Title

Ancestral diversity improves discovery and fine-mapping of genetic loci for anthropometric traits - the Hispanic/Latino Anthropometry Consortium

Authors and Affiliations

Lindsay Fernández-Rhodes,[1,2*] Mariaelisa Graff,[2*] Victoria L. Buchanan[2], Anne E. Justice,[2,3] Heather M. Highland,[2] Xiuqing Guo,[4] Kristin L. Young,[2] Kaustubh Adhikari,[5] Nicholette (Palmer) Allred,[6] Jennifer E. Below,[7] Jonathan Bradfield,[8] Alexandre C. Pereira,[9] LáShauntá Glover,[2] Daeeun Kim,[2] Adam G. Lilly,[10,11] Poojan Shrestha,[2,12] Alvin G. Thomas,[2] Xinruo Zhang,[2] Minhui Chen,[13] Charleston W. K. Chiang,[13,14] Sara Pulit,[15] Andrea Horimoto,[9] Jose E. Krieger,[9] Marta Guindo-Martinez,[16,17] Michael Preuss,[16] Claudia Schumann,[18] Roelof A.J. Smit,[16] Gabriela Torres-Mejía,[19] Victor Acuña-Alonzo,[20] Gabriel Bedoya,[21] Maria-Cátira Bortolini,[22] Samuel Canizales-Quinteros,[23] Carla Gallo,[24] Rolando González-José,[25] Giovanni Poletti,[24] Francisco Rothhammer,[26] Hakon Hakonarson,[8] Robert Igo,[27] Sharon G Adler,[28] Sudha K. Iyengar,[27] Susanne B. Nicholas,[29] Stephanie M. Gogarten,[30] Carmen R. Isasi,[31] George Papnicolaou,[32] Adrienne M. Stilp,[30] Qibin Qi,[31] Minjung Kho,[33] Jennifer A. Smith,[33] Carl Langfeld,[34] Lynne Wagenknecht,[35] Roberta Mckean-Cowdin,[36] Xiaoyi Raymond Gao,[37] Darryl Nousome,[36] David V. Conti,[13] Ye Feng,[36] Matthew A. Allison,[38] Zorayr Arzumanyan,[4] Thomas A. Buchanan,[29,39] Yii-Der Ida Chen,[4] Pauline M. Genter,[40] Mark O. Goodarzi,[41] Yang Hai,[4] Willa Hsueh,[42] Eli Ipp,[29,40] Fouad R. Kandeel,[43] Kelvin Lam,[4] Xiaohui Li,[4] Jerry L. Nadler,[44] Leslie J. Raffel,[45] Kaye Roll,[4] Kevin Sandow,[4] Jingyi Tan,[4] Kent D. Taylor,[4] Anny H. Xiang,[46] Jie Yao,[4] Astride Audirac-Chalifour,[47] Jose de Jesus Peralta Romero,[47] Fernando Hartwig,[48] Bernando Horta,[48] John Blangero,[49] Joanne E. Curran,[49] Ravindranath Duggirala,[49] Donna E. Lehman,[50] Sobha Puppala,[51] Laura Fejerman,[52] Esther John,[53] Carlos Aguilar-Salinas,[54] Noël P. Burtt,[55] Jose C. Florez,[55-57] Clicerio González-Villalpando,[58] Josep Mercader,[55-57] Lorena Orozco,[59] Teresa Tusié,[60] Estela Blanco,[61] Sheila Gahagan,[61] Nancy J. Cox,[7] Craig Hanis,[62] Nancy F. Butte,[63] Shelley A. Cole,[64] Anthony G. Commuzzie,[65] V. Saroja Voruganti,[66] Rebecca Rohde,[2] Yujie Wang,[2] Tamar Sofer,[57,67] Elad Ziv,[68] Struan F.A. Grant,[8] Andres Ruiz-Linares,[69-71] Jerome I. Rotter,[4] Christopher A. Haiman,[13] Esteban J. Parra,[72] Miguel Cruz,[47] Ruth J.F. Loos,[16] Kari E. North.[2,73]

[1]Department of Biobehavioral Health, Pennsylvania State University, University Park, PA, USA 16802
[2]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA 27599
[3]Department of Biomedical and Translational Informatics, Geisinger Health System, Danville, PA, USA 17822
[4]The Institute for Translational Genomics and Population Sciences, The Lundquist Institute for Biomedical Innovation at Harbor-University of California Los Angeles Medical Center, Torrance, CA, USA 90502
[5]School of Mathematics and Statistics, Faculty of Science, Technology, Engineering and Mathematics, The Open University, Milton Keynes, United Kingdom MK7 6AA
[6]Center for Diabetes Research, Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC, USA 27101
[7]Vanderbilt Genetics Institute, Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA 37232
[8]Center for Applied Genomics, Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA 19104
[9]Laboratory of Genetics and Molecular Cardiology, Heart Institute, University of São Paulo, Brazil 05508-220

[10]Department of Sociology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA 27599
[11]Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA 27599
[12]Division of Pediatric and Public Health, Adams School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA 27599
[13]Center for Genetic Epidemiology, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA 90033
[14]Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA, USA 90007
[15]Vertex Pharmaceuticals, Oxford, United Kingdom W2 6BD
[16]The Charles Bronfman Institutes for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA 10029
[17]The Novo Nordisk Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark 2200
[18]Hasso Plattner Institute, University of Potsdam, Digital Health Center, Potsdam, Germany 14482
[19]Department of Research in Cardiovascular Diseases, Diabetes Mellitus, and Cancer, Population Health Research Center, National Institute of Public Health, Cuernavaca, Morelos, Mexico 62100
[20]National Institute of Anthropology and History, Mexico City, Mexico 06600
[21]Molecular Genetics Investigation Group, University of Antioquia, Medellín, Colombia 1226
[22]Department of Genetics, Federal University of Rio Grande do Sul, Porto Alegre, Brazil 90040-060
[23]Population Genomics Applied to Health Unit, the National Institute of Genomic Medicine and the Faculty of Chemistry at the National Autonomous University of Mexico, Mexico City, Mexico 04510
[24]Research and Development Laboratories, Faculty of the Sciences and Philosophy, Peruvian University Cayetano Heredia, Lima, Peru 15102
[25]Patagonian Institute of the Social and Human Sciences, Patagonian National Center, Puerto Madryn, Argentina U9120
[26]Institute of High Studies, University of Tarapacá, Arica, Chile 1000000
[27]Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA 44106
[28]Division of Nephrology and Hypertension, Harbor-University of California Los Angeles Medical Center, Torrance, CA, USA 90502
[29]Department of Medicine, David Geffen School of Medicine at University of California, Los Angeles, CA, USA 90095
[30]Department of Biostatistics, University of Washington, Seattle, WA, USA 98195
[31]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA 10461
[32]National Heart, Lung and Blood Institute, Bethesda, MD, USA 20892
[33]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA 48109
[34]Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC USA 27101
[35]Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC USA 27101
[36]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA 90032

[37]Department of Opthalmology and Visual Sciences, Department of Biomedical Informatics, Division of Human Genetics, The Ohio State University, Columbus, OH, USA 43210
[38]Department of Family Medicine and Public Health, University of California, San Diego, CA, USA 92161
[39]Department of Physiology and Biophysics, Keck School of Medicine of USC, Los Angeles, CA USA 90033
[40]Department of Medicine, Division of Endocrinology, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA, USA 90502
[41]Division of Endocrinology, Diabetes, and Metabolism, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, USA 90048
[42]Department of Internal Medicine, The Ohio State University Wexner Medical Center, Columbus, OH, USA 43210
[43]Department of Translational Research & Cellular Therapeutics, Beckman Research Institute of City of Hope, Duarte, CA USA 91010
[44]New York Medical College, School of Medicine, Valhalla, NY, USA 10595
[45]Division of Genetic and Genomic Medicine, Department of Pediatrics, University of California, Irvine, CA, USA 92697
[46]Research and Evaluation Branch, Kaiser Permanente of Southern California, Pasadena, CA, USA 91101
[47]Medical Research Unit in Biochemistry, Specialty Hospital, National Medical Center of the Twenty-First Century, Mexican Institute of Social Security, Mexico City, Mexico 06725
[48]Postgraduate Program in Epidemiology, Federal University of Pelotas, Pelotas, Brazil 96010-610
[49]Department of Human Genetics and South Texas Diabetes and Obesity Institute, School of Medicine, University of Texas Rio Grande Valley, Brownsville and Edinburg, TX, USA 78520 and 78539
[50]Department of Medicine, School of Medicine, University of Texas Health San Antonio, San Antonio, TX, USA 78229
[51]Department of Internal Medicine, Wake Forest University, Winston-Salem, NC, USA 27109
[52]Department of Public Health Sciences, School of Medicine, and the Comprehensive Cancer Center, University of California Davis, Davis, CA, USA 95616
[53]Departments of Epidemiology & Population Health and Medicine-Oncology, Stanford University School of Medicine, Stanford, CA, USA 94305
[54]Division of Nutrition, Salvador Zubirán National Institute of Health Sciences and Nutrition, Mexico City, Mexico 14080
[55]Programs in Metabolism and Medical and Population Genetics, Broad Institute of the Massachusetts Institute of Technology and Harvard, Cambridge, MA, USA 02142
[56]Department of Medicine, Harvard Medical School, Boston, MA, USA 02115
[57]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA 02114
[58]Center for Diabetes Studies, Research Unit for Diabetes and Cardiovascular Risk, Center for Population Health Studies, National Institute of Public Health, Mexico City, Mexico 14080
[59]Laboratory of Immunogenomics and Metabolic Diseases, National Institute of Genomic Medicine, Mexico City, Mexico 14610
[60]Molecular Biology and Medical Genomics Unity, Institute of Biomedical Research, the National Autonomous University of Mexico and the Salvador Zubirán National Institute of Health Sciences and Nutrition, Mexico City, Mexico 14080[61]Center for Community Health, Division of Academic General Pediatrics, University of California at San Diego, San Diego, CA, USA 92093
[62]University of Texas Health Science Center at Houston, Houston, TX, USA 77030

[63]United States Department of Agriculture, Agricultural Research Service, The Children's Nutrition Research Center, and the Department Pediatrics, Baylor College of Medicine, Houston, TX, USA 77030
[64]Population Health Program, Texas Biomedical Research Institute, San Antonio, TX, USA 78227
[65]The Obesity Society, Silver Spring, MD, USA 20910
[66]Department of Nutrition, University of North Carolina Nutrition Research Institute, University of North Carolina, Kannapolis, NC, USA 28081
[67]Division of Sleep and Circadian Disorders, Brigham and Women's hospital, Boston MA, USA 02115
[68]Division of General Internal Medicine, Department of Medicine, Helen Diller Family Comprehensive Cancer Center, Institute for Human Genetics, University of California, San Francisco, San Francisco, CA, USA 94115
[69]Ministry of Education Key Laboratory of Contemporary Anthropology and Collaborative Innovation Center of Genetics and Development, School of Life Sciences and Human Phenome Institute, Fudan University, Shanghai, China 200438
[70]Department of Genetics, Evolution and Environment, and Genetics Institute of the University College London, London, UK WC1E 6BT
[71]Laboratory of Biocultural Anthropology, Law, Ethics, and Health, Aix-Marseille University, Marseille, France 13385
[72]Department of Anthropology, University of Toronto- Mississauga, Mississauga, ON, Canada L5L 1C6
[73]Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA 27514

*Co-first authors.
Present Address for Correspondence:
Lindsay Fernandez-Rhodes, 219 Biobehavioral Health Building, University Park, PA 16802, fernandez-rhodes@psu.edu
Mariaelisa Graff, 123 West Franklin Street, Chapel Hill, NC 27516
migraff@email.unc.edu

ABSTRACT

Hispanic/Latinos have been underrepresented in genome-wide association studies (GWAS) for anthropometric traits despite notable anthropometric variability with ancestry proportions, and a high burden of growth stunting and overweight/obesity in Hispanic/Latino populations. This address this knowledge gap, we analyzed densely-imputed genetic data in a sample of Hispanic/Latino adults, to identify and fine-map common genetic variants associated with body mass index (BMI), height, and BMI-adjusted waist-to-hip ratio (WHRadjBMI). We conducted a GWAS of 18 studies/consortia as part of the Hispanic/Latino Anthropometry (HISLA) Consortium (Stage 1, n=59,769) and validated our findings in 9 additional studies (HISLA Stage 2, n=9,336). We conducted a trans-ethnic GWAS with summary statistics from HISLA Stage 1 and existing consortia of European and African ancestries. In our HISLA Stage 1+2 analyses, we discovered one novel BMI locus, as well two novel BMI signals and another novel height signal, each within established anthropometric loci. In our trans-ethnic meta-analysis, we identified three additional novel BMI loci, one novel height locus, and one novel WHRadjBMI locus. We also identified three secondary signals for BMI, 28 for height, and two for WHRadjBMI. We replicated >60 established anthropometric loci in Hispanic/Latino populations at genome-wide significance—representing up to 30% of previously-reported index SNP anthropometric associations. Trans-ethnic meta-analysis of the three ancestries showed a small-to-moderate impact of uncorrected population stratification on the resulting effect size estimates. Our novel findings demonstrate that future studies may also benefit from leveraging differences in linkage disequilibrium patterns to discover novel loci and additional signals with less residual population stratification.

INTRODUCTION

A complex interplay between political, social, and economic factors has led to an increasing obesogenic global environment. In this modern context, many low- to middle- income nations have experienced a rapid transition from under-nutrition and growth stunting to over-nutrition and obesity.[1] Moreover, population-based surveys from 1975-2002 show that there is an inverse ecologic relationship between the prevalence of growth stunting and the prevalence of overweight seen among preschool children (0-5 years of age) in Latin America.[2] Growth stunting of preschool children ranges from relatively rare (7%) in the Caribbean to notably common (20%) in Central America. Moreover, it is a risk factor for overweight/obesity independent of a child's socioeconomic status.

In Latin America, by 2016 35% of the total population was overweight [body mass index (BMI) 25 to <30 kg/m^2] and another 23% was living with obesity.[3] In Mexico, more than 71% of adults are currently overweight;[4] it is projected that by 2050 only 12% of men and 9% of women will have a healthy weight (BMI <25 kg/m^2). In a recent study in Argentina, Chile, and Uruguay, the prevalence of obesity was 36%, but when using waist circumference as a measure of central obesity, it was far higher (53%).[5] Within each of these populations, there are also disparities in obesity by sex and education.

Race, ethnicity, and ancestry may play a role in anthropometric-related health disparities in Latin American. Previous studies have described the historical contexts leading to admixture in Latin American populations[6; 7] as characterized by highly diverse (variable) ancestral proportions[8-10] from any of the following regions: the Americas, Europe, Africa and East Asia.[11-16] In fact, proportion of Native American ancestry is associated with numerous biomedical traits, like obesity-related traits, and is most strongly associated with height.[17; 18] Height is inversely associated with proportion of Native American ancestry, even after taking into account the fact that globally over time populations have become taller due to mainly non-genetic nutritional factors.[16] The ultimate drivers of this association remain to be elucidated; it is possible that genetic factors and/or socio-economic factors strongly associated with Native American ancestry could be responsible for these findings. Recent studies are starting to provide relevant insights on this topic. As an example, a recent genome-wide association study (GWAS) in Peru[19] identified a missense variant in the *FBN1* gene (rs200342067) that has the largest effect size so far described for common height-associated variants in human populations (each copy of the minor allele reduces height by 2.2 cm). In the 1000 Genomes Project samples, rs200342067 is only present in two American samples (MXL: 0.78% and PEL: 4.12%), and yet

the authors reported that this missense variant shows subtle evidence of positive selection in the Peruvian population.[19]

Obesity in Latin America has quickly surpassed the levels previously seen only among adults of high-income nations, like Canada and the United States (US). In Canada the number of people reporting Latin American origins grew by 83% from the 2001 census[20] relative to the 2016 census,[21] representing 1.3% of the total Canadian population. In the US, both the population size and diversity in national origins (backgrounds) of US Hispanic/Latinos have been increasing over the past several decades.[22] If past demographic trends continue, 24% of the US adult population will identify as Hispanic/Latino by 2065.[22] Obesity-related financial costs in the US are projected to double every decade to ~$900 billion by 2030.[23; 24] US Hispanic/Latino adults and their children/adolescents face a greater burden of obesity than their non-Hispanic white counterparts.[25-28] There is a need to study Hispanic/Latino populations in order to address these disparities.[28; 29]

Given the unique historical and recent demographic shifts occurring across the Americas, there is a clear need to also understand the role that Native American or other under-studied components of admixture have on the genetic architecture of anthropometric traits in Hispanic/Latinos, and its relationship with risk of downstream poor health outcomes. Yet, to date no large-scale GWAS of anthropometric traits have been conducted among Hispanic/Latino populations. Here, we perform the largest genomic study to date of anthropometric traits, including BMI, height, and waist-to-hip ratio adjusted for BMI (WHRadjBMI) in Hispanic/Latino populations to describe what might be novel loci or signals in established loci in this population by sex and life stage.

MATERIALS AND METHODS

Hispanic/Latino Study Samples

The Hispanic/Latino Anthropometry (HISLA) Consortium is comprised of 27 studies/ consortia of adult participants. First, HISLA Stage 1 includes 17 studies and one consortium (Consortium for the Analysis of the Diversity and Evolution of Latin America, CANDELA[18]) collectively representing up to 59,771 adults, depending on the trait, from Brazil, Chile, Colombia, Mexico, Peru, or the US with self-reported heritage from across Spanish-speaking Latin America, or Native American heritage, primarily Pima and Zuni[30] (**Table S1**). HISLA Stage 2 includes nine studies with up to 10,538 adults from across Spanish-speaking Latin America or with related heritage and living in the US (**Table S1**).

This study was approved by the Institutional Review Boards of the University of North Carolina at Chapel Hill, and all contributing studies had received prior Institutional Review Boards approval for each study's activities.

Anthropometric Traits

BMI is a commonly derived index of obesity risk and is calculated as the ratio of body weight to height squared (kg/m^2). Adult height was measured/self-reported using either metric units, or US units and then converted to meters. Waist-to-hip ratio (WHR) is used to capture central fat deposition, and it is derived from the circumference of the waist at the umbilicus compared to the circumference of the hip at the maximum protrusion of the gluteal muscles.

Residuals were calculated by sex and/or case status, adjusting for age, age^2, and study-specific covariates [e.g., center, principal components of ancestry (PCA)]. For WHR, BMI was also adjusted for when creating the residuals to isolate the central deposition of fat from overall body mass. Residuals were then used to create inverse normalizations of BMI and WHRadjBMI, and z-scores of height (=residual/standard deviation for all residuals). In family-based studies the residuals were calculated in women and men together, adjusting for age and sex and other study covariates including PCs. Descriptive statistics on the covariates and anthropometric measures of are provided for each study's analytic sample in **Table S2**. Only one family-based study in Stage 1 and two non-family based studies in Stage 2 (GOLDR 0.3% <18 years, and HTN-IRS 3.9%) included a small subset of adolescents aged 15-17 years, each less than 5% of the total sample. All other study samples included individuals 18-98 years of age.

Childhood/Adolescence Study Samples, Anthropometric Traits, and Obesity

We assembled an independent sample of children/adolescents with anthropometrics, from three studies from the US, Mexico and Chile (**Table S3**). The distribution of covariates and anthropometrics of the samples of children/adolescents in each analysis are described in **Table S4.** First, childhood/adolescent obesity was defined as having a ≥95th BMI-for-age percentile versus ≤50th BMI-for-age percentiles, based on the Centers for Disease Control and Prevention growth curves,[31] as done in previous analyses of childhood obesity.[32] We used these two analyses to look up novel BMI and height findings from our adult HISLA meta-analysis and our trans-ethnic analyses. This resulted in 1,814 children/adolescents aged 2-18 years for this case-control analysis (**Tables S3-4**). Second, BMI and height-for-age z-scores were calculated in children/adolescents aged 5-18 years from the US and Chile (**Table S4**) based on the more international reference growth curves from the World Health Organization.[33] In Viva la Familia, a

family-based study,[34] these residuals were calculated adjusting for sex in the combined sample. The resulting BMI and height-for-age z-scores were available for 1,914 and 1,945 children/adolescents, respectively.

SNP Imputation and Statistical Analyses

We generated autosomal genome-wide imputed data based on 1000 Genomes Phase 1 and 3 references, with the exception of two studies that contributed Exomechip and MetaboChip (Illumina, Inc.; San Diego, CA) genotypes and one study that blended genotypes from multiple platforms (**Tables S5-6**). PCA analyses were conducted in each study to capture the main components of genetic ancestry from the Americas, Europe, Africa, and Asia. Studies with samples from related individuals accommodated this non-independence by projecting their principal component analysis from the reference to the study sample, and by accounting for relatedness using either generalized estimating equations[35] or mixed linear models.[10; 36] Assuming an additive genetic model, we tested the association of over 20 million autosomal variants on our traits, accounting for all trait or study-specific covariates (e.g., center, PCA).

Meta-Analyses of HISLA Stage 1+2

The studies of the HISLA Consortium were meta-analyzed in two stages, including discovery (Stage 1) and validation (Stage 2). Stage 1 included a total sample of 59,771 individuals with data on BMI, 56,161 with height, and 42,455 with WHRadjBMI. All Stage 1 studies/consortia provided full genome-wide analysis results. All SNPs that met our significance criteria were brought forward for validation in Stage 2, which included 10,538 individuals with data on BMI, 8,110 with height, and 4,393 with WHRadjBMI. All reported association results passed our quality control criteria; i.e., variants with low quality (info score <0.4 or Rsq<0.3), minor allele count (MAC) <5, or sample size <100 were removed. We meta-analyzed effects across all studies using a fixed-effect inverse variance weighted meta-analysis with genomic control in METAL.[37] Given the unique patterns of admixture and ancestry represented by the Brazilian or Native American samples, we conducted sensitivity analyses in Stage 1 studies (i.e., comparing the inclusion and exclusion of the Baependi Heart Study, 1982 Pelotas Birth Cohort Study, and Family Investigation of Nephropathy and Diabetes substudy of individuals of Pima and Zuni heritage) to assess the influence of the three studies on the meta-analysis results. CANDELA was retained in all analyses as <10% of the consortium's samples came from Brazil, primarily originating from the South of Brazil with wide-spread European heritage with a lesser extent Native American or African admixture.[18]

Regional plots of all GWAS-significant HISLA Stage 1 findings were plotted using LocusZoom (https://locuszoom.org). From Stage 1, we selected lead variants that met genome-wide significance ($P<5x10^{-8}$) that were independent of each other for replication. In cases where Stage 2 studies did not have the lead variant, we selected two proxies per lead variant with an $r^2 \geq 0.9$ using 1000 Genomes AMR linkage disequilibrium (LD). Stage 2 studies provided a list of the requested lead variants and/or their proxies from Stage 1 for validation. Stage 2 studies were meta-analyzed and subsequently combined with Stage 1 using METAL[25]. Effect heterogeneity was assessed through I^2 across all 27 HISLA adult studies/consortia by entering each study separately into the meta-analysis, irrespective of stage. The characteristics of the final SNP array data used in the HISLA adult studies and the children/adolescent Hispanic/Latino studies are summarized separately in **Tables S5-6.**

Meta-Analyses of HISLA Stage 1 with Other Ancestral Consortia

In addition to a Hispanic/Latino only meta-analysis, we combined the HISLA Stage 1 meta-analysis with data from previous large-scale GWAS meta-analyses from European (the Genetic Investigation of Anthropometric Traits, GIANT, Consortium[38-40], N ~ 300,000) and/or African (the African Ancestry Anthropometry Genetics Consortium, AAAGC[41], N ~ 50,000) descent populations. We used fixed-effect inverse variance weighted meta-analytic techniques in METAL to generate our trans-ethnic meta-analysis.[37] We validated our potentially novel BMI, height[42] and WHRadjBMI[43] findings from this trans-ethnic meta-analysis in either our independent sample of Hispanic/Latino children/adolescents or the British subsample GWAS of the United Kingdom Biobank (UKBB). Regional plots of these analyses of all potentially novel trans-ethnic findings are shown in the supplement (**Figures S7-52**).

Thresholds for Conditional Signals, Discovery, Validation and Transferability

We conducted approximate conditional analyses using the Genome-wide Complex Trait Analysis (GCTA, version 1.93.1) software. For the HISLA analyses, we used our Stage 1 discovery results with the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) as the LD reference dataset. For the approximate conditional trans-ethnic analyses, we used our trans-ethnic results from HISLA Stage 1, AAAGC, or GIANT and a trans-ethnic LD reference dataset of Europeans and Africans from the Atherosclerosis Risk in Communities (ARIC) cohort, and Hispanic/Latinos from HCHS/SOL, which was representative of the ancestry distribution of our meta-analysis. In both conditional analyses (HISLA only and trans-ethnic results), we first identified all independent SNPs using the --cojo-slct command. Then, we conditioned each of

these independent SNPs on all known SNPs up through December 2019 (BMI[38; 41; 42; 44-59], Height[39; 42; 51; 54; 59; 60], WHRadjBMI[40; 41; 43; 48; 50; 58; 59; 61-65]) within 10Mb of the index SNP. The trans-ethnic meta-analysis results with a *P-value*<$5x10^{-8}$ after conditioning on known SNPs were taken forward for validation in the British subsample of the UKBB.

SNP associations were then defined as either newly discovered or established, depending on their location. An established locus was defined as a SNP association within ±500 kb of at least one previously identified index SNP, otherwise the association was considered a newly-discovered locus.

We designated our Hispanic/Latino SNP-associations within either newly-discovered or established loci as novel if they met the following criteria: 1) were associated at *P-value*<$5x10^{-8}$ in HISLA Stage 1 and directionally consistent in Stage 2, and 2) the addition of Stage 2 samples improved the estimated Stage 1+2 meta-analysis. For the trans-ethnic analyses these criteria were as follows: 1) were associated at *P-value*<$5x10^{-8}$ in the combined HISLA, AAAGC and GIANT meta-analysis, and 2) were both directionally consistent and associated at *P-value*<$5x10^{-2}$ in the subsample of Hispanic/Latino children/adolescents or in the British subsample GWAS from the UKBB.

Novel Hispanic/Latino SNP effects were considered to transfer to Hispanic/Latino children/adolescents, or to African or European ancestry adults, if they were 1) directionally consistent, 2) associated at *P-value*<$5x10^{-2}$, and 3) had a heterogeneity of I^2<75% in either the Hispanic/Latino children/adolescent lookups, or either 1) the AAAGC or 2) the GIANT adult GWAS results. Conversely, SNP effects of variants previously associated with anthropometric traits in non-Hispanic/Latino populations (i.e., index published SNPs) were considered to be transferable (generalizable) to Hispanic/Latinos if they were 1) directionally consistent, 2) displayed a *P-value*<$5x10^{-2}$, and 3) had little to moderate effect heterogeneity (I^2<75%) in Stage 1.

Fine-Mapping Methods

We used FINEMAP[66] for fine-mapping analyses of the newly-discovered loci identified as part of the HISLA Stage 1 meta-analysis or trans-ethnic meta-analysis, and in established loci. For the established loci, we included index SNP-associations published as of April 2018 (BMI[38; 41; 44; 46-48; 50; 52; 55-58], Height[39; 54; 60], WHRadjBMI[40; 41; 48; 50; 67]) prior to the publications with the UKBB results.[42; 43] We used a 1Mb region subset of the summary statistics from the Stage 1 meta-analyses and HCHS/SOL[10] unrelated sample set (N ~ 7,670) to calculate the LD for each locus.

For trans-ethnic fine-mapping of the novel loci and signals identified in the trans-ethnic meta-analysis of HISLA, AAAGC, and GIANT, we used a 1Mb region defining each locus using the summary statistics of the given meta-analysis. We calculated the LD for Hispanic/Latino samples using the HCHS/SOL[10] unrelated sample (N ~ 7,670). For African and European ancestry samples, we calculated the LD using the ARIC unrelated sample that included self-reported African ancestry (N ~ 2,800) and European ancestry (N ~ 9,700). We weighted the LD matrices by the GWAS sample sizes for each trait (HISLA range: ~42,400-56,100; AAAGC: 20,300-42,700; GIANT: 210,000-330,000).

All regions allowed up to a maximum of 10 causal variants. The cumulative 95th% credible set was calculated from the estimated posterior probabilities. Convergence failed for three regions (lead SNPs: rs2902635, rs6900530, and rs4425978, all in known height loci) using the stochastic approach. For these three regions, we used the conditional approach to determine number of causal variants.

Bioinformatics Analyses

To gain further insight into the possible functional role of the identified variants and to assess their relevance to other phenotypes, we conducted bioinformatics queries of our potentially novel loci and novel signals within known loci in multiple publicly available databases, including PhenoScanner,[68] RegulomeDB,[69] Haploreg,[70] UCSC GenomeBrowser,[71] and GTEx.[72]

Trans-Ethnic Findings to Account for Population Structure in Previous GWAS

To quantify the impact of population stratification, we computed the correlation between PC loadings and beta effects estimated from GWAS. We first conducted PCA analysis on the four European populations (CEU, GBR, IBS, and TSI) from 1000 Genomes. We excluded the FIN (Finnish in Finland) population because of its known unique demographic history.[38] We only used biallelic SNPs with minor allele frequency (MAF) > 5% in the four European populations, and then pruned them by both distance and LD using PLINK 1.9.[73] Specifically, we pruned the dataset such that no two SNPs were closer than 2 kb, and then pruned using a 50 SNP LD window (moving in steps of 5 SNPs), such that no SNPs had r^2>0.2. We further removed SNPs in regions of long-range LD.[74] PCA was performed on the remaining SNPs using Eigensoft version 7.2.1(https://github.com/DReichLab/EIG/archive/v7.2.1.tar.gz).

We performed linear regressions of individual PC values on the allelic genotype count for each polymorphic variant in the four European populations from 1000 Genomes and used the resulting regression coefficients as the estimate of the variant's PC loading. For each PC,

we then computed Pearson correlation coefficients of PC loadings and effect sizes (of variants with MAF>1%) from each GWAS summary statistics. We estimated *P-values* based on Jackknife standard errors, by splitting the genome into 1,000 blocks with an equal number of variants. If there was significant correlation in either the GIANT dataset or the HISLA Stage 1, AAAGC and GIANT trans-ethnic meta-analysis, we then further evaluated the improvement of bias due to stratification in trans-ethnic meta-analysis by comparing the correlation coefficients in the trans-ethnic meta-analysis with those in GIANT. Restricting to variants shared between GIANT and the trans-ethnic meta-analysis, we computed their difference in correlation coefficients of PC loadings and effect sizes, and estimated *P-values* again based on Jackknife standard errors from 1,000 equal sized blocks.

RESULTS

One Novel BMI Locus Discovered and Validated in Hispanic/Latino Adults

The first goal of this study was to conduct a genome-wide meta-analysis of anthropometric traits in Hispanic/Latino adults to identify novel loci in an under-studied population (**Figure 1**). All regional plots of all GWAS-significant HISLA Stage 1 findings are shown in the supplement **(Figures S1-6).** No novel loci were identified in all samples combined. Yet, when excluding the Brazilian or Native American samples from Stage 1, we discovered one locus for adult BMI at *PAX3* on chromosome 2 in the HISLA Stage 1 sample (**Table S7**), and we validated this locus in HISLA Stage 2 (**Table 1**). The lead SNP*,* rs994108*,* is in moderate LD (rs7559271, r^2=0.46, D'=1.0 in 1000 Genomes phase 3 AMR) (**Figure 2**) and lies on the same haplotype as a SNP reported to influence facial morphology, including position of the nasion (the deepest point on the nasal bridge where the nose meets the forehead) in Europeans[75] and Hispanic/Latino[76] descent individuals. Other *PAX3* variants in lower LD with the lead SNP have also been associated with nasion position,[77] monobrow, and male-pattern baldness.[78; 79] *PAX3* is a well-known transcription factor in normal embryonic neural crest development and differentiation.[80] Neural crest cells can give rise to mesenchymal stem cells,[81] which can in turn give rise to adipocytes;[81-83] thus, the possible role of *PAX3* in adipogenesis may at least partially explain the association signal with BMI near this gene. Another BMI SNP (rs1505851) near *ARRDC3* on chromosome 5 found at GWAS significance in HISLA Stage 1 (**Table S7, Figure S1**) did not validate in Stage 2 (**Table 1**).

We identified two WHRadjBMI loci at *DOCK2* and *TAOK3* at GWAS significance in HISLA Stage 1 after excluding the Brazilian and Native American samples (**Table S7, Figures**

S2-S3), and neither met the p-value threshold for replication and in HISLA Stage 2. The *DOCK2* association for WHRadjBMI was observed among women in Stage 1was, however, directionally consistent among women in Stage 2. The *TAOK3* association was led by a low frequency variant (rs115981023) that was not directionally consistent across Stages. rs115981023 exhibited moderate heterogeneity across Stage 1 samples after excluding Brazilian and Native American samples (I^2=45%), and this heterogeneity remained (I^2=52%) in the combined meta-analysis of HISLA Stage 1 and 2 samples (**Table 1**).

No potentially novel loci were identified for height in HISLA Stage 1, and the exclusion of the Brazilian and Native American samples did not reveal additional novel height or WHRadjBMI loci.

Three Novel Signals in Established Loci for BMI and Height Discovered and Validated in Hispanic/Latino Adults

At two established loci for BMI, we identified new signals at *ADCY5* and near *C6orf106* (**Table S7**), which were independent of any published findings (**Table S8, Figures S4-5**). We validated these signals in Stage 2 with directional consistency and the combined Stage 1+2 meta-analysis at GWAS significance (**Table 1**). We also identified one new signal for height in an established height locus, *B4GALNT3,* which was independent of the previously reported SNPs for height (**Tables S7-8, Figure S6**). We validated this signal in Stage 2 with directional consistency and a Stage 1+2 meta-analysis that was GWAS significant (**Table 1**).

Fine-Mapping of Novel Adult Hispanic/Latino Anthropometric Findings

We fine-mapped using 1MB regions, the novel *PAX3* locus for BMI and three new signals in known loci discovered and replicated in Stages 1+2 (BMI: *ADCY5* and *C6orf106*; height: *B4GALNT3*; **Table S9**). For the three BMI loci, FINEMAP revealed one potential causal set for each locus at *PAX3, ADCY5,* and *C6orf106* locus. For the *PAX3* locus, only one causal set was proposed and the 95th% credible contained only nine plausibly causal SNPs, with lead SNP rs994108 having a very high posterior probability of being causal (0.89). However, functional annotation of this SNP was unremarkable (**Table S20**). In contrast, for *ADCY5* and *C6orf106,* FINEMAP revealed one causal configuration for each locus but with much greater uncertainty with respect to the likely functional variant given the size of the credible sets, 14 and 22 SNPs in the credible region for *ADCY5* and *C6orf106*, respectively. The posterior probability of the best lead SNP at these loci had relatively low posterior probabilities of being the causal SNP, with the best posterior probabilities of 0.23 for rs17361324 (*ADCY5*), and 0.11 for

rs73420913 (*C6orf106*), respectively. Interestingly, however, the best candidate for causality at *PAX3* and *ADCY5* loci were the lead SNPs from the HISLA meta-analysis and for *C6orf106* , the FINEMAP and HISLA SNPs were in tight LD (rs73420913 had an r^2=0.96 with the lead HISLA SNP rs148899910), providing greater support for the prioritization of these SNPs for functional interrogation. For the *B4GALNT3* locus for height, FINEMAP revealed six causal configurations. Four of the variants (rs11063185, rs215230, rs7303572, and rs11063184 with each configuration each had a posterior probability >0.99 and contained only itself in the 95% credible set. One variant (rs215223) had a posterior probability of 0.93 and thus included two variants in the 95% credible set. The sixth 95% credible set had a lead variant with a posterior probability of 45%, but contained a total of 1621 additional variants all of which had very small posterior probabilities (i.e., ≤0.05).

Transferability of Adult Novel Loci/Signals from Hispanic/Latinos to Consortia of Other Ancestral Backgrounds

To assess how well the effect estimates are transferable (generalizable) to other populations, we looked up the novel BMI and height findings from Hispanic/Latinos in the AAAGC and GIANT meta-analysis results (**Table 1**). Keeping limitations with respect to sample size, LD, allele frequency, and effect size heterogeneity in mind, we did observe directionally consistent BMI effects at the *PAX3* locus in the other consortia, although without observing nominal significance. The new BMI signals at the *ADCY5* locus (rs17361324) transferred to both AAAGC and GIANT with directional consistency (betas=0.13-0.23) and nominal significance (*P-values*<$5x10^{-2}$). The BMI lead SNP (rs148899910) representing a novel signal near *C6orf106* was available in AAAGC, the signal only appeared to be transferable to GIANT at a proxy SNP (rs1573905, r^2=0.96-1 in 1000 Genomes AMR and EUR**; Table 1**).

The new signal for height in *B4GALNT3* (rs215226) was directionally consistent and nominally significant in AAAGC. In all cases the effect sizes observed in GIANT and AAAGC were attenuated compared to the effect sizes from HISLA Stage 1.

Relevance of Novel Hispanic/Latino Anthropometric Loci/Signals from Adults to Childhood/Adolescence

We looked up our novel HISLA findings in Hispanic/Latino children/adolescents using BMI-for-age and height-for-age z-scores, as well as a case-control study of childhood obesity. Two of the three novel BMI signals were directionally consistent with the anticipated effect on the odds of obesity during childhood/adolescence, one of which was nominally significant

(rs17361324 at *ADCY5*; *P-value*=2.2x10^{-2}). None of the novel HISLA findings generalized at nominal significance with the BMI/height-for-age z-score, but were directionally consistent with the corresponding effect in adulthood (**Table S10**). This may have been due to the small available sample size of Hispanic/Latino children/adolescents.

Transferability of Established Anthropometric Loci to Hispanic/Latino Adults

Using HISLA Stage 1 results, we assessed how many established anthropometric loci, discovered in predominantly non-Hispanic/Latino samples could be transferred to Hispanic/Latino adults, given the current sample size. As shown in **Table S11**, the index SNPs at 332 of 1280 (25.9%) previously reported BMI loci were transferable to Hispanic/Latinos. Of these BMI loci, 13 SNPs in the HISLA Stage 1 displayed genome-wide significant associations with the SNP reported in the literature (**Table S7**). **Table S12** shows that a slightly higher percentage of known height loci (1177 of 3925, or 30.0%) were transferable to Hispanic/Latinos. Forty-nine height loci displayed a genome wide significant association with height in the surrounding 1 MB interval in HISLA Stage 1 (**Table S7**), with 44 of 49 SNPs being the exact index SNP from the literature (**Table S11**). Lastly, **Tables S13-15** show that 143 of 754 (19.0%) known WHRadjBMI in both sexes combined, 103 of 504 (20.4%) in women only, and 28 of 186 (15.1%) in men only loci were transferable to Hispanic/Latinos. However, none of the index SNPs from the previous literature for WHRadjBMI reached genome-wide significance. We did observe genome-wide significant evidence for association of a SNP with WHRadjBMI in the 1 MB interval of one known region (*HOXC13*), although not replicating the exact previously reported index SNP (**Table S7**).

Five Novel Loci and Thirty-three New Signals in Established Loci for Adult Anthropometric Traits Discovered and Replicated in a Trans-Ethnic Meta-Analytic Context

As shown in **Figure 1,** we pursued a secondary goal of assembling a trans-ethnic meta-analysis of HISLA Stage 1 with the AAAGC and GIANT consortia results to attempt to further leverage differences in allele frequencies across populations to identify additional novel loci and fine-map established loci. As anticipated, this trans-ethnic meta-analysis revealed eight new loci and 35 new signals in established loci that were associated at GWAS significance in the combined HISLA, AAAGC and GIANT meta-analysis (**Table S16, Figures S7-S52**), and independent of established SNPs within a 10Mb region (**Table 2**). Of this set, five new loci (3 BMI, 1 height, and 1 WHRadjBMI) and 33 new signals in established loci (3 BMI, 28 Height, and

2 WHRadjBMI) were validated using the adult British subsample of the UKBB. In some cases, the significance in the trans-ethnic results had additional signal driven more by the AAAGC and/or HISLA consortia, which could explain the lack of association in the UKBB British subsample (**Table S16, Figure S53**). We looked up the potentially novel findings from our trans-ethnic meta-analyses in the sample of Hispanic/Latino children/adolescents (**Table S17)**. Four trans-ethnic SNPs were associated at nominal significance in the child/adolescent sample, each having been already replicated in UKBB (**Table S16**). Three of these four loci were directionally consistent in the childhood/adolescence results with the trans-ethnic adult findings (**Table S17**). In summary, we found that two of the seven novel BMI/height trans-ethnic loci and 17 of the 33 new trans-ethnic BMI/height signals in established loci were directionally consistent between their adult directions of association and the BMI/height-for-age z-scores in children/adolescents. However, this directional consistency was not more than what would have been expected by chance alone ($P\text{-}values_{binomial}$>0.10).

Fine-Mapping of Trans-Ethnic Anthropometric Findings

We also fine-mapped the novel trans-ethnic findings (**Table S18**) using FINEMAP[66] to pinpoint individual variants and genes within each locus region that have a direct effect on the trait. FINEMAP uses a shotgun stochastic search algorithm[84] that iterates through causal configurations of SNPs by concentrating efforts on the configurations with non-negligible probability. Within a 1MB region which was a novel locus for the given trait, or included a new signal within a known locus, we report the causal configuration of SNPs with highest posterior probability and the posterior probability that each of these SNPs is causal.

For four of the five novel loci (three for BMI and one for WHRadjBMI) there was one SNP within the configuration with the highest posterior probability. For the novel height locus near *ANKRD36BP1*, there were two SNPs in the configuration with the highest posterior probability. In all five novel loci, the SNP with the highest posterior probability from each of these credible sets was either the exact SNP with the strongest GWAS evidence or in high LD (r^2 between 0.70 and 0.99 in each ancestry) with the lead GWAS SNP. Two of these five regions had strong prioritization given high posterior probabilities (≥0.8) and small 95th% credible sets: 1) for BMI, the *CTD-2007H13.3* region had a posterior probability of 0.88 for rs150992 with three SNPs in the credible set, and 2) for height, the *ANKRS36BP1* region had a posterior probability of 0.93 for rs10737541 with five SNPs in the credible set. From the functional annotations (**Table S21 and S22**), we find that all three of the BMI loci, the height and

WHRadjBMI loci have enhancer marks and eQTLs, most of which are in relevant tissues, e.g., adipose, muscle, thyroid, or brain.

For the other novel trans-ethnic loci, the posterior probabilities were lower, between 0.09 and 0.42; yet, four loci (rs9860730, rs17375290, rs4324883, and rs9463108) still had relatively few SNPs (<10) in the 95th% credible sets suggesting a narrow window (combination of variants) around the causal variant. For example, functional annotations of rs17375290 lead GWAS SNP in the NFIA locus associated with height, show it to have promoter markers in muscle, CADD score of 13.29 (CADD > 10 ranks variants among the top 10% potentially deleterious), and an eQTL with *FGGY* in Osteoclast tissue (**Table S21 and S22**). Three of the other SNPs in the credible set (rs599989, rs1762881, and rs17121184) have nominally significant (p-value 0.01 to 0.005) eQTLs with *FGGY* in osteoclast tissue but are not in high LD with rs17375290 (r^2 range from 0.03 to 0.1). Diseases associated with *FGGY* include Lateral Sclerosis and Spastic Paraplegia 7, Autosomal Recessive, which is known to affect height.

Within the 33 novel trans-ethnic signals in known loci, 31 had configurations with more than one putative causal SNP (e.g. more than one credible set). This made sense given these are loci with multiple independent signals, as described by our earlier conditional analyses. Among the putative causal SNPs within each locus, there were a number of SNPs that represented previously-known signals (either the exact SNP or something in high LD among all ancestries). We found that for many of these the credible sets contained <10 SNPs. Among the 33 novel signals in known loci, 26 included a putative causal SNP that is the lead GWAS SNP reported here or a SNP in high LD (r^2 > 0.75) with the lead GWAS SNP, suggesting causality for this signal in general, though perhaps maybe not initially discovered at the most-putatively-causal SNP(s). For these 24 putatively-causal SNPs, the posterior probabilities ranged from 0.09 to 1. Twenty-two of these SNPs had 95th% credible sets that contained <10 SNPs and 15 also had posterior probability ≥0.8.

Many have functional annotations that help support the fine-mapping results (**Table S21-S22**). For example, we find eQTLs for the three BMI signals (and enhancer marks for rs4807179) in relevant tissues including adipose, brain, muscle, and/or thyroid. The lead SNPs of these credible sets had posterior probabilities >0.75 and the credible sets included <10 SNPs. Of the 28 newly identified height signals, we find 13 putatively-causal SNPs that are the lead GWAS SNP or are in high LD (r^2 > 0.75) with it, have <10 SNPs in the credible set and have eQTLs in relevant tissues including muscle, thyroid, adipose, lung and osteoclasts. Some also have promoter or enhancer marks in some of the same tissues. For the two WHRadjBMI signals, both have three SNPs in the most probably causal configurations. One of these causal

SNPs for each region is either the lead GWAS SNP (rs7975017) or a SNP in high LD (rs17099388 and rs6895040 LD: AFR R2=1.0; AMR R2=1.0; EUR R2=1.0), has a posterior probability ≥0.95, and is the only SNP in the credible set. Furthermore, for rs7975017 we find eQTLs in thyroid for multiple genes (*BHLHE41, SSPN,* and *AC022509.3*) and enhancer marks in multiple tissues including those related to the WHRadjBMI trait, e.g., thyroid, muscle, fat, bone, and adrenal gland. Overall, across many of the novel loci and secondary signals, FINEMAP revealed SNPs with somewhat strong prioritization (posterior probability ≥0.8) and at some loci putatively-causal SNPs have small 95th% credible sets thus demonstrating the utility of trans-ethnic approaches to fine mapping GWAS loci.

Trans-Ethnic Findings to Account for Population Structure in Previous GWAS

The first two PCs in the PCA (**Figure S54**) reflect geographical or population structure in Europe, corresponding to the North-South and Southeast-Southwest axes of variation, respectively. We found that the bias in effect size estimates due to stratification is most obvious for height as the phenotype is known to be differentiated across Europe.[85-87] Effect sizes on height estimated from the GIANT and our trans-ethnic meta-analysis were both highly correlated with the loadings of the first PCA (rho = 0.125, *P-value*= 3.2×10^{-94} in GIANT; rho = 0.105, *P-value*= 3.4×10^{-70} in meta-analysis). The correlation was much lower in AAAGC and HISLA (rho = 0.012, *P-value*= 2.17×10^{-4} in AAAGC; rho = 0.007, *P-value*= 9.2×10^{-2} in HISLA; **Figure 4A**). Importantly, the magnitude of correlation was lessened in meta-analysis compared with GIANT (*P-value*= 6.6×10^{-9}). Other traits were not *a priori* known to be as differentiated across Europe as height, and thus the degree of correlation between effect sizes and PC loadings are much lower in GIANT (e.g. rho = -0.025 for BMI; **Figure 4B-E**).

DISCUSSION

Hispanic/Latinos are a unique population with continental admixture from the Americas, Africa and Europe[11-15] and population of great interest for anthropometric studies. Here, we present results from a large-scale meta-analysis of anthropometric traits in Hispanics/Latinos. As the first of its kind, we have assembled a large sample of Hispanics/Latinos to map a total of six novel loci and 36 novel signals using both Hispanic/Latino population-specific and trans-ethnic discovery efforts (**Figure 1**). More than 1,600 anthropometric-SNP associations were transferable at nominal significance to Hispanics/Latinos—representing between 19-30% of all index sex-combined SNP-anthropometric associations (**Tables S11-13**). Sixty-seven previously reported loci reached GWAS significance at the same index or another lead SNP in our

Hispanic/Latino adult sample (**Table S7**). Moreover, we established that four of seven of our novel HISLA findings were transferable to other ancestral populations at nominal significance. We note that even though these findings provide additional evidence for transferability of common loci for anthropometrics,[88] still a number of previously-reported anthropometric loci may not be transferable to this population in part due to variability in allele frequencies or effect sizes across ancestral populations.[59]

Our conditional and fine-mapping analyses revealed 36 novel signals in established anthropometric loci, which independently replicated in HISLA Stage 2 or the UKBB British subsample. In addition, our lead SNPs for the novel BMI signals discovered at *ADCY5* (from the HISLA meta-analysis) and *ADAMTS9-AS2* (from the trans-ethnic meta-analysis) are both nominally associated with obesity status between 2-18 years of age. Three of our new trans-ethnic signals in established height loci also displayed association with height-for-age z-scores in children/adolescents between 5-18 years of age. These observations support our premise that diverse and trans-ethnic studies represent a valuable tool for identifying multiple signals and fine-mapping in established association regions. This was done with the overarching goal of identifying putative variants that will account for some of the missing heritability of complex diseases and reveal candidate genes and SNPs for functional follow-up.

In light of the notable ancestral, geographical or environmental diversity of the studies analyzed in our meta-analyses, we observed evidence of allele frequency differences for many of our novel discoveries (**Figure 3** and **Figure S53**). Similar to reports from other diverse genome-wide analyses,[59] in some cases this allele frequency heterogeneity may drive the apparent heterogeneity effect across consortia in our HISLA, AAAGC, and GIANT meta-analysis (e.g., *IGF2BP2* I^2=78.7; *MY06* with I^2=84.4, **Tables 2 and S16**). These observations reinforce how studies of one predominant ancestry group, such as Europeans, may fail to identify novel loci or, more likely, new signals in known loci (given how many known loci there are currently) with allele frequency differences across ancestral populations.

Residual uncorrected stratification in GWAS could result in biased estimates of effect sizes.[39] For example, effect sizes on height from GIANT were reported to be significantly correlated with North-South axis of variation in Europe suggesting residual uncorrected stratification,[87; 89; 90] which we also observe here. Note that the residual stratification effect is subtle, and while the effect sizes may be biased, this does not imply the identified associations are spurious. For example, compared with effect sizes on height from UKBB, which is based on a single homogeneous population and results in better control of population stratification, the genetic correlation between GIANT and UKBB was 0.94.[87]

Of the three traits studied here, height is the most stratified in Europe. The correlation coefficient between effect sizes on height and PC loadings reached 0.125 in the GIANT only for PC1, while it was much smaller for other traits (e.g., the maximum |*rho*| = 0.042 in GIANT on WHR using only males on PC1). The decrease in bias in trans-ethnic meta-analysis was also obvious in height. The correlation with PC1 was non-significant in HISLA (*rho* = 0.007) and statistically significant but weak in AAAGC (*rho* = 0.012), consistent with a decreased impact of European population stratification on the estimate of effect size in AAAGC and HISLA. This decreased correlation could be due to large non-European ancestries known in these populations (Africans and Native Americans, respectively) that are less affected by population stratification in Europe; it could also be that by using European ancestry based loadings we are less likely to detect non-European based population stratification patterns or that smaller sample sizes in these cohorts resulting in greater noise in effect size estimates. Regardless of the reason, compared to GIANT alone, trans-ethnic meta-analysis of the three cohorts showed less impact of uncorrected stratification in effect size estimates, even though the sample size in AAAGC and HISLA are comparably small. For other traits, the conclusions are qualitatively similar: that trans-ethnic meta-analysis lessened the bias due to stratification, even though the bias in GIANT was not as strong in the first place.

As described above, in this study we were able to 1) discover six novel loci with a notably smaller analytic size than other anthropometric consortia like GIANT, 2) describe 36 new signals in established loci in HISLA or our trans-ethnic meta-analysis, and 3) generate trans-ethnic effect estimates with better control for population structure. Taken together, these findings indicate the added value of building large, more diverse GWAS in the near future.

Bioinformatic analyses of our population-specific (**Table S19-S20**) and trans-ethnic findings in newly discovered loci give us important insights into the underlying biology of obesity, bone development and growth (**Tables S21-S22**). We described any previously reported associations with other phenotypes, consequences on gene expression, and metabolites in publicly available data, as well as any eQTLs or sQTLs. Lead SNPs from our HISLA only meta-analyses appear relatively benign (not pathogenic) based on CADD and FATHMM-XF scores (**Table S20**). All SNPs potentially change motifs. Both rs17361324 (*ADCY5*) and rs215226 (*B4GALNT3*) have enhancer and promoter histone marks and eQTLs in the respective genes in relevant tissues. For BMI, there is an eQTL for rs17361324-ADCY5 in thyroid, and *ADCY5* has been previously associated with type 2 diabetes,[91] BMI,[92] central obesity traits,[43] height,[51] birth outcomes,[93-95] and a number of other phenotypes. Additionally, rs17361324 is proximate to an *ADCY5* intronic variant (rs1093467, r^2=0.3 in 1000 Genomes AMR) that is highly conserved

across species (Haploreg v4.1). For height, there is an eQTL for rs215226-B4GALNT3 in tibial, aortic, and coronary arteries, and tibial nerve. The lead SNP for the height signal in *B4GALNT3*, rs215226, has enhancer histone marks in bone and muscle, and promoter marks in muscle tissue. In addition, the variant rs215226 (B4GALNT3) has a posterior probability of 1 for causality in FINEMAP analyses (see **Table S9**). Other interesting information about these regions is provided in **Table S19.**

The lead SNPs at our newly discovered trans-ethnic loci were mainly located in intronic and intergenic regions (**Table S21**) and were benign. One exception was the novel locus *C11orf63* associated with height led by rs11605693, which showed pathogenic scores for CADD and FATHMM-XF (CADD score=17.1 and FATHMM-XF score=0.87). This lead SNP has an eQTL in *C11orf63* for adipose, tibial nerve, and testis. *C11orf63*, junctional cadherin complex regulator, is responsible for ependymal cells that line the brain and spinal cord.

Among the trans-ethnic findings, a new signal at a known locus for BMI, rs10540 at *RNH1*, has a posterior probability of 0.82 as one of two causal variants (see **Table S18**) in the locus, and is an eQTL for a wide range of tissues and genes (see **Table S22**), potentially making it relevant to body mass. A new signal in a known locus for height, led by rs12918773 that has a posterior probability of 0.98 and is one of four casual variants suggested from fine-mapping in the locus (see **Table S18**), has an eQTL (in lung, thyroid, tibial nerve and artery, breast, testis) with *CDK10*, a gene also associated with growth retardation.[96] In addition, rs1342330, another new signal in a known height locus, has a low regulomeDB score at 2b and several enhancer and promoter histone marks in relevant tissues (**Tables S21**). As an intronic variant, it is an eQTL in the pancreas with *PHACTR2* (**Tables S22**), a gene associated with body dysmorphic disorder.[97] While many of the novel loci/signals appeared to be benign based on CADD and FATHMM-XF scores, they still show enhancer and promoter histone marks in trait relevant tissues such as adipose, bone, and muscle, thymus, brain, and adrenal gland. Other information and references about the trans-ethnic findings are provided in **Table S21-22.**

Large-scale analyses of diverse populations hold great potential for advancing the field of genetic epidemiology.[59] This study illustrates how studying admixed populations, like Hispanic/Latinos, and leveraging them in trans-ethnic epidemiologic investigations, can yield additional insights into the genetic architecture of anthropometric traits. Future discovery efforts in Hispanic/Latino populations and with other diverse populations will address the research gap between who is studied and who is affected by conditions like obesity, to the benefit of both public health and precision medicine.

Supplemental Data

Supplemental Data include 22 tables and 54 figures.

Declaration of Interests

SMG and AMS receive funding from Seven Bridges Genomics to develop tools for the NHLBI BioData Catalyst consortium. All others authors declare no competing interests.

Acknowledgements

The Baependi Heart Study was supported through a collaborative effort by FAPESP and Brazil Health Ministry (PROADI). ACP was supported by NHLBI R01HL141881-01A1. The Hispanic Community Health Study/Study of Latinos was carried out as a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (N01-HC65233), University of Miami (N01-HC65234), Albert Einstein College of Medicine (N01-HC65235), Northwestern University (N01-HC65236), and San Diego State University (N01-HC65237). The following Institutes/Centers/Offices contribute 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 including (SMS, CCL, AMS) at the University of Washington was supported by NHLBI and NIDCR contracts (HHSN268201300005C AM03 and MOD03). 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, UL1-TR-001420, UL1-TR-001881, and DK063491. Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-64278. Genotyping was performed at Affymetrix (Santa Clara, California, USA) and the Broad Institute of Harvard and MIT (Boston, Massachusetts, USA) using the Affymetrix Genome-Wide Human SNP Array 6.0.

The FIND study was supported by grants U01DK57292, U01DK57329, U01DK057300, U01DK057298, U01DK057249, U01DK57295, U01DK070657, U01DK057303, and U01DK57304 from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and, in part, by the Intramural Research Program of the NIDDK. Support was also received from the National Heart, Lung and Blood Institute grants U01HL065520, U01HL041654, and U01HL041652. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health (NIH), under contract N01-CO-12400 and the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. This work was also supported by the National Center for Research Resources for the General Clinical Research Center grants: Case Western Reserve University, M01-RR-000080; Wake Forest University, M01-RR-07122; Harbor-University of California, Los Angeles Medical Center, M01-RR-00425; College of Medicine, University of California, Irvine, M01-RR-00827–29; University of New Mexico, HSC M01-RR-00997; and Frederic C. Bartter, M01-RR-01346. Computing resources were provided, in part, by the Wake Forest School of Medicine Center for Public Health Genomics. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

The Northern California Breast Cancer Family Registry (NC-BCFR) is supported by grant UM1 CA164920 from the U.S. National Cancer Institute. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the United States Government or the BCFR. The San Francisco Bay Area Breast Cancer Study (SFBCS) was supported by grants R01 CA63446 and R01 CA77305 from the National Cancer Institute, grant DAMD17-96-1-6071 from the U.S. Department of Defense, and grant 7PB-0068 from the California Breast Cancer Research Program.

Several studies were also supported by the National Institutes of Health, National Heart, Lung, and Blood Institute in collaboration with the Mexican-American Coronary Artery Disease Project (MACAD) HL-088457, the HTN-IR study HL-0697974, the Genetics of Latinos Diabetic Retinopathy (GOLDR) study grant EY14684, Hypertriglyceridemia (HyperTG) study contract R01 HL0767711, and DK-079888 (work related to insulin clearance in HTN-IR, MACAD, and NIDDM-Atherosclerosis Study (NIDDM-Athero) funded by the NHLBI. The Mexico City 1, Mexico City 2 and Mexico Kids Case control 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). The SAMAFS was supported by HL045522, DK053889, DK047482, and MH059490. The Santiago Longitudinal Study (SLS) was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development (R01 HD033487-15), National Institute on Drug Abuse (R01 DA021181-05), and National Heart Lung, and Blood Institute (T32 HL079891-11). The Viva La Familia Study was supported by R01DK59264 and R01DK080457.

AGL was supported by T32 HD091058, P2C HD050924, and P30 AG066615. AGT was supported by T32HL007055. ARL is supported by the Leverhulme Trust (F/07 134/DF), BBSRC (BB/I021213/1), the Excellence Initiative of Aix-Marseille University - A*MIDEX (a French "Investissements d'Avenir" programme), the National Natural Science Foundation of China (#31771393), Universidad de Antioquia (CODI sostenibilidad de grupos 2013- 2014 and MASO 2013-2014). The other PIs of the Consortium for the Analysis of the Diversity and Evolution of Latin America (CANDELA) were supported by multiple grants (VA-A: 0051 1 3190000, GB: 0054 (0)280 488-3184, M-CB: 0051 1 3190000, SC-Q: 0056 58 2205073). The investigators of BioME 1 and 2 were supported by the following grants (MG-M: 0052155282325, MP: 0057 320 7034343, CS: 0055 51 999523134, RAJS: 0052153501900). KA was supported by 0044 (0)20 3108 4003. LF was supported by R01CA204797. LF-R was supported by an American Heart Association (AHA) predoctoral grant (13PRE16100015). MG, KLY and KEN were supported by AHA grant 13GRNT16490017 and 15GRNT25880008, R01DK089256 and R01DK101855. In addition, LF-R, CAH, and RFJL were also supported by R01DK101855. XRG was supported by NIH grant R01EY022651. KEN was also supported by R01HD057194, R01DK122503, R01HG010297, R01HL142302, R01HL143885, and R01HG009974. LG was supported by T32 HL129982 and AGT was supported by T32HL007055. QQ was supported by R01HL060712, R01HL140976, and R01DK119268. SFAG was supported by the Daniel B. Burke Endowed Chair for Diabetes Research and R01 HD056465. In addition the above MESA and MESA SHARe funding, XG, MAA, Y-DIC, JY, and JIR were supported by EY14684, HL-0767711, HL-0697974, HL-088457, NEI EY11753, UL1-TR-001881. XL and KR were supported by R01 HL0767711 and DK-079888. ZA, TAB, JT, and AHX were supported by HTN-IR funding (HL-0697974); PMG, YH, EI, and KDT were supported by GOLDR funding (EY-14684); MOG, WH, KL, and KS were supported by MACAD funding (HL-088457). AEJ was supported in part by

K99/R00 HL130580. EMJ was supported by R01 CA063446, R01 CA077305, DOD RP9590546, CBCRP 7PB-0068. JM was supported by American Diabetes Association Innovative and Clinical Translational Award 1-19-ICTS-068 and by the National Human Genome Research Institute (U01HG011723). Computation for part of this work was supported by the University of Southern California's Center for Advanced Research Computing (https://carc.usc.edu).

References

1. . World Health Organization. (2020). Obesity and overweight, https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight.
2. Duran, P., Caballero, B., and de Onis, M. (2006). The association between stunting and overweight in Latin American and Caribbean preschool children. Food Nutr Bull 27, 300-305.
3. Organización de las Naciones Unidas para la Alimentación y la Agricultura (2017). América Latina y el Caribe: Panorama de la seguridad alimentaria y nutricional. Sistemas alimentarios sostenibles para poner fin al hambre y la malnutrición, 2016, http://iris.paho.org/xmlui/handle/123456789/33680.
4. Mueller, M., Purnell, T.S., Mensah, G.A., and Cooper, L.A. (2015). Reducing racial and ethnic disparities in hypertension prevention and control: what will it take to translate research into practice and policy? Am J Hypertens 28, 699-716.
5. Lanas, F., Bazzano, L., Rubinstein, A., Calandrelli, M., Chen, C.S., Elorriaga, N., Gutierrez, L., Manfredi, J.A., Seron, P., Mores, N., et al. (2016). Prevalence, Distributions and Determinants of Obesity and Central Obesity in the Southern Cone of America. PLoS One 11, e0163727.
6. Gravel, S., Zakharia, F., Moreno-Estrada, A., Byrnes, J.K., Muzzio, M., Rodriguez-Flores, J.L., Kenny, E.E., Gignoux, C.R., Maples, B.K., Guiblet, W., et al. (2013). Reconstructing Native American migrations from whole-genome and whole-exome data. PLoS Genet 9, e1004023.
7. Moreno-Estrada, A., Gravel, S., Zakharia, F., McCauley, J.L., Byrnes, J.K., Gignoux, C.R., Ortiz-Tello, P.A., Martinez, R.J., Hedges, D.J., Morris, R.W., et al. (2013). Reconstructing the population genetic history of the Caribbean. PLoS Genet 9, e1003925.
8. Gonzalez Burchard, E., Borrell, L.N., Choudhry, S., Naqvi, M., Tsai, H.J., Rodriguez-Santana, J.R., Chapela, R., Rogers, S.D., Mei, R., Rodriguez-Cintron, W., et al. (2005). Latino populations: a unique opportunity for the study of race, genetics, and social environment in epidemiological research. Am J Public Health 95, 2161-2168.
9. Moreno-Estrada, A., Gignoux, C.R., Fernandez-Lopez, J.C., Zakharia, F., Sikora, M., Contreras, A.V., Acuna-Alonzo, V., Sandoval, K., Eng, C., Romero-Hidalgo, S., et al. (2014). Human genetics. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. Science 344, 1280-1285.
10. Conomos, M.P., Laurie, C.A., Stilp, A.M., Gogarten, S.M., McHugh, C.P., Nelson, S.C., Sofer, T., Fernandez-Rhodes, L., Justice, A.E., Graff, M., et al. (2016). Genetic Diversity and Association Studies in US Hispanic/Latino Populations: Applications in the Hispanic Community Health Study/Study of Latinos. Am J Hum Genet 98, 165-184.
11. Bryc, K., Velez, C., Karafet, T., Moreno-Estrada, A., Reynolds, A., Auton, A., Hammer, M., Bustamante, C.D., and Ostrer, H. (2010). Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations. Proc Natl Acad Sci U S A 107 Suppl 2, 8954-8961.
12. Halder, I., Yang, B.Z., Kranzler, H.R., Stein, M.B., Shriver, M.D., and Gelernter, J. (2009). Measurement of admixture proportions and description of admixture structure in different U.S. populations. Hum Mutat 30, 1299-1309.

13. Shtir, C.J., Marjoram, P., Azen, S., Conti, D.V., Le Marchand, L., Haiman, C.A., and Varma, R. (2009). Variation in genetic admixture and population structure among Latinos: the Los Angeles Latino eye study (LALES). BMC Genet 10, 71.
14. Wang, S., Ray, N., Rojas, W., Parra, M.V., Bedoya, G., Gallo, C., Poletti, G., Mazzotti, G., Hill, K., Hurtado, A.M., et al. (2008). Geographic patterns of genome admixture in Latin American Mestizos. PLoS Genet 4, e1000037.
15. Manichaikul, A., Palmas, W., Rodriguez, C.J., Peralta, C.A., Divers, J., Guo, X., Chen, W.M., Wong, Q., Williams, K., Kerr, K.F., et al. (2012). Population structure of Hispanics in the United States: the multi-ethnic study of atherosclerosis. PLoS Genet 8, e1002640.
16. Chiang, C.W., Gajdos, Z.K., Korn, J.M., Kuruvilla, F.G., Butler, J.L., Hackett, R., Guiducci, C., Nguyen, T.T., Wilks, R., Forrester, T., et al. (2010). Rapid assessment of genetic ancestry in populations of unknown origin by genome-wide genotyping of pooled samples. PLoS Genet 6, e1000866.
17. Spear, M.L., Diaz-Papkovich, A., Ziv, E., Yracheta, J.M., Gravel, S., Torgerson, D.G., and Hernandez, R.D. (2020). Recent shifts in the genomic ancestry of Mexican Americans may alter the genetic architecture of biomedical traits. Elife 9.
18. Ruiz-Linares, A., Adhikari, K., Acuna-Alonzo, V., Quinto-Sanchez, M., Jaramillo, C., Arias, W., Fuentes, M., Pizarro, M., Everardo, P., de Avila, F., et al. (2014). Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. PLoS Genet 10, e1004572.
19. Asgari, S., Luo, Y., Akbari, A., Belbin, G.M., Li, X., Harris, D.N., Selig, M., Bartell, E., Calderon, R., Slowikowski, K., et al. (2020). A positively selected FBN1 missense variant reduces height in Peruvian individuals. Nature 582, 234-239.
20. Lindsay, C. (2007). The Latin American Community in Canada. In. (Profiles of Ethnic Communities in Canada (Statistics Canada).
21. Statistics Canada. (2017). Focus on Geography Series, 2016 Census. In. (Statistics Canada Catalogue).
22. Center, P.R. (2015). Modern Immigration Wave Brings 59 Million to U.S., Driving Population Growth and Change Through 2065. In. (
23. Zamosky, L. (2013). The obesity epidemic. While America swallows $147 billion in obesity-related healthcare costs, physicians called on to confront the crisis. Medical economics 90, 14-17.
24. Wang, Y., Beydoun, M.A., Liang, L., Caballero, B., and Kumanyika, S.K. (2008). Will all Americans become overweight or obese? estimating the progression and cost of the US obesity epidemic. Obesity 16, 2323-2330.
25. Ogden, C.L., Carroll, M.D., Kit, B.K., and Flegal, K.M. (2012). Prevalence of Obesity and Trends in Body Mass Index Among US Children and Adolescents, 1999-2010. JAMA.
26. 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.
27. Ogden, C.L., Carroll, M.D., Kit, B.K., and Flegal, K.M. (2014). Prevalence of childhood and adult obesity in the United States, 2011-2012. Jama 311, 806-814.

28. Isasi, C.R., Ayala, G.X., Sotres-Alvarez, D., Madanat, H., Penedo, F., Loria, C.M., Elder, J.P., Daviglus, M.L., Barnhart, J., Siega-Riz, A.M., et al. (2015). Is acculturation related to obesity in Hispanic/Latino adults? Results from the Hispanic community health study/study of Latinos. J Obes 2015, 186276.
29. NCD Risk Factor Collaboration. (2020). Height and body-mass index trajectories of school-aged children and adolescents from 1985 to 2019 in 200 countries and territories: a pooled analysis of 2181 population-based studies with 65 million participants. Lancet 396, 1511-1524.
30. Knowler, W.C., Coresh, J., Elston, R.C., Freedman, B.I., Iyengar, S.K., Kimmel, P.L., Olson, J.M., Plaetke, R., Sedor, J.R., Seldin, M.F., et al. (2005). The Family Investigation of Nephropathy and Diabetes (FIND): design and methods. J Diabetes Complications 19, 1-9.
31. Centers for Disease Control and Prevention. (2016). CDC Growth Charts. In. (
32. Bradfield, J.P., Vogelezang, S., Felix, J.F., Chesi, A., Helgeland, O., Horikoshi, M., Karhunen, V., Lowry, E., Cousminer, D.L., Ahluwalia, T.S., et al. (2019). A trans-ancestral meta-analysis of genome-wide association studies reveals loci associated with childhood obesity. Hum Mol Genet 28, 3327-3338.
33. World Health Organization. Growth reference data for 5-19 years. In. (
34. Butte, N.F., Cai, G., Cole, S.A., and Comuzzie, A.G. (2006). Viva la Familia Study: genetic and environmental contributions to childhood obesity and its comorbidities in the Hispanic population. Am J Clin Nutr 84, 646-654; quiz 673-644.
35. Lin, D.Y., Tao, R., Kalsbeek, W.D., Zeng, D.L., Gonzalez, F., Fernandez-Rhodes, L., Graff, M., Koch, G.G., North, K.E., and Heiss, G. (2014). Genetic Association Analysis under Complex Survey Sampling: The Hispanic Community Health Study/Study of Latinos. Am J Hum Genet 95, 675-688.
36. Chen, M.H., and Yang, Q. (2010). GWAF: an R package for genome-wide association analyses with family data. Bioinformatics 26, 580-581.
37. Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26, 2190-2191.
38. Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., et al. (2015). Genetic studies of body mass index yield new insights for obesity biology. Nature 518, 197-206.
39. Wood, A.R., Esko, T., Yang, J., Vedantam, S., Pers, T.H., Gustafsson, S., Chu, A.Y., Estrada, K., Luan, J., Kutalik, Z., et al. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 46, 1173-1186.
40. Shungin, D., Winkler, T.W., Croteau-Chonka, D.C., Ferreira, T., Locke, A.E., Magi, R., Strawbridge, R.J., Pers, T.H., Fischer, K., Justice, A.E., et al. (2015). New genetic loci link adipose and insulin biology to body fat distribution. Nature 518, 187-196.
41. Ng, M.C.Y., Graff, M., Lu, Y., Justice, A.E., Mudgal, P., Liu, C.T., Young, K., Yanek, L.R., Feitosa, M.F., Wojczynski, M.K., et al. (2017). Discovery and fine-mapping of adiposity loci using high density imputation of genome-wide association studies in individuals of African ancestry: African Ancestry Anthropometry Genetics Consortium. PLoS Genet 13, e1006719.

42. 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. Hum Mol Genet 27, 3641-3649.
43. Pulit, S.L., Stoneman, C., Morris, A.P., Wood, A.R., Glastonbury, C.A., Tyrrell, J., Yengo, L., Ferreira, T., Marouli, E., Ji, Y., et al. (2019). Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet 28, 166-174.
44. Akiyama, M., Okada, Y., Kanai, M., Takahashi, A., Momozawa, Y., Ikeda, M., Iwata, N., Ikegawa, S., Hirata, M., Matsuda, K., et al. (2017). Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. Nat Genet 49, 1458-1467.
45. Berndt, S.I., Gustafsson, S., Magi, R., Ganna, A., Wheeler, E., Feitosa, M.F., Justice, A.E., Monda, K.L., Croteau-Chonka, D.C., Day, F.R., et al. (2013). Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. Nat Genet 45, 501-512.
46. Gong, J., Nishimura, K.K., Fernandez-Rhodes, L., Haessler, J., Bien, S., Graff, M., Lim, U., Lu, Y., Gross, M., Fornage, M., et al. (2018). Trans-ethnic analysis of metabochip data identifies two new loci associated with BMI. Int J Obes (Lond) 42, 384-390.
47. Gong, J., Schumacher, F., Lim, U., Hindorff, L.A., Haessler, J., Buyske, S., Carlson, C.S., Rosse, S., Buzkova, P., Fornage, M., et al. (2013). Fine Mapping and Identification of BMI Loci in African Americans. Am J Hum Genet 93, 661-671.
48. Graff, M., Scott, R.A., Justice, A.E., Young, K.L., Feitosa, M.F., Barata, L., Winkler, T.W., Chu, A.Y., Mahajan, A., Hadley, D., et al. (2017). Genome-wide physical activity interactions in adiposity - A meta-analysis of 200,452 adults. PLoS Genet 13, e1006528.
49. Hoffmann, T.J., Choquet, H., Yin, J., Banda, Y., Kvale, M.N., Glymour, M., Schaefer, C., Risch, N., and Jorgenson, E. (2018). A Large Multiethnic Genome-Wide Association Study of Adult Body Mass Index Identifies Novel Loci. Genetics 210, 499-515.
50. Justice, A.E., Winkler, T.W., Feitosa, M.F., Graff, M., Fisher, V.A., Young, K., Barata, L., Deng, X., Czajkowski, J., Hadley, D., et al. (2017). Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. Nat Commun 8, 14977.
51. Kichaev, G., Bhatia, G., Loh, P.R., Gazal, S., Burch, K., Freund, M.K., Schoech, A., Pasaniuc, B., and Price, A.L. (2019). Leveraging Polygenic Functional Enrichment to Improve GWAS Power. Am J Hum Genet 104, 65-75.
52. Monda, K.L., Chen, G.K., Taylor, K.C., Palmer, C., Edwards, T.L., Lange, L.A., Ng, M.C., Adeyemo, A.A., Allison, M.A., Bielak, L.F., et al. (2013). A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. Nat Genet 45, 690-696.
53. Okada, Y., Kubo, M., Ohmiya, H., Takahashi, A., Kumasaka, N., Hosono, N., Maeda, S., Wen, W., Dorajoo, R., Go, M.J., et al. (2012). Common variants at

CDKAL1 and KLF9 are associated with body mass index in east Asian populations. Nat Genet 44, 302-306.
54. Tachmazidou, I., Suveges, D., Min, J.L., Ritchie, G.R.S., Steinberg, J., Walter, K., Iotchkova, V., Schwartzentruber, J., Huang, J., Memari, Y., et al. (2017). Whole-Genome Sequencing Coupled to Imputation Discovers Genetic Signals for Anthropometric Traits. Am J Hum Genet 100, 865-884.
55. Turcot, V., Lu, Y., Highland, H.M., Schurmann, C., Justice, A.E., Fine, R.S., Bradfield, J.P., Esko, T., Giri, A., Graff, M., et al. (2018). Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. Nat Genet 50, 26-41.
56. Wen, W., Cho, Y.S., Zheng, W., Dorajoo, R., Kato, N., Qi, L., Chen, C.H., Delahanty, R.J., Okada, Y., Tabara, Y., et al. (2012). Meta-analysis identifies common variants associated with body mass index in east Asians. Nat Genet 44, 307-311.
57. Wen, W., Zheng, W., Okada, Y., Takeuchi, F., Tabara, Y., Hwang, J.Y., Dorajoo, R., Li, H., Tsai, F.J., Yang, X., et al. (2014). Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. Hum Mol Genet 23, 5492-5504.
58. Winkler, T.W., Justice, A.E., Graff, M., Barata, L., Feitosa, M.F., Chu, S., Czajkowski, J., Esko, T., Fall, T., Kilpelainen, T.O., et al. (2015). The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. PLoS Genet 11, e1005378.
59. Wojcik, G.L., Graff, M., Nishimura, K.K., Tao, R., Haessler, J., Gignoux, C.R., Highland, H.M., Patel, Y.M., Sorokin, E.P., Avery, C.L., et al. (2019). Genetic analyses of diverse populations improves discovery for complex traits. Nature 570, 514-518.
60. Marouli, E., Graff, M., Medina-Gomez, C., Lo, K.S., Wood, A.R., Kjaer, T.R., Fine, R.S., Lu, Y., Schurmann, C., Highland, H.M., et al. (2017). Rare and low-frequency coding variants alter human adult height. Nature 542, 186-190.
61. 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., et al. (2009). A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nat Genet 41, 527-534.
62. Heid, I.M., Jackson, A.U., Randall, J.C., Winkler, T.W., Qi, L., Steinthorsdottir, V., Thorleifsson, G., Zillikens, M.C., Speliotes, E.K., Magi, R., et al. (2010). Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nat Genet 42, 949-960.
63. Justice, A.E., Karaderi, T., Highland, H.M., Young, K.L., Graff, M., Lu, Y., Turcot, V., Auer, P.L., Fine, R.S., Guo, X., et al. (2019). Protein-coding variants implicate novel genes related to lipid homeostasis contributing to body-fat distribution. Nat Genet 51, 452-469.
64. Scuteri, A., Sanna, S., Chen, W.M., Uda, M., Albai, G., Strait, J., Najjar, S., Nagaraja, R., Orru, M., Usala, G., et al. (2007). Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genet 3, e115.

65. Wen, W., Kato, N., Hwang, J.Y., Guo, X., Tabara, Y., Li, H., Dorajoo, R., Yang, X., Tsai, F.J., Li, S., et al. (2016). Genome-wide association studies in East Asians identify new loci for waist-hip ratio and waist circumference. Sci Rep 6, 17958.
66. Benner, C., Spencer, C.C., Havulinna, A.S., Salomaa, V., Ripatti, S., and Pirinen, M. (2016). FINEMAP: efficient variable selection using summary data from genome-wide association studies. Bioinformatics 32, 1493-1501.
67. Justice, A.E., Karaderi, T., Highland, H.M., Young, K.L., Graff, M., Lu, Y., Turcot, V., Auer, P.L., Fine, R.S., Guo, X., et al. (2019). Protein-coding variants implicate novel genes related to lipid homeostasis contributing to body-fat distribution. Nat Genet 51, 452-469.
68. 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.
69. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 22, 1790-1797.
70. Ward, L.D., and Kellis, M. (2016). HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. Nucleic Acids Res 44, D877-881.
71. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res 12, 996-1006.
72. Consortium, G.T. (2013). The Genotype-Tissue Expression (GTEx) project. Nature genetics 45, 580-585.
73. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, 7.
74. Price, A.L., Weale, M.E., Patterson, N., Myers, S.R., Need, A.C., Shianna, K.V., Ge, D., Rotter, J.I., Torres, E., Taylor, K.D., et al. (2008). Long-range LD can confound genome scans in admixed populations. Am J Hum Genet 83, 132-135; author reply 135-139.
75. Paternoster, L., Zhurov, A.I., Toma, A.M., Kemp, J.P., St Pourcain, B., Timpson, N.J., McMahon, G., McArdle, W., Ring, S.M., Smith, G.D., et al. (2012). Genome-wide association study of three-dimensional facial morphology identifies a variant in PAX3 associated with nasion position. Am J Hum Genet 90, 478-485.
76. Adhikari, K., Fuentes-Guajardo, M., Quinto-Sanchez, M., Mendoza-Revilla, J., Camilo Chacon-Duque, J., Acuna-Alonzo, V., Jaramillo, C., Arias, W., Lozano, R.B., Perez, G.M., et al. (2016). A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. Nat Commun 7, 11616.
77. Xiong, Z., Dankova, G., Howe, L.J., Lee, M.K., Hysi, P.G., de Jong, M.A., Zhu, G., Adhikari, K., Li, D., Li, Y., et al. (2019). Novel genetic loci affecting facial shape variation in humans. Elife 8.

78. Yap, C.X., Sidorenko, J., Wu, Y., Kemper, K.E., Yang, J., Wray, N.R., Robinson, M.R., and Visscher, P.M. (2018). Dissection of genetic variation and evidence for pleiotropy in male pattern baldness. Nat Commun 9, 5407.
79. Hagenaars, S.P., Hill, W.D., Harris, S.E., Ritchie, S.J., Davies, G., Liewald, D.C., Gale, C.R., Porteous, D.J., Deary, I.J., and Marioni, R.E. (2017). Genetic prediction of male pattern baldness. PLoS Genet 13, e1006594.
80. B.L., N., and P.A., L. (2010). Pax Genes. In Transcriptional Control of Neural Crest Development. (San Rafael (CA), Morgan & Claypool Life Sciences.
81. Shakhova O., and L., S. (2010). Neural crest-derived stem cells. In. (Cambridge, MA, Harvard Stem Cell Institute.
82. Lee, G., Kim, H., Elkabetz, Y., Al Shamy, G., Panagiotakos, G., Barberi, T., Tabar, V., and Studer, L. (2007). Isolation and directed differentiation of neural crest stem cells derived from human embryonic stem cells. Nat Biotechnol 25, 1468-1475.
83. Billon, N., Iannarelli, P., Monteiro, M.C., Glavieux-Pardanaud, C., Richardson, W.D., Kessaris, N., Dani, C., and Dupin, E. (2007). The generation of adipocytes by the neural crest. Development 134, 2283-2292.
84. Hans, C., Dobra, A., and West, M. (2007). Shotgun Stochastic Search for “Large p” Regression. Journal of the American Statistical Association 102, 507-516.
85. Campbell, C.D., Ogburn, E.L., Lunetta, K.L., Lyon, H.N., Freedman, M.L., Groop, L.C., Altshuler, D., Ardlie, K.G., and Hirschhorn, J.N. (2005). Demonstrating stratification in a European American population. Nat Genet 37, 868-872.
86. Robinson, M.R., Hemani, G., Medina-Gomez, C., Mezzavilla, M., Esko, T., Shakhbazov, K., Powell, J.E., Vinkhuyzen, A., Berndt, S.I., Gustafsson, S., et al. (2015). Population genetic differentiation of height and body mass index across Europe. Nat Genet 47, 1357-1362.
87. Sohail, M., Maier, R.M., Ganna, A., Bloemendal, A., Martin, A.R., Turchin, M.C., Chiang, C.W., Hirschhorn, J., Daly, M.J., Patterson, N., et al. (2019). Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. Elife 8.
88. Carlson, C.S., Matise, T.C., North, K.E., Haiman, C.A., Fesinmeyer, M.D., Buyske, S., Schumacher, F.R., Peters, U., Franceschini, N., Ritchie, M.D., et al. (2013). Generalization and dilution of association results from European GWAS in populations of non-European ancestry: the PAGE study. PLoS Biol 11, e1001661.
89. Berg, J.J., Harpak, A., Sinnott-Armstrong, N., Joergensen, A.M., Mostafavi, H., Field, Y., Boyle, E.A., Zhang, X., Racimo, F., Pritchard, J.K., et al. (2019). Reduced signal for polygenic adaptation of height in UK Biobank. Elife 8.
90. Chen, M., Sidore, C., Akiyama, M., Ishigaki, K., Kamatani, Y., Schlessinger, D., Cucca, F., Okada, Y., and Chiang, C.W.K. (2020). Evidence of Polygenic Adaptation in Sardinia at Height-Associated Loci Ascertained from the Biobank Japan. Am J Hum Genet 107, 60-71.
91. Vujkovic, M., Keaton, J.M., Lynch, J.A., Miller, D.R., Zhou, J., Tcheandjieu, C., Huffman, J.E., Assimes, T.L., Lorenz, K., Zhu, X., et al. (2020). Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. Nat Genet 52, 680-691.

92. Zhu, Z., Guo, Y., Shi, H., Liu, C.L., Panganiban, R.A., Chung, W., O'Connor, L.J., Himes, B.E., Gazal, S., Hasegawa, K., et al. (2020). Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. J Allergy Clin Immunol 145, 537-549.
93. Freathy, R.M., Mook-Kanamori, D.O., Sovio, U., Prokopenko, I., Timpson, N.J., Berry, D.J., Warrington, N.M., Widen, E., Hottenga, J.J., Kaakinen, M., et al. (2010). Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. Nat Genet 42, 430-435.
94. Zhang, G., Feenstra, B., Bacelis, J., Liu, X., Muglia, L.M., Juodakis, J., Miller, D.E., Litterman, N., Jiang, P.P., Russell, L., et al. (2017). Genetic Associations with Gestational Duration and Spontaneous Preterm Birth. N Engl J Med 377, 1156-1167.
95. Yang, X.L., Zhang, S.Y., Zhang, H., Wei, X.T., Feng, G.J., Pei, Y.F., and Zhang, L. (2019). Three Novel Loci for Infant Head Circumference Identified by a Joint Association Analysis. Front Genet 10, 947.
96. Windpassinger, C., Piard, J., Bonnard, C., Alfadhel, M., Lim, S., Bisteau, X., Blouin, S., Ali, N.B., Ng, A.Y.J., Lu, H., et al. (2017). CDK10 Mutations in Humans and Mice Cause Severe Growth Retardation, Spine Malformations, and Developmental Delays. Am J Hum Genet 101, 391-403.
97. Daelemans, C., Ritchie, M.E., Smits, G., Abu-Amero, S., Sudbery, I.M., Forrest, M.S., Campino, S., Clark, T.G., Stanier, P., Kwiatkowski, D., et al. (2010). High-throughput analysis of candidate imprinted genes and allele-specific gene expression in the human term placenta. BMC Genet 11, 25.

Figure Titles and Legends

Figure 1. Flowchart of the design and discovery of 6 novel loci and 36 novel signals in known loci in the Hispanic/Latino Anthropometry Consortium (HISLA) Meta-Analysis and the Trans-Ethnic Meta-Analysis of HISLA and Consortia of Other Ancestral Heritages

*Stage 1 maximum sample sizes varied between and 59,771 for BMI, 56,161 for height, and 42,455 for WHRadjBMI (sex combined). **Stage 2 sample sizes varied between 10,538 for BMI, 8,110 for height, and 4,393 for WHRadjBMI (sex combined). Actual sample sizes may vary by SNP. ***The BMI and height-for-age z-score models were conducted using up to 1,914 and 1,945 of children/adolescents, respectively. In contrast, the obesity case-control study compared up to 1,814 children/adolescents who were either ≥95th versus ≤50th BMI-for-age percentiles

Figure 2. Regional plot, unconditioned (A) and conditioned (B), of the novel PAX3 locus associated with body mass index (BMI) in the Hispanic/Latino Anthropometry Consortium (HISLA), excluding Brazilian and Native American samples. Linkage disequilibrium patterns are based on rs994108 (shown by the purple triangle) from the Hispanic Communities in Health Study/Study of Latinos.

Figure 3. Variability in HISLA Stage 1+2, GIANT, and AAAGC P-values, Effect Sizes and Risk Allele Frequencies.

Hispanic/Latino Anthropometry Consortium (HISLA); African American Anthropometry Genetics Consortium (AAAGC); Genetic Investigation of ANthropometric Traits (GIANT); WHRadjBMI - waist to hip ratio adjusted for BMI. *Asterisks indicating a SNPs that were significant either as a novel locus or new signals in a known locus.

Figure 4. Correlations (rho) between effect estimates and the loadings of the principal components 1-5 in each consortia (HISLA, AAAGC, GIANT) and the meta-analysis of all 3 consortia (Meta) by trait. (A) height, (B) BMI, (C) Waist-to-hip ratio adjusted for BMI (WHRadjBMI) for men and women combined, (D) WHRadjBMI for women only, (E) WHRadjBMI for men only.

Hispanic/Latino Anthropometry Consortium (HISLA); African American Anthropometry Genetics Consortium (AAAGC); Genetic Investigation of ANthropometric Traits (GIANT); WHRadjBMI - waist to hip ratio adjusted for BMI

Table Titles and Legends

Table 1. Potential novel loci and new signals in known loci from the Stage 1: Adult HISLA Discovery combined with results from the Stage 2: Adult HISLA Validation.[1] In addition, lookup of results of each locus from the AAAGC and GIANT.
Abbreviations: Chr - chromosome; EAF - effect allele frequency; HetIsq - heterogeneity I-square; N - sample size; WHRadjBMI - waist to hip ratio adjusted for BMI; AAAGC- African American Anthropometry Genetics Consotrium; GIANT- Genetic Investigation of ANthropometric Traits [1] All studies were meta-analyzed using METAL (PMID 20616382), with each study entered individuals into Stage 1+2 analyses.[2] These BMI and WHRadjBMI analyses did not include Brazilian and/or Native American samples. [3] New loci or signals are those that

were validated by HISLA stage 2 results that are directionally consistent with Stage 1 and remaining genome-wide significant after meta-analysis with Stage 1.
[4] Proxy GIANT, rs1573905 (r2= 0.96 AMR)

Table 2. Novel loci and new signals in established loci by trait from a meta-analyses of HISLA, AAAGC, and GIANT.
Abbreviations: Chr - chromosome; EAF - effect allele frequency; HetIsq - heterogeneity I-square; N - sample size; WHRadjBMI - waist to hip ratio adjusted for BMI; AAAGC- African American Anthropometry Genetics Consotrium; GIANT- Genetic Investigation of ANthropometric Traits [1] Each novel locus was defined by the absence of known (previously published) SNPs within 1Mb (+/-500 Kb) of the lead SNP.
[2] Each known locus was defined by a 1Mb region around previously identified SNP(s) for the indicated trait; the known SNP(s), P<5e-8, at each established locus can be found in Table S16.

Table 1. Potential novel loci and new signals in known loci from the Stage 1: Adult HISLA Discovery combined with results from the Stage 2: Adult HISLA Validation.[1] In addition, lookup of results of each locus from the AAAGC and GIANT.

Trait	Locus Name	SNP rsid	Genomic region	Chr	Position (hg19)	Effect/ Other Alleles	Stage	EAF	Beta	SE	P-value	HetISq	N	Yes/ No
Novel loci														
BMI[2]	*PAX3*	rs994108	intergenic	2	223057288	C/A	Stage 1: Discovery	0.390	0.041	0.007	**1.62E-08**	0	43048	**Yes**
							Stage 2: Validation	0.394	0.030	0.016	5.65E-02	8	9336	
							Stage 1 + 2	0.389	0.038	0.006	**2.19E-09**	0	52384	
							AAAGC	0.526	0.007	0.007	3.26E-01	0	42751	
							GIANT	0.342	0.0001	0.004	9.81E-01	-	233955	
BMI	*ARRDC3*	rs1505851	intronic	5	90893954	T/C	Stage 1: Discovery	0.741	0.041	0.007	**2.287E-08**	0	52365	No
							Stage 2: Validation	0.709	0.005	0.017	7.62E-01	33.3	9336	
							Stage 1 + 2	0.735	0.035	0.007	1.16E-07	14.1	61701	
							AAAGC	0.307	0.027	0.008	7.00E-04	46.6	42752	
							GIANT	0.680	0	0.004	7.90E-01	-	233999	
WHRadjBMI - Women only[2]	*DOCK2*	rs6879439	intronic	5	169314869	C/T	Stage 1: Discovery	0.520	0.060	0.010	**1.02E-08**	0	18591	No
							Stage 2: Validation	0.526	0.013	0.028	6.54E-01	28.7	2747	
							Stage 1 + 2	0.515	0.049	0.0093	1.57E-07	1.9	23382	
							AAAGC	0.440	0.012	0.012	3.09E-01	0	15600	
							GIANT	0.610	0.0025	0.005	6.30E-01	-	86317	
WHRadjBMI - Sex combined	*TAOK3*	rs115981023	intronic	12	118751105	A/G	Stage 1: Discovery	0.009	0.328	0.057	**1.08E-08**	44.8	19640	No
							Stage 2: Validation	0.004	-0.339	0.687	6.22E-01	0	1340	
							Stage 1 + 2	0.009	0.308	0.057	5.18E-08	52	20980	
							AAAGC	0.050	0.027	0.027	3.07E-01	0	15601	
							GIANT	0.002	no proxy					
New signals in known loci														
BMI	*ADCY5*	rs17361324	intronic	3	123131254	T/C	Stage 1: Discovery	0.280	0.042	0.008	**2.60E-08**	0	43333	**Yes**
							Stage 2: Validation	0.269	0.035	0.018	4.70E-02	0	9035	
							Stage 1 + 2	0.278	0.041	0.007	**2.84E-09**	0	52368	
							AAAGC	0.119	0.023	0.011	3.85E-02	0	42682	
							GIANT	0.253	0.013	0.004	9.90E-04	-	320704	
	C6orf106	rs148899910	intergenic	6	34232259	C/G	Stage 1: Discovery	0.275	0.040	0.007	**9.03E-09**	0	54105	**Yes**
							Stage 2: Validation	0.282	0.049	0.017	4.43E-03	0	9035	
							Stage 1 + 2	0.276	0.041	0.006	**1.24E-10**	0	63140	
							AAAGC	0.316	-0.016	0.008	5.02E-02	30	42750	
							GIANT[4]	0.017	0.036	0.012	3.99E-03	-	216522	
Height	*B4GALNT3*	rs215226	intronic	12	591300	A/G	Stage 1: Discovery	**0.550**	**-0.032**	**0.005**	**5.53E-09**	22.1	52156	**Yes**
							Stage 2: Validation	0.565	-0.020	0.017	2.37E-01	19.3	6906	
							Stage 1 + 2	0.551	-0.031	0.005	**1.98E-09**	21.8	59062	
							AAAGC	0.772	-0.031	0.009	8.99E-04	24	41327	
							GIANT	0.633	0.006	0.004	1.10E-01	-	220370	

Abbreviations: Chr - chromosome; EAF - effect allele frequency; HetIsq - heterogeneity I-square; N - sample size; WHRadjBMI - waist to hip ratio adjusted for BMI; African American Anthropometry Genetics Consotrium AAAGC; Genetic Investigation of ANthropometric Traits (GIANT)

1 All studies were meta-analyzed using METAL (Willer et al 2010 PMID 20616382), with each study entered individuals into Stage 1+2 analyses.

2 These BMI and WHRadjBMI analyses did not include Brazilian and/or Native American samples.

3 New loci or signals are those that were validated by HISLA stage 2 results that are directionally consistent with Stage 1 and remaining genome-wide significant after meta-analysis with Stage 1.

4 Proxy GIANT, rs1573905 (r2= 0.96 AMR)

Table 2. Novel loci and new signals in established loci by trait from a meta-analyses of HISLA, AAAGC, and GIANT.

							Unconditioned meta-analysis results					Conditioned on all known SNPs within 10Mb region			UKBB (Validation results)				
Trait	SNP rsid	Chr	Position (hg19)	Locus Name	Effect/ Other Alleles	EAF	N	Beta	SE	P-value	HetISq	EAF in LD data	Beta	SE	P-value	EAF	N	Beta	SE	P-value
Novel loci [1]																				
BMI	rs4675117	2	227769794	*RHBDD1*	T/C	0.374	343628	0.017	0.003	8.56E-08	0	0.348	0.019	0.003	2.23E-09	0.383	336107	0.006	0.002	1.82E-02
BMI	rs9860730	3	64701146	*ADAMTS9-AS2*	A/G	0.642	428763	-0.016	0.003	1.67E-08	0	0.596	-0.015	0.003	3.80E-08	0.712	336107	-0.008	0.003	4.54E-03
BMI	rs150992	5	98275197	*CTD-2007H13.3*	A/G	0.702	439077	0.018	0.003	5.40E-10	0	0.709	0.017	0.003	1.02E-08	0.693	336107	0.005	0.003	3.74E-02
Height	rs17375290	1	61334177	*NFIA*	A/G	0.806	364636	0.017	0.003	3.47E-08	0	0.801	0.017	0.003	3.00E-08	0.794	336474	0.002	0.002	4.58E-01
Height	rs10737541	1	168214098	*ANKRD36BP1*	T/G	0.286	319809	-0.018	0.003	1.60E-09	0	0.320	-0.018	0.003	3.16E-10	0.226	336474	-0.004	0.002	4.37E-02
Height	rs4618485	6	73555917	*KCNQ5*	A/G	0.634	348626	0.014	0.003	4.72E-08	39.20	0.637	0.018	0.003	4.34E-12	0.604	336474	0.003	0.002	7.02E-02
Height	rs17493997	8	82044302	*PAG1*	C/G	0.360	325906	-0.015	0.003	3.42E-08	0	0.273	-0.017	0.003	2.73E-10	0.299	336474	-0.0003	0.002	8.56E-01
WHRadjBMI sex-combined	rs16873543	6	45577134	*RUNX2*	T/C	0.703	209552	-0.018	0.004	3.20E-06	0	0.684	-0.022	0.004	9.65E-09	0.724	484563	-0.008	0.002	5.50E-04
New signals in established loci [2]																				
BMI	rs10540	11	494662	*RNH1*	A/G	0.116	470714	-0.021	0.004	1.01E-07	0	0.131	-0.023	0.004	5.75E-09	0.135	336107	-0.007	0.004	4.16E-02
BMI	rs4807179	19	1956035	*CSNK1G2*	A/G	0.484	309507	0.020	0.003	2.75E-10	0	0.523	0.018	0.003	1.52E-08	0.632	336107	0.014	0.002	1.06E-08
BMI	rs4813428	20	21451848	*NKX2-2*	T/C	0.132	321797	0.029	0.005	2.89E-10	0	0.095	0.029	0.005	1.46E-10	0.093	336107	0.013	0.004	2.47E-03
Height	rs4912122	1	19876438	*NKX2-2*	A/G	0.413	334951	-0.015	0.003	6.33E-09	0	0.423	-0.019	0.003	1.45E-13	0.350	336474	-0.012	0.002	1.58E-11
Height	rs4425978	1	42243878	*HIVEP3*	T/C	0.516	351587	0.014	0.003	2.14E-08	0	0.472	0.016	0.003	3.63E-10	0.533	336474	0.008	0.002	5.41E-06
Height	rs618555	1	86481084	*COL24A1*	T/C	0.323	320239	0.019	0.003	5.47E-12	0	0.258	0.016	0.003	2.90E-08	0.311	336474	0.008	0.002	1.67E-05
Height	rs6545538	2	56217900	*MIR216A*	A/G	0.308	305704	0.022	0.003	1.23E-13	0	0.309	0.019	0.003	2.59E-11	0.266	336474	0.011	0.002	1.25E-08
Height	rs2741311	2	233239743	*ALPP*	T/C	0.054	463609	0.046	0.005	1.41E-21	49.6	0.064	0.030	0.005	4.78E-10	0.080	336474	0.033	0.003	2.39E-24
Height	rs6935954	6	26255451	*HIST1H2BH*	A/G	0.374	345378	0.042	0.003	3.06E-59	46.5	0.359	0.018	0.003	3.29E-12	0.425	336474	-0.027	0.002	1.10E-54
Height	rs6900530	6	35280971	*DEF6*	T/C	0.207	123137	-0.057	0.005	3.09E-28	79	0.088	-0.036	0.005	6.75E-12	0.027	336474	-0.073	0.005	1.03E-42
Height	rs9472006	6	43067487	*PTK7*	A/G	0.141	212931	-0.027	0.005	4.37E-09	0	0.081	-0.034	0.005	7.04E-14	0.041	336474	-0.013	0.004	2.86E-03
Height	rs3822957	6	76607280	*MYO6*	A/G	0.235	279818	-0.023	0.003	4.88E-12	84.4	0.197	-0.023	0.003	3.44E-12	0.142	336474	-0.015	0.002	1.98E-09
Height	rs1342330	6	144065685	*PHACTR2*	A/T	0.547	353259	0.014	0.003	1.70E-08	0	0.565	0.017	0.003	6.95E-12	0.520	336474	0.006	0.002	8.84E-04
Height	rs6936615	6	154355100	*OPRM1*	A/G	0.859	415248	-0.018	0.003	2.47E-08	0	0.834	-0.020	0.003	2.33E-09	0.830	336474	-0.003	0.002	1.73E-01
Height	rs991946	6	166329862	*RP11-252P19.3*	T/C	0.495	379912	-0.019	0.002	1.28E-14	36.7	0.481	-0.019	0.002	9.96E-15	0.479	336474	-0.013	0.002	3.92E-13
Height	rs7816300	8	109787856	*TMEM74*	T/C	0.253	397735	-0.015	0.003	2.24E-08	0	0.264	-0.016	0.003	5.47E-09	0.299	336474	-0.002	0.002	3.94E-01
Height	rs4520250	9	88924057	*ZCCHC6*	A/C	0.387	295945	0.015	0.003	3.80E-08	55.9	0.322	0.016	0.003	1.48E-08	0.339	336474	0.010	0.002	4.36E-08
Height	rs7029157	9	97000863	*snoU13*	T/C	0.138	262808	0.028	0.004	5.85E-11	16.8	0.115	0.026	0.004	1.00E-09	0.088	336474	0.030	0.003	4.55E-22
Height	rs12347744	9	97575273	*C9orf3*	T/C	0.053	454111	-0.032	0.005	6.55E-11	65.8	0.056	-0.030	0.005	6.06E-10	0.061	336474	-0.031	0.004	1.04E-17
Height	rs7024254	9	109498129	*ZNF462*	A/G	0.280	324351	0.017	0.003	1.04E-08	23.7	0.278	0.036	0.003	6.10E-35	0.204	336474	0.010	0.002	6.50E-06
Height	rs10119624	9	118305438	*DEC1*	A/G	0.640	353374	0.021	0.003	2.56E-16	31.1	0.635	0.022	0.003	8.02E-17	0.671	336474	0.012	0.002	3.20E-10
Height	rs2902635	10	105476045	*SH3PXD2A*	T/G	0.729	308782	-0.021	0.003	1.59E-12	0	0.739	-0.017	0.003	1.62E-08	0.805	336474	-0.015	0.002	7.65E-12
Height	rs17659078	11	2284590	*ASCL2*	A/C	0.241	354931	0.019	0.003	1.78E-10	9	0.243	0.016	0.003	3.45E-08	0.273	336474	0.004	0.002	2.24E-02
Height	rs11605693	11	122837037	*C11orf63*	T/C	0.470	380447	-0.017	0.002	3.44E-12	0	0.449	-0.018	0.002	3.07E-14	0.447	336474	-0.013	0.002	4.45E-13
Height	rs621794	11	125849462	*CDON*	A/G	0.475	380049	-0.014	0.002	1.46E-08	0	0.438	-0.014	0.002	7.29E-09	0.429	336474	-0.009	0.002	1.76E-07
Height	rs11221442	11	128577624	*FLI1*	C/G	0.202	352360	-0.022	0.003	2.03E-12	78.8	0.207	-0.023	0.003	2.03E-13	0.252	336474	-0.008	0.002	2.59E-05
Height	rs12300112	12	103147575	*LINC00485*	C/G	0.126	154014	0.041	0.006	4.31E-13	45.2	0.063	0.038	0.006	3.11E-11	0.027	336474	0.038	0.005	4.61E-12
Height	rs11616067	12	116393174	*MED13L*	A/G	0.775	327941	0.021	0.003	6.38E-12	21	0.802	0.018	0.003	2.80E-09	0.768	336474	0.012	0.002	6.01E-09
Height	rs17197170	14	21977962	*METTL3*	A/G	0.856	310344	-0.026	0.004	5.40E-12	64.9	0.858	-0.025	0.004	3.74E-11	0.828	336474	-0.018	0.002	3.07E-14
Height	rs11076551	16	51109492	*RP11-883G14.4*	A/G	0.289	434688	0.014	0.003	2.14E-08	21.6	0.370	0.015	0.003	5.87E-09	0.376	336474	0.009	0.002	2.51E-07
Height	rs12918773	16	89741403	*C16orf55*	A/G	0.138	279104	-0.024	0.004	7.45E-09	46.1	0.100	-0.024	0.004	3.50E-09	0.112	336474	-0.024	0.003	4.71E-19
Height	rs1346490	19	7244233	*INSR*	A/C	0.521	333896	0.015	0.003	3.99E-09	12.3	0.617	0.014	0.003	4.86E-08	0.620	336474	0.009	0.002	1.67E-07
Height	rs17457472	19	17493610	*PLVAP*	A/C	0.035	356712	-0.051	0.007	5.70E-14	0	0.039	-0.042	0.007	9.89E-10	0.040	336474	-0.019	0.004	1.12E-05
WHRadjBMI sex-combined	rs17099388	5	142095250	*FGF1*	A/G	0.165	105460	0.039	0.007	3.14E-09	0	0.074	0.037	0.007	2.47E-08	0.041	484563	0.025	0.005	2.00E-07
WHRadjBMI sex-combined	rs7975017	12	26428793	*SSPN*	T/C	0.243	267044	-0.021	0.004	7.57E-09	5.50	0.263	-0.021	0.004	1.05E-08	0.239	484563	-0.014	0.002	2.00E-09

Abbreviations: Chr-chromosome; EAF-effect allele frequency; HetIsq-heterogeneity I-square; N-sample size; WHRadjBMI-waist to hip ratio adjusted for BMI; African American Anthropometry Genetics Consortium AAAGC; Genetic Investigation of ANthropometric Traits (GIANT). 1 Novel locus defined by no known published variants within 1Mb (+/-500 Kb) of the lead SNP. 2 Known locus defined by a 1Mb region with previously identified signal(s) for the indicated trait; the known SNP(s), P<5e-8, at each established locus can be found in Table S16.

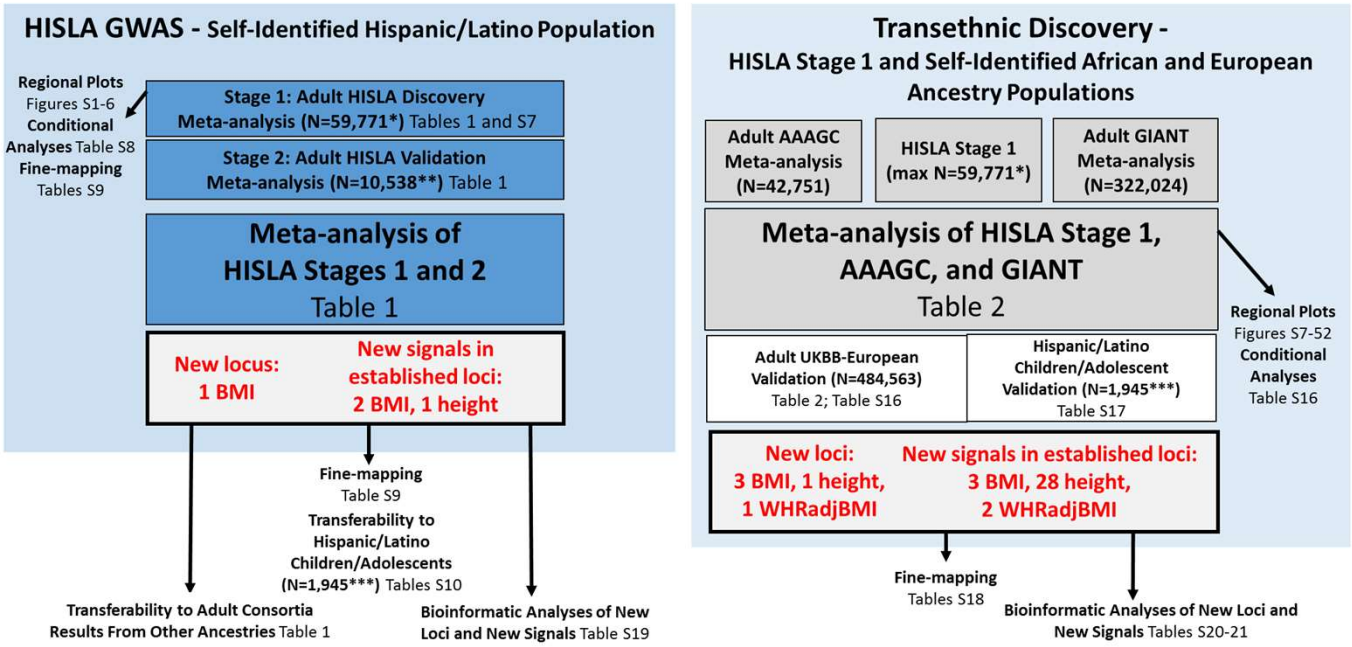
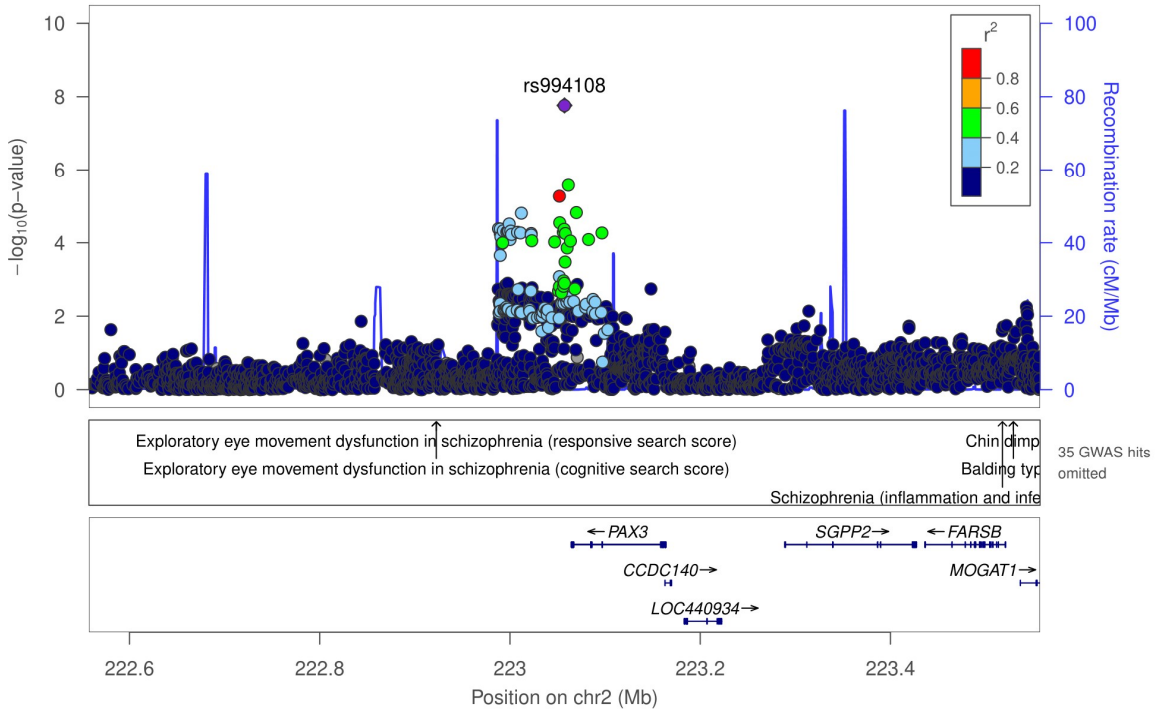


Figure 1. Flowchart of the design and discovery of 6 novel loci and 36 novel signals in known loci in the Hispanic/Latino Anthropometry Consortium (HISLA) Meta-Analysis and the Trans-Ethnic Meta-Analysis of HISLA and Consortia of Other Ancestral Heritages

*Stage 1 maximum sample sizes varied between and 59,771 for BMI, 56,161 for height, and 42,455 for WHRadjBMI (sex combined). **Stage 2 sample sizes varied between 10,538 for BMI, 8,110 for height, and 4,393 for WHRadjBMI (sex combined). Actual sample sizes may vary by SNP. ***The BMI and height-for-age z-score models were conducted using up to 1,914 and 1,945 of children/adolescents, respectively. In contrast, the obesity case-control study compared up to 1,814 children/adolescents who were either ≥95th versus ≤50th BMI-for-age percentiles.

A. *PAX3* locus associated with BMI



B.

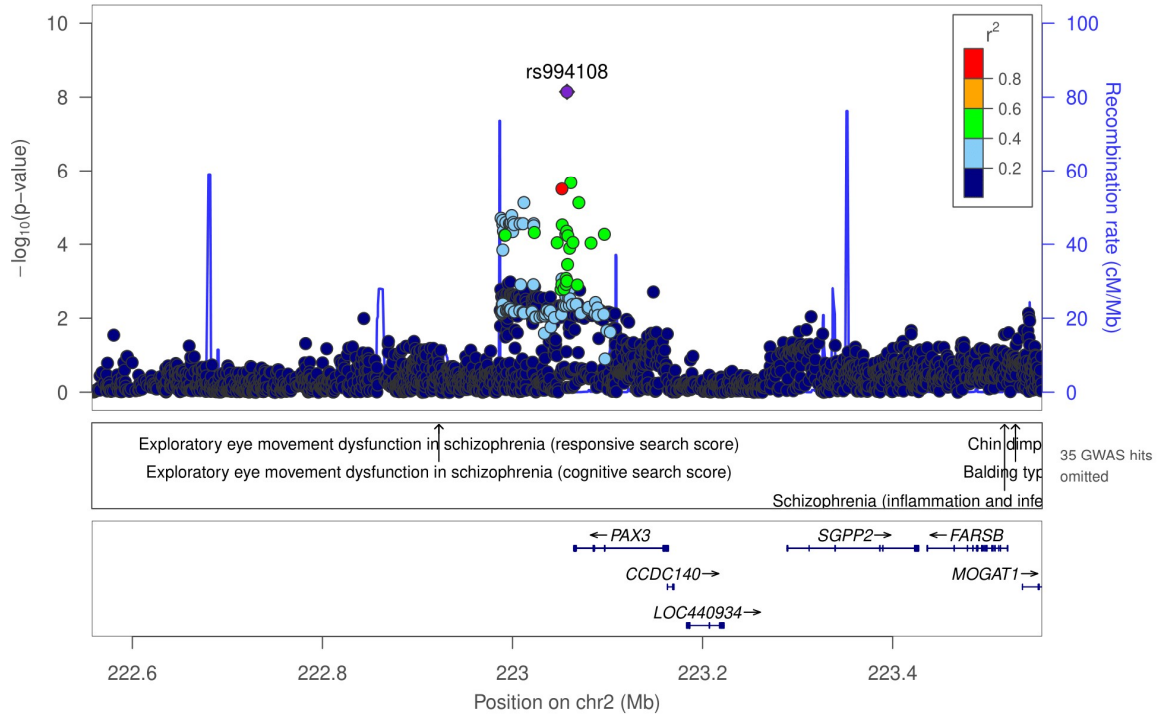


Figure 2. Regional plot, unconditioned (A) and conditioned (B), of the novel *PAX3* locus associated with body mass index (BMI) in the Hispanic/Latino Anthropometry Consortium (HISLA), excluding Brazilian and Native American samples. Linkage disequilibrium patterns are based on rs994108 (shown by the purple triangle) from the Hispanic Communities in Health Study/Study of Latinos.

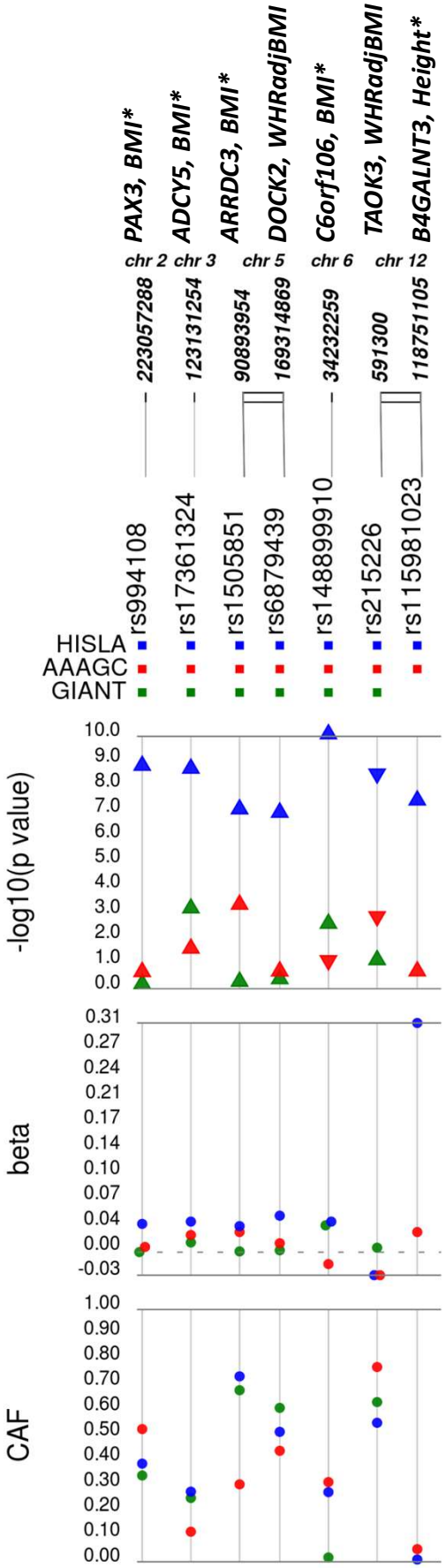


Figure 3. Variability in HISLA Stage 1+2, AAAGC, and GIANT P-values, Effect Sizes and Coded Allele Frequencies for Genome-Wide Significant Anthropometric Loci from HISLA Stage 1.

Hispanic/Latino Anthropometry Consortium (HISLA); African American Anthropometry Genetics Consortium (AAAGC); Genetic Investigation of ANthropometric Traits (GIANT); WHRadjBMI - waist to hip ratio adjusted for BMI. *Asterisks indicating a SNPs that were significant either as a novel locus or new signals in a known locus.

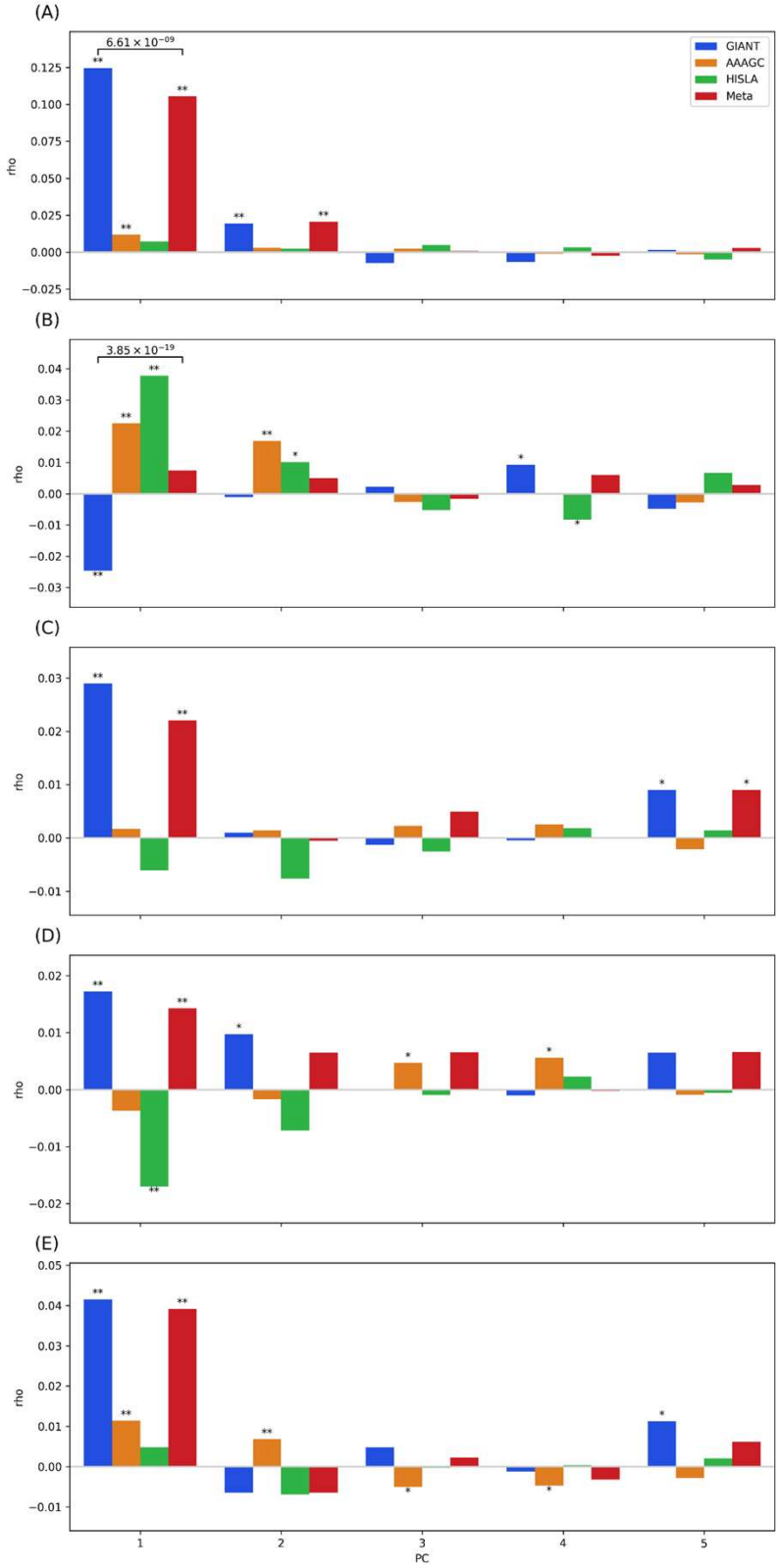


Figure 4. Correlations (rho) between effect estimates and the loadings of the principal components 1-5 in each consortia and the meta-analysis of all 3 consortia by trait. (A) height, (B) BMI, (C) Waist-to-hip ratio adjusted for BMI (WHRadjBMI) for men and women combined, (D) WHRadjBMI for women only, (E) WHRadjBMI for men only.

Hispanic/Latino Anthropometry Consortium (HISLA); African American Anthropometry Genetics Consortium (AAAGC); Genetic Investigation of ANthropometric Traits (GIANT); WHRadjBMI - waist to hip ratio adjusted for BMI