Gene 593 (2016) 315–321

Contents lists available at ScienceDirect

Gene

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

Research paper

Multi-variant study of obesity risk genes in African Americans: The Jackson Heart Study

Shijian Liu [a], James G. Wilson [b], Fan Jiang [a], Michael Griswold [c], Adolfo Correa [d], Hao Mei [a,c,*]

[a] Shanghai Children's Medical Center, School of Public Health, Shanghai Jiaotong University School of Medicine, Shanghai 200127, China
[b] Physiology & Biophysics, University of Mississippi Medical Center, Jackson, MS 39216, USA
[c] Department of Data Science, University of Mississippi Medical Center, Jackson, MS 39216, USA
[d] Jackson Heart Study, University of Mississippi Medical Center, Jackson, MS 39216, USA

ARTICLE INFO

Article history:
Received 7 June 2016
Received in revised form 28 July 2016
Accepted 25 August 2016
Available online 26 August 2016

Keywords:
Gene-gene interaction
SNP association
Genetic risk score
Haplotype
Obesity

ABSTRACT

Objective: Genome-wide association study (GWAS) has been successful in identifying obesity risk genes by single-variant association analysis. For this study, we designed steps of analysis strategy and aimed to identify multi-variant effects on obesity risk among candidate genes.
Methods: Our analyses were focused on 2137 African American participants with body mass index measured in the Jackson Heart Study and 657 common single nucleotide polymorphisms (SNPs) genotyped at 8 GWAS-identified obesity risk genes.
Results: Single-variant association test showed that no SNPs reached significance after multiple testing adjustment. The following gene-gene interaction analysis, which was focused on SNPs with unadjusted p-value < 0.10, identified 6 significant multi-variant associations. Logistic regression showed that SNPs in these associations did not have significant linear interactions; examination of genetic risk score evidenced that 4 multi-variant associations had significant additive effects of risk SNPs; and haplotype association test presented that all multi-variant associations contained one or several combinations of particular alleles or haplotypes, associated with increased obesity risk.
Conclusions: Our study evidenced that obesity risk genes generated multi-variant effects, which can be additive or non-linear interactions, and multi-variant study is an important supplement to existing GWAS for understanding genetic effects of obesity risk genes.

1. Introduction

Obesity is a risk factor of many common diseases, such as hypertension, type 2 diabetes mellitus, cardiovascular disease and sleep apnea. Over the past decades, obesity prevalence continues to increase, and more than one-third of adults and 47.8% of non-Hispanic African American adults are obese in the USA (Ogden et al., 2014). High prevalence of obesity is becoming a major public health issue and the impact is substantial. The cause of obesity is partly attributed to unhealthy life styles including excessive calorie intake and lower physical activity (Catenacci et al., 2009). However, not everyone who leads a sedentary life-style or over consumes food becomes obese, and the predisposition to obesity depends on genetic components.

Twin and family heritability studies have indicated that up to 80% of the variance in obesity, as measured by BMI, can be explained by genetic factors (Silventoinen and Kaprio, 2009; Hasselbalch, 2010; Iranzo-Tatay et al., 2015). Genome-wide association studies (GWAS) by single-variant analysis have succeeded in identifying obesity risk variants from different genes, including fat mass and obesity associated (FTO), neuronal growth regulator 1 (NEGR1), neurexin 3 (NRXN3), transmembrane protein 18 (TMEM18), transcription factor AP-2 beta (TFAP2B) and melanocortin 4 receptor (MC4R) (Welter et al., 2014a). However, risk variants identified so far have tiny effects, widely known as the missing

Abbreviations: GWAS, genome-wide association study; SNPs, single nucleotide polymorphisms; JHS, Jackson Heart Study; BMI, body mass index; MDR, multifactor dimensionality reduction; SE, standard error; MAF, minor allele frequency; HWE, Hardy-Weinberg equilibrium; LD, linkage disequilibrium; PC, principal component; PE, prediction error; GRS, genetic risk score; FTO, fat mass and obesity associated; NEGR1, neuronal growth regulator 1; NRXN3, neurexin 3; TMEM18, transmembrane protein 18; TFAP2B, transcription factor AP-2 beta; MC4R, melanocortin 4 receptor; ADCY3, adenylate cyclase 3; BDNF, brain-derived neurotrophic factor; CHRNA3, cholinergic receptor, nicotinic, alpha 3; GIPR, gastric inhibitory polypeptide receptor.

* Corresponding author at: Shanghai Children's Medical Center, School of Public Health, Shanghai Jiaotong University School of Medicine, Shanghai, 200127, China; Department of Data Science, University of Mississippi Medical Center, 2500 North State Street, Jackson, MS 39216, USA.
E-mail addresses: arrow64@163.com (S. Liu), jgwilson2@umc.edu (J.G. Wilson), fanjiang@shsmu.edu.cn (F. Jiang), mgriswold@umc.edu (M. Griswold), ACorrea@umc.edu (A. Correa), hmei@umc.edu (H. Mei).

http://dx.doi.org/10.1016/j.gene.2016.08.041
0378-1119/

heritability (Manolio et al., 2009), and it is often a challenge to detect these variants following GWAS. In contrast to a single-variant analysis, study of multi-variant effects among GWAS-identified obesity genes can help to improve test of risk variants while individual variants may not have detectable effects in a sample.

Multi-variant effects, especially gene-gene interactions, are considered as important components of the genetic architecture influencing susceptibility to common human diseases (Culverhouse et al., 2002). Investigation of obesity risk genes reported by GWAS evidenced that SNPs with effects that are weak or absent alone can jointly exert a large longitudinal effect on body mass index (BMI) from childhood to adulthood (Mei et al., 2012). Multi-variant effects are complicated, and gene-gene interaction study by multifactor dimensionality reduction (MDR) showed that genetic risk variants can jointly present additive effects, linear or nonlinear interactions in different forms (Hahn et al., 2003; Greene et al., 2010). Study of obesity risk genes and examination of their multi-variant effects will be an important supplement to existing GWAS and help to improve understanding of genetic influences on obesity development.

In the USA, African-Americans are more obese than European-American (Lewis et al., 2000). For this study, we investigated African-American sample of the Jackson Heart Study (JHS) cohort (Fuqua et al., 2005), and designed steps of analysis strategy that emphasized on test of multi-variant associations among obesity risk genes reported by previous GWAS. Our purpose is to examine if these genes generate multi-variant effects on obesity risk and measure their sizes by deliberately designed analysis strategy.

2. Materials and methods

2.1. Study participants

The JHS is a large community-based observational study of African Americans, and participants were recruited from urban and rural areas of the three counties of Hinds, Madison and Rankin in the Jackson, MS (Fuqua et al., 2005). Weight of participant was measured on a balance scale, in light clothing and without shoes, and recorded to the nearest 0.5 kg. The body mass index (BMI), as a measure of obesity, was calculated as weight in kilograms divided by height in meters squared. Obesity status was defined according to BMI (Bidulescu et al., 2011). Individuals with BMI < 30 kg/m^2 were considered as control, and those with BMI ≥ 30 kg/m^2 were considered as obese cases. The study protocol was approved by the Institutional Review Boards at the University of Mississippi Medical Center.

2.2. SNP genotyping

The Affymetrix Human SNP Array 6.0 (Affymetrix® Inc., Santa Clara, CA) was used for genome-wide SNP genotyping. Genomic DNA was quantitated and DNA quality was evaluated via gel electrophoresis. The genomic DNA samples were processed according to standard Affymetrix procedures for processing of the assay. The data were processed for genotype calling using the Affymetrix Genotypic Console software using the Birdseed calling algorithm version 2.0 (Affymetrix®, Inc., Santa Clara, CA) (Korn et al., 2008).

2.3. Obesity risk genes and quality control

Obesity risk genes were identified by searching against the Catalog of Published Genome-Wide Association Studies (www.genome.gov/gwastudies/) (Welter et al., 2014b), limited to diseases/trait categories of obesity and related traits. Top reported genes that had SNPs with association p-value ≤ 5×10^{-8} were selected as candidate genes for multi-variant study. Genomic map positions of candidate genes were identified based on Human NCBI36/hg18 Assembly. SNPs within upstream 5 kb to downstream 5 kb of candidate genes were extracted from JHS GWAS data. PLINK 1.90 (Purcell et al., 2007) was applied to calculate individual missing rate, genotype missing rate and minor allele frequency (MAF), to test Hardy-Weinberg equilibrium (HWE), and to measure pairwise linkage disequilibrium (LD). The criteria for SNP quality control included the genotype missing rate < 5%, minor allele frequency (MAF) ≥ 5% and p-value of Hardy-Weinberg equilibrium (HWE) > 0.001.

2.4. Statistical analysis

Demographic variables were analyzed with age, gender and BMI. The *t* and Chi-square test were respectively applied to examine distribution difference of continuous and category variables. The analysis was performed using Stata software (version 13.0; College Station, TX, USA), and p-values were obtained based on two-sided test.

We analyzed SNPs of JHS GWAS data, and the first 10 principal components (PCs) were calculated from the pruned genome-wide autosomal SNPs with MAF ≥ 0.05 and pair-wise $r^2 \leq 0.1$ by the EIGENSTRAT method (Price et al., 2006) , which are used as covariates to adjust for potential population structure. For every genotyped SNP of candidate genes, additive genetic model was adopted to test single-variant association with obesity by logistic regression, adjusting for age, gender and PCs. The t-statistic and odds ratio were computed with p-values obtained through a two-sided test, and potential risk SNPs were identified if they had unadjusted p-value < 0.10. The analysis was performed using PLINK software.

Potential risk SNPs were further examined and those with pairwise $r^2 \leq 0.2$ based on HapMap reference data (Yoruba in Ibadan, Nigeria) (International HapMap Consortium, 2003) were tested for gene-gene interactions by method of multifactor dimensionality reduction, MDR-Phenomics (Ritchie et al., 2001; Mei et al., 2007). MDR is a model-free approach that tests high-order gene x gene interactions without assuming any particular genetic model and estimating parameters. The method reduces dimensionality of multi-variant genotype information to a one-dimensional factor with two levels—"high-risk" and "low-risk" genotypes. The MDR analysis aimed to identify multi-variant effects and the test was based on the hypothesis that individual SNPs with weak effects, which may not be detectable in a single-variant study, can jointly exert a large effect on obesity risk. We applied the MDR method to test all 2- to 5-variant associations among the candidate SNPs that were trained and tested by 10–1 cross-validation, and measured by prediction error (PE). Association p-values, adjusted for multiple testing, were obtained by 1000 permutation tests, and the significance was defined as adjusted p-value ≤ 0.05.

For significant multi-variant associations, linear interactions between SNPs were examined by logistic regression, and additive effects of SNPs were investigated through test of genetic risk score (GRS) association. A SNP risk allele was the one that had increased odds of obesity at the single-variant logistic regression analysis, and GRS was computed as the total number of risk alleles in the multi-variant association (Cornelis et al., 2009). Association of GRS with obesity was tested by logistic regression among men, women and total sample separately.

Significant multi-variant associations were lastly explored to investigate effects for combinations of particular alleles or haplotypes across SNPs by haplotype omnibus, conditional and specific analysis using PLINK: logistic regression with omnibus test was applied to examine overall haplotype association with obesity; independent SNP effects in a haplotype association were evaluated by haplotype conditional test and effect of a particular haplotype or combination of SNP alleles was measured by the specific analysis. For both the GRS and haplotype association tests, the Bonferroni correction was applied to adjust for multiple testing, and the threshold of significant p-value was defined as 0.05 divided by the number of significant multi-variant associations identified by the MDR analysis.