Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance

Luca A. Lotta[1], Pawan Gulati[2], Felix R. Day[1], Felicity Payne[3], Halit Ongen[4], Martijn van de Bunt[5,6], Kyle J. Gaulton[7], John D. Eicher[8], Stephen J. Sharp[1], Jian'an Luan[1], Emanuella De Lucia Rolfe[1], Isobel D. Stewart[1], Eleanor Wheeler[3], Sara M. Willems[1], Claire Adams[2], Hanieh Yaghootkar[9], EPIC-InterAct Consortium[10], Cambridge FPLD1 Consortium[10], Nita G. Forouhi[1], Kay-Tee Khaw[11], Andrew D. Johnson[8], Robert K. Semple[2], Timothy Frayling[9], John R. B. Perry[1], Emmanouil Dermitzakis[4], Mark I. McCarthy[5,6], Inês Barroso[3,2]*, Nicholas J. Wareham[1]*, David B. Savage[2]*, Claudia Langenberg[1]*, Stephen O'Rahilly[2]*, Robert A. Scott[1]*.

1 MRC Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom
2 Metabolic Research Laboratories, Institute of Metabolic Science, University of Cambridge, Cambridge, United Kingdom
3 Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom
4 Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland
5 Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom
6 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom
7 Department of Pediatrics, University of California San Diego, La Jolla, USA
8 Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda, USA
9 Genetics of Complex Traits, Institute of Biomedical and Clinical Science, University of Exeter Medical School, Royal Devon and Exeter Hospital, Exeter, United Kingdom
10 A list of members and affiliations appears at the end of the manuscript
11 Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom

*these authors contributed equally

Correspondence to:
Robert A. Scott (robert.scott@mrc-epid.cam.ac.uk)
Claudia Langenberg (Claudia.langenberg@mrc-epid.cam.ac.uk)
Nicholas J. Wareham (nick.wareham@mrc-epid.cam.ac.uk)
David B. Savage (dbs23@medschl.cam.ac.uk)
Stephen O'Rahilly (so104@medschl.cam.ac.uk)
Inês Barroso (ib1@sanger.ac.uk)

Abstract

Insulin resistance is a key mediator of obesity-related cardiometabolic disease, yet the mechanisms underlying this link remain obscure. Using an integrative genomic approach, we identify 53 genomic regions associated with insulin resistance phenotypes (higher fasting insulin adjusted for BMI, lower HDL cholesterol and higher triglycerides) and provide evidence that their link with higher cardiometabolic risk is underpinned by an association with *lower* adipose mass in peripheral compartments. Using these 53 loci, we show a polygenic contribution to familial partial lipodystrophy-type 1, a severe form of insulin resistance, and highlight shared molecular mechanisms between common/mild and rare/severe insulin resistance. Population-level genetic analyses combined with experiments in cellular models implicate CCDC92, DNAH10 and L3MBTL3 as previously unrecognised molecules influencing adipocyte differentiation. Our findings support the notion that limited storage capacity of peripheral adipose tissue is an important aetiological component in insulin-resistant cardiometabolic disease and highlight genes and mechanisms underpinning this link.

Introduction

Insulin resistance, usually defined as an impaired ability of insulin to maintain normal glucose metabolism and initially manifested by higher levels of circulating insulin, is positively associated with adiposity and is a key mediator of the link between obesity and its adverse impact on metabolic and cardiovascular disease.[1-8] Given the current global epidemic of metabolic disease, there is an urgent need for improved understanding of the mechanisms that link over-nutrition to insulin resistance in the general population.[7-10]

Among individuals stratified on the basis of overall adiposity, there is considerable variation in the extent of adverse metabolic sequelae,[11] demonstrating the importance of other factors in the aetiology of insulin resistance and its complications. Indeed, while insulin resistance often coexists with obesity, severe forms of insulin resistance develop without obesity or in association with generalized or regional lack of adipose tissue, i.e. lipodystrophy.[12] In lipodystrophies,[13] it has been proposed that the impaired capacity of peripheral adipose tissue to expand under the challenge of a positive energy balance leads to lipid accumulation at ectopic sites (e.g. liver, skeletal muscle, pancreas) and eventually to overt diabetes.[12,14] The notion of "adipose overflow" or "limited adipose tissue expandability"[15-19] is supported (i) by the metabolic disturbances seen in rare, monogenic lipodystrophies and their dramatic amelioration in response to dietary calorie restriction[20] or leptin replacement,[21,22] and (ii) by a series of elegant rodent studies including those in which adipose tissue capacity was expanded by fat transplantation in lipodystrophic mice[23] or by over-expressing adiponectin,[24] or where partially lipodystrophic mice were energetically challenged by rendering them leptin deficient.[25,26] However, the relevance of this model to the general population remains uncertain.

Some initial human genetic insights into more prevalent forms of insulin resistance are available. Genome-wide studies of gold-standard measures of insulin resistance have been limited by sample size,[27] but multiple genomic loci have been associated with fasting insulin

levels, a widely-measured marker of insulin resistance.[28,29] A subset of these loci were associated with higher triglycerides and lower high-density lipoprotein (HDL) cholesterol,[28] which are hallmarks of insulin resistance.[13,30] These loci were later validated for their association with insulin resistance,[31] suggesting that the combined association with this triad of phenotypes could help identify specific genetic determinants of insulin resistance.

Given the availability of large-scale genome-wide association data on lipid traits and fasting insulin,[28,29,32] we undertook an integrative genomic approach to characterise genetic and molecular mechanisms underpinning insulin resistance at a given level of adiposity and its role in cardiometabolic disease in the general population.

Results

Associations with insulin resistance phenotypes at 53 independent genomic regions

We combined genome-wide association results for fasting insulin adjusted for body mass index,[22,23] HDL-cholesterol and triglyceride levels[32] from up to 188,577 individuals to identify loci associated with a phenotypic pattern indicative of insulin resistance (**Online Methods**, **Supplementary Figures 1-2 and Supplementary Table 1**). After aligning the association results of ~2.4 million single nucleotide polymorphisms (SNPs) to the insulin-raising allele, 630 SNPs from 53 1Mb-genomic regions were associated with higher fasting insulin, higher triglycerides and lower HDL-cholesterol ($p<0.005$ for each phenotype, expected probability of association under null hypothesis $p=3.1 \times 10^{-08}$; **Online Methods**, **Supplementary Figures 3-4**). These 53 genomic regions included 10 loci previously implicated in insulin resistance,[31] and an additional 43 loci (**Supplementary Table 2**). A subset of 25 of the 53 loci had previously been associated with HDL cholesterol or triglyceride levels at genome-wide significance,[32] while 28 had not (**Supplementary Table 2**). We first investigated the associations of these loci in a completely independent sample of 6,101 individuals and found that genetic risk scores comprising the 53 lead SNPs were strongly associated with higher fasting insulin, higher triglycerides and lower HDL-cholesterol (**Supplementary Figure 5**). We next asked whether these variants were associated with "gold-standard" measures of insulin sensitivity. Having a greater number of risk-alleles from the 53-SNP, 43-SNP or 28-SNP (excluding loci previously implicated in insulin resistance and lipid traits, respectively) genetic scores was strongly associated with (a) lower insulin sensitivity measured by euglycemic clamp or insulin suppression test in 2,764 individuals[27] (p-value for 53-SNP genetic score = 4.3×10^{-06}; **Table 1**) and (b) lower insulin sensitivity index in 4,769 individuals with a frequently-sampled oral glucose tolerance test[33] (p-value for 53-SNP genetic score = 7.3×10^{-10}; **Table 1**).

Genetic predisposition to insulin resistance via the 53 loci confers higher risk of cardiometabolic disease but lower levels of peripheral adiposity

We next investigated associations of the 53 genomic regions with a range of continuous metabolic traits and disease outcomes. In 45,836 cases and 230,358 controls, the 53-SNP genetic score was associated with a higher risk of type 2 diabetes (odds ratio [OR] per standard deviation [SD] of the genetic score [i.e. ~4.5 alleles], 1.12; 95% confidence interval [CI], 1.11-1.14; $p=9.2 \times 10^{-61}$; **Table 1**). In studies with available individual-level data, we saw no difference in associations across sex or body mass index (BMI) strata (**Supplementary Table 3**). Genetically-predicted insulin resistance was also associated with a higher risk of coronary heart disease (**Table 1**). The associations with type 2 diabetes (OR, 1.10; $p=9.0 \times 10^{-32}$) and coronary heart disease (OR, 1.04; $p=9.7 \times 10^{-07}$) remained after removing 13 loci that were previously shown to be associated with either of the two diseases at genome-wide significance.[34,35] Association estimates were also consistent after removing the 25 loci previously associated with HDL cholesterol or triglyceride levels at genome-wide significance (**Table 1**). We also observed an association with coronary heart disease in 5,369 cases of coronary heart disease and 106,969 controls from the UK Biobank study not previously included in genome-wide discovery analyses of insulin or lipid traits (OR, 1.09; $p=5.3 \times 10^{-09}$). Individually, 30 of the 53 lead SNPs were associated with higher type 2 diabetes risk ($p<0.05$; **Supplementary Table 4**), including a novel association at genome-wide significance for rs718314 near *ITPR2* (OR per allele, 1.06; $p=6.8 \times 10^{-09}$). We found an enrichment of loci associated with higher risk of both type 2 diabetes and coronary heart disease, including those encompassing the proximal insulin signalling *INSR*, *IRS1* and *PIK3R1* genes (11/53 loci associated with both diseases at $p<0.05$; two-tailed binomial probability of observing this proportion of loci by chance $p=8.5 \times 10^{-22}$; **Supplementary Table 5**).

While insulin resistance is often considered secondary to higher adiposity, at the 53 loci we observed associations with *lower* body fat percentage, BMI and hip circumference (**Figure 1A, Supplementary Figures 6 and 7**). The larger magnitude of association with measures of body fat rather than with glycaemic phenotypes is consistent with a primary effect of many variants on adipose tissue mass (**Figure 1A, Supplementary Figures 6 and 7**).

By follow-up studies using dual-energy X-ray absorptiometry (DEXA) measures in 12,848 individuals, we found that the most marked association of the genetic score was with lower levels of gynoid and leg fat mass (**Figure 1B**). Individuals in the highest quintile of the 53-SNP genetic score had an average of 712 grams *less* leg fat mass compared to individuals in the bottom quintile (**Figure 1C**), which accounted for the majority of the overall body fat association (**Supplementary Figure 8**). The association with lower levels of leg fat was accompanied by a higher hazard of incident type 2 diabetes (**Figure 1C**). In 9,150 participants from the EPIC-Norfolk cohort who gained weight during a median follow-up of 3.7 years, carrying a greater number of the 53 risk alleles was inversely associated with change in hip circumference, adjusted for the amount of weight gained (i.e. individuals carrying more alleles were less likely to deposit extra mass in their gluteal region for a given increase in body mass; β in cm of hip circumference per SD of genetic score, -0.07; p=0.027; **Supplementary Note**). Overall, these association analyses suggest that individuals genetically predisposed to insulin resistance via the 53 loci have a relative inability to expand their peripheral fat compartment when challenged by a positive energy balance and that this incapacity results in higher cardiometabolic disease risk. We also found that the 53-SNP genetic score was associated with higher levels of alanine aminotransferase and gamma-glutamyl transferase (**Supplementary Table 6**), which suggests that the failure to store lipid in gluteofemoral adipose tissue may be accompanied by hepatic lipid deposition.

The 53-SNP genetic score was associated with greater waist circumference (**Figure 1A**), but not with trunk adipose tissue (**Figure 1B**), indicating that the association with body fat distribution and cardiometabolic disease was largely driven by the association with lower levels of peripheral adipose tissue (**Figure 1 and Supplementary Figures 6-8**). Among the 53 lead SNPs, 17 were within 500kb of a waist-to-hip ratio (WHR) associated SNP[36] (**Supplementary Figure 6**). Consistent with DEXA analyses, the associations with WHR at this subset of overlapping loci were largely driven by an association with lower hip circumference, rather than higher waist circumference (**Supplementary Figure 6 and Supplementary Table 4**).

Our large-scale meta-analyses allowed the investigation of individual-SNP associations with both adiposity and metabolic risk. At eight of the 53 loci, the lead SNP was associated with *lower* total body fat percentage or hip circumference at genome-wide significance ($p<5 \times 10^{-08}$), including a novel association of the insulin-raising G allele of rs4976033 near *PIK3R1* with lower body fat percentage ($p=3.0 \times 10^{-09}$; see **Figure 1D and Supplementary Figure 9)**. Seven of the eight adiposity-lowering alleles at these loci were associated with a higher risk of type 2 diabetes ($p<0.05$; **Figure 1D**).

Role of common variants in the genetic basis of a severe form of lipodystrophy

Given the strong association with insulin resistance but lower levels of peripheral adiposity, we hypothesized that the polygenic predisposition to insulin resistance imparted by the 53 loci might contribute to the pathogenesis of familial partial lipodystrophy-type 1 (FPLD1). When compared to women from the population-based Fenland study, women diagnosed with FPLD1 displayed markedly lower levels of leg fat mass for a given fat mass in the rest of the body (**Figure 2A**). Whilst the name of the condition implies Mendelian inheritance, using exome sequencing in 9 FPLD1 cases we did not identify likely candidate causal genes (**Online Methods and Supplementary Table 7**). When compared with 5,296

unrelated women from the UKHLS study in a case-control analysis, 37 patients with FPLD1 had a higher burden of the 53 risk alleles (OR per SD of genetic score in logistic regression analyses adjusted for age and the first 10 genetic principal components, 1.70; 95% CI, 1.21-2.39; p=0.0021, **Figure 2B**; $p_{permutation}$=0.0020, see **Online Methods**). Also, the phenotypes observed in FPLD1 patients in comparison to obese women from the Fenland study mirrored the association pattern of the 53-SNP genetic score (**Supplementary Table 8**). FPLD1 women had a more severe leg fat phenotype compared to that expected from the relationship between the 53 SNP score and leg fat mass in the Fenland study (**Figure 2C**), suggesting that additional genetic and environmental factors contribute to determining this extreme phenotype.

Prioritisation of putative effector genes, cell-types and tissues

We prioritised putative effector genes at the 53 loci by integrating data about physical proximity, expression quantitative trait locus (eQTL) mapping, functional annotation and previous knowledge about genes causing monogenic forms of insulin resistance (summarised in **Supplementary Table 2;** see also **Supplementary Tables 9-11** for details). Putative effector genes included five with well-established roles in proximal insulin signalling (**Supplementary Table 2 and Figure 3A**). Other candidates included *LPL*, encoding the key lipolysis regulator lipoprotein lipase (**Supplementary Table 2 and Figure 3B**). The insulin-lowering, minor allele of rs1011685 (near *LPL*) is on the same haplotype (r^2=1) as a gain-of-function,[37] protein-truncating allele in *LPL* (p.Ser447*; rs328; minor-allele frequency [MAF], 9.9%). The p.Ser447* gain-of-function variant was recently reported to be associated with lower risk of coronary heart disease,[38] while an independent loss-of-function missense variant of *LPL*[39] (p.Asp36Asn; rs1801177; MAF, 1.9%) was associated with higher risk.[38] Here, we found that the p.Ser447* gain-of-function variant is associated with greater insulin sensitivity, lower fasting glucose, lower levels of liver markers and protection from type 2 diabetes (OR

per allele, 0.93; p=1.6 x 10^{-05}; **Figure 3B and Supplementary Figure 10**). Conversely, the p.Asp36Asn loss-of-function variant in *LPL* is associated with a higher risk of type 2 diabetes (OR per allele, 1.11; p=0.0086; **Figure 3B and Supplementary Figure 10**). Thus, recent findings of an allelic series of *LPL* variants implicating lipoprotein lipase as a putative therapeutic target in heart disease[38] are now complemented by a directionally consistent observation for type 2 diabetes, compatible with a role for impaired lipoprotein lipase-mediated lipolysis in insulin resistance and type 2 diabetes.

Among the 53 loci, three contained genes at which rare mutations have been previously implicated in severe monogenic forms of insulin resistance (i.e. *PPARG, PIK3R1, INSR;* **Figure 3C**), which is more than what expected by chance given the prevalence of monogenic insulin resistance genes in the genome (observed percentage 0.54% [3 out of a total of 553 genes in the 53 loci], expected percentage 0.064%, two-tailed binomial p=0.0056). The *PIK3R1* gene encodes regulatory subunits of a critical kinase involved in proximal insulin signalling and rare, loss-of-function mutations in this gene are associated with SHORT syndrome, a dysmorphic condition characterised by short stature, partial lipodystrophy and insulin resistance.[40-43] To date, such mutations have been identified in few families worldwide and data from Exome Aggregation Consortium show this gene to have decreased tolerance of missense variation (Z=2.42) and to be particularly intolerant of loss-of-function mutations (pLI=1; Exome Aggregation Consortium, Cambridge, MA, URL: http://exac.broadinstitute.org accessed 22nd March 2016). Here, we provide evidence that a common variant near *PIK3R1*, which accounts for almost half of all alleles in the general population (MAF=49%), is associated with subtle effects on insulin resistance, lower body fat percentage, higher risk of type 2 diabetes and coronary heart disease (**Supplementary Figure 9**). This association pattern partially overlaps with that reported for *PIK3R1* mutations and SHORT syndrome (**Supplementary Table 12**). We also found that the common rs8101064

allele T in *INSR,* encoding the insulin receptor, was associated with insulin resistance and higher risk of type 2 diabetes (OR per allele, 1.08; p=0.020), but not with body fat percentage (p=0.16), consistent with the fact that patients with heterozygous loss-of-function mutations in the *INSR* are frequently insulin resistant but are not commonly lipodystrophic.[13]

We assessed the overlap of lead SNPs and their proxies (r^2>0.8) with functional regulatory annotations across 98 cell types from the NIH Epigenome Roadmap (**Online Methods**) and, consistent with phenotypic associations, identified substantial overlap with adipose tissue active enhancer elements (31 of 53 loci overlapped adipose tissue active enhancer elements; observed percentage=58.4% of loci, expected=30.1%, binomial p=2.1 x 10^{-05}; **Online Methods and Figure 3D**). Furthermore, combined pathway analyses with integration of large-scale gene expression data[44] implicated adipocytes as likely effector cell-type underlying observed associations (**Figure 3E)**. In subcutaneous adipocyte eQTL data from 1,064 individuals of the EUROBATS and GTex projects (**Online Methods and Supplementary Table 10**), we observed evidence of eQTL associations with nearby genes ($p<10^{-06}$) at 21 loci including 14 with supportive evidence of co-localisation of lead phenotypic and eQTL associations (**Online Methods and Supplementary Table 10**).

Experimental validation of putative effector genes in cellular adipogenesis models

We sought to experimentally validate the role of five putative effector genes (*IRS1*, *CCDC92*, *DNAH10*, *L3MBTL3* and *FAM13A*) across four loci which showed associations with (a) expression of nearby transcripts in subcutaneous adipocytes, (b) lower peripheral adiposity and (c) higher metabolic disease risk (**Supplementary Note, Supplementary Figure 11, Supplementary Tables 10 and 13**). We used short interfering RNA (siRNA) to reduce mRNA levels for these five genes in OP9-K cells, an adipocyte model which shows rapid differentiation in response to adipogenic stimuli.[45] Knockdown of the *IRS1*, *CCDC92,*

DNAH10 and *L3MBTL3* genes significantly reduced both mRNA of the target gene (**Figure 4A** top graph) and lipid accumulation (**Figure 4A** bottom graph and **Figure 4B**). These results were directionally consistent with the association of the insulin-raising alleles at these loci with lower expression of these genes in subcutaneous adipocytes and with lower levels of peripheral fat (**Figure 4C**). Knockdown of *FAM13A* reduced mRNA (**Figure 4A** top graph), but did not significantly affect pre-adipocyte lipid accumulation (**Figure 4A** bottom graph and **Figure 4B**). In contrast to the other four genes, the risk allele at *FAM13A* was associated with higher mRNA expression of this gene in subcutaneous adipocytes (**Figure 4C**).

Discussion

Our data implicate the impaired capacity to adequately expand the peripheral adipose tissue compartments in human insulin resistance and related disease at the population level. These results substantially augment existing evidence[31,36,46-49] by clarifying the extent and relevance of adipose tissue dysfunction to cardiometabolic disease and by providing novel mechanistic insights into its underpinning biology.

Our results are consistent with the existence of dozens of genomic regions at which common genetic variation affects cardiometabolic disease risk via subtle "lipodystrophy-like" mechanisms. While these common genetic mechanisms have individually small effects, their cumulative effect is large and relevant to a large fraction of the population. For instance, we observed ~700 grams difference in leg fat mass between the top and bottom 20% of the population distribution of risk alleles. We also show a polygenic contribution to an extreme phenotype, referred to as FPLD1 or Köbberling-type lipodystrophy, illustrating the contribution of common alleles to severe forms of insulin resistance. At given loci (e.g. *PIK3R1*), we found that genetic variants at the two extremes of the allele frequency spectrum result in corresponding consequences at the extremes of the phenotypic severity spectrum. These findings strongly concur with the notion that molecular and pathophysiologic mechanisms first described in severe forms of lipodystrophic insulin resistance are relevant to the general population. While a centripetal distribution of body fat is a well-recognised risk factor for metabolic and cardiovascular disease,[50-54] there is considerable confusion about the underlying mechanisms and relative importance of lower peripheral fat versus higher central adiposity. Whilst supportive of a role for central fat accumulation, different lines of evidence from this study suggest a role for impaired peripheral fat deposition in insulin-resistant cardiometabolic disease in the general population. These include strong associations with gluteo-femoral adiposity, overlap with regulatory regions in adipose tissue and with genomic

loci known to cause lipodystrophies, as well as functional characterisation in adipocytes. These findings provide evidence from large-scale human genetics studies which add to a body of research about the links between subcutaneous and lower-body adipocyte phenotypes and a favourable metabolic profile.[19,55-57]

By combining population-scale association studies with eQTL data from adipose tissue and experimental evidence from murine cellular models, we provide evidence that specific risk loci influence adipose gene expression resulting in impaired adipogenesis, reduced peripheral fat depots and ultimately increased cardiometabolic disease risk. For the *L3MBTL3*, *DNAH10* or *CCDC92* genes, evidence presented in this study provides the first link with impaired adipocyte differentiation capacity. L3MBTL3 recognises methylated lysine residues on histone tails[58] and previous genome-wide anthropometric studies have implicated this locus in adult height and length at birth.[59-61] At the chromosome 12q24 locus, our analyses are consistent with the implication of both *CCDC92* and *DNAH10* in impaired adipogenesis. CCDC92 is a coiled coil domain protein which interacts with proteins in the centriole/ciliary interface.[62] *DNAH10* encodes one of the heavy chains of the dynein arms of the motile cilia and it is, therefore, surprising that its product appears to have cell autonomous effects on adipocyte biology. However, an essential splice site mutation in *DNAH10* has previously been reported to co-segregate with abnormal circulating HDL-cholesterol levels in a family,[63] while the locus containing both *CCDC92* and *DNAH10* has been associated with circulating levels of large HDL particles,[64] further supporting an unexpected role for these proteins in metabolism.

Our results have preventive and therapeutic implications for cardiovascular and metabolic disease. First, they suggest that attempts to develop pharmacological agents acting on the molecular mechanisms of obesity are likely to reduce cardiometabolic risk if they reduce calorie intake (e.g. by acting on appetite) or reduce ectopic fat deposition in tissues such as the liver, muscle and pancreas, but not if they impair adipogenesis or peripheral fat deposition.

Agents that promote adipocyte differentiation and increase peripheral adipose mass via action on the peroxisome proliferator-activated receptor gamma have powerful antidiabetic actions,[65,66] and in some cases have a beneficial effect on cardiovascular outcomes,[67,68] although some of these agents have been reported to have an adverse cardiovascular safety profile.[69] An early but vital challenge in the translation of genetic findings towards therapeutic insight is the ability to identify likely effector transcripts underlying genetic associations. In the current study, we identify putative effectors of genetic associations and the tissues in which they operate. We also demonstrate that these genetic variants often affect the risk of type 2 diabetes and coronary disease in a consistent direction, which suggests that targeting these pathways may satisfy current regulatory requirements that anti-diabetic agents should not be associated with unacceptable cardiovascular risk.[70] This is particularly true of findings from both gain- and loss-of-function variants in the *LPL* gene and risk of type 2 diabetes, which are directionally consistent with those previously reported for the same mutations and risk of heart disease.[38] Notably, the directional concordance for risk of heart disease and type 2 diabetes is in contrast to genetic evidence for other lipid-lowering agents (e.g. cholesterol-lowering variants near the molecular target of statins).[71] In the context of a growing body of evidence linking lipolysis and heart disease risk,[38,72-75] these data suggest that enhancing lipoprotein lipase activity may also become a viable preventive or therapeutic strategy in type 2 diabetes.

In interpreting the results of this study, it is important to note that combining multiple genetic-association analyses to gain insights about a latent unmeasured phenotype (i.e. insulin resistance) is not immediately comparable with a univariate genome-wide association study of a trait (e.g. fasting insulin). However, we validated these genetic variants as being strongly associated with "gold-standard" measures of insulin sensitivity, with multiple insulin-resistance related diseases, including a severe form of insulin-resistant partial lipodystrophy, and showed overlap with monogenic insulin resistance genes. Thus, approaches to leverage

additional sources of evidence to prioritise genomic variation (such as multiple phenotypes or putative functional class[76]) represent a powerful use of extant genetic association results to advance understanding of biology previously intractable to conventional strategies.

Our results were based on genome-wide analyses of fasting insulin adjusted for BMI,[28,29] and we did not identify loci with a primary effect on higher adiposity and secondary association with insulin resistance (e.g. *FTO*). Our approach was more likely to identify loci influencing insulin resistance at a given level of adiposity. Prompted by the strong association pattern of the genetic scores and variants, we focused on evaluation of mechanisms linking lower levels of peripheral adiposity with insulin resistance. The importance of adipose function as a prominent driver of common insulin resistance is highlighted by the observation that half of all variants associated with fasting insulin at genome-wide significance (without adjustment for BMI) in a previous study[28] were included in our genetic score. However, our results do not preclude the presence nor diminish the importance of other mechanisms underlying insulin resistance.[5,77] Indeed, the associations we observe of the genetic score with central fat, visceral fat and liver enzymes would be further strengthened after adjustment for overall adiposity. While we conclude that our findings implicate a primary effect on impaired adipose function and a secondary effect on insulin resistance, we cannot entirely exclude the possibility of the reverse, nor that there are pleiotropic contributions to the associations.

Conclusions

Collectively, our findings support the notion that limited capacity of peripheral adipose tissue to store surplus energy is implicated in human insulin resistance and related cardiometabolic disease in the general population. Furthermore, we highlight putative effector genes, tissues and mechanisms underpinning this link.

Data Access Statement

This research has been conducted using the UK Biobank resource. Data on glycemic traits were contributed by the MAGIC consortium investigators. Associations with type 2 diabetes were obtained from the DIAGRAM (DIAbetes Genetics Replication And Meta-analysis) consortium investigators. Data on coronary heart disease / myocardial infarction have been contributed by the CARDIoGRAMplusC4D consortium investigators. Data on body mass index, waist, hip, waist-to-hip ratio were contributed by the GIANT consortium investigators. Data about triglycerides and high-density lipoprotein cholesterol were contributed by the Global Lipids Genetics Consortium investigators. We are very grateful to the GENESIS consortium for provision of summary statistic results for clamp- or insulin suppression test-based insulin resistance. The authors would like to thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at http://exac.broadinstitute.org/about. Raw exome sequence data from FPLD1 individuals and family members is available from the European Genome-Phenome Archive (https://www.ebi.ac.uk/ega/home, see full accession codes in Supplementary Table 7). Understanding Society: The UK Household Longitudinal Study is led by the Institute for Social and Economic Research at the University of Essex and funded by the Economic and Social Research Council. The survey was conducted by NatCen and the genome-wide scan data were analysed and deposited by the Wellcome Trust Sanger Institute. Information on how to access the data can be found at https://www.understandingsociety.ac.uk/. Genome-wide genotyping data of 5,296 unrelated women from UKHLS is publically available through the European Genome-phenome Archive (Dataset Accession: EGAD00010000891).

Data download:

MAGIC consortium (http://www.magicinvestigators.org/)
GLGC consortium (http://csg.sph.umich.edu//abecasis/public/lipids2013/)
GIANT (https://www.broadinstitute.org/collaboration/giant/)
DIAGRAM consortium (http://diagram-consortium.org/)
CARDIoGRAMplusC4D (http://www.cardiogramplusc4d.org/)
Exome Aggregation Consortium (http://exac.broadinstitute.org/about)
European Genome-Phenome Archive (https://www.ebi.ac.uk/ega/home)

Study websites:

Fenland (http://www.mrc-epid.cam.ac.uk/research/studies/fenland/)
EPIC-Norfolk (http://www.srl.cam.ac.uk/epic/)
EPIC-InterAct (http://www.inter-act.eu/)
UK Biobank (http://www.ukbiobank.ac.uk/)
Understanding Society: The UK Household Longitudinal Study (https://www.understandingsociety.ac.uk/)

Acknowledgements

This study was funded by the United Kingdom's Medical Research Council through grants MC_UU_12015/1, MC_PC_13046, MC_PC_13048 and MR/L00002/1. This work was supported by the MRC Metabolic Diseases Unit (MC_UU_12012/5) and the Cambridge NIHR Biomedical Research Centre and EU/EFPIA Innovative Medicines Initiative Joint Undertaking (EMIF grant: 115372). EPIC-InterAct Study funding: funding for the InterAct project was provided by the EU FP6 programme (grant number LSHM_CT_2006_037197). This work was funded, in part, through an EFSD Rising Star award to R.A.S. supported by Novo Nordisk. D. B. S. is supported by the Wellcome Trust grant n. 107064. M. McC. is a Wellcome Trust Senior Investigator and is supported by the following grants from the Wellcome Trust: 090532 and 098381. M. vdB. is supported by a Novo Nordisk postdoctoral fellowship run in partnership with the University of Oxford. I. B. is supported by the Wellcome Trust grant WT098051. S. O. R. acknowleges funding from the Wellcome Trust (Wellcome Trust Senior Investigator Award : 095515/Z/11/Z and Wellcome Trust Strategic Award: 100574/Z/12/Z).

We are grateful for OP9-K cells kindly shared by Alan Kopin's laboratory, Tufts Medical Center, Boston, USA. The authors gratefully acknowledge the help of the MRC Epidemiology Unit Support Teams, including Field Teams, Laboratory Team and Data Management Team, and of the staff of the Wellcome Trust Clinical Research Facility.

Author Contributions

Concept and design: L. A. Lotta, I. Barroso, N. J. Wareham, D. B. Savage, C. Langenberg, S. O'Rahilly, R. A. Scott. *Generation, acquisition, analysis, or interpretation of data:* all authors. *Drafting of the manuscript:* L. A. Lotta, I. Barroso, N. J. Wareham, D. B. Savage, C. Langenberg, S. O'Rahilly, R. A. Scott. Critical review of the manuscript for important intellectual content and approval of the final version of the manuscript: all authors.

Competing Financial Interests Statement

The authors report no conflict of interest relative to this study.

EPIC-InterAct consortium authors

Claudia Langenberg(1), Robert A Scott(1), Stephen J Sharp(1), Nita G Forouhi(1), Nicola D Kerrison(1), Matt Sims(1), Debora ME Lucarelli(1), Inês Barroso(2,3), Panos Deloukas(2), Mark I McCarthy(5,6), Larraitz Arriola(12,13,14), Beverley Balkau(15,16), Aurelio Barricarte(14,17,18), Heiner Boeing(19), Paul W Franks(20,21), Carlos Gonzalez(22), Sara Grioni(23), Rudolf Kaaks(24), Timothy J Key(25), Carmen Navarro(14,26,27), Peter M Nilsson(28), Kim Overvad(29,30), Domenico Palli(31), Salvatore Panico(32), J. Ramón Quirós(33), Olov Rolandsson(21), Carlotta Sacerdote(34,35), Elena Salamanca-Fernández(14,36,37), Nadia Slimani(38), Anne Tjonneland(39), Rosario Tumino(40), Annemieke MW Spijkerman(41), Daphne L van der A(41), Yvonne T van der Schouw(42), Elio Riboli(43), Nicholas J Wareham(1).

EPIC-InterAct consortium affiliations: (12) Public Health Division of Gipuzkoa, San Sebastian, Spain, (13) Instituto BIO-Donostia, Basque Government, San Sebastian, Spain, (14) CIBER Epidemiología y Salud Pública (CIBERESP), Spain, (15) Inserm, CESP, U1018, Villejuif, France, (16) Univ Paris-Sud, UMRS 1018, Villejuif, France, (17) Navarre Public Health Institute (ISPN), Pamplona, Spain, (18) Navarra Institute for Health Research (IdiSNA) Pamplona, Spain, (19) German Institute of Human Nutrition Potsdam-Rehbruecke, Germany, (20) Lund University, Malmö, Sweden, (21) Umeå University, Umeå, Sweden, (22) Catalan Institute of Oncology (ICO), Barcelona, Spain, (23) Epidemiology and Prevention Unit, Milan, Italy, (24) German Cancer Research Centre (DKFZ), Heidelberg, Germany, (25) Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom, (26) Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain, (27) Unit of Preventive Medicine and Public Health, School of Medicine, University of Murcia, Murcia, Spain, (28) Department of Clinical Sciences, Lund University, Skane University Hospital, Malmo, Sweden, (29) Department of Public Health, Section for Epidemiology, Aarhus University, Aarhus, Denmark, (30) Aalborg University Hospital, Aalborg, Denmark, (31) Cancer Research and Prevention Institute (ISPO), Florence, Italy, (32) Dipartimento di Medicina Clinica e Chirurgia, Federico II University, Naples, Italy, (33) Public Health Directorate, Asturias, Spain, (34) Unit of Cancer Epidemiology, Citta' della Salute e della Scienza Hospital-University of Turin and Center for Cancer Prevention (CPO), Torino, Italy, (35) Human Genetics Foundation (HuGeF), Torino, Italy, (36) Andalusian School of Public Health, Granada, Spain, (37) Instituto de Investigación Biosanitaria de Granada (Granada.ibs), Granada, Spain, (38) International Agency for Research on Cancer, Lyon, France, (39) Danish Cancer Society Research Center, Copenhagen, Denmark, (40) ASP Ragusa, Ragusa, Italy, (41) National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands, (42)

University Medical Center Utrecht, Utrecht, the Netherlands, (43) School of Public Health, Imperial College London, London, United Kingdom.

Cambridge FPLD1 consortium authors

Robert Semple(2), Claire Adams(2), Anna Stears(44), Stella George(45), Mark Walker(46), Ellie Gurnell(44), Deirdre Maguire(47), Rasha Mukhtar(48), Sath Nag(48), Amanda Adler(44), Maarten Soeters(49), Ken Laji(50), Alistair Watt(51), Simon Aylwin(52), Andrew Johnson(53), Gerry Rayman(54), Fahmy Hanna(55), Sian Ellard(56), Richard Ross(57), Kristina Blaslov(58), Lea Smirčić Duvnjak(58), Stephen O'Rahilly(2), David B. Savage(2).

Cambridge FPLD1 consortium affiliations: (44) Wolfson Diabetes and Endocrine Clinic, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom, (45) East and North Herts NHS Trust, Lister Hospital, Herts, United Kingdom, (46) Institute of Cellular Medicine (Diabetes), Newcastle University Medical School, Newcastle upon Tyne, United Kingdom, (47) Harrogate and District Hospital, Harrogate, United Kingdom, (48) James Cook University Hospital, Middlesborough, United Kingdom, (49) Department of Endocrinology & Metabolism, Internal Medicine, Academic Medical Center, Amsterdam, The Netherlands, (50) St Richard's Hopsital, Spitalfield Lane, Chichester, United Kingdom, (51) North Devon District Hospital, Raleigh Park, Barnstaple, United Kingdom, (52) King's College Hospital, London, United Kingdom, (53) Department of Diabetes and Endocrinology, Southmead Hospital, Bristol, United Kingdom, (54) Ipswich Hospital, Ipswich, United Kingdom, (55) University Hospital of North Midlands NHS Trust, Royal Stoke University Hospital, Stoke-on-Trent, United Kingdom, (56) Royal Devon and Exeter NHS Foundation Trust, Exeter, United Kingdom, (57) The University of Sheffield, The Medical School, Sheffield, United Kingdom, (58) Vuk Vrhovac University Clinic for Diabetes, Endocrinology and Metabolic diseases, Zagreb, Croatia.

References

1 Samuel, V. T. & Shulman, G. I. The pathogenesis of insulin resistance: integrating signaling pathways and substrate flux. *The Journal of clinical investigation* **126**, 12-22, doi:10.1172/JCI77812 (2016).

2 Ginsberg, H. N. Insulin resistance and cardiovascular disease. *The Journal of clinical investigation* **106**, 453-458, doi:10.1172/JCI10762 (2000).

3 Shulman, G. I. Ectopic fat in insulin resistance, dyslipidemia, and cardiometabolic disease. *The New England journal of medicine* **371**, 1131-1141, doi:10.1056/NEJMra1011035 (2014).

4 Lillioja, S. & Bogardus, C. Obesity and insulin resistance: lessons learned from the Pima Indians. *Diabetes/metabolism reviews* **4**, 517-540 (1988).

5 Perry, R. J., Samuel, V. T., Petersen, K. F. & Shulman, G. I. The role of hepatic lipids in hepatic insulin resistance and type 2 diabetes. *Nature* **510**, 84-91, doi:10.1038/nature13478 (2014).

6 Arner, P. The adipocyte in insulin resistance: key molecules and the impact of the thiazolidinediones. *Trends in endocrinology and metabolism: TEM* **14**, 137-145 (2003).

7 Friedman, J. M. Obesity in the new millennium. *Nature* **404**, 632-634, doi:10.1038/35007504 (2000).

8 Guilherme, A., Virbasius, J. V., Puri, V. & Czech, M. P. Adipocyte dysfunctions linking obesity to insulin resistance and type 2 diabetes. *Nature reviews. Molecular cell biology* **9**, 367-377, doi:10.1038/nrm2391 (2008).

9 Friedman, J. M. A war on obesity, not the obese. *Science* **299**, 856-858, doi:10.1126/science.1079856 (2003).

10 Hardy, O. T., Czech, M. P. & Corvera, S. What causes the insulin resistance underlying obesity? *Current opinion in endocrinology, diabetes, and obesity* **19**, 81-87, doi:10.1097/MED.0b013e3283514e13 (2012).

11 Stefan, N., Haring, H. U., Hu, F. B. & Schulze, M. B. Metabolically healthy obesity: epidemiology, mechanisms, and clinical implications. *Lancet Diabetes Endocrinol* **1**, 152-162, doi:10.1016/S2213-8587(13)70062-7 (2013).

12 Robbins, A. L. & Savage, D. B. The genetics of lipid storage and human lipodystrophies. *Trends in molecular medicine* **21**, 433-438, doi:10.1016/j.molmed.2015.04.004 (2015).

13 Semple, R. K., Savage, D. B., Cochran, E. K., Gorden, P. & O'Rahilly, S. Genetic syndromes of severe insulin resistance. *Endocrine reviews* **32**, 498-514, doi:10.1210/er.2010-0020 (2011).

14 Samuel, V. T., Petersen, K. F. & Shulman, G. I. Lipid-induced insulin resistance: unravelling the mechanism. *Lancet* **375**, 2267-2277, doi:10.1016/S0140-6736(10)60408-4 (2010).

15 Danforth, E., Jr. Failure of adipocyte differentiation causes type II diabetes mellitus? *Nature genetics* **26**, 13, doi:10.1038/79111 (2000).

16 Unger, R. H. Lipid overload and overflow: metabolic trauma and the metabolic syndrome. *Trends in endocrinology and metabolism: TEM* **14**, 398-403 (2003).

17 Virtue, S. & Vidal-Puig, A. Adipose tissue expandability, lipotoxicity and the Metabolic Syndrome--an allostatic perspective. *Biochimica et biophysica acta* **1801**, 338-349, doi:10.1016/j.bbalip.2009.12.006 (2010).

18 Shulman, G. I. Cellular mechanisms of insulin resistance. *The Journal of clinical investigation* **106**, 171-176, doi:10.1172/JCI10583 (2000).

19 Karpe, F. & Pinnick, K. E. Biology of upper-body and lower-body adipose tissue--link to whole-body phenotypes. *Nature reviews. Endocrinology* **11**, 90-100, doi:10.1038/nrendo.2014.185 (2015).

20 Robbins, D. C. *et al.* The effect of diet on thermogenesis in acquired lipodystrophy. *Metabolism: clinical and experimental* **28**, 908-916 (1979).

21 Shimomura, I., Hammer, R. E., Ikemoto, S., Brown, M. S. & Goldstein, J. L. Leptin reverses insulin resistance and diabetes mellitus in mice with congenital lipodystrophy. *Nature* **401**, 73-76, doi:10.1038/43448 (1999).

22 Oral, E. A. *et al.* Leptin-replacement therapy for lipodystrophy. *The New England journal of medicine* **346**, 570-578, doi:10.1056/NEJMoa012437 (2002).

23 Gavrilova, O. *et al.* Surgical implantation of adipose tissue reverses diabetes in lipoatrophic mice. *The Journal of clinical investigation* **105**, 271-278, doi:10.1172/JCI7901 (2000).
24 Kim, J. Y. *et al.* Obesity-associated improvements in metabolic profile through expansion of adipose tissue. *The Journal of clinical investigation* **117**, 2621-2637, doi:10.1172/JCI31021 (2007).
25 Gray, S. L. *et al.* Leptin deficiency unmasks the deleterious effects of impaired peroxisome proliferator-activated receptor gamma function (P465L PPARgamma) in mice. *Diabetes* **55**, 2669-2677, doi:10.2337/db06-0389 (2006).
26 Medina-Gomez, G. *et al.* PPAR gamma 2 prevents lipotoxicity by controlling adipose tissue expandability and peripheral lipid metabolism. *PLoS genetics* **3**, e64, doi:10.1371/journal.pgen.0030064 (2007).
27 Knowles, J. W. *et al.* Identification and validation of N-acetyltransferase 2 as an insulin sensitivity gene. *The Journal of clinical investigation* **126**, 403, doi:10.1172/JCI85921 (2016).
28 Scott, R. A. *et al.* Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nature genetics* **44**, 991-1005, doi:10.1038/ng.2385 (2012).
29 Manning, A. K. *et al.* A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nature genetics* **44**, 659-669, doi:10.1038/ng.2274 (2012).
30 Salazar, M. R. *et al.* Comparison of the abilities of the plasma triglyceride/high-density lipoprotein cholesterol ratio and the metabolic syndrome to identify insulin resistance. *Diabetes & vascular disease research* **10**, 346-352, doi:10.1177/1479164113479809 (2013).
31 Scott, R. A. *et al.* Common genetic variants highlight the role of insulin resistance and body fat distribution in type 2 diabetes, independent of obesity. *Diabetes* **63**, 4378-4387, doi:10.2337/db14-0319 (2014).
32 Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. *Nature genetics* **45**, 1274-1283, doi:10.1038/ng.2797 (2013).
33 Prokopenko, I. *et al.* A central role for GRB10 in regulation of islet function in man. *PLoS genetics* **10**, e1004235, doi:10.1371/journal.pgen.1004235 (2014).
34 Morris, A. P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature genetics* **44**, 981-990, doi:10.1038/ng.2383 (2012).
35 Nikpay, M. *et al.* A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nature genetics* **47**, 1121-1130, doi:10.1038/ng.3396 (2015).
36 Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187-196, doi:10.1038/nature14132 (2015).
37 Kozaki, K. *et al.* Mutational analysis of human lipoprotein lipase by carboxy-terminal truncation. *Journal of lipid research* **34**, 1765-1772 (1993).
38 Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators. Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *The New England journal of medicine*, doi:10.1056/NEJMoa1507652 (2016).
39 Mailly, F. *et al.* A common variant in the gene for lipoprotein lipase (Asp9-->Asn). Functional implications and prevalence in normal and hyperlipidemic subjects. *Arteriosclerosis, thrombosis, and vascular biology* **15**, 468-478 (1995).
40 Avila, M. *et al.* Clinical reappraisal of SHORT syndrome with PIK3R1 mutations: towards recommendation for molecular testing and management. *Clinical genetics*, doi:10.1111/cge.12688 (2015).
41 Thauvin-Robinet, C. *et al.* PIK3R1 mutations cause syndromic insulin resistance with lipoatrophy. *American journal of human genetics* **93**, 141-149, doi:10.1016/j.ajhg.2013.05.019 (2013).
42 Chudasama, K. K. *et al.* SHORT syndrome with partial lipodystrophy due to impaired phosphatidylinositol 3 kinase signaling. *American journal of human genetics* **93**, 150-157, doi:10.1016/j.ajhg.2013.05.023 (2013).

43 Dyment, D. A. *et al.* Mutations in PIK3R1 cause SHORT syndrome. *American journal of human genetics* **93**, 158-166, doi:10.1016/j.ajhg.2013.06.005 (2013).

44 Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nature communications* **6**, 5890, doi:10.1038/ncomms6890 (2015).

45 Lane, J. M., Doyle, J. R., Fortin, J. P., Kopin, A. S. & Ordovas, J. M. Development of an OP9 derived cell line as a robust model to rapidly study adipocyte differentiation. *PloS one* **9**, e112123, doi:10.1371/journal.pone.0112123 (2014).

46 Yaghootkar, H. *et al.* Genetic evidence for a normal-weight "metabolically obese" phenotype linking insulin resistance, hypertension, coronary artery disease, and type 2 diabetes. *Diabetes* **63**, 4369-4377, doi:10.2337/db14-0318 (2014).

47 Lu, Y. *et al.* New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. *Nature communications* **7**, 10495, doi:10.1038/ncomms10495 (2016).

48 Kilpelainen, T. O. *et al.* Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. *Nature genetics* **43**, 753-760, doi:10.1038/ng.866 (2011).

49 Yaghootkar, H. *et al.* Genetic evidence for a link between favorable adiposity and lower risk of type 2 diabetes, hypertension and heart disease. *Diabetes*, doi:10.2337/db15-1671 (2016).

50 Biggs, M. L. *et al.* Association between adiposity in midlife and older age and risk of diabetes in older adults. *JAMA : the journal of the American Medical Association* **303**, 2504-2512, doi:10.1001/jama.2010.843 (2010).

51 Pischon, T. *et al.* General and abdominal adiposity and risk of death in Europe. *The New England journal of medicine* **359**, 2105-2120, doi:10.1056/NEJMoa0801891 (2008).

52 Vague, J. The degree of masculine differentiation of obesities: a factor determining predisposition to diabetes, atherosclerosis, gout, and uric calculous disease. *The American journal of clinical nutrition* **4**, 20-34 (1956).

53 Smith, U. Abdominal obesity: a marker of ectopic fat accumulation. *The Journal of clinical investigation* **125**, 1790-1792, doi:10.1172/JCI81507 (2015).

54 Despres, J. P. & Lemieux, I. Abdominal obesity and metabolic syndrome. *Nature* **444**, 881-887, doi:10.1038/nature05488 (2006).

55 Dahlman, I. *et al.* Numerous Genes in Loci Associated With Body Fat Distribution Are Linked to Adipose Function. *Diabetes* **65**, 433-437, doi:10.2337/db15-0828 (2016).

56 Ryden, M., Andersson, D. P., Bergstrom, I. B. & Arner, P. Adipose tissue and metabolic alterations: regional differences in fat cell size and number matter, but differently: a cross-sectional study. *The Journal of clinical endocrinology and metabolism* **99**, E1870-1876, doi:10.1210/jc.2014-1526 (2014).

57 Pinnick, K. E. *et al.* Distinct developmental profile of lower-body adipose tissue defines resistance against obesity-associated metabolic complications. *Diabetes* **63**, 3785-3797, doi:10.2337/db14-0385 (2014).

58 Baughman, B. M., Pattenden, S. G., Norris, J. L., James, L. I. & Frye, S. V. The L3MBTL3 Methyl-Lysine Reader Domain Functions As a Dimer. *ACS chemical biology* **11**, 722-728, doi:10.1021/acschembio.5b00632 (2016).

59 Randall, J. C. *et al.* Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS genetics* **9**, e1003500, doi:10.1371/journal.pgen.1003500 (2013).

60 Paternoster, L. *et al.* Adult height variants affect birth length and growth rate in children. *Human molecular genetics* **20**, 4069-4075, doi:10.1093/hmg/ddr309 (2011).

61 Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature genetics* **46**, 1173-1186, doi:10.1038/ng.3097 (2014).

62 Gupta, G. D. *et al.* A Dynamic Protein Interaction Landscape of the Human Centrosome-Cilium Interface. *Cell* **163**, 1484-1499, doi:10.1016/j.cell.2015.10.065 (2015).

63 Singaraja, R. R. *et al.* Identification of four novel genes contributing to familial elevated plasma HDL cholesterol in humans. *Journal of lipid research* **55**, 1693-1701, doi:10.1194/jlr.M048710 (2014).

64 Chasman, D. I. *et al.* Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS genetics* **5**, e1000730, doi:10.1371/journal.pgen.1000730 (2009).
65 DeFronzo, R. A. *et al.* Pioglitazone for diabetes prevention in impaired glucose tolerance. *The New England journal of medicine* **364**, 1104-1115, doi:10.1056/NEJMoa1010949 (2011).
66 DREAM (Diabetes REduction Assessment with ramipril and rosiglitazone Medication) Trial Investigators. Effect of rosiglitazone on the frequency of diabetes in patients with impaired glucose tolerance or impaired fasting glucose: a randomised controlled trial. *Lancet* **368**, 1096-1105, doi:10.1016/S0140-6736(06)69420-8 (2006).
67 Lincoff, A. M., Wolski, K., Nicholls, S. J. & Nissen, S. E. Pioglitazone and risk of cardiovascular events in patients with type 2 diabetes mellitus: a meta-analysis of randomized trials. *JAMA : the journal of the American Medical Association* **298**, 1180-1188, doi:10.1001/jama.298.10.1180 (2007).
68 Kernan, W. N. *et al.* Pioglitazone after Ischemic Stroke or Transient Ischemic Attack. *The New England journal of medicine*, doi:10.1056/NEJMoa1506930 (2016).
69 Nissen, S. E. & Wolski, K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *The New England journal of medicine* **356**, 2457-2471, doi:10.1056/NEJMoa072761 (2007).
70 Food and Drug Administration. Guidance for Industry. Diabetes Mellitus — Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes (2008). URL: http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm071627.pdf.
71 Swerdlow, D. I. *et al.* HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* **385**, 351-361, doi:10.1016/S0140-6736(14)61183-1 (2015).
72 Dewey, F. E. *et al.* Inactivating Variants in ANGPTL4 and Risk of Coronary Artery Disease. *The New England journal of medicine*, doi:10.1056/NEJMoa1510926 (2016).
73 Do, R. *et al.* Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature* **518**, 102-106, doi:10.1038/nature13917 (2015).
74 Jorgensen, A. B., Frikke-Schmidt, R., Nordestgaard, B. G. & Tybjaerg-Hansen, A. Loss-of-function mutations in APOC3 and risk of ischemic vascular disease. *The New England journal of medicine* **371**, 32-41, doi:10.1056/NEJMoa1308027 (2014).
75 TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *The New England journal of medicine* **371**, 22-31, doi:10.1056/NEJMoa1307095 (2014).
76 Myocardial Infarction Genetics Consortium Investigators. Inactivating mutations in NPC1L1 and protection from coronary heart disease. *The New England journal of medicine* **371**, 2072-2082, doi:10.1056/NEJMoa1405386 (2014).
77 Moltke, I. *et al.* A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. *Nature* **512**, 190-193, doi:10.1038/nature13425 (2014).

Figure Legends

Figure 1. Combined associations with detailed anthropometry and metabolic disease risk at the 53 genomic loci. *Panel A:* association of the 53-SNP genetic score with anthropometric and glycaemic traits in meta-analyses of genetic association studies. Body mass index (BMI), waist-to-hip ratio (WHR), waist and hip circumference data are from the GIANT consortium and the UK Biobank study. Body fat percentage data are from the UK Biobank, EPIC-Norfolk and Fenland studies. Fasting plasma glucose (FPG), 2 hour glucose and HbA1c data are from the MAGIC consortium. Squares with error bars represent the per-allele beta coefficients in standard deviation units and their 95% confidence intervals. *Panel B:* association of genetic scores with compartmental body masses. Data are from 12,848 participants of the Fenland and EPIC-Norfolk studies who underwent a DEXA scan. Squares with error bars represent the per-allele beta coefficients in standard deviation units and their 95% confidence intervals. *Panel C*: association with lower levels of leg fat mass and higher hazard of incident type 2 diabetes by quintiles of the 53-SNP genetic risk scores. Associations are reported for individuals in the exposed category compared with the bottom quintile (reference category). Associations with leg fat mass are from 9,747 participants of the Fenland study and are reported on the left. Associations with incident type 2 diabetes are from 7,420 incident cases and 9,267 controls of the InterAct study and are reported on the right. Squares represent the beta coefficients in grams of leg fat mass (left plot) or the hazard ratio (HR) for incident type 2 diabetes (right plot) in each category compared with the lowest quintile. Error bars represent the 95% confidence intervals of these estimates. *Panel D*: associations of individual lead SNPs at eight loci with waist, hip circumference (left) and type 2 diabetes (right). Loci were selected on the basis of their genome-wide significant association with hip circumference or body fat percentage (i.e. *PIK3R1*). Waist and hip circumference analyses are from a meta-analysis of the GIANT and UK Biobank studies. Type 2 diabetes analyses are from a meta-analysis of the DIAGRAM, InterAct and UK Biobank studies. Squares with error bars represent the per-allele beta coefficients in standard deviation units of waist and hip circumference (left plot) or the per-allele odds ratio (OR) of type 2 diabetes (right plot). Error bars represent the 95% confidence intervals of these estimates. *Detailed associations at the *PIK3R1* locus, which was primarily associated with lower body fat percentage, are presented in *Supplementary Figure 9*.

Figure 2. Associations at the 53 genomic loci with familial partial lipodystrophy type 1 (FPLD1). *Panel A*: distribution of leg fat mass as a function of the fat mass of the rest of the body (from DEXA) in women of the Fenland study at the extreme quintiles (Q) of the 53-SNP genetic score and in 14 FPLD1 subjects. Q1 represents a low genetic burden, whereas Q5 a high genetic burden. Lines of fit are plotted for each group. *Panel B*: histograms of the distribution of risk alleles in the FPLD1 subjects and in control women from the UKHLS study. *Panel C*: bi-dimensional box plots of the distribution of leg fat mass as a function of the distribution of the number of risk alleles in women of the Fenland study at the extreme quintiles (Q) of the 53-SNP genetic score and in FPLD1 subjects. Q1 represents a low genetic burden, whereas Q5 a high genetic burden. Each rectangle represents a group of individuals. For each dimension, the two sides of the rectangle represent the interquartile range and the central line the median. Data for obese women from Fenland were plotted to show the relationship between genetic risk and levels of leg fat in a group of women with a similar body mass index to that of FPLD1 patients.

Figure 3. Putative effector genes, tissues and cell types. *Panel A*: schematic representation of some established components of the insulin signalling pathway with stars reporting the location in the pathway of putative effector genes, with their respective lead single nucleotide polymorphism listed. *Panel B*: associations of gain- and loss-of-function genetic variants in the *LPL* gene with type 2 diabetes. The reference number in parenthesis refers to the study reporting the association with triglycerides and coronary heart disease (see reference number 38 of this manuscript).[38] *Panel C*: summary of evidence about links between genetic variants, lipodystrophy, insulin resistance, and type 2 diabetes at different levels of the population phenotypic distribution. *Rare syndromes caused by autosomal dominant *INSR* mutations are not usually associated with lipodystrophy and the *INSR* rs8101064 polymorphism is not associated with body fat percentage. *Panel D*: overlap of the 53 loci (lead SNPs plus proxy variants in $r^2>0.8$) with chromatin state annotations from the NIH Roadmap. *Panel E*: DEPICT's annotation of cell types and tissues on the basis of expression patterns in 37,427 human microarray samples. The y-axis represents the –log10(p-value) for enrichment of signal in a cell or tissue type attributed by DEPICT. The horizontal broken line represents the multiple-test corrected threshold of statistical significance (Bonferroni p=0.00072).

Figure 4. Experimental knockdown of putative effector genes in cellular adipogenesis models and comparisons with phenotypic associations. *Panel A:* results of experimental knockdown in OP9-K cells. Full circles represent the difference of the means from knockout experiments of a given gene compared with control experiments (n=4-7). Error bars represent the 95% confidence intervals of the difference of the means. Top graph: effect on mRNA levels of knockdown experiments of target genes using short interfering RNA (siRNA) in OP9-

K cells. The two-tailed t-test p-values for differences in means were: *IRS1*, p=4.6 x 10^{-06}; *CCDC92*, p=2.7 x 10^{-09}, *DNAH10,* p=2.4 x 10^{-06}; *L3MBTL3*, p=4.6 x 10^{-06}; *FAM13A*, p=2.4 x 10^{-05}. Bottom graph: effect on lipid accumulation in siRNA knockdown experiments. The two-tailed t-test p-values for differences in means were: *IRS1,* p=0.0047; *CCDC92,* p=1.2 x 10^{-05}*, DNAH10,* p=0.0027*; L3MBTL3,* p=0.00013*; FAM13A,* p=0.92*. Panel B:* illustrative images showing florescence microscopy from lipid accumulation experiments. Red indicates adipored staining of neutral lipid, blue is hoechst staining of nuclei. *Panel C:* Association of the risk (insulin-raising) allele of the lead single nucleotide polymorphism in or near each of the putative effector genes with (a) the expression of the corresponding gene in subcutaneous adipocytes in the EUROBATS project (top graph in the panel); (b) hip circumference in a meta-analysis of GIANT and UK Biobank (mid graph); and (c) type 2 diabetes in a meta-analysis of InterAct, DIAGRAM and UK Biobank (bottom graph). Full circles represent the –log10(p-value) for the association of the insulin-raising allele multiplied by the direction of the beta coefficient (i.e. a "directional" –log10(p)). For graphic display purposes, the –log10(p-value) for the association with type 2 diabetes of the rs2943645-T allele near *IRS1* is represented as 10 instead of 16.9.

Tables

Table 1. Association with "gold-standard" insulin resistance measures, type 2 diabetes and coronary heart disease of genetic scores comprising lead polymorphisms at identified loci. Results are displayed for genetic scores comprising either (a) the lead SNP at each of the 53 associated loci or (b) the lead SNP at each of the 43 additional loci identified in this study after removing 10 previously implicated in insulin resistance or (c) the lead SNP at each of the 28 loci not previously associated with levels of HDL cholesterol or triglycerides.

Exposure	Outcome	Sample size, N or N cases and N controls	Beta or odds ratio	SE or 95% CI	p-value
Association with "gold-standard" measures of insulin sensitivity					
53-SNP score	Insulin sensitivity[a]	2,764	-0.09	0.019	4.3×10^{-06}
43-SNP score			-0.08	0.022	4.6×10^{-04}
28-SNP score			-0.09	0.022	2.6×10^{-05}
53-SNP score	Insulin sensitivity index[b]	4,769	-0.10	0.016	7.3×10^{-10}
43-SNP score			-0.08	0.018	4.6×10^{-05}
28-SNP score			-0.09	0.027	0.0010
Association with disease endpoints					
53-SNP score	Type 2 diabetes	45,836 cases and 230,358 controls	1.12	1.11, 1.14	9.2×10^{-61}
43-SNP score			1.09	1.08, 1.11	7.6×10^{-29}
28-SNP score			1.11	1.09, 1.13	1.9×10^{-25}
53-SNP score	Coronary heart disease	63,746 cases and 130,681 controls	1.05	1.04, 1.06	1.8×10^{-13}
43-SNP score			1.04	1.03, 1.06	5.7×10^{-08}
28-SNP score			1.04	1.02, 1.06	1.2×10^{-05}

Abbreviations: N, number of participants; SE, standard error; CI, confidence interval; SNP, single nucleotide polymorphism. Beta coefficients are in standardised units per standard deviation of the 53-SNP genetic score (i.e. 4.5 alleles); odds ratios are per standard deviation of the 53-SNP genetic score (i.e. 4.5 alleles). The association with insulin sensitivity is from 2,764 participants of the GENESIS consortium[27] and the association with the insulin sensitivity index is from 4,769 participants of the MAGIC consortium who underwent a frequently sampled oral glucose tolerance test (OGTT)[33]; the association with type 2 diabetes is from the InterAct, DIAGRAM and UK Biobank studies; the association with coronary heart disease is from the CARDIoGRAM and the C4D consortia.

a In MAGIC, insulin sensitivity index (ISI) = $10{,}000/\sqrt{}$ (fasting plasma glucose (mg/dl)×fasting insulin×mean glucose during OGTT (mg/dl)×mean insulin during OGTT).[33]

b In GENESIS, insulin sensitivity was measured by clamp or insulin suppression test using study-specific parameters (e.g. glucose disposal or M-value) which were then standardised before meta-analysis.[27]

Online Methods

Study design

We integrated the results of genome-wide analyses on insulin and lipid phenotypes with the aim of identifying genetic variants associated with an insulin resistance phenotypic pattern (**Supplementary Figures 1-4**). We then investigated the mechanistic links of genetic variation at 53 identified genomic regions with cardiometabolic diseases by integrating analyses of: (a) cardiometabolic traits and outcomes from up to 451,193 individuals; (b) detailed continuous metabolic traits from 12,848 deeply-phenotyped individuals; (c) genetic and clinical features from 37 individuals diagnosed with familial partial lipodystrophy type 1; (d) gene expression from over 100 separate eQTL datasets and (e) siRNA mediated knockdown of putative effector genes in experimental adipogenesis models (**Supplementary Table 1 and Supplementary Figures 1-2**).

Participating studies

Lists of phenotypes, participating studies and sample sizes for each analysis are in **Supplementary Table 1 and Supplementary Figures 1-2**. Details about participants and cohorts with individual-level genotype data are in **Supplementary Table 14**. Ethical approvals were obtained at each study site and informed consent was obtained from all participants.

The Fenland study is a population-based cohort study of 12,435 participants without diabetes born between 1950 and 1975. Participants were recruited from general practice surgeries in Cambridge, Ely and Wisbech (United Kingdom) and underwent detailed metabolic phenotyping and genome-wide genotyping.

EPIC-Norfolk is a prospective cohort study of 25,639 individuals aged between 40 and 79 and living in the Norfolk county in the United Kingdom[78] at recruitment. EPIC-Norfolk is a constituent cohort of the European Prospective Investigation of Cancer (EPIC).[79] A total of

3,101 participants with available dual energy X-ray absorptiometry (DEXA) were included in analyses of detailed anthropometry, while 9,150 participants were included in analyses of change in hip or waist circumference in individuals who gained weight during follow-up.

EPIC-InterAct is a case-cohort study nested within the EPIC study, a cohort study of 519,978 European participants.[80] During an average of 8 years of follow-up, 12,403 individuals who were free of diabetes at baseline were identified incident type 2 diabetes cases.[80] InterAct has also defined a randomly-selected subcohort of 16,154 individuals free of diabetes at baseline.[80] Data on 15,357 individuals with available genotyping and not overlapping with DIAGRAM[34] were included.

UK Biobank is a population-based cohort study of ~500,000 people aged between 40-69 years who were recruited in 2006-2010 from several centres across the United Kingdom.[81] Associations with prevalent type 2 diabetes were estimated in 111,016 individuals (4,586 cases and 106,430 controls) of the initial UK Biobank dataset. We also used the UK Biobank data for anthropometry analyses and for a sensitivity analysis of prevalent coronary heart disease (i.e. self-reported myocardial infarction or angina) in 5,369 cases and 106,969 controls.

The United Kingdom Household Longitudinal Study (UKHLS; also known as Understanding Society) is a longitudinal panel survey of 40,000 households representative of the population of the United Kingdom. Participants were surveyed annually since 2009 and contributed information relating to their socioeconomic circumstances, attitudes, and behaviours via a computer assisted interview. For a subset of individuals who took part in a nurse health assessment, blood samples were taken and genomic DNA analysed.

In addition to individual-level genotyping data, we used genome-wide meta-analyses results on a variety of cardiometabolic traits and disease endpoints (**Supplementary Table 1 and Supplementary Figures 1-2**).

Detailed anthropometric analyses

In the Fenland and EPIC-Norfolk studies, body composition was determined by dual-energy X-ray absorptiometry (DEXA) using a Lunar Prodigy advanced fan beam scanner (GE Healthcare, Bedford, UK) using the encore software version 14.10.022 (GE Healthcare, Bedford UK). Participants were scanned by trained operators using standard imaging and positioning protocols. The coefficient of variation for scanning precision, calculated from 30 consecutive scans, was 2% for total fat mass. The enCORE software was used to demarcate regional boundaries. All the images were manually processed by one trained researcher, who corrected demarcations according to a standardized procedure. Boundaries of body regions are described in details in the **Supplementary Note**. In the UK Biobank study, body fat percentage was estimated by bio-impedance using the Tanita BC418MA body composition analyser.

Association of genetic variants with insulin resistance phenotypes

A dyslipidaemic pattern with higher triglyceride levels and lower HDL cholesterol is considered characteristic of the clinical presentation of insulin resistance [13] and has been used to specifically identify individuals with insulin resistance.[30] In a previous large-scale genome-wide discovery of genetic determinants of fasting insulin levels, among 19 fasting insulin-associated loci we identified ten that were strongly associated with higher triglycerides and lower HDL cholesterol.[28] This pattern of association was used to refine loci which were then validated as being associated with "gold-standard" measures of insulin resistance.[31] Loci associated with insulinaemia but not with lipid traits included *TCF7L2*,[28] which is primarily implicated in insulin secretion, rather than resistance.[82] These findings suggested that the combined association with this triad of phenotypes could help identify specific genetic determinants of insulin resistance.

With this background, we systematically triangulated the results of the association of ~2.4M single nucleotide polymorphisms (SNPs) with fasting insulin adjusted for body mass index (FIadjBMI; from up to 108,557 participants of the MAGIC consortium),[28,29] HDL-cholesterol, and triglycerides (from up to 188,577 participants of the Global Lipids Genetics Consortium)[32] using publicly available genome-wide results (**Supplementary Figures 1-4)**. For FIadjBMI analyses, we used metabochip association results[28] when available. We aligned alleles across the three phenotypes such that the effect allele was the insulin-raising allele. We took forward for further analysis all SNPs associated with higher FIadjBMI, higher triglycerides and lower HDL cholesterol at $p<0.005$ for each of the three traits. The prior probability for association of a given SNP with the three traits and in the pre-specified directional concordance under the null hypothesis corresponds to $0.005*0.0025*0.0025 = 3.1 \times 10^{-08}$. In each 1 Mb locus, we retained the lead SNP for association with fasting insulin for further analysis (**Supplementary Figure 4**).

Fasting insulin analyses adjusted for BMI were chosen because we were interested in identifying genetic determinants of insulin resistance for a given level of adiposity. It has been proposed that adjusting genetic association analyses for covariates such as BMI might result in a bias known as "collider bias".[83,84] Therefore, we assessed the association of the 53 lead SNPs for a bias in the identification of variants primarily associated with BMI and artificially associated with fasting insulin, but found no evidence on such bias (**Supplementary Figure 12**).

Statistical methods

We studied the association of individual SNPs and of genetic scores with continuous metabolic traits and endpoints. Where individual level genotype data was available, the associations of genetic variants or scores (i.e. the sum of effect alleles) with outcomes were

estimated using multivariable linear, logistic or Cox regression models. For result-level association data, we used the inverse-variance weighted method developed by Burgess et al., assigning a weight of 1 to each SNP, to approximate the association of an unweighted genetic risk score.[85] Statistical analyses were conducted using STATA v14.1 (StataCorp, College Station, Texas 77845 USA), R v3.2.2 (The R Foundation for Statistical Computing), and METAL.[29] All p-values presented in the study are two-tailed p-values.

Analyses in a severe form of partial lipodystrophy and insulin resistance

We studied the clinical and genetic characteristics of 37 women with familial partial lipodystrophy type 1 (FPLD1; also called Köbberling-type familial partial lipodystrophy).

All cases were referred to the insulin resistance/lipodystrophy specialist centre led by Drs Semple, Savage and O'Rahilly in Cambridge. FPLD1 is currently a clinical diagnosis used to describe predominantly women with selective paucity of limb adipose tissue, central obesity, severe insulin resistance, and a higher risk of type 2 diabetes.[86,87] As some women with lipodystrophy due to loss-of-function mutations in *PPARG* manifest a similar phenotype (i.e. FPLD3), mutations in this gene were excluded in all FPLD1 cases included in this study. The biochemical and anthropometric phenotype of the 37 FPLD1 patients was compared with that of female participants of the Fenland study. In these analyses we compared the phenotypes of the 37 FPLD1 patients with those of all Fenland study women (**Figure 2**) and with those of obese Fenland study women (**Figure 2 and Supplementary Table 8**), who have a similar BMI to that of FPLD1 women.

To understand the genetic basis of FPLD1, we carried out exome sequencing analyses in 18 individuals from 9 pedigrees with FPLD1 without identifying any clear candidate mutations or genes. Sequencing, variant calling and annotation were performed as described previously (Family 2;[88] Families 1 and 3–9 as part of the UK10K project).[89] Calls were annotated with

1000 Genomes allele frequencies and the NCBI dbSNP database build 132 (ftp://ftp.ncbi.nih.gov/snp/organisms/human_9606/). Variants were defined as potentially functional if they were non-synonymous, resulted in loss or gain of a stop codon or a frameshift, or occurred within essential splice sites. Those unlikely to have a functional impact were removed, as were all variants found with a MAF >1% in individuals of European descent from the 1000 Genomes Phase 1v3 (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20110521/) or the NHLBI ESP Exomes (URL: http://evs.gs.washington.edu/EVS/; January 2012). Further filtering was then implemented to retain only variants in genes seen in multiple patients.

We compared the burden of the 53-SNP genetic score in the 37 FPLD1 patients with that of 5,296 unrelated control women from UKHLS. Genome-wide genotyping of UKHLS women was performed using the Illumina Infinium HumanCoreExome-12 v1.0 BeadChip. Genotype calling was performed using the Illumina GenCall software. Genome-wide genotyping of FPLD1 patients was performed using the Illumina® Infinium CoreExome-24v1.0 chip. Prior to imputation, the following quality control criteria were applied for exclusion of SNPs in PLINK[90] (version 1.07): (1) minor allele frequency <0.01; (2) Hardy-Weinberg equilibrium $p<1 \times 10^{-06}$; (3) call rate <99%; (4) differential missingness between cases and controls $p<1 \times 10^{-06}$; (5) SNPs showing differential genotyping between the CoreExome-24v1.0 and CoreExome-12v1.0 chips. Samples were excluded prior to imputation in PLINK based on the following criteria: (1) call rate <95%; (2) autosomal heterozygosity >3 standard deviations from the mean; (3) pairwise identity by descent was calculated and one individual was removed for every pair of individuals with a pi-hat >0.05, with preference given to retaining female samples; (4) identity assessed by the concordance between the genome-wide and Fluidigm® genotypes at 24 sites (excluding individuals with concordance <90%); (5) ethnic outliers based on a principal components analysis. Imputation was performed using the 1000 Genomes Phase 3 reference panel using SHAPEIT2[91] (version 2.r644) and IMPUTE2[92]

(version 2.3.1). FPLD1 cases and UKHLS controls were imputed together. In an analysis of the genetic principal components derived from genome-wide genotyping (defined based on the combined sample), the 37 FPLD1 patients clustered with UKHLS control women (**Supplementary Figure 13**). Association analyses were performed in R by logistic regression adjusting for age and the first 10 genetic principal components. We also derived a permutation-based null comparator by performing 100,000 permutations of randomly selecting 53 SNPs (>1Mb apart) from genome-wide analyses of FPLD1 status adjusted for age and principal components and performing summary statistic Mendelian randomisation.[85] Of 100,000 iterations, 201 had a p-value less than our observed association ($p_{permutation}$= 2.01 x 10^{-03}).

Prioritisation of putative effector genes

We sought to determine the putative effector genes at the 53 loci. We combined information on (a) physical proximity, (b) eQTL data from over 100 repositories, (c) functional and regulatory annotations and (d) previous knowledge about causal genes for monogenic forms of insulin resistance.

For physical proximity analysis, we reviewed genes within a 1 Mb-window of each lead SNP and generated regional association plots using LocusZoom.[93] For eQTL analysis, we analysed both publicly available and unpublished datasets (see below). For functional annotation, we used the gene and the tissue/cell type prioritisation functions of the integrative software DEPICT,[44] in order to gain insights about putative effector genes, tissues and cell types. We also looked for nonsynonymous variants in linkage disequilibrium with the lead SNP (r^2>0.8) in European ancestry populations using Haploreg v3.[94]

We assessed the overlap of identified loci with chromatin state definitions of active enhancers and active promoters for 98 cell types from the NIH Epigenome Roadmap project, including a total of 196 genomic annotations. For each of the 53 lead SNPs, we identified the

set of proxy variants in high linkage disequilibrium (CEU r^2>0.8) using SNAP[95] and defined a 'locus' as a lead SNP plus its proxies. We then calculated the number of loci where at least one variant at the locus overlapped a given annotation. We determined enrichment in the overlap at the 53 loci compared to an expected distribution built from randomly selected matched loci. First, we identified all variants with genome-wide significant association ($p<5 \times 10^{-08}$) to any trait in Europeans from the GWAS catalog.[96] We then pruned this set of variants (using a CEU r^2 threshold of 0.1), resulting in a list of independent trait-associated variants. For each of these variants, we constructed a background 'locus' as the set of proxy variants in high linkage disequilibrium (CEU r^2<0.8). For each of the 53 loci, we then selected a locus from the background set matched on total number of proxy variants, total genomic distance covered, and distance of the midpoint to the closest gene transcription start site. We then re-calculated the overlap of each annotation in the set of matched background loci. We obtained the expected overlap for each annotation by averaging over 1 million permuted background locus sets. We then tested for enrichment of each annotation with a binomial test using the observed number of overlapping loci, total number of loci and expected percentage of overlapping loci.

We compiled a list of 13 genes (*PPARG, INSR, PIK3R1, TBC1D4, LMNA, PLIN1, AKT2, CIDEC, AGPAT2, BSCL2, CAV1, PTRF, PCYT1A*) known to cause monogenic forms of insulin resistance from the literature[13] and used that to look for overlap with genes in identified regions. Two experts in the clinical care of patients with monogenic insulin resistance (Drs Semple and Savage) reviewed the curated list.

Analysis of eQTLs in multiple tissues

Using a curated collection of over 100 separate eQTL datasets, we queried whether the 53 lead SNPs or their proxies (r^2>0.8) were associated with transcript expression in a wide range of tissues. Proxy SNPs in linkage disequilibrium in European ancestry populations were

identified using SNAP.[95] For this study, we considered all associations below a p-value cut-off of 1 x 10^{-06}. A general overview of a subgroup of >50 eQTL datasets has been published,[97] with specific citations for the >100 datasets included in the current query provided in the **Supplementary Note**.

Specific analysis of eQTLs in subcutaneous adipose tissue

We analysed in depth the association of lead SNPs with gene expression in subcutaneous adipose tissue using two subcutaneous adipose tissue cis-eQTL datasets. The first dataset was generated by the EUROBATS consortium and consists of samples from well-phenotyped healthy female twins (n=766) with eQTLs derived as described previously.[98] We also used the subcutaneous adipose tissue data (n=298) generated by the GTEx consortium (version 6),[99] which were obtained from www.gtexportal.org on 20/11/15. GTEx results were limited to GENCODE "protein_coding" and "lincRNA" biotype transcripts, and only variants with a minor allele frequency >0.01 were used. Linkage disequilibrium statistics between the index SNP (lead SNP for fasting insulin at the locus) and the most significant expression SNP for the gene were calculated in PLINK 1.9 using 1000 Genomes phase 1 version 3 European ancestry samples.[90] We also assessed the regulatory trait concordance (RTC) value for SNPs associated with gene expression in adipose tissue, in order to assess the likelihood of co-localisation of signal with the lead eQTL signal at that region.[100] In brief, if the index variant and the eQTL do tag the same causal variant, it is expected that removing the genetic effect of the index variant will have a significant consequence on the eQTL association. To this end, the RTC method assesses the likelihood of a shared functional effect between a GWAS variant and an eQTL by quantifying the change in the statistical significance of the eQTL after correcting for the genetic effect of the index variant and comparing its correction impact to that of all other SNPs in the interval. We considered an RTC of ≥ 0.8 or high linkage disequilibrium between

the lead eQTL SNP and trait-associated SNP (r^2>0.8) to be supporting evidence of co-localisation.

Functional studies in mouse OP9-K *cells*

We sought to experimentally validate candidate causal genes at loci associated with lower levels of peripheral adiposity, higher risk of type 2 diabetes and with gene expression in subcutaneous adipose tissue (**Supplementary Note and Supplementary Figure 11**). We studied the effects of gene knockdown using siRNA on adipogenic differentiation in murine OP9-K cell lines. The mouse OP9-K cell line used in this study is a model suitable for mid-throughput screening of genes influencing adipogenesis. OP9-K cells are clonal cells derived from mouse stromal OP9 cells obtained from the bone marrow, that accumulate large lipid-droplets after 72 hours of adipogenic stimulation.[45] OP9-K cells were grown and differentiated using oleic acid-containing differentiation media as described previously.[45] For siRNA transfections, 2.5 x 10^4 cells per well were cultured in the 24-well dish. After 24hrs, cells were transfected with smartpool siRNA (from Dharmacon) against each gene using Optifect reagent as per the manufacturer protocol. On the following day, differentiation of OP9-K cells into adipocytes was initiated by replacing the media with 500 µL of differentiation media. After 48hrs of differentiation induction, differentiated cells were stained with adipored to assess lipid accumulation using fluorescent spectroscopy. For quantitative-PCR, total RNA was isolated from differentiated OP9-K cells and cDNA was synthesized using TaqMan® Fast Cells-to-CT™ Kit (Applied Biosystems) according to manufacturer's instructions. Quantitative real-time PCR analysis was performed on a TaqMan ABI Prism 7900 Sequence Detector System (Applied Biosystems). Expression results were analysed relative to GAPDH mRNA content in the same sample.

1

Online-methods References

78 Day, N. *et al.* EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *British journal of cancer* **80 Suppl 1**, 95-103 (1999).

79 Riboli, E. *et al.* European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public health nutrition* **5**, 1113-1124, doi:10.1079/PHN2002394 (2002).

80 InterAct Consortium *et al.* Design and cohort description of the InterAct Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type 2 diabetes in the EPIC Study. *Diabetologia* **54**, 2272-2282, doi:10.1007/s00125-011-2182-9 (2011).

81 Collins, R. What makes UK Biobank special? *Lancet* **379**, 1173-1174, doi:10.1016/S0140-6736(12)60404-8 (2012).

82 Lyssenko, V. *et al.* Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *The Journal of clinical investigation* **117**, 2155-2163, doi:10.1172/JCI30706 (2007).

83 Aschard, H., Vilhjalmsson, B. J., Joshi, A. D., Price, A. L. & Kraft, P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *American journal of human genetics* **96**, 329-339, doi:10.1016/j.ajhg.2014.12.021 (2015).

84 Day, F. R., Loh, P. R., Scott, R. A., Ong, K. K. & Perry, J. R. A Robust Example of Collider Bias in a Genetic Association Study. *American journal of human genetics* **98**, 392-393, doi:10.1016/j.ajhg.2015.12.019 (2016).

85 Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic epidemiology* **37**, 658-665, doi:10.1002/gepi.21758 (2013).

86 Lawrence, R. D. Types of human diabetes. *British medical journal* **1**, 373-375 (1951).

87 Herbst, K. L. *et al.* Kobberling type of familial partial lipodystrophy: an underrecognized syndrome. *Diabetes care* **26**, 1819-1824 (2003).

88 Payne, F. *et al.* Hypomorphism in human NSMCE2 linked to primordial dwarfism and insulin resistance. *The Journal of clinical investigation* **124**, 4028-4038, doi:10.1172/JCI73264 (2014).

89 UK10K Consortium. The UK10K project identifies rare variants in health and disease. *Nature* **526**, 82-90, doi:10.1038/nature14962 (2015).

90 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* **81**, 559-575, doi:10.1086/519795 (2007).

91 Delaneau, O., Marchini, J. & Zagury, J. F. A linear complexity phasing method for thousands of genomes. *Nature methods* **9**, 179-181, doi:10.1038/nmeth.1785 (2012).

92 Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature genetics* **39**, 906-913, doi:10.1038/ng2088 (2007).

93 Pruim, R. J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-2337, doi:10.1093/bioinformatics/btq419 (2010).

94 Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic acids research* **40**, D930-934, doi:10.1093/nar/gkr917 (2012).

95 Johnson, A. D. *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24**, 2938-2939, doi:10.1093/bioinformatics/btn564 (2008).

96 Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic acids research* **42**, D1001-1006, doi:10.1093/nar/gkt1229 (2014).

97 Zhang, X. *et al.* Synthesis of 53 tissue and cell line expression QTL datasets reveals master eQTLs. *BMC genomics* **15**, 532, doi:10.1186/1471-2164-15-532 (2014).

98 Buil, A. *et al.* Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. *Nature genetics* **47**, 88-91, doi:10.1038/ng.3162 (2015).

99 GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-660, doi:10.1126/science.1262110 (2015).

100 Nica, A. C. *et al.* Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. *PLoS genetics* **6**, e1000895, doi:10.1371/journal.pgen.1000895 (2010).

Online Supplement

Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance

Supplementary Note

Associations at the 53 loci with changes in hip circumference in weight gainers

The association with lower levels of peripheral fat mass but higher cardiometabolic risk of the 53-SNP genetic score suggested that individuals with a greater number of alleles are unable to expand their peripheral fat compartment. To corroborate this finding, we used weight gain as a surrogate measure for a positive energy balance and change in hip circumference in weight-gainers as a surrogate measure for the changes in peripheral fat compartments. We tested the associations of the 53-SNP genetic score in longitudinal data from 9,150 participants of the EPIC-Norfolk cohort study who gained weight during a median follow-up of 3.7 years. In these individuals, the 53-SNP genetic score was not associated with the amount of weight gained during follow-up (beta coefficient [standard error] in kg of weight change per SD of genetic score, -0.026 [0.029]; p=0.37). However, in analyses adjusted for age, sex, hip circumference at baseline and weight at baseline and the amount of weight gained during follow-up, the 53-SNP genetic score was negatively associated with change in hip circumference (beta coefficient [standard error] in cm of hip circumference change per SD of genetic score, -0.069 [0.031]; p=0.027). In the same participants, in analyses adjusted for age, sex, waist circumference at baseline and weight at baseline and the amount of weight gained during follow-up, the 53-SNP genetic score was not associated with the change in waist circumference during follow-up (beta coefficient [standard error] in cm of waist circumference change per SD of genetic score, 0.055 [0.042]; p=0.20). These results support the notion that individuals with greater burden of the 53 alleles have a relative incapacity of expanding their peripheral fat compartment when challenged by a positive energy balance.

Selection of putative effector genes for experimental validation

In light of (a) the enrichment for loci overlapping adipose tissue active enhancer elements and affecting adipocyte gene expression and (b) the association of risk alleles with lower peripheral adiposity, but higher cardiometabolic risk, we hypothesised that some of the risk alleles may act via impaired adipogenesis. We further hypothesised that the effects on adipogenesis could be caused by the altered expression of an effector gene in peripheral adipose tissue (**Supplementary Figure 11A**). Therefore, to test this hypothesis, we selected five genes at four loci associated with (a) expression of a putative effector gene in subcutaneous adipocytes ($p<5 \times 10^{-08}$), (b) lower levels of peripheral fat ($p<5 \times 10^{-05}$ for hip circumference) and (c) higher risk of metabolic disease ($p<0.05$ for type 2 diabetes; see **Supplementary Tables 10 and 13 and Supplementary Figure 11** for details). For the *IRS1* (RTC=0.86) and *L3MBTL3* genes (r^2 between lead and best expression SNPs=0.83) there was evidence supporting co-localisation of phenotypic and expression signals. For the *CCDC92*, *DNAH10* and *FAM13A* genes, the lead expression SNPs (eSNPs) at the locus (rs825452 for *CCDC92*, rs78985577 for *DNAH10* and rs13149209 for *FAM13A*) were not captured by the HapMap-imputed FIadjBMI association data,[21, 22] meaning they could not be captured by our triangulation of fasting insulin and lipid data. However, the best HapMap proxy for each of those eSNPs was also associated with FIadjBMI in MAGIC,[1,2] further supporting our prioritisation of those genes (see below).

For *CCDC92*, the lead eSNP (rs825452, $p_{expression}=8.3 \times 10^{-31}$) was in very low linkage disequilibrium ($r^2=0.001$) with our lead SNP for association with FIadjBMI (rs7973683). However, rs7973683 was also strongly associated with *CCDC92* expression in adipocytes ($p_{expression}=2.1 \times 10^{-29}$), indicative of two distinct signals of association with *CCDC92* expression levels. Furthermore, while the lead eSNP was not available in FIadjBMI results, a strong proxy (rs825453; $r^2=1$) for the lead eSNP was also associated with FIadjBMI

(p=0.0053). In the same locus, we found that our lead SNP for association with FIadjBMI levels (rs7973683) was also associated with expression of *DNAH10* ($p_{expression}$=1.9 x 10^{-08}). This was in modest linkage disequilibrium (r^2=0.27) with the lead eSNP for *DNAH10* expression (rs78985577, $p_{expression}$=4.8 x 10^{-12}). While the lead eSNP was not available in FIadjBMI data, a strong proxy (rs1316952; r^2=0.83) was also associated with FIadjBMI levels (p=0.000086). At *FAM13A*, our lead SNP for association with FIadjBMI levels (rs3822072) was also associated with expression of *FAM13A* in subcutaneous adipocytes ($p_{expression}$=7.6 x 10^{-12}). Our lead SNP was in low linkage disequilibrium (r^2=0.038) with the lead eSNP (rs13149209, $p_{expression}$=4.5 x 10^{-21}), a modest proxy for which (rs2085600; r^2=0.72) was also associated with FIadjBMI levels (p=3.5 x 10^{-06}). These results suggest that multiple independent eQTLs of those genes in adipose tissue are also associated with insulin levels and therefore further support our prioritisation of these genes. Genes at all the loci showing the pre-specified pattern of association were studied experimentally, with the exception of *KLF14*. We did not seek to experimentally validate the *KLF14* gene, because it has been studied previously and previous studies suggest complex aetiologic mechanisms at this locus, including a potential parent-of-origin effect.[3] In dedicated figures and tables, we report association criteria (**Supplementary Figure 11B**), selection flow-chart (**Supplementary Figure 11C**), association estimates at loci with an eQTL signal in subcutaneous adipocytes (**Supplementary Table 10**) and at the prioritised loci (**Supplementary Table 13**). Loci that did not meet the criteria were not prioritised for experimental validation of putative effector genes (**Supplementary Figure 11 and Supplementary Table 10**). Finally, on the basis of our hypothesis, we expected that the siRNA knockdown of the candidate causal gene would have effects on adipogenesis in the direction predicted by the adipose tissue eQTL.

Whole and regional body composition analysis

Before scanning, the DEXA system was calibrated according to the manufacturer's guidelines using a spine phantom made of calcium hydroxyapatite, embedded in a lucite block. The enCORE software automatically demarcates the regional boundaries. A protocol was established to manually refine these demarcations and all the images were processed by one trained researcher, who corrected the demarcations according to a standardized procedure. The arm region was derived by positioning a line from the crease of the axilla and through the glenohumeral. The trunk region includes the neck, chest, abdominal and pelvic areas. The leg region includes all of the area below the lines that form the lower borders of the trunk. The android region was defined as the area between the ribs and the pelvis, and is enclosed by the trunk region. This region is outlined by iliac crest and with a superior height equivalent to 20% of the distance from the top of the iliac crest to the base of the skull. The gynoid region includes the hips and upper thighs, and overlaps both the leg and trunk regions. The upper demarcation is below the top of the iliac crest at a distance of 1.5 times the android height. The total height of the gynoid region is two times the height of the android region. Estimates of overall and regional body fat, lean and bone masses were derived using the DEXA software. The software also uses an inbuilt algorithm to determine visceral adipose tissue (in grams) within the android region. The subcutaneous abdominal adipose tissue (in grams) was calculated as android fat mass minus visceral abdominal adipose tissue.

List of sources for eQTL analyses

A general overview of a subset of eQTL datasets interrogated in this study has been published.[4] Specific citations for all >100 datasets included in the current query are provided below.

Tissues (PubMed ID): blood cell related eQTL studies included fresh lymphocytes (17873875), fresh leukocytes (19966804), leukocyte samples in individuals with Celiac disease (19128478), whole blood samples (18344981, 21829388, 22692066, 23818875, 23359819, 23880221, 24013639, 23157493, 23715323, 24092820, 24314549, 24956270, 24592274, 24728292, 24740359, 25609184, 22563384, 25474530, 25816334, 25578447), lymphoblastoid cell lines (LCL) derived from asthmatic children (17873877, 23345460), HapMap LCL from 3 populations (17873874), a separate study on HapMap CEU LCL (18193047), additional LCL population samples (19644074, 22286170, 22941192, 23755361, 23995691, 25010687, 25951796), neutrophils (26151758, 26259071), CD19+ B cells (22446964), primary PHA-stimulated T cells (19644074, 23755361), CD4+ T cells (20833654), peripheral blood monocytes (19222302,20502693,22446964, 23300628, 25951796, 26019233), long non-coding RNAs in monocytes (25025429) and CD14+ monocytes before and after stimulation with LPS or interferon-gamma (24604202), CD11+ dendritic cells before and after Mycobacterium tuberculosis infection (22233810) and a separate study of dendritic cells before or after stimulation with LPS, influenza or interferon-beta (24604203). Micro-RNA QTLs (21691150, 26020509), DNase-I QTLs (22307276), histone acetylation QTLs (25799442), and ribosomal occupancy QTLs (25657249) were also queried for LCL. Splicing QTLs (25685889) and micro-RNA QTLs (25791433) were queried in whole blood. Non-blood cell tissue eQTLs searched included omental and subcutaneous adipose (18344981, 21602305, 22941192, 23715323, 25578447), visceral fat (25578447) stomach (21602305), endometrial carcinomas (21226949), ER+ and ER- breast cancer tumor

cells (23374354), liver (18462017,21602305,21637794, 22006096, 24665059, 25578447), osteoblasts (19654370), intestine (23474282) and normal and cancerous colon (25079323, 25766683), skeletal muscle (24306210, 25578447), breast tissue (normal and cancer)(24388359, 22522925), lung (23209423, 23715323, 24307700, 23936167, 26102239), skin (21129726, 22941192, 23715323, 25951796), primary fibroblasts (19644074, 23755361, 24555846), sputum (21949713), pancreatic islet cells (25201977), prostate (25983244), rectal mucosa (25569741), arterial wall (25578447) and heart tissue from left ventricles (23715323, 24846176) and left and right atria (24177373). Micro-RNA QTLs were also queried for gluteal and abdominal adipose (22102887) and liver (23758991). Methylation QTLs were queried in pancreatic islet cells (25375650). Further mRNA and micro-RNA QTLs were queried from ER+ invasive breast cancer samples, colon-, kidney renal clear-, lung- and prostate-adenocarcinoma samples (24907074). Brain eQTL studies included brain cortex (19222302, 19361613, 22685416, 25609184, 25290266), cerebellar cortex (25174004), cerebellum (20485568, 22685416, 22212596, 22832957, 23622250), frontal cortex (20485568, 22832957, 25174004), gliomas (24607568), hippocampus (22832957, 25174004), inferior olivary nucleus (from medulla) (25174004), intralobular white matter (25174004), occiptal cortex (25174004), parietal lobe (22212596), pons (20485568), pre-frontal cortex (22031444, 20351726, 22832957, 23622250), putamen (at the level of anterior commussure) (25174004), substantia nigra (25174004), temporal cortex (20485568, 22685416, 22832957, 25174004), thalamus (22832957) and visual cortex (23622250).

Additional eQTL data was integrated from online sources including ScanDB, the Broad Institute's GTEx Portal, and the Pritchard Lab (eqtl.uchicago.edu). Cerebellum, parietal lobe and liver eQTL data was downloaded from ScanDB. Results for GTEx Analysis V4 for 13 tissues were downloaded from the GTEx Portal and then additionally filtered as described below (www.gtexportal.org: thyroid, leg skin [sun exposed], tibial nerve, aortic artery, tibial

artery, skeletal muscle, esophagus mucosa, esophagus muscularis, lung, heart (left ventricle), stomach, whole blood, and subcutaneous adipose [23715323]). Splicing QTL (sQTL) results generated with sQTLseeker with false discovery rate $p \leq 0.05$ were retained.

Supplementary Tables

Supplementary Table 1. Phenotypes, participating studies and maximum sample size.

Analysis	Phenotype	Participating studies (N; PMID)	Maximum sample size, N
Identification of 53 loci	FIadjBMI	MAGIC (N=108,557; PMID: 22885924, 22581228)	108,557
	HDL cholesterol	GLGC (N=188,577; PMID: 24097068)	188,577
	Triglycerides	GLGC (N=188,577; PMID: 24097068)	188,577
Validation of genetic scores	FIadjBMI	Fenland (N=4,694; this study)	4,694
	HDL cholesterol	Fenland (N=6,101; this study)	6,101
	Triglycerides	Fenland (N=6,101; this study)	6,101
	Insulin sensitivity index	MAGIC (N=4,769; PMID: 24699409)	4,769
	Insulin sensitivity	GENESIS (N=2,764; PMID: 25798622)	2,764
Association with intermediate traits	DEXA	Fenland (N=9,747; this study); EPIC-Norfolk (N=3,101; this study)	12,848
	Body fat percentage	UK Biobank (N=110,358; this study); Fenland (N=9,747; this study); EPIC-Norfolk (N=3,101; this study)	123,206
	BMI	GIANT (N=339,198; 25673413); UK Biobank (N=111,995; this study)	451,193
	Waist circumference	GIANT (N=244,419; 25673412); UK Biobank (N=112,180; this study)	356,599
	Hip circumference	GIANT (N=227,412; 25673412); UK Biobank (N=112,172; this study)	339,584
	Waist-to-hip ratio	GIANT (N=226,586; 25673412); UK Biobank (N=112,158; this study)	338,744
	Fasting plasma glucose	MAGIC (N=133,010; PMID: 22885924, 22581228)	133,010
	2 hour glucose	MAGIC (N=42,854; PMID: 22885924, 20081857)	42,854
	HbA1c	MAGIC (N=46,368; PMID: 20858683)	46,368
	Alanine aminotransferase and gamma glutamyl transferase	Fenland (N=10,330; this study)	10,330
Association with disease	Type 2 diabetes	DIAGRAM (cases=34,840; controls=114,981; PMID: 22885922); InterAct (cases=6,410; controls=8,947; this study); UK Biobank (cases=4,586; controls=106,430; this study)	45,836 cases 230,358 controls
	Coronary heart disease	CARDIoGRAMplusC4D (cases=63,746; controls=130,681; PMID: 21378988, 23202125, 21378990);	63,746 cases 130,681 controls
	FPLD1	Cambridge FPLD1 consortium (cases=37; this study); UKHLS (controls=5,296)	37 cases 5,296 controls

Supplementary Table 2. List of the 53 genomic regions associated with insulin resistance phenotypes.

SNP	Genomic coordinate	Alleles (effect / other)	Beta FIadjBMI per allele[a]	FIadjBMI p-value	Beta triglycerides per allele[b]	Triglycerides p-value	Beta HDL cholesterol per allele[b]	HDL cholesterol p-value	Locus name	Putative effector genes[c]
Loci previously implicated in insulin resistance										
rs4846565	Chr1:219722104	G / A	0.022	1.76E-09	0.014	0.00019	-0.013	0.00078	*RNU5F-1/LYPLAL1*	*RNU5F-1*[N]
rs10195252*	Chr2:165513091	T / C	0.029	1.26E-16	0.028	6.99E-15	-0.025	3.49E-11	*COBLL1/GRB14*	*COBLL1*[N], *GRB14*[E]
rs2943645*	Chr2:227099180	T / C	0.032	2.26E-19	0.028	3.76E-15	-0.032	4.16E-17	*IRS1*	*IRS1*[E, EA]
rs308971	Chr3:12116620	G / A	0.036	2.97E-11	0.021	3.51E-05	-0.016	0.0033	*SYN2/PPARG*	*SYN2*[N, E], *PPARG*[MF]
rs3822072*	Chr4:89741269	A / G	0.020	1.80E-08	0.018	5.74E-07	-0.025	4.06E-12	*FAM13A*	*FAM13A*[N, E, EA, D]
rs6822892	Chr4:157734675	A / G	0.024	2.58E-10	0.012	0.00084	-0.019	1.93E-07	*PDGFC*	*PDGFC*[N, E, EA, D]
rs4865796*	Chr5:53272664	A / G	0.025	2.16E-12	0.010	0.0030	-0.013	0.00030	*ARL15/FST*	*ARL15*[N], *FST*[EA]
rs459193*	Chr5:55806751	G / A	0.025	1.15E-10	0.018	1.31E-05	-0.024	8.10E-09	*ANKRD55*	*ANKRD55*[N]
rs2745353*	Chr6:127452935	T / C	0.019	4.10E-07	0.017	1.18E-06	-0.020	7.42E-10	*RSPO3*	*RSPO3*[N, E, EA]
rs731839*	Chr19:33899065	G / A	0.025	5.13E-12	0.022	2.65E-09	-0.022	3.44E-09	*PEPD*	*PEPD*[N, E]
Additional loci										
rs683135*	Chr1:39895460	A / G	0.014	0.00024	0.017	6.18E-07	-0.027	7.09E-12	*MACF1*	*MACF1*[N, NS, E, EA]
rs17386142	Chr1:50815783	C / T	0.024	0.00092	0.022	0.0033	-0.022	0.00060	*DMRTA2*	*DMRTA2*[N], *CDKN2C*[D]
rs11577194	Chr1:110500175	T / C	0.011	0.0025	0.011	0.0011	-0.019	1.34E-07	*CSF1*	*CSF1*[N]
rs9425291	Chr1:172312769	A / G	0.015	2.71E-05	0.016	1.9E-05	-0.014	0.00046	*DNM3*	*DNM3*[N, E], *PIGC*[E, EA]
rs2249105	Chr2:65287896	A / G	0.016	1.04E-05	0.016	2.35E-06	-0.016	0.00016	*CEP68*	*CEP68*[N, E, EA]
rs492400	Chr2:219349752	T / C	0.010	0.0038	0.018	2.25E-06	-0.011	0.0049	*USP37*	*USP37*[N, E], *ZNF142*[NS, E, EA]
rs3864041	Chr3:15185634	T / C	0.011	0.0038	0.009	0.0038	-0.013	0.00028	*COL6A4P1*	*COL6A4P1*[N]
rs295449*	Chr3:47375955	A / G	0.011	0.0020	0.014	0.00093	-0.019	0.00011	*KLHL18*	*KLHL18*[N, E], *SCAP*[NS, E], *SETD2*[E, D]
rs11130329*	Chr3:52896855	A / C	0.024	0.00051	0.020	0.0018	-0.024	0.0010	*TMEM110-MUSTN1*	*TMEM110-MUSTN1*[N]
rs9881942	Chr3:123082416	A / G	0.013	0.00014	0.010	0.0036	-0.015	4.79E-06	*ADCY5*	*ADCY5*[N]
rs645040*	Chr3:135926622	T / G	0.014	0.0012	0.029	1.83E-12	-0.031	1.53E-12	*MSL2*	*MSL2*[N, D]
rs2699429*	Chr4:3480136	C / T	0.011	0.0037	0.025	1.15E-11	-0.013	0.0042	*DOK7*	*DOK7*[N, E]
rs4976033	Chr5:67714246	G / A	0.015	0.00013	0.014	0.00020	-0.022	6.42E-08	*PIK3R1*	*PIK3R1*[N, MF, D]
rs6887914	Chr5:112711486	C / T	0.013	0.0037	0.014	0.0024	-0.017	0.00039	*MCC*	*MCC*[N]
rs1045241	Chr5:118729286	C / T	0.012	0.0020	0.015	4.06E-05	-0.014	0.00040	*TNFAIP8*	*TNFAIP8*[N, E, EA]
rs2434612	Chr5:158022041	G / A	0.016	0.00034	0.015	0.00025	-0.020	0.000015	*EBF1*	*EBF1*[N, D]
rs966544	Chr5:173350405	G / A	0.012	0.0010	0.016	1.63E-06	-0.013	0.0018	*CPEB4*	*CPEB4*[N, E]
rs12525532	Chr6:35004819	T / C	0.019	8.95E-08	0.011	0.0026	-0.015	0.000040	*ANKS1A*	*ANKS1A*[N, EA]
rs6937438*	Chr6:43815364	A / G	0.013	0.0011	0.014	0.00034	-0.019	1.94E-06	*LOC100132354*	*LOC100132354*[N]
rs9492443	Chr6:130398731	C / T	0.014	0.0004	0.016	7.43E-05	-0.013	0.0042	*L3MBTL3*	*L3MBTL3*[N, E, EA]
rs3861397*	Chr6:139828916	G / A	0.014	0.00011	0.024	1.08E-10	-0.024	8.40E-11	*LOC645434*	*LOC645434*[N], *CITED2*[D]
rs17169104	Chr7:15883727	G / C	0.020	1.52E-06	0.017	7.62E-05	-0.017	0.00028	*MEOX2*	*MEOX2*[N]

rs972283*	Chr7:130466854	G / A	0.022	4.41E-06	0.017	2.34E-07	-0.029	4.60E-16	*KLF14*	*KLF14*[N, E, EA]
rs2126259*	Chr8:9185146	T / C	0.041	3.30E-13	0.017	0.0020	-0.075	1.53E-42	*PPP1R3B*	*LOC157273*[N], *PPP1R3B*[E]
rs1011685*	Chr8:19830769	C / T	0.019	0.00098	0.168	6.12E-197	-0.156	8.65E-150	*LPL*	*LPL*[N, NS, E, D]
rs4738141	Chr8:72469742	G / A	0.014	0.0014	0.020	2.45E-05	-0.019	0.00030	*EYA1*	*EYA1*[N, EA], *LOC105375892*[D]
rs7005992*	Chr8:126528955	C / G	0.016	0.0014	0.021	5.87E-06	-0.016	0.0011	*TRIB1*	*TRIB1*[N, D]
rs498313	Chr9:78034169	A / G	0.013	0.00073	0.011	0.0026	-0.014	0.00033	*MIR548H3*	*MIR548H3*[N]
rs10995441*	Chr10:64869239	G / T	0.014	0.00085	0.017	1.19E-06	-0.018	0.00011	*NRBF2*	*NRBF2*[N, E]
rs11231693	Chr11:63862612	A / G	0.036	7.19E-07	0.030	0.00012	-0.029	0.000069	*MACROD1*	*MACROD1*[N]
rs17402950	Chr12:14571671	G / A	0.027	0.0047	0.032	0.0049	-0.034	0.0028	*ATF7IP*	*ATF7IP*[N]
rs718314	Chr12:26453283	G / A	0.017	3.65E-05	0.012	0.0015	-0.020	5.88E-06	*ITPR2*	*ITPR2*[N, E, D]
rs7973683*	Chr12:124449223	C / A	0.019	6.99E-07	0.025	4.67E-12	-0.029	5.26E-14	*CCDC92/DNAH10*	*CCDC92*[N, NS, E, EA], *DNAH10*[E, EA]
rs7323406	Chr13:111628195	A / G	0.015	0.0027	0.014	0.0044	-0.016	0.0032	*ANKRD10*	*ANKRD10*[N, E, D]
rs7176058	Chr15:39464167	A / G	0.013	0.0036	0.016	0.00094	-0.015	0.00028	*C15orf54*	*C15orf54*[N], *THBS1*[EA]
rs8032586	Chr15:73081067	C / T	0.019	0.0046	0.025	0.0010	-0.021	0.0030	*LOC100287559*	*LOC100287559*[N]
rs754814	Chr17:4657034	T / C	0.011	0.0042	0.011	0.0022	-0.013	0.0014	*ZMYND15*	*ZMYND15*[N, E]
rs7227237*	Chr18:47174679	C / T	0.017	0.0013	0.017	0.00088	-0.020	0.00041	*LIPG*	*LIPG*[N, E]
rs8101064*	Chr19:7293119	T / C	0.042	0.00062	0.069	1.91E-06	-0.066	0.000022	*INSR*	*INSR*[N, MF, D]
rs4804833*	Chr19:7970635	A / G	0.016	7.11E-06	0.015	9.90E-06	-0.022	9.89E-08	*MAP2K7*	*MAP2K7*[N, E, D]
rs4804311*	Chr19:8615589	A / G	0.019	0.0026	0.039	1.49E-09	-0.051	3.74E-14	*MYO1F*	*MYO1F*[N, E, EA]
rs6066149	Chr20:45602638	G / A	0.013	0.0019	0.018	5.22E-06	-0.010	0.0037	*EYA2*	*EYA2*[N]
rs132985*	Chr22:38563471	C / T	0.016	4.69E-06	0.022	6.65E-11	-0.015	0.000017	*PLA2G6*	*PLA2G6*[N], *MAFF*[E, EA, D]

Genomic coordinates refer to human genome build 37 (hg19). Beta coefficients are in standardised units, fasting insulin beta coefficients were standardised using the standard deviation in 8,917 participants of the Fenland study. The gene column reports the nearest gene and/or additional candidate effector genes at the locus.

*polymorphism within 500 kb of a lead SNP for HDL cholesterol or triglyceride levels reported by the Global Lipids Genetics Consortium (PubMed ID: 24097068).

a From up to 108,557 participants of the MAGIC consortium (PubMed ID: 22885924, 22581228)

b From up to 188,577 participants of the Global Lipids Genetics Consortium (PubMed ID: 24097068)

c Assigned on the basis of the following criteria: N, nearest gene; NS, nonsynonymous variant in linkage disequilibrium with lead SNP (r^2 >0.8); E, evidence of association with gene expression in surveyed eQTL repositories; AE, evidence of association with gene expression in subcutaneous adipose tissue; MF, monogenic insulin resistance forms associated with mutations in this gene; D, gene prioritised by DEPICT software as likely causal (significant p-value after accounting for false discovery rate). Relevant criteria are reported as superscript near each gene. Further details about methodology for the adjudication of these criteria are reported in the Online Methods sections dedicated to prioritisation of putative effector genes.

Abbreviations: SNP, single nucleotide polymorphism; FIadjBMI, fasting insulin adjusted for body mass index; HDL, high-density lipoprotein cholesterol.

Supplementary Table 3. Association with type 2 diabetes of the 53-polymorphism genetic score in analyses stratified by sex or body mass index. Results are scaled per 4.5 alleles, i.e. a standard deviation of genetic risk score. Results are from the EPIC-InterAct and the UK Biobank studies.

Stratum	Participants, type 2 diabetes cases / non-cases	OR (95% CI)	p-value	p-interaction
Sex-stratified analysis				
Men	6,588 / 52,887	1.12 (1.09 – 1.16)	4.13E-14	0.90
Women	5,418 / 62,811	1.12 (1.08 – 1.16)	1.48E-11	
*BMI-stratified analysis**				
BMI < 25	1,298 / 39,930	1.16 (1.09 – 1.23)	3.71E-06	0.16
BMI ≥ 25 and BMI < 30	4,663 / 49,317	1.17 (1.13 – 1.22)	1.73E-16	
BMI ≥ 30	5,945 / 26,090	1.12 (1.08 – 1.16)	7.86E-10	

Abbreviations: OR, odds ratio; CI, confidence interval; BMI, body mass index.
*Pairwise category heterogeneity tests: lean vs overweight: p=0.73; lean vs obese: p=0.31; overweight vs obese, p=0.061.

Supplementary Table 4. Associations of lead single nucleotide polymorphisms at the 53 loci with glycaemic, anthropometric traits and disease endpoints.
(see Supplementary Table Excel file)

Supplementary Table 5. Single nucleotide polymorphisms associated with higher risk of type 2 diabetes (45,836 cases 230,358 controls) and of coronary heart disease (63,746 cases 130,681 controls).

SNP	Locus name	Per allele OR type 2 diabetes (95% CI)	p-value	Per allele OR coronary heart disease (95% CI)	p-value
rs2943645	*IRS1*	1.09 (1.07-1.11)	1.1E-17	1.03 (1.01-1.05)	0.0010
rs6822892	*PDGFC*	1.04 (1.02-1.07)	2.1E-05	1.02 (1.00-1.04)	0.035
rs459193	*ANKRD55*	1.08 (1.06-1.11)	8.0E-13	1.02 (1.00-1.04)	0.025
rs4976033	*PIK3R1*	1.03 (1.01-1.05)	0.0022	1.07 (1.02-1.12)	0.0080
rs9492443	*L3MBTL3*	1.05 (1.03-1.07)	3.1E-05	1.02 (1.00-1.04)	0.028
rs3861397	*LOC645434*	1.03 (1.01-1.05)	0.0026	1.02 (1.00-1.04)	0.037
rs972283	*KLF14*	1.04 (1.02-1.06)	1.2E-05	1.03 (1.01-1.04)	0.0037
rs1011685	*LPL*	1.07 (1.04-1.10)	3.4E-05	1.09 (1.05-1.13)	8.7E-06
rs7973683	*CCDC92 / DNAH10*	1.03 (1.01-1.05)	0.0037	1.02 (1.00-1.04)	0.019
rs8101064	*INSR*	1.08 (1.01-1.16)	0.020	1.13 (1.02-1.25)	0.016
rs731839	*PEPD*	1.04 (1.02-1.06)	0.00021	1.03 (1.01-1.04)	0.0090

Abbreviations: SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval.

Supplementary Table 6. Association of the genetic scores with alanine aminotransferase and gamma glutamyl transferase in 10,330 participants of the Fenland study.

Exposure	Outcome	Beta per SD of genetic score in SDs of biomarker	SE	P-Value
53-SNP score	Alanine aminotransferase	0.054	0.009	1.37E-09
43-SNP score		0.049	0.010	1.83E-06
53-SNP score	Gamma glutamyl transferase	0.054	0.009	1.04E-09
43-SNP score		0.045	0.010	6.39E-06

Abbreviations: SNP, single nucleotide polymorphism; SD, standard deviation; SE, standard error. Beta coefficients are in standardised units per SD of genetic score (4.5 alleles).

Supplementary Table 7. European Genome-Phenome Archive Study, Dataset and Sample IDs for the raw, whole exome sequence data for 9 FPLD1 individuals and their family members.

Family ID	Sample	EGA Study ID	EGA Dataset ID	EGA Sample ID
1	Proband	EGAS00001000130	EGAD00001000419	EGAN00001015630
2	Proband	EGAS00001000025	EGAD00001000380	EGAN00001001252
2	Mother	EGAS00001000025	EGAD00001000380	EGAN00001001832
2	Father	EGAS00001000025	EGAD00001000380	EGAN00001001825
3	Proband	EGAS00001000130	EGAD00001000419	EGAN00001015627
4	Proband	EGAS00001000130	EGAD00001000419	EGAN00001015629
5	Proband	EGAS00001000130	EGAD00001000419	EGAN00001015628
6	Proband	EGAS00001000130	EGAD00001000419	EGAN00001015631
7	Proband	EGAS00001000130	EGAD00001000419	EGAN00001015624
7	Father	EGAS00001000130	EGAD00001000419	EGAN00001015625
8	Proband	EGAS00001000130	EGAD00001000419	EGAN00001015633
8	Father	EGAS00001000130	EGAD00001000419	EGAN00001015651
8	Sister 1	EGAS00001000130	EGAD00001000419	EGAN00001015652
8	Sister 2	EGAS00001000130	EGAD00001000419	EGAN00001015653
8	Sister 3	EGAS00001000130	EGAD00001000419	EGAN00001015654
8	Sister 4	EGAS00001000130	EGAD00001000419	EGAN00001015656
8	Brother	EGAS00001000130	EGAD00001000419	EGAN00001015655
9	Proband	EGAS00001000130	EGAD00001000419	EGAN00001015632

Supplementary Table 8. Characteristics of women with FPLD1 compared with obese women of the Fenland study. The *phenotype comparison* column summarises the results of comparisons between FPLD1 women and obese (BMI ≥30) Fenland study women for a given clinical variable (Student's t-test), whereas the *genetic score association pattern* column summarises the association of the 53-SNP genetic score with a given phenotype in our genetic association analyses (see Figure 1A and Supplementary Table 6).

	FPLD1 women			Fenland - obese women			Phenotype comparison, FPLD1 vs Fenland obese women, direction of association (p-value)[a]	Pattern of association of genetic score, direction of association (p-value)[a]
Variable, units	**N**	**Mean (SD)**	**Median (range)**	**N**	**Mean (SD)**	**Median (range)**		
Age, years	37	49 (12)	48 (22 - 75)	1171	49 (7)	50 (29 - 64)	N/A	N/A
Body mass index, kg/m^2	35	33 (4)	33 (26 - 47)	1171	35 (5)	33 (30 - 60)	N/A*	↓ (p=4.5E-08)
Waist circumference, cm	27	111 (12)	112 (78 - 143)	1170	103 (10)	102 (79 - 154)	↑ (p=4.6E-05)	↑ (p=0.0028)
Hip circumference, cm	27	108 (12)	108 (84 - 147)	1163	118 (10)	116 (96 - 178)	↓ (p=3.7E-07)	↓ (p=2.3E-34)
Waist-to-hip ratio	27	1.04 (0.09)	1.02 (0.91 - 1.22)	1163	0.87 (0.07)	0.87 (0.63 - 1.17)	↑ (p=3.1E-33)	↑ (p=3.3E-88)
Fasting plasma glucose, mmol/L	33	11 (5)	10.7 (4.9 - 21.9)	1163	5.0 (0.7)	4.9 (3.4 - 12.3)	↑ (p=8.5E-161)	↑ (p=2.9E-10)
HbA1c, %	35	9 (2)	8.4 (4.9 - 13.2)	789	5.6 (0.5)	5.6 (3.9 - 9.7)	↑ (p=6.9E-140)	↑ (p=2.3E-06)
Fasting insulin, pmol/L	29	208 (230)	151 (7.3 - 1210)	987	71.0 (49.4)	60.4 (2.6 - 702.0)	↑ (p=6.0E-30)	↑**
Triglycerides, mmol/L	34	2.8 (2.1)	2.6 (0.6 - 12.6)	1168	1.3 (0.7)	1.2 (0.2 - 7.4)	↑ (p=1.5E-27)	↑**
HDL cholesterol, mmol/L	34	1 (0.3)	1 (0.5- 2)	1168	1.5 (0.3)	1.4 (0.6 - 3.4)	↓ (p=5.4E-21)	↓**
Alanine aminotransferase, U/L	34	39 (25)	31 (13 - 126)	1168	29 (17)	25 (6 - 236)	↑ (p=0.00090)	↑ (p=1.4E-09)
Gamma glutamyl transferase, U/L	34	67 (51)	50 (12 - 212)	1168	35 (32)	27 (7 - 530)	↑ (p=2.2E-08)	↑ (p=1.0E-09)

a ↑ indicates associations (p<0.05) with higher levels of a given phenotype in FPLD1 (compared with obese women from the Fenland study) or for a greater number of risk alleles of the genetic score; ↓ indicates an association (p<0.05) with lower levels.

Abbreviations: N, number of participants; SD, standard deviation; FPLD1, familial partial lipodystrophy type 1; N/A not assessed. *matching variable **not reported, genetic score selection variable

Supplementary Table 9. Associations at the 53 loci with gene expression from eQTL repositories of multiple tissues.
(see Supplementary Table Excel file)

Supplementary Table 10. Associations of lead polymorphisms at the 53 loci in subcutaneous adipose tissue eQTL datasets.
(see Supplementary Table Excel file)

Supplementary Table 11. DEPICT annotation of putative effector genes.
(see Supplementary Table Excel file)

Supplementary Table 12. Associations at the *PIK3R1* locus. Comparison between phenotypic association patterns of the common single nucleotide polymorphism rs4976033 (effect allele: G; minor allele: G; minor allele frequency: 49.6%) at the *PIK3R1* locus (this study) and of rare loss-of-function mutations in *PIK3R1* (literature).

Phenotype	N of individuals or N of cases / N of controls	Beta in SDs or ln(OR) per G allele of rs4976033	P-Value	Association of rare loss of function mutations	Pubmed ID for rare loss-of-function mutation association
Height	358,297	-0.003 (0.0027)	0.32	Reduced	26497935; 23810378; 23810379; 23810382
Body mass index	447,441	-0.006 (0.0026)	0.02	Reduced*	26497935; 23810378; 23810379; 23810382
Body fat percentage	123,206	-0.020 (0.0033)	3.04E-09	Lipoatrophy or lipodistrophy*	26497935; 23810378; 23810379; 23810382
Waist-to-hip circumference	337,859	-0.003 (0.0027)	0.27	N/A	
LDL cholesterol	172,987	-0.001 (0.0040)	0.87	Normal*	23810379
HDL cholesterol	187,060	-0.021 (0.0037)	6.42E-08	Normal°	23810379
Triglycerides	177,755	0.0141 (0.0036)	0.00020	Normal°	26497935
Fasting glucose	133,010	0.002 (0.0023)	0.39	Raised°	23810379
2 hour glucose	42,854	0.008 (0.0120)	0.52	N/A	
Fasting insulin adjusted BMI	108,557	0.009 (0.0023)	0.00013	Insulin resistance*	26497935; 23810378
Type 2 diabetes	45,836 cases 230,358 controls	1.03 (1.01-1.05)	0.0022	High prevalence of early onset type 2 diabetes*	26497935; 23810378; 23810379
Coronary heart disease	8,660 cases 47,121 controls	1.07 (1.02-1.12)	0.0080	N/A	

Height data were from a meta-analysis of UK Biobank and GIANT data.
*Alignment between phenotypes associated with common and rare variants
°Lack of alignment between phenotypes associated with common and rare variants

Supplementary Table 13. Association estimates at loci selected for experimental validation of putative effector genes in cellular adipogenesis models.

SNP genomic coordinates	Insulin-raising / other allele	Putative effector gene	Direction of association with expression of the putative effector gene in subcutaneous adipocytes	p-value for expression of putative effector gene in subcutaneous adipocytes	Beta for hip circumference in standardised units (p-value)	OR of type 2 diabetes (p-value)
rs2943645 Chr2:227099180	T / C	*IRS1*	Lower expression	5.2E-09	-0.014 (9.4E-07)	1.09 (1.1E-17)
rs7973683 Chr12:124449223	C / A	*CCDC92*	Lower expression	2.1E-29	-0.014 (1.3E-06)	1.03 (3.7E-03)
rs7973683 Chr12:124449223	C / A	*DNAH10*	Lower expression	1.9E-08	-0.014 (1.3E-06)	1.03 (3.7E-03)
rs9492443 Chr6:130398731	C / T	*L3MBTL3*	Lower expression	9.1E-17	-0.021 (1.2E-11)	1.05 (3.1E-05)
rs3822072 Chr4:89741269	A / G	*FAM13A*	Higher expression	7.6E-12	-0.017 (5.5E-10)	1.04 (1.6E-05)

Genomic coordinates refer to build 37 (hg19). All association results are aligned to the insulin-raising (risk) allele. The co-localisation between association signals for gene expression in subcutaneous adipocytes and associations with fasting insulin are discussed in the Supplementary Note. Hip circumference association results are from a meta-analysis of the UK Biobank study and of the GIANT consortium. Type 2 diabetes association results are from a meta-analysis of DIAGRAM, InterAct and UK Biobank.

Supplementary Table 14. Characteristics of the participants with individual-level genotype data included in this study.

Study	Fenland	EPIC-Norfolk	InterAct	UK Biobank	UKHLS
Country	United Kingdom	United Kingdom	Multiple European countries	United Kingdom	United Kingdom
Participants	10,351	9150	15357	111016	5296
Cases / Controls	N/A	N/A	6410 / 8947	4586 / 106430	N/A
Age	48 (7)	58 (9)	53 (9)	57 (8)	53 (16)
Female sex, N (%)	5506 (53)	5015 (55)	9162 (60)	58390 (53)	5296 (100)
Genotyping chip	Affymetrix genome-Wide Human SNP Array 5.0 and Affymetrix UK Biobank Axiom Array	Affymetrix UK Biobank Axiom Array	Illumina 660w quad and Illumina CoreExome chip	Affymetrix UK Biobank Axiom Array	Illumina CoreExome Chip
Imputation panel	1000 Genomes Phase 1v3 and Phase 3	1000 Genomes Phase 3	1000 Genomes Phase 1v3	1000 Genomes Phase 3 plus UK10K	1000 Genomes Phase 3

References

1 Manning, A. K. *et al.* A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nature genetics* **44**, 659-669, doi:10.1038/ng.2274 (2012).

2 Scott, R. A. *et al.* Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nature genetics* **44**, 991-1005, doi:10.1038/ng.2385 (2012).

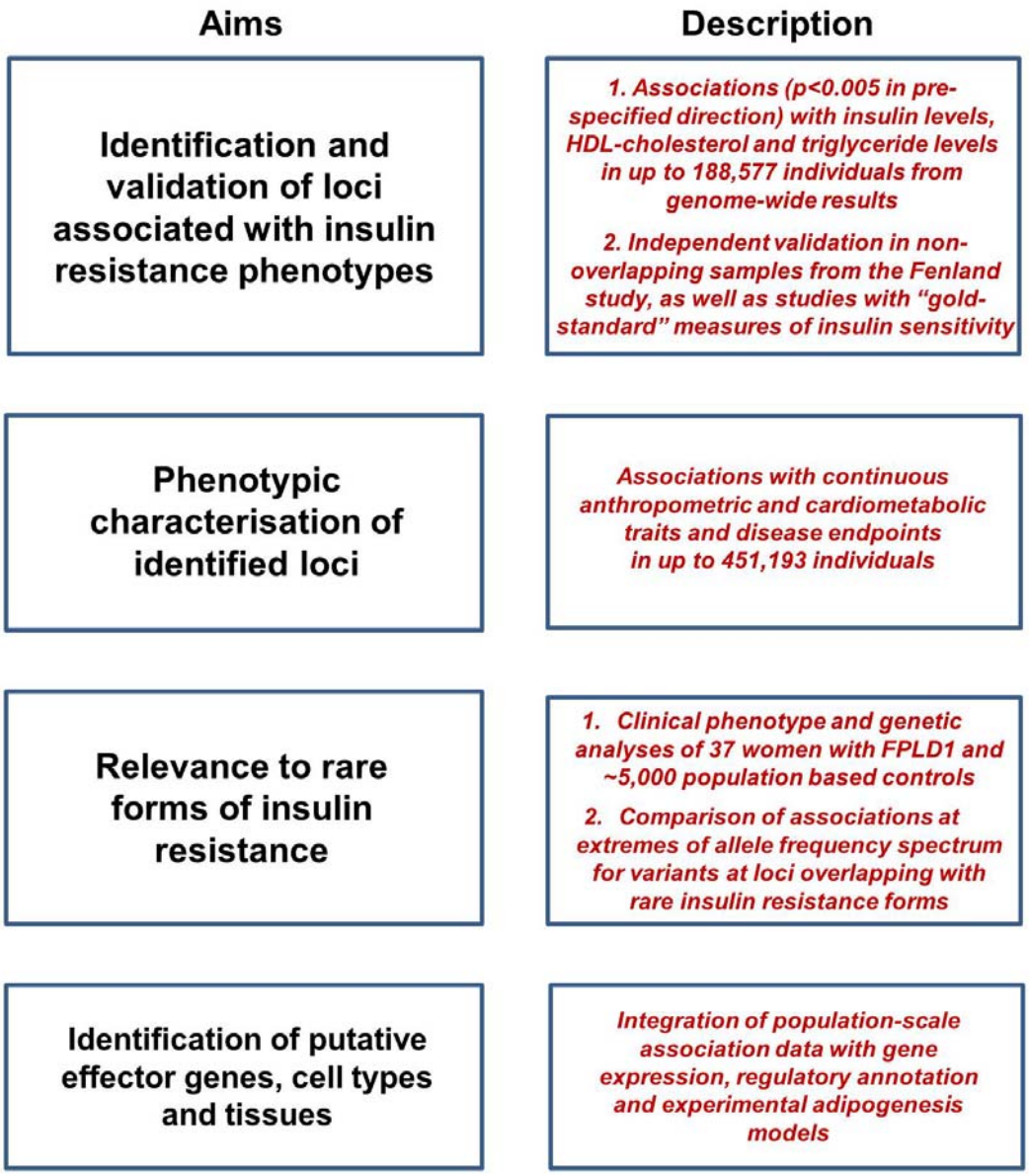
3 Small, K. S. *et al.* Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nature genetics* **43**, 561-564, doi:10.1038/ng.833 (2011).

4 Zhang, X. *et al.* Synthesis of 53 tissue and cell line expression QTL datasets reveals master eQTLs. *BMC genomics* **15**, 532, doi:10.1186/1471-2164-15-532 (2014).

5 Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators. Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *The New England journal of medicine*, doi:10.1056/NEJMoa1507652 (2016).

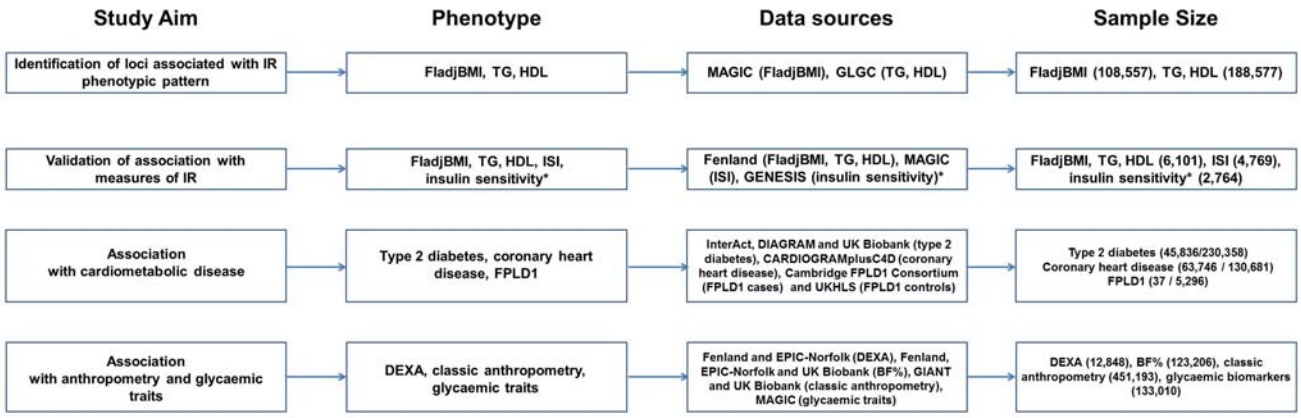
Table	Caption
Supplementary Table 4	Associations of lead single nucleotide polymorphisms at the 53 loci
Supplementary Table 9	Associations at the 53 loci with gene expression from eQTL reposit
Supplementary Table 10	Associations of lead polymorphisms at the 53 loci in subcutaneous
Supplementary Table 11	DEPICT annotation of putative effector genes.

with glycaemic, anthropometric traits and disease endpoints.
ɔries of multiple tissues.
adipose tissue eQTL datasets.



Supplementary Figure 1

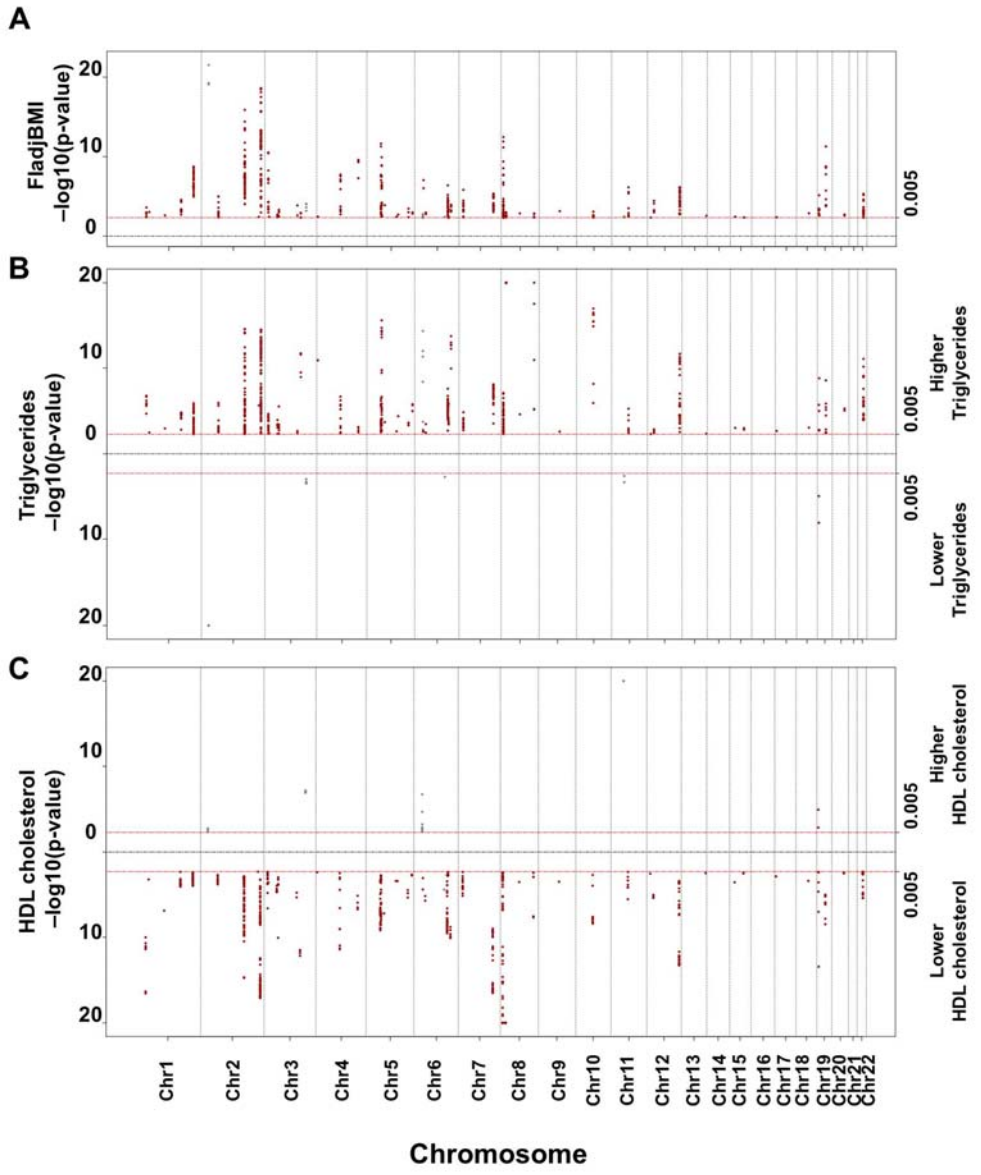
Design and scope of the study.



Supplementary Figure 2

Design of the study, investigated phenotypes, sources of data and sample size.

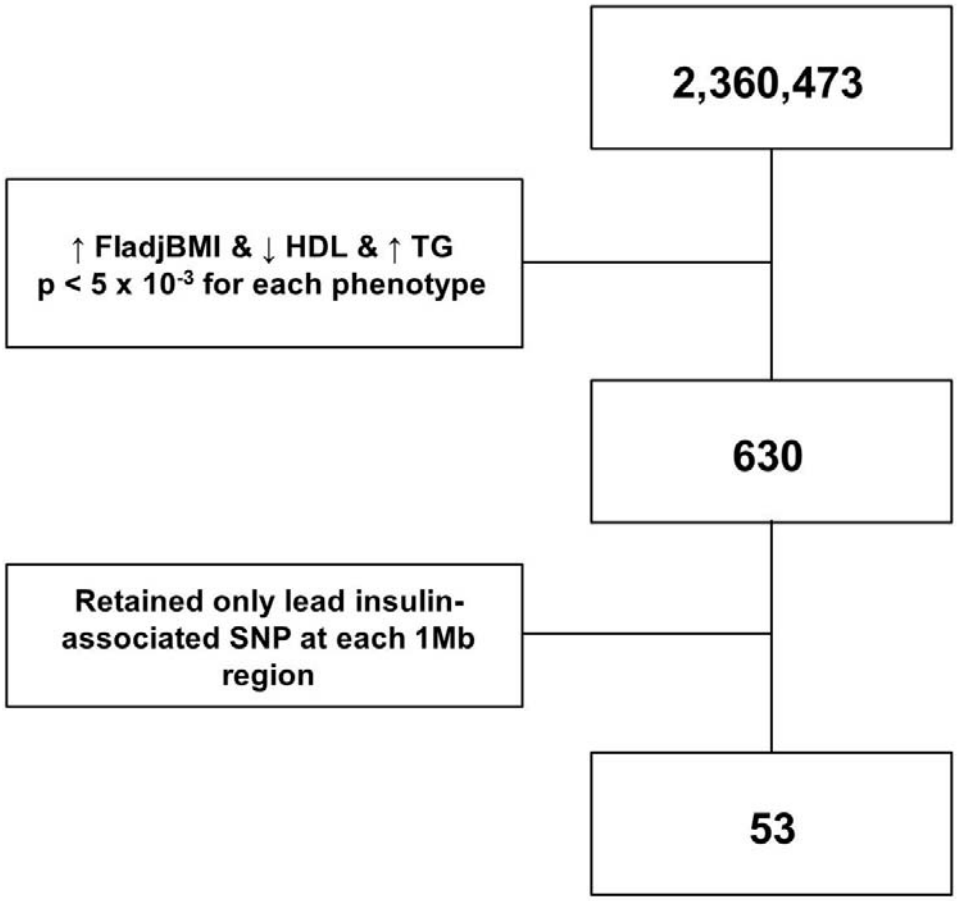
The reported sample size is the maximum available for a given trait or set of traits in this study. *In the study by Knowles and colleagues (Pubmed ID: 25798622), insulin sensitivity was measured by euglycaemic clamp or insulin suppression test in 2,764 European individuals from four cohorts. Abbreviations: IR, insulin resistance; SNP, single nucleotide polymorphism; FIadjBMI, fasting insulin levels adjusted for body mass index; TG, triglyceride levels; HDL, high-density lipoprotein cholesterol levels; ISI, insulin sensitivity index; DEXA, dual-energy X-ray absorptiometry; BF%, body fat percentage; FPLD1, familial partial lipodystrophy type 1; MAGIC, Meta-Analyses of Glucose and Insulin-related traits Consortium; GLGC, Global Lipids Genetics Consortium; GIANT, Genetic Investigation of ANthropometric Traits; DIAGRAM, DIAbetes Genetics Replication And Meta-analysis; CARDIOGRAM, Coronary ARtery DIsease Genome wide Replication and Meta-analysis; C4D, Coronary Artery Disease Genetics consortium.



Supplementary Figure 3

Combined directional Manhattan plots of the association with insulin resistance related phenotypes.

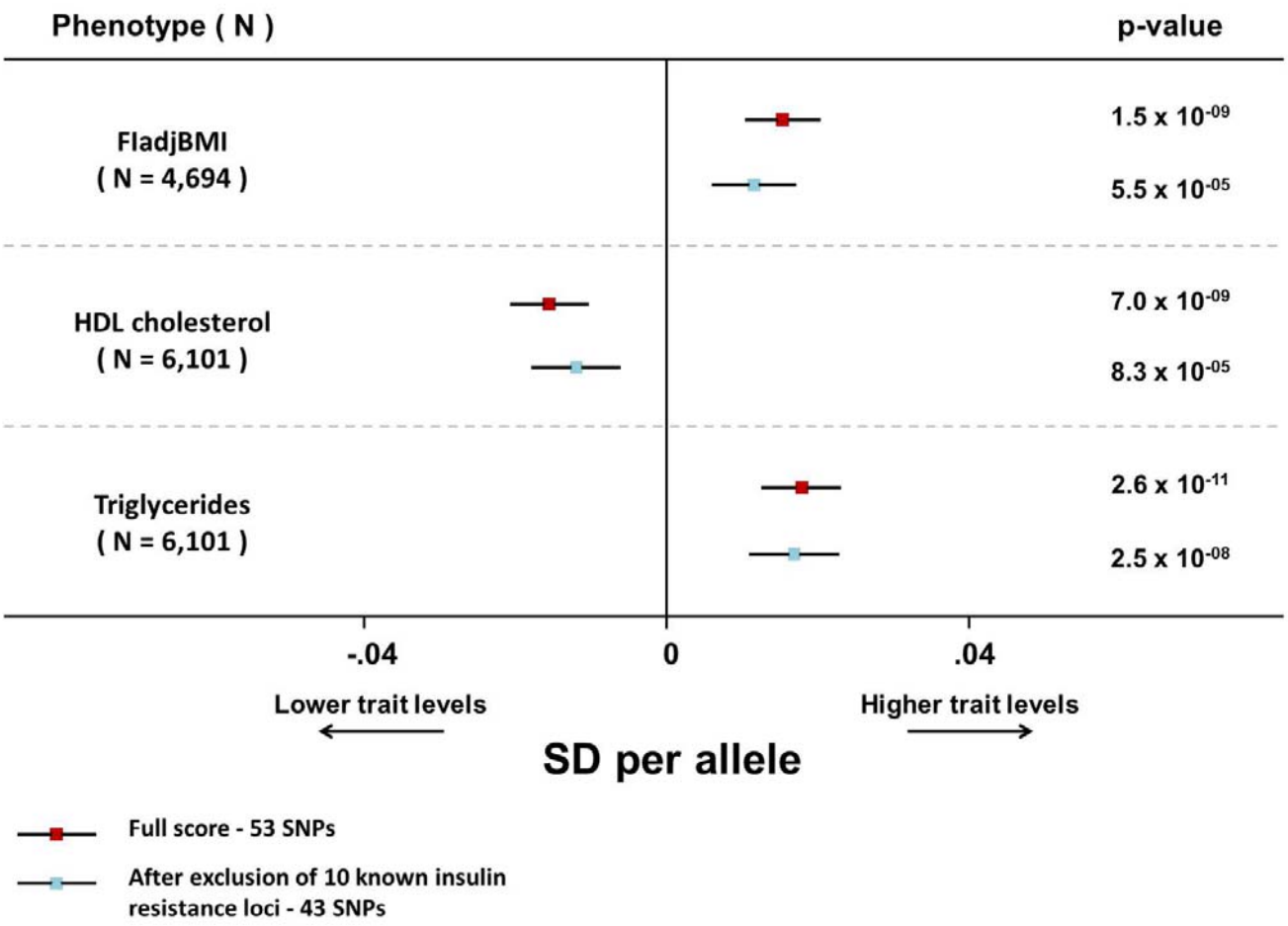
The figure represents Manhattan plots of the association of single nucleotide polymorphisms with fasting insulin adjusted for body mass index (FIadjBMI; Panel A), triglycerides (Panel B) and HDL cholesterol (Panel C). We plotted only variants with FIadjBMI, triglycerides and HDL cholesterol ($p<0.005$ for each phenotype). All associations are represented for the FIadjBMI-raising allele. The 630 alleles associated with higher FIadjBMI, higher triglycerides and lower HDL cholesterol are plotted in dark red. The graph also plots 21 variants that meet the p-value threshold for the three phenotypes but were not associated in the required direction (grey). For graphic display purposes, p-values below 10^{-20} were represented as 10^{-20}.



Supplementary Figure 4

Flowchart of the identification of insulin resistance loci.

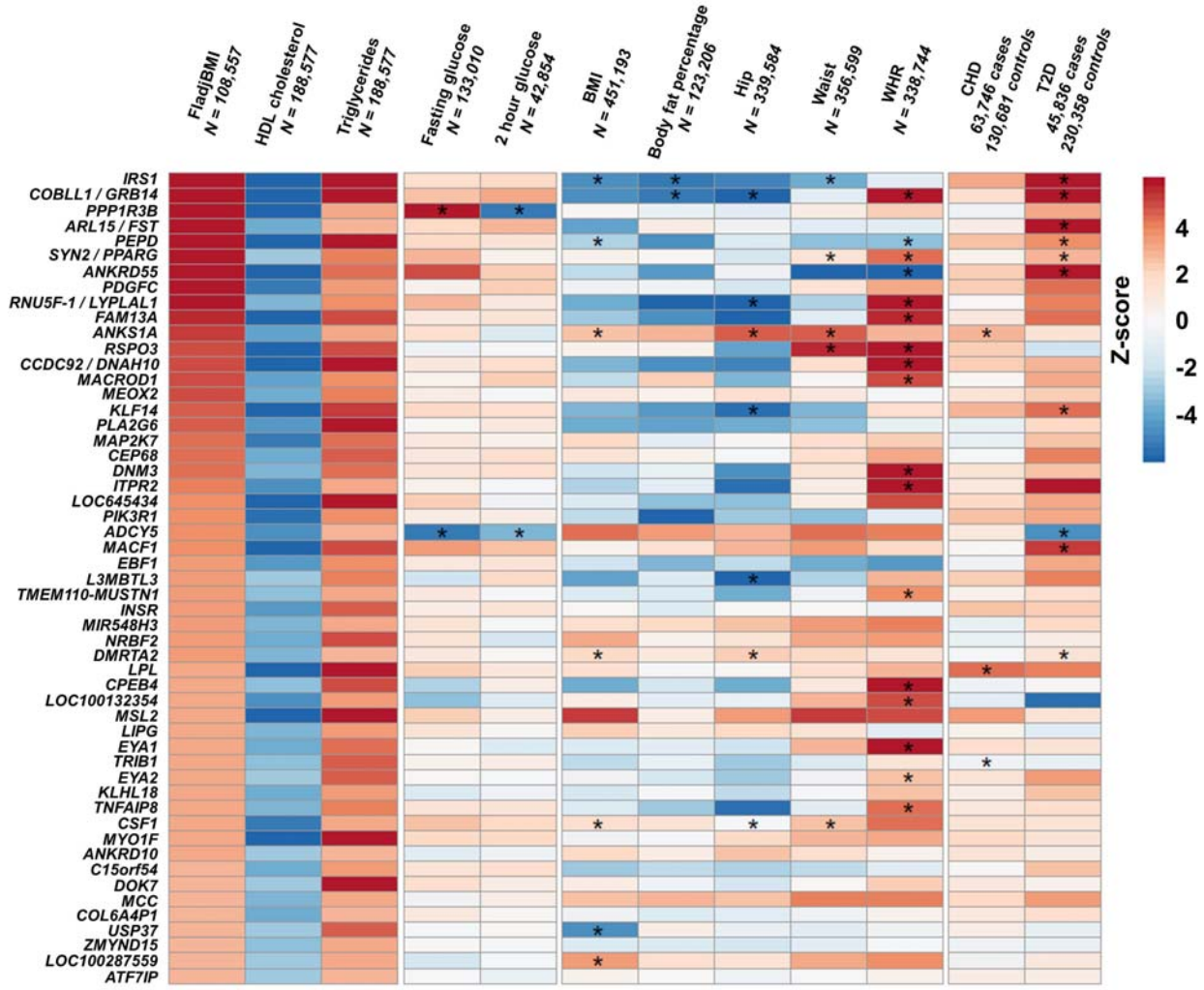
Numbers refer to single nucleotide polymorphisms. Abbreviations: FIadjBMI, fasting insulin adjusted for body mass index; HDL, high-density lipoprotein cholesterol; TG, triglycerides; SNP, single nucleotide polymorphism.



Supplementary Figure 5

Associations with insulin resistance phenotypes in an independent dataset.

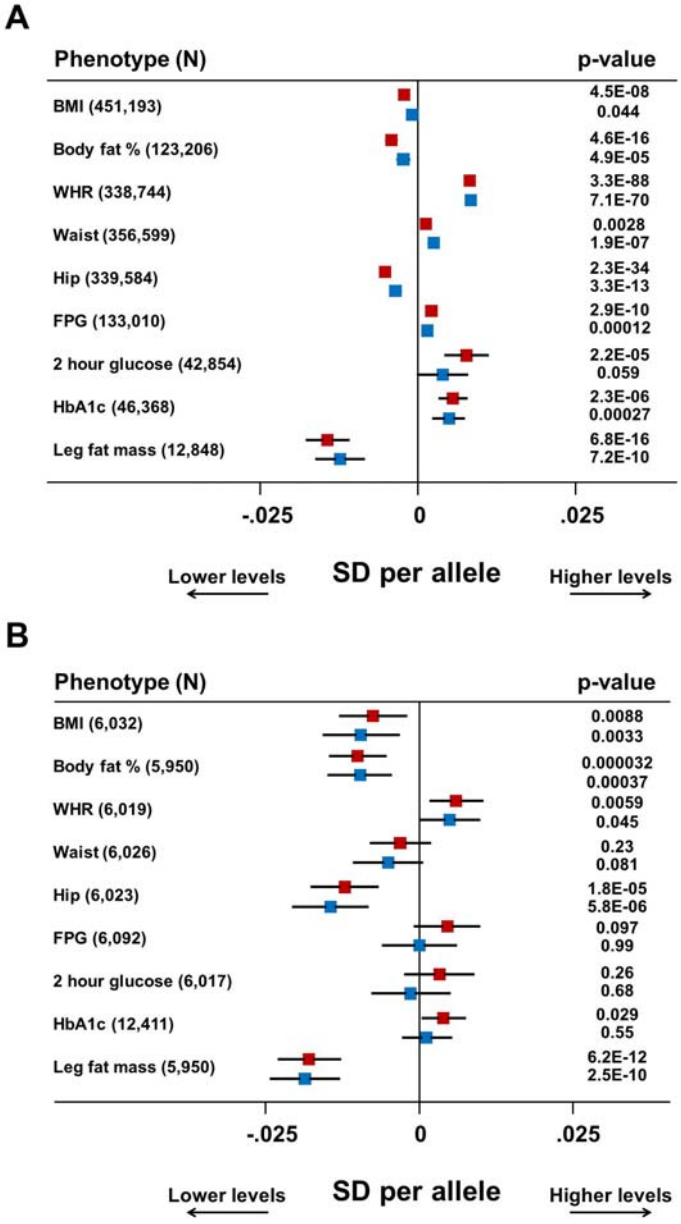
The figure reports associations of the genetic scores comprising the 53 or 43 SNPs with fasting insulin adjusted for body mass index, triglycerides and HDL cholesterol in up to 6,101 participants of the Fenland study who were not included in any of the discovery efforts used for the identification of the 53 loci. Squares indicate the central estimate of the beta coefficient; error bars the 95% confidence interval. Abbreviations: N, number of participants; FIadjBMI, fasting insulin adjusted for body mass index; HDL, high-density lipoprotein; SD, standard deviation.



Supplementary Figure 6

Associations with glycaemic, anthropometric traits and disease endpoints at the 53 genomic loci.

The heatmap represents Z-scores for the association of the lead insulin raising allele at each locus. Loci are ranked on the basis of their Z-score for fasting insulin (largest to smallest). With the exception of fasting insulin, none of the association analyses was adjusted for body mass index. Abbreviations: N, maximum sample size; FIadjBMI, fasting insulin adjusted for body mass index; HDL, high density lipoprotein cholesterol; BMI, body mass index; WHR, waist-to-hip ratio; CHD, coronary heart disease; T2D, type 2 diabetes. Colour scale: red indicates positive associations for the insulin-raising allele at each locus, while blue indicates negative associations. Asterisks indicate known loci for the traits, i.e. those for which our lead SNP is within 500 kb either side of a lead SNP from the largest GWAS of that trait.

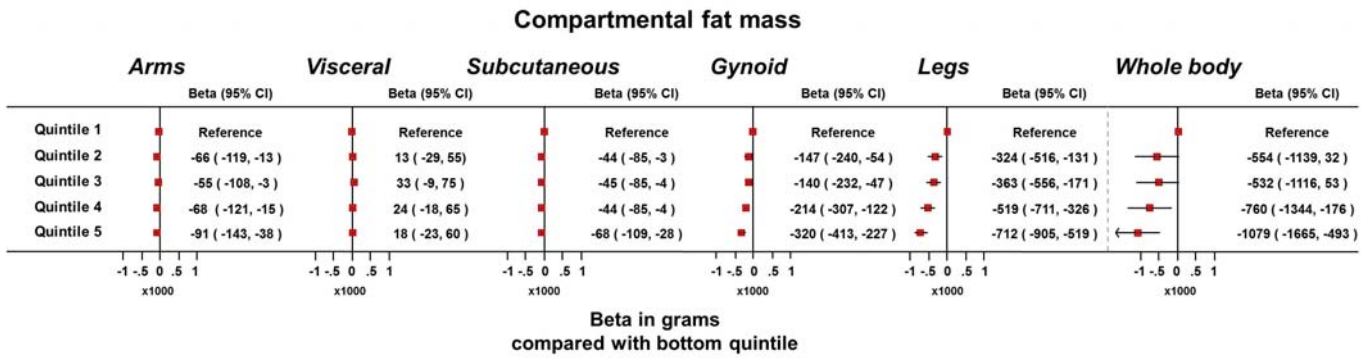


Supplementary Figure 7

Associations of the genetic scores comprising the 53 or 43 SNPs with glycaemic and anthropometric traits in large-scale meta-analyses and in the Fenland study.

Panel A shows the association of the genetic scores with anthropometric and glycaemic traits in meta-analyses of genetic association studies. Body mass index, waist-to-hip ratio, waist and hip circumference data are from the GIANT consortium and the UK Biobank study. Body fat percentage data are from the UK Biobank, EPIC-Norfolk and Fenland studies. Fasting plasma glucose, 2 hour glucose and HbA1c data are from the MAGIC consortium. Leg fat mass data are from the EPIC-Norfolk and Fenland studies. Squares with error

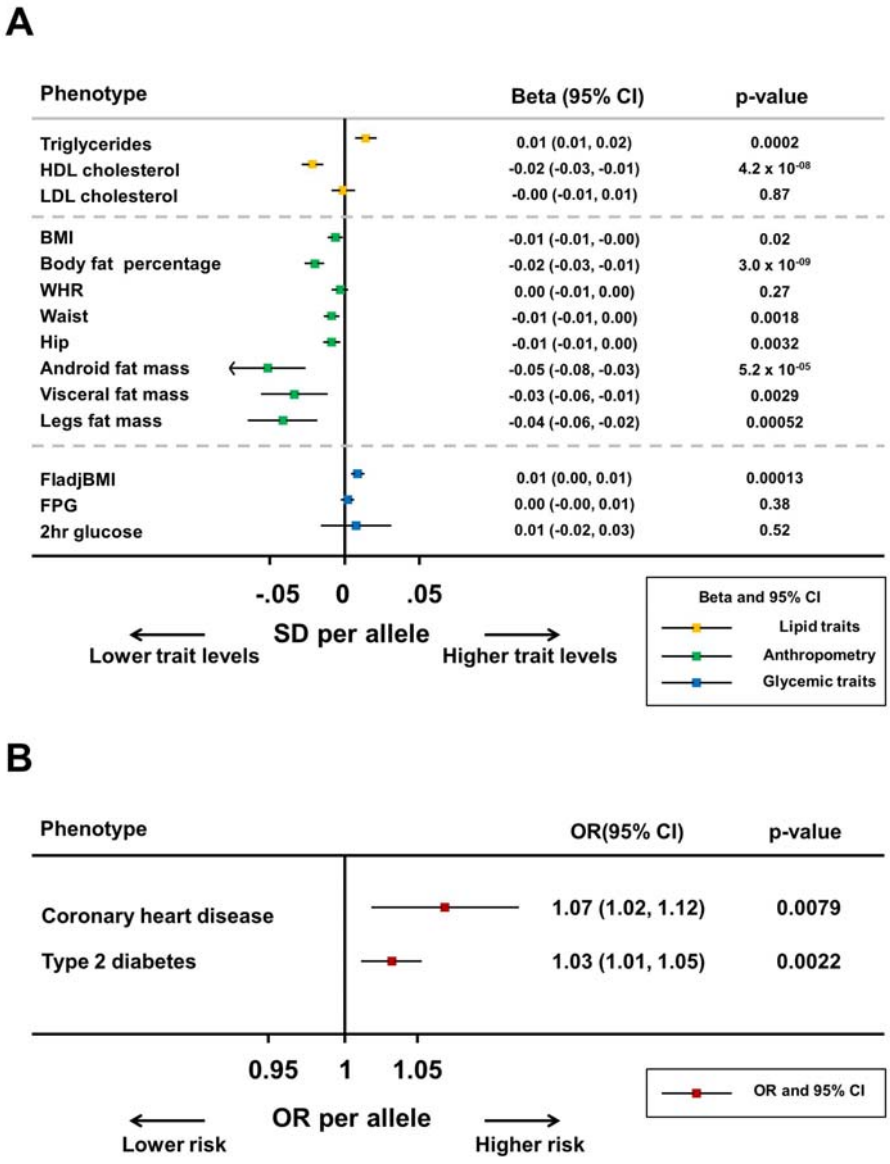
bars represent the per-allele beta coefficients in standard deviation units and their 95% confidence intervals. Panel B shows the association with the same traits in participants of the Fenland study not included in discovery efforts which contributed to the identification of the 53 loci. Since HbA1c has been measured only in a subset of Fenland, the HbA1c analysis includes also individuals from the InterAct study subcohort who did not take part in discovery efforts which contributed to the identification of the 53 loci. Squares with error bars represent the per-allele beta coefficients in standard deviation units and their 95% confidence intervals. Red and blue squares represent results of the 53-SNP and 43-SNP genetic scores, respectively. None of the results presented in the figure is adjusted for body mass index. Abbreviations: N, number of participants; SD, standard deviation; BMI, body mass index; WHR, waist-to-hip ratio; FPG, fasting plasma glucose.



Supplementary Figure 8

Associations of the 53-SNP genetic score with detailed anthropometric variables from dual energy X-ray absorptiometry.

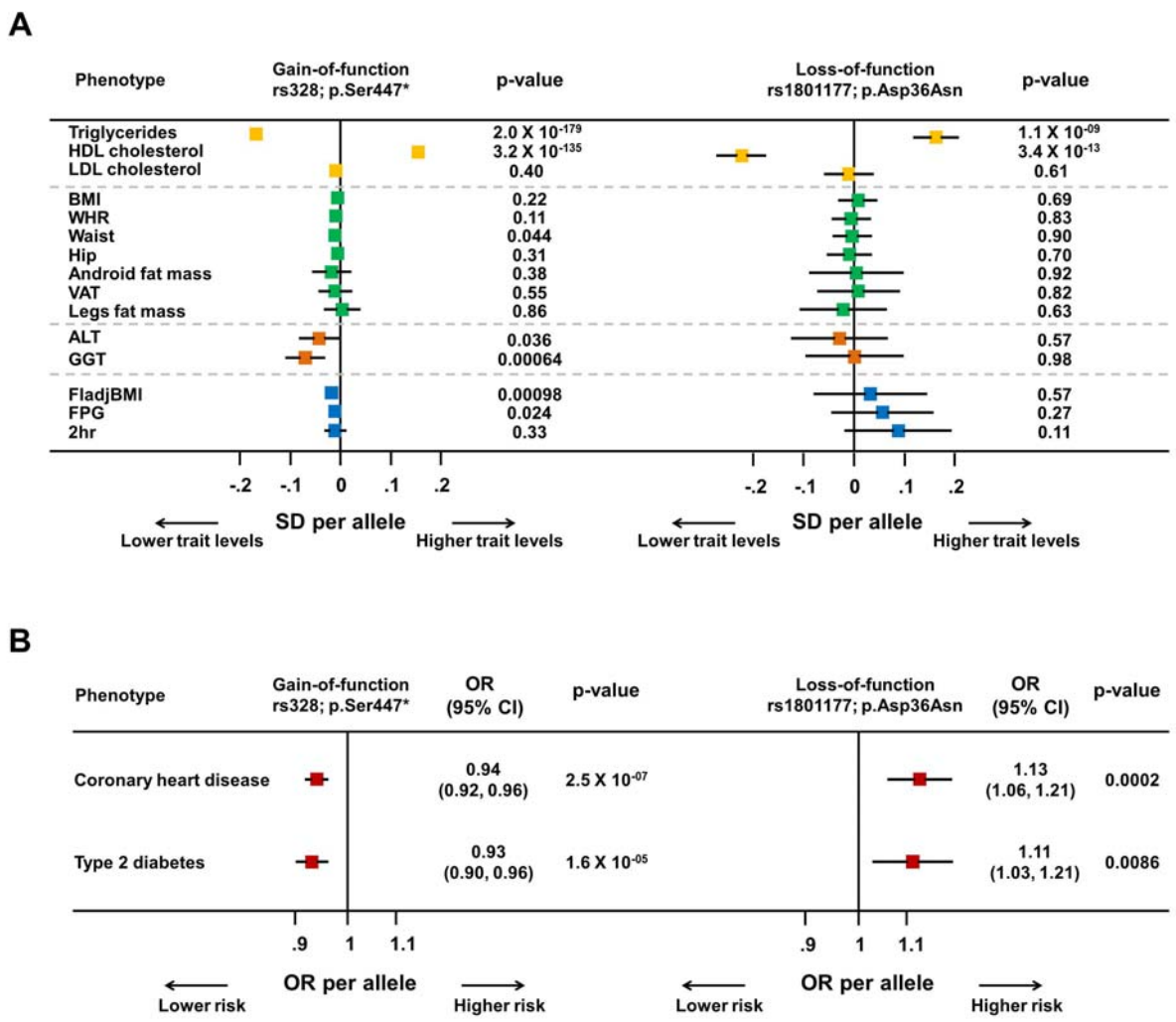
The figure represents the association of quintiles of the 53-SNP genetic score with the absolute values of compartmental and total fat mass. Data are from 9,747 participants of the Fenland study. The Fenland population was divided into quintiles of the distribution of the genetic score and each quintile was compared with the bottom (reference category). Squares with error bars represent the beta coefficients in grams for individuals in the exposure category compared with the reference category and their 95% confidence intervals.



Supplementary Figure 9

Associations of the rs4976033-G allele near *PIK3R1* with continuous metabolic traits and cardiometabolic disease endpoints.

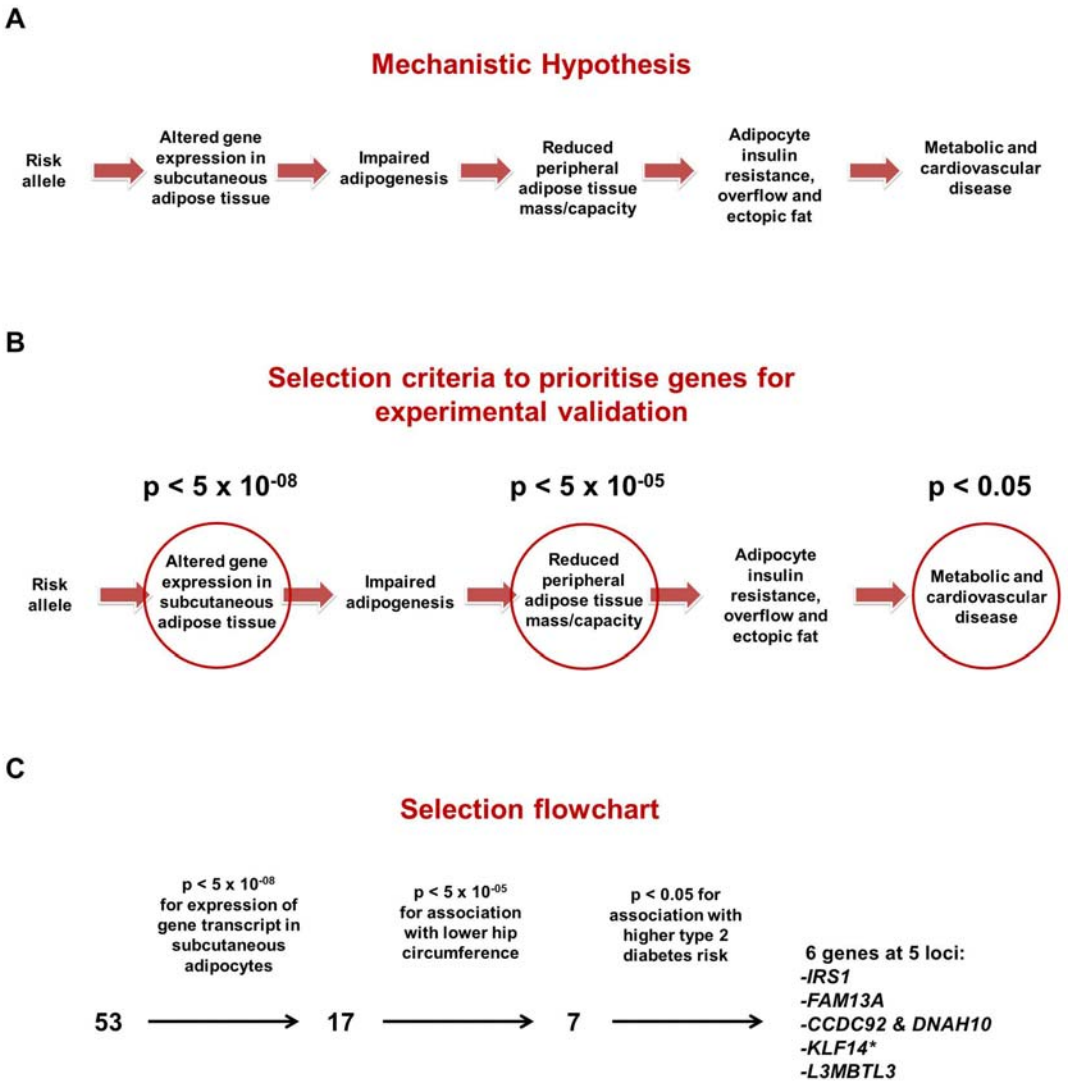
Panel A represents associations with continuous traits, while Panel B those with disease endpoints. Squares with error bars represent the beta coefficients (Panel A) or odds ratios (Panel B) and their 95% confidence intervals. Abbreviations: HDL, high density lipoprotein; LDL, low density lipoprotein; BMI, body mass index; WHR, waist-to-hip ratio; FIadjBMI, fasting insulin adjusted for BMI; FPG, fasting plasma glucose; 2hr glucose, glucose at two hours during an oral glucose challenge; SD, standard deviation; OR, odds ratio.



Supplementary Figure 10

Associations of functional variants in *LPL* with cardiometabolic traits and disease endpoints.

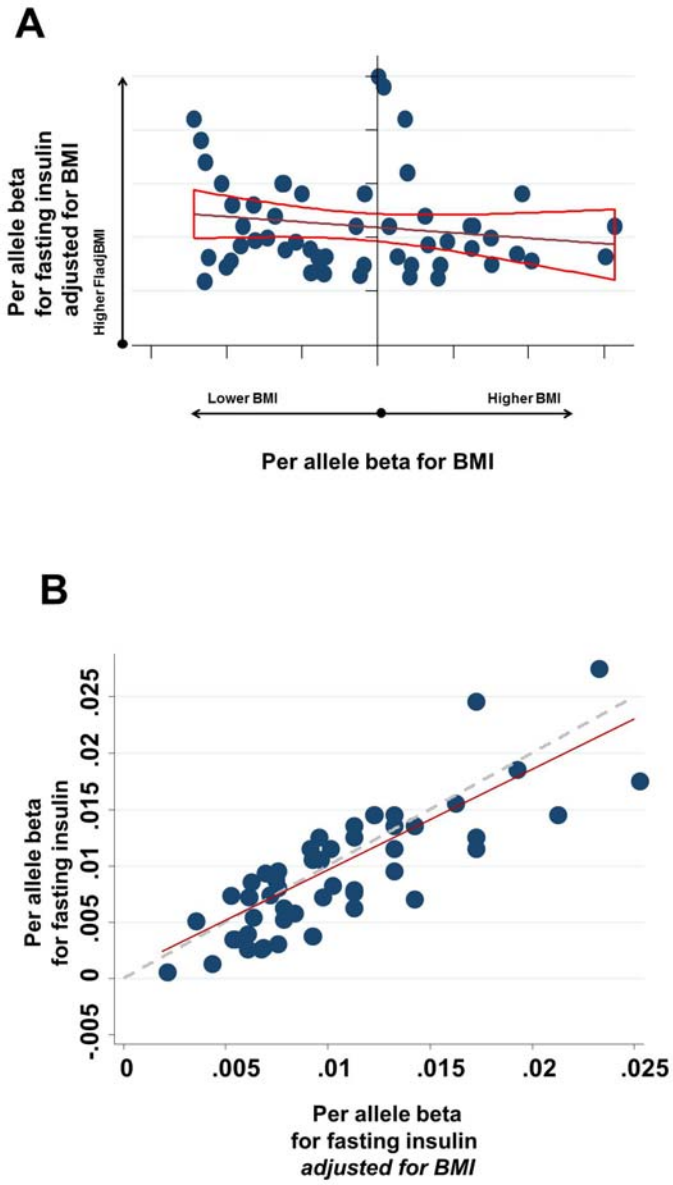
Panel A represents the association of the gain-of-function p.Ser447* (rs328; left) and of the loss-of-function p.Asp36Asn (rs1801177; right) in *LPL* with lipid levels, anthropometric traits, liver markers and glycemic traits. Panel B represents the association of the two variants with the risk of coronary heart disease (from the Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators; PubMed ID: 26934567)[5] and that of type 2 diabetes. Squares with error bars represent the beta coefficients (Panel A) or odds ratios (Panel B) and their 95% confidence intervals. Abbreviations: HDL, high density lipoprotein; LDL, low density lipoprotein; BMI, body mass index; WHR, waist-to-hip ratio; VAT, visceral adipose tissue; ALT, alanine aminotransferase; GGT, gamma-glutamyl transferase; FIadjBMI, fasting insulin adjusted for BMI; FPG, fasting plasma glucose; 2hr, glucose at two hours during an oral glucose challenge; SD, standard deviation; OR, odds ratio.



Supplementary Figure 11

Mechanistic hypothesis for the implication of putative effector genes in the observed associations and selection of genes for experimental validation in cellular models of adipogenesis.

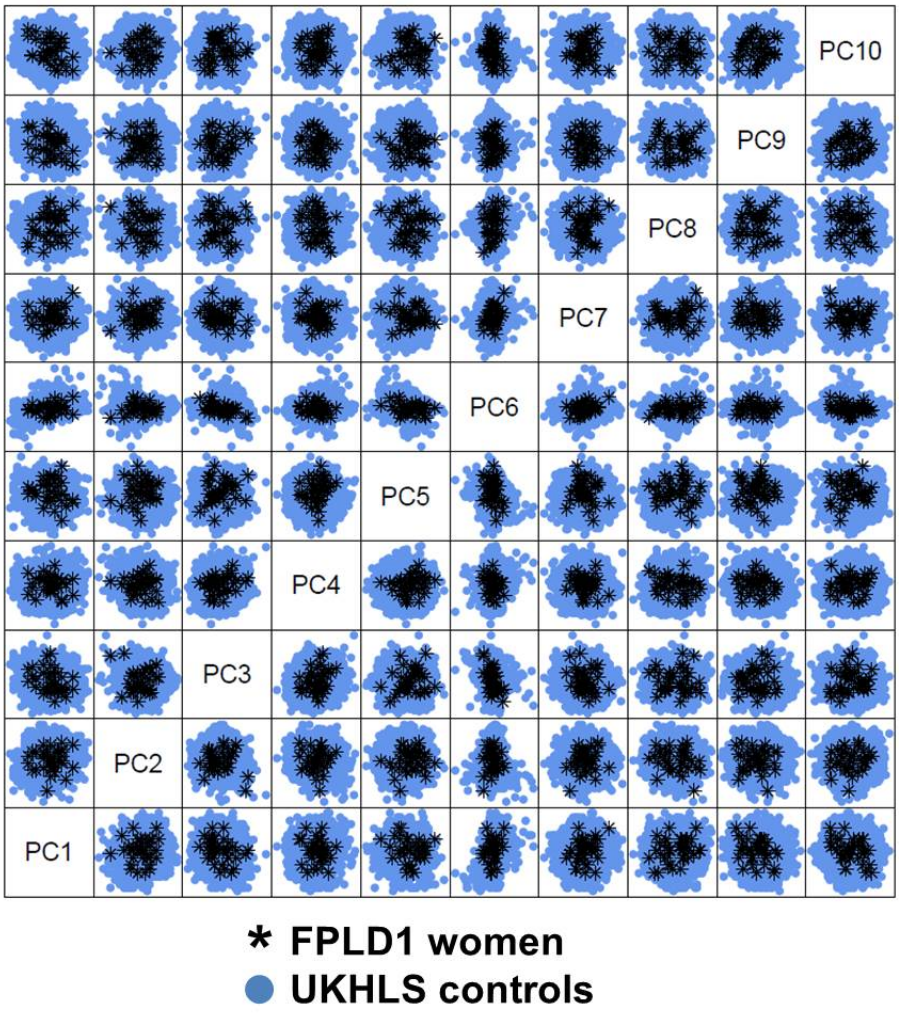
In Panel C, numbers refers to loci meeting certain selection criteria. *We did not take forward the *KLF14* gene to experimental validation because previous studies about the role of the gene in metabolic disease suggest complex aetiologic mechanisms at this locus including a possible parent-of-origin effect.[3]



Supplementary Figure 12

Associations with fasting insulin adjusted for body mass index (FIadjBMI), body mass index (BMI), or fastin insulin (FI) of the 53 polymorphisms identified in this study.

Panel A shows the association of the 53 lead polymorphisms from our study with FIadjBMI as a function of the association with BMI. There was no clear bias in the association with FIadjBMI (linear regression between beta coefficients of 53 polymorphisms, p=0.26). Panel B depicts the association with FI (unadjusted for BMI) of the lead 53 polymorphisms as a function of the association with FIadjBMI. The line of fit was aligned with the line of equality consistent with no bias. In Panel A, the dark red line and surrounding areas represent the lines of fit with 95% confidence areas. The dashed grey line in Panel B represents the line of equality. Data about fasting insulin associations is from the MAGIC consortium; data about BMI associations is from the GIANT consortium.



Supplementary Figure 13

Scatter plot matrix of the top ten genetic principal components in FPLD1 women and UKHLS control women.