The relationship between apolipoprotein E gene ε2/ε3/ε4 polymorphism and breast cancer risk: a systematic review and meta-analysis

Yun-Long Liu1
Hao-Min Zhang2
Hong-Ming Pan2
Yu-Hang Bao2
Jing Xue2
Tian-Chang Wang1
Xiao-Cheng Dong1
Xiao-Ling Li3
Hong-Guang Bao1

1Department of Chest Surgery, The Second Affiliated Hospital of Qiqihar Medical University, Qiqihar, Heilongjiang, People’s Republic of China
2Department of Anatomy, Basic Medical Science College, Qiqihar Medical University, Qiqihar, Heilongjiang, People’s Republic of China
3Department of Anatomy, Basic Medical Science College, Qiqihar Medical University, Qiqihar, Heilongjiang, People’s Republic of China

Correspondence: Hong-Guang Bao
Department of Chest Surgery, The Second Affiliated Hospital of Qiqihar Medical University, Zhonghua West Road 37, Jianhua District, Qiqihar City 161006, Heilongjiang Province, People’s Republic of China
Tel +86 452 2736 380
Email bhg28@163.com

Xiao-Ling Li
Department of Anatomy, Basic Medical Science College, Qiqihar Medical University, Bukul North Street 333, Jianhua District, Qiqihar City 161006, Heilongjiang Province, People’s Republic of China
Tel +86 452 2663 182
Email lixiaoling_lsm@126.com

Objective: We conducted a systematic review and meta-analysis aiming to assess the relationship between apolipoprotein E (APOE) gene ε2/ε3/ε4 polymorphism and breast cancer risk.

Methods: Yun-Long Liu and Hao-Min Zhang independently completed literature retrieval and data collection, and statistical analyses were performed by Statia. Individual odds ratio (OR) and 95% confidence interval (CI) were pooled in a random-effects model using the DerSimonian–Laird method. Heterogeneity was evaluated by F statistic at a significance level of 50%. Publication bias was assessed by Egger’s test.

Results: Eleven articles including 2,074 breast cancer patients and 2,372 controls were summarized. Using the most common allele ε3 as a reference, the ε2 (OR = 0.87, 95% CI = 0.72–1.05, P = 0.154, F = 0.0%) and ε4 (OR = 1.07, 95% CI = 0.80–1.42, P = 0.654, F = 71.8%) alleles were not found to be significantly associated with breast cancer risk in the overall analyses. Subgroup analyses revealed that the comparison of allele ε4 with ε3 was significant in Asians (OR = 1.58, 95% CI = 1.17–6.32, P = 0.003, F = 12.1%) and in studies that used the restriction fragment length polymorphism (RFLP) genotyping method (OR = 1.27; 95% CI = 1.01–1.61, P = 0.045, F = 34.3%), and was marginally significant in hospital-based studies (OR = 1.33; 95% CI = 0.98–1.79, P = 0.065, F = 30.2%), without heterogeneity. Moreover, the presence of the ε2 allele was significantly associated with breast cancer in small studies (total sample size < 500) (OR = 0.73, 95% CI = 0.54–1.00, P = 0.052, F = 0.0%) without heterogeneity. The Egger’s test indicated low probabilities of publication bias.

Conclusion: We observed a significant association between APOE gene ε4 allele and breast cancer risk in Asian populations. Moreover, the findings of our subgroup analyses suggest that source of controls, genotyping platform, and sample size might be the potential causes of heterogeneity.

Keywords: apolipoprotein E, breast cancer, polymorphism, meta-analysis

Introduction
Breast cancer is a heterozygous disease with a strong heritable component, and it is estimated that ~15% of breast cancer cases have at least one affected first-degree relative.1 Despite the tremendous advancement gained in medical sciences, to unravel the genetic architecture of breast cancer still remains a challenge. So far, a number of susceptibility genes have been suggested to underlie breast carcinogenesis.2,3 One of the promising candidate genes is apolipoprotein E (APOE) and its expression is proposed as a tumor-associated biomarker.4 APOE is a class of apolipoprotein and its primary function is to mediate cholesterol metabolism and catabolize triglyceride-rich lipoprotein constituents. Beyond this well-known function, evidence is amassing to reveal a different role for APOE in carcinogenesis and tumor development.5 The gene
encoding APOE is mainly inherited as one of the three alleles: ε2 (Arg148→Cys), ε3 (Cys112), and ε4 (Cys112→Arg), which generate three homozygous genotypes (ε2/2, ε3/3, and ε4/4) and three heterozygous genotypes (ε2/3, ε2/4, and ε3/4), each with a diverse receptor-binding capacity.6 In addition, there are several other minor alleles such as ε5 and ε7 in APOE gene; due to the methodological drawbacks of restriction fragment length polymorphism (RFLP), these minor alleles could be easily missed.7,8 In medical research literature, the ε2, ε3, and ε4 alleles in APOE gene are given the most attention, and their susceptibility toward breast cancer risk has been widely evaluated by genetic association studies;9–11 however, the results of only a few studies are reproducible. A previous meta-analysis of eight articles by Saadat12 found that APOE gene ε4 allele or ε4 positive genotypes was exclusively associated with breast cancer in Asians. Further meta-analysis by Anand et al13 failed to validate this significant finding for ε4 positive genotypes in overall susceptibility to breast cancer. Besides unavoidable genetic heterogeneity across races and ethnicities, a possible explanation for such divergence is inadequate statistical rigor to quantify the effect size reliably. To obtain a more precise estimate, we conducted an updated systematic review and meta-analysis of previous studies by incorporating additional articles and seeking other potential causes of heterogeneity in both subgroup and meta-regression analyses.

Methods

This systematic review and meta-analysis was conducted based on the guidelines listed in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.

Literature retrieval

A comprehensive literature search was conducted in the Medline, Embase, and Web of Science electric databases. Key subjects used for retrieval included apolipoprotein E or APOE or APO E and breast cancer or breast carcinoma. The last search was on July 5, 2015. Each retrieved article was reviewed from the title or abstract to definitively exclude the one with irrelevant content. In case of uncertainty, the full text was downloaded for further review. To avoid missing hits, the bibliographies of original articles and reviews were manually searched.

Eligibility appraisal

The eligibility of each retrieved article was appraised according to the following criteria: 1) the clinical endpoint must be breast cancer with validated clinical diagnoses; 2) it should be designed as a case–control association study; 3) the genotypes or alleles of APOE gene ε2/ε3/ε4 polymorphism or its associated risk estimates – odds ratio (OR) or beta and its 95% confidence interval (95% CI) or standard error – should be provided; and 4) conference abstracts or proceedings (due to insufficient necessary information) and non-English articles are not considered. In case of more than one independent study in a single article, each study was summarized and analyzed separately.

Information extraction

Relevant information from each eligible article was extracted separately by two authors (Yun-Long Liu and Hao-Min Zhang) according to a predefined table designed by all responsible authors, including authors, the year of publication, race/ethnicity, study design, source of controls, genotyping method, sample size, age, body mass index (BMI), and the genotype or allele counts of APOE gene ε2/ε3/ε4 polymorphism or the associated ORs and 95% CIs. In case of unresolved-by-discussion disagreements, a third senior author (Hong-Guang Bao) was involved until a consensus was reached.

Statistical analysis

The risk estimate for APOE gene ε2/ε3/ε4 polymorphism in susceptibility to breast cancer was expressed as OR and its 95% CI, which were pooled in a random-effects model using the method proposed by DerSimonian and Laird.14 The heterogeneity among the results of different studies was weighed by the I² statistic, which is defined as the percentage of observed between-study variability that stems from heterogeneity rather than chance. As recommended by Higgins et al, an I² statistic of 50% or greater is generally interpreted as significant heterogeneity.15 To explore possible causes of heterogeneity, both subgroup and meta-regression analyses were conducted. Publication bias was assessed by the trim-and-fill method and the Egger’s test.16 The significance of Egger’s test is set at 10% or less. The trim-and-fill method can infer the existence of unpublished hidden articles from a filled Funnel plot and correct the meta-analysis by imputing the presence of missing studies to yield an unbiased pooled estimate.17 The aforementioned statistical analyses were completed with the Stata software release 12 (StataCorp, College Station, TX, USA).

Results

Article selection

After a careful literature retrieval and stringent eligibility appraisal, eleven out of 43 initially retrieved articles were
included and summarized in this systematic review and meta-analysis.9–11,18–25 The PRISMA flowchart schematizing the selection of qualified studies is presented in Figure 1. The total sample size was 4,446, including 2,074 breast cancer patients and 2,372 control subjects. Two studies conducted in different locations were reported by Menzel et al,21 and they were handled separately, leading to a total of 12 studies in the final analysis.

**Study characteristics**
Table 1 shows the baseline characteristics of all included studies. Out of 12 independent studies, four were performed in Caucasians, three in Asians, and five in mixed-ethnic populations. Two studies were designed longitudinally and ten cross-sectionally. Of the 12 studies, six studies had controls collected from hospitals and six from populations. The RFLP genotyping method was adopted in seven studies. Breast cancer patients were slightly older than control subjects (54.95 vs 50.94 years, \( P=0.070 \)). Mean BMI levels were comparable between the two groups (\( P>0.999 \)).

**Overall estimates**
As the study by Mandelblatt et al25 provided only \( \varepsilon 4 \) positive/negative genotypes, overall allele comparisons were based on the remaining eleven studies (Figure 2). In overall analyses, using the most common \( \varepsilon 3 \) allele as reference, the \( \varepsilon 2 \) (OR =0.87, 95% CI =0.72–1.05, \( P=0.154 \), \( F=0.0\% \)) and \( \varepsilon 4 \) (OR =1.07, 95% CI =0.80–1.42, \( P=0.654 \), \( F=71.8\% \)) alleles were found not to be significantly associated with breast cancer risk in a random-effects model using the DerSimonian and Laird method. The probability of publication bias was low according to the Egger’s test (for \( \varepsilon 2 \) vs \( \varepsilon 3 \), \( P=0.270 \), and for \( \varepsilon 4 \) vs \( \varepsilon 3 \), \( P=0.186 \)); however, in the trim-and-fill method, three missing studies were filled for the comparison of alleles \( \varepsilon 4 \) with \( \varepsilon 3 \), in order to make the filled funnel plot symmetrical (Figure 3). After taking these three studies into consideration, the odds of breast cancer conferred by \( \varepsilon 4 \) allele
was 0.90 (95% CI =0.68–1.20, \( P=0.487 \)), with significant heterogeneity (\( F=75.4\% \)).

In addition, for the comparison of \( \varepsilon 4 \) positive with \( \varepsilon 4 \) negative genotypes, the odds of having breast cancer was 1.05 (95% CI =0.75–1.11, \( P=0.782 \)) and heterogeneity was not significant (\( F=46.1\% \)) (Figure S1).

**Stratified estimates**

To examine whether effect estimates were homogeneous in subgroups by race, study design, source of controls, genotyping method, and sample size, as well as whether heterogeneity is improved, a set of subgroup analyses were carried out in a random-effects model using the DerSimonian and Laird method (Table 2). By race, no significance was noted for the comparison of alleles \( \varepsilon 2 \) with \( \varepsilon 3 \) in Caucasians, Asians, and mixed-ethnic populations, while the comparison of alleles \( \varepsilon 4 \) with \( \varepsilon 3 \) was significant in Asians only (OR =1.58, 95% CI =1.17–6.32, \( P=0.003 \)) without heterogeneity (\( F=12.1\% \)). Grouping studies by study design and source of controls revealed only marginal significance for the comparison of alleles \( \varepsilon 4 \) with \( \varepsilon 3 \) in hospital-based studies (OR =1.33, 95% CI =0.98–1.79, \( P=0.065 , F=30.2\% \)). Similarly, studies with the RFLP genotyping method showed a significant effect estimate of 1.27 (95% CI =1.01–1.61, \( P=0.045 , F=34.3\% \)) for the comparison of alleles \( \varepsilon 4 \) with \( \varepsilon 3 \).

**Meta-regression analysis**

Two steps were taken by modeling potential confounders in a meta-regression analysis. First, potential confounders including age, BMI, race, study design, source of controls, genotyping method, and sample size were modeled one at a time, and unfortunately there was no indication of significance in all regression models. Second, age, BMI, race, study design, source of controls, genotyping method, and sample size were modeled altogether and still no significant confounding effects were observed (all \( P>0.05 \)).

**Discussion**

This is an updated systematic review and meta-analysis to assess the susceptibility of APOE gene \( \varepsilon 2/\varepsilon 3/\varepsilon 4 \)
polymorphism to breast cancer on the basis of eleven articles and 4,446 study subjects. Consistent with the meta-analytical results by Saadat, we in a larger sample size observed a significant association between APOE gene ε4 allele and breast cancer risk in Asians. In addition, our findings in subgroup analyses suggest that source of controls, genotyping platform, and sample size might be potential causes of heterogeneity.

There were experimental data reporting a high expression level of APOE gene in many forms of cancer, including breast cancer. Elevated APOE level was postulated to trigger tumor growth by providing lipid substrates for tumor cells. It was reported that different APOE isoforms affected tumor growth and proliferation at different levels, with ε2 allele having a largest protective effect, ε3 allele the moderate, and ε4 allele the least. Consistent with this

<table>
<thead>
<tr>
<th>Study ID</th>
<th>ε2 vs ε3 (P=0.154)</th>
<th>OR (95% CI)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Niemi et al19</td>
<td>1.09 (0.57, 2.08)</td>
<td>8.66</td>
<td></td>
</tr>
<tr>
<td>Moysirsch et al18</td>
<td>0.99 (0.62, 1.59)</td>
<td>16.66</td>
<td></td>
</tr>
<tr>
<td>Yaylim et al20</td>
<td>3.29 (0.15, 70.38)</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Menzel et al21 (Prague)</td>
<td>0.84 (0.44, 1.64)</td>
<td>8.29</td>
<td></td>
</tr>
<tr>
<td>Menzel et al21 (Tyrol)</td>
<td>0.78 (0.49, 1.24)</td>
<td>16.83</td>
<td></td>
</tr>
<tr>
<td>Chang et al22</td>
<td>1.06 (0.70, 1.61)</td>
<td>20.89</td>
<td></td>
</tr>
<tr>
<td>Chang et al23</td>
<td>0.73 (0.42, 1.26)</td>
<td>12.31</td>
<td></td>
</tr>
<tr>
<td>Surekha et al11</td>
<td>0.57 (0.21, 1.57)</td>
<td>3.54</td>
<td></td>
</tr>
<tr>
<td>Porrata-Doria et al12</td>
<td>0.70 (0.35, 1.43)</td>
<td>7.29</td>
<td></td>
</tr>
<tr>
<td>Mcdonald et al14</td>
<td>1.46 (0.37, 5.71)</td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td>Cibeira et al9</td>
<td>0.40 (0.14, 1.17)</td>
<td>3.18</td>
<td></td>
</tr>
<tr>
<td>Overall (P=0.0%, P=0.766)</td>
<td>0.87 (0.72, 1.05)</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study ID</th>
<th>ε4 vs ε3 (P=0.654)</th>
<th>OR (95% CI)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Niemi et al19</td>
<td>0.96 (0.69, 1.33)</td>
<td>12.35</td>
<td></td>
</tr>
<tr>
<td>Moysirsch et al18</td>
<td>0.89 (0.62, 1.26)</td>
<td>12.08</td>
<td></td>
</tr>
<tr>
<td>Yaylim et al20</td>
<td>1.98 (0.08, 49.70)</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Menzel et al21 (Prague)</td>
<td>0.72 (0.39, 1.32)</td>
<td>8.81</td>
<td></td>
</tr>
<tr>
<td>Menzel et al21 (Tyrol)</td>
<td>0.48 (0.35, 0.67)</td>
<td>12.45</td>
<td></td>
</tr>
<tr>
<td>Chang et al22</td>
<td>1.39 (1.00, 1.93)</td>
<td>12.32</td>
<td></td>
</tr>
<tr>
<td>Chang et al23</td>
<td>1.68 (0.94, 3.00)</td>
<td>9.17</td>
<td></td>
</tr>
<tr>
<td>Surekha et al11</td>
<td>2.73 (1.18, 6.32)</td>
<td>6.51</td>
<td></td>
</tr>
<tr>
<td>Porrata-Doria et al12</td>
<td>1.23 (0.81, 1.87)</td>
<td>11.21</td>
<td></td>
</tr>
<tr>
<td>Mcdonald et al14</td>
<td>1.54 (0.57, 4.17)</td>
<td>5.30</td>
<td></td>
</tr>
<tr>
<td>Cibeira et al9</td>
<td>1.01 (0.56, 1.82)</td>
<td>9.06</td>
<td></td>
</tr>
<tr>
<td>Overall (P=71.8%, P=0.000)</td>
<td>1.07 (0.80, 1.42)</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Forest plots of APOE gene ε2 and ε4 alleles relative to ε3 allele in overall analyses.
Note: Weights are from random effects analysis.
Abbreviations: OR, odds ratio; CI, confidence interval.
tendency, in overall analyses, we observed a reduced risk of breast cancer for ε2 allele and an increased risk for ε4 allele with reference to ε3 allele; yet no statistical significance was detected. However, after restricting the analysis to populations of Asian descent, the presence of ε4 allele remarkably rendered the study subjects 58% more likely to develop breast cancer, suggesting a race-specific role of APOE gene in breast carcinogenesis. Besides genetic makeup, lifestyle backgrounds such as physical activity and dietary habits could also account for this race-specific divergence. Due to the limitation of unavailable lifestyle data, there is always a possibility of residual confounding.

Heterogeneity is a potential issue when interpreting the results of all meta-analyses in medical literature. In general, a meta-analysis can at least in part confirm the involvement of a genetic variant if heterogeneity is sufficiently recognized and taken into account. As revealed by our subgroup analyses, source of controls, genotyping platform, and sample size might be potential causes of heterogeneity. Taking source of controls as an example, it is widely believed that there is poor comparability between cases and controls in hospital-based studies given a regional specialty for the disease under study and differential hospitalization rates between cases and controls. In contrast, study subjects drawn from a community or a general population might be representative of a true population. In view of this divergence, the results from population-based studies might be convincing. Moreover, sometimes genotyping errors can induce biases in

Table 2 Subgroup analyses of APOE gene ε2 and ε4 alleles relative to ε3 allele on breast cancer risk

<table>
<thead>
<tr>
<th>Subgroups</th>
<th>Number of studies</th>
<th>ε2 vs ε3</th>
<th>ε4 vs ε3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR 95% CI</td>
<td>P-value</td>
<td>I²</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>P_reg</td>
</tr>
<tr>
<td>Caucasian</td>
<td>4</td>
<td>0.91</td>
<td>0.472</td>
</tr>
<tr>
<td>Asian</td>
<td>3</td>
<td>0.88</td>
<td>0.430</td>
</tr>
<tr>
<td>Mixed</td>
<td>4</td>
<td>0.72</td>
<td>0.252</td>
</tr>
<tr>
<td>Study design</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>9</td>
<td>0.90</td>
<td>0.296</td>
</tr>
<tr>
<td>Longitudinal</td>
<td>2</td>
<td>0.66</td>
<td>0.236</td>
</tr>
<tr>
<td>Source of controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital-based</td>
<td>6</td>
<td>0.83</td>
<td>0.255</td>
</tr>
<tr>
<td>Population-based</td>
<td>5</td>
<td>0.89</td>
<td>0.349</td>
</tr>
<tr>
<td>Genotyping method</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFLP</td>
<td>7</td>
<td>0.87</td>
<td>0.230</td>
</tr>
<tr>
<td>Others</td>
<td>4</td>
<td>0.88</td>
<td>0.438</td>
</tr>
<tr>
<td>Total sample size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥500 (large studies)</td>
<td>4</td>
<td>0.96</td>
<td>0.760</td>
</tr>
<tr>
<td>&lt;500 (small studies)</td>
<td>7</td>
<td>0.73</td>
<td>0.052</td>
</tr>
</tbody>
</table>

Abbreviations: OR, odds ratio; CI, confidence interval; P_reg, P-value from meta-regression analysis; RFLP, restriction fragment length polymorphism.
frequency estimates for APOE gene $\varepsilon2/\varepsilon3/\varepsilon4$ polymorphism. For example, partial digestion of PCR product in RFLP genotyping method may introduce a misclassification bias, and it is essential to have RFLP images read independently by two persons. Further we know that individual studies with small sample sizes may have not sufficiently statistical power to detect a small effect and yield a fluctuated estimate, and often wide CIs may indicate lack of precision. On the basis of these observations, it is high time to confirm the findings of this systematic review and meta-analysis in more population-based, larger association studies with validated genotyping methods.

This systematic review and meta-analysis is limited due to the following aspects. First, only articles in English language are included, and a selection bias cannot be excluded, although no observable publication bias was recorded by the Egger’s test. Second, only one polymorphism in APOE gene is analyzed and other polymorphisms in APOE gene and other susceptibility genes are not considered in this study, leading to the possibility that the potential role of APOE gene $\varepsilon2/\varepsilon3/\varepsilon4$ polymorphism is diluted or masked by other gene–gene and gene–environment interactions. Third, this systematic review and meta-analysis is not based on individual participant data and, therefore, adjusted effect-size estimates cannot be derived. Fourth, because mainly populations from Asia and Western countries were available, the present data did not allow us to extrapolate our findings to other racial or ethnic groups.

In conclusion, we in this updated systematic review and meta-analysis observed a significant association between APOE gene $\varepsilon4$ allele and breast cancer risk in Asians. Moreover, in our subgroup analyses, source of controls, genotyping platform, and sample size might be potential causes of heterogeneity. Further epidemiologic and functional validation is necessitated to provide a convincing effect estimate.

Acknowledgments
This study was supported by the Qiqihar Medical University Inner Project (QY2016B-06) and the Science and Technology Planning Project of Qiqihar City (SFGG-201205).

Disclosure
The authors report no conflicts of interest in this work.

References


Supplementary material

Figure S1 Forest plot of APOE gene ε4 positive genotypes relative to ε4 negative genotypes in overall analysis.

Note: Weights are from random effects analysis.

Abbreviations: OR, odds ratio; CI, confidence interval.

References