A Trans-ancestral Meta-Analysis of Genome-Wide Association Studies Reveals Loci Associated with Childhood Obesity

Jonathan P. Bradfield[1,2†], Suzanne Vogelezang[3,4,5†], Janine F. Felix[3,4,5], Alessandra Chesi[6,7], Øyvind Helgeland[8,9], Momoko Horikoshi[10,11,12], Ville Karhunen[13], Estelle Lowry[14], Diana L. Cousminer[6,7,15], Tarunveer S. Ahluwalia[16,17,18], Elisabeth Thiering[19,20], Eileen Tai-Hui Boh[21], Mohammad H. Zafarmand[22,23], Natalia Vilor-Tejedor[24,25,26,27,28], Carol A. Wang[29], Raimo Joro[30], Zhanghua Chen[31], William J. Gauderman[31], Niina Pitkänen[32,33], Esteban J. Parra[34], Lindsay Fernandez-Rhodes[35,36], Akram Alyass[37], Claire Monnereau[3,4,5], John A. Curtin[38], Christian T. Have[18], Shana E. McCormack[39,40], Mette Hollensted[18], Christine Frithioff-Bøjsøe[18,41], Adan Valladares-Salgado[42], Jesus Peralta-Romero[42], Yik-Ying Teo[21,43,44,45,46], Marie Standl[19], Jaakko T. Leinonen[47], Jens-Christian Holm[18,41,48], Triinu Peters[49], Jesus Vioque[26,50,51], Martine Vrijheid[24,25,26], Angela Simpson[38], Adnan Custovic[52], Marc Vaudel[8], Mickaël Canouil[53], Virpi Lindi[54], Mustafa Atalay[30], Mika Kähönen[55,56], Olli T. Raitakari[32,33,57], Barbera D.C. van Schaik[23], Robert I. Berkowitz[58], Shelley A. Cole[59], V. Saroja Voruganti[60], Yujie Wang[61], Heather M. Highland[61], Anthony G. Comuzzie[62], Nancy F. Butte[63], Anne E. Justice[61,64], Sheila Gahagan[65], Estela Blanco[65], Terho Lehtimäki[66,67], Timo A. Lakka[30,68,69], Johannes Hebebrand[49], Amélie Bonnefond[53,70], Niels Grarup[18], Philippe Froguel[53,70], Leo-Pekka Lyytikäinen[66,67,71], Miguel Cruz[42], Sayuko Kobes[72], Robert L. Hanson[72], Babette S. Zemel[39,73], Anke Hinney[49], Koon K. Teo[37,74], David Meyre[37,75], Kari E. North[61,76], Frank D. Gilliland[31], Hans Bisgaard[16], Mariona Bustamante[24,25,26], Klaus Bonnelykke[16], Craig E. Pennell[29], Fernando Rivadeneira[3,4,77], André G. Uitterlinden[4,77], Leslie J. Baier[72], Tanja G.M. Vrijkotte[22], Joachim Heinrich[19,78,79], Thorkild I.A. Sørensen[18,80,81], Seang-Mei Saw[21,82], Oluf Pedersen[18], Torben Hansen[18], Johan Eriksson[83,84,85], Elisabeth Widén[47], Mark I. McCarthy[10,11,86], Pål R. Njølstad[8,87], Christine Power[88], Elina Hyppönen[88,89,90], Sylvain Sebert[14,91,92], Christopher D. Brown[93], Marjo-Riitta Järvelin[13,14,91,94.95], Nicholas J. Timpson[81], Stefan Johansson[8,96], Hakon Hakonarson[1,97], Vincent W.V. Jaddoe[3,4,5‡] and Struan F.A. Grant*[1,6,7,15,39,40,97‡] for the Early Growth Genetics Consortium

† These authors contributed equally to this work.

‡ These authors jointly directed this work.

***To whom correspondence should be addressed:**
Struan F.A. Grant
1102D ARC, Children's Hospital of Philadelphia Research Institute
3615 Civic Center Boulevard, Philadelphia, PA 19104, USA
Tel: 267-426-2795; E-mail: grants@email.chop.edu

1 Center for Applied Genomics, Children's Hospital of Philadelphia, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.
2 Quantinuum Research LLC, San Diego, CA, USA.
3 The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
4 Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
5 Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
6 Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, United States of America
7 Center for Spatial and Functional Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, United States of America
8 KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway
9 Department of Genetics and Bioinformatics, Health Data and Digitalization, Norwegian Institute of Public Health, Oslo, Norway
10 Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK;
11 Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK
12 RIKEN Center for Integrative Medical Sciences, Yokohama, Kanagawa 230-0045, Japan
13 Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, W2 1PG, United Kingdom
14 Center for Life Course Health Research, Faculty of Medicine, University of Oulu, PO Box 8000, FI-90014 Oulun yliopisto, Finland
15 Department of Genetics, University of Pennsylvania, Philadelphia
16 Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital 2820, University of Copenhagen, Copenhagen, Denmark
17 Steno Diabetes Center Copenhagen, Gentofte, Denmark
18 Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
19 Helmholtz Zentrum München - German Research Center for Environmental Health, Institute of Epidemiology, Neuherberg, Germany
20 Ludwig Maximilians University of Munich, Dr. von Hauner Children's Hospital, Munich, Germany
21 Saw Swee Hock School of Public Health, National University of Singapore, National University Health System, Singapore, Singapore
22 Department of Public Health, Amsterdam Public Health Research Institute, Amsterdam UMC, University of Amsterdam, The Netherlands
23 Department of Clinical Epidemiology and Biostatistics, Amsterdam Public Health Research Institute,

Amsterdam UMC, University of Amsterdam, The Netherlands
24 ISGlobal, Institute for Global Health, Barcelona, Spain
25 Universitat Pompeu Fabra (UPF), Barcelona, Spain
26 CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain.
27 Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr. Aiguader 88, 08003 Barcelona, Spain
28 BarcelonaBeta Brain Research Center (BBRC), Pasqual Maragall Foundation, Barcelona, Spain
29 School of Medicine and Public Health, Faculty of Medicine and Health, The University of Newcastle
30 Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio Campus, Finland
31 Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA
32 Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku 20521, Finland
33 Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20014, Finland
34 Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada
35 Department of Biobehavioral Health, Pennsylvania State University
36 Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
37 Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
38 Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, The University of Manchester, Manchester Academic Health Science Centre, Manchester University NHS Foundation Trust
39 Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States of America
40 Division of Endocrinology and Diabetes, The Children's Hospital of Philadelphia, Philadelphia, PA, United States of America
41 The Children’s Obesity Clinic, Department of Pediatrics, Copenhagen University Hospital Holbæk, Holbæk, Denmark
42 Unidad de Investigacion Medica en Bioquımica, Hospital de Especialidades, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Mexico City, Mexico
43 NUS Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore, 119077, Singapore
44 Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore
45 Life Sciences Institute, National University of Singapore, Singapore, 117456, Singapore
46 Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, 138672, Singapore
47 Institute for Molecular Medicine Finland, FIMM, University of Helsinki, Finland
48 University of Copenhagen, Faculty of Health and Medical Sciences, Copenhagen, Denmark

49 Department of Child and Adolescent Psychiatry, University Hospital Essen, University of Duisburg-Essen, Germany
50 University Miguel Hernandez, Alicante, Spain
51 ISABIAL–FISABIO Foundation, Alicante, Spain.
52 Department of Paediatrics, Imperial College London, UK
53 CNRS UMR 8199, European Genomic Institute for Diabetes (EGID), Institut Pasteur de Lille, University of Lille, Lille, France
54 University of Eastern Finland Library, Kuopio, Finland
55 Department of Clinical Physiology, Tampere University Hospital, Tampere 33521, Finland
56 Department of Clinical Physiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland
57 Centre for Population Health Research, University of Turku and Turku University Hospital, Finland
58 Department of Child and Adolescent Psychiatry and Behavioral Sciences, Children's Hospital of Philadelphia, Philadelphia PA 19104, USA
59 Texas Biomedical Research Institute, San Antonio, TX
60 Department of Nutrition and Nutrition Research Institute, University of North Carolina at Chapel Hill
61 Department of Epidemiology, University of North Carolina at Chapel Hill
62 The Obesity Society
63 Department of Pediatrics, Baylor College of Medicine
64 Biomedical and Translational Informatics, Geisinger Health System
65 Center for Community Health, Department of Pediatrics, University of California at San Diego
66 Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland
67 Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland
68 Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland
69 Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland
70 Department of Medicine, Section of Genomics of Common Disease, Imperial College London, London, United Kingdom
71 Department of Cardiology, Heart Center, Tampere University Hospital, Tampere 33521, Finland
72 Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes Digestive and Kidney Diseases, NIH USA
73 Division of Gastroenterology, Hepatology and Nutrition, The Children's Hospital of Philadelphia, Philadelphia, PA, United States of America
74 Department of Medicine, McMaster University, Hamilton, Ontario, Canada
75 Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario, Canada
76 Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill

77 Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands
78 Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Ludwig Maximilians University of Munich, University Hospital Munich, Munich, Germany
79 Allergy and Lung Health Unit, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia
80 Department of Public Health, Section of Epidemiology, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
81 Medical Research Council Integrative Epidemiology Unit, Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, BS8 2BN, UK
82 Singapore Eye Research Institute, Singapore, Singapore.
83 Department of General Practice and Primary Health Care, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
84 Folkhälsan Research Center, Helsink Finland
85 Chronic Disease Prevention Unit, National Institute for Health and Welfare, Helsinki, Finland
86 Oxford NIHR Biomedical Research Centre, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Oxford OX3 9DU
87 Department of Pediatrics and Adolescents, Haukeland University Hospital, Bergen, Norway
88 Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, University College London, London, UK
89 Australian Centre for Precision Health, University of South Australia Cancer Research Institute, Adelaide, Australia;
90 South Australian Health and Medical Research Institute, Adelaide, Australia
91 Biocenter Oulu, University of Oulu, Aapistie 5, 90220 Oulu, Finland
92 Department for Genomics of Common Diseases, School of Medicine, Imperial College London, UK
93 Department of Genetics, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, USA
94 Unit of Primary Health Care, Oulu University Hospital, OYS, Kajaanintie 50, 90220 Oulu, Finland
95 Department of Life Sciences, College of Health and Life Sciences, Brunel University London, Kingston Lane, Uxbridge, Middlesex UB8 3PH, United Kingdom
96 Department of Medical Genetics, Haukeland University Hospital, Bergen, Norway
97 Division of Human Genetics, Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

Abstract

Although hundreds of GWAS-implicated loci have been reported for adult obesity-related traits, less is known about the genetics specific for early-onset obesity, and with only a few studies conducted in non-European populations to date. Searching for additional genetic variants associated with childhood obesity, we performed a trans-ancestral meta-analysis of thirty studies consisting of up to 13,005 cases ($\geq 95^{th}$ percentile of BMI achieved 2-18 years old) and 15,599 controls (consistently $<50^{th}$ percentile of BMI) of European, African, North/South American and East Asian ancestry. Suggestive loci were taken forward for replication in a sample of 1,888 cases and 4,689 controls from seven cohorts of European and North/South American ancestry. In addition to observing eighteen previously implicated BMI or obesity loci, for both early and late onset, we uncovered one completely novel locus in this trans-ancestral analysis (nearest gene: *METTL15*). The variant was nominally associated in only the European subgroup analysis but had a consistent direction of effect in other ethnicities. We then utilized trans-ancestral Bayesian analysis to narrow down the location of the probable causal variant at each genome-wide significant signal. Of all the fine-mapped loci, we were able to narrow down the causative variant at four known loci to fewer than ten SNPs (*FAIM2, GNPDA2, MC4R* and *SEC16B* loci). In conclusion, an ethnically diverse setting has enabled us to both identify an additional pediatric obesity locus and further fine-map existing loci.

Introduction

Obesity is having a dramatic impact on modern societies, leading to substantial health issues, with an overall prevalence among children already greater than 20% in many populations, including the USA(1). Obesity, considerably contributes to mortality in the United States, representing a key risk factor for cardiometabolic and other chronic diseases.

The complex trait of obesity is the outcome of an interaction between environmental and genetic risk components(2). An excess in adipose tissue is commonly seen as an imbalance between energy uptake and utilization, and although now viewed as a disease may have historically conferred an advantage when food availability was restricted and high levels of physical activity were normal(3). Overall, obesity affects approximately 50 million girls and 74 million boys worldwide(1); most crucially, the prevalence of childhood obesity is on the increase worldwide(1), meaning that the known comorbidities are also on the rise across many ethnicities(2).

While environmental factors clearly play a role in the pathogenesis of childhood obesity, there is also strong evidence for a genetic component to obesity risk from twin and family studies, with heritability estimates for BMI being as high as 70%(4) . Large-scale genome-wide association studies (GWAS) have now reported many hundreds of loci associated with BMI/obesity in adults, and principally in populations of European ancestry(6). However, some studies have investigated the genome-wide genetics of obesity and/or BMI in children(7-12), but these did not address sex-specific or trans-ancestral associations.

In childhood and adolescence, BMI varies widely with age. To that end, working with

the Center for Disease Control and Prevention definition of childhood obesity as being at or above the 95th percentile of BMI for age(13), we conducted a large-scale trans-ancestral GWAS meta-analysis of the trait to uncover additional loci in order to provide further biological insight into this condition.

Results

In order to identify novel genetic variants associated with childhood obesity, we performed a two-stage trans-ancestral meta-analysis consisting of: Stage 1) thirty genome-wide genotyped cohorts augmented with genetic data imputed to the 1000G-reference panel for discovery efforts, and Stage 2) seven genotyped cohorts queried for SNPs which attained suggestive association in Stage 1 for the replication effort. The Stage 1 effort consisted of 13,005 cases ($\geq 95^{th}$ percentile of BMI achieved between 2 and 18 years old) and 15,599 controls ($<50^{th}$ percentile of BMI consistent throughout all measures during childhood). Stage 2 consisted of 1,888 cases and 4,489 controls. Each cohort was classified into four different groups based on ancestral makeup (either self-report or determined by PCA): European (Stage 1: 8,613 cases and 12,696 controls; Stage 2: 921 cases and 1,930 controls), African (Stage 1: 3,282 cases and 1,456 controls), American/Hispanic (Stage 1: 986 cases and 993 controls; Stage 2: 967 cases and 2759 controls) and East Asian group (Stage 1: 124 cases and 454 controls - consisting of East Asian ancestry samples from the United States and Singapore). The study characteristics are outlined in **Table S1**.

Stage 1: primary meta-analysis

Inverse-variance weighted fixed-effects meta-analyses, as implemented with METAL, within each of the four major continental ancestries was used to estimate effect sizes for the input into the trans-ancestral analysis using MANTRA. Sentinel SNPs were chosen by examining blocks of associated SNPs and choosing the SNP with the maximum Bayes factor (BF) in each block. New blocks were determined by distance greater than 100Kb

between successive SNPs with a $\log_{10}$ BF >= 4. The trans-ancestral analysis yielded a total of 82 independent loci reaching suggestive association ($\log_{10}$ BF >= 4.0) while there were 11 independent loci reaching genome-wide association ($\log_{10}$BF >= 6.0) (**Table S2)**. A $\log_{10}$ BF of 6.0 is equivalent to a p-value of 5.0×10^{-8}. A $\log_{10}$ BF of 4.0 is equivalent to a p-value of 5.0×10^{-6}. The Manhattan plot of the trans-ancestral meta-analysis is shown in **Figure 1**.

Stage 2: replication

The 82 independent SNPs found in the first stage of the analysis were taken forward and genotyped in the Stage 2 cohorts. In total, following the combined Stage 1 and Stage 2 effort, eighteen loci achieved genome-wide significance ($\log_{10}$ Bayes Factor >= 6.0) in the meta-analysis (**Table 1**). Of the eighteen genome-wide significant loci found in the analysis, eight SNPs (*TNNI3K, SEC16B, TMEM18, ADCY3, FAIM2, FTO, HOXB5* and *MC4R*) were found to be in linkage disequilibrium (LD) (r^2 >= 0.2, European 1000 genomes project phase 3) with variants previously shown to be associated with childhood obesity(7). Two SNPs at the *GNPDA2* and *TFAP2B* loci were in LD (r^2 >= 0.2, European 1000 genomes project phase 3) with variants previously shown to be associated with childhood BMI(9). Six of the SNPs at loci (*RANBP17, CALCR, BDNF, ADCY9*, and both variants near *CBLN4*) are in LD (r^2 >= 0.2, European 1000 Genomes Project Phase 3) with variants associated in the most recent adult BMI meta-analysis(6). After a search of the GWAS catalog, we found that two of the SNPs at two loci (*GPR1* and *METTL15*) were not in LD (r^2 < 0.2) with any variant known to be associated with childhood or adult BMI or related traits in the GWAS catalogue. But it is noted that the *GPR1* variant had an r^2 = 0.19

with a variant we reported on previously(9) (rs13387838) as associated with childhood BMI. To further assess the novelty of the *GPR1* variant, we performed an approximate conditional regression analysis of rs114670539 conditioning on rs13387838. The *P*-value of rs114670539 changed from 4.52×10^{-8} pre-conditioning to 5.94×10^{-8} post-conditioning in the Stage 1 European samples, suggesting that it is indeed independent of rs13387838. With a subsequent search of Phenoscanner, however, we found that the *GPR1* variant (rs114670539) yielded a genome-wide association to "comparative body size at age 10" in an unpublished UK Biobank GWAS (https://www.nealelab.is/uk-biobank). The novel *METTL15* variant (rs10835310) showed a genome-wide significant association to "comparative height size at age 10" in the same unpublished UK BioBank GWAS, but no genome-wide association to any metabolic traits. A regional association plot for the novel locus in the European sub-analysis for the genome-wide Stage 1 analysis is shown in **Figures S1**.

Subsequent conditional analyses revealed a novel independent signal at *TMEM18* (rs62104180, r^2=0.0008 with the previously reported rs7579427; MAF<5%) **Table** 1. **A** review of Phenoscanner revealed this variant to be associated with a number of metabolic traits in the UK Biobank, including BMI.

Heritability and Genetic Correlation Analyses

We sought to estimate the genome-wide common SNP heritability of childhood obesity and to calculate the genetic correlation of childhood obesity to other diseases. We used the LD score regression web interface called LDhub(14) to measure the common SNP heritability of childhood obesity (h^2 = 0.33) in the European summary statistics only, given

that it was the only dataset of sufficient sample size. Out of 219 traits with measured heritability, childhood obesity was ranked in the top 10% of traits. Childhood obesity had a similar common SNP heritability to three pubertal growth traits (Difference in height between adolescence and adulthood, age 14, $h^2 = 0.45$; Height, Females at age 10 and males at age 12, $h^2 = 0.43$; Difference in height between childhood and adulthood, age 8, $h^2 = 0.33$) but adult BMI, $h^2 = 0.19$, had a lower heritability. We also used LD score regression to assess the degree of genetic correlation between the European meta-analysis and other traits. The European meta-analysis summary statistics were uploaded to LDhub and compared to 235 other traits that were present on the file server. Statistical significance and genetic correlation were assessed with LDSC. Out of the 235 traits comparisons, 32 were significant after Bonferroni correction ($P < 0.00021$). There were traits that were positively or negatively genetically correlated with childhood obesity. While the most significant positive genetic correlation was with adult BMI ($r_g = 0.84$, $p = 3.4\times10^{-91}$) and the most significant negative genetic correlation was with age at menarche ($r_g = -0.40$, $p = 1.5\times10^{-24}$, **Table S3**), there were other less obvious genetic correlations such as negative genetic correlations with college completion and years of schooling and positive genetic correlations with excessive daytime sleepiness and squamous cell lung carcinoma.

We also compared our results to the largest adult BMI GWAS dataset currently available. We used 698 independently associated SNPs from Yengo *et al*(6) to compare the effect sizes between adult BMI and childhood obesity. We leveraged SNPs that were genome-wide significant in single SNP analyses. We extracted the effect sizes for these SNPs from our European Stage 1 analysis and compared them to the adult BMI effect sizes (correlation = 0.76) Figure S2. 562 out of 698 SNPs associated with adult BMI had the

same direction of effect in childhood obesity.

Functional Analysis and Fine Mapping

The trans-ancestral meta-analysis results were subsequently used to fine-map the genome-wide significant loci through credible set analysis. Four loci had 99% credible sets with fewer than ten SNPs (*FAIM2, GNPDA2, MC4R* and *SEC16B* loci). Even though the non-European samples formed a minority in the analysis, they enabled refinement of the interval within each of the 99% credible sets; indeed, none of the four loci with 99% credible sets of fewer than ten SNPs in the trans-ancestral analysis had credible sets fewer than ten SNPs in the European-only analysis. The *FAIM2* locus was refined to six SNPs, two of which are in the 3' untranslated region of the gene, and all residing within a 17kb region on chromosome 12 (hg19: 50,246,252-50,263,148). The *GNPDA2* locus also yielded six SNPs in the 99% credible set, all residing within 12kb of each other on chromosome 4 (hg19: 4,175,691-45,187,622). The signal near *MC4R* yielded four SNPs in the 99% credible set residing within 31kb of each other on chromosome 18 (hg19: 57,824,038-57,854,694). Finally, the *SEC16B* locus had five SNPs in the 99% credible set, which were all within 11kb of each other on chromosome 1 (hg19: 177,889,025-177,899,121) (**Table S4**).

All 21 of the variants in the four 99% credible sets were analyzed with the Ensembl Variant Effect Predictor(15) to access the enrichment of various functional groups in these sets. Intergenic variants were the most common predicted category with 43% of variants, 21% of variants were labeled as downstream gene variants which lie 3' of a gene. The downstream variants were concentrated around *SEC16B* and *FAIM2*. Variants located in

regulatory regions accounted 15% of the variants intronic variants represented 9% of variants. 3' untranslated region variants of *FAIM2* represented 9% of variants and one variant was in a transcription factor binding site.

Lastly, in order to attempt to place these signals in to a functional context, we investigated whether the suggestively associated variants were likely to share the same causal variant as an expression quantitative trait loci (eQTLs) of a nearby gene. We conducted colocalization analyses with GTEx v7 for all loci with $\log_{10}$BF>=4 (**Table S5**). This analysis yielded significant colocalizations at two loci across a range of tissues. The sentinel variant rs2206277 yielded a colocalization with an eQTL of *TFAP2B* in tibial nerve tissue, while rs4077678 showed significant colocalizations in numerous tissues. The most significant eQTL and tissue pair for rs4077678 was *DNAJC27* in whole blood, *ADCY3* in whole blood, *CENPO* in whole blood and *DNAJC27-AS1* in brain cerebellum. The additional significant colocalizations can be found in **Table S5**.

Discussion

Our trans-ancestral GWAS meta-analysis represents a large genome-wide survey of childhood obesity and allowed for the detection of loci not readily picked up in European only ancestral populations. We confirmed eighteen loci previously reported for childhood obesity or other metabolic phenotypes and identified one novel locus, namely at *METTL15,* associated with childhood obesity. Furthermore, the large overlap of at least nominally significant SNPs in both meta-analyses of pediatric obesity and adult BMI points to a shared genetic basis of these traits, at different times in the life course. The genetic correlation between childhood obesity and adult BMI was confirmed using LD-score regression, along with a negative genetic correlation between childhood obesity and age at menarche.

Although functional efforts are required to identify the actual effector genes at these loci, using similar approaches to what were applied to *FTO* locus which led to the implication of *IRX3* and *IRX5*(16-19), no inferences could be made from eQTLs for our novel childhood obesity loci. For the novel locus *METTL15*, the actual effector gene may be the well-established adult obesity *BDNF* gene that resides in the same topologically associating domain (TAD). Furthermore, rs2749808 near *CBLN4* gene is intergenic and may influence *MC3R*, given that it has already been strongly implicated in the pathogenesis of obesity(20, 21). We also further implicated *TMEM18* as the effector gene at this locus given the independent signal plus the rarer variants (MAF<5%) in the same neighborhood.

Trans-ancestral meta-analysis is particularly valuable in fine-mapping loci to narrow down the area harboring the causal variant. This is due to the different LD patterns present

in different ancestral populations. Despite known limitations to various fine-mapping approaches (such as whether or not the same set of variants were present in all input datasets), using MANTRA and credible set analysis we were able to narrow down the potential causal variant to fewer than ten variants at four different loci (*FAIM2, GNPDA2, MC4R* and *SEC16B*). Using the colocalization method, we were able to narrow down the putative causal variants and causal tissues for the *ADCY3 and TFAP2B* loci. There are colocalized eQTLs for various tissues with these associated loci that will need to be followed up in the future. The *ADCY3* locus is interesting in that there seems to be multiple genes (*DNAJC2*, *ADCY3*, *CENPO* and *DNAJC27-AS1*) colocalizing with the rs4077678 locus in multiple tissues (Whole Blood, Tibial Nerve, Skin, Adipose, Lung, Pituitary, Esophagus and Cerebellum). Whether this is due to coordination in all the genes in these tissues is an open question.

As with our previous GWAS of childhood obesity, we continued to use the Center for Disease Control and Prevention (CDC) definition as at or above the 95th percentile of BMI for age(22), and indeed represents the general guide for clinical practice(23). This is driven by the fact that there is a complex relationship between BMI and body fat in childhood, where it varies over time and especially during puberty. The larger heritability of childhood obesity compared to adult BMI, along with the correlation of the effects of the two traits, suggests that childhood obesity is an effective proxy trait to find variants associated with adult BMI but at smaller sample sizes.

We have conducted a large-scale trans-ancestral two-stage GWAS for childhood obesity, where we robustly identified a novel childhood obesity. We have also shown that childhood is genetically very similar to adult BMI and with far greater numbers of samples

1 we would most likely see more significant loci in common with the two phenotypes. As
2 such, we have gained greater insights in the biology of obesity in the pediatric setting and
3 these loci warrant further functional follow up in order to provide greater potential
4 therapeutic insights.
5

Materials and Methods

Research Subjects

The Stage 1 dataset consisted of thirty genome-wide genotyped studies from various ethnicities with BMI measured in childhood (2-18 years old) except GOYA which included some time points between 18-19 years old. The participating cohorts in these analyses were: the Children's Hospital of Philadelphia (CHOP) Study, the Generation R Study (GENR), the Singapore Cohort study Of the Risk factors for Myopia (SCORM), the Avon Longitudinal Study of Parents and Children (ALSPAC), the Western Australian Pregnancy Cohort (Raine) Study, the Amsterdam Born Children and their Development-Genetic Enrichment (ABCD-GE) Study, the Copenhagen Prospective Study on Asthma in Childhood (COPSAC2000), the French Obesity of the Youth (OBE) Study, the German Infant Study on the influence of Nutrition Intervention PLUS environmental and genetic influences on allergy development (GINIplus) / the Influence of life-style factors on the development of the immune system and allergies in East and West Germany (LISA) Study, the Genetics of Overweight Young Adults (GOYA) Study, the Helsinki Birth Cohort Study (HBCS), the HOLBAEK Study, the INfancia y Medio Ambiente [Environment and Childhood] (INMA) Project, the Manchester Asthma and Allergy Study (MAAS), Northern Finland Birth Cohort 1986 (NFBC86), Northern Finland Birth Cohort 1966 (NFBC66), the Physical Activity and Nutrition in Children (PANIC) Study, 1958 British Birth Cohort (1958BC), Young Finns Study (YFS), the Children's Health Study (CHS), and the MEXICO Study. Further information on the 1st stage cohorts is found in **Table S1**.

The Stage 2 dataset consisted of seven targeted genotype studies with BMI measured in childhood (ages 2-18 years) except the FAMILY study which included some time points less than 2 years of age. These studies were derived from the following participating cohorts: the Children's Health Study (CHS), the FAMILY study, The Norwegian Mother and Child Cohort Study (MoBa), the Santiago Longitudinal Study (SLS), the American Indians from Arizona Study and the VIVA la Familia Study (VIVA).

Trait Definition

Case and control definitions were based on national standard growth curves of BMI versus age for children from 2 to 18 years old. For instance, CHOP used the CDC standard growth curves (as featured in previous papers(13, 23)). The exception to this is the HBCS and 1958BC, as pediatric measures were made over two or six decades ago respectively so contemporary curves are not appropriate – in this case they generated their own reference curves. Cases were defined as an individual whose BMI is greater than or equal to the 95th percentile at any point in childhood. Controls were defined as an individual whose BMI was less than or equal to the 50th percentile consistently throughout childhood for all available measures.

Statistical Analysis

Each cohort was analyzed independently using a logistic regression framework (using an additive genetic model) where samples of different ancestry and samples genotyped on different SNP microarrays were analyzed separately. Eigenvectors calculated from

principal components analysis were used as covariates in the logistic regression by each cohort where appropriate.

For the discovery stage of the meta-analysis, data from high-density SNP arrays in each cohort were imputed to the 1000 Genomes integrated variant Phase 1 release v3 reference panel. Individual cohorts were responsible for their own pre-imputation sample exclusion criteria. Pre-imputation SNP quality control was applied by each individual cohort and it was recommended to remove SNPs with call rate < 95%, Hardy-Weinberg equilibrium $P < 1x10^{-4}$, and a minor allele frequency (MAF) filter that incorporated the accuracy of the genotyping of lower frequency SNPs. Cohort specific quality control and deviations from the recommended analysis parameters can be found in **Table S6.** Post-imputation quality control consisted of removing SNPs with MAF < 0.01, minor allele count < 10, r2_Hat < 0.3, proper_info < 0.4, or plink_info < 0.8 (depending on the software used for the statistical association analysis), as well as removing insertions and deletions.

Ancestral-specific inverse variance weighted fixed-effect meta-analysis was performed using METAL. Genomic control was applied to each cohort prior to meta-analysis and to the final meta-analysis statistics. SNPs were filtered out of the ancestral specific meta-analysis if the heterogeneity i-squared > 0.5 or if they were present in fewer than 50% of the total samples in the meta-analysis. Trans-ancestral meta-analysis was performed using MANTRA on the summary statistics obtained from the ancestral-specific meta-analyses (**Figure S3**).

Sentinel SNPs were selected at each locus from the suggestively associated results (log_{10} Bayes' factor > 4) as the SNP at each locus with the largest Bayes factor in the trans-ancestral results to maximize reproducibility across ethnicities. A locus was defined as a

collection of SNPs whose next physically closest suggestively associated SNP was within 100kb. This collection of SNPs were tested for association in the Stage 2 dataset.

The Stage 2 dataset was then combined with the Stage 1 dataset to test for association in the ancestral specific analyses and in the overall trans-ancestral analysis. The combined Stage 1 + Stage 2 results which resulted in a genome-wide significant results ($\log_{10}$ Bayes' factor > 6) are shown in Table 1. Stage 2 findings were only evaluated when combined with Stage 1, and not independently given the small sample size relative to Stage 1.

Sentinel SNPs that achieved genome-wide significance were queried against the GWAS catalogue and other available studies within Phenoscanner(25). A sentinel variant achieving $P< 5.0x10^{-8}$ in a prior metabolic GWAS was considered already discovered.

Conditional Regression

GCTA was used for pseudo-conditional regression analysis to identify variants independently associated with childhood obesity at the genome-wide significance level (trans-ancestral $\log_{10}$ Bayes factor > 6). The CHOP African American, European American, Hispanic, and East Asian samples were used to estimate the LD in GCTA. The genome-wide significant sentinel SNPs from the Stage 1 analysis were used as conditioning variants for the Stage 1 summary statistics. The ancestral-specific conditional analysis results were then analyzed in MANTRA to identify trans-ancestral significance. The top genome-wide significant SNP in the resulting conditional analysis results was then added into the list of conditioning SNPs to be analyzed again. When there were no more genome-wide significant SNPs, the conditional regression was then halted. A separate

pseudo-conditional regression analysis was carried out by conditioning rs114670539 on rs13387838 using the CHOP European American cohort to estimate LD.

LD Score Regression

LD score regression was performed using the LD Hub website interface (http://ldsc.broadinstitute.org/ldhub). The results from the European only meta-analysis were used for the LD score regression. Childhood obesity was compared against every phenotype available on LD Hub with the exception of the UK Biobank phenotypes and the previous childhood obesity meta-analysis.

eQTL Analysis Colocalization

We used *coloc* (with default parameters) to perform a Bayesian colocalization analysis comparing the meta-analysis results with GTEX version 7. We used variants with a $\log_{10}$ Bayes' factor >= 4 in the stage 1 analysis with 47 tissues from GTEX in the colocalization analysis. GWAS Bayes factors were used directly as input, while eQTL effect sizes and standard errors were used to estimate approximate Bayes factors for input. A significant colocalization was defined as PP.H3.abf + PP.H4.abf > 0.99 and PP.H4.abf / PP.H3.abf > 5(26). PP.H3.abf is defined as the posterior probability of 2 distinct causal variants. PP.H4.abf is defined as the posterior probability of 1 common causal variant.

Credible Set Analysis

The script credible_set_analysis.py located at https://github.com/edm1/Credible-set-analysis/blob/master/credible_set_analysis.py was used to calculate the 99% credible

sets for every genome-wide significant locus. The sum of the posterior probabilities was calculated from a sorted list of the most significant Bayes' factors until the cumulative sum was equal to or greater than 0.99. This set of SNPs was then considered the 99% credible set.

Acknowledgements

Cohort:

1958BC: 1958 British Birth Cohort acknowledges use of phenotype and genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping for the B58C-WTCCC subset was funded by the Wellcome Trust grant 076113/B/04/Z. The B58C-T1DGC genotyping utilised resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical 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. B58C-T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (CIMR), University of Cambridge, which is funded by Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of a Wellcome Trust Strategic Award (079895). The B58C-GABRIEL genotyping was supported by a contract from the European Commission Framework Programme 6 (018996) and grants from the French Ministry of Research.

ABCD-GE: We thank all participating hospitals, obstetric clinics, general practitioners and primary schools for their assistance in implementing the ABCD study. We also gratefully

acknowledge all the women and children who participated in this study for their cooperation.

The ABCD study has been supported by grants from The Netherlands Organisation for Health Research and Development (ZonMW) and The Netherlands Heart Foundation. Genotyping was funded by the BBMRI-NL grant CP2013-50. Dr M.H. Zafarmand was supported by BBMRI-NL (CP2013-50). Dr. T.G.M. Vrijkotte was supported by ZonMW (TOP 40–00812–98–11010).

AI-AZ: American Indian studies were funded by the Intramural Program of NIDDK, NIH.

*ALSPAC***:** We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. We would also like to acknowledge 23andMe for our genotyping collaboration. The UK Medical Research Council (Grant ref: 74882) the Wellcome Trust (Grant ref: 076467) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and JPB and SFAG will serve as guarantors for the contents of this paper.

CHOP: The authors thank the network of primary care clinicians and the patients and families for their contribution to this project and to clinical research facilitated by the Pediatric Research Consortium [PeRC]-The Children's Hospital of Philadelphia. R. Chiavacci, E. Dabaghyan, A. [Hope] Thomas, K. Harden, A. Hill, C. Johnson-Honesty, C.

Drummond, S. Harrison, F. Salley, C. Gibbons, K. Lilliston, C. Kim, E. Frackelton, F. Mentch, G. Otieno, K. Thomas, C. Hou, K. Thomas and M.L. Garris provided expert assistance with genotyping and/or data collection and management. The authors would also like to thank S. Kristinsson, L.A. Hermannsson and A. Krisbjörnsson of Raförninn ehf for extensive software design and contributions. This research was financially supported by an Institute Development Award from the Children's Hospital of Philadelphia, a Research Development Award from the Cotswold Foundation, the Daniel B. Burke Endowed Chair for Diabetes Research, the Children's Hospital of Philadelphia Endowed Chair in Genomic Research and NIH grant R01 HD056465.

CHS: The CHS was supported by the Southern California Environmental Health Sciences Center (grant P30ES007048); National Institute of Environmental Health Sciences (grants 5P01ES011627, ES021801, ES023262); P01ES009581, P01ES011627, P01ES022845, R01 ES016535, R03ES014046, P50 CA180905, R01HL061768, R01HL076647, R01HL087680, and RC2HL101651), the Environmental Protection Agency (grant #s RD83544101, R826708, RD831861, and R831845), and the Hastings Foundation.

$COPSAC_{2000}$: All funding received by COPSAC is listed on www.copsac.com. The Lundbeck Foundation (Grant no R16-A1694); The Ministry of Health (Grant no 903516); Danish Council for Strategic Research (Grant no 0603-00280B) and The Capital Region Research Foundation have provided core support to the COPSAC research center.

We express our deepest gratitude to the children and families of the $COPSAC_{2000}$ cohort study for all their support and commitment. We acknowledge and appreciate the unique efforts of the COPSAC research team.

ESSEN: We (JH, AH, TP) thank the following sources for funding or research: the German Ministry for Education and Research (National Genome Research Net-Plus 01GS0820), the German Research Foundation (DFG; HI865/2-1), the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreements n°245009 and n°262055.

FAMILY: We thank the participants of this study for their contribution. The FAMILY genetic study has been funded by the Heart and Stroke Foundation of Ontario (grant # NA 7293 "Early genetic origins of cardiovascular risk factors").D.M. is supported by a Canada Research Chair in Genetics of Obesity.

Generation R: The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the School of Law and Faculty of Social Sciences of the Erasmus University Rotterdam, the Municipal Health Service Rotterdam area, Rotterdam, the Rotterdam Homecare Foundation, Rotterdam and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond [STAR-MDC], Rotterdam. We gratefully acknowledge the contribution of children and parents, general practitioners, hospitals, midwives and pharmacies in Rotterdam. The generation and management of GWAS genotype data for the Generation R Study was done at the Genetic Laboratory of the Department of Internal

Medicine, Erasmus MC, the Netherlands. We would like to thank Karol Estrada, Dr. Tobias A. Knoch, Anis Abuseiris, Luc V. de Zeeuw, and Rob de Graaf, for their help in creating GRIMP, BigGRID, MediGRID, and Services@MediGRID/D-Grid, [funded by the German Bundesministerium fuer Forschung und Technology; grants 01 AK 803 A-H, 01 IG 07015 G] for access to their grid computing resources. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Manoushka Ganesh, Lizbeth Herrera and Marjolein Peters for their help in creating, managing and QC of the GWAS database.

The general design of Generation R Study is made possible by financial support from the Erasmus Medical Center, Rotterdam, the Erasmus University Rotterdam, the Netherlands Organization for Health Research and Development (ZonMw), the Netherlands Organisation for Scientific Research (NWO), the Ministry of Health, Welfare and Sport and the Ministry of Youth and Families. VWJ received an additional grant from the Netherlands Organization for Health Research and Development (VIDI 016.136.361) and a Consolidator Grant from the European Research Council (ERC-2014-CoG-648916). FR received an additional grant from the Netherlands Organization for Health Research and Development (VIDI 016.136.367). This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 633595 (DynaHEALTH) and No 733206 (LifeCycle).

GOYA: This genotyping of the study was conducted as part of the activities of the Danish Obesity Research Centre (DanORC) and the MRC centre for Causal Analyses in Translational Epidemiology (MRC CAiTE).

HBCS: We thank all study participants as well as everybody involved in the Helsinki Birth Cohort Study. Helsinki Birth Cohort Study has been supported by grants from the Academy of Finland (JGE grant no. 129369, 129907, 135072, 129255 and 126775), the Finnish Diabetes Research Society, Samfundet Folkhälsan, Novo Nordisk Foundation, Finska Läkaresällskapet, Juho Vainio Foundation, Signe and Ane Gyllenberg Foundation, Liv och Hälsa, University of Helsinki and Ministry of Education.

HOLBAEK: This study is part of the research activities in TARGET (The Impact of our Genomes on Individual Treatment Response in Obese Children, www.target.ku.dk), BIOCHILD (Genetics and Systems Biology of Childhood Obesity in India and Denmark, www.biochild.ku.dk), and the MicribLiver project funded by a grant from the Novo Nordisk Foundation (NNF15OC0016692). The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center at the University of Copenhagen partially funded by an unrestricted donation (Grant number NNF18CC0034900) from the Novo Nordisk Foundation (www.metabol.ku.dk). The study is part of The Danish Childhood Obesity Biobank; ClinicalTrials.gov ID-no.: NCT00928473, retrospectively registered June 25 2009.

The authors wish to thank all children and adolescents participating in the present study as well as their families. Additionally, we wish to thank Mrs. Oda Troest and Mrs. Birgitte Holløse for their invaluable assistance with blood samples and database.

This study received funding from the Innovation Fund Denmark (grant numbers 0603-00484B (TARGET) and 0603-00457B (BIOCHILD)) and The Novo Nordisk Foundation

(grant numbers NNF15OC0016544, NNF15OC0016692, and NNF18CC0034900) and The Region Zealand Health Scientific Research Foundation.

INMA: This study was funded by grants from Instituto de Salud Carlos III (CB06/02/0041, FIS PI041436, PI081151, PI041705, PS09/00432, FIS-FEDER 03/1615, 04/1509, 04/1112, 04/1931, 05/1079, 05/1052, 06/1213, 07/0314, 09/02647, PI06/0867, PI 07/0314 PI09/0009, PI 11 / 01007), Spanish Ministry of Science and Innovation (SAF2013-49108-R), European Commission (ENGAGE project and grant agreement HEALTH-F4-2007-201413), Fundació La Marató de TV3, Generalitat de Catalunya-CIRIT AGAUR 2014 SGR-1138, Conselleria de Sanitat Generalitat Valenciana. Part of the DNA extractions and genotyping was performed at the Spanish National Genotyping Centre (CEGEN-Barcelona). The authors would particularly like to thank all the participants for their generous collaboration. A full roster of the INMA Project Investigators can be found at http://www.proyectoinma.org/presentacion-inma/listado-investigadores/en_listado-investigadores.html.

MAAS: We would like to thank the children and their parents for their continued support and enthusiasm. We greatly appreciate the commitment they have given to the project. We would also like to acknowledge the hard work and dedication of the study team (post-doctoral scientists, research fellows, nurses, physiologists, technicians and clerical staff). This report includes independent research supported by National Institute for Health Research Respiratory Clinical Research Facility at Manchester University NHS Foundation Trust (Wythenshawe). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research

or the Department of Health. MAAS was supported by the Asthma UK Grants No 301 (1995-1998), No 362 (1998-2001), No 01/012 (2001-2004), No 04/014 (2004-2007), the BMA James Trust and Medical Research Council, UK (G0601361) and The Moulton Charitable Foundation (2004-current); the Medical Research Council (MRC) Grants G0601361, MR/K002449/1 and MR/L012693/1, and Angela Simpson is supported by the NIHR Manchester Biomedical Research Centre. The authors would like to acknowledge the North West Lung Centre Charity for supporting this project.

MEXICO: We thank the participants of this study for their contribution. This work was supported by the Consejo Nacional de Ciencia y Tecnología (CONACYT-México) with the grant SALUD-2013-C01-201471 (FONSECSSA/IMSS/ISSSTE). Computations were performed on the GPC supercomputer at the SciNet HPC Consortium, Canada. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund-Research Excellence; and the University of Toronto.

MOBA: The Norwegian Mother and Child Cohort Study was supported by grants from the European Research Council (AdG #293574), the Bergen Research Foundation ("Utilizing the Mother and Child Cohort and the Medical Birth Registry for Better Health"), Stiftelsen Kristian Gerhard Jebsen (Translational Medical Center), the University of Bergen, the Research Council of Norway (FRIPRO grant #240413), the Western Norway Regional Health Authority (Strategic Fund "Personalized Medicine for Children and Adults", and Open Grants («*Understanding infant weight biology through genomics and deep*

phenotyping» grant #912250), and the Norwegian Diabetes Foundation; the Research Council of Norway through its Centres of Excellence funding scheme (#262700), Better Health by Harvesting Biobanks (#229624) and The Swedish Research Council, Stockholm, Sweden (2015-02559), The Research Council of Norway, Oslo, Norway (FRIMEDBIO ES547711, March of Dimes (#21-FY16-121). The Norwegian Mother and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research, NIH/NIEHS (contract no N01-ES-75558), NIH/NINDS (grant no.1 UO1 NS 047537-01 and grant no.2 UO1 NS 047537-06A1).

NFBC: We thank Professor Paula Rantakallio (launch of NFBC1966 and initial data collection). We gratefully acknowledge the contributions of the participants in the Northern Finland Birth Cohort 1966 study and the Northern Finland Birth Cohort 1986. We also thank all the field workers and laboratory personnel for their efforts. NFBC1966 received financial support from University of Oulu Grant no. 65354, Oulu University Hospital Grant no. 2/97, 8/97, Ministry of Health and Social Affairs Grant no. 23/251/97, 160/97, 190/97, National Institute for Health and Welfare, Helsinki Grant no. 54121, Regional Institute of Occupational Health, Oulu, Finland Grant no. 50621, 54231. NFBC1986 received financial support from EU QLG1-CT-2000-01643 (EUROBLCS) Grant no. E51560, NorFA Grant no. 731, 20056, 30167, USA / NIHH 2000 G DF682 Grant no. 50945. MW was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 633212. MRJ and SS are supported by H2020-633595 DynaHEALTH action and academy of Finland EGEA-project (285547). VK is supported by H2020-721567 CAPICE project.

PANIC: The PANIC Study has financially been supported by grants from Ministry of Education and Culture of Finland, Ministry of Social Affairs and Health of Finland, Research Committee of the Kuopio University Hospital Catchment Area (State Research Funding), Finnish Innovation Fund Sitra, Social Insurance Institution of Finland, Finnish Cultural Foundation, Foundation for Paediatric Research, Diabetes Research Foundation in Finland, Finnish Foundation for Cardiovascular Research, Juho Vainio Foundation, Paavo Nurmi Foundation, Yrjö Jahnsson Foundation, and the city of Kuopio. Moreover, he PhD students and postdoctoral researchers of The PANIC Study have financially been supported by personal grants from the doctoral schools of Finnish universities and Finnish foundations. We are grateful to the members of the PANIC research team for their contribution in acquisition of data. We are also indebted to all children and their parents participating in the PANIC study.

Raine study: The Raine Study acknowledges the National Health and Medical Research Council (NHMRC) for their long term contribution to funding the study over the last 29 years. Core Management of the Raine study has been funded by the University of Western Australia (UWA), Curtin University, the UWA Faculty of Medicine, Dentistry and Health Sciences, the Raine Medical Research Foundation, the Telethon Kids Institute, the Women's and Infants Research Foundation, Edith Cowan University, Murdoch University, and the University of Notre Dame. This study was supported by the National Health and Medical Research Council of Australia [grant numbers 572613, 403981 and 003209] and the Canadian Institutes of Health Research [grant number MOP-82893]. The authors

gratefully acknowledge the assistance of the Western Australian DNA Bank (National Health and Medical Research Council of Australia National Enabling Facility). All analytic work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia.

SLS: The obesity component of the Santiago Longtiduainal Study (SLS) was funded by 5R01HL088309. The SLS genetic work was funded in part by a University of North Carolina Nutrition Research Institute internal pilot grant, and AHA grant 15GRNT25880008. Other components of the SLS were funded by NICHD and NIDA. We would like to thank the participants and their family members from the SLS.

VIVA: The Viva la Familia Study was supported by the National Institutes of Health (NIH) [DK080457], and the USDA/ARS [Cooperative Agreement 6250-51000-053]. Work performed at the Texas Biomedical Research Institute in San Antonio, Texas was conducted in facilities constructed with support from the Research Facilities Improvement Program of the National Center for Research Resources, NIH [C06 RR013556, C06 RR017515].

YFS: The Young Finns Study has been financially supported by the Academy of Finland: grants 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Competitive State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish

Foundation for Cardiovascular Research ; Finnish Cultural Foundation; The Sigrid Juselius Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; Diabetes Research Foundation of Finnish Diabetes Association; EU Horizon 2020 (grant 755320 for TAXINOMISIS); European Research Council (grant 742927 for MULTIEPIGEN project); and Tampere University Hospital Supporting Foundation.

Personal:
LFR is grateful to the Carolina Population Center for training support (T32 HD007168) and for general support (P2C HD050924).

AEJ was funded under NIH award K99HL130580-02

M.I.M. is a Wellcome Investigator (grant numbers 098381, 212259) and a NIHR Senior Investigator. Additional funding support from the Wellcome (090532, 203141) and the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

C.Power at UCL Institute of Child Health, with support from the National Institute for Health Research Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London.

D.L.C. is supported by American Diabetes Association grant #1-17-PDF-077.

Conflict of Interest Statement

Shana McCormack has participated in advisory boards for Rhythm Pharmaceuticals and Reata Pharmaceuticals. She is a site PI for a clinical trial supported by Levo Pharmaceuticals.

References

1 Collaboration, N.C.D.R.F. (2017) Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. *Lancet*, **390**, 2627-2642.

2 Heymsfield, S.B. and Wadden, T.A. (2017) Mechanisms, Pathophysiology, and Management of Obesity. *The New England journal of medicine*, **376**, 254-266.

3 Eckel, R.H. (2003) Obesity: a disease or a physiologic adaptation for survival? *Obesity Mechanisms and Clinical Management*, in press., 3–30.

4 Bray, M.S., Loos, R.J., McCaffery, J.M., Ling, C., Franks, P.W., Weinstock, G.M., Snyder, M.P., Vassy, J.L., Agurs-Collins, T. and Conference Working, G. (2016) NIH working group report-using genomic information to guide weight management: From universal to precision treatment. *Obesity (Silver Spring, Md*, **24**, 14-22.

5 Silventoinen, K., Jelenkovic, A., Sund, R., Hur, Y.M., Yokoyama, Y., Honda, C., Hjelmborg, J., Moller, S., Ooki, S., Aaltonen, S. *et al.* (2016) Genetic and environmental effects on body mass index from infancy to the onset of adulthood: an individual-based pooled analysis of 45 twin cohorts participating in the COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) study. *The American journal of clinical nutrition*, **104**, 371-379.

6 Yengo, L., Sidorenko, J., Kemper, K.E., Zheng, Z., Wood, A.R., Weedon, M.N., Frayling, T.M., Hirschhorn, J., Yang, J., Visscher, P.M. *et al.* (2018) Meta-analysis of genome-wide association studies for height and body mass index in approximately 700000 individuals of European ancestry. *Human molecular genetics*, **27**, 3641-3649.

7 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. *Nature genetics*, **44**, 526-531.

8 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 genetics*, **6**, e1000916.

9 Felix, J.F., Bradfield, J.P., Monnereau, C., van der Valk, R.J., Stergiakouli, E., Chesi, A., Gaillard, R., Feenstra, B., Thiering, E., Kreiner-Moller, E. *et al.* (2016) Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Human molecular genetics*, **25**, 389-403.

10 Zandona, M.R., Sangalli, C.N., Campagnolo, P.D., Vitolo, M.R., Almeida, S. and Mattevi, V.S. (2017) Validation of obesity susceptibility loci identified by genome-wide

association studies in early childhood in South Brazilian children. *Pediatr Obes*, **12**, 85-92.

11 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. *Nature genetics*, **41**, 157-159.

12 Meng, X.R., Song, J.Y., Ma, J., Liu, F.H., Shang, X.R., Guo, X.J. and Wang, H.J. (2014) Association study of childhood obesity with eight genetic variants recently identified by genome-wide association studies. *Pediatric research*, **76**, 310-315.

13 Flegal, K.M., Wei, R. and Ogden, C. (2002) Weight-for-stature compared with body mass index-for-age growth charts for the United States from the Centers for Disease Control and Prevention. *The American journal of clinical nutrition*, **75**, 761-766.

14 Zheng, J., Erzurumluoglu, A.M., Elsworth, B.L., Kemp, J.P., Howe, L., Haycock, P.C., Hemani, G., Tansey, K., Laurin, C., Early, G. *et al.* (2017) LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics (Oxford, England)*, **33**, 272-279.

15 McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R., Thormann, A., Flicek, P. and Cunningham, F. (2016) The Ensembl Variant Effect Predictor. *Genome biology*, **17**, 122.

16 Smemo, S., Tena, J.J., Kim, K.H., Gamazon, E.R., Sakabe, N.J., Gomez-Marin, C., Aneas, I., Credidio, F.L., Sobreira, D.R., Wasserman, N.F. *et al.* (2014) Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature*, **507**, 371-375.

17 Claussnitzer, M., Dankel, S.N., Kim, K.H., Quon, G., Meuleman, W., Haugen, C., Glunk, V., Sousa, I.S., Beaudry, J.L., Puviindran, V. *et al.* (2015) FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N Engl J Med*, **373**, 895-907.

18 Rosen, C.J. and Ingelfinger, J.R. (2015) Unraveling the Function of FTO Variants. *The New England journal of medicine*, **373**, 964-965.

19 Hunt, L.E., Noyvert, B., Bhaw-Rosun, L., Sesay, A.K., Paternoster, L., Nohr, E.A., Davey Smith, G., Tommerup, N., Sorensen, T.I. and Elgar, G. (2015) Complete re-sequencing of a 2Mb topological domain encompassing the FTO/IRXB genes identifies a novel obesity-associated region upstream of IRX5. *Genome Med*, **7**, 126.

20 Tao, Y.X. (2010) Mutations in the melanocortin-3 receptor (MC3R) gene: Impact on human obesity or adiposity. *Curr Opin Investig Drugs*, **11**, 1092-1096.

21 Koya, C., Yu, T., Strong, C. and Tsai, M.C. (2018) Association between Two Common Missense Substitutions, Thr6Lys and Val81Ile, in MC3R Gene and Childhood Obesity: A Meta-Analysis. *Child Obes*, **14**, 218-226.

22 Koplan, J.P., Liverman, C.T. and Kraak, V.I. (2005) Preventing childhood obesity: health in the balance: executive summary. *J Am Diet Assoc*, **105**, 131-138.

23 Himes, J.H. and Dietz, W.H. (1994) Guidelines for overweight in adolescent preventive services: recommendations from an expert committee. The Expert Committee on Clinical Guidelines for Overweight in Adolescent Preventive Services. *The American journal of clinical nutrition*, **59**, 307-316.

24 Cole, T.J., Freeman, J.V. and Preece, M.A. (1995) Body mass index reference curves for the UK, 1990. *Archives of disease in childhood*, **73**, 25-29.

25 Staley, J.R., Blackshaw, J., Kamat, M.A., Ellis, S., Surendran, P., Sun, B.B., Paul, D.S., Freitag, D., Burgess, S., Danesh, J. *et al.* (2016) PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics (Oxford, England)*, **32**, 3207-3209.

26 Guo, H., Fortune, M.D., Burren, O.S., Schofield, E., Todd, J.A. and Wallace, C. (2015) Integration of disease association and eQTL data using a Bayesian colocalisation approach highlights six candidate causal genes in immune-mediated diseases. *Human molecular genetics*, **24**, 3305-3313.

27 Zhu, Z., Zheng, Z., Zhang, F., Wu, Y., Trzaskowski, M., Maier, R., Robinson, M.R., McGrath, J.J., Visscher, P.M., Wray, N.R. *et al.* (2018) Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun*, **9**, 224.

Table 1: Top independent novel and known SNPs that reached genome-wide significance (log_{10} Bayes Factor >= 6) in the conditional or trans-ancestral meta-analyses. Betas and standard errors (SE) are shown for each ancestral specific sub-analysis. The heterogeneity (Het) of the Bayes' Factor (BF) is also shown. If the variant (or in LD ($r^2 > 0.2$)) was previously found in a metabolic phenotype, that phenotype is shown. "--" indicates that the variant did not pass quality control in that ancestral grouping. The first allele is the effect allele for which the beta applies.

Chr	Position	Marker	Nearest Gene	Analysis	Alleles	African			Asian			European			Hispanic			Trans-ancestral		Previously Known
						Beta	SE	P	Beta	SE	P	Beta	SE	P	Beta	SE	P	BF	Het	Metabolic Phenotype
1	74,983,835	rs10493544	*TNNI3K*	Full	t/c	0.18	0.06	3.86E-03	-0.36	0.26	1.62E-01	0.14	0.02	1.14E-13	0.02	0.05	6.45E-01	11.81	0.35	Childhood Obesity
1	177,889,025	rs539515	*SEC16B*	Full	a/c	-0.19	0.05	2.77E-04	0.08	0.25	7.37E-01	-0.18	0.02	2.68E-14	-0.24	0.06	4.06E-05	18.07	0.16	Childhood Obesity
2	466,003	rs62104180	*TMEM18*	Conditional	a/g	--	--	--	--	--	--	-0.32	0.06	4.52E-09	--	--	--	7.10	0.00	Adult BMI
2	631,183	rs7579427	*TMEM18*	Full	a/c	0.26	0.07	2.98E-04	-0.25	0.29	3.93E-01	0.21	0.02	8.54E-18	0.25	0.07	5.96E-04	20.25	0.20	Childhood Obesity
2	25,122,840	rs4077678	*ADCY3*	Full	c/g	-0.16	0.06	1.58E-02	-0.13	0.17	4.35E-01	-0.14	0.02	1.44E-13	-0.11	0.06	7.42E-02	13.38	0.10	Childhood Obesity
2	207,064,335	rs114670539	*GPR1*	Full	t/c	0.14	0.19	4.57E-01	--	--	--	0.26	0.05	2.14E-08	0.03	0.17	8.79E-01	6.12	0.23	Comp. body size at age 10
4	45,187,622	rs925494	*GNPDA2*	Full	t/c	0.24	0.06	4.21E-05	-0.02	0.21	9.25E-01	0.10	0.02	4.04E-07	0.19	0.08	1.50E-02	8.57	0.37	Childhood BMI
5	170,599,327	rs2053682	*RANBP17*	Full	a/c	0.15	0.05	1.94E-03	0.27	0.22	2.15E-01	0.09	0.02	6.76E-06	0.11	0.05	3.64E-02	6.73	0.13	Adult BMI
6	50,798,526	rs2206277	*TFAP2B*	Full	t/c	0.13	0.06	4.95E-02	0.02	0.20	9.15E-01	0.14	0.02	5.93E-10	0.21	0.05	5.39E-05	11.63	0.14	Childhood BMI
7	93,269,367	rs10224397	*CALCR*	Full	a/g	0.18	0.05	7.05E-04	0.07	0.18	6.83E-01	0.09	0.02	2.18E-06	0.08	0.07	2.51E-01	6.53	0.15	Adult BMI
11	27,667,236	rs17309874	*BDNF*	Full	a/g	0.12	0.08	1.13E-01	--	--	--	0.12	0.02	2.59E-08	0.20	0.07	2.82E-03	8.52	0.11	Adult BMI
11	28,355,657	rs10835310	*METTL15*	Full	t/c	0.10	0.05	5.41E-02	0.05	0.19	7.79E-01	0.10	0.02	3.90E-08	0.04	0.08	6.25E-01	6.26	0.13	Novel
12	50,263,148	rs7132908	*FAIM2*	Full	a/g	--	--	--	0.19	0.20	3.33E-01	0.15	0.02	4.00E-16	0.23	0.07	5.70E-04	16.39	0.14	Childhood Obesity
16	4,017,567	rs2540031	*ADCY9*	Full	a/t	0.12	0.06	3.93E-02	0.46	0.19	1.51E-02	0.08	0.02	2.75E-05	0.17	0.06	2.64E-03	6.33	0.30	Adult BMI
16	53,806,453	rs56094641	*FTO*	Full	a/g	-0.17	0.07	2.02E-02	-0.48	0.24	4.16E-02	-0.21	0.02	1.31E-28	-0.28	0.06	6.55E-06	31.88	0.19	Childhood Obesity
17	46,664,608	rs2740752	*HOXB5*	Full	t/c	0.18	0.05	1.06E-03	-0.07	0.22	7.52E-01	0.11	0.03	1.34E-04	0.20	0.06	7.89E-04	6.81	0.15	Childhood Obesity
18	57,829,135	rs6567160	*MC4R*	Full	t/c	-0.22	0.06	8.66E-05	--	--	--	-0.15	0.02	1.16E-11	-0.19	0.11	6.97E-02	13.20	0.14	Childhood Obesity
20	54,149,014	rs2749808	*CBLN4*	Full	t/c	-0.12	0.05	1.28E-02	-0.14	0.17	4.20E-01	-0.10	0.02	1.12E-06	-0.08	0.05	1.44E-01	6.47	0.09	Adult BMI
20	54,482,276	rs1437206	*CBLN4*	Full	t/c	0.18	0.05	9.67E-05	0.23	0.28	4.13E-01	-0.10	0.02	3.23E-06	-0.21	0.08	1.16E-02	7.43	1.00	Adult BMI

Legends to Figures

Figure 1: Manhattan plot of the trans-ancestral meta-analysis of the childhood obesity Stage 1 results. Bayes' factors (BF) less than 0 have been represented by a value of 0. The y-axis is the $\log_{10}$ of the BF. Sentinel SNPs from loci that achieved at least $\log_{10}$ BF >= 4 were taken forward to Stage 2.