



Tinder in mice: A match made with the sense of smell

Christelle Fraïsse based on reviews by Ludovic Claude Maisonneuve, Angeles de Cara and 1 anonymous reviewer

A recommendation of:

Divergence of olfactory receptors associated with the evolution of assortative mating and reproductive isolation in mice

Carole M. Smadja, Etienne Loire, Pierre Caminade, Dany Severac, Mathieu Gautier, Guila Ganem(2022) *bioRxiv*, 2022.07.21.500634, ver. 3 peer-reviewed and recommended by Peer Community in Evolutionary Biology <https://doi.org/10.1101/2022.07.21.500634>

Open Access

Recommendation

Differentiation-based genome scans lie at the core of speciation and adaptation genomics research. Dating back to Lewontin & Krakauer (1973), they have become very popular with the advent of genomics to identify genome regions of enhanced differentiation relative to neutral expectations. These regions may represent genetic barriers between divergent lineages and are key for studying reproductive isolation. However, genome scan methods can generate a high rate of false positives, primarily if the neutral population structure is not accounted for (Bierne et al. 2013). Moreover, interpreting genome scans can be challenging in the context of secondary contacts between diverging lineages (Bierne et al. 2011), because the coupling between different components of reproductive isolation (local adaptation, intrinsic incompatibilities, mating preferences, etc.) can occur readily, thus preventing the causes of differentiation from being determined.

Smadja and collaborators (2022) applied a sophisticated genome scan for trait association (BAYPASS, Gautier 2015) to underlie the genetic basis of a polygenetic behaviour: assortative mating in hybridizing mice. My interest in this neat study mainly relies on two reasons. First, the authors used an ingenious geographical setting (replicate pairs of “Choosy” versus “Non-Choosy” populations) with multi-way comparisons to narrow down the list of candidate regions resulting from BAYPASS. The latter corrects for population structure, handles cost-effective pool-seq data and allows for gene-based analyses that aggregate SNP signals within a gene. These features reinforce the set of outlier genes detected; however, not all are expected to be associated with mating preference.

The second reason why this study is valuable to me is that Smadja et al. (2022) complemented the population genomic approach with functional predictions to validate the genetic signal. In line with previous behavioural and chemical assays on the proximal mechanisms of mating preferences, they identified multiple olfactory and vomeronasal receptor genes as highly significant candidates. Therefore, combining genomic signals with functional analyses is a clever way to provide insights into the causes of reproductive isolation, especially when multiple barriers are involved. This is typically

Published: 2022-11-16

Copyright: This work is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

true for reinforcement (Butlin & Smadja 2018), suspected to occur in these mice because, in that case, assortative mating (a prezygotic barrier) evolves in response to the cost of hybridization (for example, due to hybrid inviability).

As advocated by the authors, their study paves the way for future work addressing the genetic basis of reinforcement, a trait of major evolutionary importance for which we lack empirical data. They also make a compelling case using complementary approaches that olfactory and vomeronasal receptors have a central role in mammal speciation.

References:

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol Ecol* 20: 2044–2072. <https://doi.org/10.1111/j.1365-294X.2011.05080.x>

Bierne N, Roze D, Welch JJ (2013) Pervasive selection or is it...? why are F_{ST} outliers sometimes so frequent? *Mol Ecol* 22: 2061–2064. <https://doi.org/10.1111/mec.12241>

Butlin RK, Smadja CM (2018) Coupling, Reinforcement, and Speciation. *Am Nat* 191:155–172. <https://doi.org/10.1086/695136>

Gautier M (2015) Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates. *Genetics* 201:1555–1579. <https://doi.org/10.1534/genetics.115.181453>

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of selective neutrality of polymorphisms. *Genetics* 74: 175–195. <https://doi.org/10.1093/genetics/74.1.175>

Smadja CM, Loire E, Caminade P, Severac D, Gautier M, Ganem G (2022) Divergence of olfactory receptors associated with the evolution of assortative mating and reproductive isolation in mice. *bioRxiv*, 2022.07.21.500634, ver. 3 peer-reviewed and recommended by Peer Community in Evolutionary Biology. <https://doi.org/10.1101/2022.07.21.500634>

Cite this recommendation as:

Christelle Fraïsse (2022) Tinder in mice: A match made with the sense of smell. *Peer Community in Evolutionary Biology*, 100157. [10.24072/pci.evolbiol.100157](https://doi.org/10.24072/pci.evolbiol.100157)

Conflict of interest:

The recommender in charge of the evaluation of the article and the reviewers declared that they have no conflict of interest (as defined in [the code of conduct of PCI](#)) with the authors or with the content of the article.

Review

Evaluation round #1

DOI or URL of the preprint: <https://doi.org/10.1101/2022.07.21.500634>

Version of the preprint: 1

Author's Reply, 30 Oct 2022

[Download author's reply](#)[Download tracked changes file](#)

Decision by [Christelle Fraïsse](#), posted 18 Sep 2022

Thank you for your patience. Three reviewers have provided comments on your preprint. All generally agree that the work is sound – after my reading of the manuscript, I fully concur. I greatly appreciated the

integrative approach that combined population genomics with functional predictions. Congratulations on this exciting work.

The reviewers make a number of helpful suggestions for improvements that need to be addressed in a revision. In particular:

1) Reviewer 1 criticizes the methodology applied to identify candidate loci underlying assortative mating as it may underestimate the variability of mating preferences between individuals of the same locality. As it stands, assortative mating is considered a binary trait (Choosy vs Non-Choosy populations), while the genetic architecture suggests a polygenic trait. More information on the genetic variability of this trait in each of the four populations would be valuable.

Besides, I have a suggestion of my own. While the multi-way comparisons across populations are a bright setting to uncover the genetic basis of assortative mating, all differentiated genomic regions between Choosy and Non-Choosy populations are not involved in this trait. For example, neutral clines could produce genetic structure at specific loci between the Choosy populations (introgressed as they are near the hybrid zone) and Non-Choosy populations (not introgressed as they stand in allopatry). The authors could be a bit more critical on this point (it is acknowledged in the Discussion section but not really reflected elsewhere), and shed light on the importance of the functional predictions to narrow down the set of candidate variants.

2) All three reviewers would like clarifications on the BayPass methodology (see their specific comments). This is key for the reader because it is the first step to identifying putative variants of assortative mating. In particular, Reviewer 3 would like to see more precise expectations for the genomic patterns left by various scenarios (neutrality, selection on a single variant, polygenic selection) on the C2 statistics. There is no need to run simulations; just clarify your expectations, please.

As your genome scan on assortative mating indicates a complex trait, I recommend (if possible) running a method that can detect footprints of polygenic selection around the two candidate clusters on chromosomes 7 and 9. For example, diploS/HIC is a machine learning method that uses unphased genotypes to classify genomic windows in hard and soft sweeps: <https://doi.org/10.1534/g3.118.200262>.

3) I agree with Reviewer 3 that the Discussion section on the role of receptor genes is a bit long. I would recommend ending the manuscript as they suggest with an outlook on what sort of analyses can be done in the future.

As there are no major criticisms (and suggestions are reasonable to tackle), I believe this can lead to a recommendation as soon as it has been revised in response to the points raised. I am looking forward to receiving your revised preprint.

With best regards,

Christelle Fraïsse.

Reviewed by Ludovic Claude Maisonneuve, 29 Aug 2022

The preprint presents a population genomics study of mate preferences suspected to be implied reinforcement. The authors studied the correlation between SNPs/genes and the presence or absence of mate preference leading to reproductive isolation with a closely related subspecies (Choosy vs Non choosy). The authors identify olfactory receptors genes that may implied in reinforcements. Those genes mainly clusters in two areas in the genome.

Reinforcement is an important mechanism because it may be involved in speciation and explain why there is

so many species. However we lack of empirical data to determine the prevalence of this mechanism in natural populations. I then think this study has value because it investigates the genetic basis of mate preference suspected to be implied reinforcement. However I am not an expert on genomics, so I cannot judge the pertinence of the method used in this preprint.

My main concern about this study is that the authors did not study the correlation between the genotype and the mating behavior but between the genotype and the locality. If I understood well the study, the mating behavior was inferred from individual locally and not properly tested. This method may underestimate the variability of mating preference between individuals of a same locality.

Another concern is about the meaning of the C2 statistic. Much of the study is based on this statistic. However I did not understand how it is computed. At least it would be nice to have an idea of what it represents.

Also the authors did not explain the relatedness observed in figure 1b. I am surprised by this relatedness as it did not correspond to the spatial structure. It would be nice if the authors discuss about this.

I also have minor comments:

I27: what is a functional prediction ?

I30-31: 'which contrasts allele frequencies' it may be because I am unfamiliar with genomics technique but I don't understand what it means

I50: 'to understand how behavioural traits evolve' it is a bit vague, what do you mean ? To identify the mechanisms underlying behaviors ?

I52: maybe replace 'play' but 'may play' is more correct, see Servedio 2004 the what and why of research on reinforcement.

I62: why these cases are 'relatively simple' ?

I63: similarly what are 'more complex traits' ? Do you mean that the genetic basis implied a larger number of genes ?

I66: similarly what is a 'complex behavioral trait' ? Does that mean what there are simple behavioral trait ?

I166-168: 'a situation that makes the identification of outlier divergent loci in the Focal Test comparing Choosy and Non-Choosy populations conservative.' Can you explain in few words why ?

I212: I don't understand what are 'already known variants'

I512-516: this part is unclear for me. I do not understand why it would correspond to a matching rule.

Reviewed by anonymous reviewer, 13 Sep 2022

This paper investigated the genetic basis of an olfactory-based behavioral divergence in the model organism *Mus musculus*. To do this, this study compared populations with choosy vs. non-choosy behavior, while also contrasting populations with the same behavior as a control. To further establish the link between genetic divergence and phenotypic divergence, the study also examines divergence across all annotated genes. Moreover, this study also examines possible genomic signature of reinforcement, including selection sweeps among genetic differentiation outliers, and association between outlier genes and candidate hybrid sterility regions. I think this is a well-designed, well-executed, extensive research. The intro and discussion are well-written and make extensive connection to relevant literatures in the field. It is an important contribution to the methodology of uncovering genetic basis of behaviors, which is notoriously hard in natural systems. It serves as a rare study that examined genomic signature of reinforcement. I only have a few minor suggestions:

1. The exact comparisons in Focal vs. control test in Figure 1 C is not clear. My take is that Focal test means comparing choosy 1 vs. Non choosy 1; choosy 1 vs. Non-choosy 2; choosy 2 vs. Non choosy 1; choosy 2 vs.

non-choosy 2 , and that Control test is choosy 1 vs. choosy 2; Non-choosy 1 vs. Non-choosy 2. However, in the figure C, “vesus” symbol is used to show the comparisons, but it means different things in different tests. I suggest to write down each comparison in the figure legend and the result session.

2. Line 295-306. the title of the section states no signatures of selective sweeps confuses me. In the text, there are some genes showed selective sweeps in both Choosy populations.

3. Figure 4. I am a bit overwhelmed by different font sizes and different colors of genes in the figure. I understand different color means different functional receptors. However, what do different sizes font mean? The symbol “<”, “>” means where the gene starts and ends? No annotations of these symbol meanings are found in the figure legend.

4. Because the key findings in this study is built upon program BAYPASS to tease apart the effect of population structure on signals of genetic divergence related to choosy behavior, I hope there is a bit more mechanistic explanation about how this program corrects for the population structure in the introduction. Right now, this information is buried in the method section.

Reviewed by Angeles de Cara, 14 Sep 2022

This is a very nice study and I have enjoyed and learnt while reading it. The goal of the study is to find the genetic basis of olfactory-based mating preferences in mice. For that purpose, the authors sampled and sequenced individuals from the border of a hybrid zone, where strong assortment occurs, and from other locations away from the border of this zone, where no assortment has been identified.

By splitting and contrasting individuals into choosy (from the border of the hybrid zone) and non-choosy, the authors obtain candidate regions and genes and then do functional analyses where they identify multiple olfactory and vomeronasal receptor genes as outliers.

Overall, I do think the goals are clearly stated, the paper is well written, and the methods are appropriate and well used.

I really enjoyed the discussion, although the section on the role of receptor genes seemed a bit long. I believe a bit more insight into what sort of analyses can be done in the future, especially for those readers not familiar with the work of Isogai et al (2011), would be useful to close down the manuscript.

Some issues I have:

It seems there is no reduced diversity in the outlier regions, and no signs of selective sweeps, and all analyses are indicative of polygenic selection. That seems to be the case both for outlier genes in C2_max or in C2_mean, while I would have expected some different in diversity on those two scenarios. On the whole, I would appreciate whether some findings, like no reduced diversity and no overlap with candidate hybrid sterility genes were expected or unexpected, or what could be the explanation for either of these findings.

Are there simulations to know which scenarios of selection / genetic architecture result in outliers of C2_max and which in outliers of C2_mean?

Minor comments:

- I suppose the order of methods at the end corresponds to the authors' final choice of journal, but I personally find it easier to follow when methods are described prior to results.
- l. 215 and l. 221: coding consequences?
- The sentence finishing in l. 270 should be rewritten ("were less found in the very most significantly")
- l. 749: with -> which

In Fig. 3, the three clusters could be magnified to better appreciate the structure.

In Supp. Fig.1, text size is too small.