OBESITY STUDIES

A Candidate-Gene Approach Identifies Novel Associations Between Common Variants in/Near Syndromic Obesity Genes and BMI in Pediatric and Adult European Populations

Dominic X. Wang,[1,2] Yuvreet Kaur,[1,3] Akram Alyass,[1] and David Meyre[1,4]

Diabetes 2019;68:724–732 | https://doi.org/10.2337/db18-0986

We hypothesized that monogenic syndromic obesity genes are also involved in the polygenic variation of BMI. Single-marker, tag single nucleotide polymorphism (tagSNP) and gene-based analysis were performed on common variants near 54 syndromic obesity genes. We used publicly available data from meta-analyses of European BMI genome-wide association studies conducted by the Genetic Investigation of ANthropometric Traits (GIANT) Consortium and the UK Biobank (UKB) (*N* = 681,275 adults). A total of 33 loci were identified, of which 19 of 33 (57.6%) were located at SNPs previously identified by the GIANT Consortium and UKB meta-analysis, 11 of 33 (33.3%) were located at novel SNPs, and 3 of 33 (9.1%) were novel genes identified with gene-based analysis. Both single-marker and tagSNP analyses mapped the previously identified 19 SNPs by the GIANT Consortium and UKB meta-analysis. Gene-based analysis confirmed 15 of 19 (78.9%) of the novel SNPs' associated genes. Of the 11 novel loci, 8 were identified with single-marker analysis and the remaining 3 were identified with tagSNP analysis. Gene-based analysis confirmed 4 of 11 (36.3%) of these loci. Meta-analysis with the Early Growth Genetics (EGG) Consortium (*N* = 35,668 children) was conducted post hoc for top SNPs, confirming 17 of 33 (51.5%) loci, of which 5 were novel. This study supports evidence for a continuum between rare monogenic syndromic and common polygenic forms of obesity.

Obesity has become a global epidemic, with the most recent data from the World Health Organization estimating a worldwide prevalence of 650 million cases in adults. It is a well-established risk factor for numerous comorbidities as well as increased mortality (1). With 65–80% of new cases being accounted for by overweight or obesity, excessive body weight is the main risk factor for incident type 2 diabetes (2). To date, most prevention programs have emphasized an "eat less, move more" global approach, but this has had limited success (3). Numerous behavioral (e.g., diet, exercise) and pharmacological treatments have been developed and implemented clinically, but have had only modest effects (4). Currently, bariatric surgery is the most effective treatment for severe obesity, but it is an invasive procedure with numerous complications and only a small fraction of eligible patients are referred (5). A better understanding of the underlying causes of obesity may be useful in developing more effective prediction, prevention, and treatment (6).

Genetics accounts for 40–75% of BMI interindividual variability (7). Recent estimates suggest that single nucleotide polymorphisms (SNPs) alone may explain up to 30–40% of BMI variance (8). Genome-wide association studies (GWAS) have identified a substantial number of SNPs associated with BMI (9). These results have helped determine causal genes and significantly expanded our understanding of obesity pathophysiology (6).

The Genetic Investigation of ANthropometric Traits (GIANT) Consortium and the UK Biobank (UKB) recently released results from the largest BMI GWAS meta-analysis to date, which identified 941 near-independent SNPs,

[1]Department of Health Research, Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
[2]Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario, Canada
[3]Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada
[4]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario, Canada

Corresponding author: David Meyre, meyred@mcmaster.ca

Received 12 September 2018 and accepted 24 January 2019

This article contains Supplementary Data online at http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db18-0986/-/DC1.

© 2019 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at http://www.diabetesjournals.org/content/license.

accounting for 6% of BMI variance (9). Although gene prioritization analysis was conducted, the individual genes were not identified. Furthermore, only single-marker analysis was conducted, causing a substantial reduction of statistical power because of the multiple testing burden (10). Power can be increased by *1*) targeting SNPs only at genes with strong pre-existing evidence for a role in obesity, *2*) conducting tagSNP analysis, and *3*) conducting gene-based analysis (10).

Although obesity in the general population is polygenic, it can also present with monogenic nonsyndromic mutations or monogenic syndromic mutations. Monogenic syndromic forms of obesity present with striking obesity along with features such as intellectual disability, dysmorphic features, and organ-specific abnormalities (11). These forms of obesity are often considered separate entities (6). However, accumulating evidence suggests that these forms exist on a continuum. For example, many loci at monogenic nonsyndromic obesity genes have also been associated with polygenic obesity (e.g., *POMC*, *PCSK1*, *MC4R*, and *ADCY3*) (6). Furthermore, several SNPs located in and near syndromic obesity genes have been reported by previous GWAS for BMI: *BDNF*, *NTRK2*, *SIM1*, *BBS2*, *BBS4*, *SH2B1*, and *SDCCAG8* (9). We hypothesized that additional syndromic obesity genes were associated with polygenic variation of BMI.

The aim of this study was to identify loci associated with BMI variation at highly relevant candidate genes missed by hypothesis-free GWAS. We recently published a systematic review describing 79 genetic syndromes with obesity, of which 30 have been fully or partially elucidated genetically (11). We conducted single-marker, tagSNP, and gene-based analysis of these identified genes to investigate the role of syndromic obesity genes in the most recent BMI GWAS meta-analysis conducted by the GIANT Consortium and UKB, as well as meta-analysis conducted by the Early Growth Genetics (EGG) Consortium (9,12).

RESEARCH DESIGN AND METHODS

Sample Population

We used publicly available data from the GIANT Consortium and UKB European meta-analysis, which consisted of European cohorts only (N = 681,275; participants on average per SNP) (9). We also used publicly available data from the EGG Consortium meta-analysis of European children in a post hoc meta-analysis (N = 35,668) (12).

Gene and SNP Selection

Sixty-three genes had been previously identified in a systematic review of obesity syndromes (11). Nine genes were excluded since they were located on the X chromosome, which was not included in the GIANT Consortium and UKB meta-analysis (9). See Supplementary Table 1 for a summarized list of the syndromes and the associated genes.

The summary statistics of all SNPs within 200 kb upstream and downstream of the remaining 54 genes were extracted from files released online by the GIANT Consortium using the R package biomaRt (13). GWAS-significant SNPs have been shown to have the greatest number of significant pathway-phenotype associations when using a gene definition of ±200 kb compared with others between 0 and 1,500 kb (14). Sensitivity analyses using different gene boundaries were conducted using 50, 20, 10, and 0 kb.

Yengo et al. (9) selected and analyzed 711,933 SNPs overlapping the GIANT Consortium and UKB SNP sets after imputation and quality control. Briefly, Yengo et al. (9) analyzed a subset of SNPs with consistent alleles between the GIANT Consortium and UKB and after further elimination using a linkage disequilibrium (LD) threshold of $r^2 > 0.9$. To increase power, we further pruned SNPs by selecting tagSNPs, variants that can infer nearby SNPs because of limited historical recombination in these regions (15). We selected tagSNPs genotyped in the largest sample from SNPs with $r^2 > 0.8$ calculated from 1000 Genomes Project Consortium phase 1 European data based on greatest sample size, then greatest minor allele frequency, then greatest number of SNPs tagged, then selecting the SNP closest to the gene, and finally by random selection (16,17).

Data Analysis

The GIANT Consortium transformed BMI residuals using an inverse standard normal function in linear regression models after adjustment for age, age squared, and/or other study-specific covariates (18). In the UKB, linear mixed model association testing was conducted under an infinitesimal model and adjusted for age, sex, recruitment center, genotyping batches, and the top 10 principal components.

All gene-based tests in this study were applied using Versatile Gene-based test for Genome-wide Association Studies 2 (VEGAS2) (19). VEGAS2 uses a simulation approach to calculate gene-based empirical association P values by pooling P values of individual SNP associations to χ^2 statistics after correcting for LD (19). Since a 200-kb gene definition was not an available option, the largest alternative was used: 50 kb.

Meta-analysis was performed using a random-effects model in PLINK version 1.90 software. Study weights were generated using the inverse variance method. Heterogeneity between studies was assessed using the χ^2 test for homogeneity, the I^2 statistic (20).

The Bonferroni correction was applied to all tests (α = 0.05). Applying an experiment-wide Bonferroni correction reduces the chance of making type I errors but increases the chance of making type II errors. Therefore, we followed the strategy reported previously by Feise (21) and considered independent Bonferroni corrections for each question asked. We considered each type of analysis (single-marker, tagSNP, or gene-based analysis) using a specific gene definition as separate questions (e.g., single-marker analysis with gene boundaries of ±200 kb). A total of 15 questions

was tested, of which 3 are the primary hypotheses of the study, 11 are sensitivity analyses, and 1 is the meta-analysis with the EGG Consortium cohort.

LocusZoom was used to create regional association plots (22). LDlink was used to determine LD between SNPs (using r^2 values), where 1000 Genomes Project Consortium phase 3 (version 5) European data were used (23).

Fine-Mapping

To further delineate the potential causative SNPs, functional annotation of SNPs was conducted using the Ensembl Variant Effect Predictor (VEP) tool (https://useast.ensembl.org/Tools/VEP). VEP predicts potential functional consequences of variants using the position of each variant. We tested all fine-mapping SNPs, which were defined as lead signals from single-marker, tagSNP, and gene-based analysis as well as SNPs in moderate LD ($r^2 > 0.5$) (24), which had also been tested by Yengo et al. (9).

RESULTS

Single-Marker Analysis

Using ±200 kb gene boundaries, 18,140 SNPs were extracted from 54 genes (Supplementary Tables 1 and 2). A total of 997 of 18,140 SNPs (5.5%) located at 27 loci in/near 29 genes were significantly associated with BMI ($P < 2.80 \times 10^{-6}$). Of these SNPs, 38 of 997 (3.8%) were in/near eight novel loci: *MRAP2*, *BBS1*, *GNAS*, *IFT27*, *LZTFL1*, *NIPBL*, *MAGEL2/MKRN3/NDN*, and *BBS9* (Table 1 and Supplementary Table 2).

To test for independence from previously identified loci, the r^2 value was determined between top SNPs at novel loci and top SNPs at loci previously identified in the study by Yengo et al. (9) as well as genome-wide significant SNPs identified for BMI in the GWAS Catalog (25). The top SNP at *BBS1* (rs12805133) was in LD with previously identified SNPs: rs682842, which was previously identified at *C11orf80* (r^2 = 0.564) (9). There was no LD ($r^2 > 0.1$) between any of the other top SNPs at novel loci and previously identified SNPs.

Of the remaining 959 of 997 significant SNPs, 520 of 997 (52.2%) were GWAS significant ($P < 5 \times 10^{-8}$) and were located at the following 19 genes: *BDNF*, *SH2B1*, *CREBBP*, *BBS4*, *KCTD13*, *SDCCAG8*, *BBS2*, *SIM1*, *NTRK2*, *IFT172*, *EHMT1*, *KMT2D*, *SOX10*, *WDPCP*, *MKS1*, *EP300*, *FGFR1*, *RAB23*, and *RAI1* (Supplementary Table 2). The remaining 439 of 997 (44.0%) SNPs were mapped to the 19 genes with genome-wide significant loci.

tagSNP Analysis

Of the SNPs identified with single-marker analysis, 4,942 of 18,140 (27.2%) were selected as tagSNPs (Supplementary Table 3). A total of 224 of 4,942 (4.5%) tagSNPs were significantly associated with BMI ($P < 1.01 \times 10^{-5}$) of which 3 of 224 (1.3%) were novel association signals in/near *BBS12*, *TUB*, and *RAD21* (Table 2). None of these SNPs were in LD ($r^2 > 0.1$) with previously identified SNPs. The remaining 221 of 224 tagSNPs were mapped to the same 27 loci identified in single-marker analysis (Supplementary Table 3).

Gene-Based Analysis

All 54 selected genes were analyzed with VEGAS2, and 22 of 54 (40.7%) were significantly associated with BMI ($P < 9.26 \times 10^{-4}$) (Supplementary Table 4). Three novel genes were identified: *GHR*, *PROKR2*, and *CHD2* (Table 3). Top SNPs at these loci were not in LD ($r^2 > 0.1$) with previously identified SNPs. Of the remaining genes, 15 of 54 (27.8%) were gene candidates for GWAS-significant SNPs identified in single-marker and tagSNP analysis, whereas 4 of 54 (7.4%) were candidates for novel SNPs identified in single-marker analysis: *BBS1*, *GNAS*, *IFT27*, and *LZTFL1* (Supplementary Table 4).

Sensitivity Analyses

Single-marker, tagSNP, and gene-based analyses were repeated using different gene boundaries: ±0, ±10, and ±20 kb, as well as ±50 kb for single-marker and tagSNP analyses. SNPs at novel genes were not identified, although tagSNP analysis using ±10- and ±20-kb gene definitions identified SNPs at loci that had been identified previously by gene-based analysis (Supplementary Table 5).

Meta-Analysis with the EGG Consortium Cohort

We conducted a meta-analysis of the top SNPs from single-marker, tagSNP, and gene-based analyses from the GWAS meta-analysis by the EGG Consortium of BMI in children, investigating the pooled results of the GIANT Consortium and the UKB as well as the meta-analyses of the EGG Consortium (Table 4). Of the SNPs investigated, 17 of 33 (51.5%) were significantly associated ($P < 1.52 \times 10^{-3}$): *BDNF*, *SH2B1*, *KCTD13*, *CREBBP*, *BBS4*, *SIM1*, *BBS2*, *EP300*, *SOX10*, *WDPCP*, and *RAB23*. Five of those 17 SNPs were in/near novel loci: *MRAP2*, *BBS9*, *BBS12*, *GHR*, and *TUB*. Heterogeneity between the studies for these SNPs was not significant ($I^2 < 5$).

Fine-Mapping

Fine-mapping SNPs were identified and VEP analysis was conducted for 715 variants (Supplementary Table 6). Most SNPs had multiple predicted functions, including nonsense-mediated decay transcript variants and/or noncoding transcript variants. The majority were intronic (501 of 715; 70.1%) and intergenic variants (113 of 715; 15.8%). Others were related to regulatory function, with 73 of 715 (10.2%) at regulatory regions, 22 of 715 (3.1%) at 3′ untranslated regions, and 5 of 15 (0.7%) at 5′ untranslated regions. Several SNPs were identified as coding variants, with 10 of 715 (1.4%) identified as missense variants and 1 of 715 (0.1%) identified as a stop-gained variant. Of the 10 missense variants, 1 was identified for *BDNF* (rs6265), 2 for *SH2B1* (rs7498665, rs2904880), 1 for *BBS4* (rs2277598), 1 for *SIM1* (rs240780), 1 for *BBS2* (rs11373), 3 for *BBS1*

Table 1—The top significant SNPs of their associated gene from single-marker analysis ($P < 2.8 \times 10^{-6}$)

SNP	Chr:position	Gene	Alleles	EAF	β	SE	*P*	*N*
rs6265	11:27679916	*BDNF*	T/C	0.1951	−0.0412	0.0021	1.0×10^{-86}	795,458
rs7498665	16:28883241	*SH2B1*	A/G	0.5962	−0.0271	0.0017	5.6×10^{-60}	790,299
rs3814883	16:29994922	*KCTD13*	T/C	0.4764	0.0232	0.0017	1.1×10^{-40}	685,519
rs879620	16:4015729	*CREBBP*	T/C	0.6179	0.0231	0.0018	5.3×10^{-38}	688,377
rs7164727	15:73093991	*BBS4*	T/C	0.681	0.0182	0.0017	3.3×10^{-25}	791,156
rs946824	1:243684019	*SDCCAG8*	T/C	0.141	0.0206	0.0026	1.1×10^{-15}	689,849
rs12448738	16:56489343	*BBS2*	A/C	0.8629	−0.0168	0.0025	2.8×10^{-11}	691,932
rs12206564	6:100987009	*SIM1*	T/C	0.505	−0.0113	0.0017	5.4×10^{-11}	689,836
rs1187352	9:87293457	*NTRK2*	T/C	0.3482	−0.0119	0.0018	6.0×10^{-11}	688,522
rs1260326	2:27730940	*IFT172*	T/C	0.4027	−0.0105	0.0017	3.9×10^{-10}	784,462
rs11792069	9:140646121	*EHMT1*	A/G	0.83	0.0145	0.0024	6.5×10^{-10}	644,252
rs7975791	12:49413486	*KMT2D*	T/C	0.03601	−0.0264	0.0043	1.1×10^{-9}	770,139
rs12628891	22:38317137	*SOX10*	T/C	0.3169	−0.0112	0.0019	3.0×10^{-9}	686,575
rs12478556	2:63341711	*WDPCP*	A/T	0.4177	0.0103	0.0018	4.5×10^{-9}	683,709
rs11649864	17:56093061	*MKS1*	A/G	0.09127	0.0178	0.0031	6.7×10^{-9}	663,638
rs139531	22:41676176	*EP300*	A/G	0.7197	0.0112	0.0019	9.5×10^{-9}	690,810
rs881301	8:38332318	*FGFR1*	T/C	0.5818	−0.0097	0.0017	2.4×10^{-8}	691,753
rs6901944	6:56862805	*RAB23*	A/G	0.8344	−0.0126	0.0023	2.7×10^{-8}	692,035
rs6910117	**6:84971679**	***MRAP2***	**A/T**	**0.9099**	**−0.0157**	**0.0029**	$\mathbf{9.1 \times 10^{-8}}$	**643,225**
rs12805133	**11:66483265**	***BBS1***	**A/G**	**0.5426**	**−0.0091**	**0.0018**	$\mathbf{3.5 \times 10^{-7}}$	**635,477**
rs6026567	**20:57444915**	***GNAS***	**A/G**	**0.571**	**−0.0092**	**0.0018**	$\mathbf{4.4 \times 10^{-7}}$	**659,708**
rs9610560	**22:37064378**	***IFT27***	**A/G**	**0.7375**	**−0.0099**	**0.002**	$\mathbf{5.1 \times 10^{-7}}$	**671,468**
rs17765088	**3:45943595**	***LZTFL1***	**C/G**	**0.8817**	**−0.0136**	**0.0027**	$\mathbf{5.7 \times 10^{-7}}$	**666,309**
rs13171414	**5:37232079**	***NIPBL***	**A/G**	**0.7836**	**0.0103**	**0.0021**	$\mathbf{7.2 \times 10^{-7}}$	**692,376**
rs11161347	**15:23948049**	***MAGEL2/MKRN3/NDN***	**A/G**	**0.4831**	**0.0085**	**0.0018**	$\mathbf{1.1 \times 10^{-6}}$	**672,682**
rs13311608	**7:32999601**	***BBS9***	**A/G**	**0.5147**	**−0.0083**	**0.0017**	$\mathbf{1.3 \times 10^{-6}}$	**688,219**
rs752579	17:17660347	*RAI1*	T/C	0.6077	0.0082	0.0017	1.6×10^{-6}	781,905

Novel SNPs are shown in bold. Chromosome (Chr):position, Alleles (effect/other), effect allele frequency (EAF), beta (β), SE, *P* values, and number of observations (*N*) are reported. SNP positions are reported according to Build 37. See Supplementary Table 2 for the complete results from single-marker analysis.

(rs618838, rs540874, rs1671064), and 1 for *RAB23* (rs3800023). The stop-gained variant was also identified for *BBS1* (rs1815739).

DISCUSSION

We analyzed highly relevant candidate genes linked to syndromic obesity in publicly available data from the latest meta-analysis of BMI by the GIANT Consortium and UKB. Single-marker, tagSNP, and gene-based analyses revealed 14 novel associations at 16 of 54 (29.6%) candidate genes, and genome-wide significant SNPs were mapped to 19 of 54 (35.2%) syndromic obesity genes. A significant association for 17 of 33 (51.5%) loci was also observed with BMI in children.

Our data confirm the importance of conducting candidate gene approaches to identify SNPs with modest genetic effects missed by conventional hypothesis-free GWAS. As an illustration, the average effect size of the 11 novel SNPs identified in this study on BMI was two times lower in comparison with the 941 nearly independent SNPs identified in the most recent GWAS for BMI (β = 0.0105 [95% CI 0.0089, 0.0122] vs. β = 0.0205 [95% CI 0.0195, 0.0215] (9).

Biological Role of the Associated Genes

Most of the identified genes are related to leptin signaling and primary cilia function in neurons. Leptin signaling produces an anorexigenic effect by activating proopiomelanocortin (POMC)-expressing neurons, whereas dysfunction in primary cilia causes ciliopathies, which are well-established causes of complex syndromes that often present with obesity (6,11).

Eleven of the identified genes are related to the BBSome: *BBS4*, *BBS2*, *MKS1*, *WDPCP*, *SDCCAG8*, *IFT172*, *BBS1*, *BBS9*, *BBS12*, *LZTFL1*, and *IFT27*. The BBSome is a protein complex in primary cilia with a key role in leptin receptor (LepR) trafficking (26). Mutations in these genes cause Bardet-Biedl syndrome, which is characterized by truncal obesity as

Table 2—SNPs from tagSNP analysis associated with genes that were not identified by single SNP analysis ($P < 1.01 \times 10^{-5}$)

SNP	Chr:position	Gene	Alleles	EAF	β	SE	*P*	*N*
rs3891400	4:123736349	*BBS12*	T/G	0.8777	−0.0123	0.0027	6.4×10^{-6}	688,812
rs7484168	11:8316829	*TUB*	A/C	0.5499	−0.0078	0.0018	8.5×10^{-6}	684,328
rs12675038	8:117706218	*RAD21*	T/C	0.8855	0.0113	0.0025	8.9×10^{-6}	794,582

Chromosome (Chr):position, Alleles (effect/other), effect allele frequency (EAF), beta (β), SE, *P* values, and number of observations (*N*) are reported. SNP positions are reported according to Build 37. See Supplementary Table 3 for the complete results from tagSNP analysis.

well as intellectual disability, retinal dystrophy, polydactyly, microorchidism in men, and renal malformations (11). Postnatal disruption of the primary cilia function in POMC neurons of mice results in hyperphagia and obesity (27). Previous evidence has demonstrated a potential role for the BBSome in common obesity, as higher BMI in heterozygous carriers of rare mutations in BBS genes without additional clinical features has been observed (28).

TUB and *RAB23* may also be related to neuronal cilia. Mutations in *TUB*, which encodes tubby, produce a phenotype similar to that caused by mutations in the BBSome, including obesity and retinopathy (11). Tubby has demonstrated a critical role in trafficking G-protein–coupled receptors in neuronal cilia (29). The protein product of *RAB23* also has a role in trafficking in primary cilia (30).

MRAP2, *MAGEL2*, *NDN*, and *GNAS* are also related to the leptin signaling pathway. Mutations in *MRAP2* alter melanocortin receptor 4 (MC4R) signaling, a receptor downstream of POMC (31). Although large deletions involving *MRAP2* have been reported in cases of syndromic obesity, pathogenic variants also have been identified in several cases of early onset obesity without additional features (11,31). Moreover, Magel2 (encoded by *MAGEL2*) promotes the cell surface presence of LepR (32). LepR subsequently interacts with necdin, encoded by *NDN* (32). POMC neurons in Magel2-deficient mice have impaired leptin response, which may be caused by reduced LepR trafficking to the cell surface (33). Finally, *GNAS* encodes G-protein α subunits, including the G-protein α subunits integral in MC4R (34). Mutations in *SOX10*, *PROKR2*, and *FGFR1* may also affect leptin signaling via their effect on the migration of gonadotropin-releasing hormone neurons. Gonadotropin-releasing hormone is modulated by leptin and LepR, and inactivating mutations in leptin and LepR causes an absence of pubertal development as well as hypogonadotropic hypogonadism (35).

The *KCTD13* locus lies within the well-described 16p11.2 interval, along with *SH2B1*. Deletion of the region is associated with macrocephaly, hyperphagic obesity, psychosis, epilepsy, schizophrenia, and autism spectrum disorder (ASD) (11). Although *SH2B1* has been identified as the primary driver for the changes in feeding behaviors and body weight, *KCTD13* has been identified as the primary candidate for the brain-related phenotypes (36,37). However, *KCTD13* interacts with ciliopathy-associated genes, and *KCTD13* knockdown-induced macrocephaly can be rescued by *BBS7* overexpression (38).

SIM1 has been linked to obesity through its critical role in the development of key components of the hypothalamic-pituitary axis (39). Some evidence suggests that SIM1 may be downstream of MC4R signaling. For example, the effect of a melanocortin receptor agonist that reduces food intake in wild-type mice is blunted in Sim1 heterozygous mice (40).

Another pathway heavily implicated in obesity is the brain-derived neurotrophic factor (BDNF)-tyrosine-related kinase B (TrkB) axis. *BDNF* encodes BDNF, which binds to the TrkB receptor encoded by *NTRK2* to produce an anorexigenic effect (41). Strong evidence has implicated the BDNF-TrkB axis in common obesity (42). There is also an interaction between the BDNF-TrkB pathway and the leptin signaling pathway, as the coadministration of a TrkB antagonist attenuates the anorexigenic effects of leptin (43).

Kleefstra, Kabuki, and Smith-Magenis syndromes are linked to mutations in *EHMT1*, *KMT2D*, and *RAI1*, respectively (11). Individuals with any of these three syndromes can present with ASD as well as obesity. Pathway analysis has demonstrated potential functional convergence, and these syndromes may be involved in common and/or overlapping pathways (44). One of those pathways may involve the BDNF-TrkB axis. *EHMT1*, a histone methyltransferase, has been shown to mediate *BDNF* expression during synaptic scaling (45). However, the observed association with BMI in our study may result from an

Table 3—Significant genes from gene-based analysis with VEGAS2 that had not been not identified by single-marker and tagSNP analysis ($P < 9.26 \times 10^{-4}$)

Gene	nSNPs	nSims	Chr:start–stop	*P*	Top SNP	Top SNP *P* value
PROKR2	174	1,000,000	20:5232685–5345015	1.90×10^{-4}	rs6116750	1.90×10^{-5}
CHD2	180	1,000,000	15:93393550–93621237	1.91×10^{-4}	rs1881837	2.50×10^{-5}
GHR	267	1,000,000	5:42373876–42771980	4.04×10^{-4}	rs11950813	1.90×10^{-5}

See Supplementary Table 4 for the complete results from gene-based analysis. Chr, chromosome; nSNP, number of SNPs used for analysis.

Table 4—Meta-analysis of the UKB and GIANT Consortium and the EGG Consortium with a random-effects model of the top SNPs of genes identified by single-marker, tagSNP, and gene-based analyses in the EGG Consortium

SNP	Chr:position	Gene	Alleles	β	*P*	I^2
rs6265	**11:27636492**	***BDNF***	**T/C**	**−0.0415**	**1.14E−90**	**0**
rs7498665	**16:28790742**	***SH2B1***	**A/G**	**−0.0272**	**4.39E−60**	**0**
rs3814883	**16:29902423**	***KCTD13***	**T/C**	**0.0231**	**1.78E−43**	**0**
rs879620	**16:3955730**	***CREBBP***	**T/C**	**0.0233**	**8.09E−40**	**0**
rs7164727	**15:70881044**	***BBS4***	**T/C**	**0.0183**	**5.98E−28**	**0**
rs12206564	**6:101093730**	***SIM1***	**T/C**	**−0.0115**	**4.86E−12**	**0**
rs12448738	**16:55046844**	***BBS2***	**A/C**	**−0.0166**	**1.28E−11**	**0**
rs139531	**22:40006122**	***EP300***	**A/G**	**0.0115**	**6.52E−10**	**0**
rs12628891	**22:36647083**	***SOX10***	**T/C**	**−0.0113**	**1.38E−09**	**0**
rs12478556	**2:63195215**	***WDPCP***	**A/T**	**0.0102**	**6.14E−09**	**0**
rs6901944	**6:56970764**	***RAB23***	**A/G**	**−0.013**	**7.87E−09**	**0**
rs6910117	**6:85028398**	***MRAP2***[a]	**A/T**	**−0.0153**	**8.77E−08**	**0**
rs1187352	**9:86483277**	***NTRK2***	**T/C**	**−0.0113**	**2.72E−07**	**4.69**
rs13311608	**7:32966126**	***BBS9***[a]	**A/G**	**−0.0083**	**5.85E−07**	**0**
rs3891400	**4:123955799**	***BBS12***[b]	**T/G**	**−0.0119**	**7.29E−06**	**0**
rs11950813	**5:42778662**	***GHR***[c]	**A/G**	**0.0083**	**8.44E−06**	**0**
rs7484168	**11:8273405**	***TUB***[b]	**A/C**	**−0.0076**	**1.68E−05**	**0**
rs946824	1:241750642	*SDCCAG8*	T/C	0.0174	5.87E−03	34.7
rs11792069	9:139765942	*EHMT1*	A/G	0.0125	1.15E−02	21.87
rs881301	8:38451475	*FGFR1*	T/C	−0.0131	1.72E−02	51.45
rs13171414	5:37267836	*NIPBL*[a]	A/G	0.0145	3.06E−02	52.97
rs9610560	22:35394324	*IFT27*[a]	A/G	−0.0083	3.71E−02	21.08
rs752579	17:17601072	*RAI1*	T/C	0.0169	9.22E−02	84.53
rs7975791	12:47699753	*KMT2D*	T/C	−0.0146	4.00E−01	60.08
rs12675038	8:117775399	*RAD21*[b]	T/C	0.006	4.63E−01	55.33
rs6116750	20:5251842	*PROKR2*[c]	T/C	0.0037	5.39E−01	56.01
rs1881837	15:91266191	*CHD2*[c]	A/G	−0.0041	6.26E−01	62.2
rs12805133	11:66239841	*BBS1*[a]	A/G	0.005	7.40E−01	92.07
rs17765088	3:45918599	*LZTFL1*[a]	C/G	−0.004	7.60E−01	69.18
rs11161347	15:21499142	*MAGEL2/MKRN3/NDN*[a]	A/G	−0.003	8.15E−01	88.91
rs1260326	2:27584444	*IFT172*	T/C	0.0034	8.19E−01	92.65
rs6026567	20:56878310	*GNAS*[a]	A/G	−0.0017	8.57E−01	77.9
rs11649864	17:53448060	*MKS1*	A/G	0.0028	8.78E−01	81.01

Significant SNPs ($P < 1.52 \times 10^{-3}$) are shown in bold. Chr, chromosome. Heterogeneity between studies was assessed using the I^2 statistic. [a]Novel genes identified by single-marker analysis. [b]Novel genes identified by tagSNP analysis. [c]Novel genes identified by gene-based analysis.

interaction with ASD, which has been associated with a higher BMI (46). Mutations in *CHD2* have also been associated with ASD as well as intellectual disability and photosensitive epilepsy. However, only one case report (11) has described obesity with a *CHD2* mutation. Two patients harbored a 15q26.1 microdeletion of *CHD2* and *RGMA*, and exhibited truncal obesity (11).

The remaining genes identified in this study relate to various syndromes whose underlying mechanisms are not well understood. Mutations in cohesin-related proteins rad21 and nipbl cause Cornelia de Lange syndrome (CdLS), which presents with characteristic facial features and variations of weight (11,47). However, none of the 23 reported cases of CdLS with Rad21 mutations are obese (47). Among the genes tested for in a suspected diagnosis of CdLS are *CREBBP* and *EP300*, which are causative genes for a similar disorder: Rubinstein-Taybi syndrome (RTS). RTS is characterized by features including growth restriction, microcephaly, and broad thumbs and toes (11). Case series have demonstrated that obesity presents in 30–40% of patients with RTS caused by mutations in *CREBBP* or *EP300* (48). Finally, the deficiency of *GHR* produces Laron

syndrome, which is characterized by abnormal growth with child-like proportions in adulthood (e.g., greater deviation of stature than head size). Obesity is observed in early childhood and progresses into adulthood as total body fat reaches 30–60% of weight (11). This is not attributed to increased caloric intake or reduced energy expenditure (49).

Study Implications

Our results confirm a large continuum between monogenic syndromic, monogenic nonsyndromic, oligogenic, and polygenic obesity (6). Substantial evidence, including candidate gene studies as well as independent GWAS, has previously suggested a role for many of the genes identified in this study in polygenic obesity (6).

In addition to the candidate-gene single-marker analysis, tagSNP and gene-based approaches were also conducted to explore the sensitivity of our analyses after further reduction of the Bonferroni correction. Although the original analysis conducted by the GIANT Consortium and UKB had pruned SNPs in high LD ($r^2 > 0.9$), additional pruning ($r^2 > 0.8$) in our tagSNP analysis identified all 27 loci from single-marker analysis as well as three additional SNPs at novel loci. Future GWAS may identify additional SNPs by using a lower threshold for LD. Furthermore, gene-based analysis rescued three loci where no individual SNP was significantly associated with BMI, but rather several variants had a moderate effect on BMI. We recommend that future GWAS systematically conduct gene-based analysis in tandem with classic single-marker analysis (50).

This study defined gene boundaries as ±200 kb. We found that this boundary captured the greatest number of significant SNPs after Bonferroni correction compared with boundaries of ±50, ±20, ±10, and ±0 kb. Although larger gene boundaries were not tested, this finding is consistent with a separate study that found the greatest number of significant pathway-phenotype associations for GWAS-significant SNPs when using a gene definition of ±200 kb compared with others between 0 and 1,500 kb (14).

Some of the included syndromic obesity genes in this study have not been replicated or confirmed by a second independent study (see Supplementary Table 1) yet have been previously confirmed as GWAS significant ($P < 5 \times 10^{-8}$): *NTRK2*, *SH2B1*, and *SDCCAG8*. Therefore, there is no direct correlation between the strength of known evidence for the role of a gene in a genetic obesity syndrome and its effects on common BMI variation. However, this study provides some additional support for a role for these genes in genetic obesity syndromes. This is particularly impactful for genes that have not been replicated or confirmed: *SH2B1*, *SDCCAG8*, and *NTRK2*, as well as *LZTFL1*, *TUB*, and *CHD2*.

Meta-analysis with the EGG Consortium cohort demonstrated significant signals in more than half of the top SNPs from single-marker, tagSNP, and gene-based analysis (17 of 33; 51.5%) after Bonferroni correction ($P < 1.52 \times 10^{-3}$). Since the EGG Consortium cohort was composed of children rather than adults, significant genetic associations reflect those that play a role in anthropometry throughout life. In contrast, SNPs that did not demonstrate a significant association all showed significant heterogeneity between the studies ($I^2 > 20$). This may suggest age-dependent effects for these SNPs. However, some SNPs showing significant heterogeneity are located in/near genes that have been previously associated with childhood obesity syndromes (e.g., *EHMT1* in Kleefstra syndrome, *GNAS* in Allbright Hereditary Osteodystrophy) or have shown previous evidence of associations with obesity in children and adolescents (e.g., *SDCCAG8*).

Strengths and Limitations

This study demonstrates that candidate gene studies can be highly successful, with 64.8% of the candidate genes (35 of 54) demonstrating significant associations and 51.5% of significant loci (17 of 33) showing association after meta-analysis with a separate pediatric cohort. Two main components of this study differ significantly from most candidate-gene studies: *1*) the selected candidate genes have strong evidence for a causal role in obesity (i.e., clinically evident cases of morbid obesity with evidence of drastic genetic mutations); and *2*) we used a very well-powered BMI data set (up to 681,275 participants). Other candidate gene studies often use indirect evidence such as transcriptome and methylome analysis, which do not strongly distinguish obesity as a cause or consequence (10). We recommend that future well-powered GWAS analyses supplement hypothesis-free analysis with candidate-gene approaches using highly causal evidence, including novel syndromic obesity genes. Other potential sources of highly causal evidence include mouse models of obesity (6).

The genes analyzed in this study are likely a limited selection of potential candidates since only a proportion of the known genetic syndromes with obesity have been fully genetically elucidated (19 of 79; 24%) (11). Some have only been partially genetically elucidated (11 of 79; 14%), and there may also be syndromes that have yet to be identified (11). Also, neither rare variants nor potential long-range associations were included in these analyses. These genes were identified in a population of European descent only. Therefore, transferability studies in other ethnic groups are needed to investigate the generalizability of these results. Finally, as with all GWAS, the genes identified in this study are candidates, and the SNPs identified may only be proxies for the causative SNP. Other potential causative genes may lie within the proximity of the identified SNPs. The causative SNPs may lie within the LD blocks of the identified SNPs.

Conclusion

In summary, this study identified novel common variants associated with BMI in a large sample of children and adults of European ancestry using a combination of candidate-gene, single-marker, tagSNP, and gene-based approaches.

This provides additional support to growing evidence of a continuum between monogenic syndromic obesity and common polygenic obesity.

Acknowledgments. The authors thank the members of the GIANT Consortium, the UKB, and the EGG Consortium for their data and participation.
Funding. D.M. is supported by a Canada Research Chair in Genetics of Obesity.
Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. D.X.W. designed the study, analyzed the data, wrote the paper and had responsibility for final content, and read and approved the final manuscript. Y.K. wrote the paper and had responsibility for final content, and read and approved the final manuscript. A.A. analyzed the data, and read and approved the final manuscript. D.M. designed the study, wrote the paper and had responsibility for final content, and read and approved the final manuscript. D.M. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Data Availability. The GIANT Consortium and the UKB have freely provided their data at http://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files. The EGG Consortium also provided their data freely at https://egg-consortium.org/.

References

1. Heymsfield SB, Wadden TA. Mechanisms, pathophysiology, and management of obesity. N Engl J Med 2017;376:254–266
2. Mokdad AH, Ford ES, Bowman BA, et al. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. JAMA 2003;289:76–79
3. Stice E, Shaw H, Marti CN. A meta-analytic review of obesity prevention programs for children and adolescents: the skinny on interventions that work. Psychol Bull 2006;132:667–691
4. Peirson L, Douketis J, Ciliska D, Fitzpatrick-Lewis D, Ali MU, Raina P. Treatment for overweight and obesity in adult populations: a systematic review and meta-analysis. CMAJ Open 2014;2:E306–E317
5. Kral JG, Kava RA, Catalano PM, Moore BJ. Severe obesity: the neglected epidemic. Obes Facts 2012;5:254–269
6. Pigeyre M, Yazdi FT, Kaur Y, Meyre D. Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of human obesity. Clin Sci (Lond) 2016;130:943–986
7. Stryjecki C, Alyass A, Meyre D. Ethnic and population differences in the genetic predisposition to human obesity. Obes Rev 2018;19:62–80
8. Yang J, Bakshi A, Zhu Z, et al.; LifeLines Cohort Study. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. Nat Genet 2015;47:1114–1120
9. Yengo L, Sidorenko J, Kemper KE, et al.; GIANT Consortium. Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. Hum Mol Genet 2018;27:3641–3649
10. Li A, Meyre D. Challenges in reproducibility of genetic association studies: lessons learned from the obesity field. Int J Obes 2013;37:559–567
11. Kaur Y, de Souza RJ, Gibson WT, Meyre D. A systematic review of genetic syndromes with obesity. Obes Rev 2017;18:603–634
12. Felix JF, Bradfield JP, Monnereau C, et al.; Bone Mineral Density in Childhood Study (BMDCS); Early Genetics and Lifecourse Epidemiology (EAGLE) Consortium; Early Growth Genetics (EGG) Consortium; Bone Mineral Density in Childhood Study BMDCS. Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. Hum Mol Genet 2016;25:389–403
13. Durinck S, Moreau Y, Kasprzyk A, et al. BioMart and bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics 2005;21:3439–3440
14. Brodie A, Azaria JR, Ofran Y. How far from the SNP may the causative genes be? Nucleic Acids Res 2016;44:6046–6054
15. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science 2008;322:881–888
16. Abecasis GR, Auton A, Brooks LD, et al.; 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. Nature 2012;491:56–65
17. Spencer CC, Su Z, Donnelly P, Marchini J. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. PLoS Genet 2009;5:e1000477
18. Locke AE, Kahali B, Berndt SI, et al.; LifeLines Cohort Study; ADIPOGen Consortium; AGEN-BMI Working Group; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; International Endogene Consortium. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197–206
19. Mishra A, Macgregor S. VEGAS2: software for more flexible gene-based Testing. Twin Res Hum Genet 2015;18:86–91
20. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–1558
21. Feise RJ. Do multiple outcome measures require p-value adjustment? BMC Med Res Methodol 2002;2:8
22. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336–2337
23. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics 2015;31:3555–3557
24. van de Bunt M, Cortes A, Brown MA, Morris AP, McCarthy MI; IGAS Consortium. Evaluating the performance of fine-mapping strategies at common variant GWAS loci. PLoS Genet 2015;11:e1005535
25. MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI catalog of published genome-wide association studies (GWAS catalog). Nucleic Acids Res 2017;45:D896–D901
26. Seo S, Guo DF, Bugge K, Morgan DA, Rahmouni K, Sheffield VC. Requirement of Bardet-Biedl syndrome proteins for leptin receptor signaling. Hum Mol Genet 2009;18:1323–1331
27. Davenport JR, Watts AJ, Roper VC, et al. Disruption of intraflagellar transport in adult mice leads to obesity and slow-onset cystic kidney disease. Curr Biol 2007;17:1586–1594
28. Croft JB, Morrell D, Chase CL, Swift M. Obesity in heterozygous carriers of the gene for the Bardet-Biedl syndrome. Am J Med Genet 1995;55:12–15
29. Sun X, Haley J, Bulgakov OV, Cai X, McGinnis J, Li T. Tubby is required for trafficking G protein-coupled receptors to neuronal cilia. Cilia 2012;1:21
30. Leaf A, Von Zastrow M. Dopamine receptors reveal an essential role of IFT-B, KIF17, and Rab23 in delivering specific receptors to primary cilia. eLife 2015;4:e06996
31. Asai M, Ramachandrappa S, Joachim M, et al. Loss of function of the melanocortin 2 receptor accessory protein 2 is associated with mammalian obesity. Science 2013;341:275–278
32. Wijesuriya TM, De Ceuninck L, Masschaele D, et al. The Prader-Willi syndrome proteins MAGEL2 and necdin regulate leptin receptor cell surface abundance through ubiquitination pathways. Hum Mol Genet 2017;26:4215–4230
33. Oncul M, Dilsiz P, Ates Oz E, et al. Impaired melanocortin pathway function in Prader-Willi syndrome gene-Magel2 deficient mice. Hum Mol Genet 2018;27:3129–3136
34. Chen M, Shrestha YB, Podyma B, et al. Gsα deficiency in the dorsomedial hypothalamus underlies obesity associated with Gsα mutations. J Clin Invest 2017;127:500–510
35. Strobel A, Issad T, Camoin L, Ozata M, Strosberg AD. A leptin missense mutation associated with hypogonadism and morbid obesity. Nat Genet 1998;18:213–215
36. Doche ME, Bochukova EG, Su HW, et al. Human SH2B1 mutations are associated with maladaptive behaviors and obesity [published correction appears in J Clin Invest 2013;123:526]. J Clin Invest 2012;122:4732–4736

37. Golzio C, Willer J, Talkowski ME, et al. KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. Nature 2012; 485:363–367

38. Migliavacca E, Golzio C, Männik K, et al.; 16p11.2 European Consortium. A potential contributory role for ciliary dysfunction in the 16p11.2 600 kb BP4-BP5 pathology. Am J Hum Genet 2015;96:784–796

39. Michaud JL, Rosenquist T, May NR, Fan CM. Development of neuroendocrine lineages requires the bHLH-PAS transcription factor SIM1. Genes Dev 1998;12: 3264–3275

40. Kublaoui BM, Holder JL Jr, Gemelli T, Zinn AR. Sim1 haploinsufficiency impairs melanocortin-mediated anorexia and activation of paraventricular nucleus neurons. Mol Endocrinol 2006;20:2483–2492

41. Soppet D, Escandon E, Maragos J, et al. The neurotrophic factors brain-derived neurotrophic factor and neurotrophin-3 are ligands for the trkB tyrosine kinase receptor. Cell 1991;65:895–903

42. Speliotes EK, Willer CJ, Berndt SI, et al.; MAGIC; Procardis Consortium. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 2010;42:937–948

43. Spaeth AM, Kanoski SE, Hayes MR, Grill HJ. TrkB receptor signaling in the nucleus tractus solitarius mediates the food intake-suppressive effects of hindbrain BDNF and leptin. Am J Physiol Endocrinol Metab 2012;302:E1252–E1260

44. Mullegama SV, Alaimo JT, Chen L, Elsea SH. Phenotypic and molecular convergence of 2q23.1 deletion syndrome with other neurodevelopmental syndromes associated with autism spectrum disorder. Int J Mol Sci 2015;16:7627–7643

45. Benevento M, Iacono G, Selten M, et al. Histone methylation by the Kleefstra syndrome protein EHMT1 mediates homeostatic synaptic scaling. Neuron 2016; 91:341–355

46. Zheng Z, Zhang L, Li S, et al. Association among obesity, overweight and autism spectrum disorder: a systematic review and meta-analysis. Sci Rep 2017; 7:11697

47. Kline AD, Moss JF, Selicorni A, et al. Diagnosis and management of Cornelia de Lange syndrome: first international consensus statement. Nat Rev Genet 2018; 19:649–666

48. Fergelot P, Van Belzen M, Van Gils J, et al. Phenotype and genotype in 52 patients with Rubinstein-Taybi syndrome caused by EP300 mutations. Am J Med Genet A 2016;170:3069–3082

49. Ginsberg S, Laron Z, Bed MA, Vaisman N. The obesity of patients with Laron Syndrome is not associated with excessive nutritional intake. Obes Res Clin Pract 2009;3:1–52

50. Hägg S, Ganna A, Van Der Laan SW, et al.; GIANT Consortium. Gene-based meta-analysis of genome-wide association studies implicates new loci involved in obesity. Hum Mol Genet 2015;24:6849–6860