelli Refugium in the Italian Alps) for RNASeq, a_1 and a_2 haploid strains were used separately as well as pooled in mating experiments. For extractions from haploid cells grown on PDA, we used the Zymo -----RNA Extraction Kit (Zymo Research, Irvine, California). Prior to RNA extractions, each haploid fungal strain was grown on two PDA Petri dishes at 28°C for 24 hours. 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_1 and a_2 cells were grown on YPD agar for four days prior to the mating assay. Haploid cells were then suspended in distilled and autoclaved water before adjusting to 10^9 cells/mL. Suspended fungal cells were combined in equal proportions before being spotted in 50 µl spots on water agar Petri dishes. Plates were left to incubate at 13° Celsius for 48 hours. 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 were performed on mated fungal tissue that was aggregated from the water agar petri dishes for each mated sample. RNAseq of three biological replicates per condition (mated, haploid a_1 and haploid a_2) were performed by CD Genomics (Shirley, New York). Three replicates were sequenced for each condition.

Three *D. pavonius* plants infected with *M. superbum* were obtained from Emme Bruns at the University of Maryland. For generating these diseased plants, seeds of *D. pavonius* were

collected from two populations in Aug 2016 in Italy (Duca coord. 44.1946327, 7.66067016 and Garelli coord. 44.189128, 7.688879), which is not a party of the Nagoya protocol. They were reared in greenhouse conditions planted in Sept 2016, reared and crossed within populations in 2017 to produce seeds at the University of Maryland. F1 seeds were planted and inoculated in 2018 with a strain from the Duca population, and scored as diseased 2020. The plants (DE 10B 3; HL-2 duca 1061; De 10A 8) were used for RNA-seq production. Four different flower bud size binspools were used in these analyses: 11-12 mm, 13-14 mm, 15-16 mm, and 18-20 mm. These represented different stages of both host and parasite development. RNA extractions from plant tissue were performed using Qiagen -----RNA Extraction Kit. Flower buds were harvested from plants that were maintained in a controlled growth chamber. The number of flower buds used for each extraction was dependent on the size of the buds that were collected. All extractions with flower buds that fell into the 11-12 mm bin category were completed with at least two flower buds while extractions on buds in the 13-14 mm or larger bin categories used one flower bud per reaction. In total we sent 9 samples from three plants for RNAseq (CD Genomics).

Expression of HD genes

For the gene expression analysis on HD genes, we used only the conditions under which we had multiple technical replicates and for which HD genes are expected to be expressed, i.e, the three 11-12 mm *M. superbum* replicates obtained from the same plant and the three replicates from each of the culture conditions (mated at 48 h and haploid cultures). Expression of HD genes in *M. lychnidis-dioicae* and *M. intermedium* was assessed as a control of expectations for functional HD genes, in two or three replicates for each of three culture conditions (haploid cultures of alternative mating types grown on water, *M. lychnidis-dioicae*, or PDA, *M. intermedium*, and for both under mating conditions at 48 h [40,77]) and from infected flowers at a late stage [78], see Supplementary Table 3. Reads were mapped with minimap2 2.26-r1175 [79] with the -ax sr option, depth per base extracted with samtools depth function [80], explored with jbrowse 2 [81] and visualized with plotkaryotypeR [82].

Genome assemblies and gene prediction

For *M. superbum* 1065, we sequenced separately two haploid genomes, *a priori* of compatible mating types based on their position in the isolated linear tetrad. We assembled these two genomes separately with Canu 2.2 [83] using the raw reads files (fastq) as input, and setting the -

pacbio-hifi parameter. 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 [84], 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 [85] using the basidiomycota_odb10 as reference genes.

The two alternative mating types of *M. scorzonarae* were sequenced on Amherst laboratory ONT platform. Raw ONT reads were trimmed with nanoFilt [86] to keep only reads with a length above 900 bp and minimal q-score of 7. Trimmed reads were assembled with flye v2.2 [87]. The resulting assembly was polished using the ONT read themselves with MarginPolish (https://github.com/UCSC-nanopore-cgl/MarginPolish) and with illumina 150 bp paired-end sequencing belonging to the two mating types. To that end, short-read sequences were trimmed with fastp [88] using default parameters, aligned to their respective reference genome with bwa-mem2 [89], filtered to keep mapping with a minimal quality of 20, sorted and indexed with samtools [90]. Then 3 rounds of polishing were performed with pilon [91]. We then assessed Genome completeness with BUSCO v5.4 [85] using the basidiomycota_odb10 as reference genes.

Repeated regions and low complexity DNA sequences were identified using RepeatModeler [92]. The genome was then softmasked using RepeatMasker open-4.0.9 [93] with the *de novo* detected repeats as well as custom libraries of candidate transposable elements (TEs) from our previous studies [42] and existing fungal TEs from RepBase [94]. 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 incertain. Paired-end RNAseq reads from a total of 12 *M. superbum* 1065, 12 *M. intermedium* BM12-12.1 and 17 *M. lychnidis-dioicae* 1064 samples (combinations of replicates and conditions) were used to train Braker.

Genes were then predicted on the soft-masked genome using the software Braker version 2.1.6 [95] which uses Diamond, ProtHint 2.6.0, GeneMark 4 and AUGUSTUS 3.4.0. To help gene prediction, we combined RNAseq data from *M. superbum*, *M. intermedium* and *M. lychnidis-dioicae*, 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 [18,19,66,67]. The Illumina RNAseq data from 12 *M. superbum* and 6 *M. intermedium* samples generated here together with 5 *M. intermedium* [40] and 17 *M.*

lychnidis-dioicae [77] samples previously obtained were processed with fastp [88], aligned to the reference genome with the software GSNAP and GMAP [96], combined and sorted with samtools [80] and filtered with Augustus [97] to keep only the unique best match (--uniq option) (Table S3) [77]. 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 [98] and assessed the quality of the final gene set with BUSCO using the basidiomycota_odb10 as reference. D*e novo* detection of transposable elements (TEs) was performed as described previously [42].

Orthogroup construction

We performed orthology reconstruction using a set of 47 published genomes from our previous work [40,66] and including a set of four newly assembled genomes. The *Rhodosporidium babjavae* genome was used as an outgroup [66]. We 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, 11 and 17 conditions for *M. superbum*, *M. intermedium* and *M. lychnidis-dioicae* respectively, Table S3). When multiple isoforms were present, we kept the gene prediction with the longest transcript. This extensive dataset was then fed to OrthoFinder [99] with default parameters to reconstruct orthogroups.

Identification of mating-type chromosomes

We identified the contigs belonging to the mating-type chromosomes by i) detecting the contigs carrying PR and HD genes using BLASTN 2.6.0 [100]) with the gene sequences from *M. lagerheimii* as queries [101,102], ii) plotting the ortholog links for all contigs of our focal species against the reference genome of the *M. lagerheimii* strain 1512 (a_1) [18] using Rideogram [103] iii) plotting the ortholog links using Rideogram between the a_1 and a_2 genomes of our focal species.

We also produced circos plots using circlize [104]. 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

We aligned the coding sequences of the HD2 gene for all the *Microbotryum* species with available genomes (Figure 1) using the software MACSE v2 [105]. The HD2 gene that was predicted as two separate genes in the a_2 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 [106] to detect relaxed or intensified selection on the branch of the HD2 gene in the a_2 genome of *M. superbum*, compared to the HD2 gene sequence in the a_1 and a_2 *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 [18,67]. In *M. lagerheimii,* PR and HD loci are located on different chromosomes [18]. For each species, synonymous divergence (d_S) was computed from the alignment of a_1 and a_2 allele sequences, obtained with macse [107], using the yn00 v4.9f program of the PAML package [108] and plotted using the ggplot2 library of R [109]. Pseudo-autosomal, recombining regions were identified as regions with null synonymous divergence on d_S plots, as expected in these highly selfing fungi, and with the same gene order between mating types and 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_S 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_S along the ancestral gene order along the mating-type chromosomes using a changepoint analysis performed using the R package mcp [110]. This analysis relies on Bayesian regression to infer the location of changes in means of the d_s values.

Telomeres and centromeres

To detect telomeres, we counted the occurrence of the telomere-specific motif 'TTAGGG' and its reverse complement 'CCCTAA' within 10kb windows along the contigs. Peaks occur when telomeres are present as they are constituted by multiple repeats of this motif. Centromeres were identified by searching for the centromeric repeats previously described in *Microbotryum* fungi [43,66].

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 [18,19,24].

We performed codon-based alignment with macse v2.05 [105] 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 [111] with default parameters. Tree topologies were assessed using a homemade awk command line, available on demand. Trees were separated into three categories: species tree (a_1 and a_2 are grouped together within the species node), full trans-specific polymorphism (both a_1 are grouped together and both a_2 are grouped together) and partial trans-specific polymorphism (when only the a_1 or the a_2 are grouped together, or some alleles are missing).

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 only 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, with therefore likely no trans-specific polymorphism.

Illumina sequencing and analysis

A total of 165 samples were collected on several diseased *Dianthus* plants from various species (*D. pavonius, D. seguieri, D. carthusianorum, D. hyssopifolius, D. albinos, D. superbus, D. caryophyllus, D. deltoides* 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 and under artificial light. Spores in a given flower are produced by a single diploid individual [112]. 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 populations, so that some mixtures of

recovered sporidia can include only the a_1 or a_2 mating-type chromosome [59,68]. 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 double-stranded DNA concentration. Preparation of DNA libraries for sequencing was performed in the ETH laboratory 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 its genetic resources in relation to the Nagoya Protocol, in Switzerland, where we just need to keep track of information for fundamental research, and in France, where there was an exception for micro-organisms at the time of sampling.

Illumina reads from the 165 whole Illumina genomes were mapped onto the *M. v. Dpavonius* reference genome. Two independent mapping approaches were performed for population genomic analyses. First, we mapped the reads independently on the genomes MvDp−1065−A2 and MvDp−1065−A2. Second, we constructed a hybrid reference genome by merging the whole genome MvDp−1065−A1 and the two mating-type chromosomes of the MvDp−1065−A2 genome (MvDp−1065−A2_tig00000068 and MvDp−1065−A2_tig00000031, respectively harboring the alleles PR2 and HD2). 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. We used CombineGVCFs and GenotypeGVCFs to obtain vcf files. 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 MvDp−1065−A1 and the 149 strains that passed filtering.

In our analysis of homozygosity of HD genes, we first checked the ploidy of the sequenced samples at mating-type chromosomes, as they were mixtures of meiotic products with possibly haplolethal alleles associated with mating types. For this goal, we performed a remapping of

the reads on the hybrid reference genome. Two autosomes were used as controls: tig00000012 and tig00000097. We calculated the median coverage along the two autosomes, as well as along the PR and HD chromosomes. Strains with a median coverage lower than six reads for the control autosomes were not considered further. The median coverage of the HD and PR chromosomes were divided by the median coverage of one of the control autosome, to get their relative coverage. Strains with a relative coverage higher than 0.7 for PR a_1 or higher than 0.8 for PR a_2 were considered as haploid (23 a_1 and 14 a_2 strains). Strains with a relative coverage around 0.5 (between 0.35 and 0.7) for PR a_2 also showed a copy number around 0.5 for PR a_1 and were therefore considered reliable diploid (62 strains). The 17 other strains were considered as having "unknown ploidy". The results were consistent when using the other autosome as a control. Heterozygosity of the HD chromosomes was calculated for each strain on the vcf files using the command -het of vcftools, and was defined as the observed number of heterozygous sites divided by the total number of sites.

Tetrad isolation and PCR analyses

We pulled out tetrads of *M. scorzonerae* from *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 20 tetrads from seven *M. superbum* strains 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 approximately 12 hours. In *Microbotryum* fungi, a linear basidium containing a meiotic tetrad of cells forms during germination [59]. 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 [59]. In order to assess segregation patterns of the mating type loci across

multiple meioses, sporidia cells were isolated from 22 tetrads in *M. scorzonerae* and 20 tetrads in *M. superbum* samples from seven *M. superbum* strains collected from *D. pavonius*. 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 [59]. Sporidial cultures which grew successfully were maintained continually on PDA and were stored as samples in silica powder at -18ºC.

Primers were designed, using Primer3 [113], 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 [59] (Table S1).

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 5mM in addition of bovine serum albumin (BSA) 20mg/ml, W1 detergent 1%, home made 10X Buffer (Ammonium Sulfat 1M, Tris HCl 1M, $MgCl_2 6H_2O$ 1M) and sterile deionized H_2O), 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_2O. The touchdown PCR amplification protocol consisted of 5 min at 94°C for pre-denaturation, 10 cycles of denaturation at 94°C for 30 sec, annealing at 60°C (with 1°C decrements from 60°C to 50°C at every cycle) for 30 sec, and extension at 72°C for 1 min. This was followed by a further 30 cycles of denaturation at 94°C for 30 sec, annealing at 50°C for 30 sec, 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 10000X (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.

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 10µl each, onto water agar plates with 0.5ml phytol solution and onto water agar plates with 0.5 ml sterile water. The phytol solution was 0.006% phytol and 0.8% ethanol [59]. Phytol has been shown to induce hyphal growth in *Microbotryum* fungi [114]. Plates were refrigerated at 5ºC. 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 types but identical HD alleles to inoculate seeds of *Dianthus chinesis* as previously described [59]. We grew the plants in the greenhouse and checked for symptoms in anthers. The species *D. chinesis* was used as it is easier to grow under the greenhouse and make it flower than *D. pavonius* or *D. carthusianorum* and the focal *Microbotryum* fungi are found on various Dianthus hosts in nature.

Softwares for figures

For figures illustrating synteny, we used the package circlize [115] or Rideogram [103] based on alignments performed with mummer [116]. We used ggplot2 for plotting d_S values. For RNAseq mapping illustration, we used the package karyoploteR [82].

RESULTS

A large non-recombining PR chromosome including an arm of the ancestral HD chromosome but not the HD genes

We obtained high-quality genome assemblies from one strain from each of two *Microbotryum* species parasitizing *Dianthus* plants: *M. superbum* (aka MvDsp2 [53], strain 1065) and *M. shykoffianum* (aka MvDsp1 [53], strain c212). The phylogenetic tree based on 2,391 single-copy orthologs representing 5.2 mb of nucleotide sequences placed them as sister species as expected (Figure 1). We compared their mating-type chromosomes (identified by detecting PR and HD genes) to those of *M. lagerheimii*, inferred to be a good proxy for the ancestral gene order of *Microbotryum* mating-type chromosomes on separate HD and PR chromosomes [18,66]. 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 and HD chromosomes, with however a small part of the ancestral HD chromosome, carrying the HD gene, found on a separate chromosome (Figures 2 and S1-4). This new PR chromosome 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* (Figures 2 and S1-5). Only small regions were collinear at one or the two edges of the PR chromosome, corresponding to recombining pseudo-autosomal

regions (PAR, Figures S4-S5). The centromere was identified within the non-recombining region and telomeres at the edges of the new PR and HD chromosomes (Figures S4-S5). 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 (Figures 3 and S1-5), suggesting ongoing recombination. In all four haploid genomes, the PR and HD genes were found on different contigs (Figures 2 and S1-3) and with telomeric repeats identified on almost all of these mating-type contig edges (Figure S4-6), confirming that the two mating-type loci are located on different chromosomes.

In both *M. superbum* and *M. shykoffianum*, a few genes present on the HD chromosome part carrying the HD genes in *M. lagerheimii* were however found in the non-recombining region of the PR chromosome, and conversely (genes colored in green in Figures 4-5 and S4-S5). This indicates an ancient fusion between the whole PR and HD chromosomes, followed by a few genomic rearrangements, before an excision of a part of the ancestral short HD chromosome arm (Figure 3). Such an excision may have occurred in a single step or by chromosome fissions followed by a new fusion.

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 *Microbotrum* fungi parasitizing *Dianthus* plants. Using SNPs on 149 Illumina genomes collected on various *Dianthus* plants in western Europe, high levels of LD were found across most of the PR chromosome, while LD levels were much lower between HD and PR contigs (Figure S7). The LD values within the HD chromosome, as well as against the PR chromosome, were a bit higher than those within autosomes, but were similar as those between the HD chromosomes and autosomes, thus not indicating linkage.

The high per-gene synonymous divergence (d_S) between alleles on the a_1 and a_2 mating-type chromosomes (Figures 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, recombining region shows zero d_S values between alleles in recombining regions such as autosomes, pseudo-autosomal regions of the PR chromosome and in the HD chromosome (Figures 4-5). In contrast, the rearranged region of the PR chromosome displayed high d_S values in both *M. superbum* and *M. shykoffianum* (Figures 4-5). Following previous studies on *Microbotryum* fungi, the region initially linking the PR and HD genes was called the black stratum (Figures 4-5), although this

stratum does not link the PR and HD mating-type loci any more due to the proposed fission of the HD-containing chromosome arm (Figure 3). This black stratum evolved before the divergence of the *M. superbum* and *M. shykoffianum* lineages, as 73% of the genes showed full trans-specific polymorphism between these two lineages (Figure 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; Figures 4-5), that were shown to have evolved around the HD and PR genes ancestrally in the *Microbotryum* clade [18,19,24].

In the *M. shykoffianum* genome, the HD chromosome in the a_2 assembly was much shorter than in the a_1 assembly, while still carrying the HD gene (Figure S4-5). The mapping of diploid reads on the a_1 plus a_2 genome assemblies showed double coverage in the missing regions of the HD chromosome, indicating the collapse of the two homologous chromosomes in the corresponding regions in the a_1 assembly (Figure S8). The coverage was however not doubled in a small region at the edge of one pseudo-autosomal region, and the corresponding genomic region was not found in the a_2 genome assembly, suggesting that this small region may be genuinely missing in the a_2 genome (Figure S8).

Loss-of-function in mating-type determination of HD genes and lack of PR-HD linkage established by segregation analyses

The alleles of the two tightly-linked HD genes (HD1 and HD2) appeared functional in the two haploid genomes of *M. shykoffianum* and in the a_1 genome of *M. superbum*, but not in the a_2 genome of *M. superbum*. The HD2 gene indeed appeared separated into two different protein coding sequences (CDS) in the a_2 genome in *M. superbum,* with an early stop codon and a new start codon (Figure 7). One of the two novel CDS appeared expressed during mating conditions, but not *in planta*, and we detected no expression of the HD2 gene from the a_1 in mating conditions or when cultured as haploid. In contrast, in *M. lychnidis-dioicae* (with non-recombining mating-type chromosomes) and *M. intermedium* (with almost fully recombining mating-type chromosomes) both HD genes are upregulated under mating or *in planta* conditions (Figure 7). Analyses of non-synonymous substitution proportions based on gene genealogies with all available species (Figure 1) indicated significant positive selection in the two new CDS, suggesting neofunctionalization (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. *Microbotryum* produces linear tetrads allowing assessment of first versus second division segregation [59,117]. 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 [101]. 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. In all 12 tetrads but one, two of the four meiotic products conjugated with the a_1 tester and the other half 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. If HD genes were still involved in mating-type determination, only half of conjugated pairing would produce hyphae due to independent segregation/assortment of PR and HD mating type loci inferred from genomic assemblies.

Additionally, PCR tests confirmed that the 12 meiotic tetrads analysed in *M. superbum* 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 S1). The primers for 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. 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 any more in these strains for successful mating and hyphae production. We also inoculated *D. chinesis* in the greenhouse and showed these homozygous pairs could cause disease (5 out of 24 inoculated plants, as typical under inoculations). In one tetrad, two meiotic products produced hyphae alone without any tester strains 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*) [118]. 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. 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 S2).

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, while the PR alleles showed high differentiation and trans-specific polymorphism as expected. This 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-type determination.

Polymorphism analysis of the HD genes based on Illumina sequencing confirmed that they could be homozygous in natural strains. We used SNPs called in the 149 Illumina genomes against the a_1 *M. superbum* reference plus the HD and PR chromosomes from the a_2 genome. The genomes were obtained by spreading diploid teliospores on culture medium, where they undergo meiosis and haploid meiosis product multiply. In 23 strains, we found only significant mapping on the a_1 PR chromosome and in 14 strains only on the a_2 PR chromosome. This is likely due to haplo-lethal alleles on the PR chromosome, with gene disruptions that are lethal at the haploid stage due to degeneration following recombination suppression, preventing growth of haploid sporidia on our culture plates, as identified in several *Microbotryum* species [59,68,70]. Among the 62 strains for which we were confident that they carried the two PR chromosomes (and therefore confidently sequenced as diploids), 42 exhibited polymorphism at the HD genes in SNP data, indicating heterozygosity, while 20 showed no polymorphism at the HD genes. These findings confirm that heterozygosity at HD genes is no longer required for mating compatibility or completion of the life cycle. The SNP analysis further indicated a high selfing rate in these fungi: the F_{IS} mode on autosomes was at 0.83, corresponding to a selfing rate of ca s=0.91 using the formula $s=2F_{IS}/(1+F_{IS})$.

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* (Figure 1). Oxford Nanopore sequencing of two haploid genomes isolated from the same meiotic tetrad and with contrasting PR alleles, revealed identical sequences for both HD1 and HD2 genes. We performed mating tests between meiotic products of 22 tetrads, which again indicated compatibility segregating a single mating-type

locus: when cells were compatible for PR-determined conjugation, they also produced hyphae in 100% of the cases.

The HD1 sequences from the sequenced *M. scorzonerae* a_1 and a_2 haploid genomes appeared to have an uninterrupted coding sequence. The HD2 coding sequences displayed a 19 bp gap inducing a frameshift mutation in the coding region and a stop codon which would truncate the gene to 120 amino-acids instead of 487 amino-acids in *M. lagerheimii*. Using PCR primers and Sanger sequencing, we confirmed that the frameshift is present in all the isolated cultures, as well as in the *M. scorzonerae* sample not used to produce haploid cultures. As in *M. superbum*, the analysis of 22 tetrads indicated that a single genetic factor controlled mating type in *M. scorzonerae*.

The a_1 and a_2 genome for *M. scorzonerae* did not assemble as well as for *M. superbum* and *M. shykoffianum.* 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 autosomes, indicating that mating-type chromosome fusion occurred, and also with ancestral autosomes (Figure S9A). The contigs carrying parts of the ancestral PR and HD chromosomes were nevertheless 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 (Figure S9B). 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, followed by chromosomal rearrangements and subsequent loss-of-function of HD genes in mating-type determination. We could not detect telomeres however in any of these contigs.

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 [19,66]), we found a region with significantly lower d_S values at the edge of the non-recombining region, while significantly different from zero (Figures 4 and S10). 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 (Figures S4-5 and S11) while remaining still much less rearranged than the rest of the non-recombining region, as expected for a young evolutionary stratum. This very same turquoise stratum had in contrast zero d_S values in *M. shykoffianum* and has remained in the pseudo-autosomal region (Figures 5 and S4-S5). An 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 (Figure 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, and 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 (Figure 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 (Figure 4), in the region that we called "ruby" (Figure 5).

A changepoint analysis confirmed the inference of distinct 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 (Figure 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 (Figures 4 and S4). The black, turquoise, light and dark ruby strata in fact had significant different d_S mean values than their nearby strata (Figure S10).

The dark ruby stratum was only partly present in *M. shykoffianum* (Figure 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 (Figure 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 (Figures 4-5).

In *M. scorzonareae,* genes around the PR locus were highly rearranged between mating-type chromosomes and compared to the *M. lagerheimii* gene order (Fig S9), indicating ancient

recombination suppression, which was further supported by high d_S values (Fig S13). This non-recombining region extended farther than the purple and orange strata shared among all *Microbotryum* species, indicating an extension of recombination suppression specific to *M. scorzonareae.* 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. scorzonareae* (Fig S13) and is currently located near the PR locus because of the inter-chromosomal rearrangements (Fig S9). In addition, we detected footprints of a younger specific evolutionary stratum, called quartz, with very low d_S values (Fig S13), 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 S9). 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.

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 (Figures 2, S1-S5). Similarly, expected of 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) (Figure 8). Some TE families seem to have preferentially expanded, i.e. *Copia*, Ty3 (*Gypsy)* and *Helitrons* (Figure 8). As a result, the PR chromosome has become much larger than its ancestral recombining state (Figure 8, 12,930,922 vs 942,259 bp). In contrast, the HD chromosome did not have an elevated TE load (mean 36% of bp, Figure 8), in agreement with the inference of continued recombination in this chromosome.

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 [8,15]. 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. 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 versus only 25% with two mating-type loci [57]. 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 [8]. Linkage between mating-type loci has been observed in *Ustilago hordei* [45] and multiple times independently in the *Malassezia* [46,47] and *Microbotryum* [19,66] clades, as well as in other basidiomycetes [8,119-122].

Loss-of-function at one of the mating-type loci, specifically at the PR locus, has also been observed in several basidiomycete fungi, including *Pholiota nameko* [123], *Coprinellus disseminatus* [48], and others [8]. To the best of our knowledge, this study in *Microbotryum* is the first known case of functional loss at the HD locus in natural populations of heterothallic basidiomycete fungi. 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 anther-smut fungi. Altogether, this sheds new 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 genomic mechanisms.

In addition, finding this example of an 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 dikaryotic growth, 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* [50]. The need for HD heterodimers may be bypassed for dikaryotic growth, as shown in the invasive Californian death cap mushroom, *Amanita phalloides,* that can mate as a haploid homothallic, despite the two HD genes in haploid genomes being unable to form a heterodimer [124]. 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 [125]. This hypothesis would explain the positive selection detected on the new HD coding sequences.

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. 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 [19,66]. In the *Microbotryum* species parasitizing *Dianthus* plants analysed here, a transient stage seems to have occurred, with HD-PR linkage by fusion of the whole PR and HD ancestral chromosomes, followed by a 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.* Reversion to unlinked mating-type chromosomes has been suggested in *Malassezia*-related fungi [47], 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. Alternatively, the fission may have occurred first, followed by the loss of function of the HD genes, selected for unifactoriality, being advantageous under selfing.

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. *M. superbum*, *M. shykoffianum* and *M. scorzonareae* displayed high differentiation between mating-type chromosomes and rearrangements within the non-recombining region, but lower differentiation and lower levels of gene order reshuffling in young evolutionary strata than the rest of the non-recombining region. The evolutionary cause for such stepwise and repeated recombination suppression events may be a combination of selection for less genetic loaded in non-recombining fragments and deleterious-mutation sheltering [32]. Indeed, if a non-recombining fragment traps fewer recessive deleterious alleles than average in the population, this fragment will have a selective advantage, but can only become fixed if it also captures a permanently heterozygous allele that will shelter its few recessive deleterious mutations when it becomes frequent enough to often be homozygous in populations [32]. The only other hypothesis that can explain such repeated evolutionary strata on fungal mating-type chromosomes is a neutral fixation of inversions or genetic differences impairing recombination [30], but this does not

explain the observed association between dikaryotic life cycle and evolutionary strata in fungi [34]. The proximal mechanism of recombination suppression may be the observed movements of fragments from the pseudoautosomal region into the non-recombining region, forming the young strata, as previously reported for the red stratum in *M. lychnidis-dioicae* [18]. 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 [42].

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 [33]. Our scenario actually proposes an early reversion 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 [32,34]. Recombination restoration would be even more difficult if the proximal mechanism of recombination suppression is the movement of pseudo-autosomal regions into non-recombining regions, as observed for young evolutionary strata. Recombination restoration would indeed 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 novel insights into the evolution of breeding systems and sex-related chromosomes, reporting unprecedented 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.

Data availability

RNASeq data for all experiments listed in this section are available in GEO at NCBI, Accession Numbers/Record GSE233508 for the *M. superbum* data.

Acknowledgements

This work was supported by the Louis D. Foundation award and EvolSexChrom ERC advanced grant #832352 to T. G, a National Institute of Health (NIH) grant R15GM119092 to M. E. H., and a National Science Foundation (NSF) grant #2007449 to M.H.P. We thank Klara Koupilova, Janis Antonovics, Emme Bruns and Juha Tuomola for the *M. v. Dianthus* strains, Justus & Maureen Houthuesen-Hulshoff, Julian Woodman, Juha Tuomola for the *M. scorzonerae* strains, and Emme Bruns for *Dianthus* plants. We thank Julian Woodman for the picture of *M. scorzonareae*.

Author contributions

T.G., M.E.H., A.C. and R.R.d.l.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. sequenced and analysed the *M. scorzonerae* genome. E.L., P.J., Q.R., L.B. and R.R.d.l.V. analyzed the PacBio genomes. A.C. acquired the Illumina genomes and performed SNP calling. P.J. and E.L. analysed 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.H.P. and R.K.H. acquired the expression data. R.R.d.l.V. analyzed the RNAseq data. T.G. and E.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.

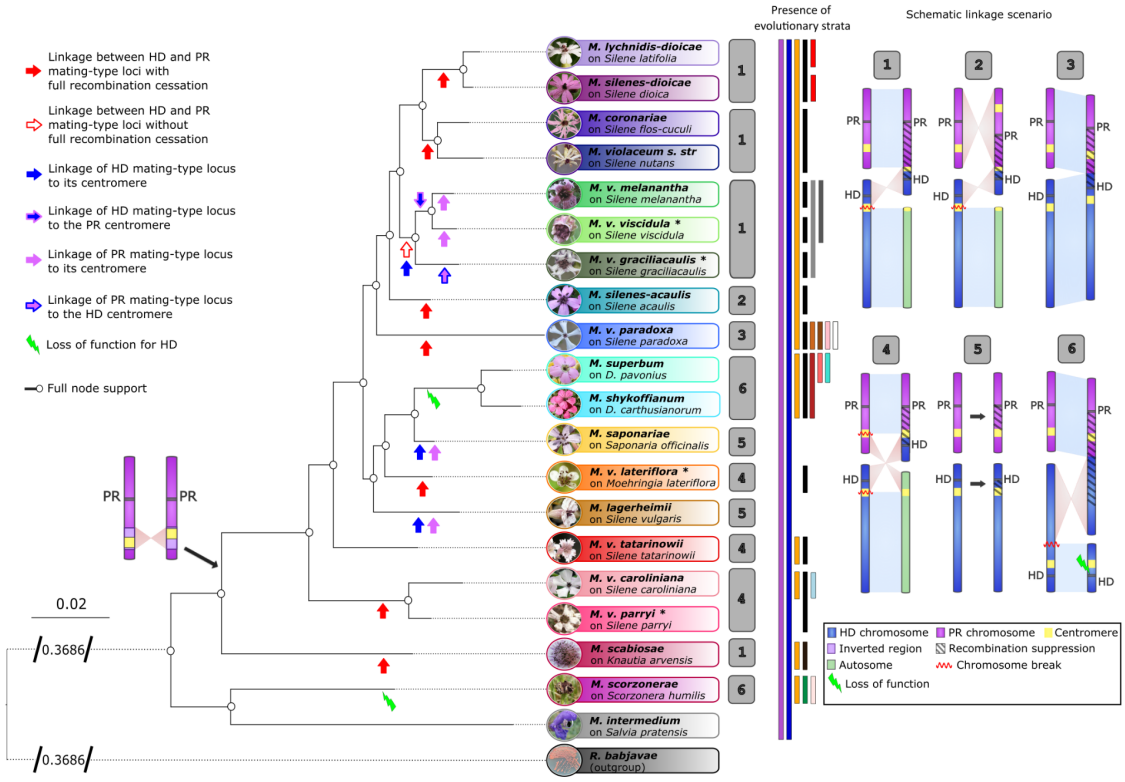

Figure 1: Species tree with pictures of diseased plants parasitized by the different *Microbotryum* species and showing chromosomal arrangements of the mating-type chromosomes compared to the ancestral state. Arrows of different colors on the tree branches indicate mating-type locus linkage, or their 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.

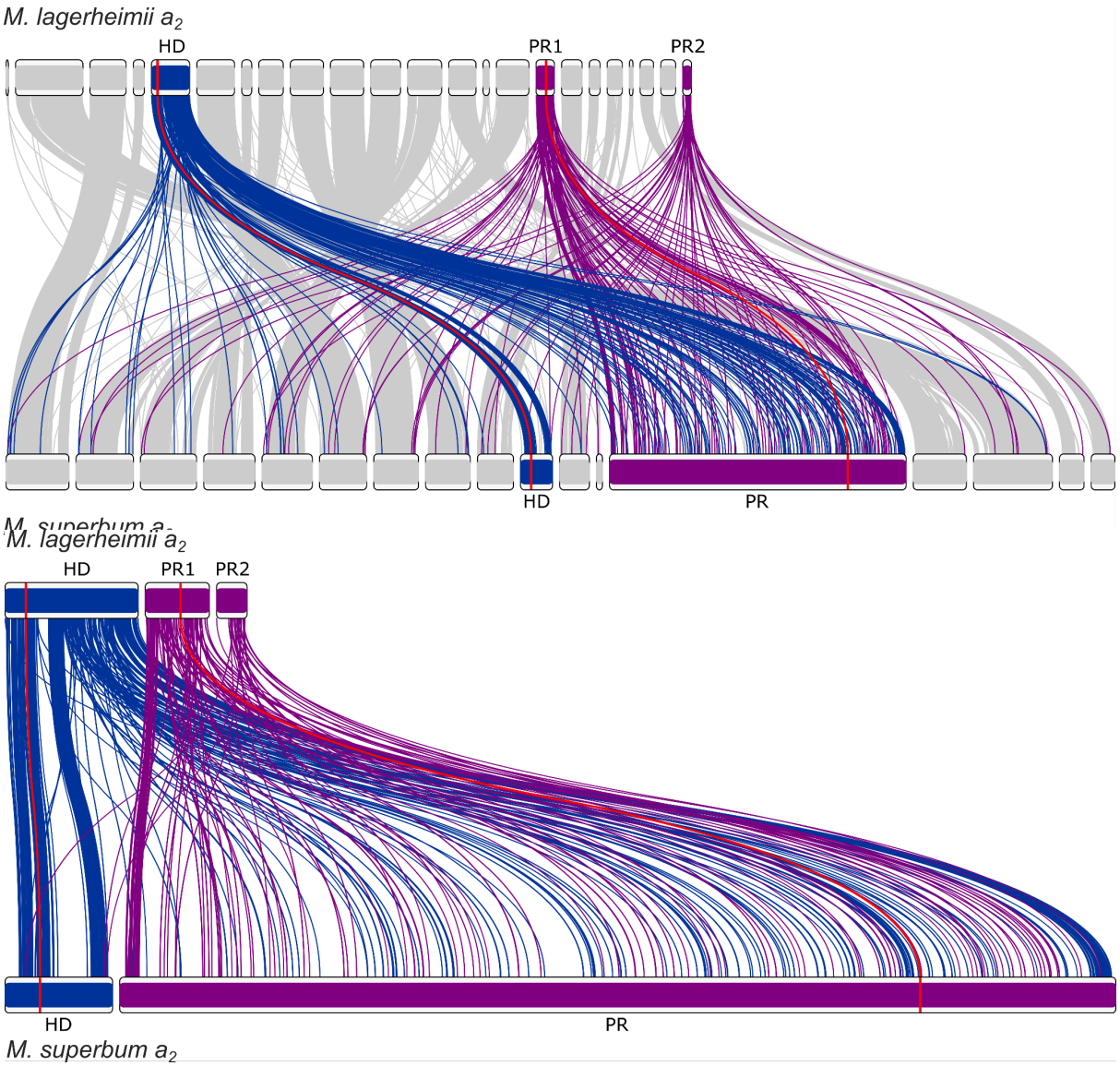

Figure 2: Synteny and rearrangements between the a_1 genomes of *Microbotryum superbum* and *M. lagerheimii*, with all contigs (top) or only the mating-type chromosomes (bottom). The *M. lagerheimii* genome represents a proxy for the ancestral state before recombination suppression and with two separate mating-type chromosomes. The HD mating-type chromosome from *M. lagerheimii* is represented in blue and the PR mating-type chromosome in purple (splitted into two scaffolds), 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. The PR chromosome shows a substantial increase in size and a chaos of rearrangements.

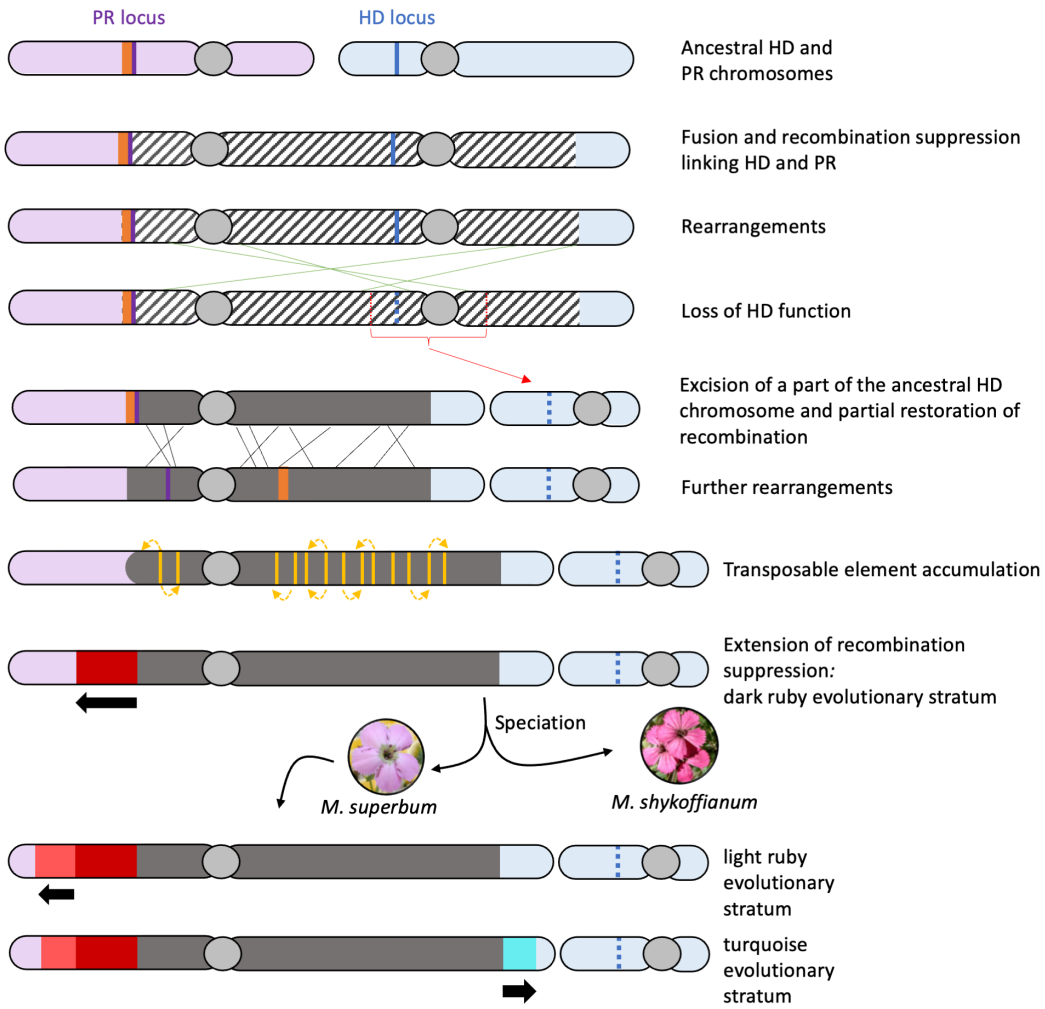

Figure 3: Inferred scenario for the evolution of the mating-type chromosomes in *Microbotryum superbum* and *M. shykoffianum*. Pictures of the diseased plants parasitized by the two species are shown. Recombination suppression is figured by colors corresponding to the different evolutionary strata.

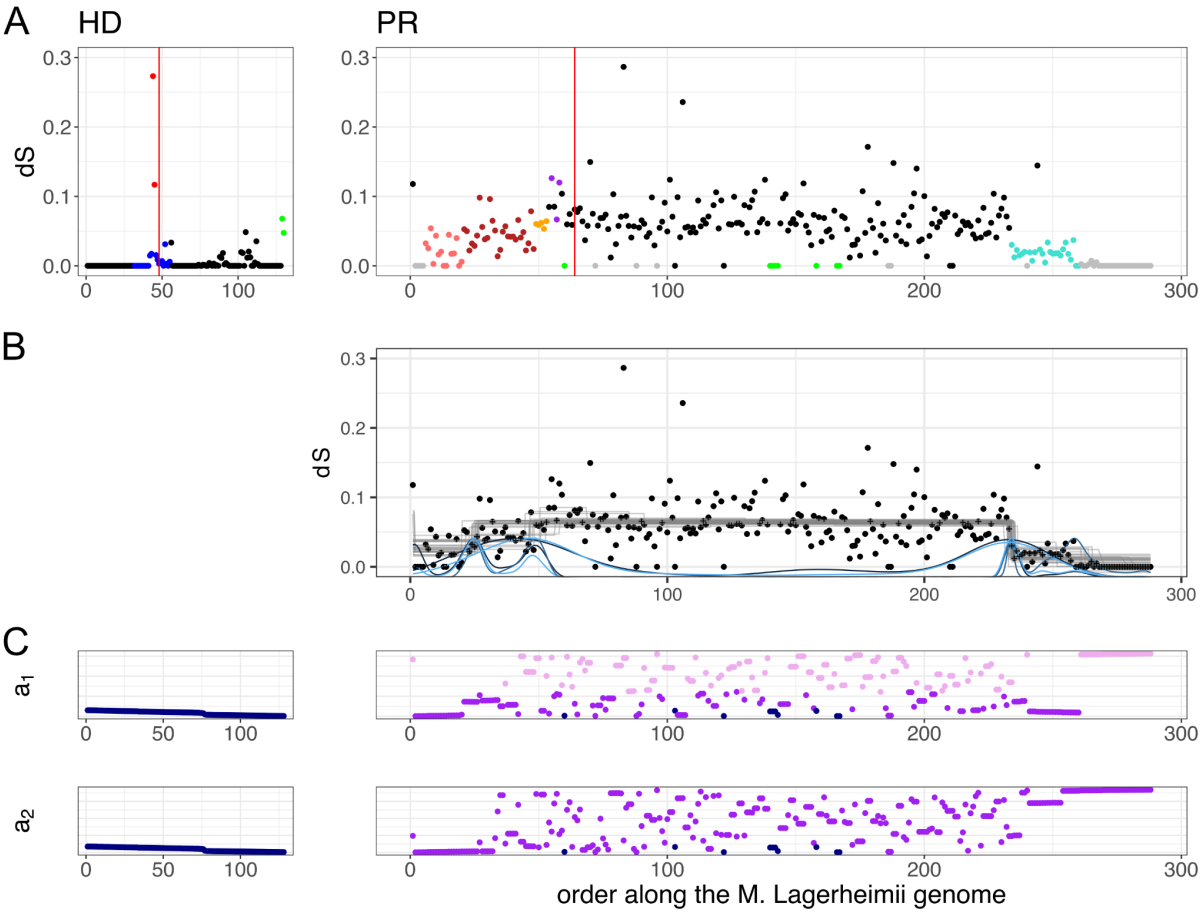

Figure 4: Differentiation and rearrangements between the mating-type chromosomes in *Microbotryum superbum*. A) Per-gene synonymous divergence (d_S) plotted along the ancestral gene order (taking as proxy the gene order along the *M. lagerheimii* mating-type chromosomes), the HD chromosome at left and the PR chromosome at right; 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 grey. 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 dS levels. **C)** Rearrangements compared to the ancestral gene order figured by plotting the gene rank in the current gene order (b1 in the a_1 genome and b_2 in the a_2 genome) as a function of the gene rank in the ancestral gene order.

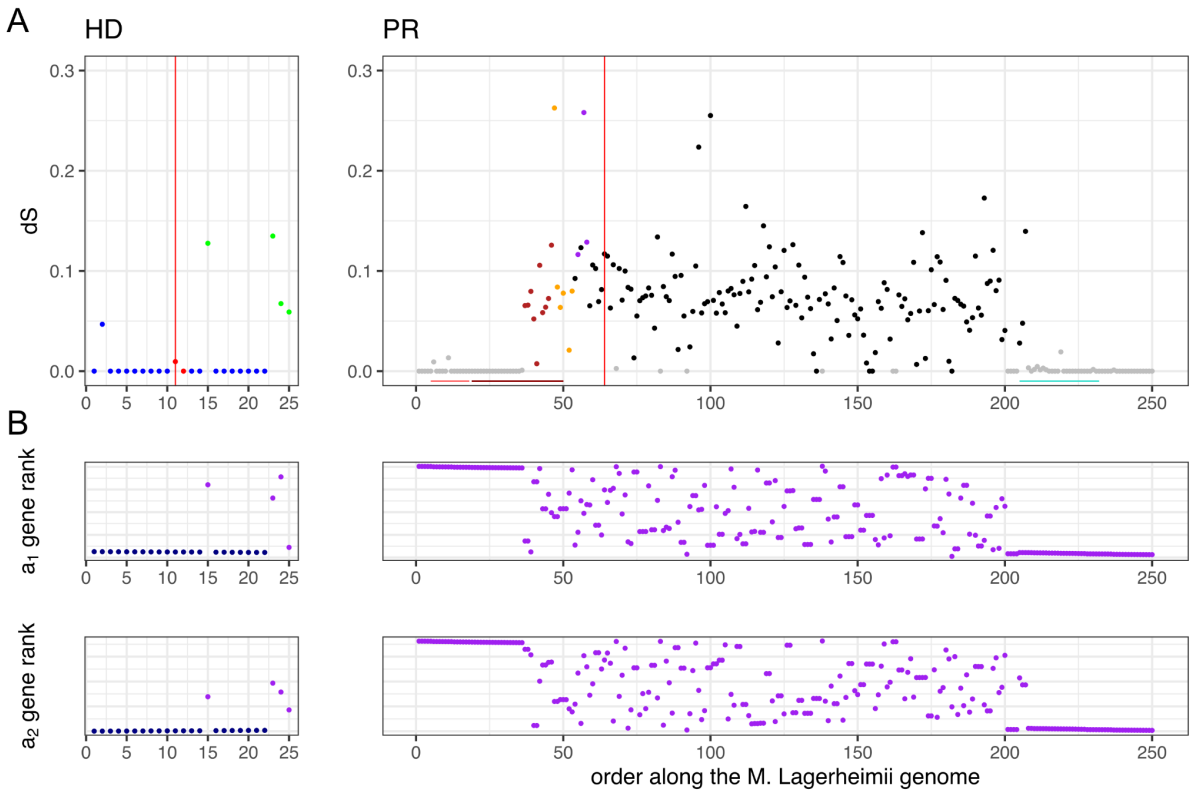

Figure 5: Differentiation and rearrangements between the mating-type chromosomes in *Microbotryum shykoffianum*. a) Per-gene synonymous divergence (d_S) plotted along the ancestral gene order (taking as proxy the gene order along the *M. lagerheimii* mating-type chromosomes)**,** the HD chromosome at left and the PR chromosome at right. The points are colored according to their evolutionary stratum assignment: turquoise, light ruby and half of the dark ruby region for the *M. superbum* specific evolutionary strata, black and the other half of the dark ruby 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 grey. The position of the *M. superbum* specific evolutionary strata are indicated by color bars at the bottom. 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)** Rearrangements compared to the ancestral gene order figured by plotting the gene rank in the current gene order (b_1 in the a_1 genome and b_2 in the a_2 genome) as a function of the gene rank in the ancestral gene order.

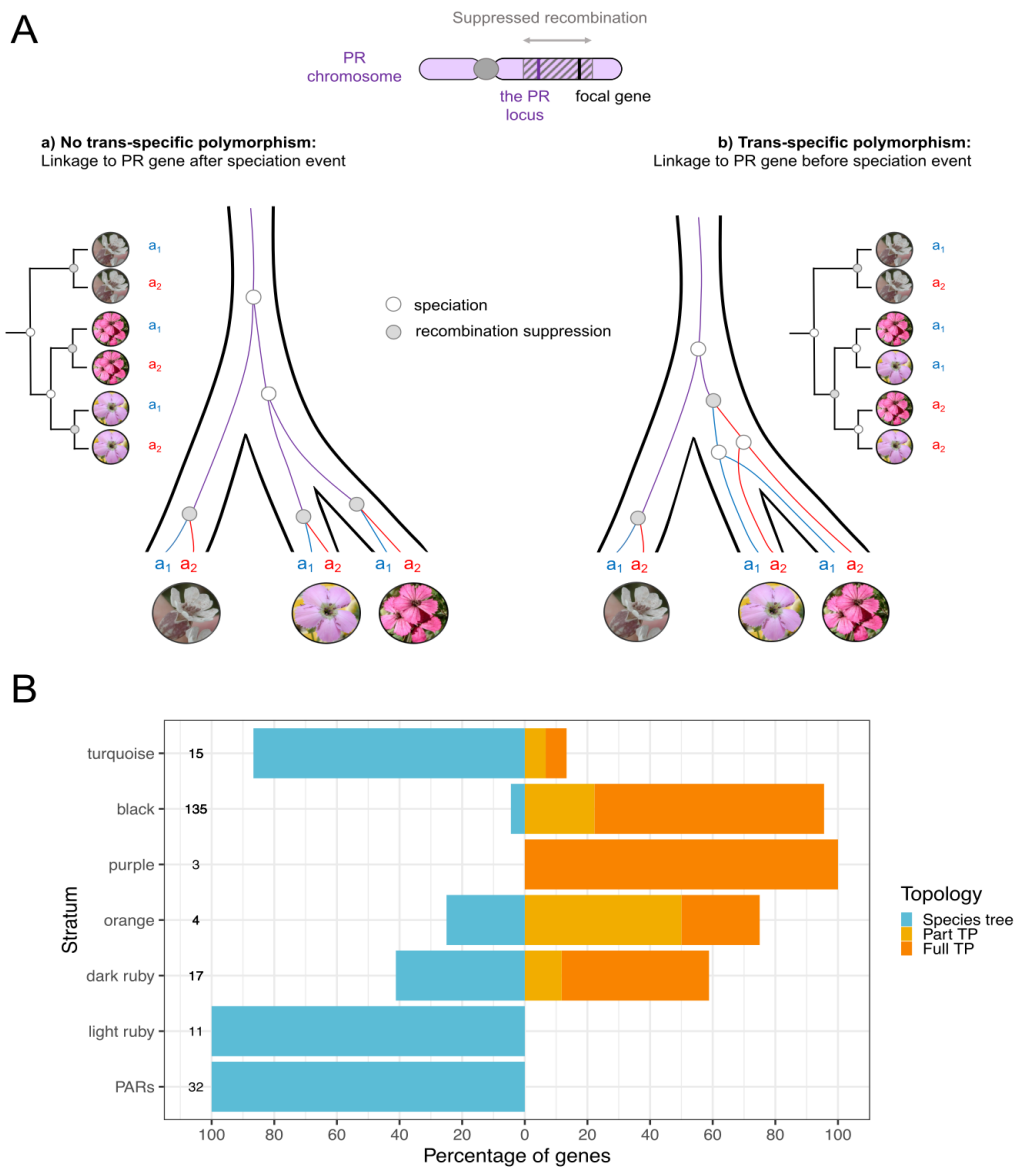

Figure 6: Trans-specific polymorphism (TP) of genes present along the mating-type chromosomes in *Microbotryum superbum*, considering also *M. shykoffianum* and *M. lagerheimii.* A) Schematic representation of trans-specific polymorphism in genes completely linked to mating-type loci by suppressed recombination, 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 no trans-specific polymorphism or unresolved. The total number of genes

analyzed per stratum is indicated at the left of each barplot. The PARs (pseudo-autosomal regions) correspond to the PR chromosome PARs pooled together.

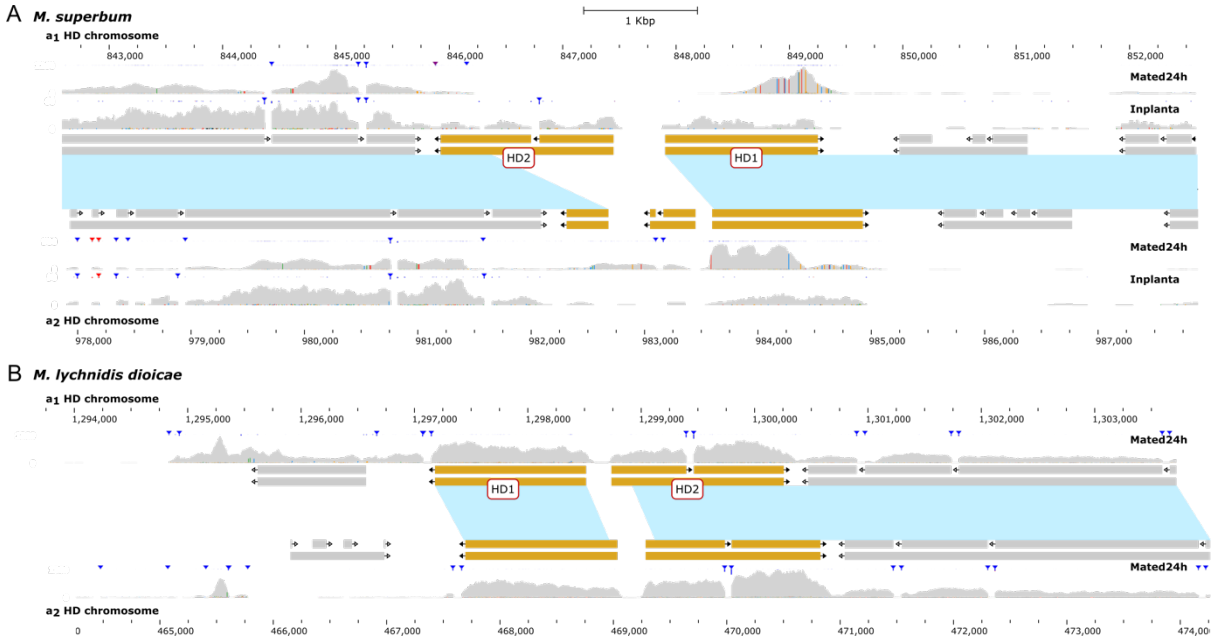

Figure 7: HD gene structure and expression for M. superbum (A) and M. lychnidis dioicae (B). In in the middle of each panel, the coding sequences of the genes present in the a_1 HD chromosome and the a_2 HD chromosome are represented by rectangles, and the direction of the transcription is indicated by an arrow. HD2 and HD1 are represented in yellow rectangles. The blue areas represent the syntenic regions between the a_1 and a_2 sequences. Above the rectangles, the level of expression is represented and annotated according to the condition of the RNAseq extraction (Mated 24h or In planta).

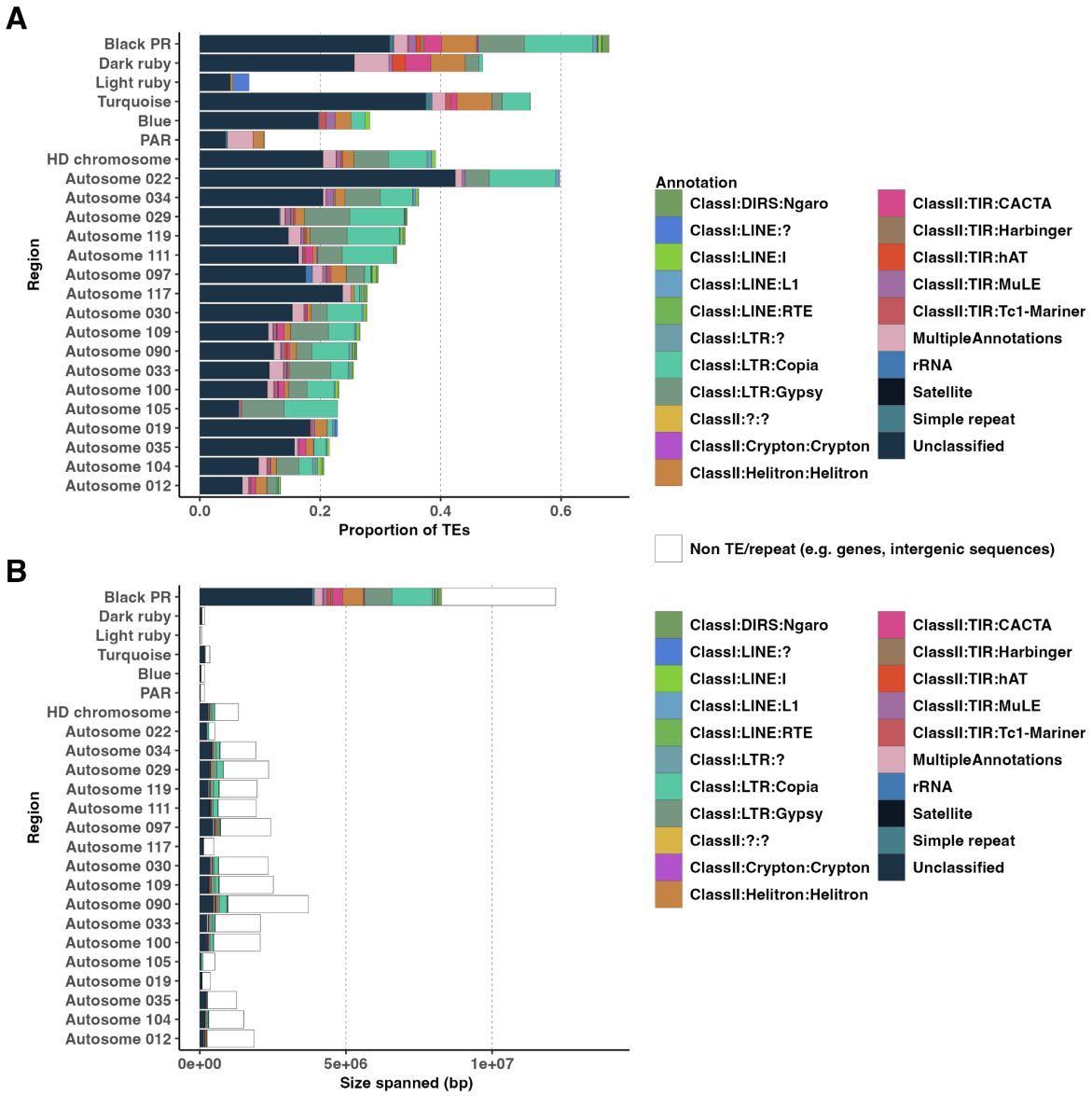

Figure 8: Scaffold size and transposable element (TE) content in the genome of *Microbotryum superbum* a1, as a proportion (A) or absolute content in base pair (bp, B). The relative proportions of transposable element classes in each scaffold is represented by the sizes of the colored bars. The PR chromosome is separated into its different evolutionary strata (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 300kb are displayed.

REFERENCES

1. Lande, R. & Schemske, D. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* **39**, 24-40 (1985).
2. Charlesworth, D. & Charlesworth, B. Inbreeding depression and its evolutionary consequences. *Annu Rev Ecol Evol Syst* **18**, 237-268 (1987).
3. 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).
4. Charlesworth, D. Plant sex determination and sex chromosomes. *Heredity* **88**, 94-101 (2002).
5. Igic, B., Lande, R. & Kohn, J. Loss of self-incompatibility and its evolutionary consequences. *Int J Plant Sci* **169**, 93-104 (2008).
6. Hereford, J. Does selfing or outcrossing promote local adaptation? . *Am J Bot* **97**, 298-302 (2010).
7. Lande, R. Evolution of phenotypic plasticity in colonizing species. *Mol Ecol* **24**, 2038-2045 (2015).
8. Nieuwenhuis, B.P.S. *et al.* Evolution of uni- and bifactorial sexual compatibility systems in fungi. *Heredity* **111**, 445-455 (2013).
9. Goldberg, E. & Igić, B. Tempo and mode in plant breeding system evolution. *Evolution* **66**, 3701-3709 (2012).
10. Goldberg, E. & Igić, B. Tempo and mode in plant breeding system evolution. *Evolution* **66**, 3701-3709 (2012).
11. 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).
12. Hanschen, 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).
13. Beukeboom, L. & Perrin, N. The Evolution of Sex Determination. 92-93 (Oxford University Press, Oxford, 2014).
14. 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).
15. 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).
16. 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).
17. Coelho, M.A., Bakkeren, G., Sun, S., Hood, M.E. & Giraud, T. Fungal Sex: The Basidiomycota. *Microbiology Spectrum* **5**, 1-30 (2017).
18. Branco, S. *et al.* Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. *Proceedings of the National Academy of Sciences* **114**, 7067-7072 (2017).
19. Branco, S. *et al.* Multiple convergent supergene evolution events in mating-type chromosomes. *Nature Communications* **9**, 2000 (2018).
20. Charlesworth, D. The status of supergenes in the 21st century: recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations. *Evolutionary Applications* **9**, 74-90 (2016).

21. Charlesworth, D. Evolution of recombination rates between sex chromosomes. *Philosophical Transactions of the Royal Society B-Biological Sciences* **372**(2017).
22. Charlesworth, D. When and how do sex-linked regions become sex chromosomes? *Evolution* **75**, 569-581 (2021).
23. Bergero, R. & Charlesworth, D. The evolution of restricted recombination in sex chromosomes. *Trends in Ecology & Evolution* **24**, 94-102 (2009).
24. Hartmann, F.E. *et al.* Recombination suppression and evolutionary strata around mating-type loci in fungi: documenting patterns and understanding evolutionary and mechanistic causes. *New Phytologist* **229**, 2470-2491 (2021).
25. Vittorelli, N. *et al.* Stepwise recombination suppression around the mating-type locus in the fungus *Schizothecium tetrasporum* (Ascomycota, Sordariales). *BioRxiv* **https://www.biorxiv.org/content/10.1101/2022.07.20.500756v2**(2022).
26. Ironside, J.E. No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *Bioessays* **32**, 718-726 (2010).
27. Charlesworth, D., Charlesworth, B. & Marais, G. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95** 118-128 (2005).
28. Wright, A.E., Dean, R., Zimmer, F. & Mank, J.E. How to make a sex chromosome. *Nature Communications* **7**, 12087 (2016).
29. 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 Genetics* **9**, 1987-1998 (2019).
30. Jeffries, D.L., Gerchen, J.F., Scharmann, M. & Pannell, J.R. A neutral model for the loss of recombination on sex chromosomes. *Philosophical Transactions of the Royal Society B-Biological Sciences* **376**(2021).
31. Kent, T.V., Uzunovic, J. & Wright, S.I. Coevolution between transposable elements and recombination. *Philosophical Transactions of the Royal Society B-Biological Sciences* **372**(2017).
32. Jay, P., Tezenas, E., Véber, A. & Giraud, T. Sheltering of deleterious mutations explains the stepwise extension of recombination suppression on sex chromosomes and other supergenes. *PLOS Biology* **20**, e3001698 (2022).
33. Lenormand, T. & Roze, D. Y recombination arrest and degeneration in the absence of sexual dimorphism. *Science* **375**, 663-+ (2022).
34. Jay, P., Jeffries, D., Hartmann, F.E. & Giraud, T. Why do sex chromosomes progressively lose recombination? . *Trends in Genetics* (2024).
35. Hill, W.G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genetics Research* **8**, 269-294 (1966).
36. Muller, H.J. Some genetic aspects of sex. *The American Naturalist* **66**, 118-138 (1932).
37. Charlesworth, B. & Charlesworth, D. The degeneration of Y chromosomes. *Philosophical Transactions of the Royal Society B-Biological Sciences* **355**, 1563-1572 (2000).
38. Charlesworth, D. The timing of genetic degeneration of sex chromosomes. *Phil. Trans. R. Soc. B* **376**, 20200093 (2021).
39. Bachtrog, D. Y-chromosome evolution : emerging insights into processes of Y-chromosome degeneration. *Nature Reviews Genetics* **14**, 113-124 (2013).
40. Carpentier, F. *et al.* Tempo of degeneration across independently evolved nonrecombining regions. *Molecular Biology and Evolution* **39**(2022).
41. Fontanillas, E. *et al.* Degeneration of the nonrecombining regions in the mating-type chromosomes of the anther-smut fungi. *Molecular Biology and Evolution* **32**, 928-943 (2015).
42. Duhamel, M., Hood, M.E., Rodriguez 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. *BioRxiv* **https://biorxiv.org/cgi/content/short/2022.08.03.502670v1**(2023).
43. 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).
44. 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.) Pp. 97–160 (Springer, Berlin, Heidelberg, 2011).
45. Bakkeren, G. & Kronstad, J.W. Linkage of mating-type loci distinguishes bipolar from tetrapolar mating in basidiomycetous smut fungi. *Proceedings of the National Academy of Sciences of the United States of America* **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. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 18730-18735 (2007).
47. Coelho, M.A. *et al.* Genetic and genomic analysis of Malassezia reveals transitions in mating-type locus chromosomal organization and early steps in sexual reproduction. *BioRxiv* **https://doi.org/10.1101/2023.03.25.534224** (2023).
48. 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).
49. 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. *G3-Genes Genomes Genetics* **6**, 2135-2146 (2016).
50. Kues, U. *et al.* A chimeric homeodomain protein causes self-incompatibility and constitutive sexual development in the mushroom *Coprinus cinereus*. *Embo Journal* **13**, 4054-4059 (1994).
51. Hood, M.E. *et al.* Distribution of the anther-smut pathogen non species of the Caryophyllaceae. *New Phytologist* **187**, 217-229 (2010).
52. Hartmann, F.E. *et al.* Understanding Adaptation, Coevolution, Host Specialization, and Mating System in Castrating Anther-Smut Fungi by Combining Population and Comparative Genomics. in *Annual Review of Phytopathology, Vol 57, 2019*, Vol. 57 (eds. Leach, J.E. & Lindow, S.E.) 431-457 (2019).
53. 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).
54. 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 Evolutionary Biology* **8**(2008).
55. Hood, M.E., Antonovics, J. & Koskella, B. Shared forces of sex chromosome evolution in haploids and diploids. *Genetics* **168**, 141-146 (2004).
56. Feurtey, A. *et al.* Strong phylogeographic co-structure between the anther-smut fungus and its white campion host. *New Phytologist* **212**, 668-679 (2016).
57. Giraud, T., Yockteng, R., Lopez-Villavicencio, M., Refregier, G. & Hood, M.E. The mating system of the anther smut fungus, *Microbotryum violaceum*: selfing under heterothallism. *Euk. Cell* **7**, 765-775 (2008).
58. 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. *Molecular Ecology* **29**, 1154-1172 (2020).
59. Hood, M.E. & Antonovics, J. Intratetrad mating, heterozygosity, and the maintenance of deleterious alleles in *Microbotryum violaceum (= Ustilago violacea)*. *Heredity* **85**, 231-241 (2000).

60. 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).
61. Antonovics, J., Iwasa, Y. & Hassel, M.P. A generalized model of parasitoid, veneral, and vector-based transmission processes. *Am. Nat.* **145**, 661-675 (1995).
62. Hood, M.E., Antonovics, J. & Heishman, H. Karyotypic similarity identifies multiple host-shifts of a pathogenic fungus in natural populations. *Infection, Genetics and Evolution* **2**, 167-172 (2003).
63. Hartmann, F.E. *et al.* Higher gene flow in sex-related chromosomes than in autosomes during fungal divergence. *Molecular Biology and Evolution* **37**, 668-682 (2020).
64. Zakharov, I.A. [Intratetrad mating and its genetic and evolutionary consequences]. *Genetika* **41**, 508-19 (2005).
65. Zakharov, I.A. Intratetrad mating as the driving force behind the formation of sex chromosomes in fungi. *Trends in Genetics and Evolution* **6**, doi: 10.24294/tge.v6i1.252 (2023).
66. Duhamel, M. *et al.* Onset and stepwise extensions of recombination suppression are common in mating-type chromosomes of *Microbotryum* anther-smut fungi. *Journal of Evolutionary Biology* **35**, 1619-1634 (2022).
67. Carpentier, F. *et al.* Convergent recombination cessation between mating-type genes and centromeres in selfing anther-smut fungi. *Genome Research* **29**, 944-953 (2019).
68. Thomas, A., Shykoff, J., Jonot, O. & Giraud, T. Sex-ratio bias in populations of the phytopathogenic fungus Microbotryum violaceum from several host species. *International Journal of Plant Sciences* **164**, 641-647 (2003).
69. Kaltz, O. & Shykoff, J.A. Sporidial mating-type ratios of teliospores from natural populations of the anther smut fungus Microbotryum (equals Ustilago) violaceum. *International Journal of Plant Sciences* **158**, 575-584 (1997).
70. 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).
71. Horns, F., Petit, E. & Hood, M.E. Massive expansion of *Gypsy*-Like retrotransposons in *Microbotryum* fungi. *Genome Biology and Evolution* **9**, 363-371 (2017).
72. de Vienne, D.M., Hood, M.E. & Giraud, T. Phylogenetic determinants of potential host shifts in fungal pathogens. *Journal of Evolutionary Biology* **22**, 2532-2541 (2009).
73. Petit, E. *et al.* Co-occurrence and hybridization of anther-smut pathogens specialized on Dianthus hosts. *Molecular Ecology* **26**, 1877-1890 (2017).
74. Denchev, C.M., Giraud, T. & Hood, M.E. Three new species of anthericolous smut fungi on Caryophyllaceae. *Mycologica Balcanica* **6**, 79-84 (2009).
75. Liu, T., Tian, H., He, S. & Guo, L. *Microbotryum scorzonerae* (Microbotryaceae), new to China, on a new host plant. *Mycotaxon* **108**, 245-247 (2009).
76. Le Gac, M., Hood, M.E. & Giraud, T. Evolution of reproductive isolation within a parasitic fungal species complex. *Evolution* **61**, 1781-1787 (2007).
77. Perlin, M.H. *et al.* Sex and parasites: Genomic and transcriptomic analysis of *Microbotryum lychnidis-dioicae,* the biotrophic and plant-castrating anther smut fungus. *BMC genomics* **16**, 1-24 (2015).
78. San Toh, S. *et al.* Pas de deux: An Intricate Dance of Anther Smut and Its Host. **8**, 505-518 (2018).
79. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018).
80. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
81. Diesh, C. *et al.* JBrowse 2: a modular genome browser with views of synteny and structural variation. *Genome Biology* **24**, 74 (2023).

82. Gel, B. & Serra, E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* **33**, 3088-3090 (2017).
83. Koren, S. *et al.* Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research* **27**, 722-736 (2017).
84. Cheng, H., Concepcion, G., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods* **18**, 170-175 (2021).
85. 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. *Molecular Biology and Evolution* **38**, 4647-4654 (2021).
86. De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M. & Van Broeckhoven, C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* **34**, 2666-2669 (2018).
87. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P.A. Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology* **37**, 540-546 (2019).
88. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884-i890 (2018).
89. Vasimuddin, M., Misra, S., Li, H. & Aluru, S. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 314-324 (2019).
90. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
91. Walker, B.J. *et al.* Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLOS ONE* **9**, e112963 (2014).
92. Flynn, J.M. *et al.* RepeatModeler2 for automated genomic discovery of transposable element families. **117**, 9451-9457 (2020).
93. Smit, A., Hubley, R. & Green, P. RepeatMasker Open-4.0. <http://www.repeatmasker.org> (2013-2015).
94. Bao, W., Kojima, K.K. & Kohany, O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA* **6**, 11 (2015).
95. 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 Genomics and Bioinformatics* **3**, lqaa108 (2021).
96. 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.) 283-334 (Springer New York, New York, NY, 2016).
97. Keller, O., Kollmar, M., Stanke, M. & Waack, S. A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics* **27**, 757-63 (2011).
98. Gabriel, L., Hoff, K.J., Brůna, T., Borodovsky, M. & Stanke, M. TSEBRA: transcript selector for BRAKER. **22**, 566 (2021).
99. Emms, D.M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**, 238 (2019).
100. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410 (1990).
101. Devier, B., Aguileta, G., Hood, M. & Giraud, T. Ancient trans-specific polymorphism at pheromone receptor genes in basidiomycetes. *Genetics* **181**, 209–223 (2009).
102. Petit, E. *et al.* Linkage to the mating-type locus across the genus *Microbotryum*: insights into non-recombining chromosomes. *Evolution* **66**, 3519-3533 (2012).
103. Hao, Z. *et al.* RIdeogram: drawing SVG graphics to visualize and map genome-wide data on the idiograms. *PeerJ Comput Sci* **6**, e251 (2020).

104. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811-2 (2014).
105. 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. *Molecular Biology and Evolution* **35**, 2582-2584 (2018).
106. Wertheim, J.O., Murrell, B., Smith, M.D., Kosakovsky Pond, S.L. & Scheffler, K. RELAX: Detecting Relaxed Selection in a Phylogenetic Framework. *Molecular Biology and Evolution* **32**, 820-832 (2015).
107. Ranwez, V., Harispe, S., Delsuc, F. & Douzery, E.J.P. MACSE: Multiple alignment of coding sequences accounting for frameshifts and stop codons. *PLOS ONE* **6**, e22594 (2011).
108. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**, 1586-1591 (2007).
109. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis (2nd. ed.)*, (Springer Publishing Company, 2009).
110. Lindeløv, J.K. Mcp: An R package for regression with multiple change points. *https://doi.org/10.31219/osf.io/fzqxv* (2020).
111. Minh, B.Q. *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**, 1530-1534 (2020).
112. Lopez-Villavicencio, M. *et al.* Multiple infections by the anther smut pathogen are frequent and involve related strains. *PloS Pathogens* **3**, e176 (2007).
113. Untergasser, A. *et al.* Primer3--new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).
114. Castle, A.J. & Day, A.W. Isolation and Identification of α -Tocopherol as an Inducer of the Parasitic Phase of *Ustilago violacea Phytopathology* **74**, 1194-1200 (1984).
115. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811-2812 (2014).
116. Kurtz, S. *et al.* Versatile and open software for comparing large genomes. *Genome Biology* **5**, R12 (2004).
117. Hood, M.E. & Antonovics, J. Mating Within the Meiotic Tetrad and the Maintenance of Genomic. **1759**, 1751-1759 (2004).
118. Day, A. & Jones, J. The production and characteristics of diploids in *Ustilago violacea*. *Genet. Res., Camb.* **11**, 63-81 (1968).
119. 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. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 15026-15031 (1999).
120. Fraser, J.A. *et al.* Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms. *Plos Biol.* **2**, 2243-2255 (2004).
121. Rabe, F. *et al.* A complete toolset for the study of *Ustilago bromivora* and *Brachypodium sp* as a fungal-temperate grass pathosystem. *Elife* **5**(2016).
122. Sun, S., Coelho, M.A., Heitman, J. & Nowrousian, M. Convergent evolution of linked mating-type loci in basidiomycete fungi. *PLoS Genet.* **15**, e1008365 (2019).
123. Yi, R. *et al.* Genomic structure of the A mating-type locus in a bipolar basidiomycete, Pholiota nameko. *Mycol Res.* **113**, 240-248 (2009).
124. Wang, Y.-W. *et al.* Invasive Californian death caps develop mushrooms unisexually and bisexually. *Nature Communications* **14**, 6560 (2023).
125. Cabrita, A. *et al.* Multiple pathways to homothallism in closely related yeast lineages in the Basidiomycota. *mBio* **12**, 10.1128/mbio.03130-20 (2021).