Journal of Psychiatric Research 47 (2013) 1717–1724

Contents lists available at ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/psychires

Genome-wide association studies of maximum number of drinks

Yue Pan [a,b], Xingguang Luo [c], Xuefeng Liu [a], Long-Yang Wu [d], Qunyuan Zhang [e], Liang Wang [a], Weize Wang [b], Lingjun Zuo [c], Ke-Sheng Wang [a,*]

[a] Department of Biostatistics and Epidemiology, College of Public Health, East Tennessee State University, Johnson City, TN, USA
[b] Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, FL, USA
[c] Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA
[d] The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA
[e] Division of Statistical Genomics, Washington University School of Medicine, St. Louis, MO, USA

ARTICLE INFO

Article history:
Received 16 May 2013
Received in revised form
25 June 2013
Accepted 18 July 2013

Keywords:
Maximum number of drinks
Genome-wide association
Meta-analysis
SGOL1
DTWD2
NDST4
KCNB2
DDC

ABSTRACT

Maximum number of drinks (MaxDrinks) defined as "Maximum number of alcoholic drinks consumed in a 24-h period" is an intermediate phenotype that is closely related to alcohol dependence (AD). Family, twin and adoption studies have shown that the heritability of MaxDrinks is approximately 0.5. We conducted the first genome-wide association (GWA) study and meta-analysis of MaxDrinks as a continuous phenotype. 1059 individuals were from the Collaborative Study on the Genetics of Alcoholism (COGA) sample and 1628 individuals were from the Study of Addiction – Genetics and Environment (SAGE) sample. Family sample with 3137 individuals was from the Australian twin-family study of alcohol use disorder (OZALC). Two population-based Caucasian samples (COGA and SAGE) with 1 million single-nucleotide polymorphisms (SNPs) were used for gene discovery and one family-based Caucasian sample was used for replication. Through meta-analysis we identified 162 SNPs associated with MaxDirnks ($p < 10^{-4}$). The most significant association with MaxDrinks was observed with SNP rs11128951 ($p = 4.27 \times 10^{-8}$) near SGOL1 gene at 3p24.3. Furthermore, several SNPs (rs17144687 near DTWD2, rs12108602 near NDST4, and rs2128158 in KCNB2) showed significant associations with MaxDrinks ($p < 5 \times 10^{-7}$) in the meta-analysis. Especially, 8 SNPs in DDC gene showed significant associations with MaxDrinks ($p < 5 \times 10^{-7}$) in the SAGE sample. Several flanking SNPs in above genes/regions were confirmed in the OZALC family sample. In conclusions, we identified several genes/regions associated with MaxDrinks. These findings can improve the understanding about the pathogenesis of alcohol consumption phenotypes and alcohol-related disorders.

1. Introduction

Maximum number of drinks (MaxDrinks) defined as "Maximum number of drinks consumed in a 24-h period" is an alcoholism-related phenotype that could be a surrogate to alcohol dependence (AD) and a quantitative measure to grade non-alcoholic individuals (Bierut et al., 2002). There have been an increasing number of reports on binges, alcohol-related life problems such as physiological complications, alcohol-related emotional/psychiatric symptoms in the groups with larger maximum number of drinks (Schuckit et al., 1998). Family, twin and adoption studies have shown that the heritability of MaxDrinks (Saccone et al., 2000) is approximately 0.5. Studies of the genetic basis of MaxDrinks can provide more information for the understanding of AD, and enhance the development of efficient prevention strategies and personalized treatments. Saccone et al. (2000) conducted genome-wide linkage analysis of MaxDrinks and detected linkage region of the alcohol dehydrogenase gene cluster on chromosome 4 (LOD = 3.5). Furthermore, Saccone et al. (2005) detected linkage to chromosomes 2 and 7 using a two-stage method. de Andrade et al. (2005) identified suggestive linkage on chromosomes 1, 4, 10 and 13. Linkage for MaxDrinks was also detected on chromosomes 7 (Chen et al., 2005; Saccone et al., 2005). In an Irish affected sib pair study, Kuo et al. (2006) reported that MaxDrinks was associated with regions on chromosomes 12 and 18.

---

[*] Corresponding author. Department of Biostatistics and Epidemiology, College of Public Health, East Tennessee State University, PO Box 70259, Lamb Hall, Johnson City, TN 37614-1700, USA. Tel.: +1 423 439 4481; fax: +1 423 439 4606.
E-mail address: wangk@etsu.edu (K.-S. Wang).

0022-3956/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jpsychires.2013.07.013

The genome-wide association (GWA) study has been successfully used as an important tool for identifying regions of human genome that are associated with more than 40 different common diseases. This approach has provided new insights into pathophysiology and suggested previously unsuspected etiologic pathways for common diseases that could be of use in identifying new therapeutic targets and in developing targeted interventions based on genetically defined risk (Manolio et al., 2008). Recently, there are several completed GWA studies for alcohol dependence (Bierut et al., 2010; Edenberg et al., 2010; Lind et al., 2010; Treutlein and Rietschel, 2011). However, based on our knowledge, no GWA study has been conducted on MaxDrinks as a quantitative phenotype in the literature. In this study, we performed a meta-analysis using two genome-wide data to detect genetic variants that may influence MaxDrinks in both Caucasian samples. We also used the dataset from The Australian twin-family study of alcohol use disorder (OZALC study) for replication.

2. Materials and methods

2.1. Study samples

2.1.1. The COGA sample

The Collaborative Study on the Genetics of Alcoholism (COGA) Case Control Study is a case–control GWA study of AD. It contains about 1 million Illumina SNPs (1,069,796 SNPs), and 1234 cases with AD and 711 controls (Edenberg et al., 2010). Phenotypes include AD as a binary trait according to DSM-IV diagnosis. Besides, another quantitative and heritable phenotype MaxDrinks that measured the maximum number of drinks a person has consumed in a 24-h period has been included in the study which increased the study power (Edenberg, 2002). In the present study, we chose 1059 Caucasian (non-Hispanic) individuals (572 males and 487 females) with MaxDrinks.

2.1.2. The SAGE sample

The Study of Addiction – Genetics and Environment (SAGE) is a comprehensive GWA study using approximately 4000 unrelated subjects of European and African-American descent. Cases with AD include 1944 subjects with the primary phenotype having been DSM-IV AD (Bierut et al., 2010). Controls consist of 1965 subjects who have used alcohol, but have never been addicted to alcohol or other illicit substances. In order to screen for substance dependence, MaxDrinks were asked to potential control subjects. In the present study, we used 1628 Caucasian individuals with MaxDrinks (629 males and 999 females) from the combined data of the Family Study of Cocaine Dependence (FSCD), and the Collaborative Genetic Study of Nicotine Dependence (COGEND) in the SAGE study. It contains about 1 million Illumina SNPs (1,069,796 SNPs).

2.1.3. The OZALC sample

The Australian twin-family study of alcohol use disorder (OZALC study) derives from telephone diagnostic interview studies of two general population volunteer cohorts of Australian twins (cohort 1, mostly born 1940–1964; cohort 2, born 1964–1971) and the spouses of the former cohort – a total of over 11,000 families. The data used in the present study is from the publicly available data from the Genome wide Association Study of Alcohol Use and Alcohol Use Disorder in Australian Twin-Families (OZALC GWAS) – Study Accession: phs000181.v1.p1. The details about these subjects were described elsewhere (Grant et al., 2009; Lind et al., 2010). Genotyping data using the ILLUMINA HumanCNV370v1 (total 343,955 SNPs) are available for 4119 individuals in this dataset. After merging with pedigree and phenotypes, we removed one from each of 44 MZ twins and 72 outliers based on the data description, and 669 individuals with unknown case status. Consequently, there were 3137 individuals with MaxDrinks left for further analysis (1708 males and 1429 females).

The basic characteristics of the subjects in these 3 samples are presented in Table 1.

2.2. Statistical analyses

2.2.1. Genome-wide association analysis

For the initial GWA analysis, HelixTree Software (http://www.goldenhelix.com/SNP_Variation/HelixTree/index.html) was used to assess control genotype data for conformity with Hardy–Weinberg equilibrium (HWE). To test for association with MaxDrinks as a quantitative trait, linear regression, adjusted for age and sex, was performed by PLINK 1.07 (Purcell et al., 2007) to obtain the regression coefficient and standard error as well as Wald test asymptotic *p*-value. For statistical significance, we used a significance level of $\alpha = 5 \times 10^{-7}$ (Wellcome Trust Case Control Consortium, 2007). At the same time, we also used a less stringent criterion of “suggestive association” with a cut-off of $\alpha = 10^{-4}$. In addition to obtaining nominal *p*-values, empirical *p*-values were generated by 100,000 permutation tests using the Max (T) permutation procedure implemented in PLINK. In this procedure, two sets of empirical significance values were calculated: pointwise estimates of an individual SNP’s significance (empirical pointwise *p*-values) and corrected values for multiple testing (corrected empirical *p*-values).

2.2.2. Meta-analysis

The COGA and SAGE samples used the same genotyping platform: Illumina Human 1M BeadChips (both datasets have 1,069,796 SNPs). Results from the two GWA analyses were meta-analyzed by combining the separate results of COGA and SAGE samples (regression coefficient and standard error) into one meta-analysis of overall effects. For meta-analysis of two datasets, the basic meta-analysis function in PLINK was applied. The between-study heterogeneity was tested by the χ^2-based Cochrane’s *Q* statistic. Fixed-effect model was used due to the homogeneity *Q* statistic and its associated *p*-value was reported to provide support for this decision. Fixed-effect meta-analysis *p*-values and fixed-effect ORs were estimated.

2.2.3. Family-based association analysis in the OZALC sample

In this study, family-based association analysis was performed in the OZALC sample by using the PBAT v3.61 (Hoffmann and Lange, 2006), which can handle nuclear families, as well as extended pedigrees. For family-based association analysis, the additive model was applied. To deal with the multiple testing in the family-based association study, we used software QVALUE (http://genomics.princeton.edu/storeylab/qvalue/) to calculate the false discovery rate (FDR) (Storey, 2002).

Table 1
Descriptive characteristics of the subjects with MaxDrinks.

Parameter	COGA	SAGE	OZALC
Number	1059	1628	3137
Women (%)	487 (46)	999 (61)	1429 (46)
Age (years)			
Mean ± SD	40.8 ± 11.4	35.7 ± 7.7	42.9 ± 8.2
MaxDrinks			
Mean ± SD	23.9 ± 21.1	17.5 ± 17.7	21.0 ± 14.9