



**HAL**  
open science

## Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism

Sara E Branco, H el ene E Badouin, Ricardo E Rodr iguez de La Vega, Jerome Gouzy, Fantin E Carpentier, Gabriela E Aguilera, Sophie A Siguenza, Jean-Tristan E Brandenburg, Marco E Coelho, Michael E Hood, et al.

### ► To cite this version:

Sara E Branco, H el ene E Badouin, Ricardo E Rodr iguez de La Vega, Jerome Gouzy, Fantin E Carpentier, et al.. Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. *Proceedings of the National Academy of Sciences of the United States of America*, 2017, 114 (27), pp.7067-7072. 10.1073/pnas.1701658114 . hal-01547790

**HAL Id: hal-01547790**

**<https://hal.science/hal-01547790v1>**

Submitted on 27 Jun 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

Copyright

# Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism

Sara Branco<sup>a,1</sup>, H el ene Badouin<sup>a,b,1</sup>, Ricardo C. Rodr iguez de la Vega<sup>a,1</sup>, J er ome Gouzy<sup>b</sup>, Fantin Carpentier<sup>a</sup>, Gabriela Aguilera<sup>a</sup>, Sophie Siguenza<sup>b</sup>, Jean-Tristan Brandenburg<sup>a</sup>, Marco A. Coelho<sup>c</sup>, Michael E. Hood<sup>d</sup>, and Tatiana Giraud<sup>a,2</sup>

<sup>a</sup>Ecologie Syst ematique Evolution, Univ. Paris Sud, AgroParisTech, CNRS, Universit  Paris-Saclay, 91400 Orsay, France; <sup>b</sup>Laboratoire des Interactions Plantes-Microorganismes, Universit  de Toulouse, Institut National de la Recherche Agronomique, CNRS, 31326 Castanet-Tolosan, France; <sup>c</sup>UCIBIO-REQUIMTE, Departamento de Ci ncias da Vida, Faculdade de Ci ncias e Tecnologia, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal; and <sup>d</sup>Department of Biology, Amherst College, Amherst, MA 01002-5000

Edited by James J. Bull, The University of Texas at Austin, Austin, TX, and approved May 23, 2017 (received for review February 9, 2017)

Sex chromosomes can display successive steps of recombination suppression known as “evolutionary strata,” which are thought to result from the successive linkage of sexually antagonistic genes to sex-determining genes. However, there is little evidence to support this explanation. Here we investigate whether evolutionary strata can evolve without sexual antagonism using fungi that display suppressed recombination extending beyond loci determining mating compatibility despite lack of male/female roles associated with their mating types. By comparing full-length chromosome assemblies from five anther-smut fungi with or without recombination suppression in their mating-type chromosomes, we inferred the ancestral gene order and derived chromosomal arrangements in this group. This approach shed light on the chromosomal fusion underlying the linkage of mating-type loci in fungi and provided evidence for multiple clearly resolved evolutionary strata over a range of ages (0.9–2.1 million years) in mating-type chromosomes. Several evolutionary strata did not include genes involved in mating-type determination. The existence of strata devoid of mating-type genes, despite the lack of sexual antagonism, calls for a unified theory of sex-related chromosome evolution, incorporating, for example, the influence of partially linked deleterious mutations and the maintenance of neutral rearrangement polymorphism due to balancing selection on sexes and mating types.

evolutionary strata | chromosomal rearrangements | fungi | genomic degeneration | mating-type chromosomes

Chromosomes carrying genes controlling mating compatibility may display suppressed recombination, as seen in sex chromosomes in some animals and plants (1, 2). When recombination is suppressed and one of the sex chromosomes is always in a heterozygous state (e.g., the Y chromosome in humans), extensive rearrangements of that chromosome typically occur. Our understanding of how sex chromosomes evolve remains incomplete, with many unanswered questions (3), including the reasons why these chromosomes display such large regions of suppressed recombination, extending well beyond the sex-determining genes. Sex chromosomes may bear discrete “evolutionary strata” of differentiation between alleles, with divergence decreasing with distance from sex-determining genes in the ancestral gene order, as inferred from the nonrearranged sex chromosome (1). It is generally accepted that such strata result from successive rounds of selection favoring the linkage of sexually antagonistic genes (with alleles beneficial in one sex but detrimental in the other) and sex-determining loci (1, 4). However, there is no definitive evidence to support a prominent role for sexually antagonistic genes in the formation of evolutionary strata in sex chromosomes (3). Other hypotheses have been put forward, including the permanent sheltering of deleterious alleles, which may accumulate in the margins of nonrecombining regions, provided that recombination rates are low in these regions. Another hypothesis is the fixation of neutral rearrangements

by drift in one of the two sex chromosomes with automatic recombination arrest and the maintenance of a polymorphic state due to balancing selection on sexes (2, 5–8) (*SI Appendix, SI Text and Fig. S1*).

Recombination suppression has been documented in fungal mating-type chromosomes (9, 10), which carry key loci determining mating compatibility. In basidiomycetes, mating type is typically controlled by two loci: (i) the pheromone receptor (PR) locus, which carries one gene encoding a mating-type-specific PR and one to several genes encoding pheromones, and (ii) the homeodomain gene (HD) locus, encoding two homeodomain-type transcription factors controlling postmating growth. Recombination suppression ensures full linkage of the mating-type genes within each of these two loci, which is required for correct mating-type determination (11), and typically does not extend beyond the mating-type genes (12, 13). Recombination suppression

## Significance

Sex chromosomes can display divergent evolution, as seen in humans, in which the Y chromosome underlying maleness is smaller and contains much less information than the X chromosome. The differentiation between sex chromosomes can occur stepwise along their length, which is thought to result from the successive beneficial linkage of genes with different phenotype optima in the two sexes to sex-determining genes. However, there is little evidence to support this hypothesis. Here, we recovered ancestral chromosome structures and gathered evidence for stepwise differentiation between fungal mating-type chromosomes despite the absence of male/female roles. Our results suggest that the analogous features of sex chromosomes may not be due to differences in selection between males and females.

Author contributions: M.E.H. and T.G. designed research; M.E.H. and T.G. jointly supervised the study; S.B., H.B., R.C.R.d.l.V., J.G., F.C., G.A., S.S., J.-T.B., M.A.C., M.E.H., and T.G. performed research; S.B., H.B., R.C.R.d.l.V., F.C., M.E.H., and T.G. analyzed data; and S.B., M.E.H., and T.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The genome assemblies are available at the European Nucleotide Archive. Accession numbers are as follows: PRJEB12080, ERS459551, and ERZ250722 for *Microbotryum lychnidis-dioicae* 1064 a2; PRJEB12080, ERS1013679, and ERZ250721 for *M. lychnidis-dioicae* 1064 a1; PRJEB12080, ERS1013678, and ERZ250720 for *Microbotryum lagerheimii* 1253 a2; PRJEB12080, ERS1013677, and ERZ250719 for *M. lagerheimii* 1253 a1; PRJEB12080, ERS1013672, and ERZ250714 for *Microbotryum violaceum* s. str. 1249 a2; PRJEB12080, ERS1013671, and ERZ250713 for *M. violaceum* s. str. 1249 a1; PRJEB16471 and ERZ369313 for *Microbotryum intermedium*; PRJEB16741, ERS1436593, and ERZ370138 for *Microbotryum silenes-dioicae* 1303 a2; and PRJEB16741, ERS1436592, and ERZ369313 for *M. silenes-dioicae* 1303 a1.

<sup>1</sup>S.B., H.B., and R.C.R.d.l.V. contributed equally to this work.

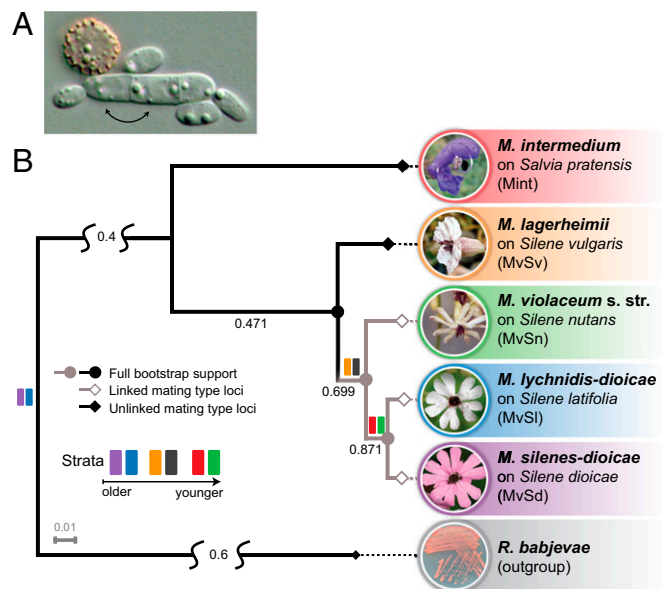
<sup>2</sup>To whom correspondence should be addressed. Email: tatiana.giraud@u-psud.fr.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1701658114/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1701658114/-DCSupplemental).

in fungal mating-type chromosomes can further link the two mating-type loci, HD and PR, to each other and/or the centromere (9, 14–18). Such linkage is beneficial in species with selfing-based mating systems, as it increases the odds of compatibility between gametes from a diploid parent (16, 19, 20) (*SI Appendix, SI Text and Fig. S2*). It has been repeatedly suggested that some fungal mating-type chromosomes have additional evolutionary strata extending the suppression of recombination beyond mating-type genes (9, 18, 21, 22) (*SI Appendix, SI Text*). However, no evidence has been provided of multiple discrete regions in which divergence between fungal mating-type-associated alleles decreases with increasing distance from sexual compatibility loci. Uncertainty about fungal evolutionary strata persists due to difficulties in inferring the ancestral gene order, which is essential for reconstructions of the evolution of mating-type chromosomes. If fungi have evolutionary strata not involving the linkage of mating-type genes, their presence cannot be explained by sexually antagonistic selection, as fungal mating types are not associated with male/female functions. Moreover, there are virtually no differences in ecological or life-history traits between fungal mating types (23), making it unlikely that there is any analogous mating-type-antagonistic selection improving the function of one mating type while having deleterious effects on the function of the other.

We investigated the existence of evolutionary strata in *Microbotryum lychnidis-dioicae*, a castrating anther-smut fungus. This species is particularly suited for determining whether the stepwise evolution of suppressed recombination can occur without sexual antagonism. It has two alternative mating-type chromosomes with a large region of suppressed recombination (10, 17, 24, 25), complete absence of male/female function, and virtually no opportunity for mating-type-antagonistic selection. Indeed, almost all matings occur within the tetrad (26–28) with the isogamous gametes fusing rapidly after meiosis without mitotic proliferation in a free haploid stage (Fig. 1*A*). There has been much debate about the existence of evolutionary strata in the *M. lychnidis-dioicae* mating-type chromosomes (17, 21, 22, 25) (*SI Appendix, SI Text*). Genes with different levels of synonymous divergence between alleles linked to the two mating types (called  $a_1$  and  $a_2$ ) are widely dispersed over 90% of the chromosome length, suggesting that, if there were any evolutionary strata, they have been erased by massive rearrangements (17). As in all fungi with distinct mating types, and in contrast to plant and animal sex chromosomes, *M. lychnidis-dioicae* mating-type chromosomes are always heterozygous in the diploid stage, so neither conserves the ancestral gene order (17). This precludes the typical use of the nonrearranged chromosome for the detection of evolutionary strata (1). The recombining regions in *M. lychnidis-dioicae*, such as the autosomes and the pseudoautosomal regions (PARs), are, however, highly collinear (17) (*SI Appendix, Fig. S3*), suggesting that the ancestral gene order should be retained in mating-type chromosomes undergoing recombination, even in selfing species. We use the term “autosome” as a contrast to “dimorphic” chromosomes, as originally intended (29, 30), rather than in opposition to sex chromosomes per se. Similarly, PAR is used to name the genomic regions at the edges of the mating-type chromosomes that are collinear and recombine like autosomes.

We compiled chromosome-length genome assemblies for five *Microbotryum* species and adopted an approach in which species carrying mostly recombining mating-type chromosomes were identified and used to recover the ancestral chromosomal state and gene order. Using this method, we were able to reconstruct the rearrangements leading to the current nonrecombining mating-type chromosomes in the selfing *M. lychnidis-dioicae*. One important step involved chromosomal rearrangements and fusion of the two separate ancestral mating-type chromosomes, linking the two mating-type loci in the same chromosome, which is beneficial under selfing (although not required) (*SI Appendix,*



**Fig. 1.** Anther-smut fungi. (*A*) *M. lychnidis-dioicae* tetrad with a cell in the spore and three cells in a horizontal promycelium; mating occurs rapidly after meiosis within the tetrad (black arrows). Some haploid cells are budding, but almost 100% of matings occur within the promycelium. (*B*) *Microbotryomycete* phylogenetic tree based on 4,000 orthologous genes, including the studied *Microbotryum* species (shown in the anthers of their host plants) and the outgroup *Rhodotorula babjevae*. Branch color and symbol indicate linked (gray branches and open diamonds) or unlinked (black branches and closed diamonds) mating-type loci. The closed circles indicate full bootstrap support. Internode certainty with no conflict bipartitions is given below the branches (i.e., the normalized frequency of the most frequent bipartition across gene genealogies relative to the summed frequencies of the two most frequent bipartitions), indicating good support for the nodes. Relative tree certainty of the tree was 0.68. Colored rectangles along the branches indicate the evolution of the various evolutionary strata (see Fig. 3).

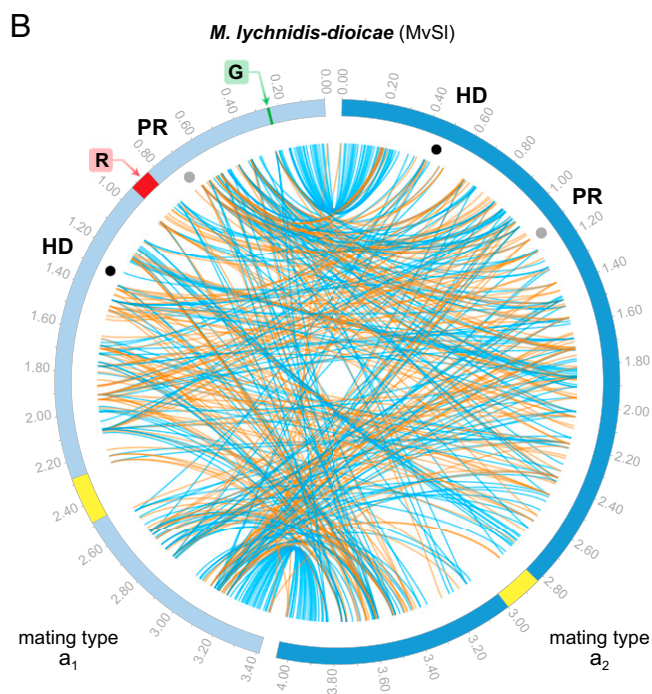
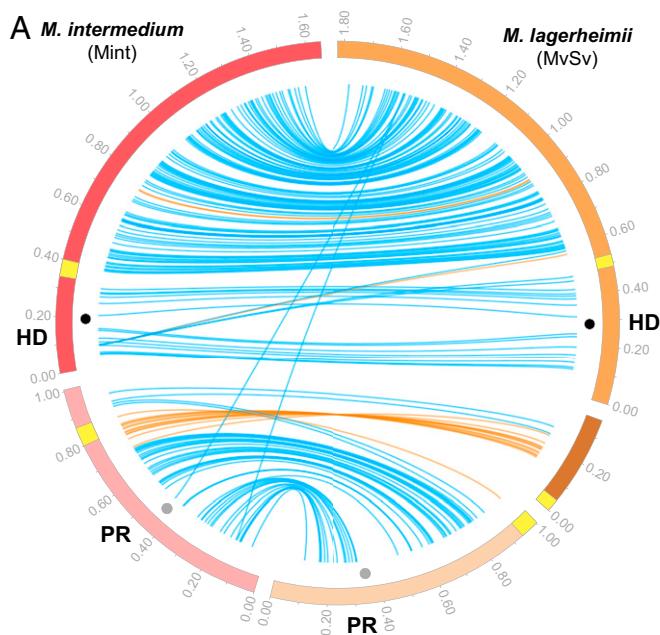
*Fig. S2*) (16). This finding sheds light on the genomic evolutionary steps leading to changes in breeding systems in fungi. We also provide compelling evidence for multiple, clearly differentiated and ancestrally adjacent evolutionary strata in fungal mating-type chromosomes with several strata devoid of genes controlling mating types. The existence of successive extensions of the region of recombination suppression beyond the linkage of mating-type genes in fungi without male/female functions indicates that evolutionary strata can occur without sexual antagonism. This finding has profound implications for our understanding of the evolutionary genomics of sexual eukaryotes across three kingdoms, highlighting the need for more serious consideration of alternative theoretical models of sex-related chromosome evolution beyond sexually antagonistic processes.

## Results and Discussion

### Ancestral Gene Order in *Microbotryum* Mating-Type Chromosomes.

We obtained high-quality assemblies of haploid genomes for two *Microbotryum* species expected to have retained the ancestral gene order: *Microbotryum intermedium*, a distant relative of *M. lychnidis-dioicae* (Fig. 1*B*) (31, 32), and *Microbotryum lagerheimii*, a species known to have unlinked HD and PR mating-type loci (32). As hypothesized, *M. intermedium* was found to have the PR and HD loci in two different chromosomes that were highly collinear with the mating-type chromosomes of *M. lagerheimii* (Fig. 2*A*). The collinearity between these distantly related species indicated a high degree of gene order conservation across multiple speciation events (31), therefore closely reflecting the





**Fig. 2.** Comparison of gene order across *Microbotryum* mating-type chromosomes. The mating-type HD and PR genes are indicated by small black and gray circles, respectively. Blue and orange lines link regions with collinearity extending over 2 kb, the latter corresponding to inversions. Areas without links correspond to highly rearranged regions. Yellow regions on the outer track indicate centromere-specific repeats (17). (A) Comparison of gene order between the mating-type chromosomes of *M. intermedium* (Left) and *M. lagerheimii* (Right). The darkest contig on the Right corresponds to MC16, as discussed in *SI Appendix, SI Text*. (B) Comparison of  $a_1$  (light contig) and  $a_2$  (dark contig) *M. lychnidis-doicae* mating-type chromosomes. The red and green genes in Fig. 4B are indicated in the  $a_1$  genome.

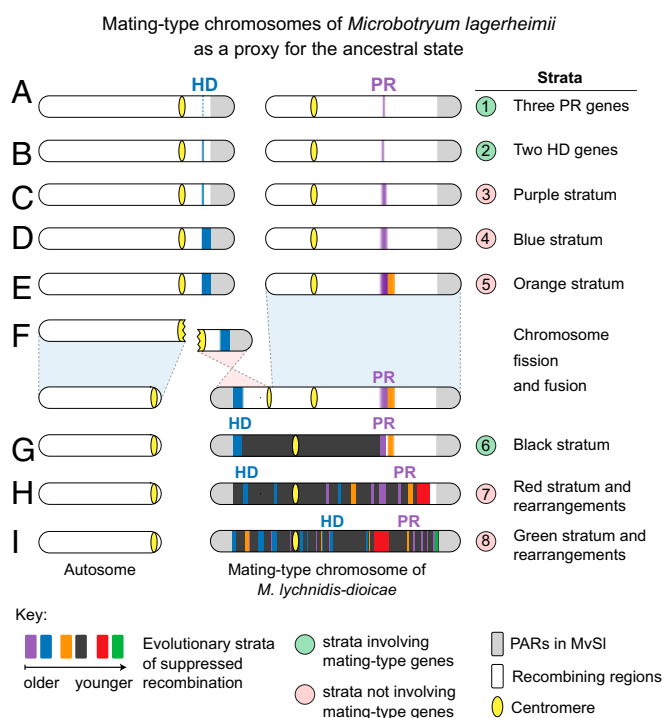
ancestral state in this clade. The ancestral gene order has probably been maintained by regular recombination, as also indicated by the collinearity between the  $a_1$  and  $a_2$  mating-type chromosomes in *M. lagerheimii* (*SI Appendix, Fig. S4*), typical of recombining genomic regions (*SI Appendix, Fig. S3*).

### Ancient Linkage of Mating-Type Genes at Each of the HD and PR Loci.

We found footprints of ancient recombination suppression linking the essential pairs of interacting mating-type genes at each of the HD and PR loci, as is typical and ancestral in basidiomycete fungi (13). The nucleotide sequences of the  $a_1$  and  $a_2$  alleles of the PR gene were too differentiated to be aligned (Figs. 3A and 4) and a previous study inferred that these alleles had been diverging for about 370 million years (33). Recombination suppression linking the two HD mating-type genes was also very old (Fig. 3B), as shown by the second-highest level of synonymous divergence ( $d_s$ ) between the  $a_1$  and  $a_2$  mating types at these genes in both *M. lagerheimii* and *M. lychnidis-doicae* (Fig. 4) and by gene genealogies (34).

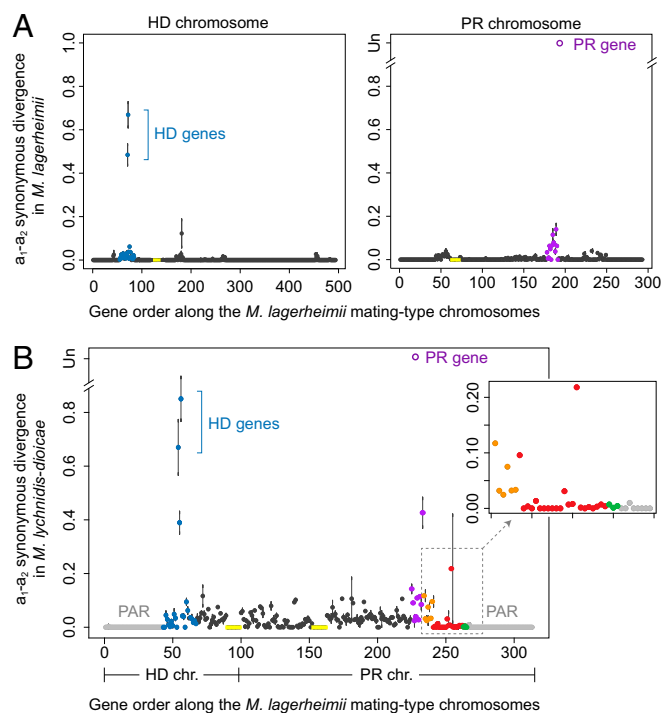
### Chromosomal Fusion Led to the Large Nonrecombining Region Linking HD and PR Mating-Type Genes in *M. lychnidis-doicae*.

A more recent recombination suppression event involved the linkage of the HD and PR mating-type loci to each other (Fig. 3), which is favorable under selfing (15, 16) (*SI Appendix, Fig. S2*), the predominant mating system in *Microbotryum* (26, 28). Genome comparisons indicated that the *M. lychnidis-doicae* mating-type chromosome was



**Fig. 3.** Scenario of mating-type chromosome evolution in the genus *Microbotryum*. Strata involving the linkage of mating-type genes are numbered in green circles, whereas strata devoid of mating-type genes are numbered in red circles. (A) Ancestral state with separate HD and PR chromosomes and suppressed recombination restricted to the PR mating-type genes. (B) Suppressed recombination at all mating-type genes (as typically seen in basidiomycete fungi, with linked PR and pheromone genes at one locus and the two linked HD genes linked at a second locus). (C) Extension of suppressed recombination to purple nonmating-type genes. (D) Extension of suppressed recombination to blue nonmating-type genes. (E) Extension of suppressed recombination distal to the purple stratum, forming the orange stratum, devoid of mating-type genes. (F) Fission of the HD chromosome and fusion of one arm with the PR chromosome. (G) Complete linkage between the HD and PR mating-type loci, forming the black stratum. (H) Extension of suppressed recombination to nonmating-type genes forming the red stratum and rearrangements shuffling older evolutionary strata. (I) Further extension of suppressed recombination forming the green stratum and further rearrangements, leading to the current *M. lychnidis-doicae* (MvSI) mating-type chromosome with the red stratum located between the PR and HD genes.





**Fig. 4.** Synonymous divergence ( $d_s$ ) between alleles associated with the  $a_1$  and  $a_2$  mating types along mating-type chromosomes, as found in the PacBio-sequenced diploid individual in each *Microbotryum* species. Per-gene mean  $d_s$  values and SE values. The locations of the mating-type genes (PR and HD) are indicated. The divergence between the  $a_1$  and  $a_2$  PR alleles was too extensive (33) and could not be computed. It is plotted as an “unaligned” (Un) open purple circle. The *M. lagerheimii* centromeres are shown in yellow. (A) Synonymous divergence for all single-copy genes common to the *M. lagerheimii*  $a_1$  and  $a_2$  mating-type chromosomes plotted according to the genomic coordinates of its  $a_1$  mating-type chromosomes. The HD chromosome is displayed on the left and the PR chromosome on the right. Genes with within-individual  $a_1$ - $a_2$   $d_s > 0$  are shown in blue around HD and purple around PR. The mating-type loci are partially linked to the centromere in *M. lagerheimii* (32). (B) Synonymous divergence for all single-copy genes common to the *M. lychnidis-dioicae* mating-type chromosomes and *M. lagerheimii*, plotted in the ancestral gene order, as inferred from the genomic coordinates of the *M. lagerheimii*  $a_1$  mating-type chromosomes. The limits of the *M. lagerheimii* PR and HD mating-type chromosomes are indicated, together with the positions of their centromeres. Blue and purple genes correspond to genes orthologous to the blue and purple regions in A. Other evolutionary strata are indicated in orange, black, red, and green. (Inset) Magnification of the youngest strata. The gray gene with nonzero  $d_s$  value at the right of the green stratum was located within the collinear PAR in Fig. 2B and was separated from the green genes by a small noncollinear region. The PARs in *M. lychnidis-dioicae* are shown in gray.

formed by the fusion of the complete ancestral PR chromosome with one arm of the ancestral chromosome carrying the HD locus (Fig. 3 D–F and *SI Appendix*, Fig. S5A). The remaining ancestral HD chromosome arm became an autosomal component in *M. lychnidis-dioicae* (Fig. 3 D–F and *SI Appendix*, Fig. S5A). Based on the inferred ancestral state, the two mating-type loci were initially some 800 kb apart following the fusion event and the fused chromosome was about 2 Mb long. The subsequent cessation of recombination between the PR and HD mating-type loci led to extensive rearrangements (*SI Appendix*, Fig. S3) and divergence between alleles associated with the  $a_1$  and  $a_2$  mating types in this region (in black in Fig. 4B and Fig. 3 G–I). Conversely, the chromosome regions ancestrally distal to the PR and HD loci largely constituted PARs that remained collinear (*SI Appendix*, Fig. S1), with a within-individual  $d_s$  of 0 (gray in Fig. 4B), as expected for recombining regions in organisms with high selfing rates.

For confirmation of the complete and ancient nature of recombination suppression in the region ancestrally between the HD and PR loci (black in Fig. 4B), we obtained and analyzed high-quality genome assemblies of *M. silenes-dioicae* and *M. violaceum* s. str., two species closely related to *M. lychnidis-dioicae* (Fig. 1B). All three species had homologous mating-type chromosomes (*SI Appendix*, Fig. S5 B and C), each including the large black nonrecombining region ancestrally located between PR and HD, with a high degree of rearrangement between mating types and between species (*SI Appendix*, Fig. S5). Most of the genes (66%) in the black region displayed *trans*-specific polymorphism across *M. lychnidis-dioicae* and *M. silenes-dioicae*, with alleles associated with the  $a_1$  mating type of both species branching together rather than clustering with the allele associated with the  $a_2$  mating type from the same species. *Trans*-specific polymorphism also extended to the divergence from *M. violaceum* s. str. in 20% of the genes (*SI Appendix*, Figs. S6 and S7), which were dispersed along the ancestral gene order in the black stratum (corresponding to the black points with the highest  $d_s$  values in Fig. 4B). Such old *trans*-specific polymorphism associated with mating type could not have persisted without an ancient and complete cessation of recombination preceding the divergence of these three species, as even very low levels of recombination would have had time to break down the association between alleles at genes in the black stratum and mating-type genes. The less marked or absent *trans*-specific polymorphism for some of the genes within the black stratum probably results from occasional localized gene conversion events, as demonstrated for mating-type chromosomes in *Microbotryum* (35) and other fungi (36, 37). *Trans*-specific polymorphism in the genealogies of genes in the black stratum never extended to *M. lagerheimii*, consistent with the lack of linkage between the PR and HD mating-type loci in this species. When calibrated by the previous estimate for the date of the speciation event between *M. lychnidis-dioicae* and *M. silenes-dioicae* to 420,000 years ago (38), the divergence between  $a_1$ - and  $a_2$ -associated alleles in the black region led to an estimate for the date of linkage between the HD and PR loci of about 1.3 million years ago (95% confidence interval between 1.1 and 1.6 million years) (*SI Appendix*, *SI Text* and Table S1).

#### Evolutionary Strata Extending Suppressed Recombination Around Each of the PR and HD Mating-Type Loci.

We detected further independent evolutionary strata, extending recombination suppression to nonmating-type genes in several successive steps (Fig. 3 C and D and *SI Appendix*, Table S1). Three strata (blue, purple, and orange strata, Figs. 1, 3, and 4), devoid of mating-type genes, evolved independently after the recombination suppression events at each of the PR and HD mating-type loci but before the HD and PR linkage via chromosome fusion (black stratum, Figs. 1, 3, and 4). Indeed, *M. lagerheimii*, which has unlinked HD and PR mating-type loci and therefore no black stratum, showed high divergence levels between alleles associated with the  $a_1$  mating type and those associated with the  $a_2$  mating type for 14 HD-proximal (blue) genes and 31 PR-proximal (purple) genes, none of which are involved in mating-type determination (Fig. 4A and *SI Appendix*, Tables S1, S2, and S3). These findings are in stark contrast with the near-zero within-individual  $d_s$  values along the rest of the mating-type chromosomes in this species. These same genes ancestrally located near the PR and HD loci also had the next-highest level of divergence after the mating-type genes themselves in *M. lychnidis-dioicae*, *M. silenes-dioicae*, and *M. violaceum* s. str. (purple and blue in Fig. 4B; *SI Appendix*, Tables S1 and S3 and Fig. S8). Most of these genes also displayed high levels of synonymous divergence between mating types in three other *Microbotryomycetes* (*SI Appendix*, Table S2 and figure 2B in ref. 39) and were also localized within noncollinear regions in these distant outgroups (figure 1A in ref. 39). These results

suggest that recombination suppression linking nonmating-type genes to key mating-type determinants is very old and occurred before these Microbotryomycetes species diverged (Fig. 1B).

Further supporting this inference, all genes in the purple stratum displayed *trans*-specific polymorphism with 50% displaying *trans*-specific polymorphism extending up to the divergence of *M. lychnidis-dioicae* and *M. lagerheimii* and 12.5% back to the root of the Microbotryomycetes (SI Appendix, Figs. S6 and S7). This confirms that a set of genes has been linked to the mating-type genes since long before the black stratum evolved, at least before the split between the common ancestor of *Microbotryum* and these outgroups. Indeed, such *trans*-specific polymorphisms separating alleles associated with the  $a_1$  and  $a_2$  mating types across species cannot be maintained without a full recombination suppression preceding all those speciation events. Additional strong evidence for an ancient and complete recombination cessation in the purple stratum is provided by the rearrangements between mating types present in *M. lagerheimii*, with the fully linked PR and pheromone genes of the  $a_1$  genome located at opposite edges of the purple region (SI Appendix, Fig. S4). Recombination was suppressed in the purple stratum about 2.1 million years ago (95% confidence interval between 1.6 and 2.7 million years), as estimated based on the divergence between alleles associated with the alternative mating types (SI Appendix, SI Text and Table S1). This date confirms that the suppression of recombination predates the black stratum. We also found *trans*-specific polymorphism in the blue stratum in 75% of the genes (SI Appendix, Fig. S7). Recombination suppression in the purple and blue nonmating-type genes was more recent than that between the genes at each of the HD and PR mating-type loci, as shown by the lower  $d_S$  than for the mating-type genes and by the ancestral situation in basidiomycetes, with only mating-type genes linked at each of the PR and HD loci. Furthermore, the estimated dates of recombination suppression and *trans*-specific polymorphism in the purple stratum were much more recent than those for the PR gene (SI Appendix, Table S1 and Fig. S7) (33). The blue stratum may be more recent than the purple stratum, as suggested by its lower  $d_S$  values and less deep *trans*-specific polymorphism. However, even if equally ancient, the blue and purple strata constitute independent events of recombination suppression, given that they evolved in separate chromosomes, as seen in *M. lagerheimii* (Figs. 3 and 4).

We found another independent stratum ancestrally located distal to the purple stratum, toward the PARs (orange in Fig. 4B), with high  $d_S$  values in *M. lychnidis-dioicae*, *M. silenens-dioicae*, and *M. violaceum s. str.* (Fig. 4 and SI Appendix, Fig. S8). Genes in the orange stratum were not involved in mating-type determination and also displayed deep *trans*-specific polymorphism, indicating an ancient complete suppression of recombination (SI Appendix, Figs. S6 and S7). The orange stratum most likely evolved before the black stratum linking the HD and PR mating-type genes (Fig. 3E), as it showed a much higher mean  $d_S$  value (Figs. 4B and SI Appendix, Table S1). The orange stratum genes had zero  $d_S$  in *M. lagerheimii* (Fig. 4 and SI Appendix, Fig. S8), indicating that recombination suppression in this region evolved after the divergence with this species and therefore after the purple stratum (Fig. 3). Consistent with this inference, the genes of the orange stratum had lower  $d_S$  values than those of the purple stratum.

**Additional, Younger Evolutionary Strata Evolved After the Linkage Between HD and PR Mating-Type Loci in *M. lychnidis-dioicae*.** We found evidence for two additional and younger evolutionary strata in *M. lychnidis-dioicae*. These strata did not involve mating-type genes either and arose much later than the linkage between HD and PR mating-type genes (red and green in Fig. 4B and SI Appendix, Tables S1 and S3). The red stratum contains genes that were ancestrally located distal to the orange stratum

in the PAR and displayed intermediate  $d_S$  values in *M. lychnidis-dioicae* (SI Appendix, Table S1 and Fig. 4B, Inset). These findings indicate a further distal expansion of recombination suppression, after the evolution of the black stratum, in the lineage leading to *M. lychnidis-dioicae* (Fig. 3H and I). The red stratum is currently located between the HD and PR loci in *M. lychnidis-dioicae* (Fig. 2B), where recombination has been shown to be completely suppressed using segregation analyses (10, 24). This additional recombination suppression step occurred before the divergence of *M. lychnidis-dioicae* and *M. silenens-dioicae*, as shown by the nonzero within-individual  $d_S$  values found in *M. silenens-dioicae* (SI Appendix, Fig. S8A) and the *trans*-specific polymorphism observed in almost 30% of the genes across these species (SI Appendix, Fig. S7). In *M. silenens-dioicae*, the red stratum genes have remained in close proximity but are separated from the PAR by a small noncollinear region in *M. silenens-dioicae* (SI Appendix, Fig. S5E). In contrast, the *trans*-specific polymorphism in the red stratum genes never extended to *M. violaceum s. str.* (SI Appendix, Fig. S7). In this later species, the red stratum genes were located in the PAR and had zero  $d_S$  values within the sequenced diploid strain (SI Appendix, Figs. S5B and D and S7), demonstrating the regular occurrence of recombination. This indicates that the red stratum, which contains no mating-type genes, was formed more recently than the black stratum, in the lineage leading to *M. lychnidis-dioicae* and *M. silenens-dioicae*, after its divergence with *M. violaceum s. str.* (Fig. 1). We estimated the date of recombination suppression in the red stratum at about 0.9 millions years ago (95% confidence interval between 0.7 and 1.1 million years), which is much more recent than the black stratum linking the HD and PR loci (SI Appendix, SI Text and Table S1).

We found yet another small, very recent putative evolutionary stratum, also common only to *M. lychnidis-dioicae* and *M. silenens-dioicae* (green in Figs. 2B and 4B). In both species, this green stratum was located close to the PAR, but was separated from it by a region with an inversion between  $a_1$  and  $a_2$  (Fig. 2B and SI Appendix, Fig. S5E), and the genes had low but nonzero  $d_S$  values. These genes probably correspond to a very recent expansion of the nonrecombining region after the formation of the red stratum (Fig. 3I), as suggested by their lower  $d_S$  values (Fig. 4B and SI Appendix, Table S1), distinctive current location in the *M. lychnidis-dioicae* genome, and lack of *trans*-specific polymorphism (SI Appendix, Fig. S7). Alternatively, the red and green strata may have been generated in a single event, followed by physical separation due to rearrangements.

We further confirmed recombination suppression in the red and green strata using multiple available resequenced genomes for *M. lychnidis-dioicae* and *M. silenens-dioicae* (22, 40). Gene genealogies revealed that the alleles associated with the  $a_1$  and  $a_2$  mating types formed two different clades for at least one species in all of the gene genealogies, and *trans*-specific polymorphism was observed across all individuals in 23% of cases (SI Appendix, Fig. S9). These patterns can be explained only by full linkage to mating type before the two species split. Furthermore, the mean levels of polymorphism per mating type and per species were significantly lower in all of the evolutionary strata than in the PARs (Student's *t* tests,  $P < 0.05$ ) (SI Appendix, Fig. S10), as expected in regions without recombination due to the lower effective population size. These findings indicated that the nonzero  $d_S$  in the evolutionary strata within the sequenced individuals was due to recombination suppression rather than elevated substitution rates. Maximum-likelihood Hudson–Kreitman–Agade tests further supported that the higher  $d_S$  levels in the purple, blue, orange, black, and red strata compared with the PARs in *M. lychnidis-dioicae* and *M. silenens-dioicae* were due to balancing selection (because of linkage to mating type) rather than to elevated mutation rates (SI Appendix, SI Text).



**Conclusion.** We found multiple evolutionary strata in *Microbotryum* mating-type chromosomes, several of which included no genes related to mating-type determination (*SI Appendix, Tables S2 and S3*). Only the two oldest strata (linking pairs of mating-type genes within each of the separate PR and HD loci, respectively) and the black stratum (linking the HD and PR loci together) involved mating-type genes. The linkage between HD and PR genes probably evolved because it increased the likelihood of compatibility between gametes in a selfing system (16) (*SI Appendix, SI Text and Fig. S2*). Mating-type loci linkage has been reported in several fungi (9, 14, 41, 42), and our study provides clues to the chromosomal rearrangements underlying this phenomenon. The five other *M. lychnidis-dioicae* evolutionary strata (purple, blue, orange, red, and green strata) did not involve mating-type genes. As this species has no male/female traits, our results indicate that sexual antagonism is insufficient to account for the occurrence of these evolutionary strata. Furthermore, given the *M. lychnidis-dioicae* life cycle, the existence of ecological differences between mating types that would enhance alternative mating-type functions is highly unlikely. Alternative hypotheses should therefore be considered, including the capture and shelter of deleterious alleles in a permanently

heterozygous state, the fixation of neutral rearrangements by drift in one gametolog (see *SI Appendix, SI Text and Fig. S1* for more details concerning these hypotheses). These general mechanisms apply to both sex and mating-type chromosomes across the plant, animal, and fungal kingdoms. Our results thus call for an expansion of theories concerning the evolution of eukaryotic sex-related chromosomes to forces beyond sexual antagonism.

## Materials and Methods

Haploid genomes were sequenced with P6/C4 Pacific Biosciences SMRT technology (University of California, San Diego IGM Genomics Facility) (*SI Appendix, Table S4*). Orthologs were identified by orthomcl (43). Trees were inferred with RAxML (44). Strata divergence times were inferred with BEAST2 (45). More details about the materials and methods used are provided in *SI Appendix, SI Text*.

**ACKNOWLEDGMENTS.** We thank Stéphanie Le Prieur, Alodie Snirc, Cécile Fairhead, and Gilles Deparis for help with DNA extraction. This work was supported by European Research Council GenomeFun Grant 309403 and a grant from Institut Diversité Ecologie et Evolution du Vivant (to T.G.); NSF Grant DEB-1115765 and NIH Grant R15GM119092 (to M.E.H.); the Marie Curie European Grant 701646 (to S.B.); and a postdoctoral fellowship (SFRH/BPD/79198/2011) from Fundação para a Ciência e a Tecnologia, Portugal (to M.A.C.).

- Bergero R, Charlesworth D (2009) The evolution of restricted recombination in sex chromosomes. *Trends Ecol Evol* 24:94–102.
- Charlesworth D (2015) The status of supergenes in the 21st century: Recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations. *Evol Appl* 9:74–90.
- Wright AE, Dean R, Zimmer F, Mank JE (2016) How to make a sex chromosome. *Nat Commun* 7:12087.
- Charlesworth B (1991) The evolution of sex chromosomes. *Science* 251:1030–1033.
- Ironsides JE (2010) No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *BioEssays* 32:718–726.
- Charlesworth B, Wall JD (1999) Inbreeding, heterozygote advantage and the evolution of neo-X and neo-Y sex chromosomes. *Proc Biol Sci* 266:51–56.
- Antonovics J, Abrams JY (2004) Intratetrad mating and the evolution of linkage relationships. *Evolution* 58:702–709.
- Johnson LJ, Antonovics J, Hood ME (2005) The evolution of intratetrad mating rates. *Evolution* 59:2525–2532.
- Menkis A, Jacobson DJ, Gustafsson T, Johannesson H (2008) The mating-type chromosome in the filamentous ascomycete *Neurospora tetrasperma* represents a model for early evolution of sex chromosomes. *PLoS Genet* 4:e1000030.
- Hood ME (2002) Dimorphic mating-type chromosomes in the fungus *Microbotryum violaceum*. *Genetics* 160:457–461.
- Kües U (2000) Life history and developmental processes in the basidiomycete *Coprinus cinereus*. *Microbiol Mol Biol Rev* 64:316–353.
- Coelho MA, Bakkeren G, Sun S, Hood ME, Giraud T (2017) Fungal sex: The basidiomycota. *Microbiol Spectr*, 10.1128/microbiolspec.FUNK-0046-2016.
- Kües U (2015) From two to many: Multiple mating types in Basidiomycetes. *Fungal Biol Rev* 29:126–166.
- Bakkeren G, Kronstad JW (1994) Linkage of mating-type loci distinguishes bipolar from tetrapolar mating in basidiomycetous smut fungi. *Proc Natl Acad Sci USA* 91:7085–7089.
- Idnurm A, Hood ME, Johannesson H, Giraud T (2015) Contrasted patterns in mating-type chromosomes in fungi: Hotspots versus coldspots of recombination. *Fungal Biol Rev* 29:220–229.
- Nieuwenhuis BPS, et al. (2013) Evolution of uni- and bifactorial sexual compatibility systems in fungi. *Heredity (Edinb)* 111:445–455.
- Badouin H, et al. (2015) Chaos of rearrangements in the mating-type chromosomes of the anther-smut fungus *Microbotryum lychnidis-dioicae*. *Genetics* 200:1275–1284.
- Fraser JA, et al. (2004) Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms. *PLoS Biol* 2:e384.
- James TY (2015) Why mushrooms have evolved to be so promiscuous: Insights from evolutionary and ecological patterns. *Fungal Biol Rev* 29:167–178.
- Giraud T, Yockteng R, López-Villavicencio M, Refrégier G, Hood ME (2008) Mating system of the anther smut fungus *Microbotryum violaceum*: Selfing under heterothallism. *Eukaryot Cell* 7:765–775.
- Votintseva AA, Filatov DA (2009) Evolutionary strata in a small mating-type-specific region of the smut fungus *Microbotryum violaceum*. *Genetics* 182:1391–1396.
- Whittle CA, Votintseva A, Ridout K, Filatov DA (2015) Recent and massive expansion of the mating-type-specific region in the smut fungus *Microbotryum*. *Genetics* 199:809–816.
- Billiard S, et al. (2011) Having sex, yes, but with whom? Inferences from fungi on the evolution of anisogamy and mating types. *Biol Rev Camb Philos Soc* 86:421–442.
- Hood ME, Antonovics J, Koskella B (2004) Shared forces of sex chromosome evolution in haploid-mating and diploid-mating organisms: *Microbotryum violaceum* and other model organisms. *Genetics* 168:141–146.
- Hood ME, Petit E, Giraud T (2013) Extensive divergence between mating-type chromosomes of the anther-smut fungus. *Genetics* 193:309–315.
- Hood ME, Antonovics J (2000) Intratetrad mating, heterozygosity, and the maintenance of deleterious alleles in *Microbotryum violaceum* (= *Ustilago violacea*). *Heredity (Edinb)* 85:231–241.
- Hood ME, Antonovics J (2004) Mating within the meiotic tetrad and the maintenance of genomic heterozygosity. *Genetics* 166:1751–1759.
- Giraud T, Jonot O, Shykoff JA (2005) Selfing propensity under choice conditions in a parasitic fungus, *Microbotryum violaceum*, and parameters influencing infection success in artificial inoculations. *Int J Plant Sci* 166:649–657.
- Montgomery T (1911) Are particular chromosomes sex determinants? *Biol Bull* 19:1–17.
- Montgomery TH, Jr (1906) The terminology of aberrant chromosomes and their behavior in certain Hemiptera. *Science* 23:36–38.
- Kemler M, Göker M, Oberwinkler F, Begerow D (2006) Implications of molecular characters for the phylogeny of the Microbotryaceae (Basidiomycota: Urediniomycetes). *BMC Evol Biol* 6:35.
- Hood ME, Scott M, Hwang M (2015) Breaking linkage between mating compatibility factors: Tetrapolarity in *Microbotryum*. *Evolution* 69:2561–2572.
- Devier B, Aguilera G, Hood ME, Giraud T (2009) Ancient trans-specific polymorphism at pheromone receptor genes in basidiomycetes. *Genetics* 181:209–223.
- Petit E, et al. (2012) Linkage to the mating-type locus across the genus *Microbotryum*: Insights into nonrecombining chromosomes. *Evolution* 66:3519–3533.
- Fontanillas E, et al. (2015) Degeneration of the nonrecombining regions in the mating-type chromosomes of the anther-smut fungi. *Mol Biol Evol* 32:928–943.
- Menkis A, Whittle CA, Johannesson H (2010) Gene genealogies indicates abundant gene conversions and independent evolutionary histories of the mating-type chromosomes in the evolutionary history of *Neurospora tetrasperma*. *BMC Evol Biol* 10:234.
- Sun S, Hsueh Y-P, Heitman J (2012) Gene conversion occurs within the mating-type locus of *Cryptococcus neoformans* during sexual reproduction. *PLoS Genet* 8:e1002810.
- Gladieux P, et al. (2011) Maintenance of fungal pathogen species that are specialized to different hosts: Allopatric divergence and introgression through secondary contact. *Mol Biol Evol* 28:459–471.
- Maia TM, et al. (2015) Evolution of mating systems in Basidiomycetes and the genetic architecture underlying mating-type determination in the yeast *Leucosporidium scottii*. *Genetics* 201:75–89.
- Badouin H, et al. (2017) Widespread selective sweeps throughout the genome of model plant pathogenic fungi and identification of effector candidates. *Mol Ecol* 26:2041–2062.
- Fraser JA, Heitman J (2005) Chromosomal sex-determining regions in animals, plants and fungi. *Curr Opin Genet Dev* 15:645–651.
- Grognet P, et al. (2014) Maintaining two mating types: Structure of the mating type locus and its role in heterokaryosis in *Podospira anserina*. *Genetics* 197:421–432.
- Li L, Stoeckert CJ, Jr, Roos DS (2003) OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res* 13:2178–2189.
- Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
- Bouckaert R, et al. (2014) BEAST 2: A software platform for Bayesian evolutionary analysis. *PLOS Comput Biol* 10:e1003537.



# SI Appendix

## Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism

Sara Branco<sup>1,a</sup>, H el ene Badouin<sup>1,a,b</sup>, Ricardo C. Rodr iguez de la Vega<sup>1,a</sup>, J er ome Gouzy<sup>b</sup>, Fantin Carpentier<sup>a</sup>, Gabriela Aguilera<sup>a</sup>, Sophie Siguenza<sup>b</sup>, Jean-Tristan Brandenburg<sup>a</sup>, Marco A. Coelho<sup>c</sup>, Michael E. Hood<sup>2,d</sup>, Tatiana Giraud<sup>2,a,\*</sup>

<sup>1</sup> These authors contributed equally to the study

<sup>2</sup> These authors jointly supervised the study

<sup>a</sup> Ecologie Syst ematique Evolution, B atiment 360, Univ. Paris-Sud, AgroParisTech, CNRS, Universit e Paris-Saclay, 91400 Orsay, France.

<sup>b</sup> LIPM, Universit e de Toulouse, INRA, CNRS, 31326 Castanet-Tolosan, France.

<sup>c</sup> UCIBIO-REQUIMTE, Departamento de Ci ncias da Vida, Faculdade de Ci ncias e Tecnologia, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

<sup>d</sup> Department of Biology, Amherst College, 01002-5000 Amherst, Massachusetts, United State of America.

\* Corresponding author: Tatiana Giraud (tatiana.giraud@u-psud.fr)

### SI Appendix Includes:

SI Text

SI Materials and Methods

SI References

SI Figures S1 – S10

SI Tables S1 – S7

## SI Text

### Mating-type chromosomes in fungi

There are several hypotheses for the evolution of recombination suppression in fungal mating-type chromosomes, some also holding for animal, plant and algal sex chromosomes. These include:

1) Linkage of two mating-type loci: Most basidiomycete fungi (i.e., mushrooms, rusts and smut fungi) contain two loci determining pre- and post-mating compatibility, respectively. The PR locus controls gamete fusion, with pheromone receptor and pheromone precursor genes (the latter sometimes being present in multiple copies), and the HD locus controls compatibility during growth of after fusion of haploids cells, with two adjacent homeodomain genes (1, 2, 3). Both PR and HD mating-type loci in basidiomycetes encompass two to several genes, and recombination is suppressed within each of these two loci, in all species, ensuring proper mating-type determination (3, 4). Linkage of these two mating-type loci is beneficial (but not required) under selfing mating systems in basidiomycete fungi as it increases the probability of compatibility among gametes from the same diploid parent (5-7) (Fig. S2). In contrast, the segregation of the two mating-type loci in association with multi-allelism increases discrimination against selfing and mating among closely related individuals, which promotes outcrossing (5-9). The PR and HD loci segregate independently in most basidiomycetes, but these loci are linked in some species (6, 10-13). Independent meiotic segregation of the two mating-type loci in basidiomycetes is called tetrapolarity (because meiotic segregation can result in four possible mating types among gametes of a diploid parent, Fig. S2A), while linkage of the mating-type loci in basidiomycetes leads to bipolarity (segregation of only two mating types among gametes of a diploid parent, Fig. S2C).

2) A second explanation for suppressed recombination in fungal mating-type chromosomes is the linkage of mating-type genes to the centromere. In ascomycete and basidiomycete automictic species (i.e., undergoing selfing among products of the same meiotic tetrad (11, 14, 15)), such linkage is particularly common and thought to be favorable (16, 17), resulting in mating-type segregation and reformation of the diploid between meiotic products separated at the first meiotic division. In basidiomycetes, linkage of both HD and PR mating-type genes to their respective centromeres increases the odds of compatibility within tetrads (Fig. S2B). Note that in basidiomycete fungi with linked HD and PR genes, linkage to the centromere is not beneficial with regard to the odds of compatibility between gametes, as it will not further increase mating compatibility among gametes in a diploid individual (Fig. S2C).

3) In analogy to the sexually antagonistic selection hypothesis, which is the main theory for explaining evolutionary strata in sex chromosome evolution, there could be “mating-type antagonistic selection” (18) if some phenotypes were beneficial for improving the  $a_1$  function while being detrimental

to the  $a_2$  function, or vice versa. However, there are not many traits that could have different optima in the two mating types.

4) Epistatic interactions can also favor linkage for protecting beneficial combinations of alleles, unrelated to mating type or sex functions. Unlike hypothesis 3 above that also deals with epistasis (19), this fourth hypothesis postulates that beneficial allelic combinations do not improve the  $a_1$  and  $a_2$  mating type functions in themselves, and therefore does not correspond to “mating-type antagonistic selection”. Epistatic interactions can also occur on autosomes but the maintenance of alternative alleles should be facilitated when linked to the mating-type or sex-determining genes, which are under strong balancing selection.

5) Another evolutionary explanation for suppressed recombination on fungal mating-type chromosomes, that could also apply to sex chromosomes, involves linkage of deleterious alleles to the mating-type loci, favoring permanent sheltering in an heterozygous state, as has been theoretically modeled (20, 21). This may arise specifically in sex and mating-type chromosomes if recombination frequency gradually decreases from the non-recombining into the PAR, so that partial linkage (linkage disequilibrium) to mating-type or sex-determining genes makes selection against recessive deleterious alleles less efficient. This would allow deleterious alleles to increase in frequency in these PAR edges at the margin of the non-combining region. Rare recombination events would then generate individuals homozygous for deleterious alleles and could therefore be selected against. Complete linkage of these PAR margins in disequilibrium with mating type may thus be favorable and selected for (Fig. S1A). Permanent sheltering would thus be more easily achieved than purging if recombination is rare at the PAR margins. Further theoretical models are needed to explore the general conditions under which such a mechanism can generate evolutionary strata.

6) A last mechanism explaining suppressed recombination involves neutral inversions or rearrangements extending the region of suppressed recombination (22); these could be fixed by drift in one of the sex or mating-type chromosomes and arrest recombination between sex or mating-type chromosomes. For example, an inversion trapping a part of the PAR at the margin of the non-recombining region could be neutral in the short term, and could be fixed by drift in one of the sex or mating-type chromosomes, and not the other, because of the recombination suppression and because of mating occurring only between mating types (Fig. S1B and C). Such phenomena can occur on any chromosome, but linkage in mating-type and sex chromosomes could be maintained longer in a polymorphic state (and therefore be more detectable), due to balancing selection retaining alternate alleles determining sexes and mating types (22). As for the hypothesis above, theoretical models could be developed for validating whether such a mechanism can generate evolutionary strata.



## The *Microbotryum lychnidis-dioicae* system

Contributing to understanding of these various evolutionary models, the anther-smut fungus *Microbotryum lychnidis-dioicae* has been important for the study of fungal mating-type chromosomes. It was central to the first studies describing bipolar breeding systems in fungi (23) and the first case of size-dimorphic mating-type chromosomes in fungi (24). Suppressed recombination encompasses ca. 90% of the mating-type chromosome lengths (sized 3.5 Mb for  $a_1$  and 4.0 Mb for  $a_2$ ) with 614 and 683 predicted non-transposable element genes, respectively (12, 24-27). The non-recombining region shows chaos of rearrangements and gene losses, transposable element accumulation, non-synonymous substitutions and impaired gene expression (12, 27). The non-recombining region is flanked at both sides by two small recombining and collinear regions, therefore called pseudo-autosomal regions (PARs).

The existence of evolutionary strata has been debated in *M. lychnidis-dioicae*. Varying levels of synonymous divergence between  $a_1$  and  $a_2$  alleles in a handful of genes in the unassembled non-recombining region led to the claim that evolutionary strata existed (28). Based on highly fragmented assemblies, it was then speculated that an oldest stratum evolved for linking pheromone precursor and pheromone receptor genes, a second small (60 kb) old stratum for linking the HD and PR locus, and a third recent and large stratum for linking the MAT genes to the centromeres (29). However, the full assemblies of the mating-type chromosomes recently showed that the HD and PR locus are 600 kb apart, that no clear discrete strata patterns were discernable using the current gene orders, as genes with different  $a_1$ - $a_2$  synonymous divergence levels (a proxy for time since recombination suppression) were dispersed chaotically along the chromosomes (12). These observations could be explained by a lack of evolutionary strata, extensive rearrangements of original strata, and/or gene conversion, which is known to occur in fungal mating-type chromosomes (30, 31). A recent study suggested the existence of five evolutionary strata along the mating-type chromosomes in *M. lychnidis-dioicae* using a clustering method based on putative oligonucleotide compositional differences between strata (32); however, this method did not seem to yield reliable inference because it could not retrieve the pseudo-autosomal boundary in one mating-type chromosome. Furthermore, each inferred stratum included genes with highly heterogeneous levels of divergence (12). Our study shows that the recovered putative strata in this previous study (32) in fact did not correspond to genuine evolutionary strata.

Elucidating whether evolutionary strata exist and which genomic regions they include is important for understanding how and why suppressed recombination evolved. The different mechanisms for explaining suppression of recombination outlined above do not all predict the existence of evolutionary strata. For example, linking the two mating-type loci would not generate multiple strata, while the successive linkage of deleterious alleles to the mating-type locus would (20). In this regard, an interesting feature of *M. lychnidis-dioicae* is that frequent deleterious alleles linked to mating types have been reported, that prevent haploid growth (14, 33, 34). In fact, many genes have been lost from

one or the other mating type (12) and transposable elements and non-synonymous substitutions have accumulated in non-recombining regions (12, 27).

### Genome assemblies and comparisons

We used Pacific Bioscience sequencing technologies to generate very high quality genome assemblies for the mating-type chromosomes of several *Microbotryum* species. The  $a_1$  and  $a_2$  haploid genomes of *M. lagerheimii* and one haploid genome of *M. intermedium* were well assembled, with 42 and 37 contigs for  $a_1$  and  $a_2$  *M. lagerheimii*, respectively, 24 contigs for *M. intermedium*, for an estimated number of chromosomes of 12 for haploid *Microbotryum* genomes (24). The mating-type chromosomes were well assembled; the HD chromosomes of *M. intermedium* and *M. lagerheimii* appeared assembled in a single contig each, and were homologous and syntenic between the species; the complete PR chromosome of *M. intermedium* matched two contigs in the  $a_1$  *M. lagerheimii* assembly, with a small inversion internal to one contig (the MC16 contig, in dark purple in Figure 2A; the inversion is represented by the orange links).

In addition to collinearity with *M. intermedium*, collinearity between the  $a_1$  and  $a_2$  mating-type chromosomes within *M. lagerheimii* (Fig. S4) indicates that recombination is frequent enough to conserve gene order along most of their lengths. In *M. lychnidis-dioicae* the recombining pseudo-autosomal regions of the mating-type chromosomes and autosomes are also highly collinear between  $a_1$  and  $a_2$  genomes, while the non-recombining regions of the mating-type chromosomes are thoroughly rearranged (Fig. S3).

The two distinct contigs carrying the PR and HD genes in *M. lagerheimii* (Fig. S4) are consistent with the species being tetrapolar, as inferred earlier based on meiotic segregation analysis of recombination between these mating-type loci (35). The high degree of collinearity between the mating-type chromosomes of *M. lagerheimii* and *M. intermedium*, and the finding of PR and HD genes in different contigs, each harbouring centromere-specific repeats (12) (Fig. 2A), indicate that *M. intermedium* is also tetrapolar.

We included comparisons of the mating-type chromosomes of *M. lychnidis-dioicae* and *M. violaceum s. str.* (Fig. S5B), whose haploid genomes were also well assembled (Table S4). Three contigs in the  $a_1$  *M. violaceum sens. str.* genome matched the  $a_1$  mating-type chromosomes of *M. lychnidis-dioicae*. In addition, these three contigs had some fragments joined together in the assembly of the *M. violaceum sens. str.*  $a_2$  genome (Fig. S5D), indicating that all three belonged to the mating-type chromosome. The non-recombining region was highly rearranged between *M. lychnidis-dioicae* and *M. violaceum s. str.* (Fig. S5B) in all four possible comparisons, as well as between  $a_1$  and  $a_2$  in both species (Fig. S5D).

We also compared the mating-type chromosomes of the sister species *M. lychnidis-dioicae* and *M. silenes-dioicae* (Fig. S5C), the latter being assembled in more contigs than other species (Table S4;

6 contigs for both  $a_1$  and  $a_2$  mating-type chromosomes). These different contigs had some fragments joined together in the assembly of the alternative mating type (Fig. S5E) and in *M. lychnidis-dioicae* (Fig. S5C), allowing assigning to mating-type chromosomes. The non-recombining region appeared homologous but already highly rearranged between *M. lychnidis-dioicae* and *M. silenes-dioicae* (Fig. S5C) in all four possible comparisons, as well as between  $a_1$  and  $a_2$  in *M. silenes-dioicae* (Fig. S5E).



## SI Materials and Methods

### DNA extraction and sequencing

DNA was extracted using the Qiagen Kit 10243 (Courtaboeuf, France) following manufacturer instructions and using a Carver hydraulic press (reference 3968, Wabash, IN, USA) for breaking cell walls. A haploid genome of *M. intermedium* (strain 1389-BM-12-12, collected on *Salvia pratensis*, Italy, Coord. GPS : 44.33353 & 7.13637), as well as the genomes of  $a_1$  and  $a_2$  haploid cells of *M. lagerheimii* (strain 1253, collected on *S. vulgaris*, near Chambéry, France, Coord. GPS : 45.4 & 6.11), *M. violaceum s. str.* (strain 1249, collected on *S. nutans*, Guarda, Switzerland, Coord. GPS : 46.777 & 10.16) and *M. silenes-dioicae* (strain 1303, collected on *S. dioica*, Bois Carre, Lac de Puy Vachier 45.026485 & 6.276612) were sequenced using the P6/C4 Pacific Biosciences SMRT technology (UCSD IGM Genomics Facility La Jolla, CA, USA) (Table S4).

### Assembly and annotation

We assembled the haploid genomes of  $a_1$  and  $a_2$  *M. lagerheimii*, *M. violaceum s. str.*, *M. silenes-dioicae*, and  $a_1$  *M. intermedium*, and re-assembled the reference genome of *M. lychnidis-dioicae* Lamole (12) into separated  $a_1$  and  $a_2$  haploid genomes. Assemblies of the genomes were generated with the wgs-8.2 version of the PBcR assembler (36) with the following parameters: genomeSize=30000000, assembleCoverage=50. Assemblies were polished with quiver software (<https://github.com/PacificBiosciences/GenomicConsensus>). A summary of raw data and assembly statistics is reported in Table S4.

The protein-coding gene models were predicted with EuGene (37), trained for *Microbotryum*. Similarities to the fungi subset of the uniprot database (38) plus the *M. lychnidis-dioicae* Lamole proteome (12) were integrated into EuGene for the prediction of gene models.

Mating-type chromosomes were identified by: 1) finding the contigs carrying the PR or HD mating-type genes in both  $a_1$  and  $a_2$  haploid genomes, 2) using BLAST to search the  $a_1$  against the  $a_2$  haploid genomes and visualizing the output using Circos (39) for identifying contigs with lack of collinearity, 3) blasting the haploid genomes against the completely assembled mating-type chromosomes of *M. lychnidis-dioicae* (12), 4) blasting the  $a_1$  contigs identified in the first steps to the whole  $a_2$  haploid genome, and reciprocally, to identify additional mating-type contigs, 5) repeating this process from step 3 until no additional contig was identified. These contigs were then orientated by: 1) using the centromere-specific repeats (12), as initial assemblies often yielded chromosome arms broken at their centromeres, with identifiable centromere-specific repeats on each separated contig (e.g., Fig. 2A); and 2) blasting the  $a_1$  and  $a_2$  mating-type contigs against each other for identifying the PARs as the collinear regions, that were then assumed to be at the edges of the chromosomes. When the mating-type

chromosomes were fragmented and internal contigs did not include centromeric repeats at one of its edges, it was impossible to recover contig orientation (as indicated in the figure captions).

### **Figures and statistical tests**

The Figures 2, S3, S4 and S5 were prepared using Circos (39). We analyzed gene order to identify larger blocks of synteny. Alleles were assigned between the two mating-type chromosomes and between species by applying orthomcl (40) to the protein data sets. The Student t test was performed using JMP v7 (SAS Institute).

### **Species tree**

We compared the translated gene models of five *Microbotryum* species and the red yeast *Rhodotorula babjevae* with blastp+ 2.30, and the output was used to obtain orthologous groups by Markov clustering (41) as implemented in OrthoMCL v1.4 (40). We aligned the protein sequences of 4,000 single-copy autosomal genes with MUSCLE v3.8.31 (42), and obtained the codon-based CDS alignments with TranslatorX (43). We used RAxML 8.2.7 (44) to obtain maximum likelihood gene trees for all 4,000 fully conserved single-copy genes and a species tree with the concatenated alignment of 2,100,301 codons with no gaps (trimal -nogaps option (45)) under the GTRGAMMA substitution model. We estimated the branch support values by bootstrapping the species tree based on the concatenated alignment and by estimating the relative internode and tree certainty scores based on the frequency of conflicting bipartitions for each branch in the species tree among the fully conserved single-copy genes (46).

### **Gene genealogies**

Gene genealogies were inferred for codon-based alignments of genes in the different strata using RAxML (44) version 8.2.7, assuming the GTRGAMMA model and rapid bootstrap (options: -f a and -# 100). We analyzed for trans-specific polymorphism levels all single-copy genes for which we had both alleles and all species.

### **Date estimates for recombination suppression**

In order to estimate dates of recombination suppression, we leveraged the codon-based alignments of single-copy orthologous groups yielding gene trees displaying trans-specific polymorphism in each evolutionary stratum, as the divergence between alleles associated to the  $a_1$  versus  $a_2$  mating types then corresponds to the time since recombination suppression. We could use three gene families in the purple stratum (1,744 aligned codons), 22 gene families in the black stratum (17,776 aligned codons) and four

gene families in the red stratum (2,726 aligned codons), with alignments including the five *Microbotryum* species and the red yeast used as outgroup, and gene genealogies showing trans-specific polymorphism. There was not enough single-copy gene families showing trans-specific polymorphism for reliable date estimate in the other evolutionary strata. Divergence times were estimated using BEAST v2.4.0 (47), the xml inputs being generated using BEAUTi, and setting the following parameters (others left as default values): unlinked substitution (HKY+G with empirical frequencies for each codon position) and clock models, Yule process to model speciation, and 10,000,000 mcmc generations sampled every 1,000. For all runs we used a single calibration prior at 0.42 million years, corresponding to the divergence between *M. lychnidis-dioicae* and *M. silenes-dioicae* (48), with a normal distribution and a sigma of 0.05.

### **Transposable element identification**

Repetitive DNA content was analyzed with RepeatMasker (49), using REPBASE v19.11 (50). For plotting  $d_s$  along chromosomes, these repeats were removed and further filtering of repeats was performed by blasting (tBLASTx) and removing those matching at more than five locations in the genome.

### **Polymorphism data and analyses**

To rule out high levels of polymorphism as the cause for the observed high  $d_s$  values in the *Microbotryum* mating-type chromosomes, we assessed the level of polymorphism across *M. silenes-dioicae* and *M. lychnidis-dioicae* by using multiple available genomes (Illumina paired-end 60x-100x), from individuals collected across Europe (29, 51). We only used genomes including a single mating-type chromosome ( $a_1$  or  $a_2$ ) either because a haploid strain was sequenced or because a lethal allele was linked to one mating type in a diploid strain so that only haploid cells carrying the alternative mating-type chromosome could be cultivated *in vitro* and sequenced (34, 51, 52). We compiled sequences for a total of ten  $a_1$  and fourteen  $a_2$  mating-type chromosomes of *M. lychnidis-dioicae*, and six  $a_1$  and eight  $a_2$  mating-type chromosomes of *M. silenes-dioicae* (Table S5).

We used two approaches to determine which mating-type chromosome(s) were present in genomes sequences ( $a_1$  or  $a_2$ ), by checking the presence/absence of the  $a_1$  and  $a_2$  pheromone receptor alleles. First, each genome was assembled *de novo* using SOAPdenovo (53) with default parameters, and searched using BLAST for the sequences of the pheromone receptor alleles of *M. lychnidis-dioicae* and *M. silenes-dioicae*, which are highly differentiated between  $a_1$  and  $a_2$  (54). In a second validation step, we counted the reads mapped against the two alleles of the mating-type pheromone receptors. Strains that displayed reads mapping against both alleles were excluded from further analyses.

We performed SNP calling as described previously (51). Reads were mapped with GLINT (T. Faraut & E. Courcelle, available at <http://lipm-bioinfo.toulouse.inra.fr/download/glint/>) with parameters

set as follows: minimum length of the high-scoring pair hsp  $\geq 90$ , with  $\leq 5$  mismatches, no gap allowed, only best-scoring hits taken into account. For each strain, we used as a reference for mapping the haploid PacBio genome assembly corresponding to its species and mating-type. Between 92% and 97% of reads were mapped as proper pairs, and the mapping coverage ranged from 50 to 171 (Table S5). SNPs were called with VarScan (55) (min-coverage 10, min-reads 5, min-avg-qual 30, min-var-freq 0.3, p-value 0.01), and filtered to remove SNPs located near indels, strand-biased SNPs and heterozygote SNPs.

To compute statistics of diversity, we generated pseudo-alignments of CDS in FASTA format from the VCF file produced by VarScan as previously (51). Pseudo-alignments are obtained by substituting reference nucleotides by their variants in the reference sequence. We then computed the  $\theta_\pi$  statistic of diversity with EggLib version 2 (56) for each species and mating-type.

We generated alignments combining the *M. lychnidis-dioicae* and *M. silenes-dioicae* sequences generated above, as well as sequences of the other species. Codon-based alignments of single-copy genes in mating-type evolutionary strata and autosomes were obtained with MUSCLE and translatorX (see above). We used a custom python script to incorporate the sequences of the multiple genomes of *M. lychnidis-dioicae* and *M. silenes-dioicae* while taking into account gaps in the translatorX alignments. Trees were computed with the RAxML rapid bootstrap mode (-f a -m GTRGAMMA -x # 100). To help visualization, trees were rooted using *R. babjevae* and *M. intermedium*, and strains were colored by mating-type with the ETE3 python toolkit (57).

For further checking that the observed differences in diversity levels between the pseudo-autosomal regions and the different strata in *M. lychnidis-dioicae* and *M. silenes-dioicae* (SI Appendix, Table S6) were not due to differences in mutation rates, we performed maximum-likelihood Hudson-Kreitman-Agade (MLHKA) tests as implemented in the MLHKA software (58) using the multiple genomes described above. The MLHKA test is designed to detect signatures of departure from neutrality (either balancing or positive selection) while controlling for mutation rate by the use of divergence data, and it allows multilocus comparisons (58). MLHKA tests can thus be used to test whether differences in diversity levels between two sets of genes is most likely due to elevated mutation rates or to balancing selection (in our case due to linkage of genes in evolutionary strata to mating-type genes), by comparing a model where all loci in the dataset evolve neutrally to an alternative model where  $N$  loci depart from neutrality (59, 60). Because MLHKA tests require a set of control loci that evolve neutrally, we applied these tests to pairwise comparisons between each stratum and the pseudo-autosomal region. To perform the MLHKA tests, we considered the pooled CDS sequences associated with the  $a_1$  and  $a_2$  mating-types for each stratum in each of *M. lychnidis-dioicae* and *M. silenes-dioicae* and used the genome of *M. violaceum sens. str.* as an outgroup. In order to improve divergence estimates, alignments were re-aligned with MACSE (61). Only synonymous sites were taken into account, and singletons were excluded. Unless stated otherwise, tests were run with 1,000,000 iterations of the Markov chain, and repeated at least three times with different random seed numbers to ensure convergence. We arbitrarily chose a starting value of 10 for divergence time, as this value is of little importance when the number of



iterations is high enough for the Markov chains to converge (58). Synonymous nucleotide diversity ( $\theta_\pi$ ) was computed with EggLib version 2 (56) and used as starting value for theta. As the test assumes independence between loci, we concatenated the alignments of the CDS for each evolutionary stratum. For the largest stratum (the black stratum), we randomly sampled and concatenated CDS for 20 genes, as the high number of genes in this stratum resulted in unreasonable running times. For the two most recent strata (red and green strata), we also ran the tests without concatenating the alignments, with 2,000,000 million of iterations of the Markov chains. This yielded similar results as concatenating the CDS. *Microbotryum lychnidis-dioicae* displays a strong population structure in Europe (51, 62), which can lead to over-estimate the diversity in neutral loci and result in false negative tests. In fact, visual inspection of the alignments showed that most polymorphic sites in the PARs in this species corresponded to fixed differences between genetic clusters. We therefore also ran tests separately on the two genetic clusters for which we had several strains of each mating-type, i.e., the Southern and North-Western clusters (51). We corrected p-values for multiple testing within each species with a Benjamini and Hochberg correction (R function p.adjust). In *M. silenes-dioicae*, tests for all strata were significant (adjusted p-values < 0.05, Table S7A), indicating that balancing selection was more likely than elevated mutation rates for explaining the higher diversity in strata than in the PARs. In *M. lychnidis-dioicae*, tests were also significant for the purple, blue, black and orange strata (adjusted p-values < 0.05, Table S7B-D). The difference in diversity between the red stratum and the PAR was marginally significant when considering the whole data set, and significant when the tests were run separately on two genetic clusters (Table S7B-D). There was no significant difference in diversity between the green stratum and the PAR in *M. lychnidis-dioicae*, suggesting that suppressed recombination is not complete there in this species, or that it is too recent to have allowed an accumulation of differences between mating types.

## SI References

1. Feldbrugge M, Kamper J, Steinberg G, & Kahmann R (2004) Regulation of mating and pathogenic development in *Ustilago maydis*. *Curr Op Microbiol* 7:666-672.
2. Kuees U (2015) From two to many: Multiple mating types in Basidiomycetes. *Fung Biol Rev* 29(3-4):126-166.
3. Casselton LA & Kues U (2007) The origin of multiple mating types in the model mushrooms *Coprinopsis cinerea* and *Schizophyllum commune*. *Sex in Fungi*, eds Heitman J, Kronstad J, Taylor J, & Casselton L (ASM Press, Washington, DC), Vol 88, pp 142 - 147.
4. Kues U (2000) Life history and developmental processes in the basidiomycete *Coprinus cinereus*. *Microbiol Mol Biol Rev* 64(2):316-453.
5. James TY (2015) Why mushrooms have evolved to be so promiscuous: Insights from evolutionary and ecological patterns. *Fung Biol Rev* 29(3-4):167-178.
6. Nieuwenhuis BPS, *et al.* (2013) Evolution of uni- and bifactorial sexual compatibility systems in fungi. *Heredity* 111:445-455.
7. Giraud T, Yockteng R, Lopez-Villavicencio M, Refregier G, & Hood ME (2008) The mating system of the anther smut fungus, *Microbotryum violaceum*: selfing under heterothallism. *Euk Cell* 7:765-775.
8. Heitman J, Sun S, & James TY (2013) Evolution of fungal sexual reproduction. *Mycologia* 105(1):1-27.
9. Kües U, James T, & Heitman J (2011 ) The mycota XIV : Mating type in basidiomycetes: Unipolar, bipolar, and tetrapolar patterns of sexuality. *Evolution of fungi and fungal-like organisms*, eds Pöggeler S & Wöstemeyer J (Springer-Verlag, Berlin, Heidelberg Germany).
10. Bakkeren G & Kronstad JW (1994) Linkage of mating type loci distinguishes bipolar from tetrapolar mating in basidiomycetous smut fungi. *Proc Natl Acad Sci USA* 91:7085-7089.
11. Idnurm A, Hood ME, Johannesson H, & Giraud T (2015) Contrasted patterns in mating-type chromosomes in fungi: Hotspots versus coldspots of recombination. *Fung Biol Rev* 29(3-4):220-229.
12. Badouin H, *et al.* (2015) Chaos of rearrangements in the mating-type chromosomes of the anther-smut fungus *Microbotryum lychnidis-dioicae*. *Genetics* 200(4):1275-1284.
13. Fraser JA, *et al.* (2004) Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms. *PLoS Biol* 2:2243-2255.
14. Hood ME & Antonovics J (2000) Intratetrad mating, heterozygosity, and the maintenance of deleterious alleles in *Microbotryum violaceum* (= *Ustilago violacea*). *Heredity* 85:231-241.
15. Menkis A, Jacobson DJ, Gustafsson T, & Johannesson H (2008) The mating-type chromosome in the filamentous ascomycete *Neurospora tetrasperma* represents a model for early evolution of sex chromosomes. *PLoS Genet* 4(3):e1000030.
16. Zakharov IA (1986) Some principles of the gene localization in eukaryotic chromosomes. Formation of the problem and analysis of nonrandom localization of the mating-type loci in some fungi. *Genetika* 22:2620-2624.
17. Zakharov IA (2005) Intratetrad mating and its genetic and evolutionary consequences. *Russ J Genet* 41:508-519.
18. Abbate JL & Hood ME (2010) Dynamic linkage relationships to the mating-type locus in automictic fungi of the genus *Microbotryum*. *J. Evol. Biol* 23(8):1800-1805.
19. Charlesworth D (2016) The status of supergenes in the 21st century: recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations. *Evol App* 9(1):74-90.

20. Antonovics J & Abrams JY (2004) Intratetrad mating and the evolution of linkage relationships. *Evolution* 58(4):702-709.
21. Johnson LJ, Antonovics J, & Hood ME (2005) The evolution of intratetrad mating rates. *Evolution* 59(12):2525-2532.
22. Ironside JE (2010) No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *Bioessays* 32(8):718-726.
23. Kniep H (1919) Untersuchungen uber den Antherenbrand (*Ustilago violacea* Pers.). Ein Beitrag zum Sexualitatsproblem. *Z Bot* 11:257-284.
24. Hood ME (2002) Dimorphic mating-type chromosomes in the fungus *Microbotryum violaceum*. *Genetics* 160:457-461.
25. Hood ME, Antonovics J, & Koskella B (2004) Shared forces of sex chromosome evolution in haploids and diploids. *Genetics* 168:141-146.
26. Hood ME, Petit E, & Giraud T (2013) Extensive divergence between mating-type chromosomes of the anther-smut fungus. *Genetics* 193(1):309-315.
27. Fontanillas E, *et al.* (2015) Degeneration of the non-recombining regions in the mating type chromosomes of the anther smut fungi. *Mol Biol Evol* 32:928-943.
28. Votintseva AA & Filatov DA (2009) Evolutionary strata in a small mating-type-specific region of the smut fungus *Microbotryum violaceum*. *Genetics* 182(4):1391-1396.
29. Whittle CA, Votintseva A, Ridout K, & Filatov DA (2015) Recent and massive expansion of the mating-type specific region in the smut fungus *Microbotryum*. *Genetics* 199:809-816.
30. Menkis A, Whittle C, & Johannesson H (2010) Gene genealogies indicates abundant gene conversions and independent evolutionary histories of the mating-type chromosomes in the evolutionary history of *Neurospora tetrasperma*. *BMC Evol Biol* 10(1):234.
31. Sun S, Hsueh Y-P, & Heitman J (2012) Gene conversion occurs within the mating-type locus of *Cryptococcus neoformans* during sexual reproduction. *PLoS Genet* 8(7):e1002810.
32. Pandey RS & Azad RK (2016) Deciphering evolutionary strata on plant sex chromosomes and fungal mating-type chromosomes through compositional segmentation. *Plant Mol Biol* 90(4-5):359-373.
33. Hood ME & Antonovics JA (2004) Mating within the meiotic tetrad and the maintenance of genomic heterozygosity. *Genetics* 166:1751-1759.
34. Thomas A, Shykoff J, Jonot O, & Giraud T (2003) Sex-ratio bias in populations of the phytopathogenic fungus *Microbotryum violaceum* from several host species. *Int J Plant Sci* 164(4):641-647.
35. Hood ME, Scott M, & Hwang M (2015) Breaking linkage between mating compatibility factors: Tetrapolarity in *Microbotryum*. *Evolution* 69(10):2561-2572.
36. Koren S, *et al.* (2012) Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nat Biotechnol* 30:693-700.
37. Foissac S, *et al.* (2008) Genome annotation in plants and fungi: EuGene as a model platform. *Current Bioinformatics* 3:87-97.
38. Consortium TU (2011) Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res* 39:D214-D219.
39. Krzywinski MI, *et al.* (2009) Circos: An information aesthetic for comparative genomics. *Genom Res* 19:1639-1645.
40. Li L, Stoeckert CJ, & Roos D (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Research* 13:2178-2189.
41. van Dongen S (2000) Graph clustering by flow simulation. PhD (University of Utrecht, The Netherlands).

42. Edgar R (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32 1792-1797.
43. Abascal F, Zardoya R, & Telford M (2010 ) TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res* 38:W7-13.
44. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690.
45. Capella-Gutierrez S, Silla-Martinez JM, & Gabaldon T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972-1973.
46. Salichos L, Stamatakis A, & Rokas A (2014) Novel Information Theory-Based Measures for Quantifying Incongruence among Phylogenetic Trees. *Molyang Biol Evol* 31(5):1261-1271.
47. Bouckaert R, *et al.* (2014 ) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 10:e1003537.
48. Gladieux P, *et al.* (2011) Maintenance of fungal pathogen species that are specialized to different hosts: allopatric divergence and introgression through secondary contact. *Mol Biol Evol* 28:459-471.
49. Smit A & Green P (RepeatMasker at <http://www.repeatmasker.org/>).
50. Jurka J (1998) Repeats in genomic DNA: mining and meaning. *Curr Opin Struct Biol* 8:333-337.
51. Badouin H, *et al.* (2017) Widespread selective sweeps throughout the genome of model plant pathogenic fungi and identification of effector candidates. *Mol Ecol* 26:2041-2062.
52. Oudemans P, *et al.* (1998) The distribution of mating type bias in natural populations of the anther smut *Ustilago violacea* (= *Microbotryum violaceum*) on *Silene alba* (= *S. latifolia*) in Virginia. *Mycologia* 90:372-381.
53. Luo R, *et al.* (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* 1(1):18.
54. Devier B, Aguilera G, Hood M, & Giraud T (2009) Ancient trans-specific polymorphism at pheromone receptor genes in basidiomycetes. *Genetics* 181:209–223.
55. Koboldt D, *et al.* (2012) VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genom Res* 22:568-576.
56. De Mita S & Siol M (2012) EggLib: processing, analysis and simulation tools for population genetics and genomics. *BMC Genetics* 13(1):27.
57. Huerta-Cepas J, Serra F, & Bork P (2016) ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol Biol Evol* 33(6):1635-1638.
58. Wright SI & Charlesworth B (2004) The HKA test revisited: A maximum-likelihood-ratio test of the standard neutral model. *Genetics* 168(2):1071-1076.
59. Qiu S, Bergero R, & Charlesworth D (2013) Testing for the Footprint of Sexually Antagonistic Polymorphisms in the Pseudoautosomal Region of a Plant Sex Chromosome Pair. *Genetics* 194(3):663-+.
60. Wright SI, *et al.* (2006) Testing for effects of recombination rate on nucleotide diversity in natural populations of *Arabidopsis lyrata*. *Genetics* 174(3):1421-1430.
61. Ranwez V, Harispe S, Delsuc F, & Douzery E (2011) MACSE: Multiple alignment of coding sequences accounting for frameshifts and stop codons. *PLoS ONE* 6:e22594.
62. Vercken E, *et al.* (2010) Glacial refugia in pathogens: European genetic structure of anther smut pathogens on *Silene latifolia* and *S. dioica*. *PloS Path* 6:e1001229.



## List of SI Figures

**Fig. S1.** Hypotheses for the generation of evolutionary strata on sex or mating-type chromosomes.

**Fig. S2.** Odds of compatibility among gametes of a diploid individual in basidiomycete fungi.

**Fig. S3.** Comparison of gene order between *Microbotryum lychnidis-dioicae* chromosomes.

**Fig. S4.** Synteny between  $a_1$  and  $a_2$  mating-type chromosomes of *Microbotryum lagerheimii*.

**Fig. S5.** Comparison of gene order between mating-type chromosomes.

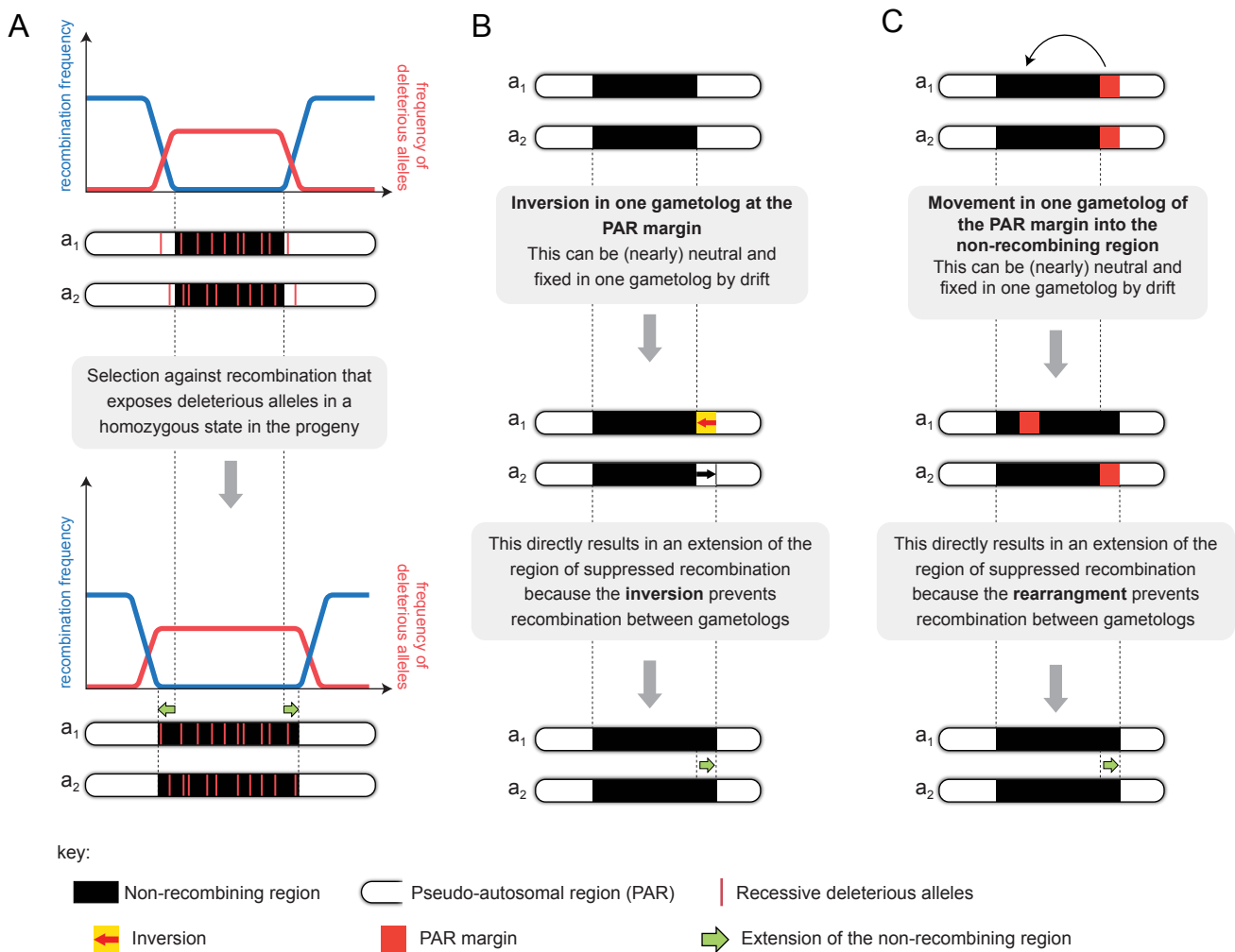
**Fig. S6.** Examples of gene genealogies at non-mating type genes showing different levels of trans-specific polymorphism, i.e., with more or less ancient linkage to mating type.

**Fig. S7.** Cumulative percentage of genes showing trans-specific polymorphism across strata and age of trans-specific polymorphism relative to speciation events.

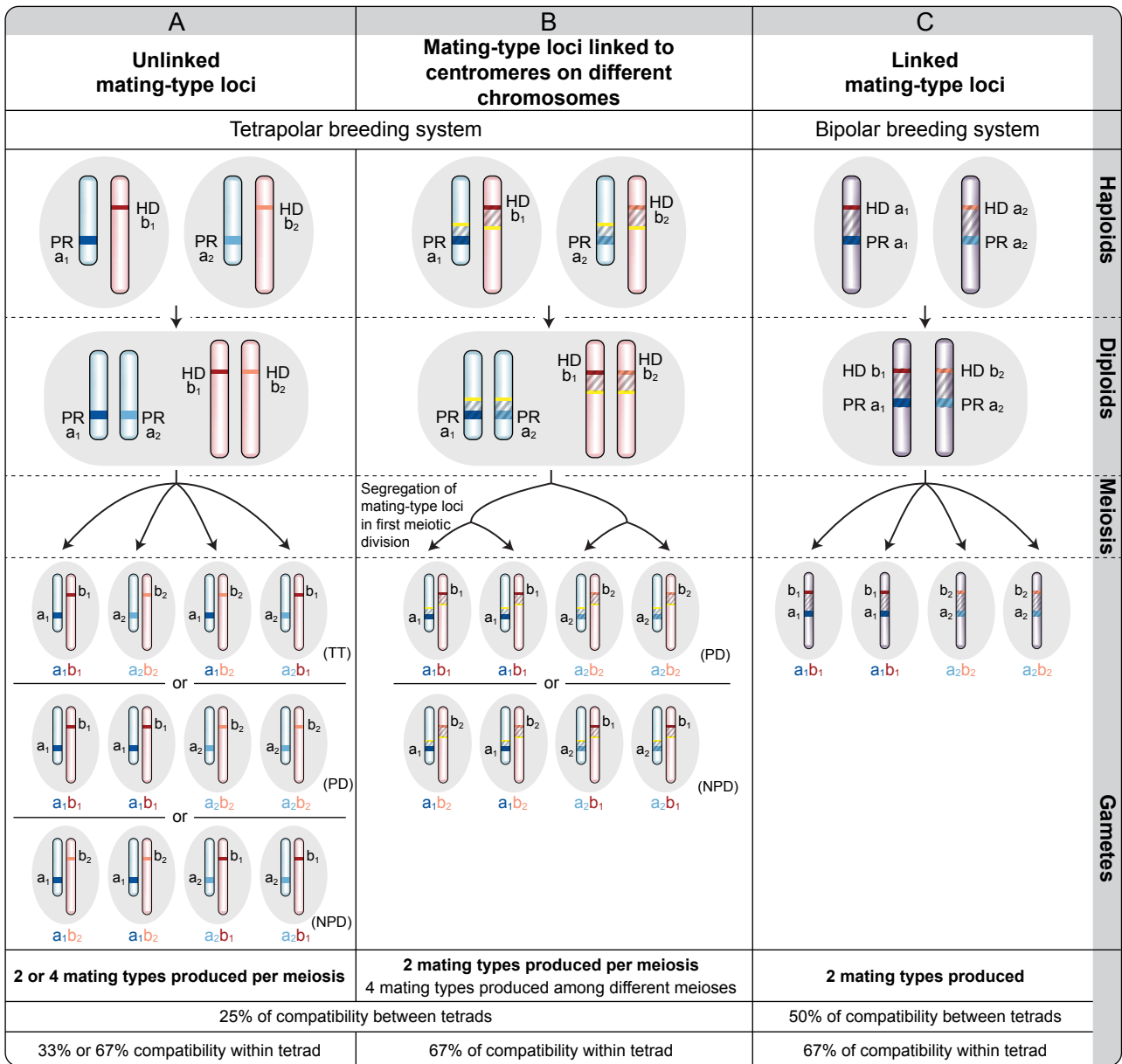
**Fig. S8.** Synonymous divergence between alleles associated with  $a_1$  and  $a_2$  mating types in the sequenced diploid individual in each of *Microbotryum silenes-dioicae* and *M. violaceum s. str.*, plotted along the ancestral gene order.

**Fig. S9.** Patterns of segregation of alleles at non-mating-type genes, associated with the  $a_1$  and  $a_2$  mating types, respectively, in gene genealogies using multiple resequenced genomes of *M. silenes-dioicae* (MvSd) and *M. lychnidis-dioicae* (MvSl).

**Fig. S10.** Genetic diversity in *Microbotryum lychnidis-dioicae* and *M. silenes-dioicae* ( $\theta\pi$ ) based on multiple genomes.

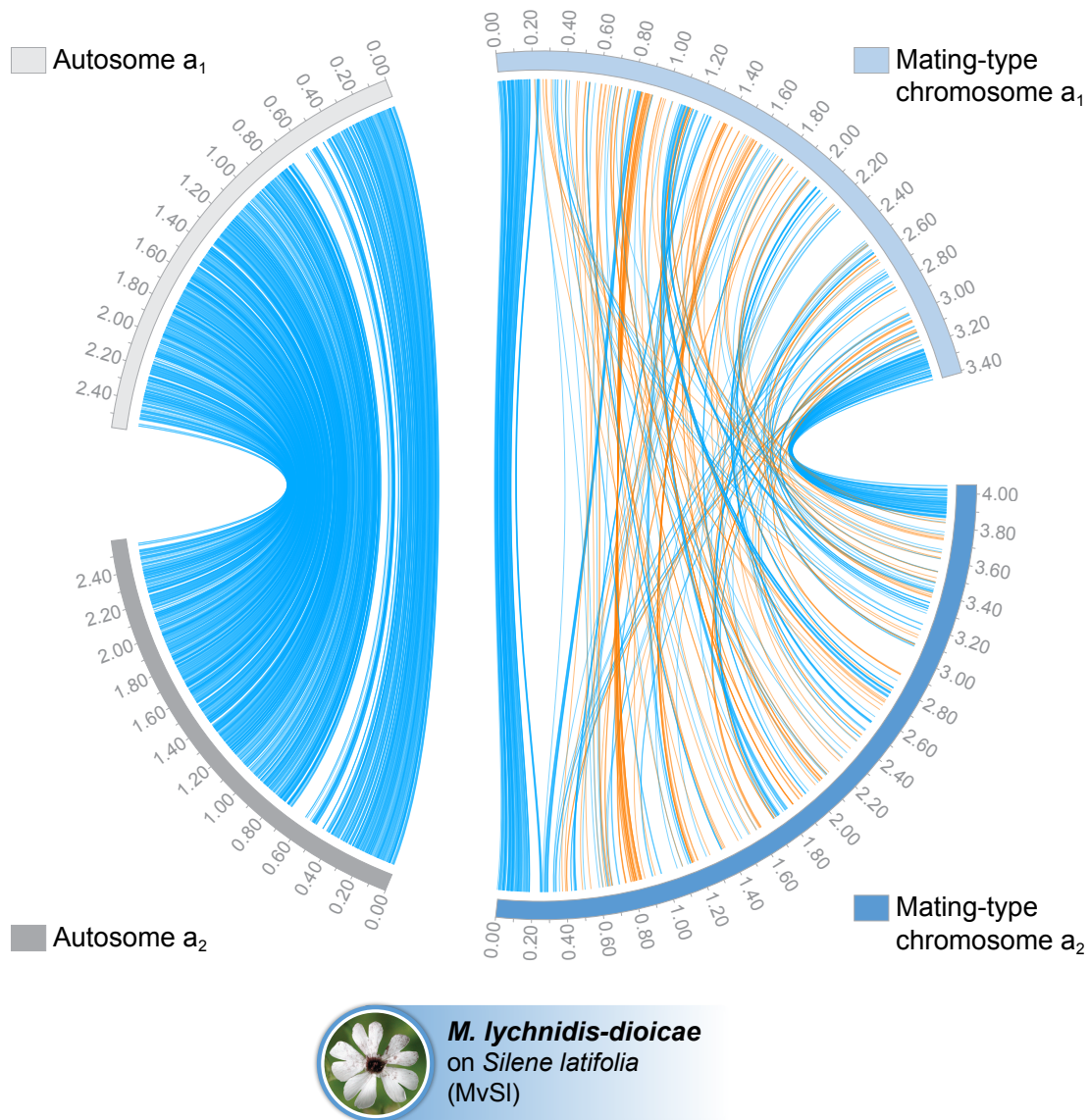


**Fig. S1. Hypotheses for the generation of evolutionary strata in sex or mating-type chromosomes.** (A) Expansion of linkage to the mating-type loci that encompasses partially linked genetic load loci, favoring permanent sheltering in a heterozygous state. This may occur if recombination frequency gradually decreases from the non-recombining region into the PAR, so that partial linkage (linkage disequilibrium) to mating-type or sex-determining genes makes selection against recessive deleterious alleles less efficient. This may allow deleterious alleles to increase in frequency in these PAR regions at the margin of the non-recombining region, further selecting for cessation of recombination. Rare recombination events would generate individuals homozygous for deleterious alleles and would therefore be selected against. Complete linkage of these PAR regions in linkage disequilibrium with mating type would be selected for. Alternative mechanisms for explaining suppressed recombination involve neutral inversions (B) or rearrangements (C) extending the region of suppressed recombination; these could be fixed by drift in one of the sex or mating-type chromosomes and arrest recombination between sex or mating-type chromosomes.



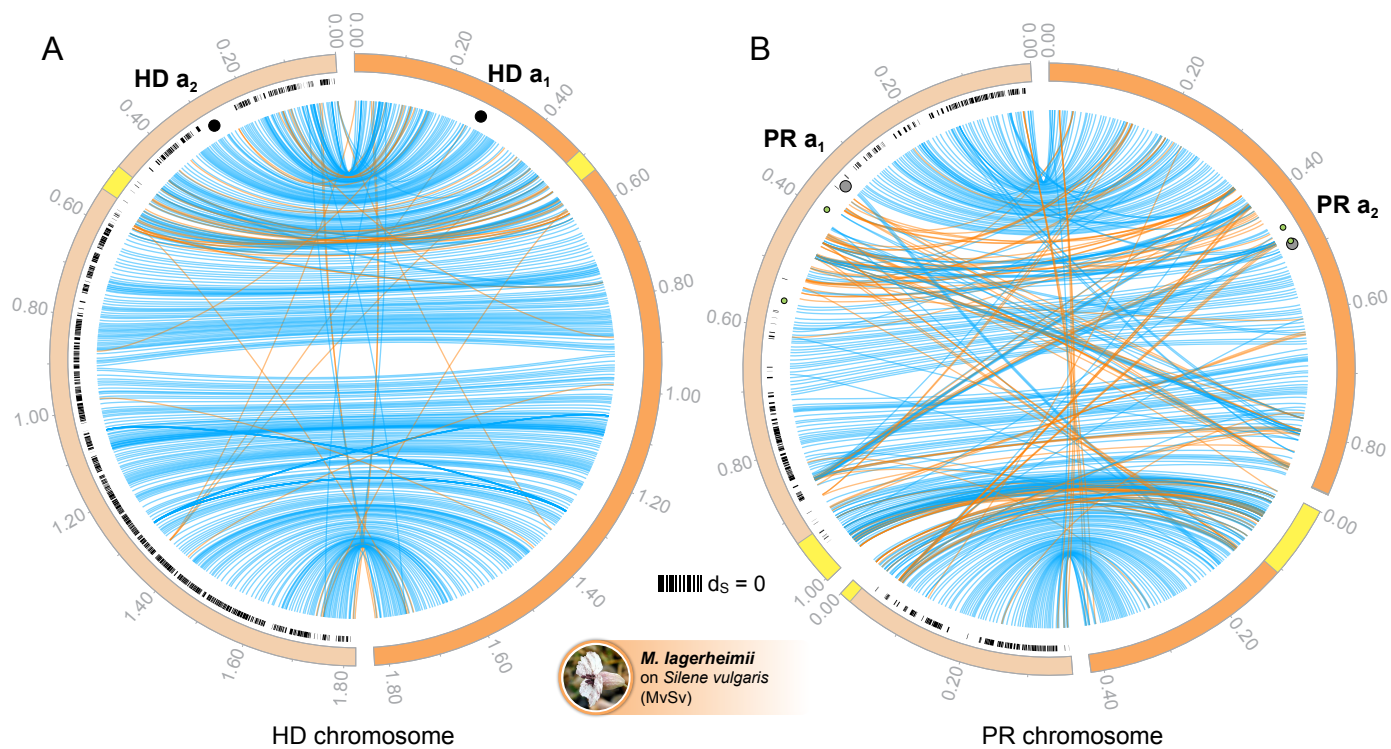
Key: Mating-type loci

**Fig. S2. Odds of compatibility among gametes of a diploid individual in basidiomycete fungi.** Gametes are fully compatible only if they carry different alleles at both mating-type loci, i.e., the PR (including pheromone receptor and pheromone genes, with  $a_1$  and  $a_2$  alleles) and HD (including homeodomain genes, with  $b_1$  and  $b_2$  alleles) loci. (A) Under tetrapolarity with PR and HD mating-type loci unlinked from each other and from the centromeres (shown here located in different chromosomes, blue and red, respectively), the percentage of compatibility of a given gamete among the other gametes produced by the same diploid individual is 25% across multiple meioses (a given gamete is compatible with one of every four gametes), and the percentage is 33% within tetrad (a given gamete is compatible with one of the other three gametes in the tetrad) or 67% (a given gamete is compatible with two of the three remaining gametes in the tetrad) depending on segregation of the mating type alleles. The different types of gametes produced are tetratypes (TT), parental ditypes (PD) or non-parental ditypes (NPD), which depends on allele segregation and on whether a crossing-over occurred between one of the two loci and the centromere. (B) Under tetrapolarity with PR and HD mating-type genes linked to the centromeres of different chromosomes (blue and red, respectively), the percentage of compatibility of a given gamete among the other gametes produced by the same diploid individual is 25% across multiple meioses but 67% within a tetrad (a given gamete is compatible with two of the three other gametes in the tetrad) due to the segregation of the variation occurring only at meiosis I for both mating type loci. The different types of gametes produced are parental ditypes (PD) or non-parental ditypes (NPD), which depends on segregation. (C) Under bipolarity, i.e., with HD and PR loci fully linked to each other on the same chromosome, the percentage of compatibility of a given gamete among the other gametes produced by the same diploid individual is 50% across multiple meioses (a given gamete is compatible with two of every four gametes), and 67% within a single meiotic tetrad (a given gamete is compatible with two of the three other gametes in the tetrad).



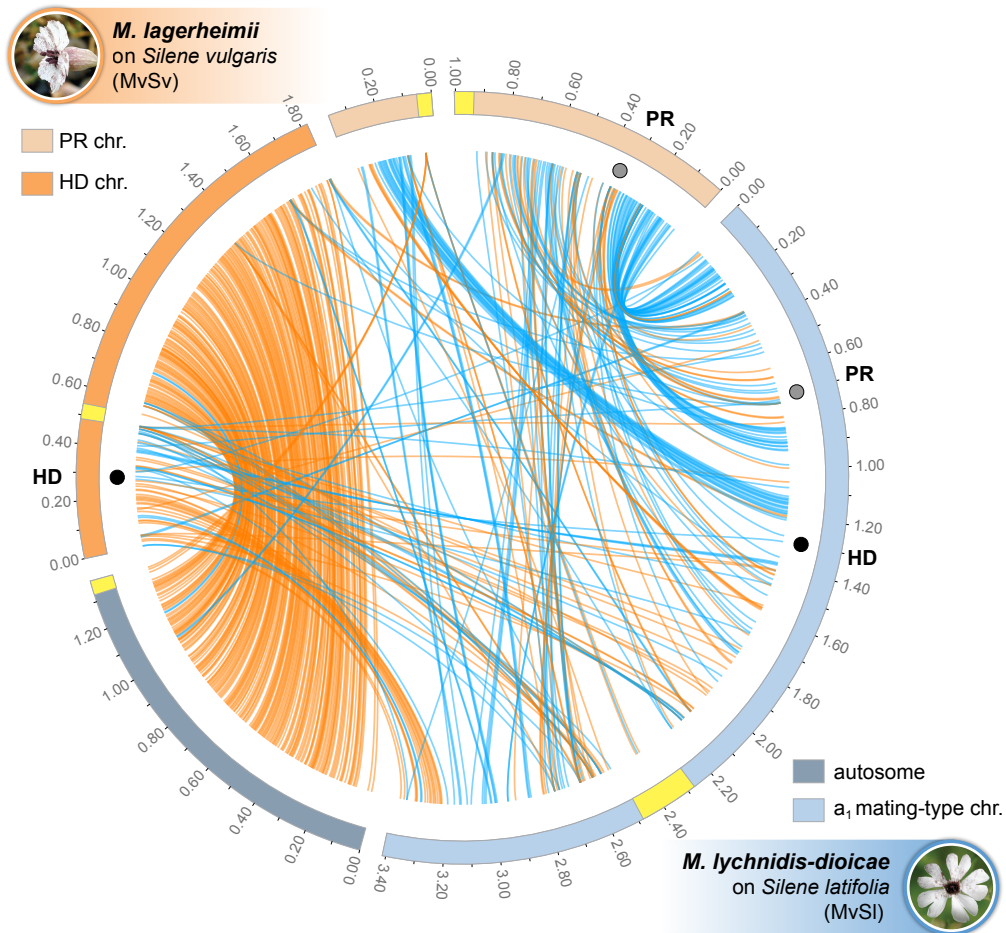
**Fig. S3. Comparison of gene order between *Microbotryum lychnidis-dioicae* chromosomes.**  $a_1$  (top) and  $a_2$  (bottom) mating chromosomes of *M. lychnidis-dioicae*; an autosome pair is shown at the left and the mating-type chromosomes at the right. Blue and orange lines link regions with collinearity across > 5 kb, with the orange links corresponding to inversions. The genomic regions without links correspond to highly rearranged regions. The recombining regions (the autosome and the pseudo-autosomal regions, PARs) are highly collinear between  $a_1$  and  $a_2$  chromosomes, despite a highly selfing mating system, while the non-recombining region on the mating-type chromosome is highly rearranged.





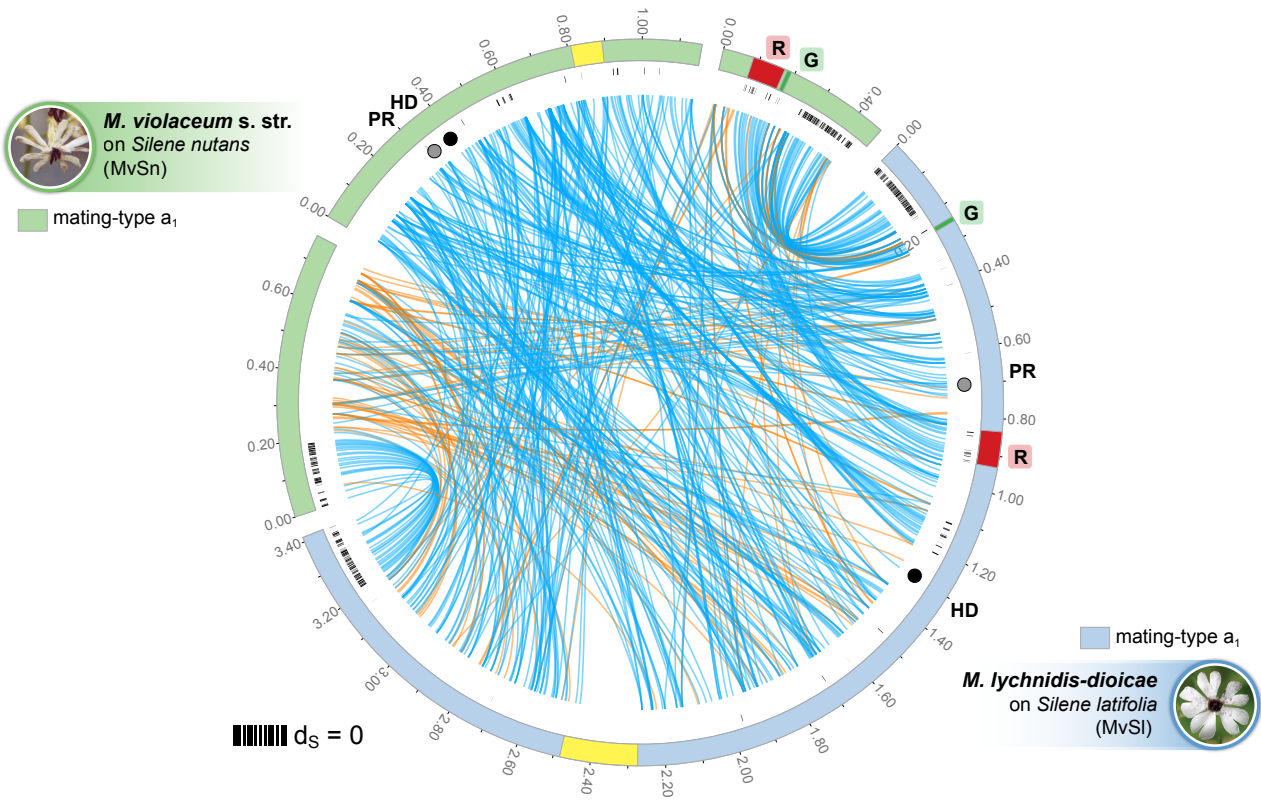
**Fig. S4. Synteny between a<sub>1</sub> and a<sub>2</sub> mating-type chromosomes of *Microbotryum lagerheimii*.** The homeodomain (HD) mating-type chromosomes are displayed in the left circle and the pheromone receptor (PR) mating-type chromosomes in the right circle. The HD genes are indicated by black circles, the pheromone receptor gene by grey circles and the two pheromone genes by small green circles. Blue and orange lines link regions with collinearity across > 1 kb, with the orange links corresponding to inversions. The genomic region between the pheromone receptor and pheromone genes in the a<sub>1</sub> mating-type chromosomes is rearranged compared to the a<sub>2</sub> mating-type chromosomes. Black marks along the contig circles indicate genes with no synonymous differentiation between a<sub>1</sub> and a<sub>2</sub> alleles within the sequenced individual ( $d_s = 0$ ). Yellow regions on the outer track indicate centromere-specific repeats (12).

A

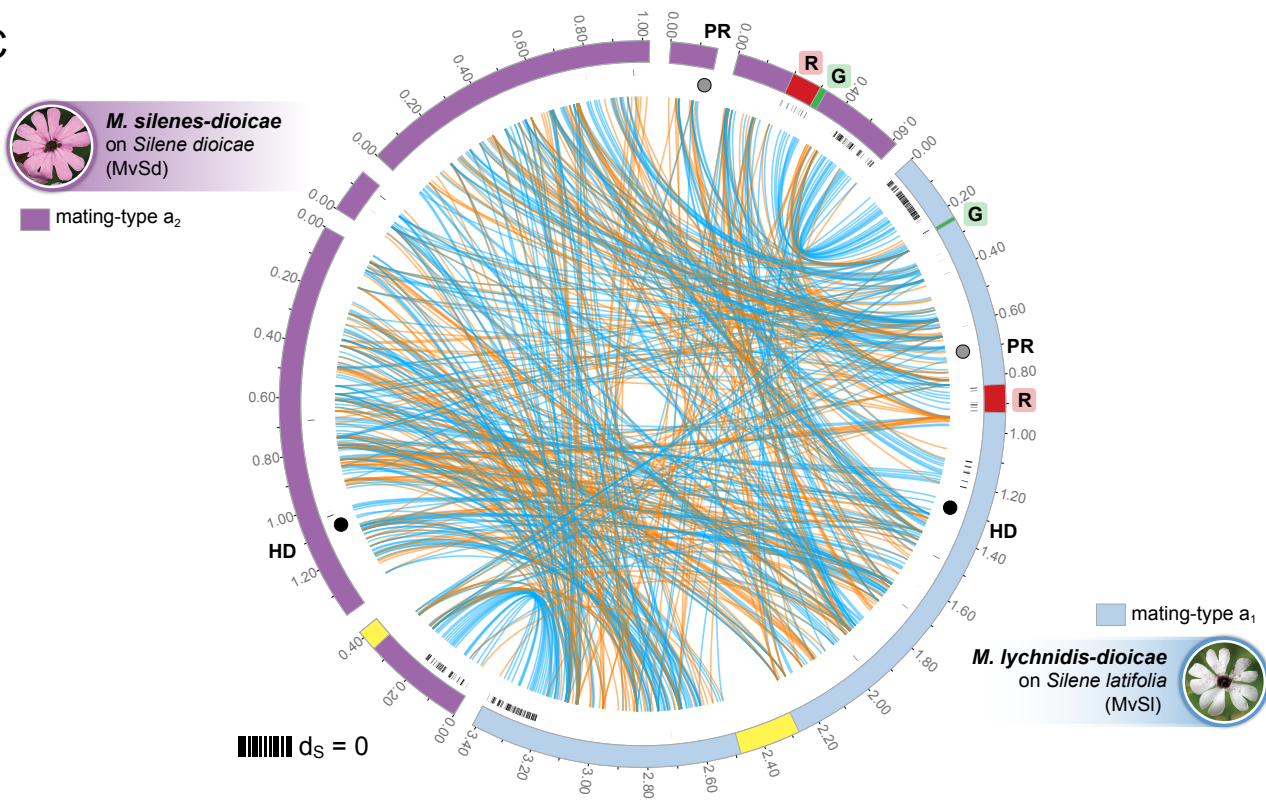


**Fig. S5. Comparison of gene order between mating-type chromosomes.** In all panels (A-E), the outer tracks illustrate assembled chromosomes/contigs and are color coded by species as in Fig. 1. The positions of HD and PR genes are indicated by the black and grey small circles, respectively. Yellow regions on the outer track indicate the centromere-specific repeats (12) and, when represented, red (“R”) and green (“G”) regions indicate the genomic location of the red and green genes as described in Fig. 4B. **(A)** Comparison between the a<sub>1</sub> mating-type chromosomes (chr.) of *Microbotryum lagerheimii* (orange contigs, with light orange for the PR mating-type chromosome and dark orange for the HD mating-type chromosome) and the mating-type chromosome of *M. lychnidis-dioicae* (light blue contig), as well as an autosomal contig of *M. lychnidis-dioicae* (dark blue contig). Blue and orange lines link regions with collinearity across > 2 kb, with orange corresponding to inversions. Genomic regions without links correspond to highly rearranged regions.

B



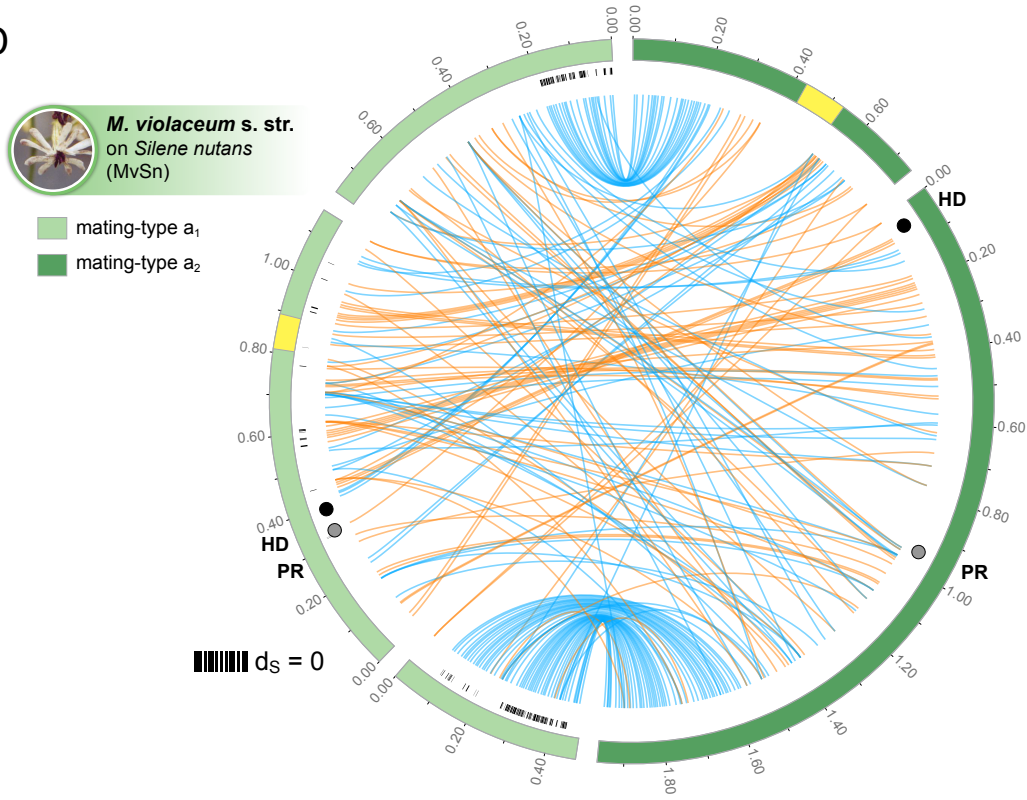
C



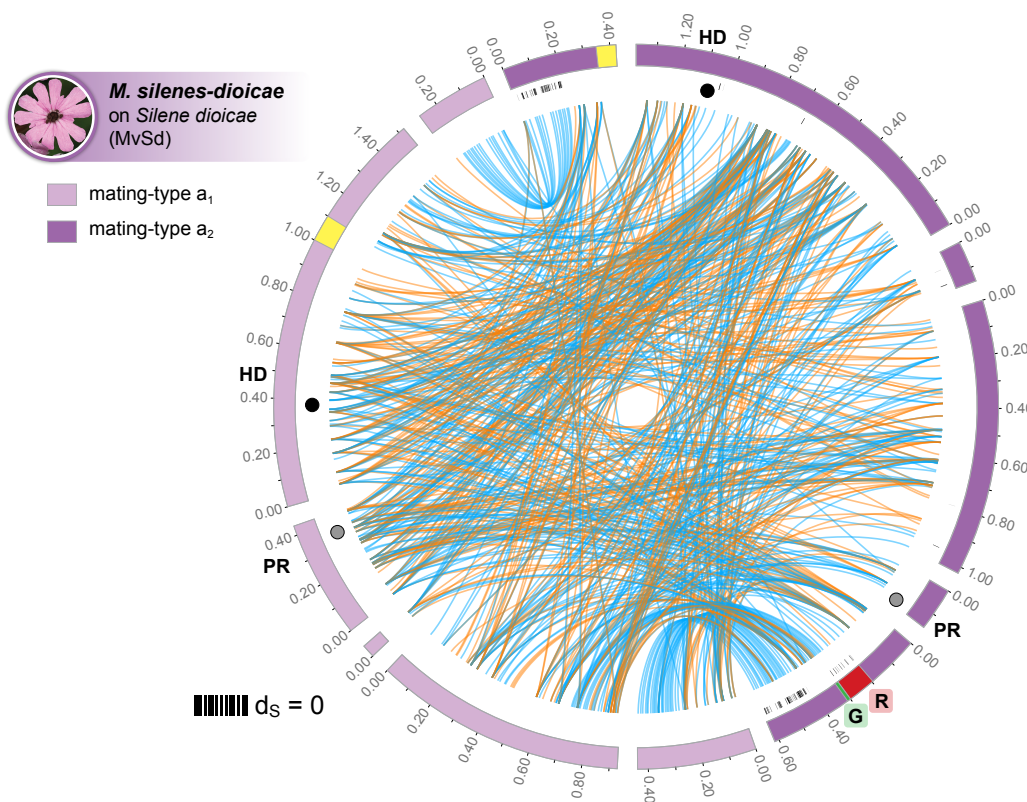
**Fig. S5.** Continued. **(B)** Comparison of  $a_1$  mating-type chromosomes of *Microbotryum lychnidis-dioicae* (right, light blue contig) and *M. violaceum s. str.* (left, light green contigs). It should be noted that the orientation of the central contig in the  $a_1$  *M. violaceum s. str.* mating-type chromosome could not be assessed, so all its links were drawn blue. The black marks along the contig circle indicate genes that have no synonymous substitution between  $a_1$  and  $a_2$  alleles ( $d_s = 0$ ) in *M. lychnidis-dioicae*. **(C)** Comparison between the  $a_1$  mating-type chromosomes of *M. silenes-dioicae* (dark purple contigs on the left) and *M. lychnidis-dioicae* (light blue contig on the right). Blue and orange lines link regions with collinearity across  $> 3$  kb, with orange corresponding to inversions (note that the *M. silenes-dioicae* central contigs could not be ordered or oriented). All the remaining features are as described in panel (A).



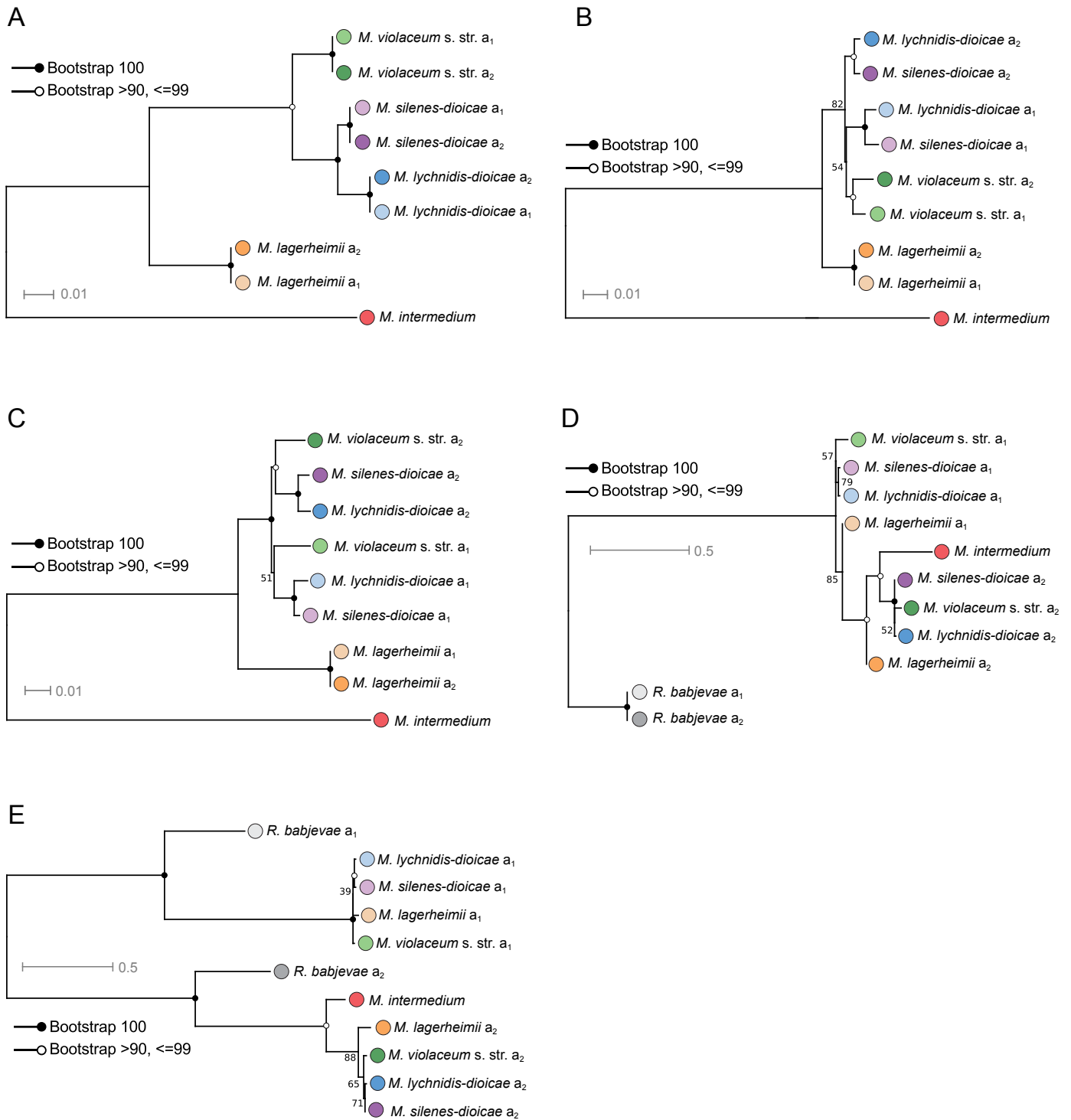
D



E

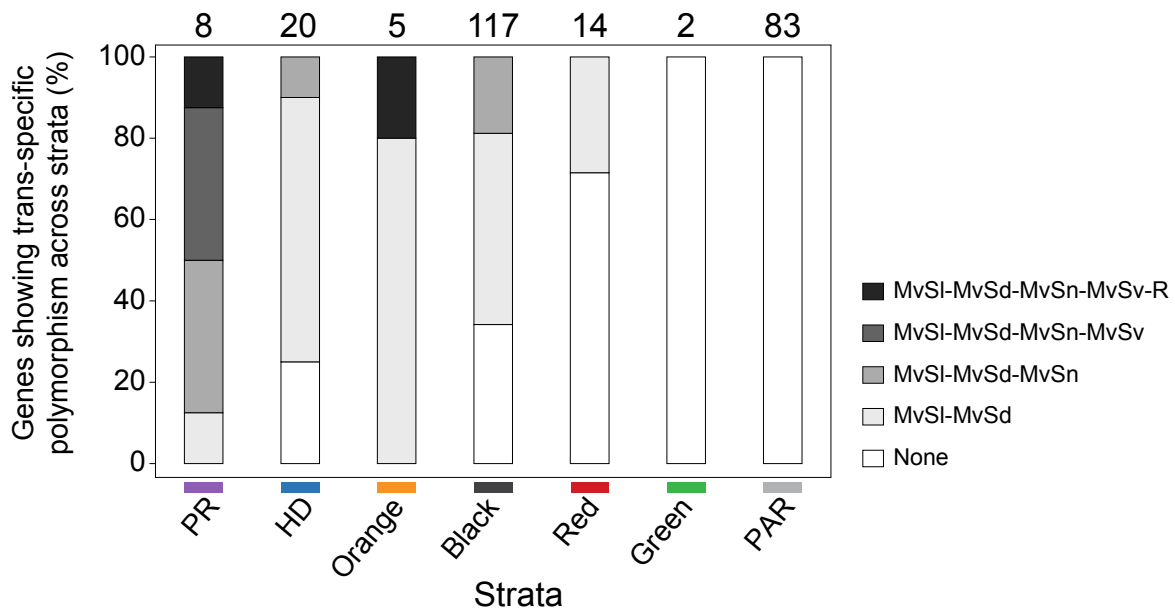


**Fig. S5.** Continued. **(D)** Comparison of a<sub>1</sub> (left, light green) and a<sub>2</sub> (right, dark green) mating-type chromosomes of *M. violaceum s. stricto*. Blue and orange lines link regions with collinearity across >3 kb, with orange corresponding to inversions. It should be noted that the orientation of the central contig in the a<sub>1</sub> mating-type chromosome could not be assessed. The black marks along the contig circle indicate the genes with no synonymous substitutions between a<sub>1</sub> and a<sub>2</sub> alleles ( $d_S=0$ ) in *M. violaceum s. stricto*. **(E)** Comparison of a<sub>1</sub> (left, light purple) and a<sub>2</sub> (right, dark purple) mating-type chromosomes of *M. silenes-dioicae*. Blue and orange lines link regions with collinearity across >3 kb, with orange corresponding to inversions. Note that the central contigs could not be ordered or oriented. The location of the red (“R”) and green (“G”) genes in Fig. 4B are indicated in the a<sub>2</sub> genome. The black marks along the contig circle indicate the genes with no synonymous substitutions between a<sub>1</sub> and a<sub>2</sub> alleles ( $d_S = 0$ ). The remaining features are as described in panel (A).

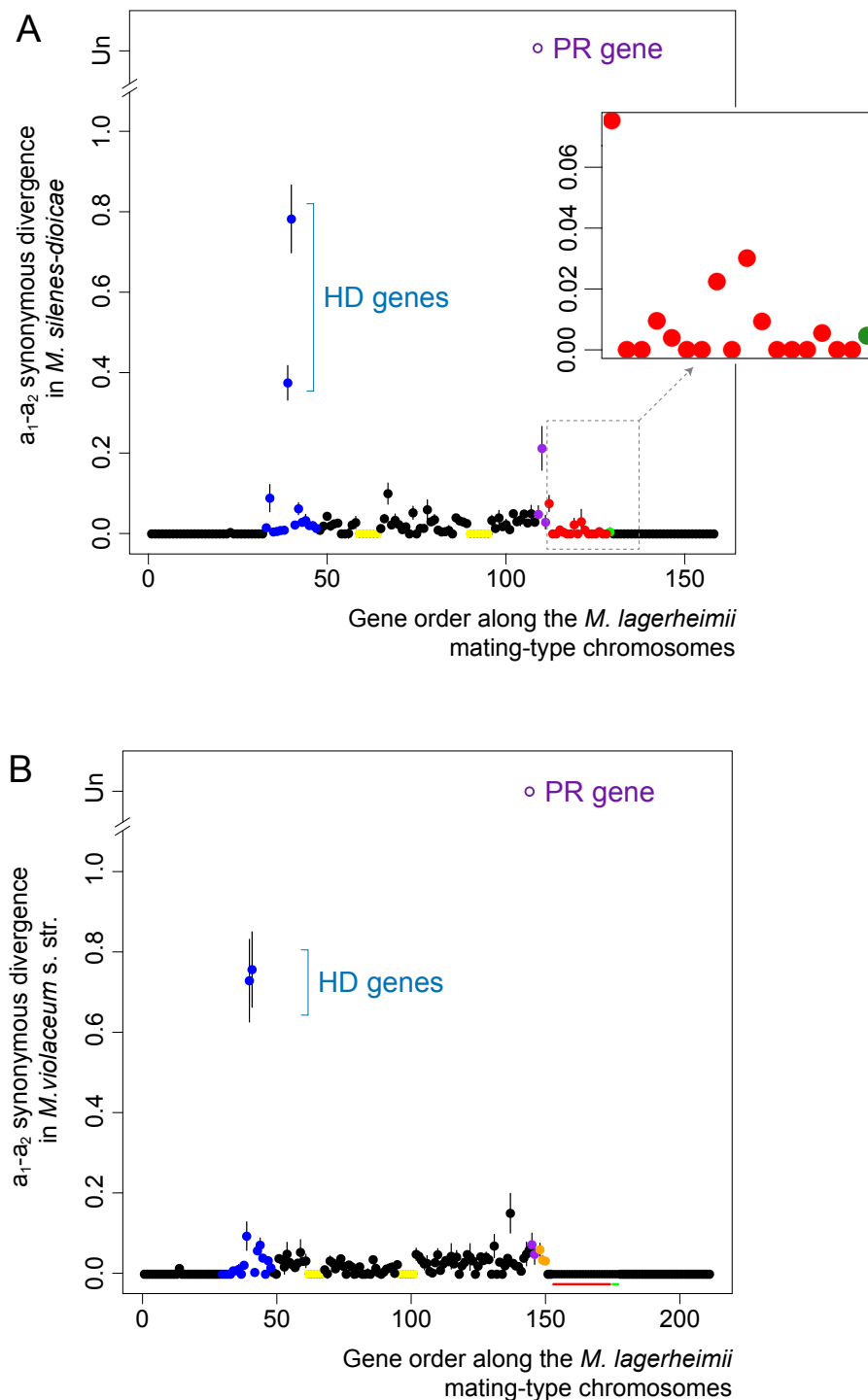


**Fig. S6. Examples of gene genealogies at non-mating type genes showing different levels of trans-specific polymorphism, i.e., with more or less ancient linkage to mating type.** Bootstraps are shown at nodes and the branch length scale is indicated. Strains of the same species are indicated by circles of the same color (a<sub>1</sub>, lighter variant; a<sub>2</sub>, darker variant) and according to Fig. 1. **(A)** No trans-specific polymorphism (orthogroup belonging to the *M. lychnidis-dioicae* PAR). **(B)** Trans-specific polymorphism across *M. lychnidis-dioicae* and *M. silenes-dioicae* (orthogroup belonging to the *M. lychnidis-dioicae* black stratum). **(C)** Trans-specific polymorphism across *M. lychnidis-dioicae*, *M. silenes-dioicae* and *M. violaceum* s. str. (orthogroup belonging to the *M. lychnidis-dioicae* black stratum). **(D)** Trans-specific polymorphism across *M. lychnidis-dioicae*, *M. silenes-dioicae*, *M. violaceum* s. str. and *M. lagerheimii* (orthogroup belonging to the *M. lychnidis-dioicae* PR stratum). **(E)** Trans-specific polymorphism across *M. lychnidis-dioicae*, *M. silenes-dioicae*, *M. violaceum* s. str. *M. lagerheimii* and *Rhodotorula babjevae* (orthogroup belonging to the *M. lychnidis-dioicae* orange stratum).

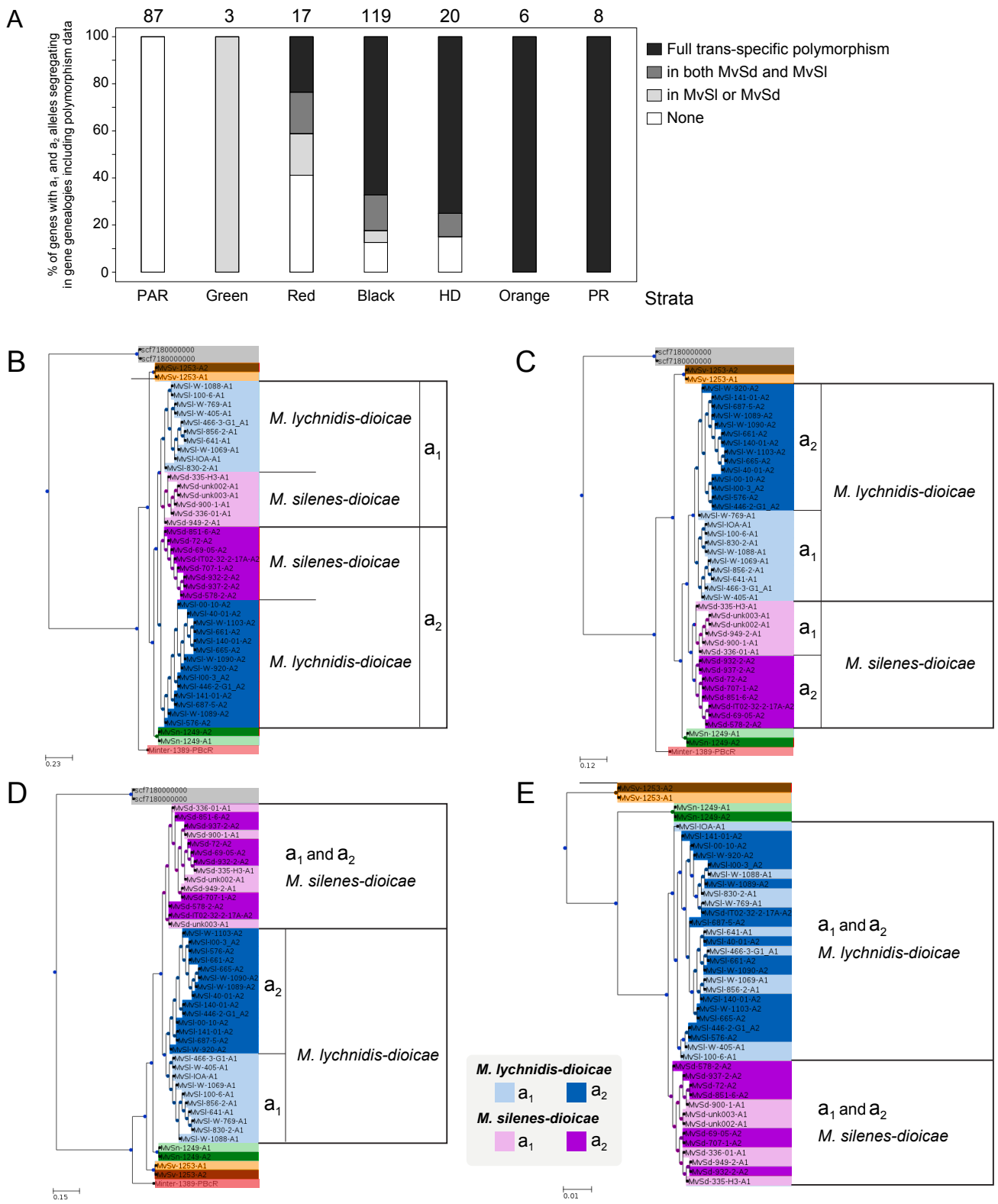




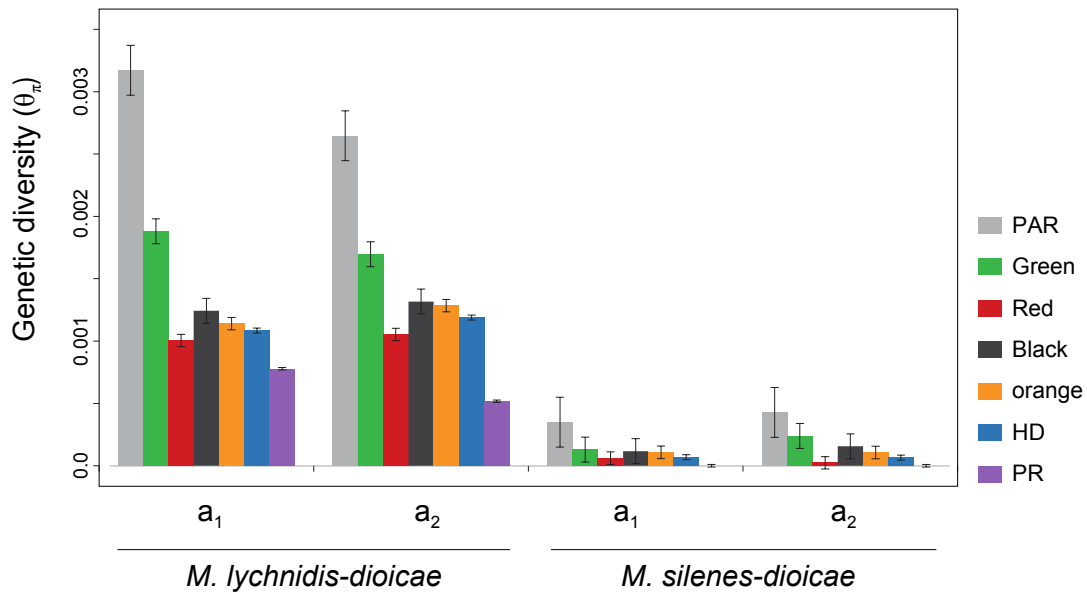
**Fig. S7. Cumulative percentage of genes showing trans-specific polymorphism across strata and age of trans-specific polymorphism relative to speciation events.** Black - single-copy genes with deep trans-specific polymorphism including all species, *M. lychnidis-dioicae* (MvSI), *M. silenes-dioicae* (MvSd), *M. violaceum* s. str. (MvSn), *M. lagerheimii* (MvSv), and even *Rhodotorula babjevae* (R); dark grey - genes showing trans-specific polymorphism including only *M. lychnidis-dioicae*, *M. silenes-dioicae*, *M. violaceum* s. str., and *M. lagerheimii*; grey - genes showing trans-specific polymorphism with only *M. lychnidis-dioicae*, *M. silenes-dioicae* and *M. violaceum* s. str.; light grey - genes with trans-specific polymorphism only between *M. lychnidis-dioicae* and *M. silenes-dioicae*; white - genes with no trans-specific polymorphism. Numbers on the top indicate genes for which gene genealogies with all species and alleles were computed. See Figure S6 for examples of gene genealogies with different levels of trans-specific polymorphism.



**Fig. S8. Synonymous divergence between alleles associated with  $a_1$  and  $a_2$  mating types in the sequenced diploid individual in each of *Microbotryum silenes-dioicae* and *M. violaceum s. stricto*, plotted along the ancestral gene order.** Synonymous divergence between  $a_1$ - $a_2$  in *M. silenes-dioicae* (A) and *M. violaceum s. str.* (B) plotted in the ancestral genomic coordinates, as inferred from the *M. lagerheimii*  $a_1$  mating-type chromosome. In both panels, points are mean  $d_s$  values per gene and error bars the standard errors. All non-transposable element genes shared by the mating-type chromosomes are depicted. Divergence between the  $a_1$  and  $a_2$  pheromone receptor (PR) was too extensive (54) and could not be computed in both species (it is plotted as an "unalignable" open circle to ease visualization of the various strata). The positions of the inferred ancestral centromeres are indicated by yellow boxes. Genes with non-zero  $d_s$  around the PR and HD mating-type loci within the sequenced individual of *M. lagerheimii* are colored in purple and blue, respectively. More recent putative evolutionary strata in *M. silenes-dioicae* (panel A) are indicated in red and green, and zoomed in on the inset shown on the upper right. In panel (B), the genes belonging to more recent evolutionary strata in *M. silenes-dioicae* and *M. lychnidis-dioicae* are indicated by red and green bars.



**Fig. S9.** Patterns of segregation of alleles at non-mating-type genes, associated with the  $a_1$  and  $a_2$  mating types, respectively, in gene genealogies using multiple resequenced genomes of *M. silenes-dioicae* (MvSd) and *M. lychnidis-dioicae* (MvSI). (A) Percentage (%) of gene genealogies showing different levels of segregation of  $a_1$  and  $a_2$  alleles: white - gene genealogies with mixed  $a_1$  and  $a_2$  *M. silenes-dioicae* and *M. lychnidis-dioicae* alleles; light grey - gene genealogies with  $a_1$  and  $a_2$  alleles separated in *M. silenes-dioicae* or *M. lychnidis-dioicae*, but not in both species; dark grey - gene genealogies with  $a_1$  and  $a_2$  alleles separated in both *M. silenes-dioicae* or *M. lychnidis-dioicae*, but without trans-specific polymorphism across the two species; black - gene genealogies with full trans-specific polymorphism. The numbers at the top indicate the number of genes for which information was available. (B-E) Examples of gene genealogies showing different levels of segregation of alleles at non-mating-type genes, associated with the  $a_1$  and  $a_2$  mating types, respectively (shown in lighter and darker color variants, respectively) based on the analyses of multiple *M. silenes-dioicae* (MvSd, purple) and *M. lychnidis-dioicae* (MvSI, blue) genomes. Alleles of *M. lagerheimii* (MvSv, orange), *M. violaceum* s. str. (MvSn, green), *M. intermedium* (Minter, red) and *Rhodotorula babjevae* (sfc, gray) are also included; (B) gene genealogy with full trans-specific polymorphism (red stratum); (C) alleles at non-mating-type genes, associated with the  $a_1$  and  $a_2$  mating types, respectively, separated in both *M. silenes-dioicae* and *M. lychnidis-dioicae*, but without trans-specific polymorphism across the two species (red stratum); (D) gene genealogy with alleles at non-mating-type genes, associated with the  $a_1$  and  $a_2$  mating types, respectively, separated in *M. lychnidis-dioicae* only (red stratum); (E) intermingled alleles at non-mating-type genes, associated with the  $a_1$  and  $a_2$  mating types, respectively, of both *M. silenes-dioicae* and *M. lychnidis-dioicae* (pseudo-autosomal region, PAR).



**Fig. S10. *Microbotryum lychnidis-dioicae* and *M. silenes-dioicae* genetic diversity ( $\theta_{\pi}$ ) based on multiple genomes.** Diversity ( $\pm$ SE) was estimated in  $a_1$  or  $a_2$  mating type separately, in the pseudo-autosomal regions (PARs) and the different evolutionary strata (green, red, black, orange, HD-proximal and PR-proximal). *M. silenes-dioicae* is known to display lower genetic diversity than *M. lychnidis-dioicae* (48, 51, 62).

## List of SI Tables

**Table S1.** Synonymous substitution ( $d_s$ ) means within diploid individuals between alleles associated to the  $a_1$  and  $a_2$  mating types for genes in the different putative strata and the pseudo-autosomal regions (PARs), number of genes and estimated dates for recombination suppression.

**Table S2.** Synonymous substitution ( $d_s$ ) between alleles at non-mating-type genes, associated with the  $a_1$  and  $a_2$  mating types, respectively, in three outgroup species for the genes belonging to the purple and blue strata in *Microbotryum lagerheimii*.

**Table S3.** Gene ID and putative function of the genes located in the different evolutionary strata in *Microbotryum lychnidis-dioicae* (MvSI) (shared with *M. lagerheimii* (MvSv) mating-type chromosomes).

**Table S4.** Raw data summary and assembly statistics for the genomes generated in this study.

**Table S5.** *Microbotryum lychnidis-dioicae* and *M. silenes-dioicae* strains used for performing polymorphism analyses.

**Table S6.** Nucleotide diversity ( $\theta_\pi$ ) and divergence (mean number of pairwise differences) estimates in pseudo-autosomal regions and evolutionary strata in *M. lychnidis-dioicae* and *M. silenes-dioicae*.

**Table S7.** Results of maximum-likelihood Hudson-Kreitman-Agade (MLHKA) tests for assessing whether differences in diversity between each evolutionary stratum and the pseudo-autosomal regions (PARs) is due to balancing selection or elevated mutation rates.



**Table S1. Synonymous substitution ( $d_s$ ) means within diploid individuals between alleles associated to the a1 and a2 mating types for genes in the different putative strata and the pseudo-autosomal regions (PARs), number of genes and estimated dates for recombination suppression.** Mean  $d_s$  and standard error (SE) across genes between the large putative strata, numbered in their presumed order of origin of suppressed recombination, whether they involve mating-type genes, the *Microbotryum* species in which recombination is suppressed in these strata, the estimated date of recombination suppression in strata where inference was possible, and number of genes ancestrally in *M. lychnidis-dioicae* in the strata (many genes have been lost from one or the other mating types following recombination suppression).

Stratum order	Stratum	Mean $d_s \pm SE$	Involving mating-type genes?	Species in which recombination is suppressed	Estimated date recombination suppression	Number of genes ancestrally in <i>M. lychnidis-dioicae</i>
1	PR mating-type gene	Not alignable	Yes	All species	370 MYA (54)	3
2	HD mating-type genes	$0.62 \pm 0.326$	Yes	All species		2
3	PR-proximal genes (purple)	$0.112 \pm 0.040$	No	All species	2.1 MYA	31
4	HD-proximal genes (blue)	$0.103 \pm 0.045$	No	All species		14
5	Black (linking PR and HD mating-type loci)	$0.032 \pm 0.002$	Yes	<i>M. lychnidis-dioicae</i> , <i>M. silenes-dioicae</i> and <i>M. violaceum</i> s. str.	1.3 MYA	85
6	Orange	$0.052 \pm 0.015$	No	<i>M. lychnidis-dioicae</i> , <i>M. silenes-dioicae</i> and <i>M. violaceum</i> s. str.		7
7	Red	$0.018 \pm 0.049$	No	<i>M. lychnidis-dioicae</i> and <i>M. silenes-dioicae</i>	0.9 MYA	22
8	Green	$0.004 \pm 0.003$	No	<i>M. lychnidis-dioicae</i> and <i>M. silenes-dioicae</i>		3
	pPAR (from the ancestral HD chromosome)	<0.0001	No	None		73
	qPAR (from the ancestral PR chromosome)	<0.0001	No	None		80

**Table S2. Synonymous substitution ( $d_s$ ) between alleles associated with the  $a_1$  and  $a_2$  mating types in three individuals of outgroup species for the genes belonging to the *Microbotryum lagerheimii* purple and blue strata. (A) *Sporobolomyces salmonicolor* (Ss) strains CBS 6832 vs. ML 2241; (B) *Rhodotorula toruloides* (Rt) strains JCM 10020 (A1B1) vs. JCM 10021 (A2B1); (C) *Leucosporidium scottii* (Ls) strains CBS 5930 vs. CBS 5931. The ID of the query gene from the reference *M. lychnidis-dioicae* genome, the ID in the outgroup genome, the  $d_s$  value and the putative function inferred from similarity are given. The only functions associated with mating or mating-type functions are the mating-type genes (highlighted in bold). The pheromone receptor genes are not shown as they could not be aligned between mating types.**

(A)

MvSI-Query ID	Paired ID in the outgroup	$d_s$	Functional annotation
MvSI-1064-A1-R4_A1g00025	Ss_CBS6832 g3490.t1;Ss_ML2241 g967.t1	0.0253	glycoside hydrolase family 28 protein
MvSI-1064-A1-R4_A1g00175	Ss_CBS6832 g4453.t1;Ss_ML2241 g4966.t1	0.0177	ABC1-A atypical protein kinase
MvSI-1064-A1-R4_A1g00230	Ss_CBS6832 g4436.t1;Ss_ML2241 g4982.t1	0.0045	ASF1-like histone chaperone
MvSI-1064-A1-R4_A1g00478	Ss_CBS6832 g2690.t1;Ss_ML2241 g5108.t1	0.0167	protein disulfide-isomerase
MvSI-1064-A1-R4_A1g00495	Ss_CBS6832 g6519.t1;Ss_ML2241 g6611.t1	0.0232	nucleotide-binding protein/RRM_SF
<b>MvSI-1064-A1-R4_A1g00591</b>	<b>Ss_CBS6832 g5519.t1;Ss_ML2241 g1781.t1</b>	<b>0.3037</b>	<b>homeodomain transcription factor HD1</b>
<b>MvSI-1064-A1-R4_A1g00592</b>	<b>Ss_CBS6832 g5518.t1;Ss_ML2241 g1780.t1</b>	<b>0.3041</b>	<b>homeodomain transcription factor HD2</b>
MvSI-1064-A1-R4_A1g00593	Ss_CBS6832 g2692.t1;Ss_ML2241 g5106.t1	0.0223	5'-3' exoribonuclease 2
MvSI-1064-A1-R4_A1g00659	Ss_CBS6832 g307.t1;Ss_ML2241 g1640.t1	0.0285	alpha-mannosyltransferase, glycosyltransferase family 71 protein
MvSI-1064-A1-R4_A1g00726	Ss_CBS6832 g4454.t1;Ss_ML2241 g4962.t1	0.0173	defective in Cullin neddylation protein 1/UBA_like_SF
MvSI-1064-A1-R4_A1g00772	Ss_CBS6832 g2693.t1;Ss_ML2241 g5105.t1	0.0260	U3 snoRNP-associated protein Esf2 / RRM_ABT1_like
MvSI-1064-A1-R4_A1g00800	Ss_CBS6832 g2586.t1 corr;Ss_ML2241 g6295.t1 corr	0.0368	S-(hydroxymethyl)glutathione dehydrogenase / alcohol dehydrogenase
MvSI-1064-A1-R4_A1g00801	Ss_CBS6832 g6218.t1;Ss_ML2241 g6054.t1	0.0308	exosome complex protein LRP1
MvSI-1064-A1-R4_A1g00844	Ss_CBS6832 g3596.t1;Ss_ML2241 g741.t1	0.0609	hypothetical protein
MvSI-1064-A1-R4_A1g00854	Ss_CBS6832 g4447.t1;Ss_ML2241 g4977.t1	0.0555	karyopherin kap95
MvSI-1064-A1-R4_A1g00970	Ss_CBS6832 g2670.t1;Ss_ML2241 g3808.t1	0.0287	T-complex protein 1 subunit beta
MvSI-1064-A1-R4_A1g00971	Ss_CBS6832 g2640.t1;Ss_ML2241 g6034.t1	0.0593	dolichyldiphosphatase / PAP2_like
MvSI-1064-A1-R4_A1g00972	Ss_CBS6832 g2641.t1;Ss_ML2241 g6035.t1	0.0449	hypothetical protein
MvSI-1064-A1-R4_A1g00997	Ss_CBS6832 g3124.t1;Ss_ML2241 g781.t1	0.0024	hypothetical protein
MvSI-1064-A1-R4_A1g00998	Ss_CBS6832 g2641.t1 corr;Ss_ML2241 g6036.t1 corr	0.0293	U3 small nucleolar ribonucleoprotein IMP3 / Ribosomal protein S4/S9
MvSI-1064-A1-R4_A1g01017	Ss_CBS6832 g4458.t1;Ss_ML2241 g4960.t1	0.0349	phosphatidylethanolamine N-methyltransferase (PEMT)
MvSI-1064-A1-R4_A1g01154	Ss_CBS6832 g2758.t1;Ss_ML2241 g2001.t1	0.0684	ribosomal protein L1
MvSI-1064-A1-R4_A1g01159	Ss_CBS6832 g5531.t1;Ss_ML2241 g1794.t1	0.0478	bZIP transcription factor (AP-1)
MvSI-1064-A1-R4_A1g01160	Ss_CBS6832 g2542.t1;Ss_ML2241 g1752.t1	0.0300	COP9 signalosome complex subunit 2
MvSI-1064-A1-R4_A1g01161	Ss_CBS6832 g2659.t1;Ss_ML2241 g3820.t1	0.0212	mitochondrial carrier protein
MvSI-1064-A1-R4_A1g01162	Ss_CBS6832 g5512.t1;Ss_ML2241 g1774.t1	0.0186	hypothetical protein
MvSI-1064-A1-R4_A1g01163	Ss_CBS6832 g2660.t1;Ss_ML2241 g3819.t1	0.0627	ADP-ribosylation factor-like protein 2
MvSI-1064-A1-R4_A1g01388	Ss_CBS6832 g1143.t1;Ss_ML2241 g6591.t1	0.0183	ribosome assembly protein Noc2
MvSI-1064-A1-R4_A1g01440	Ss_CBS6832 g2708.t1;Ss_ML2241 g5917.t1	0.0458	autophagy-related protein 7 / E1-like protein-activating enzyme Gsa7p/Apg7p
MvSI-1064-A1-R4_A1g01441	Ss_CBS6832 g2773.t1;Ss_ML2241 g1987.t1	0.0234	zinc finger, C3HC4 RING-type protein
MvSI-1064-A1-R4_A1g01457	Ss_CBS6832 g2652.t1;Ss_ML2241 g3827.t1	0.0269	unnamed protein product
MvSI-1064-A1-R4_A1g01458	Ss_CBS6832 g2652.t1;Ss_ML2241 g3827.t1	0.0269	DNA-directed RNA polymerase III subunit RPC1

**Table S2.** Continued.

(B)

MvSI-Query ID	Paired ID in the outgroup	dS	Functional annotation
MvSI-1064-A1-R4_A1g00175	Rt_JCM10020_06889-R1;Rt_JCM10021_05528-R1	0.338	ABC1-A atypical protein kinase
MvSI-1064-A1-R4_A1g00230	Rt_JCM10020_06874-R1;Rt_JCM10021_05544-R1	0.382	ASF1-like histone chaperone
MvSI-1064-A1-R4_A1g00478	Rt_JCM10020_01934-R1;Rt_JCM10021_07790-R1	0.618	protein disulfide-isomerase
MvSI-1064-A1-R4_A1g00495	Rt_JCM10020_01942-R1;Rt_JCM10021_02138-R1	0.363	nucleotide-binding protein/RRM_SF
<b>MvSI-1064-A1-R4_A1g00591</b>	<b>Rt_JCM10020_02036-R1;Rt_JCM10021_02036-R1</b>	<b>0.692</b>	<b>homeodomain transcription factor HD1</b>
<b>MvSI-1064-A1-R4_A1g00592</b>	<b>Rt_JCM10020_02035-R1;Rt_JCM10021_02037-R1</b>	<b>0.648</b>	<b>homeodomain transcription factor HD2</b>
MvSI-1064-A1-R4_A1g00593	Rt_JCM10020_01962-R1;Rt_JCM10021_02108-R1	0.423	5'-3' exoribonuclease 2
MvSI-1064-A1-R4_A1g00659	Rt_JCM10020_01652-R1;Rt_JCM10021_00547-R1	0.357	alpha-mannosyltransferase, glycosyltransferase family 71 protein
MvSI-1064-A1-R4_A1g00726	Rt_JCM10020_06891-R1;Rt_JCM10021_05538-R1	0.210	defective in Cullin neddylation protein 1/UBA_like_SF
MvSI-1064-A1-R4_A1g00772	Rt_JCM10020_01963-R1;Rt_JCM10021_02107-R1	0.554	U3 snoRNP-associated protein Esf2 / RRM_ABT1_like
MvSI-1064-A1-R4_A1g00800	Rt_JCM10020_02223-R1;Rt_JCM10021_01846-R1	0.361	S-(hydroxymethyl)glutathione dehydrogenase / alcohol dehydrogenase
MvSI-1064-A1-R4_A1g00801	Rt_JCM10020_01998-R1;Rt_JCM10021_02076-R1	0.334	exosome complex protein LRP1
MvSI-1064-A1-R4_A1g00854	Rt_JCM10020_06883-R1;Rt_JCM10021_05536-R1	0.195	karyopherin kap95
MvSI-1064-A1-R4_A1g00970	Rt_JCM10020_02405-R1;Rt_JCM10021_03563-R1	0.472	T-complex protein 1 subunit beta
MvSI-1064-A1-R4_A1g00971	Rt_JCM10020_02275-R1;Rt_JCM10021_01791-R1	0.442	dolichyldiphosphatase / PAP2_like
MvSI-1064-A1-R4_A1g00972	Rt_JCM10020_02276-R1;Rt_JCM10021_01790-R1	0.367	hypothetical protein
MvSI-1064-A1-R4_A1g00997	Rt_JCM10020_04030-R1;Rt_JCM10021_02518-R1	0.423	hypothetical protein
MvSI-1064-A1-R4_A1g00998	Rt_JCM10020_02282-R1;Rt_JCM10021_01784-R1	0.496	U3 small nucleolar ribonucleoprotein IMP3 / Ribosomal protein S4/S9
MvSI-1064-A1-R4_A1g01017	Rt_JCM10020_06899-R1;Rt_JCM10021_05518-R1	0.523	phosphatidylethanolamine N-methyltransferase (PEMT)
MvSI-1064-A1-R4_A1g01154	Rt_JCM10020_02129-R1;Rt_JCM10021_01938-R1	0.223	ribosomal protein L1
MvSI-1064-A1-R4_A1g01159	Rt_JCM10020_02186-R1;Rt_JCM10021_01883-R1	0.270	bZIP transcription factor (AP-1)
MvSI-1064-A1-R4_A1g01160	Rt_JCM10020_02367-R1;Rt_JCM10021_03528-R1	0.537	COP9 signalosome complex subunit 2
MvSI-1064-A1-R4_A1g01161	Rt_JCM10020_02382-R1;Rt_JCM10021_03542-R1	0.334	mitochondrial carrier protein
MvSI-1064-A1-R4_A1g01162	Rt_JCM10020_02317-R1;Rt_JCM10021_01748-R1	0.370	hypothetical protein
MvSI-1064-A1-R4_A1g01163	Rt_JCM10020_02358-R1;Rt_JCM10021_01709-R1	0.425	ADP-ribosylation factor-like protein 2
MvSI-1064-A1-R4_A1g01388	Rt_JCM10020_04617-R1;Rt_JCM10021_04683-R1	0.394	ribosome assembly protein Noc2
MvSI-1064-A1-R4_A1g01440	Rt_JCM10020_05792-R1;Rt_JCM10021_07432-R1	0.565	autophagy-related protein 7 / E1-like protein-activating enzyme Gsa7p/Apg7p
MvSI-1064-A1-R4_A1g01441	Rt_JCM10020_02100-R1;Rt_JCM10021_01969-R1	0.380	zinc finger, C3HC4 RING-type protein
MvSI-1064-A1-R4_A1g01458	Rt_JCM10020_02343-R1;Rt_JCM10021_01724-R1	0.559	DNA-directed RNA polymerase III subunit RPC1

**Table S2.** Continued.

(C)

MvSI-Query ID	Paired ID in the outgroup	dS	Functional annotation
MvSI-1064-A1-R4_A1g00025	Ls_CBS5930 g7897.t1;Ls_CBS5931 g4047.t1	0.0000	glycoside hydrolase family 28 protein
MvSI-1064-A1-R4_A1g00175	Ls_CBS5930 g8243.t1;Ls_CBS5931 g3668.t1	0.0205	ABC1-A atypical protein kinase
MvSI-1064-A1-R4_A1g00230	Ls_CBS5930 g8248.t1;Ls_CBS5931 g3672.t1	0.1701	ASF1-like histone chaperone
MvSI-1064-A1-R4_A1g00478	Ls_CBS5930 g3772.t1 corr;Ls_CBS5931 g4597.t1 corr	0.0000	protein disulfide-isomerase
MvSI-1064-A1-R4_A1g00495	Ls_CBS5930 g3767.t1;Ls_CBS5931 g4602.t1	0.0000	nucleotide-binding protein/RRM_SF
<b>MvSI-1064-A1-R4_A1g00591</b>	<b>Ls_CBS5930 g2009.t1;Ls_CBS5931 g4188.t1</b>	<b>0.5762</b>	<b>homeodomain transcription factor HD1</b>
<b>MvSI-1064-A1-R4_A1g00592</b>	<b>Ls_CBS5930 g2008.t1;Ls_CBS5931 g4189.t1</b>	<b>0.5376</b>	<b>homeodomain transcription factor HD2</b>
MvSI-1064-A1-R4_A1g00593	Ls_CBS5930 g3769.t1;Ls_CBS5931 g4600.t1	0.0000	5'-3' exoribonuclease 2
MvSI-1064-A1-R4_A1g00659	Ls_CBS5930 g5144.t1;Ls_CBS5931 g1185.t1	0.0117	alpha-mannosyltransferase, glycosyltransferase family 71 protein defective in Cullin neddylation protein 1/UBA_like_SF
MvSI-1064-A1-R4_A1g00726	Ls_CBS5930 g3411.t1;Ls_CBS5931 g3664.t1	0.0155	U3 snoRNP-associated protein Esf2 / RRM_ABT1_like
MvSI-1064-A1-R4_A1g00772	Ls_CBS5930 g3768.t1;Ls_CBS5931 g4601.t1	0.0000	S-(hydroxymethyl)glutathione dehydrogenase / alcohol dehydrogenase
MvSI-1064-A1-R4_A1g00800	Ls_CBS5930 g1990.t1 corr;Ls_CBS5931 g4209.t1 corr	0.0000	hypothetical protein
MvSI-1064-A1-R4_A1g00844	Ls_CBS5930 g3965.t1;Ls_CBS5931 g6949.t1	0.0058	hypothetical protein
MvSI-1064-A1-R4_A1g00854	Ls_CBS5930 g8369.t1 corr;Ls_CBS5931 g3662.t1	0.0159	karyopherin kap95
MvSI-1064-A1-R4_A1g00970	Ls_CBS5930 g6624.t1 corr;Ls_CBS5931 g3773.t1 corr	0.0000	T-complex protein 1 subunit beta
MvSI-1064-A1-R4_A1g00971	Ls_CBS5930 g5099.t1;Ls_CBS5931 g239.t1	0.0488	dolichyldiphosphatase / PAP2_like
MvSI-1064-A1-R4_A1g00972	Ls_CBS5930 g5100.t1;Ls_CBS5931 g238.t1	0.0065	hypothetical protein
MvSI-1064-A1-R4_A1g00997	Ls_CBS5930 g1916.t1;Ls_CBS5931 g98.t1	0.0062	hypothetical protein
MvSI-1064-A1-R4_A1g00998	Ls_CBS5930 g3566.t1;Ls_CBS5931 g147.t1	0.1334	U3 small nucleolar ribonucleoprotein IMP3 / Ribosomal protein S4/S9
MvSI-1064-A1-R4_A1g01017	Ls_CBS5930 g3407.t1;Ls_CBS5931 g8158.t1	0.0516	phosphatidylethanolamine N-methyltransferase (PEMT)
MvSI-1064-A1-R4_A1g01154	Ls_CBS5930 g5411.t1 corr;Ls_CBS5931 g3145.t1	0.0786	ribosomal protein L1
MvSI-1064-A1-R4_A1g01159	Ls_CBS5930 g3515.t1;Ls_CBS5931 g197.t1	0.0000	bZIP transcription factor (AP-1)
MvSI-1064-A1-R4_A1g01160	Ls_CBS5930 g5395.t1;Ls_CBS5931 g3159.t1	0.0239	COP9 signalosome complex subunit 2
MvSI-1064-A1-R4_A1g01161	Ls_CBS5930 g6628.t1;Ls_CBS5931 g3769.t1	0.0000	mitochondrial carrier protein
MvSI-1064-A1-R4_A1g01162	Ls_CBS5930 g5385.t1;Ls_CBS5931 g3169.t1	0.0986	hypothetical protein
MvSI-1064-A1-R4_A1g01163	Ls_CBS5930 g6627.t1;Ls_CBS5931 g3770.t1	0.0000	ADP-ribosylation factor-like protein 2
MvSI-1064-A1-R4_A1g01388	Ls_CBS5930 g4892.t1 corr;Ls_CBS5931 g6975.t1 corr	0.0948	ribosome assembly protein Noc2
MvSI-1064-A1-R4_A1g01440	Ls_CBS5930 g3762.t1;Ls_CBS5931 g4607.t1	0.0000	autophagy-related protein 7 / E1-like protein-activating enzyme Gsa7p/Apg7p
MvSI-1064-A1-R4_A1g01441	Ls_CBS5930 g3507.t1;Ls_CBS5931 g205.t1	0.0000	zinc finger, C3HC4 RING-type protein
MvSI-1064-A1-R4_A1g01457	Ls_CBS5930 g6847.t1;Ls_CBS5931 g3790.t1	0.0022	unnamed protein product
MvSI-1064-A1-R4_A1g01458	Ls_CBS5930 g6847.t1;Ls_CBS5931 g3790.t1	0.0022	DNA-directed RNA polymerase III subunit RPC1

**Table S3. Gene ID and putative function of the genes located in the different evolutionary strata in *Microbotryum lychnidis-dioicae* (MvSI) (shared with *M. lagerheimii* (MvSv) mating-type chromosomes). The two HD genes are highlighted in bold and grey. The  $d_s$  corresponds to synonymous divergence between alleles associated to the alternative mating types of the sequenced individual.**

MvSI gene ID ( $a_1$ )	MvSv gene ID ( $a_1$ )	$d_s$ in MvSI	$d_s$ in MvSv	Stratum	Putative function
MvSI-1064-A1-R4_A1g01509	MvSv-1253-A1-R1_MC03g02122	0	0	PAR	"AutoIPR: IPR008926:Ribonucleotide reductase R1 subunit, N-terminal; IPR013346:Ribonucleotide reductase, class I, alpha subunit; IPR013350:Ribonucleotide reductase alpha chain;"
MvSI-1064-A1-R4_A1g01508	MvSv-1253-A1-R1_MC03g02123	0	0	PAR	AutoIPR: IPR010482: Peroxin/Dysferlin domain;
MvSI-1064-A1-R4_A1g01502	MvSv-1253-A1-R1_MC03g02124	0,002787	0	PAR	AutoIPR: IPR009053: Prefoldin;
MvSI-1064-A1-R4_A1g01501	MvSv-1253-A1-R1_MC03g02125	0	0	PAR	AutoIPR: IPR001805: Adenosine kinase;
MvSI-1064-A1-R4_A1g01500	MvSv-1253-A1-R1_MC03g02126	0	0	PAR	AutoIPR: IPR016024: Armadillo-type fold; IPR027159: Nuclear cap-binding protein subunit 1;
MvSI-1064-A1-R4_A1g01498	MvSv-1253-A1-R1_MC03g02130	0	0	PAR	AutoIPR: IPR005821: Ion transport domain;
MvSI-1064-A1-R4_A1g01497	MvSv-1253-A1-R1_MC03g02134	0	0	PAR	AutoIPR: IPR011992: EF-hand domain pair;
MvSI-1064-A1-R4_A1g01492	MvSv-1253-A1-R1_MC03g02142	0	0	PAR	
MvSI-1064-A1-R4_A1g01491	MvSv-1253-A1-R1_MC03g02144	0	0	PAR	
MvSI-1064-A1-R4_A1g01489	MvSv-1253-A1-R1_MC03g02148	0	0	PAR	"AutoIPR: IPR006110: RNA polymerase, subunit omega/K/RPABC2;"
MvSI-1064-A1-R4_A1g01488	MvSv-1253-A1-R1_MC03g02149	0	0	PAR	"AutoIPR: IPR022591: Transcription initiation factor TFIID subunit 1, domain of unknown function;"
MvSI-1064-A1-R4_A1g01487	MvSv-1253-A1-R1_MC03g02150	0	0	PAR	AutoIPR: IPR021709: Protein of unknown function DUF3292;
MvSI-1064-A1-R4_A1g01478	MvSv-1253-A1-R1_MC03g02153	0	0	PAR	
MvSI-1064-A1-R4_A1g01486	MvSv-1253-A1-R1_MC03g02154	0	0	PAR	AutoIPR: IPR022440: Conserved hypothetical protein CHP03788;
MvSI-1064-A1-R4_A1g01485	MvSv-1253-A1-R1_MC03g02155	0	0	PAR	AutoIPR: IPR001237: 43kDa postsynaptic protein; IPR001623: DnaJ domain; IPR026901: DnaJ homologue subfamily C member 3;
MvSI-1064-A1-R4_A1g01474	MvSv-1253-A1-R1_MC03g02159	0	0	PAR	
MvSI-1064-A1-R4_A1g01476	MvSv-1253-A1-R1_MC03g02160	0	0	PAR	
MvSI-1064-A1-R4_A1g01477	MvSv-1253-A1-R1_MC03g02164	0	0	PAR	AutoIPR: IPR006994: Transcription factor 25;
MvSI-1064-A1-R4_A1g01484	MvSv-1253-A1-R1_MC03g02165	0	0	PAR	
MvSI-1064-A1-R4_A1g01473	MvSv-1253-A1-R1_MC03g02166	0	0	PAR	
MvSI-1064-A1-R4_A1g01472	MvSv-1253-A1-R1_MC03g02167	0	0	PAR	AutoIPR: IPR003162: Transcription initiation factor TAFII31; IPR009072: Histone-fold;
MvSI-1064-A1-R4_A1g01471	MvSv-1253-A1-R1_MC03g02168	0	0	PAR	AutoIPR: IPR007900: Transcription initiation factor TFIID component TAF4;
MvSI-1064-A1-R4_A1g01470	MvSv-1253-A1-R1_MC03g02170	0	0	PAR	
MvSI-1064-A1-R4_A1g01468	MvSv-1253-A1-R1_MC03g02172	0	0	PAR	
MvSI-1064-A1-R4_A1g01466	MvSv-1253-A1-R1_MC03g02174	0	0	PAR	
MvSI-1064-A1-R4_A1g01463	MvSv-1253-A1-R1_MC03g02180	0	0	PAR	
MvSI-1064-A1-R4_A1g01462	MvSv-1253-A1-R1_MC03g02181	0	0	PAR	"AutoIPR: IPR014624: Predicted 26S proteasome regulatory complex, non-ATPase subcomplex, subunit s5a, Plasmodium; IPR027040: Proteasome subunit Rpn10;"
MvSI-1064-A1-R4_A1g01461	MvSv-1253-A1-R1_MC03g02182	0	0	PAR	AutoIPR: IPR000456: Ribosomal protein L17;
MvSI-1064-A1-R4_A1g01460	MvSv-1253-A1-R1_MC03g02183	0	0	PAR	
MvSI-1064-A1-R4_A1g01458	MvSv-1253-A1-R1_MC03g02185	0	0,01838	PAR	"AutoIPR: IPR009010: Aspartate decarboxylase-like domain; IPR012754: DNA-directed RNA polymerase, subunit beta-prime; IPR015700: DNA-directed RNA polymerase III largest subunit;"
MvSI-1064-A1-R4_A1g01457	MvSv-1253-A1-R1_MC03g02186	0	0,02487	PAR	
MvSI-1064-A1-R4_A1g01456	MvSv-1253-A1-R1_MC03g02187	0	0	PAR	
MvSI-1064-A1-R4_A1g01455	MvSv-1253-A1-R1_MC03g02188	0	0	PAR	
MvSI-1064-A1-R4_A1g01454	MvSv-1253-A1-R1_MC03g02189	0	0	PAR	
MvSI-1064-A1-R4_A1g01453	MvSv-1253-A1-R1_MC03g02190	0	0	PAR	"AutoIPR: IPR005828: General substrate transporter; IPR016196: Major facilitator superfamily domain, general substrate transporter;"
MvSI-1064-A1-R4_A1g01451	MvSv-1253-A1-R1_MC03g02191	0	0	PAR	"AutoIPR: IPR011042: Six-bladed beta-propeller, TolB-like;"
MvSI-1064-A1-R4_A1g01450	MvSv-1253-A1-R1_MC03g02192	0	0	PAR	"AutoIPR: IPR009082: Signal transduction histidine kinase, homodimeric domain; IPR011006: CheY-like superfamily; IPR013655: PAS fold-3; IPR014285: Nitrogen fixation negative regulator NifL;"
MvSI-1064-A1-R4_A1g01447	MvSv-1253-A1-R1_MC03g02193	0	0	PAR	"AutoIPR: IPR017923: Transcription factor IIS, N-terminal;"
MvSI-1064-A1-R4_A1g01446	MvSv-1253-A1-R1_MC03g02194	0	0	PAR	
MvSI-1064-A1-R4_A1g01445	MvSv-1253-A1-R1_MC03g02195	0	0	PAR	
MvSI-1064-A1-R4_A1g01444	MvSv-1253-A1-R1_MC03g02196	0	0	PAR	"AutoIPR: IPR002226: Catalase haem-binding site; IPR010582: Catalase immune-responsive domain; IPR011614: Catalase core domain; IPR018028: Catalase, monofunctional, haem-containing; IPR020835: Catalase-like domain; IPR024708: Catalase active site;"

MvSI gene ID (a <sub>1</sub> )	MvSv gene ID (a <sub>1</sub> )	d <sub>s</sub> in MvSI	d <sub>s</sub> in MvSv	Stratum	Putative function
MvSI-1064-A1-R4_A1g01441	MvSv-1253-A1-R1_MC03g02203	0	0,00184	blue	"AutoIPR: IPR013083:Zinc finger, RING/FYVE/PHD-type;"
MvSI-1064-A1-R4_A1g01440	MvSv-1253-A1-R1_MC03g02204	0	0,008478	blue	"AutoIPR: IPR006285:Ubiquitin-like modifier-activating enzyme Atg7; IPR018075:Ubiquitin-activating enzyme, E1;"
MvSI-1064-A1-R4_A1g01163	MvSv-1253-A1-R1_MC03g02211	0,04476	0,009024	blue	"AutoIPR: IPR005225:Small GTP-binding protein domain; IPR006689:Small GTPase superfamily, ARF/SAR type; IPR027417:P-loop containing nucleoside triphosphate hydrolase;"
MvSI-1064-A1-R4_A1g01162	MvSv-1253-A1-R1_MC03g02212	0,02929	0,0101	blue	
MvSI-1064-A1-R4_A1g01161	MvSv-1253-A1-R1_MC03g02213	0,00892	0,01558	blue	AutoIPR: IPR023395:Mitochondrial carrier domain;
MvSI-1064-A1-R4_A1g01160	MvSv-1253-A1-R1_MC03g02214	0,002244	0,03034	blue	AutoIPR: IPR000717:Proteasome component (PCI) domain; IPR011990:Tetratricopeptide-like helical; IPR013143:PCI/PINT associated module;
MvSI-1064-A1-R4_A1g01159	MvSv-1253-A1-R1_MC03g02215	0,02168	0,0225	blue	AutoIPR: IPR004827:Basic-leucine zipper domain; IPR013910:Transcription factor PAP1; IPR023167:Yap1 redox domain;
MvSI-1064-A1-R4_A1g01154	MvSv-1253-A1-R1_MC03g02216	0,005776	0,0116	blue	
MvSI-1064-A1-R4_A1g00659	MvSv-1253-A1-R1_MC03g02219	0,04854	0,008824	blue	AutoIPR: IPR022751:Alpha-mannosyltransferase;
MvSI-1064-A1-R4_A1g00800	MvSv-1253-A1-R1_MC03g02223	0,03416	0,007407	blue	"AutoIPR: IPR002085:Alcohol dehydrogenase superfamily, zinc-type; IPR016040:NAD(P)-binding domain; IPR020843:Polyketide synthase, enoylreductase;"
MvSI-1064-A1-R4_A1g00801	MvSv-1253-A1-R1_MC03g02224	0	0,0222	blue	AutoIPR: IPR007146:Sas10/Utp3/C1D; IPR011082:Exosome-associated factor Rrp47/DNA strand repair C1D;
MvSI-1064-A1-R4_A1g00808	MvSv-1253-A1-R1_MC03g02226	0,6698	0,02125	blue	
<b>MvSI-1064-A1-R4_A1g00591</b>	<b>MvSv-1253-A1-R1_MC03g02227</b>	<b>0,3894</b>	<b>0,4844</b>	<b>blue</b>	<b>AutoIPR: IPR008422:Homeobox KN domain; IPR009057:Homeodomain-like; HD gene</b>
<b>MvSI-1064-A1-R4_A1g00592</b>	<b>MvSv-1253-A1-R1_MC03g02228</b>	<b>0,8506</b>	<b>0,669</b>	<b>blue</b>	<b>HD gene</b>
MvSI-1064-A1-R4_A1g00593	MvSv-1253-A1-R1_MC03g02229	0,04204	0,03775	blue	AutoIPR: IPR027073:5'-3' exoribonuclease;
MvSI-1064-A1-R4_A1g00772	MvSv-1253-A1-R1_MC03g02230	0,03975	0,005973	blue	"AutoIPR: IPR012677:Nucleotide-binding, alpha-beta plait;"
MvSI-1064-A1-R4_A1g00025	MvSv-1253-A1-R1_MC03g02235	0	0,01151	blue	"AutoIPR: IPR000743:Glycoside hydrolase, family 28; IPR011050:Pectin lyase fold/virulence factor;"
MvSI-1064-A1-R4_A1g00478	MvSv-1253-A1-R1_MC03g02236	0,09438	0,03602	blue	AutoIPR: IPR012336:Thioredoxin-like fold;
MvSI-1064-A1-R4_A1g00495	MvSv-1253-A1-R1_MC03g02237	0,06234	0,01041	blue	"AutoIPR: IPR012677:Nucleotide-binding, alpha-beta plait;"
MvSI-1064-A1-R4_A1g01431	MvSv-1253-A1-R1_MC03g02238	0,03381	0	blue	"AutoIPR: IPR001752:Kinesin, motor domain; IPR008984:SMAD/FHA domain; IPR011993:Plckstrin homology-like domain; IPR022164:Kinesin-like; IPR027417:P-loop containing nucleoside triphosphate hydrolase; IPR027640:Kinesin-like protein;"
MvSI-1064-A1-R4_A1g00997	MvSv-1253-A1-R1_MC03g02240	0,03601	0,01086	blue	
MvSI-1064-A1-R4_A1g00998	MvSv-1253-A1-R1_MC03g02241	0,03152	0,004411	blue	AutoIPR: IPR022801:Ribosomal protein S4/S9;
MvSI-1064-A1-R4_A1g00972	MvSv-1253-A1-R1_MC03g02242	0,01747	0,0052	blue	
MvSI-1064-A1-R4_A1g00971	MvSv-1253-A1-R1_MC03g02243	0,02441	0,02234	blue	AutoIPR: IPR000326:Phosphatidic acid phosphatase type 2/haloperoxidase;
MvSI-1064-A1-R4_A1g00970	MvSv-1253-A1-R1_MC03g02244	0,01335	0,01239	blue	AutoIPR: IPR002423:Chaperonin Cpn60/TCP-1; IPR027409:GroEL-like apical domain; IPR027410:TCP-1-like chaperonin intermediate domain; IPR027413:GroEL-like equatorial domain;
MvSI-1064-A1-R4_A1g00960	MvSv-1253-A1-R1_MC03g02248	0,03457	0	black	AutoIPR: IPR005024:Snf7;
MvSI-1064-A1-R4_A1g00958	MvSv-1253-A1-R1_MC03g02249	0,0192	0	black	AutoIPR: IPR004000:Actin-related protein;
MvSI-1064-A1-R4_A1g00957	MvSv-1253-A1-R1_MC03g02250	0,0344	0	black	AutoIPR: IPR013926:CGI121/TPRKB;
MvSI-1064-A1-R4_A1g00956	MvSv-1253-A1-R1_MC03g02251	0,05355	0	black	
MvSI-1064-A1-R4_A1g00955	MvSv-1253-A1-R1_MC03g02252	0,1164	0	black	
MvSI-1064-A1-R4_A1g00922	MvSv-1253-A1-R1_MC03g02253	0,04901	0	black	
MvSI-1064-A1-R4_A1g00921	MvSv-1253-A1-R1_MC03g02254	0,02868	0,003722	black	AutoIPR: IPR008942:ENTH/VHS;
MvSI-1064-A1-R4_A1g00920	MvSv-1253-A1-R1_MC03g02255	0,03374	0	black	
MvSI-1064-A1-R4_A1g00913	MvSv-1253-A1-R1_MC03g02256	0,06733	0,000814	black	AutoIPR: IPR012943:Spindle associated; IPR024545:Mto2p-binding domain;
MvSI-1064-A1-R4_A1g00912	MvSv-1253-A1-R1_MC03g02260	0	0	black	AutoIPR: IPR003782:Copper chaperone SCO1/SenC; IPR012336:Thioredoxin-like fold;
MvSI-1064-A1-R4_A1g00911	MvSv-1253-A1-R1_MC03g02261	0,01205	0	black	"AutoIPR: IPR008698:NADH:ubiquinone oxidoreductase, B18 subunit;"
MvSI-1064-A1-R4_A1g00794	MvSv-1253-A1-R1_MC03g02262	0,02735	0	black	AutoIPR: IPR009846:Splicing factor 3B subunit 5/RDS3 complex subunit 10;
MvSI-1064-A1-R4_A1g00644	MvSv-1253-A1-R1_MC03g02267	0,03817	0	black	"AutoIPR: IPR000571:Zinc finger, CCCH-type;"
MvSI-1064-A1-R4_A1g00645	MvSv-1253-A1-R1_MC03g02268	0,0177	0	black	AutoIPR: IPR001138:Zn(2)-C6 fungal-type DNA-binding domain;
MvSI-1064-A1-R4_A1g00646	MvSv-1253-A1-R1_MC03g02269	0,01013	0	black	AutoIPR: IPR012908:GPI inositol-deacylase PGAP1-like;
MvSI-1064-A1-R4_A1g00647	MvSv-1253-A1-R1_MC03g02270	0,07508	0	black	"AutoIPR: IPR007541:Uncharacterised protein family, basic secretory protein;"
MvSI-1064-A1-R4_A1g00648	MvSv-1253-A1-R1_MC03g02271	0,02229	0	black	"AutoIPR: IPR019339:CBF1-interacting co-repressor CIR, N-terminal domain; IPR022209:Pre-mRNA splicing factor;"
MvSI-1064-A1-R4_A1g01129	MvSv-1253-A1-R1_MC03g02273	0,05174	0	black	AutoIPR: IPR008978:HSP20-like chaperone;
MvSI-1064-A1-R4_A1g00311	MvSv-1253-A1-R1_MC03g02276	0,02807	0	black	AutoIPR: IPR015915:Kelch-type beta propeller;
MvSI-1064-A1-R4_A1g00312	MvSv-1253-A1-R1_MC03g02277	0,02654	0	black	"AutoIPR: IPR019191:Essential protein Yae1, N-terminal;"
MvSI-1064-A1-R4_A1g00886	MvSv-1253-A1-R1_MC03g02280	0,03442	0	black	
MvSI-1064-A1-R4_A1g00869	MvSv-1253-A1-R1_MC03g02282	0,04332	0	black	"AutoIPR: IPR006003:Carbohydrate kinase, FGGY-related;"



MvSI gene ID (a <sub>1</sub> )	MvSv gene ID (a <sub>1</sub> )	d <sub>s</sub> in MvSI	d <sub>s</sub> in MvSv	Stratum	Putative function
MvSI-1064-A1-R4_A1g01135	MvSv-1253-A1-R1_MC16g09520	0,05381	0	black	"AutoIPR: IPR019400:Peptidase C65, otubain;"
MvSI-1064-A1-R4_A1g01136	MvSv-1253-A1-R1_MC16g09519	0,0594	0	black	
MvSI-1064-A1-R4_A1g01137	MvSv-1253-A1-R1_MC16g09518	0,02957	0	black	
MvSI-1064-A1-R4_A1g00732	MvSv-1253-A1-R1_MC16g09516	0,002396	0	black	
MvSI-1064-A1-R4_A1g01407	MvSv-1253-A1-R1_MC16g09514	0,02369	0	black	
MvSI-1064-A1-R4_A1g01406	MvSv-1253-A1-R1_MC16g09512	0,002257	0	black	AutoIPR: IPR007757:MT-A70-like;
MvSI-1064-A1-R4_A1g01405	MvSv-1253-A1-R1_MC16g09511	0,01406	0	black	AutoIPR: IPR027450:Alpha-ketoglutarate-dependent dioxygenase AlkB-like;
MvSI-1064-A1-R4_A1g01404	MvSv-1253-A1-R1_MC16g09510	0,05471	0	black	
MvSI-1064-A1-R4_A1g00326	MvSv-1253-A1-R1_MC16g09508	0,04389	0	black	
MvSI-1064-A1-R4_A1g00327	MvSv-1253-A1-R1_MC16g09507	0,02724	0	black	
MvSI-1064-A1-R4_A1g00448	MvSv-1253-A1-R1_MC16g09503	0,06255	0	black	
MvSI-1064-A1-R4_A1g00749	MvSv-1253-A1-R1_MC16g09501	0,02509	0	black	
MvSI-1064-A1-R4_A1g00750	MvSv-1253-A1-R1_MC16g09500	0,002806	0	black	
MvSI-1064-A1-R4_A1g00751	MvSv-1253-A1-R1_MC16g09499	0,01061	0	black	
MvSI-1064-A1-R4_A1g00752	MvSv-1253-A1-R1_MC16g09498	0,07687	0	black	
MvSI-1064-A1-R4_A1g00755	MvSv-1253-A1-R1_MC16g09497	0,07173	0	black	
MvSI-1064-A1-R4_A1g00757	MvSv-1253-A1-R1_MC16g09496	0,05103	0	black	AutoIPR: IPR001461:Peptidase A1; IPR021109:Aspartic peptidase;
MvSI-1064-A1-R4_A1g00544	MvSv-1253-A1-R1_MC16g09494	0,007251	0	black	"AutoIPR: IPR010032:FAD-linked oxidoreductase; IPR016166:FAD-binding, type 2; IPR023595:L-gulonolactone/D-arabinono-1,4-lactone oxidase;"
MvSI-1064-A1-R4_A1g00543	MvSv-1253-A1-R1_MC16g09493	0	0	black	AutoIPR: IPR016460:Coatomer beta subunit (COPB1);
MvSI-1064-A1-R4_A1g00542	MvSv-1253-A1-R1_MC16g09492	0,002593	0	black	IPR026739:AP complex subunit beta;
MvSI-1064-A1-R4_A1g00541	MvSv-1253-A1-R1_MC16g09488	0,01331	0	black	AutoIPR: IPR001737:Ribosomal RNA adenine methylase transferase; IPR025814:18S rRNA dimethylase DIM1;
MvSI-1064-A1-R4_A1g00534	MvSv-1253-A1-R1_MC16g09487	0	0	black	AutoIPR: IPR024990:Anaphase-promoting complex subunit 1;
MvSI-1064-A1-R4_A1g00533	MvSv-1253-A1-R1_MC16g09486	0,07944	0	black	AutoIPR: IPR001557:L-lactate/malate dehydrogenase;
MvSI-1064-A1-R4_A1g00532	MvSv-1253-A1-R1_MC16g09485	0,0261	0	black	"AutoIPR: IPR011234:Fumarylacetoacetase, C-terminal-related;"
MvSI-1064-A1-R4_A1g00528	MvSv-1253-A1-R1_MC16g09482	0,02956	0	black	
MvSI-1064-A1-R4_A1g00525	MvSv-1253-A1-R1_MC16g09479	0,02199	0,01648	black	
MvSI-1064-A1-R4_A1g00522	MvSv-1253-A1-R1_MC16g09478	0,04318	0,01311	black	"AutoIPR: IPR011701:Major facilitator superfamily; IPR016196:Major facilitator superfamily domain, general substrate transporter;"
MvSI-1064-A1-R4_A1g00521	MvSv-1253-A1-R1_MC16g09477	0,002238	0,002248	black	AutoIPR: IPR013126:Heat shock protein 70 family;
MvSI-1064-A1-R4_A1g00520	MvSv-1253-A1-R1_MC16g09476	0	0,02046	black	
MvSI-1064-A1-R4_A1g00519	MvSv-1253-A1-R1_MC16g09475	0	0,01423	black	"AutoIPR: IPR011016:Zinc finger, RING-CH-type; IPR013083:Zinc finger, RING/FYVE/PHD-type;"
MvSI-1064-A1-R4_A1g00518	MvSv-1253-A1-R1_MC16g09474	0,002431	0,01173	black	AutoIPR: IPR011047:Quinonoprotein alcohol dehydrogenase-like superfamily; IPR015943:WD40/YVTN repeat-like-containing domain;
MvSI-1064-A1-R4_A1g00517	MvSv-1253-A1-R1_MC16g09473	0	0	black	
MvSI-1064-A1-R4_A1g00516	MvSv-1253-A1-R1_MC16g09472	0,02563	0,01878	black	"AutoIPR: IPR007087:Zinc finger, C2H2; IPR015940:Ubiquitin-associated/translation elongation factor EF1B, N-terminal, eukaryote; IPR017346:Uncharacterised conserved protein UCP037991, UAS/UBX;"
MvSI-1064-A1-R4_A1g00515	MvSv-1253-A1-R1_MC16g09471	0,01326	0,02447	black	"AutoIPR: IPR006293:DNA helicase, ATP-dependent, RecQ type, bacterial; IPR027417:P-loop containing nucleoside triphosphate hydrolase;"
MvSI-1064-A1-R4_A1g00514	MvSv-1253-A1-R1_MC16g09470	0,02296	0,02073	black	AutoIPR: IPR008936:Rho GTPase activation protein;
MvSI-1064-A1-R4_A1g00513	MvSv-1253-A1-R1_MC16g09469	0,0009531	0,01438	black	AutoIPR: IPR008521:Magnesium transporter NIPA;
MvSI-1064-A1-R4_A1g00512	MvSv-1253-A1-R1_MC16g09468	0	0,02597	black	"AutoIPR: IPR007305:Vesicle transport protein, Got1/SFT2-like;"
MvSI-1064-A1-R4_A1g00511	MvSv-1253-A1-R1_MC16g09467	0	0,03902	black	AutoIPR: IPR023395:Mitochondrial carrier domain;
MvSI-1064-A1-R4_A1g00510	MvSv-1253-A1-R1_MC16g09466	0	0,01317	black	"AutoIPR: IPR003754:Tetrapyrrole biosynthesis, uroporphyrinogen III synthase;"
MvSI-1064-A1-R4_A1g00509	MvSv-1253-A1-R1_MC16g09465	0,09326	0,02499	black	AutoIPR: IPR003817:Phosphatidylserine decarboxylase-related;
MvSI-1064-A1-R4_A1g00508	MvSv-1253-A1-R1_MC16g09464	0,1019	0,01383	black	"AutoIPR: IPR002938:Monooxygenase, FAD-binding; IPR020946:Flavin monooxygenase-like;"
MvSI-1064-A1-R4_A1g00506	MvSv-1253-A1-R1_MC16g09461	0,004505	0	black	AutoIPR: IPR002791:Domain of unknown function DUF89;
MvSI-1064-A1-R4_A1g00505	MvSv-1253-A1-R1_MC16g09460	0	0,01273	black	"AutoIPR: IPR005097:Saccharopine dehydrogenase / Homospermidine synthase; IPR007698:Alanine dehydrogenase/PNT, NAD(H)-binding domain; IPR007886:Alanine dehydrogenase/pyridine nucleotide transhydrogenase, N-terminal;"
MvSI-1064-A1-R4_A1g00503	MvSv-1253-A1-R1_MC16g09459	0,009071	0,01064	black	AutoIPR: IPR008942:ENTH/VHS; IPR018205:VHS subgroup;
MvSI-1064-A1-R4_A1g01351	MvSv-1253-A1-R1_MC16g09458	0,003649	0	black	AutoIPR: IPR001189:Manganese/iron superoxide dismutase;
MvSI-1064-A1-R4_A1g01350	MvSv-1253-A1-R1_MC16g09457	0,01323	0	black	AutoIPR: IPR024338:Stretch-activated cation channel Mid1;
MvSI-1064-A1-R4_A1g01348	MvSv-1253-A1-R1_MC16g09456	0,0105	0	black	AutoIPR: IPR000387:Protein-tyrosine/Dual specificity phosphatase; IPR008343:Mitogen-activated protein (MAP) kinase phosphatase; IPR024950:Dual specificity phosphatase;

MvSI gene ID (a <sub>1</sub> )	MvSv gene ID (a <sub>1</sub> )	d <sub>s</sub> in MvSI	d <sub>s</sub> in MvSv	Stratum	Putative function
MvSI-1064-A1-R4_A1g01347	MvSv-1253-A1-R1_MC16g09455	0,02952	0	black	"AutoIPR: IPR012954:BP28, C-terminal domain; IPR016024:Armadiolo-type fold; IPR022125:U3 small nucleolar RNA-associated protein 10;"
MvSI-1064-A1-R4_A1g01415	MvSv-1253-A1-R1_MC16g09444	0,02235	0	black	"AutoIPR: IPR008352:Mitogen-activated protein (MAP) kinase, p38; IPR020777:Tyrosine-protein kinase, neurotrophic receptor;"
MvSI-1064-A1-R4_A1g01393	MvSv-1253-A1-R1_MC16g09443	0,03313	0	black	
MvSI-1064-A1-R4_A1g01392	MvSv-1253-A1-R1_MC16g09442	0,02447	0	black	
MvSI-1064-A1-R4_A1g01479	MvSv-1253-A1-R1_MC16g09439	0	0	black	
MvSI-1064-A1-R4_A1g00211	MvSv-1253-A1-R1_MC12g07980	0,01398	0	black	"AutoIPR: IPR008010:Membrane protein, Tapt1/CMV receptor;"
MvSI-1064-A1-R4_A1g00779	MvSv-1253-A1-R1_MC12g07965	0,04375	0	black	AutoIPR: IPR005378:Vacuolar protein sorting-associated protein 35;
MvSI-1064-A1-R4_A1g00778	MvSv-1253-A1-R1_MC12g07964	0,04144	0	black	"AutoIPR: IPR015408:Zinc finger, Mcm10/DnaG-type;"
MvSI-1064-A1-R4_A1g00241	MvSv-1253-A1-R1_MC12g07960	0,03193	0	black	
MvSI-1064-A1-R4_A1g00242	MvSv-1253-A1-R1_MC12g07959	0,033	0	black	AutoIPR: IPR007653:Signal peptidase 22kDa subunit;
MvSI-1064-A1-R4_A1g00245	MvSv-1253-A1-R1_MC12g07956	0,07408	0	black	
MvSI-1064-A1-R4_A1g00269	MvSv-1253-A1-R1_MC12g07952	0,04395	0	black	"AutoIPR: IPR003111:Peptidase S16, lon N-terminal; IPR014252:Sporulation protease LonC; IPR014721:Ribosomal protein S5 domain 2-type fold, subgroup; IPR015947:PUA-like domain; IPR027065:Lon protease; IPR027417:P-loop containing nucleoside triphosphate hydrolase;"
MvSI-1064-A1-R4_A1g00270	MvSv-1253-A1-R1_MC12g07950	0,04929	0	black	"AutoIPR: IPR006219:DHAP synthase, class 1; IPR013785:Aldolase-type TIM barrel;"
MvSI-1064-A1-R4_A1g00292	MvSv-1253-A1-R1_MC12g07947	0,02908	0	black	AutoIPR: IPR001394:Ubiquitin carboxyl-terminal hydrolases family 2; IPR019955:Ubiquitin supergroup;
MvSI-1064-A1-R4_A1g00290	MvSv-1253-A1-R1_MC12g07946	0,02661	0	black	AutoIPR: IPR027815:Domain of unknown function DUF4463;
MvSI-1064-A1-R4_A1g00693	MvSv-1253-A1-R1_MC12g07943	0	0	black	AutoIPR: IPR002042:Uricase;
MvSI-1064-A1-R4_A1g00694	MvSv-1253-A1-R1_MC12g07942	0,01544	0	black	"AutoIPR: IPR002100:Transcription factor, MADS-box;"
MvSI-1064-A1-R4_A1g00719	MvSv-1253-A1-R1_MC12g07935	0,1062	0	black	"AutoIPR: IPR004589:DNA helicase, ATP-dependent, RecQ type; IPR018973:DEAD/DEAH-box helicase, putative; IPR027417:P-loop containing nucleoside triphosphate hydrolase;"
MvSI-1064-A1-R4_A1g00718	MvSv-1253-A1-R1_MC12g07934	0,03731	0	black	"AutoIPR: IPR007577:Glycosyltransferase, DXD sugar-binding motif;"
MvSI-1064-A1-R4_A1g00254	MvSv-1253-A1-R1_MC12g07933	0,03514	0	black	AutoIPR: IPR003386:Lecithin:cholesterol/phospholipid:diacylglycerol acyltransferase;
MvSI-1064-A1-R4_A1g00359	MvSv-1253-A1-R1_MC12g07932	0,02437	0	black	AutoIPR: IPR020831:Glyceraldehyde/Erythrose phosphate dehydrogenase family;
MvSI-1064-A1-R4_A1g00281	MvSv-1253-A1-R1_MC12g07931	0,04624	0	black	"AutoIPR: IPR003959:ATPase, AAA-type, core; IPR027417:P-loop containing nucleoside triphosphate hydrolase;"
MvSI-1064-A1-R4_A1g00280	MvSv-1253-A1-R1_MC12g07930	0,02891	0	black	"AutoIPR: IPR008973:C2 calcium/lipid-binding domain, CaLB; IPR019558:Mammalian uncoordinated homology 13, subgroup, domain 2;"
MvSI-1064-A1-R4_A1g00433	MvSv-1253-A1-R1_MC12g07928	0,02267	0,001815	black	"AutoIPR: IPR022042:snRNA-activating protein complex, subunit 3;"
MvSI-1064-A1-R4_A1g00432	MvSv-1253-A1-R1_MC12g07927	0,1063	0	black	AutoIPR: IPR009071:High mobility group box domain;
MvSI-1064-A1-R4_A1g00349	MvSv-1253-A1-R1_MC12g07926	0,09006	0	black	AutoIPR: IPR000791:GPR1/FUN34/yaaH;
MvSI-1064-A1-R4_A1g00350	MvSv-1253-A1-R1_MC12g07925	0,02792	0	black	AutoIPR: IPR015915:Kelch-type beta propeller;
MvSI-1064-A1-R4_A1g00351	MvSv-1253-A1-R1_MC12g07924	0,03002	0,0007955	black	"AutoIPR: IPR007320:Programmed cell death protein 2, C-terminal; IPR015940:Ubiquitin-associated/translation elongation factor EF1B, N-terminal, eukaryote; IPR017346:Uncharacterised conserved protein UCPO37991, UAS/UBX;"
MvSI-1064-A1-R4_A1g00352	MvSv-1253-A1-R1_MC12g07923	0,03106	0,006247	black	"AutoIPR: IPR016656:Transcription initiation factor TFIIE, beta subunit;"
MvSI-1064-A1-R4_A1g01272	MvSv-1253-A1-R1_MC12g07922	0,01626	0	black	AutoIPR: IPR008417:B-cell receptor-associated 31-like;
MvSI-1064-A1-R4_A1g01271	MvSv-1253-A1-R1_MC12g07921	0,01543	0	black	AutoIPR: IPR008011:Complex 1 LYR protein;
MvSI-1064-A1-R4_A1g01173	MvSv-1253-A1-R1_MC12g07920	0,03657	0,007907	black	"AutoIPR: IPR012677:Nucleotide-binding, alpha-beta plait;"
MvSI-1064-A1-R4_A1g01174	MvSv-1253-A1-R1_MC12g07919	0,02616	0	black	AutoIPR: IPR014729:Rossmann-like alpha/beta/alpha sandwich fold;
MvSI-1064-A1-R4_A1g01220	MvSv-1253-A1-R1_MC12g07918	0,09423	0	black	"AutoIPR: IPR001781:Zinc finger, LIM-type;"
MvSI-1064-A1-R4_A1g01219	MvSv-1253-A1-R1_MC12g07917	0,06024	0	black	AutoIPR: IPR018803:Stress-responsive protein Ish1;
MvSI-1064-A1-R4_A1g00165	MvSv-1253-A1-R1_MC12g07913	0,03698	0,00352	black	"AutoIPR: IPR006153:Cation/H <sup>+</sup> exchanger; IPR018422:Cation/H <sup>+</sup> exchanger, CPA1 family;"
MvSI-1064-A1-R4_A1g00157	MvSv-1253-A1-R1_MC12g07911	0,03635	0,01021	black	"AutoIPR: IPR016196:Major facilitator superfamily domain, general substrate transporter;"
MvSI-1064-A1-R4_A1g00155	MvSv-1253-A1-R1_MC12g07910	0,04423	0	black	AutoIPR: IPR001645:Folypolyglutamate synthetase;
MvSI-1064-A1-R4_A1g00153	MvSv-1253-A1-R1_MC12g07908	0,03147	0,003478	black	"AutoIPR: IPR010222:RNA helicase, ATP-dependent DEAH box, HrpA-type; IPR022307:DEAD/DEAH-box helicase, putative, actinobacteria; IPR027417:P-loop containing nucleoside triphosphate hydrolase;"
MvSI-1064-A1-R4_A1g01212	MvSv-1253-A1-R1_MC12g07907	0,03602	0,002255	black	"AutoIPR: IPR001208:Mini-chromosome maintenance, DNA-dependent ATPase; IPR027417:P-loop containing nucleoside triphosphate hydrolase; IPR027925:MCM N-terminal domain;"
MvSI-1064-A1-R4_A1g01184	MvSv-1253-A1-R1_MC12g07904	0,03576	0,006311	black	AutoIPR: IPR001708:Membrane insertase OXA1/ALB3/YidC;
MvSI-1064-A1-R4_A1g01189	MvSv-1253-A1-R1_MC12g07903	0,02386	0	black	"AutoIPR: IPR011598:Myc-type, basic helix-loop-helix (bHLH) domain;"
MvSI-1064-A1-R4_A1g00016	MvSv-1253-A1-R1_MC12g07897	0	0	black	
MvSI-1064-A1-R4_A1g00017	MvSv-1253-A1-R1_MC12g07896	0	0,01191	black	"AutoIPR: IPR011701:Major facilitator superfamily; IPR016196:Major facilitator superfamily domain, general substrate transporter;"
MvSI-1064-A1-R4_A1g00144	MvSv-1253-A1-R1_MC12g07894	0,05262	0	black	"AutoIPR: IPR006789:ARP2/3 complex, 16kDa subunit (p16-Arc);"

MvSI gene ID (a <sub>1</sub> )	MvSv gene ID (a <sub>1</sub> )	d <sub>s</sub> in MvSI	d <sub>s</sub> in MvSv	Stratum	Putative function
MvSI-1064-A1-R4_A1g01389	MvSv-1253-A1-R1_MC12g07893	0,03083	0	black	
MvSI-1064-A1-R4_A1g01388	MvSv-1253-A1-R1_MC12g07892	0,04141	0,00527	black	AutoIPR: IPR005343:Nucleolar complex protein 2;
MvSI-1064-A1-R4_A1g01385	MvSv-1253-A1-R1_MC12g07891	0,02823	0,001958	black	"AutoIPR: IPR009668:RNA polymerase I associated factor, A49-like;"
MvSI-1064-A1-R4_A1g01384	MvSv-1253-A1-R1_MC12g07890	0,0293	0,00419	black	AutoIPR: IPR001060:FCH domain; IPR018808:Muscin C-terminal mu homology domain;
MvSI-1064-A1-R4_A1g01369	MvSv-1253-A1-R1_MC12g07886	0,02333	0,001781	black	AutoIPR: IPR005792:Protein disulphide isomerase;
MvSI-1064-A1-R4_A1g01368	MvSv-1253-A1-R1_MC12g07885	0,02638	0,005801	black	IPR012336:Thioredoxin-like fold;
MvSI-1064-A1-R4_A1g01367	MvSv-1253-A1-R1_MC12g07884	0,01551	0,00982	black	AutoIPR: IPR019021:Methyl methanesulphonate-sensitivity protein 22;
MvSI-1064-A1-R4_A1g01366	MvSv-1253-A1-R1_MC12g07883	0,0362	0,00599	black	AutoIPR: IPR006565:Bromodomain transcription factor;
MvSI-1064-A1-R4_A1g00845	MvSv-1253-A1-R1_MC12g07881	0,03648	0,001712	black	"AutoIPR: IPR008937:Ras guanine nucleotide exchange factor; IPR023578:Ras guanine nucleotide exchange factor, domain;"
MvSI-1064-A1-R4_A1g00844	MvSv-1253-A1-R1_MC12g07880	0,008035	0	black	
MvSI-1064-A1-R4_A1g00843	MvSv-1253-A1-R1_MC12g07879	0,0116	0,002791	black	AutoIPR: IPR002060:Squalene/phytoene synthase;
MvSI-1064-A1-R4_A1g00842	MvSv-1253-A1-R1_MC12g07878	0,06322	0	black	"AutoIPR: IPR002018:Carboxylesterase, type B;"
MvSI-1064-A1-R4_A1g00840	MvSv-1253-A1-R1_MC12g07874	0,04014	0,004762	black	
MvSI-1064-A1-R4_A1g00832	MvSv-1253-A1-R1_MC12g07872	0,02415	0,001244	black	AutoIPR: IPR011009:Protein kinase-like domain;
MvSI-1064-A1-R4_A1g00575	MvSv-1253-A1-R1_MC12g07870	0,02905	0,003192	black	AutoIPR: IPR002058:PAP/25A-associated; IPR002934:Nucleotidyl transferase domain;
MvSI-1064-A1-R4_A1g00569	MvSv-1253-A1-R1_MC12g07869	0,1057	0,04197	black	
MvSI-1064-A1-R4_A1g00093	MvSv-1253-A1-R1_MC12g07864	0,04154	0,006273	black	
MvSI-1064-A1-R4_A1g00186	MvSv-1253-A1-R1_MC12g07863	0,03298	0,01127	black	AutoIPR: IPR005013:Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit WBP1;
MvSI-1064-A1-R4_A1g00185	MvSv-1253-A1-R1_MC12g07862	0,02174	0,002111	black	AutoIPR: IPR011047:Quinonprotein alcohol dehydrogenase-like superfamily; IPR026895:ER membrane protein complex subunit 1;
MvSI-1064-A1-R4_A1g00184	MvSv-1253-A1-R1_MC12g07861	0	0	black	"AutoIPR: IPR002222:Ribosomal protein S19/S15; IPR023575:Ribosomal protein S19, superfamily;"
MvSI-1064-A1-R4_A1g01038	MvSv-1253-A1-R1_MC12g07859	0,01992	0	black	AutoIPR: IPR000836:Phosphoribosyltransferase domain; IPR023031:Orotate phosphoribosyltransferase;
MvSI-1064-A1-R4_A1g01021	MvSv-1253-A1-R1_MC12g07856	0,02939	NA	black	
MvSI-1064-A1-R4_A1g00175	MvSv-1253-A1-R1_MC12g07851	0,1428	0,06338	purple	AutoIPR: IPR004147:UbiB domain;
MvSI-1064-A1-R4_A1g00279	MvSv-1253-A1-R1_MC12g07848	0,09067	NA	purple	
MvSI-1064-A1-R4_A1g00854	MvSv-1253-A1-R1_MC12g07846	0,02438	0,04877	purple	AutoIPR: IPR016024:Armado-type fold; IPR027140:Importin subunit beta;
MvSI-1064-A1-R4_A1g01227	MvSv-1253-A1-R1_MC12g07836	0,03955	0	purple	AutoIPR: IPR021138:Ribosomal protein L18a; IPR023573:Ribosomal protein L18a/LX;
MvSI-1064-A1-R4_A1g00230	MvSv-1253-A1-R1_MC12g07828	0,109	0,1148	purple	"AutoIPR: IPR006818:Histone chaperone, ASF1-like;"
MvSI-1064-A1-R4_A1g01017	MvSv-1253-A1-R1_MC12g07823	0,02769	0,07915	purple	AutoIPR: IPR007318:Phospholipid methyltransferase;
MvSI-1064-A1-R4_A1g00726	MvSv-1253-A1-R1_MC12g07815	0,1145	0,07136	purple	AutoIPR: IPR009060:UBA-like; IPR014764:Defective-in-cullin neddylation protein;
MvSI-1064-A1-R4_A1g00624	MvSv-1253-A1-R1_MC12g07792	0,08473	0,03743	purple	"AutoIPR: IPR016071:Staphylococcal nuclease (SNase-like), OB-fold;"
MvSI-1064-A1-R4_A1g00126	MvSv-1253-A1-R1_MC12g07780	0,4265	0,1397	purple	AutoIPR: IPR019166:Apolipoprotein O;
MvSI-1064-A1-R4_A1g00306	MvSv-1253-A1-R1_MC12g07758	NA	NA	orange	
MvSI-1064-A1-R4_A1g00307	MvSv-1253-A1-R1_MC12g07757	0,1172	0,004983	orange	AutoIPR: IPR000702:Ribosomal protein L6;
MvSI-1064-A1-R4_A1g00441	MvSv-1253-A1-R1_MC12g07755	0,03188	0,004631	orange	"AutoIPR: IPR009025:DNA-directed RNA polymerase, RBP11-like dimerisation domain; IPR01262:DNA-directed RNA polymerase, insert domain;"
MvSI-1064-A1-R4_A1g01253	MvSv-1253-A1-R1_MC12g07753	0,02466	0	orange	"AutoIPR: IPR009053:Prefoldin; IPR016655:Prefoldin, subunit 3;"
MvSI-1064-A1-R4_A1g01252	MvSv-1253-A1-R1_MC12g07752	0,07508	0,006777	orange	
MvSI-1064-A1-R4_A1g00669	MvSv-1253-A1-R1_MC12g07750	0,03217	0,002827	orange	"AutoIPR: IPR001876:Zinc finger, RanBP2-type; IPR007286:EAP30; IPR011991:Winged helix-turn-helix DNA-binding domain; IPR021648:Vacuolar protein sorting protein 36, GLUE domain;"
MvSI-1064-A1-R4_A1g01261	MvSv-1253-A1-R1_MC12g07749	0,03357	0,0004392	orange	"AutoIPR: IPR007309:B-block binding subunit of TFIIC; IPR017956:AT hook, DNA-binding motif;"
MvSI-1064-A1-R4_A1g00407	MvSv-1253-A1-R1_MC12g07746	0,09576	0	red	AutoIPR: IPR000999:Ribonuclease III domain;
MvSI-1064-A1-R4_A1g00406	MvSv-1253-A1-R1_MC12g07745	0	0,01146	red	AutoIPR: IPR000307:Ribosomal protein S16; IPR023803:Ribosomal protein S16 domain;
MvSI-1064-A1-R4_A1g00405	MvSv-1253-A1-R1_MC12g07744	0,003614	0	red	AutoIPR: IPR015943:WD40/YVTN repeat-like-containing domain;
MvSI-1064-A1-R4_A1g00404	MvSv-1253-A1-R1_MC12g07742	0	0	red	AutoIPR: IPR013216:Methyltransferase type 11;
MvSI-1064-A1-R4_A1g00402	MvSv-1253-A1-R1_MC12g07740	0,01347	0,0006517	red	AutoIPR: IPR025279:Stress response protein NST1;
MvSI-1064-A1-R4_A1g00401	MvSv-1253-A1-R1_MC12g07739	0	0	red	AutoIPR: IPR002942:RNA-binding S4 domain; IPR022801:Ribosomal protein S4/S9;
MvSI-1064-A1-R4_A1g00399	MvSv-1253-A1-R1_MC12g07737	0	0	red	AutoIPR: IPR001147:Ribosomal protein L21e;
MvSI-1064-A1-R4_A1g00398	MvSv-1253-A1-R1_MC12g07736	0	0	red	"AutoIPR: IPR019711:ATPase, F0 complex, subunit H;"
MvSI-1064-A1-R4_A1g00397	MvSv-1253-A1-R1_MC12g07735	0	0	red	"AutoIPR: IPR007648:ATPase inhibitor, IATP, mitochondria;"
MvSI-1064-A1-R4_A1g00396	MvSv-1253-A1-R1_MC12g07734	0	0	red	"AutoIPR: IPR007482:Protein-tyrosine phosphatase-like, PTPLA;"
MvSI-1064-A1-R4_A1g00394	MvSv-1253-A1-R1_MC12g07732	0	0	red	

MvSI gene ID (a <sub>1</sub> )	MvSv gene ID (a <sub>1</sub> )	d <sub>s</sub> in MvSI	d <sub>s</sub> in MvSv	Stratum	Putative function
MvSI-1064-A1-R4_A1g00392	MvSv-1253-A1-R1_MC12g07731	0,03094	0	red	
MvSI-1064-A1-R4_A1g00391	MvSv-1253-A1-R1_MC12g07730	0,006953	0	red	"AutoIPR: IPR001392:Clathrin adaptor, mu subunit; IPR027059:Coatomer delta subunit;"
MvSI-1064-A1-R4_A1g00390	MvSv-1253-A1-R1_MC12g07729	0,008075	0	red	
MvSI-1064-A1-R4_A1g00384	MvSv-1253-A1-R1_MC12g07725	0,2181	0		
MvSI-1064-A1-R4_A1g00374	MvSv-1253-A1-R1_MC12g07723	0,001469	0,008644	red	AutoIPR: IPR009075:Acyl-CoA dehydrogenase/oxidase C-terminal; IPR012258:Acyl-CoA oxidase;
MvSI-1064-A1-R4_A1g00373	MvSv-1253-A1-R1_MC12g07722	0	0,02706	red	"AutoIPR: IPR020869:Exosome complex exonuclease 2, probable; IPR027408:PNase/RNase PH domain;"
MvSI-1064-A1-R4_A1g00372	MvSv-1253-A1-R1_MC12g07721	0,002337	0,002394	red	"AutoIPR: IPR000924:Glutamyl/glutaminyl-tRNA synthetase, class Ib; IPR010987:Glutathione S-transferase, C-terminal-like;"
MvSI-1064-A1-R4_A1g00370	MvSv-1253-A1-R1_MC12g07719	0	0,01147	red	"AutoIPR: IPR007246:Gaa1-like, GPI transamidase component;"
MvSI-1064-A1-R4_A1g00369	MvSv-1253-A1-R1_MC12g07718	0,003175	0,001599	red	AutoIPR: IPR016161:Aldehyde/histidinol dehydrogenase;
MvSI-1064-A1-R4_A1g00368	MvSv-1253-A1-R1_MC12g07717	0,007024	0,002525	red	AutoIPR: IPR003358:tRNA (guanine-N-7) methyltransferase;
MvSI-1064-A1-R4_A1g00367	MvSv-1253-A1-R1_MC12g07716	0,003783	0	red	"AutoIPR: IPR011701:Major facilitator superfamily; IPR016196:Major facilitator superfamily domain, general substrate transporter;"
MvSI-1064-A1-R4_A1g00113	MvSv-1253-A1-R1_MC12g07715	0,007056	0	green	"AutoIPR: IPR002290:Serine/threonine- / dual specificity protein kinase, catalytic domain; IPR020777:Tyrosine-protein kinase, neurotrophic receptor;"
MvSI-1064-A1-R4_A1g00110	MvSv-1253-A1-R1_MC12g07712	0,001184	0	green	
MvSI-1064-A1-R4_A1g00109	MvSv-1253-A1-R1_MC12g07711	0,004364	0,04003	green	"AutoIPR: IPR012677:Nucleotide-binding, alpha-beta plait;"
MvSI-1064-A1-R4_A1g00080	MvSv-1253-A1-R1_MC12g07706	0	0	PAR	
MvSI-1064-A1-R4_A1g00078	MvSv-1253-A1-R1_MC12g07704	0	0	PAR	AutoIPR: IPR002672:Ribosomal protein L28e;
MvSI-1064-A1-R4_A1g00077	MvSv-1253-A1-R1_MC12g07703	0,009973	0,03064	PAR	AutoIPR: IPR004087:K Homology domain;
MvSI-1064-A1-R4_A1g00076	MvSv-1253-A1-R1_MC12g07702	0	0,002377	PAR	
MvSI-1064-A1-R4_A1g00075	MvSv-1253-A1-R1_MC12g07701	0	0,004902	PAR	"AutoIPR: IPR008928:Six-hairpin glycosidase-like; IPR010905:Glycosyl hydrolase, family 88;"
MvSI-1064-A1-R4_A1g00074	MvSv-1253-A1-R1_MC12g07700	0	0,0205	PAR	AutoIPR: IPR002791:Domain of unknown function DUF89;
MvSI-1064-A1-R4_A1g00073	MvSv-1253-A1-R1_MC12g07699	0	0,01967	PAR	AutoIPR: IPR013763:Cyclin-like;
MvSI-1064-A1-R4_A1g00072	MvSv-1253-A1-R1_MC12g07698	0	0,01032	PAR	AutoIPR: IPR002123:Phospholipid/glycerol acyltransferase;
MvSI-1064-A1-R4_A1g00071	MvSv-1253-A1-R1_MC12g07697	0	0	PAR	"AutoIPR: IPR012951:Berberine/berberine-like; IPR016166:FAD-binding, type 2;"
MvSI-1064-A1-R4_A1g00070	MvSv-1253-A1-R1_MC12g07696	0	0	PAR	AutoIPR: IPR009449:GDPGTP exchange factor Sec2p;
MvSI-1064-A1-R4_A1g00069	MvSv-1253-A1-R1_MC12g07695	0	0,0039	PAR	"AutoIPR: IPR020461:Tyrosine-protein kinase, neurotrophic receptor, type 1;"
MvSI-1064-A1-R4_A1g00068	MvSv-1253-A1-R1_MC12g07694	0	0	PAR	
MvSI-1064-A1-R4_A1g00067	MvSv-1253-A1-R1_MC12g07693	0	0,0214	PAR	
MvSI-1064-A1-R4_A1g00065	MvSv-1253-A1-R1_MC12g07691	0	0	PAR	AutoIPR: IPR012696:Phosphonate metabolism PhnM;
MvSI-1064-A1-R4_A1g00064	MvSv-1253-A1-R1_MC12g07683	0	0	PAR	
MvSI-1064-A1-R4_A1g00063	MvSv-1253-A1-R1_MC12g07682	0	0	PAR	"AutoIPR: IPR009244:Mediator complex, subunit Med7;"
MvSI-1064-A1-R4_A1g00062	MvSv-1253-A1-R1_MC12g07681	0	0	PAR	"AutoIPR: IPR012677:Nucleotide-binding, alpha-beta plait;"
MvSI-1064-A1-R4_A1g00061	MvSv-1253-A1-R1_MC12g07680	0	0	PAR	AutoIPR: IPR001806:Small GTPase superfamily; IPR002041:Ran GTPase; IPR005225:Small GTP-binding protein domain; IPR027417:P-loop containing nucleoside triphosphate hydrolase;
MvSI-1064-A1-R4_A1g00057	MvSv-1253-A1-R1_MC12g07678	0	0	PAR	
MvSI-1064-A1-R4_A1g00056	MvSv-1253-A1-R1_MC12g07677	0	0	PAR	"AutoIPR: IPR005828:General substrate transporter; IPR016196:Major facilitator superfamily domain, general substrate transporter;"
MvSI-1064-A1-R4_A1g00055	MvSv-1253-A1-R1_MC12g07676	0	0	PAR	AutoIPR: IPR013960:DASH complex subunit Duo1;role in spindle attachment, chromosome segregation and spindle stability.
MvSI-1064-A1-R4_A1g00054	MvSv-1253-A1-R1_MC12g07675	0	0	PAR	AutoIPR: IPR001269:tRNA-dihydrouridine synthase; IPR013785:Aldolase-type TIM barrel;
MvSI-1064-A1-R4_A1g00053	MvSv-1253-A1-R1_MC12g07674	0	0	PAR	AutoIPR: IPR012336:Thioredoxin-like fold;
MvSI-1064-A1-R4_A1g00052	MvSv-1253-A1-R1_MC12g07673	0	0	PAR	
MvSI-1064-A1-R4_A1g00051	MvSv-1253-A1-R1_MC12g07672	0	0	PAR	AutoIPR: IPR000760:Inositol monophosphatase;
MvSI-1064-A1-R4_A1g00050	MvSv-1253-A1-R1_MC12g07671	0	0	PAR	"AutoIPR: IPR002092:DNA-directed RNA polymerase, phage-type; IPR024075:DNA-directed RNA polymerase, helix hairpin domain;"
MvSI-1064-A1-R4_A1g00049	MvSv-1253-A1-R1_MC12g07670	0	0	PAR	AutoIPR: IPR002761:DUF71 domain; IPR013813:Endoribonuclease L-PSP/chorismate mutase-like;
MvSI-1064-A1-R4_A1g00048	MvSv-1253-A1-R1_MC12g07669	0	0	PAR	
MvSI-1064-A1-R4_A1g00047	MvSv-1253-A1-R1_MC12g07668	0	0	PAR	
MvSI-1064-A1-R4_A1g00046	MvSv-1253-A1-R1_MC12g07667	0	0	PAR	"AutoIPR: IPR011330:Glycoside hydrolase/deacetylase, beta/alpha-barrel;"
MvSI-1064-A1-R4_A1g00044	MvSv-1253-A1-R1_MC12g07665	0	0	PAR	"AutoIPR: IPR011330:Glycoside hydrolase/deacetylase, beta/alpha-barrel;"
MvSI-1064-A1-R4_A1g00040	MvSv-1253-A1-R1_MC12g07664	0	0	PAR	"AutoIPR: IPR011330:Glycoside hydrolase/deacetylase, beta/alpha-barrel;"
MvSI-1064-A1-R4_A1g00039	MvSv-1253-A1-R1_MC12g07663	0	0	PAR	
MvSI-1064-A1-R4_A1g00038	MvSv-1253-A1-R1_MC12g07661	0	0	PAR	

MvSI gene ID (a <sub>1</sub> )	MvSv gene ID (a <sub>1</sub> )	d <sub>s</sub> in MvSI	d <sub>s</sub> in MvSv	Stratum	Putative function
MvSI-1064-A1-R4_A1g00037	MvSv-1253-A1-R1_MC12g07660	0	0	PAR	"AutoIPR: IPR001567:Peptidase M3A/M3B; IPR024077:Neurolysin/Thimet oligopeptidase, domain 2; IPR024079:Metallopeptidase, catalytic domain;"
MvSI-1064-A1-R4_A1g00036	MvSv-1253-A1-R1_MC12g07659	0	0	PAR	AutoIPR: IPR006565:Bromodomain transcription factor;This recognition is often a prerequisite for protein-histone association and chromatin remodeling.
MvSI-1064-A1-R4_A1g00035	MvSv-1253-A1-R1_MC12g07658	0	0	PAR	
MvSI-1064-A1-R4_A1g00033	MvSv-1253-A1-R1_MC12g07656	0	0	PAR	
MvSI-1064-A1-R4_A1g00032	MvSv-1253-A1-R1_MC12g07655	0	0	PAR	"AutoIPR: IPR002035:von Willebrand factor, type A;"
MvSI-1064-A1-R4_A1g00031	MvSv-1253-A1-R1_MC12g07654	0	0	PAR	"AutoIPR: IPR003083:S-crystallin; IPR017933:Glutathione S-transferase/chloride channel, C-terminal;"
MvSI-1064-A1-R4_A1g00030	MvSv-1253-A1-R1_MC12g07653	0	0	PAR	
MvSI-1064-A1-R4_A1g00029	MvSv-1253-A1-R1_MC12g07652	0	0	PAR	
MvSI-1064-A1-R4_A1g00009	MvSv-1253-A1-R1_MC12g07644	0	0	PAR	"AutoIPR: IPR006785:Peroxisome membrane anchor protein Pex14p, N-terminal; IPR025655:Peroxisomal membrane protein 14;"
MvSI-1064-A1-R4_A1g00008	MvSv-1253-A1-R1_MC12g07643	0	0	PAR	AutoIPR: IPR013087:Zinc finger C2H2-type/integrase DNA-binding domain;
MvSI-1064-A1-R4_A1g00007	MvSv-1253-A1-R1_MC12g07642	0	0	PAR	AutoIPR: IPR008984:SMAD/FHA domain;
MvSI-1064-A1-R4_A1g00006	MvSv-1253-A1-R1_MC12g07641	0	0	PAR	"AutoIPR: IPR003204:Cytochrome c oxidase, subunit Va/VI;"
MvSI-1064-A1-R4_A1g00005	MvSv-1253-A1-R1_MC12g07640	0	0	PAR	
MvSI-1064-A1-R4_A1g00004	MvSv-1253-A1-R1_MC12g07639	0	0	PAR	AutoIPR: IPR003121:SWIB/MDM2 domain;

**Table S4. Raw data summary and assembly statistics for the genomes generated in this study.**

sample	Raw data summary							Assembly statistics										
	# Smart cell	# Subreads	Max Subread Length (bp)	N50 (bp)	Mean (bp)	Median (bp)	BP	Accession numbers	# Contigs	Min length (bp)	Max length (bp)	N50 BP	L50 (#contigs)	N90 (bp)	L90 (#contigs)	mean length (bp)	Median Length (bp)	Assembly size (bp)
<i>Microbotryum lagerheimii</i> 1253 A1	4	589978	41394	9526	7290	7059	4 301 220 311	ERS1013677	42	6001	3235273	1584903	7	462124	17	614444	172652	25 806 628
<i>Microbotryum lagerheimii</i> 1253 A2	4	587933	43134	9660	7394	7202	4 347 011 348	ERS1013678	37	7917	3230111	1585144	7	534025	17	693725	474693	25 667 824
<i>Microbotryum violaceum</i> s. str. 1249 A1	4	548702	44168	11951	8759	8361	4 806 182 398	ERS1013671	166	5956	2602331	1183707	10	65951	47	191449	31698	31 780 530
<i>Microbotryum violaceum</i> s. str. 1249 A2	4	456117	46941	13250	9516	8962	4 340 608 458	ERS1013672	61	8455	2644084	1643691	7	606103	17	456993	30510	27 876 544
<i>Microbotryum lychnidis-dioicae</i> Lamole A1							Raw data from Badouin et al. 2015 (12)	ERS1013679	48	12190	3412169	1736850	6	648002	16	622941	44679	29 901 156
<i>Microbotryum lychnidis-dioicae</i> Lamole A2							Raw data from Badouin et al. 2015 (12)	ERS459551	37	15308	4051571	1730088	6	1040457	14	819414	312349	30 318 316
<i>Microbotryum intermedium</i> 1389	3	562278	62677	11040	8512	8220	4 786 123 623	ERS1324257	24	16897	2982376	1644950	6	1002499	13	977543	1002499	23 461 035
<i>Microbotryum silenes-dioicae</i> 1303_A1	7	938931	46274	10834	7946	7576	7 460 889 114	ERS1436592	144	6283	2425171	936137	12	153077	45	233285	32003	33 593 023
<i>Microbotryum silenes-dioicae</i> 1303_A2	6	490924	46480	12620	8406	7474	4 126 642 950	ERS1436593	128	8587	3251277	1321747	10	359665	30	265304	29362	33 958 966

**Table S5. *Microbotryum lychnidis-dioicae* and *M. silenes-dioicae* strains used for performing polymorphism analyses.** Reference to strain ID, species, mating-type, accession number, number and percentage of reads mapped as proper pairs and reference of the original genome publication, are given.

Strain ID	Species	Mating-type	Accession number	Number of reads mapped as proper pairs	% of reads mapped as proper pairs	Reference
MvSd-335-H3-A1	<i>M. silenes-dioicae</i>	a <sub>1</sub>	SRS1072004	102696842	96.28%	(48)
MvSd-336-01-A1	<i>M. silenes-dioicae</i>	a <sub>1</sub>	SRS1072005	127697676	96.85%	(48)
MvSd-578-2-A2	<i>M. silenes-dioicae</i>	a <sub>2</sub>	SRS1072006 SRS1072007	78559190	96.54%	(48)
MvSd-69-05-A2	<i>M. silenes-dioicae</i>	a <sub>2</sub>	SRS1072010	121301046	97.35%	(48)
MvSd-707-1-A2	<i>M. silenes-dioicae</i>	a <sub>2</sub>	SRS1072011	138917854	96.93%	(48)
MvSd-72-A2	<i>M. silenes-dioicae</i>	a <sub>2</sub>	SRS1072012 SRS1072013	181254142	97.40%	(48)
MvSd-851-6-A2	<i>M. silenes-dioicae</i>	a <sub>2</sub>	SRS1072015	113257210	97.21%	(48)
MvSd-900-1-A1	<i>M. silenes-dioicae</i>	a <sub>1</sub>	SRS1072016	143422672	97.33%	(48)
MvSd-932-2-A2	<i>M. silenes-dioicae</i>	a <sub>2</sub>	SRS1072017	107796114	96.87%	(48)
MvSd-937-2-A2	<i>M. silenes-dioicae</i>	a <sub>2</sub>	SRS1072018	134111998	97.24%	(48)
MvSd-949-2-A1	<i>M. silenes-dioicae</i>	a <sub>1</sub>	SRS1072019	105576460	96.66%	(48)
MvSd-IT02-A2	<i>M. silenes-dioicae</i>	a <sub>2</sub>	SRS1072020	126925450	96.74%	(48)
MvSd-sp002-A1	<i>M. silenes-dioicae</i>	a <sub>1</sub>	SRS1072021	133835232	96.73%	(48)
MvSd-sp003-A1	<i>M. silenes-dioicae</i>	a <sub>1</sub>	SRS1072023	110615508	97.32%	(48)
MvSI-00-10-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS1072024	71471502	95.32%	(48)
MvSI-100-6-A1	<i>M. lychnidis-dioicae</i>	a <sub>1</sub>	SRS1072025	70881632	95.69%	(48)
MvSI-140-01-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS1072033	157160988	95.07%	(48)
MvSI-141-01-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS1072034	66844444	94.89%	(48)
MvSI-40-01-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS1072035	65532992	94.18%	(48)
MvSI-446-2-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS1072038 SRS1072039	68358472	94.97%	(48)
MvSI-466-3-A1	<i>M. lychnidis-dioicae</i>	a <sub>1</sub>	SRS1072042 SRS1072043	75508212	94.92%	(48)
MvSI-576-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS1072044 SRS1072045	56128954	95.40%	(48)
MvSI-641-A1	<i>M. lychnidis-dioicae</i>	a <sub>1</sub>	SRS1072046 SRS1072047	87723406	95.25%	(48)
MvSI-661-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS1072050	84294192	94.85%	(48)
MvSI-665-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS1072048 SRS1072049	37283112	94.87%	(48)
MvSI-687-5-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS1072051	46854784	94.81%	(48)
MvSI-830-2-A1	<i>M. lychnidis-dioicae</i>	a <sub>1</sub>	SRS1072056	81814412	95.81%	(48)
MvSI-856-2-A1	<i>M. lychnidis-dioicae</i>	a <sub>1</sub>	SRS1072057	37685654	94.57%	(48)
MvSI-I00-3_A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS1072064	78297378	95.19%	(48)
MvSI-IOA-A1	<i>M. lychnidis-dioicae</i>	a <sub>1</sub>	SRS1072065	61186726	95.57%	(48)
MvSI-W-1069-A1	<i>M. lychnidis-dioicae</i>	a <sub>1</sub>	SRS781520	46860062	92.35%	(26)
MvSI-W-1088-A1	<i>M. lychnidis-dioicae</i>	a <sub>1</sub>	SRS781522	38830336	92.42%	(26)
MvSI-W-1089-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS781488	38918966	91.94%	(26)
MvSI-W-1090-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS781491	31476556	93.72%	(26)
MvSI-W-1103-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS781483	45555798	91.93%	(26)
MvSI-W-405-A1	<i>M. lychnidis-dioicae</i>	a <sub>1</sub>	SRS780332	88330498	94.39%	(26)
MvSI-W-769-A1	<i>M. lychnidis-dioicae</i>	a <sub>1</sub>	SRS781521	42520222	91.74%	(26)
MvSI-W-920-A2	<i>M. lychnidis-dioicae</i>	a <sub>2</sub>	SRS781485	42908076	92.50%	(26)



**Table S6. Nucleotide diversity ( $\theta_{\pi}$ ) and divergence (mean number of pairwise differences) estimates in pseudo-autosomal regions (PARs) and evolutionary strata in:**

(A) *Microbotryum silenes-dioicae*

Gene set	Mean diversity $\pm$ SE		Divergence from <i>M. violaceum</i> s. str.	
	All sites	Silent sites	All sites	Silent sites
Purple stratum	3.24e-02 $\pm$ 1.53e-02	5.23e-02 $\pm$ 1.98e-02	4.71e-02 $\pm$ 2.04e-02	6.09e-02 $\pm$ 1.42e-02
Blue stratum	2.43e-02 $\pm$ 1.09e-02	3.72e-02 $\pm$ 1.72e-02	5.61e-02 $\pm$ 2.09e-02	6.69e-02 $\pm$ 1.45e-02
Orange stratum	6.25e-02 $\pm$ 4.75e-02	8.88e-02 $\pm$ 6.2e-02	7.87e-02 $\pm$ 5.16e-02	6.25e-02 $\pm$ 2.02e-02
Black stratum	9e-03 $\pm$ 8.79e-04	1.45e-02 $\pm$ 1.29e-03	2.33e-02 $\pm$ 1.13e-03	3.57e-02 $\pm$ 1.11e-03
Red stratum	2.47e-03 $\pm$ 1.56e-03	4.03e-03 $\pm$ 2.41e-03	2.26e-02 $\pm$ 3.12e-03	3.88e-02 $\pm$ 3.8e-03
Green stratum	2.22e-03 $\pm$ 3.6e-04	3.4e-03 $\pm$ 8.92e-04	2.39e-02 $\pm$ 6.44e-03	4.53e-02 $\pm$ 1.27e-02
PARs	2.46e-04 $\pm$ 3.37e-05	4.11e-04 $\pm$ 8.56e-05	4.1e-02 $\pm$ 2.11e-03	8.18e-02 $\pm$ 3.54e-03

(B) *Microbotryum lychnidis-dioicae* (whole data set)

Gene set	Mean diversity $\pm$ SE		Divergence from <i>M. violaceum</i> s. str.	
	All sites	Silent sites	All sites	Silent sites
Purple stratum	3.6e-02 $\pm$ 1.48e-02	5.53e-02 $\pm$ 2.02e-02	5.28e-02 $\pm$ 2.05e-02	6.46e-02 $\pm$ 1.47e-02
Blue stratum	2.44e-02 $\pm$ 9.91e-03	3.92e-02 $\pm$ 1.58e-02	5.66e-02 $\pm$ 1.98e-02	7.01e-02 $\pm$ 1.39e-02
Orange stratum	6.16e-02 $\pm$ 4.51e-02	8.94e-02 $\pm$ 6.01e-02	8.06e-02 $\pm$ 5.18e-02	6.5e-02 $\pm$ 2.04e-02
Black stratum	9.88e-03 $\pm$ 6.78e-04	1.61e-02 $\pm$ 1.08e-03	2.57e-02 $\pm$ 1.02e-03	3.88e-02 $\pm$ 1.13e-03
Red stratum	3.11e-03 $\pm$ 1.34e-03	3.99e-03 $\pm$ 2.36e-03	2.32e-02 $\pm$ 2.88e-03	4.03e-02 $\pm$ 4.13e-03
Green stratum	1.47e-03 $\pm$ 4.39e-04	2.49e-03 $\pm$ 1.2e-04	2.48e-02 $\pm$ 6.92e-03	4.75e-02 $\pm$ 1.12e-02
PARs	2.61e-03 $\pm$ 2.22e-04	3.16e-03 $\pm$ 2.64e-04	4.41e-02 $\pm$ 2.24e-03	8.78e-02 $\pm$ 3.79e-03

(C) the North-Western cluster of *Microbotryum lychnidis-dioicae*

Gene set	Mean diversity $\pm$ SE		Divergence from <i>M. violaceum</i> s. str.	
	All sites	Silent sites	All sites	Silent sites
Purple stratum	3.92e-02 $\pm$ 1.61e-02	5.77e-02 $\pm$ 2.03e-02	5.23e-02 $\pm$ 1.98e-02	6.34e-02 $\pm$ 1.36e-02
Blue stratum	2.68e-02 $\pm$ 1.09e-02	3.99e-02 $\pm$ 1.6e-02	5.69e-02 $\pm$ 1.98e-02	7.03e-02 $\pm$ 1.38e-02
Orange stratum	6.72e-02 $\pm$ 4.95e-02	9.03e-02 $\pm$ 6.11e-02	7.8e-02 $\pm$ 4.93e-02	6.41e-02 $\pm$ 1.93e-02
Black stratum	1.05e-02 $\pm$ 7.57e-04	1.63e-02 $\pm$ 1.12e-03	2.58e-02 $\pm$ 1.02e-03	3.89e-02 $\pm$ 1.15e-03
Red stratum	3.4e-03 $\pm$ 1.45e-03	4.85e-03 $\pm$ 2.41e-03	2.36e-02 $\pm$ 2.93e-03	4.07e-02 $\pm$ 4.18e-03
Green stratum	4.72e-04 $\pm$ 3.69e-04	9.37e-04 $\pm$ 9.37e-04	2.5e-02 $\pm$ 7.24e-03	4.75e-02 $\pm$ 1.22e-02
PARs	9.05e-04 $\pm$ 1.23e-04	1.49e-03 $\pm$ 2.19e-04	4.45e-02 $\pm$ 2.29e-03	8.76e-02 $\pm$ 3.79e-03

(D) the Southern cluster of *Microbotryum lychnidis-dioicae*

Gene set	Mean diversity $\pm$ SE		Divergence from <i>M. violaceum</i> s. str.	
	All sites	Silent sites	All sites	Silent sites
Purple stratum	4.23e-02 $\pm$ 1.76e-02	5.54e-02 $\pm$ 2e-02	4.73e-02 $\pm$ 1.63e-02	5.64e-02 $\pm$ 1.07e-02
Blue stratum	2.85e-02 $\pm$ 1.18e-02	3.82e-02 $\pm$ 1.57e-02	5.66e-02 $\pm$ 1.99e-02	6.93e-02 $\pm$ 1.36e-02
Orange stratum	7.29e-02 $\pm$ 5.34e-02	8.82e-02 $\pm$ 5.9e-02	6.26e-02 $\pm$ 3.53e-02	5.88e-02 $\pm$ 1.41e-02
Black stratum	1.1e-02 $\pm$ 8.12e-04	1.54e-02 $\pm$ 1.07e-03	2.53e-02 $\pm$ 9.53e-04	3.85e-02 $\pm$ 1.11e-03
Red stratum	2.8e-03 $\pm$ 1.64e-03	3.57e-03 $\pm$ 2.24e-03	2.35e-02 $\pm$ 3.16e-03	4.01e-02 $\pm$ 4.28e-03
Green stratum	8.3e-04 $\pm$ 5.2e-04	1.2e-03 $\pm$ 1.2e-03	2.5e-02 $\pm$ 6.54e-03	4.81e-02 $\pm$ 9.62e-03
PARs	1.07e-03 $\pm$ 1.29e-04	1.66e-03 $\pm$ 2.58e-04	4.39e-02 $\pm$ 2.19e-03	8.71e-02 $\pm$ 3.78e-03

**Table S7. Results of maximum-likelihood Hudson-Kreitman-Agade (MLHKA) tests for assessing whether differences in diversity between each evolutionary stratum and the pseudo-autosomal regions (PARs) is due to balancing selection or elevated mutation rates.** Genes were concatenated per stratum unless specified otherwise. Statistically significant results before and after correction for multiple testing at the significance level of 0.05 are indicated in bold; d.f.: degrees of freedom; Selection coefficient: maximum likelihood estimate of the selection coefficient.

(A) In *Microbotryum silenes-dioicae*

Test	d.f.	Selection coefficient	ln likelihood null model	ln likelihood alternative model	Chi2 statistics	p-values	Adjusted p-values
Purple stratum vs PAR	1	60.4123	-284.798	-370.048	170.5	<b>5.75e-39</b>	<b>4.03e-38</b>
Blue stratum vs PAR	1	44.6019	-286.815	-370.684	167.74	<b>2.31e-38</b>	<b>1.38e-37</b>
Orange stratum vs PAR	1	75.4768	-286.583	-391.715	210.26	<b>1.20e-47</b>	<b>9.62e-47</b>
Black stratum vs PAR	1	24.844	-287.805	-340.116	104.62	<b>1.48e-24</b>	<b>7.39e-24</b>
Red stratum vs PAR	1	7.6661	-288.224	-300.015	23.58	<b>1.20e-06</b>	<b>3.59e-06</b>
Green stratum vs PAR	1	4.74454	-285.191	-289.094	7.81	<b>5.21e-03</b>	<b>5.21e-03</b>
Red stratum vs PAR (no concatenation)	16	56.0665	-335.885	-380.386	89	<b>3.82e-12</b>	<b>1.53e-11</b>
Green stratum vs PAR (no concatenation)	3	10.3826	-289.177	-297.46	16.57	<b>8.68e-04</b>	<b>1.74e-03</b>

(B) In *Microbotryum lychnidis-dioicae* (whole data set)

Test	d.f.	Selection coefficient	ln likelihood null model	ln likelihood alternative model	Chi2 statistics	p-values	Adjusted p-values
Purple stratum vs PAR	1	12.1267	-383.916	-412.85	57.87	<b>2.80e-14</b>	<b>1.40e-13</b>
Blue stratum vs PAR	1	10.7939	-386.523	-412.316	51.59	<b>6.85e-13</b>	<b>2.74e-12</b>
Orange stratum vs PAR	1	18.0431	-384.52	-424.17	79.3	<b>5.34e-19</b>	<b>3.20e-18</b>
Black stratum vs PAR	1	6.39375	-385.591	-402.312	33.44	<b>7.34e-09</b>	<b>2.20e-08</b>
Red stratum vs PAR	1	1.94212	-383.672	-385.361	3.38	6.61e-02	1.32e-01
Green stratum vs PAR	1	0.795155	-384.517	-380.695	-7.64	1.00e+00	1.00e+00

(C) In the North-Western cluster of *Microbotryum lychnidis-dioicae*

Test	d.f.	Selection coefficient	ln likelihood null model	ln likelihood alternative model	Chi2 statistics	p-values	Adjusted p-values
Red stratum vs PAR	1	4.96955	-344.795	-336.841	15.91	<b>6.65e-05</b>	<b>1.99e-04</b>
Green stratum vs PAR	1	0.941793	-333.064	-336.68	-7.23	1.00e+00	1.00e+00
Red stratum vs PAR (no concatenation)	16	2.29673	-417.742	-388.587	58.31	<b>1.01e-06</b>	<b>4.02e-06</b>
Green stratum vs PAR (no concatenation)	3	0	-332.85	-333.992	-2.28	1.00e+00	1.00e+00

(D) In the Southern cluster of *Microbotryum lychnidis-dioicae*

Test	d.f.	Selection coefficient	ln likelihood null model	ln likelihood alternative model	Chi2 statistics	p-values	Adjusted p-values
Red stratum vs PAR	1	4.0771	-348.271	-341.431	13.68	<b>2.17e-04</b>	<b>8.67e-04</b>
Green stratum vs PAR	1	1.1329	-339.258	-342.942	-7.37	1.00e+00	1.00e+00
Red stratum vs PAR (no concatenation)	16	1.48588	-410.369	-392.762	32.65	<b>8.22e-03</b>	<b>2.46e-02</b>
Green stratum vs PAR (no concatenation)	3	0.639097	-338.134	-342.151	-8.03	1.00e+00	1.00e+00