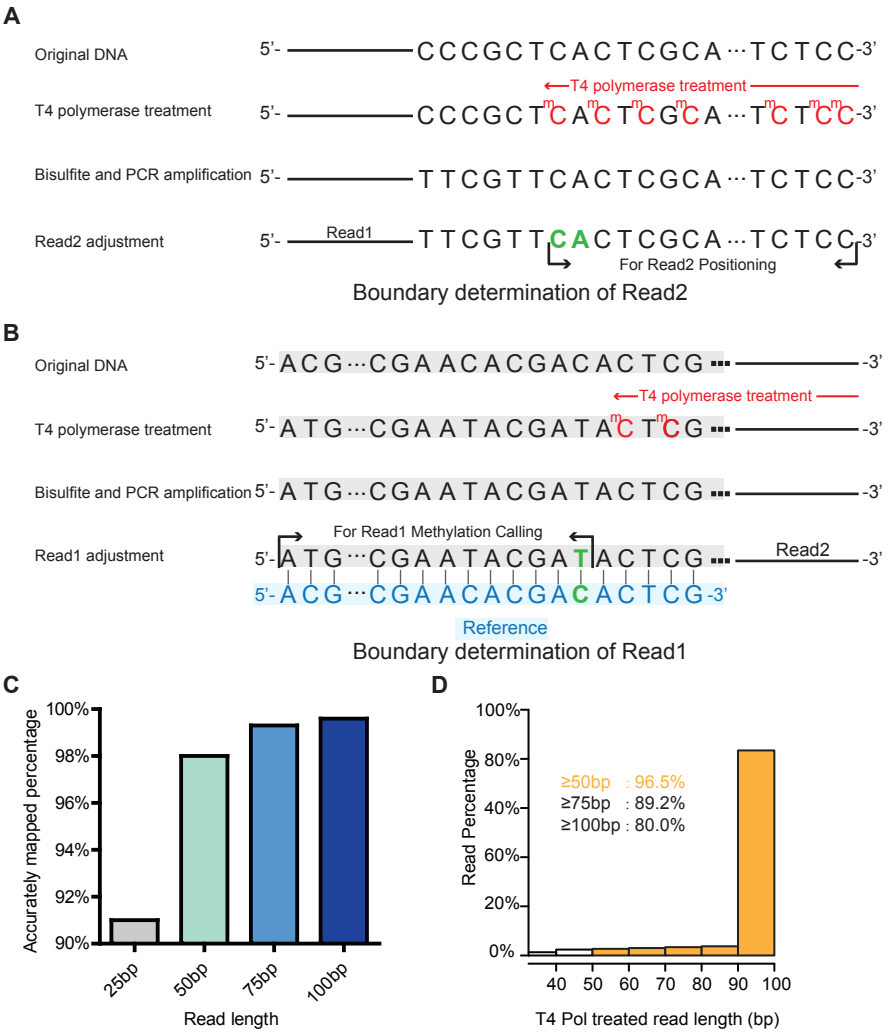

Supplementary Figure S15. Normalization of Read2 and Read1 and efficiency of T4 polymerase treatment.

(A) Read2 adjustment was performed when DNA was not sufficiently treated by T4 DNA polymerase. Red ^{m}C represents the location of cytosine replaced by methyl-cytosine after T4 DNA polymerase treatment. Green CA represents the farthest CH from 3' end of Read2, which was identified as the boundary of T4 DNA polymerase treatment. Fragment between boundary and 3' end of Read2 was picked out for alignment. **(B)** Read1 adjustment was performed when DNA was over-treated by T4 DNA polymerase. Green C->T conversion represents first C->T conversion from 3' end of Read1, which was identified as the boundary of T4 DNA polymerase treatment. **(C)** Mapping accuracy of simulated reads with different length. Simulated single-end reads were aligned to reference genome with Bowtie 2. The percentage of mapped accuracy was increased as read length becomes longer, and it clearly shows that 50bp reads are long enough to map to the correct genome position. **(D)** The performance of T4 DNA polymerase treatment in Read2 of GPS. About 80% of Read2 are completely treated by T4 DNA polymerase treatment. Meanwhile, more than 96.5% of Read2 were at least 50bp treated by T4 DNA polymerase (Yellow bar), indicating the high efficiency of GPS. Taken **C** and **D** together, the Read2 of GPS will map to the correct position with high efficiency, which further ensures to guide Read1 to the right position for DNA methylation calling.