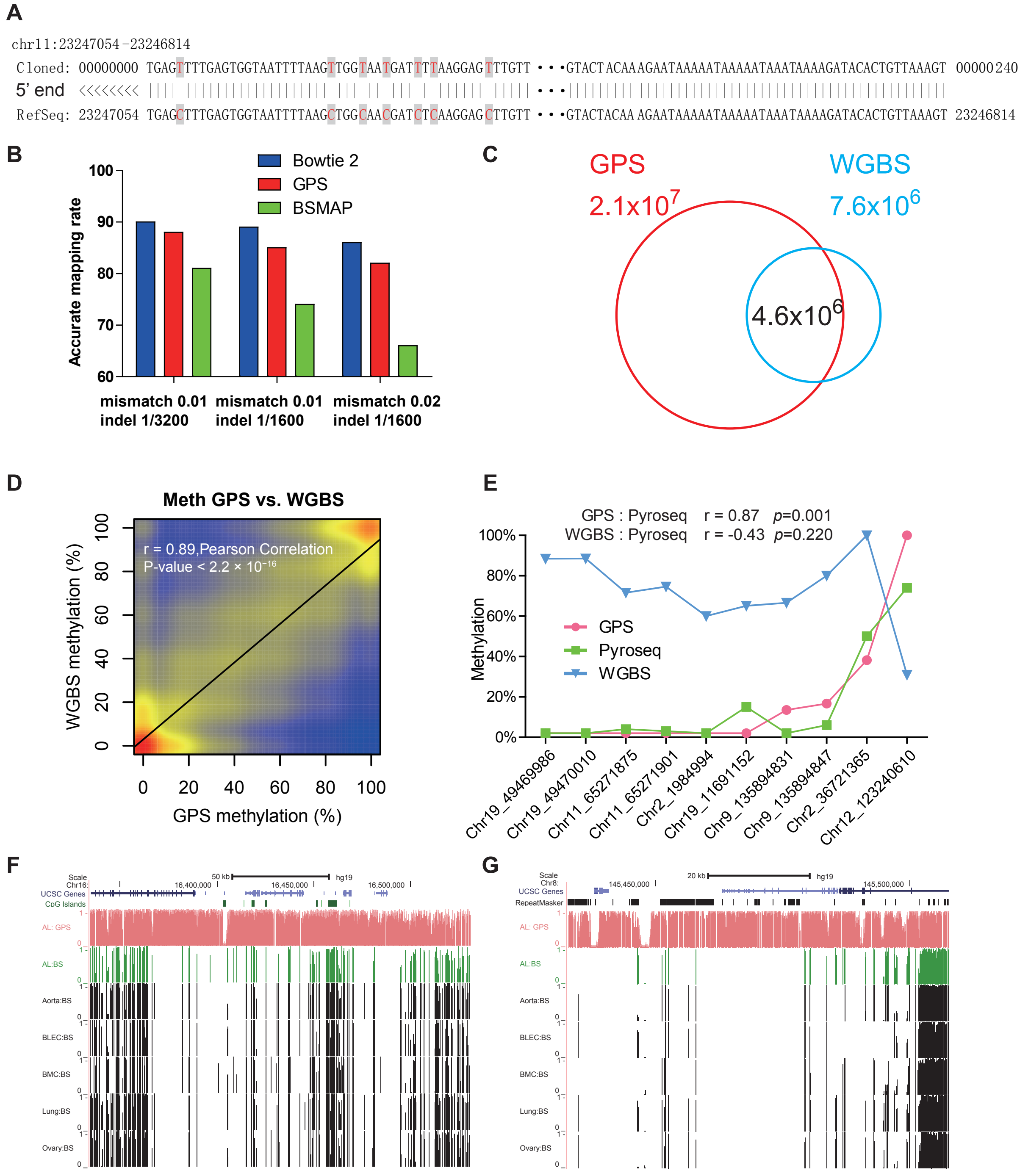

Supplementary Figure S1. DNA methylation detection performance of GPS comparing to WGBS.

(A) DNA fragment with cytosines (Cs) to thymines (Ts) conversion in the 5’ end and unconverted Cs in the 3’ end by GPS. **(B)** Distinct performance of Bowtie 2, GPS and BSMAP in simulated sequence with mismatch or Indel. Mismatch 0.01 represents that the mutation occurrence probability of base pairs except C->T conversion is 1%. (**C**) GPS detected more CpG sites than WGBS at ≥5x coverage. There are $2.1x10^7$ CpG sites detected by GPS and $7.6x10^6$ by WGBS, with $4.6x10^6$ CpG sites detected by both methods under similar amount of bisulfite-converted data (375M reads). (**D**) The methylation level in $4.6x10^6$ CpG sites identified by both methods is strongly correlated (R=0.89, Pearson correlation). Linear correlation analysis is shown (black line). (**E**) Methylation of CpG sites detected by GPS and WGBS verified by bisulfite pyrosequencing. Bisulfite pyrosequencing performed on 10 CpG sites whose methylation levels had the biggest difference between WGBS and GPS (see details in Supplementary Methods). Methylation levels verified by bisulfite pyrosequencing had highly positive correlation with GPS results (r = 0.87, P-value=0.001, Pearson correlation), but not with WGBS (r = -0.43, P-value=0.220, Pearson correlation). (**F,G**) UCSC Genome Browser screenshots showing methylation signals by GPS and WGBS on two representative genomic regions. Left panel (**F**) is covered on CpG islands, and right (**G**) on repetitive elements (RepeatMasker). The WGBS datasets were obtained from NIH Roadmap Epigenomics Mapping Consortium (AL: Adult Liver; BLEC: Breast Luminal Epithelial Cell; BMC: Breast Myoepithelial Cells; BS: WGBS).