Open access • Posted Content • DOI:10.1101/2021.02.23.21251958

Genome-wide association study of body fat distribution traits in Hispanics/Latinos from the HCHS/SOL Study — Source link

Anne E. Justice, Anne E. Justice, Kristin L. Young, Stephanie M. Gogarten ...+30 more authors

Institutions: University of North Carolina at Chapel Hill, Geisinger Health System, University of Washington, University of Arizona ...+14 more institutions

Published on: 25 Feb 2021 - medRxiv (Cold Spring Harbor Laboratory Press)

Genome-wide association study of body fat distribution traits in Hispanics/Latinos from the HCHS/SOL Study

Anne E. Justice[1,2,¶,*], Kristin Young[2,¶], Stephanie M. Gogarten[3], Tamar Sofer[3], Misa Graff[2], Shelly Ann Love[2], Yujie Wang[2], Yann C. Klimentidis[4], Miguel Cruz[5], Xiuqing Guo[6], Fernando Hartwig[7], Lauren Petty[8], Jie Yao[6], Matthew A. Allison[9], Jennifer E. Below[8], Thomas A. Buchanan[10], Yii-Der Ida Chen[6], Mark O. Goodarzi[11], Craig Hanis[12], Heather M. Highland[2], Willa A. Hsueh[13], Eli Ipp[14], Esteban Parra[15], Walter Palmas[16], Leslie J. Raffel[17], Jerome I. Rotter[6], Jingyi Tan[6], Kent D. Taylor[6], Adan Valladares[5], Anny H. Xiang[18], Lisa Sánchez-Johnsen[19], Carmen R. Isasi[20], Kari E. North[2]

[1] Population Health Sciences, Geisinger Health System, Danville, PA, USA

[2] Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[3] Department of Biostatistics, University of Washington, Seattle, WA

[4] Department of Epidemiology and Biostatistics, University of Arizona, Tucson, AZ, USA

[5] Unidad de Investigacion Medica en Bioquimica, Hospital de Especialidades, CMNSXX1-IMSS, Mexico City, Mexico

[6] The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA USA

[7] Center for Epidemiological Research, Universidade Federal de Pelotas, Pelotas, Brazil

[8] Vanderbilt Genetics Institute, Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA

[9] Division of Preventive Medicine, Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA USA

[10] Department of Physiology and Biophysics and Department of Medicine, Keck School of Medicine of USC, Los Angeles, CA USA

[11] Division of Endocrinology, Diabetes, and Metabolism, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA USA

[12]Human Genetics Center, Department of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX 77030 USA

[13] Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Wexner Medical Center, The Ohio State University, Columbus, OH USA

[14] Department of Medicine, Endocrinology, Diabetes & Metabolism, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA USA

[15]Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada

[16] Department of Medicine, Columbia University Medical Center, New York, NY USA

[17] Department of Pediatrcs, Division of Genetic and Genomic medicine, University of California, Irvine, CA USA

[18] Department of Research and Education, Kaiser Permanente Southern California, Pasadena, CA, USA

[19] Department of Family Medicine, Rush University Medical Center, Chicago, Illinois, USA

[20] Department of Epidemiology & Population Health, Department of Pediatrics, Albert Einstein College of Medicine, Bronx, NY

* Corresponding author

E-mail: aejustice1@geisinger.edu

¶ The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

Author contributions

Conceived of Study Design: AEJ, KY, KEN; Drafted Manuscript: AEJ, KY, KEN, MG, SAL; Acquisition of Data: AEJ, KEN, CL, CI; Data Preparation: SMG, TS; Performed Statistical Analyses: SMG, TS, AEJ, LEP, YW; Prepared Figures and Table: AEJ, KL, SAL, MG, SMG, TS; Performed Look-ups: AEJ; Supervised the work: KEN, CL, CI. All authors revised and approved the manuscript, assisted in interpretation of results, and agree to be held accountable for the content.

Abstract

Central obesity is a leading health concern with a great burden carried by ethnic minority populations, and especially Hispanics/Latinos. Genetic factors contribute to the obesity burden overall and to inter-population differences. We aim to: 1) identify novel loci associated with central adiposity measured as waist-to-hip ratio (WHR), waist circumference (WC), and hip circumference (HIP), all adjusted for body mass index (adjBMI), using the Hispanic Community Health Study/Study of Latinos (HCHS/SOL); 2) determine if differences in genetic associations differ by background group within HCHS/SOL; 3) determine whether previously reported association regions generalize to HCHS/SOL. Our analyses included 7,472 women and 5,200 men of mainland (Mexican, Central and South American) and Caribbean (Puerto Rican, Cuban, and Dominican) background residing in the US, with genome-wide array data imputed to the 1000 genomes Phase I multiethnic reference panel. We analyzed associations stratified by sex in addition to sexes combined using linear mixed-model regression. We identified 16 variants for WHRadjBMI, 22 for WCadjBMI, and 28 for HIPadjBMI that reached suggestive significance ($P<1x10^{-6}$). Many of the loci exhibited differences in strength of associations by ethnic background and sex. We brought a total of 66 variants forward for validation in nine cohort studies (N=34,161) with participants of Hispanic/Latino, African and European descent. We confirmed four novel loci (ancestry-specific $P<0.05$ in replication, consistent direction of effect with HCHS/SOL, and $P<5x10^{-8}$ after meta-analysis with HCHS/SOL), including rs13301996 in the sexes-combined analysis, and rs79478137 for women-only for WHRadjBMI; rs28692724 in women-only for HIPadjBMI; and rs3168072 in the sexes combined analysis for WCadjBMI. Also, a total of eight previously reported WHRadjBMI association regions, 12 for HIPadjBMI, and 10 for WCadjBMI generalized to HCHS/SOL. Our study findings highlight the importance of large-scale genomic studies in ancestrally diverse Hispanic/Latino populations for identifying and characterizing central obesity-susceptibility that may be ancestry-specific.

Introduction

Obesity, and especially central obesity, is a leading risk factor for metabolic and cardiovascular diseases, with the greatest burden carried by minority populations (1-4), particularly Hispanic/Latino Americans and African Americans (5). Emerging evidence suggests that genetic factors may contribute not only to the obesity burden overall, explaining 40% to 70% of the inter-individual variation (6), but also to population-specific differences in obesity susceptibility (7-12). For example, although a majority of the >1000 genome-wide association study (GWAS)-identified obesity (body mass index (BMI), waist-to-hip ratio (WHR), waist circumference (WC), hip circumference (HIP), and body fat percentage) loci generalize across populations (13-20), recent studies in populations of Asian (19, 20) and African (16, 21) ancestry have revealed a number of novel and population-specific loci. These observations highlight the importance of large-scale genomic studies in ancestrally diverse populations, including Hispanic/Latinos, to identify obesity-susceptibility loci, and more specifically, alleles that are ancestry-specific and may thus partly explain disparities. However, no large-scale GWAS for any obesity-related traits has been performed in Hispanic/Latino populations, despite their increased prevalence of obesity.

While obesity is commonly assessed by BMI, measures of central adiposity such as WHR and WC are predictors of increased cardiometabolic risk independent of BMI (22-25). Here we consider three measures of central obesity: WHR, WC, and HIP after accounting for overall body size, measured as body mass index (BMI) (WHRadjBMI, WCadjBMI, HIPadjBMI). Larger WHR indicates increased abdominal fat and is associated with increased risk for type 2 diabetes (T2D) and cardiovascular disease (CVD) (26-28), while smaller WHR indicates a proportionately greater fat accumulation around the hips, and is associated with lower risk for T2D, hypertension and dyslipidemia (29). Previous GWAS have identified WHR, WC, and HIP loci that are enriched for association with other cardiometabolic traits and suggested that different fat distribution patterns can have distinct genetic underpinnings (30-32). Identifying

104 genetic risk variants across these traits in Hispanic/Latinos may provide insights into these mechanisms
105 and highlight population-specific variants that increase susceptibility to obesity in specific groups.
106 We aimed to: 1) identify novel genetic loci associated with central obesity, measured here as
107 WHRadjBMI, WCadjBMI, and HIPadjBMI, in Hispanics/Latinos; 2) determine if differences in genetic
108 associations by background group (mainland or Caribbean) and sex exist in HCHS/SOL; and 3) assess
109 generalization of central adiposity associated loci discovered in European, African and multi-ethnic
110 studies to Hispanics/Latinos.
111

112 Results

113

114 Discovery

115 We identified 16 loci for WHRadjBMI, 22 for WCadjBMI, and 28 for HIPadjBMI that exhibited
116 suggestive evidence of association in the HCHS/SOL in at least one stratum (**Table 1**, **Figures 1-3**,
117 **Supplementary Tables 2-4**, **Supplementary Figures 1-12**). For WHRadjBMI, we identified four loci that
118 reach suggestive significance (P<1x10^{-6}) in the combined sexes, including rs12435790 near *KIAA0391*,
119 which is within a previously reported WHRadjBMI locus (+/- 500 Kb from tag SNP) (33). We also
120 identified five loci for men only, including one reaching genome-wide significance (GWS, P<5x10^{-8}). A
121 total of eight suggestive loci were identified in the women-only analyses, including one, rs115981023 in
122 *TAOK3*, which also reached suggestive significance in the combined sexes analysis, and rs79478137 in
123 *SLC22A18AS* near a previously implicated WHRadjBMI locus(34). For WCadjBMI, we identified nine loci,
124 including one GWS locus in the combined sexes; 11 for men only, including two SNPs that reach GWS;
125 and two for women only. Of the WCadjBMI loci identified, two were nearby previously-reported
126 WCadjBMI loci, rs6809759 near *PROK2* (men-only) (15, 17, 35) and rs77319470 near *ADAMTS3* (sexes-
127 combined) (15, 17, 36). For HIPadjBMI, we identified eight loci that reach p<1x10^{-6} for the combined

sexes; nine for men only, including one in a locus that reached suggestive significance for the combined sexes as well (near *ANO10*); and 12 for women only, including one SNP in a locus that reached suggestive significance for the combined sexes as well (near *LPPR4*). Of the WCadjBMI loci, rs10818474 near *MEGF9* was within 500 Kb of a recently reported WHRadjBMI association in women (35).

Association Differences by Genetic Ancestry

All of the top loci were directionally consistent in each background group, yet many of the loci exhibited effect heterogeneity by background group (**Table 2**, **Figures 1-3**, **Supplementary Tables 5-7**), as exhibited by moderate to high I^2 values (ISQ>65%) and/or significant interaction across subgroups (P_{diff}<0.05). For example, rs113818604 (β=0.0269, *P*= 5.47 x 10^{-8}), I^2=78.5%, P_{diff}=0.38) in *NTM* is significantly associated with WHRadjBMI in women from the mainland background groups (N=4220, MAF=0.014, β=0.0343, *P*=1.63x10^{-8}), but not in women from Caribbean background groups (N=3238, MAF=0.013, β=0.0144, *P*=0.08) **(Supplementary Table 5**). Also, for the women-only primary analysis, the rs77186623 in *LOC105375745* locus associated with HIPadjBMI (β=-0.006, P=1.74x10^{-7}) exhibited nominally significant interaction by background group (I^2=55.3%, P_{diff}= 0.042) , and was GWS in the Caribbean group (N=3231, MAF=0.041, β=-0.0078, *P*=3.05x10^{-8}), but not significant in the mainland group (N=4216, MAF=0.008, β=-0.0015, *P*=0.567, **Supplementary Table 7**). Additional examples that cannot be explained due to power (i.e. MAF and sample size are similar) for WHRadjBMI include rs77377042 near *MARCKSL1* and rs61305557 in *C19orf67* for women, and rs16977373 near *RIT2* for men; for WCadjBMI in women-only include rs76842062 in *MAP4K4* and rs76941364 near *COBL;* and for HIPadjBMI rs6860625 near *NREP* for women and rs145815581 in *ANO10* for the combined sexes.

For other loci, allele frequency and linkage disequilibrium (LD) differences across Hispanic/Latino populations likely contributed to observed differences in magnitude of effect and significance levels (**Supplementary Table 8**). For example, while the magnitude of effect for the

rs115981023 *TAOK3* association with WHRadjBMI in women (β=-0.029, P=8.88x10^{-7}, I^2=0, P_{diff}=0.391) was similar across background groups, the P-value was far more significant in the Caribbean background group (MAF=0.016, β=-0.030, P=2.72x10^{-5}) compared to the mainland (MAF=0.003, β=-0.027, P=0.025), likely due to the higher MAF in the Caribbean group. Of note, the minor allele at this SNP is more common in the 1000 Genomes AFR compared to the EUR and AMR reference samples (**Supplementary Table 5**) and the local ancestry of participants at this locus indicate that those with African ancestry exhibit the highest MAF (**Supplementary Table 8**). Additional loci where the significance level differences between Caribbean and mainland background groups appear to be driven by increased MAF due to African ancestry in Caribbean populations include the *SLC22A18AS* and *CDH4* loci for WHRadjBMI; *LOC102723448*, *FZD7*, *WSB2*, and *ACTRT2* loci for WCadjBMI; and *COQ2*, *LPPR4*, *TMEM63A*, and *FHIT* loci for HIPadjBMI (**Supplementary Tables 5-8**). Rs12478843 in *HEATR5B* (β=-0.002, P=8.2x10^{-8}, I^2=1.7%, P_{diff}=0.385) is more significantly associated with HIPadjBMI in mainland (MAF=0.320, β=-0.002, *P*=6.50x10^{-6}) women compared to Caribbean (MAF=0.154, β=-0.002, *P*=6.03x10^{-3}), likely reflecting the higher MAF among those from mainland Latin America with greater Native American ancestry (**Supplementary Table 8**). Similarly, differences in effect magnitude between mainland and Caribbean background groups for the *TAF4* (HIPadjBMI in women) and the *ESRRG* (WCadjBMI in men) loci may also be due to higher MAF in the mainland group because of a greater proportion of Native American ancestry (**Supplementary Tables 6-8**).

Replication

We brought 66 variants forward for replication in nine cohort studies (N up to 34,161) with participants of Hispanic/Latino, African and European descent, and for further examination of replication by ancestral background (**Table 1**, **Supplementary Tables 1-4, Supplementary Figures 10-12**). Our criteria for replication included both nominal evidence of an association (*P*<0.05), consistent

direction of effect between the replication results and the HCHS/SOL discovery results for any ancestry/sex stratum, and genome-wide significance ($P<5x10^{-8}$) when meta-analyzed together with HCHS/SOL. Based on these criteria, we were able to replicate four novel loci (**Table 1**) after combining our HCHS/SOL discovery sample with specific ancestry results. For WHRadjBMI in men and women combined, rs13301996 was significant after meta-analyzing HCHS/SOL with the African American replication sample ($P=2.88x10^{-8}$). For WHRadjBMI in women only, rs79478137 was GWS after combining HCHS/SOL with the Hispanic/Latino replication sample ($P=3.64x10^{-9}$). For HIPadjBMI in women only, rs28692724 was significant after meta-analyzing HCHS/SOL with the European American replication sample ($P=4.02x10^{-8}$). For WCadjBMI in men and women combined, rs3168072 was significant after combining HCHS/SOL with the European American replication sample ($P=4.21x10^{-8}$).

Of note, for rs13301996, which only replicated in African Americans, we see a larger effect size in the Caribbean background group compared to the mainland, although this is not a significant difference (**Table 2, Supplementary Table 5 Supplementary Figure 10**). This finding may provide insight into why the variants were more successful upon replication with a particular ancestry. For the remaining loci there is little difference in effect magnitude between the Caribbean and the mainland background groups that could explain differences in replication by ancestral group.

Generalization of Previous Loci

We examined previously reported association regions from the GIANT Consortium (35) to assess generalization to the HCHS/SOL (**Supplementary Tables 9-11, Supplementary Figures 13-21**). To account for differences in LD between GIANT (primarily European descent populations) and HCHS/SOL (highly admixed Hispanic/Latino populations), we report generalization results based on the lead generalized SNP (the SNP with lowest r-value in the region of the previously reported variant in GIANT). In sex combined analyses, there were a total of 12 association regions across the genome that generalized to HCHS/SOL for WHRadjBMI ($r<0.05$), including three for both women-only and sexes combined, three for

women-only, and six for the sexes-combined analysis (**Supplementary Table 9**). A total of 15 association regions generalized to HCHS/SOL for WCadjBMI, including seven sex-specific loci (two for men, five for women, **Supplementary Table 10**), one for the sexes-combined only, and seven for more than one stratum. Of note, we identified rs6809759 near *PROK2,* which was significantly associated with WCadjBMI in HCHS/SOL for men-only and sexes-combined and was within 500 kb (+/-) of rs12330322, identified in Shungin et al (35). However, this previously-identified locus did not generalize to HCHS/SOL ($r>0.05$); and may represent an independent association signal in a known region (i.e. all GIANT variants at this locus with $P<1x10^{-6}$ exhibit $r>0.05$ in HCHS/SOL and rs6809759 had a $P>1x10^{-6}$ in Shungin et al. (35) [$P=1.4x10^{-1}$]). A total of 40 regions generalized to HCHS/SOL for HIPadjBMI, including 29 for sexes-combined, three of which were significant for both women-only and sexes combined analyses (**Supplementary Table 11**).

Because some of the SNPs previously reported by GIANT may not have generalized due to lack of power in HCHS/SOL, we calculated individual-level genetic scores based on trait-increasing alleles for each central adiposity phenotype (**Supplementary Table 12**) and sex stratum (three association tests per phenotype). For genetic scores based on SNPs with p-value$<1x10^{-7}$ in GIANT, all association tests were significant (p-value<0.05). For genetic scores calculated from GIANT SNPs with $1x10^{-7}<P<1x10^{-6}$, six of the nine association tests were significant. Given that only three out of 27 analyses had a p-value>0.05, there is considerable overlap in the association results of Hispanics/Latinos to those previously reported in the GIANT multiethnic analysis.

Biological Curation

We examined the four SNPs (i.e. rs13301996, rs79478137, rs28692724, and rs3168072) in novel loci identified in the replication analyses (**Table 1**) for association with other phenotypes, gene expression, and metabolites in publicly available data using Phenoscanner (37, 38), and assessed the

potential regulatory role of these variants and those in LD using publicly available databases, including RegulomeDB (39), Haploreg (40), UCSC GenomeBrowser (41), and GTeX (42). Known associations with these variants meeting Bonferroni-corrected significance after correcting for number of reported associations in Phenoscanner for the four variants within each category ($P<0.05/7631=6.55x10^{-5}$ for GWAS; $P<0.05/88=5.68x10^{-4}$ for gene expression; $P<0.05/488=P<1.02x10^{-4}$ for metabolites) are provided in **Supplementary Tables 13-15**.

WHRadjBMI-associated variant, rs13301996, which is intronic to *CDK5RAP2* (CDK5 [cyclin-dependent kinase 5] Regulatory Subunit Associated Protein 2), was significantly associated with expression of 15 genes and one lncRNA across 17 tissue types (**Supplementary Table 14**). The most significant of these associations was with *MEGF9* (Multiple Epidermal Growth Factor-Like Domains 9) in whole blood ($P=1.8x10^{-149}$), a gene that rests 30 Kb upstream of rs1330996. This SNP is also significantly associated with expression of *MEGF9* in subcutaneous adipose tissue, sun-exposed skin, and T-cells. Additionally, our lead variant in *CDK5RAP2* is associated with expression of *MEGF9* in whole blood and the testis; and with expression of *PSMD5* (proteasome [prosome, macropain] 26S subunit, non-ATPase, 5) and/or *PSMD5-AS1* in several relevant tissues, including whole blood, tibial artery, tibial nerve, lung, thyroid, esophagus muscle, skeletal muscle, liver, cerebellum, and subcutaneous adipose tissues, among others. Although rs13301996 is associated with gene expression for several genes, there is additional support for a regulatory role of this SNP and those with which it is in high LD ($r^2>0.8$). For example, our lead SNP lies just outside of a DNase hypersentivitiy cluster, lies within a region with evidence of histone modification in nine tissues including brain, skin, muscle, and heart; and likely falls in a transcription factor binding site active in skeletal and lung tissue; etc. (**Supplementary Table 16**) (39-41) . While there are multiple lines of evidence for a regulatory role of this variant and multiple genes, rs13301996 has RegulomeDB Score of 6, indicating little evidence of binding.

WHRadjBMI-associated SNP, rs79478137, is a low frequency variant (MAF=1.6%) intronic in *SLC22A18AS* (Solute Carrier Family 22 (Organic Cation Transporter), Member 18 Antisense). This region is subject to genomic imprinting (43) which has been linked with Beckwith-Wiedemann syndrome, a disease caused by increased rate of growth in children (44-46). Our lead variant is associated with two Electronic Health Record (EHR)-derived phenotypes (cause of death: multisystem degeneration; and cause of death: tongue, unspecified) (**Supplementary Table 13**) in Phenoscanner. There is limited evidence of a regulatory role for our lead SNP (RegulomeDB Score = 4), but rs79478137 is in perfect LD with several variants with evidence of regulation (histone modification, open chromatin, DNAse hypersensitivity, transcription factor binding) in more than 50 tissues, including blood, pancreas, liver, and skeletal muscle, and hippocampal tissues, etc. (**Supplementary Table 16)** (39-41).

Our lead SNP associated with HIPadjBMI in women, rs28692724 (NC_000014.9:g.77027445C>T), is a synonymous variant exonic to *IRF2BPL* (interferon regulatory factor 2 binding protein like) that is significantly associated with expression of the same gene in whole blood (**Supplementary Table 14**). Additionally, this variant lies in a known CTCF binding site (RegulomDB Score = 2b), among other transcription factors, and a DNAse Hypersentivity cluster (**Supplementary Table 16)** (39-41).

WCadjBMI-associated SNP, rs3168072, was significantly associated with existing GWAS traits present in Phenoscanner, including “cause of death: other specified respiratory disorders” (**Supplementary Table 13**). Additionally, rs3168072 is significantly associated with expression of several genes in whole blood, but most significantly associated with expression of *TMEM258* (Transmembrane Protein 258) (**Supplementary Table 14**). Rs3168072 is ~95 Kb upstream of *TMEM258*. Our lead variant is likely to play a role in gene expression regulation (RegulomeDB score= 2b, “likely to affect binding”) (39). Additionally, our lead variant and those in high LD (R^2>0.8) lie within known DNase hypersentivitiy regions and within active areas of histone modification, open chromatin, and likely gene enhancer regions (**Supplementary Table 16)** (39-41). Our lead SNP associated with WCadjBMI, rs3168072, is

significantly associated with five lipid-related metabolites (**Supplementary Table 15**), including "Other polyunsaturated fatty acids than 18:2," "CH2 groups in fatty acids," "Ratio of bis allylic bonds to double bonds in lipids," "CH2 groups to double bonds ratio," and "Ratio of bis allylic bonds to total fatty acids in lipids."

Discussion

We performed the first large-scale GWAS of 3 central adiposity traits (i.e., WHRadjBMI, WCadjBMI, and HIPadjBMI) in a sample of approximately 12,672 Hispanic/Latino individuals. We identified 16 variants that were suggestively associated ($P<1x10^{-6}$) with WHRadjBMI, 22 for WCadjBMI, and 28 for HIPadjBMI. Of these 66 variants that were suggestively associated with the three central adiposity traits, four novel loci replicated after meta-analysis with replication samples. Additionally, we demonstrated that eight previously identified GWAS loci generalized to Hispanic/Latino study participants for WHRadjBMI, 10 for WCadjBMI, and 12 for HIPadjBMI in HCHS/SOL.

Discovery of Four Novel Loci

Given the large number of published GWAS on central adiposity measures, it may seem surprising that four novel loci (rs13301996, rs79478137, rs28692724, and rs3168072) were mapped. We ascribe this to (1) previous GWAS were primarily conducted in European populations. Indeed, all four novel SNPs were absent from previous GIANT HapMap imputed analyses (35), and one (rs28692724) of the four absent from a more recent GWAS that included Europeans from the UK Biobank (33); (2) the consideration of a broad spectrum of ancestrally diverse Hispanic/Latino populations, including not just those of Mexican ancestry, but also those with ancestry from the Caribbean, Central, and/or South America (47); (3) the use of the entire 1,000 Genomes Phase I Reference panel, including populations with Native American ancestry: MXL (Mexico), CLM (Colombia), and PUR (Puerto Rico); (4)

295 demonstrated differences in the patterning of body composition by ancestry (48, 49). More specifically,
296 African ancestry populations have lower body fat percentages than men and women of non-Hispanic
297 European, Native American, and East Asian ancestry at the same BMI. Additionally, non-Hispanic African
298 ancestry men and women have greater skeletal and muscle mass than their non-Hispanic European
299 ancestry counterparts, who in turn have greater skeletal and muscle mass than men and women of East
300 Asian origin (48, 50-52)

301 Recent GWAS for coding variation of waist circumference traits identified the importance of
302 central adiposity genes in lipid regulation, storage, and homeostasis (53). Similarly, we found a novel
303 association of a variant in *FADS2* (rs3168072) with WCadjBMI following meta-analysis of HCHS/SOL
304 results with results from an independent sample of European descent individuals, which further implies
305 a role of this locus in central adiposity and lipid homeostasis. Genetic variations in the *FADS2* gene has
306 been associated with several traits related to obesity and cardiometabolic health, including fatty acid
307 metabolism and adipose tissue inflammation, leading to an interaction between weight loss and *FADS2*
308 genes in the regulation of adipose tissue inflammation (54). A nearby variant, rs174546 (R^2=0.3523,
309 D'=0.916 in AMR), in *FADS1* has previously been associated with four lipid traits (55). The A allele
310 (MAF=38%) is associated with greater waist circumference in our samples, and is nearly fixed among
311 sub-Saharan African populations (99% in 1000 Genomes AFR), at very high frequency in European
312 populations (97% in EUR), and at a lower frequency in East Asian (75% in EAS) and Native American
313 populations (63% in AMR). Rs3168072 is intronic to *FADS2* - a member of the fatty acid desaturase
314 (FADS) gene family and is involved in the endogenous conversion of short-chain polyunsaturated fatty
315 acids to long chain fatty acids. The *FADS* cluster of genes appears to have been under strong selection in
316 several human populations, which likely explains the large differences in allele frequencies across global
317 populations (56-59), and why previous GWAS of waist traits primarily focused on European descent
318 populations did not detect an association signal in this region.

We identified a novel association for WHRadjBMI with rs13301996 following meta-analysis with an independent sample of African descent individuals. Rs13301996 is intronic to *CDK5RAP2*, which encodes a regulator of CDK5 (cyclin-dependent kinase 5) activity (60), interacts with CDK5R1 and pericentrin (PCNT) (60), plays a role in centriole engagement and microtubule nucleation (61), and has been linked to primary microcephaly and Alzheimer's disease (62, 63). In addition, we identified a novel association for WHRadjBMI with rs79478137 (p-value= $3.64E^{-9}$) in Hispanic/Latino women. Rs79478137 is intronic to the antisense *SLC22A18AS* gene, which is highly expressed in the liver and kidney, as well as the gastrointestinal tract and placenta. Very little is known of the biological role of this gene (64), and *SLC22A18AS* has no counterpart in mice or other rodents (65). Thus, although its genomic organization is known, the regulation and function of this gene is not understood (66).

Lastly, we identified a novel association for HIPadjBMI at rs28692724 following meta-analysis with an independent sample of European women. Rs28692724 is a synonymous variant in *IRF2BPL*, which encodes a transcription factor that, acting within the neuroendocrine system, plays a role in regulating female reproductive function (67).

Differences in Association by Background Group

Many of the loci mapped in this study displayed effect heterogeneity by background group. For example, the *NTM* locus associated with WHRadjBMI in women, displayed nearly threefold the effect size in the mainland background group compared to the Caribbean background group. Also, for the women-only primary analysis, rs77186623 in the *LOC105375745* locus displayed a fourfold greater effect in the Caribbean background group compared to the mainland group. These and other loci displaying heterogeneity by background group (i.e. *MARCKSL1*, *C19orf67*, *RIT2*, *MAP4K4*, *COBL*, *NREP*, and *ANO10*) were not validated in replication analyses, possibly due in part to heterogeneity by background group.

Limitations

A limitation of this study was the small sample size within each HCHS/SOL background group. However, the use of genetic-analysis groups in our analyses accounted for heterogeneity of genetic effects among ethnic groups often ignored in GWAS studies. Compared to self-identified background groups, genetic-analysis groups are more genetically homogeneous and lack principal component outliers in stratified analysis, which may hinder detection of and adjustment for important population structure when ignored (68). In addition, genetic-analysis groups allow all individuals to be classified in a specific group, whereas many individuals in HCHS/SOL have a missing or non-specific self-identified background (68). Therefore, by using genetic-analysis groups in our analysis rather than self-identified background groups, we have increased our study's power to detect novel and previously documented associations with central adiposity traits (68). Due to the diverse background of our discovery population, another limitation was the lack of an ideal replication study. We attempted to overcome this limitation by focusing on both multiethnic meta-analyses, which would validate those associations that generalize across ancestries, and meta-analyses stratified by ancestry, which may allow for validation of more population-specific associations. However, it is possible that the limited Native American ancestry present across our replication cohorts may have hindered replication, and further analyses in more diverse Hispanic/Latino populations are needed to confirm the relevance of promising central adiposity associated loci identified in our study.

Conclusion

We identified 4 novel loci for central adiposity traits in a large population of Hispanic/Latino Americans. We also found that several previously identified central adiposity loci discovered in

European American populations generalized to Hispanic / Latino Americans. Many of the loci interrogated exhibit subgroup-specific effects, likely due to population history (admixture, natural selection), that have resulted in changes in LD, or allele frequency differences, or due to variation in etiology. These observations highlight the importance of large-scale genomic studies in ancestrally diverse populations for identifying obesity-susceptibility loci that generalize and those that are ancestry-specific.

Materials and Methods

Study Sample

Details on the study and sampling design of the Hispanic Community Health Study /Study of Latinos (HCHS/SOL) have been previously described (69). Briefly, HCHS/SOL is a community based prospective cohort study of 16,415 self-identified Hispanic/Latino adults aged 18-74 years at screening from randomly selected households in four US field centers (Chicago, IL; Miami, FL; Bronx, NY; San Diego, CA) with baseline examination (2008 to 2011) and yearly telephone follow-up assessment for at least three years. The HCHS/SOL cohort includes participants who self-identified as being Central American (n=1,732), Cuban (n=2,348), Dominican (n=1,473), Mexican (n=6,472), Puerto-Rican (n=2,728), and South American (n=1,072). The goals of the HCHS/SOL are to describe the prevalence of risk and protective factors for chronic conditions (e.g. cardiovascular disease (CVD), diabetes and pulmonary disease), and to quantify all-cause mortality, fatal and non-fatal CVD and pulmonary disease, and pulmonary disease exacerbation over time. The baseline clinical examination (70) included comprehensive biological (e.g., anthropometrics, blood draw, oral glucose tolerance test, ankle brachial pressure index, electrocardiogram), behavioral (e.g. dietary intake assessed with two 24-hour recalls, physical activity assessment by accelerometer and self-report, overnight sleep exam for apneic events,

391 tobacco and alcohol assessed by self-report), and socio-demographic (e.g., socioeconomic status,
392 migration history) assessments. This study was approved by the institutional review boards at each field
393 center, where all subjects gave written informed consent.

394 Participants in HCHS/SOL self-identified their background as Mexican, Central American, South
395 American (mainland), Puerto Rican, Cuban, or Dominican (Caribbean). Some participants chose "more
396 than one," "other," or chose not to self-identify. We addressed the missing or inconsistent data in self-
397 identified background groups by defining "genetic analysis groups," described in Conomos et al (68). To
398 increase power in this analysis, we chose to stratify by the broader mainland or Caribbean categories
399 rather than more specific groups. In this paper, we will use the term "background group" to refer to a
400 super-group of genetic analysis groups by geographic region, mainland or Caribbean. Hispanics/Latinos
401 have admixed ancestry from three continents: Africa, America, and Europe. In general, participants from
402 the mainland group have higher proportions of American ancestry and lower African ancestry, while
403 participants in the Caribbean group have higher proportions of African ancestry (68).
404

405 **Phenotypes**

406 All measurements were taken from the baseline visit. Participants were dressed in scrub suits or
407 light non-constricting clothing and shoes were removed for weight and height measurements. WC and
408 HIP were measured using Gulick II 150 and 250 cm anthropometric tape and rounded to the nearest
409 centimeter (cm). Height was measured using a wall mounted stadiometer and rounded to the nearest
410 cm, and weight measured with a Tanita Body Composition Analyzer, TBF-300Ato the nearest tenth of a
411 kilogram (kg). Height and weight were used to calculate BMI (kg/m^2). We applied a log10 transformation
412 on HIP, due to its non-normal trait distribution.
413

414 **Genotyping**

Our analyses included 7,472 women and 5,200 men of mainland (Mexican, Central and South American) or Caribbean (Puerto Rican, Cuban, and Dominican) ancestry residing in the U.S. All participants were genotyped on the Illumina SOL Omni2.5M custom content array, which was subsequently used to impute millions of additional variants, based on the entire 1,000 Genomes Phase I Reference panel, including populations with Native American ancestry: MXL (Mexico), CLM (Colombia), and PUR (Puerto Rico). Pre-phasing was performed using SHAPEIT, followed by imputation with IMPUTE2 (71, 72).

Discovery Analyses

Due to known differences in genetic effects on waist and hip traits between men and women (35, 73, 74), we analyzed associations stratified by sex for each trait, in addition to the entire sample. We used linear mixed-model regression, assuming an additive genetic model adjusted for age, age^2, study center, sample weights, genetic analysis subgroup (68, 75), principal components to account for ancestry, population structure using kinship coefficients and sample eigenvectors, household, census block group, and sex in the combined analysis. Kinship, household, and block group were treated as random effects in each model. Sample weights were incorporated in our models as a fixed effect to account for oversampling of the communities in the 45-74 age group (n=9,714, 59.2%) which was intended to facilitate the examination of HCHS/SOL target outcomes. HCHS/SOL sampling weights are the product of a "base weight" (reciprocal of the probability of selection) and three adjustments: 1) non-response adjustments made relative to the sampling frame, 2) trimming to handle extreme values (to avoid a few weights with extreme values being overly influential in the analyses), and 3) calibration of weights to the 2010 U.S. Census according to age, sex, and Hispanic background. We used genetic-analysis groups in our analyses accounted for heterogeneity of genetic effects among ethnic groups. Compared to self-identified background groups, genetic-analysis groups are more genetically

439 homogeneous and lack principal component outliers in stratified analysis, which may hinder detection of
440 and adjustment for important population structure when ignored (68). In addition, genetic-analysis
441 groups allow all individuals to be classified in a specific group, whereas many individuals in HCHS/SOL
442 have a missing or non-specific self-identified background (68). Also, we conducted stratified analyses by
443 region (mainland vs. Caribbean) to identify potential heterogeneity in effect by background group. We
444 examined heterogeneity across background group using I^2 statistics calculated using METAL (76) and
445 tested for significant interaction (P_{diff}<0.05) by background group using EasyStrata (77).
446 To decrease the number of spurious associations, we filtered all results on minor allele
447 frequency (MAF) < 0.5%, Hardy-Weinberg Equilibrium (HWE) $P < 1x10^{-7}$, minor allele count (MAC
448 [effective N]) < 30 (68). Additionally, we categorized suggestive loci as those with variants reaching
449 $P<1x10^{-6}$ and with at least one additional variant within 500 kb+/- with a $P<1x10^{-5}$. We used regional
450 association plots produced in LocusZoom to visualize association regions using 1000 Genomes Admixed
451 American (AMR) reference population for LD (http://locuszoom.sph.umich.edu/).
452

453 **Local Ancestry Estimation**

454 We estimated local ancestry (African, Native American, and European) using RFMix (78), which
455 applies a conditional-random-field-based approach for estimation, to inform differences by background
456 group. We used 236,456 genotyped SNPs available in both HCHS/SOL and reference-panel datasets in
457 the Human Genome Diversity Project (HGDP) (79), HapMap 3 (80), and 1000 Genomes phase 1 for
458 detecting African, Native American, and European ancestry. We used BEAGLE (v.4) to phase and impute
459 sporadic missing genotypes in the HCHS/SOL and reference-panel datasets (81).
460

461 **Replication and Meta-Analyses**

An aim of our study was to identify genetic variants that contribute to central adiposity which may vary by ancestry; therefore, we sought to replicate our association findings using 1000 Genomes imputed GWAS data available in independent cohorts including eight studies with Hispanics/Latinos (HL: N up to 12,341), three studies with African-Americans (AA; N up to 12,496), and one study with European-Americans (EUR: N up to 8,845). Study design and descriptive statistics for each replication study are provided in **Supplementary Table 1**. Each replication study excluded individuals that were pregnant or exhibited extreme values for waist or hip measures (outside of ±4 SD from the mean). Each study used measures from a single visit with the greatest sample size. We used linear regression (or linear mixed effects models if the study had related individuals) association analyses on the trait residuals after adjustment for age, age^2, principal components to account for ancestry, BMI, other study specific factors (e.g. study center), and sex in the sex combined analysis, stratified by race/ethnicity where applicable for each SNP that reached suggestive significance (P<1x10^{-6}) in the discovery analysis.

We employed a fixed-effects meta-analysis using the inverse variance-weighted method for WHRadjBMI and WCadjBMI. For HIPadjBMI, due to trait transformations, we used sample size weighted meta-analysis. All meta-analyses were implemented in METAL (82). We conducted meta-analyses stratified by race/ethnicity group and combined across groups. We included SNPs with a study and stratum specific imputation quality (Rsq) greater than 0.4, Hardy-Weinberg Equilibrium P-value greater than 1x10^{-7}, and a minor allele count (MAC) greater than five. To declare statistical significance for replicated loci, we required in each replication sample a trait and stratum-specific P<0.05 with a consistent direction of effect with discovery, and genome-wide significance (P<5x10^{-8}) when meta-analyzed together with HCHS/SOL.

Generalization

To examine whether previously reported association regions generalized to the HCHS/SOL, we downloaded the publicly available multi-ethnic (European, Asian, and African ancestry) GWAS results from the Genetic Investigation of Anthropometric Traits (GIANT) consortium (14) for HIPadjBMI, WHRadjBMI, and WCadjBMI (https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files#GIANT_consortium_2012-2015_GWAS_Metadata_is_Available_Here_for_Download), in men, women, and sexes combined, and applied the framework of Sofer et al. (2016) for generalization testing (83). We took all variant associations with $P<1x10^{-6}$ in GIANT and identified the matching association test in HCHS/SOL. For each such association, we calculated a directional FDR r-value, by combining the *P*-values from GIANT and HCHS/SOL, while accounting for multiple testing and for the direction of estimated associations in each of the studies. Then, an association was declared as generalized, while controlling for the False Discovery Rate (FDR) at the 0.05 level, if its r-value was smaller than 0.05. Note that multiple SNPs from the same region were tested. Therefore, in an iterative procedure, we pruned the results list by identifying the SNP with the lowest r-value in an analysis, then finding all SNPs in a 1MB region around it and removing them from the list. Thus, the number of generalized regions is the number of generalized SNPs in the pruned list.

We also hypothesized that some regions did not generalize due to lack of power (the HCHS/SOL sample size is much smaller than the GIANT sample size). To test this, we took all tested SNPs from the non-generalized regions and considered the GIANT multi-ethnic GWAS results. In an iterative procedure, we pruned the list by first identifying the SNP with lowest GIANT *P*-value in the analysis, then found all SNPs in a 1MB region around it and removed them from the list. We repeated until no SNPs remained. All the SNPs in the pruned list were selected solely based on their GIANT *P*-values. Since there were many such variants, we further grouped them according to their *P*-values. Groups were formed by trait, sex (men, women, combined), and GIANT P-value (between 10^{-6} to 10^{-7}, between 10^{-7} to 10^{-8}, and

509 smaller than 10^{-8}). For each such group of SNPs, we created a genetic risk score (GRS) in HCHS/SOL. For
510 each sex stratum and each group of SNPs, the value of the GRS was the sum of all trait increasing alleles
511 in that group. We tested the GRS in the appropriate analysis group (men, women, combined). A low *P*-
512 value implies that some of the SNPs in the group are likely associated with the trait in HCHS/SOL.
513
514 **Biological Curation**
515 To gain further insight into the possible functional role of the identified variants and to assess
516 the relevance of our identified variants with other phenotypes, we conducted lookups of our replicated
517 variants in multiple publicly available databases, including PhenoScanner (37), RegulomeDB (39),
518 Haploreg (40), and UCSC GenomeBrowser (41). Additionally, we conducted lookups of nearby genes in
519 GTeX (42). The R package HaploR was used to query HaploReg and RegulomeDB (https://cran.r-
520 project.org/web/packages/haploR/vignettes/haplor-vignette.html).
521

522 # Acknowledgements

523 AEJ was supported in part by National Institute of Health (NIH), National Heart, Lung, and Blood
524 Institutes (NHLBI) grant K99/R00 HL 130580 and American Heart Association (AHA) grant
525 13POST16500011. YCK is supported by the National Heart, Lung, and Blood Institutes (R01-HL136528).
526 HMH was supported in part by NIH NHLBI grant T32 HL007055.
527 **HCHS/SOL:** The authors thank the staff and participants of HCHS/SOL for their important contributions.
528 (Investigators website - http://www.cscc.unc.edu/hchs/). The Hispanic Community Health Study/Study
529 of Latinos is a collaborative study supported by contracts from the National Heart, Lung, and Blood
530 Institute (NHLBI) to the University of North Carolina (HHSN268201300001I / N01-HC-65233), Universit y
531 o f Miami (HHSN268201300004I / N01-HC-65234), Albert Einstein College of Medicine
532 (HHSN268201300002I / N01-HC-65235), University of Illinois at Chicago – HHSN268201300003I / N01-

HC-65236 Northwestern Univ), and San Diego State University (HHSN268201300005I / N01-HC-65237). The following Institutes/Centers/Offices have contributed to the HCHS/SOL through a transfer of funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, NIH Institution-Office of Dietary Supplements. The Genetic Analysis Center at the University of Washington was supported by NHLBI and NIDCR contracts (HHSN268201300005C AM03 and MOD03).

ARIC: The Atherosclerosis Risk in Communities Study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I and HHSN268201700005I), R01HL087641, R01HL059367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. The authors thank the staff and participants of the ARIC study for their important contributions.

GUARDIAN, MACAD, HTN-IR, GOLDR: This research was supported by the GUARDIAN (Genetics Underlying Diabetes in Hispanics) Study, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) contract DK085175; (Mexican-American Coronary Artery Disease (MACAD) National Heart, Lung, and Blood Institute, contracts R01-HL088457, R01-HL-60030; National Heart, Lung, and Blood Institute, Hypertension and Insulin Resistance (HTN-IR) contracts R01-HL067974, R01-HL-55005, R01-HL 067974; National Institutes of Health, Genetics of Latino Diabetic Retinopathy (GOLDR) contract EY014684.

MESA: MESA and the MESA SHARe projects are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts 75N92020D00001, HHSN268201500003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420. Also supported by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

The Mexico City 1 and Mexico City 2 studies were supported in Mexico by the Fondo Sectorial de Investigación en Salud y Seguridad Social (SSA/IMSS/ISSSTECONACYT, project 150352), Temas Prioritarios de Salud Instituto Mexicano del Seguro Social (2014-FIS/IMSS/PROT/PRIO/14/34), and the Fundación IMSS. We thank Miguel Alexander Vazquez Moreno, Daniel Locia and Araceli Méndez Padrón for technical support in Mexico. In Canada, this research was enabled in part by two CIHR Operating grants to EJP, a CIHR New Investigator Award to EJP and by support provided by Compute Ontario (www.computeontario.ca), and Compute Canada (www.compute.canada.ca).

WHI: Funding support for the "Epidemiology of putative genetic variants: The Women's Health Initiative" study is provided through the NHGRI grants HG006292 and HL129132. The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSC271201100004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whiscience.org/publications/ WHI_investigators_shortlist.pdf.

Data Availability Statement

Raw genetic data used in the discovery analysis for this project is available through request on dbGAP (dbGaP Study Accession: phs000810.v1.p1). Additionally, the full GWAS summary results for common variants with a MAF>1% will be made available upon request to the author and on the NHGRI-GWAS Catalog (upon publication).

Conflict of Interest Statement

The authors have no conflicts of interest to disclose.

References

1 Flegal, K.M., Carroll, M.D., Kit, B.K. and Ogden, C.L. (2012) Prevalence of Obesity and Trends in the Distribution of Body Mass Index Among US Adults, 1999-2010. *JAMA: The Journal of the American Medical Association*, **307**, 491-497.

2 Ogden, C.L., Carroll, M.D., Curtin, L.R., McDowell, M.A., Tabak, C.J. and Flegal, K.M. (2006) Prevalence of overweight and obesity in the United States, 1999-2004. *JAMA*, **295**, 1549-1555.

3 Stevens, G.A., Singh, G.M., Lu, Y., Danaei, G., Lin, J.K., Finucane, M.M., Bahalim, A.N., McIntire, R.K., Gutierrez, H.R., Cowan, M. *et al.* (2012) National, regional, and global trends in adult overweight and obesity prevalences. *Population health metrics*, **10**, 22.

4 Finucane, M.M., Stevens, G.A., Cowan, M.J., Danaei, G., Lin, J.K., Paciorek, C.J., Singh, G.M., Gutierrez, H.R., Lu, Y., Bahalim, A.N. *et al.* (2011) National, regional, and global trends in body-mass

603 index since 1980: systematic analysis of health examination surveys and epidemiological studies with
604 960 country-years and 9.1 million participants. *Lancet*, **377**, 557-567.
605 5 Hales, C.M., Carroll, M.D., Fryar, C.D. and Ogden, C.L. (2017) Prevalence of Obesity Among
606 Adults and Youth: United States, 2015-2016. *NCHS Data Brief*, in press., 1-8.
607 6 Elks, C.E., den Hoed, M., Zhao, J.H., Sharp, S.J., Wareham, N.J., Loos, R.J. and Ong, K.K. (2012)
608 Variability in the heritability of body mass index: a systematic review and meta-regression. *Frontiers in*
609 *endocrinology*, **3**, 29.
610 7 Norden-Krichmar, T.M., Gizer, I.R., Libiger, O., Wilhelmsen, K.C., Ehlers, C.L. and Schork, N.J.
611 (2014) Correlation analysis of genetic admixture and social identification with body mass index in a
612 Native American community. *American journal of human biology : the official journal of the Human*
613 *Biology Council*, **26**, 347-360.
614 8 Ravussin, E., Valencia, M.E., Esparza, J., Bennett, P.H. and Schulz, O. (1994) Effects of a
615 traditional lifestyle on obesity in Pima Indians. *Diabetes Care*, **17** 1067-1074.
616 9 Bogardus, C., Lillioja, S., Ravussin, E., Abbott, W.G., Zawadzki, J.K., Young, A., Knowler, W.C.,
617 Jacobowitz, R. and Moll, P.P. (1986) Familial dependence of the resting metabolic rate. *New England*
618 *Journal of Medicine*, **315**, 96-100.
619 10 Knowler, W.C., Pettitt, D.J., Saad, M.F., Charles, M.A., Nelson, R.G., Howard, B.V., Bogardus, C.
620 and Bennett, P.H. (1991) Obesity in the Pima Indians: its magnitude and relationship with diabetes.
621 *American Journal of Clinical Nutrition*, **53**, 1543S-1551S.
622 11 Abate, N. and Chandalia, M. (2007) Ethnicity, type 2 diabetes & migrant Asian Indians. *The*
623 *Indian journal of medical research*, **125**, 251-258.
624 12 Nassir, R., Qi, L., Kosoy, R., Garcia, L., Allison, M., Ochs-Balcom, H.M., Tylavsky, F., Manson, J.E.,
625 Shigeta, R., Robbins, J. *et al.* (2012) Relationship between adiposity and admixture in African-American
626 and Hispanic-American women. *Int J Obes (Lond)*, **36**, 304-313.

627 13 Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S.,
628 Buchkovich, M.L., Yang, J. *et al.* (2015) Genetic studies of body mass index yield new insights for obesity
629 biology. *Nature*, **518**, 197-206.
630 14 Shungin, D., Winkler, T.W., Croteau-Chonka, D.C., Ferreira, T., Locke, A.E., Magi, R., Strawbridge,
631 R.J., Pers, T.H., Fischer, K., Justice, A.E. *et al.* (2015) New genetic loci link adipose and insulin biology to
632 body fat distribution. *Nature*, **518**, 187-196.
633 15 Justice, A.E., Winkler, T.W., Feitosa, M.F., Graff, M., Fisher, V.A., Young, K., Barata, L., Deng, X.,
634 Czajkowski, J., Hadley, D. *et al.* (2017) Genome-wide meta-analysis of 241,258 adults accounting for
635 smoking behaviour identifies novel loci for obesity traits. *Nature communications*, **8**, 14977.
636 16 Ng, M.C.Y., Graff, M., Lu, Y., Justice, A.E., Mudgal, P., Liu, C.T., Young, K., Yanek, L.R., Feitosa,
637 M.F., Wojczynski, M.K. *et al.* (2017) Discovery and fine-mapping of adiposity loci using high density
638 imputation of genome-wide association studies in individuals of African ancestry: African Ancestry
639 Anthropometry Genetics Consortium. *PLoS genetics*, **13**, e1006719.
640 17 Graff, M., Scott, R.A., Justice, A.E., Young, K.L., Feitosa, M.F., Barata, L., Winkler, T.W., Chu, A.Y.,
641 Mahajan, A., Hadley, D. *et al.* (2017) Genome-wide physical activity interactions in adiposity - A meta-
642 analysis of 200,452 adults. *PLoS genetics*, **13**, e1006528.
643 18 Berndt, S.I., Gustafsson, S., Magi, R., Ganna, A., Wheeler, E., Feitosa, M.F., Justice, A.E., Monda,
644 K.L., Croteau-Chonka, D.C., Day, F.R. *et al.* (2013) Genome-wide meta-analysis identifies 11 new loci for
645 anthropometric traits and provides insights into genetic architecture. *Nat Genet*, **45**, 501-512.
646 19 Wen, W., Cho, Y.S., Zheng, W., Dorajoo, R., Kato, N., Qi, L., Chen, C.H., Delahanty, R.J., Okada, Y.,
647 Tabara, Y. *et al.* (2012) Meta-analysis identifies common variants associated with body mass index in
648 east Asians. *Nat Genet*, **44**, 307-311.

649 20 Cho, Y.S., Go, M.J., Kim, Y.J., Heo, J.Y., Oh, J.H., Ban, H.J., Yoon, D., Lee, M.H., Kim, D.J., Park, M.
650 *et al.* (2009) A large-scale genome-wide association study of Asian populations uncovers genetic factors
651 influencing eight quantitative traits. *Nat Genet*, **41**, 527-534.
652 21 Monda, K.L., Chen, G.K., Taylor, K.C., Palmer, C., Edwards, T.L., Lange, L.A., Ng, M.C., Adeyemo,
653 A.A., Allison, M.A., Bielak, L.F. *et al.* (2013) A meta-analysis identifies new loci associated with body mass
654 index in individuals of African ancestry. *Nat Genet*, in press.
655 22 Folsom, A.R., Kushi, L.H., Anderson, K.E., Mink, P.J., Olson, J.E., Hong, C.P., Sellers, T.A., Lazovich,
656 D. and Prineas, R.J. (2000) Associations of general and abdominal obesity with multiple health outcomes
657 in older women: the Iowa Women's Health Study. *Archives of internal medicine*, **160**, 2117-2128.
658 23 Lavie, C.J., Milani, R.V. and Ventura, H.O. (2009) Obesity and cardiovascular disease: risk factor,
659 paradox, and impact of weight loss. *Journal of the American College of Cardiology*, **53**, 1925-1932.
660 24 Huxley, R., Mendis, S., Zheleznyakov, E., Reddy, S. and Chan, J. (2010) Body mass index, waist
661 circumference and waist:hip ratio as predictors of cardiovascular risk--a review of the literature.
662 *European journal of clinical nutrition*, **64**, 16-22.
663 25 Zyriax, B.C., Schoeffauer, M., Klipstein-Grobusch, K., Boeing, H. and Windler, E. (2011)
664 Differential association of anthropometric parameters with coronary risk in women--data of the CORA
665 study. *Obesity facts*, **4**, 358-364.
666 26 Carey, V.J., Walters, E.E., Colditz, G.A., Solomon, C.G., Willett, W.C., Rosner, B.A., Speizer, F.E.
667 and Manson, J.E. (1997) Body fat distribution and risk of non-insulin-dependent diabetes mellitus in
668 women. The Nurses' Health Study. *American journal of epidemiology*, **145**, 614-619.
669 27 Wang, Y., Rimm, E.B., Stampfer, M.J., Willett, W.C. and Hu, F.B. (2005) Comparison of abdominal
670 adiposity and overall obesity in predicting risk of type 2 diabetes among men. *Am J Clin Nutr*, **81**, 555-
671 563.

672 28 Canoy, D. (2008) Distribution of body fat and risk of coronary heart disease in men and women.
673 *Curr Opin Cardiol*, **23**, 591-598.
674 29 Snijder, M.B., Dekker, J.M., Visser, M., Yudkin, J.S., Stehouwer, C.D., Bouter, L.M., Heine, R.J.,
675 Nijpels, G. and Seidell, J.C. (2003) Larger thigh and hip circumferences are associated with better glucose
676 tolerance: the Hoorn study. *Obesity research*, **11**, 104-111.
677 30 Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J.,
678 Jackson, A.U., Vedantam, S., Raychaudhuri, S. *et al.* (2010) Hundreds of variants clustered in genomic
679 loci and biological pathways affect human height. *Nature*, **467**, 832-838.
680 31 Fox, C.S., White, C.C., Lohman, K., Heard-Costa, N., Cohen, P., Zhang, Y., Johnson, A.D., Emilsson,
681 V., Liu, C.T., Chen, Y.D. *et al.* (2012) Genome-wide association of pericardial fat identifies a unique locus
682 for ectopic fat. *PLoS genetics*, **8**, e1002705.
683 32 Randall, J.C., Winkler, T.W., Kutalik, Z., Berndt, S.I. and et_al. (2013) Sex-stratified genome-wide
684 association studies including 270,000 individuals show sexual dimorphism in genetic loci for
685 anthropometric traits. *PLoS genetics*, in press., Epub_2013_Jun_2016.
686 33 Pulit, S.L., Stoneman, C., Morris, A.P., Wood, A.R., Glastonbury, C.A., Tyrrell, J., Yengo, L.,
687 Ferreira, T., Marouli, E., Ji, Y. *et al.* (2019) Meta-analysis of genome-wide association studies for body fat
688 distribution in 694 649 individuals of European ancestry. *Hum Mol Genet*, **28**, 166-174.
689 34 Tachmazidou, I., Suveges, D., Min, J.L., Ritchie, G.R.S., Steinberg, J., Walter, K., Iotchkova, V.,
690 Schwartzentruber, J., Huang, J., Memari, Y. *et al.* (2017) Whole-Genome Sequencing Coupled to
691 Imputation Discovers Genetic Signals for Anthropometric Traits. *Am J Hum Genet*, **100**, 865-884.
692 35 Shungin, D., Winkler, T.W., Croteau-Chonka, D.C., Ferreira, T., Locke, A.E., Magi, R., Strawbridge,
693 R.J., Pers, T.H., Fischer, K., Justice, A.E. *et al.* (2015) New genetic loci link adipose and insulin biology to
694 body fat distribution. *Nature*, **518**, 187-196.

695 36 Wen, W., Kato, N., Hwang, J.Y., Guo, X., Tabara, Y., Li, H., Dorajoo, R., Yang, X., Tsai, F.J., Li, S. *et*
696 *al.* (2016) Genome-wide association studies in East Asians identify new loci for waist-hip ratio and waist
697 circumference. *Sci Rep*, **6**, 17958.
698 37 Staley, J.R., Blackshaw, J., Kamat, M.A., Ellis, S., Surendran, P., Sun, B.B., Paul, D.S., Freitag, D.,
699 Burgess, S., Danesh, J. *et al.* (2016) PhenoScanner: a database of human genotype-phenotype
700 associations. *Bioinformatics (Oxford, England)*, **32**, 3207-3209.
701 38 Kamat, M.A., Blackshaw, J.A., Young, R., Surendran, P., Burgess, S., Danesh, J., Butterworth, A.S.
702 and Staley, J.R. (2019) PhenoScanner V2: an expanded tool for searching human genotype-phenotype
703 associations. *Bioinformatics (Oxford, England)*, **35**, 4851-4853.
704 39 Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J.,
705 Park, J., Hitz, B.C., Weng, S. *et al.* (2012) Annotation of functional variation in personal genomes using
706 RegulomeDB. *Genome Res*, **22**, 1790-1797.
707 40 Ward, L.D. and Kellis, M. (2016) HaploReg v4: systematic mining of putative causal variants, cell
708 types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res*, **44**, D877-
709 881.
710 41 Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D.
711 (2002) The human genome browser at UCSC. *Genome Res*, **12**, 996-1006.
712 42 GTEx Consortium. (2013) The Genotype-Tissue Expression (GTEx) project. *Nat Genet*, **45**, 580-
713 585.
714 43 Gallagher, E., Mc Goldrick, A., Chung, W.Y., Mc Cormack, O., Harrison, M., Kerin, M., Dervan,
715 P.A. and Mc Cann, A. (2006) Gain of imprinting of SLC22A18 sense and antisense transcripts in human
716 breast cancer. *Genomics*, **88**, 12-17.

717 44 Gardiner, K., Chitayat, D., Choufani, S., Shuman, C., Blaser, S., Terespolsky, D., Farrell, S., Reiss,
718 R., Wodak, S., Pu, S. *et al.* (2012) Brain abnormalities in patients with Beckwith-Wiedemann syndrome.
719 *Am J Med Genet A*, **158A**, 1388-1394.
720 45 Weksberg, R., Shuman, C., Caluseriu, O., Smith, A.C., Fei, Y.L., Nishikawa, J., Stockley, T.L., Best,
721 L., Chitayat, D., Olney, A. *et al.* (2002) Discordant KCNQ1OT1 imprinting in sets of monozygotic twins
722 discordant for Beckwith-Wiedemann syndrome. *Hum Mol Genet*, **11**, 1317-1325.
723 46 Brioude, F., Kalish, J.M., Mussa, A., Foster, A.C., Bliek, J., Ferrero, G.B., Boonen, S.E., Cole, T.,
724 Baker, R., Bertoletti, M. *et al.* (2018) Expert consensus document: Clinical and molecular diagnosis,
725 screening and management of Beckwith-Wiedemann syndrome: an international consensus statement.
726 *Nat Rev Endocrinol*, **14**, 229-249.
727 47 Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M.,
728 Marth, G.T. and McVean, G.A. (2012) An integrated map of genetic variation from 1,092 human
729 genomes. *Nature*, **491**, 56-65.
730 48 Rao, G., Powell-Wiley, T.M., Ancheta, I., Hairston, K., Kirley, K., Lear, S.A., North, K.E.,
731 Palaniappan, L. and Rosal, M.C. (2015) Identification of Obesity and Cardiovascular Risk in Ethnically and
732 Racially Diverse Populations: A Scientific Statement From the American Heart Association. *Circulation*,
733 **132**, 457-472.
734 49 Heymsfield, S.B., Peterson, C.M., Thomas, D.M., Heo, M. and Schuna, J.M., Jr. (2016) Why are
735 there race/ethnic differences in adult body mass index-adiposity relationships? A quantitative critical
736 review. *Obesity reviews : an official journal of the International Association for the Study of Obesity*, **17**,
737 262-275.
738 50 Wagner, D.R. and Heyward, V.H. (2000) Measures of body composition in blacks and whites: a
739 comparative review. *Am J Clin Nutr*, **71**, 1392-1402.

740 51 Rush, E.C., Goedecke, J.H., Jennings, C., Micklesfield, L., Dugas, L., Lambert, E.V. and Plank, L.D.
741 (2007) BMI, fat and muscle differences in urban women of five ethnicities from two countries. *Int J Obes*
742 *(Lond)*, **31**, 1232-1239.
743 52 Hull, H.R., Thornton, J., Wang, J., Pierson, R.N., Jr., Kaleem, Z., Pi-Sunyer, X., Heymsfield, S., Albu,
744 J., Fernandez, J.R., Vanitallie, T.B. *et al.* (2011) Fat-free mass index: changes and race/ethnic differences
745 in adulthood. *Int J Obes (Lond)*, **35**, 121-127.
746 53 Justice, A.E., Karaderi, T., Highland, H.M., Young, K.L., Graff, M., Lu, Y., Turcot, V., Auer, P.L.,
747 Fine, R.S., Guo, X. *et al.* (2019) Protein-coding variants implicate novel genes related to lipid homeostasis
748 contributing to body-fat distribution. *Nat Genet*, **51**, 452-469.
749 54 Vaittinen, M., Walle, P., Kuosmanen, E., Mannisto, V., Kakela, P., Agren, J., Schwab, U. and
750 Pihlajamaki, J. (2016) FADS2 genotype regulates delta-6 desaturase activity and inflammation in human
751 adipose tissue. *Journal of lipid research*, **57**, 56-65.
752 55 Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M.,
753 Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J. *et al.* (2010) Biological, clinical and population
754 relevance of 95 loci for blood lipids. *Nature*, **466**, 707-713.
755 56 Fumagalli, M., Moltke, I., Grarup, N., Racimo, F., Bjerregaard, P., Jorgensen, M.E., Korneliussen,
756 T.S., Gerbault, P., Skotte, L., Linneberg, A. *et al.* (2015) Greenlandic Inuit show genetic signatures of diet
757 and climate adaptation. *Science*, **349**, 1343-1347.
758 57 Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E.,
759 Stewardson, K., Fernandes, D., Novak, M. *et al.* (2015) Genome-wide patterns of selection in 230 ancient
760 Eurasians. *Nature*, **528**, 499-503.
761 58 Mathias, R.A., Fu, W., Akey, J.M., Ainsworth, H.C., Torgerson, D.G., Ruczinski, I., Sergeant, S.,
762 Barnes, K.C. and Chilton, F.H. (2012) Adaptive evolution of the FADS gene cluster within Africa. *PLoS*
763 *One*, **7**, e44926.

764 59 Mathias, R.A., Sergeant, S., Ruczinski, I., Torgerson, D.G., Hugenschmidt, C.E., Kubala, M.,
765 Vaidya, D., Suktitipat, B., Ziegler, J.T., Ivester, P. *et al.* (2011) The impact of FADS genetic variants on
766 omega6 polyunsaturated fatty acid metabolism in African Americans. *BMC Genet*, **12**, 50.
767 60 Sukumaran, S.K., Stumpf, M., Salamon, S., Ahmad, I., Bhattacharya, K., Fischer, S., Müller, R.,
768 Altmüller, J., Budde, B., Thiele, H. *et al.* (2017) CDK5RAP2 interaction with components of the Hippo
769 signaling pathway may play a role in primary microcephaly. *Molecular genetics and genomics : MGG*,
770 **292**, 365-383.
771 61 Barrera, J.A., Kao, L.R., Hammer, R.E., Seemann, J., Fuchs, J.L. and Megraw, T.L. (2010)
772 CDK5RAP2 regulates centriole engagement and cohesion in mice. *Developmental cell*, **18**, 913-926.
773 62 Erten-Lyons, D., Wilmot, B., Anur, P., McWeeney, S., Westaway, S.K., Silbert, L., Kramer, P. and
774 Kaye, J. (2011) Microcephaly genes and risk of late-onset Alzheimer disease. *Alzheimer disease and*
775 *associated disorders*, **25**, 276-282.
776 63 Miron, J., Picard, C., Nilsson, N., Frappier, J., Dea, D., Theroux, L. and Poirier, J. (2018) CDK5RAP2
777 gene and tau pathophysiology in late-onset sporadic Alzheimer's disease. *Alzheimer's & dementia : the*
778 *journal of the Alzheimer's Association*, **14**, 787-796.
779 64 Schwienbacher, C., Sabbioni, S., Campi, M., Veronese, A., Bernardi, G., Menegatti, A., Hatada, I.,
780 Mukai, T., Ohashi, H., Barbanti-Brodano, G. *et al.* (1998) Transcriptional map of 170-kb region at
781 chromosome 11p15.5: identification and mutational analysis of the BWR1A gene reveals the presence of
782 mutations in tumor samples. *Proceedings of the National Academy of Sciences of the United States of*
783 *America*, **95**, 3873-3878.
784 65 Bajaj, V., Markandaya, M., Krishna, L. and Kumar, A. (2004) Paternal imprinting of the SLC22A1LS
785 gene located in the human chromosome segment 11p15.5. *BMC Genet*, **5**, 13.
786 66 Bajaj, V., Singhmar, P. and Kumar, A. (2008) Promoter characterization and regulation of
787 expression of an imprinted gene SLC22A18AS. *Gene*, **424**, 40-47.

788 67 Heger, S., Mastronardi, C., Dissen, G.A., Lomniczi, A., Cabrera, R., Roth, C.L., Jung, H., Galimi, F.,
789 Sippell, W. and Ojeda, S.R. (2007) Enhanced at puberty 1 (EAP1) is a new transcriptional regulator of the
790 female neuroendocrine reproductive axis. *The Journal of clinical investigation*, **117**, 2145-2154.
791 68 Conomos, M.P., Laurie, C.A., Stilp, A.M., Gogarten, S.M., McHugh, C.P., Nelson, S.C., Sofer, T.,
792 Fernandez-Rhodes, L., Justice, A.E., Graff, M. *et al.* (2016) Genetic Diversity and Association Studies in US
793 Hispanic/Latino Populations: Applications in the Hispanic Community Health Study/Study of Latinos. *Am*
794 *J Hum Genet*, **98**, 165-184.
795 69 Lavange, L.M., Kalsbeek, W.D., Sorlie, P.D., Aviles-Santa, L.M., Kaplan, R.C., Barnhart, J., Liu, K.,
796 Giachello, A., Lee, D.J., Ryan, J. *et al.* (2010) Sample design and cohort selection in the Hispanic
797 Community Health Study/Study of Latinos. *Annals of epidemiology*, **20**, 642-649.
798 70 Sorlie, P.D., Aviles-Santa, L.M., Wassertheil-Smoller, S., Kaplan, R.C., Daviglus, M.L., Giachello,
799 A.L., Schneiderman, N., Raij, L., Talavera, G., Allison, M. *et al.* (2010) Design and implementation of the
800 Hispanic Community Health Study/Study of Latinos. *Annals of epidemiology*, **20**, 629-641.
801 71 Howie, B., Marchini, J. and Stephens, M. (2011) Genotype imputation with thousands of
802 genomes. *G3 (Bethesda)*, **1**, 457-470.
803 72 Delaneau, O., Marchini, J. and Zagury, J.F. (2011) A linear complexity phasing method for
804 thousands of genomes. *Nature methods*, **9**, 179-181.
805 73 Winkler, T.W., Justice, A.E., Graff, M., Barata, L., Feitosa, M.F., Chu, S., Czajkowski, J., Esko, T.,
806 Fall, T., Kilpelainen, T.O. *et al.* (2015) The Influence of Age and Sex on Genetic Associations with Adult
807 Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLoS genetics*, **11**, e1005378.
808 74 Randall, J.C., Winkler, T.W., Kutalik, Z., Berndt, S.I., Jackson, A.U., Monda, K.L., Kilpelainen, T.O.,
809 Esko, T., Magi, R., Li, S. *et al.* (2013) Sex-stratified Genome-wide Association Studies Including 270,000
810 Individuals Show Sexual Dimorphism in Genetic Loci for Anthropometric Traits. *PLoS genetics*, **9**,
811 e1003500.

812 75 Conomos, M.P., Miller, M.B. and Thornton, T.A. (2015) Robust inference of population structure
813 for ancestry prediction and correction of stratification in the presence of relatedness. *Genet Epidemiol*,
814 **39**, 276-293.
815 76 Willer, C.J., Li, Y. and Abecasis, G.R. (2010) METAL: fast and efficient meta-analysis of
816 genomewide association scans. *Bioinformatics (Oxford, England)*, **26**, 2190-2191.
817 77 Winkler, T.W., Kutalik, Z., Gorski, M., Lottaz, C., Kronenberg, F. and Heid, I.M. (2015) EasyStrata:
818 evaluation and visualization of stratified genome-wide association meta-analysis data. *Bioinformatics*
819 *(Oxford, England)*, **31**, 259-261.
820 78 Maples, B.K., Gravel, S., Kenny, E.E. and Bustamante, C.D. (2013) RFMix: a discriminative
821 modeling approach for rapid and robust local-ancestry inference. *Am J Hum Genet*, **93**, 278-288.
822 79 Cavalli-Sforza, L.L. (2005) The Human Genome Diversity Project: past, present and future. *Nat*
823 *Rev Genet*, **6**, 333-340.
824 80 International HapMap Consortium, Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M.,
825 Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F. *et al.* (2010) Integrating common and rare
826 genetic variation in diverse human populations. *Nature*, **467**, 52-58.
827 81 Browning, S.R. and Browning, B.L. (2007) Rapid and accurate haplotype phasing and missing-
828 data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum*
829 *Genet*, **81**, 1084-1097.
830 82 Heid, I.M., Jackson, A.U., Randall, J.C., Winkler, T.W., Qi, L., Steinthorsdottir, V., Thorleifsson, G.,
831 Zillikens, M.C., Speliotes, E.K., Magi, R. *et al.* (2010) Meta-analysis identifies 13 new loci associated with
832 waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet*, **42**, 949-
833 960.

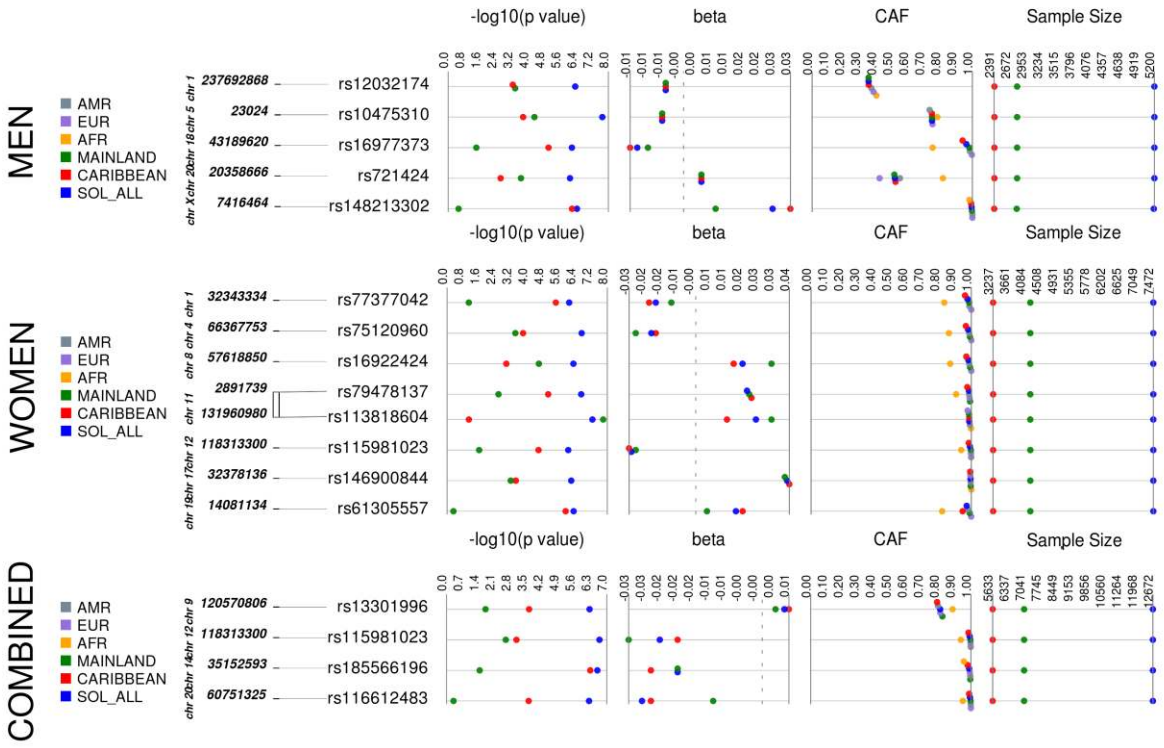
834 83 Sofer, T., Heller, R., Bogomolov, M., Avery, C.L., Graff, M., North, K.E., Reiner, A.P., Thornton,
835 T.A., Rice, K., Benjamini, Y. *et al.* (2017) A powerful statistical framework for generalization testing in
836 GWAS, with application to the HCHS/SOL. *Genet Epidemiol*, **41**, 251-258.

837

838 Figures

839 **Figure 1.** WCadjBMI Synthesis View plot that shows -log10 *P*-values, beta (effect estimate), effect/coded
840 allele frequency (CAF), and sample size across analysis samples for all loci that reached suggestive
841 significance in one or more of our discovery strata. This chart also shows the coded allele frequency
842 (CAF) of each of our top loci by background group and by 1000 genomes reference panel (European -
843 EUR, Latin American - AMR, and African - YRI).

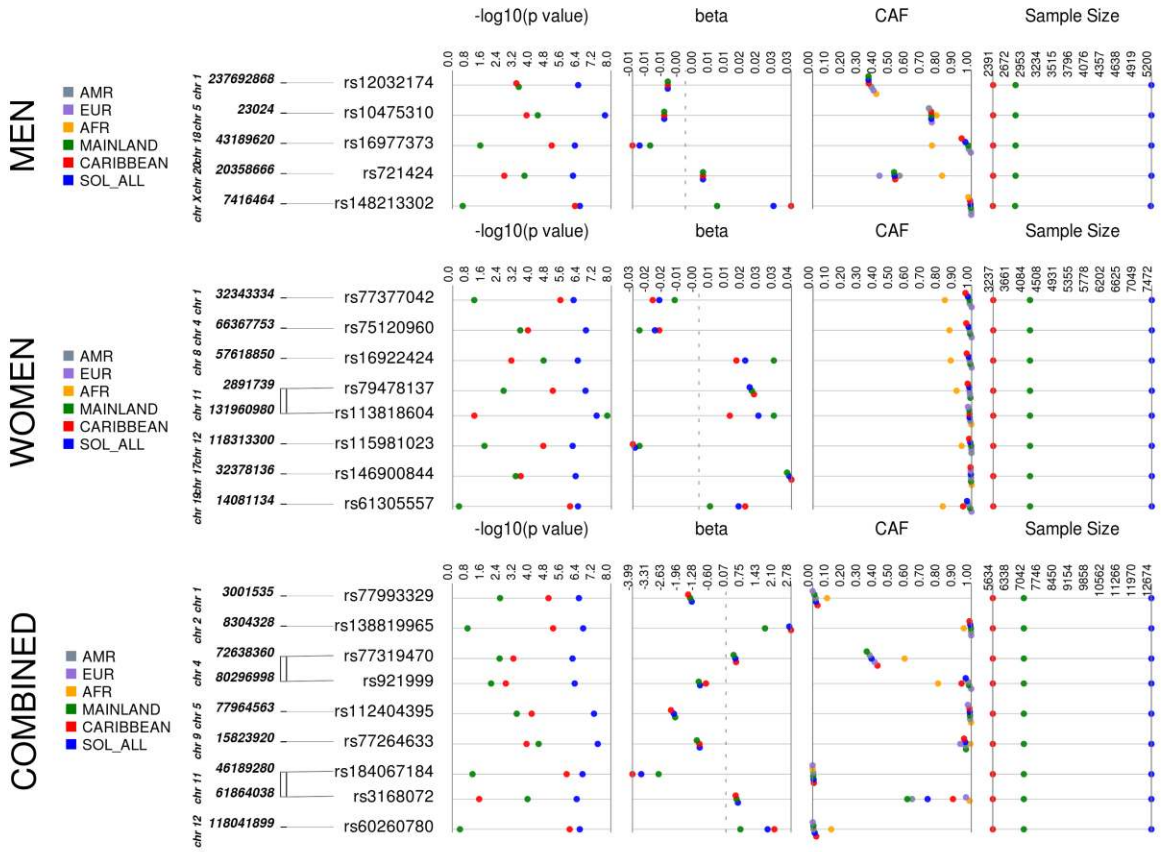
844



845

846

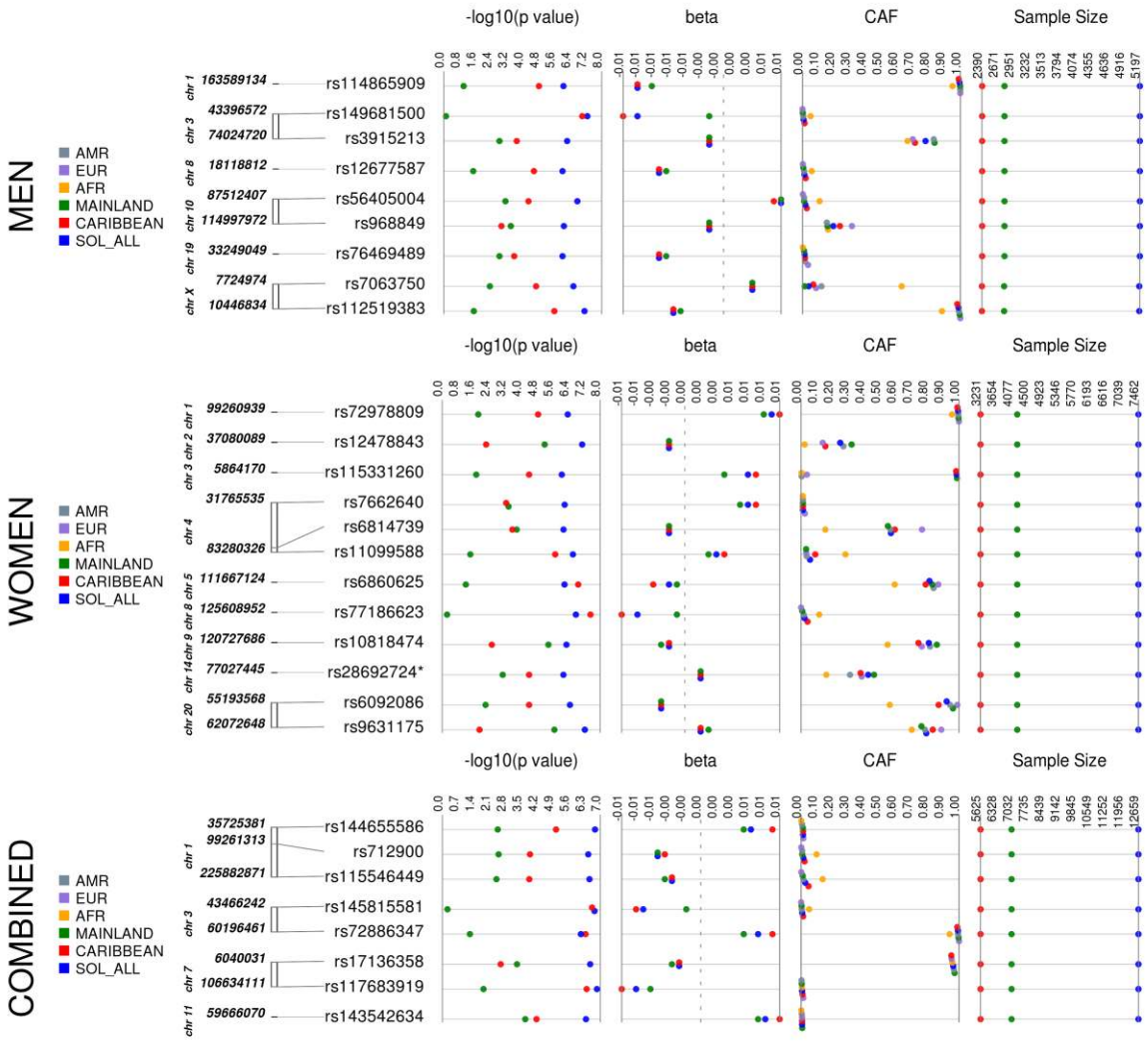
847 **Figure 2.** WHRadjBMI Synthesis View plot that shows -log10 *P*-values, beta (effect estimate),
848 effect/coded allele frequency (CAF), and sample size across analysis samples for all loci that reached
849 suggestive significance in one or more of our discovery strata. This chart also shows the coded allele
850 frequency (CAF) of each of our top loci by background group and by 1000 genomes reference panel
851 (European - EUR, Latin American - AMR, and African - YRI).



852

853

854 **Figure 3**. HIPadjBMI Synthesis View plot that shows the -log10 *P*-values, beta (effect estimate),
855 effect/coded allele frequency (CAF), and sample size across analysis samples for all loci that reached
856 suggestive significance in one or more of our discovery strata. This chart also shows the coded allele
857 frequency (CAF) of each of our top loci by background group and by 1000 genomes reference panel
858 (European - EUR, Latin American - AMR, and African - YRI).



859

860

Table 1. Summary of association results for all loci that passed replication criteria. EAF-estimated allele frequency, chr- chromosome, pos- position, SE- standard error, ISQ- I squared heterogeneity.

Stratum	dbSNPID	CHR	POS (GRCh38)	Nearest Gene	Effect Allele	Other Allele	Stage	EAF	Beta	SE	P	N
WHRadjBMI												
AA[ł] - Combined	rs13301996	9	120570806	*CDK5RAP2*	T	G	SOL	0.8080	0.0050	0.0010	5.69E-07	12,672
							Replication	0.8720	0.0036	0.0014	1.10E-02	12,496
							SOL + Replication	0.8295	0.0045	0.0008	**2.88E-08**	25,168
HL[‡] - Women	rs79478137	11	2891739	*SLC22A18AS*	T	C	SOL	0.0150	-0.0230	0.0040	2.03E-07	7,472
							Replication	0.0169	-0.0116	0.0054	3.12E-02	6,582
							SOL + Replication	0.0157	-0.0189	0.0032	**3.64E-09**	14,054
HIPadjBMI												
EUR[¥] - Women	rs28692724*	14	77027445	*IRF2BPL*	T	C	SOL	0.4250	0.0020	0.0004	7.32E-07	7462
							Replication	0.303	0.789	0.305	9.62E-03	4,678
							SOL + Replication	0.3781	5.4900		**4.02E-08**	12,140
WCadjBMI												
EUR[¥] - Combined	rs3168072	11	61864038	*FADS2*	A	T	SOL	0.7250	0.5140	0.1020	5.28E-07	12,674
							Replication	0.9750	2.0132	0.6323	1.45E-03	8,845
							SOL + Replication	0.7313	0.5520	0.1007	**4.21E-08**	21,519

*For rs28692724, SOL analyses were performed on log10 transformed hip circumference, while replication analyses in european-descent population used untransformed hip measurements. In the SOL + Replication analyses, a z-score is provided instead of a beta.

łAfrican American (AA) replication samples included : Atherosclerosis Risk in Communities(ARIC) Study, Multi-Ethnic Study of Atherosclerosis(MESA), Women's Health Initiative(WHI)

‡Hispanic Latino (HL) replication samples included : Genetics of Latinos Diabetic Retinopathy(GOLDR), Hispanic Community Health Study / Study of Latinos(HCHS/SOL), Mexican–American Hypertension Study(HTN), Mexican-American Coronary Artery Disease(MACAD), Multi-Ethnic Study of Atherosclerosis(MESA), Mexico-City, 1982 Pelotas Birth Cohort(PELOTAS), Starr County Health

¥European American (EA) replication samples included : Atherosclerosis Risk in Communities(ARIC)

** rs28692724 is <500 Kb from a previously reported SNP nominally associated with WHRadjBMI (PMID: 28552196).

Table 2. Summary of top association results in SOL subgroup analyses. EAF-effect allele frequency, CHR- chromosome, POS- position, SE- standard error, ISQ- I squared heterogeneity. *EAF for reference population obtained from ExAC; all other estimated are from 1000 Genomes Project Phase 3.

Stratum	dnSNPID	CHR	POS (GRCh38)	Nearest Gene	Effect Allele	Other Allele	Subgroup	EAF	Beta	SE	P	N	Pdiff	EAF		
														AFR	EUR	AMR
WHRadjBMI																
AA - Combined	rs13301996	9	120570806	*CDK5RAP2*	T	G	SOL	0.8080	0.0050	0.0010	5.69E-07	12,672	1.88E-01	0.886	0.809	0.794
							Mainland	0.8220	0.0034	0.0014	1.95E-02	7,013				
							Caribbean	0.7890	0.0061	0.0017	2.52E-04	5,633				
HL - Women	rs79478137	11	2891739	*SLC22A18AS*	T	C	SOL	0.0150	-0.0230	0.0040	2.03E-07	7,472	3.97E-01	0.095	0.011	0.010
							Mainland	0.0080	-0.0241	0.0080	2.72E-03	4,220				
							Caribbean	0.0250	-0.0250	0.0056	8.87E-06	3,238				
HIPadjBMI																
EUR - Women	rs28692724*	14	77027445	*IRF2BPL*	T	C	SOL	0.4250	0.0020	0.0004	7.32E-07	7,462	2.31E-01	0.160	0.384	0.310
							Mainland	0.462	0.0017	0.001	8.70E-04	4,216				
							Caribbean	0.3765	0.0025	0.0006	4.02E-05	3,232				
WCadjBMI																
EUR - Combined	rs3168072	11	61864038	*FADS2*	A	T	SOL	0.7250	0.5140	0.1020	5.28E-07	12,674	3.94E-01	0.990	0.967	0.630
							Mainland	0.5969	0.4552	0.1204	1.57E-04	7,013				
							Caribbean	0.8846	0.4170	0.2075	4.45E-02	5,635				

863