School of Medicine Publications and Presentations

School of Medicine

3-2017

Analysis of whole exome sequencing with cardiometabolic traits using family-based linkage and association in the IRAS Family Study

Keri L. Tabb

Jacklyn N. Hellwege

Nicholette D. Palmer

Latchezar Dimitrov

Satria Sajuthi

See next page for additional authors

Authors

Keri L. Tabb, Jacklyn N. Hellwege, Nicholette D. Palmer, Latchezar Dimitrov, Satria Sajuthi, Kent D. Taylor, Maggie C.Y. NG, Gregory A. Hawkins, Yii-Der Ida Chen, and John Blangero

Analysis of whole exome sequencing with cardiometabolic traits using family-based linkage and association in the IRAS Family Study

Keri L. Tabb[1,2,3], **Jacklyn N. Hellwege**[4,5], **Nicholette D. Palmer**[1,2,3,6], **Latchezar Dimitrov**[2], **Satria Sajuthi**[6,7], **Kent D. Taylor**[8], **Maggie C.Y. NG**[2,3], **Gregory A. Hawkins**[2,4], **Yii-Der Ida Chen**[8], **W. Mark Brown**[6,7], **David McWilliams**[6,7], **Adrienne Williams**[6,7], **Carlos Lorenzo**[9], **Jill M. Norris**[10], **Jirong Long**[4,5], **Jerome I. Rotter**[8], **Joanne E. Curran**[11], **John Blangero**[11], **Lynne E. Wagenknecht**[6,12], **Carl D. Langefeld**[6,7], and **Donald W. Bowden**[1,2,3]

[1]Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC

[2]Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC

[3]Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC

[4]Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN

[5]Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN

[6]Center for Public Health Genomics, Wake Forest School of Medicine, Winston-Salem, NC

[7]Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC

[8]Institute for Translational Genomics and Population Sciences and Department of Pediatrics, Los Angeles BioMedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA

[9]Department of Medicine, University of Texas Health Science Center, San Antonio, TX

[10]Department of Epidemiology, Colorado School of Public Health, University of Colorado Denver, Aurora, CO

[11]South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley School of Medicine, Brownsville, TX

[12]Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC

Summary

Family-based methods are a potentially powerful tool to identify trait-defining genetic variants in extended families, particularly when used to complement conventional association analysis. We utilized two-point linkage analysis and single variant association analysis to evaluate whole exome

Corresponding Author: Donald W. Bowden, Telephone: 336-713-7507, FAX: 336-713-7566, dbowden@wakehealth.edu, Wake Forest School of Medicine, Medical Center Blvd., Winston-Salem, NC 27157, USA.

sequencing (WES) data from 1,205 Hispanic Americans (78 families) from the Insulin Resistance Atherosclerosis Family Study. WES identified 211,612 variants above the minor allele frequency threshold of 0.005. These variants were tested for linkage and/or association with 50 cardiometabolic traits after quality control checks. Two-point linkage analysis yielded 10,580,600 LOD scores with 1,148 LOD scores 3, 183 LOD scores 4, and 29 LOD scores 5. The maximal novel LOD score was 5.50 for rs2289043:T>C, in *UNC5C* with subcutaneous adipose tissue volume. Association analysis identified 13 variants attaining genome-wide significance ($p<5\times10^{-08}$), with the strongest association between rs651821:C>T in *APOA5*, and triglyceride levels ($p=3.67\times10^{-10}$). Overall, there was a 5.2-fold increase in the number of informative variants detected by WES compared to exome chip analysis in this population, nearly 30% of which were novel variants relative to dbSNP build 138. Thus, integration of results from two-point linkage and single-variant association analysis from WES data enabled identification of novel signals potentially contributing to cardiometabolic traits.

Keywords

cohort study; genetic variance; Hispanic; novel variants

Introduction

Despite its success in the study of Mendelian disorders, family-based linkage analysis has shown a limited ability to identify genetic variants underlying complex traits or disease. Genome-wide association studies (GWAS) have largely become the discovery method of choice in the search for variants associated with complex traits or diseases (Ott et al., 2015). However, GWAS approaches have limitations. Notably, the majority of GWAS have been performed in European-derived populations, and thus far, the loci identified by these studies have, in many cases, provided limited information about trait- or disease-associated variants in other ethnicities (Bowden, 2011, Rosenberg et al., 2010). In addition, GWAS requires very large sample sizes to achieve sufficient power and, with few exceptions, primarily identifies common genetic variants which account for a small proportion of the heritability of most complex diseases (Manolio et al., 2009). A major advantage of family-based linkage analysis is its inherent potential to identify high impact variants, especially low frequency (i.e., minor allele frequency [MAF] > 0.005) variants in moderately-sized familial cohorts (Bowden, 2011). Additionally, family-based linkage can be a powerful tool even in moderately sized families for detecting loci near causal variants. The benefits of gene discovery efforts in families are well-known, particularly the ability to limit the number of causative genes and/or pathways involved in disease (Borecki and Province, 2008).

Whole exome sequencing (WES) has recently become a practical approach for identifying coding variants regardless of their frequency within a population (Albrechtsen et al., 2013, Li et al., 2010). When combined with family-based linkage analysis and complemented with association analysis, it can be a powerful, cost-effective method for elucidating variants of biomedical relevance (Gazal et al., 2016). Drawing upon the strengths of both methods, two-point linkage analysis and conventional association analysis have been implemented in parallel to search for coding variants which substantially contribute to the variance within

traits. The ability of this integrated approach to precisely align SNP results for the two analyses is a significant advantage, highlighting variants which may show evidence of both nominal significance and moderate linkage at a given locus, uncovering variants that may not have been identified by either analysis alone.

Previously these approaches have been applied to exome chip data from 130 African American and Hispanic American families comprising the Insulin Resistance Atherosclerosis Family Study (IRASFS) (Hellwege et al., 2014) with a panel of cardiometabolic phenotypes. Here, we have explored the value of WES in 78 Hispanic American families from IRASFS. We hypothesized that the WES dataset would significantly expand the number of coding variations for analysis and identify novel coding variants associated and/or linked to cardiometabolic disease.

Materials and Methods

Samples

The Insulin Resistance Atherosclerosis Family Study (IRASFS) is a family-based study designed to identify genetic and environmental determinants of insulin resistance and visceral adiposity in Hispanic American and African American populations (Henkin et al., 2003). This report involved DNA samples (N=1,221) from 78 families in the Hispanic American cohort from two locations: San Antonio, TX and San Luis Valley, CO. An extensive number of measurements relevant to cardiometabolic disease had previously been collected for these subjects, including those pertaining to glucose homeostasis, blood lipids, anthropometric traits, and fat deposition. All clinical and analysis sites secured IRB approval, and informed consent was obtained from all participants.

Exome sequencing

Exome sequencing was performed using the Illumina Nextera Expanded Exome Enrichment kit in conjunction with an Illumina HiSeq 2500 sequencer. This platform covered 62Mb of exonic sequence and, in comparison to other panels, increased the coverage of untranslated regions (UTRs) and micro RNAs (miRNAs), targeting 20,794 genes. Exome-enriched samples were amplified and the resulting library was loaded onto an Illumina flow cell for cluster generation (standard clonal amplification) using the Illumina TruSeq paired-end cluster kit v2. The flow cell was then transferred to the HiSeq 2500 instrument for parallel sequencing by synthesis using the Illumina TruSeq SBS kit for the HiSeq for 200 cycles. Forty-eight samples per flow cell were sequenced using paired-end reads and multiplexing six samples per lane. Additionally, a PhiX control was spiked into a flow cell lane.

Data processing

All sequence reads were passed through the Illumina Data Analysis Pipeline. Raw intensity files were converted to sequences with preliminary quality scores using intensity and phasing correction, base determination and preliminary quality score estimation. Sequencing image analysis and base calling of sequences into FASTQ files were performed using the instrument's Sequencing Control Software Real Time Analysis (SCS/RTA) software. Participant samples were demultiplexed using CASAVA v1.8. Quality control (QC) metrics

of the unmapped sequence reads were collected using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). All sequence reads from samples passing QC criteria were mapped to the human genome reference sequence (hg19) using the BWA (Li and Durbin, 2009) software which was implemented in parallel on an 8,000 processor computing cluster. BWA-generated alignments in Sequence Alignment/Map (SAM) format were converted to Binary SAM (BAM) format, sorted, and indexed using SAMtools (Li et al., 2009). Alignments from samples of the same subject were merged into a single BAM file to which the base call quality recalibration and indel realignment methods from the Genome Analysis Toolkit (GATK) (DePristo et al., 2011, McKenna et al., 2010) were applied. Coverage and other post-alignment QC metrics were collected using BEDTools (Quinlan and Hall, 2010), SAMtools, and GATK. SNP and INDEL variant discovery were carried out across several related and unrelated subjects simultaneously. Both mpileup from SAMtools and the UnifiedGenotyper method from GATK were used, together a combination of parameter-based hard-filtering and variant quality score recalibration (DePristo et al., 2011). Known annotations anchored to the proper human genome reference used for the alignments were added to the detected variant sites using ANNOVAR (Wang et al., 2010) and/or variant tools (San Lucas et al., 2012). Annotated variants and genotype calls were stored in bgzip (SAMtools)-compressed VCF files.

Low quality samples (N=3) with mean depth <20 were removed prior to analysis, as well as SNPs with fewer than 3000 reads. Restricting to variants with mean sample depth >5 resulted in 555,651 SNPs of high quality. The final mean depth of the sequencing was 60.3X (range=4.2–106.4X). Supplementary Table 1 provides the transition/transversion (Ti/Tv) ratios broken down by SNP category for all variants included in the final analysis. The mean concordance with SNPs genotyped on the exome array (N=49,714) (Hellwege et al., 2014) was 99.93%. During quality control, one sample was removed for low concordance with SNPs on the exome array, one sample was removed for a substantial number of Mendelian inconsistencies, and an additional five samples were removed for inconsistencies between X and Y marker calls and participant gender. SHAPEIT2 (O'Connell et al., 2014) was used to perform Mendelian error checking with regard to established pedigree structures. Discordant genotypes were imputed to genotypes consistent with the pedigree structure, and individual SNPs with <95% efficiency were zeroed out. Samples found to have <95% overall efficiency by SHAPEIT2 (N=6) were excluded from the analysis.

Statistical analysis

SNP data were analyzed for both two-point family-based linkage and single variant association using Sequential Oligogenic Linkage Analysis Routines (SOLAR) (Almasy and Blangero, 1998). Briefly, both analyses were performed using the variance components method implemented in SOLAR, with age, sex, recruitment center (in the Hispanic cohorts), ancestry proportions (1-3 principal components depending on ethnic group) and BMI as covariates. The measured genotype analysis, which accounts for the non-independence of family members, involves incorporating each variant separately in a model as a measured covariate (the number of copies of the minor allele) evaluating genotype-specific differences in the trait means. These approaches have been documented in detail previously (Hellwege et al., 2016, 2014). Waist-to-hip ratio (WHR), waist circumference, visceral adipose tissue

area (VAT), and visceral to subcutaneous tissue ratio (VSR) were evaluated both with and without adjustment for BMI. Other measures of adiposity, including body adiposity index (BAI), subcutaneous adipose tissue area (SAT), and percent body fat, did not include BMI as a covariate. Supplementary Table 2 shows the phenotypes analyzed in this study, and where relevant, the transformations used to approximate normality. The presence of a low frequency variant encoding the G45R missense mutation in *ADIPOQ* (rs200573126:G>C) was included as a covariate for analyses involving adiponectin (ADP_G45R), as this variant is known to have a strong influence on adiponectin levels in this population (Hellwege et al., 2015, Bowden et al., 2010). Additionally, three admixture proportions were included as covariates in the association analysis. Estimates of admixture proportion had previously been computed by maximum likelihood estimation (MLE) of individual ancestries in ADMIXTURE after pruning for linkage disequilibrium (LD) and assuming five ancestral populations (K=5) to produce admixture estimates for the largest number of samples (Hellwege et al., 2014). Three of the five variables considered were selected as representative of the variation in these Hispanic samples, as they encompassed the majority of the variability.

Results

Demographic information, along with biometric and lipid characteristics of the samples relevant to this analysis, is shown in Supplementary Table 3. Overall, 555,651 variants were identified from whole exome sequencing with average depth of 60.3X in up to 1,205 Hispanic individuals from 78 families. To exclude variants found only in 12 individuals or less, a minor allele frequency (MAF) threshold was set at 0.5%, resulting in a set of 211,612 SNPs for analysis. Of the variants meeting the MAF threshold, 11,973 (5.66%) were previously unknown (relative to dbSNP build 138).

Linkage

Linkage analysis for each variant with 50 metabolic traits yielded 10,580,600 LOD scores with 1,148 LOD scores greater than 3, 183 LOD scores greater than 4, and 29 LOD scores greater than 5. The highest LOD score was in a previously identified broad linkage peak on chromosome 3 for adiponectin levels (Bowden et al., 2010). Excluding variants linked with adiponectin (not adjusted for the G45R variant), there were 1,060 LOD scores greater than or equal to 3, 68 LOD scores greater than 4, and 20 LOD scores greater than 4.5 (Table 1). Of these variants, the most biologically relevant result was evidence of linkage with SAT on chromosome 4 (rs2289043:T>C; LOD=5.49; Figure 1). This missense mutation (Met721Thr) is within *UNC5C*, which encodes a member of the UNC5H netrin receptor family. Additional strong linkage signals included: rs35705:G>A (intronic) in *GAS2L3* with gamma-glutamyl transpeptidase levels (LOD=5.33), rs116505219:A>C (missense; Phe51Cys) in *AQP12B* with adiponectin levels (LOD=5.30), and rs974334:G>C (intronic; *GPX6*) and rs139032867 (NP_653272.2:p.Gly133_Gly135del; *FAM109A*) with WHR adjusted for BMI (WHR_BMI; LOD=5.13 and 5.02, respectively).

In addition to specific variants, there were several notable broad linkage peaks. Most strikingly, a 67.5 Mb region of chromosome 10, including the majority of the p arm and a

large portion of the proximal q arm, showed evidence of linkage (LOD>3.0) with systolic blood pressure (SBP). The highest LOD score in this peak was 4.41 at rs74832669:G>T, a variant within the 3′-UTR of *SLC18A3* (Solute Carrier Family 18, Member 3). This gene encodes a transmembrane protein which transfers acetylcholine into vesicles for secretion into extracellular space. Three additional variants within *SLC18A3* (rs2269338:G>T, rs8175353:C>T, and rs1880675:C>T) had LOD scores above 4.0.

Additional broad linkage peaks included a region on chromosome 1 with acute insulin response (33 variants with LOD>3 in the peak and 3 with LOD>4; Supplementary Table 4) and a region on chromosome 12 with adiponectin levels (adjusted for G45R), both consistent with our previous findings (Hellwege et al., 2016). A smaller region on chromosome 6 was observed to show evidence of linkage with WHR_BMI. A number of variants within this linkage peak were located within human leukocyte antigen genes, including *HLA-C*, *HLA-DQA2*, and *HLA-DQB2*. The *GPX6* variant described above was also located within this region.

Association

Thirteen variants attained conventional genome-wide significance (p<5.0×10^{-8}; Table 2), with the strongest association between rs651821:C>T, a 5′-UTR variant in *APOA5*, and triglyceride levels (p=3.67×10^{-10}). Apolipoprotein AV is known to be involved in the regulation of plasma triglyceride levels (Pennacchio et al., 2002, 2001). Other genes containing variants significantly associated with cardiometabolic phenotypes included the well-documented *PNPLA3* gene association with measures of hepatic steatosis adjusted for BMI (rs738408:C>T, p=1.92×10^{-9}; rs738409:C>G, p=2.16×10^{-9}) (Palmer et al., 2013, Cox et al., 2011, Wagenknecht et al., 2011), *IMP4* with glucose effectiveness (rs72854959:G>T; p=4.25×10^{-9}), and *IDH1* with waist circumference (rs11554137:G>A; p=2.45×10^{-8}) and BMI (rs11554137:G>A; p=3.86×10^{-8}). The *IDH1* associations have been observed previously in this dataset (Gao et al., 2015).

Variants with evidence of both linkage and association

One variant, rs71508052:G>A, was found to have evidence of both linkage (LOD>3.0) and genome-wide significant association *(p<5.0×10^{-8})* with SBP (LOD=3.74, p=2.52×10^{-9}; Table 3). This variant maps to a locus on chromosome 10 purported to be a pseudogene predicted to have a regulatory function; however, this has not been functionally confirmed. Two additional genes contained variants showing both strong linkage and additional strong evidence of association with triglyceride levels: an intronic variant (rs189547099:G>C) in *FNIP2* and a novel intronic variant (g.157997598C>G) in *GLRB* (p=6.31×10^{-8}, LOD=3.13; both variants). Evaluation of suggestive variants (p=5.0×10^{-7}; LOD>2.0) revealed six additional SNPs: rs2072560:C>T and rs651821:T>C in *APOA5* with triglyceride levels (p=5.14×10^{-10}, LOD=2.06; p=3.67×10^{-10}, LOD=2.36, respectively), rs11554137:G>A in *IDH1* with two adiposity-related traits (p=2.45×10^{-8}, LOD=2.06, waist circumference; p=3.83×10^{-7}, LOD=2.19, WHR_BMI), and three SNPs in *MGRN1*, *SEC14L5*, and *ALG1* (rs748293549:C>T, rs763273802:G>A, and rs780440168:C>A, respectively) with percent body fat (p=1.67×10^{-7}, LOD=3.26; all variants).

Adiponectin

The strongest evidence of both linkage and association in this sample was with the G45R variant in *ADIPOQ* with adiponectin levels (rs200573126:G>C; *p*=4.53×10^{-41}, LOD=17.26). This finding is consistent with our previous results reporting a broad linkage peak on the distal q arm of chromosome 3 (Hellwege et al., 2015, Bowden et al., 2010, Guo et al., 2006). As expected, including the presence of the functional variant G45R as a covariate in the analyses greatly diminished the linkage peak (Supplementary Figure 1) and decreased the magnitude of association for variants within the region. This adjusted analysis identified rs116505219:A>C in *AQP12B*, an aquaporin, which showed strong linkage with adiponectin levels after adjustment for G45R (LOD=5.30; Table 1) as well as loci on chromosomes 10, 12, and 21 with LOD scores ≥ 4.5.

Comparison to exome chip

Of the 555K variants identified through sequencing of the exonic regions, 49,714 had previously been genotyped in Hispanics on the Illumina HumanExome Beadchip v1 (Hellwege et al., 2014). Whole exome sequencing identified 505,937 variants which were not on the exome chip; however, the chip included an additional 31,845 variants not seen in the WES sequencing, 14,218 of which were non-exonic. Although there were over 30,000 variants found in the exome chip dataset that were missing from the exome sequencing dataset, nearly half of these variants were non-exonic. In addition, the chip contains 4,761 GWAS-enriched common variants, 3,468 ancestry informative markers, and 3,369 markers for identity-by-descent. A sizeable percentage of these subsets may not have been detectable by exome sequencing due to intergenic positioning.

Overall, there was a 5.2-fold increase in the number of informative variants detected by exome sequencing relative to the exome chip. Among the SNPs unique to the exome sequencing dataset, 151,175 (29.9%) were previously unidentified variants (dbSNP build 138). Comparison of the top 25 variants showing evidence of strong linkage or association with results from the exome chip, as well as variants showing evidence of both linkage and association, revealed very little overlap, i.e. WES largely revealed new results. The only variants among the top signals also found on the exome chip were rs200573126:G>C (*ADIPOQ*), rs2289043:T>C (*UNC5C*), rs738409:C>G (*PNPLA3*), and rs4917:T>C (*AHSG*).

Discussion

In this study, we performed a family-based linkage and association analysis of WES data from 78 Hispanic American families with cardiometabolic traits. We were able to evaluate the performance of WES data compared to an earlier analysis of exome chip data in this Hispanic sample. The primary goal of this work was to identify novel loci, as well as known variants that were not included on the exome chip. Thus, we have evaluated 50 phenotypes related to anthropometry, glucose homeostasis, lipids, blood pressure, adiposity, hepatic fat and enzymes, and biomarkers. Considering the volume of results, some important observations can be noted. Most importantly, it is clear that many more exonic variants were identified using WES than the restricted catalog of SNPs included in chip-based genotyping.

This greatly enhances the capability to identify strong candidate SNPs for cardiometabolic traits. Consequently, this analysis greatly expands the results from this discovery analysis for potential replication in other cohorts.

A major advantage of using an approach which integrates results from both linkage and association analyses is the opportunity to identify variants within a population that, although they may not reach genome-wide significance, still show a strong signal of both linkage and association. Here, two such variants are found within *FNIP2* and *GLRB* with triglyceride levels. FNIP2 suppresses *PPARGC1a* RNA expression and is expressed in liver, pancreas, and adipose tissue (Hasumi et al., 2008); *GLRB* encodes the β unit of the glycine receptor, a neurotransmitter-gated chloride channel. The presence of insulin increases the effectiveness of glycine at its receptor (Caraiscos et al., 2007). The variants within the two genes appear to be in LD ($r^2=1$). It is noteworthy that the variant in *GLRB* is a novel SNP identified in this population by WES.

One of the more notable linkage results was a broad peak spanning the majority of the p arm and a large region of the proximal q arm of chromosome 10, showing evidence of linkage with SBP. Located within this peak, on the q arm, is *SLC18A3*, encoding a 57 kDa vesicular acetylcholine transporter (VAchT) family protein which assists in moving acetylcholine into secretory vesicles for extracellular release. The four variants showing strong linkage with SBP are located throughout the gene (two in the 3′-UTR, one in the 5′-UTR, and one synonymous coding variant). Previous reports have indicated that peripheral cholinergic systems are down-regulated during systemic inflammation (Lips et al., 2007), a co-morbidity of cardiometabolic disease, and that deterioration of cholinergic anti-inflammatory pathways promotes the development of hypertensive target organ damage, suggesting dysfunction of the VAchT (Li et al., 2011). We further explored the linkage between variants within *SLC18A3* and SBP by performing family-based linkage analysis. Of the two families showing nominal evidence of linkage, one appeared to show a distinct pattern of segregation of the minor allele with higher SBP values (mean SBP: major allele = 133.56, minor allele = 182.83). The number of carriers in the first family was low (N = 6), and the pattern was not repeated in the second family. However, the strongest evidence of linkage in our prior publication on exome chip data from this cohort, (rs7412 in *APOE* with plasma apolipoprotein B) also did not show strong family-specific linkage (Hellwege et al., 2014). It is also noteworthy that within this broad linkage peak, there is an intronic variant (rs71508052:G>A) showing evidence of both linkage and genome-wide significant association with SBP. Interestingly, the variant is located within a predicted pseudogene of unknown function, *NUTM2HP*.

While we did observe linkage peaks on chromosomes 1 (acute insulin response) and 12 (G45R-adjusted adiponectin levels) consistent with our previous report (Hellwege et al., 2016), the signals in the current report were not found to be as strong. This is likely due to the difference in sample sizes between the two studies; the exome chip analysis included up to 1,414 participants from 90 families compared with 1,205 samples from 78 families in the exome sequencing analysis. Nonetheless, the long arm of chromosome 1 is of particular interest, as linkage of this region with type 2 diabetes and metabolic syndrome has been documented in multiple populations, including Hispanics (Prokopenko et al., 2009,

Langefeld et al., 2004). Supplementary Table 4 lists the variants included within the chromosome 1q linkage peak. Although none of the variants listed in the table were found in previous reports of linkage with acute insulin response, a different variant within *F5* was reported to show genome-wide association with fasting serum levels of ferritin, one of a number of biomarkers included in a model to predict the 5-year risk of type 2 diabetes (Ahluwalia et al., 2015).

The single variant showing the strongest evidence of linkage with SAT, rs2289043:T>C, is located within *UNC5C*. This gene encodes a member of the UNC5H family of receptors which has high protein expression levels in the luminal membranes of the gall bladder, intestine, and kidney (Uhlén et al., 2015). UNC5C receptors bind netrin-1, a protein involved in both cell migration during neural development and macrophage retention in adipose tissue. Netrins have cell type-specific attractant and repellant properties; the repellant response is mediated by UNC5H and adenosine A2 B receptors (Corset et al., 2000, Ackerman et al., 1997, Leonardo et al., 1997). Recently, a suggestive association between an intronic variant (rs11097470:T>C) in *UNC5C* and energy balance ($p=5.14\times10^{-06}$) was reported in Hispanic children (Comuzzie et al., 2012). High expression levels of netrin-1 have also been found in obese, but not lean, adipose tissue in both humans and mice (Ramkhelawon et al., 2014). Taken together, this information suggests that alterations in UNC5C expression or structure could affect subcutaneous adipose tissue volume.

Several variants of biological relevance reached genome-wide significant association. As mentioned above, specific variants in *APOA5* are well known to have substantial effects on plasma triglyceride levels, particularly in African American and Hispanic populations (Pennacchio et al., 2002). Consistent with our previous reports and those of others, variants within *PNPLA3*, a triacylglycerol lipase, were found to have a significant association with measures of hepatic density (Palmer et al., 2013, Cox et al., 2011, Wagenknecht et al., 2011, Romeo et al., 2008). PNPLA3 is thought to be membrane-bound and may mediate energy storage and usage in adipocytes. *IDH1* (isocitrate dehydrogenase 1) is a gene associated with waist circumference in a previous report analyzing GWAS and exome chip data (Gao et al., 2015), although the variant identified was different than that reported in this study. Mutations in *IDH1* have primarily been examined for their role in the pathogenesis of cancers, particularly glioma and acute myeloid leukemia (Fujii et al., 2016). However, results from rodent models suggest that there may be a direct correlation between IDPc (cytosolic NAPD$^+$-dependent isocitrate dehydrogenase) and adipose tissue levels and that IDPc may be a necessary cofactor during adipogenesis (Koh et al., 2004).

Compared with our previous work using the HumanExome Bead Chip (Hellwege et al., 2014), exome sequencing identified more than five times the number of polymorphic SNPs in this population. More importantly, use of whole exome sequencing allowed for discovery of novel variants which accounted for nearly 30% of identified variants. While the exome chip did provide some insight into variants linked or associated with cardiometabolic disease, exome sequencing largely broadens that knowledge by vastly increasing the number of potential variants.

This study targeted genetic factors related to metabolic and cardiovascular disease in a Hispanic population. The majority of cardiometabolic genomic studies have been carried out in European-derived populations (Bowden, 2011). Hispanics are underrepresented in the type 2 diabetes literature, despite an increased prevalence of type 2 diabetes in the United States for Hispanics compared with non-Hispanic whites (12.8% vs. 7.6% as of 2014) (CDC/NCHS, 2016). In 2013, the most current data available, cardiovascular disease was the 2nd leading cause of death among Hispanics, with diabetes ranked as the 5th leading cause (Below and Parra, 2016, Heron, 2016). This report adds to the available literature on cardiometabolic disease in Hispanic populations.

This study, however, was not without limitations. Due to the cost of whole exome sequencing and the size of the study population, the sample used was relatively smaller in size, which can decrease statistical power. The analysis was limited to exons, miRNA, and untranslated regions; however, it is becoming clearer that many causal variants are intergenic or may be located in various non-coding RNAs other than miRNAs. It would be beneficial for future experiments to take advantage of family-based linkage and single-variant association analyses of whole genome sequencing data to identify the specific variants which significantly contribute to cardiovascular and metabolic diseases within this population.

In conclusion, WES data provided insights that were not able to be observed from exome chip data analysis and thus enabled identification of novel signals contributing to cardiometabolic traits. Further, the integration of results from two-point linkage and single-variant association analysis allowed for recognition of variants that may otherwise have been missed with either individual analysis separately. This study provided evidence that whole exome sequencing data, in combination with family-based linkage and association analyses, is a valuable tool in the discovery and identification of novel variants in complex disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Ackerman SL, Kozak LP, Przyborski SA, Rund LA, Boyer BB, Knowles BB. The mouse rostral cerebellar malformation gene encodes an UNC-5-like protein. Nature. 1997; 386:838–42. [PubMed: 9126743]

Ahluwalia TS, Allin KH, Sandholt CH, Sparsø TH, Jørgensen ME, Rowe M, Christensen C, Brandslund I, Lauritzen T, Linneberg A, Husemoen LL, Jørgensen T, Hansen T, Grarup N, Pedersen O. Discovery of coding genetic variants influencing diabetes-related serum biomarkers and their

impact on risk of type 2 diabetes. J Clin Endocrinol Metab. 2015; 100:E664–71. [PubMed: 25599387]

Albrechtsen A, Grarup N, Li Y, Sparsø T, Tian G, Cao H, Jiang T, Kim SY, Korneliussen T, Li Q, Nie C, Wu R, Skotte L, Morris AP, Ladenvall C, Cauchi S, Stan áková A, Andersen G, Astrup A, Banasik K, Bennett AJ, Bolund L, Charpentier G, Chen Y, Dekker JM, Doney AS, Dorkhan M, Forsen T, Frayling TM, Groves CJ, Gui Y, Hallmans G, Hattersley AT, He K, Hitman GA, Holmkvist J, Huang S, Jiang H, Jin X, Justesen JM, Kristiansen K, Kuusisto J, Lajer M, Lantieri O, Li W, Liang H, Liao Q, Liu X, Ma T, Ma X, Manijak MP, Marre M, Mokrosi ski J, Morris AD, Mu B, Nielsen AA, Nijpels G, Nilsson P, Palmer CN, Rayner NW, Renström F, RIBEL-Madsen R, Robertson N, Rolandsson O, Rossing P, Schwartz TW, Slagboom PE, Sterner M, Tang M, Tarnow L, Tuomi T, Van't Riet E, VAN Leeuwen N, Varga TV, Vestmar MA, Walker M, Wang B, Wang Y, Wu H, Xi F, Yengo L, Yu C, Zhang X, Zhang J, Zhang Q, Zhang W, Zheng H, Zhou Y, Altshuler D, 'T Hart LM, Franks PW, Balkau B, Froguel P, Mccarthy MI, Laakso M, Groop L, Christensen C, Brandslund I, Lauritzen T, Witte DR, et al. Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. Diabetologia. 2013; 56:298–310. [PubMed: 23160641]

Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet. 1998; 62:1198–211. [PubMed: 9545414]

Below JE, Parra EJ. Genome-Wide Studies of Type 2 Diabetes and Lipid Traits in Hispanics. Curr Diab Rep. 2016; 16:41. [PubMed: 27007718]

Borecki IB, Province MA. Genetic and genomic discovery using family studies. Circulation. 2008; 118:1057–63. [PubMed: 18765388]

Bowden DW. Will Family Studies Return to Prominence in Human Genetics and Genomics? Rare Variants and Linkage Analysis of Complex Traits. Genes & Genomics. 2011:1–8.

Bowden DW, An SS, Palmer ND, Brown WM, Norris JM, Haffner SM, Hawkins GA, Guo X, Rotter JI, Chen YD, Wagenknecht LE, Langefeld CD. Molecular basis of a linkage peak: exome sequencing and family-based analysis identify a rare genetic variant in the ADIPOQ gene in the IRAS Family Study. Hum Mol Genet. 2010; 19:4112–20. [PubMed: 20688759]

Caraiscos VB, Bonin RP, Newell JG, Czerwinska E, Macdonald JF, Orser BA. Insulin increases the potency of glycine at ionotropic glycine receptors. Mol Pharmacol. 2007; 71:1277–87. [PubMed: 17308032]

CDC/NCHS. Health, United States, 2015: With Special Feature on Racial and Ethnic Health Disparities. Hyattsville, MD: Centers for Disease Control and Prevention; 2016.

Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, Butte NF. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PLoS One. 2012; 7:e51954. [PubMed: 23251661]

Corset V, NGUYEN-BA-Charvet KT, Forcet C, Moyse E, Chédotal A, Mehlen P. Netrin-1-mediated axon outgrowth and cAMP production requires interaction with adenosine A2b receptor. Nature. 2000; 407:747–50. [PubMed: 11048721]

Cox AJ, Wing MR, Carr JJ, Hightower RC, Smith SC, Xu J, Wagenknecht LE, Bowden DW, Freedman BI. Association of PNPLA3 SNP rs738409 with liver density in African Americans with type 2 diabetes mellitus. Diabetes Metab. 2011; 37:452–5. [PubMed: 21665509]

Depristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, DEL Angel G, Rivas MA, Hanna M, Mckenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43:491–8. [PubMed: 21478889]

Fujii T, Khawaja MR, Dinardo CD, Atkins JT, Janku F. Targeting isocitrate dehydrogenase (IDH) in cancer. Discov Med. 2016; 21:373–80. [PubMed: 27355333]

Gao C, Wang N, Guo X, Ziegler JT, Taylor KD, Xiang AH, Hai Y, Kridel SJ, Nadler JL, Kandeel F, Raffel LJ, Chen YD, Norris JM, Rotter JI, Watanabe RM, Wagenknecht LE, Bowden DW, Speliotes EK, Goodarzi MO, Langefeld CD, Palmer ND. A Comprehensive Analysis of Common and Rare Variants to Identify Adiposity Loci in Hispanic Americans: The IRAS Family Study (IRASFS). PLoS One. 2015; 10:e0134649. [PubMed: 26599207]

Gazal S, Gosset S, Verdura E, Bergametti F, Guey S, Babron MC, Tournier-Lasserve E. Can whole-exome sequencing data be used for linkage analysis? Eur J Hum Genet. 2016; 24:581–6. [PubMed: 26173971]

Guo X, Saad MF, Langefeld CD, Williams AH, Cui J, Taylor KD, Norris JM, Jinagouda S, Darwin CH, Mitchell BD, Bergman RN, Sutton B, Chen YD, Wagenknecht LE, Bowden DW, Rotter JI. Genome-wide linkage of plasma adiponectin reveals a major locus on chromosome 3q distinct from the adiponectin structural gene: the IRAS family study. Diabetes. 2006; 55:1723–30. [PubMed: 16731835]

Hasumi H, Baba M, Hong SB, Hasumi Y, Huang Y, Yao M, Valera VA, Linehan WM, Schmidt LS. Identification and characterization of a novel folliculin-interacting protein FNIP2. Gene. 2008; 415:60–7. [PubMed: 18403135]

Hellwege JN, Palmer ND, Brown WM, Ziegler JT, An SS, Guo X, Chen YDI, Chen IY, Taylor K, Hawkins GA, Ng MC, Speliotes EK, Lorenzo C, Norris JM, Rotter JI, Wagenknecht LE, Langefeld CD, Bowden DW. Empirical characteristics of family-based linkage to a complex trait: the ADIPOQ region and adiponectin levels. Hum Genet. 2015; 134:203–13. [PubMed: 25447270]

Hellwege JN, Palmer ND, Dimitrov L, Keaton JM, Tabb KL, Sajuthi S, Taylor KD, Ng MCY, Speliotes EK, Hawkins GA, Long J, IDA Chen YD, Lorenzo C, Norris JM, Rotter JI, Langefeld CD, Wagenknecht LE, Bowden DW. Genome-wide linkage and association analysis of cardiometabolic phenotypes in Hispanic Americans. J Hum Genet. 2016

Hellwege JN, Palmer ND, Raffield LM, Ng MC, Hawkins GA, Long J, Lorenzo C, Norris JM, IDA Chen YD, Speliotes EK, Rotter JI, Langefeld CD, Wagenknecht LE, Bowden DW. Genome-wide family-based linkage analysis of exome chip variants and cardiometabolic risk. Genet Epidemiol. 2014; 38:345–52. [PubMed: 24719370]

Henkin L, Bergman RN, Bowden DW, Ellsworth DL, Haffner SM, Langefeld CD, Mitchell BD, Norris JM, Rewers M, Saad MF, Stamm E, Wagenknecht LE, Rich SS. Genetic epidemiology of insulin resistance and visceral adiposity. The IRAS Family Study design and methods. Ann Epidemiol. 2003; 13:211–7. [PubMed: 12684185]

Heron M. Deaths: Leading Causes for 2013. Natl Vital Stat Rep. 2016; 65:1–95.

Koh HJ, Lee SM, Son BG, Lee SH, Ryoo ZY, Chang KT, Park JW, Park DC, Song BJ, Veech RL, Song H, Huh TL. Cytosolic NADP+-dependent isocitrate dehydrogenase plays a key role in lipid metabolism. J Biol Chem. 2004; 279:39968–74. [PubMed: 15254034]

Langefeld CD, Wagenknecht LE, Rotter JI, Williams AH, Hokanson JE, Saad MF, Bowden DW, Haffner S, Norris JM, Rich SS, Mitchell BD, Study IRASF. Linkage of the metabolic syndrome to 1q23-q31 in Hispanic families: the Insulin Resistance Atherosclerosis Study Family Study. Diabetes. 2004; 53:1170–4. [PubMed: 15047638]

Leonardo ED, Hinck L, Masu M, KEINO-Masu K, Ackerman SL, Tessier-Lavigne M. Vertebrate homologues of C. elegans UNC-5 are candidate netrin receptors. Nature. 1997; 386:833–8. [PubMed: 9126742]

Li DJ, Evans RG, Yang ZW, Song SW, Wang P, Ma XJ, Liu C, Xi T, Su DF, Shen FM. Dysfunction of the Cholinergic Anti-Inflammatory Pathway Mediates Organ Damage in Hypertension. Hypertension. 2011; 57:298–307. [PubMed: 21173343]

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–60. [PubMed: 19451168]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup GPDP. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–9. [PubMed: 19505943]

Li Y, Vinckenbosch N, Tian G, Huerta-Sanchez E, Jiang T, Jiang H, Albrechtsen A, Andersen G, Cao H, Korneliussen T, Grarup N, Guo Y, Hellman I, Jin X, Li Q, Liu J, Liu X, Sparsø T, Tang M, Wu H, Wu R, Yu C, Zheng H, Astrup A, Bolund L, Holmkvist J, Jørgensen T, Kristiansen K, Schmitz O, Schwartz TW, Zhang X, Li R, Yang H, Wang J, Hansen T, Pedersen O, Nielsen R. Resequencing of 200 human exomes identifies an excess of low-frequency non-synonymous coding variants. Nat Genet. 2010; 42:969–72. [PubMed: 20890277]

Lips KS, Lührmann A, Tschernig T, Stoeger T, Alessandrini F, Grau V, Haberberger RV, Koepsell H, Pabst R, Kummer W. Down-regulation of the non-neuronal acetylcholine synthesis and release

machinery in acute allergic airway inflammation of rat and mouse. Life Sci. 2007; 80:2263–9. [PubMed: 17328924]

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, Mccarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, Mccarroll SA, Visscher PM. Finding the missing heritability of complex diseases. Nature. 2009; 461:747–53. [PubMed: 19812666]

Mckenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, Depristo MA. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20:1297–303. [PubMed: 20644199]

O'connell J, Gurdasani D, Delaneau O, Pirastu N, Ulivi S, Cocca M, Traglia M, Huang J, Huffman JE, Rudan I, Mcquillan R, Fraser RM, Campbell H, Polasek O, Asiki G, Ekoru K, Hayward C, Wright AF, Vitart V, Navarro P, Zagury JF, Wilson JF, Toniolo D, Gasparini P, Soranzo N, Sandhu MS, Marchini J. A general approach for haplotype phasing across the full spectrum of relatedness. PLoS Genet. 2014; 10:e1004234. [PubMed: 24743097]

Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. Nat Rev Genet. 2015; 16:275–84. [PubMed: 25824869]

Palmer ND, Musani SK, YERGES-Armstrong LM, Feitosa MF, Bielak LF, Hernaez R, Kahali B, Carr JJ, Harris TB, Jhun MA, Kardia SL, Langefeld CD, Mosley TH, Norris JM, Smith AV, Taylor HA, Wagenknecht LE, Liu J, Borecki IB, Peyser PA, Speliotes EK. Characterization of European ancestry nonalcoholic fatty liver disease-associated variants in individuals of African and Hispanic descent. Hepatology. 2013; 58:966–75. [PubMed: 23564467]

Pennacchio LA, Olivier M, Hubacek JA, Cohen JC, Cox DR, Fruchart JC, Krauss RM, Rubin EM. An apolipoprotein influencing triglycerides in humans and mice revealed by comparative sequencing. Science. 2001; 294:169–73. [PubMed: 11588264]

Pennacchio LA, Olivier M, Hubacek JA, Krauss RM, Rubin EM, Cohen JC. Two independent apolipoprotein A5 haplotypes influence human plasma triglyceride levels. Hum Mol Genet. 2002; 11:3031–8. [PubMed: 12417524]

Prokopenko I, Zeggini E, Hanson RL, Mitchell BD, Rayner NW, Akan P, Baier L, Das SK, Elliott KS, Fu M, Frayling TM, Groves CJ, Gwilliam R, Scott LJ, Voight BF, Hattersley AT, Hu C, Morris AD, Ng M, Palmer CN, TELLO-Ruiz M, Vaxillaire M, Wang CR, Stein L, Chan J, Jia W, Froguel P, Elbein SC, Deloukas P, Bogardus C, Shuldiner AR, Mccarthy MI, Consortium ITDQ. Linkage disequilibrium mapping of the replicated type 2 diabetes linkage signal on chromosome 1q. Diabetes. 2009; 58:1704–9. [PubMed: 19389826]

Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26:841–2. [PubMed: 20110278]

Ramkhelawon B, Hennessy EJ, Ménager M, Ray TD, Sheedy FJ, Hutchison S, Wanschel A, Oldebeken S, Geoffrion M, Spiro W, Miller G, Mcpherson R, Rayner KJ, Moore KJ. Netrin-1 promotes adipose tissue macrophage retention and insulin resistance in obesity. Nat Med. 2014; 20:377–84. [PubMed: 24584118]

Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, Boerwinkle E, Cohen JC, Hobbs HH. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet. 2008; 40:1461–5. [PubMed: 18820647]

Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, Boehnke M. Genome-wide association studies in diverse populations. Nat Rev Genet. 2010; 11:356–66. [PubMed: 20395969]

SAN Lucas FA, Wang G, Scheet P, Peng B. Integrated annotation and analysis of genetic variants from next-generation sequencing studies with variant tools. Bioinformatics. 2012; 28:421–2. [PubMed: 22138362]

Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, VON Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, VON Heijne G, Nielsen J, Pontén F. Proteomics. Tissue-based map of the human proteome. Science. 2015; 347:1260419. [PubMed: 25613900]

Wagenknecht LE, Palmer ND, Bowden DW, Rotter JI, Norris JM, Ziegler J, Chen YD, Haffner S, Scherzinger A, Langefeld CD. Association of PNPLA3 with non-alcoholic fatty liver disease in a minority cohort: the Insulin Resistance Atherosclerosis Family Study. Liver Int. 2011; 31:412–6. [PubMed: 21281435]

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38:e164. [PubMed: 20601685]
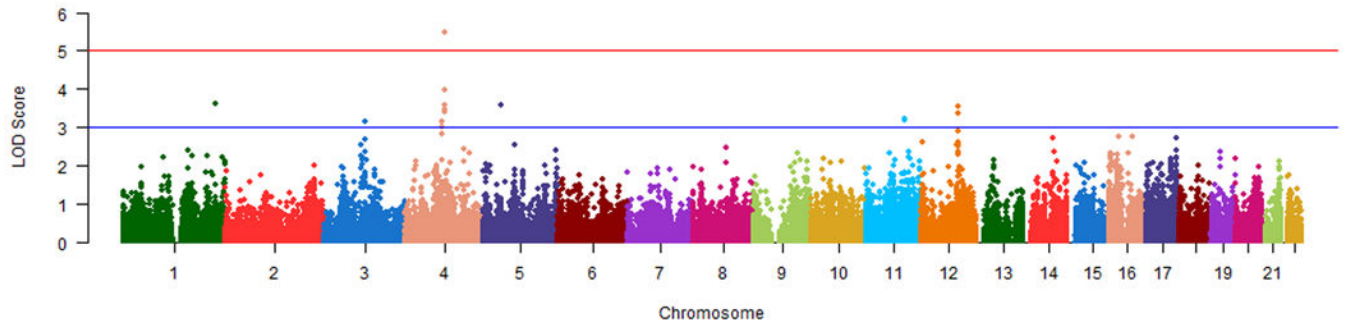
Figure 1.
Linkage plot for subcutaneous adipose tissue measurements. Note the peak on chromosome 4; the strongest linkage signal is rs2289043 in *UNC5C*.

Table 1

Variants with LOD score ≥ 4.5[*]

SNP	Chr	Position[1]	MAF	Gene	Annotation	Trait	P-value	LOD	Beta Value	N
rs116505219	2	241622103	0.056	AQP12B	Phe51Cys	ADP_G45R	0.296	5.30	-0.043	1097
rs1047369	10	122348969	0.160	PLPP4	Synonymous	ADP_G45R	1	4.68	7.70E-06	1097
rs3832810	12	26385301	0.120	SSPN	3′-UTR	ADP_G45R	0.305	4.65	-0.029	1097
rs112835312	12	88437346	0.007	C12orf29	Intronic	ADP_G45R	1.45E-06	4.76	-0.613	1097
rs67506729	21	15313088	0.205	C21orf81	Intronic	ADP_G45R	0.143	4.75	0.034	1097
rs414685	21	15327273	0.201	ANKRD20A11P	Intronic	ADP_G45R	0.075	4.75	0.042	1097
rs413639	21	15333206	0.203	ANKRD20A11P	Intronic	ADP_G45R	0.097	4.71	0.039	1097
rs7281846	21	15755349	0.486	HSPA13	Intronic	ADP_G45R	0.476	4.66	0.014	1097
rs11648905	16	425298	0.409	TMEM8A	Intronic	ApoB	0.066	4.68	-0.096	1104
rs35719	12	100967660	0.396	GAS2L3	Intronic	GGT	0.113	4.75	0.057	1029
rs35705	12	100974894	0.393	GAS2L3	Intronic	GGT	0.027	5.33	0.080	1029
rs738409	22	44324727	0.403	PNPLA3	Ile148Met	Percent Body Fat	0.580	4.84	0.162	912
rs738408	22	44324730	0.404	PNPLA3	Synonymous	Percent Body Fat	0.584	4.75	0.160	912
rs2289043	4	96106322	0.222	UNC5C	Met721Thr	SAT	0.223	5.49	-0.267	1150
rs1523519	3	121631841	0.394	SLC15A2	Intronic	TNFaR2	0.535	4.65	0.009	946
rs2026724	4	159212904	0.277	TMEM144	NA	TNFaR2	0.148	4.81	0.023	946
rs2026725	4	159213138	0.273	TMEM144	NA	TNFaR2	0.132	4.65	0.024	946
rs76613066	12	42478689	0.096	GXYLT1	3′-UTR	TNFaR2	0.040	4.80	0.045	946
rs974334	6	28474218	0.220	GPX6	Intronic	WHR_BMI	0.245	5.13	0.004	1203
rs139032867	12	111800826	0.374	FAM109A	p.del133-135	WHR_BMI	0.159	5.02	0.004	1203

[*] Linkage results for adiponectin levels not adjusted for the G45R variant were excluded.

[1] Relative to build GRCh37/hg19

Abbreviations: ApoB, apolipoprotein B; GGT, gamma glutamyl transferase; TNFaR2, TNFα receptor 2

Table 2

Variants with P-value ≤ 5.0 × 10⁻⁰⁸[*]

SNP	Chr	Position[1]	MAF	Gene	Annotation	Trait	P-value	LOD	Beta Value	Variance[2]	N
rs11554137	2	209113192	0.065	*IDH1*	Synonymous	BMI	3.86E-08	1.68	-0.101	0.02	1203
rs7083870	10	25144390	0.009	*PRTFDC1*	Intronic	Fasting Insulin	3.54E-08	0.19	-0.808	0.03	1041
rs199620366	1	16268435	0.034	*ZBTB17*	3′-UTR	GGT	4.01E-08	0.06	0.511	0.03	1029
rs72854959	2	131100415	0.013	*IMP4*	5′-UTR	Glucose Effectiveness	4.25E-09	0.01	-0.006	0	983
rs738409	22	44324727	0.400	*PNPLA3*	Ile148Met	Liver Density	2.34E-09	0.08	-311.741	0.05	884
rs738408	22	44324730	0.401	*PNPLA3*	Synonymous	Liver Density	2.17E-09	0.09	-312.240	0.05	884
rs3762329	1	102252905	0.005	*LINC01307*	NA	Normalized Liver	2.52E-09	1.19	1.878	0.04	884
rs738409	22	44324727	0.400	*PNPLA3*	Ile148Met	Normalized Liver	2.16E-09	0.18	-0.292	0.06	884
rs738408	22	44324730	0.401	*PNPLA3*	Synonymous	Normalized Liver	1.92E-09	0.20	-0.292	0.06	884
rs71508052	10	52445346	0.008	*NUTM2HP*	NA	SBP	2.52E-09	3.74	0.166	0.03	1202
rs2072560	11	116661826	0.132	*APOA5*	Intronic	Triglycerides	5.14E-10	2.06	0.235	0.03	1195
rs651821	11	116662579	0.142	*APOA5*	5′-UTR	Triglycerides	3.67E-10	2.36	0.231	0.03	1195
rs11554137	2	209113192	0.065	*IDH1*	Synonymous	Waist Circumference	2.45E-08	2.06	-0.073	0.02	1202

[*]Association results for adiponectin levels not adjusted for the G45R variant were excluded

[1]Relative to build GRCh37/hg19

[2]Proportion of total variance of trait attributed to SNP by association analysis

Table 3

Variants showing evidence suggestive of both linkage (LOD ≥ 2) and association (p ≤ 5.0 × 10⁻⁰⁷)[*]

SNP	Chr	Position[1]	MAF	Gene	Annotation	Trait	P-value	LOD	Beta Value	Variance[2]	N
rs748293549	16	4715100	0.013	*MGRN1*	Ala210Val	Percent Body Fat	1.67E-07	3.25	-6.856	0.036	912
rs763273802	16	5041850	0.013	*SEC14L5*	Synonymous	Percent Body Fat	1.67E-07	3.26	-6.856	0.036	912
rs780440168	16	5105390	0.013	*ALG1*	Intronic	Percent Body Fat	1.67E-07	3.26	-6.856	0.036	912
rs71508052	10	52445346	0.008	*NUTM2HP*	NA	SBP	2.52E-09	3.74	0.166	0.030	1202
c4_157997598[3]	4	157997598	0.005	*GLRB*	Intronic	Triglycerides	6.31E-08	3.13	1.039	0.031	1195
rs189547099	4	159814881	0.005	*FNIP2*	Intronic	Triglycerides	6.31E-08	3.13	1.039	0.031	1195
rs2072560	11	116661826	0.132	*APOA5*	Intronic	Triglycerides	5.14E-10	2.06	0.235	0.034	1195
rs651821	11	116662579	0.142	*APOA5*	5′-UTR	Triglycerides	3.67E-10	2.36	0.231	0.033	1195
rs11554137	2	209113192	0.065	*IDH1*	Synonymous	Waist Circumference	2.45E-08	2.06	-0.073	0.022	1202
rs11554137	2	209113192	0.065	*IDH1*	Synonymous	WHR	3.83E-07	2.19	-0.030	0.020	1203

[*] Linkage and association results for adiponectin levels not adjusted for the G45R variant were excluded

[1] Relative to build GRCh37/hg19

[2] Proportion of total variance of trait attributed to SNP by association analysis

[3] Nomenclature is provided as chr_position (relative to hg19)