Title

The Lipid Droplet Knowledge Portal: A resource for systematic analyses of lipid droplet biology

Journal

Developmental Cell, 57(3)

ISSN

1534-5807

Authors

Mejhert, Niklas

Gabriel, Katlyn R

Frendo-Cumbo, Scott

et al.

Publication Date

2022-02-01

DOI

10.1016/j.devcel.2022.01.003

HHS Public Access

Author manuscript

Dev Cell. Author manuscript; available in PMC 2022 May 24.

Published in final edited form as:

Dev Cell. 2022 February 07; 57(3): 387–397.e4. doi:10.1016/j.devcel.2022.01.003.

The Lipid Droplet Knowledge Portal: A resource for systematic analyses of lipid droplet biology

Niklas Mejhert[1,2,3,§], **Katlyn R. Gabriel**[1,2,4,§], **Scott Frendo-Cumbo**[3], **Natalie Krahmer**[5,6], **Jiunn Song**[1,2], **Leena Kuruvilla**[7], **Chandramohan Chitraju**[1,2], **Sebastian Boland**[1,2], **Dong-Keun Jang**[8], **Marcin von Grotthuss**[8], **Maria C. Costanzo**[8], **Mikael Rydén**[3], **James A. Olzmann**[9,10,11,12], **Jason Flannick**[8,13,14], **Noël P. Burtt**[8], **Robert V. Farese Jr.**[1,2,15,16,§,*], **Tobias C. Walther**[1,2,4,15,16,†,§,*]

[1]Department of Molecular Metabolism, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

[2]Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA

[3]Department of Medicine (H7), Karolinska Institutet, Huddinge, 141 86 Stockholm, Sweden.

[4]Howard Hughes Medical Institute, Boston, MA 02115, USA

[5]Institute for Diabetes and Obesity, Helmholtz Zentrum München, 85764 Neuherberg, Germany

[6]German Center for Diabetes Research, Neuherberg, Germany

[7]Primary Pharmacology Group, Discovery Sciences, Pfizer Inc., Groton, CT 06340, USA

[8]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA

[9]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA

[10]Department of Nutritional Sciences and Toxicology, University of California, Berkeley, Berkeley, CA 94720, USA

[11]Miller Institute for Basic Research in Science, University of California, Berkeley, Berkeley, CA 94720, USA

[12]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

[13]Division of Genetics and Genomics, Boston Children's Hospital, MA 02115, USA

[14]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA

[*]Correspondence: twalther@hsph.harvard.edu and robert@hsph.harvard.edu.

AUTHOR CONTRIBUTIONS

N.M., K.R.G., N.K., R.V.F., and T.C.W. started the project and wrote the first version of the manuscript. N.M., K.R.G., N.K., J.S., L.K., and J.A.O. generated the screen and proteomics data based on *Drosophila* R+ cells, Huh7 cells, U2OS cells, SUM159 cells, THP-1 macrophages and mouse liver. C.C. and S.B. performed the cholesterol esterification assays in lysates and cells, respectively. N.M., S.F-C., and M.R. determined the knockdown efficiencies of *C16orf53* and *MSRB3*. The portal was built by D-K.J., M.v.G., M.C.C., J.F., and N.P.B. All co-authors read, commented, and approved the final version of the manuscript.

[†]Lead contact

[§]These authors contributed equally

DECLARATION OF INTERESTS

The authors declare no competing interests.

[15]Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA

[16]Center on the Causes and Prevention of Cardiovascular Disease (CAP-CVD), Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

SUMMARY

Lipid droplets (LDs) are organelles of cellular lipid storage with fundamental roles in energy metabolism and cell membrane homeostasis. There has been an explosion of research into the biology of LDs, in part due to their relevance in diseases of lipid storage, such as atherosclerosis, obesity, type 2 diabetes, and hepatic steatosis. Consequently, there is an increasing need for a resource that combines datasets from systematic analyses of LD biology. Here we integrate high-confidence, systematically generated human, mouse and fly data from on studies of LDs in the framework of an online platform named the *Lipid Droplet Knowledge Portal* (https://lipiddroplet.org/). This scalable and interactive portal includes comprehensive datasets, across a variety of cell types, for LD biology, including transcriptional profiles of induced lipid storage, organellar proteomics, genome-wide screen phenotypes, and ties to human genetics. This resource is a powerful platform that can be utilized to identify determinants of lipid storage.

Keywords

inflammation; triacylglycerol; sterol ester; proteasome; C16orf54; MSRB3; proximity labeling; protein targeting

INTRODUCTION

Lipid droplets (LDs) are phospholipid monolayer–bound organelles found in most eukaryotes and some prokaryotes. These organelles store neutral lipids, such as triacylglycerols (TGs) and cholesterol esters (CE), that can be used to generate metabolic energy or cell membranes. Specific proteins, including many important lipid metabolism enzymes (*e.g.*, TG synthesis and degradation enzymes) bind to LD surfaces. Due to their important function in metabolism, alterations in LD biology are causal or implicated in diseases, such as lipodystrophy, atherosclerosis, obesity and related disorders (*e.g.*, type 2 diabetes mellitus (T2D), non-alcoholic fatty liver disease (NAFLD), and nonalcoholic steatohepatitis (NASH)). Moreover, alterations in LD metabolism are implicated in cancer, neurodegeneration, and immune function (Cruz et al., 2020; Gluchowski et al., 2017; Pereira-Dutra et al., 2019; Seebacher et al., 2020; Walther and Farese, 2012).

Despite the relevance of LDs to metabolic diseases, many aspects of their biology remain unclear, which has led to a recent surge of research into the biology of this organelle. In particular, systematic, unbiased approaches to studying LDs, including genome-wide screens to identify genes governing LD biology (Beller et al., 2008; Guo et al., 2008; Mejhert et al., 2020; Scott et al., 2015), and LD proteomics in different cells and tissues (Bersuker et al., 2018; Krahmer et al., 2013; Krahmer et al., 2018; Mejhert et al., 2020), have been instrumental to progress in our understanding of LDs. However, the results from these various large-scale experiments are currently fragmented, limiting the integration and interrogation of data from various experiments. Specifically, unlike for other organelles,

such as mitochondria (Rath et al., 2021), there is no comprehensive and scalable repository for integrating large data sets relevant to LD biology.

To address this deficiency and provide a resource for investigators of LD biology, we have created the *Lipid Droplet Knowledge Portal* (LD-Portal, https://lipiddroplet.org), an online resource that includes data from systematic research in LD biology. In this resource paper, we describe the initial version of the LD-Portal and, by highlighting several genes with phenotypes in these datasets that were previously not linked to LDs, we provide examples of how the LD-Portal can be used for discovering new facets of LD biology.

RESULTS

Overview of the Lipid Droplet Knowledge Portal

A conceptual content map and detailed overview of the available data in the initial version of the LD-Portal are shown in Figure 1 and Figure S1. The initial datasets integrated in the LD-Portal include a comprehensive dataset for RNA expression in human THP-1 macrophages (without and with lipid (Mejhert et al., 2020), LD proteomics for a variety of cell types (Bersuker et al., 2018; Krahmer et al., 2018; Mejhert et al., 2020) and high-content imaging screens of genes governing LD biology in human THP-1 macrophages and *Drosophila* S2 R+ cells (Mejhert et al., 2020; Song et al., 2021). In addition, the LD-Portal includes datasets for LD proteomics and phosphoproteomics of murine liver from mice fed chow or high-fat diets (Krahmer et al., 2018). To enable efficient data mining of LD biology relevant to human physiology and disease, we integrated the LD-Portal data with human genetics data from the Common Metabolic Diseases Knowledge Portal (https://HuGeAMP.org/). The LD-Portal resource allows both gene and phenotype-centric (“Gene Finder”) queries.

Transcriptional Response to Increased Lipid Storage in Macrophages

Cells store excess lipids, such as fatty acids or sterols, as neutral lipids in LDs, a process that we named the “lipid storage response” (LSR, (Mejhert et al., 2020)). One component of the LSR is a rewiring of transcription to facilitate LD formation and lipid storage and utilization. To enable discovery of LSR mechanisms and integration with other aspects of LD biology, the LD-Portal includes information on gene expression changes in differentiated human THP-1 macrophages, under conditions promoting lipid storage, in this case by incubation of cells with lipoproteins that induce the formation of LDs. Cells were cultured with acetylated low-density lipoproteins (ac-Lipo) (Mejhert et al., 2020) that contain both CEs and TGs (4.82 mg/dL total cholesterol and 4.12 mg/dL triglyceride). As also shown in Mejhert *et al.* (Mejhert et al., 2020), the uptake and degradation of ac-Lipo in the endo-lysosomal pathway resulted in LD storage of both CEs and TGs (Figure 2A). Formation of CE- and TG-containing LDs enabled us to probe pathways for the intracellular storage of either neutral lipid; in contrast, incubation with oleic acid resulted primarily in TG accumulation in LDs (Figure 2A).

Our RNA sequencing studies of the LSR in THP-1 cells showed pronounced changes in gene expression for 2414 genes (1289 up-regulated, and 1125 down-regulated, adj. p-value<0.01) after culturing cells in the presence of ac-Lipo (Mejhert et al., 2020). Re-

analysis of these data with gene set enrichment showed expected changes, such as the downregulation of cholesterol homeostasis genes and the induction of expression of the unfolded protein response and inflammatory genes (Figure 2B). However, genes in many other categories were also significantly altered during the LSR of this cell type. For many of these genes, the relationship between changes in their expression and ac-Lipo treatment is currently unknown. Further analyses of this dataset revealed that SREBP2 target genes were predominantly downregulated, explaining the changes in cholesterol homeostasis gene expression, and NF-kB target genes were upregulated, explaining increased expression of inflammation-related genes (Figure 2C–D). This dataset, accessible on the LD-portal interface, therefore provides a rich resource for probing the cellular response to ac-Lipo loading in macrophages.

Lipid Droplet Proteomes of Human Cells and Murine Liver

The LD-Portal also includes data on the subcellular localization of proteins and particularly highlights the propensity of proteins to localize to LDs under different conditions. To collect comprehensive information on the LD proteome of several model systems, we integrated data from the proteomic analyses of LDs from human THP-1 macrophages (Mejhert et al., 2020), human SUM159 triple-negative breast cancer cells (Mejhert et al., 2020)), human hepatoma Huh7 and human osteosarcoma U2OS cells (Bersuker et al., 2018), and a large-scale *in vivo* murine liver proteomic and phosphoproteomic organellar-localization atlas (Krahmer et al., 2018). These protein localization data are now collectively available for analysis on the LD-Portal.

For Huh7 and U2OS cells, LD proteins were identified using a proximity-labeling strategy (Bersuker et al., 2018). In this approach, the LD proteins ATGL or PLIN2 were tagged with the promiscuous biotinylating enzyme APEX2. Upon addition of APEX2 substrates, proteins in the immediate vicinity of either LD protein were modified with a biotin, isolated, and identified by proteomics (Bersuker et al., 2018). For each cell type, proteins that displayed a normalized confidence score >1 were identified as LD proteins, resulting in 77 and 152 LD proteins in Huh7 and U2OS cells, respectively (Bersuker et al., 2018).

For the SUM159 and THP-1 cultured cell lines, we measured the enrichment of proteins in the LD fraction compared with the total input fraction to identify proteins enriched on LDs. Based on the enrichment of *bona fide* LD proteins identified in Bersuker *et al.* (Bersuker and Olzmann, 2019), we calculated enrichment scores and used these as cut-offs for classifying protein localization. For THP-1 cells, a total of 5801 proteins were detected in the whole-cell lysate, with 1412 proteins in the LD fraction, of which 75 were enriched therein (enrichment score threshold, 3.07). For SUM159 cells, 5708 proteins were detected in the whole-cell lysate, with 629 proteins in the LD fraction, and 64 enriched in this fraction (enrichment score threshold, 2.02). To assess the robustness of the THP-1 and SUM159 proteomes, we compared the protein intensities in the LD fraction of THP-1 macrophages to that of SUM159 cells and found that these datasets were well correlated (R^2=0.76, p<0.0001), with 35 specific LD proteins in common, including MLX, VPS13C, RAB18, FAF2, RAB4A, and DHDDS (Figure 2E). Of the 35 LD proteins common to both cells, 28 were also reported in

U2OS or HuH7 cells (Bersuker et al., 2018). A list of the LD-enriched proteins found in the four cell types is provided in Table S1.

The LD-Portal also includes data on subcellular protein localization based on protein correlation profiling for the majority of proteins across organelles in C57BL/6J murine liver (Krahmer et al., 2018). These studies were performed in mice fed chow or high-fat diets (HFD), and by examining proteins across different cell fractions, they revealed how nutrient overload leads to organellar reorganization (Krahmer et al., 2018). Of the 6163 proteins quantified across cellular fractions, 5878 gave reproducible profiles for organelle assignment. Diet-dependent re-localization was found for 901 proteins, and protein expression changes for 258. The reproducibility of this dataset was assessed by calculating Pearson correlation of profiles, derived from the same biological conditions and between different diets, and this revealed Pearson's coefficients of 0.86 and 0.78 for protein levels and re-localization patterns, respectively (Figure S2A). For the murine liver samples, 787 protein profiles showed a characteristic peak in the top fraction after organelle separation by density centrifugation, indicating localization on the LDs or in LD-associated membranes. Most of these proteins localized to multiple organelles, and only 94 had a unique LD localization (Figure S2B). Of the 787 LD proteins, 308 showed a significant profile shift under HFD feeding. For instance, the proteins RAB7A, HSD17B11, DHRS1 and RAB1A undergo HFD-induced relocalization to LDs (Figure 2F).

We also utilized the LD-portal data to compare the LD proteins detected in cells (THP-1, SUM159, U2OS and Huh7) and murine liver. Based on these data, we identified 12 LD proteins that were common to all datasets (Table S1). These include well-known LD proteins (*e.g.*, AUP1, ACSL3, and RAB18) but also several for which data on LD localization was previously sparse (*e.g.*, NSDHL and LSS, two proteins regulating cholesterol biosynthesis). Notably, the function in this organelle is unclear for many of the proteins that were reproducibly and robustly identified in at least one of the LD fractions. Similarly, the LD-portal allows for comparison of LD proteins identified in murine liver and human Huh7 hepatoma cells, which yielded 30 proteins in common, including a number of proteins not well characterized as LD proteins (*e.g.*, the autophagy receptor SQSTM1 and the putative methyltransferases METTL7A and METTL7B). We anticipate that these proteomic datasets will open numerous lines of investigation.

To facilitate integrative analyses of genes and proteins that were identified for example as LD proteins, as part of the transcriptional LSR, or required for normal LDs in cells, the LD-portal includes a "Gene Finder" module. This tool allows for the easy analysis of queries for hits identified in multiple assays, and also enables definition of custom criteria, such as adjusting the cut-off values for significance in a particular assay.

LD Proteins of Murine Liver That Are Phosphorylated

The LD-Portal also includes comprehensive data on the localization of phosphorylated forms of proteins within murine liver (Krahmer et al., 2018). From 24,524 phosphosites detected, 11,712 gave reproducible profiles, and 1676 phosphorylation levels changed with HFD after normalization to protein levels. Analyzing specifically the LD proteins, 3037 had partial and 229 had unique LD localization, as assigned by support vector machine-based organelle

assignments. Among all proteins targeted from other compartments to LDs under HFD, almost half of the re-localizations (133) were accompanied with phosphorylation changes, indicating that this might be an important regulatory mechanism for targeting of many proteins to LDs.

Overlaying the protein and phosphosite profiles enabled identification of localization-specific phosphosites and phosphorylation events that are independent of protein localization. For example, the profile for AUP1, a protein identified as an LD protein (Klemm et al., 2011) shows dual LD and ER localization (Figure 2G). Yet, the same protein had phosphosites (S287/S290) that appeared only in the LD part of the profile, indicating that this site is phosphorylated only in the LD fraction and not the ER pool of the protein. Similarly, S33 of HSD17B13, S180/S187 of TPD52l2, and S435/S439 of LPIN1 had LD-specific phospho-signatures (Figure 2G). These examples illustrate how the portal can be used in future studies to identify phosphosites that might regulate protein localization to LDs.

Genome-Perturbation Screens for Lipid Droplet Phenotypes

The LD-Portal additionally features large datasets from high-content, imaging-based genome perturbation screens (Mejhert et al., 2020). In one of these screens, LDs were induced by ac-Lipo and stained with BODIPY, and LD information was collected using automated imaging and extraction of multiple image parameters at the single-cell level by image segmentation (Mejhert et al., 2020) (Figure 3A). With this pipeline, we disrupted expression of essentially all genes one-by-one in triplicate experiments and analyzed the effects on LDs in THP-1 macrophages. This dataset was utilized, for example, to discover that the MLX family of transcription factors (*e.g.*, MLX, MLXIP, MLXIPL/ChREBP) bind LDs and modulate their transcriptional activity (Mejhert et al., 2020).

Our analyses of the data from this screen yielded 21 non-redundant image parameters that describe LD size, number, dispersion, shape, and intensity in the screen images (Mejhert et al., 2020). Using tools from the LD-Portal, we performed additional analyses of these screen data. When we clustered genes with similar effects on LD parameters, we identified clusters containing genes with similar biological functions (Figure 3B, Table S2). For example, cluster 1 (c1) contains 17 genes involved with proteasome function (Table S2). The similarity of phenotype associated with different proteasomal (PSM) subunit genes is also apparent in a network plot that shows the genes with the most similar phenotype for each proteasome subunit in c1 (Figure 3C). By this analysis, each of the *PSM* genes, except *PSMB4*, are interconnected. Finding LD phenotypes for disruptions of *PSM* genes is consistent with our previous RNAi screen in *Drosophila* S2 R+ cells (Guo et al., 2008). In addition, 12 more genes are part of this proteasome network. Among these, several genes (*e.g.*, *DDI1*) have been directly implicated in the ubiquitin-proteasome system (Yip et al., 2020). In addition, *TMEM61*, which encodes an unknown ER protein, was reported in human genetic datasets as being highly associated with cholesterol metabolism and has limited homology to scavenger receptors (HHPRED, (Söding et al., 2005)).

For the four other clusters (c2, c3, c4, and c5) highlighted in Figure 3B, the biological underpinnings for similar LD phenotypes are unknown and not readily apparent.

Nonetheless, these clusters are likely to be informative for LD biology. For instance, independent replicates of *BSCL2*, encoding the protein seipin—a key factor in LD formation (Chung et al., 2019; Fei et al., 2008; Sui et al., 2018; Szymanski et al., 2007; Wang et al., 2016)—clustered tightly, validating the approach, and this phenotype (found in c5) identified several genes whose depletion phenotypes were highly similar (Figure 3D). The phenotype for *BSCL2*/seipin depletion was also similar to that of *LDAF1* knockdown (not classified as a "hit" by stringent criteria, but shown in Figure S3A), and the proteins encoded by these two genes function together in an LD formation complex (Chung et al., 2019). Another hit tightly correlated with *BSCL2*/seipin is the uncharacterized *C16orf54* (Figure 3D, upper panel). This open-reading frame is predicted to encode a protein with one transmembrane domain and is expressed highly in hematopoietic cells (Figure 3D, lower panels). Thus, the encoded protein may be functionally related to BSCL2/seipin, possibly with a functional role in blood cells. Representative RNAi screen images and image analysis results of cells depleted of *BSCL2* and *C16orf54* are displayed in Figure 3E. As seipin regulates LD formation induced by fatty acid supplementation, we tested if knockdown of *C16orf54* in THP-1 macrophages changed LD morphology when incubating the cells with oleic acid. Our results show that *C16orf54* depletion resulted in larger LDs with lower eccentricity compared to control cells (Figure S3B-C). Furthermore, our initial analyses of these screen data reveal groups of genes with similar LD depletion phenotypes and suggest that further mining of correlated genes may yield many mechanistic discoveries of machinery or pathways affecting LD biology. A tool to identify highly correlated genes for RNAi depletion studies in macrophages is included in the LD-portal.

LD processes and consequently LD phenotypes are determined in part by proteins on the LD surface. Therefore, datasets that address protein targeting to LDs are useful aspects to include in the LD-portal. One such example is a genome-wide screen for LD morphology and protein targeting phenotypes in *Drosophila* S2 R+ cells, using the metabolic enzyme GPAT4 as a model cargo for ER-to-LD targeting (Figure S1) (Song et al., 2021). For each human or *Drosophila* gene page, the LD-portal contains a link to the homologous protein(s), facilitating comparisons of phenotypes in both systems. As an example of such an analysis, *RAB1A* (homologous to Rab1 in *Drosophila*) was identified as an LD associated protein (present in all proteomic datasets included in the portal) required for GPAT4 targeting to LDs in *Drosophila* S2 R+ cells (Song et al., 2021).

Importantly, the LD-portal is scalable and as more genetic perturbation screens become available, they will be incorporated for analysis (see https://lipiddroplet.org/about for the data submission process).

Data Mining of the LD-Portal Identifies MSRB3 as a Determinant of Cholesterol Ester Storage

As an example of how the LD-Portal can be mined for new insights to LD biology, we performed a secondary screen in which we compared LD phenotypes in response to gene knockdowns when LD formation was driven by cholesterol (via culture with ac-LDL) and fatty acids (via culture with oleic acid). As displayed in Figure 4A, we selected 19 genes with different depletion effects on LD size, and our focused re-screen revealed that some

genes, such as *BCSL2*, exhibited robust phenotypes for either culture condition. Other genes, such as *C9orf16*, were more important for storage of one of the excess lipids, in this case TG that was induced by oleic acid.

This secondary screen identified *MSRB3* as a gene with a striking depletion phenotype characterized by small and dispersed LDs (Figure 4B). In the LD-Portal (see below), *MSRB3* was associated with the diagnosis of type 2 diabetes (T2D) (p=3.8e-5) and adiponectin levels (p=7.7e-5), among other traits. *MSRB3* encodes an ER methionine sulfoxide reductase of unclear function. Mutations in human *MSRB3* lead to deafness, and it has been associated with progression of renal clear cell carcinoma, gastric cancer, and Alzheimer disease (Ahmed et al., 2011; Conner et al., 2019; Kwon et al., 2014; Ma et al., 2019; Ye et al., 2020) Most often, sulfoxide reductases are thought to help to maintain protein folding, structure, and activity. To determine the biochemical basis of the LD phenotype, we investigated synthesis of cholesterol esters in cells and lysates depleted for *MSRB3* (Figure S3D). Cholesterol ester synthesis was significantly increased when *MSRB3* was absent (Figure 4C). Increased levels of the main cholesterol ester synthesis enzyme in THP-1 macrophages, ACAT1 (encoded by *SOAT1*), were also found (Figure 4D). However, *SOAT1* mRNA levels were not affected by *MSR3B3* silencing (Figure S3E). These findings suggest that the methionine sulfoxide reducatase MSRB3 is required to control ACAT1 protein turnover or activity, a hypothesis that can now be investigated in mechanistic detail.

Integration of Lipid Droplet Biology Datasets with Human Genetics

The LD-Portal also contains human genetic gene and gene-set association analyses for many complex traits, calculated using the *Multi-Marker Analysis of GenoMic Annotation* (MAGMA) algorithm (de Leeuw et al., 2015). To explore these connections, we determined if our datasets of LD proteins or gene hits associated with an LD-phenotype had preferential association scores in the MAGMA dataset. We defined a subset of MAGMA traits as metabolically associated, including BMI, coronary artery disease, child obesity, inflammatory bowel disease, T2D, waist circumference, visceral adipose tissue volume, and total cholesterol, ALT, adiponectin, oleic acid, palmitic acid, palmitoleic acid, fasting glucose, fasting insulin, HDL-cholesterol, LDL-cholesterol, and TGs.

We examined MAGMA metabolic association scores for RNAi THP-1 screen hits, genes whose RNA expression levels were modulated by ac-Lipo, and LD-localized proteins. We found a number of expected associations in each set of genes (Figure 5A). For example, we detected strong associations of *APOB* and *APOE* with cholesterol and LDL phenotypes; scavenger receptor class B type 1 (*SCARB1*) with HDL cholesterol; insulin growth factor 1 (*IGF1*) with fasting insulin, and *FTO* with BMI (Dina et al., 2007; Frayling et al., 2007; Scuteri et al., 2007). Within each of these datasets, we detected significantly more associations with metabolic traits than expected for a random sample of genes (Figure 5B) (p-values=0.025, 6.72e-5, and 3.09e-5, respectively).

Analyzing the hits of the genome-perturbation screen in THP-1 cells, we detected many highly significant associations with human metabolic traits. For instance, we identified *ABHD16A* associated with five different metabolic traits, including BMI, cholesterol levels, T2D, TGs, and ulcerative colitis. We plotted the MAGMA score percentile

for each gene, highlighting *ABHD16A* for nine total traits (Figure S4). Eight traits had significant associations, indicating a p-value less than 2.5 E-6 (6.4 on a $\log_{10}$ scale). The function of ABHD16A is not well-understood, but molecularly it encodes a phosphatidylserine hydrolase of the ER (Kamat et al., 2015). Our data suggest that modulating phosphatidylserine levels is important to maintain normal LDs, and interference with normal phosphatidylserine levels can lead to metabolic complications.

DISCUSSION

The LD-Portal provides a rich open-source platform for mining biological databases related to LD biology. The current version of the LD-Portal provides several searchable databases that can be mined to query genes or phenotypes and discover connections for further mechanistic exploration. Additionally, integration of LD-Portal data with other platforms, such as human genetic MAGMA data from the *Common Metabolic Diseases Knowledge Portal*, allows filtering of queries to discern connections with human disease. In this description of the LD-Portal resource, we highlighted several examples, based on our initial analysis of the data sets, illustrating how mining of the LD-Portal resources will undoubtedly advance discoveries in LD biology.

Limitations of the study

The current version of the LD-Portal contains only selected datasets. However, the portal is scalable, allowing for integration of data from various sources, including genome perturbation screens, proteomic studies, and gene expression analyses. In the future, lipidomics, metabolomics, or other types of datasets from important tissues, such as adipose tissue can be integrated. We expect that over time, the LD-portal will contain many more systematically generated datasets from the scientific community. A simple process for data submission is outlined on the LD-portal page (https://lipiddroplet.org/about). The LD-Portal will enable insights into the basic biology of important genetic risk factors for diseases associated with prevalent public health problems, such as obesity, hepatic steatosis and NAFLD/NASH, and cardiovascular disease.

STAR METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contacts, Robert V. Farese Jr. (rfarese@hsph.harvard.edu) and Tobias C. Walther (twalther@hsph.harvard.edu).

Materials availability—This study did not generate new unique reagents.

Data and code availability—Transcriptomic and proteomic data have been deposited at GEO/PRIDE, respectively, and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Microscopy data reported in this paper will be shared by the lead contact upon request.

This paper does not report original code.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell studies—THP-1 monocyte/macrophage and SUM159 cell-culture conditions are described in Mejhert *et al.* (Mejhert et al., 2020). For induction of lipid storage, cells were incubated in the presence of ac-Lipo (A6961, PanReac Applichem), ac-LDL (BT-906, Alfa Aesar), or OA (O1383, Sigma-Aldrich). Human ac-Lipo was acetylated as described (Basu et al., 1976) and OA was complexed with essentially fatty acid–free BSA (A6003, Sigma-Aldrich) at a fatty acid/albumin molar ratio of 3:1.

Animal studies—Mice were handled as described in Krahmer *et al.* (Krahmer et al., 2018). In brief, 4-week-old male C57BL/6J mice were fed either a low-fat (D12331, Research Diets) or high-fat diet (D12329, Research Diets) for 12 weeks. In accordance with an approved protocol (Animal Protection Institute of Upper Bavaria 55.2-1-54-2532-164-2015), mice were sacrificed in an *ad-libitum*-fed state, and the livers dissected for proteomic analyses. Ethical approval was received for all animal work.

METHOD DETAILS

RNA isolation and sequencing

Total RNA isolation and sequencing procedures are described in Mejhert *et al.* (Mejhert et al., 2020). In brief, total RNA was isolated from THP-1 macrophages using the QIAshredder and RNeasy Mini kits (79656 and 74106, QIAGEN). Samples were submitted to the Genomics Core at Tufts University for RNA sequencing. After quality controls and library preparation, samples were sequenced on a HiSeq 2500 using V4 chemistry (Illumina). Data analyses are described under “Processing of RNA sequencing data” below.

Lipid extraction and thin layer chromatography

Details on lipid extraction and thin layer chromatography are described in Mejhert *et al.* (Mejhert et al., 2020). In brief, lipids were extracted from THP-1 macrophages incubated with ac-Lipo or OA using Folch’s extraction (Folch et al., 1957). Lipids were subsequently separated by thin layer chromatography using a neutral lipid solvent (heptane/isopropyl ether/acetic acid, 60:40:4, v/v/v) as described (Lehner and Vance, 1999) and detected by cerium molybdate staining. Quantifications were performed in Fiji (Schindelin et al., 2012). For cholesterol ester quantifications, lipids were collected from lower organic phase and separated by TLC using a hexane:diethyl ether:acetic acid (80:20:1) solvent system. TLC plates were exposed to a phosphor-imaging cassette overnight and revealed by Typhoon FLA 7000 phosphor imager. Band intensities were quantified using Fiji.

Organellar proteomics

Mouse liver protein correlation profiles were generated as described in Krahmer *et al.* (Krahmer et al., 2018). THP-1 and SUM159 LD proteomes were generated and described as in Mejhert *et al.* (Mejhert et al., 2020).

Genome-wide and secondary RNAi screens

The THP-1 macrophage RNAi screen was completed and described in Mejhert *et al.* (Mejhert et al., 2020). In brief, the screen was run with samples in triplicate using an siRNA library comprising 18,119 target genes with 4 oligos per target gene. THP-1 cells were plated and differentiated for 1 day in RPMI medium containing phorbol 12-myristate 13-acetate, then transfected using Lipofectamine RNAiMAX. Subsequently, cells were grown in serum-free RPMI medium for 3 days, followed by incubations with 25 μg/mL of ac-Lipo for 2 days except for controls not containing lipids. To stain LDs and nuclei, cells were fixed with 4% paraformaldehyde and then incubated with Hoechst and BODIPY stains. 7 images per well were acquired for each channel using an Opera High Content microscope (PerkinElmer). To extend the genome-wide RNAi screen, a secondary screen was performed. Genes were randomly selected and re-screened using pools of 4 siRNAs. Cells were incubated with OA or ac-LDL to induce storage of TGs or CEs, respectively. Results were generated as described above, and results were compared among the genome-wide RNAi screen and the validation studies by performing pair-wise correlations using the set of selected image features. Data analyses are described under “Image analyses” below.

Structure predictions

To predict transmembrane helices potentially present in C16orf54, the TMHMM Server v. 2.0 was used with default settings (Krogh et al., 2001). The FASTA sequence of C16orf54 was obtained from UniProt database (2021).

Cholesterol esterification assays

Cholesterol ester formation was measured in cells and lysates. For both assays, samples from RISC-free control or siMSRB3 transfected macrophages were included, and pharmacological inhibition of cholesterol ester formation was used as an assay control (S9318, Sigma-Aldrich). The assay performed in cells was originally described in (Goldstein et al., 1983). In brief, cells were starved from serum for 3 days. 16 hours prior to performing the assay, compactin and mevalonate were added to the cell-culture media to inhibit endogenous cholesterol biosynthesis. Cells were subsequently incubated with the indicated amount of ac-LDL for 7 hours. To determine cholesterol esterification levels, [^{14}C]oleate was added to the media for the last 2 hours. After thorough washes with cold PBS, lipids were extracted and quantified as described under “Lipid extraction and thin layer chromatography”. The *in vitro* acyl CoA:cholesterol acyltransferase (ACAT) activity in cell lysates was measured as described (Meiner et al., 1996) with some modifications. In brief, THP-1 cells were lysed in lysis buffer (50 mM Tris Cl, pH 7.4, 250 mM sucrose, with protease inhibitors (11873580001, Roche)). Cell pellets were resuspended in ice-cold lysis buffer and sonicated using ultrasonic homogenizer (Biologics, Inc., model 3000MP) for 10 sec with 30% amplitude. Cell homogenate was centrifuged at 3000 × g at 4°C for 5 min and supernatant was used as enzyme source. Total ACAT was measured at Vmax substrate concentrations. Assay mixture contained 20 μg of proteins, 200 μM of cholesterol (dissolved in ethanol), 25 μM of oleoyl-CoA, which contained [^{14}C] oleoyl-CoA as tracer, and 1 mM $MgCl_2$ in an assay in buffer containing 100 mM Tris-HCl (pH 7.4) and protease inhibitors. Total reaction volume was 200 μl, and the reaction was performed in 2 ml tubes.

The reaction was carried out for 30 min at 37°C in a water bath with shaking. Reaction was stopped by adding 1 ml chloroform and methanol (2:1) and acidified water (2% orthophosphoric acid). After stopping the reaction, tubes were vortexed well and centrifuged 10,000 × g at room temperature for 10 min. Quantifications were performed as described under “Lipid extraction and thin layer chromatography”.

SDS page and western blot

Details on SDS page and western blotting, are described in Mejhert *et al.* (Mejhert et al., 2020). Primary antibodies targeting calnexin (sc-46669, Santa Cruz Biotechn.), MSRB3 (ab180584, Abcam), C16orf54 (HPA060546, Sigma-Aldrich), GAPDH (2118S, Cell Signaling Techn.) and ACAT1 (kindly provided by Drs. Ta-Yuan and Catherine Chung-Yao Chang, Department of Biochemistry, Dartmouth Medical School) were used.

cDNA synthesis and qPCR

Total RNA was isolated as described under “RNA isolation and sequencing”, cDNA was synthesized using iScript cDNA Synthesis Kit (1708891, Bio-Rad) and real-time qPCR was performed with Power SYBR Green PCR Master Mix (4367659, Applied Biosystems). Forward and reverse primers were as follows: *MSRB3* (Fw 5′-AAC TGA GGA AGC GGC TAA CA-3′, Rv 5′-ACA AGG CAG CCG AAT TTA TG-3′), *C16orf54* (Fw 5′-CTT ACT TAT AAT GCT CCA CCC TAC-3′, Rv 5′-AGG GAA ATG GAA ACT ACA TCT G-3′), *SOAT1* (Fw 5′-CTC TCT CTT AGA TGA ACT GCT TG-3′, Rv 5′-CTA CAA GTG TGC TGA GGA TAA A-3′) and *GAPDH* (Fw 5′-ACA GTT GCC ATG TAG ACC-3′, Rv 5′-TTT TTG GTT GAG CAC AGG-3′). Results were normalized to the reference gene *GAPDH* and evaluated using the delta-delta Ct method.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistics

Visualizations and statistical analyses of results were performed using appropriate packages in RStudio (version 1.0.143) as described under each subheading.

Processing of RNA sequencing data

Raw sequencing data were analyzed as described in Mejhert *et al.* (Mejhert et al., 2020). Briefly, transcript abundance was quantified using Salmon (Patro et al., 2017), results were imported into RStudio using tximport (Soneson et al., 2015), and differentially expressed genes were identified using DESeq2 (Love et al., 2014). Gene set–enrichment analysis was performed to identify gene sets regulated by lipid storage. For this, the Hallmark track was used from the Molecular Signatures Database. As a proxy for SREBP2 and NFκB activity, pathways were created based on target genes identified in published studies (Horton et al., 2003; Lim et al., 2007)

Processing of proteomic data

Details on the processing of proteomic data are described in Mejhert *et al.* (Mejhert et al., 2020) and Krahmer *et al.* (Krahmer et al., 2018). In brief, correlation profiling

was applied to map the cellular localization of proteins and phosphopeptides in mouse liver. Cellular localizations are assigned by support-vector machine-based learning on the generated profiles. For THP-1 and SUM159 data, fold changes comparing LD fractions with total cell lysates were based on label-free quantification. To calculate 99% confidence intervals for canonical LD proteins, the top 50 high-confidence proteins targeting to LDs were extracted (Bersuker et al., 2018) and overlapped with the results presented in this study. From this analysis, the lower boundary of the confidence interval was used as a cut-off for our classification.

Image analyses

Details on the image analyses were described in Mejhert *et al.* (Mejhert et al., 2020). In brief, CellProfiler was used to extract features from the images. For each extracted image feature, the median rz-score was calculated per gene. Image feature replicates were compared pairwise across the screen, and non-reproducible parameters were excluded. After this, a correlation matrix was generated by correlating all included image features with each other, and the dimensionality of the matrix was tested using hierarchical clustering. Features were excluded if they covaried, and genes were classified as hits if they were distributed top/bottom 15 for one image parameter and/or top/bottom 50 for more than one of the remaining high-confidence image parameters. The RNAi screen hits were pairwise correlated, based on the filtered image features, and the resulting matrix containing Spearman's rho values was clustered using the pheatmap package (with default clustering methods and cutree_rows/cols set to five). All steps downstream of the CellProfiler analysis were performed in RStudio. The top-three and top-10 neighbors for proteasome subunits and *BSCL2* were extracted and highlighted using Cytoscape or the pheatmap package, respectively. For images from THP-1 macrophages transfected with control or *C16orf54*-targeting siRNAs, the filtered features from the RNAi screen were extracted, z-scored and clustered using the pheatmap package.

Expression analyses of FANTOM CAT browser data

C16orf54 expression levels across 571 human cells and tissues were extracted from the FANTOM CAT browser (Hon et al., 2017). Samples were ranked from high to low abundance of *C16orf54* expression levels.

Human genetic association analysis

Genetic association results in the LD-Portal are derived from the Common Metabolic Diseases Knowledge Portal (CMDKP; cmdkp.org), a public resource that aggregates genetic association results for >300 metabolic diseases and traits. In the CMDKP, genetic association results are meta-analyzed using the METAL algorithm, accounting for sample overlap between datasets, to generate "bottom-line" single-variant associations for each disease and trait. These are then analyzed with the MAGMA method (using default parameters) to generate gene-level association scores for each trait and each gene (de Leeuw et al., 2015). MAGMA gene-level association, scores are calculated based on the average association Z-scores for SNPs within a fixed window of the gene, after correcting for correlations among single-nucleotide polymorphism.

ADDITIONAL RESOURCES

Lipid droplet knowledge portal: https://lipiddroplet.org/

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Drs. Ta-Yuan and Catherine Chung-Yao Chang (Department of Biochemistry, Dartmouth Medical School) for kindly providing a custom-made ACAT1 antibody. This work was supported in part by the Howard Hughes Medical Institute (to T.C.W.) and NIH grants (R01DK124913 to R.V.F. and T.C.W., 5TL1TR001101/T32GM007753 to J.S., and UM1DK105554 to J.F. and N.P.B.), the Margareta af Uggla's foundation (M.R.), Knut & Alice Wallenberg's foundation (M.R.), the Swedish Research Council (M.R., N.M.), ERC-SyG SPHERES (856404 to M.R.), the NovoNordisk Foundation (including the Tripartite Immuno-metabolism Consortium Grant Number NNF15CC0018486, the MSAM consortium NNF15SA0018346 and the MeRIAD consortium Grant number 0064142, all three M.R. and NNF20OC0061149 to N.M.), CIMED (N.M. and M.R.), the Swedish Diabetes Foundation (M.R.), the Stockholm County Council (M.R.), the Strategic Research Program in Diabetes at Karolinska Institutet (M.R.), and the American Heart Association Predoctoral Fellowship and the Aramont Fund for Emerging Science Research (to J.S.).

REFERENCES

(2021). UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res 49, D480–d489. [PubMed: 33237286]

Ahmed ZM, Yousaf R, Lee BC, Khan SN, Lee S, Lee K, Husnain T, Rehman AU, Bonneux S, Ansar M, et al. (2011). Functional null mutations of MSRB3 encoding methionine sulfoxide reductase are associated with human deafness DFNB74. Am J Hum Genet 88, 19–29. [PubMed: 21185009]

Basu SK, Goldstein JL, Anderson GW, and Brown MS (1976). Degradation of cationized low density lipoprotein and regulation of cholesterol metabolism in homozygous familial hypercholesterolemia fibroblasts. Proc Natl Acad Sci U S A 73, 3178–3182. [PubMed: 184464]

Beller M, Sztalryd C, Southall N, Bell M, Jäckle H, Auld DS, and Oliver B (2008). COPI complex is a regulator of lipid homeostasis. PLoS Biol 6, e292. [PubMed: 19067489]

Bersuker K, and Olzmann JA (2019). Identification of Lipid Droplet Proteomes by Proximity Labeling Proteomics Using APEX2. Methods Mol Biol 2008, 57–72. [PubMed: 31124088]

Bersuker K, Peterson CWH, To M, Sahl SJ, Savikhin V, Grossman EA, Nomura DK, and Olzmann JA (2018). A Proximity Labeling Strategy Provides Insights into the Composition and Dynamics of Lipid Droplet Proteomes. Dev Cell 44, 97–112.e117. [PubMed: 29275994]

Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 7, R100. [PubMed: 17076895]

Chang CC, Chen J, Thomas MA, Cheng D, Del Priore VA, Newton RS, Pape ME, and Chang TY (1995). Regulation and immunolocalization of acyl-coenzyme A: cholesterol acyltransferase in mammalian cells as studied with specific antibodies. J Biol Chem 270, 29532–29540. [PubMed: 7493995]

Chung J, Wu X, Lambert TJ, Lai ZW, Walther TC, and Farese RV Jr. (2019). LDAF1 and Seipin Form a Lipid Droplet Assembly Complex. Dev Cell 51, 551–563.e557. [PubMed: 31708432]

Conner SC, Benayoun L, Himali JJ, Adams SL, Yang Q, DeCarli C, Blusztajn JK, Beiser A, Seshadri S, and Delalle I (2019). Methionine Sulfoxide Reductase-B3 Risk Allele Implicated in Alzheimer's Disease Associates with Increased Odds for Brain Infarcts. J Alzheimers Dis 68, 357–365. [PubMed: 30775993]

Cruz ALS, Barreto EA, Fazolini NPB, Viola JPB, and Bozza PT (2020). Lipid droplets: platforms with multiple functions in cancer hallmarks. Cell Death Dis 11, 105. [PubMed: 32029741]

de Leeuw CA, Mooij JM, Heskes T, and Posthuma D (2015). MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 11, e1004219. [PubMed: 25885710]

Dina C, Meyre D, Gallina S, Durand E, Körner A, Jacobson P, Carlsson LM, Kiess W, Vatin V, Lecoeur C, et al. (2007). Variation in FTO contributes to childhood obesity and severe adult obesity. Nat Genet 39, 724–726. [PubMed: 17496892]

Fei W, Shui G, Gaeta B, Du X, Kuerschner L, Li P, Brown AJ, Wenk MR, Parton RG, and Yang H (2008). Fld1p, a functional homologue of human seipin, regulates the size of lipid droplets in yeast. J Cell Biol 180, 473–482. [PubMed: 18250201]

Folch J, Lees M, and Sloane Stanley GH (1957). A simple method for the isolation and purification of total lipides from animal tissues. J Biol Chem 226, 497–509. [PubMed: 13428781]

Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JR, Elliott KS, Lango H, Rayner NW, et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316, 889–894. [PubMed: 17434869]

Gluchowski NL, Becuwe M, Walther TC, and Farese RV Jr. (2017). Lipid droplets and liver disease: from basic biology to clinical implications. Nat Rev Gastroenterol Hepatol 14, 343–355. [PubMed: 28428634]

Goldstein JL, Basu SK, and Brown MS (1983). Receptor-mediated endocytosis of low-density lipoprotein in cultured cells. Methods Enzymol 98, 241–260. [PubMed: 6321901]

Guo Y, Walther TC, Rao M, Stuurman N, Goshima G, Terayama K, Wong JS, Vale RD, Walter P, and Farese RV (2008). Functional genomic screen reveals genes involved in lipid-droplet formation and utilization. Nature 453, 657–661. [PubMed: 18408709]

Hon CC, Ramilowski JA, Harshbarger J, Bertin N, Rackham OJ, Gough J, Denisenko E, Schmeier S, Poulsen TM, Severin J, et al. (2017). An atlas of human long non-coding RNAs with accurate 5' ends. Nature 543, 199–204. [PubMed: 28241135]

Horton JD, Shah NA, Warrington JA, Anderson NN, Park SW, Brown MS, and Goldstein JL (2003). Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. Proc Natl Acad Sci U S A 100, 12027–12032. [PubMed: 14512514]

Kamat SS, Camara K, Parsons WH, Chen DH, Dix MM, Bird TD, Howell AR, and Cravatt BF (2015). Immunomodulatory lysophosphatidylserines are regulated by ABHD16A and ABHD12 interplay. Nat Chem Biol 11, 164–171. [PubMed: 25580854]

Klemm EJ, Spooner E, and Ploegh HL (2011). Dual role of ancient ubiquitous protein 1 (AUP1) in lipid droplet accumulation and endoplasmic reticulum (ER) protein quality control. J Biol Chem 286, 37602–37614. [PubMed: 21857022]

Krahmer N, Hilger M, Kory N, Wilfling F, Stoehr G, Mann M, Farese RV Jr., and Walther TC (2013). Protein correlation profiles identify lipid droplet proteins with high confidence. Mol Cell Proteomics 12, 1115–1126. [PubMed: 23319140]

Krahmer N, Najafi B, Schueder F, Quagliarini F, Steger M, Seitz S, Kasper R, Salinas F, Cox J, Uhlenhaut NH, et al. (2018). Organellar Proteomics and Phospho-Proteomics Reveal Subcellular Reorganization in Diet-Induced Hepatic Steatosis. Dev Cell 47, 205–221.e207. [PubMed: 30352176]

Krogh A, Larsson B, von Heijne G, and Sonnhammer EL (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305, 567–580. [PubMed: 11152613]

Kwon TJ, Cho HJ, Kim UK, Lee E, Oh SK, Bok J, Bae YC, Yi JK, Lee JW, Ryoo ZY, et al. (2014). Methionine sulfoxide reductase B3 deficiency causes hearing loss due to stereocilia degeneration and apoptotic cell death in cochlear hair cells. Hum Mol Genet 23, 1591–1601. [PubMed: 24191262]

Lehner R, and Vance DE (1999). Cloning and expression of a cDNA encoding a hepatic microsomal lipase that mobilizes stored triacylglycerol. Biochem J 343 *Pt* 1, 1–10. [PubMed: 10493905]

Lim CA, Yao F, Wong JJ, George J, Xu H, Chiu KP, Sung WK, Lipovich L, Vega VB, Chen J, et al. (2007). Genome-wide mapping of RELA(p65) binding identifies E2F1 as a transcriptional

activator recruited by NF-kappaB upon TLR4 activation. Mol Cell 27, 622–635. [PubMed: 17707233]

Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550. [PubMed: 25516281]

Ma X, Wang J, Zhao M, Huang H, and Wu J (2019). Increased expression of methionine sulfoxide reductases B3 is associated with poor prognosis in gastric cancer. Oncol Lett 18, 465–471. [PubMed: 31289518]

Meiner VL, Cases S, Myers HM, Sande ER, Bellosta S, Schambelan M, Pitas RE, McGuire J, Herz J, and Farese RV Jr. (1996). Disruption of the acyl-CoA:cholesterol acyltransferase gene in mice: evidence suggesting multiple cholesterol esterification enzymes in mammals. Proc Natl Acad Sci U S A 93, 14041–14046. [PubMed: 8943057]

Mejhert N, Kuruvilla L, Gabriel KR, Elliott SD, Guie MA, Wang H, Lai ZW, Lane EA, Christiano R, Danial NN, et al. (2020). Partitioning of MLX-Family Transcription Factors to Lipid Droplets Regulates Metabolic Gene Expression. Mol Cell 77, 1251–1264.e1259. [PubMed: 32023484]

Patro R, Duggal G, Love MI, Irizarry RA, and Kingsford C (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14, 417–419. [PubMed: 28263959]

Pereira-Dutra FS, Teixeira L, de Souza Costa MF, and Bozza PT (2019). Fat, fight, and beyond: The multiple roles of lipid droplets in infections and inflammation. J Leukoc Biol 106, 563–580. [PubMed: 31121077]

Rath S, Sharma R, Gupta R, Ast T, Chan C, Durham TJ, Goodman RP, Grabarek Z, Haas ME, Hung WHW, et al. (2021). MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. Nucleic Acids Res 49, D1541–d1547. [PubMed: 33174596]

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676–682. [PubMed: 22743772]

Scott CC, Vossio S, Vacca F, Snijder B, Larios J, Schaad O, Guex N, Kuznetsov D, Martin O, Chambon M, et al. (2015). Wnt directs the endosomal flux of LDL-derived cholesterol and lipid droplet homeostasis. EMBO Rep 16, 741–752. [PubMed: 25851648]

Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, Najjar S, Nagaraja R, Orrú M, Usala G, et al. (2007). Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genet 3, e115. [PubMed: 17658951]

Seebacher F, Zeigerer A, Kory N, and Krahmer N (2020). Hepatic lipid droplet homeostasis and fatty liver disease. Semin Cell Dev Biol 108, 72–81. [PubMed: 32444289]

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, and Ideker T (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498–2504. [PubMed: 14597658]

Söding J, Biegert A, and Lupas AN (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33, W244–248. [PubMed: 15980461]

Soneson C, Love MI, and Robinson MD (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res 4, 1521. [PubMed: 26925227]

Song J, Mizrak A, Lee C-W, Cicconet M, Lai ZW, Lu C-H, Mohr SE, Farese RV, and Walther TC (2021). Identification of two pathways mediating protein targeting from ER to lipid droplets. bioRxiv, 2021.2009.2014.460330.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102, 15545–15550. [PubMed: 16199517]

Sui X, Arlt H, Brock KP, Lai ZW, DiMaio F, Marks DS, Liao M, Farese RV Jr., and Walther TC (2018). Cryo-electron microscopy structure of the lipid droplet-formation protein seipin. J Cell Biol 217, 4080–4091. [PubMed: 30327422]

Szymanski KM, Binns D, Bartz R, Grishin NV, Li WP, Agarwal AK, Garg A, Anderson RG, and Goodman JM (2007). The lipodystrophy protein seipin is found at endoplasmic reticulum lipid droplet junctions and is important for droplet morphology. Proc Natl Acad Sci U S A 104, 20890–20895. [PubMed: 18093937]

Walther TC, and Farese RV Jr. (2012). Lipid droplets and cellular lipid metabolism. Annu Rev Biochem 81, 687–714. [PubMed: 22524315]

Wang H, Becuwe M, Housden BE, Chitraju C, Porras AJ, Graham MM, Liu XN, Thiam AR, Savage DB, Agarwal AK, et al. (2016). Seipin is required for converting nascent to mature lipid droplets. Elife 5.

Ye X, Liang T, Deng C, Li Z, and Yan D (2020). MSRB3 promotes the progression of clear cell renal cell carcinoma via regulating endoplasmic reticulum stress. Pathol Res Pract 216, 152780. [PubMed: 31889586]

Yip MCJ, Bodnar NO, and Rapoport TA (2020). Ddi1 is a ubiquitin-dependent protease. Proc Natl Acad Sci U S A 117, 7776–7781. [PubMed: 32193351]

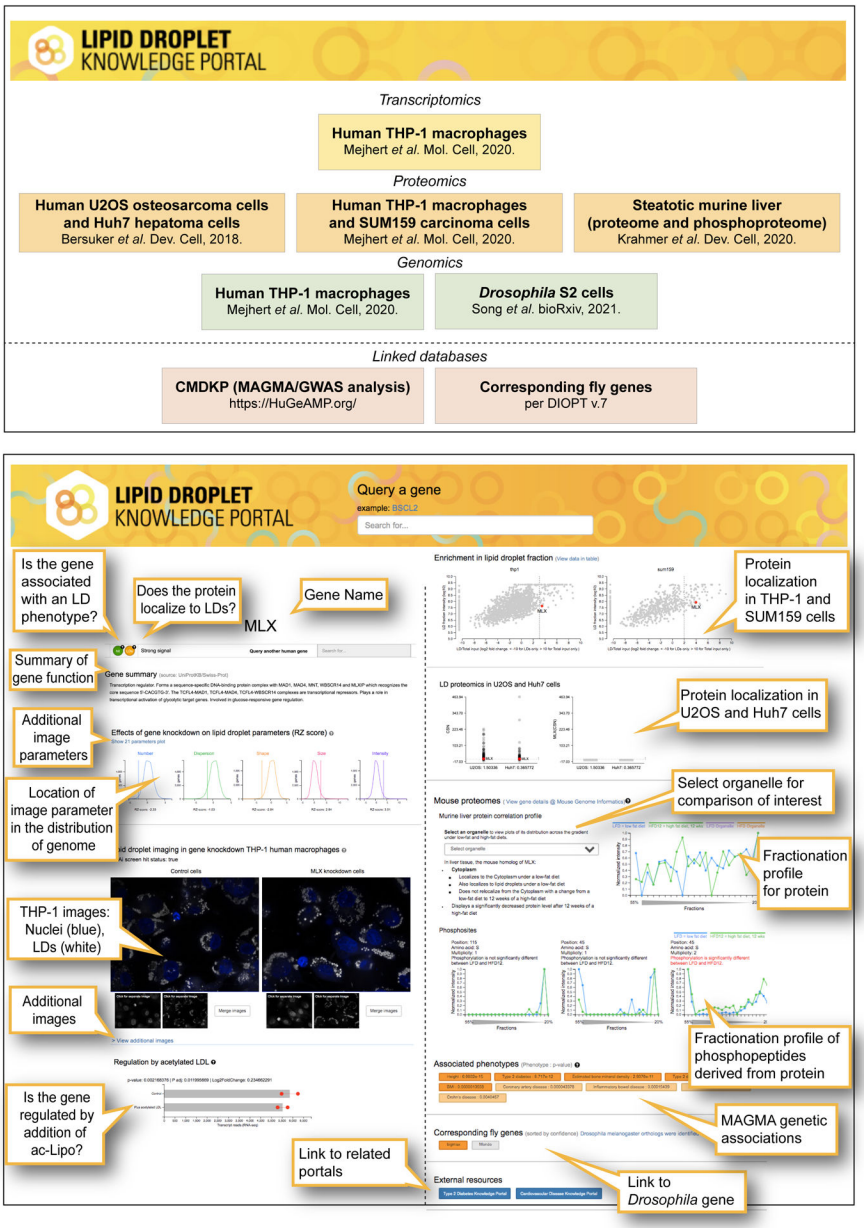

Figure 1: Overview of the *Lipid Droplet Knowledge Portal*

Content of the LD-Portal. Original publications that contributed data to the initial version of the LD-Portal are listed (upper panel) and a graphical summary of the LD-Portal interface with key to the data modules is displayed (lower panel).

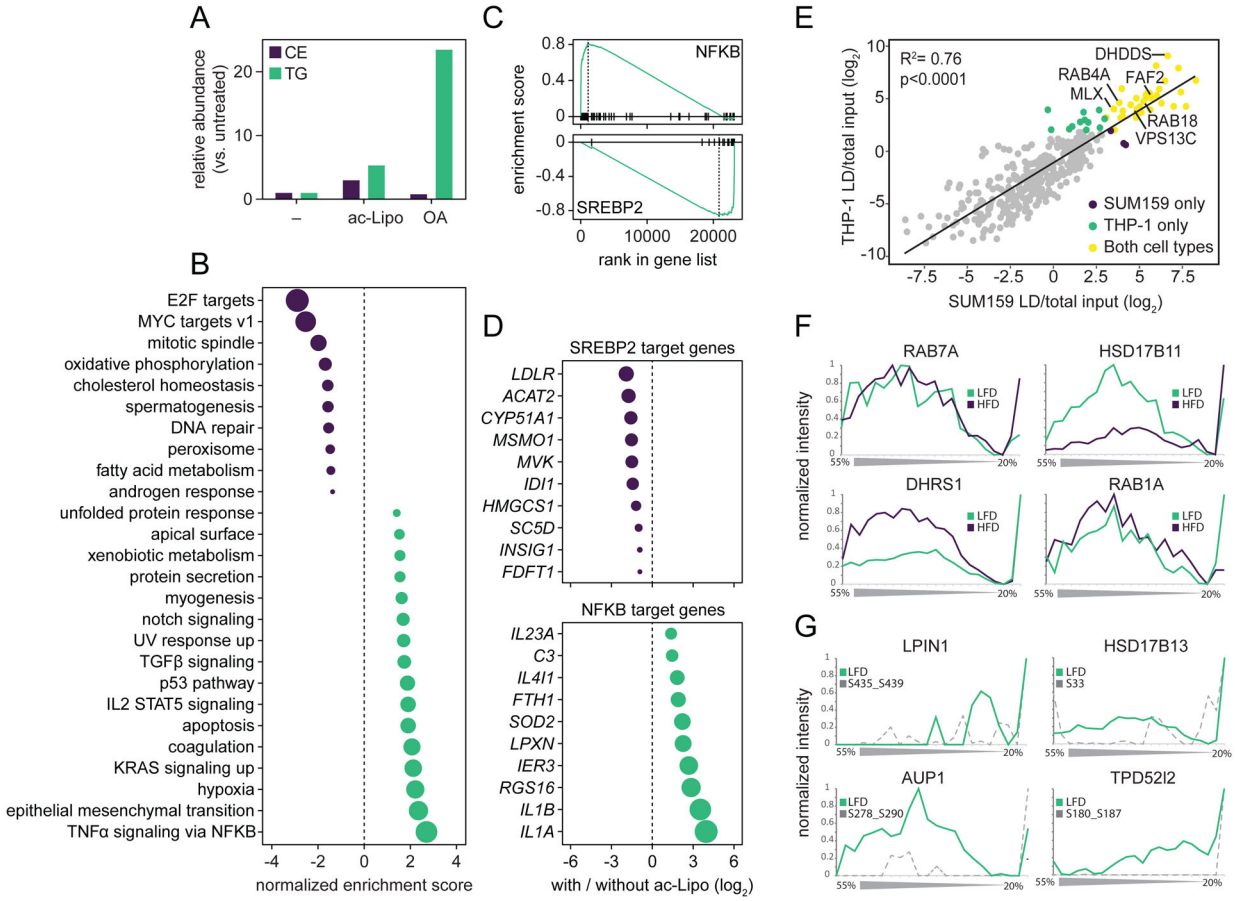

Figure 2: Lipid storage induction regulates metabolic and inflammatory pathways and protein localization to LDs

(A) Incubation with ac-Lipo induces TG and CE storage in THP-1 macrophages. Cellular lipid storage was induced by incubating differentiated THP-1 macrophages in the absence/presence of ac-Lipo (100 μg/mL) or OA (0.5 mM) for 1 day, followed by determination of lipid composition by thin layer chromatography. Results from one representative experiment are shown.

(B-D) Metabolic and inflammatory pathways are regulated by macrophage lipid storage. Transcriptional profiles of THP-1 macrophages incubated in the presence/absence of ac-Lipo (50 μg/mL) for 2 days were determined using RNA sequencing. (B) Pathways regulated by lipid storage were identified using gene set enrichment analysis based on the hallmark gene set. (C) As a proxy for transcriptional activities, validated SREBP2 and NFκB target genes were ranked across the RNA sequencing results and enrichment scores calculated. (D) Top 10 regulated SREBP2/NFκB target genes from panel C are displayed. Results are based on two replicates per condition, and in panels (B) and (D), the size of each circle is scaled to match the absolute value of the respective x-axes.

(E) Multiple proteins are localized to LDs in both SUM159 and THP-1 cells. The $\log_2$ fold-change of the LD fraction/total input was plotted for proteins common to the LD fraction of both SUM159 and THP-1 cells. Intensity cutoffs of 3.066 and 2.02 were used for THP-1 and SUM159 cells, respectively. Proteins that were over the threshold in both cell

types are labeled in yellow, and proteins that fulfilled the criteria in only one of the cell types are highlighted in purple (SUM159) and blue (THP-1).

(F) High-fat diet induced protein relocalization. PCP profiles of RAB7A, HSD17B11, DHRS1, and RAB1A show strong signals in the LD fraction, indicating a relocalization of the protein to LDs under HFD conditions.

(G) Specific phosphosites of proteins are associated with increased LD targeting. Protein profiles of LPIN1, HSD17B13, AUP1, and TPD52l2 overlaid with profiles of an LD localization-specific phosphorylation event.

Abbreviations: ac-Lipo, acetylated apolipoprotein B-containing lipoprotein; CE, cholesterol ester; LD, lipid droplet; OA, oleic acid; TG, triacylglycerol.

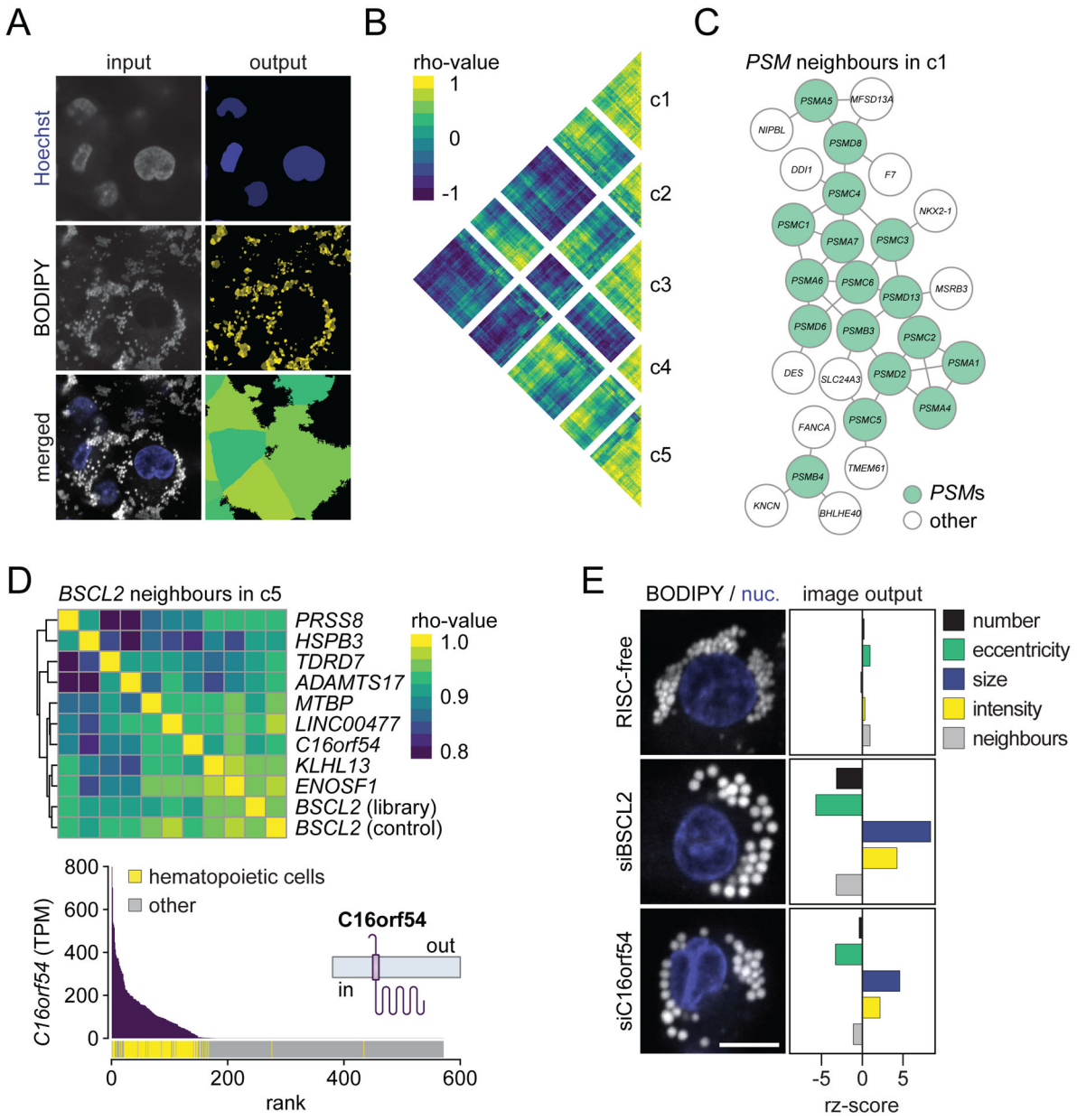

Figure 3: Clustering of RNAi screen image results identifies classes of hits with similar LD morphologies

(A) Image analysis extracts LD information at the single-cell level. After lipid storage induction by ac-Lipo, nuclei and LDs of THP-1 macrophages were stained using Hoechst and BODIPY, respectively. Images were acquired using a high-throughput confocal microscope, and image analyses were performed using CellProfiler. Segmented nuclei, LDs and cells are shown in the output column.

(B) Five classes of macrophage lipid determinants. RNAi screen hits (n=558) were pair-wise correlated, based on image information, and the resulting matrix was classified (c1-5) by hierarchical clustering. c1 contains predominantly proteasome genes. c5 is the *BSCL2*/seipin cluster.

(C) Knockdown of proteasomal subunits results in small and dispersed LDs. Network displaying the 17 proteasomal subunits and their closest neighbors identified in cluster 1. The three closest neighbors of each proteasomal gene in cluster 1 were extracted from the correlation matrix presented in panel (B) and added to the network. Genes are presented as nodes, and top three neighbors are connected by edges.

(D-E) Depleting *C16orf54* and *BSCL2* results in similar macrophage LD morphology. (D) In the upper panel, correlation scores for the 10 closest neighbors of *BSCL2* in cluster 5 are displayed. *BSCL2* occurs two times: once from the genome-wide library (library) and once as a median score of all BSCL2 control wells present on each plate (control). In the lower panel to the left, *C16orf54* gene expression data across 571 human cells and tissues were extracted from the FANTOM5 database. *C16orf54* transcript abundance was ranked from high to low, and samples from hematopoietic cells were highlighted. In the lower panel to the right, transmembrane helices in C16orf54 were predicted using TMHMM Server v. 2.0. (E) Representative confocal images and image analysis ouput for five features of RISC-free, siBSCL2, or siC16orf54 transfected macrophages from the original RNAi screen. Scale bar, 5 μm.

Abbreviations: c1-5, cluster 1–5; nuc., nucleus; PSMs, proteasomal subunits; TPM, tags per million.

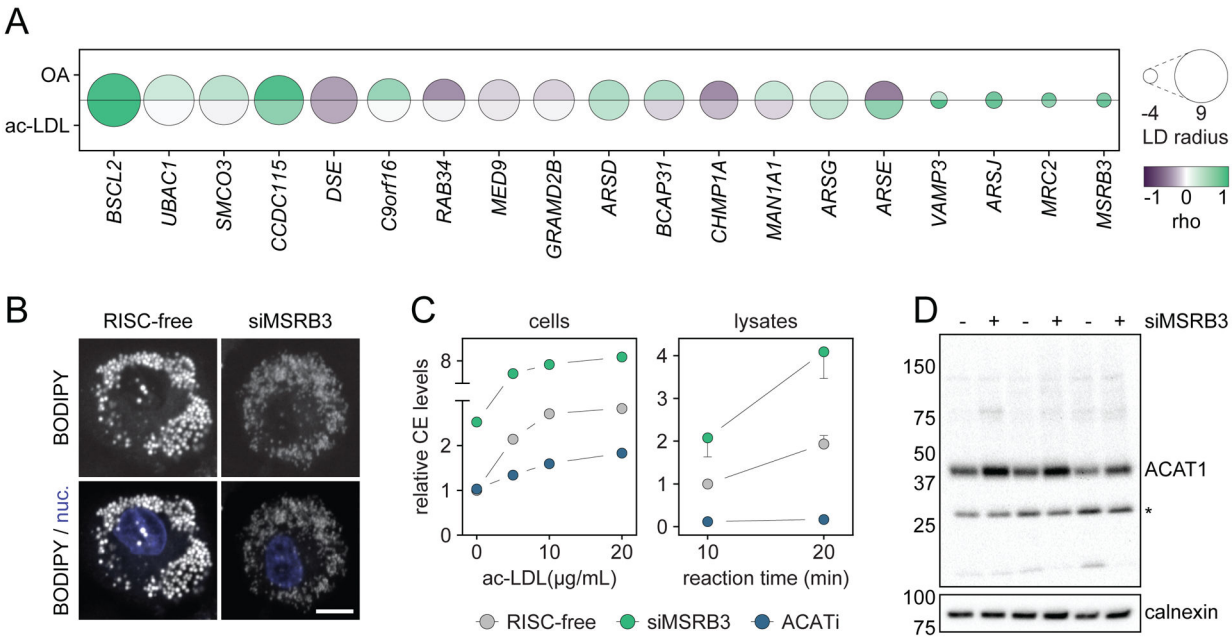

Figure 4: Secondary screening identifies genetic determinants of cholesterol ester versus triacylglycerol storage and identifies MSRB3 as a regulator of cholesterol storage

(A) 19 genes were re-screened in THP-1 macrophages with ac-LDL or OA. The size and color of the circle are proportional to the effect on LD size in the original screen and reproducibility in the two secondary screens, respectively.

(B) Macrophage *MSRB3* knockdown results in small and dispersed LDs. Representative confocal images of RISC-free or siMSRB3 transfected macrophages from the original RNAi screen. Scale bar, 5 μm.

(C) Cholesterol esterification is increased in *MSRB3*-depleted macrophages. THP-1 macrophage cholesterol esterification assays were performed in live cells (left panel) or lysates (right panel) 3 days post-transfection with siRNAs targeting RISC-free or MSRB3. ACAT inhibition was used as a control for the assays, and levels of extracted radiolabeled CE were determined by thin layer chromatography. In the left panel, endogenous production and exogenous uptake of cholesterol were reduced in cells by culturing them with compactin and without FBS, respectively. Subsequently, ac-LDL was added to the media for 7 hours out of which the 2 last hours were in the presence of radiolabeled OA. One representative experiment is shown. In the right panel, radiolabeled cholesterol was added to cell lysates, and CE formation was allowed for the indicated reaction times. Results are based on four replicates and data are represented as mean ± SD.

(D) ACAT1 protein levels are increased in *MSRB3*-depleted macrophages. Protein levels of ACAT1 and calnexin were determined by western blotting in THP-1 macrophages 3 days post-transfection of RISC-free or MSRB3 siRNAs. In addition to the ACAT1 band (approximately 45 kDa), an unspecific band marked by an asterisk was detected (approximately 30 kDa). Results display three independent experiments and molecular weight markers are indicated on the left side of the membranes.

Abbreviations: ACATi, ACAT inhibition; ac-LDL, acetylated low-density lipoprotein; LD, lipid droplet; min, minute; nuc, nucleus; OA, oleic acid.

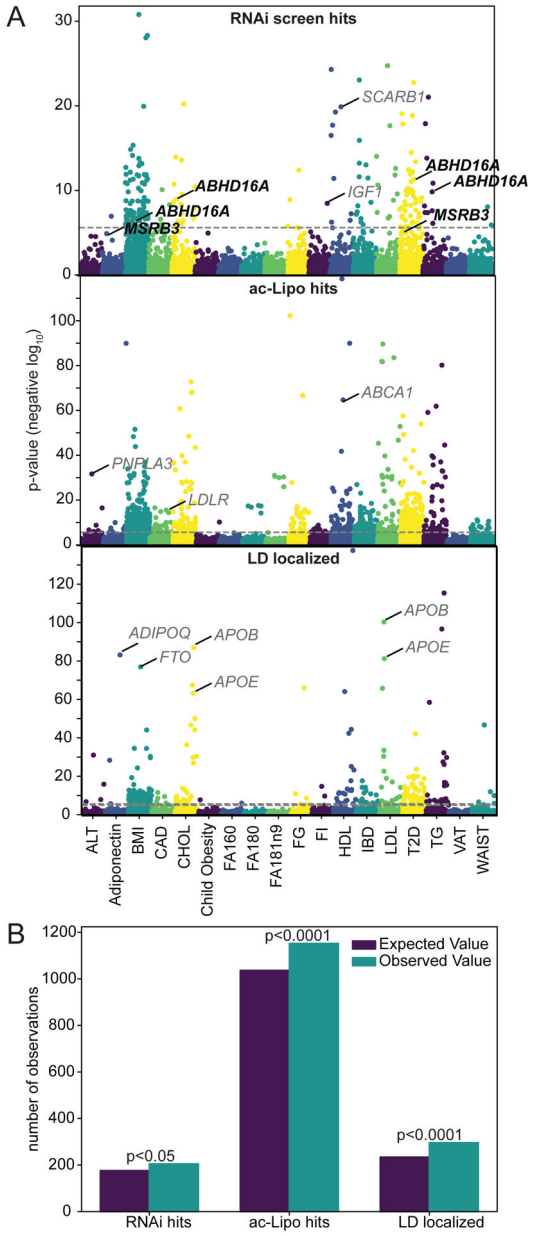

Figure 5: The LD portal gene sets are associated with human genetic traits

(A) LD portal data sets contain hits associated with metabolic phenotypes. Metabolic associations ($-\log_{10}$ p values) of genes that were either RNAi screen hits, transcriptionally regulated by ac-Lipo (adj. p-value<0.05), or LD localized proteins (from SUM159, THP-1, U2OS, HuH7 and murine proteomes) were analyzed for 19 metabolic phenotypes. A significance threshold value of p = 2.5e-6 was used (dotted line).

(B) LD portal screen hits select for genes associated with metabolic disorders. Chi-squared results of RNAi screen hits, genes transcriptionally regulated by ac-Lipo, and LD localized proteins that have significant associations with metabolic phenotypes vs. random selection. Chi-squared statistics= 5.00, 15.89, and 17.36, respectively.

Abbreviations: ALT, alanine transaminase; BMI, body mass index; CAD, coronary artery disease; CHOL, cholesterol; ChildObesity, child obesity; FA160, palmitic acid; FA180, palmitoleic acid; FA181n9, oleic acid; FG, fasting glucose; FI, fasting insulin; HDL, HDL cholesterol; IBD, inflammatory bowel disease; LDL, LDL cholesterol; T2D, type 2 diabetes; TG, triglyceride; VAT, visceral adipose tissue volume; WAIST, waist circumference