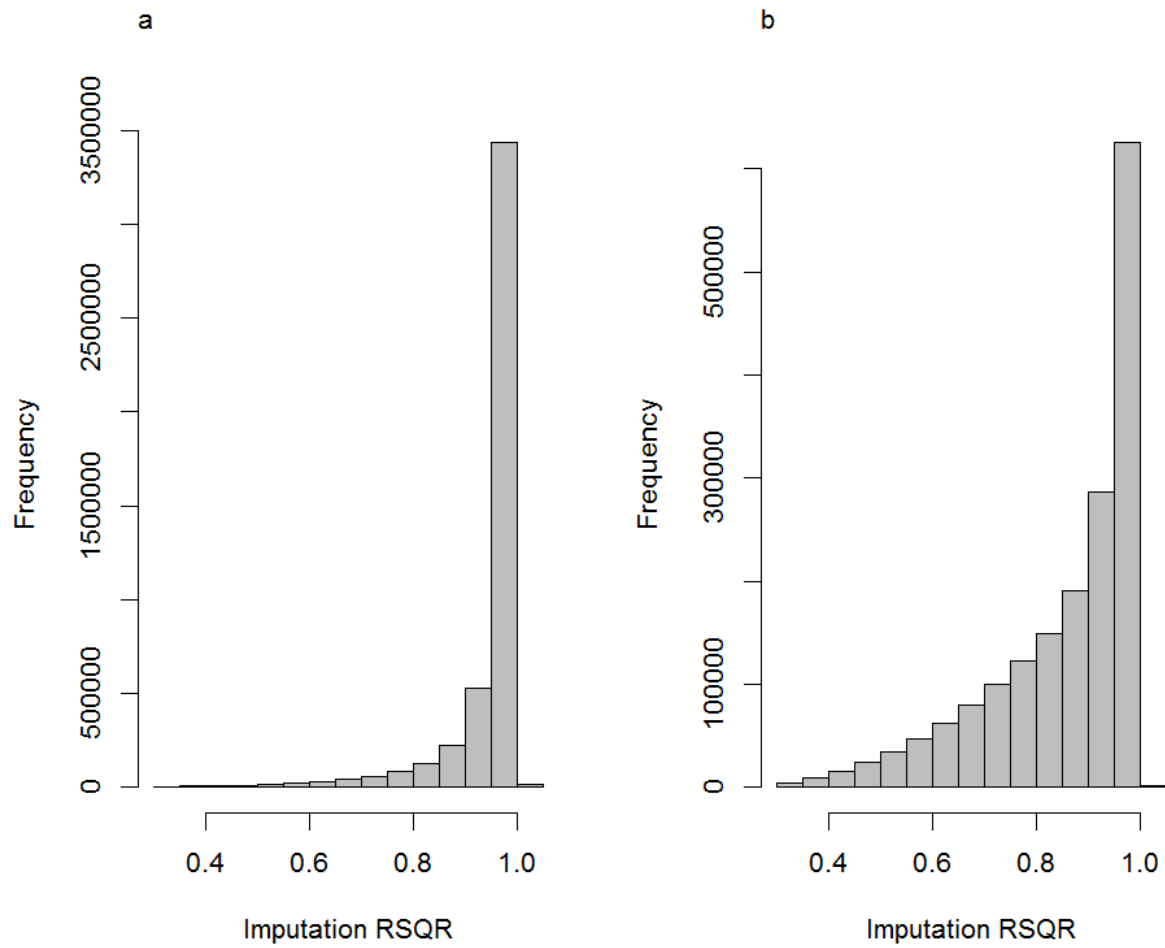


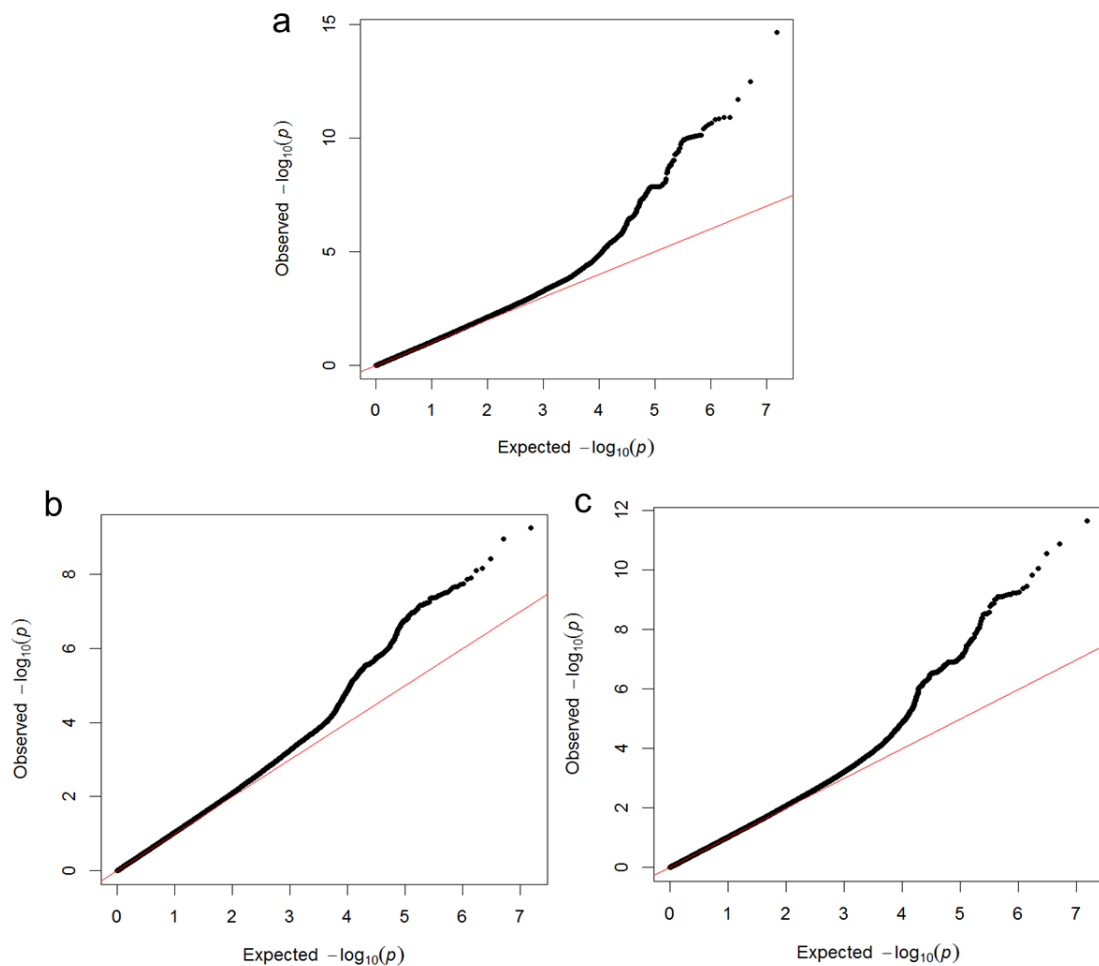
## Supplementary Figure 1

Histograms of imputation  $r^2$  (a) Variants with MAF  $\geq 0.05$  (b) Variants with MAF  $< 0.05$ .



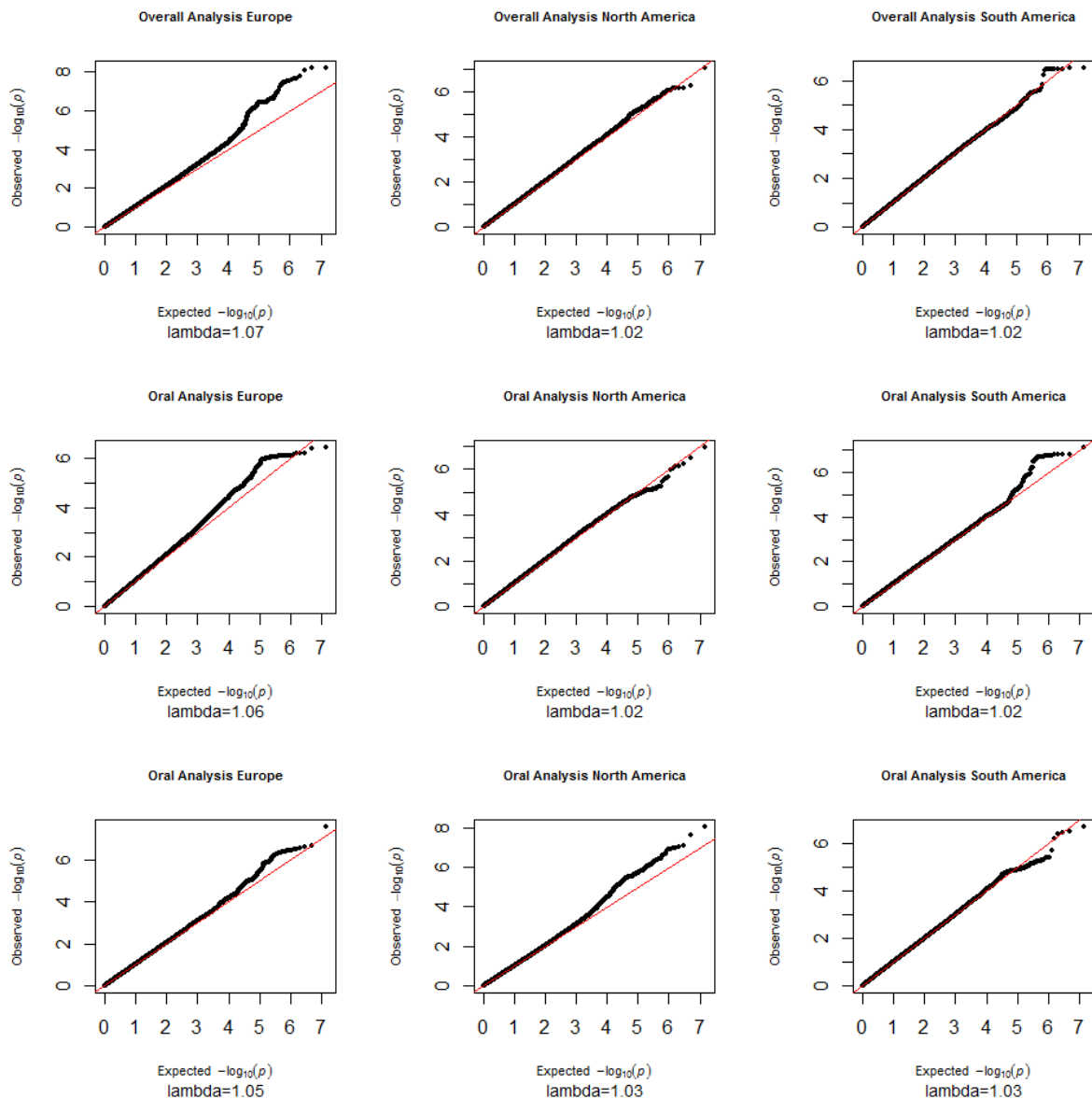
## Supplementary Figure 2

Quantile-quantile plots of the GWAS meta-analyses analyses of GWAS by geographical region Europe, South America and North America (adjusted by age, sex and region eigenvectors) (a) Overall oral cavity and pharynx cancer ( $\lambda=1.06$ ), (b) oral cancer ( $\lambda=1.05$ ), (c) oropharyngeal cancer ( $\lambda=1.04$ ).



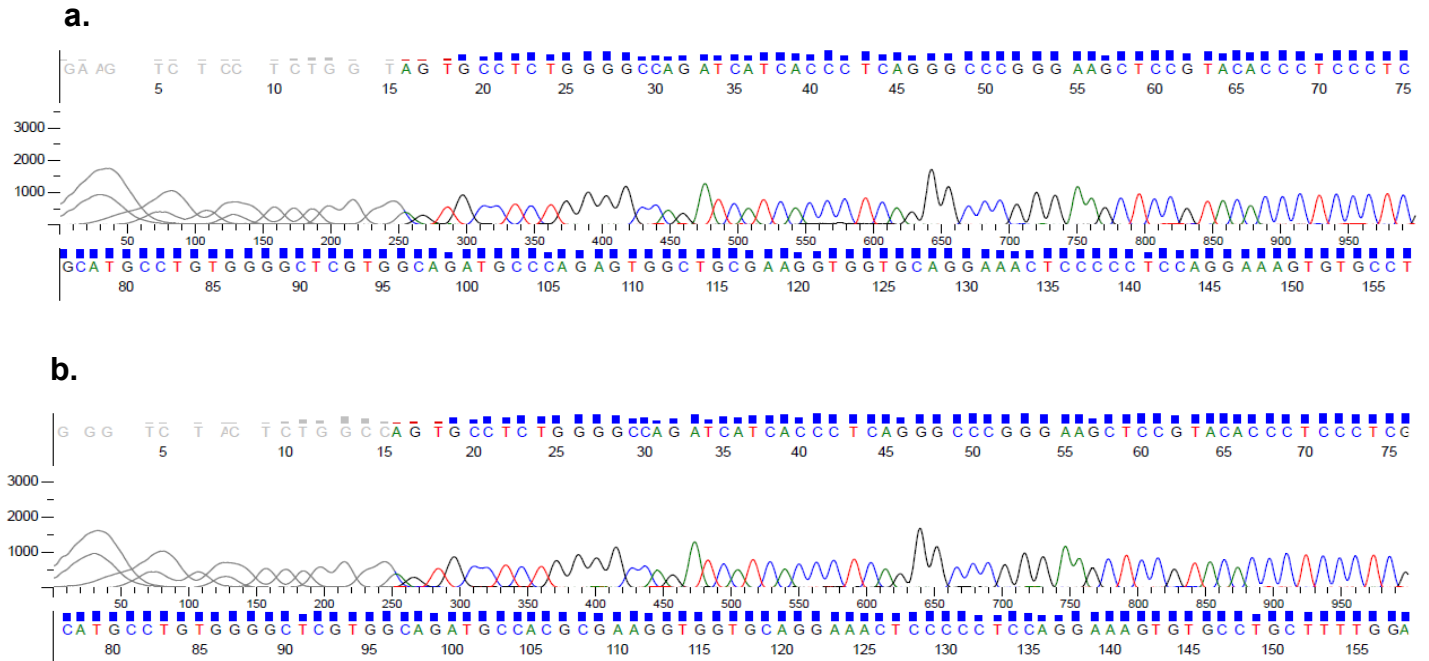
### Supplementary Figure 3

Quantile-quantile plots of by region GWAS. Top panel: overall oral and pharyngeal cancer; middle panel: oral cancer; bottom panel: oropharyngeal cancer. From left to right Europe, North America and South America. All analyses adjusted by age, sex and eigenvectors.



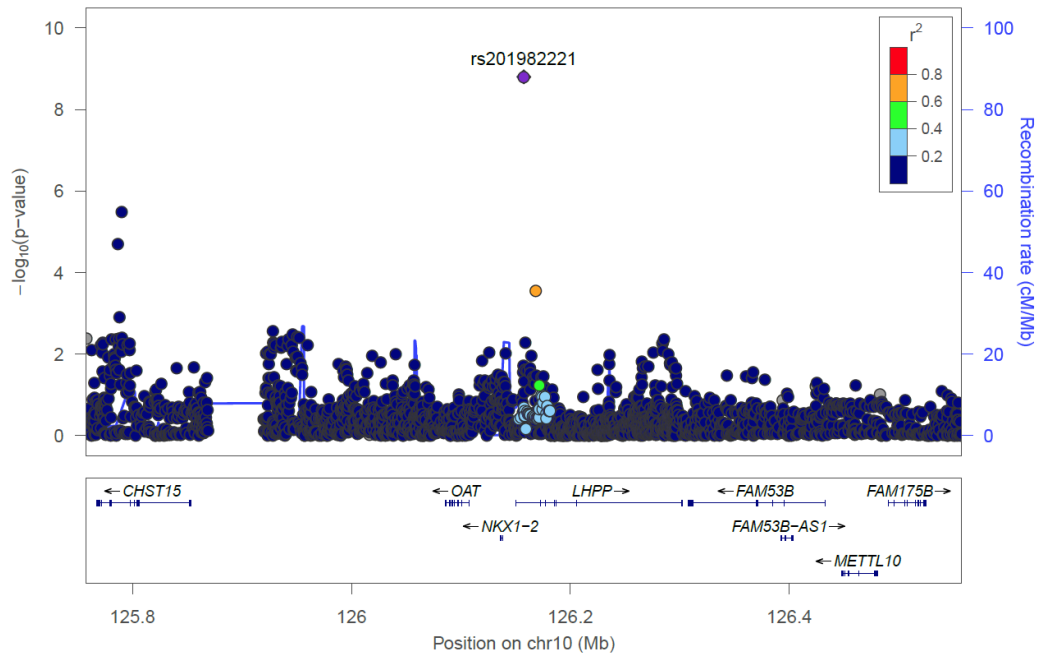
## Supplementary Figure 4

Sequence chromatogram of rs201982221 (10q26.13). a. Example of wild type insertion. b. Example of homozygous deletion. The deletion start corresponds to nucleotide 105 in the chromatogram.



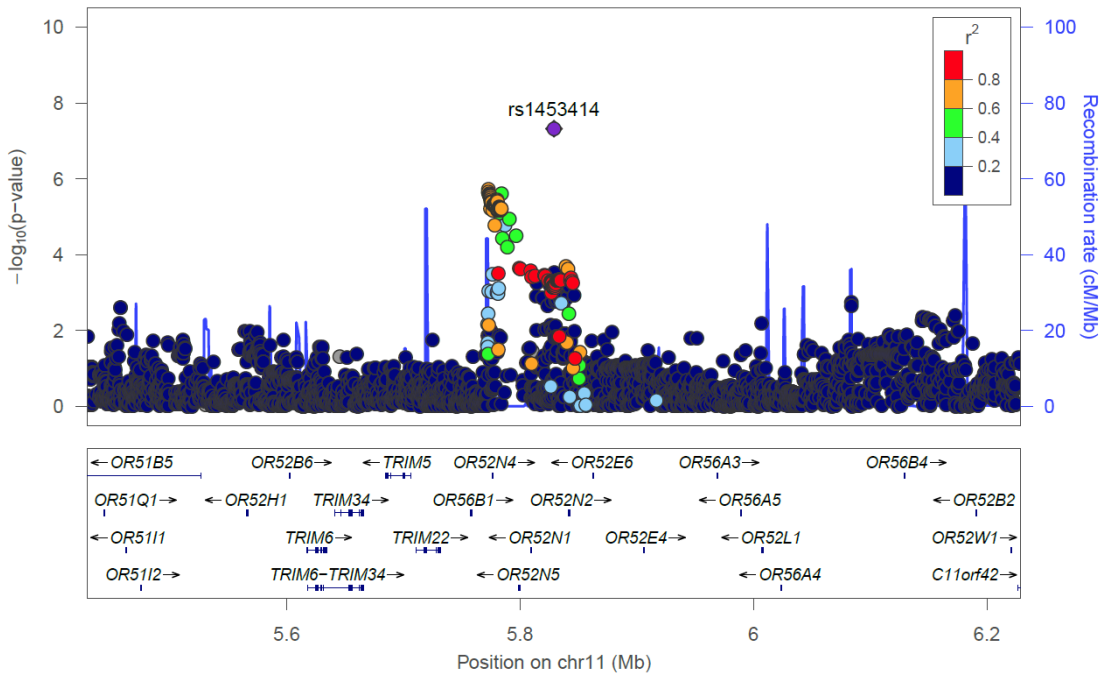
## Supplementary Figure 5

Regional association plot at 10q26.13. Chromosome positions (x-axis) and  $-\log_{10} P$ -value (y-axis) for oral cavity and pharynx cancer. LD information and recombination rates from 1000 Genomes Project November 2014 release (EUR population). Genome coordinates according to NCBI genome build 37 (Hg19). Genotyped and imputed variants colored according to their LD with the labeled SNP (purple diamond).



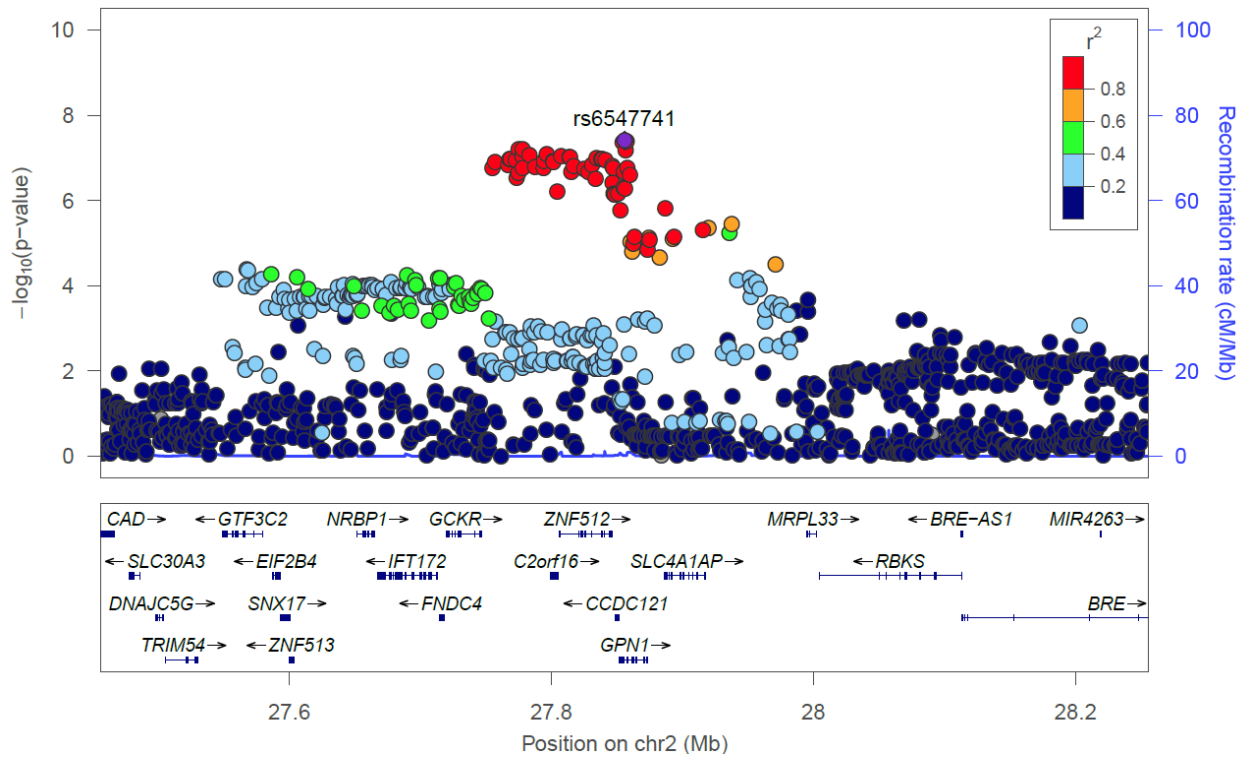
## Supplementary Figure 6

Regional association plot at 11p15.4. Chromosome positions (x-axis) and  $-\log_{10} P$ -value (y-axis) for oral cavity and pharynx cancer. LD information and recombination rates from 1000 Genomes Project November 2014 release (EUR population). Genome coordinates according to NCBI genome build 37 (Hg19). Genotyped and imputed variants colored according to their LD with the labeled SNP (purple diamond).



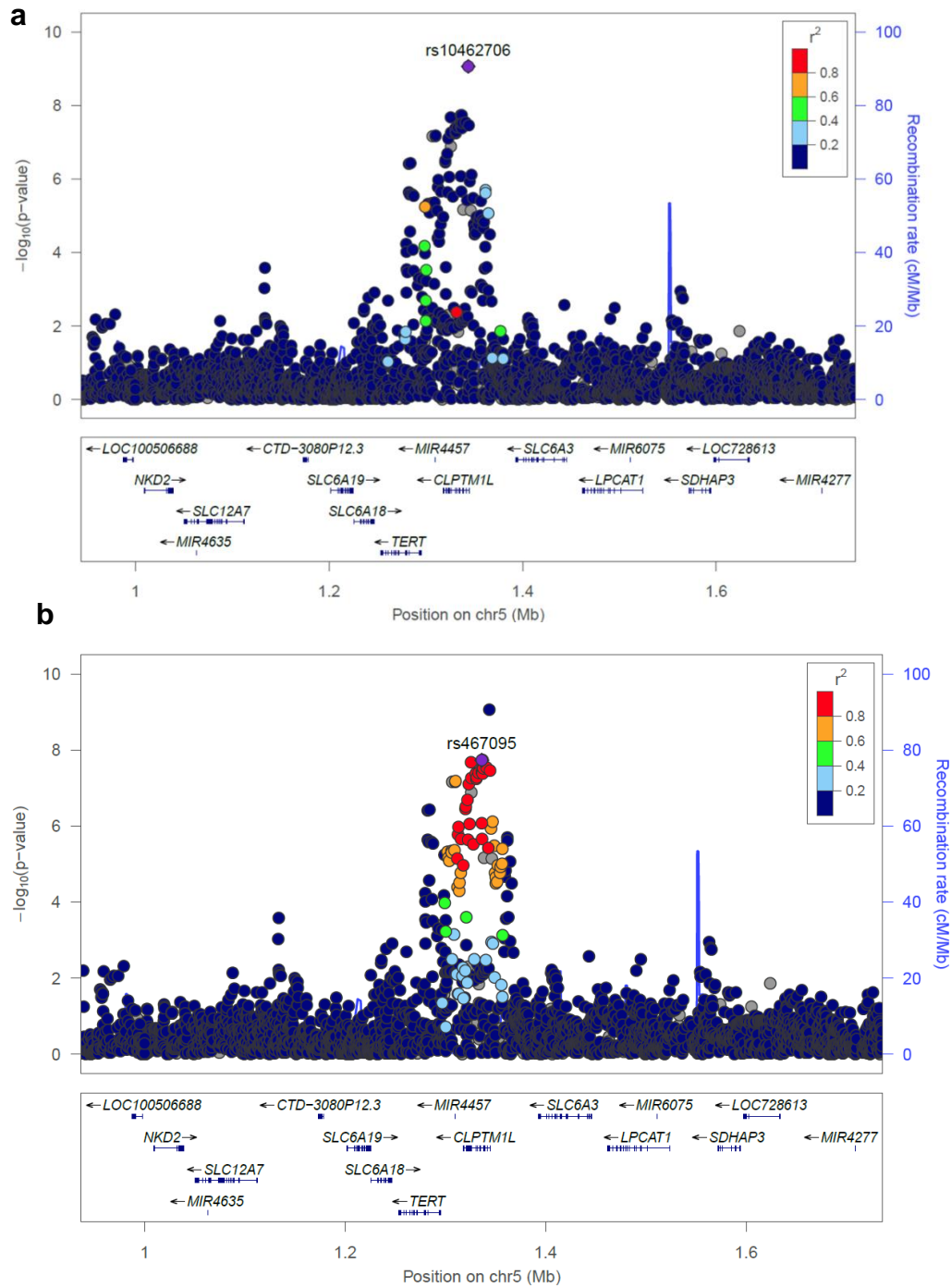
## Supplementary Figure 7

Regional association plot at 2p23.3, Chromosome positions (x-axis) and  $-\log_{10} P$ -value (y-axis) for oral cancer LD information and recombination rates from 1000 Genomes Project November 2014 release (EUR population). Genome coordinates according to NCBI genome build 37 (Hg19). Genotyped and imputed variants colored according to their LD with the labeled SNP (purple diamond).



## Supplementary Figure 8

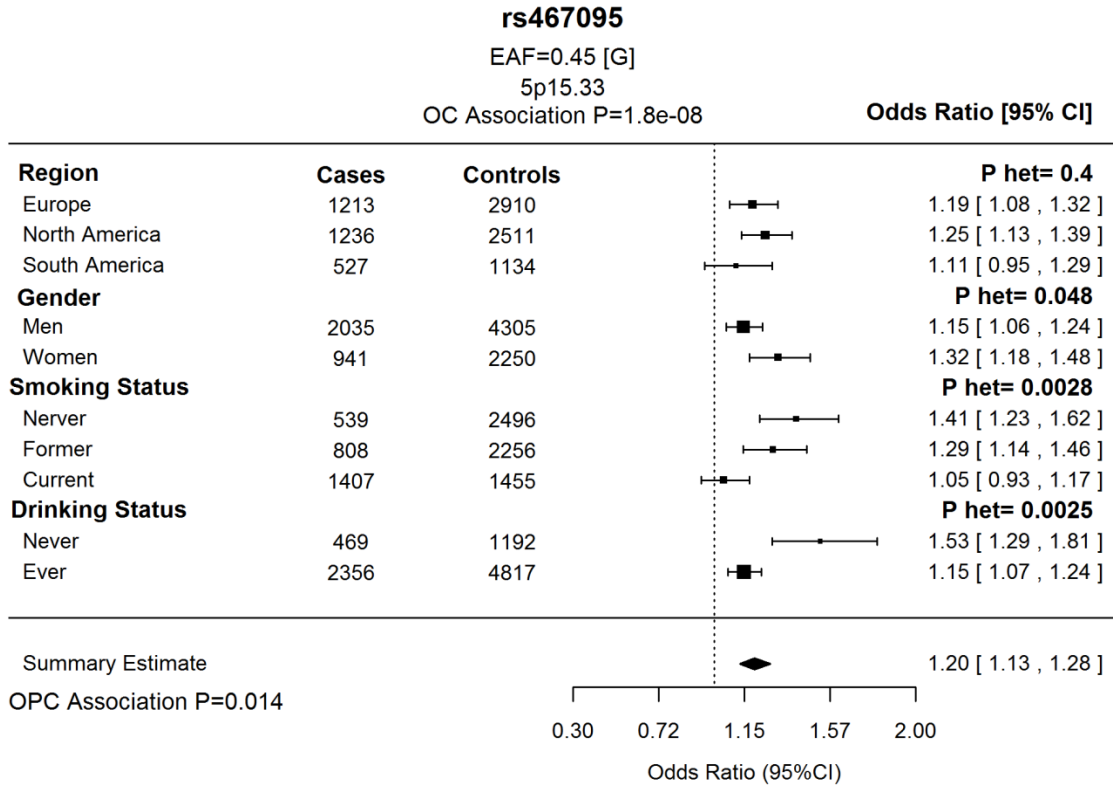
Regional association plots at 5p15.33. (a) rs10462706 (b) rs467095 (2<sup>nd</sup> strongest association). Chromosome positions (x-axis) and  $-\log_{10} P$ -value (y-axis) plotted for oral cancer. LD information and recombination rates from 1000 Genomes Project November 2014 release (EUR population). Genome coordinates according to NCBI genome build 37 (Hg19). Genotyped and imputed variants colored according to their LD with the labeled SNP (purple diamond).





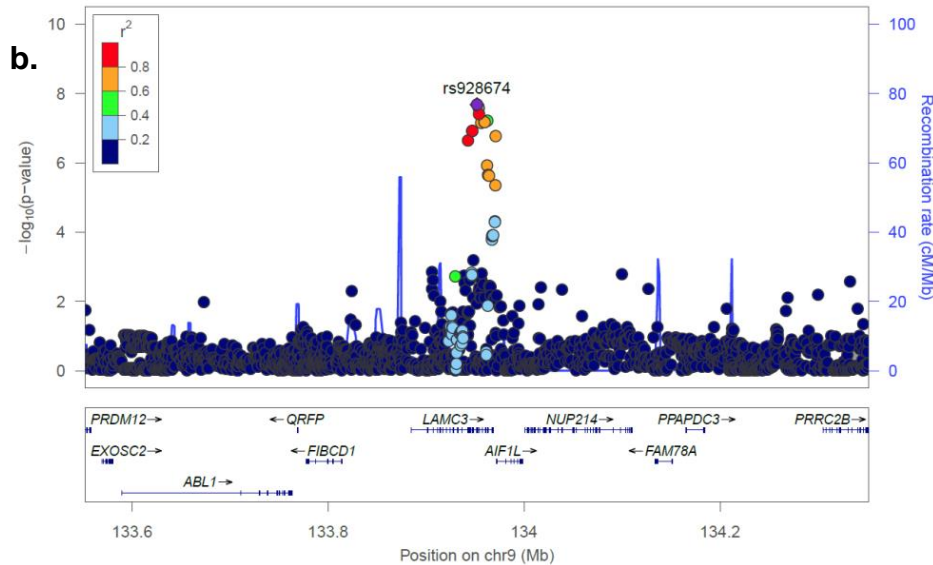
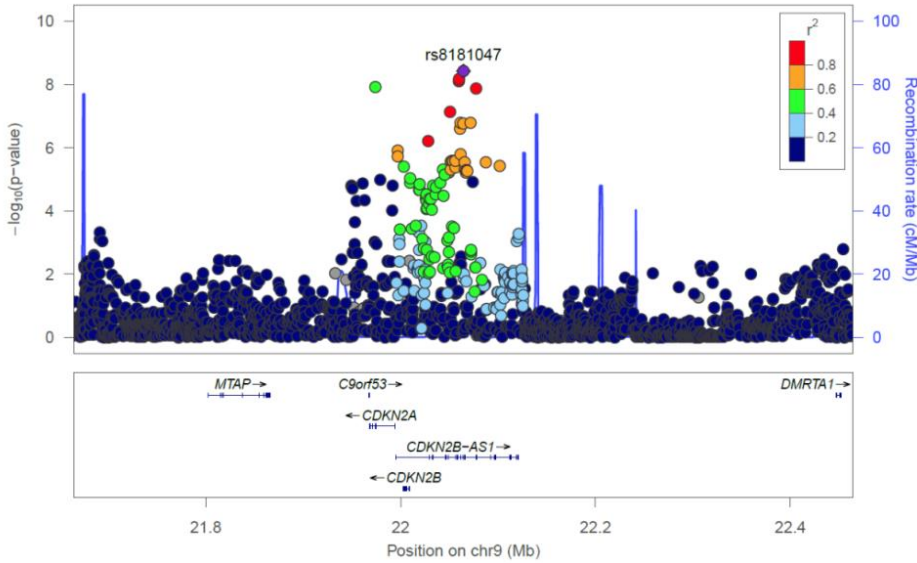
### Supplementary Figure 9

Forest plot of odds ratios for oral cancer association and 5q15.3 rs467095 *CLPTM1*, EAF=effect allele frequency in 6585 controls.



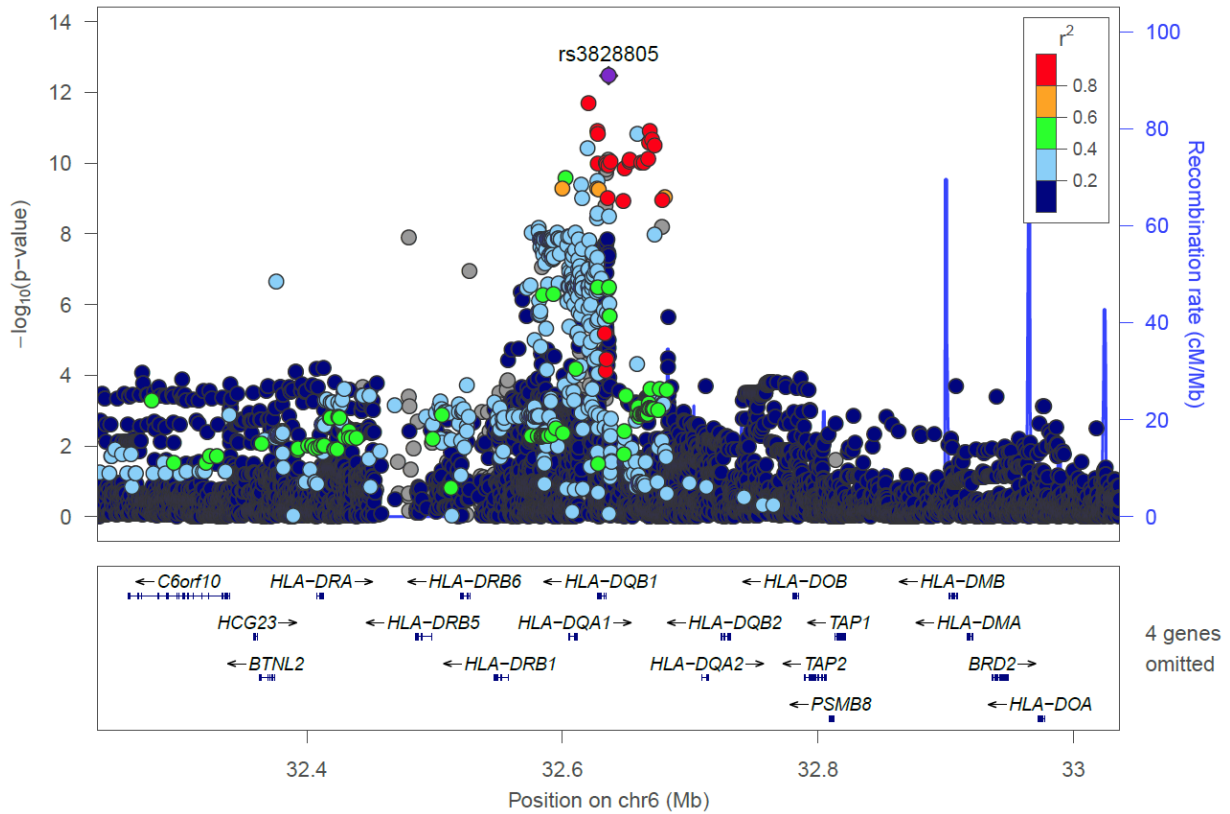
### Supplementary Figure 10

Regional association plots. a) rs8181047 at 9p21.3, b) rs928674 at 9q34.12. Chromosome positions (x-axis) and the  $-\log_{10} P$ -values (y-axis) for oral cancer. LD and recombination rates from 1000 Genomes Project November 2014 release (EUR population). Genome coordinates according to NCBI genome build 37 (Hg19). The plots show genotyped and imputed variants colored according to their LD with the labeled SNP (purple diamond).  
**a.**



## Supplementary Figure 11

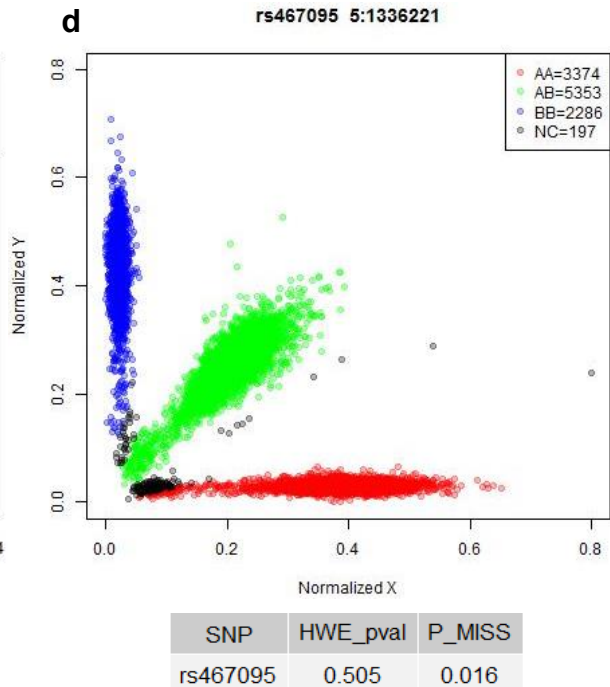
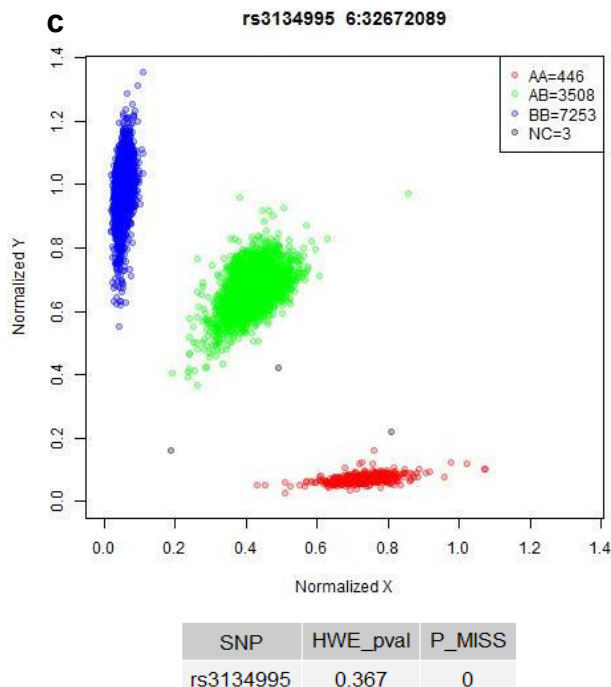
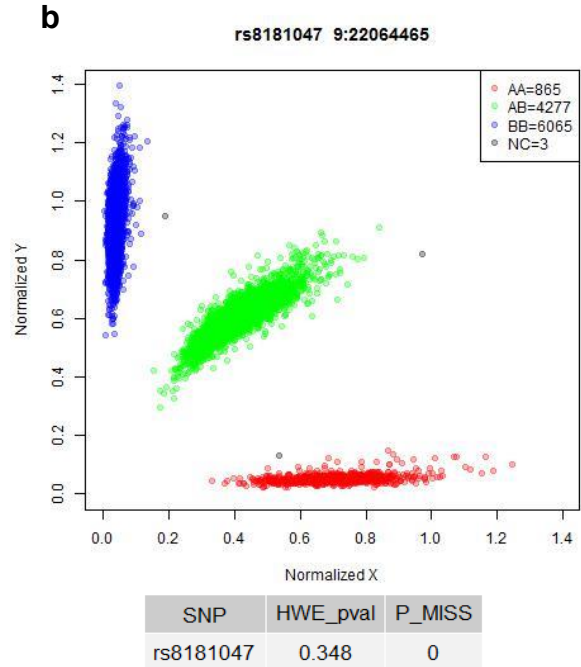
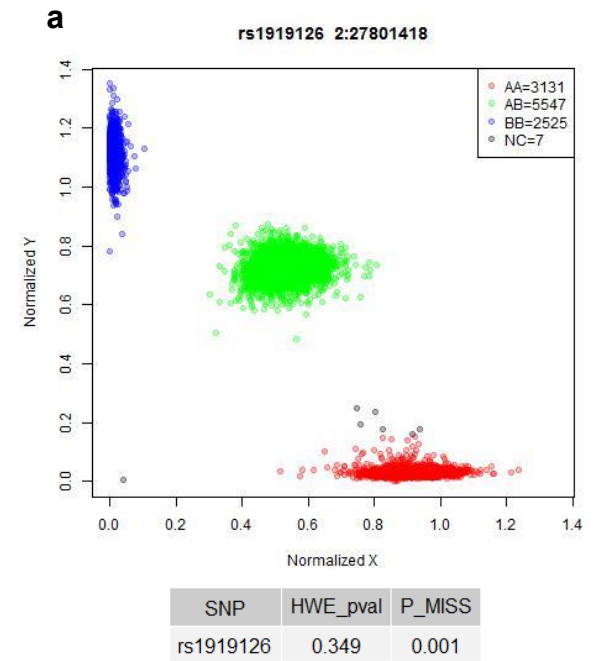
Regional association plot at 6p21.32. Chromosome positions (x-axis) and  $-\log_{10}$  P-values (y-axis) for combined oral and pharyngeal cancer. LD and recombination rates from 1000 Genomes Project November 2014 release (EUR population). Genome coordinates according to NCBI genome build 37 (Hg19). Genotyped and imputed variants colored according to their LD with the labeled SNP rs3828805 the lead variant (purple diamond).

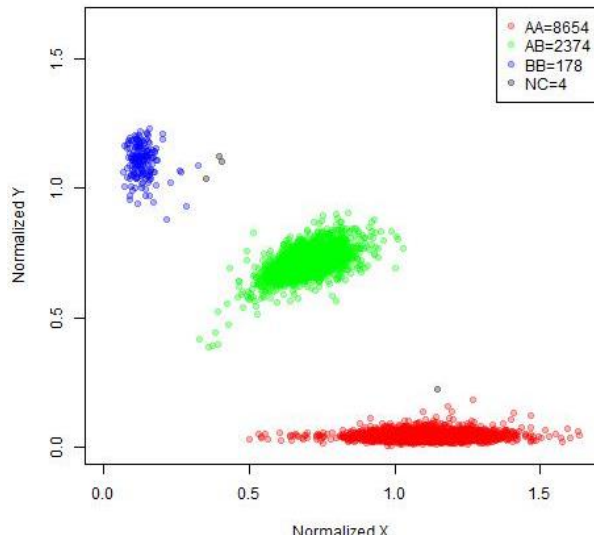


## Supplementary Figure 12

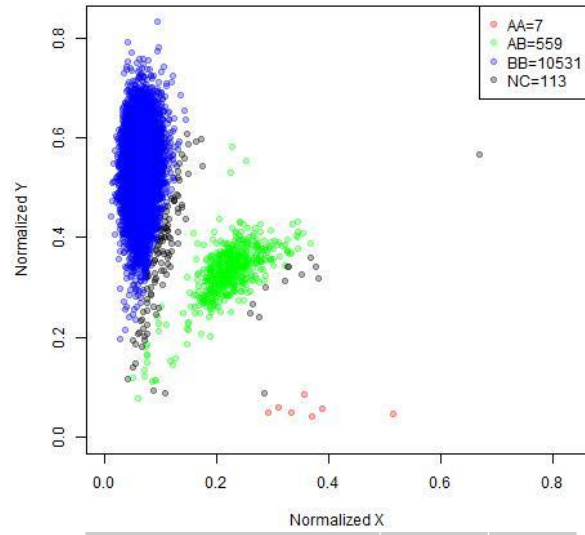
### Genotype cluster plots of top loci

(a) 2p23 rs1919126; (b) 9p21.3, rs8181047; (c) 6p21.32, rs3134995; (d) 5p15.3, rs467095; (e) 9q34, rs199717881/chr9\_133953882\_A\_C; (f) 10q26, rs201982221/chr10\_126157446\_CAG\_INDEL; (g) 11p15, rs1453414/chr11\_5829084\_G\_T; (h) 5p14, rs79767424/chr5\_19108690\_G\_T, (SNP not validated by Taqman). In tables, Hardy Weinberg *P*-value (controls) and percent of genotypes with no call (*P*\_MISS).

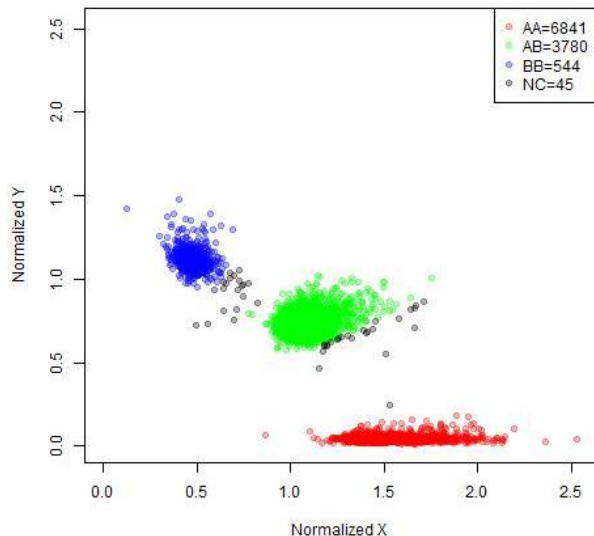


**e****chr9\_133953882\_A\_C 9:133953882**

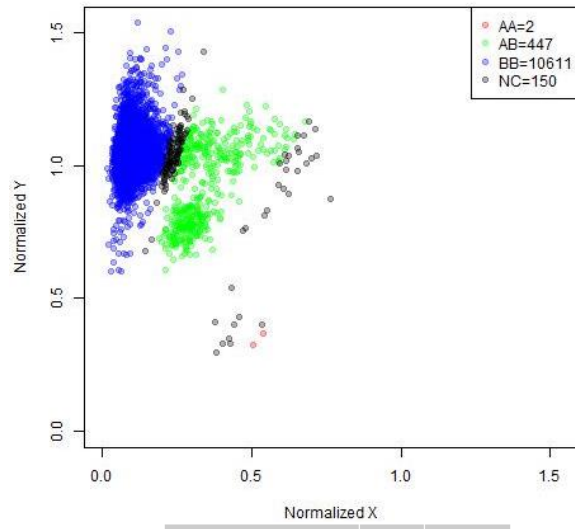
SNP	HWE_pval	P_MISS
chr9_133953882_A_C	0.4666	2e-04

**f****chr10\_126157446\_CAG\_INDEL 10:126157446**

SNP	HWE_pval	P_MISS
chr10_126157446_CAG_INDEL	0.638	0.009

**g****chr11\_5829084\_G\_T 11:5829084**

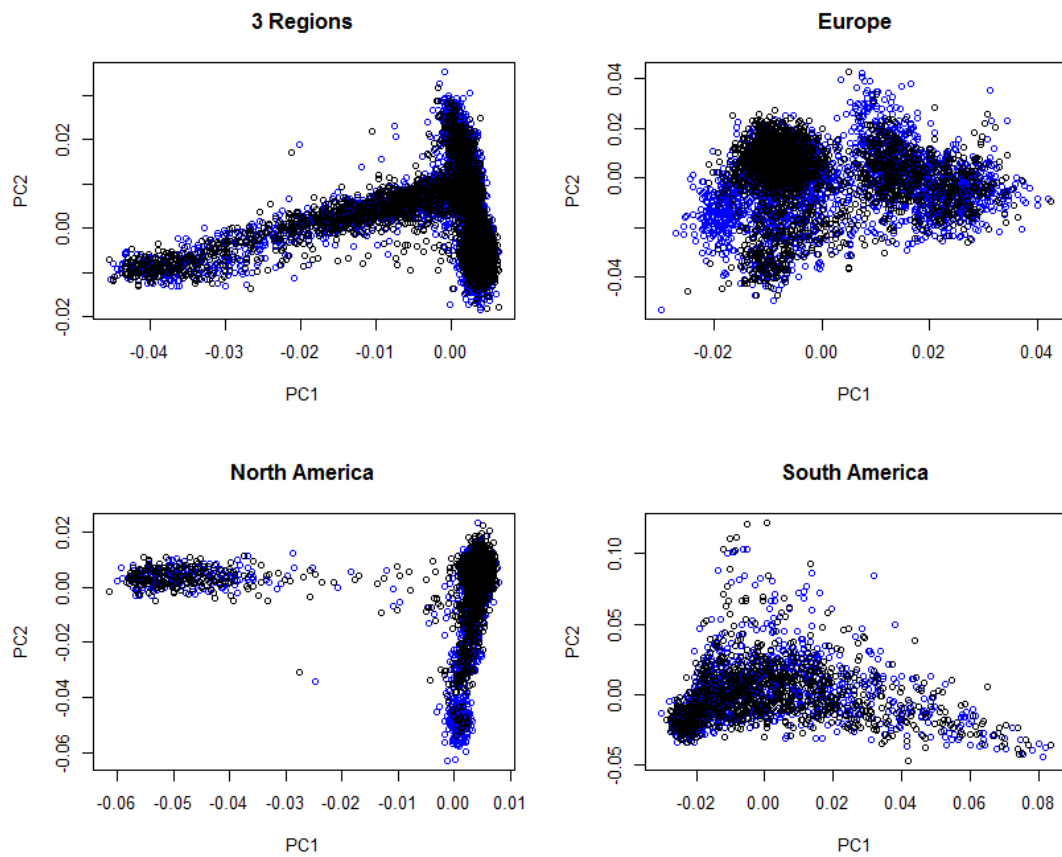
SNP	HWE_pval	P_MISS
chr11_5829084_G_T	0.863	0.002

**h****chr5\_19108690\_G\_T 5:19108690**

SNP	P_av	P_MISS
chr5_19108690_G_T	0.438	0.02

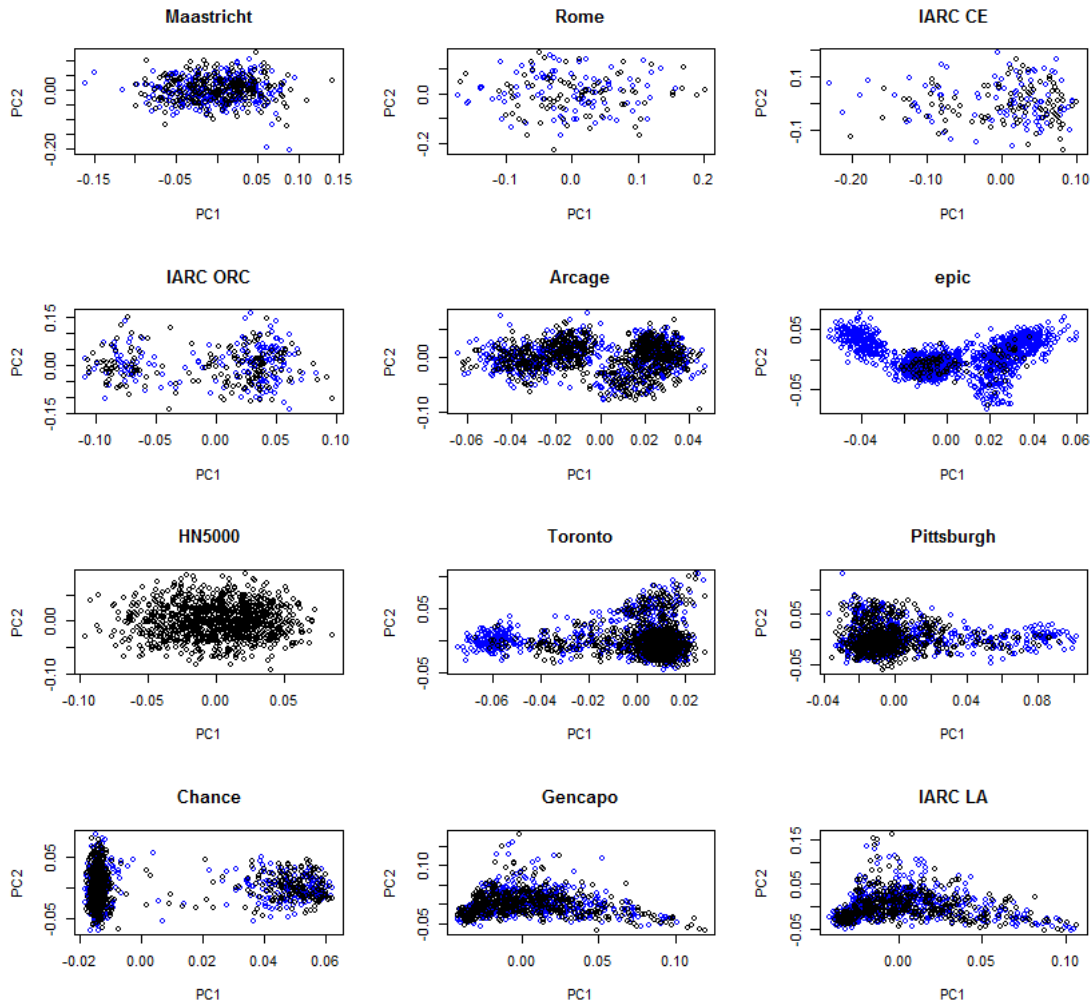
### Supplementary Figure 13

**Principal Components Analyses (PCA) of all the study participants and within regions.** Principal component 1 is displayed in the x-axis and principal component 2 in the y-axis. Blue dots are cases and black dots are controls.



## Supplementary Figure 14

**Principal Components Analyses (PCA) within epidemiological study.** Principal component 1 is displayed in the x-axis and principal component 2 in the y-axis. Blue dots are cases and black dots are controls.



**Supplementary Table 1**

Study Name	Control/Case	Tumor site*	Location	Principal Investigator	Study Design	Reference
Carolina Head and Neck Cancer Epidemiology (Chance)	736/727	OC n=352 OPC n=345 HPC n=30	USA (North Carolina)	AF Olshan	Population-based case-control	Divaris et al. Oral health and risk for head and neck squamous cell carcinoma: the Carolina Head and Neck Cancer Study Cancer Causes Control (2010) 21:567–575 PMID: 20049634
Gencapo	535/529	OC n=291 OPC n=187 HPC n=29	Brazil	V Wunsch-Filho	Multicenter Hospital-based case-control	Lopez et al. Human papillomavirus (HPV) 16 and the prognosis of head and neck cancer in a geographical region with a low prevalence of HPV infection Cancer Causes Control (2014) 25:461–471 PMID: 24474236
Alcohol-Related Cancers and Genetic Susceptibility in Europe (Arcage)	729/735	OC n=414 OPC n=239 HPC n=52	Croatia Czech Republic Germany Greece Ireland Italy Norway Spain UK	P. Brennan	Multicenter Hospital-based case-control	Lagiou et al. Alcohol-related cancers and genetic susceptibility in Europe: the ARCAGE project: study samples and data collection. European Journal of Cancer Prevention 2009, 18:76–84 PMID: 18830131
IARC Central Europe	100/102	OC n=61 OPC n=14 HPC n=5	Poland Romania Russia Slovakia	P. Brennan	Multicenter Hospital-based case-control	Hashibe et al. Evidence for an Important Role of Alcohol- and Aldehyde-Metabolizing Genes in Cancers of the Upper Aerodigestive Tract. Cancer Epidemiol Biomarkers Prev 2006;15(4). April 2006 PMID: 16614111
European Prospective Investigation into Cancer and Nutrition	1570/106	OC n=56 OPC n=42 HPC n=5		P. Brennan	Multicenter cohort study	Riboli et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutrition: 5(6B),



(EPIC)						
IARC Latin America	600/457	OC n=236 OPC n=157 HPC n=51	Brazil, Argentina	P. Brennan	Multicenter hospital-based case-control	1113–1124.PMID: 12639222  Szymanaska et al. Alcohol and tobacco, and the risk of cancers of the upper aerodigestive tract in Latin America: a case-control study. <i>Cancer Causes Control</i> (2011) 22:1037–1046. PMID: 21607590  Herrero et al. Human Papillomavirus and Oral Cancer: The International Agency for Research on Cancer Multicenter Study. <i>JNCI</i> , Vol. 95, No. 23, Dec 2003. PMID: 14652239
IARC oral cancer (ORC)	175/168	OC n=100 OPC n=68 HPC n=0	Canada Italy Poland Spain	S. Franceschi	Multicenter hospital-based case-control	Lacko et al. Genetic polymorphism in the conjugating enzyme UGT1A1 and the risk of head and neck cancer. <i>Int. J. Cancer</i> : 127, 2815–2821 (2010). PMID: 21351260  Troy et al. Childhood passive smoke exposure is associated with adult head and neck cancer. <i>Cancer Epidemiology</i> 37 (2013) 417–423. PMID: 23619143.
Maastricht	275/280	OC n =103 OPC n=161 HPC n=11	Netherlands	M. Lacko, W. Peters	Hospital-based case-control	De Feo et al. A Case-Control Study on the Effect of Apolipoprotein E Genotype on Head and Neck Cancer Risk. <i>Cancer Epidemiol Biomarkers Prev</i> ; 19(11) November 2010. PMID: 20861397
University of Pittsburgh	820/786	OC n=464 OPC n=306 HPC n=16	USA (Pittsburgh)	B Diergaard	Hospital-based case-control	Wong et al 2014 Two BRM promoter insertion polymorphisms increase the risk of early-stage upper
Rome	93/87	OC n=56 OPC n=17 HPC n=11	Italy	S. Boccia	Hospital-based case-control	
Toronto	952/1023	OC n=412 OPC n=553 HPC n=58	Canada	R. Hung	Hospital-based case-control	

Head and neck 5000 (HN5000)	0/1034	OC n=445 OPC n=552 HPC n=37	UK	A. Ness	Multicenter HNC clinical cohort	aerodigestive tract cancers. Cancer Medicine 2014; 3(2): 426–433. PMID: 24519853 Ness et al. Establishing a large prospective clinical cohort in people with head and neck cancer as a biomedical resource: head and neck 5000. Ness et al. BMC Cancer 2014, 14:973 PMID: 25519023
-----------------------------	--------	-----------------------------------	----	---------	---------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

\*OC=Oral cavity cancer; OPC= Oropharyngeal cancer; HPC=hypopharynx cancer

## Supplementary Table 2

Individual exclusions in HNC OncoArray quality control (QC)

<b>Exclusion criteria</b>	<b>Cases</b>	<b>Controls</b>	<b>Total</b>
Initial dataset	6289	6818	13107
Low success rates (<95%)	0	15	15
Sex mismatch	18	34	52
Outlier heterozygosity	39	22	61
Expected duplicate	36	67	103
Unexpected duplicates	32	42	74
Unexpected Relatives IBD >0.3	4	40	44
<b>By Geographic Region QC</b>			
Europe PCA outliers	17	2	19
North America PCA outliers	108	4	112
South America PCA outliers	1	7	8
<b>Included in analysis</b>			
Europe	2499	2928	5427
North America	2549	2522	5071
South America	986	1135	2121
Total	6034	6585	12619

### Supplementary Table 3

Genetic ancestry<sup>a</sup> of individuals by region

Region	>70% CEU		>70%YRI		>30 admix		total
	n	%	n	%	%		
Europe	5425	100	0	0	2	0.04	5427
North America	4671	92	298	5.9	102	2	5071
South America	1340	63	63	3	718	34	2121
Total	11436	90.6	361	2.9	822	6.5	12619

<sup>a</sup> Analysis in STRUCTURE 2.3.4 software

Supplemental tables 4-11 (excel files)

## Supplementary Table 12

Expression Quantitative Trait Loci (eQTL)<sup>1</sup> for lead SNPs within each associated region ( $P < 5 \times 10^{-8}$ )

Region	rs	Gene	Functional class	Allele	Tissue	eQTL Gene	Effect Size	P
2p23.3	rs6547741	GPN1	intronic	A/G	Esophagus - Mucosa	AC074117.10	0.23	0.000018
					Muscle - Skeletal	SNX17	0.33	8.60E-15
					Spleen	FNDC4	-0.69	3.60E-12
					Spleen	GCKR	-0.81	9.50E-11
					Skin - Sun Exposed (Lower leg)	NRBP1	0.29	6.50E-07
					Lung	NRBP1	0.22	0.000014
					Cells - Transformed fibroblasts	AC074117.10	0.23	0.000035
					Esophagus - Mucosa	TERT	-0.3	1.70E-06
5p15.33	rs467095	CLPTM1L	intronic	G/A	Esophagus - Mucosa	TERT	-0.3	1.70E-06
9q34.12	rs928674	LAMC3	intronic	G/A	Esophagus - Mucosa	AIF1L	0.25	4.2E-06
11p15.4	rs1453414	OR52N2/ TRIM5	intronic	C/A	cerebellar cortex	TRIM5, TRIM22	NA	2.53E-06

<sup>1</sup>eQTL surveyed from the Genotype-Tissue Expression (GTEx) project and HaploReg.  
NA=not available

### Supplementary Table 13

Sensitivity analysis of lead SNP per locus in individuals with >70% Caucasian ancestry

Region	rs	Europe		North America		South America		Frq. <sup>b</sup> case	Frq. <sup>b</sup> contr ol	Meta-analysis <sup>a</sup>	
		OR	P	OR	P	OR	P			OR	P
<b>Overall oral and pharynx cancer</b>											
4q23	rs1229984	0.50	1.0E-09	0.75	1.4E-02	0.43	3.8E-05	0.03	0.06	0.58	7.2E-13
5p14.3	rs79767424	0.42	2.7E-06	0.66	1.7E-02	0.63	1.2E-01	0.01	0.02	0.55	2.3E-07
6p21.32	rs3828805	1.39	6.2E-09	1.29	7E-07	1.11	3E-01	0.75	0.72	1.30	9.8E-14
11p15.4	rs1453414	1.22	1.5E-04	1.21	3.5E-04	1.15	1.5E-01	0.22	0.19	1.21	7.4E-08
10q26.1	rs201982221	1.87	1.3E-06	1.51	1.5E-03	1.33	3.0E-01	0.03	0.02	1.64	1.2E-08
<b>Oral cancer</b>											
2p23.3	rs6547741	0.84	6.7E-04	0.81	8.9E-05	0.97	7.8E-01	0.50	0.54	1.23	2.5E-08
5p15.33	rs10462706	0.75	1.7E-04	0.74	1.3E-04	0.64	2.3E-03	0.12	0.16	0.73	1.1E-09
5p15.33	rs467095	1.19	6.9E-04	1.26	1.7E-05	1.25	2.2E-02	0.49	0.44	1.22	3.9E-09
9p21.3	rs8181047	1.18	2.3E-03	1.26	5.6E-05	1.34	9.6E-03	0.31	0.26	1.23	2.5E-08
9q34.12	rs928674	1.30	9.2E-04	1.33	4.2E-04	1.22	2.2E-01	0.13	0.11	1.30	6.8E-07
<b>Oropharyngeal cancer</b>											
6p21.32	rs3828805	1.47	5E-07	1.42	1.8E-07	1.02	0.89	0.75	0.72	1.39	7.1E-12

<sup>a</sup> Fixed-effects meta-analysis

<sup>b</sup> AF= average allele frequency of the effect allele.

## Supplementary Table 14

Validation of risk loci by TaqMan genotyping in ~ 700 individuals

rs	Region	Gene	# Genotypes <sup>a</sup>	% Concordance <sup>b</sup>	Type <sup>c</sup>	Meta-analysis results		
						OR	P	Analysis <sup>d</sup>
rs9270911	6p21.3	<i>HLA-DRB1</i>	758	100.0	Imputed	1.16	3.50E-07	overall
rs35975014	6p21.3	<i>HLA-DQA1</i> <i>OR52N4</i>	791	97.1	Genotyped	0.73	2.80E-08	overall
rs12363178 <sup>e</sup>	11p15	<i>TRIM5</i>	759	99.9	Genotyped	1.17	6.95E-06	overall
rs1919126	2p23.3	<i>C2orf16</i>	677	100.0	Genotyped	1.19	1.28E-07	OC
rs467095	5p15.3	<i>CLPTM1L</i>	747	100.0	Genotyped	1.20	1.77E-08	OC
rs3130614	6p21.3	<i>MICB</i>	684	99.7	Imputed	1.37	2.49E-07	OC
rs3731239	9p21.3	<i>CDKN2A</i>	736	100.0	Genotyped	1.21	1.22E-08	OC
rs73658919	9q34.1	<i>LAMC3</i>	725	99.2	Imputed	1.33	3.96E-08	OC
rs12910284	15q21.	<i>LOC105370811</i>	762	99.9	Imputed	1.19	4.28E-07	OC
rs2398180	15q26.	<i>NR2F2-AS1</i>	689	93.8	Imputed	1.20	2.77E-07	OC
rs1961637	2q36.1		665	100.0	Imputed	1.21	3.49E-07	OPC
rs12524487	6p21.3	<i>HLA-S</i>	791	97.1	Genotyped	0.66	1.68E-07	OPC

<sup>a</sup> Number of shared genotypes between array/imputed and TaqMan genotyped samples

<sup>b</sup> Concordance between array genotyped or imputed variants and TaqMan genotypes in a subset of individuals from ARCAGE, IARC Latin America, EPIC, and IARC oral CA study.

<sup>c</sup> Genotyped in the OncoArray or imputed variant

<sup>d</sup> Overall=oral and pharynx cancer, OC=oral cancer, OPC= oropharyngeal cancer

<sup>e</sup> Correlated with rs1453414 in European samples ( $r^2=0.85$ )

### Supplementary Table 15

Unconditional and conditional analyses (meta-analysis from 3 regions) for overall oral and pharynx cancer variants ( $P < 5 \times 10^{-8}$ ) at 4q23, conditioning was performed on the lead SNP rs1229984. Analyses adjusted by age, sex and eigenvectors.

rs	Chr:pos <sup>a</sup>	Effect allele	Other allele	Unconditional		Conditional on rs1229984	
				OR <sup>b</sup>	P	OR <sup>b</sup>	P
rs1229984	4:100239319	A	G	0.5587	2.29E-15		
rs116203444	4:100068114	C	T	0.5936	2.51E-07	0.91	0.44
rs138495951	4:100249726	A	G	0.5185	1.75E-09	0.91	0.50
rs141973904	4:100262242	T	C	0.5188	1.98E-09	0.90	0.47
rs145452708	4:100248642	C	G	0.5181	1.64E-09	0.91	0.50
rs146788033	4:99941138	G	A	0.5522	5.61E-08	0.90	0.42

<sup>a</sup> SNP position according to NCBI genome build 37 (Hg19)

<sup>b</sup> Fixed-effects meta-analysis OR and p-values



## Supplementary Table 16

Oral and oropharynx cancer results for previously identified loci in the upper aerodigestive tract GWAS (McKay et al 2011)

OC OncoArray results*						Previous UADT GWAS results**			
rs	EA/OA	case/control	OR	95%CI	P	case/control	OR	95%CI	P
rs1229984 <i>ADH1B</i>	A/G	2170/4950	0.65	0.53-0.81	0.0001	1980/9932	0.55	0.46-0.67	4.6E-10
rs1573496 <i>ADH7</i>	G/C	2179/4987	0.86	0.75-0.98	0.028	1975/9939	0.72	0.64-0.82	2.0E-07
OPC OncoArray Results*									
rs gene	EA/OA	case/control	OR	95%CI	P	case/control	OR	95%CI	P
rs1229984 <i>ADH1B</i>	A/G	2160/4950	0.56	0.44-0.71	2.3E-6	1897/9741	0.59	0.49-0.72	7.6E-08
rs1573496 <i>ADH7</i>	G/C	2166/4987	0.87	0.76-1.01	0.06	1922/9752	0.73	0.64-0.83	2.1E-06

\*Results for non-overlapping individuals between this study and UADT GWAS; OC=Oral cancer; OPC= Oral Cavity Cancer; EA=effect allele; OA=other allele.

\*\***Note:** rs1494961, rs1789924 and rs4767364 were significantly associated with all upper aerodigestive tract cancers at  $P < 5 \times 10^{-8}$  (Ref. 7) but not with OC or OPC alone.

### Supplementary Table 17

Unconditional and conditional analyses (meta-analysis from 3 regions) for overall oral and pharynx cancer variants ( $P < 5 \times 10^{-6}$ ) at 11p15, conditioning on the lead SNP rs1453414. Analyses adjusted by age, sex and eigenvectors.

rs	Chr:pos <sup>a</sup>	Effect allele	Other allele	Unconditional		Conditional on rs1453414	
				OR <sup>b</sup>	P	OR <sup>b</sup>	P
rs1453414	11:5829084	C	A	1.19	4.78E-08		
rs10838607	11:5778538	C	T	1.18	4.73E-06	1.02	0.80
rs11038937	11:5777248	C	T	1.17	4.96E-06	1.02	0.81
rs11038939	11:5777395	A	G	1.18	4.88E-06	1.02	0.81
rs11038940	11:5777739	C	T	1.18	4.82E-06	1.02	0.80
rs11038942	11:5778010	G	A	1.18	4.77E-06	1.02	0.80
rs11038944	11:5778930	T	C	1.18	4.71E-06	1.02	0.80
rs11038945	11:5779431	G	A	1.18	4.66E-06	1.02	0.79
rs11038946	11:5779859	T	C	1.18	4.27E-06	1.02	0.77
rs11038949	11:5780198	T	A	1.18	4.16E-06	1.02	0.76
rs11038950	11:5780334	T	C	1.18	3.74E-06	1.02	0.73
rs11038951	11:5780343	C	G	1.18	3.63E-06	1.02	0.72
rs11038952	11:5780354	C	T	1.18	3.8E-06	1.02	0.73
rs11038960	11:5780740	C	T	1.18	4.04E-06	1.02	0.76
rs11826964	11:5775269	C	T	1.18	4.23E-06	1.02	0.75
rs11828490	11:5775148	A	G	1.18	4.07E-06	1.02	0.74
rs12362742	11:5774732	G	C	1.17	6.08E-06	1.03	0.67
rs12362750	11:5774765	T	C	1.18	2.92E-06	1.03	0.64
rs12363545	11:5777096	C	T	1.17	4.94E-06	1.02	0.81
rs12364223	11:5774253	T	A	1.18	2.73E-06	1.03	0.62
rs12364265	11:5774653	C	A	1.18	3.12E-06	1.03	0.66
rs12364563	11:5774429	C	T	1.18	2.92E-06	1.03	0.64
rs12364723	11:5775788	G	A	1.18	4.33E-06	1.02	0.76
rs12365564	11:5774436	C	G	1.18	2.95E-06	1.03	0.64
rs12365568	11:5774455	A	G	1.18	2.97E-06	1.03	0.64
rs12793267	11:5774025	T	C	1.18	2.79E-06	1.03	0.62
rs1498557	11:5773051	T	A	1.18	2.2E-06	1.05	0.39
rs1498558	11:5772818	A	T	1.18	1.95E-06	1.05	0.38
rs4414254	11:5774814	T	A	1.18	3.29E-06	1.03	0.68
rs4597095	11:5775049	T	C	1.18	3.54E-06	1.03	0.70
rs5012841	11:5775730	C	T	1.18	4.18E-06	1.02	0.75
rs5012842	11:5775362	G	A	1.18	4.17E-06	1.02	0.75
rs5012843	11:5775360	C	G	1.18	4.12E-06	1.02	0.75
rs61875860	11:5773663	C	T	1.18	2.72E-06	1.03	0.60
rs61875861	11:5773735	C	T	1.18	2.49E-06	1.04	0.58
rs61875862	11:5773781	A	T	1.18	2.66E-06	1.04	0.60
rs7126020	11:5784187	T	C	0.86	2.45E-06	1.05	0.63
rs7395852	11:5777210	A	G	0.86	3.08E-06	1.05	0.56

<sup>a</sup> SNP position according to NCBI genome build 37 (Hg19)

<sup>b</sup> Fixed-effects meta-analysis OR and p-values

**Supplementary Table 18**

Unconditional and conditional meta-analysis from 3 regions for oral cancer variants ( $P < 5 \times 10^{-8}$ ) at 2p23, conditioning was performed on the lead SNP rs6547741. Analyses adjusted by age, sex and eigenvectors.

rs	Chr:pos <sup>a</sup>	Effect allele	Other allele	Unconditional		Conditional on rs6547741	
				OR <sup>b</sup>	P	OR <sup>b</sup>	P
rs6547741	2:27855924	A	G	0.8345	3.97E-08		
rs1528402	2:27857517	A	G	0.8347	4.15E-08	3955.671	0.1828
rs4666007	2:27856495	C	T	0.8347	4.16E-08	145398.4	0.1004
rs6547740	2:27854369	A	G	0.8349	4.33E-08	0.0239	0.47

<sup>a</sup> SNP position according to NCBI genome build 37 (Hg19)

<sup>b</sup> Fixed-effects meta-analysis OR and p-values

## Supplementary Table 19

Unconditional and conditional meta-analysis from 3 regions for oral cancer variants at 5p15.3, conditioning was performed on the lead SNP rs10462706 and rs467095. Analyses adjusted by age, sex and eigenvectors.

rs	Chr:pos <sup>a</sup>	Effect allele	Other allele	Unconditional		Conditioned		conditioned on
				OR <sup>b</sup>	P	OR <sup>b</sup>	P	
rs10462706	5:1343794	T	C	0.74	5.54E-10	0.74	5.54E-10	
rs467095	5:1336221	C	T	1.20	2.28E-08	1.13	5.14E-04	rs10462706
rs112155367	5:1337070	G	A	1.20	1.9E-08	1.16	1.16E-05	rs10462706
rs201708486	5:1334508	I	D	1.20	2.91E-08	1.16	1.58E-05	rs10462706
rs2447853	5:1333077	G	A	1.19	4.52E-08	1.15	1.92E-05	rs10462706
rs31484	5:1337906	T	A	1.20	3.68E-08	1.16	1.97E-05	rs10462706
rs31487	5:1341101	C	G	1.20	3.25E-08	1.16	1.46E-05	rs10462706
rs31490	5:1344458	A	G	1.20	3.6E-08	1.16	1.72E-05	rs10462706
rs421284	5:1325590	G	A	1.20	2.14E-08	1.16	9.88E-06	rs10462706
rs452932	5:1330253	G	A	1.19	4.31E-08	1.15	2.23E-05	rs10462706
rs455433	5:1336243	G	A	1.20	3.77E-08	1.16	2.01E-05	rs10462706
rs459961	5:1337106	A	T	1.19	4.52E-08	1.15	1.84E-05	rs10462706
rs460073	5:1336459	G	A	1.20	3.05E-08	1.16	1.69E-05	rs10462706
rs10462706	5:1336221	C	T	0.74	5.54E-10	0.78	4.24E-06	rs467095

<sup>a</sup> SNP position according to NCBI genome build 37 (Hg19)

<sup>b</sup> Fixed-effects meta-analysis OR and p-values

### Supplementary Table 20

Unconditional and conditional meta-analysis from 3 regions for oral cancer variants ( $P < 5 \times 10^{-8}$ ) at 9p21, conditioning was performed on the lead SNP lead rs8181047. Analyses adjusted by age, sex and eigenvectors.

rs	Chr:pos <sup>a</sup>	Effect allele	Other allele	Unconditional		Conditional on rs8181047	
				OR <sup>b</sup>	P	OR <sup>b</sup>	P
rs8181047	9:22064465	A	G	1.24	3.80E-09		
rs10757268	9:22059905	C	T	0.81	7.78E-09	1.03	0.87
rs1412832	9:22077543	G	A	1.23	1.34E-08	0.94	0.74
rs1537378	9:22061614	A	G	1.20	1.68E-07	1.04	0.51
rs2095144	9:22060136	G	A	0.81	6.93E-09	1.02	0.94
rs2383205	9:22060935	A	G	1.20	2.55E-07	1.05	0.43
rs35307545	9:22071750	G	A	1.19	1.65E-07	1.05	0.33
rs3731239	9:21974218	G	A	1.21	1.22E-08	1.11	0.03
rs8181050	9:22064391	G	A	1.20	1.74E-07	1.04	0.54
rs944800	9:22050898	A	G	1.21	7.42E-08	0.96	0.71

<sup>a</sup> SNP position according to NCBI genome build 37 (Hg19)

<sup>b</sup> Fixed-effects meta-analysis OR and p-values

### Supplementary Table 21

Unconditional and conditional meta-analysis from 3 regions for oral cancer variants ( $P < 5 \times 10^{-8}$ ) at 9q34, conditioning was performed on the lead SNP lead rs928674. Analyses adjusted by age, sex and eigenvectors.

rs	Chr:pos <sup>a</sup>	Effect allele	Other allele	Unconditional		Conditional on rs928674	
				OR <sup>b</sup>	P	OR <sup>b</sup>	P
rs928674	9:133952024	G	A	1.33	2.09E-08		
rs10901348	9:133956465	A	G	1.31	7.12E-08	1.09	0.47
rs11791030	9:133971049	G	T	1.31	1.68E-07	1.09	0.41
rs199717881	9:133953882	C	A	1.32	2.62E-08	1.14	0.32
rs3765566	9:133942766	G	A	1.31	2.28E-07	1.01	0.92
rs72768533	9:133959518	G	C	1.31	6.82E-08	1.10	0.40
rs72768534	9:133959740	C	G	1.31	6.77E-08	1.10	0.39
rs73658919	9:133953983	G	A	1.33	3.96E-08	1.08	0.68
rs77452476	9:133962573	G	T	1.29	5.97E-08	1.13	0.11
rs7858204	9:133947345	A	G	1.32	1.2E-07	1.04	0.79
rs7875478	9:133947180	T	A	1.32	1.23E-07	1.04	0.79

<sup>a</sup> SNP position according to NCBI genome build 37 (Hg19)

<sup>b</sup> Fixed-effects meta-analysis OR and p-values

### Supplementary Table 25

Overall oral and pharyngeal analyses (6,034 cases and 6,585 controls) top loci stratified analyses (region, tumor site, gender, tobacco smoking and alcohol drinking status). Data for forest plots in Figure 2.

<b>a. rs1229984 4q23</b>		<b>case</b>	<b>control</b>	<b>OR</b>	<b>P</b>	<b>95% CI</b>	
<b>Region</b>	Europe	2490	2905	0.50	9.74E-10	0.40	0.62
	North America	2541	2503	0.74	9.59E-03	0.59	0.93
	South America	977	1116	0.38	3.21E-08	0.27	0.53
<b>Site</b>	Oral	2975	6524	0.57	1.09E-09	0.48	0.68
	Oropharynx	2632	6524	0.55	8.53E-09	0.45	0.67
<b>Gender</b>	Men	4506	4283	0.51	2.64E-14	0.43	0.60
	Women	1501	2241	0.70	8.57E-03	0.53	0.91
<b>Smoking Status</b>	Never	1052	2483	0.93	5.87E-01	0.72	1.20
	Former	1780	2244	0.71	8.70E-03	0.55	0.92
	Current	2615	1451	0.33	5.76E-15	0.25	0.43
<b>Drinking Status</b>	Never	816	1183	0.63	1.67E-04	0.50	0.80
	Ever	4820	4797	0.49	1.36E-15	0.42	0.58

<b>b. rs201982221 10q26.13</b>		<b>case</b>	<b>control</b>	<b>OR</b>	<b>P</b>	<b>95% CI</b>	
<b>Region</b>	Europe	2462	2910	1.87	1.16E-06	1.46	2.42
	North America	2531	2504	1.53	8.81E-04	1.19	1.96
	South America	982	1129	1.55	7.85E-02	0.95	2.51
<b>Site</b>	Oral	2961	6543	1.71	1.05E-07	1.41	2.08
	Oropharynx	2613	6543	1.70	7.98E-07	1.38	2.09
	Men	4484	4299	1.66	1.31E-06	1.35	2.03
<b>Smoking Status</b>	Women	1490	2244	1.81	1.05E-04	1.34	2.43
	Never	1049	2494	1.17	2.83E-03	1.19	2.32
	Former	1768	2251	1.35	6.02E-03	1.13	2.09
<b>Drinking Status</b>	Current	2597	1452	1.32	1.72E-02	1.07	2.03
	Never	814	1189	1.63	3.07E-04	1.25	2.13
	Ever	4789	4809	1.60	1.43E-06	1.33	1.93

<b>c. rs1453414 11p15</b>		<b>case</b>	<b>control</b>	<b>OR</b>	<b>P</b>	<b>95% CI</b>	
<b>Region</b>	Europe	2493	2915	1.22	1.56E-04	1.10	1.35
	North America	2541	2509	1.21	1.16E-04	1.10	1.34
	South America	983	1129	1.11	1.80E-01	0.95	1.29
<b>Site</b>	Oral	2981	6553	1.19	1.65E-05	1.10	1.28
	Oropharynx	2635	6553	1.22	4.26E-06	1.12	1.32
<b>Gender</b>	Men	4514	4307	1.18	2.29E-05	1.09	1.27
	Women	1502	2246	1.24	3.47E-04	1.10	1.40
<b>Smoking Status</b>	Never	1053	2498	1.17	2.24E-02	1.02	1.33
	Former	1785	2248	1.33	1.75E-06	1.18	1.49
	Current	2617	1461	1.18	6.80E-03	1.05	1.34
<b>Drinking Status</b>	Never	819	1193	1.11	5.03E-02	1.00	1.23
	Ever	4825	4817	1.19	2.92E-06	1.11	1.28

<b>d. rs3828805 6p21.32</b>		<b>case</b>	<b>control</b>	<b>OR</b>	<b>P</b>	<b>95% CI</b>	
<b>Region</b>	Europe	2499	2928	1.39	6.09E-09	1.24	1.56
	North America	2549	2522	1.30	8.63E-08	1.18	1.43
	South America	986	1135	1.04	6.18E-01	0.89	1.21
<b>Site</b>	Oral	2990	6585	1.23	2.83E-07	1.14	1.34
	Oropharynx	2641	6585	1.37	2.21E-12	1.25	1.50
<b>Gender</b>	Men	4527	4325	1.31	8.38E-12	1.22	1.42
	Women	1507	2260	1.22	1.43E-03	1.08	1.38
<b>Smoking Status</b>	Never	1057	2508	1.32	1.30E-02	1.06	1.65
	Former	1792	2263	1.36	5.28E-07	1.20	1.53
<b>Drinking Status</b>	Current	2623	1466	1.27	1.07E-04	1.13	1.44
	Never	820	1199	1.17	8.55E-02	0.98	1.39
	Ever	4840	4840	1.30	6.97E-12	1.20	1.40



### Supplementary Table 26

Oral analyses (2,990 cases and 6,585 controls) top loci stratified analyses (region, gender, tobacco smoking and alcohol drinking status). Data for forest plots in Figure 3.

<b>a. rs6547741 2p23</b>		<b>case</b>	<b>control</b>	<b>OR</b>	<b>P</b>	<b>95% CI</b>	
<b>Region</b>	Europe	1225	2928	0.84	6.46E-04	0.76	0.93
	North America	1238	2522	0.80	1.03E-05	0.72	0.88
	South America	527	1135	0.92	2.64E-01	0.79	1.07
<b>Gender</b>	Men	2045	4325	0.84	9.42E-06	0.77	0.90
	Women	945	2260	0.84	2.81E-03	0.75	0.94
<b>Smoking Status</b>	Never	543	2508	0.87	5.89E-02	0.76	1.01
	Former	815	2263	0.79	1.19E-04	0.69	0.89
<b>Drinking Status</b>	Current	1410	1466	0.81	2.84E-04	0.72	0.91
	Never	469	1199	0.86	8.18E-02	0.72	1.02
	Ever	2369	4840	0.82	2.62E-07	0.77	0.89

<b>b. s10462706 5p15.3</b>		<b>case</b>	<b>control</b>	<b>OR</b>	<b>P</b>	<b>95% CI</b>	
<b>Region</b>	Europe	1225	2928	0.75	1.87E-04	0.65	0.87
	North America	1238	2522	0.74	7.04E-05	0.63	0.86
	South America	527	1135	0.69	2.62E-03	0.55	0.88
<b>Gender</b>	Men	2045	4325	0.74	7.64E-07	0.66	0.84
	Women	945	2260	0.71	1.44E-04	0.60	0.85
<b>Smoking Status</b>	Never	543	2508	0.60	7.27E-06	0.48	0.75
	Former	815	2263	0.71	3.98E-04	0.59	0.86
<b>Drinking Status</b>	Current	1410	1466	0.83	3.48E-02	0.70	0.99
	Never	469	1199	0.53	2.02E-06	0.41	0.69
	Ever	2369	4840	0.76	1.21E-06	0.68	0.85

<b>c. rs8181047 9p21.3</b>		<b>case</b>	<b>control</b>	<b>OR</b>	<b>P</b>	<b>95% CI</b>	
<b>Region</b>	Europe	1225	2926	1.18	2.28E-03	1.06	1.32
	North America	1237	2522	1.25	7.27E-05	1.12	1.39
	South America	527	1135	1.39	7.01E-04	1.15	1.68
<b>Gender</b>	Men	2044	4324	1.25	5.96E-07	1.15	1.37
	Women	945	2259	1.21	2.84E-03	1.07	1.37
<b>Smoking Status</b>	Never	543	2507	1.25	3.40E-03	1.08	1.46
	Former	814	2262	1.17	2.10E-02	1.02	1.34
<b>Drinking Status</b>	Current	1410	1466	1.36	4.50E-06	1.19	1.55
	Never	469	1199	1.29	9.34E-03	1.06	1.55
	Ever	2368	4839	1.24	3.54E-07	1.14	1.34

<b>d. rs928674 9q34</b>		<b>case</b>	<b>control</b>	<b>OR</b>	<b>P</b>	<b>95% CI</b>	
<b>Region</b>	Europe	1225	2928	1.31	6.48E-04	1.12	1.53
	North America	1238	2522	1.36	5.63E-05	1.17	1.58
	South America	527	1135	1.27	5.07E-02	1.00	1.62
<b>Gender</b>	Men	2045	4325	1.33	4.92E-06	1.18	1.50
	Women	945	2260	1.30	3.09E-03	1.09	1.55
<b>Smoking Status</b>	Never	543	2508	1.32	1.30E-02	1.06	1.65
	Former	815	2263	1.45	6.81E-05	1.21	1.73
<b>Drinking Status</b>	Current	1410	1466	1.38	4.00E-04	1.16	1.66
	Never	469	1199	1.43	7.18E-03	1.10	1.85
	Ever	2369	4840	1.33	9.94E-07	1.18	1.49

