

























Supplemental Figure Legends

Figure S1 Associations of 1,732 SNPs conditional on rs1830298. The most significantly associated SNP after conditioning on rs1830298 is indicated. The upper panel plots SNPs based on their chromosomal coordinates on the x axis and their p values on the $-\log_{10}$ scale on the y axis. Circle and diamond symbols represent typed and imputed SNPs respectively. The colours indicate the pairwise r^2 with the best hit (highlighted in purple); r^2 are calculated based on the European panel in the 1000 genomes project. Genes within the region are indicated in the lower panel, with arrows indicating transcript direction, dense blocks for exons and lines for introns. The highlighted region contains the SNPs with p values $\leq 1 \times 10^{-4}$. The plot was generated using LocusZoom¹⁷.

Figure S2. Associations of family history, tumour stage, tumour grade, tumour size and lymph node status by SNP.

Squares denote the individual study per-allele OR with the size of the symbol indicating sample size. Notation: famHisPos indicates at least one first degree relative with breast cancer, famHisNeg indicates no first degree relatives with breast cancer. Tumours were grouped into stages 2,3,4 or stages 0,1. Tumours were grouped by grade into grade 3 or grades 2 and 1). sizeCat 23 indicates tumour size ≥ 2 cm, sizeCat1 indicates < 2 cm. NodePos indicates the presence of positive lymph nodes and NodeNeg indicates no positive lymph nodes. Cases in each of the above groups were compared to all controls.

Figure S3. Associations of ER positive, ER negative, triple negative, invasive and non-invasive breast cancer by SNP.

Squares denote the individual study per-allele OR with the size of the symbol indicating sample size. Triple negative indicates the absence of tumour receptors for ER, PR and HER2. Cases in each subtype group were compared to all controls.