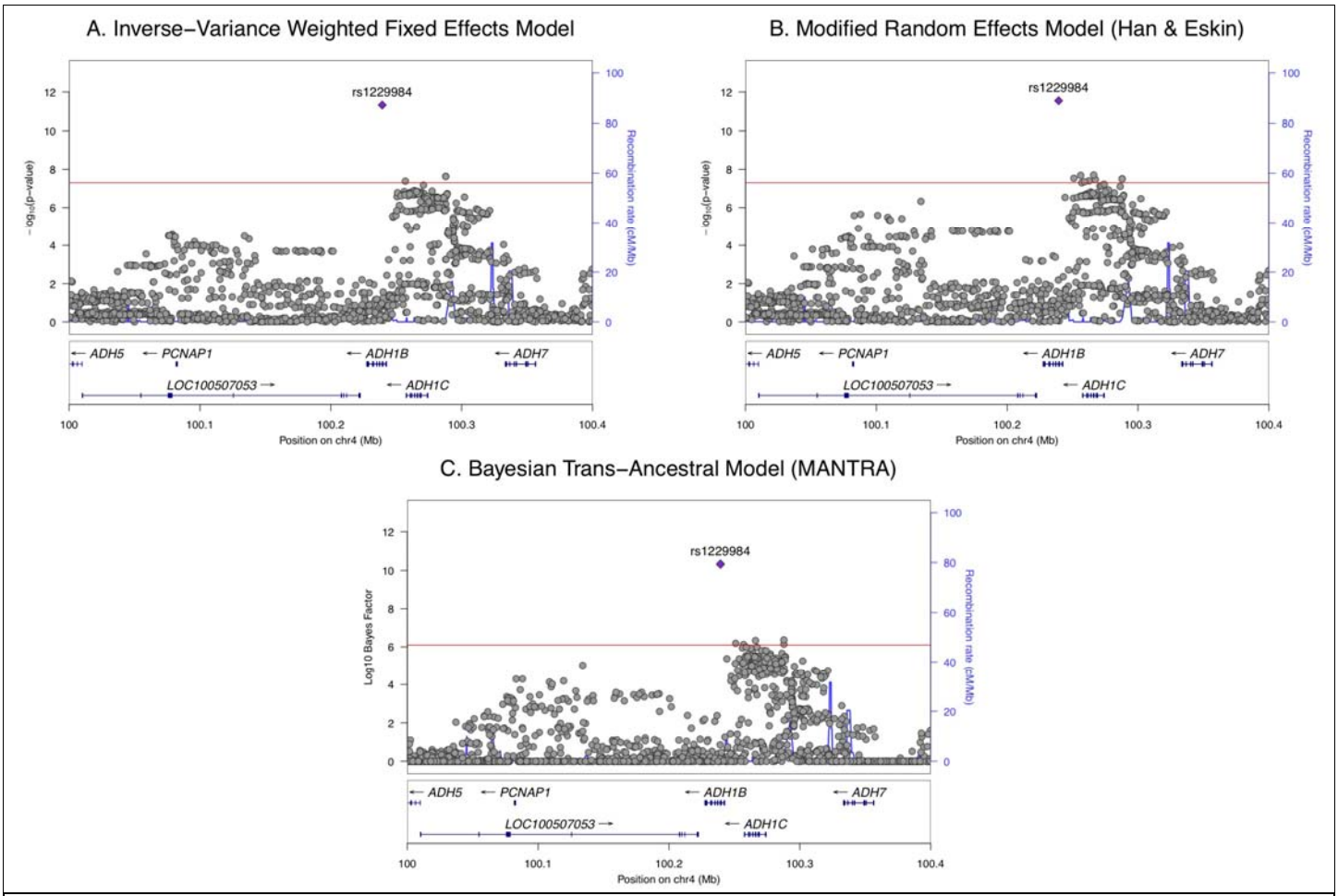
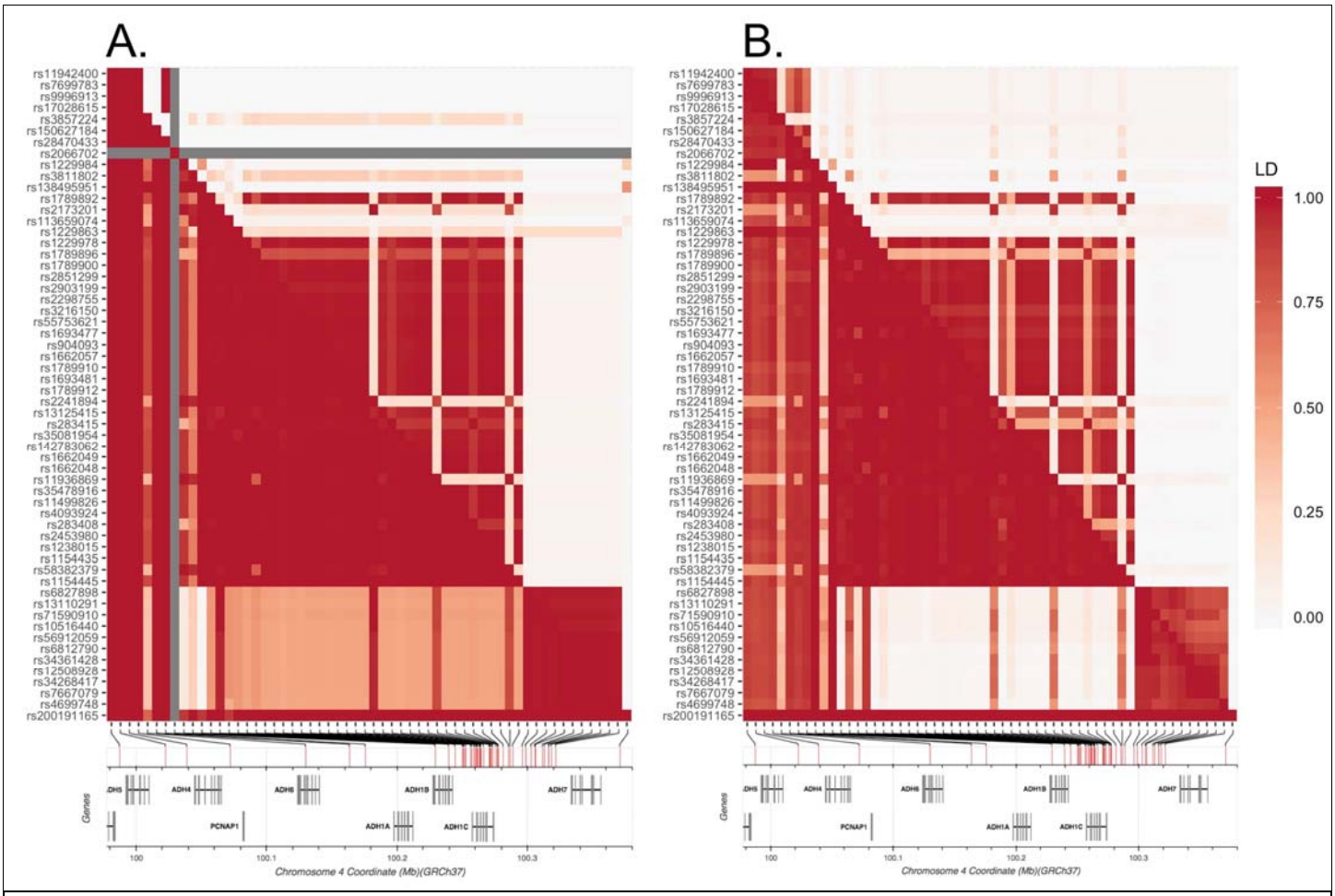


Supplementary Figure 1

Results for *ADH* locus from transancestral modeling of unrelated individuals

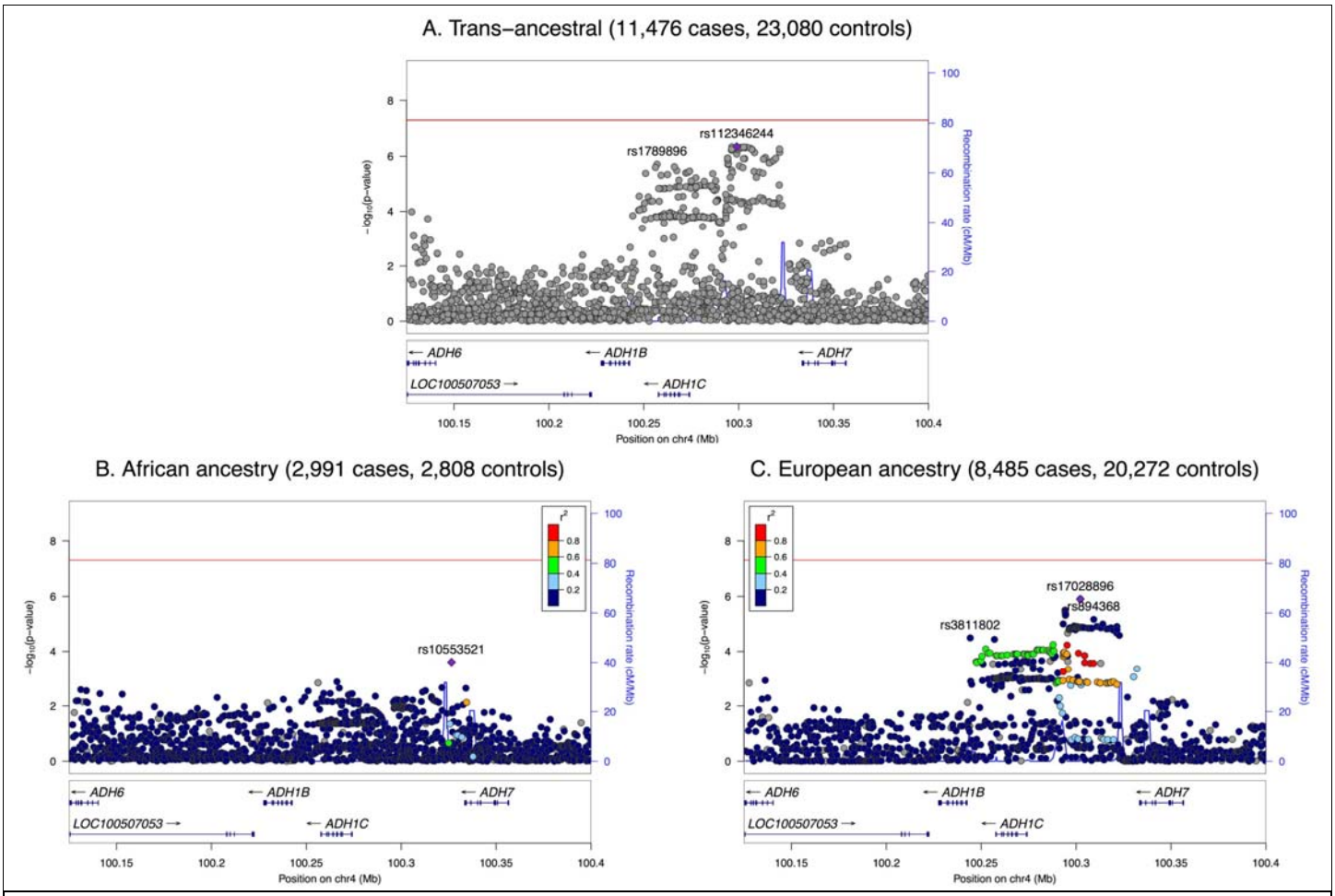
Regional association plot for the *ADH* gene region in trans-ancestral meta-analysis of unrelated genotyped individuals from AA and EU cohorts (n_{case}=11,476, $n_{control}$=23,080). Association with AD across ancestries was evaluated using (A) an inverse-variance weighted fixed-effects model, (B) the modified random-effects model[92], and (C) the Bayesian trans-ancestral model[93]. The Bayesian model is fit with default priors and Metropolis-Hastings Markov chain Monte Carlo algorithm in MANTRA. Due to LD differences across ancestral groups combined in the trans-ancestral models LD is not indicated. The fixed-effects and random-effects models report conventional p-values, while the Bayesian model reports the Bayes factor for comparison of the null and alternative hypotheses. A $\log_{10}$ Bayes factor > 6.1 roughly corresponds the $P < 5 \times 10^{-8}$ significance threshold[94]. Plots generated with LocusZoom[170] (http://locuszoom.sph.umich.edu/). See Supplementary Information for references.



Supplementary Figure 2

LD structure of *ADH* locus

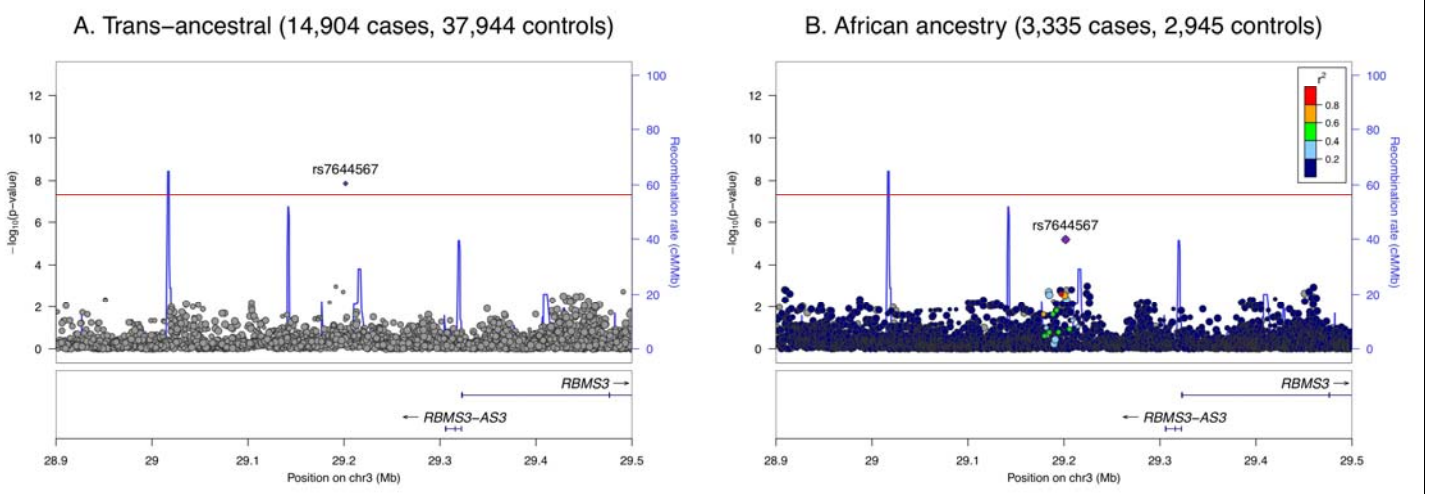
Linkage disequilibrium (LD) among top variants in the *ADH* gene region in (A) European ancestry and (B) African ancestry populations from the 1000 Genomes Project. Within each plot, the upper triangle displays Pearson correlation (r^2) for each pair of markers, and the lower triangle displays *D*'. Missing values are indicated in gray. LD is reported for variants in the region with two-tailed $P < 1 \times 10^{-7}$ in the full discovery meta-analysis weighted by effective sample size (14,904 individuals with AD, 37,944 controls). Variants that are effectively perfectly correlated ($r > 0.995$) in both European and African populations are thinned to improve legibility. Plot generated with the assistance of LD Link[97] (https://ldlink.nci.nih.gov/). See Supplementary Information for references.



Supplementary Figure 3

Locus plots of conditional analysis of *ADH1B* region

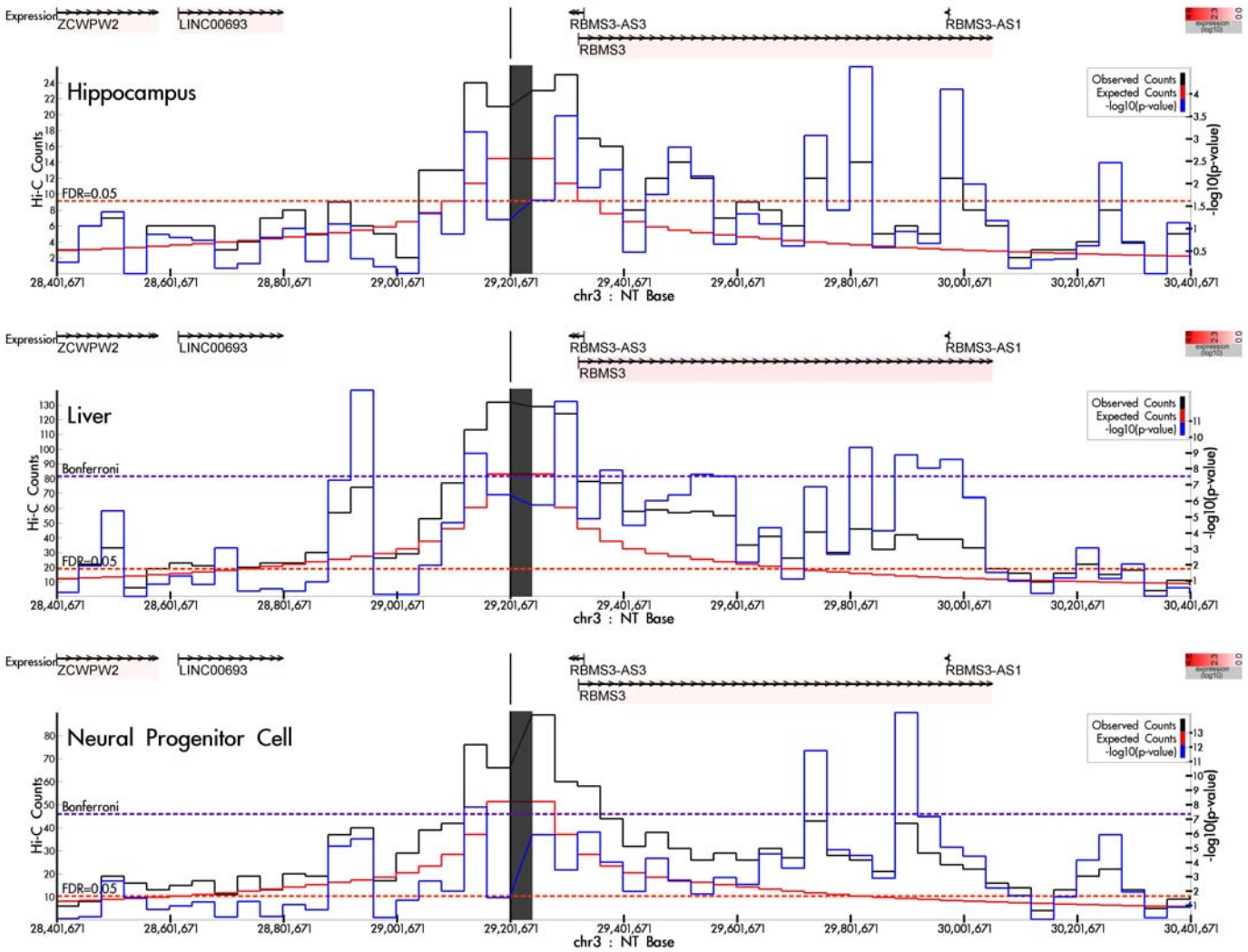
Regional association plot for the *ADH* gene region in inverse-variance weighted fixed effects meta-analysis of logistic regression of unrelated genotyped individuals within each cohort conditional on genotype for top *ADH1B* associations. AA cohorts where conditioned on rs2066702 and, where possible, rs1229984; EU cohorts were conditioned on rs1229984. Results are shown for (A) meta-analysis across EU and AA cohorts (n_{case}=11,476, $n_{control}$=23,080), (B) meta-analysis of the AA cohorts only (n_{case}=2,991, $n_{control}$=2,808), and (C) meta-analysis of the EU cohorts (n_{case}=8,485, $n_{control}$=20,272) only. The red reference line indicates the two-tailed $P < 5 \times 10^{-8}$ threshold for genome-wide significance within each analysis. Colored points indicate LD to the index variant in individuals of (B) African ancestry or (C) European ancestry, respectively, from the 1000 Genomes Project reference data[69]. LD information is not shown for the trans-ancestral results (A) because the results reflect a combination of European and African ancestries and thus don't have a well-defined population LD reference panel. Plot generated with LocusZoom[170] (http://locuszoom.sph.umich.edu/). See Supplementary Information for references.



Supplementary Figure 4

Locus plots of chromosome 3 locus (rs7644567)

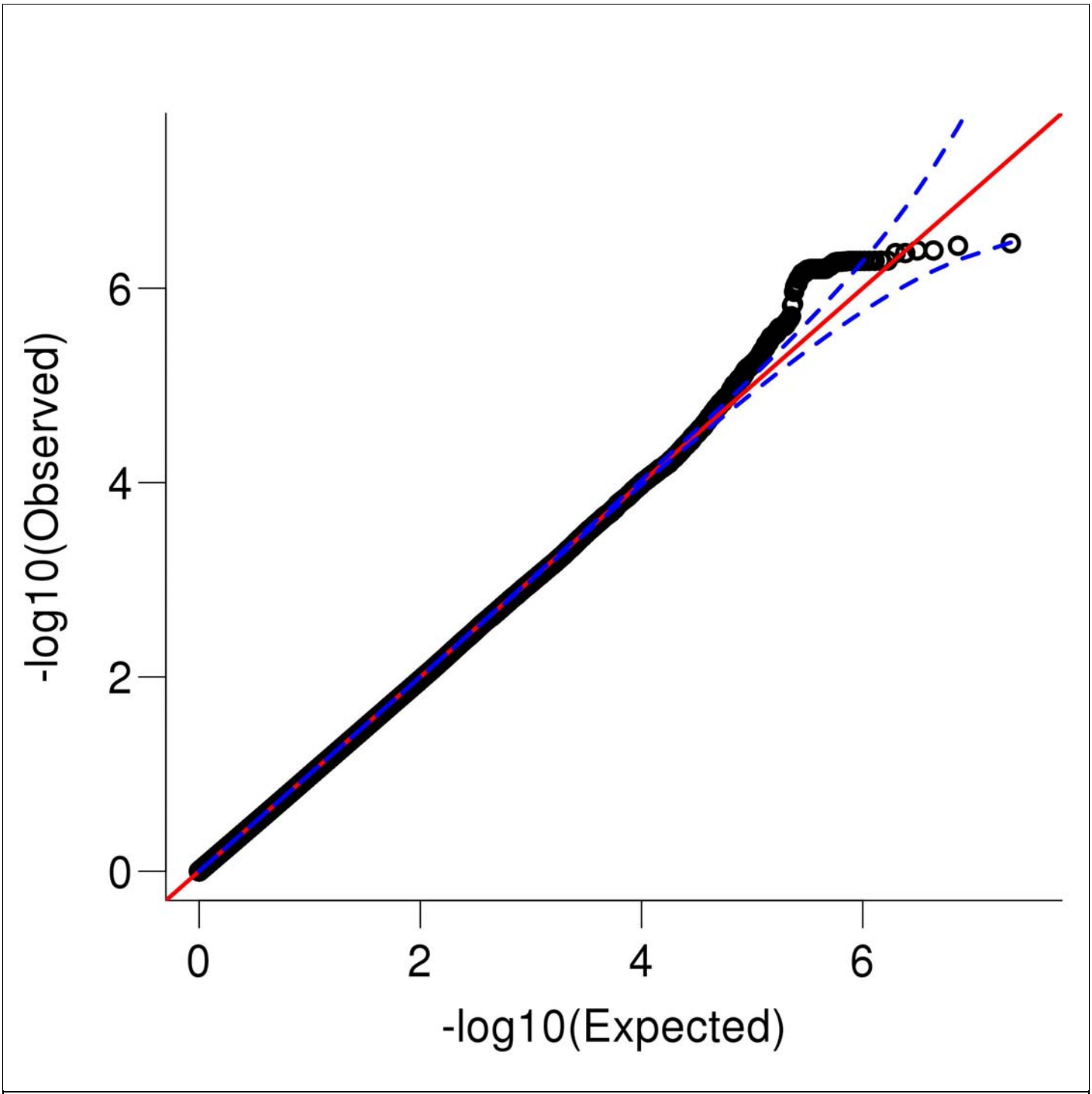
Regional association plot for rs7644567 in (A) the discovery meta-analysis of AA and EU cohorts (n_{case}=14,904; $n_{control}$=37,944) and (B) the AA discovery meta-analysis (n_{case}=3,335; $n_{control}$=2,945) under a fixed effects model with effective sample size weighting. Red reference line indicates the genome-wide significance threshold of two-tailed $P < 5 \times 10^{-8}$. Colored points in panel B indicate LD to the index variant in individuals of African ancestry in the 1000 Genomes Project reference data[69]. LD information is not plotted in panel A due to the lack of the population LD reference panel for the combined European and African ancestries. Plot generated with LocusZoom[170] (http://locuszoom.sph.umich.edu/). See Supplementary Information for references.



Supplementary Figure 5

Regions of chromatin contact for rs7644567

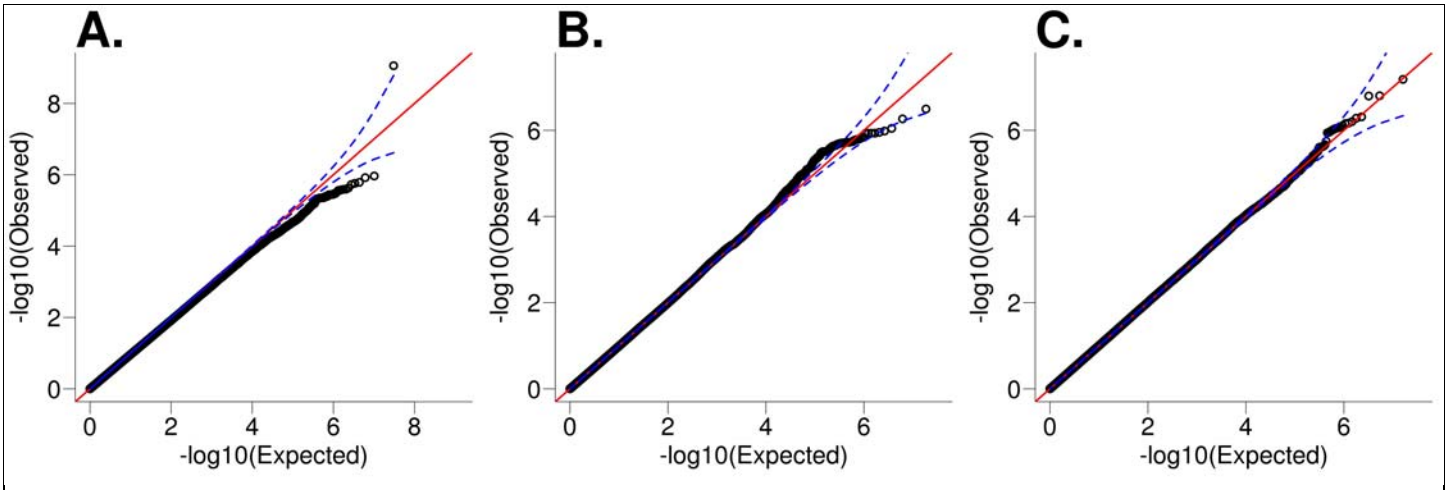
HUGIn[101] results identifying chromatin contacts with the region containing rs7644567 on chromosome 3. Blue lines reflect –log(p-value) for the one-tailed test, using Fit-Hi-C[100], of whether observed Hi-C counts (black lines) were greater than the expected number of Hi-C counts (solid red lines) in previously reported Hi-C data[99]. Note differences in scale of the plots for the 3 tissue types. Tissue-specific False Discovery Rate (FDR, dashed red) and Bonferroni (dashed purple) corrections are shown for each tissue type. See Supplementary Information for references.



Supplementary Figure 6

QQ plot for omnibus test of heterogeneity across all cohorts

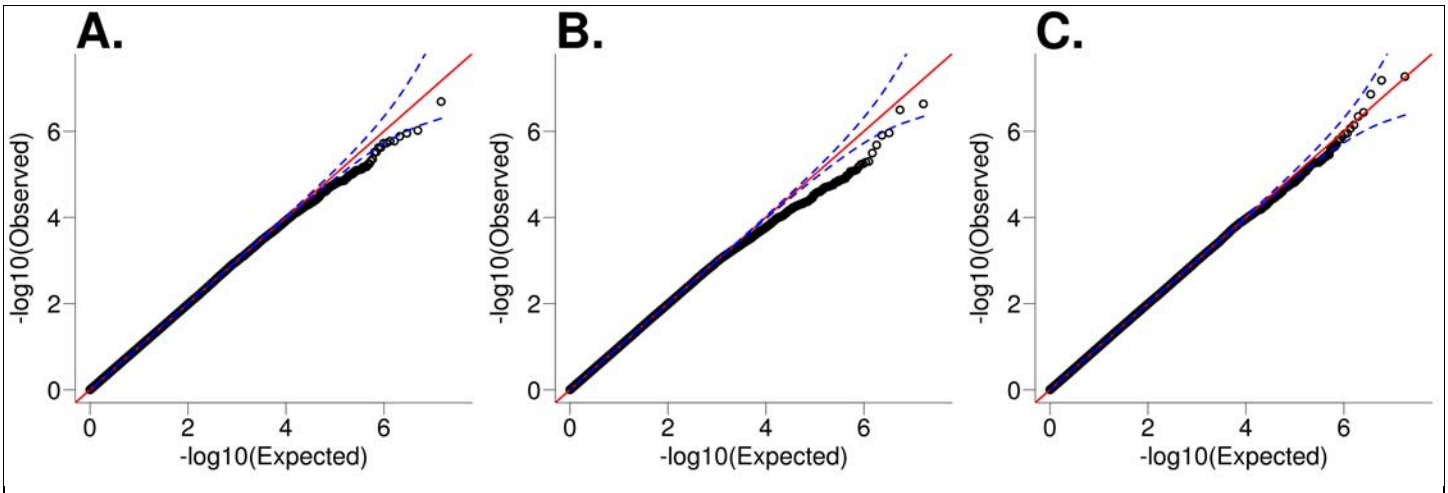
QQ plot of p-values for the omnibus (34 degree of freedom) test of heterogeneity across all AA and EU cohorts in the discovery meta-analysis (n_{case}=14,904; $n_{control}$=37,944). Heterogeneity is tested with respect to a fixed effects model for meta-analysis of p-values with effective sample size-based weights. Little deviation is observed from the expected null distribution, suggesting limited heterogeneity across cohorts.



Supplementary Figure 7

QQ plots for tests of heterogeneity within and between ancestries

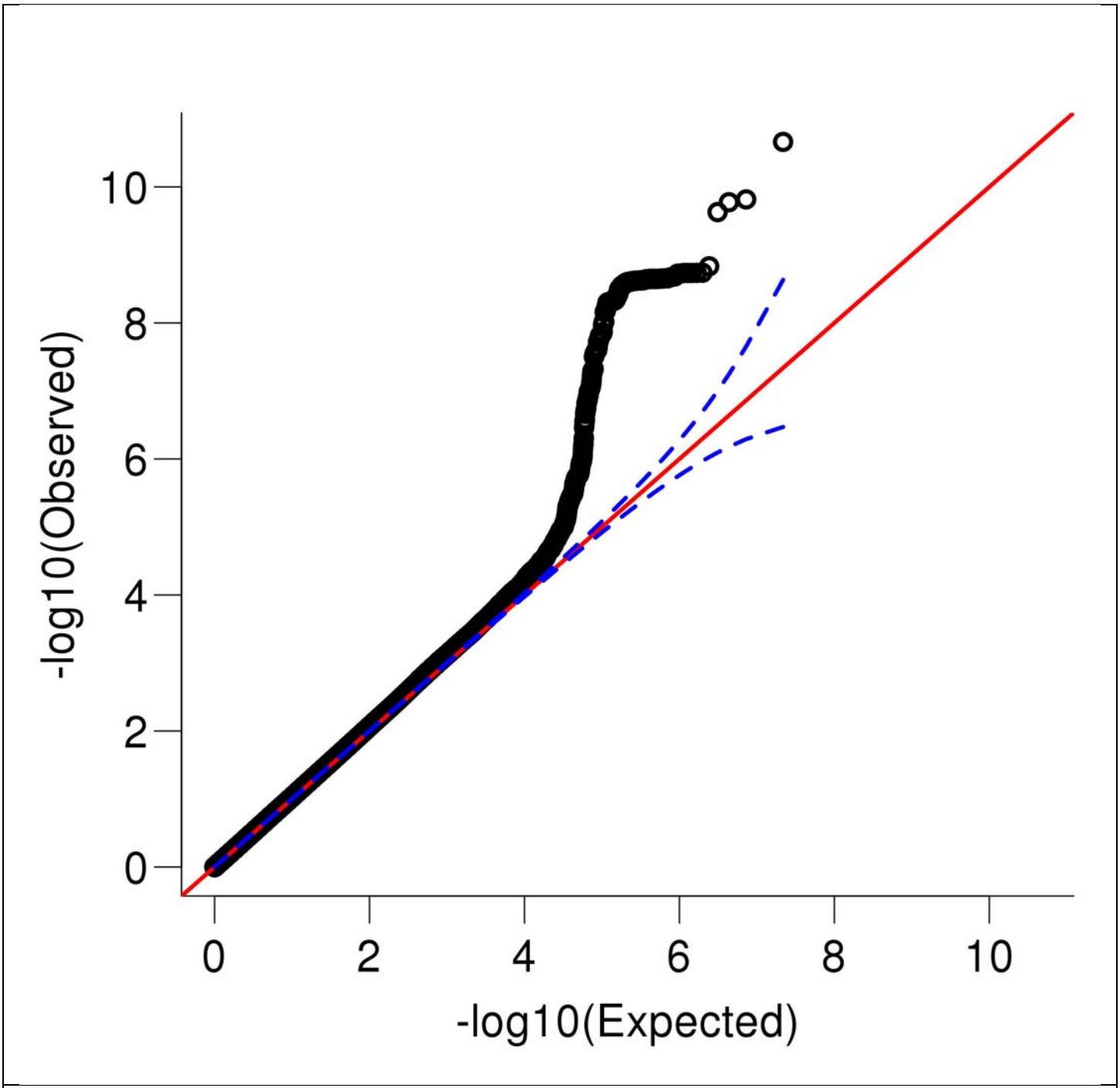
QQ plot of p-values for (A) the omnibus (7 degree of freedom) test of heterogeneity across all AA cohorts in the discovery meta-analysis (n_{case}=3,335; $n_{control}$=2,945); (B) the omnibus (26 degree of freedom) test of heterogeneity across all EU cohorts in the discovery meta-analysis (n_{case}=11,569; $n_{control}$=34,999); and (C) the 1 degree of freedom test of heterogeneity between EU cohorts (n_{case}=11,569, $n_{control}$=34,999) and AA cohorts (n_{case}=3,335, $n_{control}$=2,945). Heterogeneity is tested with respect to a fixed effects model for meta-analysis of p-values with effective sample size-based weights. Little deviation is observed from the expected null distribution, suggesting limited overall heterogeneity within or between ancestry. The exception is one variant (rs4673609) with genome-wide significant heterogeneity ($P = 8.78 \times 10^{-10}$) among AA cohorts.



Supplementary Figure 8

QQ plots for tests of heterogeneity between study designs

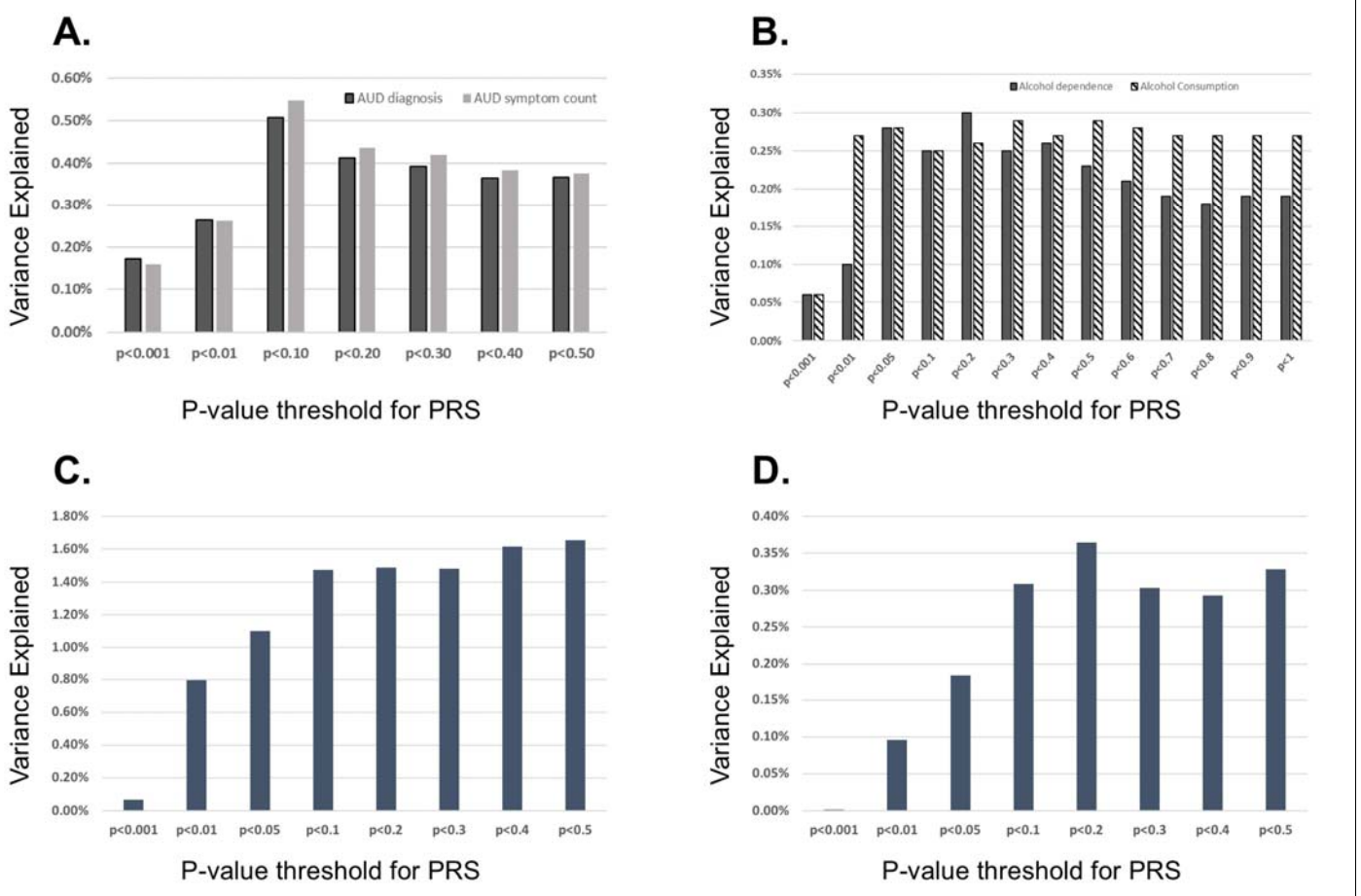
QQ plot of p-values for (A) the 1 degree of freedom test of heterogeneity between simple family-based EU cohorts tested using the GEE model (n_{case}=2,107, $n_{control}$=12,353) and complex family-based EU cohorts tested using the logistic mixed model (n_{case}=2,897, $n_{control}$=5,565); (B) the 1 degree of freedom test of heterogeneity between genotyped family-based EU cohorts (n_{case}=5,004, $n_{control}$=17,918) and genotyped unrelated case/control EU cohorts (n_{case}=4,844, $n_{control}$=8,873); and (C) the 1 degree of freedom test of heterogeneity between genotyped EU cohorts (n_{case}=9,848, $n_{control}$=26,791) and EU cohorts included with summary statistics only (n_{case}=1,721, $n_{control}$=8,208). Heterogeneity is tested with respect to a fixed effects model for meta-analysis of p-values with effective sample size-based weights. Little deviation is observed from the expected null distribution, suggesting limited heterogeneity between study designs.



Supplementary Figure 9

QQ plot for discovery meta-analysis of AD

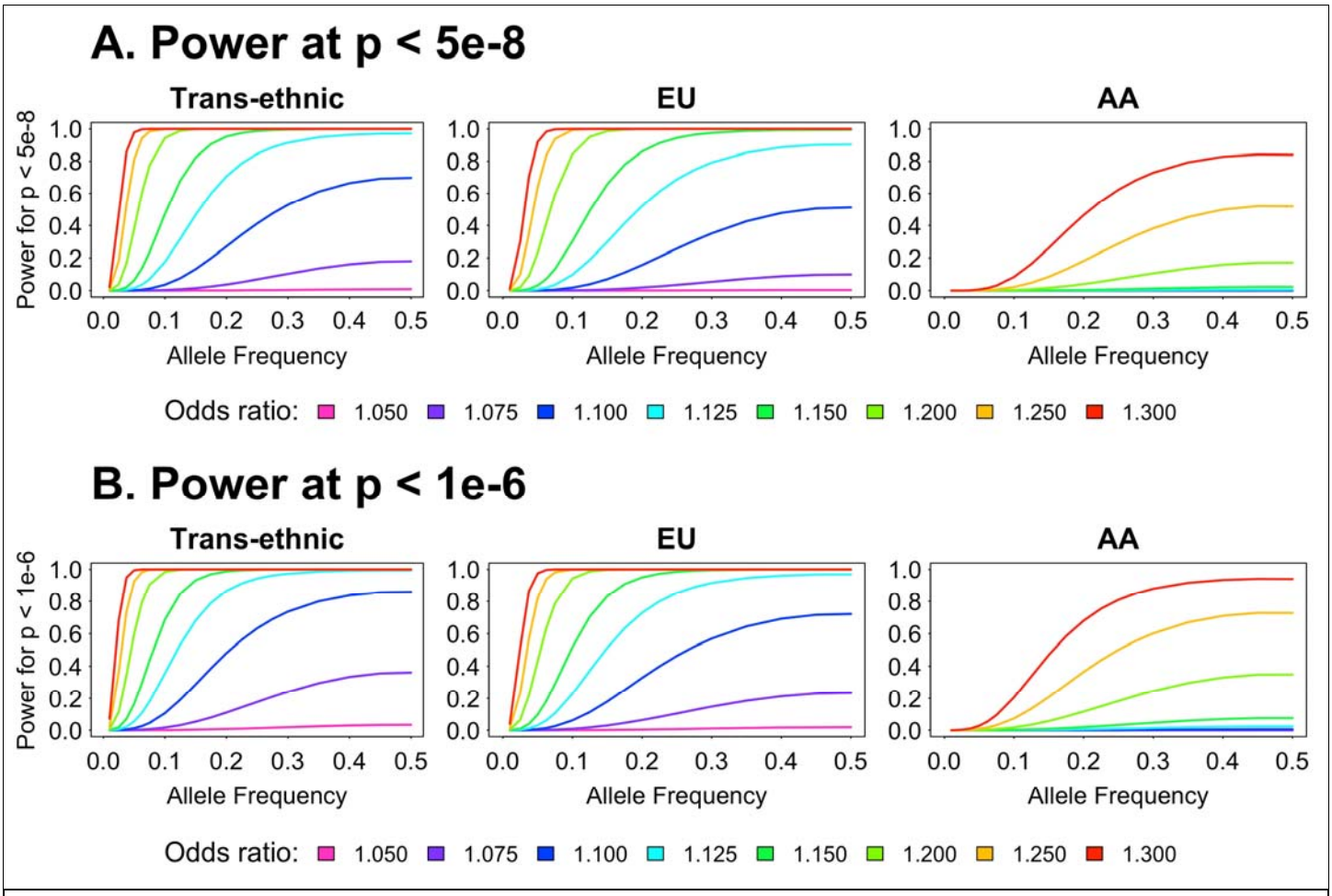
QQ plot of two-tailed p-values for association with AD in the discovery meta-analysis of AA and EU cohorts (n_{case}=14,904, $n_{control}$=37,944). Meta-analysis is performed using effective sample size-based weights in a fixed effects model. Moderate deviation from the expected null distribution is observed, but this inflation is restricted to the upper tail of results (lambda=0.962). LD score regression within each ancestry suggests that true polygenic effects for AD are the primary source of deviations from the null distribution within both the EU (lambda=1.053, intercept=1.018, ratio=0.298) and AA (lambda=1.007, intercept=0.991-0.997; see Supplementary Information) ancestry analyses.



Supplementary Figure 10

PRS prediction using weights derived from alcohol dependence GWAS of unrelated EU and AA individuals

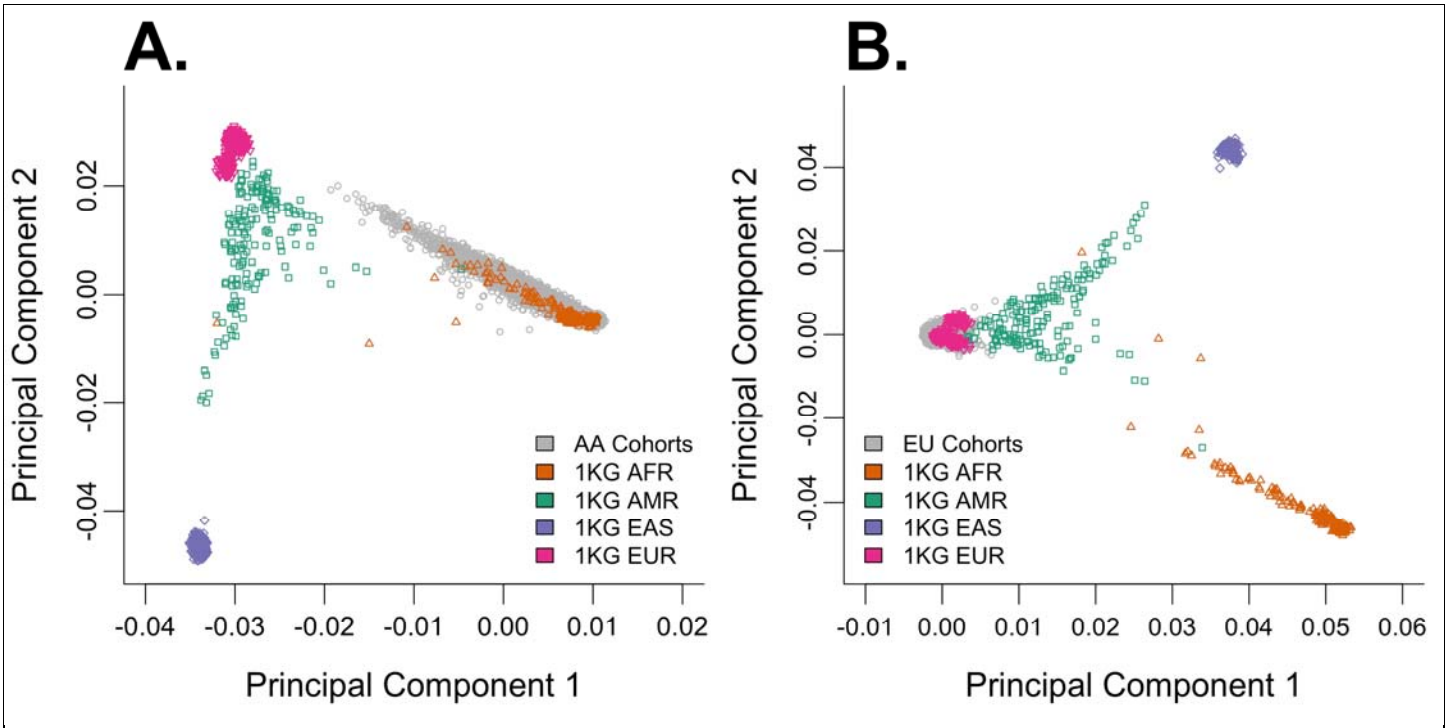
Variance in alcohol phenotypes explained by polygenic risk scores (PRS) derived from the alcohol dependence GWAS meta-analysis of unrelated EU (panels A, B and D; n_{case}=8,485, $n_{control}$=20,272) and AA (panel C; n_{case}=2,991, $n_{control}$=2,808) individuals. The y-axis is pseudo-R^2 for ordinal traits or R^2 for continuous traits, reported as a percentage; note that the scale of the y-axis differs between plots. Panel A shows the association between EU alcohol dependence PRS and alcohol use disorder (AUD) diagnosis (dark gray) and symptom count (light gray) in the Avon Longitudinal Study of Parents and Children (ALSPAC; n_{case}=337, $n_{controls}$=2,386); Panel B shows the association of CAGE alcohol screener scores with EU alcohol dependence PRS (solid bar), as well as PRS derived for alcohol consumption (striped bar), in Generation Scotland (GS; N=6,906); Panel C shows the prediction of DSM-IV alcohol dependence in the COGA AAfGWAS cohort (N=2,828) by PRS derived from the AA GWAS of alcohol dependence conducted in this study; Panel D shows the prediction of DSM-IV alcohol dependence in the COGA AAfGWAS cohort by PRS derived from the EU GWAS of alcohol dependence conducted in this study. Results are uncorrected for multiple testing.



Supplementary Figure 11

Power analysis for current meta-analysis

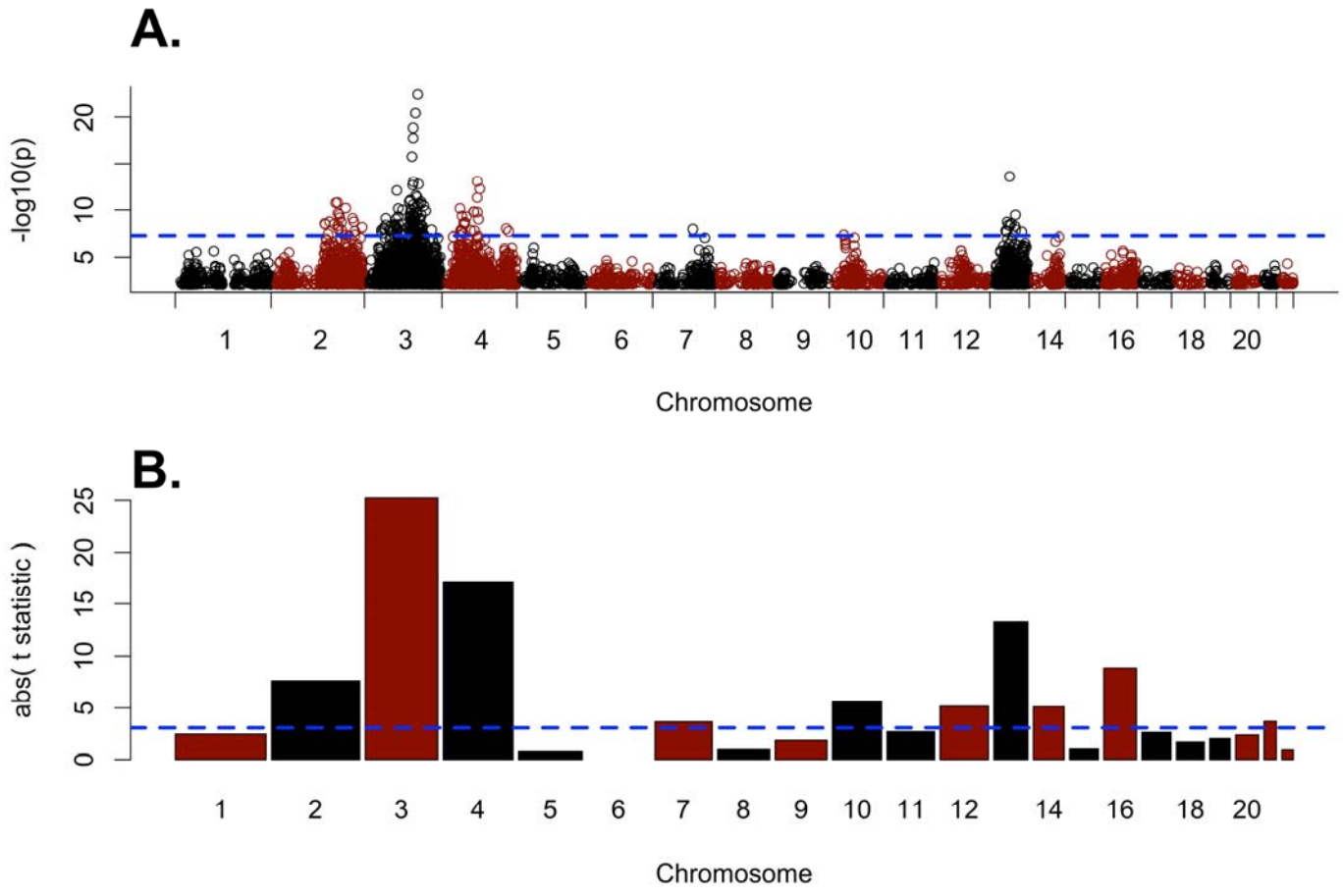
Analysis of power to detect variants associated with AD at thresholds of (A) $P < 5 \times 10^{-8}$ and (B) $P < 1 \times 10^{-6}$ in the current study, conditional on allele frequency and effect size (odds ratio) using CaTS[108]. Power calculated based on an effective sample sizes of n=31,844 for the trans-ancestral discovery meta-analysis, n=26,853 for the EU meta-analysis, and n=4,991 for the AA meta-analysis. See Supplementary Information for references.



Supplementary Figure 12

PCA comparison to 1000 Genomes Project populations

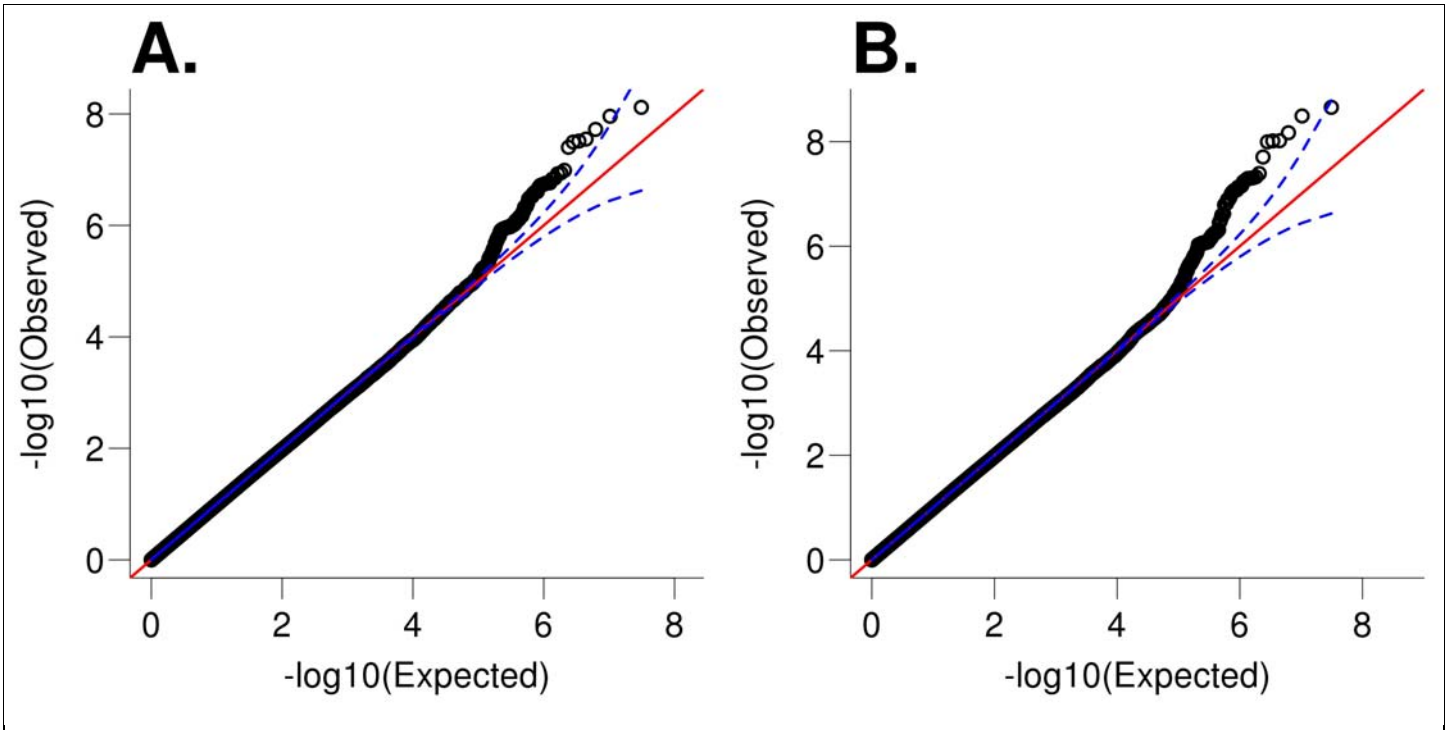
PCA of (A) AA (n_{case}=2,991; $n_{control}$=2,808) and (B) EU (n_{case}=8,485; $n_{control}$=20,272) unrelated genotyped individuals merged across cohorts and with 1000 Genomes Project[69] population reference samples (AFR: African; AMR: Admixed American; EAS: East Asian; EUR: European). Orientation of the ancestry axes differs between the two plots due to the inclusion of the AA and EU samples in the respective PCA calculations. Results confirm that the EU and AA cohorts are consistent with the expected population ancestries, and that the admixed AA samples have the expected cline of admixture of African and European ancestry. See Supplementary Information for references.



Supplementary Figure 13

Example of PCA results tagging local ancestry in ADAA cohort

Association of the 8th principle component (PC) in the ADAA cohort (*n*=1,813) from linear regression with (A) each SNP genome-wide and (B) estimated proportion of African ancestry on each chromosome conditional on genome-wide ancestry proportions. Panel A reports two-tailed p-values for each SNP, with the dashed blue reference line indicating the $P < 5 \times 10^{-8}$ genome-wide significance threshold, and illustrates the characteristic pattern of PCs associated with localized regions of the genome that is observed in multiple AA cohorts. Bars in Panel B reflect the *t* statistic for the two-sided test of association with the 8th PC in linear regression, colored according to the sign of the effect and with bar widths proportional to the size of the chromosome. The dashed blue reference line in panel B indicates Bonferroni-adjusted significance ($P < 2.27 \times 10^{-3}$ = .05/22 autosomal chromosomes); results for chromosome 6 are omitted due to computational complexity. Comparison of Panel A and Panel B suggests that the localized association with the PC strongly corresponds to differences in local ancestry across chromosomes.



Supplementary Figure 14

Comparison of AA meta-analysis results by number of PCA covariates

QQ plots for association with AD in effective sample size weighted meta-analysis of all AA cohorts (n_{case}=3,335, $n_{control}$=2,945) (A) controlling for a full 5 principle components (PCs) in each cohort based on sample size, or (B) controlling for 1-5 PCs in each cohort, restricting to PCs that are associated with variants genome-wide rather than specific genomic regions. Compared to the basic analysis in Panel A, Panel B shows little evidence that the reduced number of PC covariates yields inflation from population stratification. The meta-analysis of AA cohorts reflected in Panel B is used as the primary analysis for the current paper.