Reviewer Report

Title: GuideMaker: Software to design CRISPR-Cas guide RNA pools in non-model genomes

Version: Original Submission **Date:** 7/21/2021

Reviewer name: Wen Wang, Ph.D

Reviewer Comments to Author:

The author developed a software, GuideMaker, for designing CRISPR-Cas guide RNA pools in non-model genomes. Three bacterial genomes, a fungal genome, and a plant genome were used in performance benchmarking, which proves that the software supports the design of gRNAs in non-standard Cas enzymes for non-model organisms at the genome-scale. However, the advantages of this software are not well estimated nor presented compared to other tools like CHOPCHOP. Also, the software was mainly evaluated in three bacteria genomes, one fungus and Arabidopsis genome. There are no tests for non-model plant or animal genomes. Therefore, the "non-model genomes" in the title are exaggerated. I list more problems as follows.

Major comments:

1. The authors did not compare the computation resources and performance (running time, memory) with existing softwares like CHOPCHOP. Also, the authors need to compare the score rankings with CHOPCHOP to present the relative power of GuideMaker. Is there any score rankings concerning efficiency or off-target possibilities for the designed Guide RNAs
2. It is better to add support for gff formated annotation input files since many non-model species do not have GenBank annotations.
3. The authors mentioned GuideMaker can design gRNAs for any small to medium size genome (up to about 500 megabases). The maximum genome used in the article was Arabidopsis thaliana (114.1MB), which is obviously smaller than the described (up to about 500 megabases). We couldn't find the description whether the authors had investigated the larger genomes. Therefore, the detailed analysis or discussion of this problem is needed.
4. The authors stated GuideMaker to design CRISPR-Cas guide RNA pools in non-model genomes. Arabidopsis thaliana is a model organism and test in a non-model plant genome will be highly valuable.
5. It is also stated that GuideMaker can design gRNAs for any PAM sequence from any Cas system but the results of SaCas and StCad was described in only one sentence.
6. The source of the genomes was missing in the manuscript. In particular, some species have multiple genome versions in the same database. Therefore, to make the results more repeatable, the specific website and version number for each species are needed.

Minor comments:

There are many typos. I give some examples here.

1. Line 11, "bacteria" should be "bacterias".
2. Line 38, delete the", including non-model organisms"ï¼Œprokaryotic and eukaryotic organisms include the non-model organisms.
3. Line 111, "candidates guides" should be "candidate guides".

4. Line154, "gRNA identify with GuideMaker" should be "gRNA identified with GuideMaker".
5. Line 195, "The second way GuideMaker reduces…" should be "The second way that GuideMaker reduces…".
6. Line 204, "and", no need for italics.
7. Line 207, "gRNA's" should be "gRNAs".
8. Lines 209-210, "we anticipate performance will…" should be "we anticipate that performance will…".
9. Figure. 1. It seems that the font size of the description of Control gRNAs is inconsistent with others, please check.
10. Line 22,55,98,159,175,187,219 and 247, "Guidemaker" should be "GuideMaker".
11. Line 262, "CAS" should be "Cas".
12. Supplementary Figure 4. Grammar mistake in sentence "the different number of logical cores with or without AVX2 settings are available". It should be "the different number of logical cores with or without AVX2 settings is available".

Level of Interest

Please indicate how interesting you found the manuscript: Choose an item.

Quality of Written English

Please indicate the quality of language in the manuscript: Choose an item.

Declaration of Competing Interests

Please complete a declaration of competing interests, considering the following questions:

- Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
- Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
- Do you hold or are you currently applying for any patents relating to the content of the manuscript?
- Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?
- Do you have any other financial competing interests?
- Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

Choose an item.

To further support our reviewers, we have joined with Publons, where you can gain additional credit to further highlight your hard work (see: https://publons.com/journal/530/gigascience). On publication of this paper, your review will be automatically added to Publons, you can then choose whether or not to claim your Publons credit. I understand this statement.

Yes Choose an item.