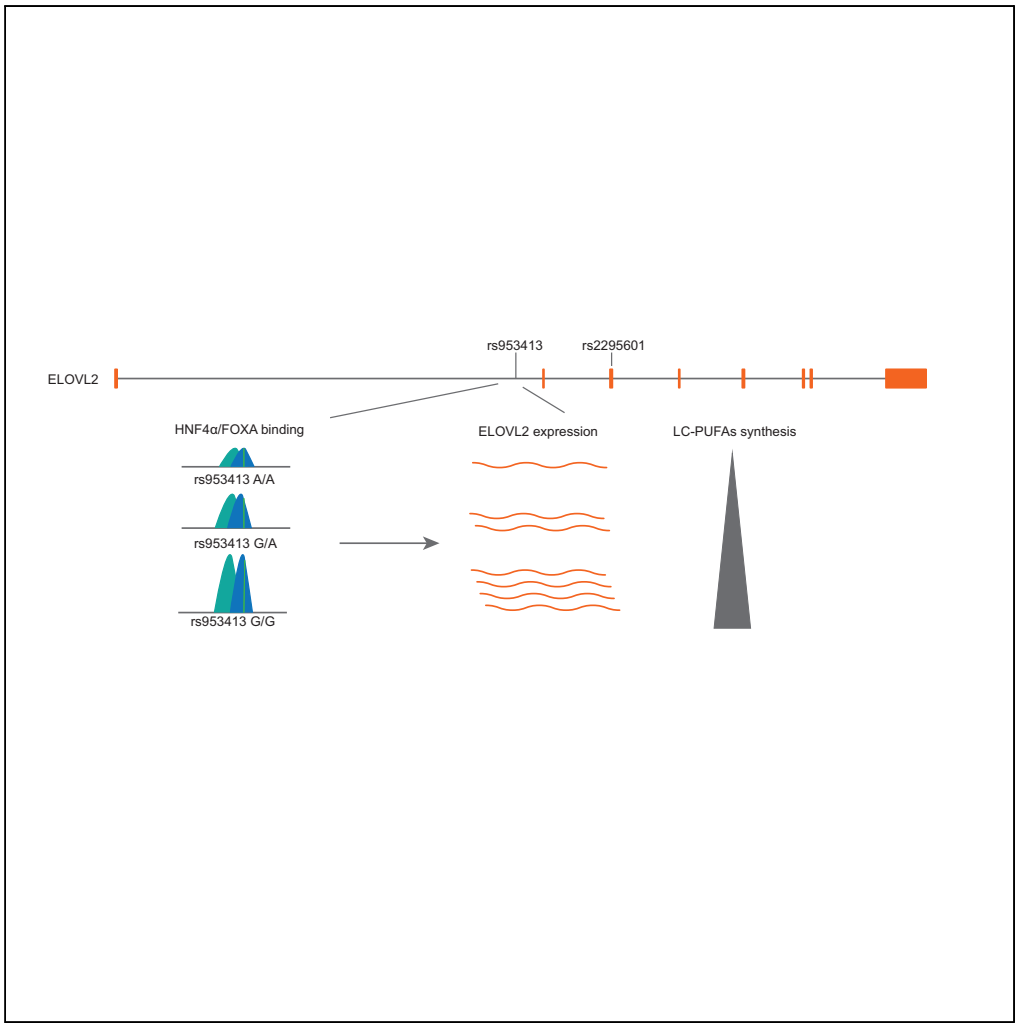
iScience

CellPress

Article

# rs953413 Regulates Polyunsaturated Fatty Acid Metabolism by Modulating *ELOVL2* Expression




Gang Pan, Marco Cavalli, Björn Carlsson, Stanko Skrtic, Chanchal Kumar, Claes Wadelius

claes.wadelius@igp.uu.se

## HIGHLIGHTS

rs953413 resides in an evolutionarily conserved enhancer region

rs953413 mediates the cooperative binding of FOXA and HNF4α to the enhancer region

The rs953413 locus plays a key role in regulating *ELOVL2* expression

rs953413 is implicated in PUFA metabolism by regulating *ELOVL2* expression

Pan et al., iScience *23*, 100808 February 21, 2020 
https://doi.org/10.1016/j.isci.2019.100808

iScience

CellPress

Article

# rs953413 Regulates Polyunsaturated Fatty Acid Metabolism by Modulating *ELOVL2* Expression

Gang Pan,[1] Marco Cavalli,[1] Björn Carlsson,[2] Stanko Skrtic,[3,4] Chanchal Kumar,[5,6] and Claes Wadelius[1,7,*]

## SUMMARY

**Long-chain polyunsaturated fatty acids (LC-PUFAs) influence human health in several areas, including cardiovascular disease, diabetes, fatty liver disease, and cancer. *ELOVL2* encodes one of the key enzymes in the *in vivo* synthesis of LC-PUFAs from their precursors. Variants near *ELOVL2* have repeatedly been associated with levels of LC-PUFA-derived metabolites in genome-wide association studies (GWAS), but the mechanisms behind these observations remain poorly defined. In this study, we found that rs953413, located in the first intron of *ELOVL2*, lies within a functional FOXA and HNF4α cooperative binding site. The G allele of rs953413 increases binding of FOXA1/FOXA2 and HNF4α to an evolutionarily conserved enhancer element, conferring allele-specific upregulation of the rs953413-associated gene *ELOVL2*. The expression of *ELOVL2* was significantly downregulated by both *FOXA1* and *HNF4α* knockdown and CRISPR/Cas9-mediated direct mutation to the enhancer element. Our results suggest that rs953413 regulates LC-PUFAs metabolism by altering *ELOVL2* expression through FOXA1/FOXA2 and HNF4α cooperation.**

## INTRODUCTION

LC-PUFAs, including ω-3 and ω-6, are essential fatty acids to mammals that cannot be synthesized *de novo*. The ω-6 arachidonic acid, ω-3 eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA) are obtained through the diet either directly or as their precursors, linoleic acid and alpha-linolenic acid. Bioconversion of LC-PUFAs from their precursors is mediated by desaturases encoded by fatty acid desaturase 1 (*FADS1*) and *FADS2* in the *FADS* gene cluster and elongases encoded by *ELOVL* fatty acid elongase 2 (*ELOVL2*) and *ELOVL5* (Nakamura and Nara, 2004; Zhang et al., 2016).

Disrupting any of the desaturases and elongases blocked the *in vivo* synthesis of LC-PUFAs, which significantly affected normal growth and development in mouse models (Fan et al., 2012; Moon et al., 2009; Stroud et al., 2009; Zadravec et al., 2011). Recently, the ELOVL2 enzyme, in addition to FADS2, was verified to be another key enzyme in the synthesis of DHA *in vivo* (Gregory et al., 2011; Pauter et al., 2014, 2017). SNPs in the *ELOVL2* locus are strongly associated with levels of LC-PUFA-derived metabolites in blood in many GWAS and meta-analyses (Table S1). Minor alleles of SNPs in the *ELOVL2* locus were associated with higher levels of EPA and DPA and lower levels of DHA in plasma phospholipids, suggesting carriers of the minor alleles have lower efficiency in the bioconversion cascade from EPA to DPA and DHA catalyzed by the ELOVL2 enzyme (Lemaitre et al., 2011; Suhre et al., 2011). The identification of genetic determinants of plasma and tissue lipid levels is a key enabling factor for successful future personalized medicine interventions with EPA/DHA supplements (Chilton et al., 2017). The recent successful outcome of REDUCE-IT illustrates the value of evaluating the effects of EPA/DHA treatment on specific patient segments (hypertriglyceridemia and related lipoprotein abnormalities selected biochemically or genetically) with elevated risk that is believed to be at least partly attributable to an elevated level of the target of the intervention (Bhatt et al., 2019).

The identification of functional variants behind GWAS is hindered by the linkage disequilibrium (LD) in associated regions. To prioritize the variants in LD with reported SNPs in GWAS for functional validation, we developed an allele-specific (AS)-SNPs pipeline (Cavalli et al., 2016b). This pipeline utilizes the available chromatin immunoprecipitation coupled with massively parallel sequencing (ChIP-seq) data from the Encyclopedia of DNA Elements (ENCODE) project to detect variants located in regulatory regions and that showed allelic imbalance in chromatin binding by different transcription factors (TFs) (The Encode Project Consortium, 2012).

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

[2]Research and Early Development, Cardiovascular, Renal and Metabolism, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

[3]Pharmaceutical Technology & Development, AstraZeneca AB, Gothenburg, Sweden

[4]Department of Medicine, Sahlgrenska University Hospital, Gothenburg, Sweden

[5]Translational Science & Experimental Medicine, Research and Early Development, Cardiovascular, Renal and Metabolism (CVRM), BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

[6]Karolinska Institutet/AstraZeneca Integrated CardioMetabolic Center (KI/AZ ICMC), Department of Medicine, Novum, Huddinge, Sweden

[7]Lead Contact

*Correspondence: claes.wadelius@igp.uu.se

https://doi.org/10.1016/j.isci.2019.100808

Liver is the most important organ for LC-PUFAs synthesis and *ELOVL2* is also highly expressed in the liver. We hypothesized that the observed association of LC-PUFA profiles in blood to the *ELOVL2* locus is mainly reflecting the differences in LC-PUFAs metabolism in the liver. Among all the AS-SNPs identified in human liver HepG2 cells, rs953413 located in the first intron of *ELOVL2* showed significant allelic imbalance in chromatin binding. In the present study, we found that rs953413, located in an evolutionarily conserved enhancer element, directly regulates *ELOVL2* expression by mediating the cooperative binding between FOXA and HNF4α factors and that it is an important contributor to the observed difference in LC-PUFAs levels among carriers of different alleles in GWAS.

## RESULTS

### Prioritization of Candidate Regulatory Variations in the *ELOVL2* Locus

The reported SNPs in the *ELOVL2* locus span a region of ~90kb that totally covers the gene body of *ELOVL2* and are observed to be in high LD in the European populations (Figure S1). The alleles of the reported SNPs in LD with the major G allele of rs953413 were generally associated with a more efficient synthesis of LC-PUFAs from their precursors, e.g. synthesis of DHA from EPA and DPA (Draisma et al., 2015; Illig et al., 2009; Lemaitre et al., 2011; Li et al., 2018; Suhre et al., 2011). To pinpoint the true causal variants in the *ELOVL2* locus, fine mapping of variants in this region to a 99% credible set was carried out in a trans-ethnic meta-analysis in Chinese- and European-ancestry populations (Hu et al., 2016). A total of 123 potential regulatory variations were identified in the HaploReg database that are in LD ($r^2 > 0.8$ in 1,000 Genomes Pilot I) with the SNPs in the credible set led by rs3798713 for EPA, rs9393915 for DPA, and rs953413 for DHA (Hu et al., 2016). As liver is the most important organ for *in vivo* synthesis of LC-PUFAs, we hypothesized that the observed association of LC-PUFA profiles in blood to the *ELOVL2* locus is mainly reflecting the differences in LC-PUFAs metabolism in the liver. Only the potential regulatory variants that are annotated to be located in active promoter or enhancer regions in the liver that consist of 47 variations were kept for further consideration (Figure 1A; detailed information on variations is listed in Table S2).

To further prioritize the identified regulatory variants, we intersected the 47 regulatory variants with SNPs in the *ELOVL2* locus that showed allelic imbalance in chromatin binding identified by the AS-SNPs pipeline in HepG2 cells. A total of three SNPs in the *ELOVL2* locus including rs953413, rs3798713, and rs17675073 were identified to be AS-SNPs in HepG2 cells that showed allelic imbalance in chromatin binding for FOXA1 and FOXA2 (Table S3) (Cavalli et al., 2016a). As both rs953413 and rs3798713 were identified to be lead SNPs in the calculated credible set and showed allelic imbalance in chromatin binding for FOXA1 and FOXA2, they were selected for further functional validation.

### rs953413 Resides in a Liver-Specific Enhancer

The *ELOVL2* locus has copy number variation (CNV) in HepG2 cells, which has two copies of chromosome bearing the G allele of rs953413 (Lopez-Terrada et al., 2009; Zhou et al., 2018). To verify that the observed allelic imbalance in rs953413 and rs3798713 is not only caused by the CNV and to identify candidate causal TFs, we employed a more stringent strategy that filtered out PCR duplicates and tolerated no mismatch other than the SNP itself to each TF that mapped to these two regions in ChIP-seq experiments in HepG2 cells. Among all the TFs that have peaks overlapping or nearby rs953413 and rs3798713, only FOXA1 and FOXA2 showed reproducible allelic imbalance in chromatin binding with a preference for the G alleles of rs953413 and rs3798713 (Table S3). In accordance with occupation by multiple TFs, both the rs953413 region and rs3798713 region are highly enriched with H3K27ac signals and are also located in active DNase I hypersensitivity sites in HepG2 cells (Figure 1A). Additionally, both the rs953413 region and rs3798713 region are observed to directly interact with the promoter region of *ELOVL2* identified by Capture Hi-C in GeneHancer, suggesting that these regions may be actively involved in *ELOVL2* regulation (Figure 1A) (Fishilevich et al., 2017).

We next performed luciferase assay to test the enhancer activity of the two selected regions. As rs17675073 is in proximity to rs3798713, we amplified a 287-bp fragment encompassing both SNPs and inserted upstream of the minimal promoter (MP) to test its ability to drive luciferase expression. This region did show enhancer activity compared with the control plasmid without insertion. However, neither of the SNPs showed significant differences between alleles conditional on the other SNP (Figure S2). This region was then discarded for further analysis. For the rs953413 region, the enhancer activity was tested with a 529-bp fragment with rs953413 located in the middle. Both alleles of rs953413 were observed to have strong enhancer activity in cooperation with the MP in driving luciferase expression. The G allele of rs953413

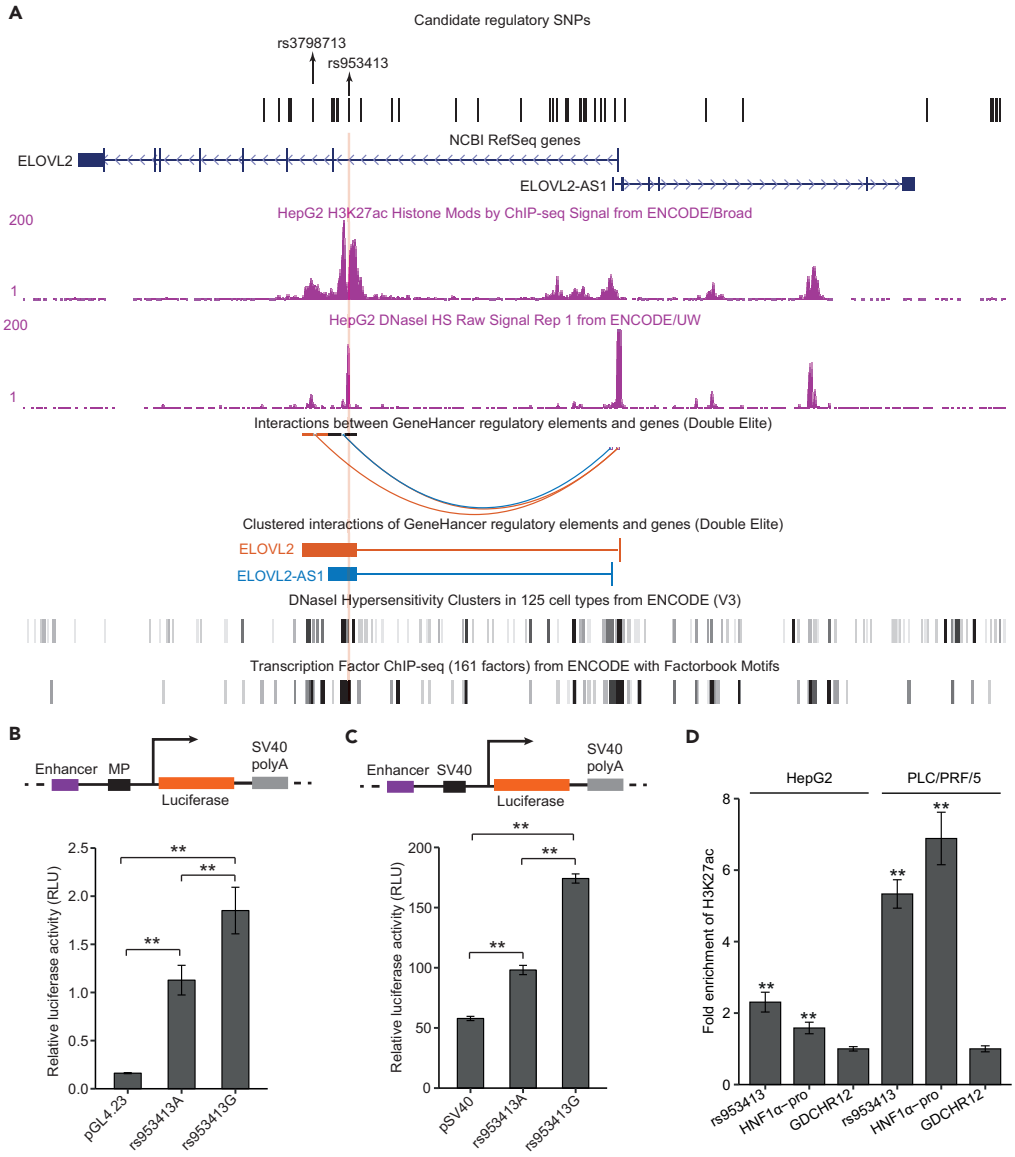



**Figure 1. SNP rs953413 Locates in an Active Enhancer Region to Regulate Its Enhancer Activity**

(A) Both rs953413 and rs3798713 are located in highly active DNase I hypersensitivity sites that are also enriched with H3K27ac signals in HepG2 cells. The candidate regulatory SNPs in the *ELOVL2* locus selected from the credible set that are located in regulatory regions in the liver are listed.

(B and C) The G allele of rs953413 showed significant higher enhancer activity compared with the A allele in luciferase assay coupled with both the MP (B) and SV40 promoter (C). Error bars, s.d. $n = 6$ from two independent plasmid extractions and transfections with each transfection had three technical replicates.

(D) The rs953413 region is enriched with H3K27ac signal in both HepG2 and PLC/PRF/5 cells compared with the negative control region GDCHR12 in ChIP-qPCR. The HNF1α-pro is employed as the positive control. Error bars, s.d. $n = 3$ technical replicates. In B–D, **$p < 0.01$ evaluated using two-tailed Student's *t* tests.

See also Figure S2.

had a significantly higher enhancer activity compared with the A allele (Figure 1B). This difference in enhancer activity between alleles of rs953413 was further validated with the SV40 promoter in the other luciferase construct (Figure 1C).

The meta-analysis carried out in the CHARGE Consortium provided us with the comprehensive association results for ω-3 fatty acids in plasma phospholipid in the *ELOVL2* locus (Lemaitre et al., 2011). rs953413 and

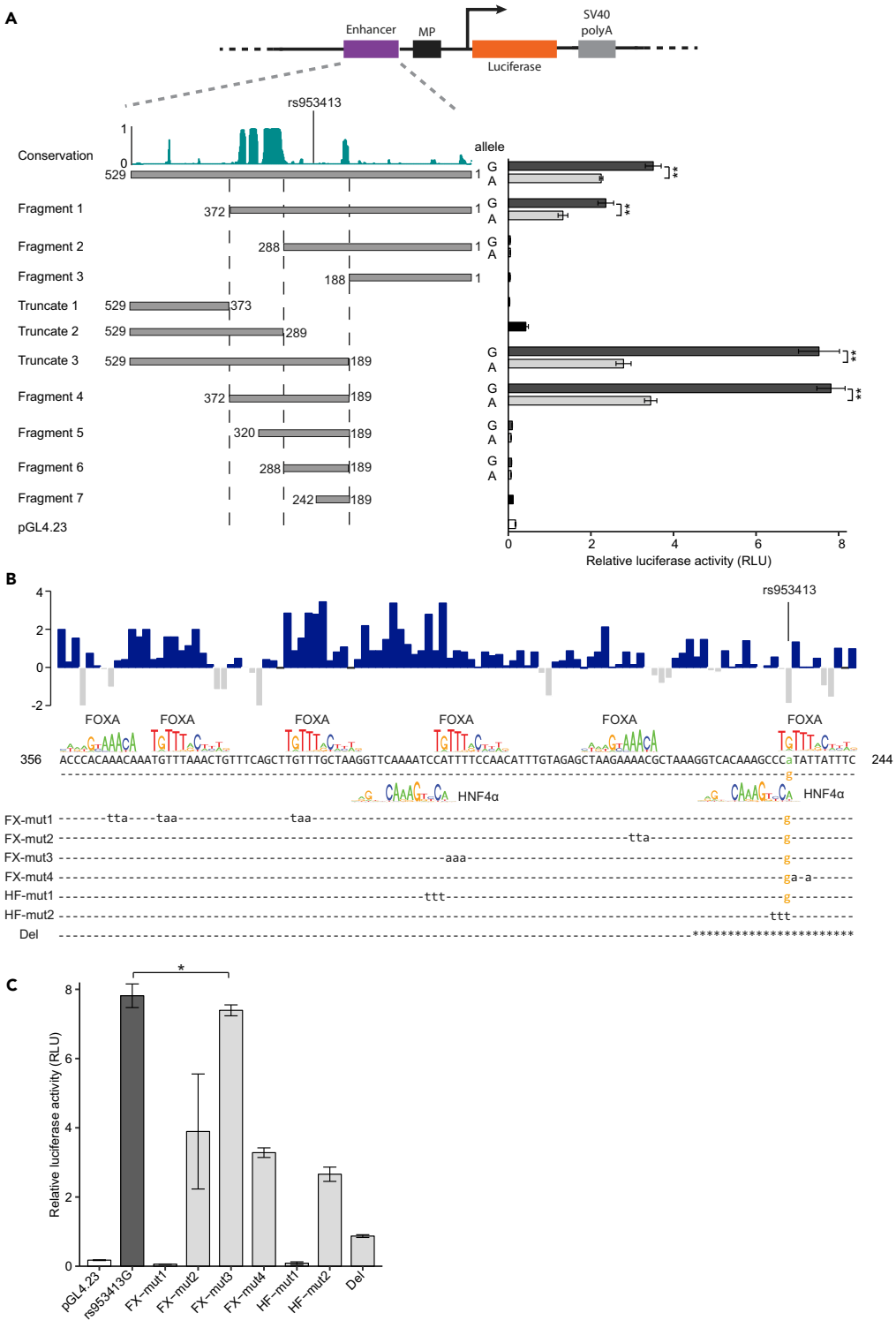



**Figure 2. Fine Mapping of the ME Element of the rs953413 Region to an Evolutionarily Conserved Region that Is Enriched with Putative Binding Sites for Both FOXA and HNF4α**

(A) The ME element is fine mapped to an evolutionarily conserved region encompassing rs953413 (Fragment 4) by luciferase assay with a series of truncation and deletion constructs from the original 529-bp fragment. The Conservation track showed the phastCons score for 100 species from the ENCODE project.

**Figure 2. *Continued***

(B) The ME element was predicted to have multiple conserved binding sites for both FOXA and HNF4α. The basewise conservation score for 100 species from the ENCODE project is displayed on top of the sequence. The luciferase constructs with mutations introduced into the predicted motif sequences are generated by site-directed mutagen and listed at the bottom. The introduced mutations for each mutation construct are listed as lower-case letters with dash line standing for identical nucleotides. For the deletion construct (Del), the deleted region is marked with star character.

(C) The enhancer activity of the rs953413 region is significantly decreased by mutations introduced into the predicted binding sites for HNF4α and FOXA factors (as illustrated in B) in luciferase assay. Luciferase construct Fragment 4 bearing the G allele of rs953413 was employed as template for mutation. For A and C, error bars, s.d. $n = 6$ technical replicates from two independent plasmid extractions and transfections with each transfection had three technical replicates. *$p < 0.05$; **$p < 0.01$ calculated by two-tailed Student's *t* tests.

its proxies ($r^2 > 0.8$ in 1,000 Genomes phase 3) generally showed stronger association with EPA, DPA, and DHA levels compared with other SNPs (Figure S3). In addition, the rs953413 region showed tissue-specific H3K27ac signals with enrichment of H3K27ac binding only observed in HepG2, A549, and PANC-1 cells from all the available cell lines in the ENCODE project (Figure S4). The enrichment of H3K27ac over the rs953413 region was further validated with ChIP followed by quantitative real time (ChIP-qPCR) in liver cancer cell lines HepG2 and PLC/PRF/5 (Figure 1D). The enrichment of H3K27ac signal together with enrichment of TFs binding to the rs953413 region only in liver cells support the idea that the rs953413 region shows liver-specific enhancer feature (Figure S4). Interestingly, although not coming to genome-wide significant level, carriers of the G allele of rs953413 were observed to have higher expression of *ELOVL2* ($p = 1.4 \times 10^{-4}$) in liver tissues from the GTEx project (Figure S5) (Battle et al., 2017). This observation coupled with the preferential binding of FOXA1 and FOXA2 to the G allele of rs953413 suggests that rs953413 may regulate *ELOVL2* expression by modulating the enhancer activity of the rs953413 region in liver, which further affects the synthesis of LC-PUFAs.

## The rs953413 Region Is Evolutionarily Conserved

To identify the causal TFs mediating the difference in enhancer activity in the rs953413 region, we first isolated the minimal enhancer (ME) element for this region. A series of truncation and deletion luciferase constructs were made with both genotypes of the original 529-bp fragments as templates and inserted upstream of the MP sequence as in Figure 1B. The ME element was mapped to a 184-bp fragment (Fragment 4) that showed strong enhancer activity while keeping the sequence at minimum (Figure 2A). This ME element includes a region of around 80-bp in length that is highly conserved in vertebrates and is essential for the observed enhancer activity as only luciferase constructs containing the whole conserved region (the original 529-bp fragment, Fragment 1, 4, and Truncate 3) showed strong enhancer activity. The sequences surrounding rs953413 are less conserved in vertebrates but are also essential for the enhancer activity as the luciferase construct containing only the evolutionarily conserved region (Truncate 2) had drastically decreased enhancer activity compared with luciferase construct containing both the conserved element and the rs953413 surrounding sequence (Truncate 3) (Figure 2A).

We then searched for TFs predicted to bind differently to the rs953413 region with the ME element as template and using the TRAP tool with position weight matrices (PWMs) from both the JASPAR database and literature as references (Mathelier et al., 2014; Thomas-Chollier et al., 2011). In accordance with the observed allelic imbalance in chromatin binding for FOXA1 and FOXA2 in ChIP-seq, we identified multiple highly conserved FOXA motifs together with two conserved HNF4α motifs in the ME element (Figure 2B). The predicted motifs are located in evolutionarily conserved sequences denoted by the basewise conservation in 100 vertebrates from the ENCODE project (Figure 2B). The variant rs953413 is predicted to be located in two closely located motifs bound by FOXA and HNF4α indicative of cooperative binding by these TFs (Jolma et al., 2015).

We next examined the effect of the predicted TF binding sites on the enhancer activity of the rs953413 region by site-directed mutagenesis. Fragment 4 bearing the G allele of rs953413 and containing the whole ME element was employed as the template for construction of mutants (Figure 2B). We observed that introducing mutations at any of the predicted TF binding sites significantly decreased the enhancer activity of the rs953413 region (Figure 2C). The enhancer activity was completely depleted when mutations were introduced to the predicted binding sites in highly conserved region (constructs FX-mut1 and HF-mut1). Notably, introducing mutations to the predicted binding sites overlapping with rs953413 (constructs FX-mut4, HF-mut2, and Del) also severely decreased the enhancer activity, suggesting that rs953413 is functional in determining the enhancer activity of this region (Figures 2B and 2C).

### FOXA1 and FOXA2 Regulates *ELOVL2* Expression by Binding to the rs953413 Region

We next examined if both *FOXA1* and *FOXA2* are directly involved in the regulation of the enhancer activity of the rs953413 region. In accordance with our motif prediction, the minimal region that responds to both *FOXA1* and *FOXA2* induction is mapped to the luciferase construct containing the whole ME element (Fragment 4) from a series of deletion constructs in luciferase assay (Figure S6). Both alleles of rs953413 were highly induced by both *FOXA1* and *FOXA2* overexpression (Figure 3A). The difference in enhancer activity between alleles of rs953413 was significantly compromised by *FOXA* overexpression probably due to presence of multiple conserved FOXA binding sites in this region. The responses to both *FOXA1* and *FOXA2* overexpression were significantly decreased by introducing mutations to the predicted FOXA binding sites in luciferase construct Fragment 4 bearing the G allele of rs953413. Introducing mutations to the predicted HNF4α binding sites also made this region less responsive to both *FOXA1* and *FOXA2* overexpression with significantly decreased enhancer activities compared with the wild type (Figure 3B). Especially, the response to *FOXA1*/*FOXA2* overexpression was completely depleted in luciferase construct HF-mut1 with one of the predicted HNF4α binding sites being mutated, suggesting that binding of both FOXA1/FOXA2 and HNF4α to the rs953413 region is essential for the enhancer activity.

We next assessed the *in vivo* binding of both FOXA1 and FOXA2 to the rs953413 region through ChIP-qPCR. In accordance with the enrichment of ChIP-seq signals for both FOXA1 and FOXA2 in this region in HepG2 cells from the ENCODE project, the rs953413 region was observed to be highly enriched with FOXA1 and FOXA2 binding in parallel with the positive control region HNF1α-pro after normalization with the negative control region GDCHR12 in both HepG2 and PLC/PRF/5 cells (Figure 3C). To address whether FOXA1/FOXA2 showed allele-specific DNA binding at rs953413 *in vivo*, we conducted ChIP followed by allele-specific qPCR (AS-qPCR). The allele-specific amplification of each allele was achieved with primers designed following mismatch amplification mutation assay (MAMA) and validated with genomic DNA from samples bearing different genotypes at rs953413 as templates (Figure S7) (Cha et al., 1992; Li et al., 2004). In accordance with the observed preference for the G allele in ChIP-seq experiments, both FOXA1 and FOXA2 were identified to be preferentially bound to the G allele of rs953413 in ChIP followed by AS-qPCR in both HepG2 and PLC/PRF/5 cells (Figure 3D). This allelic imbalance in chromatin binding for both FOXA1 and FOXA2 at the rs953413 region was further confirmed with Sanger sequencing analysis of the rs953413-containing region in chromatin fragments immunoprecipitated with antibody against both FOXA1 and FOXA2 compared with that in the input genomic DNA as control (Figure 3E). Additionally, the G allele of rs953413 was significantly more enriched with H3K27ac signal compared with the A allele in ChIP samples from Figure 1D in both HepG2 and PLC/PRF/5 cells, suggesting more of FOXA binding in this region leading to stronger enhancer activity (Figure 3F).

To determine if *FOXA1* and *FOXA2* are directly involved in regulation of *ELOVL2* expression, we next performed lentiviral-mediated *FOXA1* knockdown through short hairpin RNA (shRNA) and artificial microRNA (amiRNA) in HepG2 and PLC/PRF/5 cells (Liang et al., 2012). Knockdown of *FOXA1* significantly decreased the expression of *ELOVL2* in both cells after normalization to reference genes *RSP18*, *ACTB*, and *GAPDH*, respectively (Figures 3G and S8). The expression of *FOXA2* is significantly increased by *FOXA1* knockdown in HepG2 cells, whereas it is observed to be significantly downregulated by *FOXA1* knockdown in PLC/PRF/5 cells, indicative of cell heterogeneity between these two cells. The A allele of exonic SNP rs2295601 of *ELOVL2* is on the same haplotype as the A allele of rs953413 evaluated with the data from the 1,000 Genomes Project (Machiela and Chanock, 2015). The efficiency and specificity of AS-qPCR assay targeting rs2295601 was validated with genomic DNA from the same set of samples for rs953413 AS-qPCR as templates (Figure S7). We have evidence that FOXA1/FOXA2 favor binding to the G allele of rs953413 (Figures 3D and 3E). We therefore evaluated whether this allele-specific binding caused allelic imbalance in *ELOVL2* expression with rs2295601 AS-qPCR in HepG2 cells, which is heterozygous at this location, whereas PLC/PRF/5 was identified to be homozygous with the G allele. This analysis identified that the G allele of rs2295601 in linkage with the G allele of rs953413 was selectively transcribed in HepG2 cells. Knockdown of *FOXA1* significantly diminished the allelic imbalance in *ELOVL2* expression (Figure 3H).

### HNF4α Regulates *ELOVL2* Expression through Interacting with the rs953413 Region

We next assessed whether HNF4α is also directly involved in *ELOVL2* regulation through interacting with the rs953413 region. We observed that luciferase construct (Fragment 5) encompassing both of the predicted HNF4α binding sites was highly induced by *HNF4α* overexpression (Figure 4A). To refine the

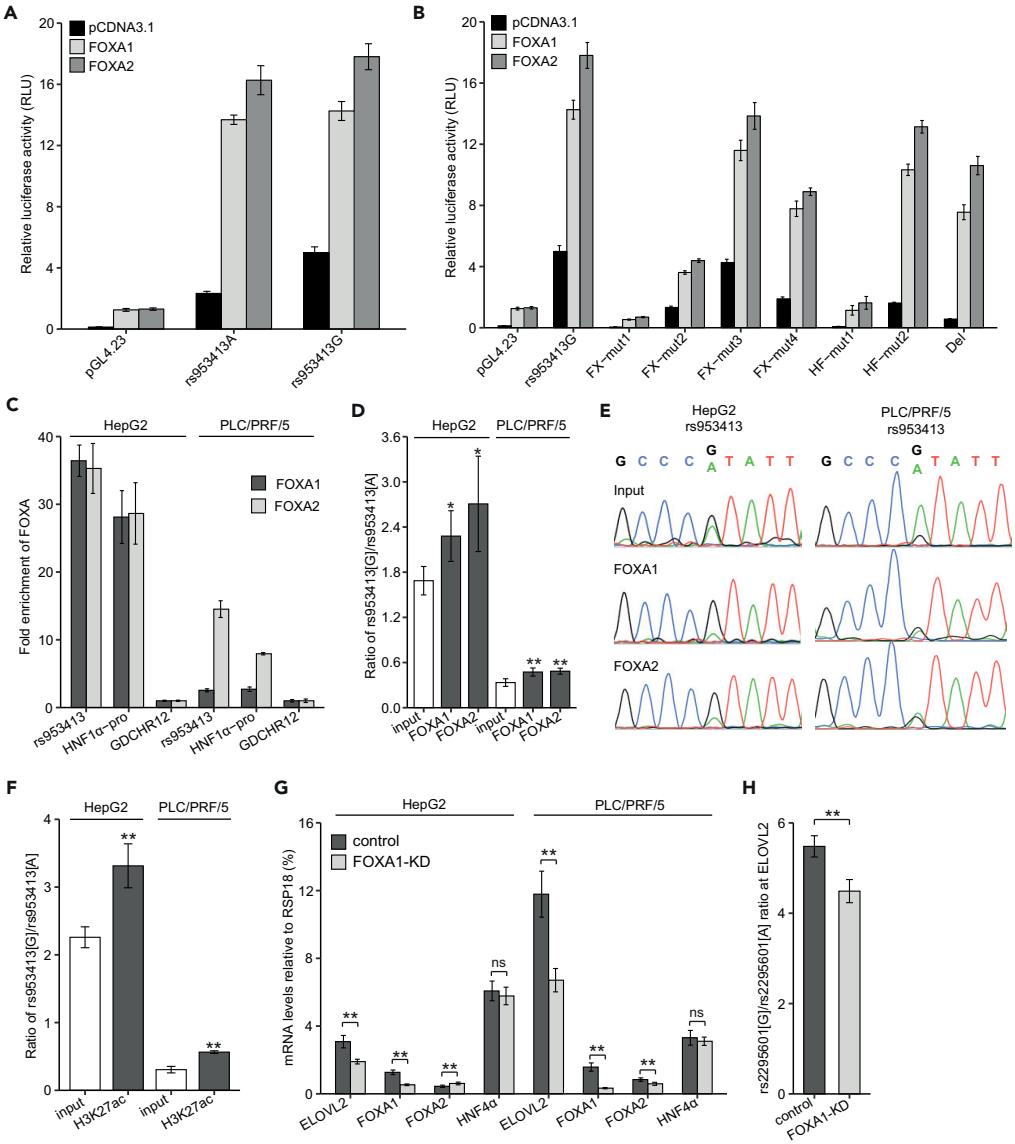



**Figure 3. FOXA Factors Directly Bind to the ME Element and Confer Allelic Imbalance in *ELOVL2* Expression Induced by rs953413**

(A) The enhancer activity of the rs953413 region is highly induced by both *FOXA1* and *FOXA2* overexpression in luciferase assay.

(B) The response to either *FOXA1* or *FOXA2* overexpression is significantly compromised by the mutations introduced into the predicted binding sites for HNF4α and FOXA factors in luciferase assay. In A and B, error bars, s.d. $n = 6$ technical replicates from two independent plasmid extractions and transfections with each transfection had three technical replicates.

(C) Both FOXA1 and FOXA2 are highly enriched in the rs953413 region verified by ChIP-qPCR in both HepG2 and PLC/PRF/5 cells. Error bars, s.d.

(D and E) FOXA1 and FOXA2 favor binding to the G allele of rs953413 as determined by ChIP followed by AS-qPCR (D) and ChIP followed by PCR amplification and Sanger sequencing (E) in both HepG2 and PLC/PRF/5 cells. The input for ChIP was used as the control. *$p < 0.05$ and **$p < 0.01$ calculated by two-tailed Student's *t* tests. Error bars, s.d. In C and D, $n = 4$ for HepG2 cells and $n = 3$ for PLC/PRF/5 cells.

(F) The G allele of rs953413 is also significantly more enriched with H3K27ac signal compared with the A allele determined by ChIP followed by AS-qPCR. $n = 3$. Error bars, s.d. **$p < 0.01$ calculated by two-tailed Student's *t* tests.

(G) The expression of *ELOVL2* is significantly downregulated by *FOXA1* knockdown in HepG2 and PLC/PRF/5 cells.

(H) The allelic imbalance in *ELOVL2* expression is significantly decreased by *FOXA1* knockdown in HepG2 cells. The raw ratios are shown here without correcting for the CNV. In G and H, **$p < 0.01$ and ns, not significant, calculated by two-tailed Student's *t* tests. Error bars, s.d. $n = 8$ technical replicates.

See also Figures S6–S8.

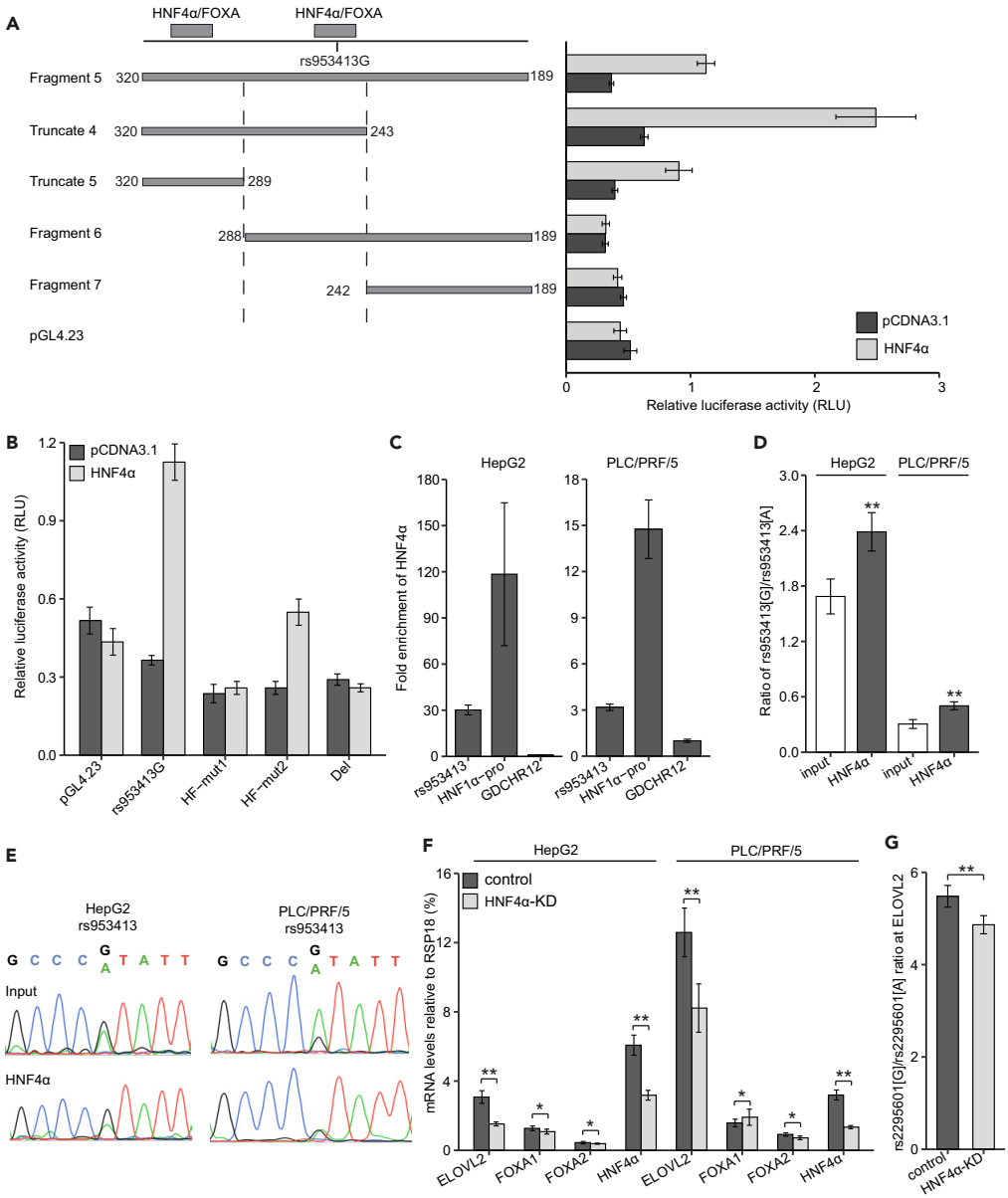



**Figure 4. HNF4α regulates *ELOVL2* Expression by Directly Binding to the ME Element and Confers Allelic Imbalance in *ELOVL2* Expression by rs953413**

(A) Both of the predicted HNF4α binding sites in the ME element are essential for the maximal induction by *HNF4α* in luciferase assay.

(B) Disruption either of the predicted HNF4α-binding sites significantly decreased the enhancer activity induced by *HNF4α* in luciferase assay. In A and B, error bars, s.d. $n = 6$ technical replicates from two independent plasmid extractions and transfections with each transfection had three technical replicates.

(C) The rs953413 region is highly enriched with HNF4α binding compared with GDCHR12 in ChIP-qPCR. Error bars, s.d.

(D and E) HNF4α favors binding to the G allele of rs953413 relative to the A allele as determined by ChIP followed by AS-qPCR (D) and ChIP followed by PCR amplification and Sanger sequencing (E) in both HepG2 and PLC/PRF/5 cells. The input for ChIP was used as the control. **$p < 0.01$ calculated by two-tailed Student's *t* tests. Error bars, s.d. In C and D, $n = 4$ for HepG2 cells and $n = 3$ for PLC/PRF/5 cells.

(F) The expression of *ELOVL2* is significantly decreased by *HNF4α* knockdown in both HepG2 and PLC/PRF/5 cells.

(G) The allelic imbalance in *ELOVL2* expression is significantly decreased by *HNF4α* knockdown in HepG2 cells. The raw ratios are shown here without correcting for the CNV. In F and G, error bars, s.d. $n = 8$ technical replicates.

See also Figures S7 and S9.

*HNF4α* responsive regions, truncation and deletion luciferase constructs were further constructed and subjected to *HNF4α* overexpression in luciferase assay. In accordance with our prediction, only constructs containing the predicted HNF4α-binding sites (Fragment 5, Truncate 4 and 5) were induced by *HNF4α* overexpression. To map the sequences needed for *HNF4α* induction, we introduced mutations to either of the predicted HNF4α sites with Fragment 5 as template that led to drastically diminished *HNF4α* induction (Figure 4B). Intriguingly, the Del construct (as illustrated in Figure 2B) with the predicted binding sites for HNF4α and FOXA factors surrounding rs953413 being removed completely abolished the *HNF4α* induction, suggesting rs953413 is directly located in one functional HNF4α-binding site.

We next performed ChIP-qPCR assays with antibody against HNF4α in both HepG2 and PLC/PRF/5 cells. In accordance with the data from the ENCODE project, we confirmed HNF4α chromatin binding at the rs953413 region and at the positive HNF1α-pro region after normalization with GDCHR12 (Figure 4C). To address whether HNF4α showed allele-specific DNA binding at rs953413 *in vivo*, we conducted ChIP followed by rs953413 AS-qPCR. This analysis showed that HNF4α was preferentially bound to the G allele of rs953413 in both HepG2 and PLC/PRF/5 cells (Figure 4D). Consistent with this finding, Sanger sequencing analysis of the rs953413-containing region showed that the G allele was enriched in chromatin fragments immunoprecipitated with antibody against HNF4α compared with the input genomic DNA as control in both cells (Figure 4E). We next performed lentiviral-mediated *HNF4α* knockdown through shRNA and amiRNA to test its effect on *ELOVL2* expression (Liang et al., 2012). Knockdown of *HNF4α* resulted in significantly decreased expression of *ELOVL2* after normalization to reference genes *RSP18*, *ACTB*, and *GAPDH*, respectively (Figures 4F and S9). We then performed rs2295601 AS-qPCR to evaluate if the allelic imbalance in *ELOVL2* expression was affected by *HNF4α* knockdown. After *HNF4α* knockdown, the allelic imbalance in *ELOVL2* expression was also observed to be significantly decreasing (Figure 4G).

## Cooperation between HNF4α and FOXA Factors Determine the Enhancer Activity of the rs953413 Region

As both HNF4α and FOXA factors are bound to the rs953413 region and essential for its enhancer activity, we hypothesized that HNF4α and FOXA factors may cooperate with each other in a complex to determine the enhancer activity of the rs953413 region, which further regulates *ELOVL2* expression. To test this hypothesis, we first carried out luciferase assay overexpressing both *HNF4α* and *FOXA* factors together with luciferase construct Fragment 4 containing the whole ME element. Compared with overexpression of *HNF4α* alone or either *FOXA1* or *FOXA2* alone, the enhancer activities of both alleles of rs953413 were significantly increased by overexpression of *HNF4α* together with either *FOXA1* or *FOXA2,* indicative of cooperation between these TFs in determining the enhancer activity of the rs953413 region (Figures 5A and S10).

To determine the effect of the cooperation between HNF4α and FOXA factors on *ELOVL2* expression, we generated HepG2 and PLC/PRF/5 cells with both *HNF4α* and *FOXA1* being knocked down. This double knockdown was achieved with either sequential or simultaneous transduction of lentivirus for *HNF4α* and *FOXA1* knockdown. The double knockdown cells produced with simultaneous knockdown were named FOXA1 + HNF4α KD. FOXA1-KD + HNF4α-KD denotes cells first transduced with lentivirus for *FOXA1* knockdown followed by *HNF4α* knockdown and vice versa for HNF4α-KD + FOXA1-KD. To be comparable, the control cells and cells with knockdown of only *FOXA1* or *HNF4α* were also subjected to two rounds of lentiviral transduction. Compared with knockdown of only *HNF4α* or *FOXA1*, this double knockdown strategy led to significant downregulation of not only *HNF4α* and *FOXA1* but also *FOXA2* in both cells after normalization to reference genes *RSP18*, *ACTB,* and *GAPDH*, respectively (Figures 5B–5D and S11). It is noteworthy that the degree of knockdown for all the three TFs is much stronger compared with single knockdown in HepG2 cells (Figure 5C). As *HNF4α* and *FOXA1*/*FOXA2* are all members of the hepatocyte nuclear factor family, the cross-regulation of one another and the synergy in gene regulation has been well studied (Lau et al., 2018). We investigated ChIP-seq signals for histone modifications (H3K4me3 and H3K27ac) in HepG2 cells indicating promoters and enhancers over the gene bodies of all the three TFs. We found that ChIP-seq signals for all the three TFs are enriched over the active chromatin regions of each TF's gene body indicative of coregulation of all the three TFs by themselves (Figure S12). Coupling to the knockdown of all the three TFs, the expression of *ELOVL2* was significantly downregulated compared with knockdown of only *HNF4α* or *FOXA1* in both HepG2 and PLC/PRF/5 cells (Figures 5B and S11). In addition, contrary to the observed decrease in allelic imbalance in *ELOVL2* expression after knockdown of only *HNF4α* or *FOXA1 in* HepG2 cells, the triple knockdown led to a significant increase in the allelic imbalance in *ELOVL2* expression (Figure 5E). Interestingly, the degree of allelic imbalance in *ELOVL2*

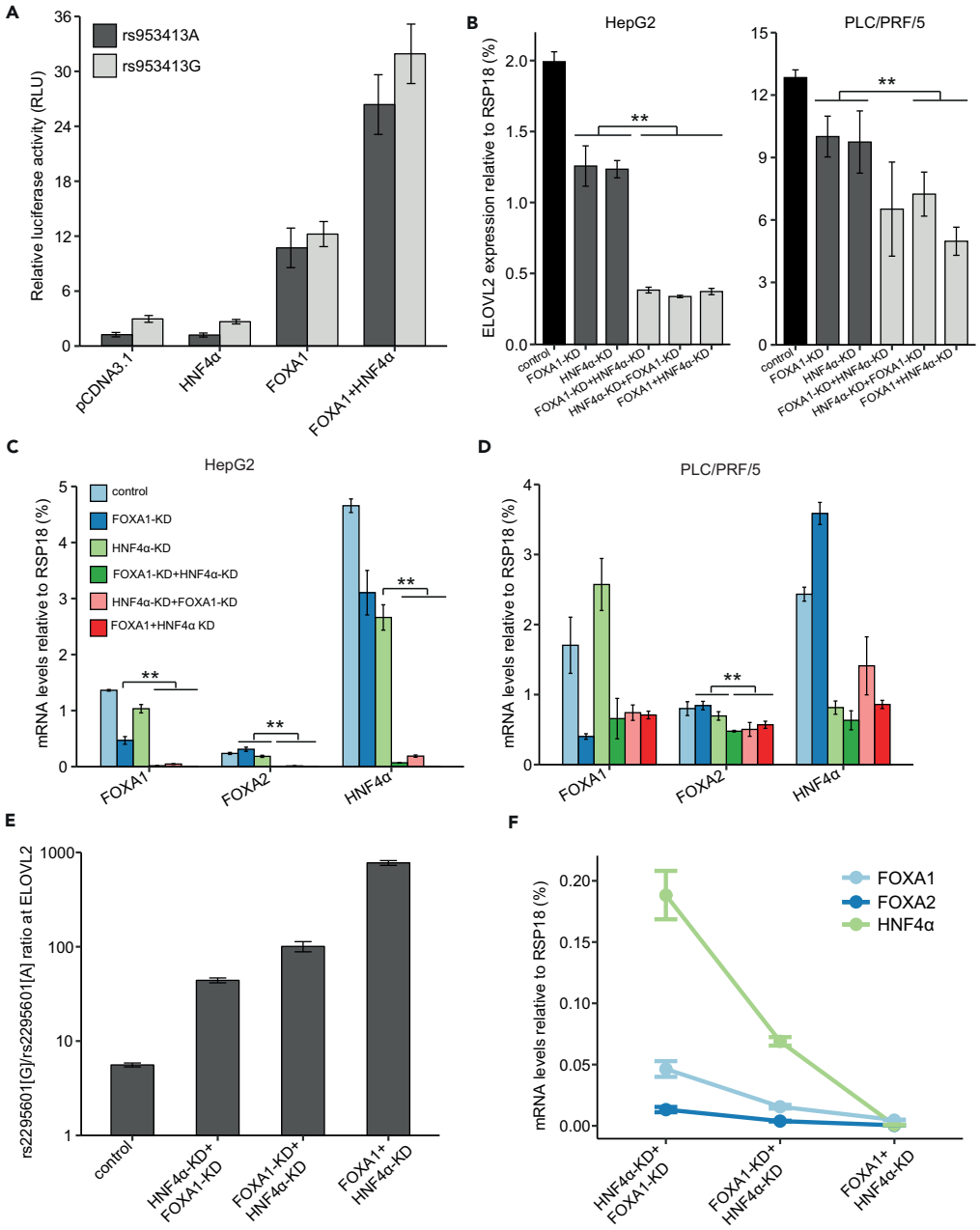



**Figure 5. Cooperation between HNF4α and FOXA Factors Determines the Enhancer Activity of the rs953413 Region and *ELOVL2* Expression**

(A) The enhancer activity of the rs953413 region was significantly increased by simultaneous overexpression of *HNF4α* and *FOXA1* compared with overexpression of only *HNF4α* or *FOXA1* in luciferase assay. Error bars, s.d. $n = 6$ technical replicates from two independent plasmid extractions and transfections with each transfection had three technical replicates.

(B–D) Double knockdown of both *HNF4α* and *FOXA1* led to a significant decrease in *ELOVL2* expression compared with knockdown only *HNF4α* or *FOXA1* (B). This double knockdown also significantly downregulated *FOXA2* expression in both HepG2 cells (C) and PLC/PRF/5 cells (D).

(E and F) The allelic imbalance in *ELOVL2* expression shown in E is gradually increased by the gradual decreases in expression of *HNF4α* and *FOXA1*/*FOXA2* in different double-knockdown HepG2 cells as shown in F. In B–F, $^{**}p < 0.01$ calculated by two-tailed Student's *t* tests. Error bars, s.d. $n = 4$ technical replicates.

See also Figures S10 and S11.

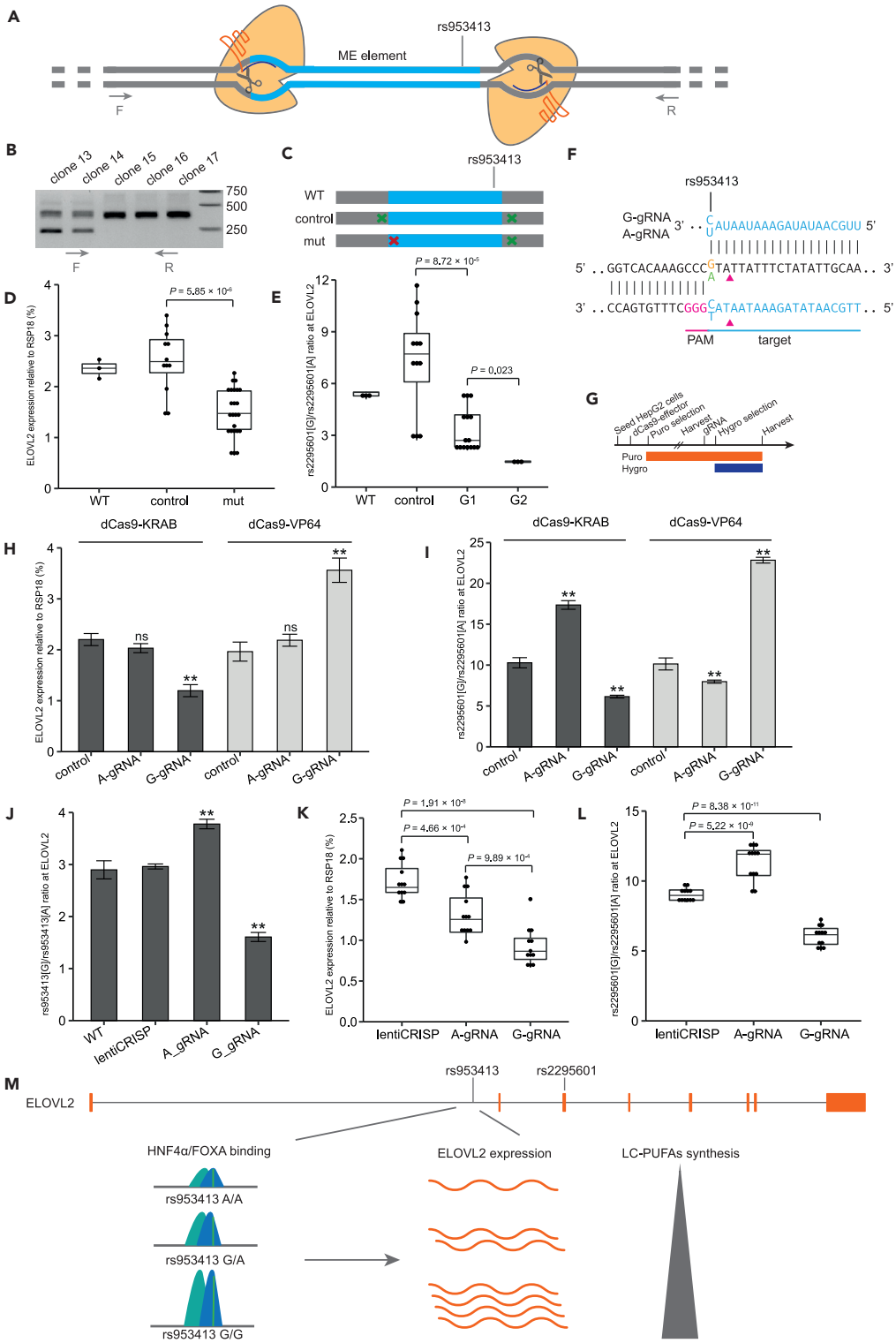



**Figure 6. CRISP/Cas9-mediated Mutation of the ME Element Impairs *ELOVL2* Expression in an Allele-Specific Manner**

(A) Schematic illustrating CRISP/Cas9-mediated mutation of the ME element by double gRNA. The ME element is highlighted in blue color. The binding sites for the primer pair used to validate mutations are indicated.

**Figure 6. *Continued***

(B) Representative gel image showing PCR amplification of the rs953413 region in selected individual clones displayed varieties of mutations introduced by the CRISP/Cas9 system.

(C) Individual clones are grouped by the position of mutations introduced.

(D) Clones with mutations introduced into the ME element of the rs953413 region (mut, $n = 8$) had significantly decreased expression level of *ELOVL2* compared with clones without key mutations introduced (control, $n = 4$). WT denotes the *ELOVL2* expression level in normal HepG2 cells.

(E) The allelic imbalance in *ELOVL2* expression is decreased in a dosage dependent manner by mutations introduced into the ME element bearing the G allele of rs953413. For D and E, each clone had three technical replicates.

(F) Schematic illustrating AS-gRNA targeting rs953413 locus.

(G) Experimental time line for lentiviral transduction and antibiotic selection of HepG2 cells with lentiviral dCas9-effector and AS-gRNA.

(H–L) (H and I) Changes in *ELOVL2* expression (H) and the coupled changes in allelic imbalance in *ELOVL2* expression (I) following transduction of dCas9-KRAB and dCas9-VP64 together with AS-gRNA. In H and I, $^{**}p < 0.01$ and ns, not significant, calculated by two-tailed Student's *t* tests. Error bars, s.d. $n = 4$ technical replicates (J–L). Each AS-gRNA preferentially mutated respective allele of rs953413 when coupled with wild-type Cas9 (J), which significantly downregulated *ELOVL2* expression (K) in an allele-specific manner (L). In J–L, $^{**}p < 0.01$ calculated by two-tailed Student's *t* tests. Error bars, s.d. $n = 4$ in J and $n = 12$ technical replicates in K and L.

(M) Graphic representation of the regulatory relationships between HNF4α/FOXA, rs953413, and *ELVOL2*. HNF4α/FOXA binds to the evolutionarily conserved rs953413 region in the intron of *ELOVL2* and regulates the expression of *ELVOL2*. The G allele of rs953413 results in increased HNF4α/FOXA binding and thus upregulates *ELVOL2* expression, which consequently leads to increased synthesis of LC-PUFAs.

See also Figures S16–S19.

expression was tightly associated with the degree of downregulation of the three TFs, indicating that the G allele of rs953413 is preferentially bound by these TFs when the availability of these TFs is limited, which further determines the allelic imbalance in *ELOVL2* expression (Figure 5F).

## The rs953413 Region Works as a Key Enhancer Region for *ELOVL2* Regulation

To provide more direct evidence that the rs953413 region is involved in *ELOVL2* regulation, we first conducted CRISPR/Cas9-mediated mutagenesis of this region in HepG2 cells. Two guide RNAs were designed to cut the flanking sequences on both sides of the ME element to delete the whole rs953413 region (Figure 6A). After selecting multiple colonies, we did not get single colonies homozygous with each allele of the ME element being deleted due to the existence of CNV in this region. However, we observed that the expression of *ELOVL2* together with the allelic imbalance in *ELOVL2* expression was significantly different among different single colonies with different mutation profiles in the rs953413 region (Figure 6B). To determine the causal relationship between the mutations introduced into the rs953413 region with *ELOVL2* expression, we selected nineteen single colonies with different mutation profiles for further characterization. Sanger sequencing on direct PCR product amplified from the rs953413 region was carried out from both sides to determine the mutation pattern for each colony together with multiple (from 5 to 26) TA clones of PCR product to determine the genotype of each colony. From all the selected colonies, we successfully identified the genotypes of twelve colonies, with the others either having multiple alleles detected or missing the A allele of rs953413 detected.

Our previous findings suggest that the predicted binding sites for FOXA factors are essential for the enhancer activity of the rs953413 region as mutation construct FX-mut1 completely lost the enhancer activity and its response to both *FOXA1* and *FOXA2* induction in luciferase assay (Figure 3B). The twelve colonies can generally be categorized into two groups based on the location of the mutations introduced. The alleles of the control group only have mutations flanking the ME element and the mut group have mutations to the ME element that also disrupted the predicted binding sites for FOXA factors as in mutation construct FX-mut1 (Figure 6C). Four colonies were categorized as control (Figure S13) and eight colonies fell into the category of mut with five colonies having critical mutations in one copy of the ME element bearing the G allele of rs953413 (G1; Figure S14). One colony had mutations in both copies of ME element bearing the G allele of rs953413 (G2; Figure S15A) and two colonies had all the alleles of the ME elements mutated (all; Figure S15B). In accordance with our previous results, the colonies with mutations in the predicted FOXA-binding sites (mut) had significantly decreased expression of *ELOVL2* compared with the control group (control) after normalization with reference genes *RSP18*, *ACTB*, and *GAPDH*, respectively (Figures 6D and S16). Accordingly, the allelic imbalance in *ELOVL2* expression detected by rs2295601 AS-qPCR was significantly decreased in both group G1 and G2 compared with the control group in a dosage-dependent manner (Figure 6E). In addition, the two colonies grouped with all the copies of the

ME elements being mutated (all) had relatively more conserved FOXA binding sites in the ME elements bearing the G allele of rs953413 compared with that of the A allele of rs953413 that led to significantly increased allelic imbalance in *ELOVL2* expression (Figure S17). This suggested that the ME element encompassing rs953413 is a key regulatory region for *ELOVL2* expression.

The double gRNA strategy failed to directly mutate rs953413. We searched the rs953413 region again and designed allele-specific gRNA (AS-gRNA) that directly targets rs953413 to provide complementary evidence that rs953413 is essential for *ELOVL2* expression (Figure 6F). As rs953413 is the first nucleotide 3′ adjacent to PAM, which is part of the seed sequence essential for efficient gRNA binding, the designed AS-gRNA targeting rs953413 in theory should show high allelic preference in binding to the rs953413 region (Semenova et al., 2011). To prove our hypothesis, each AS-gRNA targeting rs953413 was coupled with the catalytically inactive Cas9 (dCas9) fused with either the activating VP64 domain (dCas9-VP64) or the suppressive KRAB domain (dCas9-KRAB) to be delivered into HepG2 cells by lentiviral transduction to test its effect on *ELOVL2* expression *in trans* (Ho et al., 2017) (Figure 6G). The AS-gRNA targeting the G allele of rs953413 (G-gRNA) potently suppressed *ELOVL2* expression when coupled with the suppressive dCas9-KRAB, whereas the *ELOVL2* expression was significantly increased by G-gRNA coupled with dCas9-VP64 after normalization with reference genes *RSP18*, *ACTB*, and *GAPDH*, respectively (Figures 6H and S18). The AS-gRNA targeting the A allele of rs953413 (A-gRNA) is less effective in regulating *ELOVL2* expression, which was only observed to significantly increase *ELOVL2* expression when coupled with dCas9-VP64 after normalization with *GAPDH* (Figure S18B). However, the rs2295601 AS-qPCR clearly showed that both AS-gRNA displayed high allelic preference in affecting *ELOVL2* expression (Figure 6I). When coupled with dCas9-KRAB, each AS-gRNA preferentially suppressed the transcription of *ELOVL2* located on the same chromosome with target allele of rs953413. While coupling with dCas9-VP64, the *ELOVL2* transcript in linkage with respective allele of rs953413 was selectively activated by AS-gRNA.

We next coupled each AS-gRNA with wild-type Cas9 to directly introduce mutations to rs953413 to determine its effect on *ELOVL2* expression in HepG2 cells. The allelic imbalance in mutating rs953413 by each AS-gRNA was determined with rs953413 AS-qPCR, which can specifically quantify the relative amount of mutations introduced into each allele of rs953413. In accordance with our previous observations, the A allele of rs953413 was preferentially mutated by A-gRNA coupled with wild-type Cas9, whereas the G allele of rs953413 was preferentially mutated by G-gRNA compared with cells transduced with lentiCRISP v2 virus as control (Figure 6J). The expression of *ELOVL2* was observed to be significantly downregulated by mutations introduced to rs953413 by each AS-gRNA after normalization with reference genes *RSP18*, *ACTB*, and *GAPDH*, respectively (Figures 6K and S19). Accordingly, preferentially introducing mutations to either allele of rs953413 significantly decreased the expression of its linked *ELOVL2* transcript determined by rs2295601 AS-qPCR (Figure 6L). These observations together with the results from double gRNA clearly demonstrated that rs953413 determines the enhancer activity of the identified ME element which further regulates *ELOVL2* expression.

## DISCUSSION

We here demonstrate that rs953413 mediates the cooperative binding of HNF4α and FOXA1/FOXA2 to the evolutionarily conserved enhancer region, which acts as a key cis-regulatory element for *ELOVL2* expression. The G allele of rs953413 is preferentially bound by the complex formed by these TFs and increases the expression of *ELOVL2* that further increases the amount of LC-PUFAs synthesized (Figure 6M).

As members of the hepatocyte nuclear factor family, *HNF4α* and *FOXA1/FOXA2* are all essential TFs for normal liver function and their expression profile also showed high tissue specificities with a preference for the liver, pancreas, and kidney (Lau et al., 2018). In accordance with the enhancer activity tightly regulated by the cooperation between FOXA1/FOXA2 and HNF4α, the rs953413 region showed tissue-specific enhancer features with both active chromatin marker H3K27ac and different TFs enriched mainly in liver HepG2 cells (Figure S4). As liver is the most important organ for LC-PUFAs synthesis, the rs953413 region is supposed to be a key enhancer region for *ELOVL2* regulation in the liver, which further determines the systemic LC-PUFAs profiles. In accordance with the G allele of rs953413 upregulating *ELOVL2* expression in our study, SNPs in LD with the G allele of rs953413 were generally associated with a more efficient conversion of DHA from its precursors in GWAS (Draisma et al., 2015; Illig et al., 2009; Lemaitre et al., 2011; Li et al., 2018; Suhre et al., 2011). The expression of *ELOVL2* in other tissues might be regulated by different mechanisms other than the rs953413 region illustrated by the observation that the A allele of rs953413 was associated with higher expression of *ELOVL2* in transformed fibroblasts ($p = 1.5 \times 10^{-6}$) from the GTEx project

(Figure S5) (Battle et al., 2017). Additionally, it should be noted that an independent association between SNPs indexed by rs2281591 in the *ELOVL2* locus with DPA levels was observed in conditional analyses in both European- and Chinese-ancestry populations (Hu et al., 2016; Lemaitre et al., 2011). As these SNPs are not in LD with rs953413, this independent association signal suggests that there are other functional variants in this locus involved in *ELOVL2* regulation.

Previous studies have shown that ω-3 PUFAs, particularly DHA, may influence human health by exerting beneficial effects on many diseases such as cardiovascular disease, diabetes, cancer, depression, nonalcoholic fatty liver disease (NAFLD), and rheumatoid arthritis (Jump et al., 2018; Lopez-Vicario et al., 2014; Shahidi and Ambigaipalan, 2018). Recently, rs2236212, an intronic SNP of *ELOVL2*, was shown to be associated with NALFD in obese subjects (Zusi et al., 2019). Patients and animal models of nonalcoholic steatohepatitis (NASH), a severe form of NAFLD, showed marked decrease in hepatic ω-3 PUFAs levels, which may play a role in the development and progression of NASH (Jump et al., 2018; Lopez-Vicario et al., 2014). Dietary intervention with ω-3 PUFAs DHA alone or EPA/DHA have shown indications of improvement in biomarkers related to NASH. The identified TFs in this study, including *FOXA1*/*FOXA2* and *HNF4α*, are tightly involved in NAFLD progression and are significantly downregulated in NAFLD (Lake et al., 2016; Moya et al., 2012; Wang et al., 2017; Weiss et al., 2017). This implies that downregulation of *ELOVL2* together with its upstream regulators *FOXA1*/*FOXA2* and *HNF4α* impairs hepatic DHA synthesis, which may play a role in the pathogenesis of NAFLD.

As a key enzyme in the *in vivo* synthesis of DHA, dysregulation of *ELOVL2* may also be involved in impairment of the systemic inflammatory process and diabetes progression (Bellini et al., 2018; Cruciani-Guglielmacci et al., 2017; Talamonti et al., 2017; Tikhonenko et al., 2010). However, further studies are needed to investigate the key TFs regulating *ELOVL2* in other cell types. It is interesting to note that *HNF4α*, *FOXA1*, and *FOXA2* have been extensively studied and appears to be involved in the progression of diabetes, and variations nearby *HNF4α* and *FOXA2* have been reported to be associated with diabetes related traits in GWAS (Lau et al., 2018; MacArthur et al., 2017). *ELOVL2* may be an important downstream target in diabetes caused by dysregulation of these TFs.

In conclusion, we show that rs953413 affects LC-PUFAs levels by altering *ELOVL2* expression through FOXA1/FOXA2 and HNF4α cooperation. The results provide important mechanistic insights to the transcriptional machinery regulating *ELOVL2* in the liver and thereby circulating levels of PUFAs. Further studies are needed to elucidate the role of this pathway in diseases, liver diseases in particular.

### Limitations of the Study

Due to lack of detailed genotype and phenotype data at individual level in reported GWAS in this locus, the exact effect of rs953413 and other variations conditional on rs953413 on the reported phenotypes cannot be determined. As the exonic SNP rs2295601 is homozygous in PLC/PRF/5 cells, the changes in allelic imbalance in *ELOVL2* expression upon modulating the rs953413 locus cannot be determined in this cell line.

## 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.2019.100808.

## ACKNOWLEDGMENTS

The study was supported by grants from AstraZeneca, The Swedish Diabetes Foundation (DIA 2015-064, 2017-269), The Family Ernfors Fund, and EXODIAB to CW and the Borgström Foundation to MC. Sequencing was performed at the SNP&SEQ platform of SciLifeLab, Uppsala. Funding for open access charge: The Swedish Diabetes Foundation (2017-269).

## AUTHOR CONTRIBUTIONS

C.W. conceived and oversaw this study. G.P. performed the experiments. G.P. and C.W. analyzed the data and wrote the manuscript with input from M.C., B.C., S.S., and C.K.. All authors approved the manuscript.

## DECLARATION OF INTERESTS

B.C., S.S., and C.K. are employees of AstraZeneca and may own stock or stock options. The remaining authors declare no competing interests.

Received: March 22, 2019
Revised: November 26, 2019
Accepted: December 23, 2019
Published: January 9, 2020

## REFERENCES

Battle, A., Brown, C.D., Engelhardt, B.E., and Montgomery, S.B. (2017). Genetic effects on gene expression across human tissues. Nature *550*, 204–213.

Bellini, L., Campana, M., Rouch, C., Chacinska, M., Bugliani, M., Meneyrol, K., Hainault, I., Lenoir, V., Denom, J., Veret, J., et al. (2018). Protective role of the ELOVL2/docosahexaenoic acid axis in glucolipotoxicity-induced apoptosis in rodent beta cells and human islets. Diabetologia *61*, 1780–1793.

Bhatt, D.L., Steg, P.G., Miller, M., Brinton, E.A., Jacobson, T.A., Ketchum, S.B., Doyle, R.T., Juliano, R.A., Jiao, L., Granowitz, C., et al. (2019). Cardiovascular risk reduction with icosapent ethyl for hypertriglyceridemia. N. Engl. J. Med. *380*, 11–22.

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

Cavalli, M., Pan, G., Nord, H., Wallerman, O., Wallen Arzt, E., Berggren, O., Elvers, I., Eloranta, M.L., Ronnblom, L., Lindblad Toh, K., et al. (2016b). Allele-specific transcription factor binding to common and rare variants associated with disease and gene expression. Hum. Genet. *135*, 485–497.

Cha, R.S., Zarbl, H., Keohavong, P., and Thilly, W.G. (1992). Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene. PCR Methods Appl. *2*, 14–20.

Chilton, F.H., Dutta, R., Reynolds, L.M., Sergeant, S., Mathias, R.A., and Seeds, M.C. (2017). Precision nutrition and omega-3 polyunsaturated fatty acids: a case for personalized supplementation approaches for the prevention and management of human diseases. Nutrients *9*, 1165.

Cruciani-Guglielmacci, C., Bellini, L., Denom, J., Oshima, M., Fernandez, N., Normandie-Levi, P., Berney, X.P., Kassis, N., Rouch, C., Dairou, J., et al. (2017). Molecular phenotyping of multiple mouse strains under metabolic challenge uncovers a role for Elovl2 in glucose-induced insulin secretion. Mol. Metab. *6*, 340–351.

Draisma, H.H.M., Pool, R., Kobl, M., Jansen, R., Petersen, A.-K., Vaarhorst, A.A.M., Yet, I., Haller, T., Demirkan, A., Esko, T., et al. (2015). Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels. Nat. Commun. *6*, 7208.

Fan, Y.Y., Monk, J.M., Hou, T.Y., Callway, E., Vincent, L., Weeks, B., Yang, P., and Chapkin, R.S. (2012). Characterization of an arachidonic acid-deficient (Fads1 knockout) mouse model. J. Lipid Res. *53*, 1287–1295.

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., et al. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database *2017*, 1–17.

Gregory, M.K., Gibson, R.A., Cook-Johnson, R.J., Cleland, L.G., and James, M.J. (2011). Elongase reactions as control points in long-chain polyunsaturated fatty acid synthesis. PLoS One *6*, e29662.

Ho, S.M., Hartley, B.J., Flaherty, E., Rajarajan, P., Abdelaal, R., Obiorah, I., Barretto, N., Muhammad, H., Phatnani, H.P., Akbarian, S., et al. (2017). Evaluating synthetic activation and repression of neuropsychiatric-related genes in hiPSC-derived NPCs, neurons, and astrocytes. Stem Cell Reports *9*, 615–628.

Hu, Y., Li, H.X., Lu, L., Manichaikul, A., Zhu, J.W., Chen, Y.D.I., Sun, L., Liang, S., Siscovick, D.S., Steffen, L.M., et al. (2016). Genome-wide meta-analyses identify novel loci associated with n-3 and n-6 polyunsaturated fatty acid levels in Chinese and European-ancestry populations. Hum. Mol. Genet. *25*, 1215–1224.

Illig, T., Gieger, C., Zhai, G., Römisch-Margl, W., Wang-Sattler, R., Prehn, C., Altmaier, E., Kastenmüller, G., Kato, B.S., Mewes, H.-W., et al. (2009). A genome-wide perspective of genetic variation in human metabolism. Nat. Genet. *42*, 137–141.

Jolma, A., Yin, Y., Nitta, K.R., Dave, K., Popov, A., Taipale, M., Enge, M., Kivioja, T., Morgunova, E., and Taipale, J. (2015). DNA-dependent formation of transcription factor pairs alters their binding specificity. Nature *527*, 384–388.

Jump, D.B., Lytle, K.A., Depner, C.M., and Tripathy, S. (2018). Omega-3 polyunsaturated fatty acids as a treatment strategy for nonalcoholic fatty liver disease. Pharmacol. Ther. *181*, 108–125.

Lake, A.D., Chaput, A.L., Novak, P., Cherrington, N.J., and Smith, C.L. (2016). Transcription factor binding site enrichment analysis predicts drivers of altered gene expression in nonalcoholic steatohepatitis. Biochem. Pharmacol. *122*, 62–71.

Lau, H.H., Ng, N.H.J., Loo, L.S.W., Jasmen, J.B., and Teo, A.K.K. (2018). The molecular functions of hepatocyte nuclear factors - in and beyond the liver. J. Hepatol. *68*, 1033–1048.

Lemaitre, R.N., Tanaka, T., Tang, W., Manichaikul, A., Foy, M., Kabagambe, E.K., Nettleton, J.A., King, I.B., Weng, L.-C., Bhattacharya, S., et al. (2011). Genetic loci associated with plasma phospholipid n-3 fatty acids: a meta-analysis of genome-wide association studies from the charge Consortium. PLoS Genet. *7*, e1002193.

Li, B., Kadura, I., Fu, D.-J., and Watson, D.E. (2004). Genotyping with TaqMAMA. Genomics *83*, 311–320.

Li, Y., Sekula, P., Wuttke, M., Wahrheit, J., Hausknecht, B., Schultheiss, U.T., Gronwald, W., Schlosser, P., Tucci, S., Ekici, A.B., et al. (2018). Genome-wide association studies of metabolites in patients with CKD identify multiple loci and illuminate tubular transport mechanisms. J. Am. Soc. Nephrol. *29*, 1513–1524.

Liang, G., He, H., Li, Y., and Yu, D. (2012). A new strategy for construction of artificial miRNA vectors in Arabidopsis. Planta *235*, 1421–1429.

Lopez-Terrada, D., Cheung, S.W., Finegold, M.J., and Knowles, B.B. (2009). Hep G2 is a hepatoblastoma-derived cell line. Hum. Pathol. *40*, 1512–1515.

Lopez-Vicario, C., Gonzalez-Periz, A., Rius, B., Moran-Salvador, E., Garcia-Alonso, V., Lozano, J.J., Bataller, R., Cofan, M., Kang, J.X., Arroyo, V., et al. (2014). Molecular interplay between Delta5/Delta6 desaturases and long-chain fatty acids in the pathogenesis of non-alcoholic steatohepatitis. Gut *63*, 344–355.

MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Junkins, H., McMahon, A., Milano, A., Morales, J., et al. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. *45*, D896–D901.

Machiela, M.J., and Chanock, S.J. (2015). LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics *31*, 3555–3557.

Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D.J., Buchman, S., Chen, C.Y., Chou, A., Ienasescu, H., et al. (2014). Jaspar 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res. *42*, D142–D147.

Moon, Y.-A., Hammer, R.E., and Horton, J.D. (2009). Deletion of ELOVL5 leads to fatty liver through activation of SREBP-1c in mice. J. Lipid Res. *50*, 412–423.

Moya, M., Benet, M., Guzmán, C., Tolosa, L., García-Monzón, C., Pareja, E., Castell, J.V., and Jover, R. (2012). Foxa1 reduces lipid accumulation in human hepatocytes and is down-regulated in nonalcoholic fatty liver. PLoS One *7*, e30014.

Nakamura, M.T., and Nara, T.Y. (2004). Structure, function, and dietary regulation of delta6, delta5, and delta9 desaturases. Annu. Rev. Nutr. *24*, 345–376.

Pauter, A.M., Olsson, P., Asadi, A., Herslof, B., Csikasz, R.I., Zadravec, D., and Jacobsson, A. (2014). Elovl2 ablation demonstrates that systemic DHA is endogenously produced and is essential for lipid homeostasis in mice. J. Lipid Res. *55*, 718–728.

Pauter, A.M., Trattner, S., Gonzalez-Bengtsson, A., Talamonti, E., Asadi, A., Dethlefsen, O., and Jacobsson, A. (2017). Both maternal and offspring Elovl2 genotypes determine systemic DHA levels in perinatal mice. J. Lipid Res. *58*, 111–123.

Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc. Natl. Acad. Sci. U S A *108*, 10098–10103.

Shahidi, F., and Ambigaipalan, P. (2018). Omega-3 polyunsaturated fatty acids and their health benefits. Annu. Rev. Food Sci. Technol. *9*, 345–381.

Stroud, C.K., Nara, T.Y., Roqueta-Rivera, M., Radlowski, E.C., Lawrence, P., Zhang, Y., Cho, B.H., Segre, M., Hess, R.A., Brenna, J.T., et al. (2009). Disruption of FADS2 gene in mice impairs male reproduction and causes dermal and intestinal ulceration. J. Lipid Res. *50*, 1870–1880.

Suhre, K., Shin, S.-Y., Petersen, A.-K., Mohney, R.P., Meredith, D., Wägele, B., Altmaier, E., CARDIoGRAM, Deloukas, P., Erdmann, J., et al. (2011). Human metabolic individuality in biomedical and pharmaceutical research. Nature *477*, 54–60.

Talamonti, E., Pauter, A.M., Asadi, A., Fischer, A.W., Chiurchiu, V., and Jacobsson, A. (2017). Impairment of systemic DHA synthesis affects macrophage plasticity and polarization: implications for DHA supplementation during inflammation. Cell Mol. Life Sci. *74*, 2815–2826.

The Encode Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Thomas-Chollier, M., Hufton, A., Heinig, M., O'Keeffe, S., Masri, N.E., Roider, H.G., Manke, T., and Vingron, M. (2011). Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. Nat. Protoc. *6*, 1860–1869.

Tikhonenko, M., Lydic, T.A., Wang, Y., Chen, W.Q., Opreanu, M., Sochacki, A., McSorley, K.M., Renis, R.L., Kern, T., Jump, D.B., et al. (2010). Remodeling of retinal fatty acids in an animal model of diabetes A decrease in long-chain polyunsaturated fatty acids is associated with a decrease in fatty acid elongases Elovl2 and Elovl4. Diabetes *59*, 219–227.

Wang, W., Yao, L.-J., Shen, W., Ding, K., Shi, P.-M., Chen, F., He, J., Ding, J., Zhang, X., and Xie, W.-F. (2017). FOXA2 alleviates CCl4-induced liver fibrosis by protecting hepatocytes in mice. Sci. Rep. *7*, 15532.

Weiss, T.S., Lupke, M., Ibrahim, S., Buechler, C., Lorenz, J., Ruemmele, P., Hofmann, U., Melter, M., and Dayoub, R. (2017). Attenuated lipotoxicity and apoptosis is linked to exogenous and endogenous augmenter of liver regeneration by different pathways. PLoS One *12*, e0184282.

Zadravec, D., Tvrdik, P., Guillou, H., Haslam, R., Kobayashi, T., Napier, J.A., Capecchi, M.R., and Jacobsson, A. (2011). ELOVL2 controls the level of n-6 28:5 and 30:5 fatty acids in testis, a prerequisite for male fertility and sperm maturation in mice. J. Lipid Res. *52*, 245–255.

Zhang, J.Y., Kothapalli, K.S.D., and Brenna, J.T. (2016). Desaturase and elongase-limiting endogenous long-chain polyunsaturated fatty acid biosynthesis. Curr. Opin. Clin. Nutr. Metab. Care *19*, 103–110.

Zhou, B., Ho, S.S., Greer, S.U., Spies, N., Bell, J.M., Zhang, X., Zhu, X., Arthur, J.G., Byeon, S., Pattni, R., et al. (2018). Haplotype-resolved and integrated genome analysis of ENCODE cell line HepG2. bioRxiv. https://doi.org/10.1101/378497.

Zusi, C., Mantovani, A., Olivieri, F., Morandi, A., Corradi, M., Miraglia Del Giudice, E., Dauriz, M., Valenti, L., Byrne, C.D., Targher, G., et al. (2019). Contribution of a genetic risk score to clinical prediction of hepatic steatosis in obese children and adolescents. Dig. Liver Dis. *51*, 1586–1592.

**iScience, Volume *23***

## Supplemental Information

# rs953413 Regulates Polyunsaturated Fatty Acid Metabolism by Modulating *ELOVL2* Expression

**Gang Pan, Marco Cavalli, Björn Carlsson, Stanko Skrtic, Chanchal Kumar, and Claes Wadelius**

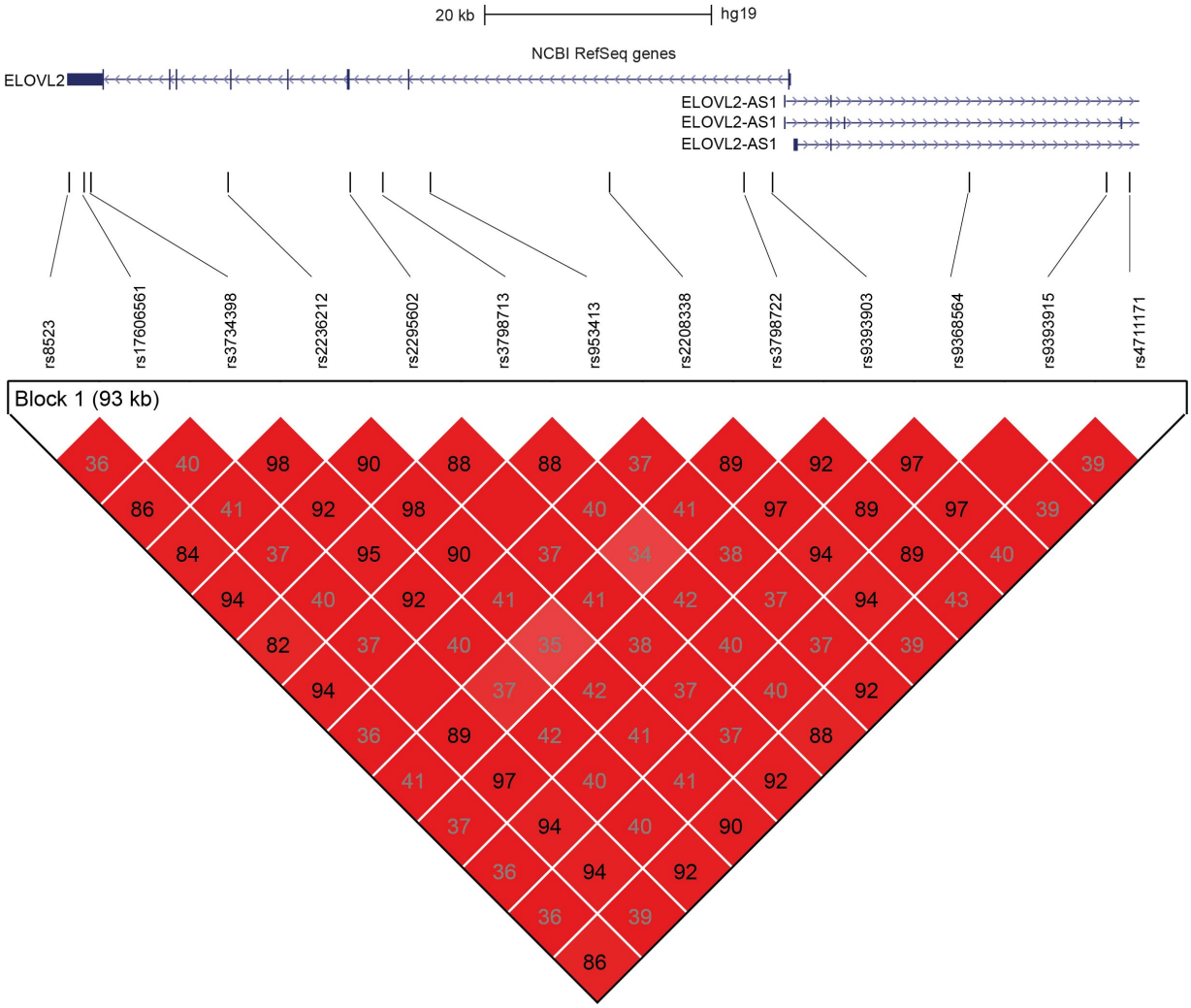



**Figure S1. The reported tag SNPs in the *ELOVL2* locus associated with LC-PUFAs derived metabolites in GWAS are in high LD.** Related to Figure 1. The tag SNPs are from the GWAS listed in Table S1. The color in each well denotes LOD value while the number denotes $r^2$ value between pair of SNPs in European population (CEU in 1000 Genomes Phase 3).

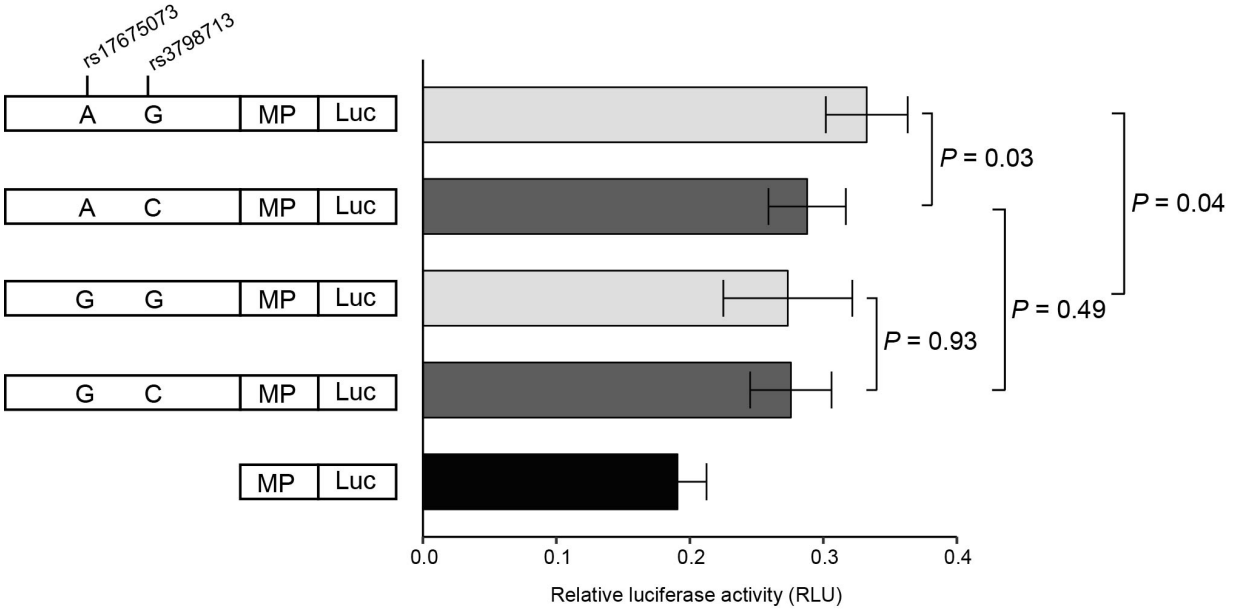



**Figure S2. Luciferase assay on the rs3798713 region.** Related to Figure 1. Both rs3798713 and rs17675073 are AS-SNPs in HepG2 cells that are in proximity. To evaluate the effect of each SNP, luciferase constructs containing different genotypes were constructed by direct PCR and site-directed mutagen with primers listed in Table S4. Error bars, s.d. $n = 6$ from two independent plasmid extractions and transfections with each transfection had three technical replicates. *P* values were calculated using two-tailed Student's *t* tests.

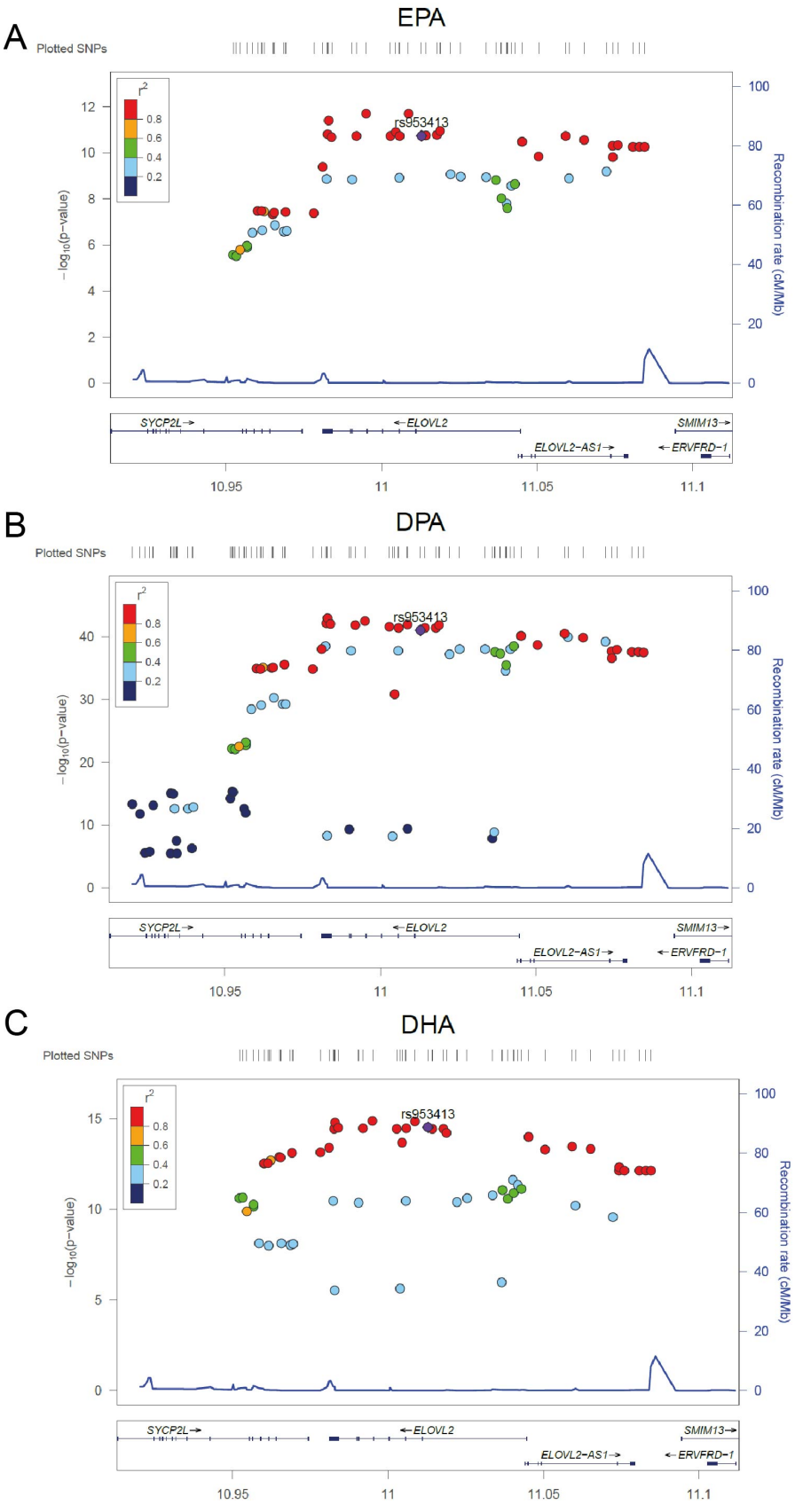

A
EPA
Plotted SNPs
rs953413
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
SYCP2L
ELOVL2
SMIM13
ELOVL2−AS1
ERVFRD−1
10.95
11
11.05
11.1
B
DPA
Plotted SNPs
rs953413
C
DHA
Plotted SNPs
rs953413

**Figure S3. rs953413 and its proxies showed more statistical powder in associating with levels of ω-3 fatty acids in plasma phospholipid.** Related to Figure 1. All the significant SNPs in the *ELOVL2* locus associated with EPA (A), DPA (B) and DHA (C) in the meta-analysis carried out in the CHARGE Consortium were plotted (Lemaitre et al., 2011). The plots were generated with rs953413 as the reference and LD is indicated by color scale in relation to rs953413 ($r^2$ values calculated in EUR in 1000 Genomes Phase 3).

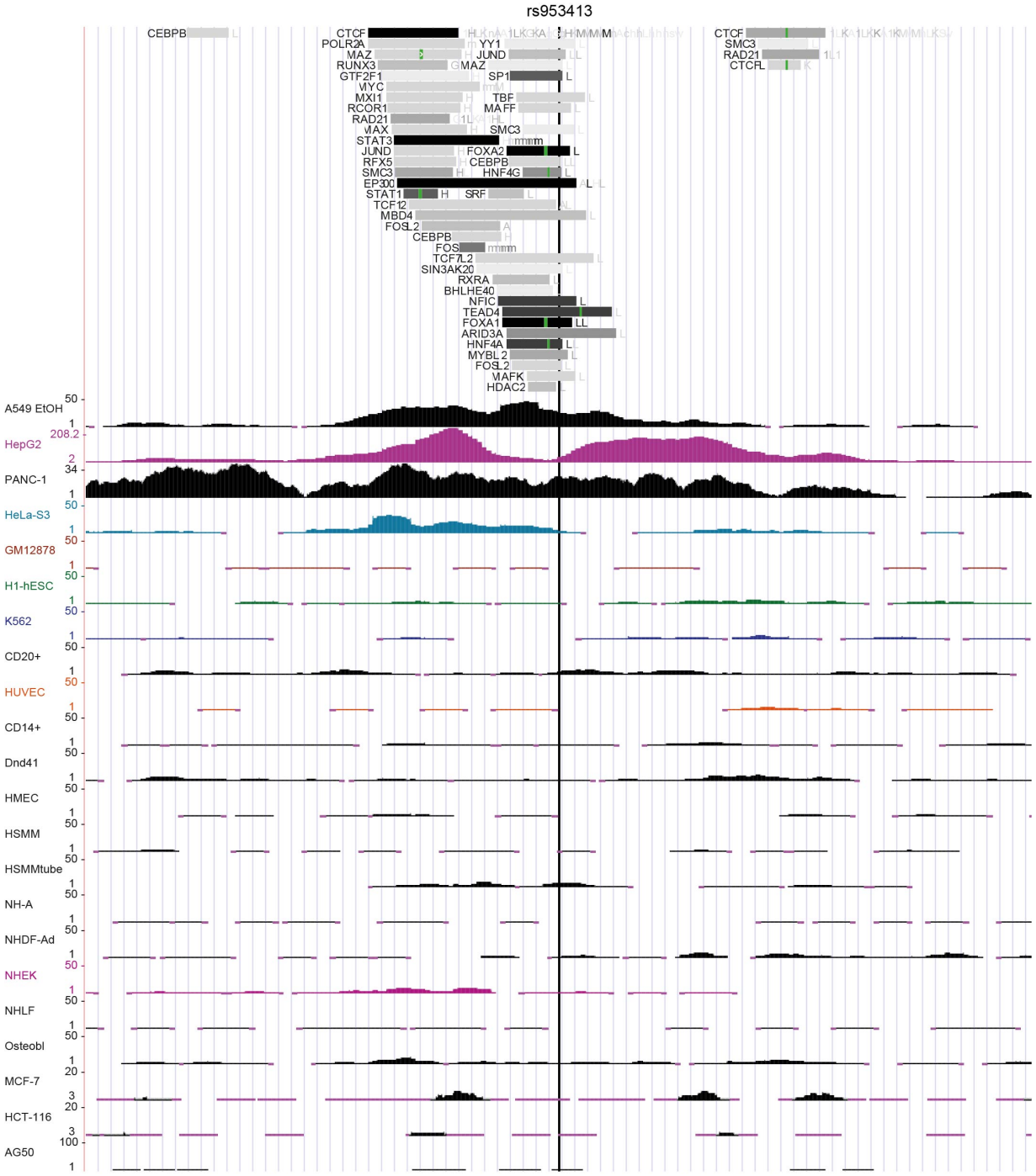


**Figure S4. rs953413 is located in a liver-specific enhancer region.** Related to Figure 1. A 5-kb window with rs953413 in the center is displayed here. The H3K27ac signals in HepG2 and other cell lines from the ENCODE project are shown for this region (The Encode Project Consortium, 2012). The exact location of rs953413 is highlighted by a vertical black line. The rs953413-containing region is also bound by many TFs in the liver which is identified by ChIP-seq experiments with antibodies against different TFs from the ENCODE project and is shown at the top.

| Tissue | Samples | NES | p-value | m-value |
|---|---|---|---|---|
| Colon - Sigmoid | 203 | 0.266 | 3.2e-5 | 1.00 |
| Cells - Transformed fibroblasts | 300 | 0.258 | 1.5e-6 | 1.00 |
| Esophagus - Gastroesophageal Junction | 213 | 0.226 | 2.1e-3 | 0.998 |
| Brain - Anterior cingulate cortex (BA24) | 109 | 0.223 | 1.1e-5 | 1.00 |
| Heart - Left Ventricle | 272 | 0.206 | 3.2e-4 | 1.00 |
| Cells - EBV-transformed lymphocytes | 117 | 0.204 | 0.08 | 0.911 |
| Brain - Caudate (basal ganglia) | 144 | 0.189 | 4.0e-4 | 1.00 |
| Brain - Cerebellum | 154 | 0.183 | 3.2e-5 | 1.00 |
| Brain - Cortex | 136 | 0.180 | 3.8e-5 | 1.00 |
| Esophagus - Muscularis | 335 | 0.164 | 1.2e-3 | 1.00 |
| Brain - Amygdala | 88 | 0.163 | 0.02 | 0.950 |
| Pancreas | 220 | 0.160 | 0.02 | 0.956 |
| Brain - Frontal Cortex (BA9) | 118 | 0.157 | 1.0e-3 | 1.00 |
| Ovary | 122 | 0.154 | 0.2 | 0.792 |
| Brain - Nucleus accumbens (basal ganglia) | 130 | 0.154 | 4.7e-3 | 0.991 |
| Prostate | 132 | 0.137 | 0.07 | 0.942 |
| Thyroid | 399 | 0.132 | 0.01 | 0.985 |
| Testis | 225 | 0.123 | 0.05 | 0.944 |
| Muscle - Skeletal | 491 | 0.121 | 0.02 | 0.961 |
| Brain - Cerebellar Hemisphere | 125 | 0.114 | 0.01 | 0.947 |
| Brain - Hippocampus | 111 | 0.113 | 0.06 | 0.844 |
| Minor Salivary Gland | 85 | 0.110 | 0.4 | 0.682 |
| Nerve - Tibial | 361 | 0.100 | 0.03 | 0.940 |
| Brain - Putamen (basal ganglia) | 111 | 0.0995 | 0.1 | 0.832 |
| Adrenal Gland | 175 | 0.0985 | 0.4 | 0.611 |
| Brain - Spinal cord (cervical c-1) | 83 | 0.0915 | 0.3 | 0.764 |
| Esophagus - Mucosa | 358 | 0.0757 | 0.1 | 0.672 |
| Small Intestine - Terminal Ileum | 122 | 0.0744 | 0.4 | 0.752 |
| Adipose - Subcutaneous | 385 | 0.0699 | 0.1 | 0.529 |
| Artery - Coronary | 152 | 0.0694 | 0.3 | 0.602 |
| Skin - Not Sun Exposed (Suprapubic) | 335 | 0.0588 | 0.4 | 0.431 |
| Colon - Transverse | 246 | 0.0501 | 0.3 | 0.274 |
| Brain - Hypothalamus | 108 | 0.0463 | 0.6 | 0.447 |
| Skin - Sun Exposed (Lower leg) | 414 | 0.0265 | 0.6 | 0.145 |
| Brain - Substantia nigra | 80 | 0.0252 | 0.8 | 0.517 |
| Uterus | 101 | 0.0250 | 0.9 | 0.682 |
| Artery - Tibial | 388 | 0.0146 | 0.7 | 0.00200 |
| Stomach | 237 | 0.0143 | 0.8 | 0.0830 |
| Heart - Atrial Appendage | 264 | 0.00896 | 0.9 | 0.135 |
| Pituitary | 157 | 0.00159 | 1 | 0.162 |
| Whole Blood | 369 | - | - | - |
| Spleen | 146 | - | - | - |
| Adipose - Visceral (Omentum) | 313 | -0.00256 | 1 | 0.126 |
| Breast - Mammary Tissue | 251 | -0.0178 | 0.7 | 0.00500 |
| Lung | 383 | -0.0270 | 0.6 | 0.0100 |
| Artery - Aorta | 267 | -0.0371 | 0.5 | 0.00800 |
| Vagina | 106 | -0.0636 | 0.5 | 0.143 |
| Liver | 153 | -0.212 | 1.4e-4 | 0.00 |

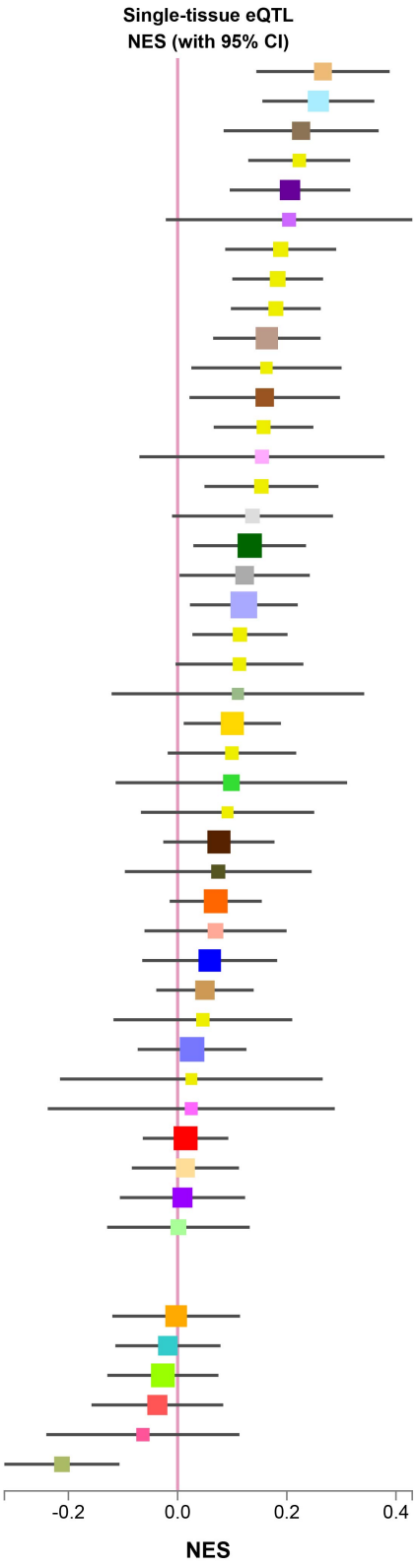


**Figure S5. The eQTL analysis of rs953413 on *ELOVL2* expression in multiple tissues from the GTEx portal.** Related to Figure 1. The figure was directly downloaded from the GTEx portal with minor modifications (Battle et al., 2017). NES, normalized effect size. In the liver tissue, the A allele of rs953413 is associated with decreased expression of *ELOVL2* with NES of -0.212.

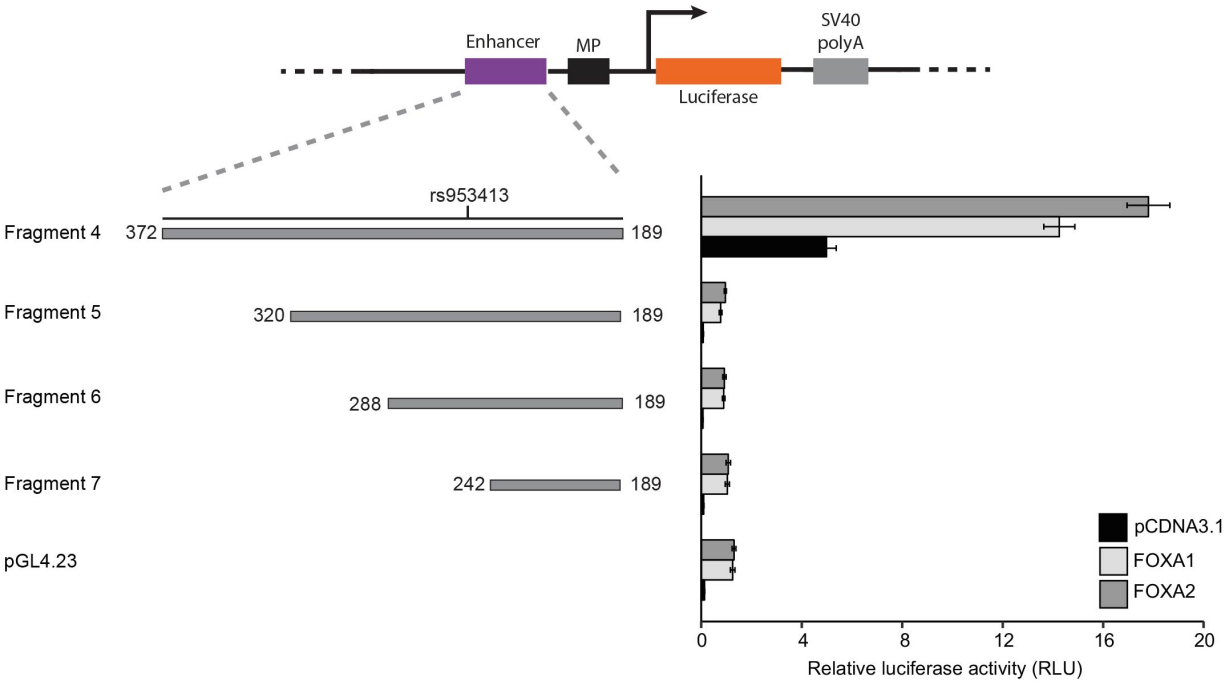



**Figure S6. The whole ME element is essential for the enhancer activity of the rs953413 region induced by FOXA1 and FOXA2.** Related to Figure 3. Only luciferase construct contains the whole ME element (Fragment 4) could be induced by FOXA1 and FOXA2 overexpression. Error bars, s.d. *n* = 6 from two independent plasmid extractions and transfections with each transfection had three technical replicates.

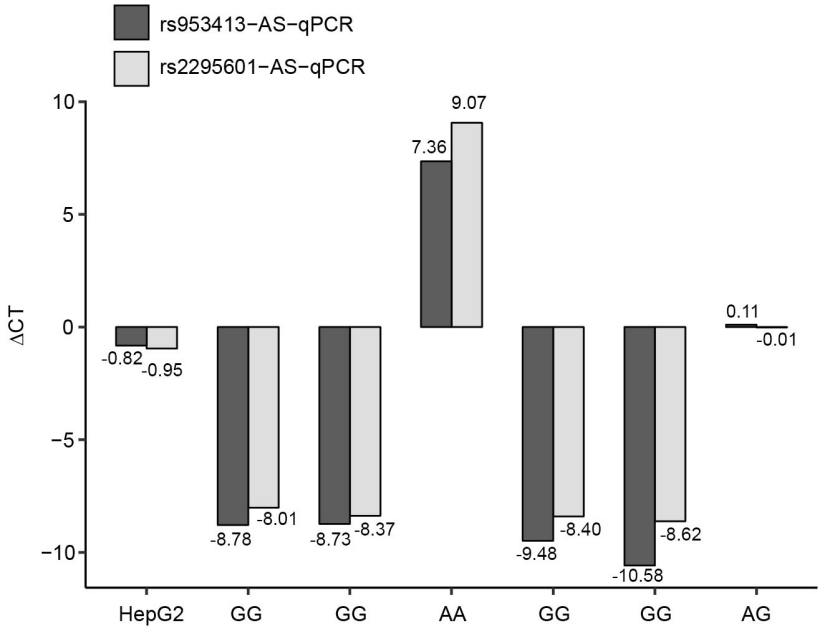



**Figure S7. Validation of AS-qPCR primers for rs953413 and rs2295601 with genomic DNA from different samples as templates.** Related to Figure 3-6. The *y* axis shows the $C_T$ value differences in qPCR with different genomic DNA as templates for both rs953413[G] versus rs953413[A] (dark grey bar) and rs2295601[G] versus rs2295601[A] (light grey bar). The exact $\Delta C_T$ values are displayed for each experiment. The determined genotypes for both rs953413 and rs2295601 are the same in each sample and are listed on *x* axis. Due to existence of two copies of the G allele in HepG2 cells, the $\Delta C_T$ values in HepG2 cells for both SNPs are different from the normal heterozygous sample.

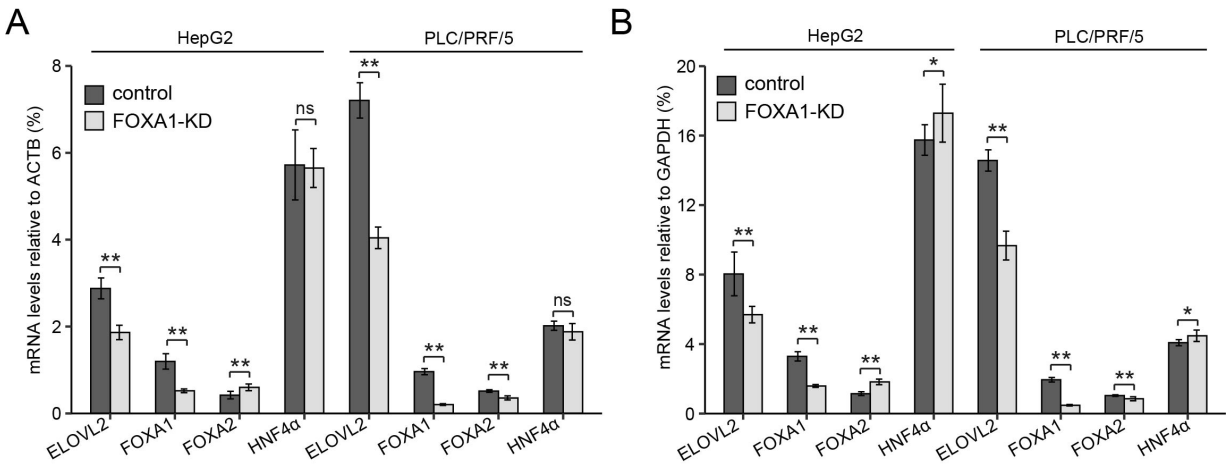



**Figure S8. The expression of *ELOVL2* is significantly decreased by *FOXA1* knockdown.** Related to Figure 3. The relative expression of each target gene was normalized with the expression of *ACTB* (A) and *GAPDH* (B). **P* < 0.05; ***P* < 0.01 and ns, not significant calculated by two-tailed Student's *t* tests. Error bars, s.d. *n* = 8 technical replicates from two independent experiments with each experiment had four replicates.

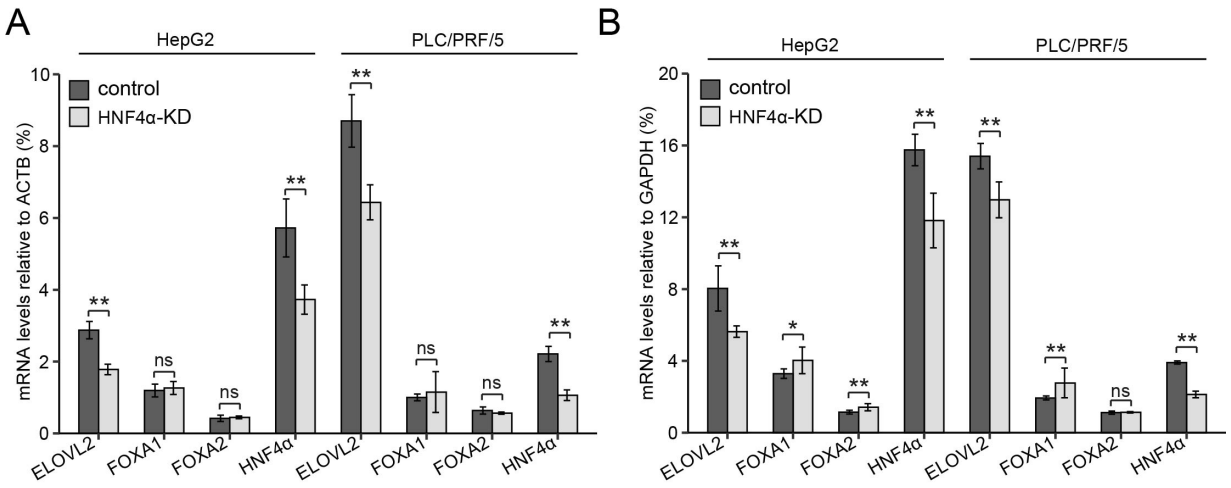



**Figure S9. The expression of *ELOVL2* is significantly decreased by *HNF4α* knockdown.** Related to Figure 4. The relative expression of each detected gene was normalized with the expression of *ACTB* (A) and *GAPDH* (B). **P* < 0.05; ***P* < 0.01 and ns, not significant calculated by two-tailed Student's *t* tests. Error bars, s.d. *n* = 8 technical replicates from two independent experiments with each experiment had four replicates.

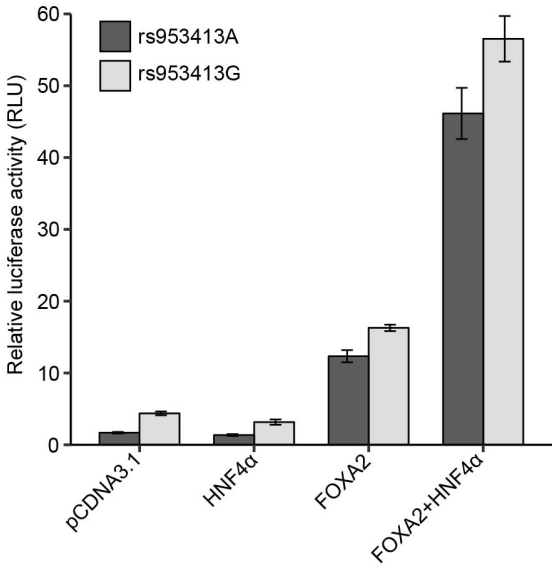



**Figure S10. Cooperation between HNF4α and FOXA2 significantly increased the enhancer activity of the rs953413 region in luciferase assay.** Related to Figure 5. The enhancer activity of the rs953413 region is significantly increased by simultaneous overexpression of HNF4α and FOXA2 compared with overexpression only HNF4α or FOXA2 in luciferase assay. Error bars, s.d. *n* = 6 technical replicates from two independent plasmid extractions and transfections with each transfection had three technical replicates.

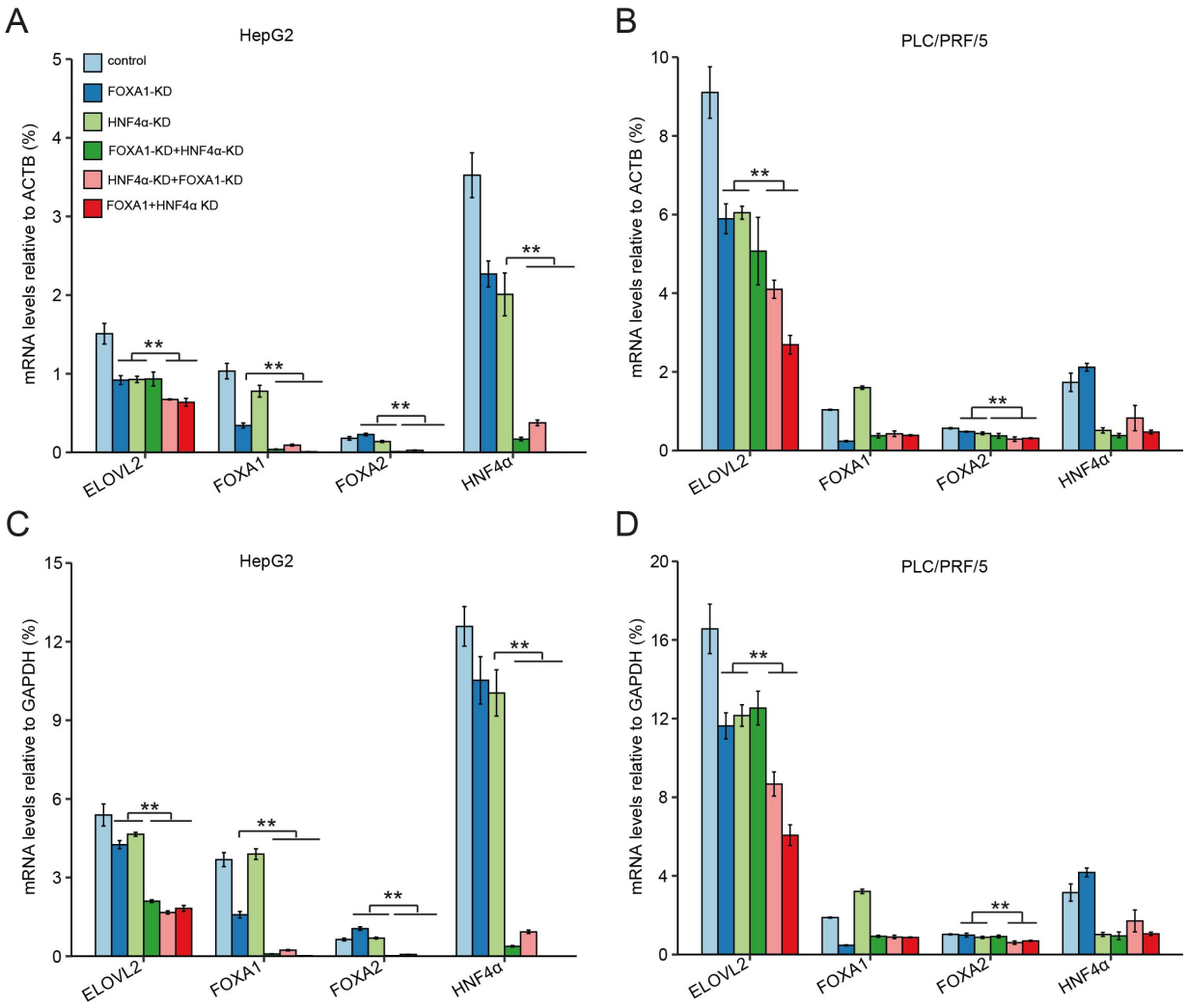



**Figure S11. Double knockdown of both *HNF4α* and *FOXA1* leads to a significant decrease in *ELOVL2* expression compared with knockdown only *HNF4α* or *FOXA1*.** Related to Figure 5. The relative expression of each gene is normalized to the expression of *ACTB* (A, B) and *GAPDH* (C, D) in both HepG2 and PLC/PRF/5 cells. The double knockdown cell lines were generated by either sequential or simultaneous lentivirus transduction. FOXA1-KD+HNF4α-KD denotes the cell line first transduced with the virus for *FOXA1* knockdown and then transduced the virus for *HNF4α* knockdown and vice versa for HNF4α-KD+FOXA1-KD. FOXA1+HNF4α KD denotes the cell line with simultaneously transduced virus for both *HNF4α* and *FOXA1* knockdown. $^{**}P < 0.01$ calculated by two-tailed Student's *t* tests. Error bars, s.d. $n = 4$ technical replicates.

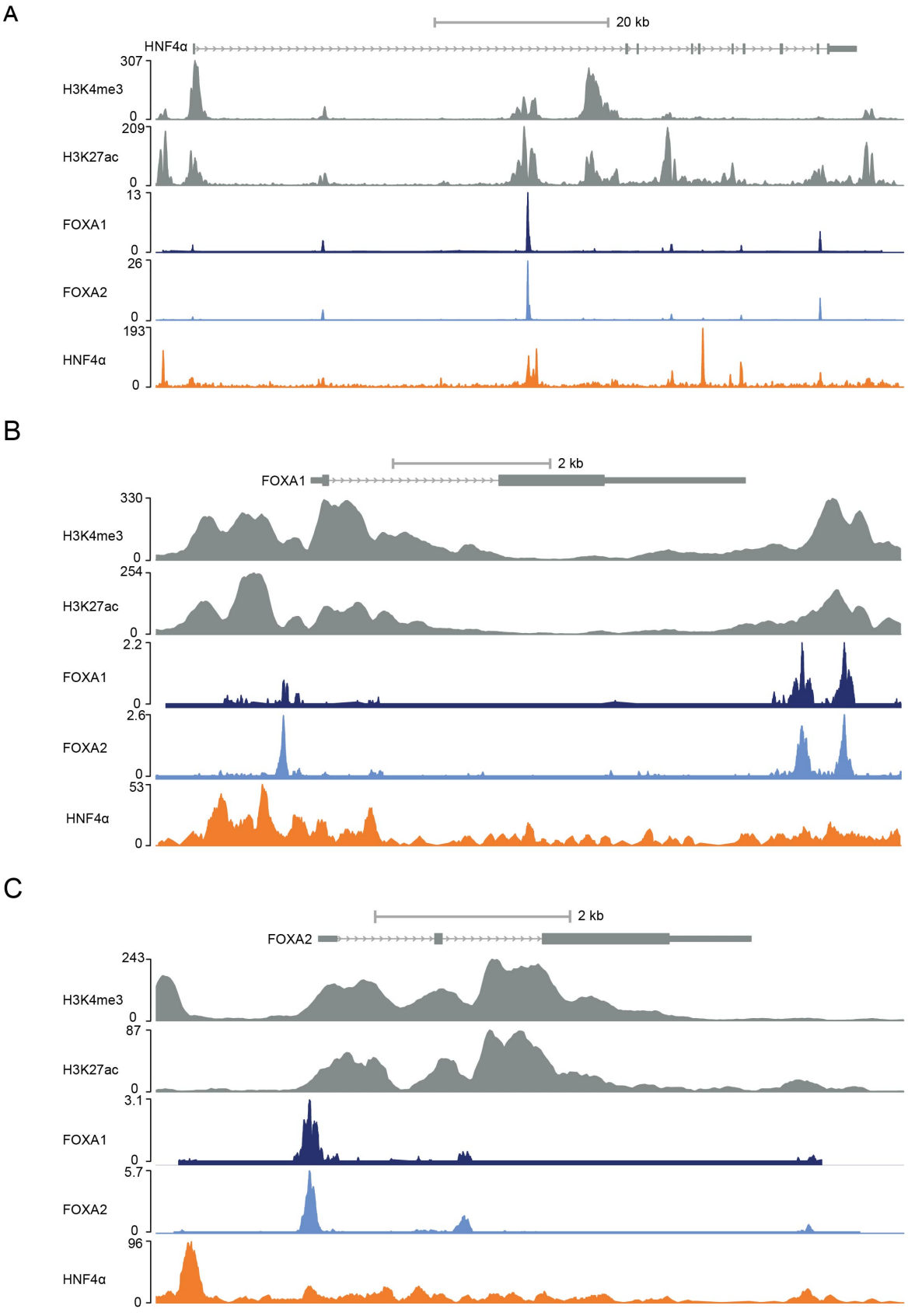

A
20 kb
HNF4α
307
H3K4me3
0
209
H3K27ac
0
13
FOXA1
0
26
FOXA2
0
193
HNF4α
0
B
2 kb
FOXA1
330
H3K4me3
0
254
H3K27ac
0
2.2
FOXA1
0
2.6
FOXA2
0
53
HNF4α
0
C
2 kb
FOXA2
243
H3K4me3
0
87
H3K27ac
0
3.1
FOXA1
0
5.7
FOXA2
0
96
HNF4α
0

**Figure S12. ChIP-seq signals of HNF4α, FOXA1 and FOXA2 from the ENCODE project are all enriched over the gene body of *HNF4α*, *FOXA1* and *FOXA2*, respectively in HepG2 cells.** Related to Figure 5. The ChIP-seq signals of HNF4α, FOXA1 and FOXA2 together with H3K27ac and H3K4me3 signals in HepG2 cells are displayed over the gene body of *HNF4α* (A), *FOXA1* (B) and *FOXA2* (C).

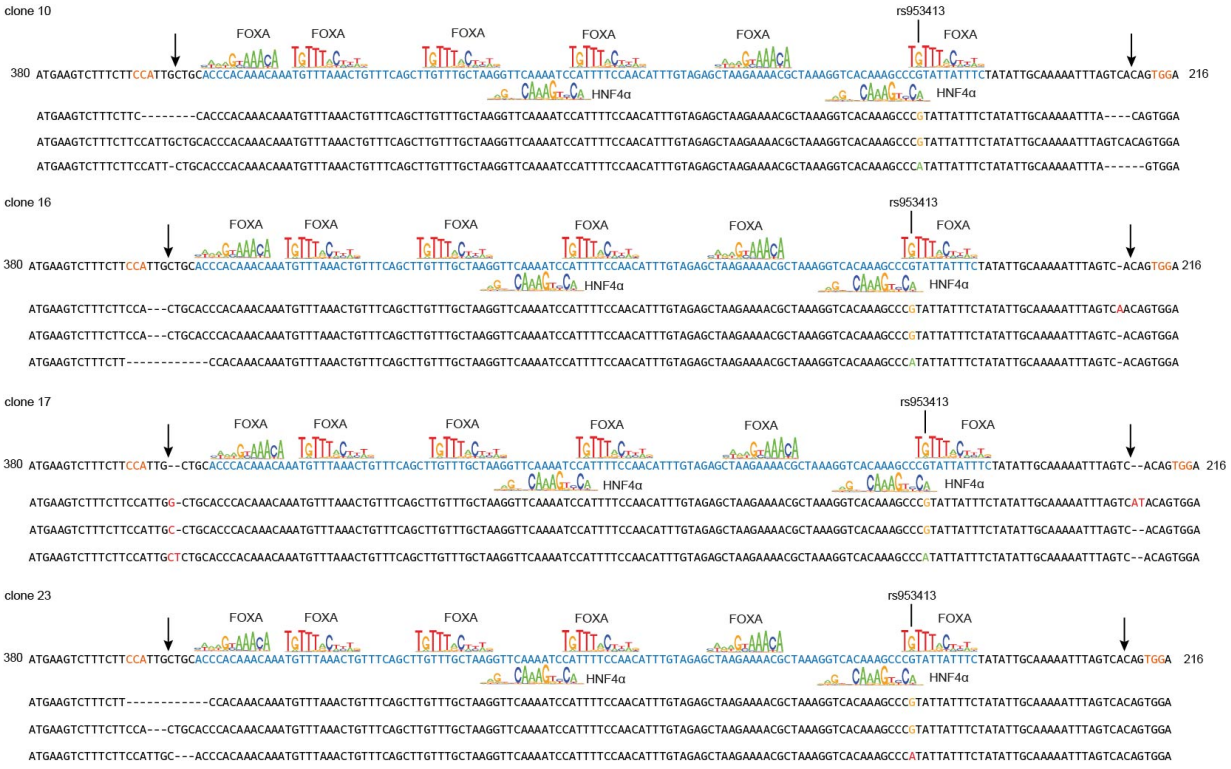



**Figure S13. Individual clones bearing no mutations to the ME element identified by Sanger sequencing.** Related to Figure 6. For each clone, genomic DNA was PCR-amplified and subjected to both direct PCR product sequencing and TA cloning sequencing. The ME element is highlighted in blue. The arrow denotes the expected Cas9 cut sites. Base-pair changes are shown in red and deletions are represented by dashes.


**Figure S14. Individual clones with mutations introduced to one copy of the ME element bearing the G allele of rs953413 identified by Sanger sequencing.** Related to Figure 6. For each clone, genomic DNA was PCR-amplified and subjected to both direct PCR product sequencing and TA cloning sequencing. The ME element is highlighted in blue. The arrow denotes the expected Cas9 cut sites. Base-pair changes are shown in red and deletions are represented by dashes.

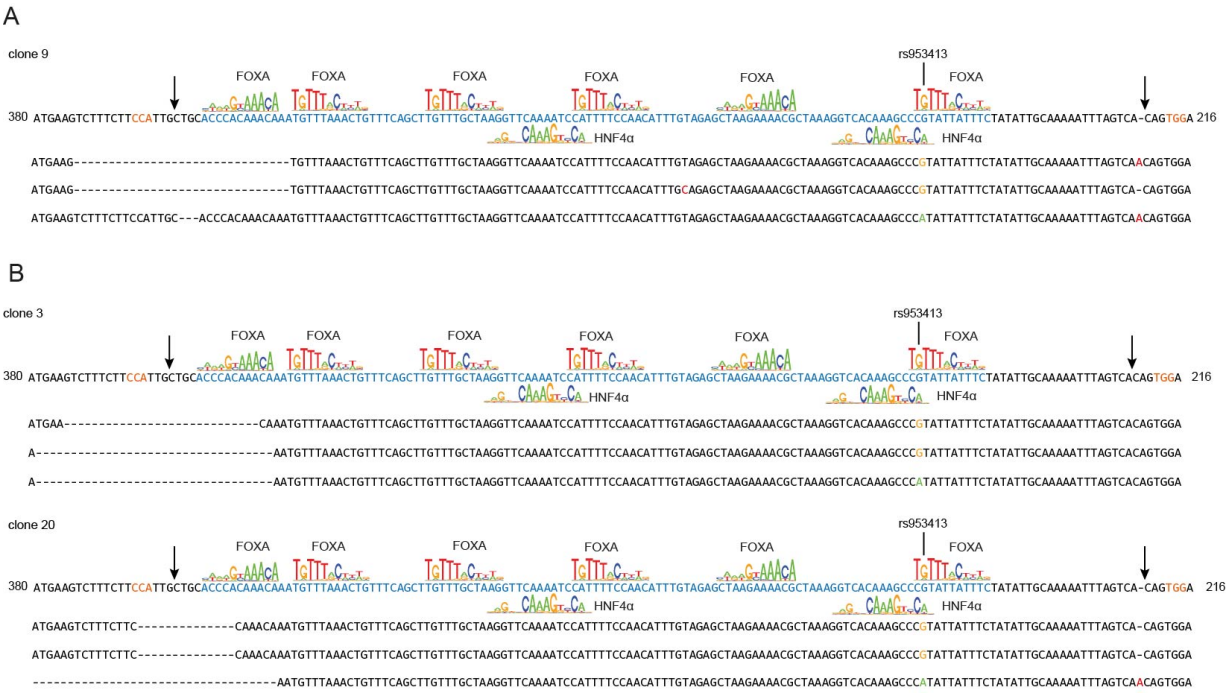



**Figure S15. Individual clones with mutations introduced to multiple copies of the ME element identified by Sanger sequencing.** Related to Figure 6. (A) clone 9 introduced mutations to both copies of the ME element bearing the G allele of rs953413. (B) clone 3 and clone 20 are verified to have mutations to all copies of the ME element. For each clone, genomic DNA was PCR-amplified and subjected to both direct PCR product sequencing and TA cloning sequencing. The ME element is highlighted in blue. The arrow denotes the expected Cas9 cut sites. Base-pair changes are shown in red and deletions are represented by dashes.

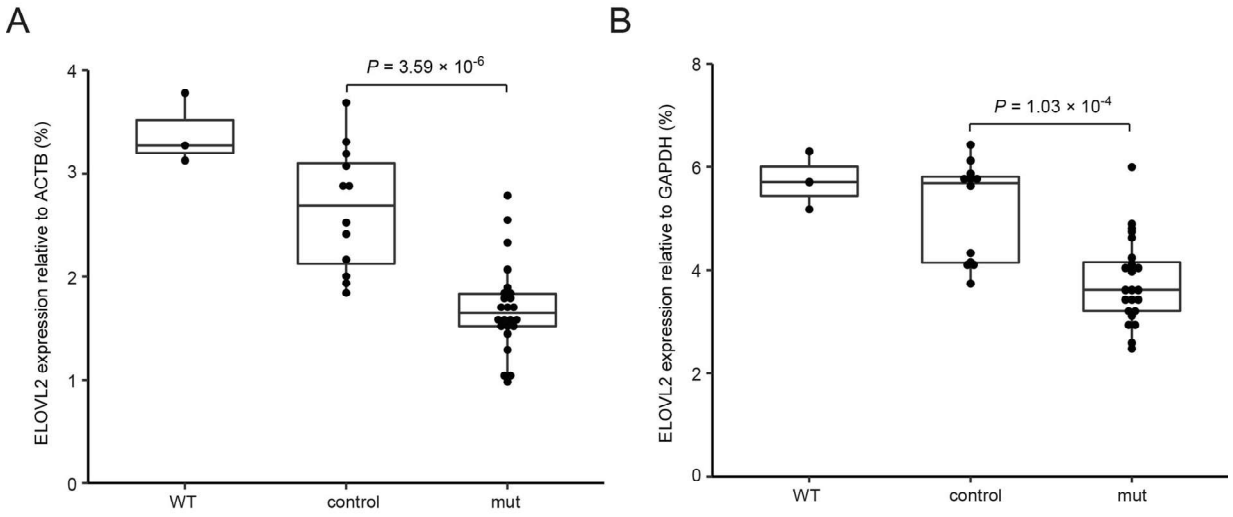



**Figure S16. The clones bearing mutations in the ME element (mut) had impaired *ELOVL2* expression compared with clones without mutations introduced in the ME element (control).** Related to Figure 6. The expression of *ELOVL2* is either normalized to the expression of *ACTB* (A) or *GAPDH* (B). Each clone has three technical replicates. *P* values were evaluated using two-tailed Student's *t* tests. WT denotes *ELOVL2* expression from normal HepG2 cells (*n*=3).

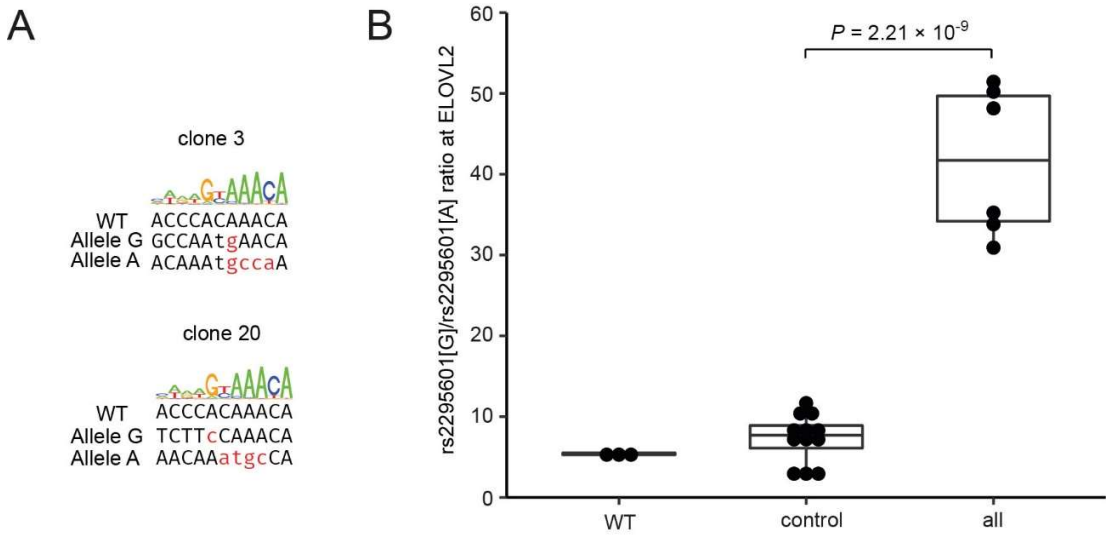



**Figure S17. Increased allelic imbalance in *ELOVL2* expression in clone 3 and clone 20.** Related to Figure 6**.** (A) In both clone 3 and clone 20, the ME element of each allele is mutated and the ME elements bearing the G allele of rs953413 have a more conserved FOXA motif compared to the A allele (see also Figure S15B). WT represents the wild type sequence while allele G and allele A represent the mutated ME sequences in linkage with rs953413[G] and rs953413[A], respectively. The mutated key motif sequences are displayed in lower-case letters and highlighted in red. (B) The increased allelic imbalance in *ELOVL2*

expression is detected by rs2295601 AS-qPCR. A significant increase in allelic imbalance of *ELVOL2* expression is observed in clone 3 and clone 20 (all) when compared with clones without mutations introduced to the FOXA motif (control, as listed in Figure S13). WT represents the observed allelic imbalance in *ELOVL2* expression in normal HepG2 cells (*n*=3).

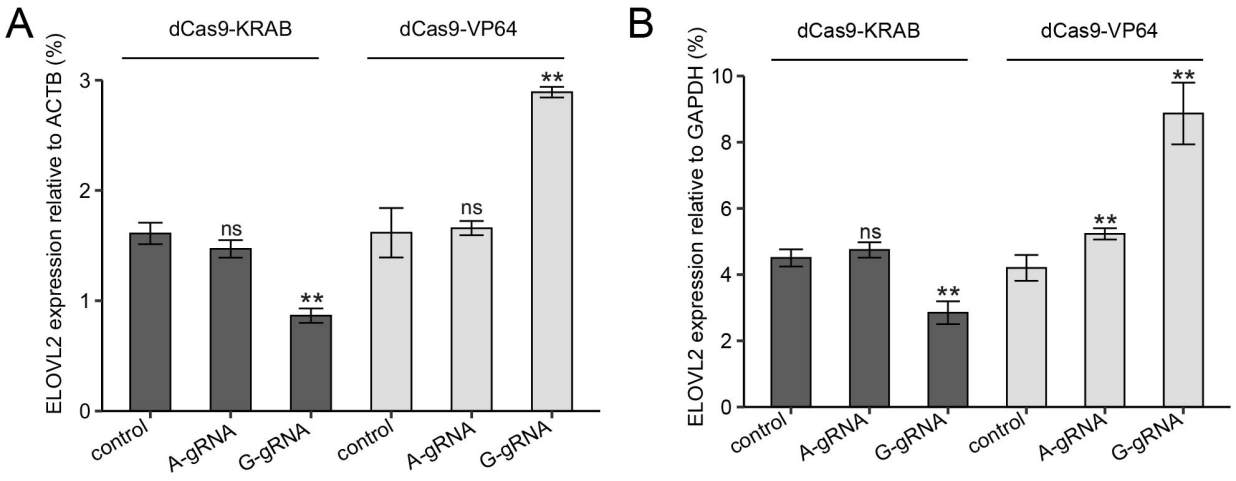



**Figure S18. The effect of AS-gRNA targeting rs953413 coupled with dCas9-effector on *ELOVL2* expression.** Related to Figure 6. The expression of *ELOVL2* was normalized to the expression of *ACTB* (A) and *GAPDH* (B), respectively. HepG2 cells transduced with only dCas9-effector virus was employed as the control. ***P* < 0.01 and ns, not significant calculated by two-tailed Student's *t* tests. Error bars, s.d. *n* = 4 technical replicates.

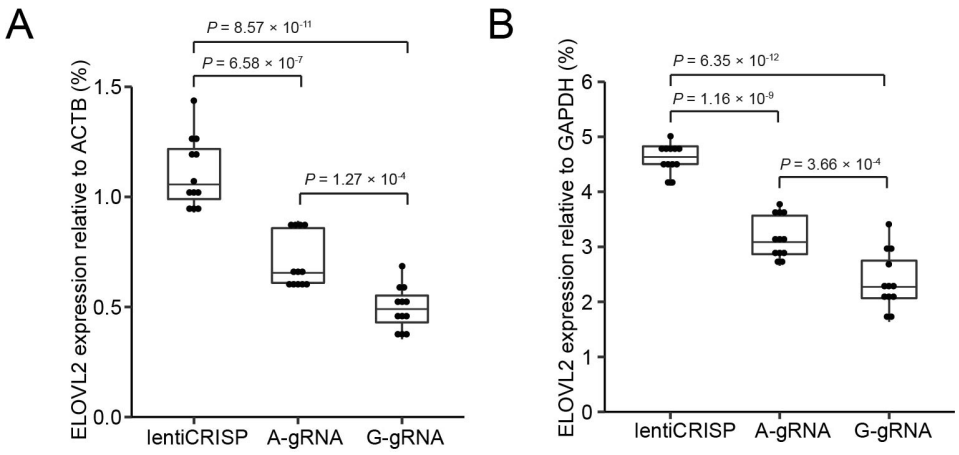



**Figure S19. The expression of *ELOVL2* was significantly decreased by mutations introduced to rs953413 by AS-gRNA targeting rs953413 coupled with wide type Cas9.** Related to Figure 6. The

expression of *ELOVL2* was normalized to the expression of *ACTB* (A) and *GAPDH* (B). The expression of *ELOVL2* in HepG2 cells transduced with lentiCRISP v2 virus was used as the control. *n* = 12 technical replicates. The *P* values were calculated by two-tailed Student's *t* tests.

## Supplemental Tables

**Table S3. AS-SNPs in the *ELOVL2* locus**. Related to Figure 1.

| SNP | allele | AF (EUR) | FOXA1 (sc101058) | | FOXA1 (sc6553) | | FOXA2 (sc6554) | |
|---|---|---|---|---|---|---|---|---|
| | | | Rep 1 | Rep 2 | Rep1 | Rep2 | Rep1 | Rep2 |
| rs953413 | rs953413[G] | 0.56 | 22 | 12 | 15 | 15 | 4 | 28 |
| | rs953413[A] | 0.44 | 5 | 4 | 14 | 4 | 1 | 9 |
| rs3798713 | rs3798713[G] | 0.58 | 10 | 11 | 17 | 4 | NA | 16 |
| | rs3798713[C] | 0.42 | 2 | 1 | 5 | 3 | NA | 2 |
| rs17675073 | rs17675073[G] | 0.86 | 0 | 2 | 3 | 1 | NA | 2 |
| | rs17675073[A] | 0.14 | 9 | 9 | 13 | 2 | NA | 4 |

Note: The number of uniquely mapped reads covering each allele in ChIP-seq was reported for each AS-SNP. NA, no reads covered. AF, allele frequency in European population.

## Transparent Methods

### Allelic imbalance in chromatin binding of the rs953413 and rs3798713 region in ChIP-seq experiments

The bam files for each TF mapped to regions where AS-SNPs are located were acquired from the ENCODE project (The Encode Project Consortium, 2012). Duplicate reads were removed with picard MarkDuplicates (http://broadinstitute.github.io/picard). Aligned reads with mismatch other than the SNP were discarded with bamtools (Barnett et al., 2011). The reads mapped to each allele were separated with splitSNP (https://github.com/astatham/splitSNP) and counted.

### Expression plasmids construction

The human *FOXA1* gene was amplified from HepG2 cDNA and cloned into pCDNA3.1 through SLiCE cloning method (Zhang et al., 2012). The ORF clone of *FOXA2* was purchased from Sino Biological (HG14174-G). Plasmid pCDNA3.1-FOXA2 was constructed by inserting the coding region of *FOXA2* into pCDNA3.1 digested with HindIII and EcoRV through SLiCE cloning method (Zhang et al., 2012). FR_HNF4A2 was a gift from Gerhart Ryffel (Addgene plasmid # 31100) (Thomas et al., 2004). Plasmid

pCDNA3.1-HNF4α was constructed by inserting the coding sequence of *HNF4α* isoform 2 digested with KpnI and NotI from FR_HNF4A2 into pCDNA3.1 digested with the same enzymes. All the plasmids are validated by Sanger sequencing. The detailed primer sequences for each expression construct is listed in Table S5.

**Luciferase constructs and reporter assay**

Plasmid pSV40 was constructed by transfer the SV40 promoter sequence from pGL3-promoter vector into pGL4.10 (Promega) through BglII and HindIII double digestion. The 287-bp (chr6: 11,008,440-11,008,726; hg19) fragment containing both rs3798713 and rs17675073 was used for enhancer validation. The 529-bp (chr6: 11,012,583-11,013,111; hg19) rs953413-centered region was also selected for enhancer validation based on the ChIP-seq signals for different TFs from the ENCODE project (Figure S4). The regions were amplified from HepG2 genomic DNA and inserted upstream of either the minimal promoter (MP) sequence of pGL4.23 (Promega) or the SV40 promoter sequence of pSV40 to test their enhancer activities. To identify the ME region encompassing rs953413, a series of truncation and deletion luciferase constructs were generated by SLiCE cloning method (Zhang et al., 2012). A three fragments In-fusion cloning system (Takara) with mutations introduced through primer sequences was employed to introduce desired mutations into luciferase constructs. The detailed primer information and cloning method for each construct is listed in Table S4. All the resulting plasmids were verified by Sanger sequencing.

HepG2 cells were plated one day before transfection in 96-well plates. The confluency was 50-70% on transfection. For normal luciferase assay, each well was transfected with 0.3 μl X-tremeGENE HP DNA transfection reagent (Roche) and 100 ng of experimental firefly luciferase reporter plasmid, and 1 ng of pGL4.74 renilla luciferase reporter vector as internal control for monitoring transfection and lysis efficiency. For luciferase assay overexpressing HNF4α in Figure 4A and 4B, each well was transfected with 0.3 μl X-tremeGENE HP DNA transfection reagent (Roche), 40 ng of firefly luciferase reporter plasmid, 50 ng pCDNA3.1-HNF4α and 10 ng of pRL-MP described earlier (Pan et al., 2017). For other cotranfection experiments, each well was transfected with 0.3 μl X-tremeGENE HP DNA transfection reagent (Roche), 50 ng of firefly luciferase reporter plasmid, 50 ng of expression plasmid and 1 ng of pGL4.74. For experiments cotransfecting HNF4α and FOXA1 in Figure 5A and cotransfecting HNF4α and FOXA2 in Figure S10, 25 ng of each expression plasmid was used with plasmid pCDNA3.1 used as control.

Cells were harvested 24 h after transfection and assayed with the Dual-Luciferase Reporter Assay System (Promega) on an Infinite M200 pro reader (Tecan). All the results are expressed directly as the ratio of firefly luciferase activity from experiment plasmids to renilla luciferase activity from control plasmids. All the luciferase experiments came from two independent transfections i.e., independent plasmid preparations and transfections each with three technical replicates. Luciferase values are expressed as

averages with error bars representing standard deviations (s.d) from all technical replicates and statistical analyses were performed by two-tailed Student's *t* tests.

**Cell culture**

HepG2 cells were originally purchased from the American Type Culture Collection (ATCC) and maintained in RPMI1640 basal medium supplemented with 10% fetal bovine serum (FBS) and 2 mM L-glutamine. Human PLC/PRF/5 cells were purchased from European Collection of Authenticated Cell Cultures (ECACC 85061113) and maintained in high glucose DMEM medium supplemented with 10% FBS, 1 mM sodium pyruvate. 293T cells were grown in DMEM supplemented with 10% FBS, 1 mM sodium pyruvate, and 500 μg/ml Geneticin. All the cells were also supplemented with 100 units of penicillin and 100 μg of streptomycin per 1 ml of culture medium.

**Lentiviral-mediated gene knockdown**

Both shRNA and artificial miRNA expression cassettes were utilized to knockdown target genes. The shRNA sequence was directly synthesized and inserted into pGreenPuro shRNA expression lentivector (SI505A-1, System Biosciences) digested with BamHI and EcoRI. The artificial miRNA precursor was constructed from the natural miR-124-1 precursor sequence introducing two suitable restriction sites recognized by EcoRV and EcoRI, respectively (Liang et al., 2012). The constructed artificial miRNA precursor was inserted downstream of the puromycin resistance gene through SLiCE cloning method. The target sequences for *FOXA1* knockdown are GATGTGTAGACATCCTCCGTATATT and GGCGTACTACCAAGGTGTGTA. For *HNF4α* knockdown, the target sequences are GAACCACATGTACTCCTGCAGATT and TCAGCACTCGAAGGTCAAGCTAT. The detailed primer information for each lentiviral construct is listed in Table S5.

Lentivirus was produced in 293T cells by transfecting the lentiviral plasmid together with packaging plasmids pLP1, pLP2 and envelope plasmid pLP/VSVG (Life Technologies) using polyethylenimine (Polysciences) following the manufacturer's instructions. Cells were plated in 24-well plates and transduced with virus supernatant together with sequabrene (Sigma) at final concentration of 8 μg/ml. The cell lines with double knockdown of *FOXA1* and *HNF4α* were achieved by sequential or simultaneous transduction of virus knocking down *FOXA1* and *HNF4α*. The cell lines with target gene knockdown were selected by puromycin (Life Technologies) at concentration of 1 μg/ml for HepG2 cells and 4 μg/ml for PLC/PRF/5 cells. The selected cells were maintained with 0.5 μg/ml of puromycin and regularly passaged for further analysis.

**Quantitative real time PCR**

At least three replicates were studies by quantitative real time PCR (qPCR) on cDNA samples and on genomic DNA samples from ChIP experiments. List of primers used in this study is shown in Table S6. The qPCR reactions were performed with JumpStart Taq ReadyMix (Sigma) coupled with EvaGreen dye (Biotium). For gene expression analysis, equal number of cells ($8 \times 10^5$ for HepG2 cells and $2.5 \times 10^5$ for PLCR/PRF/5 cells) were plated in 12-well plates 24 h before harvesting. Total RNA was extracted from cells with TRIzol (Life Technologies) according to the manufacturer's instructions. A total of 1 μg total RNA was reverse transcribed into cDNA with Maxima First Strand cDNA synthesis kit (Life Technologies). The expression was normalized to three control transcripts including *RSP18*, *ACTB* and *GAPDH*. Statistical analyses were carried out by two-tailed Student's *t* tests.

For ChIP experiments, a standard curve was generated for each target with serial-diluted input DNA as templates. The amounts of DNA pulled down for each target was first normalized to the input based on the standard curve. Then the relative enrichment of target DNA was displayed as fold changes over the negative control region GDCHR12, located in one gene desert region on chromosome 12. The promoter region of *HNF1α* (HNF1α-pro) was used as the positive control. All the primer sequences used for qPCR analysis are listed in Table S6.

**Allele-specific quantitative PCR**

Primers for AS-qPCR were designed based on mismatch amplification mutation assay (MAMA) to accurately quantify single nucleotide mutations (Cha et al., 1992; Li et al., 2004). Allele-specific amplification of the rs953413 region was employed to validate the allelic imbalance in chromatin binding of both TFs and H3K27ac in ChIP experiments. The exonic rs2295601 AS-qPCR was used to evaluate the allelic imbalance in *ELOVL2* expression. The PCR reactions were carried out with JumpStart Taq ReadyMix (Sigma) coupled with EvaGreen dye (Biotium). The specificity of the AS-qPCR primers was validated with genomic DNA from different samples as templates from the Excellence of Diabetes Research in Sweden (EXODIAB) biobank and shown in Figure S7. The detailed primer sequences used in AS-qPCR are listed in Table S6.

**Chromatin immunoprecipitation**

ChIP experiments were carried out as previously described with modifications (Blecher-Gonen et al., 2013; Pan et al., 2017). Cells were cultured in T175 flask until reaching confluency of >80% and crosslinked with 1% formaldehyde on a shaking platform for 10 min at room temperature and quenched with 125 mM glycine for 5 min. Cells were collected, washed twice with ice-cold PBS and resuspended in cell lysis buffer (10 mM Tris–HCl, pH 8.0, 10 mM NaCl and 0.2% NP-40) supplemented with protease inhibitor (Roche) to isolate nuclei. RIPA buffer (1× PBS, 1%NP-40, 0.1% SDS, 0.5% sodium deoxycholate and 0.004% sodium azide) supplemented with protease inhibitor was used to isolate the cross-linked

chromatin from nuclei. The isolated chromatin was sonicated to an average size of 250 bp using the Bioruptor Pico sonication device (Diagenode). An aliquot was saved each time to be subjected to DNA extraction and used as input. The leftover was used for ChIP assay.

For ChIP experiments with antibody against histone modification H3K27ac, 4 μg of antibody from Abcam (ab4729) together with 40 μl of Dynabeads protein G (Life Technologies) was added simultaneously into sonicated chromatin from 1 × $10^6$ cells and incubated on a rotating platform overnight at 4°C. For ChIP experiments against TFs, chromatin from 3 × $10^6$ cells was first precleared by incubating with 30 μl Dynabeads for 2 h at 4°C. The supernatant was separated and incubated overnight at 4°C with 8 μg of antibody against FOXA1 (ab5089, Abcam), FOXA2 (SC-6554 for HepG2 and SC-374376 for PLC/PRF/5 due to discontinuation of antibody used in HepG2 cells, Santa Cruz Biotechnology) or HNF4α (SC-6556X for HepG2 cells and SC-374229 for PLC/PRF/5 due to discontinuation of antibody used in HepG2 cells, Santa Cruz Biotechnology). Normal rabbit IgG (12-370, Millipore) was also included to check the background of antibody nonspecific binding in each batch of experiments. The immune complexes were captured by incubation with 75 μl of Dynabeads protein G at room temperature for 1 h. For all ChIP assays, a series of washing steps was applied to Dynabeads protein G after immunoprecipitation including four times of washing with RIPA buffer, two times of washing with ChIP wash buffer (10 mM Tris–HCl, pH 8.0, 0.25 M LiCl, 10 mM EDTA, 1% NP-40 and 1% sodium deoxycholate) and once with TE buffer. The DNA–protein complexes were eluted in direct elution buffer (10 mM Tris–HCl, pH 8.0, 5 mM EDTA, 0.3 M NaCl and 0.5% SDS) and together with the previous saved sonicated lysates to be used as input were treated with RNase A and proteinase K respectively and incubated at 65°C overnight to reverse crosslink. DNA was purified with Agencourt AMPure XP beads (Beckman Coulter) following the manufacturer's instructions, eluted in 10 mM Tris–HCl (pH 8.0) and ready to be analyzed by qPCR. Input DNA was quantified by Nanodrop 2000 (Life Technologies) and serial diluted to be used as templates for setting up stand curves of different primers targeting candidate regions.

**CRISPR/Cas9-mediated genome editing**

Lentivector lentiCRISPR v2 (Addgene plasmid # 52961) was a gift from Feng Zhang (Sanjana et al., 2014). In order to knock out the ME element, two gRNAs adjacent to the ME element were designed. The expression of the two gRNAs are driven by human and mouse U6 promoter, respectively. The expression cassette is inserted in lentiCRISPR v2 digested with BsmBI by In-Fusion cloning system (Takara). The AS-gRNA targeting each allele of rs953413 was directly synthesized and inserted into lentiCRISP v2 digested with BsmBI through T4 ligation. The identity of all the constructs were validated by Sanger sequencing. The detailed primer information is listed in Table S5. The lentivirus was produced in 293T cells and transduced into HepG2 cells coupled with sequabrene (Sigma) at final concentration of 8 μg/ml in 12-well plate. The positive cells were selected by puromycin (Life Technologies) at concentration of 1 μg/ml. Individual colonies were generated by splitting the positive cells to 0.5~3 cells per well in 96-well

plates and culturing in normal medium until single colony is formed. The genomic DNA for each single colony in 96-well plate is extracted by QuickExtract DNA Extraction Solution (QE09050, Lucigen). The introduced mutations for each colony are verified by both direct PCR sequencing and TA cloning followed by Sanger sequencing with the primers listed in Table S6. The selected colonies are passed regularly until reaching T75 flask and ready for expression study.

**Transactivation and transrepression of the rs953413 locus through dCas9-effector coupled with rs953413 AS-gRNA**

Lentiviral plasmid lenti-EF1a-dCas9-KRAB-Puro (Addgene plasmid # 99372), lenti-EF1a-dCas9-VP64-Puro (Addgene plasmid # 99371) and lentiGuide-Hygro-eGFP (Addgene plasmid # 99375) were gifts from Kristen Brennand (Ho et al., 2017). The AS-gRNA targeting each allele of rs953413 was directly synthesized and inserted into lentiGuide-Hygro-eGFP digested with BsmBI through T4 ligation. The HepG2 cells stably expressing dCas9-KRAB and dCas9-VP64 were acquired by lentiviral transduction followed by puromycin selection at 1 μg/ml. The cells expressing dCas9-effector were further transduced with lentivirus expressing AS-gRNA targeting either the A allele or the G allele of rs953413 and selected with medium containing 1 μg/ml puromycin and 0.2 mg/ml of hygromycin B (Life Technologies) until ready for expression study.

## Supplemental References


Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P., and Marth, G.T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics *27*, 1691-1692.

Battle, A., Brown, C.D., Engelhardt, B.E., and Montgomery, S.B. (2017). Genetic effects on gene expression across human tissues. Nature *550*, 204-213.

Blecher-Gonen, R., Barnett-Itzhaki, Z., Jaitin, D., Amann-Zalcenstein, D., Lara-Astiaso, D., and Amit, I. (2013). High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. Nat Protoc *8*, 539-554.

Cha, R.S., Zarbl, H., Keohavong, P., and Thilly, W.G. (1992). Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene. PCR Methods Appl *2*, 14-20.

Ho, S.M., Hartley, B.J., Flaherty, E., Rajarajan, P., Abdelaal, R., Obiorah, I., Barretto, N., Muhammad, H., Phatnani, H.P., Akbarian, S.*, et al.* (2017). Evaluating Synthetic Activation and Repression of Neuropsychiatric-Related Genes in hiPSC-Derived NPCs, Neurons, and Astrocytes. Stem cell reports *9*, 615-628.

Lemaitre, R.N., Tanaka, T., Tang, W., Manichaikul, A., Foy, M., Kabagambe, E.K., Nettleton, J.A., King, I.B., Weng, L.-C., Bhattacharya, S.*, et al.* (2011). Genetic Loci Associated with Plasma Phospholipid n-3 Fatty Acids: A Meta-Analysis of Genome-Wide Association Studies from the CHARGE Consortium. PLoS Genet *7*, e1002193.

Li, B., Kadura, I., Fu, D.-J., and Watson, D.E. (2004). Genotyping with TaqMAMA. Genomics *83*, 311-320.

Liang, G., He, H., Li, Y., and Yu, D. (2012). A new strategy for construction of artificial miRNA vectors in Arabidopsis. Planta *235*, 1421-1429.

Pan, G., Ameur, A., Enroth, S., Bysani, M., Nord, H., Cavalli, M., Essand, M., Gyllensten, U., and Wadelius, C. (2017). PATZ1 down-regulates FADS1 by binding to rs174557 and is opposed by SP1/SREBP1c. Nucleic Acids Res *45*, 2408-2422.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods *11*, 783-784.

The Encode Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Thomas, H., Senkel, S., Erdmann, S., Arndt, T., Turan, G., Klein-Hitpass, L., and Ryffel, G.U. (2004). Pattern of genes influenced by conditional expression of the transcription factors HNF6, HNF4alpha and HNF1beta in a pancreatic beta-cell line. Nucleic Acids Res *32*, e150.

Zhang, Y., Werling, U., and Edelmann, W. (2012). SLiCE: a novel bacterial cell extract-based DNA cloning method. Nucleic Acids Res *40*, e55-e55.