Peer Community In Genomics

# Anopheles coluzzii, a new system to study how transposable elements may foster adaptation to urban environments

***Anne Roulin*** *based on peer reviews by* ***Yann Bourgeois*** *and 1 anonymous reviewer*

Carlos Vargas-Chavez, Neil Michel Longo Pendy, Sandrine E. Nsango, Laura Aguilera, Diego Ayala, and Josefa González (2021) Uncovering transposable element variants and their potential adaptive impact in urban populations of the malaria vector Anopheles coluzzii. Missing preprint_server, ver. Missing article_version, peer-reviewed and recommended by Peer Community in Genomics. https://doi.org/10.1101/2020.11.22.393231

Submitted: 02 December 2020, Recommended: 27 April 2021

**Cite this recommendation as:**
Roulin, A. (2021) Anopheles coluzzii, a new system to study how transposable elements may foster adaptation to urban environments. *Peer Community in Genomics*, 100006. 10.24072/pci.genomics.100006

Published: 27 April 2021


---

Transposable elements (TEs) are mobile DNA sequences that can increase their copy number and move from one location to another within the genome [1]. Because of their transposition dynamics, TEs constitute a significant fraction of eukaryotic genomes. TEs are also known to play an important functional role and a wealth of studies has now reported how TEs may influence single host traits [e.g. 2–4]. Given that TEs are more likely than classical point mutations to cause extreme changes in gene expression and phenotypes, they might therefore be especially prone to produce the raw diversity necessary for individuals to respond to challenging environments [5,6] such as the ones found in urban area.

In their study [7], Vargas et al. establish the foundation to investigate how TEs may help Anopheles coluzzii - the primary vectors of human malaria in sub-Saharan Africa - adapt to urban environments. To cover natural breeding sites in major Central Africa cities, they made use of the previously available An. coluzzii genome from Yaoundé (Cameroon) and sequenced with long-read technology six additional ones originating from Douala (Cameroon) and Libreville (Gabon). The de novo annotation of TEs in these genomes revealed 64 new anopheline TE families and allowed to identify seven active families. As a first step towards characterizing the potential role of TEs in the adaptation of An. coluzzii to urban environments, they further analyzed the distribution of TEs across the seven genomes. By doing so, they identified a significant number of polymorphic or fixed TE insertions located in the vicinity of genes involved in insecticide resistance and immune response genes.

The availability of seven An. coluzzii genomes allowed the authors to explore how TE diversity may affect genes functionally relevant for the adaptation to urban environments and provide ground for further functional validation studies. More and more studies have demonstrated the impact of TEs on adaptation and as such, the work of Vargas et al. contributes to fostering our understanding of the link between TEs and gain of function in a species facing strong anthropogenic pressures. **References**


[1] Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. Nature Reviews Genetics, 8, 973–982. https://doi.org/10.1038/nrg2165

[2] van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ (2016) The industrial melanism mutation in British peppered moths is a transposable element. Nature, 534, 102–105. https://doi.org/10.1038/nature17951

[3] González J, Karasov TL, Messer PW, Petrov DA (2010) Genome-wide patterns of adaptation to temperate environments associated with transposable elements in Drosophila. PLOS Genetics, 6, e1000905. https://doi.org/10.1371/journal.pgen.1000905

[4] Lisch D (2013) How important are transposons for plant evolution? Nature Reviews Genetics, 14, 49–61. https://doi.org/10.1038/nrg3374

[5] Bonchev G, Parisod C (2013) Transposable elements and microevolutionary changes in natural populations. Molecular Ecology Resources, 13, 765–775. https://doi.org/10.1111/1755-0998.12133

[6] Casacuberta E, González J (2013) The impact of transposable elements in environmental adaptation. Molecular Ecology, 22, 1503–1517. https://doi.org/10.1111/mec.12170

[7] Vargas-Chavez C, Pendy NML, Nsango SE, Aguilera L, Ayala D, González J (2021). Uncovering transposable element variants and their potential adaptive impact in urban populations of the malaria vector Anopheles coluzzii. bioRxiv, 2020.11.22.393231, ver. 3 peer-reviewed and recommended by Peer community in Genomics. https://doi.org/10.1101/2020.11.22.393231


# Reviews

## Evaluation round #1


DOI or URL of the preprint: **10.1101/2020.11.22.393231**


### Authors' reply, 02 April 2021

**Download author's reply**
**Download tracked changes file**

## Decision by Anne Roulin, posted 27 January 2021

**Decision on your manuscript "Impact of transposable elements on the genome of the urban malaria vector Anopheles coluzzii" - revision required**

Dear Dr. Gonzalez,

Thank you for submitting your work entitled "Impact of transposable elements on the genome of the urban malaria vector Anopheles coluzzii" for consideration by PCI Genomics. Your article has been reviewed by 2 peer reviewers.

Based on these 2 reviews, I regret to inform you that your work can not be accepted for recommendation in PCI genomics as such. I would like nonetheless to invite you to revise your manuscript based on the reviewers comments.

While both reviewers acknowledge that the data produced are extremely valuable and open new avenues of research to study the adaptation of An. coluzzii, they also fear that the study does not demonstrate the role of TEs in this process per se and therefore remains descriptive. I understand that a functional validation is beyond the scope of the study but Reviewer1 suggested some analyses to include in order to sustain the claims you made. They have also made additional comments which I hope will help you to improve the manuscript.

Looking forward to reading a revised version of your manuscript

Best wishes,

Anne Roulin

## Reviewed by anonymous reviewer 1, 20 January 2021

**Download the review**

## Reviewed by Yann Bourgeois, 11 January 2021

This paper constitutes a valuable resource to understand the genomic bases of local adaptation to the urban environment in Anopheles mosquitoes. It represents a substantial sequencing and analytical effort. I also appreciated the care brought to provide a detailed functional annotation of TEs regulatory and promoter sequences. My main comments are as follows:

1) The authors acknowledge and discuss this, but it remains difficult to draw strong conclusions about the role of TEs in adaptation due to the rather low sample size. However, the discussion on potential mechanisms is rather long. It may be seen as hand-waving, and I worry that it is edging toward story-telling as defined in (Pavlidis, Jensen, Stephan, & Stamatakis, 2012). It may be worth toning down these parts.

2) However, a possible way to provide more support for a role of TEs in adaptation may be to examine patterns of SNP variation at their flanking genomic sides. This may require substantial analytical time, and I would understand it if the authors declined to go forward with this suggestion. A possible way would be to run a software such as ARGWeaver (Hubisz, Williams, & Siepel, 2020; Rasmussen, Hubisz, Gronau, & Siepel, 2014) to extract times since coalescence (TMRCA) along the genome, and test how these times differ at the vicinity of insertions that may be candidate for local adaptation. With six genomes, the algorithm should run rather fast, and this would bring a lot of information to at least check whether TMRCAs drop drastically near a candidate insertion.

3) In relation to point 2), the authors do not genotype insertions nor SNPs but look at presence/absence and reconstruct a haploid genome assembly if my understanding is correct. It might be useful to call variants and TEs (maybe with MELT? (Gardner et al., 2017)) to examine, for example, variation in the allele frequency spectrum of TEs in relation with features such as recombination. Given the potential importance of background/linked selection in shaping the average frequency of TEs, examining correlations between TE density/frequency with recombination might be relevant. This would help discussing whether indirect effects of selection may impact the identification of TEs under positive selection.

I also have a few minor comments listed below.

L100: citation not in the same format than other citations

L194: It may be worth explaining what sort of threshold is used to group TEs in a family (check Methods section)

L165: I am a bit lost here, sorry. Should not there be an overlap between these 85 and the 435 families identified previously?

L182: Can you perhaps examine whether these TRIM elements have the said impact on Anophele genomes, since you have long reads?

L204: What about other features such as recombination rate?

L208: I understand why, but it may be worth explaining that you want to focus on well-assembled regions, where TEs may have a functional impact. Especially because you just said that most of the variation is in heterochromatin, so why looking at euchromatin?

L209: p-value -> p in italics

L281: I feel like this whole part could come earlier. For example the sentence L285 makes it clear that LTRs have been recently active and this echoes their large variance in copy number shown L210.

L302: This is where I thought that it would be nice to have the frequency spectrum/heterozygosity for these TEs rather than presence-absence data.

L548: Maybe explain whether these are the default parameters, or why you chose these specific ones.

L692: A/B scores > support? Maybe provide more details.

Figure 1: It would be interesting to see how abundant each new family is when adding new genomes. In other words, are we mostly adding families with a single element? Maybe provide a graph with the total length or number of TEs, in relation to the number of genomes added, irrespective of family?

Figure 2: I do not find this representation very clear nor informative. In panel A, maybe provide the length or a scale for each element? Panel B is hard to interpret. Panel C could benefit from a phylogenetic representation of relationships between species. In general, the figures I had were too compressed, hard to decipher.

Thank you for your work, I hope you find the comments constructive.

All the best,

References

Gardner, E. J., Lam, V. K., Harris, D. N., Chuang, N. T., Scott, E. C., Pittard, W. S., … Devine, S. E. (2017). The Mobile Element Locator Tool (MELT): Population-scale mobile element discovery and biology. Genome Research, (410), gr.218032.116. doi: 10.1101/gr.218032.116

Hubisz, M. J., Williams, A. L., & Siepel, A. (2020). Mapping gene flow between ancient hominins through demography-aware inference of the ancestral recombination graph. PLoS Genetics, 16(8), 1–24. doi: 10.1371/JOURNAL.PGEN.1008895

Pavlidis, P., Jensen, J. D., Stephan, W., & Stamatakis, A. (2012). A critical assessment of storytelling: Gene ontology categories and the importance of validating genomic scans. Molecular Biology and Evolution, 29(10), 3237–3248. doi: 10.1093/molbev/mss136

Rasmussen, M. D., Hubisz, M. J., Gronau, I., & Siepel, A. (2014). Genome-Wide Inference of Ancestral Recombination Graphs. PLoS Genetics, 10(5). doi: 10.1371/journal.pgen.1004342