Common 5p15.33 and 6p21.33 variants influence lung cancer risk

Yufei Wang[1], Peter Broderick[1], Emily Webb[1], Xifeng Wu[2], Jayaram Vijayakrishnan[1], Athena Matakidou[1], Mobshra Qureshi[1], Qiong Dong[2], Xiangjun Gu[2], Wei Vivien Chen[2] Margaret Spitz[2], Timothy Eisen[3$], Christopher I. Amos[2$], Richard S. Houlston[1$+]

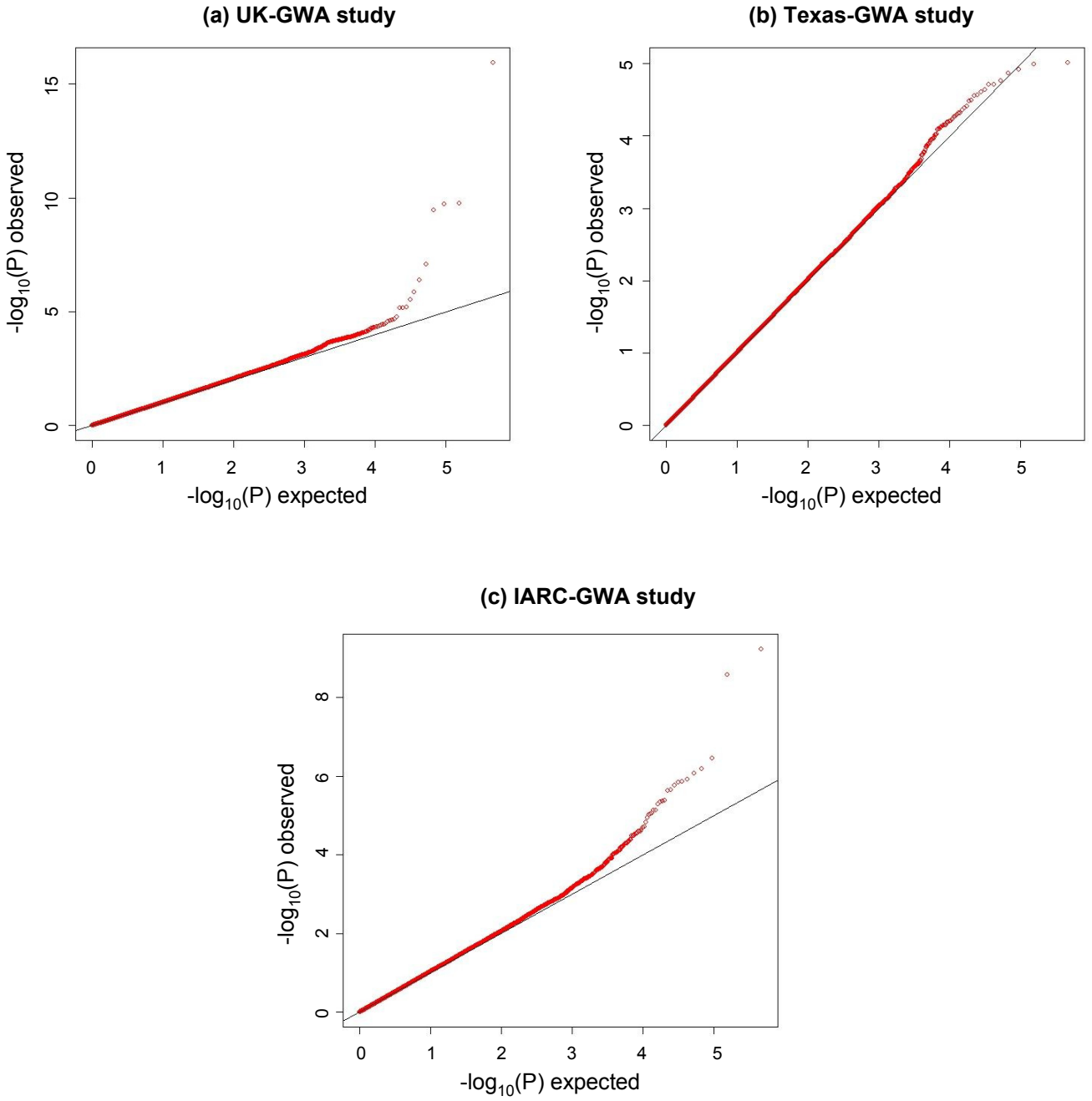
1. Section of Cancer Genetics, Institute of Cancer Research, SM2 5NG. UK.
2. Department of Epidemiology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
3. Department of Oncology, University of Cambridge, Cambridge CB2 2RE. UK.

[$]Joint principal authors
[+]Corresponding author

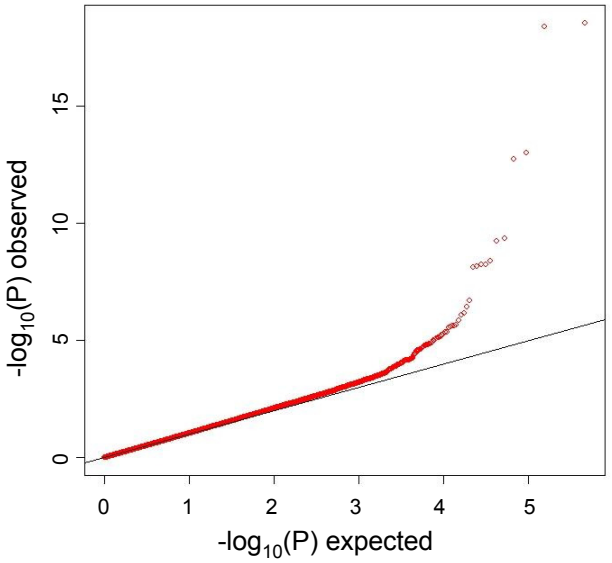
Correspondence should be addressed to:
Richard S Houlston, Section of Cancer Genetics, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey. SM2 5NG. UK.
Tel: +44 (0) 208 722 4175; Fax: +44 (0) 208 722 4365; e-mail: richard.houlston@icr.ac.uk

Supplementary Figure 1: Quantile-quantile (QQ) plots.

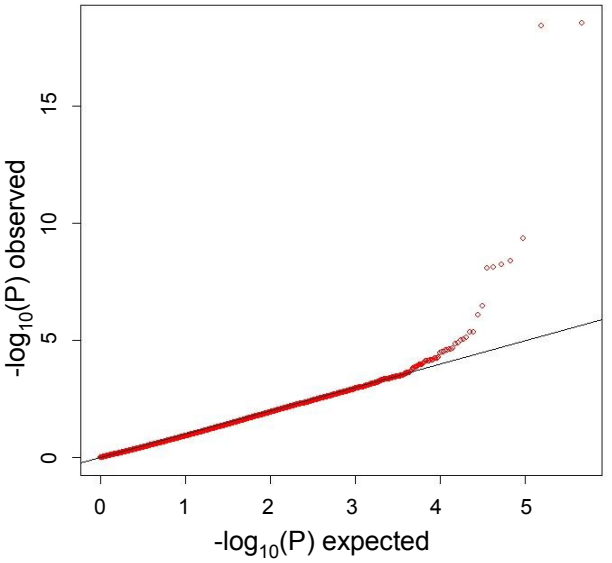
The blue line represents the null hypothesis of no true association. (a) UK-GWA study; (b) Texas-GWA study; (c) IARC-GWA study; (d) Pooled analysis, under fixed effects model; (e) Pooled analysis, under random effects model.



(d) Meta analysis, fixed effects model

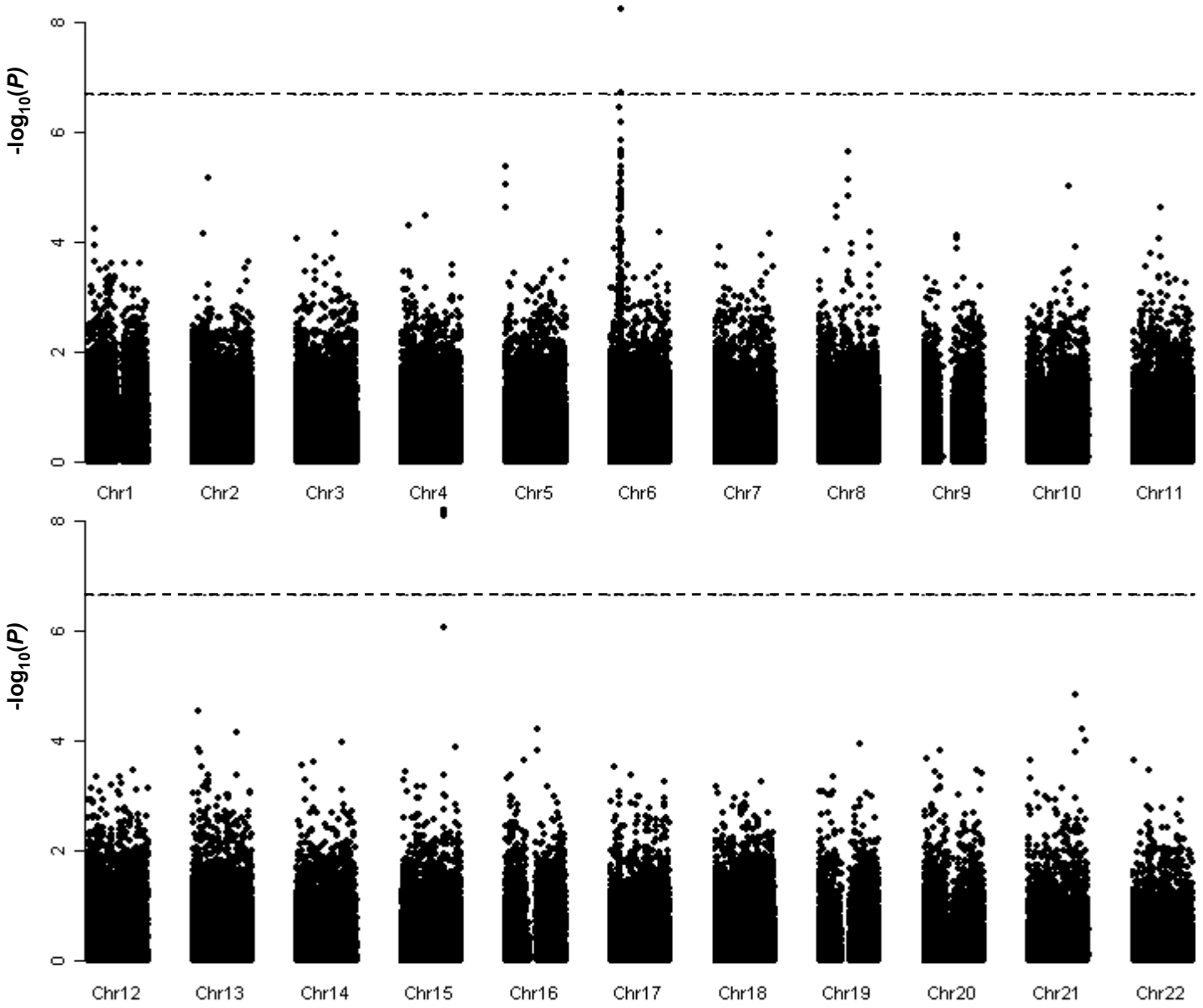


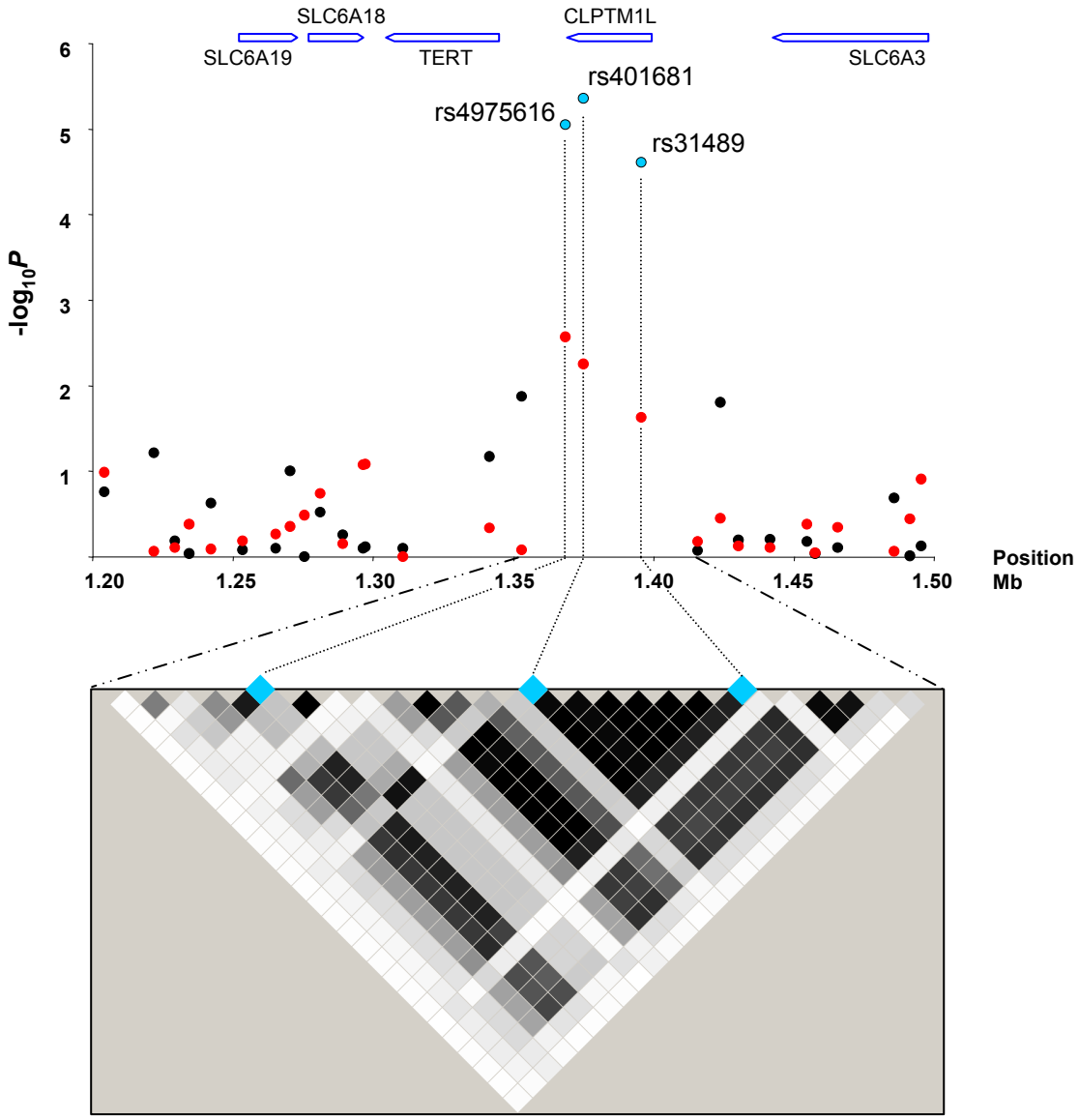
(e) Meta analysis, random effects model



Supplementary Figure 2: Results from meta-analysis of the three lung cancer GWA studies by chromosome.

Single marker association statistics (-log10*P*) are shown as a function of position.





Supplementary Figure 3. LD structure and association results for the 5p15.33 association.

Chromosomal positions based on NCBI Build 36 coordinates, showing Ensembl (release 48) genes. *P* values (as $-\log_{10}$ values; y axis) are shown for SNPs analyzed in UK-GWA study (red circles) and Mantel-Haenzel association test *P* values (black and blue circles). Also shown are the relative position of genes mapping to the region of association,. In the lower panel are the estimated statistics of the square of the correlation coefficient (r^2), derived from Haploview software (v3.2) using HapMap CEU genotypes. The values indicate the LD relationship between each pair of SNPs; the darker the shading, the greater extent of LD.

Supplementary Table 2: Clinico-pathological association testing

A. Association between rs3117582 genotype and gender, histology, age and family history status in cases from UK-GWA, UK replication and Texas GWA studies.

Gender

UK GWA, UK replication

	Male	Female	OR (95% CI)	*P*	P_{trend}
AA	2029 (72.3%)	1114 (71.5%)	1.00 (ref)		
AC	699 (24.9%)	400 (25.7%)	0.96 (0.83-1.11)	0.57	0.65
CC	78 (2.8%)	43 (2.8%)	1.00 (0.67-1.49)	0.98	

Texas GWA

	Male	Female	OR (95% CI)	*P*	P_{trend}
AA	515 (78.3%)	391 (78.8%)	1.00 (ref)		
AC	134 (20.4%)	95 (19.2%)	1.07 (0.80-1.44)	0.65	0.98
CC	9 (1.4%)	10 (2.0%)	0.68 (0.28-1.70)	0.41	

UK GWA, UK replication, Texas GWA

	Male	Female	OR (95% CI)	P^{**}	P_{trend}^{**}
AA			1.00 (ref)		
AC			0.98 (0.86-1.11)	0.76	0.67
CC			0.94 (0.66-1.34)	0.74	

Histology

UK GWA, UK replication

	SCLC	NSCLC	OR (95% CI)	*P*	P_{trend}
AA	762 (73.8%)	2380 (71.5%)	1.00 (ref)		
AC	246 (23.8%)	853 (23.6%)	0.90 (0.76-1.06)	0.21	0.15
CC	25 (2.4%)	95 (2.9%)	0.82 (0.50-1.30)	0.39	

UK GWA, UK replication

	Squamous	Adenocarcinoma	OR (95% CI)	*P*	P_{trend}
AA	1079 (70.9%)	745 (71.6%)	1.00 (ref)		
AC	397 (26.1%)	263 (25.3%)	1.04 (0.87-1.26)	0.66	0.75
CC	46 (3.0%)	32 (3.1%)	0.99 (0.67-1.63)	0.97	

Texas GWA

	Squamous	Adenocarcinoma	OR (95% CI)	*P*	P_{trend}
AA	243 (79.2%)	490 (79.0%)	1.00 (ref)		
AC	57 (18.6%)	121 (19.5%)	0.95 (0.67-1.35)	0.77	0.82
CC	7 (2.3%)	9 (1.5%)	1.57 (0.58-4.26)	0.38	

UK GWA , UK replication, Texas GWA

	Squamous	Adenocarcinoma	OR (95% CI)	P^{**}	P_{trend}^{**}
AA			1.00 (ref)		
AC			1.02 (0.87-1.20)	0.79	0.70
CC			1.07 (0.71-1.63)	0.74	

Age

UK GWA, UK replication

	< 60 years	≥ 60 years	OR (95% CI)	*P*	P_{trend}
AA	742 (72.2%)	2401 (72.0%)	1.00 (ref)		
AC	259 (25.2%)	840 (25.2%)	1.00 (0.85-1.17)	0.98	0.84
CC	27 (2.6%)	94 (2.8%)	0.93 (0.60-1.44)	0.74	

Texas GWA study

	< 60 years	≥ 60 years	OR (95% CI)	*P*	P_{trend}
AA	360 (80.9%)	546 (77.0%)	1.00 (ref)		
AC	76 (17.1%)	153 (21.6%)	0.75 (0.55-1.02)	0.07	0.24
CC	9 (2.0%)	10 (1.4%)	1.37 (0.55-3.39)	0.50	

UK GWA, UK replication, Texas GWA

	< 60 years	≥ 60 years	OR (95% CI)	P^{**}	P_{trend}^{**}
AA			1.00 (ref)		
AC			0.94 (0.81-1.08)	0.37	0.47
CC			1.00 (0.67-1.48)	0.99	

Family history*

UK GWA, UK replication

	Positive	Negative	OR (95% CI)	*P*	P_{trend}
AA	430 (68.5%)	2713 (72.6%)	1.00 (ref)		
AC	174 (27.7%)	925 (24.8%)	1.19 (0.98-1.44)	0.08	0.02
CC	24 (3.8%)	97 (2.6%)	1.56 (0.94-2.49)	0.05	

Texas GWA

	Positive	Negative	OR (95% CI)	*P*	P_{trend}
AA	192 (78.7%)	711 (78.6%)	1.00 (ref)		
AC	48 (19.7%)	179 (19.8%)	0.99 (0.70-1.42)	0.97	0.97
CC	4 (1.64%)	15 (1.7%)	0.99 (0.32-3.01)	0.98	

UK GWA, UK replication, Texas GWA

	Positive	Negative	OR (95% CI)	P^{**}	P_{trend}^{**}
AA			1.00 (ref)		
AC			1.14 (0.96-1.35)	0.13	0.03
CC			1.45 (0.95-2.21)	0.08	

*Defined by having at least one first-degree relative affected with lung cancer
**Adjusted by study

B. Association between rs3117582 genotype and number of cigarettes consumed per day (CPD) in cases and controls from UK-GWA, UK replication and Texas GWA studies.

Cases

UK GWA, UK replication

CPD	AA	AC	CC	OR_{trend} (95% CI)	*P*
≤ 10	393 (71.3%)	146 (26.5%)	12 (2.2%)	1.00 (ref)	
11-20	1281 (71.4%)	452 (25.2%)	60 (3.3%)	1.04 (0.87-1.25)	0.68
21-30	647 (74.5%)	200 (23.0%)	21 (2.4%)	0.89 (0.72-1.10)	0.28
> 30	556 (71.0%)	205 (26.2%)	22 (2.8%)	1.04 (0.84-1.28)	0.74

Texas GWA

CPD	AA	AC	CC	OR_{trend} (95% CI)	*P*
≤ 10	90 (78.9%)	23 (20.2%)	1 (0.9%)	1.00 (ref)	
11-20	347 (79.0%)	83 (18.9%)	9 (2.1%)	1.05 (0.67-1.66)	0.82
21-30	202 (77.4%)	51 (19.5%)	8 (3.1%)	1.18 (0.74-1.89)	0.49
> 30	267 (78.5%)	72 (21.2%)	1 (0.3%)	0.99 (0.60-1.63)	0.97

UK GWA study, UK replication,Texas GWA

CPD	AA	AC	CC	OR_{trend} (95% CI)	P^{**}
≤10				1.00 (ref)	
11-20				1.04 (0.88-1.23)	0.64
21-30				0.93 (0.77-1.13)	0.49
>30				1.03 (0.85-1.25)	0.77

Controls

UK GWA, UK replication

CPD	AA	AC	CC	OR_{trend} (95% CI)	*P*
≤ 10	197 (75.2%)	59 (22.5%)	6 (2.3%)	1.00 (ref)	
11-20	320 (72.9%)	111 (25.3%)	8 (1.8%)	1.08 (0.79-1.48)	0.63
21-30	98 (83.1%)	19 (16.1%)	1 (0.8%)	0.63 (0.38-1.05)	0.08
> 30	58 (70.7%)	21 (25.6%)	3 (3.7%)	1.24 (0.78-2.00)	0.36

Texas GWA

CPD	AA	AC	CC	OR_{trend} (95% CI)	*P*
≤ 10	108 (79.4%)	26 (19.1%)	2 (1.5%)	1.00 (ref)	
11-20	367 (78.9%)	95 (20.4%)	3 (0.6%)	0.98 (0.63-1.53)	0.94
21-30	193 (81.1%)	42 (17.6%)	3 (1.3%)	0.91 (0.56-1.46)	0.69
> 30	226 (76.1%)	67 (22.6%)	4 (1.3%)	1.17 (0.74-1.83)	0.50

UK GWA study, UK replication, Texas GWA

CPD	AA	AC	CC	OR_{trend} (95% CI)	P^{**}
≤ 10				1.00 (ref)	
11-20				1.05 (0.81-1.35)	0.73
21-30				0.76 (0.54-1.07)	0.12
> 30				1.20 (0.87-1.67)	0.27

Cases and Controls

UK GWA, UK replication, Texas GWA

CPD	AA	AC	CC	OR_{trend} (95% CI)	P^{**}
≤ 10				1.00 (ref)	
11-20				1.02 (0.89-1.18)	0.76
21-30				0.89 (0.75-1.04)	0.14
> 30				1.01 (0.86-1.19)	0.87

**Adjusted by study

C. Association between rs401681 genotype and gender, histology, age and family history status in cases from UK-GWA, UK replication and Texas GWA studies.

Gender

UK GWA, UK replication

	Male	Female	OR (95% CI)	*P*	P_{trend}
GG	1006 (35.9%)	551 (35.6%)	1.00 (ref)		
GA	1310 (46.8%)	751 (48.6%)	0.96 (0.84-1.10)	0.52	0.39
AA	483 (17.3%)	245 (15.8%)	1.08 (0.91-1.29)	0.42	

Texas GWA study

	Male	Female	OR (95% CI)	*P*	P_{trend}
GG	209 (31.8%)	187 (37.8%)	1.00 (ref)		
GA	342 (52.0%)	228 (46.1%)	1.34 (1.04-1.74)	0.03	0.14
AA	107 (16.3%)	80 (16.2%)	1.20 (0.84-1.70)	0.32	

UK GWA, UK replication, Texas GWA

Male	Female	OR (95% CI)	P^{**}	P_{trend}^{**}
		1.00 (ref)		
		1.03 (0.91-1.16)	0.64	0.26
		1.10 (0.94-1.30)	0.24	

Histology

UK GWA, UK replication

	SCLC	NSCLC	OR (95% CI)	*P*	P_{trend}
GG	355 (34.5%)	1202 (36.3%)	1.00 (ref)		
GA	495 (48.2%)	1566 (47.2%)	1.07 (0.91-1.25)	0.39	0.58
AA	178 (17.3%)	548 (16.5%)	1.10 (0.89-1.36)	0.37	

UK GWA, UK replication

	Squamous	Adenocarcinoma	OR (95% CI)	*P*	P_{trend}
GG	553 (36.5%)	385 (37.1%)	1.00 (ref)		
GA	721 (47.7%)	479 (46.1%)	1.05 (0.88-1.25)	0.60	0.68
AA	239 (15.8%)	175 (16.8%)	0.95 (0.75-1.20)	0.67	

Texas GWA

	Squamous	Adenocarcinoma	OR (95% CI)	*P*	P_{trend}
GG	100 (32.7%)	216 (34.9%)	1.00 (ref)		
GA	155 (50.7%)	304 (49.0%)	1.10 (0.81-1.50)	0.54	0.57
AA	51 (16.7%)	100 (16.1%)	1.10 (0.73-1.66)	0.65	

UK GWA, UK replication, Texas GWA

		OR (95% CI)	P^{**}	P_{trend}^{**}
		1.00 (ref)		
		1.06 (0.91-1.23)	0.44	0.91
		0.98 (0.80-1.21)	0.89	

Age

UK GWA, UK replication

	< 60 years	≥ 60 years	OR (95% CI)	*P*	P_{trend}
GG	353 (34.4%)	1204 (36.3%)	1.00 (ref)		
GA	495 (48.2%)	1566 (47.2%)	1.08 (0.92-1.26)	0.34	0.26
AA	179 (17.4%)	549 (16.5%)	1.11 (0.90-1.37)	0.31	

Texas GWA

	< 60 years	≥ 60 years	OR (95% CI)	*P*	P_{trend}
GG	148 (33.3%)	248 (35.0%)	1.00 (ref)		
GA	224 (50.3%)	346 (48.9%)	1.08 (0.83-1.41)	0.55	0.62
AA	73 (16.4%)	114 (16.1%)	1.07 (0.75-1.53)	0.70	

UK GWA, UK replication, Texas GWA

	< 60 years	≥ 60 years	OR (95% CI)	P^{**}	P_{trend}^{**}
GG			1.00 (ref)		
GA			1.08 (0.94-1.24)	0.26	0.22
AA			1.10 (0.92-1.32)	0.29	

Family history*

UK GWA, UK replication

	Positive	Negative	OR (95% CI)	*P*	P_{trend}
GG	236 (37.8%)	1321 (35.5%)	1.00 (ref)		
GA	287 (45.9%)	1774 (47.7%)	0.91 (0.75-1.11)	0.30	0.55
AA	102 (16.3%)	626 (16.8%)	0.91 (0.71-1.17)	0.47	

Texas GWA

	Positive	Negative	OR (95% CI)	*P*	P_{trend}
GG	76 (31.3%)	317 (35.0%)	1.00 (ref)		
GA	122 (50.2%)	447 (49.4%)	1.14 (0.83-1.57)	0.43	0.18
AA	45 (18.5%)	141 (15.6%)	1.33 (0.88-2.02)	0.18	

UK GWA, UK replication, Texas GWA

	Positive	Negative	OR (95% CI)	P^{**}	P_{trend}^{**}
GG			1.00 (ref)		
GA			0.96 (0.82-1.13)	0.62	0.92
AA			1.00 (0.81-1.25)	0.96	

*Defined by having at least one first-degree relative affected with lung cancer
**Adjusted by study

D. Association between rs401681 genotype and number of cigarettes consumed per day (CPD) in cases and controls from UK-GWA, UK replication and Texas GWA studies.

Cases

UK GWA, UK replication

CPD	AA	AC	CC	OR_{trend} (95% CI)	*P*
≤ 10	194 (34.7%)	270 (48.3%)	95 (17.0%)	1.00 (ref)	
11-20	646 (35.8%)	850 (47.1%)	307 (17.0%)	0.98 (0.85-1.12)	0.75
21-30	288 (34.6%)	395 (47.5%)	149 (17.9%)	1.02 (0.88-1.19)	0.79
> 30	283 (36.3%)	376 (48.3%)	120 (15.4%)	0.94 (0.80-1.09)	0.40

Texas GWA

CPD	AA	AC	CC	OR_{trend} (95% CI)	*P*
≤ 10	35 (31.0%)	56 (49.6%)	22 (19.5%)	1.00 (ref)	
11-20	166 (37.8%)	209 (47.6%)	64 (14.6%)	0.78 (0.58-1.06)	0.11
21-30	83 (31.8%)	133 (51.0%)	45 (17.2%)	0.94 (0.68-1.29)	0.69
> 30	112 (32.9%)	172 (50.6%)	56 (16.5%)	0.90 (0.66-1.23)	0.51

UK GWA, UK replication, Texas GWA

CPD	AA	AC	CC	OR_{trend} (95% CI)	*P***
≤ 10				1.00 (ref)	
11-20				0.94 (0.83-1.07)	0.34
21-30				1.00 (0.88-1.15)	0.95
> 30				0.93 (0.81-1.07)	0.30

Controls

UK GWA, UK replication

CPD	AA	AC	CC	OR_{trend} (95% CI)	*P*
≤ 10	99 (37.6%)	125 (47.5%)	39 (14.8%)	1.00 (ref)	
11-20	130 (29.4%)	231 (52.3%)	81 (18.3%)	1.29 (1.03-1.61)	0.03
21-30	44 (37.6%)	52 (44.4%)	21 (17.9%)	1.07 (0.78-1.46)	0.68
> 30	22 (26.2%)	51 (60.7%)	11 (13.1%)	1.24 (0.86-1.78)	0.25

Texas GWA

CPD	AA	AC	CC	OR_{trend} (95% CI)	*P*
≤ 10	37 (27.2%)	81 (59.6%)	18 (13.2%)	1.00 (ref)	
11-20	142 (30.5%)	240 (51.5%)	84 (18.0%)	1.03 (0.78-1.38)	0.82
21-30	67 (28.2%)	117 (49.2%)	54 (22.7%)	1.20 (0.88-1.64)	0.25
> 30	90 (30.3%)	158 (53.2%)	49 (16.5%)	1.00 (0.74-1.37)	0.98

UK GWA, UK replication, Texas GWA

CPD	AA	AC	CC	OR_{trend} (95% CI)	*P***
≤ 10				1.00(ref)	
11-20				1.18 (0.99-1.41)	0.06
21-30				1.13 (0.91-1.41)	0.27
> 30				1.10 (0.87-1.39)	0.44

Cases and Controls

UK GWA, UK replication, Texas GWA

CPD	AA	AC	CC	OR_{trend} (95% CI)	P^{**}
≤ 10				1.00 (ref)	
11-20				1.01 (0.92-1.12)	0.79
21-30				1.07 (0.96-1.20)	0.23
> 30				1.00 (0.89-1.12)	0.95

**Adjusted by study

SUPPLEMENTARY METHODS

Study participants

UK-GWA study: Cases with pathologically confirmed lung cancer were ascertained through the Genetic Lung Cancer Predisposition Study (GELCAPS). A standardized proforma was used to collect information on demographic information and information on smoking. Further details about the design and conduct of this population-based study are described in published material[1]. The current analysis is based on 1,952 patients (1,166 male, 786 female; mean age at diagnosis 62 years, SD 12). All were British residents and self reported to be of European Ancestry. Individuals from the 1958 Birth cohort served as source of controls. Comprehensive information on the 1958 Birth Cohort can be obtained through: http://www.cls.ioe.ac.uk/studies.asp?section=000100020003

Details of case and control ascertainment and matching criteria, as well as the genotyping of Texas and IARC GWA studies have been published previously[2,3]. Briefly, Texas-GWA study: Cases and controls were ascertained from a case-control study that has been ongoing at the U.T. M.D. Anderson Cancer Center since 1991. Cases are newly diagnosed, histologically-confirmed NSCLC patients presenting at M.D. Anderson Cancer and who had not previously received treatment other than surgery. Controls are healthy individuals seen for routine care at Kelsey-Seybold Clinics in the Houston Metropolitan area, frequency matched to cases according to their smoking behaviour, age in 5 year categories, ethnicity, and sex. Former smoking controls were further frequency matched to former smoking cases according to the number of years since smoking cessation (in 5 year categories). The IARC-GWA study was based on a lung cancer case-control study conducted in 6 central European countries (Czech Republic, Hungary, Poland, Romania, Russia and Slovakia) between 1998 and 2002. Cases were individuals with newly diagnosed lung cancer. Controls (hospital patients or population controls) were frequency matched to cases by sex, age, geographical area and period of recruitment.

UK-Replication: An additional series of 2,484 cases (1,690 male, 794 female; mean age at diagnosis 72 years, SD 7) with pathologically confirmed lung cancer were ascertained through GELCAPS. Blood samples were obtained from 3,036 healthy individuals (1,497 male, 1,539 female; mean age 61 years, SD 11) recruited to the National Cancer Research Network genetic epidemiological studies, the National Study of Colorectal Cancer (NSCCG; 1999-2006; n=541), GELCAPS (1999-2004; n=1,520); and the Royal Marsden Hospital Trust/Institute of Cancer Research Family History and DNA Registry (1999-2004; n=975). These controls were the spouses or unrelated friends of patients with malignancies. None had a personal history of

malignancy at time of ascertainment. All were British residents and self reported to be of European Ancestry.

Ethical approval for the UK study was obtained from the London Multi-Centre Research Ethics Committee (MREC/98/2/67) in accordance with the tenets of the Declaration of Helsinki. All participants provided informed consent.

Genotyping

DNA was extracted from samples using conventional methodologies and quantified using PicoGreen (Invitrogen, Carlsbad, USA). A GWA study of tag SNPs was conducted using the Illumina Human550 BeadChips according to the manufacturer's protocols (Illumina, San Diego, USA). DNA samples with GenCall scores <0.25 at any locus were considered "no calls". A DNA sample was deemed to have failed if it generated genotypes at <95% of loci. A SNP was deemed to have failed if fewer than 95% of DNA samples generated a genotype at the locus. To ensure quality of genotyping, a series of duplicate samples were genotyped and cases and controls were genotyped in the same batches.

Genotyping of rs3117582 and rs401681 in the UK-Replication series was conducted by competitive allele-specific PCR KASPar chemistry (KBiosciences Ltd, Hertfordshire, UK); primers and probes used are available on request. Genotyping quality control was tested using duplicate DNA samples within studies and SNP assays. For all SNPs, >99.9% concordant results were obtained.

Statistical analysis

Statistical analysis was undertaken using S Plus v7.0 (Insightful, New York, US), R v2.6 and STATA v8.0 (Station College, Texas, US) Software. Genotype data were used to search for duplicates and closely related individuals amongst all samples in Phase 1. Identity by state values were calculated for each pair of individuals, and for any pair with allele sharing >80%, the sample generating the lowest call rate was removed from further analysis. In Phase 1, genotyped samples were excluded from further analyses for the following reasons: gender discrepancy (n=6), duplicated (n=0), relatedness (n=0).

The adequacy of the case-control matching and possibility of differential genotyping of cases and controls were formally evaluated using Q-Q plots of $-\log_{10}(P)$ values (based on the 90% least significant SNPs). Deviation of the genotype frequencies in the controls from those expected under Hardy-Weinberg Equilibrium (HWE) was assessed by χ^2 test (1 degree of

freedom, d.f.), or Fisher's exact test where an expected cell count was <5. Comparison of the difference in number of associations observed and expected was made using the binomial test.

The association between each SNP and risk was assessed by the allele test. Odds ratios (ORs) and associated 95% confidence intervals (CIs) were calculated by unconditional logistic regression. Associations by gender, histology (NSCLC, SCLC) and age were examined by logistic regression in case-only analyses. We examined the relationship between genotype and smoking in two ways, firstly number cigarettes smoked per day within each genotype group was assessed by χ^2 test and secondly risks associated with genotype were adjusted by logistic regression for number of pack years.

Meta-analysis was conducted using standard methods for combining raw data based on the Mantel-Haenszel method and weighted average of study-specific estimates of the ORs, using inverse variance weights[4]. Cochran's Q statistic to test for heterogeneity and the I^2 statistic to quantify the proportion of the total variation due to heterogeneity were calculated.

The sibling relative risk attributable to a given SNP was calculated using the formula[5]:

$$\lambda^* = \frac{p(pr_2 + qr_1)^2 + q(pr_1 + q)^2}{[p^2 r_2 + 2pqr_1 + q^2]^2}$$

where p is the population frequency of the minor allele, q=1-p, and r_1 and r_2 are the relative risks (estimated as OR) for heterozygotes and rare homozygotes, relative to common homozygotes. Assuming a multiplicative interaction the proportion of the familial risk attributable to a SNP was calculated as $\log(\lambda^*)/\log(\lambda_0)$, where λ_0 is the overall familial relative risk estimated from epidemiological studies, assumed to be 1.8[6].

Bioinformatics

We used Haploview software (v3.2) to infer the LD structure of the genome in the regions containing loci associated with disease risk.

URLs

Online Inheritance in Man: http://www.ncbi.nlm.nih.gov/sites/entrez

The R suite can be found at http://www.r-project.org/

Detailed information on the tag SNP panel can be found at http://www.illumina.com/

dbSNP: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=snp

HAPMAP: http://www.hapmap.org/

http://pipeline.lbl.gov/cgi-bin/gateway2

GELCAPS: http://pfsearch.ukcrn.org.uk/StudyDetail.aspx?TopicID=1&StudyID=781

- http://www.dh.gov.uk/assetRoot/04/01/45/13/04014513.pdf

National Study of Colorectal Cancer Genetics (NSCCG): http://pfsearch.ukcrn.org.uk/StudyDetail.aspx?TopicID=1&StudyID=1269

ICR-RMH Family history and DNA resource: http://intratest.icr.ac.uk/tissueres/index.htm

MACH1: http://www.sph.umich.edu/csg/abecasis/MACH/

1958 Birth Cohort: http://www.cls.ioe.ac.uk/studies.asp?section=000100020003

Central Europe data from IARC-GWAS: http://www.ceph.fr/cancer

KBiosciences: http://www.kbioscience.co.uk/

REFERENCES

1. Matakidou, A. et al. Case-control study of familial lung cancer risks in UK women. *Int J Cancer* **116**, 445-50 (2005).
2. Amos, C.I. et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* **40**, 616-22 (2008).
3. Hung, R.J. et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* **452**, 633-7 (2008).
4. Petitti, D. Meta-analysis Decision Analysis and Cost-Effectiveness Analysis. *Oxford, New York, Oxford***.**(1994).
5. Cox, A. et al. A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet* **39**, 352-8 (2007).
6. Matakidou, A., Eisen, T. & Houlston, R.S. Systematic review of the relationship between family history and lung cancer risk. *Br J Cancer* **93**, 825-33 (2005).