Supplementary Material for

“Genome-wide association studies of gastric adenocarcinoma and esophageal squamous cell carcinoma identify a shared susceptibility locus in *PLCE1* at 10q23”

Christian C Abnet, Neal D Freedman, Nan Hu, Zhaoming Wang, Kai Yu, Xiao-Ou Shu, Jian-Min Yuan, Wei Zheng, Sanford M Dawsey, Linda M Dong, Maxwell P Lee, Ti Ding, You-Lin Qiao, Yu-Tang Gao, Woon-Puay Koh, Yong-Bing Xiang, Ze-Zhong Tang, Jin-Hu Fan, Chaoyu Wang, William Wheeler, Mitchell H Gail, Meredith Yeager, Jeff Yuenger, Amy Hutchinson, Kevin B Jacobs, Carol A Giffen, Laurie Burdett, Joseph F Fraumeni Jr, Margaret A Tucker, Wong-Ho Chow, Alisa M Goldstein, Stephen J Chanock & Philip R Taylor

Supplementary Table 1. Description of studies and number of subjects included in the genome-wide scan phase, second phase, and combined phase by study and phenotype

			Shanxi UGI Cancer Genetics Project	**Linxian Nutrition Intervention Trial**	**Shanghai Men's Health Study**	**Shanghai Women's Health Study**	**Singapore Chinese Health Study**
Abbreviation			Shanxi	NIT	SMHS	SWHS	SCHS
Design			Case-Control; Case only	Cohort	Cohort	Cohort	Cohort
Location			Shanxi Province, PRC	Henan Province, PRC	Shanghai, PRC	Shanghai, PRC	Singapore
Nested sampling method			Not applicable	Case-Cohort	Incidence density	Incidence density	Incidence density
Matching			Sex Age (± 5 years) *For case-control portion*	Sex Age (10 year groups)	Age (± 2 years) Biosample collection: Time of day (AM:PM) Month (± 1)	Age (± 2 years) Menstrual period (± 1 day) Biosample collection: Time of day (AM:PM) Month (± 1) Last meal (≤2 hours)	Sex Age (± 2 years) Dialect Month of DNA collection (± 6)
Phase	**Phenotype**	**Total**					
Genome-wide	Controls	2100	1650	450			
	Gastric	1625	1368	257			
	Cardia	1110	853	257			
	Noncardia	515	515	0			
	Unclassified	0	0	0			
	ESCC	1898	1399	499			
Second	Controls	1202	104	330	477	54	237
	Gastric	615	50	145	205	0	215
	Cardia	103	28	27	19	0	29
	Noncardia	402	22	118	147		115
	Unclassified	110	0	0	39	0	71
	ESCC	217	82	52	34	27	22
Combined	Controls	3302	1754	780	477	54	237
	Gastric	2240	1418	402	205	0	215
	Cardia	1213	881	284	19	0	29
	Noncardia	917	537	118	147	0	115
	Unclassified	110	0	0	39	0	71
	ESCC	2115	1481	551	34	27	22
	total	7657	4653	1733	716	81	474

Supplementary Table 2. Associations between rs3781264 and gastric cardia cancer or ESCC separately in the Shanxi, NIT, and all studies combined using per allele models

Study	**Cancer**	**SNP**	**MAF Controls**	**MAF Cases**	**# controls**	**# cases**	***P*-value**	**OR**	**(95% CI)**
Shanxi	Cardia	rs3781264	15%	23%	1648	853	1.02×10^{-10}	1.65	(1.41-1.92)
NIT	Cardia	rs3781264	14%	22%	450	257	1.87×10^{-4}	1.72	(1.29-2.30)
Combined	Cardia	rs3781264	15%	22%	3302	1213	1.06×10^{-13}	1.60	(1.41-1.81)
Shanxi	ESCC	rs3781264	15%	20%	1648	1466	4.26×10^{-7}	1.40	(1.23-1.59)
NIT	ESCC	rs3781264	13%	17%	450	547	0.038	1.28	(1.01-1.61)
Combined	ESCC	rs3781264	15%	19%	3302	2115	7.30×10^{-9}	1.38	(1.23-1.53)

Supplementary Table 3. Association between new and *a priori* SNPs reported by other GWAS for gastric cancer in all cases and by anatomic location within the stomach and for esophageal squamous cell carcinoma in the genome-wide scan phase

		Per Allele	**Genotype models**		
			Heterozygous	**Homozygous**	
Cancer group and chromosome location	**NCI dbSNP identifier (major, minor allele)**	**OR (95%CI)**	**OR (95%CI)**	**OR (95%CI)**	P_{2df}
Total Gastric					
10q23	rs2274223 (A, G)	1.40 (1.25-1.56)	1.36 (1.19-1.57)	2.05 (1.53-2.76)	1.55×10^{-8}
	rs3765524 (C, T)	1.39 (1.25-1.56)	1.36 (1.19-1.57)	2.05 (1.52-2.76)	1.84×10^{-8}
	rs3781264 (T, C)	1.51 (1.33-1.70)	1.54 (1.33-1.78)	2.08(1.41-3.07)	2.80×10^{-10}
	rs11187842 (C, T)	1.48 (1.31-1.68)	1.53 (1.32-1.77)	1.96 (1.31-2.92)	2.23×10^{-9}
	rs753724 (G, T)	1.49 (1.31-1.68)	1.52 (1.32-1.76)	2.02 (1.36-3.00)	1.80×10^{-9}
1q22	rs4072037 (A, G)	0.71 (0.62-0.82)	0.68 (0.58-0.79)	0.66 (0.40-1.10)	3.04×10^{-6}
	rs4460629 (C, T)	0.68 (0.59-0.79)	0.68 (0.58-0.80)	0.49 (0.26-0.92)	2.18×10^{-6}
22q12	rs738722 (C, T)	1.01 (0.91-1.13)	0.96 (0.83-1.10)	1.16 (0.88-1.53)	0.37
	rs1547014 (C, T) [a]	1.01 (0.89-1.13)	0.90 (0.78-1.04)	1.46 (1.03-2.07)	0.022
8q24	rs2920297 (T, C) [b]	1.07 (0.97-1.19)	1.16 (1.01-1.33)	1.03 (0.81-1.31)	0.10
	rs2294008 (C, T)	1.07 (0.97-1.19)	1.16 (1.01-1.33)	1.03 (0.80-1.31)	0.091
Cardia					
10q23	rs2274223 (A, G)	1.59 (1.41-1.80)	1.64 (1.40-1.91)	2.39 (1.74-3.30)	5.23×10^{-13}
	rs3765524 (C, T)	1.59 (1.40-1.79)	1.63 (1.40-1.90)	2.38 (1.72-3.28)	8.73×10^{-13}
	rs3781264 (T, C)	1.66 (1.45-1.90)	1.73 (1.48-2.04)	2.38 (1.57-3.61)	5.32×10^{-13}
	rs11187842 (C, T)	1.63 (1.42-1.87)	1.72 (1.46-2.02)	2.19 (1.42-3.36)	7.09×10^{-12}
	rs753724 (G, T)	1.63 (1.42-1.87)	1.72 (1.46-2.02)	2.19 (1.42-3.36)	7.96×10^{-12}
1q22	rs4072037 (A, G)	0.75 (0.65-0.88)	0.72 (0.61-0.86)	0.72 (0.41-1.26)	9.86×10^{-4}
	rs4460629 (C, T)	0.72 (0.61-0.85)	0.73 (0.61-0.87)	0.49 (0.23-1.04)	5.77×10^{-4}
22q12	rs738722 (C, T)	1.06 (0.94-1.19)	0.96 (0.82-1.12)	1.33 (0.99-1.79)	0.10
	rs1547014 (C, T) [a]	1.09 (0.96-1.24)	0.96 (0.82-1.13)	1.71 (1.18-2.49)	0.011
8q24	rs2920297 (T, C) [b]	1.00 (0.89-1.12)	1.06 (0.91-1.23)	0.92 (0.69-1.21)	0.56
	rs2294008 (C, T)	1.00 (0.89-1.12)	1.06 (0.91-1.24)	0.92 (0.69-1.21)	0.53
Noncardia					
10q23	rs2274223 (A, G)	1.02 (0.86-1.22)	0.90 (0.73-1.12)	1.41 (0.91-2.21)	0.15
	rs3765524 (C, T)	1.03 (0.87-1.22)	0.91 (0.73-1.13)	1.42 (0.91-2.22)	0.16
	rs3781264 (T, C)	1.18 (0.97-1.42)	1.18 (0.94-1.47)	1.38 (0.74-2.54)	0.25
	rs11187842 (C, T)	1.17 (0.97-1.42)	1.17 (0.93-1.47)	1.41 (0.76-2.61)	0.26
	rs753724 (G, T)	1.19 (0.99-1.44)	1.16 (0.92-1.45)	1.60 (0.88-2.89)	0.16
1q22	rs4072037 (A, G)	0.60 (0.47-0.75)	0.55 (0.43-0.71)	0.59 (0.26-1.37)	1.63×10^{-5}
	rs4460629 (C, T)	0.59 (0.47-0.75)	0.57 (0.44-0.74)	0.52 (0.20-1.36)	5.56×10^{-5}
22q12	rs738722 (C, T)	0.93 (0.78-1.10)	0.96 (0.78-1.18)	0.79 (0.49-1.27)	0.61
	rs1547014 (C, T) [a]	0.83 (0.69-1.01)	0.77 (0.62-0.96)	0.99 (0.54-1.81)	0.069
8q24	rs2920297 (T, C) [b]	1.27 (1.09-1.47)	1.45 (1.18-1.79)	1.35 (0.94-1.94)	1.87×10^{-3}

	rs2294008 (C, T)	1.27 (1.09-1.47)	1.46 (1.18-1.80)	1.35 (0.94-1.94)	1.63 x 10^{-3}
ESCC					
10q23	rs2274223 (A, G)	1.33 (1.20-1.48)	1.32 (1.15-1.50)	1.84 (1.37-2.46)	7.75 x 10^{-7}
	rs3765524 (C, T)	1.34 (1.20-1.49)	1.32 (1.16-1.51)	1.83 (1.37-2.47)	6.32 x 10^{-7}
	rs3781264 (T, C)	1.38 (1.22-1.55)	1.35 (1.17-1.55)	2.03 (1.39-2.97)	7.14 x 10^{-7}
	rs11187842 (C, T)	1.37 (1.21-1.54)	1.35 (1.17-1.55)	1.99 (1.35-2.94)	1.66 x 10^{-6}
	rs753724 (G, T)	1.37 (1.22-1.55)	1.35 (1.17-1.55)	2.02 (1.37-2.98)	1.35 x 10^{-6}
22q12	rs738722 (C, T)	1.32 (1.19-1.45)	1.22 (1.07-1.39)	1.96 (1.53-2.50)	9.94 x 10^{-8}
	rs1547014 (C, T) [a]	1.33 (1.19-1.48)	1.23 (1.08-1.40)	2.22 (1.62-3.06)	2.79 x 10^{-7}
1q22	rs4072037(A, G)	0.85 (0.75-0.97)	0.80 (0.69-0.93)	0.99 (0.64-1.55)	0.012
	rs4460629(C, T)	0.86 (0.75-0.99)	0.82 (0.71-0.96)	1.01 (0.61-1.69)	0.038
4q23	rs1042026 (G, A) [c]	1.03 (0.94-1.14)	0.98 (0.86-1.12)	1.15 (0.90-1.46)	0.45
	rs2051428 (C, T)	0.98 (0.89-1.07)	0.96 (0.84-1.10)	0.97 (0.79-1.19)	0.83
	rs3819197 (T, C)	1.02 (0.92-1.12)	0.96 (0.84-1.10)	1.12 (0.89-1.41)	0.43
	rs1159918 (G, T)	1.05 (0.95-1.17)	0.99 (0.87-1.13)	1.25 (0.97-1.62)	0.21
12q24	rs671 (G, A)	0.91 (0.80-1.03)	0.96 (0.83-1.11)	0.65 (0.43-0.98)	0.11
	rs3782886 (A, G)	0.92 (0.82-1.04)	0.96 (0.83-1.11)	0.71 (0.48-1.06)	0.22

Results were derived from logistic regression models adjusted for age (10 year categories), sex, and study. There were 1,625 gastric cancer cases (1,110 cardia and 515 noncardia) 1,898 ESCC cases, and 2,100 control subjects. [a] Since the TaqMan assay for this SNP failed validation, we present results for this SNP here, but not in Table 2. [b] This SNP is presented as a proxy for rs2976392, the two SNPs have pair-wise r^2 of 1.0 for the CHB ethnicity in HapMap 2. [C] This SNP is presented as a proxy for rs1229984 the two SNPs have pair-wise r^2 of 0.68 for the CHB ethnicity in HapMap 2.

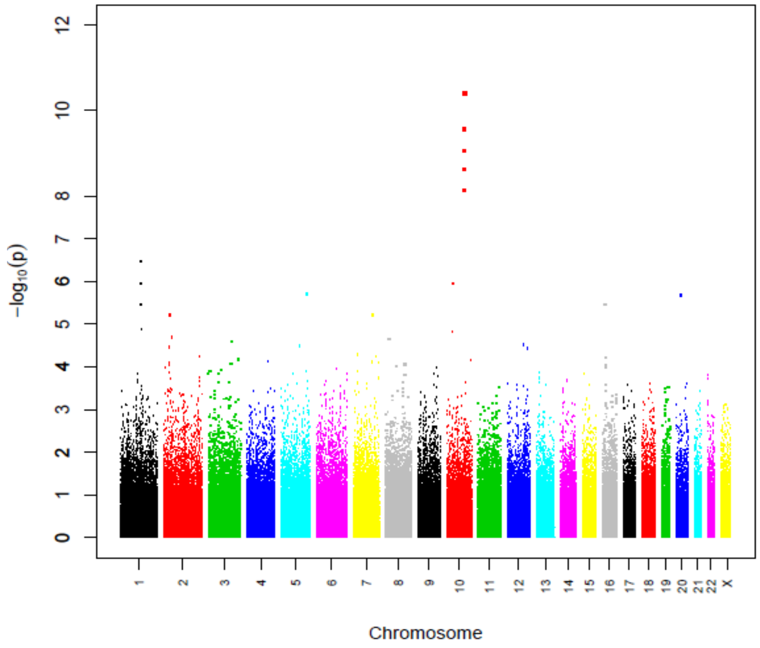
Supplementary Table 4. Association between SNPs at loci with genes involved in alcohol metabolism and risk of esophageal squamous cell carcinoma overall and stratified by alcohol or tobacco use based on subjects in the genome-wide phase

			Heterozygote	**Homozygote**	
Locus	**SNP**	**Strata**	**OR (CI)**	**OR (CI)**	$P_{\text{2 df score}}$
ADH1B	rs1159918 (G, T)	*Overall* [a]	0.99 (0.87-1.13)	1.24 (0.96, 1.61)	0.24
		Alcohol [b]			
		Never	0.94 (0.79, 1.12)	1.13 (0.80, 1.61)	0.54
		Ever	1.05 (0.86, 1.28)	1.37 (0.93, 2.03)	0.28
		Tobacco [c]			
		Never	0.88 (0.72, 1.07)	0.89 (0.60, 1.33)	0.41
		Ever	1.09 (0.91, 1.30)	1.54 (1.09, 2.18)	0.046
	rs1042026 (G, A)[d]	*Overall* [a]	0.98 (0.86, 1.12)	1.14 (0.89, 1.44)	0.49
		Alcohol [b]			
		Never	0.94 (0.79, 1.12)	1.01 (0.73, 1.40)	0.78
		Ever	1.02 (0.84, 1.25)	1.28 (0.90, 1.84)	0.39
		Tobacco [c]			
		Never	0.90 (0.74, 1.10)	0.84 (0.58, 1.20)	0.44
		Ever	1.05 (0.88, 1.26)	1.41 (1.02, 1.94)	0.11
ALDH2	rs3782886 (A, G)	*Overall* [a]	0.99 (0.85, 1.14)	0.75 (0.50, 1.12)	0.37
		Alcohol [b]			
		Never	0.87 (0.73, 1.05)	0.74 (0.48, 1.13)	0.18
		Ever	1.26 (0.99, 1.61)	0.69 (0.20, 2.41)	0.15
		Tobacco [c]			
		Never	0.82 (0.66, 1.02)	0.86 (0.48, 1.54)	0.19
		Ever	1.21 (0.98, 1.48)	0.71 (0.41, 1.25)	0.063
	rs671 (G, A)	*Overall* [a]	0.99 (0.85, 1.14)	0.69(0.45, 1.04)	0.21
		Alcohol [b]			
		Never	0.86 (0.72, 1.04)	0.67 (0.43, 1.04)	0.095
		Ever	1.28 (1.00, 1.64)	0.69 (0.20, 2.41)	0.12
		Tobacco [c]			
		Never	0.83 (0.66, 1.03)	0.83 (0.45, 1.53)	0.21
		Ever	1.20 (0.97, 1.47)	0.63 (0.35, 1.13)	0.042

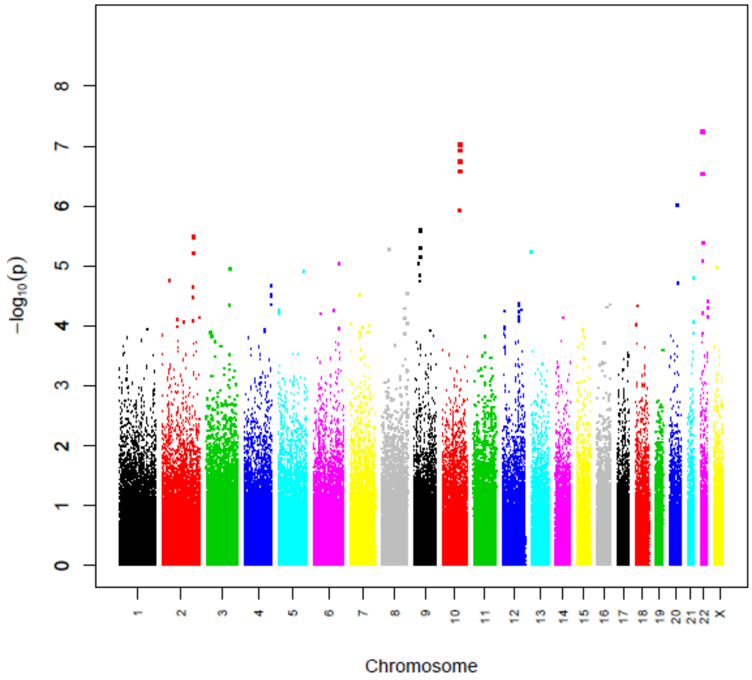
[a] Results were derived from the overall logistic regression model adjusted for age, sex, study, alcohol, and tobacco use. [b] The proportion of cases and controls that ever consumed alcohol was 59% and 58%, respectively. [c] The proportion of cases and controls that ever smoked tobacco was 48% and 45%, respectively. [d] This SNP is presented as a proxy for rs1229984 the two SNPs have pair-wise r^2 of 0.68 for the CHB ethnicity in HapMap 2.

Supplementary Figure 1. Manhattan plots based on *P*-values derived from 1 df trend tests from logistic regression models adjusted for age, sex, and study using subjects in the genome-wide phase. These associations were based on 1,625 gastric cancer cases or 1,898 esophageal squamous cell carcinoma cases and 2,100 common control subjects scanned using the Illumina 660W Quad chip. **a.** gastric cancer; **b.** esophageal squamous cell carcinoma

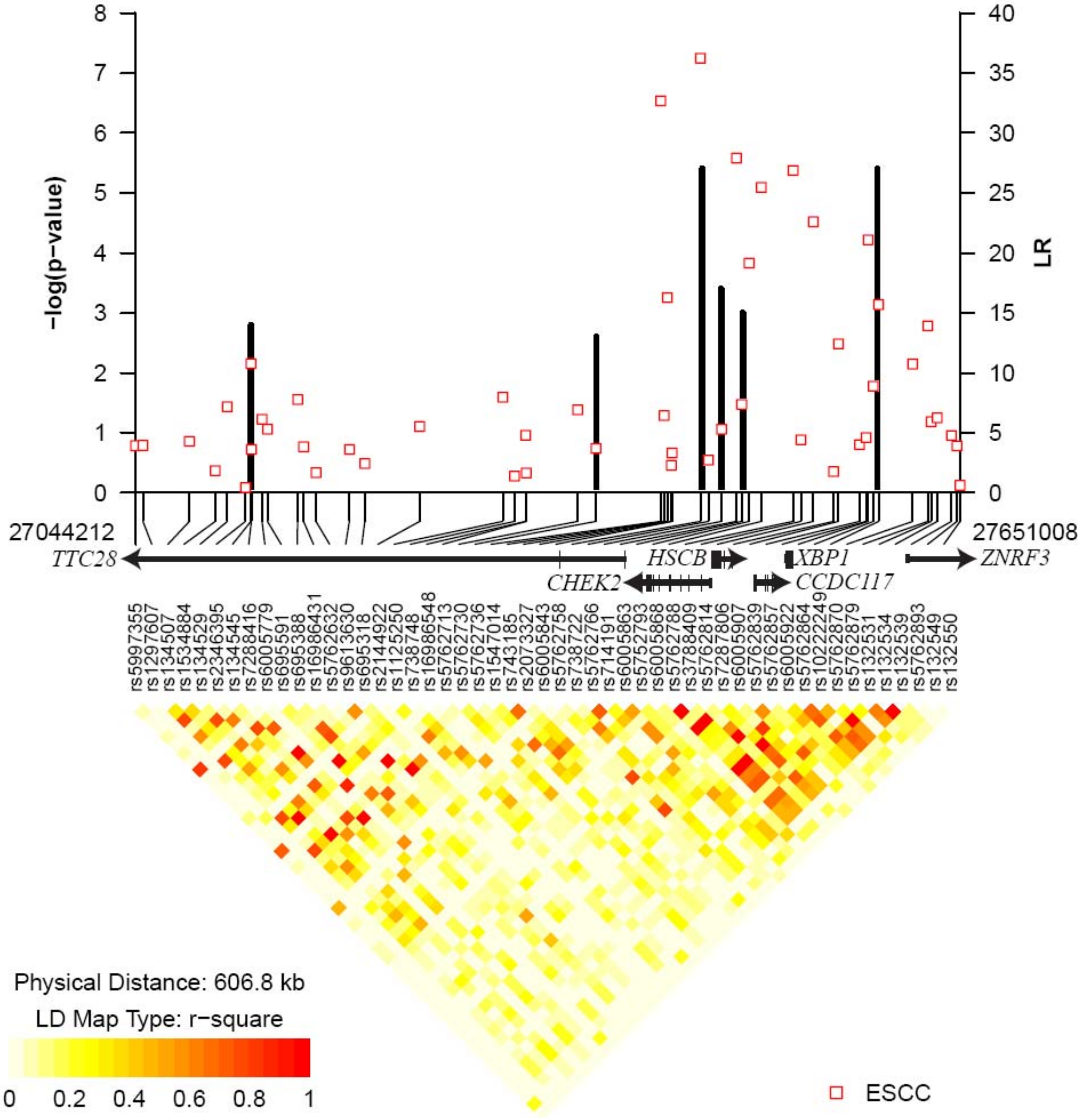
a. gastric cancer



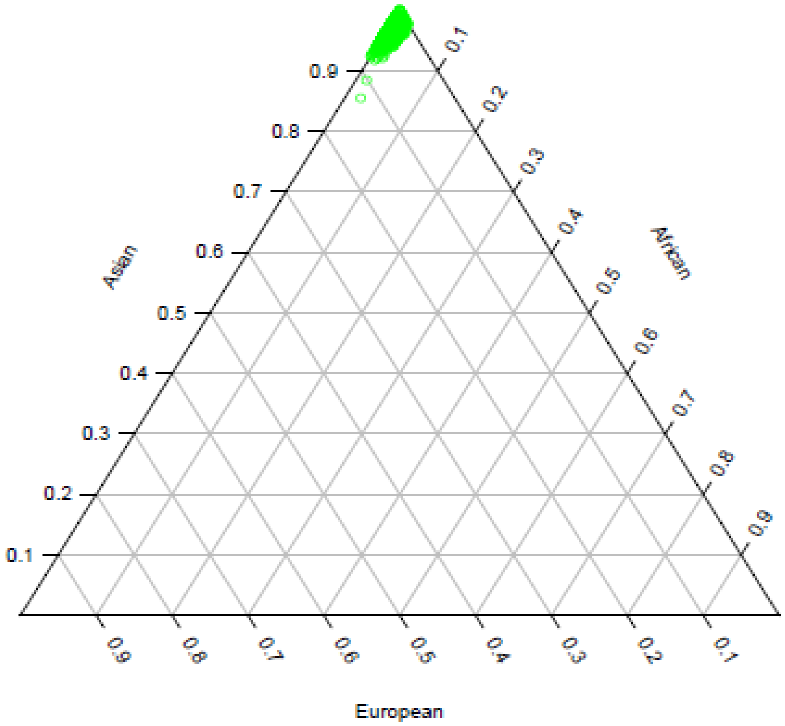
b. esophageal squamous cell carcinoma



Supplementary Figure 2. Association results, recombination, and linkage disequilibrium plots for the region of 22q12 in the genome-wide phase for esophageal squamous cell carcinoma (ESCC). *P*-values were derived from 1 df trend tests in logistic regression models adjusted for age, sex, and study and are shown with the LR values for putative recombination hotspots using SequenceLDhot (vertical bars). Pair-wise r^2 are displayed at the bottom for all SNPs included in the GWAS analysis. Coordinates refer to genome Build 36.1. The figure depicts chromosome 22q12 (27,044,212 – 27,651,008) including the *CHEK2* gene.

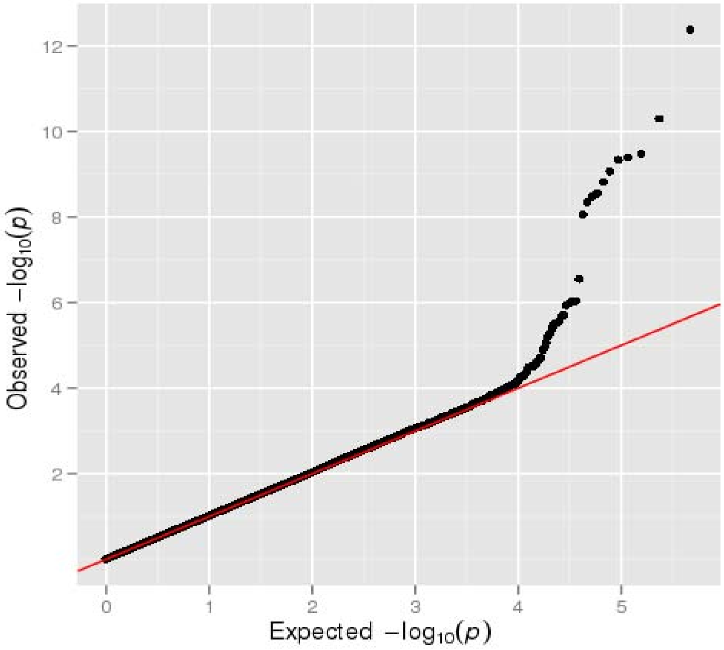


Supplementary Figure 3. Plot of admixture, defined by analysis with STRUCTURE, of subjects in the genome-wide phase. The two individuals with <90% Asian ethnicity were removed from further analyses.

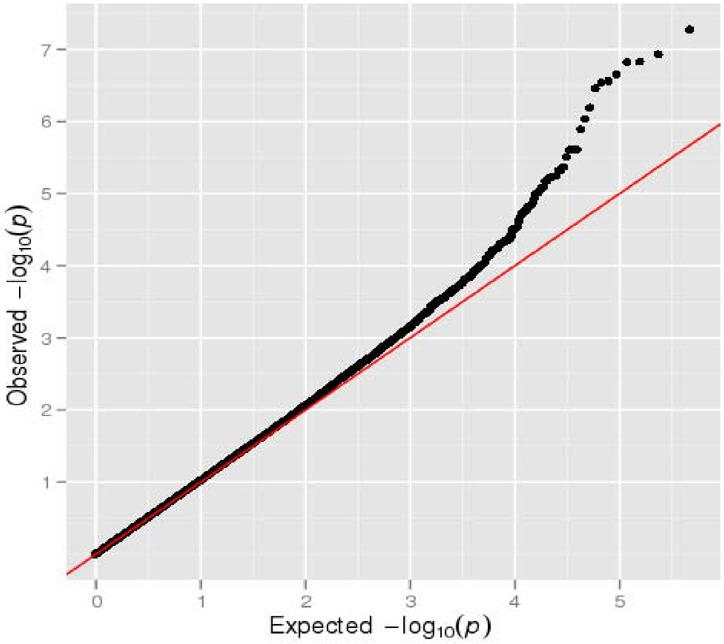


Supplementary Figure 4. Quantile-Quantile (Q-Q) plot of observed versus expected P values for the genome-wide association phase. **a.** gastric cancer; **b.** esophageal squamous cell carcinoma.

a. Gastric cancer



b. Esophageal squamous cell carcinoma



Supplementary Figure 5. Illumina plots for two SNPs tested in replication that achieved genome-wide significance in the overall combined analysis. Sample genotype clusters are shown for two SNPs achieving genome-wide significance in the combined analysis: rs2274223 and rs3781264.