Peer Community In Evolutionary Biology

DNA transposons drive genome evolution of the root-knot nematode Meloidogyne incognita

Ines Alvarez *based on reviews by Daniel Vitales and 2 anonymous reviewers*

A recommendation of:

Open Access

Submitted: 04 May 2020, Recommended: 04 August 2020

Published: 5 August 2020

Duplications, mutations and recombination may be considered the main sources of genomic variation and evolution. In addition, sexual recombination is essential in purging deleterious mutations and allowing advantageous allelic combinations to occur (Glémin et al. 2019). However, in parthenogenetic asexual organisms, variation cannot be explained by sexual recombination, and other mechanisms must account for it. Although it is known that transposable elements (TE) may influence on genome structure and gene expression patterns, their role as a primary source of genomic variation and rapid adaptability has received less attention. An important role of TE on adaptive genome evolution has been documented for fungal phytopathogens (Faino et al. 2016), suggesting that TE activity might explain the evolutionary dynamics of this type of organisms.
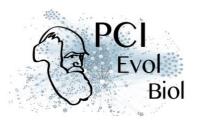
The phytopathogen nematode *Meloidogyne incognita* is one of the worst agricultural pests in warm climates (Savary et al. 2019). This species, as well as other root-knot nematodes (RKN), shows a wide geographical distribution range infecting diverse groups of plants. Although allopolyploidy may have played an important role on the wide adaptation of this phytopathogen, it may not explain by itself the rapid changes required to overcome plant resistance in a few generations. Paradoxically, *M. incognita* reproduces asexually via mitotic parthenogenesis (Trudgill and Blok 2001; Castagnone-Sereno and Danchin 2014) and only few single nucleotide variations were identified between different host races isolates (Koutsovoulos et al. 2020). Therefore, this is an interesting model to explore other sources of genomic variation such the TE activity and its role on the success and adaptability of this phytopathogen.

To address these questions, Kozlowski et al. (2020) estimated the TE mobility across 12 geographical isolates that presented phenotypic variations in *Meloidogyne incognita*, concluding that recent activity of TE in both genic and regulatory regions might have given rise to relevant functional differences between genomes. This was the first estimation of TE activity as a mechanism probably involved in genome plasticity of this root-knot nematode. This study also shed light on evolutionary mechanisms of asexual organisms with an allopolyploid origin. These authors re-annotated the 185 Mb triploid genome of *M. incognita* for TE content analysis using stringent filters (Kozlowski 2020a), and estimated activity by their distribution using a population genomics approach including isolates from different crops and locations. Canonical TE represented around 4.7% of the *M. incognita* genome of which mostly correspond to TIR (Terminal Inverted Repeats) and MITEs (Miniature Inverted repeat Transposable Elements) followed by Maverick DNA transposons and LTR (Long Terminal Repeats) retrotransposons. The result that most TE found were represented by DNA transposons is similar to the previous studies with the nematode species model *Caenorhabditis elegans* (Bessereau 2006; Kozlowski 2020b) and other nematodes as well. Canonical TE annotations were highly similar to their consensus sequences containing transposition machinery when TE are autonomous, whereas no genes involved in transposition were found in non-autonomous ones. These findings suggest recent activity of TE in the *M. incognita* genome. Other relevant result was the significant variation in TE presence frequencies found in more than 3,500 loci across isolates, following a bimodal distribution within isolates. However, variation in TE frequencies was low to moderate between isolates recapitulating the phylogenetic signal of isolates DNA sequences polymorphisms. A detailed analysis of TE frequencies across isolates allowed identifying polymorphic TE loci, some of which might be neo-insertions mostly of TIRs and MITEs (Kozlowski 2020c). Interestingly, the two thirds of the fixed neo-insertions were located in coding regions or in regulatory regions impacting expression of specific genes in *M. incognita*. Future research on proteomics is needed to evaluate the functional impact that these insertions have on adaptive evolution in *M. incognita*. In this line, this pioneer research of Kozlowski et al. (2020) is a first step that is also relevant to remark the role that allopolyploidy and reproduction have had on shaping nematode genomes.

References

[1] Bessereau J-L. 2006. Transposons in C. elegans. WormBook. https://doi.org/10.1895/wormbook.1.70.1
[2] Castagnone-Sereno P, Danchin EGJ. 2014. Parasitic success without sex - the nematode experience. J. Evol. Biol. 27:1323-1333. https://doi.org/10.1111/jeb.12337
[3] Faino L, Seidl MF, Shi-Kunne X, Pauper M, Berg GCM van den, Wittenberg AHJ, Thomma BPHJ. 2016. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. Genome Res. 26:1091-1100. https://doi.org/10.1101/gr.204974.116
[4] Glémin S, François CM, Galtier N. 2019. Genome Evolution in Outcrossing vs. Selfing vs. Asexual Species. In: Anisimova M, editor. Evolutionary Genomics: Statistical and Computational Methods. Methods in Molecular Biology. New York, NY: Springer. p. 331-369. https://doi.org/10.1007/978-1-4939-9074-0_11
[5] Koutsovoulos GD, Marques E, Arguel M-J, Duret L, Machado ACZ, Carneiro RMDG, Kozlowski DK, Bailly-Bechet M, Castagnone-Sereno P, Albuquerque EVS, et al. 2020. Population genomics supports clonal reproduction and multiple independent gains and losses of parasitic abilities in the most devastating nematode pest. Evol. Appl. 13:442-457. https://doi.org/10.1111/eva.12881
[6] Kozlowski D. 2020a. Transposable Elements prediction and annotation in the M. incognita genome. Portail Data INRAE. https://doi.org/10.15454/EPTDOS
[7] Kozlowski D. 2020b. Transposable Elements prediction and annotation in the C. elegans genome. Portail Data INRAE. https://doi.org/10.15454/LQCIW0
[8] Kozlowski D. 2020c. TE polymorphisms detection and analysis with PopoolationTE2. Portail Data INRAE. https://doi.org/10.15454/EWJCT8
[9] Kozlowski DK, Hassanaly-Goulamhoussen R, Da Rocha M, Koutsovoulos GD, Bailly-Bechet M, Danchin EG (2020) Transposable Elements are an evolutionary force shaping genomic plasticity in the parthenogenetic root-knot nematode Meloidogyne incognita. bioRxiv, 2020.04.30.069948, ver. 4 peer-reviewed and recommended by PCI Evolutionary Biology. https://doi.org/10.1101/2020.04.30.069948
[10] Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. 2019. The global burden of pathogens and pests on major food crops. Nat. Ecol. Evol. 3:430-439. https://doi.org/10.1038/s41559-018-

0793-y

[11] Trudgill DL, Blok VC. 2001. Apomictic, polyphagous root-knot nematodes: exceptionally successful and damaging biotrophic root pathogens. Annu Rev Phytopathol 39:53-77. https://doi.org/10.1146/annurev.phyto.39.1.53

Revision round #2

2020-08-03
Thanks to the authors.

I'm satisfied with their modifications and just ask them to make the minor modifications of reviewer 3 and to include the link to their dataverse website (https://data.inrae.fr/dataverse/TE-mobility-in-MiV3) in the main manuscript.

A message will be sent to the authors with the formatting instructions.

Best,

Inez Alvarez

Preprint DOI: 10.1101/2020.04.30.069948

Reviewed by anonymous reviewer, 2020-07-13 12:21

I thank the authors for adressing all the raised points. The revised manuscript is ok for me.

Reviewed by anonymous reviewer, 2020-08-02 19:23

The authors have thoroughly revised their manuscript and have done an excellent job of addressing all the reviewer concerns. The authors have also added new data on HGT likelihood of various TE associated loci, which is very welcome. Both the introduction and discussion sections are substantially enhanced and provide a comprehensive review of the research field. I strongly support that the manuscript be officially recommended on PCI-EvolBio.

One minor point: please include the link to their dataverse website (https://data.inrae.fr/dataverse/TE-mobility-in-MiV3) in the main manuscript, currently I could only see it in the response letter from the authors.

Congratulations on a rigorous piece of work.

Reviewed by Daniel Vitales, 2020-07-17 17:37

I read with great interest this new version of the manuscript prepared by Kozlowski and collaborators. In my opinion, the study looks now clearer and more appealing after the revision. I would like to thank the authors for being so careful addressing my previous comments as well as those form the other reviewers. I think the manuscript could be recommendable at its current state, but I would also like to add some further suggestions that the authors might consider to incorporate in the final paper.

In this new version of the manuscript, the authors provide a table summarising the whole (i.e. unfiltered) TE characterization of M. incognita (Table S1). This is a very interesting information allowing the comparison among whole repeatome annotation and the TE landscape represented by canonical (filtered) TEs. To me, the most unexpected result from this comparison is related to the abundance of Maverick elements, being by far the most abundant elements according to the whole repeatome characterization but showing much lower abundances when canonical elements are considered. I missed an explanation to these somewhat contradictory outputs and perhaps a discussion on how this could bias (or not) other results obtained.

The Table S5 is another very informative addition (I don't remember having seen it in the previous version) to the manuscript. There, we observe that LTRs, LINEs and Maverick are those elements showing a larger number of substantially expressed putative transposition machinery genes (i.e. suggesting TE activity). Conversely, according to TE-polymorphisms analyses (i.e the number of neoinsertions), MITEs and TIRs are reported as elements that "might have been more active in the genome of M. incognita than elements from other orders". To me, both results seem somewhat contradictory, so perhaps the authors could try to explain this better too.

Finally, the criteria employed to select the 5 HCPTE loci validated by PCR were clearly explained in the point-by-point response letter to the reviewers. In my opinion, this explanation should also be included within the manuscript for a better interpretation of the results obtained. Without this information, the readers could understand that the validated "locus 1" being inserted in an expressed Meloidogyne-specific gene is a signal of adaptation that could be extrapolated to the rest of HCPTEs. However, this locus 1 was specifically selected among the 22 HCPTEs for having those precise characteristics (e.g. being expressed in Morelos transcriptome and being Meloidogyne-specific). To me, explaining these criteria would make reading easier.

Other minor points:

L39. As a taxonomic name of a kingdom, I think that "Metazoa" should be written in capital letters.

L100, L105. "Arabidopsis" and "Drosophila" should be written in italics.

L168. The taxonomic name of the phylum should be "Nematoda".

L188. Regarding the differences of TE-content estimations among filtered and unfiltered approaches, here I would specify that "..almost two-thirds of the M. incognita canonical TE content".

L235. I am not sure whether here should be cited Table S2 or Table S4.

L311. The isolate from Morelos is named as "morelos" and Morelos" indistinctly along the manuscript. I think this should be unified.

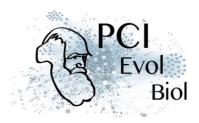L489. I think that "22 out of 33" should be better described as "majority" rather than "vast majority".

Author's reply:

[Download author's reply (PDF file)](#)

Revision round #1

2020-06-02

This is a valuable piece of work, very relevant in knot-root nematode M. incognita genome evolution. It is interesting for the community and I think it could be recommendable after addressing reviewers main concerns, other minor questions, and incorporating their suggestions. Three reviewers coincided in the good quality and relevance of the study and therefore, it has a positive feedback, but I also agree that there are several points that should be clarified before its recommendation.

Major concerns are related with the recapitulation of the actual knowledge on TEs in nematodes in order to highlight the relevance of the present study, as well as other study cases in which TEs activity directly affect regulatory and coding regions. Other major concern is that TEs activity drives adaptive evolution seems not conclusive here. This is not demonstrated with the data presented here and it should be noticed. There are several paragraphs dificult to follow and understand (see reviewers anotations). Please, respond to each question of all reviewers and make changes in the text accordingly in order to produce a recomendable new version of your manuscript.

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."
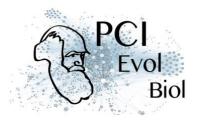
Preprint DOI: [10.1101/2020.04.30.069948](10.1101/2020.04.30.069948)

Reviewed by anonymous reviewer, 2020-05-31 05:34

In the current manuscript, the authors have surveyed the variation in the presence and frequencies of various transposable elements (TEs) in population isolates of the root-knot nematode Meloidogyne incognita. The authors have performed a comprehensive and careful analysis and reported the results quite clearly. The manuscript may be published as it is. Some minor comments and suggestions are provided below if the authors wish to include them in their analysis or discussions.

Strengths of the manuscript:

1. The root-knot nematode M. incognita is a major pest of agricultural plants. The flexibility of this pest to adapt to various plant host across wide geographical areas, despite being a clonal allopolyploid species without sexual recombination is particularly intriguing. Studies of its genome evolution might provide insights into its biology and potential ways of combating it.

2. Population genetic analysis of sequence evolution is a powerful approach to find genomic regions associated with a given phenotype, with many methods focusing on SNPs. However, the variation in TEs is hard to study from a technical as well as a theoretical perspective. Therefore studies into documenting TE variations are welcome as they may spur the development of new, appropriate methods.

3. The authors have used state-of-the-art computational methods for analyzing their data and used stringent quality filters. The methods are described in sufficient detail and most likely the scripts will be provided as supplementary materials.

4. The authors have experimentally validated some of the TE insertions predicted from their computational analysis.

Open questions / Suggestions:

1. Although authors include "adaptability" in their title, and one of their hypotheses is that TE activity might generate some adaptive variation in M. incognita genomes, very few examples of direct effects on protein-coding genes were observed. The authors have discussed the multiple reasons that could explain this (stringent filters, un-annotated genome, other adaptive effects e.g. recombination). (i) If the authors could discuss results from similar studies in other organisms with respect to the number of protein-coding and regulatory changes caused by TE, it will be a valuable information.

2. Since TEs can transfer genetic material via horizontal gene transfer, the authors might want to discuss this aspect as a potential contributor to adaptive functions of TEs. This could ve particularly interesting as the authors do observe some examples of TE insertions in Meloidogyne-specific genes. Could these species- or genus-specific genes arise from HGT via TE insertions?

3. Some TEs, most famously the P-elements in Drosophila melanogaster and Dropshila simulans have been observed to arise and spread in wild populations incredibly fast (e.g. between 1950s and 1990s). The authors might want to consider a more recent spread of TEs in M. incognita lineage as a potential reason why not many adaptive examples are observed. It could also be informative to analyze if fast invasions of TEs can be diagnosed by some genomic signatures e.g. patterns of TE diversity and genomic hot-spots.

4. Typical lengths of various TE loci in each order : This would be a piece of useful information that can be included in a small table (main text or supplementary)

5. The authors have used %identity of TE loci with corresponding consensus sequences as a key metric. Is it possible to also provide multiple sequence alignments of at least some representative loci within each order, demonstrating various patterns of variation?
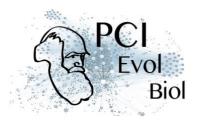
Note: Since the supplementary material and files were not available with the manuscript, they could not be reviewed.

Reviewed by anonymous reviewer, 2020-05-28 14:53

Review for PCI EVOL BIOL of the manuscript entitled "Transposable Elements activity and roles in Meloidogyne incognita genome dynamics and adaptability".
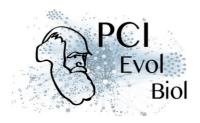
General comments: In this manuscript, the authors retrace the dynamics of transposable elements (TE) in the genome of the root-knot nematode Meloidogyne incognita, a plant pest that reproduces asexually via mitotic parthenogenesis. The authors re-annotated the latest version of the M. incognita genome for TE and took advantage of population genomics data of a dozen of geographical isolates to study TE polymorphisms as a reporter of TE activity. They showed that TE in M. incognita are mostly DNA transposons and that thousands of TE present very highly contrasted frequencies among isolates, suggesting their transposition activity. Very interestingly, few dozen correspond to neo-insertions (some being experimentally validated) that could possibly impact protein coding genes or their regulation. These results provide evidence that TE could play a significant role in the genome plasticity and adaptive evolution of Meloidogyne incognita and will be of interest for the community. There are however some points that should be addressed before.

Major points - In the introduction, the authors should recapitulate the actual knowledge about TE in nematodes and in C. elegans, in particular. This would further highlight the importance of the present study.

- lines 181-192. This paragraph is hard to understand and more quantitative details should be given to help the reader. The authors should better explain why they consider that HELITRON and MAVERICK elements do not share a high identity level with their consensus and they should better substantiate their argument that SINE and CLASS2LIKE distributions are similar. Statistical comparisons would clear the picture. o It is not clear to me why annotations sharing the highest similarity with their consensus are found among DNA-transposon elements. o The authors should explain what they mean by "sufficient evidence" to consider class1like as retros and class2like as DNA-transposons. There is no conclusion about class1like and class2like elements at the end of the section and they are no longer mentioned in the rest of the text. Why are they important? o Is there a bias in the coverage of the consensus sequence length between retrotransposons and DNA-transposons, and between the different groups therein, that could partly explain the differences of the distributions of the per-copy identity percentages? If yes, the per-copy identity would not directly reflect a biological signal but would be determined by the sequence length. o According to Table S2, the TIR annotations, but also the CLASS2LIKE and SINE ones share above 99% identity with their consensus. What is the rationale for highlighting only TIR annotations on the main text?

- Figure 4. Why is the Morelos isolate absent from the ML tree in Fig 4A? The authors say that clade 2 is identical in both trees, including branching, but clade 2 contains only two isolates unless I missed something.

- 4 consensuses are involved in 24 out of 33 HCPTE. Is there a bias among MITEs, TIRs and LINEs?

- The 5 HCPTE for experimental validation all impact Meloidogyne-specific genes, but others HCPTE also share this property according to Table S4. What is the rationale for selecting those 5?

- It would be worth mentioning whether or not such highly contrasted polymorphism in TE has already been observed in C. elegans, and to briefly remind the reader of what (if any) has been performed to estimate the impact of TE in C. elegans.

- In the discussion, when referring to their results, authors should refer to the corresponding figures and Tables. This is especially true for arguments on lines 730 to 736. Based on Figure 2, it is not obvious that the behavior of MITEs and TIRs are different in function of the identity rate with their consensuses. Also, the comment that "TIR neo-insertions are less numerous than expected owing to their abundance in the genome" is supported by Fig S4 but Fig S4 is never cited in the text.

- Methods: Diagram recapitulating all the annotation steps would greatly help the reader. A decision tree would also help for the polymorphism characterization. The bootstrap approach for the TE-frequency NJ tree should be explained, at least briefly.

- Data : The authors should also give the project accession number of the M. incognita reference genome they used from Blanc-Mathieu et al. 2017. Minor points:

- A definition of autonomous and non-autonomous TE would be nice and the "autonomous/non-autonomous" status for each order of transposons in the M. incognita annotations could be given in Table 1.

- Table 1. What is the median of median identity with consensus (%)? Tipo?
- Figure 3A. It is very hard to see something at that small size.
- Figure 5. A title (the type of TE locus described) for each panel would help. The number of loci used in each panel could also be given in the panel, not only in the legend.
- Table S4 title: correct "ortologs"

Reviewed by Daniel Vitales, 2020-05-22 18:10

The manuscript "Transposable Elements activity and role in Meloidogyne incognita genome dynamics and adaptability" by Kozlowski et al. presents a thorough study of TE content, frequency variability and activity among several isolates of M. incognita, a parasitic root-knot nematode showing significant impact as an agricultural pest. The whole genome sequencing data used by the authors, which was already analysed in a population genomics study (Koutsovoulos et al. 2020), is here nicely re-analysed to characterize the repeatome of this species, to identify potentially active TE loci and compare their frequencies among some M.incognita isolates with different hosts and geographic origins. Additionally, they study the impact that transposition of some elements could have in some transcriptionally active genes. I would like to say I really enjoyed reading this work, which allowed me to know new methodologies and points of view in the study of repetitive elements evolution.

First of all, I must make clear that I have never used neither the general approaches nor the specific pipelines here employed by the authors to analyse the repeatome. Consequently, although the procedures seem correct to me, I cannot (and will not) evaluate the details of the methodology. In this case, my review will focus on general points related to the hypotheses tested, the results obtained and some of the interpretations carried out. I hope my comments will be helpful for the authors:

1) As stated in the Introduction section, the main goal of this study is to test "whether the TE activity could represent a mechanism supporting genome plasticity and eventually adaptive evolution in M. incognita". The authors do provide solid evidences to prove the hypothesis that TE played an important role in the intraspecific genomic diversity of M. incognita. They also clearly show the potential impact of some TE in the activity of certain genes. However, while reading the Introduction and the Discussion, my impression was that the most important goal of the manuscript was on testing the adaptive role of TE in the evolution of the species. The first section of the Discussion is indeed devoted to expose potential effects of TE activity in adaptive evolution or the impact on the function of certain genes in different species. Certainly, some genic regions appear to be impacted by TEs in M. incognita and they were confirmed to be expressed ¬— according to the transcriptome data they got – but I am not sure that the experimental design enables to test here the adaptive role of TE activity. As the authors comment in the first Discussion section, "functional impact itself would need to be evaluated in the future (L591-L592)" and "no evident role in adaptive evolution for the M. incognita genes impacted by TE insertions could be reported so far (L611-L613)". In my opinion the data and the experimental design here employed works very well to characterize and compare the repeatome content among different isolates of the species, whereas they do not allow to test whether TE activity drives adaptive evolution. I think the message of the manuscript would be stronger if the authors focus the objectives (mainly reformulating the Introduction and the Discussion) on the intraspecific dynamism of TEs in M. incognita, where they could show clear results and obtain interesting conclusions.

2) At first glance, I was happy to see that the second section of the Discussion was devoted to the "TE-load and composition" of M. incognita. Certainly, the analyses performed by the authors provides remarkable information on the TE content of the species, as well as its intraspecific variability. However, I just found five lines at the end of this section (L699-L703) commenting the finding that DNA transposons are the most abundant element of M. incognita repeatome, while the rest of the section mainly presents relatively disconnected study cases where the TE content was affected (or not) by hybridization, polyploidy or asexual reproduction. Probably I was expecting that the authors compare the repeatome characterization of M. incognita isolates they obtained with the repeatome of other Meloidogyne species. For this purpose, in case they need a different (broader) perspective of the repeatome landscape of this species (and other they could study) I would suggest to analyse their genomic data with other de-novo approaches (e.g. RepeatExplorer).

3) Finally, regarding the structure of the Discussion, I found that the third section entitled "TE show signs of recent activity in M. incognita and they might still be active" presents some of the most solid and remarkable results of the paper. I would recommend to start the Discussion with the results commented in this section, then going on with the section devoted to explain the impact of TE in the gene(s).

Some other points:

L55. This could be a good place to explain for the first time that M. incognita is a triploid, parthenogenic species with a hybrid origin.

L79. I think the abbreviation of "id est" should be (i.e.)

L99. The expression "'novelty' / plasticity" could be changed by "genomic novelty or plasticity"

L105. The "arms race" was written between quotation marks three line before.

L115-L117. To me, this statement sounds as a Result or as a Discussion.

L137. The abbreviation "cf." is employed throughout the text to point the methods as a source of information. I think the word ("see XXX") would be more appropriate.

L146-L147. There is stated that "Retro-transposons and DNA-transposons respectively cover 0.94 and 3.78 % of the genome". I understand these values correspond to the retrotransposons and DNA transposons within the canonical TE (not the whole repeatome, which could include "non-canonical TE). If I am right, this clarification should be specified here. In addition, I think it would be interesting to know whether the composition of canonical TEs is comparable to the composition of the whole repeatome. As commented above, the characterization of the whole repeatome of M. incognita using a de novo approach such as RepeatExplorer could be easy and very informative.

L147-L148. Being the first time the acronyms TIR and MITEs are used, I would suggest to mention them as "Terminal Inverted Repeats (TIR) and Miniature Inverted repeat Transposable Elements (MITEs)..."

L151-L164. I must recognize I got difficulties to understand some of the information in this paragraph. I got that REPET pipeline estimated a repetitive content in C. elegans genome very similar to that obtained by Bessereau 2006). Were the predictions of 1.8% and 0.2% of MITEs and LTR obtained with or without the filtering protocol? Regarding Bessereau estimations, do they correspond to canonical (i.e. "potentially active" TEs) or to fossil + active TEs? To be comparable, both predicitions should consider the same type (canonical or all kind) of TEs. The authors might consider to rephrase the paragraph for a better understanding.

L158. Supplementary materials indicate that MITEs compose 0.7% of C. elegans genomes, while here it is mentioned "1.8%". This should be revised. I would also recommend to construct a table where one could easily compare TE estimations among Bessereau 2006 and your approach.

L174. Please cite a reference for this genome size value.

L184. Should "HELITRON" and "MAVERICK" be written in capital letters?

L183. I was not able to find A and B parts in Figure 2.

L206-L208. I was not able to explain why only 6.26% of canonical TEs (i.e. complete elements and suposedly being potentially active) contain a protein coding gene. Is this value something expected? What those canonical elements lacking transposition genes correspond to? Could this percentage be compared to the "autonomous:non-autonomous ratio" of elements found in the repetome of M. incognita?

L211. I was not able to find this supplementary information on the proportion of LTRs, LINES, TIRs, etc among the 111 canonical TE with transposition genes.

L240. If I understood well, this corresponds only to the 3,524 variable TE loci. Perhaps this could be better explained.

L238-L250. I would suggest to use "within" for intra-isolate variability and "between" for inter-isolate variables. Or another nomenclature but using it consistently throughout the text. In my opinion, this will make easier for the reader to understand the values/indexes you are talking about.

L264-L265. It would be nice to state how many "loci frequencies" values were employed to construct the distance matrix.

L276-L277. Here, I would also take into account that the branches leading to R2-1 and R1-6 do not show high support values in TE-based tree. This low resolution obtained in that parts of the TE-based trees - which could be caused by several factors (e.g. hybridisation; polyploidy) - should probably be mentioned in the results.

L278-L279. I was delighted to see the beautiful phylogenetic signal the authors obtained from their repeatomic data. However, I would rather expose that "TE-frequencies between isolates contain a valid phylogenetic signal". I would have also enjoyed some comments in the Discussion section about these results. For instance, the authors could consider potential biases caused by the loci selection they performed. This tree was based on the 3524 significantly variable loci, which is only a part from the 9000 canonical TE loci.

L302. As this is important to understand the text below, I would suggest to include the name of the different categories beside the codes A, B, C and D of Fig. 5.

L331-L355. I am not sure to be properly understanding the categorization procedure for polymorphic TEs. Could the "truncated or diverged versions of TE" be also present in the other isolates (excluding Morelos)? In other words, I was not able to see how authors can be sure that "neo-insertions" do not in fact correspond to "extra-detections" occurring in the rest of isolates (but not in Morelos). Perhaps this part could be better explained.

L367. Could Figure 6 include the information of non-polymorphic loci?

L452-L453. The authors might consider to compare the positions of "neo-insertions" with the position of other "polymorphic TE" categories. Is the proportion of elements inside a gene or a regulatory region the same in other "polymorphic TE" categories?

L470. I was not able to find why specifically these 5 HCPTE were studied but not the rest of them.

L481. I was not able to find the Supplementary material 4 the authors mention here. Is this the table S4?

L482. I missed some more details about the PCR validation results on the rest of HCPTE.

L590. According to the transcriptome data, some of these regions were certainly confirmed to be expressed, but can you really state they are functionally important? As the authors comment in the following lines, "functional impact itself would need to be evaluated in the future".

L722-L723. As the authors admit in L737-747, there should be additional data and stronger evidences to confirm a TE burst during the evolution of the species.

L751-L752. I am not sure if the authors can conclude this statement without considering other alternatives. Could these "neo-insertions" be present in the original pre-agriculture genomic pool and later on being fixed or erased from some isolates?

Author's reply:

[Download author's reply (PDF file)](#)