

RESEARCH ARTICLE

Regional fat depot masses are influenced by protein-coding gene variants

Matt J. Neville^{1,2}, Laura B. L. Wittemans^{3,4}, Katherine E. Pinnick¹, Marijana Todorčević¹, Risto Kaksonen⁵, Kirsi H. Pietiläinen^{5,6}, Jian'an Luan³, Robert A. Scott³, Nicholas J. Wareham³, Claudia Langenberg³, Fredrik Karpe^{1,2*}

1 Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom, **2** Oxford NIHR Biomedical Research Centre, Oxford University Hospitals Trust, Oxford, United Kingdom, **3** MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Cambridge, United Kingdom, **4** Wellcome Centre for Human Genetics, University of Oxford, Oxford, United Kingdom, **5** Obesity Research Unit, Research Programs Unit, Diabetes and Obesity, University of Helsinki, Helsinki, Finland, **6** Abdominal Center, Endocrinology, Helsinki University Hospital, Helsinki, Finland

☞ These authors contributed equally to this work.

* Fredrik.Karpe@ocdem.ox.ac.uk



OPEN ACCESS

Citation: Neville MJ, Wittemans LBL, Pinnick KE, Todorčević M, Kaksonen R, Pietiläinen KH, et al. (2019) Regional fat depot masses are influenced by protein-coding gene variants. PLoS ONE 14(5): e0217644. <https://doi.org/10.1371/journal.pone.0217644>

Editor: Yvonne Böttcher, University of Oslo, NORWAY

Received: January 31, 2019

Accepted: May 15, 2019

Published: May 30, 2019

Copyright: © 2019 Neville et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: FK and MN were funded by British Heart Foundation (BHF) (RG/17/1/32663). The Oxford Biobank volunteer recruitment and recall process was funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). LW, RS, NW JL and CI and work performed at the MRC Epidemiology Unit, Cambridge was funded by the United Kingdom's

Abstract

Waist-to-hip ratio (WHR) is a prominent cardiometabolic risk factor that increases cardio-metabolic disease risk independently of BMI and for which multiple genetic loci have been identified. However, WHR is a relatively crude proxy for fat distribution and it does not capture all variation in fat distribution. We here present a study of the role of coding genetic variants on fat mass in 6 distinct regions of the body, based on dual-energy X-ray absorptiometry imaging on more than 17k participants. We find that the missense variant *CCDC92*_{S70C}, previously associated with WHR, is associated specifically increased leg fat mass and reduced visceral but not subcutaneous central fat. The minor allele-carrying transcript of *CCDC92* is constitutively more highly expressed in adipose tissue samples. In addition, we identify two coding variants in *SPATA20* and *UQCC1* that are associated with arm fat mass. *SPATA20*_{K422R} is a low-frequency variant with a large effect on arm fat only, and *UQCC1*_{R51Q} is a common variant reaching significance for arm but showing similar trends in other subcutaneous fat depots. Our findings support the notion that different fat compartments are regulated by distinct genetic factors.

Introduction

While generalised adiposity, as measured by body mass index (BMI), is a well-established and major risk factor for cardio-metabolic diseases, body fat distribution is increasingly being recognised as an even stronger determinant of metabolic risk[1, 2]. For example, Yusuf *et al.*[1] showed that waist-to-hip ratio (WHR) is a stronger predictor of myocardial infarction than BMI. To date, the overwhelming majority of genome and exome-wide association studies on fat distribution have focussed on waist and hip circumference and WHR[3, 4]. While these measures are easy and cheap to obtain on a large scale, they do not capture all variation in fat

Medical Research Council through grants MC_UU_12015/1, MC_PC_13046, MC_PC_13048 and MR/L00002/1. Work by RK and KP on the Finnish adipose tissue Cohort was funded by the Academy of Finland, Centre of Excellence in Research on Mitochondria, Metabolism and Disease (FinMIT), grant number: 272376; Academy of Finland, grant numbers: 314383, 266286; Finnish Medical Foundation (Suomen Lääketieteen Säätiö), Gyllenberg Foundation, Finnish Foundation for Cardiovascular Research (Sydäntutkimussäätiö); Novo Nordisk Foundation; University of Helsinki; Helsinki University Hospital; Finnish Diabetes Research Foundation; Diabetesliitto; and the Wellcome Trust (grant number WT083442AIA). EPIC-Norfolk has received funding from the Medical Research Council (MR/N003284/1) and Cancer Research United Kingdom (14136). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

distribution. For example, WHR does not capture peripheral fat stored in the upper limbs and the distribution of overall central fat over the subcutaneous and visceral compartments, of which the latter have been suggested to have discordant effects on cardio-metabolic risk[5–8]. Furthermore, circumference-based estimates of fat accumulation do not take into account differences in lean mass and bone structure and mass. Therefore, genetic association studies based on direct measures of regional fat mass may help unpick mechanisms underlying the expansion of distinct fat depots.

Quantification of fat mass in distinct regions of the body requires imaging methods with post-image processing to derive delineation of tissues, such as magnetic resonance imaging or dual x-ray absorptiometry (DXA). DEXA-based measures of regional body fat have been shown to correlate more strongly with metabolic risk factors than traditional anthropometric risk factors, such as BMI and waist and hip circumference[9]. We therefore investigated associations of common and rare non-synonymous variants covered on the exome chip with 6 distinct fat compartments, derived from DXA scans on 17,212 participants from 3 cohorts. We hypothesise that by identifying fat depot-specific genetic loci we may gain better insight into the site-specific role of adipose tissue on disease aetiology.

Results and discussion

We tested the associations of all non-synonymous genetic variants covered on the Illumina Human Exome Bead chip with DXA-derived fat mass in the 6 body fat regions: arm fat, leg fat, gynoid fat, total android fat, visceral abdominal fat and subcutaneous abdominal fat (S1 Table). The regional fat phenotypes were adjusted for the first 4 genetic principal components, age and % body fat and the residuals were rank-based inverse normally transformed for men and women separately. The exome chip meta-analyses included up to 17,212 participants of European ancestry from the Oxford Biobank[10], Fenland[11] and EPIC-Norfolk[12] cohorts (Table 1 and S1 Table). Three non-synonymous variants reached exome-wide significance ($p < 2 \times 10^{-7}$) (Fig 1, Table 2 and S2 Table): rs11057401, a common missense variant in Coiled-Coil Domain Containing 92 (*CCDC92*_{S70C}); rs62621401, a novel low-frequency missense variant in Spermatogenesis Associated 20 (*SPATA20*_{K422R}) and rs4911494, a common missense variant in Ubiquinol-Cytochrome C Reductase Complex Assembly Factor 1 (*UQCCL1*_{R51Q}). Heterogeneity tests based on Cochran's Q statistics indicate no significant differences among the effect estimates in the 6 contributing datasets (p for Cochran's Q statistic for rs11057401 = 0.19, for rs4911494 = 0.61 and rs62621401 = 0.42). The three loci identified here show the same direction of effect in both men and women. However, the effect size of rs11057401 (*CCDC92*) on visceral fat was two-fold stronger in men than in women whereas the association with lower body fat mass showed no difference between the sexes. An additional 30 non-synonymous variants reached suggestive significance across 38 tests ($p < 10^{-6}$, S3 Table), including

Table 1. Study cohorts.

Study name	Sample size (% men)	Genotyping array	QC Passed Variants. Total N ^a	Polymorphic variants. N
Fenland-ExomeChip	1,145 (45.8%)	Illumina Exome BeadChip v1.0	240,859	95,739
Fenland-CoreExome	997 (45.2%)	Illumina Infinium Core Exome 24 v1 array	234,179	85,218
Fenland-Axiom	7,363 (47.5%)	Affymetrix UK Biobank Axiom array	58,240	57,864
EPIC-Norfolk	3,101 (45.0%)	Affymetrix UK Biobank Axiom array	56,837	52,020
Oxford Biobank Exome Chip	3,281 (43.7%)	Illumina Exome BeadChip v1.0	245,138	125,912
Oxford Biobank Axiom	1,325 (41.4%)	Affymetrix UK Biobank Axiom array	62,732	56,820

^a Counts represent the number of variants in each dataset that overlap with the Illumina Exome Beadchip v1.0 content after standard QC metrics are applied.

<https://doi.org/10.1371/journal.pone.0217644.t001>

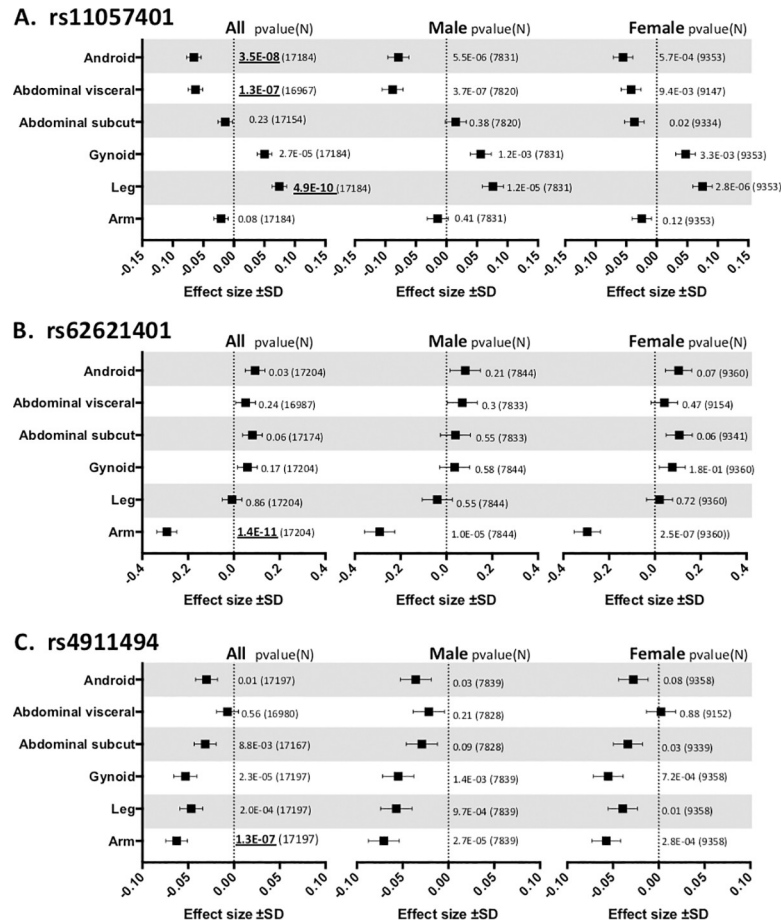


Fig 1. The effect size and direction of effect of meta-analysis findings. Effect size and direction of effect of the three exome-wide significant missense variants: **A.** rs11057401 in *CCDC92* (EAF = 0.32), **B.** rs62621401 in *SPATA20* (EAF = 0.016) and **C.** rs4911494 in *UQCC1* (EAF = 0.62). Data is presented for the 6 DXA measures under investigation and is presented as the beta value \pm SD. The meta-analysis significance level using an additive model for gender combined (All) as well as for gender stratified analysis, together with the N indicated to the right of the data in parentheses. DXA measures are Arm fat mass (Arm), Total android fat mass (Android), Subcutaneous android fat mass (Abdominal subcut), Visceral android fat mass (Abdominal visceral), Gluteal fat mass (Gynoid) and Leg fat mass (Leg). Exome-wide significant data ($p < 2E^{-7}$) are in bold and underlined.

<https://doi.org/10.1371/journal.pone.0217644.g001>

a large haplotype block on chromosome 17 containing 8 missense variants across the *SPPL2C*, *MAPT*, *KANSL1* genes and *GDF5*_{S276A} in LD with *UQCC1*_{R51Q}.

The rs11057401 *CCDC92*_{S70C} variant (EAF = 0.32) is predicted to cause a deleterious amino acid change as assessed by predictSNP[13]. The minor allele of rs11057401 shows significant opposing effects on android, specifically visceral fat mass, and lower body fat, in the sex-combined additive model (total abdominal fat mass: $\beta = -0.065$, SE = 0.012, $p = 3.5 \times 10^{-8}$; visceral abdominal fat mass: $\beta = -0.063$, SE = 0.012, $p = 1.3 \times 10^{-7}$; leg fat mass: $\beta = 0.075$, SE = 0.020, $p = 4.9 \times 10^{-10}$) (Table 2). These data extend WHR associations reported by Justice, *et al.*[3] and Lotta *et al.*[11] at this locus. Lotta *et al.*[11] describe the contribution of leg fat mass; we here demonstrate an additional opposing effect specifically on abdominal visceral fat mass but not for abdominal subcutaneous fat mass, which would correspond to the already observed association with increased waist circumference but with the present analysis showing that the effect is confined to the intra-abdominal fat depot only. Found on chromosome 12q24, *CCDC92* is ubiquitously expressed with highest levels in adipose tissue, brain and testes. It is a nuclear

Table 2. Primary exome-wide significant findings.

rsID	Chr:Position (GRCH37)	Gene	Amino acid Change	Ref Allele	Alt Allele	DXA derived Fat Depot	N	Alternate Allele Frequency	Effect Size	standard error	Pvalue ^a	Haplotype region (GRCH37) ^b	Number of SNPs in LD with index SNP ^c
rs11057401	12:124427306	CCDC92	S70C	T	A	Total Android Fat	17184	0.321	-0.065	0.012	3.5E-08	chr12:124403769–124495203	99
						Visceral Fat	16967	0.321	-0.063	0.012	1.3E-07		
						Leg Fat	17184	0.321	0.075	0.012	4.9E-10		
rs62621401	17:48628160	SPATA20	K422R	A	G	Arm Fat	17204	0.016	-0.293	0.043	1.5E-11	chr17:48623825–48633043	7
rs4911494	20:33971914	UQCC1	R51Q	C	T	Arm Fat	17197	0.616	-0.063	0.012	1.3E-07	chr20:33887955–34194423	129

^a Exome-wide significance was set to 2×10^{-7}

^b The haplotype region is defined as the furthest 3' and 5' SNPs with a $R^2 > 0.9$ with the index SNP

^c SNP count is based on 1000 genome SNP data with SNPs in high LD ($R^2 > 0.9$) with the index SNP

<https://doi.org/10.1371/journal.pone.0217644.t002>

protein interacting with the centriole-ciliary interface[11] and may also be involved in DNA repair[14]. The lead variant tags a large haplotype of at least 99 SNPs ($r^2 > 0.9$) across a number of genes including the putative transcription factor Zinc Finger Protein 664 (*ZNF664*) and Dynein Axonemal Heavy chain 10 (*DNAH10*) (Table 2). There is also strong evidence for multiple eQTL signals across this haplotype which includes three genes, i.e. *CCDC92*, *DNAH10* and *ZNF664*[15]. Previous GWAS studies have also associated SNPs in this haplotype with a reduction in insulin resistance[11], improvements in metabolic syndrome[16], reduced WHRadjBMI[3, 4], increased adiponectin levels[17] and with increased plasma HDL-cholesterol and reduced triglyceride concentrations[18–20]. Ablation of *CCDC92* and *DNAH10* in mouse OP9-K cells impairs adipogenesis and reduced lipid accumulation[11]. To further define the likely causative gene or genes in this complex region we undertook a number of gene expression studies in human regional adipocytes and whole adipose tissue. *CCDC92* and *ZNF664* showed very similar expression profiles between abdominal subcutaneous adipose tissue (ASAT), gluteal subcutaneous adipose tissue (GSAT) and arm subcutaneous adipose tissue derived (ARM) cDNA (Figs 2 and 3, S1–S3 Figs), whilst *DNAH10* expression could not be detected in cDNA from either of the diverse human adipose tissues or cultured primary human preadipocytes, making it an unlikely effector transcript. Across a panel of 52 paired ASAT and GSAT cDNA samples (S1 Fig), qPCR showed small differences in expression of *CCDC92* and *ZNF664* between ASAT and GSAT as well as between lean and obese individuals. In a cultured human primary pre-adipocyte differentiation time course experiment, both *CCDC92* and *ZNF664* showed a significant upregulation by day 4 of differentiation (Fig 2A and 2B) but no difference in expression levels was observed between preadipocytes of ASAT and GSAT origin. Whilst this study focusses on exonic coding variants, recent studies have highlighted that coding variants are not always the causal variant at the locus, but may be associated with the trait simply because they are correlated with the causal (non-coding) variant [21]. To that end, we also sought to investigate the reported eQTL signals at this locus, for both *CCDC92* and *ZNF664* (GTEx project[15] and[11]) using allele-specific qPCR; a method that allows us to assess expressed allelic imbalance in heterozygous individuals and thus an eQTL. This showed a highly statistically significant increased expression of transcripts found on the minor allele haplotype for both genes (ASAT 5.8%, GSAT 4.9%, Fig 3A and 3B). Of functional

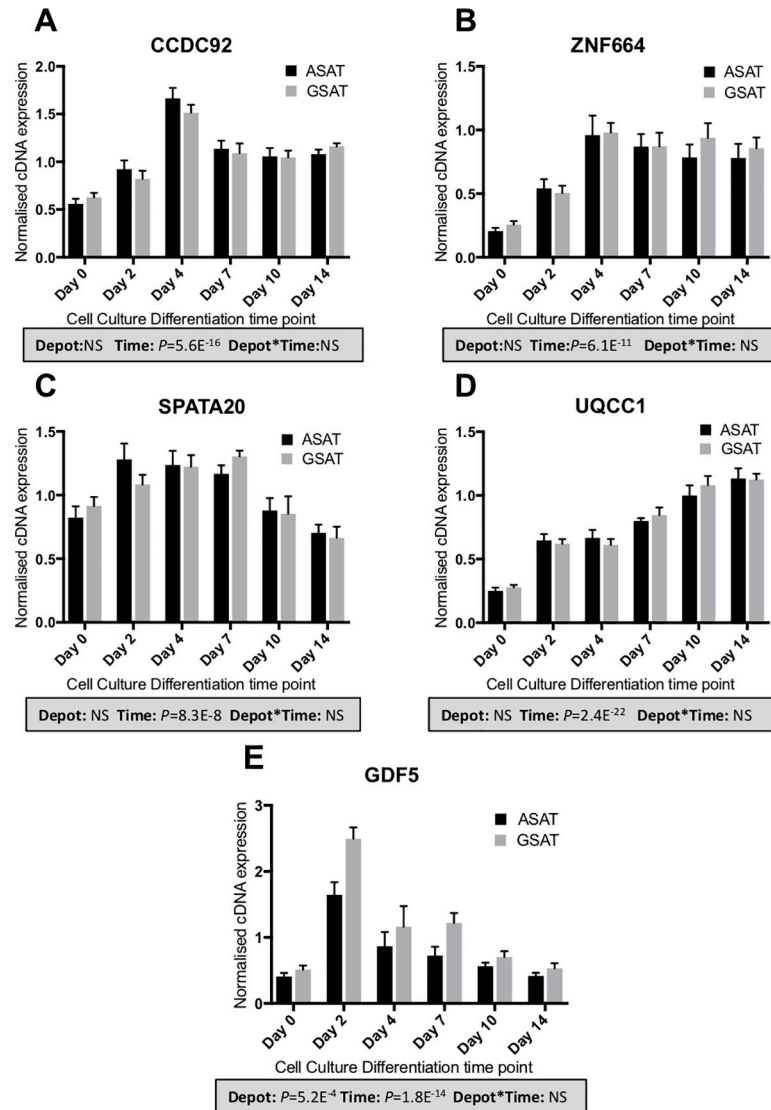


Fig 2. Expression of candidate genes across a human primary adipocyte differentiation time course. cDNA expression of *CCDC92* (A), *ZNF664* (B), *SPATA20* (C), *UQCC1* (D) and *GDF5* (E) was measured over a 14-day adipogenic differentiation time-course using primary preadipocytes from abdominal subcutaneous adipose tissue (ASAT) and gluteal subcutaneous adipose tissue (GSAT) fat depots [40]. Data are shown as DDCt values (normalized to *PPIA* and *PGK1*; $n = 6$, mean \pm SEM). A multivariate general linear model was used to test for statistical significance between depots and time, and to assess depot x time interactions. p -values are presented in the shaded boxes, NS: non-significant.

<https://doi.org/10.1371/journal.pone.0217644.g002>

importance is that this allelic expression imbalance would result in the increased expression of the predicted deleterious serine-70-cysteine amino acid substitution in the *CCDC92* protein. Interestingly, zinc finger proteins such as *ZNF664* have been suggested to regulate the expression of near-by genes [16]. The observed co-regulatory expression pattern of *ZNF664* and *CCDC92* could then possibly be due to the eQTL acting on *ZNF664* which then upregulates the *CCDC92* transcript carrying the deleterious coding variant. Further work needs to be done to investigate this.

The rs62621401 *SPATA20*_{K422R} is a rare variant (EAF 0.016), which has to our knowledge not been reported for any other phenotypes. This amino acid substitution is predicted to be

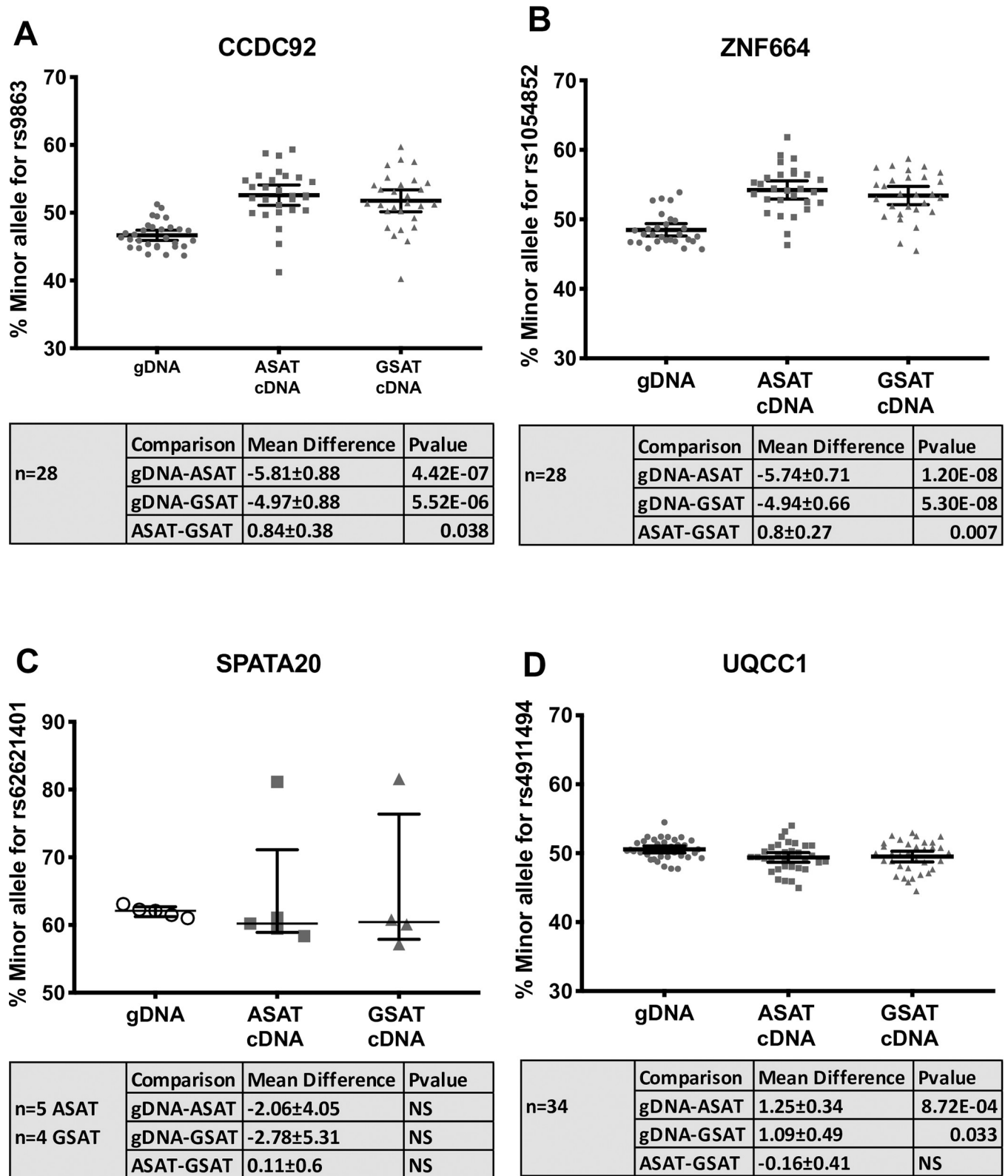


Fig 3. eQTL assessment of exome-wide significant loci by allele-specific qPCR expression. Allelic expression was measured on 4 candidate transcripts in our three exome-wide significant regions using allele-specific qPCR. Data is presented as the % of the minor allele detected compared to the major allele, as described in the methods, with a line indicating the mean and 95% CIs. To assess the rs11057401 eQTL haplotype the proxy SNP rs9863 was assessed for

CCDC92 (A) and the transcribed region proxy SNP rs1054852 for *ZNF664* (B). The index SNP rs62621401 was used to assess the *SPATA20* transcript (C) and the index SNP rs4911494 for *UQCC1*. Paired samples were compared between abdominal subcutaneous adipose tissue (ASAT) and gluteal subcutaneous adipose tissue (GSAT) and genomic DNA (gDNA). For each transcript ABI Taqman genotyping assays were selected that fall within the transcribed sequence. gDNA selected from the same individuals as the cDNAs acts as a paired control with presumed equal allele expression. Deviation from 50% for gDNA, particularly pronounced in *SPATA20* (C), represents inherent imbalance in assay technical performance and position of optimal Ct between Vic and Fam fluorescence. By using paired gDNAs to select cDNAs allelic expression imbalance can be resolved by comparing cDNA to its paired gDNA. Significance was assessed with paired t-test in SPSSv24. Mean differences between comparisons and statistical significance is presented in shaded boxes. NS: Non-significant. The single outlier seen for *SPATA20* (C) was replicated in a second cDNA synthesis and both ASAT and GSAT. No phenotype differences were observed for this individual and no obvious genetic differences were observed.

<https://doi.org/10.1371/journal.pone.0217644.g003>

benign and allele-specific qPCR on paired ASAT and GSAT cDNA samples of five Oxford Biobank participants heterozygous for rs62621401 did not reveal any association with expression of *SPATA20* (Fig 3C). This variant shows a large effect size and is the first locus to be associated with arm fat mass ($\beta = -0.29$, SE = 0.04, $p = 1.5 \times 10^{-11}$, Table 2 and Fig 1B) with an estimated per-allele effect size in the Oxford Biobank (n = 4,606) of 125g less arm fat mass (-5.8%, 95% confidence interval (CI) = [-9.3%, -2.3%]) in men and 67g (-2.6%, 95% CI = [-5.3%, 0.4%]) in women (approximate fat mass (grams) per allele after adjusting for covariates with % change in parentheses, S1 Table). *SPATA20*, linked to spermatogenesis in mice[22], is highly expressed in human testes but also ubiquitously expressed, including in adipose tissue. *SPATA20* is a putative member of the thioredoxin family and other members of this family have been shown to be involved in preadipocyte proliferation[23] and pro-adipogenic *Wnt* signalling[24]. *SPATA20* expression was higher in men than women, although this was only significant for GSAT ($p = 0.01$) (S1 Fig). Expression of *SPATA20* during adipocyte differentiation showed an increase between days 0 to 7 of adipogenesis then a drop back to pre-differentiation levels between days 7 and 14 ($p = 8.3 \times 10^{-8}$, Fig 2C) suggesting a role for this gene in adipocyte development. Surprisingly, despite the arm-specific association, expression of *SPATA20* was similar between arm fat, ASAT and GSAT (Panel A in S2 Fig). However, qPCR assessment of a number of developmental genes (Homeobox genes) in arm adipose tissue compared to ASAT and GSAT showed significant differences (Panel B in S2 Fig), which suggests that arm fat is developmentally distinct from the other fat depots assessed. It is therefore possible that *SPATA20* is involved in an arm fat-specific developmental pathway.

The rs4911494 *UQCC1*_{R51Q} (EAF = 0.62) variant was also inversely associated with arm fat mass (Table 2 and Fig 1C) but was not predicted as damaging by predictSNP. Whilst exome-wide significance is only observed for arm fat, the arm fat-lowering allele is suggestively associated with less fat in all peripheral and subcutaneous fat depots and in both men and women (Fig 1C and S1 Table). *UQCC1* plays a role in mitochondrial respiratory chain complex III protein expression[25] and is structurally similar to the mouse *Bfzb* controlling mouse brown fat [26]. Previous associations at this locus, either the arm fat reducing allele of rs4911494 reported here or SNPs in LD have shown; reduced height [27], reduced body weight[28], increased WHRadjBMI[3] and increased osteoarthritis risk [29]. Another missense variant in the nearby canonical *Wnt* signalling gene *GDF5* (rs224331) is in LD with rs4911494 ($r^2 > 0.9$) and reaches suggestive significance with arm fat mass in women ($\beta = 0.13$, SE = 0.03, $p = 8.97 \times 10^{-7}$, S3 Table). During adipogenesis, expression of *UQCC1* increases ($p = 2.4 \times 10^{-22}$, Fig 2D) but with no difference between ASAT and GSAT in cultured pre-adipocytes. Allele-specific qPCR showed that the minor allele (rs4911494) was associated with a small but statistically significant decrease in expression of *UQCC1* in ASAT (per-allele percentage change in expression = -1.25%, $p = 8.7 \times 10^{-4}$; Fig 3D). It is unclear whether this small change is biologically relevant. *GDF5* expression was not detected in adipose tissue cDNA samples. However, during adipogenesis *GDF5* showed transient expression at day 2, highlighting the possibility that *GDF5* is a regulator of early adipocyte differentiation.

Conclusion

This study represents the largest exome chip meta-analysis on DXA-derived discrete fat depot masses to date. The value of better-defined fat depot regions is illustrated at the *CCDC92* locus. The *CCDC92*_{S70C} variant shows a clear effect on visceral fat mass, but not on subcutaneous abdominal fat mass, while an opposing effect on lower body fat mass was observed. These findings refine our understanding of the role of the *CCDC92* locus on distinct fat compartments compared to what could be derived from genetic analyses on circumference-based measures of fat distribution. Whether the opposing effects on visceral versus leg fat mass are because genetic variation at the *CCDC92* locus plays a direct role in both depots, or because the opposite effect on one depot is simply a consequence of a primary effect on the other depot is unclear and will require further investigation. A previous study investigating computerised tomography (CT) scan-derived visceral and subcutaneous fat mass also found associations at the *CCDC92* locus with *ZNF664* (rs1048497) and *DNAH10* (rs1316952)[30] but these SNPs are both in low LD with the index variant in this study (r^2 0.34 and 0.18, respectively) and may represent independent signals. No other previously reported loci associated with CT-derived visceral fat measures[30] were replicated to suggestive significance. The limitation of using an algorithm-based approximation of visceral fat mass, as presented here with the DXA platform should be acknowledged in comparison with the direct measurements provided by CT or MRI methods. However, DXA-derived estimation of visceral fat mass is very strongly correlated with CT-derived measures ($r^2 = 0.96$)[31] as has associations with T2D and CVD risk profiles comparable with CT and MRI methods[9]. These data and the depth of previous GWAS findings at the *CCDC92-ZNF664* locus highlight this as an important region in regulating adipose tissue distribution. In addition, we report two coding variants associated with arm fat mass: a novel low-frequency variant in *SPATA20* with a large effect size that seems to only affect arm fat mass and a common variant in *UQCC1* that has additional but weaker effects on other subcutaneous fat depots. Overall, when comparing these data to exome chip analyses on standard anthropometric traits, such as WHR[3], we replicate few of the reported loci, likely due to the relatively small sample size of our analyses. However, we have identified a locus, not found for any traditional anthropometric traits, for arm fat mass and refined the tissue-specific association for another locus (*CCDC92*), highlighting the value of more defined regional fat measures.

Regional fat depots have distinct physiological regulation with an impact on the whole body metabolic homeostasis, with distinct transcriptomes demonstrating functional differences and differences in origin[7]. This study further supports genetic evidence for overall, distinct regulation of regional fat depots.

Materials and methods

Population cohorts

Oxford biobank. The Oxford Biobank (OBB) cohort (<http://www.oxfordbiobank.org.uk>) consists of an age-stratified random sample of men and women (aged 30 to 50 years) of European ancestry resident in Oxfordshire, UK, as described previously[10]. All participants gave written, informed consent to participate, and studies were approved by the Oxfordshire Research Ethics Committee (08/H0606/107+5). A total of 3,281 individuals from the Oxford Biobank had both measures of fat mass with GE Lunar iDXA[9] and Illumina Human Exome Beadchip genotypes after QC checks. An additional 1,325 individuals contained DXA data and Affymetrix UK Biobank Axiom array genotype data from which overlapping ExomeChip data was extracted for the purposes of this study (Table 1).

Fenland. The Fenland study is a longitudinal cohort consisting of 12,435 participants born between 1950 and 1975. Participants were recruited from the general population through GP surgeries in Cambridge, Ely and Wisbech (Cambridgeshire, East Anglia)[32] and underwent detailed metabolic phenotyping and genome-wide genotyping. Ethical approval for the study was given by the Cambridge Local Ethics committee (ref. 04/Q0108/19) and all participants gave their written consent prior to entering the study.

A total of 1,145 individuals from the Fenland cohort had both measures of fatmass with GE Lunar iDXA[9] and Illumina Human Exome BeadChip genotypes after QC checks (Fenland-ExomeChip, Table 1), a further 997 had Illumina Infinium Core Exome data and 7,363 had Affymetrix UK Biobank Axiom array genotype data from which overlapping ExomeChip data was extracted for the purpose of this study (Table 1).

EPIC-Norfolk. EPIC-Norfolk is one of the two UK-based centres of the European Prospective Investigation into Cancer and nutrition (EPIC) study—an ongoing multi-centre European cohort study set up to study the role of nutrition in cancer risk [33] The EPIC-Norfolk study is a cohort of 25,000 participants, aged 45–74 at baseline, who were recruited through GP surgeries in Norfolk (East Anglia, UK) from the general population between 1993 and 1997[12]. The study was approved by the Norfolk Research Ethics Committee (ref. 05/Q0101/191) and all participants gave their written consent before entering the study. A total of 3,101 individuals had Affymetrix UK Biobank Axiom array genotype data from which overlapping ExomeChip data was extracted for the purposes of this study (Table 1).

DXA-derived depot-specific fat mass measures

For all cohorts depot-specific fat mass was quantified using GE Lunar iDXA (GE Healthcare, Bucks, UK). As previously described[9] these give high precision estimates of body composition. The standard setting of the Encore software (version 14.0; GE Healthcare, Bucks, UK) was used to automatically define regions of interest ensuring that boundaries were consistent between cohorts. The descriptives for the DXA measures used are presented in S1 Table. Visceral fat mass and android subcutaneous fat mass were not measured directly. Visceral fat mass was calculated using an algorithm within the Encore software as described elsewhere[9, 31] and the android subcutaneous fat mass was calculated by subtracting the visceral fatmass from total android fat mass. The DXA scanning was calibrated as per manufacturer's instructions. For arm and leg fat masses the combined masses of the left and right limbs were used, as calculated in the software.

Exome-wide genotype analysis

Datasets. Six data sets from three cohorts, Oxford Biobank[10], Fenland[34] and EPIC--Norfolk[12] (Table 1), equalling a total of 17,212 individuals of European ancestry were compiled for this analysis. The Illumina Exome BeadChip v1.0 genotype content was used as the base content. Where other genotype arrays were used (see Table 1) only the content overlapping with the Illumina ExomeChip were selected. The breakdown of descriptives for each of the 6 datasets can be found in S1 Table. Standardized, between cohorts, Exome-Chip protocols and quality control (QC) metrics were employed for genotype calling for each cohort separately using zCall [35] and individuals and loci that failed QC removed before association analysis. Allele calls, chromosome positions and strand were harmonised between cohorts with a common reference files using PLINK and converted to .vcf files in plinkseq. A standard protocol was also shared between cohorts for harmonised phenotype transformations.

single-variant analysis. Principal component analysis was performed within each study on LD-pruned genotype data with $MAF > 5\%$ obtained using GWAS arrays. Models were

adjusted for the first 4 genetic principal components (PCs) in order to adjust for population stratification and relatedness, which is commonly done in exome and genome-wide association studies[36]. All DXA-derived phenotypes were log-transformed, adjusted for the first 4 PCs, age and percentage total fat mass (calculated as the percentage of total fat mass (grams) to total mass (grams)) and the residuals inverse normal transformed in the R statistical environment to remove systematic differences between studies/DXA instruments/operators. Percentage total fat mass adjusted for age and PC1-4 was also included in the analysis to assess collider bias. Individual datasets were analysed separately in sex-combined and sex-specific analyses using RAREMETALWORKER[37] (<http://genome.sph.umich.edu/wiki/RAREMETALWORKER>). To account for cryptic relatedness, kinship matrices were first calculated and added into the analysis. Single-variant analysis was performed with, additive, recessive and dominant models.

Meta-analysis. Meta-analysis was carried out centrally using RAREMETAL[38]. Variants were excluded if they had a call rate <90%, Hardy-Weinberg equilibrium $p < 10^{-7}$ and markers on Y chromosome or mitochondrial genome. Exome-wide significance for the single-variant analysis was set, based on the full ExomeChip content, as $p < 2 \times 10^{-7}$. A suggestive significance was set to $p < 10^{-6}$.

For this analysis we focussed on non-synonymous variants only, therefore all non-coding variants and synonymous variants were filtered out post meta-analysis together with loci that had a total heterozygous count of less than 10. The exome-wide significant findings are presented in Fig 1 and S1 Table; the additional suggestive significant findings are presented in S3 Table.

Additional informatics

For the three exome-wide significant loci the amino-acid substitutions was assessed for functional significance using the predictSNP online consensus tool[13] (<https://loschmidt.chemi.muni.cz/predictsnp1/>). This allows for assessment across a number of different tools to generate a consensus assessment. For *CCDC92* the S70C missense variant was assessed; for *UQCC1* the R51Q was assessed and for *SPATA20* three different proteins as products of different splice variants were assessed (K422R, K406R and K362R).

Adipose tissue gene expression panels

Six genes found within the three index SNP LD boundaries (Table 2) were assessed for expression levels across a collection of human adipose tissue gene expression panels. Applied Biosystems Taqman assay-on-demand qPCR assays were selected for each gene that also avoid the index SNPs presented here, for *CCDC92* (ABI assay, hs01556139), *ZNF664* (ABI assay, hs00921074), *DNAH10* (ABI assay, hs1387352), *SPATA20* (ABI assay, hs00256188), *UQCC1* (ABI assay hs00921074) and *GDF5* (ABI assay, hs00167060).

For tissue panels, subcutaneous adipose tissue biopsies were collected by needle biopsy as previously described[39]. For cell-cultured human primary preadipocytes, of both abdominal subcutaneous adipose tissue (ASAT) and gluteal subcutaneous adipose tissue (GSAT) origin, a differentiation time course ($n = 6$) was performed as described in Todorčević, *et al.*[40]. All biopsies and cells were homogenized in Tri-reagent (cat. no. T9424, Sigma-Aldrich, UK) and RNA was extracted with a standard Tri-reagent protocol. A total of 500ng RNA was used for cDNA synthesis following standard protocols and random hexamer primers using the cDNA Reverse Transcription Kit (Life Technologies, UK). Real-time PCR reactions were performed on a 1/40 cDNA dilution using Taqman Assays-on-Demand (Applied Biosystems) and Kapa Probe Fast Mastermix (Kapa Biosystems) in triplicate in a 6 μ l final volume and run on an

Applied Biosystems 7900HT machine. Expression was assessed within each panel using a relative qPCR approach[41] and normalised using the previously assessed stably expressed endogenous control genes[39]. For the Lean/Obese Oxford Biobank panel (S1 Fig) the geometric mean of *PPIA*, *PGK1*, *PSMB6* and *IPO8* were used. *IPO8* was not used in a paired arm, ASAT and GSAT panel (S2 Fig) as it was not stably expressed between arm and the other depots. *PPIA* and *PGK1* were used as endogenous controls for primary cell culture experiments.

Neither *DNAH10* nor *GDF5* could be detected above background in whole tissue cDNA panels. *GDF5* was however detected in a 14-day in vitro adipocyte differentiation time course.

Data for a panel of 52 paired ASAT and GSAT biopsy samples was used to assess expression between sexes, between ASAT and GSAT fat depots, and between lean and obese individuals. Descriptives for this panel are presented in S1 Fig. As both *SPATA20* rs62621401 and *UQCC1* rs4911494 were associated with arm fat mass their expression, along with *CCDC92* and *ZNF664* was assessed in a paired arm, ASAT and GSAT cDNA panel. As there is no published data on arm subcutaneous adipose tissue transcriptomics the additional *HOX* gene transcripts *HOXA5*, *HOXB8*, *HOXC8*, *HOXC9* and *HOXC11* were assessed as these are known to be differentially expressed between ASAT, GSAT and visceral fat (Data comparing arm subcutaneous adipose tissue(ARM), ASAT and GSAT gene expression are presented in S2 Fig).

The setup of a human primary adipocyte differentiation time course is described elsewhere [40]. Relative qPCR was run as above on the adipocyte panel for *CCDC92*, *ZNF664*, *SPATA20*, *UQCC1* and *GDF5*. Data is presented in Fig 2.

Allele-specific qPCR

Both the *CCDC92* and the *UQCC1* loci are associated with multiple eQTL signals. Whilst we only consider non-synonymous variants in this analysis this does not discount that the coding locus is also under the influence of an eQTL. To assess the available data from resources such as the GTEx portal and to assess any eQTL effect between ASAT and GSAT fat depots we used the combined resources available within the Oxford Biobank.

Allele specific qPCR was run essentially as described in Fogarty *et al.*[42].

Taqman genotyping assays (Applied Biosystems) were selected to fall within the transcripts under investigation. For *CCDC92* the index SNP assay performed poorly so the Proxy SNP rs9863 (ABI assay, C_206415_30) was selected. To assess the nearby gene *ZNF664* a SNP in high LD with the *CCDC92* index SNP that fell within the *ZNF664* transcript, rs1054852 (ABI assay, C_1169873_10), was selected. For *SPATA20* the index SNP was used (rs62621401, ABI assay C_25983779_10) as was for *UQCC1* (rs4911494, ABI assay, C_25472999_10). As was previously stated neither *DNAH10*, nor *GDF5* could be detected in whole adipose tissue panels. Therefore, allele specific qPCR could not be assessed for these two genes.

From a panel of 200 paired ASAT and GSAT cDNA samples available from the Oxford Biobank, heterozygous individuals were selected. For *CCDC92* and *ZNF664* 28 paired ASAT and GSAT samples were selected, for *UQCC1* there were 34 and for *SPATA20* there were 5. Genomic DNA (gDNA) for these individuals were also retrieved and diluted to 1.5ng/μl. The gDNA is used as the control comparison to the cDNA samples as there is an equal quantity of both alleles in heterozygous gDNA samples. By comparing the ratio of the Ct values from each allele (the ratio of the genotype assay Vic or Fam fluorophore signals) between cDNA and gDNA any allelic expression differences observed in the cDNA samples can be resolved. This is particularly relevant as technical variation exists with each genotyping assay; particularly pronounced in *SPATA20* (Fig 3C).

Data are presented as the percentage of the minor allele Ct value compared to the major allele Ct. This is calculated by first generating a standard curve and regression statistic for each

assay. A standard curve is generated from genomic DNA for individuals homozygous for the major allele (BB) and minor allele (bb). Genomic DNAs are diluted to 1.5ng/μl then BB and bb homozygotes are combined to ratios 80:20, 60:40, 50:50, 40:60, 80:20. Following qPCR analysis using the dual-labelled TaqMan Genotyping assays the ratio of the B to b Ct values are calculated (Ct B minus Ct b) then plotted against the percentage of the minor allele in the dilution series. The linear regression statistic from this standard curve is then used to calculate the percentage minor allele expression of the unknown heterozygous individuals. The standard curves are presented in S3 Fig panels A-D and allele-specific qPCR data for heterozygous individuals are presented in Fig 3.

For *CCDC92*, *ZNF664* and *UQCC1* there were sufficient cDNAs in the 200 panel, however for *SPATA20* there were only 5 individuals. Therefore, to improve the accuracy of the *SPATA20* analysis, each sample was run in triplicate 4x across the assay plate and the average of all 4 sets of triplicates calculated. A single outlier in the *SPATA20* data was followed up in a second cDNA synthesis and persisted in both ASAT and GSAT samples. No phenotype differences were observed for this individual and no obvious genetic differences were found.

Statistical analysis

Statistical significance was assessed for each experiment in SPSS v24. For estimates of per-allele grams fat mass change, log phenotype data was analysed in a general linear model and adjusted for age, PC1-4 and total %fat mass then estimated marginal means were calculated (S1 Table).

Supporting information

S1 Fig. mRNA expression of candidate genes across a lean/obese adipose tissue gene expression panel. mRNA expression of the genes *CCDC92*, *ZNF664*, *UQCC1* and *SPATA20* across a panel of paired abdominal subcutaneous adipose tissue (ASAT) and gluteal subcutaneous adipose tissue (GSAT) cDNA samples from the Oxford Biobank. The panel consisted of 25 male and 29 female healthy individuals selected for either high or low BMI (Lean male, n = 13, age 44.5±0.9 yrs, BMI 22.7±0.3 kg/m², fasting blood glucose 5.2±0.1 mmol/l; Obese males, n = 12, age 43.4±1.2 yrs, BMI 34.9±5.2 kg/m², fasting blood glucose 5.6±0.1 mmol/l; Lean females, n = 15, age 44±1.0 yrs, BMI 21.2±0.2 kg/m², fasting blood glucose 4.8±0.1 mmol/l; Obese Females, n = 14, age 44±1.0 yrs, BMI 33.6±0.6 kg/m², fasting blood glucose 5.2±0.1 mmol/l—data expressed as mean ±SEM). Data are shown as the mean ± SEM DDcT values (normalized to the geometric mean of the endogenous control genes *PPIA*, *PGK1*, *IPO8* and *PSMB6*) as described previously [39, 41]. A multivariate general linear model was used to test for statistical significance between gender, fat depots and obesity and to assess interactions. P-values are presented in the shaded box, NS: non-significant.

There were small but significant differences in expression of *CCDC92*, *ZNF664* and *UQCC1* between fat depots in lean individuals but this difference was lost and expression was significantly reduced, in obese individuals. This is in keeping with a general quiescent state observed in transcripts associated with adipocyte metabolic activity in obesity.

(TIF)

S2 Fig. mRNA expression of candidate genes and homeobox genes across a panel of 22 paired arm, abdominal subcutaneous adipose tissue (ASAT) and gluteal subcutaneous adipose tissue (GSAT). mRNA expression of the candidate genes **A:** *CCDC92*, *ZNF664*, *UQCC1* and *SPATA20* and a selection of developmental *HOX* genes **B:** *HOXA5*, *HOXB8*, *HOXC8*, *HOXC9* and *HOXC11* were determined by real-time qPCR. Data are shown as the mean ±SEM DDcT values (normalized to *PPIA*, *PGK1* and *PSMB6*; n = 22). A univariate general

linear model was used to test for statistical significance between depots. P-values for the HOX genes in **B** are presented in the shaded box.

(TIF)

S3 Fig. Allele specific qPCR standard curves and *CCDC92-ZNF664* regression analysis. The standard curve and regression statistic used to calculate the percentage minor allele expression with allele-specific qPCR is shown above for *CCDC92* (A), *ZNF664* (B), *SPATA20* (C) and *UQCC1* (D). To quantify any allelic expression imbalance for the four genes a standard curve was generated from genomic DNA for individuals homozygous for the Major allele (BB) and Minor allele (bb). Genomic DNAs are diluted to 1.5ng/μl then BB and bb homozygotes were combined to ratios 80:20, 60:40, 50:50, 40:60, 80:20 to generate a standard curve. Following qPCR analysis using dual labelled TaqMan Genotyping assays the ratio of the B to b allele Ct values are calculated (Ct B minus Ct b) then plotted against the percentage of the minor allele in the dilution series. The linear regression statistic from this (A, B, C and D above) is then used to calculate the percentage minor allele expression of our unknown individuals. For *CCDC92* (A), *ZNF664* (B) and *UQCC1* (D) three different pairs of homozygote individuals were used to generate each standard curve and a Mean ± SEM plotted for each dilution (A, B and D). For *SPATA20* only one genomic DNA homozygote minor allele individual was available so an error bar cannot be displayed.

As discussed in the main text there was an observed co-regulatory pattern of expression between *CCDC92* and *ZNF664* across different cDNA panels. To assess any correlation between these two genes within the samples, the allele-specific qPCR paired data points were plotted and regression statistic calculated (Graphs E and F). For both abdominal subcutaneous adipose tissue (ASAT) (E) and gluteal subcutaneous adipose tissue (GSAT) (F) there was a significant correlation, further supporting the co-regulatory pattern of expression.

(TIF)

S1 Table. Population cohort descriptives.

(DOCX)

S2 Table. Exome-wide significant loci. Detailed data on the three exome-wide significant loci described. DXA parameters are included for all measures and meta-analysis statistics for the additive model. DXA measures are arm fatmass (Arm), Total android fat mass (Android), Subcutaneous android fat mass (Subcut), Visceral android fat mass (Visceral), Gluteal fat mass (Gluteal) and Leg fat mass (Leg). Effect size data for suggestive exome-wide significance ($p < 10^{-6}$) is shown in bold. Exome-wide significant data ($p < 2E^{-7}$) are in bold and underlined.

^a The impact of missense variants were assessed using the PREDICTsnP online consensus tool [13] (<https://loschmidt.chemi.muni.cz/predictsnp1/>).

^b Approximate fat mass (grams) changes per allele is shown where test reaches suggestive significance and were calculated as marginal means after adjusting for age, PCs1-4 and %fatmass as covariates in a general linear model, implemented in SPSS v24

(DOCX)

S3 Table. Exome-wide loci showing suggestive level of statistical significance. Additional non-synonymous loci where statistical tests did not reach exome-wide significance but did reach a suggestive significance cut off of $p < 10^{-6}$ are included above.

^a Where it reaches suggestive significance the model is shown as Additive (add), Recessive(rec) or Dominant (Dom).

^b The impact of missense variants were assessed using the predictSNP online consensus tool [13] (<https://loschmidt.chemi.muni.cz/predictsnp1/>).

^c The cluster of 8 Missense SNPs found at the SPPL2C-MAPT-KANSL1 locus on chromosome

17 are part of a single haplotype that extends across ~400kb in this region containing >2300 SNPs ($r^2 > 0.9$), rather than independent signals. (DOCX)

Acknowledgments

We thank the volunteers from the Oxford Biobank (www.oxfordbiobank.org.uk) for their participation. The volunteer recruitment and recall process was funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health recalling process of the volunteers.

The Fenland Study is supported by the MRC (MC_UU_12015/1). Genetic analyses were funded by the MRC (MC_PC_13046). We are grateful to all the volunteers for their time and help, and to the General Practitioners and practice staff for assistance with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams.

EPIC-Norfolk has received funding from the Medical Research Council (MR/N003284/1) and Cancer Research United Kingdom (14136). The genetics analyses in EPIC-Norfolk were funded by the Medical Research Council (MC_PC_13048). We thank all participants and staff for their contribution to the study.

We particularly thank Dr Emanuella De Lucia Rolfe and Richard Powell for processing and conducting the quality control of the DEXA images for the Fenland and EPIC-Norfolk studies.

L. B. L. W acknowledges funding by the Wellcome Trust (WT083442AIA). N. J. W. and C. L. acknowledge funding from the Medical Research Council Epidemiology Unit (MC_UU_12015/1).

Author Contributions

Conceptualization: Matt J. Neville, Robert A. Scott, Claudia Langenberg, Fredrik Karpe.

Data curation: Matt J. Neville, Laura B. L. Wittemans, Jian'an Luan, Robert A. Scott, Nicholas J. Wareham, Claudia Langenberg.

Formal analysis: Matt J. Neville, Laura B. L. Wittemans, Jian'an Luan, Robert A. Scott, Claudia Langenberg.

Funding acquisition: Robert A. Scott, Nicholas J. Wareham, Claudia Langenberg, Fredrik Karpe.

Investigation: Matt J. Neville, Laura B. L. Wittemans, Katherine E. Pinnick, Marijana Todorčević, Risto Kaksonen, Kirsi H. Pietiläinen, Claudia Langenberg, Fredrik Karpe.

Methodology: Laura B. L. Wittemans, Katherine E. Pinnick, Marijana Todorčević, Risto Kaksonen, Kirsi H. Pietiläinen, Jian'an Luan, Fredrik Karpe.

Project administration: Matt J. Neville, Laura B. L. Wittemans, Claudia Langenberg, Fredrik Karpe.

Resources: Laura B. L. Wittemans, Katherine E. Pinnick, Marijana Todorčević, Risto Kaksonen, Kirsi H. Pietiläinen, Jian'an Luan, Nicholas J. Wareham, Claudia Langenberg, Fredrik Karpe.

Software: Matt J. Neville, Laura B. L. Wittemans, Jian'an Luan, Robert A. Scott.

Supervision: Claudia Langenberg, Fredrik Karpe.

Visualization: Matt J. Neville.

Writing – original draft: Matt J. Neville, Fredrik Karpe.

Writing – review & editing: Matt J. Neville, Laura B. L. Wittemans, Katherine E. Pinnick, Marijana Todorčević, Risto Kaksonen, Kirsi H. Pietiläinen, Jian'an Luan, Robert A. Scott, Nicholas J. Wareham, Claudia Langenberg, Fredrik Karpe.

References

1. Yusuf S, Hawken S, Ounpuu S, Bautista L, Franzosi MG, Commerford P, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case-control study. *Lancet*. 2005; 366(9497):1640–9. [https://doi.org/10.1016/S0140-6736\(05\)67663-5](https://doi.org/10.1016/S0140-6736(05)67663-5) PMID: 16271645
2. Prentice AM, Jebb SA. Beyond body mass index. *Obes Rev*. 2001; 2(3):141–7. <https://doi.org/10.1046/j.1467-789x.2001.00031.x> PMID: 12120099.
3. Justice AE, Karaderi T, Highland H, Mahajan A, Gan W, Karpe F, et al. Protein-coding variants implicate novel genes related to lipid homeostasis contributing to body fat distribution. *Nature Genetics* 2018. *In Press*.
4. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Mägi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature*. 2015; 518(7538):187–96. <https://doi.org/10.1038/nature14132> PMID: 25673412
5. Cantile M, Procino A, D'Armiento M, Cindolo L, Cillo C. HOX gene network is involved in the transcriptional regulation of in vivo human adipogenesis. *J Cell Physiol*. 2003; 194(2):225–36. <https://doi.org/10.1002/jcp.10210> PMID: 12494461
6. Min JL, Nicholson G, Halgrimsdottir I, Almstrup K, Petri A, Barrett A, et al. Coexpression network analysis in abdominal and gluteal adipose tissue reveals regulatory genetic loci for metabolic syndrome and related phenotypes. *PLoS genetics*. 2012; 8(2). <https://doi.org/10.1371/journal.pgen.1002505> PMID: 22383892
7. Karpe F, Pinnick KE. Biology of upper-body and lower-body adipose tissue—link to whole-body phenotypes. *Nat Rev Endocrinol*. 2015; 11(2):90–100. Epub 2014/11/04. <https://doi.org/10.1038/nrendo.2014.185> PMID: 25365922.
8. Hilton C, Karpe F, Pinnick KE. Role of developmental transcription factors in white, brown and beige adipose tissues. *Biochim Biophys Acta*. 2015; 1851(5):686–96. Epub 2015/02/08. <https://doi.org/10.1016/j.bbailip.2015.02.003> PMID: 25668679.
9. Vasan SK, Osmond C, Canoy D, Christodoulides C, Neville MJ, Di Gravio C, et al. Comparison of regional fat measurements by dual-energy X-ray absorptiometry and conventional anthropometry and their association with markers of diabetes and cardiovascular disease risk. *Int J Obes (Lond)*. 2017; 74:315. <https://doi.org/10.1038/ijo.2017.289>
10. Karpe F, Vasan SK, Humphreys SM, Miller J, Cheeseman J, Louise Dennis A, et al. Cohort Profile: The Oxford Biobank. *Int J Epidemiol*. 2017. <https://doi.org/10.1093/ije/dyx132> PMID: 29040543
11. Lotta LA, Gulati P, Day FR, Payne F, Ongen H, van de Bunt M, et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nat Genet*. 2017; 49(1):17–26. <https://doi.org/10.1038/ng.3714> PMID: 27841877
12. Day N, Oakes S, Luben R, Khaw KT, Bingham S, Welch A, et al. EPIC-Norfolk: study design and characteristics of the cohort. *European Prospective Investigation of Cancer*. *Br J Cancer*. 1999; 80 Suppl 1:95–103. PMID: 10466767.
13. Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, et al. PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. *PLoS Comput Biol*. 2014; 10(1): e1003440. <https://doi.org/10.1371/journal.pcbi.1003440> PMID: 24453961
14. Chaki M, Airik R, Ghosh AK, Giles RH, Chen R, Slaats GG, et al. Exome capture reveals ZNF423 and CEP164 mutations, linking renal ciliopathies to DNA damage response signaling. *Cell*. 2012; 150(3):533–48. <https://doi.org/10.1016/j.cell.2012.06.028> PMID: 22863007
15. Ward MC, Gilad Y. Human genomics: Cracking the regulatory code. *Nature*. 2017; 550(7675):190–1. <https://doi.org/10.1038/550190a> PMID: 29022577.
16. Kraja AT, Chasman DI, North KE, Reiner AP, Yanek LR, Kilpeläinen TO, et al. Pleiotropic genes for metabolic syndrome and inflammation. *Mol Genet Metab*. 2014; 112(4):317–38. <https://doi.org/10.1016/j.mgme.2014.04.007> PMID: 24981077
17. Wu Y, Gao H, Li H, Tabara Y, Nakatochi M, Chiu Y-F, et al. A meta-analysis of genome-wide association studies for adiponectin levels in East Asians identifies a novel locus near WDR11-FGFR2. *Hum Mol Genet*. 2014; 23(4):1108–19. <https://doi.org/10.1093/hmg/ddt488> PMID: 24105470

18. Chasman DI, Paré G, Mora S, Hopewell JC, Peloso G, Clarke R, et al. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS genetics*. 2009; 5(11):e1000730. <https://doi.org/10.1371/journal.pgen.1000730> PMID: 19936222
19. Liu DJ, Peloso GM, Yu H, Butterworth AS, Wang X, Mahajan A, et al. Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet*. 2017; 49(12):1758–66. <https://doi.org/10.1038/ng.3977> PMID: 29083408
20. Singaraja RR, Tietjen I, Hovingh GK, Franchini PL, Radomski C, Wong K, et al. Identification of four novel genes contributing to familial elevated plasma HDL cholesterol in humans. *J Lipid Res*. 2014; 55(8):1693–701. <https://doi.org/10.1194/jlr.M048710> PMID: 24891332
21. Mahajan A, Wessel J, Willems SM, Zhao W, Robertson NR, Chu AY, et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. *Nat Genet*. 2018; 50(4):559–71. Epub 2018/04/09. <https://doi.org/10.1038/s41588-018-0084-1> PMID: 29632382; PubMed Central PMCID: PMC5898373.
22. Shi H-J, Wu AZ, Santos M, Feng Z-M, Huang L, Chen Y-M, et al. Cloning and characterization of rat spermatid protein SSP411: a thioredoxin-like protein. *J Androl*. 2004; 25(4):479–93. <https://doi.org/10.1002/j.1939-4640.2004.tb02819.x> PMID: 15223837
23. Song JS, Cho HH, Lee B-J, Bae YC, Jung JS. Role of thioredoxin 1 and thioredoxin 2 on proliferation of human adipose tissue-derived mesenchymal stem cells. *Stem cells and development*. 2011; 20(9):1529–37. <https://doi.org/10.1089/scd.2010.0364> PMID: 21158569
24. Bahn YJ, Lee K-P, Lee S-M, Choi JY, Seo Y-S, Kwon K-S. Nucleoredoxin promotes adipogenic differentiation through regulation of Wnt/ β -catenin signaling. *J Lipid Res*. 2015; 56(2):294–303. <https://doi.org/10.1194/jlr.M054056> PMID: 25548260
25. Tucker EJ, Wanschers BFJ, Szklarczyk R, Mountford HS, Wijeyeratne XW, van den Brand MAM, et al. Mutations in the UQCC1-Interacting Protein, UQCC2, Cause Human Complex III Deficiency Associated with Perturbed Cytochrome b Protein Expression. *PLoS genetics*. 2013; 9(12):e1004034. <https://doi.org/10.1371/journal.pgen.1004034> PMID: 24385928
26. Vetter K, Wurst W. Expression of a novel mouse gene 'mbFZb' in distinct regions of the developing nervous system and the adult brain. *Mechanisms of Development*. 2001; 100(1):123–6. [https://doi.org/10.1016/S0925-4773\(00\)00511-6](https://doi.org/10.1016/S0925-4773(00)00511-6) PMID: 11118897
27. Sanna S, Jackson AU, Nagaraja R, Willer CJ, Chen WM, Bonnycastle LL, et al. Common variants in the GDF5-UQCC region are associated with variation in human height. *Nat Genet*. 2008; 40(2):198–203. Epub 2008/01/13. <https://doi.org/10.1038/ng.74> PMID: 18193045; PubMed Central PMCID: PMC2914680.
28. Randall JC, Winkler TW, Kutalik Z, Berndt SI, Jackson AU, Monda KL, et al. Sex-stratified Genome-wide Association Studies Including 270,000 Individuals Show Sexual Dimorphism in Genetic Loci for Anthropometric Traits. *PLoS genetics*. 2013; 9(6). <https://doi.org/10.1371/journal.pgen.1003500> PMID: 23754948
29. Lettre G. The osteoarthritis and height GDF5 locus yields its secrets. *Nat Genet*. 2017; 49(8):1165–6. <https://doi.org/10.1038/ng.3924> PMID: 28747750
30. Fox CS, Liu Y, White CC, Feitosa M, Smith AV, Heard-Costa N, et al. Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. *PLoS genetics*. 2012; 8(5):e1002695. <https://doi.org/10.1371/journal.pgen.1002695> PMID: 22589738
31. Kaul S, Rothney MP, Peters DM, Wacker WK, Davis CE, Shapiro MD, et al. Dual-energy X-ray absorptiometry for quantification of visceral fat. *Obesity (Silver Spring)*. 2012; 20(6):1313–8. Epub 2012/01/26. <https://doi.org/10.1038/oby.2011.393> PMID: 22282048; PubMed Central PMCID: PMC3361068.
32. Lotta LA, Scott RA, Sharp SJ, Burgess S, Luan J, Tillin T, et al. Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. *PLoS Med*. 2016; 13(11):e1002179. Epub 2016/11/29. <https://doi.org/10.1371/journal.pmed.1002179> PMID: 27898682; PubMed Central PMCID: PMC5127513.
33. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr*. 2002; 5(6B):1113–24. <https://doi.org/10.1079/PHN2002394> PMID: 12639222.
34. Lotta LA, Gulati P, Day FR, Payne F, Ongen H, van de Bunt M, et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nat Genet*. 2017; 49(1):17–26. Epub 2016/11/14. <https://doi.org/10.1038/ng.3714> PMID: 27841877; PubMed Central PMCID: PMC5774584.
35. Goldstein JL, Crenshaw A, Carey J, Grant GB, Maguire J, Fromer M, et al. zCall: a rare variant caller for array-based genotyping: genetics and population analysis. *Bioinformatics*. 2012; 28(19):2543–5. Epub 2012/07/27. <https://doi.org/10.1093/bioinformatics/bts479> PMID: 22843986; PubMed Central PMCID: PMC3463112.

36. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006; 38(8):904–9. Epub 2006/07/23. <https://doi.org/10.1038/ng1847> PMID: 16862161.
37. Feng S, Liu D, Zhan X, Wing MK, Abecasis GR. RAREMETAL: fast and powerful meta-analysis for rare variants. *Bioinformatics.* 2014; 30(19):2828–9. Epub 2014/06/03. <https://doi.org/10.1093/bioinformatics/btu367> PMID: 24894501; PubMed Central PMCID: PMC4173011.
38. Liu DJ, Peloso GM, Zhan X, Holmen OL, Zawistowski M, Feng S, et al. Meta-analysis of gene-level tests for rare variant association. *Nat Genet.* 2014; 46(2):200–4. Epub 2013/12/15. <https://doi.org/10.1038/ng.2852> PMID: 24336170; PubMed Central PMCID: PMC3939031.
39. 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.* 2010; 19(4):888–92. <https://doi.org/10.1038/oby.2010.257> PMID: 20948521
40. Todorčević M, Hilton C, McNeil C, Christodoulides C, Hodson L, Karpe F, et al. A cellular model for the investigation of depot specific human adipocyte biology. *Adipocyte.* 2017; 6(1):40–55. <https://doi.org/10.1080/21623945.2016.1277052> PMID: 28452592
41. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 2001; 29(9):e45. <https://doi.org/10.1093/nar/29.9.e45> PMID: 11328886 PubMed Central PMCID: PMC55695.
42. Fogarty MP, Xiao R, Prokunina-Olsson L, Scott LJ, Mohlke KL. Allelic expression imbalance at high-density lipoprotein cholesterol locus MMAB-MVK. *Hum Mol Genet.* 2010; 19(10):1921–9. <https://doi.org/10.1093/hmg/ddq067> PMID: 20159775