nature portfolio

Peer Review File

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

Corresponding Author: Dr Elise Lucotte

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I read the manuscript by Lucotte et al. with pleasure and a fair bit of excitement, since they present the first case of loss of mating-type determination at the HD locus in a basidiomycete fungus. This scenario has been thought of for a long time, but in every previously studied case of transitions from a bi- to a unifactorial mating system, the loci have either been linked, or the PR locus has lost mating type determination function. The authors base this novel finding on thorough genomic analyses of different lineages of anther-smut fungi belonging to the genus Microbotryum, and accompany their bioinformatics methods with experimental work that bring a crucial component to their study. The experimental and genomics methods are robust and detailed, and I cannot identify any major methodological issues. However, the presentation of these findings can be improved by clarification in several sections of the text, and some details should be added to the figures to help the reader interpret them.

In particular, the presentation of how the HD2 gene in M. superbum was deemed to be non-functional requires some additional work. Viewing Figure 7, I agree that the HD2 gene(s) look suspicious. However, the authors need to make sure the splitting of the gene is not an artifact from gene annotation. In the “Mated24” track of RNA-seq coverage, there is actually coverage across the whole site, which may indicate an intact gene. The authors explain in the text that they identified new start/stop codons, but the annotated reading frame starts before any RNA-seq coverage, which leads me to question that the correct reading frame is annotated. A track of mapped DNA-seq reads across this whole region would also help the reader ensure that there is no problem with the assembly.

The main text (lines 226-239) describing this crucial result is also in need of improvement. I think a better term for the two HD2 fragments would be gene models, since CDS implies that they are exons of the same gene, which is not the case according to the current gene annotation. “One of the two novel CDS appeared expressed during mating conditions”, actually there is RNA-seq coverage across both gene models. “This may be because the strain is homozygous”, how can a haploid strain be homozygous? “and M. intermedium”, this species is not in Figure 7. Is it supposed to be? Otherwise the authors should refer to the figure after M. lychnidis-dioicae, and put “data not shown” after M. intermedium. If they have the data, which the manuscript gives the impression of, they should show it to corroborate their results.

Furthermore, the authors identify much larger PR chromosomes in M. superbum and M. shykoffianum, and state that these chromosomes have accumulated specific families of TEs. That the TE load is higher than on the autosomes is quite clear, but to be able to say that specific TE families have preferentially accumulated, the authors need statistical support. On lines 482-485 preferential accumulation seems to be taken as a fact, but seems to be based only on what is shown in Figure 8. From the figure I can accept that LINEs may have preferentially accumulated in the light ruby stratum, which is not mentioned in the text, but the other TE families are also common on the autosomes. The authors have to test this statistically before they can say anything about preferential accumulation of specific families.

The authors need to show data from the crossing experiments. As it stands, these results are only described in text (in the paragraph starting on line 240, and on lines 302-305). To enable the reader to make their own interpretation of the data, a supplementary table or figure should be included showing the results from the crosses.

To make the message of the paper as clear as possible, some improvements should be made in how the results are visually presented. Near the end, I outline some issues for each figure.

Minor comments:
Line 40 and throughout the text: Microbotryum scorzonerae is often misspelled. I assume it should be the taxon with taxid 72746 on NCBI throughout the text (I find no taxon called M. scorzonarae or M. scorzoraneae). Please check the text and correct this.
Line 159: The term ”high-quality” regarding genome assemblies can mean almost anything depending on the organism and context. A pointer towards what the authors mean here would help the reader orient themselves. I assume the authors mean to say that the assemblies are of high contiguity, allowing the analysis of physical linkage between MAT loci, but some numbers (maybe N90?) would help.
Line 159: “from one strain from each of two” is this not just the same as “from two species”?
Line 162: “as expected”, based on what? It is not necessary to mention any expectation here.
Line 173: What does a1 and a2 mean? I assume they refer to mating types. A short note in the results about how they were obtained would help the reader, without having to go into methods details. I suggest adding this explanation on line 163, before “In all four haploid genomes”.
Line 176: How was the centromere identified? This objective is non-trivial in most species, and some explanation should be given. Centromeres are known to be able to change even in short evolutionary time. A better, more careful, wording would be “predicted” instead of “identified”. Precise determination of the centromere would require additional data.
Lines 192-193: The wording of this sentence is unclear. “SNPs on 149 Illumina genomes” makes little sense to me. I think the authors mean to say that SNPs were identified from 149 genomes sequenced with Illumina technology. They should also state which species have been investigated. In the previous sentence, they refer to these samples as “collected on Dianthus plants”, but it is not clear whether these samples correspond to one or several species of Microbotryum, and how they are related to the strains they based most of their study on.
Lines 195-196: Here some numbers would help interpretation. “A bit higher” is not very specific.
Line 216: “the mapping of diploid reads”, this wording is incorrect. The reads are not diploid, but originate from a diploid genome.
Line 292: what does “exhibited polymorphism” actually mean? Would a single SNP in the HD locus count? Some numbers are necessary here, I suggest pi for average nucleotide divergence, but raw SNP number per strain would also help.
Line 307: What is meant by “gap” here? I think the authors mean a deletion, but gap brings to mind an assembly gap in the form of a stretch of N’s. They also do not refer to any figure describing this finding. The finding of convergent non-sense mutations in the HD2 gene in these different lineages is highly important and should be shown. The authors should consider adding another panel in Fig. 7, or at least a supplementary figure if this is not possible (e.g., maybe they do not have RNA seq of M. scorzonerae).
Line 400: which figures?
Line 450: Scorzonera is misspelled.
There are some typos, faulty sentences, and missing references in the Methods section. I point out some errors I find below, but I suggest also another round of proofreading before the next submission.
Line 586: Some detail is missing about the Zymo kit.
Line 611: Some detail is missing about the Qiagen kit.
Lines 622-628: This sentence is impossible to read. Please rephrase.
Line 624: Here it is stated that mating was assessed after 48h. In Figure 7 the legend says after 24h. Which is correct, or is this not the same data? If not, where can the reader find the methods underlying Fig. 7?
Line 641 and in various places in the text: Make sure species names are correctly italicized.
Line 647: Missing reference/link.
Lines 746-748: This sentence is speculative. The authors could address the matter of divergence between a1 and a2 HD alleles by mapping back the raw reads and calling variants in this region.

Figure 2: Reading the figure as it is first referred to in the main text, “a1” or “a2” have not been explained. The title of the figure says a1, but in the figure itself a2 is annotated. Which is correct? Is the figure drawn to scale? If so, the legend should mention this, and there should be a scale bar to orient the reader. Can the telomeres be indicated in the figure? This would give the reader a sense of the assembly quality.
Figure 4: The red vertical lines in panel A are not explained. Panel C: it is unclear to me what is shown here. What is on the y-axis? What are b1 and b2, referred to in the legend but not present in the figure? In the legend, b1 is not written in subscript but a1, a2, and b2 are. Is this correct? The x-axis label: the species name should be written as M. lagerheimii, i.e. with italics and lower-case “l”.
Figure 6: Panel A: an insert showing which photograph belongs to which species would be a useful reminder here.
Figure 7: Some cosmetic improvement would help the reader. Flipping the direction of the panel for M. superbum (upper) would improve the ability to interpret the figure, so that the HD1 and HD2 genes are in the same orientation as for M. lychnidis-dioicae. The details and text are also quite small, so I suggest removing some of the flanks around the HD locus, perhaps 1.5kb on each side, and zooming in on the HD genes.
Figure 8: Some modifications would improve clarity here. It is hard to get a general idea about the PR chromosome since the different strata are divided up. My interpretation is that it would be very similar to “Black PR” since it is so big compared to the other strata and the PAR. At least, the categories making up the PR chromosome should be indicated as belonging to this chromosome. The contrast between “Unclassified” and “Satellite” is very low, so it is hard to determine the category of the largest proportion of TEs. Also, satellites, rRNA, and simple repeats are not TE categories and should not be called as such on the x-axis of panel A.
Figure S9: The authors should state which species is infecting the flower in the figure legend.
Table S1: The orientation of the primer sequences should be stated in the legend or directly in the table.

Table S3 contains “TBD” for sequences generated during this study. Please update with current accession numbers. There are also empty cells for samples generated during previous studies. Were these data not deposited? If so, this could be indicated in the table. The authors should also consider submitting these data for transparency.
A recurring language issue throughout the manuscript: in figure legends the authors use the word “figured” as a synonym of “visualized”. I have not encountered this use of the word before; please confirm with a native English speaker that it is linguistically correct.

(Remarks on code availability)
There is a vast resource of open access code in the specified repository. I have visually reviewed some randomly selected scripts, but not tested the code myself. There is ample documentation about the code.

Reviewer #2

(Remarks to the Author)
This study by Lucotte et al. investigates the mating systems and genomic architecture of anther-smut Microbotryum fungi, with a specific focus on the lack of linkage between the HD and PR mating-type determining loci, possibly associated with the repeated loss of function at HD mating-type genes. By leveraging high-quality genome assemblies of additional sister Microbotryum species isolated from Dianthus and a distantly related species, they provide a comprehensive analysis of how these genomic configurations may have evolved. Specifically, their research reveals that the HD and PR mating-type loci, currently residing on different chromosomes, have evolved from an ancestral state where these loci were likely linked, undergoing subsequent chromosomal breakage. These findings underscore a complex evolutionary landscape where mating compatibility is apparently controlled by a single genetic factor due to the loss of function in HD genes, which is a novel observation in heterothallic basidiomycete fungi.

The manuscript offers valuable insights into the evolutionary mechanisms that shape fungal mating systems, elucidating the complex interplay between genomic architecture and reproductive strategies and should be of broad interest to the evolutionary genomics community. While the findings are significant, the presentation could benefit from improved clarity and streamlined exposition. Specific areas needing attention include typographical errors, inconsistent formatting, and the need for better-detailed explanations in the results section that are currently detailed in the methods section. Additionally, while some figures offer valuable information, others require enhancement to effectively convey the underlying data. Addressing these points, along with my specific comments provided below, will greatly enhance the manuscript's readability and scientific impact.

- Main comments -

On the formation of the large non-recombining PR chromosome:

1) The authors propose that the ancestral PR and HD chromosomes in M. superbum and M. shykoffianum fully fused, followed by extensive recombination suppression and intrachromosomal rearrangements linking the PR and HD loci. However, the observed conservation of synteny in the extant HD chromosomes compared to M. lagerheimii (representing the ancestral state) raises some questions about this model. It could suggest a potential alternative scenario where the “fusion” might have involved only a segment of the ancestral HD chromosome that excludes the HD genes. Such a scenario could involve a fission in the ancestral HD chromosome, followed by the fusion of the acentric fragment with the PR chromosome, while the remaining portion could have been stabilized by the de novo formation of telomeric sequences. This hypothesis could potentially explain the conservation of synteny of the extant HD chromosome.

2) Assuming the proposed fusion resulted in a dicentric chromosome, the likely genomic instability would necessitate resolution mechanisms. These might include breakage-fusion-bridge cycles, centromere inactivation (either epigenetic or structural), or chromosome breakage followed by repair to restore monocentric chromosomes. Given the hypothesis of HD and PR loci linkage, it seems plausible that one of the centromeres was inactivated immediately following fusion. Subsequent to any proposed fission event, centromere function would need to be re-established. Could the authors provide insights on how these dicentric conflicts might have been resolved in these species?

3) In basidiomycetes, the PR locus often expands, possibly due to the duplication or even quadruplication of pheromone genes, which can promote intrachromosomal rearrangements (e.g. inversions). Therefore, I think it would be informative to examine the position of the pheromone genes within the PR chromosome. Clarifying whether these genes contribute to the expansion of the PR locus, even in the absence of linkage to the HD locus, could provide additional insight into the genomic dynamics and evolutionary pressures shaping these regions.

On the RNA-seq and polymorphism analysis of the HD genes:

1) The RNA-seq read coverage presented in Fig. 7 leaves it unclear whether the HD2 gene of M. superbum is expressed as a single transcript. Furthermore, the lack of synteny between HD2 gene fragments in a1 and a2 strains requires clarification. Could you specify the method used to assess synteny in these plots?

2) Figure 7 references “Mated24”, suggesting RNA samples were obtained from a1 and a2 cells undergoing mating for 24

hours. However, Table S3 and the methods section refer to “Mating 48h”, indicating a potential inconsistency. Additionally, the presence of SNPs and unequal read coverage along the HD1 gene in the “mated24” condition is suggestive of allelic variation. Could you please clarify these observations?

3) A few enhancements in displaying RNA-seq data are warranted, particularly regarding consistent scaling or clear indication of scale used. Further details are also needed on whether the RNA-seq data was mapped to a combined reference genome encompassing both PR and HD mating-type chromosomes. Additionally, please clarify whether the RNA-seq data were paired-end or single-end and the choice of minimap2 for read mapping over other RNA-seq mappers that are more adept at handling splice junctions.

4) It should be noted that even if the HD2 gene in the a2 mating type of M. superbum is non-functional, the HD2 gene from the a1 mating type could still potentially form a functional heterodimer with the HD1 gene of the a2 mating type. This phenomenon has been observed in other species with PR-HD linked loci, where typically one mating type loses one HD gene, while the other retains the opposite HD gene, demonstrating an important aspect of mating-type gene functionality and evolution.

5) I may have missed it, but I couldn't find any results related to RNA-seq data analysis for M. superbum haploid condition or for M. intermedium. Additionally, regarding M. shykoffianum c212, the genome appears to have been sequenced from a mixture of meiotic products from presumably the same diploid individual, showing that the HD1 and HD2 sequences are intact and seemingly 100% identical. Was this observation independently validated through PCR/Sanger sequencing using separate a1 and a2 haploid sporidia? Alternatively, examining HiFi reads might reveal any variation when mapping to the HD chromosome.

6) Could you please clarify whether the 149 Illumina genomes were obtained from multiple diploid individuals collected from different flowers? Given the considerable polymorphism observed at the HD genes, it would be informative to understand how the SNP patterns are distributed along these genes, particularly whether there is more variation at the corresponding N-terminal regions, as is common in other basidiomycetes. Additionally, contrasting the level of heterozygosity of HD genes in these species with that of species where HD genes still apparently determine mating could also be relevant.

7) In M. scorzonerae the 19bp gap noted in the HD2 coding sequences needs a baseline for comparison. Additionally, has there been any attempt to predict an intact CDS or to rule out additional introns?

8) While not essential for the conclusions of this study, future research could benefit from considering targeted gene deletions of the HD genes in some of these species. This approach would not only definitively confirm whether these genes are necessary for mating-type determination but could also elucidate their roles in other aspects of sexual reproduction, i.e. maybe the HD genes are not involved in mating-type identity but are still required for sexual reproduction and thus retained to some extent.

On the tetrad analysis section:

1) I think it would be beneficial to include mating test results for all 48 meiotic products from the 12 tetrads collected from D. pavonius. Presenting this data in a supplementary table could provide a better overview of the mating compatibility outcomes.

2) The authors should also consider including supplementary data displaying the results from PCR analyses that confirm the presence of varying PR alleles and uniform HD alleles among the meiotic products. Additionally, providing photo examples of conjugated (mating compatible) and non-conjugated (mating incompatible) cases, along with visualizations of hyphal production, would effectively support the assertion that HD genes no longer influence mating-type determination in these species. Likewise, a detailed account of disease symptoms in plants inoculated with strains that have identical HD alleles would also be helpful. I think this would strengthen this unique observation among basidiomycetes for all readers but also the study's groundbreaking conclusions.

3) For the two meiotic products that produced hyphae without tester strains, have the authors assessed if they are fully diploids or aneuploids? Interesting that they don’t behave as the “solopathogenic” strains of certain Ustilago spp. Any idea why?

4) For the 6 additional isolates, where 2 strains are homozygous and 4 are heterozygous for the HD gene, you could could consider inspecting the SNP distribution along the HD genes in the heterozygous. It would be informative to enumerate the changes in the predicted protein and compare these with typical patterns observed in basidiomycetes and in other Microbotryum species where the HD genes are still involved in mating-type determination. In other words, would be great to have an idea about how many changes are sufficient for HD alleles to be functionally different.

- Minor comments -

Lines 66-68: Consider breaking up and rephrasing for clarity: “In fungi with mating types, gamete compatibility is controlled during the haploid stage, where only cells of different mating types can successfully mate. Mating types are determined by either one or two loci, classifying the systems as uni- and bifactorial, respectively.”

Lines 72-74: The statement “Recombination is typically suppressed at loci determining sexes and mating types, which ensures proper function in these complex phenotypes” could benefit from further clarification on how exactly recombination suppression contributes to the proper function of these phenotypes.

Lines 100-101: Introduction of unifactorial basidiomycetes resulting from HD-PR linkage could include references to Cryptococcus pathogenic species where extensive studies exist to provide a broader perspective on the diversity across basidiomycetes.

Line 101: Please add a period after 'spp' to correctly punctuate the abbreviation for species.

Line 105-106: Suggest adding a comma after “plants” to improve the clarity of the sentence structure.

Line 117: Should it be intra-tetrad mating?

Line 136: Consider changing to “carry” and "yet display"

Line 141: Consider rephrasing “issued from the fusion” to “resulting from the fusion”.

Line 142: A "the" is missing before the HD locus.

Lines 144-145: The description of the translocation process as occurring “via a transient stage of whole chromosome fusion” is not clear.

Line 159: It would be helpful to have some statistics on the genome assemblies produced in this study (genome size, number of contigs, etc.) as it is currently unclear how well assembled they turn out to be. This could be provided as supplementary data.

Lines 164-166: Although it is stated that telomeric repeats were identified in almost all of the mating-type contigs, only 5/8 telomeric repeats for M. superbum and 4/8 for M. shykoffianum are accounted for. Therefore, maybe consider instead emphasize that at least one set of PR or HD chromosomes were completely assembled telomere-to-telomere in each species. Was there any technical reason for not having the mating-type chromosomes assembled into single contigs (read length, low coverage, too many repeats, etc.)? At least in some cases (e.g. M-sup-PR a1) the break is not within the centromere, which is usually more repetitive.

Lines 176-179: Please clarify this sentence: “The centromere was identified within the non-recombining region and telomeres at the edges of the new PR and HD chromosomes (Figures S4-S5).” As noted above, telomeres were not found at all the chromosome ends. It is also not completely clear at this stage if the HD-a2 chromosome of M. shykoffianum is substantially smaller than the respective HD-a1 chromosome or is not well assembled (this is only clarified further down in the text). Finally, from Figure S5B, it seems there are a few links between the a1-HD chromosome and the a2-PR chromosome in M. shykoffianum: are these genes or e.g. transposable elements?

Line 182: Please check if Figure 3 here is appropriate, as I don’t see any colored green genes in Fig. 3

Line 296: Maybe consider explaining in detail the inbreeding coefficient (FIS) and what values it can take.

Lines 315-320: Why not compare M. scorzonerae with M. intermedium instead, given they seem to be more closely related? Is it because M. intermedium and M. lagerheimii are completely collinear across the genome? Additionally, since M. scorzonerae was sequenced using ONT, I’m curious if the authors considered inspecting the 5mC methylation landscape in the MAT chromosome compared to autosomes.

Lines 367-379: It is very hard to follow all these proposed rearrangements given the provided figures, in particular Fig S9. How small is the predicted non-recombining region around the PR in M. scorzonerae? Would be great to have comparisons such as those shown in the circos plots for this species as well, compared to M. lagerheimii (or M. intermedium). In general, I think the provided circos plots are much more informative as they also display other chromosomal features, e.g., centromeres, position of the mating-type determining genes, etc. On this note, would also be great if you could plot the position of the pheromone genes in these plots as the definition/delimitation of the PR loci includes both the pheromone and receptor genes.

Lines 414-415: As far as I'm aware, current research on Malassezia suggests that the transition from two MAT loci on separate chromosomes to a configuration where both loci are on the same chromosome (though not necessarily genetically linked) likely occurred only once. This event was followed by secondary rearrangements that relocated the loci back onto separate chromosomes.

Line 556: Is M. intermedium really an outgroup? Is the current model that this species also harbors PR and HD loci linked to the centromeres? Otherwise it is intriguing why this species retained a bifactorial system with both mating-type loci on separate chromosomes.

Line 586: check the extra “----”

Line 1434: There is a panel or figure after Figure S8 that has no legend?

Figure 2: As indicated above, would it be possible to also add the location of the pheromone genes since it is important to understand if they are close to the pheromone receptor or not? Additionally, the legend references a1 genomes, yet the figure appears to display mating type a2 genomes. Please verify and correct this discrepancy. Furthermore, the PR chromosome of M. lagerheimii is depicted as split into two contigs; please clarify if these are indeed scaffolds if gaps are present, otherwise they should be referred to as contigs. Also, could you clarify the nature of the additional links going from M. lagerheimii HD and PR chromosomes into other autosomes in M. superbum and M. shykoffianum? Are these links representing genes (e.g. duplications) or are perhaps annotations of transposable elements (TEs) as many appear to target the ends of the contigs?

Figure 4 and legend: Please revise this figure and legend. Consider rephrasing to "…the HD chromosome on the left and the PR chromosome on the right". Additionally, please indicate what the two red dots represent in the HD chromosome plot as well as the two the red vertical lines in both the HD and PR chromosome plots.

Data availability: please ensure that all genomic data and utilized code, including that used for tree topology assessment, are made available in a publicly accessible database/repository.

Strain nomenclature consistency: I detected a few instances of inconsistency in the naming of strains, with "MvDp−1065" used in some instances (e.g., Methods, line 795) and "Mv-Sup-1065" in others (e.g., Fig. 7). Please consider standardizing the nomenclature across the manuscript to improve readability.

References: please double-check the references, as some, such as Reference 33, have been published and may require updating to reflect their current status.

(Remarks on code availability)
I have not been able to test the code, but seems to be well explained.

Reviewer #3

(Remarks to the Author)
Regarding the Lucotte et al manuscript titled "Repeated loss of function at HD mating-type genes and of recombination suppression without mating-type locus linkage in anther-smut fungi", the manuscript represents a tremendous amount of research aimed at providing insight into the genetic changes associated with mating types in the Microbotryum anther smut fungi. The work convincingly shows the rearrangement and loss of mating-type function for the HD genes. This represents the first instance in which the ancestral two mating type loci system in basidiomycete fungi is converted to a single locus system by the loss of the HD genes function and not the PR function. The manuscript also uncovers further suppression of recombination in various Microbotrym lineages. As a body of research this is well executed and appropriately controlled with the appropriate follow-on experiments. As such I have no criticism of the work, however, deciding whether the findings are obvious or novel is what I am struggling with. If the evolutionary pressure is there to favor a single locus mating type in a primarily selfing population then it is perhaps obvious that convergent evolution would lead to this situation, that is the situation could arise by loss of one or the other mating type loci. The loss of PR gene function in mating has been previously reported and this represents the first described instance where a single mating type locus arose by loss of the HD function as expected. Another interpretation is that of course this is obvious now that it has been found and I must admit that I cannot separate whether this is would have been obvious without this observation. Regardless it is an important finding and very worthy of note. I commend the authors for this solid investigation

In reading the manuscript I found several areas where wording could be edited to improve clarity. I indicate these below
Lines 52-54 The sentence beginning "The control of mating …" should be broken in two or otherwise edited since including several points in a run-together manner leads to a weak statement and breaking it up could strengthen this important point.
Line 59 – The sentence beginning "Mating systems .." is awkwardly worded and editing could improve this. Consider, is 'corresponding' is the right word in this context
Line 60 Sentence beginning "A wide diversity …" contains 'systems" three times, edit to reduce the redundancy in word usage.
Line 68 the phrase "Being then called" is awkward and could be edited to also remove te redundancy in the use of 'being'
Line 71 suggests that recombination ensures proper function but I do not think the authors intend to say this since many factors are involved in the proper function of mating types (expression, protein structure, protein-protein interactions etc), do they mean ensures that expression of the loci are properly maintained or something else, edit to clarify or delete
Line 74 delete "still"
Line 108 "as diverse fields as" seems unnecessary and could be deleted.
Line 142-43 edit to clarify eg. " and a part of the ancestral HD chromosome that does not contain the HD locus."
Line 157-158 Consider editing the title to make a statement not an open-ended comment
Lines 181-182 Delete "however", clarify what you mean by "and conversely"
Lines 200-203 In the sentence starting "Indeed, …" consider editing the sentence to remove stating 'recombining region' twice and to improve clarity
Lines 208-212 The structure of the sentence starting "The black stratum.." is awkward and thus not clear Edit to improve

clarity
Line 292 what does this confidently sequenced mean? Edit to clarify, researcher 'confidence' should be based on data so state the data that provides the confidence instead of just stating confidence.
Line 317 the phrase ‘and also with’ does not make sense in this context edit for improved clarity
Line 443-444 The hypothesis indicates that heterodimers are no longer needed since a homodimer can provide the appropriate function, does the structure (sequence) of the HD proteins shed any information as to how to determine if a homodimer would have a different function than a heterodimer? Are there perhaps differences in DNA binding domain sequences or other regions?

Line 632 the use of ‘a priori’ seems out of place in this sentence, consider editing for improved clarity.
Line 747 the phrase “with therefore likely” is awkward, edit for improved flow and clarity.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The revised manuscript by Lucotte et al. is much improved. All my concerns have been addressed. In particular, the addition of results from experimental work was essential, and the reworked Figure 7 is much clearer than the previous version. The addition of statistical testing of TE family enrichment made the conclusions by the authors more robust. Perhaps these results could have been summarized in Figure 8 or a table, since they take up a lot of space in the main text, but this solution is also acceptable.

(Remarks on code availability)
See initial review for my comment of the code.

Reviewer #2

(Remarks to the Author)
In reviewing the revised manuscript, I find that the authors have generally addressed the reviewers’ comments, resulting in a clearer and more well-structured presentation of their findings. The paper is methodologically sound, with robust genomic analyses that are well-documented and should be reproducible by others in the field. The results are generally interpreted thoughtfully, with appropriate support from the data, and are placed into the context of existing literature on fungal mating systems, highlighting the study’s contributions and implications for further research. The figures and tables now better support the narrative. Overall, the manuscript makes a relevant contribution to our understanding of mating-type gene evolution in basidiomycete fungi and represents a novel observation in the field, particularly regarding the loss of function of HD genes in mating-type determination. I only have one main comment and a few minor points the authors may want to consider addressing at their discretion. The line numbers correspond to the tracked version of the manuscript.

Main Comment

The proposed hypothesis of a complete fusion between the ancestral PR and HD chromosomes, followed by rearrangements and a subsequent fission event, raises questions that, in my opinion, are difficult to reconcile. Specifically, if the rearrangements involved relocating genes from the ancestral PR chromosome into regions corresponding to the ancestral HD chromosome, this would imply significant restructuring after the proposed fission event. It remains unclear how collinearity of the HD chromosome could have been re-established to match the ancestral organization following extensive rearrangements. This aspect of the hypothesis seems counterintuitive, as a fission event after such rearrangements would likely leave evidence of disrupted synteny rather than a return to ancestral collinearity. Could the observed reshuffling of some of the genes instead be explained by their already different, possibly lineage-specific, positions on the ancestral chromosomes prior to fusion, without requiring a complete chromosome fusion?
This interpretation would also avoid the complexities of resolving a dicentric chromosome upon fusion and re-establishing a centromere after fission.
Additionally, the pseudoautosomal region (PAR) at the right end of the PR chromosome aligns intriguingly with the right end of the ancestral HD chromosome. This observation raises questions about whether more localized events, such as translocations, rather than complete chromosome fusion, could better explain the observed gene distributions. Finally, the apparent expansion of recombination suppression within the PR-containing chromosome, as suggested by the young evolutionary strata, might indicate that recombination suppression is actively extending in this region, despite the absence of linkage to the HD locus. This possibility could provide an alternative explanation for the observed patterns without invoking complete linkage to the HD locus or a fusion-fission cycle. Could the authors briefly explore these alternative scenarios or provide evidence to clarify how the current data support or rule out these mechanisms?

Minor Comments

Line 41: Consider changing “with a part of” to “with a segment of”

Line 44: Remove the comma after “fusion” (should read: “whole chromosome fusion and, unexpectedly, the…”).

Line 45: Consider changing this sentence for smoother flow to: “…the HD genes lost their function in mating compatibility, with natural diploid strains often being homozygous at the HD locus”.

The loss of function of HD genes in mating compatibility is stated three times in the abstract (lines 44-45, 47-48, 49-50). Consider streamlining and integrating these statements, perhaps by decoupling the fusion events from the loss of HD gene function in mating compatibility.

Line 86: I think that Ref. 34 is not appropriately cited here.

Line 95: Consider adding that the products of HD1 and HD2 heterodimerize exclusively with non-allelic counterparts from different strains to form an active transcription factor. This mechanism avoids self-recognition, as HD1 and HD2 proteins from the same strain cannot interact, thereby promoting non-self-interactions and ensuring compatibility between genetically distinct strains.

Lines 104-118: Streamline this section to avoid repetition of “Microbotryum fungi.” Use alternatives like “these species” or “these fungi.”

Lines 119-120: The phrase “as neither ever recombines” could be misleading, as it might suggest that the entire chromosomes do not undergo any recombination. To avoid confusion, consider rephrasing to: “as neither of these regions ever recombines.” As indicated somewhere in the manuscript, pseudoautosomal regions of these chromosomes do undergo recombination, and this is typically required for proper chromosomal alignment and segregation during meiosis.

Lines 135-138: Suggested phrasing: “Focusing on two of these species, M. superbum and M. shykoffianum, our aim was to characterize their mating-type chromosomes, elucidating whether they displayed recombination suppression and, if so, to ascertain whether the mating-type loci were linked together or if each was linked to its respective centromere on separate chromosomes.”

Lines 144-145: Consider revising to: “The HD2 gene appeared disrupted in several strains of M. superbum, was homozygous in diploids, and showed low expression levels under mating conditions”. Additionally, it is interesting to see some level of expression of the full transcript, suggesting, for example, incomplete degradation of transcripts by NMD.

Lines 191-192: Clarify whether the “missing” region refers to a failure in assembly or a genuine difference between the two chromosomes.

Lines 204-212: Explicitly state that the LD analysis was only done for one species (as indicated in Figure S8) and clarify its presentation as an example. Rephrase “We used SNPs on 149 genomes” to: “We used SNPs identified across 149 genomes…”

Lines 303-305: Suggested simplification: “The HD2 gene contained an additional intron compared to the functional M. lychnidis-dioicae HD2 gene (Figure 7A) and was shorter due to a deletion in the highly variable N-terminal heterodimerization domain, which is challenging to align across species”.

Line 306: Clarify why the authors suggest the M. scorzonerae genome assembly is poorer than the other species despite metrics in Table S1. Also, do you know whether the ~50% smaller genome size reflects contraction in M. scorzonerae or expansion in the other species.

Line 410: Simplify “…encompassing multiple genes linked one to each other” to “…linked together”. I guess it depends on how you define the HD locus, but typically it comprises only HD1 and HD2 genes.

Line 412-414: Rephrase to: “One mating-type locus yields 50% compatibility among gametes of a given diploid individual under random mating”.

Figure 1: Not sure if I’m interpreting this correctly, but the schematics in Figure 1 seem to suggest that only a segment of the HD chromosome (excluding the HD genes) fused with the PR chromosome. This seems to contradict Figure 3, which proposes whole-chromosome fusion followed by subsequent fission. Could the authors clarify whether these two figures represent different interpretations of the fusion and fission events, or if additional details are needed to reconcile these depictions?

Figure 7: Revise the figure legend for consistency in panel labeling (e.g., “A)” vs. “A:” vs. “B.”). Correct typos (e.g., “t-wo”) and missing spaces (e.g., “M. lychnidis-dioicae.The diploid”).

Figure S2: Add a scale bar to maintain consistency with other similar figures.

Figure S4: Clarify in the legend whether the PR gene location in the a2 M. superbum strain overlaps with the pheromone gene and is therefore not visible.

Data Availability, Please ensure that the BioProject and raw data associated with this study are publicly available, as I could not locate them.

(Remarks on code availability)
The specified repository provides open-access code, which appears to be well-documented, but I have not personally tested or run the code.

RESPONSE TO REVIEWERS

Reviewer #1 (Remarks to the Author):

I read the manuscript by Lucotte et al. with pleasure and a fair bit of excitement, since they present the first case of loss of mating-type determination at the HD locus in a basidiomycete fungus. This scenario has been thought of for a long time, but in every previously studied case of transitions from a bi- to a unifactorial mating system, the loci have either been linked, or the PR locus has lost mating type determination function. The authors base this novel finding on thorough genomic analyses of different lineages of anther-smut fungi belonging to the genus Microbotryum, and accompany their bioinformatics methods with experimental work that bring a crucial component to their study. The experimental and genomics methods are robust and detailed, and I cannot identify any major methodological issues.

>>We thank the referee very much for the positive comments!

However, the presentation of these findings can be improved by clarification in several sections of the text, and some details should be added to the figures to help the reader interpret them.

In particular, the presentation of how the HD2 gene in M. superbum was deemed to be non-functional requires some additional work. Viewing Figure 7, I agree that the HD2 gene(s) look suspicious. However, the authors need to make sure the splitting of the gene is not an artifact from gene annotation. In the "Mated24" track of RNA-seq coverage, there is actually coverage across the whole site, which may indicate an intact gene. The authors explain in the text that they identified new start/stop codons, but the annotated reading frame starts before any RNA-seq coverage, which leads me to question that the correct reading frame is annotated. A track of mapped DNA-seq reads across this whole region would also help the reader ensure that there is no problem with the assembly.

>>The strain used for RNAseq was different from the reference genome strain, which indeed rendered the figure unclear. We have produced a new assembly for this strain, based on ONT reads, and predicted the genes with the same pipeline. RNAseq reads indeed covered the entire region, but the intron appeared perfectly processed and thus the coding sequence will stop prematurely. We have redrawn figure 7 and added the corresponding explanations in the manuscript (L232-246). We have removed analysis of RNAseq *in planta*, as they were performed with different strains than the reference genomes and the read mapping appeared less reliable. We now focus on the 'mating' condition, which also allows clarifying the figure and does not change the conclusions.

The main text (lines 226-239) describing this crucial result is also in need of improvement. I think a better term for the two HD2 fragments would be gene models, since CDS implies that they are exons of the same gene, which is not the case according to the current gene annotation. "One of the two novel CDS appeared expressed during mating conditions", actually there is RNA-seq coverage across both gene models.

>>We have changed the terms CDS to gene models and we have clarified the text. In particular, the two predicted gene models at the locus homologous to HD2 are expressed at much lower levels than the HD1 gene in the "mating" condition, while they are expressed in even amounts in *M. lychnidis-dioicae* (L232-246).

"This may be because the strain is homozygous", how can a haploid strain be homozygous?

>>We meant the diploid strain used for inoculation, but we have completely rewritten this section (L232-246).

"and M. intermedium", this species is not in Figure 7. Is it supposed to be? Otherwise the authors should refer to the figure after M. lychnidis-dioicae, and put "data not shown" after M. intermedium. If they have the data, which the manuscript gives the impression of, they should show it to corroborate their results.

>>We have deleted this statement. We do have expression data in *M. intermedium*, but not of opposite HD alleles, so not really useful to support these sentences.

Furthermore, the authors identify much larger PR chromosomes in M. superbum and M. shykoffianum, and state that these chromosomes have accumulated specific families of TEs. That the TE load is higher than on the autosomes is quite clear, but to be able to say that specific TE families have preferentially accumulated, the authors need statistical support. On lines 482-485 preferential accumulation seems to be taken as a fact, but seems to be based only on what is shown in Figure 8. From the figure I can accept that LINEs may have preferentially accumulated in the light ruby stratum, which is not mentioned in the text, but the other TE families are also common on the autosomes. The authors have to test this statistically before they can say anything about preferential accumulation of specific families.

>>We thank the reviewer for this comment. We have added statistical tests on the density of TEs in the PR chromosome versus in the autosomes. To perform such a comparison of an individual value (the PR chromosome) to a small sample (the different autosomes), we used the Crawford and Howell's t-test, and we computed z-scores to illustrate the deviation of the density in the PR chromosome from the autosomal average. We have indicated in the text the transposable element families for which the PR chromosome had the highest density, and the families that do not show any expansion in the PR chromosomes, and for each of these families, we have provided the corresponding z-scores and Crawford and Howell's t-test p-values (L389-400).

The authors need to show data from the crossing experiments. As it stands, these results are only described in text (in the paragraph starting on line 240, and on lines 302-305). To enable the reader to make their own interpretation of the data, a supplementary table or figure should be included showing the results from the crosses.

>>We have added the results in Supplementary Table S4.

To make the message of the paper as clear as possible, some improvements should be made in how the results are visually presented. Near the end, I outline some issues for each figure.

>>We thank the referee very much for the thorough comments and we tried to improve the figures as much as possible.

Minor comments:

Line 40 and throughout the text: Microbotryum scorzonerae is often misspelled. I assume it should be the taxon with taxid 72746 on NCBI throughout the text (I find no taxon called M. scorzonarae or M. scorzoraneae). Please check the text and correct this.

>>Thanks for pointing to these errors, they have been corrected.

Line 159: The term ”high-quality” regarding genome assemblies can mean almost anything depending on the organism and context. A pointer towards what the authors mean here would help the reader orient themselves. I assume the authors mean to say that the assemblies are of high contiguity, allowing the analysis of physical linkage between MAT loci, but some numbers (maybe N90?) would help.

>>Thank you for this suggestion, we have added statistics about the assemblies in Table S1 including N90. We have changed “high-quality” by "highly contiguous” L157.

Line 159: “from one strain from each of two” is this not just the same as “from two species”?

>>We have clarified that we meant here to specify that we had two haploid genomes from a single strain per species (L161).

Line 162: “as expected”, based on what? It is not necessary to mention any expectation here.

>>This statement has been deleted.

Line 173: What does a1 and a2 mean? I assume they refer to mating types. A short note in the results about how they were obtained would help the reader, without having to go into methods details. I suggest adding this explanation on line 163, before “In all four haploid genomes”.

>>Yes, a_1 and a_2 refer to mating types. We have added an explanation (L161).

Line 176: How was the centromere identified? This objective is non-trivial in most species, and some explanation should be given. Centromeres are known to be able to change even in short evolutionary time. A better, more careful, wording would be “predicted” instead of “identified”. Precise determination of the centromere would require additional data.

>>We have clarified that centromeres were only predicted, L175 and in the M&M where the method for their prediction is explained (L675).

Lines 192-193: The wording of this sentence is unclear. “SNPs on 149 Illumina genomes” makes little sense to me. I think the authors mean to say that SNPs were identified from 149 genomes sequenced with Illumina technology. They should also state which species have been investigated. In the previous sentence, they refer to these samples as “collected on Dianthus plants”, but it is not clear whether these samples correspond to one or several species of Microbotryum, and how they are related to the strains they based most of their study on.

>>We have clarified the text here (L206-207) and in the introduction (L134-135). The full analysis of this Illumina dataset will be reported in a separate paper, with analyses on the population structure, species distinction, host species, selection and diversity, which is beyond the scope of this manuscript. Note that we now use SNPs only for linkage disequilibrium analyses, and we have assembled three genomes for the HD gene analyses to phase haplotypes.

Lines 195-196: Here some numbers would help interpretation. “A bit higher” is not very specific.

>>We have added the mean r^2 value for each of the genomic regions mentioned to the text (L208-212).

Line 216: “the mapping of diploid reads”, this wording is incorrect. The reads are not diploid, but originate from a diploid genome.

>>We have modified this sentence (L186), thank you.

Line 292: what does “exhibited polymorphism” actually mean? Would a single SNP in the HD locus count? Some numbers are necessary here, I suggest pi for average nucleotide divergence, but raw SNP number per strain would also help.

>>We have finally only retained three strains from the Illumina dataset for this analysis, as the mapping of Illumina reads did not allow distinguishing alleles clearly enough. We only retained the three strains that we could *de novo* assemble reasonably well to obtain phased haplotypes and therefore reliable allele sequences; see L790-802 and 290-296.

Line 307: What is meant by “gap” here? I think the authors mean a deletion, but gap brings to mind an assembly gap in the form of a stretch of N’s. They also do not refer to any figure describing this finding. The finding of convergent non-sense mutations in the HD2 gene in these different lineages is highly important and should be shown. The authors should consider adding another panel in Fig. 7, or at least a supplementary figure if this is not possible (e.g., maybe they do not have RNA seq of M. scorzonerae).

>>We have rephrased these sentences as there is indeed a very small intron in *M. scorzonerae*. Thanks for this suggestion! The protein is shorter and there may well be a deletion, but in a variable part of the protein that is impossible to align across species, so it is difficult to be certain. We have added the panel Figure 7A to show the structures of the HD1 and HD2 gene models in the various species, and changed the text accordingly L301-305. We do not have any RNAseq data for *M. scorzonerae.*

Line 400: which figures?

>>We have added the figure number, thanks for spotting this mistake (L403).

Line 450: Scorzonera is misspelled.

>>This has been corrected.

There are some typos, faulty sentences, and missing references in the Methods section. I point out some errors I find below, but I suggest also another round of proofreading before the next submission.

>>We have carefully checked the text.

Line 586: Some detail is missing about the Zymo kit.

>>We have added the information of the Zymo kit reference (L583).

Line 611: Some detail is missing about the Qiagen kit.

>>We have removed this part because we do not use RNAseq *in planta* data anymore (see answer above: "We have removed analysis of RNAseq *in planta*, as they were performed with different strains than the reference genomes and the read mapping appeared less reliable.".

Lines 622-628: This sentence is impossible to read. Please rephrase.

>>Sorry about this copy-paste mistake, we have corrected the sentence (L602-604).

Line 624: Here it is stated that mating was assessed after 48h. In Figure 7 the legend says after 24h. Which is correct, or is this not the same data? If not, where can the reader find the methods underlying Fig. 7?

>>We have corrected the legend of Figure 7, 48 h is correct, thanks for spotting this mistake.

Line 641 and in various places in the text: Make sure species names are correctly italicized.

>>Corrected here and in other various places.

Line 647: Missing reference/link.

>>We have added the reference.

Lines 746-748: This sentence is speculative. The authors could address the matter of divergence between a1 and a2 HD alleles by mapping back the raw reads and calling variants in this region.

>>Mapping the reads to the genome confirmed that there were no SNPs in this region, thanks for the suggestion. We have added this information in the methods (L732-734) and the main text (L285).

Figure 2: Reading the figure as it is first referred to in the main text, "a1" or "a2" have not been explained.

>>We have added an explanation of a_1 and a_2 at the beginning of the result section.

The title of the figure says a1, but in the figure itself a2 is annotated. Which is correct? Is the figure drawn to scale? If so, the legend should mention this, and there should be a scale bar to

orient the reader. Can the telomeres be indicated in the figure? This would give the reader a sense of the assembly quality.

>> Thank you for spotting this mistake. It has been corrected. The figure is indeed drawn to scale, which has been added to the legend. We have also added the predicted telomeres and centromeres, as well as the position of the HD and PR genes and the pheromone genes. We have added a scale bar to each Rideogram, and we have added a supplementary table with the length of all mating-type chromosomes included in this study (Table S8).

Figure 4: The red vertical lines in panel A are not explained. Panel C: it is unclear to me what is shown here. What is on the y-axis? What are b1 and b2, referred to in the legend but not present in the figure? In the legend, b1 is not written in subscript but a1, a2, and b2 are. Is this correct? The x-axis label: the species name should be written as M. lagerheimii, i.e. with italics and lower-case "l".

>> Thank you for spotting these mistakes. We have now more explicitly explained the legends of Figures 4 and 5. Indeed, panel C represents: "Rearrangements compared to the ancestral gene order illustrated by plotting the gene rank in the current gene order as a function of the gene rank in the ancestral gene order." The b1 term was an old annotation that has since been removed. We have italicised the species name.

Figure 6: Panel A: an insert showing which photograph belongs to which species would be a useful reminder here.

>> Thank you for this suggestion, we have added a panel to figure 6.

Figure 7: Some cosmetic improvement would help the reader. Flipping the direction of the panel for M. superbum (upper) would improve the ability to interpret the figure, so that the HD1 and HD2 genes are in the same orientation as for M. lychnidis-dioicae. The details and text are also quite small, so I suggest removing some of the flanks around the HD locus, perhaps 1.5kb on each side, and zooming in on the HD genes.

>> We have completely redrawn the figure 7 on expression of HD genes.

Figure 8: Some modifications would improve clarity here. It is hard to get a general idea about the PR chromosome since the different strata are divided up. My interpretation is that it would be very similar to "Black PR" since it is so big compared to the other strata and the PAR. At least, the categories making up the PR chromosome should be indicated as belonging to this chromosome. The contrast between "Unclassified" and "Satellite" is very low, so it is hard to determine the category of the largest proportion of TEs. Also, satellites, rRNA, and simple repeats are not TE categories and should not be called as such on the x-axis of panel A.

>> We have improved Figure 8 accordingly. In addition to the proportion of repeated elements in each stratum and autosome, the new version shows the proportion of repeated elements in the entire PR chromosomes and in average in autosomes. The color code has been modified so that "Unclassified" and "Satellite" elements are distinguishable. The X axis of panel A has also been changed.

Figure S9: The authors should state which species is infecting the flower in the figure legend.

>> We have added a description of the picture in both the figure and figure legend (now Figure S10).

Table S1: The orientation of the primer sequences should be stated in the legend or directly in the table.

>> The orientation was added to the Table (now Table S5).

Table S3 contains "TBD" for sequences generated during this study. Please update with current accession numbers. There are also empty cells for samples generated during previous studies. Were these data not deposited? If so, this could be indicated in the table. The authors should also consider submitting these data for transparency.

>> All sequences will be under the same Genbank bioproject PRJNA1169769, but we are currently experiencing delays in the submission because of the hacking that disrupted all University Paris Saclay services. The authentification by Genbank was blocked and therefore it was impossible to submit the sequences on time. Details on the samples will still be available in Table S7.

A recurring language issue throughout the manuscript: in figure legends the authors use the word "figured" as a synonym of "visualized". I have not encountered this use of the word before; please confirm with a native English speaker that it is linguistically correct.

>> This term has been replaced by indicated or illustrated.

Reviewer #1 (Remarks on code availability):

There is a vast resource of open access code in the specified repository. I have visually reviewed some randomly selected scripts, but not tested the code myself. There is ample documentation about the code.

Reviewer #2 (Remarks to the Author):

This study by Lucotte et al. investigates the mating systems and genomic architecture of anther-smut Microbotryum fungi, with a specific focus on the lack of linkage between the HD and PR mating-type determining loci, possibly associated with the repeated loss of function at HD mating-type genes. By leveraging high-quality genome assemblies of additional sister Microbotryum species isolated from Dianthus and a distantly related species, they provide a comprehensive analysis of how these genomic configurations may have evolved. Specifically, their research reveals that the HD and PR mating-type loci, currently residing on different chromosomes, have evolved from an ancestral state where these loci were likely linked, undergoing subsequent chromosomal breakage. These findings underscore a complex evolutionary landscape where mating compatibility is apparently controlled by a single genetic

factor due to the loss of function in HD genes, which is a novel observation in heterothallic basidiomycete fungi.

The manuscript offers valuable insights into the evolutionary mechanisms that shape fungal mating systems, elucidating the complex interplay between genomic architecture and reproductive strategies and should be of broad interest to the evolutionary genomics community. While the findings are significant, the presentation could benefit from improved clarity and streamlined exposition. Specific areas needing attention include typographical errors, inconsistent formatting, and the need for better-detailed explanations in the results section that are currently detailed in the methods section. Additionally, while some figures offer valuable information, others require enhancement to effectively convey the underlying data. Addressing these points, along with my specific comments provided below, will greatly enhance the manuscript's readability and scientific impact.

>> We thank the referee for the positive assessment and we have tried to clarify the manuscript.

- Main comments –

On the formation of the large non-recombining PR chromosome:

1) The authors propose that the ancestral PR and HD chromosomes in M. superbum and M. shykoffianum fully fused, followed by extensive recombination suppression and intrachromosomal rearrangements linking the PR and HD loci. However, the observed conservation of synteny in the extant HD chromosomes compared to M. lagerheimii (representing the ancestral state) raises some questions about this model. It could suggest a potential alternative scenario where the “fusion” might have involved only a segment of the ancestral HD chromosome that excludes the HD genes. Such a scenario could involve a fission in the ancestral HD chromosome, followed by the fusion of the acentric fragment with the PR chromosome, while the remaining portion could have been stabilized by the de novo formation of telomeric sequences. This hypothesis could potentially explain the conservation of synteny of the extant HD chromosome.

>>We had considered this scenario, but it seems contradicted by the presence of genes on the current HD chromosome that were ancestrally in the PR chromosome, and vice versa, which suggest a brief linkage (and a few rearrangements) followed by a fission. We have tried to clarify this point, while presenting the suggested alternative scenario (L469-471).

2) Assuming the proposed fusion resulted in a dicentric chromosome, the likely genomic instability would necessitate resolution mechanisms. These might include breakage-fusion-bridge cycles, centromere inactivation (either epigenetic or structural), or chromosome breakage followed by repair to restore monocentric chromosomes. Given the hypothesis of HD and PR loci linkage, it seems plausible that one of the centromeres was inactivated immediately following fusion. Subsequent to any proposed fission event, centromere function would need to be re-established. Could the authors provide insights on how these dicentric conflicts might have been resolved in these species?

>>We agree this is a very interesting question, but we think that adding possible explanations at this point would be highly speculative, as we have no indication supporting any particular scenario or mechanism. Functional evidence is needed to favour one or the other scenario, but is out of scope of this paper.

3) In basidiomycetes, the PR locus often expands, possibly due to the duplication or even quadruplication of pheromone genes, which can promote intrachromosomal rearrangements (e.g. inversions). Therefore, I think it would be informative to examine the position of the pheromone genes within the PR chromosome. Clarifying whether these genes contribute to the expansion of the PR locus, even in the absence of linkage to the HD locus, could provide additional insight into the genomic dynamics and evolutionary pressures shaping these regions.

>> Thank you for this suggestion. We have added the pheromone gene positions to the Rideogram figures (figures 2, S1, S2 and S3) and in the circos (Figures S4 and S5). We find two copies of each pheromone gene in each of the *M. lagerheimii*, *M. shykoffianum* and *M. superbum* PR chromosomes. We could not identify the pheromone genes in *M. scorzonerae*, most probably because it is distantly related to the other species studied, and they are small, rapidly evolving genes. A more complete study focusing on the identification and the evolutionary dynamics of these genes would be needed to address this question adequately. However, the pheromone genes alone are unlikely to contribute to expansions or the PR locus, as they are ancestrally linked to the pheromone receptor genes in basidiomycetes, so no further recombination suppression is expected to be selected for. We have added a sentence in the result section.

On the RNA-seq and polymorphism analysis of the HD genes:

1) The RNA-seq read coverage presented in Fig. 7 leaves it unclear whether the HD2 gene of M. superbum is expressed as a single transcript. Furthermore, the lack of synteny between HD2 gene fragments in a1 and a2 strains requires clarification. Could you specify the method used to assess synteny in these plots?

>> We agree that the figure was not clear and we have performed further analyses and completely redrawn figure 7.

2) Figure 7 references “Mated24”, suggesting RNA samples were obtained from a1 and a2 cells undergoing mating for 24 hours. However, Table S3 and the methods section refer to “Mating 48h”, indicating a potential inconsistency. Additionally, the presence of SNPs and unequal read coverage along the HD1 gene in the “mated24” condition is suggestive of allelic variation. Could you please clarify these observations?

>> For the RNAseq experiments we used a different strain, which was found homozygous at the HD genes. We have assembled its genome and now present the results of read mapping on the new assembly. We have completely redrawn figure 7. The “mating” sample presented was obtained at 48h, sorry about this mistake.

3) A few enhancements in displaying RNA-seq data are warranted, particularly regarding consistent scaling or clear indication of scale used. Further details are also needed on whether the RNA-seq data was mapped to a combined reference genome encompassing both PR and HD mating-type chromosomes.

>> We have completely redrawn figure 7, we hope this is clearer now. Details of the mapping are provided in the method section, we have mapped reads on a single haploid genome per

species, to which we added contigs imputed to the mating-type chromosomes of the opposite mating-type haploid genome.

Additionally, please clarify whether the RNA-seq data were paired-end or single-end and the choice of minimap2 for read mapping over other RNA-seq mappers that are more adept at handling splice junctions.

>>We have clarified in M&M that all RNAseq were paired end. We have remapped the reads with STAR, which specifically handles splice alignments.

4) It should be noted that even if the HD2 gene in the a2 mating type of M. superbum is non-functional, the HD2 gene from the a1 mating type could still potentially form a functional heterodimer with the HD1 gene of the a2 mating type. This phenomenon has been observed in other species with PR-HD linked loci, where typically one mating type loses one HD gene, while the other retains the opposite HD gene, demonstrating an important aspect of mating-type gene functionality and evolution.

>> We agree in principle, but *M. superbum* strains with homozygous disrupted HD sequences can mate and produce disease, so this hypothesis seems unlikely here. We have clarified this point L469.

5) I may have missed it, but I couldn't find any results related to RNA-seq data analysis for M. superbum haploid condition or for M. intermedium.

>> We have deleted this statement. We now focus on the "mating" conditions and compare the expression level of HD genes from *M. superbum* (with nonfunctional HD genes) and *M. lychnidis-dioicae* (with functional HD genes).

Additionally, regarding M. shykoffianum c212, the genome appears to have been sequenced from a mixture of meiotic products from presumably the same diploid individual, showing that the HD1 and HD2 sequences are intact and seemingly 100% identical. Was this observation independently validated through PCR/Sanger sequencing using separate a1 and a2 haploid sporidia? Alternatively, examining HiFi reads might reveal any variation when mapping to the HD chromosome.

>> We have checked in the HiFi reads and found no polymorphism at all (we have added this point L285). We do not have the separated haploids for this strain, this is why we sequenced this strain as a mixture of haploids.

6) Could you please clarify whether the 149 Illumina genomes were obtained from multiple diploid individuals collected from different flowers?

>> We have clarified L207 and L749 that they were collected from different plant individuals, but we have finally retained only three genomes from this dataset for the HD polymorphism analysis.

Given the considerable polymorphism observed at the HD genes, it would be informative to understand how the SNP patterns are distributed along these genes, particularly whether there is more variation at the corresponding N-terminal regions, as is common in other basidiomycetes. Additionally, contrasting the level of heterozygosity of HD genes in these

species with that of species where HD genes still apparently determine mating could also be relevant.

>>Thanks for this interesting suggestion. We fully agree this should be the next step, but we would need phased SNP data for a thorough analysis. We have tried here to *de novo* assemble the genomes sequenced with Illumina, but found only three that could be assembled well enough for such analyses. We have changed the manuscript accordingly.

7) In M. scorzonerae the 19bp gap noted in the HD2 coding sequences needs a baseline for comparison. Additionally, has there been any attempt to predict an intact CDS or to rule out additional introns?

>>We thank the referee very much for this very useful suggestion. The gap was indeed a very small intron, so we have rephrased this sentence. The protein is still shorter than the functional one from *M. lychnidis-dioicae*, but it is difficult to say whether the new intron corresponds to a deletion in the protein as this part is challenging to align. We therefore remained cautious, only stating that the protein was shorter and with a new intron. Many thanks for the comment!

8) While not essential for the conclusions of this study, future research could benefit from considering targeted gene deletions of the HD genes in some of these species. This approach would not only definitively confirm whether these genes are necessary for mating-type determination but could also elucidate their roles in other aspects of sexual reproduction, i.e. maybe the HD genes are not involved in mating-type identity but are still required for sexual reproduction and thus retained to some extent.

>>We agree this would be highly informative, but this is still not technically possible in *Microbotryum* for now.

On the tetrad analysis section:

1) I think it would be beneficial to include mating test results for all 48 meiotic products from the 12 tetrads collected from D. pavonius. Presenting this data in a supplementary table could provide a better overview of the mating compatibility outcomes.

>> We have added a Table (Table S4).

2) The authors should also consider including supplementary data displaying the results from PCR analyses that confirm the presence of varying PR alleles and uniform HD alleles among the meiotic products. Additionally, providing photo examples of conjugated (mating compatible) and non-conjugated (mating incompatible) cases, along with visualizations of hyphal production, would effectively support the assertion that HD genes no longer influence mating-type determination in these species. Likewise, a detailed account of disease symptoms in plants inoculated with strains that have identical HD alleles would also be helpful. I think this would strengthen this unique observation among basidiomycetes for all readers but also the study's groundbreaking conclusions.

>> We thank the referee for the suggestion. We have added the available pictures: a PCR gel picture and hyphae in the new Figure S9, but we did not have pictures of symptoms for the strains used in this experiment; we have added a sentence stating that symptoms looked perfectly normal, with anthers full of spores and ovary reduction L273-274. We do not have

pictures of either conjugating or unconjugating spores from this very experiment. We have also added pictures of the meiotic products that produced hyphae without tester strains (Figure S9).

3) For the two meiotic products that produced hyphae without tester strains, have the authors assessed if they are fully diploids or aneuploids? Interesting that they don't behave as the "solopathogenic" strains of certain Ustilago spp. Any idea why?

>>This is indeed highly interesting, but we have not checked the ploidy and do not know why they do not behave as full solopathogenic.

4) For the 6 additional isolates, where 2 strains are homozygous and 4 are heterozygous for the HD gene, you could could consider inspecting the SNP distribution along the HD genes in the heterozygous. It would be informative to enumerate the changes in the predicted protein and compare these with typical patterns observed in basidiomycetes and in other Microbotryum species where the HD genes are still involved in mating-type determination. In other words, would be great to have an idea about how many changes are sufficient for HD alleles to be functionally different.

>>We thank the referee for this interesting suggestion. As the HD genes have likely lost their mating-type function before the divergence between *M. superbum* and *M. shykoffianum,* the HD genes have probably accumulated many substitutions, as suggested by the selection analyses. The heterozygous versus homozygous status of strains is likely not an indication of some alleles being functional and others not. It is therefore not possible to identify the precise changes or the number of changes for the loss of function with the available data. The loss of function may even be due to changes in other genes that could have allowed to bypass the need of HD genes, that would then have accumulated mutations only by relaxed selection, without selection for loss of function at the HD genes themselves. We will try to analyse in more detail in a future study the patterns of polymorphism and selective pressures on the HD genes, and the changes on the functional domains, but this will require a substantial amount of work beyond the scope of the present study.

- Minor comments –

Lines 66-68: Consider breaking up and rephrasing for clarity: "In fungi with mating types, gamete compatibility is controlled during the haploid stage, where only cells of different mating types can successfully mate. Mating types are determined by either one or two loci, classifying the systems as uni- and bifactorial, respectively."

>> This has been corrected L66-69.

Lines 72-74: The statement "Recombination is typically suppressed at loci determining sexes and mating types, which ensures proper function in these complex phenotypes" could benefit from further clarification on how exactly recombination suppression contributes to the proper function of these phenotypes.

>> We have clarified this sentence L73-74.

Lines 100-101: Introduction of unifactorial basidiomycetes resulting from HD-PR linkage

could include references to Cryptococcus pathogenic species where extensive studies exist to provide a broader perspective on the diversity across basidiomycetes.

>> We have added a reference L100, thanks for spotting this.

Line 101: Please add a period after 'spp' to correctly punctuate the abbreviation for species.

>> This has been corrected L100.

Line 105-106: Suggest adding a comma after “plants” to improve the clarity of the sentence structure.

>> This has been corrected L105.

Line 117: Should it be intra-tetrad mating?

>> “Intra-tetrad selfing” is also correct, but we have changed to “intra-tetrad mating” to be consistent L115.

Line 136: Consider changing to “carry” and "yet display"

>> This has been changed L133.

Line 141: Consider rephrasing “issued from the fusion” to “resulting from the fusion”.

>> This has been changed L139.

Line 142: A "the" is missing before the HD locus.

>> This sentence was modified (L138-141).

Lines 144-145: The description of the translocation process as occurring “via a transient stage of whole chromosome fusion” is not clear.

>> We have clarified L141-143.

Line 159: It would be helpful to have some statistics on the genome assemblies produced in this study (genome size, number of contigs, etc.) as it is currently unclear how well assembled they turn out to be. This could be provided as supplementary data.

>> We have added the Table S1 with statistics on genome assemblies.

Lines 164-166: Although it is stated that telomeric repeats were identified in almost all of the

mating-type contigs, only 5/8 telomeric repeats for M. superbum and 4/8 for M. shykoffianum are accounted for. Therefore, maybe consider instead emphasize that at least one set of PR or HD chromosomes were completely assembled telomere-to-telomere in each species.

> We have clarified this sentence L178-182. Thank you for this suggestion.

Was there any technical reason for not having the mating-type chromosomes assembled into single contigs (read length, low coverage, too many repeats, etc.)? At least in some cases (e.g. M-sup-PR a1) the break is not within the centromere, which is usually more repetitive.

>> This may indeed be because of repeated elements; mating-type chromosomes are often not assembled telomere-to-telomere in *Microbotryum* fungi with large non-recombining regions, with breaks not only in centromeres. We have clarified this point L180-182.

Lines 176-179: Please clarify this sentence: "The centromere was identified within the non-recombining region and telomeres at the edges of the new PR and HD chromosomes (Figures S4-S5)." As noted above, telomeres were not found at all the chromosome ends.

> We have clarified this part L175-180.

It is also not completely clear at this stage if the HD-a2 chromosome of M. shykoffianum is substantially smaller than the respective HD-a1 chromosome or is not well assembled (this is only clarified further down in the text).

>> Right, we have moved the corresponding paragraph much earlier in the text, when describing the genome assemblies.

Finally, from Figure S5B, it seems there are a few links between the a1-HD chromosome and the a2-PR chromosome in M. shykoffianum: are these genes or e.g. transposable elements?

>> We agree that these links between the a_1 HD chromosome and the a_2 PR chromosome are puzzling, but they are most likely artefacts. Indeed, part of the HD chromosome is collapsed in the a_1 genome, so that no ortholog can be detected in the a_2 genome on the HD chromosome even if it existed (and it probably does as the a_1 and a_2 haplotypes are collapsed because of the high sequence identity in this region). Therefore, potential multicopy genes that would be removed by our pipeline are not removed in this case (i.e. with copies present both on the HD and PR chromosomes). We have therefore chosen to remove all the links from the collapsed part of the HD chromosome on the circos comparing a_1 and a_2 in *M. shykoffianum*, as they are unreliable. We have added a sentence in the figure legend to explain this point and we thank the referee for spotting this issue.

Line 182: Please check if Figure 3 here is appropriate, as I don't see any colored green genes in Fig. 3

>> We have clarified the figure and the legend.

Line 296: Maybe consider explaining in detail the inbreeding coefficient (FIS) and what values it can take.

>>This part has been deleted.

Lines 315-320: Why not compare M. scorzonerae with M. intermedium instead, given they seem to be more closely related? Is it because M. intermedium and M. lagerheimii are completely collinear across the genome?

>>Indeed, the genomes of *M. intermedium* and *M. lagerheimii* are highly collinear (Branco et al 2017), this is why they can reliably be considered as proxies for the ancestral state. We have clarified this point L166.

Additionally, since M. scorzonerae was sequenced using ONT, I'm curious if the authors considered inspecting the 5mC methylation landscape in the MAT chromosome compared to autosomes.

>> This is an interesting question, and this is a whole current project of a PhD student in the team. Such inferences require a lot of new data (as we found the available ONT data to be not reliable enough regarding methylation signals) as well as analyses that are beyond the scope of this study. We are not sure yet about the methylation patterns.

Lines 367-379: It is very hard to follow all these proposed rearrangements given the provided figures, in particular Fig S9. How small is the predicted non-recombining region around the PR in M. scorzonerae? Would be great to have comparisons such as those shown in the circos plots for this species as well, compared to M. lagerheimii (or M. intermedium). In general, I think the provided circos plots are much more informative as they also display other chromosomal features, e.g., centromeres, position of the mating-type determining genes, etc. On this note, would also be great if you could plot the position of the pheromone genes in these plots as the definition/delimitation of the PR loci includes both the pheromone and receptor genes.

>> Thank you for this useful suggestion. We have now added a circos representation of the genomic rearrangements for *M. scorzonerae* (Figure S11). Additionally, we have added the pheromone gene positions to the Rideogram figures (figures 2, S1, S2, S3) and the circos (figures S4, S5). To clarify the Rideograms, we have also added the predicted centromeres and telomeres, as well as the position of HD and PR genes.

Lines 414-415: As far as I'm aware, current research on Malassezia suggests that the transition from two MAT loci on separate chromosomes to a configuration where both loci are on the same chromosome (though not necessarily genetically linked) likely occurred only once. This event was followed by secondary rearrangements that relocated the loci back onto separate chromosomes.

>> This has been corrected L417.

Line 556: Is M. intermedium really an outgroup? Is the current model that this species also harbors PR and HD loci linked to the centromeres? Otherwise it is intriguing why this species retained a bifactorial system with both mating-type loci on separate chromosomes.

>> We have removed *M. intermedium* RNAseq data from our paper. To answer your question, from genomic data (Duhamel et al. 2022), we found no evidence of recombination suppression linking mating-type loci to centromeres in *M. intermedium*. We did not analyse segregation in tetrads, however, to check linkage to centromeres. We did not analyse either selfing rates in natural populations in this species either, it may be outcrossing, in which case there would be no selection to link HD and PR loci.

Line 586: check the extra "----"

>> This has been replaced by the catalogue number for the kit.

Line 1434: There is a panel or figure after Figure S8 that has no legend?

>>Indeed, thank you for spotting this mistake, we have deleted this panel.

Figure 2: As indicated above, would it be possible to also add the location of the pheromone genes since it is important to understand if they are close to the pheromone receptor or not?

>> It has been added.

Additionally, the legend references a1 genomes, yet the figure appears to display mating type a2 genomes. Please verify and correct this discrepancy.

>> Thank you for spotting that mistake. We have corrected it.

Furthermore, the PR chromosome of M. lagerheimii is depicted as split into two contigs; please clarify if these are indeed scaffolds if gaps are present, otherwise they should be referred to as contigs.

>> They are indeed separate contigs in the assembly of this particular genome, but they have been shown to constitute a single chromosome, we have clarified this point in the legend.

Also, could you clarify the nature of the additional links going from M. lagerheimii HD and PR chromosomes into other autosomes in M. superbum and M. shykoffianum? Are these links representing genes (e.g. duplications) or are perhaps annotations of transposable elements (TEs) as many appear to target the ends of the contigs?

>> The additional links going from *M. lagerheimii* HD and PR chromosomes into other autosomes are supposed to be predicted genes, as the transposable elements were removed from the dataset. However, it is possible that some of them remain after the filter. The fact that they are at the ends of contigs indeed suggests they are actually repeats. We have added a sentence in the Figure 2 legend. It could also be genes that moved from mating-type chromosomes to autosomes.

Figure 4 and legend: Please revise this figure and legend. Consider rephrasing to "…the HD chromosome on the left and the PR chromosome on the right". Additionally, please indicate what the two red dots represent in the HD chromosome plot as well as the two the red vertical lines in both the HD and PR chromosome plots.

>> This has been corrected and clarified in the legends of Figures 4 and 5.

Data availability: please ensure that all genomic data and utilized code, including that used for tree topology assessment, are made available in a publicly accessible database/repository.

>> For genome submission we experienced some delays because of the hacking that blocked the University Paris authentication by Genebank, and this is still in progress. All new data will be available upon publication under the Genbank bioproject PRJNA1169769. Homemade codes

used for the study, including the one used for tree topology assessment, have been deposited at this address: https://github.com/eliselucotte/MicrobotryumOnDianthus.

Strain nomenclature consistency: I detected a few instances of inconsistency in the naming of strains, with "MvDp−1065" used in some instances (e.g., Methods, line 795) and "Mv-Sup-1065" in others (e.g., Fig. 7). Please consider standardizing the nomenclature across the manuscript to improve readability.

>> Thank you for pointing out inconsistencies. We have standardized the nomenclature and used only the species name and their strain number (for exemple *M. superbum* 1065).

References: please double-check the references, as some, such as Reference 33, have been published and may require updating to reflect their current status.

>> We have checked carefully the references, thank you for spotting that.

Reviewer #2 (Remarks on code availability):

I have not been able to test the code, but seems to be well explained.

Reviewer #3 (Remarks to the Author):

Regarding the Lucotte et al manuscript titled "Repeated loss of function at HD mating-type genes and of recombination suppression without mating-type locus linkage in anther-smut fungi", the manuscript represents a tremendous amount of research aimed at providing insight into the genetic changes associated with mating types in the Microbotryum anther smut fungi. The work convincingly shows the rearrangement and loss of mating-type function for the HD genes. This represents the first instance in which the ancestral two mating type loci system in basidiomycete fungi is converted to a single locus system by the loss of the HD genes function and not the PR function. The manuscript also uncovers further suppression of recombination in various Microbotrym lineages. As a body of research this is well executed and appropriately controlled with the appropriate follow-on experiments. As such I have no criticism of the work, however, deciding whether the findings are obvious or novel is what I am struggling with. If the evolutionary pressure is there to favor a single locus mating type in a primarily selfing population then it is perhaps obvious that convergent evolution would lead to this situation, that is the situation could arise by loss of one or the other mating type loci. The loss of PR gene function in mating has been previously reported and this represents the first described instance where a single mating type locus arose by loss of the HD function as expected. Another interpretation is that of course this is obvious now that it has been found and I must admit that I cannot separate whether this is would have been obvious without this observation. Regardless it is an important finding and very worthy of note. I commend the authors for this solid investigation

>> We thank the Referee for these positive comments. Actually, it was far from expected that HD genes could lose their function, as a compatible heterodimer from alternative mating types was thought so far to be required for proper dikaryotic growth in basidiomycetes, as you rightly pointed out. While there have been multiple independent events of HD-PR linkage, or of loss

of PR function, there had been no exemple so far of HD loss of function.

In reading the manuscript I found several areas where wording could be edited to improve clarity. I indicate these below.

Lines 52-54 The sentence beginning "The control of mating …" should be broken in two or otherwise edited since including several points in a run-together manner leads to a weak statement and breaking it up could strengthen this important point.

>>We have clarified this sentence at the end of the abstract.

Line 59 – The sentence beginning "Mating systems .." is awkwardly worded and editing could improve this. Consider, is 'corresponding' is the right word in this context

>>We have replaced "corresponding to" by "controlling" L61.

Line 60 Sentence beginning "A wide diversity …" contains 'systems" three times, edit to reduce the redundancy in word usage.

>> We have changed terms L62.

Line 68 the phrase "Being then called" is awkward and could be edited to also remove te redundancy in the use of 'being'

>> This has been rephrased L68-69.

Line 71 suggests that recombination ensures proper function but I do not think the authors intend to say this since many factors are involved in the proper function of mating types (expression, protein structure, protein-protein interactions etc), do they mean ensures that expression of the loci are properly maintained or something else, edit to clarify or delete

>> This has been clarified L72-74.

Line 74 delete "still"

>> This word has been deleted L76.

Line 108 "as diverse fields as" seems unnecessary and could be deleted.

>> These words have been deleted L106.

Line 142-43 edit to clarify eg. " and a part of the ancestral HD chromosome that does not contain the HD locus."

>> We have changed the sentence as suggested L140.

Line 157-158 Consider editing the title to make a statement not an open-ended comment

>> We have rephrased the title L155-156

Lines 181-182 Delete “however”, clarify what you mean by "and conversely"

>> We have deleted “however” and replaced “conversely” by “reciprocally” L195.

Lines 200-203 In the sentence starting “Indeed, …” consider editing the sentence to remove stating ‘recombining region’ twice and to improve clarity

>> The sentence has been rephrased L215.

Lines 208-212 The structure of the sentence starting “The black stratum..” is awkward and thus not clear Edit to improve clarity

>> The sentence has been rephrased L223-226.

Line 292 what does this confidently sequenced mean? Edit to clarify, researcher 'confidence' should be based on data so state the data that provides the confidence instead of just stating confidence.

>> We have removed this analysis and therefore this sentence.

Line 317 the phrase ‘and also with’ does not make sense in this context edit for improved clarity

>> The sentence has been rephrased L308.

Line 443-444 The hypothesis indicates that heterodimers are no longer needed since a homodimer can provide the appropriate function, does the structure (sequence) of the HD proteins shed any information as to how to determine if a homodimer would have a different function than a heterodimer? Are there perhaps differences in DNA binding domain sequences or other regions?

>> As highlighted above, this would need functional characterization of the HD sequences across various strains, and will require substantial amounts of additional work, and is beyond the scope of this study. We have nevertheless added some pieces of discussion L449-458.

Line 632 the use of ‘a priori’ seems out of place in this sentence, consider editing for improved clarity.

>> The sentence has been rephrased L614.

Line 747 the phrase “with therefore likely” is awkward, edit for improved flow and clarity.

>> The sentence has been rephrased L733.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The revised manuscript by Lucotte et al. is much improved. All my concerns have been addressed. In particular, the addition of results from experimental work was essential, and the reworked Figure 7 is much clearer than the previous version. The addition of statistical testing of TE family enrichment made the conclusions by the authors more robust. Perhaps these results could have been summarized in Figure 8 or a table, since they take up a lot of space in the main text, but this solution is also acceptable.

>> We have moved all tests in the new Table S8, which indeed renders the text easier to read, thanks for the suggestion.

Reviewer #1 (Remarks on code availability):

See initial review for my comment of the code.

Reviewer #2 (Remarks to the Author):
In reviewing the revised manuscript, I find that the authors have generally addressed the reviewers' comments, resulting in a clearer and more well-structured presentation of their findings. The paper is methodologically sound, with robust genomic analyses that are well-documented and should be reproducible by others in the field. The results are generally interpreted thoughtfully, with appropriate support from the data, and are placed into the context of existing literature on fungal mating systems, highlighting the study's contributions and implications for further research. The figures and tables now better support the narrative. Overall, the manuscript makes a relevant contribution to our understanding of mating-type gene evolution in basidiomycete fungi and represents a novel observation in the field, particularly regarding the loss of function of HD genes in mating-type determination. I only have one main comment and a few minor points the authors may want to consider addressing at their discretion. The line numbers correspond to the tracked version of the manuscript.

Main Comment

The proposed hypothesis of a complete fusion between the ancestral PR and HD chromosomes, followed by rearrangements and a subsequent fission event, raises questions that, in my opinion, are difficult to reconcile. Specifically, if the rearrangements involved relocating genes from the ancestral PR chromosome into regions corresponding to the ancestral HD chromosome, this would imply significant restructuring after the proposed fission event. It remains unclear how collinearity of the HD chromosome could have been re-established to match the ancestral organization following extensive rearrangements. This aspect of the hypothesis seems counterintuitive, as a fission event after such rearrangements would likely leave evidence of disrupted synteny rather than a return to ancestral collinearity. Could the observed reshuffling of some of the genes instead be explained by their already different, possibly lineage-specific, positions on the ancestral chromosomes prior to fusion, without requiring a complete chromosome fusion?
This interpretation would also avoid the complexities of resolving a dicentric chromosome upon fusion and re-establishing a centromere after fission.

>> We agree that this scenario would seem simpler regarding chromosomal fusions and fissions. Note that the few genes reshuffled indeed are shared between *M. shykoffianum* and *M. superbum*, and not lineage-specific, as indicated in the results: "the nine genes found reshuffled between HD and PR chromosomes were located on the same chromosomes in the two species". They could have moved before the speciation between *M. shykoffianum* and *M. superbum* without HD-PR fusion, and we have added a sentence in the text to nuance our argument L471. However, this scenario would not explain the selective forces driving the chromosomal rearrangements: there is no selective advantage in terms of gamete compatibility in linking the HD chromosome fragment without the HD genes to the PR chromosome. In contrast, in our scenario, all intermediates are advantageous. It implies very few rearrangements before the fission (*i.e.* only the nine "green" genes), so there would be no reversion to collinearity, synteny would just be conserved. We have tried to better explain these points L193-202, L467-480 and on figures 3 and S6. We have nevertheless added the scenario suggested by the referee as a supplemental figure (Figure S18), we discuss this scenario a bit more L467-480, and we have deleted the chromosome fusion statement from the abstract.

Additionally, the pseudoautosomal region (PAR) at the right end of the PR chromosome aligns intriguingly with the right end of the ancestral HD chromosome. This observation raises questions about whether more localized events, such as translocations, rather than complete chromosome fusion, could better explain the observed gene distributions.

>> In all scenarios, we expect the end of the ancestral HD chromosome to remain an edge of the new mating-type chromosome. As said above, the idea is that very few rearrangements occurred before the chromosomal fission, we have tried to clarify this point.

Finally, the apparent expansion of recombination suppression within the PR-containing chromosome, as suggested by the young evolutionary strata, might indicate that recombination suppression is actively extending in this region, despite the absence of linkage to the HD locus. This possibility could provide an alternative explanation for the observed patterns without invoking complete linkage to the HD locus or a fusion-fission cycle. Could the authors briefly explore these alternative scenarios or provide evidence to clarify how the current data support or rule out these mechanisms?

>> We agree that the recombination suppression on the PR chromosome could partly be evolutionary strata expanding. However, this would not explain how a fragment of the ancestral HD chromosome would be incorporated into the ancestral PR chromosome. It is true that fusions between sex-related chromosomes and autosomes can occur by the same mechanism of deleterious mutation sheltering as evolutionary strata expansion, but it would be a striking coincidence that this would precisely occur with a fragment of the ancestral HD chromosome. We have added a sentence in the discussion on this scenario L471.

Minor Comments

Line 41: Consider changing "with a part of" to "with a segment of"

>> This has been changed as suggested.

Line 44: Remove the comma after “fusion” (should read: “whole chromosome fusion and, unexpectedly, the…”).

>> The comma has been deleted.

Line 45: Consider changing this sentence for smoother flow to: “…the HD genes lost their function in mating compatibility, with natural diploid strains often being homozygous at the HD locus”.

>> This has been changed as suggested.

The loss of function of HD genes in mating compatibility is stated three times in the abstract (lines 44-45, 47-48, 49-50). Consider streamlining and integrating these statements, perhaps by decoupling the fusion events from the loss of HD gene function in mating compatibility.

>> We have streamlined the abstract.

Line 86: I think that Ref. 34 is not appropriately cited here.

>> Indeed, thanks for checking, the reference has been deleted from there.

Line 95: Consider adding that the products of HD1 and HD2 heterodimerize exclusively with non-allelic counterparts from different strains to form an active transcription factor. This mechanism avoids self-recognition, as HD1 and HD2 proteins from the same strain cannot interact, thereby promoting non-self-interactions and ensuring compatibility between genetically distinct strains.

>> We have added this clarification.

Lines 104-118: Streamline this section to avoid repetition of “Microbotryum fungi.” Use alternatives like “these species” or “these fungi.”

>> We have reduced the number of repeats of the word *Microbotryum*.

Lines 119-120: The phrase “as neither ever recombines” could be misleading, as it might suggest that the entire chromosomes do not undergo any recombination. To avoid confusion, consider rephrasing to: “as neither of these regions ever recombines.” As indicated somewhere in the manuscript, pseudoautosomal regions of these chromosomes do undergo recombination, and this is typically required for proper chromosomal alignment and segregation during meiosis.

>> We have specified that we meant only these regions.

Lines 135-138: Suggested phrasing: “Focusing on two of these species, M. superbum and M. shykoffianum, our aim was to characterize their mating-type chromosomes, elucidating whether they displayed recombination suppression and, if so, to ascertain whether the mating-type loci were linked together or if each was linked to its respective centromere on separate chromosomes.”

>> This sentence has been changed as suggested.

Lines 144-145: Consider revising to: “The HD2 gene appeared disrupted in several strains of M. superbum, was homozygous in diploids, and showed low expression levels under mating conditions”.

>> This sentence has been changed as suggested.

Additionally, it is interesting to see some level of expression of the full transcript, suggesting, for example, incomplete degradation of transcripts by NMD.

>> We thank the referee for this insight, and we have added a sentence in the manuscript L455.

Lines 191-192: Clarify whether the “missing” region refers to a failure in assembly or a genuine difference between the two chromosomes.

>>We said in the text L190-191 “suggesting that this small region may be genuinely missing in the *M. shykoffianum* a_2 genome”.

Lines 204-212: Explicitly state that the LD analysis was only done for one species (as indicated in Figure S8) and clarify its presentation as an example. Rephrase “We used SNPs on 149 genomes” to: “We used SNPs identified across 149 genomes…”

>> We computed LD on all 149 strains, including the two closely related species studied here. Some LD may be due to population structure; however, this should not bias the inference that HD and PR are unlinked. Indeed, this should not bias the differences of LD within the PR chromosome one the one hand and either within autosomes or between the HD and PR chromosomes on the other hand. We have corrected the legend of the Figure S8 (now S9). We have also changed the text as suggested L205.

Lines 303-305: Suggested simplification: “The HD2 gene contained an additional intron compared to the functional M. lychnidis-dioicae HD2 gene (Figure 7A) and was shorter due to a deletion in the highly variable N-terminal heterodimerization domain, which is challenging to align across species”.

>> This sentence has been changed as suggested.

Line 306: Clarify why the authors suggest the M. scorzonerae genome assembly is poorer than the other species despite metrics in Table S1.

>> The *M. scorzonerae* assembly is more fragmented that the other ones. For example, the largest contig and the N50 statistic is twice as big in *M. superbum* and *M. shykoffianum* compared to *M. scorzonerae*. We have clarified this in the text L306.

Also, do you know whether the ~50% smaller genome size reflects contraction in M. scorzonerae or expansion in the other species.

>>The genome size of *M. superbum* is indeed larger than that of *M. scorzonerae*, and we suspect that it is because of an expansion in *M. superbum* and *M. shykoffianum*. Indeed, the genome

sizes of other *Microbotryum* species are closer to that of *M. scorzonerae* (around 30Mb). We have added a sentence in the manuscript L398.

Line 410: Simplify "…encompassing multiple genes linked one to each other" to "…linked together". I guess it depends on how you define the HD locus, but typically it comprises only HD1 and HD2 genes.

>> This sentence has been simplified as suggested.

Line 412-414: Rephrase to: "One mating-type locus yields 50% compatibility among gametes of a given diploid individual under random mating".

>> We have added "across tetrads" instead of "random mating" L410 as this later is used typically for outcrossing, while this sentence is about selfing.

Figure 1: Not sure if I'm interpreting this correctly, but the schematics in Figure 1 seem to suggest that only a segment of the HD chromosome (excluding the HD genes) fused with the PR chromosome. This seems to contradict Figure 3, which proposes whole-chromosome fusion followed by subsequent fission. Could the authors clarify whether these two figures represent different interpretations of the fusion and fission events, or if additional details are needed to reconcile these depictions?

>> The figure 1 only illustrates the comparison between the ancestral state and the current state, which is more related to a synteny plot than a detailed scenario. We have clarified the legend of figure 1.

Figure 7: Revise the figure legend for consistency in panel labeling (e.g., "A)" vs. "A:" vs. "B."). Correct typos (e.g., "t-wo") and missing spaces (e.g., "M. lychnidis-dioicae.The diploid").

>> This has been corrected as suggested.

Figure S2: Add a scale bar to maintain consistency with other similar figures.

>>We thank the reviewer for spotting this mistake. We have updated the figure to the version with a scale bar.

Figure S4: Clarify in the legend whether the PR gene location in the a2 M. superbum strain overlaps with the pheromone gene and is therefore not visible.

>> The genes do not overlap but are very close to each other (2831 bp exactly, see Table S3), which does not allow to distinguish them on the circos. We have added a sentence in the legend.

Data Availability, Please ensure that the BioProject and raw data associated with this study are publicly available, as I could not locate them.

>> Whole genome shotgun sequencing, RNAseq data and genome assemblies generated for this work will be available at GenBank BioProject PRJNA1169769, which is still under processing. Accession numbers are detailed in Table S9 for the RNAseq data, and in Table S11 for whole genome sequences.

Reviewer #2 (Remarks on code availability):

The specified repository provides open-access code, which appears to be well-documented, but I have not personally tested or run the code.