Supplemental figures legend

Supplemental Figure 1. Power and tagging analysis.

a) Power to detect association for varying genotype relative risk (GRR; 1.2-1.6) and risk allele frequencies (RAF; 0.1-0.9) at significance levels defined by a Bonferroni correction for the number of SNPs tested in the candidate gene regions (0.0004 for *FLJ22536*; 0.0016 for *LMO1*; 0.0015 for *BARD1* (not shown) was identical to power for *LMO1*). b) Percentage of HapMap Phase II SNPs in the three candidate gene regions in the YRI and CEU populations tagged by the genotyped SNPs included in the analysis.

Supplemental Figure 2. MDS plots and distribution of African ancestry.

Multidimensional scaling plots of cases (a) and controls (b) against the four major HapMap populations. C1: first MDS component (X axis); C2: second MDS component (Y axis).

c) Distribution of proportion of African ancestry in cases and controls.

Supplemental Figure 3. Regional association plots for *BARD1* SNPs.

a) Negative $\log_{10}$ p-values from logistic regression analysis for all genotyped SNPs tested in the *BARD1* region plotted against their genomic location. b) Negative $\log_{10}$ p-values from logistic regression analysis conditional on association at rs7587476.

Color shading for r^2 is relative to the most significant SNP (rs7587476) based on data from the YRI population. Blue lines are estimates of recombination rates (cM/Mb) (right Y axis).

Supplemental Figure 4. Regional LD plots in *BARD1* genomic region.

r^2 values relative to rs7587476 in the YRI (a) and CEU (b) populations for SNPs in a 200Kb window around it plotted against their genomic location. Data are from the 1000 Genomes Project Pilot 1. Dotted vertical lines delimit regions including SNPs with $r^2>0.5$. Size and color

intensity of markers are proportional to r^2 with rs7587476. Blue lines are estimates of recombination rates (cM/Mb) (right Y axis).

Supplemental Figure 5. Regional association plots after imputation.

Negative $\log_{10}$ p-values from SNPTEST frequentist test of association under the additive model for all imputed and genotyped SNPs in the *BARD1* (a), *FLJ22536* (b), and *LMO1* (c) gene regions plotted against their genomic location. In each plot, the most significant SNP is represented by a red diamond. All other SNPs are represented by diamonds if genotyped, and squares if imputed. For *BARD1*, color shading for r^2 is relative to the most significant SNP (rs35953323) based on data from the YRI population from the 1000 Genomes Project. No LD data for the two most significant *FLJ22536* and *LMO1* SNPs are available. Blue lines are estimates of recombination rates (cM/Mb) (right Y axis).