nature portfolio

Peer Review File

The Landscape and Evolutionary Dynamics of Derived lncRNAs in Insects

Corresponding Author: Dr Kang He

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors performed a comprehensive evolutionary and functional analysis of lncRNAs across 115 insect species, utilizing publicly available genomic data from the Insectbase2.0 database and NCBI. Their findings demonstrate that lncRNA abundance is positively correlated with both genome size and transposon content. Notably, lncRNAs derived from protein-coding genes exhibit distinct characteristics compared to conventional lncRNAs, including differences in expression levels, genomic distribution, and regulatory complexity. To further validate these findings, the study highlights Msex-lnc001918 from Manduca sexta as a case study, supporting the hypothesis that the functional roles of protein-coding gene-derived lncRNAs are closely linked to the biological functions of their ancestral genes.

In summary, this large-scale comparative analysis provides systematic insights into the genomic features of lncRNAs in insects. It elucidates the molecular mechanisms by which protein-coding genes give rise to novel regulatory elements through sequence transformation, offering a valuable theoretical framework for understanding the evolution and functional diversification of lncRNAs. However, the present manuscript requires revisions to address the following issues prior to publication.

1. The authors used known RNA-seq data from 115 insect species for comparative analysis of lncRNAs across different insects. Given the inconsistencies in the quality or quantity of genomic data and RNA-seq data among many species, how did the authors ensure the reliability of the analysis results and conclusions based on these data?
2. The authors proposed that Msex-lnc-001918 may regulate spermatogenesis based on bioinformatic analysis, but this hypothesis lacks experimental validation. Functional studies using RNA interference (RNAi) or CRISPR/Cas9-mediated knockout of Msex-lnc-001918 and its host gene would be necessary to confirm their roles in spermatogenesis.
3. In L53-54, references 14 and 16 are inappropriately placed at the end of the sentence. They should be moved after "...in vertebrates," as these two references are unrelated to lncRNAs in insects.
4. L247, please change "supplementally table2" to "Supplementary Table 2". In the Manuscript Items of the submission system, there are two entries labeled "Supplementary Table 1", with no "Supplementary Table 2". The second "Supplementary Table 1" should be "Supplementary Table 2".
5. L93-95, "The substantial variation (~36-fold) in lncRNA numbers across species with Lepidoptera showing the highest average number, followed by Hymenoptera and Diptera, …." This statement lacks support from any results or literature citation. It is recommended to cite relevant results or references.
6. The figure legend for Figure 1C contains an inconsistency regarding the x-axis label. While the main text clearly indicates that the x-axis represents genome size (Lines 111-113), the figure legend incorrectly describes it as both transcriptome size and sample number (Lines 650-652). This discrepancy should be corrected for accuracy.
7. L127-136, the results are only described without citing the corresponding figures or tables. Supporting data needs to be provided.
8. L282-283: "Increasing numbers of functional studies on lncRNAs are revealing the importance, however, limited data are available for insects." Reference 34 does not support this conclusion.
9. L297-299, a reference citation is needed after "…findings in mammals".
10. The punctuation marks in sentences L131 and L163 need to be corrected.

Reviewer #2

(Remarks to the Author)
This study provides a comprehensive analysis of drived lncRNAs from evolutionary and functional perspectives across 115 insect species. These derived lncRNAs show distinct features including lower length, lower transposable element fraction, higher expression levels, broader tissue distribution, and increased regulatory complexity compared to conventional lncRNAs. Functional analysis has revealed that Msex-lnc-001918 has higher expression levels, more transposable elements and similar biological functions compared to its syntenic genes. In summary, this study provides fundamental insights into insect derived lncRNAs and demonstrates how protein-coding gene to lncRNA transformation creates novel regulatory innovations in insect genomes.

Besides these, there are some limitations in the study.

1 The authors identified massive multi-exonic lncRNAs. I'm particularly curious about the characteristics of single-exonic lncRNAs in insects.

2 The authors suggest that 500 diversified transcriptomes represent the minimum threshold for comprehensively identifying lncRNAs. This claim is questionable because discussing sample size independently of data size is methodologically unsound.

3 Given the well-documented pseudogene transcripts in genomes, the authors are essentially studying pseudogene transcripts through homologous protein screening for lncRNAs. The manuscript would benefit from incorporating pseudogene-related perspectives into the Introduction and Discussion sections.

4 Evolutionary conservation is critical for studying gene function. I recommend conducting a detailed comparative visualization between protein-coding genes and lncRNAs based on homologous families in results subsection "Evolutionary analysis of lncRNA orthologous families". For instance, a plot with the X-axis representing cross-species conservation levels of homologous families and the Y-axis showing the percentage representation of conserved families would visually demonstrate their conservation divergence.

5 Figure 1 legend needs revision carefully.

6 The text needs to be proofread and revised. Some of the examples are listed as follows (please note that this is not a thorough list of all sentences and phrases to be revised):

Line 95-100: This sentence suffers from logical flaws. First, there is no direct link between the amount of transcriptome data and the quantity of lncRNAs, nor does it address the uneven distribution of lncRNAs across insect species. Second, direct numerical comparisons are problematic – lncRNA counts should be normalized against factors like genome size and protein-coding gene numbers.

Line 106-111: The full scientific name should be used upon first mention of a species: "A. aegypti" should be "Anopheles aegypti" for the first time.

Line 131: Remove redundant punctuation marks.

Line 156-157: Please rephrase this sentence. For instance: "X and Y represent flanking genes of syntenic genes and lncRNA genes (located within 100 kb), respectively, with confirmed orthology."

Line 180: "(Figure3 B)" should be "(Figure 3B)".

Line 181-182: The authors conclude that derived lncRNAs are shorter than their homologous protein-coding genes, yet omit comparative analysis of structural features. This gap particularly interests me, as our previous study demonstrates that non-coding genes exhibit larger gene and intron sizes—but smaller exons—compared to protein-coding genes with identical exon numbers. (Zhao, Y., Liu, W., Zeng, J., et al. (2016). Identification and analysis of mouse non-coding RNA using transcriptome data. Sci China Life Sc.)

Line 188: "from protein-coding gene to lncRNAs" should be "from protein-coding genes to lncRNAs".

Line 200-204: This part could benefit from conciseness refinements.

Line 228: In molecular biological and genetic contexts, cis, as a Latin-derived term, must be italicized.

Line 316-318: This statement suffers from a logical flaw: the study fails to establish a causal relationship between the reduced length of long non-coding RNAs and their novel regulatory functions or associated regulatory elements.

Line 321-323: The previous sentence only demonstrates an association between transposable elements (TEs) and transition of protein-coding genes to lncRNAs. There is no direct links between TEs and role in the regulatory landscape of lncRNAs. These conclusions warrant caution.

Line 366: Remove "over" in this sentence.

Line 639: There is a discrepancy between the graphical content and its accompanying legend in Figure 1B, 1C and 1D. Moreover, the lncRNA count in Figure 1D is significantly lower than the quantity reported in the text.

Line 690: Please redraw Figure 4A to visually demonstrate distributional differences between the two groups.

Line 702: Please separate two sentences with a space.

Line 739: Figure S1C duplicates the content of Figure 1D. In addition, there is a discrepancy between the graphical content and its accompanying legend in Figure S1C.

Line 757: The sense and antisense transcripts may be mislabeled in Figure S3A.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have revised the manuscript based on the reviewers'suggestions, resulting in some improvement in its quality. Necessary explanations were provided for aspects that could not be further improved, such as the lack of experimental evidence for the function of Msex-lnc-001918. However, there remain significant issues with the writing, primarily involving numerous grammatical errors and imprecise or unclear expressions in certain parts. Further revisions in writing are still required. Below, I only list some of these issues—all of the mentioned sections contain grammatical errors.

Line 50-53, Similarity, JPX,51 a lncRNA derived from Uspl, plays a critical role in regulating Xist, with deletion of...;
Line 93-95: The substantial variation (~36-fold) in lncRNA numbers across species with Lepidoptera ...and quantity effects.

Line 162-163: This approach identified 4,806 derived lncRNAs across six orders, with D. melanogaster
163 has derived lncRNAs, making it the species with the highest abundance (Figure 2B).

Line 164-165:Since Blattodea and Coleoptera have a small amount of derived lncRNAs, we thus only retained data from Hemiptera, Hymenoptera, Lepidoptera, and Diptera for further analysis.

Line 166-168: Among these four orders, the number of derived lncRNAs in Hymenoptera was significantly higher than in Hemiptera and Lepidoptera (**P < 0.01, Wilcoxon’s test), while there were no significant differences compared to Diptera (Figure S5A).

Line 168-170: Furthermore, when we analyzed the rates of genome alignment between different species, revealing that Hymenoptera insects exhibited higher alignment rates than the other orders.

Line 356-376:Basically, there are grammatical errors throughout the entire paragraph.

Line 377-379. Focusing on the specific derived lncRNA Msex-lnc-001918 in Manduca sexta, our analysis revealed testis-specific expression pattern, with higher expression than its syntenic protein-coding homologs.

Reviewer #2

(Remarks to the Author)
The authors concluded that the derived lncRNAs were shorter than their homologous protein-coding genes, as demonstrated by a global comparative analysis. However, the authors did not provide a more exhaustive comparative analysis of the structural characteristics between the derived lncRNAs and their homologous protein-coding genes. Let me just list a few. Is there a difference in the number of exons and introns between the derived lncRNAs and their corresponding homologous proteins? What is the trend in exon and intron length? Is the length change caused by repetitive sequences?

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
The authors made a more detailed comparison between derived lncRNAs and their syntenic genes.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.
In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.
The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.
To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

Point-by-point responses to Editor and Reviewers' comments

Reviewer #1

R1-1: The authors performed a comprehensive evolutionary and functional analysis of lncRNAs across 115 insect species, utilizing publicly available genomic data from the Insectbase2.0 database and NCBI. Their findings demonstrate that lncRNA abundance is positively correlated with both genome size and transposon content. Notably, lncRNAs derived from protein-coding genes exhibit distinct characteristics compared to conventional lncRNAs, including differences in expression levels, genomic distribution, and regulatory complexity. To further validate these findings, the study highlights Msex-lnc001918 from Manduca sexta as a case study, supporting the hypothesis that the functional roles of protein-coding gene-derived lncRNAs are closely linked to the biological functions of their ancestral genes.

In summary, this large-scale comparative analysis provides systematic insights into the genomic features of lncRNAs in insects. It elucidates the molecular mechanisms by which protein-coding genes give rise to novel regulatory elements through sequence transformation, offering a valuable theoretical framework for understanding the evolution and functional diversification of lncRNAs. However, the present manuscript requires revisions to address the following issues prior to publication.

Response:

We sincerely thank the reviewer for the thorough and insightful summary of our work. Your comments have helped us identify key areas for clarification and improvement. In response, we have carefully revised the manuscript to address the specific concerns raised, including adjustments to overly strong interpretations, enhancement of figure clarity, and expansion of relevant analyses and discussions. We believe these revisions have strengthened the clarity, rigor, and overall impact of our study.

R1-2: The authors used known RNA-seq data from 115 insect species for comparative analysis of lncRNAs across different insects. Given the inconsistencies in the quality or quantity of genomic data and RNA-seq data among many species, how did the authors ensure the reliability of the analysis results and conclusions based on these data?

Response:

Thanks for raising this critical concern about data quality consistency across species. To address this, we included only species with >50 transcriptomes and high-quality genome assemblies (BUSCO completeness > 90%). Moreover, we employed a standardized bioinformatics pipeline across all species to minimize technical variation. For downstream analyses such as correlation and saturation estimation, we further restricted to 14 species with >500 transcriptomes to reduce data bias. We acknowledge that transcriptome coverage varies significantly across species, and we have provided comprehensive metadata in **Supplementary Table 1** detailing sample numbers, and the SRA accession for each species. Our analyses and key biological conclusions were based on subsets of data with sufficient transcriptome coverage, ensuring the robustness and reliability of our findings across evolutionary comparisons.

R1-3. The authors proposed that *Msex-lnc-001918* may regulate spermatogenesis based on bioinformatic analysis, but this hypothesis lacks experimental validation. Functional studies using RNA interference (RNAi) or CRISPR/Cas9-mediated knockout of *Msex-lnc-001918* and its host gene would be necessary to confirm their roles in spermatogenesis.

Response:

We fully agree. Due to limited access to *Manduca sexta* in Chinese laboratories, and few groups currently maintain this species, to explore possible alternatives, we attempted to identify homologs of *Msex-lnc-001918* in other species, however, as low sequence conservation for lncRNA, and no clear orthologous sequences were identified, so we are presently unable to conduct the functional experiments. We rewrote the sentence in the manuscript and plan to actively seek international collaborations in future work to experimentally validate the role of *Msex-lnc-001918*.

Despite these limitations, we discussed the need for experimental validation in the revised **Discussion section**.

R1-4. In L53-54, references 14 and 16 are inappropriately placed at the end of the sentence. They should be moved after "...in vertebrates," as these two references are unrelated to lncRNAs in insects.

Response:

Thank you. We have moved references 14 and 16 to follow "in vertebrates" to maintain contextual accuracy.

R1-5. L247, please change "supplementally table2" to "Supplementary Table 2". In the Manuscript Items of the submission system, there are two entries labeled "Supplementary Table 1", with no "Supplementary Table 2". The second "Supplementary Table 1" should be "Supplementary Table 2".

Response:

We corrected both the reference in the main text (L247) and the label in the submission system. "Supplementally table2" has been revised to "Supplementary Table 2".

R1-6: L93-95, "The substantial variation (~36-fold) in lncRNA numbers across species with Lepidoptera showing the highest average number, followed by Hymenoptera and Diptera, …." This statement lacks support from any results or literature citation. It is recommended to cite relevant results or references.

Response:

We have added a citation to Figure S1C to support this observation and revised in Result section.

R1-7: The figure legend for Figure 1C contains an inconsistency regarding the x-axis label. While the main text clearly indicates that the x-axis represents genome size (Lines 111-113), the figure legend incorrectly describes it as both transcriptome size and sample number (Lines 650-652). This discrepancy should be corrected for accuracy.

Response:

Thanks for your suggestion. We corrected the figure legend of Figure 1C to clearly indicate the x-axis represents the number of transcriptome samples.

R1-8: L127-136, the results are only described without citing the corresponding figures or tables. Supporting data needs to be provided.

Response:

Thanks for your suggestion. We added explicit references to Figure 2A in this section to support

the described observations.

R1-9: L282-283: "Increasing numbers of functional studies on lncRNAs are revealing the importance, however, limited data are available for insects." Reference 34 does not support this conclusion.

Response:

We agree and have removed reference 34 and replaced it with a more relevant citation (Carrieri et al. 2012; Chen & Kim, 2024).

R1-10: L297-299, a reference citation is needed after "…findings in mammals".

Response:

We have revised the content (Lines 309 – 311) and added appropriate references (Kapusta et al. 2013) to support this statement.

R1-11: The punctuation marks in sentences L131 and L163 need to be corrected.

Response:

We have corrected the redundant and misplaced punctuation in both lines.

Reviewer #2

R2-1: This study provides a comprehensive analysis of drived lncRNAs from evolutionary and functional perspectives across 115 insect species. These derived lncRNAs show distinct features including lower length, lower transposable element fraction, higher expression levels, broader tissue distribution, and increased regulatory complexity compared to conventional lncRNAs. Functional analysis has revealed that *Msex-lnc-001918* has higher expression levels, more transposable elements and similar biological functions compared to its syntenic genes. In summary, this study provides fundamental insights into insect derived lncRNAs and demonstrates how protein-coding gene to lncRNA transformation creates novel regulatory innovations in insect genomes.

The authors identified massive multi-exonic lncRNAs. I'm particularly curious about the characteristics of single-exonic lncRNAs in insects.

Response:

We appreciate the reviewer's interest in the characteristics of single-exonic lncRNAs. While single-exon lncRNAs are indeed abundant in many transcriptomes, most current studies (Sarropoulos et al. 2019; Ulitsky et al. 2016), including ours, have focused primarily on multi-exonic lncRNAs. We focused on multi-exonic lncRNAs due to higher confidence in their annotation and lower likelihood of assembly artifacts. Single-exon transcripts are likely to result from transcriptional noise or incomplete transcript assembly. So, in our study, we aimed to minimize noise and ensure high-confidence lncRNA annotations across a broad taxonomic range. However, we fully acknowledge that single-exon lncRNAs may also play important regulatory roles. We have noted this in the revised Discussion(Lines 291 – 299) as a direction for future research, especially as long-read sequencing and improved transcriptome annotation methods become more widely available in insect species.

R2-2: The authors suggest that 500 diversified transcriptomes represent the minimum threshold for comprehensively identifying lncRNAs. This claim is questionable because discussing sample size independently of data size is methodologically unsound.

Response:

We thank the reviewer for this important comment. We agree that transcriptome coverage is influenced not only by the number of samples but also by sequencing depth, tissue diversity, and technical quality. In our analysis, we observed that lncRNA discovery tends to plateau in species with approximately 500 transcriptome datasets, particularly when these represent a broad range of developmental stages and tissues. However, we do not consider this number as a strict threshold. To clarify this point, we have revised the relevant section in the Results (Lines 100 – 103) and Discussion (Lines 301 – 304) to explicitly state that the "500-sample" mark represents an empirical benchmark rather than an absolute minimum, and that data volume and diversity are equally critical factors for comprehensive lncRNA identification.

R2-3: Given the well-documented pseudogene transcripts in genomes, the authors are essentially studying pseudogene transcripts through homologous protein screening for lncRNAs. The manuscript would benefit from incorporating pseudogene-related perspectives into the

Introduction and Discussion sections.

Response:

Thank you for the insightful suggestion. Pseudogenization of coding genes can serve as a substrate for regulatory innovation through lncRNAs. So derived lncRNAs may represent a functional subset of processed pseudogene transcripts. We have added a discussion on the relationship between pseudogenes and derived lncRNAs (Duret et al. 2016; Milligan &Lipovich, 2014; Wen et al. 2023) in Discussion (Lines 323 – 326) sections.

R2-4: Evolutionary conservation is critical for studying gene function. I recommend conducting a detailed comparative visualization between protein-coding genes and lncRNAs based on homologous families in results subsection "Evolutionary analysis of lncRNA orthologous families". For instance, a plot with the X-axis representing cross-species conservation levels of homologous families and the Y-axis showing the percentage representation of conserved families would visually demonstrate their conservation divergence.

Response:

We appreciate this suggestion and have added a new **supplementary figure (Figure S2)** to illustrate the comparative conservation distribution across lncRNA and protein-coding gene orthogroups.

R2-5: Figure 1 legend needs revision carefully.

Response:

We have revised the entire Figure 1 legend to ensure consistency with figures and clarity in variable definitions.

R2-6: The text needs to be proofread and revised. Some of the examples are listed as follows (please note that this is not a thorough list of all sentences and phrases to be revised):

Line 95-100: This sentence suffers from logical flaws. First, there is no direct link between the amount of transcriptome data and the quantity of lncRNAs, nor does it address the uneven distribution of lncRNAs across insect species. Second, direct numerical comparisons are problematic-lncRNA counts should be normalized against factors like genome size and protein-

coding gene numbers.

Response:

Thank you for pointing this out. We agree that the original statement lacked sufficient logical support and could be misleading. In response, we have removed this sentence from the revised manuscript to avoid overinterpretation and to maintain clarity in data interpretation.

R2-7: Line 106-111: The full scientific name should be used upon first mention of a species: "*A. aegypti*" should be "*Anopheles aegypti*" for the first time.

Response:

Thank you for the suggestion. We confirm that the full scientific name *Aedes aegypti* was already introduced upon first mention in Line 81 of the manuscript. Subsequent references use the abbreviated form "*A. aegypti*" in accordance with standard taxonomic convention.

R2-8: Line 131: Remove redundant punctuation marks.

Response:

Thank you for your careful review. We have thoroughly checked the manuscript and removed the redundant punctuation marks as suggested.

R2-9: Line 156-157: Please rephrase this sentence. For instance: "X and Y represent flanking genes of syntenic genes and lncRNA genes (located within 100 kb), respectively, with confirmed orthology."

Response:

Thank you for the helpful suggestion. We have revised the sentence on Lines 156–157 to improve clarity and align with the reviewer's proposed structure. The updated sentence now clearly defines the roles of X and Y and outlines the orthologous criteria more precisely.

R2-10: Line 180: "(Figure3 B)" should be "(Figure 3B)".

Response:

Thank you for noticing this formatting issue. We have corrected "(Figure3 B)" to "(Figure 3B)" in the revised manuscript.

R2-11: Line 181-182: The authors conclude that derived lncRNAs are shorter than their homologous protein-coding genes, yet omit comparative analysis of structural features. This gap particularly interests me, as our previous study demonstrates that non-coding genes exhibit larger gene and intron sizes—but smaller exons—compared to protein-coding genes with identical exon numbers. (Zhao, Y., Liu, W., Zeng, J., et al. (2016). Identification and analysis of mouse non-coding RNA using transcriptome data. Sci China Life Sc.)

Response:

Thank you for this valuable comment and for highlighting your relevant study. We have reviewed the structural characteristics of derived lncRNAs in more detail and expanded the corresponding section in the revised manuscript. Specifically, we now include a comparative analysis of exon and intron lengths between derived lncRNAs and their syntenic protein-coding genes, which reveals a significant reduction in both features (**Figure S6A, B**). To provide broader context, we have also cited Zhao et al. (2016), which reported contrasting patterns in mammals. This comparison highlights potential lineage-specific differences in the structural evolution of non-coding genes.

R2-12: Line 188: "from protein-coding gene to lncRNAs" should be "from protein-coding genes to lncRNAs".

Response:

Thank you for pointing this out. We corrected it accordingly.

R2-13: Line 200-204: This part could benefit from conciseness refinements.

Response:

Thank you for the suggestion. We have revised Lines 200–204 to improve conciseness and clarity while preserving the original meaning. The revised version enhances readability and aligns with the overall tone of the manuscript.

R2-14: Line 228: In molecular biological and genetic contexts, cis, as a Latin-derived term, must be italicized.

Response:

Thank you for your suggestion. We have italicized cis throughout the manuscript where used in molecular and genetic contexts, as per standard formatting conventions.

R2-15: Line 316-318: This statement suffers from a logical flaw: the study fails to establish a causal relationship between the reduced length of long non-coding RNAs and their novel regulatory functions or associated regulatory elements.

Response:

Thank you for highlighting this important point. We have revised the text to clarify that the observed reduction in lncRNA length is correlated with, but not necessarily causally linked to, increased regulatory complexity. The revised wording avoids overinterpretation and more accurately reflects the scope of our findings.

R2-16: Line 321-323: The previous sentence only demonstrates an association between transposable elements (TEs) and transition of protein-coding genes to lncRNAs. There is no direct links between TEs and role in the regulatory landscape of lncRNAs. These conclusions warrant caution.

Response:

Thank you for this thoughtful comment. We agree that our original statement may have overstated the functional implications of transposable elements (TEs) in the regulatory landscape of lncRNAs. We have revised the relevant sentence to clarify that our findings demonstrate an association between TE insertions and the structural transition from protein-coding genes to lncRNAs, but do not establish a direct regulatory role. The updated wording adopts a more cautious tone and avoids overinterpretation.

R2-17: Line 366: Remove "over" in this sentence.

Response:

Thank you for pointing this out. We have corrected it in the revised manuscript.

R2-18: Line 639: There is a discrepancy between the graphical content and its accompanying legend in Figure 1B, 1C and 1D. Moreover, the lncRNA count in Figure 1D is significantly lower

than the quantity reported in the text.

Response:

Thank you for carefully reviewing the figures. We have now corrected these issues by regenerating the affected panels and updating both the figure legends and corresponding descriptions in the main text to ensure consistency and accuracy.

R2-19: Line 690: Please redraw Figure 4A to visually demonstrate distributional differences between the two groups.

Response:

Thank you for the helpful suggestion. In response, we have replaced Figure 4A with a violin plot to more effectively visualize the distributional differences between derived lncRNAs and conventional lncRNAs. This revised format provides clearer insights into the density and variability of genomic positioning across the two groups and enhances interpretability compared to the previous version.

R2-20: Line 702: Please separate two sentences with a space.

Response:

Thank you for your attention to detail. We have added the missing space to separate the two sentences on Line 702 in the revised manuscript.

R2-21: Line 739: Figure S1C duplicates the content of Figure 1D. In addition, there is a discrepancy between the graphical content and its accompanying legend in Figure S1C.

Response:

Thank you for pointing this out. Figure S1C corresponded to the content previously described in Lines 95 – 100, which we have now removed based on earlier reviewer suggestions regarding logical clarity and normalization. Accordingly, Figure S1C has also been deleted from the revised manuscript to eliminate redundancy with Figure 1D and to ensure consistency between figures and the main text.

R2-22: Line 757: The sense and antisense transcripts may be mislabeled in Figure S3A.

Response:

Thank you for carefully examining Figure S3A. The sense and antisense designations have now been corrected in the updated version of Figure S4A, and the revised figure has been included in the resubmitted manuscript.

Point-by-point responses to Editor and Reviewers' comments

Reviewer #1

The authors have revised the manuscript based on the reviewers' suggestions, resulting in some improvement in its quality. Necessary explanations were provided for aspects that could not be further improved, such as the lack of experimental evidence for the function of *Msex-lnc-001918*. However, there remain significant issues with the writing, primarily involving numerous grammatical errors and imprecise or unclear expressions in certain parts. Further revisions in writing are still required. Below, I only list some of these issues—all of the mentioned sections contain grammatical errors.

Response:

We sincerely thank the reviewer for carefully pointing out the grammatical issues. We have revised the manuscript extensively to improve the English writing, grammar, and clarity. Beyond these specific examples, we carefully reviewed the entire manuscript, corrected grammatical mistakes, improved sentence structure, and ensured consistent use of tense and terminology. We believe these revisions have significantly improved the manuscript's presentation and clarity, and we thank the reviewer for this helpful feedback.

R1-1: Line 50-53, Similarity, JPX,51 a lncRNA derived from Uspl, plays a critical role in regulating Xist, with deletion of...;

Response:

We corrected this sentence to improve clarity and grammar.

R1-2: Line 93-95: The substantial variation (~36-fold) in lncRNA numbers across species with Lepidoptera ...and quantity effects.

Response:

We revised this sentence for precision.

R1-3: Line 162-163: This approach identified 4,806 derived lncRNAs across six orders, with D. melanogaster 163 has derived lncRNAs, making it the species with the highest abundance (Figure 2B).

Response:

We corrected the grammar and restructured the sentence.

R1-4: Line 164-165: Since Blattodea and Coleoptera have a small amount of derived lncRNAs, we thus only retained data from Hemiptera, Hymenoptera, Lepidoptera, and Diptera for further analysis.

Response:

We corrected the grammar and wording.

R1-5: Line 166-168: Among these four orders, the number of derived lncRNAs in Hymenoptera was significantly higher than in Hemiptera and Lepidoptera (**P < 0.01, Wilcoxon's test), while there were no significant differences compared to Diptera (Figure S5A).

Response:

We revised the sentence to improve clarity.

R1-6: Line 168-170: Furthermore, when we analyzed the rates of genome alignment between different species, revealing that Hymenoptera insects exhibited higher alignment rates than the other orders.

Response:

We corrected this incomplete sentence.

R1-7: Line 356-376: Basically, there are grammatical errors throughout the entire paragraph.

Response:

We carefully revised this paragraph, corrected grammatical errors, and improved sentence structure to ensure clarity and readability.

R1-8: Line 377-379. Focusing on the specific derived lncRNA Msex-lnc-001918 in Manduca sexta, our analysis revealed testis-specific expression pattern, with higher expression than its syntenic protein-coding homologs.

Response:

We revised this sentence for clarity.

Reviewer #2

R2-1: The authors concluded that the derived lncRNAs were shorter than their homologous protein-coding genes, as demonstrated by a global comparative analysis. However, the authors did not provide a more exhaustive comparative analysis of the structural characteristics between the derived lncRNAs and their homologous protein-coding genes. Let me just list a few. Is there a difference in the number of exons and introns between the derived lncRNAs and their corresponding homologous proteins? What is the trend in exon and intron length? Is the length change caused by repetitive sequences?

Response:

We appreciate the reviewer’s insightful suggestion regarding the comparison between derived lncRNAs and their syntenic protein-coding genes. In the revised manuscript, we have added a detailed comparative analysis of exon and intron numbers and length distributions. Our results show that derived lncRNAs consistently have fewer exons and introns, and that both exons and introns are significantly shorter compared to their homologous protein-coding genes (new Figure S6A-D). Furthermore, we examined repetitive sequence content and found that the shortening trend may associate with increased TE insertions during the transition from coding genes to lncRNAs (see revised Results, lines 181–185; revised Discussion, lines 334–344). These additional analyses strengthen our conclusion that derived lncRNAs represent structurally reduced but functionally adapted transcripts.

Point-by-point responses to Editor and Reviewers' comments

Reviewer #2

R2-1: The authors made a more detailed comparison between derived lncRNAs and their syntenic genes.

Response:

We sincerely thank the reviewer for the comment. We are pleased that the reviewer recognized the expanded and detailed comparison between derived lncRNAs and their syntenic genes in our revised manuscript. We truly appreciate this acknowledgment, which confirms that the additional analyses and clarifications have strengthened the clarity of our study.