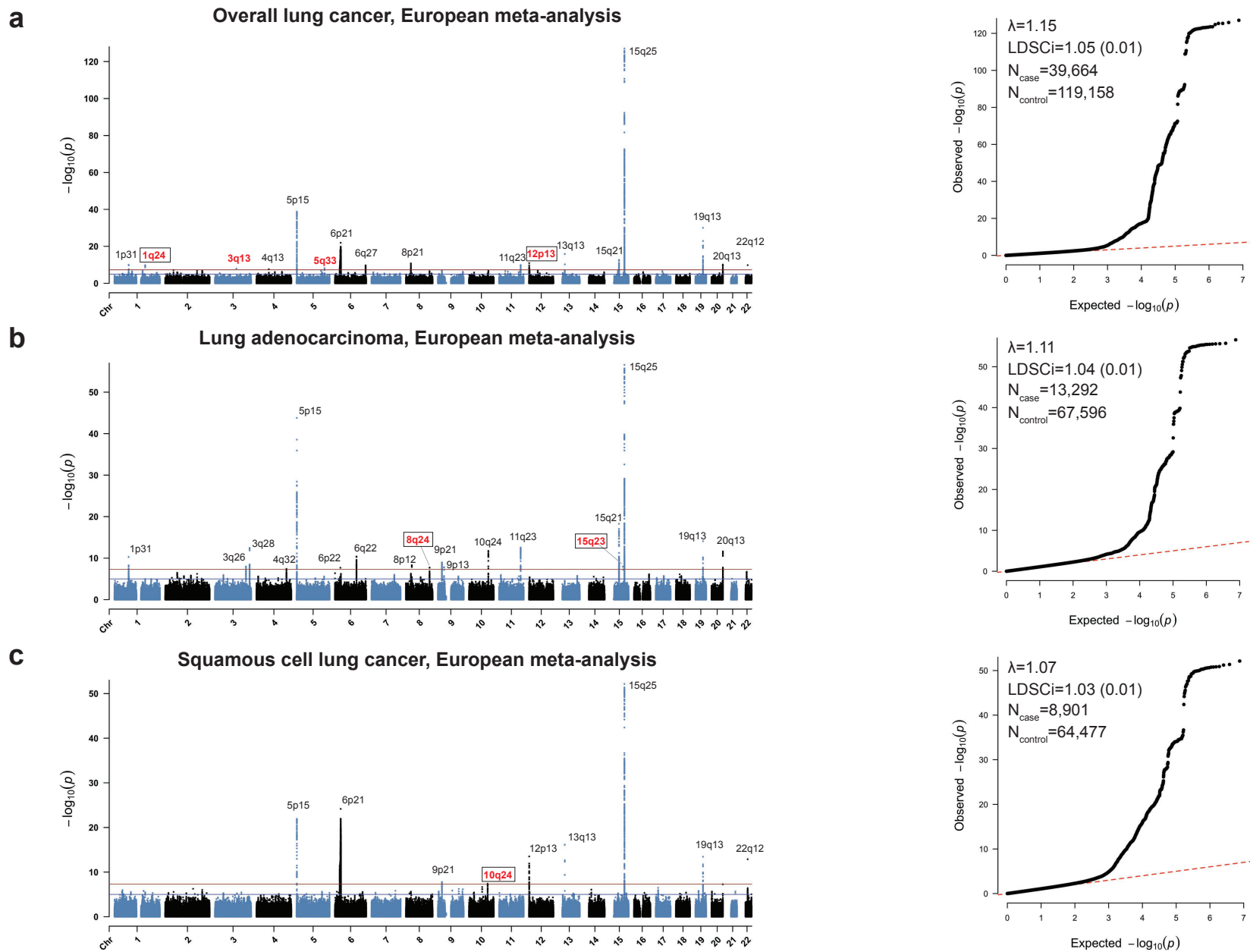
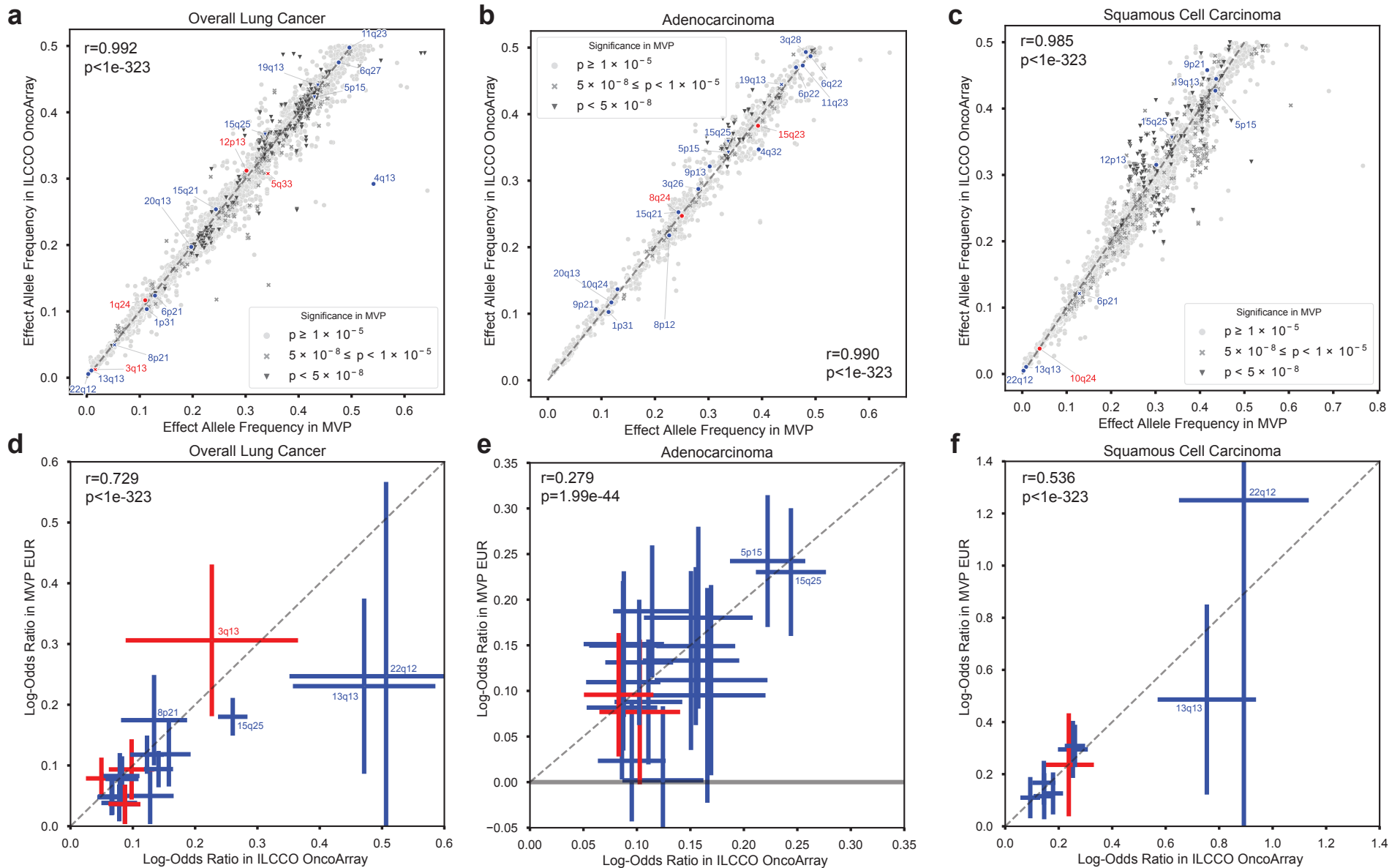


**Supplementary Fig. 1. Study overview.** Genome-wide association studies were performed in Million Veteran Program (MVP) European and African ancestry (AA) cohorts for overall lung cancer, adenocarcinoma, and squamous cell carcinoma. MVP and International Lung Cancer Consortium OncoArray (ILCCO) European cohorts were meta-analyzed, and further meta-analyzed with AA for multi-ancestry meta-analysis. Multi-trait conditional meta-analysis was performed on EA using average cigarettes per day from Liu et al. (2019). Replication and combined meta-analysis was performed using external OncoArray cohorts.



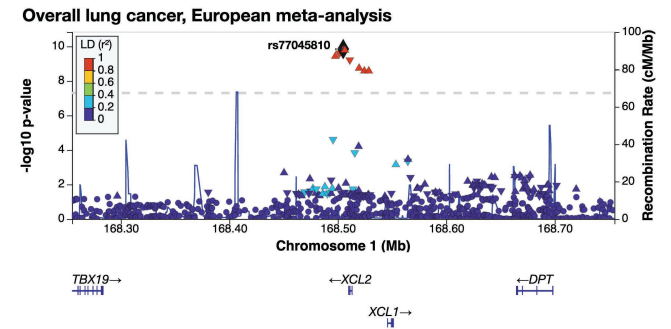
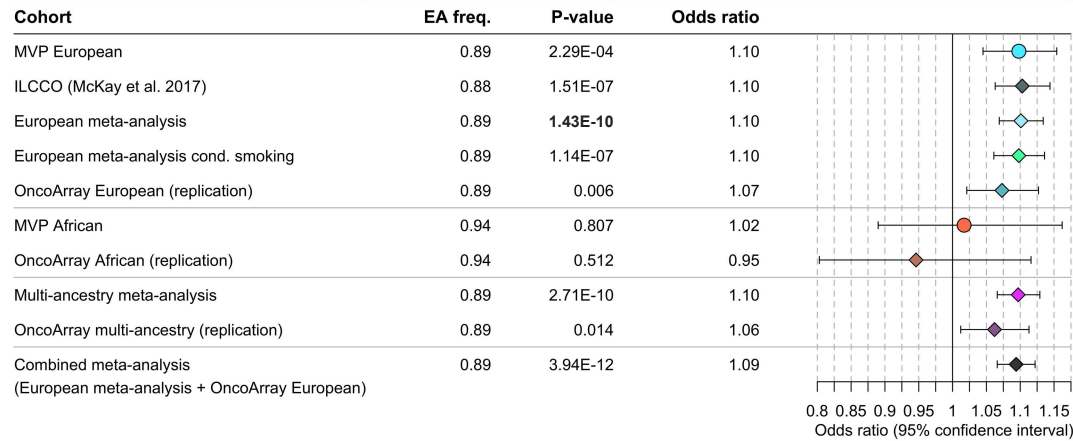
**Supplementary Fig. 2. Manhattan plots and quantile-quantile (QQ) plots for European meta-analyses.** Manhattan and QQ plots are shown for **a)** overall lung cancer; **b)** lung adenocarcinoma (LUAD); and **c)** squamous cell lung carcinoma (LUSC). Cytoband positions for significant loci are noted in each Manhattan plot; putatively novel loci identified in this study are in red; externally replicated novel loci are indicated by a box. Genomic control ( $\lambda$ ) values, LDSC intercepts, and sample sizes are inset in QQ plots.



**Supplementary Fig. 3. Effect allele frequency concordance between International Lung Cancer Consortium OncoArray (ILCCO) and Million Veteran Program European ancestry (EA) GWAS. (a-c)** Effect allele frequency concordance for all variants tested in both studies with  $P < 1 \times 10^{-5}$  in ILCCO for **a)** overall lung cancer, **b)** lung adenocarcinoma, and **c)** squamous cell lung carcinoma. Points are styled based on significance level in MVP. **(d-f)** Effect size concordance for genome-wide significant variants in **d)** overall lung cancer, **e)** lung adenocarcinoma, and **f)** squamous cell lung carcinoma. One-to-one concordance is shown as a dashed line. Index variants from the EA meta-analysis between ILCCO and MVP are annotated by locus. Novel significant loci after meta-analysis are annotated in red.

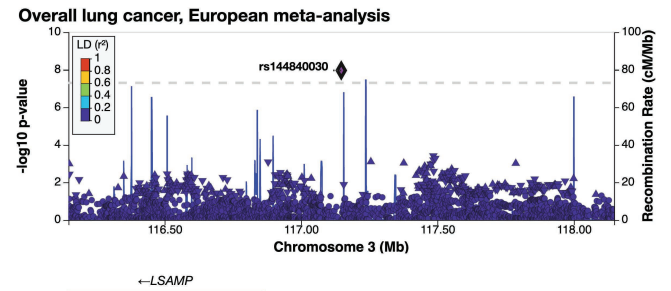
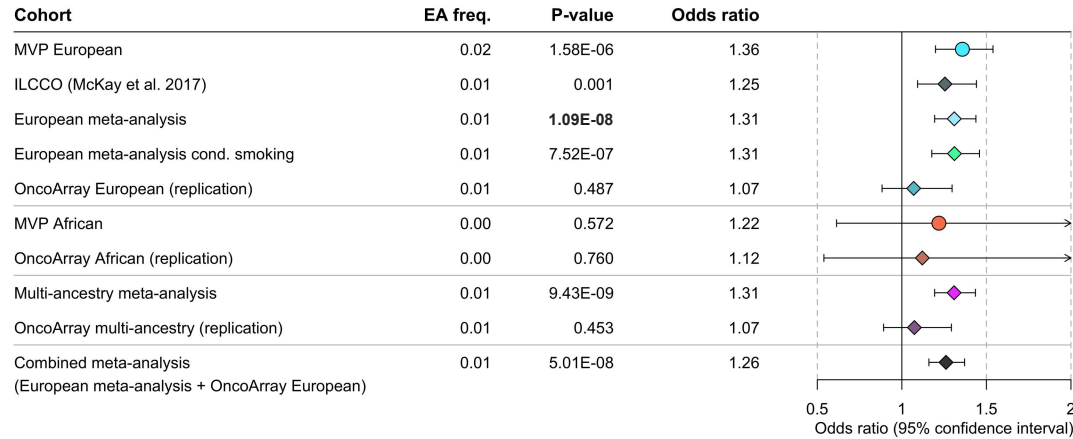
S4a

Overall lung cancer, rs77045810 (XCL2; 1q24.2) : Novel and replicated



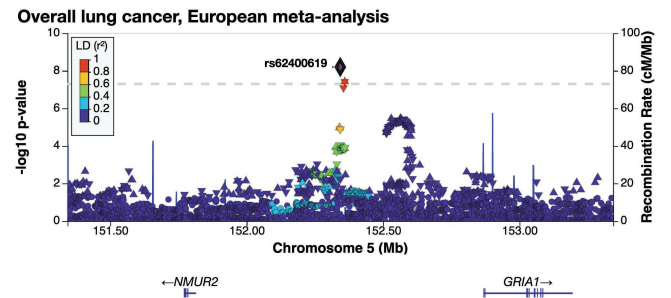
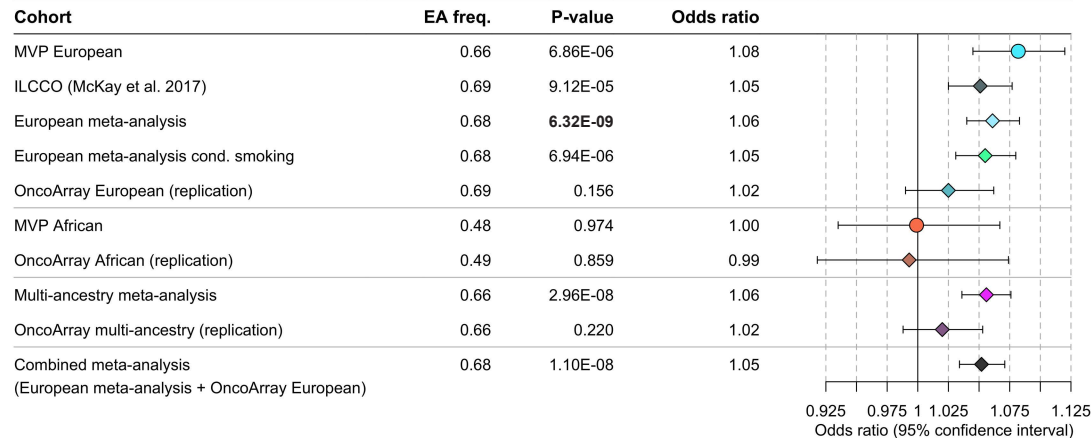
b

Overall lung cancer, rs144840030 (LSAMP; 3q13.31) : Putatively novel



c

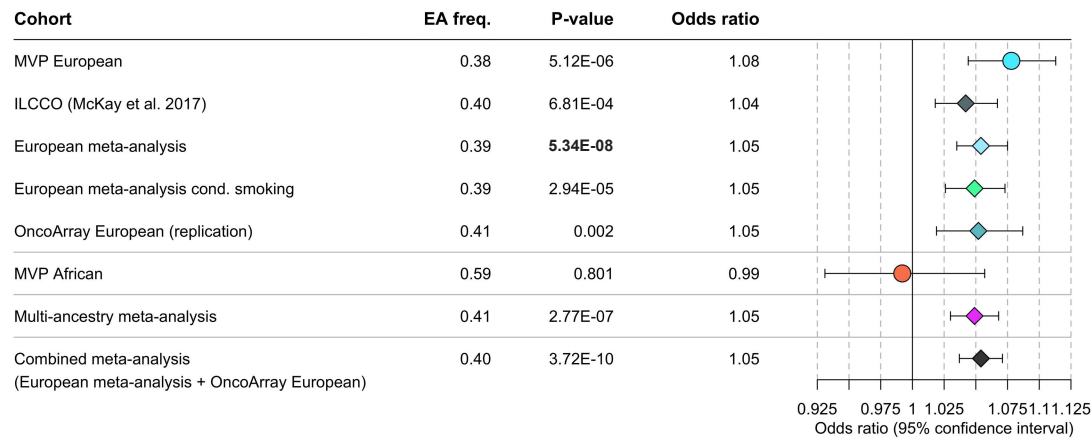
Overall lung cancer, rs62400619 (NMUR2; 5q33.1) : Putatively novel



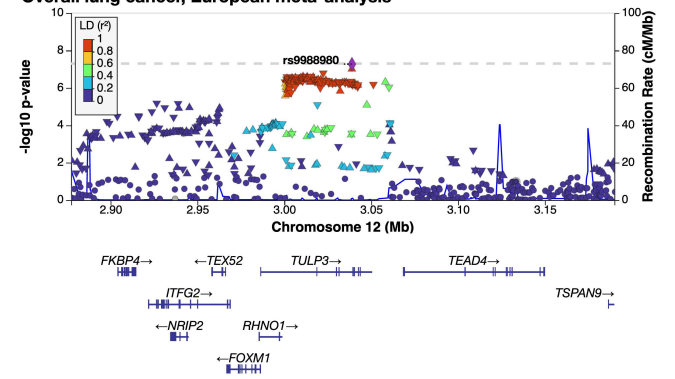


S4d

Overall lung cancer, rs9988980 (TULP3; 12p13.33) : Novel and replicated

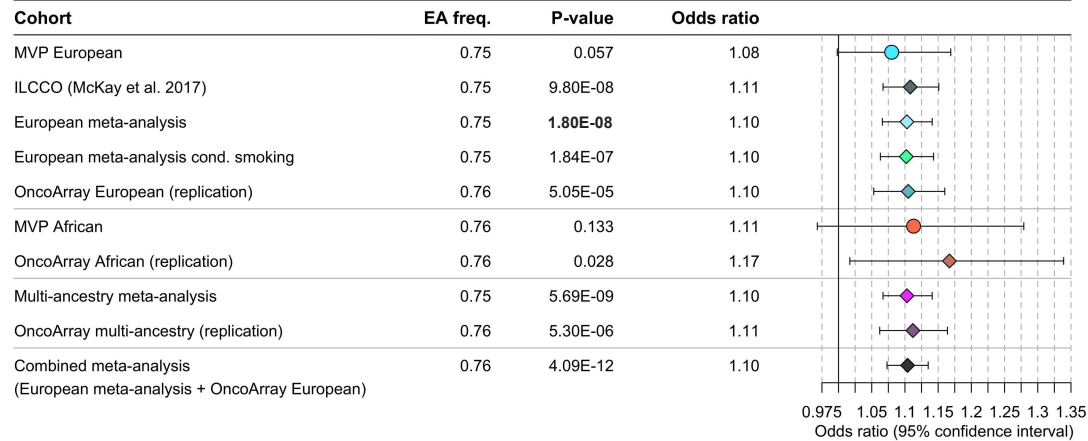


Overall lung cancer, European meta-analysis

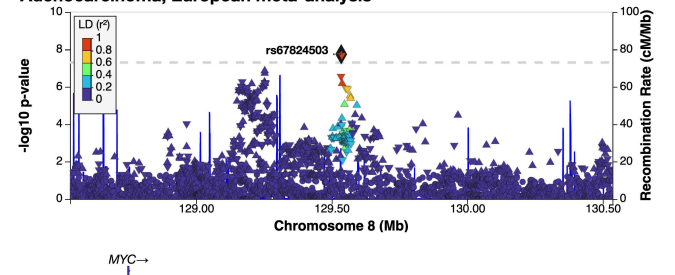


e

Lung adenocarcinoma, rs67824503 (MYC; 8q24.21) : Novel and replicated

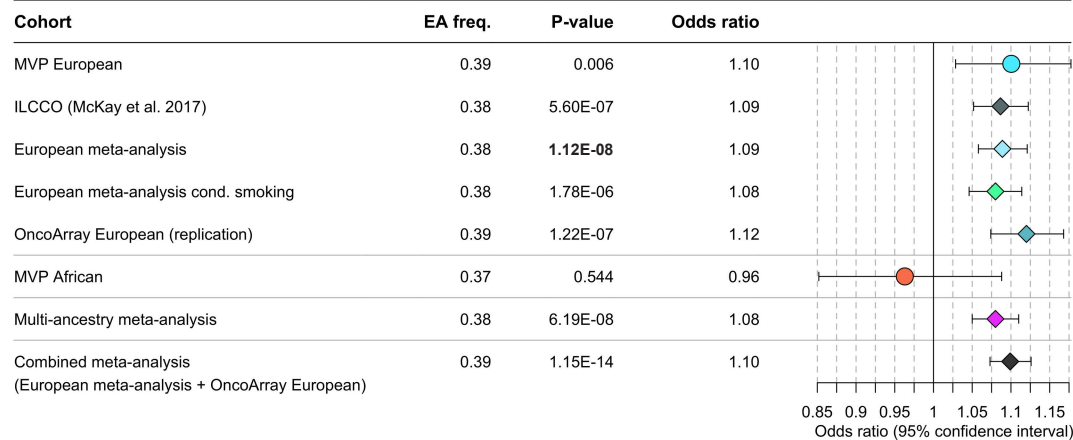


Adenocarcinoma, European meta-analysis

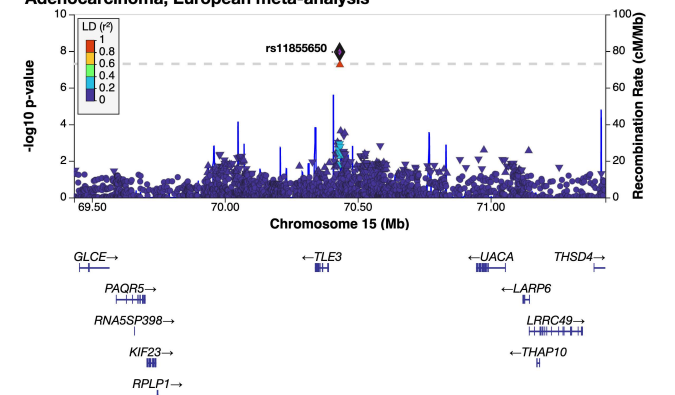


f

Lung adenocarcinoma, rs11855650 (TLE3; 15q23) : Novel and replicated

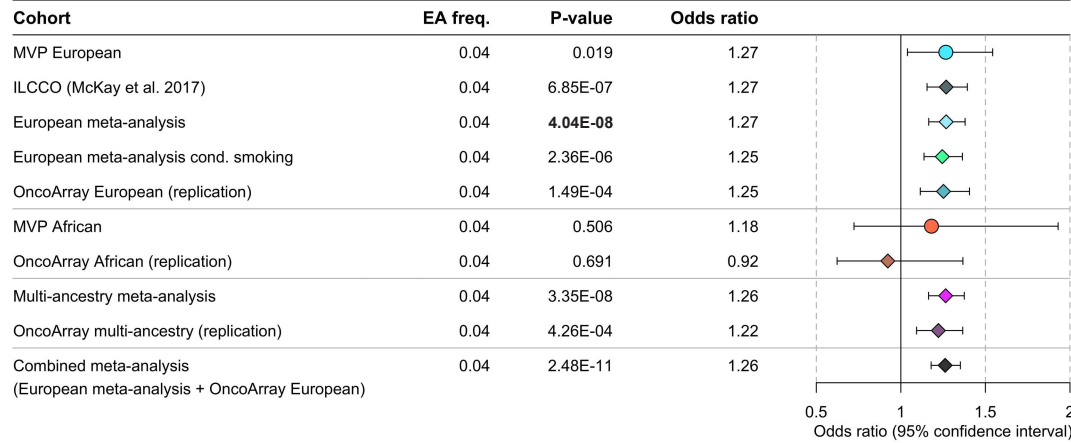


Adenocarcinoma, European meta-analysis

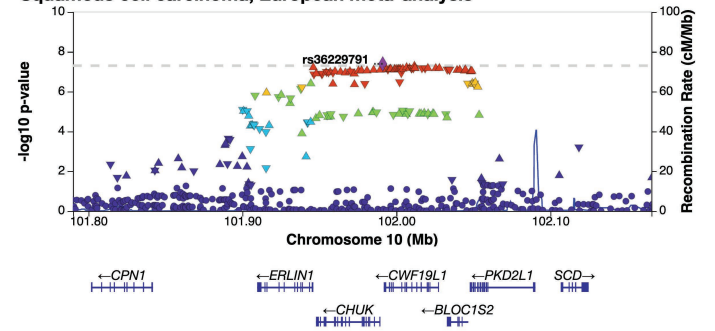


S4g

**Squamous cell lung carcinoma, rs36229791 (CHUK/BLOC1S2; 10q24.31) : Novel and replicated**

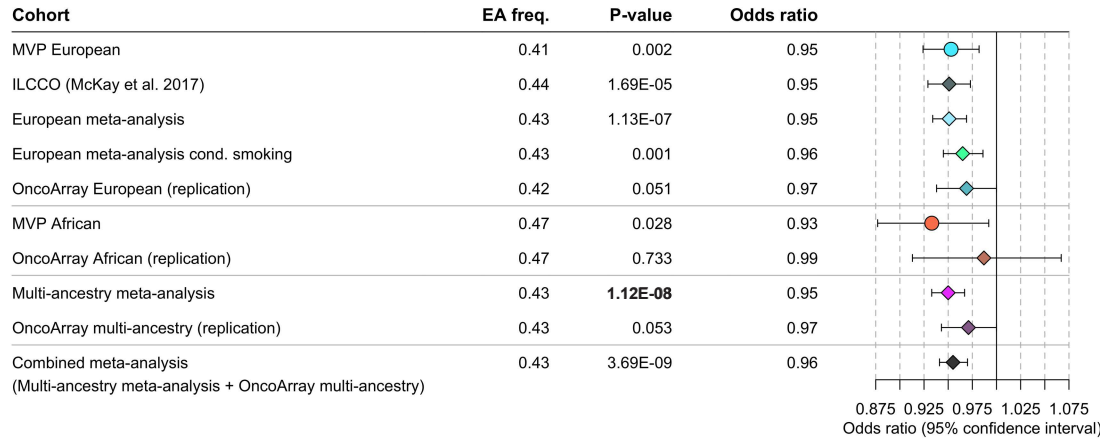


**Squamous cell carcinoma, European meta-analysis**

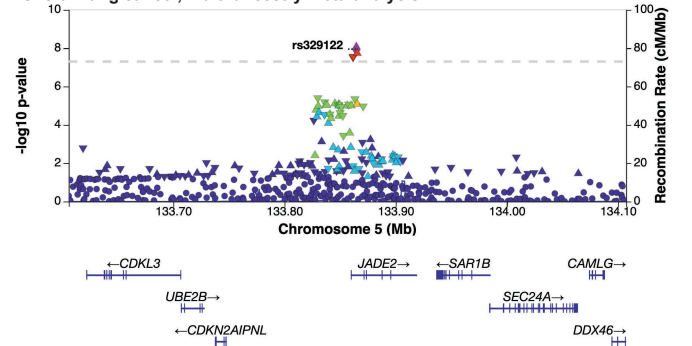


h

**Overall lung cancer, rs329122 (JADE2; 5q31.1) : Novel and replicated**

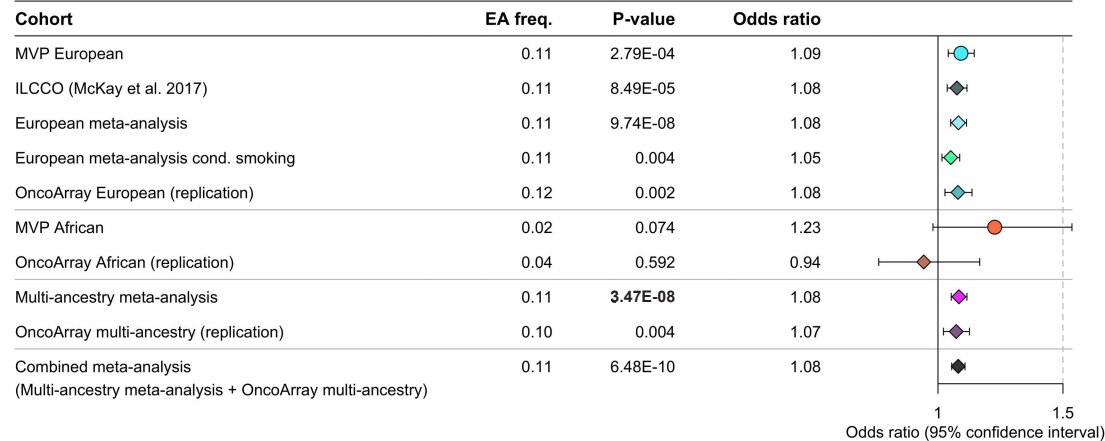


**Overall lung cancer, multi-ancestry meta-analysis**

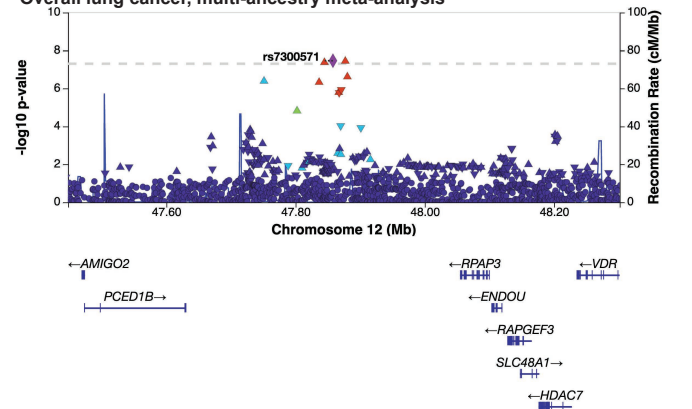


i

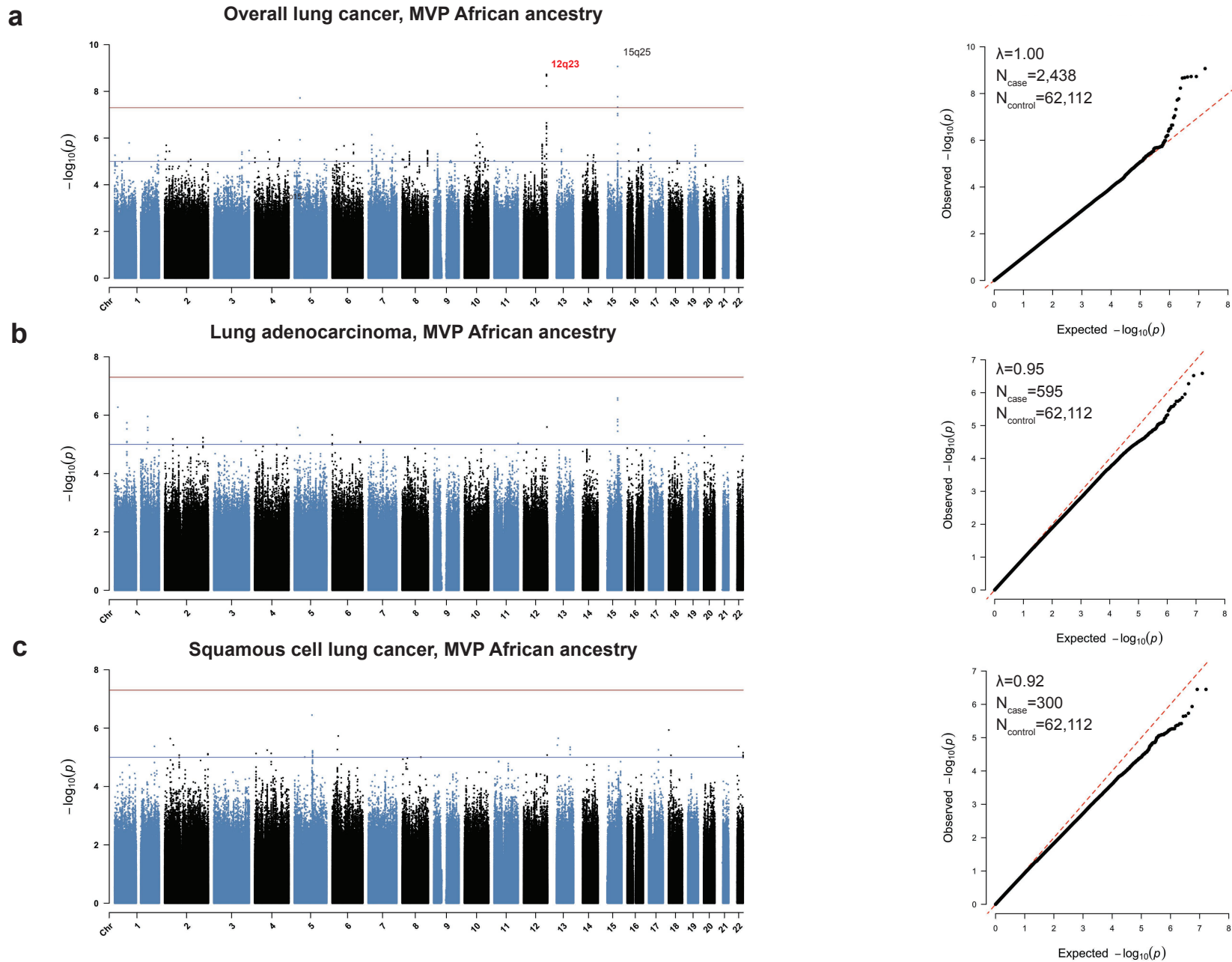
**Overall lung cancer, rs7300571 (RPAP3; 12q13.11) : Novel and replicated**



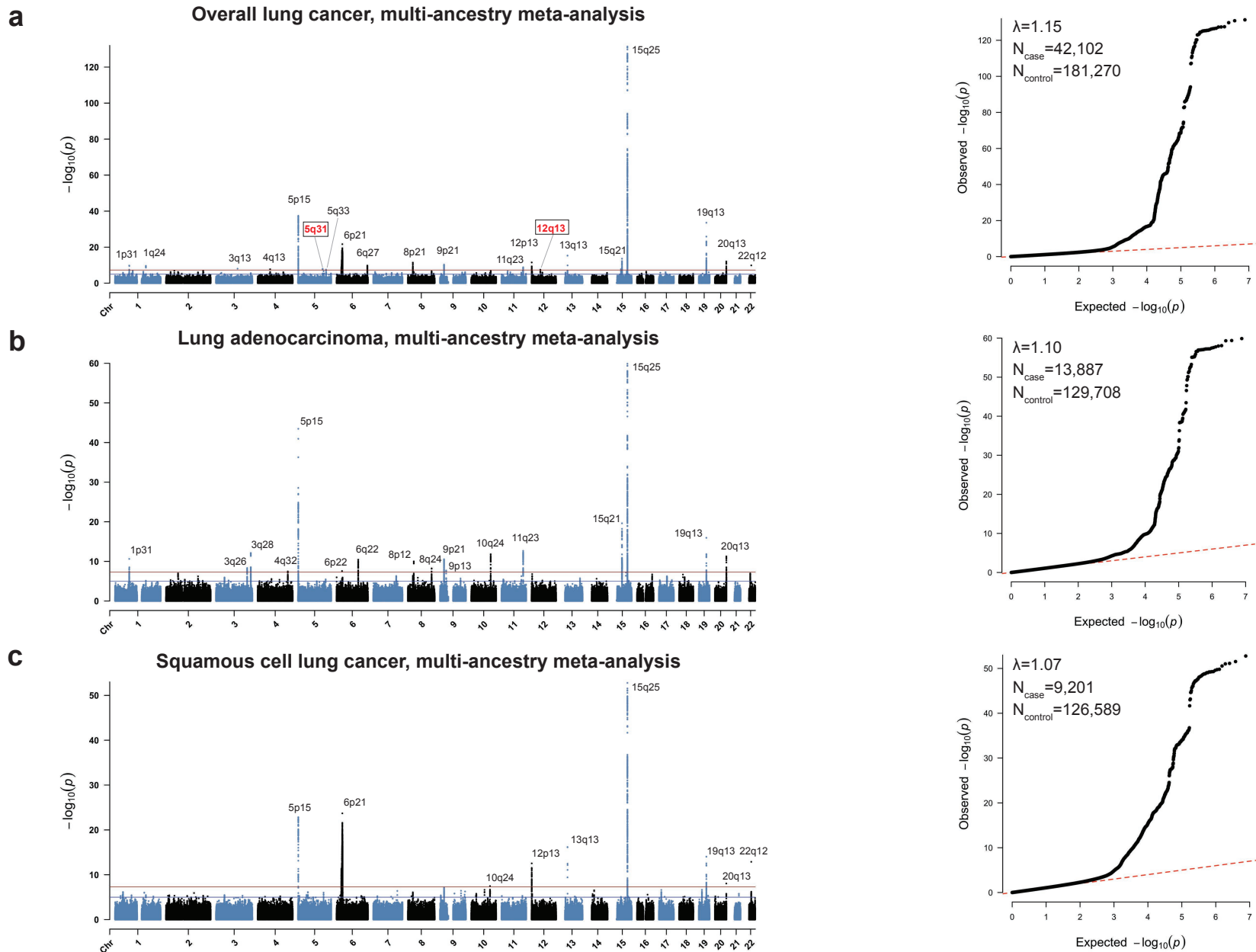
**Overall lung cancer, multi-ancestry meta-analysis**



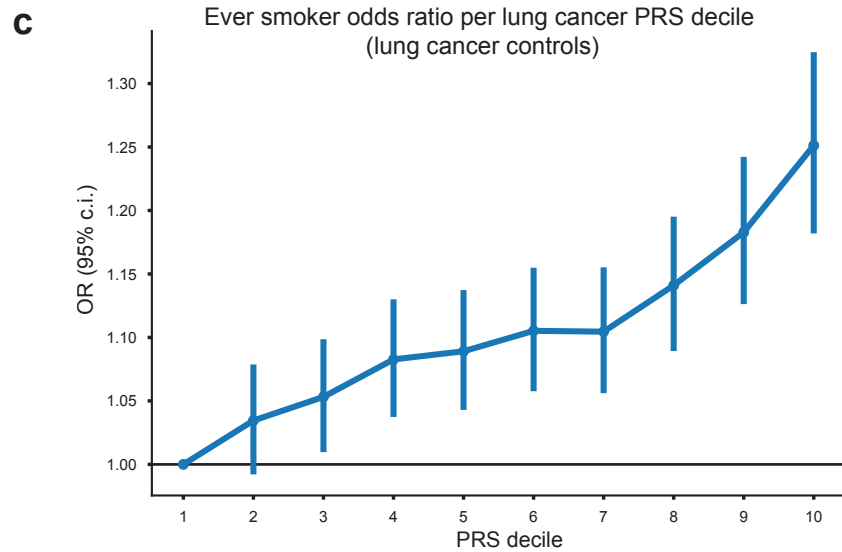
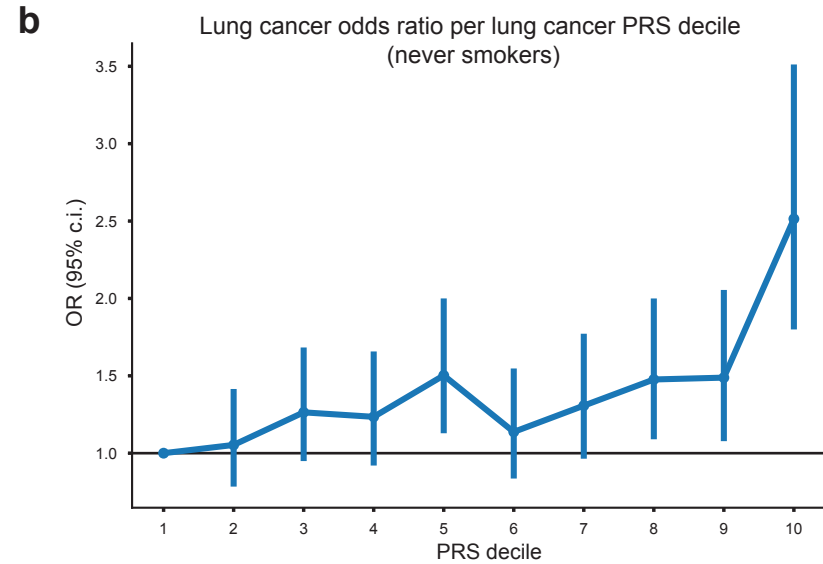
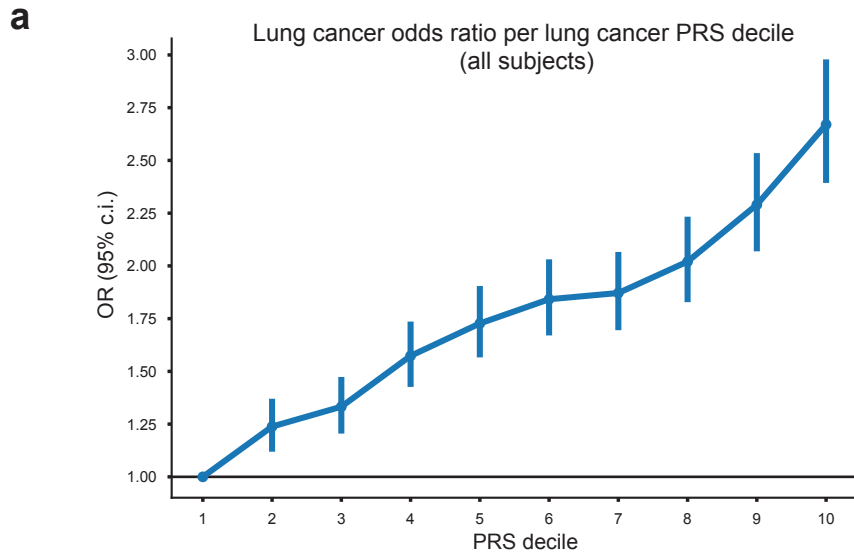
**Supplementary Fig. 4. Genome-wide significant novel lung cancer loci.** Forest plots (left) and regional Manhattan plots (right) for novel loci from European meta-analysis: **a)** *XCL2*, **b)** *LSAMP*, **c)** *NMUR2*, **d)** *TUPL3*, **e)** *MYC*, **f)** *TLE3*, and **g)** *BLOC1S2*; and from multi-ancestry meta-analysis: **h)** *JADE2*; **i)** *RPAP3*. Manhattan plot SNP data triangles pointing upwards or downwards correspond to positive or negative effect direction, based on the effect allele defined in summary statistics.



**Supplementary Fig. 5. Manhattan plots and quantile-quantile (QQ) plots for MVP African ancestry.** Manhattan and QQ plots are shown for **a**) African ancestry overall lung cancer; **b**) lung adenocarcinoma (LUAD); and **c**) squamous cell lung carcinoma (LUSC). Cytoband positions for significant loci are noted in each Manhattan plot; putatively novel loci identified in this study are in red. Genomic control ( $\lambda$ ) values and sample sizes are inset in QQ plots.

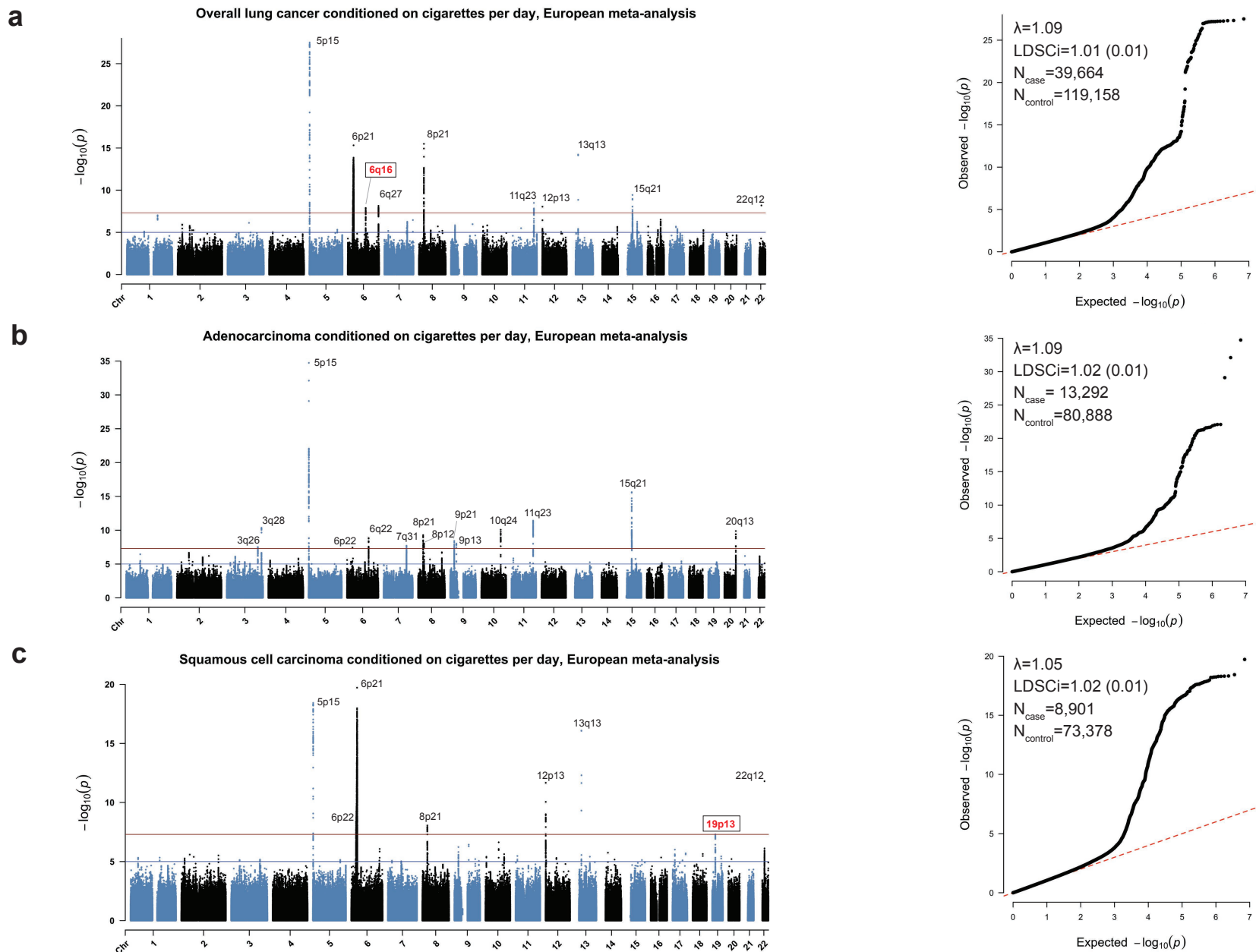


**Supplementary Fig. 6. Manhattan plots and quantile-quantile (QQ) plots for multi-ancestry meta-analyses.** Manhattan and QQ plots are shown for **a)** the multi-ancestry meta-analysis in overall lung cancer; **b)** lung adenocarcinoma (LUAD); and **c)** squamous cell lung carcinoma (LUSC). Cytoband positions for significant loci are noted in each Manhattan plot; novel loci not identified in the European meta-analysis are in red; externally replicated novel loci are indicated by a box. Genomic control ( $\lambda$ ) values and sample sizes are inset in QQ plots.



**Supplementary Fig. 7. Association of the lung cancer polygenic risk score (PRS) with lung cancer by smoking status. a)** Association of the lung cancer PRS with overall lung cancer risk. The risk of lung cancer reached an odds ratio (OR) of 2.51 (95% confidence interval: 1.80, 3.51) in the top decile. **b)** Association of the lung cancer PRS with lung cancer risk in never-smokers. Among never-smokers, lung cancer risk reached an OR of 2.67 (2.40, 2.98) in the top decile. **c)** Association of the lung cancer PRS with lung cancer risk in ever-smokers with no history of lung cancer. The top PRS decile was associated with an OR of 1.25 (1.18, 1.32).

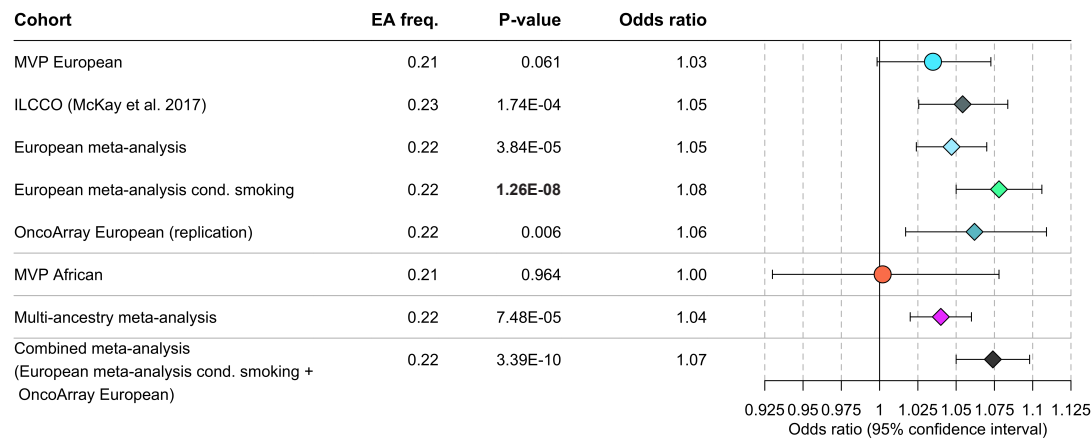




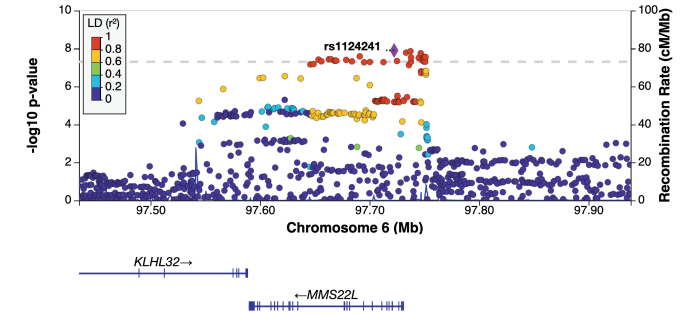
**Supplementary Fig 8. Manhattan plots and quantile-quantile (QQ) plots for European meta-analyses conditioned on cigarettes per day** Manhattan and QQ plots for **a)** overall lung cancer conditioned on cigarettes per day; **b)** lung adenocarcinoma (LUAD) conditioned on cigarettes per day; and **c)** squamous cell lung carcinoma (LUSC) conditioned on cigarettes per day. Cytoband positions for significant loci are noted in each Manhattan plot; novel loci not identified in the European meta-analysis are in red; externally replicated novel loci are indicated by a box. Genomic control ( $\lambda$ ) values, LDSC intercepts, and sample sizes are inset in QQ plots.

S9a

Overall lung cancer, rs1124241 (MMS22L; 6q16.1) : Novel and replicated

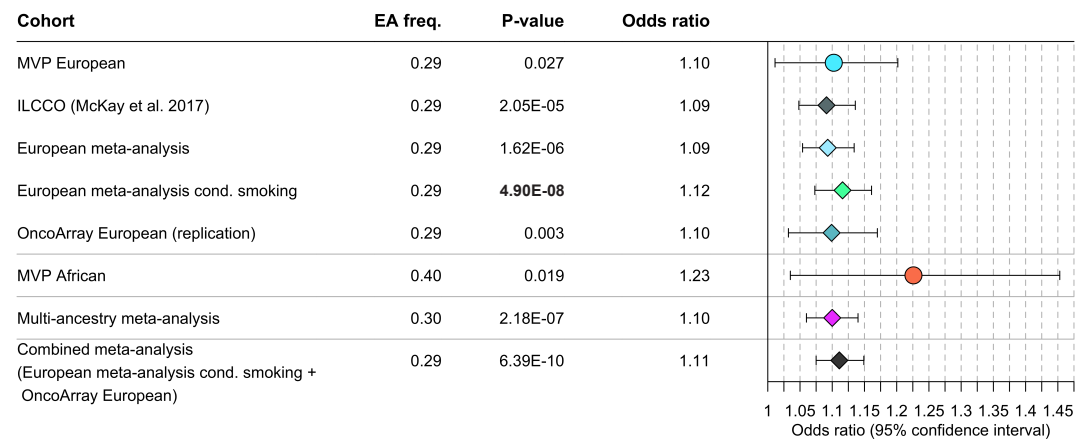


Overall lung cancer conditioned on cigarettes per day, European meta-analysis

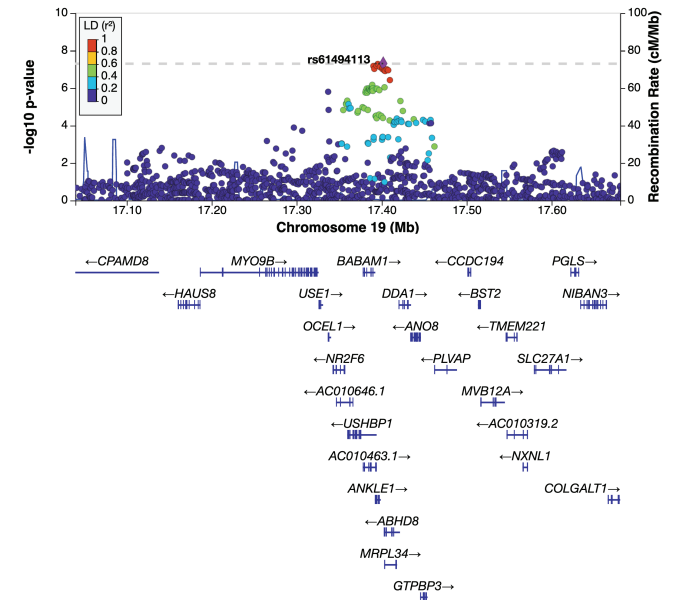


b

Squamous cell lung carcinoma, rs61494113 (ABHD8; 19p13.11) : Novel and replicated



Squamous cell lung cancer conditioned on cigarettes per day, European meta-analysis



Supplementary Fig. 9. Novel loci for overall lung cancer and squamous cell carcinoma conditioned on smoking. Locus zoom plots for the novel loci in the European meta-analysis conditioned on cigarettes per day: a) 6q16 in overall lung cancer and b) 19p13 in squamous cell lung cancer.