Subject: point-by-point letter from a previous resubmission

Dear PCI recommender, dear reviewers,

Firstly, if you are reading this message, it means that you have kindly agreed to serve as a recommender or reviewer for PCI Evolutionary Biology on our manuscript. We are extremely grateful for the time you will dedicate to reviewing our manuscript.

To be fully transparent, this manuscript was initially submitted to a society journal in plant science, and thus, we did not originally anticipate to submit it to PCI Evolutionary Biology for consideration. In September 2023, the society journal decided to rejected our ms after peer-reviewing, despite relatively positive reviews, but invited us to resubmit. After resubmission, the editor has finally decided to not send our manuscript for peer-reviewing.

From my perspective, this situation is ethically questionable and disrespectful for both the work of the authors and the reviewers. I indeed invested significant time and effort in revising the manuscript and composing the letter. Consequently, I consider that these long weeks of work should not go to waste. Therefore, I have decided to (i) move to a more ethical peer-reviewing process and (ii) provide you access to our 20-page point-by-point response letter, after removing any information that could identify the journal, editor name or other personal information.

This point-by-point letter is available at :
https://github.com/roseGWASbrowser/PopGen_GWAS_19thcentury_roses/PDF/Leroy_PCIEvolBiol_letter.pdf

Given the relevance and valuable insights provided in the previous review process, we believe that this information could offer valuable context and aid in your assessment of our manuscript. More globally, we are committed to transparency and ensuring that all relevant information is readily accessible to you.

Thank you again for your time and attention to our manuscript.

Best regards,

Thibault Leroy, on behalf of all authors

Dear editor, dear reviewers,

We would like to thank you, as well as the reviewers, for having pointed out the originality and excellence of our work and for all the suggestions. Before providing a point-by-point response, we would like to emphasize a few main points:

First, and most importantly, we sincerely apologize for the considerable delay between the initial submission and the resubmission. As the first author of the study, I experienced an overwhelming workload during the second semester of 2023, preventing me from promptly implementing the necessary changes requested by the reviewers and composing this letter, despite the relatively limited amount of work required. On behalf of all co-authors, I extend my sincere apologies for the additional effort required to revisit this manuscript after such an extended period of time.

Second, as later discussed in the letter, ploidy level variation in roses was a crucial point of our sampling design and our methodological strategy in order to be as precise as possible in our genomic investigations, *e.g.* explaining why we almost exclusively focused on diploid and tetraploid roses. We have amended the text in order to make more explicit the particularly high attention we made to perform analyses, in this specific study, by either performing analysis explicitly accounting for this bias and/or performing analysis that are known to be relatively insensitive to the ploidy level, in order to provide results as accurate as possible. That being said, ploidy variation in roses remains an inherent difficulty of the biological model and it would be presumptuous to affirm that we successfully solved/circumvented all the challenges.

Third, regarding the ratios at diagnostic markers, as pointed out by reviewers 1 and 2 in their reviews, I have performed additional analyses and added an additional supplementary note covering these reanalyses. Importantly, as a consequence of a crash of our HPC cluster in November 2023 that has induced a massive loss of information, I was only able to perform this complementary analysis on a subset of a few hundreds diagnostic SNPs spread throughout the rose genome. This analysis is however sufficient to conclude that the Early x Asian roses and Tea hybrid roses are not direct first-generation hybrids and BC1. This information is now provided in the Supplementary Note S2.

Four, a sentence regarding the cultural importance of roses appears to have been misunderstood or interpreted differently than originally intended. This sentence has been removed. It is crucial to emphasize that in the previous version, we did not intend to convey any "*socially motivated interpretations*". We sincerely apologize for the generated confusion, which is probably associated with our non-native speaker status.

Thank you very much for your understanding and best regards,

Thibault Leroy, on behalf of all authors

Specific note regarding line numbers for PCI Evol Biol reviewers and editor :
Line numbers indicated by the reviewers (in black)are those of the v1 :
https://www.biorxiv.org/content/10.1101/2023.06.22.546162v1.full.pdf
Line numbers indicated in our replies (blue) are those of the v3 :
https://www.biorxiv.org/content/10.1101/2023.06.22.546162v3.full.pdf

Decision: Reject, resubmission encouraged

Referee: 1

Comments to the Author
Dear authors,
I congratulate you on conducting this interesting and impressive piece of work on the genetic diversity in Rosa. I still have some comments and corrections that hopefully will help in improving the manuscript further.

We thank the reviewer for the supportive comment and constructive feedback below.

Main manuscript
L. 119 Why cite the Smulders paper for the array rather than Koning Boucourain?
Thank you for spotting this mistake. The correct paper is now cited (l. 142-143).

L. 131-132 could you have performed imputation from the array data to WGS data?
It is important to indicate that the two datasets have been studied independently, with very different objectives, with very little overlap in the rose varieties. Our objective was not to intersect the two datasets. However, we agree that the two datasets open new avenues and represent two interesting resources to the rose community.

L. 182 "phenotypic selection": you mean phenotype perhaps?
We agree that the wording was insufficiently clear. In this revised version, we are more explicit "*emphasizing the importance of selecting for this trait at the beginning of the 19th century*" (l. 482-483).

L.184-185 Could this later reduction also have been due to selection for another trait that is negatively correlated e.g. for increased scent production, which might be less abundant in flowers with very many petals?
Evaluating whether the reduction in the number of petals is direct or indirect is challenging, and we prefer to exercise caution in making hypotheses. Regarding the reviewer's hypothesis suggesting an overall negative correlation between the quantity of the components of rose scent and the number of petals, our dataset does not support such a correlation for either 2-PE (p-value=0.78) or geraniol (p-value=0.63). However, it should be noted that the end of the 19th century corresponds to a period of introduction of new wild species (such as *R. multiflora* or *R. rugosa*), these roses have simple flowers, in such a way that a repeated backcrossing of these new progenitors could explain the reduction in the petal number. Then, whether this reduction in petal counts was really desired, consistent with a change in aesthetic preferences, or was rather a by-product of selecting for some other traits remains difficult to evaluate. Historical sources attest that from 1880 onwards the selection of garden roses included single-flowered roses (5 petals) and the selection for cut flowers concentrated on semi-

double flowers (an undetermined number of petals, but a bud shape that is not compatible with full flowers).

L.213 "Fig S1": What I miss in Figure S1 is some connection between the points on the PCA plots and the classification of the samples (as given in legend of Fig S2 for example)

We are not totally sure that we have fully understood the comment from the Reviewer 1. We apologize if the corrected changes are not satisfactory. As far as we have understood, the reviewer 1 asked us to add the different periods of rose breeding (Fig. 1 below).

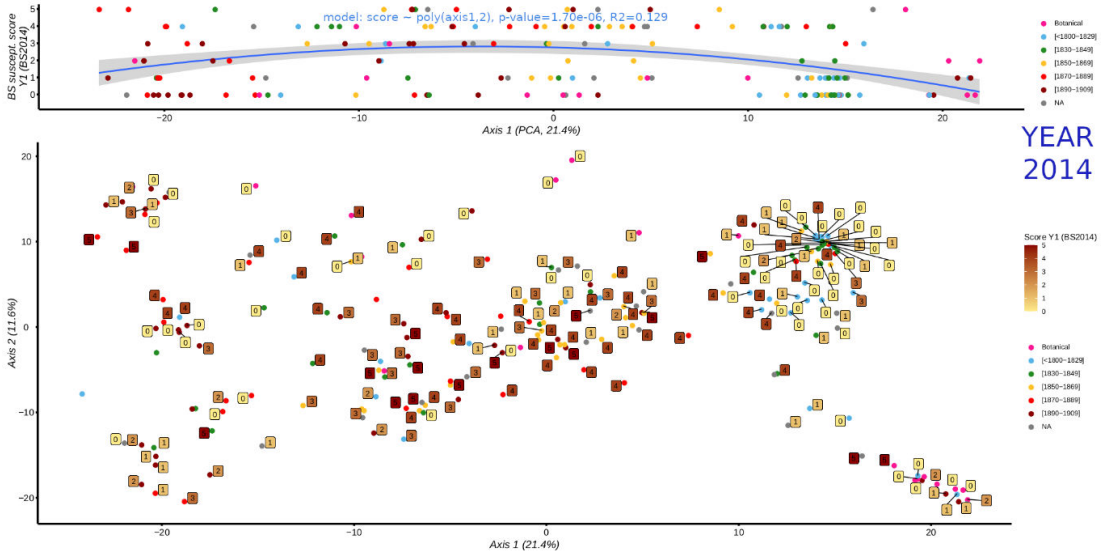

Figure 1: Edited version of the previous Fig. S1 that now includes the different periods of selection as different colors (see Fig. S4 for the complete figure)

L.228 (figure 2A caption) "PCA": how did you handle variable ploidy in the PCA?

As introduced in the Materials and Methods (previously l. 684-688, now l. 249-253), all the genotyping was performed assuming a ploidy of 4. " *To take into account the variable ploidy level between individuals, with the vast majority of the accessions used expected to be either diploid or tetraploid, we generated a final genotyping matrix using FitPoly assuming ploidy=4 for all individuals, which means that diploid individuals with heterozygous alleles are generally expected to be called AABB."* An important point to highlight again here is that we made a particular effort to genotype roses for which we have previous in-lab information about their potential ploidy level. Even if there is no variable ploidy in our PCA, it is also important to note that it would not have been a problem since PCA is independent from a specific population genetics model, making it particularly relevant to a variable ploidy level (see also below).

L.238 (figure 2C) Not obvious what the numbers represent in the plot. I assume these are the 32 sequenced accessions but needs to be stated.

We apologize for this lack of clarity. We now clearly state that the numbers correspond to the accessions and that the information can be found on Table S2: "*See Table S2 for a correspondence table between accession numbers and cultivar IDs*"

L.244 (figure D) I am wondering how useful is it to colour the dots by chromosome in fig 2D, as it is very unclear whether this is to do with the colour of the dots, or perhaps the means? If the dots as I guess, they are far too small to be distinguished by colour from this figure. I would say if you think it is important to provide a per-chromosome picture then do that in the Supp Figs, and drop the chr legend here as it only confuses. The mean lines are coloured by the groupings from B I think, but now the Botanical group has been replaced with all accessions. All accessions is fine to include, but where are Botanicals? It is quite confusing.
Thank you very much for this comment and suggestion. First, we followed your suggestion and decided to drop the information of the chromosomes, see the new Fig. 2D. Second, the botanical group is not shown in the analysis, since this group is composed of different species and can be considered as too extremely diverse to be considered as a single group (Fig. 2A & B). We have now made this information explicit (l. 392-393): "*Botanical roses were not included in this part of the study because their population structure is not consistent with a single group.*"

L. 248 here the correct reference to the rose WagRhSNP array is given
Thank you for your understanding.

L. 260 Is the reseq data (and the SNP array data) publicly available? Should mention in a data availability statement somewhere. Also, what were the other samples used in this part of the study? Give reference to supplementary table somewhere in main text here I think.
The newly sequenced data are publicly available on SRA (PRJNA997103). At the time of the first submission, all the data were already submitted on the NCBI clusters, but not yet available publicly. We apologize for this inconvenience. This issue is now fixed and this information is explicitly indicated in Table S2, as well as in the dedicated data availability section.

L.267 "ascertainment bias": could also be because of the extra noise in the WGS data rather an array data issue?
Unfortunately, we kindly request clarification on the specific additional noise the reviewer is referring to. One important aspect here is associated with the number of PCs that isolate a single botanical group from the rest of the samples. For instance, three of the five first PCs (PC2, PC4 and PC5) exhibit this Botanical-specific pattern, plus PC3 to a lesser extent. Even if the results are not totally comparable and that we have not performed detailed comparisons, we have (i) observed a huge difference in the number of SNPs between the WGS dataset with or without the botanical roses (~78 vs. 54 million SNPs) and ii) not observed such a clear botanical pattern on the first axes of the PCA based on the SNP array, which would be consistent with the reference panel used for the design of the array. The choices made at the time of the SNP array are however fully justified. It just means that the SNP array and the WGS data are two complementary approaches, each with their own advantages.

L.275 "calculated on the seq data": calculated on all 77M snps, or the 50K selection?
KING was used on the whole SNP data. We now made this information explicit.

L.294 "3:1 ratio": Do you discuss the significance of this ratio more in the paper? Looks like the result of a single backcross to Asian, possibly then all subsequent crosses were tea x tea that (without selection pressure favouring one founder background) would conserve a 3:1 ratio. Interesting to speculate about this, could you test whether they are all BC1 accessions in your study?
We fully agree with the reviewer 1 that this ratio is astonishing and we would have been definitively very happy to discover that all Hybrid Tea are BC1, but this was unlikely to be the case according to us. In this revised manuscript version, we have provided a dedicated supplementary note (S2), for

which we investigated this hypothesis based on the observed genotypes at diagnostic markers. Our results are inconsistent with a simple F1, followed by a BC1, and is more consistent with the pattern observed on Fig. 2C, which suggests that selection was based on a very limited number of generations, but more than one per group. Future large genotyping or sequencing projects will probably contribute to recovering the pedigrees, at least partly, and contribute to a clearer picture.

L.302 "16 chromosome": I don't understand where the nr.16 comes from if the base chromosome number of rose is 7.
We are referring here to the number of chromosome sets among the individuals. For instance, in the Ancient Asian group, we included 7 individuals for the nucleotide diversity (see Table S2), six inferred as diploids and one tetraploid. Thus, the cumulative count of complete chromosome sets within this group amounts to sixteen. All groups have exactly 16 complete chromosome sets. This information is now provided more explicitly (l. 615-618).

L.420 portail -> portal
This second typo has been removed.

L.462 "Figure 4": Although I understand that circular plots are very visually attractive, for Manhatten plots they do not work as well as ordinary flat plots, it becomes a bit disorienting. Could I suggest that the developers also provide a linear version of these plots on the website. This could be laid out in a 2 x 4 grid for example, including chr00. The key could be a pop-out on hovering over some element.
Unfortunately, here it is not possible. First, because it would require considerable work to redo all the work, reshape the website etc and that there is no staff available to work on that (the contract of the first author finished in July 2022). Second, because our objective was to only provide access to all the results, not to build an advanced website. It does not mean that we do not take the suggestion of the reviewer into account, but it would require another dedicated project. In this new version, we have however agreed to provide a linear Manhattan plot corresponding to figure 4 (See Fig. 2 below, a figure which is now part of the SI, see Figure S14).

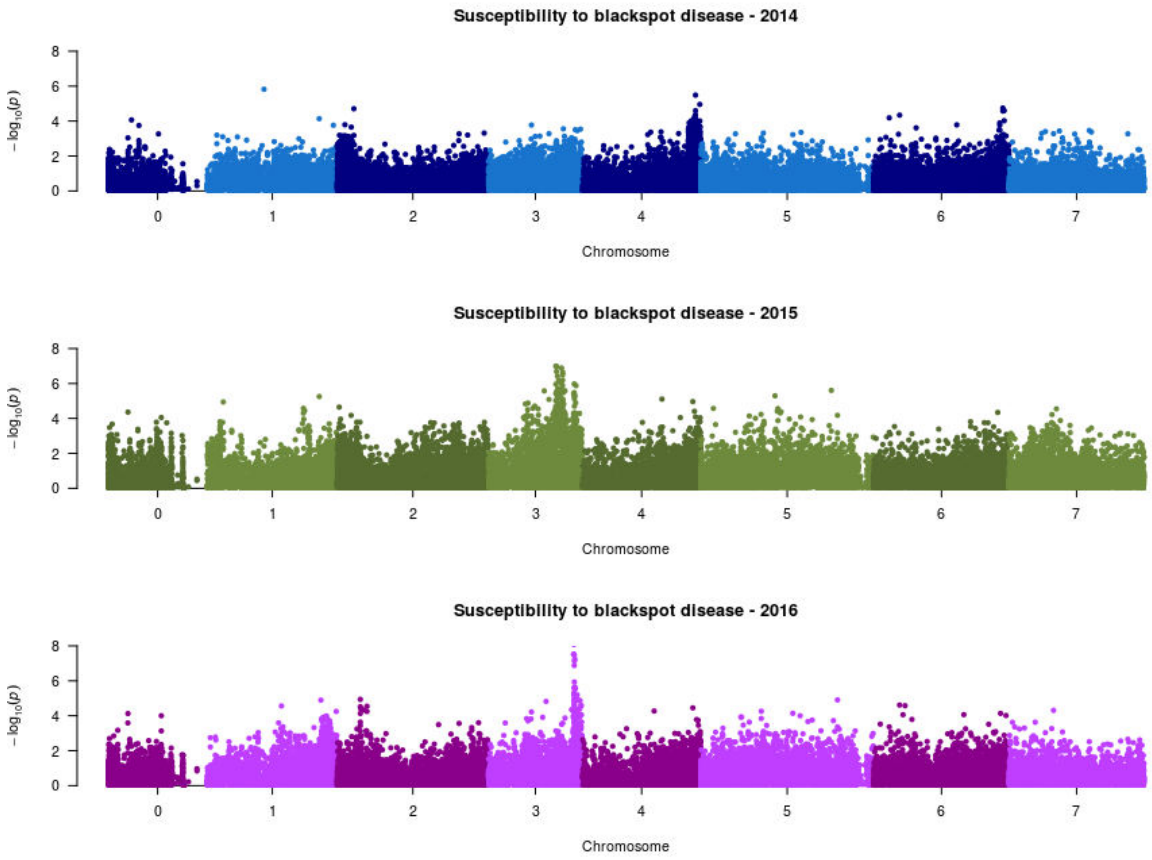

Figure 2: Linear Manhattan plots corresponding to the results of FIgure 4 of the main text. Manhattan plot generated with the qqman R package.

L.481 "food security: I agree with these sentiments, but rose is of far less vital important to global food security than wheat.

This sentence does not refer to roses at all, since this sentence is associated with the loss of diversity in bread wheat. If we, of course, agree that the observed loss of diversity in roses is more minor, since it cannot generate catastrophic consequences for humankind as compared to a cereal that is central in the human diet, the reduction of diversity in roses remains a worrying situation as rose can provide numerous services for mankind (food, medicine, hedonic value…).. Our objective here was to make this point explicit.

L.492 "anecdotal": Anecdotal is probably not what you mean here (= a story / single example).

Ok,we have now preferred the use of “minor”.

L.494 "monotheistic religions": Not sure why this is relevant

L.495 "political parties": You mean the labour movement??

This whole section / line of argumentation is becoming increasingly obscure. I don't think you need to try to justify the value of genetic diversity by drawing on the symbolic importance of a species etc. Unraveling the breeding history of the world's favourite flower is fascinating enough as a subject.

First of all, we apologize for the misunderstanding, which had a cascading effect on the editor, who misunderstood here and considered that we made “*socially motivated interpretations*”! We fully agree with the point of the reviewer. We have toned down this part and now only refers to “*the cultural and economical significance of rose cultivation*”, without providing any information regarding the importance of the rose symbolism.

L.535 "allele is not present": Did you also investigate using the Raymond et al assembly which might have had the copia allele? Although both assemblies were of the same genotype (OB) so I guess they are the same allele, unless there were assembly issues on that region in one or the other assembly. Please check.

As mentioned in our Supplementary Note S1, two lost-of-function alleles exist at the *RoKSN* locus: in the reference genome (Hibrand et al., 2018) used in this study, the null allele is associated with a large rearrangement. In the other reference genome (Raymond et al., 2018), the *copia* allele is present. As proposed by the reviewer, it could be interesting to test the signal on the genome with the *copia* allele (from Raymond et al, 1998). However, this analysis would represent several complete months of work since it would require redoing all the steps on this new reference genome, including the SNP calling, the fasta sequence reconstruction and the genome scanning for nucleotide diversity. Given our objective was not to perform a detailed analysis of the *RoKSN* locus here, such an analysis is not mandatory to conclude here. We are however confident that the strong signals detected in the breakpoint regions of the rearrangement are consistent with the selection of the genomic region of *RoKSN*. Given that the raw data are now publicly available through SRA, future investigations on this specific question will be possible.

L.574 "precise": narrow down

We substituted the word “precise” by “describe”.

L.577 "the chromosome 3": remove 'the'

Done

L.577 "associated to": associated with (French à = to usually, but in French your construction would sound something like "associé avec" instead of "associé à" I am guessing. In any case, 'with' is better than 'to' here)

Done

L.597 "RoKSN -, the": delete ', the'

Done

L.611 "presumably": presumed

Done

L.618 "bais": bias

Done

L.648 "higher is value, higher is the reblooming": this should read "the higher the value, the higher the reblooming capacity". This mistake was also made somewhere else, I think in the supplementary file, please check.

Done

L. 676 "fitPoly": cite fitPoly: Voorrips, R.E., Gort, G. & Vosman, B. Genotype calling in tetraploid species from bi-allelic marker data using mixture models. BMC Bioinformatics 12, 172 (2011). https://doi.org/10.1186/1471-2105-12-172

Done

L. 684 "unpublished R package": true, but you can download it from polyploids.org. Perhaps mention?

We would be very happy to cite the R package fitPolyTools, could the reviewer provide more information regarding the precise location of this package on the polyploids website? We have not identified this package among the different softwares: https://polyploids.r-universe.dev/builds

L.688 "AABB": did you look at the called dosages of the diploid samples and check for presence of 1 or 3 scores? In my opinion these should not be present at any great frequency, and if present should be made NA

Here, we would like to clarify our views regarding the allowed quality of the calls and the expected power in the underlying analysis. In our case, we have decided to not go in this direction, even if this strategy was evaluated at the beginning of the postdoc of the first author. The rationale is the following: if a genotype is called ABBB, it could be either AB or BB, but the probability to be AA is ~0. So, when the number of individuals can be considered as relatively limiting for a GWAS analysis, as typically expected in our study (~200 individuals only), removing all the information associated with that call is expected to more strongly penalize the GWAS than just considering an imprecision on the call.

To be more explicit here, if we consider Old Blush as a typical example of a diploid rose, we observed a clear excess of AAAA (26,9%), AABB (26,1%) and BBBB (25,7%), with regards to the AAAB (10,4%) and ABBB (10,9%) genotypes. It should be noted that this result is very good given the continuous nature of the signal of fluorescence. One remarkable result we obtained based on the analysis of replicates is that the genotyping error is highly repeatable (97,3% of the calls are similar between two replicates, Fig. S1). It suggests that the error is not at all stochastic and that two diploid genotypes tend to be similarly called AAAB or ABBB for the two replicates, which is probably linked to the SNP array and/or complexity of the rose genome.

Finally, it is also important to indicate that considering all AAAB and ABBB as heterozygotes would be unsatisfactory as well, since it would generate a large excess of heterozygotes. According to us, the strategy used is expected to offer the best balance between accuracy and power.

L.694-695 this description is not clear, especially the number 35540 in brackets, suggest to remove this caveat and the next sentence ("this choice"). I would say, state what you did. Most readers will not even be aware that you could / should only include SNPs where both probes were successful, high quality and matching. It is debatable whether this double probe approach was a good strategy or just a waste of space on the snp chip! Your other QC steps should deal with any quality issues so I see no need to "defend" your approach here in the M+M section.

We agree with the reviewer, this section was removed accordingly.

L.706 "this metric": it is called 'simple matching'

We have preferred to be even more explicit here by saying "*the observed distribution of the proportion of similar calls based on all the pairwise comparisons*".

L.713 "92007 markers": You could also have done an LD analysis first to remove markers that were in LD. This can be done e.g. using the SNPRelate::snpgdsLDpruning function, from the SNPRelate package you already use. Using collinear SNPs in PCA could cause unwanted artefacts.

Thank you for this suggestion. The PCA in Figure 2A is now based on a pruned SNP set of 11,848 SNPs, that we generated with plink (--index-pairwise 20 5 0.2), in order to be consistent with the WGS data. On this subset of markers, PCA was again performed. As expected, we observed very little deviation of the coordinates of the samples between the two analyses (PC1: R^2=0.98, p-value < 2e-16; PC2, R^2=0.95, p-value<2e.16, see Fig. 3 below). The main difference is associated with the proportion of explained variance, especially on PC1 (21.4% rather than 29.3%). Even if our analysis was indeed robust to this bias, we have decided to update all the figures as suggested.

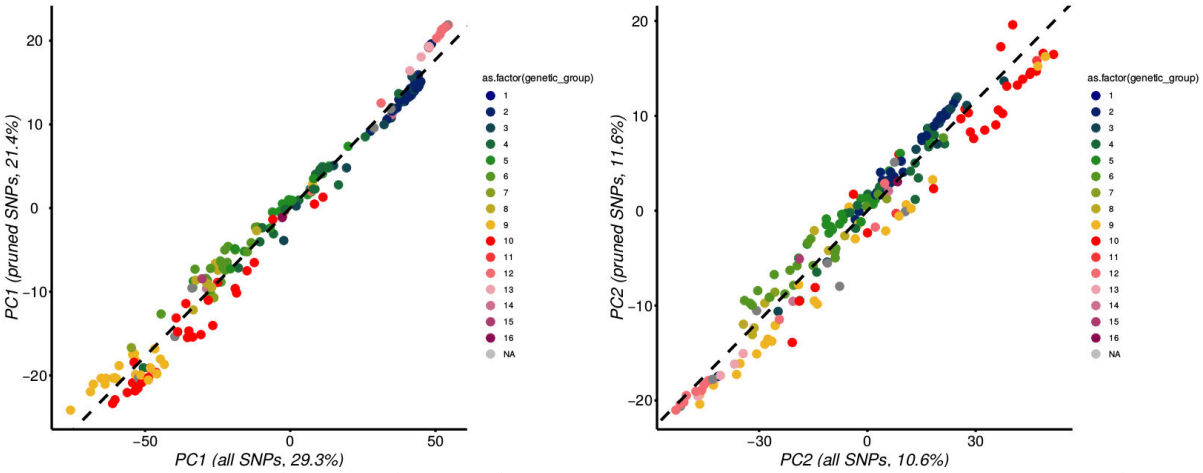

Figure 3: comparison between the PCA before (97,007) or after SNP pruning (11,848) on the SNP array data. The dotted lines correspond to the best linear models.

L.744 "SRA": explain acronym
Done

L.775 "randomly selecting": you randomly selected 50K snps from 78M. I would be concerned that the quality of these 50K snps would not be high.
Also: how does this package deal with correlations between SNPs? see my previous comment about LD analysis to remove possible collinear SNPs (in LD) before performing the PCA.
And does it standardise the data before performing the PCA?
Also interested in your thoughts regarding PCA analysis across multiple ploidy levels.
One main advantage of PCA, and its derived approaches, *e.g.* DAPC, is associated with its independence on a specific population genetics model, making it free from assumptions regarding Hardy-Weinberg equilibrium or linkage disequilibrium. Consequently, PCA is known to be a valuable tool, regardless of their ploidy and rate of genetic recombination (e.g. see Jombart et al. BMC Genomics). This result explains why we have decided to highlight the results of PCA in the ms (Figs. 2A and 2B) and shows the results of the model-based methods as supplementary figures (Figs. S7 and S9). That being said, it is even more important to understand the rationale behind our first round of WGS analysis here. We performed this analysis to (i) verify the consistency of the population structure observed with the NGS data with the previously detected using the SNP array and (ii) exclude individuals that appear particularly closely-related and (iii), even more importantly, empirically estimate the ploidy level of each individual. So, by definition, this approach was performed to evaluate the ploidy level. This is why, we then performed a second round of SNP calling,assuming the exact ploidy level of each individual (except those previously excluded) using the "--ploidy" option HaplotypeCaller, which was then used for the rest of the analyses, including to ensure unbiased estimates of nucleotide diversity. For this latter analysis, we reconstructed fasta sequences with positions explicitly considering the level of ploidy.

Regarding the SNP pruning, we agree that LD remains substantial in our dataset and that pruning SNPs for LD could induce an artifact, albeit small, especially with regards to the proportion of explained variance. To take this into account, we used PLINK to remove SNP exhibiting too high LD (--index-pairwise 20 5 0.2) from the previous SNP set used for population structure, leading to a new pruned SNP set of 17,669 SNPs. As expected, the PCA after SNP pruning (bottom, Fig. 4 below) is very similar to the previous one (top, Fig. 4 below), and the most noticeable difference is proportion

of variance explained by the first axes as compared to the total inertia, a proportion which is now smaller on all the PCs shown.

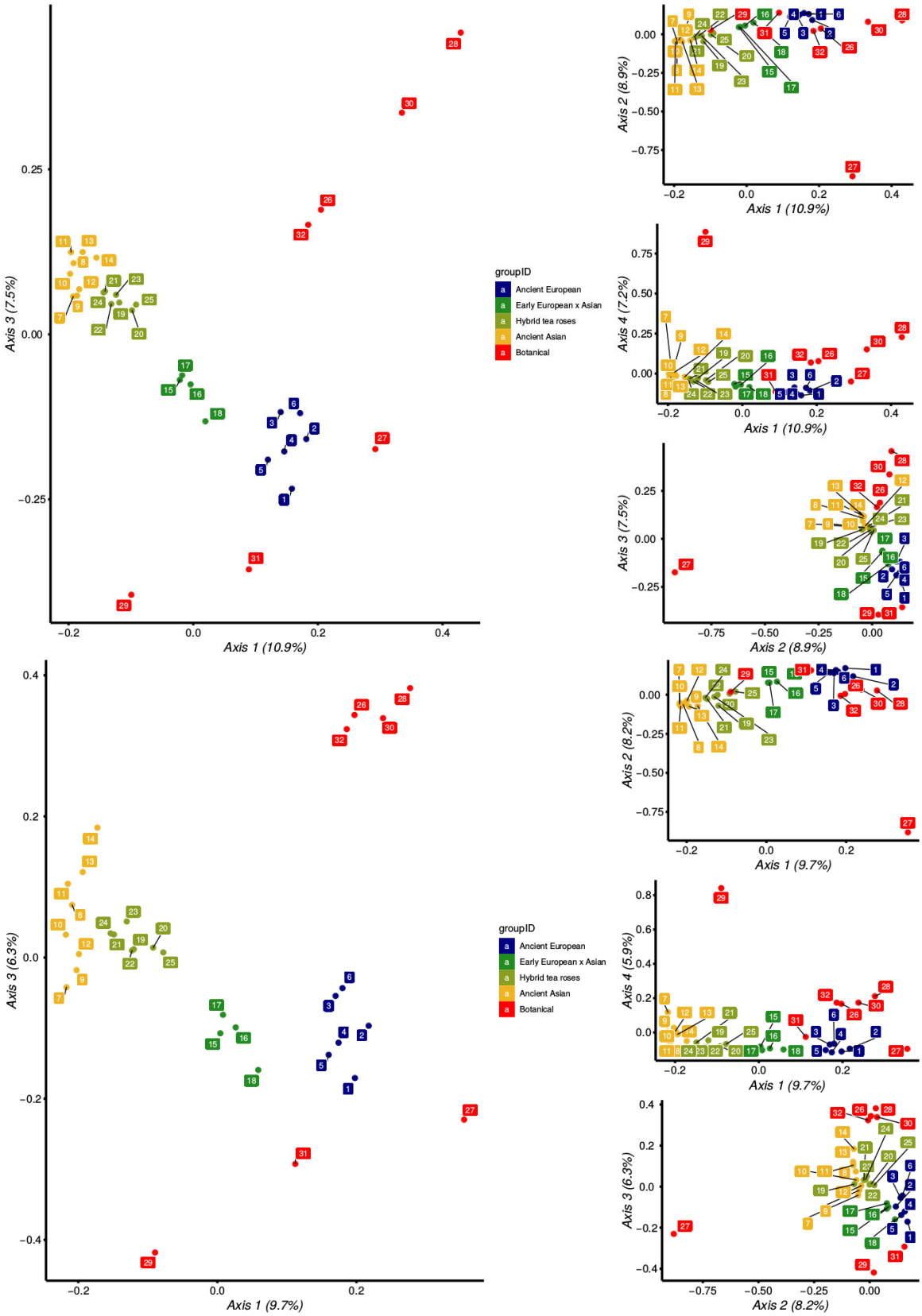

Fig. 4: PCA before and after SNP pruning based on linkage disequilibrium on the WGS data.

Supplementary File:
L.71 looking at figure S2, it appears that the gypsy allele of AP2 actually results in fewer petals, which is not what I would expect with a double flower.
We agree with the reviewer that this result is not in full agreement with what was previously published. We mentioned this discrepancy in the Supplementary Note S1 (l. 83-96). This suggests that the genetic basis of double-flowers might be different between Chinese and European roses. The previous studies were mainly made on Chinese-related roses (old Chinese and modern roses), whereas in our study, we have a lot of European-related roses. We think that this work opens a new avenue to investigate this in detail in the future.

L.179 "K = 6": K=2 to k=5 shown only, not k=6
Thank you for pointing out this typo.

L.265 "shown in yellow": anna_maria duplicate is not coloured in yellow. comtesse de murinais is also a duplicate, but not shown? Why grouped with Zoe as a triplicate in yellow? I would recommend to tidy up this table.
We have reorganized the Supplementary Table 1.

Also, 4 seasons variants are probably mutants or sports so would be likely to show up as duplicates without necessarily being identical clones I would imagine. You do mention this as a possibility in the main text but here is a clear example of it I think that you do not even refer to!
Our analysis is indeed consistent with the fact that the cultivars referred to as “4 saisons blanc mousseux” and “4 saisons continues” in the Loubert collection are clones (mutants or sports). But we do not want to highlight this result too much, since we also observed that the names of the genotypes can often be misleading, that mislabeling is frequent in the collections etc, in such a way that we encourage rose geneticists to always check that the expected clones are indeed true ones based on their genotypes.

Referee: 2

Comments to the Author
In this manuscript Leroy et al. collect phenotypic and genetic data from 204 roses on 69K SNP array plus 32 whole genomes. Leroy et al. use these data to conclude three main things. 1) Like in other crops, Roses have suffered a monumental reduction in genome wide diversity when compared to their wild ancestors. 2) Genomic footprints of artificial selection can be found on four of the seven rose chromosomes. 3) phenotypes important to rose breeders will map to the rose genome with GWAS approaches

The paper is well written and the science is indeed sound. I do have three points of criticism.
We thank the reviewer for the positive assignment and constructive feedback below.

First, If I am understanding the manuscript correctly, hybrid tea roses in Europe are the result of a Europe x Asia cross followed by one backcross to Asian roses. Essentially, AA x BB = AB; AB x BB = 50% AB and 50% BB leading to the 3:1 allele frequencies reported on line 294. The reduction in diversity therefor is not the classic domestication syndrome, but instead just the result of an F1 backcross. You're free to disagree with me, but I don't think this result is particularly interesting. I think the paper would be better served focusing on the genomic or GWAS portions.

Thank you for this suggestion. As previously discussed with reviewer #1 (see above), we agree that the values are intriguing and that the F1/BC1 hypothesis is particularly tempting. Our manuscript now includes a new Supplementary Note that investigates this hypothesis in detail based on the observed genotypes at diagnostic markers (Supplementary Note S2). Our results are inconsistent with the simple hypothesis. That being said, our results are consistent with an extremely limited number of generations of breeding. It seems also important to mention that the observed reduction of diversity is more pronounced than expected based on the 3:1 ratios (expected: $1.32x10^{-2}$ observed: $1.21x10^{-2}$, see lines 628-631) suggesting that selection for a few key traits likely contributed to this reduction of diversity levels, which is an important information to mention according to us.

Second, and much, much, less important is that, the way I understand it, while asexual propagation does preserve ancient genotypes much better than sexual reproduction ever could, it that does not mean they're perfect clones. The accumulation of somatic mutations can alter phenotypic expression - as without recombination selection has no good way of purging deleterious mutations. I don't see this mentioned at all. Is it something you were concerned with? Tested for? Could this elevate the diversity measures of the ancient samples?

Thank you very much for this interesting comment. In our study, none of the sequenced genotypes exhibited consistency with clones, as determined by KING analysis, rendering it impossible to conduct the suggested investigation. If we were to consider the possibility that clones were identified in the analysis, we might have examined the proportion of mismatches, anticipating that some of them would be indeed linked to true somatic mutations. However, it is important to note that this number of mutations is anticipated to explain only a fraction of the overall mismatches, given that the number of somatic mutations is expected to be limited (not at all at the scale of the levels of the observed diversity) and because there are many biological and methodological limitations to consider (see several papers on this topic that involved the first author, including Plomion et al. 2018 Nature Plants, Schmitt et al. 2023 Peer Community Journal, and Schmitt et al. 2024 PNAS). Having said that, we agree that the exploration of this aspect could be a promising avenue for future research, particularly in the context of roses where naturally occurring mutants with different phenotypes ("sports") are subsequently vegetatively propagated. This suggests that analyzing clones with distinct phenotypes in the future could potentially unveil variations in crucial genes related to aesthetic traits.

Lastly, the exclusion of any of the GWAS results is strange. I know you probably don't want to discuss all of them for length reasons but was there a particularly strong one? Or one on a charismatic phenotype? If the editor is okay with just pointing to the online GWAS explorer, than you can ignore me.

The reviewer is definitively right. It was a difficult choice to make. We performed hundreds of genome-wide associations for tens of traits and unfortunately, due to space constraints, we found no better way of delivering the result than (i) focusing in the main text on the results of one single trait and (ii) building a dedicated website that could host all the results. We found that the susceptibility to blackspot disease was a particularly interesting trait since (1) we observe a clear increase of the susceptibility throughout the breeding period, (2) this trait represents a typical new target for breeding programs in the context of the reduction of pesticide sprays and (3) we found particularly clear signal for this trait.

I enjoyed reading this paper. Great work! Please find below a list of minor comments

line 80 - confusing sentence. reword.

We agree. We have revised this sentence to clarify our views. "*Despite roses being cultivated since antiquity, both independently in China and the Mediterranean region, the number of varieties has long remained particularly limited.*"

line 84 - add comma to 6000
Done
line 139 - is "largest collection of GWAS analyses" the right term? Maybe "the largest GWAS" or "the GWAS with the most number of phenotypes"
We agree that this work represents a single GWA study, even if many genome-wide associations have been performed in this study, so the former wording was awkward. We changed the text accordingly.
line 176 - rouge "the" please remove
Done
line 234 - generally axes 1&2 are shown. Why here is it axes 1&3
Among the five first axes of the PCA, three (2, 4 and 5) only isolate a single botanical individual and are therefore not meaningful with regards to the population structure among non-botanical samples. To be explicit, the previous version of the ms included a supplementary figure showing the results for the first four PCs. This remains true on this new version (Fig. S8).
figure 2D - confused why this is included in a main text figure
We assume that the reviewer is referring here to the different colors for the different chromosomes, an issue that was similarly reported by reviewer 1, we followed the suggestion and removed this information from the figure.

line 249 - what is clone correction? Can you just say removal of clonal genotypes?
Done
line 257 - I think "consistent" is to strong for this. Some Ancient Asian samples (group 9) are placed well within the PCA space you're claiming are hybrids
Thank you very much for pointing it out. Accordingly, we changed to "mostly consistent with". Due to the limiting number of SSRs used by Liorzou et al. (32), some groups were less resolved than in our study. For instance the cultivars "la reine" or "enfant de France", which are known hybrids, clustered within the group 9 in Liorzou et al., while our analysis based on tens of thousands SNPs allows us to place them as hybrids between the ancient European and the ancient Asian rose clusters (Fig. S6).
line 273 - replace "thanks" with "due"
Done
line 273 - If the botanical samples are skewing the PCA that much you might redo the PCA without them for a clearer picture.
Here, we have decided to not follow this suggestion. According to us, it remains better to provide all the information to the reader. This analysis highlights the high private diversity of the botanical roses.

line 283 - Is it possible to get the Liorzou et al. 2016 groupings into this paper as I'd like to see if these results are consistent with your SNP array PCA but can't since they're just named 1-16.
We are unsure to have correctly understood the comment of the reviewer here. As far as we understood, the reviewer suggested adding the labeling of Mathilde Liorzou et al. on the PCA based on the WGS data. This is unfortunately impossible since a substantial number of individuals we sequenced were not all part of Mathilde Liorzou's work. The WGS work was indeed based on the historical writings ("star varieties") and also on the rose data already available through SRA (most of them collected in different rose collections). As a consequence, our list of genotypes poorly intersects

with the one from Liorzou. We however consider that the two PCA are sufficient to highlight that the two axes 1 capture the same component of population structure and are consistent with each other.

line 290-295 - Remarkable! this suggests just a single generation post hybridization for the early AxE and two generations post hybridization for the hybrid tea roses.
Thank you for this suggestion. We have indeed worked on this hypothesis based on the calls at diagnostic SNPs, but our results appear inconsistent with such a simple hypothesis (see also above).

line 372 - The few generations since hybridization would only show the traits with the highest strength of selection. You would miss many of the smaller effect loci.
We are fully in line with the reviewer #2 regarding this point. Breeders, during the period covered by the study (19th century), but very likely later too, probably targeted a relatively limited number of high-effect mutations because of the limited number of generations. This opens new avenues for future breeding.
line 314 - I don't think you meant "population structure" here. I think "genomic composition", "genomic make-up", or "genomic architecture" work better.
Agreed, “genomic composition” was preferred.

line 327 - I think this is just consistent with a reduction in diversity due to inbreeding
We only partly agree with the reviewer here. As stated in this section, inbreeding alone cannot explain all the patterns, since we observed a lower diversity in the Hybrid tea group as compared to the diversity that could be expected based on the genomic composition.

line 331 - how many would you expect by chance?
Without knowing the number of generations of breeding and the recombination landscape, it seems impossible to derive an expectation here. Rather than inferring a number of “significant” outlying windows, we rather consider these windows as candidate windows. Assuming that the detection was random, we could have expected to see these windows spread across the whole genome, while most of the signal is observed on some specific regions of only two chromosomes (chrs 3 and 5).

line 431 - most "impactful"
Done

Referee: 3

Comments to the Author
In this work, Leroy et al. investigated the population structure of many rose varieties with mixed ploidy. SNPs of more than 200 accessions were genotyped in a SNP array, and with ~30 accessions with whole-genome re-sequencing data. The phenotypes, population structure, and GWAS results were described.

While being a study with large amounts of data, one major issue in this study and the set of samples is the uncertainty of accession ploidy and whether correct methods have been used to analyze them. The previous study (Liorzou et al. 2016, Fig 4D) showed ploidy ranging from 2x, 3x, 4x, 5x, to 6x. Yet in the current study, for whole-genome sequenced data, the authors assumed the samples are either 2x, 3x, or 4x without other possibilities (for example, in lines 760-770, 838-839). For the SNP array data, it appears all individuals were being treated as 4x (lines 710-740) in the key analyses. Likely due to

this limitation, population structure was only described in terms of the similarity among individuals, rather than detailed dissection of the history of ploidy change and intra-/inter-ploidy hybridization. From Liorzou et al. 2016, the history of hybridization could be very complex, with different levels of ploidy within the same "genetic group", yet the current study did not investigate beyond what was already known from Liorzou et al. 2016.

We sincerely appreciate the reviewer's valuable feedback. Firstly, it is crucial to acknowledge that many of the authors (including the senior ones) were actively involved in the study conducted by Mathilde Liorzou and collaborators (2016). We are therefore especially pleased that the reviewer #3 recognizes the high quality of our previous work. However, our current objective was not to replicate the study by Liorzou and collaborators but rather to expand upon it. We aimed to explore the genetic foundations of significant traits through Genome-Wide Association Studies (GWAS) and investigate potential footprints of selection (selective sweeps) that breeders have induced in the rose genome during that period—two aspects that are challenging to investigate thoroughly with SSRs.

While we are delighted that the reviewer acknowledges the excellence of the work carried out in our lab a decade ago, we believe our current study has significantly advanced beyond the scope of Liorzou et al. 2016.

It should be noted that given that most authors were also involved in the work of Mathilde, we explicitly excluded some cultivars that could have been particularly difficult to identify clusters on fluorescence data and generate high quality genotyping on polyploid species (here cultivars with ≥5x ploidy levels in Liorzou were excluded). We mostly focused on cultivars with 2x and 4x ploidy levels, only allowing a few individuals with 3x, because they were known to be quite pivotal in rose breeding during the 19th century (e.g. cultivar ‘La France’). It was also discussed in Liorzou et al. (2016) that 5x and 6x clustered mostly to the groups 14-16 that were very isolated from most of the other groups and that it can be interpreted by a limited role in rose breeding during the nineteenth century. Therefore, due to the technical difficulties for their analysis and their limited interest in understanding the process of genetic improvement in the 19th century, it was preferable not to include them in our analysis.

For the genomic part, one major finding of this study is that the population structure is consistent with the previous SSR-based study and previous pedigree information of Asian-European rose hybridization. While being consistent is good, one would wonder what is the novel finding and insight from this study. With the genotyping efforts in this work, except the same general description as previously known, what more biology can one infer? While the whole-genome scan for traces of artificial selection might not have been done before, this was based on the hybrid tea rose population with 7 accessions (Fig 2). The accuracy of such scan based on 7 individuals remains to be discussed. In addition, instead of identifying novel target genes or inferring hidden traits under selection from gene annotation, the authors focused on describing the co-localization of selection windows with a previously identified chromosomal inversion containing the RoKSN gene. Therefore, the value of the selection scan part seems to be in the findings of Kawamura et al. 2022 rather the the current analyses. Similarly, for the GWAS part, this manuscript sends the message of "GWAS were done and results are in this website" without further investigating the GWAS results.

We concur with the observations made by reviewer #3 regarding the ways of improvements with more data in the future. Our primary aim was to contribute to the field by transparently presenting the available information on the data at hand. Even if pi is independent of the sampling sizes, local pi can

vary and therefore its use to identify candidate selective sweeps could be limited by the limited number of sequences (n=16 sets of chromosomes per group in our study), as a consequence, our results should be interpreted as a first evidence and we indeed encourage forthcoming studies in roses to pursue a similar approach with more individuals.
Concerning *RoKSN*, a well-recognized target in breeding, our goal was to showcase that our methodology is proficient in identifying the genomic region associated with this gene. However, it's essential to emphasize that our approach is comprehensive, extending beyond *RoKSN* to identify potential candidate targets. Further details can be found in the report accessible here: https://github.com/roseGWASbrowser/PopGen_GWAS_19thcentury_roses/blob/main/popgenomics/NucleotideDiversity_DTaj_ROD/diversity_ROD_KperK_050422_withcandidategenes.pdf
Importantly, all the data are publicly available to the reader in order to allow them to browse the results:
https://github.com/roseGWASbrowser/PopGen_GWAS_19thcentury_roses/tree/main/popgenomics/NucleotideDiversity_DTaj_ROD

It remains unclear what the purpose is to include 32 whole-genome-sequenced accessions. For these accessions, the analyses still used sub-sampled 50K SNPs, even less than the 92K SNPs from the SNP array dataset. How do these samples help us to get more insight on the history of rose breeding?
Sorry for the lack of clarity. All the analyses based on the WGS were performed on the whole set of SNPs (round 1 with botanicals: 77,862,879; round 2: 54,481,222). However, given that population structure inferences do not require tons of SNPs, we initially subsampled our dataset to 50,369 randomly chosen SNPs (now reduced to 17,669 SNPs after LD-pruning, see reply to reviewer's 1 comment). The consistency of our population structure with the one observed by Mathilde Liorzou et al. at only 32 SSRs is an empirical proof of how accurate population structure can be inferred with limited data. The rest of the analyses, including the unbiased estimates of diversity, requires whole-genome data.

It is difficult to see whether the genetic grouping or PCA is consistent between Fig 1A vs. Fig 1B, because the readers have no background information of what the 16 SSR groups are. A related issue is that it is unclear why the authors seem to stick with the SSR genetic grouping from Liorzou et al. 2016 despite the current dataset should be more accurate. Despite the authors have done fastStructure for this dataset (Fig S5), accessions in all PCA figures of this manuscript was separated based on the previous SSR study. The authors should provide some means to help readers judge what is the reasonable K values from fastSTRUCTURE, check whether it is consistent with PCA, and discuss the meaning of these groups (and whether it is consistent with the SSR grouping or not).
Our initial attempt was to provide the following information in the figure caption: *"Roses with ancient European and ancient Asian backgrounds are shown in blue and yellow, respectively (typically corresponding to groups 1-3 and 9, respectively; see also Fig. S6 and S7)."* So to indicate that the color in Figure 1A is consistent with categories in Figure 1B. We agree with the reviewer #3 that it is not explicit enough. We have tried to make this information more explicit.
The number highlighted in Liorzou et al. 2016 corresponds to different clusters in a continuous distribution, similarly to what we observe in the PCA. This number of clusters should be considered very cautiously. Here, we do not want to provide too many details regarding the model-based population structure as inferred with fastStructure . If by "reasonable K values" the reviewer #3 refers to the identification of a "best K" with an ad hoc method (e.g. Evanno et al. 2005), it is important to consider that these criteria also exhibit important biases (Waples & Gaggiotti, 2006 Mol Ecol; Verity & Nichols 2016 Genetics; Janes et al. 2017 Mol Ecol, among others), this is expected to be especially

true here given the deviations from the models (see also our replies to the reviewer #1). Consequently, we prefer to provide the barplots for inferences between K=2 to K=5, without providing the information of a best fit to the core fastStructure model, for a model we know to be partially wrong with our model and dataset. Furthermore, it is crucial to understand that our objective in this paper was not simply to infer population structure (Fig. 2A & 2B), but also to investigate the relatedness of the samples and how this aligns with the history of rose breeding (2C), as well as the global evolution of the genetic diversity (2D), the footprints of artificial sweeps (Fig. 3), as well as the genetic basis for some important traits (Fig. 4).

Other specific comments:

Abstract: I would suggest not using the term "Thanks to this study".
Done

Line 473: On rose GWAS browser, the GWAS result pages have no information.
We do not understand the comment of the reviewer here. Could the reviewer please provide additional information regarding any issues encountered while browsing the website? Specifically, we would appreciate details about the web browser used (e.g., Mozilla Firefox, Chrome, Safari, etc.). The reviewer may anticipate that a PDF containing the results will open on each webpage. Despite its simplicity (none of the authors are web designers), our rose GWAS browser website offers a viable alternative to a repository containing all the results (and consequently, PDFs). Without this browser, readers would need to sift through hundreds of PDFs to locate the desired information (all PDFs are available at: https://github.com/roseGWASbrowser/rosegwasbrowser.github.io/tree/main/PDF). We consider that our website effectively fulfills this role.

Line 610: It's unclear "newly sequenced samples were selected" for what?
We agree, "*selected*" was probably a confusing term- here. Thank you for pointing it out. The new sentence is: "*...each of the newly sequenced samples was chosen based on its presumed importance during the 19th century...* (l. 175-178)
Line 648: "higher is value" - check the grammar
Done (l. 212-214).

Line 663: The "aggressiveness factor" of a rose plant needs to be explained.
The "aggressiveness factor" does not refer to the roses but to the fungal strains, responsible for the blackspot disease (*D. rosae*). We made the hypothesis that the fungal strain diversity present during the 19th century is different from the present diversity in terms of virulence and aggressiveness, for the sake of clarity, we now refer to the evolution of the pathogen.

Line 765: Using sequencing depth (?) of alleles is a good idea to identify polyploids. As mentioned, the peak should be around 0.25, 0.33, or 0.5, but almost half of the samples have their peak around 0.15. How do the authors confirm their ploidy? The previous study (Lioorzou et al. 2016) showed ploidy ranging from 2x to 6x. Did the authors only pick the previously identified 2x & 4x individuals in this study?

It is a bit more complex than suggested by the reviewer. In brief, we used the minor allele frequency, which is expected to tend toward 0.25, 0.33 and 0.5. However, to observe such precise values it would require sequencing at very high depth of coverage (e.g. 200X or so). Indeed, with this high coverage, we would expect to empirically recover a MAF very close to 0.5 (e.g. a median value near 0.48). Why would this value not be exactly 0.5? Because a MAF is by definition ≤ 0.5, in such a way that the deviation in the frequency of the two sequenced alleles that occur by chance should be considered. With high coverage data, this deviation is limited. However, here we used relatively moderate coverage data, so the deviations from these theoretical classes are substantial. As a consequence, we cannot use this simple criteria, but we should rather consider the distributions for some individuals for which we have considerable knowledge regarding their ploidy level (e.g. typically, Old Blush for a textbook diploid example) and use this as an empirical scale to provide meaningful results. It is also crucial to understand here that we used information from Liorzou et al. 2016, as well as additional sources of information, to select the individuals for sequencing. Among the criteria, a ploidy level ≤ 4 was a main criteria.

Line 792: "higher than >0.354", "coefficients ranging from [0.177, 0.354], ..."
We have changed this sentence to “coefficients within the intervals [0.177, 0.354], [0.0884, 0.177], and [0.0442, 0.0884] correspond to 1st-degree, 2nd-degree, and 3rd-degree relationships, respectively” (l. 358-361).

Line 800-830: Since pi and many other population genetics parameters could be directly estimated from vcf, it is unclear why the authors wish to obtain "perfectly aligned sequence blocks" for pi estimation. Please explain more.

Without going too to much details, it is important to note that several methods have been developed to generate unbiased estimates of nucleotide diversity, all requires to generate a VCF with invariant sites (a “complete vcf”), e.g. pixy (Korunes & Samuk, 2021 Mol Ecol Res) or the approach developed by Leroy et al. (2020 Peer Community Journal; 2021 Current Biology). Here, the approach by Leroy et al. was extended to explicitly account for the variation in ploidy levels among individuals (the scripts were made available on github; https://github.com/roseGWASbrowser/PopGen_GWAS_19thcentury_roses/blob/main/popgenomics/NucleotideDiversity_DTaj_ROD/VCF2Fasta_fast_withcovquantiles_polyploid_outputfilenames.py). Both the approach developed by Korunes & Samuk and Leroy et al. are expected to provide very consistent values, since the rationale is the same, but to strongly deviate from values derived from variant-only vcf (e.g. vcftools, for more details regarding the methodological issue and empirical evidences, see Korunes & Samuk, 2021 Mol Ecol Res).

Line 840-850: If the ref genome is an Asian accession, the Asian population should mostly have reference allele, and therefore with alternative allele frequency 0 (corresponding to them having genotype calls of "allele 0" in the vcf). The description here seems opposite - is this actually referring to the frequency of "reference allele"?
We apologize for the lack of clarity, by convention the allele “0” is the reference allele, consequently, given that the reference genome is an Asian genotype, the Ancient Asian genotypes are expected to have a frequency of 1 for the allele “0”), i.e. f(“allele 0”)=1. Reciprocally, the Ancient European genotypes are expected to have a frequency of 0 for the allele “0”, i.e. f(“allele 0”)=0. We tried to make this information more explicit in the new version of the ms.