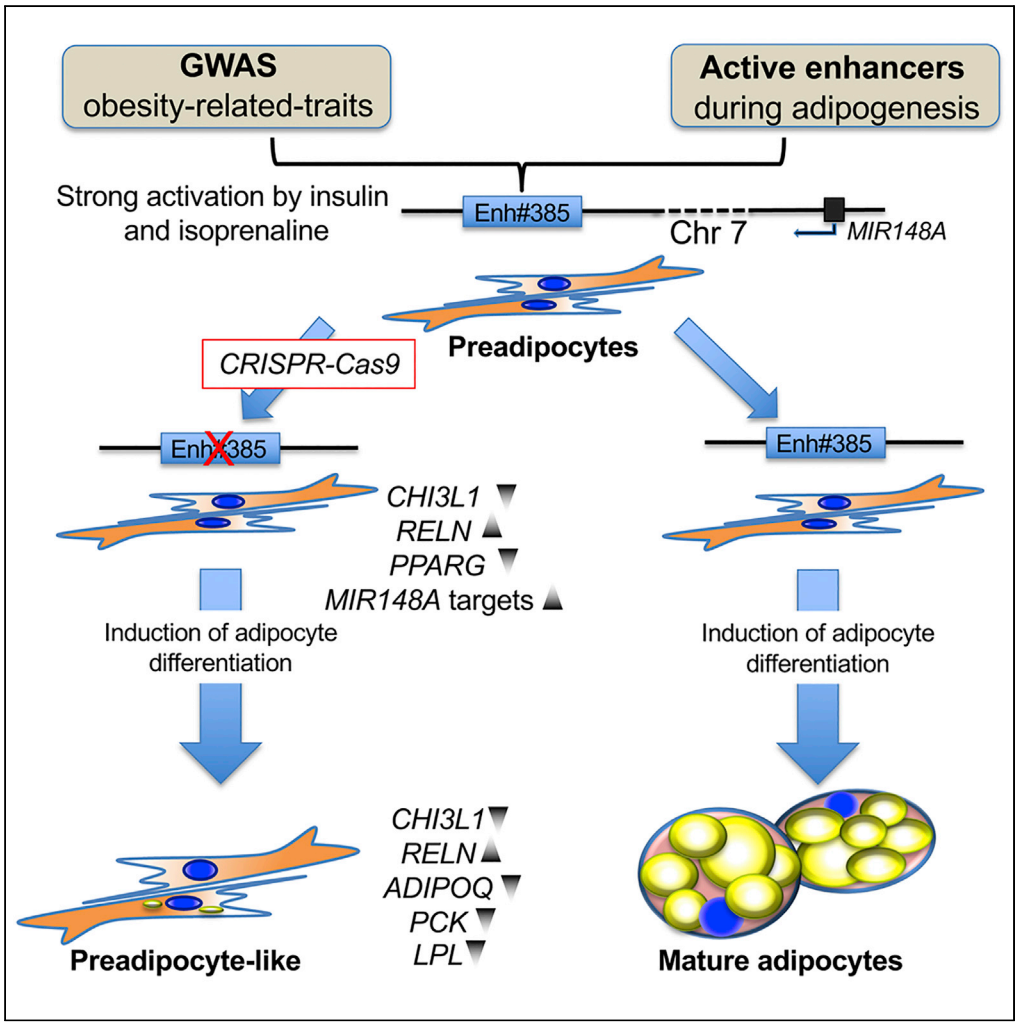
iScience

CellPress

Article

Detailed Functional Characterization of a Waist-Hip Ratio Locus in 7p15.2 Defines an Enhancer Controlling Adipocyte Differentiation



Casimiro Castillejo-Lopez, Milos Pjanic, Anna Chiara Pirona, ..., Thomas Quertermous, Erik Arner, Erik Ingelsson

eriking@stanford.edu

HIGHLIGHTS

An enhancer active during adipogenesis is located in an obesity GWAS locus

The enhancer responded strongly to insulin and isoprenaline

Mutation of the enhancer by CRISPR-Cas9 decreased adipocyte differentiation

Knockout of *CHI3L1* decreased adipocyte differentiation

Castillejo-Lopez et al., iScience *20*, 42–59
October 25, 2019
https://doi.org/10.1016/j.isci.2019.09.006

iScience

CellPress

Article

Detailed Functional Characterization of a Waist-Hip Ratio Locus in 7p15.2 Defines an Enhancer Controlling Adipocyte Differentiation

Casimiro Castillejo-Lopez,[1] Milos Pjanic,[2] Anna Chiara Pirona,[3,4] Susanne Hetty,[3] Martin Wabitsch,[5] Claes Wadelius,[1] Thomas Quertermous,[2,6] Erik Arner,[7] and Erik Ingelsson[2,3,6,8,9,*]

SUMMARY

We combined CAGE sequencing in human adipocytes during differentiation with data from genome-wide association studies to identify an enhancer in the *SNX10* locus on chromosome 7, presumably involved in body fat distribution. Using reporter assays and CRISPR-Cas9 gene editing in human cell lines, we characterized the role of the enhancer in adipogenesis. The enhancer was active during adipogenesis and responded strongly to insulin and isoprenaline. The allele associated with increased waist-hip ratio in human genetic studies was associated with higher enhancer activity. Mutations of the enhancer resulted in less adipocyte differentiation. RNA sequencing of cells with disrupted enhancer showed reduced expression of established adipocyte markers, such as *ADIPOQ* and *LPL*, and identified *CHI3L1* on chromosome 1 as a potential gene involved in adipocyte differentiation. In conclusion, we identified and characterized an enhancer in the *SNX10* locus and outlined its plausible mechanisms of action and downstream targets.

INTRODUCTION

According to estimations by the World Health Organization, more than 1.9 billion adults were overweight and ~13% of the world's adult population was obese in 2016. In addition, 41 million children younger than 5 years were overweight or obese. In parallel with this, dramatic increases of downstream consequences of obesity, such as insulin resistance (IR), type 2 diabetes (T2D), and cardiovascular disease are also expected over the next decade, which will have a devastating impact on global health. Once considered as a problem limited to high-income countries, obesity, IR, and associated conditions are now dramatically increasing also in low- and middle-income countries, especially in urban settings. Heart disease and stroke are the global leading causes of morbidity and mortality, being the underlying cause of death for ~13 million people in 2010, or one in four deaths worldwide (Lozano et al., 2012). Hence, identifying molecular mechanisms underlying obesity and risk of cardiovascular disease and T2D are important public health priorities.

Genome-wide association studies (GWAS) have been tremendously successful in identifying loci associated with complex traits (Visscher et al., 2017). As a result of large meta-analyses of GWAS led by the GIANT consortium, more than 500 loci have been associated with body mass index (BMI), waist-hip ratio (WHR), and other obesity traits (Locke et al., 2015; Lu et al., 2016; Shungin et al., 2015; Turcot et al., 2018; Winkler et al., 2016; Pulit et al., 2019). However, the causal variants, genes, or downstream mechanisms have been established for very few of these loci; hence, the transformative potential of human genomics to unravel the molecular mechanisms underlying obesity, IR, and adipocyte biology remains to the realized.

In the present study, we characterized an active enhancer located in one of the first chromosomal regions to be associated with WHR (Heid et al., 2010), most strongly in women (Shungin et al., 2015; Winkler et al., 2015), and also associated with BMI (Monda et al., 2013), endometriosis (Nyholt et al., 2012; Sapkota et al., 2017), and kidney function (Pattaro et al., 2016) in several GWAS (Figure 1A). The GWAS associations have mostly been annotated as *SNX10* based on expression quantitative trait loci (eQTL) analyses (Graff et al., 2017; Justice et al., 2017; Shungin et al., 2015), and also as *NFE2L3*, *MIR148A*, or *CBX3* based on the gene nearest to the lead GWAS variant from the study in question. The enhancer is localized in an independent haplotype block tagged by the lead GWAS variant rs3902751 that spans 48 kb from rs10245353 to rs1451385 (Figure 1A).

[1]Department of Immunology, Genetics and Pathology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

[2]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

[3]Department of Medical Sciences and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

[4]German Cancer Research Center (DKFZ), Heidelberg, Germany

[5]Department of Pediatrics and Adolescent Medicine, Division of Pediatric Endocrinology and Diabetes, University of Ulm, Ulm, Germany

[6]Stanford Cardiovascular Institute, Stanford University, Stanford, CA 94305, USA

[7]Laboratory for Applied Regulatory Genomics Network Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Kanagawa 230-0045 Japan

[8]Stanford Diabetes Research Center, Stanford University, Stanford, CA 94305, USA

[9]Lead Contact

*Correspondence: eriking@stanford.edu

https://doi.org/10.1016/j.isci.2019.09.006

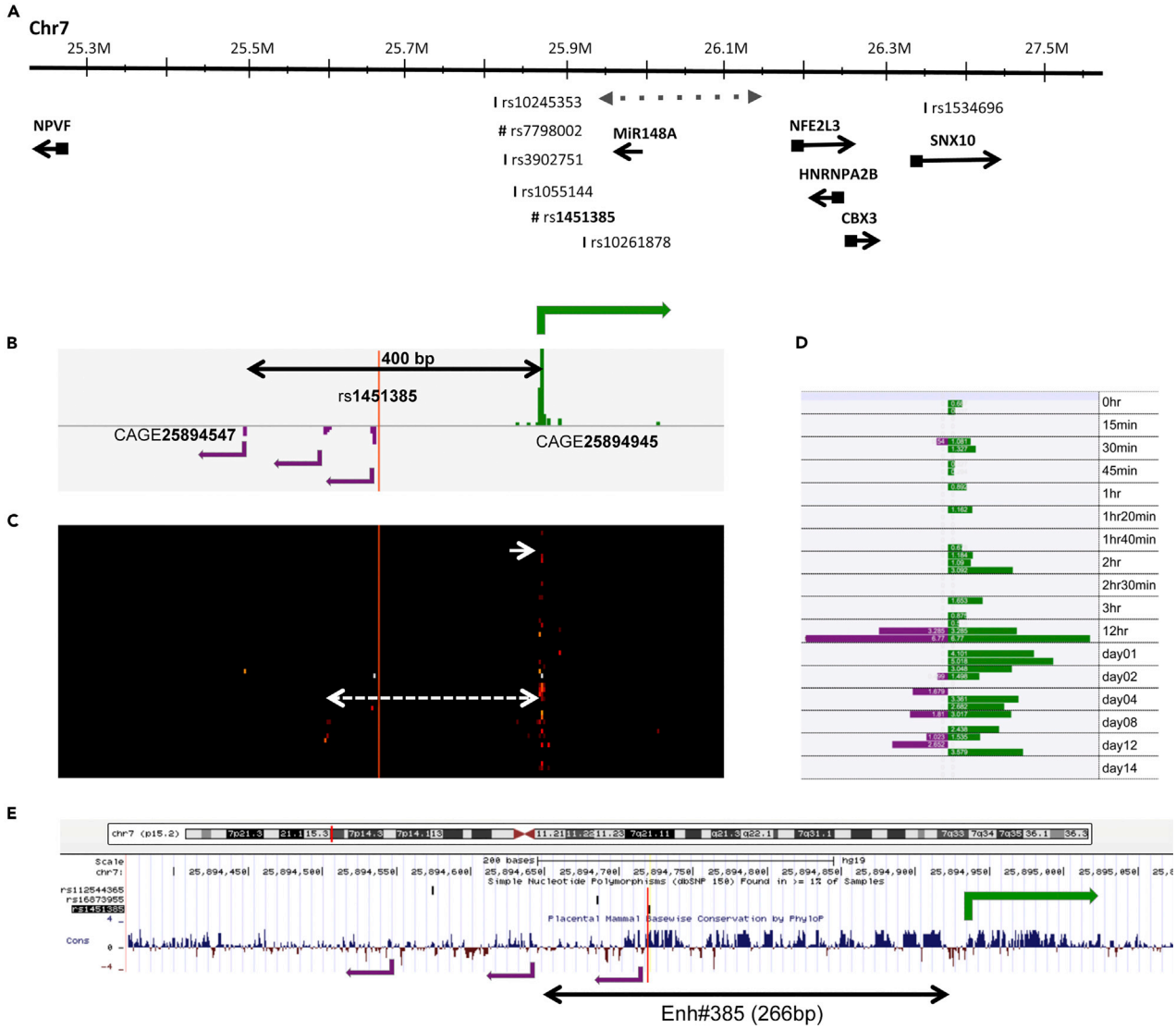


Figure 1. Genetic Features of the 7p15.2 Region Tagged by GWAS Potentially Involved in Adipogenesis Centered around rs1451385

(A) Physical map of the region showing the position of annotated genes, *MIR148A*, and genetic polymorphisms (SNPs, annotated with their rs identifiers) associated with obesity and other related phenotypes. The region is delimited by the distal gene *NPVF* and the proximal gene *SNX10*. Genome-wide significant SNPs include lead variants associated with WHR (rs10245353, rs3902751, rs1534696, Shungin et al., 2015; rs1055144, rs7798002, rs1451385, Heid et al., 2010) and BMI (rs10261878; Monda et al., 2013). rs1451385 is located in a haplotype block of 48 kb (LD $r^2 > 0.8$) that is delimited by the rs10245353 and rs1451385. rs10261878 (associated with BMI) is not included in this haplotype block; instead it is linked to a haplotype block that includes *MIR148A*. The SNP rs1534696 at *SNX10* is located in an independent haplotype. SNPs indicated by # are located at transcriptional peaks as indicated by CAGE. The dotted line represents 17 SNPs with a weak association with BMI (-log p values >3.5; Speliotes et al., 2010).

(B) Transcriptional signals during adipogenesis of mesenchymal stem cells visualized using the ZENBU browser (http://fantom.gsc.riken.jp/zenbu/). The variant rs1451385 is located in a transcriptionally active region of 400 bp delimited by CAGE25894547 and CAGE25894945. The directions of the transcription are indicated with green and purple arrows for forward and reverse directions, respectively. The intensities of the green and purple peaks correspond to the amount of CAGE-RNAs captured during the entire induction of adipogenesis.

(C) CAGE transcriptional signals colored by intensity on the black background and ordered from top to bottom from 1 h after induction of adipogenesis to day 14. The signals are aligned with the corresponding peaks in (B). A transcriptional promoter-like and an enhancer RNA (eRNA) region are predicted by FANTOM5 CAGE data and indicated with white filled and dotted arrows, respectively. The transcriptional like-promoter sequence is annotated as a POL2-binding site by ENCODE chromatin immunoprecipitation sequencing data, and the enhancer region is annotated as an enhancer according to the Regulatory Elements from ORegAnno at the UCSC Genome Browser in various cell lines.

(D) Numerical quantification of the CAGE signals as in (C) showing the three biological replicates for each time point. Maximal transcription is scored in both directions at 12 h post-induction.

Figure 1. *Continued*

(E) Chromosomal position at the short arm of chromosome 7 and conservation of the region flanking rs1451385 (marked by a red line) depicted by the vertebrate Multiz Alignment at the UCSC Genome Browser. Based on this phylogenetic conservation in mammals, an arbitrary sequence of 266 bp was chosen for further molecular characterization. The CAGE transcriptional signals are displayed as in (B).

In this study, we identified and characterized this active enhancer. Disrupting it causes the loss of differentiation capacity from precursor cells to mature adipocytes. Furthermore, combining CRISPR-Cas9 gene editing with global RNA sequencing (RNA-seq), we identified and characterized potential downstream targets of the enhancer. We aimed to disentangle the molecular mechanisms behind one of the first and strongest GWAS signals associated with obesity-related traits.

RESULTS

In Silico Analyses Highlight a Regulatory Region Tagged by GWAS Potentially Involved in Adipocyte Biology

Using the GWAS catalog (www.ebi.ac.uk/gwas/), we identified all lead variants associated with obesity-related traits (WHR, BMI, lipid levels, glucose-related traits) at the time of study initiation (October 2015). Next, we selected all common single nucleotide polymorphisms (SNPs) in high linkage disequilibrium (LD; $r^2 \geq 0.8$) with the lead variants and searched for overlaps with annotated enhancers using FANTOM5 CAGE data from adipogenesis (Arner et al., 2015; Ehrlund et al., 2017) favoring (1) enhancers with high gene expression during induction of adipogenesis from mesenchymal stem cells and (2) enhancers showing higher expression levels in adipocytes and preadipocytes than in other cell types. This approach identified 14 putative adipocyte-specific enhancers with supporting evidence from GWAS of obesity-related traits. Of these, nine had weak human genetics support (non-genome-wide significant association and/or reported in a small GWAS without further replication), two were associated with proinsulin and low-density lipoprotein cholesterol, respectively (making them less attractive for studies in adipocytes), and two were upstream of *NEGR1* (which is an already well-studied gene). In contrast, the enhancer in 7p15.2 had strong evidence for involvement in adipogenesis from FANTOM5, represented a very strong GWAS signal associated with WHR adjusted for BMI, and had supporting evidence from transcription-binding motifs and chromatin immunoprecipitation sequencing from ENCODE, but little was known about its function.

In this GWAS locus, the FANTOM5 enhancer data highlighted a bidirectional transcription start site active during adipocyte differentiation flanking rs1451385, suggesting that this SNP is located in a functional element of relevance for adipogenesis, as well as fat distribution (Figures 1B–1D). Furthermore, data from ENCODE and the Roadmap Epigenomics Project show that rs1451385 colocalizes with a DNaseI hypersensitive site and with dynamic changes of chromatin acetylation of histone 3 lysine 27 (H3K27ac), suggesting an active regulatory element in this region. The sequence comprising rs1451385 co-immunoprecipitates with the following transcription factors: CTCF (CCCTC-binding factor), USF1 (upstream transcription factor 1), IRF1 (interferon regulatory factor 1), and POLR2A (RNA polymerase II subunit A). USF1 is a transcription factor controlling expression of several genes involved in lipid and glucose homeostasis (Putt et al., 2004) that has been linked to familial combined hyperlipidemia (Pajukanta et al., 2004). USF1 binds to E-box motifs (5′-CACGTG-3′), and it is positioned about 23 nucleotides from a pyrimidine-rich region (Massari and Murre, 2000). The rs1451385 polymorphism is located 2 bp upstream of the E-box motif consensus sequence (5′-ACACGTGA-3′), which is located 23 nucleotides upstream of a 31-nucleotide-long pyrimidine rich-region (23 pyrimidines of a total of 31 nucleotides). Furthermore, analyses of nucleotide-binding sequences using PROMO (Messeguer et al., 2002) show differences in putative binding sites for several transcription factors. Specifically, the T allele of rs1451385, which is in perfect LD with the risk allele (A) at rs3902751 associated with increased WHR (Shungin et al., 2015), confers an extra VDR (vitamin D receptor)-binding site, a FOXP3 site, and a PXR-1:RXR-alpha (a nuclear receptor involved in metabolism sensing)-binding site. Finally, rs1451385 is a borderline significant eQTL ($p = 6.7 \times 10^{-6}$) in subcutaneous adipose tissue in GTEx (www.gtexportal.org/home) for AC003090.1 (ENSG00000223561.2), a long intergenic noncoding RNA (lincRNA) with unknown function, located 100 kb from rs1451385. However, due to its low expression in adipose tissue, it is unlikely that this lincRNA plays an important role in adipogenesis. There were no other eQTLs for subcutaneous or visceral fat in this locus.

The Sequence Comprising the GWAS Signal Shows High Enhancer Activity

We used a luciferase reporter assay to test the regulatory function of the region surrounding rs1451385. Based on phylogenetic conservation (Figure 1E), we cloned a 266-bp fragment (from now on called

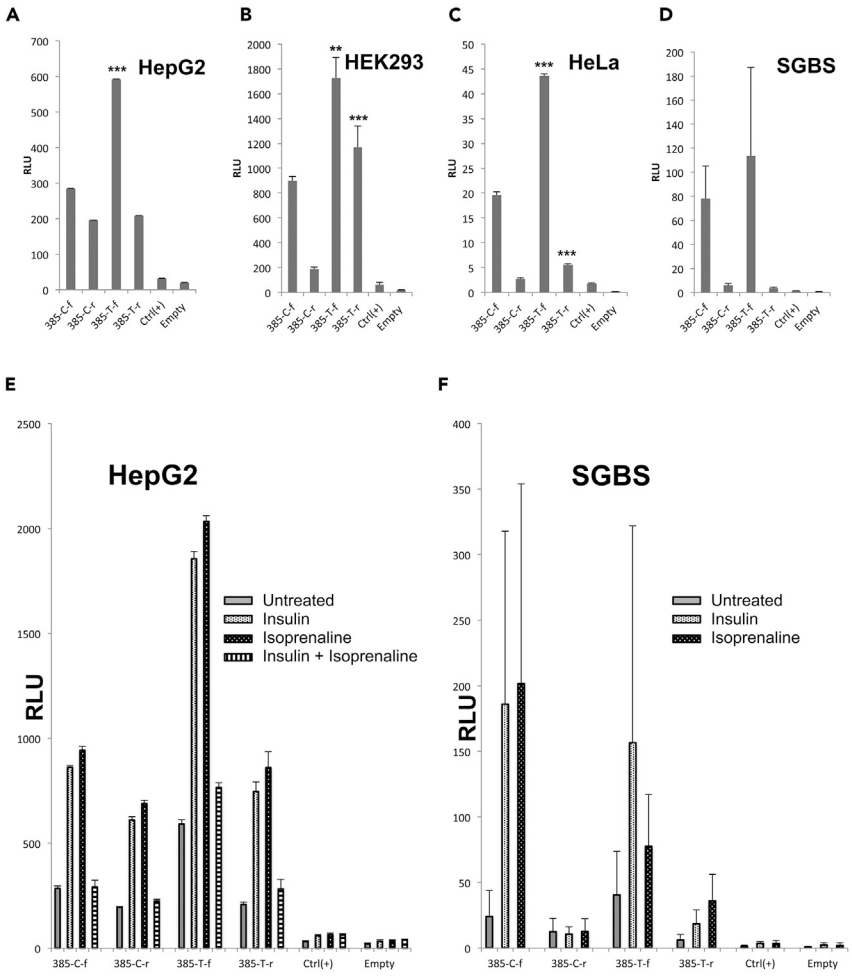


Figure 2. Enhancer Tagged by GWAS Signal Shows High Activity and Response to Insulin

(A–D) rs1451385 shows allelic differences and is located in a strong enhancer. DNA fragments of 266 bp (Enh#385) containing the two allelic variants of rs1451385 (C or T) were tested for enhancer activity in (A) HepG2, (B) HEK293, (C) HeLa, and (D) SGBS human cell lines in transient reporter assays using 12–14 biological replicates. Both forward (f) and reverse (r) directions were tested along with a positive control (Cavalli et al., 2016) and empty vector as negative control. (E and F) The enhancer is inducible by insulin and isoprenaline, compounds that have profound effects on adipocytes and are involved in insulin resistance. Reporter assays comparing luciferase expression after treatment of the indicated cell lines with insulin (100nM) and/or isoprenaline (3 μM) for 24 h. Data are presented as means ± SD. **$p < 0.01$, ***$p < 0.001$ comparing the C allele to T allele cloned in the same direction using a two-tailed unequal variance Student's t test. All the Enh#385 construct variants increased luciferase expression when compared with the empty vector ($p < 0.01$), and the positive control consistently showed a significant increase in expression when compared with the empty vector ($p < 0.05$). (E) Significant induction with insulin and isoprenaline was observed in HepG2 cells for all constructs when compared with non-treated cells ($p < 0.001$). (F) In SGBS cells, significant induction was only achieved with the constructs 385-C-f and 385-T-r ($p < 0.05$; Figures 2E and 2F).

Enh#385) into the reporter vector pGL4.10 (Promega) with a minimal promoter upstream of the luciferase gene. The two variants (C/T) of rs1451385 were cloned in both directions in relation to the luciferase gene. As a positive control, we chose a human 411-bp enhancer from chromosome 12 that consistently increases luciferase expression 3-fold in a variety of cells lines (Cavalli et al., 2016).

We tested the four variants of Enh#385, the positive and empty controls in several cell lines. In all tested cell lines, we found a strong luciferase activity indicating a functional DNA fragment (Figures 2A–2D). The activity varied depending on the cell line and the Enh#385 variant used. The positive control enhancer increased luciferase activity to about 4-fold on average in all the cell lines tested, whereas Enh#385 increased luciferase activity up to 300-fold when compared with the empty vector. We observed an overall

higher activity of the T allele, which is on the same haplotype as the WHR-increasing variant (Shungin et al., 2015).

The Enhancer Enh#358 Is Potentiated by Insulin and Isoprenaline

Next, we tested the transcriptional response of Enh#385 to insulin and isoprenaline in the reporter assay. Such a response would indicate a role of the enhancer in adipocyte biology and/or IR given that both compounds have profound effects on adipocytes and are involved in IR. Isoprenaline, also called isoproterenol, is a non-selective β-adrenoreceptor agonist that increases intracellular cAMP activity, stimulates lipolysis, and inhibits insulin-stimulated glucose transport, whereas insulin stimulates glucose uptake via Glut4 translocation and activation.

In both human hepatocytes (HepG2) and preadipocytes (SGBS), we observed a strong induction of luciferase activity with both substances (Figures 2E and 2F). The induction was suppressed with simultaneous treatment in HepG2 (Figure 2E). This suppression might reflect the opposite physiological effect of the lipogenic insulin and the lipolytic isoprenaline. At a molecular level, this type of suppression could be explained by competitive binding mechanisms, squelching, or transcriptional interference (Kamei et al., 1996; Levine and Manley, 1989; Manna and Stocco, 2007; Step et al., 2014; Yang-Yen et al., 1990; Zhang and Teng, 2001), resulting in insulin-induced factors competing with isoprenaline-induced factors for binding to Enh#385 (Figure S1). The formation of heterogeneous complexes could block the enhancer activity; a similar model has been proposed for scaffold protein complexes (Ferrell, 2000).

Molecular Dissection of the Enhancer

To define the boundaries and functional components of the enhancer, we used a luciferase assay to assess 20 independent constructs corresponding to different parts of Enh#385 with the T allele in forward direction (Figure 3A). Owing to poor transfection efficiency of SGBS cells, we restricted these experiments to HepG2 cells. To assess potential biotechnological applications of Enh#385, we included the enhancer of human CMV (cytomegalovirus), which is the most powerful enhancer currently used in commercial applications as a comparison. Based on phylogenetic conservation, we defined six sequential blocks (Figure 3A). Luciferase expression showed construct Enh#385-17 as having the highest enhancer activity (Figure 3B). The intensity of this construct was twice that of the original Enh#385 (Enh#385-11) but only 36% of the CMV enhancer. Furthermore, we tested the inducible properties of the most potent constructs (Figure 3C). Enh#385-11 showed the highest response to insulin and isoprenaline, whereas the CMV enhancer was not affected by either substance. In summary, our results indicate the presence of a core enhancer region centered around rs1451385.

CRISPR-Cas9 Mutation of the Enhancer Impairs Lipid Accumulation in Differentiating SGBS Cells

To study the function of the enhancer, we generated two independent mutations in undifferentiated preadipocytes using CRISPR-Cas9 (Figure S2).

After supplementation with adipogenic agents, we compared adipocyte differentiation of wild-type (WT) and mutated cultures after knockout (KO) of Enh#385 (KO-385-V11 and KO-385-V3). The WT cells differentiated to mature adipocytes in a normal fashion as evidenced by accumulation of lipid droplets, whereas the mutated cells showed highly reduced differentiation capacity (Figure 4A). Quantification of lipid droplets showed a 5-fold reduction of droplets in Enh#385 KO cells after 19 days of differentiation when compared with cells edited with a guide targeting a random intergenic region (Ctrl-24) or WT cells (Figure 4B). Additional experiments with independent CRISPR-Cas9 transductions provided similar results when measuring lipid content with oil red O staining (Ramirez-Zacarias et al., 1992) (Figure S3) or staining with the fluorophore BODIPY 493/503 (Majithia et al., 2014; Warnke et al., 2011) (data not shown).

Markers of Adipocyte Differentiation Are Reduced in Cell Cultures with Disrupted Enhancer

The impairment in differentiation was confirmed at the gene expression level by RT-qPCR of three adipocyte markers with divergent functions: the phosphoenolpyruvate carboxykinase 1 kinase (*PCK1*), involved in glucose metabolism; the lipoprotein lipase (*LPL*), involved in metabolism of fat; and the secreted adipokine adiponectin (*ADIPOQ*), involved in the control of fat metabolism and insulin sensitivity. In addition, insulin receptor (*INSR*) was included in this hypothesis-driven experiment owing to the potent induction by insulin of Enh#385. After 8 days of differentiation measured by RT-qPCR, all three adipocyte markers were strongly

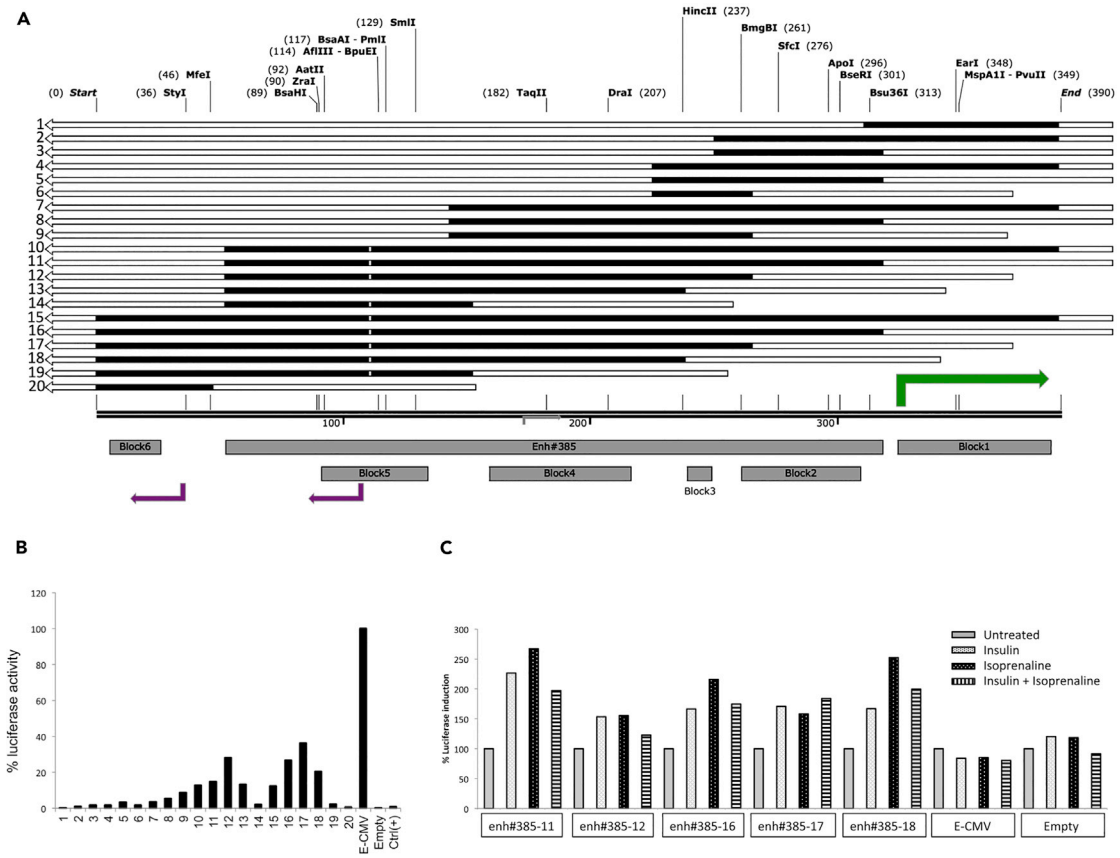


Figure 3. Molecular Dissection of the Enhancer

(A) Restriction map of the constructs used for the luciferase assay. The original Enh#385 sequence and the phylogenetic blocks are marked at the bottom. The SNP rs1451385 is marked as a white square centered in Block5. CAGE transcription signals are indicated with green and purple arrows as in Figures 1B and 1C.

(B) Comparison of luciferase activity between constructs containing variations of Enh#385 and the CMV enhancer (E-CMV). The luciferase activity is displayed as percentage activity of E-CMV. The length of the fragments varies between the 62 bp of the construct Enh#385-20 to the 404 bp of the construct Enh#385-16; the E-CMV is 382 bp in length.

(C) Comparison of the inducible properties of the most potent enhancers from (B) and normalized to the luciferase activity of untreated parallel cultures.

reduced and INSR was slightly reduced in the KO-385-V3 culture (Figure 4C; Figure S4); these results were subsequently confirmed by RNA-seq (next section).

RNA Sequencing Highlights Several Genes Showing Differential Expression after Mutation of the Enhancer

To understand the downstream consequences of the editing of Enh#385, we compared gene expression between WT and Enh#385-mutated cells after 8 days of induced differentiation by global RNA-seq. We included a culture transduced with a single-guide RNA targeting an unrelated intergenic region as a comparison to cope with the effect of transduction and the ectopic expression of the Cas9 nuclease. We focused on the most differentially expressed genes across the whole transcriptome to address downstream effects of Enh#385 and to help disentangle the role of this element on adipocyte differentiation (Figure 5A; Table S1). Among the top differentially expressed genes, we selected 20 genes for technical and biological validations (Table 1; Figure 5A). Technical validations were done by RT-qPCR (i.e., using a different technique than RNA-seq). The technical replication was high: differential gene expression assessed by RT-qPCR was confirmed for 19 of 20 genes when using cells mutated with the same CRISPR-Cas 9 vector that was used in cell cultures undergoing RNA-seq (downstream mutant [M(ds)] KO-385-V3). One gene, *HYDIN*, was impossible to consistently amplify by RT-qPCR, probably owing to its low expression in SGBS cells. Although the two methods correlated very well, we noted that especially for genes with low expression,

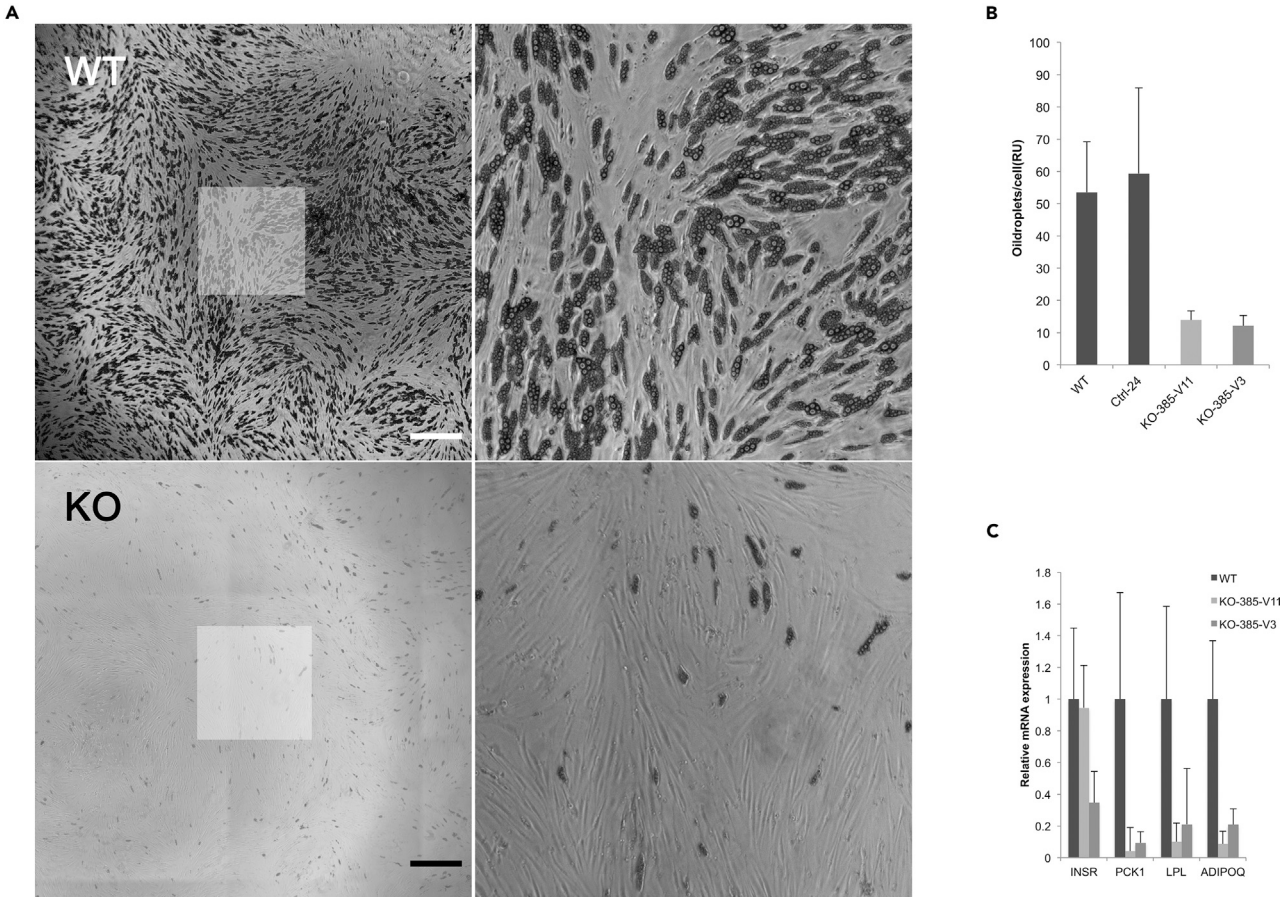


Figure 4. Mutations of the Enhancer Impair Lipid Accumulation in Differentiating SGBS Cells

(A) Representative photographs of wild-type (WT) and mutated (KO-385-V3) differentiated adipocytes cultures. The faded areas in the pictures on the left are enlarged on the right, where the oil droplets are visible in the differentiated adipocytes. Scale bars, 200 μm.

(B) Quantification of adipogenesis by measuring oil droplets in the cell culture normalized by the total cell number after 19 days of differentiation. The KO-385-V11 and KO-385-V3 are cultures targeted with CRISPR-Cas9 upstream and downstream of the rs1451385, respectively, whereas Ctrl-24 targeted a random intergenic region. Data are presented as means ± SD.

(C) Quantitative PCR of WT and enhancer mutants for insulin receptor (INSR), phosphoenolpyruvate carboxykinase 1 (PCK1), lipoprotein lipase (LPL), and adiponectin (ADIPOQ). Data are presented as means ± SD.

the RNA-seq technology was more accurate than RT-qPCR. Two types of biological validations were done: first, using an independent mutation generated with a different CRISPR-Cas9 vector (KO-385-V11) targeting the sequence just upstream of rs1451385 (hence, annotated M(us) in Table 1) and second, using an innocuous lentivirus containing only GFP to assess a potential effect of the transduction protocol on gene expression (annotated as GFP in Table 1). Biological validation was confirmed for 11 of the remaining 19 genes when using the independent CRISPR-Cas9 upstream mutation, but when we evaluated the potential effects of the lentiviral transduction protocol *per se* on gene expression, we found that *HEPH*, *NDN*, and *RCAN2* were downregulated in transduced cells (without the Enh#385 mutation), suggesting that the differential expression of these genes was, at least in part, due to the transduction protocol, rather than due to the mutation in Enh#385 (Table 1 and Figure 5B).

Next, to investigate a potential role in adipogenesis of the eight differentially expressed genes passing both technical and biological validation, we examined their expression in undifferentiated preadipocytes (Table 1). We assumed that only genes that were differentially expressed already in preadipocytes could be responsible for the impairment in adipogenesis, whereas a large number of differentially expressed genes represent the consequence of altered adipogenesis of the mutated cultures rather than causal genes. After comparisons of gene expression by RT-qPCR between undifferentiated downstream mutants, upstream mutants, GFP-transduced cells, and WT cells, we found that only chitinase-3-like protein 1 (*CHI3L1*) on

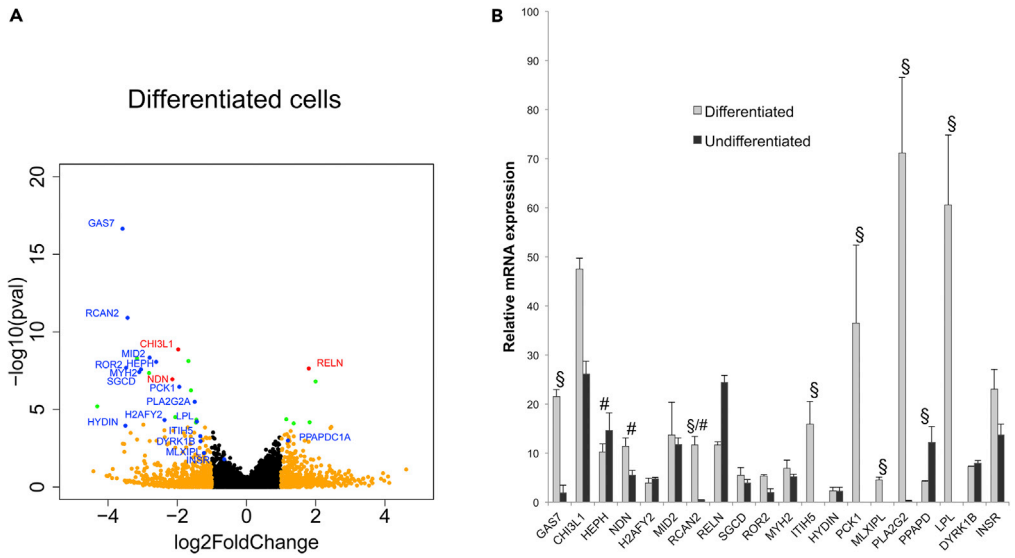


Figure 5. Differential Gene Expression in Differentiated Adipocytes

(A) Volcano plot comparing gene expression in differentiated adipocytes between wild-type and Enh#385 mutated cell cultures. Genes with significant differential expression (DE; $p_{adj} < 0.05$ and abs(log_2 fold change) > 1) are in green, whereas those with abs(log_2 fold change) > 1 are in orange. The 20 genes selected for subsequent validation by RT-qPCR are in blue, and the three genes subject to further investigation (*CHI3L1*, *RELN*, and *NDN*) are in red. Positive values mean that those genes are upregulated in mutants (*RELN* and *PPAPDC1A*), whereas negative values mean that they are downregulated. These results correspond to those shown in Table 1 (column 3).

(B) Comparison of relative gene expression from qPCR between differentiated SGBS adipocytes and undifferentiated SGBS preadipocytes for the 20 differentially expressed genes selected for validation with RT-qPCR using three biological replicates (Table 1). Genes wherein the expression was significantly changed more than 2-fold by the transduction protocol itself are marked with #. Adipocyte-specific genes, wherein the expression changed more than 2-fold between undifferentiated and differentiated cells, are marked with §. *RCAN2* was both sensitive to transduction and adipocyte-specific and hence marked with both § and #. Data are presented as means ± SD.

chromosome 1 and reelin (*RELN*) on chromosome 7 were differentially expressed in undifferentiated pre-adipocytes (at least 2-fold difference; Table 1). Several genes with strong prior evidence for involvement in metabolic disease, such as *ITIH5*, *PCK1*, *MLXIPL*, and *LPL*, were only expressed in mature adipocytes, and hence unlikely to be causal in driving the decrease in adipogenesis, but rather be a result of the difference in cell fate between the Enh#385 KO and WT cells (Table 1; Figure 5B). In subsequent experiments using global gene expression analyses, we were able to extend our search of causal genes and identified altered pathways important for adipogenesis (see below, Table S4 and Figure 7).

Ablation of *CHI3L1* Decreases Adipogenesis in Differentiating SGBS Cells

To evaluate the direct role of the two consistently differentially expressed genes on adipogenesis, we used CRISPR-Cas9 to disrupt *CHI3L1* and *RELN* in SGBS cells (Figure S9). After induction of adipogenesis, the CHI3L1-KO cells showed a significant reduction in lipid accumulation, whereas RELN-KO cells did not differ when compared with controls with regard to differentiation capacity (Figure 6). The phenotype similarity obtained by direct mutation of *CHI3L1* and reduction of CHI3L1 via mutation of Enh#385 indicates that *CHI3L1* and the enhancer are linked by functional interaction circuits. RELN was upregulated in Enh#385 mutant cells; therefore an increase of lipid accumulation in RELN-KO cells could be expected. We did not detect any differences in lipid accumulation between the WT and RELN-KO cells, but we cannot rule out that the overexpression of this gene could reduce the rate of adipocyte differentiation.

CHI3L1, also known as YKL-40, is a secreted glycoprotein coupled with stress-induced cellular responses (Ling and Recklies, 2004). Protein levels of CHI3L1 have been associated with several pathogenic processes including schizophrenia, asthma, obesity, and cancer, but the biological function of *YKL-40* in specific tissues is largely unknown (Kyrgios et al., 2012; Ober et al., 2008; Zhao et al., 2007). Experimental data have shown that mice deficient in *Chi3l1* develop less visceral obesity and have smaller adipocytes; in contrast,

		Differentiated SGBS Cells									Undifferentiated Cells						
		RNA-Seq			RT-qPCR Validation						RT-qPCR						
		M(ds)/WT			M(ds)/WT		M(us)/WT		GFP/WT		M(ds)/WT		M(us)/WT		GFP/WT		
Gene Symbol	Full Gene Name	Log2FC	p Value*	Adjusted p	Log2FC	p Value*	Log2FC	p Value*	Log2FC	p Value*	Log2FC	p Value*	Log2FC	p Value*	Log2FC	p Value*	Comment
GAS7	Growth arrest-specific 7	−3.45	4.2×10^{-9}	3.8×10^{-13}	−3.19	1.2×10^{-2}	0.67	2.1×10^{-1}									Not validated M(us)
CHI3L1	Chitinase 3-like 1 cartilage glycoprotein-39	−2.03	1.9×10^{-8}	5.7×10^{-6}	−3.46	4.0×10^{-3}	−1.57	8.0×10^{-3}	0.88	3.9×10^{-1}	−2.12	5.4×10^{-3}	−1.05	2.4×10^{-2}	−0.12	7.4×10^{-1}	Putative causal gene
HEPH	Hephaestin	−2.87	2.3×10^{-7}	1.8×10^{-5}	−5.47	4.9×10^{-3}	−2.24	1.4×10^{-2}	−2.82	1.6×10^{-2}							Downregulated by transduction
NDN	Necdin melanoma antigen MAGE family member	−2.15	1.9×10^{-6}	1.4×10^{-4}	−4.71	3.8×10^{-4}	−1.88	4.2×10^{-3}	−2.70	9.8×10^{-4}							Downregulated by transduction
H2AFY2	H2A histone family member Y2	−2.62	2.0×10^{-6}	3.4×10^{-2}	−2.70	4.5×10^{-2}	0.92	3.6×10^{-2}									Not validated M(us)
MID2	Midline 2	−2.67	1.1×10^{-5}	1.4×10^{-5}	−4.35	3.4×10^{-4}	0.60	6.8×10^{-2}									Not validated M(us)
RCAN2	Regulator of calcineurin 2	−2.63	1.5×10^{-5}	1.1×10^{-7}	−4.19	2.5×10^{-2}	−3.20	3.9×10^{-2}	−3.35	3.1×10^{-2}							Downregulated by transduction
RELN	Reelin	1.82	1.7×10^{-5}	3.9×10^{-5}	1.41	4.2×10^{-2}	1.92	1.5×10^{-3}	0.44	4.2×10^{-1}	1.48	2.4×10^{-2}	1.06	3.4×10^{-2}	0.83	7.8×10^{-2}	Putative causal gene
SGCD	Sarcoglycan delta 35 kDa dystrophin-associated glycoprotein	−3.43	2.4×10^{-5}	5.5×10^{-5}	−2.03	4.4×10^{-2}	0.10	8.9×10^{-2}									Not validated M(us)
ROR2	Receptor tyrosine kinase-like orphan receptor 2	−2.91	2.8×10^{-5}	3.9×10^{-5}	−1.39	5.2×10^{-2}	0.63	5.0×10^{-2}									Not validated M(us)

Table 1. Validation and Filtering of Candidate Genes for Further Studies among 20 Top-Ranked Differentially Expressed Transcripts

(Continued on next page)

		Differentiated SGBS Cells									Undifferentiated Cells						
		RNA-Seq			RT-qPCR Validation						RT-qPCR						
		M(ds)/WT			M(ds)/WT		M(us)/WT		GFP/WT		M(ds)/WT		M(us)/WT		GFP/WT		
Gene Symbol	Full Gene Name	Log2FC	p Value*	Adjusted p	Log2FC	p Value*	Log2FC	p Value*	Log2FC	p Value*	Log2FC	p Value*	Log2FC	p Value*	Log2FC	p Value*	Comment
MYH2	Myosin heavy chain 2 skeletal muscle adult	−2.77	2.8×10^{-5}	4.1×10^{-5}	−1.65	3.9×10^{-2}	0.47	3.7×10^{-1}									Not validated M(us)
ITIH5	Inter-alpha-trypsin inhibitor heavy chain family member 5	−1.32	2.9×10^{-5}	1.9×10^{-1}	−3.17	4.6×10^{-5}	−5.01	1.0×10^{-5}	−1.56	8.1×10^{-3}	Not expressed						Adipocyte specific
HYDIN	HYDIN axonemal central pair apparatus protein	−3.22	3.2×10^{-5}	6.3×10^{-2}	ND												
PCK1	Phosphoenol-pyruvate carboxykinase 1 soluble	−1.90	4.1×10^{-5}	3.7×10^{-4}	−2.03	7.7×10^{-3}	−2.94	2.9×10^{-3}	0.70	4.5×10^{-1}	Not expressed						Adipocyte specific
MLXIPL	MLX-interacting protein-like	−1.67	5.3×10^{-5}	8.6×10^{-1}	−1.51	2.7×10^{-2}	−0.73	5.3×10^{-2}	−0.62	2.8×10^{-1}	Not expressed						Adipocyte-specific
PLA2G2A	Phospholipase A2 group IIA platelets synovial fluid	−1.51	6.1×10^{-5}	2.9×10^{-3}	−1.32	2.2×10^{-3}	−2.81	3.1×10^{-4}	1.25	8.4×10^{-3}	−0.17	8.0×10^{-1}	−0.42	4.1×10^{-1}	0.34	4.2×10^{-1}	Adipocyte specific
PPAPDC1A	Phosphatidic acid phosphatase type 2 domain-containing 1A	1.31	7.9×10^{-5}	2.7×10^{-1}	2.18	4.7×10^{-4}	1.45	3.8×10^{-2}	0.55	3.0×10^{-1}	−0.05	9.2×10^{-1}	−0.75	5.3×10^{-2}	−0.45	1.8×10^{-1}	Preadipocyte specific
LPL	Lipoprotein lipase	−1.40	1.3×10^{-4}	4.1×10^{-2}	−2.04	2.5×10^{-9}	−2.90	2.1×10^{-11}	−0.07	4.7×10^{-1}	Not expressed						Adipocyte specific

Table 1. ***Continued***

(Continued on next page)

		Differentiated SGBS Cells										Undifferentiated Cells						
		RNA-Seq			RT-qPCR Validation							RT-qPCR						
		M(ds)/WT			M(ds)/WT		M(us)/WT		GFP/WT			M(ds)/WT		M(us)/WT		GFP/WT		
Gene Symbol	Full Gene Name	Log2FC	p Value*	Adjusted p	Log2FC	p Value*	Log2FC	p Value*	Log2FC	p Value*		Log2FC	p Value*	Log2FC	p Value*	Log2FC	p Value*	Comment
DYRK1B	Dual-specificity tyrosine-Y-phosphorylation regulated kinase 1B	−1.61	1.8×10^{-4}	2.7×10^{-1}	−2.43	8.9×10^{-6}	−0.15	5.7×10^{-1}										Not validated M(us)
INSR	Insulin receptor	−0.83	8.1×10^{-3}	2.7×10^{-1}	−0.89	2.6×10^{-2}	0.56	8.4×10^{-2}										Not validated M(us)

Table 1. Continued

Genes were selected for confirmatory RT-qPCR analyses based on differential expression (DE) in RNA-seq analyses, excluding genes wherein DE was caused by the transduction procedure and genes with very low gene expression (see Transparent Methods supplemental file for details). The criteria for filtering out false-positives and selecting candidate genes for downstream experiments based on the RT-qPCR validation were three-fold: (1) A technical validation using RNA from the same cells as those used for the global transcriptome analysis requiring significant differentially expressed in the same direction as in RNA-seq for successful validation (columns 6 and 7 [M(ds)/WT]). (2) A biological validation using RNA from an independent mutation (columns 8 and 9 [M(up)/WT]). (3) A control of potential transduction effects in which the candidate genes should not show differential expression comparing non-transduced WT cells and cells transduced with a lentivirus expressing only GFP (columns 10 and 11 [GFP/WT]). Validated genes were further interrogated for possible involvement in adipogenesis by studying their expression in undifferentiated cells. Genes not expressed in preadipocytes or with a >2-fold difference in expression between the differentiated and undifferentiated state were considered adipocyte specific (Figure 5B), and likely a consequence of the phenotype and not causative genes. RT-qPCR data of three independent experiments; p values comparing expression in mutated cells with WT cells (as $\log_2$ fold change [FC]) were calculated by Student's t test with Benjamini-Hochberg correction (columns 7, 9, 11, 13, 15, and 17).
* P-values are from two-sample Welch t-tests, unadjusted for multiple testing.

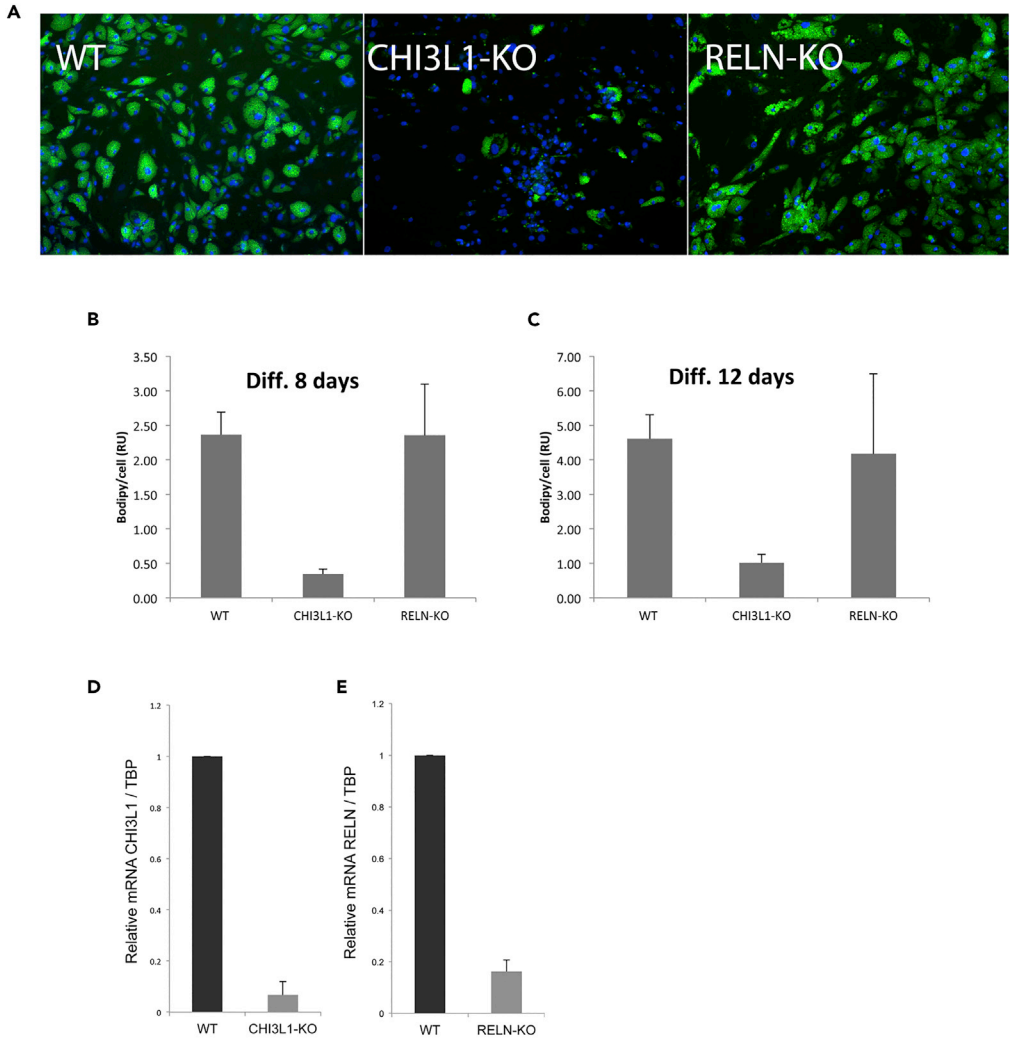


Figure 6. Adipocyte Differentiation Capacity of Wild-Type and Mutant SGBS Cells Estimated as Accumulation of Neutral Lipids

(A) Representative images of BODIPY and Hoechst staining, in green (lipids) and blue (nuclei), respectively, of differentiating SGBS cells. From left to right: wild-type (WT), CRISPR-edited *CHI3L1* (CHI3L1-KO), and CRISPR-edited *RELN* (RELN-KO) cells.

(B) Quantification of neutral lipids after 8 days of differentiation.

(C) Quantification of neutral lipids after 12 days of differentiation. Both at 8 and 12 days after induction of differentiation, the lipid accumulation was significantly reduced in *CHI3L1* mutants ($p < 0.001$), whereas *RELN* mutants were no different from WT. The values in (B and C) represent the mean ± SD of a minimum of three cell cultures.

(D) Relative *CHI3L1* expression comparing wild-type cultures and CRISPR-edited SGBS cells (knocking down *CHI3L1*).

(E) Relative *RELN* expression comparing wild-type cultures and CRISPR-edited SGBS cells (knocking down *RELN*). The data in (D and E) are presented as mean ± SD.

its overexpression induced adiposity (Ahangari et al., 2015). The precise molecular mechanism of *CHI3L1* on adipogenesis has not been determined, but a direct role on the extracellular matrix through type I collagen has been proposed (Iwata et al., 2009; Mariman and Wang, 2010). It is possible that *CHI3L1* is required for the remodeling of the extracellular matrix that precedes the adipocyte differentiation process. In a mouse model of osteomyelitis, partial restoration of osteogenesis has been achieved by suppressing *Chi3l1* (Chen et al., 2017). This suggests that *CHI3L1* might play a key role in cell fate decision (Figure 7).

Enh#385 Regulates Genes Involved in Adipocyte Biology and Nearby Genes

We specifically examined the expression of genes linked to adipocyte biology in our RNA-seq data from differentiated cells. We confirmed the reduction in expression of the markers previously measured by

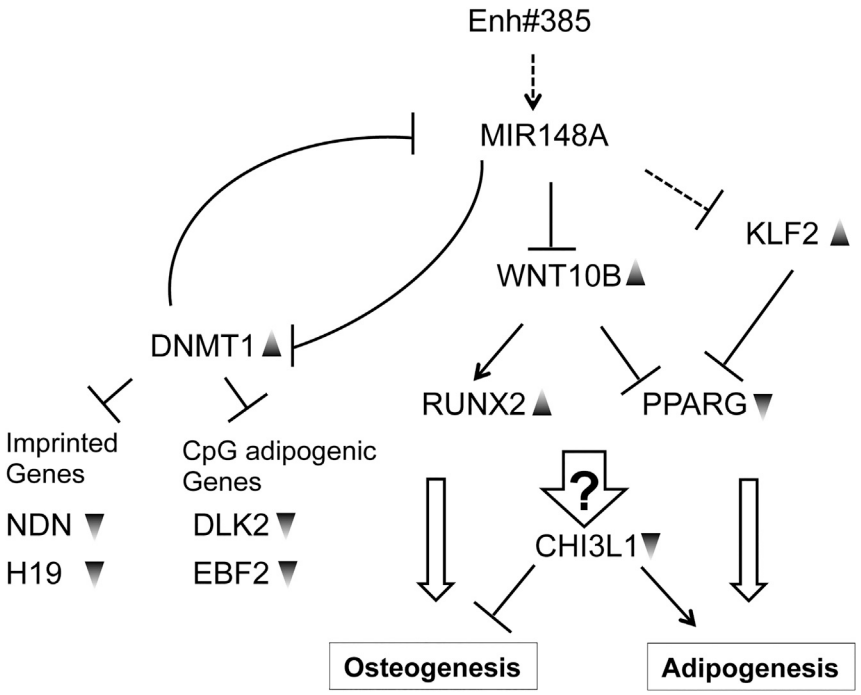


Figure 7. A Schematic Illustration of Key Adipogenic Factors that We Propose as Being Regulated by the Enhancer (Enh#385)

Arrows with dotted lines annotate predictions based on the present work, whereas solid arrows indicate associations experimentally demonstrated and reported in prior literature. Arrowhead lines indicate transcriptional activation; blunt-head lines indicate inhibition. Gradient arrowheads after the name of the gene indicate the direction of the expression measured in our mutated cultures when compared with wild-type cell cultures. The upstream regulation of *CHI3L1* is uncertain and indicated with a question mark.

RT-qPCR and observed reduction of several additional adipocyte markers. For example, *LEP*, *PPARGC1A*, *PLIN1*, *CD36*, and *SLC2A4* (GLUT4) were reduced in mutant differentiated cultures. We noted that expression of some key proteins involved in fat synthesis (*FASN*), transport of lipoproteins (*APOE*), and the receptors for *ADIPOQ* (*ADIPOR1* and *ADIPOR2*) were unaffected by the mutations of the enhancer (Table S2).

Next, we explored whether there were any putative *cis*-regulated genes among the differentially expressed genes. We scanned the chromosomal region surrounding rs1451385 (±5Mb) for differentially expressed genes. All genes annotated in the UCSC Browser (https://genome.ucsc.edu/), including protein-coding genes and non-coding RNA genes within this interval, were identified and expression levels were compared between mutant and WT cells. This region included a microRNA (miRNA) (*MIR148A*) and a few genes that are good candidates for involvement in metabolic disease based on prior literature, such as *IGF2BP3* (insulin-like growth factor 2 mRNA-binding protein 3) at chr7:23,349,828 (7p15.3) and *CREB5* (cAMP responsive element-binding protein 5) at chr7:28,865,511 (7p14.1). The only transcript with a significant differential expression in this large region was *NFE2L3* (2.3-fold reduction, $p = 0.009$). However, as *NFE2L3* is expressed at very low levels in SGBS cells, and because we did not find any differential expression in similar experiments using undifferentiated SGBS cells (see below), we did not consider this association further. These results suggest that the enhancer does not act in *cis* on protein-coding genes, but it should be noted that our global RNA sequence method did not allow quantification of short RNAs, such as miRNAs.

Causative Genes Regulated by Enh#385 Uncovered by Differential Gene Expression in Undifferentiated Cells

To extend the list of potential downstream causal genes regulated by Enh#385, we performed another transcriptomics analysis by RNA-seq, this time in undifferentiated SGBS preadipocytes. We compared the two independent mutants of Enh#385 (KO-385-V3 and KO-385-V11) with WT cells and a GFP transduction control (Figure S5; Table S3). In these analyses, we confirmed the downregulation of *CHI3L1* (fold change,

0.49; p_{adj}, 0.0002) and the upregulation of *RELN* (fold change, 1.55; p_{adj}, 0.041) in mutated cells corroborating a potential causative role in driving the decreased adipogenesis noted after mutation of Enh#385. However, *RELN* was also upregulated in GFP-transduced cells, which may indicate that this differential expression was, at least partly, driven by the transduction procedure. In addition, *NDN* (Necdin; fold change, 0.42; p_{adj}, 0.018), associated with the Prader-Willi syndrome in humans, characterized by severe obesity, showed significant DE in undifferentiated preadipocytes, indicating a role as effector transcript downstream of Enh#385. This gene was discarded in the first qPCR validation analysis due to concerns regarding unspecific differential expression induced by the transduction protocol. Comparison of differentiated and undifferentiated WT cells revealed several adipocyte-specific genes (Figure 5B). These genes show higher (or even exclusive) expression in differentiated adipocytes; hence, it is likely that their differential expression in the Enh#385 mutants is a consequence of the lack of adipogenesis, rather than being involved in the causative process.

The most differentially expressed gene was H19, a maternally imprinted lincRNA that interacts with insulin-like growth factor 2 (*IGF2*), and has been suggested to be a tumor suppressor (Bartolomei et al., 1991; Bell and Felsenfeld, 2000). Defects in the H19/IGF2 imprinting have been associated with Beckwith-Wiedemann syndrome (DeBaun et al., 2002), Silver-Russell syndrome (Bartholdi et al., 2009; Bliek et al., 2006), and Wilms tumor 2 (Steenman et al., 1994). Other top differentially expressed genes included several plausible candidates for involvement in adiposity and/or IR, such as *ADH1B*, *IL8*, *ACE*, *IL1B*, and *IL6* (Figure S5; Table S3).

Next, we focused on genes linked to adipocyte biology and adipogenesis based on prior literature (Table S4). We found that *PPARG* was reduced (0.7-fold) in mutant cells, whereas *WNT10B*, encoding a molecular switch that inhibits adipogenesis, was significantly increased (2.5-fold) (Christodoulides et al., 2006; Ross et al., 2000). It has been shown that *WNT10B* inhibits *PPARG* and promotes the expression of the osteoblastogenic transcription factor *RUNX2* (Bennett et al., 2005). Consistent with a potential pivotal role of *WNT10B* in the observed phenotype, we found an increase (1.3-fold) of *RUNX2* in the mutated cells. The anti-adipogenic transcription factor *KLF2* has also been suggested as contributor in this pathway (Banerjee et al., 2003). Again, consistent with this hypothesis, *KLF2* was significantly upregulated (1.4-fold) in our enhancer mutants, thus blocking adipogenesis. The increase of these two important anti-adipogenic regulators, *WNT10B* and *KLF2*, acting upstream of *PPARG* suggests that Enh#385 might function very early in the initiation of the adipogenesis program.

DISCUSSION

Our study offers a detailed characterization of a GWAS locus (previously annotated as *SNX10*) associated with WHR and other obesity-related traits in a series of functional experiments. Our main conclusions are severalfold. First, we identified an enhancer active during adipogenesis that responds strongly to external metabolic stimuli, such as insulin and isoprenaline. The known WHR-raising allele (A at rs3902751) (Shungin et al., 2015) is on the same haplotype as the allele being associated with higher enhancer activity (T at rs1451385), suggesting that higher activity of this enhancer is associated with higher WHR. Second, mutation of the enhancer using CRISPR-Cas9 in preadipocytes dramatically impairs adipocyte differentiation. RNA-seq followed by replication using RT-qPCR in mature adipocytes demonstrated that several adipocyte genes were downregulated in cells with disrupted Enh#385 and highlighted *CHI3L1* as a potential downstream target, also downregulated. Third, disruption of *CHI3L1* caused a decrease in adipogenesis.

Independently of the allele tested, the enhancement of luciferase expression was higher with Enh#385 cloned in forward direction (Figure 2), suggesting that the cloned DNA fragments may have a significant promoter-like activity. Consistent with this, the signal from the CAGE-sequencing experiment was generally stronger in the forward direction throughout adipogenesis (Figure 1D). However, given the long physical distance to the nearest gene, the regulatory element is most likely to be an enhancer with promoter-like activity. This complex nature of the regulatory element has been observed in a previous study, wherein the sequence of Enh#385 was included in a DNA fragment (976 bp) functionally classified as weak anti-repressor element (Kwaks et al., 2003).

CHI3L1 could be a downstream effector of Enh#385 having a role in adipocyte differentiation, but the mechanistic circuits between the Enh#385 and *CHI3L1* function are unclear, especially as they are not located on the same chromosome. We hypothesized that the mediator of the link from Enh#385 to *CHI3L1* and other downstream genes involved in adipogenesis could be *MIR148A*, a non-protein-coding

gene encoding an miRNA localized 95 kb from the enhancer (Figure 1). miRNAs are known to be master regulators of groups of proteins (Stefani and Slack, 2008), and indeed, several miRNAs have been reported to regulate adipogenesis and lipid metabolism (Ahn et al., 2013; Esau et al., 2004; Yang et al., 2011). Indeed, our analysis of enhancer-promoter interactions by HiChiP in a human coronary artery smooth muscle cell line shows connections between the Enh#385 and *MIR148A* (Figure S6).

Searching for *MIR148A* targets using TargetScan (http://www.targetscan.org), we found two conserved sequences at the 3′ UTR of *WNT10B* and two sequences of *KLF2* that acts as binding targets for miR-148a (Figure S7). Consistent with this, we observed a significant increase of *WNT10B* in Enh#385 mutants (Table S4). There is also support in prior literature for involvement of *MIR148A* and *WNT10B* in adipogenesis. In mice, miR-148a is upregulated during adipogenesis and downregulated in mature adipocytes of obese animals (Xie et al., 2009). In a recent report using differentiating human adipose-derived mesenchymal stem cells, it was shown that *MIR148A* induces adipogenesis by suppressing its target gene, *WNT10B*, an endogenous inhibitor of adipogenesis (Shi et al., 2015). It is tempting to speculate that the increase of *WNT10B* and *KLF2,* two anti-adipogenic factors, in our mutants is the result of the reduction of MIR148A (Table S4). This scenario is consistent with the observed increase of *RUNX2* and decrease of *PPARG* in our DE data and with the lack of adipogenesis in the Enh#385 mutant cells. A schematic illustration of this proposed model is shown in Figure 7.

Experimentally, we were able to measure an increase of expression of *MIR148A* between 0 and 8 days of differentiation, but we could not detect significant differences in expression of *MIR148A* between WT and Enh#385 KO cells before or 8 days after induction of differentiation. It is possible that a role of *MIR148A* in regulating adipogenesis is transient, and can only be recorded in a short time window during the differentiation process (and we analyzed differential expression at just two time points). Supporting this interpretation, CAGE-sequencing data (from the FANTOM consortium) measuring the activity of the *MIR148A* promoter during MSC adipocyte differentiation show only a consistent increase in activity at 4 and 14 days, but not at the rest of the 17 time points, indicating a fluctuating and transient regulation of *MIR148A* (Figure S8). The inconsistent results in detection of *MIR148A* in mutant cells could be also due to the negative regulatory feedback loop between *MIR148A* and one of its targets, the methylase *DNMT1* (Figure 7). The methylase downregulates the expression of *MIR148A* by hypermethylation of its promoter (Hong et al., 2018; Long et al., 2014).

The consistent downregulation of the imprinted gene *NDN* (Necdin) leads us to hypothesize that imprinting by methylation of *NDN* could be mediated by the *DNMT1* (DNA methyltransferase 1)-*MIR148A* circuit. *NDN* is a particularly interesting gene for adipocyte biology as it is one of the candidate genes in the chromosomal region 15q11-15q13, which results in Prader-Willi syndrome when the paternal copy is deleted. In addition to mild to moderate intellectual impairment and behavioral problems, obesity and T2D are the most common symptoms of Prader-Willi syndrome, owing to insatiable appetite and chronic overeating (MacDonald and Wevrick, 1997). *DNMT1* is a predicted target of *MIR148A,* and its expression is elevated in adipocytes of obese individuals (Braconi et al., 2010). DNMT1 maintains methylation pattern during adipocyte differentiation, and its silencing accelerated adipogenesis (Londono Gentile et al., 2013). It has been proposed that DNMT1 and *MIR148A* are regulated by a negative feedback loop that could explain the discrepancy in the published reports (Xu et al., 2013). In fact, the promoter of *MIR148A* is in proximity of CpG islands, and it is silenced by hypermethylation (Hanoun et al., 2010). Our differential expression analysis shows that *DNMT1* expression is only slightly increased in mutant cultures; however, it has been reported that the targeting of this gene by *MIR148A* occurs exclusively at the protein level (Pan et al., 2010), which makes the differences in mRNA a less ideal indicator of enzyme methylation activity. Imprinted genes (identified using the Genomic Imprinting website: http://www.geneimprint.com) were generally not affected by the mutations of Enh#385. However, in addition to *NDN*, the adipogenesis-linked and imprinted transcript *H19* was the most differentially expressed gene in preadipocytes, showing a consistent downregulation in mutated cells (Han et al., 2018; Huang et al., 2016). It is plausible that only loci involved in adipogenesis are accessible to *DNMT1*, so it should not be expected that all imprinted genes and genes regulated by CpG methylation are downregulated in Enh#385 mutants. The increased activity of this enzyme may lead to the repression of active adipogenic loci controlled by methylation. Our expression analysis shows that the imprinted gene *NDN* and genes regulated by methylation such as *DLK2* (Delta like non-canonical Notch ligand 2) and *EBF2* (Early B-cell factor 2), both promoting adipogenesis (Jimenez et al., 2007; Nueda et al., 2007), are consistently downregulated in cells with disrupted Enh#385 (Tables S2 and S4). Thus, despite being limited to a few genes, this methylation link reinforces our model in which *MIR148A* is the most likely target of the enhancer Enh#385 (Figure 7).

GWAS have been remarkably successful in discovering loci associated with complex traits, but to unlock the transformative potential of these findings, there is a strong need for detailed studies of the molecular mechanisms underlying these signals. We have performed such a study that improves the understanding of adipocyte biology and fat distribution. Although the exact molecular events that link genetic variation in Enh#385 (previously annotated as the *SNX10* locus) with the impairment of adipocyte differentiation still need to be characterized in detail, one potential explanation for link of the enhancer with adipogenesis could be *cis*-acting regulation through the nearby *MIR148A*. We hypothesize that allelic variants that reduce the activity of Enh#385 could lead to the downregulation of MIR148A. This reduction of *MIR148A* may produce an increase of the cell fate determinant *WNT10B* and the subsequent inhibition of adipogenesis. These initial and complex alterations of gene activity could lead to the reduction of *CHI3L1* expression, which, as we demonstrate in this work, is essential for the adipocyte maturation. Further studies are needed to establish these potential mechanisms linking Enh#385 to adipocyte differentiation.

Limitations of the Study

The model we propose in which the enhancer Enh#385 acts on *MIR148A* is based on indirect observations and is not directly proved in this study. Hence, it should be viewed as hypothesis generating, and we do not discard the possibility that the enhancer also acts on other genes or genetic elements.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2019.09.006.

ACKNOWLEDGMENTS

This project was supported by the National Institutes of Health (R01DK106236), the Swedish Research Council (2015-02907), the Knut and Alice Wallenberg Foundation (2013.0126), the Exploratory Funding Program of Uppsala University Innovation, Exodiab (Excellence of Diabetes Research in Sweden) and a research grant from MEXT to the RIKEN Center for Integrative Medical Sciences. The authors would like to acknowledge support of the National Genomics Infrastructure (NGI) / Uppsala Genome Center (Sweden).

AUTHOR CONTRIBUTIONS

C.C-L., E.A., and E.I. conceived and designed the study. C.C-L., A.P., and S.H. designed, performed, and analyzed the experiments. C.C-L., M.P., and E.I. analyzed gene expression data. M.W. and C.W. contributed with cell lines and reagents. T.Q. contributed with the chromatin interactome data. C.C-L. and E.I. wrote the manuscript with input from the other authors. All authors approved of the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 18, 2019
Revised: July 10, 2019
Accepted: September 5, 2019
Published: October 25, 2019

REFERENCES

Ahangari, F., Sood, A., Ma, B., Takyar, S., Schuyler, M., Qualls, C., Dela Cruz, C.S., Chupp, G.L., Lee, C.G., and Elias, J.A. (2015). Chitinase 3-like-1 regulates both visceral fat accumulation and asthma-like Th2 inflammation. Am. J. Respir. Crit. Care Med. *191*, 746–757.

Ahn, J., Lee, H., Jung, C.H., Jeon, T.I., and Ha, T.Y. (2013). MicroRNA-146b promotes adipogenesis by suppressing the SIRT1-FOXO1 cascade. EMBO Mol. Med. *5*, 1602–1612.

Arner, E., Daub, C.O., Vitting-Seerup, K., Andersson, R., Lilje, B., Drablos, F., Lennartsson, A., Ronnerblad, M., Hrydziuszko, O., Vitezic, M., et al. (2015). Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. Science *347*, 1010–1014.

Banerjee, S.S., Feinberg, M.W., Watanabe, M., Gray, S., Haspel, R.L., Denkinger, D.J., Kawahara, R., Hauner, H., and Jain, M.K. (2003). The Kruppel-like factor KLF2 inhibits peroxisome proliferator-activated receptor-gamma expression and adipogenesis. J. Biol. Chem. *278*, 2581–2584.

Bartholdi, D., Krajewska-Walasek, M., Ounap, K., Gaspar, H., Chrzanowska, K.H., Ilyana, H., Kayserili, H., Lurie, I.W., Schinzel, A., and Baumer, A. (2009). Epigenetic mutations of the imprinted IGF2-H19 domain in Silver-Russell syndrome (SRS): results from a large cohort of patients with SRS and SRS-like phenotypes. J. Med. Genet. *46*, 192–197.

Bartolomei, M.S., Zemel, S., and Tilghman, S.M. (1991). Parental imprinting of the mouse H19 gene. Nature *351*, 153–155.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482–485.

Bennett, C.N., Longo, K.A., Wright, W.S., Suva, L.J., Lane, T.F., Hankenson, K.D., and MacDougald, O.A. (2005). Regulation of osteoblastogenesis and bone mass by Wnt10b. Proc. Natl. Acad. Sci. U S A *102*, 3324–3329.

Bliek, J., Terhal, P., van den Bogaard, M.J., Maas, S., Hamel, B., Salieb-Beugelaar, G., Simon, M., Letteboer, T., van der Smagt, J., Kroes, H., et al. (2006). Hypomethylation of the H19 gene causes not only Silver-Russell syndrome (SRS) but also isolated asymmetry or an SRS-like phenotype. Am. J. Hum. Genet. *78*, 604–614.

Braconi, C., Huang, N., and Patel, T. (2010). MicroRNA-dependent regulation of DNA methyltransferase-1 and tumor suppressor gene expression by interleukin-6 in human malignant cholangiocytes. Hepatology *51*, 881–890.

Cavalli, M., Pan, G., Nord, H., Wallen Arzt, E., Wallerman, O., and Wadelius, C. (2016). Allele-specific transcription factor binding in liver and cervix cells unveils many likely drivers of GWAS signals. Genomics *107*, 248–254.

Chen, X., Jiao, J., He, X., Zhang, J., Wang, H., Xu, Y., and Jin, T. (2017). CHI3L1 regulation of inflammation and the effects on osteogenesis in a Staphylococcus aureus-induced murine model of osteomyelitis. FEBS J. *284*, 1738–1747.

Christodoulides, C., Scarda, A., Granzotto, M., Milan, G., Dalla Nora, E., Keogh, J., De Pergola, G., Stirling, H., Pannacciulli, N., Sethi, J.K., et al. (2006). WNT10B mutations in human obesity. Diabetologia *49*, 678–684.

DeBaun, M.R., Niemitz, E.L., McNeil, D.E., Brandenburg, S.A., Lee, M.P., and Feinberg, A.P. (2002). Epigenetic alterations of H19 and LIT1 distinguish patients with Beckwith-Wiedemann syndrome with cancer and birth defects. Am. J. Hum. Genet. *70*, 604–611.

Ehrlund, A., Mejhert, N., Bjork, C., Andersson, R., Kulyte, A., Astrom, G., Itoh, M., Kawaji, H., Lassmann, T., Daub, C.O., et al. (2017). Transcriptional dynamics during human adipogenesis and its link to adipose morphology and distribution. Diabetes *66*, 218–230.

Esau, C., Kang, X., Peralta, E., Hanson, E., Marcusson, E.G., Ravichandran, L.V., Sun, Y., Koo, S., Perera, R.J., Jain, R., et al. (2004). MicroRNA-143 regulates adipocyte differentiation. J. Biol. Chem. *279*, 52361–52365.

Ferrell, J.E., Jr. (2000). What do scaffold proteins really do? Sci. STKE *2000*, pe1.

Graff, M., Scott, R.A., Justice, A.E., Young, K.L., Feitosa, M.F., Barata, L., Winkler, T.W., Chu, A.Y., Mahajan, A., Hadley, D., et al. (2017). Genome-wide physical activity interactions in adiposity - a meta-analysis of 200,452 adults. PLoS Genet. *13*, e1006528.

Han, Y., Ma, J., Wang, J., and Wang, L. (2018). Silencing of H19 inhibits the adipogenesis and inflammation response in ox-LDL-treated Raw264.7 cells by up-regulating miR-130b. Mol. Immunol. *93*, 107–114.

Hanoun, N., Delpu, Y., Suriawinata, A.A., Bournet, B., Bureau, C., Selves, J., Tsongalis, G.J., Dufresne, M., Buscail, L., Cordelier, P., et al. (2010). The silencing of microRNA 148a production by DNA hypermethylation is an early event in pancreatic carcinogenesis. Clin. Chem. *56*, 1107–1118.

Heid, I.M., Jackson, A.U., Randall, J.C., Winkler, T.W., Qi, L., Steinthorsdottir, V., Thorleifsson, G., Zillikens, M.C., Speliotes, E.K., Magi, R., et al. (2010). Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nat. Genet. *42*, 949–960.

Hong, L., Sun, G., Peng, L., Tu, Y., Wan, Z., Xiong, H., Li, Y., and Xiao, W. (2018). The interaction between miR148a and DNMT1 suppresses cell migration and invasion by reactivating tumor suppressor genes in pancreatic cancer. Oncol. Rep. *40*, 2916–2925.

Huang, Y., Zheng, Y., Jin, C., Li, X., Jia, L., and Li, W. (2016). Long non-coding RNA h19 inhibits adipocyte differentiation of bone marrow mesenchymal stem cells through epigenetic modulation of histone deacetylases. Sci. Rep. *6*, 28897.

Iwata, T., Kuwajima, M., Sukeno, A., Ishimaru, N., Hayashi, Y., Wabitsch, M., Mizusawa, N., Itakura, M., and Yoshimoto, K. (2009). YKL-40 secreted from adipose tissue inhibits degradation of type I collagen. Biochem. Biophys. Res. Commun. *388*, 511–516.

Jimenez, M.A., Akerblad, P., Sigvardsson, M., and Rosen, E.D. (2007). Critical role for Ebf1 and Ebf2 in the adipogenic transcriptional cascade. Mol. Cell. Biol. *27*, 743–757.

Justice, A.E., Winkler, T.W., Feitosa, M.F., Graff, M., Fisher, V.A., Young, K., Barata, L., Deng, X., Czajkowski, J., Hadley, D., et al. (2017). Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. Nat. Commun. *8*, 14977.

Kamei, Y., Xu, L., Heinzel, T., Torchia, J., Kurokawa, R., Gloss, B., Lin, S.C., Heyman, R.A., Rose, D.W., Glass, C.K., et al. (1996). A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. Cell *85*, 403–414.

Kwaks, T.H., Barnett, P., Hemrika, W., Siersma, T., Sewalt, R.G., Satijn, D.P., Brons, J.F., van Blokland, R., Kwakman, P., Kruckeberg, A.L., et al. (2003). Identification of anti-repressor elements that confer high and stable protein production in mammalian cells. Nat. Biotechnol. *21*, 553–558.

Kyrgios, I., Galli-Tsinopoulou, A., Stylianou, C., Papakonstantinou, E., Arvanitidou, M., and Haidich, A.B. (2012). Elevated circulating levels of the serum acute-phase protein YKL-40 (chitinase 3-like protein 1) are a marker of obesity and insulin resistance in prepubertal children. Metabolism *61*, 562–568.

Levine, M., and Manley, J.L. (1989). Transcriptional repression of eukaryotic promoters. Cell *59*, 405–408.

Ling, H., and Recklies, A.D. (2004). The chitinase 3-like protein human cartilage glycoprotein 39 inhibits cellular responses to the inflammatory cytokines interleukin-1 and tumour necrosis factor-alpha. Biochem. J. *380*, 651–659.

Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., et al. (2015). Genetic studies of body mass index yield new insights for obesity biology. Nature *518*, 197–206.

Londono Gentile, T., Lu, C., Lodato, P.M., Tse, S., Olejniczak, S.H., Witze, E.S., Thompson, C.B., and Wellen, K.E. (2013). DNMT1 is regulated by ATP-citrate lyase and maintains methylation patterns during adipocyte differentiation. Mol. Cell. Biol. *33*, 3864–3878.

Long, X.R., He, Y., Huang, C., and Li, J. (2014). MicroRNA-148a is silenced by hypermethylation and interacts with DNA methyltransferase 1 in hepatocellular carcinogenesis. Int. J. Oncol. *44*, 1915–1922.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., Abraham, J., Adair, T., Aggarwal, R., Ahn, S.Y., et al. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet *380*, 2095–2128.

Lu, Y., Day, F.R., Gustafsson, S., Buchkovich, M.L., Na, J., Bataille, V., Cousminer, D.L., Dastani, Z., Drong, A.W., Esko, T., et al. (2016). New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. Nat. Commun. *7*, 10495.

MacDonald, H.R., and Wevrick, R. (1997). The necdin gene is deleted in Prader-Willi syndrome and is imprinted in human and mouse. Hum. Mol. Genet. *6*, 1873–1878.

Majithia, A.R., Flannick, J., Shahinian, P., Guo, M., Bray, M.A., Fontanillas, P., Gabriel, S.B., GoT2D Consortium, NHGRI JHS/FHS Allelic Spectrum Project, SIGMA T2D Consortium, et al.. (2014). Rare variants in PPARG with decreased activity in adipocyte differentiation are associated with increased risk of type 2 diabetes. Proc. Natl. Acad. Sci. U S A *111*, 13127–13132.

Manna, P.R., and Stocco, D.M. (2007). Crosstalk of CREB and Fos/Jun on a single cis-element: transcriptional repression of the steroidogenic acute regulatory protein gene. J. Mol. Endocrinol. *39*, 261–277.

Mariman, E.C., and Wang, P. (2010). Adipocyte extracellular matrix composition, dynamics and role in obesity. Cell. Mol. Life Sci. *67*, 1277–1292.

Massari, M.E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. Mol. Cell. Biol. *20*, 429–440.

Messeguer, X., Escudero, R., Farre, D., Nunez, O., Martinez, J., and Alba, M.M. (2002). PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics *18*, 333–334.

Monda, K.L., Chen, G.K., Taylor, K.C., Palmer, C., Edwards, T.L., Lange, L.A., Ng, M.C., Adeyemo, A.A., Allison, M.A., Bielak, L.F., et al. (2013). A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. Nat. Genet. *45*, 690–696.

Nueda, M.L., Baladron, V., Garcia-Ramirez, J.J., Sanchez-Solana, B., Ruvira, M.D., Rivero, S., Ballesteros, M.A., Monsalve, E.M., Diaz-Guerra, M.J., Ruiz-Hidalgo, M.J., et al. (2007). The novel gene EGFL9/Dlk2, highly homologous to Dlk1, functions as a modulator of adipogenesis. J. Mol. Biol. *367*, 1270–1280.

Nyholt, D.R., Low, S.K., Anderson, C.A., Painter, J.N., Uno, S., Morris, A.P., MacGregor, S., Gordon, S.D., Henders, A.K., Martin, N.G., et al. (2012). Genome-wide association meta-analysis identifies new endometriosis risk loci. Nat. Genet. *44*, 1355–1359.

Ober, C., Tan, Z., Sun, Y., Possick, J.D., Pan, L., Nicolae, R., Radford, S., Parry, R.R., Heinzmann, A., Deichmann, K.A., et al. (2008). Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. N. Engl. J. Med. *358*, 1682–1691.

Pajukanta, P., Lilja, H.E., Sinsheimer, J.S., Cantor, R.M., Lusis, A.J., Gentile, M., Duan, X.J., Soro-Paavonen, A., Naukkarinen, J., Saarela, J., et al. (2004). Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1). Nat. Genet. *36*, 371–376.

Pan, W., Zhu, S., Yuan, M., Cui, H., Wang, L., Luo, X., Li, J., Zhou, H., Tang, Y., and Shen, N. (2010). MicroRNA-21 and microRNA-148a contribute to DNA hypomethylation in lupus CD4+ T cells by directly and indirectly targeting DNA methyltransferase 1. J. Immunol. *184*, 6773–6781.

Pattaro, C., Teumer, A., Gorski, M., Chu, A.Y., Li, M., Mijatovic, V., Garnaas, M., Tin, A., Sorice, R., Li, Y., et al. (2016). Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. Nat. Commun. *7*, 10023.

Pulit, S.L., Stoneman, C., Morris, A.P., Wood, A.R., Glastonbury, C.A., Tyrrell, J., Yengo, L., Ferreira, T., Marouli, E., Ji, Y., et al. (2019). Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum. Mol. Genet. *28*, 166–174.

Putt, W., Palmen, J., Nicaud, V., Tregouet, D.A., Tahri-Daizadeh, N., Flavell, D.M., Humphries, S.E., Talmud, P.J., and group, E. (2004). Variation in USF1 shows haplotype effects, gene : gene and gene : environment associations with glucose and lipid parameters in the European Atherosclerosis Research Study II. Hum. Mol. Genet. *13*, 1587–1597.

Ramirez-Zacarias, J.L., Castro-Munozledo, F., and Kuri-Harcuch, W. (1992). Quantitation of adipose conversion and triglycerides by staining intracytoplasmic lipids with Oil red O. Histochemistry *97*, 493–497.

Ross, S.E., Hemati, N., Longo, K.A., Bennett, C.N., Lucas, P.C., Erickson, R.L., and MacDougald, O.A. (2000). Inhibition of adipogenesis by Wnt signaling. Science *289*, 950–953.

Sapkota, Y., Steinthorsdottir, V., Morris, A.P., Fassbender, A., Rahmioglu, N., De Vivo, I., Buring, J.E., Zhang, F., Edwards, T.L., Jones, S., et al. (2017). Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism. Nat. Commun. *8*, 15539.

Shi, C., Zhang, M., Tong, M., Yang, L., Pang, L., Chen, L., Xu, G., Chi, X., Hong, Q., Ni, Y., et al. (2015). miR-148a is associated with obesity and modulates adipocyte differentiation of mesenchymal stem cells through WNT signaling. Sci. Rep. *5*, 9930.

Shungin, D., Winkler, T.W., Croteau-Chonka, D.C., Ferreira, T., Locke, A.E., Magi, R., Strawbridge, R.J., Pers, T.H., Fischer, K., Justice, A.E., et al. (2015). New genetic loci link adipose and insulin biology to body fat distribution. Nature *518*, 187–196.

Speliotes, E.K., Willer, C.J., Berndt, S.I., Monda, K.L., Thorleifsson, G., Jackson, A.U., Allen, H.L., Lindgren, C.M., Luan, J., Magi, R., et al. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat. Genet. *42*, 937–948.

Steenman, M.J., Rainier, S., Dobry, C.J., Grundy, P., Horon, I.L., and Feinberg, A.P. (1994). Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumour. Nat. Genet. *7*, 433–439.

Stefani, G., and Slack, F.J. (2008). Small non-coding RNAs in animal development. Nat. Rev. Mol. Cell Biol. *9*, 219–230.

Step, S.E., Lim, H.W., Marinis, J.M., Prokesch, A., Steger, D.J., You, S.H., Won, K.J., and Lazar, M.A. (2014). Anti-diabetic rosiglitazone remodels the adipocyte transcriptome by redistributing transcription to PPARgamma-driven enhancers. Genes Dev. *28*, 1018–1028.

Turcot, V., Lu, Y., Highland, H.M., Schurmann, C., Justice, A.E., Fine, R.S., Bradfield, J.P., Esko, T., Giri, A., Graff, M., et al. (2018). Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. Nat. Genet. *50*, 26–41.

Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., and Yang, J. (2017). 10 Years of GWAS discovery: biology, function, and translation. Am. J. Hum. Genet. *101*, 5–22.

Warnke, I., Goralczyk, R., Fuhrer, E., and Schwager, J. (2011). Dietary constituents reduce lipid accumulation in murine C3H10 T1/2 adipocytes: a novel fluorescent method to quantify fat droplets. Nutr. Metab. (Lond.) *8*, 30.

Winkler, T.W., Justice, A.E., Graff, M., Barata, L., Feitosa, M.F., Chu, S., Czajkowski, J., Esko, T., Fall, T., Kilpelainen, T.O., et al. (2015). The influence of age and sex on genetic associations with adult body size and shape: a large-scale genome-wide interaction study. PLoS Genet. *11*, e1005378.

Winkler, T.W., Justice, A.E., Graff, M., Barata, L., Feitosa, M.F., Chu, S., Czajkowski, J., Esko, T., Fall, T., Kilpelainen, T.O., et al. (2016). Correction: the influence of age and sex on genetic associations with adult body size and shape: a large-scale genome-wide interaction study. PLoS Genet. *12*, e1006166.

World Health Organization. Fact sheet on obesity and overweight. http://www.who.int/mediacentre/factsheets/fs311/en/.

Xie, H., Lim, B., and Lodish, H.F. (2009). MicroRNAs induced during adipogenesis that accelerate fat cell development are downregulated in obesity. Diabetes *58*, 1050–1057.

Xu, Q., Jiang, Y., Yin, Y., Li, Q., He, J., Jing, Y., Qi, Y.T., Xu, Q., Li, W., Lu, B., et al. (2013). A regulatory circuit of miR-148a/152 and DNMT1 in modulating cell transformation and tumor angiogenesis through IGF-IR and IRS1. J. Mol. Cell Biol. *5*, 3–13.

Yang, Z., Bian, C., Zhou, H., Huang, S., Wang, S., Liao, L., and Zhao, R.C. (2011). MicroRNA hsa-miR-138 inhibits adipogenic differentiation of human adipose tissue-derived mesenchymal stem cells through adenovirus EID-1. Stem Cells Dev. *20*, 259–267.

Yang-Yen, H.F., Chambard, J.C., Sun, Y.L., Smeal, T., Schmidt, T.J., Drouin, J., and Karin, M. (1990). Transcriptional interference between c-Jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein-protein interaction. Cell *62*, 1205–1215.

Zhang, Z., and Teng, C.T. (2001). Estrogen receptor alpha and estrogen receptor-related receptor alpha1 compete for binding and coactivator. Mol. Cell. Endocrinol. *172*, 223–233.

Zhao, X., Tang, R., Gao, B., Shi, Y., Zhou, J., Guo, S., Zhang, J., Wang, Y., Tang, W., Meng, J., et al. (2007). Functional variants in the promoter region of Chitinase 3-like 1 (CHI3L1) and susceptibility to schizophrenia. Am. J. Hum. Genet. *80*, 12–18.

ISCI, Volume *20*

Supplemental Information

Detailed Functional Characterization of a Waist-Hip Ratio Locus in 7p15.2 Defines an Enhancer Controlling Adipocyte Differentiation

Casimiro Castillejo-Lopez, Milos Pjanic, Anna Chiara Pirona, Susanne Hetty, Martin Wabitsch, Claes Wadelius, Thomas Quertermous, Erik Arner, and Erik Ingelsson

Supplementary Information

Detailed functional characterization of a waist-hip ratio locus in 7p15.2 defines an enhancer controlling adipocyte differentiation

Casimiro Castillejo-Lopez, Milos Pjanic, Anna Pirona, Susanne Hetty, Martin Wabitsch, Claes Wadelius, Thomas Quertermous, Erik Arner, Erik Ingelsson

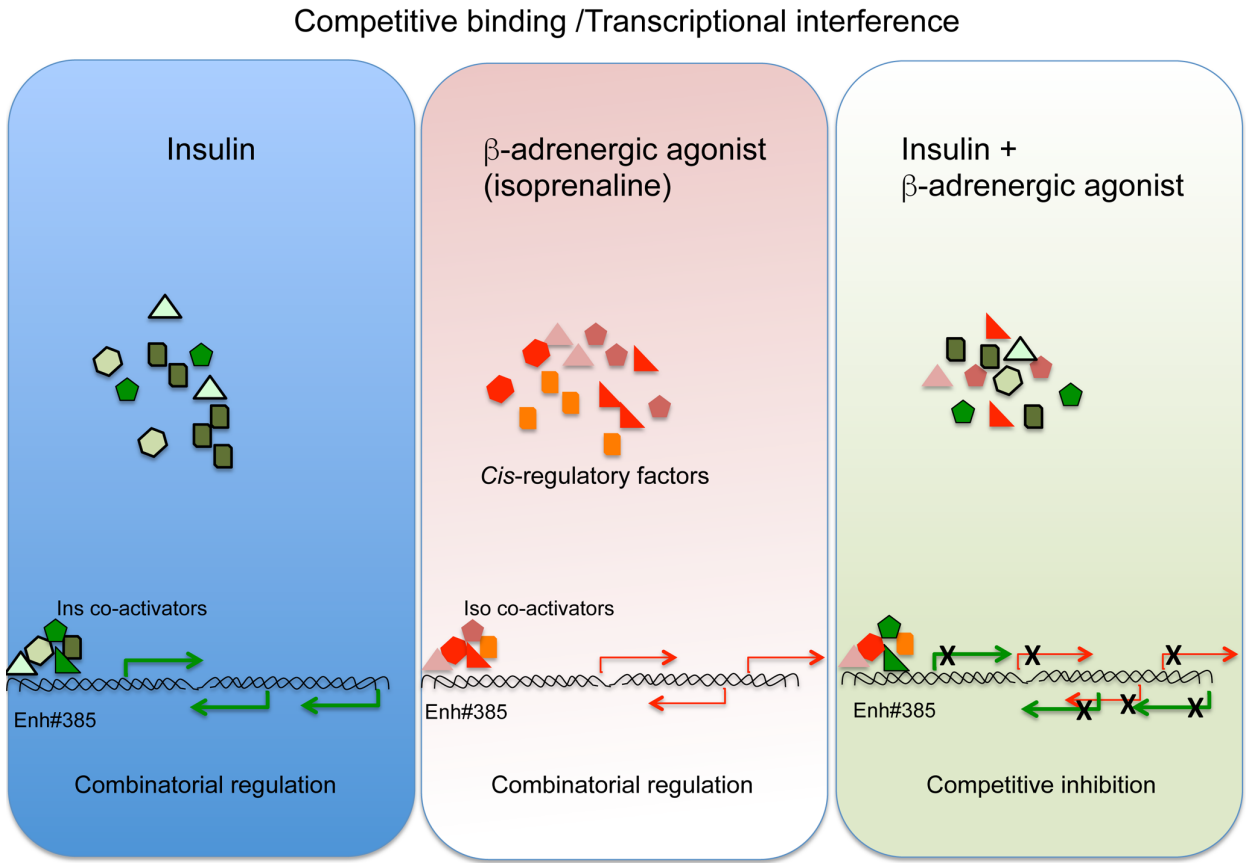


Figure S1. Potential model of inhibition by competitive binding. Related to Figures 2 and 3. Both the stimulation with insulin and isoprenaline (in blue and pink, respectively) potentiated the transcriptional activity of the enhancer. A potential model explaining our findings is that each stimulation generates a specific set of co-activators, and that simultaneous stimulation (in green) results in competitive inhibition of co-activators that annulate the potentiation of the transcription.

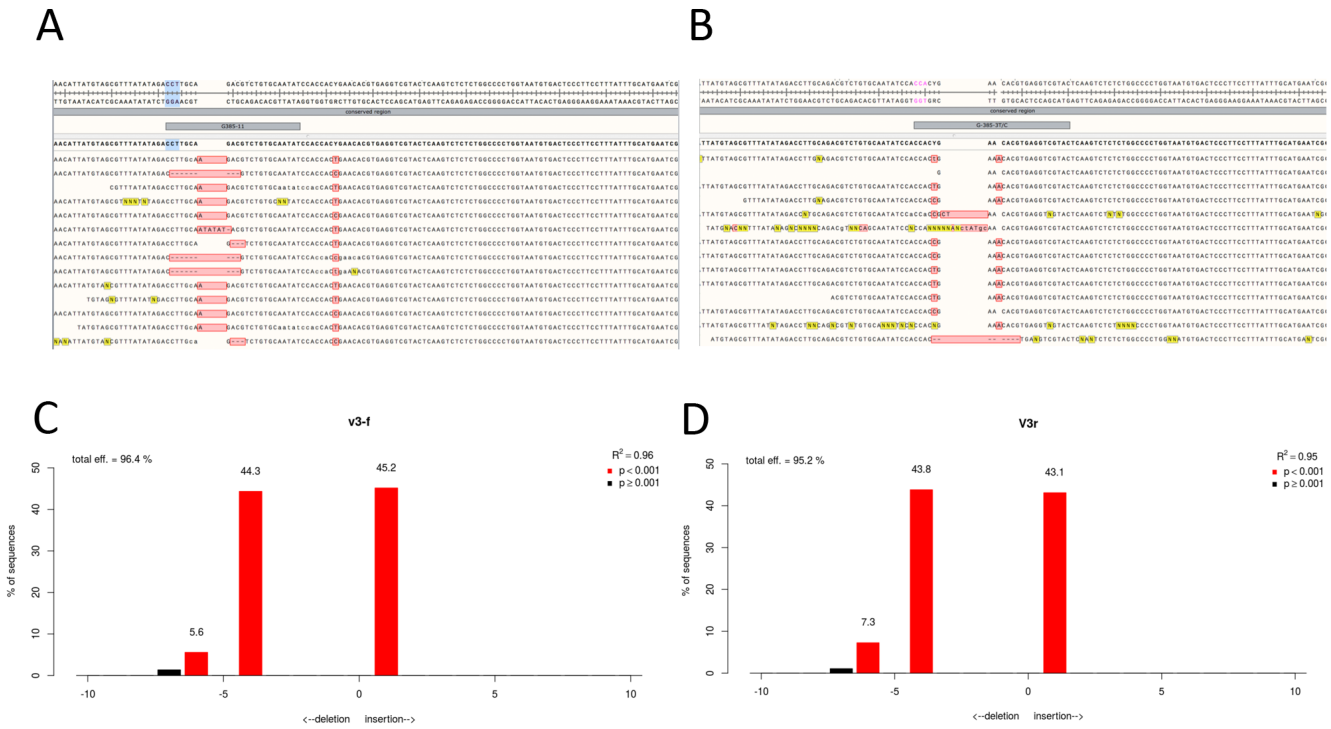


Figure S2. Estimation of editing efficiency by allelic isolation and TIDE. Related to Figures 4 and 5, and Table 1.

(A-B) Sequences alignments representing isolated alleles of SGBS cell cultures treated with CRISPR-Cas9 editing vectors; and (C-D) graphs showing the estimation of In-Dels in mutated cultures by Tracking of Indels by Decomposition (TIDE). (A) Sequences from KO-385-V11 cultures targeted upstream the rs1451385 polymorphism (T/C). Each row corresponds to an isolated allele. The guide RNA is delineated at the top in the wildtype sequence and the PAM sequence is indicated in blue. The cutting site of the Cas9 nuclease is three nucleotides upstream of the PAM (note that this construct is targeting the lagging strand) as predicted. The efficiency of mutation in this culture was estimated as 100% with the most abundant allele carrying an A insertion. (B) Isolated alleles from KO-385-V3 cultures targeted downstream of the polymorphism. The PAM sequence is indicated in pink letters in the reference sequence. The cutting site of the Cas9 nuclease is precisely three nucleotides upstream of the PAM sequence. Only one of the recovered alleles was wild type. The alignments were done using the SnapGene software. (C) TIDE analysis of the KO-385-V3 culture sequenced with the forward genomic primer; and (D) sequenced with the reverse genomic primer (**Supplementary Table 5**). The x-axis indicates the type of mutation (insertion or deletion), and the number of nucleotides inserted or deleted. The efficiency of both mutations (KO-385-V3 and KO-385-V11[not shown]) reached more that 95% according to both evaluations. The most abundant alleles were a four-nucleotide deletion and a one-nucleotide insertion, respectively.

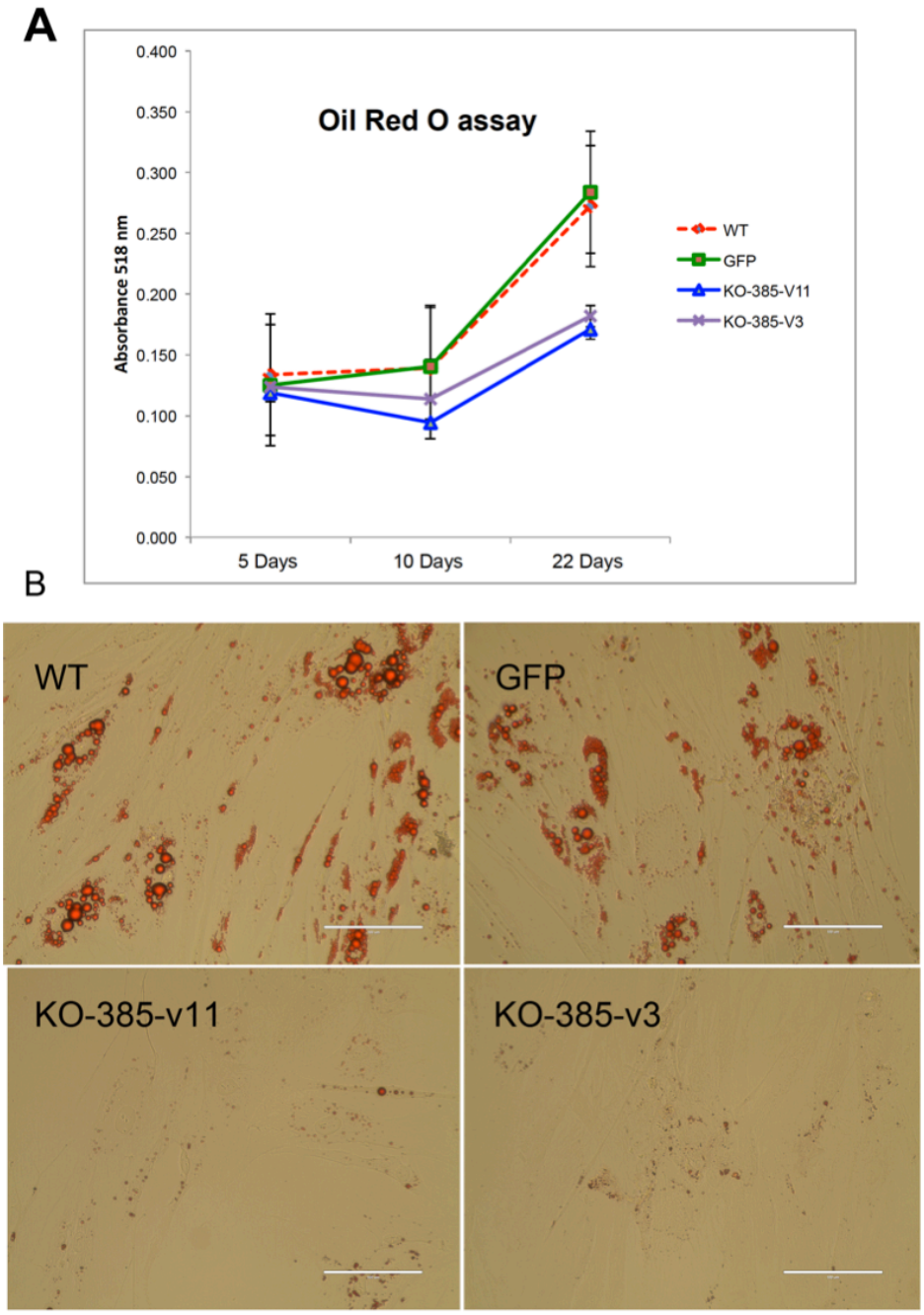


Figure S3. Quantification of lipids by Oil Red O staining. Related to Figure 4. Oil Red O, a lipid-specific dye, was used to determine lipid accumulation in differentiating cells cultures.(Ramirez-Zacarias et al., 1992) Cells were fixed for 1 hour with 10% formalin solution, washed with water, and stained for 2 hours with Oil Red O, followed by exhaustive rinsing with water. The dye was extracted with isopropyl alcohol and its absorbance was read at 510 nm. (A) Absorbance quantification at 5, 10 and 22 days after induced differentiation. Each point represents the mean values of three independent cultures. (B) Representative photographs of wild type, GFP transduced and mutant cultures at day 12. Scale bar, 100 micrometers.

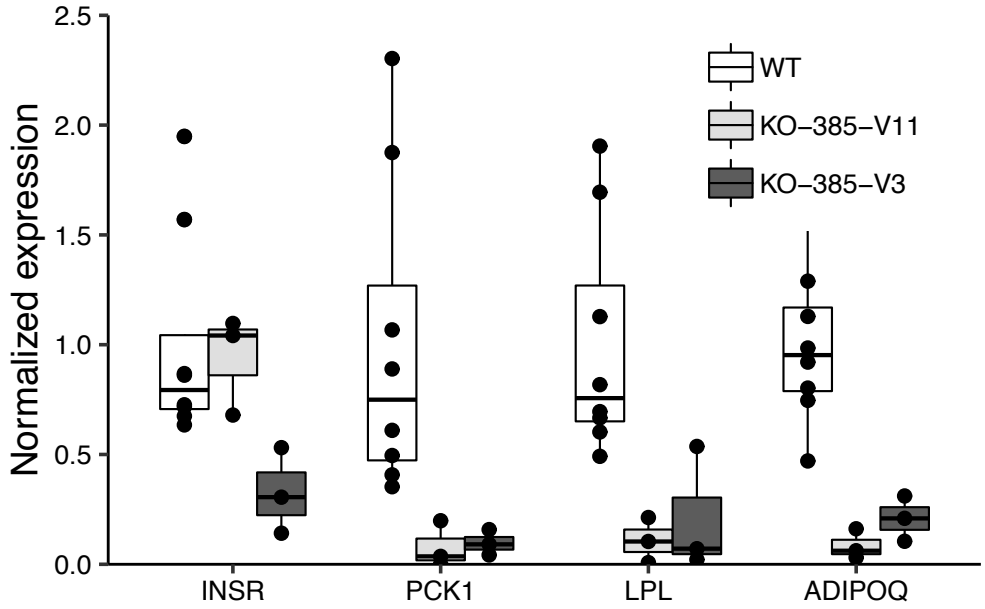


Figure S4. Relative RNA expression by RT-qPCR in differentiated SGBS cells. Related to Figure 4C and Table 1. The expression of the selected genes was done comparing three mutated cultures (K0-385-v11) propagated and differentiated in parallel with four WT cultures and three mutated (KO-385-v3) against four independent WT cultures. Each dot represents the mean of three technical replicates of the same culture. The normalization was done with the housekeeping gene TBP.

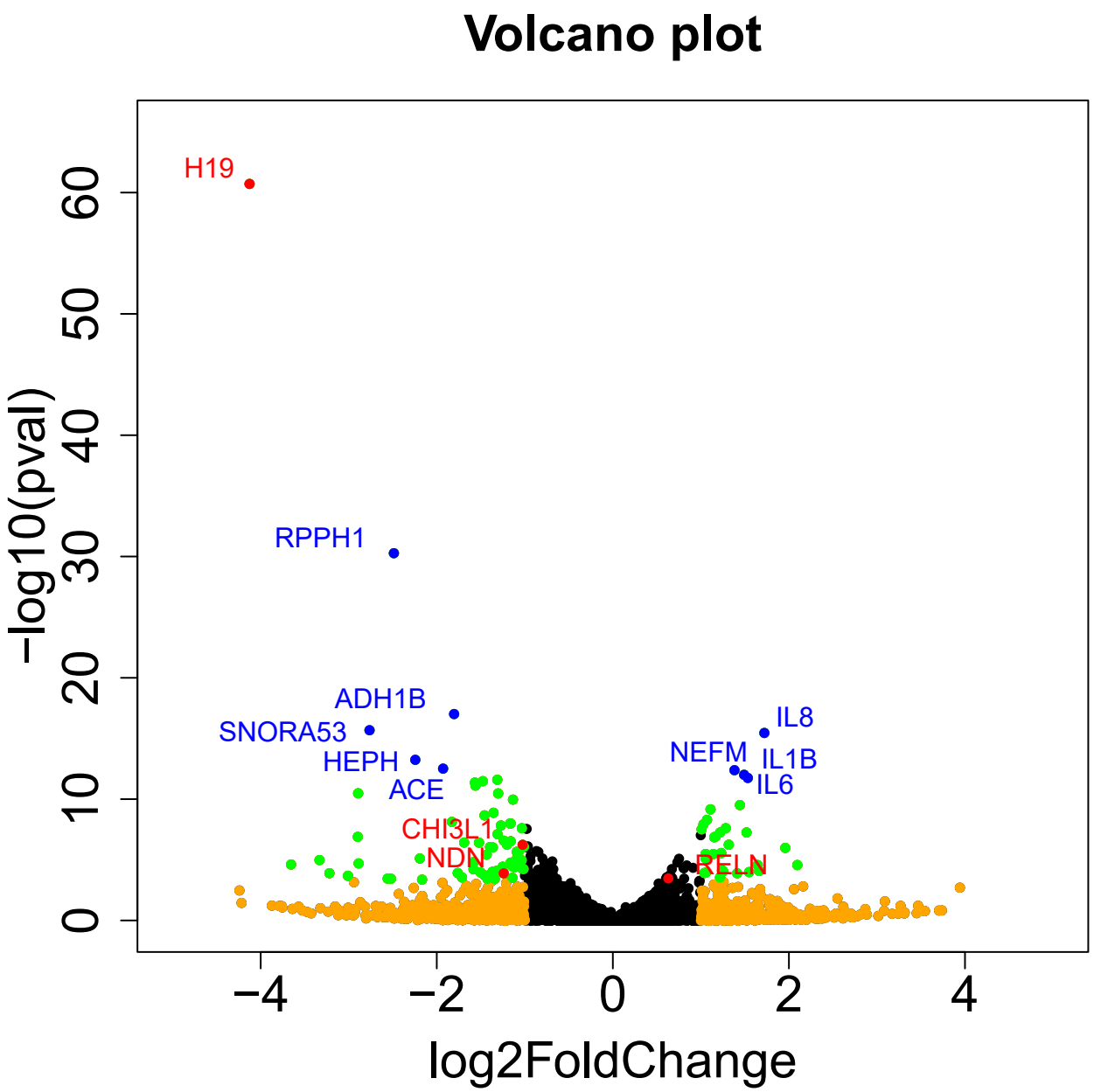


Figure S5. Differential gene expression in preadipocytes. Related to Figure 5B. Volcano plot comparing gene expression in preadipocytes between Enh#385 mutated (KO-385-V3 and KO-385-V11) and control (WT and GFP-transduced) cell cultures. Gene with significant differential expression (DE; P_{adj}<0.05 and abs(log_2 fold change)>1) are in green, while those with abs(log_2 fold change)>1 are in orange. The top 10 differentially expressed genes are in blue, and four highlighted genes (*CHI3L1*, *RELN*, *NDN* and *H19*) are in red.

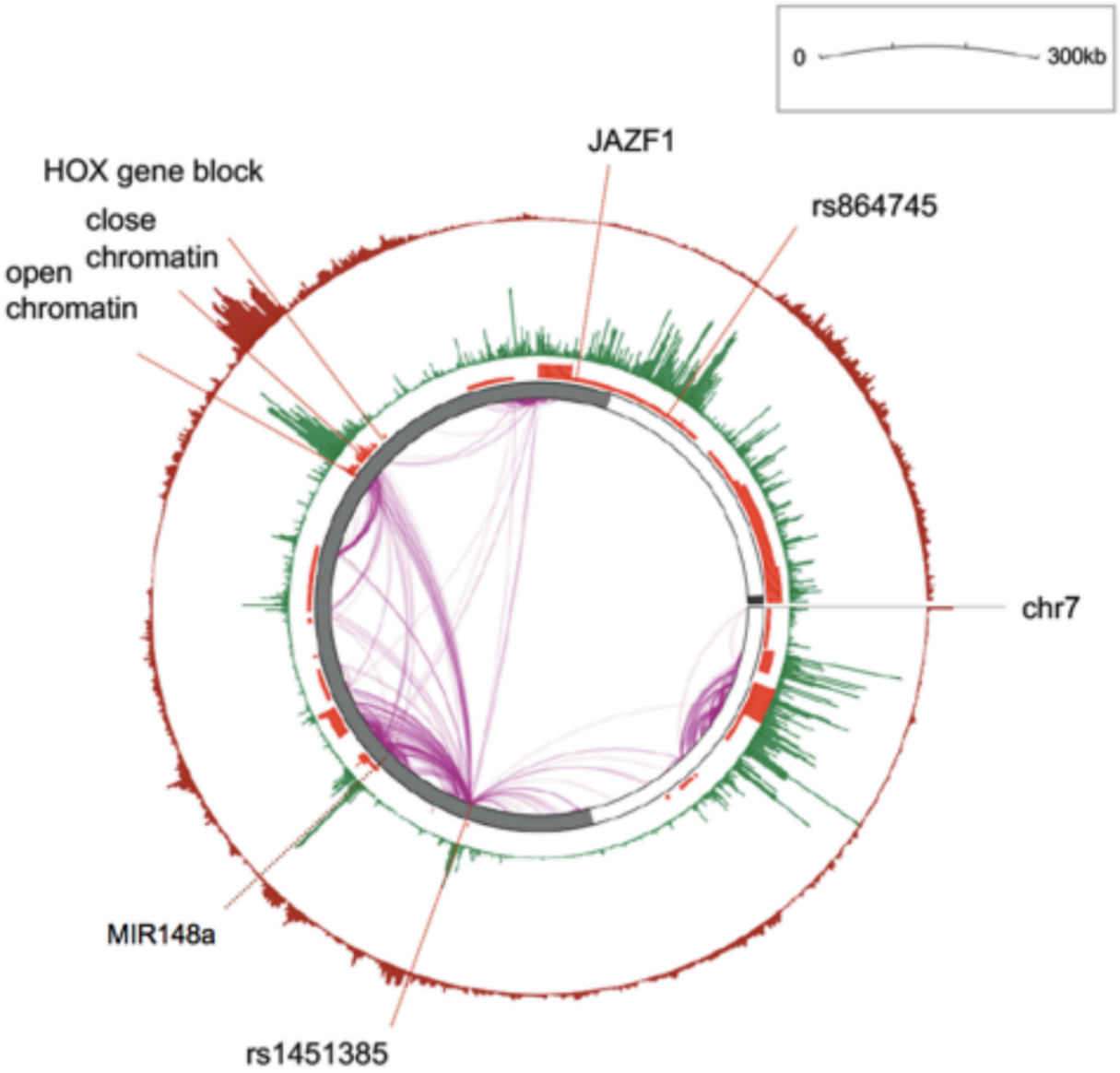


Figure S6. Circos plot of the chromatin interactome in the cis region around Enh#385. Related to Figure 7. The plot was generated using the Bioconductor function *FitHiC* with HiChIP data obtained from human coronary artery smooth muscle cells. The positions of the Enh#385 defined by the SNP rs1451385 and the *MIR148A* are marked. Chromatin looping tracks are in purple, and open chromatin are shown as bars. Raw and processed data are available at the Gene Expression Omnibus (GEO) under accession GSE101498. The data can be visualized in the WashU Epigenome Browser (https://epigenomegateway.wustl.edu/) with the following session ID: 98051079.

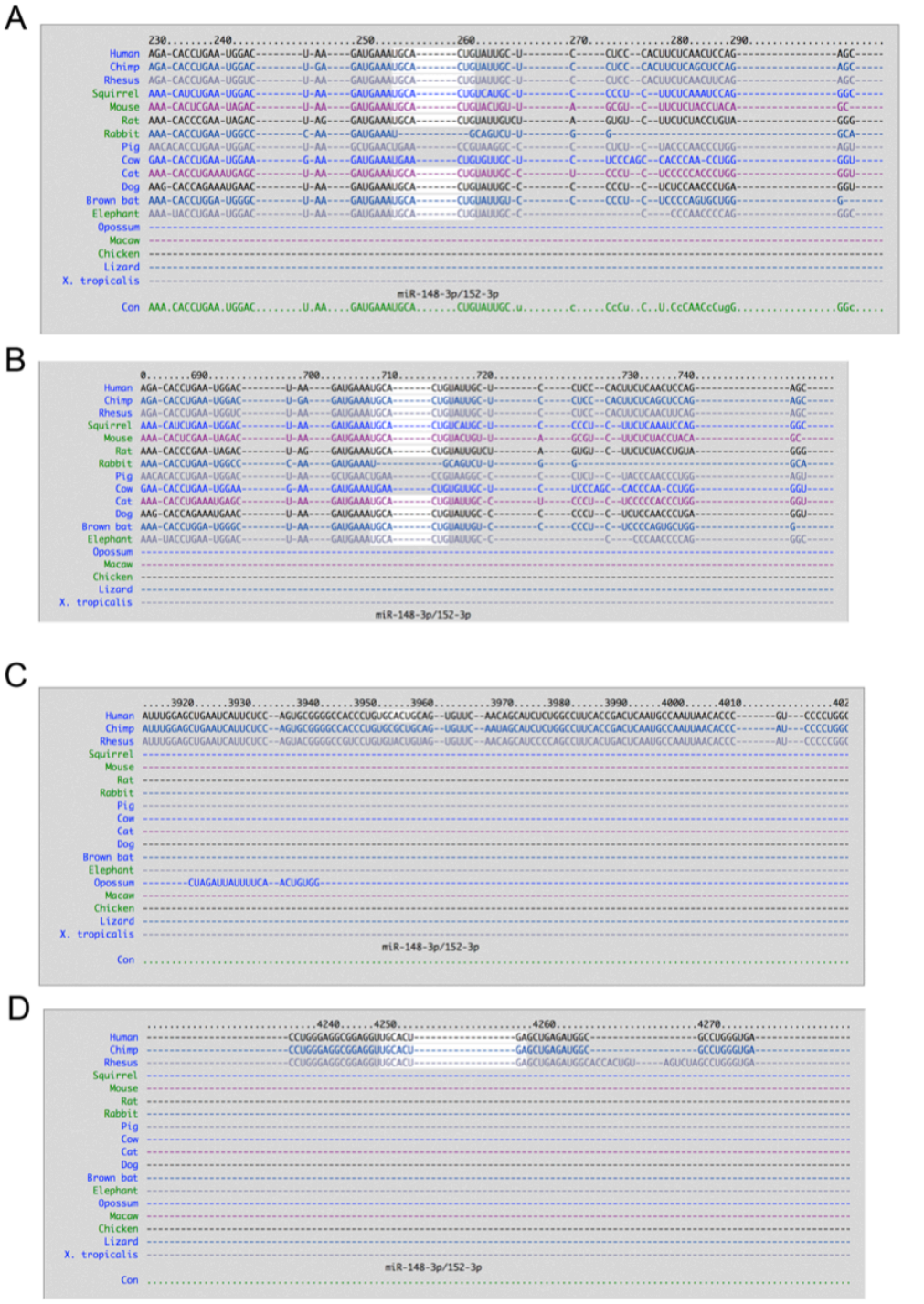


Figure S7. Predicted targets of miR148a that may be involved in initiation of adipogenesis. Related to Figure 7. Conserved sequences at the 3′ UTR regions of *WNT10B* (A and B) and *KLF2* (C and D) that were predicted as targets of *MIR148A* using Targetscan (http://www.targetscan.org). The predicted targets were found by searching for the presence of conserved 8mer, 7mer, and 6mer sites at the 3′UTR of the genes matching the seed region of *MIR148A*. Matches to the human *WTNT10B* and *KLF2* 3′ UTR are shadowed white. The targets in the 3′ UTR of *WNT10B* are broadly conserved among vertebrates, while the conservation is restricted to primates in the case of *KLF2*.

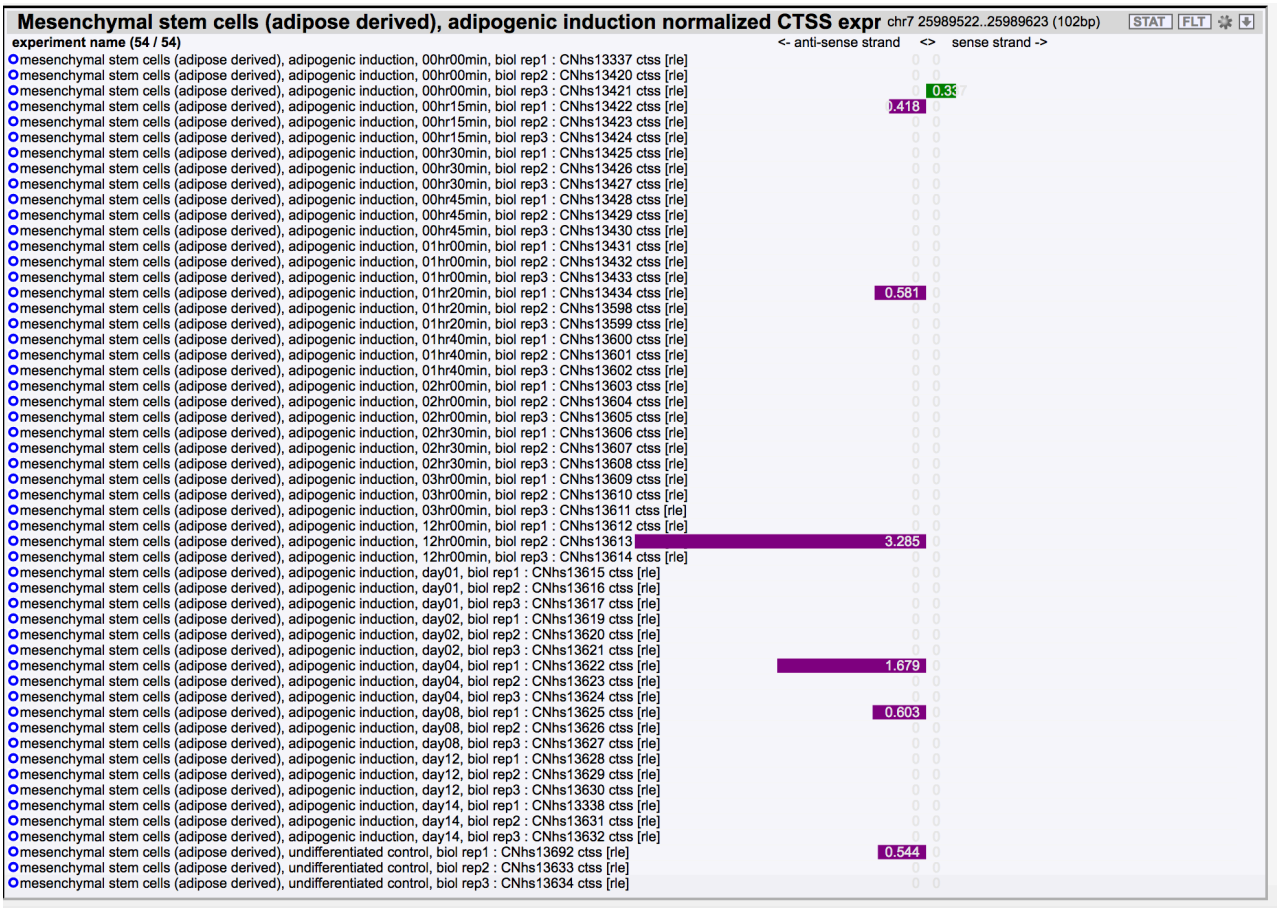


Figure S8. Activity of the *MIR148A* promoter during mesenchymal stem cell adipocyte differentiation. Related to Figure 7. Transcriptional signal of *MIR148A* during adipogenesis of mesenchymal stem cells visualized using the ZENBU browser (http://fantom.gsc.riken.jp/zenbu/). Signals are ordered from top to bottom from one hour after induction of adipogenesis to day 14.

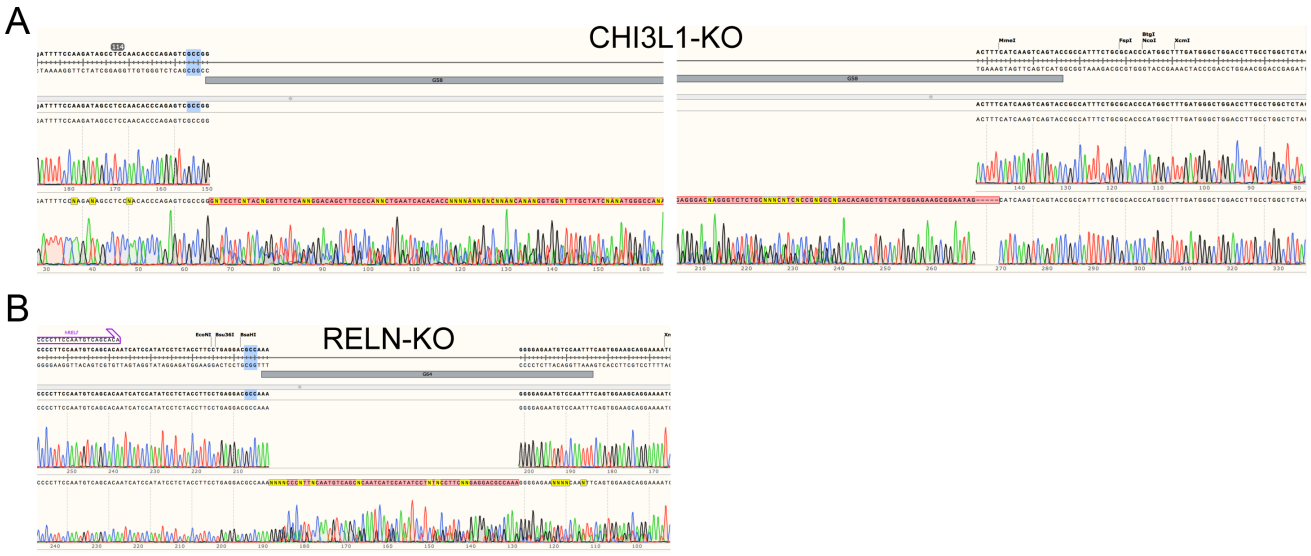


Figure S9. Sequencing chromatograms of wild type cell cultures and cultures mutated with CRISPR-Cas9. Related to Figure 6. (A) Sequencing chromatograms of the amplified genomic DNA of the target region of CHI3L1: Sequence of wild type cell culture (above) and of CRISPR-Cas9 treated cells (below). The more frequent mutated allele contains a 199 bp insertion at the predicted cutting site of the nuclease. The PAM sequence corresponding to the guide G58 is colored in blue. The central part of the chromatogram is not shown due to space restrictions. (B) Sequences of the RELN targeted region. The more frequent mutated allele contains a 60 bp insertion at the expected site of the PAM sequence corresponding to the guide G64.

Differentiated SGBS cells, DE of adipocytes linked genes						
Adipocytes markers genes						
		Relative expression				
Gene symbol	**NCBI_name**	**Wt**	**Mutant**	**Fold C**	**Class**	**Function**
LEP	leptin	3.6	0.9	0.3	AR	Adipokine key player in the regulation of energy balance and body weight
ADIPOQ	adiponectin__C1Q_and_collagen_domain_containing	45.0	15.7	0.3	AR	Adipokine involved in the control of fat metabolism and insulin sensitivity
ADIPOR2	adiponectin_receptor_2	34.1	27.7	0.8	AR	Adiponectin receptor
ADIPOR1	adiponectin_receptor_1	81.3	78.2	1.0	AR	Adiponectin receptor
SLC2A4	solute_carrier_family_2__facilitated_glucose_transporter__member_4	0.2	0.0	0.0	GM	Insulin-regulated glucose transporter (GLUT4)
PCK1	phosphoenolpyruvate_carboxykinase_1__soluble	36.5	9.4	0.3	GM	Regulates carbohydrate biosynthesis
PDHA1	pyruvate_dehydrogenase__lipoamide__alpha_1	69.6	59.3	0.9	GM	Link between glycolysis and the TCA cycle
SLC2A4RG	SLC2A4_regulator	199.4	212.8	1.1	GM	Regulator of the glucose trnasporter GLUT4
PDHB	pyruvate_dehydrogenase__lipoamide__beta	13.1	14.6	1.1	GM	Forms complex with PDHA1
PDK4	pyruvate_dehydrogenase_kinase__isozyme_4	110.3	152.4	1.4	GM	Key regulator of glucose and fatty acid metabolism
RASGRP1	RAS_guanyl_releasing_protein_1__calcium_and_DAG-regulated	5.5	4.4	0.8	IG	Immune activation
LPL	lipoprotein_lipase	60.6	22.9	0.4	LM	Triglyceride lipase activity
PLIN1	perilipin_1	30.4	13.3	0.4	LM	Adipocyte specific phosphoprotein that encircles the lipid droplet
CD36	CD36_molecule__thrombospondin_receptor	67.9	40.2	0.6	LM	Fatty acid multifunctional receptor
FABP4	fatty_acid_binding_protein_4__adipocyte	258.6	181.2	0.7	LM	Cytosolic lipid chaperone
SLC27A5	solute_carrier_family_27__fatty_acid_transporter__member_5	2.7	2.0	0.7	LM	FATP5
LPIN1	lipin_1	19.0	15.5	0.8	LM	Modulates lipid metabolism
DGAT2	diacylglycerol_O-acyltransferase_2	23.8	20.6	0.9	LM	DAG metabolism
SLC27A3	solute_carrier_family_27__fatty_acid_transporter__member_3	19.8	18.2	0.9	LM	FATP3
LPIN2	lipin_2	41.0	37.2	0.9	LM	Modulates lipid metabolism
FITM2	fat_storage-inducing_transmembrane_protein_2	23.9	22.4	0.9	LM	Involved in fat storage
FASN	fatty_acid_synthase	161.0	159.8	1.0	LM	Key enzyme for fat synthesis
SLC27A1	solute_carrier_family_27__fatty_acid_transporter__member_1	40.4	40.4	1.0	LM	FATP1
APOE	apolipoprotein_E	2887.4	2898.2	1.0	LM	Mediates the binding, internalization, and catabolism of lipoprotein particles
DIO2	deiodinase__iodothyronine__type_II	0.2	0.2	1.0	LM	Involved in the activation of the thyroid hormone T3
AGPAT2	1-acylglycerol-3-phosphate_O-acyltransferase_2	18.7	20.2	1.1	LM	Acts in the triacylglycerol synthesis
CPT1A	carnitine_palmitoyltransferase_1A__liver	36.0	43.8	1.2	LM	Involved in the transport to the fatty acid oxidation
DGAT1	diacylglycerol_O-acyltransferase_1	5.1	6.3	1.2	LM	DAG metabolism
PLIN2	perilipin_2	584.6	784.5	1.3	LM	Colocalisation with lipid droplets
Adipogenic genes						
DLK2	delta-like_2_homolog__Drosophila	0.6	0.3	0.5	ST	Promotes adipogenesis (CpG rich promoter)
STAT3	signal_transducer_and_activator_of_transcription_3__acute-phase_response_factor	110.6	107.1	1.0	ST	Involved in adipogenesis via CEBPB
SIRT1	sirtuin_1	12.8	14.0	1.1	ST	Promotes browning
SPARC	secreted_protein__acidic__cysteine-rich__osteonectin	6296.2	5484.5	0.9	ST	Functions as an inhibitor of adipogenesis
WNT10B	wingless-type_MMTV_integration_site_family__member_10B	0.2	0.1	0.5	ST	Promotes osteoblastogenesis, suppress PPARG expression
EBF2	early_B-cell_factor_2	24.0	26.7	1.1	TF	Promotes adipogenesis (CpG rich promoter)
PPARG	peroxisome_proliferator-activated_receptor_gamma	41.8	43.9	1.1	TF	Master TF required for adipogenesis
SREBF1	sterol_regulatory_element_binding_transcription_factor_1	11.0	9.6	0.9	TF	Induce adipocyte specific genes
EBF1	early_B-cell_factor_1	21.6	22.7	1.0	TF	Promotes adipogenesis
GATA2	GATA_binding_protein_2	13.9	13.9	1.0	TF	Suppress adipocyte differentiation
CEBPB	CCAAT_enhancer_binding_protein__C_EBP__beta	1.8	2.8	1.5	TF	Induces PPARG
CEBPD	CCAAT_enhancer_binding_protein__C_EBP__delta	35.8	46.2	1.3	TF	Induces CEBPA
RUNX2	runt-related_transcription_factor_2	10.0	7.2	0.7	TF	Inhibits adipogenesis
KLF2	Kruppel-like_factor_2	23.8	38.5	1.6	TF	Inhibit adipogenesis
ZIC1	Zic_family_member_1	0.2	0.4	1.8	TF	TF binds to APOE promoter
CEBPA	CCAAT_enhancer_binding_protein__C_EBP__alpha	6.0	4.9	0.8	TF	Induces PPARG and SREBF1
PPARGC1A	peroxisome_proliferator-activated_receptor_gamma__coactivator_1_alpha	0.4	0.1	0.2	TF	Brown fat fate regulator (PGC1a)
PRDM16	PR_domain_containing_16	0.6	0.3	0.5	TF	Brown fat fate regulator
H19	H19__imprinted_maternally_expressed_transcript__non-protein_coding	14530.4	12119.1	0.8	LR	Promotes adipogenesis (imprinted)

Table S2. Differential expression of genes known to be involved in adipocyte biology after Enh#385 KO in differentiated adipocytes. Related to Figure 7. Genes were selected from published literature and divided in two categories “Adipocyte marker genes” and “ Adipogenic genes” based on their function as a constituent part of the adipocyte; or as a gene necessary for the process of differentiation to mature adipocyte, respectively. The first category was subdivided in Adipokine Related genes (AR), genes involved in Glucose Metabolism (GM); Immune-related Genes (IG) and genes involved in Lipid Metabolism (LM). The second category was subdivided in genes involved in Signal Transduction (ST), Transcription Factors (TF) and Long-non-coding RNAs (LR). The relative expression represents the number of transcripts divided by the total number of counts; and provides an estimate of the abundance of the transcript. Fold change (Fold C) represents the difference in relative expression between the Enh#385 KO and WT cultures.

		Undifferentiated SGBS cells, DE adipocytes linked genes				
Adipocytes markers genes						
		Relative expression				
Gene symbol	**NCBI_name**	**Wt**	**Mutant**	**Fold C**	**Class**	**Function**
LEP	leptin	0.0	0.0	NA	AR	Adipokine key player in the regulation of energy balance and body weight
ADIPOQ	adiponectin__C1Q_and_collagen_domain_containing	0.0	0.0	NA	AR	Adipokine involved in the control of fat metabolism and insulin sensitivity
ADIPOR2	adiponectin_receptor_2	50.8	56.8	1.1	AR	Adiponectin receptor
ADIPOR1	adiponectin_receptor_1	126.0	132.8	1.1	AR	Adiponectin receptor
SLC2A4	solute_carrier_family_2__facilitated_glucose_transporter__member_4	0.0	0.0	NA	GM	Insulin-regulated glucose transporter (GLUT4)
PCK1	phosphoenolpyruvate_carboxykinase_1__soluble	0.0	0.0	NA	GM	Regulates carbohydrate biosynthesis
PDHA1	pyruvate_dehydrogenase__lipoamide__alpha_1	61.9	58.8	0.9	GM	Link between glycolysis and the TCA cycle
SLC2A4RG	SLC2A4_regulator	210.8	176.9	0.8	GM	Regulator of the glucose trnasporter GLUT4
PDHB	pyruvate_dehydrogenase__lipoamide__beta	10.3	9.8	0.9	GM	Forms complex with PDHA1
PDK4	pyruvate_dehydrogenase_kinase__isozyme_4	11.7	10.4	0.9	GM	Key regulator of glucose and fatty acid metabolism
RASGRP1	RAS_guanyl_releasing_protein_1__calcium_and_DAG-regulated	11.0	22.6	2.1	IG	Immune activation
LPL	lipoprotein_lipase	0.0	0.0	NA	LM	Triglyceride lipase activity
PLIN1	perilipin_1	0.3	0.3	0.8	LM	Adipocyte specific phosphoprotein that encircles the lipid droplet
CD36	CD36_molecule__thrombospondin_receptor	2.3	2.3	1.0	LM	Fatty acid multifunctional receptor
FABP4	fatty_acid_binding_protein_4__adipocyte	0.0	0.0	NA	LM	Cytosolic lipid chaperone
SLC27A5	solute_carrier_family_27__fatty_acid_transporter__member_5	1.4	0.9	0.6	LM	FATP5
LPIN1	lipin_1	28.5	23.3	0.8	LM	Modulates lipid metabolism
DGAT2	diacylglycerol_O-acyltransferase_2	10.9	16.3	1.5	LM	DAG metabolism
SLC27A3	solute_carrier_family_27__fatty_acid_transporter__member_3	15.5	9.4	0.6	LM	FATP3
LPIN2	lipin_2	42.8	45.8	1.1	LM	Modulates lipid metabolism
FITM2	fat_storage-inducing_transmembrane_protein_2	20.6	20.7	1.0	LM	Involved in fat storage
FASN	fatty_acid_synthase	258.1	224.6	0.9	LM	Key enzyme for fat synthesis
SLC27A1	solute_carrier_family_27__fatty_acid_transporter__member_1	25.0	20.9	0.8	LM	FATP1
APOE	apolipoprotein_E	72.3	47.4	0.7	LM	Mediates the binding, internalization, and catabolism of lipoprotein particles
DIO2	deiodinase__iodothyronine__type_II	0.0	0.7	14.8	LM	Involved in the activation of the thyroid hormone T3
AGPAT2	1-acylglycerol-3-phosphate_O-acyltransferase_2	18.2	15.5	0.9	LM	Acts in the triacylglycerol synthesis
CPT1A	carnitine_palmitoyltransferase_1A__liver	52.3	51.3	1.0	LM	Involved in the transport to the fatty acid oxidation
DGAT1	diacylglycerol_O-acyltransferase_1	4.8	5.1	1.1	LM	DAG metabolism
PLIN2	perilipin_2	440.7	542.6	1.2	LM	Colocalisation with lipid droplets
Adipogenic genes						
DLK2	delta-like_2_homolog__Drosophila	0.2	0.1	0.3	ST	Promotes adipogenesis (CpG rich promoter)
STAT3	signal_transducer_and_activator_of_transcription_3__acute-phase_response_factor	132.0	113.7	0.9	ST	Involved in adipogenesis via CEBPB
SIRT1	sirtuin_1	12.3	11.9	1.0	ST	Promotes browning
SPARC	secreted_protein__acidic__cysteine-rich__osteonectin	4964.1	6632.1	1.3	ST	Functions as an inhibitor of adipogenesis
WNT10B	wingless-type_MMTV_integration_site_family__member_10B	0.2	0.5	2.5	ST	Promotes osteoblastogenesis, suppress PPARG expression
EBF2	early_B-cell_factor_2	14.4	9.5	0.7	TF	Promotes adipogenesis (CpG rich promoter)
PPARG	peroxisome_proliferator-activated_receptor_gamma	51.2	34.5	0.7	TF	Master TF required for adipogenesis
SREBF1	sterol_regulatory_element_binding_transcription_factor_1	16.4	13.5	0.8	TF	Induce adipocyte specific genes
EBF1	early_B-cell_factor_1	12.7	10.7	0.8	TF	Promotes adipogenesis
GATA2	GATA_binding_protein_2	28.1	27.4	1.0	TF	Suppress adipocyte differentiation
CEBPB	CCAAT_enhancer_binding_protein__C_EBP__beta	2.5	2.7	1.1	TF	Induces PPARG
CEBPD	CCAAT_enhancer_binding_protein__C_EBP__delta	46.3	56.4	1.2	TF	Induces CEBPA
RUNX2	runt-related_transcription_factor_2	23.3	29.4	1.3	TF	Inhibits adipogenesis
KLF2	Kruppel-like_factor_2	12.1	17.1	1.4	TF	Inhibit adipogenesis
ZIC1	Zic_family_member_1	0.8	1.6	1.9	TF	TF binds to APOE promoter
CEBPA	CCAAT_enhancer_binding_protein__C_EBP__alpha	0.0	0.0	NA	TF	Induces PPARG and SREBF1
PPARGC1A	peroxisome_proliferator-activated_receptor_gamma__coactivator_1_alpha	0.0	0.0	NA	TF	Brown fat fate regulator (PGC1a)
PRDM16	PR_domain_containing_16	0.0	0.0	NA	TF	Brown fat fate regulator
H19	H19__imprinted_maternally_expressed_transcript__non-protein_coding	42.5	2.4	0.1	LR	Promotes adipogenesis (imprinted)

Table S4. Differential expression of genes known to be involved in adipocyte biology after Enh#385 KO in differentiated adipocytes. Related to Figure 7. Genes were selected from published literature and divided in two categories “Adipocyte marker genes” and “ Adipogenic genes” based on their function as a constituent part of the adipocyte; or as a gene necessary for the process of differentiation to mature adipocyte, respectively. The first category was subdivided in Adipokine Related genes (AR), genes involved in Glucose Metabolism (GM); Immune-related Genes (IG) and genes involved in Lipid Metabolism (LM). The second category was subdivided in genes involved in Signal Transduction (ST), Transcription Factors (TF) and Long-non-coding RNAs (LR). The relative expression represents the number of transcripts divided by the total number of counts; and provides an estimate of the abundance of the transcript. Fold change (Fold C) represents the difference in relative expression between the Enh#385 KO and WT cultures.

Primers used for genome amplification of the CRISPR-Cas9 targets and TIDE sequencing		
Target	**Forward primer**	**Reverse primer**
Enh#385	TTGAGTTGAACATTATGTAGCGTTT	TGAAACAGTAAACATTACGAAAATTAG
RELN	CCCCTTCCAATGTCAGCACA	GGAGATGAAGGATTAAAACAACCCA
CHI3L1	CCCAAGGCACCTAGCATATCA	CTTCTAGCCCACCCCATTCC

Primers used for RT-qPCR validation		
Gene symbol	**Forward primer**	**Reverse primer**
GAS7	ACGTTCCCTCACCCAGACA	TTCAGCTGTTTCTGGAGCAGTAG
CHI3L1	CAACCTGAAGACTCTCTTGTCTGT	GCAAGGTCCAGCCCATCAAA
HEPH	TGATGGAGCCACTCGAGTCTA	CCCCCTATCCGGTTCTTGTC
NDN	CGTGCGGAAGCTCATCACT	CCAGGAACTCCATGATTTGCATCTT
LPL	GTGATTCATACTTTAGCTGGTCAGACT	CGCAGGTGCCTTTCCTTTCT
H2AFY2	AGAACAGCTTGAAGAGACCATCAAAA	TGGGCTGAGATGGCTTTGAG
MID2	CACGGTTTCTACAGTCTGCAAAA	ACCCCTCTAGCAGTTTCTTTTCTCT
RCAN2	GTGGTGGATGTCGAGGTCTT	TCGGGCTGCAGATTTAGGATTG
RELN	CCCAAAATGCTCTGGAAGTGTTT	ATTGCCCACTGCATGTCTGA
ITIH5	CCGAGGCAAGTCAGACTGT	GAACTCAATGTCCTGGTCTTCAGAA
SGCD	TGCAGAAGCTGGCAATATGGAA	AGACCTTCTGCCTCGTTCCT
MYH2	GCCGGACTCTAGAGGACCAA	ACACCAGAGCTTCCTTTTCATCAAG
ROR2	GATGGGTTCTGCCAGCCTTA	CGACAGGTGCGTAGACGTG
PCK1	GGAAGAATAAGGAGTGGAGCTCAGA	GGCCTCCAAAGATAATGCCTTCAA
MLXIPL	GGGTACCCATCACACACCAG	TGGACACCATCCCGTTGAAG
PLA2G2A	CTGCATTTGTCACCCAAGAACTC	ACTGAGTGCGGCTTCCTTT
PPAPDC1A	GCATGTGCGACTACAAGCAT	GCTGGGACTCGCAGACTAA
HYDIN	TGCTGGATATTGGGAAAGTTTTCACT	GATGCCTTCCTTGGGACTGAA
DYRK1B	ATCAAGACCTACAAGCACATCAATGA	GCGCACGATGTAGTCATGG
INSR	CCTGAAGAGCTACCTCCGTTCT	AGGTCCCGATGCACAAACTT
TBP	GCGCAAGGGTTTCTGGTTTG	CCAGTGCCATAAGGCATCATTG
ADIPOQ	CCTGGTGAGAAGGGTGAGAAAG	CATAGGCACCTTCTCCAGGTT

Table S5. Primers sequences for genome amplification of the CRISPR-Cas9 targets and for RT-qPCR. Related to Figures 4, 5 and 6, and Table 1.

Transparent Methods

Cell cultures and differentiation

The human preadipocyte SGBS cell line was cultured as previously described(Fischer-Posovszky et al., 2008) in OF-serum containing medium composed of DMEM/F12 (1:1) medium (Life Technologies) supplemented with 10% FCS, 17 µM biotin, 33 µM pantothenic acid and 1% penicillin/streptomycin/amphotericin B (Gibco). Cells were maintained at 37°C and 5% CO_2.

To promote adipose differentiation, cells were seeded into a 24-well plate at a density of $2.5x10^5$ cells per well and cultured in Quick-Diff medium consisting of serum-free medium supplemented with transferrin (0.01 mg/ml), insulin (20 nM), cortisol (100 nM), triiodothyronine (0.2 nM), dexamethasone (25 nM), isobutyl-1-methylxanthine (IBMX) (250 μM) and the PPARG-agonist rosiglitazone (2 μM) (BioNordika, Sweden). After 4 days, the medium was removed and replaced by 3FC medium composed of serum-free medium supplemented with transferrin (0.01 mg/ml), insulin (20 nM), cortisol (100 nM) and triiodothyronine (0.2 nM). The medium was changed every 4 days, and the differentiation process was documented by microscopy every 4 days.

The HEK293T (ATTC Cat# CRL-3216, RRID: CVCL_0063), HELA (CLS Cat# 300194/p772_HeLa, RRID: CVCL_0030) and HepG2 (CLS Cat# 300198/p2277_Hep-G2, RRID: CVCL_0027) cells were cultured in RPMI 1640 medium supplemented with 10% FBS, 2 mM L-glutamine and 1% penicillin/streptomycin/ amphotericin B (Gibco).

Evaluation of enhancer activity

Cloning of reporter constructs. A 266 bp genomic fragment containing the polymorphism rs1451385 was amplified with the primers (5´caggtaccGTTGAACATTATGTAGC 3´) and (5´ggagatcTTGAACATTATGTAGC 3´) and cloned into the KpnI and BglII sites of the pGL4.10 firefly luciferase reporter (Promega). The minimal promoter of pGL4.26 (Promega) was inserted upstream of the luciferase gene. The 266 bp fragment was cloned in inverse position by amplification with the primers (5´cgggtaCCTGAGGAGTAGTATGC 3´) and (5´tcagatctCCTGAGGAGTAGTATGC 3´) and inserted into the same KpnI and BglII sites. In order to test both alleles, the cloned allele (T) was mutated to (C) by conventional PCR mutagenesis using the primers (5´CAATATCCACCACtGAACACGTGAGGTCGTAC 3´) and (5´CGACCTCACGTGTTCaGTGGTGGATATTGCACAGA 3´). The vectors used in our typical reporter assay were the following: forward direction-allele T, forward direction-allele C, reverse

direction-allele T, reverse direction-allele C, empty vector containing only the minimal promoter, and a positive control vector. As a positive control, we chose a human 411-bp enhancer from chromosome 12 that consistently increases luciferase expression about three-fold in a variety of cells lines (Cavalli et al., 2016). The positive enhancer control was amplified with the primers 5´-GCTGTGGGGGAAGAGGTATTC-3´and 5´-GAGAGACTCCCCCACCTTCA-3´.

Quantification of enhancer activity. Enhancer activity was examined by luciferase reporter gene assays in the following human cells lines: HepG2 (hepatocytes), HeLa (cervical cancer cells), HEK293 (embryonic kidney) and SGBS (preadipocytes).
Co-transfections were done in 48-well plate at 80 % confluence using Lipofectamine 3000 (Life Technologies) with 150 ng of reporter construct and 10 ng Renilla luciferase reporter vector to normalize the transfection efficiency. Cells were lysed 24 h after transfection, and ratios between Firefly luciferase and Renilla luciferase activity were measured with a dual luciferase assay (Promega) using a Varioskan Lux Microplate Reader (Thermo). For reporter experiments where inducible reagents were used (insulin and/or isoprenaline), the measurements were done 48 hours after transfection, corresponding to 24 hours of inducible treatment.

Replicates. Each mean value is the result of measuring 12-14 biological replicates, using three different cell cultures and two different plasmid preparations.

Genome editing

We performed CRISPR-Cas9 genome editing in the human preadipocyte SGBS cell line to alter the sequence adjacent to rs1451385. The guide RNAs (gRNA) were designed using the online tools: http://www.broadinstitute.org/rnai/public/analysis-tools/sgrna-design and http://crispr.mit.edu/. The DNA oligos were purchased from IDT (Integrated DNA Technologies), annealed and cloned into the BsmBI site of the lentiCRISPRv2 lentiviral vector as previously described.(Shalem et al., 2014) Correct insertion was verified by Sanger sequencing. To rule out off-target effects, we induced mutations of the enhancer with two separate guide RNAs. We also searched for potential off-target sites in the human genome for both guides.(Bae et al., 2014) For guide G385-11, there were two possible off-targets with two mismatches; while for guide G385-3T/C, we found seven possible off-targets with two mismatches. No sites with precise match or with one mismatch were observed. We further searched the regions of these nine potential off-target sites (localized in intergenic regions) for known genes involved in adipocyte biology, but we did not find any adipocyte-related genes close to

these potential off-target sites. Using our differential expression data, we have also checked for alteration of gene expression in the proximity of these potential off-target sites, but we were unable to detect any such effects.

The gRNA sequences were: G385-11, atattgcacagacgtctgcaAGG for construct V11-385; G385-3T/C, gtacgacctcacgtgttcagTGG for construct V3-385; aaattggacattctcccttTGG for *RELN* and gtactgacttgatgaaagtcCGG for *CHI3L1* (PAM sequences in capitals). Lentiviruses expressing Cas9 and the gRNAs were generated in HEK293T cells by co-transfection of the packaging plasmids psPAX.2 and psMD2.G (Addgene). Supernatants containing lentiviruses were harvested 24 and 48 h post-transfection. Confluent SGBS cells were subsequently exposed to lentiviruses (with polybrene; 8 μg/ml) and selected with puromycin (1 μg/ml) for 72 h.

Assessment of editing efficacy

Mutation efficacy was estimated by two method: first, cloning and sequencing of isolated alleles; and second, by analysis of the sequence chromatograms of the amplified target region using Tracking of Indels by DEcomposition (TIDE). The first method gives information about the frequency and the type of mutations, while the second method gives information of indel mutations but neither the nature of the indel or possible substitutions are counted. We verified that that both methods correlated well in our CRISPR-Cas9 transductions (**Supplementary Figure 2**). We chose TIDE as the standard method for our mutation assessments of independent cell cultures.

For isolation of mutated alleles, cells resistant to puromycin were propagated and an aliquot corresponding to 1×10^5 cells was withdrawn for genomic DNA preparation. PCR amplifications of the target sequence were carried out with 20 ng genomic DNA in 50 μl reaction using Platinum Taq polymerase (Thermo Fisher) and the primers (5´ TTCCATGCGTTGGAAGTCCATT-3´and 5´ CGGCCTGAGGAGTAGTATGC-3´). PCR conditions were 3 min at 94°C (1x), followed by (2x) touchdown cycles of 30 s at 94°C, 20 s at 59°C, 58°C, 57°C and 30 s at 72°C and ending with (27x) of 30 s at 94°C, 20 s at 56°C and 30 s at 72°C, and a final extension at 72°C for 5 min. The PCR products were purified using the PCR DNA and Gel Band purification kit (GE Healthcare) and cloned using the *TOPO TA Cloning kit* in the pCR™ 2.1 vector (ThermoFisher). Individual bacteria clones were Sanger sequenced using universal primers present in the vector. Sequences were aligned to a wild type clone and the frequency of editing was estimated. Independent transductions were assessed for efficiency of genome editing using Tracking of Indels by DEcomposition (TIDE) as described previously(Brinkman et al., 2014) (**Supplementary Figure 2**). PCR products were purified using the SequalPrep™

Normalization Plate (Invitrogen). An aliquot of 30 ng DNA was Sanger sequenced in both directions using the amplification primers (**Supplementary Figure 2**). For primers sequences, see **Supplementary Table 5**. The chromatograms were analyzed using the TIDE tool, available at https://tide.nki.nl.

Image quantification of lipid accumulation

We observed some limitations using the classical Oil Red O staining method for quantification of lipid content in our SGBS cell cultures. During the rinsing step, small oil droplets detached from some cells and floated to the surface where they were discarded during subsequent washing steps. This loss was more pronounced when cultures were well-differentiated. To cope with this problem, we decided to measure the oil droplets by cell imaging and normalize by the number of cell nuclei. Lipid content was estimated using phase contrast images and/or images stained with Bodipy 493/503 (4, 4-difluro-1, 3, 5, 7, 8- pentamethyl-4-bora-3a, 4a-diaza-s-indacene) and captured with the Phase or GFP channel, respectively. To quantify the number of cells, the nuclei were stained with Hoechst 33342 (Thermo Fisher). Before imaging, the culture medium was replaced by PBS containing a solution of Bodipy®493/503 (2 μM) and Hoechst® 33342 (1.6 μM). The EVOS FL Auto-fluorescent microscope was used to image the cells on the phase contrast channel and with the EVOS® DAPI Light Cube to capture the nuclei and GFP Light Cube to capture stained lipids. Cell counting was performed on the DAPI channel. Cultures were differentiated in triplicates and at least 25% of each culture well was automatically recorded and the pictures analyzed at batch mode using ImageJ 1.50i (ImageJ, RRID: SCR_003070).

The following macros for ImageJ were used for quantification of lipids stained with Bodipy 493/503 and recorded with the GFP channel; or without staining and recorded using phase contrast. The number of cells was determined with Hoechst staining. The photographs were taken automatically using 10X magnification in an EVOS FL Auto2 fluorescence microscope.

Bodipy 10x:

```
run("8-bit");
run("Auto Local Threshold...", "method=Bernsen radius=15 parameter_1=0 parameter_2=0 white");
run("Measure");
```

Phase contrast:

```
run("8-bit");
run("Auto Local Threshold...", "method=Sauvola radius=15 parameter_1=0 parameter_2=0 white");
setOption("BlackBackground", false);
run("Dilate");
run("Analyze Particles...", "show=Ellipses");
run("Fill Holes");
run("Measure");
```

Hoechst staining:

```
run("Smooth");
setOption("BlackBackground", false);
run("8-bit");
run("Auto Local Threshold...", "method=Bernsen radius=15 parameter_1=0 parameter_2=0 white");
run("Analyze Particles...", "size=20-700 circularity=0.50-1.00 show=Outlines display clear
summarize");
```

RNA sequencing

Total RNA was isolated with RNeasy Mini kit (Qiagen). Three independent RNA samples were analyzed from each type of differentiated SGBS cells: Wild type cells (WT), cells transduced with lentivirus expressing a gRNA targeting an unrelated intergenic region and cells transduced with lentivirus expressing Cas9 and a sgRNA targeting the sequence downstream of the polymorphism rs1451385 (V3-385). To assess the effect of insulin, we treated all our fully differentiated cell cultures with or without 100 nM insulin for 24h. This resulted in a total of 18 RNA samples of differentiated cells being used in the first experiment. Additionally, to study differential gene expression in undifferentiated preadipocytes, we used duplicate samples of wild type cells, cells transduced with lentivirus expressing GFP (GFP), cells mutated downstream of rs1451385 (V3-385) and cells mutated upstream of rs1451385 (V11-385), i.e. 8 samples in total. The total RNAs were subjected to amplification using The Ion AmpliSeq Transcriptome Human Gene Expression Mini Kit (Thermo Fisher) and run in an IonProton instrument (Thermo Fisher) at the Uppsala Genome Center. The average number of reads per sample was > $1.2x10^7$. RNA-Seq reads were quality-controlled and analyzed using Torrent Suite Software (Thermo Fisher) with standard settings. Differential expression of genes between KO and control conditions was assessed using the DESeq2 R package from

Bioconductor (http://bioconductor.org/packages/release/bioc/html/DESeq2.html).

Validation using RT-qPCR

RNA was extracted using RNeasy and cDNA synthesized using the SuperScript IV Reverse Transcriptase (Fisher Scientific) kit according to the manufacturer's instructions. Real-time PCR was performed using the MX 3005 Strata quantitative PCR (Stratagene) according to the supplier's manual (Stratagene). PCR amplification was detected using the SYBR™ Green master mix (Fisher Scientific). Three biological replicates were used and converted to cDNA. Each cDNA sample was run in triplicate and raw data were normalized to the house keeping gene TBP4. Primers sequences are described in **Supplementary Table 5**.

We selected genes for confirmatory RT-qPCR analyses using the following criteria:

1) Differential expression in mature adipocytes: Genes whose expression was significantly altered after differentiation to adipocytes comparing three mutated cultures (M(ds)) against three wildtype (WT) cultures analyzed with RNA-seq (**Table 1**, column 4[adjP]) were further investigated. In order to increase statistical power and discard false positives due to transduction effects, we used the ranking of DE genes obtained from the comparison of *six* mutated cultures (M (ds)) with *twelve* control cultures (*six* wildtype cultures and *six* cultures with mutants cells from a CRISPR-Cas9 mutation of a regulatory element on chromosome 1 with very low effect on global gene expression). The significance of this comparison was estimated by Student's t-test with unpaired values and the Welch correction for unequal variance (**Table 1**, column 5 [p-val*]).
2) Discarding genes whose differential expression was potentially caused by the transduction procedure, i.e. genes differentially expressed only in the transduced cultures (e.g. *EMX2*, *EMX2OS*).
3) Discarding genes with an overall very low gene expression (e.g. *MXRA5*).
4) Adding a few wildcard genes from the top 50 DE genes from RNA-seq with high biological potential (e.g. *ITIH5*, *MLXIPL*, *PPAPDC1A*, *DYRK1B*, *INSR*).

The criteria for filtering out false positives and selecting candidate genes for downstream experiments based on the RT-qPCR validation were three-fold: (1) A technical validation using RNA from the same cells as those used for the global transcriptome analysis requiring significant DE in the same direction as in RNA-seq for successful validation (**Table 1**, columns 6-7 [M(ds)/WT]). (2) A biological

validation using RNA from an independent mutation (**Table 1**, columns 8-9 [M(up)/WT]). (3) A control of potential transduction effects in which the candidate genes should not show differential expression comparing non-transduced WT cells and cells transduced with a lentivirus expressing only GFP (**Table 1**, columns 10-11 [GFP/WT]). Finally, validated genes were further interrogated for possible involvement in adipogenesis by studying their expression in undifferentiated cells. Genes not expressed in preadipocytes or if there was a >2-fold difference in expression between the differentiated and undifferentiated state were considered adipocyte-specific (**Figure 5B**), and likely a consequence of the phenotype and not causative genes. RT-qPCR data of three independent experiments, p-values (p-val) comparing expression in mutated cells with WT cells were calculated by Student's t-test with Benjamini-Hochberg correction (**Table 1**, columns 7, 9, 11, 13, 15 and 17).

Supplemental References

Bae, S., Park, J., and Kim, J.S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics *30*, 1473-1475.

Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res *42*, e168.

Cavalli, M., Pan, G., Nord, H., Wallen Arzt, E., Wallerman, O., and Wadelius, C. (2016). Allele-specific transcription factor binding in liver and cervix cells unveils many likely drivers of GWAS signals. Genomics *107*, 248-254.

Fischer-Posovszky, P., Newell, F.S., Wabitsch, M., and Tornqvist, H.E. (2008). Human SGBS cells - a unique tool for studies of human fat cell biology. Obes Facts *1*, 184-189.

Ramirez-Zacarias, J.L., Castro-Munozledo, F., and Kuri-Harcuch, W. (1992). Quantitation of adipose conversion and triglycerides by staining intracytoplasmic lipids with Oil red O. Histochemistry *97*, 493-497.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science *343*, 84-87.