# Review of: ""polishCLR: a Nextflow workflow for polishing PacBio CLR genome assemblies""

weiwen wang[1]

1 Australian National University

Potential competing interests: The author(s) declared that no potential competing interests exist.

This manuscript demonstrated an automated assembly polish pipeline (polishCLR) for PacBio long reads. It is a Nextflow workflow that is convenient for users to use. This workflow provides assessments for each polish step. In short, polishCLR provides a solution for researchers to conveniently polish PacBio assemblies.

I thought polishCLR was haplotype-aware, which provided different polish strategies for haplotype-resolved or unresolved assemblies, but it seems not (maybe Figure S1 contains more details, but the figure resolution is too low). From my understanding, the process difference between Case 1,2, and 3 is the number of rounds of Arrow polish in step 1. If I also apply one round of Arrow polish on Case 3 in step 1, will it cause some errors? Or whether it is just useless because the Case 3 assemblies have been polished before?

Additionally, I also have some small comments as follows:

1. Can users define how many rounds of polish? Such as 3 rounds of Arrow polish and 3 rounds of FreeBayes polish?

2. I found the Dockerfile on Github, maybe it is better to also describe the installation through Dokcer.

3. Although polishCLR provided a test sample, the test sample is too large (>20 GB). I suggest authors provide a small sample to enable users to test polishCLR in a short time.

4. Whether polishCLR supports the plasmid assembly as input? If yes, I suggest authors change the name of "--mitochondrial_assembly" to "--organellar_assembly", or add an argument --plasmid_assembly"? Also, why the "--mitochondrial_assembly" is one of the mandatory arguments? Some projects may not assemble the mitochondrial genome.