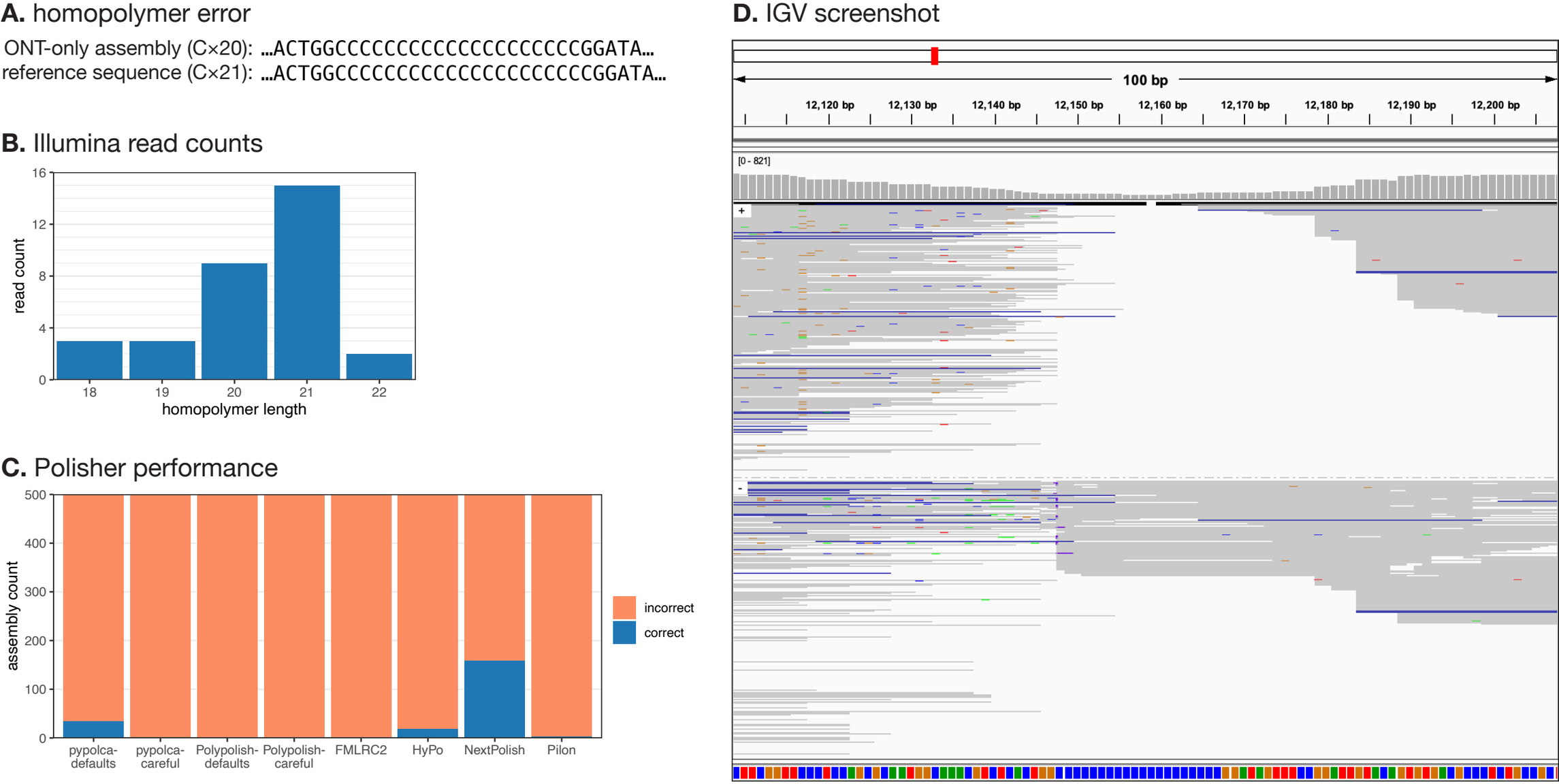

Figure SX: long homopolymer in *Salmonella enterica* ATCC-10708.

A: This genome's plasmid contains a long homopolymer at position 12148, and there is a discrepancy between the ONT-only assembly (C×20) and the polished reference sequence (C×21).

B: Illumina reads counts with exact matches to the homopolymer plus five bp of adjacent sequence on both sides. There were 32 matching reads in total, 15 of which supported C×21. All matching reads were on the G-strand and none were on the C-strand. This relative lack of reads and strand bias creates uncertainty in the true homopolymer length for this genome, and the C×21 in the reference sequence is only a best guess.

C: Assuming that C×21 is the correct homopolymer length, this plot shows assembly counts with the correct (blue) and incorrect (orange) length at this homopolymer, for each polisher tested. NextPolish performed best, but none of the polishers were able to reliably fix this error.

D: Integrative Genomics Viewer (IGV) screenshot of the relevant region of the plasmid, with Illumina read alignments grouped by strand. This was generated using the full set of Illumina reads (367× depth).