

## Review

# Association of the *CHRNA5-A3-B4* Gene Cluster With Heaviness of Smoking: A Meta-Analysis

Jennifer J. Ware, B.Sc.,<sup>1,2</sup> Marianne B. M. van den Bree, Ph.D.,<sup>1</sup> & Marcus R. Munafò, Ph.D.<sup>2</sup>

<sup>1</sup> Department of Psychological Medicine, Cardiff University, Heath Park Campus, Cardiff, UK

<sup>2</sup> School of Experimental Psychology, University of Bristol, Bristol, UK

Corresponding Author: Jennifer Ware, B.Sc., Department of Psychological Medicine, Cardiff University, 1<sup>st</sup> Floor Neuadd Merionydd, Heath Park Campus, Cardiff CF14 4YS, UK. Telephone: +44-0-29 2068 7910; Fax: +44-0-29 2068 7915; E-mail: warejj@cardiff.ac.uk

Received February 24, 2011; accepted May 05, 2011

## Abstract

**Introduction:** Variation in the *CHRNA5-A3-B4* gene cluster is a promising candidate region for smoking behavior and has been linked to multiple smoking-related phenotypes (e.g., nicotine dependence) and diseases (e.g., lung cancer). Two single nucleotide polymorphisms (SNPs), rs16969968 in *CHRNA5* and rs1051730 in *CHRNA3*, have generated particular interest.

**Methods:** We evaluated the published evidence for association between rs16969968 ( $k = 27$  samples) and rs1051730 ( $k = 44$  samples) SNPs with heaviness of smoking using meta-analytic techniques. We explored which SNP provided a stronger genetic signal and investigated study-level characteristics (i.e., ancestry, disease state) to establish whether the strength of association differed across populations. We additionally tested for small study bias and explored the impact of year of publication.

**Results and Conclusions:** Meta-analysis indicated compelling evidence of an association between the rs1051730/rs16969968 variants and daily cigarette consumption (fixed effects:  $B = 0.91$ , 95%  $CI = 0.77, 1.06$ ,  $p < .001$ ; random effects:  $B = 1.01$ , 95%  $CI = 0.81, 1.22$ ,  $p < .001$ ), equivalent to a per-allele effect of approximately 1 cigarette/day. SNP rs1051730 was found to provide a stronger signal than rs16969968 in stratified analyses ( $p_{diff} = .028$ ), although this difference was only qualitatively observed in the subset of samples that provided data on both SNPs. While the functional relevance of rs1051730 is unknown, it may be a strong tagging SNP for functional haplotypes in this region.

## Introduction

Tobacco use is one of the greatest public health concerns facing modern society, currently accounting for the deaths of 5.4 million people a year (World Health Organization [WHO], 2008). While overall prevalence rates suggest that tobacco use is now falling in high-income countries, the epidemic has shifted to the developing world where tobacco use is increasing (WHO, 2008). Current predictions estimate that tobacco use will

account for the deaths of 8 million individuals a year by 2030 (Mathers & Loncar, 2006). Furthering our understanding of the genetic contribution of smoking-related behaviors may ultimately facilitate advances in cessation treatment through the identification of novel treatment targets, which may help to reduce the substantial health concern associated with tobacco use.

There is consistent evidence from twin and adoption studies that genetic factors contribute to the etiology of cigarette smoking, playing an important role in smoking initiation, progression to heavy use, and persistence (Fowler et al., 2007; Kendler et al., 1999; Lessov et al., 2004; Munafò & Johnstone, 2008; Sullivan & Kendler, 1999). A meta-analysis (Li, Cheng, Ma, & Swan, 2003) reported that genetic factors were responsible for approximately 50% of the variation noted in both initiation and persistence. However, despite a large number of candidate gene studies (focusing primarily on targets in relevant neurotransmitter pathways and enzymes associated with nicotine metabolism), few reported associations between gene variants and smoking-related phenotypes have proven to replicate reliably. Recently, however, variation in the 15q24 nicotinic acetylcholine receptor (nAChR) gene cluster *CHRNA5-A3-B4* (responsible for encoding  $\alpha 5$ ,  $\alpha 3$ , and  $\beta 4$  nAChR subunits) has shown promise as a candidate region for smoking behavior. Polymorphisms in this cluster have been linked to multiple smoking-related phenotypes, such as nicotine dependence (Bierut et al., 2008; L. S. Chen, Johnson, et al., 2009; Grucza et al., 2008; S. F. Saccone et al., 2007; Spitz, Amos, Dong, Lin, & Wu, 2008; Thorgeirsson et al., 2008), smoking quantity (Amos et al., 2008; Berrettini et al., 2008; Keskitalo et al., 2009; Lips et al., 2009; Stevens et al., 2008; Thorgeirsson et al., 2008), and smoking cessation (Freathy et al., 2009) as well as smoking-related diseases, such as lung cancer (Amos et al., 2008; Hung et al., 2008; Lips et al., 2009; P. Liu et al., 2008; Spitz et al., 2008; Thorgeirsson et al., 2008), chronic obstructive pulmonary disease (Pillai et al., 2009; Young et al., 2008), peripheral arterial disease (Thorgeirsson et al., 2008), and upper aerodigestive tract cancers (Lips et al., 2009). Whether the associations noted between these variants and diseases are direct or mediated via the variants' association with smoking-related behaviors, or both, remains a topic of debate (see Thorgeirsson & Stefansson, 2010).

doi: 10.1093/ntr/ntr118

Advance Access published on November 9, 2011

© The Author 2011. Published by Oxford University Press on behalf of the Society for Research on Nicotine and Tobacco. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited

Two single nucleotide polymorphisms (SNPs) in the *CHRNA5-A3-B4* region, rs16969968 in *CHRNA5* and rs1051730 in *CHRNA3*, have generated particular interest with respect to smoking-related behaviors. The SNP rs16969968 is notable as a missense mutation, resulting in an amino acid change (aspartate to asparagine) at position 398 in the  $\alpha$ -5 subunit protein (Bierut, 2010). This polymorphism appears to be of functional significance—in vitro studies have demonstrated that  $\alpha$ 4 $\beta$ 2 $\alpha$ 5 receptors with the aspartic acid variant exhibit a greater response to a nicotine agonist than  $\alpha$ 4 $\beta$ 2 $\alpha$ 5 receptors containing the asparagine substitution (Bierut et al., 2008). Reduced nAChR function may therefore predispose to nicotine dependence. This finding complements evidence from a number of population-based genetic studies/candidate gene association studies linking this specific polymorphism to phenotypes, such as nicotine dependence and smoking quantity (Gruza et al., 2008; Lips et al., 2009; S. F. Saccone et al., 2007, N. L. Saccone et al., 2009). An association between this variant and smoking quantity was also identified in a recent meta-analysis of primarily new unpublished data conducted by N. L. Saccone et al. (2010). The SNP rs1051730 (a polymorphism in strong linkage disequilibrium [LD] with rs16969968 and frequently used interchangeably with this SNP in the literature) has also shown promise as a strong candidate for further research. Associations between this variant and both nicotine dependence and smoking quantity have been replicated in several independent samples (X. Chen, Chen, et al., 2009; S. F. Saccone et al., 2007; Thorgeirsson et al., 2008). Moreover, three recent genome-wide meta-analyses have highlighted an association between this locus and smoking quantity (Furberg et al., 2010; J. Z. Liu et al., 2010; Thorgeirsson et al., 2010).

The *CHRNA5-A3-B4* cluster clearly shows promise as a candidate region for smoking behavior, with an established link between SNPs rs16969968 and rs1051730 and smoking quantity. To date, however, no study-level meta-analysis has been conducted to determine the strength of association between rs1051730 specifically and smoking quantity, to determine whether the strength of association between both SNPs and smoking quantity differs according to sample ancestry or disease status, or to assess small study bias. It is of note that typically only one of these two SNPs tends to be used in analyses in the wider literature for practical reasons. It is therefore of interest to determine whether they should continue to be used interchangeably. We sought to (a) evaluate the strength of evidence for the association between the rs16969968 and rs1051730 SNPs and heaviness of smoking (both in pooled and in independent analyses), as measured by daily cigarette consumption, using meta-analytic techniques to synthesize existing published data; (b) explore which SNP provides a stronger genetic signal; (c) test for the possibility that small study sample sizes may have biased findings; (d) explore the impact of year of publication; and (e) investigate the impact of ancestry and disease state as potential moderating variables.

## Methods

### Selection of Studies for Study Inclusion

Studies were included that reported data on the *CHRNA5* polymorphism rs16969968 and/or the *CHRNA3* polymorphism rs1051730 and smoking quantity. If data regarding smoking

quantity were presented categorically, or were available but not reported by genotype, we contacted the authors to determine whether data in an appropriate format for inclusion were available. Three attempts were made to contact study authors. If these attempts did not result in the provision of data, the study was included but coded as “data not available.” Studies in any language reporting data on samples of any ethnic origin were included as were studies reporting data on either single-sex samples or samples including both males and females.

Studies were excluded if no data on smoking quantity were available, neither of the SNPs of interest was investigated, or if extreme smoking quantity phenotypes had been selected for analysis. Reviews, letters to the editor, and editorials were excluded if these did not present new or relevant data. Family-based studies were also excluded. Additionally, studies were excluded if an inappropriate study design was employed (e.g., DNA pooling).

### Search Strategy

The search was performed in Scopus and PubMed. These databases were searched from the first date available in each database up to May 12, 2010, using the following search terms: “*CHRNA5* or *CHRNA3* or *CHRNA4*”; “rs16969968 or rs1051730”; “smok\* and 15q2\*.” Once articles had been collected, references were hand searched for additional studies of interest.

The titles and abstracts of studies identified by these search strategies were examined, and those clearly fitting the inclusion or exclusion criteria were retained or excluded, respectively (initial screening conducted by JW). Of the remaining studies, a more thorough examination of the full text and supplementary material (if available) was required to determine retention or rejection (full-text assessment conducted by JW). All duplications were deleted. Where studies reported previously published data, we included data from only one of the publications, namely that reporting the largest sample. Ten percent of all studies identified by the search strategy were additionally assessed for eligibility by a second reviewer (interrater agreement >90%). Disagreements between reviewers were resolved by mutual consent.

### Data Extraction

For each study the following data were extracted: (a) authors and year of publication, (b) sample characteristics (ancestry and disease state), (c) SNP(s) studied, and (d) *M*, *SD*, and *N* for cigarettes per day by genotype. Genotype frequencies were used to calculate deviation from Hardy–Weinberg equilibrium (HWE). Ancestry was coded as European or “other,” given the paucity of studies reporting data on non-European samples. To be coded as European, a sample had to be comprised of at least 95% European individuals.

### Data Analyses

Given the high LD between rs16969968 and rs1051730 (European:  $r^2 = .902$ , Japanese/Chinese:  $r^2 = 1.000$ , African:  $r^2 =$  unavailable; calculated using HapMap data in conjunction with SNAP [<http://www.broadinstitute.org/mpg/snap/ldsearchpw.php>]), we initially conducted pooled analyses incorporating data from all samples, regardless of SNP studied, omitting one dataset if data on both SNPs had been collected for a sample. The standard additive model of genetic action was used for evaluation. Small study bias was assessed using the Egger test (Egger, Davey Smith, Schneider, & Minder, 1997) for both pooled and independent

SNP analyses. The impact of year of publication on effect size estimate was also examined. Data were analyzed within both a fixed- and random-effects framework. Individual study effect sizes were pooled to generate a summary effect estimate and 95% *CI*, the significance of which was determined using a Z test. Stratified analyses by sample ancestry (European vs. other) and disease state (control/population vs. disease/partial disease) were conducted to ascertain the potential moderating effects of these variables. We also explored which SNP provided a stronger genetic signal. The differences in pooled effect sizes were determined using a Z test.

Between-study heterogeneity was examined using a chi-square test and quantified through calculation of  $I^2$ —the conventional bounds for low, medium, and high heterogeneity based on the  $I^2$  statistic being 25%, 50%, and 75%, respectively.

Data were analyzed with Comprehensive Meta-Analysis Version 2 statistical software (Biostat, Englewood, NJ).

## Results

### Study Selection

The search of Scopus and PubMed databases provided 585 records. Two additional records were identified through other

sources (hand-searching references of identified papers). After adjusting for duplications, 432 records remained. Of these, 325 were discarded because after reviewing the abstracts, it appeared that these papers clearly did not meet the required criteria. The full texts of the remaining 107 studies were examined in detail (Figure 1). Of these, 37 were identified for inclusion in the meta-analysis (Supplementary Tables S1 and S2).

### Characteristics of Included Studies

A total of 37 studies published between 2006 and 2010 were identified for inclusion in the meta-analysis. Of these, 19 studies (comprising  $k = 57$  independent samples and a further  $k = 15$  duplicate samples) provided data contributing to the meta-analysis (Amos et al., 2008; Breiting et al., 2009; Broderick et al., 2009; X. Chen, Chen, et al., 2009; Etter et al., 2009; Freathy et al., 2009; Greenbaum, Rigbi, Teltsh, & Lerer, 2009; Greenbaum et al., 2006; Grucza et al., 2008; Keskitalo et al., 2009; Lambrechts et al., 2010; Landi et al., 2009; Le Marchand et al., 2008; Lips et al., 2009; Schwartz, Cote, Wenzlaff, Land, & Amos, 2009; Shiraishi et al., 2009; Spitz et al., 2008; Young et al., 2008; Zienolddiny et al., 2009). The remaining 18 studies identified for inclusion did not contribute data as data from these studies were not available or the sample(s) featured had been included in another study, which we had already included in our analyses (refer to Supplementary Tables S1 and S2; Baker et al., 2009; Caporaso et al., 2009; Conti et al., 2008; Hung et al., 2008; P. Liu et al.,

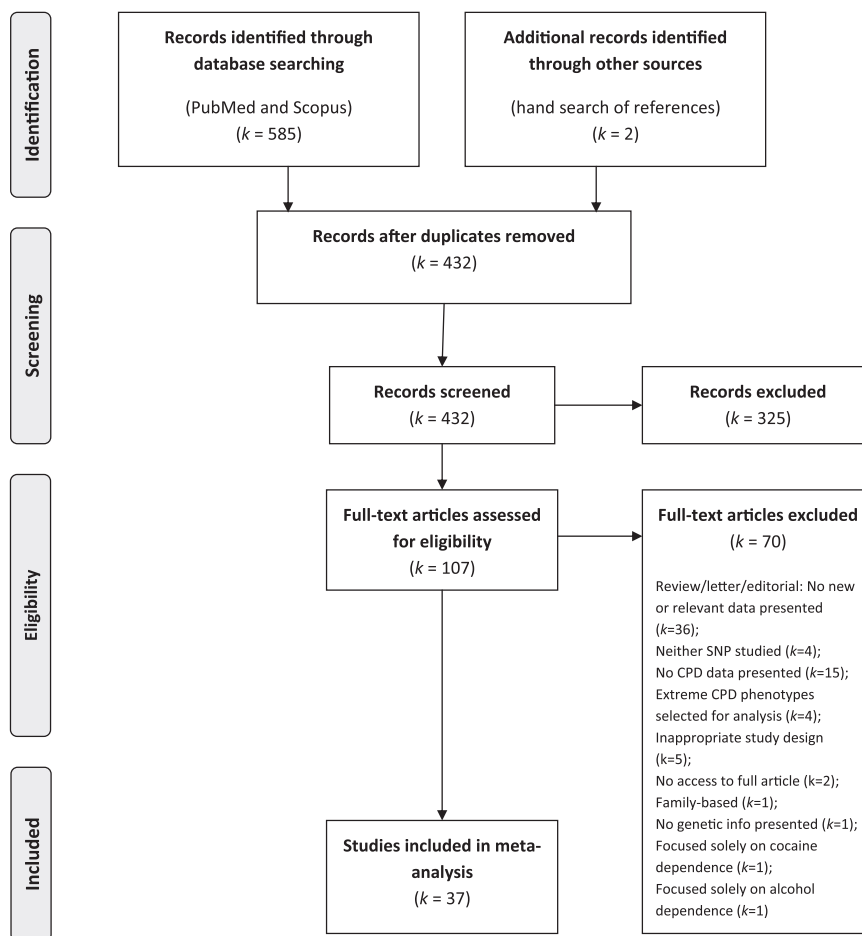


Figure 1. Flow diagram of study selection.

2008, 2010; McKay et al., 2008; Pillai et al., 2009; Ray et al., 2010; Rigbi et al., 2008; Sherva et al., 2008; Thorgeirsson et al., 2008, 2010; Furberg et al., 2010; Weiss et al., 2008; Wu et al., 2009; Yang et al., 2010; Young et al., 2009).

A total of 50 samples provided data on participants of predominantly European ancestry and 7 on participants of other ancestry. Twenty-one samples reported data on control/population samples and 36 on disease/partial disease samples (e.g., lung cancer cases). Forty-four samples reported data on rs1051730 and 27 on rs16969968 (NB:  $k = 15$  samples reported data on both SNPs). Two samples reported genotype frequencies that deviated substantially from HWE (NB: one additional non-HWE sample was excluded from analyses as the homozygous risk genotype group contained only one participant). Minor allele frequencies of rs16969968 (A) and rs1051730 (T) ranged from 0.03 to 0.43 (median = 0.35). The wide ranges were primarily driven by the inclusion of non-European samples in which the minor alleles were rare.

### Smoking Quantity

#### Primary Analyses

Meta-analysis indicated strong evidence of association between the rs1051730/rs16969968 variants and daily cigarette consumption (fixed effects:  $B = 0.91$ , 95%  $CI = 0.77, 1.06$ ,  $p < .001$ ; random effects:  $B = 1.01$ , 95%  $CI = 0.81, 1.22$ ,  $p < .001$ ; Table 1, Figure 2). There was evidence of moderate between-study heterogeneity,  $Q(56) = 85.46$ ,  $p = .007$ ,  $I^2 = 34\%$ .

Egger’s test indicated weak evidence of small study bias,  $t(55) = 1.84$ ,  $p_{\text{one tailed}} = .036$ . To adjust for this, we utilized Duval and Tweedie’s (2000) “trim and fill” method. This method removes studies with outlying effect size values identified on funnel plots until symmetry is achieved and then replaces these along with imputed “mirror” values in order to retain symmetry. This correction had minimal effect on the overall effect estimate (adjusted value:  $B = 0.85$ , 95%  $CI = 0.62, 1.07$ ). There was no evidence of an association between effect size estimate and year of publication ( $p = .27$ ).

### Stratified Analyses

Results from all stratified analyses, under both fixed- and random-effects models, are displayed in Table 1. Random-effects model outcomes are presented here.

Evidence for an association between rs1051730/rs16969968 variants and heaviness of smoking was observed irrespective of stratification by study level characteristics.

Two of the 57 samples included in our analysis deviated from HWE. We compared this pair of samples to the group of 55 samples, which did not deviate from HWE. In both groups, there was strong evidence to suggest an association between rs1051730/rs16969968 and daily cigarette consumption (HWE:  $B = 0.99$ , 95%  $CI = 0.78, 1.20$ ,  $p < .001$ ; non-HWE:  $B = 1.89$ , 95%  $CI = 0.76, 3.03$ ,  $p = .001$ ). There was no clear evidence to suggest a difference in effect size estimates between groups ( $p_{\text{diff}} = .12$ ).

There was strong evidence of an association between rs1051730/rs16969968 and daily cigarette consumption in both European and other groups (European:  $B = 0.97$ , 95%  $CI = 0.76, 1.18$ ,  $p < .001$ ; other:  $B = 1.58$ , 95%  $CI = 0.71, 2.46$ ,  $p < .001$ ). There was no clear evidence to suggest that this effect size differed between groups ( $p_{\text{diff}} = .18$ ).

There was strong evidence of an association between rs1051730/rs16969968 and daily cigarette consumption in both the control/population group and the disease/partial disease group (control:  $B = 0.97$ , 95%  $CI = 0.55, 1.39$ ,  $p < .001$ ; disease/partial:  $B = 1.04$ , 95%  $CI = 0.80, 1.27$ ,  $p < .001$ ). There was no evidence for a difference in effect size estimates between groups ( $p_{\text{diff}} = .79$ ).

There was strong evidence of an association between both rs1051730 and rs16969968 SNPs and daily cigarette consumption (rs1051730:  $B = 1.17$ , 95%  $CI = 0.95, 1.39$ ,  $p < .001$ ; rs16969968:  $B = 0.77$ , 95%  $CI = 0.50, 1.05$ ,  $p < .001$ ), and this effect size appeared to differ between groups ( $p = .028$ ). However, although this difference was qualitatively observed in the

**Table 1. Meta-Analysis of rs1051730/rs16969968 and Heaviness of Smoking: Full and Stratified Analyses**

	<i>k</i>	Fixed effects					Random effects					
		Effect size	95% <i>CI</i>	<i>p</i> value	<i>I</i> <sup>2</sup> (%)	<i>p</i> <sub>diff</sub>	Effect size	95% <i>CI</i>	<i>p</i> value	<i>p</i> <sub>diff</sub>		
Full model	57	0.915	0.769	1.060	<.001	34	NA	1.012	0.806	1.218	<.001	NA
HWE												
Yes	55	0.898	0.752	1.045	<.001	34		0.989	0.781	1.196	<.001	
No	2	1.894	0.757	3.032	.001	0	.089	1.894	0.757	3.032	.001	.12
Ancestry												
European	50	0.887	0.739	1.036	<.001	34		0.971	0.763	1.178	<.001	
Other	7	1.634	0.874	2.394	<.001	20	.059	1.584	0.712	2.456	<.001	.18
Disease state												
No	21	0.838	0.533	1.143	<.001	33		0.968	0.547	1.390	<.001	
Yes/partial	36	0.937	0.772	1.103	<.001	37	.57	1.035	0.798	1.273	<.001	.79
SNP												
rs1051730	44	1.144	0.964	1.323	<.001	19		1.170	0.952	1.388	<.001	
rs16969968	27	0.648	0.444	0.852	<.001	30	<.001	0.775	0.499	1.051	<.001	.028

Note. Results under an additive model of genetic action displayed. HWE = Hardy–Weinberg equilibrium; SNP = single nucleotide polymorphism.

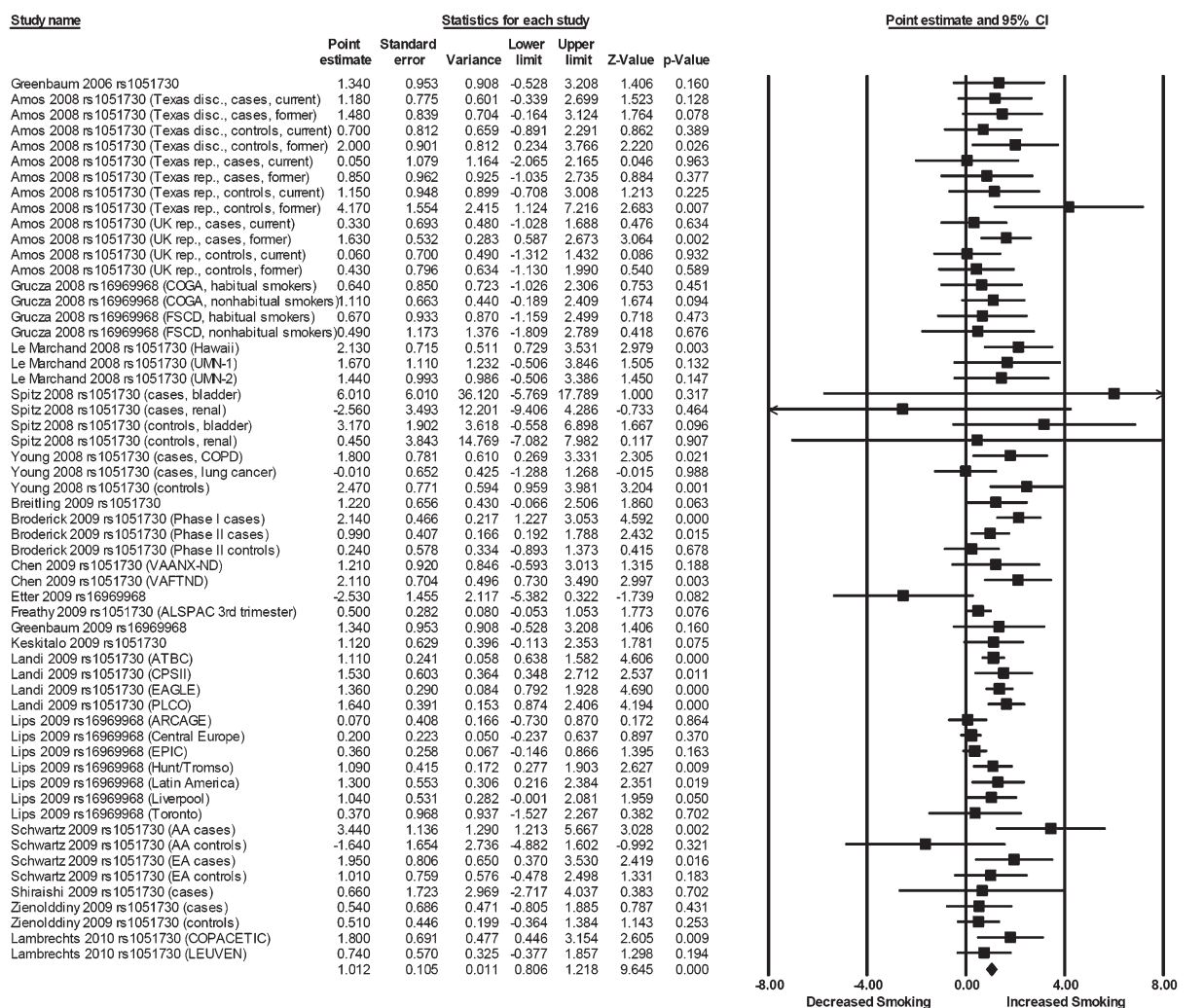


Figure 2. Meta-analysis of association of rs1051730/rs16969968 single nucleotide polymorphisms with heaviness of smoking.

subset of samples ( $k = 14$ ) that contained data on both SNPs, with a slightly larger effect size observed for rs1051730 ( $B = 1.09$ ,  $95\% CI = 0.72, 1.46$ ,  $p < .001$ ) compared with rs16969968 ( $B = 1.05$ ,  $95\% CI = 0.65, 1.46$ ,  $p < .001$ ), this difference did not achieve statistical significance ( $p_{diff} = .89$ ), suggesting that the observed difference in the full meta-analysis may be due to confounding arising from other study- or sample-level differences. One sample reporting data on both SNPs was excluded from analyses as the homozygous risk genotype group contained only one participant.

Egger's test indicated no evidence of small study bias for SNP rs1051730,  $t(42) = 0.92$ ,  $p_{one\text{-tailed}} = .18$ . Evidence of small study bias was observed for SNP rs16969968, however,  $t(25) = 2.01$ ,  $p_{one\text{-tailed}} = .028$ . We utilized Duval and Tweedie's "trim and fill" method to adjust for this, which led to a reduction in the overall effect estimate for this SNP (adjusted value:  $B = 0.49$ ,  $95\% CI = 0.19, 0.79$ ).

## Discussion

Our data suggest compelling evidence for a small effect of the rs16969968/rs1051730 SNPs on daily cigarette consumption,

equivalent to a per-allele effect of approximately 1 cigarette/day. Interestingly, SNP rs1051730 may provide a stronger signal than rs16969968, although evidence for this is indirect and should therefore be treated with caution. No evidence for a difference in effect size between groups was observed in other stratified analyses (i.e., ancestry, disease state). Strong evidence for an association between rs16969968/rs1051730 SNPs and daily cigarette consumption was observed irrespective of study level characteristics, suggesting that the association is robust.

The nicotinic acetylcholine receptor (nAChR), to which nicotine binds, is a plausible and biologically relevant candidate for smoking etiology. Neuronal nAChRs are widely distributed throughout the central and peripheral nervous system. They are ligand-gated ion channels composed of five transmembrane subunit proteins arranged around a central pore. Neuronal nAChRs consist of  $\alpha$  ( $\alpha 2$ - $\alpha 10$ ) and  $\beta$  ( $\beta 2$ - $\beta 4$ ) subunits (Gotti, Zoli, & Clementi, 2006), each of which is encoded for by a single *CHRN* gene, and may be homomeric or heteromeric in terms of subunit composition. Different combinations of subunits result in receptors differing in biological function (Bierut, 2010), and specific combinations of subunits may be found in specific locations throughout the brain. Expression of nAChRs in brain

areas implicated in drug addiction lends further credence for their study in this field. Of particular relevance to the current study, Berrettini et al. (2008) describe expression of both *CHRNA5* and *CHRNA3* in multiple brain areas relevant to addiction, including the amygdala and nucleus accumbens.

Although the missense mutation rs16969968 appears to be of functional significance, a larger effect was observed for the synonymous SNP rs1051730 in full stratified analyses. While this SNP appears to be of no functional significance, which to some extent limits the interpretation of the observed association, as a single marker, it appears superior to rs16969968 with regards to determining variation in smoking quantity. It is also noteworthy that evidence of small study bias was observed for SNP rs16969968 but not for rs1051730. Adjusting for this bias increased the difference in effect estimate between SNPs. It is possible that rs1051730 is a strong tagging SNP for functional haplotypes in this region, and it would therefore be important to focus research efforts on identifying these. It is crucial to note, however, that the difference in effect size estimates between SNPs was only qualitatively observed in the subset of samples which contained data on both rs1051730 and rs16969968. This may be due to the limited number of studies examined ( $k = 14$ ) or confounding arising from other study- or sample-level differences. A large-scale study directly comparing both SNPs would be required to answer this question definitively.

The primary limitation of our meta-analysis was that we did not have the data necessary to perform a joint SNP analysis in which the effects of one SNP were conditioned on the other. This analysis would have enabled us to comment more authoritatively on the difference in genetic signal between these two SNPs, if any, which are known to be in LD. An additional limitation of our meta-analysis was that the procedures used allowed only comparable data to be combined. As such a number of studies that would have ideally been included in our analysis had to be excluded, such as those examining extreme smoking quantity phenotypes (e.g., Stevens et al., 2008). We were also unable to include data from studies reporting only categorical smoking quantity data by genotype (e.g., Thorgeirsson et al., 2008). An additional shortcoming was that we were only able to investigate a limited number of study-level characteristics. It is of note, however, that the analysis of study-level characteristics is indirect and may lead to ecological fallacy. Any differences observed should be considered hypothesis generating to be followed up in appropriately designed primary studies. Finally, it is of note that methods employed to correct for publication bias, such as Duval and Tweedie's "trim and fill" approach as utilized here, are not widely accepted and rest on certain assumptions (see Munafo, Clark, & Flint, 2004). As such, corrected findings should be interpreted with caution.

In conclusion, our analyses confirm that two SNPs (rs16969968 and rs1051730) located in the nicotinic acetylcholine receptor gene cluster *CHRNA5-A3-B4* are robustly associated with heaviness of smoking. Interestingly, SNP rs1051730 may provide a stronger signal than rs16969968, although evidence for this is indirect. Much variability in this phenotype remains to be determined, however. Smoking is a complex behavior determined by both genetic and environmental factors. It is likely that many other loci will contribute to this phenotype as will multiple environmental factors. Further

research into gene–environment interactions as well as gene–gene interactions is also called for.

## Supplementary Material

Supplementary Tables S1 and S2 can be found online at <http://www.ntr.oxfordjournals.org>.

## Funding

This work was funded primarily by a Wellcome Trust PhD studentship to JW. MRM is a member of the UK Centre for Tobacco Control Studies, a UKCRC Public Health Research: Centre of Excellence. Funding from British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, and the National Institute for Health Research, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. This research was funded in part by the Wellcome Trust (086684).

## Declaration of Interests

*The authors have no competing interests to declare. This publication is the work of the authors and JW will serve as guarantor for the contents of this paper.*

## Acknowledgments

*The authors thank Lutz Breitling, Richard Houlston, Yufei Wang, Xiangning Chen, Jingchun Chen, Bernard Lerer, Lior Greenbaum, Richard Gruzca, Laura Bierut, Sarah Bertelsen, Maria Teresa Landi, Loic Le Marchand, Sharon Murphy, Elizabeth Thompson, Jun Yokota, Takashi Kohno, Robert Young, Shan Zienolddiny, Chris Amos, Margaret Spitz, Qiong Dong, Xifeng Wu, Maosheng Huang, Andrew Hattersley, Rachel Freathy, Kaisu Kesitalo, Ann Schwartz, and Angie Wenzlaff for kindly releasing data in a format that enabled their inclusion in the meta-analysis. We also thank James McKay, Esther Lips, and Diether Lambrechts for their advice and correspondence and Sean David and George Davey-Smith for kindly reviewing this manuscript prior to submission.*

## References

- Amos, C. I., Wu, X., Broderick, P., Gorlov, I. P., Gu, J., Eisen, T., et al. (2008). Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nature Genetics*, 40, 616–622. doi:10.1038/ng.109
- Baker, T. B., Weiss, R. B., Bolt, D., von Niederhausern, A., Fiore, M. C., Dunn, D. M., et al. (2009). Human neuronal acetylcholine receptor A5-A3-B4 haplotypes are associated with multiple nicotine dependence phenotypes. *Nicotine & Tobacco Research*, 11, 785–796. doi:10.1093/ntr/ntp064
- Berrettini, W., Yuan, X., Tozzi, F., Song, K., Francks, C., Chilcoat, H., et al. (2008). Alpha-5/alpha-3 nicotinic receptor subunit alleles increase risk for heavy smoking. *Molecular Psychiatry*, 13, 368–373. doi:10.1038/sj.mp.4002154

- Bierut, L. J. (2010). Convergence of genetic findings for nicotine dependence and smoking related diseases with chromosome 15q24-25. *Trends in Pharmacological Sciences*, 31, 46–51. doi:10.1016/j.tips.2009.10.004
- Bierut, L. J., Stitzel, J. A., Wang, J. C., Hinrichs, A. L., Gruzca, R. A., Xuei, X., et al. (2008). Variants in nicotinic receptors and risk for nicotine dependence. *American Journal of Psychiatry*, 165, 1163–1171. doi:10.1176/appi.ajp.2008.07111711
- Breitling, L. P., Dahmen, N., Mittelstraß, K., Illig, T., Rujescu, D., Raum, E., et al. (2009). Smoking cessation and variations in nicotinic acetylcholine receptor subunits  $\alpha$ -5,  $\alpha$ -3, and  $\beta$ -4 genes. *Biological Psychiatry*, 65, 691–695. doi:10.1016/j.biopsych.2008.10.004
- Broderick, P., Wang, Y., Vijaykrishnan, J., Matakidou, A., Spitz, M. R., Eisen, T., et al. (2009). Deciphering the impact of common genetic variation on lung cancer risk: A genome-wide association study. *Cancer Research*, 69, 6633–6641. doi:10.1158/0008-5472.CAN-09-0680
- Caporaso, N., Gu, F., Chatterjee, N., Sheng-Chih, J., Yu, K., Yeager, M., et al. (2009). Genome-wide and candidate gene association study of cigarette smoking behaviors. *PLoS One*, 4, Retrieved from <http://plosone.org/>
- Chen, L. S., Johnson, E. O., Breslau, N., Hatsukami, D., Saccone, N. L., Gruzca, R. A., et al. (2009). Interplay of genetic risk factors and parent monitoring in risk for nicotine dependence. *Addiction*, 104, 1731–1740. doi:10.1111/j.1360-0443.2009.02697.x
- Chen, X., Chen, J., Williamson, V. S., An, S. S., Hettema, J. M., Aggen, S. H., et al. (2009). Variants in nicotinic acetylcholine receptors  $\alpha$ 5 and  $\alpha$ 3 increase risks to nicotine dependence. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, 150, 926–933. doi:10.1002/ajmg.b.30919
- Conti, D. V., Lee, W., Li, D., Liu, J., Van den Berg, D., Thomas, P. D., et al. (2008). Nicotinic acetylcholine receptor  $\beta$ 2 subunit gene implicated in a systems-based candidate gene study of smoking cessation. *Human Molecular Genetics*, 17, 2834–2848. doi:10.1093/hmg/ddn181
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56, 455–463. doi:10.1111/j.0006-341X.2000.00455.x
- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315, 629–634. Retrieved from <http://www.bmj.com/>
- Etter, J. F., Hoda, J. C., Perroud, N., Munafo, M., Buresi, C., Duret, C., et al. (2009). Association of genes coding for the  $\alpha$ -4,  $\alpha$ -5,  $\beta$ -2 and  $\beta$ -3 subunits of nicotinic receptors with cigarette smoking and nicotine dependence. *Addictive Behaviors*, 34, 772–775. doi:10.1016/j.addbeh.2009.05.010
- Fowler, T., Lifford, K., Shelton, K., Rice, F., Thapar, A., Neale, M. C., et al. (2007). Exploring the relationship between genetic and environmental influences on initiation and progression of substance use. *Addiction*, 102, 413–422. doi:10.1111/j.1360-0443.2006.01694.x
- Freathy, R. M., Ring, S. M., Shields, B., Galobardes, B., Knight, B., Weedon, M. N., et al. (2009). A common genetic variant in the 15q24 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRNA4) is associated with a reduced ability of women to quit smoking in pregnancy. *Human Molecular Genetics*, 18, 2922–2927. doi:10.1093/hmg/ddp216
- Furberg, H., Kim, Y., Dackor, J., Boerwinkle, E., Franceschini, N., Ardissino, D., et al. (2010). Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nature Genetics*, 42, 441–447. doi:10.1038/ng.571
- Gotti, C., Zoli, M., & Clementi, F. (2006). Brain nicotinic acetylcholine receptors: Native subtypes and their relevance. *Trends in Pharmacological Sciences*, 27, 482–491. doi:10.1016/j.tips.2006.07.004
- Greenbaum, L., Kanyas, K., Karni, O., Merbl, Y., Olender, T., Horowitz, A., et al. (2006). Why do young women smoke? I. Direct and interactive effects of environment, psychological characteristics and nicotinic cholinergic receptor genes. *Molecular Psychiatry*, 11, 312–322. doi:10.1038/sj.mp.4001774
- Greenbaum, L., Rigbi, A., Teltsh, O., & Lerer, B. (2009). Role of genetic variants in the CHRNA5-CHRNA3-CHRNA4 cluster in nicotine dependence risk: Importance of gene-environment interplay. *Molecular Psychiatry*, 14, 828–830. doi:10.1038/mp.2009.25
- Gruzca, R. A., Wang, J. C., Stitzel, J. A., Hinrichs, A. L., Saccone, S. F., Saccone, N. L., et al. (2008). A risk allele for nicotine dependence in CHRNA5 is a protective allele for cocaine dependence. *Biological Psychiatry*, 64, 922–929. doi:10.1016/j.biopsych.2008.04.018
- Hung, R. J., McKay, J. D., Gaborieau, V., Boffetta, P., Hashibe, M., Zaridze, D., et al. (2008). A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature*, 452, 633–637. doi:10.1038/nature06885
- Kendler, K. S., Neale, M. C., Sullivan, P., Corey, L. A., Gardner, C. O., & Prescott, C. A. (1999). A population-based twin study in women of smoking initiation and nicotine dependence. *Psychological Medicine*, 29, 299–308. doi:10.1017/S0033291798008022
- Keskitalo, K., Broms, U., Heliövaara, M., Ripatti, S., Surakka, I., Perola, M., et al. (2009). Association of serum cotinine level with a cluster of three nicotinic acetylcholine receptor genes (CHRNA3/CHRNA5/CHRNA4) on chromosome 15. *Human Molecular Genetics*, 18, 4007–4012. doi:10.1093/hmg/ddp322
- Lambrechts, D., Buyschaert, I., Zanen, P., Coolen, J., Lays, N., Cuppens, H., et al. (2010). The 15q24/25 susceptibility variant for lung cancer and chronic obstructive pulmonary disease is associated with emphysema. *American Journal of Respiratory and Critical Care Medicine*, 181, 486–493. doi:10.1164/rccm.200909-1364OC
- Landi, M. T., Chatterjee, N., Yu, K., Goldin, L. R., Goldstein, A. M., Rotunno, M., et al. (2009). A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated

- with risk for adenocarcinoma. *American Journal of Human Genetics*, 85, 679–691. doi:10.1016/j.ajhg.2009.09.012
- Le Marchand, L., Derby, K. S., Murphy, S. E., Hecht, S. S., Hatsukami, D., Carmella, S. G., et al. (2008). Smokers with the CHRNA lung cancer-associated variants are exposed to higher levels of nicotine equivalents and a carcinogenic tobacco-specific nitrosamine. *Cancer Research*, 68, 9137–9140. doi:10.1158/0008-5472.CAN-08-2271
- Lessov, C. N., Martin, N. G., Statham, D. J., Todorov, A. A., Slutske, W. S., Bucholz, K. K., et al. (2004). Defining nicotine dependence for genetic research: Evidence from Australian twins. *Psychological Medicine*, 34, 865–879. doi:10.1017/S0033291703001582
- Li, M. D., Cheng, R., Ma, J. Z., & Swan, G. E. (2003). A meta-analysis of estimated genetic and environmental effects on smoking behavior in male and female adult twins. *Addiction*, 98, 23–31. doi:10.1046/j.1360-0443.2003.00295.x
- Lips, E. H., Gaborieau, V., McKay, J. D., Chabrier, A., Hung, R. J., Boffetta, P., et al. (2009). Association between a 15q25 gene variant, smoking quantity and tobacco-related cancers among 17 000 individuals. *International Journal of Epidemiology*, 39, 563–577. doi:10.1093/ije/dyp288
- Liu, J. Z., Tozzi, F., Waterworth, D. M., Pillai, S. G., Muglia, P., Middleton, L., et al. (2010). Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nature Genetics*, 42, 436–440. doi:10.1038/ng.572
- Liu, P., Vikis, H. G., Wang, D., Lu, Y., Wang, Y., Schwartz, A. G., et al. (2008). Familial aggregation of common sequence variants on 15q24-25.1 in lung cancer. *Journal of the National Cancer Institute*, 100, 1326–1330. doi:10.1093/jnci/djn268
- Mathers, C. D., & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Medicine*, 3, 2011–2030. doi:10.1371/journal.pmed.0030442
- McKay, J. D., Hung, R. J., Gaborieau, V., Boffetta, P., Chabrier, A., Byrnes, G., et al. (2008). Lung cancer susceptibility locus at 5p15.33. *Nature Genetics*, 40, 1404–1406. doi:10.1038/ng.254
- Munafo, M. R., Clark, T. G., & Flint, J. (2004). Assessing publication bias in genetic association studies: Evidence from a recent meta-analysis. *Psychiatry Research*, 129, 39–44. doi:10.1016/j.psychres.2004.06.011
- Munafo, M. R., & Johnstone, E. C. (2008). Genes and cigarette smoking. *Addiction*, 103, 893–904. doi:10.1111/j.1360-0443.2007.02071.x
- Pillai, S. G., Ge, D., Zhu, G., Kong, X., Shianna, K. V., Need, A. C., et al. (2009). A genome-wide association study in chronic obstructive pulmonary disease (COPD): Identification of two major susceptibility loci. *PLoS Genetics*, 5, e1000421. doi:10.1371/journal.pgen.1000421
- Ray, R., Mitra, N., Baldwin, D., Guo, M., Patterson, F., Heitjan, D. F., et al. (2010). Convergent evidence that choline acetyltransferase gene variation is associated with prospective smoking cessation and nicotine dependence. *Neuropsychopharmacology*, 35, 1374–1382. doi:10.1038/npp.2010.7
- Rigbi, A., Kanyas, K., Yakir, A., Greenbaum, L., Pollak, Y., Ben-Asher, E., et al. (2008). Why do young women smoke? V. Role of direct and interactive effects of nicotinic cholinergic receptor gene variation on neurocognitive function. *Genes, Brain and Behavior*, 7, 164–172. doi:10.1111/j.1601-183X.2007.00329.x
- Saccone, N. L., Culverhouse, R. C., Schwantes-An, T. H., Cannon, D. S., Chen, X., Cichon, S., et al. (2010). Multiple independent loci at chromosome 15q25.1 affect smoking quantity: A meta-analysis and comparison with lung cancer and COPD. *PLoS Genetics*, 6, e1001053. doi:10.1371/journal.pgen.1001053
- Saccone, N. L., Wang, J. C., Breslau, N., Johnson, E. O., Hatsukami, D., Saccone, S. F., et al. (2009). The CHRNA5-CHRNA3-CHRNA4 nicotinic receptor subunit gene cluster affects risk for nicotine dependence in African-Americans and in European-Americans. *Cancer Research*, 69, 6848–6856. doi:10.1158/0008-5472.CAN-09-0786
- Saccone, S. F., Hinrichs, A. L., Saccone, N. L., Chase, G. A., Konvicka, K., Madden, P. A. F., et al. (2007). Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Human Molecular Genetics*, 16, 36–49. doi:10.1093/hmg/ddl438
- Schwartz, A. G., Cote, M. L., Wenzlaff, A. S., Land, S., & Amos, C. I. (2009). Racial differences in the association between SNPs on 15q25.1, smoking behavior, and risk of non-small cell lung cancer. *Journal of Thoracic Oncology*, 4, 1195–1201. doi:10.1097/JTO.0b013e3181b244ef
- Sherva, R., Wilhelmsen, K., Pomerleau, C. S., Chasse, S. A., Rice, J. P., Snedecor, S. M., et al. (2008). Association of a single nucleotide polymorphism in neuronal acetylcholine receptor subunit alpha 5 (CHRNA5) with smoking status and with 'pleasurable buzz' during early experimentation with smoking. *Addiction*, 103, 1544–1552. doi:10.1111/j.1360-0443.2008.02279.x
- Shiraishi, K., Kohno, T., Kunitoh, H., Watanabe, S. I., Goto, K., Nishiwaki, Y., et al. (2009). Contribution of nicotine acetylcholine receptor polymorphisms to lung cancer risk in a smoking-independent manner in the Japanese. *Carcinogenesis*, 30, 65–70. doi:10.1093/carcin/bgn257
- Spitz, M. R., Amos, C. I., Dong, Q., Lin, J., & Wu, X. (2008). The CHRNA5-A3 region on chromosome 15q24-25.1 is a risk factor both for nicotine dependence and for lung cancer. *Journal of the National Cancer Institute*, 100, 1552–1556. doi:10.1093/jnci/djn363
- Stevens, V. L., Bierut, L. J., Talbot, J. T., Wang, J. C., Sun, J., Hinrichs, A. L., et al. (2008). Nicotinic receptor gene variants influence susceptibility to heavy smoking. *Cancer Epidemiology Biomarkers and Prevention*, 17, 3517–3525. doi:10.1158/1055-9965.EPI-08-0585
- Sullivan, P. F., & Kendler, K. S. (1999). The genetic epidemiology of smoking. *Nicotine & Tobacco Research*, 1(Suppl. 2), S51–S57. doi:10.1080/14622299050011811



- Thorgeirsson, T. E., Geller, F., Sulem, P., Rafnar, T., Wiste, A., Magnusson, K. P., et al. (2008). A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature*, *452*, 638–642. doi:10.1038/nature06846
- Thorgeirsson, T. E., Gudbjartsson, D. F., Surakka, I., Vink, J. M., Amin, N., Geller, F., et al. (2010). Sequence variants at CHRN3-CHRNA6 and CYP2A6 affect smoking behavior. *Nature Genetics*, *42*, 448–453. doi:10.1038/ng.573
- Thorgeirsson, T. E., & Stefansson, K. (2010). Commentary: Gene-environment interactions and smoking-related cancers. *International Journal of Epidemiology*, *39*, 577–579. doi:10.1093/ije/dyp385
- Weiss, R. B., Baker, T. B., Cannon, D. S., Von Niederhausern, A., Dunn, D. M., Matsunami, N., et al. (2008). A candidate gene approach identifies the CHRNA5-A3-B4 region as a risk factor for age-dependent nicotine addiction. *PLoS Genetics*, *4*, e1000125. doi:10.1371/journal.pgen.1000125
- World Health Organization. (2008). *WHO Report on the Global Tobacco Epidemic, 2008: The MPOWER package*. Geneva, Switzerland: Author. Retrieved from <http://www.who.int/en/>
- Wu, C., Hu, Z., Yu, D., Huang, L., Jin, G., Liang, J., et al. (2009). Genetic variants on chromosome 15q25 associated with lung cancer risk in Chinese populations. *Cancer Research*, *69*, 5065–5072. doi:10.1158/0008-5472.CAN-09-0081
- Yang, P., Li, Y., Jiang, R., Cunningham, J. M., Zhang, F., & De Andrade, M. (2010). A rigorous and comprehensive validation: Common genetic variations and lung cancer. *Cancer Epidemiology Biomarkers and Prevention*, *19*, 240–244. doi:10.1158/1055-9965.EPI-09-0710
- Young, R. P., Hopkins, R. J., Hay, B. A., Epton, M. J., Black, P. N., & Gamble, G. D. (2008). Lung cancer gene associated with COPD: Triple whammy or possible confounding effect? *European Respiratory Journal*, *32*, 1158–1164. doi:10.1183/09031936.00093908
- Young, R. P., Hopkins, R. J., Hay, B. A., Epton, M. J., Mills, G. D., Black, P. N., et al. (2009). A gene-based risk score for lung cancer susceptibility in smokers and ex-smokers. *Postgraduate Medical Journal*, *85*, 515–524. doi:10.1136/pgmj.2008.077107
- Zienolddiny, S., Skaug, V., Landvik, N. E., Ryberg, D., Phillips, D. H., Houlston, R., et al. (2009). The TERT-CLPTM1L lung cancer susceptibility variant associates with higher DNA adduct formation in the lung. *Carcinogenesis*, *30*, 1368–1371. doi:10.1093/carcin/bgp131