Peer Community In Evolutionary Biology

Replication-independent mutations: a universal signature ?

Nicolas Galtier based on peer reviews by *Marc Robinson-Rechavi* and *Robert Lanfear*

Guillaume Achaz, Serge Gangloff, Benoit Arcangioli (2019) The quiescent X, the replicative Y and the Autosomes. Missing preprint_server, ver. Missing article_version, peer-reviewed and recommended by Peer Community in Evolutionary Biology. 10.1101/351288

Submitted: 25 July 2018, Recommended: 05 February 2019

Cite this recommendation as:
Galtier, N. (2019) Replication-independent mutations: a universal signature ?. *Peer Community in Evolutionary Biology*, 100066. 10.24072/pci.evolbiol.100066

Published: 05 February 2019

Mutations are the primary source of genetic variation, and there is an obvious interest in characterizing and understanding the processes by which they appear. One particularly important question is the relative abundance, and nature, of replication-dependent and replication-independent mutations - the former arise as cells replicate due to DNA polymerization errors, whereas the latter are unrelated to the cell cycle. A recent experimental study in fission yeast identified a signature of mutations in quiescent (=non-replicating) cells: the spectrum of such mutations is characterized by an enrichment in insertions and deletions (indels) compared to point mutations, and an enrichment of deletions compared to insertions [2]. What Achaz *et al.* [1] report here is that the very same signature is detectable in humans. This time the approach is indirect and relies on two key aspects of mammalian reproduction biology: (1) oocytes remain quiescent over most of a female's lifespan, whereas spermatocytes keep dividing after male puberty, and (2) X chromosome, Y chromosome and autosomes spend different amounts of time in a female *vs.* male context. In agreement with the yeast study, Achaz *et al.* show that in humans the male-associated Y chromosome, for which quiescence is minimal, has by far the lowest ratios of indels to point mutations and of deletions to insertions, whereas the female-associated X chromosome has the highest. This is true both of variants that are polymorphic among humans and of fixed differences between humans and chimpanzees. So we appear to be here learning about an important and general aspect of the mutation process. The authors suggest that, to a large extent, chromosomes tend to break in pieces at a rate that is proportional to absolute time - because indels in quiescent stage presumably result from double-strand DNA breaks. A very recent analysis of numerous mother-father-child trios in humans confirms this prediction in demonstrating an effect of maternal age, but not of paternal age, on the recombination rate [3]. This result also has important implications with respect to the interpretation of substitution rate variation among taxa and genomic compartments, particularly

mitochondrial *vs.* nuclear, and their relationship with the generation time and longevity of organisms (e.g. [4]).

References:

[1] Achaz, G., Gangloff, S., and Arcangioli, B. (2019). The quiescent X, the replicative Y and the Autosomes. BioRxiv, 351288, ver. 3 peer-reviewed and recommended by PCI Evol Biol. doi: [10.1101/351288](https://dx.doi.org/10.1101/351288)

[2] Gangloff, S., Achaz, G., Francesconi, S., Villain, A., Miled, S., Denis, C., and Arcangioli, B. (2017). Quiescence unveils a novel mutational force in fission yeast. eLife, 6:e27469. doi: [10.7554/eLife.27469](https://dx.doi.org/10.7554/eLife.27469)

[3] Halldorsson, B. V., Palsson, G., Stefansson, O. A., Jonsson, H., Hardarson, M. T., Eggertsson, H. P., ... Stefansson, K. (2019). Characterizing mutagenic effects of recombination through a sequence-level genetic map. Science, 363: eaau1043. doi: [10.1126/science.aau1043](https://dx.doi.org/10.1126/science.aau1043)

[4] Saclier, N., François, C. M., Konecny-Dupré, L., Lartillot, N., Guéguen, L., Duret, L., ... Lefébure, T. (2019). Life History Traits Impact the Nuclear Rate of Substitution but Not the Mitochondrial Rate in Isopods. Molecular Biology and Evolution, in press. doi: [10.1093/molbev/msy247](https://dx.doi.org/10.1093/molbev/msy247)

Reviews

Evaluation round #1

DOI or URL of the preprint: https://doi.org/10.1101/351288
Version of the preprint: 1

Authors' reply, 29 January 2019

Download author's reply
Download tracked changes file

Decision by Nicolas Galtier, posted 08 November 2018

Revise

Achaz et al. report a simple but highly meaningful observation: the ratio of indels to point mutations, and of deletions over insertions, differ between X, Y and autosomes in humans. Why is this meaningful? Because (1) the results are fully consistent with the hypothesis that indels are more frequent than point mutations in quiescent oocytes (hence the X>autosomes>Y ranking), and (2) this very pattern has been experimentally demonstrated to occur in yeast. So we appear to be here learning about an important and general aspect of the mutation process. The two reviewers agree that this is an important result. They provide a number of useful suggestions, which should help improve the manuscript further. In particular, the authors should cite and take into account the last publications in this rapidly moving field (I take this opportunity to insert my and PCI Evol Biol's apologies about the slowness of the process), and make sure they provide a reliable, long-term public distribution of their source code.

Additional requirements of the managing board

We ask you to carefully verify that your manuscript complies with the following requirements (indicated in the

'How does it work?' section and in the code of conduct) and to modify your manuscript accordingly:
-Data must be available to readers after recommendation, either in the text or through an open data repository such as Zenodo, 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) must be 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 must be available to readers in the text or as appendices.
-Authors must 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."
This disclosure may be completed by a sentence indicating that some of the authors are PCI recommenders: "XY is one of the PCI Evol Biol recommenders."

Reviewed by Marc Robinson-Rechavi, 27 September 2018

In this manuscript, Achaz et al present an interesting extension to sex chromosomes of their previous observation of quiescent patterns of mutation in yeasts. In yeasts, they have previously reported that mutations biased towards indels accumulate during quiescence. In human, they now report that the X chromosome, which spends more time in quiescent oocytes, has a similar indel-biased mutation pattern, whereas the Y has the opposite pattern. This is especially interesting in the context of the recent results on mutational rates and patterns in humans.

Major comments:

The Y chromosome is enriched in low complexity regions, which complicate read mapping and mutation calls. The assertion p. 4 that "sequencing errors would affect all chromosomes equally" is at best an hypothesis to be tested. Indeed, it is contradicted by the statement in the Methods that accessible sites represent "on average 90% of the chromosome size, with the exception of the Y where it is 18% of the chromosome". I recommend that the authors take this issue into account both in the analysis and in the discussion.

All over the manuscript, the term "indels" is used, but in the discussion the authors specify that the pattern is driven by deletions. If that is the case, then I recommend specify "deletions" everywhere in the manuscript where it is possible to make that call.

Be careful of distinction between frequently occurring mutation, and mutation which has increased to high frequency in the population: "Similarly, frequent mutations that are more likely to get fixed in humans (Minor Allele Frequency > 0.01) also exhibit a higher fraction of indels."

Whether the "low proportion of indels" is "due to the high density of coding sequences" could be easily tested by comparing patterns within and outside of coding sequences.

p. 3 "likely because they contain many "slightly deleterious" alleles (15–17)." The references cited do not support this claim.

p. 4 references 7 and 19 do not analyse indels, so the relation to the results of this manuscript should be clarified.

I don't understand the last sentence of the Discussion: why assume that the pattern observed should be advantageous?

For calling indels, multiple sequence alignments would be more accurate than pairwise alignments, and should be preferred.

The paragraph "Statistical significance" in the Methods is an odd mix of methods, results and interpretation.

End of Methods, the assertion "close to the observed I vector" is not very clear; what is "close"?

Additional literature which I recommend citing and discussing: Makova et al (Genome Res. 2004. 14: 567-573 10.1101/gr.1971104) report male bias in indels in rat and mouse. Jónsson et al (Nature 549, 519–522 2017)

report little relation of indel patterns with age or sex of parents; stronger paternal slope with age for indels (Sup Table 9); and strongest effect of father, most significant of father-age (sup table 11). Makova & Hardison (2015 Nature Reviews Genetics 16, 213–223) report that indels are influenced by DNA environment, including chromatin, in correlation with other mutational patterns. Are there differences between X, Y, and autosomes? What is the influence of being in spermatogenesis (see preprint Xia et al)? Xia et al (preprint https://www.biorxiv.org/content/early/2018/03/14/282129) report very important results on the role of transcription in testes for mutational rates and patterns: "Widespread transcriptional scanning in testes modulates gene evolution rates". Gao et al (https://www.biorxiv.org/content/early/2018/05/22/327098) report that the view that "germline point mutations stem primarily from DNA replication errors" should be "called into question". Finally, Thomas et al (Current Biol 28, 2018, 3193-3197.e5) have recently reported a relation between longevity and mutation rates in owl monkeys, relative to human and chimpanzee.

Minor comment:

The term gonosomes is not widely used in the literature on this topic; it would be clearer to use "sex chromosomes".

Reviewed by **Robert Lanfear**, 27 September 2018

This paper uses a beautifully simple approach to generalise a fundamentally important finding about molecular evolution from yeasts to primates. The authors had previously observed that quiescent yeast cells accumulate indels and SNVs in roughly equal proportion, but that replicating yeast cells accumulate proportionally more SNVs. Using the nature of the mammalian germline as a natural experiment, the authors very convincingly show that exactly the same pattern exists in human genomes, and that it also occurs when comparing human and chimp genomes. The broadening of the scope of the pattern from yeasts to primates suggests that, perhaps, this pattern is something that we might see conserved across much of the tree of life. The paper is written clearly, simply, and concisely. The result is both very convincing and very exciting.

The only major comment I have is that I would like to see the code used for the analyses archived somewhere with a DOI – this would assist others in replicating or building on the work presented here, and would also assist readers and reviewers in assessing the details of the way that the analyses were performed. Suitable venues for the code would be Data Dryad, FigShare, or Zenodo. Perhaps the most preferable way to archive the code is to first upload it to GitHub, and then mint a DOI for the github repository using Zenodo. Instructions for this can be found here: https://guides.github.com/activities/citable-code/. Generally though, as long as the code has a DOI, and the DOI is presented in the manuscript, that is all that is required.

Minor comments

1. I suggest using 'sex chromosomes' in place of 'gonosomes'. Both are correct of course, but my suggestion is based on the observation that evolutionary biology papers more commonly refer to the x and y as 'sex chromosomes'.

2. Similarly, I suggest replacing 'neo-mutations' with 'de-novo mutations', because the latter is more commonly used.

3. I found this sentence hard to understand: "Interestingly, at the divergence level, the degenerating Y chromosome accumulates more indels and SNVs than any other chromosome (Fig. 1), a pattern that was reported previously (12,13).". After looking at figure 1, I see that by 'at the divergence level' you mean when comparing humans and chimps. I think it would be useful to clarify this. One suggestion would be to rephrase in terms of substitutions, something like "Interestingly, when comparing humans and chimpanzees, the degenerating Y chromosome has accumulated more indel and single-nucleotide substitutions..."

4. "striking simple" should be "strikingly simple"

5. “We also noticed the negative correlation between deletions and chromosome size but did not further investigate it yet.” – why not just report the statistics of the correlation? This seems like a simple result to report.

6. I do not find this argument quite as accurate as I think it could be: “This result supports the view arguing that SNVs are more efficiently removed by purifying selection in the long run (14) likely because they contain many “slightly deleterious” alleles (15–17).”. I think the lack of accuracy stems from the lack of comparison between indel and SNV fitness effects. For indels to be removed more by purifying selection, it is necessary and sufficient to assume that they have, on average, more negative selection coefficients than SNVs. So, a statement only about the selection coefficients of indels doesn’t establish this.

7. “Because of the haploidy” should be “because of haploidy”

8. “in complete line” should be “completely in line”

Rob Lanfear