Carcinogenesis vol.31 no.2 pp.234–238, 2010
doi:10.1093/carcin/bgp287
Advance Access publication December 2, 2009

Role of 5p15.33 (*TERT-CLPTM1L*), 6p21.33 and 15q25.1 (*CHRNA5-CHRNA3*) variation and lung cancer risk in never-smokers

Yufei Wang, Peter Broderick, Athena Matakidou, Timothy Eisen[1] and Richard S.Houlston*

Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK, [1]Department of Oncology, University of Cambridge, Cambridge CB2 2RE, UK

*To whom correspondence should be addressed. Tel: +44 208 722 4175; Fax: +44 208 722 4365;
Email: richard.houlston@icr.ac.uk

Genome-wide association studies have provided evidence that common variation at 5p15.33 (*TERT-CLPTM1L*), 6p21.33 and 15q25.1 (*CHRNA5-CHRNA3*) influences lung cancer risk. To examine if variation at any of these loci influences the risk of lung cancer in never-smokers, we compared 5p15.33-*TERT* (rs2736100), 5p15.33-*CLPTM1L* (rs4975616), 6p21.33-*BAT3* (rs3117582), 15q25.1-*CHRNA3* (rs8042374) and 15q25.1-*CHRNA3* (rs12914385) genotypes in a series of 239 never-smoker lung cancer cases and 553 never-smoker controls. A statistically significant association between lung cancer risk and 5p15.33 genotypes was found: rs2736100 (odds ratio = 0.78, 95% confidence interval: 0.63–0.97; $P = 0.02$), rs4975616 (odds ratio = 0.69, 95% confidence interval: 0.55–0.85; $P = 7.95 \times 10^{-4}$), primarily for adenocarcinoma. There was no evidence of association between 6p21.33 or 15q25.1 variation and risk of lung cancer. This analysis provides evidence that *TERT-CLPTM1L* variants may influence the risk of lung cancer outside the context of tobacco smoking.

Introduction

Lung cancer is a major cause of cancer death worldwide with >1 million deaths each year (1). While the disease is largely preventable as most cases are due to tobacco smoking, global statistics estimate that ~15% of lung cancer cases in men and ~50% in women are not attributable to smoking. Thus, if considered a separate category, lung cancer in never-smokers would rank as the seventh most common cause of cancer death worldwide, with a higher incidence than cancers of the cervix, pancreas and prostate.

Although all the major histological types of lung cancer are associated with smoking, the association is strongest for small-cell lung cancer and is stronger for squamous cell than for adenocarcinoma. In contrast, adenocarcinoma is recognized to be the commonest form of lung cancer developing in never-smokers (2,3).

Recent genome-wide association studies have shown that polymorphic variation at 5p15.33, 6p21.33 and 15q25.1 influences the risk of developing lung cancer (4–8). The 5p15.33 locus contains two independent association signals, which map to *TERT* and *CLPTM1L* (7,9). Evidence for an indirect effect of polymorphic variation at 15q25.1 on lung cancer risk is provided by some of the studies that have shown a relationship between nicotinic acetylcholine receptor (*CHRN*) gene variation and smoking behavior (5,9,10), with smokers who carry *CHRNA3* and *CHRNA5* variants smoking more intensively, resulting in a higher exposure to tobacco carcinogens. Given that there is evidence of a familial basis to lung cancer in never-smokers (11), it is therefore of considerable interest to further examine if any of these loci impact on the risk of lung cancer independent of smoking.

Abbreviations: NSCLC, non-small-cell lung cancer; SNP, single-nucleotide polymorphism.

To examine this possibility, we studied the relationship between 5p15.33, 6p21.33 and 15q25.1 variants and risk in never-smokers comparing the frequency of 5p15.33 (rs2736100 and rs4975616), 6p21.33 (rs3117582) and 15q25.1 (rs8042374 and rs12914385) risk genotypes in 239 cases and 553 controls. Our findings provide evidence that 5p15.33 variation represents a risk factor for the development of lung cancer in never-smokers, primarily for adenocarcinoma.

Materials and methods

Study participants

Never-smoker status was defined on the basis of having had a lifetime exposure of <100 cigarettes. A series of 239 never-smoker lung cancer cases (138 male and 101 female; median age at diagnosis 67 years, range from 26 to 87 years) were ascertained through the Genetic Lung Cancer Predisposition Study (GELCAPS) (12) and have been the subject of previous analyses. All the cases had pathologically confirmed lung cancer: 39 small-cell lung cancer (21 males and 18 females) and 200 non-small-cell lung cancer [NSCLC; of which 112 adenocarcinoma (77 males and 35 females), 48 squamous cell (18 males and 30 females) and 40 other NSCLC histology (22 males and 18 females)]. Five hundred and fifty-three never-smoker controls (105 males and 448 females; median age at sampling 63 years, range from 21 to 91 years) were obtained through GELCAPS. These controls were the spouses or unrelated friends of patients. None had a personal history of malignancy at time of ascertainment. Both cases and controls were British residents and self-reported to be of European Ancestry.

Collection of blood samples and clinicopathologiocal information from patients and controls was undertaken with informed consent and ethical review board approval in accordance with the tenets of the Declaration of Helsinki.

Single-nucleotide polymorphism selection and genotyping

DNA was extracted from samples using conventional methodologies and quantified using PicoGreen (Invitrogen, Carlsbad, CA). To evaluate 5p15.33, 6p21.33 and 15q25.1 variation on lung cancer risk in never-smokers, we derived rs2736100, rs4975616, rs3117582, rs8042374 and rs12914385 genotypes. We selected these single-nucleotide polymorphisms (SNPs) for analysis as they provide the primary evidence for an association between 5p15.33, 6p21.33 and 15q25.1 variation and lung cancer risk in unselected cases (7,9). Genotyping was conducted using Illumina Human550 BeadChips and Illumina Infinium custom arrays according to the manufacturer's protocols as described previously (6). Briefly, 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. An SNP was deemed to have failed if <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. By virtue of the genotyping platform employed in addition to deriving rs2736100, rs4975616, rs3117582, rs8042374 and rs12914385 genotypes, we obtained data on 94 additional SNPs mapping to 5p15.33 (*TERT-CLPTM1L*; $n = 6$), 6p21.33 ($n = 53$) and 15q25.1-*CHRNA3-CHRNA5* ($n = 35$).

Statistical analysis

Statistical analyses were undertaken in R (v2.8) software. Deviation of the genotype frequencies in the controls from those expected under Hardy–Weinberg equilibrium as assessed by χ^2 test. The risk of lung cancer associated with SNP genotype was estimated by odds ratios using logistic regression with and without adjustment for age, sex and occupational exposure.

Bioinformatics

We used Haploview software (v3.2) for haplotype analysis and to infer the linkage disequilibrium structure around *IRF4*. Prediction of the untyped SNPs was carried out using IMPUTE (13) on HapMap (Jan07 on National Center for Biotechnology Information B35 assembly, dbSNPb125). Imputed data integrity was verified where possible, by cross-checking the concordance of imputed genotypes with that of available Illumina SNP genotype data.

Table I. Association between 5p15.33 (rs2736100 and rs4975616), 15q25.1 (rs8042374 and rs12914385), 6p21.33 (rs3117582) and overall lung cancer risk and stratified by lung cancer histology

Histology	Locus	Genotype	*N*		Unadjusted		Adjusted[a]	
			Case (%)	Control (%)	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
Overall	5p15.33 *TERT*	GG	82 (34.3)	158 (28.6)	1.00 (Ref)		1.00 (Ref)	
	(rs2736100)	TG	115 (48.1)	259 (46.8)	0.86 (0.61–1.21)	0.38	0.81 (0.57–1.16)	0.25
		TT	42 (17.6)	136 (24.6)	0.60 (0.38–0.92)	0.02	0.59 (0.38–0.93)	0.02
			T allele (OR_{trend}/P_{trend})		0.78 (0.63–0.97)	0.02	0.76 (0.61–0.95)	0.02
	5p15.33 *CLPTM1L*	AA	107 (44.8)	179 (32.4)	1.00 (Ref)		1.00 (Ref)	
	(rs4975616)	AG	100 (41.8)	266 (48.1)	0.63 (0.45–0.88)	6.18×10^{-3}	0.56 (0.40–0.80)	1.17×10^{-3}
		GG	32 (13.4)	108 (19.5)	0.50 (0.31–0.79)	2.88×10^{-3}	0.53 (0.33–0.85)	8.40×10^{-3}
			G allele (OR_{trend}/P_{trend})		0.69 (0.55–0.85)	7.95×10^{-4}	0.68 (0.54–0.86)	1.10×10^{-4}
	15q25.1 *CHRNA3*	AA	135 (56.5)	314 (56.8)	1.00 (Ref)		1.00 (Ref)	
	(rs8042374)	AG	86 (35.98)	206 (37.3)	0.97 (0.70–1.34)	0.86	1.01 (0.73–1.41)	0.94
		GG	18 (7.5)	33 (6.0)	1.27 (0.69–2.33)	0.44	1.37 (0.73–2.59)	0.33
			G allele (OR_{trend}/P_{trend})		1.05 (0.82–1.34)	0.70	1.10 (0.85–1.42)	0.47
	15q25.1 *CHRNA3*	CC	100 (41.8)	217 (39.2)	1.00 (Ref)		1.00 (Ref)	
	(rs12914385)	TC	109 (45.6)	260 (47.0)	0.91 (0.66–1.26)	0.57	0.90 (0.64–1.27)	0.56
		TT	30 (12.6)	76 (13.7)	0.86 (0.53–1.39)	0.53	0.81 (0.49–1.34)	0.41
			T allele (OR_{trend}/P_{trend})		0.92 (0.74–1.15)	0.47	0.91 (0.72–1.15)	0.43
	6p21.33 *BAT3*	AA	172 (72.0)	413 (74.7)	1.00 (Ref)		1.00 (Ref)	
	(rs3117582)	AC	63 (26.4)	124 (22.4)	1.22 (0.86–1.73)	0.27	1.22 (0.85–1.76)	0.28
		CC	4 (1.7)	16 (2.9)	0.60 (0.20–1.82)	0.37	0.61 (0.19–1.92)	0.40
			C allele (OR_{trend}/P_{trend})		1.06 (0.78–1.45)	0.7	1.05 (0.77–1.43)	0.78
SCLC	5p15.33 *TERT*	GG	10 (25.6)	158 (28.6)	1.00 (Ref)		1.00 (Ref)	
	(rs2736100)	TG	18 (46.2)	259 (46.8)	1.10 (0.49–2.44)	0.82	1.06 (0.47–2.41)	0.88
		TT	11 (28.2)	136 (24.6)	1.28 (0.53–3.10)	0.59	1.40 (0.56–3.51)	0.47
			T allele (OR_{trend}/P_{trend})		1.13 (0.72–1.77)	0.59	1.21 (0.76–1.94)	0.42
	5p15.33 *CLPTM1L*	AA	16 (41.0)	179 (32.4)	1.00 (Ref)		1.00 (Ref)	
	(rs4975616)	AG	16 (41.0)	266 (48.1)	0.67 (0.33–1.38)	0.28	0.64 (0.31–1.32)	0.23
		GG	7 (18.0)	108 (19.5)	0.73 (0.29–1.82)	0.49	0.91 (0.35–2.37)	0.84
			G allele (OR_{trend}/P_{trend})		0.81 (0.51–1.30)	0.39	0.85 (0.52–1.40)	0.53
	15q25.1 *CHRNA3*	AA	25 (64.1)	314 (56.8)	1.00 (Ref)		1.00 (Ref)	
	(rs8042374)	AG	13 (33.3)	206 (37.3)	0.79 (0.40–1.58)	0.51	0.82 (0.40–1.67)	0.59
		GG	1 (2.6)	33 (6.0)	0.38 (0.05–2.90)	0.35	0.46 (0.06–3.58)	0.46
			G allele (OR_{trend}/P_{trend})		0.73 (0.41–1.30)	0.29	0.76 (0.42–1.39)	0.38
	15q25.1 *CHRNA3*	CC	16 (41.0)	217 (39.2)	1.00 (Ref)		1.00 (Ref)	
	(rs12914385)	TC	19 (48.7)	260 (47.0)	0.99 (0.50–1.97)	0.98	0.89 (0.44–1.81)	0.75
		TT	4 (10.3)	76 (13.7)	0.71 (0.23–2.20)	0.56	0.57 (0.18–1.83)	0.34
			T allele (OR_{trend}/P_{trend})		0.89 (0.55–1.45)	0.64	0.82 (0.50–1.35)	0.44
	6p21.33 *BAT3*	AA	27 (69.2)	413 (74.7)	1.00 (Ref)		1.00 (Ref)	
	(rs3117582)	AC	11 (28.2)	124 (22.4)	1.36 (0.65–2.81)	0.41	1.37 (0.65-2.87)	0.41
		CC	1 (2.6)	16 (2.9)	0.96 (0.12–7.48)	0.97	0.96 (0.12-7.75)	0.97
			T allele (OR_{trend}/P_{trend})		1.20 (0.66–2.18)	0.55	1.19 (0.64–2.20)	0.59
NSCLC	5p15.33 *TERT*	GG	72 (36.0)	158 (28.6)	1.00 (Ref)		1.00 (Ref)	
	(rs2736100)	TG	97 (48.5)	259 (46.8)	0.82 (0.57–1.18)	0.29	0.78 (0.54–1.14)	0.20
		TT	31 (15.5)	136 (24.6)	0.50 (0.31–0.81)	4.62×10^{-3}	0.50 (0.30–0.81)	5.53×10^{-3}
			T allele (OR_{trend}/P_{trend})		0.72 (0.58–0.91)	5.71×10^{-3}	0.71 (0.56–0.90)	4.12×10^{-3}
	5p15.33 *CLPTM1L*	AA	91 (45.5)	179 (32.4)	1.00 (Ref)		1.00 (Ref)	
	(rs4975616)	AG	84 (42.0)	266 (48.1)	0.62 (0.44–0.88)	8.00×10^{-3}	0.55 (0.38–0.80)	1.57×10^{-3}
		GG	25 (12.5)	108 (19.5)	0.46 (0.28–0.75)	2.17×10^{-3}	0.50 (0.30–0.83)	7.73×10^{-3}
			G allele (OR_{trend}/P_{trend})		0.66 (0.52–0.84)	6.17×10^{-4}	0.66 (0.51–0.84)	9.68×10^{-4}
	15q25.1 *CHRNA3*	AA	110 (55.0)	314 (56.8)	1.00 (Ref)		1.00 (Ref)	
	(rs8042374)	AG	73 (36.5)	206 (37.3)	1.01 (0.72–1.43)	0.95	1.03 (0.73–1.48)	0.85
		GG	17 (8.5)	33 (6.0)	1.47 (0.79–2.74)	0.23	1.54 (0.81–2.95)	0.19
			G allele (OR_{trend}/P_{trend})		1.12 (0.86–1.45)	0.40	1.15 (0.88–1.51)	0.30
	15q25.1 *CHRNA3*	CC	84 (42.0)	217 (39.2)	1.00 (Ref)		1.00 (Ref)	
	(rs12914385)	TC	90 (45.0)	260 (47.0)	0.89 (0.63–1.27)	0.53	0.89 (0.62–1.28)	0.54
		TT	26 (13.0)	76 (13.7)	0.88 (0.53–1.47)	0.64	0.84 (0.49–1.43)	0.52
			T allele (OR_{trend}/P_{trend})		0.93 (0.73–1.18)	0.53	0.92 (0.72–1.18)	0.52
	6p21.33 *BAT3*	AA	145 (72.5)	413 (74.7)	1.00 (Ref)		1.00 (Ref)	
	(rs3117582)	AC	52 (26.0)	124 (22.4)	1.19 (0.82–1.74)	0.35	1.21 (0.82–1.78)	0.33
		CC	3 (1.5)	16 (2.9)	0.53 (0.15–1.86)	0.32	0.57 (0.16–2.06)	0.39
			T allele (OR_{trend}/P_{trend})		1.03 (0.75–1.42)	0.85	1.03 (0.74–1.44)	0.85

Table I. *Continued*

Histology	Locus	Genotype	*N*		Unadjusted		Adjusted[a]	
			Case (%)	Control (%)	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
Adenocarcinoma	5p15.33 *TERT*	GG	39 (34.8)	158 (28.6)	1.00 (Ref)		1.00 (Ref)	
	(rs2736100)	TG	60 (53.6)	259 (46.8)	0.94 (0.60–1.47)	0.78	0.91 (0.58-1.44)	0.70
		TT	13 (11.6)	136 (24.6)	0.39 (0.20–0.76)	5.40×10^{-3}	0.40 (0.21–0.79)	8.26×10^{-3}
		T allele (OR_{trend}/P_{trend})			0.68 (0.51–0.91)	0.01	0.68 (0.51–0.91)	0.01
	5p15.33 *CLPTM1L*	AA	59 (52.7)	179 (32.4)	1.00 (Ref)		1.00 (Ref)	
	(rs4975616)	AG	42 (37.5)	266 (48.1)	0.48 (0.31–0.74)	1.01×10^{-3}	0.45 (0.29–0.71)	5.00×10^{-4}
		GG	11 (9.8)	108 (19.5)	0.31 (0.16–0.61)	7.99×10^{-4}	0.33 (0.16–0.66)	1.61×10^{-3}
		G allele (OR_{trend}/P_{trend})			0.53 (0.39–0.72)	6.04×10^{-5}	0.52 (0.38–0.72)	6.42×10^{-5}
	15q25.1 *CHRNA3*	AA	58 (51.8)	314 (56.8)	1.00 (Ref)		1.00 (Ref)	
	(rs8042374)	AG	42 (37.5)	206 (37.3)	1.10 (0.71–1.70)	0.66	1.13 (0.73–1.74)	0.59
		GG	12 (10.7)	33 (6.0)	1.97 (0.96–4.03)	0.06	2.11 (1.01–4.39)	0.05
		G allele (OR_{trend}/P_{trend})			1.28 (0.93–1.75)	0.13	1.30 (0.95–1.80)	0.10
	15q25.1 *CHRNA3*	CC	49 (43.8)	217 (39.2)	1.00 (Ref)		1.00 (Ref)	
	(rs12914385)	TC	51 (45.5)	260 (47.0)	0.87 (0.56–1.34)	0.52	0.88 (0.57–1.36)	0.55
		TT	12 (10.7)	76 (13.7)	0.70 (0.35–1.38)	0.30	0.67 (0.33–1.34)	0.26
		T allele (OR_{trend}/P_{trend})			0.85 (0.62–1.15)	0.28	0.85 (0.63–1.16)	0.30
	6p21.33 *BAT3*	AA	82 (73.2)	413 (74.7)	1.00 (Ref)		1.00 (Ref)	
	(rs3117582)	AC	28 (25.0)	124 (22.4)	1.14 (0.71–1.83)	0.59	1.14 (0.71–1.83)	0.60
		CC	2 (1.8)	16 (2.9)	0.63 (0.14–2.79)	0.54	0.66 (0.15–2.95)	0.58
		T allele (OR_{trend}/P_{trend})			1.01 (0.68–1.51)	0.95	1.01 (0.67–1.52)	0.96
Squamous	5p15.33 *TERT*	GG	17 (35.4)	158 (28.57)	1.00 (Ref)		1.00 (Ref)	
	(rs2736100)	TG	23 (47.9)	259 (46.84)	0.83 (0.43–1.59)	0.57	0.73 (0.37–1.47)	0.38
		TT	8 (16.7)	136 (24.59)	0.55 (0.23–1.31)	0.17	0.57 (0.23–1.41)	0.23
		T allele (OR_{trend}/P_{trend})			0.75 (0.50–1.14)	0.18	0.74 (0.47–1.17)	0.20
	5p15.33 *CLPTM1L*	AA	19 (39.6)	179 (32.37)	1.00 (Ref)		1.00 (Ref)	
	(rs4975616)	AG	21 (43.8)	266 (48.10)	0.74 (0.39–1.42)	0.37	0.61 (0.31–1.22)	0.17
		GG	8 (16.7)	108 (19.53)	0.70 (0.30–1.65)	0.41	0.88 (0.35–2.21)	0.79
		G allele (OR_{trend}/P_{trend})			0.82 (0.53–1.25)	0.35	0.84 (0.53–1.25)	0.48
	15q25.1 *CHRNA3*	AA	30 (62.5)	314 (56.8)	1.00 (Ref)		1.00 (Ref)	
	(rs8042374)	AG	15 (31.3)	206 (37.3)	0.76 (0.40–1.45)	0.41	0.76 (0.39–1.49)	0.43
		GG	3 (6.3)	33 (6.0)	0.95 (0.28–3.29)	0.94	1.11 (0.30–4.02)	0.88
		G allele (OR_{trend}/P_{trend})			0.86 (0.52–1.42)	0.55	0.88 (0.52–1.51)	0.65
	15q25.1 *CHRNA3*	CC	19 (39.6)	217 (39.24)	1.00 (Ref)		1.00 (Ref)	
	(rs12914385)	TC	21 (43.8)	260 (47.02)	0.92 (0.48–1.76)	0.81	0.75 (0.38–1.51)	0.43
		TT	8 (16.7)	76 (13.7)	1.20 (0.51–2.86)	0.68	0.83 (0.32–2.17)	0.71
		T allele (OR_{trend}/P_{trend})			1.06 (0.69–1.62)	0.80	0.90 (0.57–1.43)	0.66
	6p21.33 *BAT3*	AA	38 (79.2)	413 (74.7)	1.00 (Ref)		1.00 (Ref)	
	(rs3117582)	AC	9 (18.8)	124 (22.4)	0.79 (0.37–1.68)	0.54	0.81 (0.36–1.79)	0.60
		CC	1 (2.1)	16 (2.9)	0.68 (0.09–5.26)	0.71	0.84 (0.10–7.12)	0.87
		T allele (OR_{trend}/P_{trend})			0.80 (0.43–1.50)	0.49	0.83 (0.42–1.62)	0.58

OR, odds ratio; CI, confidence interval.
[a]Adjusted for age, gender and occupational exposure.

Results

In keeping with published data, the proportion of never-smoker cases diagnosed with NSCLC was significantly higher than that seen in lung cancer cases developing in smokers/former smokers in GELCAPS (84 and 47%, respectively).

Genotypes were obtained for >95% of cases and controls for all five SNPs (Table I); hence there was no evidence of any systematic bias in genotyping. There was complete concordance between duplicate samples. The frequency of alleles of each SNP in controls in our study was similar to previously published data on the Northern European population. Furthermore, there was no evidence of population stratification as the genotype distribution in controls for each of the five SNPs satisfied Hardy–Weinberg equilibrium.

The two 5p15.33 SNPs rs2736100 and rs4975616, which annotate to *TERT* and *CLPTM1L*, respectively, showed a statistically significant association with lung cancer risk in a dose-dependent fashion ($P = 0.02$ and 7.95×10^{-4}, respectively; Table I); the rs4975616 association being statistically significant even with adjustment for multiple testing. Given the biological differences between the histological forms of lung cancer, we examined the association between SNP genotypes and risk by subtype. In this restricted analysis, there was enhanced evidence of an association between 5p15.33 rs2736100 and rs4975616 genotypes and risk of NSCLC, especially with the risk of adenocarcinoma (Table I). There was no evidence that either SNP was associated with small-cell lung cancer risk although the number of cases with this histology was small thereby limiting power to demonstrate a relationship. We sought to establish whether we could identify SNPs better correlated with disease risk at 5p15.33 (1,333,028-1,412,838 bp, encompassing *TERT* and *CLPTM1L*) through imputation of untyped SNPs by referencing HapMap. The linkage disequilibrium structure encompassing the *TERT* locus precludes imputation of any additional SNPs at this locus. In total, an 11 additional HapMap SNPs mapping to the remainder of the interval were successfully imputed (Figure 1). Using imputed data did not, however, provide for a significantly stronger association at the *CLPM1L* locus (Figure 1).

In contrast to 5p15.33, there was no evidence that variation at either 6p21.33 or 15q25.1 influenced risk of all lung cancer or any of the different subtypes in never-smokers either through analysis of directly typed SNPs or imputed SNP data (Table I; Figure 1).

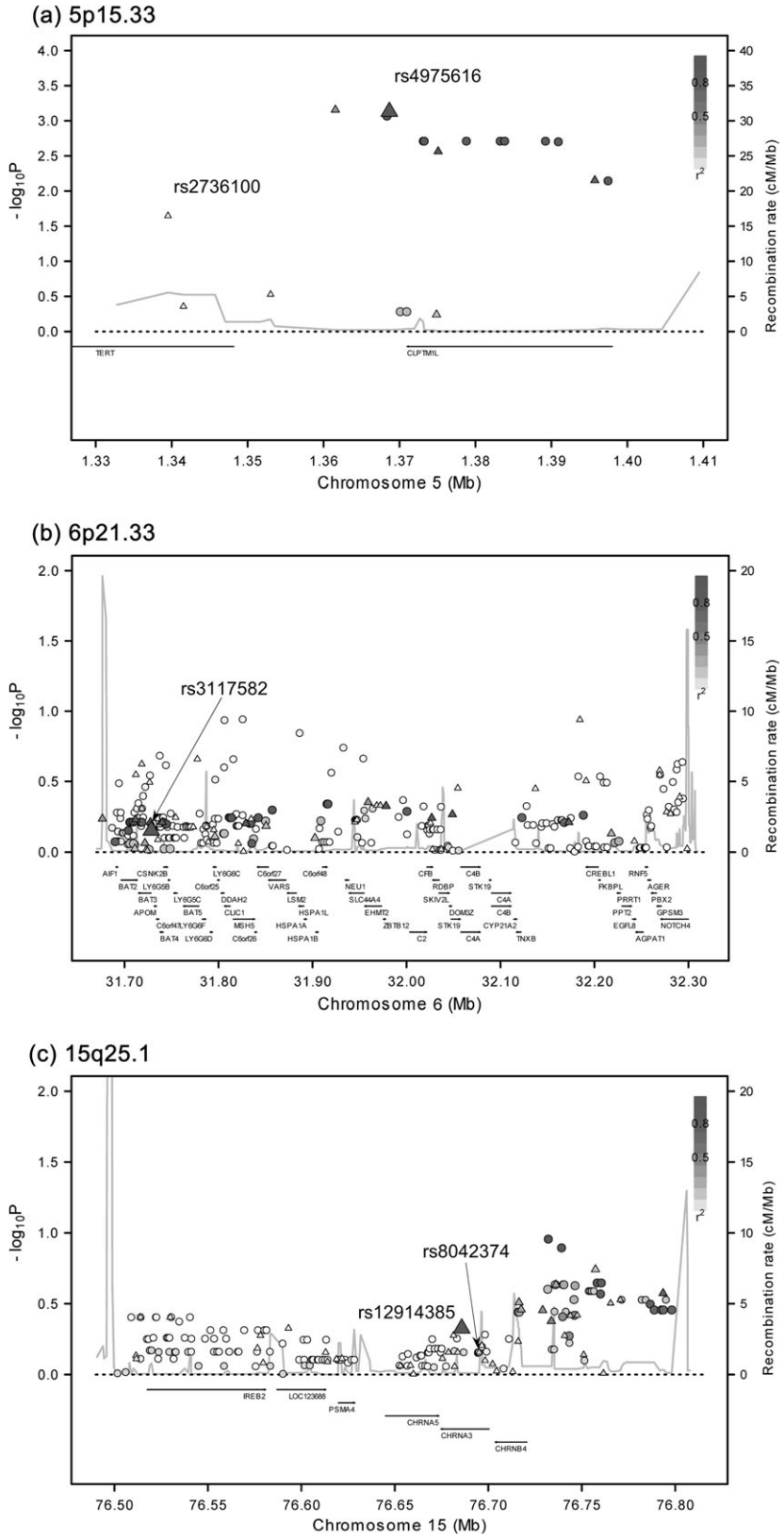

Fig. 1. Regional plots of the 5p15.33, 6p21.33 and 15q25.1 lung cancer loci. (**a**) Illustration of the 5p15.33 locus, with the local recombination rate plotted in light gray over this 80 kb chromosomal segment encompassing rs2736100 and rs4975616. Triangles represent SNPs directly genotyped and circles represent imputed SNPs. The color intensity of each symbol reflects the extent of linkage disequilibrium (LD) with rs4975616—dark-gray ($r^2 > 0.8$) through to white ($r^2 < 0.3$). Physical positions are based on build 36 of the human genome. rs2736100 is located in intron 2 of *TERT* and rs4975616 localizes 2.5 kb telomeric of *CLPTM1L*. (**b**) The 6p21.33 locus. We present all SNPs, directly genotyped and imputed, which localize within a 600 kb window defined by recombination peaks. The color intensity of each symbol reflects the extent of LD with rs3117572, which maps to *BAT3*. (**c**) The 15q25.1 locus. We present all SNPs, directly genotyped and imputed, which localize within a 300 kb window defined by recombination peaks. The color intensity of each symbol reflects the extent of LD with rs12914385, which maps to *CHRNA3*.

Discussion

Our findings provide evidence that polymorphic variation at 5p15.33 plays a role in determining the risk of developing lung cancer in never-smokers. Specifically, we observed an association with the variants of *TERT* and *CLPTM1L*. Moreover, despite the relatively modest size of our case series, we were able to demonstrate that the association is primarily mediated through propensity to develop adenocarcinoma. This observation is in keeping with epidemiological data showing that this form of NSCLC is the principal lung cancer developing in never-smokers.

A major strength of our study is that these data, although a subgroup analysis of a larger parent study, have been systematically ascertained in a consistent fashion and by making use of genome-wide association data bias from population stratification as a source of confounding has been avoided. It is, however, fully acknowledged that as our study was small, replication of findings is required.

The definition of what constitutes a never-smoker varies considerably in the literature. Here we have adopted the conservative criterion, which has general acceptance, as being those individuals who have smoked $<$100 cigarettes in their lifetime. A potential limitation of the study is that we did not collect data on environmental tobacco smoke. While environmental tobacco smoke has been demonstrated to confer an increased risk of lung cancer, environmental tobacco smoke is, however, a relatively weak carcinogen (14) and is therefore unlikely to have biased studied findings significantly. Population stratification is a concern in all association studies as a source of bias as the frequency of genotypes for many polymorphic variants differ markedly between ethnic groups. We have sought to further minimize this form of bias by excluding subjects with non-Western European ethnicity. Moreover, the frequency of SNP genotypes in controls were directly comparable with those seen in previously published data on the UK population (9).

In contrast to 5p15.33, we did not find any significant association between 6p21.33 or 15q25.1 variation and lung cancer risk. Our findings are therefore in keeping with the assertion that the *CHRNA* association is probably to be indirect, mediated primarily by the propensity to smoke. Failure to demonstrate an association with 6p21.33 may be a function of study power, as even in the genome-wide association study datasets the association was very modest.

Both *CLPTM1L* (cisplatin resistance-related protein 9, alias *CRR9*; MIM 612585) and *TERT* (telomerase reverse transcriptase; MIM 187270), which map to 5p15.33, represent attractive candidates for lung cancer susceptibility *a priori* assuming that the causal variant exerts an influence through a *cis* effect. The biology of TERT makes it an attractive candidate for a gene that influences lung cancer risk and moreover association between rs2736100 risk allele and shorter telomere length has recently been reported (15). High levels of polycyclic aromatic hydrocarbons adducts correlate with lung cancer risk and as a major effect of platinum on cells is through adduct formation, *CLPTM1L* is also an attractive lung cancer susceptibility gene as it encodes a transcript whose overexpression has been linked to *cis*-platinum resistance (16). Hence, while it will be challenging to identify the precise mechanism by which the 5p15.33 variants affects lung cancer development, determining the causal basis may prove highly informative, endorsing etiological hypotheses or suggesting new ones that merit testing via gene/environment-specific studies relevant to lung cancer development unrelated to smoking.

Web resources (URLs)

Illumina: http://www.illumina.com/; dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP/; HAPLOVIEW: http://www.broadinstitute.org/haploview/haploview; University of California Santa Cruz: http://genome.ucsc.edu/; HapMap: http://hapmap.ncbi.nlm.nih.gov/; IMPUTE: http://mathgen.stats.ox.ac.uk/impute/impute.html; Online Mendelian Inheritance in Man: http://www.ncbi.nlm.nih.gov/omim

Funding

Cancer Research UK (C1298/A8780 and C1298/A8362—Bobby Moore Fund for Cancer Research UK). A.M. was the recipient of a clinical research fellowship from the Allan J Lerner Fund. We are also grateful to NCRN, HEAL and Sanofi-Aventis. We acknowledge National Health Service funding for the Royal Marsden Biomedical Research Centre.

Acknowledgements

We would like to thank all individuals who participated in this study and the clinicians who took part in the GELCAPS consortium. This study made use of genotyping data on the 1958 Birth Cohort; these data were generated and generously supplied to us by Panagiotis Deloukas of the Wellcome Trust Sanger Institute. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk.

Conflict of Interest Statement: None declared.

References

1. Parkin,D.M. *et al*. (2005) Global cancer statistics, 2002. *CA Cancer J. Clin.*, **55**, 74–108.
2. Toh,C.K. *et al*. (2006) Never-smokers with lung cancer: epidemiologic evidence of a distinct disease entity. *J. Clin. Oncol.*, **24**, 2245–2251.
3. Sun,S. *et al*. (2007) Lung cancer in never smokers—a different disease. *Nat. Rev. Cancer*, **7**, 778–790.
4. Amos,C.I. *et al*. (2008) Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat. Genet.*, **40**, 616–622.
5. Thorgeirsson,T.E. *et al*. (2008) A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature*, **452**, 638–642.
6. Wang,Y. *et al*. (2008) Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat. Genet.*, **40**, 1407–1409.
7. McKay,J.D. *et al*. (2008) Lung cancer susceptibility locus at 5p15.33. *Nat. Genet.*, **40**, 1404–1406.
8. Hung,R.J. *et al*. (2008) A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature*, **452**, 633–637.
9. Broderick,P. *et al*. (2009) Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. *Cancer Res.*, **69**, 6633–6641.
10. Freathy,R.M. *et al*. (2009) A common genetic variant in the 15q24 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRNB4) is associated with a reduced ability of women to quit smoking in pregnancy. *Hum. Mol. Genet.*, **18**, 2922–2927.
11. Matakidou,A. *et al*. (2005) Systematic review of the relationship between family history and lung cancer risk. *Br. J. Cancer*, **93**, 825–833.
12. Eisen,T. *et al*. (2008) Identification of low penetrance alleles for lung cancer: the GEnetic Lung CAncer Predisposition Study (GELCAPS). *BMC Cancer*, **8**, 244.
13. Marchini,J. *et al*. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906–913.
14. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. (2004) Tobacco smoke and involuntary smoking. *IARC Monogr. Eval. Carcinog. Risks Hum.*, **83**, 1–1438.
15. Rafnar,T. *et al*. (2009) Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat. Genet.*, **41**, 221–227.
16. Yamamoto,K. *et al*. (2001) A novel gene, CRR9, which was up-regulated in CDDP-resistant ovarian tumor cell line, was associated with apoptosis. *Biochem. Biophys. Res. Commun.*, **280**, 1148–1154.

Received August 11, 2009; revised November 10, 2009; accepted November 12, 2009