



Peer Community In Evolutionary Biology

Multi-gene and lineage comparative assessment of the strength of selection in Hymenoptera

Bertanne Visser based on peer reviews by **Michael Lattorff** and 1 anonymous reviewer

Arthur Weyna, Jonathan Romiguier (2021) Relaxation of purifying selection suggests low effective population size in eusocial Hymenoptera and solitary pollinating bees. Missing preprint_server, ver. Missing article_version, peer-reviewed and recommended by Peer Community in Evolutionary Biology. <https://doi.org/10.1101/2020.04.14.038893>

Submitted: 21 April 2020, Recommended: 05 February 2021

Cite this recommendation as:

Visser, B. (2021) Multi-gene and lineage comparative assessment of the strength of selection in Hymenoptera. *Peer Community in Evolutionary Biology*, 100120. [10.24072/pci.evolbiol.100120](https://doi.org/10.24072/pci.evolbiol.100120)

Published: 05 February 2021

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

Genetic variation is the raw material for selection to act upon and the amount of genetic variation present within a population is a pivotal determinant of a population's evolutionary potential. A large effective population size, i.e., the ideal number of individuals experiencing the same amount of genetic drift and inbreeding as an actual population, N_e (Wright 1931, Crow 1954), thus increases the probability of long-term survival of a population. However, natural populations, as opposed to theoretical ones, rarely adhere to the requirements of an ideal panmictic population (Sjödín et al. 2005). A range of circumstances can reduce N_e , including the structuring of populations (through space and time, as well as age and developmental stages) and inbreeding (Charlesworth 2009). In mammals, species with a larger body mass (as a proxy for lower N_e) were found to have a higher rate of nonsynonymous nucleotide substitutions (that alter the amino acid sequence of a protein), as well as radical amino acid substitutions (altering the physicochemical properties of a protein) (Popadin et al. 2007). In general, low effective population sizes increase the chance of mutation accumulation and drift, while reducing the strength of selection (Sjödín et al. 2005). In this paper, Weyna and Romiguier (2021) set out to test if parasitism, body size, geographic range, and/or eusociality affect the strength of selection in Hymenoptera. Hymenoptera include the bees, wasps and ants and is an extraordinarily diverse order within the insects. It was recently estimated that Hymenoptera is the most speciose order of the animal kingdom (Forbes et al. 2018). Hymenoptera are further characterized by an impressive radiation of parasitic species, mainly parasitoids, that feed in or on a single host individual to complete their own development (Godfray 1994). All hymenopterans share the same sex determination system: haplo-diploidy, where unfertilized eggs

are haploid males and fertilized eggs are diploid females. Compared to other animals, Hymenoptera further contain an impressive number of clades that evolved eusociality (Rehan and Toth 2015), in which societies show a clear division of labor for reproduction (i.e., castes) and cooperative brood care. Hymenopterans thus represent a diverse and interesting group of insects to investigate potential factors affecting strength of selection and N_e . Using a previously published phylogenomic dataset containing 3256 genes and 169 hymenopteran species (Peters et al. 2017), Weyna and Romiguier (2021) estimated mean genomic dN/dS ratios (nonsynonymous to synonymous substitution rates) for each species and compared these values between parasitic and non-parasitic species, eusocial and solitary species, and in relation to body size, parasitoid-specific traits and geographic range, thought to affect the effective population size and strength of selection. The use of a large number of species, as well as several distinct traits is a clear asset of this study. The authors found no effect of body size, geographic range or parasitism (including a range of parasitoid-specific traits). There was an effect, however, of eusociality where dN/dS increased in three out of four eusocial lineages. Future studies including more independent evolutionary transitions to eusociality can lend further support that eusocial species indeed reduces the efficiency of selection. The most intriguing result was that for solitary and social bees, with high dN/dS ratios and a strong signature of relaxed selection (i.e., the elimination or reduction of a source of selection (Lahti et al. 2009). The authors suggest that the pollen-collecting behaviors of these species can constrain N_e , as pollen availability varies at both a spatial and temporal scale, requiring a large investment in foraging that may in turn limit reproductive output. It would be interesting to see if other pollen feeders, such as certain beetles, flies, butterflies and moths, as well as mites and spiders, experience relaxed selection as a consequence of the trade-off between energy investment in pollen foraging versus fecundity.

References:

- Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, 10(3), 195-205. doi: [\[https://doi.org/10.1038/nrg2526\]](https://doi.org/10.1038/nrg2526)(<https://doi.org/10.1038/nrg2526>)
- Crow, J. F. (1954) *Statistics and Mathematics in Biology* (eds Kempthorne, O., Bancroft, T. A., Gowen, J. W. & Lush, J. L.) 543-556 (Iowa State Univ. Press, Ames, Iowa)
- Forbes, A. A., Bagley, R. K., Beer, M. A., Hippee, A. C., and Widmayer, H. A. (2018). Quantifying the unquantifiable: why Hymenoptera, not Coleoptera, is the most speciose animal order. *BMC ecology*, 18(1), 1-11. doi: [\[https://doi.org/10.1186/s12898-018-0176-x\]](https://doi.org/10.1186/s12898-018-0176-x)(<https://doi.org/10.1186/s12898-018-0176-x>)
- Godfray, H. C. J. (1994) *Parasitoids: Behavioral and Evolutionary Ecology*. Vol. 67, Princeton University Press, 1994. doi: [\[https://doi.org/10.2307/j.ctvs32rmp\]](https://doi.org/10.2307/j.ctvs32rmp)(<https://doi.org/10.2307/j.ctvs32rmp>)
- Lahti et al. (2009). Relaxed selection in the wild. *Trends in ecology & evolution*, 24(9), 487-496. doi: [\[https://doi.org/10.1016/j.tree.2009.03.010\]](https://doi.org/10.1016/j.tree.2009.03.010)(<https://doi.org/10.1016/j.tree.2009.03.010>)
- Peters et al. (2017). Evolutionary history of the Hymenoptera. *Current Biology*, 27(7), 1013-1018. doi: [\[https://doi.org/10.1016/j.cub.2017.01.027\]](https://doi.org/10.1016/j.cub.2017.01.027)(<https://doi.org/10.1016/j.cub.2017.01.027>)
- Popadin, K., Polishchuk, L. V., Mamirova, L., Knorre, D., and Gunbin, K. (2007). Accumulation of slightly deleterious mutations in mitochondrial protein-coding genes of large versus small mammals. *Proceedings of the National Academy of Sciences*, 104(33), 13390-13395. doi: [\[https://doi.org/10.1073/pnas.0701256104\]](https://doi.org/10.1073/pnas.0701256104)(<https://doi.org/10.1073/pnas.0701256104>)

Rehan, S. M., and Toth, A. L. (2015). Climbing the social ladder: the molecular evolution of sociality. Trends in ecology & evolution, 30(7), 426-433. doi: <https://doi.org/10.1016/j.tree.2015.05.004>

Sjödín, P., Kaj, I., Krone, S., Lascoux, M., and Nordborg, M. (2005). On the meaning and existence of an effective population size. Genetics, 169(2), 1061-1070. doi: <https://doi.org/10.1534/genetics.104.026799>

Weyna, A., and Romiguier, J. (2021) Relaxation of purifying selection suggests low effective population size in eusocial Hymenoptera and solitary pollinating bees. bioRxiv, 2020.04.14.038893, ver. 5 peer-reviewed and recommended by PCI Evol Biol. doi: <https://doi.org/10.1101/2020.04.14.038893>

Wright, S. (1931). Evolution in Mendelian populations. Genetics, 16(2), 97-159.

Reviews

Evaluation round #3

DOI or URL of the preprint: <https://doi.org/10.1101/2020.04.14.038893>

Version of the preprint: <https://www.biorxiv.org/content/10.1101/2020.04.14.038893v2>

Authors' reply, 12 January 2021

[Download author's reply](#)

[Download tracked changes file](#)

Decision by Bertanne Visser, posted 24 December 2020

Minor revisions.

Dear Drs. Weyna and Romiguier,

Thank you for submitting your revisions. I appreciate the changes made, but some of the suggestions that I provided in the PDF of the last version were not included in the current file (changes that I suggested with regard to the parasitoid table for example). Please have a look at my questions and suggestions in the previous version and incorporate these in your manuscript (or if you do not wish to do so, it would be helpful to know why). For the file containing track changes, it would be useful to see what was actually changed (as now often the entire paragraph is marked as deleted/changed). I have also included some additional suggestions in your latest version, please find this attached. Thanks in advance.

All the best, Bertanne Visser [Download recommender's annotations](#)

Evaluation round #2

DOI or URL of the preprint: <https://doi.org/10.1101/2020.04.14.038893>

Version of the preprint: <https://www.biorxiv.org/content/10.1101/2020.04.14.038893v2>

Authors' reply, 15 December 2020

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Bertanne Visser](#), posted 22 October 2020

Decision for Weyna & Romiguier, bioRxiv doi: <https://doi.org/10.1101/2020.04.14.038893>, for PCI Evolutionary Biology

Dear Drs Weyna and Romiguier,

Thank you for submitting your revision, which has now been reviewed again. I think you did a good job addressing the comments and suggestions in your rebuttal and revision. There is one major point that still requires some attention: One of the reviewers is still concerned with the number of eusocial species included in your analysis and the conclusions that you draw based on that limited number of species. While I understand the rationale for sticking to the dataset of Peters (which you have explained very clearly in your rebuttal) I agree with the reviewer that a thorough test would require a larger number of species. I think you should explicitly mention that in your manuscript, make the proposition on how to test this based on the reviewer's suggestion, and put less emphasis on this finding throughout your manuscript (e.g., the largest part of your discussion still focuses mainly on eusociality). As you mention in your rebuttal, you consider the high dN/dS ratio in social and non-social pollinating bees, as well as the *Antophila*, as your most novel findings. I think it will be worthwhile to spend more time on this in your discussion and to provide a perspective for future studies.

It seems to me that you wrote this paper (at least initially) within a context that fit with the results that you obtained, even though I think you started out this project with less specific expectations (i.e., you included parasitism for which you collected an incredible amount of data). Once you obtained these first results, I think you started testing more specific ideas and making appropriate corrections where needed. And that is all fine. If this is indeed the case, I think your paper would be stronger if you bluntly state that and take the reader through the successive steps and the thought process that led you to your final results.

I have further carefully read your revision and provided minor comments in the PDF.

Thanks again for your thorough revision and rebuttal and I look forward to receiving a new version.

All the best, Bertanne Visser

Additional requirements of the managing board:

As indicated in the 'How does it work?' section and in the code of conduct, please make sure that:

- Data are available to readers, either in the text or through an open data repository such as Zenodo (free), Dryad or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data.
- Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, codes) are available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or codes must be carefully described so that they can be reused.
- Details on experimental procedures are available to readers in the text or as appendices.
- Authors have no financial conflict of interest relating to the article. The article must contain a "Conflict of interest disclosure" paragraph before the reference section containing this sentence: "The authors of this preprint declare that they have no financial conflict of interest with the content of this article." If appropriate, this disclosure may be completed by a sentence indicating that some of the authors are PCI recommenders: "XXX is one of the PCI XXX recommenders." [Download recommender's annotations](#)

Reviewed by [Michael Lattorff](#), 26 September 2020

The revised manuscript "Relaxation of purifying selection suggests low effective population size in eusocial Hymenoptera and pollinating bees" by Weyna & Romiguier uploaded to bioRxiv and under review for PCI Evolutionary Biology is a bit difficult to review as no responses to reviewers were delivered.

However, it seems that the author addressed all major conceptual issues I raised during the first review. There are just a few minor issues listed below that I could point out now, as line numbers were used in the

revised version of the manuscript.

L29, L47 L61, L74, L79, L240, L251, L252, L253, L262, L277 Ne in italics

L75 ...expected to rely...

L89, L101 hymenopteran species or species of Hymenoptera

L154 literature

L166 R2, btw, what is the justification of using that many digits? Can you justify that type of precision and accuracy by the methods used to calculate, I think it implies higher precision than you actually have. I think you can stick with 2-3 significant figures.

L216 relaxation of purifying selection

L270 too many brackets for the citation

In some place you might have used two subsequent spaces. Check and replace with single space.

Reviewed by anonymous reviewer 1, 06 September 2020

The authors have done a decent job of addressing most of my previous comments. However, I still remain concerned that they have such few eusocial species in this dataset and are trying to write a paper about eusociality. The genomes you need to do this exist and while I realize it would take a lot of work to redo all of your analyses, I still worry that your results do not reflect what is really going on. Perhaps someone will use your paper as a hypothesis to test this idea across a much larger dataset.

Evaluation round #1

DOI or URL of the preprint: <https://doi.org/10.1101/2020.04.14.038893>

Authors' reply, 27 August 2020

[Download author's reply](#)

[Download tracked changes file](#)

Decision by Bertanne Visser, posted 12 June 2020

Decision on your manuscript submitted to the Peer Community in Evolutionary Biology.

Dear Drs. Weyna and Romiguier,

I have now received two reviews for your manuscript 'Relaxation of purifying selection suggests low effective population size in eusocial Hymenoptera and pollinating bees'. One of the reviewers highlights that your study is very intriguing, and I agree, but both reviewers raise several points that need to be addressed in your revision. In particular, a key question raised by both reviewers is why only 12 eusocial taxa were included in your analyses. Furthermore, the role of inbreeding as a prerequisite for the evolution of eusociality was also raised by both reviewers. Aside from these two major points, please address each of the reviewers' comments carefully in your rebuttal and revision. Please also include line numbers in your revision.

Thanks for submitting your work to the Peer Community in Evolutionary Biology and I look forward to receiving your rebuttal and revision.

Best wishes,

Bertanne Visser.

Additional requirements of the managing board:

As indicated in the 'How does it work?' section and in the code of conduct, please make sure that:

-Data are available to readers, either in the text or through an open data repository such as Zenodo (free), Dryad or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data.

-Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, codes) are available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or codes must be carefully described so that they can be reused.

-Details on experimental procedures are available to readers in the text or as appendices.

-Authors have no financial conflict of interest relating to the article. The article must contain a "Conflict of interest disclosure" paragraph before the reference section containing this sentence: "The authors of this preprint declare that they have no financial conflict of interest with the content of this article." If appropriate, this disclosure may be completed by a sentence indicating that some of the authors are PCI recommenders: "XXX is one of the PCI XXX recommenders."

Reviewed by **Michael Lattorff**, 26 May 2020

[Download the review](#)

Reviewed by anonymous reviewer 1, 01 June 2020

This study examines the relationship between eusociality, selection, and effective population size (N_e) across 169 Hymenoptera, including about half of which are bees. The authors demonstrate a pattern of relaxed selection in eusocial lineages, which they attribute to differences in N_e . I think the data presented in this study are very intriguing. However, I have some concerns with the way the data are analyzed and interpreted. Some of the modeling is a bit repetitive and could be simplified to make things more transparent. I also worry that the patterns are being driven by one subtaxa, given the sparse sampling of eusocial species in other lineages. Some of the decisions to categorize these lineages as eusocial or not is also a bit concerning. More importantly, I am not sure the authors are correctly portraying hypotheses about social evolution. They seem a bit confused by differences between inbreeding and kin selection (the maintenance of high relatedness in groups). With the exception of some social spider species, most insects (and vertebrates and other taxa), AVOID inbreeding because the costs of doing so generally outweigh the benefits. N_e is low in eusocial species because of a reproductive division of labor and high reproductive skew where most individuals in the population will never breed. This is a consequence of the social structure, not a cause as the authors imply. In general, while this study presents some interesting results, I think some more work is needed on the analysis, the interpretation of the results, and the overall framework being tested.

Without line numbers, it is hard to make specific comments, so I have made some general ones below:

Introduction

I am not sure it is fair to say most invertebrates have larger and less variable long-term N_e . I would wager that eusocial insects have the most genetic data, and they are also likely to have the lowest N_e because of their reproductive division of labor.

Inbreeding has not been invoked to explain the evolution of eusociality. The maintenance of high relatedness within groups has been raised as a prerequisite, but not inbreeding.

Hymenoptera are also "special" because they are haplodiploid, which could result in lower N_e . You mention this later in introduction, but it is a bit misleading to not mention this when you introduce Hymenoptera earlier.

I think the end of the Introduction should briefly describe the approach taken to compare these species and look at purifying selection. You should also make it clear that only 12 of the 169 species examined are eusocial. I am actually surprised by how low this number. Certainly, there are many more than 12 complete genomes published for ants, bees, wasps. Why have the authors only chosen 12 of these is unclear. It would certainly improve your paper to use genetic data from more eusocial species, which is definitely available as of 2020.

Results

In the analyses in "Effect of body size and parasitism on relative protein evolution rates" it is unclear how you control for phylogenetic non-independence. In the methods you only say that you use PICs, but no information

on how you actually do this. Currently, it seems like these are straight linear models, but these should all be done in a phylogenetic framework. I would only present the results from the phylogenetic analyses.

In the analyses in the "Eusociality is associated to accelerated relative protein evolution rates" the presentation of the results is a bit unclear if not misleading. You can't claim a branch is eusocial because sample you have not sampled are eusocial. You really should work with the data (and species) you have here. Adding new taxa might change your genetic results, not just which branches you label as eusocial or not. I suggest presenting one more in the beginning of the paragraph and then one set of results that takes into account species you have sampled.

In the analyses in the "Antophila bees displays highly accelerated relative protein evolution rates", are you basically saying that this group is driving all of your results? I have trouble following this section, and not sure it needs to be its own sub-section.

Rather than having an entire section "Controlling for branch lengths, biased gene conversion or species sampling", why not just include branch length in all of your prior models? Similarly, why not use corrected genomic dN/dS values from the outset, rather than saying doing so gives similar results? In other words, rather than replicating analyses multiple times throughout this paper, I think it would be much improved if you simplify this by explaining all of this in the methods and then presenting fewer models that control for various factors and use standardized values.

Discussion

I think you are again confused by the role of inbreeding in the evolution of eusociality. Everywhere you discuss inbreeding, I think you mean kin selection and/or high relatedness among group members.

I also think you might have the order of events backwards when you discuss the relationship between N_e and eusociality. Eusocial species have low N_e because of a reproductive division of labor and high reproductive skew within groups. N_e is therefore a consequence of the social system rather than a cause. When you say things like "...eusociality can be seen as a complex trait that can evolve only in taxa with low N_e , where selection is not strong enough to maintain simpler organisations.", I think your logic is flawed. I don't know of a single person studying social evolution in insects or any other organism that would agree with this statement. I think your directionality is backwards and you need to think that high skew results in low N_e . The selective forces that driving group influence skew, which in turn influences N_e .