Common variants at 12q24 are associated with drinking behavior in Han Chinese1–3

Xueli Yang, Xiangfeng Lu, Laiyuan Wang, Shufeng Chen, Jianxin Li, Jie Cao, Jichun Chen, Yongchen Hao, Ying Li, Liancheng Zhao, Hongfan Li, Donghua Liu, Liguig Wang, Fanghong Lu, Chong Shen, Lin Yu, Xianping Wu, Qi Zhao, Xu Ji, Dongshuang Guo, Xiaozhong Peng, Jianfeng Huang, and Dongfeng Gu

ABSTRACT

Background: Alcohol consumption is heritable, but genetic susceptibility to drinking behavior has not been investigated widely in genome-wide association studies.

Objective: We aimed to identify susceptibility loci for drinking behavior (drinkers compared with nondrinkers) in Han Chinese.

Design: We performed 2 genome-wide association studies including 1420 drinkers and 3590 nondrinkers in discovery, followed by a de novo replication analysis comprising 4896 drinkers and 13,293 nondrinkers. DNA samples of the subjects were collected for genotyping.

Results: The association results of drinking behavior (drinkers or nondrinkers) showed a cluster of single nucleotide polymorphisms at 12q24 in discovery ($P < 5 \times 10^{-8}$), with the strongest association for rs11066280 near C12orf51 ($P$-combined = $3.26 \times 10^{-215}$). Moreover, we observed the association with drinking behavior for a functional variant in ALDH2 at 12q24 (rs671, $P$-discovery = $5.17 \times 10^{-35}$). We also identified the association between rs11066280 and daily alcohol intake among drinkers ($P$-combined = $4.01 \times 10^{-21}$).

Conclusion: Our data indicate that common variants at 12q24 may contribute to the susceptibility of drinking behavior in Han Chinese. Am J Clin Nutr 2013;97:545–51.

INTRODUCTION

The harmful use of alcohol results in ~2.5 million deaths worldwide each year (1), and alcohol consumption has been identified as an avoidable risk factor for chronic disease and injury, especially in middle-income countries (1, 2). In China, there has been a rapid increase in the consumption of alcohol, and the overall prevalence of current drinking was recently reported to be 35.7% in adults aged 15–69 y (3).

Although alcohol drinking can be largely influenced by nongenetic factors, such as occupation, religion, and familial environment, there is also an important genetic component in population variance of alcohol consumption, with a heritability of ~50% (4). Previous findings from candidate-gene approaches suggested that many genes may play a role in alcohol drinking and in alcohol dependence (5–7). In contrast with alcohol dependence, which has been investigated in several genome-wide association studies (GWASs)4 to identify genetic loci (8–11), genetic susceptibility to alcohol consumption in general populations has not been explored widely. Recently, a GWAS in male Korean drinkers identified loci at 12q24 related to daily alcohol intake (12), whereas another GWAS in Japanese confirmed ALDH2 at 12q24 for drinking behavior (drinkers or nondrinkers) (13). In European ancestry, AUTS2 at 7q11 was newly identified with daily alcohol intake (14). However, environmental factors such as lifestyles vary across different population groups, and genetic heterogeneity may influence the susceptibility of alcohol drinking; thus, GWASs conducted in different ethnic groups will improve our understanding of genetic mechanism of drinking behavior. Here, we performed a 2-stage GWAS to identify genetic loci for drinking behavior in Han Chinese drinkers and nondrinkers, which was followed by an examination of genetic association with daily alcohol intake among drinkers.

1 From State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center of Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China (XY, XL, Laiyuan Wang, SC, JL, J Cao, J Chen, YH, YL, LZ, DL, QZ, JH, and D Gu); the National Human Genome Center at Beijing, Beijing, China (Laiyuan Wang and HL); the Key Laboratory for Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Sciences and Technology, Wuhan, China (Liguig Wang); the Cardio-Cerebrovascular Control and Research Center, Institute of Basic Medicine, Shandong Academy of Medical Sciences, Jinan, China (FL); the Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, Nanjing, China (CS); the Department of Cardiology, Fujian Provincial People’s Hospital, Fuzhou, China (LY); the Center for Chronic and Noncommunicable Disease Control and Prevention, Sichuan Center of Disease Control and Prevention, Chengdu, China (XW); the Department of Internal Medicine, Xilin Red Cross Hospital, Xilin, China (XJ); the Department of Internal Medicine, Xuyian Renmin Hospital, Xuyian, China (D Gu); and the Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China (XP).

2 Supported by the National Basic Research Program of China (973 Plan) (2011CB503901) from the Ministry of Science and Technology of China, and the National Science Foundation of China (30930047).

3 Address correspondence to D Gu, State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center of Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, 167 Beilishitu, Beijing 100037, China. E-mail: gudongfeng@vip.sina.com.

4 Abbreviations used: GWAS, genome-wide association study; LD, linkage disequilibrium; SNP, single nucleotide polymorphism.

Received July 5, 2012. Accepted for publication December 7, 2012.

First published online January 30, 2013; doi: 10.3945/ajcn.112.046482.
SUBJECTS AND METHODS

Study populations

In the discovery stage, a meta-analysis of 2 independent GWASs was conducted in Han Chinese populations. The GWAS 1 screening genetic association with drinking behavior consisted of 3990 subjects (993 drinkers and 2997 nondrivers), who were from the International Collaborative Study of Cardiovascular Disease in Asia (InterASIA in China) as previously reported (15). The GWAS 2 involved 1020 individuals (427 drinkers and 593 nondrivers), who were randomly selected from a community-based survey in Beijing, China (16). All the participants aged 35–74 y completed an interview-based questionnaire that assessed demographic information, medical histories, and lifestyle behaviors, including current alcohol intake and cigarette smoking. The subjects were judged to be free of coronary artery disease by medical histories, clinical examinations, and Rose questionnaire (17). Subjects with renal or hepatic diseases were also excluded. In the replication stage, 18,189 subjects (4896 drinkers and 13,293 nondrinkers) were selected from the China Cardiovascular Health Study project, which was a population-based investigation of risk factors for cardiovascular diseases in China. Briefly, the data (anthropometric measures, blood pressure, and other risk factors based on questionnaires and blood samples) were collected through in-person interviews and examinations including. The subjects with coronary artery disease, renal disease, or hepatic disease were excluded.

This study was conducted according to the principles expressed in the Declaration of Helsinki. All participants gave written informed consent, and the study was approved by the institutional review boards of Fuwai Hospital, Chinese Academy of Medical Sciences, and Peking Union Medical College.

Phenotypes and alcohol consumption

In this study, 2 phenotypes were used: drinking behavior and daily alcohol intake. Drinking behavior was defined as drinkers or nondrinkers, which was assessed on the basis of the interviewer-administered lifestyle questionnaire (18). Briefly, participants were asked whether they drank ≥12 drinks in the previous year, and individuals who gave a positive answer were defined as drinkers, whereas those who gave a negative answer were nondrinkers. For the measurement of daily alcohol intake as a quantitative trait, drinkers were asked to specify the types (beer, liquor, wine, and rice wine) and quantities of alcohol consumed, including the number of months, the number of times per month, and the amount per time they consumed each type of alcohol in the previous year. To obtain the actual grams of daily alcohol intake from beer, liquor, wine, or rice wine, the corresponding amount of each type of alcohol per day was multiplied by the average alcohol content of each product, according to the Chinese Food Composition table (beer, 3.4%; wine, 9.3%; rice wine, 13.4%; liquor, 44.4%) (19).

Genotyping and quality control

Genomic DNA was extracted from peripheral blood leukocytes by using standard procedures. In the discovery stage, genotyping and quality-control criteria were described elsewhere (20). For GWAS 1, genotyping was performed by using the Axiom Genome-Wide CHB 1 Array Plate; 3990 samples and 613,724 autosomal single nucleotide polymorphisms (SNPs) were retained for subsequent association analyses. For GWAS 2, genotyping was performed by using the Affymetrix GeneChip Human Mapping 500K Array Set. After the same quality-control procedures were used, 1020 samples and 367,129 autosomal SNPs remained for the subsequent association analyses.

In the replication stage, rs11066280 was chosen for genotyping with the iPLEX MassARRAY platform (Sequenom). To assess genotyping reproducibility, 380 duplicate samples were genotyped, and the concordance rate was 99.7%.

Genotype imputation

Genotypes were imputed for SNPs that were not presented in the genome-wide arrays or when genotyping had failed to meet the quality-control criteria. Imputation of allele dosage of missing-genotype SNPs was carried out by using Markov Chain Haplotype software (MaCH; 21) with haplotype reference panels from the phased Han Chinese in Beijing, China (CHB), and Japanese in Tokyo, Japan (JPT), HapMap (release 22) reference data set (22). After exclusion of imputed SNPs with low quality scores (Rsq < 0.3), call rate < 0.90, minor allele frequency < 0.01, and Hardy–Weinberg equilibrium $P < 1 \times 10^{-5}$, a total of ~2.2 million autosomal SNPs remained for subsequent analysis.

Statistical analysis

A quantile-quantile plot generated by using R software was used to evaluate the overall significance of the GWAS results and the potential effect of population stratification. The genomic inflation factor ($\lambda$) was estimated from the median of the chi-square statistic divided by 0.456 (23).

For a qualitative trait, SNPs were tested for the association with drinking behavior (drinkers compared with nondrinkers) in each GWAS panel and replication panel by using logistic regression analyses after adjustment for age, sex, and smoking status. For a quantitative trait, the average daily alcohol intake was transformed by the natural logarithm, and then multivariate linear regression analyses were conducted with adjustment for age, sex, and smoking status. All the association tests were performed by using PLINK v1.07 (24). A fixed-effect model with inverse variance weighting by METAL software (25) was used to perform the meta-analysis of association results from the 2 GWAS in discovery and those in replication stage.

A conditional analysis was performed to test the independence of significant SNPs in the region of 12q24 conditioning on rs11066280. The analysis was carried out by using PLINK with the –logistic and –condition options.

RESULTS

The characteristics of the participants are summarized in Table 1. In the discovery stage, we performed a meta-analysis of 2 GWAS of drinking behavior with ~2.2 million imputed and genotyped SNPs, consisting of a total of 1420 drinkers and 3590 nondrinkers. The value of the genomic inflation factor was 1.022 for the discovery meta-analysis, which indicated that population stratification effects were negligible in our study samples.
The Manhattan plot of initial meta-analysis for drinking behavior illustrated that there was a prominent cluster of SNPs at 12q24 over the genome-wide significance level \( (P < 5 \times 10^{-8}) \), whereas none of SNPs in other chromosomes reached the significance level (Figure 1). Furthermore, a quantile-quantile plot of the observed \( P \) values showed a clear deviation at the tail of the distribution from the expected distribution under the null hypothesis of no association. After exclusion of 493 SNPs within the region of 12q24 that were associated with drinking behavior at \( P < 0.05 \), the distribution of observed \( P \) values only slightly deviated from those expected in the extreme tail (see Supplemental Figure 1 under “Supplemental data” in the online issue), which suggested that the observed results probably reflected a true genetic association, and additional drinking-related loci remained to be uncovered in the Chinese population.

A total of 279 SNPs that all clustered at 12q24 showed significant associations with drinking behavior (\( P \) values from \( 4.33 \times 10^{-8} \) to \( 1.95 \times 10^{-47} \)) (Table 2). Detailed findings from the region of interest at 12q24 showed that multiple genes harbored extremely significant signals, including rs2074356 in C12orf51, rs4646776 and rs671 in ALDH2, and rs3782886 in BRAP, all of which exceeded the significance level of \( P < 1.0 \times 10^{-30} \) (Figure 2 and see Supplemental Table 1 under “Supplemental data” in the online issue). For instance, the ALDH2 gene encodes the enzyme in alcohol metabolism, and rs671 in ALDH2 was associated with drinking behavior (OR: 0.26; \( P = 5.17 \times 10^{-35} \)). In linkage disequilibrium (LD) analysis, these SNPs were almost in perfect LD with the proxy SNP rs11066280 (\( r^2 = 0.95 \) \( \pm 1.00 \); see Supplemental Table 2 and Supplemental Figure 2 under “Supplemental data” in the online issue). Furthermore, to test the independence of these SNPs for the association with drinking behavior, we conducted association analyses conditioning on the genetic effect of rs11066280. As a result, most of the other association signals on 12q24 were abolished or the strength of association became substantially diminished after conditional analysis (\( P > 0.005 \)), which indicated that these

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Characteristics of the study populations in GWASs and replication (^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GWAS 1 (( n = 3990 ))</td>
</tr>
<tr>
<td>Sample (( n ))</td>
<td>Drinker</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male (( n ))</td>
<td>929</td>
</tr>
<tr>
<td>Female (( n ))</td>
<td>64</td>
</tr>
<tr>
<td>Male (%)</td>
<td>94</td>
</tr>
<tr>
<td>Smoking (%)</td>
<td>69</td>
</tr>
<tr>
<td>Age (y)</td>
<td>51.5 ± 7.5(^2)</td>
</tr>
<tr>
<td>BMI (kg/m(^2))</td>
<td>25.2 ± 3.6</td>
</tr>
</tbody>
</table>

\(^1\) GWAS, genome-wide association study.
\(^2\) Mean ± SD (all such values).

FIGURE 1. Manhattan plot for meta-analysis of 2 genome-wide association studies for 1420 drinkers and 3590 nondrinkers. The x axis represents the genome in physical order; the y axis shows –log\(_{10}\)\( P \) for all single nucleotide polymorphisms. Each chromosome is depicted in a different color. Chr, chromosome.
TABLE 2

<table>
<thead>
<tr>
<th>Samples</th>
<th>SNP2</th>
<th>Chr</th>
<th>Position</th>
<th>Minor allele/major allele</th>
<th>Discovery (GWAS 1 + GWAS 2)</th>
<th>Replication (discovery + replication)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>rs11066280</td>
<td>12</td>
<td>111302166</td>
<td>A/T</td>
<td>0.16 (0.25, 0.30)</td>
<td>0.28 (0.25, 0.30)</td>
</tr>
<tr>
<td>Male</td>
<td>rs11066280</td>
<td>12</td>
<td>111302166</td>
<td>A/T</td>
<td>0.16 (0.25, 0.30)</td>
<td>0.27 (0.26, 0.28)</td>
</tr>
<tr>
<td>Female</td>
<td>rs11066280</td>
<td>12</td>
<td>111302166</td>
<td>A/T</td>
<td>0.15 (0.26, 0.30)</td>
<td>0.27 (0.26, 0.30)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A1/A2, minor allele/major allele</th>
<th>Chr</th>
<th>Position</th>
<th>Nearby gene</th>
<th>A1/A2 MAF</th>
<th>Combined OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11066280</td>
<td>12</td>
<td>111302166</td>
<td>C12orf51</td>
<td>0.16</td>
<td>0.28 (0.25, 0.30)</td>
<td>1.95</td>
</tr>
</tbody>
</table>

**DISCUSSION**

The current study was the first GWAS published to date in the Han Chinese population to investigate genetic susceptibility to alcohol consumption; the study consisted of 5010 individuals in the GWAS scan with replication in an additional 18,189 individuals. The significant signal at 12q24 was observed in the discovery stage, and rs11066280 near C12orf51 was successfully replicated for drinking behavior and daily alcohol intake in an independent population.

A recent GWAS in Korean men, involving a total of 2834 drinkers, reported 8 SNPs in or near C12orf51, CCDC63, and MYL2 at 12q24 that highly associated with daily alcohol intake, and rs11066280 was one of these SNPs ($P = 9.34 \times 10^{-54}$) (12). We checked all of the reported 8 SNPs in our discovery data for both drinking behavior (including drinkers and nondrinkers) and daily alcohol intake (among drinkers only) (see Supplemental Table 3 under “Supplemental data” in the online issue). Besides the associations with alcohol intake supported by our GWAS analysis ($P < 0.05$), we provided new highly significant evidence of associations with drinking behavior at the 8 SNPs ($P = 1.01 \times 10^{-11} - 1.95 \times 10^{-15}$). The GWAS in the Korean population mentioned that SNPs in CCDC63 and MYL2 associations were not independent of rs11066280 (Figure 2); rs11066280 is located to the 74 kb upstream of C12orf51, and the function of C12orf51 has not been elucidated.

We further validated the association between rs11066280 and drinking behavior in an independent population for replication, and the association remained significant (OR: 0.31; $P = 5.01 \times 10^{-170}$). When we combined the results of the discovery and replication stages, the evidence of C12orf51 associated with drinking behavior became much stronger (OR: 0.30; $P = 3.26 \times 10^{-215}$) (Table 2).

In addition, to examine whether the association was consistent for sex, we divided the combined population into a male population (6074 drinkers and 7390 nondrinkers) and a female population (242 drinkers and 9493 nondrinkers). The association between rs11066280 and drinking behavior remained significant in the male subjects ($P = 3.05 \times 10^{-210}$). Although the sample size of female drinkers was small, so that it was difficult to detect the association because of insufficient power, rs11066280 showed nearly genome-wide significant association with drinking behavior ($P = 2.61 \times 10^{-7}$). Moreover, the marker had similar effects in the same direction for the male and female populations (OR: 0.30 in male subjects; OR: 0.43 in female subjects; Table 2), which suggests that rs11066280 near C12orf51 might be the genetic locus for drinking behavior in both men and women.

To verify the genetic effects of rs11066280 on daily alcohol intake, we performed association analyses with the amount of alcohol intake per day among drinkers only. As the minor allele dosage increased (A-allele for rs11066280), a clear decrease in daily alcohol intake was observed in discovery and was successfully validated in replication. The effect sizes were consistent in the same direction, and $P$ values were $1.10 \times 10^{-5}$ in discovery, $7.80 \times 10^{-17}$ in replication, and $4.01 \times 10^{-21}$ in combined data (Table 3). These results strengthened the evidence of the susceptibility locus at 12q24 for drinking behavior in the Chinese population.
were in a different LD block far from the region of C12orf51. Nevertheless, after conditional analysis of rs11066280, the associations for drinking behavior were obviously attenuated at the rest of 7 SNPs, which showed low or moderate LD with rs11066280 ($r^2 = 0.25 \sim 0.58$) in our discovery data (see Supplemental Table 4 under “Supplemental data” in the online issue). It suggested that SNPs in or near CCDC63 and MYL2 were not totally independent of rs11066280 in the Chinese population.

In addition, one GWAS performed in the Japanese population confirmed ALDH2 as a major locus regulating drinking behavior (13), and rs671 in the ALDH2 gene showed the strongest association between 4780 drinkers and 1940 nondrinkers in their joint analysis (OR: 0.16; $P = 3.6 \times 10^{-211}$). The 504lys (A) allele of rs671 has been known to be a functional variant in ALDH2, lysine at codon position 504 instead of glutamic acid, leading to a catalytically inactive isozyme of aldehyde dehydrogenase (26, 27). Individuals homozygous for the A allele have a distinctly reduced capacity to clear acetaldehyde and show facial flushing and experience nausea after alcohol consumption (27). In our discovery data, we observed that rs671 was associated with drinking behavior (see Supplemental Table 1 under “Supplemental data” in the online issue) and was suggestively associated with daily alcohol intake ($P = 6.37 \times 10^{-3}$); rs11066280 near C12orf51 and rs671 in ALDH2 was in high LD with the pairwise $r^2 = 0.95$, which suggests that the signal at 12q24 might originate from the region of ALDH2. However, considering the complexity of biomechanisms in regulating drinking behavior, we could not exclude the possibility that C12orf51 might influence drinking behavior through other mechanisms independent of the acetaldehyde metabolism pathway. Further fine mapping and functional analyses are necessary to investigate the causal variant in the region of 12q24.

**FIGURE 2.** Regional association plots of genome-wide association studies on 12q24 before and after conditional analysis. Top panel: single nucleotide polymorphisms are plotted according to their chromosomal positions (National Center for Biotechnology Information build 36) with their $P$ values (as $-\log_{10}P$ values) from the meta-analysis results in discovery stage. The top 5 single nucleotide polymorphisms ($-\log_{10}P > 30$) are represented by triangles, and other single nucleotide polymorphisms are represented by circles. Middle panel: the meta-analysis results after conditional analysis of rs11066280. Bottom panel: genes within the ~2.0-Mb region of 12q24 are annotated and are shown as arrows.
The association of rs11066280 at 12q24 with daily alcohol intake and comparison of daily alcohol intake across genotypes

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>Position</th>
<th>Nearby gene</th>
<th>AA genotype</th>
<th>AT genotype</th>
<th>TT genotype</th>
<th>Study stage</th>
<th>Sample size</th>
<th>β ± SE</th>
<th>P</th>
<th>β ± SE</th>
<th>P</th>
<th>β ± SE</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11066280</td>
<td>12</td>
<td>11130166</td>
<td>C12orf51</td>
<td>AA</td>
<td>AT</td>
<td>TT</td>
<td>Discovery</td>
<td>1394</td>
<td>0.44 ± 0.10</td>
<td>1.10 × 10⁻⁵</td>
<td>0.27</td>
<td>0.38 ± 0.05</td>
<td>7.80 × 10⁻¹⁷</td>
<td>0.05</td>
</tr>
<tr>
<td>rs1066280</td>
<td>12</td>
<td>225376176</td>
<td>AUTS2</td>
<td>AA</td>
<td>AT</td>
<td>TT</td>
<td>Replication</td>
<td>4817</td>
<td>-0.39 ± 0.04</td>
<td>4.01 × 10⁻²¹</td>
<td>0.04</td>
<td>-0.38 ± 0.03</td>
<td>4.60 × 10⁻¹⁷</td>
<td>0.10</td>
</tr>
</tbody>
</table>

In European populations, there were few variants associated with alcohol consumption approaching the genome-wide significance (28). A large-scale meta-analysis (of >47,000 individuals of European ancestry) identified the SNP rs6943555 in AUTS2 to be associated with alcohol consumption (daily alcohol intake; \( P = 4.1 \times 10^{-9} \)) (14). Unfortunately, we could not validate the association of rs6943555 for drinking behavior including drinkers and nondrinkers (\( P = 0.27 \)) nor for daily alcohol intake only among drinkers (\( P = 0.33 \)). The minor allele frequency of rs6943555 was a little higher in our discovery samples (0.38) than that reported in Europeans (0.24), whereas the common variant rs11066280 associated with drinking behavior in Chinese was monomorphic in European ancestry, according to the HapMap CEU data. These disparities in GWAS findings suggest that genetic heterogeneity in drinking behavior might exist between European and Chinese populations.

In conclusion, the current GWAS identified SNPs at 12q24 as a signal of genetic susceptibility for drinking behavior in a Han Chinese population and suggested that the causal genetic variants associated with alcohol consumption may be located in the region of 12q24 in East Asians. Further large-scale meta-analyses of GWAS for alcohol consumption will have opportunities to find susceptibility loci with small to modest effect in East Asians and other ethnicities. The more comprehensive understanding of genetic mechanism of alcohol drinking will be helpful in individualized prevention and control of alcohol-associated diseases worldwide.

The authors’ responsibilities were as follows—D Gu: designed the research; Laiyuan Wang, SC, HL, and DL: conducted the experiments; JL, J Cao, J Chen, YL, LZ, Ligui Wang, FL, CS, LY, WX, XI, QZ, D Guo, XP, and JH: provided reagents and materials; XY, XL, and YH: conducted the statistical analyses; D Gu, XY, and XL: wrote the manuscript; and D Gu: had primary responsibility for the final content. All authors read and approved the final manuscript. None of the authors declared a conflict of interest.

REFERENCES


