Human Molecular Genetics, 2013, Vol. 22, No. 17 ***3597–3607***
doi:10.1093/hmg/ddt205
Advance Access published on May 12, 2013

Genome-wide analysis of BMI in adolescents and young adults reveals additional insight into the effects of genetic loci over the life course

Mariaelisa Graff[1,2,*,†], Julius S. Ngwa[3,22,†], Tsegaselassie Workalemahu[4,†], Georg Homuth[5], Sabine Schipf[6], Alexander Teumer[5], Henry Völzke[6], Henri Wallaschofski[7], Goncalo R. Abecasis[8], Lakatta Edward[9], Cucca Francesco[10], Serena Sanna[10], Paul Scheet[11], David Schlessinger[9], Carlo Sidore[8,9], Xiangjun Xiao[11], Zhaoming Wang[13,14], Stephen J. Chanock[13], Kevin B. Jacobs[13,14], Richard B. Hayes[15], Frank Hu[4,16], Rob M. Van Dam[17,4], The GIANT Consortium[12,‡], Richard J. Crout[18], Mary L. Marazita[19], John R Shaffer[20], Larry D. Atwood[21,22], Caroline S. Fox[21], Nancy L. Heard-Costa[21,22], Charles White[3], Audrey C. Choh[23], Stefan A. Czerwinski[23], Ellen W. Demerath[24], Thomas D. Dyer[25], Bradford Towne[23], Najaf Amin[26], Ben A. Oostra[26], Cornelia M. Van Duijn[26,28], M. Carola Zillikens[27], Tõnu Esko[29,30,31], Mari Nelis[29,31,32], Tit Nikopensius[30], Andres Metspalu[29,30,31], David P. Strachan[33], Keri Monda[1,34], Lu Qi[4,†], Kari E. North[1,35,†], L. Adrienne Cupples[3,22,†], Penny Gordon-Larsen[2,36,†] and Sonja I. Berndt[13,†]

[1]Department of Epidemiology and [2]Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA [3]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA [4]Department of Nutrition, Harvard School of Public Health, Boston, MA, USA [5]Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany [6]Institute for Community Medicine and [7]Institute for Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany [8]Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA [9]Intramural Research Program, National Institute on Aging, Baltimore, MD, USA [10]Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Cagliari, Italy [11]Department of Epidemiology, University of Texas M. D. Anderson, Houston, TX, USA [12]The Genetic Investigation of ANthropometric Traits (GIANT), [13]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA [14]Core Genotyping Facility, SAIC-Frederick, Inc., NCI-Frederick, Frederick, MD 21702, USA [15]New York University Medical Center, New York, NY 10016, USA [16]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA [17]Saw Swee Hock School of Public Health and Yong Loo Lin School of Medicine, National University of Si ngapore, Singapore [18]Department of Periodontics, West Virginia University, Morgantown, WV, USA [19]Center for Craniofacial and Dental Genetics, Department of Oral Biology and [20]Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA, USA [21]NHLBI's Framingham Heart Study, Framingham, MA, USA [22]Department of Neurology, Boston University School of Medicine, Boston, MA, USA [23]Wright State University School of Medicine, Dayton, OH, USA [24]School of Public Health, University of Minnesota, Minneapolis, MN, USA [25]Department of Genetics, Texas Biomedical Research Institute, Houston, TX, USA [26]Department of Epidemiology and [27]Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands [28]Centre of Medical Systems Biology, Leiden, the Netherlands [29]Estonian Genome Center and [30]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia [31]Estonian Biocentre, Tartu, Estonia [32]Department of Medical Genetics and Development, University of Geneva Medical School, Geneva, Switzerland [33]Division of Population

[*]To whom correspondence should be addressed at: Department of Epidemiology, University of North Carolina, Bank of America Center, 137 East Franklin Str., Suite 306, Chapel Hill, NC 27517, USA. Tel: 919-966-8491; Fax: 919-966-9800; Email: migraff@email.unc.edu
[†]These authors contributed jointly to this work.
[‡]See Acknowledgements in the Supplementary Material.

Health Sciences and Education, St George's University of London, London, UK [34]The Center for Observational Research, Amgen, Inc., Thousand Oaks, CA, USA [35]Carolina Center for Genome Sciences and [36]Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Received March 13, 2013; Revised April 18, 2013; Accepted May 3, 2013

Genetic loci for body mass index (BMI) in adolescence and young adulthood, a period of high risk for weight gain, are understudied, yet may yield important insight into the etiology of obesity and early intervention. To identify novel genetic loci and examine the influence of known loci on BMI during this critical time period in late adolescence and early adulthood, we performed a two-stage meta-analysis using 14 genome-wide association studies in populations of European ancestry with data on BMI between ages 16 and 25 in up to 29 880 individuals. We identified seven independent loci ($P < 5.0 \times 10^{-8}$) near *FTO* ($P = 3.72 \times 10^{-23}$), *TMEM18* ($P = 3.24 \times 10^{-17}$), *MC4R* ($P = 4.41 \times 10^{-17}$), *TNNI3K* ($P = 4.32 \times 10^{-11}$), *SEC16B* ($P = 6.24 \times 10^{-9}$), *GNPDA2* ($P = 1.11 \times 10^{-8}$) and *POMC* ($P = 4.94 \times 10^{-8}$) as well as a potential secondary signal at the *POMC* locus (rs2118404, $P = 2.4 \times 10^{-5}$ after conditioning on the established single-nucleotide polymorphism at this locus) in adolescents and young adults. To evaluate the impact of the established genetic loci on BMI at these young ages, we examined differences between the effect sizes of 32 published BMI loci in European adult populations (aged 18–90) and those observed in our adolescent and young adult meta-analysis. Four loci (near *PRKD1*, *TNNI3K*, *SEC16B* and *CADM2*) had larger effects and one locus (near *SH2B1*) had a smaller effect on BMI during adolescence and young adulthood compared with older adults ($P < 0.05$). These results suggest that genetic loci for BMI can vary in their effects across the life course, underlying the importance of evaluating BMI at different ages.

INTRODUCTION

The period of adolescence and young adulthood is recognized as a period of elevated risk for excess weight gain (1–3). In the past 5 years, genome-wide association studies (GWASs) have identified over 30 common genetic loci associated with body mass index (BMI, kg/m^2) mainly in European adult samples, with an average age often greater than 50 (4–10). Whereas several loci identified in adults have also been found to be associated with BMI in childhood (6,11–20) and two loci were recently identified in childhood obesity (17), little is known about these obesity susceptibility loci across high-risk periods for weight gain, such as adolescence and young adulthood. The influence of these loci in adolescence and young adulthood remains largely speculative from previously established association studies that illustrate the association of these loci on BMI during middle-aged adulthood and/or childhood.

The purpose of the current study was to conduct a two-stage GWAS for genetic loci influencing BMI during late adolescence and early adulthood (aged 16–25). Furthermore, we sought to compare estimates of effect sizes on BMI for the 32 BMI loci previously identified in European middle-aged adults (4) to effect sizes observed in adolescent and young adults of European descent.

RESULTS

In the discovery meta-analysis of 10 GWASs, we observed an excess of small *P*-values compared with chance alone, an excess that was somewhat diminished when the known BMI loci were removed (Supplementary Material, Fig. S1a, Q–Q plot). Although our initial analyses were stratified by gender, Q–Q plots of the observed versus expected *P*-values for heterogeneity between men and women suggested similar results by gender; thus, all results are presented for the combined sample. A total of four independent loci (as defined as separated by at least 1 Mb), rs9940128 (near *FTO*), rs12463617 (near *TMEM18*), rs7234864 (near *MC4R*) and rs12142020 (near *TNNI3K*), reached genome-wide significance ($P < 5.0 \times 10^{-8}$) in the discovery sample (Supplementary Material, Fig. S1b), and after filtering the results for the single-nucleotide polymorphisms (SNPs) using a distance criteria of $\pm$ 500 kb and linkage disequilibrium threshold of $r^2 < 0.1$, the top 76 SNPs with $P < 5 \times 10^{-5}$ were taken forward for follow-up in four studies. From the combined meta-analysis of the discovery and follow-up results, seven independent loci ($P < 5.0 \times 10^{-8}$; Table 1 and Supplementary Material, Fig. S2) near *FTO* (rs9940128, $P = 3.72 \times 10^{-23}$), *TMEM18* (rs12463617, $P = 3.24 \times 10^{-17}$), *MC4R/PMAIP1* (rs7234864 $< P = 4.41 \times 10^{-17}$), *TNNI3K* (rs12142020, $P = 4.32 \times 10^{-11}$), *SEC16B* (rs591120, $P = 6.24 \times 10^{-9}$), *GNPDA2* (rs13130484, $P = 1.11 \times 10^{-8}$) and *POMC* (rs1561288, $P = 4.94 \times 10^{-8}$) reached genome-wide statistical significance. There was little heterogeneity between the studies for all seven SNPs (Supplementary Material, Fig. S3a–g, $P_{\text{heterogeneity}} \geq 0.06$). All seven SNPs fell within 1 Mb of previously established loci, and five of the seven were highly correlated ($r^2 > 0.7$) with the reported SNP from the Genetic Investigation of ANThropometric Traits (GIANT) consortium (4). Two of the seven SNPs (rs591120 near *SEC16B* and rs1561288 near *POMC*) and a second possible signal near *MC4R* (rs17066846) were only weakly correlated ($r^2 \leq 0.3$) with the previously published loci (4). We performed conditional analyses within a 1 Mb region on either side of these three SNPs to identify additional independent SNPs.

Results for meta-analysis of region 1 (chromosome 1, 175.8–176.8 Mb, captures *SEC16B*) conditioned on our most

Table 1. Loci that reach genome-wide significance ($P < 5 \times 10^{-8}$) in the joint analysis of discovery and follow-up studies in young adults and adolescents

SNP information		Discovery sample				Follow-up sample		Discovery and follow-up		Discovery and follow-up						Previously identified loci in adults[a] and D' (R^2)	
SNP name	Chr	Position	Closest gene	BMI increasing/ other allele	FEA	*P*-value	*n*	*P*-value	*n*	Beta[b]	SE[b]	Estimated change in BMI (kg/m^2) per effect allele	*P*-value	*n*	I^2 (het *P*-value)	SNP name	D' (R^2)[c]
rs9940128	16	52358255	*FTO*	A/G	0.43	6.96E−12	13 503	8.29E−13	15 443	0.083	0.008	0.33	3.72E−23	28 946	3.5 (0.41)	rs1558902	1.0 (0.90)
rs12463617	2	619244	*TMEM18*	C/A	0.85	3.95E−13	13 356	2.20E−06	15 446	0.100	0.012	0.40	3.24E−17	28 802	24.9 (0.14)	rs2867125	1.0 (0.92)
rs7234864	18	55885837	*MC4R, PMAIP1*	T/C	0.26	2.40E−12	13 506	8.00E−07	15 444	0.081	0.010	0.32	4.41E−17	28 950	23.5 (0.14)	rs571312	0.87 (0.73)
rs12142020	1	74772599	*TNNI3K*	T/A	0.41	5.52E−12	12 237	6.13E−03	15 445	0.056	0.009	0.22	4.32E−11	27 682	14.6 (0.25)	rs1514175	0.93 (0.80)
rs591120	1	176169376	*SEC16B*	C/G	0.44	3.37E−05	12 238	5.49E−04	15 444	0.050	0.009	0.19	6.24E−09	27 682	31.5 (0.06)	rs543874	1.0 (0.31)
rs13130484	4	44870448	*GNPDA2*	T/C	0.43	1.42E−06	13 607	5.56E−05	15 436	0.048	0.009	0.19	1.11E−08	29 043	0 (0.80)	rs10938397	1.0 (1.0)
rs1561288	2	25222506	*POMC*	C/T	0.78	8.23E−07	13 562	4.38E−03	15 436	0.055	0.010	0.22	4.94E−08	28 998	26.1 (0.11)	rs713586	0.92 (0.24)

BMI, body mass index; FEA, frequency of effect allele (i.e. BMI increasing allele); SE, standard error.
[a]Published in Speliotes *et al.*[4]
[b]Beta and SE based on inverse normal transformed BMI phenotype.
[c]D' (R^2) linkage between the SNP reported by GIANT in adults and the SNP we report here in young adults for the specified locus.

significant SNP, rs591120, in the adolescent/young adult sample or the GIANT SNP rs543874 suggest that the identified *SEC16B* SNPs are not independent (Table 2 and Supplementary Material, Fig. S4a and c). Meta-analysis of region 3 (chromosome 18, 55.5–56.6 Mb, captures *MC4R* and *PMAIP1*) conditioned on the GIANT SNP rs571312 indicated that rs17066846 is simply another marker tagging this locus ($P = 0.5$ after conditioning on rs571312) and that SNPs in that region are also not independent (Table 2 and Supplementary Material, Fig. S4b and d). In contrast, meta-analysis of region 2 (chromosome 2, 24.8–25.8 Mb, captures *ADCY3*, *POMC* and *DTNB*) conditioned either on rs1561288, the most significant SNP identified in the primary adolescent/young adult meta-analysis, or on the GIANT SNP rs713586 suggests that rs1561288 is a possible second locus in the chromosome 2 region (Table 2 and Fig. 1A and B). Because the studies used for the unconditional and conditional analysis comparison were slightly different than for the primary meta-analysis, we also show conditional results for rs2118404, which was the most significant SNP in this region before conditioning and remained the most significant SNP after conditioning in this sample. Rs2118404 is highly correlated with rs1561288 ($r^2 = 1.0$), and thus, these two SNPs represent the same signal.

We evaluated the 32 loci that had been previously published for BMI in middle-aged European adults from the GIANT consortium (4). Of the 32 published SNPs representing these loci, we observed consistent direction of effects with BMI in the adolescent/young adult meta-analysis for all SNPs except one, *SH2B1* (Table 3), and nominal statistical significance ($P < 0.05$) was achieved for 27 SNPs. A comparison of the published effect sizes for BMI in middle-aged adults (GIANT) and the observed effect sizes in our adolescent/young adult meta-analysis for these 32 SNPs showed that at $P < 0.05$, four SNPs, rs11847697 (near *PRKD1*), rs1514175 (near *TNNI3K*), rs543874 (near *SEC16B*) and rs13078807 (near *CADM2*), had larger effect estimates in the adolescent/young adult meta-analysis and rs7359397 (near *SH2B1*) had a smaller estimate compared with the GIANT analysis based on middle-aged adults. Even though we had greater than 85% power to detect the previously published effect sizes (4), we failed to find a nominally significant effect on BMI for rs7359397 (near *SH2B1*) as well as for rs10150332 (near *NRXN3*), rs10968576 (near *LRRN6NC*) and rs2287019 (near *QPCTL*). We also failed to detect a nominally significant effect for rs887912 (near *FANCL)*, but this could have been because we lacked sufficient power.

A genome-wide comparison of the differences in effect sizes between middle-aged adults (GIANT) and adolescent/young adults did not identify any SNPs with differences in effect size that were genome-wide significant ($P < 5 \times 10^{-8}$). After filtering the results based on $r^2 < 0.1$ and a distance of $\pm$ 500 kb, we identified 23 SNPs with differences between effect sizes in young adult and middle-aged adults at $P < 5 \times 10^{-5}$ (Supplementary Material, Table S7). Only one of the 23 SNPs, rs8055138 (near *SH2B1*, $r^2 = 1.0$ with the identified GIANT SNP rs7359397) had an effect size of greater magnitude in middle-aged adults compared with young adults. The other 22 SNPs had greater effect size in adolescents and young adults compared with middle-aged adults from GIANT. Among these 22 SNPs was rs2118404 (near *POMC*), which was identified from conditional analyses as described above. The other 21

Table 2. Association results for BMI in young adults and adolescents for three separate 1 Mb regions conditioned on the BMI SNP identified by the GIANT[a] consortium

Gene (Chr: region)	SNP(s) of interest (effect/other allele)	FEA	Unconditioned results Beta (SE)[b]	*P*-value	GIANT locus conditioned on[a]	*D*‘ (R^2)	Conditioned results Beta (SE)[b]	*P*-value
SEC16B (Chr1: 175.8–176.8)	rs591120 (C/G)	0.44	0.18 (0.03)	6.8E−08	rs543874	1.0 (0.31)	0.03 (0.03)	0.3
	rs12728890 (T/G)	0.06	0.28 (0.07)	8.3E−05		0.13 (0.005)	0.21 (0.07)	2.8E−03
POMC (Chr2: 24.8–25.8)	rs1561288 (C/T)	0.22	0.22 (0.04)	4.0E−08	rs713586	0.92 (0.24)	0.15 (0.04)	2.3E−04
	[c]rs2118404 (C/T)	0.23	0.26 (0.04)	2.0E−09		0.85 (0.21)	0.18 (0.04)	2.4E−05
MC4R (Chr 18: 55.5–56.6MB)	rs7234864 (T/C)	0.26	0.32 (0.04)	2.0E−17	rs571312	0.87 (0.73)	0.05 (0.04)	0.2
	rs17066846 (G/T)	0.8	0.20 (0.04)	5.9E−06		0.37 (0.11)	0.03 (0.04)	0.5
	rs11873305 (A/C)	0.96	0.36 (0.09)	4.6E−05		0.12 (0.001)	0.29 (0.09)	4.90E−04

BMI, body mass index; FEA, frequency of effect allele (i.e. BMI increasing allele); SE, standard error.
[a]SEC16B, rs543874; POMC, rs713586; MC4R, rs571312 (published in Speliotes *et al.*[4]).
[b]BMI (kg/m^2) per effect allele.
[c]Because the studies used for the unconditioned versus conditioned analyses were slightly different than for the primary meta-analysis, we also show rs2118404 that was the most significant SNP in this region before conditioning and remained the most significant SNP after conditioning in this subset. This SNP is in high LD with rs1561288 ($R^2 = 1$) and thus essentially the same signal.

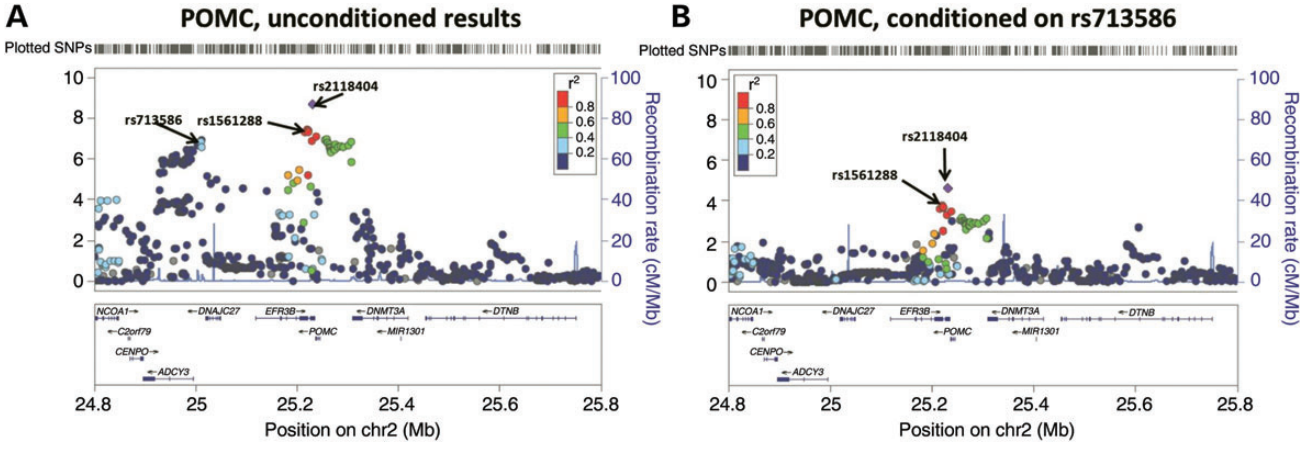

Figure 1. Potential secondary signal at POMC contributing to BMI in young adults and adolescents. (**A**) Plot with unconditioned results illustrating the most significant SNPs (rs2118404 and rs1561288) and the previously reported SNP (rs713586) in the unconditional analysis. Rs1561288 was the most significant SNP in this region in the primary analysis, but in the subset of studies used for the conditional analysis, rs2118404 (which has $r^2 = 1$ with rs1561288) was the most significant SNP. (**B**) Plot highlighting relevant SNPs after conditioning on previous reported locus, rs713586.

SNPs have not been previously reported in the literature for BMI or obesity.

We also evaluated the 13 SNPs representing 12 loci previously identified in the literature for obesity in children (Supplementary Material, Table S6) (9–11,13,17,20,21). For these 13 SNPs, we observed 10 SNPs with consistent directions of effect with BMI and five SNPs that displayed Bonferroni corrected (assuming 13 tests) significant associations ($P < 0.0038$), including the recently identified SNP near HOX5 (rs9299) (17). The SNP near *OLFM4* (rs9568856) (17) also displayed a nominally significant effect in our young adult sample.

Using data from the British 1958 Birth Cohort with BMI at 16 years and the Atherosclerosis Risk in Communities (ARIC) cohort with BMI at 25 years, we evaluated the proportion of variance explained by the seven loci identified in our study for adolescents/young adults as well as the 32 previously identified loci for BMI from GIANT. The proportion of BMI variance explained by the two cohorts combined was 1.23% for the 7 loci that reached genome-wide significance in the adolescent/young adult meta-analysis and 2.89% for the 32 BMI-associated loci found in GIANT (Supplementary Material, Fig. S5). Within ARIC, the proportion of variance explained by the seven loci that reached genome-wide significance in the adolescent/young adult meta-analysis was 1.17% in young adulthood (i.e. 25 years of age) compared with 0.72% in middle-age adulthood (i.e. 45–64 years of age). Similarly, the proportion of variance explained by the 32 established BMI-associated loci was 2.71% for BMI in young adulthood compared with 2.08% for BMI in middle-aged adulthood in ARIC. To evaluate the contribution of all common SNPs to the heritability of BMI at both life stages, we utilized the method proposed by Yang *et al.* (22) and implemented in the Genome-wide Complex Trait Analysis (GCTA) software. In ARIC, we found that the proportion of genetic variance explained for BMI based on all common SNPs was 20% in both young adulthood and middle-aged adulthood.

DISCUSSION

Although GWASs have identified over 40 common genetic loci associated with BMI in European adult (4–6,9,10,12,13) and

Table 3. Comparison of association results for the 32 loci associated with BMI in European adults in the GIANT consortium[a] with the association results for these SNPs in adolescent/young adults

SNP	Nearest gene[b]	BMI increasing allele	Other allele	FEA[c]	Results for young adults (discovery and replication)							Results for adults (stage 2, reported by GIANT)				Difference in effect size, per allele[h]		
					Beta (SE) for inverse normal BMI units per allele	Estimated effect size for BMI (kg/m^2) per allele[d]	*n*	*P*-value	I^2	P_{het}[e]	Power (%)[f]	Beta (SE) for inverse normal BIV1I units per allele	Estimated effect size for BMI (kg/m^2) per allele[g]	*n*	*P*-value	Beta (SE) for difference per allele using inverse normal BMI units	*P*-value for effect difference	Estimated effect[i] size difference per allele using BMI (kg/m^2)
rs11847697	*PRKD1*	T	C	0.04	0.101 (0.024)	0.40	28 414	8.43E−06	28.5	0.11	51.0	0.041 (0.011)	0.17	85 213	2.25×10^{-4}	0.06 (0.026)	2.18E−02	0.24
rsl514175	*TNNI3K*	A	G	0.43	0.060 (0.009)	0.24	29 040	3.40E−11	8.8	0.33	55.1	0.019 (0.004)	0.07	118 952	7.04×10^{-6}	0.041 (0.01)	2.79 E−05	0.16
rs543874	*SEC16B*	G	A	0.19	0.081 (0.011)	0.32	28 762	4.11E−11	35.1	0.05	99.9	0.051 (0.007)	0.22	55 551	2.41×10^{-11}	0.03 (0.013)	0.026	0.12
rsl3078807	*CADM2*	G	A	0.2	0.049 (0.011)	0.20	28 918	2.86E−05	21.2	0.17	67.2	0.021 (0.005)	0.10	113 596	5.32×10^{-5}	0.028 (0.012)	2.35E−02	0.11
rs571312	*MC4R*	A	C	0.23	0.078 (0.010)	0.31	29 044	8.69E−16	0	0.66	99.9	0.057 (0.006)	0.23	79 788	3.19×10^{-21}	0.022 (0.012)	0.064	0.09
rsl3107325	*SLC39A8*	T	C	0.07	0.067 (0.018)	0.27	28 555	4.34E−05	0	0.87	82.6	0.047 (0.009)	0.19	104 065	1.93×10^{-7}	0.02 (0.012)	0.10	0.08
rs206936	*NUDT3*	G	A	0.2	0.036 (0.011)	0.14	28 865	7.75E−03	0	0.67	30.3	0.016 (0.004)	0.06	125 912	7.39×10^{-4}	0.02 (0.02)	0.298	0.08
rs4836133	*ZNF608*	A	C	0.48	0.034 (0.009)	0.14	24 804	3.31E−04	25	0.14	49.5	0.015 (0.004)	0.07	119 157	1.88×10^{-4}	0.019 (0.01)	0.057	0.08
rs4771122	*MTIF3*	G	A	0.24	0.038 (0.011)	0.15	28 911	1.36E−04	0	0.84	63.7	0.020 (0.006)	0.09	118 733	8.24×10^{-4}	0.018 (0.01)	0.08	0.07
rs3817334	*MTCH2*	T	C	0.41	0.036 (0.009)	0.14	28 951	9.23E−05	1.7	0.44	42.7	0.018 (0.005)	0.06	68 128	1.10×10^{-3}	0.018 (0.012)	0.136	0.07
rs2112347	*FU35779*	T	G	0.63	0.039 (0.009)	0.16	28 953	5.00E−05	4	0.41	82.8	0.022 (0.004)	0.10	93 055	8.29×10^{-7}	0.017 (0.01)	0.092	0.07
rs2867125	*TMEM18*	C	T	0.83	0.095 (0.011)	0.38	28 954	2.74E−16	8.5	0.34	99.9	0.078 (0.006)	0.31	73 973	4.42×10^{-30}	0.017 (0.013)	0.208	0.07
rs12444979	*GPRC5B*	C	T	0.87	0.053 (0.013)	0.21	28 797	1.36E−05	2.7	0.42	92.9	0.041 (0.006)	0.17	107 030	8.13×10^{-12}	0.012 (0.014)	0.39	0.05
rs3810291	*TMEM160*	A	G	0.67	0.033 (0.010)	0.13	28 811	2.59E−04	18.8	0.20	71.9	0.021 (0.004)	0.09	122 030	1.59×10^{-6}	0.011 (0.011)	0.33	0.05
rs713586	*RBJ POMC*	C	T	0.47	0.046 (0.009)	0.18	28 951	9.35E−08	32.8	0.06	98.8	0.035 (0.004)	0.14	115 850	1.44×10^{-16}	0.01 (0.01)	0.29	0.04
rs2815752	*NEGR1*	A	G	0.61	0.041 (0.009)	0.16	29 050	1.36E−05	13.1	0.28	96.9	0.032 (0.005)	0.13	74 531	2.29×10^{-9}	0.01 (0.01)	0.34	0.04
rs29941	*KCTD15*	G	A	0.67	0.021 (0.009)	0.08	29 049	6.28E−03	0	0.73	39.8	0.012 (0.005)	0.05	69 030	2.40×10^{-2}	0.009 (0.011)	0.43	0.04
rs1555543	*PTBP2*	C	A	0.58	0.024 (0.009)	0.10	28 949	2.94E−03	0.5	0.45	42.9	0.017 (0.004)	0.06	74 716	4.48×10^{-5}	0.007 (0.01)	0.45	0.03
rs987237	*TFAP2B*	G	A	0.18	0.038 (0.011)	0.15	28 964	1.70E−04	7.5	0.36	85.2	0.032 (0.006)	0.13	71 916	2.40×10^{-6}	0.006 (0.013)	0.62	0.03
rs2241423	*MAP2K5*	G	A	0.77	0.038 (0.010)	0.15	29 011	1.68E−04	0	0.84	90.9	0.032 (0.005)	0.13	104 115	1.59×10^{-9}	0.006 (0.012)	0.61	0.02
rs9816226	*ETV5*	T	A	0.82	0.038 (0.012)	0.15	28 866	1.75E−03	0	0.99	89.8	0.033 (0.006)	0.14	72 362	1.15×10^{-6}	0.005 (0.013)	0.69	0.02
rsl0938397	*GNPDA2*	G	A	0.43	0.046 (0.009)	0.18	28 951	B.42E−08	0	0.59	99.9	0.043 (0.005)	0.18	73 160	1.45×10^{-15}	0.003 (0.01)	0.76	0.01
rs4929949	*RPL27A*	C	T	0.52	0.014 (0.009)	0.06	28 951	0.02	0	0.63	43.8	0.013 (0.004)	0.06	125 931	1.00×10^{-3}	0.002 (0.01)	0.87	0.01
rs7138803	*FAIM2*	A	G	0.38	0.029 (0.009)	0.12	29 046	1.66E−04	0	0.66	93.9	0.028 (0.005)	0.12	76 265	7.82×10^{-8}	0.001 (0.01)	0.90	0.01
rs887912	*FANCL*	T	C	0.29	0.023 (0.009)	0.09	28 911	0.14	32.1	0.06	77.9	0.024 (0.004)	0.10	113 981	1.72×10^{-7}	0 (0.011)	0.98	0.00
rs2890652	*LRP1B*	C	T	0.17	0.023 (0.012)	0.09	28 866	0.03	20.6	0.18	52.8	0.024 (0.006)	0.09	121 816	9.47×10^{-5}	0 (0.013)	0.98	0.00
rs2287019	*QPCTL*	C	T	0.8	0.034 (0.014)	0.14	27 781	0.20	32.2	0.07	94.4	0.037 (0.005)	0.15	94 091	1.40×10^{-10}	−0.003 (0.015)	0.86	−0.01
rs1558902	*FTO*	A	T	0.42	0.084 (0.009)	0.34	28 249	1.69E−22	2.4	0.43	99.9	0.091 (0.005)	0.39	68 498	1.01×10^{-60}	−0.007 (0.01)	0.50	−0.03
rsl0150332	*NRXN3*	C	T	0.21	0.020 (0.011)	0.08	28 864	0.06	8	0.35	88.9	0.029 (0.007)	0.13	59 157	2.86×10^{-5}	−0.009 (0.013)	0.50	−0.04
rsl0968576	*LRRN6C*	G	A	0.31	0.010 (0.009)	0.04	29 052	0.20	14.4	0.25	86.6	0.023 (0.004)	0.11	107 866	3.19×10^{-6}	−0.013 (0.011)	0.23	−0.05
rs10767664	*BDNF*	A	T	0.78	0.027 (0.011)	0.11	28 865	0.01	0	0.48	99.7	0.047 (0.006)	0.19	80 293	1.17×10^{-14}	−0.021 (0.012)	0.089	−0.08
rs7359397	*SH2B1*	T	C	0.61	−0.002 (0.009)	−0.01	28 947	0.74	9	0.33	99.3	0.034 (0.005)	0.15	80 445	7.89×10^{-12}	−0.037 (0.01)	3.15E−04	−0.15

Loci are shown from largest to smallest difference in effect size. BMI, body mass index; FEA, frequency of effect allele (i.e. BMI increasing allele); SE, standard error.

[a]Published in Speliotes *et al.*[4]

[b]Nearest gene within 500 kb of the indicated SNP.

[c]Frequency of the effect allele, similar in GIANT adult cohorts.

[d]Calculated from effect size of stage 1 and 2 results and SD of BMI in all cohorts (SD = 4.0 kg/m^2).

[e]Heterogeneity quantified the effect of inconsistency across the studies (all results files) and was calculated using Cochrane *Q* and comparing in two ways: (i) with the χ^2 distribution with $k - 1$ degrees of freedom, where $k = 27$ input results files (P_{het} = *P*-value using χ^2); and (ii) with the I^2, the percentage of total variation across studies that is due to heterogeneity, that is calculated as $I^2 = 100\% \times (Q - \text{df})/Q$, where *Q* is the Cochrane *Q* and df the degrees of freedom.

[f]Power estimated using QUANTO. Power estimates per locus included allele frequencies (FEA), sample size and mean BMI (mean ± SD) = 23.0 kg/m^2 ± 4.0 from young adult data in current study, and effect estimates from those reported for adults in the GIANT study. The type 1 error rate set at $\alpha = 0.05$, to detect nominally significant effect estimates.

[g]Reported effects from stage 2 cohorts only; calculated using effect size × SD of mean BMI.

[h]Four larger effect sizes [+Beta (SE) for difference and $P < 0.05$] in young adults and one smaller effect size in young adults [−Beta (SE) for difference and $P < 0.05$] compared with GIANT stage 2 results in adults. Beta difference calculated using *z*-statistic including Goncalo's correction for correlated data with Pearson $r = 0.08$: (Beta_A − Beta_B)/SQRT[(SE_A)^2 + (SE_B)^2] − 2 × 0.08 × SE_A × SE_B, where Beta_A and SE_A are from the young adult sample and Beta_B and SE_B are from the GIANT stage 2 sample.

[i]Calculated using effect size × SD of mean BMI.

pediatric (6,9,10,13,15–20) studies, the extent to which these findings can be generalized to other life-cycle periods is an open question. In our analysis of 16–25-year olds of European descent, a life-cycle period of elevated risk for weight gain, we found seven genome-wide significant loci (*FTO*, *TMEM18*, *MC4R*, *TNNI3K*, *SEC16B*, *GNDPA2* and *POMC*) associated with BMI. All seven loci were previously identified in European adult populations (4). Furthermore, all have been identified as being associated with obesity in children (6,15,17). Interestingly, although SNP rs1561288 showed partial dependence on the previously identified index SNP in the region, we found suggestive evidence for a secondary signal in the POMC region from the conditional analyses. Except for *SH2B1*, the other BMI loci identified in European adults displayed directionally consistent effect estimates during the adolescent/young adult period. Further, 27 of 32 loci were nominally significant ($P < 0.05$). For the 13 loci identified for obesity in children, 10 showed consistent direction of effect and 6 were nominally significant ($P < 0.05$). Taken together, our results support the generalization of at least most of the genetic obesity loci across the lifespan.

Differences in the effect sizes for the 32 known BMI loci in adults (average age 55 years) relative to the adolescents and young adults were statistically significant (at $P < 0.05$) for five loci. For four loci (*PRKD1*, *TNNI3K*, *SEC16B* and *CADM2*), the effect sizes were larger in adolescence/young adulthood compared with middle-age adulthood, but for the locus near *SH2B1* (rs7359397), the effect size was larger in later adulthood. Even after correction for multiple testing, rs1514715 (near *TNNI3K*) and rs7359397 (near *SH2B1*) remained significantly different. Although the differences in size for the other 27 loci were not nominally significant, absolute differences in effect estimates showed that 20 of 27 SNPs had larger effect sizes in adolescent/young adults compared with middle-aged adults (Table 3). Thus, we found a greater number of effect estimates that were larger in the adolescent/young adults compared with the older adults than expected by chance alone ($P = 0.007$). Further, we found no association between BMI and rs7359397 (near *SH2B1*) in the adolescent/young adult meta-analysis, even though we had greater than 80% power to detect an effect. Possibly, this locus contributes more to BMI during a later time point in the lifespan. Although the genome-wide comparison of the effect sizes for BMI in middle-aged adults (GIANT) and the adolescent/young adult sample did not identify any loci with genome-wide significant differences, it is notable that 22 of the 23 SNPs that displayed *P*-values of $<5 \times 10^{-5}$ had larger magnitude of effects in the adolescents and young adults compared with the middle-aged adults. Twenty-one of the 22 SNPs have not been previously reported in the literature for BMI or obesity. Notably, rs9391253 is highly correlated ($r^2 > 0.8$) with SNPs near *LIN28B* (rs7759938, rs314268 and rs314276), which have been associated with height and menarche. Another SNP, rs9923856, is moderately ($r^2 = 0.4$) correlated with SNPs near *CLEC16A/KIAA0350* that have been associated with type 1 diabetes.

The proportion of variance explained by the 7 loci showing genome-wide significance in young adults and the 32 established BMI loci was slightly (but not significantly) larger for ARIC individuals in young adulthood than middle-aged adulthood. The analysis of all common SNPs using GCTA suggested that the total proportion of variance explained by common SNPs for BMI was similar between young adulthood and middle-age; however, differences in individual SNPs or rare variants may still exist.

Studies in European pediatric populations have shown associations of childhood obesity with *TNNI3K* and *SEC16B* (*17*). *TNNI3K* is a novel cardiac troponin I-interacting kinase gene (23), but its role in body weight and obesity is unclear. The lack of association observed in the adolescent/young adult period for *SH2B1* is supported by the published literature. Studies in European pediatric populations have also failed to find an association between *SH2B1* and BMI or obesity (6,14,15). It is possible that *SH2B1*, a neuronal gene implicated in glucose homeostasis and leptin signaling (24,25), fails to play a measurable effect on BMI until other endogenous or exogenous factors that only manifest or accumulate later in life occur. Alternatively, perhaps it takes a long time for the cumulative effects of the genetic variant to substantially alter BMI. Consistent with a potential cumulative effect, studies in mice show that a deficiency of *SH2B1* results in metabolic disorders such as obesity and diabetes; and when *SH2B1* expression is restored, it protects against obesity and diabetic phenotypes in a dose-dependent manner (26).

Distinct genetic effects during adolescence/young adulthood are also more broadly supported by the literature. A study in European adolescents and children found that effect sizes for BMI were more pronounced in children than in adults for *TMEM18*, *SEC16B* and *KCTD15*, whereas a locus near *BNDF* was comparatively smaller in children (27). In addition, a study of variation in the association between *FTO* and *MC4R* gene variants with body size over the life course from birth to age 53 years showed that the association reached peak strength at age 20 and then weakened during later adulthood (28). We did not find any nominally significant differences in young adult and adult effect estimates for these two loci. However, the effect estimates for the *MC4R* locus (rs571312) was slightly larger in young adults compared with adults [Supplementary Material, Table S4, beta (SE) difference = 0.022 (0.012), $P = 0.07$]. *MC4R* are expressed in hypothalamic tissue, the metabolic control center, which is outstandingly active during adolescence and young adulthood (29–31). In a previous study, odds ratios for eight loci (*FTO*, *TMEM18*, *MC4R*, *TNNI3K*, *SEC16B*, *GNDPA2*, *QPCTL* and *BNDF*) nominally associated with obesity in children (15) also appeared to be larger than those reported for adult obesity in the GIANT study (4). Among the possible biological mechanisms, larger effect sizes of loci on BMI earlier in life may be due to alterations in factors that, for example, regulate metabolism and glycemic homeostasis. These factors may be more sensitive to genetic alterations early in life, whereas at older ages, other environmental elements may play a larger role. We can speculate that we observed larger effect estimates during adolescence and young adulthood because this group captures a relatively narrow age range marked by post-pubertal growth and body composition changes. In contrast, the adult sample spanned a large age range (aged 18–90), with a comparatively stable period of weight change. It is possible that the current generation of young adults having matured in a comparatively more obesogenic environment might have experienced differential contribution of genetic influences on weight. It is also possible that over time, and with age, the obesogenic

environment plays a comparatively stronger role in body weight than genetic factors.

We acknowledge that there are limitations in our study, including the range of years in data collection between cohorts across the meta-analysis sample. In addition, although we did not observe substantial heterogeneity in effects across study samples, it is possible that some gender differences may have been overlooked by combining women and men. The year of data collection was controlled for in family cohorts where data collection spanned across generations. However, secular differences between cohorts still may influence our results. The use of combined self-report and measured height could also be an issue, as self-reported height and weight may be subject to under-reporting bias (32,33), particularly in individuals with higher BMI. Although our sample is the heretofore largest to span the adolescent to young adult years, an even larger sample would have permitted the detection of smaller BMI–SNP associations.

Despite the limitations, the current study strengthens the current understanding of genetic influences on BMI during a narrow period of the life cycle. Even with a limited sample size, we detected associations with previously reported genetic loci for BMI. Further, we provide evidence for an underlying genetic predisposition to obesity that may have greater influence on body weight during the period of adolescence compared with adulthood. The current understanding of the epidemiological architecture of these genetic effects across the life cycle, as well as gene–gene and gene–environment effects in developmental trajectories across lifespan remains incomplete with many aspects still to be studied; however, our study does provide additional insight into the role of genetic factors during the critical time period from late adolescence and early adulthood.

MATERIALS AND METHODS

Study design

We designed a two-stage study comprising a genome-wide association meta-analysis (discovery sample) of data on up to 13 627 genotyped individuals from 10 studies and selected the most promising SNPs after filtering ($n = 76$) for follow-up analysis in up to 16 253 additional genotyped individuals from four studies. Subjects were excluded if they were not of European ethnicity, lacked information for BMI between 16 and 25 years or had extreme values for BMI that were outside of ± 4 SD from the cohort-specific mean. We removed potential outliers for BMI (e.g. >4 SD) in order to ensure that BMI was normally distributed for the analysis. All cohorts for discovery sample included observations for BMI measured or reported between ages 16 and 21 ($n = 13\,627$) and follow-up cohorts included observations for BMI between the ages of 16 and 25 ($n = 16\,253$).

Study population and genotyping

Discovery sample of GWASs

The discovery meta-analysis samples came from10 GWASs, including the British 1958 Birth Cohort (B58C), Framingham Heart Study (FHS), GENEVA Dental Caries Study, Sardinian Study on Aging (SardiNIA), Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO), Nurses' Health Study (NHS), Health Professionals Follow-up Study (HPFS), Study of Health in Pomerania (SHIP), Estonian Genome Center, University of Tartu (EGCUT), and Erasmus Rupchen Family (ERF) (Supplementary Material, Tables S1–S3). The sample size for these studies ranged from 65 to 4171 individuals, with a combined sample size of 13 627. All studies were genotyped using Affymetrix or Illumina whole genome genotyping arrays. To allow for meta-analysis across different marker sets, imputation of the polymorphic SNPs from the HapMap2 European CEU population was performed using MACH (34) or IMPUTE (35) or BimBam (36).

Follow-up sample of GWASs

The follow-up sample consisted of four additional GWAS: ARIC, Fels Longitudinal Study (Fels) and additional participants from the B58C and PLCO studies (i.e. PLCO2; Supplementary Material, Tables S1–S3). Data were available for 16 253 individuals that were genotyped using Affymetrix or Illumina whole genome genotyping arrays. For ARIC, B58C and PLCO, the polymorphic SNPs from the HapMap2 European CEU population were imputed using MACH (34).

Statistical analyses

Association analyses with BMI

Each study performed single-marker association analyses with BMI calculated from measured or self-reported height and weight at adolescence or early adulthood, as the dependent variable, using an additive genetic model implemented in MACH2QTL (Y. Li *et al*., unpublished data), PLINK (37), SNPTEST (35), ProbAbel (38), SOLAR (39), Merlin (40) or linear-mixed effects models in R (lmekin) (41). Prior to analysis, BMI was adjusted for age, study center and principal components to correct for population substructure using linear regression. The adjusted BMI residuals were then transformed using an inverse normal transformation with a mean of zero and standard deviation of 1, which in turn was used as the phenotype outcome for association analyses with each SNP. Analyses were stratified by sex and by case status, when applicable. FHS, SardiNIA, ERF and Fels used a linear mixed-effects model clustering individuals in families so as to account for relatedness between these individuals, using Merlin (38), lmekin in R (39) or proc mixed in SAS with a random effect for family. For each study, we excluded SNPs with poor imputation quality scores (r^2.hat < 0.3 in MACH or BimBam or proper info <0.4 in SNPTEST) and a minor allele count (MAC $= 2N \times$ minor allele frequency) <20. Genomic control was applied to each study which had lambda values ranging from 0.984 to 1.034. The results of the discovery meta-analysis were followed by an additional overall genomic control correction in METAL. Before the correction, the genomic control lambda for the discovery results was 1.061. After correction, the genomic control lambda was 0.999.

Meta-analysis

The meta-analyses for discovery and follow-up stages and the combined discovery and follow-up stage used a fixed effects approach by weighting the effect size estimates with the inverse of the standard errors. We validated effect estimates from the standard error weighted method with the sample size weighted

z-score method, which is based on the direction of association and *P*-values of each individual study. Both meta-analyses were performed using METAL (http://www.sph.umich.edu/csg/abecasis/metal) and the correlation between the resulting $-\log_{10}$ *P*-values was high ($r > 0.99$). All SNP association analyses were tested for between sample heterogeneity.

To assess the validity of discovery findings, the most promising SNPs from the discovery meta-analysis were taken forward into the four follow-up studies. After filtering based on $r^2 < 0.1$ and a distance of ± 500 kb using PLINK v1.07 (http://pngu.mgh.harvard.edu/purcell/plink/), we identified 76 SNPs with $P < 5 \times 10^{-5}$ for follow-up. The results from the follow-up studies were meta-analyzed as described above. We considered SNPs that reached genome-wide significance at $P < 5 \times 10^{-8}$ in the combined meta-analysis of the discovery and follow-up studies to be noteworthy. Although we did not require the SNPs to reach a particular significance level in the follow-up analyses, it should be noted that all genome-wide significant results from discovery and follow-up stages combined were at least nominally significant in the meta-analysis of the follow-up studies only.

In a sensitivity analyses in the combined discovery and follow-up samples, we ran association analyses of BMI with 76 loci of interest separately on those with self-reported BMI or with measured BMI (Supplementary Material, Table S4b). Accounting for sample size differences, the effect sizes were similar.

Conditional meta-analysis

We performed conditional analysis for three 1 Mb regions of the genome where we identified significant associations for SNPs (rs591120, near *SEC16B*; rs1561288, near *POMC*; rs17066846, near *MC4R*) that had low linkage disequilibrium ($r^2 \leq 0.3$) with the previously published SNPs in the GIANT consortium (4). We tested the independence of the indicated SNPs within the following 1 Mb regions: (i) chromosome 1, 175.8–176.8 Mb, which captures our SNP, rs591120, and the GIANT SNP, rs543874 (*SEC16B*); (ii) chromosome 2, 24.8–25.8 Mb, which captures our SNP, rs1561288, and the GIANT SNP, rs713586 (*ADCY3*, *POMC*, *DTNB*) and (iii) chromosome 18, 55.5–56.6 Mb, which captures our most significant SNP, rs7234864, and a possible secondary locus, rs17066846, and the GIANT SNP, rs571312 (*MC4R*, *PMAIP1*). These conditional analyses were conducted using the following equation, where SNP_1 and SNP_2 are the two SNPs of interest in the region:

$$\text{transformed BMI residual} = \alpha + \beta_1 SNP_1 + \beta_2 SNP_2 + \text{error}.$$

Comparison of effect sizes for known BMI loci

The 32 loci associated with BMI in Europeans identified from the GIANT consortium (4) were meta-analyzed in the combined discovery and follow-up studies as described above. We compared our effect sizes from inverse normally transformed BMI with effect sizes in middle-aged adults of European ancestry from GIANT replication studies (4), also inverse normally transformed. Nine of the 14 studies included in our analysis had overlapping samples with GIANT. Although the BMI measurements utilized were different (i.e. adolescence/early adulthood versus middle-aged adulthood), we accounted for the correlation due to overlapping samples in our comparison. We used *z*-tests to compare effect estimates (β) from our study (A) and the GIANT study (B) adjusting for the correlation due to overlapping samples such that:

$$\frac{\beta_A - \beta_B}{\sqrt{(SE_A)^2 + (SE_B)^2 - 2(r)(SE_A)(SE_B)}},$$

where SE is the standard error and *r* the Pearson correlation coefficient between the effect estimates. We calculated the Pearson correlation coefficient between our study and GIANT using the discovery stage data from both studies. The significance level (*P*-value) was based on a two-tailed *z*-test.

In a sensitivity analyses, we compared effect sizes in each young adults that did and did not contribute middle-aged BMI observations to GIANT and with effect sizes in middle-aged adults of European ancestry from GIANT replication studies (Supplementary Material, Tables S5a and S5b). Accounting for sample size differences, the differences in effect sizes between young adults and middle-aged adults are similar regardless of their contribution to the GIANT analysis.

Genome-wide comparison of effect sizes

As a comprehensive comparison of age-related differences in effects, we compared effect sizes adolescent/young adult samples with effect sizes in middle-aged adults of European ancestry from GIANT stage 1 cohorts (4) on a genome-wide level. Estimates from inverse normally transformed BMI were used in the *z*-test comparisons, as described above, to compare the effect estimates across the genome. To assess the number of loci from the *z*-test comparisons, all results were filtered based on a linkage disequilibrium of $r^2 < 0.1$, a distance of ± 500 kb and $P < 5 \times 10^{-5}$ using PLINK v1.07 (http://pngu.mgh.harvard.edu/purcell/plink/). SNPs were excluded if the sample size for each the adolescent/young adult sample or the GIANT sample fell below the mean.

Proportion of variance explained and GCTAs

Using two of the largest cohorts in our sample, British 1958 Birth Cohort and the ARIC cohort, we estimated the proportion of variance explained by the 7 loci that were genome-wide significant in the combined meta-analyzed young adults and also for the 32 GIANT BMI-associated loci at late adolescence/early adulthood. These two cohorts also spanned the adolescent/young adult age range; the British 1958 Birth Cohort participants had BMI at 16 years and the ARIC cohort sample provided self-reported heights and weights at 25 years of age. As a comparison, we again estimated the proportion of variance explained by the 7 loci and the 32 GIANT BMI-associated loci for middle-aged adults using BMI measurements at 45–64 years of age (mean age 54.3 years) from the ARIC cohort. Untransformed BMI residuals, adjusted for sex and age, were linearly regressed on each set of SNPs (i.e. the genome-wide significant 7 loci or 32 GIANT BMI-associated loci).

To explore the contribution of common SNPs to BMI, we estimated the variance explained by all the autosomal SNPs for BMI in early adulthood (25 years of age) and also middle-age (mean

age 54.3 years) in ARIC, using the method proposed by Yang *et al*. (40) and implemented in the GCTA software package (http://www.complextraitgenomics.com/software/gcta/).

Power

We estimated the power to detect a genome-wide significant association for our two-stage study assuming an additive genetic model, a combined sample size of 29 880 young adults and an alpha level of 5×10^{-8}. Based on effect sizes published in European populations, we had at least 80% power for to detect an effect genome-wide for SNPs where the allele frequency was at least 0.2 and the effect size was >0.25 kg/m^2 (Supplementary Material, Fig. S6). We also calculated power to detect associations in the 32 previously identified GIANT SNPs assuming an additive genetic model, a sample size of 29 880 young adults and a α level of 0.05. As shown in Table 3, we have $>80\%$ power to detect over half (19 of 32) of the loci. Power estimates were calculated using QUANTO v1.2.4 (http://hydra.usc.edu/gxe/).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

Complete lists of acknowledgements are provided in the Supplementary Material.

Conflict of Interest statement. None declared.

FUNDING

This work was funded by National Institutes of Health (R01HD057194).

Atherosclerosis Risk in Communities (ARIC). The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute (N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC- 55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367 and R01HL086694); National Human Genome Research Institute (U01HG004402); and National Institutes of Health (HHSN268200625226C). Infrastructure was partly supported by a component of the National Institutes of Health and NIH Roadmap for Medical Research (UL1RR025005). The project described was supported by the National Center for Research Resources (NCRR) (UL1 RR 025005).

British 1958 Birth Cohort (B58C). We acknowledge the use of phenotype and genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council (G0000934) and the Wellcome Trust (068545/Z/02) (http://www.b58cgene.sgul.ac.uk/). Genotyping for the B58C-WTCCC subset was funded by the Wellcome Trust (076113/B/04/Z). The B58C-T1DGC genotyping utilized resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418.

Erasmus Rucphen Family (EUROSPAN) (ERF). The genotyping for the ERF study was supported by EUROSPAN (European Special Populations Research Network) and the European Commission FP6 STRP (018947; LSHG-CT-2006-01947). The ERF study was further supported by grants from the Netherlands Organization for Scientific Research, Erasmus MC, the Centre for Medical Systems Biology (CMSB) and the Netherlands Brain Foundation (HersenStichting Nederland).

Estonian Genome Center, University of Tartu (EGCUT). EGCUT received financing by (ENGAGE, OPENGENE), targeted financing from Estonian Government (funding no: SF0180142s08), Estonian Research Roadmap through Estonian Ministry of Education and Research (3.2.0304.11-0312), Center of Excellence in Genomics (EXCEGEN) and Development Fund of University of Tartu (funding no: SP1GVARENG).

Framingham Heart Study (FramHS). This research was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (contract no: N01-HC-25195) and its contract with Affymetrix, Inc., for genotyping services (contract no. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. Analyses of the Framingham data were supported by subcontract to Boston University under NIH (R01HD057194).

Dental Caries: Whole Genome Association and Gene x Environment Studies (GENEVA Dental Caries). Funding support for the GWAS project was provided by the National Institutes of Dental and Craniofacial Research (U01-DE018903). Data and samples were provided by the Center for Oral Health Research in Appalachia, a collaboration of the University of Pittsburgh and West Virginia University funded by NIDCR (R01-DE 014899); genotyping was done by the Johns Hopkins University Center for Inherited Disease Research (CIDR) which is fully funded through a federal contract from the National Institutes of Health (NIH) to The Johns Hopkins University (contract no: HHSN268200782096C). Funds for this project's genotyping were provided by the NIDCR through CIDR's NIH contract. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (www.genevastudy.org; U01-HG004446).

Fels Longitudinal Study (Fels). This work was supported by the National Institutes of Health (R01-HD012252, R01-HD053685, R01-DK064391, R01-AR052147).

The Nurses' Health Study/Health Professionals Follow up Study (NHS/HPFS). This study was supported by National Institutes of Health grants [HL71981 and DK46200 (Boston Obesity Nutrition Research Center)] and by an American Heart Association Scientist Development Award (0730094N).

Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO/PLCO2). The Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial was funded by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH.

Study of Health in Pomerania (SHIP). SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (01ZZ9603, 01ZZ0103 and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED)' funded by the Federal Ministry of Education and Research (03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg-West Pomerania.

Study of Health in Sardinia, Italy (SardNIA). This work was supported in part by the Intramural Research Program of the National Institute on Aging (NIA) and National Institutes of Health (NIH) (NO1-AG-1-2109).

REFERENCES

1. Dietz, W. (1994) Critical periods in childhood for the development of obesity. *Am. J. Clin. Nutr.*, **59**, 955–959.
2. Gordon-Larsen,, P., Adair, L.S., Nelson, M.C. and BM, P. (2004) Five-year obesity incidence in the transition period between adolescence and adulthood: the National Longitudinal Study of Adolescent Health. *Am. J. Clin. Nutr.*, **80**, 569–575.
3. The, N.S., Suchindran, C., North, K.E., Popkin, B.M. and Gordon-Larsen, P. (2010) Association of adolescent obesity with risk of severe obesity in adulthood. *JAMA*, **304**, 2042–2047.
4. Speliotes, E.K., Willer, C.J., Berndt, S.I., Monda, K.L., Thorleifsson, G., Jackson, A.U., Allen, H.L., Lindgren, C.M., Luan, J., Magi, R. *et al.* (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.*, **42**, 937–948.
5. Thorleifsson, G., Walters, G., Gudbjartsson, D., Steinthorsdottir, V., Sulem, P., Helgadottir, A. and Styrkarsdottir, U. (2009) Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat. Genet.*, **41**, 18–24.
6. Willer, C.J., Speliotes, E.K., Loos, R.J., Li, S., Lindgren, C.M., Heid, I.M., Berndt, S.I., Elliott, A.L., Jackson, A.U., Lamina, C. *et al.* (2009) Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat. Genet.*, **41**, 25–34.
7. Okada, Y., Kubo, M., Ohmiya, H., Takahashi, A., Kumasaka, N., Hosono, N., Maeda, S., Wen, W., Dorajoo, R., Go, M.J. *et al.* (2012) Common variants at CDKAL1 and KLF9 are associated with body mass index in east Asian populations. *Nat. Genet.*, **44**, 302–306.
8. Wen, W., Cho, Y.S., Zheng, W., Dorajoo, R., Kato, N., Qi, L., Chen, C.H., Delahanty, R.J., Okada, Y., Tabara, Y. *et al.* (2012) Meta-analysis identifies common variants associated with body mass index in east Asians. *Nat. Genet.*, **44**, 307–311.
9. Scherag, A., Dina, C., Hinney, A., Vatin, V., Scherag, S., Vogel, C.I., Muller, T.D., Grallert, H., Wichmann, H.E., Balkau, B. *et al.* (2010) Two new Loci for body-weight regulation identified in a joint analysis of genome-wide association studies for early-onset extreme obesity in French and German study groups. *PLoS Genet.*, **6**, e1000916.
10. Meyre, D., Delplanque, J., Chevre, J.C., Lecoeur, C., Lobbens, S., Gallina, S., Durand, E., Vatin, V., Degraeve, F., Proenca, C. *et al.* (2009) Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat. Genet.*, **41**, 157–159.
11. Dina, C., Meyre, D., Gallina, S., Durand, E., Korner, A., Jacobson, P., Carlsson, L.M., Kiess, W., Vatin, V., Lecoeur, C. *et al.* (2007) Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat. Genet.*, **39**, 724–726.
12. Frayling, T.M., Timpson, N.J., Weedon, M.N., Zeggini, E., Freathy, R.M., Lindgren, C.M., Perry, J.R., Elliott, K.S., Lango, H., Rayner, N.W. *et al.* (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*, **316**, 889–894.
13. Loos, R.J., Lindgren, C.M., Li, S., Wheeler, E., Zhao, J.H., Prokopenko, I., Inouye, M., Freathy, R.M., Attwood, A.P., Beckmann, J.S. *et al.* (2008) Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat. Genet.*, **40**, 768–775.
14. Zhao, J., Bradfield, J.P., Li, M., Wang, K., Zhang, H., Kim, C.E., Annaiah, K., Glessner, J.T., Thomas, K., Garris, M. *et al.* (2009) The role of obesity-associated loci identified in genome-wide association studies in the determination of pediatric BMI. *Obesity (Silver Spring)*, **17**, 2254–2257.
15. Zhao, J., Bradfield, J.P., Zhang, H., Sleiman, P.M., Kim, C.E., Glessner, J.T., Deliard, S., Thomas, K.A., Frackelton, E.C., Li, M. *et al.* (2011) Role of BMI-associated loci identified in GWAS meta-analyses in the context of common childhood obesity in European Americans. *Obesity (Silver Spring)*, **19**, 2436–2439.
16. Zhao, J. and Grant, S.F. (2011) Genetics of childhood obesity. *J. Obes.*, **2011**, 845–148.
17. Bradfield, J.P., Taal, H.R., Timpson, N.J., Scherag, A., Lecoeur, C., Warrington, N.M., Hypponen, E., Holst, C., Valcarcel, B., Thiering, E. *et al.* (2012) A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat. Genet.*, **44**, 526–531.
18. Hinney, A., Nguyen, T.T., Scherag, A., Friedel, S., Bronner, G., Muller, T.D., Grallert, H., Illig, T., Wichmann, H.E., Rief, W. *et al.* (2007) Genome wide association (GWA) study for early onset extreme obesity supports the role of fat mass and obesity associated gene (FTO) variants. *PLoS One*, **2**, e1361.
19. Lasky-Su, J., Lyon, H.N., Emilsson, V., Heid, I.M., Molony, C., Raby, B.A., Lazarus, R., Klanderman, B., Soto-Quiros, M.E., Avila, L. *et al.* (2008) On the replication of genetic associations: timing can be everything! *Am. J. Hum. Genet.*, **82**, 849–858.
20. Paternoster, L., Evans, D.M., Nohr, E.A., Holst, C., Gaborieau, V., Brennan, P., Gjesing, A.P., Grarup, N., Witte, D.R., Jorgensen, T. *et al.* (2011) Genome-wide population-based association study of extremely overweight young adults–the GOYA study. *PLoS One*, **6**, e24303.
21. Benzinou, M., Creemers, J.W., Choquet, H., Lobbens, S., Dina, C., Durand, E., Guerardel, A., Boutin, P., Jouret, B., Heude, B. *et al.* (2008) Common nonsynonymous variants in PCSK1 confer risk of obesity. *Nat. Genet.*, **40**, 943–945.
22. Yang, J., Lee, S.H., Goddard, M.E. and Visscher, P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.*, **88**, 76–82.
23. Zhao, Y., Meng, X.M., Wei, Y.J., Zhao, X.W., Liu, D.Q., Cao, H.Q., Liew, C.C. and Ding, J.F. (2003) Cloning and characterization of a novel cardiac-specific kinase that interacts specifically with cardiac troponin I. *J. Mol. Med. (Berl.)*, **81**, 297–304.
24. Li, Z., Zhou, Y., Carter-Su, C., Myers, M.G. Jr and Rui, L. (2007) SH2B1 enhances leptin signaling by both Janus kinase 2 Tyr813 phosphorylation-dependent and -independent mechanisms. *Mol. Endocrinol.*, **21**, 2270–2281.
25. Morris, D.L., Cho, K.W., Zhou, Y. and Rui, L. (2009) SH2B1 enhances insulin sensitivity by both stimulating the insulin receptor and inhibiting tyrosine dephosphorylation of insulin receptor substrate proteins. *Diabetes*, **58**, 2039–2047.
26. Ren, D., Zhou, Y., Morris, D., Li, M., Li, Z. and Rui, L. (2007) Neuronal SH2B1 is essential for controlling energy and glucose homeostasis. *J. Clin. Invest.*, **117**, 397–406.
27. den Hoed, M., Ekelund, U., Brage, S., Grontved, A., Zhao, J.H., Sharp, S.J., Ong, K.K., Wareham, N.J. and Loos, R.J. (2010) Genetic susceptibility to obesity and related traits in childhood and adolescence: influence of loci identified by genome-wide association studies. *Diabetes*, **59**, 2980–2988.
28. Hardy, R., Wills, A.K., Wong, A., Elks, C.E., Wareham, N.J., Loos, R.J., Kuh, D. and Ong, K.K. (2010) Life course variations in the associations between FTO and MC4R gene variants and body size. *Hum. Mol. Genet.*, **19**, 545–552.
29. Gunnar, M.R., Wewerka, S., Frenn, K., Long, J.D. and Griggs, C. (2009) Developmental changes in hypothalamus-pituitary-adrenal activity over the transition to adolescence: normative changes and associations with puberty. *Dev. Psychopath.*, **21**, 69–85.
30. Spear, L.P. (2000) The adolescent brain and age-related behavioral manifestations. *Neurosci. Biobehav. Rev.*, **24**, 417–463.
31. Stroud, L.R., Papandonatos, G.D., Williamson, D.E. and Dahl, R.E. (2004) Sex differences in the effects of pubertal development on responses to a corticotropin-releasing hormone challenge: the Pittsburgh psychobiologic studies. *Ann. N. Y. Acad. Sci.*, **1021**, 348–351.

32. Gorber, S.C., Tremblay, M., Moher, D. and Gorber, B. (2007) A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes. Rev.*, **8**, 307–326.
33. Jain, R.B. (2010) Regression models to predict corrected weight, height and obesity prevalence from self-reported data: data from BRFSS 1999–2007. *Int. J. Obes. (Lond.)*, **34**, 1655–1664.
34. Li, Y., Willer, C., Sanna, S. and Abecasis, G. (2009) Genotype imputation. *Annu. Rev. Genomics Hum. Genet.*, **10**, 387–406.
35. Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906–913.
36. Guan, Y. and Stephens, M. (2008) Practical issues in imputation-based association mapping. *PLoS Genet.*, **4**, e1000279.
37. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
38. Aulchenko, Y.S., Struchalin, M.V. and van Duijn, C.M. (2010) ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics*, **11**, 134.
39. Almasy, L. and Blangero, J. (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.*, **62**, 1198–1211.
40. Yang, J., Lee, S.H., Goddard, M.E. and Visscher, P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.*, **88**, 76–82.
41. Bailey, D. and Alimadhi, F. (2007) In Kosuke Imai, G.K. and Olivia, L. (eds), *Zelig: Everyone's Statistical Software*, http://gking.harvard.edu/zelig.