UCLA
UCLA Previously Published Works

Title

Regulatory variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition.

Permalink

https://escholarship.org/uc/item/3v41156f

Journal

Nature Genetics, 50(4)

Authors

Small, Kerrin
Todorčević, Marijana
Civelek, Mete
et al.

Publication Date

2018-04-01

DOI

10.1038/s41588-018-0088-x

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Europe PMC Funders Group
Author Manuscript
Nat Genet. Author manuscript; available in PMC 2018 October 09.

Published in final edited form as:
Nat Genet. 2018 April ; 50(4): 572–580. doi:10.1038/s41588-018-0088-x.

Regulatory variants at *KLF14* influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition

Kerrin S. Small[#1,†], **Marijana Todorčević**[#2], **Mete Civelek**[#3,4], **Julia S. El-Sayed Moustafa**[#1], **Xiao Wang**[#5], **Michelle M. Simon**[6], **Juan Fernandez-Tajes**[7], **Anubha Mahajan**[7], **Momoko Horikoshi**[2,7], **Alison Hugill**[8], **Craig A. Glastonbury**[1], **Lydia Quaye**[1], **Matt J. Neville**[2,9], **Siddharth Sethi**[6], **Marianne Yon**[8], **Calvin Pan**[18], **Nam Che**[4], **Ana Viñuela**[1], **Pei-Chien Tsai**[1], **Abhishek Nag**[1], **Alfonso Buil**[10], **Gudmar Thorleifsson**[11], **Avanthi Raghavan**[12], **Qiurong Ding**[13], **Andrew P. Morris**[7,14], **Jordana T. Bell**[1], **Unnur Thorsteinsdottir**[11,15], **Kari Stefansson**[11,15], **Markku Laakso**[16], **Ingrid Dahlman**[17], **Peter Arner**[17], **Anna L. Gloyn**[2,7,9], **Kiran Musunuru**[#5], **Aldons J. Lusis**[#4,18,19], **Roger Cox**[#8], **Fredrik Karpe**[#2,9], and **Mark I. McCarthy**[2,7,9,†]

[1]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK [2]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, UK [3]Center for Public Health Genomics, Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia, USA [4]Department of Medicine, University of California, Los Angeles, California, USA [5]Cardiovascular Institute, Department of Medicine, Department of Genetics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA [6]Biocomputing, Medical Research Council Harwell Institute, Oxford, UK [7]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK [8]Genetics of type 2 diabetes, Medical Research Council Harwell Institute, Oxford, UK [9]Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK [10]Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland [11]deCODE Genetics, Reykjavik, Iceland [12]Harvard Medical School, Boston, Massachusetts, USA [13]CAS Key Laboratory of Nutrition and Metabolism, Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, PR China [14]Department of Biostatistics, University of Liverpool, Liverpool, UK [15]Faculty of Medicine, University of Iceland, Reykjavik Iceland [16]Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland [17]Department of Medicine, Huddinge, Karolinska

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

[†]Corresponding authors: Correspondence should be addressed to K.S.S. (Kerrin.small@kcl.ac.uk) or M.I.M (mark.mccarthy@drl.ox.ac.uk).

Author contributions

K.S.S., A.L.G., K.M., A.J.L., R.C., F.K., and M.I.M. conceived and designed the project M.T., M.C., J.S.E.M., M.M.S., J.F., C.A.G., L.Q., C.P., G.T., P.C.T. and A.N. analysed data. M.T., X.W, Q.D., A.R., A.H., S.S., M.Y. and N.C. performed experiments. A.M., M.H., M.J.N. and A.V. contributed experimental and technical support as well as discussion. A.B., J.T.B., I.D., P.A., M.L., U.T. and K.S. contributed data. K.S.S., M.T., M.C., J.S.E.M., M.M.S., K.M., A.J.L., R.C., F.K., and M.I.M. wrote the manuscript. All authors read and approved the manuscript.

Competing financial interests statment

M.I.M. has received consulting and advisory board honoraria from Pfizer, Lilly, and NovoNordisk. G.T., U.T. and K.S. are employees of deCODE Genetics/Amgen.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Institutet, Stockholm, Sweden [18]Department of Human Genetics, University of California, Los Angeles, California, USA [19]Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, California, USA

[#] These authors contributed equally to this work.

Abstract

Individual risk of type 2 diabetes (T2D) is modified by perturbations of adipose mass, distribution and function. To investigate mechanisms responsible, we explored the molecular, cellular, and whole-body effects of T2D-associated alleles near *KLF14*. We show that *KLF14* diabetes-risk alleles act in adipose tissue to reduce *KLF14* expression, and modulate, in *trans*, expression of 385 genes. We demonstrate that, in human cellular studies, reduced *KLF14* expression increases pre-adipocyte proliferation but disrupts lipogenesis, and, in mice, adipose-specific deletion of *Klf14* partially recapitulates the human phenotype of insulin resistance, dyslipidemia and T2D. We show that *KLF14* T2D risk-allele carriers shift body fat from gynoid to abdominal stores, and display a marked increase in adipocyte cell size: these effects on fat distribution, and the T2D-association, are female-specific. Metabolic risk associated with variation at this imprinted locus depends on both the sex of the subject, and of the parent from whom the risk-allele derives.

The replicated genome-wide significant T2D association signal at chr7q32.3 maps to a 45kb recombination interval, extending from 3kb to 48kb upstream of *KLF14*1,2 (Figure 1a-c). In previous work based on microarray-derived RNA expression data, *KLF14*, which encodes an imprinted transcription factor, was exposed as the likely cis-effector gene for this locus in subcutaneous adipose tissue1 and revealed to be a trans-regulator of a programme of adipose tissue expression3. The KLF family of zinc-finger binding proteins have wide-ranging regulatory roles in biological processes such as proliferation, differentiation and growth4,5. However, little is known about *KLF14*, a single exon gene whose transcription is limited to the maternally inherited chromosome in embryonic, extra-embryonic, and adult tissue in humans and mice6.

Results

Adipose-specific regulation of KLF14 mediates the T2D association

Using RNAseq data from subcutaneous adipose biopsies collected from 776 female twin members of the TwinsUK cohort7, we confirmed the *cis*-expression quantitative trait locus (eQTL) (using rs4731702 T2D risk allele C as the reference, β=-0.56, P=1.8x10^{-36}) and identified a 38-fold-expanded *trans*-network of 385 genes (false discovery rate (FDR) <5%) (Supplementary Table 1, Figure 2). The *KLF14 cis* and *trans* effects were robust, replicating in three independent studies of subcutaneous adipose expression (MGH8, METSIM9 and deCODE10) (Figure 2c; Supplementary Table 1, Supplementary Table 2). Despite detectable *KLF14* expression in a range of tissues, these *cis*- and *trans*-associations were completely adipose-specific with no eQTL signal evident in (a) skin, whole blood and lymphoblastoid cell lines (LCL) from the same TwinsUK individuals7 (Figure 1f, Figure 2d); (b) T2D-relevant tissues such as muscle11, liver8,12, and islet13; or (c) the broader coverage

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

represented in eQTL datasets such as GTEx14 (Supplementary Table 3). *KLF14* is imprinted, and as with the T2D-association1, the *KLF14 cis* and *trans* eQTLs were maternal-specific – the paternally inherited allele had no effect on *KLF14* expression (Supplementary Figure 1, Supplementary Table 4). All other genes within 300kb of the interval are also imprinted, but paternally-expressed, confirming *KLF14* as the likely mediator of the T2D association at this locus.

We first sought to refine the location of the causal variant(s) responsible for these associations. The T2D association could not be resolved beyond a set of 29 SNPs in high mutual linkage disequilibrium (LD, r^2>0.94 in UK10K); the apparent European-specificity of the T2D association15 precludes trans-ethnic fine-mapping. The ancestral T2D-risk allele varies in global frequency (EUR = 54%, ASN=69%, AMR=58%, AFR=73%) but we could detect no evidence of positive selection (Supplementary Figure 2, Supplemental Note). We utilized chromatin state maps to annotate the associated interval, identifying a ~1.6kb enhancer ~5kb upstream of *KLF14*. This enhancer encompasses five of the 29 associated variants and shows marked tissue-specificity: in ChromHMM predictions from ENCODE16 and Roadmap17, it is annotated as active in 10 (out of 127) cell types and tissues, three of them derived from adipose tissue, including adipocytes and adipose-derived mesenchymal stem cells (Figure 1d). The enhancer is also active in certain blood cells, but there was no corresponding *cis*-eQTL18–20, making it unlikely that transcriptional regulation in blood cells contributes to the GWAS phenotypes. There were no other adipose active enhancers in the region.

As methylation is one of the key processes related to enhancer function, we explored population-level methylation data at this locus in many of the same TwinsUK individuals. The T2D-risk haplotype was associated with increased methylation levels at Illumina 450K array probe cg02385110, ~3kb upstream of *KLF14,* in subcutaneous adipose tissue *(n*=603, P=2.2x10^{-7}, β=0.01 for rs4731702) but not whole blood (*n*=309, P=0.69) or skin (n=437, P=0.39) (Figure 1e-i). There was a consistent direction of effect across these analyses in adipose: the T2D risk-haplotype was associated with increased methylation and decreased RNA expression, and cg02385110 methylation and *KLF14* expression were negatively correlated. A second probe (cg08097417), located at the *KLF14* transcription start site has been associated with age in whole blood21 and adipose tissue samples22. However, this probe lies outside the T2D-association interval, and cg08097417 methylation was not related to *KLF14* expression (P=0.36) or risk-haplotype (P=0.99), indicating that age-related variability of cg08097417 is unrelated to *KLF14* expression or disease pathogenesis (Supplementary Figure 3). Instead, we conclude that T2D-associated risk attributable to this locus is likely to be consequence of sequence variation at the adipose enhancer upstream of *KLF14* and is marked by altered methylation.

The KLF14 variants regulate a large adipose-specific trans network

The *trans*-network regulated by the *KLF14* variants is remarkable both for its size and robust replication. Consistent with the known function of *KLF14* as both transcriptional activator and repressor, the *trans*-associations included both positive and negative effects (Figure 2a, Supplementary Table 1). Mediation analysis applied to the RNA expression data supported a

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

causal role for *KLF14* expression in regulation of many of the *trans*-genes (161/385 genes passed Bonferroni-corrected Sobel's mediation $P<1.3\text{x}10^{-4}$).

The principal *trans*-regulatory mechanism appears to be direct interaction of KLF14 with *trans*-gene *cis*-regulatory elements; the 20kb regions upstream of the 385 *trans* genes were enriched both for *KLF14* binding peaks in empirical CHiPSeq data23 from HEK293T cells ($P=2\text{x}10^{-4}$) and the presence of the proposed *KLF14* binding motif (Normalized Enrichment Score=4.17, $P=1.5\text{x}10^{-5}$). We assessed functional annotation of the *trans*-genes using ToppGene24, and found the subset of 177 *trans*-genes with KLF14 binding-motifs to be enriched for "metabolic pathways" ($P_{\text{Bonferroni}}=1.2\times10^{-3}$) and "binding by *PPARG* and *RXRA* during adipocyte differentiation" ($P_{\text{Bonferroni}}=3.9\text{x}10^{-8}$) (Supplementary Table 5). Further subsetting, defining 122 *trans*-genes enriched for both KLF14 binding motifs and *KLF14* ChIP-seq peaks revealed broadly similar functional enrichment ("binding by PPARG and RXRA during adipocyte differentiation": $P_{\text{Bonferroni}}=0.04$; "metabolic pathways": $P_{\text{Bonferroni}}=2.6\text{x}10^{-4}$).

A sub-network of 18 *trans*-genes showed enrichment for *SREBF1* binding motifs (NES=4.22, $P=1.2\text{x}10^{-5}$), 11 of which did not have *KLF14* motifs (Figure 2b). The sub-network of *SREBF1* motif-containing *trans*-genes was enriched for "cholesterol biosynthesis and lipid metabolism" ($P_{\text{Bonferroni}}=2.8\times10^{-7}$) (Supplementary Table 5). *SREBF1*, a transcription factor involved in cholesterol homeostasis, is itself a *trans*-gene directly regulated by *KLF14* and appears to act as an intermediary in the regulation of this sub-network.

The human *trans*-network includes several genes with functions that make them attractive candidates for mediating the range of *KLF14*-associated phenotypes. These include *SLC2A4* and *IDE*, which encode the GLUT4 transporter and insulin degrading enzyme, respectively. GLUT4 mediates glucose uptake in adipose tissue and skeletal muscle, among other tissues, and adipose-specific reduction in *SLC2A4* expression levels has been reported in T2D25. IDE degrades peptides including insulin, glucagon, and amylin, and maps to a long-established T2D genome-wide association study (GWAS) interval15. Notably, at multiple *trans*-genes, we identified adipose *cis*-eQTLs coincident with metabolic-trait GWAS SNPs ($r^2>0.8$) including three known T2D loci (*STARD10*, *C6orf57* and *CDK2AP1*) (Supplementary Table 6). This shows that local regulation of these genes can mediate T2D-susceptibility independent of *trans*-regulation via *KLF14,* and indicates that the phenotypic consequences of *KLF14* variation are likely mediated by multiple genes within the *trans* network.

Sex-specific metabolic trait associations implicate insulin action

Previous studies have demonstrated that T2D-risk alleles at *KLF14* are, in non-diabetic individuals, associated with increased fasting insulin26 and reduced high density lipoprotein (HDL)-cholesterol27,28, emphasizing a primary impact on insulin action rather than insulin secretion. By collecting data from the largest-available GWA meta-analyses for multiple traits26,28–31, we have extended the spectrum of *KLF14* association to encompass a broad range of insulin resistant, "metabolic syndrome" phenotypes including low density lipoprotein (LDL)-cholesterol, triglycerides, waist-hip ratio and the HOMA-IR measure of

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

insulin sensitivity32 (Table 1). Crucially, as demonstrated in Shungin et al31, the impact of *KLF14* variation was far more marked on fat distribution rather than overall adiposity (as measured by body mass index (BMI)), with a particularly strong association between the T2D-risk allele and reduced hip circumference (β=-0.017, P=1.6x10^{-6}).

Sex-stratified reanalysis of these GWAS data revealed that, for all traits with available data, effect sizes in females substantially exceeded those in males (Table 1). The associations with hip circumference, triglycerides, LDL and fasting insulin were exclusive to females, while the HDL, WHR and T2D associations displayed a strong female-bias (e.g. T2D: female P=2.2x10^{-6}, odds ratio (OR) 1.14 [1.08-1.20]; male P=0.002, OR 1.08 [1.03-1.14]). In UKBiobank (N=118,192), we found significant SNPxSex interactions for Hip Circumference (P=3.1x10^{-7}) and Waist Hip Ratio adjusted for BMI (P=3.9 x 10^{-6}), but not for Waist Circumference (P=0.06) or BMI (P=0.43). In a combination of UKBiobank, Wellcome Trust Case Control consortium and The Resource for Genetic Epidemiology on Adult Health and Aging (GERA) samples (13,728 cases; 129,911 controls), rs4731702 and sex interacted to influence T2D-risk (P=6.6x10^{-3}).

These analyses demonstrate that the metabolic and anthropometric consequences of *KLF14* variation are dependent on both the parental origin of the risk-allele and the sex of the recipient of that allele. Based on combining these sex-specific ORs with the known imprinting-related maternal-specificity of these associations1, we estimate that the point estimate for the per-allele OR for T2D is ˜1.28 in women inheriting the risk-allele from their mother. In contrast, although *KLF14* expression in adipose tissue was higher in females than males across all genotype classes, we found no equivalent sex-difference in the magnitude of the adipose *cis*-eQTL effect (Figure 3a, Supplementary Table 4). To further explore drivers of the sexual dimorphism in KLF14 expression levels, we compared *KLF14* expression level between 86 pre-menopausal and 487 post-menopausal females from TwinsUK. Individuals taking hormone replacement therapy were excluded from this analysis. There was no association between *KLF14* expression and menopausal status (P=0.81) and the *cis*-eQTL effect size in each group was comparable, suggesting gonadal steroids do not underlie the observed sexual dimorphism in *KLF14* expression levels. We conclude that the sex-dependency of the whole body phenotype is not the direct consequence of sexual dimorphism at the *cis*-regulatory level.

The genetic associations were consistent with our observations using gene expression levels in the TwinsUK cohort. *KLF14* gene expression in adipose tissue was associated with a combined insulin resistance phenotype characterized by increased fasting insulin and triglycerides, and reduced HDL cholesterol (P=1.08x10^{-3}). *KLF14 trans*-genes also showed enrichment for association between *trans*-gene expression levels and the same combined insulin resistance phenotype (P=1.82x10^{6}) (Supplementary Table 7- 8). We used Dual energy X-ray Absorptiometry (DXA)-derived body composition data from the same TwinsUK individuals to further dissect the impact on fat distribution. Whilst there was no association between *KLF14* expression in subcutaneous adipose tissue and total fat volume (P=0.56) or BMI (P=0.27), there were differences with respect to the distribution of that fat, including an inverse association with abdominal visceral fat (P=0.02) and with the ratio of android:gynoid fat (P=1.7x10^{-3}). In many other insulin-resistant settings, abnormalities of

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

fat distribution are associated with ectopic deposition of fat including the liver, leading to non-alcoholic fatty liver disease. However, we found no evidence that *KLF14* risk variants influence liver fat deposition, whether measured directly (P=0.26, N=7,176)33 or indirectly via disturbed liver function (alanine aminotransferase (ALT) P=0.79, gamma-glutamyl transferase (GGT) P=0.89, N=61,089)34.

These data indicate that genetically-determined reduction in *cis*-expression of *KLF14* in adipose results in an insulin-resistant and T2D-predisposing phenotype, which, in women only, is characterized by, and potentially mediated through, a shift in fat distribution from relatively inert gynoid stores to more metabolically-active abdominal deposition.

Conserved mouse phenotypes

We generated mice carrying an adipose-specific *Klf14* deletion allele through the combination of a CRISPR-Cas9-generated conditional allele with *Klf14*-flanking loxP sites and adipose-specific (*Adipoq*) Cre recombinase and compared these mice with wildtype Cre-expressing colony mates. Adipose-specific knockout mice of both sexes (null on both maternal and paternal chromosomes) displayed a broad pattern of insulin resistance phenotypes: HDL-C was reduced (16 weeks, significant only in females), triglycerides increased (16 weeks, significant only in males) (Figure 4a,b), and both glucose tolerance and insulin sensitivity were impaired in both sexes at 12 weeks (Figure 4c-f).

We also characterised two mouse-lines in which *Klf14* expression was reduced globally. The first mouse (*Klf14*$^{tm1(KOMP)Vlcg}$; generated within KOMP) featured deletion of the entire single exon of *Klf14.* The second mouse, designed to minimise any disruption of adjacent regulatory sequences, used CRISPR-Cas9 genome editing to induce a 7bp frameshift indel allele. There were some inconsistencies between the models, but the predominant phenotype was a reduction in HDL-C (Supplementary Figures 4a and 5a) with a modest and/or transitory effect on glucose homeostasis (Supplementary Figures 4b,c,d and 5d,e). The HDL-C phenotype seen in global deletion mice mirrors the phenotype reported following hepatic-specific deletion of *Klf14*35.

RNASeq analysis of subcutaneous fat taken from both the adipose *Klf14* KO and the *Klf14*$^{tm1(KOMP)Vlcg}$ global deletion mouse recovered some, but not all, features of the *trans*-network seen in humans. In the adipose-specific *Klf14* KO mouse (N=8), 18 genes were differentially expressed at FDR q<0.05 (or 1286 with p<0.05). The latter set (of 1286 nominally significant mouse *trans*-genes) was significantly enriched for overlap with the human trans-network ($P<2x10^{-4}$). In the *Klf14*$^{tm1(KOMP)Vlcg}$ global deletion mouse (N=16), we confirmed this overlap ($P<1x10^{-6}$ with the human trans network) in the 285 genes differentially expressed compared to controls (FDR q<0.05): 5' noncoding sequences for these genes were also enriched for the presence of the *Klf14* binding motif (P<0.001). Although *Srebf2* (the mouse ortholog of *SREBF1*) is directly regulated by *Klf14,* we did not detect enrichment of the murine *trans*-genes for the *Srebf2* motif. The murine *trans*-genes were enriched for relevant functional categories, including "response to lipid" ($q=4.4x10^{-5}$), "triglyceride metabolic process" ($q=2.3x10^{-3}$), "response to steroid hormone" ($q=2.7x10^{-3}$) and "regulation of cell proliferation" ($q=3.2x10^{-3}$) (Supplementary Table 9).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Thus, adipose-specific deletion of *Klf14* in mice recapitulates many aspects of the phenotype associated with variants that influence adipose-specific *cis*-regulation of *KLF14* in humans, but the absence of sex specificity, and differences in the molecular consequences of *Klf14* perturbation place limits on the relevance of murine models for this gene.

KLF14 impacts glucose uptake, lipogenesis and cell size

Members of the KLF family are involved in transcriptional control of adipocyte development and function5, but the role of *KLF14* in adipogenesis is unknown. To evaluate the role of *KLF14* in adipocyte development, we measured *KLF14* expression in primary pre-adipocytes isolated from human abdominal subcutaneous adipose tissue biopsies (female N=4, male N=4) during proliferation and subsequent differentiation over 14 days. Adipocyte expression of *KLF14* was higher in females compared to males at all time-points (Figure 3b), demonstrating that the sex-differential expression of *KLF14* observed in adult biopsies is present throughout adipogenesis.

We next investigated the link between reduced *KLF14* expression and adipocyte development. Compared to tissue from non-risk allele homozygotes, fresh adipose tissue explants from female risk allele homozygotes showed a 44% reduction (P=0.001, N=132) in lipogenesis as measured by the incorporation of ^{14}C- glucose label into the triacylglycerol backbone (Figure 5e). There was no significant difference in males (P=0.49, N=32) (Figure 5e). We confirmed these findings *in vitro* through shRNA knockdown (40% at day 1 and 33% at day 14) of *KLF14* (Figure 5a) in primary pre-adipocytes isolated from abdominal subcutaneous biopsies (N=7 females): this resulted in 60% reduction in TAG accumulation after 14 days of differentiation (P=0.03) (Figure 5c). The adipocyte differentiation wholly relies on *de novo* lipogenesis as the cell culture medium contains glucose but not fatty acids. Knowing that *SLC2A4* (which encodes GLUT4) is a *KLF14* transgene, we first tested the expression of GLUT4 in the KLF14 shRNA differentiated adipocytes. This showed prominent reduction (by 64±7% (P=0.02)) (Figure 5a). To investigate the functional consequences of the reduction in GLUT4 expression, we quantified the insulin-stimulated glucose uptake, finding a 50% reduction in *KLF14* shRNA differentiated adipocytes (Figure 5d). Concomitant with the lower glucose uptake and lipogenesis, the expression of adipocyte maturation factors *PLIN1* (-42%±8, P=0.02) *LEP (*-54%±13, P=0.03), and *FITM2* (-12%±4, P=0.04) was significantly ablated in the knockdown cells, whilst the levels of the classical early-differentiation, pro-adipogenic transcription factors (*CEBPA*, *PPARG2*) were not changed (Figure 5a). *KLF14* knockdown also increased cellular proliferation, observed as reduced doubling time (P=0.02) (Figure 5b). We infer that reduced levels of *KLF14* result in a defect in glucose uptake resulting in impaired adipocyte lipogenesis and maturation, and propose that the increase in proliferation results from fruitless cycling at earlier stages of differentiation.

Adipocyte size is an important marker of adipose tissue dysfunction and metabolic disease and, in some studies, has been predictive of insulin resistance and T2D independent of obesity36–40. We assessed adipocyte size and cell number distribution by histological analysis of subcutaneous abdominal and gluteal adipose tissue biopsies from sex-, age- and BMI-matched pairs of individuals from the Oxford Biobank, homozygous for either the risk

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

or non-risk haplotypes (N=18 males, 18 females). At both sites, there was a marked shift towards increased adipocyte size in T2D-risk haplotype homozygotes: this was only observed in women (comparison of medians: abdominal P=0.008, gluteal P=0.02) (Figure 6b,d,e; Supplementary Figure 6a,b,c). We estimated, assuming spherical cell morphology, that there was a two-fold difference in adipocyte volume between the genotypic groups in females. Gluteal and abdominal tissue from female T2D-risk haplotype homozygotes also contained fewer total adipocytes (comparison of medians: Abdominal TT=$2.21\pm0.31\text{x}10^{10}$, GG=$0.66\pm0.19\text{x}10^{10}$, P=0.04; Gluteal TT=$4.87\pm0.93\text{x}10^{10}$, GG=$2.25\pm0.58\text{x}10^{10}$, P=0.04). The cell size findings and their sex-specificity were confirmed using a different technique (measuring diameter of isolated adipocytes from collagenase digested tissue) in an independent cohort of 28 females (P=0.004), and 16 males (P=0.78) also matched for age and BMI (Figure 6c).

Discussion

T2D risk variants that lead to reduced expression of *KLF14* in adult adipose tissue are therefore also implicated in a defect of adipogenesis that is likely to reflect impaired glucose uptake. The consequence of this defect appears to be a profound effect on subcutaneous adipose tissue structure characterised by the presence of fewer but larger adipocytes. We infer that *KLF14* risk-allele carriers are obliged to favour expansion of existing adipocytes to meet lipid storage needs, resulting in suboptimal fat storage and an increased risk of insulin resistance and T2D. This hypothesis of impaired fat storage is consistent with the elevated circulating triglyceride levels observed in *KLF14* risk haplotype carriers. Reduced expression of *KLF14* in female carriers of the risk-allele is likely to drive a shift in the distribution of adipose tissue between subcutaneous fat depots, favouring abdominal over gluteal deposition. This results in loss of the relative protection against the metabolic consequences of adipose tissue accumulation associated with adipose deposition in gynoid sites which has been widely observed in epidemiological studies41,42 and is consistent with the observed association of the *KLF14* T2D-risk allele and human fat distribution.

We note that the sex-specificity of the *KLF14* risk alleles varies across biological domains: effects on adipocyte development and whole body phenotypes are more marked in females despite the fact that *cis* and *trans* regulatory effects are shared between the sexes. The simplest explanation for this discordance is a threshold effect driven by the higher baseline expression levels of *KLF14* in females, though it is also possible that males lack downstream processes (e.g. related to hormonal environment) that mediate the response to *KLF14*. While the *KLF14* T2D association was discovered through typical GWAS approaches, the remarkable specificity of this association – with respect to sex, parent of origin, ethnicity and tissue of action - is a reminder that risk prediction, or targeted medical treatment, based on genotype alone may fail to capture highly-relevant aspects of biological complexity.

Online Methods

Choice of lead SNP

The T2D association at the *KLF14* locus consists of 29 SNPs in near perfect linkage disequilibrium (r^2>0.94). The lead SNP in T2D GWAS has been reported as rs4731702

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

(Kong, et al1) rs972283 (Voight, et al2) and rs10954284 (Morris et al30 and Mahajan, et al15), all three of which are in perfect linkage disequilibrium in Europeans (r^2=1). The lead SNP for the HDL GWAS is rs4731702 (Teslovich et al28, Willer et al27). As rs4731702 is present on all commonly used genotyping arrays we chose to report all results in this manuscript with respect to rs4731702, where C is the T2D risk allele and T is the non-risk allele.

TwinsUK gene expression

Biopsies and blood samples from 856 healthy female twins from the TwinsUK cohort were collected within the MuTHER project44 and RNA sequenced as previously described7. See Supplementary Note for further details of the TwinsUK data.

Cis and *trans* eQTL analysis

Cis and trans eQTL analysis was conducted in the TwinsUK RNAseq data as follows. Exon read counts were corrected for technical covariates and family structure using a mixed effects model including insert size and mean GC content as fixed effects and primer index, sample processing date, family and zygosity as random effects. Corrected residuals were used for all eQTL analyses. *Cis*- and *trans*-eQTL associations were conducted using the MatrixeQTL package45 using a standard additive linear model, with BMI, age and age-squared included as covariates. A window size of 1Mb was used for *cis*-eQTL analyses. A FDR46 threshold of 5% was applied to *trans*-eQTL results, with FDR calculated using the qvalue package47 in R version 3.1.1. The deCODE, METSIM and MGH cohorts were analyzed as previously described8–10.

Functional fine mapping

We used the chromatin states predicted by ChromHMM43 utilizing 127 reference epigenomes from the ENCODE16 and Roadmap Epigenomics projects55. ChromHMM segments the genome into 15 states using 5 chromatin marks (H3K4me3, H3K4me1, H3K36me3, H3K27me3, H3K9me3) with predictions ranging from active transcription start sites to repressed Polycomb complexes. The ChromHMM predicted enhancer spans a 1.6 kb region, ˜4kb upstream of *KLF14* transcription start site (TSS) and harbors 5 of the T2D GWAS SNPs (rs12154627, rs6973807, rs6974400, rs6974288, rs11762784).

TwinsUK methylation data

Adipose tissue sample in 603 individuals (taken from the same biopsy as the RNAseq data) and 309 whole blood samples from the TwinsUK study were profiled on the Illumina Infinium HumanMethylation450 BeadChip previously48,49. The following covariates were included in analyses, batch (beadchip), BS conversion efficiency (as assessed using the built-in BS conversion efficiency controls) and BS-treated DNA input. The association of age with methylation was examined using a linear mixed effect regression model (LME) adjusting for batch effects, BMI, and, family and zygosity structure. Blood methylation was additionally corrected for estimated cell counts. To compare methylation and expression levels, each was first adjusted for covariates, and Pearson's correlation performed on the residuals.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Trans-network mediation analysis

Significant mediation was determined by computing Sobel's test statistic50. To calculate the mediation score, the three following models were implemented:

Model 1:

$$Y \sim \beta_1 A + \beta_2 P + \beta_3 G + \varepsilon$$

Model 2:

$$Y \sim \beta_1 E + \beta_2 A + \beta_3 P + \beta_4 G + \varepsilon$$

Model 3:

$$\text{Mediation score} = \frac{\beta_3 - \beta_4}{\beta_4}$$

Where Y = *trans*-gene expression, A = age, E = *cis*-gene expression, P = BMI, G = *cis*-genotype.

By conditioning on *cis*-gene expression (the mediator: E) we can determine if each individual association detected in *trans* is regulated in *cis* or is independent by quantifying $\Delta\beta G$.

Enrichment of *KLF14* Chipseq binding in *trans*-genes

We utilized the HEK293 KLF14 Chipseq data from Najafabadi *et al* (2015)51. Genomic co-ordinates of the peaks of *KLF14* binding (along with the summit position for each peak) identified in the Chip-seq experiments were available in the supplementary data of Najafabadi *et al* (2015) (http://hugheslab.ccbr.utoronto.ca/supplementary-data/C2H2_B1H/). For the enrichment analysis for *KLF14* binding, we utilised the summit position of the 18,652 distinct Chip-seq peaks of *KLF14* binding (filtered on a minimum Phred-like quality score of call confidence of 50) provided in this data. An enrichment of *KLF14* binding sites for the 385 genes with *trans*-expression association with the *KLF14* locus (FDR < 5%) was evaluated by comparing the proportion of genes for which *KLF14* binding sites were identified in the following two groups of genes: first, the 385 genes with *trans*-expression association with the *KLF14* locus; and second, the remaining genes that were tested in the *trans*-expression analysis (N=18,678). An enrichment of *KLF14* binding sites for the 385 genes with *trans*-expression association was further evaluated by performing 10,000 permutations of a random set of 385 genes sampled from the entire set of genes that was tested in the *trans*-expression analysis, and estimating the number of genes with *KLF14* binding sites for each permutation.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

iRegulon transcription factor binding analyses

In order to find the *KLF14* regulon (a transcription factor (TF) and its direct transcriptional targets, which contain common TF binding sites in their *cis*-regulatory control elements) we use the iRegulon cytoscape plugin52. The prediction of regulons consists of four steps: 1) motif detection, this process relies on an offline scoring step of a sequence search space (10kb or 20kb around the TSS) whereby every gene in the human genome, along with orthologous sequences in ten other vertebrate genomes, is scanned with *Cluster-Buster*53 for homotypic clusters of motifs using a library of nearly 10,000 position weight matrices (PWMs), resulting in a gene-ranking list for each PWM. Enriched motifs are those motifs for which the input genes are enriched at the top of the ranking, using the Area Under the Curve (AUC) of the cumulative curve; 2) Track discovery, this step also relies on an offline scoring step whereby every gene in the human genome is scored with around 1,120 ChiP-seq tracks, generating a list of TFs according to the highest ChiP peak within the regulatory space; 3) Motif2TF mapping, the candidate TFs are prioritized by finding the optimal path from a motif to a TF, in a motif-TF network; 4) Target detection, for each enriched motif, the candidate targets are selected as the significantly highly ranked genes compared to the genomic background and to the entire motif collection as background.

Functional enrichment analysis

KLF14 trans-genes were subdivided into three subsets based on the presence of bindings motifs for *KLF14*, *SREBF1* or neither. The three subsets were assessed for functional enrichment using ToppFun, part of the ToppGene package24. Results were corrected for multiple testing using the Bonferroni method, with enrichments considered significant at Bonferroni-corrected $P < 0.05$.

Overlap between *KLF14 trans*-regulated genes and published GWAS results

The NHGRI-EBI GWAS catalog54 (Available at: www.ebi.ac.uk/gwas, accessed November 2015) was filtered so as to retain signals meeting genome-wide significance ($P<5.0\text{x}10^{-8}$). Corresponding *cis*-eQTLs for each of the 385 *KLF14* trans-genes, including secondary, tertiary and quaternary *cis*-eQTLs sequentially conditioned on the lead *cis*-eQTLs at each locus (a total of 465 *cis*-eQTLs), were then cross-referenced against the NHGRI GWAS results, retaining at each *KLF14* trans-gene locus those reported NHGRI associations within +/-250kb of the relevant *cis*-eQTLs that showed moderate to high LD ($r^2 \geq 0.8$) between the reported GWAS SNP and the respective *cis*-eQTL.

Association between *KLF14* expression and TwinsUK metabolic traits

To determine if gene expression was associated to a concurrently measured phenotype of interest, expression of each exon was treated as a quantitative trait in a linear mixed effects model implemented using the lme4 package51. The linear mixed effects model was adjusted for age, BMI and technical covariates (mean GC content and insert size mode) (fixed effects), family relationship (twin pairing), primer index and zygosity (random effects). Phenotypic data were treated as continuous independent traits, and were inverse normal transformed. A full model with the phenotype fitted was compared to a null model (no phenotype) using a 1d.f. ANOVA. A FDR was estimated using the package QVALUE to

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

obtain *q-values* that correspond to a controlled FDR 5%. TwinsUK metabolic phenotypes were measured at the same time-point as the biopsy and were collected as previously described3. Body-fat distribution traits were measured using dual-energy X-ray absorptiometry (DXA; Hologic QDR 4500 plus) with the standard protocol.

Trait-Expression Association Enrichment Analyses

The most significantly-associated exon per gene was retained for all genes. Enrichment was assessed by comparing the number of genes significantly associated with each phenotype (at q<0.05) within the set of *KLF14 trans*-genes ($N_{trans\text{-}genes}$=385) to the remaining genes not within the *KLF14 trans*-network (N_{genes}=18,716), using a Fisher's exact test.

Sex x SNP interaction analysis in large cohorts

Details of the data preparation, quality control and cohort specific analysis covariates for the UK Biobank, GERA and WTCCC/T2D data are provided in the Supplementary Methods. For evaluating the SNP x Sex interaction effect of rs4731702 on anthropometric traits in the UK Biobank, we tested for association of the genotype with each inverse normalized variable using linear regression with a main effect for SNP and SNP-by-sex interaction terms using an additive model. We adjusted for covariates: age, age2, sex, six (within UK) ancestry principal components, and array used to measure genotypes. A maximum of 118,193 individuals (62,165 females and 56,027 males) with genotype and valid BMI, height, waist and hip were available for analysis. For evaluating the SNP-by-sex interaction effect of rs4731702 on risk of T2D, we considered a total of 13,728 T2D cases and 129,911 controls from 3 studies from populations of European ancestry. Sample and variant quality control was performed within each study (see Supplementary Methods). Within each study, the variant was tested for T2D association under an additive model after adjustment for study-specific covariates, including principal components to adjust for population structure and with a main effect for SNP and SNP-by-sex interaction terms. We then combined association summary statistics for the variant across studies via fixed-effects inverse-variance weighted meta-analysis.

Animal experiments information

CRISPR-Cas9 mice were generated in the Harvard University Genome Modification Facility and were kept and studied in accordance with Harvard University's Faculty of Arts and Sciences Institutional Animal Care and Use Committee (IACUC) protocol 14-05-202. *Klf14*$^{tm1(KOMP)Vlcg}$ mice were kept and studied in accordance with UK Home Office legislation and local ethical guidelines issued by the Medical Research Council (Responsibility in the Use of Animals for Medical Research, July 1993; home office license 30/3146).

Experimental design of in vivo mouse experiments in CRISPR-Cas9 mouse models

The CRISPR-Cas9 knockout mouse models were generated as described in the Supplementary Note. All procedures used for CRISPR-Cas9 animal studies were approved by Harvard University's Faculty of Arts and Sciences Institutional Animal Care and Use Committee and were consistent with local, state, and federal regulations as applicable. All

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

animals were housed in a 12-hour dark-light cycle and fed standard chow *ad libitum* in conditions similar to those described above. The control mice and KO mice were littermates from several litters that were timed to be the same age; the mice were kept in the same cages, in as few cages as permitted by the restrictions set forth by institutional policy (typically no more than three adult mice per cage).

For lipid measurements, blood samples were collected from the lateral tail vein following an overnight fast. Samples were kept on ice then centrifuged for 10 minutes at 2,000xg in a centrifuge at 4°C. The plasma levels of triglyceride, total cholesterol, and HDL-C were measured using Infinity Triglycerides Reagent (Thermo Fisher), Infinity Cholesterol Reagent (Thermo Fisher) and HDL-Cholesterol E (Wako Diagnostics) independently according to the manufacturers' instructions. For the intraperitoneal glucose tolerance test (IPGTT), the mice were fasted overnight, and glucose concentrations were measured immediately before and at 20, 40, 60, and 120 minutes after intraperitoneal injection of glucose (1g/kg body weight). For the insulin tolerance test (ITT), the mice were fasted for 4 hours, and glucose concentrations were measured immediately before and at 20, 40, 60, and 120 minutes after intraperitoneal injection with human insulin (Sigma) at 1U/kg body weight. Whole blood glucose levels were measured using a glucose meter and test strips (OneTouch). For histological examination, fat tissue was fixed in 4% paraformaldehyde and embedded in paraffin. Consecutive sections were cut and stained by hematoxylin and eosin for light microscopy examination and evaluation.

Experimental design of in vivo mouse experiments in deletion mouse models

Klf14$^{tm1(KOMP)Vlcg}$ deletion mice were generated as described in the Supplemental Note. Deletion mice were kept under controlled light (light 7am–7pm, dark 7pm–7am), temperature (21±2°C) and humidity (55±10%) conditions. They had free access to water (9–13 ppm chlorine) and were fed *ad libitum* on a commercial diet (SDS Rat and Mouse No. 3 Breeding diet, RM3) until 18 weeks of age when they were then transferred to a high fat diet (45% kcal from fat; D12451; Research Diets).

Cohorts of male and female mice were bred for longitudinal blood and body composition based phenotyping tests. These included four groups: 1) MAT (heterozygotes (hets) inheriting the allele maternally), 2) PAT (hets inheriting the allele paternally), 3) WT MAT controls and 4) WT PAT controls. Sample size estimates were based on previous experience and data from other mouse models in which the relevant traits were measured. These data were used in power calculations to help in deciding cohort sizes. Mice were housed in single sex groups of mixed genotype across multiple litters and were not randomised into groups. Mouse IDs and genotypes were recorded on the cages and were not blinded to the operator carrying out the animal procedure although subsequent tests only include animal ID information and clinical chemistry was carried out by a core service without knowledge of the genotypes.

Echo-MRI analysis and fasted blood sampling from the lateral tail vein (local anaesthetic) were carried out at 8, 12, 22 and 27 weeks of age. Whole blood glucose was measured using an AlphaTRAK meter and test strips (Abbott). Plasma insulin was assayed using a Mouse insulin ELISA kit (Mercodia). Terminal blood samples were collected from 33 week old

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

mice by intraperitoneal anaesthesia and cardiac puncture following an overnight fast. Samples were kept on ice then centrifuged for 10 minutes at 8,000xg in a centrifuge set to room temperature. The resulting plasma was analysed on board a Beckman Coulter AU680 clinical chemistry analyser using reagents and settings recommended by the manufacturer. Clinical biochemistry for cholesterol analysis was conducted on 8, 22, 27 and 33 week blood samples.

Statistical analysis of mouse data

Data collection, summary calculations and descriptive statistics were carried out using Microsoft Excel 2010. Statistical analysis was carried out using software Graphpad Prism v6. Effects of genotype at different time-points on clinical chemistry parameters were determined by two way ANOVA with repeated measures and Bonferroni's correction for multiple comparisons. Unpaired 2-tailed t tests were used for pairwise comparisons as appropriate and variances were equal, other than in the case of figure 4a male data where a Welch's correction was applied for unequal variance. Data sets were tested for normal distribution and equal variance according to a D'Agostino-Pearson omnibus K2 normality test and a Brown-Forsythe test respectively in Graphpad Prism. Consequently, non-parametric tests were applied as necessary and indicated in the figures. In these cases area under the curve was calculated (baselined to t=0 values) and tested with either a 1-way ANOVA Kruskal-Wallis test and Dunns multiple comparison test where there were multiple groups (*Klf14*$^{tm1(KOMP)Vlcg}$ knockout mice) or Mann-Whitney 2-tailed t-test when there were only two groups (CRISPR-Cas KO). Further as indicated, individual pairwise comparisons were made using Mann-Whitney 2-tailed t-tests and shown in the figures.

Gene expression in mouse models

Mouse RNAseq data from was collected as described in the Supplemental Note and analysed by the in house bioinformatics team at MRC Harwell using a previously described pipeline55. Functional enrichment of mouse differentially expressed genes was assessed using ToppFun24, with enrichments considered significant after FDR multiple testing correction ($q<0.05$). MEME_chip was used to identify motifs that were significantly over-represented in DNase1 hypersensitivity hotspots upstream of mouse differentiated genes.

Isolation, culture, and differentiation of pre-adipocytes

Abdominal subcutaneous adipose tissue (ASAT) biopsies were obtained from 8 healthy subjects (four females and four males) recruited from the Oxford BioBank (OBB) (http://www.oxfordbiobank.org.uk), aged 30-50 years with BMI ranging from 22-27kg/m^2. The study was approved by Oxfordshire Clinical Research Ethics Committee (08/H0606/107+5) and all subjects gave written informed consent. Primary pre-adipocytes were isolated as previously described (Collins *et al.*56) and cultured in DMEM/F12 Ham nutrient mixture (DMEM/F12), 10% FBS, 2 mmol/L glutamine, 0.25 ng/mL fibroblast growth factor, 100 units/mL penicillin, and 0.1 mg/mL streptomycin. Confluent pre-adipocytes were then stimulated for 14 days with an adipogenic cocktail comprising DMEM/F12, 2mmol/L glutamine, 17mmol/L pantothenate, 100nmol/L human insulin, 10nmol/L triiodo-L-thyronine, 33mmol/L biotin, 10mg/mL transferrin, 1mmol/L dexamethasone, 100 units/mL penicillin, and 0.1mg/mL streptomycin. For the first 4 days, 0.25mmol/L 3-isobutyl-1-

methylxanthine and 4mmol/L troglitazone were added. The cells were harvested every day during proliferation and 14 days during differentiation.

RNA isolation and Real-time PCR in pre-adipocytes

Total RNA was extracted from pre-adipocytes as described by Collins et al56. Real time PCR was performed on an Applied Biosystems 7900HT, using TaqMan Assays-on-Demand (Applied Biosystems) and Klear Kall Master Mix (KBiosciences). mRNA expression values for target genes were calculated using the DCt transformation method57. The DCt was calculated as follows: DCt = efficiency (minimum Ct–sample Ct). Values were normalized to endogenous control genes (PPIA and UBC)58.

Explant lipogenesis experiments

Adipocytes isolated from abdominal subcutaneous adipose tissue were used to measure activation of insulin-stimulated lipogenesis, as described in detail elsewhere59. In brief, isolated human adipocytes were incubated at a concentration of 2% (v/v) in Krebs–Ringer phosphate buffer (pH 7.4) containing albumin (40mg/ml), [3-3H] glucose (5×10^5dpm/ml), unlabelled glucose (1μmol/l) and varying concentrations of human insulin (0-70nmol/l). The incubations were conducted for 2 h at 37°C with air as the gas phase. Incubations were stopped by rapidly chilling the incubation vials to 4°C and the incorporation of radiolabelled glucose into adipocyte lipids (i.e. lipogenesis) was determined. Lipogenesis was expressed as the amount of glucose incorporated either per lipid weight of fat cells or per fat cell number, as described previously59.

Short Hairpin RNA–Mediated Silencing of *KLF14*

KLF14 was silenced in primary pre-adipocytes derived from female ASAT. Lentiviral particles were produced by transient co-transfection of HEK293 cells, using *KLF14*-short hairpin RNA (shRNA) lentiviral transduction construct (SHCLNV-NM_138693.1-255s1c1; Sigma-Aldrich) and non-target shRNA lentiviral constructs (SHC002V; Sigma-Aldrich) with packaging vectors (MISSION [Sigma-Aldrich]). Cells were stably transfected by transduction of pre-adipocytes with lentiviral particles followed by selection in growth media containing 2mg/ml puromycin.

Quantification of intracellular lipid content

Control and *KLF14* shRNA cell lysates were prepared in lysis buffer containing 1% IGEPAL-630, 150mM NaCl and 50mM Tris HCl. Lysates were sonicated and an aliquot was collected for protein quantification using the BioRad Dc Protein Assay kit. The remaining lysate was heated at 95°C for 30 minutes, allowed to cool to room temperature, and then centrifuged at 12,000xg for 10 minutes. Triacylglycerol concentration in the supernatant was determined using a commercially available enzymatic kit (Instrumentation Laboratory UK) on an iLAB 650 Chemistry Analyser. Total triacylglycerol was normalised to protein concentration.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Estimation of Cell Doubling Time

Equal number of control and *KLF14* shRNA cells were seeded in T75 flasks, $1.5x10^5$ cells. Cells were trypsinized and double counted every 5 days. Doubling time was calculated using the formula $T_d = (t_2 - t_1) \times [\log(2) \div \log(q_2 \div q_1)]$, where t = time (days) and q = cell number.

Insulin stimulated glucose uptake—Glucose uptake was assayed according to the established protocol from a commercial glucose uptake kit (J1342; Promega). The luminescent glucose uptake assay was applied to control and KLF14 shRNA differentiated adipocytes in 96-well plates. In brief, differentiated cells were starved in hormones-free DMEM/F12 medium overnight. The medium was removed and the cells were washed with 100μl of phosphate-buffered saline (PBS) followed with 10nM insulin incubation for 1h at 37°C in 5%CO_2. To initiate glucose uptake, 50μl of 2DG (1mM) in PBS was added to the cells. The uptake reaction was stopped, and the samples were processed as described in the protocol. All assay steps were performed at room temperature. All data were acquired on a PerkinElmer EnSpire 2300 multimode plate reader instrument, with an integration time of 0.5s.

Adipocyte Cell Size and Number Study Population and Sample Collection

Nine females and 9 males carrying *KLF14* risk allele and paired control subjects age (30-50) and BMI (22-27kgm^{-2}) matched were recruited from the OBB. Paired abdominal and gluteal subcutaneous adipose specimens were obtained by gun and needle biopsies. The gun biopsies were fixed in 10% paraformaldehyde, embedded in paraffin wax, cut into 5 μm sections, and stained with hematoxylin and eosin. Sections were viewed at 20x magnification, and adipocyte cross-sectional area was calculated using Adobe Photoshop 5.0.1 (Adobe Systems, San Jose, CA) and Image Processing Tool Kit (Reindeer Games, Gainesville, FL). As previously described60, in order to accurately determine the minimum number of cells required for measurement of cell size distribution in a sample, we took 4 samples and counted 1,000 cells in each. Data were removed by 100 at a time and it was observed that the coefficient of variation started to increase when fewer than 100 cells were included in each biopsy. Therefore, we included only biopsies with more than 100 cells available for quantification (N=9 pairs from each genotype). Statistical significance was assessed using a Wilcoxon signed-rank test. Replication of the cell size was undertaken in a Swedish cohort61. Abdominal subcutaneous adipose specimens were obtained by needle biopsy. Adipocytes were separated from stroma cells by treatment in a shaking bath at 37°C for 60 min with collagenase (0.5mg/L) in 5ml Krebs Ringer phosphate buffer (pH7.4) with purified BSA (40g/L) as previously described59. Adipocyte suspensions were then rinsed three times in collagenase-free buffer using nylon filters, and the cell sizes were measured by direct microscopy. The mean adipocyte diameter was calculated from measurements of 100 cells and statistical significance assessed using a Wilcoxon signed-rank test. Cell numbers were calculated using the formula described by Hirsch, *et al*62 where DXA measured adipose depot size was divided by mean cell volume.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Date Availability Statement

TwinsUK RNAseq data is available from EGA (Accession number: EGAS00001000805). TwinsUK adipose methylation data is available from ArrayExpress (E-MTAB-1866), and blood methylation data from GEO (GSE50660). TwinsUK genotypes are available upon application to the TwinsUK cohort. METSIM adipose array data is available from GEO(GSE70353).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was supported by MRC grant MR/J010642/1 to K.S.S., R.D.C., F.K. and M.I.M.. K.S.S. is supported by a MRC New Investigator Award (MR/L01999X/1). M.I.M. is a Wellcome Senior Investigator and is supported by Wellcome (090532, 106130, 098381, 203141), NIDDK (U01-DK105535) and the MRC (MR/L020149/1). F.K. is supported by British Heart Foundation (RG/17/1/32663). M.C. is supported by NIH Award R00 HL121172. A.L.G. is a Wellcome Senior Fellow in Basic Biomedical Science (095101/Z/10/Z and 200837/Z/16/Z). K.M. is supported by NIH award R01-DK099571 and A.J.L. by NIH award NIH P01HL28481. R.D.C. is supported by MRC MC_U142661184. J.F. was a Marie Curie Fellow. M.L. is supported by Academy of Finland grants 77299 and 124243; the Finnish Heart Foundation; the Finnish Diabetes Foundation; Commission of the European Community HEALTH-F2-2007-201681. A.V. and A.B. were supported by the European Union Framework Programme 7 grant EuroBATS (259749). Some computations were performed at the Vital-IT center for high-performance computing of the Swiss Institute of Bioinformatics (SIB; http://www.vital-it.ch/). The TwinsUK study was funded by Wellcome and European Community's Seventh Framework Programme (FP7/2007-2013). The TwinsUK study also receives support from the National Institute for Health Research (NIHR)- funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. We acknowledge excellent technical support for animal husbandry (Mary Lyon Centre), genotyping, histology and pathology. We thank the volunteers from the Oxford Biobank, NIHR Oxford Biomedical Research Centre, for their participation. The Oxford Biobank (www.oxfordbiobank.org.uk) is also part of the NIHR National Bioresource which supported the recalling process of the volunteers. GERA data came from a grant, the Resource for Genetic Epidemiology Research in Adult Health and Aging (RC2 AG033067; Schaefer and Risch, PIs) awarded to the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH) and the UCSF Institute for Human Genetics. The RPGEH was supported by grants from the Robert Wood Johnson Foundation, the Wayne and Gladys Valley Foundation, the Ellison Medical endonte

References

1. Kong A, et al. Parental origin of sequence variants associated with complex diseases. Nature. 2009; 462:868–874. doi:nature08625 [pii]. DOI: 10.1038/nature08625 [PubMed: 20016592]
2. Voight BF, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet. 42:579–589. doi:ng.609 [pii]. DOI: 10.1038/ng.609
3. Small KS, et al. Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. Nat Genet. 43:561–564. doi:ng.833 [pii]. DOI: 10.1038/ng.833
4. Dang DT, Pevsner J, Yang VW. The biology of the mammalian Kruppel-like family of transcription factors. Int J Biochem Cell Biol. 2000; 32:1103–1121. doi:S1357-2725(00)00059-5 [pii]. [PubMed: 11137451]
5. Wu Z, Wang S. Role of kruppel-like transcription factors in adipogenesis. Developmental biology. 2013; 373:235–243. DOI: 10.1016/j.ydbio.2012.10.031 [PubMed: 23142072]
6. Parker-Katiraee L, et al. Identification of the imprinted KLF14 transcription factor undergoing human-specific accelerated evolution. PLoS Genet. 2007; 3:e65. doi:06-PLGE-RA-0297R3 [pii]. doi: 10.1371/journal.pgen.0030065 [PubMed: 17480121]
7. Buil A, et al. Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. Nat Genet. 2015; 47:88–91. DOI: 10.1038/ng.3162 [PubMed: 25436857]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

8. Greenawalt DM, et al. A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. Genome Res. 2011; 21:1008–1016. DOI: 10.1101/gr.112821.110 [PubMed: 21602305]
9. Civelek M, et al. Genetic Regulation of Adipose Gene Expression and Cardio-Metabolic Traits. American journal of human genetics. 2017; 100:428–443. DOI: 10.1016/j.ajhg.2017.01.027 [PubMed: 28257690]
10. Emilsson V, et al. Genetics of gene expression and its effect on disease. Nature. 2008; 452:423–428. doi:nature06758 [pii]. DOI: 10.1038/nature06758 [PubMed: 18344981]
11. Keildson S, et al. Expression of phosphofructokinase in skeletal muscle is influenced by genetic variation and associated with insulin sensitivity. Diabetes. 2014; 63:1154–1165. DOI: 10.2337/db13-1301 [PubMed: 24306210]
12. Innocenti F, et al. Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. PLoS Genet. 2011; 7:e1002078.doi: 10.1371/journal.pgen.1002078 [PubMed: 21637794]
13. van de Bunt M, et al. Transcript Expression Data from Human Islets Links Regulatory Signals from Genome-Wide Association Studies for Type 2 Diabetes and Glycemic Traits to Their Downstream Effectors. PLoS Genet. 2015; 11:e1005694.doi: 10.1371/journal.pgen.1005694 [PubMed: 26624892]
14. Consortium GT. et al. Genetic effects on gene expression across human tissues. Nature. 2017; 550:204–213. DOI: 10.1038/nature24277 [PubMed: 29022597]
15. Replication DIG. et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat Genet. 2014; 46:234–244. DOI: 10.1038/ng.2897 [PubMed: 24509480]
16. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. DOI: 10.1038/nature11247 [PubMed: 22955616]
17. Roadmap Epigenomics C. et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015; 518:317–330. DOI: 10.1038/nature14248 [PubMed: 25693563]
18. Fairfax BP, et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. Nat Genet. 2012; 44:502–510. DOI: 10.1038/ng.2205 [PubMed: 22446964]
19. Naranbhai V, et al. Genomic modulators of gene expression in human neutrophils. Nature communications. 2015; 6:7545.doi: 10.1038/ncomms8545
20. Rotival M, et al. Integrating genome-wide genetic variations and monocyte expression data reveals trans-regulated gene modules in humans. PLoS Genet. 2011; 7:e1002367.doi: 10.1371/journal.pgen.1002367 [PubMed: 22144904]
21. Hannum G, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Molecular cell. 2013; 49:359–367. DOI: 10.1016/j.molcel.2012.10.016 [PubMed: 23177740]
22. Ronn T, et al. Impact of age, BMI and HbA1c levels on the genome-wide DNA methylation and mRNA expression patterns in human adipose tissue and identification of epigenetic biomarkers in blood. Human molecular genetics. 2015; 24:3792–3813. DOI: 10.1093/hmg/ddv124 [PubMed: 25861810]
23. Najafabadi HS, et al. C2H2 zinc finger proteins greatly expand the human regulatory lexicon. Nature biotechnology. 2015; 33:555–562. DOI: 10.1038/nbt.3128
24. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. 2009; 37:W305–311. DOI: 10.1093/nar/gkp427 [PubMed: 19465376]
25. Garvey WT, et al. Pretranslational suppression of a glucose transporter protein causes insulin resistance in adipocytes from patients with non-insulin-dependent diabetes mellitus and obesity. The Journal of clinical investigation. 1991; 87:1072–1081. DOI: 10.1172/JCI115068 [PubMed: 1999488]
26. Horikoshi M, et al. Discovery and Fine-Mapping of Glycaemic and Obesity-Related Trait Loci Using High-Density Imputation. PLoS Genet. 2015; 11:e1005230.doi: 10.1371/journal.pgen.1005230 [PubMed: 26132169]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

27. Global Lipids Genetics C. et al. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013; 45:1274–1283. DOI: 10.1038/ng.2797 [PubMed: 24097068]
28. Teslovich TM, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 466:707–713. doi:nature09270 [pii]. DOI: 10.1038/nature09270
29. Locke AE, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015; 518:197–206. DOI: 10.1038/nature14177 [PubMed: 25673413]
30. Morris AP, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet. 2012; 44:981–990. DOI: 10.1038/ng.2383 [PubMed: 22885922]
31. Shungin D, et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature. 2015; 518:187–196. DOI: 10.1038/nature14132 [PubMed: 25673412]
32. Dupuis J, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet. 42:105–116. doi:ng.520 [pii]. DOI: 10.1038/ng.520
33. Speliotes EK, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. PLoS Genet. 2011; 7:e1001324.doi: 10.1371/journal.pgen.1001324 [PubMed: 21423719]
34. Chambers JC, et al. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet. 2011; 43:1131–1138. DOI: 10.1038/ng.970 [PubMed: 22001757]
35. Guo Y, et al. Perhexiline activates KLF14 and reduces atherosclerosis by modulating ApoA-I production. The Journal of clinical investigation. 2015; 125:3819–3830. DOI: 10.1172/JCI79048 [PubMed: 26368306]
36. Arner E, et al. Adipocyte turnover: relevance to human adipose tissue morphology. Diabetes. 2010; 59:105–109. DOI: 10.2337/db09-0942 [PubMed: 19846802]
37. Hammarstedt A, Graham TE, Kahn BB. Adipose tissue dysregulation and reduced insulin sensitivity in non-obese individuals with enlarged abdominal adipose cells. Diabetology & metabolic syndrome. 2012; 4:42.doi: 10.1186/1758-5996-4-42 [PubMed: 22992414]
38. Lonn M, Mehlig K, Bengtsson C, Lissner L. Adipocyte size predicts incidence of type 2 diabetes in women. FASEB journal: official publication of the Federation of American Societies for Experimental Biology. 2010; 24:326–331. DOI: 10.1096/fj.09-133058 [PubMed: 19741173]
39. Lundgren M, et al. Fat cell enlargement is an independent marker of insulin resistance and 'hyperleptinaemia'. Diabetologia. 2007; 50:625–633. DOI: 10.1007/s00125-006-0572-1 [PubMed: 17216279]
40. Weyer C, Foley JE, Bogardus C, Tataranni PA, Pratley RE. Enlarged subcutaneous abdominal adipocyte size, but not obesity itself, predicts type II diabetes independent of insulin resistance. Diabetologia. 2000; 43:1498–1506. DOI: 10.1007/s001250051560 [PubMed: 11151758]
41. Snijder MB, et al. Independent and opposite associations of waist and hip circumferences with diabetes, hypertension and dyslipidemia: the AusDiab Study. International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity. 2004; 28:402–409. DOI: 10.1038/sj.ijo.0802567
42. Yusuf S, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case-control study. Lancet. 2005; 366:1640–1649. DOI: 10.1016/S0140-6736(05)67663-5 [PubMed: 16271645]
43. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nature methods. 2012; 9:215–216. DOI: 10.1038/nmeth.1906 [PubMed: 22373907]
44. Grundberg E, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat Genet. 2012; 44:1084–1089. DOI: 10.1038/ng.2394 [PubMed: 22941192]
45. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics. 2012; 28:1353–1358. DOI: 10.1093/bioinformatics/bts163 [PubMed: 22492648]
46. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A. 2003; 100:9440–9445. DOI: 10.1073/pnas.1530509100 [PubMed: 12883005]
47. qvalue: Q-value estimation for false discovery rate control. v. R package version 1.40.0

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

48. Grundberg E, et al. Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. American journal of human genetics. 2013; 93:876–890. DOI: 10.1016/j.ajhg.2013.10.004 [PubMed: 24183450]
49. Tsaprouni LG, et al. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. Epigenetics. 2014; 9:1382–1396. DOI: 10.4161/15592294.2014.969637 [PubMed: 25424692]
50. Pierce BL, et al. Mediation analysis demonstrates that trans-eQTLs are often explained by cis-mediation: a genome-wide analysis among 1,800 South Asians. PLoS Genet. 2014; 10:e1004818.doi: 10.1371/journal.pgen.1004818 [PubMed: 25474530]
51. Najafabadi HS, Albu M, Hughes TR. Identification of C2H2-ZF binding preferences from ChIP-seq data using RCADE. Bioinformatics. 2015; 31:2879–2881. DOI: 10.1093/bioinformatics/btv284 [PubMed: 25953800]
52. Janky R, et al. iRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS computational biology. 2014; 10:e1003731.doi: 10.1371/journal.pcbi.1003731 [PubMed: 25058159]
53. Frith MC, Li MC, Weng Z. Cluster-Buster: Finding dense clusters of motifs in DNA sequences. Nucleic Acids Res. 2003; 31:3666–3668. [PubMed: 12824389]
54. Welter D, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014; 42:D1001–1006. DOI: 10.1093/nar/gkt1229 [PubMed: 24316577]
55. Parsons MJ, et al. The Regulatory Factor ZFHX3 Modifies Circadian Function in SCN via an AT Motif-Driven Axis. Cell. 2015; 162:607–621. DOI: 10.1016/j.cell.2015.06.060 [PubMed: 26232227]
56. Collins JM, Neville MJ, Hoppa MB, Frayn KN. De novo lipogenesis and stearoyl-CoA desaturase are coordinately regulated in the human adipocyte and protect against palmitate-induced cell injury. The Journal of biological chemistry. 2010; 285:6044–6052. DOI: 10.1074/jbc.M109.053280 [PubMed: 20032470]
57. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001; 29:e45. [PubMed: 11328886]
58. Neville MJ, Collins JM, Gloyn AL, McCarthy MI, Karpe F. Comprehensive human adipose tissue mRNA and microRNA endogenous control selection for quantitative real-time-PCR normalization. Obesity. 2011; 19:888–892. DOI: 10.1038/oby.2010.257 [PubMed: 20948521]
59. Lofgren P, Hoffstedt J, Naslund E, Wiren M, Arner P. Prospective and controlled studies of the actions of insulin and catecholamine in fat cells of obese women following weight reduction. Diabetologia. 2005; 48:2334–2342. DOI: 10.1007/s00125-005-1961-6 [PubMed: 16222518]
60. Marinou K, et al. Structural and functional properties of deep abdominal subcutaneous adipose tissue explain its association with insulin resistance and cardiovascular risk in men. Diabetes care. 2014; 37:821–829. DOI: 10.2337/dc13-1353 [PubMed: 24186879]
61. Dahlman I, et al. Numerous Genes in Loci Associated With Body Fat Distribution Are Linked to Adipose Function. Diabetes. 2016; 65:433–437. DOI: 10.2337/db15-0828 [PubMed: 26798124]
62. Hirsch J, Gallian E. Methods for the determination of adipose cell size in man and animals. Journal of lipid research. 1968; 9:110–119. [PubMed: 4295346]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

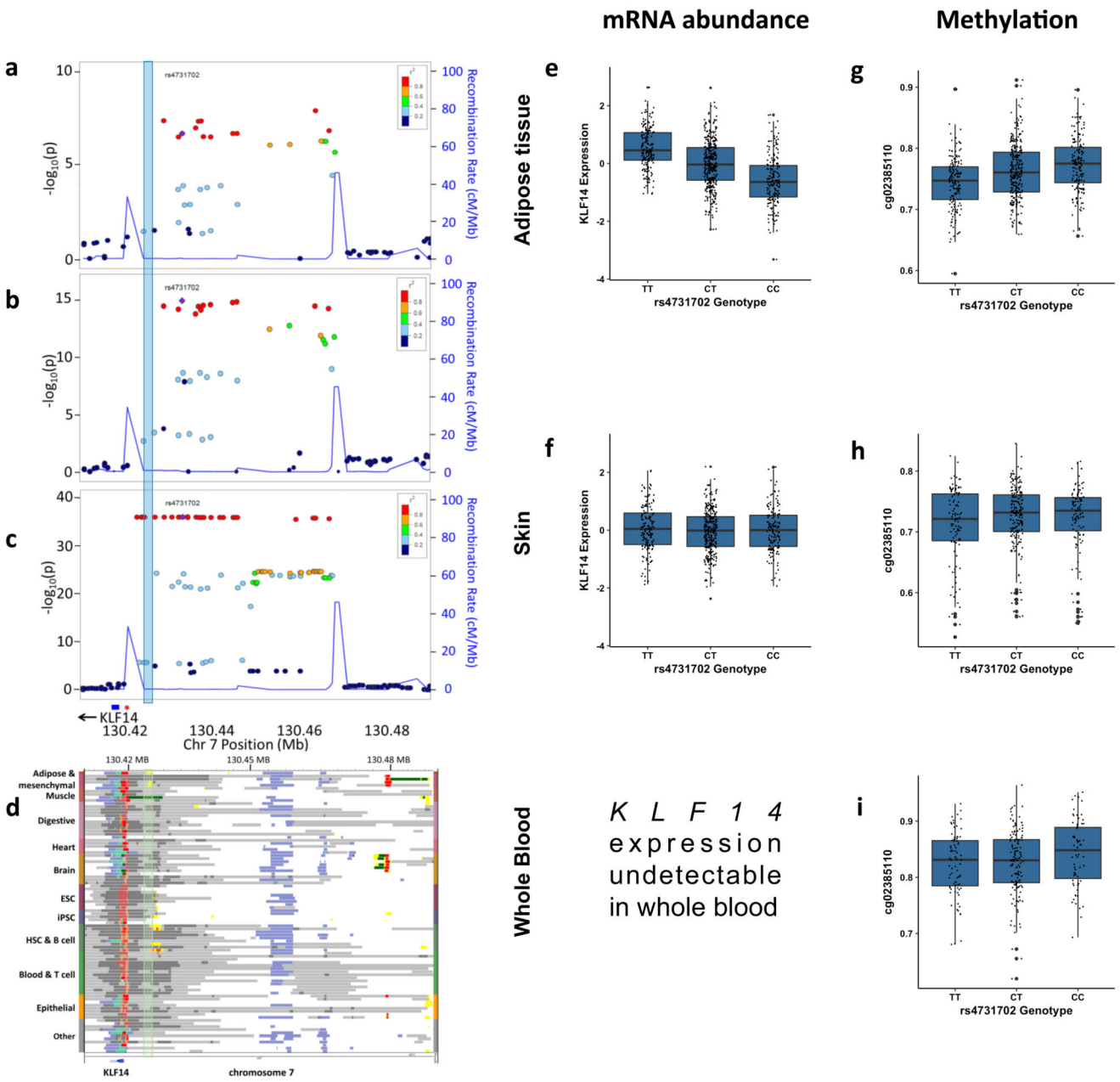

Figure 1. Cell-type-specific enhancer in the risk haplotype regulates *KLF14* expression.

Association of *KLF14* locus variants with (a) type 2 diabetes (N=69,033) (Morris et al39), (b) high density lipoprotein levels (N=99,900) (Teslovich et al27); (c) *KLF14* mRNA abundance in adipose tissue of 776 women (TwinsUK). Circles represent genotyped and imputed DNA variants and are colored by LD r^2 values with the index SNP rs4731702. The red asterisk indicates the position of methylation probe cg02385110. Association between genotype and *KLF14* expression was assessed using mixed effects models, correcting for family structure and relevant covariates (methods). Uncorrected *P*-values displayed. (d) Chromatin state annotations for the locus across 93 reference epigenomes (rows) for cell and tissue types profiled by the Roadmap Epigenomics and ENCODE Projects. Colors

correspond to chromatin states; yellow=enhancer, red=promoter as detailed in Ernst et al43. ESC, embryonic stem cell; HSC, hematopoietic stem cell; iPSC, induced pluripotent stem cell. Blue box contains the putative enhancer. Chromosome coordinates correspond to UCSC Genome Browser build hg19. (e-i) Boxplots of *KLF14* mRNA abundance (e-f) and methylation of probe cg02385110 (g-i) in concurrently-sampled adipose ($N_{Expression}$=720; $N_{Methylation}$=595), skin ($N_{Expression}$=606; $N_{Methylation}$=414) and whole blood ($N_{Expression}$=368; $N_{Methylation}$=289) from the TwinsUK study participants, based on rs4731702 genotype (CC=risk-allele homozygotes). *KLF14* expression was undetectable in whole blood. Boxplots display the median, with hinges corresponding to the first and third quartiles. Whiskers extend from the hinge to the largest value no further than 1.5xinter-quartile range, with values beyond shown as outliers. Associations assessed using linear regression, correcting for family structure and relevant covariates (methods).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

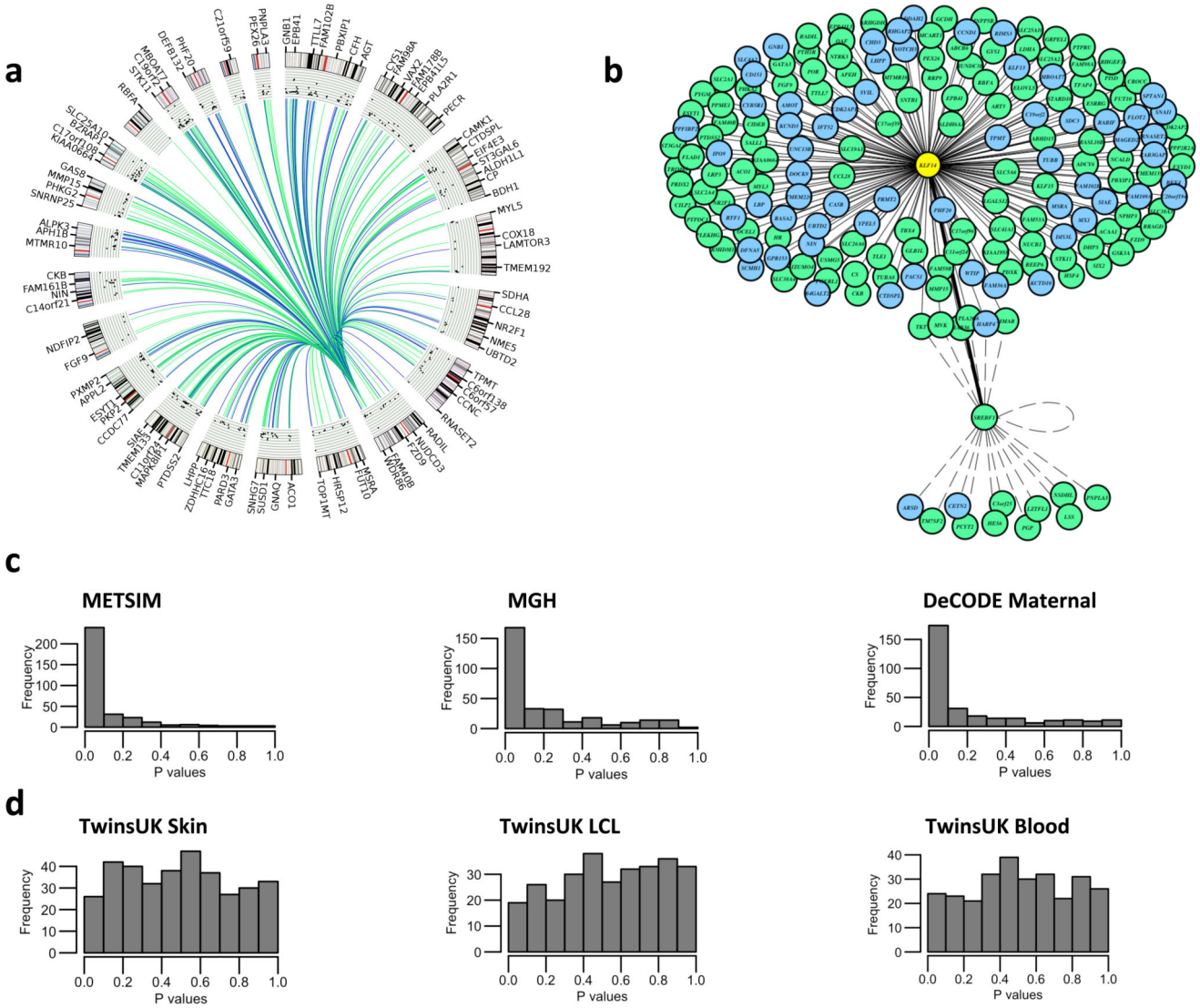

Figure 2. *KLF14 cis*-eQTL is an adipose-specific *trans*-regulator of a large network of genes. **a,** Genomic location of the 385 *trans*-genes. Line colour indicates direction of *trans*-effect; blue, positive association with T2D risk-allele; green, negative. **b,** 177 *trans*-genes are enriched for *KLF14* motifs and 18 *trans*-genes form a sub-network regulated by *SREBF1*. Genes are colored as in a) to indicate direction of *trans*-effect. Solid lines link *KLF14* to trans genes containing KLF14 binding sites, dashed lines link *SREBF1* to trans-genes containing SREBF1 binding sites **c,** *KLF14 trans*-network replicates in three independent adipose cohorts, METSIM (N=770), MGH (N=701) and deCODE (N=589). Histograms show p-value distribution of *trans*-genes in each replication cohort. **d,** *Trans*-network is not present in other tissues in the same TwinsUK samples as adipose discovery. Histograms show P-value distribution of *trans*-genes in skin (N=716), LCLs (N=814), and whole blood (N=384). In TwinsUK, METSIM and DeCODE, association between *KLF14* lead SNP rs4731702 genotype and gene expression was assessed using linear models with correction for relevant covariates (see methods), and using the Kruskal-Wallis test in the MGH dataset.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Uncorrected *P*-values displayed. *Trans*-gene associations were considered replicated at $P<0.05$.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

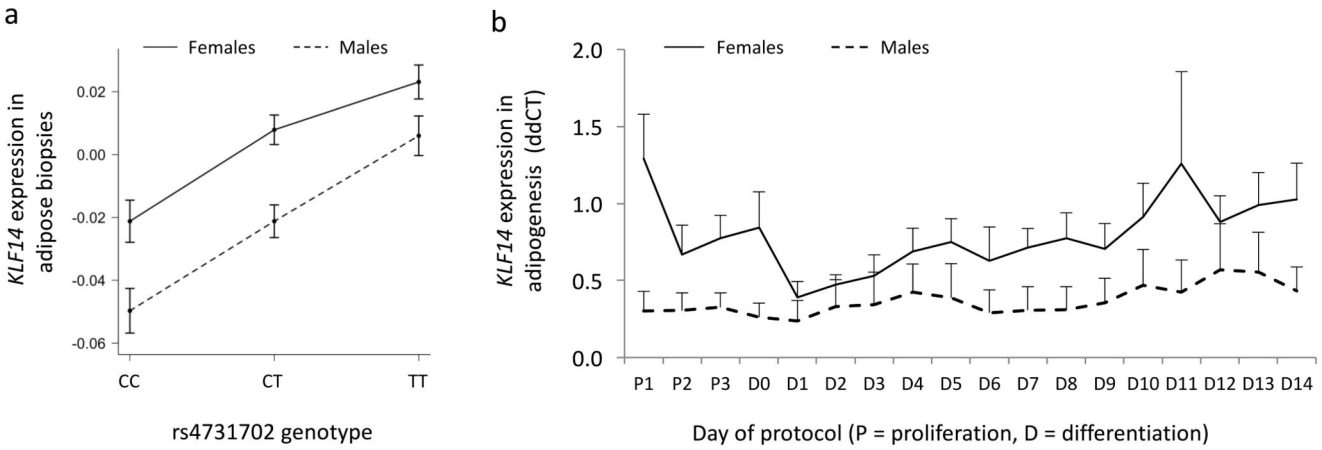

Figure 3. *KLF14* expression is sex-differentiated in biopsies and throughout adipocyte differentiation.

a, Adipose expression of *KLF14* in the deCODE cohort. The *KLF14 cis*-eQTL has a similar effect size in males (N=265) and females (N=376), but *KLF14* expression is higher in females than males across all genotype classes. Plots display *KLF14* mean relative expression ±SEM. **b,** Pre-adipocytes derived from abdominal biopsies and subjected to a 14 day-differentiation protocol show higher *KLF14* expression in females (N=4, mean±SEM) compared to males (N=4, mean±SEM) at all time points.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

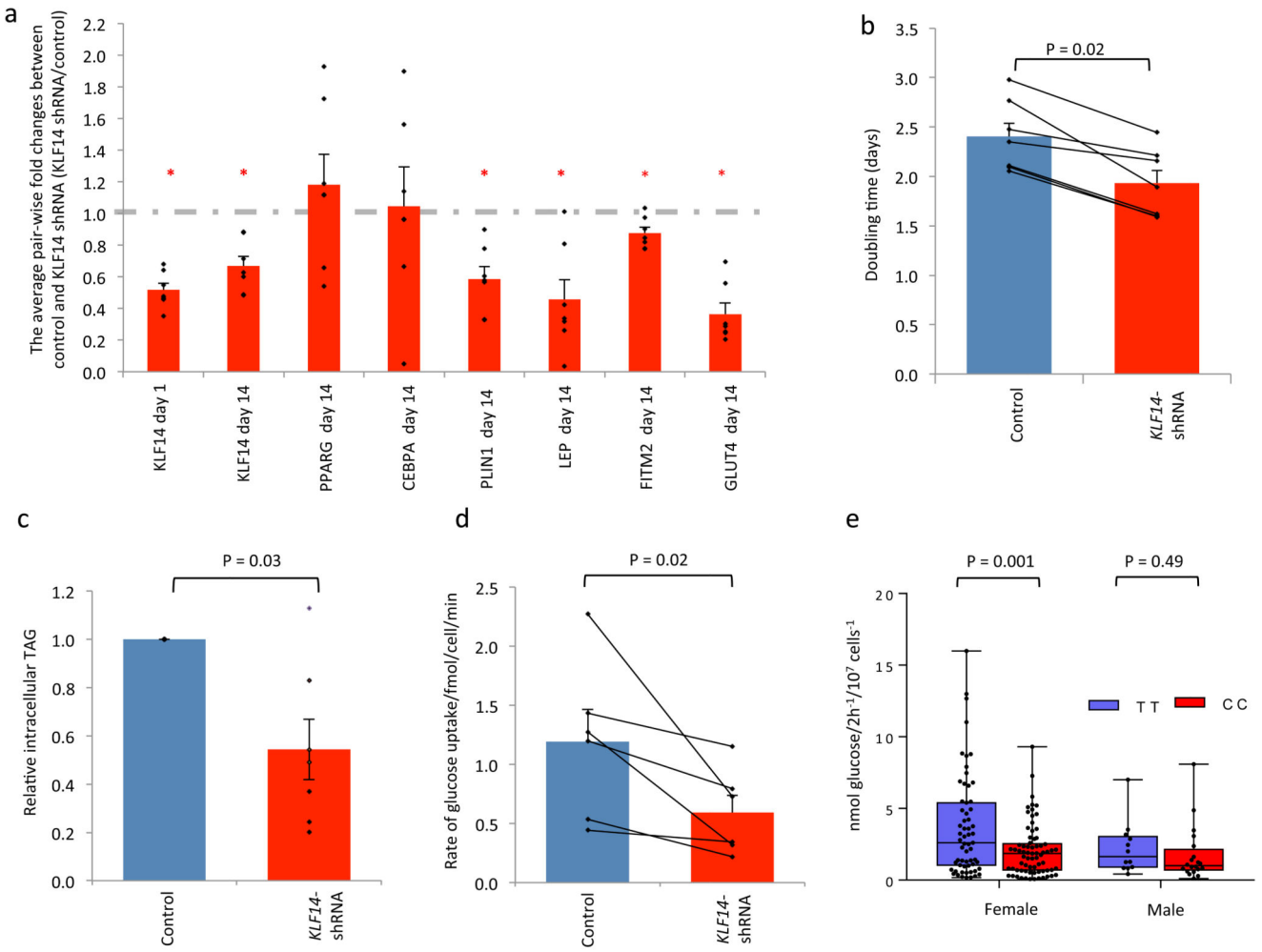

Figure 4. Adipose-specific knockout of *Klf14* in the mouse.

Clinical chemistry parameters were measured in female and male adipose-specific (*Adipoq*-Cre) knockout (KO) C57BL/6J *Klf14* mice and their *Adipoq*-Cre-expressing, wildtype (Wt) controls. Mice were fed a standard diet throughout their lifetimes. **a**, HDL-C at 16 weeks significantly reduced in female KO group mice compared to wildtype (P= 0.04); **b**, TG at 16 weeks significantly increased in the male KO group compared to wildtype (P=0.03 unpaired two-tailed t-test); **c and d**, glucose levels in an IPGTT at 12 weeks of age significantly increased in female (c) and male (d) KO compared to wildtype controls at multiple time-points; **e and f**, in an ITT at 12 weeks of age glucose reduced to a lesser extent in KO mice compared to their wildtype controls at multiple time-points in both females (e) and males (f). Pairwise significance as compared to controls is shown. Analysis in (a) and (b) use unpaired two-tailed t-test (for male HDL-C, with Welch's correction), and in (c) to (f) by 2-way ANOVA with repeated measures and Bonferroni correction (GraphPad Prism 6). Wildtype mice shown as blue lines and fill, females (N=8) and males (N=8). Adipose KO mice shown in red lines and fill, females (N=5) and males (N=6). All error bars are plotted as a mean central value plus or minus the standard deviation.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

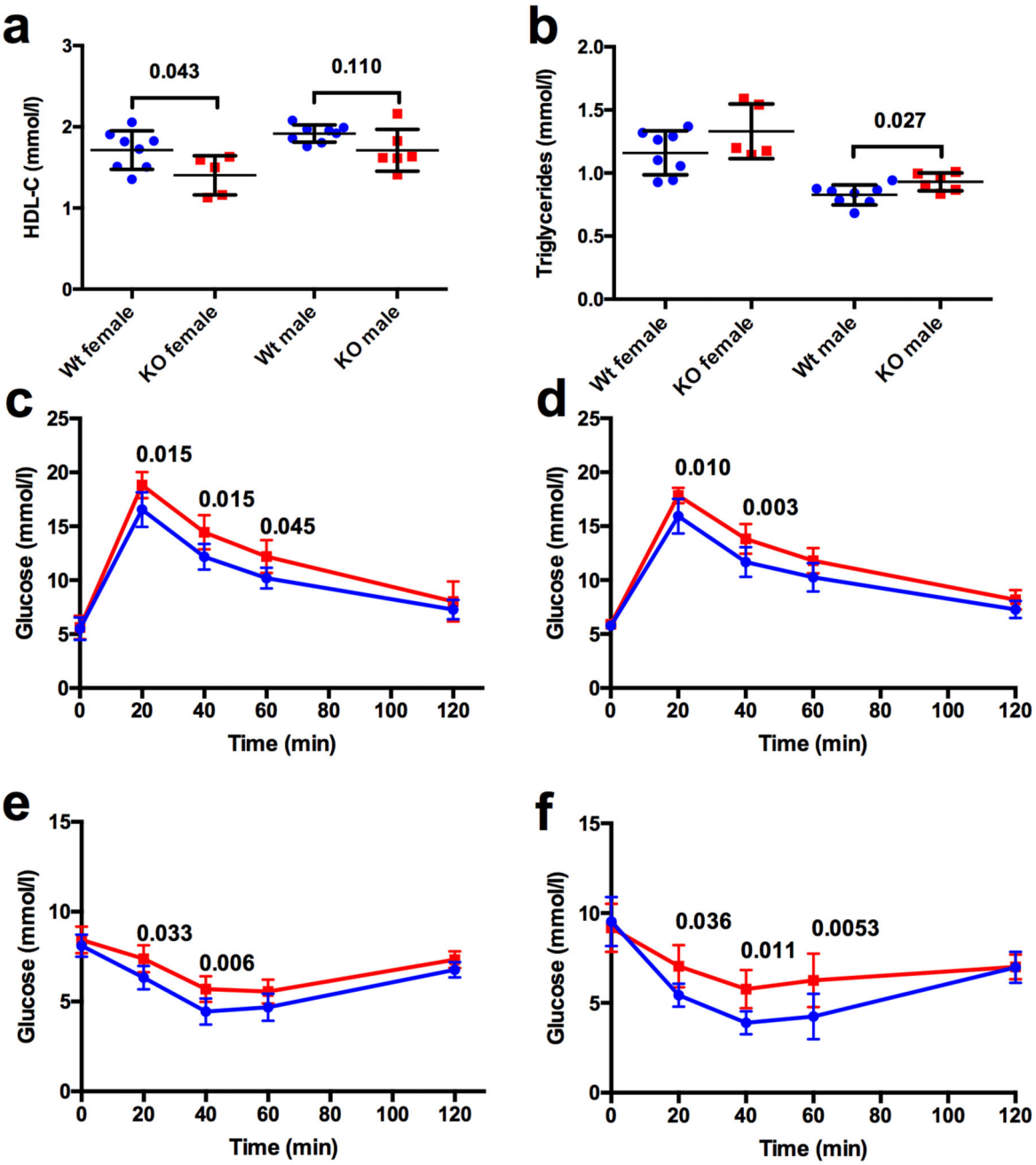

Figure 5. *KLF14* expression effects expression of adipocyte maturation marker genes and adipocyte function.

a, The mean fold change in expression of marker genes following *KLF14* shRNA knockdown. Mean fold change is plotted as the expression of the labeled gene in *KLF14* shRNA adipocytes divided by expression in control adipocytes. Expression of all genes was measured with Real Time PCR in seven paired knockdown/control replicates (N=7 ±SEM). A significant decrease in expression was seen for *KLF14*, measured at days 1 and 14. There was no change in early differentiation markers *PPARG2* and *CEBPA* at either day 1 or 14. A

significant decrease in expression was seen for late-stage differentiation markers *LEP* and *PLIN1* and late-stage adipogenic trans-genes *FITM2* and *GLUT4* at day 14. Statistical significance was assessed using a Wilcoxon signed-rank two-sided test. (**)* marks statistical differences between knockdown and control adipocytes for each gene ($p \leq 0.05$). **b,** *KLF14* shRNA pre-adipocytes isolated from female abdominal adipose tissue biopsies (N=7, mean ±SEM) have a significant decrease in doubling time compared to non-target shRNA controls. **c,** *KLF14* shRNA differentiated adipocytes have a significant decrease in TAG accumulation (N=7, mean±SEM) and **d,** in insulin stimulated glucose uptake (N=6, mean ±SEM) compared to non-target shRNA controls. Statistical significance was assessed using a Wilcoxon signed-rank two-sided test. **e,** Female *KLF14* risk allele homozygotes (CC) have a significant decrease in lipogenesis, as measured in adipose explants, compared to non-risk allele homozygotes (TT) (Female (N=132), Male (N=32), mean±SEM). The minimum and maximum values of lipogenesis are presented with their whiskers. The line in the box represents the median. Statistical significance was assessed using unpaired two-sided t-test.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

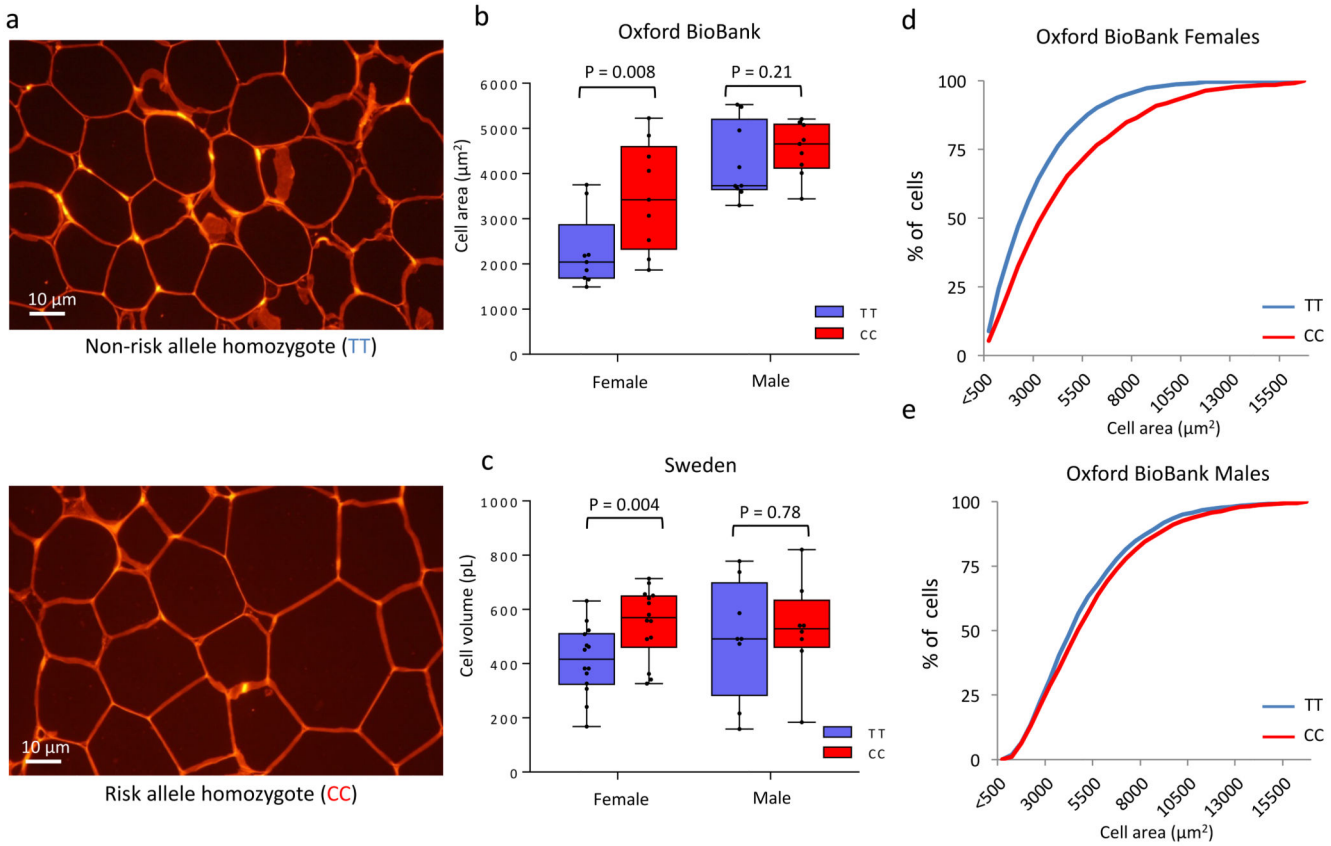

Figure 6. Adipose tissue of Type 2 diabetes risk allele homozygotes contains fewer, larger mature adipocytes compared to non-risk allele homozygotes.

a, Sections of subcutaneous abdominal adipose tissue biopsies from two age- and BMI-matched female volunteers. Histological sections were stained with hematoxylin and eosin. **b,** The minimum and maximum values of adipocyte cell surface area in abdominal adipose histological sections stratified by sex and genotype are presented with their whiskers. The line in the box represents the median. >100 cells were measured for each biopsy (N=18 per sex). Statistical significance was assessed using a Wilcoxon signed-rank two-sided test. **c,** The minimum and maximum values of adipocyte cell volume measured in dispersed adipocytes from collagenase-dispersed abdominal adipose tissue stratified by sex and genotype are presented with their whiskers (female N=28, male N=16, mean ± SEM). The line in the box represents the median. Statistical significance was assessed using a Wilcoxon signed-rank two-sided test. **d,** Cumulative frequency distribution of adipocyte cell surface area in females (N=18), measured as in **b**. **e,** Cumulative frequency distribution of adipocyte cell surface area in males (N=18), measured as in **b**.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Table 1

rs4731702 is associated with insulin resistance, lipid and body-composition phenotypes

HDL, high density lipoprotein; TG, Triglycerides; LDL, low density lipoprotein; BMI, body mass index; WHR, waist-hip ratio. Morris *et al*30 results are taken from the Stage 1 Discovery as Stage 2 included non-European samples. HDL and LDL effect sizes are reported in mmol/l, TG was log-transformed. Type 2 Diabetes effect sizes are Odds Ratios, not Betas.

Trait		Type 2 diabetes	Fasting insulin	Fasting glucose	HDL	TG	LDL	BMI	WHR	Waist	Hip
GWAS Citation		Morris, 2012	Horikoshi, 2015	Horikoshi, 2015	Teslovich, 2010	Teslovich, 2010	Teslovich, 2010	Locke, 2015	Shungin 2015	Shungin, 2015	Shungin, 2015
Sex-combined	β	1.10 (1.06-1.15)	0.019	0.007	-0.015	0.016	0.009	0.009	0.006	-0.009	-0.017
	P	3.6×10^{-8}	1.2×10^{-5}	0.07	1.2×10^{-15}	1.1×10^{-6}	0.02	0.002	0.08	0.009	1.6×10^{-6}
	N	69,033	24,243	46,656	99,900	96,598	95,454	322,022	212,127	230,394	211,022
Female	β	1.14 (1.08-1.20)	0.031	0.011	-0.042	0.036	0.018	0.010	0.021	-0.008	-0.033
	P	1.4×10^{-6}	2.4×10^{-7}	0.07	3.5×10^{-11}	1.0×10^{-8}	0.002	0.01	3.6×10^{-6}	0.08	9.9×10^{-14}
	N	40,413	13,073	23,663	62,816	59,473	61,803	171,899	117,967	126,971	117,288
Male	β	1.08 (1.03-1.14)	0.006	0.004	-0.034	0.013	0.010	0.009	0.011	-0.009	0.003
	P	0.003	0.36	0.58	9.9×10^{-7}	0.04	0.19	0.03	0.02	0.04	0.50
	N	28,620	11,158	17,731	37,745	35,288	36,840	152,830	94,344	103,616	93,919