A

unpolished assembly \ polished assembly	no error	error
error	True positive (**TP**): the polisher correctly fixed an error	False negative (**FN**): the polisher failed to correct an error
no error	True negative (**TN**): the polisher did not change a correct base	False positive (**FP**): the polisher introduced an error

B

```
                    TN
reference:          AACACCCATCAACAT-CGGTATCCGTTATAG
before polishing:   AACACC-ATCAACAT-CGGTATCCATTATAG
after polishing:    AACACCCATCAACATACGGTATCCATTATAG
                       TP       FP         FN
```

C

Polisher	TP	TN	FP	FN
HyPo	406804	366150176	51	4766
NextPolish	408177	366150103	102	3415
ntEdit	376561	366149853	371	35012
Pilon	406393	366147448	2760	5196
POLCA	404541	366150151	37	7068
Polypolish	406627	366150226	2	4942
Racon	403936	366145001	5161	7699
wtpoa	397472	357964313	8183772	16240

D

```
                tandem repeat
reference:      AACAAGCAACAAGCAACAAGCAACAAGCAACAAGCAACAAGCAACAAGCAACAAGCAACAAGCAACAAGCAACAAGC
before polishing: AACAAGCAACAAGCAACAAGCAACAAGCAACAAGCAACAA---GCAAGCAACAAGCAACAAGCAACAAGCAACAAGC
after polishing:  AACAAGCAACAAGCAACAAGCAACAAGCAACAAGCAAC-----ACAAGCAACAAGCAACAAGCAACAAGCAACAAGC
                                                            FP FN
```

Figure S7: confusion matrices for the simulated-read tests.

A: Definitions of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) classifications in the context of assembly polishing.

B: Labelled sample sequence. A result (TP, TN, FP or FN) is given to each base in the reference, which can include multiple assembly bases in the case of insertion errors.

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 positions of the assembly).

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.