Supplementary figure 1
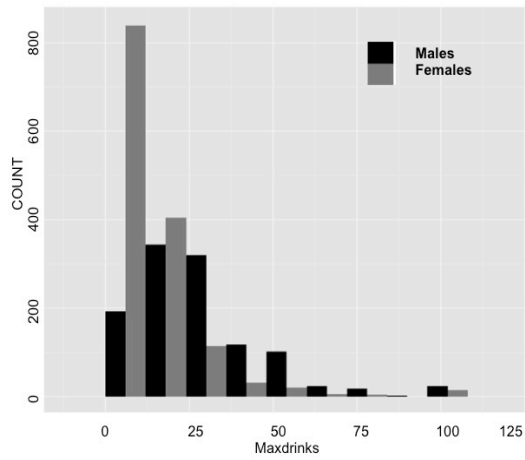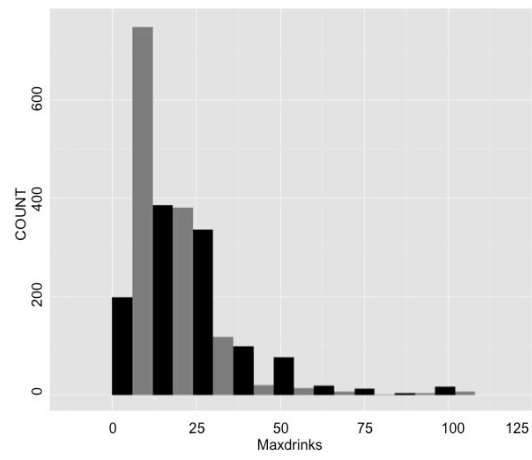
Distribution of maxdrinks and ln (maxdrinks) among males and females in COGA and SAGE

A & B shows maxdrinks and ln (maxdrinks) distribution in COGA C & D shows distribution maxdrinks and ln (maxdrinks) distribution in SAGE dataset. Black and grey color bars represent distributions in males and females respectively.
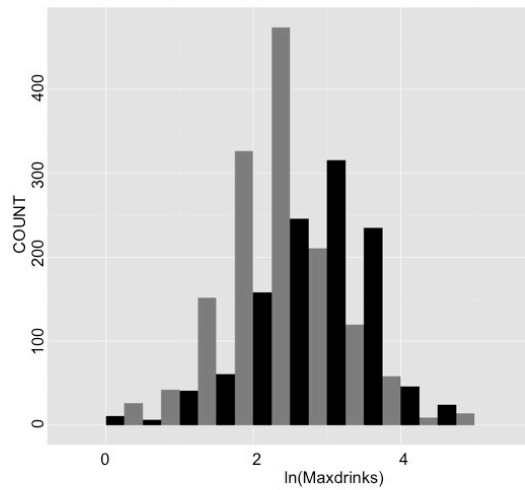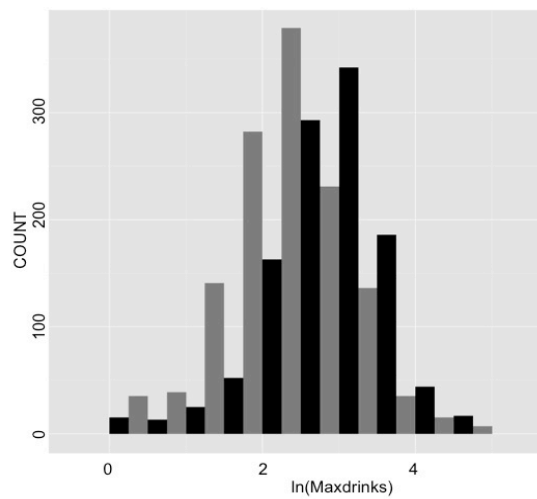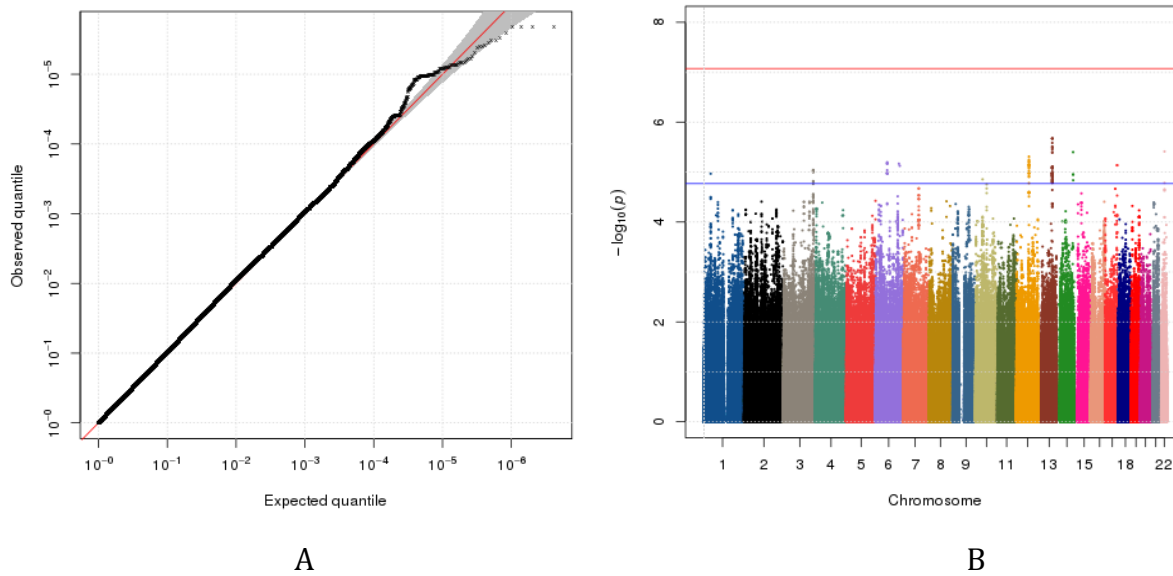
Supplementary figure 2

**QQ and Manhattan plot for the genome-wide association analysis in COGA dataset**

A   B

(A) Q–Q plot for association analysis of the alcohol dependence phenotype. Observed l p values for the 4,058,415 SNPs (black dots) and expected p values (red line) are plotted against the expected p value (x-axis). The genomic inflation factor value (lambda) was 0.99.

(B) Observed log p values for the 4,058,415 SNPs were plotted for each of the 22 autosomal chromosome in Manhattan plot

**Supplementary figure 3**

**QQ Manhattan plots for the genome-wide analysis of maxdrinks phenotype in SAGE dataset**

A   B
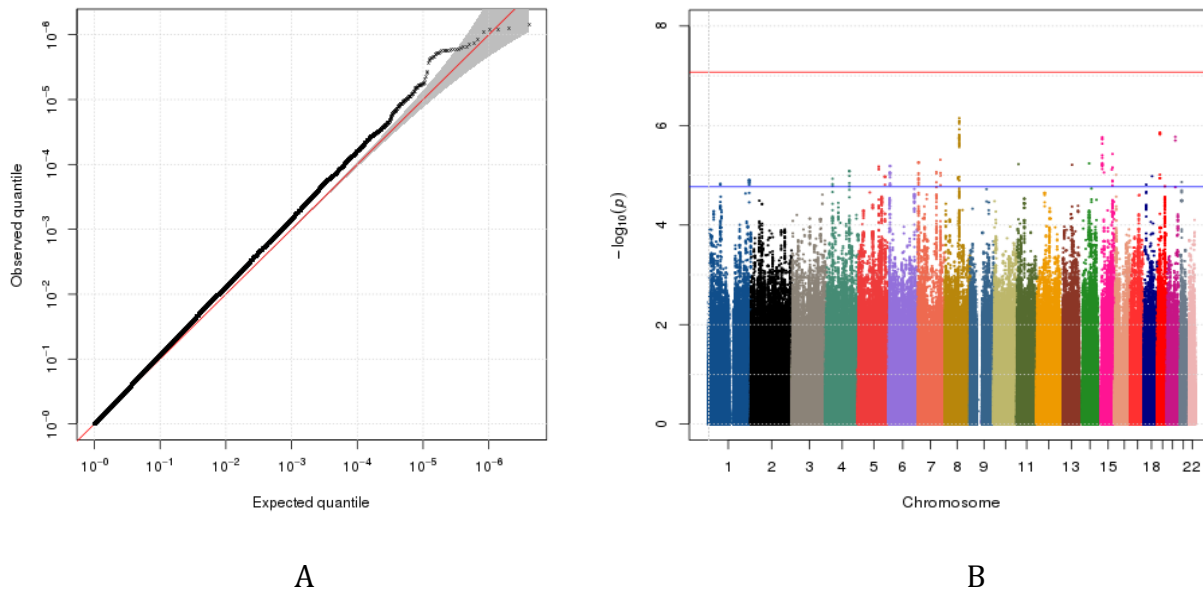
(A) Q–Q plot for association analysis of the maxdrinks phenotype. Observed l p values for the 4,058,415 SNPs (black dots) and expected p values (red line) are plotted against the expected p value (x-axis). The genomic inflation factor value (lambda) was 1.07.

(B) Observed log p values for the 4,058,415 SNPs were plotted for each of the 22 autosomal chromosome in Manhattan plot
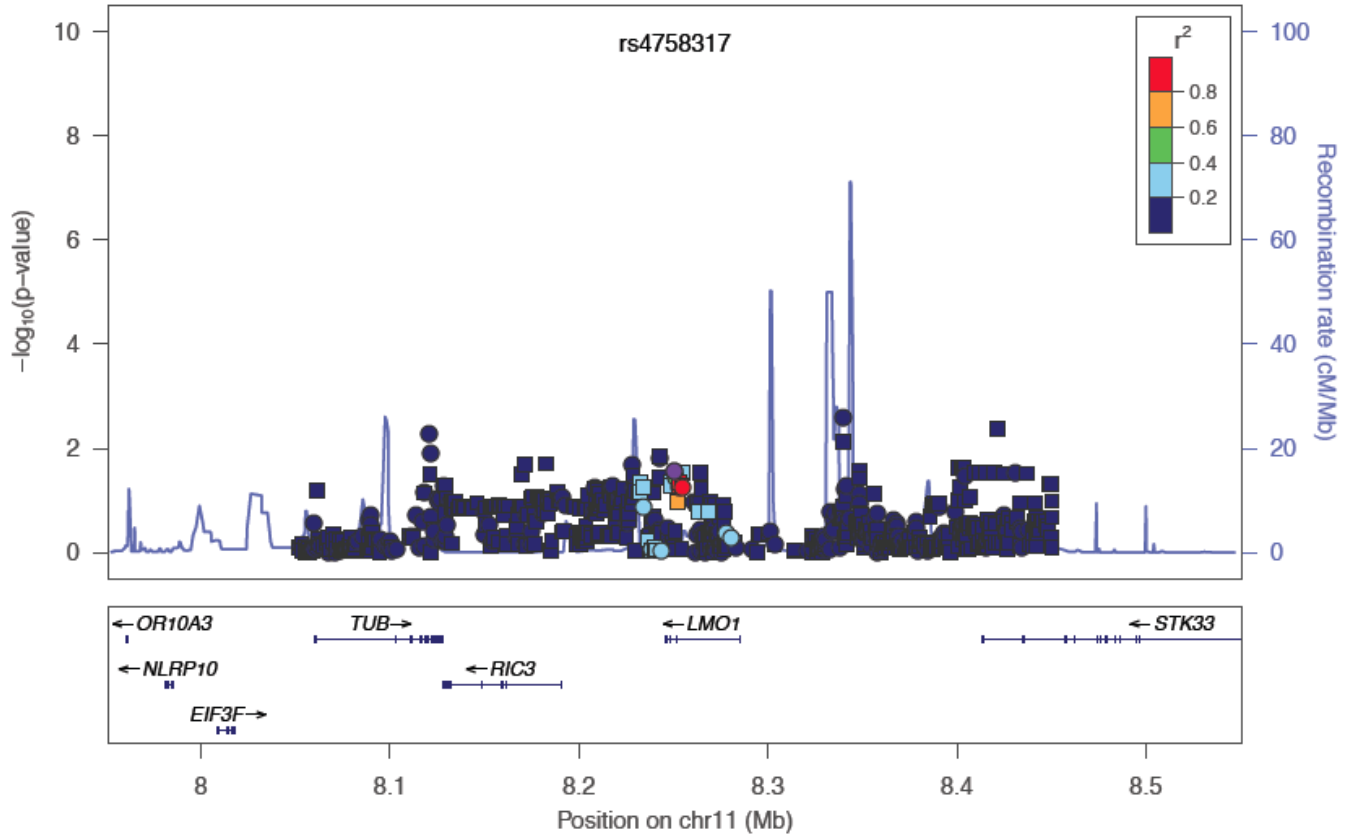
Supplementary Figure 4

Regional plot for the variants identified in meta-analysis near in and around *LMO1* gene.
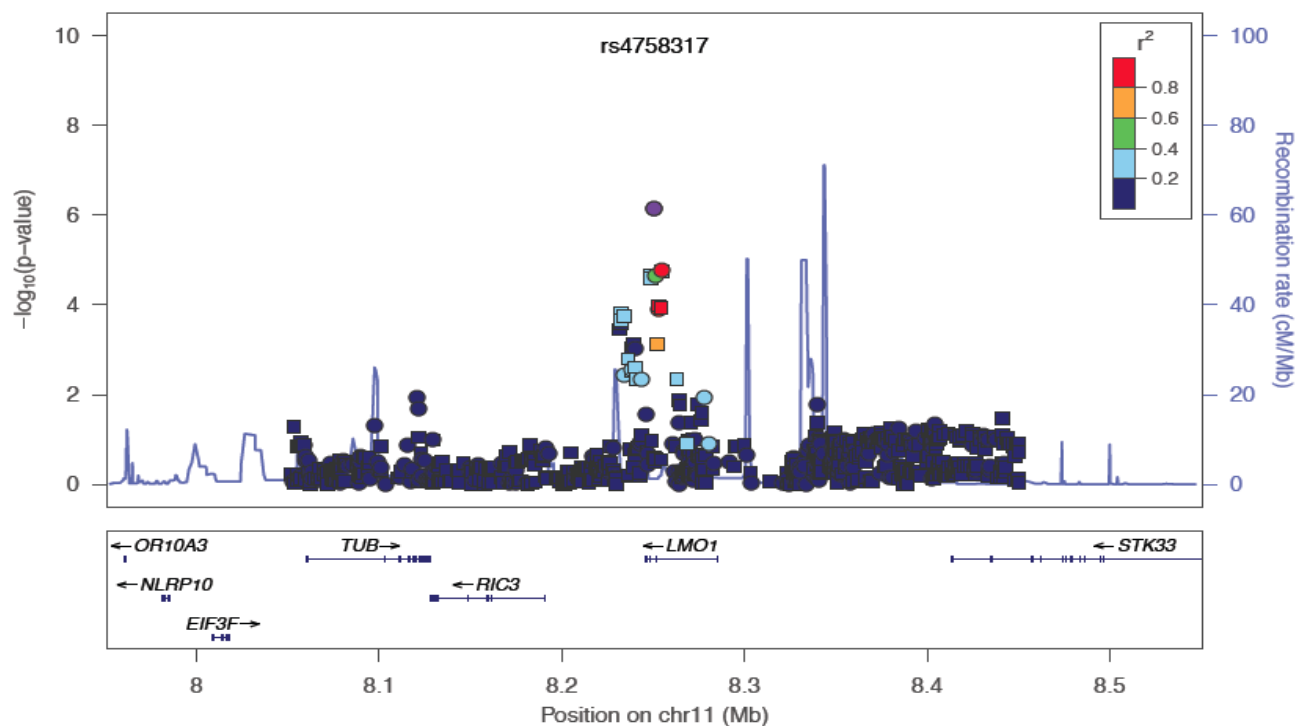
A
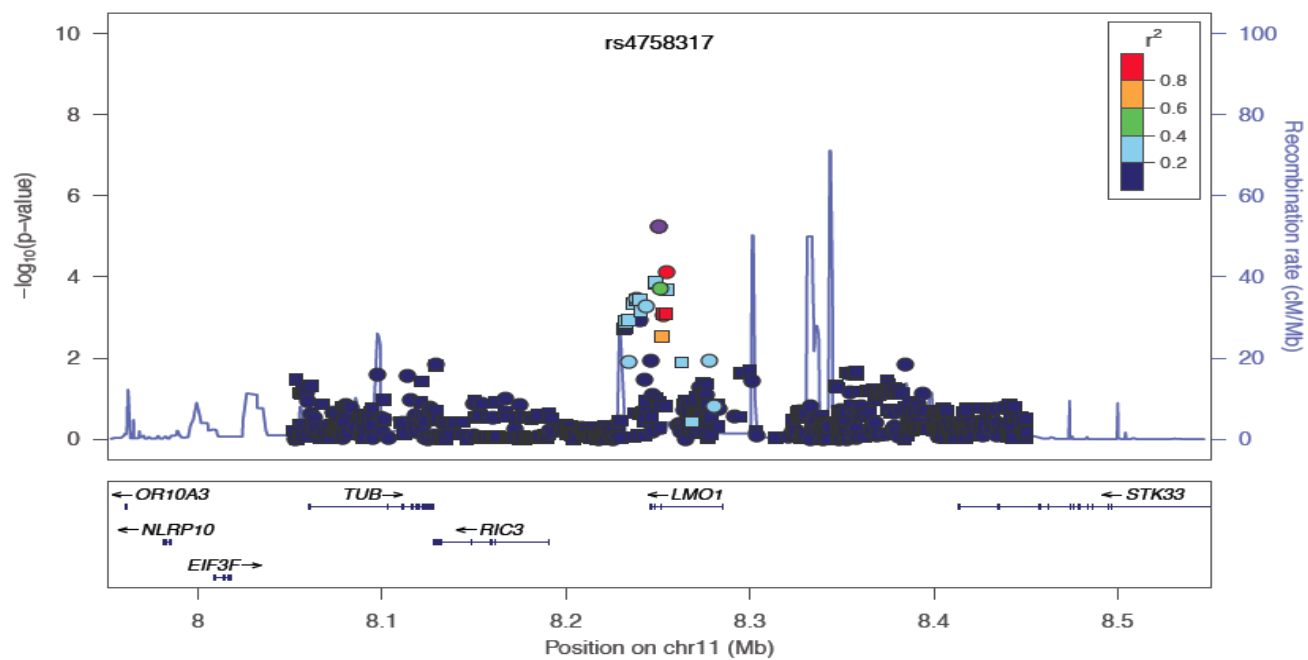
B
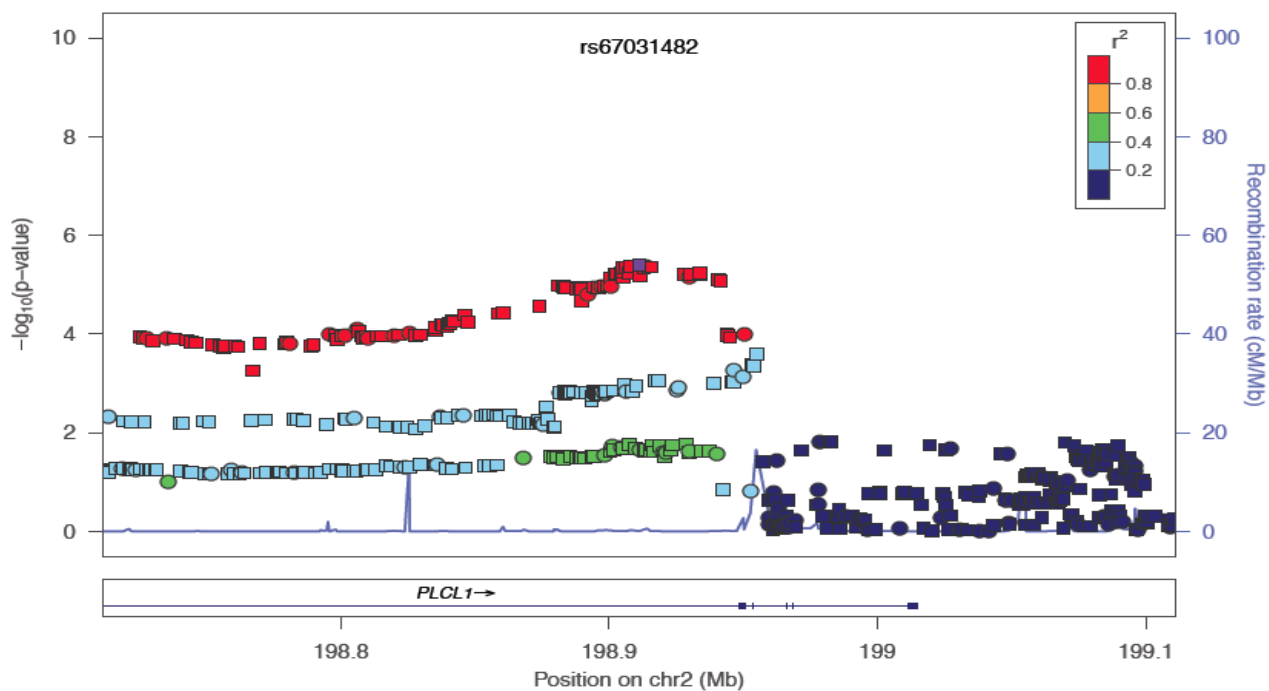
C

Y axis represent –log10 P values in (A) Meta-analysis (B) SAGE and (c) COGA datasets. Square blue dot represent the strongest signal identified in meta-analysis. The most significantly associated SNP is shown in purple. The extent of linkage disequilibrium (as measured by r2) between each SNP and the most significantly associated SNP in meta-analysis is indicated by the color scale at top right. Larger values of r2 indicate greater linkage disequilibrium. Circles represent the genotyped SNPs and squares represent the imputed SNPs.

Supplementary Figure 5

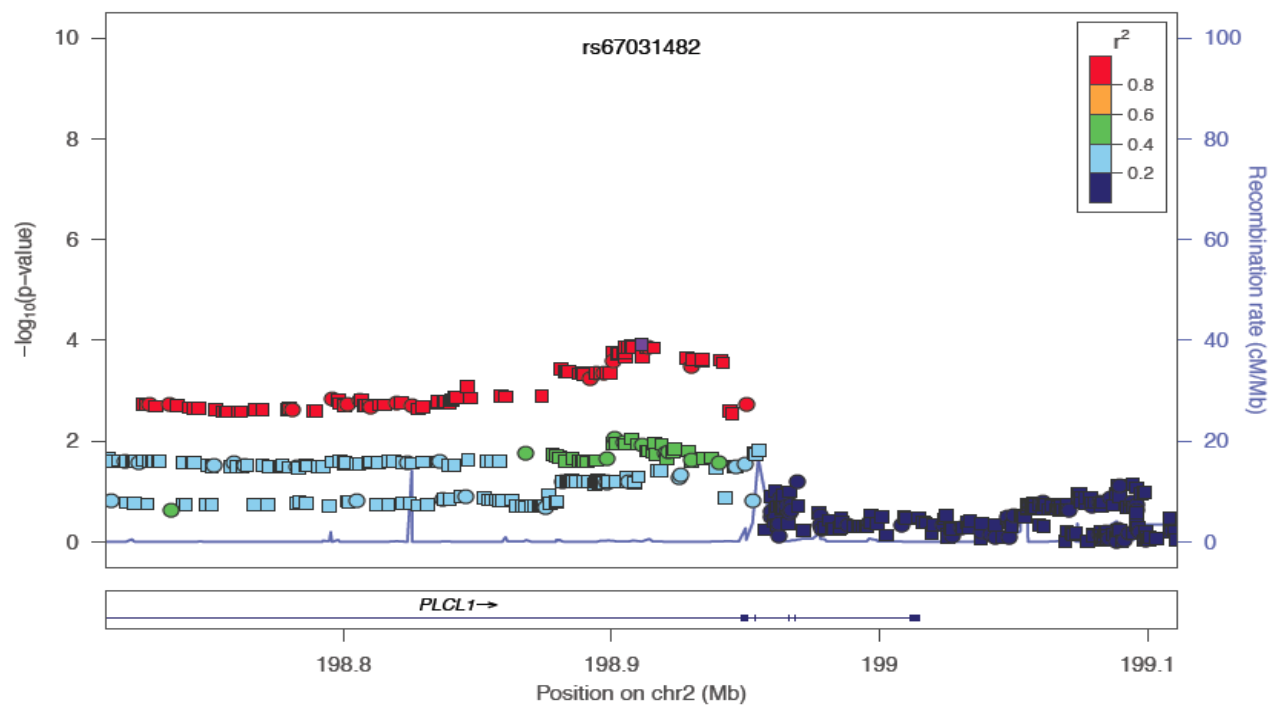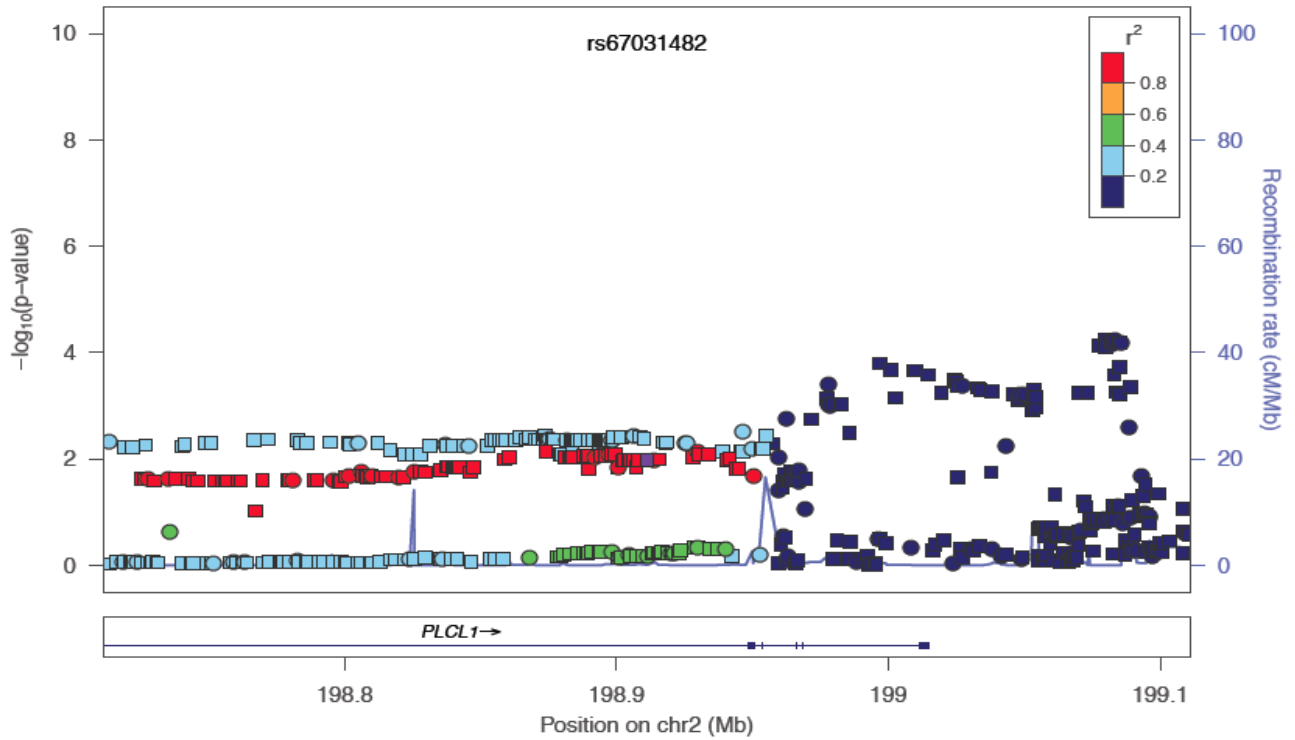## Regional plot for the variants identified in meta-analysis in and around *PLCL1* gene

A

B

C

Y axis represent –log10 P values in (A) Meta-analysis (B) SAGE and (c) COGA datasets. Square blue dot represent the strongest signal identified in meta-analysis. (A) Meta-analysis (B) SAGE and (c) COGA. Square blue dot represent the strongest signal identified in meta-analysis. The most significantly associated SNP in meta-analysis is shown in purple. The extent of linkage disequilibrium (as measured by r2) between each SNP and the most significantly associated SNP is indicated by the color scale at top right. Larger values of r2 indicate greater linkage disequilibrium. Circles represent the genotyped SNPs and squares represent the imputed SNPs.