
* GEP Finishing Checklist for Hybrid Assemblies

Project Name: _____
Student Finisher: _____

Goal: all sequence in one contig, all bases in the consensus having Phred >25, optimally both strands sequenced. For regions covered by a single strand (very unlikely), bases must have a Phred score >30. All locations with potential issues carefully evaluated and consensus corrected if indicated.

Project Status as submitted:

- Completely finished, all potential problem sites evaluated and resolved, all gaps filled, suspicious areas tagged for follow-up assessment
 Project has 1 or more gaps that need PCR/Sanger data
 No PCR/Sanger data needed, but sites needing careful inspection remain

Single Nucleotide Polymorphisms (SNPs)

- Tagged all putative SNPs (i.e. regions with a high number of HQ discrepancies not associated with mononucleotide repeats and not associated with repeats with improperly mapped reads)

Verify the Assembly

- Project is in a single contig (i.e. any gaps are resolved)
 Comment tag on any "PCR only" regions

Use "search for string" to check the following:

- X's in consensus (none should be present - "X" indicates vector!)
 N's in consensus (none should be present - "N" indicates gap or ambiguous base)

Final Checks

- Add comment tag at the left end of the contig with your name, professor and school
 Save final version with the suffix ace.0