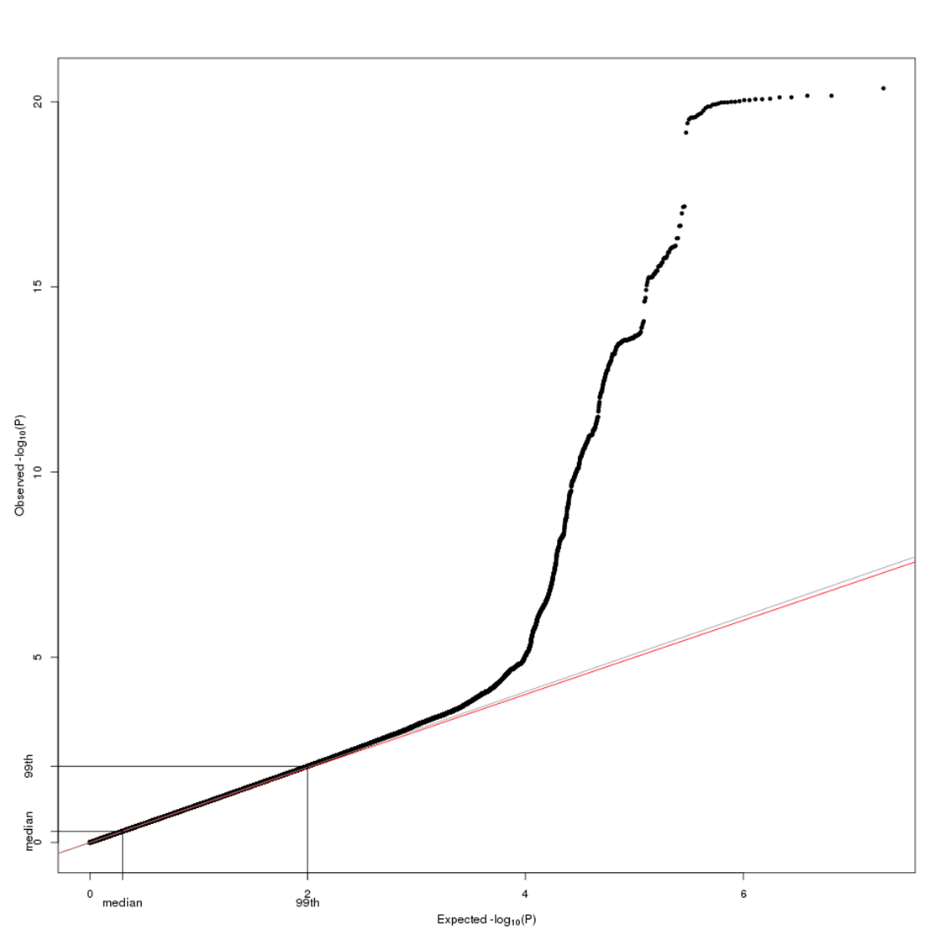
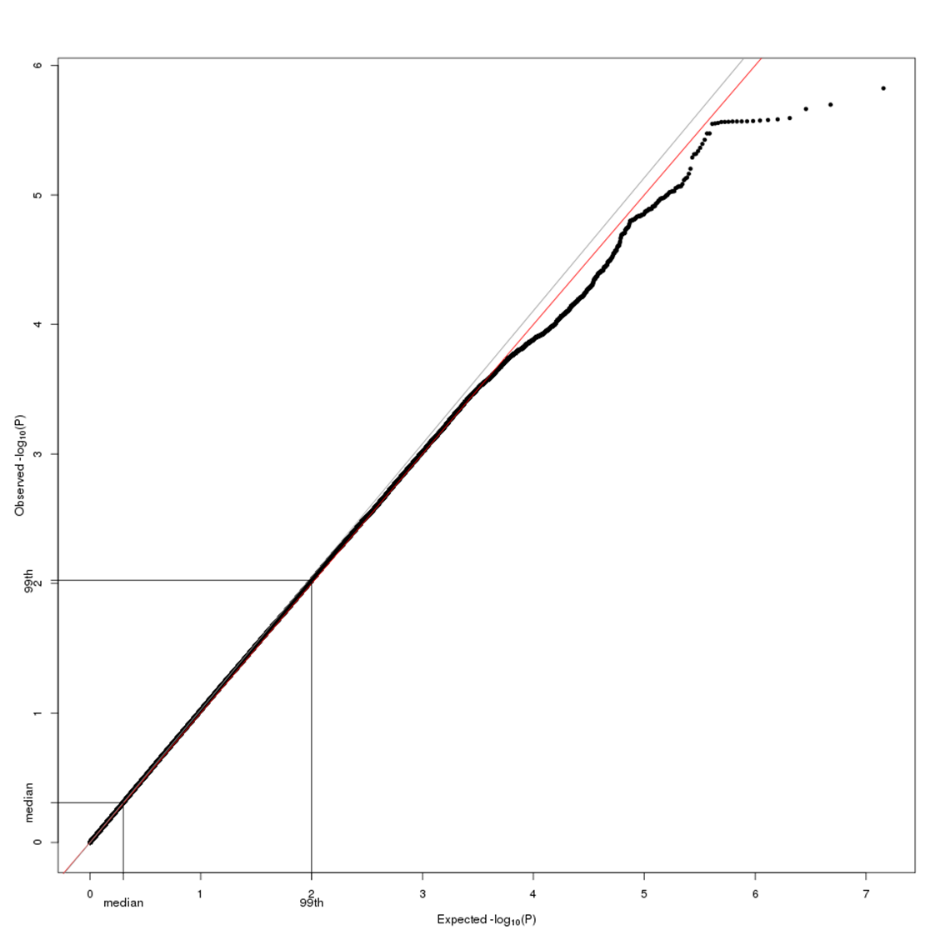
**Supplementary Figure 1 – Quantile-Quantile plots of GWAS among (a) Europeans, (b) Oceanians. (c) Pooled analysis combining both populations.**

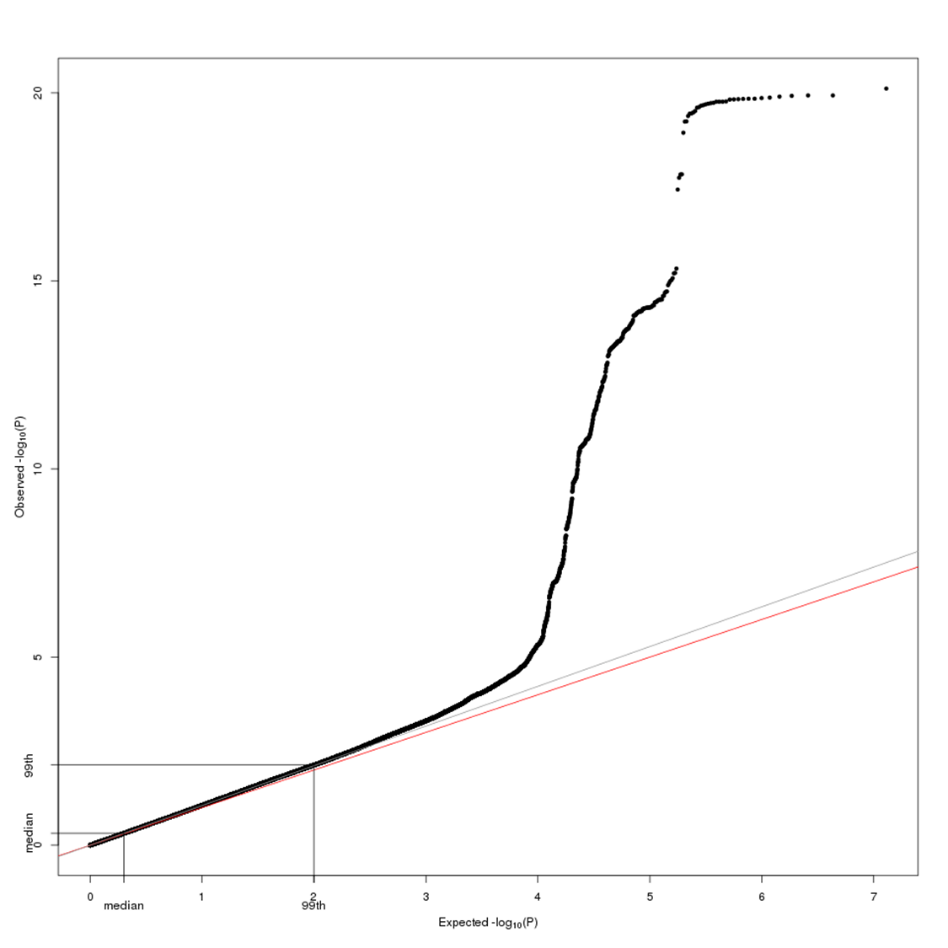
1. Inflation factor lambda = 1.0211

****

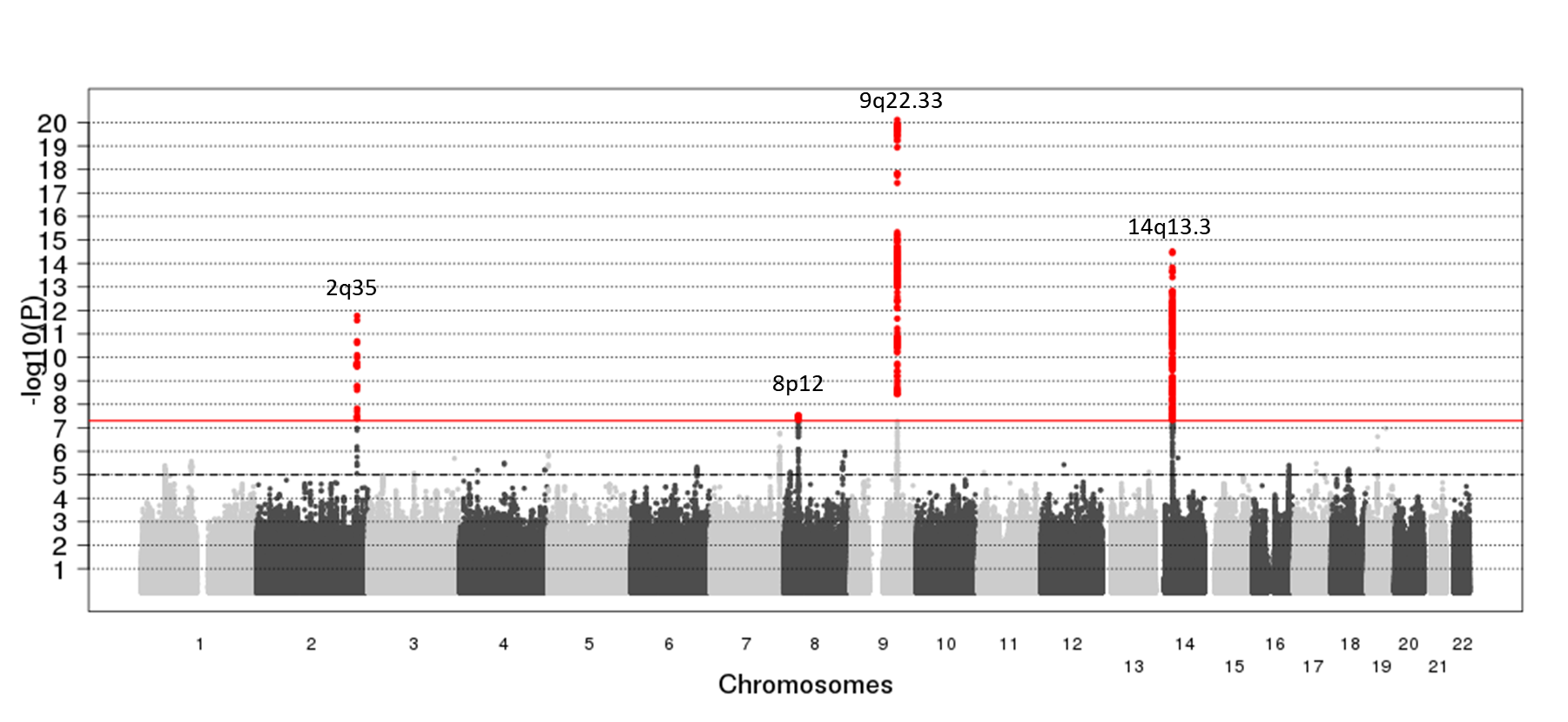
1. Inflation factor lambda = 1.0401



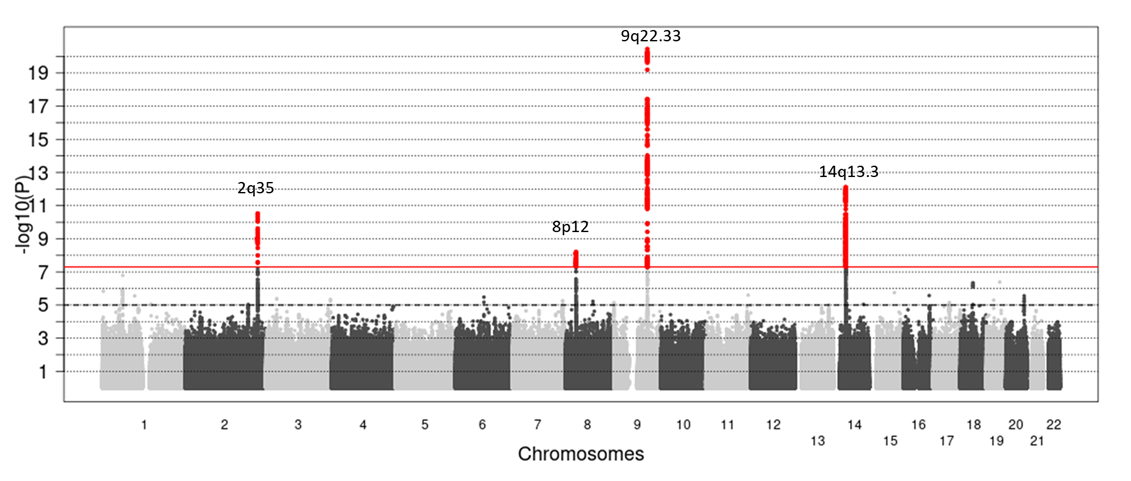
1. Inflation factor lambda = 1.0776

****

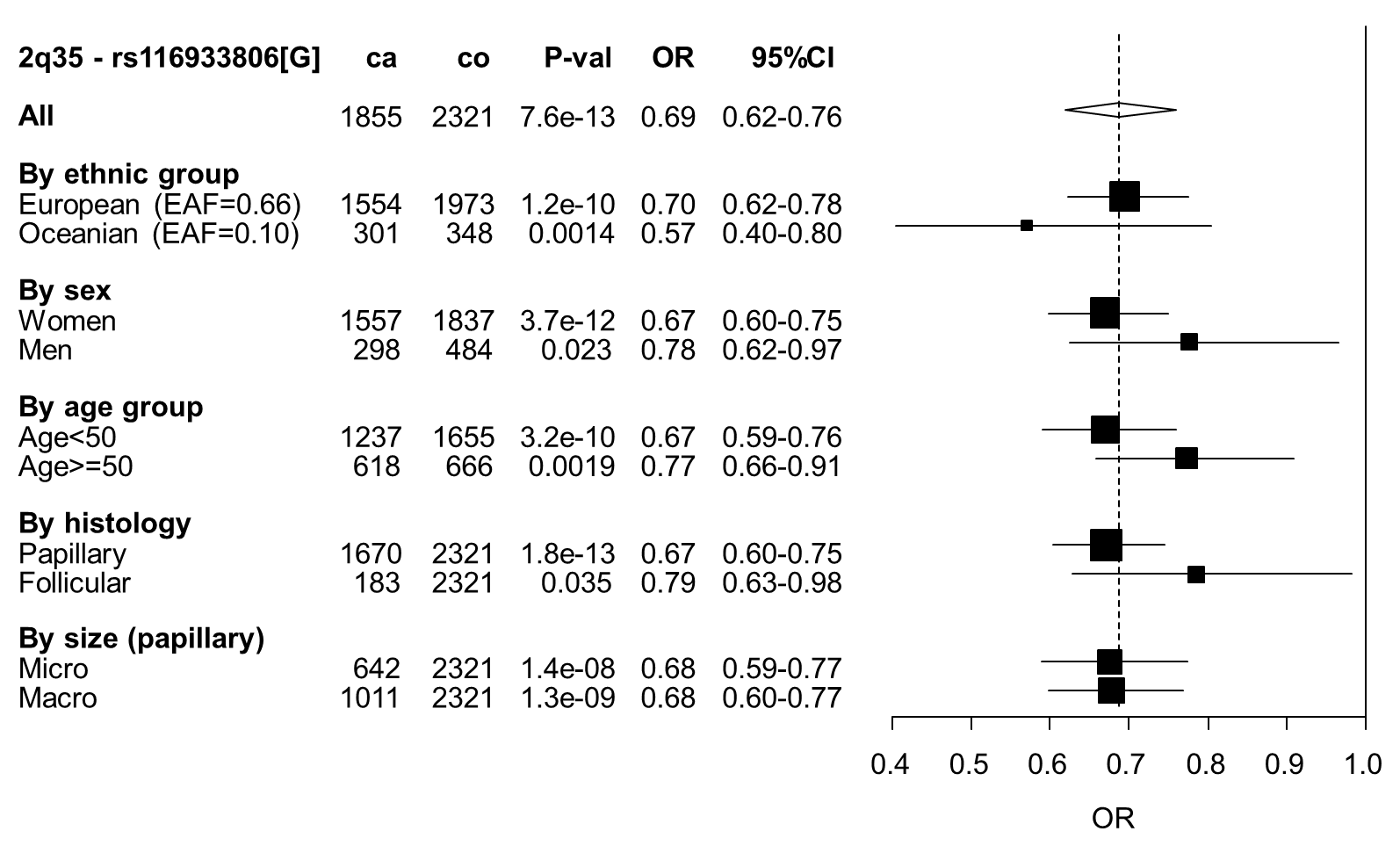
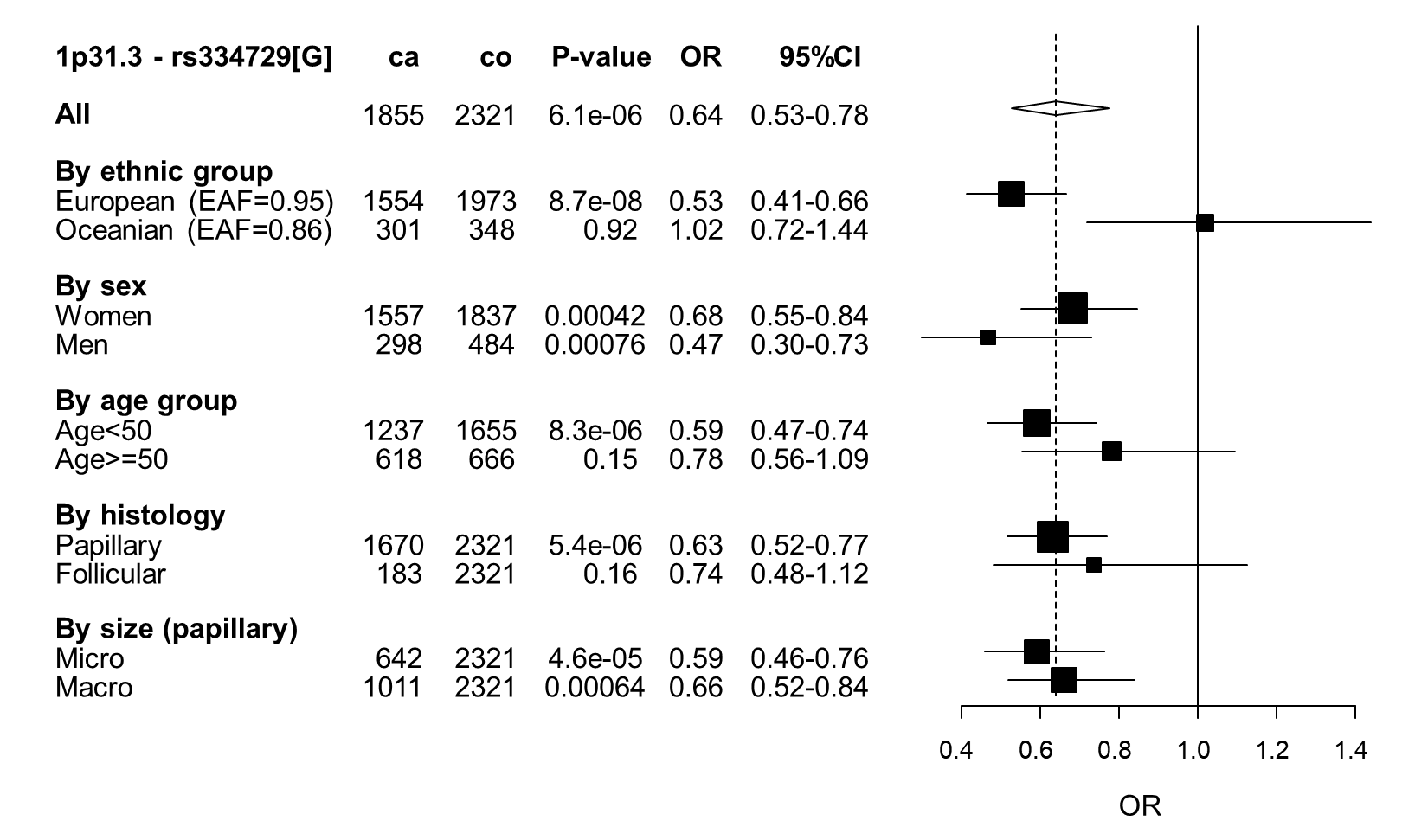
**Supplementary Figure 2 – Manhattan plot of the DTC GWAS on the pooled data including European and Oceanian populations.** Results on 6,479,975 SNPs from 1,855 cases and 2,321 controls. The x axis represents chromosomal location and the y axis represents –log10 (P-value) calculated using the score test. The red line indicates the genome-wide significance threshold (p<5x10-8) while the horizontal dashed line indicates the 10-5 threshold. Red points denote SNPs that are significantly associated to DTC.

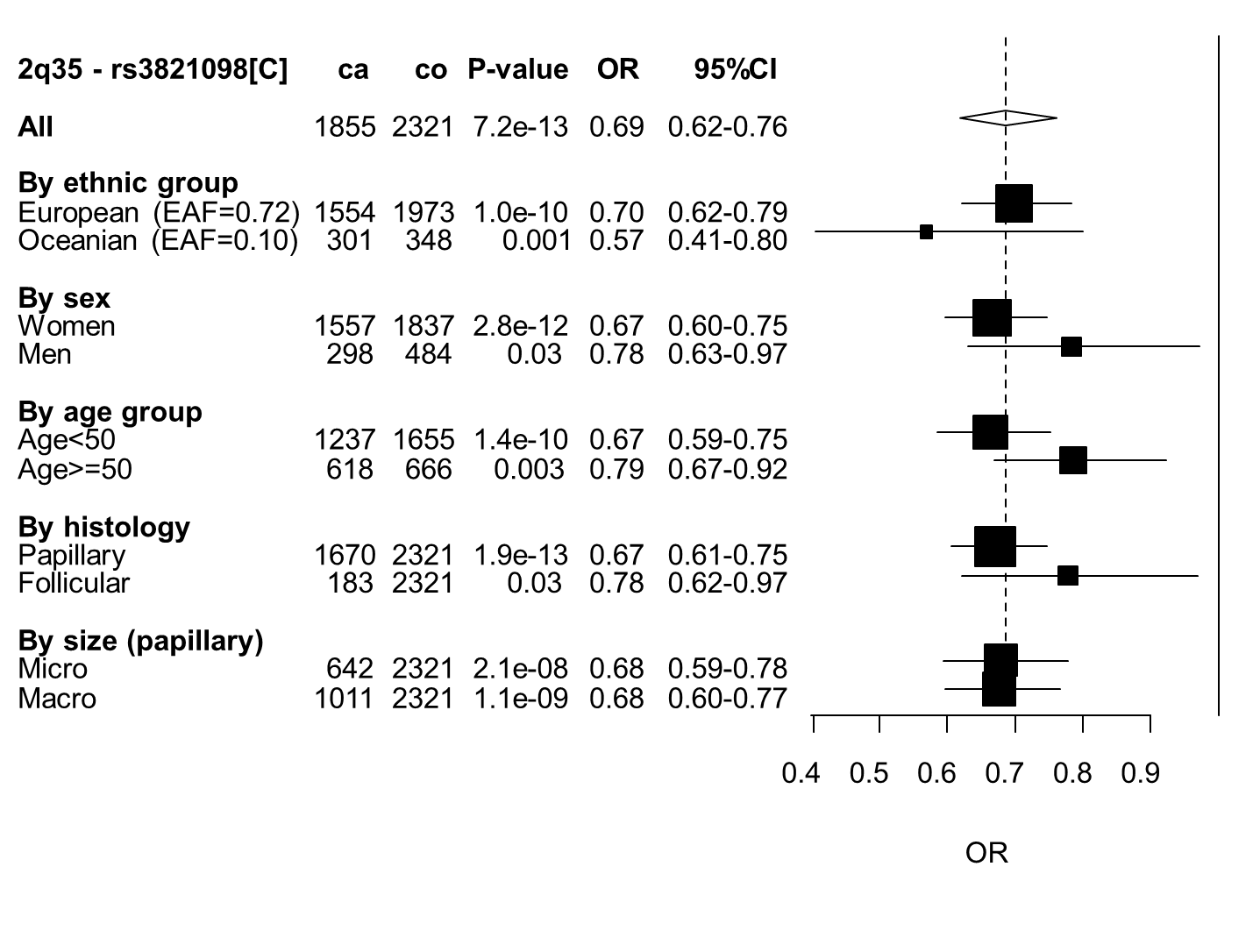
****

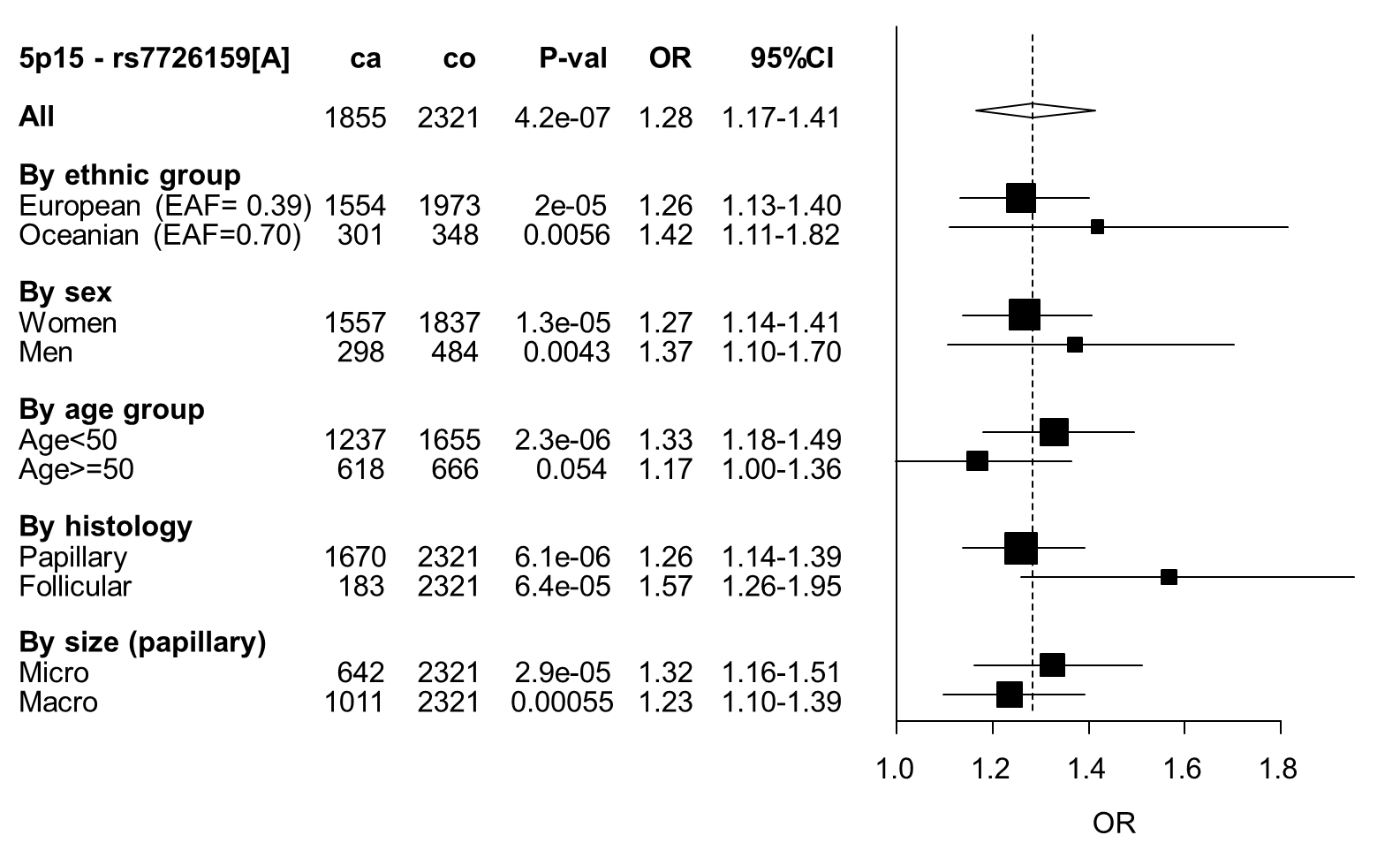
**Supplementary Figure 3 – Manhattan plot of the DTC GWAS restricted to papillary cancer in Europeans.** Results on 6,479,975 SNPs from 1,417 cases and 1,973 controls. The x axis represents chromosomal location and the y axis represents –log10 (P-value) calculated using the score test. The red line indicates the genome-wide significance threshold (p<5x10-8) while the horizontal dashed line indicates the 10-5 threshold. Red points denote SNPs that are significantly associated to DTC.

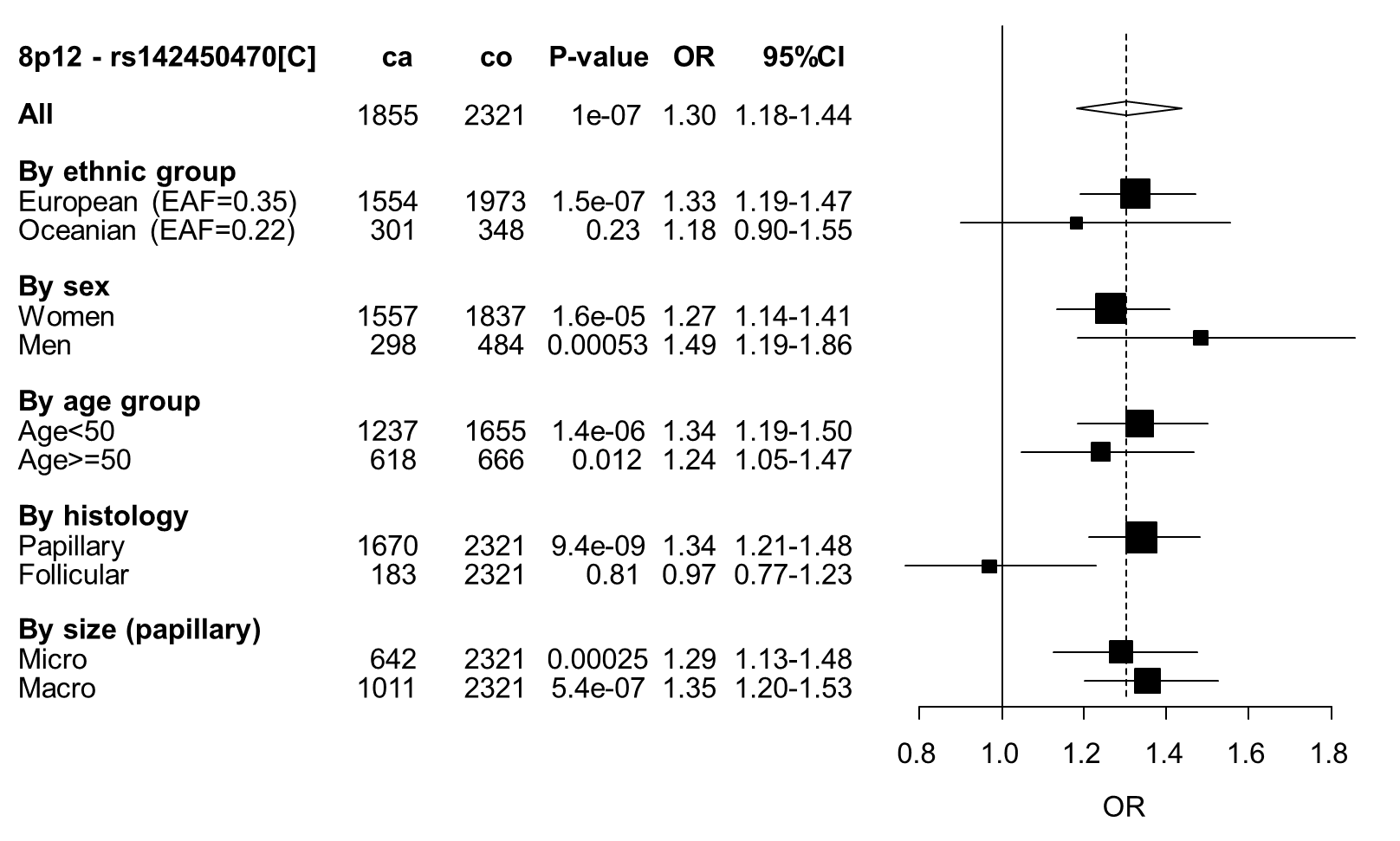


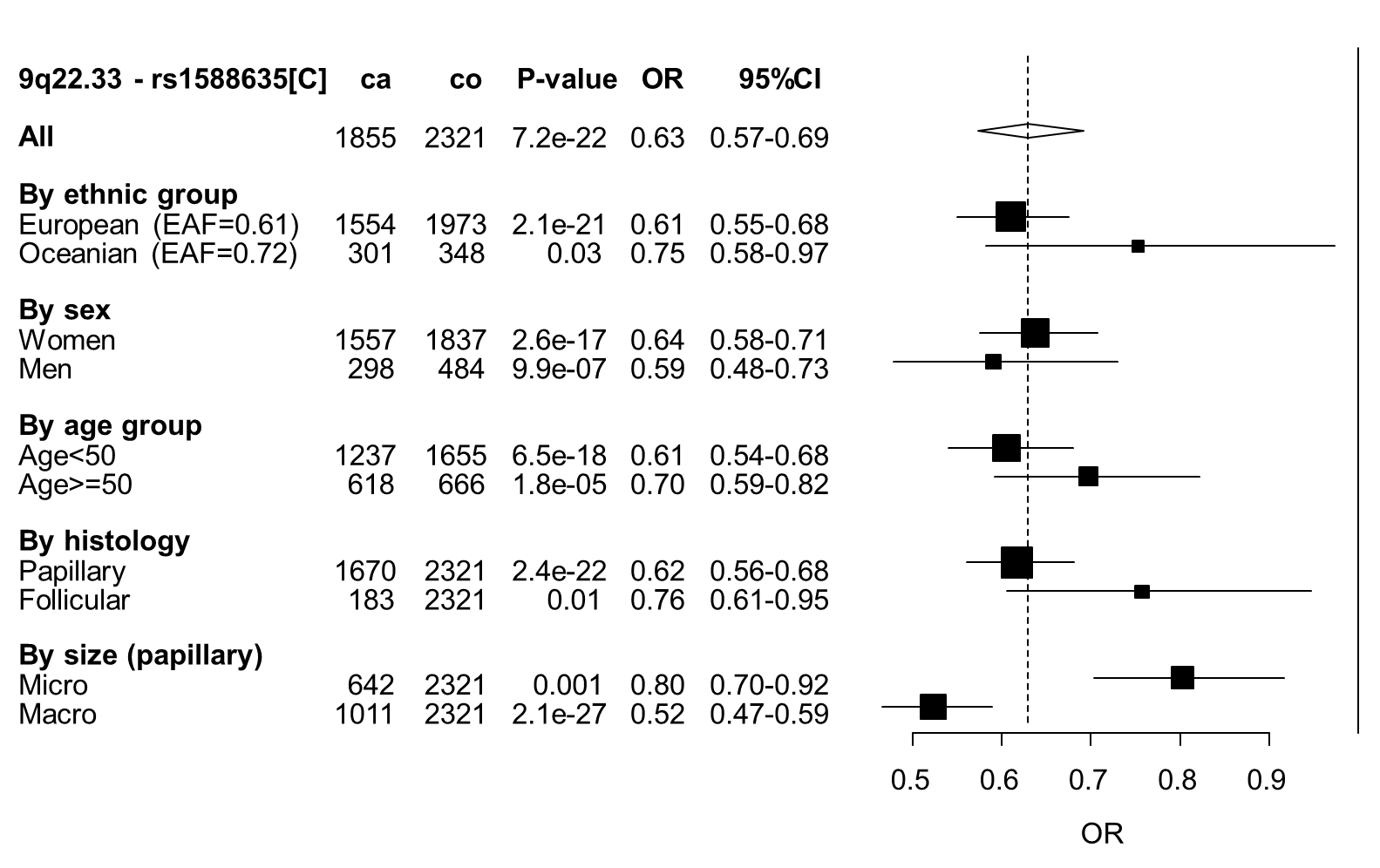
**Supplementary figures 4 – Stratified analyses of lead SNPs at known and suggestive loci in the pooled analysis.** Squares represent odds ratios; size of the square represents inverse of the variance of the log odds ratio; horizontal lines represent 95% confidence intervals; diamonds represent estimate from the pooled analysis; solid vertical line represents an odds ratio of 1; dashed vertical line represents the overall odds ratio. Ca: number of cases, co: number of controls, EAF: effect allele frequency, OR: Odds ratio, CI: confidence interval.

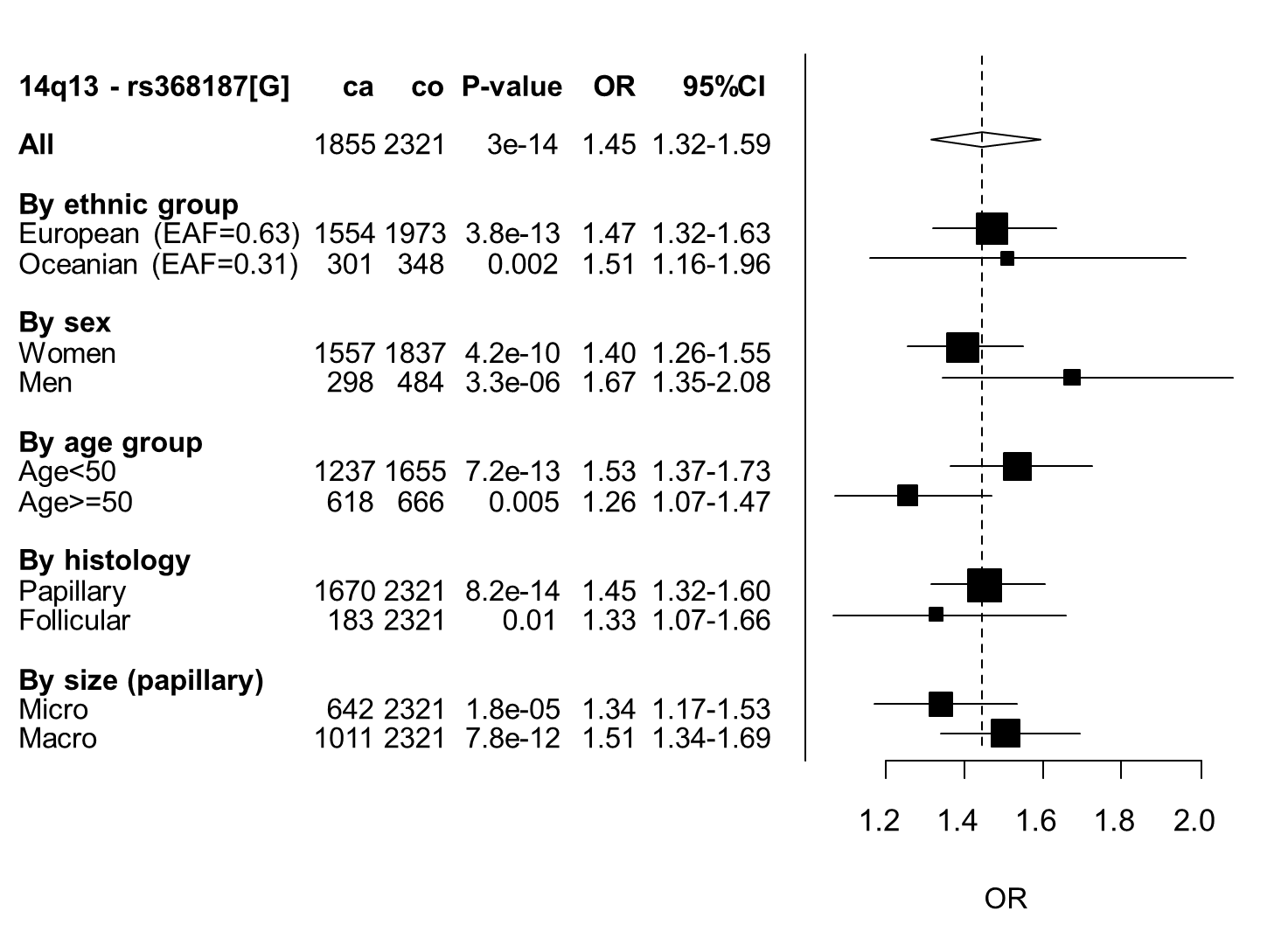
****

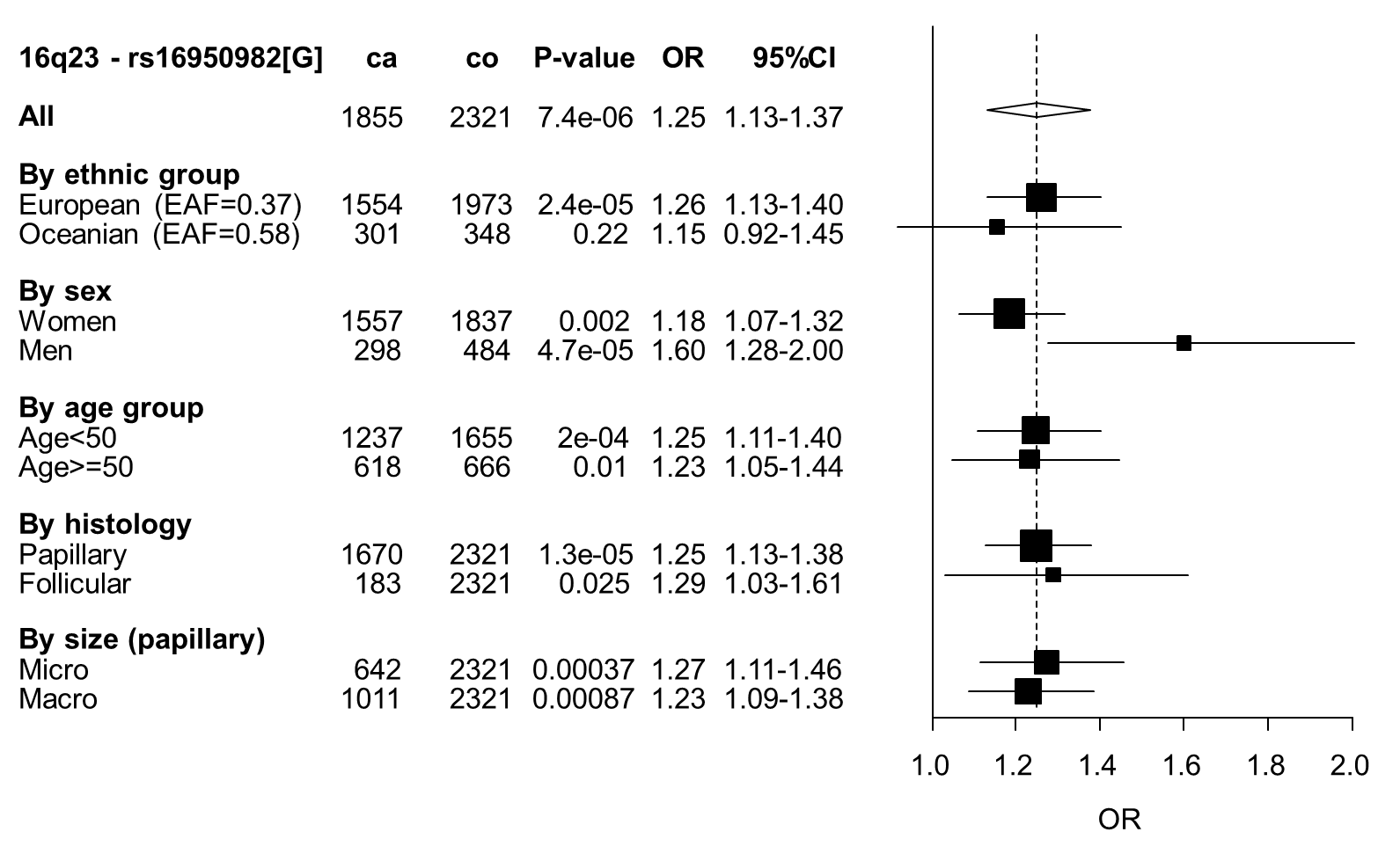


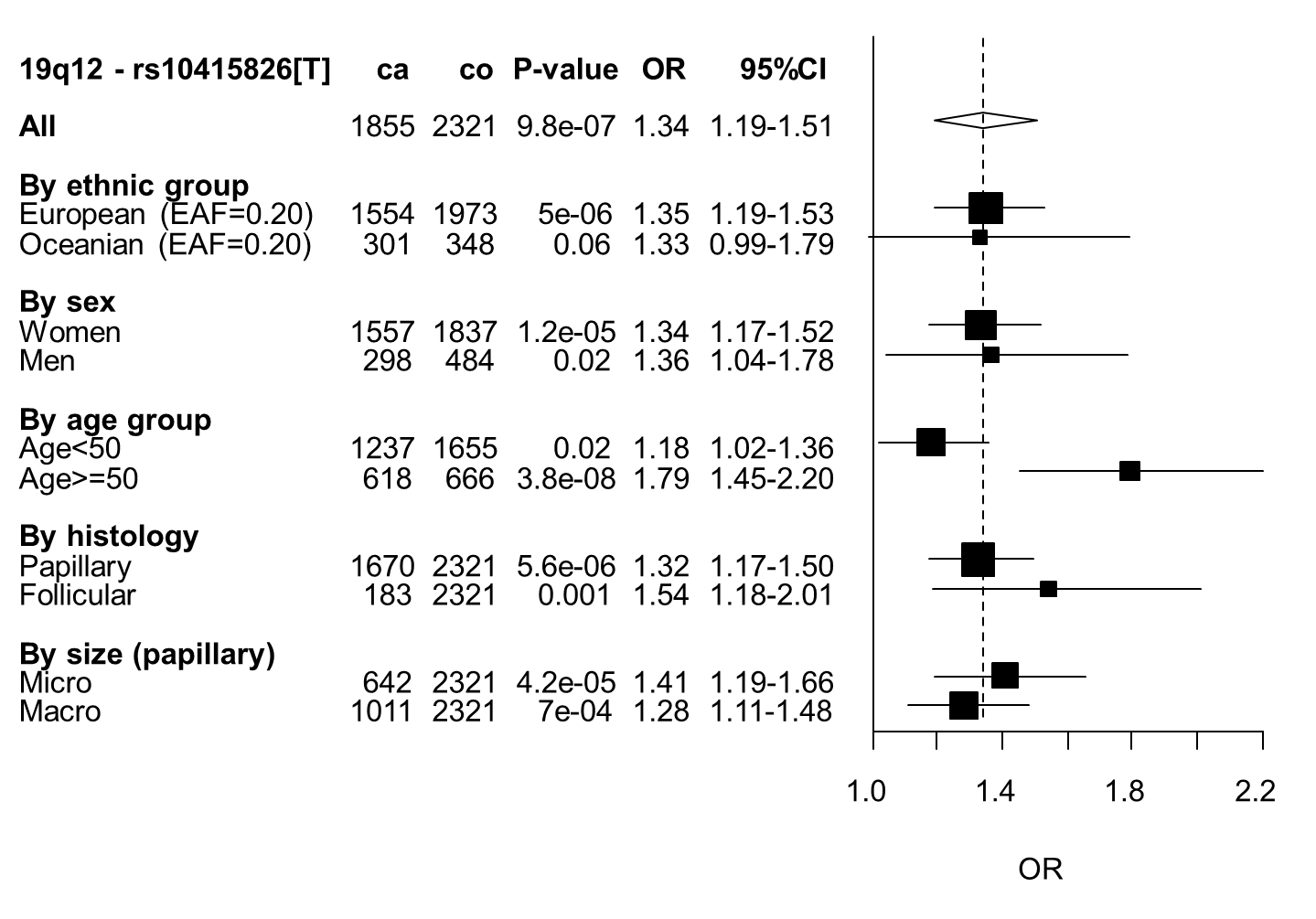
****

****

****

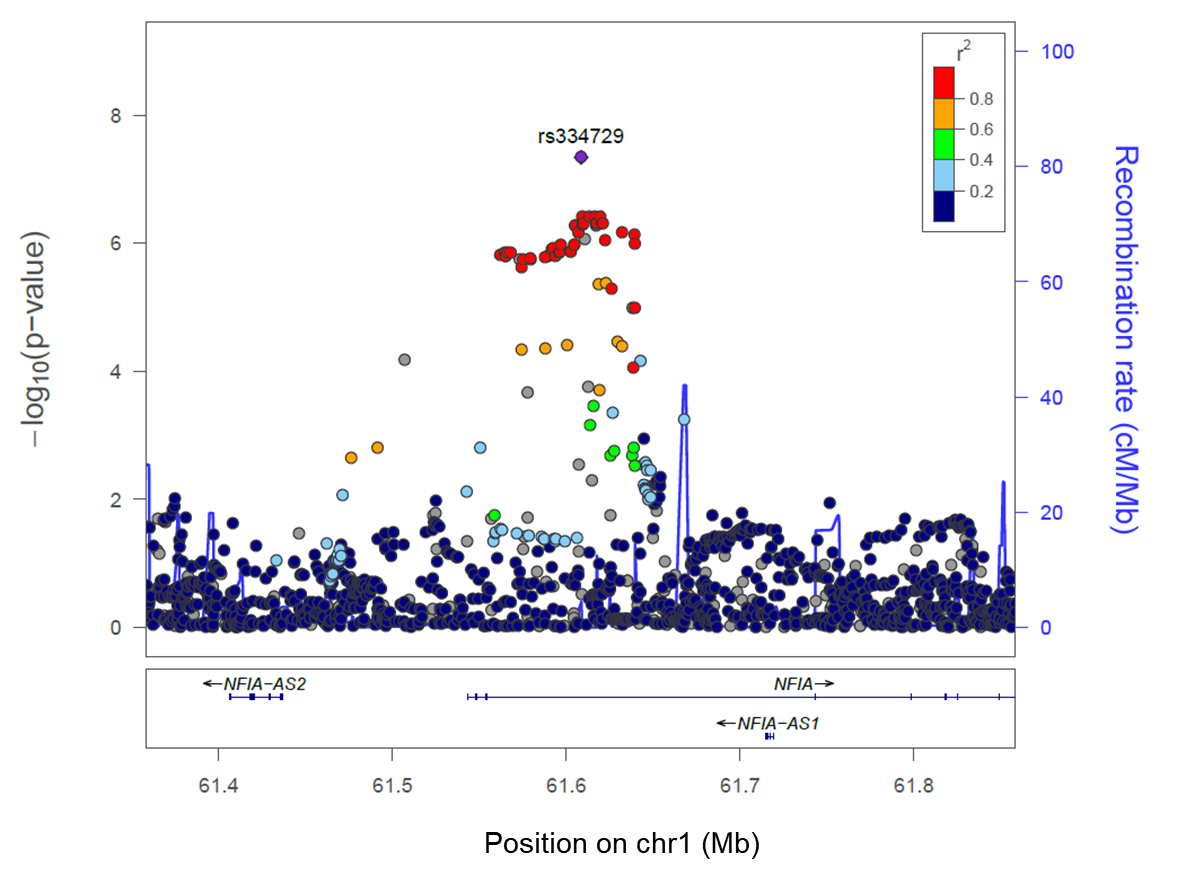
****

****

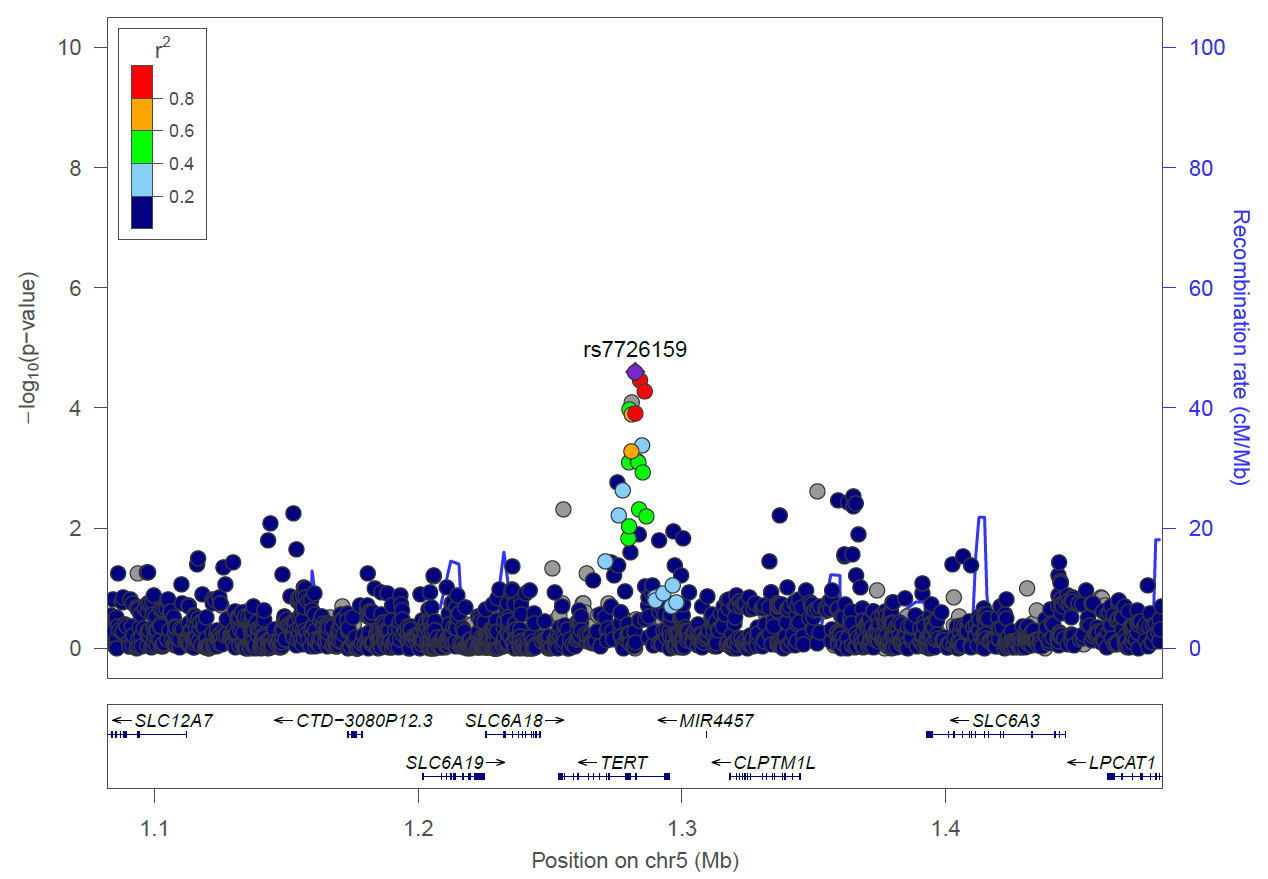
****

**Supplementary Figures 5. Regional plot of the DTC GWAS in Europeans at (a) 1p13.3, (b) 5p15.33, (c) 16q23.2 and (d) 19p12.** The x axis represents the position of SNPs and gene on the chromosome, the y axis represents the – log10 (P-value) and the recombination rate. The color of each SNP spot reflects its correlation (r2) with the most associated SNP in the region (in purple). The correlation between SNP is ranged from high (in red) to low (in blue).

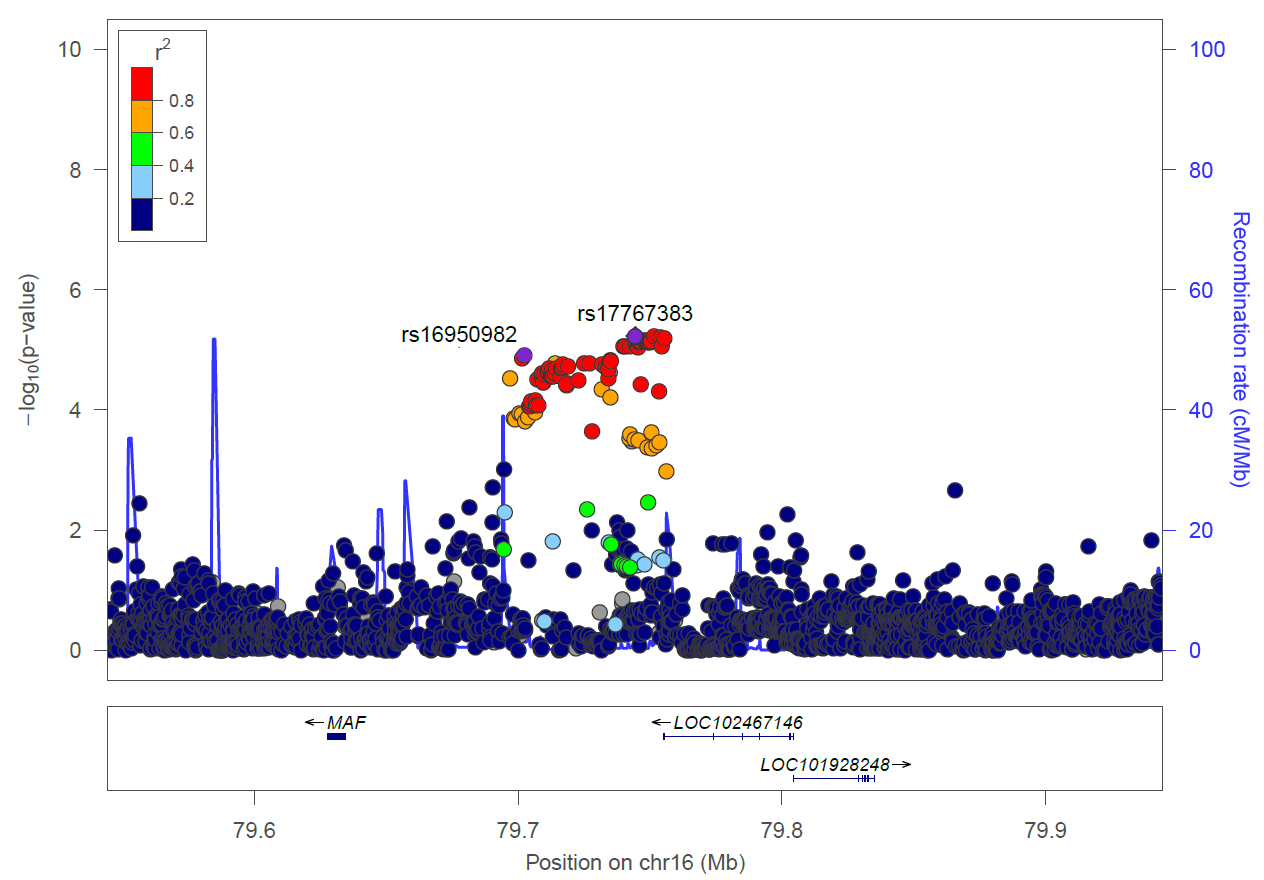
**(a)**

****

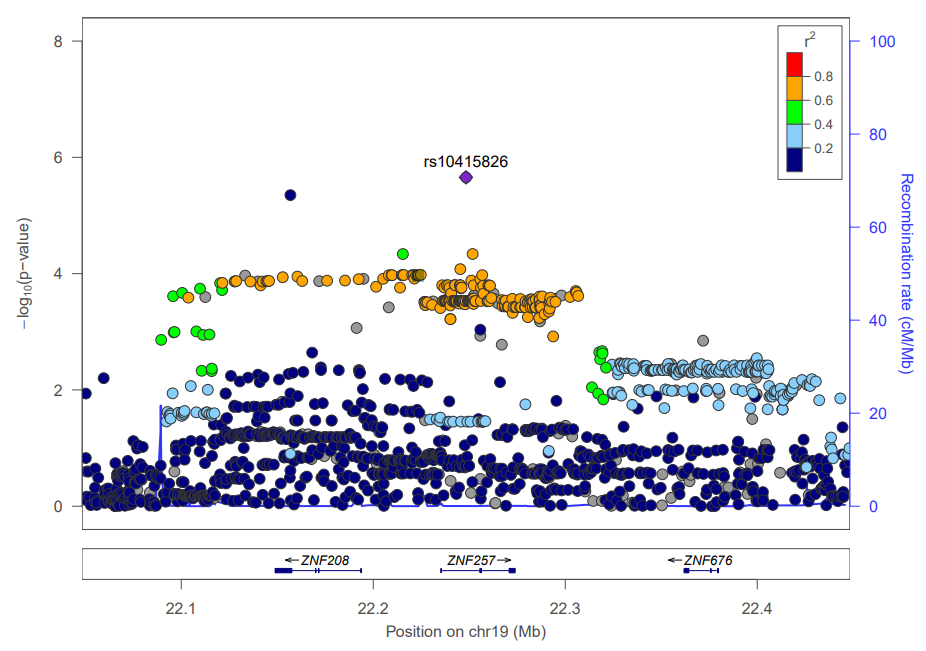
**(b)**

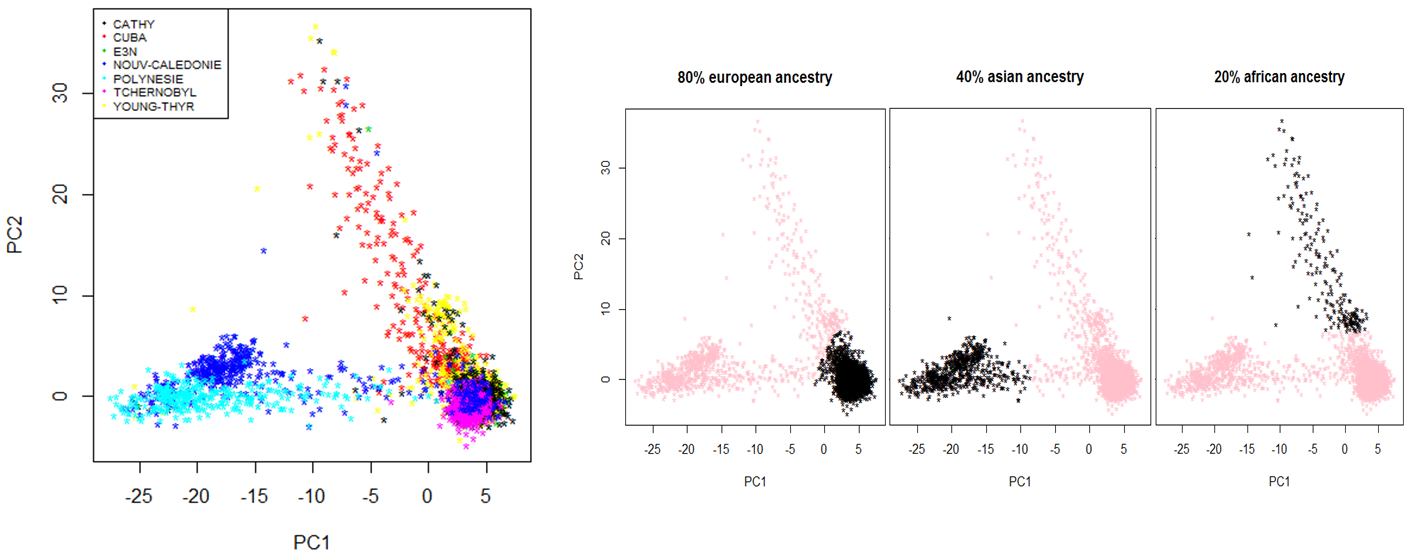


**(c)**



**(d)**



**Supplementary figure 6 – Population structure of the EPITHYR datasets by study.** Proportion of ancestry inference using first and second PCA score. X and Y axis denote the first and second PCA score generated by FastPop.The centroids of each population were computed for Hapmap samples with European, African and Asian ancestry**.**

**Supplementary Figure 7 - Summary of the quality control procedures of the genotyped data**

512,929 SNPs were manufactured

(Illumina OncoArray chip: 499,170 SNPs

+ Add-ons: 13,759)

Genotypes received from CNRGH:

493,303 SNPs (call>90%)

4,553 individuals

-814 Duplicate SNPs

-8,327 SNPs with call <95% by study

-13,832 monomorphic SNPs

-4,120 SNPs with non-ideal cluster plots

466,210 SNPs

4,553 individuals

-17 individuals with call<95%

-12 individuals with sex discordance

-20 duplicates within or across studies

-5 pairs of individuals with duplicate genetic data but different epidemiological data

**466,210 SNPs**

**4,494 individuals**

Definition of ethnic groups using Fastpop based on 2,318 markers

|  |  |  |  |
| --- | --- | --- | --- |
| European  466,210 SNPs  3,538 individuals | | Asian  466,210 SNPS  649 individuals | |
|  | -5,210 monomorphic SNPs  -58,912 SNPs 0<MAF<1%  -563 SNPs not in HWE (p<10-7 in co, p<10-12 in ca)  -11 individuals with extreme heterozygosity |  | -37,513 monomorphic SNPs  -75,177 SNPs 0<MAF<1%  -417 SNPs not in HWE |
| **401,525 SNPs**  **3,527 individuals** | | **353,103 SNPS**  **649 individuals** | |

**Supplementary Figure 8 –** **Principal component analysis of the Oceanian population of EPITHYR, the Human Genome Diversity Project and Oceanians from the Simons Diversity Project.** The x and y axes represent the two first principal components. POL: individuals from French Polynesia study, NC: individuals from New Caledonia study, YT : individuals from Young-Thyr study.