

Figure S1. Linkage disequilibrium pattern surrounding rs1011970 in (A) Utah Residents with European Ancestry, (B) Han Chinese in Beijing and (C) Yoruba in Ibadan populations from the 1000 Genomes Project. Each square indicates one single nucleotide polymorphism. The color from white to black corresponds to the  $r^2$  value (from 0 to 1). The vertical arrow indicates the position of rs1011970.

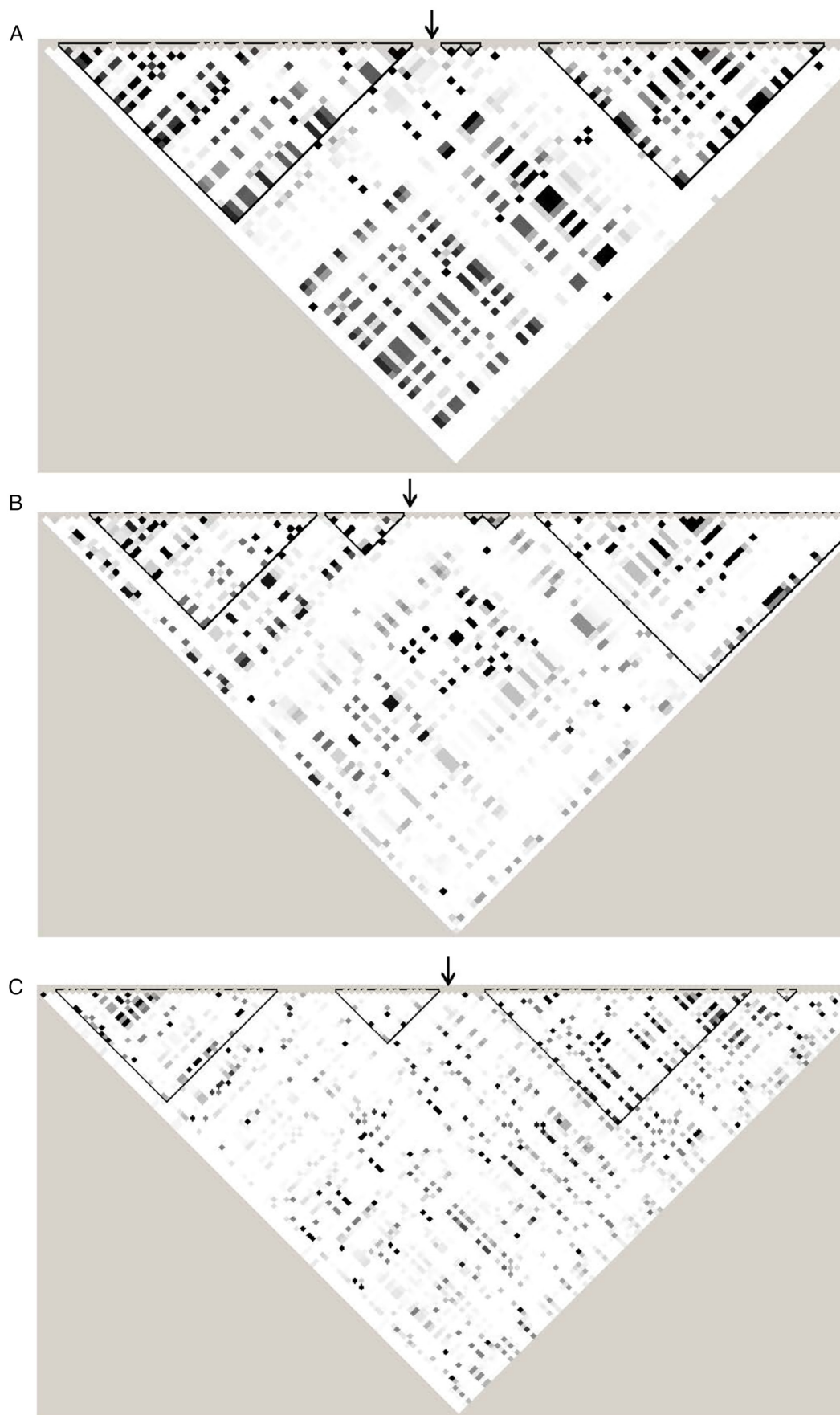
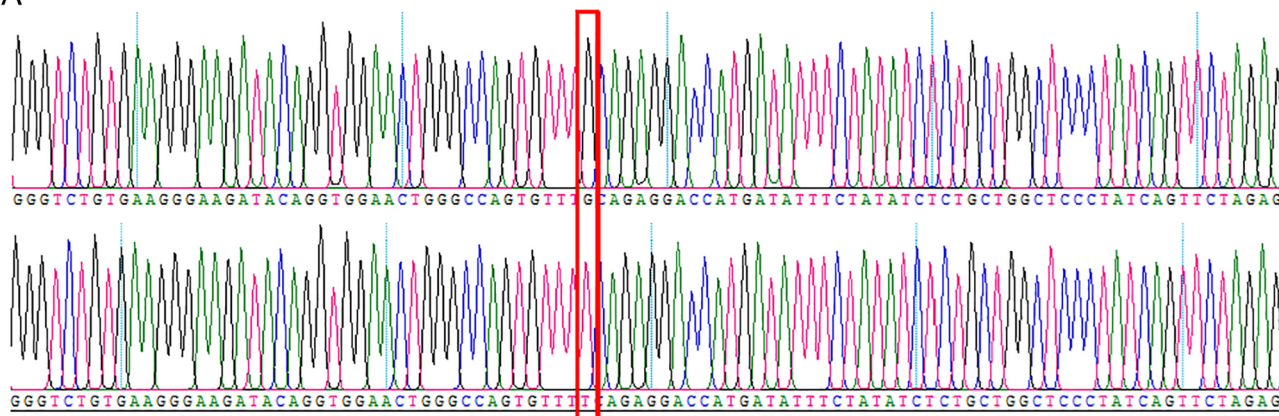


Figure S2. The sequence trace of the plasmids containing rs1011970 (A) and rs77283072 (B) alleles. For each part, the above plasmid is from original construction while the below one from mutagenesis. The red box indicates the position of rs1011970 or rs77283072.

A



B

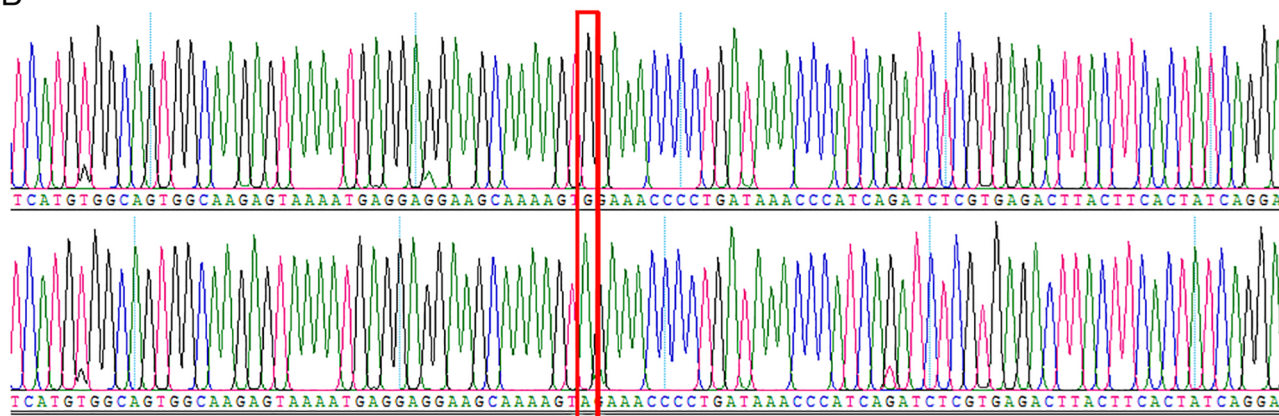


Figure S3. Histone modification in breast cancer HMEC cell line for the segment surrounding rs77283072 (chr9:22060733-22060933). The vertical bold turquoise line indicates the location of rs77283072. Each row represents one type of histone modification. The x-axis indicates the genome coordinate in chromosome 9. The y-axes each indicate the signal magnitude of chromatin immunoprecipitation sequencing. The pink line indicates the edge of each histone modification enrichment peak.

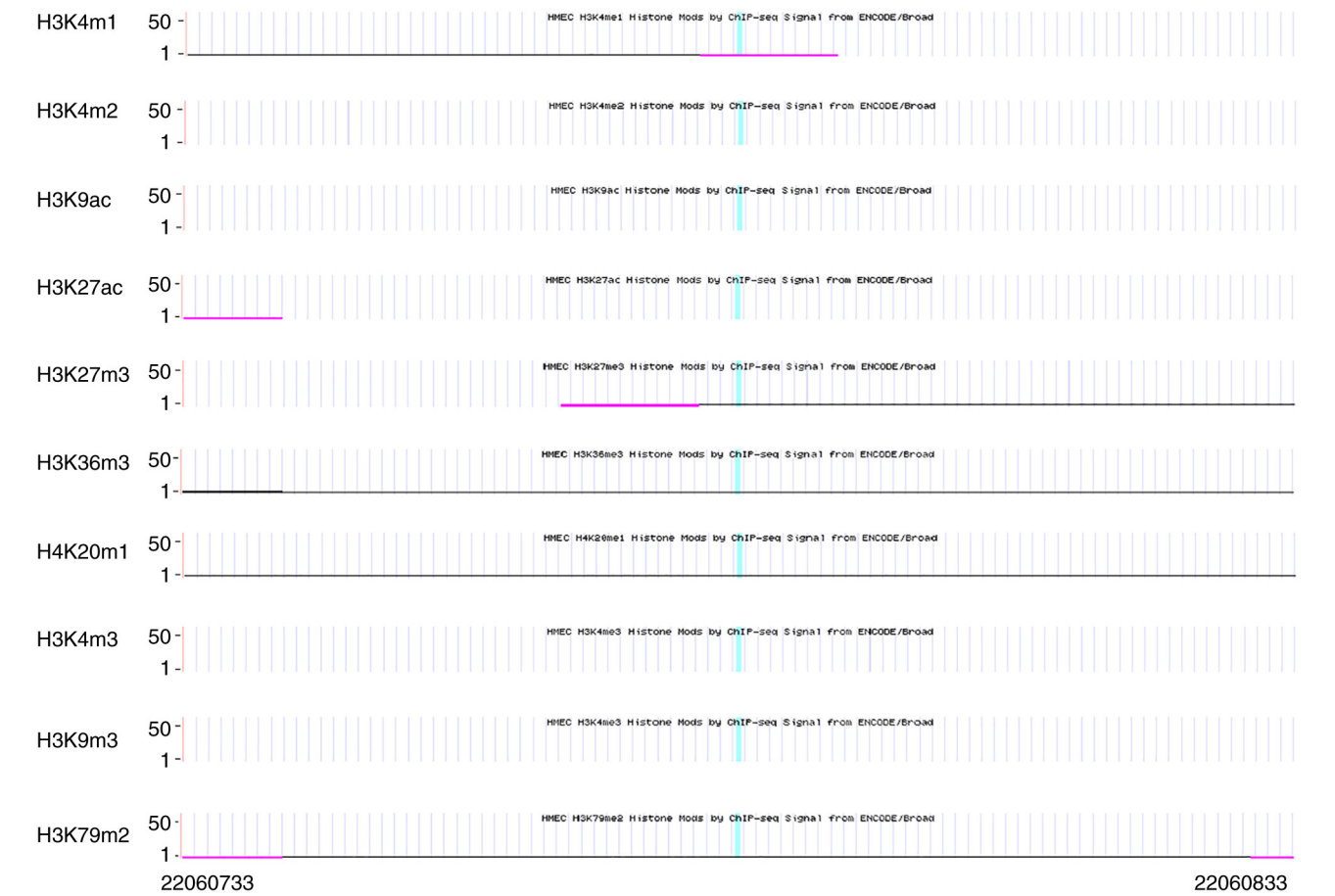


Figure S4. Results of chromatin immunoprecipitation positive controls for (A) REL ( $P=0.0049$ ), (B) ELK1 ( $P=0.037$ ), (C) POU2F1 ( $P=0.00052$ ) or (D) PAX6 ( $P=0.00050$ ) antibodies in MCF-7 cells. The data are presented as mean  $\pm$  standard deviation. \* $P<0.05$ . REL, REL proto-oncogene; ELK1, transcription factor ELK1; POU2F1, POU class 2 homeobox 1; PAX6, paired box 6.

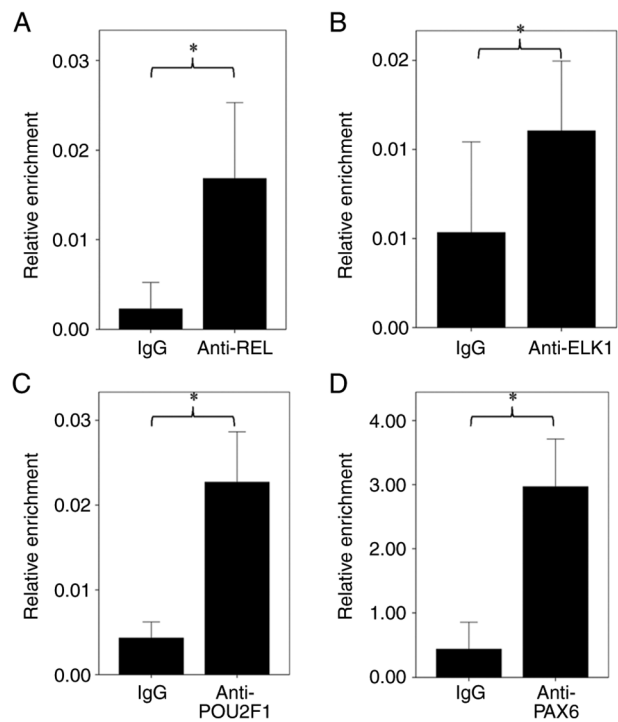


Figure S5. Chromatin enrichment of the region surrounding rs77283072 by (A) REL (P=0.51), (B) ELK1 (P=0.50), (C) POU2F1 (P=0.051) or (D) PAX6 (P=0.25) antibodies in MCF-7 cells. The data are presented as mean  $\pm$  standard deviation. REL, REL proto-oncogene; ELK1, transcription factor ELK1; POU2F1, POU class 2 homeobox 1; PAX6, paired box 6; ns, not significant.

