nature metabolism

ARTICLES

https://doi.org/10.1038/s42255-020-0211-z

Check for updates

Systematic mapping of genetic interactions for de novo fatty acid synthesis identifies *C12orf49* as a regulator of lipid metabolism

Michael Aregger[1,8], Keith A. Lawson[1,2,3,8], Maximillian Billmann[4,8], Michael Costanzo[1], Amy H. Y. Tong[1], Katherine Chan[1], Mahfuzur Rahman[4], Kevin R. Brown[1], Catherine Ross[1], Matej Usaj[1], Lucy Nedyalkova[1], Olga Sizova[1], Andrea Habsid[1], Judy Pawling[5], Zhen-Yuan Lin[5], Hala Abdouni[5], Cassandra J. Wong[5], Alexander Weiss[1], Patricia Mero[1], James W. Dennis[5], Anne-Claude Gingras[2,5], Chad L. Myers[4,6]✉, Brenda J. Andrews[1,2]✉, Charles Boone[1,2]✉ and Jason Moffat[1,2,7]✉

The de novo synthesis of fatty acids has emerged as a therapeutic target for various diseases, including cancer. Because cancer cells are intrinsically buffered to combat metabolic stress, it is important to understand how cells may adapt to the loss of de novo fatty acid biosynthesis. Here, we use pooled genome-wide CRISPR screens to systematically map genetic interactions (GIs) in human HAP1 cells carrying a loss-of-function mutation in fatty acid synthase (*FASN*), whose product catalyses the formation of long-chain fatty acids. *FASN*-mutant cells show a strong dependence on lipid uptake that is reflected in negative GIs with genes involved in the LDL receptor pathway, vesicle trafficking and protein glycosylation. Further support for these functional relationships is derived from additional GI screens in query cell lines deficient in other genes involved in lipid metabolism, including *LDLR*, *SREBF1*, *SREBF2* and *ACACA*. Our GI profiles also identify a potential role for the previously uncharacterized gene *C12orf49* (which we call *LUR1*) in regulation of exogenous lipid uptake through modulation of SREBF2 signalling in response to lipid starvation. Overall, our data highlight the genetic determinants underlying the cellular adaptation associated with loss of de novo fatty acid synthesis and demonstrate the power of systematic GI mapping for uncovering metabolic buffering mechanisms in human cells.

It has long been recognized that cancer cells exploit lipid metabolism to fuel their proliferative demands and support oncogenic signalling. Notably, alterations in lipid metabolism, including the uptake of lipids and/or synthesis of fatty acids, not only are recognized hallmarks of cancer but also commonly occur in diverse pathologic states such as fatty liver disease and metabolic syndrome, underscoring the importance of understanding this metabolic process[1]. De novo fatty acid synthesis in particular has gained substantial traction as a targetable pathway, following observations that overexpression of *FASN*, which encodes fatty acid synthase and catalyses the formation of long-chain fatty acids, and *ACACA*, which encodes acetyl-CoA carboxylase alpha and acts directly upstream of FASN, are associated with decreased survival rates for numerous solid malignancies[2–6]. Efforts to develop and translate small-molecule inhibitors of FASN (for example, TVB-2640) have helped to validate this enzyme as a targetable liability in cancer[7,8], and have led to several clinical trials (for example, NCT02223247, NCT02948569, NCT03179904 and NCT02980029). Given that metabolic pathways are highly buffered to cope with environmental change, genetic-screening approaches are a powerful strategy to reveal metabolic regulatory mechanisms that underscore metabolic redundancy, cross-talk and plasticity[9,10]. An understanding of how cells adapt to perturbation of de novo fatty acid synthesis could help to identify new targetable vulnerabilities that may inform new therapeutic strategies or biomarker approaches.

Mapping GI networks provides a powerful approach for identifying the functional relationships between genes and their corresponding pathways. The systematic exploration of pairwise GIs in model organisms has revealed that GIs often occur among functionally related genes, and that GI profiles organize a hierarchy of functional modules[11,12]. Thus, GI mapping has become an effective strategy for identifying functional modules and annotating the roles of previously uncharacterized genes. Model-organism GI mapping has also provided insight into the mechanistic basis of cellular plasticity or phenotypic switching that occurs as cells evolve within their environments[13,14]. Accordingly, the insights gained through systematic interrogation of GIs have fuelled considerable interest in leveraging these approaches for functionally annotating the human genome.

Recent technological advances using CRISPR–Cas enable the systematic mapping of GIs in human cells[15,16]. Here, we explore

[1]Donnelly Centre, University of Toronto, Toronto, Ontario, Canada. [2]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. [3]Division of Urology, Department of Surgery, University of Toronto, Toronto, Ontario, Canada. [4]Department of Computer Science and Engineering, University of Minnesota – Twin Cities, Minneapolis, MN, USA. [5]Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada. [6]Bioinformatics and Computational Biology Graduate Program, University of Minnesota – Twin Cities, Minneapolis, MN, USA. [7]Institute for Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada. [8]These authors contributed equally: Michael Aregger, Keith A. Lawson, Maximillian Billmann. ✉e-mail: cmyers@cs.umn.edu; brenda.andrews@utoronto.ca; charlie.boone@utoronto.ca; j.moffat@utoronto.ca

genome-wide GI screens within the context of human 'query' mutant cells defective in de novo fatty acid synthesis. We systematically mapped genome-wide GI profiles for six genes involved in lipid metabolism, revealing cellular processes that pinpoint genetic vulnerabilities associated with defects in de novo fatty acid synthesis. In particular, negative GIs with known fatty acid synthesis genes tend to identify other genes that are associated with this process, including a previously uncharacterized gene called chromosome 12 open reading frame 49 (*C12orf49*, which we name *LUR1*), which appears to function as a regulator of exogenous lipid uptake. Collectively, our data support the strategy of systematically mapping digenic interactions by using knockout (KO) query cell lines for identifying buffering mechanisms for a given biological process (such as lipid metabolism).

Results

Systematic identification of GIs for de novo fatty acid synthesis. De novo fatty acid synthesis is a multi-step enzymatic process that converts cytosolic acetyl-CoA, malonyl-CoA and NADPH to palmitate. Palmitate can be used directly or can be further elongated and/or undergo desaturation to form alternative lipid species. To systematically identify GIs associated with this metabolic process, we performed genome-wide CRISPR screens in co-isogenic cell lines that were either wild type (WT) or deficient in FASN, an enzyme involved in de novo fatty acid synthesis that is frequently overexpressed in malignancies[6,17] (Fig. 1a). We chose the human near-haploid cell line HAP1 as a model system, given the relative ease of generating KO mutations in this background[18]. We first validated our clonal *FASN*-KO cells by confirming loss of FASN protein levels by western blot (Extended Data Fig. 1a). We also performed targeted metabolite profiling of our parental HAP1 and *FASN*-KO cells, which revealed a significant increase in the FASN substrate malonyl-CoA in the *FASN*-KO cells, demonstrating their suitability as a model system for defective de novo fatty acid synthesis ($P=0.029$; Extended Data Fig. 1b).

To map *FASN* GIs, we performed genome-wide CRISPR screens using the sequence-optimized TKOv3 guide RNA (gRNA) library[19] in both the *FASN*-KO query cell line and control WT HAP1 cells, and we compared the relative abundance of individual gRNAs between the screen start (T0) and end (T18) time points (Fig. 1a,b). The relative abundance of gRNAs targeting each of ~18,000 protein-coding genes in WT cells provides an estimate of single-mutant fitness, whereas the relative abundance of gRNAs in a query mutant cell line provides an estimate of double-mutant fitness. Because mutant phenotypes can strongly depend on culture conditions[20] and most standard cell-culture medium contains supra-physiological nutrient levels that could mask phenotypic effects of perturbing certain metabolic pathways, we performed our screens utilizing medium conditions containing the minimum amounts of glucose and glutamine required to sustain proliferation of HAP1 cells (termed 'minimal medium'; Extended Data Fig. 1c and Methods).

We developed a quantitative GI (qGI) score that measures the strength and significance of GIs by comparing the relative abundance of gRNAs in a given query mutant cell line to the relative abundance of gRNAs targeting the corresponding genes in an extensive panel of 21 genome-wide WT HAP1 screens (Fig. 1b and Methods). In this context, negative interactions are identified as genes whose corresponding gRNAs exhibit significantly decreased abundance in a mutant KO background relative to that in the control WT HAP1 cell line, whereas positive interactions reflect genes with increased gRNA abundance in a mutant cell line relative to that in the parental line. The qGI score represents mean-summarized gRNA-level interactions after removal of disagreeing reagents (Fig. 1b and Methods).

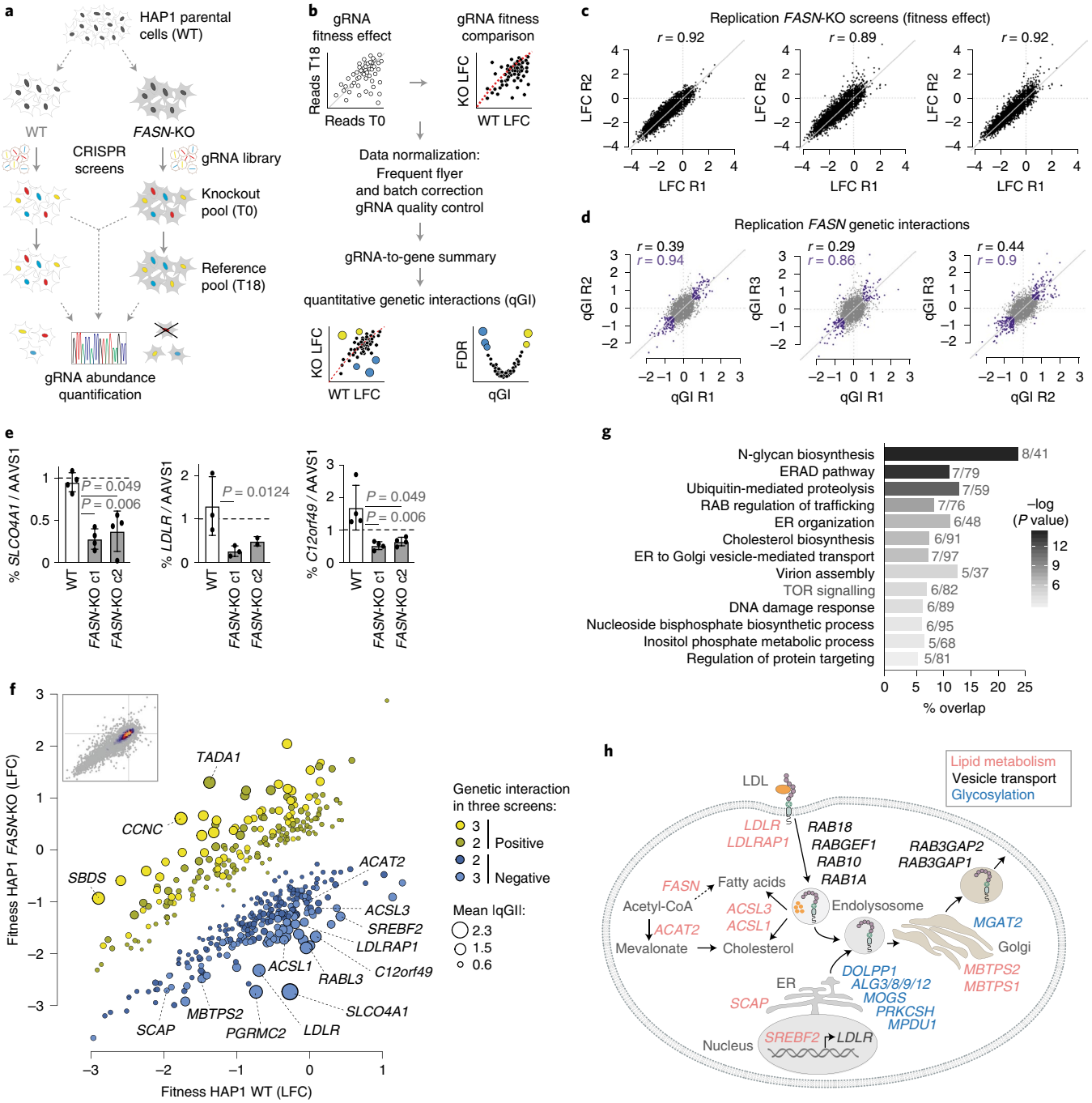
We performed three independent genome-wide, GI screens using our *FASN*-KO query mutant cell line. Because GIs rely on accurate measurement of single- and double-mutant phenotypes, we first examined the reproducibility of our single- and double-mutant fitness measurements (Methods). We observed a strong agreement of single-mutant fitness effects (LFC) among 21 WT HAP1 replicate

Fig. 1 | Genome-scale identification of digenic interactions with *FASN*. **a**, Schematic of the identification of GIs in co-isogenic HAP1 cell lines. *FASN*-KO and WT parental cells were infected with a lentiviral genome-wide CRISPR gene KO library (TKOv3), and gRNA abundance was determined using Illumina sequencing of gRNA sequences amplified from extracted genomic DNA from the starting cell population (T0) and end time point (day 18, T18) of the screen. **b**, Schematic outline for scoring qGI across co-isogenic query cell lines. First, the $\log_2$(fold change) (LFC) for each gRNA, comparing sequence abundance at the starting (T0) and end time point (T18), in a given query KO or WT cell population was computed. The differential LFC for each gRNA was then estimated by comparing its LFC in WT and query KO cells. A series of normalization steps and statistical tests were applied to these data to generate gene-level qGI scores and FDRs (Methods). The LFC scatterplot (bottom left graph) visualizes differential fitness defects in specific query KO and WT cells, whereas the volcano plot (bottom right graph) visualizes qGI scores for a specific query. **c**, Replicate analysis of gene loss-of-function fitness phenotypes in FASN screens. Scatterplots of LFC associated with perturbation of 17,804 individual genes derived from a query *FASN*-KO mutant screen conducted in triplicate. Reproducibility of fitness effects was determined by measuring the Pearson correlation coefficients (r) between all possible pairwise combinations of *FASN*-KO replicate screens. **d**, Evaluation of *FASN* qGIs. qGI scores were measured by comparing the LFC for every gene represented in the TKOv3 gRNA library in a *FASN*-KO cell line with those observed in a WT cell line, as described. Scatterplots show *FASN* GIs (qGI scores) derived from all possible pairwise combinations of three biological replicate screens. r was based on comparison of all qGI scores (r shown in grey, calculated on all the grey and purple data points in the scatterplots), or only GIs that exceed a given significance threshold ($|qGI| > 0.5$, FDR < 0.5) in both screens (purple). **e**, Validation of FASN negative GIs. Bar plots depict the ratio of WT and *FASN*-KO (2 independent clones, c1 and c2) cells carrying a gRNA targeting *SLCO4A1*, *LDLR* or *C12orf49*, all of which showed a negative interaction with *FASN*, compared with a gRNA targeting *AAVS1* (intergenic control). Experiments were performed with three independent gRNAs targeting each GI screen hit. All data are represented as means ± s.d.; $n=3$ (*LDLR*) or 4 (*SLCO4A1*, *C12orf49*) biologically independent experiments; P values are from one-way analysis of variance (ANOVA) (Kruskal–Wallis test). **f**, *FASN* negative and positive GIs. A scatterplot illustrating the fitness effect (LFC) of 450 genes in a *FASN*-KO versus WT parental HAP1 cell line, which exhibited a significant GI in at least 2 out of 3 *FASN*-KO replicate screens ($|qGI| > 0.5$, FDR < 0.5). Negative (blue) and positive (yellow) *FASN* GIs are shown. Node size corresponds to the mean absolute qGI score derived from three biologically independent replicate screens. Genes with mean absolute qGI score > 1.5, as well as selected negative interactions involving genes with established roles in lipid metabolism, are indicated. The inset is a scatterplot of the fitness effect (LFC) of all 17,804 genes targeted by the TKOv3 gRNA library in a *FASN*-KO cell line versus a WT parental HAP1 cell line. The colour indicates the density of genes. **g**, Enrichment for Gene Ontology (GO) molecular function, GO bioprocesses and Reactome terms among genes that exhibited a significant negative GI with *FASN* (significant in at least 2 *FASN* replicates, $|qGI| > 0.5$, FDR < 0.5). The number of genes overlapping a particular term and term size are indicated next to each bar. The greyscale colour legend for P values is indicated on the right; P values were calculated using gProfileR[69]. **h**, Schematic depicting the function of selected *FASN* negative interactions known to be involved in lipid uptake and homeostasis pathways (red), vesicle transport (black) and glycosylation (blue).

screens ($r > 0.87$) (Extended Data Fig. 1d) and double-mutant fitness effects derived from independent *FASN*-KO replicate screens ($r > 0.89$) (Fig. 1c). Moreover, all three *FASN* screens robustly distinguished a set of reference essential genes from non-essential genes (Extended Data Fig. 1e,f).

The identification of qGI scores depends on comparison of single-mutant fitness measurements in a WT HAP1 cell screen and double-mutant fitness measurements in a query mutant screen, both of which have inherent variability associated with them; therefore, the reproducibility of qGIs is expected to be more challenging than the measurement of either single- or double-mutant fitness phenotypes. Indeed, modest agreement was observed between qGI scores of the three *FASN*-KO replicate screens prior to filtering for significant interactions (pairwise $r = 0.29$ to 0.44) (Fig. 1d). The pairwise correlation between replicate screens increased substantially when we considered GIs found to be significant ($|qGI| > 0.5$; false-discovery rate (FDR) < 0.5) in at least one ($r = 0.52$–0.69) or two ($r = 0.86$–0.94) *FASN*-KO replicate screens (Fig. 1d, Extended Data Fig. 1g,h and Supplementary Table 1).

Leveraging all three *FASN*-KO replicates, we developed a reproducibility score that measures each gene's contribution to the covariance within two replicate screens and summarizes the resulting values across all available screen pairs (replicate 1–2, 1–3, 2–3) (Methods and Supplementary Table 1). This analysis confirms that both the strongest positive and negative qGI scores were highly reproducible across independent screens (Extended Data Fig. 1i). In particular, the most reproducible negative GIs with *FASN* were interactions with *SLCO4A1*, *PGRMC2*, *LDLR*, *RABL3* and *C12orf49* (Extended Data Fig. 1i and Supplementary Table 1). We tested three of these top five strongest negative GIs by independent validation



assays and confirmed all three by examining WT and *FASN*-KO HAP1 cells expressing gRNAs against *SLCO4A1*, *LDLR* and *C12orf49* (Fig. 1e and Extended Data Fig. 1j).

To generate an aggregate set of *FASN* GIs, we mean-summarized qGI scores across the three replicate screens (Fig. 1f and Supplementary Table 2). At a pathway level, significant negative GIs (qGI < −0.5; FDR < 0.5) with *FASN* were strongly enriched for genes annotated with roles in protein glycosylation, vesicle transport and cholesterol metabolism (FDR < 0.05) (Fig. 1g and Supplementary Table 3). In the global yeast genetic network, negative GIs often connect functionally related genes[11,21], and we observed a similar general trend for the *FASN* negative GIs. For example, the *FASN* negative GIs included genes with established roles in the uptake, transport and breakdown of low-density lipoprotein (LDL), a major extracellular source of lipids, including the LDL receptor (*LDLR*) itself and its coreceptor adaptor protein (*LDLRAP1*). We also observed negative GIs between *FASN* and the transcription factor *SREBF2*, which controls expression of *LDLR*, as well as *SCAP*, *MBTPS1* and *MBTPS2*, all of which are important in the activation and nuclear translocation of SREBF2 upon cholesterol depletion (Fig. 1h). Moreover, we observed negative GIs with additional lipid metabolic processes, such as cholesterol biosynthesis (*ACAT2*), with genes functioning in long-chain fatty acid activation and β-oxidation (*ACSL1*, *ACSL3*) and with vesicle-trafficking genes (*RAB18*, *RAB10*, *RAB1A*, *RABGEF1*, *RAB3GAP2* and *RAB3GAP1*) (Fig. 1h and Extended Data Fig. 1k), as well as a positive GI with the stearoyl-CoA desaturase gene (*SCD*), the product of which catalyses the rate-limiting step in the biosynthesis of monounsaturated fatty acids. SCD acts downstream of FASN in the elongation and desaturation of fatty acids, and loss of SCD seems to be buffered by the upstream perturbation of fatty acid synthesis (Supplementary Table 2).

The *FASN* screen highlighted an enrichment for genes functioning in protein N-linked glycosylation (for example, *ALG3*, *ALG8*, *ALG9*, *ALG12*, *MOGS*, *DOLPP1*, *PRKCSH*, *MGAT2*) (Fig. 1g,h and Extended Data Fig. 1k). Interestingly, the hexosamine biosynthetic and N-linked glycosylation pathways have been implicated in facilitating lipid accumulation from environmental sources through direct modulation of N-glycan branching on fatty acid transporters, possibly explaining the strong GIs we observed[22]. N-linked glycosylation is also known to have an important role in the activity of LDLR and activation of the SREBP transcriptional programs, providing a potential explanation for the interaction between loss of *FASN* and the glycosylation pathway[23,24]. Currently, we cannot exclude the notion that glycosylation of additional targets may play a role in the adaptation to loss of de novo fatty acid biosynthesis; nonetheless, our data highlight a role for N-linked glycosylation in fatty acid biosynthesis. Finally, we observed a significant negative GI between *FASN* and *SLCO4A1* (qGI < −1.84; FDR < 5.34×10^{-6}; Fig. 1f and Extended Data Fig. 1i). *SLCO4A1* encodes a member of the organic anion-transporting polypeptides (OATPs), which can transport a wide range of structurally unrelated compounds, including hormones, bile acids and lipid species (prostaglandins)[25]. Which of these compounds may buffer loss of FASN and how they might do so, or whether SLCO4A1 may be involved in the transport of additional metabolites, await further investigation. In summary, these results suggest that, in the absence of cell-autonomous de novo fatty acid synthesis, cells depend on uptake and breakdown of lipids from the environment or the synthesis of sterols. Furthermore, our data illuminate the genetic determinants of how proliferating cells rewire to meet the demand for lipids.

Expanding the GI landscape of de novo fatty acid synthesis. To better understand the GI landscape of de novo fatty acid synthesis, we next performed pooled genome-wide CRISPR screens using the TKOv3 library in five additional co-isogenic cell lines harbouring genetic KO of genes that exhibited significant negative GIs with our *FASN*-KO query, including *LDLR*, *C12orf49* and *SREBF2* (qGI < −0.5; FDR < 0.5; Supplementary Table 2), as well as two genes that did not show a negative GI with *FASN*, including *SREBF1*, which regulates the expression of *FASN* and other de novo fatty acid synthesis genes, and *ACACA*, which functions in the same pathway and immediately upstream of *FASN* (Fig. 2a)[6,17,26]. Each of these five query gene screens was performed in technical triplicate (that is, parallel cultures from a common infection). Because these GI screens were performed under the same conditions that we used for the *FASN*-KO screens, we applied the same confidence threshold on the derived qGI scores (|qGI| > 0.5, FDR < 0.5; Methods) (Fig. 2b–f, Extended Data Fig. 2a,b and Supplementary Table 2). At this confidence threshold, we estimated a per-screen FDR of ~0.3 and a false-negative rate of ~0.6 (Methods and Extended Data Fig. 1l). Overall, 169 (36%, $P < 2 \times 10^{-55}$) and 69 (49%, $P < 2 \times 10^{-9}$) of the genes that interacted with FASN in 2 or 3 screens, respectively, also interacted with at least 1 of the 5 queries (Fig. 2g and Supplementary Table 2).

We next analysed the functional enrichment across all GIs identified by our query screens related to fatty acid synthesis. Although we did not detect strong functional enrichment among the positive GIs in our dataset, we observed a clear fivefold enrichment of negative GIs for genes annotated to functionally relevant pathways, which were defined by the metabolism-focused HumanCyc standard[27] (Extended Data Fig. 2c). We further quantified enrichment for pathways annotated at different levels of the HumanCyc database hierarchy, including gene sets corresponding to general metabolic reaction categories, sub-categories and finally specific metabolic pathways (Supplementary Table 4). At the most general level of the HumanCyc pathway hierarchy, negative GIs from all six genome-wide screens were most enriched for genes annotated to the biosynthesis and macromolecule-modifications pathway categories (Fig. 3a). Through further analysis of these terms at a more specific level of the HumanCyc hierarchy (the subcategory level), we found that genes exhibiting negative GIs were associated with functions related to the roles of our six query genes, including fatty acid, lipid and carbohydrate biosynthesis (Fig. 3b and Extended Data Fig. 3a). At a more refined level of functional specificity within the fatty acid and lipid biosynthesis pathway, we found that each query gene was associated with a significant enrichment for negative GIs with functionally related genes of distinct pathways ($P < 0.05$). For example, the *LDLR* GI profile includes negative GIs with genes in the cholesterol/epoxysqualene biosynthesis pathway (*HMGCS1*, *MSMO1*, *HMGCR*, *FDFT1*, *NSDHL*, *HSD17B7*, *SQLE*, *HSD17B7*, *ACT2*, *SQLE*, *LSS*), and the *ACACA*, *LDLR* and *SREBF2* GI profiles include negative GIs with fatty acid elongation and biosynthesis pathway genes (*FASN*, *ACACA*, *OXSM*) (Fig. 3c,d). Notably, the *FASN* GI profile, and to a lesser extent the *ACACA* and *LDLR* GI profiles, revealed negative GIs with pathways and genes involved in N-glycosylation initiation (*ALG6*, *ALG13*, *ALG11*, *ALG1*, *ALG2*, *ALG8*, *ALG5*, *ALG3*, *ALG12*, *ALG9*), processing (*MOGS*, *PRKCSH*), dolichol monophosphate mannose synthase activity (*DPM2*, *DPM3*, *DPM1*) and glycan transfer (*STT3A*, *STT3B*) (Fig. 3c,e and Supplementary Table 4).

Our survey of GIs related to perturbation of de novo fatty acid synthesis or exogenous fatty acid uptake pathways provided unique insight into the genetic regulation of these processes. Specifically, for the *SREBF2* screen, although we observed negative GIs with lipid-uptake genes, such as *LDLR* and *LDLRAP1* (Fig. 3f and Supplementary Table 2), none were observed with the cholesterol biosynthesis pathway (Figs. 2d and 3d). This observation is consistent with *SREBF2* being the predominant transcriptional regulator of cholesterol homeostasis[26], as its perturbation does not further reduce cellular fitness in cells deficient in cholesterol biosynthesis. In addition, we detected a strong positive GI between *SREBF2* and

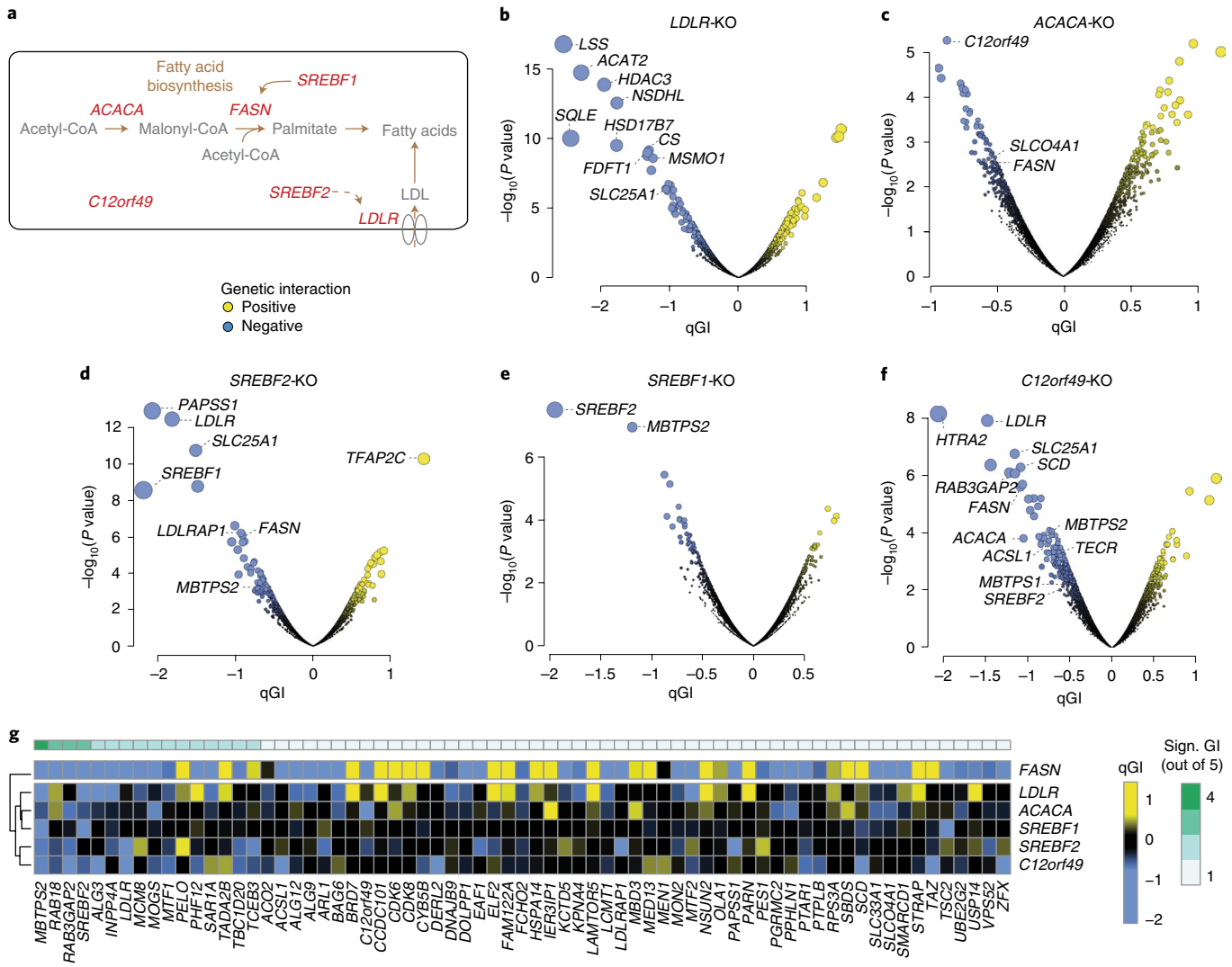


Fig. 2 | Querying five additional lipid metabolism genes for digenic interactions. a, Schematic diagram showing key steps in fatty acid metabolism. The genes encoding the proteins mediating these key steps, which are also query genes for GI screens described in the main text, are labelled in red. **b–f**, Volcano plots showing qGI scores and associated significance ($-\log_{10}$(*P* value)) for the 17,804 genes targeted by the TKOv3 gRNA library in the *LDLR*-KO (**b**), *ACACA*-KO (**c**), *SREBF2*-KO (**d**), *SREBF1*-KO (**e**) and *C12orf49*-KO (**f**) screens. Coloured dots indicate genes that meet the standard threshold of |qGI| > 0.5 and FDR < 0.5; positive GIs are indicated in yellow and negative GIs in blue. The dot size is proportional to both qGI and FDR, calculated as described in the Methods. Genes with |qGI| > 1.5, as well as selected top negative GI hits associated with lipid metabolism, citrate synthesis and transport, are indicated. **g**, Heat map showing overlapping GIs across the six interrogated queries. Heat map displaying genes (*x* axis) with significant interaction with FASN across all three FASN replicate screens and at least one additional screened query gene (*y* axis) (|qGI| > 0.5, FDR < 0.5). Positive and negative qGI scores are indicated in yellow and blue, respectively. The FASN qGI is represented as the mean qGI from the three FASN screens (same data as in Fig. 1f). The green bar above the heat map indicates the number of screens, out of the five screens in *LDLR*, *ACACA*, *SREBF1*, *SREBF2* and *C12orf49*-KO backgrounds, in which a significant GI was measured (|qGI| > 0.5, FDR < 0.5).

TFAP2C (Fig. 2d). Indeed, the *TFAP2* transcription factor family has recently been proposed to be a 'master' regulator of lipid droplet biogenesis[28], and as such our data suggest that reduced sequestration of lipids into lipid droplets may benefit *SREBF2*-KO cells by mitigating lipid starvation.

In contrast, *SREBF1* did not show enrichment for GIs for either the cholesterol or fatty acid synthesis pathways (Fig. 3c and Supplementary Table 2). Instead, this query was found to show only a strong reciprocal negative GI with its paralog *SREBF2*, highlighting the functional redundancy between the paralog pair (Fig. 2e and Supplementary Table 2), and suggesting that *SREBF2* may regulate some of the transcriptional targets of *SREBF1*, as previously described[26,29]. Furthermore, the imbalanced number of GIs between *SREBF1* and *SREBF2* may point towards asymmetric paralog evolution, whereby duplicated genes gain or lose functional roles at different rates while maintaining partially redundant functions, a process that has been previously observed in yeast and human cells[30–32].

A new role for *C12orf49* in lipid biosynthesis. One of the strongest negative GIs identified in both the *FASN* and the *ACACA* profiles involved the uncharacterized gene *C12orf49*, suggesting that this gene may have a role in lipid metabolism (Figs. 1f and 2c and Supplementary Table 2). C12orf49 is a 23.5-kDa protein that is part of the UPF0454 family of uncharacterized proteins. It contains an amino-terminal transmembrane sequence, a single

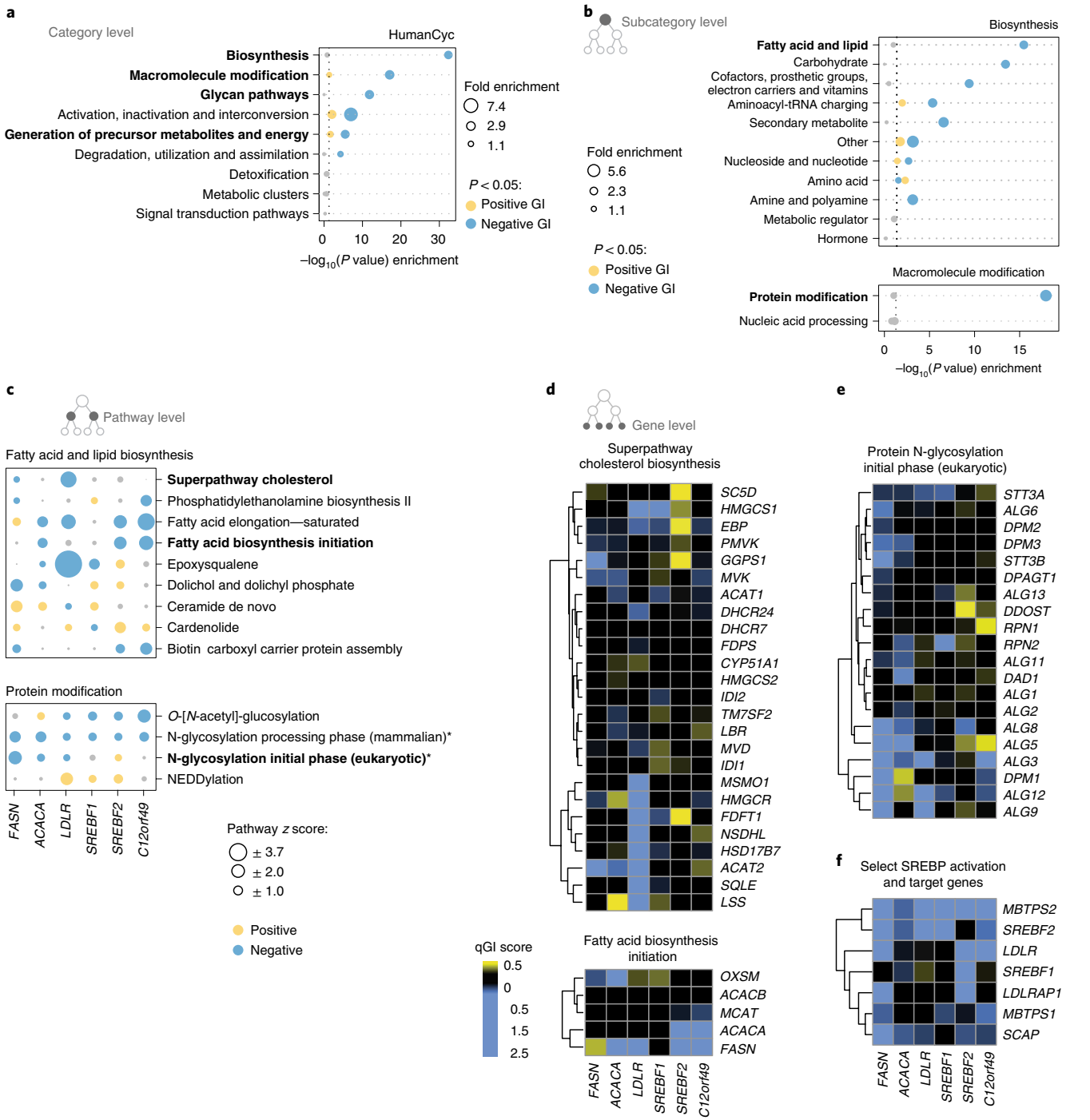


Fig. 3 | GIs reveal multiple levels of functional enrichment. a, Dot plot of normalized pathway enrichment scores on the HumanCyc-category level, calculated from qGIs across all six query genes (*FASN*, *ACACA*, *LDLR*, *SREBF1*, *SREBF2*, *C12orf49*). A GI was identified for a query-library pair if the |qGI| > 0.5 and FDR < 0.5. Enrichment for positive (yellow) and negative (blue) GIs is tested in each of the 10 HumanCyc main pathway categories using a hypergeometric test, considering all 17,804 genes targeted in the TKOv3 library as the universe. Enrichment with $P < 0.05$ is in blue (negative GI) and yellow (positive GI). Dot size is proportional to the fold enrichment in the indicated categories, and the black dotted line indicates the significance threshold. Categories in bold are further expanded in **b** and in Extended Data Fig. 3a. **b**, Dot plot of normalized pathway enrichment of GIs on a subcategory level, calculated as described in **a**, except that sub-categories were examined inside the biosynthesis and macromolecule-modification HumanCyc branches. Enrichment with $P < 0.05$ are blue (negative GI) and yellow (positive GI). Dot size is proportional to the fold enrichment in the indicated categories, and the black dotted line indicates the significance threshold. Categories indicated in bold are further expanded in **c**. **c**, Matrix dot plot of pathway enrichments of GIs for the fatty acid and lipid biosynthesis and protein modification sub-categories. Dots show positive (yellow) or negative (blue) *z*-transformed qGI scores summarized at the pathway level. qGI scores were first *z*-score transformed at a gene level for each genome-wide query screen separately. Then, a mean *z* score was calculated for each pathway for a given query screen. Dot size corresponds to the absolute *z*-transformed mean qGI score, and grey dots represent $|z| < 0.5$. Pathways marked with an asterisk are annotated to both protein modification and carbohydrate biosynthesis pathways. Bold pathways are shown in **d**–**e**. Pathways were displayed if they shared an absolute *z* score larger than 1.5 with any query gene. **d**–**f**, Gene-level heat maps for genes involved in enriched pathways. qGI scores between query genes and all genes from the selected pathways. Positive and negative qGI scores are indicated in yellow and blue, respectively.

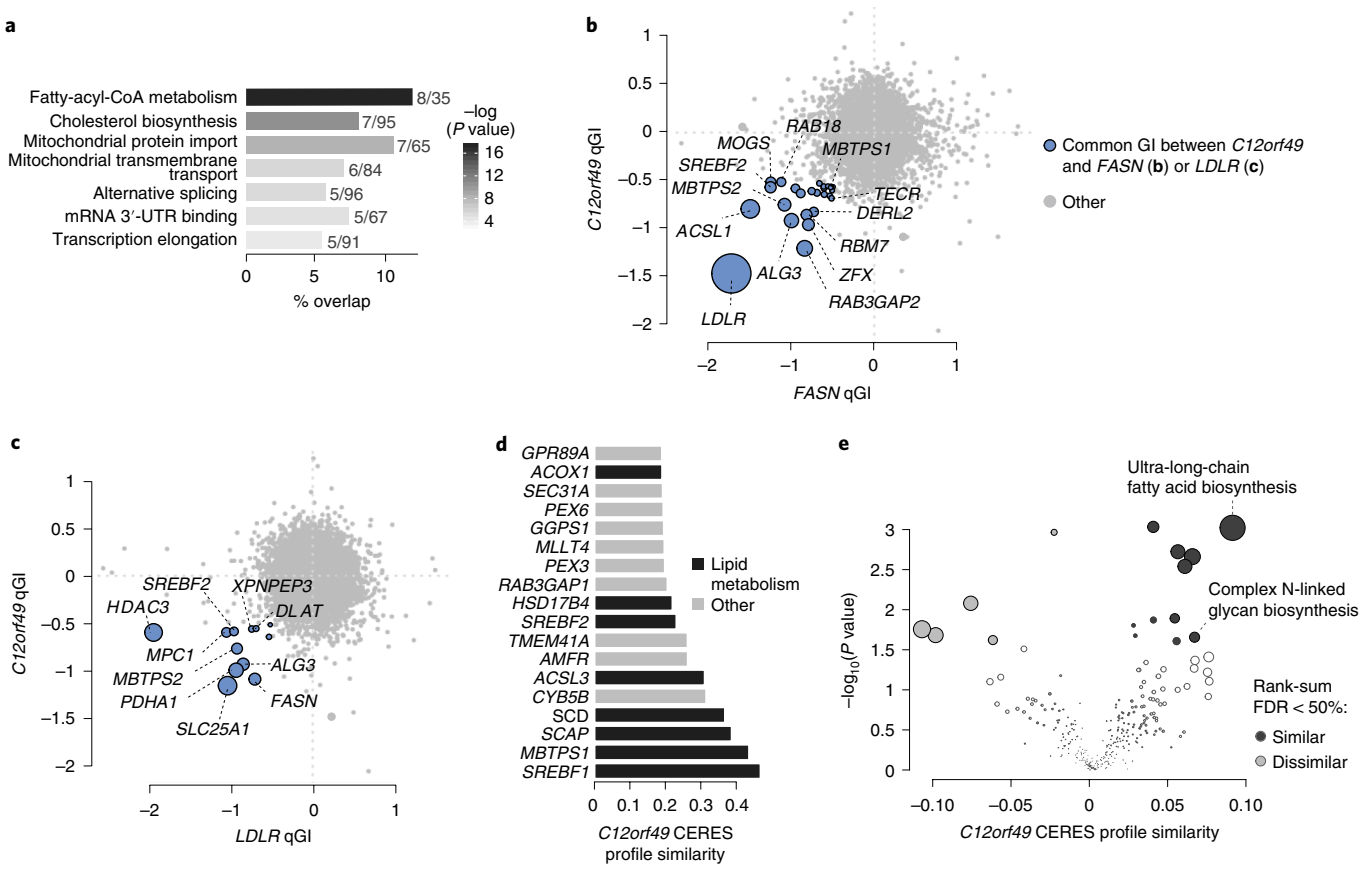


Fig. 4 | *C12orf49*'s GI profile suggests it has a functional role in lipid metabolism. a, Bar plot depicting pathway enrichment of negative GIs with *C12orf49* (|qGI| > 0.5; FDR < 0.5) using GO molecular functions, GO bioprocesses and Reactome standards. Significantly enriched gene sets are shown ($P < 0.05$; maximum term size, 100). Bars depict the mean percentage overlap with the indicated term, and the numbers next to each bar indicate the number of genes overlapping a particular term and term size. The greyscale colour legend for *P* values is indicated on the right; *P* values were calculated using gProfileR[69]. **b**, Scatterplot of *C12orf49* and *FASN* qGIs depicting GI overlap between *C12orf49* and *FASN* qGI scores. *FASN* qGI scores are represented as the mean of three independent screens. A common negative GI is determined if it is significant (qGI < −0.5; FDR < 0.5) in the *C12orf49*-KO screen and significant in 2 of 3 *FASN*-KO screens (indicated in blue). The top ten strongest common GIs, lipid metabolism and vesicle trafficking genes are labelled. **c**, Scatterplot of *C12orf49* and *LDLR* qGIs depicting GI overlap between *C12orf49* and *LDLR* qGI scores. A common negative GI is determined if it is significant (qGI < −0.5; FDR < 0.5) in both screens (indicated in blue). The top ten strongest common GIs and lipid metabolism genes are labelled. **d**, Bar plot indicating the *C12orf49* profile similarity across genome-wide DepMap CRISPR-Cas9 screens. Similarity (that is, co-essentiality) was quantified by taking all pairwise gene–gene Pearson correlation coefficients of CERES score profiles across 563 screens (19Q2 DepMap data release). The top 18 out of 17,633 gene profiles most similar to *C12orf49* are shown. Genes associated with lipid metabolism are indicated in black. **e**, Volcano plot of pathway enrichment for *C12orf49* co-essential genes. *C12orf49* co-essentiality profile scores for all 17,634 genes represented in the DepMap were mean-summarized by pathway, as defined in the HumanCyc standard[27]. Tendencies towards pathway-level similarity (co-essentiality) and dissimilarity (exclusivity) with *C12orf49* were tested using a two-sided Wilcoxon rank-sum test, followed by multiple-hypothesis correction with the Benjamini–Hochberg procedure.

uncharacterized DUF2054 domain of approximately 200 amino acid residues, 14 conserved cysteines (three of which are annotated to form cytosine–cytosine dimers) and a predicted glycosylation site[33] (Extended Data Fig. 4a). In some plant proteins, the uncharacterized UPF0454 is found in juxtaposition with a glycosyltransferase domain, and thus may be targeted into the lumen of the endoplasmic reticulum (ER) or Golgi[34]. By extension, the bulk of C12orf49 protein may reside in the lumen of the ER or Golgi. In addition, C12orf49 is ubiquitously expressed across tissues and cell lines (http://www.proteinatlas.org)[35]. Notably, expression of *C12orf49* is associated with differential prognoses on univariate analysis of The Cancer Genome Atlas data across multiple tumour types, including kidney, breast, liver and sarcoma[36] (Extended Data Fig. 4b–e; $P < 0.05$), which further motivated us to study the functional role of this previously uncharacterized gene.

GIs derived from a genome-wide screen using a *C12orf49*-KO query cell line further supported a role for this gene in lipid biogenesis. Consistent with the results described above, *C12orf49* showed a strong negative GI with both *FASN* and *ACACA* (Fig. 2f). *C12orf49* also showed negative GIs with *LDLR*, *ACSL1* (encoding acyl-CoA synthase), *SLC25A1* (encoding mitochondrial citrate transporter), *SCD* and *SREBF2*, further supporting a role for this gene in fatty acid biosynthesis (Fig. 2f). Consistently, *C12orf49* negative GIs were enriched for genes involved in fatty acid metabolism, cholesterol biosynthesis and additional metabolic pathways (FDR < 0.05) (Fig. 4a and Supplementary Table 3). Moreover, as observed for the *FASN* GI profile, *C12orf49* negative GIs involved genes functioning in vesicle-mediated trafficking and endocytosis, including *RAB3GAP2*, *RABIF*, *RAB18*, *VPS18*, *VPS419* and *VPS39* (Supplementary Table 2). Furthermore, many of the genes that showed a negative GI with *C12orf49* also had negative GIs with other query genes in our lipid metabolism panel (for example, *LDLR*, *ALG3*, *ASCL1*, *MBTPS2*, *SLC25A1*, *PDHA1*), supporting the functional relatedness of these genes (Fig. 4b,c and Extended Data

Fig. 4f–h). Thus, our lipid metabolism GI map strongly implicates *C12orf49* as having a functional role in lipid metabolism.

To further confirm the predictions about *C12orf49*'s function based on our HAP1 GI data, we examined publicly available data from the 19Q2 DepMap release and observed that *C12orf49* is essential for fitness in 120 out of 563 cell lines, with the highest dependencies observed for lung, ovarian, pancreatic, colon and bile-duct origins[37,38]. Other genes that shared similar cell line essentiality profiles to that of *C12orf49* included *SREBF1*, *SREBF2*, *MBTPS1*, *SCAP*, *SCD* and *ACSL3* (Fig. 4d and Extended Data Fig. 4i). The association of *C12orf49* with lipid metabolism genes was corroborated by a pathway enrichment analysis of the co-essentiality profiles, which revealed strong enrichment for genes annotated to ultra-long-chain fatty acid biosynthesis (Fig. 4e and Extended Data Fig. 4j). Furthermore, we found that cell lines that depend on *C12orf49* had missense mutations more frequently in *FASN* (FDR < 20%). Interestingly, germline variants in *C12orf49* have been reported to associate with serum lipid abnormalities in high-density lipoprotein (HDL) in a multi-ancestry cohort of the Million Veteran Program, further supporting a role for this gene in lipid metabolism[39]. Overall, these observations support a new function for *C12orf49* in lipid metabolism that is conserved across diverse cell types.

C12orf49 is a previously unrecognized regulator of lipid uptake. Given that *FASN*-deficient cells have a strong dependency on exogenous lipid uptake, we hypothesized that *C12orf49* might have a role in this biological process. To explore this, we measured uptake of labelled LDL particles, which represent one of the major sources of extracellular lipids, upon serum starvation across several HAP1 KO lines. As expected, loss of *LDLR* resulted in abolishment of LDL staining, whereas *FASN*-KO cells displayed increased uptake of exogenous lipid (Fig. 5a and Extended Data Fig. 5a). In contrast, loss of *C12orf49* caused a significant reduction of LDL uptake, which was rescued by the exogenous expression of C12orf49 ($P = 0.004$; Fig. 5a and Extended Data Fig. 5a). This phenotype was also observed in *SREBF1*- and *SREBF2*-deficient cells. Notably, this reduction of LDL was not secondary to a generalized impairment in receptor-mediated endocytosis, as loss of *C12orf49* did not impair uptake of labelled transferrin (Extended Data Fig. 5b). We further confirmed a strong reduction in LDL uptake in *C12orf49*-KO cells upon lipoprotein starvation (Extended Data Fig. 5c). Overall, these results support the hypothesis that *C12orf49* participates in lipid homeostasis through regulation of lipid uptake.

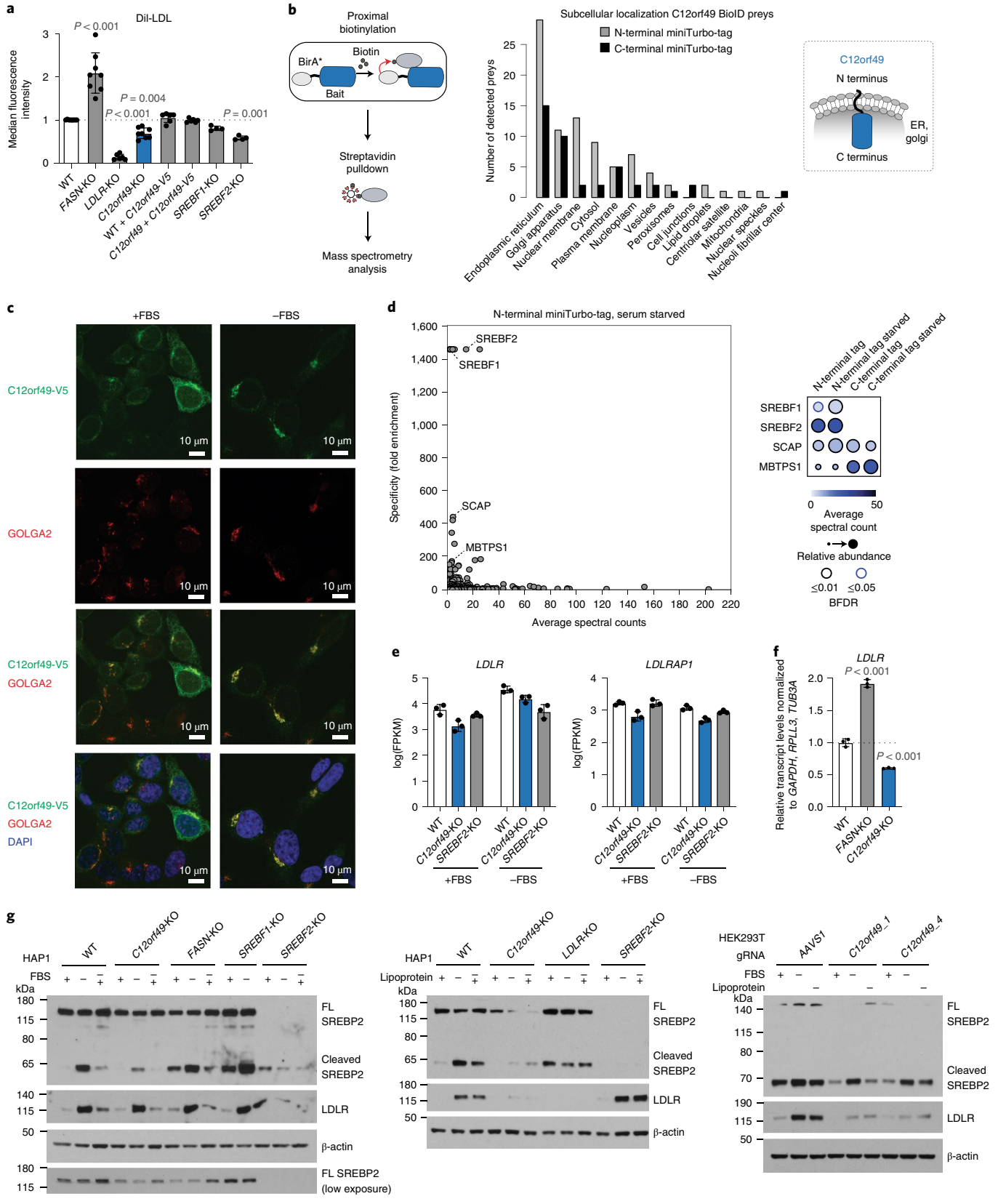
To investigate why *C12orf49* is required for optimal LDL uptake, we performed proximity-dependent biotinylation of proteins coupled to mass spectrometry (BioID) to reveal the physical neighbourhood in which C12orf49 resides. Because the C12orf49 single predicted N-terminal transmembrane domain may direct the carboxy-terminal DUF2054 domain into the lumen of the secretory pathway, leaving the N terminus facing the cytoplasm, BioID-MS was performed separately with both N- and C-terminal miniTurbo BirA*-tagged C12orf49 open reading frames (ORFs) expressed in HEK293 cells. Proximity-based labelling with the N-terminal construct captured proteins localizing to various cellular compartments, including the ER, Golgi apparatus, plasma membrane and cytosol, whereas the C-terminal miniTurbo construct revealed a strong enrichment of proteins localizing to the ER lumen (Fig. 5b and Supplementary Table 5). Furthermore, the BirA* ligase fused to the N-terminus captured proximal interactions that are enriched for proteins functioning in cholesterol biosynthesis and vesicle-mediated ER–Golgi transport, whereas the C-terminus-labelled proteins enriched for functions related to protein glycosylation (Extended Data Fig. 5d and Supplementary Table 6).

To study the subcellular localization of C12orf49, we performed immunofluorescence analysis under normal and serum-starved conditions, considering that SREBPs, the master regulators of lipid homeostasis, are known to travel between the ER and Golgi in response to lipid deprivation[26,40]. Under normal growth conditions (with serum), C12orf49 containing a C-terminal V5 tag (C12orf49–V5) was localized throughout the ER–Golgi network (Fig. 5c), consistent with our BioID results. Strikingly, C12orf49–V5 accumulated in the Golgi apparatus under serum starvation, as demonstrated by co-staining with GOLGA2, a Golgi membrane marker protein (Fig. 5c). These data thus suggest that localization of C12orf49 is regulated in a growth-condition-dependent manner, involving shuttling between the ER and the Golgi apparatus.

To gain insight into potential protein interactions of C12orf49, we further explored our BioID data performed under normal and serum-starved conditions (Extended Data Fig. 5e). Although 1,688 proteins passed the high-confidence criteria (≤1% Bayesian FDR) against our negative controls across the four tested conditions

Fig. 5 | C12orf49 shuttles between the ER and Golgi, and regulates lipid uptake through modulation of SREBP2 processing. a, Bar plots showing the results of LDL-uptake assays in the indicated cells using the Dil-LDL probe. All data are represented as means ± s.d.; $n = 4$ (*SREBF1*, *SREBF2*), 6 (*LDLR*, *WT* + *C12orf49-V5*, *C12orf49* + *C12orf49-V5*) or 8 (*WT*, *FASN*, *C12orf49*) independent biological replicates; *P* values are from one-way ANOVA. **b**, Left, Schematic outlining proximal protein capture using BioID mass spectrometry analysis (upper panel) and analysis of subcellular localization of C12orf49 BioID preys (lower panel). Right, Bar plots depicting the fraction of proteins localizing to indicated cellular compartments for preys captured with N-terminal (grey) or C-terminal (black) miniTurbo-tagged C12orf49 in HEK293 cells. The inset shows a schematic representation of the predicted topology and orientation of C12orf49 with respect to the cytoplasm and ER. **c**, Immunofluorescence microscopy analysis of C-terminal V5-tagged C12orf49 in HAP1 cells under the normal (left) or serum-starved (right) growth condition. C12orf49–V5 localization is shown in green, GOLGA2 is a marker of the Golgi apparatus and shown in red and DAPI (blue) marks the nuclei. Scale bars, 10 µm. Data are representative of two independent biological replicates. **d**, Scatterplot displaying the specificity of detected preys captured with BioID. Average spectral counts of preys captured in proximity to N-terminal miniTurbo BirA-tagged C12orf49 exposed to serum starvation are plotted against their specificity across hundreds of baits listed in the Human Cell Map BioID dataset (https://humancellmap.org/)[41] (left). The inset shows a comparison of the average spectral counts for preys involved in the regulation of SREBPs across the different miniTurbo-tagged constructs (N- and C-terminal tags) and growth conditions (normal (+FBS) and serum starved (−FBS)) (right). Bayesian FDR (BFDR) was calculated using SAINTexpress[62] v3.6.1. Data are representative of three biologically independent experiments. **e**, Bar plots indicating fragments per kilobase of transcript per million reads mapped (FPKM) expression values from RNA-sequencing data for *LDLR* and *LDLRAP1* in WT, *C12orf49*-KO and *SREBF2*-KO cells under normal and serum-starved growth conditions. All data are represented as means ± s.d.; $n = 3$ independent biological replicates. **f**, Bar plot of relative messenger RNA expression of *LDLR* across HAP1 WT, *FASN*-KO and *C12orf49*-KO cells. All data are represented as means ± s.d.; $n = 3$ independent biological replicates; *P* values from one-way ANOVA. **g**, Western blotting analysis of SREBP2, LDLR and β-actin levels across the indicated HAP1 co-isogenic KO cell lines cultured in the presence of FBS or exposed to overnight serum (left panel) or lipoprotein (middle panel) withdrawal and a short re-feeding period (indicated by the − and + symbols). Unprocessed full-length (FL) and processed C-terminal SREBP2 products are indicated. The same analysis was repeated in HEK293T cells (right panel). Prior to protein extraction, HEK293T cells were stably transduced with lentiviral Cas9 and gRNA expression cassettes, selected and cultured for 4 d. Data are representative of three independent biological replicates.

(Supplementary Table 5), distinguishing between specific proximity partners versus the general footprint of the ER–Golgi neighbourhood is challenging. To do this, we leveraged a recently generated reference map of a human cell (https://humancellmap.org/)[41], containing 192 BioID experiments in HEK293 cells, and compared our quantitative profiles with those of these baits. This revealed high-specificity interactions with the master lipid homeostatic transcription factor, SREBF2, in addition to SREBF1, SCAP and MBTPS1 (Fig. 5d). Notably, GIs were also observed between *C12orf49* and *SREBF2*, as well as the SREBF2-regulatory protease



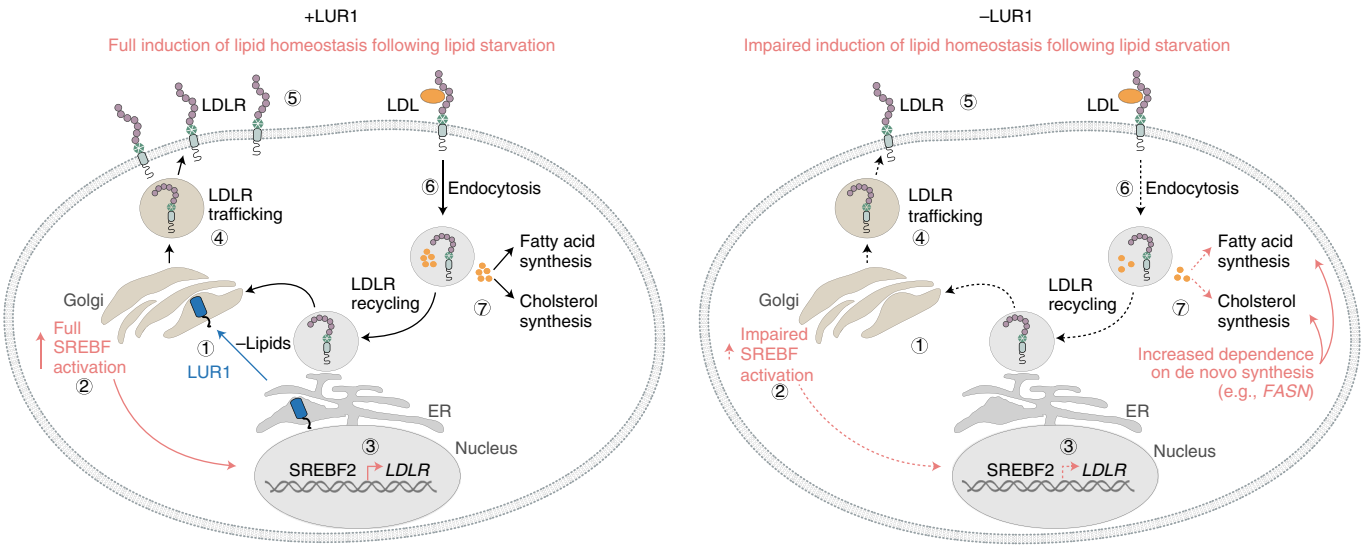


Fig. 6 | LUR1 shuttles between the ER and Golgi and regulates SREBF2 activation and lipid uptake. Our proposed model summarizing functions and locations of key players in lipid metabolism, including LUR1, and highlighting the processes induced upon lipid deprivation in presence (left) or upon loss of LUR1 (right). Left, (1) upon lipid deprivation (for example, –LDL), LUR1 and SREBF2 relocate from the ER to the Golgi; (2) SREBF2 is activated in the Golgi through proteolytic cleavage; (3) the processed, transcriptionally active domain shuttles to the nucleus, where it induces expression of target genes required for lipid homeostasis, such as *LDLR*; (4) newly synthesized and recycled LDLRs shuttle through the ER–Golgi network, where they are post-translationally modified (including glycosylation) and traffic to the cell surface; (5) on the cell surface, LDLR binds LDL particles; (6) LDL particles are taken up through receptor-mediated endocytosis; and (7) LDL particles are degraded, and lipoprotein becomes available for metabolic processes including the synthesis, modification or storage of fatty acids and cholesterol. Right, (1) loss of LUR1 results in (2) impaired SREBF2 processing and (3–5) subsequently reduced expression of LDLR; (6) LDL uptake levels are decreased as a consequence of reduced LDLR expression and uptake activity; and (7) decreased availability of extracellular lipoprotein leads to increased dependence on de novo synthesis pathways, such as de novo fatty acid and cholesterol synthesis, explaining the negative GI between *LUR1* and *FASN* and other members of lipid metabolic pathways.

genes *MBTPS1* and *MBTPS2* and the major SREBF2 transcriptional target *LDLR*. Further, similar to C12orf49, SREBF2 is known to traffic to the Golgi upon serum starvation, is cleaved and subsequently translocated to the nucleus and then activates the transcription of genes regulating lipid homeostasis[26,40]. Given this, we hypothesized that C12orf49 could have a role in the activation of the SREBF2.

To investigate a role for C12orf49 in the regulation of SREBF2, we performed RNA-sequencing experiments under normal and serum-starved conditions across HAP1 WT, *C12orf49*-KO and *SREBF2*-KO cells (Supplementary Table 7). As expected, serum starvation resulted in the induction of a cholesterol biosynthetic transcriptomic signature in HAP1 WT cells, but not in *SREBF2*-KO cells (Fig. 5e and Extended Data Fig. 6a,b). In *C12orf49*-KO cells, we observed a SREBF2-mediated transcriptional response similar to that of WT cells, suggesting that C12orf49 is not absolutely required for the activation of SREBF2 upon serum starvation (Fig. 5e and Extended Data Fig. 6a,b). However, we did notice a trend of lower expression of cholesterol biosynthesis and LDL-uptake genes in *C12orf49*-KO cells, which was confirmed by analysis of LDLR levels by quantitative reverse-transcription PCR (qRT–PCR) and western blot (Fig. 5f,g). To more directly test the hypothesis that C12orf49 regulates SREBF2 processing, we performed western blots to assess cleavage of SREBF2 into its active form following serum and lipid starvation. This demonstrated a clear decrease in overall SREBF2 expression and its processing in *C12orf49*-deficient versus WT HAP1 and HEK293T cells, a phenotype that was not rescued by inhibition of the proteasome (Fig. 5g and Extended Data Fig. 6c), confirming the requirement of this gene for optimal SREBF2 activation.

In summary, our unbiased GI screens and transcriptomic, proteomic and biochemical investigations reveal a new role for the uncharacterized gene *C12orf49* in the maintenance of lipid homeostasis. Our data indicate that C12orf49 localizes in a growth-condition-dependent manner throughout the ER–Golgi network, in parallel with SREBF2, and is required for full SREBF2 activation as a major lipid homeostasis transcription factor (Fig. 6). On the basis of our findings, we suggest that *C12orf49* be named *LUR1* for its role in lipid uptake regulation.

Discussion

The systematic mapping of GIs in model organisms such as yeast has provided a detailed view into the functional organization of eukaryotic cells[42]. Recent advances in CRISPR-based genome-engineering technologies provide a path for similar systematic GI studies in human cells[43–48]. Here, we apply genome-wide CRISPR-based fitness screens using query mutant HAP1 cell lines to systematically map GIs with a focus on lipid metabolism. Our data revealed a strong interaction between de novo fatty acid synthesis and lipid-uptake processes, highlighting a system that balances synthesizing lipids intracellularly with their uptake from the extracellular environment. More generally, this analysis confirms that relatively strong negative GIs identify functionally related genes, mapping a functional wiring diagram for a particular cellular process.

We screened a *FASN*-mutant query cell line multiple times and identified highly confident negative GIs, many of which were involved in lipid metabolism. Perturbation of de novo fatty acid synthesis has been suggested to be a prominent cancer therapeutic approach, and multiple compounds targeting FASN are currently being tested in clinical trials; for example, TVB-2640 is a FASN inhibitor that is being tested in solid tumours in phase 2 trials, and both fatostatin and betulin are inhibitors of the SREBP–SCAP interaction in preclinical development[6,49]. Because single-agent therapies often lead to emergence of resistance and tumour relapse, it makes sense to pursue therapeutic targets that are synergistic with FASN inhibition. Thus, the strong GIs detected in our FASN screen may

be informative for future investigations of combinatorial targets or biomarkers to treat diseases that would benefit from disruption of de novo fatty acid biosynthesis.

Our focused GI landscape related to de novo fatty acid biosynthesis provides unique insight into the genetic dependencies required for response to perturbation of lipid metabolism. Several pathways emerge as being most commonly utilized to adapt to perturbations, including those involved in alternative fatty acid and cholesterol biosynthesis processes, as well as lipid uptake. Interestingly, although our screens revealed strong negative GIs between de novo fatty acid synthesis and uptake of LDL, we failed to detect interactions with transporters of fatty acids. This may be a consequence of the genetic redundancy inherent among the *SLC27A* (FATP) fatty acid transporter family[50]. As previously shown in yeast[31], functional redundancy between paralogs can mask GIs associated with perturbation of a single gene of a duplicated pair, highlighting an important need for systems targeting multiple genes to survey complex GIs involving more than two genes[48]. Nonetheless, our data suggest a strong functional relationship between de novo fatty acid synthesis and glycosylation, and may involve a mechanism wherein cells modify the FATP transporters through N-glycosylation, thereby enhancing lipid uptake, as suggested by Ryczko et al.[22] As such, this pathway serves as an obvious focal point for not only ongoing mechanistic investigation, but also therapeutic development for anti-cancer strategies targeting de novo fatty acid synthesis.

Genome-wide GI profiling revealed an important role for *LUR1* (*C12orf49*) in lipid uptake. Interestingly, analysis of the DepMap data revealed that *LUR1* is essential in the same set of cancer cell lines that depend on lipid-biosynthesis-related genes for viability, including *SREBF1*, *MBTPBS1*, *SCAP* and *SCD*. Similarly, two recent studies identifying co-functional gene clusters support a functional role for *LUR1* in lipid metabolism across diverse genetic backgrounds[51,52]. Furthermore, genome-wide association studies in large cohorts have found *LUR1* variants linked to abnormal HDL profiles[39], neuroticism[53–55] and body height[54], all phenotypes that could have root causes in lipid metabolism defects.

In summary, we present an unbiased and genome-wide approach for uncovering genetic vulnerabilities related to lipid metabolism in human cells, which led us to identify a function for *LUR1* in regulating SREBF2 activity (Fig. 6). Our GI profiles for de novo fatty acid synthesis and related lipid-uptake genes provide a resource for studying metabolic rewiring and disease phenotypes linked to lipid metabolism. We also demonstrate the power of systematic GI profiling using query mutants in a co-isogenic cell line, an approach that can be applied to other bioprocesses and expanded to begin generating more comprehensive GI maps for human genes.

Methods

Cell culture. Human HAP1 WT cells were obtained from Horizon Genomics (clone C631; sex, male with lost Y chromosome; RRID: CVCL_Y019). The following HAP1 gene-KO cell lines were obtained from Horizon: *FASN* (HZGHC003700c006, used for GI screening; HZGHC003700c011, used for hit validation), *ACACA* (HZGHC004903c002), *LDLR* (HZGHC003978c007), *SREBF1* (HZGHC001361c012) and *SREBF2* (HZGHC000683c004). All gene-KO cell lines were confirmed to carry the expected out-of-frame insertions or deletions by Sanger sequencing of PCR products. HAP1 cells were maintained in low glucose (10 mM), low glutamine (1 mM) DMEM (Wisent no. 319-162-CL) supplemented with 10% FBS (Life Technologies) and 1% penicillin–streptomycin (Life Technologies). This culture medium is referred to as 'minimal medium'. Cells were dissociated with trypsin (Life Technologies) and all cells were maintained at 37 °C and 5% CO_2. Cells were regularly monitored for mycoplasma infection.

HAP1 KO cell line generation. The HAP1 *C12orf49* gene-KO cell line was constructed by first cloning a gRNA targeting *C12orf49* (Supplementary Table 8) into the pX459v2 backbone (Addgene no. 62988), which was modified to carry the same restriction overhangs as the pLCKO vector (Addgene no. 73311). Then, 350,000 HAP1 WT cells were seeded into a 6-well plate, and 24 h later, cells were transfected with a mix of 2 µg pX459v2 plasmid carrying a gRNA, 6 µl X-treme Gene transfection reagent (Roche) and 100 µL Opti-MEM media (Life Technologies). Twenty-four hours after transfection, cells were selected in medium containing 1 µg mL^{-1} puromycin for 3 d, and single cells were sorted onto 96-well plates by manual seeding of a single-cell suspension at 0.6 cells per well. Following amplification of cells from individual wells, genomic DNA was extracted with Extracta DNA Prep (Quanta Bio), Sanger sequencing was performed across the gRNA target sites following PCR amplification and successful gene KO was identified following sequence analysis.

Library virus production and multiplicity of infection determination. For CRISPR library virus production, 8 million HEK293T cells were seeded per 15-cm plate in DMEM containing high glucose, pyruvate and 10% FBS. Twenty-four hours after seeding, the cells were transfected with a mix of 8 µg lentiviral lentiCRISPRv2 vector containing the TKOv3 gRNA library[19] (Addgene no. 90294), 4.8 µg packaging vector psPAX2, 3.2 µg envelope vector pMD2.G, 48 µl X-treme gene transfection reagent (Roche) and 1.4 mL Opti-MEM medium (Life Technologies). Twenty-four hours after transfection, the medium was replaced with serum-free, high-BSA growth medium (DMEM, 1.1 g per 100 mL BSA, 1% penicillin–streptomycin). Virus-containing medium was collected 48 h after transfection, centrifuged at 1,500 r.p.m. for 5 min, aliquoted and frozen at −80 °C.

For determination of viral titres, 3 million HAP1 cells seeded in 15-cm plates were transduced with different dilutions of the TKOv3 lentiviral gRNA library along with polybrene (8 µg mL^{-1}), in a total of 20 mL medium. After 24 h, the virus-containing medium was replaced with 25 mL of fresh medium containing puromycin (1 µg mL^{-1}), and cells were incubated for an additional 48 h. Multiplicity of infection (MOI) of the titrated virus was determined 72 h post-infection by comparing percentage survival of puromycin-selected cells with that of cells that were infected but not selected with puromycin (puro-minus controls).

Pooled CRISPR dropout screens. For pooled CRISPR dropout screens, 3 million HAP1 cells were seeded in 15-cm plates in 20 mL of specified medium. A total of 90 million cells were transduced with the lentiviral TKOv3 library at a MOI of ~0.3, such that each gRNA was represented in about 200–300 cells. Twenty-four hours after infection, transduced cells were selected with 25 mL medium containing 1 µg mL^{-1} puromycin for 48 h. Cells were then collected and pooled, and 30 million cells were collected for subsequent gDNA extraction and determination of the library representation at day 0 (that is, T0 reference). The pooled cells were then seeded into three replicate plates, each containing 18 million cells (>200-fold library coverage), which were passaged every 3 d and maintained at >200-fold library coverage until T18. Genomic DNA pellets from each replicate were collected at each day of cell passage.

Preparation of sequencing libraries and Illumina sequencing. Genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega). The gDNA pellets were resuspended in TE buffer, and the concentration was estimated by Qubit using double-stranded DNA (dsDNA) Broad Range Assay reagents (Invitrogen). Sequencing libraries were prepared from 50 µg of the extracted gDNA in two PCR steps, the first to enrich gRNA regions from the genome, and the second to amplify gRNA and attach Illumina TruSeq adapters with i5 and i7 indices, as described previously, using staggered primers aligning in both orientations to the gRNA region (Supplementary Table 8)[56]. Barcoded libraries were gel purified, and final concentrations were estimated by qRT–PCR. Sequencing libraries were sequenced on an Illumina HiSeq2500 using single-read sequencing and were completed with standard primers for dual indexing with HiSeq SBS Kit v4 reagents. The first 21 cycles of sequencing were dark cycles, or base additions without imaging. The actual 36-base read begins after the dark cycles and contains 2 index reads, in which i7 is read first, followed by the i5 sequences. The T0 and T18 time point samples were sequenced at 400- and 200-fold library coverage, respectively.

Construction of colour-coded lentiCRISPRv2 vectors for co-culture assay. The colour-coded lentiCRISPRv2 vectors were derived from the lentiCRISPRv2 vector (Addgene no. 52961) by inserting mCherry (Addgene no. 36084) or mClover3 (Addgene no. 74236) open reading frames between the Cas9 and PuroR expression cassette. To this end, the lentiCRISPRv2 vector was digested with BamHI, PCR products coding for the respective fluorescent protein flanked by T2A and P2A self-cleaving peptides were ligated into the vector using Gibson assembly. The two forward primers (Supplementary Table 8) were used at a 1:0.1:1 (P233:P234:P235) ratio in the same PCR reaction with the reverse primer (primers bind to both fluorescent proteins mCherry and mClover3).

Validation of GIs using co-culture assays. For validation of GIs, HAP1 parental and gene-KO clones were transduced with colour-coded lentiCRISPRv2 vectors targeting either an intergenic site in the AAVS1 locus (negative control), or a specific target gene hit (for example, *LDLR*). Each gene was targeted with three independent and unique gRNAs. Twenty-four hours after transduction, cells were selected with 1 µg mL^{-1} puromycin for 48 h and seeded for co-culture proliferation assays as follow: 50,000 of green (for example lentiCRISPRv2-mClover3 *AAVS1* gRNA) and red (for example lentiCRISPRv2-mCherry hit gene gRNA) cells were mixed (total, 100,000 cells) in a 6-well plate in both colour orientations for both

parental and gene-KO cells. Cells were passaged every 4 d until day ~12 (T12). Cells were trypsinized, washed and stained for dead cells using Zombie NIR (BioLegend). The relative proportions of red and green cells in the co-culture were assessed using an LSR Fortessa or X20 flow cytometer (BD Bioscience). The relative ratio of Hit:AAVS1 was calculated and averaged for the three gene-targeting guides and two-colour orientations.

LDL- and transferrin-uptake assay. For uptake experiments with labelled probes, 150,000 HAP1 cells were seeded in a 12-well plate. After 48 h, cells were serum starved or deprived of lipoprotein overnight in minimal medium (described above) complemented with 0.3% BSA (BioShop) or lipoprotein-deprived FBS (MilliporeSigma no. S5394) instead of standard FBS, respectively. After 16 h, cells were labelled with Dil-LDL (Invitrogen no. L3482), pHrodo Red LDL (Invitrogen no. L34356) or pHrodo Red Transferin (Invitrogen no. P35376) at 2 μg mL^{-1} (1:500) in minimal medium plus 0.3% BSA for 15 min at 37 °C. Cells were washed in PBS, trypsinized and stained with 7-AAD (BioLegend no. 420404) or Zombie NIR (BioLegend no. 423105) cell viability solution at 25 ng mL^{-1} (1:2,000) for 5 min at room temperature. Staining was measured using an LSR Fortessa or X20 flow cytometer (BD Bioscience). The gating strategy is outlined in Extended Data Fig. 7.

Proximity-based labelling of proteins captured with mass spectrometry. BioID-MS analysis was performed essentially as described previously[57], with minor modifications. In brief, HEK293 Flp-In T-REx lines expressing inducible N- or C-terminal miniTurbo-FLAG-tagged C12orf49 open reading frames were generated[58]. Five (for normal growth condition) and 12.5 (for serum-starvation condition) million cells were seeded on 15-cm plates, and after 24 h, cells were treated with 1 μg mL^{-1} tetracycline to induce expression of baits. Twenty-four hours later, 50 μM biotin was added for labelling of proximal proteins for 3.5 h. Cell pellets were collected and lysed in RIPA lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% (wt/vol) SDS, 1% NP-40, 1 mM EDTA, 1 mM $MgCl_2$; 0.5% deoxycholate and Sigma protease inhibitors were added right before cell lysis) at a 1:10 (g:mL) ratio, and were sonicated 3 times for 5 s with 2-s breaks. Then, 1 μL per sample TurboNuclease (BioVision) and 1 μL per sample RNAse (Sigma) were added, and samples were incubated at 4 °C for 30 min; 20% SDS was added to bring the sample's final SDS concentration to 0.25%, and samples were mixed well and centrifuged at 14,000 r.p.m. (Microfuge) for 20 min at 4 °C. The supernatant was added to streptavidin resin (pre-washed with lysis buffer) using 30-μL bed volume and was rotated at 4 °C for 3 h. Beads were washed after binding as following: (1) 1 × 1 mL of 2% SDS buffer (2% SDS, 50 mM Tris-Hcl pH 7.5); (2) 1 × 1 mL of lysis buffer; (3) 1 × 1 mL of HEK293 lysis buffer (with 0.1% NP-40); and (4) 3 × 1 mL of 50 mM ammonium bicarbonate (made fresh). After purification of biotinylated preys using streptavidin sepharose, samples were digested on beads using trypsin. Samples were separated by liquid chromatography and analysed by tandem mass spectrometry (LC–MS/MS) on a AB Sciex TripleTOF 5600 mass spectrometer.

The LC–MS/MS setup consisted of a TripleTOF 5600 (SCIEX) equipped with a nanoelectrospray ion source connected in-line to an AS-2 Nano-HPLC system (Eksigent Technologies). The fused silica column (10 cm × ID 75 μm, OD 360 μm) had an integrated emitter tip prepared in-house using a laser puller (Sutter Instrument Co.). The column was packed with ~10 cm of C18 resin (Reprosil-Pur, 3.5 μm, Dr.Maisch HPLC). Next, 5 μL sample was loaded onto the column using the autosampler at 400 nL min^{-1}, and the LC delivered the organic phase gradient at 200 nL min^{-1} over 90 min (2–35% acetonitrile with 0.1% formic acid). The MS instrument was operated in data-dependent acquisition mode with 1-MS scan (250 ms; mass range, 400–1,250 *m/z*) followed by up to 20 MS/MS scans (100 ms each). Only candidate ions between 2 and 4 charge states were considered, and ions were dynamically excluded for 15 s with a 50-mDa window. The isolation width was 0.7 *m/z*, and the minimum threshold was set to 250. Between sample injections, 2 blank samples were injected (0.1% formic acid), each with 3 rapid gradient cycles at 300 nL min^{-1} over 60 min. Before another sample was injected, system performance was verified with a 30-min BSA quality control run, and a 30-min BSA mass-calibration run.

Raw files (.WIFF and.WIFF.SCAN) were converted to an MGF format and to an mzML format using ProteoWizard (v3.0.4468) and the AB Sciex MS Data Converter (v1.3 beta), as implemented within ProHits[59]. For human samples, the database used for searches consisted of the human and adenovirus sequences in the RefSeq protein database (version 57). The database was supplemented with 'common contaminants' from the Max Planck Institute (http://141.61.102.106:8080/share.cgi?ssid=0f2gfuB) and the Global Proteome Machine (http://www.thegpm.org/crap/index.html), and with commonly used epitope tags. The search databases consisted of forward and reverse sequences (labelled 'gi 9999' or 'DECOY'); in total, 72,481 entries were searched for the human database. Spectra were analysed separately using Mascot (2.3.02; Matrix Science) and Comet (2018.01 rev.4) with trypsin specificity and up to two missed cleavages; deamidation (Asn or Gln) and oxidation (Met) were selected as variable modifications. The fragment mass tolerance was 0.15 Da, and the mass window for the precursor was ±35 ppm with charges of 2+ to 4+ (both monoisotopic mass). The resulting Comet and Mascot results were individually processed by PeptideProphet and combined into a final iProphet output using the Trans-Proteomic Pipeline (TPP; Linux version, v5.2.1-dev Flammagenitus, Build-201906251008-exported). TPP options were as follows: peptide prophet were --minprob 0.05 --ppm --decoy DECOY --nonparam --accmass --expectscore --decoyprobs; and iProphet options were --nonsp --nonrs --nonsi --nonsm --nonse. All proteins with a minimal iProphet probability of 0.95 and 2 unique peptides were used for analysis.

Data processing and analysis was performed within the ProHits LIMS[60] searched against the RefSeq human and adenovirus database, version 57, forward and reverse. Mascot and Comet search results were jointly analysed using the iProphet component of the Trans Proteomic Pipeline[61].

High-confidence interactions were determined by scoring bait samples against negative control samples (6 runs of miniTurbo-FLAG-EGFP) using the statistical tool SAINTexpress[62] v3.6.1 with default parameters. Preys with a SAINT score (FDR) of <1% were considered high-confidence hits. To calculate specific enrichment of preys against a reference dataset, we leveraged the humancellmap.org, a compendium of 192 BioID baits profiled in HEK293s[41]. Each SAINTexpress-filtered list was uploaded to the 'Analyze' module of the humancellmap, and processed using default options. Default specificity plots (calculating the specific enrichment of spectral counts for a prey with a bait against that across the entire database) were generated within the humancellmap site and visualized in ProHits-viz[59].

Western blotting. HAP1 and 293T cells were lysed in buffer F (10 mM Tris pH 7.05, 50 mM NaCl, 30 mM sodium pyrophosphate, 50 mM NaF, 10% glycerol, 0.5% Triton X-100) and centrifuged at 14,000 r.p.m. for 10 min. The supernatant was collected, and protein concentration was determined using Bradford reagent (Bio-Rad). Next, 10–30 μg protein was resolved on 4–12% Bis-Tris gels (Life Technologies) and transferred to Immobilon-P nitrocellulose membrane (Millipore) at 66 V for 90 min. Subsequently, proteins were detected using anti-FASN (1:2,000, Abcam no. ab128870), anti-SREBP2 (1:250, BD Biosciences no. 557037), anti-LDLR (1:250, Abcam no. ab52818), anti-GAPDH (1:10,000, Santa Cruz no. 166574) and anti-β-actin (1:10,000, Abcam no. ab8226) antibodies and proteins were visualized on X-ray film using Super Signal chemiluminescence reagent (Thermo Fisher Scientific) and Western Lightning ECL Pro (PerkinElmer). Scans of uncropped western blots are provided in Source Data for Fig. 5 and Extended Data Figs. 1 and 6.

Immunofluorescence. Cells were seeded on cover slips and fixed with 4% paraformaldehyde in PBS for 10 min at room temperature. Cells were permeabilized with 1% NP-40 in antibody dilution solution (PBS, 0.2% BSA, 0.02% sodium azide) for 10 min and blocked with 1% goat serum for 45 min. Cells were incubated with anti-V5 (1:250, Abcam no. ab27671) and anti-GOLGA2 antibodies (1:250, Sigma no. HPA021799) for 1 h at room temperature. Subsequently, cells were incubated with Alexa Fluor 488 goat anti-mouse (1:500, Invitrogen no. A-11001) or Alexa Fluor 647 anti-rabbit antibodies (1:500, Invitrogen no. A-21245) and counterstained with 1 μg mL^{-1} DAPI (Cell Signaling Technology no. 4083S) for 45 min in the dark. Cells were visualized by confocal microscopy (Zeiss LSM 880).

RNA sequencing. *Sample preparation.* HAP1 WT cells and *C12orf49*- and *SREBF2*-KO cells were cultured in minimal DMEM for 48 h and either control treated or serum starved for 4 h, as indicated. Each cell line was cultured and processed in three biological replicates. RNA was extracted using the RNeasy Kit (Qiagen), according to the manufacturer's instructions. Eighteen total RNA samples were DNase-treated using RNase-free DNase Set (Qiagen no. 79254). Samples were submitted for mRNA-Seq at the Donnelly Sequencing Centre at the University of Toronto (http://ccbr.utoronto.ca/donnelly-sequencing-centre). RNA was quantified using Qubit RNA BR (Thermo Fisher Scientific no. Q10211) fluorescent chemistry, and 1 ng was used to obtain the RNA Integrity Number (RIN) using the Bioanalyzer RNA 6000 Pico kit (Agilent Technologies no. 5067-1513). The lowest RIN was 9.5; the median RIN score was 9.8. Then, 1,000 ng per sample was processed using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs no. E7760L) and included poly(A) enrichment using NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs no. E7490L), fragmentation for 15 min at 94 °C prior to first-strand synthesis, and 8 cycles of amplification after adapter ligation. For each purified final library, 1 μL top stock was run on an Agilent Bioanalyzer dsDNA High Sensitivity chip (Agilent Technologies no. 5067-4626). The libraries were quantified using the Quant-iT dsDNA high-sensitivity (Thermo Fisher Scientific no. Q33120) and were pooled at equimolar ratios after size adjustment. The final pool was run on an Agilent Bioanalyzer dsDNA High Sensitivity chip and quantified using NEBNext Library Quant Kit for Illumina (New England Biolabs no. E7630L). The quantified pool was hybridized at a final concentration of 400 pM and sequenced paired-end on the Illumina NovaSeq6000 platform using a S2 flowcell at 2 × 151-bp read lengths.

Data processing. Samples were mixed to obtain an average of 35 million clusters that passed filtering. Reads shorter than 36 bp on either read 1 or read 2 were removed prior to mapping. Reads were aligned to reference genome hg38 and Gencode V25 gene models using the STAR short-read aligner (v2.6.0a)[63]. Approximately 80% of the filtered reads mapped uniquely, and the read counts from each sample, computed by STAR, were merged into a single matrix using R.

Differential expression. Differentially expressed genes were identified using the Bioconductor packages limma (v3.32.10) and edgeR (v3.24.3). The read-count matrix was filtered using the filterByExpr() function using default parameters. Principal-components analysis was performed to examine the main treatment effects, and to exclude the presence of confounding batch effects, using the base R function prcomp(). Samples were normalized using calcNormFactors(method = 'TMM') from edgeR and $\log_2$-transformed using voom(). Next, a design matrix was specified to fit coefficients for the CRISPR knockouts and presence or absence of FBS, and an interaction term to examine differences in the FBS effect in the mutant backgrounds. Differentially expressed genes were extracted using topTable() with $\log_2$(fold change) > 0.58 and adjusted $P < 0.05$.

qRT–PCR analysis. HAP1 WT, *FASN*-KO and *C12orf49*-KO cells were cultured in minimal DMEM for 48 h and either control treated or serum starved for 4 h, as indicated. RNA was extracted using the RNeasy Kit (Qiagen), according to the manufacturer's instructions. RNA was converted into complementary DNA using the cVilo master mix (Thermo Fisher Scientific) according to the manufacturer's instructions. The cDNA was amplified and quantified by qPCR using a Bio-Rad CFX96 real-time PCR detection system (Bio-Rad) and using the Maxima SYBR Green PCR master mix (Thermo Fisher Scientific) according to manufacturer's instructions. Transcript levels were normalized to *GAPDH* (see Supplementary Table 8 for primer sequences).

Metabolite profiling. HAP1 WT and *FASN*-KO cells were cultured in minimal medium for 3 d. Cells were washed twice in warm PBS and subsequently flash frozen on liquid nitrogen. Cells were scraped in chilled extraction solvent (40% acetonitrile:40% methanol:20% water, all high-performance liquid chromatography (HPLC) grade), transferred to clean tubes and shaken for 1 h at 4 °C, and were subsequently centrifuged at 4 °C at 14,000 r.p.m. for 10 min. The supernatants were transferred to a clean tube and dried in a SpeedVac, then stored at −80 °C until MS analysis. Samples were reconstituted in water containing internal standards [D_7]glucose and [^{13}C,^{15}N]tyrosine and were subjected twice to HPLC (Dionex Corporation) for positive- and negative-mode analysis using a reverse-phase column (Inertsil ODS-3, 4.6-mm internal diameter, 150-mm length and 3-µM particle size). In positive-mode analysis, the mobile-phase gradient ramped from 5% to 90% acetonitrile in 16 min, remained for 1 min at 90% and then returned to 5% acetonitrile in 0.1% acetic acid over 2 min. In negative mode, the acetonitrile composition ramped from 5% to 90% in 10 min, remained for 1 min at 90% and then returned to 5% acetonitrile in mobile phase (0.1% tributylamine, 0.03% acetic acid and 10% methanol). The total runtime in both the positive and negative modes was 20 min, the samples were maintained at 4 °C and the injection volume was 10 µL. An automated washing procedure was included before and after each sample to avoid any sample carryover.

The eluted metabolites were analysed at the optimum polarity in MRM mode on an electrospray ionization (ESI) triple-quadrupole mass spectrometer (AB Sciex 5500 Qtrap). The MS data-acquisition time for each run was 20 min, and the dwell time for each MRM channel was 10 ms. Mass spectrometric parameters were as previously published[64]. Metabolite peak areas were determined using Multiquant software (SCIEX), normalized to internal standard in each mode yielding an area ratio and then further normalized to total cell number for each sample; malonyl-CoA levels were further normalized to those in WT cells.

Guide mapping and quantification. FASTQ files from single-read sequencing runs were first trimmed by locating constant sequence anchors and extracting the 20-bp gRNA sequence preceding the anchor sequence. Pre-processed paired reads were aligned to a FASTA file containing the TKOv3 library sequences using Bowtie (v0.12.8), allowing up to 2 mismatches and 1 exact alignment (specific parameters: -v2 -m1 -p4 --sam-nohead). Successfully aligned reads were counted and merged along with annotations into a matrix.

Scoring of quantitative GIs with the qGI score. To identify and quantify GIs, genome-wide CRISPR–Cas9 screens were performed using the TKOv3 gRNA library in HAP1 co-isogenic cell lines. Co-isogenic KO query cell lines were obtained from Horizon Genomics (see 'Cell culture') or were generated by introducing mutations in target genes of interest (see above) in the parental HAP1 cells, which we consider to be WT. The TKOv3 library contains 71,090 gRNAs that target ~18,000 human protein-coding genes, most of them with four sequence-independent gRNAs[19]. To quantify GIs, LFCs between read-depth-normalized gRNA abundance in the starting population (T0) and at the end point (T18) were computed. Matched T0 measurement assured that differences between screens during library infection and puromycin selection would not result in false-positive GIs. Matched T0 were stabilized using the median across many T0 measurements (common T0), and those two estimates were combined in a weighted fashion to minimize correlation between GI scores and residual T0 (matched T0 – common T0).

gRNA-level residual scores were derived for a given genetic background by estimating a non-interacting model between LFC values in this background and 21 WT HAP1 backgrounds. To do so, for each WT-KO screen pair, the population of LFC values were *M*-*A*-transformed, which contrasts the per-gRNA LFC difference (*M*) with per-gRNA mean (*A*). A LOESS regression was fitted, which was additionally locally stabilized by binning the data along *A*, and considered equal bin sizes and equal numbers of data points in every bin. For each gRNA, this resulted in 21 residual scores, which represent the contrasts of a given KO with the 21 WT HAP1 screen. Under the assumption that GIs are sparse and that experimental artefacts such as batch effects would introduce additional signal into the population of residual values, we computed a weighted mean of its 21 residual scores by giving a higher weight to WT HAP1 screens with lower absolute residual mean of all 71,000 gRNAs. We refer to the resulting value for each gRNA as the 'guide-level' GI score. Those guide-level GI scores were further normalized. First, locally defined shifts towards negative or positive scores were identified and normalized on the basis of genome location of the target genes. Next, to remove unwanted effects that would arise from screen-to-screen variability, we quantified guide-level GI scores for each of the 21 WT HAP1 screens by contrasting a given WT screen with the remaining WT screens (as described for the KO–WT comparison above). Patterns that explain substantial variance among these WT guide-level GI scores are likely to correspond with unwanted experimental artefacts. To remove these artefacts from the GI data, we performed singular value decomposition (SVD) on guide-level GI scores of the HAP1 WT screens only. We then projected guide-level GI scores onto the left singular vectors, and subtracted the resulting signal from the GI scores.

Finally, we computed gene-level GI scores. First, gRNAs were excluded when their guide-level GI profile disagreed with those of the remaining gRNAs against the same gene. Specifically, the mean within-gene guide-level GI profile Pearson correlation coefficient was computed, which we call the gRNA quality score (QS) and a per-gene quality score (gQS) was computed using the median of the gRNA QS for a given gene. For each gene, we tested three criteria and excluded the gRNA with the lowest QS, if true. First, a gQS above a selected threshold (th1) indicated that sufficient signal was present in the guide-level GI profiles. Second, a difference between the lowest QS and the gQS above a selected threshold (th2) indicated sufficient within-gene disagreement. Third, a QS of the gRNA with the lowest QS for a given gene below a selected threshold (th3) indicated a lack of gene-specific signal. On the basis of these criteria, we excluded a gRNA for 648 out of the 17,804 genes. All remaining guide-level GI values per gene were mean-summarized, and their significance was computed using limma's moderated *t*-test followed by Benjamini–Hochberg multiple-testing correction.

Screen reproducibility analysis. Reproducibility of the gRNA library screening data in FASN-KO cells was tested across three independent screens (referred to as A, B and C). The three screens were started from independent infections with lentivirus packaged gRNA library and performed as described above. To assess reproducibility of fitness effects, a LFC quantifying the drop out between T0 (after puromycin selection) and T18 (end point) was computed for each gene by mean-summarizing the respective four gRNA LFC values (the TKOv3 library targets each gene with four gRNAs). The Pearson correlation coefficients (PCCs) were computed between LFC values of all three pairs of independent replicates.

Our experiments were designed to quantify fitness-effect differences due to the introduction of a specific mutation into an otherwise isogenic background (that is, GIs). To assess reproducibility of GIs, PCCs were computed between qGI values of all pairs of independent replicates. Although we assessed screen reproducibility primarily at the gene level, we also computed the gRNA-level GI PCC between the replicates as reference (Extended Data Fig. 1g).

To test reproducibility of genes, each gene's contribution to the covariance between a pair of *FASN*-KO screens was computed and divided by the product of s.d. of both given screens. The resulting three pairwise (for replicates A–B, A–C, B–C) gene-level scores were mean-summarized to a FASN qGI reproducibility score.

Reproducibility analysis of FASN interactions. We used an Markov chain Monte Carlo (MCMC)-based approach to measure the reproducibility of *FASN* GIs. Specifically, we first independently scored the three independent *FASN* replicate screens and applied an FDR threshold of FDR 50% to generate positive and negative GI profiles for each of the three screens. MCMC was then used to jointly infer false-negative and false-positive rates, as well as a binary consensus *FASN* GI profile (separately for positive and negative GI). Then, using this consensus profile as a standard for evaluation (assuming pairs with posterior probability of interaction of >0.5 as positives), we measured precision and recall statistics (averaged across the three screens) at two different cut-offs: a 'standard' cut-off (absolute qGI score > 0.5 and FDR 50%) and a 'stringent' cut-off (absolute qGI score > 0.7; FDR 20%).

Precision–recall analysis. To control quality of genome-wide gRNA screens, gene-level fitness effects were estimated by computing a LFC quantifying the drop out between T0 (after puromycin selection) and T18 (end point) for each gene and mean-summarizing the respective four gRNA LFC values (the TKOv3 library targets each gene with four gRNAs). Gold-standard essential (reference) and non-essential (background) gene sets were taken from Hart et al.[65] and Hart et al.[19]. Per screen, the separation between reference and background genes according to the LFC values was assessed by computing precision over true positive statistics.

Functional evaluation of GIs. To calculate the enrichment of metabolic GIs in different functional standards, we separated the metabolic GIs in two different sets: all (background) GI scores and high-confidence (reference) GI (FDR < 0.5; |qGI| ≥ 0.5). Then we calculated the fold enrichment of the reference set against the background set in a particular functional standard. First, we computed the overlap of metabolic GI pairs as co-annotations in the standard. Then, we divided the overlap density of the background set into the overlap density of the reference set to determine the fold enrichment. Once we got the fold enrichments, we calculated *P* values on the actual overlap counts of the reference and background sets according to hypergeometric tests. We used four different functional standards: Human functional network[66], GO biological processes[67], Pathway (Canonical pathways from ref. [68]) and HumanCyc[27].

Gene ontology enrichment analysis. GO enrichment analyses for the *FASN* and *C12orf49* GI screen, RNA sequencing and the BioID experiments were performed using the gProfileR R package (v 0.7.0) using the GO-Bioprocesses, GO-Molecular Function and Reactome pathway standards[69]. For the GI screens, enrichment analysis was performed for significant negative GIs (qGI < −0.5; FDR < 0.5) and enriched pathways ($P < 0.05$; minimum term (gene set) size, 3; maximum term size, 100) with a similarity of >50% were collapsed using the Cytoscape Enrichment Map function (v 3.5.1), with the AutoAnnotate plugin (v1.3) used to find pathway themes. For RNA sequencing, GO enrichment was performed as described above for differentially expressed genes (see 'RNA sequencing'), and enriched pathways ($P < 0.01$) with a similarity > 75% were collapsed. For the BioID experiments, GO-Cellular compartment, and KEGG standards were also included in the analysis. Enrichment analysis was performed for significant hits (FDR < 0.01), and enriched pathways ($P < 0.05$) with a similarity of >70% were collapsed. For pathways with at least five collapsed terms, the mean percentage overlap of hits within the theme were visualized on a bar plot.

Co-occurrence of missense mutations and gene essentiality. Starting from the DepMap (19Q2) dataset, we first filtered down to a subset of cell lines (523) that had no more than 1,000 called missense mutations. Among this set, we identified the subset of cell lines (6 total) that exhibited dependency on *C12orf49* (dependency scores threshold of <−1.0). Then, for each gene with at least one missense mutation with the candidate cell lines, we tested for association between the presence of a missense mutation in the corresponding gene and the dependency on *C12orf49*, using the hypergeometric distribution to assess significance and Benjamini–Hochberg multiple-testing correction. For *FASN* mutations, we observed 30 cell lines with at least 1 mutation with *FASN*, and 3 of this set were also dependent on *C12orf49* (out of 6) (FDR = 18.9%).

Statistical analysis. For all experiments, the number of technical and/or biological replicates is listed in the Figure legends or text. Unless otherwise indicated, statistical significance was assessed via one- or two-way ANOVA with Fisher's least significant difference test. Statistical analyses were performed using GraphPad Prism 8 (GraphPad Software) or the R language programming environment.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The datasets generated and analysed in this study are included in the manuscript. The raw FASTQ files for the sequencing data are available upon request and have also been deposited to the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/): RNA-sequencing data, GSE147770; CRISPR screen sequencing data, GSE148627. All mass spectrometry data have been deposited to the MassIVE repository (https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp) and assigned the accession number MSV000085005. The ProteomeXchange accession is PXD017719. Descriptions of the analyses, tools and algorithms are provided in the Methods and the Reporting Summary. Custom code for generating gRNA counts from FASTQ files and code for generating qGI scores is available on GitHub (https://github.com/csbio/metabolicGIN). Source data for Fig. 5 and Extended Data Figs. 1 and 6 are presented with the paper.

Received: 8 November 2019; Accepted: 23 April 2020;
Published online: 1 June 2020

References

1. Chen, M. & Huang, J. The expanded role of fatty acid metabolism in cancer: new aspects and targets. *Precis. Clin. Med* **2**, 183–191 (2019).
2. Chen, R. R. et al. Targeting of lipid metabolism with a metabolic inhibitor cocktail eradicates peritoneal metastases in ovarian cancer cells. *Commun. Biol* **2**, 281 (2019).
3. Imoto, M. Chemistry and biology for the small molecules targeting characteristics of cancer cells. *Biosci. Biotechnol. Biochem.* **83**, 1–10 (2018).
4. Menendez, J. A. & Lupu, R. Fatty acid synthase (FASN) as a therapeutic target in breast cancer. *Expert Opin. Ther. Targets* **21**, 1001–1016 (2017).
5. Garber, K. Cancer anabolic metabolism inhibitors move into clinic. *Nat. Biotechnol.* **34**, 794–795 (2016).
6. Röhrig, F. & Schulze, A. The multifaceted roles of fatty acid synthesis in cancer. *Nat. Rev. Cancer* **16**, 732–749 (2016).
7. Jones, S. F. & Infante, J. R. Molecular pathways: fatty acid synthase. *Clin. Cancer Res.* **21**, 5434–5438 (2015).
8. Benjamin, D. I. et al. Diacylglycerol metabolism and signaling is a driving force underlying FASN inhibitor sensitivity in cancer cells. *ACS Chem. Biol.* **10**, 1616–1623 (2015).
9. Birsoy, K. et al. Metabolic determinants of cancer cell sensitivity to glucose limitation and biguanides. *Nature* **508**, 108–112 (2014).
10. Birsoy, K. et al. An essential role of the mitochondrial electron transport chain in cell proliferation is to enable aspartate synthesis. *Cell* **162**, 540–551 (2015).
11. Costanzo, M. et al. A global genetic interaction network maps a wiring diagram of cellular function. *Science* **353**, aaf1420 (2016).
12. Fischer, B. et al. A map of directional genetic interactions in a metazoan cell. *eLife* **4**, e05464 (2015).
13. Harrison, R., Papp, B., Pál, C., Oliver, S. G. & Delneri, D. Plasticity of genetic interactions in metabolic networks of yeast. *Proc. Natl Acad. Sci. USA* **104**, 2307–2312 (2007).
14. Szappanos, B. et al. An integrated approach to characterize genetic interaction networks in yeast metabolism. *Nat. Genet.* **43**, 656–662 (2011).
15. Wright, A. V., Nuñez, J. K. & Doudna, J. A. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* **164**, 29–44 (2016).
16. Doench, J. G. Am I ready for CRISPR? A user's guide to genetic screens. *Nat. Rev. Genet.* **19**, 67–80 (2018).
17. Currie, E., Schulze, A., Zechner, R., Walther, T. C. & Farese, R. V. Cellular fatty acid metabolism and cancer. *Cell Metab.* **18**, 153–161 (2013).
18. Carette, J. E. et al. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature* **477**, 340–343 (2011).
19. Hart, T. et al. Evaluation and design of genome-wide CRISPR/SpCas9 knockout screens. *G3 (Bethesda)* **7**, 2719–2727 (2017).
20. Billmann, M., Chaudhary, V., ElMaghraby, M. F., Fischer, B. & Boutros, M. Widespread rewiring of genetic networks upon cancer signaling pathway activation. *Cell Syst.* **6**, 52–64(2018).
21. Costanzo, M. et al. The genetic landscape of a cell. *Science* **327**, 425–431 (2010).
22. Ryczko, M. C. et al. Metabolic reprogramming by hexosamine biosynthetic and Golgi N-glycan branching pathways. *Sci. Rep.* **6**, 23043 (2016).
23. Cheng, C. et al. Glucose-mediated *N*-glycosylation of SCAP is essential for SREBP-1 activation and tumor growth. *Cancer Cell* **28**, 569–581 (2015).
24. Wang, S. et al. Site-specific *O*-glycosylation of members of the low-density lipoprotein receptor superfamily enhances ligand interactions. *J. Biol. Chem.* **293**, 7408–7422 (2018).
25. Obaidat, A., Roth, M. & Hagenbuch, B. The expression and function of organic anion transporting polypeptides in normal tissues and in cancer. *Annu. Rev. Pharmacol. Toxicol.* **52**, 135–151 (2012).
26. Horton, J. D., Goldstein, J. L. & Brown, M. S. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J. Clin. Invest.* **109**, 1125–1131 (2008).
27. Romero, P. et al. Computational prediction of human metabolic pathways from the complete human genome. *Genome Biol.* **6**, R2 (2005).
28. Scott, C. C., Vossio, S., Rougemont, J. & Gruenberg, J. TFAP2 transcription factors are regulators of lipid droplet biogenesis. *eLife* **7**, e36330 (2018).
29. Shimano, H. & Sato, R. SREBP-regulated lipid metabolism: convergent physiology — divergent pathophysiology. *Nat. Rev. Endocrinol.* **13**, 710–730 (2017).
30. Zhou, Z., Zhou, J., Su, Z. & Gu, X. Asymmetric evolution of human transcription factor regulatory networks. *Mol. Biol. Evol* **31**, 2149–2155 (2014).
31. VanderSluis, B. et al. Genetic interactions reveal the evolutionary trajectories of duplicate genes. *Mol. Syst. Biol.* **6**, 429 (2010).
32. Ascencio, D., Ochoa, S., Delaye, L. & DeLuna, A. Increased rates of protein evolution and asymmetric deceleration after the whole-genome duplication in yeasts. *BMC Evol. Biol.* **17**, 40 (2017).
33. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).
34. Mitchell, A. L. et al. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **47**, D351–D360 (2019).
35. Uhlen, M. et al. Tissue-based map of the human proteome. *Science* **347**, 1260419–1260419 (2015).
36. Nagy, Á., Lánczky, A., Menyhárt, O. & Győrffy, B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Sci. Rep.* **8**, 9227 (2018).
37. Meyers, R. M. et al. Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (2017).
38. Behan, F. M. et al. Prioritization of cancer therapeutic targets using CRISPR–Cas9 screens. *Nature* **568**, 511–516 (2019).

39. Klarin, D. et al. Genetics of blood lipids among ~300,000 multi-ethnic participants of the million veteran program. *Nat. Genet.* **50**, 1514–1523 (2018).
40. Brown, M. S. & Goldstein, J. L. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* **89**, 331–340 (1997).
41. Go, C. D. et al. A proximity biotinylation map of a human cell. Preprint at *bioRxiv* https://doi.org/10.1101/796391(2019).
42. Costanzo, M. et al. Global genetic networks and the genotype-to-phenotype relationship. *Cell* **177**, 85–100 (2019).
43. Horlbeck, M. A. et al. Mapping the genetic landscape of human cells. *Cell* **174**, 953–967.e22 (2018).
44. Najm, F. J. et al. Orthologous CRISPR–Cas9 enzymes for combinatorial genetic screens. *Nat. Biotechnol.* **36**, 179–189 (2018).
45. Han, K. et al. Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. *Nat. Biotechnol.* **35**, 463–474 (2017).
46. Norman, T. M. et al. Exploring genetic interaction manifolds constructed from rich single-cell phenotypes. *Science* **365**, 786–793 (2019).
47. Shen, J. P. et al. Combinatorial CRISPR–Cas9 screens for de novo mapping of genetic interactions. *Nat. Methods* **14**, 573–576 (2017).
48. Gonatopoulos-Pournatzis, T. et al. Genetic interaction mapping and exon-resolution functional genomics with a hybrid Cas9–Cas12a platform. *Nat. Biotechnol.* **38**, 638–648 (2020).
49. Brenner, A. et al. Heavily pre-treated breast cancer patients show promising responses in the first in human study of the first-In-class fatty acid synthase (FASN) inhibitor, TVB-2640 in combination with paclitaxel. in *Cancer Res.* **77**, abstr. P6-11-09 (2017).
50. Gimeno, R. E. Fatty acid transport proteins. *Curr. Opin. Lipidol.* **18**, 271–276 (2007).
51. Boyle, E. A., Pritchard, J. K. & Greenleaf, W. J. High-resolution mapping of cancer cell networks using co-functional interactions. *Mol. Syst. Biol.* **14**, e8594 (2018).
52. Kim, E. et al. A network of human functional gene interactions from knockout fitness screens in cancer cells. *Life Sci. Alliance* **2**, e201800278 (2019).
53. Luciano, M. et al. Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nat. Genet.* **50**, 6–11 (2018).
54. Kichaev, G. et al. Leveraging polygenic functional enrichment to improve GWAS power. *Am. J. Hum. Genet.* **104**, 65–75 (2019).
55. Nagel, M., Watanabe, K., Stringer, S., Posthuma, D. & van der Sluis, S. Item-level analyses reveal genetic heterogeneity in neuroticism. *Nat. Commun.* **9**, 905 (2018).
56. Aregger, M., Chandrashekhar, M., Tong, A. H. Y., Chan, K. & Moffat, J. Pooled lentiviral CRISPR–Cas9 screens for functional genomics in mammalian cells. *Methods Mol. Biol.* **1869**, 169–188 (2019).
57. Hesketh, G. G., Youn, J. Y., Samavarchi-Tehrani, P., Raught, B. & Gingras, A. C. Parallel exploration of interaction space by BioID and affinity purification coupled to mass spectrometry. *Methods Mol. Biol* **1550**, 115–136 (2017).
58. Branon, T. C. et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887 (2018).
59. Knight, J. D. R. et al. ProHits-viz: a suite of web tools for visualizing interaction proteomics data. *Nat. Methods* **14**, 645–646 (2017).
60. Liu, G. et al. Data independent acquisition analysis in ProHits 4.0. *J. Proteomics* **149**, 64–68 (2016).
61. Deutsch, E. W. et al. Trans-Proteomic pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. *Proteomics. Clin. Appl.* **9**, 745–754 (2015).
62. Teo, G. et al. SAINTexpress: improvements and additional features in significance analysis of INTeractome software. *J. Proteomics* **100**, 37–43 (2014).
63. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
64. Abdel Rahman, A. M., Ryczko, M., Pawling, J. & Dennis, J. W. Probing the hexosamine biosynthetic pathway in human tumor cells by multitargeted tandem mass spectrometry. *ACS Chem. Biol.* **8**, 2053–2062 (2013).
65. Hart, T. et al. High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* **163**, 1515–1526 (2015).
66. Greene, C. S. et al. Understanding multicellular function and disease with human tissue-specific networks. *Nat. Genet.* **47**, 569–576 (2015).
67. Ashburner, M. et al. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
68. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
69. Raudvere, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).

Acknowledgements

We thank members of the Moffat lab for helpful discussions. Q. Huang, M. Olivieri, C. Sheene, R. Akthar and S. Sidhu are gratefully acknowledged for assistance with molecular biology experiments. Next-generation sequencing was performed at the Donnelly Sequencing Centre at the University of Toronto. Proteomics work was performed at the Network Biology Collaborative Centre at the Lunenfeld-Tanenbaum Research Institute, a facility supported by Canada Foundation for Innovation funding, by the Ontarian Government and by Genome Canada and Ontario Genomics (OGI-097, OGI-139). M.A. was supported by a Swiss National Science Foundation Postdoctoral Fellowship; K.A.L. was supported by a Vanier Canada Graduate Scholarship and Studentship award from the Kidney Cancer Research Network of Canada. M.B. was supported by a DFG Fellowship (Bi 2086/1-1). This research was funded by grants from the Canadian Institutes for Health Research (J.M., C.B, B.J.A. and A.-C.G.), Ontario Research Fund (B.J.A., C.B. and J.M) and Canada Research Chairs Program (J.M., C.B. and A.-C.G.). C.L.M., M.B. and M.R. are partially supported by grants from the National Science Foundation (MCB 1818293) and the National Institutes of Health (R01HG005084, R01HG005853).

Author contributions

Conceptualization and design of the study: M.A., K.A.L. and J.M.; Experimental investigation: M.A., K.A.L., A.H.Y.T., K.C., L.N., O.S., A.H., J.P., Z.-Y.L., H.A., C.J.W. and A.W. Data analysis: M.A., K.L., M.B., M.C., M.R., K.R.B., C.R., M.U., P.M., J.W.D., A.C.G, J.L.M. and J.M.; Writing and editing: M.A., K.A.L., M.B., M.C., B.J.A., C.B., C.L.M. and J.M. with input from other authors; Supervision: J.W.D., A.C.G., C.L.M., B.J.A., C.B. and J.M.; Funding acquisition: C.L.M., B.J.A., C.B. and J.M.

Competing interests

J.M., B.J.A. and C.B are shareholders in Northern Biologics. J.M. is a shareholder in Pionyr Immunotherapeutics, is acting chief scientific officer and shareholder in Empirica Therapeutics and is an SAB member and shareholder of Aelian Biotechnology. C.B. is an SAB member of Yumanity Therapeutics.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s42255-020-0211-z.

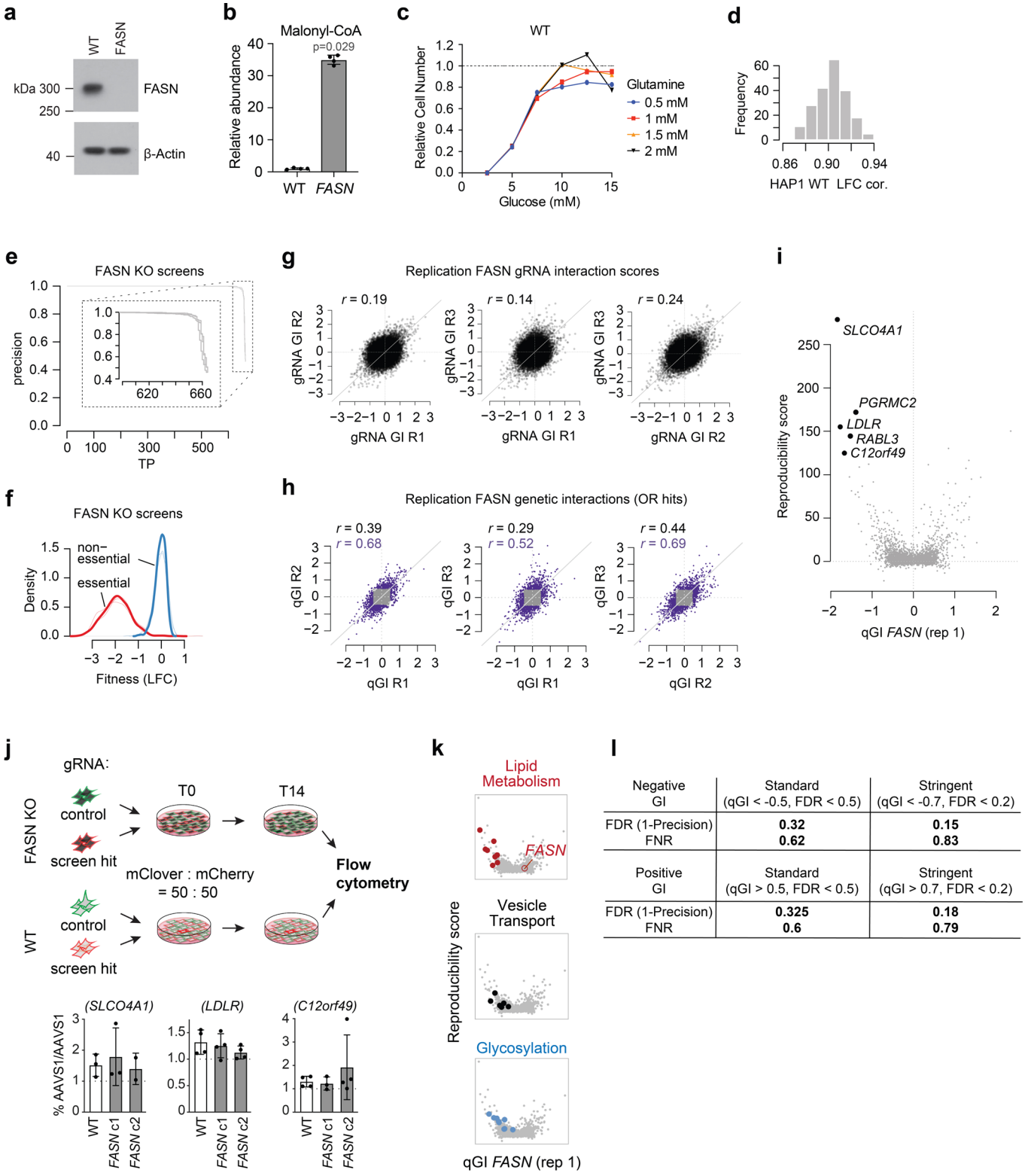
Supplementary information is available for this paper at https://doi.org/10.1038/s42255-020-0211-z.

Correspondence and requests for materials should be addressed to C.L.M., B.J.A., C.B. or J.M.

Peer review information Primary Handling Editor: Pooja Jha.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

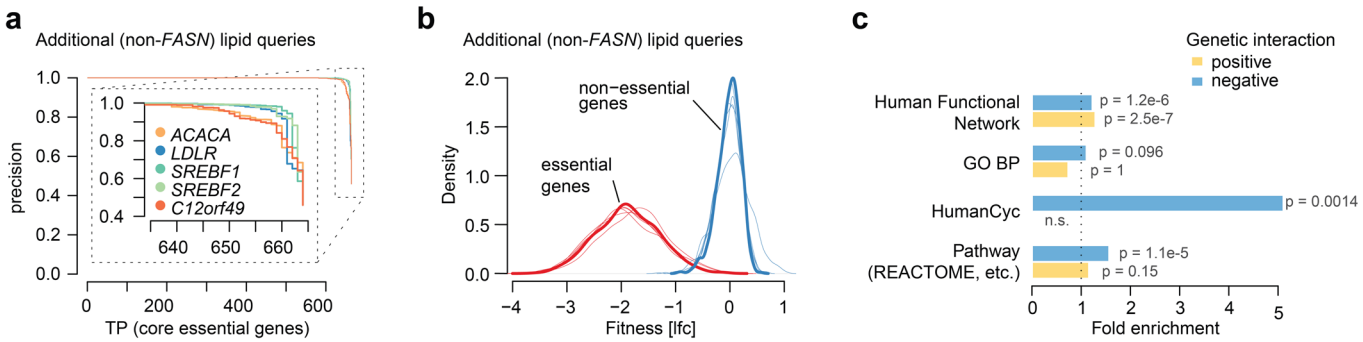


l

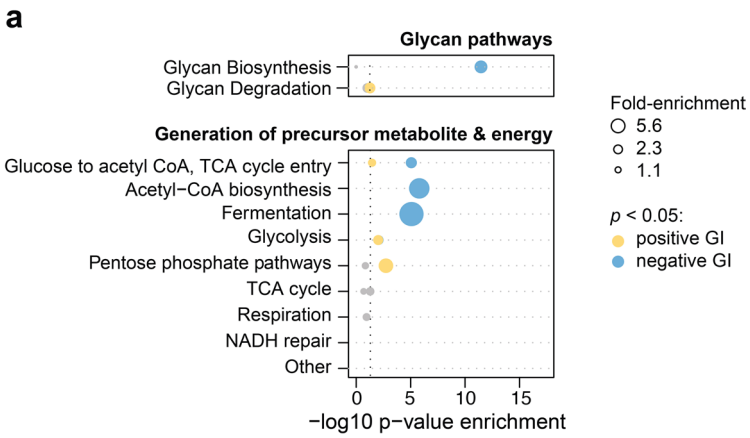
Negative GI	Standard (qGI < -0.5, FDR < 0.5)	Stringent (qGI < -0.7, FDR < 0.2)
FDR (1-Precision)	**0.32**	**0.15**
FNR	**0.62**	**0.83**
Positive GI	**Standard (qGI > 0.5, FDR < 0.5)**	**Stringent (qGI > 0.7, FDR < 0.2)**
FDR (1-Precision)	**0.325**	**0.18**
FNR	**0.6**	**0.79**

Extended Data Fig. 1 | See next page for caption.

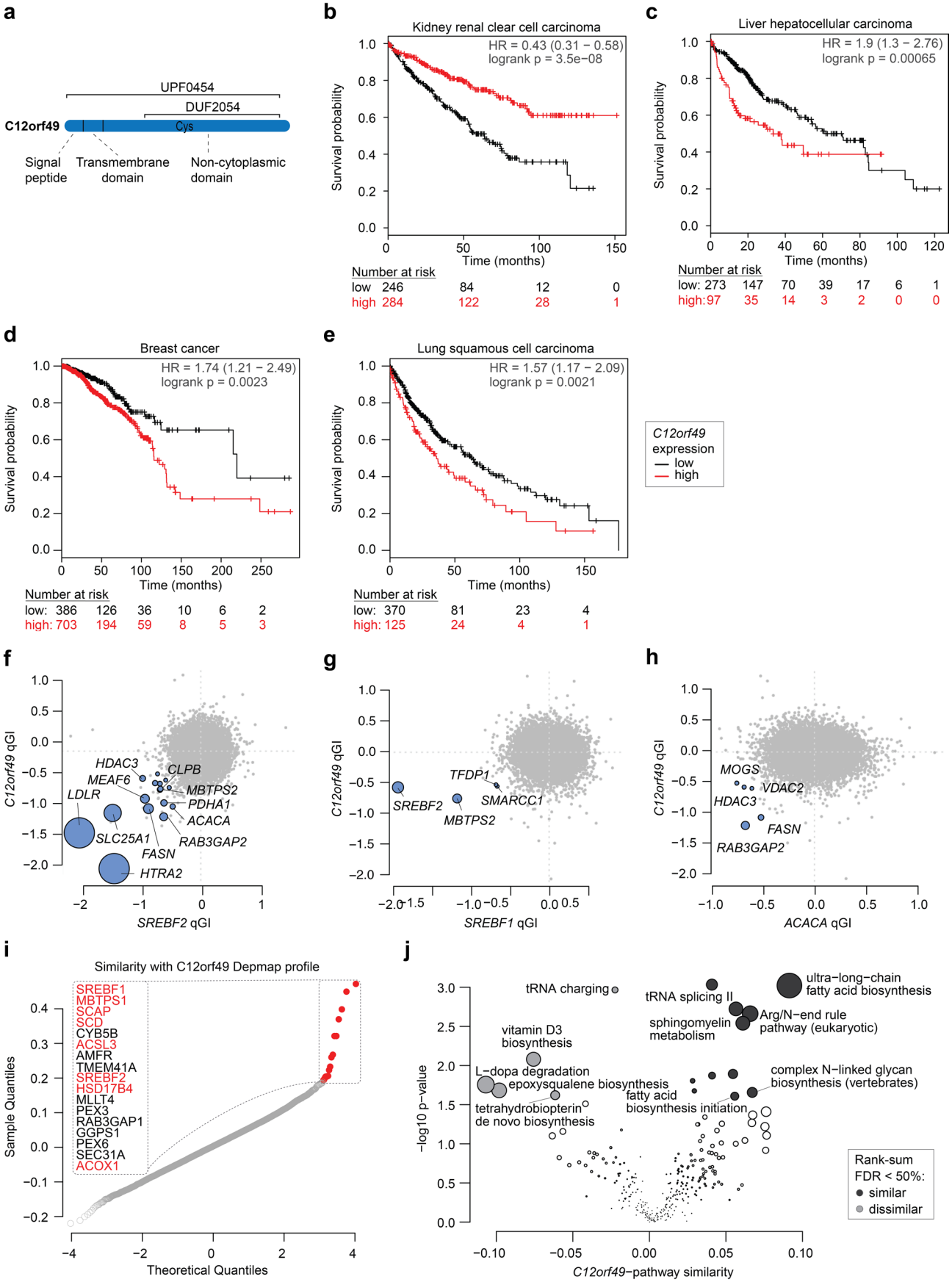
Extended Data Fig. 1 | Validation of *FASN*-KO cells and genetic interactions screens. a, Western blot depicting FASN and β-Actin levels in HAP1 parental wildtype (WT) and *FASN*-KO cells. Representative data from three biologically independent experiments. **b**, Bar plot depicting malonyl-CoA levels in HAP1 WT and *FASN*-KO cells as detected by mass spectrometry-based metabolite profiling, normalized to parent HAP1 WT cells. Data are represented as means ± standard deviation; n = four biologically independent experiments; two-tailed Mann Whitney U test. **c**, Growth curves of HAP1 WT cells depicting relative cell numbers over 3 days, plotted as a function of glucose concentration in mM, in either 0.5 mM (blue), 1 mM (red), 1.5 mM (yellow), or 2 mM (black) glutamine. **d**, Histogram showing a frequency distribution of all pairwise Pearson correlation coefficients for LFC values (T0/T18) of the 21 WT HAP1 screens. **e**, Precision–recall curves for the three CRISPR replicate screens in HAP1 *FASN*-KO cells using the reference core essential gene set (CEG2) defined in Hart et al., 2017[19]. **f**, Fitness effect (log2 fold-change, LFC) distributions for reference core essential (CEG2) and non-essential gene sets defined in Hart et al., 2017[19] across the three *FASN*-KO query screens. **g**, Agreement of gRNA-level genetic interaction scores with *FASN*. Scatter plots show all possible pairwise combinations of three biological replicate screens. The Pearson correlation coefficient (*r*), based on comparison values for 70,152 gRNAs. **h**, Agreement of *FASN* quantitative genetic interactions (qGIs). Scatter plots show gene-level *FASN* genetic interactions (qGI scores) derived from all possible pairwise combinations of three biological replicate screens. The Pearson correlation coefficient (*r*), based on comparison of all qGI scores (r shown in grey, calculated on all the grey and purple data points in the scatter plots), or only genetic interactions that exceed a given significance threshold (|qGI| > 0.5, FDR < 0.5) in either one screen (logic OR; purple). **i**, Scatter plot showing reproducibility scores as a function of qGI scores for a single FASN-KO screen (replicate A). Pairwise reproducibility of a qGI score was calculated by computing the contribution of each of the 17,804 genes to the covariance between a pair of screens divided by the sum of standard deviations. The reproducibility score represents the sum of those values across the three pairwise comparisons. Five genes with highest reproducibility scores and the most negative qGI scores with the *FASN*-KO screen (replicate A) are labelled. **j**, Establishing the *AAVS1* target locus as a good negative control site in HAP1 WT and *FASN*-KO cells. Schematic depicting co-culture validation assays (upper panel). Parental WT and *FASN*-KO cells were stably transduced with color-coded gRNA expression vectors carrying an intergenic control or screen hit-targeting gRNA. Color-coded cells are mixed at an equal ratio, cultured over two weeks and the relative proportion of green and red cells was quantified by flow cytometry. Control co-culture experiments performed in parallel to the validation of hit genes depicted in Fig. 1e as indicated above each barplot (lower panel). Bar plots are depicting the color ratio of cells carrying two colour-coded gRNAs targeting *AAVS1* (intergenic control) across WT and two *FASN*-KO clones as indicated. Experiments were performed with three independent gRNA targeting AAVS1 and using both color orientations. All data are represented as means ± standard deviation; n = three (*LDLR*) or four (*SLCO4A1, C12orf49*) biologically independent experiments. **k**, Scatter plots reproducibility scores as a function of qGI scores for the negative genetic interaction hits depicted in Fig. 1h functioning in lipid uptake and homeostasis (red), vesicle transport genes (black) and glycosylation (blue). **l**, Precision and recall values for GIs with *FASN* measured at the standard (|qGI| > 0.5, FDR < 0.5) and stringent (|qGI| > 0.7, FDR < 0.2) thresholds. Precision and recall values were computed using an MCMC-based approach (see Methods).



Extended Data Fig. 2 | Quality control of genetic interaction screens for fatty acid synthesis-related query genes. a, Precision-recall curves distinguishing the reference core essential gene set (CEG2) defined in Hart et al., 2017[19] and a non-essential gene set in CRISPR screens in five HAP1 knockout query cell lines (*LDLR, C12orf49, SREBF2, ACACA, SREBF1*-KO). **b**, Fitness effect (LFC) distributions for reference core essential (CEG2) and non-essential gene sets defined in Hart et al., 2017[19] across CRISPR screens in five HAP1 KO cell lines (*LDLR, C12orf49, SREBF2, ACACA, SREBF1*). **c**, Bar plot of enrichment of co-annotation as defined by the Human Functional Network, Gene Ontology Bioprocesses (BP), HumanCyc or and aggregation of pathway standards (including REACTOME, KEGG or BIOCARTA) for genetic interactions identified across all six query genome-wide screens (*FASN, LDLR, C12orf49, SREBF2, ACACA, SREBF1*). Enrichment was tested using a hypergeometric test. See methods for details of analysis.

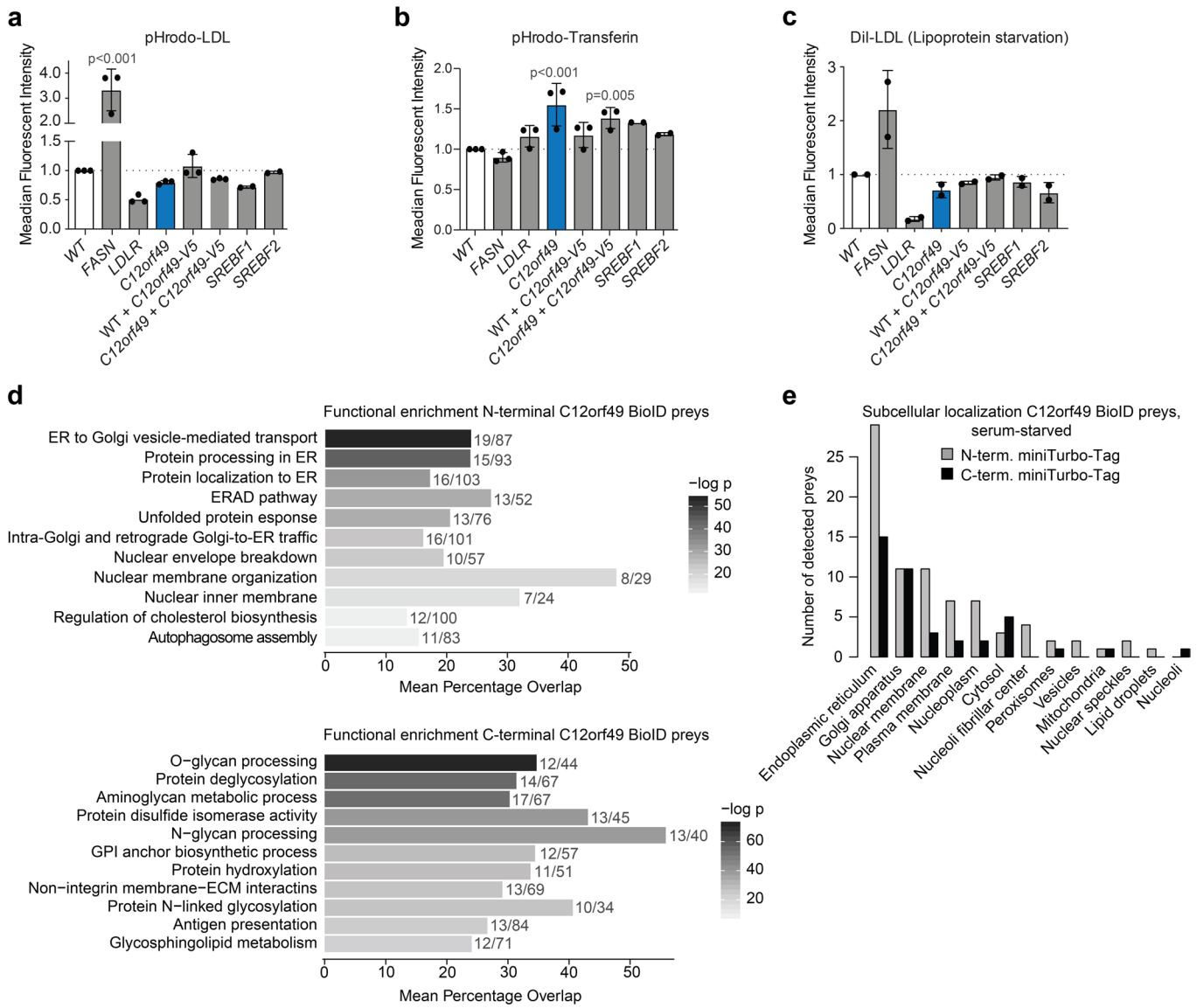


Extended Data Fig. 3 | Pathway enrichment analysis of genetic interactions for fatty acid synthesis-related query genes in additional HumanCyc sub-categories. a, Dot plot of normalized pathway enrichment values for aggregate GIs across the six query genes (*FASN, C12orf49, LDLR, SREBF2, ACACA, SREBF1)* with sub-categories from HumanCyc are indicated. A GI is identified for a query-library pair if the |qGI| > 0.5 and FDR < 0.5. Enrichment for positive (yellow) and negative (blue) GI is tested inside Glycan Pathways and Generation of precursor metabolite and energy HumanCyc branches using a hypergeometric test. Enrichment with p-value < 0.05 are blue (negative GI) and yellow (positive GI). Dot size is proportional to the fold-enrichment in the indicated categories and specified in the legend.

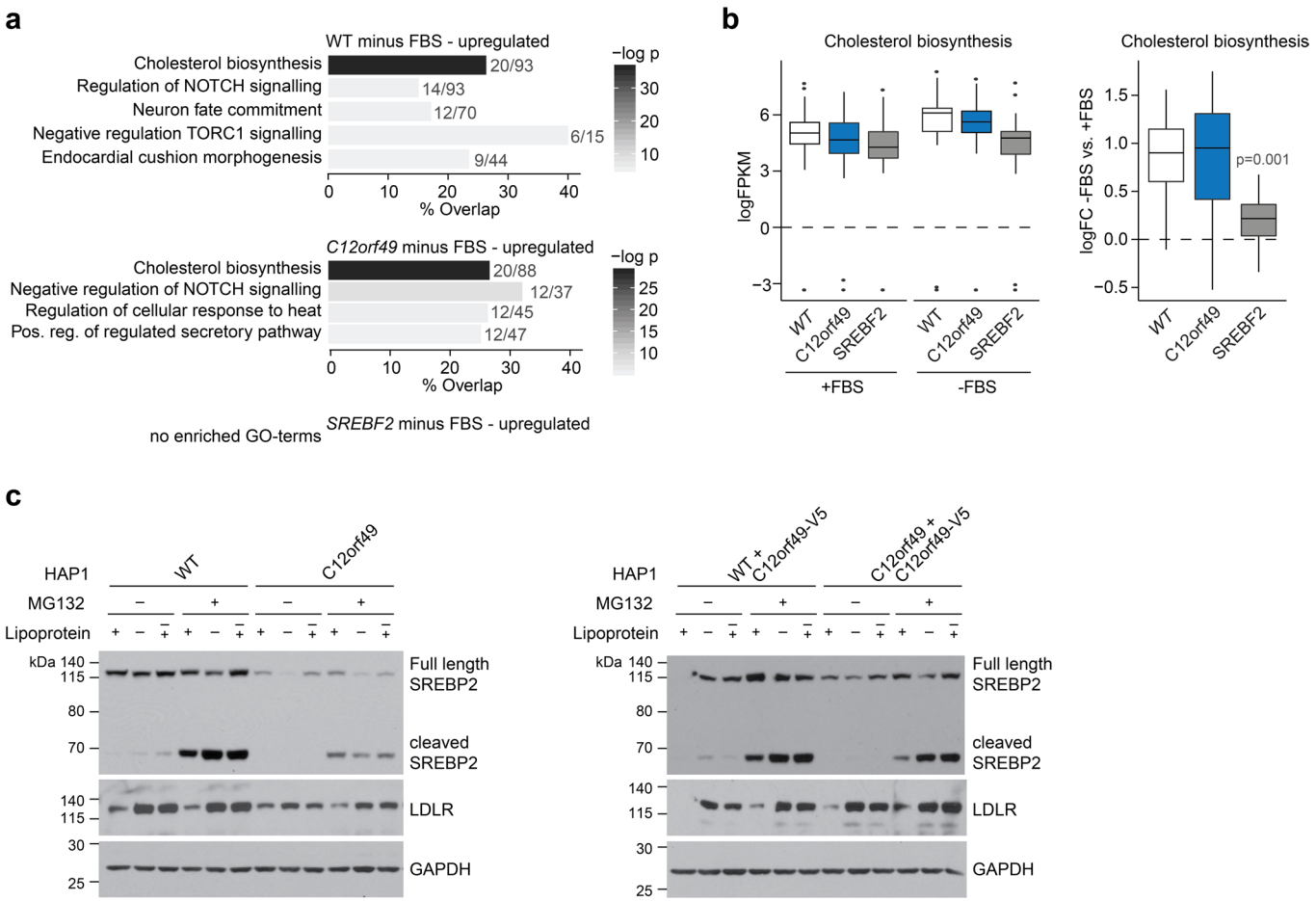


Extended Data Fig. 4 | See next page for caption.

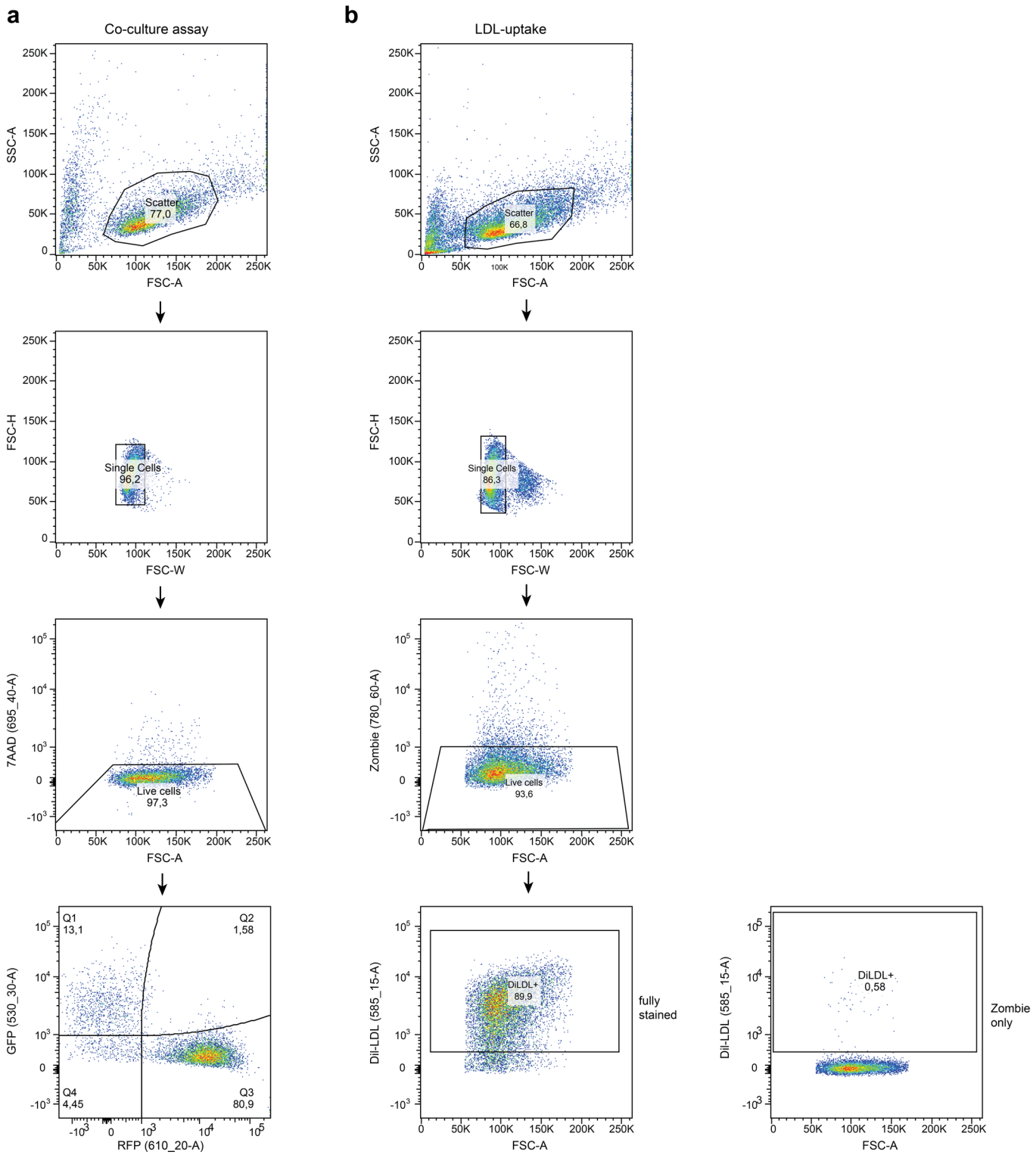
Extended Data Fig. 4 | Overview of C12orf49, cancer associations, and functional correlates. a, Cartoon of C12orf49 protein sequence features and domains. **b–e**, Kaplan Meier survival plots displaying univariate analysis of TCGA data across multiple tumor types including kidney, breast, liver and sarcoma for *C12orf49* high vs. low expressing tumor tissue (www.kmplot.com)[36]. Patient numbers at risk (n) are indicated below each plot; two-sided logrank test. **f-h**, GI overlap between the 17,804 *C12orf49* and *SREBF2, SREBF1* and *ACACA* qGI scores shown as pairwise scatter plots with *C12orf49* as function of *SREBF2* (**f**), *SREBF1* (**g**) and *ACACA* (h). A common negative GI is called if it is significant (qGI < -0.5, FDR < 0.5) in both screens (indicated in blue). The top 10 strongest common GIs and lipid metabolism genes are labelled. **i**, Profile similarity of *C12orf49* across genome-wide DepMap CRISPR/Cas9 screens. Similarity was quantified by taking all pairwise gene-gene Pearson correlation coefficients of CERES score profiles across 563 screens (19Q2 DepMap data release). The distribution of 17,633 CERES profile similarity is plotted as a quantile-quantile plot, and the top 18 most similar out of 17,633 genes are labelled. Genes associated with lipid metabolism are indicated in red. **j**, Pathway analysis of *C12orf49* profile similarity. *C12orf49* profile similarity scores for all 17,634 genes represented in the DepMap were mean-summarized by pathway as defined in the HumanCyc standard[27]. Tendencies towards pathway-level similarity (co-essentiality) and dissimilarity (exclusivity) with *C12orf49* were tested using a two-sided Wilcoxon rank-sum test with multiple-hypothesis correction using the Benjamini and Hochberg procedure.



Extended Data Fig. 5 | Regulation of LDL uptake and LDLR expression by *C12orf49*. a, Bar plots showing the results of a low density lipoprotein (LDL) uptake assay across the indicated HAP1 cell lines using pHrodo-LDL probe. All data are represented as means ± standard deviation; n = two (*SREBF1, SREBF2*) or three (*WT, FASN, LDLR, C12orf49, WT + C12orf49-V5, C12orf49 + C12orf49-V5*) biologically independent experiments; one-way ANOVA. **b**, Bar plots showing the results of a transferin uptake assay across the indicated HAP1 cell lines using pHrodo-Transferin probe. All data are represented as means ± standard deviation; n = two (*SREBF1, SREBF2*) or three (*WT, FASN, LDLR, C12orf49, WT + C12orf49-V5, C12orf49 + C12orf49-V5*) biologically independent experiments; one-way ANOVA. **c**, Bar plots showing the results of a low density lipoprotein (LDL) uptake assay across the indicated in lipoprotein-deprived HAP1 cell lines using Dil-LDL probe. All data are represented as means ± standard deviation; n = two biologically independent experiments. **d**, Pathway enrichment analysis of BioID preys captured with N-terminal (top panel) or C-terminal (bottom panel) miniTurbo-tagged C12orf49 under normal growth conditions using the GO molecular function, biological process, cellular compartments, KEGG and Reactome standards. Terms for significantly enriched gene sets ($p < 0.05$, maximum term size 105) are indicated and bars depict mean percentage overlap with the indicated term. Numbers next to each bar indicate the mean overlap and term sizes, respectively. The greyscale color legend for p-values is indicated on the right. P-values were calculated using gProfileR[69]. **e**, Barplots depicting the number of proteins localizing to indicated cellular compartments for preys captured with N-terminal (grey) or C-terminal (black) miniTurbo-tagged C12orf49 in 293 cells under serum-starvation.



Extended Data Fig. 6 | RNAseq and western blot analysis in response to serum or lipoprotein starvation and upon inhibition of the proteasome. a, Gene ontology enrichment analysis of genes upregulated under serum starvation in HAP1 wildtype (WT), *C12orf49* or *SREBF2*-KO cells using GO molecular functions, GO bioprocesses and Reactome standards. Gene sets with overlapping members have been merged and bars depict mean percentage overlap with the indicated term. Numbers next to each bar indicate the mean overlap and term sizes, respectively. The greyscale color legend for p-values is indicated on the right; p-values were calculated using gProfileR[69]. **b**, Boxplots depicting mean expression and induction of genes assigned with the indicated term across HAP1 WT, *C12orf49* and *SREBF2*-KO cells under normal (+FBS) and serum-starved (-FBS) conditions; n = three biologically independent experiments, two-tailed student's t test. Boxes show IQR, 25th to 75th percentile, with the median indicated by a horizontal line. Whiskers extend to the quartile ± 1.5 x IQR. **c**, Western blotting analysis of SREBP2, LDLR and β-Actin levels across the indicated HAP1 co-isogenic knockout cell lines in response to overnight lipoprotein withdrawal (-) and a short refeeding period (-/+) in presence or absence of the proteasomal inhibitor MG132 (10μM MG132 for 5 hours) as indicated. Unprocessed full length and processed C-terminal SREBP2 products are indicated. Representative data from two biologically independent experiments.



Extended Data Fig. 7 | Gating strategy for flow cytometry experiments. Gating strategies for flow cytometry experiments for **a**, co-culture assays and **b**, LDL/Transferrin uptake assays. In all cases the following steps were taken: 1. Forward scatter area (FSC-A) vs. side scatter area (SSC-A) were used to separate cell events from debris. 2. Forward scatter height (FSC-H) vs. forward scatter width (FSC-W) was used separate single cells from aggregates. 3. Forward scatter area (FSC-A) vs. viability stain (7AAD/B695-40 or Zombie NIR/R780-60) was utilized to gate live cells. For co-culture assays, gating scheme for separating GFP/B530-30 vs RFP/YG610-20 positive cells including steps 1-3 are shown in panel (**a**). For Dil-LDL/YG585-15 quantification, marker-positive live cells were quantified relative to unstained controls following steps 1-3 as displayed panel (**b**).

nature research

Corresponding author(s): Jason Moffat, Charles Boone, Brenda J. Andrews, Chad L. Myers

Last updated by author(s): Mar 30, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Illumina HiSeq2500, Novaseq 6000 (https://www.illumina.com/) Zeiss LSM880 confocal microscope, Zeiss (http://www.zeiss.com/) LSR Fortessa flow cytometer, BD Biosciences (http://www.bdbiosciences.com/) TripleTOF 5600 mass spectrometer, SCIEX (www.sciex.com/) QTRAP 5500 mass spectromter, SCIEX (www.sciex.com/) BioRad CFX96 real-time PCR detection system, BioRad (www.biorad.com/)
Data analysis	Bowtie 0.12.7 & 1.1.1 (https://en.wikipedia.org/wiki/Bowtie_(sequence_analysis)), Bioconductor EdgeR 3.24.3 & 3.26.8 (https://bioconductor.org/packages/release/bioc/html/edgeR.html) Bioconductor Limma v3.32.10 (https://bioconductor.org/packages/release/bioc/html/limma.html) R 3.5 & 3.6 (various functions; https://www.r-project.org/) GraphPad Prism 8.1.1 (https://www.graphpad.com/scientific-software/prism/) ProHits LIMS (https://prohits-web.lunenfeld.ca/) ProHits-viz (https://prohits-viz.lunenfeld.ca/) Mascot 2.3.02 (http://www.matrixscience.com/) Comet 2018.01 rev.4 (http://comet-ms.sourceforge.net/) PeptideProphet (http://peptideprophet.sourceforge.net/) SAINTexpress v3.6.1 (http://saint-apms.sourceforge.net/Main.html) STAR short-read aligner v2.6.0a (https://github.com/alexdobin/STAR) Maltiquant (https://sciex.com/products/software/multiquant-software)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All raw data has been provided in supplemental tables or will be made available upon publication or on request. There is no restriction on data availability.
Links to publicaly available datasets analyzed in this manuscript:
DepMap data (https:/depmap.org)
Tissue expression and subcellular localization (http://www.proteinatlas.org)
Kaplan Meier survival plots of TCGA data (www.kmplot.com)
GWAS catalog (www.ebi.ac.uk/gwas/)
Protein domains (www.ebi.ac.uk/interpro/)

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculations were performed and sample sizes were arbitrarily chosen according to convention in the field. For small-scale experiments, the number of replicates exceed at least 2 biological replicates (= independent experiments) and/or at least 3 technical replicates (= repeated measurements of the same original sample), as indicated in each figure legend. Screens were performed at >200-fold library representation and the mutagenized cell pool was split into 3 replicates post-selection and processed independently in all downstream steps.
Data exclusions	No data were excluded from any experiments and figures shown.
Replication	We present no experimental results that were not reproducible.
Randomization	Screen samples were processed and sequenced in a randomized manner and labelled with numbers instead of sample names.
Blinding	No sample allocation to groups was performed, so blinding was not relevant to this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	anti-FASN (1:2,000, Abcam ab128870, Lot GR196759-6) anti-SREBP2 (1:250, BD Biosciences 557037, Lot 8184855) anti-LDLR (1:250, Abcam ab52818, Lot GR295148-22) anti-β-Actin (1:10,000, Abcam ab8226, Lot GR3206282-1) anti-GAPDH (1:10,000, Santa Cruz 166574, Lot P2018)

anti-V5 (1:1250, Abcam ab27671, Lot GR322548-5)
anti-GOLGA2 (1:250, Sigma HPA021799, Lot A89838)
anti-Rabbit HRP-conjugated (1:5,000, Cell Signalling Technology 7074, Lot 05/2017/27)
anti-Mouse HRP-conjugated (1:5,000, Jackson Immuno Research 715-035-151, Lot NA)
Alexa Fluor488 goat anti-mouse (1:500, Invitrogen A-11001, Lot 1834337)
Alexa Fluor647 anti-rabbit antibodies (1:500, Invitrogen A-21245, Lot 1037285)

Validation

All antibodies are commercially available and tested by manufacturers. In addition, Anti-FASN, SREBP2 and LDLR antibodies were validated by western blotting using HAP1 FASN knockout cells. Anti-V5 antibodies were validated experimentally in presence or absence of V5-tagged constructs by western blotting and immunofluorescence microscopy. GOLGA2 antibodies confirmed expected localization to the Gogli apparatus as determined by immunofluorescence microscopy. The β-Actin and GAPDH antibodies have not been experimentally validated in this study but are routinely used in our lab and across the scientific community.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Human HAP1 wild type cells were obtained from Horizon Genomics (clone C631, sex: male with lost Y chromosome, RRID: CVCL_Y019). The following HAP1 gene knockout cell lines were obtained from Horizon: FASN (HZGHC003700c006 & HZGHC003700c011), ACACA (HZGHC004903c002), LDLR (HZGHC003978c007), SREBF1 (HZGHC001361c012), SREBF2 (HZGHC000683c004). HEK 293T cells were obtained from ATCC. HEK293 Flp-In T-REx 293 were obtained from Invitrogen.

Authentication

HAP1 and 293T cells were authenticated by STR profiling at the Centre for Applied Genomics (TCAG) at the Hospital for Sick Children (SickKids) in Toronto. HAP1 cells were also whole-genome sequenced. All gene knockout cell lines were confirmed to carry the expected out-of-frame insertions or deletions by Sanger Sequencing of PCR products. HEK293 Flp-In T-REx 293 cells were not authenticated.

Mycoplasma contamination

All cell lines were routinely tested and confirmed negative for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

None of the cell lines used in this study is listed as commonly misidentified.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

For co-culture assays of color-coded cell lines, cells were trypsinized, washed and stained for dead cells using Zombie NIR (BioLegend). The relative proportion of red and green cells in the co-culture were assessed using an LSR Fortessa flow cytometer (BD Bioscience).
For particle uptake assays, cells incubated with Dil-LDL (Thermo L3482) for 15 minutes then washed in PBS, trypsinized and stained with 7-AAD (BioLegend 420404) or Zombie NIR (BioLegend 423105) cell viability solution at 25 ng/ml (1:2,000) for 5 minutes at room temperature. Staining was measured using an LSR Fortessa flow cytometer (BD Bioscience).

Instrument

LSR Fortessa: 3 laser (488/640/405) configuration (BD Biosciences) or iQue Screener PLUS Blue-Red configuration (IntelliCyt)

Software

Data were acquired using FACSDIVA software v8.0-8.1 (BD Biosciences) or ForeCyt software v6.2.6752 (IntelliCyt) with automated compensation (performed independently for each experiment with single-stained samples following the software assistant) and analyzed using FlowJo software (RRID:SCR_008520) v10.4.-v10.6.0 (TreeStar).

Cell population abundance

No sorting was applied.

Gating strategy

Gating strategy is described in Figure S6 for (a) co-culture assays, (b) LDL/Transferrin uptake assays and (c) quantification of LDLR surface expression. In all cases the following steps were taken: 1. Forward scatter area (FSC-A) vs. side scatter area (SSC-A) were used to separate cell events from debris. 2. Forward scatter height (FSC-H) vs. forward scatter width (FSC-W) was used separate single cells from aggregates. 3. Forward scatter area (FSC-A) vs. viability stain (7AAD/B695-40 or Zombie NIR/R780-60) was utilized to gate live cells. For co-culture assays, gating scheme for separating GFP/B530-30 vs RFP/YG610-20 positive cells including steps 1-3 are shown in panel (a). For Dil-LDL/YG585-15 quantification, marker-positive live cells were quantified relative to unstained controls following steps 1-3 as displayed panel (b).

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.