HHS Public Access

Author manuscript

Int J Obes (Lond). Author manuscript; available in PMC 2019 March 07.

Published in final edited form as:

Int J Obes (Lond). 2019 March ; 43(3): 450–456. doi:10.1038/s41366-018-0069-2.

Detecting epistasis within chromatin regulatory circuitry reveals *CAND2* as a novel susceptibility gene for obesity

Shan-Shan Dong[#1], **Shi Yao**[#1], **Yi-Xiao Chen**[1], **Yan Guo**[1], **Yu-Jie Zhang**[1], **Hui-Min Niu**[1], **Ruo-Han Hao**[1], **Hui Shen**[2], **Qing Tian**[2], **Hong-Wen Deng**[2], and **Tie-Lin Yang**[1]

[1]Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China

[2]School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA 70112, USA

[#] These authors contributed equally to this work.

Abstract

Background: Genome-wide association studies have identified many susceptibility loci for obesity. However, missing heritability problem is still challenging and ignorance of genetic interactions is believed to be an important cause. Current methods for detecting interactions usually do not consider regulatory elements in non-coding regions. Interaction analyses within chromatin regulatory circuitry may identify new susceptibility loci.

Methods: We developed a pipeline named interaction analyses within chromatin regulatory circuitry (IACRC), to identify genetic interactions impacting body mass index (BMI). Potential interacting SNP pairs were obtained based on Hi-C datasets, PreSTIGE (Predicting Specific Tissue Interactions of Genes and Enhancers) algorithm, and super enhancer regions. SNP × SNP analyses were next performed in three GWAS datasets, including 2286 unrelated Caucasians from Kansas City, 3062 healthy Caucasians from the Gene Environment Association Studies initiative, and 3164 Hispanic subjects from the Women's Health Initiative.

Results: A total of 16,643,227 SNP × SNP analyses were performed. Meta-analyses showed that two SNP pairs, rs6808450–rs9813534 (combined $P = 2.39 \times 10^{-9}$) and rs6808450–rs3773306 (combined $P = 2.89 \times 10^{-9}$) were associated with BMI after multiple testing corrections. Single-SNP analyses did not detect significant association signals for these three SNPs. In obesity relevant cells, rs6808450 is located in intergenic enhancers, while rs9813534 and rs3773306 are located in the region of strong transcription regions of *CAND2* and *RPL32*, respectively. The expression of *CAND2* was significantly downregulated after the differentiation of human Simpson–Golabi–Behmel syndrome (SGBS) preadipocyte cells ($P = 0.0241$). Functional validation in the International Mouse Phenotyping Consortium database showed that *CAND2* was associated with increased lean body mass and decreased total body fat amount.

Correspondence to: Tie-Lin Yang.

Conflict of interest The authors declare that they have no conflict of interest.

Electronic supplementary material The online version of this article (https://doi.org/10.1038/s41366-018-0069-2) contains supplementary material, which is available to authorized users.

Conclusions: Detecting epistasis within chromatin regulatory circuitry identified *CAND2* as a novel obesity susceptibility gene. We hope IACRC could facilitate the interaction analyses for complex diseases and offer new insights into solving the missing heritability problem.

Introduction

Obesity is a worldwide health problem. It is associated with various diseases, particularly cardiovascular disease, type 2 diabetes and certain types of cancer. In addition, obesity is becoming more widespread and it is estimated that the number of obese individuals would be more than 1.12 billion by 2030 [1]. Genetic studies aimed to discover genetic variations that can be used in the prevention and treatment of obesity have drawn wide public concern.

Genome-wide association studies (GWASs) are extremely powerful to investigate the genetic architecture of complex diseases [2]. With the help of GWASs, many loci associated with obesity have been identified. However, missing heritability, which is defined as the fact that variants identified so far can explain only a small proportion of the observed heritability of diseases, is still a big challenge. For example, with up to 339,224 individuals, Locke et al. [3] identified 97 body mass index (BMI)-associated loci, which only account for 2.7% of BMI variation. In contrast, the heritability of BMI was estimated as 40–70% [4, 5]. One significant limitation of GWASs is that it examines the effect of each locus independently. However, complex diseases often arise from genetic interactions [6]. In addition, ignorance of genetic interactions is believed to be an important cause for a substantial proportion of the missing heritability in studies of complex diseases/traits [7]. Therefore, researchers have tried to detect epistasis. However, interaction analyses using SNPs in the whole genome usually suffer from the problem of a very stringent significance [8]. Therefore, it is important to select regions of interest to discover the association signal. To solve this problem, previous studies usually restricted the search to a few candidate genes [9, 10], risk SNPs that had weak marginal effects [11, 12], known GWAS hits, protein–protein interactions, or pathway information [13]. However, none of them have specifically considered the interactions between regulatory elements in non-coding regions and their target genes.

Recently, with the release of regulatory data from Encyclopedia of DNA Elements (ENCODE) [14] and Roadmap Epigenomics Project [15], researchers began to realize the importance of regulatory elements in non-coding regions. Diseases associated with susceptibility SNPs identified by GWASs are found to be enriched in cellspecific regulatory elements, mostly enhancers [16, 17]. Of the heritability explained by susceptibility SNPs identified by GWASs, SNPs in regulatory elements were estimated to account for 79% of the variance [18]. Variants in enhancers often regulate target gene expression through long-range interaction [19]. In addition, analyses of 3D genomic architecture showed that enhancers within a gene’s regulatory circuitry can physically interact with one another and collectively regulate the expression of the target gene [20]. Considering the important roles of enhancers for genetic predisposition to diseases, it is possible that interaction analyses within chromatin regulatory circuitry may identify new susceptibility loci for obesity.

In this study, we developed a pipeline named interaction analyses within chromatin regulatory circuitry (IACRC), to identify genetic interactions impacting BMI. IACRC would automatically select chromatin regulatory circuit regions with Hi-C datasets, PreSTIGE (Predicting Specific Tissue Interactions of Genes and Enhancers) algorithm [21], or super enhancer regions [22]. SNP × SNP interaction analyses were then performed in regions within chromatin regulatory circuits. Our results would identify novel genes that may contribute to the development of obesity. The pipeline can also be used to detect epistasis within chromatin regulatory circuitry for other complex diseases.

Materials and methods

GWAS datasets

Three GWAS datasets were used. Basic characteristics for samples in all datasets are listed in Table 1. The first dataset (KCS) comprises 2286 unrelated US Caucasians of Northern European origin living in Kansas City and its surrounding areas. The description of this study has been detailed in our previous study [23]. The second dataset was downloaded from the Database of Genotypes and Phenotypes (dbGaP) with the accession number of phs000091.v2. p1. This study is part of the Gene Environment Association Studies initiative that aimed to identify genetic factors that contribute to type 2 diabetes mellitus (GENEVA_T2D). Data from 3062 healthy Caucasian controls were used in our analyses. The third dataset was also downloaded from the dbGaP database and the accession number is phs000386.v7.p3. This dataset belongs to the Women's Health Initiative (WHI) project and we used the data from 3164 Hispanic subjects in our analyses. The genotyping platform for all three datasets was the Affymetrix 6.0 array.

Acquisition of potential interacting regions

The outline of the current study is shown in Fig. 1. An inhouse python script, which can be freely downloaded from the web resource (https://github.com/studentyaoshi/IACRC/), was used to obtain potential interacting regions within chromatin regulatory circuits meeting one of the following conditions. Firstly, Hi-C chromatin interacting regions for the obesity relevant cells were downloaded from published articles and the 4DGenome chromatin interaction database (https://4dgenome.research.chop.edu/). The significance threshold of interaction bins with unusually high counts was set as $P < 0.05$. Chromatin states of the obesity relevant cells/tissues (supplementary Table S1) were obtained from the Roadmap database to get the enhancer region information and annotate the interacting regions. We collected Hi-C interactions for which there was a protein coding gene in one locus that was paired to an enhancer in the other. These pairs were further used to collect enhancers that were associated with the same target gene. We also collected interactions with both loci annotated as enhancers. Secondly, The PreSTIGE algorithm [21] was used to predict enhancer-gene interactions from disease relevant cells. We set the threshold that maximizes the number of predictions made while maintaining an estimated positive prediction rate >60%, i.e., the Shannon entropy Q score had to be below 6.1 for the enhancer and below 6.8 for the gene paired to that enhancer [21]. Enhancer pairs with same target genes were also collected. Thirdly, super-enhancers previously defined for disease relevant cells/tissues [22] were also used. All genes within 100 kb of super-enhancers were defined as putative target

genes. Interactions may also occur within the super-enhancer region. We used data from the GM12878 cell line to obtain Hi-C [24–27] and PreSTIGE-based interacting regions. Data for obesity relevant tissues/cells (supplementary Table S2) were used to obtain super-enhancer-based interacting regions. The interacting regions supported by at least one method were used in subsequent analyses.

Acquisition of SNP pairs and SNP pruning

After interacting region collections, SNP pairs were obtained from these interacting regions by using an inhouse script that can also be downloaded from the web resource. SNP quality control was then performed with the following criteria: individual missing rate <5%, SNP call rate >95%, minor allele frequency (MAF)>5%, and Hardy–Weinberg equilibrium (HWE) *P*-value> 0.001. SNP pairs with two SNPs in linkage disequilibrium (LD, r^2>0.1) with each other were removed.

SNP × SNP interaction analyses and meta-analyses

SNP × SNP interaction analyses were finally performed by using the regression function (linear regression for quantitative trait and logistic regression for qualitative trait) in R (version 3.3.2) for the remaining SNP pairs. For each population, the principal components were calculated by using the GCTA software [28]. The first five principal components, age, and sex were used as potential covariates. Finally, the METAL software [29] was used to combine SNP × SNP interaction results obtained from different populations. We used the Bonferroni correction method [30] to account for the multiple testing problems, ie., the significant threshold was set as 0.05/(number of SNP pairs).

Differential expression analysis

The Gene Expression Omnibus (GEO) dataset GSE76131 was used to check whether the identified susceptibility genes were differentially expressed during the adipogenesis of human Simpson–Golabi–Behmel syndrome (SGBS) preadipocyte cells (0 vs. 384 h).

Genetically modified animals used for functional validation

We used the International Mouse Phenotyping Consortium (IMPC) database, which aims to discover the biological functions of every protein-coding gene through generating null alleles in mice on a C57BL/6 genetic background [31], for the functional investigation of identified susceptibility genes by our pipeline.

Results

SNP × SNP interaction analyses

Detail statistics of intermediate stages are shown in Fig. 2. A total of 16,643,227 SNP × SNP analyses were performed and the significant threshold after multiple testing correction was set as combined $P < 3.0 \times 10^{-9}$.

Meta-analyses results showed that two SNP pairs, rs6808450–rs9813534 (combined $P = 2.39 \times 10^{-9}$) and rs6808450–rs3773306 (combined $P = 2.89 \times 10^{-9}$) were associated with BMI after multiple testing corrections (Table 2). However, single-SNP analyses did not

detect significant association signals for these three SNPs ($P > 0.05$, supplementary Table S3). Rs9813534 and rs3773306 are in high LD with each other, with the r^2 of 0.97, 0.97, and 0.80 in KCS, GENEVA_T2D, and WHI, respectively. We further checked whether the effect of the minor allele "C" of rs9813534 on BMI was different between subjects carrying different rs6808450 genotypes by using the beta coefficient. As shown in supplementary Figure S1A, in all three populations, the "C" allele was positively associated with BMI in subjects carrying "CC" genotypes of rs6808450. However, in subjects carrying "TT" genotypes of rs6808450, the "C" allele of rs9813534 was negatively associated with BMI. Similarly, the minor "C" allele of rs3773306 was also positively associated with BMI in subjects carrying "CC" genotypes of rs6808450 but negatively associated with BMI in subjects carrying "TT" genotypes of rs6808450 (supplementary Figure S1B).

Differential expression analysis

As shown in Fig. 3, in obesity relevant cells, rs6808450 is located in intergenic enhancers, while rs9813534 and rs3773306 are located in the region of strong transcription regions of *CAND2* and *RPL32*, respectively. Therefore, *CAND2* and *RPL32* are the target genes in regulatory circuits and they may be new obesity susceptibility genes.

We next checked whether *CAND2* and *RPL32* were differentially expressed during the differentiation of SGBS cells. The results showed that only the expression of *CAND2* was significantly downregulated after the differentiation of SGBS (Fig. 4, $P = 0.0241$).

Functional validation in IMPC

Information from the IMPC database further supported the functional involvement of *CAND2* in obesity. This gene is associated with increased lean body mass and decreased total body fat amount (http://www.mousephenotype.org/data/genes/MGI:1914338#section-associations) in mice model.

Code availability

IACRC: https://github.com/studentyaoshi/IACRC.

Discussion

Considering the important roles of enhancers for genetic predisposition to diseases, here we developed a pipeline named IACRC, to perform interaction analyses within chromatin regulatory circuits for BMI.

We identified two significant interaction pairs associated with BMI, rs6808450–rs9813534 and rs6808450–rs3773306. The relationships between these SNPs and obesity or other diseases have not been reported by previous GWASs before. According to the annotation results, *CAND2* and *RPL32* are the target genes in regulatory circuits. Further differential expression analyses in SGBS cells and functional validation in the mice model suggested *CAND2* may be a novel obesity susceptibility gene. However, the relationship between *RPL32* and obesity were not further validated. Since rs9813534 and rs3773306 are in high LD with each other, it is possible that the interaction between rs6808450 and rs3773306 was

detected due to LD. *CAND2* is located on chromosome 3p25.2 (chr3:12838171-12876313, Reference GRCh37.p13 Assembly) and encodes TATA-binding protein-interacting protein 120B (TIP120B). The association between *CAND2* and obesity or BMI has not been reported before. A previous study revealed that TIP120B could inhibit the ubiquitination of myogenin through binding to Cullin1, resulting in the stabilization of myogenin and acceleration of the myogenic differentiation process [32]. In addition, our analyses using SGBS cells showed that *CAND2* was downregulated after the differentiation of human SGBS cells, suggesting that it might suppress the adipogenesis process. Since the imbalance between adipose and muscle mass is a hallmark of obesity [33], it is possible that *CAND2* might be associated with obesity through promoting myogenesis and suppressing adipogenesis.

Limitations of the current study should be addressed. Since the current Hi-C and PreSTIGE data for obesity relevant cell/tissues are still limited, so we only used the Hi-C and PreSTIGE data from the GM12878 cell line to collect interacting regions in these two conditions. When data for other obesity relevant cell/tissues are available, other interactions associated with BMI may be detected with the IACRC pipeline.

In summary, in this study, we developed IACRC to perform interaction analyses within chromatin regulatory circuitry and identify genetic interactions impacting BMI. We identified two SNP pairs, rs6808450-rs9813534 and rs6808450-rs3773306 were associated with BMI after multiple testing corrections. Further annotation, differential expression analysis, and functional validation results supported that *CAND2* may be a novel obesity susceptibility gene. We hope IACRC could facilitate interaction analyses and identify new genetic interactions for other complex diseases.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the Gene Environment Association Studies initiative (GENEVA) which aimed to identify genetic factors that contribute to type 2 diabetes mellitus. We also thank the Women's Health Initiative (WHI) project. When we performed the current study, we did not collaborate with the investigators of these projects. Therefore, our study does not necessarily reflect the opinions of them. The datasets we used were obtained through dbGaP authorized access with the accession number of phs000091.v2.p1 and phs000386.v7.p3.

Funding: This study is supported by National Natural Science Foundation of China (31371278, 81573241, 31701095, and 31771399); China Postdoctoral Science Foundation (2016M602797); Natural Science Basic Research Program Shaanxi Province (2016JQ3026); and the Fundamental Research Funds for the Central Universities.

References

1. Kelly T, Yang W, Chen CS, Reynolds K, He J. Global burden of obesity in 2005 and projections to 2030. Int J Obes. 2008;32:1431–7.
2. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS discovery: biology, function, and translation. Am J Hum Genet. 2017;101:5–22. [PubMed: 28686856]

3. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518:197–206. [PubMed: 25673413]
4. Stunkard AJ, Foch TT, Hrubec Z. A twin study of human obesity. JAMA. 1986;256:51–4. [PubMed: 3712713]
5. Turula M, Kaprio J, Rissanen A, Koskenvuo M. Body weight in the Finnish Twin Cohort. Diabetes Res Clin Pract. 1990;10(Suppl 1):S33–6. [PubMed: 2286148]
6. Cordell HJ. Detecting gene-gene interactions that underlie human diseases. Nat Rev Genet. 2009;10:392–404. [PubMed: 19434077]
7. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: genetic interactions create phantom heritability. Proc Natl Acad Sci USA. 2012;109:1193–8. [PubMed: 22223662]
8. Wei WH, Hemani G, Haley CS. Detecting epistasis in human complex traits. Nat Rev Genet. 2014;15:722–33. [PubMed: 25200660]
9. Dong SS, Hu WX, Yang TL, Chen XF, Yan H, Chen XD, et al. SNP-SNP interactions between WNT4 and WNT5A were associated with obesity related traits in Han Chinese Population. Sci Rep. 2017;7:43939. [PubMed: 28272483]
10. Wang MH, Li J, Yeung VS, Zee BC, Yu RH, Ho S, et al. Four pairs of gene-gene interactions associated with increased risk for type 2 diabetes (CDKN2BAS-KCNJ11), obesity (SLC2A9-IGF2BP2, FTO-APOA5), and hypertension (MC4R-IGF2BP2) in Chinese women. Meta Gene. 2014;2:384–91. [PubMed: 25606423]
11. Lucas G, Lluis-Ganella C, Subirana I, Musameh MD, Gonzalez JR, Nelson CP, et al. Hypothesis-based analysis of gene-gene interactions and risk of myocardial infarction. PLoS ONE. 2012;7:e41730. [PubMed: 22876292]
12. Yang TL, Guo Y, Li J, Zhang L, Shen H, Li SM, et al. Gene-gene interaction between RBMS3 and ZNF516 influences bone mineral density. J Bone Miner Res. 2013;28:828–37. [PubMed: 23045156]
13. Ma L, Brautbar A, Boerwinkle E, Sing CF, Clark AG, Keinan A Knowledge-driven analysis identifies a gene-gene interaction affecting high-density lipoprotein cholesterol levels in multiethnic populations. PLoS Genet. 2012;8:e1002714. [PubMed: 22654671]
14. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489:57–74. [PubMed: 22955616]
15. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518:317–30. [PubMed: 25693563]
16. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012;337:1190–5. [PubMed: 22955828]
17. Trynka G, Sandor C, Han B, Xu H, Stranger BE, Liu XS, et al. Chromatin marks identify critical cell types for fine mapping complex trait variants. Nat Genet. 2013;45:124–30. [PubMed: 23263488]
18. Gusev A, Lee SH, Trynka G, Finucane H, Vilhjalmsson BJ, Xu H, et al. Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. Am J Hum Genet. 2014;95:535–52. [PubMed: 25439723]
19. Krijger PH, de Laat W. Regulation of disease-associated gene expression in the 3D genome. Nat Rev Mol Cell Biol. 2016;17:771–82. [PubMed: 27826147]
20. Ing-Simmons E, Seitan VC, Faure AJ, Flicek P, Carroll T, Dekker J, et al. Spatial enhancer clustering and regulation of enhancerproximal genes by cohesin. Genome Res. 2015;25:504–13. [PubMed: 25677180]
21. Corradin O, Saiakhova A, Akhtar-Zaidi B, Myeroff L, Willis J, Cowper-Sal lari R, et al. Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. Genome Res. 2014;24:1–13. [PubMed: 24196873]
22. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, et al. Super-enhancers in the control of cell identity and disease. Cell. 2013;155:934–47. [PubMed: 24119843]

23. Yang TL, Guo Y, Li SM, Li SK, Tian Q, Liu YJ, et al. Ethnic differentiation of copy number variation on chromosome 16p12.3 for association with obesity phenotypes in European and Chinese populations. Int J Obes. 2013;37:188–90.
24. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159:1665–80. [PubMed: 25497547]
25. Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, et al. CTCFmediated human 3D genome architecture reveals chromatin topology for transcription. Cell. 2015;163:1611–27. [PubMed: 26686651]
26. Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L, et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat Genet. 2015;47:598–606. [PubMed: 25938943]
27. Martin P, McGovern A, Orozco G, Duffus K, Yarwood A, Schoenfelder S, et al. Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. Nat Commun. 2015;6:10069. [PubMed: 26616563]
28. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88:76–82. [PubMed: 21167468]
29. Willer CJ, Li Y, Abecasis GR METAL: fast and efficient metaanalysis of genomewide association scans. Bioinformatics. 2010;26:2190–1. [PubMed: 20616382]
30. Bonferroni CE. Teoria statistica delle classi e calcolo delle probabilità. Pubbl Del R Ist Super di Sci Econ e Commer di Firenze. 1936;8:3–62.
31. International Mouse Knockout C, Collins FS, Rossant J, Wurst W. A mouse for all reasons. Cell. 2007;128:9–13. [PubMed: 17218247]
32. Shiraishi S, Zhou C, Aoki T, Sato N, Chiba T, Tanaka K, et al. TBP-interacting protein 120B (TIP120B)/cullin-associated and neddylation-dissociated 2 (CAND2) inhibits SCF-dependent ubiquitination of myogenin and accelerates myogenic differentiation. J Biol Chem. 2007;282:9017–28. [PubMed: 17242400]
33. Kahn BB, Flier JS. Obesity and insulin resistance. J Clin Invest. 2000;106:473–81. [PubMed: 10953022]

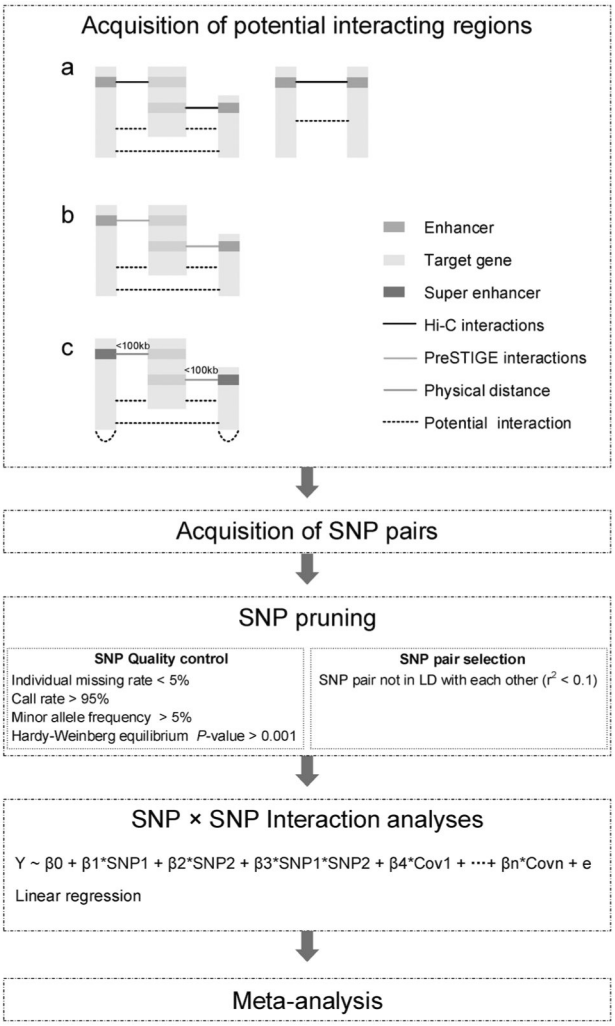

Fig. 1.

Schematic diagram of IACRC. Interacting regions were firstly obtained based on Hi-C data (panel a), PreSTIGE (Predicting Specific Tissue Interactions of Genes and Enhancers) algorithm (panel b), and super-enhancer data (panel c). SNP pairs were next extracted from these interacting regions. SNP pairs with SNPs that failed quality control were removed. SNP pairs with two SNPs in linkage disequilibrium (LD, r^2>0.1) with each other were also removed. SNP × SNP interaction analyses were performed by using the linear regression function in R (version 3.3.2) for the remaining SNP pairs. The METAL software were used to combine SNP × SNP interaction results obtained from different populations

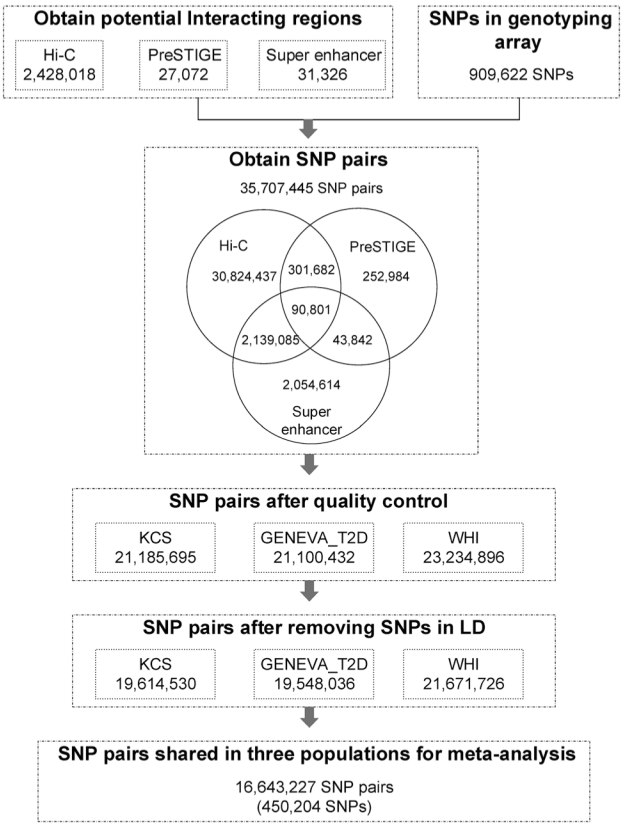

Fig. 2.
Statistics of intermediate stages in the analyses for BMI

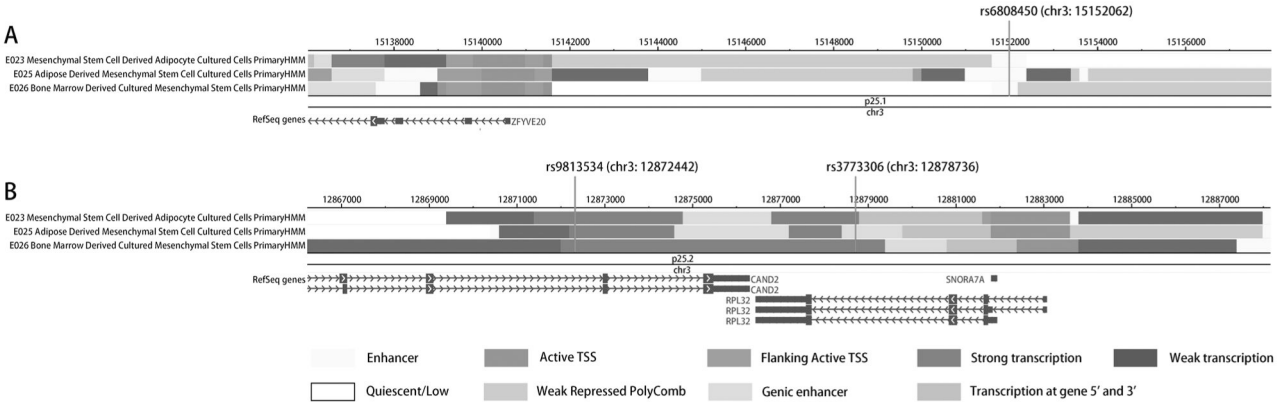

Fig. 3.

a Annotation of rs6808450. The vertical red line refers to rs6808450. **b** Annotation of rs9813534 and rs3773306. The two vertical red lines refer to rs9813534 and rs3773306, respectively. HMM refers to the chromatin states predicted by hidden Markov model based on combinations of histone modification marks

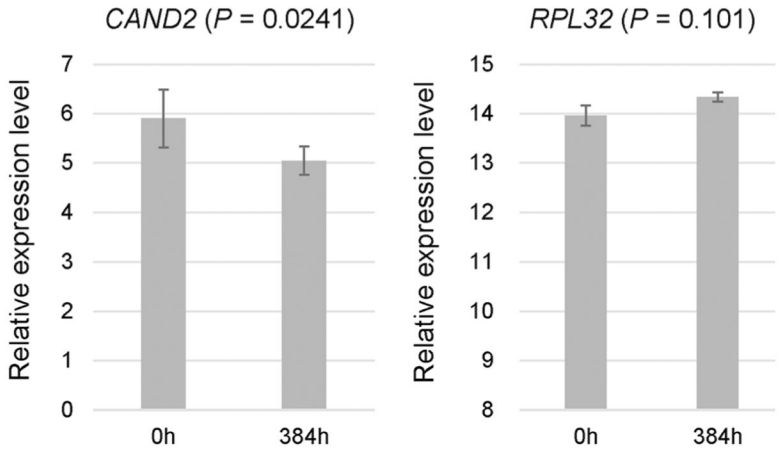

Fig. 4.

Differential expression analyses results (384 vs. 0 h) in the adipogenesis of human Simpson–Golabi–Behmel syndrome (SGBS) preadipocyte cell for *CAND2* and *RPL32*

Table 1

Basic characteristics of all subjects

Population	KCS Caucasian	GENEVA_T2D Caucasian	WHI Caucasian
Sample size	2286	3062	3164
Age (years)	51.37 ± 13.76	57.11 ± 7.65	60.07 ± 6.67
Height (cm)	166.35 ± 8.47	170.14 ± 9.54	156.79 ± 5.66
Weight (kg)	75.27 ± 17.54	73.45 ± 13.90	71.90 ± 14.57
BMI (kg/m^2)	27.14 ± 5.75	25.31 ± 4.10	29.20 ± 5.78

Note: Data are shown as mean ± standard deviation

KCS Kansas City study, *GENEVA_T2D* Gene Environment Association Studies initiative that aimed to identify genetic factors that contribute to type 2 diabetes mellitus, *WHI* Women's Health Initiative, *BMI* body mass index

Table 2

The interaction SNP pairs identified by IACRC for BMI

SNP1	Location1	SNP2	Location2	KCS		GENEVA_T2D		WHI		Combined *P*
				Beta	*P*	Beta	*P*	Beta	*P*	
rs6808450	Intergenic	rs9813534	CAND2	0.1852	5.22×10^{-3}	0.2177	1.57×10^{-4}	0.1996	2.22×10^{-4}	2.39×10^{-9}
rs6808450	Intergenic	rs3773306	RPL32	0.1799	6.86×10^{-3}	0.2241	1.07×10^{-4}	0.2132	2.91×10^{-4}	2.89×10^{-9}

Note: Only the significant SNP pairs after multiple testing corrections are shown

Location1 the located region/gene of SNP1, *Location2* the located region/gene of SNP2, *KCS* Kansas City study, *GENEVA_T2D* Gene Environment Association Studies initiative that aimed to identify genetic factors that contribute to type 2 diabetes mellitus, *WHI* Women's Health Initiative, *Combined P* meta-analyses results by using all three GWAS samples