RESEARCH ARTICLE

Meta-Analysis of the Association between the rs8034191 Polymorphism in AGPHD1 and Lung Cancer Risk

Le Zhang1,2*, Tian-Bo Jin1,2,3*, Ya Gao4, Hui-Juan Wang1,2, Hua Yang1,2, Tian Feng2, Chen Chen2, Long-Li Kang3, Chao Chen1,2 &

Abstract

Background: Possible associations between the single nucleotide polymorphism (SNP) rs8034191 in the aminoglycosidephosphotransferase domain containing 1 (AGPHD1) gene and lung cancer risk have been studied by many researchers but the results have been contradictory. Materials and Methods: A computerized search for publications on rs8034191 and lung cancer risk was performed. Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated to assess the association between rs8034191 and lung cancer risk with 13 selected case-control studies. Sensitivity analysis, test of heterogeneity, cumulative meta-analysis, and assessment of bias were also performed. Results: A significant association between rs8034191 and lung cancer susceptibility was found using the dominant genetic model (OR=1.344, 95% CI: 1.285-1.406), the additive genetic model (OR=1.613, 95% CI: 1.503-1.730), and the recessive genetic model (OR=1.408, 95% CI: 1.319-1.503). Moreover, an increased lung cancer risk was found with all genetic models after stratification of ethnicity. Conclusions: The association between rs8034191 and lung cancer risk was significant using multiple genetic models, suggesting that rs8034191 is a risk factor for lung cancer. Further functional studies of this polymorphism and lung cancer risk are warranted.

Keywords: Lung cancer - single nucleotide polymorphism - AGPHD1 - genetic polymorphism - meta-analysis

Asian Pac J Cancer Prev, 16 (7), 2713-2717

Introduction

Lung cancer is the leading cause of cancer-related death throughout the world with an estimated 1.3 million new cases diagnosed annually (Shibuya et al., 2002; Herbst et al., 2008). In many countries, the morbidity and mortality of lung cancer have increased rapidly in recent years (Bhat et al., 2013; Shukla et al., 2013; Yilmaz et al., 2014). Well-known risk factors for lung cancer include cigarette smoking and exposure to ionizing radiation. Although over 80% of lung cancer cases are related to the use of tobacco (Parkin et al., 1994), only a small percentage of smokers (<20%) develop this disease. A accumulating evidence suggests that genetic factors may contribute to variation in susceptibility to lung cancer. It is widely accepted that lung cancer is a complex multifactorial disease that is attributed to the interaction of genetic factors with environmental factors (Amos et al., 2008; Heller et al., 2010; ). Despite intensive efforts devoted to investigating the genetic factors associated with lung cancer, the genes and genetic variants that drive the development of lung cancer remain unclear. Recently, the chromosome 15q24-25.1 region has been identified as a hot spot for lung cancer susceptibility by genome-wide association (GWA) studies (Hung et al., 2008; Thorgerisson et al., 2008; Broderick et al., 2009; Wei et al., 2011). Genes that map to this region include aminoglycoside phosphotransferase domain containing 1 (AGPHD1); cholinergic receptor, nicotinic, alpha 3 (CHRNA3); cholinergic receptor, nicotinic, alpha 4 (CHRNA4); cholinergic receptor, nicotinic, alpha 5 (CHRNA5); PSM4; and LOC123688. In particular, the relationship between the single nucleotide polymorphism (SNP) rs8034191 in AGPHD1 and lung cancer risk has been widely investigated but the results have been inconclusive (Mantel et al., 1959; Amos et al., 2008; Schwartz et al., 2009; Zienolddiny et al., 2009; Truong et al., 2010; Chen et al., 2011; Jaworowska et al., 2011; Sakoda et al., 2011; Wang et al., 2013). Due to insufficient sample size, these previous studies have lacked statistical power to detect common variants that have minor effects on lung carcinogenesis. Furthermore, the results of these studies are not reproducible. To address the heterogeneity and publication bias among
Le Zhang et al

previous studies and better understand the effect of the SNP rs8034191 on the risk of lung cancer we performed a meta-analysis of 13 selected case-control studies.

Materials and Methods

Publication search and inclusion/exclusion criteria

In August 2013, we searched PubMed, Google Scholar, EMBASE, and the China National Knowledge Infrastructure using the following search terms: AGPHD1, rs8034191, lung cancer, gene, genotype, mutation, and polymorphism. Articles identified in the primary literature met our initial criteria for inclusion in the meta-analysis if they were published in English, focused on humans, and were free of obvious overlap with other studies.

From the publications identified above, two investigators independently selected articles containing information on the association between AGPHD1 and lung cancer morbidity and checked the corresponding reference lists. If multiple studies were published on the same population or subpopulation, only the most recent or informative study was included in the meta-analysis.

Articles were included in this meta-analysis if they 1) examined the hypothesis that rs8034191 is associated with lung cancer risk, 2) followed a nested case-control, case-control, or cross-sectional study design, and 3) provided estimates of ORs and corresponding 95% CIs or sufficient information on genotype/allele counts between cases and controls to calculate the ORs and 95% CIs. Articles were excluded if they included non-case-control studies, a control population containing patients with malignant tumor, or were redundant with other published studies. A flow chart outlining the selection process for inclusion in the meta-analysis is shown in Figure 1.

Data extraction

The following information was extracted from each study: the first author’s name, the year of publication, the country in which the study was performed, ethnicities of subjects, and the number of cases and controls with the TT, TC, and CC rs8034191 genotypes. Two investigators independently extracted the data from all eligible publications, and any inconsistencies were resolved by discussion.

All statistical analyses were performed using STATA software (version 11.0: Stata Corporation, College Station, TX). Two-sided p-values less than 0.05 were considered statistically significant. We calculated the allelic frequencies for the case and control groups in each study and assessed them for Hardy-Weinberg equilibrium (HWE) using an chi-square test (Egger et al., 1997). The OR and 95% CI values were determined to assess the strength of the association between each rs8034191 polymorphism and lung cancer risk. For each study, the OR and 95% CI were assessed in an additive model, a recessive model, and an additive model. Subgroup analyses were performed based on the source of the controls and the ethnicity of the study participants. The chi-squared based Q-statistic was calculated to test for heterogeneity among the studies. If the studies were found to be heterogeneous (p<0.05) the pooled ORs were analyzed using a random-effects model (Higgins et al., 2002); otherwise, a fixed-effects model was used (Egger et al., 1997). The I² statistic was then used to quantitatively estimate heterogeneity, with Fless than 25%, between 25% and 75%, and greater than 75% representing low, moderate, and high degrees of inconsistency, respectively (Higgins et al., 2003; Hemminki et al., 2006). The significance of the combined OR was determined using a Z test (p<0.05 was considered statistically significant). Cumulative meta-analyses were performed on all eligible cancer studies according to case sample size. Additionally, sensitivity of the meta-analysis was evaluated through the sequential removal of each study.

Finally, we produced a Begg’s funnel plot and performed an Egger’s test to statistically assess publication bias.

Table 1. Characteristics of the studies on the association between AGPHD1 rs8034191 polymorphisms and cancer risk included in the meta-analysis

<table>
<thead>
<tr>
<th>ID</th>
<th>Author</th>
<th>Year</th>
<th>Ethnic group</th>
<th>Sample Size (Case/Control)</th>
<th>Source of controls</th>
<th>Case Alleles</th>
<th>Control Alleles</th>
<th>p-value (HWE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amos et al. (Texas discovery)</td>
<td>2008</td>
<td>Caucasian</td>
<td>1153/1137</td>
<td>HB</td>
<td>426</td>
<td>536/191</td>
<td>0.522</td>
</tr>
<tr>
<td>2</td>
<td>Amos et al. (Texas discovery)</td>
<td>2008</td>
<td>Caucasian</td>
<td>698/591</td>
<td>HB</td>
<td>259</td>
<td>328/111</td>
<td>0.69</td>
</tr>
<tr>
<td>3</td>
<td>Amos et al. (USA)</td>
<td>2008</td>
<td>Caucasian</td>
<td>1831/960</td>
<td>HB</td>
<td>670</td>
<td>858/303</td>
<td>0.345</td>
</tr>
<tr>
<td>4</td>
<td>Schwartz et al. (Caucasian)</td>
<td>2009</td>
<td>Caucasian</td>
<td>809/539</td>
<td>HB</td>
<td>185</td>
<td>264/90</td>
<td>0.46</td>
</tr>
<tr>
<td>5</td>
<td>Schwartz et al. (African American)</td>
<td>2009</td>
<td>African American</td>
<td>421/360</td>
<td>PB</td>
<td>231</td>
<td>119/10</td>
<td>0.148</td>
</tr>
<tr>
<td>6</td>
<td>Zienolddiny et al. (Caucasian)</td>
<td>2008</td>
<td>Caucasian</td>
<td>1831/960</td>
<td>HB</td>
<td>670</td>
<td>858/303</td>
<td>0.345</td>
</tr>
<tr>
<td>7</td>
<td>Truong et al. (Caucasian)</td>
<td>2010</td>
<td>Caucasian</td>
<td>7259/9463</td>
<td>HB</td>
<td>2586</td>
<td>3488/1185</td>
<td>0.344</td>
</tr>
<tr>
<td>8</td>
<td>Truong et al. (Asian)</td>
<td>2010</td>
<td>Asian</td>
<td>1690/2117</td>
<td>-</td>
<td>1583</td>
<td>104/3</td>
<td>0.3</td>
</tr>
<tr>
<td>9</td>
<td>Sakoda et al. (non-Hispanic white)</td>
<td>2011</td>
<td>Caucasian</td>
<td>746/1475</td>
<td>PB</td>
<td>258</td>
<td>369/199</td>
<td>0.159</td>
</tr>
<tr>
<td>10</td>
<td>Wei et al. (Caucasian)</td>
<td>2011</td>
<td>Caucasian</td>
<td>198/295</td>
<td>HB</td>
<td>64</td>
<td>100/34</td>
<td>0.31</td>
</tr>
<tr>
<td>11</td>
<td>Jaworewika et al. (Polish)</td>
<td>2011</td>
<td>Caucasian</td>
<td>833/831</td>
<td>HB</td>
<td>286</td>
<td>419/128</td>
<td>0.351</td>
</tr>
<tr>
<td>12</td>
<td>Chen et al. (Caucasian)</td>
<td>2011</td>
<td>Caucasian</td>
<td>487/974</td>
<td>HB</td>
<td>222</td>
<td>212/53</td>
<td>0.373</td>
</tr>
<tr>
<td>13</td>
<td>Wang et al. (Chinese)</td>
<td>2012</td>
<td>Asian</td>
<td>381/410</td>
<td>HB</td>
<td>350</td>
<td>29/2</td>
<td>0.524</td>
</tr>
</tbody>
</table>
Table 2. Stratified Analyses of the Association between AGPHD1 rs8034191 Polymorphisms and Lung Cancer Risk

<table>
<thead>
<tr>
<th>Variables</th>
<th>Dominant model</th>
<th>Recessive model</th>
<th>Additive model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR(95% CI)</td>
<td>p*</td>
<td>OR(95% CI)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>1.127(0.888-1.432)</td>
<td>0.447</td>
<td>1.887(0.489-7.274)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>1.352(1.291-1.417)</td>
<td>0.107</td>
<td>1.412(1.323-1.508)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1.344(1.285-1.406)</td>
<td>0.142</td>
<td>1.408(1.319-1.503)</td>
</tr>
</tbody>
</table>

*p-value from chi-square test for heterogeneity; *A fixed-effects model was used when the p value from the heterogeneity test was <0.05; otherwise, a random-effects model was used.

Results

Eligible studies and quality assessment

Through the procedure outlined in Figure 1 we identified nine articles that met the inclusion criteria. These articles covered 13 case-control studies including 16,858 cases and 19,576 controls. The characteristics of the studies are listed in Table 1. The studies focused solely on lung cancer and represented multiple ethnic populations. The alleles at rs8034191 were in HWE in all of the studies. After evaluating these studies, all 13 were deemed to be of sufficient quality to be included in our analysis.

Meta-analysis results

After pooling the 16,858 cases and 19,576 controls in the meta-analysis we found a significant association between AGPHD1 rs8034191 polymorphisms and lung cancer risk using the dominant model (OR=1.344, 95% CI: 1.285-1.406), the additive model (OR=1.613, 95% CI: 1.503-1.730), and the recessive model (OR=1.408, 95% CI: 1.319-1.503).

There was no significant heterogeneity in the rs8034191 variant genemod. We stratified the data by dividing the participants into two subgroups based on ethnicity: Asian and Caucasian. The pooled ORs for the recessive model were 1.887 (95% CI 0.489-7.274) for the Asian subgroup and 1.412 (95% CI 1.323-1.508) for the Caucasian subgroup (Table 2).

Test for heterogeneity

Based on the dominant model, there were significant heterogeneities associated with the study of Chen et al. (OR=0.955; 95% CI: 0.768-1.189) (Figure 2). Likewise, heterogeneity was detected in the study of African Americans by Schwartz et al. when using either the recessive model (OR=0.773; 95% CI: 0.343-1.743).
Discussion

We performed a comprehensive meta-analysis to evaluate the association of a common polymorphism 15q25.1 with the risk of lung cancer. By performing subgroup analyses, we identified ethnicity as a potential source of inconsistency between studies. This is not surprising, as it is well established that genetic heterogeneity is inevitable in disease identification strategies (Higgins et al., 2003). Specifically, the overall results of this study demonstrated that the rs8034191-T allele of the AGPHD1 gene might be a risk factor for the development of lung cancer in Caucasians, but not in Asians (Table 2). We also noticed remarkable differences in the rs8034191-C allele between Caucasians and Asians, making it very difficult to detect weak associations in Asians unless examining a very large population. This suggests that different genetic backgrounds may differentially affect this allele or that different populations may have different linkage disequilibrium patterns; for example, the studied polymorphisms may be in linkage with another causative variant in one ethnic population but not in another (Hemminki et al., 2006). Considering the divergent genetic backgrounds, it is necessary to construct a database of polymorphisms related to lung cancer in each ethnic/racial group.

To the best of our knowledge, the present study is the only meta-analysis to date investigating the association of the rs8034191 polymorphism with lung cancer susceptibility. Although potential sources of heterogeneity cannot be easily eliminated, the strengths of this study include the relatively large sample size, the lack of deviation from Hardy-Weinberg equilibrium, and the high quality of the studies involved. However, this study should be interpreted with several technical limitations in mind. First, most of the studies in this meta-analysis were case-control studies, which are susceptible to selection bias by including only non-fatal cases. Second, because only studies in the English language were considered, a publication bias may have been introduced. To address this, we performed a funnel plot and an accompanying Egger’s test, which did not reveal an obvious bias. Moreover, any asymmetry in the funnel plot, either through visual interpretation or statistically testing, may result from a fundamental difference between small and large studies that arises from inherent inter-study heterogeneity. There is no perfect method to test for publication bias, and the validity of the funnel plot and Egger’s test have been challenged (Yu et al., 2010). Thus, we cannot completely rule out the low probability that relevant studies (for example small negative studies) are missing from the plot although the trim and fill method suggested that no missing studies were required to make the funnel plot symmetrical for either polymorphism. Third, the single-locus-based nature of this meta-analysis precluded the possibility of investigating gene-gene and gene-environment interactions, as well as haplotype-based effects. In particular, further studies should investigate other markers adjacent to 15q25.1 to clarify whether the observed association is causal or due to linkage disequilibrium. It is likely that the rs8034191 SNP alone makes a minor contribution to risk prediction in lung cancer patients, and further studies are needed to determine whether multiple polymorphisms integrated with other risk factors will enhance the predictive capabilities. Additional studies are necessary to fully understand the relationship between the rs8034191 SNP and lung cancer susceptibility.

In conclusion, we have expanded previous studies by providing evidence that the rs8034191-T allele of the AGPHD1 gene might be a risk factor for the development of lung cancer in Caucasians, but not in Asians. Functional studies of the association between this polymorphism and cancer risk are warranted.

Acknowledgements

This work is supported by National 863 High Technology Research and Development Program (No. 2012AA02A519). We are grateful to all the patients and individuals in the studies who made this work possible. We would also like to thank the clinicians and other hospital staff who contributed to the data collection for this study.

References

Egger M, Davey SG, Schneider M, Minder C (1997). Bias in
Meta-analysis detected by a simple, graphical test. *BMJ*, 315, 629-34.


