Meta-review – main suggestions

- **Accessibility (explain pangenomes/graphs for NLP readers) and Motivation**
  Change: We have largely improved writing. We added more details in the introduction, introducing the motivation around the classical tokenization and pangenome graph usage; We expanded DNA basics and pangenome graph explanation; clarified advantages and linkage to our method; refined figures with clearer examples/notations.
  Where: §1, §2.2, Appendix A, Figure 2,3

- **Model choice & compute limits stated upfront.**
  Change: Stated model choices and computational constraints clearly in the main text.
  Where: §2.1.

- **Remove privacy/DP distractions.**
  Change: delete all DP related mentions/graphics;
  Where: Fig.1, Appendix B2.

Other reviewer points

Y4qk

- **Task choices**
  Change: We added an important explanation to our task choice for the experiment and emphasized why generation + alignment is a better task-metric combination compared with classification tasks.
  Where: §4.2.

- **value of synthetic data.**
  Change: We added a detailed downstream tasks discussion. We in last round proposed the variant calling task as an extra experiment and asked if the reviewers would want to see it. And although not required in the meta review, we added the variant calling experiment to show the utility of synthetic data besides the alignment score.
  Where: Appendix C.2

- **Acronyms/typos.**
  Change: We did proofreading and corrected all the acronyms without explanation and typos we found.

mspM

- **PNT limitation emphasize.**
  Change: We make this clear in the main body and limitation section.
  Where: §3.2.1 , §7.

zGYB

- **Metric fairness across tokenizations unclear.**
  Change: We explain how Character-level prediction accuracy is more fair, and we emphasize that the alignment score is the most important metric in the paper.
  Where: §4.1, §4.2.

- **Alignment percentage**
  Change: We added an explanation of the alignment percentage.
  Where: §4.2.

- **Interpolation/extrapolation/memorization.**
  Change: We added a discussion on how they are reflected, and discussed how memorization is not an issue, as shown in our experiment.
  Where: §5.1, Appendix C.