5-2015

GENETICS OF OBESITY IN STARR COUNTY, TEXAS MEXICAN AMERICANS

Heather M. Highland

GENETICS OF OBESITY IN STARR COUNTY, TEXAS MEXICAN AMERICANS

by

Heather Michelle Highland

APPROVED:

_____
Craig L. Hanis, Ph.D., Supervisory Professor

_____
Carol J. Etzel, Ph.D.

_____
Xiaoming Liu, Ph.D.

_____
Paul A. Scheet, Ph.D.

_____
Elaine Symanski, Ph.D.

APPROVED:

_____
Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston

GENETICS OFOBESITY IN STARR COUNTY, TEXAS MEXICAN AMERICANS

A
DISSERTATION

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
and
The University of Texas
MD Anderson Cancer Center
Graduate School of Biomedical Sciences
in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by
Heather Michelle Highland
Houston, Texas

May 2015

Acknowledgements

This work would not have been possible without the involvement of many people. Thank you to each and every one of them.

First and foremost, my heartfelt thank you to the residents of Starr County, Texas who have participated in research studies over the years.

My advisor, Craig Hanis, has provided the mentorship that has helped me develop my scientific thinking, and has mentored me through not only my dissertation, but also working in a consortia setting. In addition to his guidance, Craig is responsible for the Starr County Health Studies and the collection of the data set utilized in this dissertation.

Thank you to Piper Below, who has provided scientific guidance along the way, encouraged me on hard days, and celebrated successes.

Many colleagues and friends have helped along the road. Collaborators at the University of Chicago, Graeme Bell, Nancy Cox and members of their groups, who helped me transition from a novice graduate student to a scientist in a matter of months. The T2D-GENES Consortium provided exome sequencing data, but more importantly a network of geneticists and a forum to discuss methodologies and approaches to quantitative genetics. Specifically, Xueling Sim guided me through the set up of existing tools and many useful discussions. Hao Hu helped with expanding the multi-trait analyses for gene-based tests.

Thank you to my current and former committee members, Craig Hanis, Carol Etzel, Elaine Symanski, Paul Scheet, Xiaoming Liu, Alanna Morrison, James Hixson,

and Wenyaw Chan who helped develop this project and push me along when I was distracted.

Thank you to fellow members of the Human and Molecular Genetics program who have provided insight from various perspectives and in addition to friendship.

Thank you to everyone in the Human Genetics Center. There is always someone available to offer guidance, trouble shoot a problem, or discuss the merits of different approaches.

On a personal note, thank you to my friends and family that have supported me along the way. From late nights at the office with Selina Vattathil, to brunches with Lekki Frazier-Wood (and her son) to get me up "early" on the weekends, or just a meal away from the computer with Sarah May. My friends have been here to support me and just as importantly make the journey through graduate school fun.

Thank you to my parents, who encouraged me to live up to my potential and pursue my dreams, supported me through trials and tribulations. Thank you to my Shannon, Trevor, and Christine who have pulled me away from work to remind me to take a break and live life outside of science.

GENETICS OF OBESITY IN STARR COUNTY, TEXAS MEXICAN AMERICANS

Heather Michelle Highland

Advisory Professor: Craig L. Hanis, Ph.D.

Currently, over two-thirds of Americans are classified as over-weight or obese. Obesity increases risk for many other diseases including type 2 diabetes, heart disease, stroke, and cancer, making obesity the largest public health problem in America and most other Westernized nations. Hispanics have a higher rate of both obesity and type 2 diabetes, making them a particularly interesting population in which to study obesity. For the last 33 years, the Starr County Health Studies has collected an array of phenotypes and biological samples from residents of Starr County, along Texas-Mexico border. This study includes 825 subjects who were not known to have diabetes at ascertainment. These subjects have now been seen a second time, on average 8.5 years later. At both visits we measured several aspects of obesity including BMI, bioimpedance to estimate percent body fat, and waist, hip, and arm circumferences. By using multivariate approaches to leverage the array of obesity measures, we have better captured both the amount of adipose tissue and the location of fat deposits.

To assess association of obesity related traits with genetic variation from both genome-wide array data imputed to 1000 Genomes Phase 1 integrated dataset and exome sequencing, both gene-based and single variant tests were conducted. Through these single variant tests, we identified an association with waist to hip ratio and low frequency variants, in two adjacent GABA receptor subunit genes, *GABRB2* and

GABRA6, including a nonsynonymous variant in *GABRA6*. Additional associations include an association with a composite measure of adiposity that encompasses degree of adiposity and location of excess fat above or below the waist and *TREK1,* a gene responsible for trafficking the $GABA_A$ receptor to the cell membrane. Gene based tests of rare variants yielded associations between central versus peripheral adiposity and *ACSL1*, a gene involved in triglyceride biosynthesis. Further replication is required to confirm these associations. While the importance of neuronal signaling pathways in body fat distribution has long been known, many aspects of these pathways are poorly understood. Better understanding of these pathways may identify potential pharmaceutical targets.

Table of Contents

List of Illustrations

List of Tables

Abbreviations

BMI	body mass index
WHR	waist to hip ratio
GWAS	genome wide association study
LepR	leptin receptor
POMC	proopiomelanocortin
α-MSH	alpha Melanocyte stimulating hormone
β-MSH	beta Melanocyte stimulating hormone
PC1	prohormone convertase 1
MC4R	melanocortin 4 receptor
MC3R	melanocortin 3 receptor
SIM1	Single-minded homolog 1 (Drosophila)
NUCB2	nucleobindin 2
BDNF	nucleobindin 2
TrkB	Tyrosine receptor kinase B
NTRK2	neurotrophic tyrosine kinase, receptor, type 2
AgRP	agouti related protein
ARC	Arcuate nucleus of hypothalamus
NPY	neuropeptide Y
PVN	paraventricular nucleus of hypothalamus
CRF	corticotropin-releasing factor
VMH	ventromedial nucleus of hypothalamus
LEP	Leptin

WARGO and obesity	Wilms tumor, aniridia, genitourinary anomalies, intellectual disability,
FLP	familial partial lipodystrophy
DEXA	dual-energy X-ray absorptiometry
MRI	magnetic resonance imaging
cm	centimeter
PCA	principal components analysis
VEGAS	versatile gene-based association study
ESP	exome sequencing project

Chapter 1: Background and Significance

Epidemiology and Impact of Obesity

With obesity prevalence increasing at alarming rates, it is now more important than ever to understand the risk factors and underlying biology of body fat mass and distribution. Obesity is defined as having a body mass index (BMI) of at least 30 kg/m^2(1).The prevalence of obesity has more than doubled between 1995 and 2012(1). Hispanics, particularly in Texas, consistently have a higher prevalence of obesity than the general United States population (Figure 1). The research presented here focuses on identifying genetic factors contributing to obesity specifically in the Mexican-American population of Starr County, Texas.

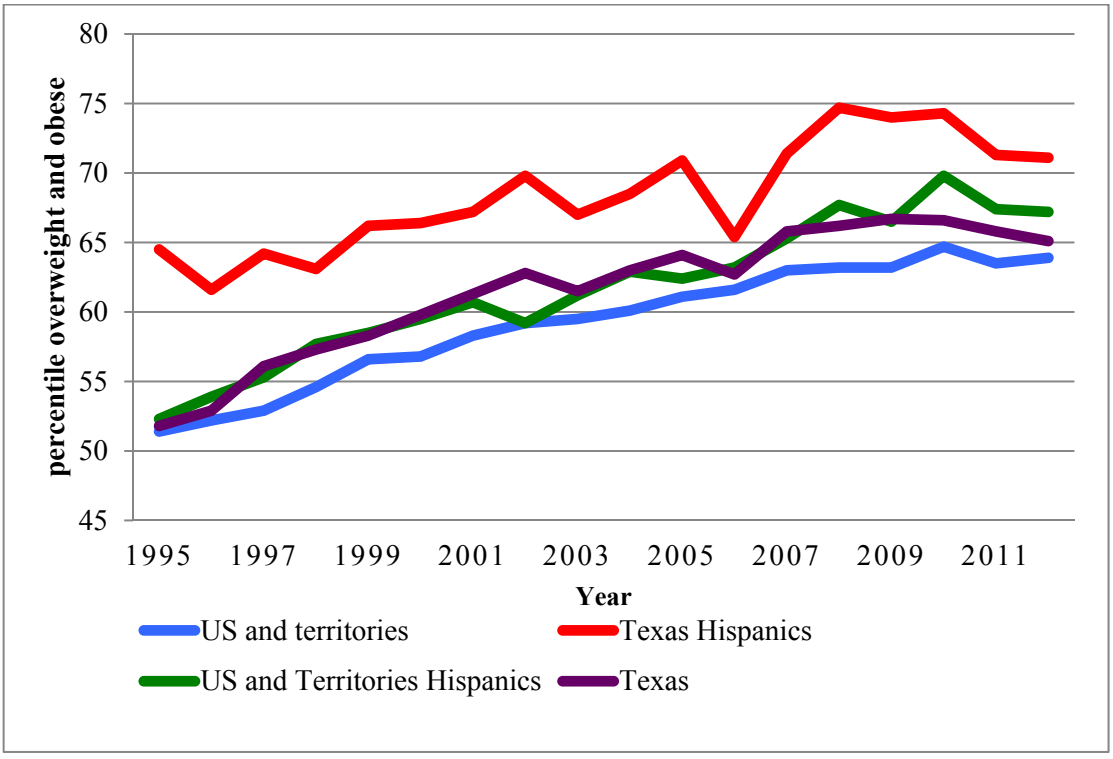


Figure 1. Prevalence of obesity and overweight. The prevalence of BMI > $25kg/m^2$ in the United States and Texas split by ancestry(1).

While one could argue whether obesity itself is a disease, it is certainly a risk factor for the development of many other diseases, including type 2 diabetes, heart disease, cancer, stroke, osteoarthritis, and sleep apnea. A decrease in body fat, lowers an individual's risk of developing these comorbidities. These comorbidities contribute to the considerable economic impact of obesity in the forms of increased medical costs, lost productivity, and the diet and exercise industry. Individuals with elevated BMIs have higher costs for medical care; the biggest absolute increases are attributable to comorbidities of circulatory diseases (10.53% increase) including cardiovascular disease and myeloproliferative diseases (10.67% increase) such as chronic myelogenous leukemia and thrombocytosis in men and musculoskeletal conditions (3.46% increase) and circulatory diseases (4.27% increase) in women(2). Someone with a BMI of 45 will average more than double the medical costs of someone with a BMI of 19(2). In total the excess medical spending attributable to obesity was $147 billion in 2008(3). Beyond increased medical costs, employees with obesity are estimated to have a 22% increase in missed work over their normal-weight counterparts(4). While employers have increased costs due to obesity in the form of higher insurance premiums, lost labor, and decreased productivity, individuals have increased spending as well, largely in the form of medical expenditures and costs associated with attempts to lose weight. The weight loss industry accounts for more than $60 billion spending annually; this spending includes gym memberships, supplements, diet food, and weight loss plans(5).

Lifestyle intervention, diet, exercise, and behavior therapy have all been shown to result in clinically significant weight loss in some people. In an 8-year behavioral intervention, 26.9% of individuals had a net weight loss of at least 10% at the end of the

study(6). This weight loss alters risk for comorbidities in a disproportionate manner; for example behavioral intervention aimed at reducing weight by 7%, also reduces risk of developing type 2 diabetes by 58%(7). The success of these studies is contrary to what was seen in early studies, where it was shown that nearly all weight was regained within 5 years(8), leading to a common perception that achieving long-term weight loss is near impossible. However, population based studies of the NHANES 1999-2006 show that 17.3% of individuals reported long term weight loss, defined as a >10% weight loss maintained for at least a year(9). In NHANES 1999-2002, 33.5% of overweight individuals that had lost at least 10% of their maximum weight experienced weight regain in the prior year(10).

The National Weight Control Registry was designed to look at the behaviors of those that successfully maintained weight loss(11); this study found at 5 and 10 years of follow up that 86.6% of participants maintained at least 10% weight loss(12). Those that started with a bigger initial weight loss showed faster regain while those who had maintained weight loss for more than 2 years at baseline experienced less regain(12). Weight loss and maintenance, which leads to decreased risk of comorbidities, were once perceived to be near impossible. Now it is recognized that while difficult, a subset of the populations can achieve long-term weight loss through lifestyle changes.

For those who cannot achieve sufficient weight loss through lifestyle changes, surgical interventions have become a viable option. Following weight-loss surgery, a majority of patients with diabetes, hypertension, hyperlipidemia, or obstructive sleep apnea had these comorbid condition(s) resolved(13, 14). Due to the overall impact on health, insurance companies now offer coverage for weight-loss surgeries such as gastric

bypass and gastric banding; however, changes in requirements, such as surgery center designation, may keep some from accessing this coverage(15). Different weight-loss surgeries have different risks and benefits. While gastric bypass (laparoscopic Roux-en-Y gastric bypass) yields higher excess weight loss than the gastric band, 69% versus 46%, gastric bypass comes with a higher risk of perioperative complications(16). The adjustable gastric band is subject to long-term complications that may result in a second operation due to band slippage, pouch dilation, or unsatisfactory weight loss(16). In 2013 approximately 179,000 bariatric surgeries were performed in the United States(17). The rate has slowed considerably from the initial exponential growth seen in the early 2000s(15). The decline in surgeries in 2011-2013 seen in Figure 2 may be due to different sources of data, but others have suggested the decline is due to combined economic recession and increased surgery center regulation(15, 18).

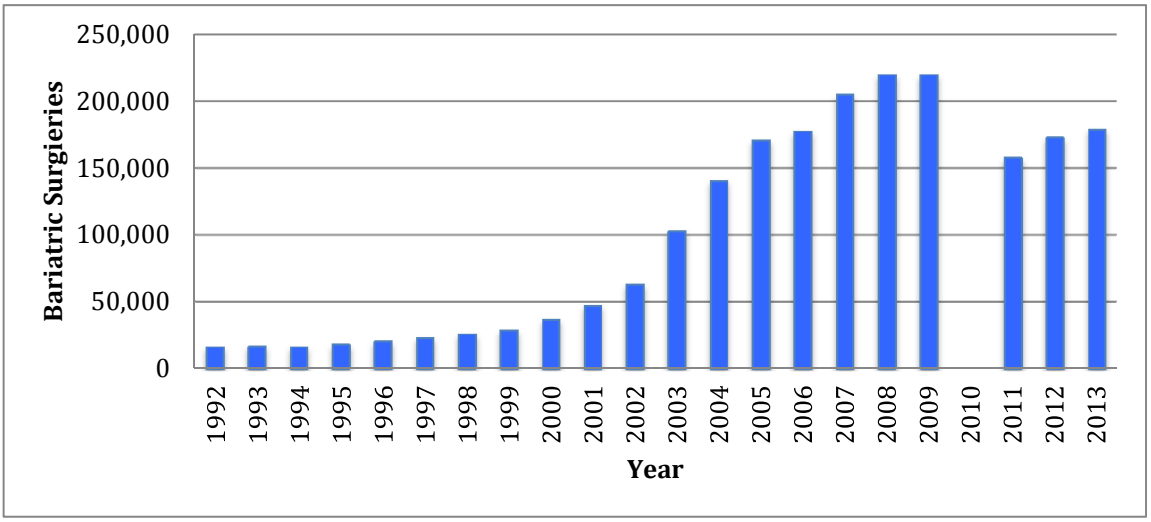


Figure 2. Incidence of Bariatric Surgery in the United States. Estimates from 1992-2009 came from (18). Estimates from 2011-2014 came from (17). No estimates were available for 2010.

Obesity Risk Factors

The body accumulates fat as a means of storing energy when more calories are consumed than are expended. In his "thrifty gene" hypothesis, Neel proposed that evolutionarily, this stored energy allows someone to survive when food resources are scarce(19, 20). In developed societies, humans no longer have extended periods of food scarcity. Furthermore, the typical Western diet now consists of processed foods that are high in fat and sugar that are inexpensive and convenient. While diet is one part of the energy balance, energy expenditure must also be considered. As computers, television and video games have become ubiquitous across America, people are spending more leisure time in a sedentary state. This, coupled with an increase in desk jobs, has led to a significant decrease in total energy expenditure(19).

Even though most members of Western society are exposed to this "obesogenic" environment, not everyone becomes obese(21). This difference may, in part, be due to differences in diet and exercise. In large part, exercise habits and food preferences are learned in childhood(22). Other factors affecting obesity risk are sex, age, smoking, education, socioeconomic status, racial group, and family history(23).

The location of stored body fat is sexually dimorphic. Men have a greater propensity to store excess body fat as visceral adipose tissue, while women are more apt to store excess fat as subcutaneous adipose tissue, meaning overweight men have more fat between their organs while overweight women carry fat just under the skin; this sexual dimorphism increases with age(24). Visceral adipose tissue is thought to have a larger effect on metabolic state and overall health(24).

Obesity prevalence increases with increasing age through the 50's, and begins to decline in the 60's(23). The decline in later years may be due to a combination of a survivor's effect and weight loss due to serious health conditions. The prevalence of obesity is lower for smokers (17.8%) than never smokers (20.9%) or former smokers (23.9%)(23). Weight gain is associated with smoking cessation and the amount of weight gained is correlated with cigarettes smoked per day(25).

Education, socioeconomic status, income, and ethnicity have complex combined relationship with obesity. Obesity prevalence decreases with increasing level of education(23). Lower income is associated with increased obesity prevalence(26). African Americans have higher obesity rates than Hispanics which have higher rates than Caucasians(26). Ancestry contributes to obesity risk through differences in culture and genetics, but is confounded by disparities in income and education(26). Some of the differences may stem from differences in food availability, neighborhood walkability, and access to exercise facilities(26).

Family history encompasses two factors, the common environment shared by people living together and genetics. Close relatives tend to have similar body shapes and sizes. This is reflected in the correlation of BMIs between different relationships as shown in Table 1. Twin studies yield higher heritability estimates (0.60-0.80) than family studies (0.30-0.60)(27); while adoption studies yield estimates similar to family studies in the absence of a common environment(28). Adoption studies, which look at monozygotic twins adopted by different families at birth, show that in the absence of a shared home environment, BMIs have a correlation of 0.66 to 0.70(29). *A priori* there are many factors and pathways likely to be involved in the risk of obesity with each

subject to alteration by genetic variation. These include genes involved in basal metabolic rate regulation, lipid metabolism, carbohydrate metabolism, and neurological factors involved in feeding and exercise behaviors(30). Identification of these genes, however, has proven to be challenging, however one example of a pathway that has repeatedly been implicated in obesity is the leptin-melanocortin pathway.

Relationship	Correlation
Monozygotic twins	0.74
Dizygotic twins	0.32
Full siblings	0.24
Parent-offspring	0.19
Spouses	0.12

Table 1. BMI correlation between relatives. Pooled estimates across studies come from Maes *et al.* (27)

Underlying Signaling Pathways

While many mechanisms are plausible, most variants identified so far associated with obesity are in genes relating to the leptin-melanocortin pathway (shown in Figure 3), which is involved in hunger and satiety. Leptin, a hormone secreted by adipose tissue, travels through blood to the arcuate nucleus of the hypothalamus where it binds to the leptin receptor (LepR) in two types of cells to decrease food intake through two distinct pathways, the anorexigenic pathway and orexigenic pathway(31). In the anorexigenic pathway, leptin binding to the leptin receptor induces transcription of proopiomelanocortin (POMC), which is then cleaved into α-MSH and β-MSH by prohormone convertase 1(PC1)(32). α-MSH and β-MSH travel to the paraventricular nucleus of the hypothalamus and signal through the melanocortin 4 (MC4R) and melanocortin 3 (MC3R) receptors to inhibit food intake and decrease fat storage(32). This is accomplished through unknown mechanisms that involve Single-minded

homolog 1 (Drosophila) (SIM1), nucleobindin 2 (NUCB2), brain derived neurotrophic factor (BDNF) and Tyrosine receptor kinase B (TrkB) that is encoded by *NTRK2*(32). In the orexigenic pathway, the binding of leptin to the leptin receptor inhibits agouti related protein (AgRP) expression in the arcuate nucleus of the hypothalamus(32). AgRP acts as an antagonist of α-MSH by binding MC3R and MC4R, inhibiting signaling of satiety through these receptors(32). NPY is produced in the same neurons as AgRP(33). In mice, the NPY is released in the paraventricular nucleus of hypothalamus and results in increased production of corticotropin-releasing factor and subsequent activation of the hypothalamic pituitary adrenal axis which controls cortisol levels and the "fight or flight response" to stress. (34). NPY is down regulated with increased leptin levels(33). Disruption of the leptin-melanocortin pathway can result in obesity, as demonstrated by several monogenic obesity disorders where mutations in these pathway genes lead to severe obesity. This pathway demonstrates the complex relationship between genes and their gene products. The complexity and redundancy of pathways contributes to the difficulty of identifying genes and variants that contribute to obesity but whose effects may be masked by the complex relationships between gene products.

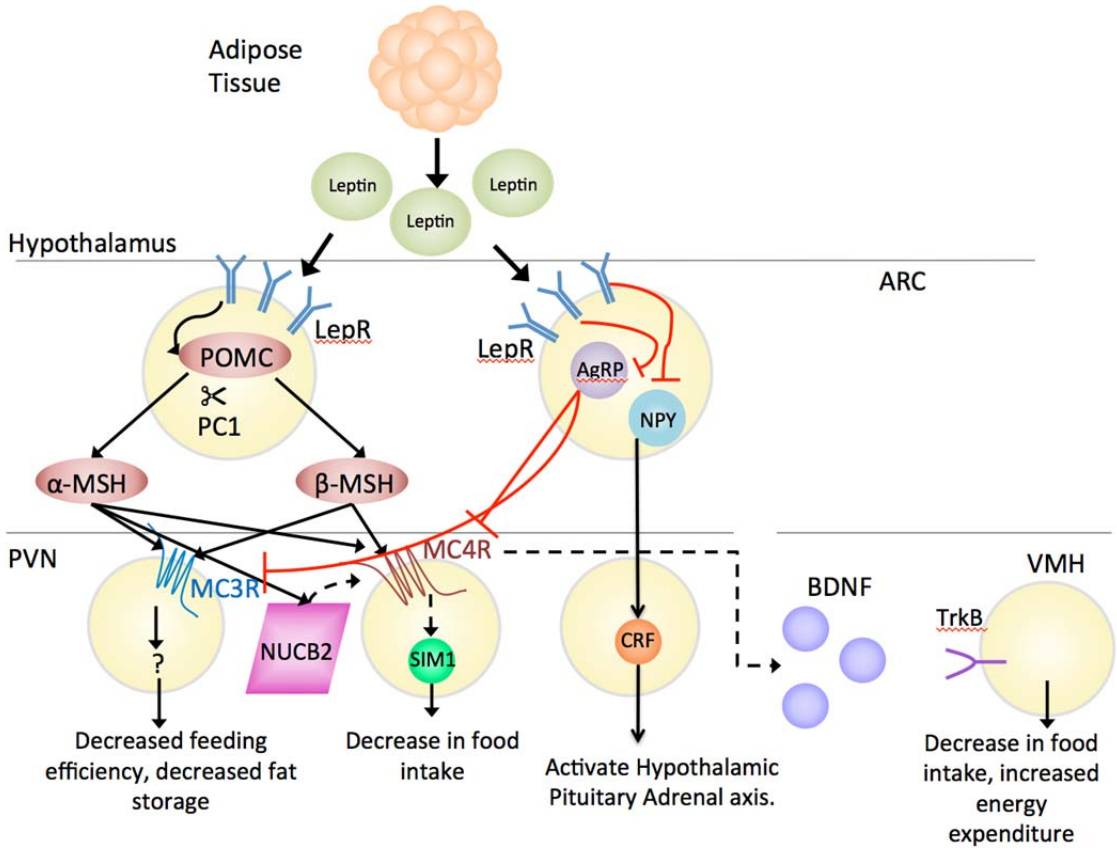


Figure 3. Leptin- Melanocortin Pathway. Image adapted and expanded from Beckers, S., D. Zegers, L. F. Van Gaal, and W. Van Hul. 2009. The role of the leptin-melanocortin signalling pathway in the control of food intake. Critical reviews in eukaryotic gene expression 19:267-287.(32). Abbreviations: ARC: Arcuate nucleus of hypothalamus, LepR: leptin receptor, POMC: proopiomelanocortin, PC1: prohormone convertase 1, AgRP: agouti related protein, NPY: neuropeptide Y, α-MSH: alpha Melanocyte stimulating hormone, β-MSH: beta Melanocyte stimulating hormone, PVN: paraventricular nucleus of hypothalamus, MC3R: melanocortin 3 receptor, MC4R: melanocortin 4 receptor, NUCB2: nucleobindin 2, SIM1: involve Single-minded homolog 1 (Drosophila), CRF: corticotropin-releasing factor, BDNF: brain derived neurotrophic factor, TrkB: Tyrosine receptor kinase B, VMH: ventromedial nucleus of hypothalamus.

Monogenic Disorders

In addition to common obesity, there are single gene disorders that cause obesity or lipodystrophy, sometimes in the context of a more complex syndrome (Table 2). Monogenic forms of obesity are typically characterized by being extreme and very early onset. Many of the genes implicated in single gene disorders are also part of the leptin melanocortin pathway and are characterized by hyperphagia. Mutations in *MC4R* result in the most common form of monogenic obesity, occurring in up to 6% of individuals with severe childhood obesity, meaning a BMI *z*-score ≥ 3 standard deviations from the age and sex specific mean (35, 36). *Leptin* (*LEP)* and *leptin receptor* (*LEPR*) mutations are perhaps the best-recognized genes due to parabiosis experiments involving the first murine obesity models. *ob/ob* (*LEPR* homolog) and *db/db* (*LEP* homolog) knock out mice are phenotypically similar; the mice are hyperphagic, develop morbid obesity and diabetes(37). When sewn together so as to share their circulatory systems with either a wild-type or *db/db* mouse, an *ob/ob* mouse decreases food intake, lose weight, and have decreased blood sugar levels; when paired with the *db/db* mouse, these effects are so severe that the *ob/ob* mouse can die of starvation without intervention(37). In contrast, *db/db* mice, paired with either wild-type or *ob/ob* mouse, continue to gain weight in the form of adipose tissue while its partner starves to death(37). Similarly, humans with mutations in both copies of *LEP*, develop severe obesity in early childhood characterized by hyperphagia, but respond to treatment with leptin injections(38).

Lipodystrophy is class of metabolic disorders characterized by the loss of body fat and sometimes localized accumulation of body fat. Both monogenic forms and acquired forms (see Table 2) have been reported(39). Lipodystrophy patients have

complications of insulin resistance, hepatic steatosis, and hypertriglyceridemia, highlighting that adipose tissues is metabolically active(39).

The third class of monogenic obesity disorders is syndromic disease. These diseases occur with an array of other traits specific to each disorder (see Table 2). Many monogenic disorders include intellectual disability, hypogonadism, malformations of organs, and bone deformities. For example Bardet-Beidel syndrome, which is a ciliopathy characterized by renal abnormalities, retinal degeneration, polydactyly, central obesity, and intellectual disability(40), is a genetically heterogeneous disease, with 19 different genes implicated thus far. Some of these genes form the BBSome, which is a molecule involved in signaling receptor trafficking to the cilia(41). Wilms tumor, aniridia, genitourinary anomalies, intellectual disability, and obesity (WARGO), is caused by a deletion in 11p13; deletions that include genes *WT1, PAX6*, and *BDNF* include an obesity component of the disease(42). These are just two examples of the 35 syndromes that include obesity as a key feature. While monogenic obesity, syndromic obesity, and genetic forms of lipodystrophy provided insight into the pathways contributing to fat mass and distribution, these account for only small fraction of obesity cases, but common variants in or near some of these genes have been implicated in common obesity as discussed below.

Table 2. Monogenic forms of obesity, lipodystrophy, and syndromic obesity.

monogenic disease	characteristics	implicated genes	mode of inheritance	references
Monogenic obesity (isolated)	hyperphagia, severe early onset obesity	*BDNF, CRHR1, CRHR2, LEP, LEPR, MCHR1, MC3R, MC4R, MRAP2, NTRK2, PCSK1, POMC, SIM1*	varies	(43)
Berardinelli-Seip congenital lipodystrophy	lipoatrophic diabetes, acanthosis nigricans, large hands and feet, lipemia, hepatosplenomegaly, insulin resistance	*AGPAT2, BSCL2, CAV1, PTRF*	AR	(39)
Familial partial lipodystrophy (FPL)	partial lipodystrophy, insulin resistance	*CIDEC*	AR	(44)
Familial partial lipodystrophy (FPL)	loss of subcutaneous fat; fat accumulation in face, insulin resistance	*LMNA, PPARG, AKT2, PLIN1*	AD	(39)
Autoinflammation, lipodystrophy, and dermatosis syndrome	annular erythematous plaques, partial lipodystrophy, immune dysregulation, recurrent fever, muscle weakness	*PSMB8*	AR	(45, 46)
Carbohydrate-deficient glycoprotein syndrome type 1a	hypotonia, hyporeflexia, trunk ataxia, growth retardation, lipodystrophy of the buttocks	*PMM2*	AD	(47)
Hutchinson-Gilford progeria	short stature, low body weight, early hair loss, lipodystrophy, scleroderma, aged facial features, decreased joint mobility	*LMNA*	AD, some AR	(48)

monogenic disease	characteristics	implicated genes	mode of inheritance	references
Mandibuloacral dysplasia (MAD), Type A	growth retardation, craniofacial anomalies, mandibular hypoplasiam, lipodystrophy with acral loss of fatty tissue	*LMNA*	AR	(49)
Mandibuloacral dysplasia (MAD), Type B	small chin, nose, and mouth, thin facial skin, skeletal anomalies, generalized lipodystrophy,	*ZMPSTE24*	AR	(50)
Achondroplasia	short-limb dwarfism, characteristic facies, obesity	*FGFR3*	AD	(51)
AHO (Pseudopseudohypoparathyroidism)	resistance to parathyroid hormone, thyroid-stimulation hormone, and gonadotropins; short stature, obesity, round facies, subcutaneous ossifications, brachydactyly, intellectual disability	*GNAS*	Maternally inherited defect	(52)
Alstrom syndrome	blindness, sensorineural hearing loss, childhood obesity, hyperinsulinemia, type 2 diabetes	*ALMS1*	AR	(53)
Angelman syndrome with obesity	developmental delays, movement/balance disorder, frequent laugher/smiling, speech impairment, microcephaly, seizures,	*UBE3A*	Maternally inherited defect	(54)
Atypical progeroid syndrome	lipodystrophy and progeroid syndrome	*LMNA*	AD	(55)

monogenic disease	characteristics	implicated genes	mode of inheritance	references
Bardet-Biedl syndrome	renal abnormalities, polydactyly, retinal degeneration, obesity	*ARL6, BBIP1, BBS1, BBS10, BBS12, BBS2, BBS4, BBS5, BBS7, BBS9, WDPCP, CCDC28B, CEP290, IFT27, LZTFL1, MKKS, MKS1, TMEM67, SDCCAG8, TRIM32, TTC8*	varies, mostly AR	(56-60)
Borjeson-Forssman-Lehmann syndrome	severe mental defect, epilepsy, hypogonadism, hypometabolism, obesity, characteristic facies,	*PHF6*	X-linked	(61)
Brachydactyly mental retardation syndrome	short stature, stocky build, intellectual disability, brachymetaphalangia, eczema, obesity	*GPR35,*	AD	(62)
Carney complex with primary pigmented nodular adrenocortical disease and Cushing's syndrome	tumors, myxomas in the heart, endocrine tumors, Cushing's syndrome, weight gain	*PRKAR1A*	AD	(63)
Carpenter Syndrome 1	acrocephaly, peculiar facies, brachydactylyl, congenital heart defects, intellectual disability, hypogenitalism, and obesity	*RAB23*	AR	(64)
Carpenter Syndrome 2	craniosynostosis, polysyndactly, obesity, umbilical hernia, cryptochidism, congenital heart disease	*MEGF8*	AR	(65)
Choroideremia with deafness and obesity		*CHM, DFN3*	X-linked	(66)

monogenic disease	characteristics	implicated genes	mode of inheritance	references
Cohen syndrome	psychomotor retardation, clumsiness, microcephaly, hypotonia and joint laxity, progressive retinochoroidal dystrophy, thick hair, short philtrum, characteristic facies, obesity	*VPS13B*	AR	(67)
Combined pituitary hormone deficiency	panhypopituitary dwarfism, deficiency of pituitary hormones, increased weight,	*PROP1*	AR	(68)
Cortisone reductase deficiency	ACTH-mediated adrenal hyperandrogenism, males: precocious pseudopuberty; females: hirsutism, oligomenorrhea, infertility, overweight	*H6PD*	complex	(69)
Fanconi-Bickel syndrome	hepatorenal glycogen accumulation, proximal renal tube dysfunction, impaired glacatose and glucose utilization, facial obesity, lipodystrophy	*SLC2A2*	AR	(70)
Fragile X syndrome with Prader-Willi-like phenotype	intellectual disability, obesity, anal atrasia	*FMR1*	X-linked	(71)
Insulin resistance syndromes	severe insulin resistance, acanthosis nigricans, accelerated growth, obesity, polycystic ovary syndrome	*INSR*	AD, some AR	(72)
Isolated growth hormone (GH) deficiency	dwarfism, delay bone maturation, micropenis, fasting hypoglycemia, truncal obesity, young facial appearance, high pitched voice	*GH1, GHRHR*	AR, some AD	(73)

monogenic disease	characteristics	implicated genes	mode of inheritance	references
MEHMO syndrome	intellectual disability, epileptic seizures, hypogonadism and hypogenitalism, microcephaly, and obesity	*MEHMO*	X-linked	(74)
Mental retardation X-linked, syndromic 16	intellectual disability, emotional disturbances, hypotonia, obesity, and gynecomastia	*MECP2*	X-linked	(75)
Mental retardation X-linked, syndromic 7	intellectual disability, obesity, hypogonadism, and tapering fingers	*MRXS7*	X-linked	(76)
Mental retardation, X-linked, syndromic 11	intellectual disability, characteristic facial dysmorphic features, obesity, large testes	*MRXS11*	X-linked	(77)
Multiple endocrine neoplasia, type 1 with Cushing's disease	tumors of endocrine tissues, including pituitary and adrenal tumors	*MEN1*	AD	(78)
Prader-Willi syndrome	decreased fetal activity, intellectual disability, short stature, hypogonadotropic hypogonadism, small hands and feet, obesity	*GABRG3, IPW, MAGEL2, MKRN3, NDN, PWCR1, SNRPN*	Paternally inherited defect	(79)
Prader-Willi-like syndrome (chromosome 6q)	hypotonia, progressive obesity, delayed developmental milestones, small extremities	*SIM1*	AD	(80)
Prader-Willi-like syndrome, X-linked	hypogenitalism, obesity, intellectual disability	*PWLSX*	X-linked	(81)
Simpson-Golabi-Behmel 1	pre- and postnatal overgrowth, congenital heart defects, coarse facies.	*GPC3, GPC4*	X-linked	(82, 83)

monogenic disease	characteristics	implicated genes	mode of inheritance	references
Simpson-Golabi-Behmel 2	developmental delay, macrocephaly, early death, intellectual disability, dysmorphic facies, obesity	*OFD1*	X-linked	(84)
Thyroid hormone resistance syndrome	Resistance to thyroid hormone, goiter, short stature, obesity	*THRB*	AR	(85)
Ulnar-Mammary (Schinzel) syndrome	ulnar-ray defects, small penis, delayed puberty, obesity, abnormal breast development	*TBX3*	AD	(86)
WAGR syndrome with obesity	Wilms tumor, aniridia, genitourinary anomalies, intellectual disability, and obesity	11p13 deletion including *PAX6, WT1, BDNF*	AD	(42)
Wilson-Turner syndrome	dysmorphic facial features, hypogonadism, short stature, truncal obesity, severe intellectual disability	*HDAC8*	X-linked	(87)

Table 2. Monogenic forms of obesity, lipodystrophy, and syndromic obesity.

Genetics of Common Obesity

Investigations of monogenic and syndromic obesity and lipodystrophy have identified several genes, including many from the leptin melanocortin pathway, capable of causing obesity. These studies certainly demonstrate that genes cause obesity, but carriers of these mutations represent only a small proportion of obese people. To identify genes involved in the common forms of obesity in humans, over 500 candidate gene studies, 95 genome-wide linkage studies, and 43 genome-wide association studies (GWAS) with obesity-related traits have been performed(88). The results of these studies have established associations with obesity-related traits on all 22 autosomes and the X chromosome(88). The linkage studies, in particular, implicate broad regions of the genome not specific variants and genes. Genome wide SNP arrays began to dominate the field in 2005. The 43 genome wide association studies published to date on obesity and adiposity related traits have identified 305 variants at 167 loci associated with an obesity related trait at $P < 5\times10^{-8}$ (Figure 4)(89-131). Key findings include associations with variants in and near genes that are part of the leptin-melanocortin pathway, including *MC4R, BDNF,* and *PCSK1* as well as many other genes, such as *LYPLAL1*, *NEGR1*, and *NRXN3*, that are also expressed in the brain and thought to be involved in hunger and satiety pathways(21).

Two additional well-replicated loci are *FTO* and *INSIG2*. A signal in the noncoding region of *FTO* was initially attributed to the *FTO* gene, due to increased fat mass in murine models in which increased *FTO* expression was induced (132, 133). Despite this, the BMI increasing variants have not been connected to changes in *FTO* expression or function. These variants are eQTLs for the neighboring gene, *IRX3*(134).

Smemo *et al.* show that the BMI associated region of *FTO* interacts with the promoter of *IRX3* altering the expression pattern(134). Further support for the role of *IRX3* in body fat mass comes from knock out murine models that demonstrate a 25-35% decrease in body mass in comparison to wild type(134). For this locus the field has been able to elucidate a mechanism, through the *IRX3* gene, albeit the field was initially dominated by *FTO* studies themselves. This is an exception, however, as generally linkage and genome wide chip arrays leave us with broad regions of the genome associated with the trait and no causal variant or gene.

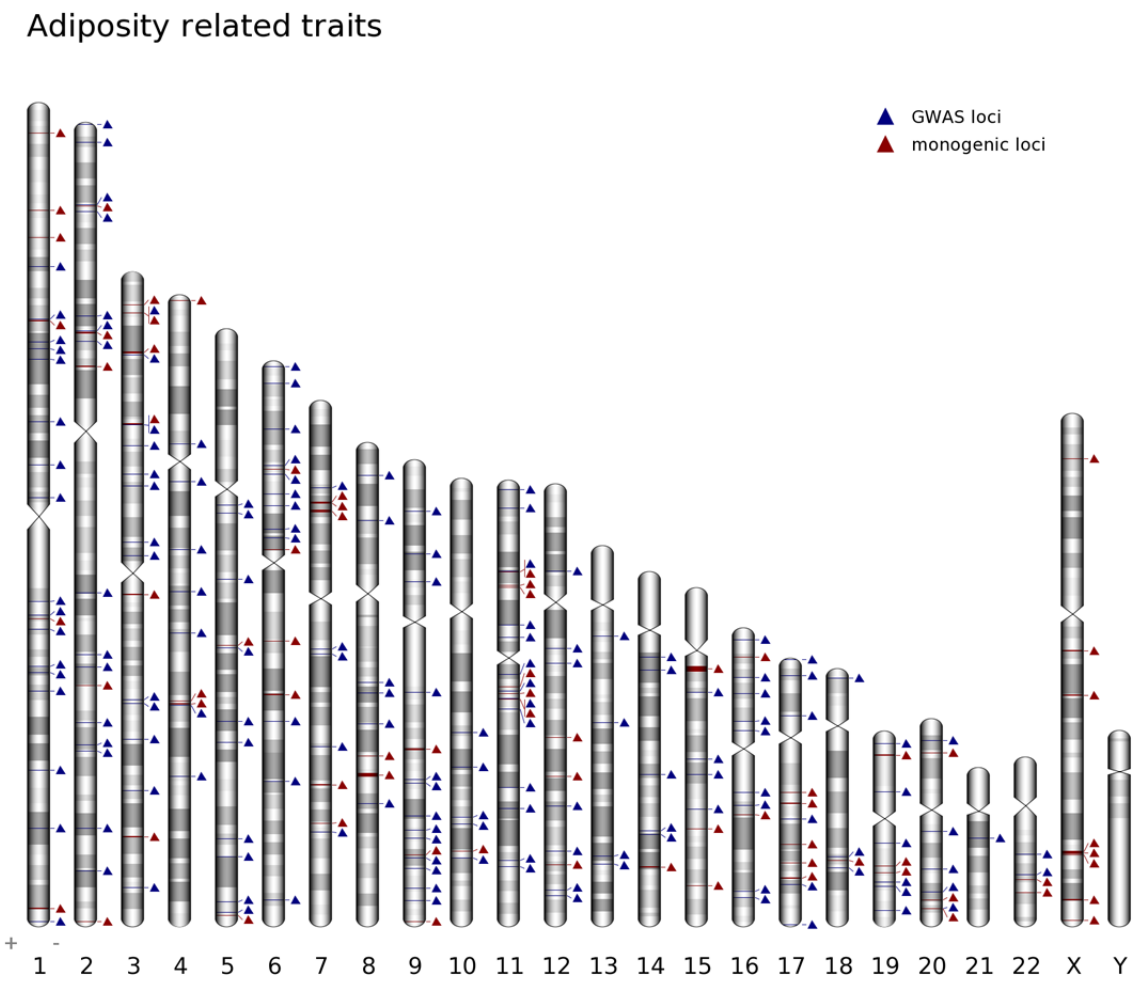


Figure 4. Obesity Loci Identified by GWAS. Loci identified as being associated with obesity-related traits at *p* value< $1x10^{-8}$ are indicated with blue triangles. Monogenic adiposity loci are indicated by red triangles. Variants within 1 mega-base are represented only once(135).While genome-wide chip data provide an agnostic look, the hypothesis that coding variants are more likely to alter gene function justifies looking more closely at coding variation. As the cost of sequencing has decreased, there has been a recent influx of data in the form of whole genome sequencing, whole exome sequencing, and an array targeting low frequency variation seen by exome sequencing data (http://genome.sph.umich.edu/wiki/Exome_Chip_Design). To date no large-scale exome

array or exome sequencing studies have been published for adiposity traits, although these studies are currently underway.

With the increasing prevalence of obesity, it is important to understand the genetic basis of disease. Each gene involved in obesity represents a potential mechanism that could be exploited to treat obesity. Heritability estimates indicate there is a substantial role for genetic variation, however currently identified genes and variants explain less than 3% of the variance in BMI(130). Large-scale studies have investigated associations of common variation and only the most basic measures of obesity, BMI and WHR, in predominately European samples. To identify novel associations, the work presented here utilizes additional adiposity measures, rare coding variants as well as common variants in Hispanics from Starr County, Texas.

Starr County Health Studies

To investigate the role of genetic variation in the amount and distribution of body fat, a sample of non-diabetic individuals from Starr County, Texas will be used. Starr County, Texas lies along the Texas-Mexico border, approximately 100 miles inland from the Gulf of Mexico. The population of Starr County is over 95% Hispanic(136).The population is overwhelmingly poor, with over 40% of individuals living below the poverty line; the median income between 2008 and 2012 being just $24,653(136). The combination of homogenous low income and single ancestry composition makes this an ideal population for genetic studies.

The contemporary Mexican-American population is the result of admixture of Native Americans, Europeans and Africans. Ancestry was computed for genome-wide array data using the 1000 Genomes reference populations using ADMIXTURE

(http://www.genetics.ucla.edu/software/admixture/)(137, 138). On average in Starr County, Mexican-Americans have 61% European Ancestry, 37% Native American Ancestry, and 2% African Ancestry. This is similar to previous reports in this population using microsatellite markers and blood groups(139, 140).

Epidemiological studies in Mexican-Americans in Starr County started in 1981. These early studies looked at familial aggregation of diseases including type 2 diabetes(141), gallbladder disease(142), and hypertension(143). These studies showed an exceptionally high prevalence of type 2 diabetes resulting in a long-term focus on diabetes, its complications, risk factors, and interventions(144-147).

From 2002-2006 a group of 1,345 individuals were sampled from the population to be representative of adults in Starr County not diagnosed with type 2 diabetes (148). During this study, subjects underwent an oral glucose tolerance test, electrocardiogram, anthropometrics measurement, blood pressure determination, and collection of blood samples. All of these individuals were invited to take part in a second study occurring from 2010-2013; 57% of individuals took part in this subsequent study. During the second visit, most of the same measures as before were obtained as well as additional measures including cardiovascular measures from an echocardiogram and an in-home sleep study. The individuals participating in these two examinations form the sample for study in this dissertation.

Because obesity is a major risk factor for developing type 2 diabetes, it has been studied through the years in the Starr County Health Studies. One of the first genetic studies of obesity in Starr County was an affected sib-pair study that looked for linkage between the *Leptin* gene and obesity(149).While there was no evidence of linkage with

obesity in this study, a later study which used families, did find linkage between *Leptin* variation and waist to hip ratio (WHR)(30). This larger study also found evidence of linkage between obesity and *NPY*, but no evidence of linkage for several other candidate genes including *LEPR, GLP1R,* and *UCP1(30)*.

More recently, association signals from large-scale genotyping studies have been replicated in samples from Starr County. For example, Herbert et al. identified rs7566605, a variant near *INSIG2,* associated with BMI in both children and adults(150). Replication efforts, including data from Starr County did not find evidence of association with this variant and an array of adiposity traits including obesity, BMI, weight, waist circumference and WHR(151). The variability of replication of this variant across studies has been the subject of much discussion. A meta-analysis looked at sources of heterogeneity across studies(152). Heid *et al.* found evidence that the effect is largest when comparing normal weight individuals to individuals with extreme obesity.

Here I will expand on the obesity related analyses in Starr County. First I will investigate measures of obesity using multivariate techniques to capture overall adiposity and distribution of fat. I will then carry these measures forward as outcome variables for genetic analyses, using both genome-wide array and whole exome sequencing data. In addition to standard single variant analyses, gene-based analyses that aggregate signal across a gene, accounting for linkage across sites in the case of common variants will be utilized. Finally I will investigate the array of variation in genes known to cause monogenic obesity disorders.

Chapter 2: Alternative obesity measures

INTRODUCTION

While obesity is defined as having a body mass index of at least 30 kg/m^2, it is actually excess adipose tissue that is associated with the risk of comorbidities. The ease with which BMI can be calculated makes it an attractive measure; however, it essentially is only weight adjusted for height and does not account for either the type of tissue or fat distribution. Two individuals with the same BMI can have very different body compositions. For example, athletes are often classified as overweight or obese despite having low amounts of adipose tissue and increased muscle mass(153). In Starr County, data show that individuals with the same BMI can have markedly different distribution/disease profiles. Therefore, it seems prudent to examine alternative measures of obesity and fat distribution that may more adequately reflects the underlying biological processes. Measures that better reflect the underlying biological processes may be more amenable to discover the genetic effects. In this chapter, alternative composite measures of obesity are explored.

Percent body fat quantifies the proportion of adipose tissue in an individual's body. The gold standard for calculating body fat is through medical imaging such as dual-energy X-ray absorptiometry (DEXA) or magnetic resonance imaging (MRI); however, both of these involve expensive medical equipment, require specialized training and expose subjects to radiation(24, 154). Bioelectrical impendence analysis is a proxy for percent body fat. Bioelectrical impendence analysis is obtained sending a small electrical current through the body and measuring resistance and reactance(155). This can be done using small portable devices that require little training. Scales utilizing

this technology can now be readily purchased for home use. Offsetting the advantages of bioimpedance is the fact that it is sensitive to factors such as dehydration. Even so estimates of percent body fat derived from bioimpedance have a correlation greater than 0.9 with measures from DEXA(155). As with BMI, quantifying the amount or percent of adipose tissue does not specify where the adipose tissue is.

Adipose location and type are important considerations, because deposits in different regions of the body have varying metabolic properties. For example, subcutaneous adipose tissue is thought to have a lesser role in the risk of developing comorbidities whereas excess visceral adipose tissue increases risk for an array of metabolic disease such as cardiovascular disease and type 2 diabetes(24). To better capture the distribution of body fat, measures such as circumferences and skin folds are used. Waist circumference and waist to hip ratio reflect central adiposity, but they are correlated with BMI(156). Adjusting for BMI has led to the successful identification loci specific to central adiposity(107, 157). Mid upper arm circumference is a measure of subcutaneous adipose tissue(158). Another measure of subcutaneous adipose tissue is skin fold thickness at a variety of anatomical sites. Skin fold measures have poor reproducibility and are more difficult to obtain in obese individuals(159).

BMI and other single measures of adiposity and adiposity distribution are limited in their ability to capture the amount of fat and fat distribution simultaneously. The confounding and correlation between these measures is likely to hinder the identification of loci and variants. Previously, principal components analyses and factor analyses of skinfold measures, percent body fat, body circumferences, BMI, and ratios between some of these measures have identified factors that have an independent genetic basis

and represent total obesity, subcutaneous fat, or fat distribution between extremities and trunk (160-163).

While the Starr County data does not contain skinfolds, and thus cannot be directly compared to any of these prior studies, the multiple circumferences, BMI, and percent body fat allow the use of principal components to create composite measures of obesity that capture both magnitude and distribution of adiposity.

METHODS

In the Starr County Health Studies, measurements of mid upper arm, hip, and waist circumferences, height, weight and bioimpendance were assessed at two visits an average of 8.5 years apart. Waist measurements were taken to the nearest 10^{th} of a centimeter (cm) while holding a tape measure horizontal to the ground at the umbilicus. An observer verified the tape measure remained horizontal. Hip measurements were taken at the widest circumference to the nearest 10^{th} of a cm. Arm circumference was measured half way between the shoulder and elbow to the nearest 10^{th} of a cm. Height was measured without shoes using a wall-mounted stadiometer to the nearest 10^{th} of a cm. Weight was measured to the nearest 10^{th} of a kilogram using a balance beam scale. These were used to calculate BMI. Weight at age 18 and weight at maximum were both self reported, and measured height was used to calculate BMI. Bioimpedance was measured using a bioimpedance device from RJL Systems (Clinton Township, NJ). Fat free mass, and subsequently percent body fat, were then calculated using the equations of Segal *et al.*(155). The average and standard error for each of these measurements at each visit and the correlation between the two visits are presented in Table 3, excluding individuals diagnosed with type 2 diabetes prior to the visit or without genetic data.

Individuals that meet diagnostic criteria for type 2 diabetes at the study visit are retained in analyses, because they have not been exposed to diabetes treatments that alter body composition. On average, individuals gained 3 kg, 1.6 cm on the hips and 5.1 cm on the waist between visits.

	VISIT 1	VISIT2	correlation
Sample size	825	438	-
Newly diagnosed T2D	5.40%	9.50%	-
Percent female	71%	76%	-
Age (years)	39.6 (9.6)	48.5 (8.9)	-
Weight (kg)	79.3 (18.5)	81.3 (18.1)	0.90
Height (cm)	161.6 (8.7)	160.3 (8.2)	0.98
BMI (kg/m^2)	30.3 (6.3)	31.7 (6.6)	0.89
BMI at age 18	-	23.4 (4.9)	-
BMI at maximum	-	35.1 (7.7)	-
Waist circumference (cm)	97.1 (15.0)	105.8 (15.5)	0.82
Hip circumference (cm)	109.2 (12.5)	112.1 (13.9)	0.87
WHR	0.89 (0.08)	0.91 (0.07)	0.72
Percent body fat	32.1 (10)	34.8 (9.7)	0.86
Arm circumference (cm)	32.9 (4.6)	33.8 (4.9)	0.80

Table 3. Characteristics of samples at each visit. Means and standard deviations are given for each anthropometric measure at each of two study visits. The Pearson's correlation coefficients between measures from the two visits are given in the last column.

Each measure of adiposity has its own strengths and weaknesses. To capture a composite of these measures, I combined all the measures at each visit using principal components analysis (PCA). Here, the correlation matrix was used since variables have different scales. PCA transforms the original set of N measures into N uncorrelated linear combinations with the first accounting for the largest proportion of variation and the Nth accounts for the least variation. Each principal component is orthogonal

(uncorrelated) to each prior component. For each time point I utilized five measures: BMI, percent body fat, waist, hip, and arm circumferences.

RESULTS

The proportion of variance explained by each principal component is in Table 4. The similarity between the principal components analysis for the two visits is striking. For visit 1 and visit 2, the first principal component, which accounts for 82.3% and 82.6% of the variation, respectively, is strongly correlated with BMI (r^2 = -0.98 and -0.97, respectively) as shown in Tables 5 and 6 and plotted in Appendix Figures 1 and 2. The first principal component has similar loadings, the weight each variable has in the component, for all five of the correlated obesity measures Tables 5 and 6.

	proportion of variance explained	
	VISIT 1	**VISIT 2**
PC1	0.823	0.826
PC2	0.092	0.089
PC3	0.050	0.047
PC4	0.025	0.026
PC5	0.010	0.012

Table 4. Proportion of variance explained by each principal component.

		Comp.1		**Comp.2**		**Comp.3**		**Comp.4**		**Comp.5**	
	scaling	loading	corr.	loading	corr.	loading	corr.	loading	corr.	loading	corr.
BMI	6.292	-0.483	-0.980	0.000	0.010	0.000	-0.007	0.000	0.019	0.874	0.200
waist	15.016	-0.440	-0.890	0.469	0.320	-0.525	-0.260	0.476	0.170	-0.289	-0.065
hip	12.484	-0.466	-0.950	0.000	-0.030	-0.271	-0.140	-0.814	-0.290	-0.212	-0.048
PBF	0.100	-0.405	-0.820	-0.822	-0.560	0.000	0.014	0.327	0.120	-0.230	-0.052
arm	4.641	-0.437	-0.890	0.320	0.220	0.806	0.400	0.000	0.009	-0.236	-0.053

Table 5. Principal component loadings and correlations for visit 1.

		Comp.1		Comp.2		Comp.3		Comp.4		Comp.5	
	scaling	loading	corr.	loading	corr.	loading	corr.	loading	corr.	loading	corr.
BMI	6.500	-0.480	-0.970	0.000	-0.028	0.000	0.008	-0.144	-0.052	0.867	0.210
waist	14.348	-0.440	-0.890	-0.456	-0.300	0.572	0.280	0.482	0.180	-0.217	-0.049
hip	12.826	-0.467	-0.950	0.000	0.066	0.226	0.110	-0.757	-0.280	-0.369	-0.096
PBF	0.097	-0.408	-0.830	0.809	0.540	0.000	-0.024	0.403	0.150	-0.134	-0.030
arm	4.659	-0.439	-0.890	-0.355	-0.240	-0.786	-0.380	0.105	0.380	-0.215	-0.057

Table 6. Principal component loadings and correlations for visit 2.

The second principal component accounts for 9.2% and 8.9% of the variance for visit 1 and visit 2 respectively; but unlike BMI is not strongly correlated with any one trait. The loadings for principal component 2 (PC2) are zero for both BMI and hip circumference. PC2 increases with increasing waist circumference, increasing arm circumference and decreasing percent body fat, resulting in a separation between males and females. The direction of the loadings is flipped for visit 2, but the magnitude is similar. PC2, although uncorrelated with hip circumference, is more correlated with waist to hip ratio than either waist circumference or arm circumference, as shown for visit 1 in Figure 5. PC2 separates men with excess abdominal fat from women with small waist to hip ratios. This indicates that PC2 appears to be capturing degree of central adiposity.

PC3 captures 5% and 4.7% of the variation at visit 1 and visit 2, respectively. PC3 increases with increasing arm circumference and decreases with both waist and hip circumferences. Unlike PC2, PC3 is not sexually dimorphic. This separates individuals that have large waist, hip, and arm circumferences from those with large waist and hip circumferences, but average arm circumference. PC3 is strongly correlated with the ratio of arm circumference to waist circumference (Figure 6).

This indicates PC3 is capturing disproportionately central versus peripheral adiposity.

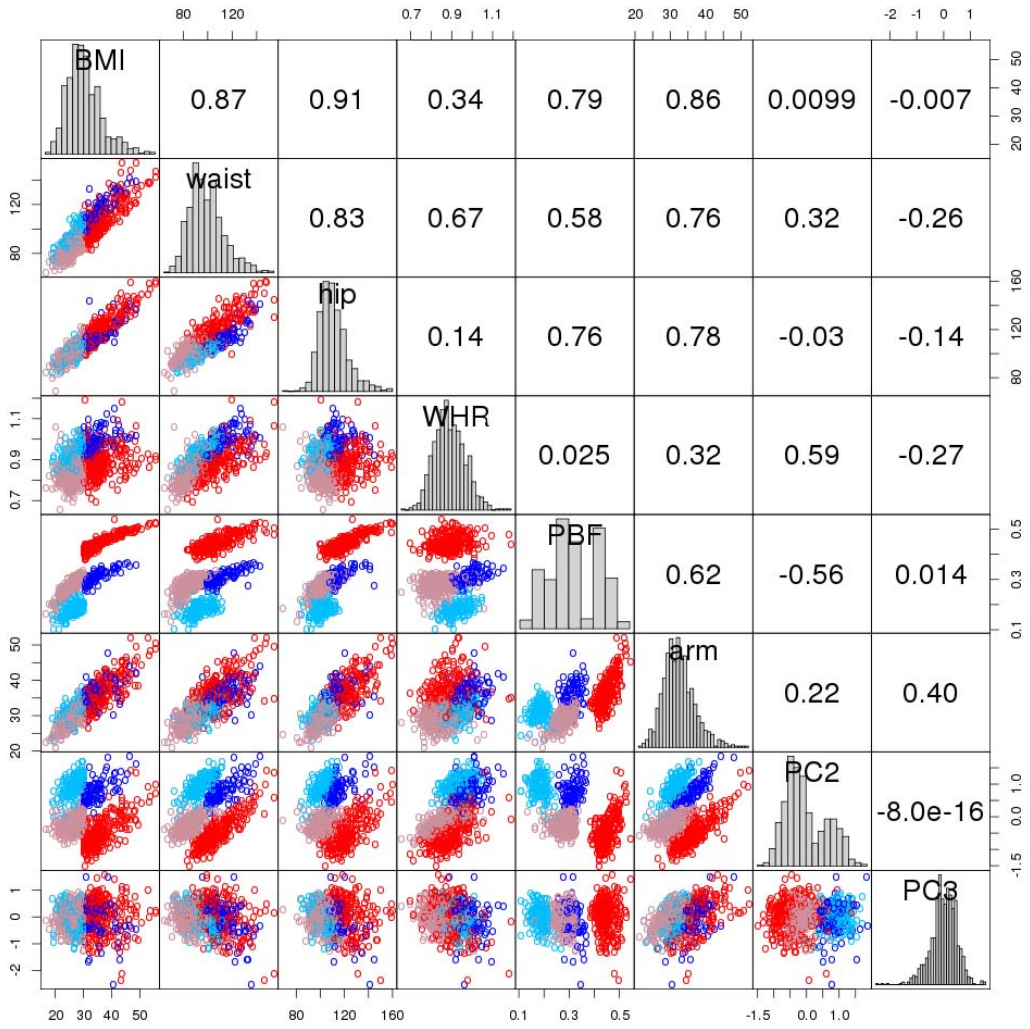


Figure 5. Relationship between PC2 and PC3 with adiposity measures for visit 1. The bottom half shows the scatter plot for each pair of traits, with PC2 in the bottom row. The upper half shows the Pearson correlation coefficient for the two measures. Histograms of the traits are displayed in the middle. Points are color coded by sex and obesity status. Obese males are dark blue; non-obese males are light blue; obese females are red, non-obese females are pink.

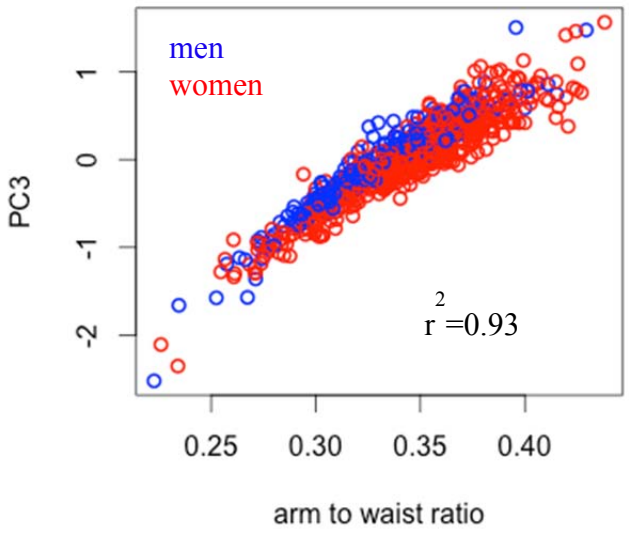


Figure 6. Correlation between PC3 and arm to waist ratio at visit 1.

DISCUSSION

Of the many measures of adiposity available, some are easier to obtain in the field than others. With the gold standards being derived from imaging that is cost-prohibitive, alternative measures are more accessible. By combining multiple measures that require minimal equipment, I have created composite measures, PC2 and PC3. PC1 is strongly correlated with BMI. Due to the loss of clinical interpretability when using the composite measure, BMI, not PC1, will be used as an outcome for genetic analyses. For PCA, BMI was not split into weight and height, because BMI represents a non-linear combination of the two traits. When including weight and height separately, the second principal captured height, instead of adiposity.

PC2, like waist to hip ratio, exhibits marked sexual dimorphism. There are two sources of sexual dimorphism in these adiposity traits. The first is that there are real differences in where men and women store excess fat, as discussed in chapter 1. The

second comes from the equations used to calculate percent body fat, with four specific equations for males and females with and without obesity(155). To account for this, sex will be used as a covariate in all genetic analyses.

PC2 polarized men with excessive abdominal fat from women with extremely low waist to hip ratios; this can be thought of as a measure of central adiposity. PC3 separates those with excessive adiposity all over from those with high central adiposity and average arm circumferences and can be thought of as a measure of central versus peripheral adiposity. Similarly Comuzzie *et al.* found that in principal components analyses using skinfolds the first PC represents overall magnitude of skin folds, the second upper vs. lower subcutaneous adiposity, and the third central versus peripheral subcutaneous adiposity(163). Livshits *et al.'s* factor analysis of BMI, body fat, skinfolds, and circumferences resulted in four primary factors with factor 1 representing the skin folds, factor 2 representing total fat and BMI, factor 3 representing central adiposity distribution and factor 4 representing the subcutaneous fat on the trunk versus extremities(160). PC2 here is analogous to Comuzzie *et al.'s* PC 2 and Livshits *et al.'s* Factor 3, while PC3 here is analogous to Comuzzie *et al.'s* PC 3 and Livshits *et al.'s* Factor 4(160, 163). Both of these studies looked at the relationship between these traits across family members and found there is a significant genetic component(160, 163)

By analyzing these composite measures, which I refer to as PC2 and PC3 throughout, in addition to standard measures of adiposity I aim to identify loci that are involved in body fat distribution and overall adiposity.

Chapter 3: Genome wide SNP analysis

INTRODUCTION

As previously discussed in chapter 1, obesity related traits have a high heritability. To better understand the genetic basis of obesity related traits at least 41 genome-wide association studies have been completed. Traditionally, these studies have looked at the evidence that each individual variant is associated with obesity, BMI, waist circumference, or waist to hip ratio. With time the density of variants has increased. Early studies genotyped a mere 100,000 SNPs; today it is common to have over 10 million variants after imputation, as we do here.

While many associations have already been successfully identified, only 2.7% of the variation in BMI is explained(130). Factors such as different allele and haplotype frequencies in different populations, interaction with the environment, and distribution of the trait can give rise to different association signals. Utilizing an array of adiposity measures increases the power of this data set. Two multi-variate approaches will be used to leverage this data set. First, the composite measures of adiposity introduced in chapter 2 will be utilized. Second, by meta-analyzing association results for array of adiposity traits while correcting for the correlation between the traits, we will be able to detect variants that have a pleiotropic effect. Further, beyond testing for association with each variant, we also test for a cumulative association across genes using versatile gene-based association study (VEGAS) (164). By utilizing VEGAS, we will be able to identify genes that have multiple variants associated with an outcome at levels lower than single variant significance thresholds.

METHODS

Genotyping: Genomic DNA isolated from whole blood from 1980 unique individuals form Starr County was genotyped on the Affymetrix Genome-Wide Human SNP Array 6.0 at the Center for Inherited Disease Research (CIDR). Genotypes were called with two algorithms, Birdseed v2(165) and corrected robust linear model with maximum likelihood classification (CRLMM)(166). Only calls that matched across the two algorithms were kept(167).

Sample subset: While samples were selected to be unrelated at close level, pair-wise identity-by-decent (IBD) estimates were calculated in PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/) and identified numerous related individuals(168). One individual from each related pair with an IBD>0.28 was removed. Individuals with type 2 diabetes were preferentially kept over controls. This unrelated subset was carried forward for additional genotyping, imputation and sequencing. For the sake of comparability across genetic datasets this unrelated subset has been used throughout.

Imputation to 1000 Genomes: To expand the number of loci we have genetic information on, we imputed our hard calls out to the 1000 Genomes haplotypes-- Phase 1 integrated variant set March 2012 release(137). Additional QC of genotyped data, including tests of Hardy-Weinberg disequilibrium and removal of ambiguous (A/T and G/C) variants were applied prior to imputation. This yielded a set of 643,446 scaffold variants. The Starr County Affymetrix genotypes were first phased using SHAPEIT(169), then imputed in 1 mega base sections with a 500 kilobase buffer using IMPUTE2 as part of the T2D-GENES project(170, 171).

Trait transformations: Within the unrelated subset of data, analyses were limited to individuals that were not diagnosed with type 2 diabetes prior to the time of the visit. This is due to the effects of many diabetes treatments, both medications and recommended changes in lifestyle, on weight and body fat. Residuals were calculated for each trait (BMI, WHR, waist circumference, hip circumference, arm circumference, percent body fat, and two composite measures of obesity, PC2 and PC3, at the two visits, and BMI at age 18 and maximum BMI) after adjusting for age at time of measurement and sex. For WHR, transformations were performed separately in males and females and with and without adjusting for BMI, resulting in a total of 28 outcomes. The residuals were rank-normalized to meet the assumption of normality, which is particularly important for low frequency and rare variant association tests. To obtain biologically meaningful effect sizes, analyses of the untransformed residuals were performed in parallel.

Population structure: To account for population structure, genetic principal components were calculated using a set of variants directly genotyped that were not in linkage disequilibrium with each other. The principal components were calculated with in the unrelated dataset using EIGENSTRAT(172).

Single variant analysis: Association of variants with phenotype residuals were analyzed using a linear regression model, which accounts for uncertainty in genotype imputation, implemented in SNPTEST v2.5 (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html) (173). To account for population structure, principal components were included as covariates in the model. Analyses were restricted to variants with an imputation information score of at least 0.7

and a minor allele count greater than five due to instability in models for very rare variants. Genome-wide significance was defined as $P < 5 \times 10^{-8}$, which corrects for the approximately one million independent genome wide tests(174). Lowering an order of magnitude to $P < 5 \times 10^{-7}$ was used as criteria for genome-wide suggestive associations.

Multi-trait analysis To capture variants' effects across the spectrum of adiposity traits, t-test statistics were meta-analyzed across all the traits correcting for the correlation between the test statistics using the software CPASSOC v.2(175, 176).

Gene-based analysis: Versatile Gene-based Association Study (VEGAS) was used to assess the association of variation across genes with adiposity traits(164). VEGAS, as implemented in FAST (https://bitbucket.org/baderlab/fast/wiki/Home), accounts for linkage between variants within each gene by first calculating the pairwise LD correlation matrix, then simulating multivariate normal vectors and calculating the simulated test statistic(177). Genes were defined as 20 kilobases in each direction from start and stop positions in NCBI build 37.3. Variants in the physical region with a minor allele frequency (MAF) > 0.01 and imputation info > 0.7 were included in the combined test. Additional gene-based tests including annotation information are utilized with the exome sequencing data in chapter 4.

RESULTS

Single Variant Associations: The single variant test statistics were well calibrated, after quality control measures of imputation information > 0.7 and minor allele count (MAC) > 5 were implemented. Quantile quantile (QQ) and Manhattan plots for the 24 traditional outcomes are in Figures 7 and 8, respectively. The genomic inflation factors (λ), which measures how well the *p*-values follow the expected

distribution under the null hypothesis of no association, range from 0.984 to 1.01, indicating that there is no systematic inflation of the test statistic and any population structure is adequately controlled for. Due to low sample size, the tendency towards deflated QQ plots, meaning there is less association than should be observed by chance alone, is expected.

Variants that are significantly ($P < 5 \times 10^{-8}$) or suggestively ($P < 5 \times 10^{-7}$) associated with any of these 24 traditional obesity related analyses are in Table 7. A total of 5 significant associations at 2 loci were observed. Significant results include an association with low frequency variants in and near *gamma-aminobutyric acid (GABA) A receptor, beta 2* (*GABRB2*) and decreased WHR at the second visit ($P = 1.66 \times 10^{-9}$, N = 434, β (SE) = -0.14 (0.02), MAF= 0.01, info=0.71). This association persists after adjusting for BMI. An association with increased waist circumference at the second visit and two rare variants in complete linkage disequilibrium (LD) upstream of *SH2 domain containing 4A* (*SH2D4A*) (index SNP rs180998363, $P = 1.25 \times 10^{-8}$, N = 497, β (SE) = 42 (6.1), MAF= 0.01, info=0.75).

There are also 56 suggestive association signals at 20 additional loci. A few highlights from biologically interesting genes including *lipin1* (*LPIN1)* and percent body fat at visit 1, neuropeptide Y (*NPY*) with BMI at visit 1, *polycystic kidney and hepatic disease 1* (*PKHD1)* and *neural precursor cell expressed, developmentally down-regulated 4-like E3 ubiquitin ligase* (*NEDD4L)* with WHR at visit 1, *SH3 and multiple ankyrin repeat domains2 (SHANK2)* with maximum BMI, *MACRO domain containing 2* (*MACROD2*) with arm circumference. The betas are directionally consistent across the two visits for all traits.

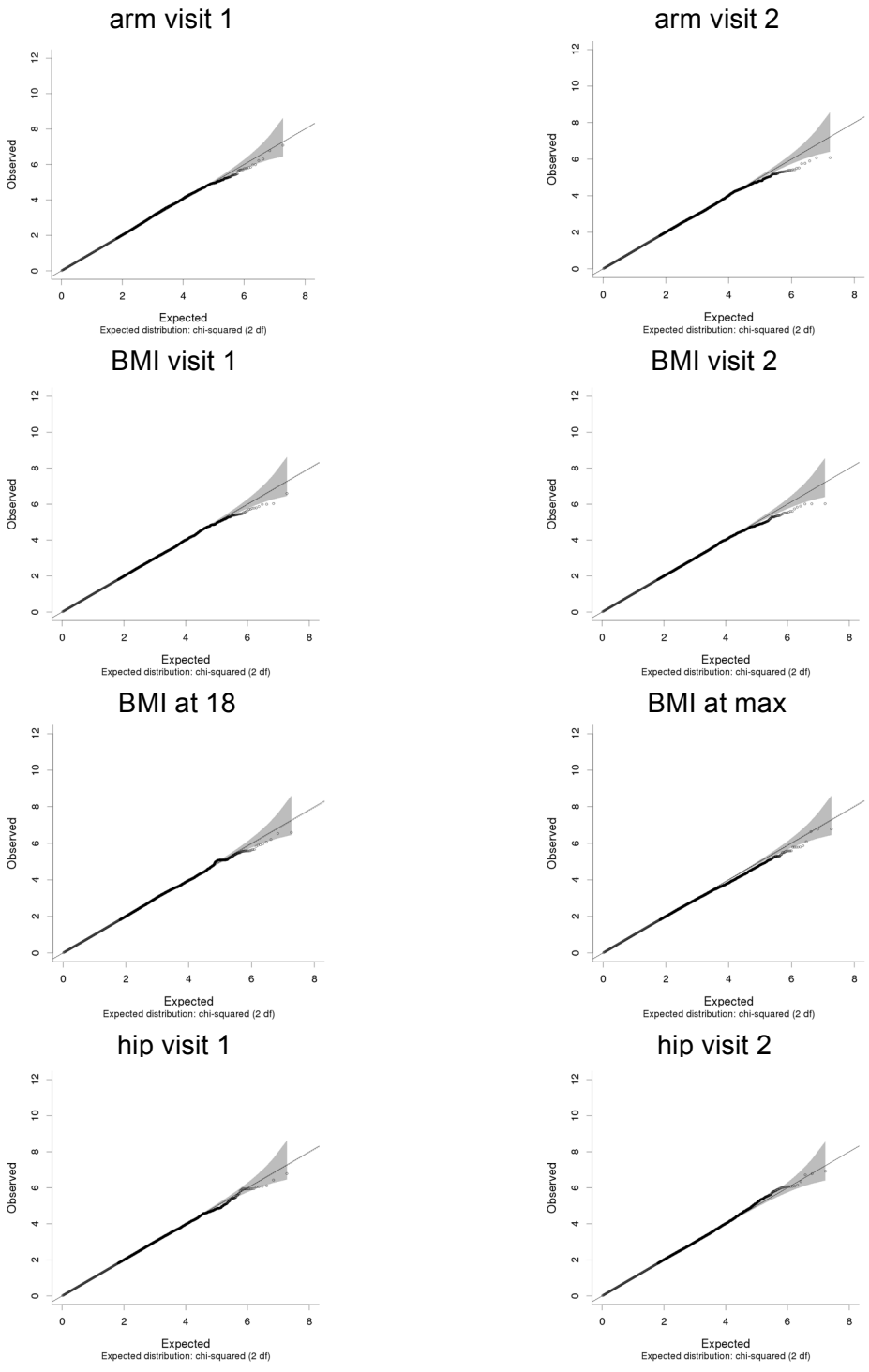


Figure 7. QQ plots for single variant association tests with imputed data.

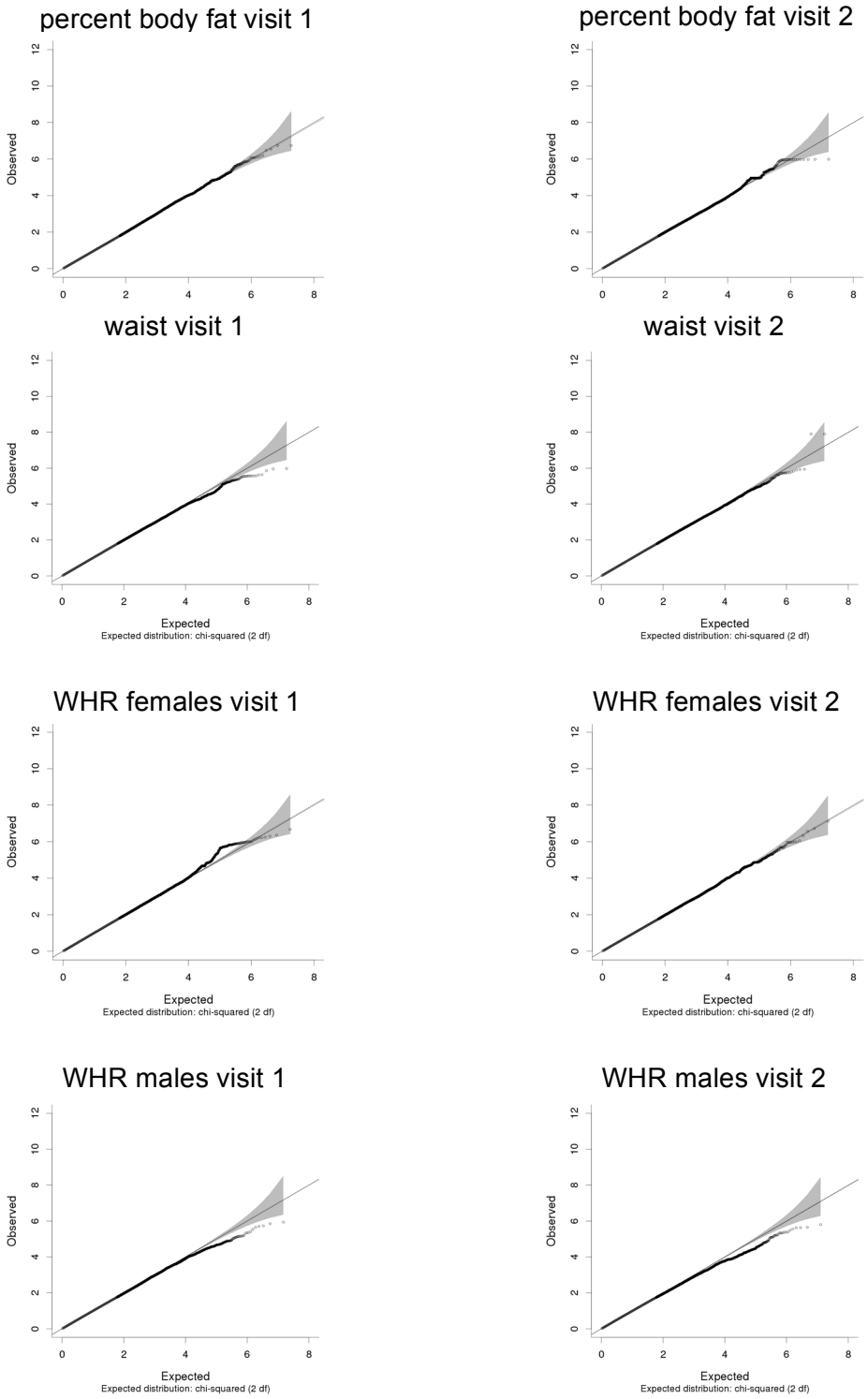


Figure 7. QQ plots for single variant association tests with imputed data.

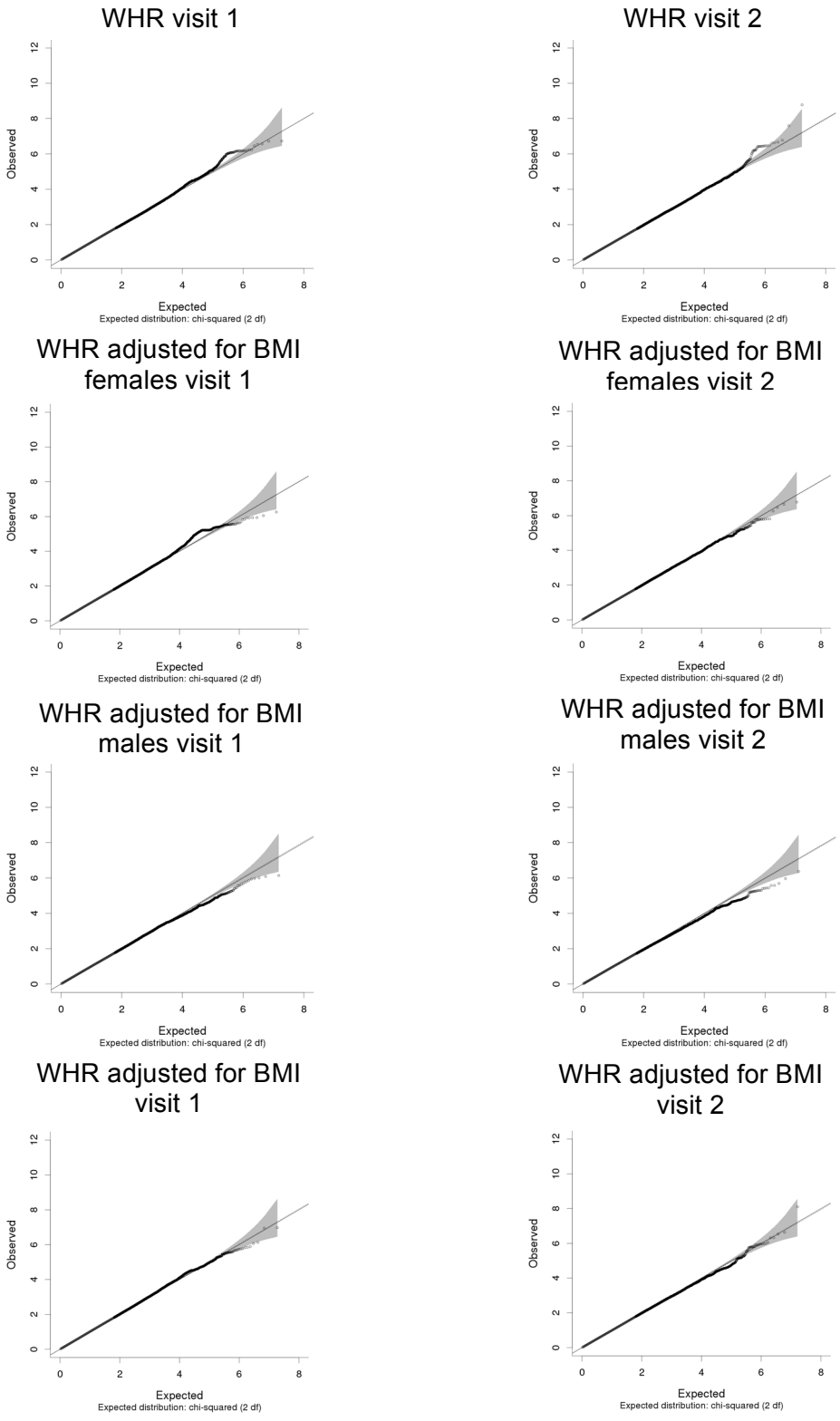


Figure 7. QQ plots for single variant association tests with imputed data.

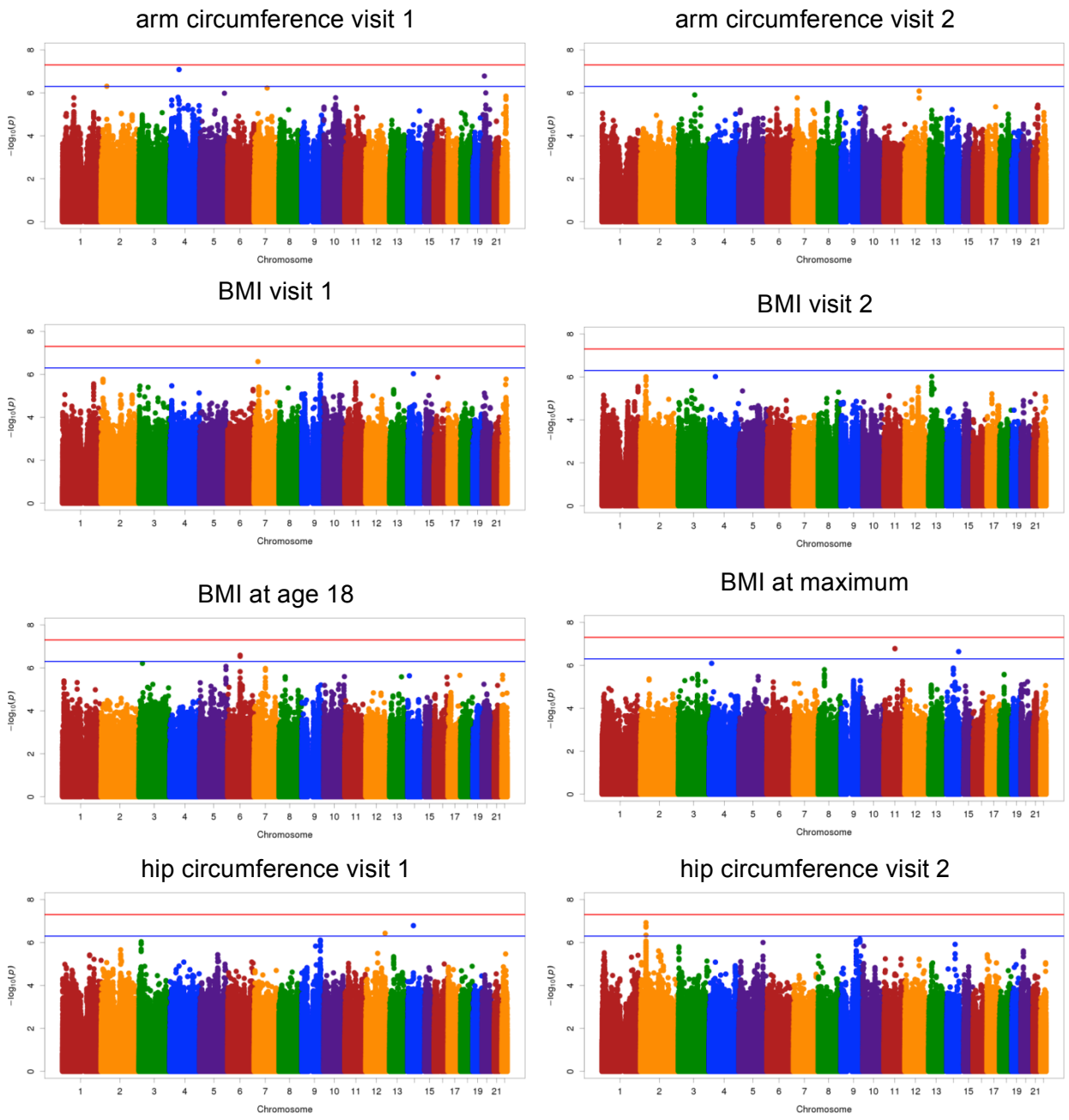


Figure 8 Single variant Manhattan plots from genome-wide imputed data

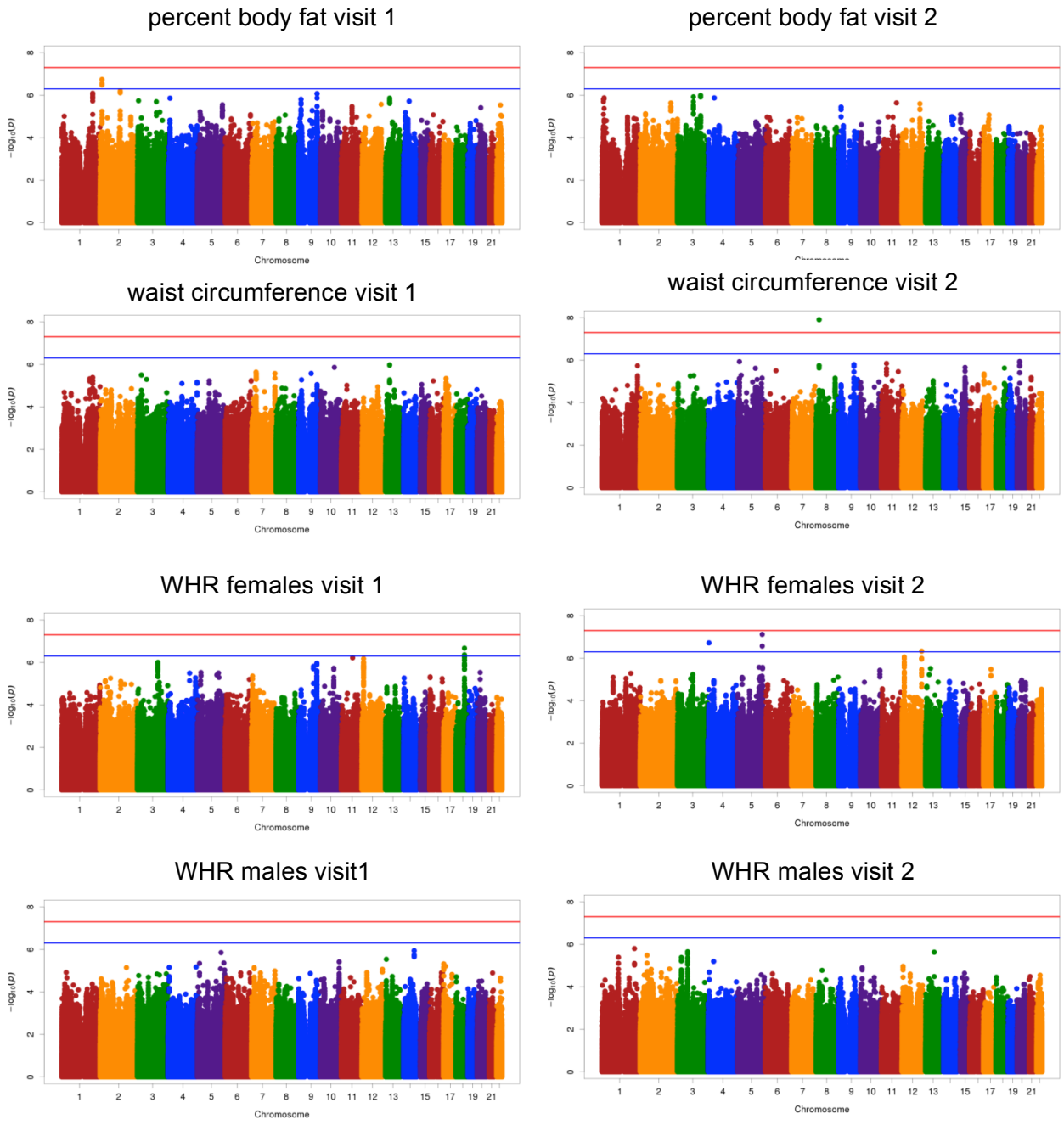


Figure 8 Single variant Manhattan plots from genome-wide imputed data

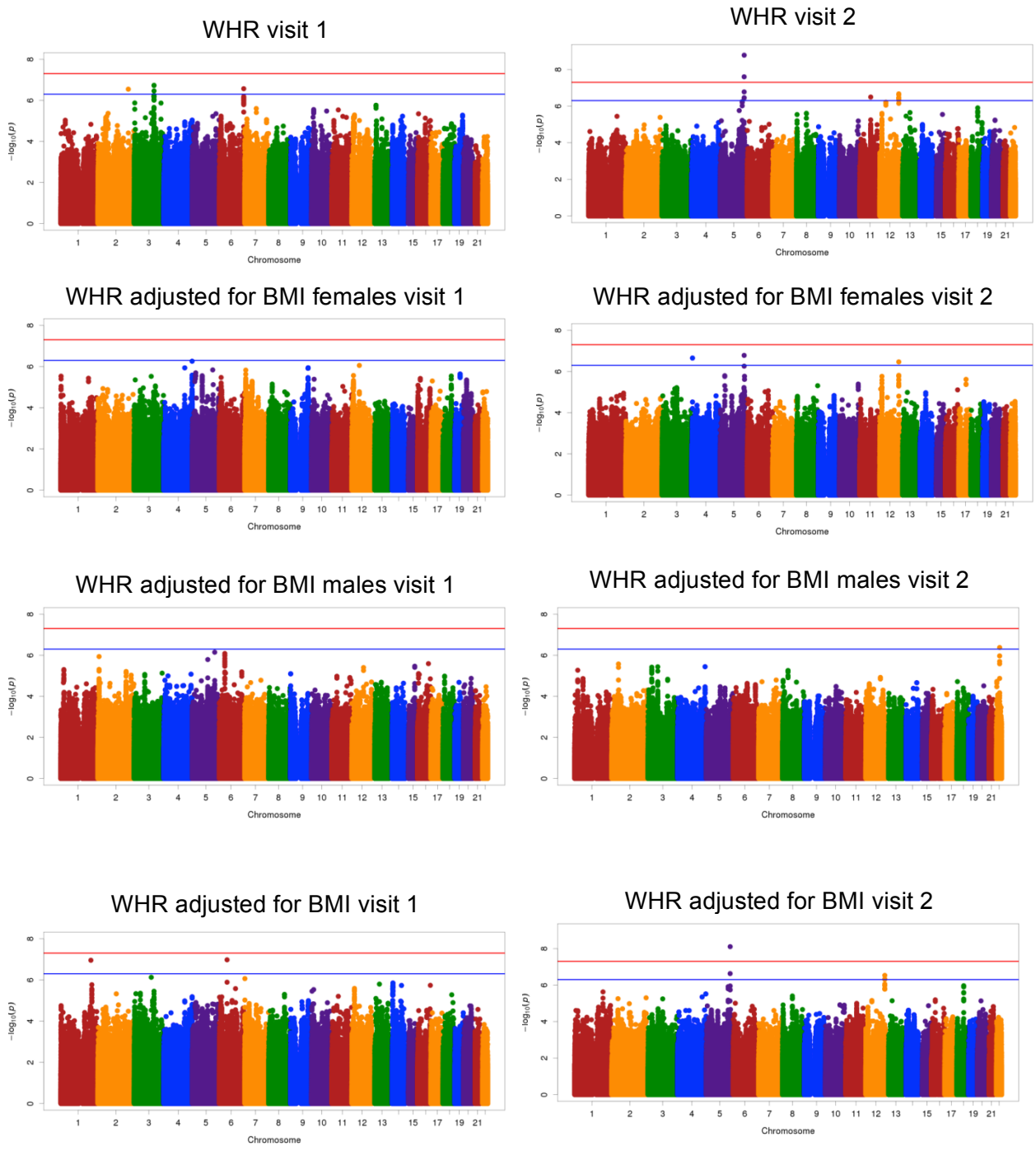


Figure 8 Single variant Manhattan plots from genome-wide imputed data

chr	position	Variant	gene	ref	alt	Trait	Visit	P	β (SE)	info	EAC	N	MAF
1	202021786	rs74862838	*ELF3*(dist=35471), *GPR37L1*(dist=70243)	T	G	WHR adj BMI	v1	1.10×10^{-7}	-0.036 (0.0073)	0.73	84	818	0.051
							v2	0.0039	-0.029 (0.0098)	0.72	50	434	0.057
2	11823403	rs75618981	*LPIN1*	G	A	percent body fat	v1	1.80×10^{-7}	0.087 (0.02)	0.87	20	824	0.012
							v2	0.22	0.042 (0.027)	0.97	10	410	0.012
2	11827803	rs191122460	*LPIN1*	C	A	percent body fat	v1	2.80×10^{-7}	0.086 (0.019)	0.93	20	824	0.012
							v2	0.21	0.042 (0.027)	0.99	10	410	0.012
2	11870002	rs113485904	*LPIN1*	G	A	percent body fat	v1	3.30×10^{-7}	0.084 (0.019)	0.97	20	824	0.012
							v2	0.21	0.042 (0.027)	0.99	10	410	0.012
2	11883170	rs112863316	*LPIN1*	T	C	percent body fat	v1	1.80×10^{-7}	0.086 (0.019)	0.91	21	824	0.013
							v2	0.31	0.038 (0.026)	0.84	11	410	0.014
2	34707113	rs6543845	*LINC01317*(dist=184300), *LINC01320*(dist=195511)	A	G	hip	v1	0.014	2 (0.82)	0.75	1267	821	0.23
							v2	1.90×10^{-7}	5.2 (1)	0.78	760	496	0.23
2	34721460	rs7557071	*LINC01317*(dist=198647), *LINC01320*(dist=181164)	A	C	hips	v1	0.015	2 (0.82)	0.75	1271	821	0.23
							v2	1.20×10^{-7}	5.3 (1.1)	0.77	763	496	0.23
2	34722105	chr2:34722105:D	*LINC01317*(dist=199292), *LINC01320*(dist=180519)	TC	T	hip	v1	0.024	1.8 (0.81)	0.75	1243	821	0.24
							v2	4.50×10^{-7}	5 (1)	0.77	748	496	0.25
2	34725422	rs13414227	*LINC01317*(dist=202609), *LINC01320*(dist=177202)	G	A	hip	v1	0.011	2.1 (0.81)	0.76	1263	821	0.23
							v2	1.70×10^{-7}	5.2 (1)	0.77	757	496	0.24
2	36152014	chr2:36152014:I	*LINC01320*(dist=1204384), *LOC100288911*(dist=429878)	T	TA	arm	v1	4.90×10^{-7}	-1.3 (0.27)	0.82	528	824	0.32
							v2	0.0023	-1.1 (0.35)	0.81	326	497	0.33
2	203424027	rs75023458	BMPR2	A	C	WHR	v1	2.90×10^{-7}	-0.02 (0.0037)	0.73	1031	818	0.37
							v2	0.0015	-0.017 (0.0051)	0.69	554	434	0.36
3	132406895	rs76967933	NPHP3-ACAD11	C	A	WHR	v1	1.90×10^{-7}	-0.035 (0.0069)	0.87	103	818	0.063
							v2	0.011	-0.022 (0.0081)	0.89	64	434	0.073
3	132415424	rs76177059	*NPHP3-ACAD11*	T	C	WHR	v1	1.90×10^{-7}	-0.036 (0.007)	0.88	99	818	0.06
							v2	0.014	-0.022 (0.0083)	0.89	61	434	0.07
3	132421998	rs79328618	*NPHP3-ACAD11*	G	A	WHR	v1	3.50×10^{-7}	-0.035 (0.0069)	0.9	99	818	0.06
							v2	0.012	-0.022 (0.0081)	0.91	61	434	0.07

chr	position	Variant	gene	ref	alt	Trait	Visit	P	β (SE)	info	EAC	N	MAF
4	5844364	rs34613066	*CRMP1*	C	T	WHR females	v1	0.2	0.014 (0.012)	0.8	42	579	0.036
							v2	1.90×10^{-7}	0.069 (0.015)	0.76	24	330	0.036
4	5844364	rs34613066	*CRMP1*	C	T	WHR adj BMI females	v1	0.11	0.016 (0.011)	0.81	42	579	0.036
							v2	2.20×10^{-7}	0.068 (0.014)	0.76	24	330	0.036
4	59573143	rs142866966	*LOC101928851*(dist=1240991), NONE(dist=NONE)	A	G	arm	v1	8.20×10^{-8}	5.2 (0.84)	0.71	48	824	0.029
							v2	4.00×10^{-5}	4.8 (1)	0.71	34	497	0.034
5	160597595	rs113244407	*LOC285629*(dist=231962), *GABRB2*(dist=117841)	T	G	WHR	v1	0.29	-0.0081 (0.014)	0.88	22	818	0.013
							v2	2.50×10^{-8}	-0.11 (0.019)	0.8	12	434	0.013
5	160597595	rs113244407	*LOC285629*(dist=231962), *GABRB2*(dist=117841)	T	G	WHR females	v1	0.45	-0.0038 (0.016)	0.89	18	579	0.015
							v2	7.60×10^{-8}	-0.12 (0.021)	0.84	10	330	0.015
5	160597595	rs113244407	*LOC285629*(dist=231962), *GABRB2*(dist=117841)	T	G	WHR adj BMI females	v1	0.65	-0.00013 (0.014)	0.9	18	579	0.015
							v2	1.70×10^{-7}	-0.11 (0.021)	0.84	10	330	0.015
5	160597595	rs113244407	*LOC285629*(dist=231962), *GABRB2*(dist=117841)	T	G	WHR adj BMI	v1	0.57	-0.0018 (0.012)	0.9	22	818	0.013
							v2	2.30×10^{-7}	-0.098 (0.018)	0.85	12	434	0.013
5	160689360	rs17456567	*LOC285629*(dist=323727), *GABRB2*(dist=26076)	G	A	WHR	v1	0.32	-0.0075 (0.014)	0.84	23	818	0.014
							v2	1.70×10^{-7}	-0.1 (0.019)	0.8	12	434	0.014
5	160689360	rs17456567	*LOC285629*(dist=323727), *GABRB2*(dist=26076)	G	A	WHR females	v1	0.49	-0.0031 (0.016)	0.85	19	579	0.016
							v2	2.70×10^{-7}	-0.11 (0.021)	0.82	10	330	0.015
5	160822038	rs112490216	*GABRB2*	G	C	WHR	v1	0.29	-0.013 (0.013)	0.83	28	818	0.017
							v2	3.60×10^{-7}	-0.11 (0.02)	0.74	13	434	0.015
5	160831588	rs111330620	*GABRB2*	A	G	WHR	v1	0.29	-0.013 (0.013)	0.83	28	818	0.017
							v2	3.50×10^{-7}	-0.11 (0.02)	0.74	13	434	0.015
5	160880879	rs150769823	*GABRB2*	C	T	WHR	v1	0.16	-0.021 (0.015)	0.88	19	818	0.011
							v2	1.70×10^{-9}	-0.14 (0.023)	0.72	10	434	0.012
5	160880879	rs150769823	*GABRB2*	C	T	WHR adj BMI	v1	0.3	-0.015 (0.013)	0.91	19	818	0.011
							v2	7.80×10^{-9}	-0.14 (0.022)	0.71	10	434	0.012
6	51934092	rs17638535	*PKHD1*	G	A	WHR adj BMI	v1	1.10×10^{-7}	-0.14 (0.028)	0.7	5	818	0.0032
							v2	0.026	-0.062 (0.031)	0.85	4	434	0.0044

chr	position	Variant	gene	ref	alt	Trait	Visit	P	β (SE)	info	EAC	N	MAF
6	80291793	rs143783440	*LCA5*(dist=44646), *SH3BGRL2*(dist=49207)	G	A	BMI at 18	v1	2.60×10^{-7}	-1.4 (0.32)	0.85	362	825	0.22
							v2						
6	80307108	chr6:80307108:D	*LCA5*(dist=59961), *SH3BGRL2*(dist=33892)	GAGGA	G	BMI at 18	v1	2.90×10^{-7}	-1.5 (0.33)	0.91	276	825	0.17
							v2						
6	163490679	chr6:163490679:D	*PACRG*	GCA	G	WHR	v1	2.70×10^{-7}	0.016 (0.0033)	0.95	756	818	0.46
							v2	0.013	0.01 (0.0042)	0.94	395	434	0.45
7	24093468	rs116467353	*STK31*(dist=221338), *NPY*(dist=230339)	C	T	BMI	v1	2.50×10^{-7}	-7.9 (2.1)	0.71	11	825	0.0064
							v2	0.061	-5.1 (3.3)	0.82	4	438	0.0051
8	19170363	rs180998363	*SH2D4A*	A	G	waist	v1	0.0099	14 (3.8)	0.84	19	822	0.011
							v2	1.20×10^{-8}	42 (6.1)	0.75	8	497	0.0085
8	19170364	rs185529143	*SH2D4A*	G	C	waist	v1	0.0099	14 (3.8)	0.84	19	822	0.011
							v2	1.20×10^{-8}	42 (6.1)	0.75	8	497	0.0085
11	71042064	rs71473821	*SHANK2*(dist=106222), *FLJ42102*(dist=74728)	G	A	BMI at max	v1	1.70×10^{-7}	-10 (2.4)	0.99	10	809	0.0062
							v2						
11	71057262	rs34212157	*SHANK2*(dist=121420), *FLJ42102*(dist=59530)	C	T	BMI at max	v1	1.70×10^{-7}	-10 (2.4)	0.99	10	809	0.0062
							v2						
11	74066524	rs61902400	*PGM2L1*	G	A	WHR	v1	0.021	0.0076 (0.0035)	0.79	785	818	0.48
							v2	3.20×10^{-7}	0.022 (0.0046)	0.76	413	434	0.48
12	124021547	rs74924067	*MIR3908*	C	T	hip	v1	3.70×10^{-7}	-6 (1.3)	0.73	150	821	0.091
							v2	0.00076	-5.4 (1.7)	0.71	88	496	0.089
12	127088883	rs1541486	*LOC100128554*(dist=131552), *LOC100996671*(dist=48609)	C	T	WHR	v1	0.032	0.0077 (0.0037)	0.85	493	818	0.3
							v2	2.10×10^{-7}	0.024 (0.0048)	0.86	256	434	0.3
12	127088883	rs1541486	*LOC100128554*(dist=131552), *LOC100996671*(dist=48609)	C	T	WHR adj BMI	v1	0.014	0.0078 (0.0033)	0.86	493	818	0.3
							v2	2.90×10^{-7}	0.023 (0.0046)	0.86	256	434	0.3
12	127088883	rs1541486	*LOC100128554*(dist=131552), *LOC100996671*(dist=48609)	C	T	WHR adj BMI females	v1	0.0082	0.01 (0.0042)	0.86	357	579	0.31
							v2	3.40×10^{-7}	0.027 (0.0056)	0.85	195	330	0.3
12	127088883	rs1541486	*LOC100128554*(dist=131552), *LOC100996671*(dist=48609)	C	T	WHR females	v1	0.022	0.01 (0.0046)	0.85	357	579	0.31
							v2	4.70×10^{-7}	0.027 (0.0057)	0.85	195	330	0.3

chr	position	Variant	gene	ref	alt	Trait	Visit	P	β (SE)	info	EAC	N	MAF
12	127092972	rs4765418	*LOC100128554*(dist=135641), *LOC100996671*(dist=44520)	T	G	WHR	v1	0.11	0.0054 (0.0035)	0.97	511	818	0.31
							v2	4.80×10^{-7}	0.021 (0.0045)	0.98	263	434	0.3
12	127093438	chr12:127093438	*LOC100128554*(dist=136107), *LOC100996671*(dist=44054)	C	T	WHR	v1	0.078	0.0059 (0.0034)	1	493	818	0.3
							v2	3.60×10^{-7}	0.021 (0.0044)	1	254	434	0.29
12	127094364	chr12:127094364:D	*LOC100128554*(dist=137033), *LOC100996671*(dist=43128)	CTT	C	WHR	v1	0.19	0.0046 (0.0037)	1	368	818	0.23
							v2	3.80×10^{-7}	0.023 (0.0048)	1	193	434	0.22
12	127094542	rs11058745	*LOC100128554*(dist=137211), *LOC100996671*(dist=42950)	T	C	WHR	v1	0.077	0.006 (0.0035)	0.98	502	818	0.31
							v2	2.30×10^{-7}	0.022 (0.0044)	0.99	259	434	0.3
12	127094542	rs11058745	*LOC100128554*(dist=137211), *LOC100996671*(dist=42950)	T	C	WHR adj BMI	v1	0.059	0.0056 (0.0031)	0.97	502	818	0.31
							v2	4.70×10^{-7}	0.02 (0.0042)	0.98	259	434	0.3
12	127094630	rs11058746	*LOC100128554*(dist=137299), *LOC100996671*(dist=42862)	T	G	WHR	v1	0.073	0.006 (0.0034)	0.99	492	818	0.3
							v2	4.10×10^{-7}	0.021 (0.0044)	0.99	254	434	0.29
12	127094761	rs11058747	*LOC100128554*(dist=137430), *LOC100996671*(dist=42731)	A	C	WHR	v1	0.074	0.006 (0.0035)	0.99	499	818	0.31
							v2	2.40×10^{-7}	0.022 (0.0044)	0.99	258	434	0.3
12	127094761	rs11058747	*LOC100128554*(dist=137430), *LOC100996671*(dist=42731)	A	C	WHR adj BMI	v1	0.05	0.0057 (0.0031)	0.98	499	818	0.31
							v2	4.90×10^{-7}	0.02 (0.0042)	0.99	258	434	0.3
12	127094795	rs11058748	*LOC100128554*(dist=137464), *LOC100996671*(dist=42697)	A	G	WHR	v1	0.19	0.0045 (0.0037)	1	368	818	0.22
							v2	3.70×10^{-7}	0.024 (0.0048)	1	193	434	0.22
12	127094931	rs79856323	*LOC100128554*(dist=137600), *LOC100996671*(dist=42561)	C	A	WHR	v1	0.19	0.0046 (0.0037)	1	368	818	0.23
							v2	3.70×10^{-7}	0.023 (0.0048)	1	193	434	0.22
12	127095200	chr12:127095200	*LOC100128554*(dist=137869), *LOC100996671*(dist=42292)	G	A	WHR	v1	0.19	0.0045 (0.0037)	1	368	818	0.22
							v2	3.70×10^{-7}	0.023 (0.0048)	1	193	434	0.22
12	127096227	rs11058749	*LOC100128554*(dist=138896), *LOC100996671*(dist=41265)	C	G	WHR	v1	0.19	0.0046 (0.0037)	0.99	367	818	0.22
							v2	3.50×10^{-7}	0.024 (0.0049)	0.99	192	434	0.22
12	127096299	rs10847184	*LOC100128554*(dist=138968), *LOC100996671*(dist=41193)	A	T	WHR	v1	0.19	0.0046 (0.0037)	0.99	367	818	0.22
							v2	3.50×10^{-7}	0.024 (0.0049)	0.99	192	434	0.22
14	57116309	rs142932442	*TMEM260*	C	G	hip	v1	1.60×10^{-7}	-12 (2.6)	0.78	27	821	0.017
							v2	0.0024	-9 (3.7)	0.78	14	496	0.014

chr	position	Variant	gene	ref	alt	Trait	Visit	P	β (SE)	info	EAC	N	MAF
14	97241615	rs117881117	*PAPOLA*(dist=208162), *VRK1*(dist=22069)	C	T	BMI at max	v1	2.30×10^{-7}	-11 (2.8)	0.82	8	809	0.0049
							v2						
18	55589140	rs9948233	*ATP8B1*(dist=118813), *NEDD4L*(dist=122470)	T	C	WHR females	v1	2.10×10^{-7}	-0.026 (0.005)	0.9	266	579	0.23
							v2	0.037	-0.011 (0.0056)	0.94	164	330	0.25
18	55601636	rs4941202	*ATP8B1*(dist=131309), *NEDD4L*(dist=109974)	T	C	WHR females	v1	4.40×10^{-7}	-0.023 (0.0046)	0.96	306	579	0.26
							v2	0.3	-0.0059 (0.0054)	0.95	175	330	0.27
20	15619338	chr20:15619338	*MACROD2*	A	C	arm	v1	1.70×10^{-7}	1.6 (0.31)	1	267	824	0.16
							v2	0.089	0.76 (0.41)	1	155	497	0.16
22	38158273	rs73168260	*TRIOBP*	A	G	WHR adj BMI males	v1	0.012	-0.012 (0.005)	0.98	95	239	0.2
							v2	4.20×10^{-7}	-0.04 (0.0081)	0.95	44	104	0.21

Table 7. Single Variant results for imputed data. All significant and suggestive single variant associations signals are shown in order of physical position. Associations are shown for both visits. The betas are all directionally consistent across visits.

Composite Measures of Adiposity: In addition to the analysis of traditional measures of obesity, associations with the composite measures of obesity, PC2 capturing central adiposity above or below the waist, and PC3, which captures truncal versus peripheral adiposity, were also analyzed. QQ and Manhattan plots are in Figures 9 and 10; the top results are in Table 8. Multiple variants on chromosome 16 are genome-wide significantly associated with PC2 at the second (index SNP rs2244324, $P = 2.81 \times 10^{-8}$, N = 408, β (SE) = -0.17 (0.03), MAF= 0.30, info=0.98). There is also a suggestive association between PC3 at visit 2 and a variant in *Down syndrome cell adhesion molecule* (*DSCAM)*. These were the only significant or suggestive results.

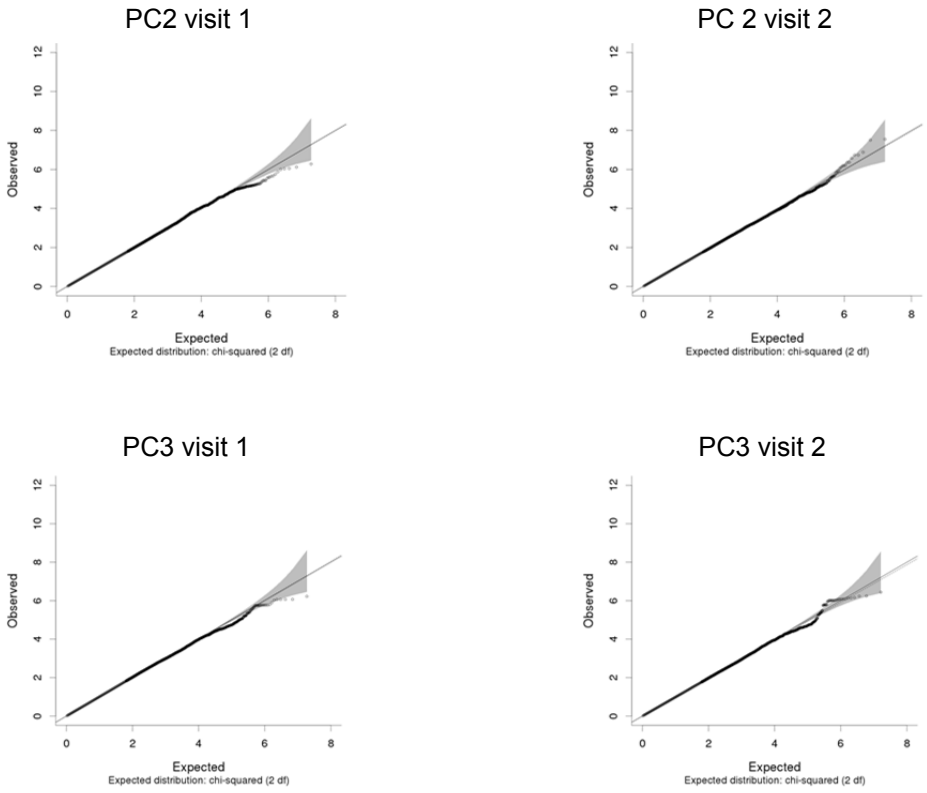


Figure 9. QQ plots for single variant association tests with composite measures of adiposity.

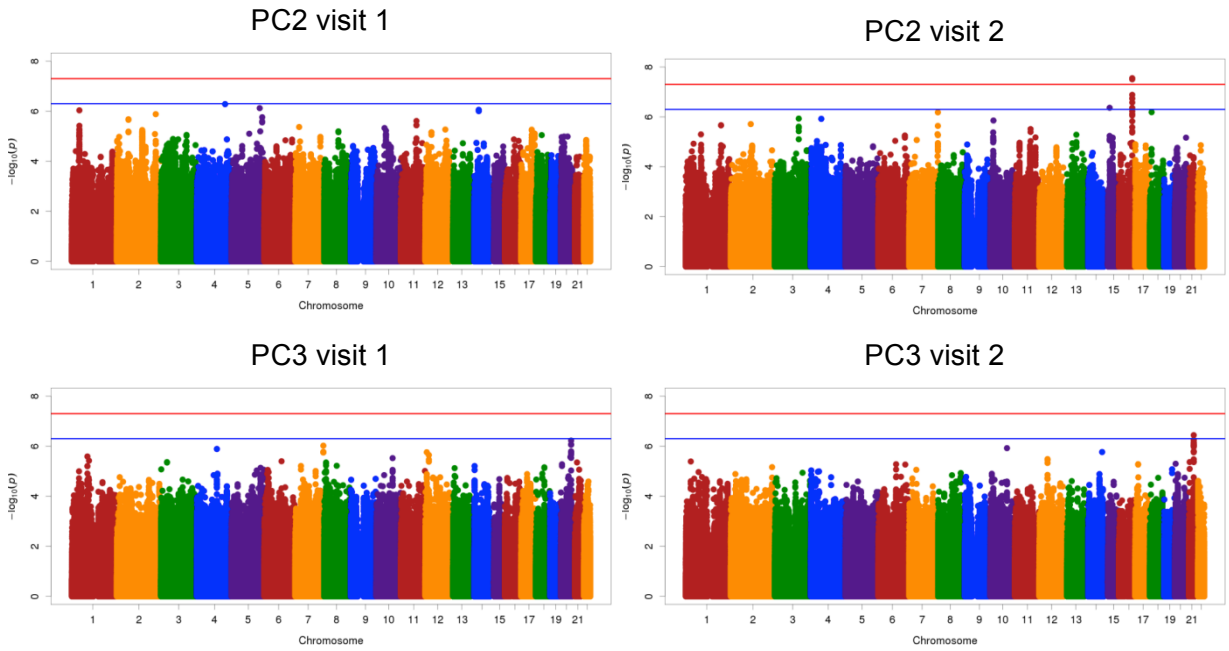


Figure 10. Manhattan plots for single variant association tests with composite measures of adiposity.

chr	position	Variant	gene	ref	alt	Trait	Visit	P	β (SE)	info	EAC	N	MAF
15	31316965	rs28697571	*TRPM1*	G	T	PC2	v1	0.008	-0.13 (0.049)	0.95	59	816	0.036
							v2	4.30×10^{-7}	0.34 (0.066)	0.96	36	408	0.044
16	73832185	rs410109	*LINC01568*(dist=376890), *LOC101928035*(dist=394106)	G	A	PC2	v1	0.69	-0.0082 (0.02)	1	474	816	0.29
							v2	1.80×10^{-7}	-0.16 (0.03)	0.99	247	408	0.3
16	73832216	rs421566	*LINC01568*(dist=376921), *LOC101928035*(dist=394075)	A	T	PC2	v1	0.69	-0.0082 (0.02)	1	474	816	0.29
							v2	1.80×10^{-7}	-0.16 (0.03)	0.99	247	408	0.3
16	73832579	rs427478	*LINC01568*(dist=377284), *LOC101928035*(dist=393712)	C	T	PC2	v1	0.59	-0.011 (0.02)	1	469	816	0.29
							v2	4.40×10^{-7}	-0.15 (0.03)	0.99	244	408	0.3
16	73832603	rs385183	*LINC01568*(dist=377308), *LOC101928035*(dist=393688)	C	A	PC2	v1	0.82	-0.0048 (0.02)	0.99	483	816	0.3
							v2	1.30×10^{-7}	-0.16 (0.03)	0.99	248	408	0.3
16	73836905	rs59844697	*LINC01568*(dist=381610), *LOC101928035*(dist=389386)	G	T	PC2	v1	1	-0.00015 (0.021)	0.99	403	816	0.25
							v2	3.10×10^{-8}	-0.18 (0.031)	0.98	211	408	0.26
16	73837290	rs328392	*LINC01568*(dist=381995), *LOC101928035*(dist=389001)	C	T	PC2	v1	0.67	-0.009 (0.02)	0.98	477	816	0.29
							v2	2.70×10^{-7}	-0.15 (0.03)	0.98	252	408	0.31
16	73847624	rs2244324	*LINC01568*(dist=392329), *LOC101928035*(dist=378667)	T	A	PC2	v1	0.76	0.0073 (0.02)	0.98	485	816	0.3
							v2	2.80×10^{-8}	-0.17 (0.03)	0.98	246	408	0.3
21	41463520	rs74762297	*DSCAM*	T	C	PC3	v1	0.0083	0.29 (0.12)	0.87	20	816	0.012
							v2	3.60×10^{-7}	-1 (0.19)	0.78	9	408	0.01

Table 8. Single variant association tests with composite measures of adiposity.

Multi-trait Analysis: Using the composite measures of obesity is useful for capturing adiposity distribution. To capture variants effects across the spectrum of adiposity traits, t-test statistics were meta-analyzed across all traditional measures of obesity at both time points while correcting for the correlation between the test statistics using the software CPASSOC(175, 176). The resulting QQ and Manhattan plots are in Figures 11 and 12; the QQ plots show a minimal inflation in the combined test statistic. While no variant reached genome-wide significance, there are 9 variants in 4 loci that are genome-wide suggestive which are shown in Table 9. These consist of low-frequency variants on chromosome 4, in the gene *sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3C* (*SEMA3C)* and in the gene *early B-cell factor 4 (EBF4),* and a common variant between the genes *growth arrest-specific 7 (GAS7)* and *myosin, heavy chain 13, skeletal muscle* (*MYH13).* The association between these variants and traditional measures of obesity are all sub-significant. The most convincing of these signals is the variant between *GAS7* and *MYH13*, which is nominally significant ($p < 0.05$) in 7 of the 16 traits.

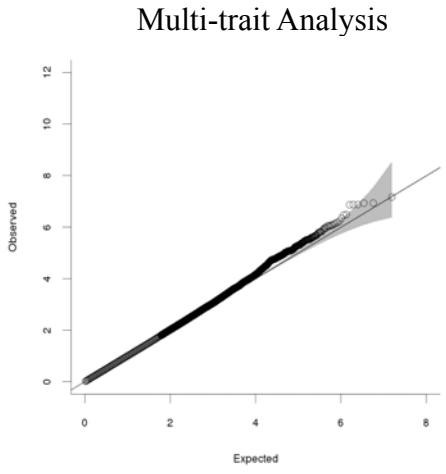


Figure 11 QQ plot of single variant multi-trait analysis.

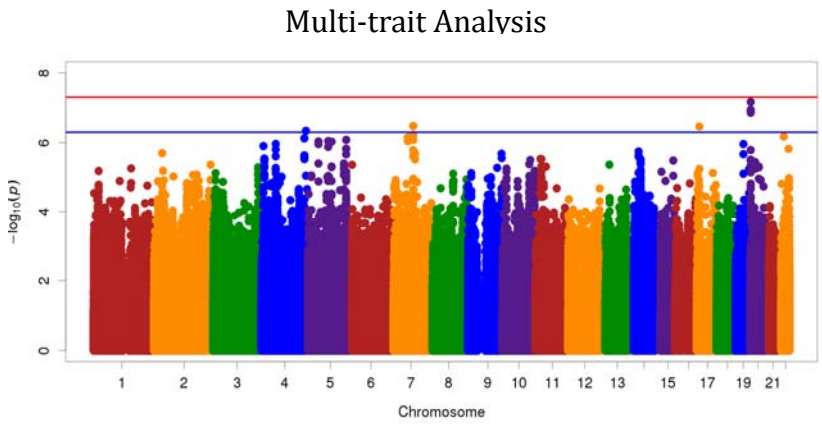


Figure 12. Manhattan plot of single variant multi-trait analysis.

chrom	position	Variant	Gene	ref	alt	P-value	MAF
4	182524751	chr4:182524751:D	*LINC00290*(dist=444449), *LOC90768*(dist=535062)	TA	T	4.55×10^{-7}	0.0098
7	80503452	rs58283863	*SEMA3C*	A	G	3.29×10^{-7}	0.0049
17	10105523	rs77233190	*GAS7*(dist=3655), *MYH13*(dist=98660)	A	G	3.42×10^{-7}	0.2488
20	2688091	rs79712868	*EBF4*	C	T	6.91×10^{-8}	0.0165
20	2691964	rs8121831	*EBF4*	A	G	1.17×10^{-7}	0.0170
20	2696131	rs80151839	*EBF4*	G	A	1.18×10^{-7}	0.0170
20	2699607	rs80023036	*EBF4*	A	G	1.37×10^{-7}	0.0173
20	2705316	rs79049619	*EBF4*	C	T	1.34×10^{-7}	0.0171
20	2705474	rs112916871	*EBF4*	C	T	1.34×10^{-7}	0.0171

Table 9. Multi-trait single variant analysis top results.

Gene based Associations: The QQ plots for the gene-based test VEGAS are systematically deflated, with genomic inflation factors ranging from 0.94 to 0.99. In the QQ plots (Figure 13) there are small shelves instead of a smooth linear line. This reflects rounding, as the number of simulations performed with in the VEGAS association test is increased only if required to establish significance. In addition to this, deflation reflects limited power due to sample size.

Using a genome-wide significance cut off of $p < 2.5 \times 10^{-6}$, which is the equivalent of correcting for the approximately 20,000 genes in the human genome, one

gene across all the traits is genome-wide significant. *C8orf4* is associated with BMI at maximum weight. Looking down an order of magnitude to 2.5×10^{-5}, an additional 15 loci (16 genes) are associated with at least one trait, as shown in Table 10. This includes both *Oxidized Low Density Lipoprotein (Lectin-Like) Receptor 1* (*OLR1*) and *transmembrane protein 52B* (*TMEM52B*) on chromosome 12 which are associated with multiple waist to hip ratio outcomes at visit 1. These two gene-based tests are overlapping with each other. Similarly, two transcripts associated with percent body fat at visit 2, *ring finger protein 13* (*RNF13*) and *LOC100422259*, represent also overlapping regions on chromosome 3.

Gene-based Associations of Composite Measures: In addition to looking at the traditional measures of obesity, we also tested for association between transcripts and the composite measures of adiposity, PC2, which captures adiposity above or below the waist, and PC3, which captures central versus peripheral adiposity. Only three transcripts are suggestively associated with PC3 at visit 1 (Figures 15 and 16, Table 11). Two of the transcripts are physically overlapping and both represent portions of the T cell receptor alpha variable region (*TRAV12-3* and *TRAV8-6*).

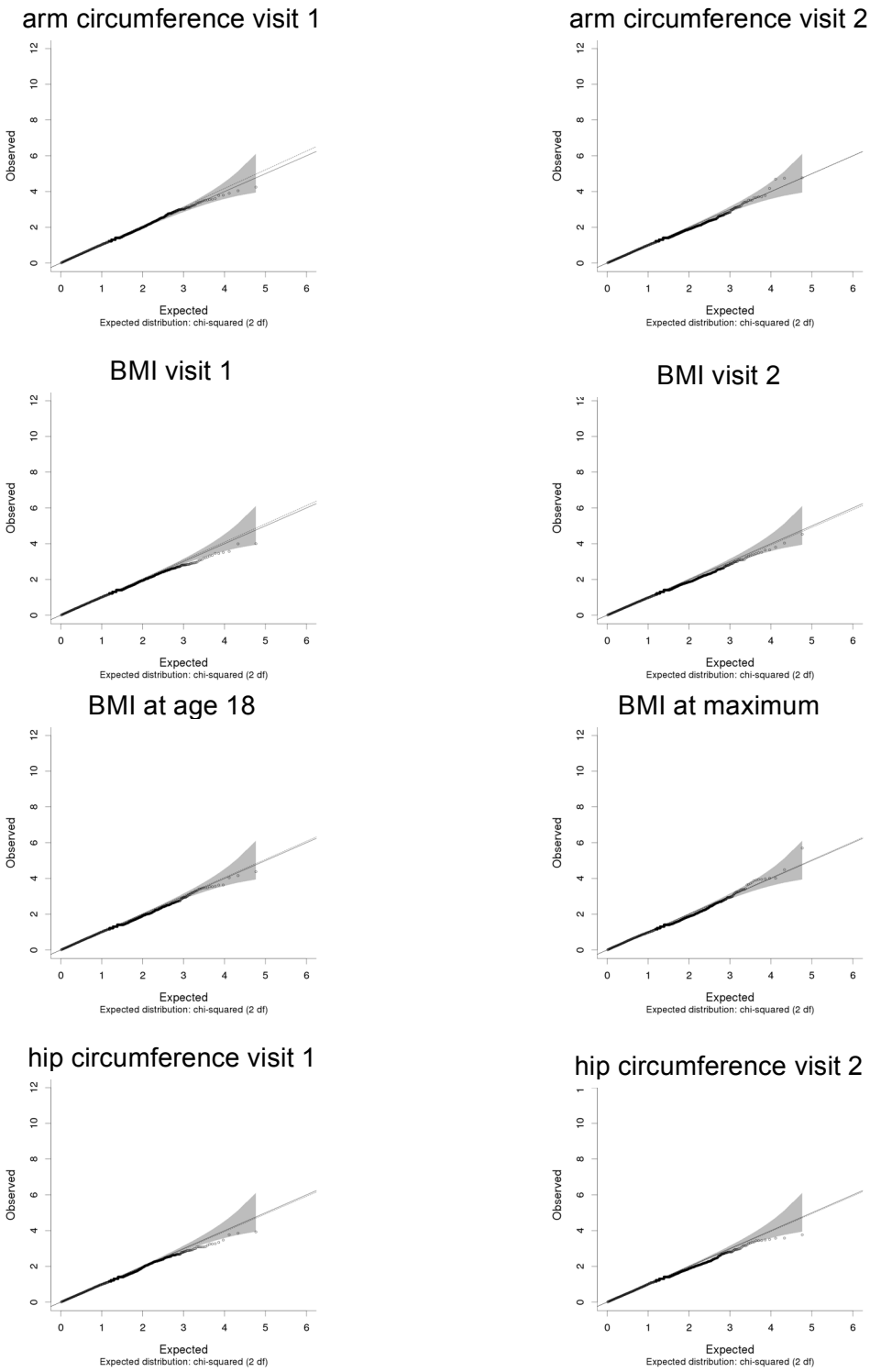


Figure 13. QQ plots for common variant gene-based association tests.

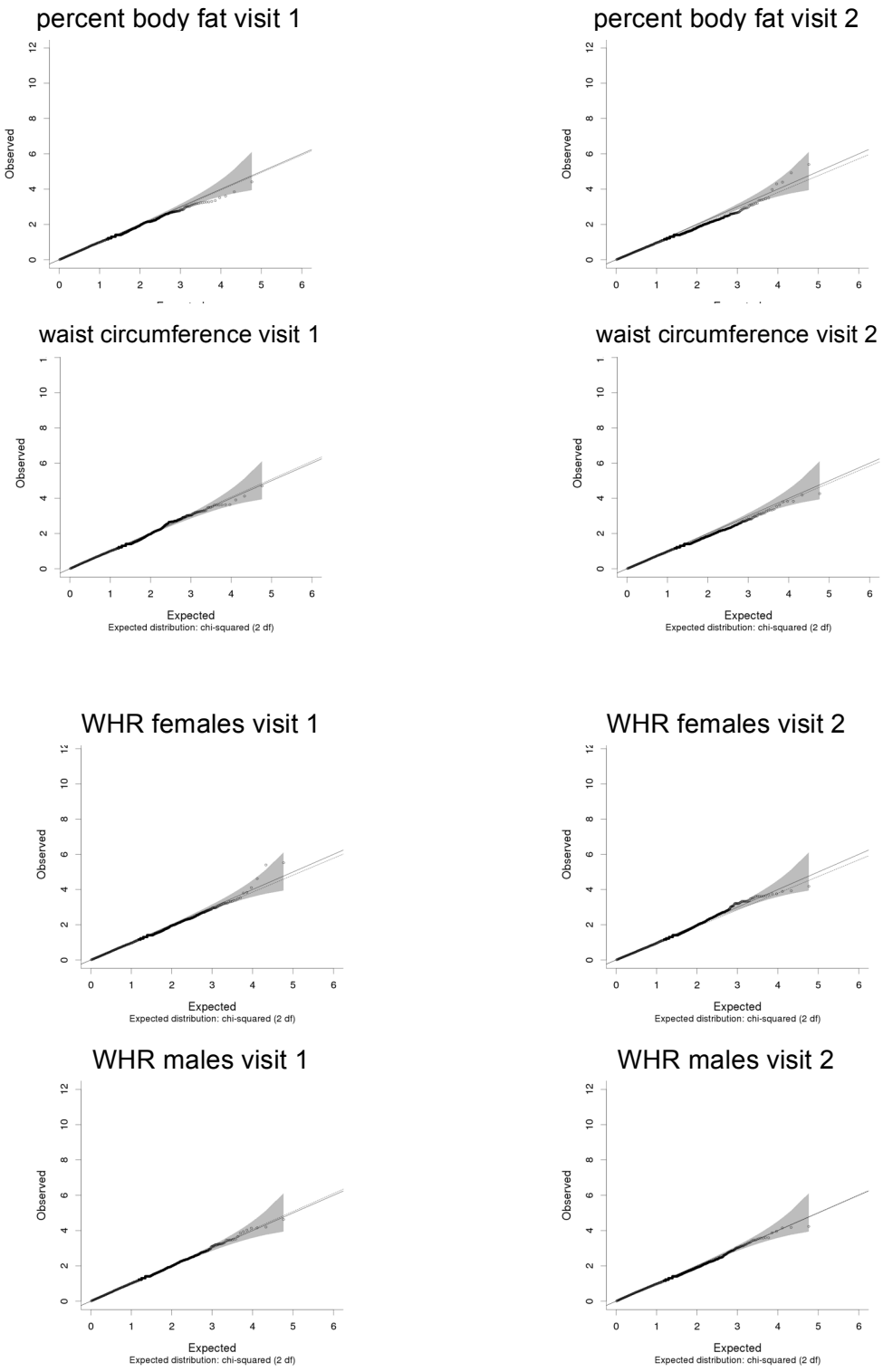


Figure 13. QQ plots for common variant gene-based association tests.

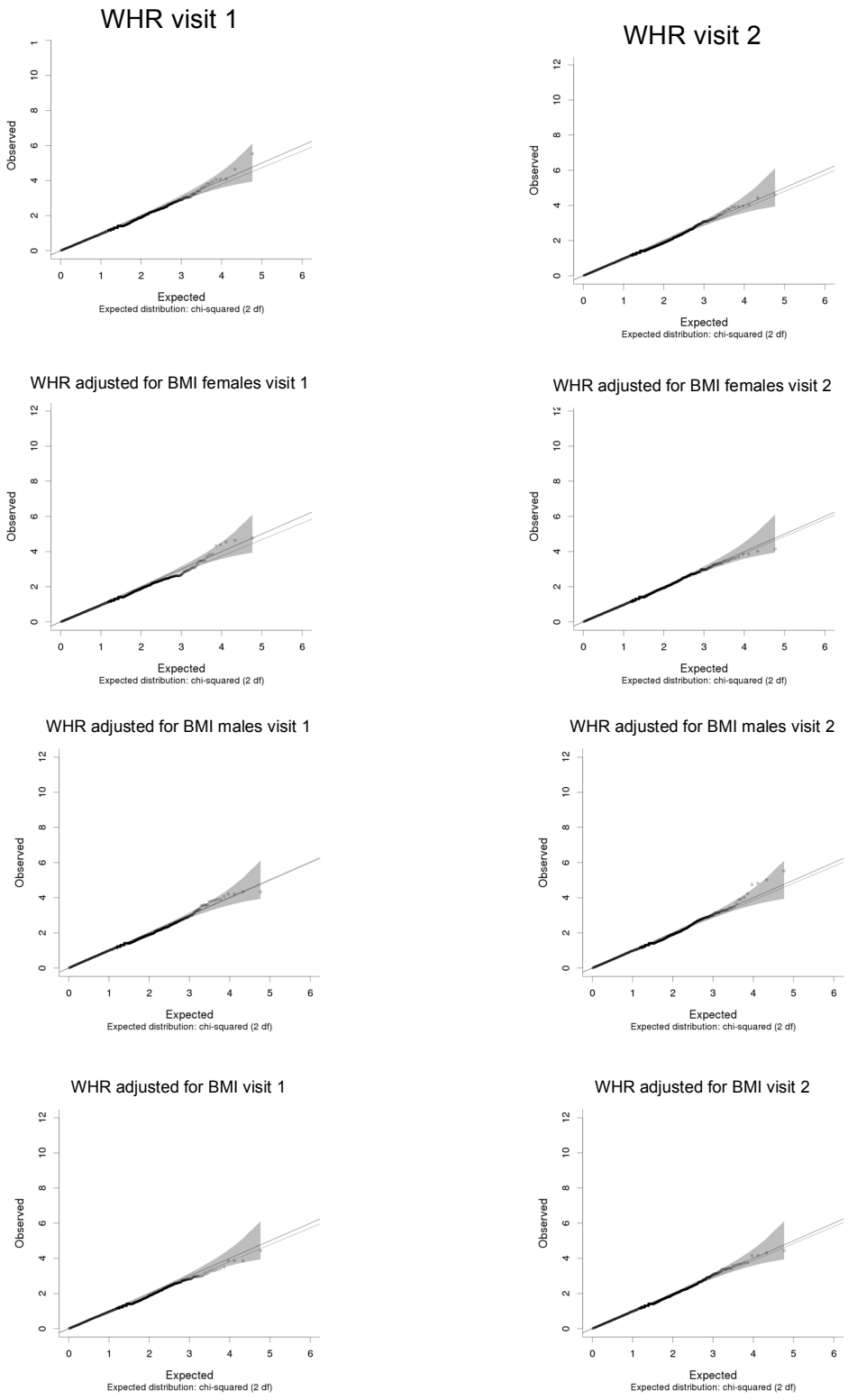


Figure 13. QQ plots for common variant gene-based association tests.

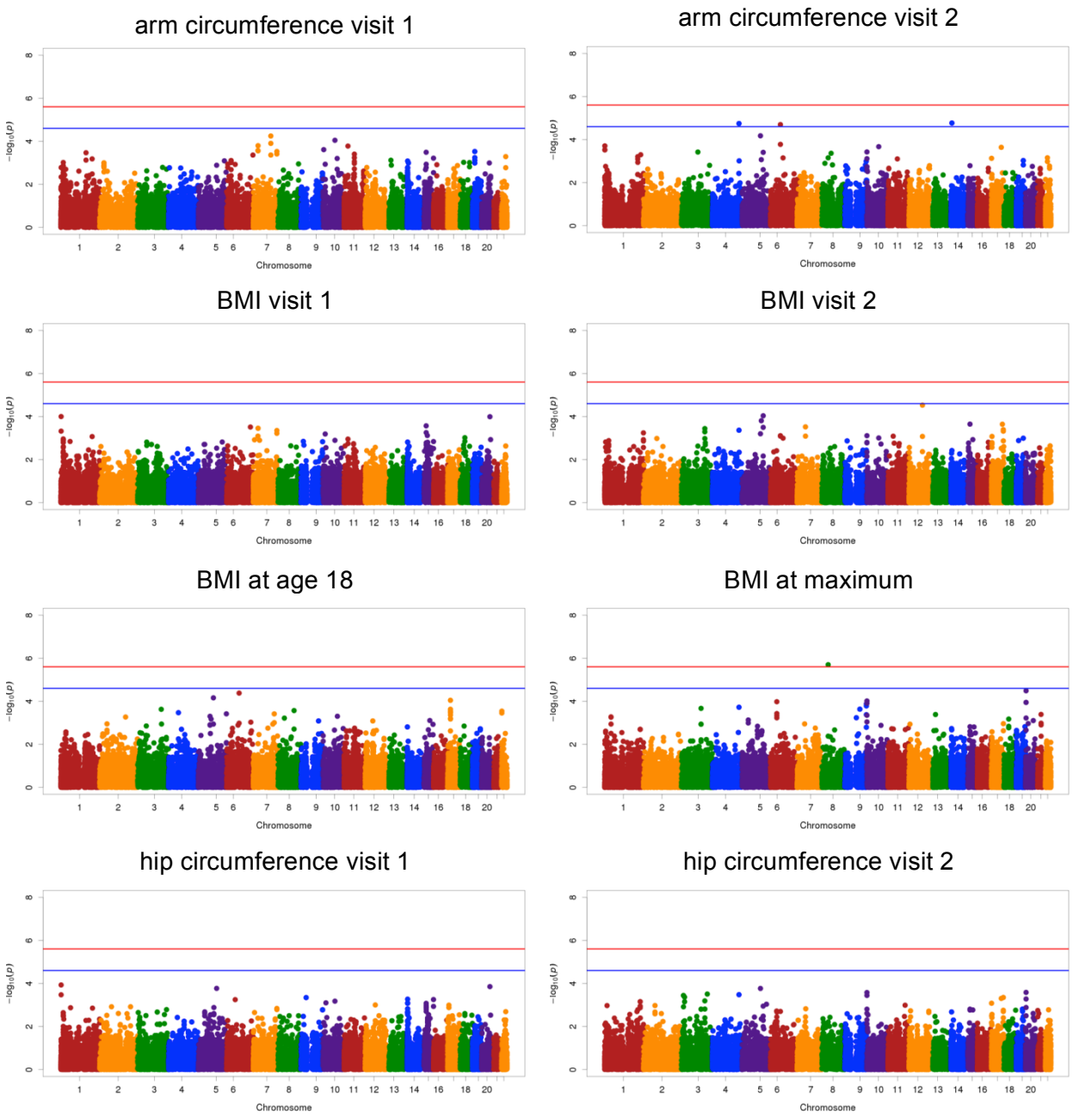


Figure 14. Manhattan plots for common variant gene-based association tests.

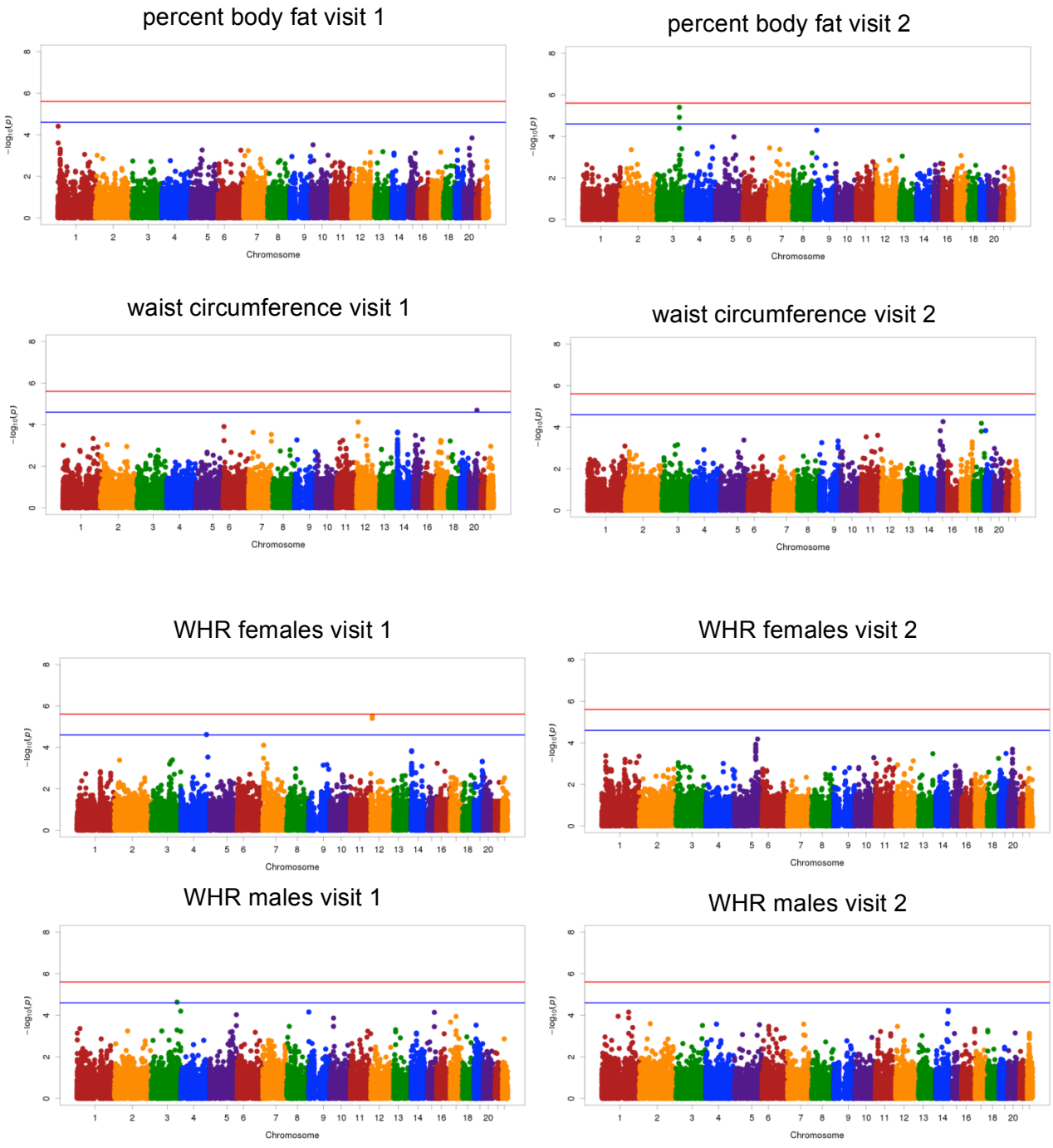


Figure 14. Manhattan plots for common variant gene-based association tests.

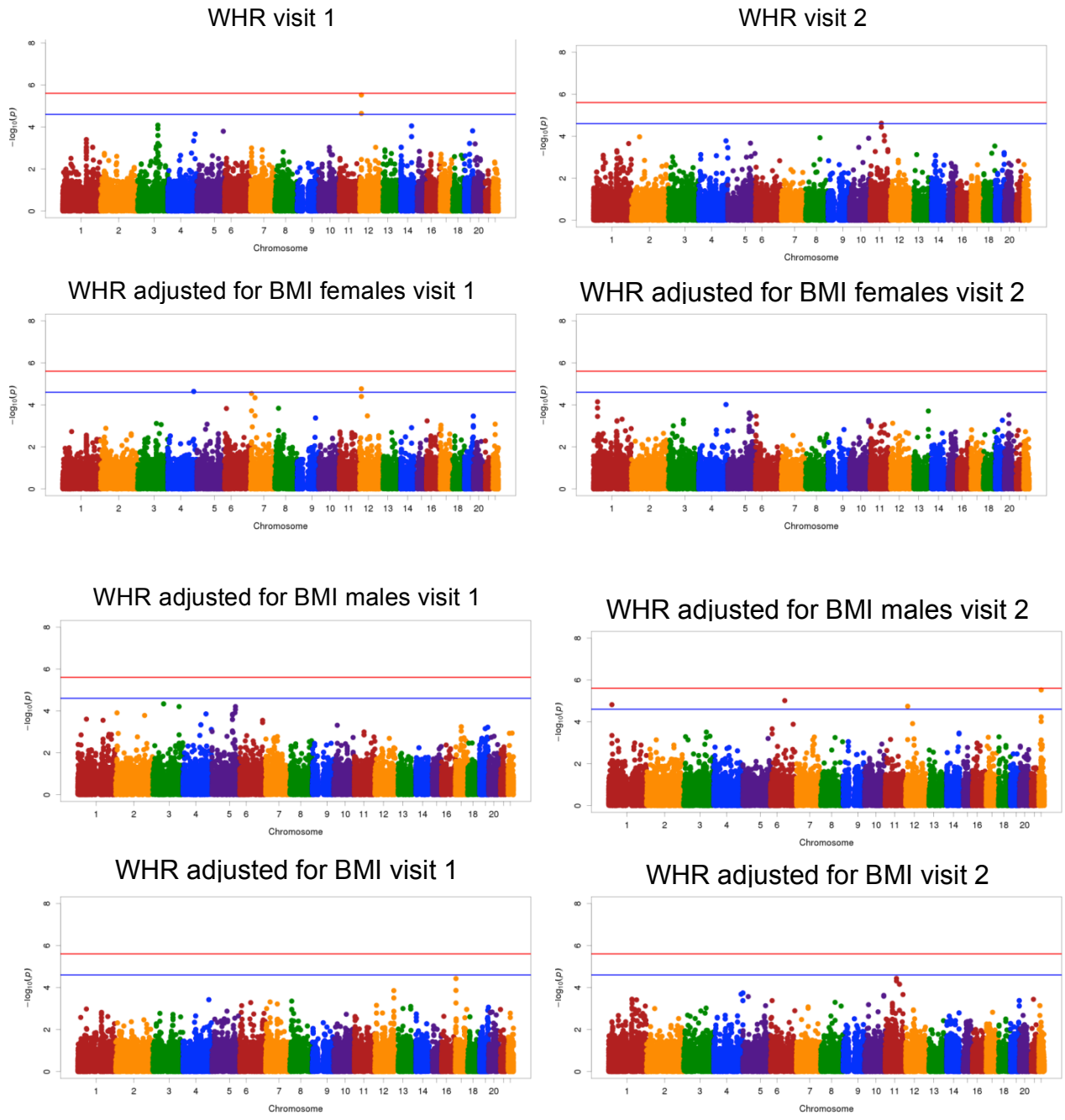


Figure 14. Manhattan plots for common variant gene-based association tests.

						Visit 1			Visit 2		
Trait	**Name**	**Chr**	**Start**	**End**	**Length**	**SNPs**	**Tests**	**Pval**	**SNPs**	**Tests**	**Pval**
WHR adj BMI males	*ECE1*	1	21543740	21672034	128295	531	29.3	0.0014	433	21.0	1.52×10^{-5}
percent body fat	*RNF13*	3	149530475	149679926	149452	405	30.7	0.040	407	28.8	4.00×10^{-6}
percent body fat	*LOC100422259*	3	149656780	149657584	805	66	14.3	0.059	66	13.9	1.20×10^{-5}
WHR males	*LOC100420939*	3	172398504	172399011	508	136	12.6	2.34×10^{-5}	109	7.9	0.0013
WHR adj BMI females	*RPL9P16*	4	169676949	169677632	684	140	11.6	2.30×10^{-5}	140	11.6	0.20
WHR females	*RPL9P16*	4	169676949	169677632	684	140	11.6	2.40×10^{-5}	140	11.6	0.119
arm	*C4orf27*	4	170650619	170679093	28475	160	18.7	0.021	146	19.5	1.80×10^{-5}
arm	*LOC642554*	6	64151128	64153818	2691	112	11.2	0.0089	111	10.1	2.02×10^{-5}
WHR adj BMI males	*COPS5P*	6	93801543	93802701	1159	108	17.2	0.84	92	10.1	9.69×10^{-6}
BMI at max	*C8orf4*	8	40010989	40012821	1833	114	22.5	2.00×10^{-6}			
WHR	*PGM2L1*	11	74041361	74109502	68142	322	24.2	0.14	325	25.6	2.41×10^{-5}
WHR adj BMI females	*OLR1*	12	10310899	10324790	13892	228	19.2	1.72×10^{-5}	228	20.2	0.75
WHR	*OLR1*	12	10310899	10324790	13892	226	20.3	2.27×10^{-5}	230	19.3	0.49
WHR females	*OLR1*	12	10310899	10324790	13892	228	19.2	4.00×10^{-6}	228	20.2	0.82
WHR	*TMEM52B*	12	10331557	10344403	12847	252	19.9	3.00×10^{-6}	258	19.2	0.168
WHR females	*TMEM52B*	12	10331557	10344403	12847	255	19.3	3.00×10^{-6}	257	19.7	0.54
WHR adj BMI males	*RPL19P17*	12	12721058	12721732	675	84	16.1	0.25	80	14.1	1.81×10^{-5}
arm	*LOC401767*	14	25894117	25901442	7326	261	26.4	0.021	249	23.9	1.71×10^{-5}
waist	*PFDN4*	20	52824502	52836492	11991	276	22.6	2.01×10^{-5}	247	22.2	0.43
WHR adj BMI males	*PLA2G3*	22	31530793	31536469	5677	116	19.1	0.59	99	13.5	3.00×10^{-6}

Table 10. Common variant gene-based association results.

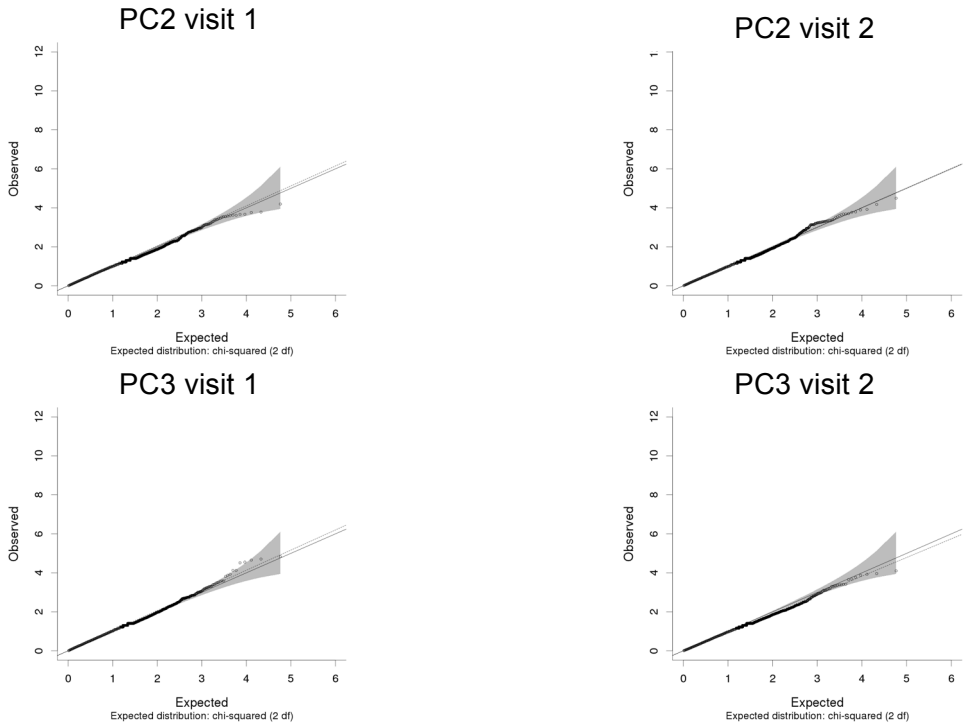


Figure 15. QQ plots for common-variant gene-based association tests for composite adiposity measures.

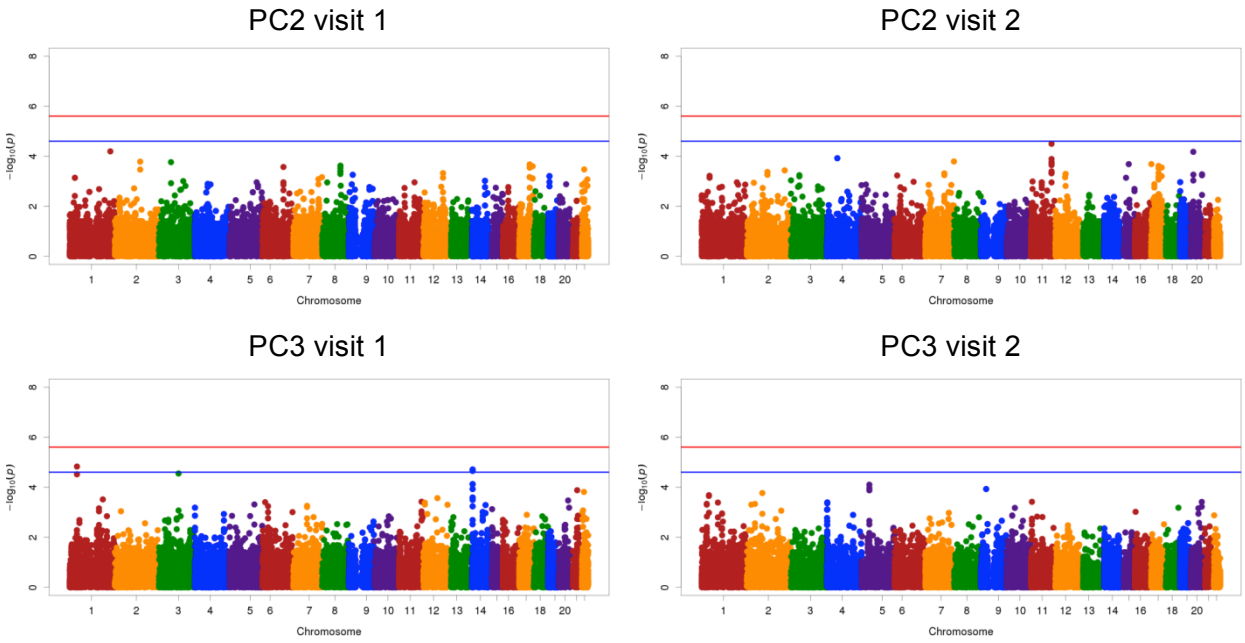


Figure 16. Manhattan plots for common-variant gene-based association tests for composite adiposity measures.

						Visit 1			Visit 2		
Trait	**Name**	**Chr**	**Start**	**End**	**Length**	**SNPs**	**Tests**	**Pval**	**SNPs**	**Tests**	**Pval**
PC3	*MYCL1*	1	40361096	40367687	6592	130	13.6	1.49×10^{-5}	131	14.6	0.089
PC3	*TRAV12-3*	14	22433736	22434290	555	199	13.4	1.96×10^{-5}	200	13.7	0.0071
PC3	*TRAV8-6*	14	22446919	22447360	442	157	10.7	2.22×10^{-5}	158	11.2	0.0090

Table 11. VEGAS association results for composite adiposity measures.

Multi-trait Gene-Based analysis: As with the single variant analyses, a multi-trait analysis was conducted for the gene-based results generated by VEGAS (Figures 17 and 18). Three transcripts come up as genome-wide significant or suggestive and are shown in Table 12. All three are located on chromosome 17p13.2 and are physically adjacent or overlapping. While no variant in these genes or the genes themselves came up as significant or suggestive for any one trait, the most significant gene-based signal is an association between *Germ Cell Associated 2* (*GSG2*) and WHR adjusted for BMI at the first visit ($p = 3.75 \times 10^{-5}$). For these three transcripts, multiple traits were nominally associated with these three genes including BMI, arm circumference, hip circumference, and percent body fat at visit 2, BMI at maximum, WHR with and without adjusting for BMI and hip circumference at visit 1, indicating these genes have a small impact on multiple traits.

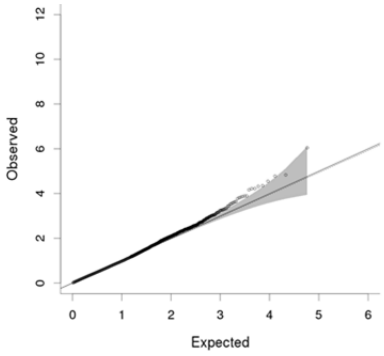


Figure 17. Multi-trait common variant gene-based analysis QQ plot.

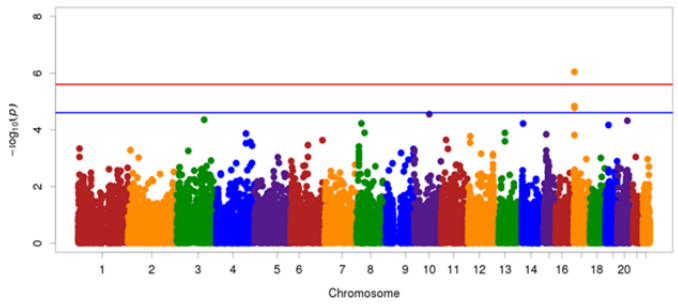


Figure 18. Multi-trait common variant gene-based analysis Manhattan plot

chr	gene	start	end	*p*-value
17	*P2RX5-TAX1BP3*	3566187	3599698	1.66×10^{-5}
17	*P2RX5*	3576521	3599698	1.46×10^{-5}
17	*GSG2*	3627197	3629993	9.02×10^{-7}

Table 12. Multi-trait common variant gene-based analysis top associations

DISCUSSION

Based on the sexual dimorphism, WHR residuals were created within sex stratum. Genetic analyses were performed both in the combined residuals and separately for males and females. Based on the top association signals shown in Table 7, there is of overlap of association signals between the female and sex-combined analyses, this is not apparent with males. This is due to the much lower male sample size resulting from the sample being about 75% female. The composite measures, PC2 is also sexually dimorphic, however in this case the residuals were highly correlated whether samples were stratified on sex or sex was adjusted for as a covariate. Genetic analyses of PC2 were done on the residuals created using sex as a covariate.

Due to the correlation between BMI and WHR, WHR analyses were done both with and without a BMI covariate. By including BMI as a covariate, the GIANT

consortium has successful identified loci that are associated with central adiposity, and not overall BMI(107). Despite the success of this approach in large consortium settings, in these data the correlation between associations with and without BMI were high (r^2>0.8), and most signals were stronger in the analysis without adjusting for BMI. However there are some signals that are stronger after adjusting out BMI, but only rs74862838 on chromosome 1 reached genome-wide suggestive criteria.

The analysis of the same adiposity traits measured in the same people 8.5 years apart should yield highly correlated results. However many of the top signals in Table 7 are not associated or are only associated at a nominal level for the other time point. The raw traits are correlated across the visits with an r^2 ranging from 0.72 to 0.90 (see Table 3 in chapter 2). The t-statistics are much less correlated (r^2 ranging from 0.03 to 0.94, excluding WHR adjusted for BMI); this is primarily due to including a non-random subset of samples from the first visit in the second visit and secondarily due to changes in BMI over time. While there was some loss to follow up, most individuals were excluded from the analysis due to being treated for type 2 diabetes prior to the second visit. Because these samples are not missing at random, the different associations may be due to differences in adiposity between those that did not develop diabetes over the 8.5 years between visits and those that did. The individuals that develop diabetes within a short time frame are the most interesting, because this is time frame in which disease development may be preventable with targeted intervention.

Despite having very few significant and suggestive signals, some of these signals are biologically interesting. To start with, variants in the same region as some associations reported here have previously been associated with obesity related traits.

These include variants in and near *NPY*, *PKHD1*, and *MACROD2*. *NPY* encodes neuropeptide Y, a signaling molecule that acts in the nervous system, including in the arcuate nucleus of hypothalamus, regulating hunger and satiety along side POMC and the leptin-melanocortin pathway(178). In Starr County, previous linkage studies have associated this gene with obesity(30). Here we report an association of a rare variant (MAF = 0.6%) 200kb from *NPY* with BMI at the first visit; this association is not seen at the second visit due to having fewer than 5 carriers.

PKHD1, which when mutated can cause autosomal recessive polycystic kidney disease(179), has previously been associated with waist circumference(101) and weight loss following bariatric surgery(180). Our associated variant *PKHD1* has a much lower MAF, indicating it is not in strong LD with previously reported signals. The initial association with smaller WHR adjusted for BMI at the first visit is diminished at visit 2, despite only having one fewer carrier. The imputation quality is at the lower boundary (info = 0.7) of variants that we examined, calling the association observed here into doubt.

MACROD2 was previously reported as associated with disordered eating(181) and suggestively associated with extreme obesity(94). The protein is expressed in many tissues including the brain, liver, muscle and pancreas(182). Our reported suggestive association with arm circumference further supports previous findings that this region of the genome plays a role in obesity.

Beyond these previously reported loci, there are also additional biologically interesting loci. *GABRB2*, encodes a subunit of gamma-aminobutyric acid (GABA) A receptor, which mediates inhibitory synaptic signal transmission in the nervous

system(183). The $GABA_A$ receptor is composed of five subunits, usually two alpha class subunits, two beta class subunits, and one gamma class subunit(184). The various homologs of each subunit class have different specificities to allosteric modulators including many pharmacological agents including benzodiazepines(184). The $GABA_A$ receptor is largely responsible for inhibitory signals to the hypothalamo–pituitary–adrenocortical axis(185). Dysregulation of this signaling pathway, as occurs under stress, is related to obesity(186, 187). Genes encoding other subunits of the receptor have previously been associated with obesity related traits including *GABRB1* with change in weight in Hispanic children(188) and *GABRA4* with type 2 diabetes(189). We reported an association of a rare variant in *GABRB2* with WHR at the second visit. This variant is imputed, however the quality at the second visit is near the minimum threshold for inclusion.

Carriers of the rare allele of rs112863316 in *LPIN1* had an average of 8.6% higher percent body fat than non-carriers at the first visit. There are far fewer carriers at the second visit (21 at visit 1 compared to 11 at visit 2) reducing power to detect an effect. Mice harboring knock out mutations in this gene have characteristics typical of human lipodystrophy, including neonatal fatty liver, hypertriglyceridemia in infancy, adipose tissue deficiency, and glucose intolerance(190). Overexpression of *Lpin1* in adipose tissue results in mice with diet-induced obesity and enhanced insulin sensitivity(191). This suggests a causal variant in the region may increase *LPIN1* expression or activity.

These handful of associations are biologically interesting candidates. While biological plausibility gives us confidence in an association results, it is neither

necessary nor sufficient to conclude a variant or gene is causative. Replication in independent studies is needed to determine if these variants truly play a role in fat distribution, or if these associations are spurious. These regions have not previously been associated with obesity related traits, however large-scale studies including very rare variants are still underway.

Most of these variants are low frequency, resulting in lower imputation quality. While strict quality control criteria were applied, imputation leaves additional uncertainty in the genotypes. If one of ten individuals called as a carrier were incorrectly imputed, then the test statistic could be drastically altered. Using sequencing data, or arrays intended to capture rare variation will give us additional certainty in the genotyping.

The multi-trait analysis will detect associations in multiple traits that are not due to the correlation across the traits. For this analysis, only four signals are suggestive (Table 9). Zhu *et al.* propose an extension that allows for heterogeneous effects across traits and studies, however the original method is more powerful when considering traits on the same scale(176). Since we are analyzing inverse normalized residuals, the original method was utilized. Two of the suggestive results are biologically interesting. rs7723319 resides between *GAS7* and *MYH13*. *MYH13* encodes a part of the myosin heavy chain, which is a key part of skeletal muscle fibers. *GAS7* encodes growth arrest-specific 7, which plays a role in neuronal development. Multiple low-frequency variants in *EBF4*, which encodes early B cell factor 4, are suggestively associated with these traits. *EBF4* plays a role in both neuronal development and B-cell maturation. Because

we know neuronal signaling and development are important for hunger and satiety, both of these genes are biologically plausible.

While gene-based analyses offer the potential to identify genes that contain multiple smaller effect size variants, the association results from VEGAS, using genes and 20 kilobases in each direction, yielded only one genome-wide significant result, in 28 different analyses. While the lack of signal may be due to not having genes with multiple variants independently associated with obesity-related traits, it could also be due to a number of considerations for gene-based tests. Including 20 kilobases in both directions from the gene may include many null variants, diminishing the combined signal. To check this, a secondary analysis was run including only three kilobases in each direction from genes. This analysis was also well calibrated, however it did not reveal stronger signal. Each window size identified some unique associations, however most signals were common to both analyses. Another option to alleviate the problem of excessive null signal is to include only variants that are associated at some preset level. This requires additional statistical considerations to account for the bias. To decrease the number of variants we can also use functional annotation to include only variants we have biological reason to think could impact protein function. Genome-wide array data has little coding variation well represented, and annotation informed approaches are better suited to coding-centric data, as will be investigated in chapter 4.

Chapter 4: Whole Exome Sequencing

INTRODUCTION

While genome-wide arrays, offer an agnostic look across the genome, it is often difficult to parse out causative transcripts and variants from these data. They do not provide sufficient coverage of the mutations and variants that are most likely to be functional. When trying to assess functionality of a variant, the most clear cut case is that of a protein altering variant, that is one that changes the amino acid sequence, length, or splicing of the gene product. Very few protein-altering variants are assayed on genome-wide chips. Some coding variants are imputed, but these are limited to common haplotypes that are represented in the 1000 genomes reference panel. The development of next generation sequencing, however, provides the opportunity to interrogate most of coding variation either through whole genome or whole exome sequencing.

While the "thrifty gene" hypothesis would support the idea that genetic variants increasing efficiency of energy storage would be common(20), this applies to older variation that has undergone generations of selection during times of famine. The Exome Sequencing Project (ESP), has shown us that while most variation within a data set is rare (86% of variants have a MAF<0.5%), most of an individual's variation is common(192). At the study level, ESP reported 58% of variants in whole exome sequencing data are nonsynonymous and 38% are synonymous, while individuals average 35 nonsense variants, 5754 missense variants, and 7652 synonymous variants(192). A considerable amount of the variation passed down from parent to offspring is rare and relatively recent variation and consequently it has not had sufficient time for selection to play its role. Both common variants that may have been selected for

when famine was a selective pressure, and more recent rare variation are shared and plausibly contribute to the genetic architecture of obesity.

To assess the role of coding variation in obesity, unrelated individuals with genome-wide array data were whole exome sequenced, although a subset were removed prior to sequencing due to insufficient DNA. More common variants were tested for associations with the same traditional measures of adiposity as in chapter 3, as well as the composite measures of adiposity derived in chapter 2. To assess the role of low frequency and rare variants, gene-based tests aggregating variants across a transcript were used for the same outcomes. As in chapter 3, the test statistics from the traditional adiposity measures were meta-analyzed correcting for the correlation between traits, to identify genes independently associated with multiple obesity measures.

METHODS

Whole Exome Sequencing: The T2D-GENES Consortium was formed to investigate the role of coding variation in type 2 diabetes across multiple ancestry groups. As part of this consortium, 13,000 individuals from 5 ancestry groups were whole exome sequenced, including all of the 1618 unrelated Starr County samples with Affymetrix 6.0 genotype data available that were analyzed in chapter 3. Sequencing was performed using Agilent SureSelect All Exon Kit v.2. Using Picard (http://picard.sourceforge.net), sequencing reads were processed and aligned to hg19 reference genome. The GATK HaplotypeCaller was used to call variant sites across 26k samples that represent the T2D-GENES consortium as well as SIGMA LuCamp, and the Exome Sequencing Project (193). A total of 1,497 unrelated Starr County samples were

successfully sequenced; however 7 failed QC metrics such as excess heterozygosity, excess non-reference alleles, or low GWAS concordance.

Single variant analysis: Tests of association with each variant with a minor allele count of at least five were performed using a linear regression model accounting for cryptic relatedness between individuals using EMMAX as implemented in the EPACTS pipeline(194). Variants with a minor allele count of less than five were excluded due to instability in model for very rare variants. To adjust for population structure two genetic principal components, as described in chapter 3, were included as covariates. Exome-wide significance was defined as $P < 5 \times 10^{-7}$, which corrects for the approximately 100,000 coding variants tested. Variants with $P < 5 \times 10^{-6}$ were considered exome-wide suggestive.

Gene grouping criteria: All variants were annotated using SnpEff v.3.1 and dbNSFP(195-197). Annotation provides information about the impact of a variant on each gene-product. For gene-based testing, variants were annotated to all possible transcripts and variants within each transcript were included based on these annotations and the minor allele frequency in the Starr County data. Transcript specific annotations were used instead of most deleterious annotation in the gene, because alternative transcripts are found in different amounts across various tissues, making them the logical biological product of interest. Criteria, number of transcripts and number of variants included are presented in Table 13.

Mask	Functional annotation	Maximum allele frequency	Variants	Transcripts with at least two variants
Truncating "PTV"	stop loss, stop gained, initiator codon, splice acceptor, splice donor, frame shift	5%	11,734	6,331
Protein altering, predicted damaging "LR"	Truncating variants and splice region, missense, miRNA, in frame mutations predicted damaging by metaLR(198)	1%	31,379	19,140
Protein altering "NS"	Truncating variants and splice region, missense, miRNA, in frame	1%	195,132	74,124

Table 13. Criteria for gene-based grouping. Variants were annotated to all possible transcripts using SnpEff(195).

Gene based association testing: For rare and low frequency variants, single variant test are underpowered and sometimes unstable. To test for an association across a transcript, variants annotated to with functional impacts as described in Table 13 were aggregated together using SKAT-O, as implemented in EPACTS, including a linear mixed model to account for cryptic relatedness (194, 199). SKAT-O optimizes between SKAT, which allows variants to have different directions of effect, and traditional burden test, which assumes that all variants have the same direction of effect(199).

Multi-trait analysis

Assessment of overall significance across the array of obesity related measures was conducted by combining the t-statistics across traits adjusting for the correlation between traits. Analysis was conducted using the software CPASSOC(175, 176). For gene-based tests, *p*-values were transformed into a multivariate normal distribution by applying the inverse of the normal cumulative probability function. These values were

than combined across traits accounting for the correlation across traits, as for single variants. The gene-based *p*-values of the resulting test statistic were calculated from the Gamma distribution, due to inflated type I error in simulated data sets when using the originally proposed Chi-squared distribution (Hao Hu personal communication, January 2, 2015).

RESULTS

Single variant associations: The single variant test statistics are well calibrated, meaning they follow the expected null distribution, after quality control. QQ and Manhattan plots for all 24 traditional adiposity measurements are in Figures 19 and 20, respectively. The genomic inflation factors (λ) range from 0.995 to 1.017, indicating that there is not systematic inflation of the test statistic and any population structure is sufficiently controlled for.

Variants that are significantly ($P < 5 \times 10^{-7}$) or suggestively ($P < 5 \times 10^{-6}$) associated with any of these 24 obesity related traits are in Table 13. A total of 3 significant associations at 3 loci were observed. A common synonymous variant (rs5996200, $p = 2.52 \times 10^{-8}$, β (SE) = 2.2 (0.37), MAF= 0.11) in *CYB5R3* was significantly associated with arm circumference at visit 1, and also suggestively associated with arm circumference at visit 2 and BMI at visit 1. A common intronic variant (rs2301922, $p = 2.71 \times 10^{-7}$, β (SE) = 4.2 (0.87), MAF= 0.24) in *WIPF3* is associated with waist circumference at visit 1; the same variant is also suggestively associated with BMI at visit 1. Lastly, a low-frequency synonymous variant (rs117042905, $p = 2.41 \times 10^{-7}$, β (SE) = -0.083 (0.8018, MAF= 0.03) in *ARHGAP39* is associated with WHR adjusted for BMI in males at visit 2; this is based on 89 men.

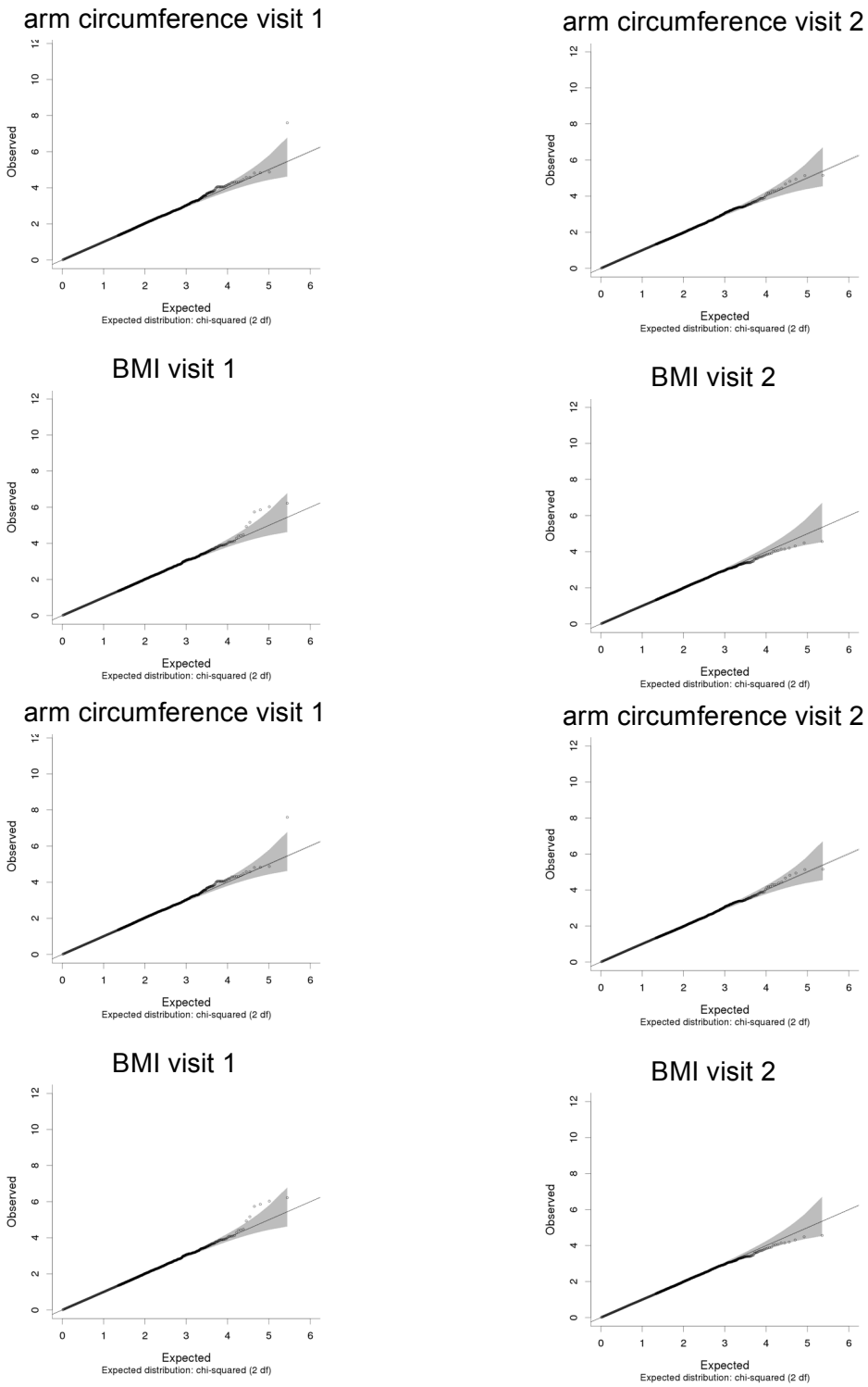


Figure 19. QQ plots for exome sequencing single variant association tests.

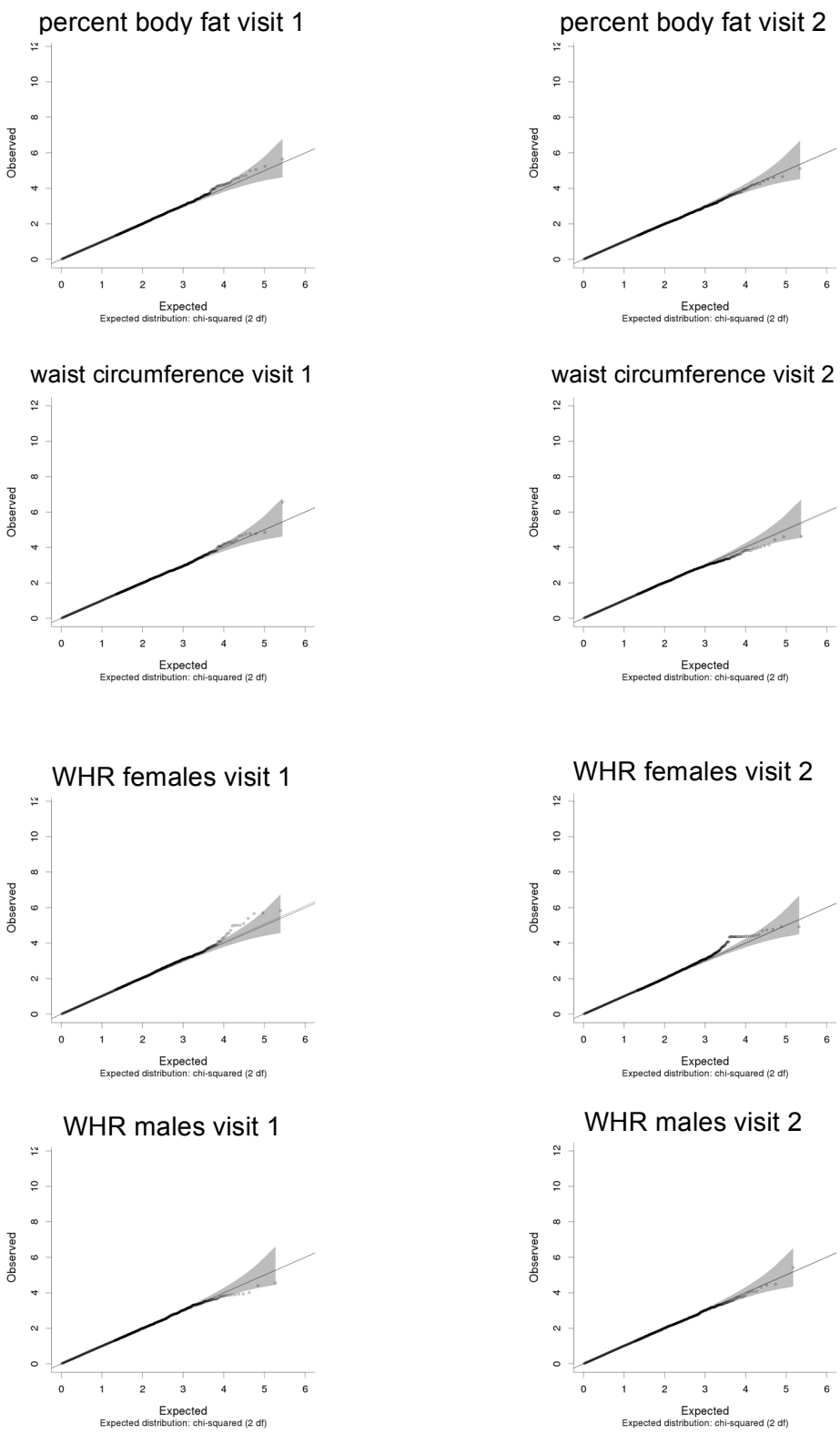


Figure 19. QQ plots for exome sequencing single variant association tests.

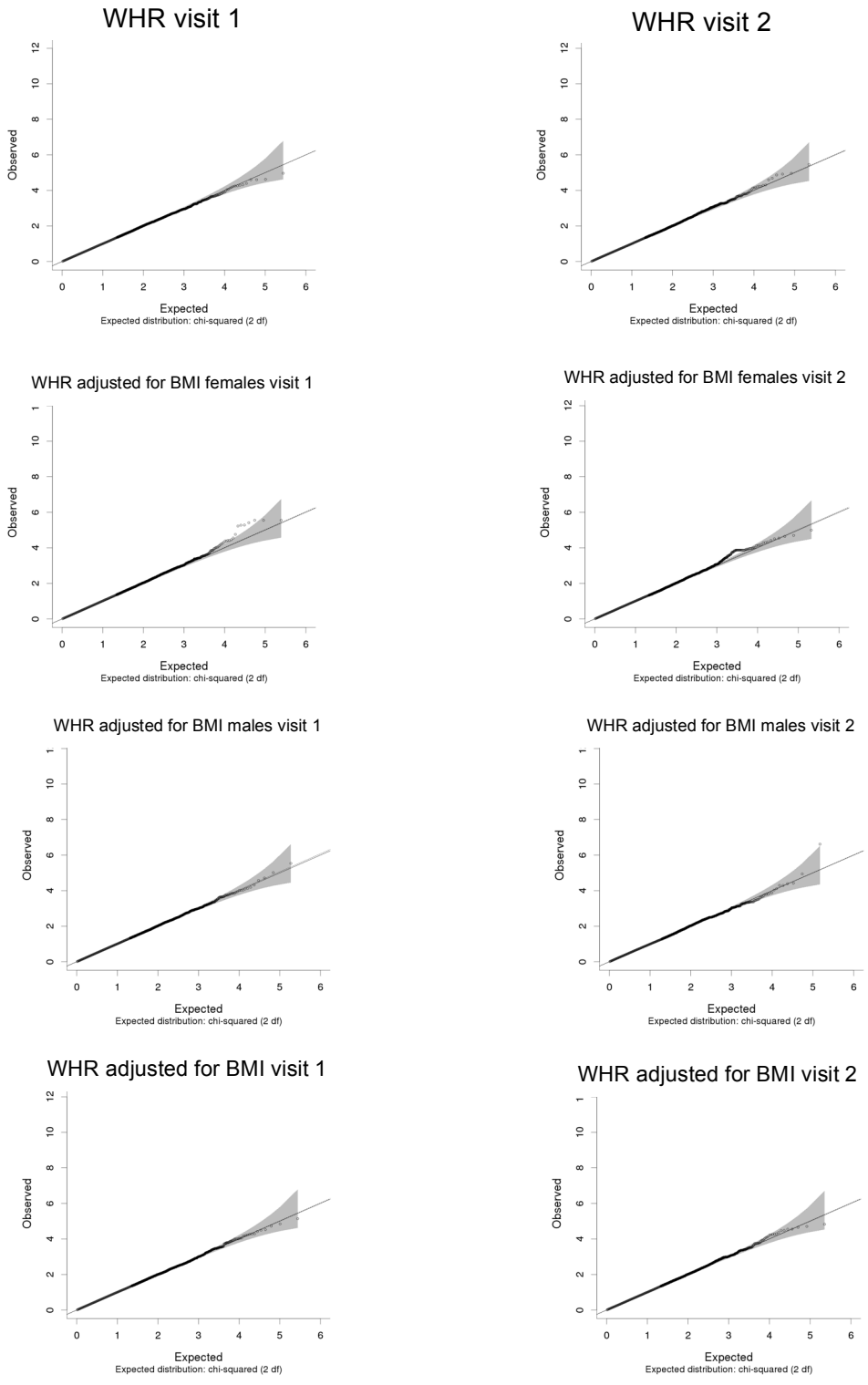


Figure 19. QQ plots for exome sequencing single variant association tests.

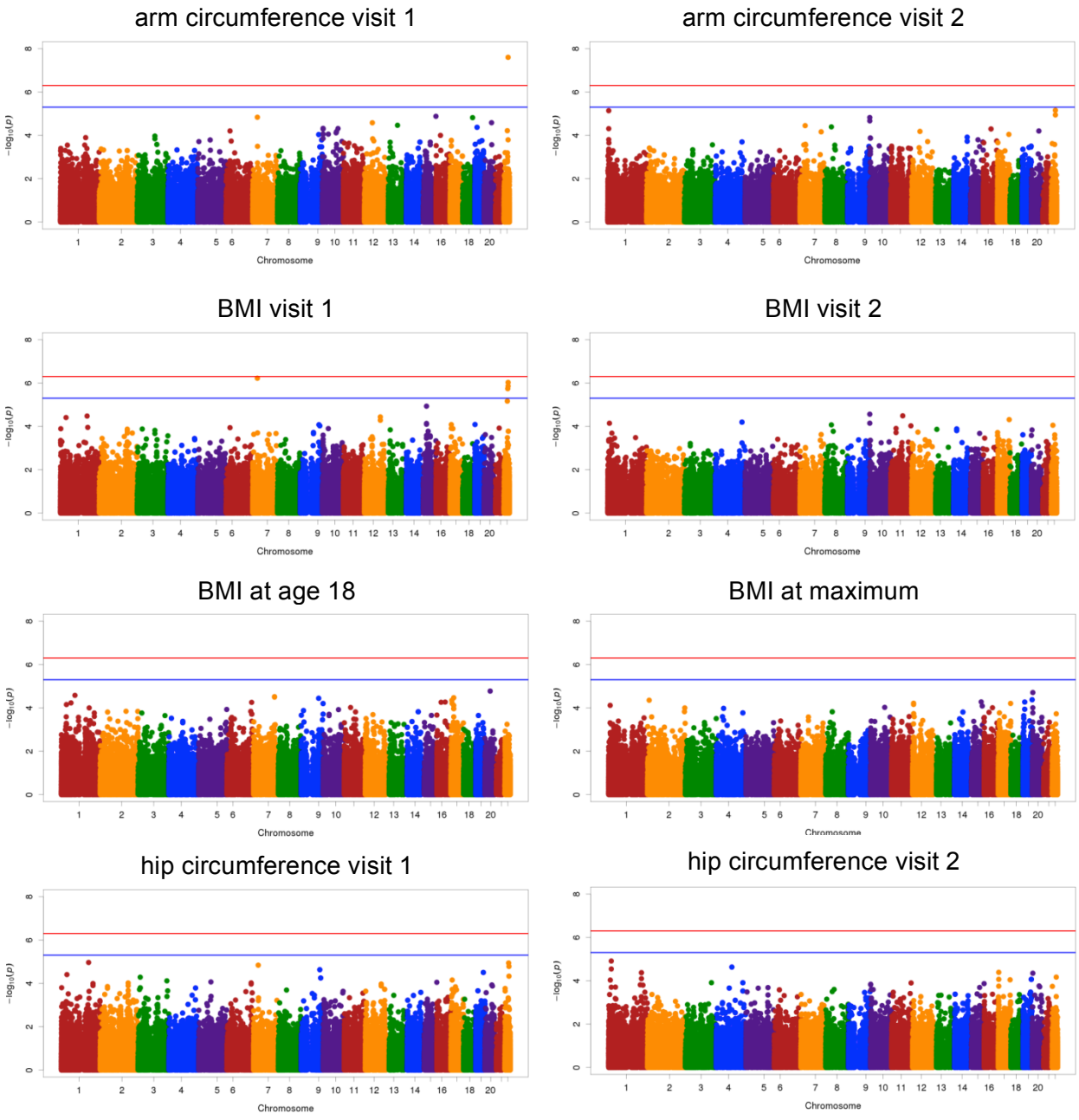


Figure 20. Manhattan plots for exome sequencing single variant association tests.

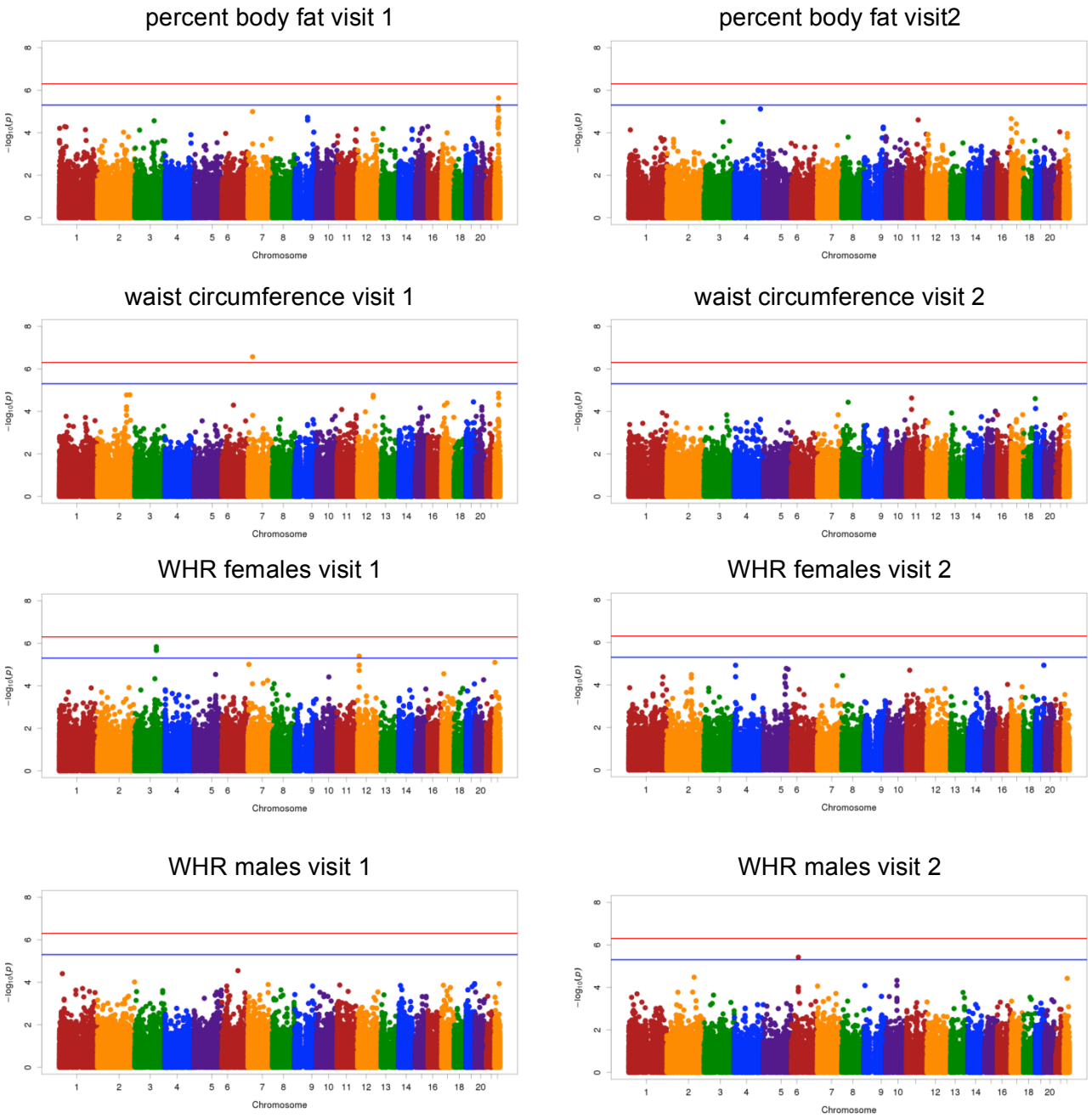


Figure 20. Manhattan plots for exome sequencing single variant association tests.

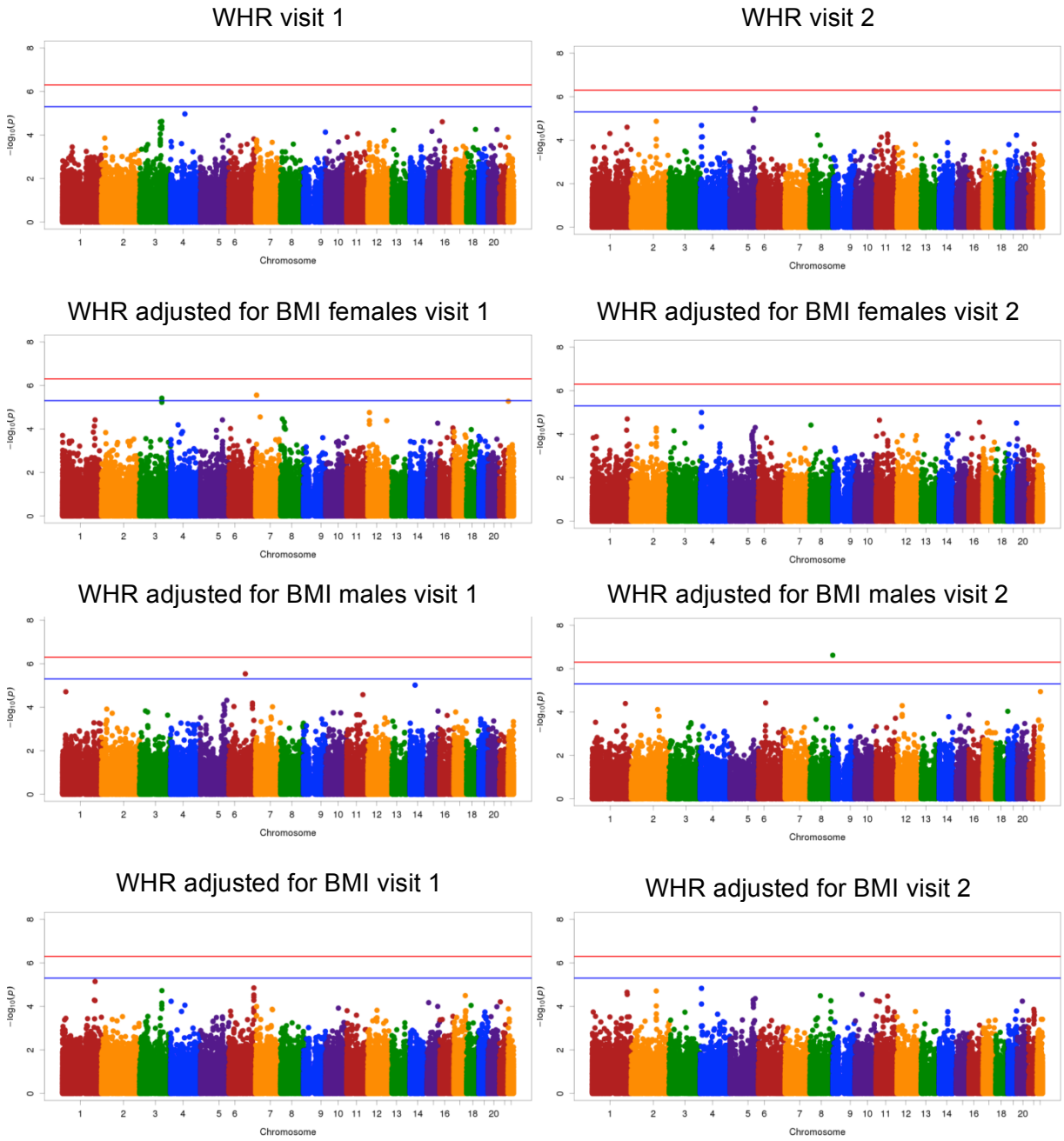


Figure 20. Manhattan plots for exome sequencing single variant association tests.

chr	pos	variant	gene	ref	alt	protein change	consequence	trait	visit	P	β (SE)	EAC	N	MAF
3	141905253	rs9837628	*GK5*	C	A	NA	intronic	WHR females	v1	1.97×10^{-6}	-0.025(0.0055)	152	515	0.14757
									v2	0.09073	-0.011(0.0068)	86	292	0.14726
3	141905259	rs9857725	*GK5*	T	G	NA	intronic	WHR adj BMI females	v1	3.86×10^{-6}	-0.023(0.0050)	152	516	0.14729
									v2	0.1219	-0.0099(0.0067)	86	291	0.14777
3	141905259	rs9857725	*GK5*	T	G	NA	intronic	WHR females	v1	1.46×10^{-6}	-0.026(0.0055)	152	516	0.14729
									v2	0.1121	-0.010(0.0069)	86	291	0.14777
3	141905261	rs9857726	*GK5*	T	G	NA	intronic	WHR females	v1	2.18×10^{-6}	-0.025(0.0055)	152	516	0.14729
									v2	0.1056	-0.010(0.0068)	86	292	0.14726
5	161116672	rs3811993	*GABRA6*	C	T	T187M	missense	WHR	v1	0.3015	-0.013(0.015)	16	755	0.0106
									v2	3.48×10^{-6}	-0.090(0.019)	10	395	0.01266
6	46684222	rs1805017	*PLA2G7*	C	T	R92H	missense	WHR males	v1	0.1453	-0.0098(0.0065)	136	216	0.31481
									v2	3.81×10^{-6}	-0.044(0.0094)	55	89	0.30899
6	105233244	var_6_105233244	*HACE1*	TAA	TA, TAAA, TAAAA, T, TAAAAA	NA	intronic	WHR adj BMI males	v1	2.92×10^{-6}	0.023(0.0047)	50	196	0.12755
									v2	0.01222	0.023(0.0081)	22	79	0.13924
7	6080686	rs34909691	*EIF2AK1*	A	T,G	L319H	missense	WHR adj BMI females	v1	2.81×10^{-6}	-0.063(0.014)	19	539	0.01763
									v2	0.02583	-0.033(0.016)	16	306	0.02614
7	6226799	rs55681139	*CYTH3*	C	T	NA	intronic	WHR adj BMI females	v1	2.81×10^{-6}	-0.063(0.014)	19	539	0.01763
									v2	0.02583	-0.033(0.016)	16	306	0.02614
7	6227341	rs41282682	*CYTH3*	A	G	Y43Y	synonymous	WHR adj BMI females	v1	2.81×10^{-6}	-0.063(0.014)	19	539	0.01763
									v2	0.02583	-0.033(0.016)	16	306	0.02614

chr	pos	variant	gene	ref	alt	protein change	consequence	trait	visit	P	β (SE)	EAC	N	MAF
7	29915593	rs2301922	*WIPF3*	G	C	NA	intronic	BMI	v1	6.01×10^{-7}	1.8(0.37)	363	762	0.23819
									v2	0.002171	1.5(0.54)	190	399	0.2381
7	29915593	rs2301922	*WIPF3*	G	C	NA	intronic	waist	v1	2.71×10^{-7}	4.2(0.87)	361	759	0.23781
									v2	0.003174	2.8(1.1)	217	455	0.23846
8	145806262	rs117042905	*ARHGAP39*	C	T	A160A	synonymous	WHR adj BMI males	v1	0.005126	-0.028(0.0098)	22	216	0.05093
									v2	2.41×10^{-7}	-0.083(0.018)	6	89	0.03371
12	10313075	rs3736235	*OLR1*	T	C	NA	intronic	WHR females	v1	4.01×10^{-6}	0.020(0.0040)	543	537	0.49441
									v2	0.7703	-0.0019(0.0052)	289	305	0.47377
22	40803186	rs2018393	*SGSM3*	G	T	NA	intronic	BMI	v1	1.82×10^{-6}	-1.5(0.33)	982	760	0.35395
									v2	0.002968	-1.3(0.48)	513	398	0.35553
22	42998902	rs113513063	*POLDIP3*	G	T	P108P	synonymous	BMI	v1	1.39×10^{-6}	8.6(1.5)	16	762	0.0105
									v2	0.0006665	8.0(2.1)	10	399	0.01253
22	42998902	rs113513063	*POLDIP3*	G	T	P108P	synonymous	percent body fat	v1	2.32×10^{-6}	0.092(0.021)	16	761	0.01051
									v2	0.001145	0.078(0.026)	10	372	0.01344
22	43032742	rs5996200	*CYB5R3*	C	T	P77P	synonymous	BMI	v1	9.41×10^{-7}	2.6(0.50)	170	762	0.11155
									v2	0.00119	2.5(0.74)	90	399	0.11278
22	43032742	rs5996200	*CYB5R3*	C	T	P77P	synonymous	arm	v1	2.52×10^{-8}	2.2(0.37)	169	761	0.11104
									v2	1.15×10^{-5}	2.2(0.49)	102	455	0.11209

Table 14. Significant and suggestive single variant associations from exome sequencing.

Composite adiposity measures: In addition to traditional measures of adiposity, associations with the composite measures of adiposity, PC2 and PC3, which capture truncal adiposity above or below the waist, and central versus peripheral adiposity respectively, were also tested. The QQ and Manhattan plots are in Figures 21 and 22, respectively. There were only two suggestive associations, as shown in Table 15. The association with PC2 at visit 1 in *TRAK1* (rs4234445, $p = 2.12 \times 10^{-6}$, β (SE) = 0.12(0.025), MAF= 0.19) replicates a prior suggestive association with visceral adipose tissue/subcutaneous adipose tissue ratio(200).

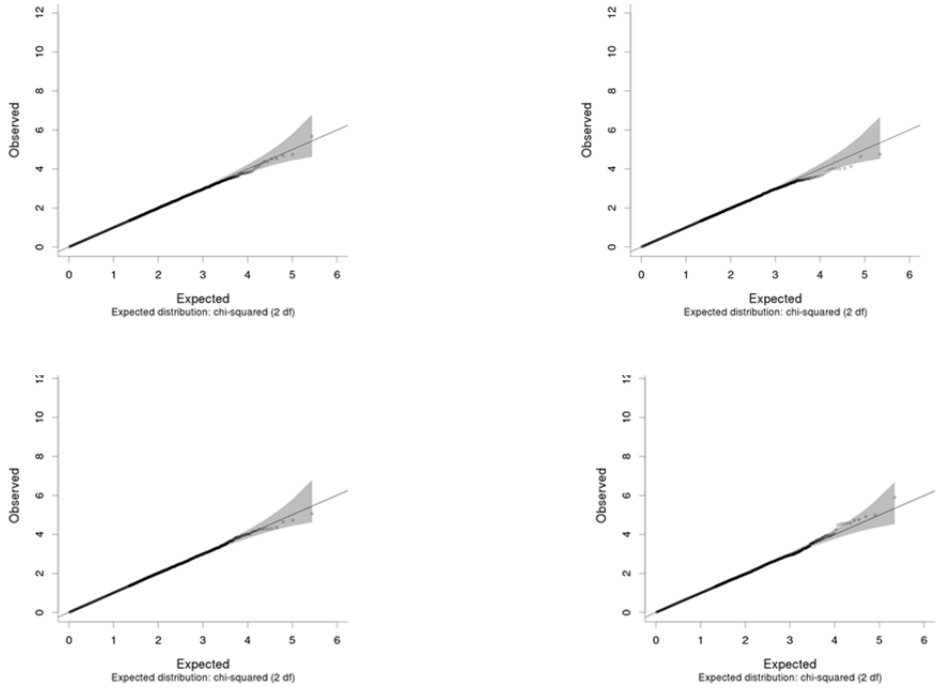


Figure 21. QQ plots for exome sequencing single variant association tests with composite measures of adiposity.

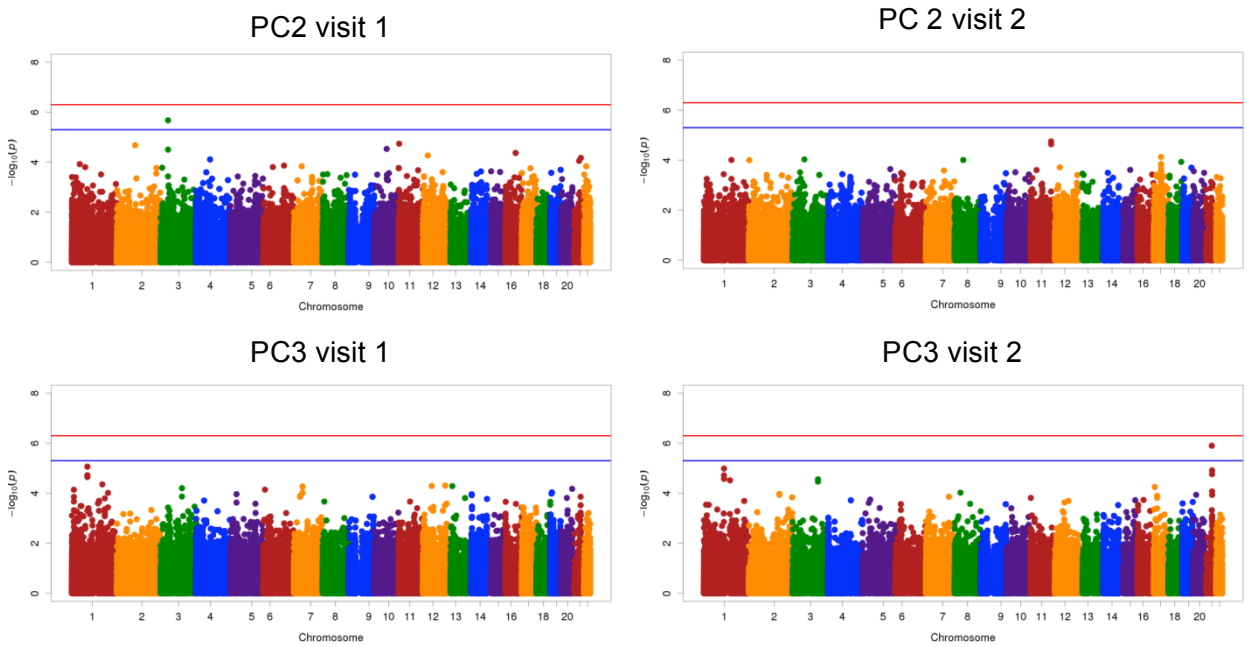


Figure 22. Manhattan plots for exome sequencing single variant association tests with composite measures of adiposity.

chr	pos	variant	gene	ref	alt	protein change	consequence	trait	visit	P	β (SE)	EAC	N	MAF
3	42226151	rs4234445	*TRAK1*	T	C	NA	intronic	PC2	v1	2.12×10^{-6}	0.12(0.025)	1218	753	0.19124
									v2	0.1475	-0.052(0.037)	586	370	0.20811
21	41465622	rs78085341	*DSCAM*	A	T	NA	intronic	PC3	v1	0.003124	0.33(0.12)	17	753	0.01129
									v2	1.26×10^{-6}	-0.82(0.16)	9	370	0.01216

Table 15. Significant and suggestive single variant associations with composite adiposity measures from exome sequencing.

Multi-trait analysis: In addition to analyzing composite measures of obesity (PC2 and PC3), analysis was conducted combining the t-test statistics across adiposity trait while accounting for correlation using the software CPASSOC(175, 176). QQ and Manhattan plots are in Figures 23 and 24 respectively. The test statistic was well calibrated with a genomic inflation factor of 1.002. Significantly and suggestively associated variants are in Table 17. Only rs57772251 was the only variant to reach exome-wide significant ($p = 2.34 \times 10^{-7}$, MAF=0.001). All significant and suggestive signals were intronic.

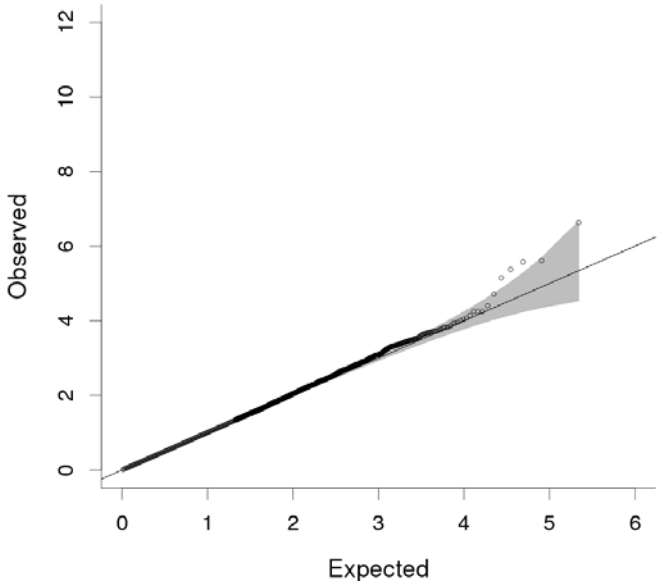


Figure 23. Multi-trait analysis QQ plot of single variant exome sequencing variants.

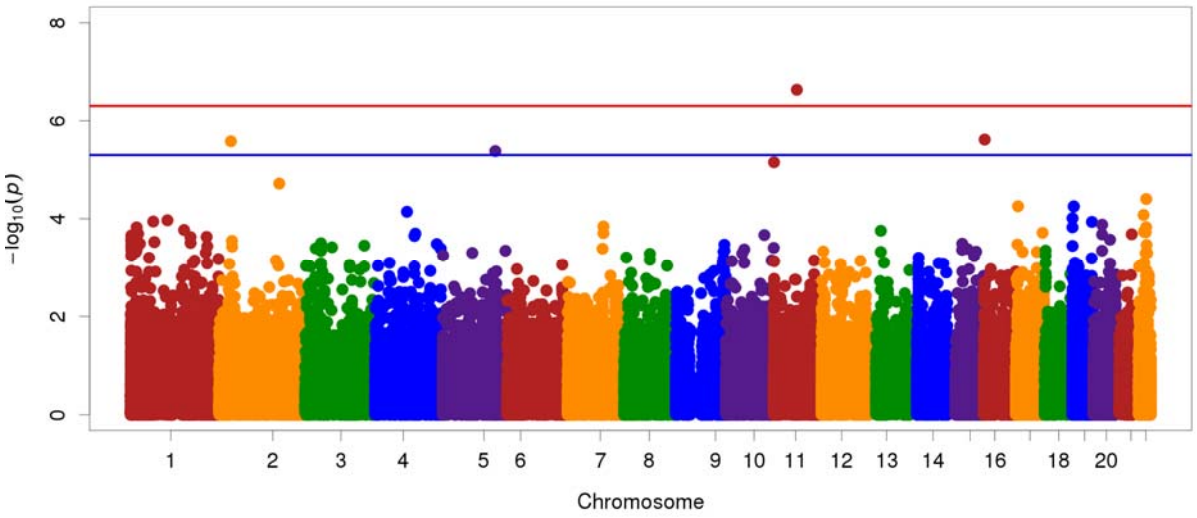


Figure 24. Multi-trait analysis Manhattan plot of single variant exome sequencing variants.

chr	position	variant	ref	alt	gene	function	MAF	*p*-value
2	33810799	var_2_33810799	A	G	*FAM98A*	INTRONIC	0.007	2.63×10^{-6}
5	147812971	rs147707335	T	C	*FBXO38*	INTRONIC	0.012	4.16×10^{-6}
11	64594490	rs57772251	A	G	*CDC42BPG*	INTRONIC	0.001	2.34×10^{-7}
16	2161887	var_16_2161887	C	T	*PKD1*	INTRONIC; CREATE SPLICE ACCEPTOR	0.0054	2.42×10^{-6}

Table 16. Significant and suggestive associations from multi-trait analysis of single variant exome sequencing variants.

Gene-based associations: The QQ plots for gene-based SKAT-O tests for traditional measures of adiposity are in Figure 25, and exhibit some shelves or plateaus, reflecting genes with multiple transcripts, that result in redundant test statistics. For the three different grouping criteria defined in Table 13 as protein truncating with a MAF < 5% (PTV), PTV plus protein altering predicted damaging with a MAF < 1% (LR), PTV plus protein altering with a MAF < 1% (NS), the test statistics were generally well calibrated, although genomic inflation factors range from 0.82 to 1.1. The small genomic inflation factors reflect phenotypes with small sample sizes, as low as 89 for male

specific analyses at visit 2, as well as the small number of tests in comparison to single variant tests. The large genomic inflation factors show early departure from the expected distribution, in part due to the redundancy of including multiple transcripts.

The Manhattan plots for the gene-base SKAT-O tests are in Figure 26. There were no exome-wide significant results ($p < 2.5 \times 10^{-6}$) for any variant grouping or trait. Ten suggestive associations ($p < 2.5 \times 10^{-5}$) with nine genes are in Table 17. Results for both time points and each of the three masks are included. In the case where multiple transcripts were redundant, only the transcript with the smallest *p*-value was included in the table. Eleven of the associations come from the NS mask, which contains low frequency protein altering variants regardless of functional prediction; all but three of these genes include fewer than two variants in the other masks.

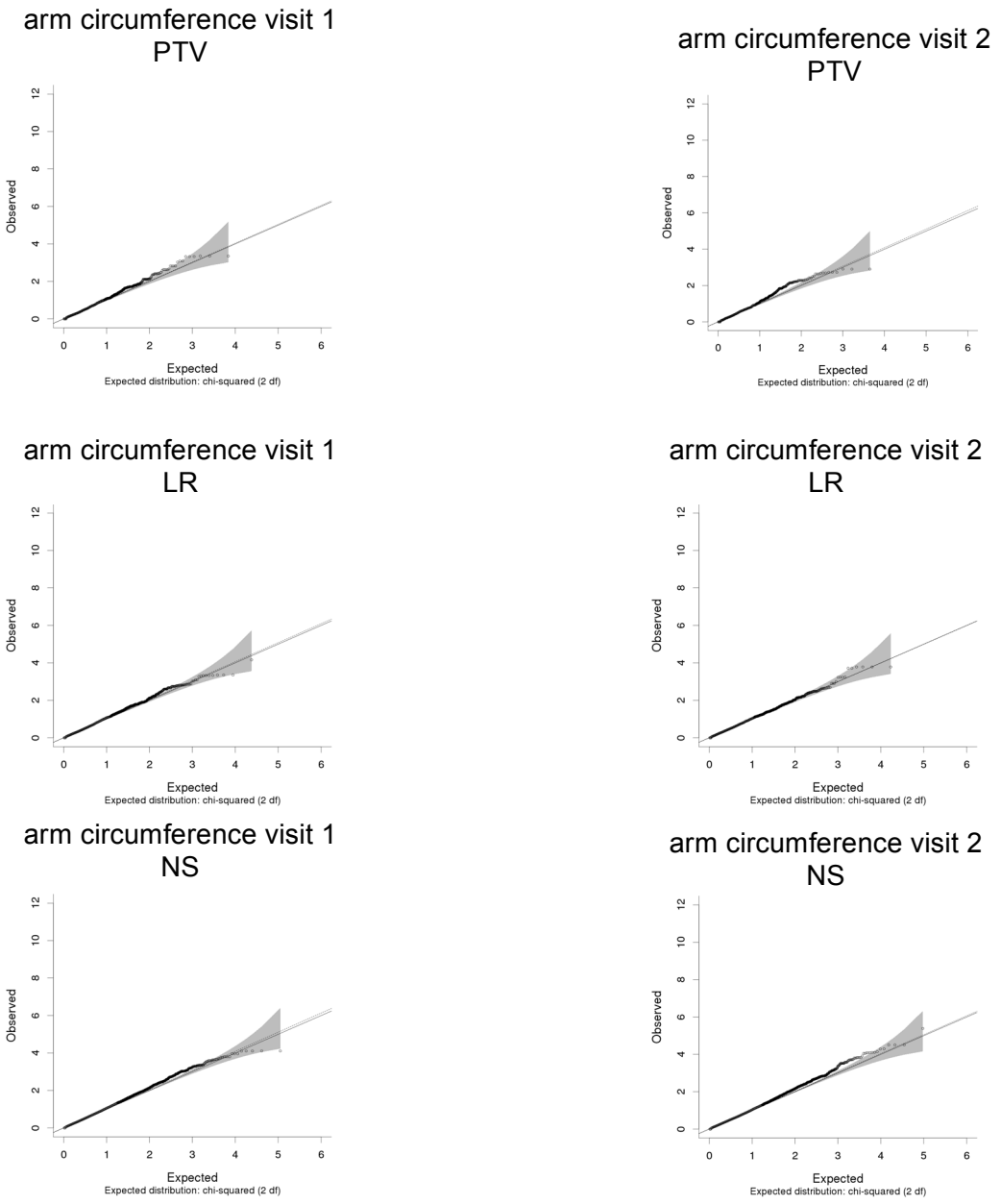


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

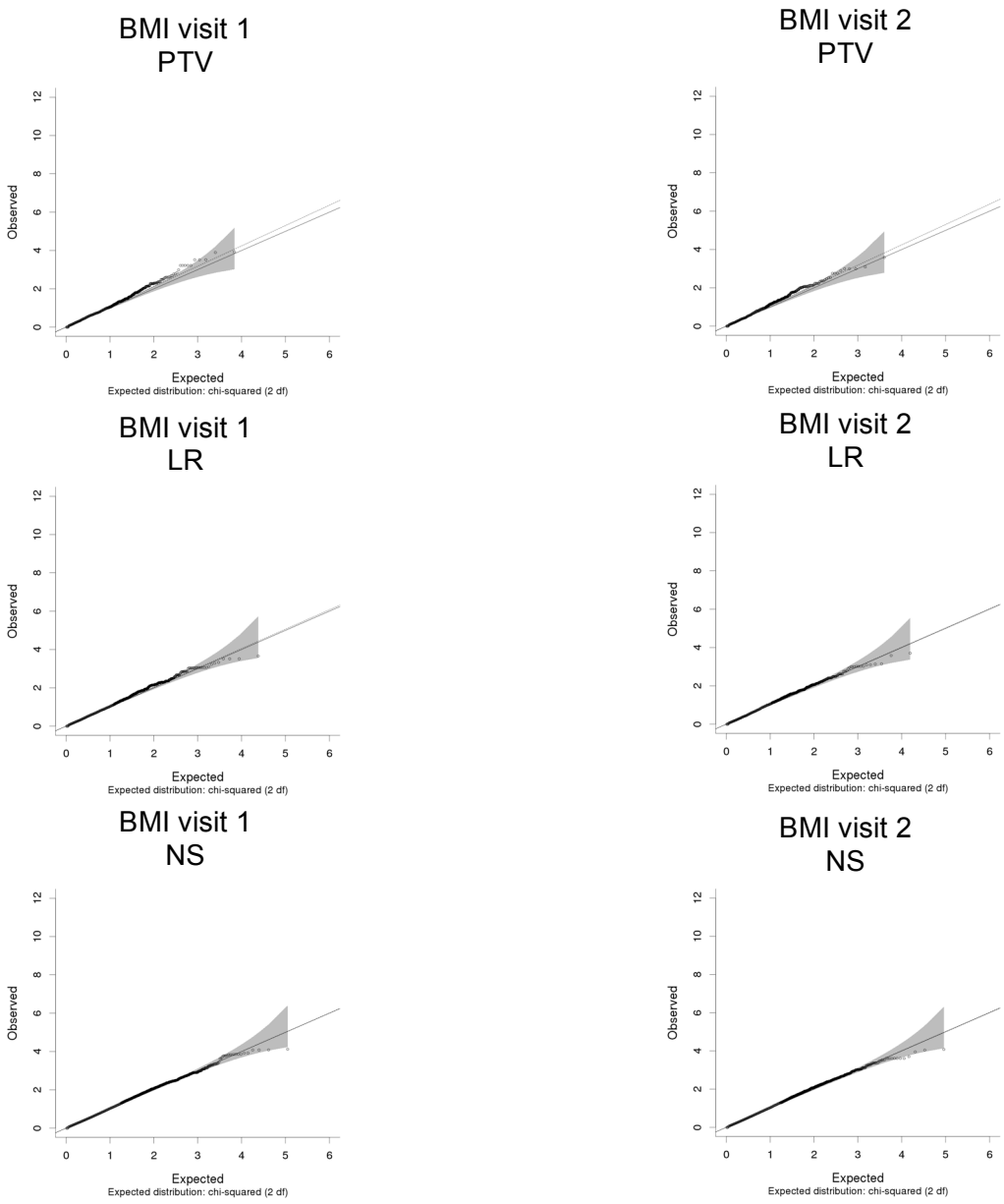


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR-PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

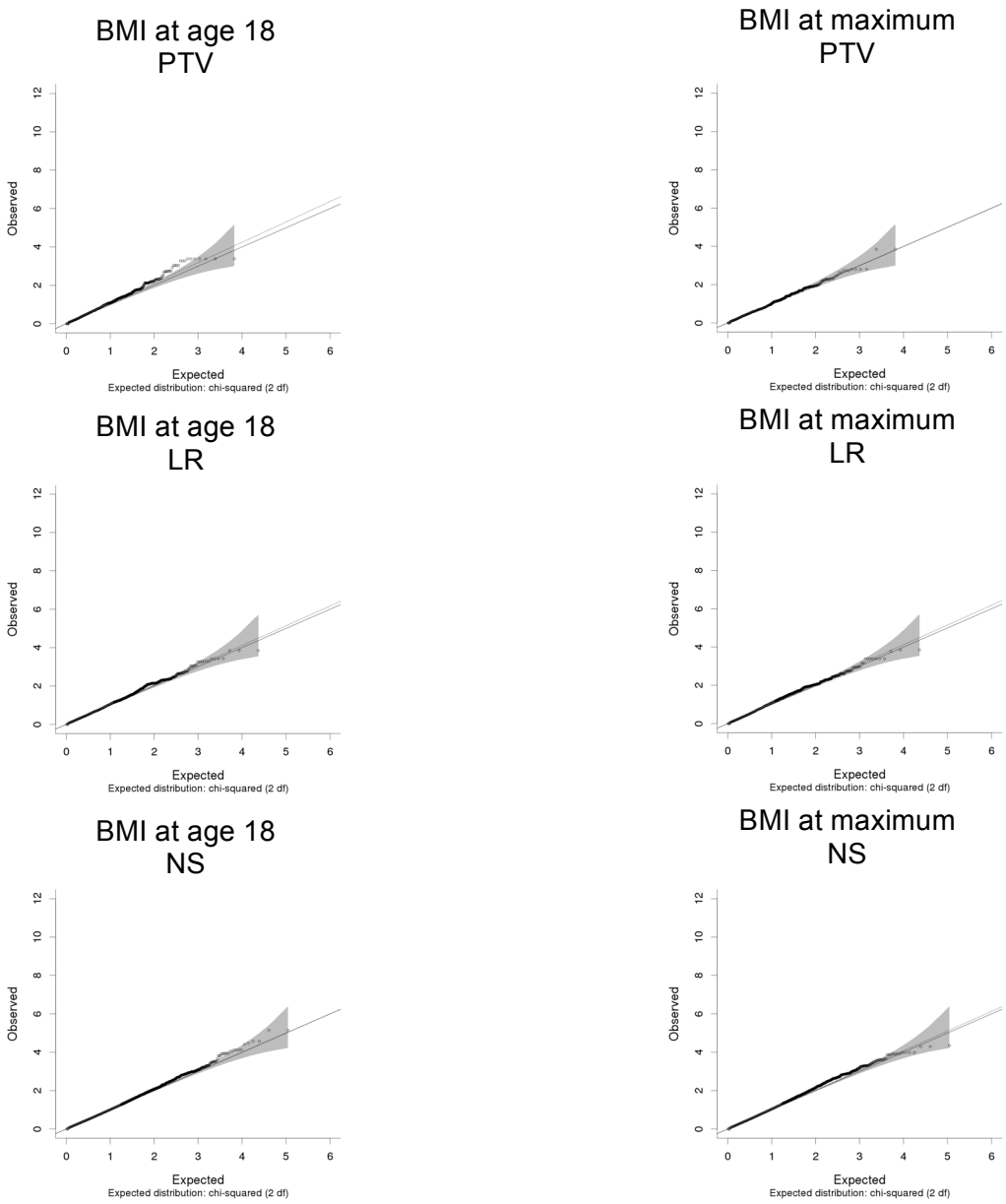


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR-PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

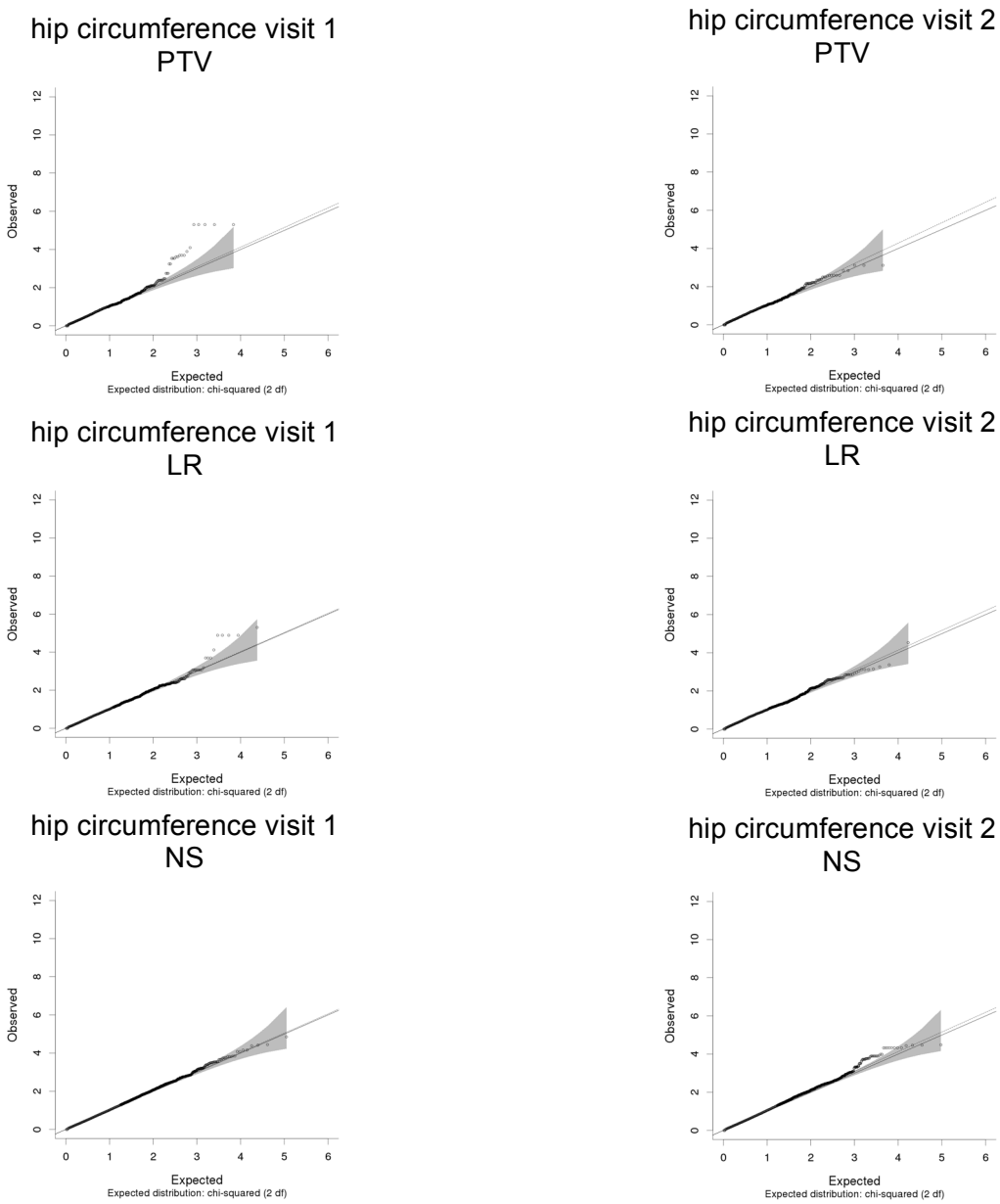


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

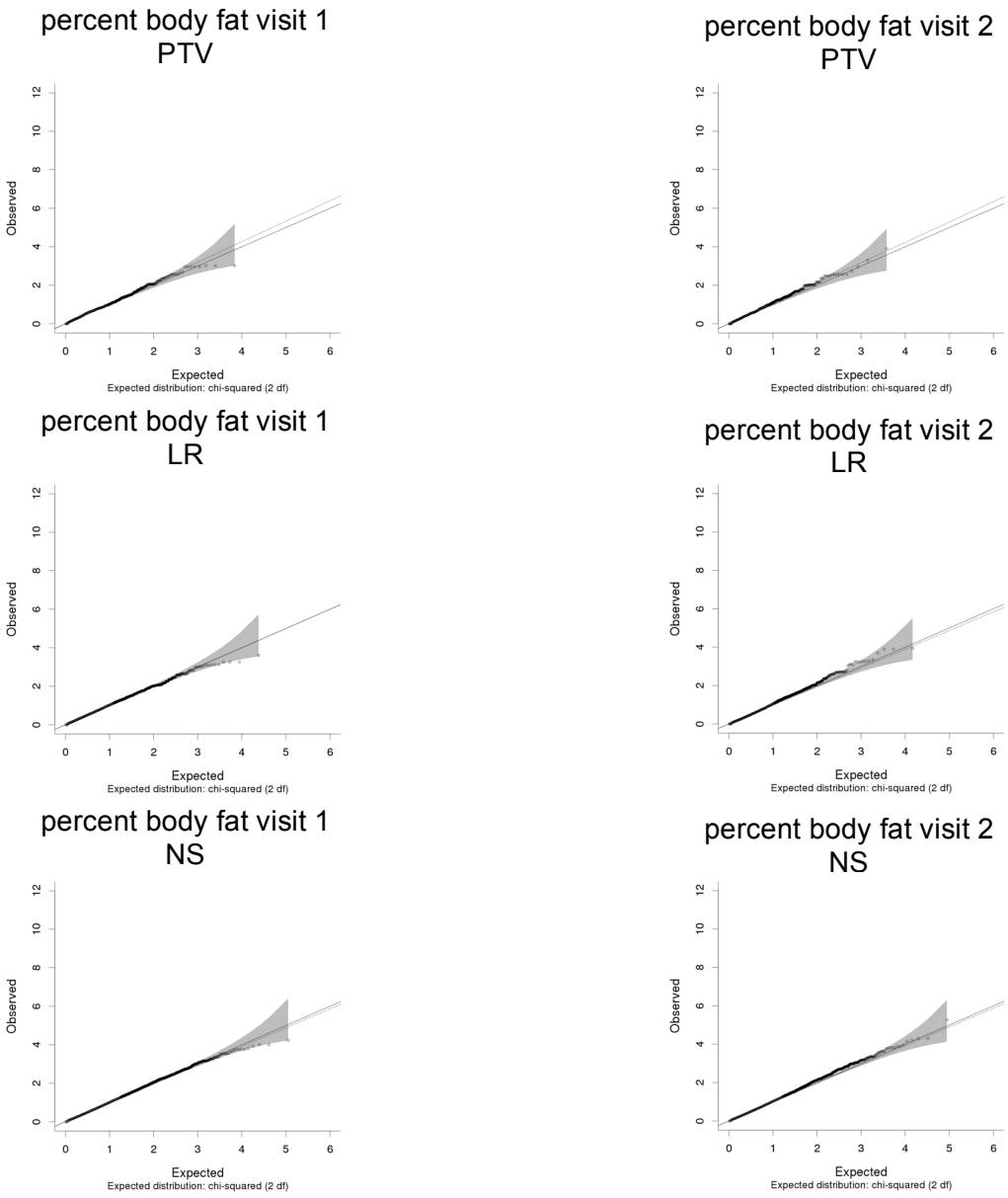


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

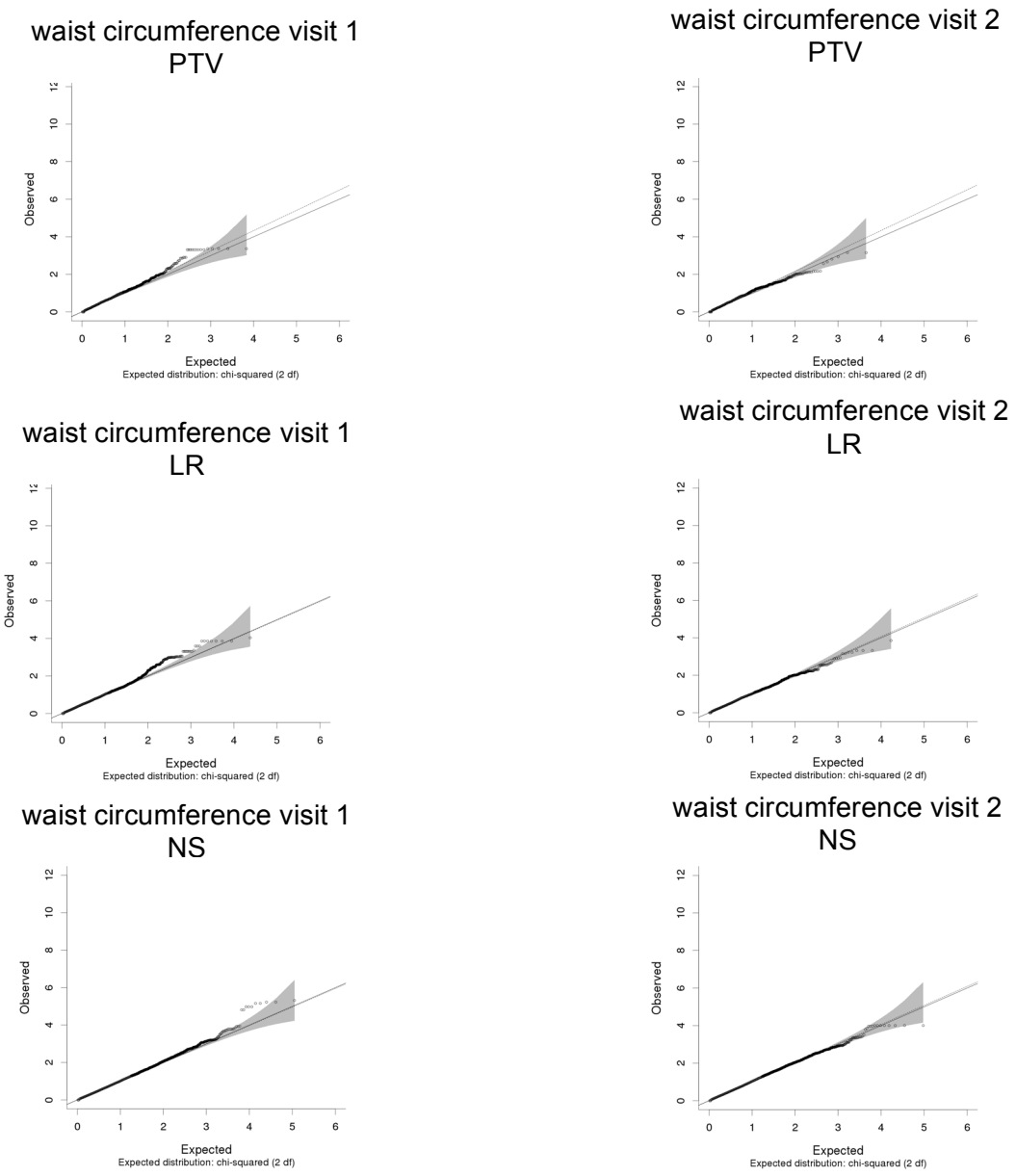


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR-PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

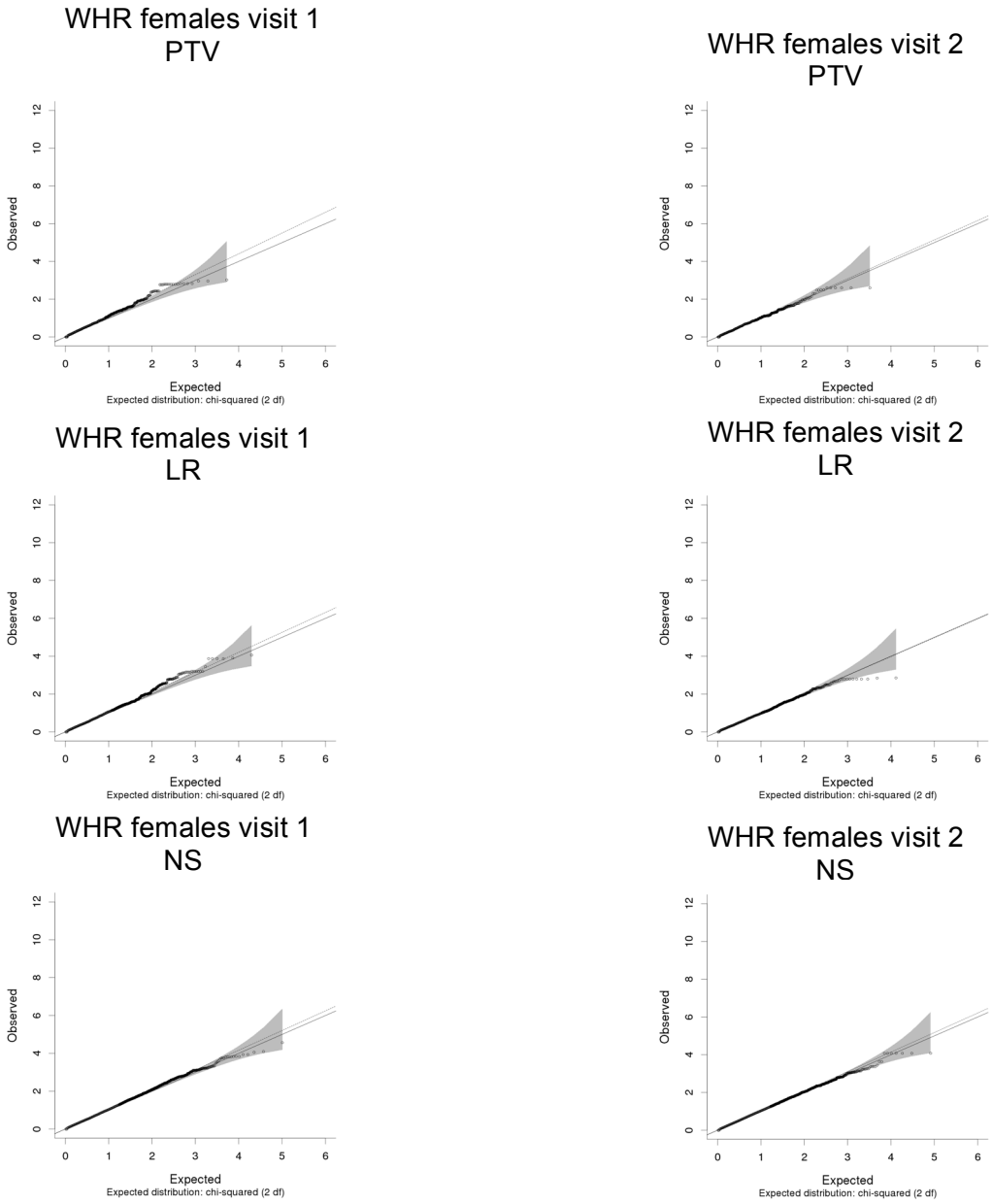


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

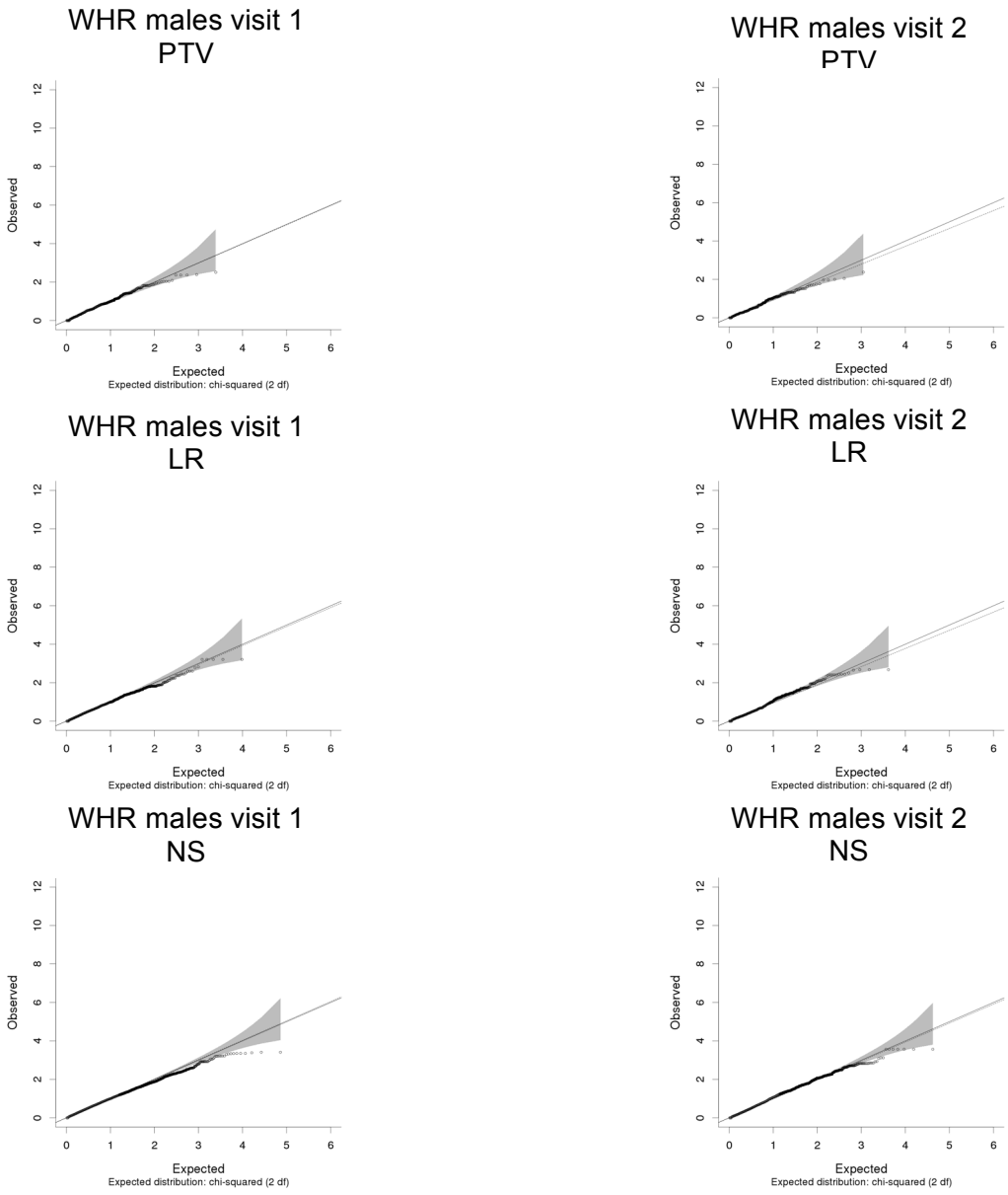


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

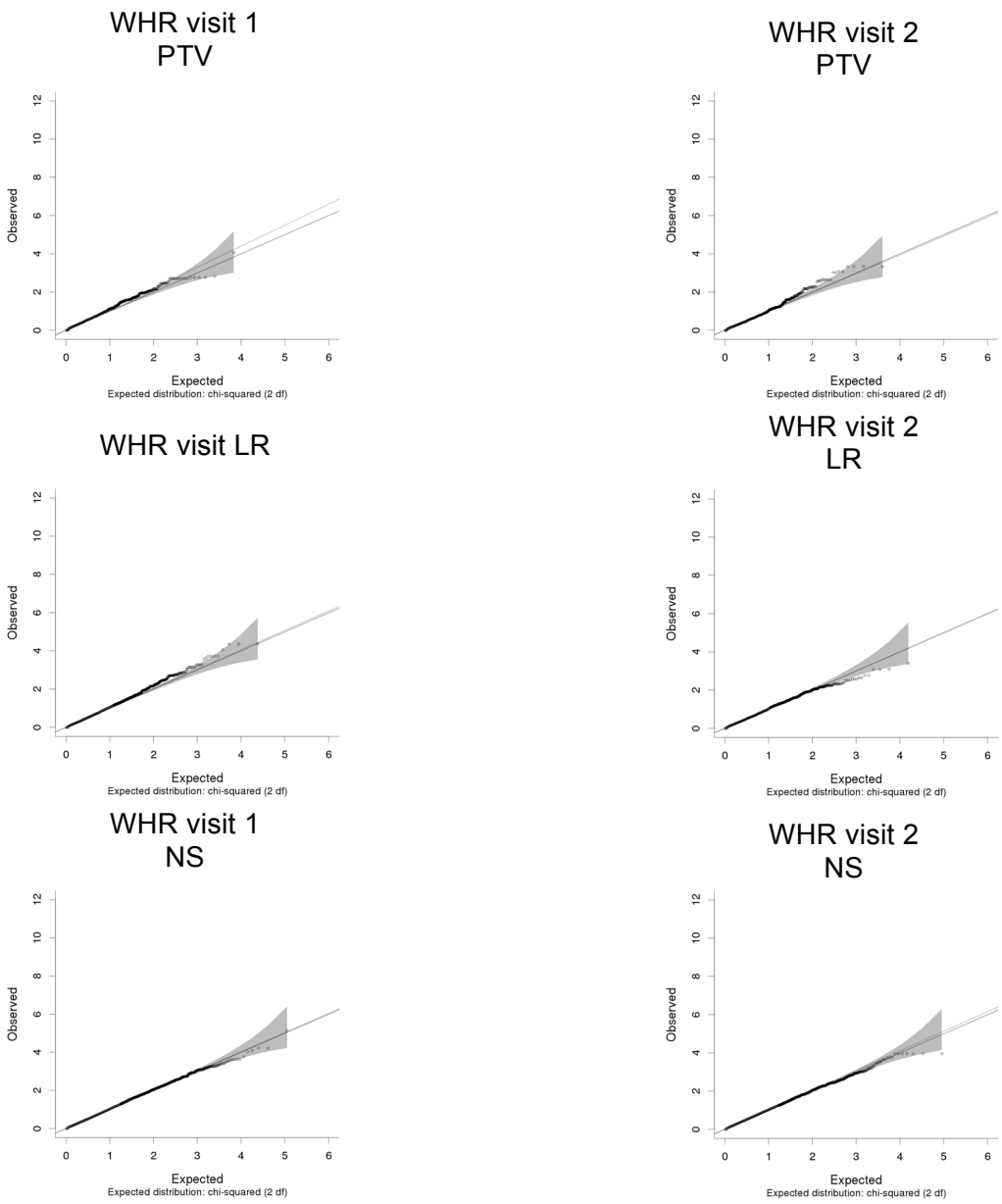


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR-PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

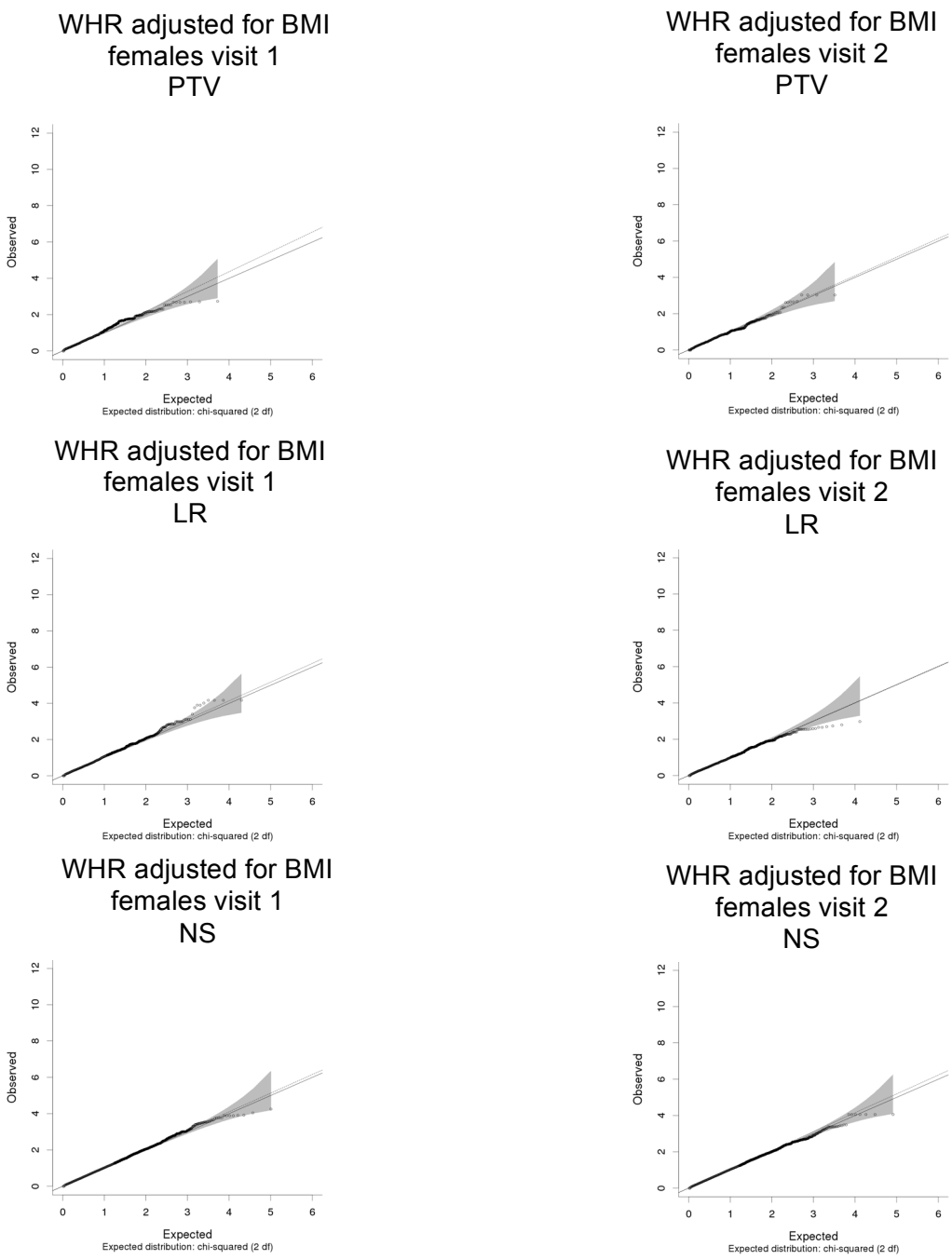


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR-PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

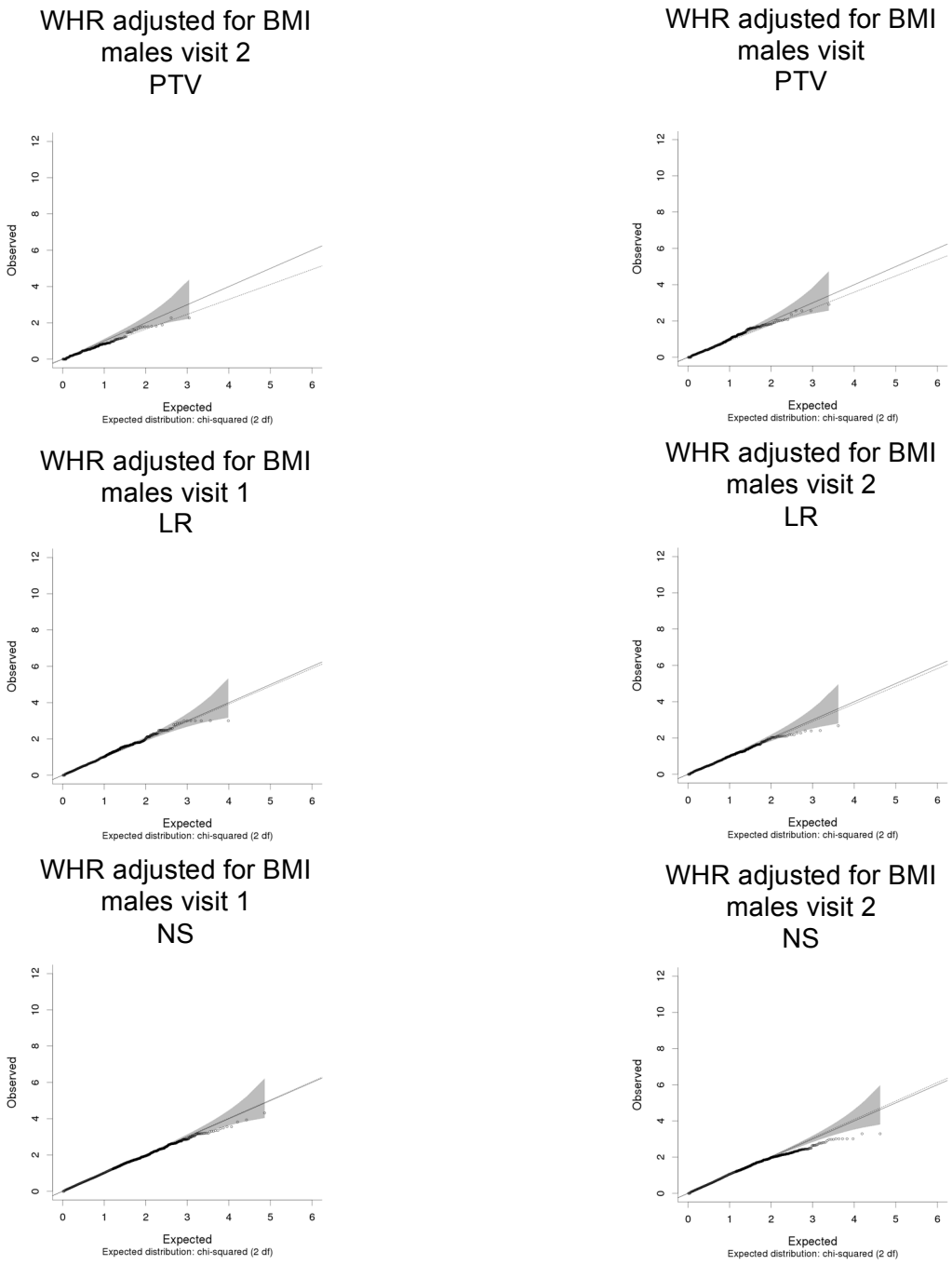


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR-PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

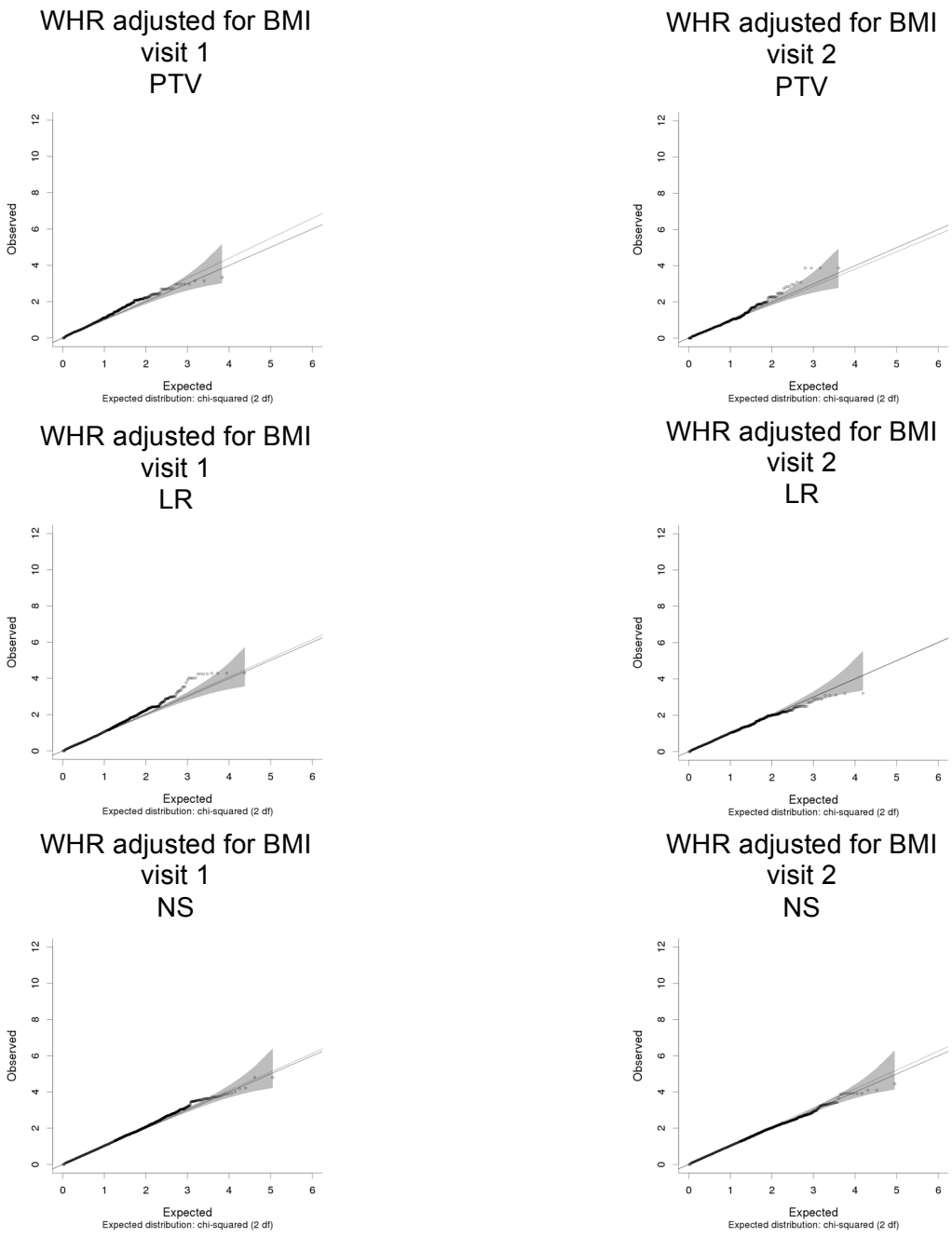


Figure 25. QQ plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

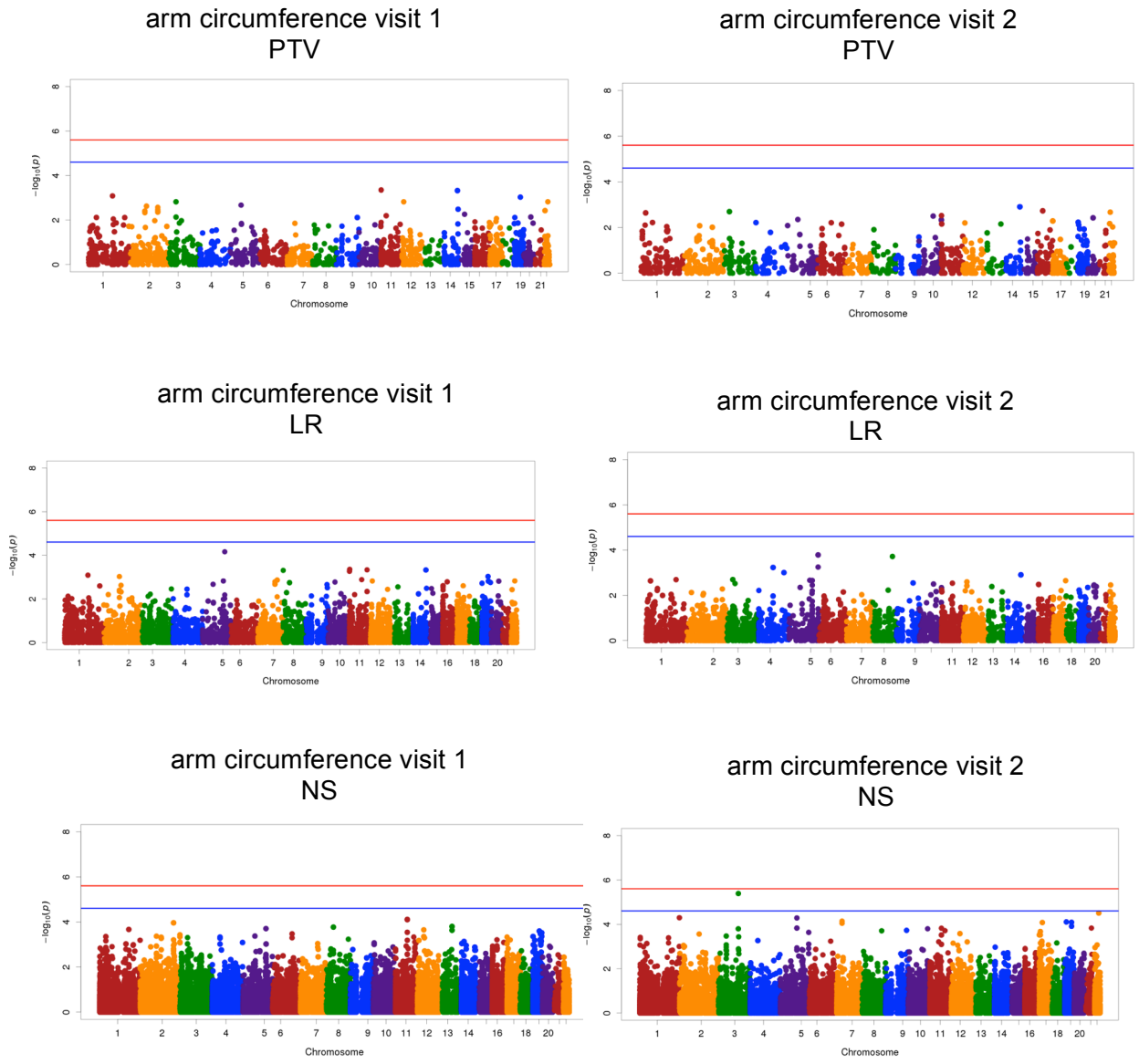


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

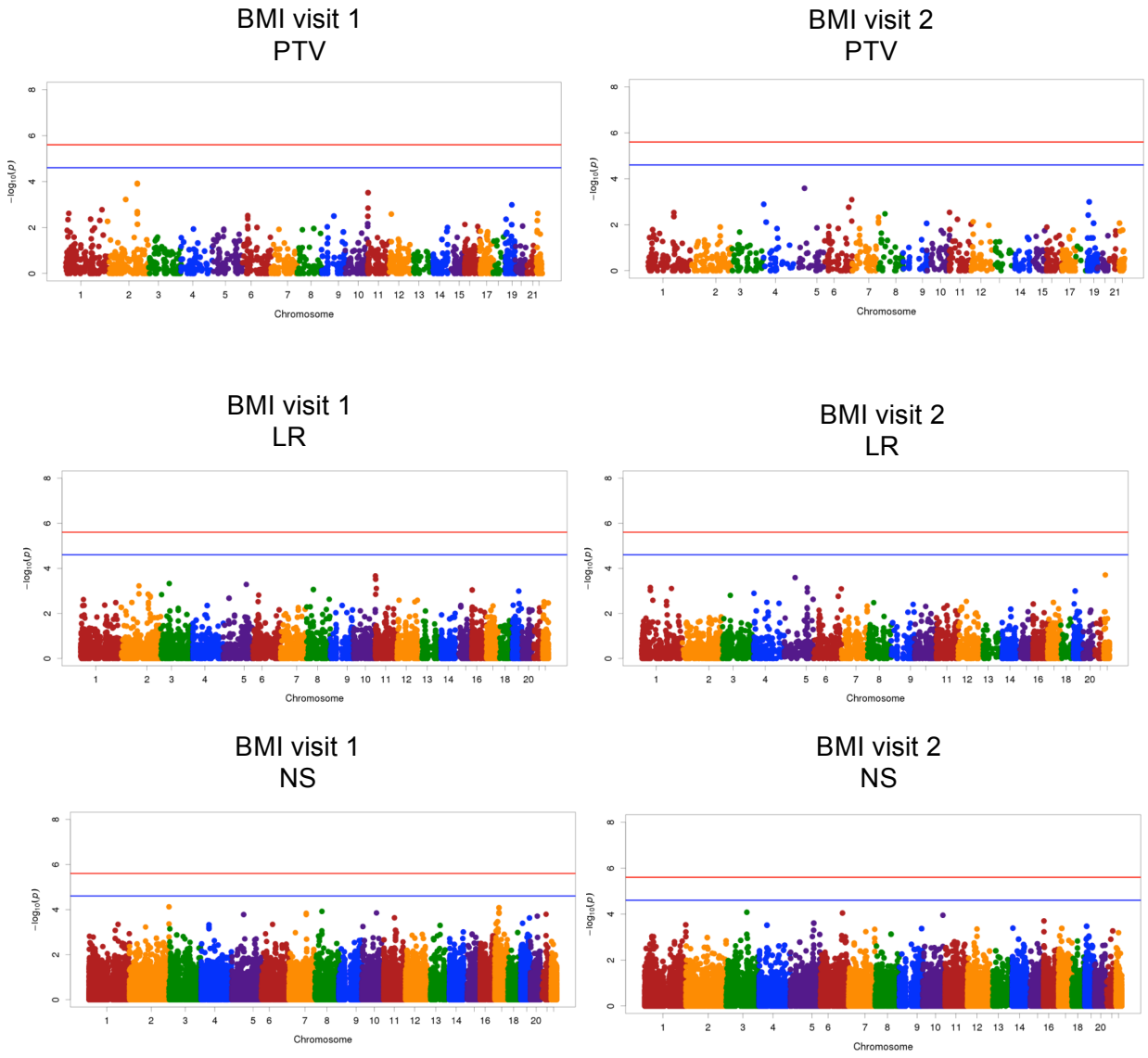


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

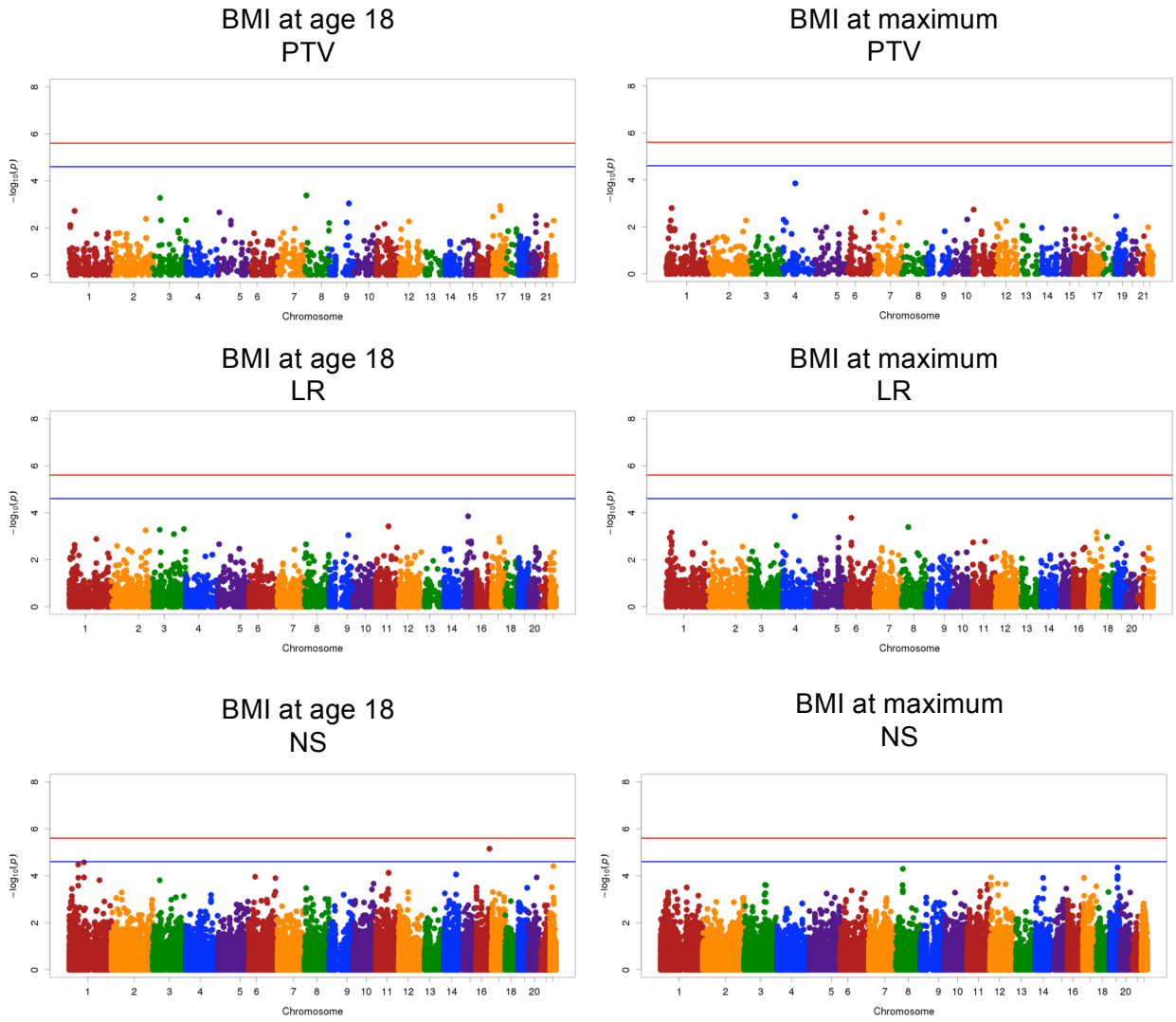


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

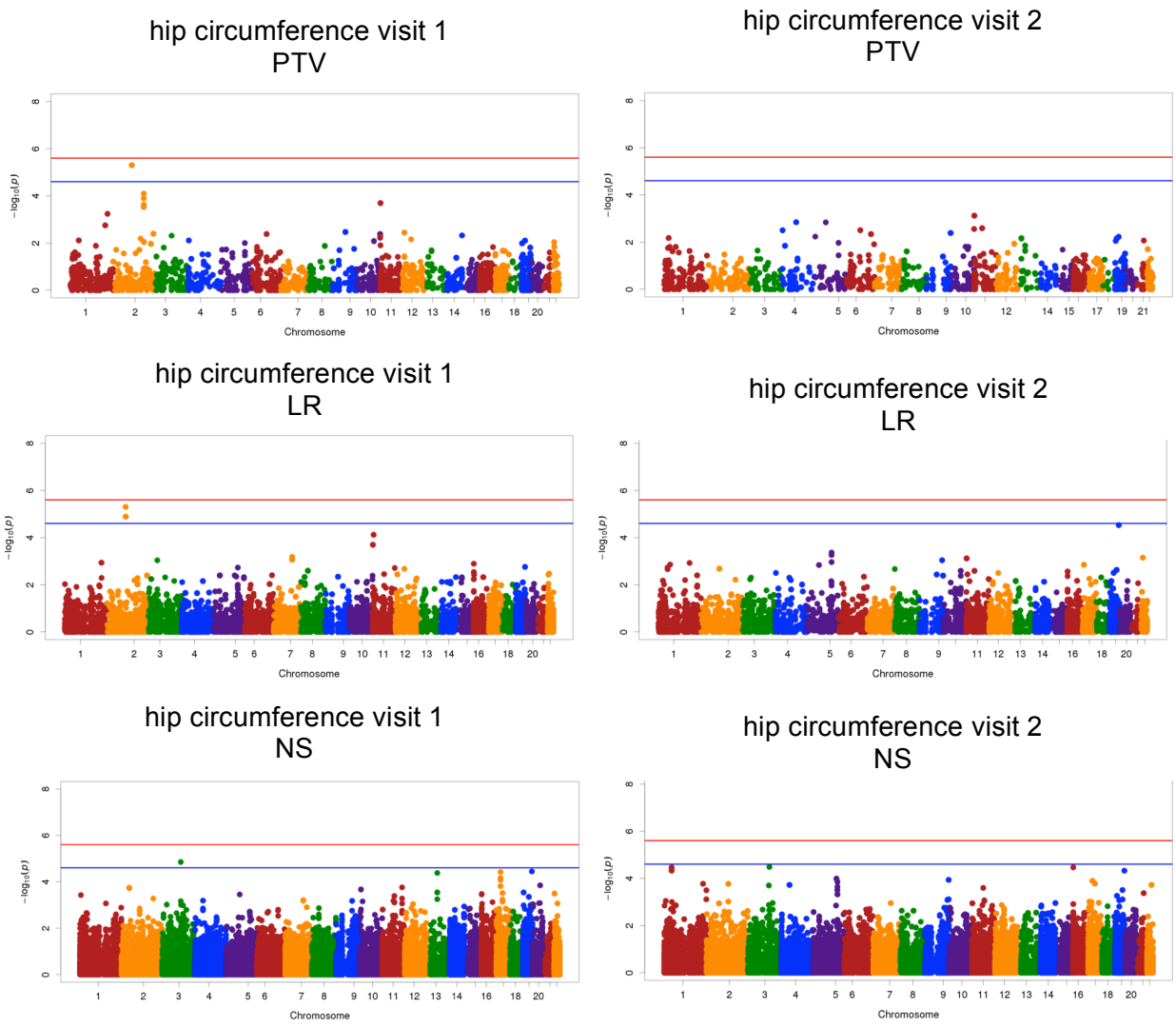


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

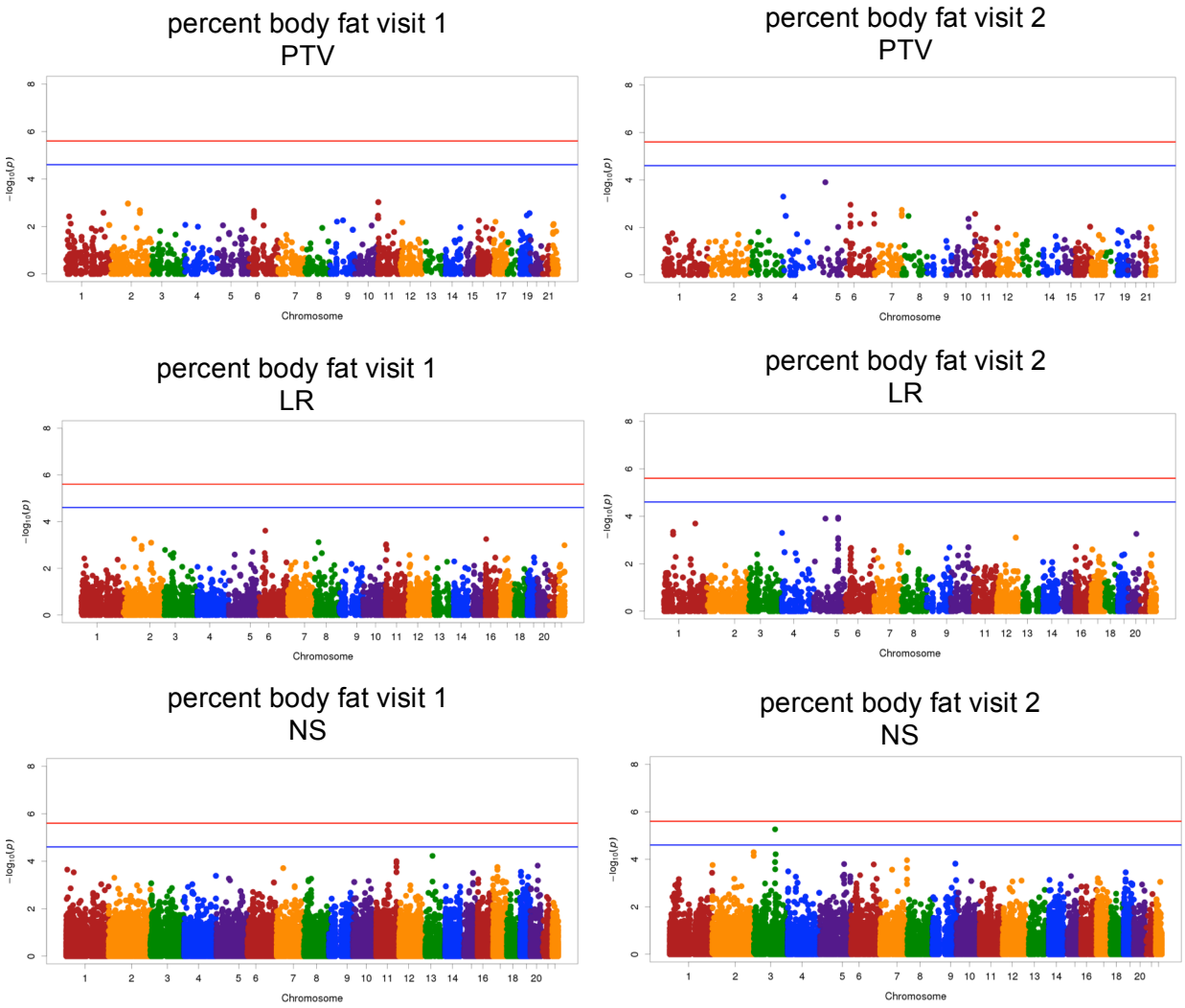


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

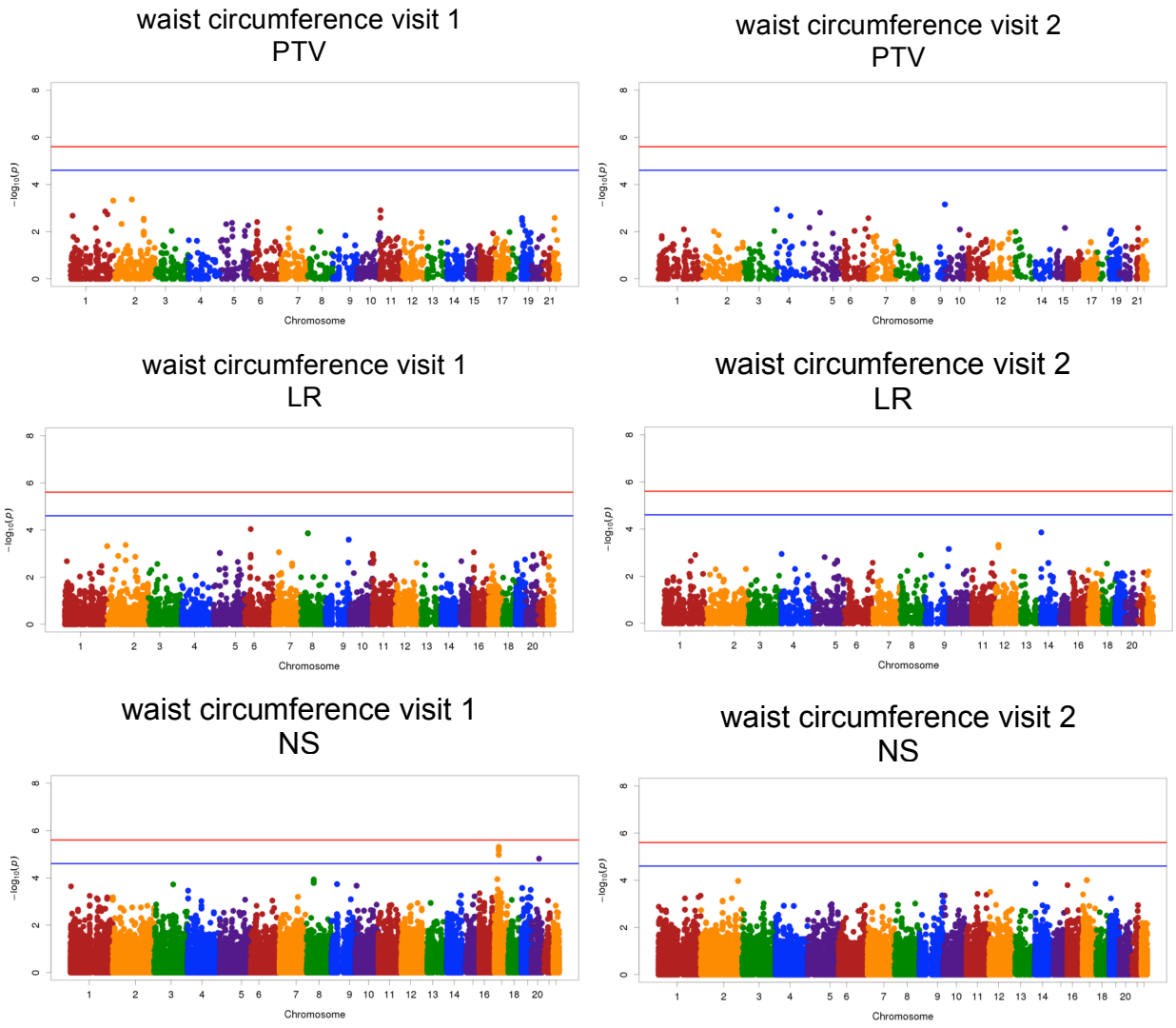


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

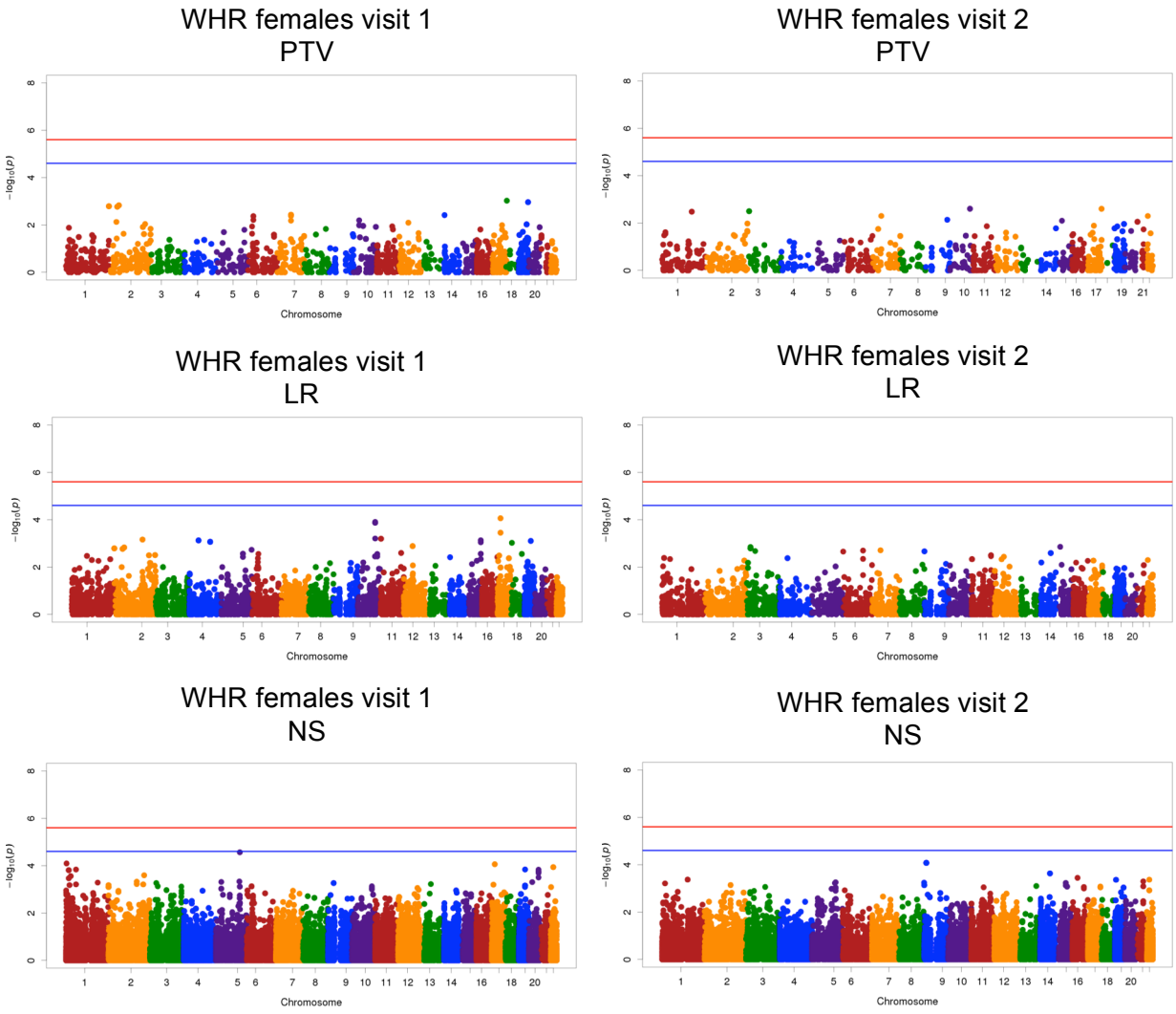


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

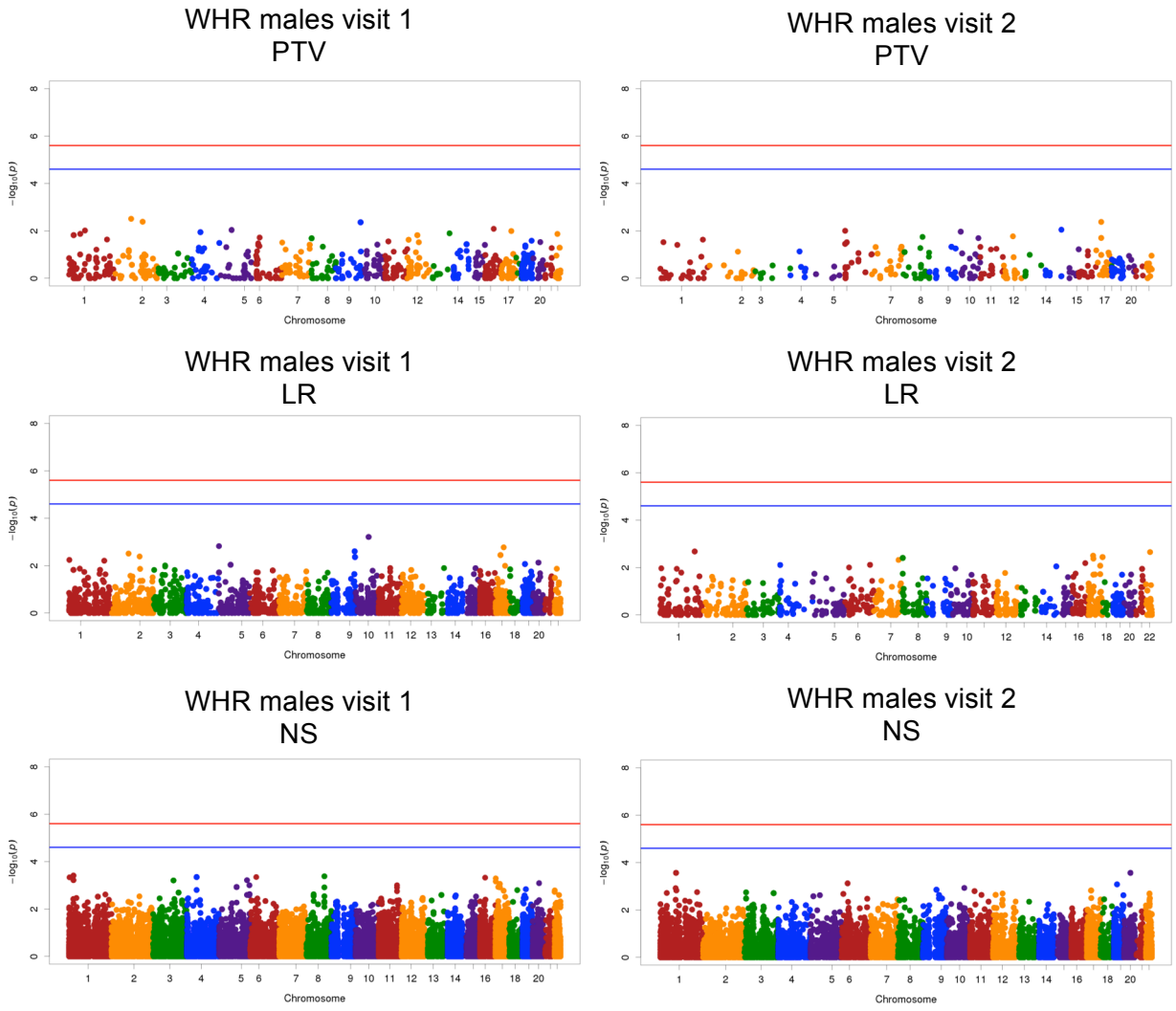


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

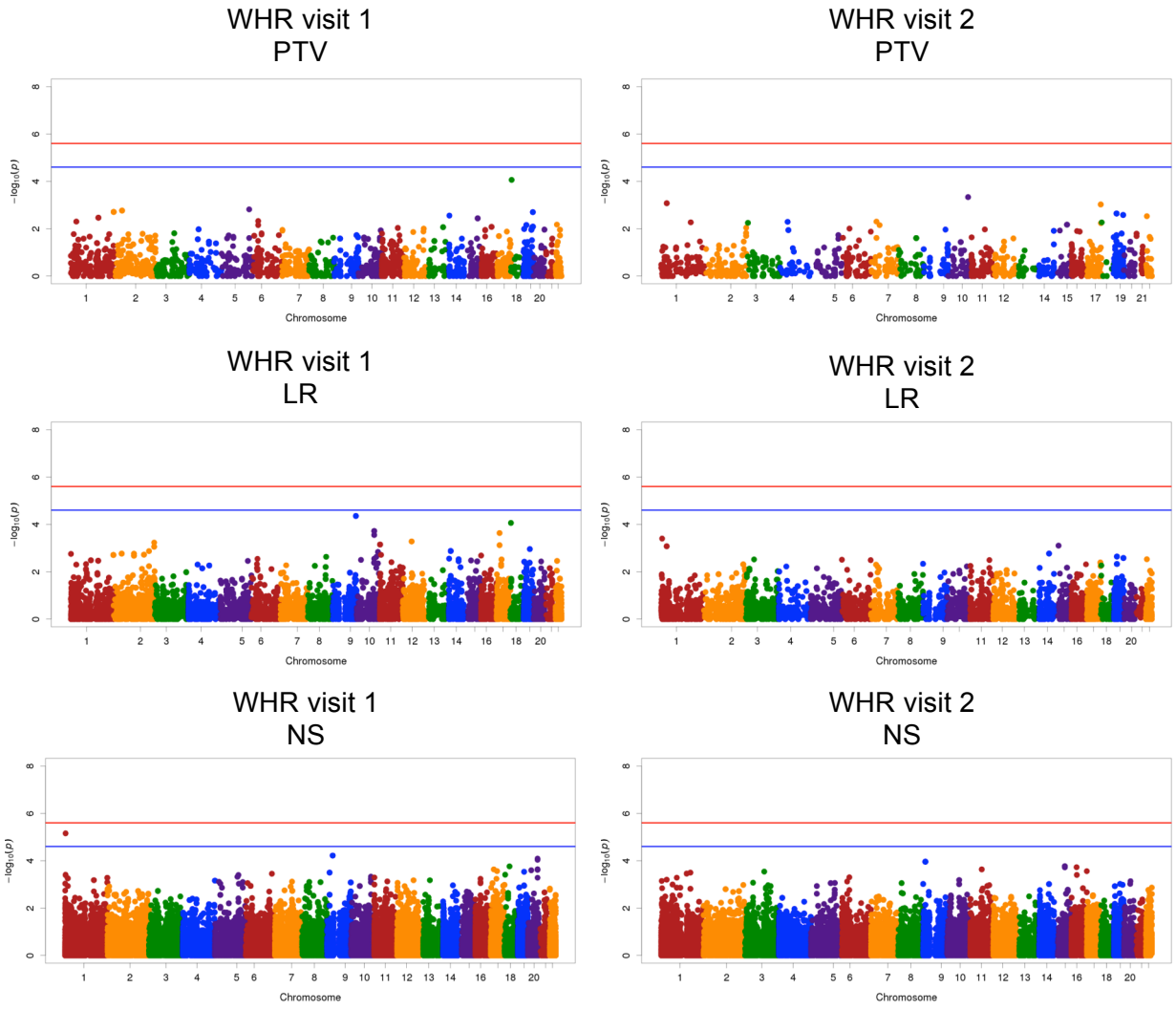


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

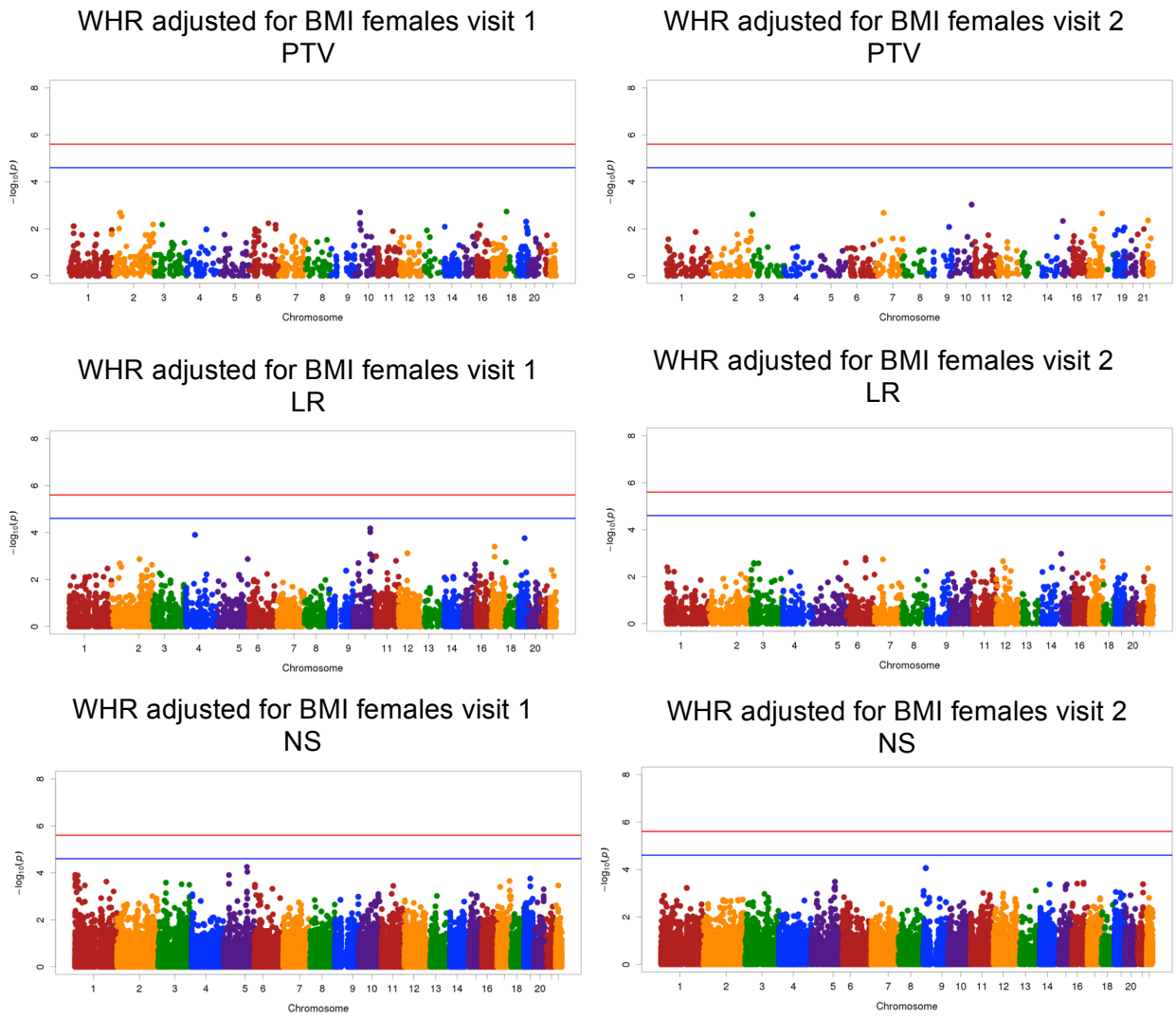


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

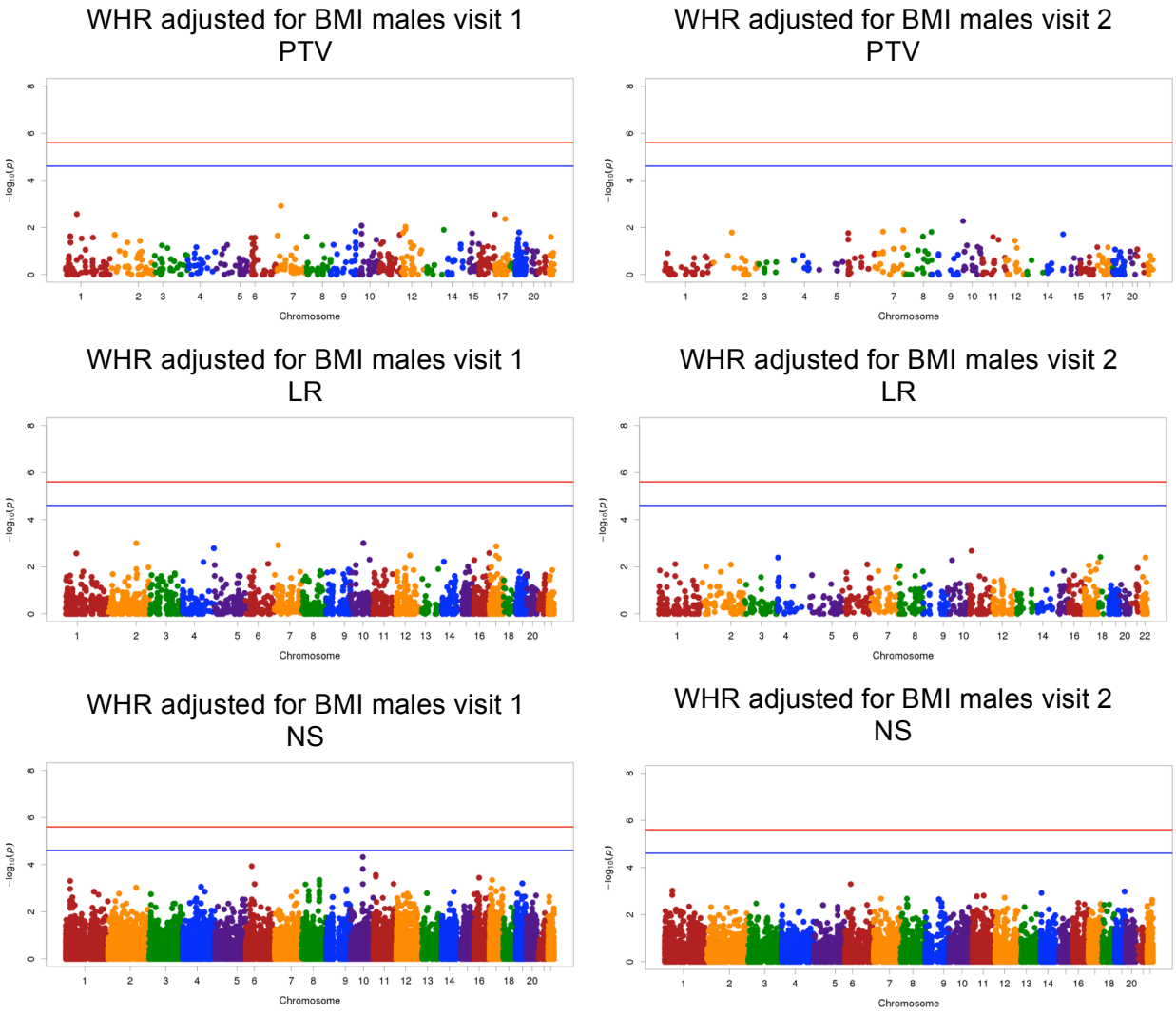


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

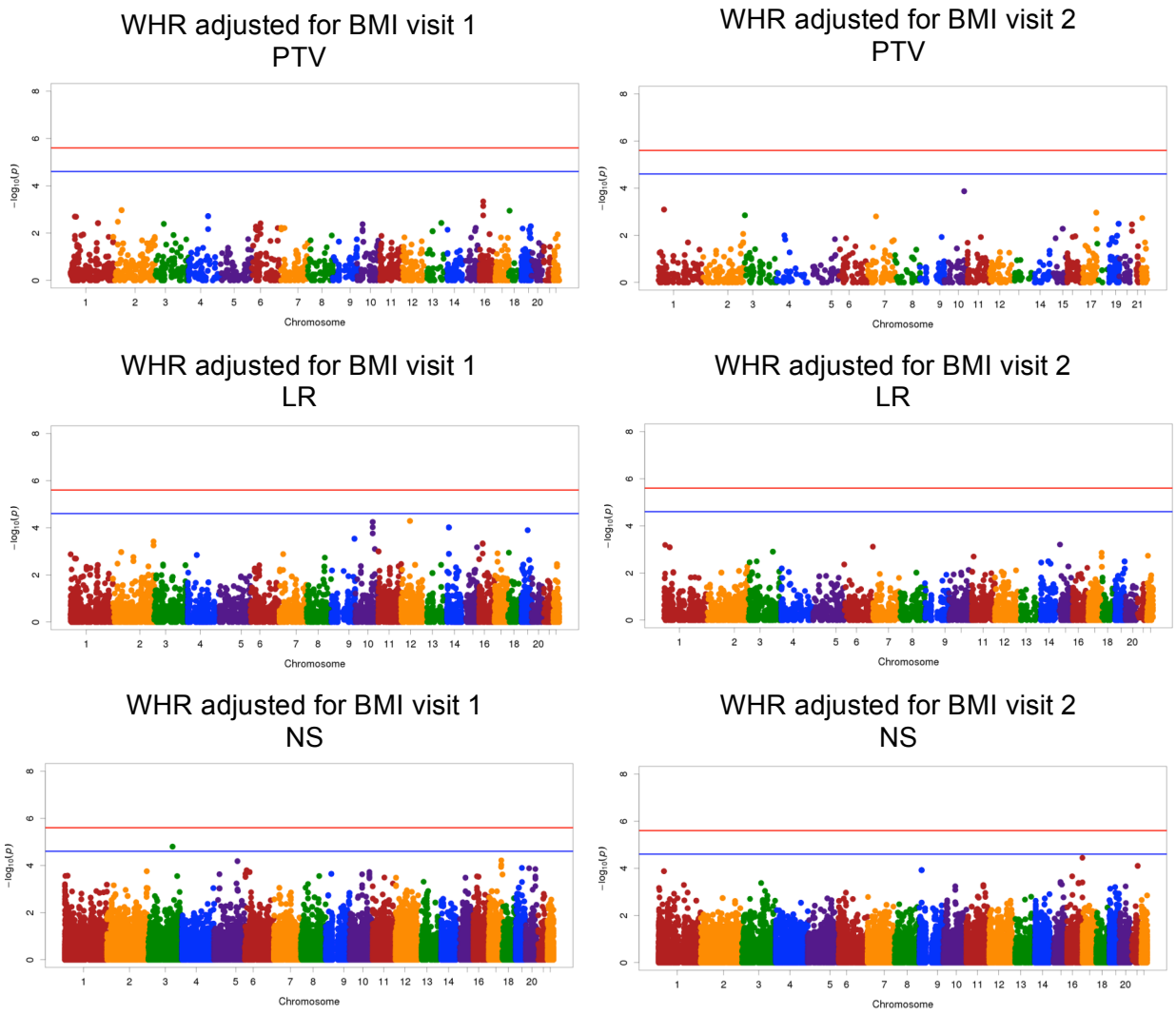


Figure 26. Manhattan plots for gene-based association tests. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS-PTV mask plus nonsynonymous variants with MAF < 1%.

trait	chr	start	end	transcript	gene	mask	visit 1			visit 2		
							P-value	# variants	cMAF	P-value	# variants	cMAF
WHR	1	3412521	3527764	ENST00000485002	*MEGF6*	PTV	0.25058	1	0.0079	0.0078885	1	0.0076
						LR	0.052417	6	0.0331	0.012905	4	0.0203
						NS	6.99×10^{-6}	27	0.0808	0.00071479	19	0.0734
hip	2	108910275	108924881	ENST00000251481	*SULT1C2*	PTV	5.00×10^{-6}	2	0.0092	0.0082261	1	0.0066
						LR	1.30×10^{-5}	3	0.0092	0.0020738	2	0.0066
						NS	0.0017541	9	0.0369	0.047537	7	0.0374
hip	3	112647783	112693652	ENST00000295863	*CD200R1*	PTV	NA	NA	NA	NA	NA	NA
						LR	NA	NA	NA	NA	NA	NA
						NS	1.41×10^{-5}	4	0.0119	0.0045153	4	0.0110
arm	3	124114153	124165606	ENST00000393501	*KALRN*	PTV	0.52068	1	0.0013	0.16982	1	0.0022
						LR	0.52068	1	0.00131	0.16982	1	0.0022
						NS	0.19228	3	0.0092	4.11×10^{-6}	3	0.0109
percent body fat	3	124114153	124165606	ENST00000393501	*KALRN*	PTV	0.52528	1	0.0013	0.091004	1	0.0027
						LR	0.52528	1	0.0013	0.091004	1	0.0027
						NS	0.047177	3	0.0092	5.48×10^{-6}	3	0.0134
WHR adjusted for BMI	3	142443450	142524981	ENST00000476941	*TRPC1*	PTV	NA	NA	NA	NA	NA	NA
						LR	NA	NA	NA	NA	NA	NA
						NS	1.59×10^{-5}	6	0.0159	0.25612	4	0.0203
BMI at 18	16	78056542	78064574	ENST00000575655	*CLEC3A*	PTV	NA	NA	NA			
						LR	NA	NA	NA			
						NS	7.04×10^{-6}	3	0.0053			
waist	17	25621468	25633916	ENST00000581185	*WSB1*	PTV	NA	NA	NA	NA	NA	NA
						LR	0.46123	2	0.0026	NA	0	NA
						NS	6.97×10^{-6}	5	0.0119	9.95×10^{-5}	2	0.0087912
waist	20	44511257	44512293	ENST00000372520	*ZSWIM1*	PTV	0.000015736	1	0.0369	0.032595	1	0.0373
						LR	0.000015736	1	0.0369	0.032595	1	0.0373
						NS	1.55×10^{-5}	3	0.0408	0.034529	3	0.0418

Table 17. Suggestive gene-based associations with traditional obesity measures.

Composite adiposity measures: In addition to looking at the traditional measures of obesity, we also tested for association between transcripts and the composite measures of adiposity, PC2, which captures adiposity above or below the waist, and PC3, which captures central versus peripheral adiposity. Three genes are suggestively associated with these composite measures (Figures 27 and 28, Table 18). This includes an association with *Long chain acyl-CoA synthetase 1 (ACSL1)* with PC3 at the first visit. At the second visit the association persists but is slightly weaker.

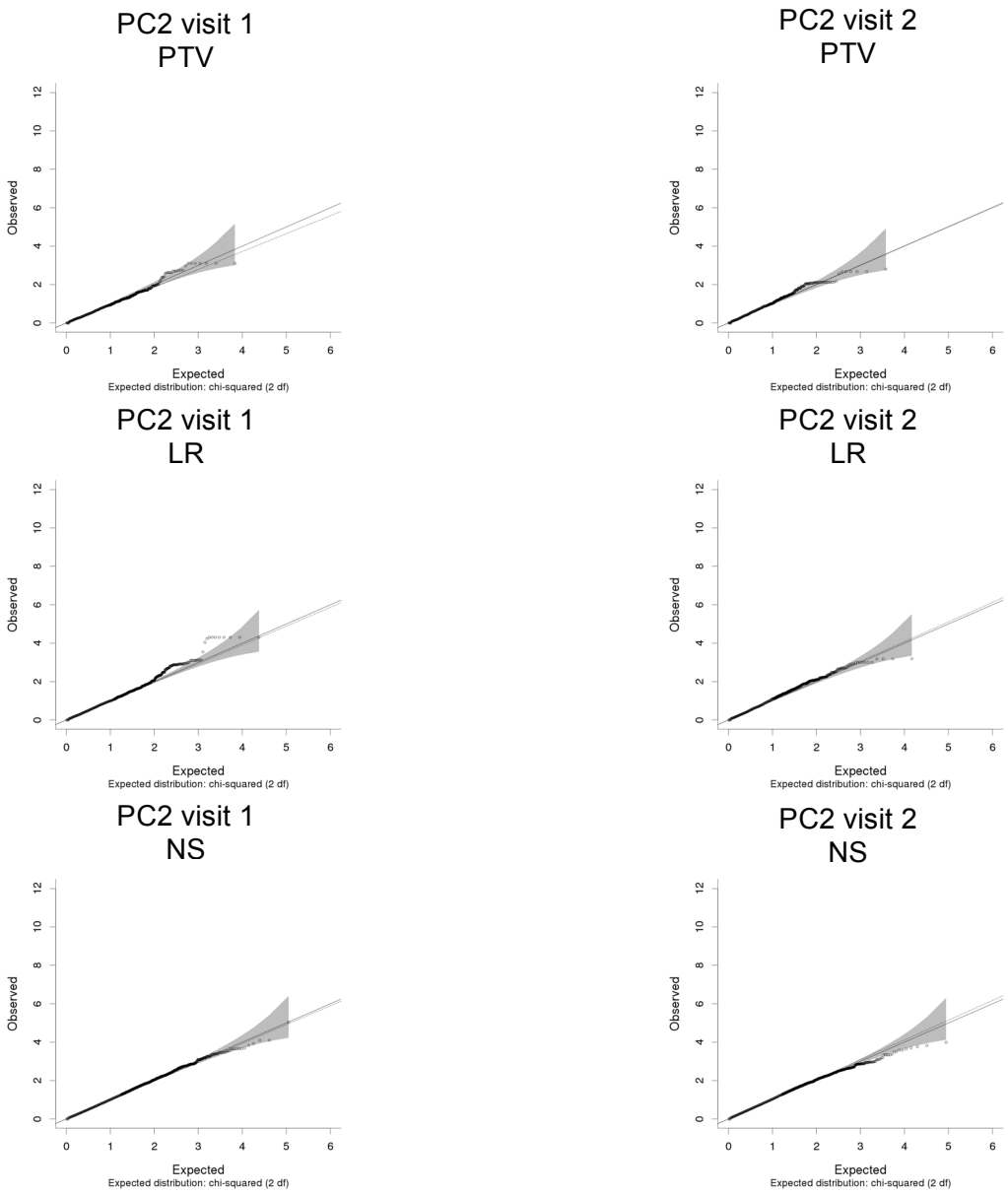


Figure 27. QQ plots for gene-based association tests with composite measures of obesity. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

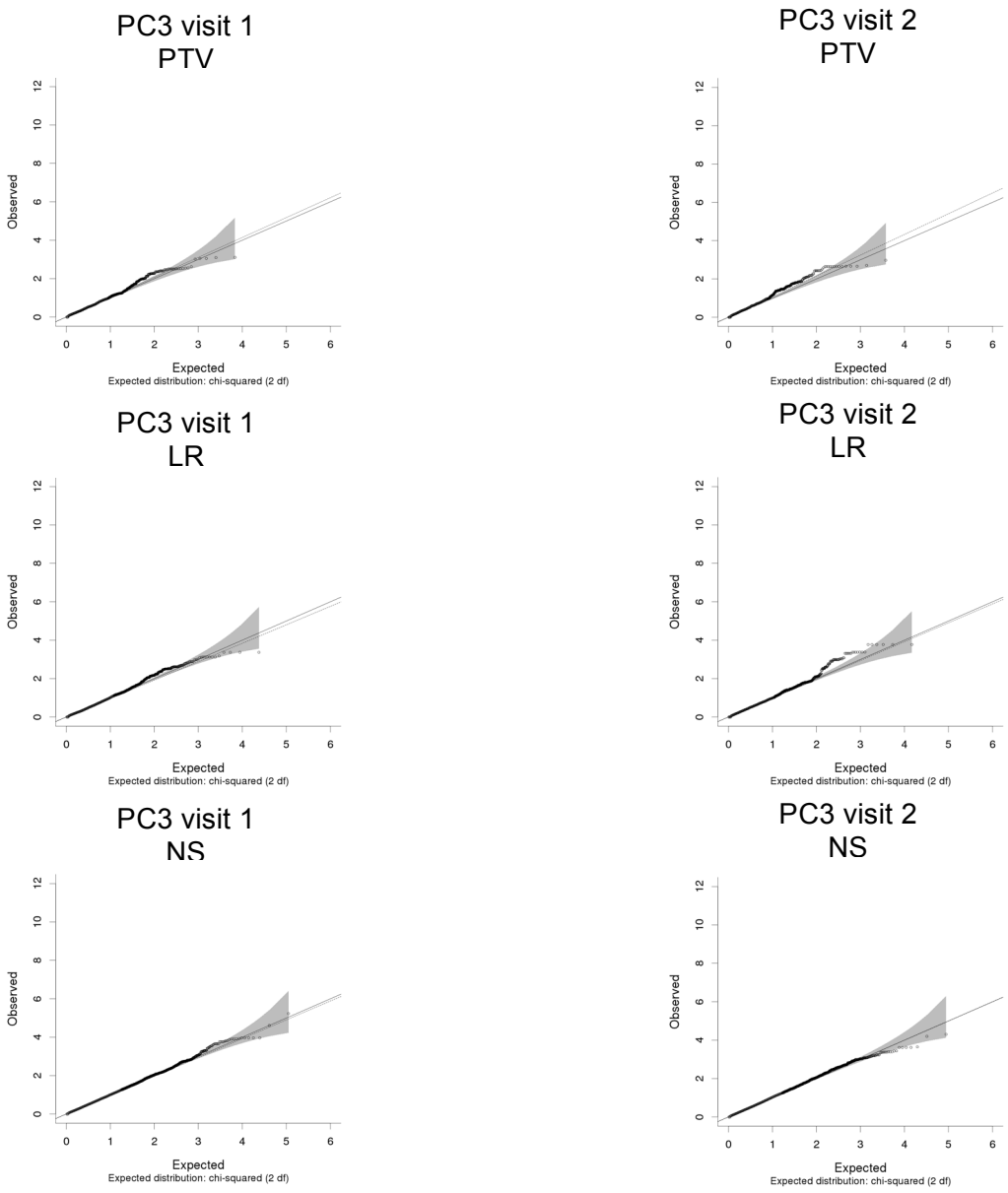


Figure 27. QQ plots for gene-based association tests with composite measures of obesity. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

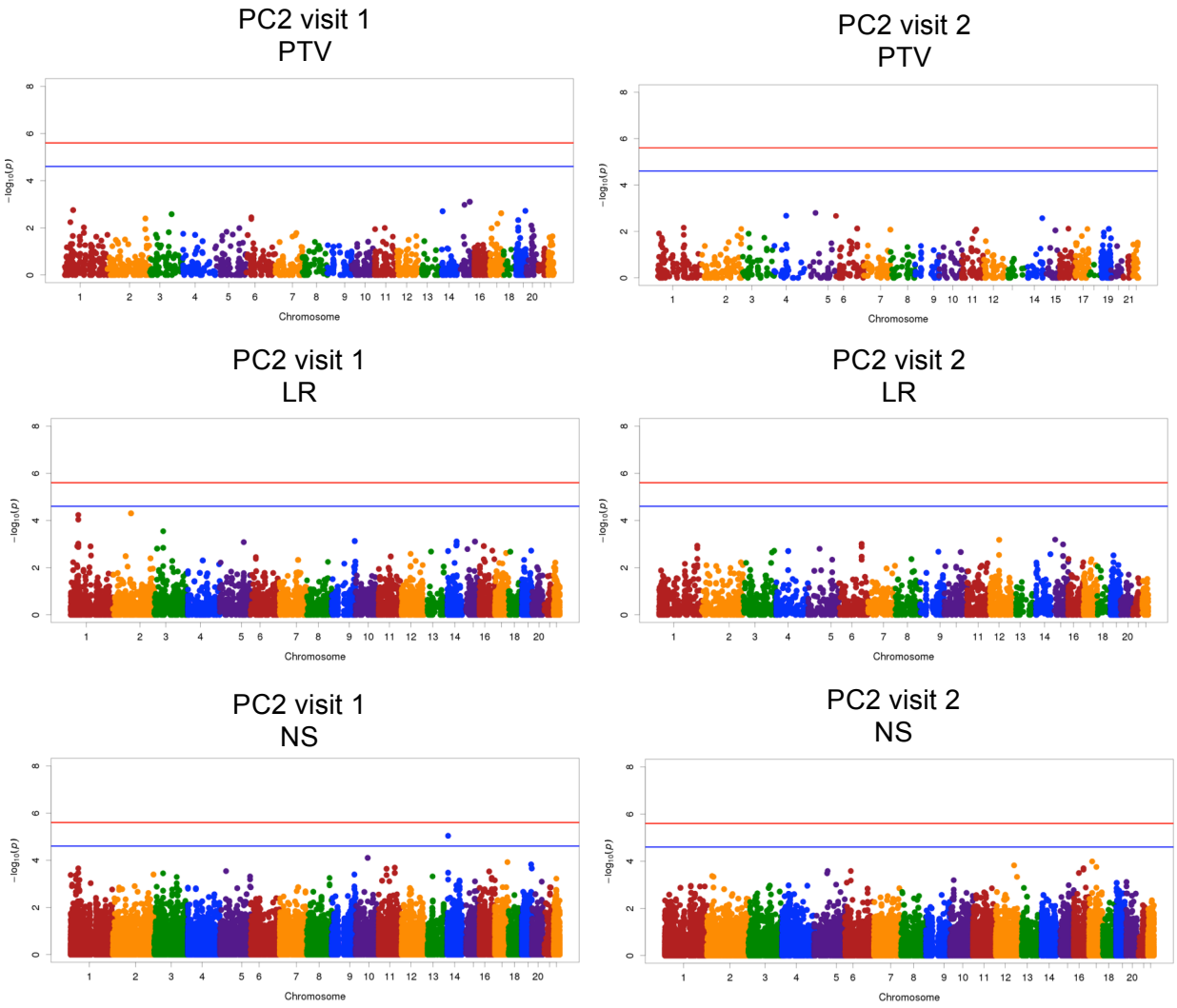


Figure 28. Manhattan plots for gene-based association tests with composite measures of obesity. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

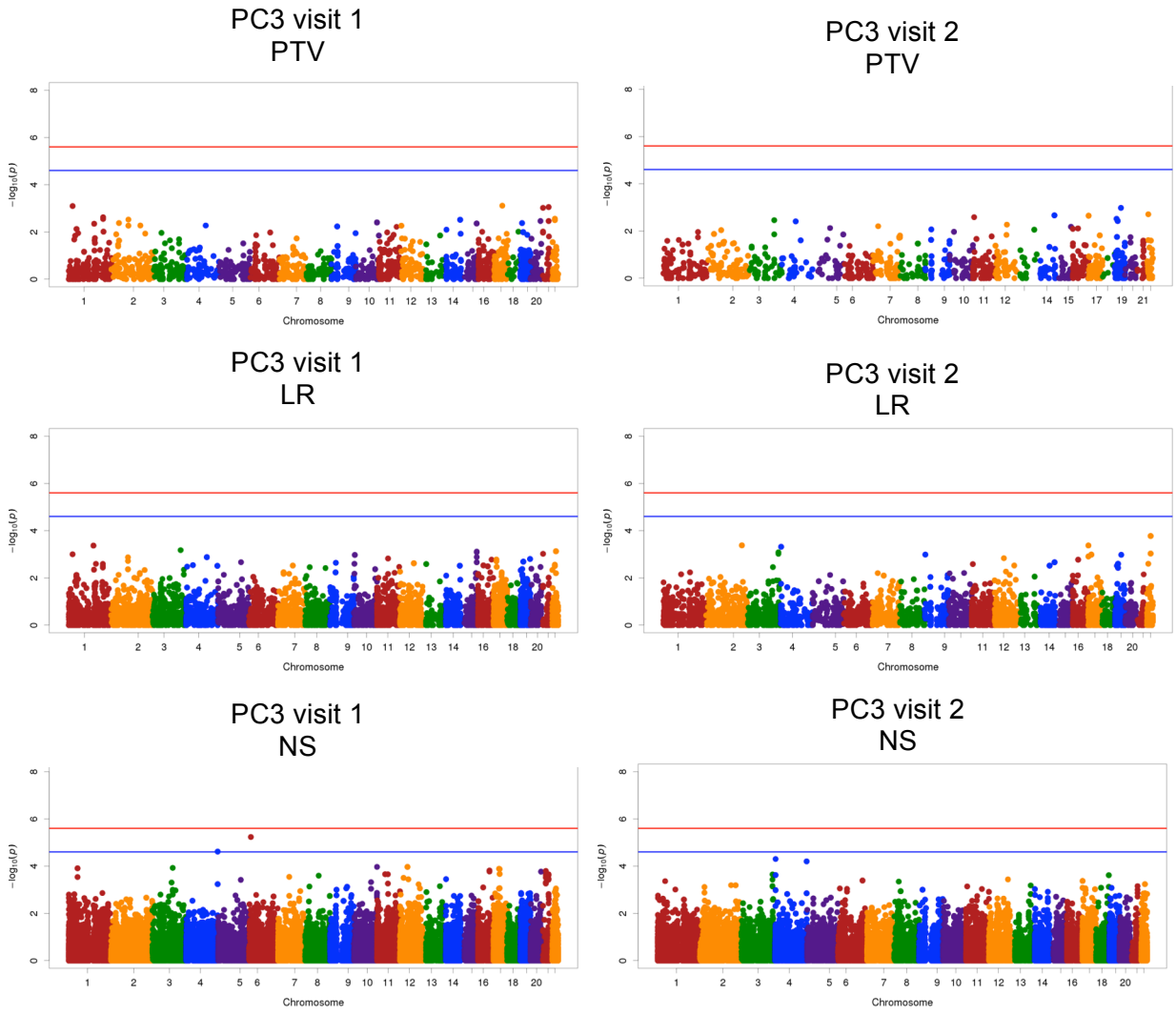


Figure 28. Manhattan plots for gene-based association tests with composite measures of obesity. Associations for each of three masks are shown. PTV- protein truncating variants with MAF < 5%, LR- PTV mask plus nonsynonymous variants with MAF < 1%. predicted damaging by metaLR, and NS- PTV mask plus nonsynonymous variants with MAF < 1%.

trait	chr	start	end	transcript	gene	mask	visit 1			visit 2		
							P-value	# variants	cMAF	P-value	# variants	cMAF
PC3	4	185678318	185678403	ENST00000503407	*ACSL1*	PTV	NA	NA	NA	NA	NA	NA
						LR	NA	NA	NA	NA	NA	NA
						NS	2.42×10^{-5}	3	0.0053	6.3×10^{-5}	2	0.0081
PC3	6	10874532	10877397	ENST00000379491	*GCM2*	PTV	NA	NA	NA	NA	NA	NA
						LR	0.00328	1	0.0027	0.00444	1	0.0027
						NS	5.88×10^{-6}	7	0.0425	0.000885	5	0.0432
PC2	14	24675130	24677343	ENST00000428351	*TSSK4*	PTV	0.00198	2	0.0106	0.608	2	0.0135
						LR	0.00198	2	0.0106	0.608	2	0.0135
						NS	9.21×10^{-6}	5	0.0186	0.0924	5	0.0216

Table 18. Suggestive gene-based associations with composite adiposity measures.

Gene-based multi-trait analysis: As with single variant analyses, transcripts were analyzed across the array of obesity related traits. The test statistics were well calibrated to a Gamma distribution as shown in the QQ plots in Figure 29; the corresponding Manhattan plots are in Figure 30. Three transcripts in the NS variant grouping attained exome-wide suggestive criteria ($p < 2.5 \times 10^{-5}$) (Table 19). None of these transcripts were strongly associated with anyone one trait. Both *DLG4* and *INPP5F* are driven by many traits with weak associations. In contrast, *RUNX3* is nominally associated with only 3 traits, BMI and arm circumference at visit 2 and hip circumference at visit 1.

chr	start	end	transcript	gene	P-value
1	25254224	25256161	ENST00000308873	*RUNX3*	6.99×10^{-6}
10	121551155	121556996	ENST00000369083	*INPP5F*	2.42×10^{-5}
17	7107367	7107559	ENST00000447163	*DLG4*	5.41×10^{-6}

Table 19. Suggestive gene-based multi-trait associations. All suggestive associations come from the NS variant grouping.

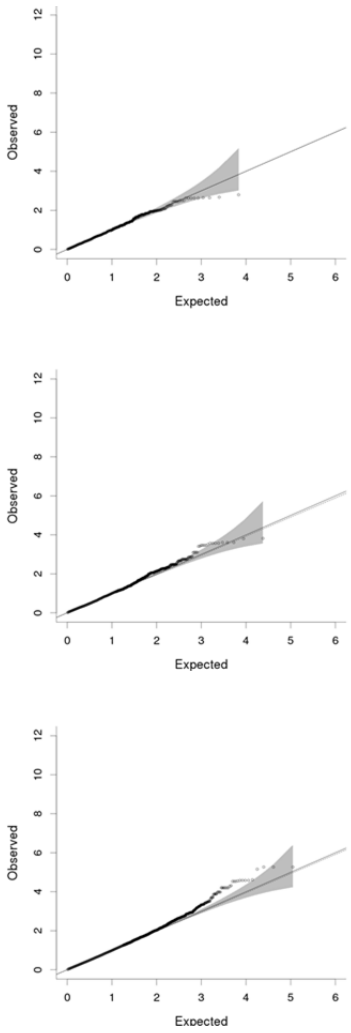


Figure 29. QQ plots for gene-based multi-trait association tests.

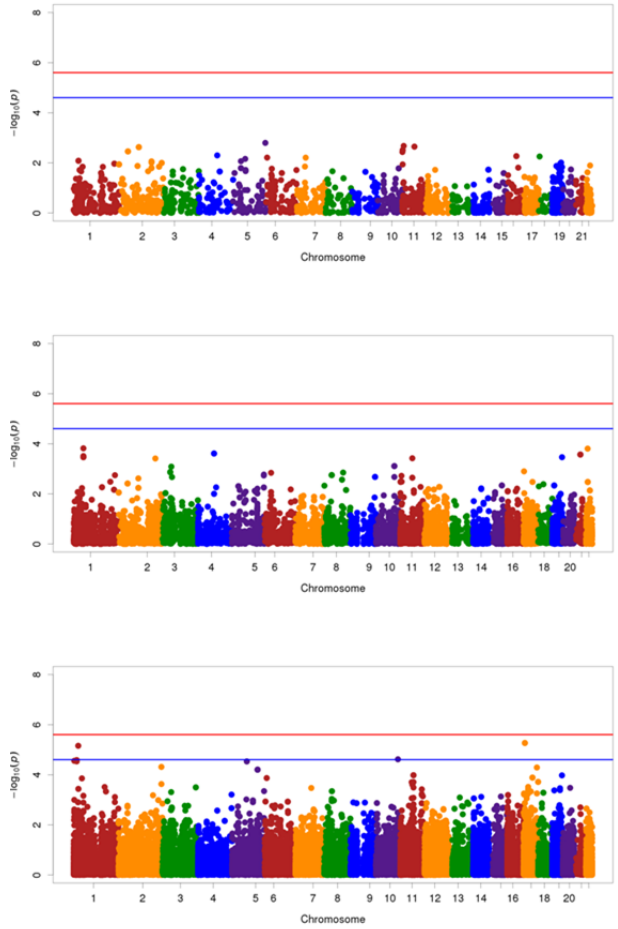


Figure 30. Manhattan plots for gene-based multi-trait association tests.

DISCUSSION

Both gene-based and single variant tests were generally well calibrated, however no transcripts and only four single variants attained exome-wide significance criteria across the 28 traits and multi-trait analyses. One of these rs117042905 in *ARHGAP39* is associated with WHR in males adjusted for BMI at the second visit. Due to the limited number of men without diabetes at visit 2, the samples size was only 89, making this test underpowered and subject to error.

Phenotypes at the two visits are correlated, and we would expect to see similar effect sizes across the visits for true associations. With the exception of *OLR1,* the top signals the betas are directionally consistent at the other time point, with the exception of PC2 and PC3 where directionality is flipped, as explained in chapter 2. Fourteen of the twenty-one suggestive single variant associations are also nominally ($p < 0.05$) associated with the same trait at the other visit. For gene-based associations, nine of eleven gene-based signals with a second time point are also nominally associated at the other time point. While this is not independent replication, lack of consistent association at the other time point casts doubt for common variants. For rare variants this is less concerning; the carriers missing from the second time point have a larger impact on the association test than for common variants.

GABRA6 p.Arg92His, which is associated with WHR at visit 2, is physically close to *GABRB2,* which contains single variants also associated with WHR at visit 2 (Chapter 3). Both *GABRA6* and *GABRB2* are subunits of the $GABA_A$ receptor. The relationship between these variants will be further investigated (Chapter 5).

Additional associations are also part of neuronal signaling or neuronal development. The gene *GCM2* was suggestively associated with PC3 at visit 1 in the gene-based association test of protein truncating variants and rare non-synonymous variant. GCM2 serves as a switch in neuronal development, which determines if cells become neurons or glial cells, the insulating cells along neural axons that make long-range signal transduction possible(201). Deletion of this gene in humans has been reported to cause familial isolated hypoparathyroidism, which is characterized by hypocalcemia(202). This gene is involved in development; genes involved in neuronal signaling throughout life were also associated with obesity related traits. A common variant in *TRAK1,* rs4234445, was suggestively associated with PC2 at visit 1. This replicates a prior association with subcutaneous adipose tissue to visceral adipose tissue ratio. The gene is known to be involved in trafficking of receptors, particularly $GABA_A$ receptors, to and from the cell membrane(203, 204). One gene from the multi-trait gene-based analysis, *DLG4,* is known to play a role in maintenance of the synaptic junction; interestingly this gene is overlapping in a head to head orientation with *ACADVL,* which is involved in β-oxidation of long-chain fatty acids(205). Coding variants in *ACADVL* were not associated with any adiposity traits, however based on these observations multiple mechanisms for variants in this locus impacting adiposity are possible.

Rare coding variants in another gene involved in fatty acid metabolism, *ACSL1*, were suggestively associated with PC3 at visit 1 and nearly attained suggestive association at the second visit as well. The gene product, long chain acyl-CoA synthetase 1, is important for both β-oxidation of long-chain fatty acids and the synthesis of cellular lipids(206). Based on the gene's function it was hypothesized to cause lipodystrophy,

however adipose tissue specific knock out mice have the same mass as their littermates, but higher percentage body fat(207). The knock out mice were unable to maintain their body temperature in a cold environment and had lower fatty acid oxidation rates(207). This indicates that disruption of fatty acid oxidation and triglycerides can change amount of body fat, as well as disrupt the normal physiological function of adipose tissue.

Both the multi-trait analysis and the composite measures of adiposity provided biologically compelling associations signals; however, the composite measures provided more associations including replication of an association with subcutaneous adipose tissue to visceral adipose tissue ratio. This is a measure that is typically expensive and time consuming to measure. This suggests that some approximation of this can be obtained from PC2, which utilizes routine non-invasive measures of adiposity.

While other significant and suggestive associations form the exome sequencing analyses lack readily apparent biological links to obesity, the top signals should be followed up in a replication sample set before further investigating the biological impact of the variants. The replication sample set should be carefully matched to the same ancestry group because many rare variants are unique to a single ancestry group. A set of 2000 individuals from Starr County is currently undergoing whole-exome sequencing and would make an ideal replication data set. The current analyses are restricted to individuals without a diagnosis of diabetes. Using individual with type 2 diabetes for replication is far from ideal, due to the impact of diabetes medications and behavioral interventions on obesity related traits, but would be easily accessible.

Chapter 5: Conditional Analysis of Overlapping Loci

INTRODUCTION

Once an association signal is identified, the next step is to develop hypotheses of what molecular changes are induced and identify an effector transcript. In the past, candidate variants have been obtained through sequencing genes in close proximity to strong signals. Since most individuals in this genome-wide chip data set also have whole exome sequencing data, we can look at these two types of data together without having to limit to the very strongest signals. Given that we have already tested for single variant associations exome wide, a first pass screen of looking for physical overlap of top signals from the imputed data and exome sequencing data is a good starting point. While this can be quickly ascertained by visual inspection of Manhattan plots, additional investigation to assess the size of regions and the linkage disequilibrium (LD) pattern between variants is required. Signals from the imputed data that are not seen in the exomes may either be spurious signals, or are not well captured by the targeted exome sequencing. Either way investigating the sequencing data is not informative.

Traditionally, a causative variant should be one of if not the strongest association. In this case, it is possible for other variants to have stronger associations due to the smaller sample size of the exome sequencing data. It is also possible for a single gene to contain multiple causative variants, which depending on the LD pattern between the causative variants and other variants in the region can create different association patterns.

METHODS

Overlap of signals across the genome wide chip data and exome sequencing data was assessed by visual comparison of Manhattan plots for the same trait and visit. To understand these loci, analyses in the region were repeated, conditioning on the lead variants in both the sequencing and chip data, as well as biologically interesting variants such as nonsynonymous variants when they are available. Regions were visualized both before and after conditioning on candidate variants using the stand-alone version of LocusZoom v1.3 (http://genome.sph.umich.edu/wiki/LocusZoom_Standalone), which creates Manhattan plots of small regions including information on LD, genes in the region, and previously published GWAS hits(208).

RESULTS

There was no over-lapping top signals between the common variant gene based test (VEGAS) applied to the genome wide data and the rare variant gene-based test (SKAT-O) results from the exomes sequencing data. The lack of common associations is reflective of the different hypotheses tested by the different methods. In VEGAS we expect to detect multiple independent common signals in the same gene, where as in SKAT-O look for association grouping rare, putatively functional variants.

Nine regions were identified as having single variant signals in both the whole exome sequencing data and the genome-wide chip data imputed to 1000 Genomes. This includes a region on chromosome 22 including *CYB5R3* and *POLDIP3* that is associated with BMI, percent body fat, and arm circumference all at the first visit; *WIPF3* associated with BMI and waist circumference at visit 1; and *GABRB2* and *GABRA6* associated with WHR with and without adjusting for BMI. Results are shown in the form

of LocusZoom regional plots. Variants from the exome sequencing data are shown as solid circles while results form the genome-wide chip data imputed to 1000 Genomes are shown as open circles. The coloring of the points represents LD with the lead variant. For example, looking at Figure 32A, the plot spans the length of the entire gene *COL24A1.* The lead SNP, rs1911545, is represented by a purple circle; other variants are colored based on their LD with this SNP, with red being the highest r^2 and dark blue being the lowest. The blue vertical lines represent recombination rates.

A region on chromosome 1, which contains *COL24A1* is near-suggestively associated with PC3 at visit 1 in both exome sequencing and chip data. A closer look revealed two protein altering variants, rs11161732 (*COL24A1* p.Pro546Ser) and rs56046090 (*COL24A1* p.Thr328Lys). The lead most significant variant, rs1911545 is in linkage disequilibrium with rs11161732 ($r^2 = 0.49$). Conditioning on rs11161732 (Figure 32 B) attenuated some of the signal, but signal remained. By conditioning on rs56934354 (Figure 32 C) it appears the complimentary signal was attenuated; this was confirmed by conditioning on both variants simultaneously (Figure 32 D). This indicates there may be two independent signals in this gene, which both include nonsynonymous variants.

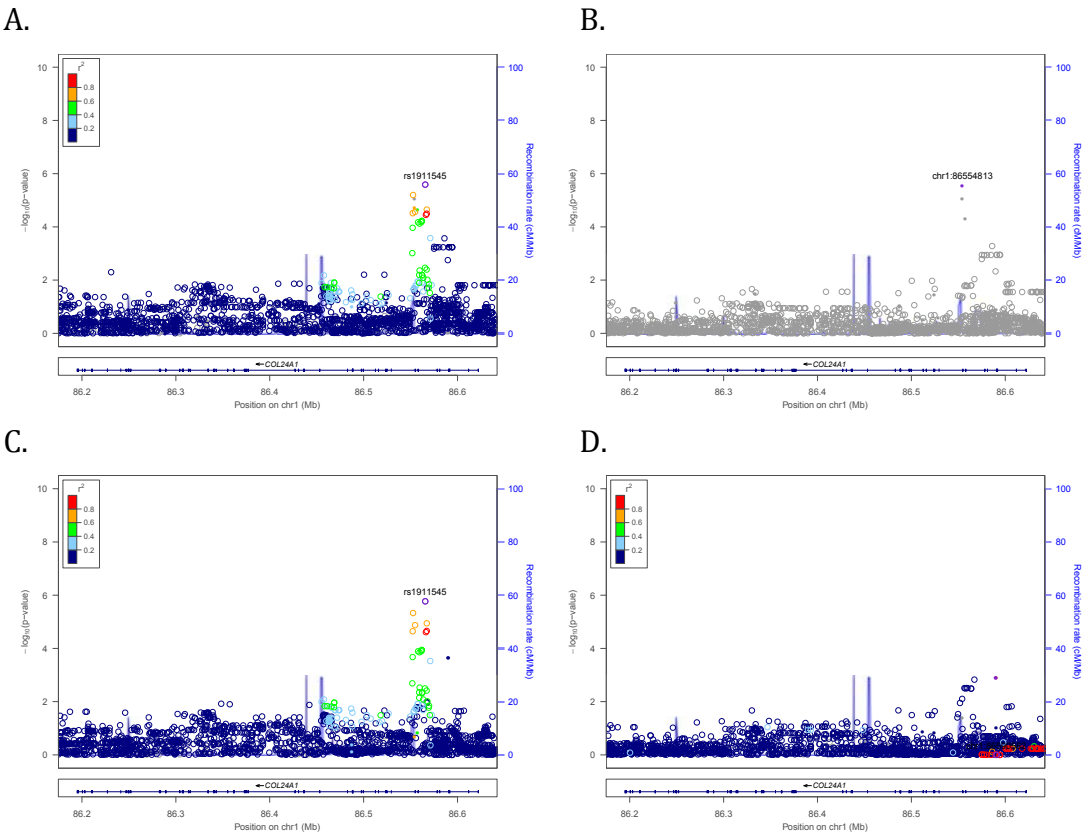


Figure 31. *COL24A1* association with PC3 at visit 1. The panels show associations of PC3 at visit 1 in the region of *COL24A1* A) without conditioning on any variants, B) conditioning on rs11161732, C) conditioning rs56046090, and D) conditioning on both rs11161732 and rs56046090. Open circles represent genome wide chip variants; solid circle represent exome sequencing variants. The point color reflects linkage disequilibrium with the top variant. All points are grey in panel B due to lack of LD information in 1000 Genomes on the top variant.

Variants on chromosome 21 near *DSCAM* are associated with PC3, which captures central verses peripheral adiposity at visit 2. Upon conditioning on the rare intronic variant, rs78085341 in *DSCAM*, (Figure 33); no association signal persists. This indicates that a single haplotype is responsible for the association with PC3 at visit 3

across both sequencing and chip data. The prior association with "Obesity related traits" shown on the plot is with arm span and height in Hispanic children. *DSCAM* is primarily expressed in the pituitary gland and brain tissues(209).

A. B.

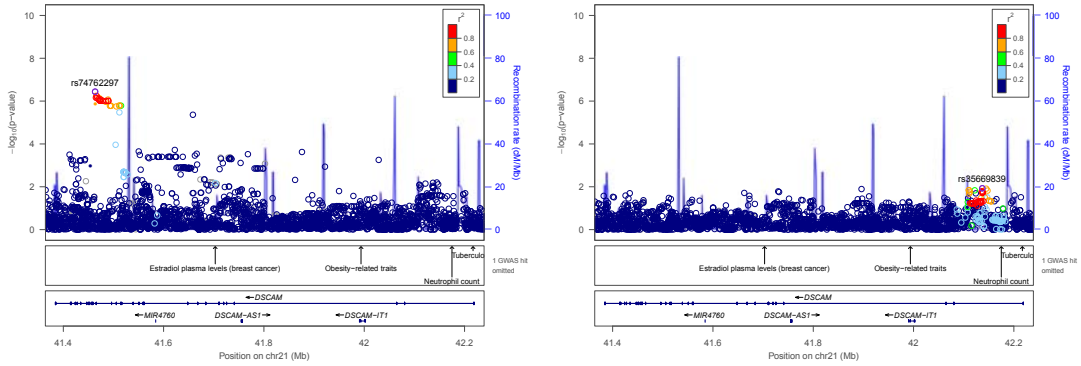


Figure 32. *DSCAM* association with PC3 at visit 2. The panels show associations of PC3 at visit 2 in the region of *DSCAM* A) without conditioning on any variants, and B) conditioning on rs78085341. Open circles represent genome wide chip variants; solid circle represent exome sequencing variants. The point color reflects linkage disequilibrium with the top variant.

A region including *WIPF3* on chromosome 7 is associated with both BMI and waist circumference at visit 1. These are highly correlated traits (r^2=0.87). For simplicity, results are shown for waist circumference, which had a slightly stronger association (Figure 34). By conditioning on the lead variant, rs2301922, there is no remaining signal, indicating one driving haplotype. The "Obesity-related traits" indicated as previously associated in the plot is actually a sleep respiratory quotient in Hispanic children(188).

A. B.

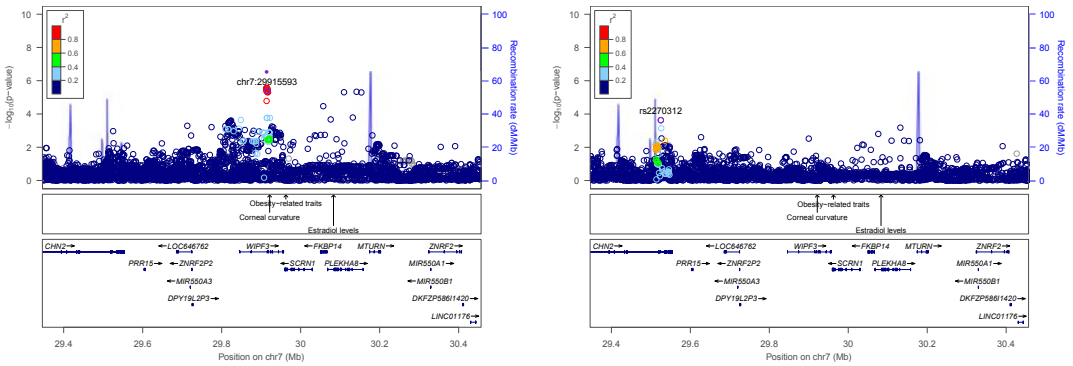


Figure 33. *WIPF3* associtation with waist circumference at visit 1. Association with waist circumference at visit 1 and variants in the region of *WIPF3* are shown A) without conditioning on any variants and B) conditioning on rs2301922. Open circles represent genome wide chip variants; solid circle represent exome sequencing variants. The point color reflects linkage disequilibrium with the top variant.

A region on chromosome 3 containing *GK5* is suggestively associated with WHR in females at visit 1 both with and without adjusting for BMI (see Table 13). Conditioning on the lead variant, rs9857725, attenuated all signal in the region (Figure 35). This lead variant is intronic and no coding variants are included in the haplotype block. Intronic variants can alter expression in many ways including altering transcription levels or disrupting or creating splice sites.

A. B.

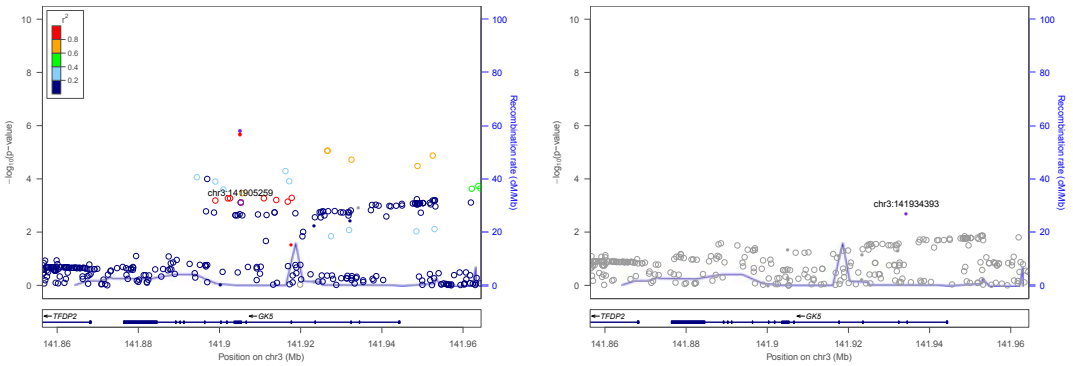


Figure 34. *GK5* association with WHR at visit 1 in females. Associations with WHR in females at visit 1 with variants near *GK5* are shown A) without conditioning on any variants and B) conditioning on rs9857725. Open circles represent genome wide chip variants; solid circle represent exome sequencing variants. The point color reflects linkage disequilibrium with the top variant. All points are grey in panel B due to lack of LD information in 1000 Genomes on the top variant.

Variants on chromosome 12 in and near *OLR1* are associated with WHR in females at visit 1. After conditioning on rs3736235, the lead variant from analysis of the exome sequencing data, the signal in attenuated (Figure 36). The common variant gene-based test for *OLR1* was also suggestively associated with WHR in females at visit1. However, conditional analysis suggests only one haplotype is responsible for this association.

A. B.

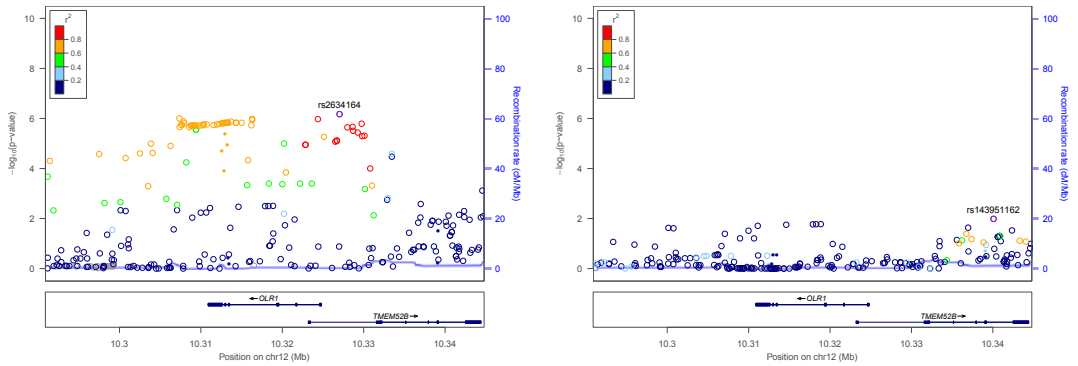


Figure 35. *OLR1* association with WHR in females at visit 1. Associations with WHR in females at visit 1 with variants near *OLR1* are shown A) without conditioning on any variants and B) conditioning on rs3736235. Open circles represent genome wide chip variants; solid circle represent exome sequencing variants. The point color reflects linkage disequilibrium with the top variant.

Despite the low number of males, two regions were associated with WHR adjusting for BMI in males in both the sequencing and chip data. At visit 1, *ECE1* on chromosome 1 is associated in both genetic data types. Conditional analyses indicate there are two regions of the gene independently associated with WHR adjusted for BMI in males. The first is represented by the lead SNP in the exome sequencing data, rs1076669; this is a nonsynonymous variant causing amino acid residue 341 to change from threonine to isoleucine. Conditioning on *ECE1* p.Thr341Ile attenuates all the signal in the 3' part of the gene, but a stronger signal about 100 kilobases away in the promoter region of the gene remains (Figure 37 B). After conditioning on the common intergenic variant, rs12137689 in the genome-wide chip data, the promoter signal is diminished (Figure 37 C). This indicates there are two separate signals, one coming form the haplotype including *ECE1* p.Thr341Ile, and another from common variation in the

promoter region of *ECE1*. The second male specific WHR association signal comes from visit 2, where we observe an association with variants in an near *TRIOBP* (Figure 38); this association is attenuated after conditioning the on rs739138, the top variant from the whole exome sequencing analyses. This variant is a nonsynonymous variant, changing residue 1300 from histidine to arginine.

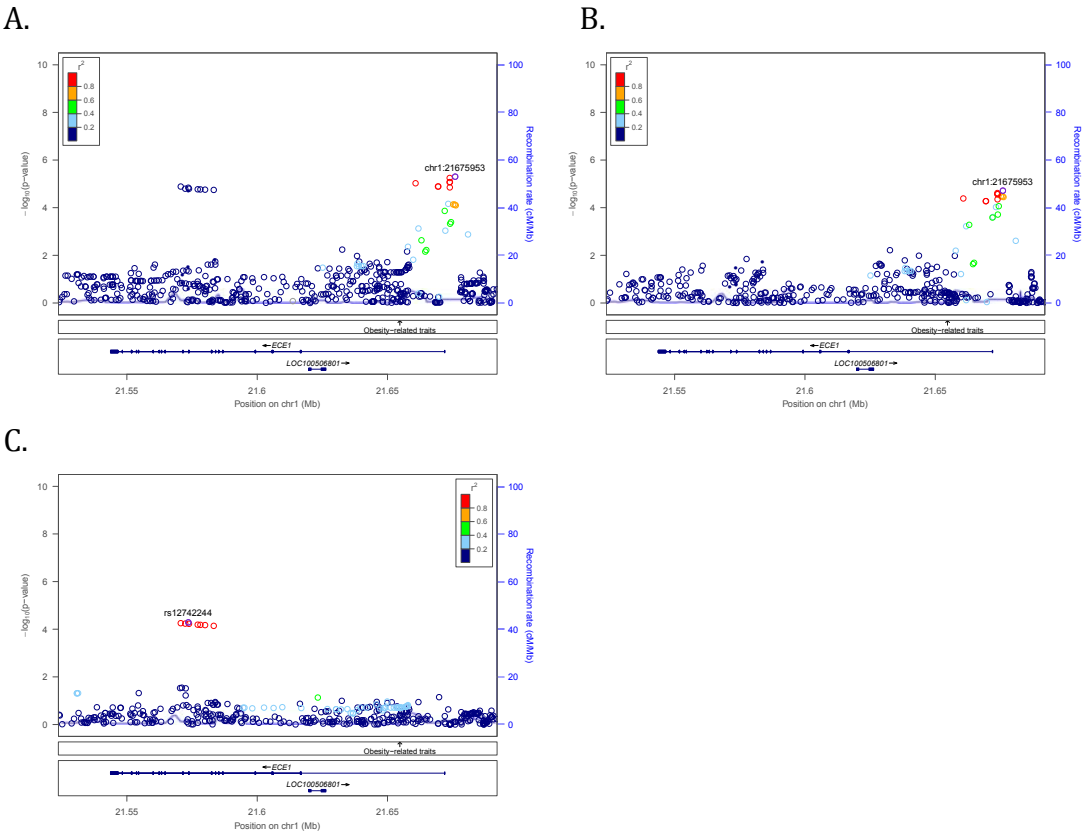


Figure 36. *ECE1* association with WHR adjusted for BMI in males at visit 1. Associations with WHR in males at visit 1 with variants near *ECE1* are shown A) without conditioning on any variants, B) conditioning on rs1076669 (*ECE1* p.Thr341Ile), and C) conditioning on rs12137689 in the imputed data. Open circles represent genome wide chip variants; solid circle represent exome sequencing variants. The point color reflects linkage disequilibrium with the top variant.

A. B.

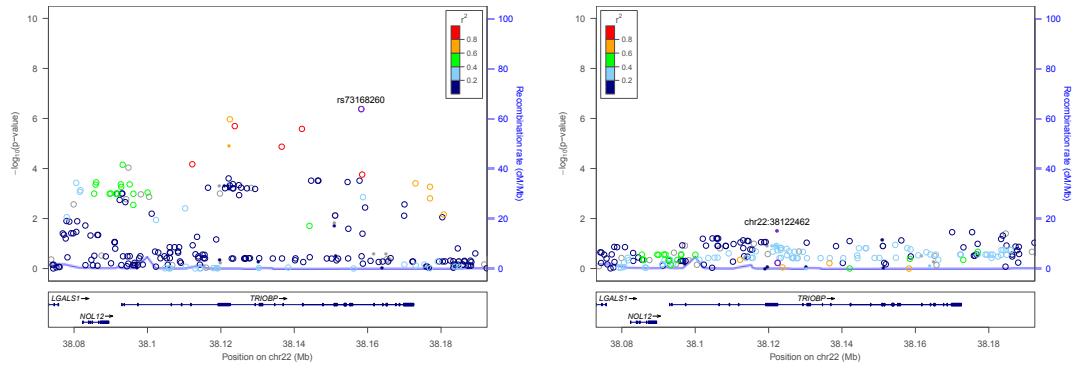


Figure 37. *TRIOBP* association with WHR adjusted for BMI in males at visit 2. Associations with WHR in males at visit 2 with variants near *TRIOBP* are shown A) without conditioning on any variants and B) conditioning on rs739138. Open circles represent genome wide chip variants; solid circle represent exome sequencing variants. The point color reflects linkage disequilibrium with the top variant.

On chromosome 22, a region containing both *CYB5R3* and *POLDIP* is associated with arm circumference, BMI, and percent body fat at visit 1. Results shown in Figure 39 are for arm circumference, where the association is strongest, but the pattern is the same in all three traits. After conditioning on rs5996200, the top signal coming from the exome sequencing data, no association signal remains. This is a synonymous variant in *CYB5R3.*

A. B.

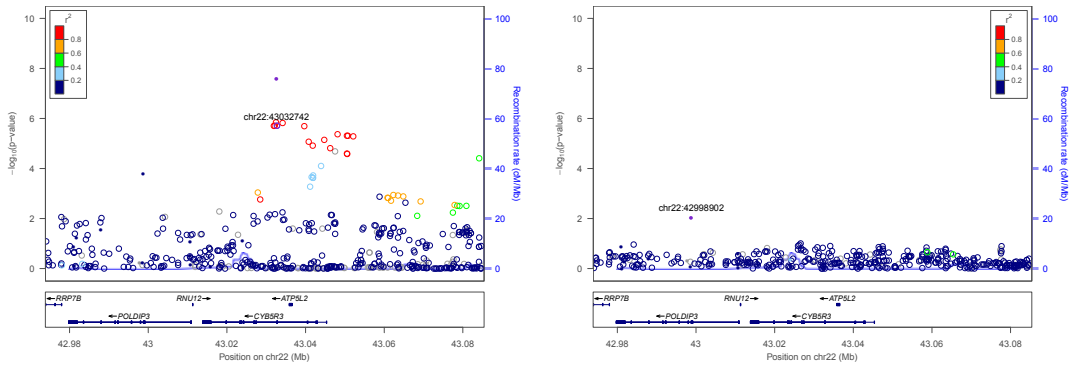
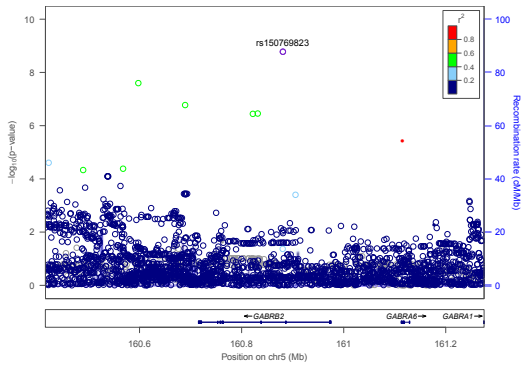


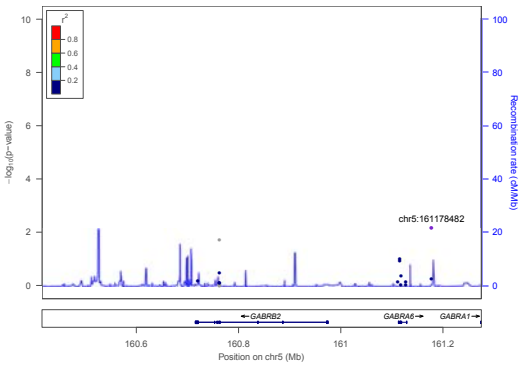
Figure 38. *CYB5R3* associations with arm circumference at visit 1. Associations with arm circumference at visit 1 with variants near *CYB5R3* are shown A) without conditioning on any variants and B) conditioning on rs5996200. Open circles represent genome wide chip variants; solid circle represent exome sequencing variants. The point color reflects linkage disequilibrium with the top variant.

Lastly, a region on chromosome 5 containing both *GABRB2* and *GABRA6* is associated with WHR at visit 2. The variant with the strongest association is rs150769823, a low-frequency intronic variant with a borderline imputation quality (info = 0.72). In the exome sequencing data, the lead variant is the low frequency nonsynonymous variant, rs3811993, which changes threonine at residue 187 of *GABRA6* to methionine (Figure 40). This variant was also imputed in the chip data, however the imputation quality excluded it from further consideration (info = 0.59) and due to poor concordance could not be used for conditional analyses. The lead variant from the imputed data, rs150769823 was not captured by the exome sequencing data. Conditioning on either of the lead variants attenuates other signal in the region for the respective datasets. The high LD between the two variants ($r^2 = 0.88$), suggests they may represent a single effector haplotype.

A.



B.



C.

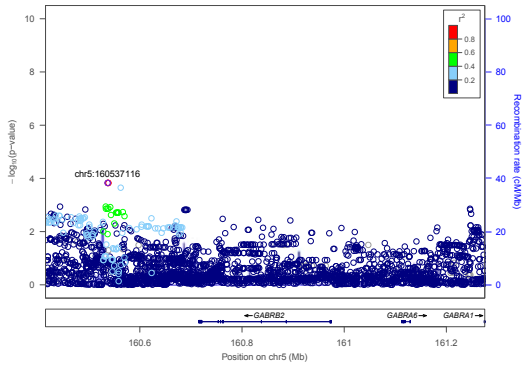


Figure 39. *GABRB2* and *GABRA1* association with WHR at visit 2. Associations with WHR visit 2 with variants near *GABRB2* and *GABRA1* are shown A) without conditioning on any variants and B) conditioning on rs3811993 in the sequencing data and C) conditioning on rs150769823 in the imputed data. Open circles represent genome wide chip variants; solid circle represent exome sequencing variants. The point color reflects linkage disequilibrium with the top variant.

DISCUSSION

Through conditional analysis of nine loci clearly associated with the same traits in exome sequencing and genome-wide chip data, nonsynonymous variants that may

account for the association signal ion the region were identified for four of the sites (Table 20).

Traits	Locus	summary
PC3 v1	*COL24A1*	Two idependent signals p.Thr328Lys and p.Pro546Ser
PC3 v2	*DSCAM*	One intronic signal
BMI v1 wasit V1	*WIPF3*	One intronic signal
WHR females v1	*GK5*	One intronic signal
WHR females v1	*OLR1*	One intronic signal
WHR adj BMI males v1	*ECE1*	Two independent singals, p.Thr341Ile and promoter region
WHR adj BMI males v2	*TRIOBP*	One signal p.His1300Arg
arm circumference v1	*CYB5R3*	One signal, synonymous variant rs5996200
WHR v2	*GABRB2/GABRA6*	High LD between GABRA6 p.The187Met, and intronic GABRB2 varaiants

Table 20. Conditional analysis conclusions.

Two nonsynonymous variants in *COL24A1* are independently associated with PC3 at visit 1. Neither of these variants quite reach exome-wide suggestive criteria, however due to the consistent effect seen in both the genome-wide chip data and exome sequencing data they are still of interest. The gene encodes a collagen that, in the mouse, is most expressed in the bone, retina, tendon, skin, and cornea (210). As discussed in chapter 2, PC3 at visit 1 measures central versus peripheral adiposity. The positive effect size for rs11161732 (β (SE) = 0.10 (0.03), MAF= 0.37) indicates that carriers have a higher in peripheral circumferences and lower central circumferences. In contrast the rare variant rs56046090 (β (SE) = -0.50 (0.13), MAF= 0.009) reflects the opposite. This may reflect differences in fat-free mass composition, such as bone density, that would require additional measures to further investigate.

Conditional analyses of regions associated WHR adjusted for BMI in males, one locus for each visit, each led to identification of putative functional nonsynonymous variants. The association with *TRIOBP* and WHR adjusted BMI for males at visit 2 may be defined by a single variant, rs739138 p.His1300Arg. This gene encodes many isoforms. While protein truncating variants in this gene are known to cause autosomal recessive nonsyndromic hearing loss (211, 212), the exon containing this variant is not part either isoform expressed in the ear, TRIOBP4 and TRIOBP5 (213). TRIOBP is involved in the regulation of cell migration, cell growth, and cytoskeletal organization(214). The association with WHR adjusted for BMI at visit 1 with *ECE1* has two independent regions, one in the promoter, the other may be defined by rs1076669, p.Thr341Ile. Mutations in this gene have been tied to Hirschprung disease and essential hypertension(215, 216). *ECE1* knock out mice have a variety of malformations including a lack of enteric neurons in the distal gut, consistent with Hirschprung disease(217). The mutation previously identified as causing Hirschprung disease is Arg754Cys, which lies on the opposite side of the protein as the variant reported here. It would be plausible that this variant has lesser effect on enteric neurons.

The final locus with a protein altering variant potentially accounting for the association with WHR at visit 2 is the *GABRA6* locus. As discussed in chapter 3 the $GABA_A$ receptor is a heteropentamer that acts as a major inhibitory synaptic signal transmission in the nervous system(183). The specific α and γ isoforms present determine the molecule's ability to bind modulators such as benzodiazepine and ethanol. The α6 isoform is classified as diazepam-insensitive(184). Mutations in some isoforms have been reported to cause epilepsy (*GABRA1, GABRG2, GABRD)*, increased

risk of alcohol dependence (*GABRA2*)(218-224). *GABRA6* has not previously been reported as associated with in any phenotypes. As discussed in chapter 3 disruption of signaling through $GABA_A$ receptor alters inhibitory signals to the hypothalamo–pituitary–adrenocortical axis, which may result in obesity. Further, recent work on mouse models shows that the neurons of lateral hypothalamus that produce GABA play a key role in compulsive sucrose seeking behaviors.(225, 226)

The other loci where conditional analyses were run did not have a nonsynonymous variant as clear candidate functional variant. However there are other means by which a variant can impact a trait. For example the intronic variants in *WIPF3* fall into an enhancer region. A variant that alters enhancer binding may subsequently impact expression levels. Investigation of the transcript levels of these genes in the GTEx portal shows that the lead variants in *COL24A1* (rs11161732 $p = 7.6 \times 10^{-9}$, tibial artery), and *TRIOBP* (rs739138 $p = 5.8 \times 10^{-7}$, whole blood), and *CYB5R3* (rs5996200 $p = 1.7 \times 10^{-6}$, esophagus mucosa) are all three *cis*-acting eQTLs, meaning they are associated with mRNA levels of the gene the variant lies in(209). Other possible mechanisms include altered promoter, repressor, or transcription factor binding. There are also encoded regulatory elements within introns, such as miRNA that are incompletely understood.

While conditional analyses clarify the relationship between variants in a region they do not indicate if the signal is real. Many of these are suggestive associations, and to establish these as obesity related variants independent replication is required. For very rare variants, ascertainment of family members will enrich the sample for carriers increasing power to detect an association. In the case that protein-altering variants have

been elucidated as likely candidates, follow up work functional work is required to establish causation. The use of existing biological information on the genes is useful for establishing the role of the gene in adiposity.

Chapter 6: Monogenic Obesity

INTRODUCTION

As reviewed in Chapter 1 and Table 2, there are three broad classes of monogenic adiposity: isolated obesity, lipodystrophy, and syndromic obesity. There are 81 genes identified as causing a monogenic disorder which features obesity or lipodystrophy. Many of these are autosomal recessive disorders that, in the heterozygous state, may have an impact on body fat amount and distribution, albeit much smaller. There is evidence that less detrimental variants in these genes, particularly *MC4R*, may have smaller effects on body fat distribution (43). With the prior information that these genes are likely to impact obesity and fat distribution, a more careful look at these genes is warranted irrespective of any other associations of lack thereof seen in the prior analyses of genome wide markers and exome sequencing. In this chapter, however, we will focus on the arrays of mutations found in these 81 genes at look at them individually and in aggregate. We will examine their impact on the full set of phenotypes as there may be effects in either the amount or distribution of body fat of both.

The exome sequencing data provides information about the coding variation, however intronic variation may alter gene expression thus having an impact on adiposity as well. Therefore we will look at the single variant associations from the imputed data as well as taking a deeper look at the whole exome sequencing data.

METHODS

Gene set definition: The monogenic adiposity genes listed in Table 2 were obtained through literature search and Online Mendelian Inheritance in Man (www.omim.org). Genes were classified as implicated in causing isolated obesity,

lipodystrophy or syndromic obesity based on published descriptions as referenced in Table 2; a single gene could fall in multiple categories (Table 21).

Array of variation: The number of variants and counts by annotation were pulled from the whole exome sequencing data using PLINK\SEQ (https://atgu.mgh.harvard.edu/plinkseq/) for all 1,490 Starr County individuals with exome sequencing data, regardless of diabetes status. The transcript with the most deleterious annotation was used for each variant. For example, rs117372135 is a missense variant in two of five *MCHR1* transcripts. For the other transcripts this variant is noncoding due to alternative splicing. We considered the most deleterious variant when looking at the amount of variation since this is the form in which the variant is most likely to have a large impact on the functionality of the gene-product.

Enrichment: To test for enrichment of association signal in monogenic adiposity genes, genomic inflation factors (λ) were calculated for variants in each of the three subsets of genes (isolated obesity, lipodystrophy, and syndromic obesity) and the full list of genes as shown in Table 21 for each obesity related traits. This was performed separately on the genome wide chip data imputed to 1000 Genomes and whole exome sequencing results. As with previous analyses, genes on chromosome X were not included. To assess the significance of inflation or deflation single variant analyses and genomic inflation factors were calculated for 1,000 simulated N~(0,1) outcomes for the same subsets individuals. This allows us to assess the frequency with which enrichment or depletion as extreme as we observed would be observed by chance alone in these particular sets of genes and individuals.

Global test of association: To evaluate the role of rare coding variants in the autosomal genes, SKAT-O analyses were conducted combining the variants across each gene set into a single gene based test. This was done for all three masks (Table 13 in Chapter 4) and each of the four gene sets.

gene	position	stop gain	missense	splice	synonymous	all	gene sets
LMNA	chr1:156052336-156109880	0 (0)	5 (4)	1 (1)	9 (4)	32 (16)	lpd
SDCCAG8	chr1:243419306-243663393	0 (0)	12 (9)	2 (2)	9 (5)	36 (24)	synd
CCDC28B	chr1:32665986-32670991	0 (0)	13 (9)	2 (1)	7 (5)	28 (21)	synd
ZMPSTE24	chr1:40723721-40759856	1 (1)	3 (3)	1 (1)	2 (1)	18 (14)	lpd
LEPR	chr1:65886334-66103176	0 (0)	19 (14)	1 (1)	13 (10)	56 (38)	ob
H6PD	chr1:9294862-9331394	0 (0)	29 (23)	2 (2)	30 (19)	66 (48)	synd
BBS5	chr2:170336005-170363165	0 (0)	9 (6)	2 (1)	5 (3)	28 (20)	synd
GPR35	chr2:241544824-241570676	2 (1)	25 (15)	0 (0)	11 (4)	47 (26)	synd
POMC	chr2:25383721-25391559	0 (0)	7 (5)	0 (0)	4 (2)	12 (8)	ob
WDPCP	chr2:63348534-63815867	0 (0)	14 (9)	5 (4)	0 (0)	35 (25)	synd
ALMS1	chr2:73612885-73837046	2 (1)	87 (57)	1 (1)	46 (31)	162 (104)	synd
PPARG	chr3:12329348-12475855	0 (0)	5 (2)	2 (2)	11 (10)	29 (21)	lpd
SLC2A2	chr3:170714136-170744768	0 (0)	8 (5)	1 (1)	6 (3)	24 (13)	synd
THRB	chr3:24158644-24536313	0 (0)	4 (2)	1 (1)	4 (2)	19 (10)	synd
LZTFL1	chr3:45864809-45957216	0 (0)	11 (8)	0 (0)	5 (4)	26 (19)	synd
ARL6	chr3:97483364-97520086	0 (0)	1 (1)	1 (1)	0 (0)	9 (6)	synd
CIDEC	chr3:9908393-9921938	1 (1)	2 (2)	0 (0)	4 (2)	11 (7)	lpd
BBS7	chr4:122745483-122791652	1 (1)	12 (11)	1 (0)	3 (2)	28 (20)	synd
BBS12	chr4:123653856-123666098	0 (0)	22 (13)	1 (1)	10 (2)	35 (17)	synd
FGFR3	chr4:1795038-1810599	0 (0)	11 (9)	7 (3)	24 (15)	87 (59)	synd
PROP1	chr5:177419235-177423243	0 (0)	6 (4)	1 (0)	4 (2)	14 (9)	synd
PCSK1	chr5:95726039-95768985	0 (0)	9 (4)	1 (1)	8 (6)	34 (26)	ob
SIM1	chr6:100836749-100911551	0 (0)	17 (10)	0 (0)	9 (7)	36 (23)	ob;synd
PSMB8	chr6:32808493-32812712	0 (0)	14 (12)	3 (1)	8 (6)	33 (25)	lpd
RAB23	chr6:57051790-57087112	1 (1)	4 (2)	1 (1)	1 (1)	17 (15)	synd
MRAP2	chr6:84743419-84800605	0 (0)	6 (6)	2 (1)	2 (0)	12 (8)	ob
CAV1	chr7:116164838-116201239	0 (0)	2 (0)	0 (0)	4 (3)	12 (7)	lpd
LEP	chr7:127881330-127897682	0 (0)	3 (2)	0 (0)	6 (5)	14 (9)	ob
CRHR2	chr7:30691558-30739719	0 (0)	8 (4)	0 (0)	3 (2)	19 (11)	ob
GHRHR	chr7:31003635-31019146	0 (0)	5 (1)	2 (2)	9 (7)	33 (17)	synd
BBS9	chr7:33169151-33645680	0 (0)	24 (16)	4 (3)	15 (13)	62 (44)	synd
VPS13B	chr8:100025493-100889814	1 (0)	64 (49)	6 (5)	42 (24)	168 (120)	synd
TMEM67	chr8:94767071-94831460	2 (2)	10 (7)	3 (2)	6 (4)	45 (32)	synd

gene	position	stop gain	missense	splice	synonymous	all	gene sets
TRIM32	chr9:119449580-119463579	0 (0)	15 (15)	0 (0)	5 (4)	21 (20)	synd
AGPAT2	chr9:139567594-139581911	0 (0)	7 (5)	1 (1)	9 (8)	36 (27)	lpd
NTRK2	chr9:87283372-87641985	0 (0)	8 (7)	4 (3)	9 (6)	42 (32)	ob
BBIP1	chr10:112658487-112679124	0 (0)	0 (0)	0 (0)	0 (0)	2 (2)	synd
BDNF	chr11:27676441-27743605	0 (0)	5 (3)	0 (0)	1 (0)	10 (6)	ob;synd
PAX6	chr11:31806340-31839509	0 (0)	4 (3)	2 (2)	9 (6)	26 (18)	synd
WT1	chr11:32409321-32457081	0 (0)	3 (3)	0 (0)	5 (2)	18 (10)	synd
BSCL2	chr11:62457733-62477091	0 (0)	9 (5)	2 (2)	11 (10)	35 (23)	lpd
MEN1	chr11:64570985-64578766	0 (0)	13 (9)	1 (1)	8 (4)	30 (17)	synd
BBS1	chr11:66278118-66301084	0 (0)	13 (9)	3 (2)	8 (5)	43 (33)	synd
TBX3	chr12:115108058-115121969	0 (0)	5 (5)	0 (0)	12 (5)	22 (13)	synd
BBS10	chr12:76738265-76742222	0 (0)	9 (7)	0 (0)	1 (1)	11 (9)	synd
CEP290	chr12:88442789-88535993	1 (1)	29 (21)	8 (7)	9 (4)	69 (45)	synd
TTC8	chr14:89290496-89344340	0 (0)	7 (6)	2 (1)	5 (5)	24 (19)	synd
MKRN3	chr15:23810453-23813166	0 (0)	10 (9)	0 (0)	7 (6)	17 (15)	synd
MAGEL2	chr15:23888695-23892993	0 (0)	12 (11)	0 (0)	5 (4)	18 (16)	synd
NDN	chr15:23930553-23932450	0 (0)	4 (4)	0 (0)	5 (4)	11 (10)	synd
SNRPN	chr15:25068793-25223729	0 (0)	1 (1)	1 (1)	5 (4)	26 (16)	synd
SNORD116-1	chr15:25296622-25296719	0 (0)	0 (0)	0 (0)	0 (0)	6 (4)	synd
IPW	chr15:25361691-25367623	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	synd
UBE3A	chr15:25582395-25684175	0 (0)	5 (5)	0 (0)	5 (2)	19 (14)	synd
GABRG3	chr15:27216428-27778373	0 (0)	4 (2)	2 (2)	11 (3)	25 (10)	synd
BBS4	chr15:72978519-73030817	0 (0)	13 (11)	3 (2)	7 (6)	35 (22)	synd
PLIN1	chr15:90207599-90222648	0 (0)	12 (7)	1 (1)	4 (1)	26 (17)	lpd
BBS2	chr16:56518258-56554008	0 (0)	4 (1)	1 (0)	8 (7)	30 (16)	synd
PMM2	chr16:8891669-8943194	1 (0)	13 (11)	1 (1)	4 (2)	35 (28)	lpd
PTRF	chr17:40554466-40575338	1 (1)	4 (4)	0 (0)	5 (4)	10 (9)	lpd
CRHR1	chr17:43861645-43913194	0 (0)	9 (7)	0 (0)	4 (1)	33 (19)	ob
MKS1	chr17:56282796-56296966	0 (0)	15 (12)	0 (0)	2 (1)	31 (22)	synd
GH1	chr17:61994552-61996212	0 (0)	5 (4)	1 (1)	9 (8)	33 (23)	synd
PRKAR1A	chr17:66409763-66547457	0 (0)	18 (15)	3 (3)	15 (12)	76 (57)	synd
MC4R	chr18:58038563-58040001	0 (0)	14 (9)	0 (0)	2 (2)	17 (12)	ob
AKT2	chr19:40736223-40791302	0 (0)	9 (9)	4 (4)	7 (6)	56 (41)	lpd
MEGF8	chr19:42829760-42882921	0 (0)	42 (33)	10 (7)	40 (25)	109 (78)	synd
INSR	chr19:7112265-7294011	0 (0)	12 (11)	6 (3)	36 (15)	81 (41)	synd
MKKS	chr20:10385427-10414887	0 (0)	13 (11)	0 (0)	6 (3)	27 (19)	synd
MC3R	chr20:54823787-54824871	1 (1)	7 (5)	0 (0)	5 (5)	18 (13)	ob
GNAS	chr20:57414794-57486250	1 (1)	19 (17)	2 (2)	26 (19)	60 (47)	synd
IFT27	chr22:37154245-37172177	0 (0)	4 (4)	0 (0)	2 (1)	12 (7)	synd
MCHR1	chr22:41075181-41078818	1 (1)	13 (9)	0 (0)	8 (6)	26 (18)	ob
GPC4	chrX:132435063-132549205	0 (0)	5 (3)	1 (1)	1 (1)	11 (5)	synd
GPC3	chrX:132669775-133119673	0 (0)	7 (5)	2 (1)	6 (4)	19 (14)	synd

gene	position	stop gain	missense	splice	synonymous	all	gene sets
PHF6	chrX:133507341-133562822	0 (0)	2 (2)	2 (1)	3 (3)	13 (9)	synd
OFD1	chrX:13752831-13787480	0 (0)	11 (11)	0 (0)	6 (4)	34 (27)	synd
FMR1	chrX:146993468-147032647	0 (0)	5 (4)	1 (0)	3 (2)	21 (12)	synd
MECP2	chrX:153287263-153363188	0 (0)	15 (13)	1 (1)	13 (9)	30 (23)	synd
HDAC8	chrX:71549365-71792953	0 (0)	2 (2)	0 (0)	5 (2)	15 (10)	synd
CHM	chrX:85116184-85302566	0 (0)	8 (5)	2 (1)	6 (4)	25 (16)	synd

Table 21. Monogenic adiposity genes. In parentheses are the numbers of variants with a MAC<6. Abbreviations: MAC: minor allele count, synd: syndromic obesity genes, lpd: lipodystrophy genes, ob: isolated obesity genes.

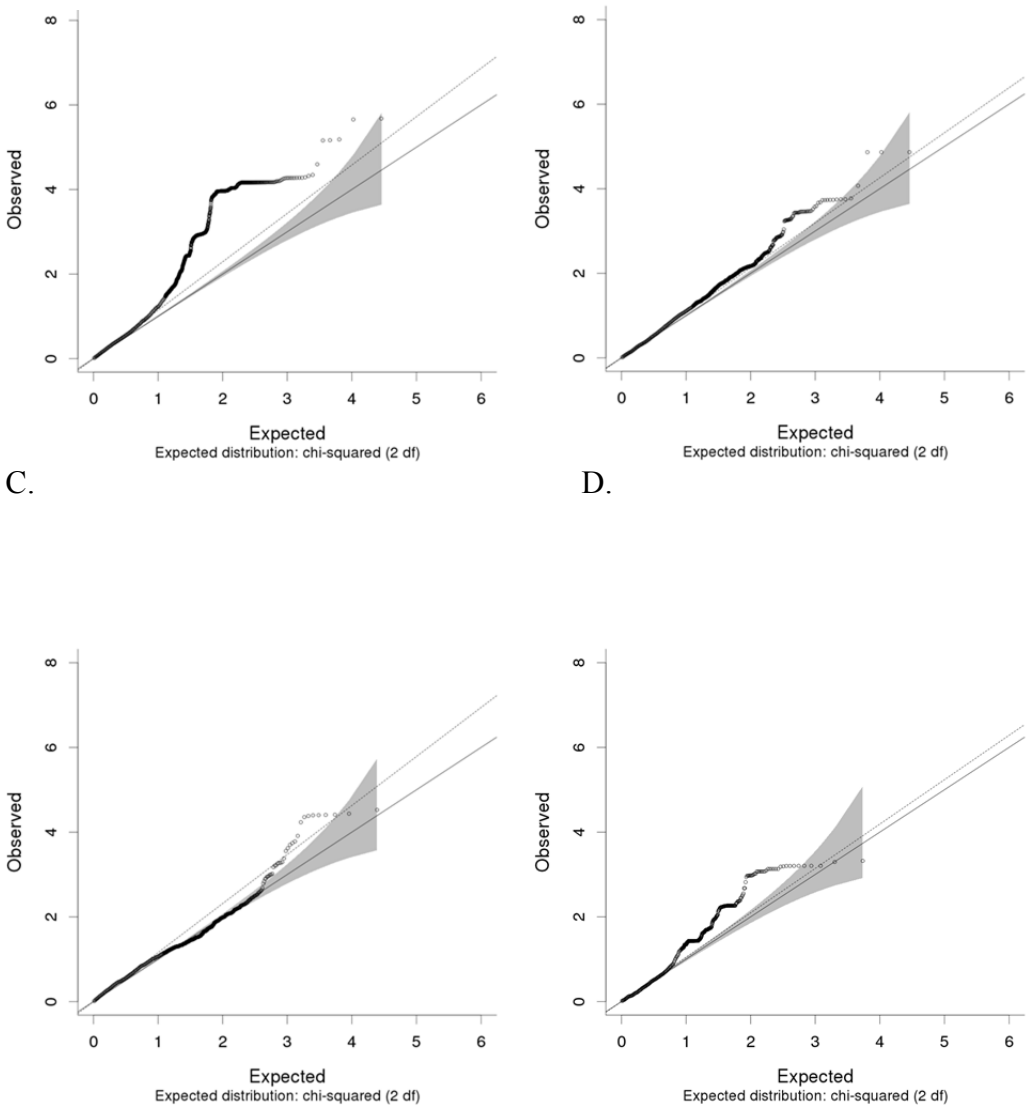
RESULTS

Across all monogenic obesity genes, very few protein-truncating variants were observed (Table 21). Phenotypic traits of carriers were assessed directly by inspection to see if carriers were obvious outliers with regard to their phenotypes. Only one man carrying a protein-truncating variant in one of these genes, *MCHR1*, stood out as having and extremely high BMI of 53.1 kg/m^2 at the age of 18. His BMI remained high throughout adulthood, with a dip down to 38.1 at age 79. Despite having an extraordinarily high BMI he was not diagnosed with diabetes until he was in his mid 70s. Mutations in *MCHR1* are reported not only to increase risk for severe early onset obesity, but also some loss of function mutations have been identified in underweight individuals(227, 228).

To assess the overall impact of these variants, the results from previous single variant analyses in both the imputed and exomes data sets were looked at for each gene set, allowing us to look outside the confidence bands instead of applying a strict significance threshold. Few variants clearly lay outside of the confidence bands on gene-set specific QQ-plots. From the genome-wide chip data imputed to 1000 Genomes, this

include variants in the syndromic obesity genes *VPS13B* and *WDPCP* for PC2 at visit 1 (Figure 41 A), *TRIM31*, *SDCCAG8, CEP290,* and *SNRPN* with percent body fat at visit 1 (Figure 41 B), *INSR* and *SNRPN* with percent body fat at visit 2 (Figure 41 C), and the isolated obesity genes *LEP* and *NTRK2* with PC3 at visit 2 (Figure 41 D).

A. B.



C. D.

Figure 40. Monogenic variants outside of confidence bands from genome-wide imputed data single variant tests.

Similar analyses in the exome sequencing data, show variants laying above the confidence bands were in the isolated obesity *MC3R* and *MCHR1* for PC2 at visit 1 (Figure 41 A) and the syndromic obesity genes *BBS2* and *VPS13B* for PC2 at visit 1

(Figure 41 B). At least one of the top variants in each of these genes from the exome sequencing data is a nonsynonymous variant.

A. B.

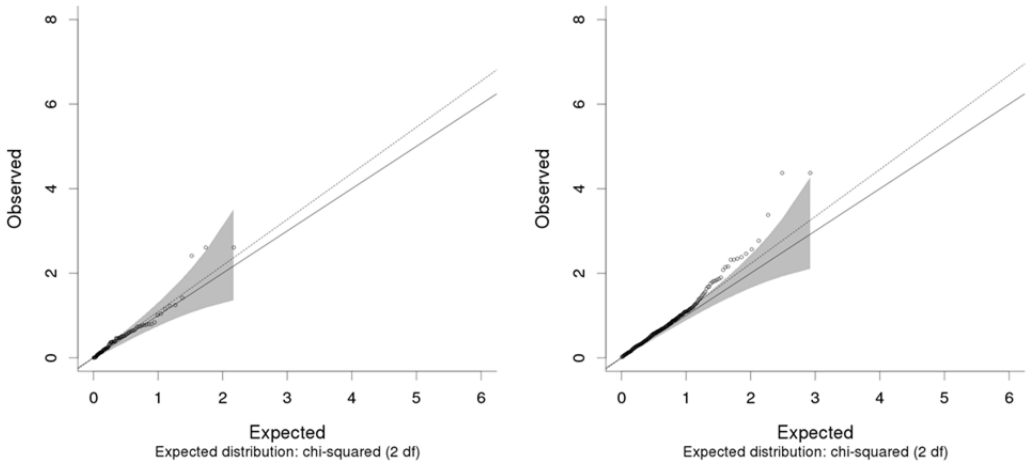
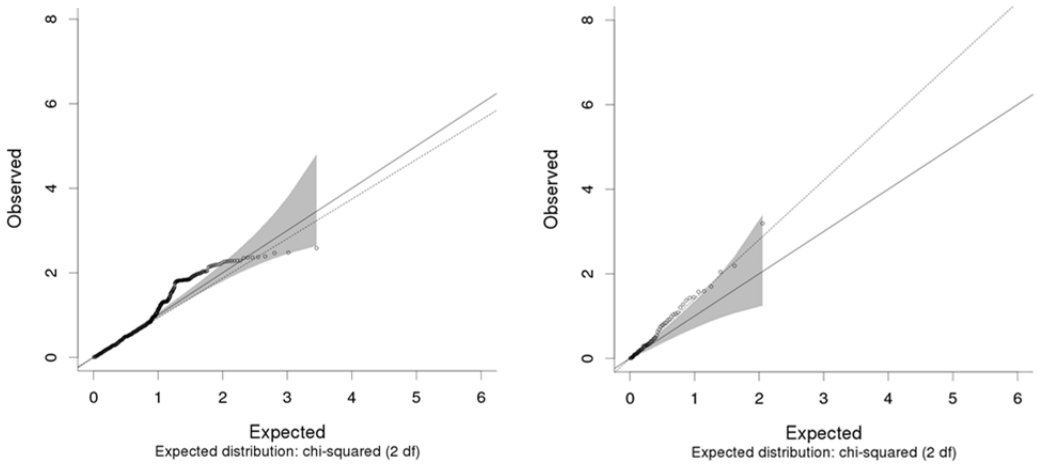


Figure 41. Monogenic variants outside of confidence bands from exome sequencing single variant tests.

In both the exome sequencing and genome-wide chip data the lipodystrophy gene *LMNA* falls above the confidence bands in the sex specific analyses, interestingly the betas for these variants are positive in females and negative in males (Figure 42). This reflects the sexually dimorphic nature of WHR where men tend to typically increase WHR with increasing weight and women gain weight on their hips, thus decreasing WHR. *LMNA* is a lipodystrophy gene affecting fat deposit location and size, making individuals have an atypical body shape for their sex.

A. C.



C. D.

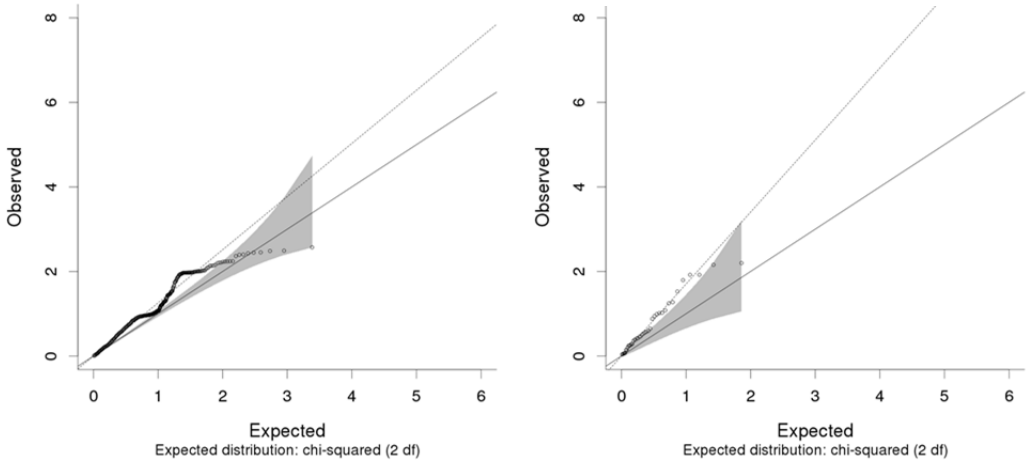


Figure 42. lipodystrophy gene variants sex-specific association with WHR.

In addition to looking for variants that fall outside the confidence bands on the QQ plots, the gene-set specific λ provides insight into systematic inflation or deflation. Across the combination of gene sets and traits, both enrichment and depletion were observed with λ ranging from 0.58 to 1.70 (Table 22), however this does not tell us if the enrichment or depletion is significant. Some of sets are small, with the lipodystrophy

gene set containing as few as 29 variants with a minor allele count of 5 or more. Further, the sample size drops to 89 for male-specific analyses at visit 2. To assess significance, analyses were repeated on 1,000 simulated N~(0,1) traits for the same individuals. Only seven gene set and trait combinations were nominally significant (*p*-value <0.05); none remained significant after correcting for multiple testing.

The nominally significant tests are in bold text in Table 22. In the genome-wide imputed data, waist circumference at visit 1 *p*-values in isolated obesity genes were deflated with a $\lambda = 0.73$. Similarly, WHR at visit 1 *p*-values for isolate obesity genes were deflated with a $\lambda = 0.77$ ($p = 0.036$). For PC2 at visit 1, *p*-values in all monogenic obesity genes are inflated, with a $\lambda = 1.12$ ($p = 0.044$). The smallest *p*-values in this gene set reside in *VPS13B,* which causes Cohen Syndrome, or *WDPCP,* which is thought to modify Bardet-Biedl syndrome(58, 67).

In the exome sequencing data, WHR in females at visit 2 *p*-values in both the isolated obesity and all monogenic obesity gene sets are deflated, with $\lambda = 0.58$ ($p = 0.014$) and $\lambda = 0.84$ ($p = 0.034$), respectively. PC3 at visit 2 *p*-values in isolated obesity genes are deflated with $\lambda = 0.61$ ($p = 0.022$). PC2 at visit 2 *p*-values were inflated with $\lambda = 1.65$ ($p = 0.020$); the variants with the smallest *p*-values are in *LMNA*.

	genome-wide imputed data								sequencing							
	all		isolated		lipodystrophy		syndromic		all		isolated		lipodystrophy		syndromic	
Trait	λ	**P-value**	λ	**P-value**	λ	**P-value**	λ	**P-value**	λ	**P-value**	λ	**P-value**	λ	**P-value**	λ	**P-value**
arm visit 1	0.99	1.000	0.83	0.142	0.80	0.186	1.06	0.464	1.02	0.696	0.83	0.380	1.15	0.352	1.03	0.636
arm visit 2	1.00	1.000	0.82	0.144	1.15	0.338	1.03	0.650	1.01	0.898	0.82	0.442	1.15	0.440	1.03	0.674
BMI visit 1	0.99	0.980	0.92	0.602	0.85	0.358	1.04	0.566	1.07	0.236	1.03	0.810	1.06	0.670	1.08	0.284
BMI visit 2	1.01	0.766	1.00	0.932	1.01	0.854	1.02	0.752	1.07	0.390	0.83	0.456	1.03	0.814	1.13	0.166
BMI at age 18	1.05	0.302	1.03	0.692	1.16	0.296	1.05	0.380	1.04	0.386	1.05	0.694	1.15	0.328	1.02	0.684
BMI at max	0.97	0.752	0.96	0.792	1.19	0.248	0.96	0.716	1.00	1.030	0.89	0.586	1.18	0.294	1.00	1.072
hip visit 1	0.98	0.854	0.90	0.444	0.88	0.492	1.01	0.780	1.05	0.410	1.09	0.590	0.96	0.848	1.06	0.390
hip visit 2	1.05	0.398	1.01	0.878	1.16	0.306	1.05	0.500	1.09	0.298	0.83	0.458	1.22	0.324	1.11	0.218
PBF visit 1	1.01	0.706	0.91	0.468	0.90	0.586	1.06	0.354	1.10	0.132	0.99	1.014	1.13	0.432	1.10	0.152
PBF visit 2	1.09	0.170	1.00	1.000	0.82	0.354	1.16	0.078	1.12	0.142	0.92	0.782	0.98	1.010	1.18	0.056
PC2 visit 1	**1.12**	**0.044**	1.06	0.596	1.08	0.510	1.14	0.060	1.11	0.090	1.09	0.588	1.05	0.722	1.11	0.120
PC2 visit 2	1.02	0.668	1.05	0.706	1.35	0.064	0.97	0.812	1.03	0.738	0.69	0.086	**1.65**	**0.020**	1.01	0.902
PC3 visit 1	0.98	0.872	0.92	0.568	0.83	0.264	1.00	0.840	0.93	0.232	0.89	0.574	0.87	0.426	0.95	0.612
PC3 visit 2	0.99	1.000	1.05	0.724	0.84	0.394	1.01	0.834	0.98	0.886	**0.61**	**0.022**	1.08	0.650	1.05	0.534
waist visit 1	0.96	0.682	**0.73**	**0.008**	0.98	0.980	1.02	0.714	1.02	0.654	0.88	0.530	1.12	0.454	1.03	0.664
waist visit 2	1.02	0.668	1.01	0.878	1.10	0.494	1.02	0.746	1.05	0.454	0.91	0.734	1.06	0.698	1.06	0.480
WHR females visit 1	0.95	0.550	0.90	0.544	0.94	0.732	0.95	0.650	0.96	0.686	1.10	0.624	1.40	0.070	0.89	0.156
WHR females visit 2	0.93	0.416	0.78	0.086	1.06	0.692	0.96	0.720	**0.84**	**0.034**	**0.58**	**0.014**	0.68	0.094	0.95	0.656
WHR adj BMI females visit 1	1.00	1.000	1.01	0.848	0.91	0.620	1.00	0.852	0.99	0.940	1.33	0.160	1.32	0.096	0.89	0.184
WHR adj BMI females visit 2	1.00	0.908	0.82	0.168	1.06	0.696	1.04	0.660	0.90	0.232	0.73	0.172	0.78	0.270	0.98	0.928
WHR males visit 1	1.03	0.690	1.25	0.172	1.26	0.208	0.96	0.690	1.01	0.900	1.39	0.254	1.70	0.110	0.86	0.290
WHR males visit 2	1.10	0.204	1.28	0.112	0.89	0.572	1.08	0.338	1.01	0.876	0.99	1.000	1.09	0.664	0.99	0.998
WHR adj BMI males visit 1	0.98	0.890	1.18	0.278	0.87	0.494	0.95	0.720	0.94	0.540	0.82	0.456	1.18	0.482	0.95	0.746
WHR adj BMI males visit 2	1.00	0.912	1.10	0.568	0.94	0.856	0.99	0.954	0.92	0.468	0.80	0.510	1.01	0.948	0.93	0.630
WHR visit 1	0.98	0.956	**0.77**	**0.036**	0.94	0.826	1.04	0.666	0.96	0.706	0.81	0.324	1.04	0.818	0.98	0.916
WHR visit 2	1.05	0.396	0.90	0.452	0.98	0.936	1.11	0.224	0.98	0.894	0.70	0.090	0.92	0.724	1.06	0.460
WHR adj BMI visit 1	1.06	0.236	0.92	0.610	0.79	0.216	1.12	0.212	1.02	0.702	1.07	0.702	1.11	0.618	1.01	0.810
WHR adj BMI visit 2	1.09	0.214	0.90	0.440	0.93	0.738	1.15	0.090	0.96	0.626	0.72	0.124	0.88	0.574	1.02	0.772

Table 22. Systematic enrichment of monogenic adiposity genes.

To test for association of rare functional variants in these genes, SKAT-O association tests were run on each trait combining all genes each gene set. Across all obesity traits, three masks, and four gene sets, no test was significant after correcting for multiple testing within trait. Conservatively treating each test as independent, Bonferroni correction provides a significance threshold of 0.004 within each trait. The strongest associations are shown in Table 23. Protein truncating variants and rare protein altering variants predicted damaging by metaLR in syndromic obesity genes are suggestively associated with PC3 at visit 2 ($p = 0.00535$), which captures central versus peripheral adiposity. Protein truncating and rare protein altering variants are suggestively associated with waist circumference at visit 2 ($p = 0.00615$).

trait	list	mask	p-value	cMAF
PC3 visit 2	syndromic	LR	0.00535	0.28649
waist visit 2	lipodystrophy	NS	0.00615	0.21538

Table 23. Global tests of rare variant association. Abbreviations: cMAF: cumulative minor allele frequency

DISCUSSION

Monogenic adiposity genes, by definition are known to have a large effect change when altered. Here we have presented some suggestive evidence that variants in *MC3R, MCHR1, BB2,* and *VPS13B* may impact central adiposity, as reflected in the composite measure PC2 at visit 1. There is also systematic inflation of association with PC2 at visit 1 for all monogenic adiposity genes across all variation from the genome-wide imputed and PC2 at visit 2 for lipodystrophy genes from the exome sequencing

data. This supports the notion that a subset of variants in monogenic obesity genes also plays a role in common obesity.

Systematic deflation of variants in these gene sets may be reflective of these analyses being limited to type 2 diabetes controls, and individuals meeting diagnostic criteria for diabetes for the first time at the current visit. Monogenic obesity patients are at high risk for developing diabetes, and many forms of lipodystrophy feature disrupted insulin signaling, resulting in exclusion bias that would prevent us from detecting an association with these genes.

The amount of variation seen in these genes shown in Table 21, includes individuals with type 2 diabetes along with the controls and newly diagnosed diabetics used for analyses. By repeating analyses in type 2 diabetes patients, we may be able to detect an effect in these monogenic genes; however, this comes with the added uncertainty of the impact of diabetes treatments on each individual's adiposity.

Monogenic adiposity genes provide attractive therapeutic targets for both the rare cases where the gene is disrupted as well as for common obesity. Leptin injections are very effective for the handful of individuals with congenital leptin deficiency(38). Recently, the leptin injects have regained popularity, as their use in individuals with leptin deficiency due to lipodystrophy has recently been approved by the FDA(229, 230). Unfortunately it is not effective for common obesity, in which a majority of people is leptin resistant. More recently, MCHR1 antagonists have been introduced as potential obesity therapeutic agents(231-234). However none of these have reached the stage of clinical trials yet.

Monogenic adiposity genes are easier to identify than common adiposity genes. The biological pathways disrupted by monogenic obesity genes are sometimes disrupted in common obesity to a lesser extent, nonetheless making the pathways plausible therapeutic targets for common obesity.

Chapter 7: Conclusions

In this dissertation we have investigated various measures of adiposity that may better capture genetic aspects of obesity and the role of common and rare genetic variation related to these measures utilizing the resources developed Starr County, Texas Mexican Americans. In this chapter we synthesize the results and highlight findings regarding potentially new variation that may be involved in adiposity both through lipid metabolism and neurological development and signaling pathways, which all play a role in hunger and satiety.

The two multivariate methods, principal components analyses and meta-analysis of multiple traits, have different genetic results. The first, principal components analysis, used the raw obesity measures as input. These components were then adjusted for age and sex. Alternatively, the age and sex adjusted residuals could have been used. However, this would not capture a person's objectively measured shape. Principal components analysis showed that BMI, the most widely used measure of obesity, captures 82% of the variation in the array of anthropometric traits used here. This indicates most variation in adiposity has been well captured in studies of adiposity, despite the small portion of heritability explained. Two additional composite measures of adiposity, PC2 and PC3 were shown to capture different aspects of adiposity, although they capture a much smaller proportion of the variation in obesity measurements. PC2 appears to capture the amount of adipose tissue and if it is located above or below the waist. It is through this measure that we replicate a prior suggestive association with subcutaneous adipose tissue to visceral adipose ratio. The later requires medical imaging to obtain making it cost prohibitive on a large scale. The use of a

principal components strategy based on more readily obtainable measures should facilitate implementing such studies in larger samples. PC3 captured peripheral versus central adiposity, which is quite similar to the ratio of arm circumference to waist circumference. PC3 identified multiple functionally plausible associations including rare variants in *ACSL1*, a gene involved triglyceride synthesis and beta oxidation; *DSCAM* which is expressed in the brain and pituitary gland, and *GCM2*, a key switch in neuronal development. The identification of these candidates suggests that by adding arm to waist ratio to the set of routinely to anthropometric measures has the potential for identifying more genes effecting fat distribution. This highlights our argument that more refined biologically relevant measures should enhance our ability to implicate genes.

The second multivariate approach was to meta-analyze the results across the traits at two time points while adjusting for the correlation between the test statistics. Essentially, this combines the evidence of a variant's or gene's involvement for each trait, but this cannot be done without considering that the traits, and therefore results, are correlated with each other. This was done using the method CPASSOC to adjust for the correlation between test-statistics in the meta-analysis(176). This approach also identified biologically plausible associations, including a variant near the gene encoding a key skeletal muscle protein, myosin heavy chain 13, and *DLG4,* which is oriented head to head with *ACADVL*. *DLG4* and *ACADVL* are both functionally interesting, as *DLG4* is involved in synaptic junction maintenance and *ACADVL* is involved in beta-oxidation of fatty acids. While this approach did have some success, it is intended to detect plieotropy. The traits studied here are all measures of the same phenotype, distribution of body fat. However each measure excels in capturing a different dimension of adiposity.

Despite having different strengths, the measures are correlated with each other. The correction for this correlation decreases the weight of each trait, hindering the power of this approach. It does have the potential to be very powerful, particularly when considering true pleiotropic effects with different metabolic outcomes.

With the same sets of obesity traits measured at two different time points there are multiple approaches for dealing with them. Here I analyzed the measured as two different outcomes. Another approach would be to look at how individuals changed over time. For example one could analyze the average annual change in each measure; however, this would limit samples size to the smaller subset that did not have diabetes at the second visit. This approach will provide a different set of association results, and may identify additional obesity loci. Furthermore initially one may think highly correlated results would be expected. This is not what we observe, with good reason. Because we have excluded individuals previously diagnosed with diabetes from these analyses, the subset of individuals analyzed from the second visit is a non-random subset. This introduces an exclusion bias by removing the subset that developed diabetes over a short time span, which had a higher average BMI at the first visit. To better understand the impact of exclusion bias as well as the adiposity profile of those most at risk for developing diabetes exploratory analyses such as looking at the correlation between genetic associations at visit 1 and visit 2 using only the subset of people that did not develop diabetes. A useful multivariate approach to understand adiposity profiles in those most likely to develop diabetes would be to make a composite measure based on linear discriminant analysis aimed at separating those that develop diabetes between visits from those that do not. This may help identify adiposity profiles at increased risk

for developing diabetes allowing for additional intervention aimed at preventing the development of diabetes. The resulting linear combinations of adiposity traits could also be analyzed for genetic associations.

Through this work a number of genes that may be involved in body fat distribution have been identified; however, with most only being suggestively associated additional work is needed to confirm these associations. First, variants that were imputed at a lower quality must undergo validation genotyping. Some variants, such those In *GABRA6* and *GABRB2* are biologically compelling; particularly in light of recent rodent models showing that GABA producing neurons are involved in compulsive sucrose seeking(225, 226).

To bolster statistical support for the involvement of these novel genes in obesity, replication in an independent data set is ideal. There are several options here, depending on the phenotype the association was observed with. BMI and WHR are readily available. With the 43 published genome wide association studies, some of these variants have been interrogated before. While these associations have not been seen before, looking at other studies' sub-significant results is one source of replication. Many of the variants of interest, particularly rare coding variants, will not have been tested because smaller arrays and imputation reference panels were used. Studies using larger data sets are still underway; in fact these data have been contributed to multiple consortia looking at these traits. BMI and WHR for the genome-wide chip data imputed to 1000 Genomes are part of the HISLA Consortium, which is meta-analyzing BMI, and WHR across many Hispanic studies. The exome sequencing data set is a portion of the

T2D-GENES consortium, which plans to analyze both BMI and WHR across 13,000 samples.

In addition to these consortia, additional individuals form Starr County are currently being genotyped: 2000 samples are undergoing exome sequencing, 900 samples are being genotyped on a chip custom designed by the PAGE consortium which is enriched for low frequency coding variation previously seen in minority samples. These would be ethnically matched to the current data, and would have the same set of anthropometric measures available.

The clearest limitation of this work is the lack of power. With a limited sample size ranging from 825 all the way down to 89 in male specific analyses at the second visit, these analyses are underpowered, particularly for rare and low frequency variants. With the largest sample size here we only have 80% power to detect an effect size of at least 0.7 standard deviations in a variant with a 1% minor allele frequency. This is a much larger effect size than has been detected thus far. There are multiple options for overcoming the lack of power when it comes to low frequency variants.

The most obvious option is increasing sample size. One consortium, GIANT is aiming to have more than 1 million individuals in their analyses of data sets imputed out to 1000 Genomes. In the GIANTexomes consortium, almost 500,000 individuals with exome chip data are being meta-analyzed, with additional studies potentially contributing to replication analyses. The exome chip captures primarily rare and low frequency coding variants and has a much-reduced per-sample cost in comparison to sequencing. However, increasing power by increasing sample size has its limitations. BMI and WHR are “extra traits” collected by most groups studying the genetic outcomes

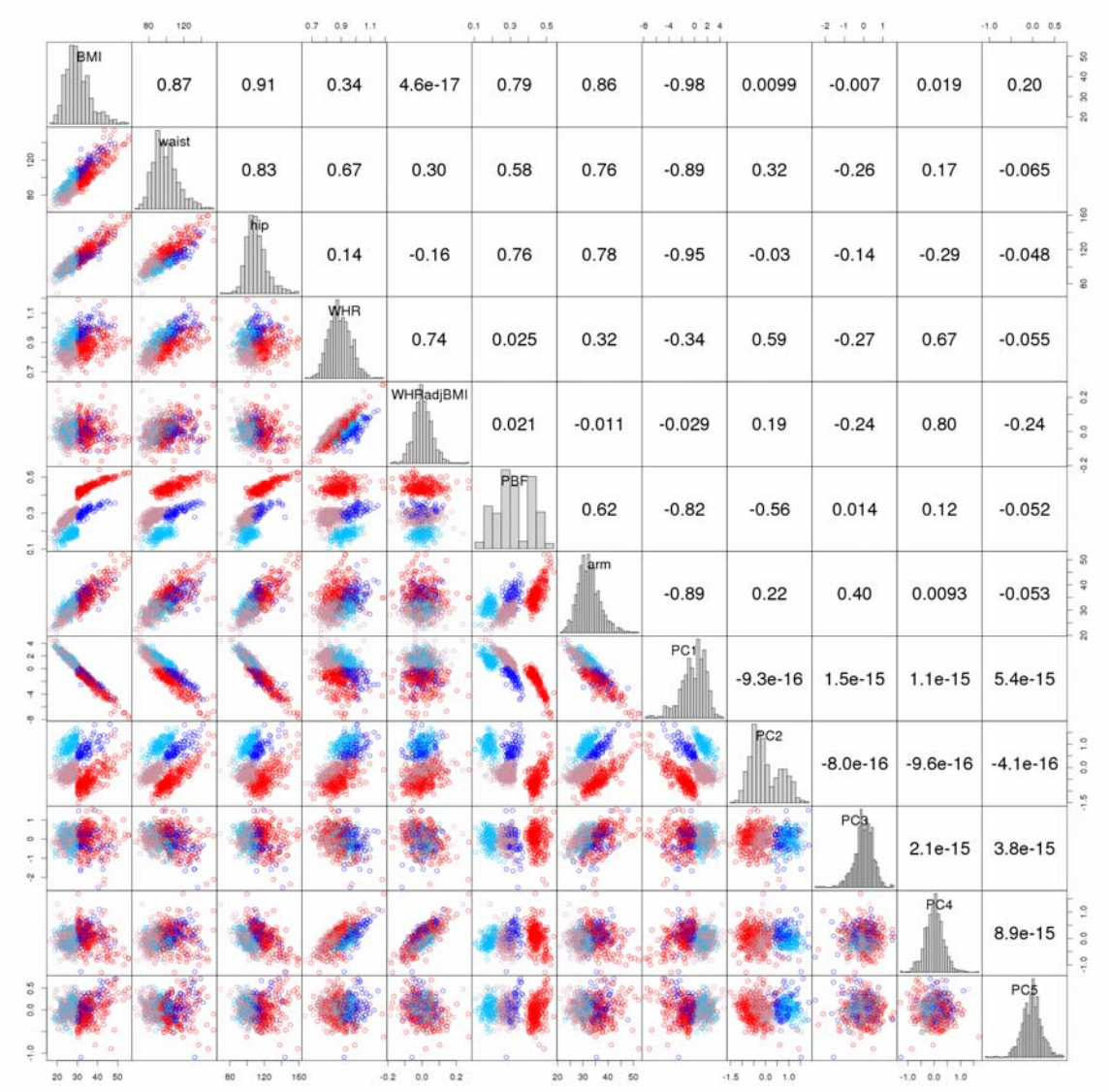
of other phenotypes and diseases; however, the expanded set of adiposity traits used here are often not measured. With many contributing studies there are sources of bias and error that can be difficult to detect and account for. Some of the outcomes data were ascertained for may alter body composition, as type 2 diabetes and subsequent treatment does. Additionally ascertainment schemes can introduce bias for add-on traits.

Another means to increase power is to bias sample collection to include carriers of rare variants of interest. For years this approach has been implemented by genotyping individuals with the most extreme outcomes. With sequencing data in hand, we can also enrich for a particular variant of interest by recruiting family members into the study. In the case of family members different analytical approaches, such as linkage, that leverage relatedness rather than adjust for relatedness are necessary.

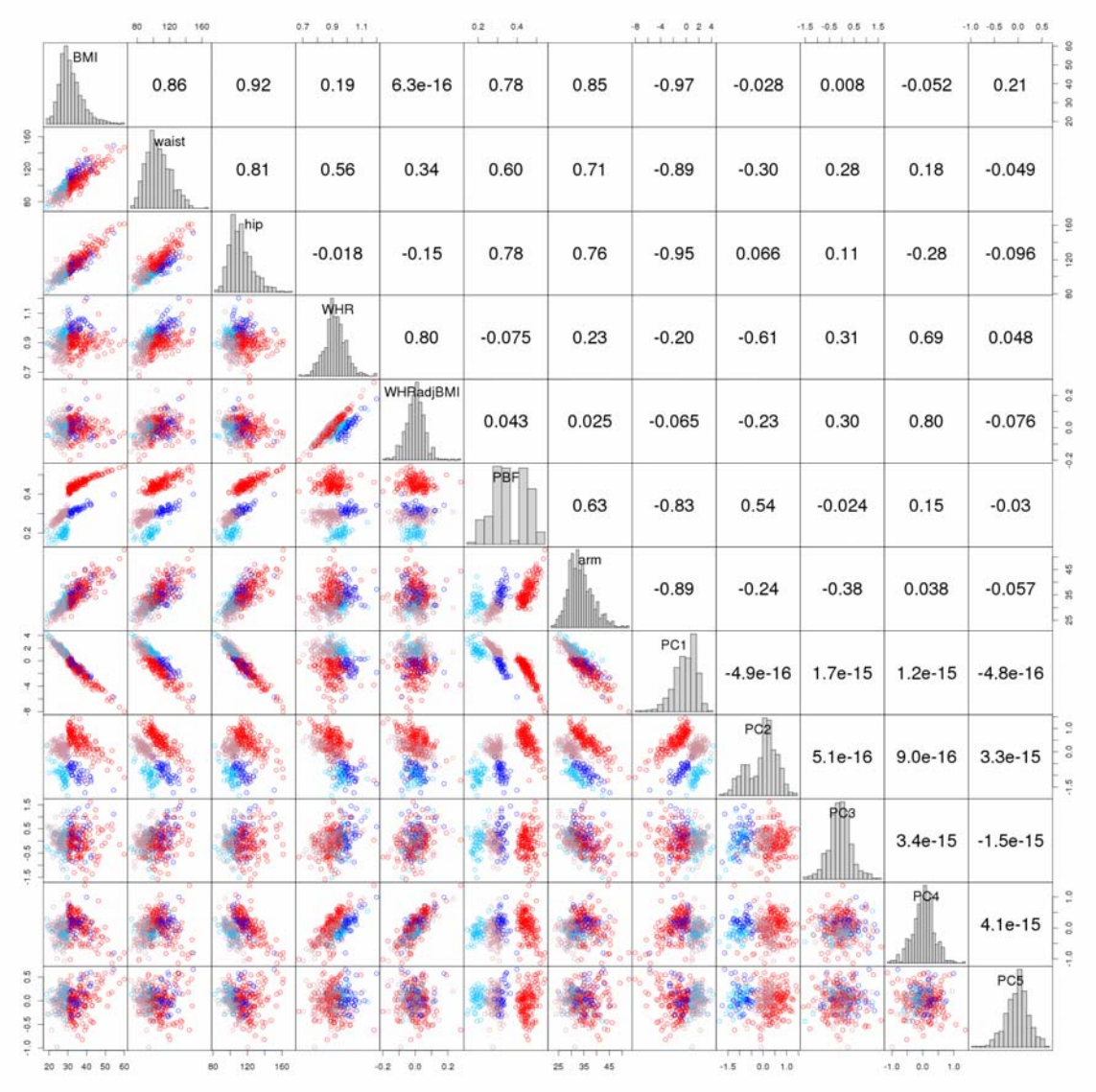
While we know that a majority of the estimated heritability is unaccounted for there are many association signals that are statistically convincing. However, there has been a lack of functional follow up work to understand the biological role of these variants and genes. Functional follow up is an expensive and time-consuming prospect. It is our responsibility as quantitative geneticists to present molecular biologists with a compelling case that these genes and variants are worth their investment. To do this, high throughput datasets, such as the GTEx project which has gene expression data for many tissues as well as genotypes, can be used to show a gene is expressed in a relevant tissue and if a variant has an impact on gene expression(209). For coding variants, the functional prediction can be taken a step beyond in the *in silico* prediction tools used to annotate variants. Existing 3D models of the protein can be utilized to view the location and predict the impact of an amino acid change on protein function. Additional data

types, such as methylation data, but can also be used to make the case that a gene and variant are involved in obesity. Once we have made this case for a gene, molecular biologists can then study the impact of a particular variant and assess the gene's biological role. The long-term goal of understanding the underlying biology is to identify therapeutic targets. Diet and exercise are only effective for long-term weight control in a subset of people. Because the obesity epidemic shows no signs of slowing down, it is more important now than ever to explore therapeutic targets to lessen the obesity crisis.

Appendix



Appendix Figure 1 Scatter plot matrix of anthropometric measures and resulting PCs for visit 1. The bottom half shows the scatter plot for each pair of traits, with PC2 in the bottom row. The upper half shows the Pearson correlation coefficient for the two measures. Histograms of the traits are displayed in the middle. Points are color coded by sex and obesity status. Obese males are dark blue; non-obese males are light blue; obese females are red, non-obese females are pink



Appendix Figure 2 Scatter plot matrix of PCs and anthropometric measures for visit 2. The bottom half shows the scatter plot for each pair of traits, with PC2 in the bottom row. The upper half shows the Pearson correlation coefficient for the two measures. Histograms of the traits are displayed in the middle. Points are color coded by sex and obesity status. Obese males are dark blue; non-obese males are light blue; obese females are red, non-obese females are pink

References

1. Centers for Disease Control and Prevention (CDC). 1995-2012. Behavioral Risk Factor Surveillance System Survey Data. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Atlanta, Georgia.
2. Ostbye, T., M. Stroo, E. L. Eisenstein, B. Peterson, and J. Dement. 2014. Is overweight and class I obesity associated with increased health claims costs? Obesity 22:1179-1186.
3. Finkelstein, E. A., J. G. Trogdon, J. W. Cohen, and W. Dietz. 2009. Annual medical spending attributable to obesity: payer-and service-specific estimates. Health affairs 28:w822-831.
4. Howard, J. T., and L. B. Potter. 2014. An assessment of the relationships between overweight, obesity, related chronic health conditions and worker absenteeism. Obesity research & clinical practice 8:e1-15.
5. Williams, G. 2013. The Heavy Price of Losing Weight. In U.S. News and World Report.
6. Look Ahead Research Group. 2014. Eight-year weight losses with an intensive lifestyle intervention: the look AHEAD study. Obesity 22:5-13.
7. Knowler, W. C., E. Barrett-Connor, S. E. Fowler, R. F. Hamman, J. M. Lachin, E. A. Walker, D. M. Nathan, and G. Diabetes Prevention Program Research. 2002. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. The New England journal of medicine 346:393-403.

8. Brownell, K. D., and R. W. Jeffery. 1987. Improving long-term weight loss: Pushing the limits of treatment. Behavior Therapy 18:353-374.

9. Kraschnewski, J. L., J. Boan, J. Esposito, N. E. Sherwood, E. B. Lehman, D. K. Kephart, and C. N. Sciamanna. 2010. Long-term weight loss maintenance in the United States. International journal of obesity 34:1644-1654.

10. Weiss, E. C., D. A. Galuska, L. Kettel Khan, C. Gillespie, and M. K. Serdula. 2007. Weight regain in U.S. adults who experienced substantial weight loss, 1999-2002. American journal of preventive medicine 33:34-40.

11. Klem, M. L., R. R. Wing, M. T. McGuire, H. M. Seagle, and J. O. Hill. 1997. A descriptive study of individuals successful at long-term maintenance of substantial weight loss. The American journal of clinical nutrition 66:239-246.

12. Thomas, J. G., D. S. Bond, S. Phelan, J. O. Hill, and R. R. Wing. 2014. Weight-loss maintenance for 10 years in the National Weight Control Registry. American journal of preventive medicine 46:17-23.

13. Brethauer, S. A., A. Aminian, H. Romero-Talamas, E. Batayyah, J. Mackey, L. Kennedy, S. R. Kashyap, J. P. Kirwan, T. Rogula, M. Kroh, B. Chand, and P. R. Schauer. 2013. Can diabetes be surgically cured? Long-term metabolic effects of bariatric surgery in obese patients with type 2 diabetes mellitus. Annals of surgery 258:628-636; discussion 636-627.

14. Buchwald, H., Y. Avidor, E. Braunwald, M. D. Jensen, W. Pories, K. Fahrbach, and K. Schoelles. 2004. Bariatric surgery: a systematic review and meta-

analysis. JAMA : the journal of the American Medical Association 292:1724-1737.

15. Livingston, E. H. 2010. The incidence of bariatric surgery has plateaued in the U.S. American journal of surgery 200:378-385.

16. Angrisani, L., P. P. Cutolo, G. Formisano, G. Nosso, and G. Vitolo. 2013. Laparoscopic adjustable gastric banding versus Roux-en-Y gastric bypass: 10-year results of a prospective, randomized trial. Surgery for obesity and related diseases : official journal of the American Society for Bariatric Surgery 9:405-413.

17. Surgery, A. S. f. M. a. B. 2014. Estimate of Bariatric Surgery Numbers || American Society for Metabolic and Bariatric Surgery.

18. Elliott, V. S. 2012. Bariatric surgery maintains, doesn't gain. In American Merical News. American Merical Societry, amednews.com.

19. Bellisari, A. 2008. Evolutionary origins of obesity. Obesity reviews : an official journal of the International Association for the Study of Obesity 9:165-180.

20. Neel, J. V. 1962. Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? American journal of human genetics 14:353-362.

21. Waalen, J. 2014. The genetics of human obesity. Translational research : the journal of laboratory and clinical medicine 164:293-301.

22. Epstein, L. H., A. M. Valoski, M. A. Kalarchian, and J. McCurley. 1995. Do children lose and maintain weight easier than adults: a comparison of child

and parent weight changes from six months to ten years. Obesity research 3:411-417.

23. Mokdad, A. H., E. S. Ford, B. A. Bowman, W. H. Dietz, F. Vinicor, V. S. Bales, and J. S. Marks. 2003. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. JAMA : the journal of the American Medical Association 289:76-79.

24. Shen, W., M. Punyanitya, A. M. Silva, J. Chen, D. Gallagher, L. B. Sardinha, D. B. Allison, and S. B. Heymsfield. 2009. Sexual dimorphism of adipose tissue distribution across the lifespan: a cross-sectional whole-body magnetic resonance imaging study. Nutrition & metabolism 6:17.

25. John, U., M. Hanke, H. J. Rumpf, and J. R. Thyrian. 2005. Smoking status, cigarettes per day, and their relationship to overweight and obesity among former and current smokers in a national adult general population sample. International journal of obesity 29:1289-1294.

26. Wang, Y., and M. A. Beydoun. 2007. The obesity epidemic in the United States--gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. Epidemiologic reviews 29:6-28.

27. Maes, H. H., M. C. Neale, and L. J. Eaves. 1997. Genetic and environmental factors in relative body weight and human adiposity. Behavior genetics 27:325-351.

28. Dina, C. 2008. New insights into the genetics of body weight. Current opinion in clinical nutrition and metabolic care 11:378-384.

29. Stunkard, A. J., J. R. Harris, N. L. Pedersen, and G. E. McClearn. 1990. The body-mass index of twins who have been reared apart. The New England journal of medicine 322:1483-1487.

30. Bray, M. S., E. Boerwinkle, and C. L. Hanis. 1999. Linkage analysis of candidate obesity genes among the Mexican-American population of Starr County, Texas. Genetic epidemiology 16:397-411.

31. Santini, F., M. Maffei, C. Pelosini, G. Salvetti, G. Scartabelli, and A. Pinchera. 2009. Melanocortin-4 receptor mutations in obesity. Advances in clinical chemistry 48:95-109.

32. Beckers, S., D. Zegers, L. F. Van Gaal, and W. Van Hul. 2009. The role of the leptin-melanocortin signalling pathway in the control of food intake. Critical reviews in eukaryotic gene expression 19:267-287.

33. Korner, J., E. Savontaus, S. C. Chua, Jr., R. L. Leibel, and S. L. Wardlaw. 2001. Leptin regulation of Agrp and Npy mRNA in the rat hypothalamus. Journal of neuroendocrinology 13:959-966.

34. Krysiak, R., E. Obuchowicz, and Z. S. Herman. 1999. Interactions between the neuropeptide Y system and the hypothalamic-pituitary-adrenal axis. European journal of endocrinology / European Federation of Endocrine Societies 140:130-136.

35. El-Sayed Moustafa, J. S., and P. Froguel. 2013. From obesity genetics to the future of personalized obesity therapy. Nature reviews. Endocrinology 9:402-413.

36. Dubern, B., K. Clement, V. Pelloux, P. Froguel, J. P. Girardet, B. Guy-Grand, and P. Tounian. 2001. Mutational analysis of melanocortin-4 receptor, agouti-related protein, and alpha-melanocyte-stimulating hormone genes in severely obese children. The Journal of pediatrics 139:204-209.

37. Coleman, D. L. 2010. A historical perspective on leptin. Nature medicine 16:1097-1099.

38. Farooqi, I. S., and S. O'Rahilly. 2008. Mutations in ligands and receptors of the leptin-melanocortin pathway that lead to obesity. Nature clinical practice. Endocrinology & metabolism 4:569-577.

39. Nolis, T. 2014. Exploring the pathophysiology behind the more common genetic and acquired lipodystrophies. Journal of human genetics 59:16-23.

40. Beales, P. L., N. Elcioglu, A. S. Woolf, D. Parker, and F. A. Flinter. 1999. New criteria for improved diagnosis of Bardet-Biedl syndrome: results of a population survey. Journal of medical genetics 36:437-446.

41. Scheidecker, S., C. Etard, N. W. Pierce, V. Geoffroy, E. Schaefer, J. Muller, K. Chennen, E. Flori, V. Pelletier, O. Poch, V. Marion, C. Stoetzel, U. Strahle, M. V. Nachury, and H. Dollfus. 2014. Exome sequencing of Bardet-Biedl syndrome patient identifies a null mutation in the BBSome subunit BBIP1 (BBS18). Journal of medical genetics 51:132-136.

42. Gul, D., G. Ogur, Y. Tunca, and O. Ozcan. 2002. Third case of WAGR syndrome with severe obesity and constitutional deletion of chromosome (11)(p12p14). American journal of medical genetics 107:70-71.

43. Girardet, C., and A. A. Butler. 2014. Neural melanocortin receptors in obesity and related metabolic disorders. Biochimica et biophysica acta 1842:482-494.

44. Rubio-Cabezas, O., V. Puri, I. Murano, V. Saudek, R. K. Semple, S. Dash, C. S. Hyden, W. Bottomley, C. Vigouroux, J. Magre, P. Raymond-Barker, P. R. Murgatroyd, A. Chawla, J. N. Skepper, V. K. Chatterjee, S. Suliman, A. M. Patch, A. K. Agarwal, A. Garg, I. Barroso, S. Cinti, M. P. Czech, J. Argente, S. O'Rahilly, D. B. Savage, and L. D. S. Consortium. 2009. Partial lipodystrophy and insulin resistant diabetes in a patient with a homozygous nonsense mutation in CIDEC. EMBO molecular medicine 1:280-287.

45. Garg, A., M. D. Hernandez, A. B. Sousa, L. Subramanyam, L. Martinez de Villarreal, H. G. dos Santos, and O. Barboza. 2010. An autosomal recessive syndrome of joint contractures, muscular atrophy, microcytic anemia, and panniculitis-associated lipodystrophy. The Journal of clinical endocrinology and metabolism 95:E58-63.

46. Agarwal, A. K., C. Xing, G. N. DeMartino, D. Mizrachi, M. D. Hernandez, A. B. Sousa, L. Martinez de Villarreal, H. G. dos Santos, and A. Garg. 2010. PSMB8 encoding the beta5i proteasome subunit is mutated in joint contractures, muscle atrophy, microcytic anemia, and panniculitis-induced lipodystrophy syndrome. American journal of human genetics 87:866-872.

47. Matthijs, G., E. Schollen, E. Pardon, M. Veiga-Da-Cunha, J. Jaeken, J. J. Cassiman, and E. Van Schaftingen. 1997. Mutations in PMM2, a phosphomannomutase gene on chromosome 16p13, in carbohydrate-

deficient glycoprotein type I syndrome (Jaeken syndrome). Nature genetics 16:88-92.

48. Eriksson, M., W. T. Brown, L. B. Gordon, M. W. Glynn, J. Singer, L. Scott, M. R. Erdos, C. M. Robbins, T. Y. Moses, P. Berglund, A. Dutra, E. Pak, S. Durkin, A. B. Csoka, M. Boehnke, T. W. Glover, and F. S. Collins. 2003. Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. Nature 423:293-298.

49. Plasilova, M., C. Chattopadhyay, P. Pal, N. A. Schaub, S. A. Buechner, H. Mueller, P. Miny, A. Ghosh, and K. Heinimann. 2004. Homozygous missense mutation in the lamin A/C gene causes autosomal recessive Hutchinson-Gilford progeria syndrome. Journal of medical genetics 41:609-614.

50. Agarwal, A. K., J. P. Fryns, R. J. Auchus, and A. Garg. 2003. Zinc metalloproteinase, ZMPSTE24, is mutated in mandibuloacral dysplasia. Human molecular genetics 12:1995-2001.

51. Hecht, J. T., O. J. Hood, R. J. Schwartz, J. C. Hennessey, B. A. Bernhardt, and W. A. Horton. 1988. Obesity in achondroplasia. American journal of medical genetics 31:597-602.

52. Mariot, V., S. Maupetit-Mehouas, C. Sinding, M. L. Kottler, and A. Linglart. 2008. A maternal epimutation of GNAS leads to Albright osteodystrophy and parathyroid hormone resistance. The Journal of clinical endocrinology and metabolism 93:661-665.

53. Collin, G. B., J. D. Marshall, A. Ikeda, W. V. So, I. Russell-Eggitt, P. Maffei, S. Beck, C. F. Boerkoel, N. Sicolo, M. Martin, P. M. Nishina, and J. K. Naggert.

2002. Mutations in ALMS1 cause obesity, type 2 diabetes and neurosensory degeneration in Alstrom syndrome. Nature genetics 31:74-78.

54. Williams, C. A., A. L. Beaudet, J. Clayton-Smith, J. H. Knoll, M. Kyllerman, L. A. Laan, R. E. Magenis, A. Moncla, A. A. Schinzel, J. A. Summers, and J. Wagstaff. 2006. Angelman syndrome 2005: updated consensus for diagnostic criteria. American journal of medical genetics. Part A 140:413-418.

55. Csoka, A. B., H. Cao, P. J. Sammak, D. Constantinescu, G. P. Schatten, and R. A. Hegele. 2004. Novel lamin A/C gene (LMNA) mutations in atypical progeroid syndromes. Journal of medical genetics 41:304-308.

56. M'Hamdi, O., I. Ouertani, and H. Chaabouni-Bouhamed. 2014. Update on the genetics of bardet-biedl syndrome. Molecular syndromology 5:51-56.

57. Leitch, C. C., N. A. Zaghloul, E. E. Davis, C. Stoetzel, A. Diaz-Font, S. Rix, M. Alfadhel, R. A. Lewis, W. Eyaid, E. Banin, H. Dollfus, P. L. Beales, J. L. Badano, and N. Katsanis. 2008. Hypomorphic mutations in syndromic encephalocele genes are associated with Bardet-Biedl syndrome. Nature genetics 40:443-448.

58. Kim, S. K., A. Shindo, T. J. Park, E. C. Oh, S. Ghosh, R. S. Gray, R. A. Lewis, C. A. Johnson, T. Attie-Bittach, N. Katsanis, and J. B. Wallingford. 2010. Planar cell polarity acts through septins to control collective cell movement and ciliogenesis. Science 329:1337-1340.

59. Badano, J. L., C. C. Leitch, S. J. Ansley, H. May-Simera, S. Lawson, R. A. Lewis, P. L. Beales, H. C. Dietz, S. Fisher, and N. Katsanis. 2006. Dissection of epistasis in oligogenic Bardet-Biedl syndrome. Nature 439:326-330.

60. Aldahmesh, M. A., Y. Li, A. Alhashem, S. Anazi, H. Alkuraya, M. Hashem, A. A. Awaji, S. Sogaty, A. Alkharashi, S. Alzahrani, S. A. Al Hazzaa, Y. Xiong, S. Kong, Z. Sun, and F. S. Alkuraya. 2014. IFT27, encoding a small GTPase component of IFT particles, is mutated in a consanguineous family with Bardet-Biedl syndrome. Human molecular genetics 23:3307-3315.

61. Lower, K. M., G. Solders, M. L. Bondeson, J. Nelson, A. Brun, J. Crawford, G. Malm, M. Borjeson, G. Turner, M. Partington, and J. Gecz. 2004. 1024C> T (R342X) is a recurrent PHF6 mutation also found in the original Borjeson-Forssman-Lehmann syndrome family. European journal of human genetics : EJHG 12:787-789.

62. Williams, S. R., M. A. Aldred, V. M. Der Kaloustian, F. Halal, G. Gowans, D. R. McLeod, S. Zondag, H. V. Toriello, R. E. Magenis, and S. H. Elsea. 2010. Haploinsufficiency of HDAC4 causes brachydactyly mental retardation syndrome, with brachydactyly type E, developmental delays, and behavioral problems. American journal of human genetics 87:219-228.

63. Kirschner, L. S., J. A. Carney, S. D. Pack, S. E. Taymans, C. Giatzakis, Y. S. Cho, Y. S. Cho-Chung, and C. A. Stratakis. 2000. Mutations of the gene encoding the protein kinase A type I-alpha regulatory subunit in patients with the Carney complex. Nature genetics 26:89-92.

64. Jenkins, D., D. Seelow, F. S. Jehee, C. A. Perlyn, L. G. Alonso, D. F. Bueno, D. Donnai, D. Josifova, I. M. Mathijssen, J. E. Morton, K. H. Orstavik, E. Sweeney, S. A. Wall, J. L. Marsh, P. Nurnberg, M. R. Passos-Bueno, and A. O. Wilkie. 2007. RAB23 mutations in Carpenter syndrome imply an unexpected role

for hedgehog signaling in cranial-suture development and obesity. American journal of human genetics 80:1162-1170.

65. Twigg, S. R., D. Lloyd, D. Jenkins, N. E. Elcioglu, C. D. Cooper, N. Al-Sannaa, A. Annagur, G. Gillessen-Kaesbach, I. Huning, S. J. Knight, J. A. Goodship, B. D. Keavney, P. L. Beales, O. Gileadi, S. J. McGowan, and A. O. Wilkie. 2012. Mutations in multidomain protein MEGF8 identify a Carpenter syndrome subtype associated with defective lateralization. American journal of human genetics 91:897-905.

66. Song, M. H., H. K. Lee, J. Y. Choi, S. Kim, J. Bok, and U. K. Kim. 2010. Clinical evaluation of DFN3 patients with deletions in the POU3F4 locus and detection of carrier female using MLPA. Clinical genetics 78:524-532.

67. Kolehmainen, J., R. Wilkinson, A. E. Lehesjoki, K. Chandler, S. Kivitie-Kallio, J. Clayton-Smith, A. L. Traskelin, L. Waris, A. Saarinen, J. Khan, V. Gross-Tsur, E. I. Traboulsi, M. Warburg, J. P. Fryns, R. Norio, G. C. Black, and F. D. Manson. 2004. Delineation of Cohen syndrome following a large-scale genotype-phenotype screen. American journal of human genetics 75:122-127.

68. Wu, W., J. D. Cogan, R. W. Pfaffle, J. S. Dasen, H. Frisch, S. M. O'Connell, S. E. Flynn, M. R. Brown, P. E. Mullis, J. S. Parks, J. A. Phillips, 3rd, and M. G. Rosenfeld. 1998. Mutations in PROP1 cause familial combined pituitary hormone deficiency. Nature genetics 18:147-149.

69. Lavery, G. G., E. A. Walker, A. Tiganescu, J. P. Ride, C. H. Shackleton, J. W. Tomlinson, J. M. Connell, D. W. Ray, A. Biason-Lauber, E. M. Malunowicz, W. Arlt, and P. M. Stewart. 2008. Steroid biomarkers and genetic studies reveal

inactivating mutations in hexose-6-phosphate dehydrogenase in patients with cortisone reductase deficiency. The Journal of clinical endocrinology and metabolism 93:3827-3832.

70. Santer, R., S. Groth, M. Kinner, A. Dombrowski, G. T. Berry, J. Brodehl, J. V. Leonard, S. Moses, S. Norgren, F. Skovby, R. Schneppenheim, B. Steinmann, and J. Schaub. 2002. The mutation spectrum of the facilitative glucose transporter gene SLC2A2 (GLUT2) in patients with Fanconi-Bickel syndrome. Human genetics 110:21-29.

71. de Vries, B. B., J. P. Fryns, M. G. Butler, F. Canziani, E. Wesby-van Swaay, J. O. van Hemel, B. A. Oostra, D. J. Halley, and M. F. Niermeijer. 1993. Clinical and molecular studies in fragile X patients with a Prader-Willi-like phenotype. Journal of medical genetics 30:761-766.

72. Odawara, M., T. Kadowaki, R. Yamamoto, Y. Shibasaki, K. Tobe, D. Accili, C. Bevins, Y. Mikami, N. Matsuura, Y. Akanuma, and et al. 1989. Human diabetes associated with a mutation in the tyrosine kinase domain of the insulin receptor. Science 245:66-68.

73. Phillips, J. A., 3rd, and J. D. Cogan. 1994. Genetic basis of endocrine disease. 6. Molecular basis of familial human growth hormone deficiency. The Journal of clinical endocrinology and metabolism 78:11-16.

74. Steinmuller, R., D. Steinberger, and U. Muller. 1998. MEHMO (mental retardation, epileptic seizures, hypogonadism and -genitalism, microcephaly, obesity), a novel syndrome: assignment of disease locus to xp21.1-p22.13. European journal of human genetics : EJHG 6:201-206.

75. Kleefstra, T., H. G. Yntema, A. R. Oudakker, T. Romein, E. Sistermans, W. Nillessen, H. van Bokhoven, B. B. de Vries, and B. C. Hamel. 2002. De novo MECP2 frameshift mutation in a boy with moderate mental retardation, obesity and gynaecomastia. Clinical genetics 61:359-362.

76. Ahmad, W., M. De Fusco, M. Faiyaz ul Haque, P. Aridon, T. Sarno, M. Sohail, S. ul Haque, M. Ahmad, A. Ballabio, B. Franco, and G. Casari. 1999. Linkage mapping of a new syndromic form of X-linked mental retardation, MRXS7, associated with obesity. European journal of human genetics : EJHG 7:828-832.

77. Shashi, V., M. N. Berry, S. Shoaf, J. J. Sciote, D. Goldstein, and T. C. Hart. 2000. A unique form of mental retardation with a distinctive phenotype maps to Xq26-q27. American journal of human genetics 66:469-479.

78. Simonds, W. F., S. Varghese, S. J. Marx, and L. K. Nieman. 2012. Cushing's syndrome in multiple endocrine neoplasia type 1. Clinical endocrinology 76:379-386.

79. Horsthemke, B., and J. Wagstaff. 2008. Mechanisms of imprinting of the Prader-Willi/Angelman region. American journal of medical genetics. Part A 146A:2041-2052.

80. Faivre, L., V. Cormier-Daire, J. M. Lapierre, L. Colleaux, S. Jacquemont, D. Genevieve, P. Saunier, A. Munnich, C. Turleau, S. Romana, M. Prieur, M. C. De Blois, and M. Vekemans. 2002. Deletion of the SIM1 gene (6q16.2) in a patient with a Prader-Willi-like phenotype. Journal of medical genetics 39:594-596.

81. Pfeiffer, R. A., L. Tschech, U. Irle, and G. F. Wundisch. 1987. [Chromosome aberrations in Prader-Willi-Labhart syndrome--critical review, documented by 4 unusual cases]. Klinische Padiatrie 199:329-335.

82. Veugelers, M., J. Vermeesch, K. Watanabe, Y. Yamaguchi, P. Marynen, and G. David. 1998. GPC4, the gene for human K-glypican, flanks GPC3 on xq26: deletion of the GPC3-GPC4 gene cluster in one family with Simpson-Golabi-Behmel syndrome. Genomics 53:1-11.

83. Pilia, G., R. M. Hughes-Benzie, A. MacKenzie, P. Baybayan, E. Y. Chen, R. Huber, G. Neri, A. Cao, A. Forabosco, and D. Schlessinger. 1996. Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. Nature genetics 12:241-247.

84. Budny, B., W. Chen, H. Omran, M. Fliegauf, A. Tzschach, M. Wisniewska, L. R. Jensen, M. Raynaud, S. A. Shoichet, M. Badura, S. Lenzner, A. Latos-Bielenska, and H. H. Ropers. 2006. A novel X-linked recessive mental retardation syndrome comprising macrocephaly and ciliary dysfunction is allelic to oral-facial-digital type I syndrome. Human genetics 120:171-178.

85. Usala, S. J., G. E. Tennyson, A. E. Bale, R. W. Lash, N. Gesundheit, F. E. Wondisford, D. Accili, P. Hauser, and B. D. Weintraub. 1990. A base mutation of the C-erbA beta thyroid hormone receptor in a kindred with generalized thyroid hormone resistance. Molecular heterogeneity in two other kindreds. The Journal of clinical investigation 85:93-100.

86. Bamshad, M., R. C. Lin, D. J. Law, W. C. Watkins, P. A. Krakowiak, M. E. Moore, P. Franceschini, R. Lala, L. B. Holmes, T. C. Gebuhr, B. G. Bruneau, A. Schinzel,

J. G. Seidman, C. E. Seidman, and L. B. Jorde. 1997. Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome. Nature genetics 16:311-315.

87. Harakalova, M., M. J. van den Boogaard, R. Sinke, S. van Lieshout, M. C. van Tuil, K. Duran, I. Renkens, P. A. Terhal, C. de Kovel, I. J. Nijman, M. van Haelst, N. V. Knoers, G. van Haaften, W. Kloosterman, R. C. Hennekam, E. Cuppen, and H. K. Ploos van Amstel. 2012. X-exome sequencing identifies a HDAC8 variant in a large pedigree with X-linked intellectual disability, truncal obesity, gynaecomastia, hypogonadism and unusual face. Journal of medical genetics 49:539-543.

88. Rankinen, T., A. Zuberi, Y. C. Chagnon, S. J. Weisnagel, G. Argyropoulos, B. Walts, L. Perusse, and C. Bouchard. 2006. The human obesity gene map: the 2005 update. Obesity 14:529-644.

89. Athanasiu, L., A. A. Brown, A. B. Birkenaes, M. Mattingsdal, I. Agartz, I. Melle, V. M. Steen, O. A. Andreassen, and S. Djurovic. 2012. Genome-wide association study identifies genetic loci associated with body mass index and high density lipoprotein-cholesterol levels during psychopharmacological treatment - a cross-sectional naturalistic study. Psychiatry research 197:327-336.

90. Boraska, V., A. Day-Williams, C. S. Franklin, K. S. Elliott, K. Panoutsopoulou, I. Tachmazidou, E. Albrecht, S. Bandinelli, L. J. Beilin, M. Bochud, G. Cadby, F. Ernst, D. M. Evans, C. Hayward, A. A. Hicks, J. Huffman, C. Huth, A. L. James, N. Klopp, I. Kolcic, Z. Kutalik, D. A. Lawlor, A. W. Musk, M. Pehlic, C. E. Pennell, J.

R. Perry, A. Peters, O. Polasek, B. St Pourcain, S. M. Ring, E. Salvi, S. Schipf, J. A. Staessen, A. Teumer, N. Timpson, V. Vitart, N. M. Warrington, H. Yaghootkar, T. Zemunik, L. Zgaga, P. An, V. Anttila, I. B. Borecki, J. Holmen, I. Ntalla, A. Palotie, K. H. Pietilainen, J. Wedenoja, B. S. Winsvold, G. V. Dedoussis, J. Kaprio, M. A. Province, J. A. Zwart, M. Burnier, H. Campbell, D. Cusi, G. D. Smith, T. M. Frayling, C. Gieger, L. J. Palmer, P. P. Pramstaller, I. Rudan, H. Volzke, H. E. Wichmann, A. F. Wright, and E. Zeggini. 2012. Genome-wide association study to identify common variants associated with brachial circumference: a meta-analysis of 14 cohorts. PloS one 7:e31369.

91. Bradfield, J. P., H. R. Taal, N. J. Timpson, A. Scherag, C. Lecoeur, N. M. Warrington, E. Hypponen, C. Holst, B. Valcarcel, E. Thiering, R. M. Salem, F. R. Schumacher, D. L. Cousminer, P. M. Sleiman, J. Zhao, R. I. Berkowitz, K. S. Vimaleswaran, I. Jarick, C. E. Pennell, D. M. Evans, B. St Pourcain, D. J. Berry, D. O. Mook-Kanamori, A. Hofman, F. Rivadeneira, A. G. Uitterlinden, C. M. van Duijn, R. J. van der Valk, J. C. de Jongste, D. S. Postma, D. I. Boomsma, W. J. Gauderman, M. T. Hassanein, C. M. Lindgren, R. Magi, C. A. Boreham, C. E. Neville, L. A. Moreno, P. Elliott, A. Pouta, A. L. Hartikainen, M. Li, O. Raitakari, T. Lehtimaki, J. G. Eriksson, A. Palotie, J. Dallongeville, S. Das, P. Deloukas, G. McMahon, S. M. Ring, J. P. Kemp, J. L. Buxton, A. I. Blakemore, M. Bustamante, M. Guxens, J. N. Hirschhorn, M. W. Gillman, E. Kreiner-Moller, H. Bisgaard, F. D. Gilliland, J. Heinrich, E. Wheeler, I. Barroso, S. O'Rahilly, A. Meirhaeghe, T. I. Sorensen, C. Power, L. J. Palmer, A. Hinney, E. Widen, I. S. Farooqi, M. I. McCarthy, P. Froguel, D. Meyre, J. Hebebrand, M. R. Jarvelin, V. W. Jaddoe, G.

D. Smith, H. Hakonarson, S. F. Grant, and C. Early Growth Genetics. 2012. A genome-wide association meta-analysis identifies new childhood obesity loci. Nature genetics 44:526-531.

92. Chambers, J. C., P. Elliott, D. Zabaneh, W. Zhang, Y. Li, P. Froguel, D. Balding, J. Scott, and J. S. Kooner. 2008. Common genetic variation near MC4R is associated with waist circumference and insulin resistance. Nature genetics 40:716-718.

93. Cho, Y. S., M. J. Go, Y. J. Kim, J. Y. Heo, J. H. Oh, H. J. Ban, D. Yoon, M. H. Lee, D. J. Kim, M. Park, S. H. Cha, J. W. Kim, B. G. Han, H. Min, Y. Ahn, M. S. Park, H. R. Han, H. Y. Jang, E. Y. Cho, J. E. Lee, N. H. Cho, C. Shin, T. Park, J. W. Park, J. K. Lee, L. Cardon, G. Clarke, M. I. McCarthy, J. Y. Lee, J. K. Lee, B. Oh, and H. L. Kim. 2009. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nature genetics 41:527-534.

94. Cotsapas, C., E. K. Speliotes, I. J. Hatoum, D. M. Greenawalt, R. Dobrin, P. Y. Lum, C. Suver, E. Chudin, D. Kemp, M. Reitman, B. F. Voight, B. M. Neale, E. E. Schadt, J. N. Hirschhorn, L. M. Kaplan, M. J. Daly, and G. Consortium. 2009. Common body mass index-associated variants confer risk of extreme obesity. Human molecular genetics 18:3502-3507.

95. Croteau-Chonka, D. C., A. F. Marvelle, E. M. Lange, N. R. Lee, L. S. Adair, L. A. Lange, and K. L. Mohlke. 2011. Genome-wide association study of anthropometric traits and evidence of interactions with age and study year in Filipino women. Obesity 19:1019-1027.

96. Dorajoo, R., A. I. Blakemore, X. Sim, R. T. Ong, D. P. Ng, M. Seielstad, T. Y. Wong, S. M. Saw, P. Froguel, J. Liu, and E. S. Tai. 2012. Replication of 13 obesity loci among Singaporean Chinese, Malay and Asian-Indian populations. International journal of obesity 36:159-163.

97. Foster, M. C., Q. Yang, S. J. Hwang, U. Hoffmann, and C. S. Fox. 2011. Heritability and genome-wide association analysis of renal sinus fat accumulation in the Framingham Heart Study. BMC medical genetics 12:148.

98. Fox, C. S., N. Heard-Costa, L. A. Cupples, J. Dupuis, R. S. Vasan, and L. D. Atwood. 2007. Genome-wide association to body mass index and waist circumference: the Framingham Heart Study 100K project. BMC medical genetics 8 Suppl 1:S18.

99. Frayling, T. M., N. J. Timpson, M. N. Weedon, E. Zeggini, R. M. Freathy, C. M. Lindgren, J. R. Perry, K. S. Elliott, H. Lango, N. W. Rayner, B. Shields, L. W. Harries, J. C. Barrett, S. Ellard, C. J. Groves, B. Knight, A. M. Patch, A. R. Ness, S. Ebrahim, D. A. Lawlor, S. M. Ring, Y. Ben-Shlomo, M. R. Jarvelin, U. Sovio, A. J. Bennett, D. Melzer, L. Ferrucci, R. J. Loos, I. Barroso, N. J. Wareham, F. Karpe, K. R. Owen, L. R. Cardon, M. Walker, G. A. Hitman, C. N. Palmer, A. S. Doney, A. D. Morris, G. D. Smith, A. T. Hattersley, and M. I. McCarthy. 2007. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316:889-894.

100. Gudbjartsson, D. F., H. Holm, O. S. Indridason, G. Thorleifsson, V. Edvardsson, P. Sulem, F. de Vegt, F. C. d'Ancona, M. den Heijer, J. F. Wetzels, L. Franzson, T. Rafnar, K. Kristjansson, U. S. Bjornsdottir, G. I. Eyjolfsson, L. A. Kiemeney,

A. Kong, R. Palsson, U. Thorsteinsdottir, and K. Stefansson. 2010. Association of variants at UMOD with chronic kidney disease and kidney stones-role of age and comorbid diseases. PLoS genetics 6:e1001039.

101. Heard-Costa, N. L., M. C. Zillikens, K. L. Monda, A. Johansson, T. B. Harris, M. Fu, T. Haritunians, M. F. Feitosa, T. Aspelund, G. Eiriksdottir, M. Garcia, L. J. Launer, A. V. Smith, B. D. Mitchell, P. F. McArdle, A. R. Shuldiner, S. J. Bielinski, E. Boerwinkle, F. Brancati, E. W. Demerath, J. S. Pankow, A. M. Arnold, Y. D. Chen, N. L. Glazer, B. McKnight, B. M. Psaty, J. I. Rotter, N. Amin, H. Campbell, U. Gyllensten, C. Pattaro, P. P. Pramstaller, I. Rudan, M. Struchalin, V. Vitart, X. Gao, A. Kraja, M. A. Province, Q. Zhang, L. D. Atwood, J. Dupuis, J. N. Hirschhorn, C. E. Jaquish, C. J. O'Donnell, R. S. Vasan, C. C. White, Y. S. Aulchenko, K. Estrada, A. Hofman, F. Rivadeneira, A. G. Uitterlinden, J. C. Witteman, B. A. Oostra, R. C. Kaplan, V. Gudnason, J. R. O'Connell, I. B. Borecki, C. M. van Duijn, L. A. Cupples, C. S. Fox, and K. E. North. 2009. NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. PLoS genetics 5:e1000539.

102. Heid, I. M., A. U. Jackson, J. C. Randall, T. W. Winkler, L. Qi, V. Steinthorsdottir, G. Thorleifsson, M. C. Zillikens, E. K. Speliotes, R. Magi, T. Workalemahu, C. C. White, N. Bouatia-Naji, T. B. Harris, S. I. Berndt, E. Ingelsson, C. J. Willer, M. N. Weedon, J. Luan, S. Vedantam, T. Esko, T. O. Kilpelainen, Z. Kutalik, S. Li, K. L. Monda, A. L. Dixon, C. C. Holmes, L. M. Kaplan, L. Liang, J. L. Min, M. F. Moffatt, C. Molony, G. Nicholson, E. E. Schadt, K. T. Zondervan, M. F. Feitosa, T. Ferreira, H. Lango Allen, R. J. Weyant, E. Wheeler, A. R. Wood, Magic, K.

Estrada, M. E. Goddard, G. Lettre, M. Mangino, D. R. Nyholt, S. Purcell, A. V. Smith, P. M. Visscher, J. Yang, S. A. McCarroll, J. Nemesh, B. F. Voight, D. Absher, N. Amin, T. Aspelund, L. Coin, N. L. Glazer, C. Hayward, N. L. Heard-Costa, J. J. Hottenga, A. Johansson, T. Johnson, M. Kaakinen, K. Kapur, S. Ketkar, J. W. Knowles, P. Kraft, A. T. Kraja, C. Lamina, M. F. Leitzmann, B. McKnight, A. P. Morris, K. K. Ong, J. R. Perry, M. J. Peters, O. Polasek, I. Prokopenko, N. W. Rayner, S. Ripatti, F. Rivadeneira, N. R. Robertson, S. Sanna, U. Sovio, I. Surakka, A. Teumer, S. van Wingerden, V. Vitart, J. H. Zhao, C. Cavalcanti-Proenca, P. S. Chines, E. Fisher, J. R. Kulzer, C. Lecoeur, N. Narisu, C. Sandholt, L. J. Scott, K. Silander, K. Stark, M. L. Tammesoo, T. M. Teslovich, N. J. Timpson, R. M. Watanabe, R. Welch, D. I. Chasman, M. N. Cooper, J. O. Jansson, J. Kettunen, R. W. Lawrence, N. Pellikka, M. Perola, L. Vandenput, H. Alavere, P. Almgren, L. D. Atwood, A. J. Bennett, R. Biffar, L. L. Bonnycastle, S. R. Bornstein, T. A. Buchanan, H. Campbell, I. N. Day, M. Dei, M. Dorr, P. Elliott, M. R. Erdos, J. G. Eriksson, N. B. Freimer, M. Fu, S. Gaget, E. J. Geus, A. P. Gjesing, H. Grallert, J. Grassler, C. J. Groves, C. Guiducci, A. L. Hartikainen, N. Hassanali, A. S. Havulinna, K. H. Herzig, A. A. Hicks, J. Hui, W. Igl, P. Jousilahti, A. Jula, E. Kajantie, L. Kinnunen, I. Kolcic, S. Koskinen, P. Kovacs, H. K. Kroemer, V. Krzelj, J. Kuusisto, K. Kvaloy, J. Laitinen, O. Lantieri, G. M. Lathrop, M. L. Lokki, R. N. Luben, B. Ludwig, W. L. McArdle, A. McCarthy, M. A. Morken, M. Nelis, M. J. Neville, G. Pare, A. N. Parker, J. F. Peden, I. Pichler, K. H. Pietilainen, C. G. Platou, A. Pouta, M. Ridderstrale, N. J. Samani, J. Saramies, J. Sinisalo, J. H. Smit, R. J. Strawbridge, H. M. Stringham,

A. J. Swift, M. Teder-Laving, B. Thomson, G. Usala, J. B. van Meurs, G. J. van Ommen, V. Vatin, C. B. Volpato, H. Wallaschofski, G. B. Walters, E. Widen, S. H. Wild, G. Willemsen, D. R. Witte, L. Zgaga, P. Zitting, J. P. Beilby, A. L. James, M. Kahonen, T. Lehtimaki, M. S. Nieminen, C. Ohlsson, L. J. Palmer, O. Raitakari, P. M. Ridker, M. Stumvoll, A. Tonjes, J. Viikari, B. Balkau, Y. Ben-Shlomo, R. N. Bergman, H. Boeing, G. D. Smith, S. Ebrahim, P. Froguel, T. Hansen, C. Hengstenberg, K. Hveem, B. Isomaa, T. Jorgensen, F. Karpe, K. T. Khaw, M. Laakso, D. A. Lawlor, M. Marre, T. Meitinger, A. Metspalu, K. Midthjell, O. Pedersen, V. Salomaa, P. E. Schwarz, T. Tuomi, J. Tuomilehto, T. T. Valle, N. J. Wareham, A. M. Arnold, J. S. Beckmann, S. Bergmann, E. Boerwinkle, D. I. Boomsma, M. J. Caulfield, F. S. Collins, G. Eiriksdottir, V. Gudnason, U. Gyllensten, A. Hamsten, A. T. Hattersley, A. Hofman, F. B. Hu, T. Illig, C. Iribarren, M. R. Jarvelin, W. H. Kao, J. Kaprio, L. J. Launer, P. B. Munroe, B. Oostra, B. W. Penninx, P. P. Pramstaller, B. M. Psaty, T. Quertermous, A. Rissanen, I. Rudan, A. R. Shuldiner, N. Soranzo, T. D. Spector, A. C. Syvanen, M. Uda, A. Uitterlinden, H. Volzke, P. Vollenweider, J. F. Wilson, J. C. Witteman, A. F. Wright, G. R. Abecasis, M. Boehnke, I. B. Borecki, P. Deloukas, T. M. Frayling, L. C. Groop, T. Haritunians, D. J. Hunter, R. C. Kaplan, K. E. North, J. R. O'Connell, L. Peltonen, D. Schlessinger, D. P. Strachan, J. N. Hirschhorn, T. L. Assimes, H. E. Wichmann, U. Thorsteinsdottir, C. M. van Duijn, K. Stefansson, L. A. Cupples, R. J. Loos, I. Barroso, M. I. McCarthy, C. S. Fox, K. L. Mohlke, and C. M. Lindgren. 2010. Meta-analysis identifies 13 new

loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nature genetics 42:949-960.

103. Hinney, A., T. T. Nguyen, A. Scherag, S. Friedel, G. Bronner, T. D. Muller, H. Grallert, T. Illig, H. E. Wichmann, W. Rief, H. Schafer, and J. Hebebrand. 2007. Genome wide association (GWA) study for early onset extreme obesity supports the role of fat mass and obesity associated gene (FTO) variants. PloS one 2:e1361.

104. Jiao, H., P. Arner, J. Hoffstedt, D. Brodin, B. Dubern, S. Czernichow, F. van't Hooft, T. Axelsson, O. Pedersen, T. Hansen, T. I. Sorensen, J. Hebebrand, J. Kere, K. Dahlman-Wright, A. Hamsten, K. Clement, and I. Dahlman. 2011. Genome wide association study identifies KCNMA1 contributing to human obesity. BMC medical genomics 4:51.

105. Johansson, A., F. Marroni, C. Hayward, C. S. Franklin, A. V. Kirichenko, I. Jonasson, A. A. Hicks, V. Vitart, A. Isaacs, T. Axenovich, S. Campbell, J. Floyd, N. Hastie, S. Knott, G. Lauc, I. Pichler, K. Rotim, S. H. Wild, I. V. Zorkoltseva, J. F. Wilson, I. Rudan, H. Campbell, C. Pattaro, P. Pramstaller, B. A. Oostra, A. F. Wright, C. M. van Duijn, Y. S. Aulchenko, U. Gyllensten, and E. Consortium. 2010. Linkage and genome-wide association analysis of obesity-related phenotypes: association of weight with the MGAT1 gene. Obesity 18:803-808.

106. Kilpelainen, T. O., M. C. Zillikens, A. Stancakova, F. M. Finucane, J. S. Ried, C. Langenberg, W. Zhang, J. S. Beckmann, J. Luan, L. Vandenput, U. Styrkarsdottir, Y. Zhou, A. V. Smith, J. H. Zhao, N. Amin, S. Vedantam, S. Y.

Shin, T. Haritunians, M. Fu, M. F. Feitosa, M. Kumari, B. V. Halldorsson, E. Tikkanen, M. Mangino, C. Hayward, C. Song, A. M. Arnold, Y. S. Aulchenko, B. A. Oostra, H. Campbell, L. A. Cupples, K. E. Davis, A. Doring, G. Eiriksdottir, K. Estrada, J. M. Fernandez-Real, M. Garcia, C. Gieger, N. L. Glazer, C. Guiducci, A. Hofman, S. E. Humphries, B. Isomaa, L. C. Jacobs, A. Jula, D. Karasik, M. K. Karlsson, K. T. Khaw, L. J. Kim, M. Kivimaki, N. Klopp, B. Kuhnel, J. Kuusisto, Y. Liu, O. Ljunggren, M. Lorentzon, R. N. Luben, B. McKnight, D. Mellstrom, B. D. Mitchell, V. Mooser, J. M. Moreno, S. Mannisto, J. R. O'Connell, L. Pascoe, L. Peltonen, B. Peral, M. Perola, B. M. Psaty, V. Salomaa, D. B. Savage, R. K. Semple, T. Skaric-Juric, G. Sigurdsson, K. S. Song, T. D. Spector, A. C. Syvanen, P. J. Talmud, G. Thorleifsson, U. Thorsteinsdottir, A. G. Uitterlinden, C. M. van Duijn, A. Vidal-Puig, S. H. Wild, A. F. Wright, D. J. Clegg, E. Schadt, J. F. Wilson, I. Rudan, S. Ripatti, I. B. Borecki, A. R. Shuldiner, E. Ingelsson, J. O. Jansson, R. C. Kaplan, V. Gudnason, T. B. Harris, L. Groop, D. P. Kiel, F. Rivadeneira, M. Walker, I. Barroso, P. Vollenweider, G. Waeber, J. C. Chambers, J. S. Kooner, N. Soranzo, J. N. Hirschhorn, K. Stefansson, H. E. Wichmann, C. Ohlsson, S. O'Rahilly, N. J. Wareham, E. K. Speliotes, C. S. Fox, M. Laakso, and R. J. Loos. 2011. Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. Nature genetics 43:753-760.

107. Lindgren, C. M., I. M. Heid, J. C. Randall, C. Lamina, V. Steinthorsdottir, L. Qi, E. K. Speliotes, G. Thorleifsson, C. J. Willer, B. M. Herrera, A. U. Jackson, N. Lim, P. Scheet, N. Soranzo, N. Amin, Y. S. Aulchenko, J. C. Chambers, A. Drong, J. Luan, H. N. Lyon, F. Rivadeneira, S. Sanna, N. J. Timpson, M. C. Zillikens, J. H.

Zhao, P. Almgren, S. Bandinelli, A. J. Bennett, R. N. Bergman, L. L. Bonnycastle, S. J. Bumpstead, S. J. Chanock, L. Cherkas, P. Chines, L. Coin, C. Cooper, G. Crawford, A. Doering, A. Dominiczak, A. S. Doney, S. Ebrahim, P. Elliott, M. R. Erdos, K. Estrada, L. Ferrucci, G. Fischer, N. G. Forouhi, C. Gieger, H. Grallert, C. J. Groves, S. Grundy, C. Guiducci, D. Hadley, A. Hamsten, A. S. Havulinna, A. Hofman, R. Holle, J. W. Holloway, T. Illig, B. Isomaa, L. C. Jacobs, K. Jameson, P. Jousilahti, F. Karpe, J. Kuusisto, J. Laitinen, G. M. Lathrop, D. A. Lawlor, M. Mangino, W. L. McArdle, T. Meitinger, M. A. Morken, A. P. Morris, P. Munroe, N. Narisu, A. Nordstrom, P. Nordstrom, B. A. Oostra, C. N. Palmer, F. Payne, J. F. Peden, I. Prokopenko, F. Renstrom, A. Ruokonen, V. Salomaa, M. S. Sandhu, L. J. Scott, A. Scuteri, K. Silander, K. Song, X. Yuan, H. M. Stringham, A. J. Swift, T. Tuomi, M. Uda, P. Vollenweider, G. Waeber, C. Wallace, G. B. Walters, M. N. Weedon, C. Wellcome Trust Case Control, J. C. Witteman, C. Zhang, W. Zhang, M. J. Caulfield, F. S. Collins, G. Davey Smith, I. N. Day, P. W. Franks, A. T. Hattersley, F. B. Hu, M. R. Jarvelin, A. Kong, J. S. Kooner, M. Laakso, E. Lakatta, V. Mooser, A. D. Morris, L. Peltonen, N. J. Samani, T. D. Spector, D. P. Strachan, T. Tanaka, J. Tuomilehto, A. G. Uitterlinden, C. M. van Duijn, N. J. Wareham, W. Hugh, C. Procardis, D. M. Waterworth, M. Boehnke, P. Deloukas, L. Groop, D. J. Hunter, U. Thorsteinsdottir, D. Schlessinger, H. E. Wichmann, T. M. Frayling, G. R. Abecasis, J. N. Hirschhorn, R. J. Loos, K. Stefansson, K. L. Mohlke, I. Barroso, M. I. McCarthy, and C. Giant. 2009. Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. PLoS genetics 5:e1000508.

108. Liu, J. Z., S. E. Medland, M. J. Wright, A. K. Henders, A. C. Heath, P. A. Madden, A. Duncan, G. W. Montgomery, N. G. Martin, and A. F. McRae. 2010. Genome-wide association study of height and body mass index in Australian twin families. Twin research and human genetics : the official journal of the International Society for Twin Studies 13:179-193.

109. Liu, X. G., L. J. Tan, S. F. Lei, Y. J. Liu, H. Shen, L. Wang, H. Yan, Y. F. Guo, D. H. Xiong, X. D. Chen, F. Pan, T. L. Yang, Y. P. Zhang, Y. Guo, N. L. Tang, X. Z. Zhu, H. Y. Deng, S. Levy, R. R. Recker, C. J. Papasian, and H. W. Deng. 2009. Genome-wide association and replication studies identified TRHR as an important gene for lean body mass. American journal of human genetics 84:418-423.

110. Liu, Y. Z., Y. F. Pei, J. F. Liu, F. Yang, Y. Guo, L. Zhang, X. G. Liu, H. Yan, L. Wang, Y. P. Zhang, S. Levy, R. R. Recker, and H. W. Deng. 2009. Powerful bivariate genome-wide association analyses suggest the SOX6 gene influencing both obesity and osteoporosis phenotypes in males. PloS one 4:e6827.

111. Loos, R. J., C. M. Lindgren, S. Li, E. Wheeler, J. H. Zhao, I. Prokopenko, M. Inouye, R. M. Freathy, A. P. Attwood, J. S. Beckmann, S. I. Berndt, L. C. Prostate, T. Ovarian Cancer Screening, K. B. Jacobs, S. J. Chanock, R. B. Hayes, S. Bergmann, A. J. Bennett, S. A. Bingham, M. Bochud, M. Brown, S. Cauchi, J. M. Connell, C. Cooper, G. D. Smith, I. Day, C. Dina, S. De, E. T. Dermitzakis, A. S. Doney, K. S. Elliott, P. Elliott, D. M. Evans, I. Sadaf Farooqi, P. Froguel, J. Ghori, C. J. Groves, R. Gwilliam, D. Hadley, A. S. Hall, A. T. Hattersley, J. Hebebrand, I. M. Heid, Kora, C. Lamina, C. Gieger, T. Illig, T. Meitinger, H. E. Wichmann, B.

Herrera, A. Hinney, S. E. Hunt, M. R. Jarvelin, T. Johnson, J. D. Jolley, F. Karpe, A. Keniry, K. T. Khaw, R. N. Luben, M. Mangino, J. Marchini, W. L. McArdle, R. McGinnis, D. Meyre, P. B. Munroe, A. D. Morris, A. R. Ness, M. J. Neville, A. C. Nica, K. K. Ong, S. O'Rahilly, K. R. Owen, C. N. Palmer, K. Papadakis, S. Potter, A. Pouta, L. Qi, S. Nurses' Health, J. C. Randall, N. W. Rayner, S. M. Ring, M. S. Sandhu, A. Scherag, M. A. Sims, K. Song, N. Soranzo, E. K. Speliotes, I. Diabetes Genetics, H. E. Syddall, S. A. Teichmann, N. J. Timpson, J. H. Tobias, M. Uda, N. I. A. S. Sardi, C. I. Vogel, C. Wallace, D. M. Waterworth, M. N. Weedon, C. Wellcome Trust Case Control, C. J. Willer, Fusion, Wraight, X. Yuan, E. Zeggini, J. N. Hirschhorn, D. P. Strachan, W. H. Ouwehand, M. J. Caulfield, N. J. Samani, T. M. Frayling, P. Vollenweider, G. Waeber, V. Mooser, P. Deloukas, M. I. McCarthy, N. J. Wareham, I. Barroso, K. B. Jacobs, S. J. Chanock, R. B. Hayes, C. Lamina, C. Gieger, T. Illig, T. Meitinger, H. E. Wichmann, P. Kraft, S. E. Hankinson, D. J. Hunter, F. B. Hu, H. N. Lyon, B. F. Voight, M. Ridderstrale, L. Groop, P. Scheet, S. Sanna, G. R. Abecasis, G. Albai, R. Nagaraja, D. Schlessinger, A. U. Jackson, J. Tuomilehto, F. S. Collins, M. Boehnke, and K. L. Mohlke. 2008. Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nature genetics 40:768-775.

112. Lowe, J. K., J. B. Maller, I. Pe'er, B. M. Neale, J. Salit, E. E. Kenny, J. L. Shea, R. Burkhardt, J. G. Smith, W. Ji, M. Noel, J. N. Foo, M. L. Blundell, V. Skilling, L. Garcia, M. L. Sullivan, H. E. Lee, A. Labek, H. Ferdowsian, S. B. Auerbach, R. P. Lifton, C. Newton-Cheh, J. L. Breslow, M. Stoffel, M. J. Daly, D. M. Altshuler, and J. M. Friedman. 2009. Genome-wide association studies in an isolated

founder population from the Pacific Island of Kosrae. PLoS genetics 5:e1000365.

113. Melka, M. G., M. Bernard, A. Mahboubi, M. Abrahamowicz, A. D. Paterson, C. Syme, A. Lourdusamy, G. Schumann, G. T. Leonard, M. Perron, L. Richer, S. Veillette, D. Gaudet, T. Paus, and Z. Pausova. 2012. Genome-wide scan for loci of adolescent obesity and their relationship with blood pressure. The Journal of clinical endocrinology and metabolism 97:E145-150.

114. Meyre, D., J. Delplanque, J. C. Chevre, C. Lecoeur, S. Lobbens, S. Gallina, E. Durand, V. Vatin, F. Degraeve, C. Proenca, S. Gaget, A. Korner, P. Kovacs, W. Kiess, J. Tichet, M. Marre, A. L. Hartikainen, F. Horber, N. Potoczna, S. Hercberg, C. Levy-Marchal, F. Pattou, B. Heude, M. Tauber, M. I. McCarthy, A. I. Blakemore, A. Montpetit, C. Polychronakos, J. Weill, L. J. Coin, J. Asher, P. Elliott, M. R. Jarvelin, S. Visvikis-Siest, B. Balkau, R. Sladek, D. Balding, A. Walley, C. Dina, and P. Froguel. 2009. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. Nature genetics 41:157-159.

115. Ng, M. C., J. M. Hester, M. R. Wing, J. Li, J. Xu, P. J. Hicks, B. H. Roh, L. Lu, J. Divers, C. D. Langefeld, B. I. Freedman, N. D. Palmer, and D. W. Bowden. 2012. Genome-wide association of BMI in African Americans. Obesity 20:622-627.

116. Okada, Y., M. Kubo, H. Ohmiya, A. Takahashi, N. Kumasaka, N. Hosono, S. Maeda, W. Wen, R. Dorajoo, M. J. Go, W. Zheng, N. Kato, J. Y. Wu, Q. Lu, G. consortium, T. Tsunoda, K. Yamamoto, Y. Nakamura, N. Kamatani, and T.

Tanaka. 2012. Common variants at CDKAL1 and KLF9 are associated with body mass index in east Asian populations. Nature genetics 44:302-306.

117. Paternoster, L., D. M. Evans, E. A. Nohr, C. Holst, V. Gaborieau, P. Brennan, A. P. Gjesing, N. Grarup, D. R. Witte, T. Jorgensen, A. Linneberg, T. Lauritzen, A. Sandbaek, T. Hansen, O. Pedersen, K. S. Elliott, J. P. Kemp, B. St Pourcain, G. McMahon, D. Zelenika, J. Hager, M. Lathrop, N. J. Timpson, G. D. Smith, and T. I. Sorensen. 2011. Genome-wide population-based association study of extremely overweight young adults--the GOYA study. PloS one 6:e24303.

118. Polasek, O., A. Marusic, K. Rotim, C. Hayward, V. Vitart, J. Huffman, S. Campbell, S. Jankovic, M. Boban, Z. Biloglav, I. Kolcic, V. Krzelj, J. Terzic, L. Matec, G. Tometic, D. Nonkovic, J. Nincevic, M. Pehlic, J. Zedelj, V. Velagic, D. Juricic, I. Kirac, S. Belak Kovacevic, A. F. Wright, H. Campbell, and I. Rudan. 2009. Genome-wide association study of anthropometric traits in Korcula Island, Croatia. Croatian medical journal 50:7-16.

119. Scherag, A., C. Dina, A. Hinney, V. Vatin, S. Scherag, C. I. Vogel, T. D. Muller, H. Grallert, H. E. Wichmann, B. Balkau, B. Heude, M. R. Jarvelin, A. L. Hartikainen, C. Levy-Marchal, J. Weill, J. Delplanque, A. Korner, W. Kiess, P. Kovacs, N. W. Rayner, I. Prokopenko, M. I. McCarthy, H. Schafer, I. Jarick, H. Boeing, E. Fisher, T. Reinehr, J. Heinrich, P. Rzehak, D. Berdel, M. Borte, H. Biebermann, H. Krude, D. Rosskopf, C. Rimmbach, W. Rief, T. Fromme, M. Klingenspor, A. Schurmann, N. Schulz, M. M. Nothen, T. W. Muhleisen, R. Erbel, K. H. Jockel, S. Moebus, T. Boes, T. Illig, P. Froguel, J. Hebebrand, and D. Meyre. 2010. Two new Loci for body-weight regulation identified in a joint

analysis of genome-wide association studies for early-onset extreme obesity in French and german study groups. PLoS genetics 6:e1000916.

120. Scuteri, A., S. Sanna, W. M. Chen, M. Uda, G. Albai, J. Strait, S. Najjar, R. Nagaraja, M. Orru, G. Usala, M. Dei, S. Lai, A. Maschio, F. Busonero, A. Mulas, G. B. Ehret, A. A. Fink, A. B. Weder, R. S. Cooper, P. Galan, A. Chakravarti, D. Schlessinger, A. Cao, E. Lakatta, and G. R. Abecasis. 2007. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS genetics 3:e115.

121. Smith, E. N., W. Chen, M. Kahonen, J. Kettunen, T. Lehtimaki, L. Peltonen, O. T. Raitakari, R. M. Salem, N. J. Schork, M. Shaw, S. R. Srinivasan, E. J. Topol, J. S. Viikari, G. S. Berenson, and S. S. Murray. 2010. Longitudinal genome-wide association of cardiovascular disease risk factors in the Bogalusa heart study. PLoS genetics 6:e1001094.

122. Speliotes, E. K., C. J. Willer, S. I. Berndt, K. L. Monda, G. Thorleifsson, A. U. Jackson, H. Lango Allen, C. M. Lindgren, J. Luan, R. Magi, J. C. Randall, S. Vedantam, T. W. Winkler, L. Qi, T. Workalemahu, I. M. Heid, V. Steinthorsdottir, H. M. Stringham, M. N. Weedon, E. Wheeler, A. R. Wood, T. Ferreira, R. J. Weyant, A. V. Segre, K. Estrada, L. Liang, J. Nemesh, J. H. Park, S. Gustafsson, T. O. Kilpelainen, J. Yang, N. Bouatia-Naji, T. Esko, M. F. Feitosa, Z. Kutalik, M. Mangino, S. Raychaudhuri, A. Scherag, A. V. Smith, R. Welch, J. H. Zhao, K. K. Aben, D. M. Absher, N. Amin, A. L. Dixon, E. Fisher, N. L. Glazer, M. E. Goddard, N. L. Heard-Costa, V. Hoesel, J. J. Hottenga, A. Johansson, T. Johnson, S. Ketkar, C. Lamina, S. Li, M. F. Moffatt, R. H. Myers, N. Narisu, J. R.

Perry, M. J. Peters, M. Preuss, S. Ripatti, F. Rivadeneira, C. Sandholt, L. J. Scott, N. J. Timpson, J. P. Tyrer, S. van Wingerden, R. M. Watanabe, C. C. White, F. Wiklund, C. Barlassina, D. I. Chasman, M. N. Cooper, J. O. Jansson, R. W. Lawrence, N. Pellikka, I. Prokopenko, J. Shi, E. Thiering, H. Alavere, M. T. Alibrandi, P. Almgren, A. M. Arnold, T. Aspelund, L. D. Atwood, B. Balkau, A. J. Balmforth, A. J. Bennett, Y. Ben-Shlomo, R. N. Bergman, S. Bergmann, H. Biebermann, A. I. Blakemore, T. Boes, L. L. Bonnycastle, S. R. Bornstein, M. J. Brown, T. A. Buchanan, F. Busonero, H. Campbell, F. P. Cappuccio, C. Cavalcanti-Proenca, Y. D. Chen, C. M. Chen, P. S. Chines, R. Clarke, L. Coin, J. Connell, I. N. Day, M. den Heijer, J. Duan, S. Ebrahim, P. Elliott, R. Elosua, G. Eiriksdottir, M. R. Erdos, J. G. Eriksson, M. F. Facheris, S. B. Felix, P. Fischer-Posovszky, A. R. Folsom, N. Friedrich, N. B. Freimer, M. Fu, S. Gaget, P. V. Gejman, E. J. Geus, C. Gieger, A. P. Gjesing, A. Goel, P. Goyette, H. Grallert, J. Grassler, D. M. Greenawalt, C. J. Groves, V. Gudnason, C. Guiducci, A. L. Hartikainen, N. Hassanali, A. S. Hall, A. S. Havulinna, C. Hayward, A. C. Heath, C. Hengstenberg, A. A. Hicks, A. Hinney, A. Hofman, G. Homuth, J. Hui, W. Igl, C. Iribarren, B. Isomaa, K. B. Jacobs, I. Jarick, E. Jewell, U. John, T. Jorgensen, P. Jousilahti, A. Jula, M. Kaakinen, E. Kajantie, L. M. Kaplan, S. Kathiresan, J. Kettunen, L. Kinnunen, J. W. Knowles, I. Kolcic, I. R. Konig, S. Koskinen, P. Kovacs, J. Kuusisto, P. Kraft, K. Kvaloy, J. Laitinen, O. Lantieri, C. Lanzani, L. J. Launer, C. Lecoeur, T. Lehtimaki, G. Lettre, J. Liu, M. L. Lokki, M. Lorentzon, R. N. Luben, B. Ludwig, Magic, P. Manunta, D. Marek, M. Marre, N. G. Martin, W. L. McArdle, A. McCarthy, B. McKnight, T. Meitinger, O. Melander, D. Meyre, K.

Midthjell, G. W. Montgomery, M. A. Morken, A. P. Morris, R. Mulic, J. S. Ngwa, M. Nelis, M. J. Neville, D. R. Nyholt, C. J. O'Donnell, S. O'Rahilly, K. K. Ong, B. Oostra, G. Pare, A. N. Parker, M. Perola, I. Pichler, K. H. Pietilainen, C. G. Platou, O. Polasek, A. Pouta, S. Rafelt, O. Raitakari, N. W. Rayner, M. Ridderstrale, W. Rief, A. Ruokonen, N. R. Robertson, P. Rzehak, V. Salomaa, A. R. Sanders, M. S. Sandhu, S. Sanna, J. Saramies, M. J. Savolainen, S. Scherag, S. Schipf, S. Schreiber, H. Schunkert, K. Silander, J. Sinisalo, D. S. Siscovick, J. H. Smit, N. Soranzo, U. Sovio, J. Stephens, I. Surakka, A. J. Swift, M. L. Tammesoo, J. C. Tardif, M. Teder-Laving, T. M. Teslovich, J. R. Thompson, B. Thomson, A. Tonjes, T. Tuomi, J. B. van Meurs, G. J. van Ommen, V. Vatin, J. Viikari, S. Visvikis-Siest, V. Vitart, C. I. Vogel, B. F. Voight, L. L. Waite, H. Wallaschofski, G. B. Walters, E. Widen, S. Wiegand, S. H. Wild, G. Willemsen, D. R. Witte, J. C. Witteman, J. Xu, Q. Zhang, L. Zgaga, A. Ziegler, P. Zitting, J. P. Beilby, I. S. Farooqi, J. Hebebrand, H. V. Huikuri, A. L. James, M. Kahonen, D. F. Levinson, F. Macciardi, M. S. Nieminen, C. Ohlsson, L. J. Palmer, P. M. Ridker, M. Stumvoll, J. S. Beckmann, H. Boeing, E. Boerwinkle, D. I. Boomsma, M. J. Caulfield, S. J. Chanock, F. S. Collins, L. A. Cupples, G. D. Smith, J. Erdmann, P. Froguel, H. Gronberg, U. Gyllensten, P. Hall, T. Hansen, T. B. Harris, A. T. Hattersley, R. B. Hayes, J. Heinrich, F. B. Hu, K. Hveem, T. Illig, M. R. Jarvelin, J. Kaprio, F. Karpe, K. T. Khaw, L. A. Kiemeney, H. Krude, M. Laakso, D. A. Lawlor, A. Metspalu, P. B. Munroe, W. H. Ouwehand, O. Pedersen, B. W. Penninx, A. Peters, P. P. Pramstaller, T. Quertermous, T. Reinehr, A. Rissanen, I. Rudan, N. J. Samani, P. E. Schwarz, A. R. Shuldiner, T. D. Spector, J.

Tuomilehto, M. Uda, A. Uitterlinden, T. T. Valle, M. Wabitsch, G. Waeber, N. J. Wareham, H. Watkins, C. Procardis, J. F. Wilson, A. F. Wright, M. C. Zillikens, N. Chatterjee, S. A. McCarroll, S. Purcell, E. E. Schadt, P. M. Visscher, T. L. Assimes, I. B. Borecki, P. Deloukas, C. S. Fox, L. C. Groop, T. Haritunians, D. J. Hunter, R. C. Kaplan, K. L. Mohlke, J. R. O'Connell, L. Peltonen, D. Schlessinger, D. P. Strachan, C. M. van Duijn, H. E. Wichmann, T. M. Frayling, U. Thorsteinsdottir, G. R. Abecasis, I. Barroso, M. Boehnke, K. Stefansson, K. E. North, M. I. McCarthy, J. N. Hirschhorn, E. Ingelsson, and R. J. Loos. 2010. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nature genetics 42:937-948.

123. Thorleifsson, G., G. B. Walters, D. F. Gudbjartsson, V. Steinthorsdottir, P. Sulem, A. Helgadottir, U. Styrkarsdottir, S. Gretarsdottir, S. Thorlacius, I. Jonsdottir, T. Jonsdottir, E. J. Olafsdottir, G. H. Olafsdottir, T. Jonsson, F. Jonsson, K. Borch-Johnsen, T. Hansen, G. Andersen, T. Jorgensen, T. Lauritzen, K. K. Aben, A. L. Verbeek, N. Roeleveld, E. Kampman, L. R. Yanek, L. C. Becker, L. Tryggvadottir, T. Rafnar, D. M. Becker, J. Gulcher, L. A. Kiemeney, O. Pedersen, A. Kong, U. Thorsteinsdottir, and K. Stefansson. 2009. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nature genetics 41:18-24.

124. Wan, E. S., M. H. Cho, N. Boutaoui, B. J. Klanderman, J. S. Sylvia, J. P. Ziniti, S. Won, C. Lange, S. G. Pillai, W. H. Anderson, X. Kong, D. A. Lomas, P. S. Bakke, A. Gulsvik, E. A. Regan, J. R. Murphy, B. J. Make, J. D. Crapo, E. F. Wouters, B. R. Celli, E. K. Silverman, D. L. DeMeo, E.-P. Evaluation of Chronic Obstructive

Pulmonary Disease Longitudinally to Identify Predictive Surrogate, c. Norway-Bergen, T. National Emphysema Treatment, and C. G. investigators. 2011. Genome-wide association analysis of body mass in chronic obstructive pulmonary disease. American journal of respiratory cell and molecular biology 45:304-310.

125. Wang, K., W. D. Li, C. K. Zhang, Z. Wang, J. T. Glessner, S. F. Grant, H. Zhao, H. Hakonarson, and R. A. Price. 2011. A genome-wide association study on obesity and obesity-related traits. PloS one 6:e18939.

126. Wang, K. S., X. Liu, S. Zheng, M. Zeng, Y. Pan, and K. Callahan. 2012. A novel locus for body mass index on 5p15.2: a meta-analysis of two genome-wide association studies. Gene 500:80-84.

127. Wen, W., Y. S. Cho, W. Zheng, R. Dorajoo, N. Kato, L. Qi, C. H. Chen, R. J. Delahanty, Y. Okada, Y. Tabara, D. Gu, D. Zhu, C. A. Haiman, Z. Mo, Y. T. Gao, S. M. Saw, M. J. Go, F. Takeuchi, L. C. Chang, Y. Kokubo, J. Liang, M. Hao, L. Le Marchand, Y. Zhang, Y. Hu, T. Y. Wong, J. Long, B. G. Han, M. Kubo, K. Yamamoto, M. H. Su, T. Miki, B. E. Henderson, H. Song, A. Tan, J. He, D. P. Ng, Q. Cai, T. Tsunoda, F. J. Tsai, N. Iwai, G. K. Chen, J. Shi, J. Xu, X. Sim, Y. B. Xiang, S. Maeda, R. T. Ong, C. Li, Y. Nakamura, T. Aung, N. Kamatani, J. J. Liu, W. Lu, M. Yokota, M. Seielstad, C. S. Fann, A. T. C. Genetic Investigation of, J. Y. Wu, J. Y. Lee, F. B. Hu, T. Tanaka, E. S. Tai, and X. O. Shu. 2012. Meta-analysis identifies common variants associated with body mass index in east Asians. Nature genetics 44:307-311.

128. Willer, C. J., E. K. Speliotes, R. J. Loos, S. Li, C. M. Lindgren, I. M. Heid, S. I. Berndt, A. L. Elliott, A. U. Jackson, C. Lamina, G. Lettre, N. Lim, H. N. Lyon, S. A. McCarroll, K. Papadakis, L. Qi, J. C. Randall, R. M. Roccasecca, S. Sanna, P. Scheet, M. N. Weedon, E. Wheeler, J. H. Zhao, L. C. Jacobs, I. Prokopenko, N. Soranzo, T. Tanaka, N. J. Timpson, P. Almgren, A. Bennett, R. N. Bergman, S. A. Bingham, L. L. Bonnycastle, M. Brown, N. P. Burtt, P. Chines, L. Coin, F. S. Collins, J. M. Connell, C. Cooper, G. D. Smith, E. M. Dennison, P. Deodhar, P. Elliott, M. R. Erdos, K. Estrada, D. M. Evans, L. Gianniny, C. Gieger, C. J. Gillson, C. Guiducci, R. Hackett, D. Hadley, A. S. Hall, A. S. Havulinna, J. Hebebrand, A. Hofman, B. Isomaa, K. B. Jacobs, T. Johnson, P. Jousilahti, Z. Jovanovic, K. T. Khaw, P. Kraft, M. Kuokkanen, J. Kuusisto, J. Laitinen, E. G. Lakatta, J. Luan, R. N. Luben, M. Mangino, W. L. McArdle, T. Meitinger, A. Mulas, P. B. Munroe, N. Narisu, A. R. Ness, K. Northstone, S. O'Rahilly, C. Purmann, M. G. Rees, M. Ridderstrale, S. M. Ring, F. Rivadeneira, A. Ruokonen, M. S. Sandhu, J. Saramies, L. J. Scott, A. Scuteri, K. Silander, M. A. Sims, K. Song, J. Stephens, S. Stevens, H. M. Stringham, Y. C. Tung, T. T. Valle, C. M. Van Duijn, K. S. Vimaleswaran, P. Vollenweider, G. Waeber, C. Wallace, R. M. Watanabe, D. M. Waterworth, N. Watkins, C. Wellcome Trust Case Control, J. C. Witteman, E. Zeggini, G. Zhai, M. C. Zillikens, D. Altshuler, M. J. Caulfield, S. J. Chanock, I. S. Farooqi, L. Ferrucci, J. M. Guralnik, A. T. Hattersley, F. B. Hu, M. R. Jarvelin, M. Laakso, V. Mooser, K. K. Ong, W. H. Ouwehand, V. Salomaa, N. J. Samani, T. D. Spector, T. Tuomi, J. Tuomilehto, M. Uda, A. G. Uitterlinden, N. J. Wareham, P. Deloukas, T. M. Frayling, L. C. Groop, R. B. Hayes, D. J. Hunter, K. L. Mohlke, L.

Peltonen, D. Schlessinger, D. P. Strachan, H. E. Wichmann, M. I. McCarthy, M. Boehnke, I. Barroso, G. R. Abecasis, J. N. Hirschhorn, and A. T. C. Genetic Investigation of. 2009. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nature genetics 41:25-34.

129. Zabaneh, D., and D. J. Balding. 2010. A genome-wide association study of the metabolic syndrome in Indian Asian men. PloS one 5:e11961.

130. Locke, A. E., B. Kahali, S. I. Berndt, A. E. Justice, T. H. Pers, F. R. Day, C. Powell, S. Vedantam, M. L. Buchkovich, J. Yang, D. C. Croteau-Chonka, T. Esko, T. Fall, T. Ferreira, S. Gustafsson, Z. Kutalik, J. Luan, R. Magi, J. C. Randall, T. W. Winkler, A. R. Wood, T. Workalemahu, J. D. Faul, J. A. Smith, J. Hua Zhao, W. Zhao, J. Chen, R. Fehrmann, A. K. Hedman, J. Karjalainen, E. M. Schmidt, D. Absher, N. Amin, D. Anderson, M. Beekman, J. L. Bolton, J. L. Bragg-Gresham, S. Buyske, A. Demirkan, G. Deng, G. B. Ehret, B. Feenstra, M. F. Feitosa, K. Fischer, A. Goel, J. Gong, A. U. Jackson, S. Kanoni, M. E. Kleber, K. Kristiansson, U. Lim, V. Lotay, M. Mangino, I. Mateo Leach, C. Medina-Gomez, S. E. Medland, M. A. Nalls, C. D. Palmer, D. Pasko, S. Pechlivanis, M. J. Peters, I. Prokopenko, D. Shungin, A. Stancakova, R. J. Strawbridge, Y. Ju Sung, T. Tanaka, A. Teumer, S. Trompet, S. W. van der Laan, J. van Setten, J. V. Van Vliet-Ostaptchouk, Z. Wang, L. Yengo, W. Zhang, A. Isaacs, E. Albrecht, J. Arnlov, G. M. Arscott, A. P. Attwood, S. Bandinelli, A. Barrett, I. N. Bas, C. Bellis, A. J. Bennett, C. Berne, R. Blagieva, M. Bluher, S. Bohringer, L. L. Bonnycastle, Y. Bottcher, H. A. Boyd, M. Bruinenberg, I. H. Caspersen, Y. D. Ida Chen, R. Clarke, E. W. Daw, A. J. de

Craen, G. Delgado, M. Dimitriou, A. S. Doney, N. Eklund, K. Estrada, E. Eury, L. Folkersen, R. M. Fraser, M. E. Garcia, F. Geller, V. Giedraitis, B. Gigante, A. S. Go, A. Golay, A. H. Goodall, S. D. Gordon, M. Gorski, H. J. Grabe, H. Grallert, T. B. Grammer, J. Grassler, H. Gronberg, C. J. Groves, G. Gusto, J. Haessler, P. Hall, T. Haller, G. Hallmans, C. A. Hartman, M. Hassinen, C. Hayward, N. L. Heard-Costa, Q. Helmer, C. Hengstenberg, O. Holmen, J. J. Hottenga, A. L. James, J. M. Jeff, A. Johansson, J. Jolley, T. Juliusdottir, L. Kinnunen, W. Koenig, M. Koskenvuo, W. Kratzer, J. Laitinen, C. Lamina, K. Leander, N. R. Lee, P. Lichtner, L. Lind, J. Lindstrom, K. Sin Lo, S. Lobbens, R. Lorbeer, Y. Lu, F. Mach, P. K. Magnusson, A. Mahajan, W. L. McArdle, S. McLachlan, C. Menni, S. Merger, E. Mihailov, L. Milani, A. Moayyeri, K. L. Monda, M. A. Morken, A. Mulas, G. Muller, M. Muller-Nurasyid, A. W. Musk, R. Nagaraja, M. M. Nothen, I. M. Nolte, S. Pilz, N. W. Rayner, F. Renstrom, R. Rettig, J. S. Ried, S. Ripke, N. R. Robertson, L. M. Rose, S. Sanna, H. Scharnagl, S. Scholtens, F. R. Schumacher, W. R. Scott, T. Seufferlein, J. Shi, A. Vernon Smith, J. Smolonska, A. V. Stanton, V. Steinthorsdottir, K. Stirrups, H. M. Stringham, J. Sundstrom, M. A. Swertz, A. J. Swift, A. C. Syvanen, S. T. Tan, B. O. Tayo, B. Thorand, G. Thorleifsson, J. P. Tyrer, H. W. Uh, L. Vandenput, F. C. Verhulst, S. H. Vermeulen, N. Verweij, J. M. Vonk, L. L. Waite, H. R. Warren, D. Waterworth, M. N. Weedon, L. R. Wilkens, C. Willenborg, T. Wilsgaard, M. K. Wojczynski, A. Wong, A. F. Wright, Q. Zhang, S. LifeLines Cohort, E. P. Brennan, M. Choi, Z. Dastani, A. W. Drong, P. Eriksson, A. Franco-Cereceda, J. R. Gadin, A. G. Gharavi, M. E. Goddard, R. E. Handsaker, J. Huang, F. Karpe, S. Kathiresan, S.

Keildson, K. Kiryluk, M. Kubo, J. Y. Lee, L. Liang, R. P. Lifton, B. Ma, S. A. McCarroll, A. J. McKnight, J. L. Min, M. F. Moffatt, G. W. Montgomery, J. M. Murabito, G. Nicholson, D. R. Nyholt, Y. Okada, J. R. Perry, R. Dorajoo, E. Reinmaa, R. M. Salem, N. Sandholm, R. A. Scott, L. Stolk, A. Takahashi, T. Tanaka, F. M. Van't Hooft, A. A. Vinkhuyzen, H. J. Westra, W. Zheng, K. T. Zondervan, A. D. Consortium, A.-B. W. Group, C. A. D. Consortium, C. K. Consortium, Glgc, Icbp, M. Investigators, T. C. Mu, M. I. Consortium, P. Consortium, C. ReproGen, G. Consortium, C. International Endogene, A. C. Heath, D. Arveiler, S. J. Bakker, J. Beilby, R. N. Bergman, J. Blangero, P. Bovet, H. Campbell, M. J. Caulfield, G. Cesana, A. Chakravarti, D. I. Chasman, P. S. Chines, F. S. Collins, D. C. Crawford, L. A. Cupples, D. Cusi, J. Danesh, U. de Faire, H. M. den Ruijter, A. F. Dominiczak, R. Erbel, J. Erdmann, J. G. Eriksson, M. Farrall, S. B. Felix, E. Ferrannini, J. Ferrieres, I. Ford, N. G. Forouhi, T. Forrester, O. H. Franco, R. T. Gansevoort, P. V. Gejman, C. Gieger, O. Gottesman, V. Gudnason, U. Gyllensten, A. S. Hall, T. B. Harris, A. T. Hattersley, A. A. Hicks, L. A. Hindorff, A. D. Hingorani, A. Hofman, G. Homuth, G. K. Hovingh, S. E. Humphries, S. C. Hunt, E. Hypponen, T. Illig, K. B. Jacobs, M. R. Jarvelin, K. H. Jockel, B. Johansen, P. Jousilahti, J. W. Jukema, A. M. Jula, J. Kaprio, J. J. Kastelein, S. M. Keinanen-Kiukaanniemi, L. A. Kiemeney, P. Knekt, J. S. Kooner, C. Kooperberg, P. Kovacs, A. T. Kraja, M. Kumari, J. Kuusisto, T. A. Lakka, C. Langenberg, L. Le Marchand, T. Lehtimaki, V. Lyssenko, S. Mannisto, A. Marette, T. C. Matise, C. A. McKenzie, B. McKnight, F. L. Moll, A. D. Morris, A. P. Morris, J. C. Murray, M. Nelis, C. Ohlsson, A. J. Oldehinkel, K. K.

Ong, P. A. Madden, G. Pasterkamp, J. F. Peden, A. Peters, D. S. Postma, P. P. Pramstaller, J. F. Price, L. Qi, O. T. Raitakari, T. Rankinen, D. C. Rao, T. K. Rice, P. M. Ridker, J. D. Rioux, M. D. Ritchie, I. Rudan, V. Salomaa, N. J. Samani, J. Saramies, M. A. Sarzynski, H. Schunkert, P. E. Schwarz, P. Sever, A. R. Shuldiner, J. Sinisalo, R. P. Stolk, K. Strauch, A. Tonjes, D. A. Tregouet, A. Tremblay, E. Tremoli, J. Virtamo, M. C. Vohl, U. Volker, G. Waeber, G. Willemsen, J. C. Witteman, M. C. Zillikens, L. S. Adair, P. Amouyel, F. W. Asselbergs, T. L. Assimes, M. Bochud, B. O. Boehm, E. Boerwinkle, S. R. Bornstein, E. P. Bottinger, C. Bouchard, S. Cauchi, J. C. Chambers, S. J. Chanock, R. S. Cooper, P. I. de Bakker, G. Dedoussis, L. Ferrucci, P. W. Franks, P. Froguel, L. C. Groop, C. A. Haiman, A. Hamsten, J. Hui, D. J. Hunter, K. Hveem, R. C. Kaplan, M. Kivimaki, D. Kuh, M. Laakso, Y. Liu, N. G. Martin, W. Marz, M. Melbye, A. Metspalu, S. Moebus, P. B. Munroe, I. Njolstad, B. A. Oostra, C. N. Palmer, N. L. Pedersen, M. Perola, L. Perusse, U. Peters, C. Power, T. Quertermous, R. Rauramaa, F. Rivadeneira, T. E. Saaristo, D. Saleheen, N. Sattar, E. E. Schadt, D. Schlessinger, P. E. Slagboom, H. Snieder, T. D. Spector, U. Thorsteinsdottir, M. Stumvoll, J. Tuomilehto, A. G. Uitterlinden, M. Uusitupa, P. van der Harst, M. Walker, H. Wallaschofski, N. J. Wareham, H. Watkins, D. R. Weir, H. E. Wichmann, J. F. Wilson, P. Zanen, I. B. Borecki, P. Deloukas, C. S. Fox, I. M. Heid, J. R. O'Connell, D. P. Strachan, K. Stefansson, C. M. van Duijn, G. R. Abecasis, L. Franke, T. M. Frayling, M. I. McCarthy, P. M. Visscher, A. Scherag, C. J. Willer, M. Boehnke, K. L. Mohlke, C. M. Lindgren, J. S. Beckmann, I. Barroso, K. E. North, E. Ingelsson, J. N.

Hirschhorn, R. J. Loos, and E. K. Speliotes. 2015. Genetic studies of body mass index yield new insights for obesity biology. Nature 518:197-206.

131. Shungin, D., T. W. Winkler, D. C. Croteau-Chonka, T. Ferreira, A. E. Locke, R. Magi, R. J. Strawbridge, T. H. Pers, K. Fischer, A. E. Justice, T. Workalemahu, J. M. Wu, M. L. Buchkovich, N. L. Heard-Costa, T. S. Roman, A. W. Drong, C. Song, S. Gustafsson, F. R. Day, T. Esko, T. Fall, Z. Kutalik, J. Luan, J. C. Randall, A. Scherag, S. Vedantam, A. R. Wood, J. Chen, R. Fehrmann, J. Karjalainen, B. Kahali, C. T. Liu, E. M. Schmidt, D. Absher, N. Amin, D. Anderson, M. Beekman, J. L. Bragg-Gresham, S. Buyske, A. Demirkan, G. B. Ehret, M. F. Feitosa, A. Goel, A. U. Jackson, T. Johnson, M. E. Kleber, K. Kristiansson, M. Mangino, I. Mateo Leach, C. Medina-Gomez, C. D. Palmer, D. Pasko, S. Pechlivanis, M. J. Peters, I. Prokopenko, A. Stancakova, Y. Ju Sung, T. Tanaka, A. Teumer, J. V. Van Vliet-Ostaptchouk, L. Yengo, W. Zhang, E. Albrecht, J. Arnlov, G. M. Arscott, S. Bandinelli, A. Barrett, C. Bellis, A. J. Bennett, C. Berne, M. Bluher, S. Bohringer, F. Bonnet, Y. Bottcher, M. Bruinenberg, D. B. Carba, I. H. Caspersen, R. Clarke, E. W. Daw, J. Deelen, E. Deelman, G. Delgado, A. S. Doney, N. Eklund, M. R. Erdos, K. Estrada, E. Eury, N. Friedrich, M. E. Garcia, V. Giedraitis, B. Gigante, A. S. Go, A. Golay, H. Grallert, T. B. Grammer, J. Grassler, J. Grewal, C. J. Groves, T. Haller, G. Hallmans, C. A. Hartman, M. Hassinen, C. Hayward, K. Heikkila, K. H. Herzig, Q. Helmer, H. L. Hillege, O. Holmen, S. C. Hunt, A. Isaacs, T. Ittermann, A. L. James, I. Johansson, T. Juliusdottir, I. P. Kalafati, L. Kinnunen, W. Koenig, I. K. Kooner, W. Kratzer, C. Lamina, K. Leander, N. R. Lee, P. Lichtner, L. Lind, J. Lindstrom, S. Lobbens, M. Lorentzon, F. Mach, P. K.

Magnusson, A. Mahajan, W. L. McArdle, C. Menni, S. Merger, E. Mihailov, L. Milani, R. Mills, A. Moayyeri, K. L. Monda, S. P. Mooijaart, T. W. Muhleisen, A. Mulas, G. Muller, M. Muller-Nurasyid, R. Nagaraja, M. A. Nalls, N. Narisu, N. Glorioso, I. M. Nolte, M. Olden, N. W. Rayner, F. Renstrom, J. S. Ried, N. R. Robertson, L. M. Rose, S. Sanna, H. Scharnagl, S. Scholtens, B. Sennblad, T. Seufferlein, C. M. Sitlani, A. Vernon Smith, K. Stirrups, H. M. Stringham, J. Sundstrom, M. A. Swertz, A. J. Swift, A. C. Syvanen, B. O. Tayo, B. Thorand, G. Thorleifsson, A. Tomaschitz, C. Troffa, F. V. van Oort, N. Verweij, J. M. Vonk, L. L. Waite, R. Wennauer, T. Wilsgaard, M. K. Wojczynski, A. Wong, Q. Zhang, J. Hua Zhao, E. P. Brennan, M. Choi, P. Eriksson, L. Folkersen, A. Franco-Cereceda, A. G. Gharavi, A. K. Hedman, M. F. Hivert, J. Huang, S. Kanoni, F. Karpe, S. Keildson, K. Kiryluk, L. Liang, R. P. Lifton, B. Ma, A. J. McKnight, R. McPherson, A. Metspalu, J. L. Min, M. F. Moffatt, G. W. Montgomery, J. M. Murabito, G. Nicholson, D. R. Nyholt, C. Olsson, J. R. Perry, E. Reinmaa, R. M. Salem, N. Sandholm, E. E. Schadt, R. A. Scott, L. Stolk, E. E. Vallejo, H. J. Westra, K. T. Zondervan, A. D. Consortium, C. A. D. Consortium, C. K. Consortium, G. Consortium, G. Consortium, Glgc, Icbp, C. International Endogene, S. LifeLines Cohort, M. Investigators, T. C. Mu, P. Consortium, C. ReproGen, P. Amouyel, D. Arveiler, S. J. Bakker, J. Beilby, R. N. Bergman, J. Blangero, M. J. Brown, M. Burnier, H. Campbell, A. Chakravarti, P. S. Chines, S. Claudi-Boehm, F. S. Collins, D. C. Crawford, J. Danesh, U. de Faire, E. J. de Geus, M. Dorr, R. Erbel, J. G. Eriksson, M. Farrall, E. Ferrannini, J. Ferrieres, N. G. Forouhi, T. Forrester, O. H. Franco, R. T. Gansevoort, C. Gieger, V.

Gudnason, C. A. Haiman, T. B. Harris, A. T. Hattersley, M. Heliovaara, A. A. Hicks, A. D. Hingorani, W. Hoffmann, A. Hofman, G. Homuth, S. E. Humphries, E. Hypponen, T. Illig, M. R. Jarvelin, B. Johansen, P. Jousilahti, A. M. Jula, J. Kaprio, F. Kee, S. M. Keinanen-Kiukaanniemi, J. S. Kooner, C. Kooperberg, P. Kovacs, A. T. Kraja, M. Kumari, K. Kuulasmaa, J. Kuusisto, T. A. Lakka, C. Langenberg, L. Le Marchand, T. Lehtimaki, V. Lyssenko, S. Mannisto, A. Marette, T. C. Matise, C. A. McKenzie, B. McKnight, A. W. Musk, S. Mohlenkamp, A. D. Morris, M. Nelis, C. Ohlsson, A. J. Oldehinkel, K. K. Ong, L. J. Palmer, B. W. Penninx, A. Peters, P. P. Pramstaller, O. T. Raitakari, T. Rankinen, D. C. Rao, T. K. Rice, P. M. Ridker, M. D. Ritchie, I. Rudan, V. Salomaa, N. J. Samani, J. Saramies, M. A. Sarzynski, P. E. Schwarz, A. R. Shuldiner, J. A. Staessen, V. Steinthorsdottir, R. P. Stolk, K. Strauch, A. Tonjes, A. Tremblay, E. Tremoli, M. C. Vohl, U. Volker, P. Vollenweider, J. F. Wilson, J. C. Witteman, L. S. Adair, M. Bochud, B. O. Boehm, S. R. Bornstein, C. Bouchard, S. Cauchi, M. J. Caulfield, J. C. Chambers, D. I. Chasman, R. S. Cooper, G. Dedoussis, L. Ferrucci, P. Froguel, H. J. Grabe, A. Hamsten, J. Hui, K. Hveem, K. H. Jockel, M. Kivimaki, D. Kuh, M. Laakso, Y. Liu, W. Marz, P. B. Munroe, I. Njolstad, B. A. Oostra, C. N. Palmer, N. L. Pedersen, M. Perola, L. Perusse, U. Peters, C. Power, T. Quertermous, R. Rauramaa, F. Rivadeneira, T. E. Saaristo, D. Saleheen, J. Sinisalo, P. E. Slagboom, H. Snieder, T. D. Spector, U. Thorsteinsdottir, M. Stumvoll, J. Tuomilehto, A. G. Uitterlinden, M. Uusitupa, P. van der Harst, G. Veronesi, M. Walker, N. J. Wareham, H. Watkins, H. E. Wichmann, G. R. Abecasis, T. L. Assimes, S. I. Berndt, M.

Boehnke, I. B. Borecki, P. Deloukas, L. Franke, T. M. Frayling, L. C. Groop, D. J. Hunter, R. C. Kaplan, J. R. O'Connell, L. Qi, D. Schlessinger, D. P. Strachan, K. Stefansson, C. M. van Duijn, C. J. Willer, P. M. Visscher, J. Yang, J. N. Hirschhorn, M. C. Zillikens, M. I. McCarthy, E. K. Speliotes, K. E. North, C. S. Fox, I. Barroso, P. W. Franks, E. Ingelsson, I. M. Heid, R. J. Loos, L. A. Cupples, A. P. Morris, C. M. Lindgren, and K. L. Mohlke. 2015. New genetic loci link adipose and insulin biology to body fat distribution. Nature 518:187-196.

132. Church, C., L. Moir, F. McMurray, C. Girard, G. T. Banks, L. Teboul, S. Wells, J. C. Bruning, P. M. Nolan, F. M. Ashcroft, and R. D. Cox. 2010. Overexpression of Fto leads to increased food intake and results in obesity. Nature genetics 42:1086-1092.

133. Fischer, J., L. Koch, C. Emmerling, J. Vierkotten, T. Peters, J. C. Bruning, and U. Ruther. 2009. Inactivation of the Fto gene protects from obesity. Nature 458:894-898.

134. Smemo, S., J. J. Tena, K. H. Kim, E. R. Gamazon, N. J. Sakabe, C. Gomez-Marin, I. Aneas, F. L. Credidio, D. R. Sobreira, N. F. Wasserman, J. H. Lee, V. Puviindran, D. Tam, M. Shen, J. E. Son, N. A. Vakili, H. K. Sung, S. Naranjo, R. D. Acemel, M. Manzanares, A. Nagy, N. J. Cox, C. C. Hui, J. L. Gomez-Skarmeta, and M. A. Nobrega. 2014. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature 507:371-375.

135. Kin, T., and Y. Ono. 2007. Idiographica: a general-purpose web application to build idiograms on-demand for human, mouse and rat. Bioinformatics 23:2945-2946.

136. QuickFacts., U. S. C. B. S. a. C. 27-Mar-2014. Data derived from Population Estimates, American Community Survey, Census of Population and Housing, State and County Housing Unit Estimates, County Business Patterns, Nonemployer Statistics, Economic Census, Survey of Business Owners, Building Permits

137. Genomes Project, C., G. R. Abecasis, A. Auton, L. D. Brooks, M. A. DePristo, R. M. Durbin, R. E. Handsaker, H. M. Kang, G. T. Marth, and G. A. McVean. 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491:56-65.

138. Alexander, D. H., J. Novembre, and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome research 19:1655-1664.

139. Tang, H., E. Jorgenson, M. Gadde, S. L. Kardia, D. C. Rao, X. Zhu, N. J. Schork, C. L. Hanis, and N. Risch. 2006. Racial admixture and its impact on BMI and blood pressure in African and Mexican Americans. Human genetics 119:624-633.

140. Hanis, C. L., D. Hewett-Emmett, T. K. Bertin, and W. J. Schull. 1991. Origins of U.S. Hispanics. Implications for diabetes. Diabetes care 14:618-627.

141. Hanis, C. L., R. E. Ferrell, S. A. Barton, L. Aguilar, A. Garza-Ibarra, B. R. Tulloch, C. A. Garcia, and W. J. Schull. 1983. Diabetes among Mexican Americans in Starr County, Texas. American journal of epidemiology 118:659-672.

142. Hanis, C. L., R. E. Ferrell, B. R. Tulloch, and W. J. Schull. 1985. Gallbladder disease epidemiology in Mexican Americans in Starr County, Texas. American journal of epidemiology 122:820-829.

143. Hanis, C. L., R. E. Ferrell, and W. J. Schull. 1985. Hypertension and sources of blood pressure variability among Mexican-Americans in Starr County, Texas. International journal of epidemiology 14:231-238.

144. Brown, S. A., and C. L. Hanis. 1995. A community-based, culturally sensitive education and group-support intervention for Mexican Americans with NIDDM: a pilot study of efficacy. The Diabetes educator 21:203-210.

145. Brown, S. A., and C. L. Hanis. 1999. Culturally competent diabetes education for Mexican Americans: the Starr County Study. The Diabetes educator 25:226-236.

146. Brown, S. A., S. L. Upchurch, A. A. Garcia, S. A. Barton, and C. L. Hanis. 1998. Symptom-related self-care of Mexican Americans with type 2 diabetes: preliminary findings of the Starr County Diabetes Education Study. The Diabetes educator 24:331-339.

147. Hanis, C. L., H. H. Chu, K. Lawson, D. Hewett-Emmett, S. A. Barton, W. J. Schull, and C. A. Garcia. 1993. Mortality of Mexican Americans with NIDDM. Retinopathy and other predictors in Starr County, Texas. Diabetes care 16:82-89.

148. Below, J. E., E. R. Gamazon, J. V. Morrison, A. Konkashbaev, A. Pluzhnikov, P. M. McKeigue, E. J. Parra, S. C. Elbein, D. M. Hallman, D. L. Nicolae, G. I. Bell, M. Cruz, N. J. Cox, and C. L. Hanis. 2011. Genome-wide association and meta-

analysis in populations from Starr County, Texas, and Mexico City identify type 2 diabetes susceptibility loci and enrichment for expression quantitative trait loci in top signals. Diabetologia 54:2047-2055.

149. Bray, M. S., E. Boerwinkle, and C. L. Hanis. 1996. OB gene not linked to human obesity in Mexican American affected sib pairs from Starr County, Texas. Human genetics 98:590-595.

150. Herbert, A., N. P. Gerry, M. B. McQueen, I. M. Heid, A. Pfeufer, T. Illig, H. E. Wichmann, T. Meitinger, D. Hunter, F. B. Hu, G. Colditz, A. Hinney, J. Hebebrand, K. Koberwitz, X. Zhu, R. Cooper, K. Ardlie, H. Lyon, J. N. Hirschhorn, N. M. Laird, M. E. Lenburg, C. Lange, and M. F. Christman. 2006. A common genetic variant is associated with adult and childhood obesity. Science 312:279-283.

151. Bressler, J., M. Fornage, C. L. Hanis, W. H. Kao, C. E. Lewis, R. McPherson, R. Dent, T. H. Mosley, L. A. Pennacchio, and E. Boerwinkle. 2009. The INSIG2 rs7566605 genetic variant does not play a major role in obesity in a sample of 24,722 individuals from four cohorts. BMC medical genetics 10:56.

152. Heid, I. M., C. Huth, R. J. Loos, F. Kronenberg, V. Adamkova, S. S. Anand, K. Ardlie, H. Biebermann, P. Bjerregaard, H. Boeing, C. Bouchard, M. Ciullo, J. A. Cooper, D. Corella, C. Dina, J. C. Engert, E. Fisher, F. Frances, P. Froguel, J. Hebebrand, R. A. Hegele, A. Hinney, M. R. Hoehe, F. B. Hu, J. A. Hubacek, S. E. Humphries, S. C. Hunt, T. Illig, M. R. Jarvelin, M. Kaakinen, B. Kollerits, H. Krude, J. Kumar, L. A. Lange, B. Langer, S. Li, A. Luchner, H. N. Lyon, D. Meyre, K. L. Mohlke, V. Mooser, A. Nebel, T. T. Nguyen, B. Paulweber, L. Perusse, L.

Qi, T. Rankinen, D. Rosskopf, S. Schreiber, S. Sengupta, R. Sorice, A. Suk, G. Thorleifsson, U. Thorsteinsdottir, H. Volzke, K. S. Vimaleswaran, N. J. Wareham, D. Waterworth, S. Yusuf, C. Lindgren, M. I. McCarthy, C. Lange, J. N. Hirschhorn, N. Laird, and H. E. Wichmann. 2009. Meta-analysis of the INSIG2 association with obesity including 74,345 individuals: does heterogeneity of estimates relate to study design? PLoS genetics 5:e1000694.

153. Suchanek, P., I. Kralova Lesna, O. Mengerova, J. Mrazkova, V. Lanska, and P. Stavek. 2012. Which index best correlates with body fat mass: BAI, BMI, waist or WHR? Neuro endocrinology letters 33 Suppl 2:78-82.

154. Flegal, K. M., J. A. Shepherd, A. C. Looker, B. I. Graubard, L. G. Borrud, C. L. Ogden, T. B. Harris, J. E. Everhart, and N. Schenker. 2009. Comparisons of percentage body fat, body mass index, waist circumference, and waist-stature ratio in adults. The American journal of clinical nutrition 89:500-508.

155. Segal, K. R., M. Van Loan, P. I. Fitzgerald, J. A. Hodgdon, and T. B. Van Itallie. 1988. Lean body mass estimation by bioelectrical impedance analysis: a four-site cross-validation study. The American journal of clinical nutrition 47:7-14.

156. Kamel, E. G., G. McNeill, and M. C. Van Wijk. 2000. Usefulness of anthropometry and DXA in predicting intra-abdominal fat in obese men and women. Obesity research 8:36-42.

157. Berndt, S. I., S. Gustafsson, R. Magi, A. Ganna, E. Wheeler, M. F. Feitosa, A. E. Justice, K. L. Monda, D. C. Croteau-Chonka, F. R. Day, T. Esko, T. Fall, T.

Ferreira, D. Gentilini, A. U. Jackson, J. Luan, J. C. Randall, S. Vedantam, C. J. Willer, T. W. Winkler, A. R. Wood, T. Workalemahu, Y. J. Hu, S. H. Lee, L. Liang, D. Y. Lin, J. L. Min, B. M. Neale, G. Thorleifsson, J. Yang, E. Albrecht, N. Amin, J. L. Bragg-Gresham, G. Cadby, M. den Heijer, N. Eklund, K. Fischer, A. Goel, J. J. Hottenga, J. E. Huffman, I. Jarick, A. Johansson, T. Johnson, S. Kanoni, M. E. Kleber, I. R. Konig, K. Kristiansson, Z. Kutalik, C. Lamina, C. Lecoeur, G. Li, M. Mangino, W. L. McArdle, C. Medina-Gomez, M. Muller-Nurasyid, J. S. Ngwa, I. M. Nolte, L. Paternoster, S. Pechlivanis, M. Perola, M. J. Peters, M. Preuss, L. M. Rose, J. Shi, D. Shungin, A. V. Smith, R. J. Strawbridge, I. Surakka, A. Teumer, M. D. Trip, J. Tyrer, J. V. Van Vliet-Ostaptchouk, L. Vandenput, L. L. Waite, J. H. Zhao, D. Absher, F. W. Asselbergs, M. Atalay, A. P. Attwood, A. J. Balmforth, H. Basart, J. Beilby, L. L. Bonnycastle, P. Brambilla, M. Bruinenberg, H. Campbell, D. I. Chasman, P. S. Chines, F. S. Collins, J. M. Connell, W. O. Cookson, U. de Faire, F. de Vegt, M. Dei, M. Dimitriou, S. Edkins, K. Estrada, D. M. Evans, M. Farrall, M. M. Ferrario, J. Ferrieres, L. Franke, F. Frau, P. V. Gejman, H. Grallert, H. Gronberg, V. Gudnason, A. S. Hall, P. Hall, A. L. Hartikainen, C. Hayward, N. L. Heard-Costa, A. C. Heath, J. Hebebrand, G. Homuth, F. B. Hu, S. E. Hunt, E. Hypponen, C. Iribarren, K. B. Jacobs, J. O. Jansson, A. Jula, M. Kahonen, S. Kathiresan, F. Kee, K. T. Khaw, M. Kivimaki, W. Koenig, A. T. Kraja, M. Kumari, K. Kuulasmaa, J. Kuusisto, J. H. Laitinen, T. A. Lakka, C. Langenberg, L. J. Launer, L. Lind, J. Lindstrom, J. Liu, A. Liuzzi, M. L. Lokki, M. Lorentzon, P. A. Madden, P. K. Magnusson, P. Manunta, D. Marek, W. Marz, I. Mateo Leach, B. McKnight, S. E. Medland, E. Mihailov, L. Milani, G.

W. Montgomery, V. Mooser, T. W. Muhleisen, P. B. Munroe, A. W. Musk, N. Narisu, G. Navis, G. Nicholson, E. A. Nohr, K. K. Ong, B. A. Oostra, C. N. Palmer, A. Palotie, J. F. Peden, N. Pedersen, A. Peters, O. Polasek, A. Pouta, P. P. Pramstaller, I. Prokopenko, C. Putter, A. Radhakrishnan, O. Raitakari, A. Rendon, F. Rivadeneira, I. Rudan, T. E. Saaristo, J. G. Sambrook, A. R. Sanders, S. Sanna, J. Saramies, S. Schipf, S. Schreiber, H. Schunkert, S. Y. Shin, S. Signorini, J. Sinisalo, B. Skrobek, N. Soranzo, A. Stancakova, K. Stark, J. C. Stephens, K. Stirrups, R. P. Stolk, M. Stumvoll, A. J. Swift, E. V. Theodoraki, B. Thorand, D. A. Tregouet, E. Tremoli, M. M. Van der Klauw, J. B. van Meurs, S. H. Vermeulen, J. Viikari, J. Virtamo, V. Vitart, G. Waeber, Z. Wang, E. Widen, S. H. Wild, G. Willemsen, B. R. Winkelmann, J. C. Witteman, B. H. Wolffenbuttel, A. Wong, A. F. Wright, M. C. Zillikens, P. Amouyel, B. O. Boehm, E. Boerwinkle, D. I. Boomsma, M. J. Caulfield, S. J. Chanock, L. A. Cupples, D. Cusi, G. V. Dedoussis, J. Erdmann, J. G. Eriksson, P. W. Franks, P. Froguel, C. Gieger, U. Gyllensten, A. Hamsten, T. B. Harris, C. Hengstenberg, A. A. Hicks, A. Hingorani, A. Hinney, A. Hofman, K. G. Hovingh, K. Hveem, T. Illig, M. R. Jarvelin, K. H. Jockel, S. M. Keinanen-Kiukaanniemi, L. A. Kiemeney, D. Kuh, M. Laakso, T. Lehtimaki, D. F. Levinson, N. G. Martin, A. Metspalu, A. D. Morris, M. S. Nieminen, I. Njolstad, C. Ohlsson, A. J. Oldehinkel, W. H. Ouwehand, L. J. Palmer, B. Penninx, C. Power, M. A. Province, B. M. Psaty, L. Qi, R. Rauramaa, P. M. Ridker, S. Ripatti, V. Salomaa, N. J. Samani, H. Snieder, T. I. Sorensen, T. D. Spector, K. Stefansson, A. Tonjes, J. Tuomilehto, A. G. Uitterlinden, M. Uusitupa, P. van der Harst, P. Vollenweider, H.

Wallaschofski, N. J. Wareham, H. Watkins, H. E. Wichmann, J. F. Wilson, G. R. Abecasis, T. L. Assimes, I. Barroso, M. Boehnke, I. B. Borecki, P. Deloukas, C. S. Fox, T. Frayling, L. C. Groop, T. Haritunian, I. M. Heid, D. Hunter, R. C. Kaplan, F. Karpe, M. F. Moffatt, K. L. Mohlke, J. R. O'Connell, Y. Pawitan, E. E. Schadt, D. Schlessinger, V. Steinthorsdottir, D. P. Strachan, U. Thorsteinsdottir, C. M. van Duijn, P. M. Visscher, A. M. Di Blasio, J. N. Hirschhorn, C. M. Lindgren, A. P. Morris, D. Meyre, A. Scherag, M. I. McCarthy, E. K. Speliotes, K. E. North, R. J. Loos, and E. Ingelsson. 2013. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. Nature genetics 45:501-512.

158. Aydogdu, A., I. Tasci, O. Kucukerdonmez, S. Tapan, S. Aydogdu, U. Aydogan, A. Sonmez, M. Yazici, and O. Azal. 2013. Increase in subcutaneous adipose tissue and fat free mass in women with polycystic ovary syndrome is related to impaired insulin sensitivity. Gynecological endocrinology : the official journal of the International Society of Gynecological Endocrinology 29:152-155.

159. Mueller, W. H., M. L. Wear, C. L. Hanis, S. A. Barton, and W. J. Schull. 1987. Body circumferences as alternatives to skinfold measurements of body fat distribution in Mexican-Americans. Int J Obes 11:309-318.

160. Livshits, G., K. Yakovenko, E. Ginsburg, and E. Kobyliansky. 1998. Genetics of human body size and shape: pleiotropic and independent genetic determinants of adiposity. Annals of human biology 25:221-236.

161. Li, Z., T. Rice, L. Pérusse, C. Bouchard, and D. C. Rao. 1996. Familial aggregation of subcutaneous fat patterning: Principal components of skinfolds in the Québec family study. American Journal of Human Biology 8:535-542.

162. Tayo, B. O., R. Harders, A. Luke, X. Zhu, and R. S. Cooper. 2008. Latent common genetic components of obesity traits. International journal of obesity 32:1799-1806.

163. Comuzzie, A. G., J. Blangero, M. C. Mahaney, B. D. Mitchell, M. P. Stern, and J. W. MacCluer. 1994. Genetic and environmental correlations among skinfold measures. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity 18:413-418.

164. Liu, J. Z., A. F. McRae, D. R. Nyholt, S. E. Medland, N. R. Wray, K. M. Brown, A. Investigators, N. K. Hayward, G. W. Montgomery, P. M. Visscher, N. G. Martin, and S. Macgregor. 2010. A versatile gene-based test for genome-wide association studies. American journal of human genetics 87:139-145.

165. Korn, J. M., F. G. Kuruvilla, S. A. McCarroll, A. Wysoker, J. Nemesh, S. Cawley, E. Hubbell, J. Veitch, P. J. Collins, K. Darvishi, C. Lee, M. M. Nizzari, S. B. Gabriel, S. Purcell, M. J. Daly, and D. Altshuler. 2008. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nature genetics 40:1253-1260.

166. Carvalho, B., H. Bengtsson, T. P. Speed, and R. A. Irizarry. 2007. Exploration, normalization, and genotype calls of high-density oligonucleotide SNP array data. Biostatistics 8:485-499.

167. Pluzhnikov, A., J. E. Below, A. Konkashbaev, A. Tikhomirov, E. Kistner-Griffin, C. A. Roe, D. L. Nicolae, and N. J. Cox. 2010. Spoiling the whole bunch: quality control aimed at preserving the integrity of high-throughput genotyping. American journal of human genetics 87:123-128.

168. Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics 81:559-575.

169. Delaneau, O., J. F. Zagury, and J. Marchini. 2013. Improved whole-chromosome phasing for disease and population genetic studies. Nature methods 10:5-6.

170. Howie, B., J. Marchini, and M. Stephens. 2011. Genotype imputation with thousands of genomes. G3 1:457-470.

171. Howie, B. N., P. Donnelly, and J. Marchini. 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS genetics 5:e1000529.

172. Price, A. L., N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, and D. Reich. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nature genetics 38:904-909.

173. Marchini, J., and B. Howie. 2010. Genotype imputation for genome-wide association studies. Nature reviews. Genetics 11:499-511.

174. Pe'er, I., R. Yelensky, D. Altshuler, and M. J. Daly. 2008. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genetic epidemiology 32:381-385.

175. Bolormaa, S., J. E. Pryce, A. Reverter, Y. Zhang, W. Barendse, K. Kemper, B. Tier, K. Savin, B. J. Hayes, and M. E. Goddard. 2014. A multi-trait, meta-analysis for detecting pleiotropic polymorphisms for stature, fatness and reproduction in beef cattle. PLoS genetics 10:e1004198.

176. Zhu, X., T. Feng, B. O. Tayo, J. Liang, J. H. Young, N. Franceschini, J. A. Smith, L. R. Yanek, Y. V. Sun, T. L. Edwards, W. Chen, M. Nalls, E. Fox, M. Sale, E. Bottinger, C. Rotimi, C. B. P. C. The, Y. Liu, B. McKnight, K. Liu, D. K. Arnett, A. Chakravati, R. S. Cooper, and S. Redline. 2014. Meta-analysis of Correlated Traits via Summary Statistics from GWASs with an Application in Hypertension. American journal of human genetics.

177. Chanda, P., H. Huang, D. E. Arking, and J. S. Bader. 2013. Fast association tests for genes with FAST. PloS one 8:e68585.

178. Escobar, C. M., S. J. Krajewski, T. Sandoval-Guzman, M. L. Voytko, and N. E. Rance. 2004. Neuropeptide Y gene expression is increased in the hypothalamus of older women. The Journal of clinical endocrinology and metabolism 89:2338-2343.

179. Ward, C. J., M. C. Hogan, S. Rossetti, D. Walker, T. Sneddon, X. Wang, V. Kubly, J. M. Cunningham, R. Bacallao, M. Ishibashi, D. S. Milliner, V. E. Torres, and P. C. Harris. 2002. The gene mutated in autosomal recessive polycystic kidney disease encodes a large, receptor-like protein. Nature genetics 30:259-269.

180. Rinella, E. S., C. Still, Y. Shao, G. C. Wood, X. Chu, B. Salerno, G. S. Gerhard, and H. Ostrer. 2013. Genome-wide association of single-nucleotide polymorphisms with weight loss outcomes after Roux-en-Y gastric bypass surgery. The Journal of clinical endocrinology and metabolism 98:E1131-1136.

181. Wade, T. D., S. Gordon, S. Medland, C. M. Bulik, A. C. Heath, G. W. Montgomery, and N. G. Martin. 2013. Genetic variants associated with disordered eating. The International journal of eating disorders 46:594-608.

182. Maas, N. M., T. Van de Putte, C. Melotte, A. Francis, C. T. Schrander-Stumpel, D. Sanlaville, D. Genevieve, S. Lyonnet, B. Dimitrov, K. Devriendt, J. P. Fryns, and J. R. Vermeesch. 2007. The C20orf133 gene is disrupted in a patient with Kabuki syndrome. Journal of medical genetics 44:562-569.

183. Whiting, P. J., T. P. Bonnert, R. M. McKernan, S. Farrar, B. Le Bourdelles, R. P. Heavens, D. W. Smith, L. Hewson, M. R. Rigby, D. J. Sirinathsinghji, S. A. Thompson, and K. A. Wafford. 1999. Molecular and functional diversity of the expanding GABA-A receptor gene family. Annals of the New York Academy of Sciences 868:645-653.

184. Derry, J. M., S. M. Dunn, and M. Davies. 2004. Identification of a residue in the gamma-aminobutyric acid type A receptor alpha subunit that differentially affects diazepam-sensitive and -insensitive benzodiazepine site binding. Journal of neurochemistry 88:1431-1438.

185. Ziegler, D. R., and J. P. Herman. 2002. Neurocircuitry of stress integration: anatomical pathways regulating the hypothalamo-pituitary-adrenocortical axis of the rat. Integrative and comparative biology 42:541-551.

186. Bose, M., B. Olivan, and B. Laferrere. 2009. Stress and obesity: the role of the hypothalamic-pituitary-adrenal axis in metabolic disease. Current opinion in endocrinology, diabetes, and obesity 16:340-346.

187. Crowley, S. K., and S. S. Girdler. 2014. Neurosteroid, GABAergic and hypothalamic pituitary adrenal (HPA) axis regulation: what is the current state of knowledge in humans? Psychopharmacology 231:3619-3634.

188. Comuzzie, A. G., S. A. Cole, S. L. Laston, V. S. Voruganti, K. Haack, R. A. Gibbs, and N. F. Butte. 2012. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PloS one 7:e51954.

189. Al Safar, H. S., H. J. Cordell, O. Jafer, D. Anderson, S. E. Jamieson, M. Fakiola, K. Khazanehdari, G. K. Tay, and J. M. Blackwell. 2013. A genome-wide search for type 2 diabetes susceptibility genes in an extended Arab family. Annals of human genetics 77:488-503.

190. Reue, K., P. Xu, X. P. Wang, and B. G. Slavin. 2000. Adipose tissue deficiency, glucose intolerance, and increased atherosclerosis result from mutation in the mouse fatty liver dystrophy (fld) gene. Journal of lipid research 41:1067-1076.

191. Fawcett, K. A., N. Grimsey, R. J. Loos, E. Wheeler, A. Daly, M. Soos, R. Semple, H. Syddall, C. Cooper, S. Siniossoglou, S. O'Rahilly, N. J. Wareham, and I. Barroso. 2008. Evaluating the role of LPIN1 variation in insulin resistance,

body weight, and human lipodystrophy in U.K. Populations. Diabetes 57:2527-2533.

192. Tennessen, J. A., A. W. Bigham, T. D. O'Connor, W. Fu, E. E. Kenny, S. Gravel, S. McGee, R. Do, X. Liu, G. Jun, H. M. Kang, D. Jordan, S. M. Leal, S. Gabriel, M. J. Rieder, G. Abecasis, D. Altshuler, D. A. Nickerson, E. Boerwinkle, S. Sunyaev, C. D. Bustamante, M. J. Bamshad, J. M. Akey, G. O. Broad, G. O. Seattle, and N. E. S. Project. 2012. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science 337:64-69.

193. DePristo, M. A., E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics 43:491-498.

194. Kang, H. M., J. H. Sul, S. K. Service, N. A. Zaitlen, S. Y. Kong, N. B. Freimer, C. Sabatti, and E. Eskin. 2010. Variance component model to account for sample structure in genome-wide association studies. Nature genetics 42:348-354.

195. Cingolani, P., A. Platts, L. Wang le, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 6:80-92.

196. Liu, X., X. Jian, and E. Boerwinkle. 2013. dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. Human mutation 34:E2393-2402.

197. Liu, X., X. Jian, and E. Boerwinkle. 2011. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. Human mutation 32:894-899.

198. Dong, C., P. Wei, X. Jian, R. Gibbs, E. Boerwinkle, K. Wang, and X. Liu. 2014. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Human molecular genetics.

199. Lee, S., M. J. Emond, M. J. Bamshad, K. C. Barnes, M. J. Rieder, D. A. Nickerson, N. G. E. S. P.-E. L. P. Team, D. C. Christiani, M. M. Wurfel, and X. Lin. 2012. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. American journal of human genetics 91:224-237.

200. Fox, C. S., Y. Liu, C. C. White, M. Feitosa, A. V. Smith, N. Heard-Costa, K. Lohman, G. Consortium, M. Consortium, G. Consortium, A. D. Johnson, M. C. Foster, D. M. Greenawalt, P. Griffin, J. Ding, A. B. Newman, F. Tylavsky, I. Miljkovic, S. B. Kritchevsky, L. Launer, M. Garcia, G. Eiriksdottir, J. J. Carr, V. Gudnason, T. B. Harris, L. A. Cupples, and I. B. Borecki. 2012. Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. PLoS genetics 8:e1002695.

201. Kanemura, Y., S. Hiraga, N. Arita, T. Ohnishi, S. Izumoto, K. Mori, H. Matsumura, M. Yamasaki, S. Fushiki, and T. Yoshimine. 1999. Isolation and expression analysis of a novel human homologue of the Drosophila glial cells missing (gcm) gene. FEBS letters 442:151-156.

202. Ding, C., B. Buckingham, and M. A. Levine. 2001. Familial isolated hypoparathyroidism caused by a mutation in the gene for the transcription factor GCMB. The Journal of clinical investigation 108:1215-1220.

203. Gilbert, S. L., L. Zhang, M. L. Forster, J. R. Anderson, T. Iwase, B. Soliven, L. R. Donahue, H. O. Sweet, R. T. Bronson, M. T. Davisson, R. L. Wollmann, and B. T. Lahn. 2006. Trak1 mutation disrupts GABA(A) receptor homeostasis in hypertonic mice. Nature genetics 38:245-250.

204. Brickley, K., M. J. Smith, M. Beck, and F. A. Stephenson. 2005. GRIF-1 and OIP106, members of a novel gene family of coiled-coil domain proteins: association in vivo and in vitro with kinesin. The Journal of biological chemistry 280:14723-14732.

205. Zhou, C., and B. Blumberg. 2003. Overlapping gene structure of human VLCAD and DLG4. Gene 305:161-166.

206. Suzuki, H., Y. Kawarabayasi, J. Kondo, T. Abe, K. Nishikawa, S. Kimura, T. Hashimoto, and T. Yamamoto. 1990. Structure and regulation of rat long-chain acyl-CoA synthetase. The Journal of biological chemistry 265:8681-8685.

207. Ellis, J. M., L. O. Li, P. C. Wu, T. R. Koves, O. Ilkayeva, R. D. Stevens, S. M. Watkins, D. M. Muoio, and R. A. Coleman. 2010. Adipose acyl-CoA

synthetase-1 directs fatty acids toward beta-oxidation and is required for cold thermogenesis. Cell metabolism 12:53-64.

208. Pruim, R. J., R. P. Welch, S. Sanna, T. M. Teslovich, P. S. Chines, T. P. Gliedt, M. Boehnke, G. R. Abecasis, and C. J. Willer. 2010. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26:2336-2337.

209. GTEx Consortium. 2013. The Genotype-Tissue Expression (GTEx) project. Nature genetics 45:580-585.

210. Koch, M., F. Laub, P. Zhou, R. A. Hahn, S. Tanaka, R. E. Burgeson, D. R. Gerecke, F. Ramirez, and M. K. Gordon. 2003. Collagen XXIV, a vertebrate fibrillar collagen with structural features of invertebrate collagens: selective expression in developing cornea and bone. The Journal of biological chemistry 278:43236-43244.

211. Shahin, H., T. Walsh, T. Sobe, J. Abu Sa'ed, A. Abu Rayan, E. D. Lynch, M. K. Lee, K. B. Avraham, M. C. King, and M. Kanaan. 2006. Mutations in a novel isoform of TRIOBP that encodes a filamentous-actin binding protein are responsible for DFNB28 recessive nonsyndromic hearing loss. American journal of human genetics 78:144-152.

212. Riazuddin, S., S. N. Khan, Z. M. Ahmed, M. Ghosh, K. Caution, S. Nazli, M. Kabra, A. U. Zafar, K. Chen, S. Naz, A. Antonellis, W. J. Pavan, E. D. Green, E. R. Wilcox, P. L. Friedman, R. J. Morell, S. Riazuddin, and T. B. Friedman. 2006. Mutations in TRIOBP, which encodes a putative cytoskeletal-organizing

protein, are associated with nonsyndromic recessive deafness. American journal of human genetics 78:137-143.

213. Kitajiri, S., T. Sakamoto, I. A. Belyantseva, R. J. Goodyear, R. Stepanyan, I. Fujiwara, J. E. Bird, S. Riazuddin, S. Riazuddin, Z. M. Ahmed, J. E. Hinshaw, J. Sellers, J. R. Bartles, J. A. Hammer, 3rd, G. P. Richardson, A. J. Griffith, G. I. Frolenkov, and T. B. Friedman. 2010. Actin-bundling protein TRIOBP forms resilient rootlets of hair cell stereocilia essential for hearing. Cell 141:786-798.

214. Seipel, K., S. P. O'Brien, E. Iannotti, Q. G. Medley, and M. Streuli. 2001. Tara, a novel F-actin binding protein, associates with the Trio guanine nucleotide exchange factor and regulates actin cytoskeletal organization. Journal of cell science 114:389-399.

215. Funke-Kaiser, H., F. Reichenberger, K. Kopke, S. M. Herrmann, J. Pfeifer, H. D. Orzechowski, W. Zidek, M. Paul, and E. Brand. 2003. Differential binding of transcription factor E2F-2 to the endothelin-converting enzyme-1b promoter affects blood pressure regulation. Human molecular genetics 12:423-433.

216. Hofstra, R. M., O. Valdenaire, E. Arch, J. Osinga, H. Kroes, B. M. Loffler, A. Hamosh, C. Meijers, and C. H. Buys. 1999. A loss-of-function mutation in the endothelin-converting enzyme 1 (ECE-1) associated with Hirschsprung disease, cardiac defects, and autonomic dysfunction. American journal of human genetics 64:304-308.

217. Yanagisawa, H., M. Yanagisawa, R. P. Kapur, J. A. Richardson, S. C. Williams, D. E. Clouthier, D. de Wit, N. Emoto, and R. E. Hammer. 1998. Dual genetic pathways of endothelin-mediated intercellular signaling revealed by targeted disruption of endothelin converting enzyme-1 gene. Development 125:825-836.

218. Wallace, R. H., C. Marini, S. Petrou, L. A. Harkin, D. N. Bowser, R. G. Panchal, D. A. Williams, G. R. Sutherland, J. C. Mulley, I. E. Scheffer, and S. F. Berkovic. 2001. Mutant GABA(A) receptor gamma2-subunit in childhood absence epilepsy and febrile seizures. Nature genetics 28:49-52.

219. Sancar, F., and C. Czajkowski. 2004. A GABAA receptor mutation linked to human epilepsy (gamma2R43Q) impairs cell surface expression of alphabetagamma receptors. The Journal of biological chemistry 279:47034-47039.

220. Edenberg, H. J., D. M. Dick, X. Xuei, H. Tian, L. Almasy, L. O. Bauer, R. R. Crowe, A. Goate, V. Hesselbrock, K. Jones, J. Kwon, T. K. Li, J. I. Nurnberger, Jr., S. J. O'Connor, T. Reich, J. Rice, M. A. Schuckit, B. Porjesz, T. Foroud, and H. Begleiter. 2004. Variations in GABRA2, encoding the alpha 2 subunit of the GABA(A) receptor, are associated with alcohol dependence and with brain oscillations. American journal of human genetics 74:705-714.

221. Dibbens, L. M., H. J. Feng, M. C. Richards, L. A. Harkin, B. L. Hodgson, D. Scott, M. Jenkins, S. Petrou, G. R. Sutherland, I. E. Scheffer, S. F. Berkovic, R. L. Macdonald, and J. C. Mulley. 2004. GABRD encoding a protein for extra- or

peri-synaptic GABAA receptors is a susceptibility locus for generalized epilepsies. Human molecular genetics 13:1315-1319.

222. Cossette, P., L. Liu, K. Brisebois, H. Dong, A. Lortie, M. Vanasse, J. M. Saint-Hilaire, L. Carmant, A. Verner, W. Y. Lu, Y. T. Wang, and G. A. Rouleau. 2002. Mutation of GABRA1 in an autosomal dominant form of juvenile myoclonic epilepsy. Nature genetics 31:184-189.

223. Bradley, C. A., C. Taghibiglou, G. L. Collingridge, and Y. T. Wang. 2008. Mechanisms involved in the reduction of GABAA receptor alpha1-subunit expression caused by the epilepsy mutation A322D in the trafficking-competent receptor. The Journal of biological chemistry 283:22043-22050.

224. Baulac, S., G. Huberfeld, I. Gourfinkel-An, G. Mitropoulou, A. Beranger, J. F. Prud'homme, M. Baulac, A. Brice, R. Bruzzone, and E. LeGuern. 2001. First genetic evidence of GABA(A) receptor dysfunction in epilepsy: a mutation in the gamma2-subunit gene. Nature genetics 28:46-48.

225. Nieh, E. H., G. A. Matthews, S. A. Allsop, K. N. Presbrey, C. A. Leppla, R. Wichmann, R. Neve, C. P. Wildes, and K. M. Tye. 2015. Decoding neural circuits that control compulsive sucrose seeking. Cell 160:528-541.

226. Jennings, J. H., R. L. Ung, S. L. Resendez, A. M. Stamatakis, J. G. Taylor, J. Huang, K. Veleta, P. A. Kantak, M. Aita, K. Shilling-Scrivo, C. Ramakrishnan, K. Deisseroth, S. Otte, and G. D. Stuber. 2015. Visualizing hypothalamic network dynamics for appetitive and consummatory behaviors. Cell 160:516-527.

227. Goldstein, C., J. C. Schroeder, J. P. Fortin, J. M. Goss, S. E. Schaus, M. Beinborn, and A. S. Kopin. 2010. Two naturally occurring mutations in the type 1

melanin-concentrating hormone receptor abolish agonist-induced signaling. The Journal of pharmacology and experimental therapeutics 335:799-806.

228. Wermter, A. K., K. Reichwald, T. Buch, F. Geller, C. Platzer, K. Huse, C. Hess, H. Remschmidt, T. Gudermann, G. Preibisch, W. Siegfried, H. P. Goldschmidt, W. D. Li, R. A. Price, H. Biebermann, H. Krude, C. Vollmert, H. E. Wichmann, T. Illig, T. I. Sorensen, A. Astrup, L. H. Larsen, O. Pedersen, D. Eberle, K. Clement, J. Blundell, M. Wabitsch, H. Schafer, M. Platzer, A. Hinney, and J. Hebebrand. 2005. Mutation analysis of the MCHR1 gene in human obesity. European journal of endocrinology / European Federation of Endocrine Societies 152:851-862.

229. 2015. Metreleptin (Myalept): a leptin analog for generalized lipodystrophy. The Medical letter on drugs and therapeutics 57:13-14.

230. Diker-Cohen, T., E. Cochran, P. Gorden, and R. J. Brown. 2015. Partial and Generalized Lipodystrophy: Comparison of Baseline Characteristics and Response to Metreleptin. The Journal of clinical endocrinology and metabolism:jc20144491.

231. Washburn, W. N., M. Manfredi, P. Devasthale, G. Zhao, S. Ahmad, A. Hernandez, J. A. Robl, W. Wang, J. Mignone, Z. Wang, K. Ngu, M. A. Pelleymounter, D. Longhi, R. Zhao, B. Wang, N. Huang, N. Flynn, A. V. Azzara, J. C. Barrish, K. Rohrbach, J. J. Devenny, S. Rooney, M. Thomas, S. Glick, H. E. Godonis, S. J. Harvey, M. J. Cullen, H. Zhang, C. Caporuscio, P. Stetsko, M. Grubb, B. D. Maxwell, H. Yang, A. Apedo, B. Gemzik, E. B. Janovitz, C. Huang, L. Zhang, C. Freeden, and B. J. Murphy. 2014. Identification of a nonbasic

melanin hormone receptor 1 antagonist as an antiobesity clinical candidate. Journal of medicinal chemistry 57:7509-7522.

232. Wu, M., Y. Li, X. Fu, J. Wang, S. Zhang, and L. Yang. 2014. Profiling the interaction mechanism of quinoline/quinazoline derivatives as MCHR1 antagonists: an in silico method. International journal of molecular sciences 15:15475-15502.

233. Johansson, A., and C. Lofberg. 2015. Novel MCH1 receptor antagonists: a patent review. Expert opinion on therapeutic patents 25:193-207.

234. Hamamoto, A., Y. Kobayashi, and Y. Saito. 2015. Identification of amino acids that are selectively involved in Gi/o activation by rat melanin-concentrating hormone receptor 1. Cellular signalling 27:818-827.

Vita

Heather Michelle Highland was born in Billings, Montana on April 14, 1985, the daughter of Glenn Dee Highland and Patricia Kay Highland. She spent a majority of her childhood in Tulsa, Oklahoma. After graduating from Denton High School in Denton, Texas in 2003, she entered Trinity University in San Antonio, Texas. She received her Bachelor of Science with a major in Biology from Trinity in 2007. In the fall of 2007 she entered The University of Texas Graduate School of Biomedical Sciences at Houston.

Permanent Address:

1404 Buena Vista Dr.

Denton, Texas 76210

Heather.Highland@gmail.com