Genetic contributions to obesity and related complex traits in the Cebu Longitudinal Health and Nutrition Survey

Damien C. Croteau-Chonka

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum in Genetics and Molecular Biology.

Chapel Hill
2012

Approved by:

Linda S. Adair, Ph.D.

Ethan M. Lange, Ph.D.

Yun Li, Ph.D.

Karen L. Mohlke, Ph.D.

Patrick F. Sullivan, M.D., FRANZCP

ABSTRACT

DAMIEN C. CROTEAU-CHONKA: Genetic contributions to obesity and related complex traits in the Cebu Longitudinal Health and Nutrition Survey
(Under the direction of Karen L. Mohlke, Ph.D.)

Obesity represents a major world health problem across socio-economic strata. Characterizing known candidate genes and identifying novel genes contributing to obesity and other related heritable risk factors may aid our understanding of its complex etiology and provide new treatment targets.

Based in the Philippines, a country experiencing a rapid obesogenic transition, the Cebu Longitudinal Health and Nutrition Survey (CLHNS) is a useful population cohort in which to study the genetics of obesity. For more than 20 years since 1983, researchers have followed a cohort of Filipino mothers and their offspring from the metropolitan area of Cebu, collecting a wide variety of anthropometric, environmental, and dietary phenotypes. In 2005, blood samples were collected from all the participants, allowing the measurement of multiple plasma biomarkers, the extraction of genomic DNA, and the genetic analysis of complex traits.

In this dissertation, I describe my genetic studies of three categories of obesity-related complex traits in the CLHNS: body size (body mass index (BMI),

weight, and waist circumference), metabolic (adiponectin), and physiological (age of menarche).

I first performed a genome-wide association (GWA) study of body size traits in Filipinos to determine the contribution of common genetic variants to those traits. I replicated three well-known BMI loci (*BDNF*, *MC4R*, and *FTO*) and further observed evidence of longitudinal changes in the effects of those genes.

Next, I sought to establish the putative causal variant(s) underlying a haplotype at the *ADIPOQ* gene identified in a previous CLHNS GWA study to be strongly associated with lower circulating plasma adiponectin level. I identified a population-specific missense variant (R221S) that affected the original measurement of the phenotype and explained the observed GWA signal.

Next, to identify novel biology underlying the etiology of central adiposity, I performed a meta-analysis of GWA studies of waist circumference (WC) in European individuals from the Genetic Investigation of Anthropometric Traits (GIANT) Consortium. Findings from this large-scale study may suggest other candidate loci for future genetic study in the CLHNS. I observed evidence of novel associations specific to WC and not other related anthropometric traits, such as BMI, height, and waist-hip ratio. These signals included the genes *NLRP3*, which is part of the obesity-related inflammasome complex, and *IRS1*, which is previously associated with body fat percentage and an adverse metabolic profile.

Finally, to explore the relationships between several obesity genes and age of menarche, I performed replication and mediation analyses in female offspring from the CLHNS. I observed that the menarche signal at *LIN28B* previously reported in

Europeans replicated in Filipinos. While this association appeared not to be mediated by childhood BMI, the results for several obesity genes were inconclusive.

Together, these studies further our understanding of the genetic architecture of several obesity traits, and suggest many promising biological candidates for further genetic and molecular study.

To my mom, the original Dr. Croteau-Chonka

ACKNOWLEDGEMENTS

Science in the 21st century is very much a team sport, and I am indebted to the many people who helped guide, support, and distract me during my doctoral studies.

First, I would like to thank my advisor, Dr. Karen Mohlke, for being a mentor *par excellence*: sending tremendous opportunities my way early and often and encouraging me to take intellectual ownership of my work. In addition, Karen's ongoing enthusiasm for the scientific endeavor was uplifting even when the scientific process was a little bit of a grind. I would also like to thank the members of my dissertation committee: Drs. Linda Adair, Ethan Lange, Yun Li, and Pat Sullivan. Their periodic reminders to really know the biology inside and out and to stick to my guns were much appreciated. Thanks also to Dr. Leslie Lange for her statistical assistance and advice on various projects.

Next, I would like to acknowledge the love and support of my family during these five years. First and foremost, I would like to acknowledge my mother, Clarisse, whose continuing commentary, insights, and advice on life in and out of science make her a committee member *ex officio* in my mind. I would also like to acknowledge my father, Phil, and my younger sisters, Alexia and Elise, for imparting

words of encouragement and for occasionally taking me down a peg to keep this process honest.

Grad school would not have been nearly as positive an experience if not for the many awesome friends I made in both the Genetics Curriculum and the Bioinformatics and Computational Biology Curriculum. For fear of leaving anybody out, I will omit an exhaustive list, but they know who they are. Thanks to all of them for their devoted comradery, for our many fun adventures enjoying North Carolina, and for all the rides before I finally got myself a car!

A special thanks, too, to Professor Susan Klebanow and the many members of the UNC Chamber Singers with whom I've had the pleasure of making music (with Karen's blessing, of course). I always told people that Chamber Singers was the four hours a week I didn't think about science unless asked a direct question. To their credit indeed, my fellow singers would often be interested in what it was that I did every day.

Finally, I would like to acknowledge the many fine people who have been a part of the Mohlke Lab during my time at UNC. Amanda Marvelle, Kyle Gaulton, and Li Qin were there at the start, helping to sell me on the lab before eventually moving on to bigger and brighter things. I've known Marie Fogarty the longest as she started her postdoctoral fellowship with Karen just as I started my first rotation. Tami Panhuis and Ghenadie Curocichin had briefer stays in the middle. I leave the lab then in the capable hands of Ying Wu, Jennifer Kulzer, Tamara Roman, Martin Buchkovich, and Rani Vadlamudi. A lab develops a collective personality reflective of its members, and the Mohlke Lab has a warm, friendly, and relaxed one. Thanks!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

1000G: 1000 Genomes Project

ADIPOQ: adiponectin, C1Q and collagen domain containing

BF%: body fat percentage

BMI: body mass index

CDH13: cadherin 13, H-cadherin

CEU: Utah residents with ancestry from northern and western Europe

CHB: Han Chinese from Beijing, China

CNV: copy number variation

CVD: cardiovascular disease

CLHNS: Cebu Longitudinal Health and Nutrition Survey

CT: computed tomography

DNA: deoxyribonucleic acid

DXA: dual-energy X-ray absorpitometry

ELISA: enzyme-linked immunosorbent assay

ETV5: ets variant gene 5

eQTL: expression quantitative trait locus

GWA: genome-wide association

h^2: narrow-sense heritability

HapMap: International Haplotype Map Project

HC: hip circumferenceJPT: Japanese from Tokyo, Japan

LD: linkage disequilibrium

MAF: minor allele frequency

MetS: metabolic syndrome

PolyPhen: Polymorphism Phenotyping

R^2: squared type II partial correlations

SAT: subcutaneous adipose tissue

SD: standard deviation

SE: standard error of the mean

SEM: standard error of the mean

SES: socio-economic status

sICAM-1: soluble circulating intracellular adhesion molecule-1

SIFT: Sorting Intolerant from Tolerant

SNP: single nucleotide polymorphism

T2D: type 2 diabetes

UNC-CH: The University of North Carolina at Chapel Hill

VAT: visceral adipose tissue

WC: waist circumference

WHO: World Health Organization

WHR: waist-hip ratio

YRI: Yorubans from Ibadan, Nigeria

CHAPTER I

Introduction

Obesity, or the accumulation of excess body fat, is a major world health issue. The World Health Organization (WHO) has estimated that in 2008 1.5 billion people worldwide were overweight, and of those, 500 million were obese (1). Despite being largely preventable, obesity and overweight together are the fifth leading risk factors for death worldwide (1). In the United States alone, medical costs in 2008 related to obesity were estimated to be $147 billion per year, representing nearly 10% of all medical costs (2). Once thought to be a disease of wealthy countries, the prevalence of obesity is also increasing in low- and middle-income nations as well (3), especially in Asian countries. A major consequence of excess body fat is the development of metabolic syndrome (MetS), a collection of maladies including dyslipidemia, insulin resistance, and hypertension that contributes to increased risk of cardiovascular disease (CVD) (4). CVDs resulted in 30% of all deaths worldwide in 2008, more than any other single cause (5).

Obesity is a textbook example of a common complex disease, meaning that genetic and environmental factors both play independent and interdependent roles in its development and widespread prevalence. Decreased physical activity and increased consumption of foods with high caloric and fat content are the two major

causes of obesity across the world (1). Socio-economic status (SES), which influences both of these two factors, has itself a complicated relationship with obesity. In developed countries, higher SES is consistently associated with decreased risk of obesity in women but not in men (6). In developing countries, higher SES is generally associated with increased risk of obesity (6), but as these countries become richer, individuals with lower SES become more obese (7). Nonetheless, heritability studies of obesity-related traits in pairs of monozygotic and dizygotic twins suggest they each have substantial genetic components (h^2 = 40-70%) (8).

MAJOR BODY FAT DEPOTS

Body fat, also known as adipose tissue, is one of the body's largest organs and is divided into two major categories: brown and white (9). Brown adipose tissue, found abundantly in babies, is highly metabolically active and thermogenic. The majority of adult fat is white adipose tissue, which serves as an energy storage depot. White adipose tissue is divided into two categories based on its location and metabolic characteristics: subcutaneous adipose tissue (SAT) and visceral adipose tissue (VAT). SAT is located beneath the skin, extending inward from the dermis to the outer bowel or abdominal wall in cross-section. SAT is further divided into subcategories of "superficial" and "deep", corresponding to their relative proximity to the outer surface of the body. VAT, often called "intra-abdominal" fat, is located inside the abdominal wall surrounding the internal organs, or extending from the outer bowel in towards the spine in cross-section. While both types of adipose

tissue are associated with several metabolic risk factors, such as increased fasting plasma glucose or high blood pressure, these associations are significantly stronger in VAT than in SAT (10). Relatedly, VAT is also associated with increased risk of CVD (11). Within SAT, deep SAT is more associated with metabolic complications than superficial SAT (12).

The amount of fat per given body size varies by sex and age and also by ethnicity, which could be a proxy for the influence of genetic and dietary factors. Initial differences in body composition between men and women manifest during puberty and are related to exposures to various endocrine factors (13). Women tend to develop more gluteo-femoreal fat (i.e., in the thighs, hips, and buttocks), whereas men develop more abdominal fat (13). In both men and women, body weight and the prevalence of obesity generally increase with age, mostly due to changes in diet and lifestyle (14, 15). After the onset of menopause, substantially more fat is distributed to the intra-abdominal area in women than before (16). Ethnic differences in obesity risk have also been observed in both children and adults, and are attributed to a wide variety of factors, including SES, physiology, and other cultural exposures (17).

MEASURING OBESITY TRAITS AND ASSOCIATED HEALTH RISKS

The gold standard for measuring VAT and SAT is computed tomography (CT), but the technology is very expensive, requires substantial training to use, exposes the patient to a non-negligible dose of radiation, and is not high-throughput enough to assess the large numbers of individuals typically involved in

epidemiological studies. Several other technological methods exist for measuring body fat, including ultrasound and dual-energy X-ray absorpitometry (DXA). Ultrasound measures body fat by applying very high frequency sound waves to the tissue and using the timing of the echoes to determine its size. DXA uses differences in X-ray absorption to help distinguish bone from soft tissue and estimate fat volume. Both of these methods are correlated with VAT measures made by CT (r = 0.79 and 0.70, respectively) (18). As succinctly summarized in (19), these methods and others not described here have advantages and disadvantages in terms of their availability, specificity, accuracy, and reproducibility that have to be carefully considered.

Anthropometric measures of obesity are most commonly used in research settings as they can be obtained in the field or the clinic with considerably less expensive equipment (e.g., a tape measure) and more basic training of staff. One of the most commonly used measures is body mass index (BMI), which is the ratio of a person's weight (kg) to the square of their height (m^2). BMI represents overall adiposity or the total amount of fat located around the body. While some body fat is necessary for proper health, excess adiposity is associated with increased cardio-metabolic risk. According to the WHO, a man or woman with a BMI ≥ 25 kg/m^2 is considered overweight and a person with a BMI ≥ 30 kg/m^2 is considered obese (1). In 2000, the WHO proposed lower thresholds in Asians for overweight and obesity of 23 and 25 kg/m^2, respectively, as they showed increased morbidity and mortality at lower BMIs than Europeans (20). Later in 2004, however, the WHO concluded instead that differences in obesity-related health risk among various Asian

populations did not suggest clear enough cut-off points different from the original guidelines (21). Although BMI is strongly correlated with VAT (r = 0.67) (18), its primary weakness as a health risk indicator is that it cannot distinguish between individuals who differ in muscle mass given the same weight and height.

Other common anthropometric measures related to health outcomes include waist circumference (WC), hip circumference (HC), and waist-hip ratio (WHR). WC, a measure of central adiposity (units in cm), is more strongly correlated with VAT (r = 0.85) (18) and with the components of MetS than BMI (22). A WC of ≥ 102 cm in men and ≥ 88 in women, often called "central obesity", is one of the key features of metabolic syndrome (4). As with BMI, the WHO recommended a lower threshold for central obesity in Asian men (90 cm) and women (80 cm) (20), but has not revised this recommendation any further. In support of this lower threshold, men and women of Chinese and south Asian descent have 12–24% greater amounts of VAT at a given WC than Europeans (23). Notably, differences in which bony or external landmarks are used for the measurement of WC appear not to substantially influence its relationship with risk of morbidity or mortality (24). In contrast to WC, increased HC (also measured in cm) has a protective effect on risk of type 2 diabetes (T2D) and CVD (25, 26). WHR, or the ratio of WC to HC (both measured in cm), is a measure of central fat distribution. WHR is considered a better predictor of CVD risk than WC alone because it also incorporates the effects of HC (27, 28). A WHR > 0.95 in men and > 0.80 in women has been designated as the cutoff for greater CVD risk (29).

Anthropometric traits exist along with other heritable obesity-related traits—both metabolic (e.g., adiponectin) and physiological (e.g., age at menarche)—as nodes in a complex biological network underlying the etiology of obesity. Their relationships will be briefly outlined below.

Excess adiposity results in an altered metabolic state that underlies the downstream co-morbidities of obesity (30). One of the most abundant proteins found in the bloodstream, adiponectin is a protein hormone primarily secreted by cells called adipocytes that comprise the bulk of fat depots (31). Adiponectin is implicated in glucose regulation and fatty acid catabolism (32) and plays a number of roles in the cardiovascular system (32). Circulating plasma levels of adiponectin are negatively correlated with BMI (33), and hypoadiponectinemia is a risk factor for MetS and related traits (34, 35). The heritability of adiponectin level ranges from 30-70% (36-38).

Obesity is also connected to physiological traits such as age of menarche or the first onset of menstruation in girls. Very early menarche (< 12 years of age) is a risk factor for adult obesity (39, 40), T2D (41, 42), MetS (43), and CVD (44). Childhood obesity, or high BMI before puberty, is linked to very early menarche and may explain its connection to adult risk of obesity (39, 45). Age of menarche is also negatively associated with HC and positively associated with WC (46). Urban life, higher SES, nutrition, and parental education are other factors correlated with earlier menarche (47). Adjusted for adult measures of BMI, WC, and other factors, early menarche is associated with higher incidence of CVDs and all-cause and cancer mortality (44). While substantively influenced by such environmental factors,

differences in menarcheal timing between racial groups suggest genetic contributions to the trait (48). Approximately 50% of phenotypic variability in age of menarche is estimated to be explained by genetic factors (49). It is yet unclear whether early menarche is a risk factor *per se* or whether it is associated with poor health outcomes through the development of obesity or hormonal changes, or possibly all three pathways concurrently.

GENETIC ASSOCIATION STUDIES OF OBESITY TRAITS

Characterizing known candidate genes and identifying novel genes contributing to obesity and related complex traits may aid our understanding of their complex etiologies and provide new treatment targets. Genetic association studies use statistical correlations between genotype and phenotype to identify variants and genes to prioritize for further functional characterization. Genetic studies of candidate genes cultivate prior biological knowledge of a trait to increase the likelihood of finding such variants. The genome-wide association (GWA) study of single nucleotide polymorphisms (SNPs) has emerged as a valuable tool for the agnostic generation of genetic hypotheses (50). By systematically testing variants across the genome, they have been useful in discovering novel genes not previously thought to be relevant to a given trait or disease (50, 51). This discovery method avoids the major disadvantage of candidate studies that rely on an incomplete knowledge of the biology underlying a trait to select specific genes for study. The first "official" GWA study was published in 2005 (52), but 2007 saw a dramatic

increase in the number of published GWA studies and is considered a turning point towards widespread adoption of the approach (50).

Two competing hypotheses exist about how associated genetic variants, such as SNPs, act to drive phenotypic variability (51). The first posits that multiple common variants with smaller effects together in aggregate have a substantial impact on a trait. Most GWA studies so far have focused on studying common SNPs ascertained by the International HapMap Project (53). The second hypothesis contends instead that a few rare variants have large individual effects on trait variability. These two hypotheses are not necessarily mutually incompatible, but without a more comprehensive catalog of known genetic variants to test for association and much functional characterization of observed association loci, it is as yet unclear how much each frequency class contributes. Towards this end, the 1000 Genomes Project was initiated to ascertain a more complete set of low frequency variants for genetic study (54).

The National Human Genome Research Institute (NHGRI) has published and continues to update a resource cataloging the results of GWA studies since 2005 (55). As of April 2012, the NHGRI GWA catalog (available on-line at http://www.genome.gov/gwastudies/) contains > 6,000 SNPs associated with >200 traits. These publications include many associations with obesity-related traits. While the success of early GWA studies prompted many similar efforts in a wide variety of complex traits, the next step of combining association studies together by meta-analysis also has really helped fuel the rapid discovery of novel loci. By combining millions of SNPs from hundreds of thousands of samples across different

studies, research groups obtain the increased sample sizes and statistical power necessary to detect additional genetic variants with even smaller contributions to traits of interest.

One such meta-analysis effort is the Genetic Investigation of Anthropometric Traits (GIANT) Consortium, which studies heritable body size traits. Previous GIANT meta-analyses have been published on SNP associations with waist circumference (56), waist-to-hip ratio (WHR) (56, 57), BMI (58, 59), and height (60). These studies observed that while they only explained a small fraction of total variation in their respective anthropometric traits, there was compelling evidence to suggest that greater sample sizes would help identify additional signals. To facilitate ongoing large-scale replication of these signals, the Cardio-MetaboChip (MetaboChip) array was designed by members of GIANT and other large disease and trait consortiums. The MetaboChip is a custom genotyping array containing >66,000 individual SNPs identified in meta-analyses as associated with one or more of 23 metabolic, atherosclerotic, and cardiovascular diseases and traits. At 257 loci previously detected for these phenotypes, >140,000 SNPs provide dense coverage of rare and common variants to allow for the fine-mapping of those association signals and the identification of potential secondary signals.

While the majority of GWA studies have so far primarily studied subjects with European ancestry, the study of cohorts with Asian and African ancestry will provide opportunities to uncover novel associated variants, especially rare ones not present in European populations (50, 51). In addition, inter-population differences in local linkage disequilibrium (LD), which are especially pronounced between

continental populations (61), may assist in fine-mapping the disease-causing variants in loci implicated across association studies. In some cases, the overlap of LD between two populations may correspond to an area of increased association signal strength, thus suggesting a narrower region of interest than in the discovery population alone (62).

THE CEBU LONGITUDINAL HEALTH AND NUTRITION SURVEY

The Cebu Longitudinal Health and Nutrition Survey (CLHNS) cohort is well suited for testing a wide range of hypotheses about the inter-relationships of genetic factors, early child exposures, and adult health outcomes because of its wealth of prospectively collected data. A comprehensive overview of the CLHNS, a birth cohort of Filipino mothers and their children, has been published previously (63). Briefly, the original purpose of the survey in 1983 was to study infant feeding patterns in the metropolitan area of Cebu, Philippines and their effects on the child, its mother, and their household. Over time, focus shifted to the impact of early life on later development of obesity and CVD risk. In the last few decades, concurrent with a general rise in economic prosperity, Filipinos have been adopting Western dietary habits and have become more sedentary in their daily lives, much like many populations in south-east Asian countries (64). As a result of this obesogenic transition, rates of overweight and obesity have increased rapidly in the Philippines (65).

During the original survey and at periodic follow-ups, trained field staff from the CLHNS conducted in-home interviews and collected anthropometrics and comprehensive environmental data, including nutritional intake and socio-economic indicators. In 2005, fasting blood samples were collected from both the mothers and their now young adult offspring. Multiple plasma biomarkers were measured and genomic DNA was extracted. These phenotypes and genotypes have allowed for the ongoing genetic study of complex traits in this population-based cohort. So far, data from the CLHNS has been utilized in a candidate gene study of obesity (66), GWA studies of plasma homocysteine (67), C reactive protein (68), and adiponectin (69), as well as in a large-scale meta-analysis of lipid trait associations (70).

In the following chapters of this dissertation, I describe my studies of the genetic bases of body size (71), circulating plasma adiponectin levels (72), and age of menarche in individuals from the CLHNS. I also describe my main contribution to the GIANT Consortium's efforts to understand the genetic basis of central obesity in individuals of European ancestry. In the concluding chapter, I discuss how together these published and unpublished data contribute to a greater understanding of several quantitative traits related to risk of overweight and obesity.

CHAPTER II

Genome-wide association study of anthropometric traits and evidence of interactions with age and study year in Filipino women[1]

OVERVIEW

Increased values of multiple adiposity-related anthropometric traits are important risk factors for many common complex diseases. We performed a genome-wide association (GWA) study for four quantitative traits related to body size and adiposity (body mass index [BMI], weight, waist circumference, and height) in a cohort of 1,792 adult Filipino women from the Cebu Longitudinal Health and Nutrition Survey. This is the first GWA study of anthropometric traits in Filipinos, a population experiencing a rapid transition into a more obesogenic environment. In addition to identifying suggestive evidence of additional SNP association signals ($P < 10^{-5}$), we replicated ($P < 0.05$, same direction of additive effect) associations previously reported in European populations of both BMI and weight with *MC4R* and *FTO*, of BMI with *BDNF*, and of height with *EFEMP1*, *ZBTB38*, and *NPPC*, but none with waist circumference. We also replicated loci reported in Japanese or Korean populations as associated with BMI (*OTOL1*) and height (*HIST1H1PS2*, *C14orf145, GPC5*). A difference in local linkage disequilibrium between European

---

[1] A version of this work was previously published as Croteau-Chonka DC *et al. Obesity (Silver Spring)*. 2011 May;**19**(5):1019-27. Epub 2010 Oct 21.

and Asian populations suggests a narrowed association region for *BDNF*, while still including a proposed functional non-synonymous amino acid substitution variant (rs6265, Val66Met). Finally, we observed significant evidence ($P < 0.0042$) for age-by-genotype interactions influencing BMI for rs17782313 (*MC4R*) and rs9939609 (*FTO*), and for a study year-by-genotype interaction for rs4923461 (*BDNF*). Our results show that several genetic risk factors are associated with anthropometric traits in Filipinos and provide further insight into the effects of *BDNF*, *FTO*, and *MC4R* on BMI.

INTRODUCTION

Increased values of adiposity-related traits are important risk factors for many common complex diseases. Understanding the basis for increased human body size may lead to insights into disease etiologies. Higher body mass index (BMI), weight, and waist circumference are associated with type 2 diabetes, cardiovascular disease (CVD), hypertension, and cancer (73, 74). Increased height, however, is inversely associated with CVDs (75).

Genome-wide association (GWA) studies have been used to identify multiple loci associated with variation in anthropometric measures (56, 58, 62, 76-91), often with more than one correlated trait. For signals that are replicated across world populations, differences in population history provide potential to further localize the associated regions. Some signals may be population-specific due to differing allele frequencies and environmental contexts.

Over the last few decades, the adoption in Asian populations of Western-style diets of increased fats and carbohydrates and of more sedentary habits has led to a marked increase in obesity (64, 65). In particular, a cohort of women from the ongoing Cebu Longitudinal Health and Nutrition Survey (CLHNS) based in the Philippines showed a six-fold increase in prevalence of overweight and obesity associated with nearly two decades of substantial and continuing socio-economic modernization (also illustrated by a increase in mean weight of 6.8 ± 7.1 kg) (65). The portion of increased prevalence due to the changes in environment versus increased age of these women is unclear.

We performed a GWA study to test for main effect SNP associations with measures of BMI, weight, waist circumference, and height in 1,792 adult Filipino women from the CLHNS. The longitudinal nature of this cohort also allowed us to examine the interactive effect of age and genotype on BMI over a 22-year period from 1983 to 2005.

METHODS AND PROCEDURES

We initially evaluated 1,895 female participants from the ongoing CLHNS, mothers of a 1983 to 1984 birth cohort (63). Trained field staff conducted in-home interviews and collected anthropometrics, blood samples, and comprehensive environmental data (publicly available at www.cpc.unc.edu/projects/cebu/). Informed consent was obtained from all CLHNS subjects, and the University of North Carolina Institutional Review Board for the Protection of Human Subjects approved the study protocol.

Most outcome and covariate measures reported in this current study were taken from the 2005 survey, except height, which is the average of the measurement during the pregnancy of the birth cohort in 1983-1984 and the first post-partum measurement. BMI, weight, and waist circumference were highly correlated with each other (Pearson $r > 0.88$), but not with average height ($r < 0.44$) (**Table 2.1**). For the longitudinal analyses, we used BMI from seven different time points in the study. The first BMI value was measured four months after birth of the index child; if data were missing, the measurement at six months or two months post-partum was substituted. Six additional measurements were taken at 1, 8, 11, 15, 19, and 22 years after the baseline survey (the final measurement is from the 2005 survey). Observations were excluded from the analysis if women were pregnant at the time of the survey. The phenotypes weight, waist circumference, and height were approximately normally distributed, and BMI values were natural log-transformed to satisfy model assumptions.

Our full methods for direct SNP genotyping and quality control (QC) as well as SNP imputation have been described previously (67). Briefly, genotyping was performed with the Affymetrix Genome-Wide Human SNP Array 5.0. Samples that failed DNA fragmentation, failed an array QC check (DM algorithm), or had a genotype call rate < 97% were excluded. We also removed one member from any likely first-degree relative pair as determined from identity-by-descent and identity-by-state estimates. We discarded individual SNPs due to poor mapping, call rates < 90%, deviation from Hardy-Weinberg equilibrium ($P < 10^{-6}$), and/or ≥ 3 discrepancies among 40 duplicate pairs. Five HapMap CEPH trios were also

genotyped for QC purposes, and we dropped any SNPs that showed ≥ 3 Mendelian inheritance errors or genotype discrepancies with known HapMap genotypes. We then used CLHNS genotypes of 352,264 SNPs and pooled reference haplotypes of 60 CEU and 90 combined CHB+JPT HapMap samples to impute the genotypes of an additional 1,878,188 SNPs in MACH (92). Imputed values were substituted for all 352,264 directly genotyped SNPs, including any missing genotypes. We then discarded SNPs with low-quality imputations (Rsq ≤ 0.3) and estimated minor allele frequencies ≤ 0.01. In total, 2,073,674 SNPs were tested for association with the four quantitative anthropometric traits in 1,792 non-pregnant CLHNS women with complete trait outcome and covariate data.

To evaluate population substructure among our CLHNS subjects, we constructed principal components (PCs) using the software EIGENSOFT (93, 94). We tested each of the first 10 PCs for association with each of the four anthropometric outcomes (**Table 2.2**). We included all PCs for which association with any trait was significant at $P < 0.05$, hence five PCs were used as covariates in the final SNP association model for all traits.

Array Studio version 3.1 was used to perform the GWA statistical analyses (Omicsoft Corporation, Research Triangle Park, NC, USA). Assuming an additive mode of inheritance, multivariable linear regression models were used to test for an association between the phenotypes and each imputed SNP genotype, with covariate adjustment for the first five PCs, age, age^2, total assets, natural log-transformed income, number of pregnancies (categorized into three groups: 0–4, 5–10, ≥11), and menopausal status. Each of these predictors was significantly

associated ($P < 0.05$) with at least one anthropometric trait in our samples (**Table 2.2**). Quanto version 1.2.3 was used for statistical power calculations (available on-line at hydra.usc.edu/gxe/).

For loci previously reported in a GWA study at $P < 5 \times 10^{-8}$ in at least 1,000 samples, we chose a single representative SNP. If this SNP was not present in our dataset, we substituted a proxy SNP in high LD ($r^2 > 0.8$ in both CEU and CHB+JPT, HapMap Release 22) when possible. For one study of an Asian population cohort, we also evaluated additional loci reported with less significant evidence of association ($P < 10^{-4}$). Conditional analyses to search for independent secondary signals were performed for all SNPs within a 2 Mb region centered on the SNP with the strongest primary signal, including the primary signal SNP as an additional covariate in the linear regression.

To detect differences in local LD structure we identified the genomic positions of the SNPs bounding a 1 LD-map unit window centered on the most strongly associated SNP in a locus, using previously constructed LD maps made from the individual CEU, CHB, and JPT HapMap populations (95).

Selected SNPs were tested for additive genotype effects on BMI in longitudinal models using SAS version 9.3 (SAS Institute, Cary, NC). General linear mixed models were adjusted for the following time-varying covariates: age, actual time in years since baseline study visit, assets, income, urbanicity index (96), menopause status, months since the previous visit spent lactating or pregnant, current lactation status and activity level. Women who were pregnant at the time of measurement were excluded from that particular time point, but included for all

other visits at which they were not pregnant. Examination of the model residuals indicated that natural log-transformation was not appropriate in this longitudinal setting, and we therefore analyzed untransformed BMI.

RESULTS

We tested 2,073,674 SNPs for association with BMI, weight, waist circumference, and height in the CLHNS cohort (**Table 2.3**, **Figure 2.1**). No evidence of residual population stratification or cryptic relatedness between samples was observed based on genomic control values (λ_{GC} = 1.00-1.03) and quantile-quantile plots (**Figure 2.2**). The most significant main effect associations ($P < 10^{-5}$) had not been previously reported (**Table 2.4**). The SNP most strongly associated with BMI was rs17124318 ($P = 5.91 \times 10^{-7}$), located downstream of *ATG4C*. For weight, the most strongly associated SNP was rs16877106 ($P = 1.44 \times 10^{-6}$), located in an intron of *ANAPC4*. The SNP most strongly associated with waist circumference was rs1440072 ($P = 7.87 \times 10^{-7}$), an intergenic SNP located downstream of *KCNE4* and in perfect LD ($r^2 = 1$ [CEU, CHB+JPT]) with a SNP in the 3'-UTR of the gene (rs3795884, $P = 1.65 \times 10^{-6}$). Our strongest height association signal (rs17818399, $P = 2.74 \times 10^{-7}$) spans the *PIGF* and *CRIPT* genes. Our study had 80% power to detect novel SNP associations that explain > 2.2% of the variation in trait outcome at $P < 5 \times 10^{-8}$.

Eight SNP-trait associations previously reported for BMI, weight, waist circumference, and height were replicated ($P < 0.05$, effect in the same direction) in the CLHNS (**Table 2.5**, **Table 2.6**). Evidence of association with BMI was observed

at the *BDNF* (rs4923461, $P = 0.00028$), *MC4R* (rs17782313, $P = 0.0073$), and *FTO* loci (rs9939609, $P = 0.0074$). Evidence from conditional analyses was consistent with a single BMI association signal at *MC4R*. For weight, we replicated previously reported associations at *FTO* (rs3751812, $P = 0.019$) and *MC4R* (rs12970134, $P = 0.041$). The CLHNS weight association for *KCTD15* (rs29941, $P = 0.034$) was in the opposite direction as previously reported, and did not meet our criteria for replication. For height, we replicated three previously reported associations with *EFEMP1* (rs3791679, $P = 0.0017$), *ZBTB38* (rs6440003, $P = 0.0048$), and *NPPC* (rs6718438, $P = 0.0096$). None of three previously reported associations with waist circumference replicated in the CLHNS. Together, these SNPs explain a small proportion of trait variation (R^2 = 1.9% [BMI], 1.1% [weight], 0.3% [waist circumference], and 2.9% [height]). Our study had 80% power to replicate ($P < 0.05$) SNPs that explained 0.44% of the total variation in anthropometric traits in 1,792 Filipino women after adjustment for covariates.

We additionally examined 20 SNPs reported in a Korean population cohort with suggestive evidence of association ($P < 10^{-4}$) with either BMI or height (83), and 10 SNPs reported in a Japanese cohort with suggestive evidence of association ($P < 10^{-5}$) with height (91). CLHNS data support evidence for association with BMI at the *OTOL1* locus (rs1399903, $P = 0.0097$) and three associations with height at the loci *GPC5* (rs8002779, $P = 0.016$), *HIST1H1PS2* (rs9393681, $P = 0.024$), and *C14orf145* (rs17110818, $P = 0.047$) (**Table 2.7**).

Visually inspecting local HapMap LD for CEU and CHB+JPT at the eight replicated loci, we observed that an inter-population difference in LD appeared to

narrow one of the association regions (**Figure 2.3**). Based on calculations from HapMap-based LD maps, the association signal at *BDNF* appears smaller in genomic size in Asian populations (115 kb [CHB], and 124 kb [JPT]) than in European ones (294 kb [CEU]). This association region contains a non-synonymous amino acid substitution SNP in *BDNF* (rs6265, Val66Met), which is in LD with the most associated CLHNS SNP rs4923461 (r^2 = 0.85 [CEU], 0.64 [CHB+JPT]).

To examine whether genetic effects changed over time due to age or increasingly obesogenic environmental conditions, SNPs at 12 loci previously reported as associated with BMI were further evaluated in longitudinal mixed models using data from seven visits spanning 22 years. Of these 12 SNPs, three were nominally associated with BMI in CLHNS cross-sectional analysis (**Table 2.5**). We tested all 12 SNPs for genotype main effects and then for age-by-genotype and study year-by-genotype interactions, both individually and jointly. Due to confounding of age effects by study year (and vice-versa), when evaluating SNPs that showed genotype interactions involving either age or study year, we performed tests of genotype main effects and age-by-genotype interactions stratified by study year. The most significantly associated SNP from the longitudinal genotype main effect analysis was the *FTO* SNP rs9939609 ($P = 2.0 \times 10^{-5}$) (**Table 2.8**). Three additional SNPs were nominally associated ($P < 0.05$) with BMI: rs4923461 (*BDNF,* $P = 0.0019$); rs17782313 (*MC4R*, $P = 0.0030$); and rs11084753 (*KCTD15*, $P = 0.027$). At all four loci, the directions of effects estimated from longitudinal models were consistent with the 2005 cross-sectional analysis. Only rs11084753 (*KCTD15*) had not shown at least nominal evidence for association in the cross-sectional analysis.

Four of the 12 SNPs showed statistically significant evidence ($P < 0.0042$, considering 12 tests) for an age-by-genotype interaction in the longitudinal analyses: rs4923461 (*BDNF*, $P = 2.4 \times 10^{-6}$); rs9939609 (*FTO*, $P = 3.8 \times 10^{-4}$); rs17782313 (*MC4R*, $P = 6.3 \times 10^{-4}$); and rs7498665 (*SH2B1*, $P = 8.2 \times 10^{-4}$) (**Table 2.9**). Three additional SNPs had nominally significant age-by-genotype interactions: rs6548238 (*TMEM18*, $P = 0.0053$), rs1093839 (*GNPDA2*, $P = 0.0086$), and rs2815752 (*NEGR1*, $P = 0.016$). Except for rs6548238 (*TMEM18*), the same pattern of significance occurred for all of these SNPs when tested for study year-by-genotype interactions in models without an age-by-genotype interaction term (**Table 2.10**). In mixed models that included both age-by-genotype and study year-by-genotype interaction terms, only the age-by-genotype interaction for rs17782313 (*MC4R*) remained nominally significant ($P = 0.046$) (**Table 2.11**).

Further analyses stratified by study year only clearly supported an age-by-genotype interaction for rs17782313 (*MC4R*) (**Figure 2.4**). Specifically, the age-by-genotype interaction coefficients for rs17782313 were consistent across study visits, resulting in slightly increasing main effect estimates of genotype over time in models absent the age-by-genotype interaction term. In contrast, age-by-genotype interaction coefficients for rs4923461 (*BDNF*) were not significantly different from zero at any single study visit, but the effect of genotype, in main effects analyses only, consistently increased over the study visits (**Figure 2.5**). The *FTO* SNP rs9939609 showed no evidence for an age-by-genotype interaction at the first four study visits, but some evidence at the latter three (**Figure 2.6**), and the main effects of genotype increased over the first four study visits, but decreased thereafter. The

patterns were less clear for the other loci that exhibited evidence for age-by-genotype and study year-by-genotype interactions (**Figure 2.7**).

DISCUSSION

We have performed the first GWA scan for anthropometric traits in a cohort from the Philippines, a country undergoing socio-economic and nutrition transition. The strongest signals with suggestive evidence of association in the CLHNS ($P < 10^{-5}$) require confirmation in other studies. Among these signals, a SNP in the *KCNE4* locus (rs1440072) was associated with both BMI and waist circumference. *KCNE4* codes for the potassium voltage-gated channel, Isk-related family, member 4 protein, which acts as an inhibitory subunit to *KCNQ1* (potassium voltage-gated channel, subfamily Q, member 1) (97). *KCNQ1* is expressed in adipose tissue and has been associated with type 2 diabetes in both European and Asian populations (98-100).

We replicated ($P < 0.05$ and consistent direction of effect) 8 of 55 non-independent previously reported SNP-trait associations ($P < 5 \times 10^{-8}$) with BMI, weight, waist circumference, and height, providing further evidence that these loci influence anthropometric trait variation across world populations. We also replicated three signals with suggestive evidence of association ($P < 10^{-4}$) with either BMI or height from another Asian population cohort (83, 91). The longitudinal main effect results for the 12 previously reported BMI SNPs were consistent with the primary cross-sectional (2005 visit) results in that the same SNPs displayed nominal evidence across both approaches. Failure to replicate

additional loci likely reflects modest power of the CLHNS study to detect the signals, but could also indicate that the loci are population-specific or influenced by environmental or dietary exposures that differ between populations.

Inter-population differences in local LD, which are especially pronounced between continental populations (53), may assist in fine-mapping the disease-causing variants in loci implicated across association studies. In some cases, the overlap of LD between two populations may correspond to an area of increased association signal strength, thus suggesting a narrower region of interest than in the discovery population alone (62). We observed an appreciable difference in LD between the European and Asian HapMap populations at the *BDNF* locus. Consistent with previous observations that the LD from the CHB and JPT populations is similar to the CLHNS (66), we observed a putatively smaller *BDNF* association region that still contained a non-synonymous amino acid substitution (rs6265, Val66Met) associated with obesity (101). Further suggestive evidence of *BDNF*'s functional relevance to BMI includes observations that heterozygous and conditional knockout mice develop hyperphagia and obesity (102-104).

Our longitudinal study also supports age-by-genotype and/or study year-by-genotype interactions for three loci previously identified as associated with BMI. Evidence for an age-by-genotype interaction for the *MC4R* SNP rs17782313 was consistent across all study visits, with additive effects of the C allele that increased with age. The effects of genotype did not appear to change over time due to study year, which would reflect the rapidly changing environment in the Philippines during the study period. Conversely, our observations in stratified analyses did not

support an age-by-genotype interaction for the widely studied *BDNF* SNP rs4923461, but we found evidence suggesting increased main effects of genotype (specifically the A allele) over the study time period. Our results suggest that rs4923461 genotype likely interacts with a factor other than age that also changed over the 22-year study period. Dietary and other environmental factors changed considerably between 1983 and 2005 and one or more of these factors may modify the effect of the rs4923461 genotype. Rates of overweight and obesity were initially low, and increased over time, providing substantially more variation in levels of body fat.

The BMI association for the *FTO* SNP rs9939609 over time is even more multifaceted, as we found evidence in the models stratified by study year supporting decreased effects of genotype with age at the later study visits, but no evidence for such an effect during the first 11 years of the study when our subjects were younger. The main effects, absent the age-by-genotype interaction term, of rs9939609 appeared to increase during the first 11 years of the study, suggesting a positive interaction between genotype and study year during this period, but then to steadily decrease to levels near baseline most recently, reflecting the decreased effects of genotype with respect to age in the latter years. Similar to our *BDNF* result, neither the age-by-genotype or study year-by-genotype interaction term was significant in the longitudinal model when both effects were included together. Our previous report of an overall survey year-by-genotype interaction at rs9939609 influencing longitudinal BMI appears to incompletely represent the genetic complexity at this locus (66).

Recently, Hardy *et al.* reported evidence for age-by-genotype interactions influencing BMI for the same two *MC4R* and *FTO* SNPs (105). They observed increasing effects of genotype for both loci through childhood and adolescence up to age 20, and then decreasing effects through adulthood. Because their participants were all the same age at any given year of the study, they could not distinguish between the effects of age and changing environment over time. The CLHNS involved participants with wide ranging ages at baseline (15–48 years) followed for 22 years. While our stratified analyses can begin to separate the nature of these putative interactions, these analyses do not completely remove the mutual confounding of age and study year in our models because the participants necessarily aged during the course of the study.

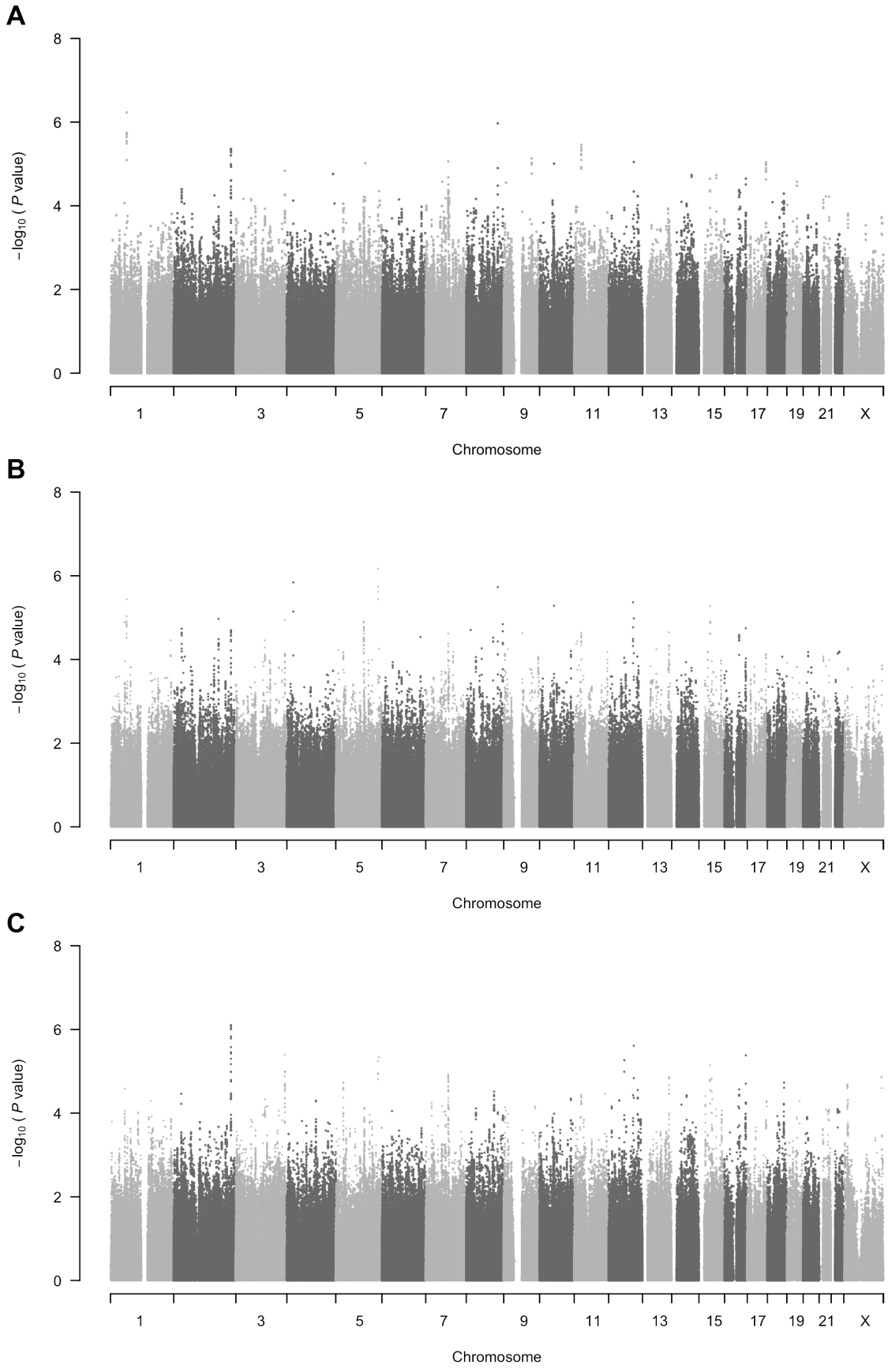
In summary, we found suggestive evidence for additional association signals in a Filipino population cohort, and replicated several previously reported SNP associations with variation in BMI, weight, and height. We also further characterized in a longitudinal setting the *MC4R*, *BDNF*, and *FTO* loci associated with BMI. Together, these results show that multiple genetic risk factors identified in other populations are also associated with anthropometric traits in Filipinos despite a transitioning nutritional environment.

ACKNOWLEDGMENTS

We thank the Office of Population Studies Foundation research and data collection teams. This work was supported by National Institutes of Health grants DK078150, TW05596, HL085144, and TW008288, pilot funds from RR20649, ES10126, and DK56350, and training grants T32 GM007092 to D.C.C.-C. and T32 HL69768 to A.F.M.

DISCLOSURE STATEMENT

The authors declared no conflict of interest.


A
B
C
8
6
4
2
0
−log10 (P value)
1
3
5
7
9
11
13
15
17
19
21
X
Chromosome

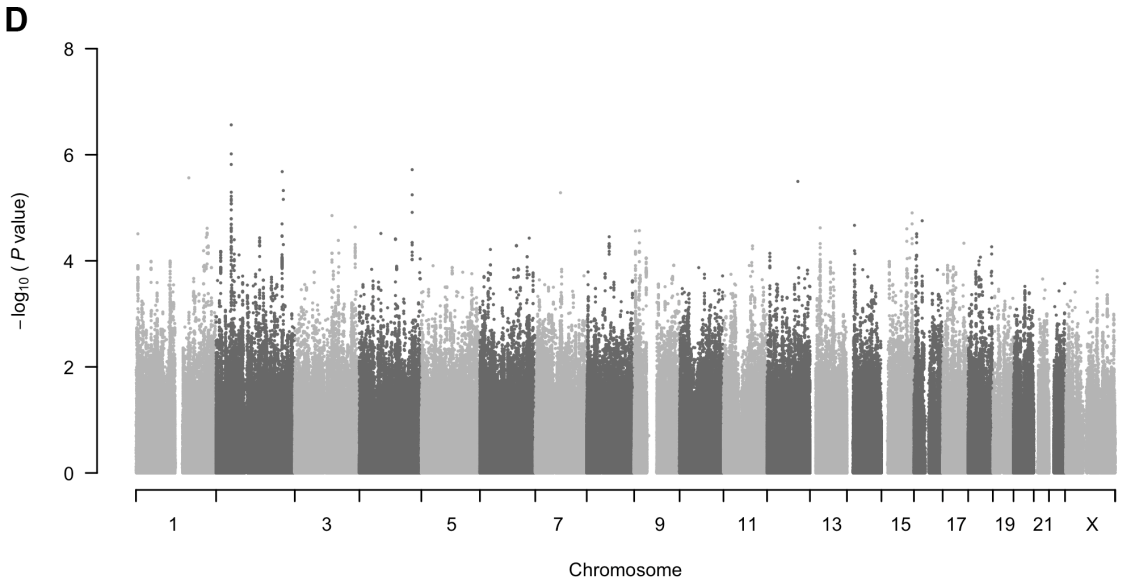


Figure 2.1 – Manhattan plots of associations with current measurements of (*A*) body mass index (natural log-transformed), (*B*) weight, (*C*) waist circumference, and (*D*) average height.

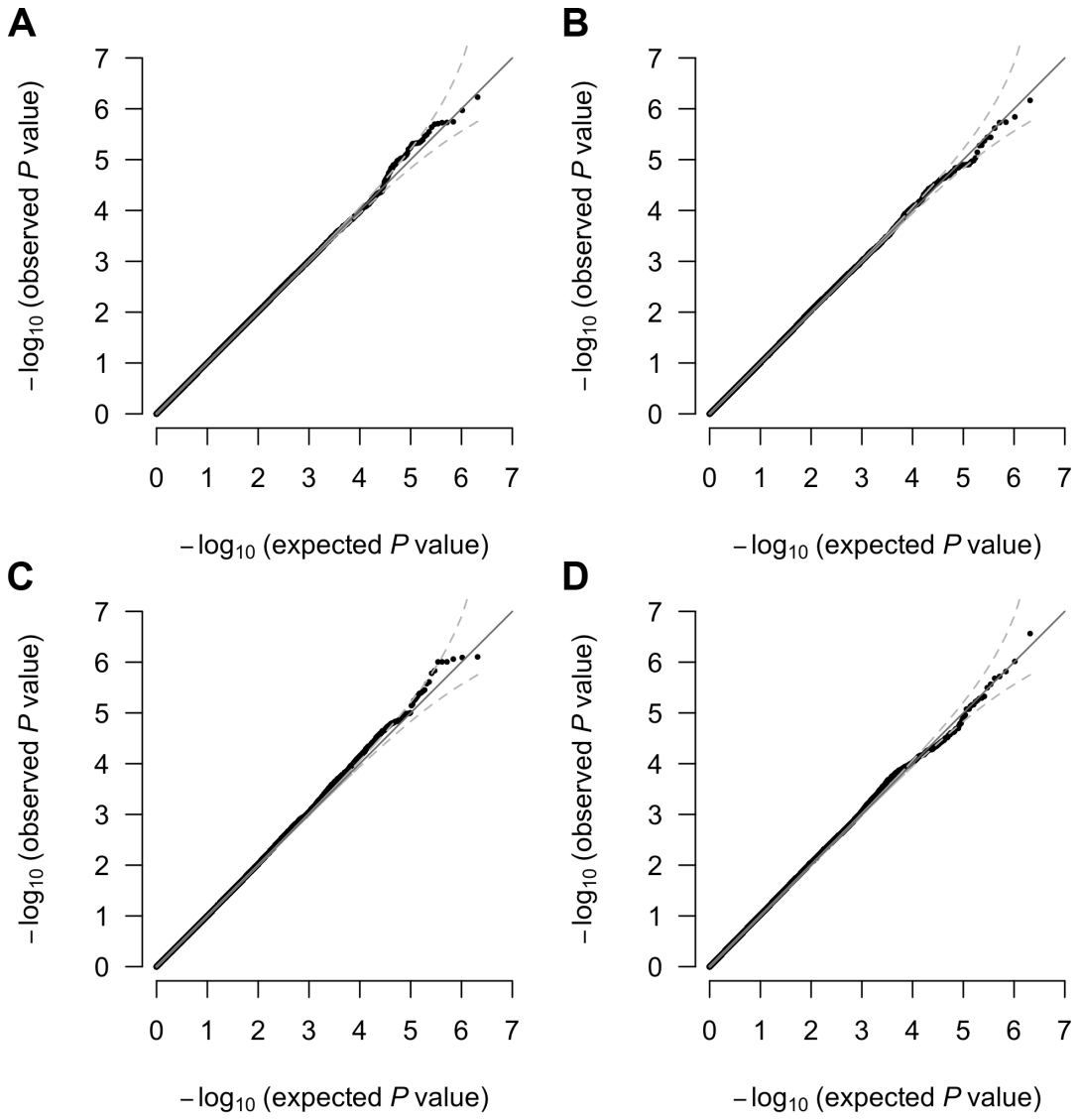


Figure 2.2 – Quantile-quantile plots of the *P* values for associations with current measurements of (*A*) body mass index (natural log-transformed), (*B*) weight, (*C*) waist circumference, and (*D*) average height. The straight dark gray line indicates the expected distribution under the null hypothesis of no association, and the two curving dashed light gray lines indicate the 95% confidence interval.

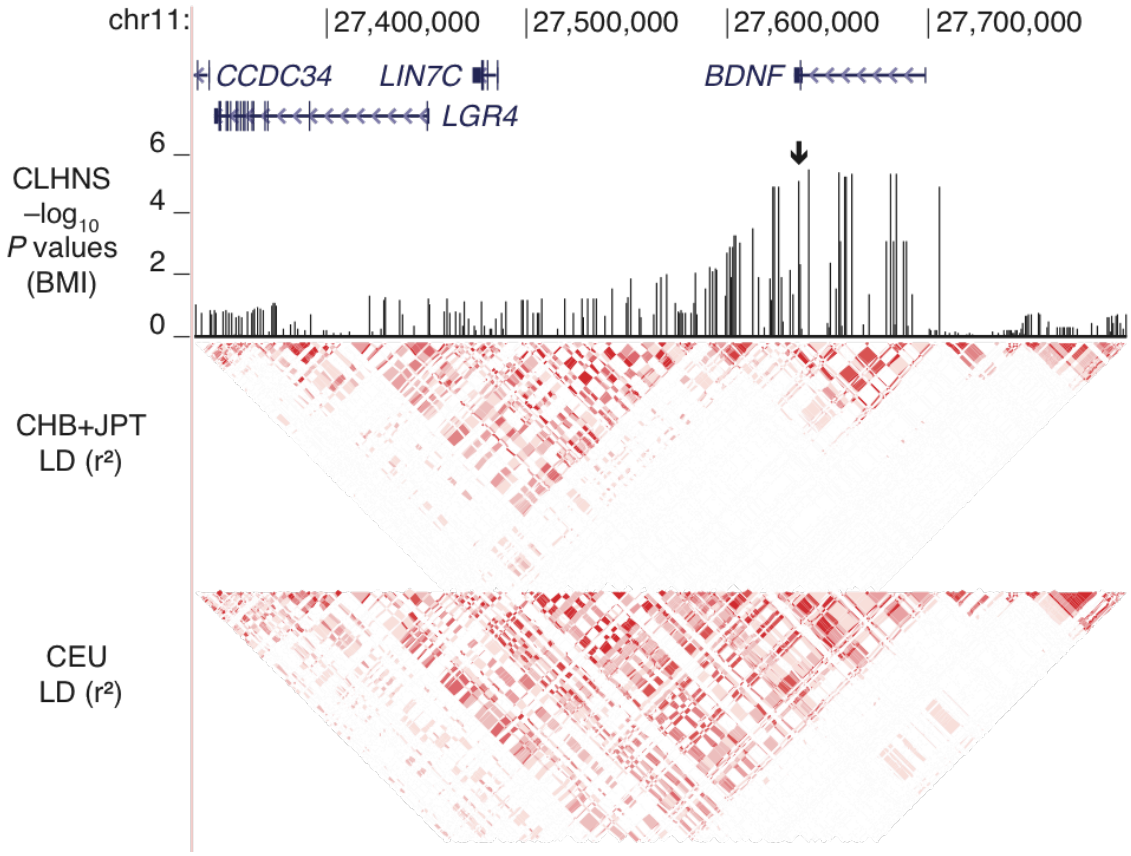


Figure 2.3 – Further localization in the CLHNS of the BMI association signal at the *BDNF* locus. CLHNS association $-\log_{10}$(*P* values) for BMI and nearby genes plotted against pair-wise HapMap Phase II linkage disequilibrium values from phased genotypes in the CEU and the CHB+JPT populations. Dark red indicates r^2 = 1 and white indicates r^2 = 0. An arrow marks the non-synonymous amino acid substitution variant Val66Met (rs6265).

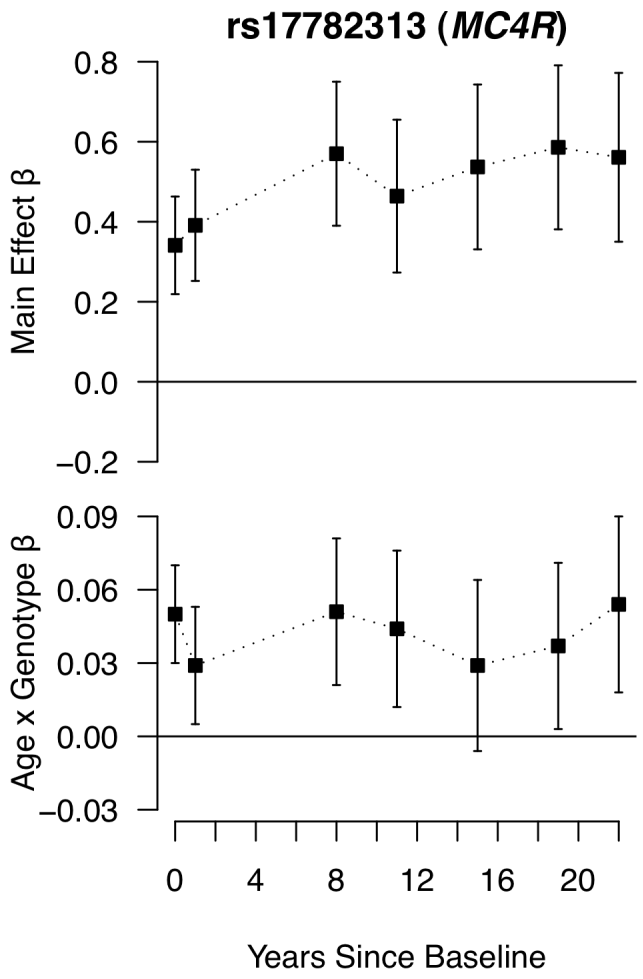


Figure 2.4 – Evidence for a consistent longitudinal age-by-genotype interaction influencing BMI at rs17782313 (*MC4R*). Cross-sectional main effect β coefficients (with no interaction term in the model) and age-by-genotype interaction β coefficients at seven time-points from baseline to 22 years afterward. Error bars represent standard errors. β coefficients are measured in untransformed BMI units (kg/m^2).

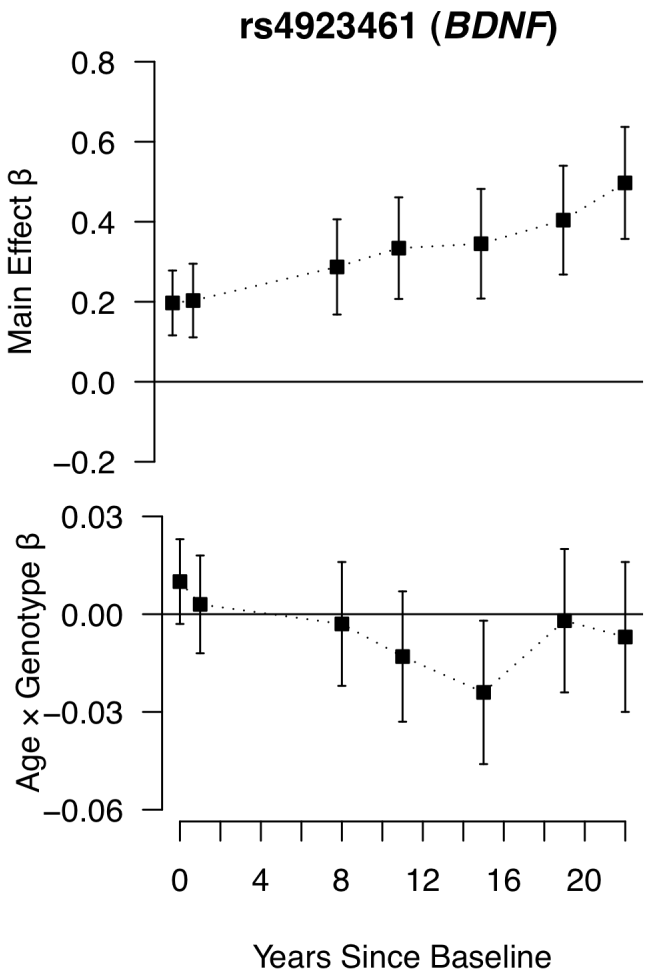


Figure 2.5 – No evidence for a longitudinal age-by-genotype interaction influencing BMI at rs4923461 (*BDNF*). Cross-sectional main effect β coefficients (with no interaction term in the model) and age-by-genotype interaction β coefficients at seven time-points from baseline to 22 years afterward. Error bars represent standard errors. Beta (β) coefficients are measured in untransformed BMI units (kg/m^2).

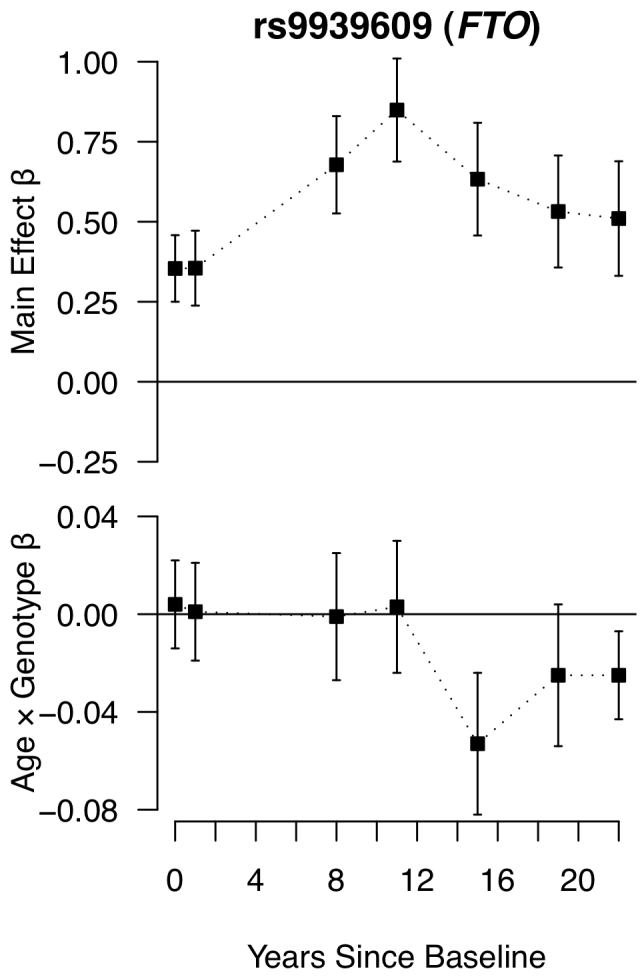


Figure 2.6 – Evidence for a longitudinal age-by-genotype interaction influencing body mass index (BMI) at rs9939609 (*FTO*) beginning after 11 years since baseline. Cross-sectional main effect β coefficients (with no interaction term in the model) and age-by-genotype interaction β coefficients at seven time-points from baseline to 22 years afterward. Error bars represent standard errors. Beta (β) coefficients are measured in untransformed BMI units (kg/m^2).

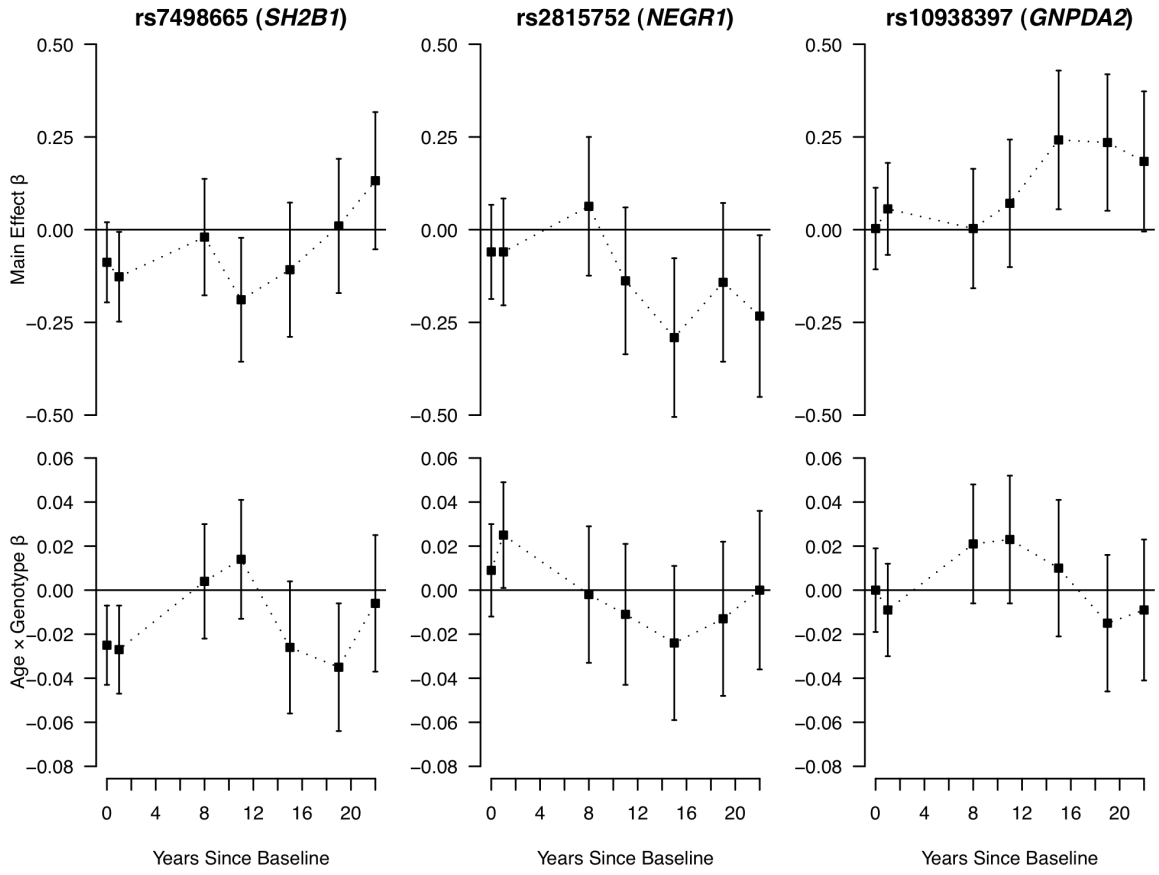


Figure 2.7 – Inconclusive evidence for longitudinal age-by-genotype interactions influencing body mass index (BMI) at rs7498665 (*SH2B1*), rs2815752 (*NEGR1*), and rs10938397 (*GNPDA2*). Cross-sectional main effect β coefficients (with no interaction term in the model) and age-by-genotype interaction Beta (β) coefficients at seven time-points from baseline to 22 years afterward. Error bars represent standard errors. β coefficients are measured in untransformed BMI units (kg/m^2).

Table 2.1 – Pair-wise Pearson correlations of anthropometric traits in the CLHNS cohort

	Waist circumference (cm)	Weight (kg)	BMI log(kg/m^2)
Average height (cm)	0.24	0.44	0.13
Waist circumference (cm)	-	0.88	0.88
Weight (kg)	-	-	0.94

BMI, natural log-transformed body mass index.

Table 2.2 – Associations of trait covariates with outcomes in the CLHNS cohort

		***P* value**			
Trait	**Covariates**	**BMI**	**Weight**	**Waist circumference**	**Average height**
Age	-	**< 0.0001**	**< 0.0001**	0.21	**0.0088**
Age^2	age	**0.010**	**0.0013**	**0.0091**	**0.0002**
Assets	age, age^2	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**
Log(income)	age, age^2	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**
Number of Pregnancies	age, age^2	**< 0.0001**	**< 0.0001**	**0.0002**	**0.016**
Menopausal status	age, age^2	**0.0005**	**0.0003**	0.11	0.075
PC1	age, age^2, all other PCs	0.67	0.069	0.35	**< 0.0001**
PC2	age, age^2, all other PCs	0.27	0.36	0.33	0.99
PC3	age, age^2, all other PCs	0.73	0.72	0.88	0.61
PC4	age, age^2, all other PCs	0.96	0.11	0.27	**< 0.0001**
PC5	age, age^2, all other PCs	**0.0092**	**0.012**	0.36	0.37
PC6	age, age^2, all other PCs	0.38	0.32	0.11	0.51
PC7	age, age^2, all other PCs	0.15	0.091	0.093	0.29
PC8	age, age^2, all other PCs	0.15	0.082	0.06	0.18
PC9	age, age^2, all other PCs	0.64	0.82	0.52	0.33
PC10	age, age^2, all other PCs	0.41	0.60	0.39	0.82

All traits are measured from the 2005 survey except average height, which is the average of the baseline measurement (during initial pregnancy) and seven post-partum measurements. Number of pregnancies is categorized into three groups: 0–4, 5–10, and ≥11. Significant associations ($P < 0.05$) are denoted in boldface. Age, age^2, household assets, natural log-transformed household income, number of pregnancies, menopausal status, and the first five PCs were used as covariates in tests of SNP association. BMI, natural log-transformed body mass index; PC, principal component of population substructure.

Table 2.3 – Demographic and descriptive statistics of the CLHNS cohort

Trait	Value	*n*
Body mass index (kg/m^2)	24.3 ± 4.4	1,780
Waist circumference (cm)	81.1 ± 10.9	1,779
Weight (kg)	55.2 ± 10.9	1,780
Average height (cm)	150.4 ± 4.9	1,792
Age (years)	48.4 ± 6.1	1,792
Number of pregnancies	6.5 ± 3.0	1,792
Menopausal status (yes/no)	687 / 1105	1,792

All values are mean ± s.d. unless specified otherwise.

Table 2.4 – SNPs with suggestive evidence of association ($P < 10^{-5}$) with four anthropometric traits in the CLHNS

SNP	Nearby gene	Chr	Position	Allele 1/2	Allele 1 Freq	Trait	Effect size (β ± s.e.m.)	*P* value
rs17124318	*ATG4C*	1	63,253,318	C/G	0.87	BMI	0.072 ± 0.014	5.91×10^{-7}
rs1440072	*KCNE4*	2	223,644,982	C/T	0.06	BMI	0.055 ± 0.012	4.37×10^{-6}
rs817858	*RAD23B*	9	109,165,632	G/C	0.13	BMI	0.042 ± 0.009	7.42×10^{-6}
rs1406503	*ZNF804B*	7	88,398,814	C/G	0.05	BMI	0.073 ± 0.016	8.69×10^{-6}
rs2373011	*ANKS1B*	12	98,485,480	C/G	0.42	BMI	0.026 ± 0.006	9.00×10^{-6}
rs9906155	*SEPT9*	17	73,213,292	T/C	0.91	BMI	0.045 ± 0.010	9.20×10^{-6}
rs2662467	*LVRN*	5	115,381,567	C/T	0.79	BMI	0.051 ± 0.011	9.62×10^{-6}
rs10740609	*PCDH15*	10	56,460,975	T/A	0.07	BMI	0.055 ± 0.012	9.85×10^{-6}
rs9313296	N/A	5	165,310,277	C/G	0.01	weight	10.004 ± 2.007	6.83×10^{-7}
rs16877106	*ANAPC4*	4	25,012,364	C/T	0.96	weight	5.420 ± 1.121	1.44×10^{-6}
rs907121	N/A	8	123,032,396	C/T	0.57	weight	1.826 ± 0.382	1.87×10^{-6}
rs17124318	*ATG4C*	1	63,253,318	C/G	0.87	weight	3.984 ± 0.858	3.63×10^{-6}
rs11108495	*C12orf55*	12	95,332,048	T/C	0.45	weight	1.554 ± 0.337	4.28×10^{-6}
rs10740609	*PCDH15*	10	56,460,975	T/A	0.07	weight	3.405 ± 0.745	5.21×10^{-6}
rs12594515	*SQRDL*	15	43,772,363	C/G	0.69	weight	1.705 ± 0.373	5.28×10^{-6}
rs1440072	*KCNE4*	2	223,644,982	C/T	0.06	waist	3.662 ± 0.739	7.87×10^{-7}
rs2373011	*ANKS1B*	12	98,485,480	C/G	0.42	waist	1.713 ± 0.363	2.46×10^{-6}
rs9290936	*IL1RAP*	3	191,735,329	G/T	0.61	waist	1.791 ± 0.387	4.03×10^{-6}
rs11647936	*KLHL36*	16	83,243,010	A/T	0.77	waist	2.294 ± 0.497	4.16×10^{-6}
rs13156607	*CCDC99*	5	168,832,565	T/C	0.11	waist	3.909 ± 0.851	4.61×10^{-6}
rs7302017	*PPM1H*	12	61,290,850	G/A	0.46	waist	1.712 ± 0.375	5.47×10^{-6}
rs9313296	N/A	5	165,310,277	C/G	0.01	waist	9.396 ± 2.065	5.72×10^{-6}
rs12594515	*SQRDL*	15	43,772,363	C/G	0.69	waist	1.727 ± 0.384	7.16×10^{-6}
rs3773996	*IL1RAP*	3	191,718,778	A/G	0.63	waist	1.616 ± 0.365	9.96×10^{-6}
rs17818399	*PIGF-CRIPT*	2	46,679,530	C/A	0.17	height	1.127 ± 0.218	2.74×10^{-7}
rs17638464	*FSTL5*	4	162,821,659	G/A	0.83	height	1.382 ± 0.289	1.91×10^{-6}
rs11888559	*CYP20A1*	2	203,873,416	T/C	0.15	height	1.051 ± 0.221	2.08×10^{-6}
rs6670655	*PBX1*	1	163,005,795	T/C	0.18	height	1.258 ± 0.267	2.73×10^{-6}
rs2660869	*LTA4H*	12	95,003,398	C/G	0.02	height	4.387 ± 0.939	3.19×10^{-6}
rs2277912	*FASTKD2*	2	207,342,481	G/T	0.15	height	1.066 ± 0.232	4.73×10^{-6}
rs2691543	*RPL13AP17*	7	77,802,724	T/C	0.97	height	3.449 ± 0.755	5.20×10^{-6}

Chromosomal positions are reported in NCBI Build 36 coordinates. Reported alleles are from the positive strand. The allele that increased the trait value is listed as allele 1. Effect sizes are reported as β coefficients per copy of allele 1 and their

associated standard errors (s.e.m.). BMI is reported in natural log-transformed units.

Chr, chromosome; BMI, natural log-transformed body mass index; N/A, no protein-coding gene within 500 kb of the SNP.

Table 2.5 – CLHNS association for SNPs previously reported to be associated with BMI, weight, and waist circumference

SNP	Nearby gene	Chr	Position	Allele 1/2	Allele 1 Freq	Trait	Effect size (β ± s.e.m.)	*P* value	Ref
rs4923461	*BDNF*	11	27,613,486	A/G	0.52	BMI	0.021 ± 0.006	**2.8×10^{-4}**	(81)
rs17782313	*MC4R*	18	56,002,077	C/T	0.12	BMI	0.023 ± 0.009	**0.0073**	(80)
rs9939609	*FTO*	16	52,378,028	A/T	0.18	BMI	0.020 ± 0.007	**0.0074**	(90)
rs2815752	*NEGR1*	1	72,585,028	A/G	0.88	BMI	-0.015 ± 0.009	0.087	(58)
rs11084753	*KCTD15*	19	39,013,977	G/A	0.30	BMI	-0.014 ± 0.008	0.095	(58)
rs7647305	*DGKG*	3	187,316,984	C/T	0.90	BMI	0.017 ± 0.012	0.16	(81)
rs10938397	*GNPDA2*	4	44,877,284	G/A	0.19	BMI	0.007 ± 0.008	0.40	(58)
rs7498665	*SH2B1*	16	28,790,742	G/A	0.18	BMI	0.006 ± 0.008	0.42	(58)
rs6548238	*TMEM18*	2	624,905	C/T	0.92	BMI	0.007 ± 0.011	0.50	(58)
rs7566605	*INSIG2*	2	118,552,495	C/G	0.45	BMI	0.004 ± 0.006	0.54	(89)
rs1569019	*GPR133*	12	130,142,144	A/C	0.01	BMI	-0.012 ± 0.026	0.64	(88)
rs10838738	*MTCH2*	11	47,619,625	G/A	0.20	BMI	0.001 ± 0.007	0.90	(58)
rs3751812	*FTO*	16	52,375,961	T/G	0.18	weight	1.034 ± 0.442	**0.019**	(81)
rs29941	*KCTD15*	19	39,001,372	A/G	0.79	weight	-0.891 ± 0.421	0.034	(81)
rs12970134	*MC4R*	18	56,035,730	A/G	0.14	weight	0.983 ± 0.480	**0.041**	(81)
rs1077393	*BAT3*	6	31,718,508	A/G	0.44	weight	-0.453 ± 0.352	0.20	(81)
rs2568958	*NEGR1*	1	72,537,704	A/G	0.88	weight	-0.691 ± 0.538	0.20	(81)
rs7561317	*TMEM18*	2	634,953	G/A	0.92	weight	0.561 ± 0.629	0.37	(81)
rs7647305	*DGKG*	3	187,316,984	C/T	0.90	weight	0.324 ± 0.702	0.64	(81)
rs4788102	*SH2B1*	16	28,780,899	A/G	0.18	weight	0.203 ± 0.459	0.66	(81)
rs10913469	*SEC16B*	1	176,180,142	C/T	0.16	weight	0.146 ± 0.505	0.77	(81)
rs7826222	*MSRA*	8	9,897,490	G/C	0.39	waist	0.624 ± 0.392	0.11	(56)
rs12970134	*MC4R*	18	56,035,730	A/G	0.14	waist	0.658 ± 0.494	0.18	(79)
rs987237	*TFAP2B*	6	50,911,009	G/A	0.21	waist	-0.037 ± 0.447	0.93	(56)

Significant associations ($P < 0.05$, consistent direction of effect) are shown in boldface.

Ref, reference; other columns and abbreviations are as described in **Table 2.4**.

Table 2.6 – CLHNS association for SNPs previously reported to be associated with height

SNP	Nearby gene	Chr	Position	Allele 1/2	Allele 1 Freq	Effect size (β ± s.e.m.)	*P* value	Ref
rs3791679	*EFEMP1*	2	55,950,396	A/G	0.29	0.590 ± 0.188	**0.0017**	(78)
rs6440003	*ZBTB38*	3	142,576,899	A/G	0.13	0.672 ± 0.238	**0.0048**	(76)
rs6718438	*NPPC*	2	232,863,810	T/C	0.62	0.427 ± 0.165	**0.0096**	(87)
rs1042725	*HMGA2*	12	64,644,614	C/T	0.18	0.381 ± 0.210	0.07	(76)
rs7846385	*PXMP3*	8	78,322,734	C/T	0.15	0.327 ± 0.219	0.14	(78)
rs678962	*DNM3*	1	170,456,512	G/T	0.13	0.357 ± 0.240	0.14	(78)
rs10946808	*HIST1H1D*	6	26,341,366	A/G	0.48	0.257 ± 0.174	0.14	(77)
rs1046934[a]	*TSEN15*	1	182,290,152	C/A	0.46	0.288 ± 0.199	0.15	(78)
rs4842838	*ADAMTSL3*	15	82,373,128	G/T	0.46	-0.234 ± 0.167	0.16	(85)
rs3118914	*DLEU7*	13	50,014,902	G/T	0.99	1.117 ± 0.823	0.17	(85)
rs4743034	*ZNF462*	9	108,672,174	A/G	0.24	0.238 ± 0.186	0.20	(78)
rs6570508[b]	*GPR126*	6	142,755,535	G/A	0.28	0.310 ± 0.246	0.21	(77)
rs13273123	*PLAG1*	8	57,263,345	A/G	0.95	0.474 ± 0.398	0.23	(83)
rs7153027	*TRIP11*	14	91,496,975	A/C	0.73	0.203 ± 0.177	0.25	(78)
rs12214804[c]	*HMGA1*	6	34,296,844	C/T	0.13	0.338 ± 0.303	0.26	(78)
rs4800148	*CABLES1*	18	18,978,326	A/G	0.88	0.322 ± 0.322	0.32	(78)
rs2282978	*CDK6*	7	92,102,346	C/T	0.11	-0.232 ± 0.248	0.35	(76)
rs4533267	*ADAMTS17*	15	98,603,794	A/G	0.39	0.152 ± 0.168	0.37	(78)
rs11205277	*SF3B4*	1	148,159,496	G/A	0.36	0.277 ± 0.322	0.39	(78)
rs5742692	*IGF1*	12	101,323,728	A/G	0.80	0.159 ± 0.203	0.43	(91)
rs12338076	*QSOX2*	9	138,261,561	C/A	0.21	0.186 ± 0.239	0.43	(91)
rs6060369	*UQCC*	20	33,370,575	C/T	0.44	0.129 ± 0.167	0.44	(62)
rs967417	*BMP2*	20	6,568,893	G/A	0.17	0.165 ± 0.218	0.45	(78)
rs1812175	*HHIP*	4	145,794,294	G/A	0.48	-0.108 ± 0.157	0.49	(78)
rs2844479	HLA class III	6	31,680,935	A/C	0.53	0.110 ± 0.176	0.53	(78)
rs3760318	*ADAP2*	17	26,271,841	G/A	0.75	0.063 ± 0.184	0.73	(78)
rs4713858	*PPARD*	6	35,510,763	G/A	0.52	0.048 ± 0.163	0.77	(78)
rs1635852	*JAZF1*	7	28,155,936	T/C	0.79	-0.037 ± 0.204	0.86	(84)
rs12986413	*DOT1L*	19	2,121,954	T/A	0.38	0.029 ± 0.184	0.87	(77)
rs12198986	*BMP6*	6	7,665,058	A/G	0.21	-0.033 ± 0.211	0.88	(78)
rs7664706[d]	*PRKG2*	4	82,437,781	T/C	0.10	0.020 ± 0.272	0.94	(85)
rs314277	*LIN28B*	6	105,514,355	A/C	0.05	0.029 ± 0.500	0.95	(77)
rs6830062	*LCORL*	4	17,626,828	T/C	0.84	0.011 ± 0.216	0.96	(78)

Columns and abbreviations are as described in **Table 2.4**.

[a]rs1046934 is a proxy for reported SNP rs2274432 (r^2 = 1 [CEU], 0.98 [CHB+JPT], HapMap Phase 2, Release 22).

[b]rs6570508 is a proxy for reported SNP rs4896582 (r^2 = 0.96 [CEU], 0.92 [CHB+JPT]).

[c]rs12214804 is a proxy for reported SNP rs1776897 (r^2 = 1 [CEU, CHB+JPT]).

[d]rs7664706 is a proxy for reported SNP rs2011962 (r^2 = 1 [CEU, CHB+JPT]).

Table 2.7 – Associations in CLHNS cohort of BMI and height SNPs previously reported ($P < 10^{-4}$) in a Korean or Japanese population cohort

SNP	Nearby gene	Chr	Position	Allele 1/2	Allele 1 Freq	Trait	Effect size (β ± s.e.m.)	*P* value	Ref
rs1399903	*OTOL1*	3	162,876,798	C / T	0.16	BMI	0.020 ± 0.008	**0.0097**	(83)
rs16953563	*MAP2K1*	15	64,473,824	A / G	0.37	BMI	0.010 ± 0.006	0.097	(83)
rs11193517	*SORCS1*	10	109,243,058	A / C	0.05	BMI	0.010 ± 0.013	0.42	(83)
rs8002779	*GPC5*	13	90,813,978	G / A	0.54	height	0.388 ± 0.161	**0.016**	(91)
rs9393681	*HIST1H1PS2*	6	26,116,239	T / C	0.27	height	0.413 ± 0.182	**0.024**	(83)
rs17110818	*C14orf145*	14	80,121,106	A / T	0.18	height	0.420 ± 0.211	**0.047**	(83)
rs6561030	*DGKH*	13	41,529,719	C / A	0.37	height	-0.281 ± 0.207	0.17	(91)
rs6563943	*CDH13*	16	82,196,836	A / G	0.40	height	-0.249 ± 0.188	0.19	(91)
rs4751815	*RPL19P16*	10	122,691,495	G / A	0.76	height	-0.243 ± 0.188	0.20	(83)
rs2305707	*CYP19A1*	15	49,356,702	A / G	0.23	height	0.240 ± 0.192	0.21	(91)
rs1136001	*NTAN1*	16	15,039,475	G / T	0.49	height	0.227 ± 0.184	0.22	(91)
rs6483645	*DBX1*	11	20,128,523	G / C	0.62	height	0.173 ± 0.168	0.30	(83)
rs7189843	*PLCG2*	16	80,459,640	G / C	0.50	height	-0.166 ± 0.164	0.31	(83)
rs11228763	*OR9G4*	11	56,267,736	G / A	0.13	height	-0.245 ± 0.242	0.31	(83)
rs162089	*DDX4*	5	55,154,432	C / T	0.37	height	0.160 ± 0.168	0.34	(83)
rs3738814	*ATP13A2*	1	17,204,263	A / G	0.16	height	0.217 ± 0.230	0.34	(91)
rs4886707	*PTPN9*	15	73,542,520	T / C	0.25	height	0.155 ± 0.185	0.40	(91)
rs2517538	*MUC21*	6	31,121,520	C / G	0.40	height	0.132 ± 0.163	0.42	(83)
rs6921309	N/A	6	18,726,714	A / C	0.25	height	-0.111 ± 0.183	0.55	(83)
rs6585827	*PLEKHA1*	10	124,155,605	A / G	0.40	height	0.090 ± 0.175	0.61	(91)
rs17016123	*INPP4B*	4	143,574,788	C / T	0.20	height	0.081 ± 0.198	0.68	(83)
rs16828478	*SHOX2*	3	159,073,925	T / C	0.10	height	-0.098 ± 0.266	0.71	(83)
rs17772163	*ZFAT*	8	135,724,428	T / C	0.52	height	0.044 ± 0.167	0.79	(83)
rs2233969	*PSORS1C1*	6	31,188,411	G / A	0.19	height	0.050 ± 0.204	0.81	(83)
rs2011603	*NCAPG*	4	17,634,582	A / G	0.68	height	0.041 ± 0.173	0.81	(83)
rs11246833	*SFRS8*	12	130,623,673	C / G	0.50	height	0.033 ± 0.163	0.84	(83)
rs2292303	*C12orf48*	12	101,037,661	C / G	0.18	height	0.034 ± 0.208	0.87	(83)
rs3013749	*GLIS1*	1	53,845,347	A / T	0.36	height	0.022 ± 0.166	0.90	(83)
rs11170631	*ATP5G2*	12	52,327,459	C / T	0.32	height	-0.007 ± 0.171	0.97	(91)
rs1175000	*CDCA7L*	7	21,990,565	C / T	0.52	height	0.003 ± 0.212	0.99	(91)

Chromosomal positions are reported in NCBI Build 36 coordinates. Reported alleles are from the positive strand. The allele that increased the trait value in the previous report is listed as allele 1. Effect sizes are additive β coefficients per copy of allele 1 with the standard error of the mean (s.e.m.). BMI is reported in natural log-

transformed units. Significant associations (CLHNS *P* values < 0.05) are shown in boldface. Chr, chromosome; Ref, reference; BMI, natural log-transformed body mass index; N/A, no protein-coding gene within 500 kb of the SNP.

Table 2.8 – Evidence in CLHNS cohort of longitudinal genotype main effect association for previously reported BMI SNPs

SNP	Gene locus	Effect β	*P* value	Reference
rs4923461	*BDNF*	0.014	**0.0019**	(81)
rs17782313	*MC4R*	0.019	**0.0030**	(80)
rs9939609	*FTO*	0.023	**2.0×10^{-5}**	(90)
rs2815752	*NEGR1*	-0.008	0.32	(58)
rs11084753	*KCTD15*	-0.015	**0.027**	(58)
rs7647305	*DGKG*	0.014	0.11	(81)
rs10938397	*GNPDA2*	0.004	0.42	(58)
rs7498665	*SH2B1*	-0.003	0.56	(58)
rs6548238	*TMEM18*	0.011	0.14	(58)
rs7566605	*INSIG2*	0.000	0.90	(89)
rs1569019	*GPR133*	-0.013	0.44	(88)
rs10838738	*MTCH2*	0.001	0.87	(58)

Significant associations ($P < 0.05$) are shown in boldface. Effect β coefficients are measured in untransformed BMI units (kg/m^2).

Table 2.9 – Evidence of longitudinal SNP associations with BMI in the CLHNS

		Age	Genotype	Genotype × Age	
SNP	**Gene locus**	**Effect size (β ± s.e.m.)**	**Effect size (β ± s.e.m.)**	**Effect size (β ± s.e.m.)**	***P* value**
rs4923461	*BDNF*	0.277 ± 0.018	-0.185 ± 0.145	0.013 ± 0.003	**2.4×10^{-6}**
rs17782313	*MC4R*	0.288 ± 0.018	-0.093 ± 0.218	0.015 ± 0.004	**6.3×10^{-4}**
rs9939609	*FTO*	0.312 ± 0.018	0.070 ± 0.185	0.013 ± 0.004	**3.8×10^{-4}**
rs2815752	*NEGR1*	0.309 ± 0.019	0.238 ± 0.226	-0.011 ± 0.004	0.016
rs11084753	*KCTD15*	0.292 ± 0.018	-0.325 ± 0.209	0.000 ± 0.004	0.96
rs7647305	*DGKG*	0.277 ± 0.020	0.039 ± 0.293	0.008 ± 0.006	0.18
rs10938397	*GNPDA2*	0.287 ± 0.018	-0.264 ± 0.196	0.010 ± 0.004	0.0086
rs7498665	*SH2B1*	0.311 ± 0.018	-0.543 ± 0.192	0.013 ± 0.004	**8.2×10^{-4}**
rs6548238	*TMEM18*	0.289 ± 0.018	0.585 ± 0.272	-0.008 ± 0.005	0.0053
rs7566605	*INSIG2*	0.291 ± 0.018	0.012 ± 0.100	0.000 ± 0.004	0.99
rs1569019	*GPR133*	0.277 ± 0.031	-0.084 ± 0.660	-0.007 ± 0.013	0.58
rs10838738	*MTCH2*	0.297 ± 0.019	-0.113 ± 0.181	0.004 ± 0.004	0.31

Beta (β) coefficients are measured in untransformed BMI units (kg/m^2). Significant associations ($P < 0.0042$, considering 12 tests) are shown in boldface.

Table 2.10 – Evidence of study year-by-genotype interactions influencing longitudinal BMI in the CLHNS cohort

		Study Year	Genotype	Study Year × Genotype	
SNP	**Gene locus**	**Effect size (β ± s.e.m.)**	**Effect size (β ± s.e.m.)**	**Effect size (β ± s.e.m.)**	***P* value**
rs4923461	*BDNF*	0.133 ± 0.011	0.182 ± 0.103	0.013 ± 0.003	**1.3×10^{-6}**
rs17782313	*MC4R*	0.143 ± 0.011	0.322 ± 0.156	0.013 ± 0.004	**0.0018**
rs9939609	*FTO*	0.166 ± 0.012	0.424 ± 0.132	0.012 ± 0.003	**3.6×10^{-6}**
rs2815752	*NEGR1*	0.164 ± 0.013	-0.051 ± 0.161	-0.011 ± 0.004	**0.012**
rs11084753	*KCTD15*	0.145 ± 0.011	-0.315 ± 0.149	0.000 ± 0.004	0.93
rs7647305	*DGKG*	0.132 ± 0.015	0.254 ± 0.210	0.007 ± 0.006	0.19
rs10938397	*GNPDA2*	0.141 ± 0.011	0.013 ± 0.140	0.010 ± 0.004	**0.0078**
rs7498665	*SH2B1*	0.167 ± 0.013	-0.204 ± 0.137	0.013 ± 0.004	**3.8×10^{-6}**
rs6548238	*TMEM18*	0.144 ± 0.011	0.358 ± 0.196	-0.008 ± 0.005	0.13
rs7566605	*INSIG2*	0.147 ± 0.011	-0.001 ± 0.103	-0.001 ± 0.003	0.67
rs1569019	*GPR133*	0.136 ± 0.027	-0.303 ± 0.472	-0.005 ± 0.012	0.71
rs10838738	*MTCH2*	0.150 ± 0.012	-0.009 ± 0.129	0.003 ± 0.003	0.36

Significant associations ($P < 0.05$) are shown in boldface. Effect β coefficients are measured in untransformed BMI units (kg/m^2).

Table 2.11 – Age-by-genotype and study year-by-genotype interaction results for longitudinal BMI in the CLHNS cohort

SNP	Gene locus	Age × Genotype Interaction *P*	Study Year × Genotype Interaction *P*
rs4923461	*BDNF*	0.81	0.27
rs17782313	*MC4R*	**0.046**	0.16
rs9939609	*FTO*	0.81	0.70
rs2815752	*NEGR1*	0.78	0.47
rs10938397	*GNPDA2*	0.98	0.66
rs7498665	*SH2B1*	0.40	0.14

Interactions for age-by-genotype effects were adjusted by study year-by-genotype effects and *vice versa*. Significant associations ($P < 0.05$) are shown in boldface

CHAPTER III

Population-specific coding variant underlies genome-wide association with adiponectin level[1]

OVERVIEW

Adiponectin is a protein hormone that can affect major metabolic processes including glucose regulation and fat metabolism. Our previous genome-wide association (GWA) study of circulating plasma adiponectin levels in Filipino women from the Cebu Longitudinal Health and Nutrition Survey (CLHNS) detected a 100 kb two-SNP haplotype at *KNG1–ADIPOQ* associated with reduced adiponectin (frequency = 0.050, $P = 1.8 \times 10^{-25}$). Subsequent genotyping of CLHNS young adult offspring detected an uncommon variant (minor allele frequency [MAF] = 0.025) located ~800 kb from *ADIPOQ* that showed strong association with lower adiponectin levels ($P = 2.7 \times 10^{-15}$, $n = 1{,}695$) and tagged a subset of *KNG1–ADIPOQ* haplotype carriers with even lower adiponectin levels. Sequencing of the *ADIPOQ* coding region detected variant R221S (MAF = 0.015, $P = 2.9 \times 10^{-69}$), which explained 17.1% of the variance in adiponectin levels and largely accounted for the initial GWA signal in Filipinos. R221S was not present in 12,630 Europeans with previously sequenced exons. To explore the mechanism of this substitution, we

---

[1] A version of this work was previously published as Croteau-Chonka DC *et al.* *Hum Mol Genet*. 2012 Jan 15;**21**(2):463-71. Epub 2011 Oct 18.

re-measured adiponectin level in 20 R221S offspring carriers and 20 non-carriers using two alternative antibodies and determined that the presence of R221S resulted in artificially low quantification of adiponectin level using the original immunoassay. These data provide an example of an uncommon variant responsible for a GWA signal and demonstrate that genetic associations with phenotypes measured by antibody-based quantification methods can be affected by uncommon coding SNPs residing in the antibody target region.

INTRODUCTION

Secreted almost exclusively from adipocytes into the bloodstream, adiponectin is an abundant protein hormone that affects metabolic processes including glucose regulation and fat metabolism (31). Circulating plasma adiponectin level is substantially heritable (36-38), and low levels are associated with increased body mass index (BMI) and risk of cardiovascular disease, type 2 diabetes, atherosclerosis, and metabolic syndrome (31, 34, 35, 106, 107). Understanding the genetic basis of adiponectin level may help unravel the etiology of these complex diseases.

The gene *ADIPOQ* that encodes adiponectin is an obvious candidate for influencing adiponectin level and its effects on target tissues. Candidate gene association studies (108-111) as well as genome-wide association (GWA) studies (112-115) of common SNPs (MAF ≥ 0.05) have identified several potentially functional *ADIPOQ* variants. *In vitro* studies have shown that rare (MAF < 0.005) and uncommon (0.005 ≤ MAF < 0.05) missense variants can disrupt the

multimerization necessary for adiponectin secretion (116) and that SNPs in the *ADIPOQ* promoter may alter its transcription (117). Nonetheless, the true causal variants responsible for the previously reported GWA signals at *ADIPOQ* have not yet been established.

We previously reported a 100-kb two-SNP haplotype spanning the *KNG1–ADIPOQ* gene region that is associated with reduced adiponectin level in 1,776 Filipino women from the CLHNS (69). Compared to the strongest single SNP associated with adiponectin in the GWA study (rs864265, MAF = 0.124, $\beta = -0.123$, $P = 3.8 \times 10^{-9}$), the C-T rs11924390-rs864265 haplotype showed substantially stronger evidence of association with a larger effect size (frequency = 0.050, $\beta = -0.385$, $P = 1.8 \times 10^{-25}$). Imputation of SNPs from the 1000 Genomes Pilot Project (June 2010) failed to identify any single SNP with association evidence at this level of significance, suggesting either that the true causal variant was poorly imputed, thus weakening its evidence of association, or that a lower frequency variant or more than one variant may be responsible for the GWA signal (69).

In the current study, we further characterized this *KNG1–ADIPOQ* haplotype association by identifying an uncommon missense variant responsible for much of the signal. In young adult offspring from the CLHNS, in whom the haplotype showed a similar effect (frequency = 0.052, $\beta = -0.386$, $P = 8.7 \times 10^{-32}$) (69), the observation of an additional putative association signal in the region led serendipitously, albeit indirectly, to an explanation for the original GWA signal at *KNG1–ADIPOQ*.

RESULTS

To study the genetic basis of circulating plasma adiponectin level and other metabolic traits, we genotyped using the MetaboChip a set of young adult offspring from the CLHNS. The offspring did not show evidence of population substructure (**Figure 3.1**), and an assessment of global ancestry showed them clustering with other Asian individuals as expected (**Figure 3.2**). Of 140,696 polymorphic MetaboChip SNPs tested for association with natural log-transformed adiponectin level in 1,695 offspring, the two most strongly associated SNPs ($P < 5.0 \times 10^{-8}$) were located approximately 800 kb apart on chromosome 3 (**Figure 3.3**). Located upstream of the *ADIPOQ* gene, the first SNP, rs864265 ($P = 1.0 \times 10^{-13}$, **Table 3.1**), had been detected by our previous GWA study (69). The second SNP, rs117016164, showed stronger evidence of association ($P = 2.7 \times 10^{-15}$, **Table 3.1**), and is located downstream of the *ETV5* gene. Originally identified by the 1000 Genomes Project as a singleton in a CHB HapMap sample (NA18593), rs117016164 had not been detected in CEU, JPT, or YRI samples. This SNP was uncommon in the CLHNS offspring (MAF = 0.025), and exhibited very low linkage disequilibrium (LD) with all other SNPs within the densely typed *ETV5* gene region ($r^2 < 0.1$, chr3:187.2–187.4 Mb in **Figure 3.3**). These data suggested a possible second and novel association signal 800 kb away from the original GWA signal.

To validate the evidence of association of rs117016164 with adiponectin in the CLHNS offspring, we re-genotyped the SNP and, observing concordant genotypes, then performed a permutation test, which generated strong and consistent evidence of association ($P_{perm} < 10^{-9}$). In addition, the 1,764 CLHNS

mothers showed directionally consistent association ($P = 3.2 \times 10^{-7}$, **Table 3.1**). Using a general linear mixed model accounting for sample relatedness, the combined set of offspring and mothers strengthened the evidence of association of rs117016164 with adiponectin levels ($P_{combined} = 7.0 \times 10^{-17}$).

We next examined the relationship between rs117016164 and the two SNPs from the previously identified adiponectin-associated *KNG1–ADIPOQ* haplotype (rs119243940 and rs864265). The SNP rs117016164 was in low LD with both rs119243940 ($D' = 0.38$, $r^2 = 0.004$) and rs864265 ($D' = 0.37$, $r^2 = 0.024$). Carrying the C-T haplotype at *KNG1–ADIPOQ* was associated with lower adiponectin levels in CLHNS offspring ($P = 1.0 \times 10^{-37}$, **Figure 3.4A**). When rs117016164 was included to form three-SNP haplotypes, all carriers of the original two-SNP C-T haplotype still had significantly lower adiponectin levels than non-carriers; they carried haplotype H4 (C-C-T, $P = 2.9 \times 10^{-16}$) and/or haplotype H6 (T-C-T, $P = 2.1 \times 10^{-39}$) (**Table 3.2**). A *post hoc* comparison of these two haplotypes showed that their strengths of association were significantly different ($P = 2.9 \times 10^{-18}$), suggesting that the minor allele (T) of rs117016164 tags a subset of *KNG1-ADIPOQ* haplotype carriers with even lower adiponectin levels. A consistent pattern of strong haplotype association also was observed in the mothers (**Table 3.3**). These data suggested that the H4 and H6 haplotypes together likely harbored two or more SNPs that might account for the observed evidence of association at the *KNG1–ADIPOQ* locus.

Three pieces of evidence led us to hypothesize that coding variants in *ADIPOQ* could explain the observed haplotype associations. First, while *ETV5* was a plausible candidate gene based on a nearby SNP previously reported to be

associated with BMI (81) and the inverse relationship between adiponectin level and BMI, rs117016164 was not located in a conserved region or a regulatory element predicted based on chromatin state. Second, the evidence of long haplotypes linking rs117016164 to the two adiponectin-associated SNPs in the *KNG1-ADIPOQ* region reinforced the possibility that causal variants could be located in other genes in the region; among these, *ADIPOQ* was the strongest biological candidate. Third, while decreased gene expression by a non-coding regulatory variant is a more likely scenario at many GWA loci, amino acid substitutions in ADIPOQ have been shown to disrupt the multimerization of the adiponectin protein necessary for its secretion into the bloodstream (116). We hypothesized then that one or more low frequency coding variants in *ADIPOQ* could have large effects on adiponectin levels consistent with the observed associations.

To identify rare or uncommon coding variants in *ADIPOQ*, we sequenced the translated exons of the gene in 47 CLHNS offspring, a sample set enriched for carriers of the H6 and H4 haplotypes. We observed three variants in the coding region, each in at least one sample: G15G (rs2241766), a common variant that previously showed modest association with adiponectin in the CLHNS mothers (P = 0.0077) (69); G48D, a novel missense variant in one individual with average adiponectin level (2.15 μg/mL); and R221S, a missense variant first reported in Japanese (118). To help predict their potential impact on protein structure and function, we performed bioinformatic annotation of the two observed missense variants. G48D was predicted by Polymorphism Phenotyping (PolyPhen) to be "likely damaging" with a score of 0.997 and by Sorting Intolerant from Tolerant

(SIFT) to "affect protein function" with a score of 0.00. R221S was predicted by PolyPhen to be "likely benign" with a score of 0.000 and by SIFT to be "tolerated" with a score of 0.41. The difference in the predicted effects of the two variants reflects that ADIPOQ amino acid 48 is more highly conserved across related protein sequences, and therefore more sensitive to substitution by an unlike residue. Eight other previously reported missense variants (108, 109, 118-121) were not found in these 47 offspring samples, including G45R, G84R, G90S, R92X, Y111H, R112C, I164T, and H241P. However, as the I164T substitution was uncommon (MAF = 0.009) and strongly associated with lower adiponectin in Japanese (111), we included this SNP in follow-up genotyping.

We directly genotyped three candidate coding variants (G48D, I164T, and R221S) in the complete sets of offspring and mothers, and tested each for association with plasma adiponectin level. No additional G48D carriers were identified, and I164T was monomorphic in both the mothers and offspring. In contrast, R221S was present in the offspring with MAF = 0.015, and carriers had significantly lower adiponectin levels than non-carriers ($P = 2.9 \times 10^{-69}$) (**Figure 3.4B**, **Table 3.1**). The *ADIPOQ* R221S variant was in low LD with the *ETV5* variant (rs117016164) ($D' = 0.54$, $r^2 = 0.17$). We observed similar evidence of R221S association in the mothers ($P = 2.0 \times 10^{-53}$) (**Figure 3.5**, **Table 3.1**), and in the combined mothers and offspring ($P_{combined} = 4.3 \times 10^{-99}$) (**Table 3.1**). R221S explained 17.1% and 13.2% of phenotypic variation in the offspring and mothers, respectively. Thus, only one of the three observed non-synonymous amino acid

substitution variants in *ADIPOQ* (R221S) was strongly associated with adiponectin in Filipinos.

To determine the relationship between the uncommon coding variant R221S and the associated haplotypes, we performed conditional analyses. Conditioning on R221S, offspring carriers of the two-SNP C-T *KNG1–ADIPOQ* haplotype still had significantly lower adiponectin levels, but the association was substantially attenuated ($P = 5.6 \times 10^{-10}$, **Figure 3.4C**). Carriers of the three-SNP H6 haplotype showed greatly attenuated association ($P = 0.030$), while carriers of H4 were still significantly associated with lower adiponectin levels ($P = 1.3 \times 10^{-8}$). The mothers showed similar attenuations (**Figure 3.5C** and data not shown). We also re-phased haplotypes in the mothers after including the R221S variant and found that it segregated very infrequently (at most twice) on any haplotypes other than H4 and H6. These results suggested not only that the H6 haplotype association was largely explained by R221S, but also that one or more additional adiponectin-associated variants remained to be identified on the H4 haplotype, as its association was only partially explained by R221S.

In the CLHNS samples, adiponectin level shows modest but significant negative correlation ($P < 0.0001$) with multiple metabolic traits, including BMI, waist circumference, fasting glucose and insulin, and homeostasis model assessment indices of insulin sensitivity and beta cell function (69). We tested R221S in both the mothers and offspring for association with these traits and lipid profiles and did not observe significant evidence of association ($P = 0.13–0.95$ in offspring and $P = 0.20–0.98$ in mothers, data not shown). In comparison, the *ADIPOQ* SNP rs864265

from the original GWA signal had shown nominal association ($P < 0.05$) in the offspring with several of these traits (69). With the comparatively smaller MAF of R221S, our study had < 10% power to detect such associations with comparable effect sizes at $P < 0.05$. Given prior *in vitro* biochemical evidence that R221S does not affect adiponectin multimer formation or secretion (116), we hypothesized that the R221S substitution affected the original antibody detection of adiponectin, resulting in artificially low levels among carriers. The original ELISA consisted of two monoclonal antibodies targeting amino acids 104–244 of *ADIPOQ*. We re-measured plasma adiponectin level in twenty R221S carriers and twenty sex- and BMI-matched non-carriers using two alternative polyclonal antibodies that recognize epitopes within amino acids 17–27 and 56–70, neither of which should be affected by amino acid 221. Adiponectin measurements were not consistent between the original ELISA and each of the alternative antibodies (**Figure 3.6** and data not shown). The subset of R221S carriers had lower adiponectin levels than non-carriers based on the original ELISA, but carriers and non-carriers showed similar distributions with the alternative antibodies. These results suggested that R221S changed the structure of adiponectin enough to alter the binding affinity of the monoclonal antibody in the original ELISA. These results confirmed that the very strong statistical association of R221S with lower adiponectin level was likely due to a methodological artifact of antibody specificity.

We then re-evaluated the evidence of adiponectin association after conditioning on R221S. Overall, there were no substantial changes in any SNP associations at loci other than *KNG1–ADIPOQ* in the offspring (data not shown) or mothers (**Figure 3.7**), including at *CDH13*, a previously reported GWA signal in the CLHNS (69); similar results were observed after removing R221S carriers from analysis (data not shown and **Figure 3.7**). To examine the associations of additional variants in the 187.2 to 188.2 Mb region of chromosome 3, we re-imputed in the mothers SNPs from the November 2010 release of the 1000 Genomes Project (**Figure 3.8**). Prior to conditioning on R221S, the most significantly associated SNP in the region was rs864265 ($P = 2.2 \times 10^{-10}$). After conditioning, this signal was highly attenuated ($P = 0.076$) and the most significantly associated SNP was rs58575091 ($P = 3.7 \times 10^{-5}$). We re-evaluated eight *ADIPOQ* SNPs that were identified as candidate functional variants in the literature (69), and of those, only rs266729 and rs182052 still showed nominal association ($P < 0.05$) after conditioning on R221S (**Table 3.4**). These associations remained significant after conditioning on both R221S and copies of the C-T haplotype (rs266729, $P = 1.8 \times 10^{-4}$ and rs182052, $P = 4.6 \times 10^{-4}$). When the two SNPs were included in the model together, rs266729 was modestly more significant than rs182052 ($P = 0.067$ versus $P = 0.22$). Comparable results were obtained when removing R221S carriers from analysis (data not shown). Thus, after conditioning on R221S, moderate evidence of association with adiponectin level remained. The conditional analyses suggested that two or more additional variants still play a role in trait variability in this study.

DISCUSSION

Follow-up of our previous GWA study in the CLHNS demonstrates that the two-SNP haplotype in the *KNG1–ADIPOQ* region associated with lower plasma adiponectin level can be attributed to a coding variant that affected antibody specificity. This uncommon non-synonymous amino acid substitution variant (R221S, MAF = 0.015) was responsible for most of the initial GWA signal, which was based on a survey of common variants. R221S explained 17.1% and 13.2% of adiponectin variation in the offspring and mothers, respectively. In comparison, the index SNP from the original GWA study (rs864265, MAF = 0.129) explained 3.2% and 2.2% of variation, respectively. We discovered R221S by detecting in the offspring a long three-SNP haplotype created by another uncommon SNP nearly 800 kb away from *ADIPOQ*. This SNP (rs117016164) was genotyped fortuitously, as it was included on the MetaboChip to fine-map the *ETV5* gene region previously associated with BMI (81). The location of rs117016164 downstream of *ETV5* and its lack of overlap with predicted regulatory elements in a relevant cell type, such as adipocytes, made *ETV5* a less compelling candidate gene for regulating adiponectin level. Furthermore, rs117016164 was the only variant in the *ETV5* region to show evidence of association comparable to that of rs864265 near *ADIPOQ*, suggesting the presence of a long haplotype spanning the entire region. Prior biochemical experiments have shown that ADIPOQ amino acid substitutions can disrupt the multimer aggregation required for adiponectin secretion (116). Targeted sequencing of *ADIPOQ* coding regions in carriers of the associated haplotypes then revealed R221S, which segregated primarily on the two long haplotypes associated

with lower adiponectin level in the entire cohort. Therefore, the causal variant was not only detectable 800 kb away, but also nearly ten times less frequent and explained a far greater proportion of phenotypic variance than the GWA index SNP, in line with a previously proposed association causation scenario (122). The proportion of reported GWAS signals explained by uncommon variants remains to be determined, but R221S is a clear example of one.

Because R221S has been reported previously only in Japanese (118), Korean (123, 124), and Japanese-Brazilian (125) cohorts, the variant may be specific to Asian-ancestry populations. This R221S variant (CGT to AGT) also was observed to be monomorphic in 12,630 individuals of European ancestry with sequenced exons (M.R. Nelson, M.G. Ehm, D. Wegmann, P. St. Jean, C. Verzilli, J. Shen, Z. Tang, S.-A. Bacanu, D. Fraser, L. Warren, J. Aponte, D. Kessner, J. Li, S. Zöllner, Y. Li, L. Li, P. Woollard, S. Topp, M.D. Hall, K. Nangle, G. Abecasis, J. Wang, L.R. Cardon, J. Novembre, J.C. Whittaker, S.L. Chissoe, V. Mooser, manuscript in preparation). A different *ADIPOQ* missense variant (I164T) with low frequency in Japanese (MAF = 0.009) has been shown to be associated with lower adiponectin levels (111), consistent with biochemical evidence that the I164T-containing adiponectin does not form trimers or high molecular weight multimers and impairs secretion from the adipocyte (116). However, we did not observe I164T or any other of seven known missense variants among the CLHNS samples.

The observed data did not support a biological role for R221S in lowering adiponectin level. Despite explaining approximately 13–17% of variation in adiponectin levels in the CLHNS, we observed no evidence of association of R221S

with metabolic traits or with waist circumference, which is a well-known predictor of adiponectin levels. This result could reflect limited power. R221S also did not show any deficiency in formation of adiponectin trimers, hexamers, or high molecular weight multimers *in vitro* (116), and bioinformatic annotations that suggested R221S would not substantially alter the protein's conformation. Among the previous reports of R221S, adiponectin levels in carriers were significantly lower (125), indistinguishable from non-carriers (118), or not reported due to low MAF (123, 124). We hypothesized that the R221S variant is located on the surface of the protein at a position that may alter the epitope recognized by the adiponectin antibody. Two previous reports support this scenario. The first describes an African-American-specific K29M amino acid substitution in the soluble circulating intracellular adhesion molecule-1 (sICAM-1) protein that resulted in the errant measurement of sICAM-1 by an ELISA using monoclonal antibodies to the region (126). The second describes a GWA study of European-ancestry samples in which a V32M substitution in the natriuretic peptide precursor A gene was strongly associated with mature plasma atrial natriuretic peptide only if measured by an immunoassay targeting the N-terminal epitope (127). Altered trait detection could be responsible for other associations of non-synonymous *ADIPOQ* variants with adiponectin or for any other genetic associations with protein traits assayed by monoclonal antibodies binding the region containing the mutation. Of further concern to the study of complex traits on a genome-wide level is that a study's power to detect associations at other loci could be affected by samples with noisy phenotypes, especially if they lie in the extreme tails of the distribution.

The novel singleton missense variant (G48D) we identified in *ADIPOQ* is located in the conserved glycine repeat motif region. Other nearby glycine-to-aspartic acid substitutions in adiponectin (116) and the homologous mannose-binding protein (128) showed deficiency in the formation of high molecular weight multimers, suggesting that G48D may also disrupt multimer formation and reduce the stability of its collagen helix domain. However, as the single CLHNS G48D carrier had only slightly below average adiponectin levels, the variant may have no effect.

We hypothesize that two or more variants of considerably smaller effect explain the remaining GWA signal at *ADIPOQ* after adjusting for R221S. The true causal variants in the remaining signal could be the same as those underlying *ADIPOQ* signals from GWA studies of European samples (112-115). The previously described candidate SNP rs266729 previously showed evidence of association in the CLHNS (69), and has been predicted to alter an SP1 transcription factor binding site (129). It is also part of a three-SNP haplotype shown to have an *in vitro* effect on *ADIPOQ* promoter activity and on DNA binding activity of nuclear proteins (117). Other putative regulatory regions near *ADIPOQ* (130-132) also may be promising targets for additional variants. Ultimately, both more comprehensive sequencing of the *ADIPOQ* gene region and improved imputation of SNPs from the 1000 Genomes Project may be needed to systematically identify additional candidate causal variants, especially of low frequency.

In summary, we characterized a GWA signal for lower plasma adiponectin level in a Filipino cohort and showed that a population-specific uncommon missense variant in the adiponectin-coding gene resulted in artificially low adiponectin measurements in a subset of samples and thus was largely responsible for the observed association signal detected with common variants. Our work highlights how genetic associations with phenotypes measured by antibody-based quantification methods can be affected by uncommon coding SNPs residing in the antibody target region. We also demonstrated the value of targeted sequencing for discovering a variant with potential functional relevance not yet present in public SNP databases. While these findings as a whole do not elucidate the biology underlying the genetics of adiponectin level, they illustrate an unusual combination of association phenomena that may inform and caution other researchers performing large-scale genetic association analyses.

MATERIALS AND METHODS

Subjects

The CLHNS participants available for this study included 1,798 mothers and 1,779 male and female young adult offspring from a 1983 to 1984 Filipino birth cohort (63). Trained field staff conducted in-home interviews and collected quantitative anthropometric measurements, blood samples, and comprehensive environmental data (available on-line at http://www.cpc.unc.edu/projects/cebu/). Outcome and covariate measures were taken from the 2005 survey. Informed consent was obtained from all CLHNS subjects, and the study protocol was approved

by the University of North Carolina Institutional Review Board for the Protection of Human Subjects. Basic descriptive characteristics of the mothers and offspring are summarized in **Table 3.5**.

Genotyping and imputation

The 1,779 CLHNS offspring were genotyped for 196,725 SNPs on the MetaboChip (Illumina, San Diego, CA, USA), a custom array of SNPs designed to replicate and fine-map loci associated with metabolic traits. Genotyping was performed by the Mammalian Genotyping Core at the University of North Carolina at Chapel Hill (UNC-CH) using the protocol recommended by the manufacturer. Individual sample success rates exceeded 98.6%. Due to poor genotype clustering, 8,610 SNPs were removed from analysis. SNP quality control filtering was then performed using PLINK v1.07 (133). We removed 1,652 SNPs for success rates ≤ 97%, 126 SNPs for deviation from Hardy-Weinberg equilibrium ($P < 10^{-6}$), 228 SNPs for Mendelian inheritance errors (combined ≥ 3 discrepancies among 79 duplicate pairs and 4 HapMap CEPH trios genotyped on the MetaboChip), 21 SNPs for ≥ 3 genotype discrepancies with available HapMap genotypes for the 4 CEPH trios, and 45,397 SNPs that were monomorphic. Five SNPs did not pass more than one of these filters, so the final set of high-quality MetaboChip data included 140,696 SNPs.

Using MACH version 1.0 (92), additional genotypes of SNPs in the 187.2 to 188.2 Mb region of chromosome 3 from the November 2010 release (20101123) of the 1000 Genomes Project were imputed in the mothers based on reference haplotypes from CEU, CHB, and JPT samples. MACH was also used to generate best-

guess phased haplotypes (those with highest posterior probabilities) of 351 directly genotyped SNPs comprising the *ETV5-KNG1-ADIPOQ* gene region in the mothers (69) and of 1,148 SNPs in the offspring. To infer high-quality haplotypes, we specified 200 rounds of Markov sampling considering 500 haplotype states when updating each individual.

Additional SNPs were genotyped using TaqMan allelic discrimination (Applied Biosystems, Foster City, CA, USA), including *ADIPOQ* missense SNPs G48D (ss469105329/rs182223755), I164T (ss469105330/rs185847354), and R221S (ss469105331/rs138773406), and MetaboChip *ETV5* SNP rs117016164 (chr3:187,239,040, hg18). Primer sequences are provided in **Table 3.6**. The two *KNG1-ADIPOQ* SNPs rs119243940 and rs864265 were previously genotyped (69). Success rates were > 95% and all SNP genotypes were consistent with Hardy-Weinberg equilibrium ($P > 0.05$). The non-missing TaqMan genotypes of rs117016164 had 99.8% concordance with the original MetaboChip genotypes. Bioinformatic annotations were assigned to missense variants in the adiponectin protein (Uniprot ID: Q15848) using PolyPhen (version 2) (134) and SIFT (135), each using default parameters.

Sequencing

We sequenced the two translated exons of *ADIPOQ* in 47 offspring samples by PCR amplification and Sanger sequencing. We chose all samples with adiponectin residuals > 2.5 standard deviations above or below the population mean. Primer sequences were previously published (118). Sequencing was

performed by the UNC-CH Genome Analysis Facility on a 3730xl DNA Analyzer (Applied Biosystems).

Statistical analyses

Adiponectin levels were natural log-transformed to approximate normality and SNPs were tested for association using multivariable linear regression models in PLINK. We assumed an additive mode of inheritance reporting β coefficients representing the estimated change in mean transformed trait value due to each additional copy of the minor allele. In the offspring, one member of each of five twin pairs were removed and 79 additional samples were excluded because they lacked non-pregnant measures of BMI in 2005. The final sample sets consisted of 1,695 non-sibling offspring and 1,764 mothers with complete sets of outcomes and covariates. Sex and BMI covariates were significantly associated ($P < 0.05$) with plasma adiponectin level in the offspring (**Table 3.7**). In the mothers, age, median-centered age^2, household assets, natural log-transformed household income and menopausal status during the 2005 survey were included as covariates, as previously described (69). For combined analysis, age, median-centered age^2, household assets, natural log-transformed household income, and BMI during the 2005 survey were used as covariates. Combined analysis of selected SNPs in the offspring and mothers together was performed in SAS version 9.3 (SAS Inc., Cary, NC) using a general linear mixed model that accounted for the relatedness between mother-child pairs. Conditional analyses of directly genotyped SNPs in the offspring and mothers were performed in SAS using linear regression. Conditional analyses of

imputed SNPs in the mothers adjusting for R221S were performed using MACH2QTL (92). Haplotype association analyses in the offspring and mothers were performed in SAS. All conditional and haplotype analyses used the corresponding outcome and covariates from the main association analyses.

To capture population substructure in the offspring, we constructed 10 principal components (PCs) using the software EIGENSOFT (93, 94). We first identified a set of 40,239 independent MetaboChip SNPs (estimated $r^2 < 0.001$ between all pairs of SNPs within 1 Mb windows) with observed MAF > 0.05. PCs were then constructed in 1,670 CHLNS offspring who had pair-wise identity-by-descent ($\hat{\pi}$) < 0.1 with all other samples, as estimated using PLINK. The remaining 104 offspring who had $\hat{\pi} \geq 0.1$ with any other sample(s) were then projected onto these PCs. **Figure 3.1** shows the first two PCs in all 1,774 offspring samples. We also performed a complementary analysis of global ancestry in these samples (**Figure 3.2**) by re-generating PCs using a subset of 24,033 SNPs shared with 165 CEU, 84 CHB, 86 JPT, and 167 YRI HapMap Phase III samples (genotypes available at ftp://ftp.ncbi.nlm.nih.gov/hapmap/genotypes/hapmap3/plink_format/draft_2/). The construction of PCs in the CLHNS mothers has been described previously (67). No PCs were associated with adiponectin level in either the offspring (**Table 3.7**) or mothers (69).

Quantification of adiponectin level

Adiponectin level in plasma originally was measured in the offspring and mothers with a commercially available enzyme-linked immunosorbent assay (ELISA) kit (R&D Systems, Minneapolis, MN, USA, #DY1065) consisting of two monoclonal antibodies that recognize epitopes within amino acids 104–244. Selected offspring samples (n = 40) re-measured by Western blot, included 20 R221S carriers and 20 sex- and BMI-matched non-carriers. The distributions of adiponectin and BMI were comparable between these two subsets and the respective groups from which they were drawn (data not shown). Plasma proteins were resolved by polyacrylamide gel electrophoresis and transferred to polyvinylidene difluoride membranes (Invitrogen, Carlsbad, CA, USA). Duplicate membranes were separately probed with rabbit polyclonal antisera generated by peptide antigens DQETTTQGPGV (136) and DGRDGTPGEKGEKGD (Sigma, St. Louis, MO, USA), corresponding to amino acids 17–27 and 56–70, respectively. Alexa Fluor 680 conjugated goat anti-rabbit IgG was then used as a secondary antibody (Invitrogen). Proteins were detected using an Odyssey imaging system (LI-COR, Lincoln, NE, USA) and the adiponectin bands (30 kDa) were quantified by densitometry using the LI-COR software. The adiponectin standards used for each assay did not allow direct comparison of the original and alternative measurements. Thus, all densitometry readings using a given antibody were natural log-transformed and zero-centered on the mean value of R221S non-carriers from that group.

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health grants [DK078150, TW05596, HL085144, K01DK075573, and P30ES10126], pilot funds [RR20649, ES10126, and DK56350], and a training grant [T32 GM007092]. We thank the research and data collection teams from the Office of Population Studies Foundation as well as the study participants who generously provided their time. We also thank Matthew Nelson, Margaret Ehm, Liling Warren, and Dawn Waterworth for sharing pre-publication sequencing data, as well as Amanda Floyd Beaty and Michael Andre of the UNC Mammalian Genotyping Core for their genotyping assistance.

DISCLOSURE STATEMENT

The authors declared no conflicts of interest.

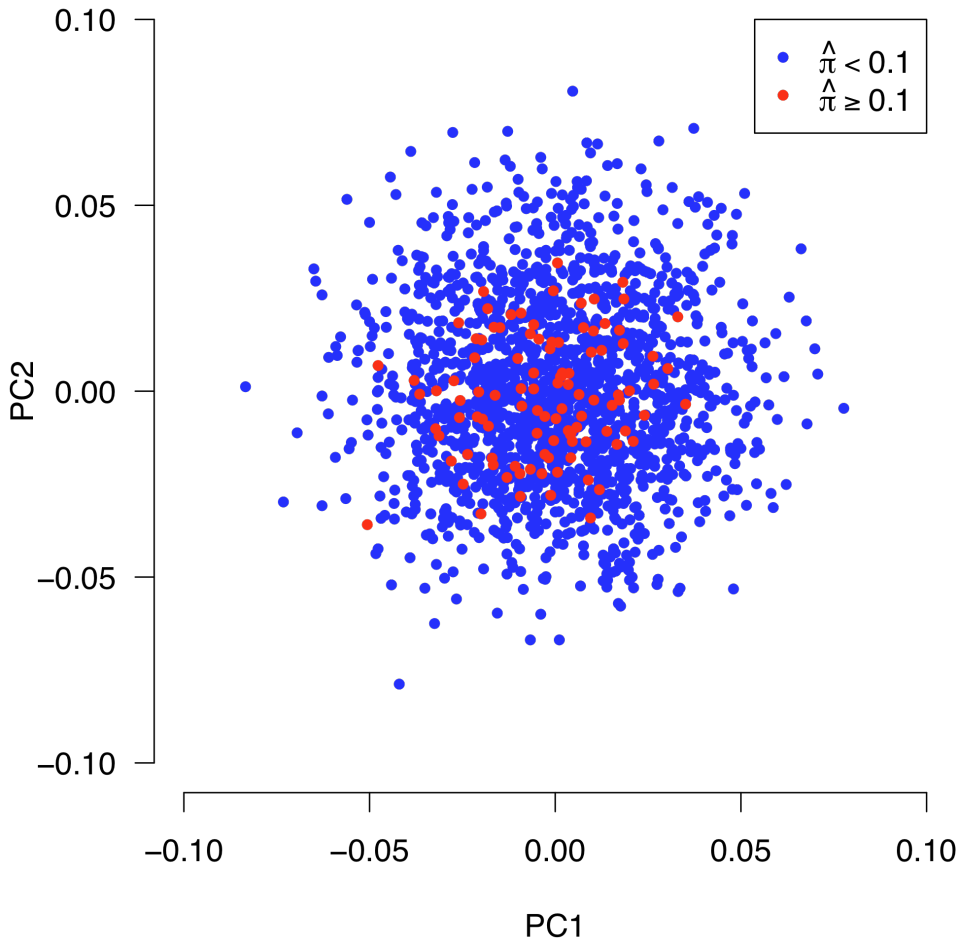


Figure 3.1 – No evidence of population substructure in CLHNS offspring. Plot of the first two principal components of population substructure in 1,774 CLHNS offspring. Samples with pair-wise identity-by-descent ($\hat{\pi}$) < 0.1 with all other samples are marked in blue and samples with $\hat{\pi} \geq 0.1$ with any other sample(s) are marked in red.

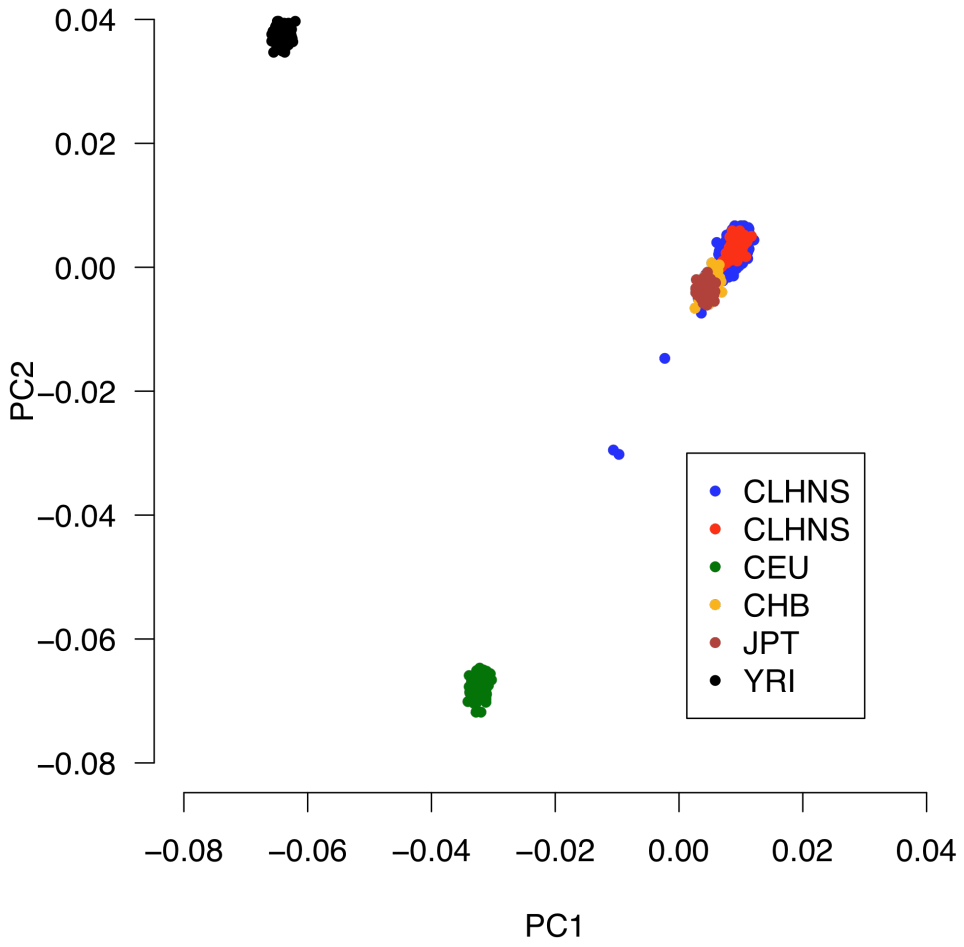


Figure 3.2 – CLHNS offspring cluster with Asian HapMap samples in assessment of global ancestry. Plot of the first two principal components of population substructure in 1,774 CLHNS offspring (blue and red, as described in Supplemental Figure 1) and 165 CEU, 84 CHB, 86 JPT, and 167 YRI HapMap Phase III samples (green, orange, brown, and black, respectively).

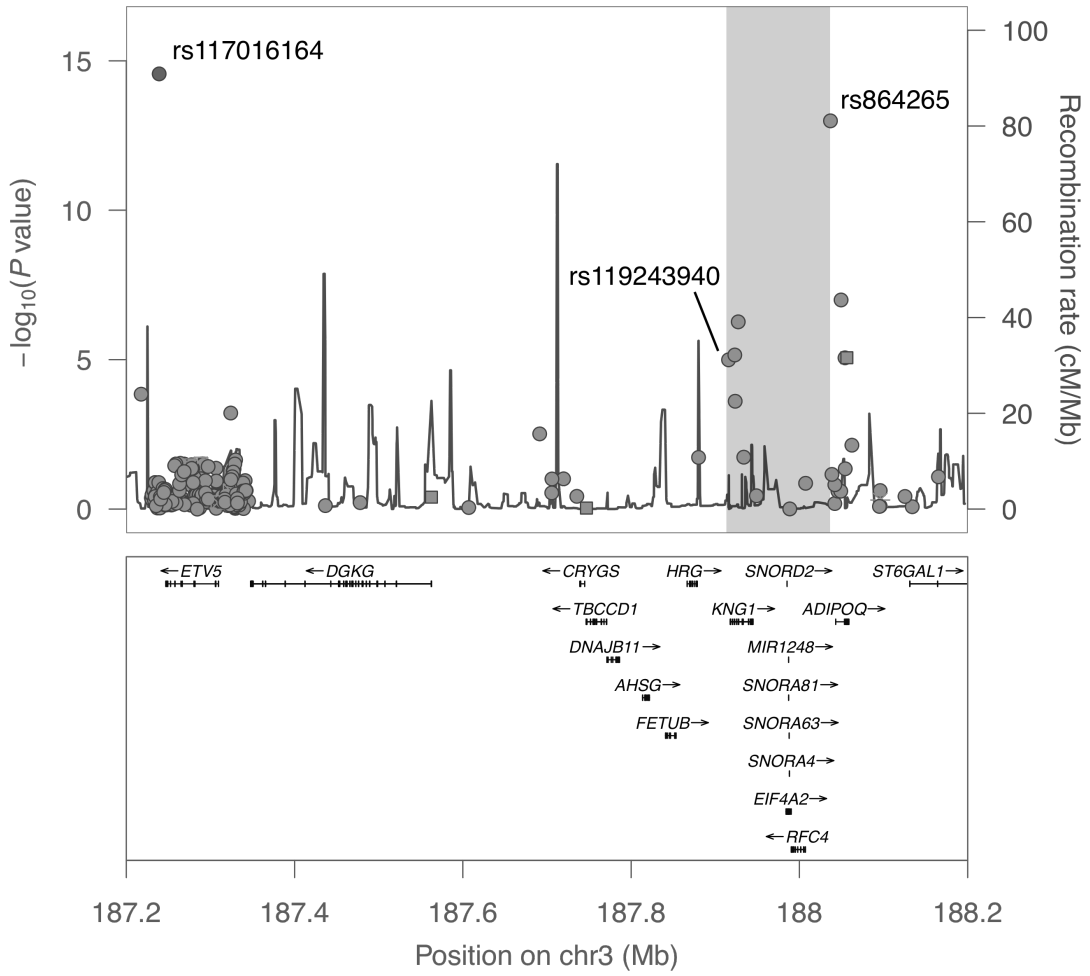


Figure 3.3 – Two adiponectin association signals ($P < 5.0 \times 10^{-8}$) on chromosome 3 in 1,695 CLHNS offspring. Regional plot of MetaboChip SNP associations with plasma adiponectin level on chromosome 3 (187.2–188.2 Mb). The previously associated 100 kb two-SNP *KNG1–ADIPOQ* haplotype of rs119243940 and rs864265 spans the region shaded gray. Plot was generated using LocusZoom (137).

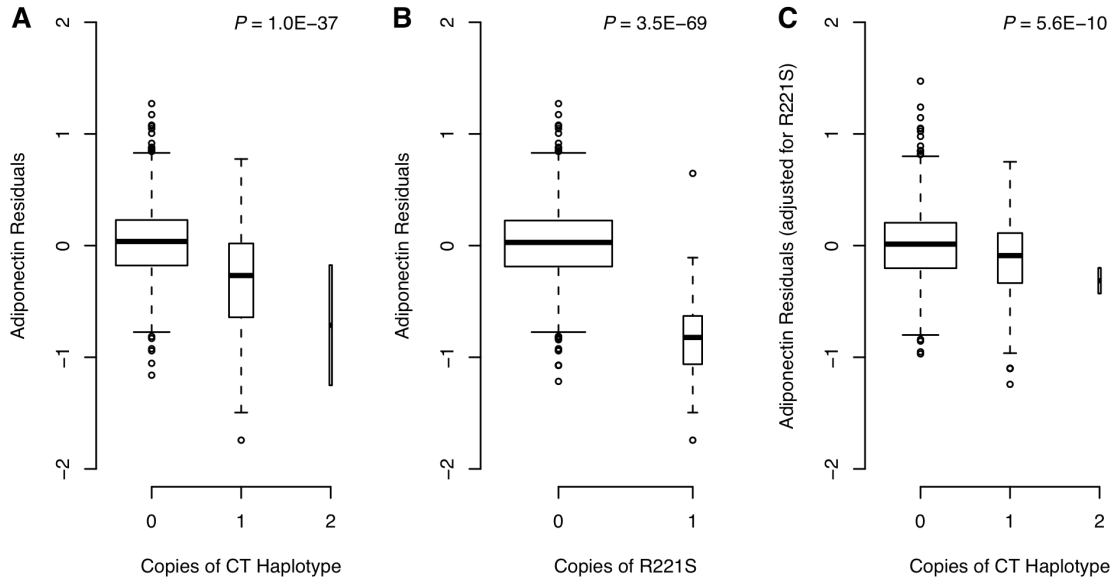


Figure 3.4 – Association of *KNG1–ADIPOQ* haplotype in CLHNS offspring is attenuated by *ADIPOQ* missense variant R221S. Adiponectin residuals plotted by number of copies of *(A)* the associated C-T *KNG1–ADIPOQ* haplotype, *(B)* the *ADIPOQ* missense variant R221S, and *(C)* the *KNG1–ADIPOQ* haplotype when conditioned on R221S.

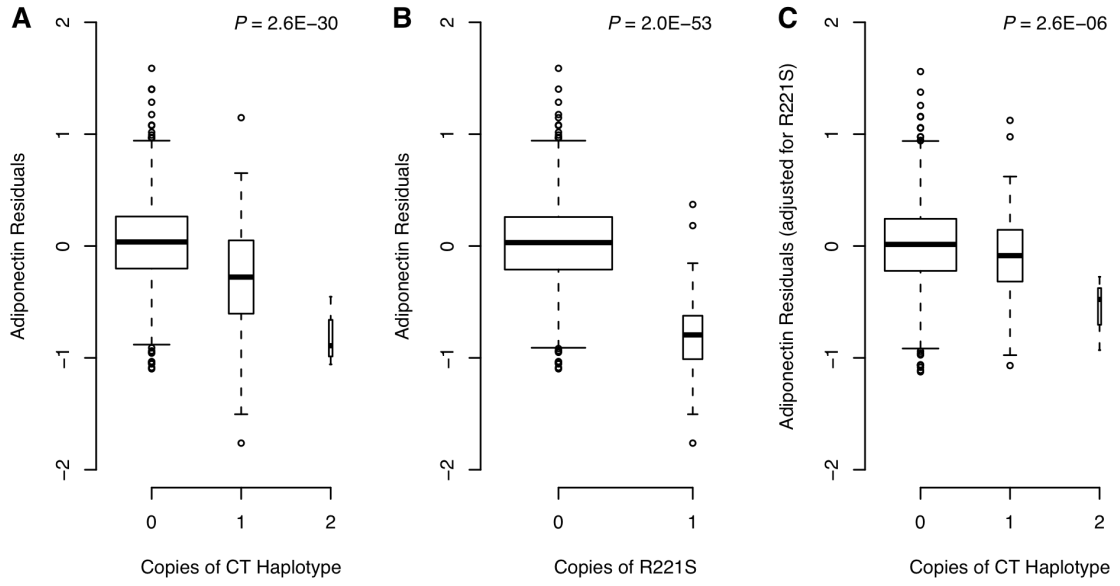


Figure 3.5 – Association of *KNG1–ADIPOQ* haplotype in CLHNS mothers is attenuated by *ADIPOQ* missense variant R221S. Adiponectin residuals plotted *(A)* by number of copies of the associated C-T *KNG1–ADIPOQ* haplotype, *(B)* by presence of the *ADIPOQ* missense variant R221S, and *(C)* by number of copies of the *KNG1–ADIPOQ* haplotype conditioned on R221S.

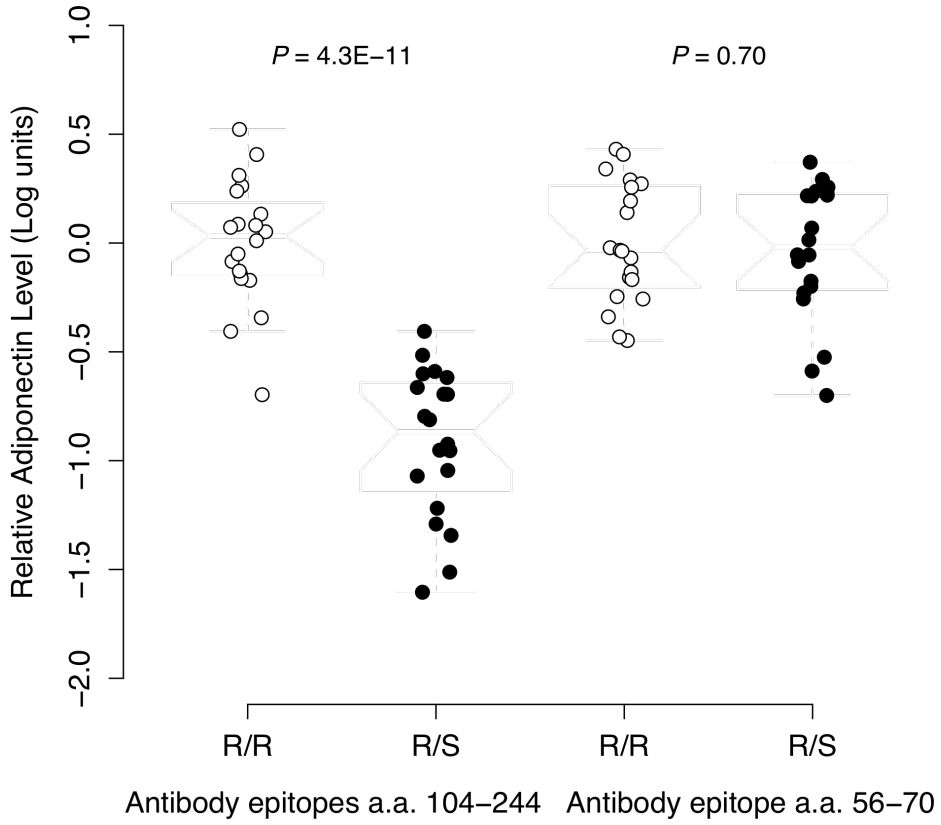


Figure 3.6 – Inconsistent plasma adiponectin measurements between assays is attributable to R221S. Plot of relative adiponectin levels of 20 R221S non-carriers (R/R: open circles) is shown compared to 20 R221S carriers (R/S: filled circles) measured by two antibody-based assays for *ADIPOQ* (left: epitopes within amino acids (a.a.) 104–244, right: 56–70). Within each assay, all adiponectin levels are natural log-transformed and zero-centered on the mean value of R221S non-carriers, and *t*-test *P* values are reported for the difference in means between carriers and non-carriers.

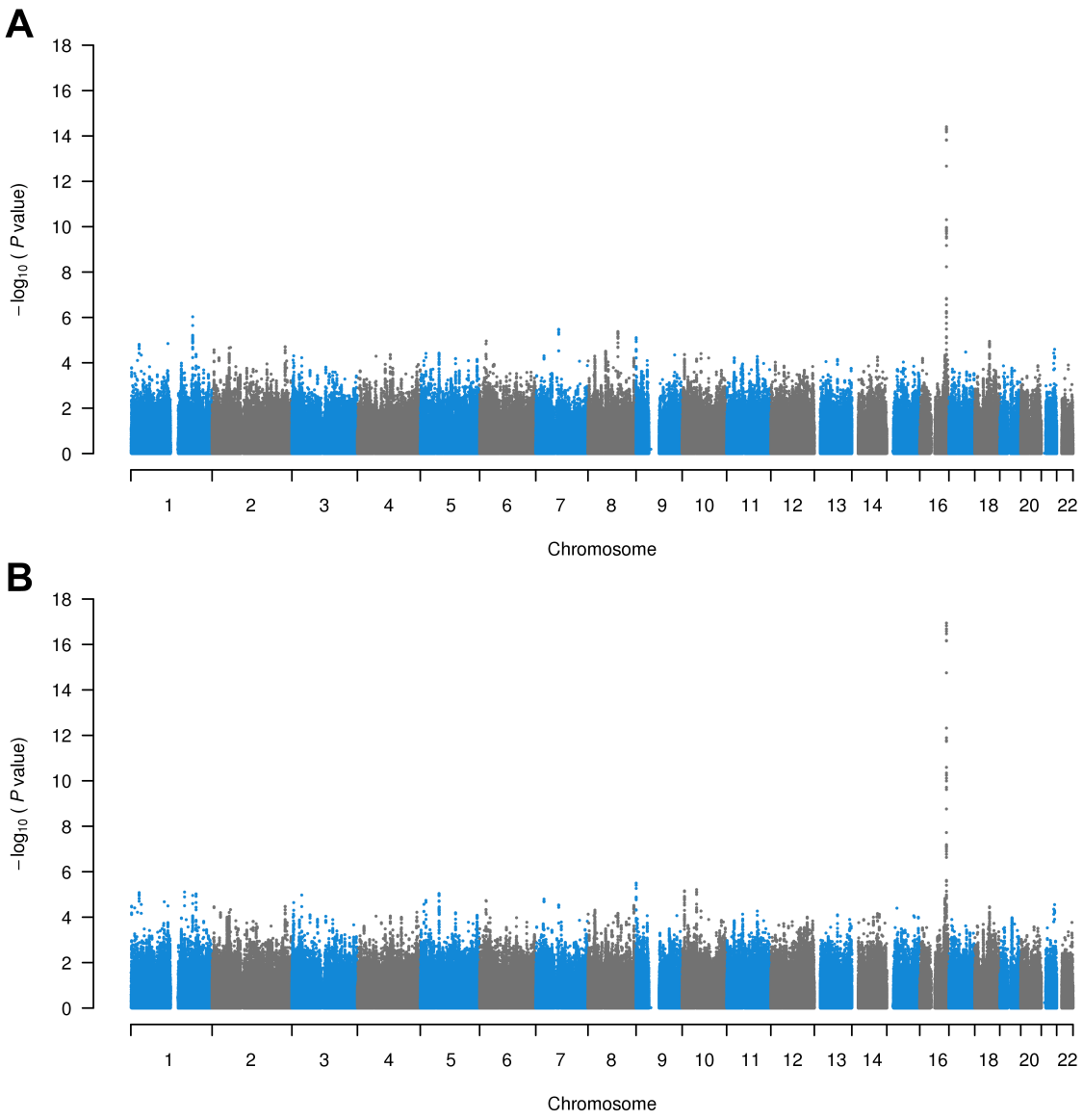


Figure 3.7 – In CLHNS mothers, only the *KNG1–ADIPOQ* association signal is substantially attenuated when taking into account *ADIPOQ* missense variant R221S. Genome-wide plot of HapMap SNP associations with plasma adiponectin level in CLHNS mothers after *(A)* conditioning on R221S (and removing samples of unknown genotype, *n* = 1,698) or *(B)* removing known R221S carriers (*n* = 1,729). While the *KNG1–ADIPOQ* association signal on chromosome 3 has been attenuated, the *CDH13* signal on chromosome 16 remains genome-wide significant.

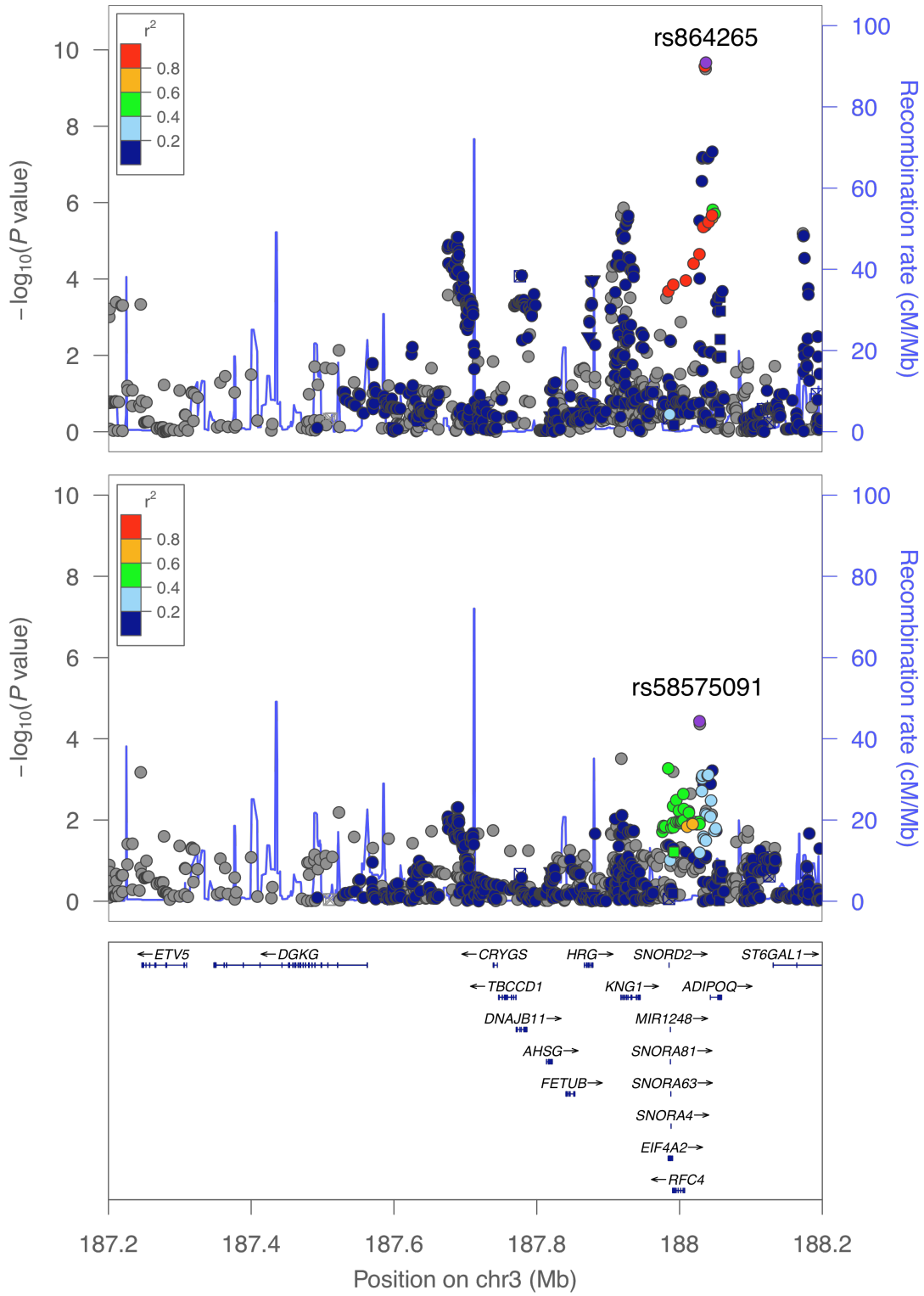

rs864265
rs58575091
r²
0.8
0.6
0.4
0.2
−log10(P value)
Recombination rate (cM/Mb)
ETV5
DGKG
CRYGS
HRG
SNORD2
ST6GAL1
TBCCD1
KNG1
ADIPOQ
DNAJB11
MIR1248
AHSG
SNORA81
FETUB
SNORA63
SNORA4
EIF4A2
RFC4
187.2
187.4
187.6
187.8
188
188.2
Position on chr3 (Mb)

Figure 3.8 – Residual association signal near *ADIPOQ* in CLHNS mothers after adjustment for R221S genotype. Regional plot of imputed SNP associations with natural log-transformed adiponectin in CLHNS mothers before and after adjustment for R221S genotype *(top and middle panels, respectively)*, along with nearby genes *(bottom panel)*. Pair-wise linkage disequilibrium (r^2) with the index SNP (purple) is reported based on data from CHB+JPT HapMap samples in the June 2010 release of the 1000 Genomes Project. The imputation quality score (Rsq) for rs58575091 was 0.79.

Table 3.1 – Association of three SNPs with plasma adiponectin level in CLHNS offspring and mothers

SNP (Gene)	Allele 1/2	CLHNS offspring (n = 1,695)				CLHNS mothers (n = 1,764)				Combined samples	
		MAF	β (SE)	P	R^2	MAF	β (SE)	P	R^2	β (SE)	P
rs864265 (*ADIPOQ*)	T/G	0.129	−0.134 (0.018)	1.0×10^{-13}	0.032	0.124	−0.124 (0.020)	4.8×10^{-10}	0.022	−0.128 (0.014)	4.2×10^{-20}
rs117016164 (*ETV5*)	T/C	0.025	−0.304 (0.038)	2.7×10^{-15}	0.036	0.030	−0.198 (0.039)	3.2×10^{-7}	0.015	−0.240 (0.028)	7.0×10^{-17}
R221S (*ADIPOQ*)	A/C	0.015	−0.853 (0.046)	2.9×10^{-69}	0.171	0.015	−0.811 (0.051)	2.0×10^{-53}	0.132	−0.823 (0.035)	4.3×10^{-99}

Effect betas (β), standard errors (SE), P values, and squared type II partial correlations (R^2) are reported in terms of the minor allele (underlined) and for natural log-transformed values of adiponectin. Minor allele frequencies (MAF) are reported for all 1,774 unrelated offspring and 1,798 mothers, regardless of covariate completeness. In the analyses of offspring, models were adjusted for sex and body mass index (BMI). In the analyses of mothers, models were adjusted for age, age^2, household assets, natural log-transformed income, menopausal status, and BMI. In the combined analyses, models accounting for sample relatedness were adjusted for sex, age, age^2, household assets, natural log-transformed income, and BMI. All covariates were from the 2005 survey.

Table 3.2 – Association of three-SNP haplotypes with plasma adiponectin level in CLHNS offspring

Haplotype	rs117016164	rs11924390	rs864265	Haplotype Frequency	*N* Carriers	β	SE	*P*
H1	C	C	G	0.435	1,196	-	-	-
H2	C	T	G	0.421	1,166	0.017	0.012	0.15
H3	C	T	T	0.075	256	0.022	0.022	0.31
H4	**C**	**C**	**T**	0.044	154	−0.237	0.029	$\mathbf{2.9 \times 10^{-16}}$
H5	T	C	G	0.010	33	−0.060	0.058	0.30
H6	**T**	**C**	**T**	0.009	31	−0.817	0.061	$\mathbf{2.1 \times 10^{-39}}$
H7	T	T	G	0.005	19	0.035	0.076	0.64
H8	T	T	T	0.001	5	−0.199	0.164	0.23

Effect betas (β), standard errors (SE), and *P* values were calculated for each haplotype relative to the most common reference haplotype. Frequencies are reported for all 1,774 unrelated offspring, regardless of covariate completeness. Significant associations ($P < 0.05$) are in boldface. Models were adjusted for sex and for body mass index from the 2005 survey. A *post hoc* comparison of H4 and H6 showed that their strengths of association were significantly different ($P = 2.9 \times 10^{-18}$).

Table 3.3 – Association of three-SNP haplotypes with plasma adiponectin level in CLHNS mothers

Haplotype	rs117016164	rs11924390	rs864265	Haplotype Frequency	*N* Carriers	β	SE	*P*
H1	C	C	G	0.429	1,203	-	-	-
H2	C	T	G	0.428	1,211	0.029	0.013	0.029
H3	C	T	T	0.066	226	0.049	0.025	0.053
H4	**C**	**C**	**T**	**0.047**	**165**	**−0.218**	**0.030**	**8.9×10^{-13}**
H5	T	C	G	0.011	40	0.149	0.057	0.0094
H6	**T**	**C**	**T**	**0.009**	**32**	**−0.787**	**0.066**	**1.9×10^{-31}**
H7	T	T	G	0.007	26	0.015	0.073	0.84
H8	T	T	T	0.002	8	−0.228	0.131	0.081

Effect betas (β), standard errors (SE), and *P* values were calculated for each haplotype relative to the most common reference haplotype. Frequencies are reported for all 1,798 mothers, regardless of covariate completeness. Significant associations ($P < 0.05$) are in boldface. Models were adjusted for gender and for body mass index from the 2005 survey. A contrast test for the difference between H4 and H6 was highly significant ($P = 2.9 \times 10^{-15}$).

Table 3.4 – Association of candidate *ADIPOQ* SNPs with plasma adiponectin level in CLHNS mothers conditioned on R221S genotype

SNP	Allele 1	Allele 2	Rsq	Freq1	β	SE	*P*
rs16861194	G	A	0.86	0.169	0.001	0.018	0.96
rs266729	G	C	0.92	0.276	−0.047	0.015	0.0017
rs182052	A	G	0.99	0.452	−0.034	0.013	0.0081
rs2241766	G	T	0.96	0.223	0.018	0.016	0.24
rs1501299	T	G	0.99	0.406	0.007	0.013	0.58
rs3774261	G	A	1.00	0.372	−0.019	0.013	0.13
rs6773957	G	A	1.00	0.372	−0.019	0.013	0.13
rs1063537	T	C	0.98	0.222	0.018	0.015	0.25

SNP frequencies (Freq1), effect betas (β), standard errors (SE), and *P* values are reported in terms of Allele 1. Imputation was performed in MACH and was assigned a quality score (Rsq).

Table 3.5 – General characteristics of the CLHNS samples

	CLHNS offspring	CLHNS mothers
N	1,695	1,764
Female (%)	45.3	100
Adiponectin (μg/ml)	2.47 (1.94, 3.10)	2.48 (1.90, 3.32)
Age (years)	21.5 ± 0.3	48.5 ± 6.1
Household income (pesos/wk)	358.6 (213.6, 572.2)	398.03 (245.0, 626.7)
Household assets (0 to 11)	5.2 ± 2.0	5.2 ± 1.9
Number of previous pregnancies	-	6.5 ± 3.0
Post-menopausal (%)	-	38.3
Body mass index (kg/m^2)*	20.2 (18.7, 22.2)	24.1 (21.4, 27.0)

Data are from the 2005 survey and are represented by mean ± SD, median (25th percentile, 75th percentile) or %.

* Women who were pregnant in the survey of 2005 were excluded from analysis.

Table 3.6 – TaqMan primers for genotyping *ADIPOQ* and *ETV5* SNPs of interest

SNP	Primer	Sequence (5'–3')
G48D	Forward	GGCCTGCACAGGTTGGAT
(rs182223755)	Reverse	GGTGCCATCTCTGCCATCA
	VIC Reporter	CATCCGGACCATAATG
	FAM Reporter	CCGGGCCATAATG
I164T	Forward	CAACATTCCTGGGCTGTACTACT
(rs185847354)	Reverse	GGCTGACCTTCACATCCTTCATAT
	VIC Reporter	TTGCCTACCACATCACAGT
	FAM Reporter	TGCCTACCACACCACAGT
R221S	Forward	GTCTGGCTCCAGGTGTATGG
(rs138773406)	Reverse	TGTGAAGGTGGAGTCATTGTCATT
	VIC Reporter	GGAGAGCGTAATGGACT
	FAM Reporter	GGAGAGAGTAATGGACTCTA
rs117016164	Forward	CTCAACCCACCCTGCTACAA
	Reverse	AGCAGCCGGCAGCTT
	VIC Reporter	AACTGATTGGAAGGGTG
	FAM Reporter	AACTGACTGGAAGGGTG

Table 3.7 – Associations of plasma adiponectin level in CLHNS offspring with non-genetic covariates and 10 principal components of population substructure

Covariate	Adjusted for	β (SE)	*P*
Age	1	−0.018 (0.029)	0.53
Age^2	2	−0.065 (0.087)	0.46
Sex†	**3**	**−0.162 (0.018)**	**< 0.0001**
Assets	4	−0.003 (0.004)	0.50
Log_income	4	0.015 (0.009)	0.084
BMI	**4**	**−0.033 (0.003)**	**< 0.0001**
PC1	5	−0.619 (0.355)	0.081
PC2	5	0.125 (0.356)	0.73
PC3	5	−0.295 (0.354)	0.40
PC4	5	−0.146 (0.359)	0.68
PC5	5	0.428 (0.357)	0.23
PC6	5	−0.670 (0.356)	0.060
PC7	5	0.660 (0.359)	0.066
PC8	5	0.019 (0.357)	0.96
PC9	5	−0.400 (0.359)	0.27
PC10	5	0.182 (0.356)	0.61

Betas (β) and standard errors (SE) are reported in natural log-transformed units of body mass index (BMI). Significant covariates ($P < 0.05$) are marked in bold. Each covariate test was additionally adjusted for (1) nothing; (2) age; (3) age and age^2; (4) age, age^2, and sex; or (5) age, age^2, sex, BMI, and all other principal components of population substructure (PC).

†Sex is coded as female = 0 and male = 1

CHAPTER IV

Meta-analysis of 231,355 European individuals identifies seven novel loci associated with waist circumference[1]

OVERVIEW

Waist circumference (WC) is a measure of central adiposity showing substantial genetic heritability and strong positive association with risk of metabolic syndrome and cardiovascular disease apart from total adiposity. In a meta-analysis of genetic associations from 82 cohorts comprising up to 231,355 European individuals, we identified 84 independent WC signals. The vast majority of the 84 signals were strongly associated ($P < 5 \times 10^{-8}$) with height and/or waist-hip ratio, but seven were not associated with either of these traits (*NLRP3*, *ZNF391*, *SCAMP5*, *IRS1*, *BTN2A1*, *CCNJL*, *LINC00240*). Of particular biological interest are loci such as *NLRP3*, which is part of the obesity-related inflammasome complex, and *IRS1*, which is previously associated with body fat percentage and an adverse metabolic profile. These results suggest the potential role of genes related to chronic inflammation and fat metabolism in the biology of central adiposity.

---

[1] This unpublished work was undertaken as part of the Genetic Investigation of ANthropometric Traits (GIANT Consortium).

INTRODUCTION

Central adiposity, as measured by waist circumference (WC) and waist-hip ratio (WHR), is an important predictor of all-cause mortality and cardiovascular disease (CVD) mortality apart from overall adiposity, as measured by body mass index (BMI) (28, 138, 139). WC is positively correlated with both hip circumference (HC) and WHR (139). While WHR is a better predictor of CVD risk because it summarizes central fat distribution more comprehensively (27, 28), WC is a strong proxy for visceral adipose tissue (18), which itself is more strongly associated with cardio-metabolic risk than subcutaneous adipose tissue (10). In particular, a WC ≥ 80 cm in women or ≥ 102 cm in men is a defining aspect of the metabolic syndrome that underlies CVD risk (140, 141).

Reports of the heritability of WC (h^2 = 31–46%) suggest there is a substantial additive genetic component (142-145). However, only five loci (*FTO*, *MC4R*, *MSRA*, *NRXN3*, and *TFAP2B*) have been previously reported as showing genome-wide significant associations ($P < 5 \times 10^{-8}$) with WC in Europeans (56, 79, 86, 146), and explain just a small portion of the phenotypic variance of WC. To discover additional associated variants, genetic association studies with larger sample sizes and greater statistical power are needed.

On behalf of the Genetic Investigation of ANthropometric Traits (GIANT) Consortium, I led a team of analysts in performing a large-scale meta-analysis of SNP associations with WC from 82 studies of European individuals. Each study contributed either genome-wide imputed SNP data or direct genotypes from a custom array called the MetaboChip, which in a cost-efficient manner targets

thousands of SNPs showing prior evidence of association (approximately $P < 10^{-4}$) with 23 cardio-metabolic and anthropometric traits, including WC, WHR, and height. Through our analyses, we implicated seven novel association loci in the development of central adiposity.

MATERIALS AND METHODS

A total of 82 studies, comprising up to 231,355 European samples with waist circumference and BMI data (104,079 men and 127,470 women) were meta-analyzed together (**Figure 4.1**). The first stage comprised 54 GWA studies with up to 153,482 samples with genome-wide SNP data (up to 61,508 men and up to 92,167 women). The second stage comprised 28 studies with up to 77,877 samples genotyped with the custom MetaboChip SNP array (up to 42,574 men and up to 35,304 women). The GWA studies each imputed up to 2,812,050 SNPs based on direct genotypes from genome-wide SNP arrays and reference haplotypes from CEU samples in the HapMap Phase 2 reference panel (147). The MetaboChip studies each directly genotyped up to 155,738 SNPs. A total of 2,867,624 SNPs were analyzed with 104,423 SNPs shared between the GWA and MetaboChip stages.

In each contributing study cohort, residuals for the WHR phenotype were computed separately for men and women and adjusted for age, BMI, and study-specific principal components of population substructure. Cohorts with related samples also calculated sex-combined residuals. Each set of residuals was then transformed using inverse normalization and tested for additive association with all available SNP genotypes.

Several quality control (QC) checks were performed on the association data contributed by each study. SNPs were dropped if they had a minor allele count (MAC) ≤ 3 (MAF × sample size), Hardy-Weinberg equilibrium $P < 10^{-6}$, call rate < 95% for directly genotyped SNPs, or low imputation quality as recommended by the authors of the imputation software used in each study.

Sex-combined and sex-specific fixed-effects meta-analyses were performed in METAL using both inverse variance and Z-score methods (148). Sensitivity analyses were performed at each meta-analysis stage to ensure concordance between these approaches. To control for potential inflation due to population stratification within each cohort and for differences among studies in study design or sample ascertainment, two genomic control (λ_{GC}) corrections were applied to the meta-analysis results. In the first stage, GWA data were meta-analyzed together applying both within-study and among-studies λ_{GC} corrections based on all SNPs (**Figure 4.1**). In the second stage, MetaboChip data were meta-analyzed together with a similar λ_{GC} correction scheme based on an expected null set of 4,427 MetaboChip SNPs included on the array to study QT interval. The GWA and MC meta-analysis results were then meta-analyzed together with no additional λ_{GC} correction. For each SNP, among-studies association heterogeneity was assessed using the I^2 statistic, and heterogeneity due to differences in sex-specific effects was assessed using a two-sided *t* test adjusting for correlation between sexes (57). Binomial sign tests were performed to determine whether directions of effect were more consistent than expected by chance alone between meta-analysis stages. Clumping of meta-analysis results was performed using PLINK (133) and linkage

disequilibrium (LD) information calculated from the genotype data of 56 unrelated CEU samples typed in the 1000 Genomes Project (1000G).

To identify evidence of potential secondary signals, we performed approximate conditional and joint analysis of the meta-analysis results using a statistical method called GCTA that uses estimated LD from a reference sample with individual genotypes to identify independent signals (149). We used 949 unrelated individuals with individual-level whole-genome genotype data (both Illumina OmniExpress and MetaboChip) selected from the PIVUS cohort to approximate the LD structure between SNPs meta-analyzed in the GIANT Consortium (150). Directly genotyped data from the PIVUS samples were imputed to the full set of variants from CEU samples in the HapMap Phase 2 reference panel (147) using IMPUTE (version 2) (151). Only "best guess" genotypes for SNPs with an imputation quality score > 0.4 and HWE $P > 1 \times 10^{-6}$ were used in the conditional analyses. Starting with approximately 2.54 million SNPs in the GIANT meta-analysis data, we removed SNPs not present in the PIVUS data after QC exclusions and only considered SNPs with an estimated sample size (N) > 50,000, thus retaining approximately 2.53 million SNPs for the conditional analysis. Assuming that the LD correlations between SNPs more than 10 Mb away or on different chromosomes are zero, we performed a genome-wide stepwise selection procedure to identify independently associated SNPs at $P < 5 \times 10^{-8}$.

To identify which SNPs were associated with central adiposity but not with other correlated anthropometric traits, preliminary meta-analysis data of sex-combined associations were obtained from other studies within the GIANT

Consortium for HC (*N* = 211,117), WHR (*N* = 210,086), height (*N* = 336,232), and BMI (*N* = 319,284). These analyses included many samples from previous meta-analyses of WHR (57), height (152), and BMI (59). BMI was included as a covariate in the associations of HC and WHR.

Finally, to learn more about the biology underlying observed association signals, SNPs were annotated for nearby gene names and related publications using SNIPPER (available on-line at http://csg.sph.umich.edu/boehnke/snipper/). We determined the LD between WC SNPs and SNPs previously reported for other diseases and quantitative traits based on pre-calculated data from phased genotypes in CEU samples from the 1000G Pilot 1 dataset queried through SNAP (153).

RESULTS

Genome-wide associations in sex-combined and sex-specific meta-analyses

We conducted multi-stage meta-analyses of associations of 2,867,624 SNPs with WC, 104,423 of which were typed in up to 231,355 individuals of European ancestry (**Figure 4.1**). To distinguish associations with the biology of central adiposity from that of overall adiposity, BMI was included as a covariate in associations with WC. We performed fixed-effects meta-analyses of study-level additive associations with waist circumference stratified by sex and adjusting for age, BMI, and population substructure. Quantile-quantile plots of the sex-combined and sex-specific meta-analysis results showed an excess of strongly associated SNPs (**Figure 4.2**, **Figure 4.3**) without strong evidence of systematic association inflation (sex-combined λ_{GC} = 1.03, female-specific λ_{GC} = 1.02, and male-specific λ_{GC} = 1.03).

Based on the sex-combined association data, we calculated λ_{GC} values for specific subsets of the 98,799 SNPs that are shared between the two meta-analysis stages and have a sample size (N) > 50,000 (**Figure 4.4**). The first subset was a null set of 4,301 SNPs included on the MetaboChip array to study QT interval, a cardiological quantitative trait unlikely to be biologically related to fat distribution, which showed a small excess of positive association (λ_{GC} = 1.13). The second subset was all 98,799 SNPs, which showed a moderate excess of positive association (λ_{GC} = 1.50). The third subset was 1,038 SNPs included on the MetaboChip specifically to target SNPs with prior evidence of association with WC, which showed a substantial excess of positive association (λ_{GC} = 15.75). Comparable results were observed in the sex-specific data (*data not shown*). Together, these results suggest that we have identified true associations with WC.

In the sex-combined meta-analysis, we identified 274 SNPs with N > 50,000 that were associated with WC at $P < 5 \times 10^{-8}$ (**Figure 4.5**). We also observed that two of the five previously reported WC GWA signals (*NRXN3* and *FTO*) showed P < 0.05 in the sex-combined analysis with consistent direction of effect (**Table 4.1**). To obtain an initial estimate of the number of genome-wide loci observed, we clumped all genome-wide significant SNPs within 1 Mb windows, regardless of LD relationships, resulting in 74 representative SNPs with the strongest P value in their respective windows. In the sex-specific meta-analyses (**Figure 4.6**), we identified 513 female-specific and 274 male-specific SNPs associated with waist circumference at $P < 5 \times 10^{-8}$ with N > 50,000, represented by 34 and 44 clump index SNPs, respectively. Many of the observed signals overlapped, so these results represented

a unique set of 69 sex-combined and six sex-specific signals (**Table 4.1**). Of these 75 signals, 29 showed nominally different effects between males and females ($P < 0.05$) and six showed substantially different effects ($P < 5 \times 10^{-8}$). Globally, there appeared to be an excess of sex-based heterogeneity at the tail end of the distribution (**Figure 4.7**), but there was no strong evidence of systemic inflation ($\lambda_{GC} = 1.018$). Eight genomic regions (*ZNT8*, *OR2W5*, *COBLL1*, *IRS1*, *BC040632*, *MAP3K1*, *CCNJL*, and *VEGFA*) showed significant sex heterogeneity at $P < 5 \times 10^{-8}$ (**Figure 4.8**). Of note, one region (*ZNT8*) was labeled as *LYPLAL1* in a meta-analysis of WHR associations and also showed significant sex-based heterogeneity (57). These results suggest limited evidence of sexual dimorphism in the tens of observed genome-wide association signals for WC.

Identification of allelic heterogeneity

To more formally identify independent signals, we pruned each of the meta-analysis result sets using conditional analyses of SNPs in 1 Mb windows (149). We observed 83, 28, and 32 independently associated SNPs in the sex-combined, female-specific, and male-specific analyses, respectively (**Table 4.2**). The vast majority of the sex-specific signals were the same as the sex-combined signals so the conditional analyses identified a total of 117 unique SNPs. In the sex-combined analysis, we observed eight instances where more than one independent SNP was located within a 1 Mb region (*TBX15*, *C5orf23*, *HIST1H2BF*, *HMGA1*, *JAZF1*, *PTCH1*, *ADAMTSL3*, and *ADAMTS17*, each named for the nearest gene to the most associated SNP). In each of the female-specific and male-specific analyses, we observed a single

pair of independent SNPs in close proximity (*KIAA2017/ZNF664* and *WARS2/SPAG17*, respectively). Only the female-specific *ZNF664* index SNP was not represented in the sex-combined joint analysis.

To determine how many additional signals were added to the 75 suggested by distance-based clumping, we checked whether the lead SNP from **Table 4.1** was represented in **Table 4.2** by the same SNP or with a proxy ($r^2 \geq 0.1$). The conditional analyses added nine additional signals. Of the 84 SNPs now representing independent signals, the joint conditional analyses identified 12 SNPs that were not genome-wide significant in the single-marker analyses. Performing conditional analyses in these data helped reveal evidence of multiple signals in regions, primarily through the sex-combined data, suggesting allelic heterogeneity in the genetic architecture of WC.

Assessment of replication patterns

To broadly evaluate how useful the set of MetaboChip SNPs was in identifying true associations with WC, we performed a series of binomial sign tests to check for consistency of effects between data from the GWA and MetaboChip meta-analysis stages. We first obtained a null set of SNPs by selecting a set of 2,000 MetaboChip QT SNPs, all of which were in low pair-wise LD with each other ($r^2 < 0.1$) to ensure that the tests would not be biased by non-independent SNPs. This set of SNPs showed greater directional consistency of effect between the two meta-analysis stages in the sex-combined analysis than expected by chance alone (1,078 SNPs compared to 1,000 expected, $P = 5.3 \times 10^{-4}$). However, results observed

separately in women (1,025 SNPs, $P = 0.27$) and in men (1,010 SNPs, $P = 0.67$) were not distinguishable from chance. We then obtained without selecting for a particular trait a set of similarly independent SNPs present in both meta-analysis stages ($n = 17,354$). This set of SNPs showed significantly greater directional consistency of effect between the two meta-analysis stages in the sex-combined analysis than expected by chance alone (9,634 SNPs compared to 8,677 expected, $P = 7.4 \times 10^{-48}$). Comparable results were observed separately in women (9,269 SNPs, $P = 2.6 \times 10^{-19}$) and in men (9,367 SNPs, $P = 1.1 \times 10^{-25}$). The strong consistency of effect between the meta-analysis stages validates our targeted SNP selection strategy for increasing sample size to observe true associations with WC.

Associations with other anthropometric measures

To identify loci associated with central adiposity but not with other correlated anthropometric traits, we examined the 84 independent WC SNPs for associations with HC, WHR, height, and BMI in data from large-scale meta-analyses by the GIANT Consortium (**Table 4.3**). The sample sizes for the HC and WHR associations were comparable to WC, but were substantially greater for both height and BMI with nearly 100,000 additional samples. We observed several broad patterns in the relationships of WC-increasing SNP alleles with other anthropometric traits. As expected, because BMI was included as a covariate in our analyses, most WC SNPs were not significantly associated with BMI ($P \geq 0.05$) or showed nominal association with decreased BMI ($P < 0.05$). However, the majority of WC SNPs (64 of 84) were nominally associated with increased height, and most

SNPs in this set (55 of 64) reached genome-wide significance ($P < 5 \times 10^{-8}$). All of these height-associated WC SNPs were also nominally associated with increased HC and/or WHR. Of the 20 remaining WC SNPs not significantly associated with increased height, 13 were strongly associated with increased WHR ($P < 0.0001$). Four of these SNPs also showed nominal associations with decreased HC and two with decreased height. We designated the remaining seven SNPs showing a $P \geq 0.0001$ with each of the other four traits as WC-specific SNPs. Nearly all of these associations were not significant at all ($P \geq 0.05$). Notably, five of these SNPs only became genome-wide significant through conditional analysis (rs4886648, rs3799380, rs2754603, rs13196692, and rs1515114) (**Table 4.2**). While we observe substantial overlap of SNP associations for WC with other anthropometric traits, seven SNPs represent WC-specific signals.

Potential biological roles

These seven novel loci distinct to WC suggest the role of several compelling new genes in the biology of central adiposity. Approximately 39 kb downstream of the first WC SNP rs10925060 is an inflammasome complex gene called *NLRP3* that plays a direct role in obesity-related metabolic pathogenesis (154, 155). Furthermore, two intronic *NLRP3* SNPs are highly associated ($P < 5 \times 10^{-8}$) with the inflammation-related traits of circulating plasma levels of C reactive protein (CRP) and fibrinogen (156, 157). Our WC index SNP is in very low LD with both the CRP index SNP (rs12239046) ($r^2 \sim 0.04$, $D' \sim 0.81$ in 1000G CEU) and the fibrinogen index SNP (rs1539019) ($r^2 \sim 0.01$, $D' \sim 0.66$). The second WC SNP rs13196692 is

located about 9.8 kb downstream of the zinc finger protein *ZNF391*, and is in very low LD (r^2 ~0.03, D' = 1) with a SNP nominally associated with rheumatoid arthritis (rs13196692, P = 0.005) (158). The third WC SNP rs4886648 is located in an intron of *SCAMP5,* which codes for a secretory carrier membrane protein that is expressed in neural and exocrine tissues and is involved in the exocytosis of signal peptide-containing cytokines (159). The fourth WC SNP rs1515114 is found about 500 kb upstream of the *IRS1* gene, which codes for the insulin receptor substrate 1 protein involved in insulin signal transduction (160). The WC index SNP is in moderate LD with another SNP (rs2943650, r^2 ~ 0.52, D' ~ 1.0) located ~490 kb upstream of *IRS1* that is associated with lower body fat percentage and an adverse metabolic profile (161). The fifth WC SNP rs3799380 is in an intron of the gene *BTN2A1*, part of the B7/butyrophilin-like group, a subset of the immunoglobulin gene superfamily, which is involved in the production of milk fat globules and the regulation of immune function (162, 163). The WC index SNP is in low LD with a *BTN2A1* SNP (rs6929846, r^2 ~ 0.04, D' ~ 0.82) associated with MetS (P < 0.05) (164). While the sixth WC SNP rs17472426 is located in an intron of *CCNJL*, an uncharacterized cyclin J-like gene, it is also an expression quantitative trait locus (eQTL) in CEU lymphoblastoid cell lines (P = 0.0001) for the *trans* expression of *CREBBP* (165). This gene codes for a transcription factor called cAMP-response element-binding protein-binding protein that co-activates peroxisome proliferator-activated receptor (PPAR)-γ in bovine adipose tissue (166). In addition, just upstream of *CCNJL* is a gene called *FABP6* that codes for the human ileal bile acid-binding protein, which aids in the digestion and absorption of dietary fat and is regulated by

PPARs (167). A rare missense mutation in that gene was also associated with lower risk of type 2 diabetes in obese individuals (168). Finally, the seventh WC SNP rs2754603 is found in the middle of a long intergenic non-protein-coding RNA (*LINC00240*), but is also located about 134 kb upstream of the *HIST1* histone gene cluster, which form the basis for nucleosomes involved in chromatin structure. These results suggest the potential role of genes related to chronic inflammation and fat metabolism in the biology of central adiposity.

DISCUSSION

Our meta-analysis comprising 231,355 European individuals represents to date the largest genetic association study of WC, identifying 84 independent signals. Of these, seven SNPs were not associated with the correlated anthropometric traits WHR and height. These meta-analysis results represent new WC loci beyond the five genes (*FTO*, *MC4R*, *MSRA*, *NRXN3*, and *TFAP2B*) from previous reports (56, 79, 86, 146), only two of which replicated in our study (*FTO* and *NRXN3*). We also identified novel evidence of nine regions comprising more than one association signal (*TBX15/WARS2/SPAG17, KIAA2017/ZNF664*, *C5orf23*, *HIST1H2BF*, *HMGA1*, *JAZF1*, *PTCH1*, *ADAMTSL3*, and *ADAMTS17*).

The seven novel association loci suggest a wide functional variety of genes play roles in WC biology. Three regions (*NLRP3*, *ZNF391*, and *SCAMP5*) include genes associated with inflammatory biology, three regions (*IRS1*, *BTN2A1*, and *CCNJL*) have biological connections to body fat and metabolism, and one region (*LINC00240*) includes a histone gene cluster whose direct connection to waist

circumference biology is unclear. The very weak LD between the WC index SNPs at *NLRP3*, *ZNF391*, and *BTN2A1* with the respective index SNPs for other inflammation- and metabolism-related traits suggests we have observed unrelated signals in these regions. However, the moderate LD between the WC and BF% index SNPs at *IRS1* may mean that the two SNPs may represent the same signal, consistent with the correlation between the traits. These results provide new candidate genes for functional characterization.

All but two of the other 77 observed WC association signals were very strongly associated with height and/or WHR ($P < 5 \times 10^{-8}$). Consistent with the strong phenotypic correlations between height and BMI (the ratio of weight to height^2) and between WC and BMI (139, 169), most of the WC-increasing association signals were observed at known height loci (152) or were very strongly associated with increased height. A formal pleiotropy analysis would be necessary to determine with more certainty whether these loci truly affect both height and WC. We also observed that many of these height-associated WC-increasing SNPs were also strongly associated with increased HC. As HC when adjusted for BMI has a phenotypic correlation with height similar to WC, it seems likely that these signals are indeed explained by height variability. Several of the observed WC signals were also found at known WHR loci (57). While prior meta-analysis results suggested that hip circumference is generally the stronger biological component in genetic associations with WHR (57), the loci we detected in this WC meta-analysis that were associated with WHR are likely driven by WC biology. The lack of significant associations with HC for some of these SNPs supports this hypothesis. A few signals,

however, were nominally associated with decreased HC, suggesting that the underlying gene might both increase WC and decrease HC resulting in increased WHR.

In summary, through a very large-scale meta-analysis of GWA studies in European populations, we have identified seven novel association loci containing genes that may play a role in biology of central adiposity. Additional molecular and genomic information, such as eQTL and regulatory data in tissues derived from central fat depots, is needed to best identify promising candidate genes at these loci and the variants that influence their expression and function. The genome-wide associations of 77 additional WC loci with strongly correlated body size phenotypes also highlight the need for careful interpretation of association results. Our work helps motivate further genetic and molecular investigations into how causal variants manifest downstream disease states, such as central obesity.

ACKNOWLEDGMENTS

I would like to acknowledge the members of the Waist Traits Working Group of the GIANT Consortium who assisted me in this work. They are T. Ferreira, D. Shungin, T.W. Winkler, A.E. Locke, R. Magi, R.J. Strawbridge, K. Fischer, T. Workalemahu, P.J. Griffin, C.C. White, A.U. Jackson, F. Day, M.C. Zillikens, I. Barroso, C.S. Fox, E. Ingelsson, M.I. McCarthy, K.E. North, E.K. Speliotes, P.W. Franks, L.A. Cupples, L. Qi, A.P. Morris, I.M. Heid, K.L. Mohlke, C.M. Lindgren, and R.J.F. Loos. I would also like to acknowledge the researchers, analysts, and subjects from the individual studies that contributed their association data.

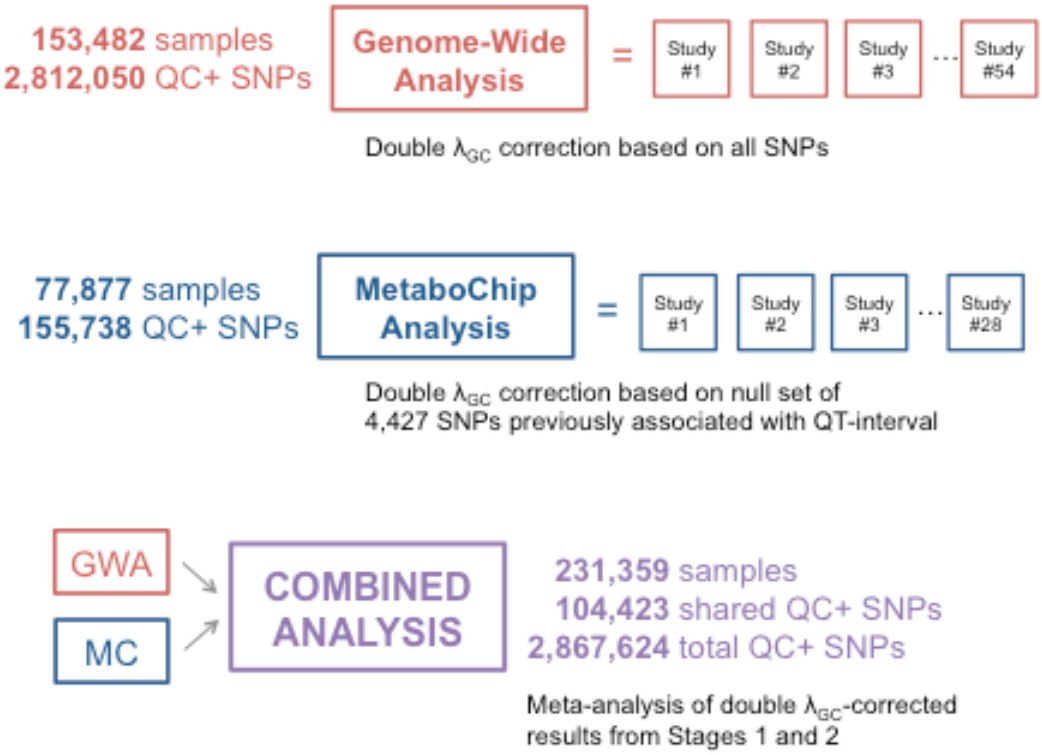


Figure 4.1 – Study design for the meta-analysis of waist circumference association data.

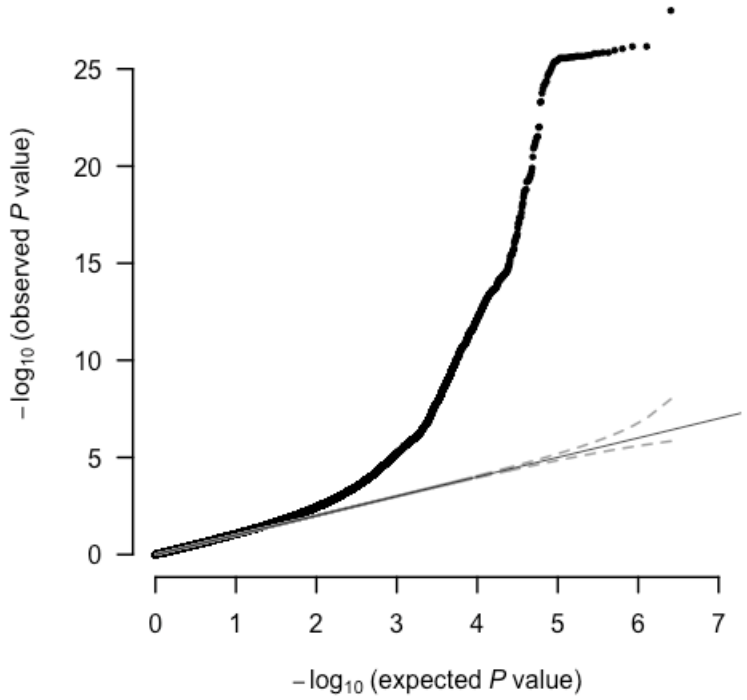


Figure 4.2 – Quantile-quantile plot of sex-combined SNP associations with waist circumference. The straight dark gray line indicates the expected distribution under the null hypothesis of no association between SNP and phenotype, and the two curving dashed light gray lines indicate the 95% confidence interval. Only SNPs with $N > 50{,}000$ are shown ($n = 2{,}545{,}722$, $\lambda_{GC} = 1.025$).

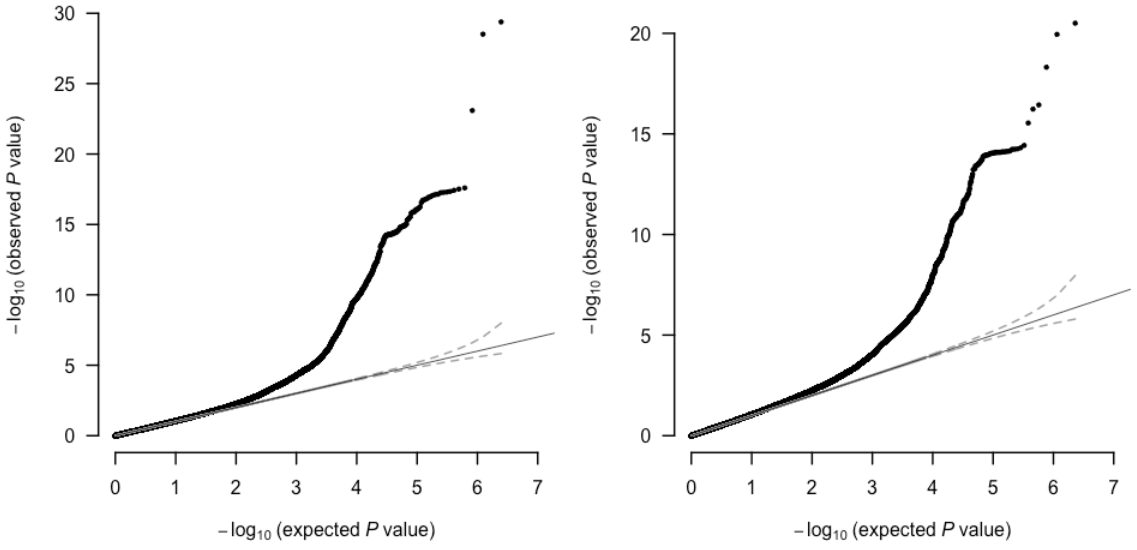


Figure 4.3 – Quantile-quantile plots of sex-specific SNP associations with waist circumference. Figure is set up as described **Figure 4.2**. Female-specific SNP associations (n = 2,473,035, λ_{GC} = 1.02) are shown in the left panel and male-specific SNP associations (n = 2,294,965, male-specific λ_{GC} = 1.03) in the right panel.

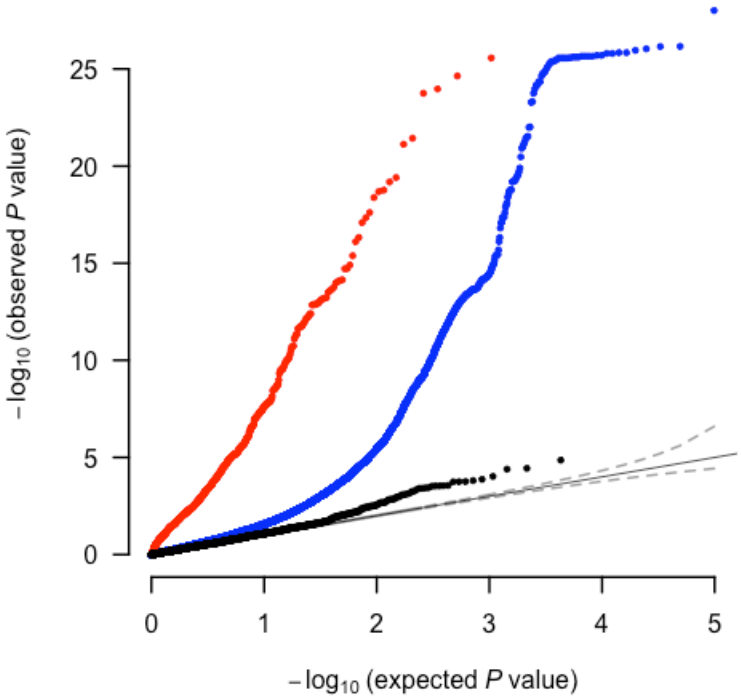


Figure 4.4 – Quantile-quantile plot of sex-combined associations with waist circumference showing excesses of positive associations. Figure is set up as described **Figure 4.2**. SNPs previously associated with QT interval (n = 4,301) and representing a null distribution are marked in black (λ_{GC} = 1.13). All SNPs present in both the GWA and MetaboChip meta-analysis stages (n = 98,799) are marked in blue (λ_{GC} = 1.50). SNPs included on the MetaboChip based on prior evidence of association with waist circumference (n = 1,038) are marked in red (λ_{GC} = 15.75).

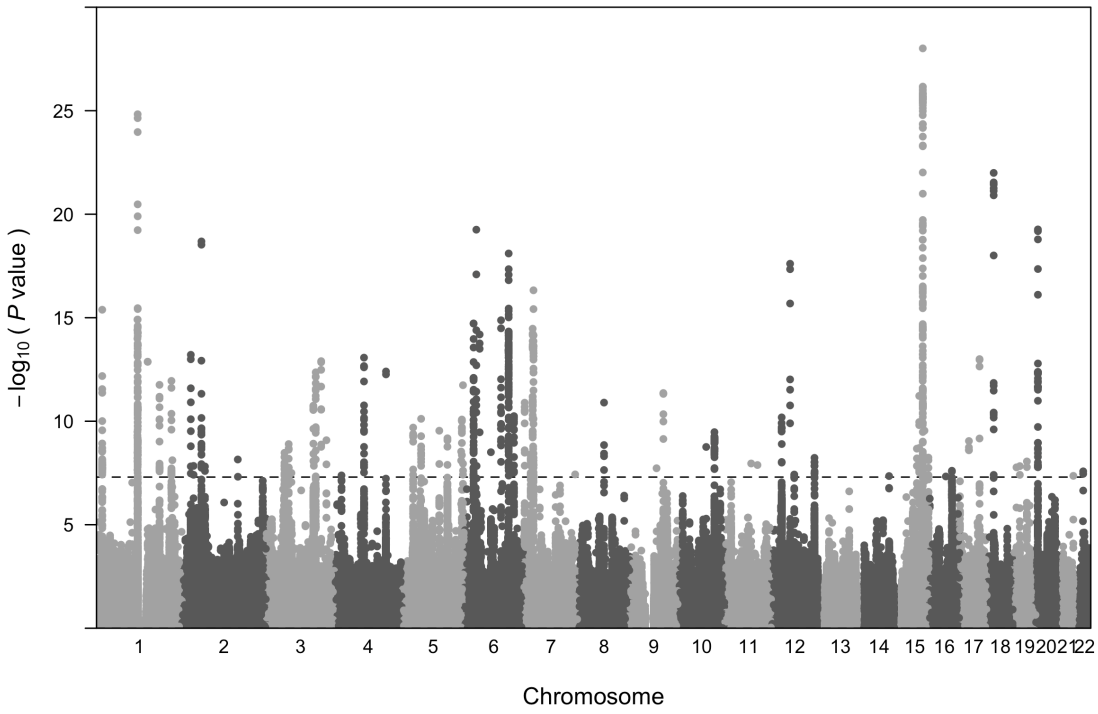


Figure 4.5 – Manhattan plot of sex-combined SNP associations with waist circumference. Only SNPs with $N > 50{,}000$ are shown ($n = 2{,}545{,}722$). The genome-wide significance threshold ($P < 5 \times 10^{-8}$) is marked with a dashed line.

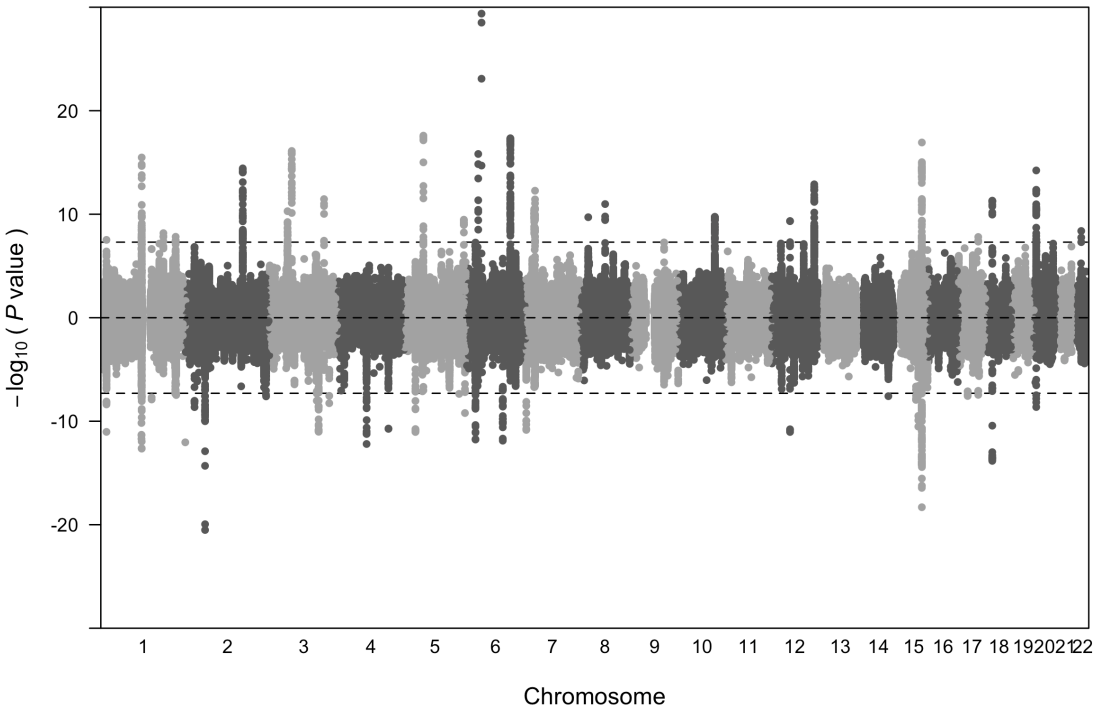


Figure 4.6 – Chicago plot of sex-specific SNP associations with waist circumference. Figure is set up as described in **Figure 4.5**. Female-specific SNP associations (n = 2,473,035) are displayed on top with positive y-axis values and male-specific SNP associations (n = 2,294,965) are displayed on bottom with negative y-axis values.

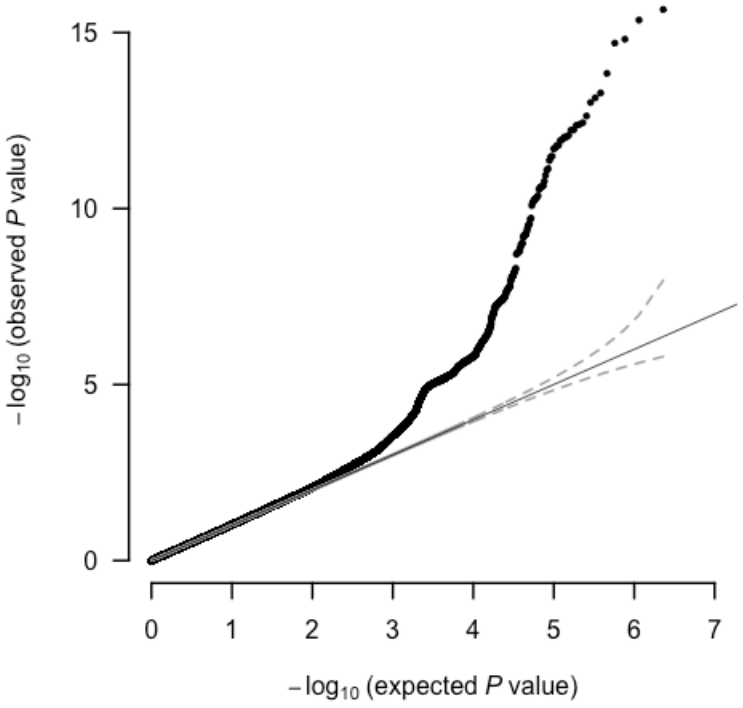


Figure 4.7 – Quantile-quantile plot of sex-based heterogeneity in SNP associations with waist circumference. Figure is set up as described **Figure 4.2**. Only SNPs with $N > 50{,}000$ between women and men are shown (n = 2,294,264, $\lambda_{GC} = 1.02$).

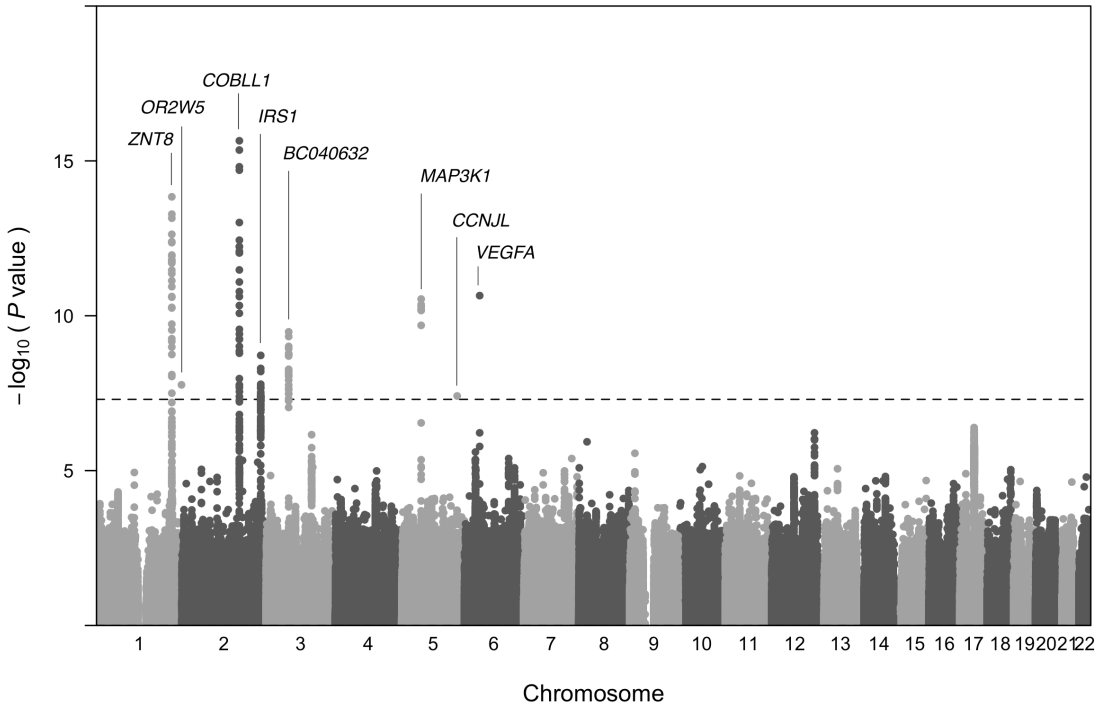


Figure 4.8 – Manhattan plot of sex-based heterogeneity in SNP associations with waist circumference. Figure is set up as described in **Figure 4.5**. Only SNPs with $N > 50{,}000$ between women and men are shown ($n = 2{,}294{,}264$) as in **Figure 4.7**. Regions with SNPs showing between-sex heterogeneity at $P < 5 \times 10^{-8}$ are labeled with the name of the nearest gene.

Table 4.1 – A total of 75 loci associated with waist circumference in sex-combined and sex-specific meta-analyses

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Effect allele freq.	Sex-combined analysis				Female-specific analysis				Male-specific analysis				Sex het. *P*
							β	SE	*P*	*N*	β	SE	*P*	*N*	β	SE	*P*	*N*	
All clumped WC loci achieving genome-wide significance in sex-combined and/or sex-specific meta-analyses																			
rs7162542	15	82,305,294	*ADAMTSL3*	C	G	0.444	-0.0379	0.0034	**9.75E-29**	229,892	-0.0369	0.0043	**1.17E-17**	126,671	-0.0391	0.0048	**2.86E-16**	103,414	7.2E-01
rs984222	1	119,305,366	*TBX15*	C	G	0.375	-0.0361	0.0035	**1.48E-25**	231,215	-0.0360	0.0044	**3.32E-16**	127,407	-0.0360	0.0049	**2.45E-13**	104,001	1.0E+00
rs4239436	18	18,985,928	*CABLES1*	A	G	0.210	-0.0405	0.0041	**1.01E-22**	229,607	-0.0364	0.0054	**1.14E-11**	126,423	-0.0450	0.0058	**1.46E-14**	103,377	2.5E-01
rs979012	20	6,571,374	*BMP2*	T	C	0.345	0.0328	0.0036	**5.43E-20**	229,815	0.0337	0.0047	**4.63E-13**	126,538	0.0303	0.0051	**2.36E-09**	103,471	6.0E-01
rs1776897	6	34,302,989	*HMGA1*	T	G	0.919	-0.0613	0.0067	**5.58E-20**	197,374	-0.0715	0.0087	**1.50E-16**	110,346	-0.0476	0.0096	7.90E-07	87,220	**4.9E-02**
rs3791679	2	55,950,396	*EFEMP1*	A	G	0.761	0.0353	0.0039	**2.06E-19**	228,968	0.0210	0.0051	3.76E-05	126,097	0.0525	0.0056	**3.12E-21**	103,063	**9.1E-06**
rs2745353	6	127,494,628	*RSPO3*	T	C	0.511	0.0293	0.0033	**7.88E-19**	231,143	0.0349	0.0044	**1.28E-15**	127,339	0.0220	0.0047	3.14E-06	103,997	**3.2E-02**
rs2071449	12	52,714,278	*HOXC5, HOXC4, HOXC5*	A	C	0.369	0.0315	0.0036	**2.47E-18**	226,567	0.0293	0.0047	**4.53E-10**	123,896	0.0337	0.0050	**1.65E-11**	102,864	4.9E-01
rs849140	7	28,150,227	*JAZF1*	T	C	0.404	0.0288	0.0034	**4.74E-17**	228,910	0.0317	0.0044	**5.41E-13**	126,079	0.0245	0.0048	3.72E-07	103,023	2.4E-01
rs9435732	1	17,180,745	*MFAP2*	T	C	0.272	-0.0309	0.0038	**4.12E-16**	228,579	-0.0272	0.0049	**3.01E-08**	125,711	-0.0365	0.0054	**9.35E-12**	103,061	1.7E-01
rs395962	6	105,504,111	*LIN28B*	T	G	0.324	0.0288	0.0036	**1.33E-15**	231,306	0.0223	0.0047	1.97E-06	127,447	0.0358	0.0051	**2.10E-12**	104,052	**3.8E-02**
rs806794	6	26,308,656	*HIST1H2BF*	A	G	0.692	0.0298	0.0037	**1.89E-15**	224,418	0.0261	0.0049	7.88E-08	124,567	0.0350	0.0053	**4.24E-11**	100,045	1.9E-01
rs4141278	7	25,824,050	*AK057379*	T	C	0.801	-0.0335	0.0043	**3.39E-15**	231,233	-0.0363	0.0055	**4.78E-11**	127,415	-0.0307	0.0059	2.06E-07	104,011	4.6E-01
rs998584	6	43,865,874	*VEGFA*	A	C	0.485	0.0293	0.0038	**6.45E-15**	210,814	0.0564	0.0049	**4.15E-30**	117,469	-0.0031	0.0053	5.57E-01	93,538	**0.0E+00**
rs12991495	2	25,340,274	*DNMT3A*	T	C	0.690	0.0281	0.0037	**6.18E-14**	229,964	0.0259	0.0049	1.52E-07	126,730	0.0318	0.0053	**2.31E-09**	103,427	3.8E-01
rs710841	4	82,368,855	*PRKG2*	T	C	0.251	0.0287	0.0038	**8.49E-14**	230,174	0.0211	0.0050	2.86E-05	126,856	0.0377	0.0055	**5.95E-12**	103,512	**1.7E-02**
rs757608	17	56,852,059	*C17orf82*	A	G	0.319	0.0268	0.0036	**9.98E-14**	229,039	0.0265	0.0047	**1.59E-08**	126,141	0.0273	0.0051	8.31E-08	103,091	9.0E-01
rs17451107	3	158,280,303	*AK311218*	T	C	0.615	0.0265	0.0036	**1.25E-13**	227,636	0.0322	0.0046	**3.51E-12**	125,592	0.0198	0.0050	7.42E-05	102,237	5.1E-02
rs11205277	1	148,159,496	*SF3B4*	A	G	0.586	-0.0270	0.0036	**1.34E-13**	215,898	-0.0244	0.0047	2.36E-07	119,052	-0.0294	0.0052	**1.25E-08**	97,039	4.5E-01
rs1812175	4	145,794,294	*HHIP*	A	G	0.165	-0.0330	0.0045	**3.95E-13**	230,433	-0.0256	0.0059	1.46E-05	126,983	-0.0432	0.0064	**1.83E-11**	103,643	**3.1E-02**

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Effect allele freq.	Sex-combined analysis				Female-specific analysis				Male-specific analysis				Sex het. *P*
							β	SE	*P*	*N*	β	SE	*P*	*N*	β	SE	*P*	*N*	
rs1344674	3	142,607,876	*ZBTB38*	A	G	0.555	-0.0240	0.0033	**4.33E-13**	231,241	-0.0180	0.0043	3.05E-05	127,417	-0.0323	0.0047	**9.50E-12**	104,018	**1.7E-02**
rs991967	1	216,682,074	*TGFB2*	A	C	0.708	-0.0260	0.0037	**1.13E-12**	230,157	-0.0232	0.0048	1.17E-06	126,792	-0.0285	0.0052	**4.55E-08**	103,557	4.2E-01
rs2274432	1	182,287,568	*TSEN15*	A	G	0.362	0.0251	0.0036	**1.75E-12**	227,843	0.0267	0.0046	**6.80E-09**	125,694	0.0230	0.0050	4.08E-06	102,341	5.6E-01
rs6556301	5	176,460,183	*FGFR4*	T	G	0.357	0.0278	0.0039	**1.82E-12**	191,245	0.0218	0.0051	2.00E-05	108,078	0.0356	0.0057	**6.23E-10**	83,358	5.4E-02
rs473902	9	97,296,056	*PTCH1*	T	G	0.916	0.0490	0.0071	**4.35E-12**	204,544	0.0504	0.0093	5.35E-08	112,714	0.0485	0.0099	8.87E-07	92,023	8.8E-01
rs4886782	15	72,015,863	*LOXL1*	A	G	0.373	-0.0245	0.0036	**5.98E-12**	228,446	-0.0177	0.0046	1.42E-04	125,776	-0.0315	0.0050	**3.22E-10**	102,863	**3.0E-02**
rs12679556	8	72,676,782	*BC048982*	T	G	0.753	-0.0263	0.0039	**1.26E-11**	225,056	-0.0345	0.0051	**1.05E-11**	125,085	-0.0157	0.0056	5.18E-03	100,164	**8.1E-03**
rs798489	7	2,768,329	*GNA12*	T	C	0.292	-0.0252	0.0037	**1.28E-11**	230,932	-0.0177	0.0048	2.60E-04	127,232	-0.0355	0.0053	**1.84E-11**	103,893	**7.9E-03**
rs6772896	3	135,686,037	*ANAPC13*	T	C	0.685	0.0242	0.0036	**1.79E-11**	231,246	0.0231	0.0047	8.33E-07	127,422	0.0247	0.0051	9.80E-07	104,017	8.1E-01
rs6570507	6	142,721,265	*GPR126*	A	G	0.286	-0.0245	0.0037	**5.73E-11**	228,993	-0.0233	0.0049	1.64E-06	126,106	-0.0268	0.0053	4.96E-07	103,080	6.0E-01
rs2638953	12	28,425,682	*CCDC91*	C	G	0.683	0.0237	0.0036	**6.53E-11**	228,074	0.0207	0.0047	9.51E-06	125,951	0.0268	0.0051	1.30E-07	102,316	3.5E-01
rs10516107	5	173,280,762	*CPEB4*	A	G	0.326	0.0232	0.0036	**8.29E-11**	231,310	0.0290	0.0046	**3.36E-10**	127,451	0.0160	0.0050	1.49E-03	104,051	**4.1E-02**
rs1879529	15	87,215,299	*ACAN*	T	G	0.284	-0.0239	0.0038	**2.86E-10**	224,276	-0.0222	0.0050	7.55E-06	124,928	-0.0235	0.0055	1.72E-05	99,541	8.5E-01
rs10041657	5	108,180,327	*FER*	A	G	0.222	0.0251	0.0040	**2.88E-10**	230,824	0.0264	0.0052	4.05E-07	127,446	0.0238	0.0057	3.26E-05	103,570	7.2E-01
rs1173770	5	32,856,968	*C5orf23*	T	C	0.402	0.0214	0.0034	**4.77E-10**	231,275	0.0118	0.0044	7.16E-03	127,421	0.0328	0.0048	**9.03E-12**	104,047	**5.8E-04**
rs3862030	10	104,317,574	*SUFU*	A	G	0.563	0.0206	0.0033	**5.83E-10**	231,146	0.0275	0.0043	**1.77E-10**	127,315	0.0126	0.0048	8.06E-03	104,024	**1.4E-02**
rs272869	5	131,705,896	*SLC22A4, BC043424*	A	G	0.404	-0.0212	0.0034	**6.68E-10**	229,935	-0.0183	0.0044	3.19E-05	126,599	-0.0250	0.0048	2.02E-07	103,528	2.7E-01
rs12493901	3	173,404,749	*FNDC3B*	A	G	0.485	-0.0209	0.0034	**8.28E-10**	230,668	-0.0164	0.0044	1.76E-04	127,119	-0.0279	0.0048	**5.47E-09**	103,742	5.9E-02
rs3760318	17	26,271,841	*ADAP2*	A	G	0.376	-0.0213	0.0035	**9.05E-10**	228,998	-0.0181	0.0045	4.78E-05	126,117	-0.0263	0.0049	7.51E-08	103,074	1.9E-01
rs13173241	5	55,897,116	*MAP3K1*	A	G	0.193	0.0259	0.0043	**1.65E-09**	229,628	0.0484	0.0055	**2.55E-18**	126,424	-0.0019	0.0061	7.59E-01	103,398	**6.3E-11**
rs780159	10	80,577,153	*ZMIZ1*	A	G	0.416	-0.0212	0.0035	**1.75E-09**	220,810	-0.0182	0.0045	5.04E-05	122,728	-0.0244	0.0050	9.35E-07	98,275	3.3E-01
rs9860730	3	64,676,186	*BC040632*	A	G	0.703	0.0219	0.0037	**2.07E-09**	231,227	0.0396	0.0048	**7.80E-17**	127,402	0.0009	0.0052	8.61E-01	104,018	**5.3E-09**
rs7166081	15	65,279,355	*DKFZp566A0646*	A	G	0.770	0.0236	0.0039	**2.12E-09**	230,255	0.0170	0.0051	9.35E-04	126,884	0.0321	0.0056	**1.14E-08**	103,561	**3.3E-02**
rs4868125	5	171,214,480	*FBXW11*	C	G	0.399	-0.0215	0.0036	**2.93E-09**	225,860	-0.0218	0.0047	4.15E-06	124,757	-0.0215	0.0052	3.11E-05	101,296	9.6E-01

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Effect allele freq.	Sex-combined analysis				Female-specific analysis				Male-specific analysis				Sex het. *P*
							β	SE	*P*	*N*	β	SE	*P*	*N*	β	SE	*P*	*N*	
rs12207675	6	76,294,461	*FILIP1*	T	C	0.883	-0.0308	0.0052	**3.14E-09**	231,322	-0.0304	0.0067	6.64E-06	127,458	-0.0307	0.0073	2.71E-05	104,058	9.7E-01
rs2214442	7	20,359,312	*ITGB8*	A	G	0.545	-0.0265	0.0045	**3.89E-09**	152,053	-0.0252	0.0056	6.80E-06	91,400	-0.0282	0.0068	3.37E-05	60,846	7.2E-01
rs648514	3	52,442,303	*SEMA3G*	A	G	0.464	0.0199	0.0034	**5.40E-09**	229,422	0.0287	0.0044	**5.11E-11**	126,328	0.0100	0.0048	3.65E-02	103,287	**2.2E-03**
rs4246302	15	98,505,490	*ADAMTS17*	A	G	0.681	-0.0216	0.0037	**5.73E-09**	227,205	-0.0248	0.0048	2.88E-07	125,027	-0.0177	0.0052	6.74E-04	102,371	2.8E-01
rs12317176	12	122,970,671	*KIAA2017*	T	C	0.663	0.0205	0.0035	**5.88E-09**	230,924	0.0342	0.0046	**1.29E-13**	127,243	0.0030	0.0050	5.47E-01	103,874	**9.5E-07**
rs2124969	2	160,697,732	*ITGB6*	T	C	0.578	-0.0199	0.0034	**7.06E-09**	231,284	-0.0157	0.0044	3.48E-04	127,437	-0.0247	0.0048	2.33E-07	104,039	1.4E-01
rs12127195	1	219,376,040	*C1orf140*	A	G	0.297	0.0211	0.0037	**7.67E-09**	231,166	0.0148	0.0048	2.00E-03	127,373	0.0283	0.0051	**3.36E-08**	103,986	**4.0E-02**
rs3786897	19	38,584,848	*PEPD*	A	G	0.582	-0.0199	0.0035	**8.77E-09**	228,567	-0.0232	0.0044	1.69E-07	125,873	-0.0152	0.0048	1.66E-03	102,887	1.9E-01
rs7621331	3	137,244,617	*PPP2R3A*	A	G	0.680	0.0207	0.0036	**9.41E-09**	231,264	0.0135	0.0047	4.03E-03	127,428	0.0276	0.0051	**4.70E-08**	104,029	**3.0E-02**
rs606452	11	74,953,826	*SERPINH1*	A	C	0.145	0.0276	0.0048	**1.10E-08**	229,378	0.0242	0.0063	1.34E-04	126,277	0.0324	0.0068	1.77E-06	103,294	3.5E-01
rs1784203	11	93,089,782	*KIAA1731*	A	G	0.010	0.0307	0.0054	**1.31E-08**	63,892	0.0004	0.0485	9.93E-01	35,539	0.0745	0.0082	**1.03E-19**	28,353	4.1E-01
rs6715793	2	33,232,767	*LTBP1*	T	C	0.513	0.0193	0.0034	**1.44E-08**	231,071	0.0163	0.0044	1.83E-04	127,302	0.0236	0.0048	7.10E-07	103,962	2.3E-01
rs12608504	19	18,250,135	*JUND*	A	G	0.357	0.0201	0.0036	**1.48E-08**	228,998	0.0182	0.0046	7.39E-05	126,113	0.0224	0.0050	6.65E-06	103,077	5.1E-01
rs2052670	2	66,071,985	*AK131224*	A	G	0.656	-0.0200	0.0035	**1.53E-08**	231,210	-0.0202	0.0046	1.23E-05	127,394	-0.0184	0.0050	2.05E-04	104,009	7.8E-01
rs4542783	19	8,548,160	*MYO1F*	T	C	0.540	0.0227	0.0040	**1.69E-08**	170,860	0.0194	0.0052	2.05E-04	96,281	0.0265	0.0057	2.93E-06	74,773	3.3E-01
rs11144688	9	77,732,106	*PCSK5*	A	G	0.126	-0.0336	0.0060	**1.87E-08**	214,513	-0.0329	0.0081	4.44E-05	118,374	-0.0358	0.0082	1.18E-05	96,332	7.9E-01
rs16957304	16	65,892,470	*KCTD19*	A	G	0.956	0.0591	0.0106	**2.47E-08**	151,917	0.0596	0.0131	5.37E-06	91,174	0.0576	0.0160	3.18E-04	60,936	9.2E-01
rs12330322	3	72,538,045	*RYBP*	T	C	0.215	-0.0223	0.0040	**3.15E-08**	229,015	-0.0165	0.0053	1.66E-03	126,120	-0.0295	0.0058	3.64E-07	103,088	7.7E-02
rs2294239	22	27,779,477	*ZNRF3*	A	G	0.585	0.0191	0.0035	**3.49E-08**	228,463	0.0261	0.0044	**4.19E-09**	125,841	0.0124	0.0049	1.09E-02	102,814	**2.6E-02**
rs822531	7	148,260,692	*RNY5*	T	C	0.782	0.0244	0.0044	**3.71E-08**	226,664	0.0196	0.0059	8.98E-04	124,836	0.0297	0.0062	1.57E-06	102,021	2.1E-01
rs7970350	12	64,646,431	*HMGA2*	T	C	0.494	-0.0188	0.0034	**3.76E-08**	229,815	-0.0142	0.0044	1.11E-03	126,536	-0.0240	0.0048	5.10E-07	103,472	1.1E-01
rs7684221	4	17,566,452	*LCORL*	A	G	0.164	-0.0259	0.0047	**4.15E-08**	229,360	-0.0181	0.0061	3.23E-03	126,312	-0.0341	0.0067	3.23E-07	103,241	5.9E-02
rs9977276	21	46,260,755	*COL6A1*	T	G	0.223	-0.0218	0.0040	**4.37E-08**	229,565	-0.0275	0.0052	1.31E-07	126,398	-0.0150	0.0057	8.50E-03	103,360	8.4E-02
rs2160077	14	91,498,163	*TRIP11*	A	G	0.421	-0.0181	0.0033	**4.46E-08**	231,255	-0.0119	0.0043	5.74E-03	127,415	-0.0264	0.0047	**2.62E-08**	104,033	**1.5E-02**
rs2047937	16	48,422,292	*ZNF423*	T	C	0.503	-0.0186	0.0034	**4.67E-08**	231,009	-0.0218	0.0044	5.49E-07	127,288	-0.0140	0.0048	3.57E-03	103,914	2.0E-01

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Effect allele freq.	Sex-combined analysis				Female-specific analysis				Male-specific analysis				Sex het. *P*
							β	SE	*P*	*N*	β	SE	*P*	*N*	β	SE	*P*	*N*	
rs6437061	2	232,893,296	*DIS3L2*	A	C	0.602	0.0166	0.0035	1.77E-06	229,808	0.0079	0.0044	7.42E-02	126,658	0.0273	0.0049	**2.50E-08**	103,343	**1.7E-03**
rs7830933	8	23,659,269	*NKX3-1*	A	G	0.768	0.0174	0.0039	9.01E-06	231,029	0.0328	0.0051	**1.93E-10**	127,293	-0.0017	0.0056	7.55E-01	103,929	**1.2E-06**
rs10925060	1	245,717,763	*OR2W5*	T	C	0.024	0.0174	0.0041	2.20E-05	140,515	0.0021	0.0051	6.81E-01	85,186	0.0450	0.0063	**9.14E-13**	55,522	**1.7E-08**
rs3897379	1	217,826,356	*ZNT8*	A	G	0.188	0.0160	0.0043	2.12E-04	231,225	0.0314	0.0055	**1.53E-08**	127,406	-0.0018	0.0061	7.66E-01	104,012	**1.6E-05**
rs10195252	2	165,221,337	*COBLL1*	T	C	0.592	0.0114	0.0034	9.44E-04	229,802	0.0345	0.0044	**3.74E-15**	126,613	-0.0164	0.0048	6.71E-04	103,381	**0.0E+00**
rs17472426	5	159,626,935	*CCNJL*	T	G	0.920	0.0144	0.0067	3.11E-02	217,564	-0.0140	0.0086	1.05E-01	119,804	0.0520	0.0095	**4.34E-08**	97,954	**3.9E-08**
Previously reported WC loci																			
rs10146997	14	79,014,915	*NRXN3*	A	G	0.782	-0.0142	0.0052	6.3E-03	151,207	-0.0201	0.0063	1.4E-03	90,855	-0.0082	0.0077	2.9E-01	60,545	**3.6E-02**
rs9939609	16	52,378,028	*FTO*	A	T	0.405	0.0078	0.0035	2.4E-02	227,252	0.0081	0.0044	6.5E-02	126,407	0.0070	0.0048	1.4E-01	101,038	7.6E-01
rs17782313	18	56,002,077	*MC4R*	T	C	0.763	-0.0048	0.0043	2.7E-01	194,234	-0.0100	0.0055	7.0E-02	109,300	0.0013	0.0062	8.3E-01	85,126	1.5E-02
rs545854[a]	8	9,897,490	*MSRA*	C	G	0.814	0.0101	0.0098	3.0E-01	54,328	0.0144	0.0131	2.7E-01	27,795	0.0056	0.0135	6.8E-01	26,533	4.0E-01
rs987237	6	50,911,009	*TFAP2B*	A	G	0.816	-0.0023	0.0044	6.0E-01	228,913	-0.0002	0.0057	9.8E-01	126,065	-0.0041	0.0062	5.1E-01	103,041	4.1E-01

SNP frequencies (freq.), effect betas (β), standard errors (SE), and *P* values are reported in terms of the effect allele. SNP sample sizes (*N*) are reported for each specific meta-analysis. Main effect $P < 5 \times 10^{-8}$ and sex heterogeneity $P < 0.05$ are marked in boldface.

[a] Renamed from rs7826222 in NCBI Build 36.

Table 4.2 – Conditional analysis identifies 117 unique SNPs independently associated with waist circumference in sex-combined and sex-specific meta-analyses

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Single-marker analysis					Joint analysis					1000G CEU r^2 with lead SNP
						Freq.	β	SE	*P*	*N*	Freq.	β	SE	*P*	LD *r*	
Independent SNPs from sex-combined conditional analysis																
rs9435732	1	17,180,745	*MFAP2*	T	C	0.272	-0.0309	0.0038	**4.12E-16**	244,275	0.269	-0.0309	0.0038	**4.28E-16**	0	same
rs7536458	1	118,666,125	***SPAG17***	T	G	0.747	0.0305	0.0038	**1.24E-15**	255,851	0.768	0.0325	0.0038	**2.77E-17**	0.128	0.107
rs6701231	1	119,297,184	***WARS2***	C	G	0.618	-0.0268	0.0035	**2.23E-14**	241,227	0.562	-0.0230	0.0037	**2.99E-10**	0.258	0.022
rs984222	1	119,305,366	***WARS2***	C	G	0.375	-0.0361	0.0035	**1.48E-25**	243,059	0.313	-0.0291	0.0036	**1.13E-15**	0	same
rs11205277	1	148,159,496	*SF3B4*	A	G	0.586	-0.0270	0.0036	**1.34E-13**	221,818	0.561	-0.0270	0.0036	**6.43E-14**	0	same
rs2274432	1	182,287,568	*C1orf19*	A	G	0.362	0.0251	0.0036	**1.75E-12**	233,238	0.381	0.0251	0.0036	**3.14E-12**	0	same
rs991967	1	216,682,074	*TGFB2*	A	C	0.708	-0.0260	0.0037	**1.13E-12**	246,424	0.707	-0.0265	0.0037	**7.52E-13**	0.025	same
rs12127195	1	219,376,040	*C1orf140*	A	G	0.297	0.0211	0.0037	**7.67E-09**	244,018	0.317	0.0218	0.0037	**4.15E-09**	0	same
rs12991495	2	25,340,274	*DNMT3A*	T	C	0.690	0.0281	0.0037	**6.18E-14**	238,298	0.693	0.0281	0.0037	**3.11E-14**	0	same
rs3791679	2	55,950,396	*EFEMP1*	A	G	0.761	0.0353	0.0039	**2.06E-19**	252,470	0.776	0.0353	0.0039	**1.43E-19**	0	same
rs2052670	2	66,071,985	*FLJ16124*	A	G	0.656	-0.0200	0.0035	**1.53E-08**	252,481	0.626	-0.0200	0.0035	**1.10E-08**	0	same
rs2124969	2	160,697,732	*ITGB6*	T	C	0.578	-0.0199	0.0034	**7.06E-09**	247,413	0.548	-0.0199	0.0034	**4.84E-09**	0	same
rs13083798	3	52,624,788	*PB1*	A	G	0.491	0.0202	0.0034	**3.37E-09**	241,521	0.485	0.0202	0.0034	**2.84E-09**	0	0.565
rs9864077	3	64,679,931	*ADAMTS9*	T	C	0.697	0.0225	0.0037	**1.27E-09**	241,109	0.705	0.0225	0.0037	**1.12E-09**	0.002	0.960
rs12330322	3	72,538,045	*RYBP*	T	C	0.215	-0.0223	0.0040	**3.15E-08**	258,482	0.208	-0.0223	0.0040	**2.33E-08**	0	same
rs6766897	3	135,673,993	*ANAPC13*	A	C	0.689	0.0241	0.0036	**2.88E-11**	251,159	0.717	0.0239	0.0036	**3.59E-11**	-0.026	1.000
rs7621331	3	137,244,617	*PPP2R3A*	A	G	0.680	0.0207	0.0036	**9.41E-09**	247,588	0.691	0.0221	0.0036	**7.94E-10**	0.033	same
rs1344674	3	142,607,876	*ZBTB38*	A	G	0.555	-0.0240	0.0033	**4.33E-13**	259,426	0.519	-0.0240	0.0033	**4.06E-13**	0	same
rs17451107	3	158,280,303	*FLJ16641*	T	C	0.615	0.0265	0.0036	**1.25E-13**	227,385	0.617	0.0265	0.0036	**1.84E-13**	0	same
rs12493901	3	173,404,749	*FNDC3B*	A	G	0.485	-0.0209	0.0034	**8.28E-10**	241,666	0.477	-0.0209	0.0034	**7.92E-10**	0	same

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Single-marker analysis					Joint analysis					1000G CEU r^2 with lead SNP
						Freq.	β	SE	*P*	*N*	Freq.	β	SE	*P*	LD *r*	
rs7684221	4	17,566,452	*LCORL*	A	G	0.164	-0.0259	0.0047	**4.15E-08**	230,275	0.110	-0.0259	0.0047	**3.58E-08**	0	same
rs710841	4	82,368,855	*PRKG2*	T	C	0.251	0.0287	0.0038	**8.49E-14**	256,871	0.231	0.0287	0.0038	**4.29E-14**	0	same
rs1812175	4	145,794,294	*HHIP*	A	G	0.165	-0.0330	0.0045	**3.95E-13**	249,713	0.182	-0.0330	0.0045	**2.26E-13**	0	same
rs1173734	5	32,802,174	***NPR3***	T	C	0.742	-0.0180	0.0038	2.94E-06	252,298	0.772	-0.0238	0.0039	**8.71E-10**	0.199	0.036
rs12656497	5	32,867,696	***C5orf23***	T	C	0.404	0.0218	0.0034	**2.03E-10**	250,633	0.427	0.0260	0.0035	**6.34E-14**	0	0.928
rs459193	5	55,842,508	*ANKRD55*	A	G	0.264	0.0247	0.0038	**7.66E-11**	248,801	0.251	0.0247	0.0038	**8.06E-11**	0	0.095
rs10041657	5	108,180,327	*FER*	A	G	0.222	0.0251	0.0040	**2.88E-10**	252,897	0.210	0.0251	0.0040	**3.51E-10**	0	same
rs272869	5	131,705,896	*SLC22A4*	A	G	0.404	-0.0212	0.0034	**6.68E-10**	250,715	0.404	-0.0212	0.0034	**4.52E-10**	0	same
rs4868125	5	171,214,480	*FBXW11*	C	G	0.399	-0.0215	0.0036	**2.93E-09**	224,570	0.417	-0.0219	0.0036	**1.10E-09**	0.024	same
rs1976074	5	173,317,269	*CPEB4*	A	T	0.328	0.0232	0.0036	**9.42E-11**	244,292	0.331	0.0248	0.0036	**5.97E-12**	0	0.964
rs6556301	5	176,460,183	*FGFR4*	T	G	0.357	0.0278	0.0039	**1.82E-12**	199,847	0.356	0.0288	0.0039	**1.66E-13**	0	same
rs4145878	6	26,306,025	***HIST1H3D***	A	T	0.500	-0.0234	0.0049	1.79E-06	116,237	0.468	-0.0903	0.0059	**3.94E-53**	0.336	0.363
rs806794	6	26,308,656	***HIST1H2BF***	A	G	0.692	0.0298	0.0037	**1.89E-15**	239,146	0.683	0.0679	0.0046	**6.88E-49**	-0.227	same
rs994379	6	26,413,386	***CR593845***	C	G	0.472	-0.0226	0.0041	**3.54E-08**	166,545	0.523	-0.3869	0.0109	**1.67E-273**	0.858	0.196
rs9358925	6	26,430,747	***CR593845***	C	G	0.625	0.0283	0.0043	**4.66E-11**	161,041	0.604	0.3890	0.0115	**4.82E-252**	-0.107	0.338
rs3799380	6	26,575,161	***BTN2A1***	T	C	0.816	0.0099	0.0044	2.44E-02	240,274	0.844	0.0473	0.0054	**2.73E-18**	-0.007	0.075
rs2754603	6	27,073,118	***MGC22265, GUSBL1***	T	C	0.680	0.0048	0.0045	2.86E-01	158,424	0.693	0.0399	0.0053	**7.38E-14**	-0.436	0.064
rs13196692	6	27,487,098	***AX747641***	T	C	0.086	-0.0133	0.0077	8.41E-02	149,422	0.072	-0.1903	0.0117	**1.28E-59**	-0.366	> 500 kb
rs276369	6	28,027,222	***OR2B6***	A	T	0.654	0.0175	0.0044	6.97E-05	159,323	0.665	0.0340	0.0047	**8.29E-13**	0	> 500 kb
rs1776897	6	34,302,989	***NR_001561***	T	G	0.919	-0.0613	0.0067	**5.58E-20**	209,464	0.929	-0.0538	0.0067	**1.55E-15**	-0.048	same
rs13210323	6	35,113,062	***ANKS1A***	A	C	0.737	0.0216	0.0038	**1.45E-08**	249,297	0.776	0.0224	0.0038	**3.91E-09**	-0.033	0.003
rs998584	6	43,865,874	*VEGFA*	A	C	0.485	0.0293	0.0038	**6.45E-15**	193,433	0.460	0.0292	0.0038	**1.51E-14**	0	same
rs12207675	6	76,294,461	*FILIP1*	T	C	0.883	-0.0308	0.0052	**3.14E-09**	248,853	0.866	-0.0308	0.0052	**3.17E-09**	0	same
rs395962	6	105,504,111	*LIN28B*	T	G	0.324	0.0288	0.0036	**1.33E-15**	245,982	0.342	0.0293	0.0036	**3.97E-16**	0.026	same

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Single-marker analysis					Joint analysis					1000G CEU r^2 with lead SNP
						Freq.	β	SE	*P*	*N*	Freq.	β	SE	*P*	LD *r*	
rs9400239	6	109,084,356	*FOXO3A*	T	C	0.315	-0.0190	0.0036	1.66E-07	249,608	0.273	-0.0198	0.0036	**4.06E-08**	0	NA
rs2745353	6	127,494,628	*RSPO3*	T	C	0.511	0.0293	0.0033	**7.88E-19**	256,380	0.535	0.0293	0.0033	**6.84E-19**	0	same
rs9389986	6	142,702,807	*DREG, GPR126*	A	T	0.285	-0.0245	0.0037	**5.73E-11**	250,006	0.288	-0.0245	0.0037	**3.57E-11**	0	1.000
rs798489	7	2,768,329	*GNA12*	T	C	0.292	-0.0252	0.0037	**1.28E-11**	246,328	0.289	-0.0252	0.0037	**9.75E-12**	0	same
rs2214442	7	20,359,312	*ITGB8*	A	G	0.545	-0.0265	0.0045	**3.89E-09**	138,954	0.608	-0.0272	0.0045	**1.49E-09**	0	same
rs7798002	7	25,827,105	*NFE2L3*	T	G	0.204	0.0327	0.0042	**6.97E-15**	243,653	0.185	0.0332	0.0042	**2.52E-15**	0.002	1.000
rs2067087	7	27,208,185	***HOXA13***	C	G	0.694	-0.0230	0.0038	**2.18E-09**	227,617	0.694	-0.0261	0.0038	**7.31E-12**	0.082	0.002
rs849140	7	28,150,227	***JAZF1***	T	C	0.404	0.0288	0.0034	**4.74E-17**	250,602	0.412	0.0314	0.0034	**3.60E-20**	0	same
rs822531	7	148,260,692	*NR_001571*	T	C	0.782	0.0244	0.0044	**3.71E-08**	211,273	0.790	0.0244	0.0044	**2.94E-08**	0	same
rs12679556	8	72,676,782	*MSC*	T	G	0.753	-0.0263	0.0039	**1.26E-11**	246,304	0.771	-0.0263	0.0039	**1.55E-11**	0	same
rs11144688	9	77,732,106	*PCSK5*	A	G	0.126	-0.0336	0.0060	**1.87E-08**	176,235	0.133	-0.0336	0.0060	**2.15E-08**	0	same
rs473902	9	97,296,056	***PTCH1***	T	G	0.916	0.0490	0.0071	**4.35E-12**	180,259	0.909	0.0455	0.0071	**1.66E-10**	0.088	same
rs7854560	9	97,422,771	***PTCH1***	T	C	0.296	0.0259	0.0037	**4.78E-12**	244,525	0.297	0.0241	0.0037	**8.25E-11**	0	0.005
rs780159	10	80,577,153	*ZMIZ1*	A	G	0.416	-0.0212	0.0035	**1.75E-09**	234,387	0.419	-0.0212	0.0035	**1.39E-09**	0	same
rs11191295	10	104,221,044	*UNQ5807, TMEM180*	A	G	0.471	-0.0226	0.0036	**5.40E-10**	216,059	0.490	-0.0226	0.0036	**3.45E-10**	0	0.799
rs606452	11	74,953,826	*SERPINH1*	A	C	0.145	0.0276	0.0048	**1.10E-08**	244,560	0.152	0.0276	0.0048	**8.94E-09**	0	same
rs2638953	12	28,425,682	*CCDC91*	C	G	0.683	0.0237	0.0036	**6.53E-11**	248,616	0.674	0.0237	0.0036	**4.62E-11**	0	same
rs2071449	12	52,714,278	*HOXC5, HOXC4*	A	C	0.369	0.0315	0.0036	**2.47E-18**	231,107	0.392	0.0315	0.0036	**2.16E-18**	0	same
rs7970350	12	64,646,431	*HMGA2*	T	C	0.494	-0.0188	0.0034	**3.76E-08**	241,476	0.463	-0.0188	0.0034	**3.22E-08**	0	same
rs12317176	12	122,970,671	*KIAA2017*	T	C	0.663	0.0205	0.0035	**5.88E-09**	254,895	0.650	0.0205	0.0035	**4.72E-09**	0	same
rs2160077	14	91,498,163	*TRIP11*	A	G	0.421	-0.0181	0.0033	**4.46E-08**	262,877	0.396	-0.0181	0.0033	**4.15E-08**	0	same
rs7166081	15	65,279,355	*FLJ11506*	A	G	0.770	0.0236	0.0039	**2.12E-09**	258,638	0.753	0.0231	0.0039	**3.05E-09**	-0.019	same
rs4886782	15	72,015,863	*LOXL1*	A	G	0.373	-0.0245	0.0036	**5.98E-12**	230,295	0.397	-0.0253	0.0036	**2.34E-12**	0.063	same

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Single-marker analysis					Joint analysis					1000G CEU r^2 with lead SNP
						Freq.	β	SE	*P*	*N*	Freq.	β	SE	*P*	LD *r*	
rs4886648	15	73,095,890	*SCAMP5*	A	G	0.434	0.0169	0.0037	4.34E-06	207,519	0.459	0.0204	0.0037	**3.76E-08**	0.041	NA
rs2562779	15	82,181,979	***ADAMTSL3***	A	G	0.818	0.0461	0.0063	**2.53E-13**	117,998	0.974	0.0391	0.0063	**6.96E-10**	0	0.027
rs7162542	15	82,305,294	***ADAMTSL3***	C	G	0.444	-0.0379	0.0034	**9.75E-29**	244,448	0.453	-0.0364	0.0034	**1.94E-26**	-0.006	same
rs1879529	15	87,215,299	*ACAN*	T	G	0.284	-0.0239	0.0038	**2.86E-10**	237,726	0.310	-0.0240	0.0038	**2.75E-10**	0	same
rs4246302	15	98,505,490	***ADAMTS17***	A	G	0.681	-0.0216	0.0037	**5.73E-09**	234,578	0.717	-0.0220	0.0037	**2.73E-09**	0.019	same
rs4567683	15	98,596,360	***ADAMTS17***	A	G	0.267	0.0219	0.0038	**7.67E-09**	246,783	0.240	0.0223	0.0038	**4.27E-09**	0	0.009
rs2047937	16	48,422,292	*ZNF423*	T	C	0.503	-0.0186	0.0034	**4.67E-08**	241,447	0.453	-0.0186	0.0034	**4.49E-08**	0	same
rs16957358	16	65,952,042	*LRRC36*	A	G	0.957	0.0463	0.0083	**2.70E-08**	244,505	0.962	0.0463	0.0083	**2.43E-08**	0	0.654
rs3760318	17	26,271,841	*CENTA2*	A	G	0.376	-0.0213	0.0035	**9.05E-10**	242,869	0.336	-0.0213	0.0035	**1.16E-09**	0	same
rs757608	17	56,852,059	*C17orf82*	A	G	0.319	0.0268	0.0036	**9.98E-14**	247,648	0.311	0.0268	0.0036	**9.80E-14**	0	same
rs4239436	18	18,985,928	*CABLES1*	A	G	0.210	-0.0405	0.0041	**1.01E-22**	250,051	0.223	-0.0405	0.0041	**5.28E-23**	0	same
rs4542783	19	8,548,160	*MYO1F*	T	C	0.540	0.0227	0.0040	**1.69E-08**	175,543	0.562	0.0229	0.0040	**1.10E-08**	-0.008	same
rs12608504	19	18,250,135	*JUND*	A	G	0.357	0.0201	0.0036	**1.48E-08**	234,572	0.317	0.0202	0.0036	**1.87E-08**	0	same
rs3786897	19	38,584,848	*PEPD*	A	G	0.582	-0.0199	0.0035	**8.77E-09**	234,057	0.576	-0.0199	0.0035	**1.31E-08**	0	same
rs1884897	20	6,560,832	*BMP2*	A	G	0.360	0.0320	0.0035	**6.56E-20**	247,111	0.380	0.0320	0.0035	**6.17E-20**	0	0.852
rs2179129	22	27,780,923	*bK175E3.6*	A	G	0.588	0.0190	0.0034	**2.64E-08**	249,105	0.607	0.0190	0.0034	**2.30E-08**	0	0.807
Independent SNPs from female-specific conditional analysis																
rs9435732	1	17,180,745	*MFAP2*	T	C	0.269	-0.0272	0.0049	**3.01E-08**	133,506	0.269	-0.0272	0.0049	**2.85E-08**	0	same
rs984222	1	119,305,366	*TBX15*	C	G	0.378	-0.0360	0.0044	**3.32E-16**	138,524	0.313	-0.0360	0.0044	**2.84E-16**	0	same
rs2274432	1	182,287,568	*C1orf19*	A	G	0.359	0.0267	0.0046	**6.80E-09**	129,472	0.381	0.0267	0.0046	**6.49E-09**	0	same
rs3897379	1	217,826,356	*ZNT8*	A	G	0.189	0.0314	0.0055	**1.53E-08**	136,092	0.190	0.0314	0.0055	**1.14E-08**	0	same
rs1128249	2	165,236,870	*COBLL1*	T	G	0.406	-0.0344	0.0044	**4.94E-15**	135,111	0.439	-0.0344	0.0044	**5.43E-15**	0	0.960
rs648514	3	52,442,303	*SEMA3G*	A	G	0.463	0.0287	0.0044	**5.11E-11**	131,041	0.462	0.0287	0.0044	**6.95E-11**	0	same
rs4616635	3	64,677,315	*ADAMTS9*	C	G	0.721	0.0409	0.0049	**8.99E-17**	130,616	0.727	0.0409	0.0049	**7.14E-17**	0	0.872

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Single-marker analysis					Joint analysis					1000G CEU r^2 with lead SNP
						Freq.	β	SE	*P*	*N*	Freq.	β	SE	*P*	LD *r*	
rs17451107	3	158,280,303	*FLJ16641*	T	C	0.612	0.0322	0.0046	**3.51E-12**	125,469	0.617	0.0322	0.0046	**2.58E-12**	0	same
rs13173241	5	55,897,116	*MEKK1*	A	G	0.195	0.0484	0.0055	**2.55E-18**	132,943	0.140	0.0484	0.0055	**1.40E-18**	0	same
rs10516107	5	173,280,762	*CPEB4*	A	G	0.323	0.0290	0.0046	**3.36E-10**	136,401	0.332	0.0290	0.0046	**2.91E-10**	0	same
rs9378213	6	32,556,376	*HLA-DRA*	T	G	0.578	-0.0236	0.0045	1.49E-07	127,738	0.604	-0.0255	0.0045	**1.61E-08**	-0.001	NA
rs1776897	6	34,302,989	*NR_001561*	T	G	0.918	-0.0715	0.0087	**1.50E-16**	111,035	0.929	-0.0865	0.0091	**2.93E-21**	-0.304	same
rs9470001	6	35,441,719	*PPARD*	C	G	0.051	-0.0382	0.0109	4.49E-04	110,115	0.049	-0.0715	0.0115	**4.62E-10**	-0.028	NA
rs998584	6	43,865,874	*VEGFA*	A	C	0.484	0.0564	0.0049	**4.15E-30**	105,085	0.460	0.0543	0.0049	**1.69E-28**	0	same
rs2745349	6	127,461,527	*RSPO3*	A	C	0.705	-0.0417	0.0048	**4.56E-18**	131,669	0.706	-0.0417	0.0048	**3.79E-18**	0	0.141
rs10232819	7	25,843,137	*NFE2L3*	T	C	0.749	-0.0331	0.0050	**4.11E-11**	134,158	0.769	-0.0344	0.0050	**6.57E-12**	0	0.693
rs849140	7	28,150,227	*JAZF1*	T	C	0.404	0.0317	0.0044	**5.41E-13**	135,270	0.412	0.0327	0.0044	**1.07E-13**	0	same
rs7830933	8	23,659,269	*NKX3-1*	A	G	0.767	0.0328	0.0051	**1.93E-10**	135,852	0.783	0.0328	0.0051	**1.27E-10**	0	same
rs12679556	8	72,676,782	*MSC*	T	G	0.753	-0.0345	0.0051	**1.05E-11**	130,230	0.771	-0.0345	0.0051	**1.35E-11**	0	same
rs7075269	10	104,355,714	*SUFU*	A	G	0.436	-0.0275	0.0043	**1.93E-10**	138,756	0.428	-0.0275	0.0043	**1.61E-10**	0	1.000
rs2071449	12	52,714,278	*HOXC5, HOXC4*	A	C	0.367	0.0293	0.0047	**4.53E-10**	122,848	0.392	0.0293	0.0047	**4.57E-10**	0	same
rs12317176	12	122,970,671	***KIAA2017***	T	C	0.661	0.0342	0.0046	**1.29E-13**	133,083	0.650	0.0300	0.0047	**1.24E-10**	0.156	same
rs863750	12	123,071,397	***ZNF664***	T	C	0.595	0.0306	0.0044	**2.52E-12**	135,187	0.600	0.0262	0.0045	**4.26E-09**	0	0.000
rs7162542	15	82,305,294	*ADAMTSL3*	C	G	0.444	-0.0369	0.0043	**1.17E-17**	138,154	0.453	-0.0369	0.0043	**9.56E-18**	0	same
rs757608	17	56,852,059	*C17orf82*	A	G	0.321	0.0265	0.0047	**1.59E-08**	131,071	0.311	0.0265	0.0047	**1.72E-08**	0	same
rs7235010	18	18,978,808	*CABLES1*	A	G	0.789	0.0369	0.0053	**4.88E-12**	134,913	0.777	0.0369	0.0053	**3.38E-12**	0	1.000
rs1884897	20	6,560,832	*BMP2*	A	G	0.362	0.0349	0.0045	**5.92E-15**	134,905	0.380	0.0349	0.0045	**8.92E-15**	0	0.852
rs2294239	22	27,779,477	*bK175E3.6*	A	G	0.583	0.0261	0.0044	**4.19E-09**	134,012	0.585	0.0261	0.0044	**3.01E-09**	0	same
Independent SNPs from male-specific conditional analysis																
rs9435732	1	17,180,745	*MFAP2*	T	C	0.274	-0.0365	0.0054	**9.35E-12**	108,287	0.269	-0.0365	0.0054	**1.40E-11**	0	same
rs7513580	1	118,651,432	***SPAG17***	A	G	0.302	-0.0390	0.0055	**1.98E-12**	98,420	0.277	-0.0378	0.0055	**6.62E-12**	-0.033	0.084

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Single-marker analysis					Joint analysis					1000G CEU r^2 with lead SNP
						Freq.	β	SE	*P*	*N*	Freq.	β	SE	*P*	LD *r*	
rs984225	1	119,305,807	***WARS2***	A	G	0.630	0.0360	0.0049	**2.25E-13**	112,150	0.687	0.0350	0.0049	**1.00E-12**	0	1.000
rs11205277	1	148,159,496	*SF3B4*	A	G	0.589	-0.0294	0.0052	**1.25E-08**	95,911	0.561	-0.0294	0.0052	**1.58E-08**	0	same
rs991967	1	216,682,074	*TGFB2*	A	C	0.706	-0.0285	0.0052	**4.55E-08**	111,886	0.707	-0.0289	0.0052	**2.93E-08**	-0.015	same
rs2820426	1	217,727,158	*ZNT8*	A	G	0.380	0.0249	0.0048	2.81E-07	115,665	0.367	0.0263	0.0048	**4.74E-08**	-0.061	0.169
rs12127195	1	219,376,040	*C1orf140*	A	G	0.297	0.0283	0.0051	**3.36E-08**	115,611	0.317	0.0307	0.0051	**1.93E-09**	0	same
rs10925060	1	245,717,763	*OR2W5*	T	C	0.025	0.0450	0.0063	**9.14E-13**	646,827	0.015	0.0450	0.0063	**9.16E-13**	0	same
rs12991495	2	25,340,274	*DNMT3A*	T	C	0.692	0.0318	0.0053	**2.31E-09**	104,863	0.693	0.0318	0.0053	**1.99E-09**	0	same
rs3791679	2	55,950,396	*EFEMP1*	A	G	0.762	0.0525	0.0056	**3.12E-21**	110,283	0.776	0.0525	0.0056	**7.17E-21**	0	same
rs1515114	2	226,806,631	*IRS1*	A	G	0.525	0.0247	0.0047	1.85E-07	114,016	0.523	0.0263	0.0047	**2.36E-08**	0	NA
rs6437061	2	232,893,296	*DIS3L2*	A	C	0.606	0.0273	0.0049	**2.50E-08**	109,521	0.644	0.0289	0.0049	**4.06E-09**	0	same
rs6775778	3	137,209,369	*PPP2R3A*	A	G	0.753	0.0316	0.0055	**9.64E-09**	111,602	0.770	0.0317	0.0055	**8.70E-09**	0.001	0.686
rs1344674	3	142,607,876	*ZBTB38*	A	G	0.555	-0.0323	0.0047	**9.50E-12**	115,099	0.519	-0.0323	0.0047	**6.03E-12**	0	same
rs12493901	3	173,404,749	*FNDC3B*	A	G	0.484	-0.0279	0.0048	**5.47E-09**	109,146	0.477	-0.0279	0.0048	**6.19E-09**	0	same
rs2197271	4	82,379,372	*PRKG2*	C	G	0.751	-0.0394	0.0055	**6.30E-13**	110,893	0.771	-0.0394	0.0055	**7.95E-13**	0	0.951
rs1812175	4	145,794,294	*HHIP*	A	G	0.166	-0.0432	0.0064	**1.83E-11**	110,638	0.182	-0.0432	0.0064	**1.49E-11**	0	same
rs1173770	5	32,856,968	*C5orf23*	T	C	0.402	0.0328	0.0048	**9.03E-12**	113,411	0.426	0.0328	0.0048	**8.38E-12**	0	same
rs17472426	5	159,626,935	*CCNJL*	T	G	0.921	0.0520	0.0095	**4.34E-08**	95,096	0.942	0.0520	0.0095	**4.43E-08**	0	same
rs6556301	5	176,460,183	*FGFR4*	T	G	0.357	0.0356	0.0057	**6.23E-10**	84,157	0.356	0.0356	0.0057	**4.26E-10**	0	same
rs7773004	6	26,375,734	*HIST1H3G*	A	G	0.501	0.0334	0.0047	**1.83E-12**	113,713	0.504	0.0334	0.0047	**1.20E-12**	0	0.173
rs7759938	6	105,485,647	*LIN28B*	T	C	0.677	-0.0362	0.0051	**1.31E-12**	110,481	0.658	-0.0362	0.0051	**1.28E-12**	0	1.000
rs798497	7	2,762,483	*GNA12*	A	G	0.690	0.0344	0.0051	**1.46E-11**	112,939	0.687	0.0344	0.0051	**1.54E-11**	0	1.000
rs10876528	12	52,707,743	*HOXC6, HOXC5, HOXC4*	A	C	0.369	0.0377	0.0055	**9.17E-12**	89,151	0.390	0.0377	0.0055	**7.24E-12**	0	0.942
rs2160077	14	91,498,163	*TRIP11*	A	G	0.420	-0.0264	0.0047	**2.62E-08**	116,728	0.396	-0.0264	0.0047	**1.95E-08**	0	same

SNP	Chr.	Position (bp) NCBI build 36	Nearest gene	Alleles (Effect, Other)		Single-marker analysis					Joint analysis					1000G CEU r^2 with lead SNP
						Freq.	β	SE	*P*	*N*	Freq.	β	SE	*P*	LD *r*	
rs7166081	15	65,279,355	*FLJ11506*	A	G	0.771	0.0321	0.0056	**1.14E-08**	113,400	0.753	0.0316	0.0056	**1.66E-08**	-0.014	same
rs893817	15	72,016,118	*LOXL1*	A	G	0.650	-0.0329	0.0050	**3.03E-11**	110,347	0.656	-0.0325	0.0050	**7.89E-11**	0	0.228
rs6602986	15	82,238,103	*ADAMTSL3*	T	C	0.314	-0.0451	0.0051	**4.86E-19**	111,999	0.297	-0.0451	0.0051	**9.57E-19**	0	0.345
rs9890032	17	26,190,060	*ATAD5*	C	G	0.613	0.0272	0.0049	**2.76E-08**	110,300	0.655	0.0272	0.0049	**2.85E-08**	0	0.933
rs882367	17	56,849,356	*C17orf82*	T	C	0.681	-0.0279	0.0051	**3.52E-08**	111,121	0.688	-0.0279	0.0051	**4.50E-08**	0	1.000
rs4239436	18	18,985,928	*CABLES1*	A	G	0.211	-0.0450	0.0058	**1.46E-14**	112,317	0.223	-0.0450	0.0058	**8.73E-15**	0	same
rs979012	20	6,571,374	*BMP2*	T	C	0.342	0.0303	0.0051	**2.36E-09**	107,270	0.350	0.0303	0.0051	**2.85E-09**	0	same

SNP frequencies (Freq.), effect betas (β), standard errors (SE), and *P* values are reported in terms of the effect allele. SNP sample sizes (*N*) are from the sex-combined meta-analysis. Main effect $P < 5 \times 10^{-8}$ are marked in boldface. For each SNP, linkage disequilibrium (LD *r*) with the SNP in next row is reported based on the PIVUS samples, and LD (1000G CEU r^2) with the corresponding lead SNP from **Table 4.2** is reported based on CEU samples from the 1000 Genomes Project (see Methods).

Table 4.3 – Associations of 84 independent waist circumference signals with other anthropometric traits

Color	Significance	Direction of effect
	$P < 5 \times 10^{-8}$	Waist circumference-increasing allele associated with higher levels of other body size trait
	$P < 0.0001$	
	$P < 0.01$	
	$P < 0.05$	
	$P \geq 0.05$	Non-significant association
	$P < 0.05$	Waist circumference-increasing allele associated with lower levels of other body size trait
	$P < 0.01$	
	$P < 0.0001$	
	$P < 5 \times 10^{-8}$	

		Other anthropometric traits			
SNP	Nearest gene	Body mass index	Height	Hip circumference	Waist-hip ratio
rs4239436	*CABLES1*				
rs395962	*LIN28B*				
rs1776897	*NR_001561*				
rs12991495	*DNMT3A*				
rs7513580	*SPAG17*				
rs757608	*C17orf82*				
rs7854560	*PTCH1*				
rs4886782	*LOXL1*				
rs1879529	*ACAN*				
rs1173734	*NPR3*				
rs822531	*NR_001571*				
rs2214442	*ITGB8*				
rs272869	*SLC22A4*				
rs780159	*ZMIZ1*				
rs648514	*SEMA3G*				
rs6556301	*FGFR4*				
rs979012	*BMP2*				
rs984222	*WARS2*				
rs6701231	*WARS2*				
rs849140	*JAZF1*				
rs2274432	*C1orf19*				
rs991967	*TGFB2*				
rs4567683	*ADAMTS17*				

		Other anthropometric traits			
SNP	**Nearest gene**	**Body mass index**	**Height**	**Hip circumference**	**Waist-hip ratio**
rs4542783	*MYO1F*				
rs7621331	*PPP2R3A*				
rs10041657	*FER*				
rs2562779	*ADAMTSL3*				
rs710841	*PRKG2*				
rs2638953	*CCDC91*				
rs4868125	*FBXW11*				
rs13210323	*ANKS1A*				
rs12127195	*C1orf140*				
rs4246302	*ADAMTS17*				
rs9400239	*FOXO3A*				
rs9470001	*PPARD*				
rs806794	*HIST1H2BF*				
rs7162542	*ADAMTSL3*				
rs11205277	*SF3B4*				
rs3791679	*EFEMP1*				
rs798489	*GNA12*				
rs9435732	*MFAP2*				
rs473902	*PTCH1*				
rs12207675	*FILIP1*				
rs1173770	*C5orf23*				
rs7970350	*HMGA2*				
rs12493901	*FNDC3B*				
rs3760318	*CENTA2*				
rs7684221	*LCORL*				
rs6437061	*DIS3L2*				
rs606452	*SERPINH1*				
rs11144688	*PCSK5*				
rs2160077	*TRIP11*				
rs12330322	*RYBP*				
rs1812175	*HHIP*				
rs1344674	*ZBTB38*				
rs7166081	*FLJ11506*				
rs276369	*OR2B6*				
rs3897379	*ZNT8*				
rs459193	*ANKRD55*				
rs2067087	*HOXA13*				
rs12317176	*KIAA2017*				
rs3786897	*PEPD*				
rs2052670	*FLJ16124*				

		Other anthropometric traits			
SNP	**Nearest gene**	**Body mass index**	**Height**	**Hip circumference**	**Waist-hip ratio**
rs2124969	*ITGB6*				
rs17451107	*FLJ16641*				
rs12608504	*JUND*				
rs12679556	*MSC*				
rs10516107	*CPEB4*				
rs2294239	*bK175E3.6*				
rs863750	*ZNF664*				
rs2745353	*RSPO3*				
rs2071449	*HOXC5,HOXC4*				
rs7830933	*NKX3-1*				
rs13173241	*MEKK1*				
rs2047937	*ZNF423*				
rs4886648	***SCAMP5***				
rs17472426	***CCNJL***				
rs3799380	***BTN2A1***				
rs10925060	***OR2W5***				
rs2754603	***MGC22265, GUSBL1***				
rs13196692	***AX747641***				
rs1515114	***IRS1***				
rs998584	*VEGFA,VEGF41*				
rs9378213	*HLA-DRA*				

SNP associations with four anthropometric traits are from sex-combined meta-analysis data (see Methods). Seven SNPs that are $P \geq 0.0001$ with all four traits are marked in boldface.

CHAPTER V

Replication of *LIN28B* SNP association with age of menarche in young Filipino women[1]

OVERVIEW

Age of menarche, or the timing of first menses in girls, is a physiological trait that shows substantial genetic heritability. Earlier age of menarche is associated with increased adult risk of obesity and cardiovascular disease. A recent genome-wide association (GWA) study identified 42 loci associated with age of menarche ($P < 5 \times 10^{-8}$) and observed nominal association for 11 obesity loci ($P < 0.05$). We tested index SNPs at 23 of these loci for association with menarche in young Filipino women from the Cebu Longitudinal Health and Nutrition Survey (CLHNS). We observed nominal association ($P < 0.05$) with consistent direction of effect for the SNP rs7759938 at *LIN28B.* To test whether childhood adiposity, as measured by body mass index (BMI) at age 8, mediates the relationship between SNPs and age of menarche, we performed mediation analyses on 13 SNPs showing consistent effect directions. We observed suggestive evidence that childhood BMI did not mediate the effect of the *LIN28B* SNP on age of menarche. These data confirm the strongest gene reported in Europeans (*LIN28B*) as a contributor to age of menarche in an Asian population.

---

[1] A version of this work will be submitted for publication.

INTRODUCTION

Earlier age of menarche in girls is associated with increased adult risk of obesity and cardiovascular disease (39, 44). While nutrition and other environmental factors play a substantial role in the timing of first menses, age of menarche is a strongly heritable trait, with approximately 50% of its variability explained by genetic factors (49, 170). Forty-two genetic loci have been strongly associated with age of menarche (32 at $P < 5 \times 10^{-8}$ and 10 at $P < 10^{-5}$) in a recent meta-analysis of GWA studies of women of European ancestry (171). Notably, four of these loci (*FTO, SEC16B, ETV5,* and *TMEM18*) have been previously associated with BMI in both adults (59) and children (58, 81, 90) at a genome-wide significance level ($P < 5 \times 10^{-8}$). Eleven obesity loci showed at least nominal association with age of menarche ($P < 0.05$), with the obesity risk alleles associated with earlier menarche (171). These results raise the question of whether obesity genes associated with age of menarche act through biological mechanisms related to increased childhood adiposity. To examine the relationships among obesity and menarche loci, childhood adiposity, and age of menarche in young adult Filipino women from the CLHNS, we tested 23 menarche and obesity SNPs for association with age of menarche, and of those, 13 SNPs showing consistent effect directions were tested for evidence of mediation by BMI at age 8.

Subjects and Methods

During the 1994–1995, 1998–1999, and 2002 surveys, while aged 10 to 11, 14 to 15, and 17 years, respectively, a total of 997 non-twin female offspring from the CLHNS were asked to report their menarcheal status and the month and year of their first menstrual period (47). Of these, 966 also were measured for BMI at age 8 during the 1991 survey, prior to the onset of menarche. Genotyping of the entire CLHNS offspring cohort (1,779 male and female samples) using the MetaboChip array (Illumina, San Diego, CA) was performed at the UNC Mammalian Genotyping Core. Quality control checks were performed on the genotype data at the sample and SNP level as described previously (72). To test the previously reported menarche loci (171), we extracted SNP genotypes for the 15 of 42 SNPs that were directly typed in the CLHNS or had usable proxies ($r^2 > 0.8$, identified using SNAP (153) and data on CHB+JPT samples from the first low-coverage sequencing pilot of the 1000 Genomes Project). Genotypes were also extracted for eight additional obesity SNPs that showed only nominal prior association ($P < 0.05$) with age of menarche (171).

All 23 candidate SNPs were tested for additive association with untransformed age of menarche using PLINK (133) (**Figure 5.1A**). Age of menarche was normally distributed and therefore not natural log-transformed. Association models were adjusted for covariates associated with age of menarche ($P < 0.05$): socioeconomic status measured during the 1991 survey and the first nine principal components of population substructure calculated previously (72). The association models were tested with and without adjustment for childhood BMI. Genotype, age

of menarche, and covariate data was available for 832 women. Significant replication was defined as $P < 0.05$ and direction of effect consistent with the previous report. Quanto v.1.2.4 (available on-line at http://hydra.usc.edu/gxe/) was used to calculate statistical power for replication.

To test whether childhood BMI mediated the effects of these SNPs on age of menarche, we performed mediation analyses using the SAS macro INDIRECT (172). Using linear regressions, this method models how an independent variable (X) affects a dependent variable (Y) through an intervening variable (M). The total effect of X on Y (c) is the sum of X's direct effect (c') and its indirect effects ($a \times b$) through a mediator M (**Figure 5.1B**). We reported the mean value of the indirect effect ($a \times b$) and its bias-corrected non-parametric bootstrap 95% confidence interval (CI) along with the unstandardized regression coefficients (a, b, c, and c'). A 95% CI that did not include zero was considered significant.

Empirical power calculations for each SNP tested in the mediation analyses were performed using published R code (173). Briefly, based on a simulated set of 832 samples derived from the observed regression coefficients (a, b, and c'), 5,000 bootstrap estimates of the mediation effect were generated and used to calculate a 95% CI ($\alpha = 0.05$). Power to detect mediation then was the number of times out of 1,000 iterations that the bootstrapped 95% CI overlapped the actual observed mediation effect ($a \times b$).

RESULTS

We performed two genetic analyses of age of menarche in 832 women from the CLHNS, whose characteristics are summarized in **Table 5.2.** We first examined 15 previously reported menarche SNPs or their proxies for evidence of association in the CLHNS. Two SNPs showed significant associations ($P < 0.05$) adjusted for BMI at age 8, rs7113874 (*TRIM66*) at $P = 0.0086$ and rs7759938 (*LIN28B*) at $P = 0.023$ (**Table 5.2**). Comparable results were observed without adjusting for childhood BMI. In both cases, the *LIN28B* association alone was in the same direction of effect as previously reported, making it the only true replication, explaining 0.5% of phenotypic variance in the CLHNS. The total amount explained previously in Europeans by the index SNPs at both *LIN28B* and *TMEM38B* was 0.6% (174). Our study had ~53% power to replicate the association of a SNP explaining 0.5% of the variability in age of menarche at $\alpha = 0.025$ in a one-sided test and ~35% power to replicate a SNP explaining 0.3% of phenotypic variability at the same threshold. We also examined the associations of an additional eight obesity SNPs that previously showed at least nominal evidence of association with age of menarche ($P < 0.05$) in Europeans with the obesity risk allele corresponding to earlier menarche (171). None were significantly associated with age of menarche in the CLHNS ($P \geq 0.05$) (**Table 5.2**).

Based only on a consistent effect direction and not requiring a significant *P* value, we then selected a subset of 13 SNPs to test for evidence of a mediating effect of a childhood adiposity measure on associations with age of menarche (**Table 5.3**). No SNP association with age of menarche showed significant evidence of mediation

by childhood BMI ($a \times b$). While childhood BMI was always significantly associated with age of menarche (b, $P < 0.05$), none of the 13 SNPs were significantly associated with childhood BMI (a, $P \geq 0.05$). Only rs7759938 (*LIN28B*) showed a significant direct effect (c', $P = 0.019$) and a significant total effect (c, $P = 0.032$). Similar results were observed when the mediation analyses were run excluding covariates (*data not shown*). Given the observed coefficients in **Table 5.3**, our study had 7.2–100% power (mean: 55%, median: 59%) to detect a mediation effect on the association between each of the 13 SNPs and age of menarche at $\alpha = 0.05$. The power for rs7759938 (*LIN28B*) was 64%. Finally, we also tested all 13 SNPs for evidence of direct interactions with childhood BMI to influence age of menarche, but observed no significant interactions ($P < 0.05$) (*data not shown*).

DISCUSSION

These genetic analyses of age of menarche in a cohort of young adult Filipino women provide supporting evidence of association of rs7759938 at *LIN28B*, a GWA locus for menarche identified in subjects of European ancestry. A recent GWA study of age of menarche in 3,468 Hispanic women showed nominal association for five of the 42 SNPs ($P < 0.05$) (175), supporting the observation that some of these signals have detectable effects across populations. Twelve of the 13 SNPs on which we then performed mediation analyses showed evidence of "no-effect non-mediation" (176), meaning that childhood BMI did not significantly mediate the relationship of each SNP with age of menarche and also that each SNP did not have a significant direct effect on age of menarche. Given our limited power, these results do not necessarily

indicate an absence of mediation. The key factor was we could not detect significant associations in the first step of the mediation model (*a*) connecting these SNPs to childhood BMI, which in the second step of the model (*b*) was always associated with age of menarche. Furthermore, a significant total effect of an independent variable on a dependent variable (*c*) is not strictly necessary to observe a significant mediation effect (176). The *LIN28B* SNP, however, showed evidence of "direct-only non-mediation" (176), meaning that while childhood BMI did not significantly mediate the relationship of this SNP with age of menarche, the SNP did have a significant direct effect on age of menarche (*c'*). This result suggests that *LIN28B* does not have an indirect effect on age of menarche through the biological pathways represented by childhood BMI.

The major limitation of this study is its small sample size, which reduces the statistical power available to detect genetic associations with age of menarche or evidence of mediation. Consistent with our expectation, this study replicated the first major menarche signal at *LIN28B*, which explained far greater phenotypic variance than all of the additional signals identified in Europeans (171). Additionally, our mediation models may be missing other important mediators representing relevant biological pathways, such as hormonal biomarkers or alternative measures of childhood adiposity. Nonetheless, our study avoids a key temporal issue in interpreting our mediation results because the measures of childhood adiposity were ascertained years before the average onset of menarche in this cohort.

In summary, these data replicated in an Asian population a *LIN28B* SNP association with age of menarche discovered in Europeans. Attempts to further characterize the connection between obesity genes and age of menarche were not successful. Larger studies of populations with similar pre-pubertal adiposity data will be necessary to more fully address the relationship between the genetics of obesity and of menarche.

ACKNOWLEDGMENTS

I would like to acknowledge Drs. Linda S. Adair, Leslie A. Lange, and Karen L. Mohlke who assisted me in this work. This work was supported by National Institutes of Health grants [DK078150, TW05596, HL085144, K01DK075573, and P30ES10126], pilot funds [RR20649, ES10126, and DK56350], and a training grant [T32 GM007092]. We thank the research and data collection teams from the Office of Population Studies Foundation as well as the study participants who generously provided their time.

DISCLOSURE STATEMENT

The authors declared no conflicts of interest.

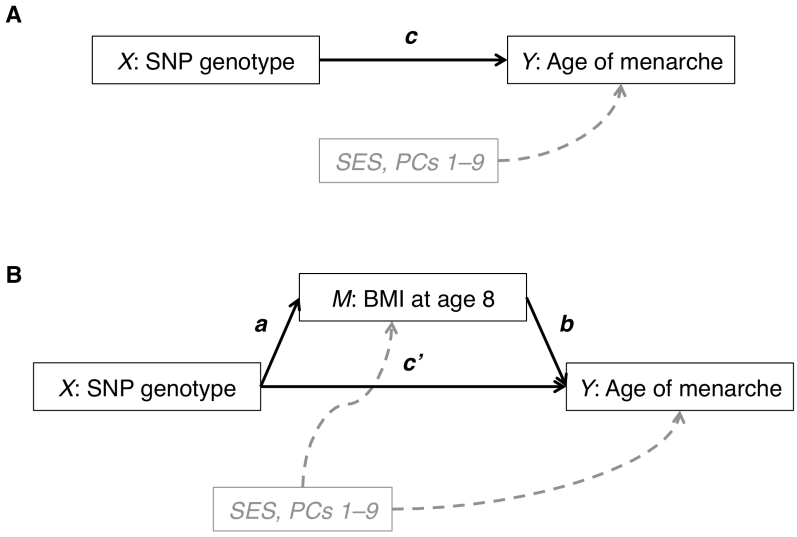


Figure 5.1 – Illustration of two genetic models for SNP-menarche associations in the CLHNS: (*A*) main effect-only model and (*B*) mediation model, both including covariates associated with age of menarche. BMI, body mass index; SES, socioeconomic status; PC, principal component of population substructure.

Table 5.1 – Basic characteristics of the CLHNS cohort

	Summaries		Correlations		
Trait	**Value**	***n***	**Age at menarche**	**BMI**	**SES**
Age at menarche (years)	13.1 ± 1.0	832	-	-0.22	-0.28
BMI at age 8 (kg/m^2)	14.7 ± 1.3	827	< 0.0001	-	0.06
SES at age 8	5 (3, 6)	832	< 0.0001	0.089	-

Values are reported as mean ± standard deviation or as median (25th percentile, 75th percentile). SES is a standardized factor score that ranges from 1.0 to 10.0. Pearson correlations are reported in the upper triangle with their significances in the lower triangle. BMI, body mass index; SES, socioeconomic status.

Table 5.2 – Associations of previously reported menarche and obesity SNPs with age of menarche in the CLHNS

									Adjusted			Unadjusted		
SNP	**Chr.**	**Position**	**Locus Name**	**Trait**	**A1**	**A2**	***n***	**Freq.**	**β**	**SE**	***P***	**β**	**SE**	***P***
rs7113874	11	8,644,592	*TRIM66*	Menarche	C	T	832	0.340	0.138	0.053	**0.0086**	0.143	0.053	**0.0073**
rs7759938	6	105,485,647	*LIN28B*	Menarche	T	C	832	0.610	-0.113	0.049	**0.023**	-0.120	0.050	**0.016**
rs10938397	4	44,877,284	*GNPDA2*	BMI	G	A	830	0.188	-0.099	0.062	0.11	-0.102	0.063	0.11
rs6473015	8	78,341,040	*PXMP3*	Menarche	A	C	832	0.835	0.096	0.064	0.13	0.091	0.065	0.16
rs889398	16	68,114,216	*NFAT5*	Menarche	C	T	832	0.907	-0.105	0.085	0.22	-0.105	0.086	0.22
rs4923461	11	27,613,486	*BDNF*	BMI	A	G	832	0.512	-0.053	0.046	0.25	-0.054	0.047	0.25
rs2090409	9	108,006,909	*TMEM38B*	Menarche	A	C	832	0.310	-0.050	0.052	0.33	-0.058	0.052	0.27
rs2815752	1	72,585,028	*NEGR1*	BMI	A	G	832	0.869	-0.089	0.074	0.23	-0.080	0.075	0.28
rs9385399	6	126,800,726	*C6orf173*	Menarche	T	G	832	0.881	-0.084	0.075	0.26	-0.080	0.076	0.29
rs1659127	16	14,295,806	*MKL2*	Menarche	G	A	832	0.436	0.058	0.049	0.24	0.051	0.049	0.31
rs633715	1	176,119,203	*SEC16B*	Menarche	C	T	832	0.079	-0.048	0.092	0.60	-0.068	0.093	0.47
rs2178403	3	185,522,360	*ECE2*	Menarche	G	A	832	0.665	-0.038	0.051	0.46	-0.037	0.052	0.48
rs987237	6	50,911,009	*TFAP2B*	WHR	G	A	832	0.216	-0.034	0.062	0.59	-0.040	0.063	0.52
rs11084753	19	39,013,977	*KCTD15*	BMI	G	A	831	0.220	0.020	0.059	0.73	0.034	0.059	0.56
rs757647	5	137,735,214	*KDM3B*	Menarche	A	G	830	0.717	0.027	0.055	0.63	0.031	0.056	0.58
rs900145	11	13,250,481	*ARNTL*	Menarche	C	T	830	0.375	-0.012	0.051	0.82	-0.020	0.052	0.69
rs1962448	8	9,909,381	*MSRA*	WHR	G	A	832	0.375	0.028	0.049	0.57	0.018	0.050	0.72
rs7647305	3	187,316,984	*ETV5*	BMI	C	T	832	0.910	-0.006	0.087	0.94	-0.025	0.088	0.78
rs7138803	12	48,533,735	*FAIM2*	BMI	A	G	832	0.150	0.014	0.070	0.84	0.018	0.071	0.80
rs2600959	3	134,098,154	*TMEM108*	Menarche	G	A	831	0.747	-0.016	0.058	0.78	-0.013	0.059	0.83
rs9939609	16	52,378,028	*FTO*	Menarche/BMI	A	T	832	0.182	0.013	0.063	0.83	0.013	0.064	0.84
rs1567890	11	77,779,348	*GAB2*	Menarche	T	C	832	0.658	0.007	0.054	0.89	0.008	0.054	0.89
rs2947411	2	604,168	*TMEM18*	Menarche/BMI	G	A	832	0.917	0.010	0.091	0.91	-0.004	0.092	0.96

SNP positions are reported in hg18/B36.3 coordinates. SNP frequencies (Freq.), effect betas (β), and standard errors (SE) are reported in terms of allele 1 (A1), which is the previously reported menarche-decreasing allele. Associations are reported with and without adjustment for body mass index (BMI) at age 8. Significant effects ($P < 0.05$) are shown in boldface.

Table 5.3 – Mediation effects of childhood adiposity on associations of menarche and obesity SNPs with age of menarche in the CLHNS

						c			***a***			***b***			***c′***			***a* × *b***		
SNP	**Locus**	**Trait**	**A1**	**A2**	***n***	**β**	**SE**	***P***	**β**	**SE**	***P***	**β**	**SE**	***P***	**β**	**SE**	***P***	**β**	**L95**	**U95**
rs7759938	*LIN28B*	Menarche	T	C	827	-0.106	0.049	**0.032**	0.075	0.064	0.24	-0.163	0.027	< 0.0001	-0.118	0.050	**0.019**	-0.012	-0.033	0.009
rs10938397	*GNPDA2*	BMI	G	A	825	-0.113	0.062	0.070	-0.062	0.080	0.44	-0.167	0.027	< 0.0001	-0.102	0.063	0.107	0.010	-0.016	0.038
rs889398	*NFAT5*	Menarche	C	T	827	-0.120	0.085	0.16	-0.107	0.110	0.33	-0.166	0.027	< 0.0001	-0.103	0.087	0.239	0.018	-0.014	0.058
rs4923461	*BDNF*	BMI	A	G	827	-0.054	0.046	0.24	0.005	0.060	0.93	-0.165	0.027	< 0.0001	-0.055	0.047	0.25	-0.001	-0.020	0.020
rs2090409	*TMEM38B*	Menarche	A	C	827	-0.048	0.052	0.36	0.072	0.067	0.29	-0.164	0.027	< 0.0001	-0.059	0.053	0.26	-0.012	-0.034	0.008
rs2815752	*NEGR1*	BMI	A	G	827	-0.094	0.074	0.20	-0.054	0.095	0.57	-0.166	0.027	< 0.0001	-0.085	0.075	0.26	0.009	-0.021	0.048
rs9385399	*C6orf173*	Menarche	T	G	827	-0.099	0.075	0.19	-0.099	0.097	0.31	-0.166	0.027	< 0.0001	-0.082	0.076	0.28	0.016	-0.015	0.054
rs633715	*SEC16B*	Menarche	C	T	827	-0.027	0.092	0.77	0.215	0.119	0.070	-0.165	0.027	< 0.0001	-0.062	0.094	0.51	-0.035	-0.091	0.004
rs2178403	*ECE2*	Menarche	G	A	827	-0.040	0.051	0.44	0.001	0.066	0.99	-0.165	0.027	< 0.0001	-0.040	0.052	0.45	0.000	-0.022	0.021
rs987237	*TFAP2B*	WHR	G	A	827	-0.033	0.062	0.59	0.061	0.080	0.45	-0.165	0.027	< 0.0001	-0.043	0.063	0.49	-0.010	-0.036	0.013
rs900145	*ARNTL*	Menarche	C	T	825	0.006	0.051	0.90	-0.076	0.066	0.25	-0.166	0.027	< 0.0001	0.019	0.052	0.72	0.013	-0.007	0.035
rs7647305	*ETV5*	BMI	C	T	827	-0.015	0.087	0.86	0.087	0.113	0.44	-0.165	0.027	< 0.0001	-0.030	0.089	0.74	-0.014	-0.052	0.021
rs2600959	*TMEM108*	Menarche	G	A	826	-0.008	0.058	0.89	0.027	0.075	0.72	-0.165	0.027	< 0.0001	-0.013	0.060	0.83	-0.005	-0.029	0.023

SNP effect betas (β) and standard errors (SE) are reported in terms of allele 1 (A1), as in Table 3. Significant effects ($P < 0.05$ or a 95% CI that does not include zero) are shown in boldface. The effect coefficients *a*, *b*, *c*, and *c′* (see **Figure 5.1**) are unstandardized. 95% CI, 95% confidence interval; L95, lower bound of 95% CI; U95, upper bound of 95% CI.

CHAPTER VI

Discussion

In the preceding chapters, I described a collection of published and unpublished findings that further our understanding of the genetic architecture of several obesity traits, primarily in an Asian population experiencing an obesogenic transition and secondarily in a set of European populations. I first performed a genome-wide association (GWA) study of body size traits in Filipinos from the Cebu Longitudinal Health and Nutrition Survey (CLHNS) to determine the contribution of common genetic variants to those traits (71). I replicated three well-known body mass index (BMI) loci (*BDNF*, *MC4R*, and *FTO*) and further observed evidence of longitudinal changes in the effects of those genes. Next, I sought to establish the putative causal variant(s) underlying a haplotype at the *ADIPOQ* gene identified in a previous CLHNS GWA study to be strongly associated with lower circulating plasma adiponectin level (72). I identified a population-specific missense variant (R221S) that affected the original measurement of the phenotype and explained the observed GWA signal. Next, to identify novel biology underlying the etiology of central adiposity, I performed a meta-analysis of GWA studies of waist circumference (WC) in European individuals from the Genetic Investigation of Anthropometric Traits (GIANT) Consortium. I observed evidence of novel

associations specific to WC and not other related anthropometric traits, such as BMI, height, and waist-hip ratio. These signals included the genes *NLRP3*, which is part of the obesity-related inflammasome complex, and *IRS1*, which is previously associated with body fat percentage and an adverse metabolic profile. Finally, to explore the relationships between several obesity genes and age of menarche, I performed replication and mediation analyses in female offspring from the CLHNS. I observed that the menarche signal at *LIN28B* previously reported in Europeans replicated in Filipinos. While this association appeared not to be mediated by childhood BMI, the results for several obesity genes were inconclusive.

Several aspects of this work are worth highlighting because they illustrate an important general genetic result, an exploration of unusual phenotypic or genotypic data, or an informative application of a particular analytical method. In Chapter II, the use of non-European samples, in this case Filipino women from the CLHNS, for the genetic study of BMI helped implicate a narrower association region at the *BDNF* gene based on population differences in local linkage disequilibrium (LD). This result emphasizes the value in studying the genetics of common complex traits and diseases across populations of different ancestries. The longitudinal nature of the CLHNS also presented some unique opportunities for genetic analysis. In Chapter II, I obtained a more detailed understanding of the contribution of three genes to BMI by identifying how their effects differed according to a time-varying factor, either the age progression of the individuals studied or an unknown dietary or environmental factor reflective of the obesogenic transition in the Philippines (64, 65). In Chapter V, having pre-pubertal adipose data in female offspring from the

CLHNS made it possible to address the temporal relationship of adiposity and the genetics of menarche, unlike in many larger studies that rely on adult proxies.

In examples like these, the data available hint substantially at the path forward, but in other cases, the researcher must personally motivate the investigation. Most often after establishing a GWA signal, more detailed characterization to find the actual causal variant(s) and determine the functional mechanism(s) is left as an exercise for the indeterminate future. For example, the *SORT1* locus modulating circulating plasma levels of low-density lipoprotein cholesterol was identified in 2008 (177), but identification of the causal non-coding single nucleotide polymorphism (SNP) did not happen until 2010 (178). While for the time being, it has been sufficient to simply make a report, the results of the adiponectin signal characterization in Chapter III speak to the need for careful, close, and timely inspection of GWA signals. This work demonstrated not only how a rare or uncommon SNP can turn out to be the basis for a common GWA signal, but also how in this case that variant can cause a phenotypic artifact that is the basis of the original association signal. Another important follow-up analysis that is becoming more commonplace is the identification of additional independent signals in an association region. Previously, this was difficult to manage either because of a lack of statistical power, or in the context of large-scale meta-analyses with sufficiently increased power, the very challenging logistics of conditional analyses. In my work with the Genetic Investigation of Anthropometric Traits (GIANT) Consortium from Chapter IV, the extraordinarily large sample size of the study and a recently developed "pain-free" conditional analysis method (149) intended for just

such kinds of summary data allowed me to move past these obstacles to identify the first evidence of additional signals at waist circumference loci.

Through a variety of approaches, my work identified examples of both common and uncommon genetic variants explaining phenotypic variability in obesity traits, providing motivation for further work to understand the specific contributions of each allelic frequency bin. Ultimately, a more complete representation of the allelic frequency spectrum in each genetic study population will be obtained through the use of denser genotyping arrays, genetic imputation, and whole-genome sequencing. Such knowledge would increase the likelihood of testing actual causal variants rather than their proxies. In the mean time, the methods and approaches in current use have been fruitful in many ways. For example, the adiponectin story from Chapter III highlights the value of small-scale targeted gene sequencing to enrich for causal variants. Additionally, the custom genotyping array known as the MetaboChip, which played a part in the three previous chapters, showed its value as a large-scale targeted follow up of thousands of SNPs with prior evidence of association with anthropometric and cardio-metabolic diseases and traits. Previously, often only a couple dozen top signals from an association meta-analysis could affordably be tested for replication in additional samples. Similar efforts targeting other trait clusters or classes of variants have been developed (e.g., the Immunochip (179), the HumanCVD BeadChip (180), exome chips, etc.) or are in development. Finally, the 1000 Genomes Project (1000G) is well under way and aims to ascertain 95% of the variants with a frequency > 1% in several world populations to serve as a comprehensive reference

resource (54). The initial pilot project set out to evaluate three methods for determining the allelic spectrum: (1) deep-coverage sequencing (42× read-depth) in a three parents-offspring trio, (2) shallow-coverage sequencing (2–6× read-depth) in 179 unrelated individuals, and (3) targeted deep-coverage exon sequencing in 697 individuals. The researchers concluded that exon-sequencing and shallow-coverage whole-genome sequencing represent the most cost-effective approaches, bearing in mind their respective limitations of smaller genomic coverage and higher error rate. This dataset will replace that of the International HapMap Project (53, 147) as the preferred reference set of variants to be imputed in genetic analysis studies using genotyping arrays with less coverage of the genome. Low-frequency variants, however, still represent an ongoing challenge for imputation because they require very large reference panels to be most accurately imputed (181). Until the costs of more comprehensive solutions for representing allelic variation, such as whole-genome sequencing, go down substantially for individual studies, the aforementioned efforts represent important intermediate approaches.

A topic of great debate in the field of human genetics is the nature of the "missing heritability" not yet explained by GWA studies of complex diseases and traits and how to determine the yet unknown causal variants (51, 182). One possible explanation for our incomplete understanding of the genetic underpinnings of phenotypic variability is that common variation, which has been the primary focus of GWA studies, is only a small subset of the actual causal variants. Rare and uncommon SNPs have become more of a focus for study recently and as the standard single-marker association approaches for common variants are severely

underpowered for rare variants, several burden tests have been developed to allow for association testing with quantitative traits (183). In these tests, rare variants are collapsed together by shared characteristics, such as allele frequency or predicted functional effect, in order to test their relative enrichment in disease cases versus controls with greater statistical power. While each of these methods works reasonably well for the particular hypothesis it is testing, none performs well as a global test, suggesting that careful hypotheses must be established prior to association testing. Larger structural variation, such as microsatellite and copy number variation (CNV), has also been implicated as important to human disease etiology (184). Notably, a GWA study of BMI identified a common SNP that tagged a candidate causal CNV at the *NEGR1* locus, a 45 kb deletion overlapping several conserved elements just upstream of the gene (58). Ultimately, explaining all of the heritability for a given trait or disease matters most when the goal is to generate predictive risk scores, but when the goal is to search out interesting new biology, the era of GWA studies has been tremendously fruitful in confirming known loci and suggesting novel loci (185). A more systematically ascertained set of candidate loci for a wide variety of complex diseases and traits now exists for more detailed genetic and molecular characterization. When more true causal variants are identified, the value of predictive risk scores will increase.

The completion of this dissertation work prompts some musings about possible future genetic analyses in the CLHNS. In 2005, when blood was first drawn from CLHNS participants for the purposes of measuring biomarkers and extracting DNA to genotype genetic markers, the era of GWA studies was in its infancy. Seven

years later, the field of human genetics continues to race forward. Large-scale meta-analyses, such as the one described in Chapter IV, will continue to play an important role in the study of complex trait genetics. The wealth of phenotypes and genotypes currently available in the CLHNS will continue to allow it to contribute to such efforts, as was done with the Global Lipids Genetics Consortium (70). Age of menarche in the female offspring (47) represents one such yet untapped resource. Future CLHNS field surveys will also undoubtedly continue to be sources of additional socio-economic, dietary, anthropometric, and biomarker data. The published longitudinal genetic analyses of BMI in the mothers (71) could be extended with such new data or some of the newest BMI association loci could be evaluated in the same way to determine if other major signals have different effects over time. Though it would be wildly implausible for financial and logistical reasons, obtaining computed tomography (CT) measurements of adipose tissue depots in CLHNS participants would allow us to examine genetic contributions to excess adiposity in a very direct way, as in a recent GWA study (186). Obtaining further genetic data would also be of great value, whether it be through improved and expanded SNP imputation based on new 1000G reference haplotypes or through work already underway to directly genotype the exome variants in the CLHNS mothers. The latter dataset will enrich for rare and common coding variants ascertained from large exome sequencing studies and help us to study how such low frequency variants affect phenotypes of interest in both the CLHNS alone and in conjunction with other similarly typed cohorts.

More broadly, the future of the genetic study of complex diseases and traits is shaping up to be a hybrid of targeted candidate gene studies based on prior biological knowledge and more agnostic genome-wide discovery analyses: layering various types of genome-wide expert data upon an increasingly comprehensive set of genetic variants to prioritize candidates for association and functional testing. New computational methods and models integrating genomic, epigenomic, and sequencing data will help motivate "wet lab" experimentation by generating testable hypotheses of molecular and biological function grounded in broad-based data. For instance, multi-layered histone modification data from a broad set of cell types can be used to predict regulatory regions with different effects on gene expression (187). Information on genomic areas of open chromatin also hints at the regulatory mechanisms of gene expression in disease-relevant tissues, such as pancreatic islets (188) and adipocytes (189). Connecting chromatin accessibility to predicted transcription factor binding sites and genetic variation could help establish further mechanistic links to gene expression that initiates downstream biological changes (190, 191). Given the tremendous volume of hypotheses that could potentially be generated, high-throughput approaches to functionally test such regulatory elements will also be essential (192). In short, the expert intersection of biology, bioinformatics, and biostatistics will be a major force in the future of human genetics.

In conclusion, as the obesity epidemic continues to grow as a major world health issue, understanding its genetic component is quite important—even as largely preventable environmental and behavioral factors, such as poor nutrition and a lack of exercise, remain its clearest causes. Together, the results of multiple genetic association studies primarily in Filipinos from the CLHNS and secondarily in Europeans from the GIANT Consortium further our understanding of the genetic architecture of several obesity traits, and suggest many promising biological targets for further genetic and molecular study.

REFERENCES

1. World Health Organization. Fact sheet on obesity and overweight. 2011.
2. Finkelstein EA, Trogdon JG, Cohen JW, Dietz W. Annual medical spending attributable to obesity: payer-and service-specific estimates. Health Aff (Millwood). 2009 Sep-Oct;28(5):w822-31.
3. Popkin BM, Adair LS, Ng SW. Global nutrition transition and the pandemic of obesity in developing countries. Nutr Rev. 2012 Jan;70(1):3-21.
4. Grundy SM. Obesity, metabolic syndrome, and cardiovascular disease. J Clin Endocrinol Metab. 2004 Jun;89(6):2595-600.
5. World Health Organization. Fact sheet on cardiovascular diseases (CVDs) 2011.
6. Sobal J, Stunkard AJ. Socioeconomic status and obesity: a review of the literature. Psychol Bull. 1989 Mar;105(2):260-75.
7. Monteiro CA, Moura EC, Conde WL, Popkin BM. Socioeconomic status and obesity in adult populations of developing countries: a review. Bull World Health Organ. 2004 Dec;82(12):940-6.
8. Herrera BM, Lindgren CM. The genetics of obesity. Curr Diab Rep. 2010 Dec;10(6):498-505.

9. Tran TT, Kahn CR. Transplantation of adipose tissue and stem cells: role in metabolism and disease. Nat Rev Endocrinol. 2010 Apr;6(4):195-213.

10. Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu CY, et al. Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham Heart Study. Circulation. 2007 Jul 3;116(1):39-48.

11. Mahabadi AA, Massaro JM, Rosito GA, Levy D, Murabito JM, Wolf PA, et al. Association of pericardial fat, intrathoracic fat, and visceral abdominal fat with cardiovascular disease burden: the Framingham Heart Study. Eur Heart J. 2009 Apr;30(7):850-6.

12. Smith SR, Lovejoy JC, Greenway F, Ryan D, deJonge L, de la Bretonne J, et al. Contributions of total body fat, abdominal subcutaneous adipose tissue compartments, and visceral adipose tissue to the metabolic complications of obesity. Metabolism. 2001 Apr;50(4):425-35.

13. Loomba-Albrecht LA, Styne DM. Effect of puberty on body composition. Curr Opin Endocrinol Diabetes Obes. 2009 Feb;16(1):10-5.

14. Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. Changes in diet and lifestyle and long-term weight gain in women and men. N Engl J Med. 2011 Jun 23;364(25):2392-404.

15. Ogden CL, Carroll MD, Curtin LR, McDowell MA, Tabak CJ, Flegal KM. Prevalence of overweight and obesity in the United States, 1999-2004. JAMA. 2006 Apr 5;295(13):1549-55.

16. Toth MJ, Tchernof A, Sites CK, Poehlman ET. Effect of menopausal status on body composition and abdominal fat distribution. Int J Obes Relat Metab Disord. 2000 Feb;24(2):226-31.

17. El-Sayed AM, Scarborough P, Galea S. Ethnic inequalities in obesity among children and adults in the UK: a systematic review of the literature. Obes Rev. 2011 May;12(5):e516-34.

18. Gradmark AM, Rydh A, Renstrom F, De Lucia-Rolfe E, Sleigh A, Nordstrom P, et al. Computed tomography-based validation of abdominal adiposity measurements from ultrasonography, dual-energy X-ray absorptiometry and anthropometry. Br J Nutr. 2010 Aug;104(4):582-8.

19. Shuster A, Patlas M, Pinthus JH, Mourtzakis M. The clinical importance of visceral adiposity: a critical review of methods for visceral adipose tissue analysis. Br J Radiol. 2012 Jan;85(1009):1-10.

20. World Health Organization. The Asia-Pacific Perspective: Redefining Obesity and Its Treatment. 2000.

21. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004 Jan 10;363(9403):157-63.

22. Shen W, Punyanitya M, Chen J, Gallagher D, Albu J, Pi-Sunyer X, et al. Waist circumference correlates with metabolic syndrome indicators better than percentage fat. Obesity (Silver Spring). 2006 Apr;14(4):727-36.

23. Lear SA, Humphries KH, Kohli S, Birmingham CL. The use of BMI and waist circumference as surrogates of body fat differs by ethnicity. Obesity (Silver Spring). 2007 Nov;15(11):2817-24.

24. Ross R, Berentzen T, Bradshaw AJ, Janssen I, Kahn HS, Katzmarzyk PT, et al. Does the relationship between waist circumference, morbidity and mortality depend on measurement protocol for waist circumference? Obes Rev. 2008 Jul;9(4):312-25.

25. Snijder MB, Dekker JM, Visser M, Bouter LM, Stehouwer CD, Kostense PJ, et al. Associations of hip and thigh circumferences independent of waist circumference with the incidence of type 2 diabetes: the Hoorn Study. Am J Clin Nutr. 2003 May;77(5):1192-7.

26. Yim JE, Heshka S, Albu JB, Heymsfield S, Gallagher D. Femoral-gluteal subcutaneous and intermuscular adipose tissues have independent and opposing relationships with CVD risk. J Appl Physiol. 2008 Mar;104(3):700-7.

27. de Koning L, Merchant AT, Pogue J, Anand SS. Waist circumference and waist-to-hip ratio as predictors of cardiovascular events: meta-regression analysis of prospective studies. Eur Heart J. 2007 Apr;28(7):850-6.

28. Czernichow S, Kengne AP, Stamatakis E, Hamer M, Batty GD. Body mass index, waist circumference and waist-hip ratio: which is the better discriminator of cardiovascular disease mortality risk?: evidence from an individual-participant meta-analysis of 82 864 participants from nine cohort studies. Obes Rev. 2011 Sep;12(9):680-7.

29. American Heart Association. Heart and stroke facts. Dallas, TX, USA1992.

30. Mathieu P, Lemieux I, Despres JP. Obesity, inflammation, and cardiovascular risk. Clin Pharmacol Ther. 2010 Apr;87(4):407-16.

31. Diez JJ, Iglesias P. The role of the novel adipocyte-derived hormone adiponectin in human disease. Eur J Endocrinol. 2003 Mar;148(3):293-300.

32. Hopkins TA, Ouchi N, Shibata R, Walsh K. Adiponectin actions in the cardiovascular system. Cardiovasc Res. 2007 Apr 1;74(1):11-8.

33. Tsilidis KK, Helzlsouer KJ, Smith MW, Grinberg V, Hoffman-Bolton J, Clipp SL, et al. Association of common polymorphisms in IL10, and in other genes related to inflammatory response and obesity with colorectal cancer. Cancer Causes Control. 2009 Nov;20(9):1739-51.

34. Gable DR, Hurel SJ, Humphries SE. Adiponectin and its gene variants as risk factors for insulin resistance, the metabolic syndrome and cardiovascular disease. Atherosclerosis. 2006 Oct;188(2):231-44.

35. Renaldi O, Pramono B, Sinorita H, Purnomo LB, Asdie RH, Asdie AH. Hypoadiponectinemia: a risk factor for metabolic syndrome. Acta Med Indones. 2009 Jan;41(1):20-4.

36. Comuzzie AG, Funahashi T, Sonnenberg G, Martin LJ, Jacob HJ, Black AE, et al. The genetic basis of plasma variation in adiponectin, a global endophenotype for obesity and the metabolic syndrome. J Clin Endocrinol Metab. 2001 Sep;86(9):4321-5.

37. Chuang LM, Chiu YF, Sheu WH, Hung YJ, Ho LT, Grove J, et al. Biethnic comparisons of autosomal genomic scan for loci linked to plasma adiponectin in populations of Chinese and Japanese origin. J Clin Endocrinol Metab. 2004 Nov;89(11):5772-8.

38. Lindsay RS, Funahashi T, Krakoff J, Matsuzawa Y, Tanaka S, Kobes S, et al. Genome-wide linkage analysis of serum adiponectin in the Pima Indian population. Diabetes. 2003 Sep;52(9):2419-25.

39. Freedman DS, Khan LK, Serdula MK, Dietz WH, Srinivasan SR, Berenson GS. The relation of menarcheal age to obesity in childhood and adulthood: the Bogalusa heart study. BMC Pediatr. 2003 Apr 30;3:3.

40. Feng Y, Hong X, Wilker E, Li Z, Zhang W, Jin D, et al. Effects of age at menarche, reproductive years, and menopause on metabolic risk factors for cardiovascular diseases. Atherosclerosis. 2008 Feb;196(2):590-7.

41. He C, Zhang C, Hunter DJ, Hankinson SE, Buck Louis GM, Hediger ML, et al. Age at menarche and risk of type 2 diabetes: results from 2 large prospective cohort studies. Am J Epidemiol. 2010 Feb 1;171(3):334-44.

42. Stöckl D, Doring A, Peters A, Thorand B, Heier M, Huth C, et al. Age at menarche is associated with prediabetes and diabetes in women (aged 32-81 years) from the general population: the KORA F4 Study. Diabetologia. 2011 Dec 15.

43. Stöckl D, Meisinger C, Peters A, Thorand B, Huth C, Heier M, et al. Age at menarche and its association with the metabolic syndrome and its components: results from the KORA F4 study. PLoS ONE. 2011;6(10):e26076.

44. Lakshman R, Forouhi NG, Sharp SJ, Luben R, Bingham SA, Khaw KT, et al. Early age at menarche associated with cardiovascular disease and mortality. J Clin Endocrinol Metab. 2009 Dec;94(12):4953-60.

45. Lee JM, Appugliese D, Kaciroti N, Corwyn RF, Bradley RH, Lumeng JC. Weight status in young girls and the onset of puberty. Pediatrics. 2007 Mar;119(3):e624-30.

46. Lassek WD, Gaulin SJ. Brief communication: menarche is related to fat distribution. Am J Phys Anthropol. 2007 Aug;133(4):1147-51.

47. Adair LS. Size at birth predicts age at menarche. Pediatrics. 2001 Apr;107(4):E59.

48. Freedman DS, Khan LK, Serdula MK, Dietz WH, Srinivasan SR, Berenson GS. Relation of age at menarche to race, time period, and anthropometric dimensions: the Bogalusa Heart Study. Pediatrics. 2002 Oct;110(4):e43.

49. Towne B, Czerwinski SA, Demerath EW, Blangero J, Roche AF, Siervogel RM. Heritability of age at menarche in girls from the Fels Longitudinal Study. Am J Phys Anthropol. 2005 Sep;128(1):210-9.

50. Ku CS, Loy EY, Pawitan Y, Chia KS. The pursuit of genome-wide association studies: where are we now? J Hum Genet. 2010 Apr;55(4):195-206.

51. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009 Oct 8;461(7265):747-53.

52. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al. Complement factor H polymorphism in age-related macular degeneration. Science. 2005 Apr 15;308(5720):385-9.

53. International HapMap Consortium. A haplotype map of the human genome. Nature. 2005 Oct 27;437(7063):1299-320.

54. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature. 2010 Oct 28;467(7319):1061-73.

55. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A. 2009 Jun 9;106(23):9362-7.

56. Lindgren CM, Heid IM, Randall JC, Lamina C, Steinthorsdottir V, Qi L, et al. Genome-wide association scan meta-analysis identifies three loci influencing adiposity and fat distribution. PLoS Genet. 2009 Jun;5(6):e1000508.

57. Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, Steinthorsdottir V, et al. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nat Genet. 2010 Nov;42(11):949-60.

58. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009 Jan;41(1):25-34.

59. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010 Nov;42(11):937-48.

60. Allen HL, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature. 2010 Oct 14;467(7317):832-8.

61. The International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007 Oct 18;449(7164):851-61.

62. Sanna S, Jackson AU, Nagaraja R, Willer CJ, Chen WM, Bonnycastle LL, et al. Common variants in the GDF5-UQCC region are associated with variation in human height. Nat Genet. 2008 Feb;40(2):198-203.

63. Adair LS, Popkin BM, Akin JS, Guilkey DK, Gultiano S, Borja J, et al. Cohort Profile: The Cebu Longitudinal Health and Nutrition Survey. Int J Epidemiol. 2010 May 27.

64. Lee CMY, Martiniuk ALC, Woodward M, Feigin V, Gu D, Jamrozik K, et al. The burden of overweight and obesity in the Asia-Pacific region. Obes Rev. 2007 May;8(3):191-6.

65. Adair LS. Dramatic rise in overweight and obesity in adult Filipino women and risk of hypertension. Obes Res. 2004 Aug;12(8):1335-41.

66. Marvelle AF, Lange LA, Qin L, Adair LS, Mohlke KL. Association of FTO With Obesity-Related Traits in the Cebu Longitudinal Health and Nutrition Survey (CLHNS) Cohort. Diabetes. 2008 July 1;57(7):1987-91.

67. Lange LA, Croteau-Chonka DC, Marvelle AF, Qin L, Gaulton KJ, Kuzawa CW, et al. Genome-wide association study of homocysteine levels in Filipinos provides evidence for CPS1 in women and a stronger MTHFR effect in young adults. Hum Mol Genet. 2010 Mar 2;19(10):2050-8.

68. Wu Y, McDade TW, Kuzawa CW, Borja J, Li Y, Adair LS, et al. Genome-wide Association with C-Reactive Protein Levels in CLHNS: Evidence for the CRP and HNF1A Loci and their Interaction with Exposure to a Pathogenic Environment. Inflammation. 2012 Apr;35(2):574-83.

69. Wu Y, Li Y, Lange EM, Croteau-Chonka DC, Kuzawa CW, McDade TW, et al. Genome-wide association study for adiponectin levels in Filipino women identifies CDH13 and a novel uncommon haplotype at KNG1-ADIPOQ. Hum Mol Genet. 2010 Oct 11.

70. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010 Aug 5;466(7307):707-13.

71. Croteau-Chonka DC, Marvelle AF, Lange EM, Lee NR, Adair LS, Lange LA, et al. Genome-Wide Association Study of Anthropometric Traits and Evidence of Interactions With Age and Study Year in Filipino Women. Obesity (Silver Spring). 2010 Oct 21.

72. Croteau-Chonka DC, Wu Y, Li Y, Fogarty MP, Lange LA, Kuzawa CW, et al. Population-specific coding variant underlies genome-wide association with adiponectin level. Hum Mol Genet. 2012 Jan 15;21(2):463-71.

73. Brown WV, Fujioka K, Wilson PW, Woodworth KA. Obesity: why be concerned? Am J Med. 2009 Apr;122(4 Suppl 1):S4-11.

74. Popkin BM. Understanding global nutrition dynamics as a step towards controlling cancer incidence. Nat Rev Cancer. 2007 Jan;7(1):61-7.

75. Lee CM, Barzi F, Woodward M, Batty GD, Giles GG, Wong JW, et al. Adult height and the risks of cardiovascular disease and major causes of death in the Asia-Pacific region: 21,000 deaths in 510,000 men and women. Int J Epidemiol. 2009 Aug;38(4):1060-71.

76. Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, et al. Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet. 2008 May;40(5):575-83.

77. Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, et al. Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet. 2008 May;40(5):584-91.

78. Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, et al. Many sequence variants affecting diversity of adult human height. Nat Genet. 2008 May;40(5):609-15.

79. Chambers JC, Elliott P, Zabaneh D, Zhang W, Li Y, Froguel P, et al. Common genetic variation near MC4R is associated with waist circumference and insulin resistance. Nat Genet. 2008 Jun;40(6):716-8.

80. Loos RJ, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, et al. Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nat Genet. 2008 Jun;40(6):768-75.

81. Thorleifsson G, Walters GB, Gudbjartsson DF, Steinthorsdottir V, Sulem P, Helgadottir A, et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nat Genet. 2009 Jan;41(1):18-24.

82. Meyre D, Delplanque J, Chevre JC, Lecoeur C, Lobbens S, Gallina S, et al. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. Nat Genet. 2009 Feb;41(2):157-9.

83. Cho YS, Go MJ, Kim YJ, Heo JY, Oh JH, Ban HJ, et al. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nat Genet. 2009 Apr 26;41(5):527-34.

84. Johansson A, Marroni F, Hayward C, Franklin CS, Kirichenko AV, Jonasson I, et al. Common variants in the JAZF1 gene associated with height identified by linkage and genome-wide association analysis. Hum Mol Genet. 2009 Jan 15;18(2):373-80.

85. Soranzo N, Rivadeneira F, Chinappen-Horsley U, Malkina I, Richards JB, Hammond N, et al. Meta-analysis of genome-wide scans for human adult stature identifies novel loci and associations with measures of skeletal frame size. PLoS Genet. 2009 Apr;5(4):e1000445.

86. Heard-Costa NL, Zillikens MC, Monda KL, Johansson A, Harris TB, Fu M, et al. NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. PLoS Genet. 2009 Jun;5(6):e1000539.

87. Estrada K, Krawczak M, Schreiber S, van Duijn K, Stolk L, van Meurs JB, et al. A genome-wide association study of northwestern Europeans involves the CNP signaling pathway in the etiology of human height variation. Hum Mol Genet. 2009 Jul 1;18(18):3516-24.

88. Tönjes A, Koriath M, Schleinitz D, Dietrich K, Bottcher Y, Rayner NW, et al. Genetic Variation in GPR133 is Associated with Height - Genome Wide Association Study in the Self-contained Population of Sorbs. Hum Mol Genet. 2009 Sep 3;18(23):4662-8.

89. Herbert A, Gerry NP, McQueen MB, Heid IM, Pfeufer A, Illig T, et al. A common genetic variant is associated with adult and childhood obesity. Science. 2006 Apr 14;312(5771):279-83.

90. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007 May 11;316(5826):889-94.

91. Okada Y, Kamatani Y, Takahashi A, Matsuda K, Hosono N, Ohmiya H, et al. A genome-wide association study in 19 633 Japanese subjects identified LHX3-QSOX2 and IGF1 as adult height loci. Hum Mol Genet. 2010 Jun 1;19(11):2303-12.

92. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol. 2010 Nov 5;34(8):816-34.

93. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet. 2006 Dec;2(12):e190.

94. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006 Aug;38(8):904-9.

95. Lau W, Kuo TY, Tapper W, Cox S, Collins A. Exploiting large scale computing to construct high resolution linkage disequilibrium maps of the human genome. Bioinformatics. 2007 Feb 15;23(4):517-9.

96. Dahly DL, Adair LS. Quantifying the urban environment: a scale measure of urbanicity outperforms the urban-rural dichotomy. Soc Sci Med. 2007 Apr;64(7):1407-19.

97. Grunnet M, Jespersen T, Rasmussen HB, Ljungstrom T, Jorgensen NK, Olesen SP, et al. KCNE4 is an inhibitory subunit to the KCNQ1 channel. J Physiol. 2002 Jul 1;542(Pt 1):119-30.

98. Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, Andersen G, et al. SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. Nat Genet. 2008 Sep;40(9):1098-102.

99. Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H, et al. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. Nat Genet. 2008 Sep;40(9):1092-7.

100. Liu Y, Zhou DZ, Zhang D, Chen Z, Zhao T, Zhang Z, et al. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes in the population of mainland China. Diabetologia. 2009 Jul;52(7):1315-21.

101. Beckers S, Peeters A, Zegers D, Mertens I, Van Gaal L, Van Hul W. Association of the BDNF Val66Met variation with obesity in women. Mol Genet Metab. 2008 Sep-Oct;95(1-2):110-2.

102. Lyons WE, Mamounas LA, Ricaurte GA, Coppola V, Reid SW, Bora SH, et al. Brain-derived neurotrophic factor-deficient mice develop aggressiveness and hyperphagia in conjunction with brain serotonergic abnormalities. Proc Natl Acad Sci U S A. 1999 Dec 21;96(26):15239-44.

103. Kernie SG, Liebl DJ, Parada LF. BDNF regulates eating behavior and locomotor activity in mice. EMBO J. 2000 Mar 15;19(6):1290-300.

104. Rios M, Fan G, Fekete C, Kelly J, Bates B, Kuehn R, et al. Conditional deletion of brain-derived neurotrophic factor in the postnatal brain leads to obesity and hyperactivity. Mol Endocrinol. 2001 Oct;15(10):1748-57.

105. Hardy R, Wills AK, Wong A, Elks CE, Wareham NJ, Loos RJ, et al. Life course variations in the associations between FTO and MC4R gene variants and body size. Hum Mol Genet. 2010 Feb 1;19(3):545-52.

106. Weyer C, Funahashi T, Tanaka S, Hotta K, Matsuzawa Y, Pratley RE, et al. Hypoadiponectinemia in obesity and type 2 diabetes: close association with insulin resistance and hyperinsulinemia. J Clin Endocrinol Metab. 2001 May;86(5):1930-5.

107. Hotta K, Funahashi T, Arita Y, Takahashi M, Matsuda M, Okamoto Y, et al. Plasma concentrations of a novel, adipose-specific protein, adiponectin, in type 2 diabetic patients. Arterioscler Thromb Vasc Biol. 2000 Jun;20(6):1595-9.

108. Takahashi M, Arita Y, Yamagata K, Matsukawa Y, Okutomi K, Horie M, et al. Genomic structure and mutations in adipose-specific gene, adiponectin. Int J Obes Relat Metab Disord. 2000 Jul;24(7):861-8.

109. Vasseur F, Helbecque N, Dina C, Lobbens S, Delannoy V, Gaget S, et al. Single-nucleotide polymorphism haplotypes in the both proximal promoter and exon 3 of the APM1 gene modulate adipocyte-secreted adiponectin hormone levels and contribute to the genetic risk for type 2 diabetes in French Caucasians. Hum Mol Genet. 2002 Oct 1;11(21):2607-14.

110. Sutton BS, Weinert S, Langefeld CD, Williams AH, Campbell JK, Saad MF, et al. Genetic analysis of adiponectin and obesity in Hispanic families: the IRAS Family Study. Hum Genet. 2005 Jul;117(2-3):107-18.

111. Tanimura D, Shibata R, Izawa H, Hirashiki A, Asano H, Murase Y, et al. Relation of a common variant of the adiponectin gene to serum adiponectin concentration and metabolic traits in an aged Japanese population. Eur J Hum Genet. 2010 Dec 8.

112. Heid IM, Henneman P, Hicks A, Coassin S, Winkler T, Aulchenko YS, et al. Clear detection of ADIPOQ locus as the major gene for plasma adiponectin: results of genome-wide association analyses including 4659 European individuals. Atherosclerosis. 2010 Feb;208(2):412-20.

113. Jee SH, Sull JW, Lee JE, Shin C, Park J, Kimm H, et al. Adiponectin concentrations: a genome-wide association study. Am J Hum Genet. 2010 Oct 8;87(4):545-52.

114. Richards JB, Waterworth D, O'Rahilly S, Hivert MF, Loos RJ, Perry JR, et al. A genome-wide association study reveals variants in ARL15 that influence adiponectin levels. PLoS Genet. 2009 Dec;5(12):e1000768.

115. Ling H, Waterworth DM, Stirnadel HA, Pollin TI, Barter PJ, Kesaniemi YA, et al. Genome-wide linkage and association analyses to identify genes influencing adiponectin levels: the GEMS Study. Obesity (Silver Spring). 2009 Apr;17(4):737-44.

116. Waki H, Yamauchi T, Kamon J, Ito Y, Uchida S, Kita S, et al. Impaired multimerization of human adiponectin mutants associated with diabetes. Molecular structure and multimer formation of adiponectin. J Biol Chem. 2003 Oct 10;278(41):40352-63.

117. Laumen H, Saningong AD, Heid IM, Hess J, Herder C, Claussnitzer M, et al. Functional characterization of promoter variants of the adiponectin gene complemented by epidemiological data. Diabetes. 2009 Apr;58(4):984-91.

118. Kondo H, Shimomura I, Matsukawa Y, Kumada M, Takahashi M, Matsuda M, et al. Association of adiponectin mutation with type 2 diabetes: a candidate gene for the insulin resistance syndrome. Diabetes. 2002 Jul;51(7):2325-8.

119. Bowden DW, An SS, Palmer ND, Brown WM, Norris JM, Haffner SM, et al. Molecular basis of a linkage peak: exome sequencing and family-based analysis identify a rare genetic variant in the ADIPOQ gene in the IRAS Family Study. Hum Mol Genet. 2010 Oct 15;19(20):4112-20.

120. Hara K, Boutin P, Mori Y, Tobe K, Dina C, Yasuda K, et al. Genetic variation in the gene encoding adiponectin is associated with an increased risk of type 2 diabetes in the Japanese population. Diabetes. 2002 Feb;51(2):536-40.

121. Schaffler A, Barth N, Palitzsch KD, Drobnik W, Scholmerich J, Schmitz G. Mutation analysis of the human adipocyte-specific apM-1 gene. Eur J Clin Invest. 2000 Oct;30(10):879-87.

122. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol. 2010 Jan;8(1):e1000294.

123. Jang Y, Lee JH, Kim OY, Koh SJ, Chae JS, Woo JH, et al. The SNP276G>T polymorphism in the adiponectin (ACDC) gene is more strongly associated with insulin resistance and cardiovascular disease risk than SNP45T>G in nonobese/nondiabetic Korean men independent of abdominal adiposity and circulating plasma adiponectin. Metabolism. 2006 Jan;55(1):59-66.

124. Jang Y, Chae JS, Koh SJ, Hyun YJ, Kim JY, Jeong YJ, et al. The influence of the adiponectin gene on adiponectin concentrations and parameters of metabolic syndrome in non-diabetic Korean women. Clin Chim Acta. 2008 May;391(1-2):85-90.

125. Vendramini MF, Pereira AC, Ferreira SR, Kasamatsu TS, Moises RS. Association of genetic variants in the adiponectin encoding gene (ADIPOQ) with type 2 diabetes in Japanese Brazilians. J Diabetes Complications. 2010 Mar-Apr;24(2):115-20.

126. Register TC, Burdon KP, Lenchik L, Bowden DW, Hawkins GA, Nicklas BJ, et al. Variability of serum soluble intercellular adhesion molecule-1 measurements attributable to a common polymorphism. Clin Chem. 2004 Nov;50(11):2185-7.

127. Newton-Cheh C, Larson MG, Vasan RS, Levy D, Bloch KD, Surti A, et al. Association of common variants in NPPA and NPPB with circulating natriuretic peptides and blood pressure. Nat Genet. 2009 Mar;41(3):348-53.

128. Wallis R, Cheng JY. Molecular defects in variant forms of mannose-binding protein associated with immunodeficiency. J Immunol. 1999 Nov 1;163(9):4953-9.

129. Zhang D, Ma J, Brismar K, Efendic S, Gu HF. A single nucleotide polymorphism alters the sequence of SP1 binding site in the adiponectin promoter region and is associated with diabetic nephropathy among type 1 diabetic patients in the Genetics of Kidneys in Diabetes Study. J Diabetes Complications. 2009 Jul-Aug;23(4):265-72.

130. Kita A, Yamasaki H, Kuwahara H, Moriuchi A, Fukushima K, Kobayashi M, et al. Identification of the promoter region required for human adiponectin gene transcription: Association with CCAAT/enhancer binding protein-beta and tumor necrosis factor-alpha. Biochem Biophys Res Commun. 2005 Jun 3;331(2):484-90.

131. Qiao L, Maclean PS, Schaack J, Orlicky DJ, Darimont C, Pagliassotti M, et al. C/EBPalpha regulates human adiponectin gene transcription through an intronic enhancer. Diabetes. 2005 Jun;54(6):1744-54.

132. Segawa K, Matsuda M, Fukuhara A, Morita K, Okuno Y, Komuro R, et al. Identification of a novel distal enhancer in human adiponectin gene. J Endocrinol. 2009 Jan;200(1):107-16.

133. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559-75.

134. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010 Apr;7(4):248-9.

135. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003 Jul 1;31(13):3812-4.

136. Combs TP, Wagner JA, Berger J, Doebber T, Wang WJ, Zhang BB, et al. Induction of adipocyte complement-related protein of 30 kilodaltons by PPARgamma agonists: a potential mechanism of insulin sensitization. Endocrinology. 2002 Mar;143(3):998-1007.

137. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010 Sep 15;26(18):2336-7.

138. Pischon T, Boeing H, Hoffmann K, Bergmann M, Schulze MB, Overvad K, et al. General and abdominal adiposity and risk of death in Europe. N Engl J Med. 2008 Nov 13;359(20):2105-20.

139. Zhang C, Rexrode KM, van Dam RM, Li TY, Hu FB. Abdominal obesity and the risk of all-cause, cardiovascular, and cancer mortality: sixteen years of follow-up in US women. Circulation. 2008 Apr 1;117(13):1658-67.

140. Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). JAMA. 2001 May 16;285(19):2486-97.

141. Grundy SM, Brewer HB, Jr., Cleeman JI, Smith SC, Jr., Lenfant C. Definition of metabolic syndrome: Report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. Circulation. 2004 Jan 27;109(3):433-8.

142. Hunt MS, Katzmarzyk PT, Perusse L, Rice T, Rao DC, Bouchard C. Familial resemblance of 7-year changes in body mass and adiposity. Obes Res. 2002 Jun;10(6):507-17.

143. Lin HF, Boden-Albala B, Juo SH, Park N, Rundek T, Sacco RL. Heritabilities of the metabolic syndrome and its components in the Northern Manhattan Family Study. Diabetologia. 2005 Oct;48(10):2006-12.

144. Pilia G, Chen WM, Scuteri A, Orru M, Albai G, Dei M, et al. Heritability of cardiovascular and personality traits in 6,148 Sardinians. PLoS Genet. 2006 Aug 25;2(8):e132.

145. Bayoumi RA, Al-Yahyaee SA, Albarwani SA, Rizvi SG, Al-Hadabi S, Al-Ubaidi FF, et al. Heritability of determinants of the metabolic syndrome among healthy Arabs of the Oman family study. Obesity (Silver Spring). 2007 Mar;15(3):551-6.

146. Kring SI, Holst C, Zimmermann E, Jess T, Berentzen T, Toubro S, et al. FTO gene associated fatness in relation to body fat distribution and metabolic traits throughout a broad range of fatness. PLoS ONE. 2008;3(8):e2958.

147. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007 Oct 18;449(7164):851-61.

148. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010 Sep 1;26(17):2190-1.

149. Yang J, Ferreira T, Morris AP, Medland SE, Madden PA, Heath AC, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet. 2012;44(4):369-75.

150. Lind L, Wallentin L, Kempf T, Tapken H, Quint A, Lindahl B, et al. Growth-differentiation factor-15 is an independent marker of cardiovascular dysfunction and disease in the elderly: results from the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) Study. Eur Heart J. 2009 Oct;30(19):2346-53.

151. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 2009 Jun;5(6):e1000529.

152. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature. 2010 Oct 14;467(7317):832-8.

153. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008 Dec 15;24(24):2938-9.

154. Vandanmagsar B, Youm YH, Ravussin A, Galgani JE, Stadler K, Mynatt RL, et al. The NLRP3 inflammasome instigates obesity-induced inflammation and insulin resistance. Nat Med. 2011 Feb;17(2):179-88.

155. Goossens GH, Blaak EE, Theunissen R, Duijvestijn AM, Clement K, Tervaert JW, et al. Expression of NLRP3 inflammasome and T cell population markers in adipose tissue are associated with insulin resistance and impaired glucose metabolism in humans. Mol Immunol. 2012 Mar;50(3):142-9.

156. Dehghan A, Dupuis J, Barbalic M, Bis JC, Eiriksdottir G, Lu C, et al. Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. Circulation. 2011 Feb 22;123(7):731-8.

157. Dehghan A, Yang Q, Peters A, Basu S, Bis JC, Rudnicka AR, et al. Association of novel genetic loci with circulating fibrinogen levels: a genome-wide association study in 6 population-based cohorts. Circ Cardiovasc Genet. 2009 Apr;2(2):125-33.

158. Orozco G, Barton A, Eyre S, Ding B, Worthington J, Ke X, et al. HLA-DPB1-COL11A2 and three additional xMHC loci are independently associated with RA in a UK cohort. Genes Immun. 2011 Apr;12(3):169-75.

159. Han C, Chen T, Yang M, Li N, Liu H, Cao X. Human SCAMP5, a novel secretory carrier membrane protein, facilitates calcium-triggered cytokine secretion by interaction with SNARE machinery. J Immunol. 2009 Mar 1;182(5):2986-96.

160. Sun XJ, Rothenberg P, Kahn CR, Backer JM, Araki E, Wilden PA, et al. Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein. Nature. 1991 Jul 4;352(6330):73-7.

161. Kilpelainen TO, Zillikens MC, Stancakova A, Finucane FM, Ried JS, Langenberg C, et al. Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. Nat Genet. 2011 Aug;43(8):753-60.

162. Ogg SL, Weldon AK, Dobbie L, Smith AJ, Mather IH. Expression of butyrophilin (Btn1a1) in lactating mammary gland is essential for the regulated secretion of milk-lipid droplets. Proc Natl Acad Sci U S A. 2004 Jul 6;101(27):10084-9.

163. Arnett HA, Escobar SS, Viney JL. Regulation of costimulation in the era of butyrophilins. Cytokine. 2009 Jun;46(3):370-5.

164. Oguri M, Kato K, Yoshida T, Fujimaki T, Horibe H, Yokoi K, et al. Association of a genetic variant of BTN2A1 with metabolic syndrome in East Asian populations. J Med Genet. 2011 Nov;48(11):787-92.

165. Gamazon ER, Zhang W, Konkashbaev A, Duan S, Kistner EO, Nicolae DL, et al. SCAN: SNP and copy number annotation. Bioinformatics. 2010 Jan 15;26(2):259-62.

166. Schmitt E, Ballou MA, Correa MN, DePeters EJ, Drackley JK, Loor JJ. Dietary lipid during the transition period to manipulate subcutaneous adipose tissue peroxisome proliferator-activated receptor-gamma co-regulator and target gene expression. J Dairy Sci. 2011 Dec;94(12):5913-25.

167. Landrier JF, Thomas C, Grober J, Zaghini I, Petit V, Poirier H, et al. The gene encoding the human ileal bile acid-binding protein (I-BABP) is regulated by peroxisome proliferator-activated receptors. Biochim Biophys Acta. 2005 Jun 15;1735(1):41-9.

168. Fisher E, Grallert H, Klapper M, Pfafflin A, Schrezenmeir J, Illig T, et al. Evidence for the Thr79Met polymorphism of the ileal fatty acid binding protein (FABP6) to be associated with type 2 diabetes in obese individuals. Mol Genet Metab. 2009 Dec;98(4):400-5.

169. Flegal KM, Shepherd JA, Looker AC, Graubard BI, Borrud LG, Ogden CL, et al. Comparisons of percentage body fat, body mass index, waist circumference, and waist-stature ratio in adults. Am J Clin Nutr. 2009 Feb;89(2):500-8.

170. Morris DH, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Familial concordance for age at menarche: analyses from the Breakthrough Generations Study. Paediatr Perinat Epidemiol. 2011 May;25(3):306-11.

171. Elks CE, Perry JR, Sulem P, Chasman DI, Franceschini N, He C, et al. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. Nat Genet. 2010 Dec;42(12):1077-85.

172. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav Res Methods. 2008 Aug;40(3):879-91.

173. Fritz MS, Mackinnon DP. Required sample size to detect the mediated effect. Psychol Sci. 2007 Mar;18(3):233-9.

174. He C, Kraft P, Chen C, Buring JE, Pare G, Hankinson SE, et al. Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. Nat Genet. 2009 Jun;41(6):724-8.

175. Chen CT, Fernandez-Rhodes L, Brzyski RG, Carlson CS, Chen Z, Heiss G, et al. Replication of loci influencing ages at menarche and menopause in Hispanic women: the Women's Health Initiative SHARe Study. Hum Mol Genet. Dec 13.

176. Zhao X, Lynch JG, Chen Q. Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. Journal of Consumer Research. 2010;37(2):197-206.

177. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat Genet. 2008 Feb;40(2):161-9.

178. Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs KV, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature. 2010 Aug 5;466(7307):714-9.

179. Cortes A, Brown MA. Promise and pitfalls of the Immunochip. Arthritis Res Ther. 2011;13(1):101.

180. Talmud PJ, Drenos F, Shah S, Shah T, Palmen J, Verzilli C, et al. Gene-centric association signals for lipids and apolipoproteins identified via the HumanCVD BeadChip. Am J Hum Genet. 2009 Nov;85(5):628-42.

181. Li L, Li Y, Browning SR, Browning BL, Slater AJ, Kong X, et al. Performance of genotype imputation for rare variants identified in exons and flanking regions of genes. PLoS ONE. 2011;6(9):e24945.

182. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, et al. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet. 2010 Jun;11(6):446-50.

183. Bansal V, Libiger O, Torkamani A, Schork NJ. Statistical analysis strategies for association studies involving rare variants. Nat Rev Genet. 2010 Nov;11(11):773-85.

184. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. Annu Rev Med. 2010;61:437-55.

185. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet. 2012 Jan 13;90(1):7-24.

186. Fox CS, Liu Y, White CC, Feitosa M, Smith AV, Heard-Costa N, et al. Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. PLoS Genet. 2012 May;8(5):e1002695.

187. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature. 2011 May 5;473(7345):43-9.

188. Gaulton KJ, Nammo T, Pasquali L, Simon JM, Giresi PG, Fogarty MP, et al. A map of open chromatin in human pancreatic islets. Nat Genet. 2010 Mar;42(3):255-9.

189. Waki H, Nakamura M, Yamauchi T, Wakabayashi K, Yu J, Hirose-Yotsuya L, et al. Global mapping of cell type-specific open chromatin by FAIRE-seq reveals the regulatory role of the NFI family in adipocyte differentiation. PLoS Genet. 2011 Oct;7(10):e1002311.

190. Pique-Regi R, Degner JF, Pai AA, Gaffney DJ, Gilad Y, Pritchard JK. Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. Genome Res. 2011 Mar;21(3):447-55.

191. Degner JF, Pai AA, Pique-Regi R, Veyrieras JB, Gaffney DJ, Pickrell JK, et al. DNase I sensitivity QTLs are a major determinant of human expression variation. Nature. 2012 Feb 16;482(7385):390-4.

192. Melnikov A, Murugan A, Zhang X, Tesileanu T, Wang L, Rogov P, et al. Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. Nat Biotechnol. 2012 Mar;30(3):271-7.