

C: TP, TN, FP and FN totals across all 100 simulated-read test genomes for each short-read polishing tool. These values compare unpolished sequences to the result of a single round of polishing. Polypolish performed best on TN and FP metrics (i.e. at error-free positions of the assembly) and NextPolish performed best on TP and FN metrics (i.e. at error-containing

D: The only instance of false positives in Polypolish assemblies occurred in the NC 002967.9 genome. The unpolished assembly contained a 3-bp deletion in a 7-bp tandem repeat, and Polypolish erroneously changed this into a 5-bp deletion,

which (on a per-reference-base basis) created two false positives and three false negatives.

Polisher

NextPolish

HvPo

TP

406804

408177

TN

366150176

366150103

FP

51

102

FN

4766

3415

polished assembly

error

False negative (FN):

the polisher failed to

False positive (**FP**):

correct an error

no error

True positive (TP):

fixed an error

positions of the assembly).

the polisher correctly

True negative (**TN**):