Published in final edited form as:
Nat Genet. ; 44(3): . doi:10.1038/ng.1086.

Common variants at CDKAL1 and KLF9 are associated with body mass index in East Asian populations

Yukinori Okada[1,2,*], **Michiaki Kubo**[3], **Hiroko Ohmiya**[1], **Atsushi Takahashi**[1], **Natsuhiko Kumasaka**[1], **Naoya Hosono**[3], **Shiro Maeda**[4], **Wanqing Wen**[5], **Rajkumar Dorajoo**[6,7], **Min Jin Go**[8], **Wei Zheng**[5], **Norihiro Kato**[9], **Jer-Yuarn Wu**[10,11], **Qi Lu**[12], **the GIANT consortium**[16], **Tatsuhiko Tsunoda**[13], **Kazuhiko Yamamoto**[2], **Yusuke Nakamura**[14], **Naoyuki Kamatani**[1], and **Toshihiro Tanaka**[15]

[1]Laboratory for Statistical Analysis, Center for Genomic Medicine (CGM), RIKEN, Yokohama, Japan

[2]Department of Allergy and Rheumatology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

[3]Laboratory for Genotyping Development, CGM, RIKEN, Yokohama, Japan

[4]Laboratory for Endocrinology and Metabolism, CGM, RIKEN, Yokohama, Japan

[5]Division of Epidemiology, Department of Medicine, Vanderbilt Center of Epidemiology and Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, USA

[6]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore

[7]Department of Genomics of Common Disease, School of Public Health, Imperial College London, Hammersmith Hospital, London, United Kingdom

[8]Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, the Republic of Korea

[9]Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo, Japan

[10]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan

[11]School of Chinese Medicine, China Medical University, Taichung, Taiwan

[*]Corresponding author: Yukinori Okada, Laboratory for Statistical Analysis, Center for Genomic Medicine, RIKEN, 4-6-1 Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan, Telephone: +81-3-5449-5509, Fax: +81-3-5449-5564, yokada@src.riken.jp.

[16]Full author list of the GIANT consortium is provided in Supplementary Information.

AUTHOR CONTRIBUTIONS

Y.O. and T. Tanaka designed the study and drafted the manuscript. N.H. and M.K. performed the genotyping. Y.O., H.O., A.T., N. Kumasaka, and T. Tsunoda performed the statistical analysis. Y.O. and M.K. managed the clinical information. W.W., R.D., M.J.G., W.Z., N. Kato., J.W., and Q.L. managed the replication study set 3. The GIANT consortium managed the association study in Europeans. S.M., K.Y., Y.N., N. Kamatani, and T. Tanaka supervised the study.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

URLs

The URLs for data presented herein are as follows.
The BioBank Japan Project, http://biobankjp.org
EIGENSTRAT software, http://genepath.med.harvard.edu/~reich/Software.htm
MACH and mach2qtl software, http://www.sph.umich.edu/csg/abecasis/MACH/index.html
International HapMap Project, http://www.hapmap.org
Quanto software, http://hydra.usc.edu/gxe
R statistical software, http://cran.r-project.org

[12]Department of Nutrition, Harvard School of Public Health, Boston 02115, Massachusetts, USA

[13]Laboratory for Medical Informatics, CGM, RIKEN, Yokohama, Japan

[14]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan

[15]Laboratory for Cardiovascular Diseases, CGM, RIKEN, Yokohama, Japan

Abstract

Obesity is a disorder with complex genetic etiology, and its epidemic is a worldwide problem. Although multiple genetic loci associated with body mass index (BMI), the most common measure of obesity, have been identified in European populations, few studies have focused on Asian populations. Here, we report a genome-wide association study (GWAS) and replication studies with 62,245 East Asian subjects, which identified two novel BMI-associated loci in the *CDKAL1* locus at 6p22 (rs2206734, $P = 1.4 \times 10^{-11}$) and the *KLF9* locus at 9q21 (rs11142387, $P = 1.3 \times 10^{-9}$), as well as previously reported loci (the *SEC16B*, *BDNF, FTO*, *MC4R*, and *GIPR* loci; $P < 5.0 \times 10^{-8}$). We subsequently performed gene–gene interaction analysis and identified an interaction ($P = 2.0 \times 10^{-8}$) between SNPs in the *KLF9* locus (rs11142387) and the *GDF8* locus at 2q32 (rs13034723). These findings should provide useful insights into the etiology of obesity.

Obesity is a major risk factor for a number of chronic diseases, and its recent rise in worldwide prevalence imposes serious medical and economic burdens[1]. It is well known that obesity is a highly heritable trait and around 40–70% of inter-individual variation is attributable to genetic factors[2]. Recently, genome-wide association studies (GWASs) have identified dozens of genetic loci associated with body mass index (BMI), the most common measure of obesity[3–12]. However, most of these studies were conducted in European populations, and few studies have assessed Asian populations[5,11], which account for two-thirds of the world's population. The degree of adiposity and the risks of diseases exacerbated by obesity are greater in Asians than in Europeans when evaluated with the same BMI[13]. Thus, the study of Asian populations might lead to the identification of novel associated loci and provide novel insight into the genetic architecture of obesity. We report herein a large-scale GWAS and replication studies of BMI examining a total of 62,245 subjects from East Asian populations.

In the GWAS for BMI, we enrolled 26,620 Japanese subjects under the support of the BioBank Japan Project[14] (Supplementary Table 1 and Supplementary Figure 1). Stringent quality control criteria, including principal component analysis (PCA) for evaluating potential population stratifications, were applied as previously described[15]. To extend coverage to the genomic region, whole-genome imputation was performed for SNPs that were not genotyped, and the genotype data of 2,178,018 autosomal SNPs with minor allele frequency (MAF) ≥ 0.01 was obtained. Each SNP was evaluated for association with BMI using a linear regression model, assuming additive effects of allele dosages on the rank-based inverse normal transformed values of BMI. Although no significant population stratification was suggested in our study population (Supplementary Figure 2) or in our previous studies for Japanese[15], for robustness we applied genomic control corrections for the results of the GWAS using inflation factor, λ_{GC}, of 1.123 (referenced $\lambda_{GC\_1000} = 1.005$)[16]. The Quantile-Quantile plot of P-values indicated remarkable discrepancy in its tail from the null hypothesis (Supplementary Figure 3), which suggested the presence of significant associations in this GWAS. We identified significant associations in three chromosomal loci (the *KLF9* locus at 9q21, the *BDNF* locus at 11p14, and the *GIPR* locus at 19q13) that satisfied the genome-wide significance threshold of $P < 5.0 \times 10^{-8}$ (Table 1 and Figure 1a).

To further validate the associations identified in the GWAS, we performed replication studies on three independent sets (Supplementary Table 1 and Supplementary Figure 1). The first and second sets consisted of 3,763 and 4,147 Japanese subjects from the BioBank Japan Project[14], respectively. The third set consisted of 27,715 subjects enrolled in the concurrently conducted meta-analysis of GWASs for BMI with cohorts of East Asians[17]. First, 36 SNPs most significantly associated in each of the loci with $P < 5.0 \times 10^{-5}$ in GWAS were evaluated in the first replication set, and then, 11 SNPs with $P < 5.0 \times 10^{-5}$ in the combined study of GWAS and replication set 1 were further evaluated in replication sets 2 and 3. Through the combined results of the GWAS and the replication studies, we identified a total of seven loci that satisfied the genome-wide significance threshold (Table 1), which included the three loci that originally satisfied the threshold in the GWAS. Among the seven identified loci, five were previously identified to be associated with BMI in Europeans (the *SEC16B*, *BDNF*, *FTO*, *MC4R*, and *GIPR* loci at 1q25, 11p14, 16q12, 18p21, and 19q13, respectively)[3–9,11,12]. Associations at the remaining two loci (6p22 and 9q21) have not been previously reported, including the large-scaled study of Europeans[12], and were novel findings to our knowledge. The landmark SNP at 6p22 (rs2206734, $P = 1.4 \times 10^{-11}$; Figure 1b) was located in the coding region of *CDKAL1*, the gene encoding cyclin-dependent kinase 5 (*CDK5*) regulatory subunit associated protein 1-like 1. The other SNP at 9q21 (rs11142387, $P = 1.3 \times 10^{-9}$; Figure 1c) was located in the promoter region of *KLF9*, the gene encoding Krüppel-like factor 9 (also known as basic transcription element-binding protein1; *BTEB1*). The LD block including rs11142387 also covered several genes, such as *MAMDC2*, *SMC5*, and *TRPM3* (Figure 1c). Although we examined tag CNVs and expression analysis of rs11142387 in the locus using publicly available database, no significant findings were observed (Supplementary Figure 4). Our study demonstrated suggestive associations ($5.0 \times 10^{-8} \leq P < 5.0 \times 10^{-5}$) in the *CCK* (cholecystokinin) locus at 3p22 (rs4377469, $P = 1.6 \times 10^{-7}$) and in the *ZNF169* (zinc finger protein 169) locus at 9q22 (rs10993160, $P = 5.5 \times 10^{-7}$). Combination of these identified loci ($P < 5.0 \times 10^{-8}$) explained 0.72% of inter-individual variance in BMI, of which the *FTO* locus explained the largest proportion (0.20%, Table 1).

To evaluate ethnic differences in the genetics of obesity, associations of the loci with confirmed or suggestive associations ($P < 5.0 \times 10^{-5}$) were further evaluated in Europeans by using the results of meta-analysis for 123,865 subjects by the GIANT consortium (Table 1)[12]. We found the same directional effects of alleles in all the nine evaluated loci. Significant associations were observed in five loci ($P < 0.028$, false discovery rate (FDR) < 0.05), including the *CDKAL1* locus ($P = 0.0049$), whereas no association could be observed in the *KLF9* locus ($P = 0.50$).

Using our data, we then evaluated the associations of the previously reported BMI-associated loci, most of which had been identified in Europeans (Supplementary Table 2)[3–12,18]. Our study replicated the associations with the same directional effects of the alleles at 10 loci, including the *TMEM18*, *RBJ-ADCY3-POMC*, *GNPDA2*, *FLJ35779-HMGCR*, *TFAP2B*, *TRHR*, *MTCH2*, *MAP2K5-LBXCOR1*, *SH2B1-ATP2A1*, and *BMP2* loci, in addition to the five loci that have been already replicated in our GWAS ($P < 0.02$, FDR < 0.05). As for the loci reported in Koreans[11], we replicated the association in *FTO*, but not in *LOC729076* at 6q24 (Supplementary Table 2).

Because obesity is a polygenic trait and epistasis may help dissect its genetic background[19], we performed gene–gene interaction analysis of BMI. We evaluated gene–gene interaction assuming additive × additive effects of SNPs[20] between each of the seven SNPs confirmed to be associated with BMI and all of the 2,178,018 genome-wide SNPs (Supplementary Figure 5) using the subjects enrolled in the GWAS, and subsequently conducted a replication study. We found an interaction that satisfied the significance threshold ($P < 5.0 \times$

10^{-8}) between rs11142387 at *KLF9* and a SNP located in the promoter region of *GDF8* (growth and differentiation factor-8, also known as myostatin; *MSTN*) at 2q32 (rs13034723, $P = 2.0 \times 10^{-8}$; Table 2 and Figure 2a). Interestingly, the association of rs13034723 at the *GDF8* locus with BMI was not significant ($P = 0.56$; Supplementary Table 3). When the association of rs11142387 was stratified by genotypes of rs13034723, a significant association of rs11142387 with BMI was observed in the subjects with AA genotypes at rs13034723 ($P = 8.7 \times 10^{-15}$; Supplementary Table 4 and 5 and Figure 2b), although no significant association was observed for the subjects with AG or GG genotypes ($P = 0.0029$ and 0.30, respectively).

Through the GWAS and the replication studies, we identified two novel loci, *CDKAL1* and *KLF9*, associated with BMI in East Asians. These two loci also demonstrated significant associations with the risk of obesity (BMI ≥ 27.5; $P = 1.1 \times 10^{-5}$ and 4.2×10^{-4}; Supplementary Table 6). Compared to the results in Europeans[12], the association in the *CDKAL1* locus was shared between East Asians and Europeans, but not in the *KLF9* locus. Because the study in Europeans[12] would have enough power to detect the *KLF9* locus (> 99%), under the assumption of the same effect size as East Asians and the allele frequencies in Europeans (= 0.51), ethnic heterogeneity in the effect of the *KLF9* locus for BMI would be suggested. We also performed gene–gene interaction analysis and demonstrated an interaction between the *KLF9* and *GDF8* loci. Although the substantial role of epistasis in polygenic traits has been recognized, approach to elucidate it has been challenging[20]. Our findings would be one of the initial pieces of evidence for epistatic associations.

Recent studies reported the associations of the *CDKAL1* locus with BMI at 8 years of age (rs4712526)[21] and birthweight (rs7756992)[22], and these SNPs indicated significant associations in our GWAS ($P < 5.0 \times 10^{-5}$). To our knowledge, our study is the first report on the association with adult BMI. The *CDKAL1* locus has been reported to be a risk locus of type 2 diabetes (T2D)[23,24]. We found that the T allele of rs2206734, which decreased BMI, significantly increased T2D risk in our study subjects ($P < 1.4 \times 10^{-18}$; Supplementary Table 6). *CDKAL1* risk variants for T2D were associated with decreased insulin secretion[23]; therefore, the observed effects of the *CDKAL1* risk variant on decreasing BMI might be mediated by decreased insulin secretion. Interestingly, a recent study identified the similar patterns of the associations in the *GIPR* locus[25], and we observed that the BMI-decreasing A allele of rs11671664 at *GIPR* increased T2D risk ($P < 1.5 \times 10^{-5}$; Supplementary Table 6). These findings suggest further studies comprehensively assessing genetic associations with T2D risk, BMI, and insulin secretion should be performed. When the subjects affected with T2D (n = 12,234) were excluded, the association of rs2206734 with BMI was not obvious (effect size = 0.031, $P = 1.4 \times 10^{-6}$). We evaluated the association of the *CDKAL1* and *GIPR* loci with other related traits, including systolic and diastolic blood pressure, and serum lipid levels (total cholesterol; TC, high density lipoprotein cholesterol; HDL-C, low density lipoprotein cholesterol; LDL-C, and triglyceride; TG), although no significant association was observed ($\alpha = 0.05$; Supplementary Table 6).

KLF9 is a member of zinc-finger transcription factors involved in various physiological processes. A recent study indicated *KLF9* as a pro-adipogenic transcription factor acting through transactivation of PPARγ[26], a key component of adipocyte differentiation implicated in obesity[27]. Zobel et al. reported the association of the *KLF7* locus with obesity in the Danish population, although we could not test its relevance because the reported SNP, rs7568369, was not polymorphic (Supplementary Table 2)[18]. It is known that *KLF5*, a gene belonging to the *KLF* family and also known as *BTEB2*, regulates adipocyte differentiation[28]. Considering these observations, the association of the *KLF9* locus with BMI would be plausible from a biological aspect. Contrary to the *CDKAL1* locus, no

significant associations of *KLF9* with T2D risk, or with other related traits, were observed ($\alpha = 0.05$; Supplementary Table 6).

GDF8 is a member of the transforming growth factor-beta (TGF-β) superfamily that regulates mesenchymal stem cell proliferation[29]. A loss-of function mutation in *GDF8* causes muscle hypertrophy and decreased body fat[29,30]. In our study, the SNP in the *GDF8* locus was not associated with BMI, but its genotypes clearly stratified the association in the *KLF9* locus. This would pose the regulatory role of *GDF8* on the effect of *KLF9* on BMI, and further studies evaluating functional epistasis would be desirable. Notably, Grade et al. identified similar conserved promoter/enhancer architecture in *KLF9* and *GDF8* through a search of evolutionary conserved regions, which suggests these two genes may form a synexpression group[31]. Other genes located near *GDF8*, such as *C2orf88*, could also be candidates, and the relatively small sample size in the replication studies provided limited evidence.

Wen et al. concurrently reported a genome-wide meta-analysis for BMI using data of eight cohorts of East Asians[17]. The subjects enrolled in these two studies overlapped due to reciprocal replication approaches, and newly identified loci were shared at *CDKAL1*, while some loci were specifically identified in each study, such as *KLF9*. This could be attributed to differences in study designs, effects of different compositions of the populations in discovery phases, and probability of study-specific bias induced by winner's curse effect.

In summary, our study identified novel associations of the *CDKAL1* and *KLF9* loci with BMI in East Asians. A gene–gene interaction between the *KLF9* and *GDF8* loci was also found. Our study should contribute to understanding of the genetic architecture of obesity.

ONLINE METHODS

Subjects

The subjects enrolled in the GWAS (n = 26,620) and the replication sets 1 and 2 (n = 3,763 and n = 4,147, respectively) were obtained from the BioBank Japan Project[14] at the Institute of Medical Science, the University of Tokyo, which consisted of patients of 32 diseases (Supplementary Table 1). Subjects with ages < 18 or > 85, with dialysis treatment, or who were determined as being of non-Japanese origin by self-report or by PCA in the GWAS or our previous study[15] were not included. Clinical information on the subjects including age (mean ± SD; 61.3 ± 12.9), gender (47.2% for female), and smoking history (49.8% for ever-smoker) were collected by a standard questionnaire. BMI (mean ± SD; 22.7 ± 3.59) were calculated based on self-reported height and weight. BMI based on self-reported data is known to be highly correlated ($r > 0.94$) with that based on measurements[32], and potential bias induced by self-reported data may have little impact on the analysis[32,33]. All participants provided written informed consent as approved by the ethical committees of the RIKEN Yokohama Institute and the Institute of Medical Science, University of Tokyo. The subjects enrolled in the replication set 3 (n = 27,715) consisted of subjects from the eight cohorts of East Asian populations, and they were enrolled in the discovery stage of the concurrent meta-analysis for BMI[17]. The subjects in our GWAS were also enrolled in the replication stage of the meta-analysis[17].

Genotyping and quality control

We used the data of 32 GWAS performed for the BioBank Japan Project, in which patients of each of the 32 diseases were genotyped (Supplementary Table 1)[14]. In the GWAS, genotyping was performed using the Illumina HumanHap610-Quad Genotyping BeadChip (Illumina, CA, USA). After excluding the subjects with call rates lower than 0.98, we excluded SNPs with call rates lower than 0.99, SNPs with ambiguous calls, or non-

autosomal SNPs. We excluded subjects in close kinships based on estimations using identity-by-state (IBS). We considered the subject pairs with an average proportion of alleles shared by IBS > 1.7 to be in first or second degree of kinship, and excluded the member of the pair with lower call rates. We also excluded subjects whose ancestries were estimated to be distinct from the other subjects using PCA performed by EIGENSTRAT version 2.0. We performed PCA for the genotype data of our study along with the genotype data of unrelated European (CEU), African (YRU), and East Asian (Japanese and Han Chinese; JPT + CHB) individuals obtained from the Phase II HapMap database (release 24)[34]. Based on the PCA plot, we excluded the outliers in terms of ancestry from JPT + CHB clusters (Supplementary Figure 1). We then excluded the SNPs with MAF < 0.01 or the SNPs with exact P-value of the Hardy-Weinberg equilibrium test < 1.0×10^{-7} and obtained genotype data of 480,103 SNPs for 26,620 subjects.

Genotype imputation was performed using MACH 1.0 in a two step procedure. The JPT and CHB individuals obtained from Phase II HapMap database (release 24)[34] were used as references. In the first step, recombination and error rate maps were estimated using 500 subjects randomly selected from the GWAS data. In the second step, genotype imputation of all subjects was conducted using the rate maps estimated in the first step. We excluded the imputed SNPs with MAF < 0.01 or *Rsq* values < 0.7, and obtained genotype data of 2,178,018 SNPs.

In the replication study sets 1 and 2, we used genotyping data which were performed using the Illumina HumanHap550v3 Genotyping BeadChip and the Illumina HumanOmniExpress Genotyping BeadChip (Illumina, CA, USA), respectively. We applied the same quality control criteria and imputation procedure as GWAS data. Details of the genotyping, quality control, imputation procedure in the replication set 3 are described elsewhere[17].

Statistical Analysis

Genome-wide association study and the replication studies of BMI—Rank-based inverse normal transformation was applied to BMI of the subjects. In the GWAS, associations of the SNPs with transformed values of BMI were assessed by linear regression assuming additive effects of allele dosages (bound between 0.0 and 2.0) using mach2qtl software, and genomic control correction was applied[35]. In the regression model, gender, age, age-squared, smoking history, the affection statuses of the diseases, and the demographic classifications of the medical institutes in Japan where the subjects were enrolled[36], were adopted as covariates. For the loci that satisfied $P < 5.0 \times 10^{-5}$ in the GWAS, replication studies were conducted consisting of three replication sets (Supplementary Table 1 and Supplementary Figure 1). In replication sets 1 and 2, the associations of the SNPs were assessed in the same manner as in the GWAS. In replication set 3, we referred the results of the discovery stage of the concurrently conducted genome-wide meta analysis of BMI[17]. The combined results of the studies were obtained using an inverse-variance method from the summary statistics β and the standard error (SE). Details of examination of tag CNVs and expression analysis in the *KLF9* locus using publicly available database for HapMap Phase II East Asian individuals[34,37] are described in Supplementary Figure 4. Associations of the SNPs that satisfied $P < 5.0 \times 10^{-5}$ in the combined study of the GWAS and the replication studies were further evaluated using the results of the meta-analysis for BMI in European populations by the GIANT consortium[12]. For the evaluation of the associations in the previously reported BMI associated loci[3–12,18], the loci that exhibited FDR < 0.05 based on the number of loci reported with non-monomorphic SNPs were considered to be significant. The statistical power of the study was estimated using Quanto version 1.2.4.

The inter-individual variance in BMI explained by each of the identified loci ($P < 5.0 \times 10^{-8}$ in the combined study) was estimated using $2f(1-f)\beta^2$, where f is the frequency of the variant in HapMap East Asian populations and β is its additive effect size on the BMI obtained from the replication studies. To estimate the variance explained by the combination of the identified loci, we calculated the genetic risk scores for the subjects in the GWAS, by summing the dosages of BMI-increasing alleles carried by the subjects, weighted by the effect sizes of the SNPs obtained from the replication studies. The explained variance was estimated from a linear regression model incorporating the score as the predictor and the covariate-adjusted inverse normal transformed BMI residuals as outcome.

Gene–gene interaction analysis of BMI—Gene–gene interactions of SNPs were evaluated using a multivariate linear regression model assuming additive × additive effects of two SNPs[20]. Allele dosages of the respective SNPs and the product of the allele dosages were involved in the model in addition to the covariates. The product of the allele dosages was denoted as interaction term. For each of the landmark SNPs in the loci confirmed to be associated with BMI, gene–gene interactions were evaluated with all of the genome-wide SNPs (7 × 2,178,018 SNP pairs; Supplementary Figure 5), and genomic control corrections were applied[35]. For SNP pairs that demonstrated $P < 5.0 \times 10^{-6}$ for the interaction term, replication studies using replication sets 1 and 2 were performed. The SNP pair that satisfied $P < 5.0 \times 10^{-8}$ in the combined study of GWAS and replication studies was considered to be significant.

Association study of metabolic traits and other related ones—Associations with obesity (BMI ≥ 27.5[38]; 3,058 cases and 31,472 controls), type 2 diabetes (T2D; 6,526 cases and 22,689 controls), systolic and diastolic blood pressure (n = 13,049), total cholesterol (TC; n = 12,565), high density lipoprotein cholesterol (HDL-C; n = 4,924), low density lipoprotein cholesterol (LDL-C; n = 4,219), and triglyceride (TG; n = 9,747) were evaluated using the subjects enrolled in the GWAS and the replication sets 1 and 2 (Supplementary Table 6). In addition to the two novel loci associated with BMI (*CDKAL1* and *KLF9*), we assessed the *GIPR* locus, where the associations with T2D and its related traits have been reported[25]. Case-control analysis and analyses of the quantitative traits were performed using logistic and linear regression models including the covariates, respectively. In the association analysis of T2D, subjects not affected with cardiovascular diseases were enrolled as controls, and BMI was additionally incorporated as a covariate.

R statistical software was used for the general analysis. Details of the study design are also indicated in Supplementary Figure 1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Kazuyuki Tobe and Minoru Iwata at the First Department of Internal Medicine, Faculty of Medicine, Toyama University, and Hiroshi Hirose at Health Center, Keio University School of Medicine, and all the staff of the Laboratory for Endocrinology and Metabolism, and Statistical Analysis at CGM, RIKEN for their assistance. This study was supported by the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1. Kopelman PG. Obesity as a medical problem. Nature. 2000; 404:635–643. [PubMed: 10766250]
2. Maes HH, Neale MC, Eaves LJ. Genetic and environmental factors in relative body weight and human adiposity. Behav Genet. 1997; 27:325–351. [PubMed: 9519560]

3. Frayling TM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007; 316:889–894. [PubMed: 17434869]
4. Liu YJ, et al. Genome-wide association scans identified CTNNBL1 as a novel gene for obesity. Hum Mol Genet. 2008; 17:1803–1813. [PubMed: 18325910]
5. Chambers JC, et al. Common genetic variation near MC4R is associated with waist circumference and insulin resistance. Nat Genet. 2008; 40:716–718. [PubMed: 18454146]
6. Loos RJ, et al. Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nat Genet. 2008; 40:768–775. [PubMed: 18454148]
7. Thorleifsson G, et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nat Genet. 2009; 41:18–24. [PubMed: 19079260]
8. Willer CJ, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009; 41:25–34. [PubMed: 19079261]
9. Meyre D, et al. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. Nat Genet. 2009; 41:157–159. [PubMed: 19151714]
10. Liu XG, et al. Genome-wide association and replication studies identified TRHR as an important gene for lean body mass. Am J Hum Genet. 2009; 84:418–423. [PubMed: 19268274]
11. Cho YS, et al. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nat Genet. 2009; 41:527–534. [PubMed: 19396169]
12. Speliotes EK, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010; 42:937–948. [PubMed: 20935630]
13. Deurenberg P, Deurenberg-Yap M, Guricci S. Asians are different from Caucasians and from each other in their body mass index/body fat per cent relationship. Obes Rev. 2002; 3:141–146. [PubMed: 12164465]
14. Nakamura Y. The BioBank Japan Project. Clin Adv Hematol Oncol. 2007; 5:696–697. [PubMed: 17982410]
15. Okada Y, et al. Genome-wide association study for C-reactive protein levels identified pleiotropic associations in the IL6 locus. Hum Mol Genet. 2011; 20:1224–1231. [PubMed: 21196492]
16. Freedman ML, et al. Assessing the impact of population stratification on genetic association studies. Nat Genet. 2004; 36:388–393. [PubMed: 15052270]
17. Wen W, et al. Meta-analysis identifies common variants associated with body mass index in East Asians. In submission.
18. Zobel DP, et al. Variation in the gene encoding Kruppel-like factor 7 influences body fat: studies of 14 818 Danes. Eur J Endocrinol. 2009; 160:603–609. [PubMed: 19147600]
19. Hinney A, Vogel CI, Hebebrand J. From monogenic to polygenic obesity: recent advances. Eur Child Adolesc Psychiatry. 2010; 19:297–310. [PubMed: 20127379]
20. Cordell HJ. Detecting gene-gene interactions that underlie human diseases. Nat Rev Genet. 2009; 10:392–404. [PubMed: 19434077]
21. Winkler C, et al. BMI at age 8 years is influenced by the type 2 diabetes susceptibility genes HHEX-IDE and CDKAL1. Diabetes. 2010; 59:2063–2067. [PubMed: 20460429]
22. Andersson EA, et al. Type 2 diabetes risk alleles near ADCY5, CDKAL1 and HHEX-IDE are associated with reduced birthweight. Diabetologia. 2010; 53:1908–1916. [PubMed: 20490451]
23. Steinthorsdottir V, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat Genet. 2007; 39:770–775. [PubMed: 17460697]
24. Yamauchi T, et al. A genome-wide association study in the Japanese population identifies susceptibility loci for type 2 diabetes at UBE2E2 and C2CD4A-C2CD4B. Nat Genet. 2010; 42:864–868. [PubMed: 20818381]
25. Saxena R, et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nat Genet. 2010; 42:142–148. [PubMed: 20081857]
26. Pei H, Yao Y, Yang Y, Liao K, Wu JR. Kruppel-like factor KLF9 regulates PPARgamma transactivation at the middle stage of adipogenesis. Cell Death Differ. 2011; 18:315–327. [PubMed: 20725087]
27. Kadowaki T, Yamauchi T. Adiponectin and adiponectin receptors. Endocr Rev. 2005; 26:439–451. [PubMed: 15897298]

28. Oishi Y, et al. Kruppel-like transcription factor KLF5 is a key regulator of adipocyte differentiation. Cell Metab. 2005; 1:27–39. [PubMed: 16054042]
29. Elkasrawy MN, Hamrick MW. Myostatin (GDF-8) as a key factor linking muscle mass and bone structure. J Musculoskelet Neuronal Interact. 2010; 10:56–63. [PubMed: 20190380]
30. Schuelke M, et al. Myostatin mutation associated with gross muscle hypertrophy in a child. N Engl J Med. 2004; 350:2682–2688. [PubMed: 15215484]
31. Grade CV, Salerno MS, Schubert FR, Dietrich S, Alvares LE. An evolutionarily conserved Myostatin proximal promoter/enhancer confers basal levels of transcription and spatial specificity in vivo. Dev Genes Evol. 2009; 219:497–508. [PubMed: 20052486]
32. Wada K, et al. Validity of self-reported height and weight in a Japanese workplace population. Int J Obes (Lond). 2005; 29:1093–1099. [PubMed: 15925952]
33. Nakamura K, Hoshino Y, Kodama K, Yamamoto M. Reliability of self-reported body height and weight of adult Japanese women. J Biosoc Sci. 1999; 31:555–558. [PubMed: 10581882]
34. The International HapMap Consortium. The International HapMap Project. Nature. 2003; 426:789–796. [PubMed: 14685227]
35. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55:997–1004. [PubMed: 11315092]
36. Yamaguchi-Kabata Y, et al. Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. Am J Hum Genet. 2008; 83:445–456. [PubMed: 18817904]
37. Stranger BE, et al. Population genomics of human gene expression. Nat Genet. 2007; 39:1217–1224. [PubMed: 17873874]
38. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004; 363:157–163. [PubMed: 14726171]

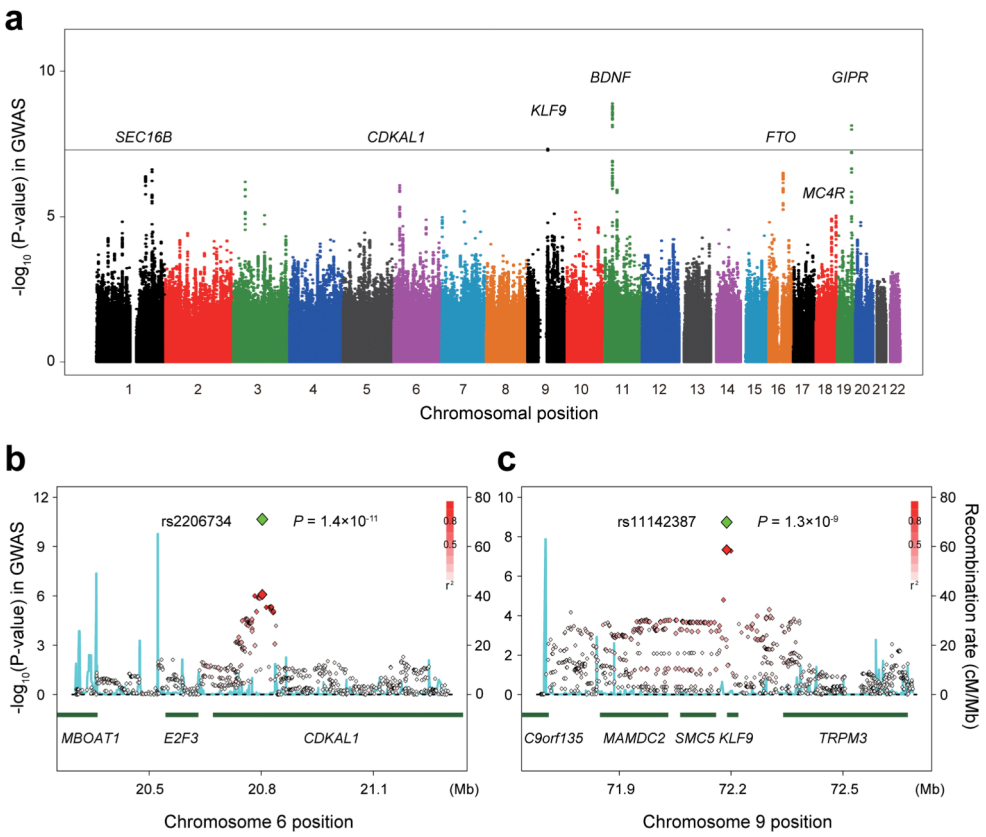

Figure 1.

Results of the genome-wide association study (GWAS) for BMI. **(a)** Manhattan plot showing the $-\log_{10}$ (P-values) of the SNPs in the GWAS for BMI in 26,620 Japanese subjects. The genetic loci that satisfied the genome-wide significance threshold of $P < 5.0 \times 10^{-8}$ in the combined study of the GWAS and the replication studies are labeled. The gray horizontal line represents the threshold of $P = 5.0 \times 10^{-8}$. Regional plots of the SNPs **(b)** in the *CDKAL1* locus and **(c)** in the *KLF9* locus. The red diamond-shaped dots represent $-\log_{10}$ (P-values) of the SNPs in the GWAS, and the green dots represent the P-value of the most significantly associated SNP in each of the loci in the combined study. The density of the red color in the small-sized dots represents the r^2 value with the most significantly associated SNP of the large-sized red dot. The blue line shows the recombination rates given by the HapMap Phase II East Asian populations (release 22). The lower part indicates the RefSeq genes in the loci.

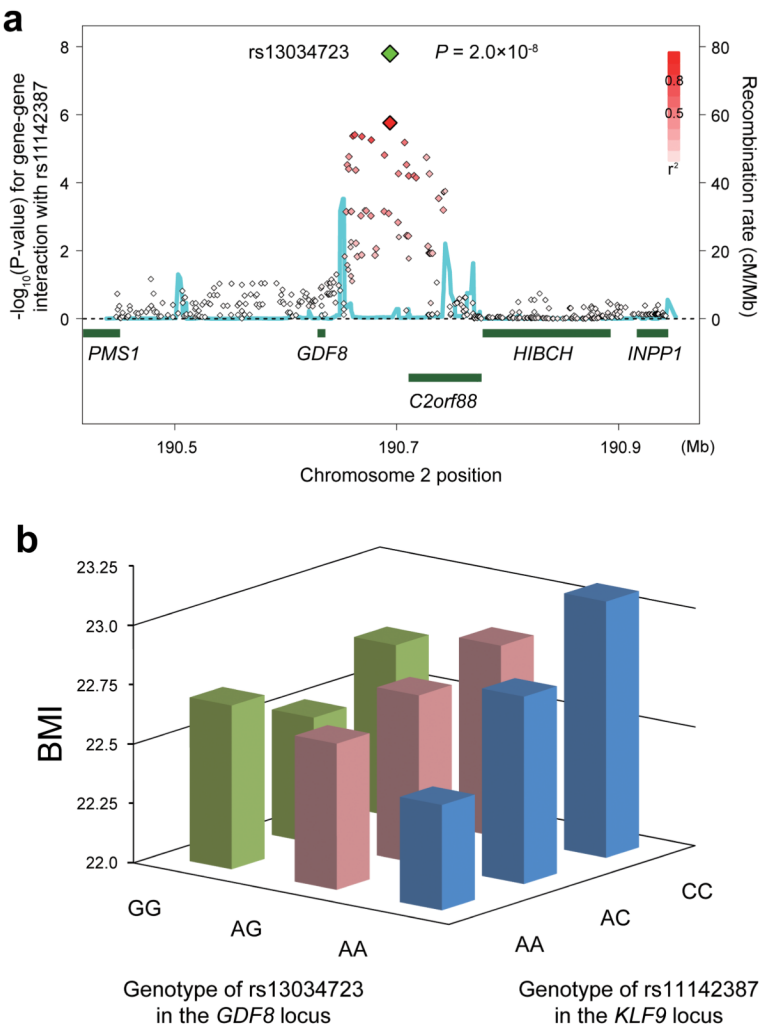

Figure 2.

Gene–gene interaction between the *KLF9* and *GDF8* loci. **(a)** Regional plots of the SNPs. Diamond-shaped dots represent $-\log_{10}$ (P-values) of the SNPs for gene–gene interaction with the landmark SNP in the *KLF9* locus (rs11142387). The green dot indicates the P-value of the most significantly associated SNP in the combined study, and the red dot indicates its P-value in the genome-wide gene–gene interaction analysis. The density of the red color in the small-sized dots represents the r^2 value with the most significantly associated SNP of the large-sized red dot. The blue line shows the recombination rates given by the HapMap database. The lower part indicates the RefSeq genes in the locus. **(b)** Mean BMI values of the subjects stratified with the genotypes of rs13034723 in the *GDF8* locus and rs11142387 in the *KLF9* locus.

Table 1

Associations of the GWAS and the replication studies for BMI.

rsID[a]	Chr	Position	Cyto band	Nearest Gene	Class	A1/A2[b]	East Asian populations								European populations (GIANT consortium)[g]	
							GWAS			Replication study[e]		Combined		Explained variance[f]		
							Freq.[c]	Beta (SE)[d]	*P*	Beta (SE)[d]	*P*	Beta (SE)[d]	*P*		Beta (SE)[d]	*P*
Significantly associated SNPs ($P < 5.0 \times 10^{-8}$)																
rs12149832	16	52,400,409	16q12	*FTO*	intron	A/G	0.20	0.056 (0.011)	3.2×10^{-7}	0.090 (0.011)	5.1×10^{-17}	0.073 (0.008)	4.8×10^{-22}	0.20%	0.077 (0.005)	5.6×10^{-58}
rs2030323	11	27,685,115	11p14	*BDNF*	intron	C/A	0.60	0.054 (0.008)	1.3×10^{-9}	0.040 (0.008)	1.8×10^{-7}	0.046 (0.006)	3.8×10^{-16}	0.08%	0.042 (0.006)	5.7×10^{-13}
rs11671664	19	50,864,118	19q13	*GIPR*	intron	G/A	0.45	0.051 (0.008)	7.4×10^{-9}	0.041 (0.009)	5.6×10^{-6}	0.046 (0.006)	6.8×10^{-14}	0.08%	0.029 (0.009)	0.0012
rs2206734	6	20,802,863	6p22	*CDKAL1*	intron	C/T	0.59	0.043 (0.008)	8.3×10^{-7}	0.035 (0.008)	6.2×10^{-6}	0.039 (0.006)	1.4×10^{-11}	0.06%	0.017 (0.006)	0.0049
rs2331841	18	55,979,617	18q21	*MC4R*	intergenic	A/G	0.25	0.045 (0.011)	1.2×10^{-5}	0.047 (0.009)	1.9×10^{-7}	0.046 (0.007)	1.8×10^{-11}	0.08%	0.035 (0.005)	1.2×10^{-13}
rs11142387	9	72,188,152	9q21	*KLF9*	intergenic	C/A	0.46	0.048 (0.008)	4.6×10^{-8}	0.028 (0.011)	0.0084	0.040 (0.007)	1.3×10^{-9}	0.04%	0.003 (0.005)	0.50
rs516636	1	176,122,140	1q25	*SEC16B*	intergenic	A/C	0.22	0.053 (0.011)	4.2×10^{-7}	0.044 (0.014)	0.0014	0.050 (0.008)	3.4×10^{-9}	0.07%	0.023 (0.027)	0.40
SNPs with suggestive associations ($5.0 \times 10^{-8} \leq P < 5.0 \times 10^{-5}$)																
rs4377469	3	42,278,078	3p22	*CCK*	intron	T/G	0.69	0.050 (0.010)	6.4×10^{-7}	0.022 (0.012)	0.058	0.039 (0.007)	1.6×10^{-7}	-	0.018 (0.033)	0.58
rs10993160	9	96,108,747	9q22	*ZNF169*	intergenic	A/G	0.83	0.061 (0.014)	7.9×10^{-6}	0.041 (0.016)	0.012	0.053 (0.011)	5.5×10^{-7}	-	0.025 (0.012)	0.035

[a] SNPs that satisfied $P < 5.0 \times 10^{-5}$ in the combined study are indicated.

[b] The allele that increased BMI is denoted as allele 1 and is indicated based on forward strand and NCBI Build 36.

[c] Frequency of allele 1.

[d] Effect size of allele 1 on the normalized BMI (mean = 0, standard deviation = 1).

[e] Combined results of three independent replication sets (Supplementary Fig. 1).

[f] Estimated based on the effect sizes in the replication studies and the allele frequencies in HapMap East Asian populations.

[g] Referenced using the results of the genome-wide meta-analysis for BMI in European populations[12].

BMI, body mass index; GWAS, Genome-wide association study; SE, standard error.

Table 2

Gene–gene interaction for BMI between the *KLF9* and *GDF8* loci

Independent variables in the regression model				Associations with BMI in the regression model					
				GWAS		Replication study[c]		Combined	
rsID	Cytoband	Gene	A1/A2[a]	Beta (SE)[b]	*P*	Beta (SE)[b]	*P*	Beta (SE)[b]	*P*
rs11142387[d]	9q21	*KLF9*	C/A	0.088 (0.012)	8.3×10^{-14}	0.117 (0.027)	1.7×10^{-5}	0.093 (0.011)	1.1×10^{-17}
rs13034723[d]	2q32	*GDF8*	A/G	−0.065 (0.015)	2.8×10^{-5}	−0.068 (0.029)	0.018	−0.065 (0.014)	1.5×10^{-6}
rs11142387 × rs13034723[e]	-	-	-	0.064 (0.013)	1.7×10^{-6}	0.073 (0.025)	0.0031	0.066 (0.012)	2.0×10^{-8}

[a]Based on forward strand and NCBI Build 36.

[b]Effect size of allele 1 on the normalized BMI (mean = 0, standard deviation = 1).

[c]Consisted of replication sets 1 and 2 (Supplementary Fig. 1).

[d]Allele dosage of allele 1 was used as an independent variable.

[e]Product of allele dosages of alleles 1 of rs11142387 and rs13034723 was used as an independent variable.

BMI, body mass index; GWAS, Genome-wide association study; SE, standard error.