ISSN: 2790-5209 DOI: 10.56028/fesr.1.1.21

Large-scale GWAS meta-analysis consisting of trans-ethnic samples identifies various genetic signals on BMI

Yiyun Chen[1,*] Zhenxiao Xu[2] Anjia Zhao[1]

[1] Unionville High school

[2] Suzhou No.10 high school of Jiangsu Province

*nicole.chen622@gmail.com

Abstract

*Due to the development of computational power and statistical theories, Genome-wide association studies (GWAS) have constantly been improved to gain higher power with reduced bias. GWAS identify hundreds of susceptibility loci body mass index in various populations such as European-ancestry, or Asian groups. Meta-analysis enables us to incorporate statistical results from various studies to detect more genetics signals in GWAS, as well as discover different signals from cis- or trans-ethnic groups. Here we combined data from three sources of large-scale genetics studies: UK Biobank, GIANT consortium, and a famous Japanese study. Among over two million candidate SNPs, we successfully detected 686 significant SNPs after Bonferroni correction (P < 2.5*10^-8), with most of them being detected previously. The top five SNPs are: "rs1558902" (P value = 2.394*10^-36), "rs1421085" (P value = 4.152*10^-36), "rs2237897" (P value = 2.542*10^-32), "rs2237896" (P value = 3.966*10^-32), "rs7202116" (P value = 2.702*10^-31). Although the total number of variants identified by the meta-analysis is lower than the Japanese population-based association study, meta-analysis successfully identifies several new loci not captured by the single-group association study. We also explored the original summary statistics datasets and conducted analysis to compare the statistical results from different populations separately.*

Keywords

*GWAS , meta analysis, BMI, trans-ethnic,Large scale***.**

1. Introduction

Obesity and body weight issues have become one of the important concerns in modern population. And it has been shown that overweighting is one risk factor for various other health diseases [1]. Body weight is regarded as a heritable and complex trait in the context of human genetics. Recent genetic studies have concluded that more than 20% genetic variance of body weight trait can be attributed to genetics factors mainly represented by variants [2]. On the other hand, this means large proportions of the heritability-deciding factors remain unknown. There have been studies focusing on genetics factors on BMI in European-ancestry population [3][4][5], as well as Asian-ancestry population [6]. Different studies show that trans-ethnic population groups tend to target different genetics factors, with partial shared genetic contributions. In 2017, a large-scale GWAS consisting of more than 170,000 Japanese subjects reveals 112 new SNPs on BMI, which is a profound discovery for east Asian groups [7]. GIANT consortium also collects information on the GWAS results. Noticing that the existing meta studies mostly focus on European ancestry populations, we also consider African-origin population.

ISSN: 2790-5209 DOI: 10.56028/fesr.1.1.21
To detect genetic signals across different ethnic groups and gain insight of obesity in these populations, we combined the large-scale Japanese-based GWAS study [7], one European-ancestry-based study retrieved from GIANT consortium [5], as well as part of UK Biobank data of African-ancestry. We utilized METAL, a classic tool for combining separate GWAS analysis summary statistics [9]. METAL incorporates information across various trans-ethnic populations, and it results in little efficiency loss compared to analysis of a combined dataset including data from all individual studies.

2. Results

The information on data source, ethnic group, sample size and number of SNP hits are listed in Table 1. The total number of SNPs in the meta-analysis is 2020153. The total sample size in this meta-analysis is around 300,000. The majority groups are Japanese and European-ancestry groups in this study. And the signals in the meta-analysis are mostly driven by Japanese group signals, because the discovery in the GIANT consortium dataset and UK Biobank African subgroup is little. The low number of hits in the UK Biobank African subgroup dataset can easily be explained by the small sample size. A possible reason of the low number of hits in the GIANT consortium dataset is that they included more than fifty studies, and therefore the signals might be masked by batch effects.

Table 1: summary of datasets information

data source	ethnic group	sample size	number of SNP hits
GIANT consortium	European	~ 122000	0
[7] Akiyama et al	Japanese	173430	1464
UK Biobank	African	6545	1
Meta-analysis	combined	~ 300000	686

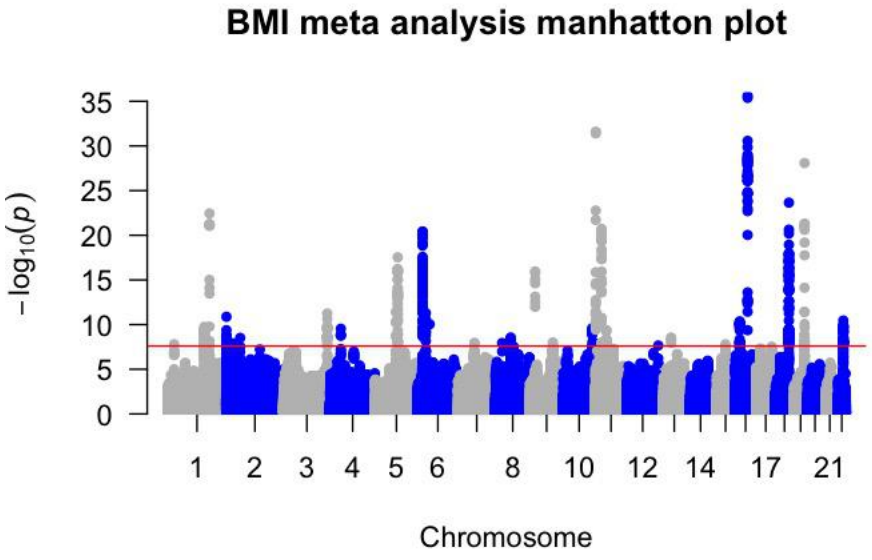

Figure 1: genome-wide Manhattan plot of the meta-analysis

ISSN: 2790-5209 DOI: 10.56028/fesr.1.1.21

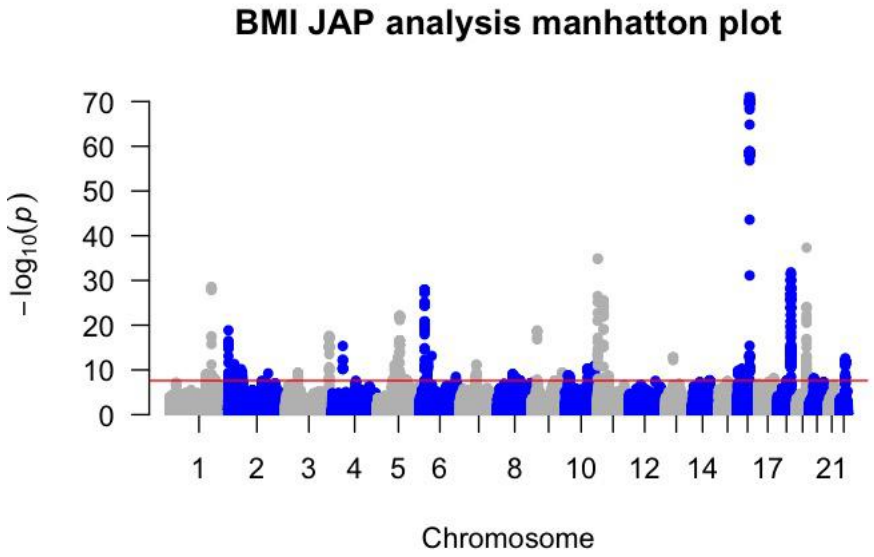

Figure 2: genome-wide Manhattan plot of the Japanese population meta-analysis

Table 2 contains information of the three top hits in the meta-analysis. Signals are concentrated in chromosome 11, 16 and 19, which can also be seen in Figure 1 and 2. There are also several neighbor SNPs with comparatively strong signals, which are the effects of genetic linkage-disequilibrium structures. These top signals are also identified by the Japanese population study. The association signal patterns for the meta-analysis and the Japanese population study are also very similar, except that the signal strength on chromosome 16 is stronger in the Japanese population study. Strong signals are present in chromosome 1, 5, 6, 11, 16, 18 and 19. Medium strong signals are present in chromosome 2, 3, 4, 12, 12.

Table 2: Top hits in the meta-analysis, excluding linkage-disequilibrium effect

RSID	chromosome	position	beta	P value	Genes
rs1558902	16	53803574	-0.0455	2.39*10^-36	FTO
rs2237897	11	2858546	-0.0421	2.54*10^-32	KCNQ1
rs11671664	19	46172278	0.0371	8.25*10^-29	GIPR

Table 3 contains information of the three top hits in the UK Biobank African ancestry-based association study. The genome-wise p value cutoff was 2.5*10^-8, and the genome suggestive cutoff was 2.5*10^-6. In African ancestry group, signals are concentrated to chromosome 6, 18 and 5. Due to the limited sample size, there is no particularly strong signals.

Table 3: Top hits in the UK Biobank African ancestry-based association study

RSID	chromosome	position	beta	P value
rs9368276	6	21101675	-0.1372	1.78*10^-8
rs1539952	18	57766512	0.1082	1.82*10^-7
rs12657213	5	85655306	0.08752	7.93*10^-7

ISSN: 2790-5209 DOI: 10.56028/fesr.1.1.21

3. Datasets and Methods

In this GWAS meta-analysis study, we focused on BMI phenotype and used three GWAS dataset result. To be more specific, the three dataset that we use are from: Akiyama et al's study [7], Yang et al's study [5] retrieved from GIANT consortium and Pan UKB team [13]. The original summary statistics can be downloaded from:

1, Japanese study: https://www.nature.com/articles/ng.3951#Ack1

2, GIANT consortium study: https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files

3, UK Biobank African ancestry study: https://pan.ukbb.broadinstitute.org/downloads

The meta-analysis was conducted using METAL [9]. METAL utilized two weighting method to combine information across different datasets: sample-size weighting or standard error weighting method. Here we adopted the sample-size weighting method in our analysis. And we filter the three original GWAS summary statistics datasets by taking their intersected SNPs. After the filtering, 2020153 SNPs remain in the dataset.

4. Discussion

Targeting on three GWAS study datasets on trans-ethnic groups including East-Asian (Japanese), Caucasian and African ancestry groups, we utilize METAL tool to successfully conduct meta-analysis and discover the genetic architecture of BMI-associated signals across different ethnic populations. The study doesn't identify new genetic signal but replicates the discoveries of previous studies [5] [7]. As table 1 listed, the top three variants in the meta-analysis are "rs1558902", "rs2237897" and "rs11671664".

The top SNP, rs1558902, has been reported as an FTO gene variant, and it has been discovered that allele A carriers have on average greater BMI and waist circumference than allele T carriers in east Asian population [10]. rs2237897 has been identified to have effects on type II diabetes in Chinese population [11], and type II diabetes tend to have higher BMI. rs11671664 was mentioned in a recent study as being associated with BMI and obesity measures [12].

The results indicate that part of genetic architecture on BMI has been shared across the east-Asian population and Caucasian population. But this pattern is not obvious in African ancestry group.

References

[1] Renehan, A. G., Tyson, M., Egger, M., Heller, R. F., & Zwahlen, M. (2008). Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. The lancet, 371(9612), 569-578.

[2] Yang, J., Bakshi, A., Zhu, Z., Hemani, G., Vinkhuyzen, A. A., Lee, S. H., ... & Visscher, P. M. (2015). Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. Nature genetics, 47(10), 1114-1120.

[3] Yengo, L., Sidorenko, J., Kemper, K. E., Zheng, Z., Wood, A. R., Weedon, M. N., ... & GIANT Consortium. (2018). Meta-analysis of genome-wide association studies for height and body mass index in~ 700000 individuals of European ancestry. Human molecular genetics, 27(20), 3641-3649.

[4] Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., ... & Lindström, J. (2015). Genetic studies of body mass index yield new insights for obesity biology. Nature, 518(7538), 197-206.

ISSN: 2790-5209 DOI: 10.56028/fesr.1.1.21
[5] Yang, J., Loos, R. J., Powell, J. E., Medland, S. E., Speliotes, E. K., Chasman, D. I., ... & Tönjes, A. (2012). FTO genotype is associated with phenotypic variability of body mass index. Nature, 490(7419), 267-272.

[6] Wen, W., Zheng, W., Okada, Y., Takeuchi, F., Tabara, Y., Hwang, J. Y., ... & Tanaka, T. (2014). Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. Human molecular genetics, 23(20), 5492-5504.

[7] Akiyama, M., Okada, Y., Kanai, M., Takahashi, A., Momozawa, Y., Ikeda, M., ... & Kamatani, Y. (2017). Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. Nature genetics, 49(10), 1458-1467.

[8] Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., ... & Collins, R. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS medicine, 12(3), e1001779.

[9] Willer, C. J., Li, Y., & Abecasis, G. R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics, 26(17), 2190-2191.

[10] Merritt, D. C., Jamnik, J., & El-Sohemy, A. (2018). FTO genotype, dietary protein intake, and body weight in a multiethnic population of young adults: a cross-sectional study. Genes & nutrition, 13(1), 1-10.

[11] Bai, H., Liu, H., Suyalatu, S., Guo, X., Chu, S., Chen, Y., ... & Narisu, N. (2015). Association analysis of genetic variants with type 2 diabetes in a Mongolian population in China. Journal of diabetes research, 2015.

[12] Reddon, H., Gerstein, H. C., Engert, J. C., Mohan, V., Bosch, J., Desai, D., ... & Meyre, D. (2016). Physical activity and genetic predisposition to obesity in a multiethnic longitudinal study. Scientific reports, 6(1), 1-10.

[13] Pan-UKB team. https://pan.ukbb.broadinstitute.org. 2020.