Establishment of a highly efficient gene disruption strategy to analyze and manipulate lipid co-regulatory networks

[1,2]Takeshi Harayama, [3]Tomomi Hashidate-Yoshida, [1,4]Auxiliadora Aguilera-Romero, [3,5]Fumie Hamano, [3]Ryo Morimoto, [3,6]Takao Shimizu, and [1]Howard Riezman

[1]Department of Biochemistry and National Centre of Competence in Research in Chemical Biology, University of Geneva, Quai Ernest-Ansermet 30, CH-1211 Geneva, Switzerland

[2]Institut de Pharmacologie Moléculaire et Cellulaire, Centre National de la Recherche Scientifique, Université Côte d'Azur, 660 Route des Lucioles, 06560 Valbonne, France

[3]Department of Lipid Signaling, National Center for Global Health and Medicine, Shinjuku-ku, Tokyo 162-8655, Japan

[4]Department of Cell Biology, Faculty of Biology, University of Seville, 41012 Seville, Spain and Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, 41012 Seville, Spain.

[5]Life Sciences Core Facility and [6]Department of Lipidomics, Graduate School of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan

Corresponding authors:

Takeshi Harayama: harayama@ipmc.cnrs.fr

Howard Riezman: howard.riezman@unige.ch

SUMMARY

Gene disruption has been drastically facilitated by recent genome editing tools. While the efficiency of gene disruption in cell culture has improved, clone isolation remains routinely performed to obtain fully mutated cells, potentially leading to artifacts due to clonal variances in cellular phenotypes. Here we report GENF, a highly efficient strategy to disrupt genes without isolating clones, which can be multiplexed. We obtained reliable lipidomics datasets from mutant cells generated with GENF, which was impossible when using clones. Through this, we found that an enzyme involved in congenital generalized lipodystrophy regulates glycerophospholipids with specific acyl-chains. We also demonstrate the possibility to dissect complex lipid co-regulatory networks and provide some common mechanisms explaining the adaptations to altered lipid metabolism. GENF is likely to contribute to all experimental setups affected by clonal variability and should be especially useful for -omics approaches.

INTRODUCTION

Gene disruption is a common strategy used in most biological fields to study the resulting phenotypic changes or to manipulate cellular functions. For this, it is obviously desirable that no phenotypic changes happen randomly during the process of gene disruption. Technologies based on the prokaryotic immune system CRISPR (clustered regularly interspaced palindromic repeats) have made gene disruption extremely easy [1]. In the commonly used CRISPR-Cas9 system, the RNA-guided Cas9 nuclease is expressed in cells to induce double-strand breaks in genomic regions determined by an engineered single guide RNA (sgRNA), eventually leading to the formation of gene disruptive indels. The sequence of the sgRNA not only determines target specificity, but also affects the efficiency of cleavage [2]. The efficiency of gene disruption by CRISPR-Cas9 has improved through advances in sgRNA design algorithms and delivery strategies of the components [2-4]. However, complete gene disruption in the whole population of cells remains difficult, and therefore many researchers isolate cellular clones after delivery of CRISPR-Cas9 component to obtain mutated cells. While it is known that different clones of cell lines can have phenotypic variabilities [5, 6], this potential pitfall of CRISPR-Cas9 technology has been discussed in the literature only rarely, and cellular phenotypes are often investigated through the analysis of a few representative clones. This can be very problematic in most fields, especially in -omics studies where a large number of variables is analyzed. Due to the presence of random variabilities introduced during clone isolation, it can be difficult to discriminate whether changes in -omics datasets are due to the genetic perturbation or due to clonal differences. In addition, clonal differences add noise to datasets, making quantitative changes between various mutants difficult to assess, and limiting the use of CRISPR-Cas9 in quantitative systems biology.

Lipidomics became a major -omics approach in biomedical research [7, 8]. It can identify hundreds of different lipid species, including the structurally diverse membrane lipids. Lipids affect the physicochemical properties of membranes as well as the activity of proteins in contact with them [9]. However, the molecular links between membrane lipid composition and biological processes are still poorly understood. It is therefore critical to accumulate more knowledge about mechanisms regulating membrane lipid composition. The metabolism of membrane lipids is complex due to the interconnection of metabolites from different lipid classes [9, 10]. It has been shown that the change in one lipid can affect the levels of other ones, including those that are metabolically distant, generating a network of co-regulated lipids [11]. This co-regulation leads to membrane lipid compositions that modulate immune responses upon various stimuli [11]. In addition, we recently found that sphingolipids (SLs) and ether phosphatidylcholine (ether PC) are co-regulated, which is critical to maintain the integrity of the secretory pathway [12]. While these studies establish the importance of lipid co-regulation, the mechanisms linking the quantitative change of one lipid to that of another one remains poorly understood. A combination of genetic disruption of lipid metabolic enzymes and lipidomics should be extremely useful to solve the problem. However, one requisite for such approaches would be that lipid co-regulation is correctly detected in the dataset, which is difficult when using clonal cells as discussed above. Therefore, a strategy that allows us to disrupt genes with an efficiency that is high enough without clone isolation is required. Here, we report a highly efficient strategy to disrupt genes in cell culture, which allowed us to generalize common mechanisms of lipid co-regulation.

RESULTS AND DISCUSSION

High mutational activity of CRISPR-Cas9 in a subset of cells

The initial idea of the gene disruption strategy came from a previous study, in which we established a PCR-based method to detect mutant clones generated with CRISPR-Cas9 [13]. In that study, we transfected cells with plasmids encoding CRISPR-Cas9 components to induce mutations in three genes (Sgpl1, Sgpp1, and Sgpp2) and analyzed 14 clones to obtain a triple mutant cell line. Only one clone was a full Sgpp2 mutant, showing that Sgpp2 sgRNA was only weakly active. Nevertheless, the other two targets were fully mutated in this clone. While this could have happened by chance, the result let us hypothesize that Sgpp2 disruption occurred in a subset of cells with high mutational activity, in which gene disruption was possible even with a poorly active sgRNA, thus enabling mutagenesis of all three targets. To test the reproducibility of the results, we further analyzed 28 clones after transfection with plasmids encoding Cas9 and sgRNAs targeting Sgpl1, Sgpp1, and Sgpp2 (Figure S1A). Transient transfection confirmed that Sgpp2 sgRNA was the least active (Figure S1B). A combination of PCR-based analysis (Figures S1C and S1D) and Sanger sequencing (Figure S1E) detected two full mutants of Sgpp2, in which Sgpl1 and Sgpp1 were also fully mutated (Figures S1E and S1F). The small number of Sgpp2 mutants obtained did not allow us to estimate precisely the statistical significance of our observations, but the result strengthened our hypothesis that a subset of cells transfected with plasmids encoding CRISPR-Cas9 components has a high mutational activity.

Isolation of highly mutated cells with a co-targeting strategy

In order to isolate cells with high mutational activity, we designed a strategy based on HPRT co-targeting [14]. In this strategy, mutation is introduced with the CRISPR-Cas9 system in the Hprt1 gene (in addition to target genes), which confers resistance against the drug 6-thioguanine (6-TG). Thus, the selection of Hprt1 mutant cells with 6-TG after co-targeting enables the selection of cells correctly transfected, which was proposed as a pre-enrichment strategy of target-mutated cells before isolating clones [14]. We hypothesized that by using conditions where Hprt1 mutation are not efficient, either due to poor sgRNA design or low plasmid amounts, 6-TG resistance would be acquired only in the subset of cells that have a high mutational activity (Figure 1A). In such conditions, low yield but high target mutation rates were expected, potentially avoiding the isolation of clones. To test this, we expressed in rat hepatoma McA-RH7777 cells the CRISPR-Cas9 components from two plasmids. One was based on pX330, encoding both Cas9 and the sgRNA against the target gene. The other, pHPRTsg, encoded only the sgRNA against Hprt1 gene. Given that Hprt1 mutation occurs only when both plasmids are delivered, this design ensured the expression of target sgRNA. In addition, this allowed us to vary the ratio between the two plasmids and reduce the efficiency of Hprt1 mutation. We performed co-targeting using six different sgRNAs against rat Hprt1 and found one (#4) with low yield of 6-TG-resistant cells (Figure 1B). Consistent with our hypothesis, target (Sphk2) gene mutation rates were highest in cells co-targeted using Hprt1 sgRNA#4, but did not reach complete deletion. By reducing the amount of pHPRTsg plasmid encoding Hprt1 sgRNA#4, we could further reduce the yield of cells, which resulted in a near-complete target gene mutation (Figure 1C). This high level of gene disruption was not achieved when using pHPRTsg plasmid encoding another sgRNA (#1) even at low amounts (Figures 1B and 1C). Thus, a combination of poor sgRNA design and low copy number of Hprt1 sgRNA was critical for achieving high rates of target mutation after 6-TG selection. We named this strategy GENF (Gene co-targeting with non-efficient conditions) as a method to achieve highly efficient gene disruption without the need to isolate clones. Throughout this study, target mutation rates are calculated from the loss of wild type signals in Sanger sequencing (analyzed by deconvolution methods [15]), which was confirmed to give more precise values than the sum of mutant signals (Figures S2A-S2C). We further tested GENF in 8 different targets, and found that single target disruption was extremely high (98.6% in the lowest case and 99.9% on average, Figure 1D). Although not being as high as for single target mutation, double or triple gene disruption GENF was highly efficient (83.8% in the lowest case and 93.5% on average, Figures 1E and 1F). To test the limit of GENF, we induced 8-plex mutations, for which day-by-day variation became larger (Figure 1G). However, in two out of three experiments high target mutation rates were obtained (71.6% in the lowest case and 92.7% in average). Thus, GENF generates very highly mutated cells even under multiplex settings.

Mismatched sgRNAs facilitate GENF

When applying GENF to a HeLa subline, HeLa MZ cells, we had difficulties to obtain an sgRNA against human HPRT1 that would reproducibly lead to near-complete target gene disruption. Therefore, we investigated whether we could use sgRNAs with mismatches as a strategy to achieve low efficiency more easily (Figure 1H). A single mismatch in HPRT1 sgRNA did not affect the yield of 6-TG-resistant cells and did not increase post-selection target mutation rates (Figures 1H and 1I). On the other hand, sgRNAs with two mismatches enabled the generation 6-TG-resistant cells only when the mismatches were in distal-most positions from the NGG proximal adjacent motif (Figures 1H and 1I). Such sgRNAs enabled complete post-selection target (CERS6) gene mutation in HeLa MZ cells (Figure 1I). Extremely high gene disruption was achieved with this protocol in multiple targets, including in multiplex experiments (Figures 1J and 1K). One of the few targets we could not mutate efficiently with GENF was HSPA9, which we found to be an essential gene according to DepMap database (Figure S3A). We note that another essential gene, TMED2, could be mutated with GENF, for which we found that indels were strongly enriched for in-frame mutations (Figures S3A and S3B). The other two targets with low mutation rates were ECI2 and HBB (Figure 1J). Transient transfection experiments revealed that the sgRNAs for these targets were extremely poorly active (Figure S3C). For the other targets, nearly perfect mutations were achieved without pre-testing of sgRNA activity. Therefore, GENF enables highly efficient gene disruption also in HeLa MZ cells, unless the sgRNA against a target is very poorly active, or when the target is essential.

Additional drug-gene pair for GENF

One limitation of GENF was that it could not be repeated multiple times to add more mutations. To overcome this, we took advantage of published genome-wide CRISPR screening results that study drug resistance. Based on a study reporting that POR gene mutation confers resistance against paraquat [16], we tested whether this drug-gene pair could be used for GENF. The co-transfection of a plasmid encoding an sgRNA against POR led to the formation of small colonies after paraquat treatment (Figures 2A and 2B). Performing POR co-targeting using an sgRNA with two mismatches enabled a reproducible near-complete depletion of a target gene (CERS6), thus establishing an additional drug-gene pair for GENF (Figures 2A-2C). The small size of colonies is consistent with a partial toxicity of paraquat in POR mutant cells [16], thus we recommend the use of HPRT1 as a primary target for GENF. Having an additional drug-gene pair for GENF, we could repeat gene disruption two times. Using this, we could overcome the slightly lower efficiency of highly multiplexed gene disruption with GENF (Figures 2D-2F). Using HPRT1 as the co-targeted gene, we generated with GENF a cell line having mutations in all six CERS genes (designated ΔCERS) with high but imperfect efficiencies (88.5% for the lowest case and 93.4% on average, Figure 2F). Another step of GENF could be repeated using POR as the co-targeted gene, resulting in paraquat-resistant cells (Figure 2E). After repeating GENF, we obtained a cell line (designated Δ/ΔCERS) having near-complete mutations in all CERS genes (95.9% for the lowest case and 99.2% on average, Figures 2D-2F). Although CERS genes are critical for the synthesis of complex sphingolipids, Δ/ΔCERS cells could be maintained in cultured similarly to control cells, probably due to the supply of lipids from the serum (as discussed later).

In another experiment, we tested whether we could manipulate metabolite levels with GENF and revert the phenotypes by adding mutations through another round of GENF. For this, we generated a cell line (designated L/P/P) lacking SGPL1, SGPP1, and SGPP2, which are enzymes that reduce the amount of sphingosine 1-phosphate (So1P) through dephosphorylation or degradation (Figure 2G) [17]. From this cell line and a control cell line (with only HPRT1 mutated, designated C1), we generated cell lines with mutations in the two kinases (SPHK1 and SPHK2) that produce So1P (Figures 2G and 2H). L/P/P cells were successfully generated with high mutation efficiency in the first round of GENF (97.1% for the lowest case and 99.0% on average, Figure 2I). In the second round, both SPHKs were completely mutated in L/P/P cells, generating L/P/P/K/K cells, or in the control cell line C1, generating C1/K/K cells (Figures 2H and 2J). Control cells of this second step, in which only POR was mutated, were designated C1/C2 and L/P/P/C2 cells (Figure 2H). So1P levels in L/P/P cells were drastically increased (>270 fold), while SGPL1 mutation alone had only a marginal effect (Figure 2K). Importantly, L/P/P/K/K cells had comparable So1P levels to control cells or C1/K/K cells, demonstrating that phenotypes of mutant cells can be reverted through repeated GENF (Figure 2L).

Characterization of mutations obtained with GENF

CRISPR-Cas9 gene deletion is known to have some limitations, such as the variability in on-target effects and the presence of off-target effects [18, 19]. Since GENF was designed specifically to address the issues related to clonal variability, we did not assume that these other limitations should be improved. Nevertheless, it was important to characterize them. We tested the on-target mutation patterns using four mutant McA-RH7777 cell lines lacking either Agpat1 or Agpat2, generated with two different sgRNAs each. The cell lines, including a control, were generated in triplicate, making a total of 15 cell lines. While different sgRNAs led to diverse indel patterns, different cell lines generated with the same sgRNA had very similar patterns (Figures 3A and 3B). This showed that GENF generates reproducible on-target effects from the same sgRNA, which should be beneficial when repeating experiments between laboratories. We also found that indel patterns are reproducible when comparing cell lines generated with the same sgRNA but on different days (Figure S4A). Our analysis was not suitable to detect kb-sized large-scale indels as well as transpositions or inversions, all having been reported with CRISPR-Cas9 experiments [18]. However, extrapolating from the extremely reproducible patterns seen in our experiments, it is likely that these adverse on-target effects are also happening to a similar extent when using the same sgRNA. We also note that such adverse mutations should be diluted in the polyclonal population when using GENF, while they should be present stochastically when using clones.

We next analyzed off-target effects using the HBB mutant cell line generated with GENF. This cell line was generated with an sgRNA having a well characterized off-target candidate list (Figure 3C), which was obtained from cleavage patterns of genomic DNA preparations from cells [20]. A previous study demonstrated that this *ex cellulo* cleavage enables a good prediction of *in cellulo* off-targets [20]. We found that GENF has off-target effects (Figure 3D), but also observed that the degree of mutation in the off-target sites were predictable from *ex cellulo* cleavage (Figure 3E). We also analyzed off-target cleavage in CERS4 mutant cells, which was generated using an sgRNA with the lowest predicted specificity score among those used in this study. Off-target mutations were seen in these cells, but only in targets with three mismatches and not in those with four mismatches (Figures S4B and S4C). In addition, the off-target mutation rates were lower than the on-target 100% mutation (Figure S4C). Therefore, GENF has off-target effects similar to other CRISPR-Cas9 protocols, but the off-targets remain predictable. We note that the Cas9 protein used in this study was from the first generation, and recent “high fidelity” versions [21] might improve the specificity of GENF. We did not test this possibility in this study, since the use of mismatched HPRT1 sgRNAs would not be possible.

GENF generates reliable lipid analysis results

We next investigated if the use of polyclonal mutant cells generated with GENF could improve phenotype analyses by eliminating clonal variabilities. For this, we analyzed phosphatidylcholine (PC, the most abundant glycerophospholipids) acyl-chains in Agpat2 mutant McA-RH7777 cells generated either by clone isolation (generated in parallel to a previous study [22]) or by GENF. This was based on another previous study where we found that enzymes from the Agpat family regulate PC acyl-chains and affect membrane physical properties [23]. Consistent with the literature describing phenotypic variabilities between clones [5, 6], we found a variability in PC acyl-chains between clones (Figures 3F and S4D). In contrast, differences between polyclonal cell lines were much smaller, and we could find a distinct PC acyl-chain profiles between control and Agpat2 mutant cells (Figures 3G and S4E). This revealed that Agpat2 is responsible for the production of PC 34:1 (34 carbons and 1 double bond in acyl-chains). Using these datasets, we performed multivariate analyses to compare the performance of both approaches (clone isolation versus GENF) to detect lipid changes induced by gene disruption. Principal component analysis (PCA) could not discriminate control and Agpat2 mutant clones with the first dimension, while mutants generated with GENF were clearly discriminated (Figures 3H and 3I). When splitting the datasets into two groups with K-means clustering, data from one control clone was clustered together with the two Agpat2 mutant clones (Figure 3H). On the other hand, datasets obtained from GENF-generated control and Agpat2 mutant cells were perfectly split (Figure 3I). Finally, when we assessed data similarity using hierarchical clustering, clones could not be correctly separated based on Agpat2 genotypes, while cell lines generated by GENF were clearly distinguished (Figures 3J and 3K). These results demonstrate that polyclonality is critical to obtain reliable quantitative datasets, and that GENF is useful for that purpose.

Agpat2 regulates linoleic acid levels in PC

Having established GENF as a strategy to obtain mutant cells for reliable lipid quantitative analysis, we applied it to ask biological questions. First, we investigated more details about the roles of Agpat2 in the regulation of PC acyl-chains. Since standard culture conditions (medium supplemented with 10% fetal calf serum) are poor in polyunsaturated fatty acid (PUFA) supply and cell lines are in general PUFA-poor [24, 25], we investigated PC acyl-chains of cells supplemented with three PUFAs: linoleic acid, arachidonic acid, and docosahexaenoic acid. In such conditions, the effect of Agpat2 mutation on PC 34:1 became smaller, and rather PC 34:2 was decreased in the mutants (Figure 4A). Fragmentation analysis in mass spectrometry revealed that this decrease in PC 34:2 is attributable to a decrease in linoleic acid (PUFA with 18 carbons and 2 double bonds) in PC (Figures S5A-S5C). Gas chromatography analysis confirmed that linoleic acid levels are specifically decreased in glycerophospholipids of Agpat2-mutant cells (Figure 4B). Since Agpat2 encodes an enzyme with lysophosphatidic acid acyltransferase (LPAAT) activity, we analyzed how Agpat2 disruption affects this activity and its substrate preference. Agpat2 mutants had decreased LPAAT activity (Figure 4C) as well as the preference for linoleic acid incorporation (Figure 4D and Figures S5D-S5G). Our previous study suggested that LPAAT activity regulates linoleic acid levels in PC [23], and this study confirms this observation and identifies Agpat2 as the responsible enzyme. We compared the selectivity of LPAAT activity and gene expression profiles of Agpat enzymes in various tissues, and found that tissues with high Agpat2 expression have an LPAAT activity with higher preference for linoleic acid (Figure S5H). In addition, cells stably overexpressing Agpat2 had increased PC 34:2 levels (Figure 4E). All these results establish Agpat2 as an enzyme regulating linoleic acid levels in PC. Of note, mutations in human AGPAT2 lead to congenital generalized lipodystrophy type 1, a genetic disease characterized by a complete loss of adipocytes [26]. Thus, our study identifies a potential link between membrane lipid acyl-chains and a genetic disease with impaired adipogenesis, which will be important to investigate in the future. A recent study identified that the PUFA docosahexaenoic acid promotes osteoblast differentiation of mesenchymal stem cells [27], which are also precursors for adipocytes. Combined with this study, our results suggest that different PUFAs work as a switch to regulate the direction of mesenchymal stem cell differentiation.

Manipulation of sphingolipid acyl-chains with GENF

Using GENF, we generated a series of mutant cells lacking individual ceramide synthases (CERS1-CERS6), which are enzymes incorporating the N-acyl-chains of sphingolipids (SLs) [28] (Figure 1J). Published RNA-seq data of HeLa MZ cells [29] revealed that only CERS2 and CERS5 are expressed in the parental cell line (Figure S6A). Therefore, mutants of the four non-expressed CERSs should not have any phenotypes unless GENF generated artifacts. We performed a lipidomic analysis of the mutant cell lines to investigate changes in SLs, while glycerophospholipids and cholesterol were also analyzed (see supplemental note for discussion about the methodologies and their limitations). CERS2 mutants had drastic decreases in C24 ceramide (C24 Cer) and compensatory increases in C16 ceramide (Figures 5A-5C). The same pattern was seen in the complex SLs sphingomyelin (SM) and hexosylceramide (HexCer) (Figures 5D-5F and S6B-S6D). On the other hand, CERS5 mutant had decreased C16 ceramide and C16 SM levels, while other mutants had no changes in SL N-acyl-chains (Figures 5A, 5B, 5D, 5E, S6B, and S6C). The results show that CERS2 and CERS5 compete for the same substrate, and that the loss of CERS2 led to an increased flux of sphingosine usage by CERS5 (Figures 5G and 5H). Indeed, the increase of C16 SLs was absent in Δ/ΔCERS cells lacking all CERSs despite the drastic loss of C24 SLs (Figures 5I, 5J, and S6E). We note a considerable amount of C16 SLs detectable in Δ/ΔCERS cells, which is probably due to the supply of SLs from the serum added in the culture medium, as has been documented [30]. The reverse reaction of ceramidases has also been reported to generate ceramides [31]. Despite the drastic reductions in SLs seen in Δ/ΔCERS cells (Figure 5K), the cells could be maintained in culture similarly to control cells with similar passage frequencies, showing the strong abilities of cell lines to survive with minimal amounts of some SLs.

We further analyzed the whole lipidomics datasets of single CERS mutants. Consistently with the lack of their expression in parental cells, CERS1, CERS3, CERS4, and CERS6 mutants had no significant changes in all lipid analyzed (Figures S6F-S6I). It is unlikely that clonal mutants, with their high variances, could lead to similar unchanged datasets comprising >230 lipids. CERS5 mutants had only two significantly changed lipids, both being C16 SLs (Figure 5L). Unexpectedly, CERS2 mutants had a drastic rearrangement of the whole lipidome (Figures 5M-5P). In addition to the changes in SL N-acyl-chain profiles (Figure 5M), we found that PC with an ether bond at the *sn*-1 position (ether PC or PC(O-)) is decreased in CERS2 mutants (Figure 5N). Multiple phosphatidylethanolamine (PE) species were increased, but the effect was stronger with PE having an ether or vinyl-ether bond at the *sn*-1 position (collectively labeled ether PE or PE(O-) in this study) (Figure 5O). Some decreases in phosphatidylserine (PS) species were also seen (Figure 5P). This revealed a whole lipid co-regulatory network associated with a decrease in C24 SL production by the loss of CERS2. Our results show the power of GENF to generate mutants and manipulate metabolite levels, and to detect co-regulatory networks associated with the manipulations. It is important to note that the data reliability of GENF is especially critical to analyze co-regulatory networks. With clonal cells, variabilities in datasets would make it difficult to analyze whether secondary changes are due to co-regulation or data variations.

Dissection of co-regulatory networks with multiplex GENF

Due to the importance of the co-regulation between SLs and ether PC to maintain the integrity of the secretory pathway [12], we further analyzed the mechanisms of ether lipid changes in CERS2 mutants. CERS2 mutants had decreased ether PC levels, while increases were seen for PE (hereafter, PE means non-ether PE) and more strongly for ether PE levels (Figures 6A-6C). The sum of ether lipids was unchanged (Figures 6D), suggesting that ether lipid biosynthesis *per se* was not affected in CERS2 mutants. Rather, it was likely that the balance of head group incorporation into the common precursor, ether diacylglycerol (ether DAG) was affected (Figure 6E). One of the degradation products of SLs is ethanolamine phosphate, which can be used for PE and ether PE synthesis after conversion into CDP-ethanolamine (Figure 6E) [17]. We hypothesized that CERS2 mutation increased the flux of SL degradation, upregulating ethanolamine phosphate production. To investigate this hypothesis, we generated a series of mutant cells lacking CERS2, PISD, or SGPL1 either alone or in combination (Figures 1J and 1K), and did lipidomics analyses. PISD encodes PS decarboxylase [32], which generates PE independently of ethanolamine phosphate, and SGPL1 is the enzyme responsible for SL degradation [17] (Figure 6E). PISD mutants had decreased levels of both PE and ether PE, while ether PC was largely increased (Figures 6F-6I). CERS2 mutation reversed these changes (compare PISD mutants and PISD/CERS2 mutants), consistent with the hypothesis that the loss of CERS2 increases ethanolamine phosphate production through SL degradation (Figures 6F-6I). In contrast to the large decrease seen in CERS2 mutants, ether PC levels were close to control levels in SGPL1/CERS2 mutants, in which SL degradation does not occur (Figure 6F). The results demonstrated that increased SL degradation in CERS2 mutants indeed contributes to the decrease in ether PC.

Importantly, the comparison between SGPL1 mutants and SGPL1/CERS2 mutants suggested an additional layer of co-regulation between SLs and ether lipids. In SGPL1/CERS2 mutants, ether PC was lower and ether PE was higher than in SGPL1 mutants (Figures 6F and 6G). A trend for decreased PE levels was also seen in SGPL1/CERS2 mutants, while the sum of PE and ether PE was only slightly increased (Figures 6H and 6I). These observations suggested that CERS2 deletion can also affect ether lipid levels independently of SL degradation, which could be through the promotion of ether PE synthesis at the expense of ether PC and PE. Although we do not know the precise mechanism of this additional layer of co-regulation, the increase in C16 SLs or the decrease in C24 SLs might affect the activity of enzymes adding head groups on ether DAG, either by reducing the transfer of phosphocholine or increasing the transfer of phosphoethanolamine (Figure 6E). Interestingly, an enzyme catalyzing the transfer of the phosphoethanolamine head group, ethanolamine phosphotransferase 1 (EPT1, encoded by SELENOI gene), is localized to the Golgi apparatus [33], where the biosynthesis of complex SLs takes place. This enzyme is more important for ether PE synthesis than for PE synthesis, and has a preference for substrates containing PUFAs [34]. We found in CERS2 mutants increased levels of multiple PE and ether PE species containing PUFAs with more than four double bonds, while some mono- and di-unsaturated species also increased (Figures 6J and 6K). Importantly, CERS2 mutations in an SGPL1 mutant background led to exclusive increases in PUFA-containing species with more than four double bonds (Figures 6L and 6M). This preference for ether PE synthesis with a strong preference for PUFA-containing species makes EPT1 a strong candidate as the enzyme regulated by SLs to mediate the co-regulation with ether lipids. As a conclusion, we could dissect the co-regulatory network between SLs and ether lipids, and revealed two distinct layers of regulation, one being due to a metabolic link and the other, which still requires further investigation, involving enzyme regulation by membrane lipid composition. This shows the power of combining GENF with lipidomics to dissect co-regulatory networks and generate hypotheses, through the analysis of a large number of mutants with single or multiplex mutations.

Lipid co-localization causes co-regulation

Cholesterol is well documented as an interaction partner of SLs [35]. The interaction with SLs has been shown to be critical to maintain cholesterol levels in the plasma membrane. We saw no change in cholesterol levels in CERS2 mutants (Figure 7A). However, a drastic decrease in cholesterol levels was seen in Δ/ΔCERS mutants (Figure 7B). Thus, our results confirm that the interaction with SLs maintains cholesterol levels, and we found that this interaction is independent of the type of SL N-acyl-chains. Thus, lipid-lipid interaction is another factor causing the co-regulation between lipids.

Another change observed in CERS2 mutants was the reduction of some PS species (Figure 5P). Interestingly, PS 36:1 (mainly with 18:0-18:1 acyl-chains) was the only reduced lipid among abundant PS species (Figure 7C). This change was seen in CERS2 mutants and Δ/ΔCERS mutants, but not in CERS5 mutants (Figures 7D and 7E). The same effect of CERS2 deletion was seen in SGPL1 or PISD mutant backgrounds (Figure 7F). These observations suggest that PS 36:1 decreased due to the reduction of C24 SLs, not the increase in C16 SLs, and that this change was not due to the metabolic link between SL degradation and glycerophospholipid synthesis. PS 36:1 was previously shown to co-localize with proteins with a glycosylphosphatidylinositol (GPI)-anchor through cholesterol-dependent trans-bilayer interactions [36]. This co-localization required actin-mediated local clustering of this PS species on the inner leaflet. Our results suggest that a similar co-localization occurs between C24 SLs and PS 36:1, thereby extending the residency time of the latter in the plasma membrane. We also observed decreases in minor saturated PS species (PS 32:0 and PS 34:0, Figures 7G and 7H). In addition, although many PE species were increased in CERS2 mutants, PE 36:1 was reduced (Figure 6J). Thus, it is likely that C24 SLs, mainly localized in the outer leaflet of the plasma membrane, generate a local environment that affect the physical properties of the opposing leaflet, which recruits PS 36:1 and other monounsaturated or saturated lipids, probably through local membrane physical properties. C16SLs, in contrast, are unable to affect the opposing leaflet similarly (PS 36:1 is reduced in CERS2 mutants despite the increase in C16 SLs). Our results provide experimental evidence for a recent concept of "hand-shaking" interactions between C24 SLs and PS 36:1, which was proposed based on correlations in lipidomics datasets and molecular dynamics simulations [37, 38]. Our results extend this concept, since they suggest that PE 36:1 and saturated PS species can also participate in the local domains generated by C24 SLs, although PS 36:1 remains the major component of the cytosolic leaflet. The nanodomains proposed here have differences from classically proposed "lipid rafts" that are generated in the plasma membrane through phase separation, which is triggered by the immiscibility between saturated lipids (plus cholesterol) and unsaturated lipids [39]. If this immiscibility was the driving force for the generation of nanodomains proposed here, C16 SLs should have behaved similarly to C24 SLs. In addition, the major lipids in the inner leaflet that contribute to the nanodomains proposed here are mono-unsaturated, which is also in contrast to classical lipid raft theories. On the other hand, the importance of SLs to maintain cholesterol levels is in line with classical theories. Thus, we propose that lipid nanodomains in the plasma membrane are more diverse and affected by the N-acyl-chains of SLs. Those containing C24 SLs are enriched with saturated or monounsaturated PS and PE species in the inner leaflet, while those containing C16 SLs are not (Figure 7I). Lipid immiscibility between saturated and unsaturated lipids is unlikely to enable the formation of nanodomains with SLs of distinct N-acyl-chains, thus the formation of nanodomains might be triggered by the local clustering of proteins that can interact with specific SLs.

Conclusions and future directions

In this study, we developed GENF as a strategy for highly efficient gene disruption. While GENF is based on the selection of cells that can mutate their targets even with a non-efficient sgRNA, the mechanism how it works remains mysterious. One mechanism could be that we are selecting cells that received more plasmids and expressed higher amounts of Cas9 and sgRNAs. However, in a study using cell sorting to isolate transfectants with high Cas9 levels, gene disruption became more efficient but not perfect [40]. In contrast, when using GENF non-disrupted alleles are barely detected. Thus, it is likely that GENF is not solely based on the selection of high expressors or CRISPR-Cas9 components, but rather involves other factors such as the epigenetic state of the genome. Once optimized, GENF is a very simple method involving only a transfection, drug selection, and occasional passages. This simplicity makes it very easy to generate multiple mutant cell lines in parallel in a time- and cost-efficient way.

The biggest advantage of GENF is the avoidance of artifacts related to clonal variances. The discovery of PC acyl-chains being regulated by Agpat2 provides a good example of a finding that would be difficult to make with clonal cell lines. The use of polyclonal cell lines is recommendable in most biological experiments where control and mutant cells are compared. We note that polyclonal mutant cells have already been used in previous studies, in which mutants were generated by lentiviral delivery of CRISPR-Cas9 components and drug selection of transduced cells [41]. While such approaches could be a good alternative to GENF, some studies revealed that the transduced cell populations are not always fully mutated [42]. This incompleteness could be problematic especially when multiplex gene disruption is performed to revert phenotypes, as has been done in our analysis of the co-regulation of ether lipids and SLs. In addition, GENF is a virus-free approach, enabling its use in most laboratories.

Through a combination of highly efficient gene disruption enabled by GENF and lipidomics analyses, we could obtain reliable datasets to propose novel concepts in lipid biology. Based on the continuity from our previous research [22, 23], we mainly focused on the roles of membrane lipid acyl-chains, and revealed a potential role of reduced linoleic acid-containing glycerophospholipids in the genetic disease congenital generalized lipodystrophy type 1. We also found that the N-acyl-chains of SLs have different roles in the maintenance of lipid homeostasis, likely through their effects on lipid metabolic enzymes as well as through the formation of distinct lipid nanodomains. Based on our results, we propose various mechanisms how lipid co-regulatory networks are shaped (Figure 7J). One mechanism involves the metabolic connection between lipids, which can be direct (when two lipids compete for the same substrate to be made) or distant (see the example of C24 SLs and ether PE, which are very indirectly connected through ethanolamine phosphate). Another mechanism involves the effect of membrane lipid composition on lipid metabolic enzymes. We propose in this study that EPT1 [34] is regulated by membrane SLs, but this enzyme being a multipass transmembrane protein, direct evaluation of this hypothesis will require extensive optimizations to purify the enzyme and reconstitute it in different liposomes, which is beyond the focus of this study. Finally, we propose that lipid co-localization is an important mechanism for lipid co-regulation.

While we used GENF to study lipid biology, the method is clearly not limited for this purpose. It is very likely that the majority of quantitative traits of cells are affected by clonal variances, and thus the accuracy of studies using mutant clones generated with CRISPR-Cas9 might be strongly affected, if not erroneous. GENF was specifically developed to avoid artifacts caused by clonal variances, which is a severe limitation of CRISPR-Cas9 if combined with clone isolation. Indeed, while the danger of using clones has already been discussed [5, 6], a large fraction of studies still uses isolated clones. GENF should be especially useful for -omics studies that analyze a large number of variables, such as transcriptomics or metabolomics. On the other hand, we note that GENF does not solve all the limitations of CRISPR-Cas9, such as off-target effects [19], deleterious on-target mutations [18], in-frame mutations, or the limitation in target sites due to restrictions of protospacer adjacent motifs [43]. However, researchers have paid more attention to these limitations, and strategies are being developed to solve them, such as the development of improved Cas9 variants [21]. A combined use of GENF with Cas9 variants will be an interesting approach for the future.

FIGURE LEGENDS

Figure 1. Establishment of GENF

(A) Design of GENF and the plasmids (pX330 and pHPRTsg) used for it. HPRT co-targeting experiments performed under standard conditions (top) or conditions where Hprt1 disruption is unfavorable (bottom, corresponding to GENF) are illustrated. (B and C) Co-targeting of Sphk2 and Hprt1 in McA-RH7777 cells using different sgRNAs against Hprt1 at various percentage of pHPRTsg plasmid. Cells were stained with Crystal Violet after culture for the indicated time with or without 6-TG selection. The non-selected cells were prepared to ensure that similar cell numbers were seeded. The percentage of Sphk2 mutation (target mutation % post-selection) is shown under the bottom panel. (D-G) Mutation rates after GENF for 8 genes targeted either alone (D), in duplex (E), in triplex (F), or in 8-plex (G). Experiments performed on different days are distinguished by colors. (H) Design of the sgRNA against HPRT1 (sgRNA with a full match, fm) or its 14 variants (mm 1-14) with mismatches (in red) used to establish GENF in HeLa MZ cells. The sgRNAs that enabled the generation of 6-TG resistant cells are labeled with a tick and the others with a cross. PAM: protospacer adjacent motif. (I) Co-targeting of CERS6 and HPRT1 in HeLa MZ cells using representative HPRT1 sgRNAs with (numbers correspond to mm# in (H)) or without (fm) mismatches. 1% (w/w) of pHPRTsg was used. Panels are organized similarly to (B) and (C). (J and K) Mutation rates after GENF for the indicated genes targeted in HeLa MZ cells either alone (J) or in multiplex (K). (K) Targets under the same lines were mutated together. See also Figures S1-S3.

Figure S1. Preliminary experiment leading to the idea of GENF.

(A) Overview of the experiment. McA-RH7777 cells were co-transfected with three pX330-based plasmids encoding Cas9 and sgRNAs against Sgpl1, Sgpp1, and Sgpp2. Transfected cells were analyzed in bulk to test the efficiencies of different sgRNAs, while one part was reseeded in dishes sparsely to obtain colonies. Colonies were directly lysed to obtain genomic DNA and screen for Sgpp2 mutants. Sgpl1 and Sgpp1 mutation was analyzed in the obtained Sgpp2 mutants. (B) In three experiments (distinguished by colors), mutation rates of Sgpp2 were always the lowest after transiently co-transfecting pX330-based plasmids. (C and D) Screening of Sgpp2 mutants by competition-based PCR. The method uses a mixture of primer that enables the detection of mutant clones by a reduction of the amplicon labeled “in” compared to the one labeled “out”, since one of the primers for “in” amplicon binds to the sgRNA target. See reference [13] for details. (C) Result of competition-based PCR. PCR patterns in the numbered lanes are from the clones that were later confirmed to be fully mutated. (D) Quantification of signal ratios between the two amplicons in (C). Five clones were selected for Sanger sequencing, and two of them were full mutants. (E) Sanger sequencing of the three targets in parental cells or clone 5. Target regions are underlined with an arrow. (F) Summary of mutations seen in clones 5 and 16. Note the absence of wild type alleles in Sgpl1 and Sgpp1, despite the fact that the clones were pre-selected based only on Sgpp2 mutations.

Figure S2. Evaluation of a method to analyze high mutation rates.

(A) Sanger sequencing of PCR amplicons from wild type or mutant HPRT1 locus, either alone or mixed at the indicated ratio. (B) TIDE (tracking indels by decomposition) analysis [15] from the Sanger sequencing results of (A). The calculated frequency of alleles with the indicated indel size is illustrated. Note that as little as 1% of wild type allele can be detected (with an indel size of zero, red bar in top panel) with TIDE. (C) Comparison of expected mutation rates (from the ratio of mixed PCR amplicons) and observed mutation rates calculated in two ways. Observed mutation rates were calculated either from the sum of mutant alleles detected with TIDE ("efficiency", which is usually used) or by subtracting the detected wild type alleles from 100 ("100-%WT"). Note that the latter gave a better estimate of mutation rates.

Figure S3. Analysis of failed cases of GENF.

(A) Essentiality of selected targets. Gene effect scores were obtained from Depmap database [44]. The scores illustrate the degree of sgRNA depletion/enrichment from cell populations during genome-wide CRISPR-Cas9 screenings in multiple cancer cell lines. A shift of the traces to the left (as seen for HSPA9 and TMED2) illustrate that the sgRNA targeting the gene is depleted upon culture, showing that the target is critical for cell survival in multiple cell lines. Note that HPRT1, which was used for co-targeting throughout this study, is not essential. (B) Indel patterns in TMED2 mutants analyzed by TIDE. Note that this target could be highly mutated with GENF with barely detectable wild type alleles (white bar), but with a high enrichment of in-frame mutations (red bars). (C) TIDE analysis of cells transiently transfected with pX330-based plasmids encoding the sgRNAs targeting the indicated genes. Note the very low levels of mutation induced with the sgRNAs targeting ECI2 or HBB. The result of a similar transient transfection to mutate LIG4 is illustrated to compare the efficiency, which shows that a slightly better target sgRNA is sufficient to obtain nearly-complete gene deletion with GENF (see LIG4 mutation rates with GENF in Figure 1J).

Figure 2. Establishment of GENF with another gene-drug pair

(A) Sequence of POR locus and the sgRNAs targeting it with (mm) or without (fm) mismatches at PAM-distal positions. (B) Cells were stained with Crystal Violet after co-targeting POR and CERS6, with or without treatment with paraquat. The colonies of drug-resistant cells are highlighted with circles in additional panels, due to their smaller size than those obtained after HPRT co-targeting. (C) Mutation rates of CERS6 after co-targeting with the indicated sgRNA and selection with paraquat in two independent experiments. (D)

Nomenclature of cell lines obtained during two-step GENF, with the different co-targeted genes color coded. (E) Survival of cells after the second round of two-step GENF (corresponding to the red arrows in (D)), which was analyzed similarly to (B). Note the absence of paraquat-resistant cells when POR was not co-targeted. (F) Mutation rates of CERS genes (see numbers) all targeted together, in the first step (ΔCERS) or the second step (Δ/ΔCERS) of GENF. See the increase in mutation rates for all targets. p value: paired t-test. (G) The metabolic pathway that was manipulated in another experiment for two-step GENF. Sphingosine (So) is generated through degradation of ceramide and is phosphorylated into sphingosine 1-phosphate (So1P) by sphingosine kinases (SPHK1, SPHK2). So1P is dephosphorylated by So1P phosphatases (SGPP1, SGPP2) or degraded by So1P lyase (SGPL1). (H) Combination of genes mutated during two-step GENF, together with the nomenclature of the mutants. (I) Mutation rates in L/P/P mutants generated in the first step illustrated in (H). (J) Mutation rates of genes targeted in the second step illustrated in (H). (K and L) So1P levels in mutants generated in the first (K) and second (L) steps. p values: one-way ANOVA followed by Dunnett's multiple comparison test to compare with controls C1 (K) or C1/C2 (L).

Figure 3. Characterization of mutants generated by GENF

(A and B) Indel patterns seen in 12 mutant cell lines generated with four different sgRNAs, and similarity between them assessed by Pearson's correlation coefficients. (C-E) Analysis of off-target mutations in HBB mutant cells. (C) Off-target candidates of HBB sgRNA reported from *ex cellulo* cleavage experiments, with mismatches highlighted in red [20]. (D) On- and off-target mutations in HBB mutant cells generated with GENF, as compared to those obtained in a published study [20]. Mutations obtained by transient transfection are also shown to demonstrate the enrichment obtained by GENF. (E) On- and off-target mutation rates in HBB mutants compared with published *ex cellulo* cleavage results [20]. Note that mutation rates are well predicted from *ex cellulo* cleavage. (F-K) Comparison of lipid analysis datasets obtained from clones and polyclonal cells generated with GENF. (F and G) Levels of the most abundant phosphatidylcholine (PC) species, PC 34:1, in different control and Agpat2 mutant clones (F) or in different polyclonal cell lines generated with GENF (G). p values: one-way ANOVA followed by Tukey's multiple comparisons test comparing all pairs. (H and I) Principal component analysis of PC acyl-chain composition datasets obtained from clonal cell lines (H) or polyclonal cell lines generated with GENF (I). Encircled are the results of K-means clustering to split datasets into two groups. (J and K) Hierarchical clustering to assess the similarity of the same datasets from clonal cell lines (J) or polyclonal cell lines generated with GENF (K). See also Figure S4.

Figure S4. Additional data to characterize GENF.

(A) Indel patterns of CERS6 mutants generated on different days, shown in different colors. (B) Off-target sites of CERS4 sgRNA predicted based on homology, with mismatches highlighted in red. (C) On- and Off-target mutations observed in CERS4 mutant cells. Non-zero values are plotted in red. (D) Comparison of PC composition between control and Agpat2 mutant clones. Different clones are shown as distinct bars. (E) Comparison of PC composition between polyclonal control and Agpat2 mutant cells generated with GENF. Different polyclonal cell lines are shown as distinct bars. We know empirically that the content of polyunsaturated glycerophospholipids vary largely between datasets obtained on different days, which is likely due to differences in various factors including the conditions of the serum, which is the sole source of polyunsaturated fatty acids in standard culture conditions. The differences in control cells between (D) and (E) are likely to be caused by this variability, since the datasets were obtained years apart with different lots of medium components.

Figure 4. Agpat2 regulates linoleic acid levels in PC

(A) PC acyl-chain composition in control or Agpat2 mutant cells supplemented with linoleic, arachidonic, and docosahexaenoic acid (PUFA supplementation). p values: two-way ANOVA followed by Dunnett's multiple comparisons test to compare with Control. Only reductions seen in both Agpat2 mutants are illustrated for simplicity. (B) Gas chromatography analysis of fatty acid methyl esters obtained from phospholipid fractions of PUFA-supplemented control or Agpat2 mutant cells. p values: two-way ANOVA followed by Dunnett's multiple comparisons test to compare with Control. Common changes are illustrated. (C and D) Lysophosphatidic acid acyltransferase (LPAAT) activity, and its selectivity for linoleoyl-CoA in membrane fractions obtained from control or Agpat2 mutant cells. p values: one-way ANOVA followed by Dunnett's multiple comparisons test to compare with Control. (E) PC acyl-chain composition in PUFA-supplemented control or Agpat2-overexpressing (O/E) cells. p values: two-way ANOVA followed by Sidak's multiple comparisons test. Only increases are illustrated for simplicity. See also Figure S5.

Figure S5. Agpat2 as a regulator of linoleic acid levels.

(A-C) Fragmentation patterns of PC 34:2 analyzed by tandem mass spectrometry. PC 34:2 from control (A) and Agpat2 mutant (B) cells was analyzed by liquid chromatography-tandem mass spectrometry as a bicarbonate adduct negative ion and fragmented to reveal the fatty acyl composition. Fragments corresponding to fatty acids are annotated. Similar results were obtained in three biological replicates and are quantified in (C). (D-G) Selectivity of lysophosphatidic acid acyltransferase (LPAAT) activity in membrane fractions obtained from control or mutant cells for the indicated substrates. (C-G) p values: one-way ANOVA followed by Dunnett's multiple comparisons test to compare with Control. (H) The correlation between LPAAT acyl-CoA selectivity [23] and mRNA expression [45] of various LPAAT enzymes in various tissue. Values above 0.4 are colored.

Figure 5. Lipidome changes induced by mutations in CERSs.

(A and B) Levels of the indicated ceramide (Cer) species in various CERS mutants. (C) Changes in Cer species in CERS2 mutants. (D and E) Changes in the indicated sphingomyelin (SM) species in various CERS mutants. (F) Changes in SM species in CERS2 mutants. (G and H) The competition between CERS2 and CERS5 explains the changes in SL profiles seen in CERS2 mutants. (I and J) Changes in Cer (I) and SM (J) species in Δ/ΔCERS mutants. (K) Changes in total levels of the indicated sphingolipid species in Δ/ΔCERS mutants. p values: (A and D) one-way ANOVA followed by Dunnett's multiple comparisons test to compare with Control, (B and E) Brown-Forsythe and Welch ANOVA tests followed by Dunnett's T3 multiple comparisons test to compare with Control, (C, F, and I-K) two-way ANOVA followed by Sidak's multiple comparisons test. (L-P) Volcano plot illustrating changes in lipid levels and their statistical significances in the indicated CERS mutants. Lipids with changes above 1.2-fold and q values (FDR-corrected multiple t-test) below 0.1 were considered as hits. (M-P) For CERS2 mutants, different lipid classes are highlighted separately. See also Figure S6.

Figure S6. Lipid changes in CERS mutants.

(A) Expression of CERSs in HeLa MZ cells obtained from RNA-seq [29]. (B and C) Changes in the indicated Hexosylceramide (HexCer) species in CERS mutants. (B) p values: one-way ANOVA followed by Dunnett's multiple comparisons test to compare with Control. (D and E) Changes in HexCer species in CERS2 and Δ/ΔCERS mutants. p values: two-way ANOVA followed by Sidak's multiple comparisons test. (F-I) Volcano plots illustrating changes in lipid levels and statistical their significances seen in the indicated CERS mutants. No lipid is identified as a hit (having changes above 1.2-fold and q values (FDR-corrected multiple t-test) below 0.1).

Figure 6. Dissection of SL and ether lipid co-regulation.

(A-D) Changes in the indicated lipid classes in CERS2 mutants. p values: unpaired t-test. (E) Metabolic pathways linking SLs and glycerophospholipids. Red arrows indicate potential metabolic regulations suggested from the whole dataset. (F-I) Changes in the indicated lipid classes in the indicated mutants. q values: one-way ANOVA followed by 5% FDR-corrected multiple t-test. Values below the bars are comparisons with Control, and those above are comparisons between the indicated pairs. (J-M) Changes in non-ether PE and ether PE in the indicated mutants. q values: two-way ANOVA followed by 5% FDR-corrected multiple t-test.

Figure 7. Co-localization as a mechanism of lipid co-regulation.

(A and B) Relative cholesterol changes in CERS2 (A) and Δ/ΔCERS (B) mutants. p values: unpaired t-test. (C) Changes in PS species in CERS2 mutants. p values: two-way ANOVA followed by Sidak's multiple comparisons test. (D-F) Changes in PS 36:1 levels in the indicated mutants. p values: one-way ANOVA followed by Dunnett's multiple comparisons test to compare with Control (D) and unpaired t-test (E); q values: one-way ANOVA followed by 5% FDR-corrected multiple t-test. Values below the bars are comparisons with Control, and those above are comparisons between the indicated pairs. (G and H) Changes in the indicated PS species in CERS2 mutants. p values: unpaired t-test. (I) Proposed model of nanodomains generated by SLs of different N-acyl-chains. (J) Generalized view of lipid co-regulatory mechanisms seen in this study.

ACKNOWLEDGEMENTS

We are grateful to all members of the Riezman laboratory and Shimizu laboratory for valuable comments and continuous support. We especially thank Hideo Shindou for providing support, Noemi Jiménez-Rojo for constructing some of the pX330-based plasmids and sharing preliminary observations, Isabelle Riezman for all technical assistance, and Keisuke Yanagida for critical reading of the manuscript. T.H. was supported by the Japan Society for the Promotion of Science (JSPS) Postdoctoral Fellowships for Research Abroad, the French Government (National Research Agency, ANR) through the "Investments for the Future" programs LABEX SIGNALIFE ANR-11-LABX-0028 and IDEX UCAJedi ANR-15-IDEX-01, and the ATIP-Avenir program (CNRS/Inserm). T.H.-Y. was supported by MEXT/JSPS KAKENHI grant 16K21651. T.S. was supported by the AMED Gapfree Program and Takeda Science Foundation 15668360. H.R. was supported by the Swiss National Science Foundation and the NCCR Chemical Biology (grants 184949 and 185898).

CONFLICTS OF INTEREST

The Department of Lipid Signaling, National Center for Global Health and Medicine, is financially supported by Ono Pharmaceutical Co., Ltd., Japan. The Department of Lipidomics, Graduate School of Medicine, The University of Tokyo, is funded by Shimadzu Corp., Japan.

METHODS

Cell culture

McA-RH7777 cells were obtained from ATCC (CRL-1601) and HeLa MZ cells were kindly provided by Jean Gruenberg (University of Geneva). Both cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM high glucose, with GlutaMAX and pyruvate), supplemented with 10% fetal bovine serum, 100 units/mL penicillin, and 100 µg/mL streptomycin (all from Thermo Fisher Scientific). McA-RH7777 were cultured on collagen-coated dishes (from AGC Techno Glass, or self-prepared using 10 µg/cm^2 of rat type I collagen from Sigma-Aldrich). Passages were performed when cells were semi-confluent, by washing them with Dulbecco's Phosphate-Buffered Saline (DPBS), followed by detachment with Trypsin-EDTA (ethylenediaminetetraacetic acid) (both from Thermo Fisher Scientific). Trypsinization was stopped by adding complete medium, and passages were performed in the range of 1:5 to 1:20 dilution. Both cell lines were tested free of mycoplasma contamination by Eurofins Genomics.

Plasmid construction

The plasmids pX330 and pX459 for expression of *Streptococcus pyogenes* Cas9 and a single guide RNA were deposited to Addgene by Feng Zhang (plasmids #42230 and #48139). Target sequences were designed using the CRISPR guide RNA design tools from Benchling or CRISPOR [46]. Selection was based on high on-target scores and low off-target scores [2], while avoiding targets upstream of potential alternate start codons and considering alternative splicing, using the information available in the UCSC genome browser. All the target sequences used in this study are available in supplemental table 1. Inserts were designed by adding appropriate overhangs (5'-CACC-3' and 5'-AAAC-3' for sense and antisense strands, respectively) on the target sequence. When the target sequence did not start with G, an additional G was added in the sense strand for the functionality of the U6 promoter, and a C was added at the end of the antisense strand make it complementary. Inserts were prepared by annealing 5 µM each of oligonucleotides (synthesized by Microsynth AG) in annealing buffer (40 mM Tris-HCl pH 8.0, 20 mM $MgCl_2$, and 50 mM NaCl), by heating them at 95ºC for 5 minutes and cooling them down slowly by reducing temperature at a speed of -0.8ºC/minute in a thermal cycler. Annealed oligonucleotides (1.5 µL) were inserted by golden gate assembly into 22.5 ng of pX330 or pX459 using 0.3 µL FastDigest Bpi I (Thermo Fisher Scientific) and 0.3 µL quick ligase in 1 x quick ligase buffer (New England Biolabs) with a total volume of 6 µL. Reactions consisted of 3 cycles of incubations at 5 minutes at 37ºC followed by 5 minutes at 25ºC. This was followed by the addition of 0.3 µL FastDigest BpiI and 0.6 µL of water and incubation at 37ºC for 1 hour to remove undigested plasmids. The plasmids pHPRTsg for expression of a single guide RNA against the rat Hprt1 or human HPRT1 gene were constructed similarly, using a previously generated pUC-U6-sg empty vector (9 ng per reaction) designed for expression of a single guide RNA under the same promoter as pX330 and pX459 [13]. Plasmids were transformed into Stbl3 chemically competent Escherichia coli (Thermo Fisher Scientific) and colonies containing correct plasmids were selected by colony PCR followed by Sanger sequencing (at Fasteris) of the PCR products (all primers used in this study are available in supplemental table 2). Plasmids were purified using the PureLink HiPure Plasmid Midiprep Kit (Thermo Fisher Scientific) or GenElute plasmid mini kit (Sigma-Aldrich). When the latter was used, plasmids were further purified by Triton X-114 isothermal extraction to obtain transfection-grade ones [47]. Plasmids were incubated in 1% (w/v) Triton X-114 and 0.5% (w/v) sodium dodecyl sulfate (SDS) for 10 minutes at room temperature, then endotoxins were precipitated by adding sodium chloride to a final concentration of 1 M and centrifuging at 12,000 x *g* for 10 minutes at room temperature. Purified plasmids were obtained from the supernatants by isopropanol precipitation, followed by two washing steps with 70% ethanol. Plasmids were dissolved in TE buffer (10 mM Tris-HCl pH 8.0 and 1 mM EDTA).

Generation and DNA analysis of mutant cell clones

To generate Sgpl1/Sgpp1/Sgpp2 mutant clones, three pX330-based plasmids encoding sgRNAs targeting each gene were transfected into McA-RH7777 cells using lipofectamine 3000 (Thermo Fisher Scientific). Transfected cells were seeded sparsely in 10 cm dishes to obtain colonies. When colonies grew enough to extract sufficient genome DNA for PCR, well-isolated colonies were directly lysed in lysis buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.67% SDS, and 125 µg/mL proteinase K) using cloning rings prepared by cutting pipette tips. Lysates were incubated at 55ºC for 4 hours and genomic DNA was obtained by isopropanol precipitation. Mutant cells were detected by competition-based PCR using ExTaq (Takara Bio), which consists of a PCR procedure using three primers; one primer overlaps with the targeted region, while the two others flank it [13]. Changed ratios of the amplicons was suggestive of mutations, which were further confirmed by Sanger sequencing of PCR amplicons treated with Exonuclease I and FastAP alkaline phosphatase (both from Thermo Fisher Scientific). To generate Agpat2 mutant clones, a pair of pX459-based plasmids was transfected into McA-RH7777 cells using lipofectamine 3000, followed by selection with puromycin (InvivoGen) for 24 hours starting the next day. Clones were obtained by limiting dilution in a 96 well plate. Mutant clones were selected by PCR, analyzing the deletions in genomic DNA generated by the pair of cuts.

Generation of polyclonal mutants

McA-RH7777 cells were co-transfected with pX330-based plasmids and pHPRTsg using lipofectamine 3000. The amount of pHPRTsg used was 10% (w/w) unless stated otherwise. Seven days after transfection, 6-

thioguanine (Sigma-Aldrich) was used at 4 μg/mL to select resistant cells lacking Hprt1 for 1 week. Mutant HeLa MZ cells were generated similarly, using 1% of pHPRTsg and starting drug selection at 6 μg/mL 5 days after transfection. When POR gene was co-targeted instead of HPRT1, plasmids based on pUC-U6-sg were used to express the sgRNA targeting POR (1% of total transfected DNA). Cells were selected with 200 μM paraquat (Sigma-Aldrich) 5 days after transfection for 4 days and replated to new dishes. The efficiency of gene mutations and the indel patterns were analyzed by isolating genomic DNA as stated above, Sanger sequencing target regions amplified by PCR (supplemental tables 1 and 2), and analyzing them with TIDE (tracking indels by decomposition) [15]. When comparing the yields of resistant cells, a similar procedure was performed, and cells were fixed with methanol. Fixed cells were stained with 0.5% (w/v) crystal violet prepared in 20% methanol.

Analysis of C18 So1P

Cells were seeded at 750,000 cells/dish in 6 cm dishes. 24 hours later, cells were scraped in ice-cold PBS, centrifuged at 2,500 rpm, and pellets were snap frozen in liquid nitrogen. Cells were resuspended in 150 μL of pyridine extraction buffer (ethanol : water : diethyl ether : pyridine = 15 : 15 : 5 : 1, supplemented with 2.1 mM ammonium hydroxide), to which internal standards (40 pmol C17 sphingosine, 40 pmol of C17 sphinganine, 400 pmol C17 sphingosine-1-phosphate, and 400 pmol C17 sphinganine-1-phosphate) were added. Samples were vortexed for 10 minutes at 4ºC, then incubated at 4ºC for 20 minutes. Samples were centrifuged at 14,000 rpm for 2 minutes at 4ºC and supernatants were collected. An additional extraction was performed from remaining pellets using the same steps except for the lack of incubation. The supernatants were combined, split into two, and dried under a stream of nitrogen. Sphingoid bases were derivatized with AQC (6-aminoquinolyl-N-hydroxysuccinimidyl carbamate) by resuspending the extracts in 10 μL of 0.1% formic acid, to which was added 70 μL of borate buffer (200 mM boric acid pH 8.8, 10 mM tris(2-carboxyethyl)-phosphine, 10 mM ascorbic acid, and 33.7 mM $^{15}N^{13}C$-valine; the heavy isotope-labeled valine was included for potential analysis of amino acids, not performed in this study) and 20 μL of 2.85 mg/mL AQC in acetonitrile. Derivatization was performed for 15 minutes at 55ºC and overnight at 24ºC. Samples were centrifuged at 14,000 rpm for 2 minutes, and supernatants were analyzed by LC-MS/MS in an Accela HPLC system coupled to a TSQ Vantage mass spectrometer (Thermo Fisher Scientific). Samples were separated on a C18 column (EC 100/2 Nucleoshell RP-18, 2.7 μm, from Macherey-Nagel) using a gradient of 0.1% formic acid / isopropanol over 0.1% formic acid / water. Multiple reaction monitoring was used to detect AQC-derivatized So1P at Q1 (*m/z* = 536.3 for C17 So1P and 550.3 for C18 So1P) and the AQC fragment at Q3 (*m/z* = 171.1). Concentrations were calculated by normalizing signals for extraction efficiency using signals of the C17 internal standard, which was further corrected for signal responses that were measured by injecting synthetic standards in another run. To normalize variations in cell numbers between samples, pellets generated during sample extractions were washed with ethanol, dried for 5 minutes at room temperature, and lysed in protein lysis buffer (100 mM Tris-HCl pH 8.0, 3 M Urea, and 1% SDS). Samples were incubated at 60ºC for 15 minutes and sonicated for 5 minutes in a bath sonicator. The incubation-sonication cycle was repeated once more. Protein concentrations were analyzed with a Bicinchoninic Acid Protein Assay Kit (Sigma-Aldrich), using bovine serum albumin as a standard, which were used to normalize variations in cell numbers.

Analysis of phosphatidylcholine (PC) by LC-MS/MS

McA-RH7777 cells were seeded on collagen-coated 96 well plates at 10,000 cells/well. When analyzing polyunsaturated fatty acid (PUFA)-supplemented cells, 10 μM each of linoleic acid, arachidonic acid, and docosahexaenoic acid was added the next day. 24 hours later, cells were washed with PBS and methanol was added to extract PC and incubated at 4ºC for >1 hour. Samples were centrifuged at 13,000 rpm for 5 minutes to remove debris, and the supernatants were analyzed by LC-MS/MS in an ACQUITY UPLC system coupled to a TSQ Vantage mass spectrometer. Samples were separated on a C8 column (ACQUITY UPLC BEH C8, 1 x 100 mm, 1.7 μm, Waters) using a gradient of acetonitrile over 20 mM ammonium bicarbonate. We noted severe peak tailing of some PC species depending on the condition of the LC-MS/MS system, and in such cases acetonitrile was replaced by isopropanol : acetonitrile = 3 : 7. PC was detected by a precursor ion scan detecting the phosphocholine fragment of m/z = 184.1 at Q3. Since the biological question was to analyze the acyl-chain balance of PC, the sum of PC signals was used for normalization.

Analysis of fatty acids by GC

McA-RH7777 were seeded on 12 well collagen-coated plates at 120,000 cells/well and PUFAs were supplemented the next day as described above. 24 hours later, cells were washed with PBS, and scraped in methanol. Methanol was further added to the wells to maximize recovery and combined with the scraped cells. Lipids were extracted by the method of Bligh and Dyer, the extracts were resuspended in chloroform, and were loaded on InertSep NH2 aminopropyl columns (GL Science). Neutral lipids were eluted with chloroform : isopropanol = 2 : 1, free fatty acids were eluted with 2% acetic acid in diethyl ether, and phospholipid were collected by eluting them in 2.8% ammonia in methanol. C23:0 (Supelco n-Tricosanic acid) was added as a standard to the phospholipid fractions. Fatty acid methyl esters were prepared with the Fatty Acid Methylation Kit, and purified using the Fatty Acid Methyl Ester Purification Kit (both from Nacalai Tesque). Fatty acid methyl esters were separated on a capillary column (FAMEWAX, 30 m, 0.25 mm ID, 0.25 μm from Restek) using a GC-2010 Plus system equipped with a flame ionization detector

(Shimadzu). Sums of abundant fatty acids were used to normalize the data and calculate the percentages of individual fatty acids. See reference for detailed procedures [48].

Lipidome analysis by nanoESI-MS

Cells were seeded on 6 cm dishes at 750,000 cells/dish. 24 hours later, cells were scraped in ice-cold PBS, pelleted down by centrifugation at 2,500 rpm at 4ºC, and pellets were snap frozen in liquid nitrogen. Lipids were extracted using a modified methyl tert-butyl ether (MTBE) extraction protocol. Pellets were resuspended in 200 µL 50% methanol/water, 100 µL of an internal standard mix (400 pmol dilauryl-PC, 1 nmol 17:0-14:1 phosphatidylethanolamine, 1 nmol 17:0-14:1 phosphatidylinositol, 3.3 nmol 17:0-14:1 phosphatidylserine, 700 pmol tetramyristoyl cardiolipin, 2.5 nmol C17 ceramide, 500 pmol C8 glucosylceramide, and 100 pmol C12 sphingomyelin, resuspended in chloroform : methanol = 1 : 1 and diluted in methanol, all from Avanti Polar Lipids) was added, followed by 180 µL methanol and 1.2 mL MTBE. Samples were vortexed for 10 minutes at 4ºC and incubated at room temperature for 1 hour. 200 µL of water was added to induce phase separation. After 10 minutes of incubation, samples were centrifuged at 10,000 x *g* for 10 minutes at room temperature. The upper layers were transferred to glass tubes, and the remaining samples were used for a re-extraction with 400 µL of artificial upper layer (MTBE : methanol : water = 10 : 3 : 1.5), which was incubated and centrifuged as described above. Upper layers were combined (designated total lipid fractions), split into three vials, and dried under a nitrogen stream. Among the split samples, one was further treated by a mild base to remove glycerophospholipids, which was critical to reduce ion suppression to hexosylceramide species and to detect sphingomyelin accurately without interference from isotopes of PC. For this, dried samples were reconstituted in 500 µL of methylamine solvent (methanol : water : water-saturated butanol : methylamine = 4 : 3 : 1 : 5), sonicated for 5 minutes in a bath sonicator, and incubated for 1 hour at 53ºC. Samples were dried under a nitrogen flux, followed by desalting. For desalting, dried samples were reconstituted in 300 µL water-saturated butanol, sonicated for 5 minutes in a bath sonicator, and 150 µL water was added to induce phase separation. Samples were centrifuged at 3200 x *g* for 10 minutes at room temperature, and upper layers were transferred to glass vials. Another extraction was repeated by adding 300 µL water-saturated butanol to the remaining samples and combined with the first extracts. Samples were dried under a nitrogen flux and were designated as sphingolipid fractions. Lipids from each fraction were reconstituted in 250 µL chloroform : methanol = 1 : 1 and diluted 4 and 10 fold for samples to be measured in the negative and positive ion mode, respectively. Dilutions were done in chloroform : methanol : water = 2 : 7 : 1 for samples analyzed in the positive ion mode and chloroform : methanol = 1 : 2 for samples analyzed in the negative ion mode, both being supplemented with 5 mM ammonium acetate. Diluted samples were directly infused in a TSQ Vantage mass spectrometer using a robotic sample delivery system equipped with a nanoESI source (TriVersa NanoMate from Advion). Lipids were detected by multiple reaction monitoring, using fragments specific for each lipid classes. A full list of transitions is available in supplemental table 3. Individual lipids were quantified by dividing signal intensities with those of corresponding internal standards, multiplying them with the amount of internal standard added. Then sums of each lipid class were calculated after removing low abundance lipids. For this filtering, the averages of the five lowest values for each lipid class were used to estimate the noise levels, and only signals at least two-fold higher than this noise were considered. Due to the low numbers of sphingolipids and sterols analyzed, this filtering was performed only for glycerophospholipids. The sums of all major lipids were calculated from phosphatidylcholine, ether phosphatidylcholine, phosphatidylethanolamine, ether phosphatidylethanolamine, phosphatidylinositol, phosphatidylserine, cardiolipin, ceramide, hexosylceramide, and sphingomyelin. Lysophospholipids and sterols were not used to calculate the sums. The sums were used to normalize data variances coming from the small differences in sample amounts used for lipid extraction, and thus lipid data are illustrated as "% of total". This normalization was used empirically as an efficient way to generate consistent datasets between replicates, making the detection of differentially abundant lipids more robust. The characteristics and limitations of different mass spectrometry approaches are discussed in supplemental note.

Measurement of lysophosphatidic acid acyltransferase (LPAAT) activity

Cells were disrupted on ice using a probe sonicator (Ohtake Works) through 2 rounds of sonications for 30 seconds each in TSC buffer (100 mM Tris-HCl pH 7.4, 300 mM sucrose, and 1 x complete protease inhibitor cocktail). Sonicated samples were centrifuged at 800 x *g* for 10 minutes, and supernatants were centrifuged at 100,000 x *g* for 1 hour to pellet membrane fractions. Pellets were resuspended in TSE buffer (20 mM Tris-HCl pH 7.4, 300 mM sucrose, and 1 mM EDTA) and protein concentrations were measured using the Protein Assay Kit (Bio-Rad). LPAAT activity was measured by incubating membrane fractions (containing 0.01 µg of protein) with substrates at 37ºC for 10 minutes. Reaction mixtures contained 25 µM deuterium-labeled lysophosphatidic acid, 1 µM each of 16:0-, 18:1-, 18:2-, 20:4-, and 22:6-CoA (Avanti Polar Lipids), 110 mM Tris-HCl pH 7.4, 150 mM sucrose, and 0.5 mM EDTA. Reactions were stopped by the addition of chloroform : methanol = 1 : 2. Dimyristoyl phosphatidic acid (Avanti Polar Lipids) was added as an internal standard, and Bligh and Dyer extraction was performed. Reaction products were quantified using LC-MS/MS by comparing peak areas of chromatograms to those of lipid standards. The system was the same as the one used for PC analysis, but the column was an ACQUITY UPLC BEH amide column (2.1 x 30 mm, 1.7 µm, from Waters). See reference for detailed procedures [23].

Informatics and Data analysis

Gene essentiality across multiple cell lines was analyzed using the depmap portal website, where data of genome-wide CRISPR screenings in 789 cell lines is available [44]. Raw data were downloaded from the website and values were extracted with R software. Gene expression profiles of Agpat genes across multiple mouse tissues were obtained from BioGPS [45]. LPAAT activity in multiple tissues was analyzed previously [23]. The correlation between datasets was calculated from the data of lung, heart, liver, spleen, kidney, stomach, small intestine, colon, testis, adipose tissue, and skeletal muscle, which are the ones overlapping between the two. Statistical analyses were performed using Prism 8 (GraphPad). Multivariate analyses (principal component analysis, K-means clustering, and hierarchical clustering) were performed using R.

SUPPLEMENTAL NOTE

In this study two approaches were used for lipid analysis. One involved the use of liquid chromatography prior to mass spectrometry (LC-MS/MS) and the other used direct infusion into the nano-electrospray ionization (nanoESI) source of the mass spectrometer (nanoESI-MS/MS). Here we discuss the characteristics and the limitations of the approaches. LC-MS/MS enables the separation of lipids before analysis, which is especially beneficial when discussing the differences in acyl-chains. This justifies its use for PC acyl-chains. Without separation with LC, signals of a certain lipid are “contaminated” by isotopic signals from a lipid with one additional double bond (e.g. signals of PC 34:1 are affected by PC 34:2 with two ^{13}C). On the other hand, nanoESI-MS/MS is useful for comprehensive analyses. In this approach, all the lipids in one sample are continuously ionized and detectable, making it possible to analyze them as long as the spray lasts. Since the analysis time becomes longer when analyzing more lipids, there is a compromise between the number of lipids analyzed and the number of data points per peak that we can acquire in LC-MS/MS, while this limitation is virtually absent in nanoESI-MS/MS. This is why the latter was used in this study for comprehensive lipidomics. In contrast to LC-MS/MS where signals are detected as peaks with retention times, only signal intensities are obtained with nanoESI-MS/MS. This characteristic of nanoESI-MS/MS can lead to false identification of lipids, especially for minor ones that are difficult to distinguish from noise signals. Due to this limitation, we filtered data to remove minor species when calculating sums of lipid classes, as explained in the methods section. Manual inspection of the data obtained from the nanoESI-MS/MS approach confirmed the consistency with expectable major lipids (e.g. major glycerophospholipids expected to have mainly C16 or C18 fatty acids at the sn-1 position and up to C22 species at the sn-2 position, mainly C16 and C24 N-acyl-chains for sphingolipids and only minor C18 and C20 species), thus justifying its use to cover a broad range of lipids, especially after filtering.

One common limitation of both approaches, which is also true for most of the published lipidomics literature, is the differences in signal responses between lipids. Lipids with different structures (head group and acyl-chains), even within the same lipid class, give different signal responses when injected at equimolar amounts into the mass spectrometer. Therefore, unless synthetic standards for all the hundreds of analyzed lipids are prepared, all the values remain good approximations and not perfect quantifications. In this study, differences in signal responses of lipids from distinct classes are normalized by the use of internal standards. Since the differences in signal responses due to different acyl-chains persist, the values reported here do not reflect amounts in mols, but they give good approximations since the differences caused by acyl-chains are smaller than those caused by lipid classes. Due to these limitations, the values reported in this study cannot be assumed to represent exact mol percentages, but they remain extremely useful for relative comparison between samples (e.g. control versus mutant) as has been done throughout the study. This limitation is especially pronounced for cholesterol. It is known that cholesterol is not efficiently ionized in ESI, and we indeed obtained weak and variable signals. For this reason, the values normalized with the total of lipid signals (see methods) had large inter-day differences and were clearly diverging from expectable mol percentages. However, they could be used for relative comparison between samples analyzed the same day, and in this study, we use control cell values as 1. Note that the results in cholesterol reported in this study (no change in CERS2 mutants and a decrease in Δ/ΔCERS mutants) were consistent in two independent triplicate experiments, thus are reproducible.

REFERENCES

1. Doudna, J.A., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.
2. Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nature Biotechnology *34*, 184-191.
3. Zuris, J.A., Thompson, D.B., Shu, Y., Guilinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.-Y., and Liu, D.R. (2014). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nature Biotechnology *33*, 73-80.
4. Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Research *24*, 1012-1019.
5. Ben-David, U., Siranosian, B., Ha, G., Tang, H., Oren, Y., Hinohara, K., Strathdee, C.A., Dempster, J., Lyons, N.J., Burns, R., et al. (2018). Genetic and transcriptional evolution alters cancer cell line drug response. Nature *560*, 325-330.
6. Dyikanov, D.T., Vasiluev, P.A., Rysenkova, K.D., Aleksandrushkina, N.A., Tyurin-Kuzmin, P.A., Kulebyakin, K.Y., Rubtsov, Y.P., Shmakova, A.A., Evseeva, M.N., Balatskiy, A.V., et al. (2019). Optimization of CRISPR/Cas9 Technology to Knock Out Genes of Interest in Aneuploid Cell Lines. Tissue Engineering Part C: Methods *25*, 168-175.
7. Shevchenko, A., and Simons, K. (2010). Lipidomics: coming to grips with lipid diversity. Nature Reviews Molecular Cell Biology *11*, 593-598.
8. Loizides-Mangold, U. (2013). On the future of mass-spectrometry-based lipidomics. FEBS Journal *280*, 2817-2829.
9. Harayama, T., and Riezman, H. (2018). Understanding the diversity of membrane lipid composition. Nature Reviews Molecular Cell Biology *19*, 281-296.
10. Rodriguez-Cuenca, S., Pellegrinelli, V., Campbell, M., Oresic, M., and Vidal-Puig, A. (2017). Sphingolipids and glycerophospholipids – The “ying and yang” of lipotoxicity in metabolic diseases. Progress in Lipid Research *66*, 14-29.
11. Köberlin, Marielle S., Snijder, B., Heinz, Leonhard X., Baumann, Christoph L., Fauster, A., Vladimer, Gregory I., Gavin, A.-C., and Superti-Furga, G. (2015). A Conserved Circular Network of Coregulated Lipids Modulates Innate Immune Responses. Cell *162*, 170-183.
12. Jiménez-Rojo, N., Leonetti, M.D., Zoni, V., Colom, A., Feng, S., Iyengar, N.R., Matile, S., Roux, A., Vanni, S., Weissman, J.S., et al. (2020). Conserved Functions of Ether Lipids and Sphingolipids in

the Early Secretory Pathway. Current Biology *30*, P3775-3787.e3777.

13. Harayama, T., and Riezman, H. (2017). Detection of genome-edited mutant clones by a simple competition-based PCR method. Plos One *12*, e0179165.
14. Liao, S., Tammaro, M., and Yan, H. (2015). Enriching CRISPR-Cas9 targeted cells by co-targeting the HPRT gene. Nucleic Acids Research *43*, gkv675.
15. Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Research *42*, e168.
16. Reczek, C.R., Birsoy, K., Kong, H., Martínez-Reyes, I., Wang, T., Gao, P., Sabatini, D.M., and Chandel, N.S. (2017). A CRISPR screen identifies a pathway required for paraquat-induced cell death. Nature Chemical Biology *13*, 1274-1279.
17. Kihara, A. (2014). Sphingosine 1-phosphate is a key metabolite linking sphingolipids to glycerophospholipids. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids *1841*, 766-772.
18. Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. Nature Biotechnology *36*, 765-771.
19. Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H.R., Hwang, J., Kim, J.-I., and Kim, J.-S. (2015). Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nature Methods *12*, 237-243.
20. Kim, D., and Kim, J.-S. (2018). DIG-seq: a genome-wide CRISPR off-target profiling method using chromatin DNA. Genome Research *28*, 1894-1900.
21. Broeders, M., Herrero-Hernandez, P., Ernst, M.P.T., van der Ploeg, A.T., and Pijnappel, W.W.M.P. (2020). Sharpening the Molecular Scissors: Advances in Gene-Editing Technology. iScience *23*, 100789.
22. Hashidate-Yoshida, T., Harayama, T., Hishikawa, D., Morimoto, R., Hamano, F., Tokuoka, S.M., Eto, M., Tamura-Nakano, M., Yanobu-Takanashi, R., Mukumoto, Y., et al. (2015). Fatty acid remodeling by LPCAT3 enriches arachidonate in phospholipid membranes and regulates triglyceride transport. eLife *4*, e06328.
23. Harayama, T., Eto, M., Shindou, H., Kita, Y., Otsubo, E., Hishikawa, D., Ishii, S., Sakimura, K., Mishina, M., and Shimizu, T. (2014). Lysophospholipid Acyltransferases Mediate Phosphatidylcholine Diversification to Achieve the Physical Properties Required In Vivo. Cell Metabolism *20*, 295-305.
24. Ichi, I., Kono, N., Arita, Y., Haga, S., Arisawa, K., Yamano, M., Nagase, M., Fujiwara, Y., and Arai,

H. (2014). Identification of genes and pathways involved in the synthesis of Mead acid (20:3n−9), an indicator of essential fatty acid deficiency. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids *1841*, 204-213.

25. Symons, J.L., Cho, K.-J., Chang, J.T., Du, G., Waxham, M.N., Hancock, J.F., Levental, I., and Levental, K.R. (2020). Lipidomic atlas of mammalian cell membranes reveals hierarchical variation induced by culture conditions, subcellular membranes, and cell lineages. Soft Matter *Advance Article*.
26. Agarwal, A.K., Arioglu, E., de Almeida, S., Akkoc, N., Taylor, S.I., Bowcock, A.M., Barnes, R.I., and Garg, A. (2002). AGPAT2 is mutated in congenital generalized lipodystrophy linked to chromosome 9q34. Nature Genetics *31*, 21-23.
27. Levental, K.R., Surma, M.A., Skinkle, A.D., Lorent, J.H., Zhou, Y., Klose, C., Chang, J.T., Hancock, J.F., and Levental, I. (2017). ω-3 polyunsaturated fatty acids direct differentiation of the membrane phenotype in mesenchymal stem cells to potentiate osteogenesis. Science Advances *3*, eaao1193.
28. Wegner, M.-S., Schiffmann, S., Parnham, M.J., Geisslinger, G., and Grösch, S. (2016). The enigma of ceramide synthase regulation in mammalian cells. Progress in Lipid Research *63*, 93-119.
29. Scott, C.C., Vossio, S., Rougemont, J., and Gruenberg, J. (2018). TFAP2 transcription factors are regulators of lipid droplet biogenesis. eLife *7*, e36330.
30. Hanada, K., Hara, T., Fukasawa, M., Yamaji, A., Umeda, M., and Nishijima, M. (1998). Mammalian Cell Mutants Resistant to a Sphingomyelin-directed Cytolysin. Journal of Biological Chemistry *273*, 33787-33794.
31. Okino, N., He, X., Gatt, S., Sandhoff, K., Ito, M., and Schuchman, E.H. (2003). The Reverse Activity of Human Acid Ceramidase. Journal of Biological Chemistry *278*, 29948-29953.
32. Vance, J.E., and Tasseva, G. (2013). Formation and function of phosphatidylserine and phosphatidylethanolamine in mammalian cells. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids *1831*, 543-554.
33. Horibata, Y., Ando, H., and Sugimoto, H. (2020). Locations and contributions of the phosphotransferases EPT1 and CEPT1 to the biosynthesis of ethanolamine phospholipids. Journal of Lipid Research *61*, 1221-1231.
34. Horibata, Y., Elpeleg, O., Eran, A., Hirabayashi, Y., Savitzki, D., Tal, G., Mandel, H., and Sugimoto, H. (2018). EPT1 (selenoprotein I) is critical for the neural development and maintenance of plasmalogen in humans. Journal of Lipid Research *59*, 1015-1026.
35. Das, A., Brown, M.S., Anderson, D.D., Goldstein, J.L., and Radhakrishnan, A. (2014). Three pools of plasma membrane cholesterol and their relation to cholesterol homeostasis. eLife *3*, e02882.

36. Sharma, P., Varma, R., Sarasij, R.C., Ira, Gousset, K., Krishnamoorthy, G., Rao, M., and Mayor, S. (2004). Nanoscale Organization of Multiple GPI-Anchored Proteins in Living Cell Membranes. Cell *116*, 577-589.

37. Róg, T., Orłowski, A., Llorente, A., Skotland, T., Sylvänne, T., Kauhanen, D., Ekroos, K., Sandvig, K., and Vattulainen, I. (2016). Interdigitation of long-chain sphingomyelin induces coupling of membrane leaflets in a cholesterol dependent manner. Biochimica et Biophysica Acta (BBA) - Biomembranes *1858*, 281-288.

38. Skotland, T., and Sandvig, K. (2019). The role of PS 18:0/18:1 in membrane function. Nature Communications *10*, 2752.

39. Simons, K., and Gerl, M.J. (2010). Revitalizing membrane rafts: new tools and insights. Nature Reviews Molecular Cell Biology *11*, 688-699.

40. Duda, K., Lonowski, L.A., Kofoed-Nielsen, M., Ibarra, A., Delay, C.M., Kang, Q., Yang, Z., Pruett-Miller, S.M., Bennett, E.P., Wandall, H.H., et al. (2014). High-efficiency genome editing via 2A-coupled co-expression of fluorescent proteins and zinc finger nucleases or CRISPR/Cas9 nickase pairs. Nucleic Acids Research *42*, e84-e84.

41. Jain, I.H., Calvo, S.E., Markhard, A.L., Skinner, O.S., To, T.-L., Ast, T., and Mootha, V.K. (2020). Genetic Screen for Cell Fitness in High or Low Oxygen Highlights Mitochondrial and Lipid Metabolism. Cell *181*, 716-727.e711.

42. He, B., Jin, J., Xu, Y., Huo, L., Ma, L., Scott, A.W., Pizzi, M.P., Li, Y., Wang, Y., Yao, X., et al. (2020). An improved strategy for CRISPR/Cas9 gene knockout and subsequent wildtype and mutant gene rescue. Plos One *15*, e0228910.

43. Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O.O., Gootenberg, J.S., Mori, H., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science *361*, 1259-1262.

44. Tsherniak, A., Vazquez, F., Montgomery, P.G., Weir, B.A., Kryukov, G., Cowley, G.S., Gill, S., Harrington, W.F., Pantel, S., Krill-Burger, J.M., et al. (2017). Defining a Cancer Dependency Map. Cell *170*, 564-576.e516.

45. Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C.L., Haase, J., Janes, J., Huss, J.W., et al. (2009). BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biology *10*, R130.

46. Haeussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J.-B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J., et al. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome

Biology *17*, 148.

47. Ma, R., Zhao, J., Du, H.-C., Tian, S., and Li, L.-W. (2012). Removing endotoxin from plasmid samples by Triton X-114 isothermal extraction. Analytical Biochemistry *424*, 124-126.

48. Hamano, F., Tokuoka, S., Hashidate-Yoshida, T., Shindou, H., Shimizu, T., and Kita, Y. (2020). Quantification of Fatty Acids in Mammalian Tissues by Gas Chromatography–Hydrogen Flame Ionization Detection. Bio-Protocol *10*, e3613.

Figure 1 (Harayama et al., 2020)

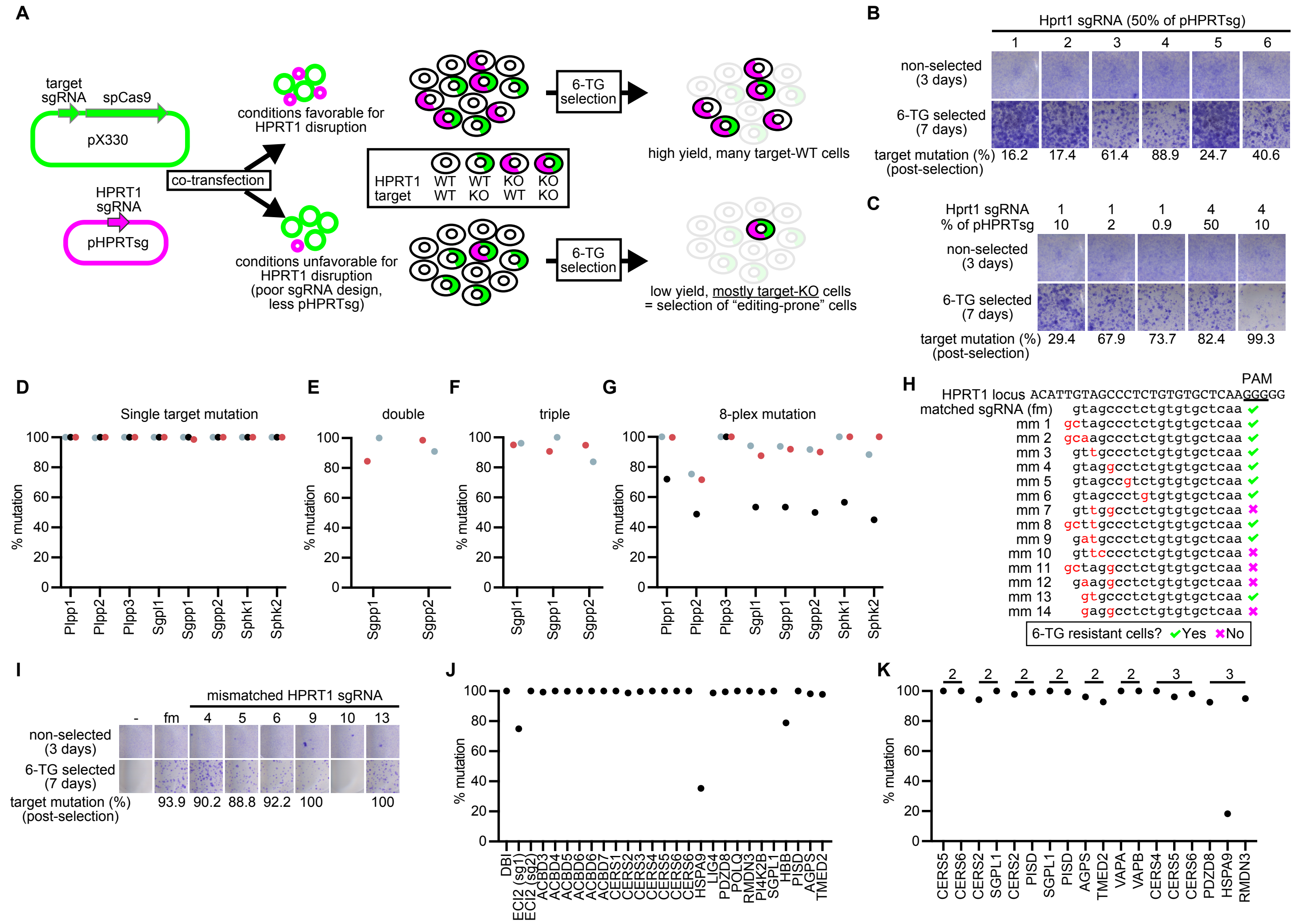


Figure 1. Establishment of GENF

(A) Design of GENF and the plasmids (pX330 and pHPRTsg) used for it. HPRT co-targeting experiments performed under standard conditions (top) or conditions where Hprt1 disruption is unfavorable (bottom, corresponding to GENF) are illustrated. (B and C) Co-targeting of Sphk2 and Hprt1 in McA-RH7777 cells using different sgRNAs against Hprt1 at various percentage of pHPRTsg plasmid. Cells were stained with Crystal Violet after culture for the indicated time with or without 6-TG selection. The non-selected cells were prepared to ensure that similar cell numbers were seeded. The percentage of Sphk2 mutation (target mutation % post-selection) is shown under the bottom panel. (D-G) Mutation rates after GENF for 8 genes targeted either alone (D), in duplex (E), in triplex (F), or in 8-plex (G). Experiments performed on different days are distinguished by colors. (H) Design of the sgRNA against HPRT1 (sgRNA with a full match, fm) or its 14 variants (mm 1-14) with mismatches (in red) used to establish GENF in HeLa MZ cells. The sgRNAs that enabled the generation of 6-TG resistant cells are labeled with a tick and the others with a cross. PAM: protospacer adjacent motif. (I) Co-targeting of CERS6 and HPRT1 in HeLa MZ cells using representative HPRT1 sgRNAs with (numbers correspond to mm# in (H)) or without (fm) mismatches. 1% (w/w) of pHPRTsg was used. Panels are organized similarly to (B) and (C). (J and K) Mutation rates after GENF for the indicated genes targeted in HeLa MZ cells either alone (J) or in multiplex (K). (K) Targets under the same lines were mutated together. See also Figures S1-S3.

Figure S1 (Harayama et al., 2020)

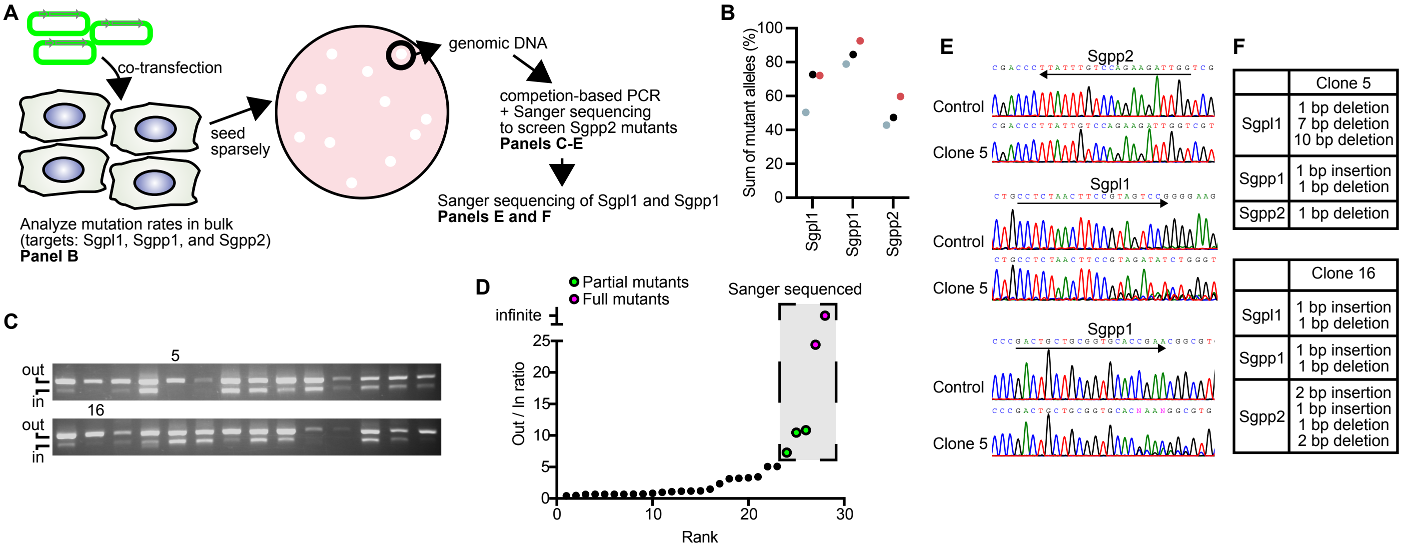


Figure S1. Preliminary experiment leading to the idea of GENF.

(A) Overview of the experiment. McA-RH7777 cells were co-transfected with three pX330-based plasmids encoding Cas9 and sgRNAs against Sgpl1, Sgpp1, and Sgpp2. Transfected cells were analyzed in bulk to test the efficiencies of different sgRNAs, while one part was reseeded in dishes sparsely to obtain colonies. Colonies were directly lysed to obtain genomic DNA and screen for Sgpp2 mutants. Sgpl1 and Sgpp1 mutation was analyzed in the obtained Sgpp2 mutants. (B) In three experiments (distinguished by colors), mutation rates of Sgpp2 were always the lowest after transiently co-transfecting pX330-based plasmids. (C and D) Screening of Sgpp2 mutants by competition-based PCR. The method uses a mixture of primer that enables the detection of mutant clones by a reduction of the amplicon labeled “in” compared to the one labeled “out” , since one of the primers for “in” amplicon binds to the sgRNA target. See reference [13] for details. (C) Result of competition-based PCR. PCR patterns in the numbered lanes are from the clones that were later confirmed to be fully mutated. (D) Quantification of signal ratios between the two amplicons in (C). Five clones were selected for Sanger sequencing, and two of them were full mutants. (E) Sanger sequencing of the three targets in parental cells or clone 5. Target regions are underlined with an arrow. (F) Summary of mutations seen in clones 5 and 16. Note the absence of wild type alleles in Sgpl1 and Sgpp1, despite the fact that the clones were pre-selected based only on Sgpp2 mutations.

Figure S2 (Harayama et al., 2020)

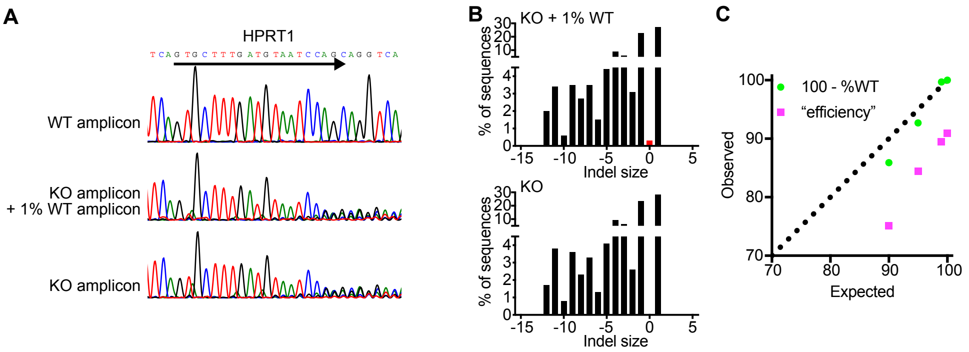


Figure S2. Evaluation of a method to analyze high mutation rates.
(A) Sanger sequencing of PCR amplicons from wild type or mutant HPRT1 locus, either alone or mixed at the indicated ratio. (B) TIDE (tracking indels by decomposition) analysis [15] from the Sanger sequencing results of (A). The calculated frequency of alleles with the indicated indel size is illustrated. Note that as little as 1% of wild type allele can be detected (with an indel size of zero, red bar in top panel) with TIDE. (C) Comparison of expected mutation rates (from the ratio of mixed PCR amplicons) and observed mutation rates calculated in two ways. Observed mutation rates were calculated either from the sum of mutant alleles detected with TIDE ("efficiency" , which is usually used) or by subtracting the detected wild type alleles from 100 ("100-%WT"). Note that the latter gave a better estimate of mutation rates.

Figure S3 (Harayama et al., 2020)

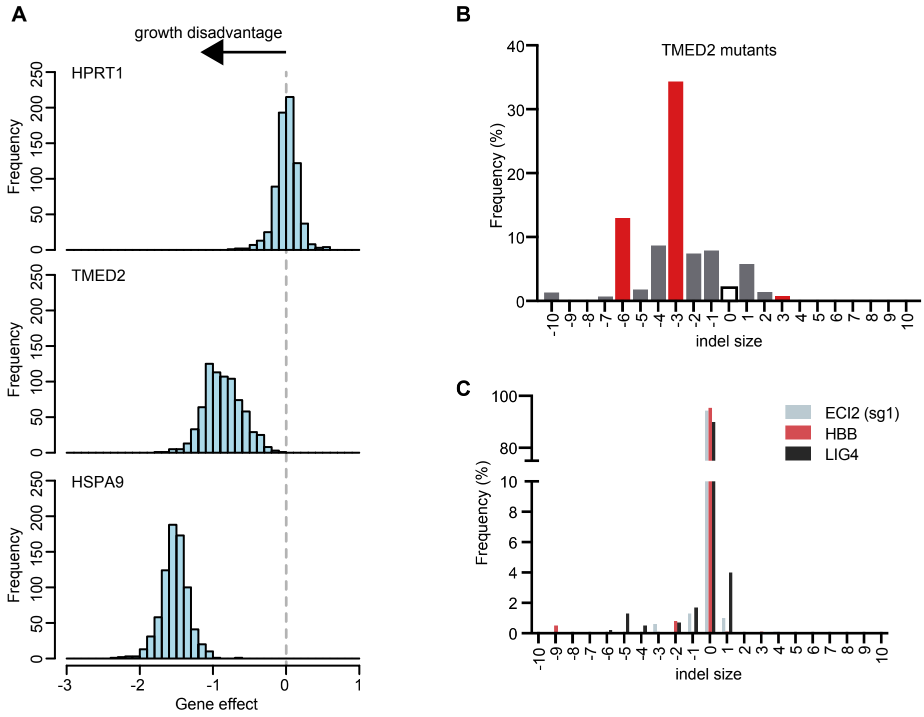


Figure S3. Analysis of failed cases of GENF.
(A) Essentiality of selected targets. Gene effect scores were obtained from Depmap database [44]. The scores illustrate the degree of sgRNA depletion/enrichment from cell populations during genome-wide CRISPR-Cas9 screenings in multiple cancer cell lines. A shift of the traces to the left (as seen for HSPA9 and TMED2) illustrate that the sgRNA targeting the gene is depleted upon culture, showing that the target is critical for cell survival in multiple cell lines. Note that HPRT1, which was used for co-targeting throughout this study, is not essential. (B) Indel patterns in TMED2 mutants analyzed by TIDE. Note that this target could be highly mutated with GENF with barely detectable wild type alleles (white bar), but with a high enrichment of in-frame mutations (red bars). (C) TIDE analysis of cells transiently transfected with pX330-based plasmids encoding the sgRNAs targeting the indicated genes. Note the very low levels of mutation induced with the sgRNAs targeting ECI2 or HBB. The result of a similar transient transfection to mutate LIG4 is illustrated to compare the efficiency, which shows that a slightly better target sgRNA is sufficient to obtain nearly-complete gene deletion with GENF (see LIG4 mutation rates with GENF in Figure 1J).

Figure 2 (Harayama et al., 2020)

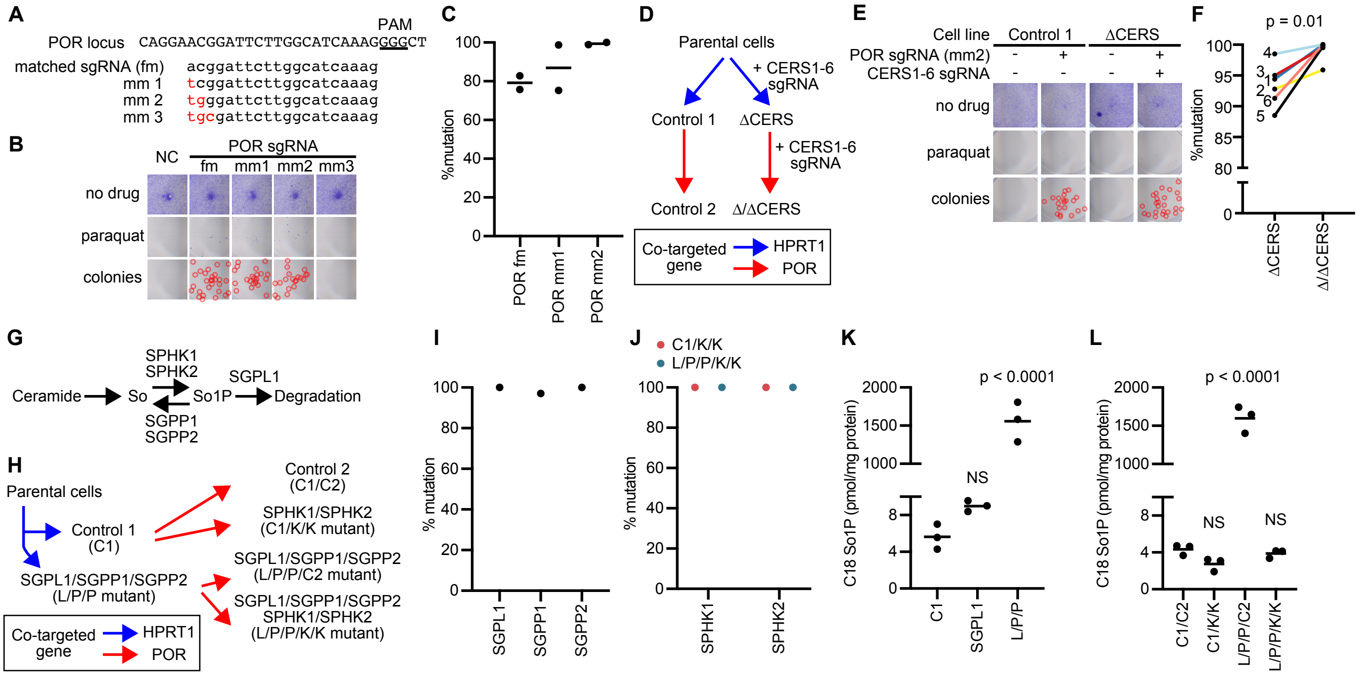


Figure 2. Establishment of GENF with another gene-drug pair

(A) Sequence of POR locus and the sgRNAs targeting it with (mm) or without (fm) mismatches at PAM-distal positions. (B) Cells were stained with Crystal Violet after co-targeting POR and CERS6, with or without treatment with paraquat. The colonies of drug-resistant cells are highlighted with circles in additional panels, due to their smaller size than those obtained after HPRT co-targeting. (C) Mutation rates of CERS6 after co-targeting with the indicated sgRNA and selection with paraquat in two independent experiments. (D) Nomenclature of cell lines obtained during two-step GENF, with the different co-targeted genes color coded. (E) Survival of cells after the second round of two-step GENF (corresponding to the red arrows in (D)), which was analyzed similarly to (B). Note the absence of paraquat-resistant cells when POR was not co-targeted. (F) Mutation rates of CERS genes (see numbers) all targeted together, in the first step (ΔCERS) or the second step (Δ/ΔCERS) of GENF. See the increase in mutation rates for all targets. p value: paired t-test. (G) The metabolic pathway that was manipulated in another experiment for two-step GENF. Sphingosine (So) is generated through degradation of ceramide and is phosphorylated into sphingosine 1-phosphate (So1P) by sphingosine kinases (SPHK1, SPHK2). So1P is dephosphorylated by So1P phosphatases (SGPP1, SGPP2) or degraded by So1P lyase (SGPL1). (H) Combination of genes mutated during two-step GENF, together with the nomenclature of the mutants. (I) Mutation rates in L/P/P mutants generated in the first step illustrated in (H). (J) Mutation rates of genes targeted in the second step illustrated in (H). (K and L) So1P levels in mutants generated in the first (K) and second (L) steps. p values: one-way ANOVA followed by Dunnett’ s multiple comparison test to compare with controls C1 (K) or C1/C2 (L).

Figure 3 (Harayama et al., 2020)

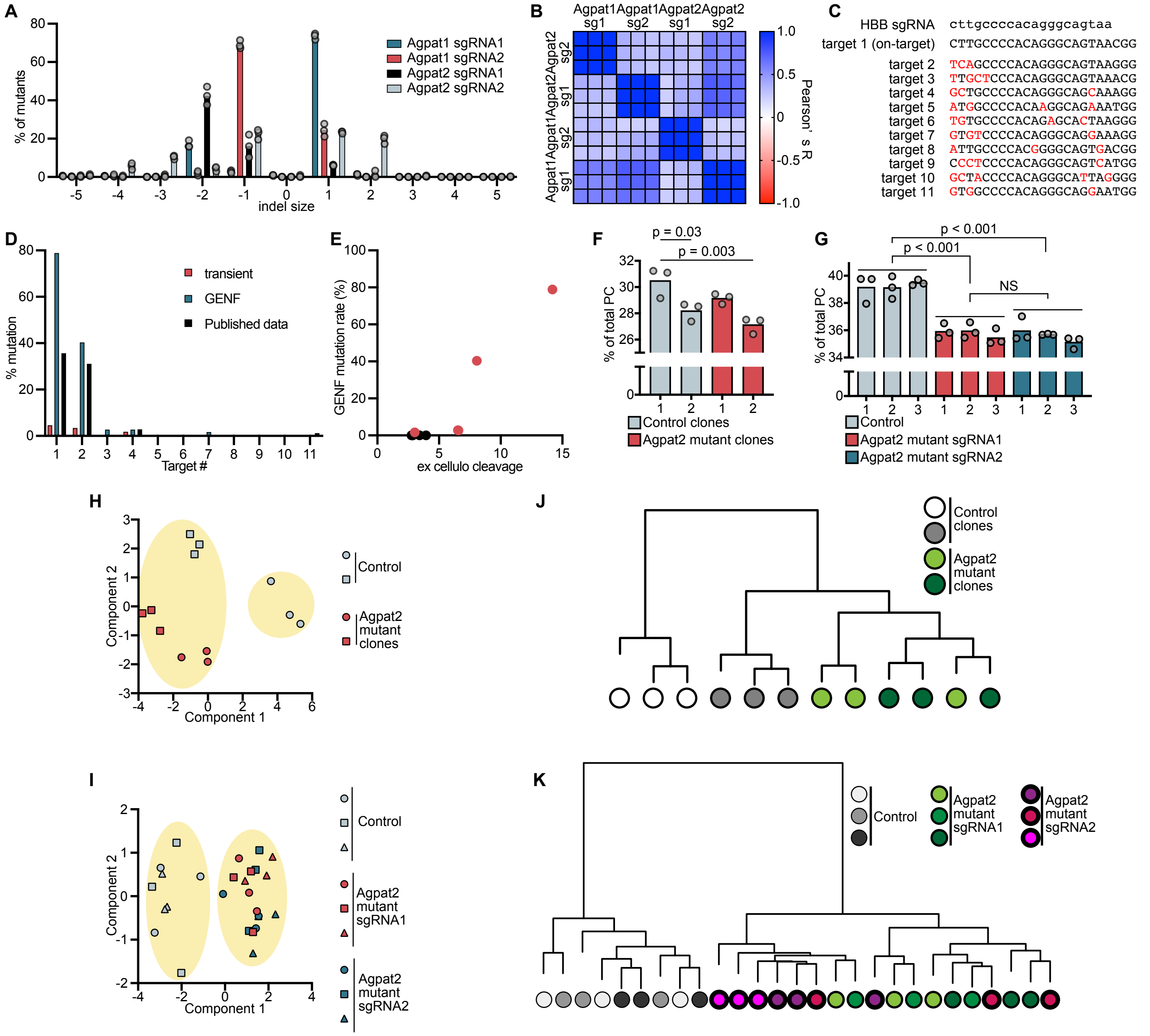


Figure 3. Characterization of mutants generated by GENF

(A and B) Indel patterns seen in 12 mutant cell lines generated with four different sgRNAs, and similarity between them assessed by Pearson's correlation coefficients. (C-E) Analysis of off-target mutations in HBB mutant cells. (C) Off-target candidates of HBB sgRNA reported from ex cellulo cleavage experiments, with mismatches highlighted in red [20]. (D) On- and off-target mutations in HBB mutant cells generated with GENF, as compared to those obtained in a published study [20]. Mutations obtained by transient transfection are also shown to demonstrate the enrichment obtained by GENF. (E) On- and off-target mutation rates in HBB mutants compared with published ex cellulo cleavage results [20]. Note that mutation rates are well predicted from ex cellulo cleavage. (F-K) Comparison of lipid analysis datasets obtained from clones and polyclonal cells generated with GENF. (F and G) Levels of the most abundant phosphatidylcholine (PC) species, PC 34:1, in different control and Agpat2 mutant clones (F) or in different polyclonal cell lines generated with GENF (G). p values: one-way ANOVA followed by Tukey's multiple comparisons test comparing all pairs. (H and I) Principal component analysis of PC acyl-chain composition datasets obtained from clonal cell lines (H) or polyclonal cell lines generated with GENF (I). Encircled are the results of K-means clustering to split datasets into two groups. (J and K) Hierarchical clustering to assess the similarity of the same datasets from clonal cell lines (J) or polyclonal cell lines generated with GENF (K). See also Figure S4.

Figure S4 (Harayama et al., 2020)

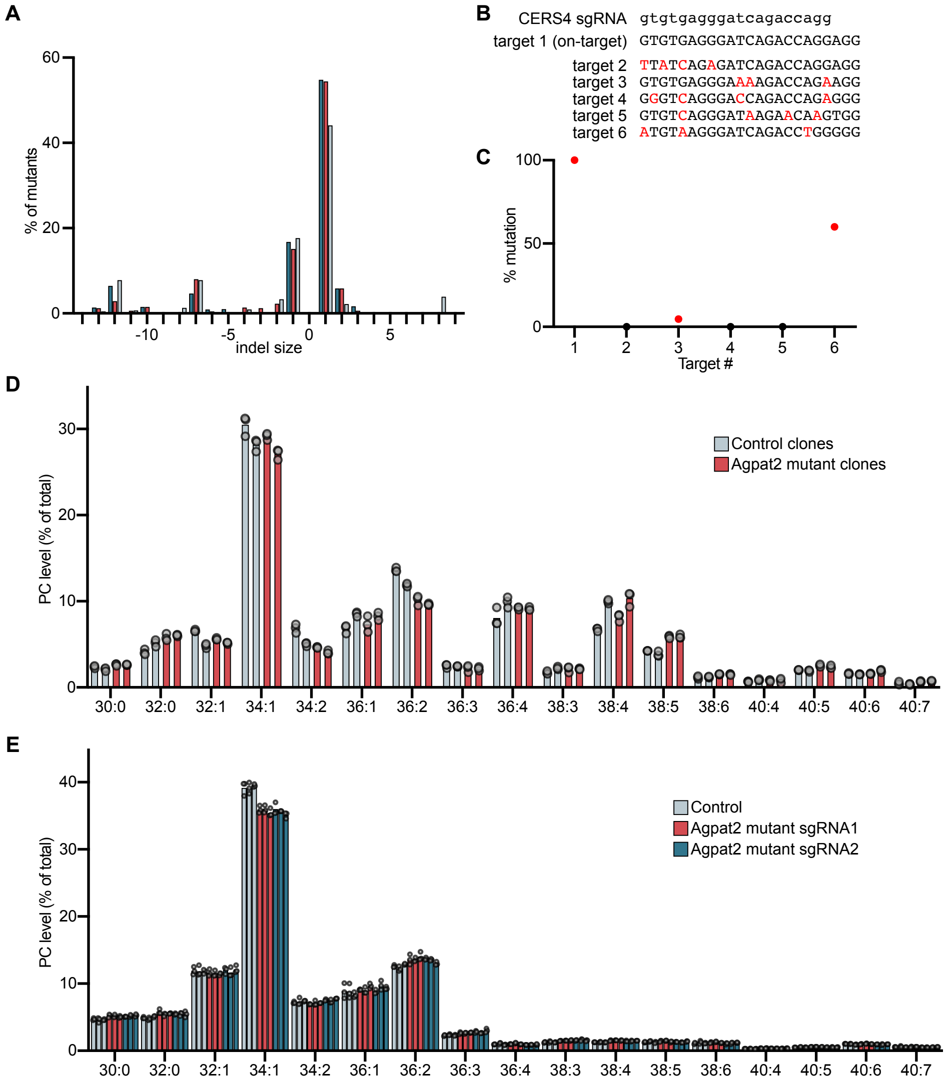


Figure S4. Additional data to characterize GENF.

(A) Indel patterns of CERS6 mutants generated at different days. (B) Off-target sites of CERS4 sgRNA predicted based on homology, with mismatches highlighted in red. (C) On- and Off-target mutations observed in CERS4 mutant cells. Non-zero values are plotted in red. (D) Comparison of PC composition between control and Agpat2 mutant clones. Different clones are shown as distinct bars. (E) Comparison of PC composition between polyclonal control and Agpat2 mutant cells generated with GENF. Different polyclonal cell lines are shown as distinct bars. We know empirically that the content of polyunsaturated glycerophospholipids vary largely between datasets obtained in different days, which is likely due to differences in various factors including the conditions of the serum, which is the sole source of polyunsaturated fatty acids in standard culture conditions. The differences in control cells between (D) and (E) are likely to be caused by this variability, since the datasets were obtained years apart.

Figure 4 (Harayama et al., 2020)

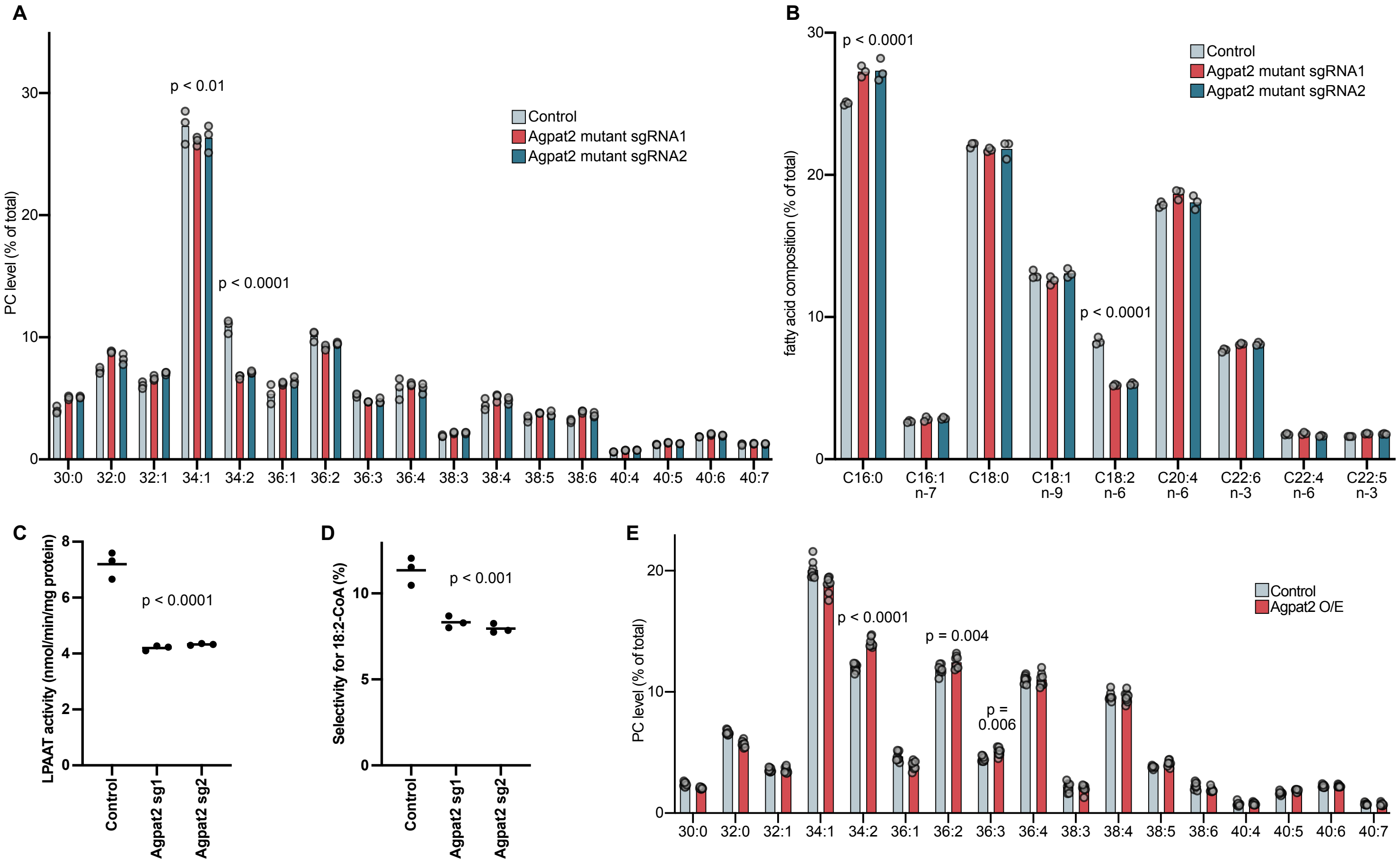
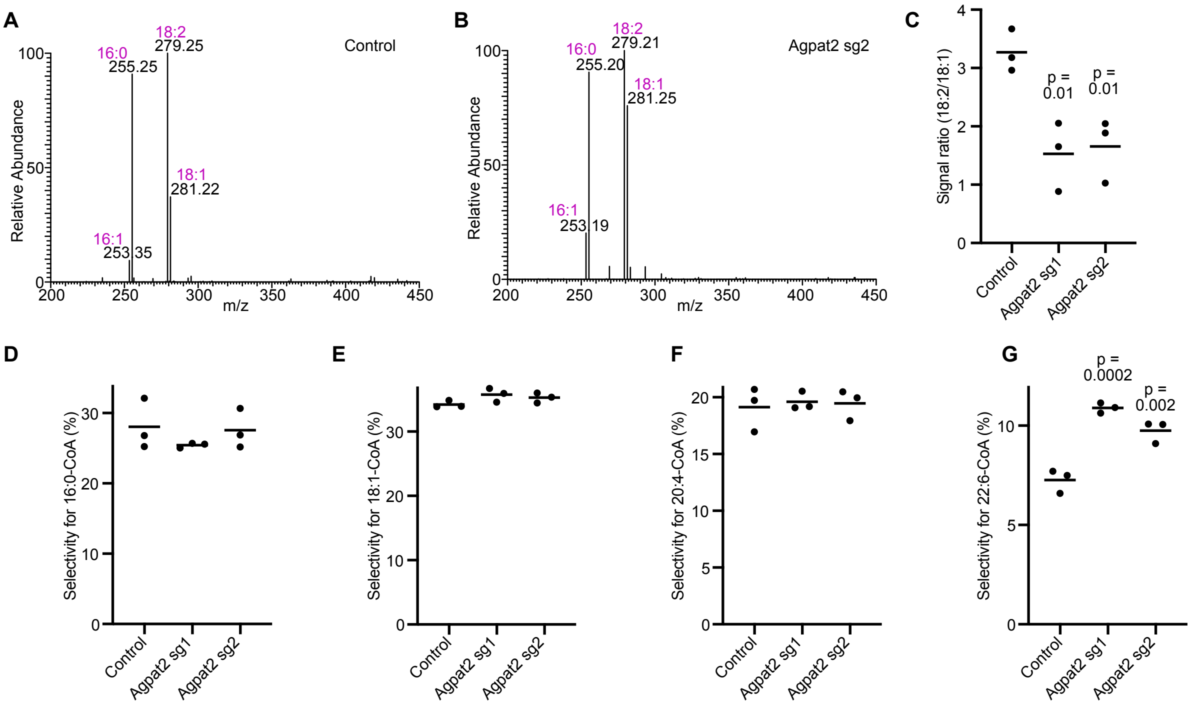


Figure 4. Agpat2 regulates linoleic acid levels in PC

(A) PC acyl-chain composition in control or Agpat2 mutant cells supplemented with linoleic, arachidonic, and docosahexaenoic acid (PUFA supplementation). p values: two-way ANOVA followed by Dunnett’ s multiple comparisons test to compare with Control. Only reductions seen in both Agpat2 mutants are illustrated for simplicity. (B) Gas chromatography analysis of fatty acid methyl esters obtained from phospholipid fractions of PUFA-supplemented control or Agpat2 mutant cells. p values: two-way ANOVA followed by Dunnett’ s multiple comparisons test to compare with Control. Common changes are illustrated. (C and D) Lysophosphatidic acid acyltransferase (LPAAT) activity, and its selectivity for linoleoyl-CoA in membrane fractions obtained from control or Agpat2 mutant cells. p values: one-way ANOVA followed by Dunnett’ s multiple comparisons test to compare with Control. (E) PC acyl-chain composition in PUFA-supplemented control or Agpat2-overexpressing (O/E) cells. p values: two-way ANOVA followed by Sidak’ s multiple comparisons test. Only increases are illustrated for simplicity. See also Figure S5.

Figure S5 (Harayama et al., 2020)



H

	Agpat1		Agpat2	Agpat3				Agpat4		Agpat5		
Dataset	1421024_at	1421025_at	1428821_at	1433817_at	1433818_at	1433819_s_at	1450504_a_at	1428336_at	1436640_x_at	1434287_at	1453257_at	1458669_at
16:0-CoA	-0.02864	-0.03261	-0.2324	-0.7118	-0.5888	-0.754	-0.3002	0.6538	0.6438	0.2957	0.1357	0.07348
18:1-CoA	-0.6519	-0.6648	0.04075	-0.455	-0.5736	-0.5859	-0.7574	0.47	0.5233	0.4542	0.3051	0.2298
18:2-CoA	-0.2396	-0.2602	0.762	-0.09605	-0.3788	-0.2518	-0.1335	0.07682	0.08686	0.06237	-0.08674	0.05403
20:4-CoA	0.05855	-0.003716	0.8044	0.2313	-0.0937	0.02114	0.253	-0.1191	-0.1341	-0.2665	-0.3687	-0.06215
22:6-CoA	0.4389	0.4763	-0.6171	0.3697	0.6497	0.5914	0.438	-0.3881	-0.4129	-0.2446	-0.02338	-0.1405

Figure S5. Agpat2 as a regulator of linoleic acid levels.
(A-C) Fragmentation patterns of PC 34:2 analyzed by tandem mass spectrometry. PC 34:2 from control (A) and Agpat2 mutant (B) cells was analyzed by liquid chromatography-tandem mass spectrometry as a bicarbonate adduct negative ion and fragmented to reveal the fatty acyl composition. Fragments corresponding to fatty acids are annotated. Similar results were obtained in three biological replicates and are quantified in (C). (D-G) Selectivity of lysophosphatidic acid acyltransferase (LPAAT) activity in membrane fractions obtained from control or mutant cells for the indicated substrates. (C-G) p values: one-way ANOVA followed by Dunnett' s multiple comparisons test to compare with Control. (H) The correlation between LPAAT acyl-CoA selectivity [23] and mRNA expression [45] of various LPAAT enzymes in various tissue. Values above 0.4 are colored.

Figure 5 (Harayama et al., 2020)

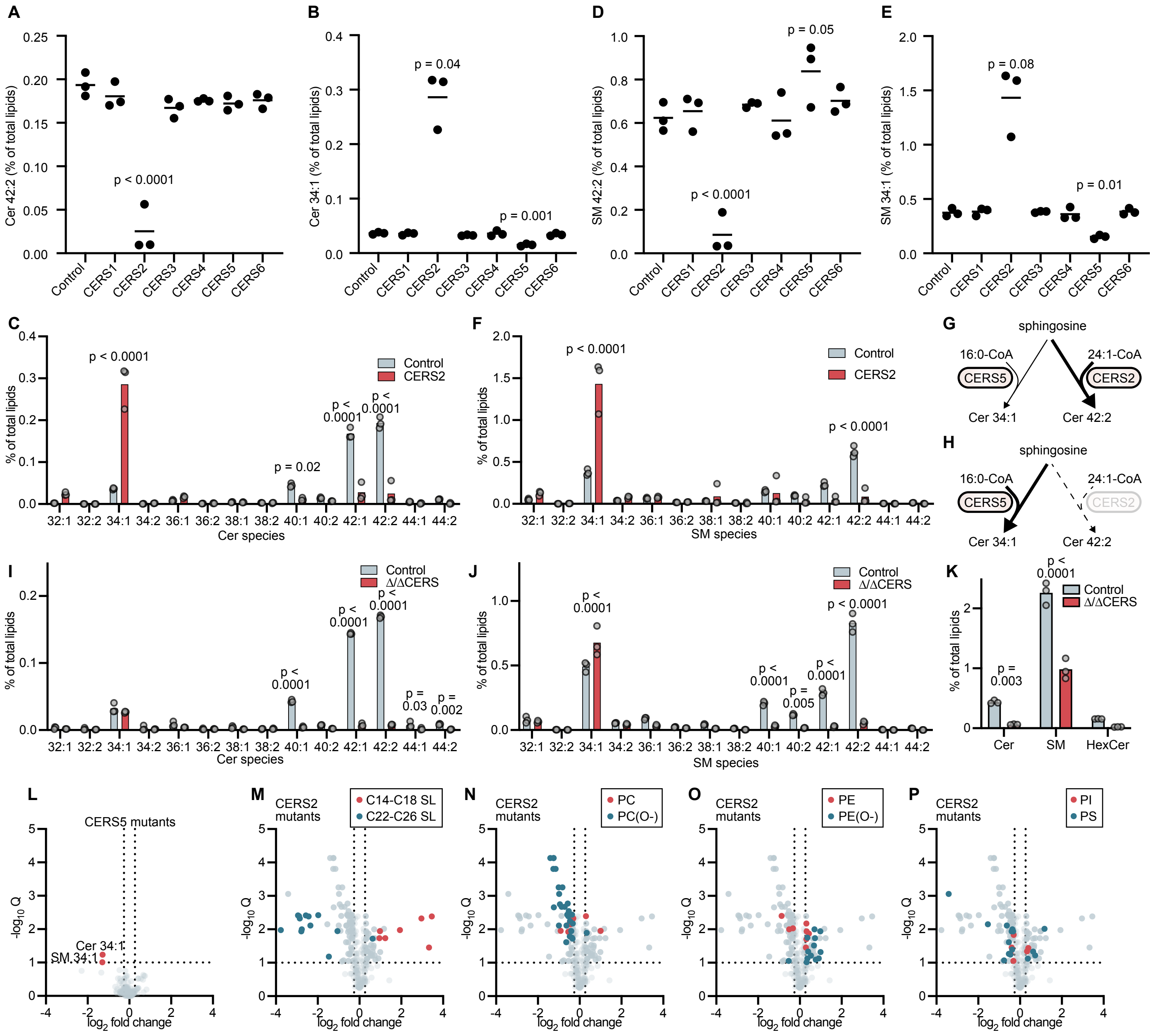


Figure 5. Lipidome changes induced by mutations in CERSs.

(A and B) Levels of the indicated ceramide (Cer) species in various CERS mutants. (C) Changes in Cer species in CERS2 mutants. (D and E) Changes in the indicated sphingomyelin (SM) species in various CERS mutants. (F) Changes in SM species in CERS2 mutants. (G and H) The competition between CERS2 and CERS5 explains the changes in SL profiles seen in CERS2 mutants. (I and J) Changes in Cer (I) and SM (J) species in Δ/ΔCERS mutants. (K) Changes in total levels of the indicated sphingolipid species in Δ/ΔCERS mutants. p values: (A and D) one-way ANOVA followed by Dunnett' s multiple comparisons test to compare with Control, (B and E) Brown-Forsythe and Welch ANOVA tests followed by Dunnett' s T3 multiple comparisons test to compare with Control, (C, F, and I-K) two-way ANOVA followed by Sidak' s multiple comparisons test. (L-P) Volcano plot illustrating changes in lipid levels and their statistical significances in the indicated CERS mutants. Lipids with changes above 1.2-fold and q values (FDR-corrected multiple t-test) below 0.1 were considered as hits. (M-P) For CERS2 mutants, different lipid classes are highlighted separately. See also Figure S6.

Figure S6 (Harayama et al., 2020)

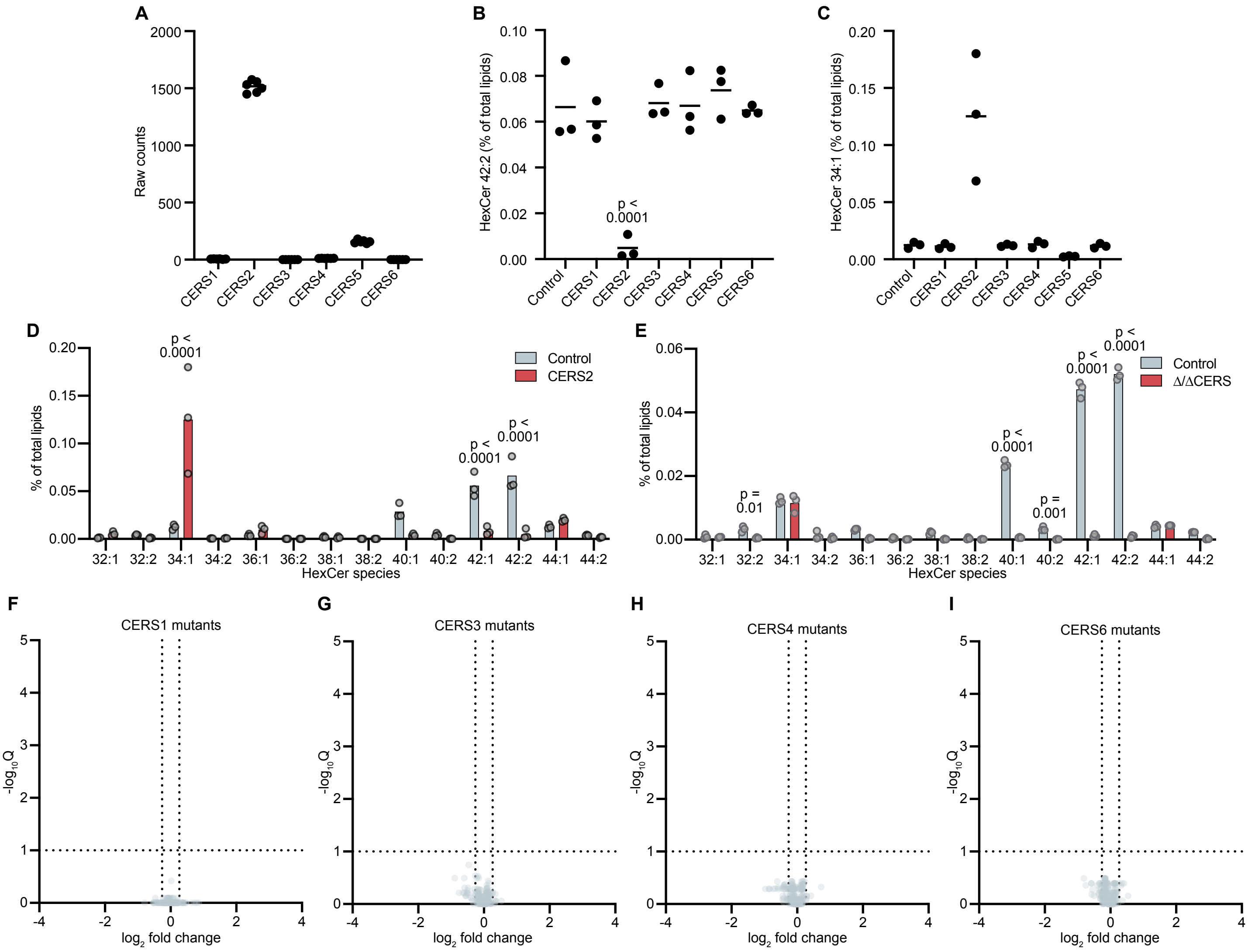


Figure S6. Lipid changes in CERS mutants.

(A) Expression of CERSs in HeLa MZ cells obtained from RNA-seq [29]. (B and C) Changes in the indicated Hexosylceramide (HexCer) species in CERS mutants. (B) p values: one-way ANOVA followed by Dunnett' s multiple comparisons test to compare with Control. (D and E) Changes in HexCer species in CERS2 and Δ/ΔCERS mutants. p values: two-way ANOVA followed by Sidak' s multiple comparisons test. (F-I) Volcano plots illustrating changes in lipid levels and statistical their significances seen in the indicated CERS mutants. No lipid is identified as a hit (having changes above 1.2-fold and q values (FDR-corrected multiple t-test) below 0.1).

Figure 6 (Harayama et al., 2020)

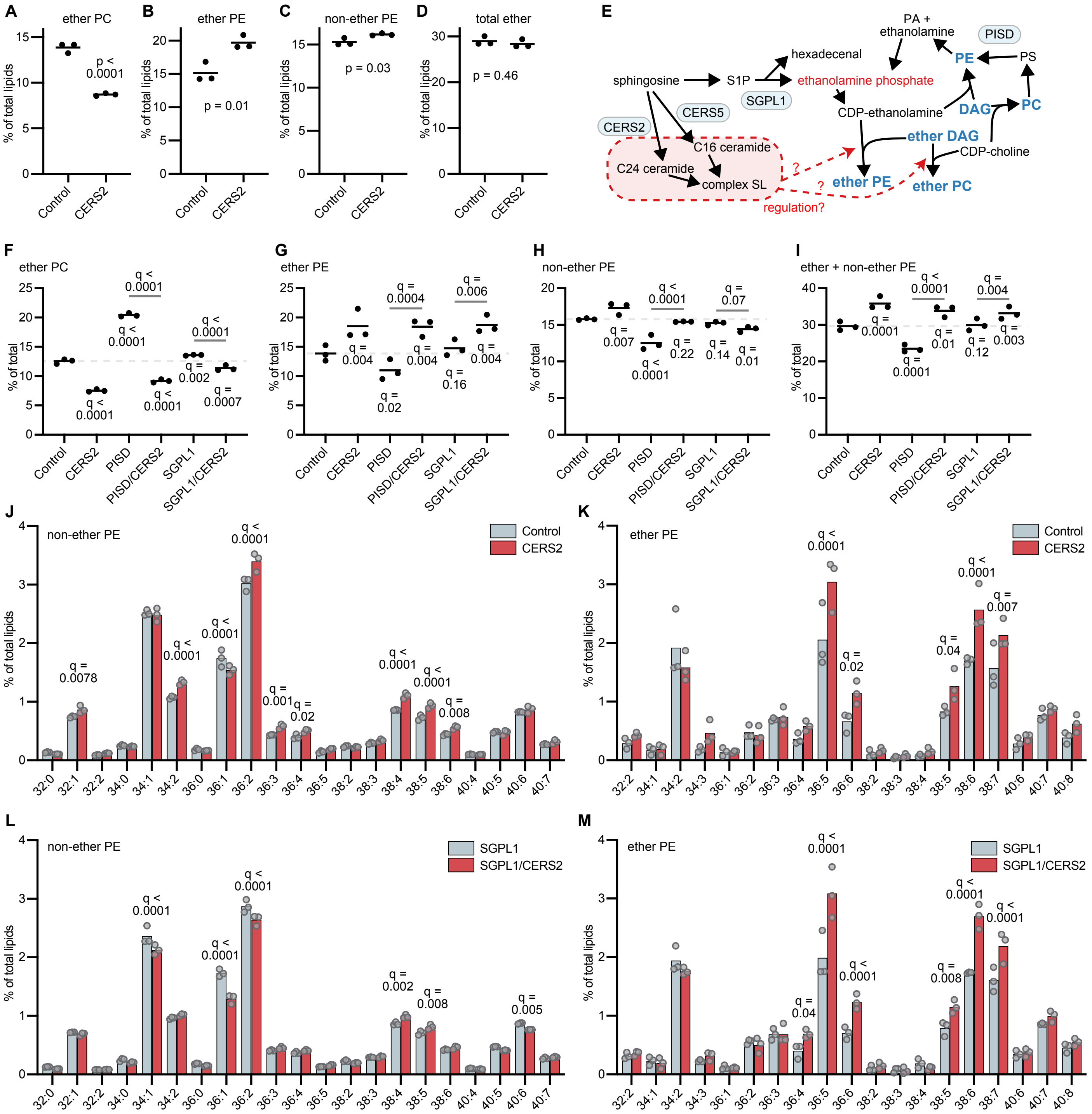


Figure 6. Dissection of SL and ether lipid co-regulation.

(A-D) Changes in the indicated lipid classes in CERS2 mutants. p values: unpaired t-test. (E) Metabolic pathways linking SLs and glycerophospholipids. Red arrows indicate potential metabolic regulations suggested from the whole dataset. (F-I) Changes in the indicated lipid classes in the indicated mutants. q values: one-way ANOVA followed by 5% FDR-corrected multiple t-test. Values below the bars are comparisons with Control, and those above are comparisons between the indicated pairs. (J-M) Changes in non-ether PE and ether PE in the indicated mutants. q values: two-way ANOVA followed by 5% FDR-corrected multiple t-test.

Figure 7 (Harayama et al., 2020)

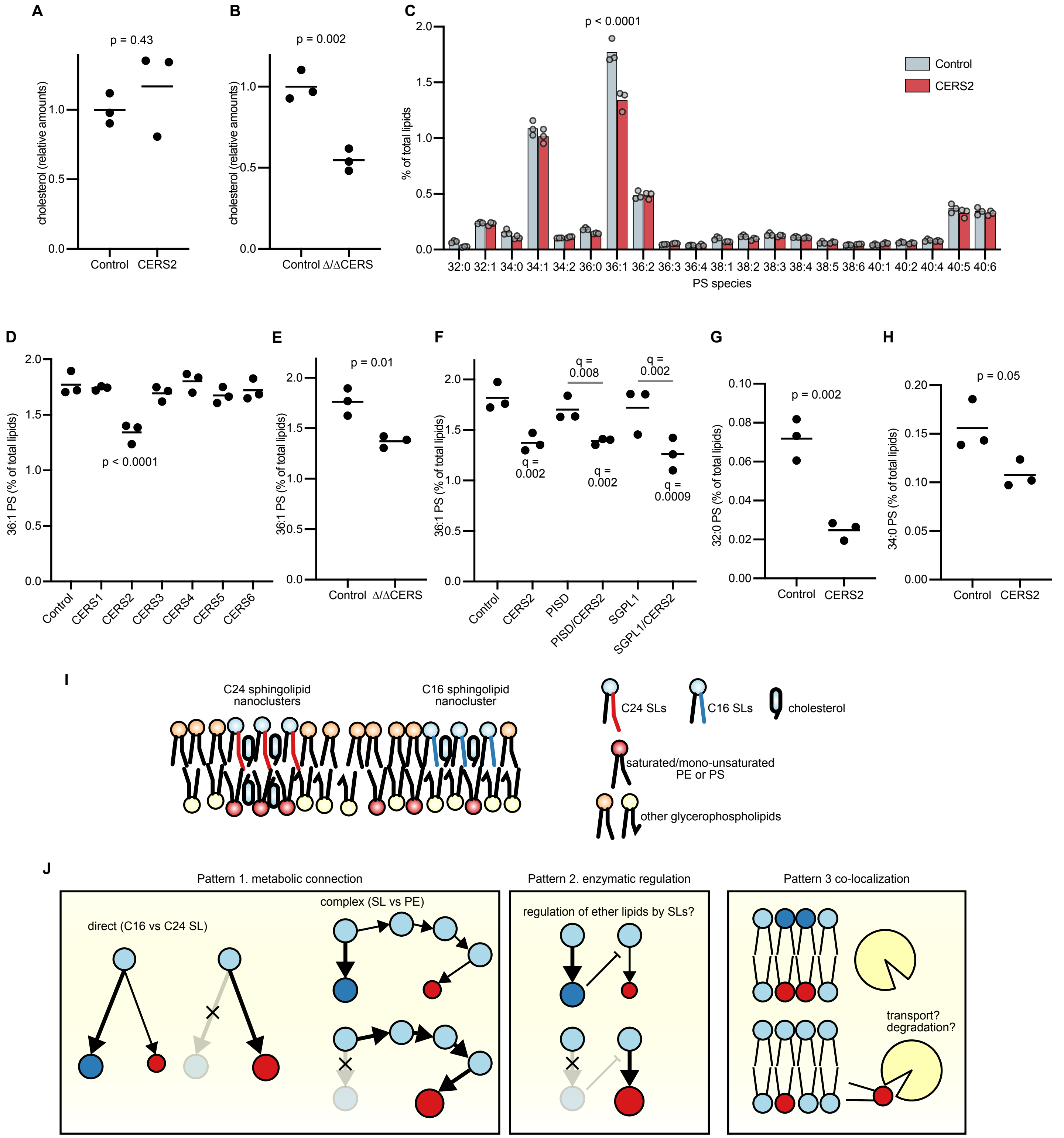


Figure 7. Co-localization as a mechanism of lipid co-regulation.

(A and B) Relative cholesterol changes in CERS2 (A) and Δ/ΔCERS (B) mutants. p values: unpaired t-test. (C) Changes in PS species in CERS2 mutants. p values: two-way ANOVA followed by Sidak' s multiple comparisons test. (D-F) Changes in PS 36:1 levels in the indicated mutants. p values: one-way ANOVA followed by Dunnett' s multiple comparisons test to compare with Control (D) and unpaired t-test (E); q values: one-way ANOVA followed by 5% FDR-corrected multiple t-test. Values below the bars are comparisons with Control, and those above are comparisons between the indicated pairs. (G and H) Changes in the indicated PS species in CERS2 mutants. p values: unpaired t-test. (I) Proposed model of nanodomains generated by SLs of different N-acyl-chains. (J) Generalized view of lipid co-regulatory mechanisms seen in this study.