nature portfolio

Peer Review File

Benchmarking Genome Choice in Functional Genomics Analyses

Corresponding Author: Dr Ting Wang

This manuscript has been previously reviewed at another journal. This document only contains information relating to versions considered at Nature Communications.

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)
With respect to my suggestion to alter the narrative to focus on benchmarking combinations of reference genomes and analytical pipelines instead of just reference genomes, the authors only made a slight adjustment to the title and didn't make any other edits.

In response to another of my previous comments the authors have this revised sentence in the abstract:
"This approach enabled quantification of CpG methylation in sequence absent from linear assemblies, revealing that non-reference Alu elements are more highly methylated than their reference counterparts, whereas L1 elements do not show a consistent statistically significant difference across individuals."

I still find it a little misleading since a reader might get the impression the sentence is trying to suggest a stronger methylation in Alu and L1, but L1 elements on average had greater methylation levels than Alu elements.

Also note the abstract word length is well over the stated limit for the journal.

Otherwise I don't have additional comments at this stage.

Reviewer #4

(Remarks to the Author)
As far as I can tell, the authors have sufficiently responded to reviewer 3's criticisms. I support publication in Nature Communications.

Minor comment: There seems to be some repetition of text in consecutive paragraphs on page 21: Increasing the threshold to 16 raised the proportion of individual-specific peaks slightly to 1.6 ± 0.27%. The impact of hg38 blacklisted regions was minimal, with 0.13% of peaks unique to the unfiltered analysis and 0.01% unique after filtering.

REVIEWERS' COMMENTS

We sincerely thank the Editor and the Reviewers for their careful reading of our manuscript and for their thoughtful, constructive comments. We are grateful for the opportunity to revise the work, and we believe the manuscript has been improved by their feedback.

Reviewer #1 (Remarks to the Author):

With respect to my suggestion to alter the narrative to focus on benchmarking combinations of reference genomes and analytical pipelines instead of just reference genomes, the authors only made a slight adjustment to the title and didn't make any other edits.

Response 1.1: We thank the reviewer for this comment. We have revised the manuscript to shift the framing in line with the reviewer's point, clarifying that several comparisons in this study reflect genome representation together with the analytical frameworks currently used to analyze each representation, rather than genome sequence alone in complete isolation. We made these revisions in the Introduction (paragraphs beginning around lines 99, 108, and 113) and in the Discussion (paragraphs beginning around lines 763 and 780). We also shortened the title to Benchmarking Genome Choice in Functional Genomics Analyses to comply with the journal's title length requirements.

In response to another of my previous comments the authors have this revised sentence in the abstract:
"This approach enabled quantification of CpG methylation in sequence absent from linear assemblies, revealing that non-reference Alu elements are more highly methylated than their reference counterparts, whereas L1 elements do not show a consistent statistically significant difference across individuals."

I still find it a little misleading since a reader might get the impression the sentence is trying to suggest a stronger methylation in Alu and L1, but L1 elements on average had greater methylation levels than Alu elements.

Also note the abstract word length is well over the stated limit for the journal.

Otherwise I don't have additional comments at this stage.

Response 1.2: We thank the reviewer for this comment. To address both points, we revised the abstract by shortening it to comply with the journal word limit and by removing the sentence comparing methylation patterns of non-reference Alu and L1 elements that the reviewer identified as potentially misleading. The revised abstract now focuses on the broader conclusion that pangenome-based analysis

reveals biologically relevant features inaccessible to linear references, without making a comparison that could be misinterpreted.

Reviewer #4 (Remarks to the Author):

As far as I can tell, the authors have sufficiently responded to reviewer 3's criticisms. I support publication in Nature Communications.

Minor comment: There seems to be some repetition of text in consecutive paragraphs on page 21: Increasing the threshold to 16 raised the proportion of individual-specific peaks slightly to 1.6 ± 0.27%. The impact of hg38 blacklisted regions was minimal, with 0.13% of peaks unique to the unfiltered analysis and 0.01% unique after filtering.

Response 2.1: We thank the reviewer for this comment. We removed the duplicated text identified in the Results section