

SCIENTIFIC REPORTS



OPEN

Analyzing large-scale samples confirms the association between the rs1051730 polymorphism and lung cancer susceptibility

Zhijie Han¹, Qinghua Jiang¹, Tianjiao Zhang², Xiaoliang Wu², Rui Ma², Jixuan Wang³, Yang Bai², Rongjie Wang², Renjie Tan² & Yadong Wang²

Received: 07 July 2015

Accepted: 24 September 2015

Published: 28 October 2015

The early genome-wide association studies (GWAS) found a significant association between lung cancer and rs1051730 (15q25) polymorphism. However, the subsequent studies reported consistent and inconsistent results in different populations. Three meta-analysis studies were thus performed to reevaluate the association. But their results remain inconsistent. After that, some new GWAS studies reported conflicting results again. We think that the divergence of these results may be due to small-scale samples or heterogeneity among different populations. Therefore, we reevaluated the association by collecting more samples (N = 33,617 cases and 116,639 controls) from 31 studies, which incorporate 8 new studies and 23 previous studies used by one or more of the three meta-analysis studies. We observed a significant association between lung cancer and rs1051730 in pooled population by using allele (OR = 1.30, 95% CI = 1.27–1.34, $P < 0.0001$), dominant (OR = 1.41, 95% CI = 1.29–1.55, $P < 0.0001$), recessive (OR = 1.53, 95% CI = 1.42–1.65, $P < 0.0001$) and additive (OR = 1.75, 95% CI = 1.61–1.90, $P < 0.0001$) models. Through the subgroup analysis, we observed a significant heterogeneity only in East Asian population ($P = 0.006$, $I^2 = 66.9\%$), and the association is significant in all subgroups (OR = 1.2976, 95% CI = 1.2622–1.3339 (European ancestry), OR = 1.5025, 95% CI = 1.2465–1.8110 (African), OR = 1.7818, 95% CI = 1.3915–2.2815 (East Asian), $P < 0.0001$). We believe that these results will contribute to understanding the genetic mechanism of lung cancer.

Lung cancer is one of the most dangerous diseases, and the morbidity and mortality are increasing constantly around the world today¹. According to the statistics from American Cancer Society (ACS), the new cases of lung cancer were 0.221 million and dead 0.158 million in the United States in 2015, which was one of the most frequently diagnosed cancers and the leading causes of cancer death in men and women². Even though, the majority of studies proved that over 80 percent lung cancer occurrence is associated with the habit of smoking, while only 20 percent of these smokers will develop into lung cancer, which suggests that lung cancer occurrence possesses genetic susceptibility^{1,3}.

Many single nucleotide polymorphism (SNP) sites in association with lung cancer were obtained by the genome-wide association studies (GWAS)⁴. For example, Amos *et al.*⁵, Thorgeirsson, *et al.*⁶ and Shiraishi *et al.*⁷ found that the rs1051730 SNP was associated with lung cancer in European ancestry and Japanese population, respectively. Amos *et al.* analyzed 956 cases and 1,830 controls from the UK, and discovered the odds ratio (OR) = 1.32, 95% confidence interval (CI) is from 1.23 to 1.39, and $P < 10E-08$. Thorgeirsson, *et al.* analyzed 665 cases and 28,752 controls from the Iceland, 269 cases and 1,474 controls from the Spain, 90 cases and 2,018 controls from the Netherlands, and the combined result showed that

¹School of Life Science and Technology, Harbin Institute of Technology, Harbin, 150001, China. ²School of Computer Science and Technology, Harbin Institute of Technology, Harbin, 150001, China. ³School of Software, Harbin Institute of Technology, Harbin, 150001, China. Correspondence and requests for materials should be addressed to Q.J. (email: qhjiang@hit.edu.cn) or Y.W. (email: ydwang@hit.edu.cn)

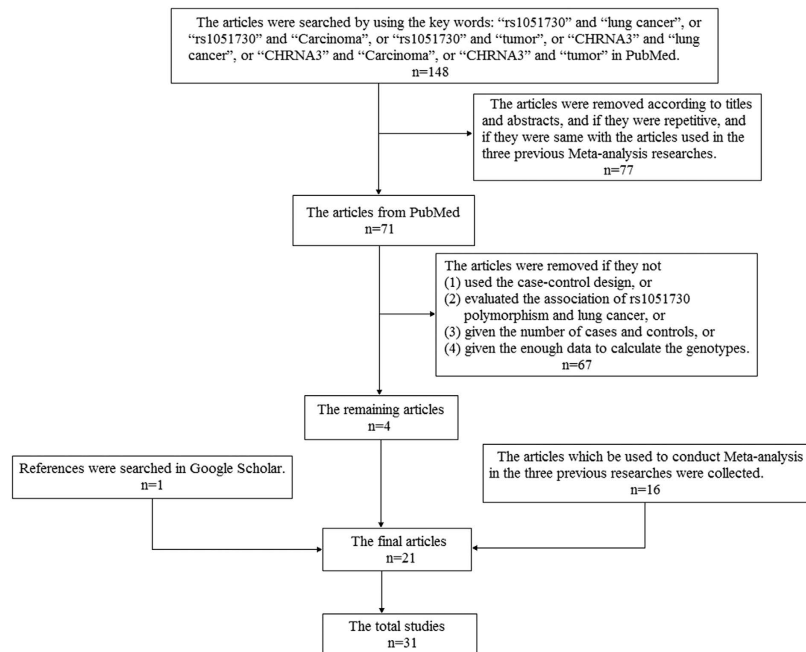


Figure 1. The flow chart of collecting articles for analyzing the association between rs1051730 polymorphism and lung cancer. The criteria of selecting articles are: (1) the study was designed according to the method of case-control. (2) the study evaluated the association of rs1051730 polymorphism and lung cancer. (3) the number of cases and controls was provided in the study. (4) the study provided the population of each individual. (5) the study provided the number of rs1051730 genotypes both in cases and controls or provided enough data to calculate the genotypes. (6) the study provided the OR value with 95% CI and the P value or provided enough data to calculate them.

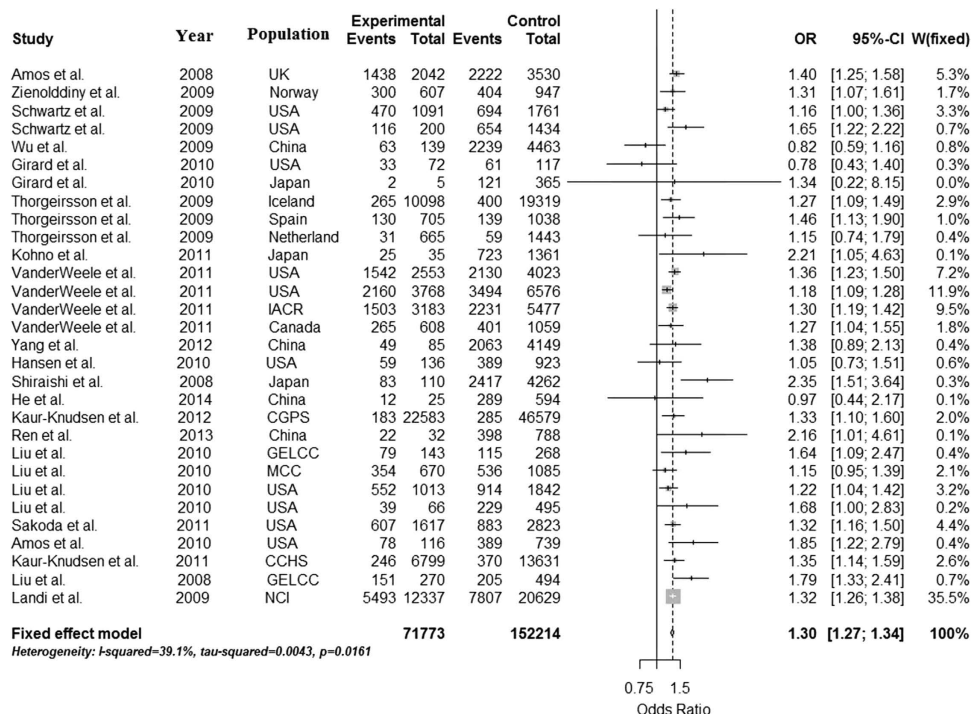


Figure 2. Forest plot for the meta-analysis of rs1051730 polymorphism using allele model. According to Table 1, 30 studies that provided the data of genotype A/G and AA/GA/GG and GA/GG were used in the allele model (A vs G).

Study	Year	Country or Institution	Ethnicity	No. of cases	No. of controls	Genotyping platform	Kind of genotype
Amos <i>et al.</i> ⁵	2008	UK	European ancestry	956	1830	PCR	AA/GA/GG
Zienolddiny <i>et al.</i> ⁸	2009	Norway	European ancestry	352	425	TaqMan	AA/GA/GG
Schwartz <i>et al.</i> ⁹	2009	USA	European ancestry	582	833	TaqMan	AA/GA/GG
Schwartz <i>et al.</i> ⁹	2009	USA	African	385	432	TaqMan	AA/GA/GG
Wu <i>et al.</i> ¹⁹	2009	China	East Asian	1151	1150	PCR	AA/GA/GG
Girard <i>et al.</i> ²⁰	2010	USA	European ancestry	94	95	Illumina 610	A/G
Girard <i>et al.</i> ²⁰	2010	Japan	East Asian	123	247	Illumina 610	A/G
Thorgeirsson <i>et al.</i> ⁶	2008	Iceland	European ancestry	665	28752	Illumina 300	A/G
Thorgeirsson <i>et al.</i> ⁶	2008	Spain	European ancestry	269	1474	Illumina 300	A/G
Thorgeirsson <i>et al.</i> ⁶	2008	Netherland	European ancestry	90	2018	Illumina 300	A/G
Kohno <i>et al.</i> ¹⁰	2011	Japan	East Asian	374	324	PCR	GA/GG
VanderWeele <i>et al.</i> ¹¹	2012	USA	European ancestry	1836	1452	Illumina 300 and TaqMan	AA/GA/GG
VanderWeele <i>et al.</i> ¹¹	2012	USA	European ancestry	2827	2345	Illumina 300	AA/GA/GG
VanderWeele <i>et al.</i> ¹¹	2012	IACR	European ancestry	1867	2463	Illumina 300	AA/GA/GG
VanderWeele <i>et al.</i> ¹¹	2012	Canada	European ancestry	333	501	Illumina 300	AA/GA/GG
Yang <i>et al.</i> ¹²	2012	China	East Asian	1056	1061	PCR	GA/GG
Hansen <i>et al.</i> ²⁸	2010	USA	African	448	611	Illumina Golden Gate	A/G
Shiraishi <i>et al.</i> ⁷	2009	Japan	East Asian	1250	936	TaqMan and the ABI Prism 7900HT	AA/GA/GG
He <i>et al.</i> ²⁷	2014	China	East Asian	301	318	Sequenom MassARRAY and PCR	A/G
Kaur-Knudsen <i>et al.</i> ²⁴	2012	CGPS	European ancestry	234	34347	TaqMan	AA/GA/GG
Ren <i>et al.</i> ¹³	2013	China	East Asian	210	200	TaqMan	GA/GG
Liu <i>et al.</i> ²⁵	2010	GELCC	European ancestry	194	217	Illumina 550	A/G
Liu <i>et al.</i> ²⁵	2010	MCC	European ancestry	890	865	Illumina 550	A/G
Liu <i>et al.</i> ²⁵	2010	USA	European ancestry	1466	1389	Illumina 550	A/G
Liu <i>et al.</i> ²⁵	2010	USA	African	268	293	Illumina 550	A/G
Sakoda <i>et al.</i> ¹⁴	2011	USA	European ancestry	745	1475	Illumina Golden Gate	AA/GA/GG
Amos <i>et al.</i> ¹⁵	2010	USA	African	467	388	TaqMan	A/G
Kaur-Knudsen <i>et al.</i> ²⁶	2011	CCHS	European ancestry	308	9907	TaqMan	AA/GA/GG
Wassenaar <i>et al.</i> ¹⁶	2013	USA	European ancestry	398	421	PCR	AA/GA and GG
Liu <i>et al.</i> ¹⁷	2008	GELCC	European ancestry	178	204	SNP Array	AA/GA/GG
Landi <i>et al.</i> ¹⁸	2009	NCI	European ancestry	13300	19666	SNP Array	A/G
All				33617	116639		

Table 1. Main information of these studies used to analyze the association between rs1051730 polymorphism and lung cancer. “AA/GA/GG” means the study have offered the data of genotypes AA/GA/GG both in cases and controls. “AA/GA and GG” means only the data of genotypes AA/GA and GG both in cases and controls have been offered in the study. “GA/GG” means only the data of genotypes GA/GG both in cases and controls have been offered in the study. “A/G” means only the data of genotypes A/G both in cases and controls have been offered in the study. All the alleles have been used A or G to express in our study to replace T or C in original articles. IARC: International Agency for Research on Cancer; CGQS: Copenhagen General Population Study; GELCC: Caucasians of Genetic Epidemiology of Lung Cancer Consortium; MCC: Mayo Clinic Caucasians; NCI: National Cancer Institute.

OR = 1.31, 95% CI is from 1.19 to 1.44, and P = 1.5E-08. Shiraishi *et al.* analyzed 1,250 cases and 936 controls from the Japan, and found the OR = 2.3, 95% CI is from 1.5 to 3.7, and P = 0.0028.

However, the subsequent studies reported consistent⁸⁻¹⁸ and inconsistent^{19,20} results. For example, Wu *et al.*¹⁹ used 1,151 cases and 1,150 controls from China, and Girard *et al.*²⁰ used 94 cases and 95 controls from America to investigate the association between rs1051730 and lung cancer, and obtained negligible or weak association.

Facing the inconsistent results above, Gu *et al.*²¹, Hu *et al.*²² and Zhan *et al.*²³ collected different samples to re-assess the association between rs1051730 and lung cancer by meta-analysis, respectively.

Subgroup	Heterogeneity test			Meta-analysis		
	I ²	P value	OR	95% CI	Z value	P value
All studies	39.10%	0.0161	1.303	[1.2682; 1.3388]	19.1601	<0.0001
European ancestry	18.40%	0.2296	1.2976	[1.2622; 1.3339]	18.473	<0.0001
African	42.20%	0.1586	1.5025	[1.2465; 1.8110]	4.2723	<0.0001
East Asian	66.90%	0.006	1.4759	[0.9957; 2.1878]	1.9383	0.0526
Japanese	0%	0.8387	2.2654	[1.5675; 3.2741]	4.3523	<0.0001
Chinese	56.50%	0.0753	1.1743	[0.7818; 1.7637]	0.774	0.4389
Removing the study of Wu <i>et al.</i> from Chinese	1.40%	0.3626	1.4235	[1.0146; 1.9972]	2.0437	0.041
Removing the study of Wu <i>et al.</i> from East Asian	12.4%	0.3356	1.7818	[1.3915; 2.2815]	4.5789	<0.0001

Table 2. The result of heterogeneity test and meta-analysis in subgroup analysis.

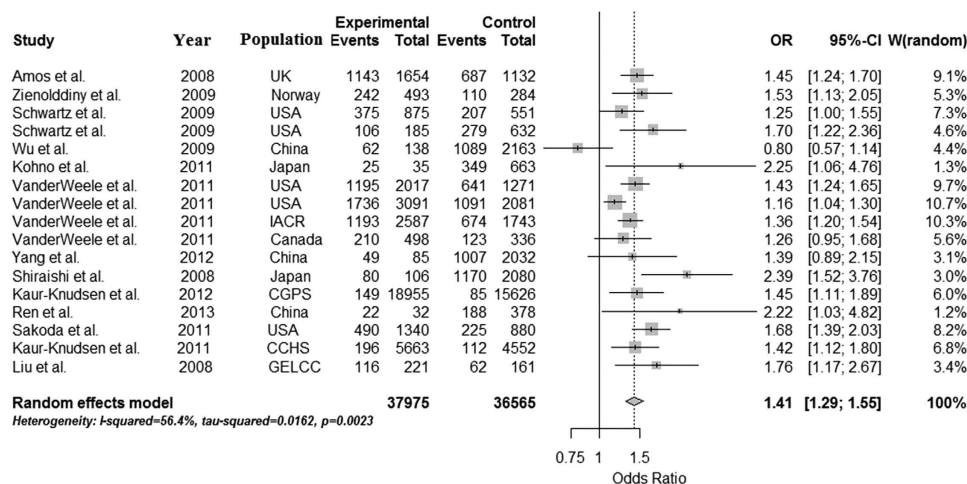


Figure 3. Forest plot for the meta-analysis of rs1051730 polymorphism using dominant model.

According to Table 1, 17 studies that provided the data of genotype AA/GA/GG and GA/GG were used in the dominant model (AA+AG vs GG).

However, the results of the three meta-analysis studies are also inconsistent. Gu *et al.*²¹ and Hu *et al.*²² found that the association between rs1051730 and lung cancer is not significant, while it is significant according to the study of Zhan *et al.*²³.

After these studies, five articles^{24–28}, which include eight studies, investigated the association between rs1051730 and lung cancer. But the results of eight studies remain inconsistent. For example, He *et al.*²⁷ used 301 cases and 318 controls from China, Hansen *et al.*²⁸ used 448 cases and 611 controls from America, and found a negligible association between rs1051730 and lung cancer, while other researchers^{24–26} reported opposite consequences.

We considered that the divergence of these results may be due to small-scale samples or heterogeneity among different populations. Here, we collected all the samples of the three meta-analysis studies and the eight studies from the five new articles, and thus obtained a larger sample sizes (33,617 cases and 116,639 controls) from 31 studies to reevaluate the association between rs1051730 polymorphism and lung cancer based on the method that was frequently used by Liu *et al.* to study Alzheimer's Disease^{29–32} and colorectal cancer³³.

Results

Literature and Study acquisition as well as Data extraction. By searching PubMed with keywords (details shown in the method section), we obtained 20 articles, which include 16 articles used in the three previous meta-analysis researches and 4 new articles. Moreover, we obtained another article through the reference search in Google Scholar. Finally, we got 31 studies from the 21 articles according to the six inclusion criteria (details shown in the method section). The workflow was showed in Fig. 1.

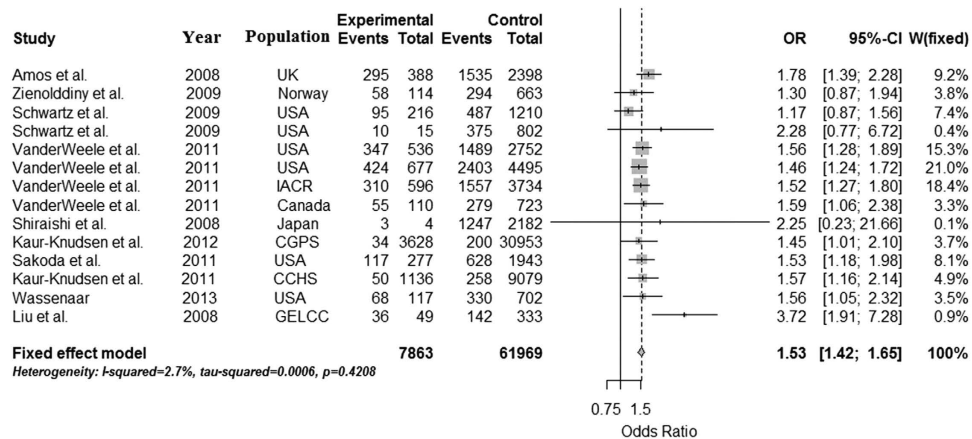


Figure 4. Forest plot for the meta-analysis of rs1051730 polymorphism using recessive model. According to Table 1, 14 studies that provided the data of genotype AA/GA/GG as well as AA/GA and GG were used in the recessive model (AA vs AG+GG).

After that, we collected the relative data according to 11 terms for each study, and 7 of the 11 terms were listed in Table 1.

Heterogeneity Test. According to the kind of genotype shown in Table 1, we used 30 studies in the allele model, 17 studies in the dominant model, 14 studies in the recessive model, and 13 studies in the additive model, respectively. We didn't get a significant heterogeneity in the allele model ($P=0.0161$ and $I^2=39.1\%$), recessive model ($P=0.4208$ and $I^2=2.7\%$), and additive model ($P=0.1791$ and $I^2=26.3\%$). However, it is significant in the dominant model ($P=0.0023$ and $I^2=56.4\%$).

Meta-analysis. Because the heterogeneity is significant in the dominant model, we used the random effect model to analyze it. And because the heterogeneity isn't significant in the allele, recessive and additive models, we used the fix effect model to analyze them, and get a relatively strong association between the rs1051730 and lung cancer in the four models (OR=1.30 and 95% CI=1.27–1.34 in the allele model, OR=1.41 and 95% CI=1.29–1.55 in the dominant model, OR=1.53 and 95% CI=1.42–1.65 in the recessive model, OR=1.75 and 95% CI=1.61–1.90 in additive model, respectively). The P values of Z test are less than 0.0001 in all models, which proved that these OR values are believable. These conclusions were described in Figs 2–5.

Subgroup Analysis. Because the allele model included the maximum number of studies, we further used it to perform the subgroup analysis. We found that there is no significant heterogeneity in European ancestry ($P=0.2296$ and $I^2=18.4\%$) and African ($P=0.1586$ and $I^2=42.2\%$) populations, while in East Asian population the heterogeneity is significant ($P=0.006$ and $I^2=66.9\%$). So we further split East Asian population into Japanese and Chinese subgroups. The heterogeneity was not found in the Japanese population ($P=0.8387$ and $I^2=0$), but it remains significant in the Chinese population ($P=0.0753$ and $I^2=56.5\%$). Then, we removed each study from Chinese population orderly, and found that there was no significant heterogeneity after the study of Wu *et al.* had been removed. After meta-analysis and Z test, we found the association between the rs1051730 and lung cancer is relatively strong in all populations. The results were described in Table 2. Forest plots of each subgroup meta-analysis were shown in Figure S1–S7.

Publication Bias Analysis and Sensitivity Analysis. We used Begg's test and Egger's test to measure the publication bias for the allele, dominant, recessive and additive models, and didn't find a significant publication bias in all these models. The P values of Begg's test and Egger's test for the allele, dominant, recessive and additive models are 0.4077 and 0.3897, 0.1061 and 0.0869, 0.1838 and 0.2044, 0.1314 and 0.1619, respectively. The funnel plot (Fig. 6) reflected the result directly. And then, the result of sensitivity analysis showed that the association between rs1051730 and lung cancer doesn't extremely change when we removed each of the studies in the four models orderly. The detail information was in supplement materials (Table S5–S8).

Discussion

Nicotine receptor protein abnormal expression is one of the reasons for lung cancer occurrence³. CHRNA3, a gene coding a part of nicotinic acetylcholine receptor protein subunits, includes a SNP rs1051730. The elder GWAS studies showed that a significant association between the rs1051730 polymorphism and

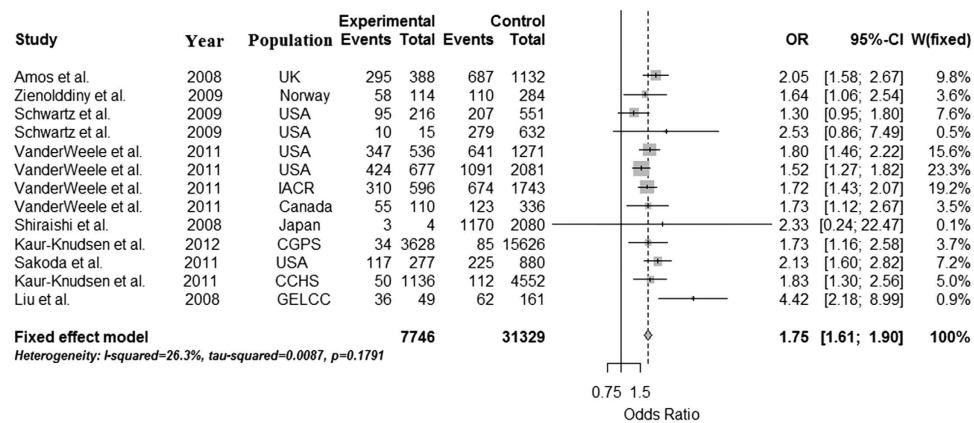


Figure 5. Forest plot for the meta-analysis of rs1051730 polymorphism using additive model. According to Table 1, 13 studies that provided the data of genotype AA/GA/GG were used in the additive model (AA vs GG).

lung cancer in the European ancestry and Japanese population⁵⁻⁷. However, the subsequent studies that researched the association in various populations or in different scales reported consistent and inconsistent results. We think the divergence of these results might be due to the small-scale samples or heterogeneity among different populations. So we summarized these studies in a larger scale and in a more comprehensive population and found that the association between rs1051730 polymorphism and lung cancer is significant by using the allelic model (OR = 1.30, 95% CI = 1.27–1.34, $P < 0.0001$), dominant model (OR = 1.41, 95% CI = 1.29–1.55, $P < 0.0001$), recessive model (OR = 1.53, 95% CI = 1.42–1.65, $P < 0.0001$) and additive model (OR = 1.75, 95% CI = 1.61–1.90, $P < 0.0001$). The farther subgroup analysis showed a significant association in European ancestry (OR = 1.2976, 95% CI = 1.2622–1.3339, $P < 0.0001$) and African (OR = 1.5025, 95% CI = 1.2465–1.8110, $P < 0.0001$) populations. But in East Asian population, we found a weak association with a non-significant result of Z test (OR = 1.4759, CI = 0.9957–2.1878, $P = 0.0526$), and the heterogeneity is significant ($P = 0.006$ and $I^2 = 66.9%$). So we further split East Asian population into Japanese and Chinese subgroups. The result indicated that there is a significant association in the Japanese population (OR = 2.2654, 95% CI = 1.5675–3.2741, $P < 0.0001$), while the result of Z test is still non-significant in the Chinese population ($P = 0.4389$). We removed each study from Chinese population orderly, and found a significant association (OR = 1.4235, 95% CI = 1.0146–1.9972, $P = 0.041$) after the study of Wu *et al.* had been removed. We think this phenomenon may be caused by a relatively small sample size in the study of Wu *et al.* Moreover, we removed the study of Wu *et al.* from East Asian population, and result of meta-analysis showed that the association is significant in East Asian population (OR = 1.7818, 95% CI = 1.3915–2.2815, $P < 0.0001$). Consequently, the association between rs1051730 polymorphism and lung cancer is significant in all of these populations. In addition, the result of sensitivity analysis reflects the conclusion is robust, and the publication bias isn't significant.

Before submitting this paper, we used keyword “rs1051730”, “lung cancer” and “meta” to search in PubMed, and obtained six articles^{21-23,34-36}, which include the researches of Gu *et al.*²¹, Hu *et al.*²² and Zhan *et al.*²³. Among these three articles, Gu *et al.*²¹ integrated 16 studies to assess the risk of rs1051730 in East Asian, European, and African populations by using allelic and dominant models. Hu *et al.*²² also collected 16 studies in the same populations, but merely the allelic model was used. They obtained a similar result that the risk is high in European and African populations, but it is weak in East Asian. Zhan *et al.*²³ only assessed the East Asian, they combined 4 studies and the result shows a significant association between the rs1051730 and lung cancer. The other three articles didn't research the association between the rs1051730 polymorphism and lung cancer. They evaluated the association between rs1051730 and habit of smoking³⁴⁻³⁶, or the association between rs1051730 and cotinine levels³⁶ by meta-analysis.

Our work is different from the others. We collected all 31 studies, which include 23 studies used by the three previous meta-analysis researches and 8 new studies. We analyzed the association between rs1051730 and lung cancer in European ancestry, African and East Asian populations. The association is significant in European ancestry and African populations, which is consistent with the result of Gu *et al.*²¹ and Hu *et al.*²². However, we found that the association remains significant in East Asian population, which is not consistent with the result of Gu *et al.*²¹ and Hu *et al.*²².

The rs1051730 polymorphism is in CHRNA3 on 15q25. CHRNA3 is one of members in a multi-gene family of nicotinic acetylcholine receptor (nAChR gene cluster) which can code various nicotinic acetylcholine receptor protein subunits include: $\alpha 3$, $\alpha 4$, $\alpha 7$, $\alpha 9$, $\alpha 10$, $\beta 2$ and $\beta 4$ nAChR subunits³⁷. These subunits are expressed on the bronchial epithelial cells of human being and primates³⁸. Through combining with the nicotinic acetylcholine receptor protein, nicotine promotes tumor cell proliferation, invasion, migration and induces blood vessel formation. At the same time, it provides a protection for the

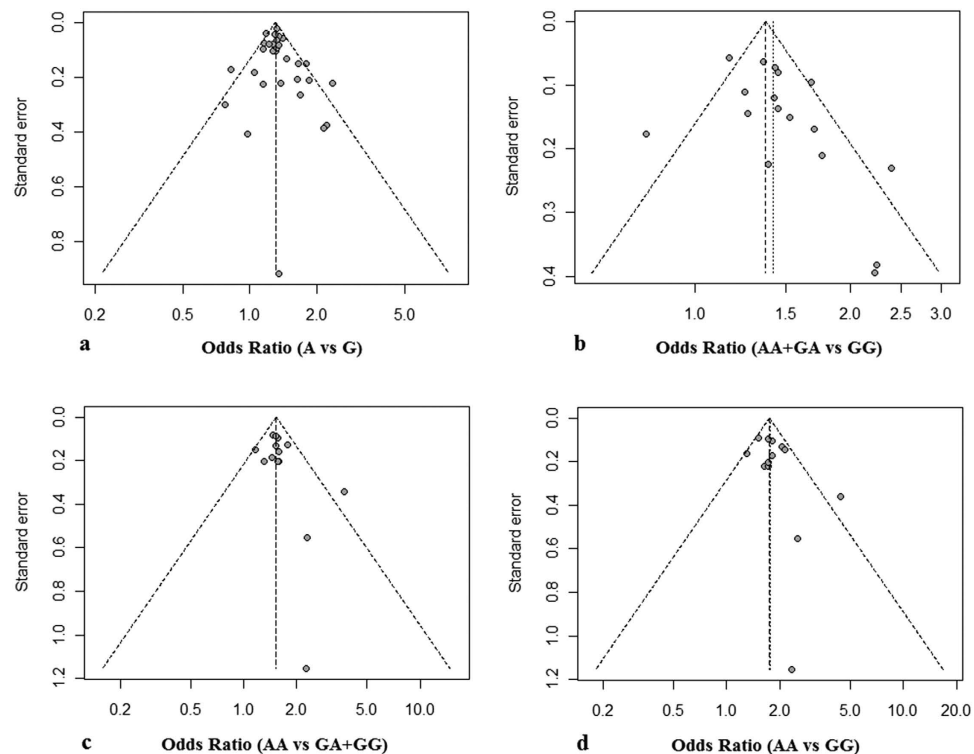


Figure 6. Funnel plot of publication bias analysis in the allele, dominant, recessive and additive model. This funnel plot shows the result of the publication bias analysis for rs1051730 polymorphism with lung cancer using allele (a), dominant (b), recessive (c) and additive (d) model. The X-axis represents OR value and the Y-axis represents standard error of each study, respectively.

tumor cell to avoid the programmed cell death³⁹. In addition, Arredondo *et al.*⁴⁰ found that $\alpha 3$, $\alpha 4$, $\alpha 7$, $\alpha 9$, $\alpha 10$, $\beta 2$ and $\beta 4$ nAChR subunits can form the high-affinity sites to bind 4-(methylnitro-samino)-1-(3-pyridyl)-1-butanone (NNK), a carcinogen produced through nicotine nitrosylation, thus to increase the risk of lung cancer. Moreover, CHRNA3 also contains other two SNP: rs578776 and rs938682 polymorphism. The rs17486278, rs11637635 as well as rs7178270 polymorphism belong to CHRNA5 and CHRNA4, respectively. They may be also associated with the susceptibility of non-small cell lung cancer was reported in a study⁴¹. In addition, another gene AGPHD1 also in 15q25 contains the rs8034191 polymorphism, and many studies reported its association with lung cancer^{5,8,9,20,24}. We expect that more research on them could be performed in the future.

Methods

Literature and Study acquisition. We collected all the articles which were used to perform meta-analysis by Gu *et al.*²¹, Hu *et al.*²² and Zhan *et al.*²³. And then, we searched all the possible articles in PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) with keywords: “rs1051730” and “lung cancer”, or “rs1051730” and “Carcinoma”, or “rs1051730” and “tumor”, or “CHRNA3” and “lung cancer”, or “CHRNA3” and “Carcinoma”, or “CHRNA3” and “tumor”. All of these literatures had been collected before the PubMed’s last update on April 7 2015. In addition, we selected the related references in these articles both from PubMed and the three meta-analysis researches by using Google Scholar (<http://scholar.google.com/>). All the selected articles were written in English.

After that, we selected the studies in all obtained articles according to the following criteria: (1) The study was designed according to the method of case-control. (2) The study evaluated the association of rs1051730 polymorphism and lung cancer. (3) The number of cases and controls was provided in the study. (4) The study provided the population of each individual. (5) The study provided the number of rs1051730 genotypes both in cases and controls or provided enough data to calculate the genotypes. (6) The study provided the OR value with 95% CI and the P value or provided enough data to calculate these.

Data extraction. We extracted the following information from each study we have selected: (1) The first author of each article. (2) The publication year of each article. (3) The population and ethnicity of individual in each study. (4) The number of cases and controls in each study. (5) The number of rs1051730 genotypes both in cases and controls. (6) The OR value with 95% CI and the P value in each study. (7) The genotyping platform. If the information didn’t be provided directly, we used program R (<http://www.r-project.org/>) to work them out.

Choice of genetic model. The rs1051730 polymorphism has two alleles A and G. Among them, A is the minor allele. We analyzed the association between rs1051730 polymorphism and lung cancer by using the allele model (A allele versus G allele), the dominant model (AA+GA versus GG), the recessive model (AA versus GA+GG), and the additive model (AA versus GG), respectively. According to Table 1, among these analyses, the studies that provided the kind of genotype A/G and AA/GA/GG and GA/GG were used in the allele model, the studies that provided the kind of genotype AA/GA/GG and GA/GG were used in the dominant model, the studies that provided the kind of genotype AA/GA/GG as well as AA/GA and GG were used in the recessive model, the studies that provided the kind of genotype AA/GA/GG were used in the additive model. These classification data according to genotype were shown in supplement materials (Table S1–S4).

Heterogeneity Test. We used two quantities, Cochran's Q and I^2 , to measure the heterogeneity among the different ethnic groups. Q approximately follows a χ^2 distribution with k-1 degrees of freedom (where k is the number of studies), and the P value can be used to measure the significance level of heterogeneity. $I^2 = (Q - (K - 1)) / Q * 100\%$, which ranges from 0 to 100%. Usually, we would tentatively assign adjectives of low, moderate, and high to I^2 values of 25%, 50%, and 75%⁴². In our study, we deemed that the different ethnic groups have significant heterogeneity when $P < 0.01$ and $I^2 > 50\%$ ^{29–33}.

Meta-analysis and Subgroup Analysis. Two models can be used in meta-analysis: the random effect model and the fix effect model⁴³. After heterogeneity test, we used the software package meta (<http://cran.r-project.org/web/packages/meta/index.html>) to perform the meta-analysis with the random effect model, if the heterogeneity is significant, and used the fix effect model, if the heterogeneity is not. And then, we used Z test to examine the significance of meta-analysis. We further conducted a subgroup analysis after meta-analysis. We split original studies into smaller groups according to population, and conducted the meta-analysis to each subgroup respectively.

Publication Bias Analysis and Sensitivity Analysis. Begg's test⁴⁴ and Egger's test⁴⁵ are two methods to evaluate the publication bias. When the P values are less than 0.05, publication bias is deemed significant. We measured the publication bias of each group, which include the allele model, the dominant model, the recessive model and the additive model, by the Begg's and Egger's test. Meanwhile, we drew the funnel plot to show the bias of each group directly. And then, we removed each study in these groups orderly to measure the influence of each study. We used the software package named meta of program R for all these calculations.

References

- Jemal, A. *et al.* Global cancer statistics. *CA Cancer J Clin* **61**, 69–90 (2011).
- Siegel, R. L. *et al.* Cancer statistics, 2015. *CA Cancer J Clin* **65**, 5–29 (2015).
- Alberg, A. J. & Samet, J. M. Epidemiology of lung cancer. *Chest* **123**, 21S–49S (2003).
- Yang, I. A., Holloway, J. W. & Fong, K. M. Genetic susceptibility to lung cancer and co-morbidities. *J Thorac Dis* **5** Suppl 5, S454–462 (2013).
- 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**, 616–622 (2008).
- Thorgeirsson, T. E. *et al.* A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* **452**, 638–642 (2008).
- Shiraishi, K. *et al.* Contribution of nicotine acetylcholine receptor polymorphisms to lung cancer risk in a smoking-independent manner in the Japanese. *Carcinogenesis* **30**, 65–70 (2009).
- Zienoldiny, S. *et al.* The TERT-CLPTM1L lung cancer susceptibility variant associates with higher DNA adduct formation in the lung. *Carcinogenesis* **30**, 1368–1371 (2009).
- Schwartz, A. G., Cote, M. L., Wenzlaff, A. S., Land, S. & Amos, C. I. Racial differences in the association between SNPs on 15q25.1, smoking behavior, and risk of non-small cell lung cancer. *J Thorac Oncol* **4**, 1195–1201 (2009).
- Kohno, T. *et al.* Contribution of the TP53, OGG1, CHRNA3, and HLA-DQA1 genes to the risk for lung squamous cell carcinoma. *J Thorac Oncol* **6**, 813–817 (2011).
- VanderWeele, T. J. *et al.* Genetic variants on 15q25.1, smoking, and lung cancer: an assessment of mediation and interaction. *Am J Epidemiol* **175**, 1013–1020 (2012).
- Yang, L. *et al.* Functional polymorphisms of CHRNA3 predict risks of chronic obstructive pulmonary disease and lung cancer in Chinese. *PLoS One* **7**, e46071 (2012).
- Ren, J. H. *et al.* Association between CHRNA3 rs1051730 genotype and lung cancer risk in Chinese Han population: a case-control study. *J Huazhong Univ Sci Technolog Med Sci* **33**, 897–901 (2013).
- Sakoda, L. C. *et al.* Chromosome 15q24-25.1 variants, diet, and lung cancer susceptibility in cigarette smokers. *Cancer Causes Control* **22**, 449–461 (2011).
- Amos, C. I. *et al.* Nicotinic acetylcholine receptor region on chromosome 15q25 and lung cancer risk among African Americans: a case-control study. *J Natl Cancer Inst* **102**, 1199–1205 (2010).
- Wassenaar, C. A., Dong, Q., Amos, C. I., Spitz, M. R. & Tyndale, R. F. Pilot Study of CYP2B6 Genetic Variation to Explore the Contribution of Nitrosamine Activation to Lung Carcinogenesis. *Int J Mol Sci* **14**, 8381–8392 (2013).
- Liu, P. *et al.* Familial aggregation of common sequence variants on 15q24-25.1 in lung cancer. *J Natl Cancer Inst* **100**, 1326–1330 (2008).
- Landi, M. T. *et al.* A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. *Am J Hum Genet* **85**, 679–691 (2009).
- Wu, C. *et al.* Genetic variants on chromosome 15q25 associated with lung cancer risk in Chinese populations. *Cancer Res* **69**, 5065–5072 (2009).

20. Girard, N. *et al.* Analysis of genetic variants in never-smokers with lung cancer facilitated by an Internet-based blood collection protocol: a preliminary report. *Clin Cancer Res* **16**, 755–763 (2010).
21. Gu, M. *et al.* Strong association between two polymorphisms on 15q25.1 and lung cancer risk: a meta-analysis. *PLoS One* **7**, e37970 (2012).
22. Hu, B. *et al.* Quantitative assessment of the influence of common variations (rs8034191 and rs1051730) at 15q25 and lung cancer risk. *Tumour Biol* **35**, 2777–2785 (2014).
23. Zhan, P. & Song, Y. CHRNA3 rs1051730 polymorphism and lung cancer susceptibility in Asian population: a meta-analysis. *Transl Lung Cancer Res* **4**, 104–108 (2015).
24. Kaur-Knudsen, D., Nordestgaard, B. G. & Bojesen, S. E. CHRNA3 genotype, nicotine dependence, lung function and disease in the general population. *Eur Respir J* **40**, 1538–1544 (2012).
25. Liu, P. *et al.* A second genetic variant on chromosome 15q24-25.1 associates with lung cancer. *Cancer Res* **70**, 3128–3135 (2010).
26. Kaur-Knudsen, D., Bojesen, S. E., Tybjaerg-Hansen, A. & Nordestgaard, B. G. Nicotinic acetylcholine receptor polymorphism, smoking behavior, and tobacco-related cancer and lung and cardiovascular diseases: a cohort study. *J Clin Oncol* **29**, 2875–2882 (2011).
27. He, P. *et al.* CHRNA3 polymorphism modifies lung adenocarcinoma risk in the Chinese Han population. *Int J Mol Sci* **15**, 5446–5457 (2014).
28. Hansen, H. M. *et al.* Fine mapping of chromosome 15q25.1 lung cancer susceptibility in African-Americans. *Hum Mol Genet* **19**, 3652–3661 (2010).
29. Liu, G. *et al.* Analyzing large-scale samples confirms the association between the ABCA7 rs3764650 polymorphism and Alzheimer's disease susceptibility. *Mol Neurobiol* **50**, 757–764 (2014).
30. Shen, N. *et al.* An Updated Analysis with 85,939 Samples Confirms the Association Between CR1 rs6656401 Polymorphism and Alzheimer's Disease. *Mol Neurobiol* **51**, 1017–1023 (2015).
31. Chen, H. *et al.* Analyzing 54,936 Samples Supports the Association Between CD2AP rs9349407 Polymorphism and Alzheimer's Disease Susceptibility. *Mol Neurobiol* **52**, 1–7 (2015).
32. Li, X. *et al.* CD33 rs3865444 Polymorphism Contributes to Alzheimer's Disease Susceptibility in Chinese, European, and North American Populations. *Mol Neurobiol* **52**, 414–421 (2015).
33. Liao, M. *et al.* Analyzing large-scale samples confirms the association between rs16892766 polymorphism and colorectal cancer susceptibility. *Sci Rep* **5**, 7957 (2015).
34. Thorgeirsson, T. E. *et al.* Sequence variants at CHRN3-CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet* **42**, 448–453 (2010).
35. Ware, J. J., van den Bree, M. B. & Munafo, M. R. Association of the CHRNA5-A3-B4 gene cluster with heaviness of smoking: a meta-analysis. *Nicotine Tob Res* **13**, 1167–1175 (2011).
36. Munafo, M. R. *et al.* Association between genetic variants on chromosome 15q25 locus and objective measures of tobacco exposure. *J Natl Cancer Inst* **104**, 740–748 (2012).
37. Fu, X. W., Lindstrom, J. & Spindel, E. R. Nicotine activates and up-regulates nicotinic acetylcholine receptors in bronchial epithelial cells. *Am J Respir Cell Mol Biol* **41**, 93–99 (2009).
38. Millar, N. S. & Gotti, C. Diversity of vertebrate nicotinic acetylcholine receptors. *Neuropharmacology* **56**, 237–246 (2009).
39. Warren, G. W. & Singh, A. K. Nicotine and lung cancer. *J Carcinog* **12**, 1 (2013).
40. Arredondo, J., Chernyavsky, A. I. & Grando, S. A. The nicotinic receptor antagonists abolish pathobiologic effects of tobacco-derived nitrosamines on BEP2D cells. *J Cancer Res Clin Oncol* **132**, 653–663 (2006).
41. Li, Z., Bao, S., Xu, X., Bao, Y. & Zhang, Y. Polymorphisms of CHRNA5-CHRNA3-CHRN4 Gene Cluster and NSCLC Risk in Chinese Population. *Transl Oncol* **5**, 448–452 (2012).
42. Higgins, J. P., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560 (2003).
43. Demets, D. L. Methods for combining randomized clinical trials: strengths and limitations. *Stat Med* **6**, 341–350 (1987).
44. Begg, C. B. & Mazumdar, M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* **50**, 1088–1101 (1994).
45. Egger, M., Davey Smith, G., Schneider, M. & Minder, C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**, 629–634 (1997).

Acknowledgements

This work was supported by supported by the National Natural Science Foundation of China (No: 61571152), the National High-tech R&D Program of China (863 Program) [Nos: 2012AA02A601, 2012AA020409, 2012AA02A604, 2014AA021505, 2015AA020101, 2015AA020108] and the National Science and Technology Major Project [No: 2013ZX03005012].

Author Contributions

Q.H.J., Y.D.W. and Z.J.H. conceived and initiated the project. Q.H.J., Z.J.H., T.J.Z. and X.L.W. searched the database and extracted the information from each study. Q.H.J., Z.J.H., R.M. and J.X.W. analyzed the data. Y.B., R.J.W. and R.J.T. prepared the Figures 1–6. Z.J.H. wrote the manuscript. Q.H.J. revised the manuscript. All authors discussed the results, and contributed to the final manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Han, Z. *et al.* Analyzing large-scale samples confirms the association between the rs1051730 polymorphism and lung cancer susceptibility. *Sci. Rep.* **5**, 15642; doi: 10.1038/srep15642 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>