Single-cell atlas of ABCA7 loss-of-function reveals impaired neuronal respiration via choline-dependent lipid imbalances

Djuna von Maydell[1,2,†], Shannon Wright[1,2,†], Julia Maeve Bonner[1,2,], Colin Staab[1,2], Andrea Spitaleri[3], Liwang Liu[1,2], Ping-Chieh Pao[1,2], Chung Jong Yu[1,2], Aine Ni Scannail[1,2], Mingpei Li[1,2], Carles A. Boix[4,5], Hansruedi Mathys[1,2,‡], Guillaume Leclerc[4], Gloria Suella Menchaca[1,2], Gwyneth Welch[1,2], Agnese Graziosi[1,2], Noelle Leary[1,2], George Samaan[1,2], Manolis Kellis[4,5], Li-Huei Tsai[1,2,*]

Affiliations:

[1]Picower Institute for Learning and Memory, Massachusetts Institute of Technology; Cambridge, MA 02139, USA.

[2]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology; Cambridge, MA 02139, USA.

[3]Department of Biotechnology and Translational Medicine University of Milan, Via F.lli Cervi, 93 - L.I.T.A., 20054 Segrate, MI, Italy

[4]MIT Computer Science and Artificial Intelligence Laboratory; Cambridge, MA 02139, USA.

[5]Broad Institute of MIT and Harvard; Cambridge, MA 02139, USA.

†These authors contributed equally to this work.

‡Present address: Department of Neurobiology, University of Pittsburgh; Pittsburgh, PA 15261

*Corresponding author. Email: lhtsai@mit.edu

Abstract

Loss-of-function (LoF) variants in the lipid transporter ABCA7 significantly increase the risk of Alzheimer's disease (odds ratio ~2), yet the pathogenic mechanisms and the neural cell types affected by these variants remain largely unknown. Here, we performed single-nuclear RNA sequencing of 36 human *post-mortem* samples from the prefrontal cortex of 12 ABCA7 LoF carriers and 24 matched non-carrier control individuals. ABCA7 LoF was associated with gene expression changes in all major cell types. Excitatory neurons, which expressed the highest

levels of ABCA7, showed transcriptional changes related to lipid metabolism, mitochondrial function, cell cycle-related pathways, and synaptic signaling. ABCA7 LoF-associated transcriptional changes in neurons were similarly perturbed in carriers of the common AD missense variant ABCA7 p.Ala1527Gly (n = 240 controls, 135 carriers), indicating that findings from our study may extend to large portions of the at-risk population. Consistent with ABCA7's function as a lipid exporter, lipidomic analysis of isogenic iPSC-derived neurons (iNs) revealed profound intracellular triglyceride accumulation in ABCA7 LoF, which was accompanied by a relative decrease in phosphatidylcholine abundance. Metabolomic and biochemical analyses of iNs further indicated that ABCA7 LoF was associated with disrupted mitochondrial bioenergetics that suggested impaired lipid breakdown by uncoupled respiration. Treatment of ABCA7 LoF iNs with CDP-choline (a rate-limiting precursor of phosphatidylcholine synthesis) reduced triglyceride accumulation and restored mitochondrial function, indicating that ABCA7 LoF-induced phosphatidylcholine dyshomeostasis may directly disrupt mitochondrial metabolism of lipids. Treatment with CDP-choline also rescued intracellular amyloid β-42 levels in ABCA7 LoF iNs, further suggesting a link between ABCA7 LoF metabolic disruptions in neurons and AD pathology. This study provides a detailed transcriptomic atlas of ABCA7 LoF in the human brain and mechanistically links ABCA7 LoF-induced lipid perturbations to neuronal energy dyshomeostasis. In line with a growing body of evidence, our study highlights the central role of lipid metabolism in the etiology of Alzheimer's disease.

Introduction

Over 50 million people worldwide have dementia, with a large fraction of cases caused by Alzheimer's disease[1]. Late-onset Alzheimer's Disease (AD) affects individuals over the age of 65 and accounts for more than 95% of all AD cases[2]. Though AD is a multifactorial disorder, twin studies suggest a strong genetic component (~70% heritability)[3] contributing to AD disease risk and progression. Large scale genome-wide association studies implicate multiple genes in AD etiology[4–11]. After APOE4, rare loss-of-function (LoF) mutations caused by premature termination codons (PTCs) in ATP-binding cassette transporter A7 (ABCA7), are among the strongest genetic factors for AD (odds ratio ~ 2)[10,12–16]. In addition to LoF variants, several common single nucleotide polymorphisms in ABCA7 - depending on the population - moderately[10,12–14,17–19] to strongly[14] increase AD risk, suggesting that ABCA7 dysfunction may play a role in a significant proportion of AD cases. Despite its prevalence and potential impact, the mechanism by which ABCA7 dysfunction increases AD risk remains poorly characterized.

ABCA7 is a member of the A subfamily of ABC transmembrane proteins[20] with high sequence homology to ABCA1, the primary lipid transporter responsible for cholesterol homeostasis and high-density lipoprotein genesis in the brain[21]. ABCA7 effluxes both cholesterol and phospholipids to APOA-I and APOE in *in vitro* studies[22–26] and has been shown to be a critical regulator of lipid metabolism, immune cell functions, and amyloid processing[27–31]. To date, study of ABCA7 LoF has been predominantly pursued in rodent knock-out models or in non-neural mammalian cell lines. These studies show that ABCA7 knock-out or missense variants cause

increased amyloid processing and deposition[32–35], reduced plaque clearance by astrocytes and microglia[36,37], and that ABCA7 LoF alters glial-mediated inflammatory responses[38,39]. While these studies shed light on potential mechanisms of ABCA7 risk in AD, studies investigating the effects of ABCA7 LoF in human cells and tissue are severely lacking, with only a small number published to date[30,35,40,41]. These human studies highlight a number of potential LoF-induced defects in human cells, including impacts on lipid metabolism and mitochondrial function[30]. However, comprehensive and unbiased profiling of multiple human neural cell types is needed to elucidate the mechanism by which ABCA7 LoF increases AD risk.

Recent work has used single-nuclear RNA sequencing of human neural tissue to reveal the cell type specific deficits associated with genetic variants in AD (including APOE4 and TREM2) and in other disorders[42–46], providing insight into disease mechanisms and possible therapeutic interventions. In this study, we generated a cell type-specific transcriptomic atlas of ABCA7 LoF effects in the human prefrontal cortex (PFC) by single-nuclear RNA sequencing (snRNAseq) of *postmortem* tissue from ABCA7 LoF variant carriers and matched control individuals. We found that ABCA7 LoF was linked to transcriptomic perturbations in all major brain cell types in the human PFC and that excitatory neurons, which expressed ABCA7 most highly, showed substantial transcriptomic perturbations to genes related to lipid metabolism, mitochondrial respiration, DNA damage, and synaptic function. We complemented our transcriptomic analysis with biochemical experiments on *postmortem* human brain and neurons derived from isogenic induced pluripotent stem cells (iPSC).

Experiments in iPSC-derived neurons (iNs) revealed a metabolic shift away from phosphatidylcholine synthesis towards triglyceride accumulation in the presence of ABCA7 LoF, trends which were also observed in the *postmortem* human brain. This shift coincided with widespread metabolic changes related to carbon catabolism through mitochondrial oxidative phosphorylation (OXPHOS), indicating that ABCA7 LoF neurons may have decreased ability to metabolize lipids. Promoting phosphatidylcholine synthesis and reducing triglyceride accumulation by supplementing iNs with CDP-choline, a rate-limiting precursor for phosphatidylcholine synthesis, ameliorated this shift, suggesting that ABCA7 LoF-induced lipid disruptions mediate mitochondrial dysfunction in neurons. Supplementation with CDP-choline also rescued intracellular Aβ42, the toxic form of amyloid-β[47–51], which was up-regulated in ABCA7 LoF neurons, suggesting a potential link between ABCA7 LoF-induced metabolic defects and AD pathology.

Together, our data suggest that neuronal ABCA7 plays an integral role in regulating intracellular lipid homeostasis critical for mitochondrial function. Our study reveals new insights into the mechanism through which damaging ABCA7 variants, including common ABCA7 missense variants, may increase AD risk by disrupting energy homeostasis in the cell.

Results

Single-nuclear transcriptomic profiling of human PFC from ABCA7 LoF-variant carriers

To investigate the cell type-specific impact of ABCA7 LoF variants in the human brain, we queried whole genome sequences of >1000 subjects from the Religious Order Study or the Rush Memory and Aging Project (collectively known as ROSMAP) for donors with Alzheimer's disease diagnoses who are carriers of rare damaging variants in ABCA7 that result in a PTC. We identified 12 heterozygous carriers of ABCA7 LoF variants, including splice region variants (c.4416+2T>G and c.5570+5G>C), frameshift variants (p.Leu1403fs and p.Glu709fs), and nonsense 'stop gained' variants (p.Trp1245* and p.Trp1085*) (**Fig. 1A; Fig. S1A,B; Data S1**). These variants have previously been associated with increased AD risk in genetic association studies (**Table S1**)[12,15] and are presumed to induce risk via ABCA7 haploinsufficiency[52]. Analysis of published proteomic data[53] confirmed that ABCA7 PTC-variant carriers indeed had lower ABCA7 protein levels in the human *postmortem* PFC compared to non-carriers (p=0.018; N = 180 Controls, 5 ABCA7 LoF; **Fig. 1B; Fig. S1C**).

We then performed snRNAseq on *postmortem* PFC samples from the 12 ABCA7 LoF variant carriers and 24 ABCA7 PTC non-carrier controls that were matched based on a number of potentially confounding variables, including AD pathology, age at death, *post-mortem* intervals, sex, APOE genotype, and cognition (**Fig. 1C;D; Fig. S1D,E; Fig. S2A,B; Data S2; Supplementary Text**). We verified that none of the 36 selected subjects carried damaging variants in other AD risk genes (*TREM2, SORL1, ATP8B4, ABCA1,* and *ADAM10*)[15] and we confirmed ABCA7 genotypes in a subset of ABCA7 LoF carriers and controls in the cohort by Sanger sequencing (**Fig. S1A**) and verified that all single-cell libraries matched their whole genome sequencing counterparts to ensure all samples were appropriately assigned to their respective genotypes (**Fig. S1B**).

For each individual, we obtained fresh frozen tissue from the BA10 region of the PFC and performed snRNAseq using the 10x Genomics Chromium platform, resulting in a total of 150,456 cells (102,710 cells after multiple rounds of quality control) (**Fig. 1E**). After quality control (qc; dimensionality reduction, batch correction, and multiple rounds of clustering; **Fig. S2C-E; Methods**) and cell type annotation with curated marker lists, we identified heterogeneous populations of inhibitory neurons (In, *SYT1 & GAD1+*), excitatory neurons (Ex, *SYT1 & NRGN+*), astrocytes (Ast, *AQP4+*), microglia (Mic, *CSF1R+*), oligodendrocytes (Oli, *MBP & PLP1+),* and oligodendrocyte precursor cells (OPCs, *VCAN+*) (**Fig. 1E; Fig. S2F-I**). The putative vascular cell cluster was small and did not meet qc cutoffs (**Methods**) and was therefore not considered for downstream analysis. After qc, individual-level gene expression profiles correlated well within cell types (mean correlation *circa* 0.95) and were well represented across individuals (**Fig. S2J-M**).

Cell type-specific perturbations in the presence of ABCA7 LoF

To explore ABCA7-LoF related gene expression changes across all major cell types, we filtered genes based on evidence for transcriptional perturbation in ABCA7 LoF ($p<0.05$, linear model, total #genes passing this cutoff = 2,389) in at least one cell type (Ex, In, Ast, Mic, Oli, or OPC), while controlling for known and unknown covariates in the model and considering only genes with cell type-specific detection rates >10% (**see Methods; Data S3**). Next, we projected perturbed genes from higher dimensional gene score space (where each dimension corresponds to cell type-specific ABCA7 LoF perturbation scores *S,* where *S=sign(log(FC))*-log10(p-value)*))) into two dimensions for visualization and clustering (**Fig. S3A; Methods**).

The 2D visual representation captured the transcriptional landscape of ABCA7 LoF gene changes across all major cell types (**Fig1F; Fig. S3B,C**). To summarize this landscape in terms of biological pathways, we assigned genes to clusters based on their projections in 2D space and quantified their enrichment for specific biological pathways (**Fig. 1G;** Gene Ontology Biological Process database; **Methods**). This analysis indicated a number of biological pathways whose perturbation was broadly associated with ABCA7 LoF in *postmortem* human PFC, including pathways related to heat shock response, inflammation, cell dynamics and adhesion, synaptic function, DNA repair, and cellular metabolism (**Fig. 1G; Data S4**).

Decomposing the ABCA7 LoF transcriptional signature into cell type-specific patterns indicated which ABCA7 LoF-associated gene clusters were commonly or uniquely perturbed across major cell types in the PFC (**Figure 1H; Fig. S3D; Data S3-5; Methods**). For example, microglia exhibited a down-regulated heat shock response (cluster 4), also present to a lesser extent in neurons and OPCs (FDR-adjusted permutation p-value<0.01 & |score|>0.25; **Fig. 1H**). Microglia and astrocytes displayed up-regulated inflammatory clusters, such as cluster 5 and cluster 10, respectively (**Fig. 1H**). OPCs demonstrated perturbed cell motility and endocytic-associated clusters (clusters 1 and 0; **Fig. 1H**). Both neuronal subtypes showed up-regulation of DNA repair-associated cluster 9 (*e.g., NUP205, POLD3, VCP*; **Fig. 1H**) and down-regulation of clusters 6 and 14, characterized by cell adhesion and synaptic genes (*e.g., SYT11, SHANK3, IFT56, GABRA3, NLGN1*; **Fig. 1H**). In addition, excitatory neurons exhibited uniquely perturbed clusters, including up-regulated cluster 8 (mitochondrial complex genes, *e.g., COX7A2*, antioxidant response *e.g., PARK7*) and down-regulated cluster 2 (lipid synthesis and transport genes, *e.g., LDLR, APOL2, NR1H3, ACLY, FITM2*; **Fig. 1H**).

Together, these findings indicate that ABCA7 LoF variants may induce widespread, cell type-specific transcriptional changes in the human PFC. This single-cell atlas provides a rich resource for future studies aiming to elucidate the contributions of individual neural cell types to ABCA7 LoF-driven forms of AD risk. This resource will be made available for exploration via the *UCSC Single Cell Browser* and for further analysis via *Synapse* (accession ID: syn53461705).

ABCA7 is expressed most highly in excitatory neurons

Our snRNAseq data suggest that excitatory neurons expressed the highest levels of ABCA7, compared to other major cell types in the brain (**Fig. 2A**). ABCA7 transcripts were detected (count>0) in ~ 30% of excitatory neurons and ~ 15% of inhibitory neurons, while the detection rate was considerably lower (<10%) for microglia and astrocytes and an order of magnitude lower (<3%) for oligodendrocytes and OPCs (**Fig. S4A, B**). We validated this expression pattern in an independent published dataset from[54] (**Table S2**), where bulk RNA sequencing of NeuN- (glial) and NeuN+ (neuronal) cell populations derived from six human *postmortem* temporal cortex samples showed significantly higher ABCA7 levels in the neuronal population *versus* the glial cell population (p=0.021, paired t-test; **Fig. S4C**). Several control genes, whose expression patterns in glial *versus* neuronal cells are well established (*ABCA1*, *APOE*, and *NEUROD1*), had expected expression patterns that matched those in the snRNAseq data (**Fig. S4B,C**). These results indicate that neurons, particularly excitatory neurons, are the primary ABCA7-expressing cell type in the aged human PFC. Given the relatively higher expression of ABCA7 in excitatory neurons and the evidence of transcriptional perturbations by ABCA7 LoF in this cell type, we focused our subsequent analysis specifically on excitatory neurons.

ABCA7 LoF perturbations in excitatory neurons

As an alternative approach to the unsupervised clustering of gene perturbation scores among all cell types, we next used prior knowledge of biological pathway structure to perform an in-depth characterization of perturbed biological processes specifically in ABCA7 LoF excitatory neurons. To this end, we first estimated statistical overrepresentation of biological gene sets (WikiPathways, N pathways = 472) among up and down-regulated genes in ABCA7 LoF excitatory neurons *vs* controls (by GSEA; **Methods**). We observed a total of 34 pathways with evidence for transcriptional perturbation at p<0.05 in excitatory neurons (**Data S6**). Enrichments of these pathways were driven by 268 unique genes ("leading edge" genes[55]; **Data S6**).

To extract unique information from leading-edge genes and limit pathway redundancy, we next separated these genes and their associated pathway annotations into non-overlapping groups, formalized as a graph partitioning problem (**Fig. 2B; Fig. S5; Methods; Supplementary Text**). Establishing gene-pathway groupings of approximately equal size revealed eight biologically interpretable "clusters" associated with ABCA7 LoF in excitatory neurons (**Fig. 2B-D**; **Data S6**). Predominantly, these gene clusters centered around two themes: (1) energy metabolism and homeostasis (C2, C7) and (2) DNA damage (C1, C4, C5), cell stress (C3, C6), and synaptic dysfunction (C0).

Clusters C2 and C7 were primarily defined by genes involved in cellular energetics, including genes related to lipid metabolism, mitochondrial function, and OXPHOS (**Fig. 2C,D**). Cluster 2, characterized by transcriptional regulators of lipid homeostasis (*e.g. NR1H3, ACLY, PPARD)*, exhibited evidence for down-regulation in ABCA7 LoF and featured pathways related to "SREBP signaling" and "lipid metabolism" (**Fig. 2C,D; Data S6**). Cluster 7 comprised multiple

227 mitochondrial complex genes (*e.g. COX7A2, NDUFV2)* responsible for ATP generation from
228 carbohydrate and lipid catabolism and showed up-regulation in ABCA7 LoF (**Fig. 2C,D; Data**
229 **S6**).The remaining clusters C0, C1, and C3-C6 were characterized by DNA damage and
230 proteasomal, inflammatory, and apoptotic mediators. Clusters C1, C3, C4, and C5 were
231 up-regulated in ABCA7 LoF excitatory neurons and characterized by pathway terms such as
232 "DNA damage response," "Parkin-Ubiquitin Proteasomal system," "Signaling pathways in
233 glioblastoma," and "Nucleotide metabolism," respectively (**Fig. 2C,D; Data S6**). They included
234 up-regulated DNA damage/repair and proteasomal genes (*e.g. RECQL, TLK2, BARD1, RBL2,*
235 *MSH6*, *PSMD5, PSMD7*) and were accompanied by increased expression of inflammatory
236 mediators (*e.g. C2, C1S, CLU, PIK3CD* in C0, C4, C6) and down-regulated gene signatures
237 related to synaptic function (C0), cytoskeletal function (C3), and apoptosis-related genes (C6)
238 (**Fig. 2C,D; Data S6**).

239 Together, these data suggest that ABCA7 LoF may disrupt energy metabolism in excitatory
240 neurons and that these disruptions coincide with a state of increased cellular stress, characterized
241 by genomic instability and neuronal dysfunction.

242 *ABCA7 LoF and common missense variants lead to overlapping neuronal perturbations*

243 ABCA7 LoF variants substantially increase AD risk (Odds Ratio = 2.03)[12] but are rare and
244 therefore only contribute to a small portion of AD cases[52]. To evaluate whether ABCA7 LoF
245 transcriptomic effects in neurons would generalize to more common, moderate-risk genetic
246 variants in ABCA7, we examined the ROSMAP WGS cohort for carriers of the prevalent
247 ABCA7 missense variant p.Ala1527Gly (rs3752246: Minor Allele Frequency ≈ 0.18; % carriers
248 >= 1 allele ≈ 30%; **Fig. 2E**), which has risen to genome-wide significance for increased AD risk
249 (Odds Ratio = 1.15 [1.11-1.18])[7,15,19], but whose molecular effects are unexplored. We identified
250 133 individuals carrying at least one copy of the p.Ala1527Gly risk variant and 227 non-carriers
251 (**Fig. 2F**) for whom snRNAseq data of the *post-mortem* PFC were available[56]. We ensured that
252 none of these 360 individuals were part of our earlier ABCA7 LoF snRNAseq cohort as control
253 samples and that none carried ABCA7 LoF variants. Using this cohort, we investigated whether
254 excitatory neurons from p.Ala1527Gly carriers exhibited evidence of transcriptomic
255 perturbations in the ABCA7 LoF-associated clusters C0-C7.

256 Remarkably, all clusters displayed directional trends in p.Ala1527Gly neurons consistent with
257 the mean effect sizes observed in ABCA7 LoF neurons (**Fig. 2C,G**), while controlling for
258 pathology, age, sex, and other covariates (**Methods**). Notably, 4 out of 8 clusters exhibited
259 substantial evidence of perturbation in p.Ala1527Gly variant carriers, with perturbation
260 directions aligning with predictions for ABCA7 LoF (**Fig. 2G,H**). Specifically, we observed an
261 up-regulation in the DNA damage cluster C1 and the proteasomal cluster C3 in p.Ala1527Gly
262 carriers compared to controls, suggesting a similar cell stress and genomic instability signature to
263 ABCA7 LoF carriers (**Fig. 2G,H**), and a borderline significant up-regulation of the
264 mitochondrial cluster C7, again consistent with ABCA7 LoF (**Fig. 2G,H**). Intriguingly, the most

265 significantly perturbed cluster was the lipid cluster C2, which exhibited down-regulation (**Fig.**
266 **2G,H**) similar to our observations in ABCA7 LoF carriers.

267 Because a missense variant can alter the dynamic structure of a protein, and mutations to glycine
268 are known to introduce significantly greater local flexibility compared to alanine, we evaluated
269 whether the convergent transcriptional signature could be attributed to changes in ABCA7
270 protein structure. Leveraging recent advances in protein structure simulation and the newly
271 crystallized structures of ABCA7 in its closed (ATP-bound) (**Fig. 2I**) and open (ATP-unbound)
272 (**Fig. S6A**) conformations[57,58] , we conducted molecular dynamics simulations of a 239 residue
273 region with and without the p.Ala1527Gly substitutions over a 300ns simulation time (**Fig. 2J**;
274 **Fig. S6B; Methods; Supplementary Text**). These simulations revealed that the AD
275 risk-associated G1527 mutation in the ATP-bound closed form induced local flexibility,
276 evidencing large structural fluctuations over time, while the A1527 variant had a stabilizing
277 effect on the ATP-bound closed conformation with only minor structural fluctuations (**Fig.**
278 **2K-M**). Both variants displayed stable conformational behavior in the unbound-open
279 conformation (**Fig. S6B-F**). Since the closed ATP-bound conformation is thought to mediate
280 lipid extrusion to the plasma membrane and apolipoproteins[57], our data indicate that the
281 p.Ala1527Gly substitution significantly impacts local secondary structure stability and
282 conformational dynamics of ABCA7, potentially affecting its lipid extrusion function.

283 Together, our molecular dynamics simulations indicate that the p.Ala1527Gly substitution likely
284 increases AD risk by perturbing ABCA7 structure, and therefore function. As such, our findings
285 imply that rare and common ABCA7 LoF variants may mediate AD risk via similar,
286 ABCA7-dependent mechanisms and that findings from an in-depth study of ABCA7 LoF may
287 extend to larger portions of the at-risk population. Moreover, the pronounced lipid signature
288 associated with both LoF and the common missense variant in excitatory neurons, which aligns
289 with ABCA7's functional role as a lipid transporter, suggests a potential upstream role of lipid
290 disruptions in ABCA7-mediated risk.

291 *Lipid disruptions in ABCA7 LoF PFC*

292 Our findings revealed a notable downregulation of various genes and pathways involved in fatty
293 acid biosynthesis in ABCA7 LoF neurons from *postmortem* human PFC (**Fig. 1G,H; Fig. 2B-D;**
294 **Fig. S7A,B**), while pathways related to lipid breakdown, particularly OXPHOS involving the
295 tricarboxylic acid (TCA) cycle via lipid oxidation, were upregulated (**Fig. S7C,D**). We validated
296 a subset of these gene expression changes by RNAscope on *postmortem* human brain samples,
297 including down-regulation of *ACLY*, a converter of mitochondria-derived citrate into acetyl-CoA
298 for fatty acid synthesis, and up-regulation of *SCP2,* a lipid transfer protein involved in lipid
299 β-oxidation (**Fig. S7E,F**). These findings suggest that ABCA7 LoF may impact the balance
300 between lipid synthesis and catabolism through mitochondrial carbon flux.

301 Since ABCA7 is known to function as a lipid transporter with an affinity for phospholipids,
302 including phosphatidylcholines[24,25], we next performed unbiased mass spectrometry-based

lipidomics on a subset of *post-mortem* human brains from the snRNAseq cohort with and without ABCA7 LoF variants (N=8 individuals per group, matched on multiple covariates; **Data S7; Methods**) to explore the effects of ABCA7 LoF on lipid composition. We noted a modest decrease of multiple phospholipid species, including a significant decrease in cardiolipin and phosphatidylcholine ($p<0.05$) (**Fig. 3A**) and a trend towards increased abundance of triglycerides (**Fig. 3A).** Notably, phospholipids (decreased) and triglycerides (increased) share a common precursor, diacylglycerol (**Fig. 3A**).

Deriving human neurons with ABCA7 LoF variants

To complement the correlative analyses in ABCA7 LoF human tissue, we next used CRISPR-Cas9 genome editing to generate two isogenic iPSC lines, each homozygous for ABCA7 LoF variants, from a parental line without ABCA7 variants (WT). The first LoF variant, ABCA7 p.Glu50fs*3, was generated by a single base-pair insertion in ABCA7 exon 3, resulting in a PTC early in the ABCA7 gene (**Fig. 3B; Fig. S8A,C**). This 5'-proximal PTC limits the possibility of producing a truncated protein with partial functionality, generating a near-complete knockout. The second LoF variant, ABCA7 p.Tyr622*, was generated by a single base-pair mutation in ABCA7 exon 15 (**Fig. 3B; Fig. S7B,C**). This PTC re-creates a variant previously observed in patients as associated with AD [12] and is meant to provide clinical context to ABCA7 dysfunction.

We differentiated the isogenic iPSCs into neurons (iNs) through lentiviral delivery of a doxycycline-inducible NGN2 expression cassette as previously described[59] (**Fig. 3C; Fig. S8D**). At 2 and 4 weeks post-NGN2 induction, cells were immunoreactive for neuronal markers TUJ1 and MAP2 and showed robust neuronal processes by pan-axonal staining (**Fig. 3D; Fig. S8E)**. Both WT and ABCA7 LoF lines had the ability to fire action potentials upon current injections (**Fig. 3E,F; Fig. S8F-J**). Though the *ABCA7* genotype did not affect resting membrane potential (**Fig. S8H)**, ABCA7 LoF iNs fired more action potentials and with lesser magnitude of current injection than WT iNs (**Fig. S8I,J)**, indicating an ABCA7 LoF hyperexcitability phenotype.

Previous studies suggest that ABCA7 dysfunction affects amyloid processing[32–35,60]. To verify that this was the case in our ABCA7 LoF lines, we examined Aβ42 levels - considered the most toxic species[47–51] - and observed increased intracellular Aβ42 by immunohistochemistry in the p.Glu50fs*3 iNs ($p = 0.016$), and a similar trend of borderline significance in p.Tyr622* iNs ($p = 0.069$) when compared to control iNs (**Fig. S9A,B**). Together, these data (1) confirm successful neuronal differentiation from iPSCs, and (2) show that ABCA7 LoF iNs recapitulate key AD-associated phenotypes of hyperexcitability and amyloid pathology.

ABCA7 LoF impacts the neuronal lipidome

Lipidomics on *postmortem* brain tissue with and without ABCA7 LoF had suggested decreases in phospholipids and increases in triglycerides (**Fig. 3A**). To investigate the effects of ABCA7 LoF variants on lipid homeostasis specifically in neurons, we performed untargeted LC-MS

lipidomics on two independent iN differentiation batches containing the p.Glu50fs*3 ABCA7 mutation (N=3 p.Glu50fs*3 and N=3 WT for each batch), selecting the more severe mutation to maximize phenotypic effects. As lipidomic profiles across differentiation batches correlated well (**Fig. S10A,B**), we pooled samples from each batch to increase statistical power and reduce the impact of potential outliers. Overall, we observed that the lipid profiles of ABCA7 p.Glu50fs*3 iNs differed substantially from their WT counterparts and were linearly separable when projected on the second principal component (**Fig. 3G**). We observed strong evidence (p<5e-3) for perturbation of sphingolipids (specifically, ceramides; (Cer), phospholipids (including phosphatidylcholine (PC), cardiolipin (CL), phosphatidylserine (PS), and phosphatidylglycerol (PG)), and neutral lipids (specifically, triglycerides (TGs) and monoglycerides (MGs)), in ABCA7 p.Glu50fs*3 vs WT iNs (**Fig. 3H,I**).

Triglycerides were in absolute and relative terms the most frequently perturbed lipid species (N=31; 56% of all detected TG species; hypergeometric enrichment p-value = 3.9e-15)(**Data S8; Fig. 2I**). All differentially detected triglyceride species were up-regulated in ABCA7 LoF iNs (**Fig. 3I, J**), indicating increased levels of neutral lipid accumulation in these cells. While none of the remaining lipid classes were significantly enriched (hypergeometric p<0.05) among the differentially detected species (**Fig. 3I**), we did note that phosphatidylcholine species emerged as the second most frequently altered lipid class in absolute terms, including a total of 21 perturbed species (p<5e-3; 16% of all detected PC species)(**Fig. 3I, K**). These changes were followed by a smaller number of perturbations to other lipid species, such as ceramides (N=5 changes) and cardiolipin (N=2 changes). Together, these data indicate a striking triglyceride accumulation accompanied by bi-directional changes for different phosphatidylcholine species, as well as a number of changes to other phospho- and sphingolipid species.

ABCA7 LoF induces a compositional shift from phospatidylcholines towards triglycerides

Phosphatidylcholines, which represent a major constituent of cellular membranes and the phospholipid envelope of lipid droplets (LDs), and triglycerides, which store fatty acids inside of lipid droplets, share diglycerides (DGs) as a common precursor[24,25]. Importantly, diglycerides are not only converted into but can also be derived from phospholipids or triglycerides and then broken down into first monoglycerides and finally fatty acids for energy (**Fig. 3L**). Breakdown of triglycerides is believed to be impacted by the phosphatidylcholine to triglyceride (PC:TG) ratio, which affects lipid droplet size and lipase access to their triglyceride core[61–64]. Therefore, the relative abundance of phosphatidylcholines and triglycerides serves as an important indicator of metabolic state in the cell. We observed a dramatically altered PC:TG ratio (**Fig. 3M**) in the ABCA7 LoF iN that was accompanied by a ~80% increase (log2FC = 0.86) in the fraction of triglycerides relative to total lipid content (**Fig. 3N**) and a circa 81% decrease (log2FC = -2.3 in the fraction of monoglycerides (**Fig. 3O**). These changes suggest an ABCA7 LoF-induced metabolic shift away from triglyceride breakdown into monoglycerides and towards triglyceride accumulation. Furthermore, the steep increase in the relative triglyceride abundance together with a more moderate change in the relative diglyceride abundance (log2FC = 0.14; +10%) (**Fig.**

379 **3P**) and a decrease in the relative phosphatidylcholine abundance (log2FC = -0.65; -36%) (**Fig.**
380 **3Q**) suggest that diglycerides are less often converted into phosphatidylcholine and more often
381 into triglycerides in the presence of ABCA LoF (**Fig. 3R**). Notably, we observed a similar trend
382 related to relatively increased triglyceride and decreased phosphatidylcholine abundance in the
383 *postmortem* human PFC of ABCA7 LoF carriers (**Fig. S9C**).

384 *ABCA7 LoF alters the neuronal metabolome*

385 To understand if these lipid changes were associated with global perturbations to the neuronal
386 metabolome in ABCA7 LoF iNs, we performed untargeted metabolomic analysis on the aqueous
387 fraction acquired during lipidomics preparations. Again, we observed consistency between
388 batches and pooled the results for further analysis (**Fig. S10D**). Projection of metabolite levels
389 onto the first principal component completely separated samples by genotype (**Fig. 4A**),
390 suggesting that the metabolomes between WT and ABCA7 LoF iNs differ substantially.

391 We observed 10 increased and 49 decreased metabolites in ABCA7 LoF vs WT iN (p-value <
392 0.05 & |log2FC|>1; **Data S9; Fig. 4B**). Nine of these differentially abundant metabolites could
393 be annotated with high confidence and were not detected in the background media (**Methods**).
394 Annotated species with differential abundance included a number of carnitine species,
395 glutamine-glutamate, inosine-inositol, and hypoxanthine, all of which were less abundant in
396 ABCA7 LoF vs WT iNs (**Fig. 4B**). Together, these data implicate a number of perturbed
397 metabolic processes in ABCA7 LoF including lipid catabolism, TCA cycle, and OXPHOS (*via*
398 carnitine[65], glutamine-glutamate[66,67]), cellular redox state (*via* inosine-inositol[68,69],
399 hypoxanthine[70,71]), and excitatory synaptic signaling (*via* glutamine-glutamate[72–74] and
400 carnitine[75–77]).

401 *ABCA7 LoF decreases carnitine abundance*

402 We were particularly intrigued by the down-regulation of carnitine and its esterified forms,
403 acetylcarnitine and propionylcarnitine, which were consistently and strongly reduced in
404 p.Glufs50*3 vs WT iNs (log2FC -1.35, -1.32, -3.55,, respectively) (**Fig. 4B, C**). Carnitine and
405 its derivatives mediate the rate-limiting transfer of medium and long chain free fatty acids across
406 the mitochondrial membrane for β-oxidation [78,79]. Reduced amounts of carnitines in the ABCA7
407 LoF iN therefore suggest a decrease in neuronal β-oxidation of lipids relative to WT iN. While
408 the primary substrates for neuronal OXPHOS are lactate and glucose[80–82], a recent study indicates
409 that up to 20% of basal neuronal OXPHOS may be fueled by lipid β-oxidation[83] Indeed, our
410 snRNAseq data from *postmortem* human PFC show that excitatory neurons express several
411 regulators of β-oxidation at detectable levels (>10%), including *HADHA,* which catalyzes the
412 ultimate steps in β-oxidation, and carnitine palmitoyltransferases (*CPT1C, CPT1B*), which
413 mediate the carnitine-dependent transport of fatty acids into mitochondria for oxidation (**Fig.**
414 **S10E**). We also observed a significant positive correlation between carnitine abundance and
415 monoglycerides, an intermediate in the hydrolysis of triglycerides into free fatty acids, in both
416 WT and ABCA7 LoF iNs (**Fig. 4D**). Together with the transcriptional changes to

417 mitochondrial-related genes (**Fig. 2C; Fig. S7C,D**), the observed decrease in carnitines points to
418 altered mitochondrial bioenergetics in ABCA7 LoF neurons and suggests a possible connection
419 to the lipid disruptions detected in these cells.

420 *ABCA7 LoF impairs mitochondrial uncoupling in neurons*

421 To directly assess mitochondrial function in ABCA7 LoF neurons, we next measured the oxygen
422 consumption rate (OCR) of WT and ABCA7 LoF iNs over time using the live-cell Seahorse
423 metabolic flux assay (Agilent). The mitochondrial membrane potential (ΔΨm) is maintained via
424 the oxygen-dependent movement of protons across the inner mitochondrial membrane during
425 OXPHOS and dissipated through the membrane-bound ATP synthase and by regulated proton
426 leakage along the gradient (**Fig. 4E**). The OCR, in the presence of various mitochondrial
427 inhibitors, therefore allows several readouts of mitochondrial function[84].

428 Since measures of oxygen consumption are sensitive to differences in viable cell number and
429 mitochondrial abundance[84,85], parameters which may be affected by ABCA7 LoF mutations, we
430 report internally-normalized measures of mitochondrial function derived from the OCR curves[86]
431 of WT, ABCA7 p.Glu50fs*30, and ABCA7 p.Tyr622* iNs (**Fig. S11A,B**). First, we report the
432 measured increase in oxygen consumption upon pharmacologically collapsing the proton
433 gradient, which indicates the cellular spare respiratory capacity [86]**(Fig. S11C).** Secondly, we
434 report the proportion of basal oxygen consumption attributed to ATP synthesis *vs* proton leak
435 across the membrane (*i.e. "relative uncoupling";* colored panels in **Fig. 4F**) [84].

436 While we did not observe a significant difference in spare respiratory capacity between WT and
437 ABCA7 LoF iNs (**Fig. S11D**), we did observe a substantial decrease in relative uncoupling in
438 ABCA7 LoF *vs* WT iNs (**Fig. 4G**; **Fig. S11E**). Levels of relative uncoupling observed in our WT
439 iNs (*circa* 20%) (**Fig. 4G**) are in line with previous reports under basal conditions in neurons and
440 other cell types[87,88], indicating that ABCA7 LoF neurons have abnormally low levels of
441 uncoupling. Impaired mitochondrial uncoupling is generally associated with a hyperpolarized
442 mitochondrial membrane potential, ΔΨm[89,90]. To verify that this was also the case in ABCA7
443 LoF iNs, we incubated iNs with a fluorescent dye that accumulates in mitochondria
444 proportionally to ΔΨm. Indeed, we observed an increase in ΔΨm assessed by mean fluorescence
445 intensity per NeuN+ surface in ABCA7 LoF vs WT iNs (**Fig. 4H; Fig. S11F**).

446 Cells actively control their ΔΨm through expression of mitochondrial uncoupling proteins[91,92],
447 which limit the formation of damaging reactive oxygen species induced by high ΔΨm[93] (**Fig.**
448 **4E**). Free fatty acids, in particular, are known to promote mitochondrial uncoupling[94–96], perhaps
449 allowing cells to safely metabolize lipids through mitochondrial β-oxidation. Together, these data
450 suggest that ABCA7 LoF neurons may have decreased ability to dissipate their ΔΨm and
451 maintain basal levels of carbon oxidation via regulated uncoupling.

452

Treatment with CDP-choline rescues lipid and mitochondrial defects

Our data indicate a relative metabolic shift away from phosphatidylcholines towards triglycerides in ABCA7 LoF neurons, accompanied by changes in mitochondrial metabolites and function. Decreasing phosphatidylcholine synthesis can indirectly increase triglyceride levels, given their common precursor[97]. In addition, deficient phosphatidylcholine at the mitochondrial membrane may directly impair mitochondrial lipid breakdown, perhaps further compounding the triglyceride accumulation[98–101]. Therefore, enhancing phosphatidylcholine synthesis in ABCA7 LoF neurons may rescue triglyceride accumulation and mitochondrial function in these neurons.

To test this, we treated ABCA7 p.Tyr622* iNs with CDP-choline, the rate-limiting precursor for phosphatidylcholine synthesis through the Kennedy pathway [102,103], for 2 weeks and then evaluated neutral lipid accumulation and mitochondrial bioenergetics. Because triglycerides are stored primarily in the form of neutral lipid droplets[104], we measured neutral lipid accumulation by staining with a lipid-droplet specific dye, LipidSpot. Fluorescence intensity of LipidSpot and an antibody against perilipin 2 (PLIN2), a protein embedded in lipid-droplet envelopes[105], correlated well, indicating that the LipidSpot stain labels *bonafide* lipid droplets in iNs (**Fig. 5A; Fig. S12A**). Treatment of ABCA7 p.Tyr622* iNs with CDP-choline reduced the neutral lipid burden per NeuN+ surface, as measured by LipidSpot fluorescence intensity, compared to vehicle treatment (**Fig. 5A; Fig. S12B).**

By measuring the oxygen consumption rate in ABCA7 p.Tyr622* iNs with and without CDP-choline, we found that CDP-choline increased relative uncoupling in p.Tyr622* iNs back to WT levels (**Fig. 5B; Fig. S12C,D**), with no significant effect on SRC (**Fig. S12E**). In line with this increased uncoupling, CDP-choline supplementation decreased $\Delta\Psi m$ in p.Tyr622* iNs *vs* vehicle treated cells, as shown by reduced MitoHealth dye accumulation (**Fig. 5C**). These data indicate that supplementation with CDP-choline rescued both triglyceride accumulation and re-established control of the mitochondrial membrane potential.

Treatment with CDP-choline decreases amyloid-β production

Changes in membrane phospholipid content influence membrane properties including curvature and fluidity[106]. Because changes to membrane fluidity are important regulators of amyloid precursor protein (APP) processing[107,108], we wondered whether phosphatidylcholine defects in ABCA7 LoF iNs were also responsible for increased levels of Aβ42 processing described earlier (**Fig. S9A,B**) and in previous literature[32–35]. To test this, we measured intracellular Aβ42 levels in p.Tyr622* iNs treated with CDP-choline. Indeed we found a significant reduction in Aβ42 fluorescence per NeuN+ volume after treatment with CDP-choline (**Fig. 5D**). Other studies suggest that the activity of secretases, which cleave APP to produce amyloid, is also modulated by membrane fluidity or by phospholipid binding[109–111]. CDP-choline treatment also decreased fluorescence intensity of the β-amyloid secretase enzyme BACE (**Fig. 5E**), indicating reduced amyloidogenic processing in treated p.Tyr622* iNs. Together, our data suggest that treatment

with CDP-choline rescues increased amyloid processing in ABCA7-LoF iNs, linking the aforementioned ABCA7 LoF-induced metabolic defects in neurons (**Fig. 5F**) to AD pathology.

Discussion

Loss-of-function (LoF) mutations in the lipid transporter ABCA7 are among the strongest genetic risk factors for late-onset AD. Here, we generated a transcriptional atlas of ABCA7 LoF effects across all major brain cell types in the human prefrontal cortex. Our dataset showed the highest levels of ABCA7 expression in excitatory neurons and strong evidence that ABCA7 LoF led to transcriptional perturbation in pathways related to lipid biosynthesis, mitochondrial respiration, and cellular stress, including up-regulation of DNA repair pathways, and changes to inflammatory and synaptic genes. Using iPSC-derived isogenic neuronal lines with and without ABCA7 LoF variants, we show that ABCA7 LoF leads to intracellular triglyceride accumulation, decreased phosphatidylcholine levels, and a reduction in mitochondrial uncoupling.

Multiple lines of evidence highlight a causal relationship between lipid metabolism and mitochondrial uncoupling[101,112–115]. While neurons rely on astrocytes to help metabolize excess lipids[116,117118,119], neurons can also perform endogenous lipid breakdown through β-oxidation[120,121]. Rates of β-oxidation are tightly linked to mitochondrial uncoupling; Cells that specialize in fatty acid oxidation have among the highest uncoupling ratios (e.g. Brown adipose tissue)[122–124] and pharmacologically stimulating mitochondrial uncoupling increases lipid catabolism[101,112–114]. We find that treatment with CDP-choline, a phosphatidylcholine precursor that bypasses the rate limiting step in phosphatidylcholine synthesis through the Kennedy pathway[125], reverts uncoupling to normal levels and reverses triglyceride accumulation, suggesting that in addition to increasing the availability of triglyceride precursors, phosphatidylcholine deficiency may exacerbate triglyceride accumulation by impairing mitochondrial lipid catabolism in these cells (**Fig. 5G**). Further studies will be needed to distinguish if increased phosphatidylcholine synthesis restores mitochondrial uncoupling directly by modifying the mitochondrial membrane, in line with a recent study [101], or indirectly by increasing the release of free fatty acids from lipid droplets[61–64].

While mitochondrial dysfunction is associated with aging, AD, and other neurodegenerative diseases - mitochondrial uncoupling was recently linked to frontotemporal dementia variants[126] -, little is known about the role of mitochondrial uncoupling in AD etiology [89,127–130]. Neurons maintain very high levels of mitochondrial OXPHOS to meet their energy demands[81,83] . Mitochondrial uncoupling, which is actively regulated by mitochondrial proteins[131,132], may help sustain this high aerobic capacity by actively controlling the mitochondrial membrane potential, managing associated levels of reactive oxygen species[89,133], and promoting mitochondrial biogenesis[134–137]. Impaired uncoupling can be neurotoxic[134–137], for example, by impairing synaptic and calcium signaling, or increasing oxidative stress, a phenotype observed recently in ABCA7 LoF neurospheroids[30]. Oxidative stress has multiple toxic downstream consequences, including the induction of DNA damage and inflammation, as suggested by ABCA7 LoF

transcriptomic signatures[138–141]. Neurons afflicted by these phenotypes also impact surrounding glial cells [54,142].

In line with our findings linking phosphatidylcholine-triglyceride imbalances to mitochondrial impairments in ABCA7 LoF neurons, a recent study in ABCA7 LoF neurospheroids independently revealed a link between phosphatidylglycerol deficiency, observed simultaneously with reduced phosphatidylcholine, and mitochondrial function[30], further highlighting the importance of lipid-centric therapeutic interventions for ABCA7 LoF. Here, we offer a therapeutic strategy to reverse these dysfunctions through CDP-choline treatment, a readily available and safe dietary supplement[102,143,144]. CDP-choline treatment also reduces toxic Aβ42 levels in ABCA7 LoF neurons, indicating that lipid metabolic defects may contribute to Alzheimer's disease pathology in ABCA7 LoF neurons. This is consistent with previous studies linking membrane lipids and lipid modifications to APP processing [109,145].

Recent work from our lab implicates choline in APOE4 dysfunction[146], and in cognitive resilience to AD pathology (Boix *et al*, 2024, *Nature*, in press), suggesting that phosphatidylcholine disruptions may be central to AD risk in large fractions of the population. Indeed, our work suggests that the common missense variant p.Ala1527Gly likely has convergent effects with ABCA7 LoF. Genetic interactions with other risk factors, including APOE4, may exacerbate otherwise subtle ABCA7 dysfunction, and contribute to risk in a significant subset of AD cases[147–150]. As such, our study supports a growing body of literature, including recent studies on APOE4 [151,152], implicating lipid disruptions in the etiology of AD and pinpoints additional genotypes that may benefit from phosphatidylcholine and triglyceride-targeting interventions.

Acknowledgments: We thank the individuals who donated *postmortem* brain samples, and their families, for enabling this research; Y. Zhou, E. McNamara and T. Garvey for administrative support and animal care; U. Geigenmüller for reviewing and editing the manuscript; J. Davila-Velderrain and R. Firenze for providing input on the manuscript; the MIT SuperCloud and Lincoln Laboratory Supercomputing Center for providing HPC and consultation resources that have contributed to the research results reported within this paper; the Harvard Center for Mass Spectrometry (HCMS) for lipidomic sample runs and Charles Vidoudez (HCMS) for consultation on the lipidomic data analysis. ROSMAP resources can be requested at https://www.radc.rush.edu. Graphic illustrations were generated using BioRender under agreement # X25SQ61W7.

Funding:

This work was supported in part by the Cure Alzheimer's Fund, the JBP Foundation and NIH grants R56-AG081376, RF1-AG062377, RF1-AG054321, RO1-AG054012 (L.-H.T.), and RO1-AG075901 (P.I. Ernest Fraenkel).

Author contributions:

Conceptualization: DVM, JMB, L-HT, SEW

Data curation: DVM, JMB, SEW, HM, CAB

Formal analysis: DVM, AS

Methodology: DVM, JMB, SEW, HM, GL, AS, LL

Investigation: DVM, JMB, SEW, CS, HM, ANS, GS, GSM, GW, NL, AS, LL, P-CP, CJY, ML, AG

Visualization: DVM, SEW, AS

Funding acquisition: L-HT, MK

Writing – original draft: DVM, SEW

Writing – review & editing: DVM, SEW, JMB, L-HT, HM

Competing interests: Authors declare that they have no competing interests.

Ethics Statement: The study protocol involving the use of human stem cells was approved by the Coriell Institutional Review Board (Coriell IRB) in compliance with DHHS regulations (45 CFR Part 46). The initial cell lines were obtained from the Coriell Institute, which ensured that informed consent was received from all donors. Donors were informed that their tissue donations would be used for the creation of cell lines intended for educational and research purposes, and that all biological materials would be anonymized. More information can be found here. For *postmortem* human brain samples, informed consent was obtained from each subject, and the Religious Orders Study and Rush Memory and Aging Project were approved by an Institutional Review Board (IRB) of Rush University Medical Center.

Data availability: All *postmortem* human data can be accessed through the Synapse AD Knowledge Portal (syn53461705), which also includes associated ROSMAP metadata. These data are subject to controlled access in compliance with human privacy regulations. To obtain the data, a data use agreement (DUA) must be completed. This requirement ensures the anonymity of ROSMAP study participants. A DUA can be established with either the Rush University Medical Center (RUMC) or SAGE, the organization that manages Synapse. The necessary forms are available for download on their respective websites. All iPSC-related data are accessible through links provided in our code repositories. For a complete list of data availability and download links, please refer to the code repositories listed below. Additionally, relevant processed datasets are available in the supplementary files of this manuscript.

Code availability: All code used in this study is available on GitHub. This includes code to replicate the analyses and figure panels presented in the paper (https://github.com/djunamay/ABCA7lof2), descriptions and code for accessing whole-genome sequencing data (https://github.com/djunamay/ROSMAPwgs), code for performing gene-pathway clustering (https://github.com/djunamay/geneclusters), and code for processing confocal images (https://github.com/djunamay/confocalQuant).

Fig. 1. Single-nuclear RNA-sequencing Atlas of Human Post-mortem Prefrontal Cortex Reveals Cell Type-specific Gene Changes in ABCA7 LoF. (A) Overview of ABCA7 gene structure with the location of variants represented in this study (average minor allele frequency for depicted variants is < 1%). Exons are depicted as black rectangles, and introns as black lines. The pie chart indicates the frequency of ABCA7 PTC-variant-carriers within the ROSMAP cohort. (B) ABCA7 protein levels (log2(abundance)) from post-mortem human prefrontal cortex in all available controls ($N = 180$) vs. ABCA7 LoF carriers ($N = 5$). P-value computed by Wilcoxon rank sum test. Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). (C) Overview of human cohort for snRNA-seq (Created with BioRender.com). (D) Overview of snRNA-seq cohort metadata for 32 individuals. (E) 2D UMAP projection of per-cell gene expression values and their transcriptionally defined cell type. (F) 2D UMAP projection of ABCA7 LoF gene perturbation scores ($S = -\log_{10}(\text{p-value}) \times \text{sign}(\log_2(\text{fold change}))$); Red = $S > 1.3$, Blue = $S < -1.3$; Point size indicates $|S|$). Up to top 20 genes by $|S|$ are labeled. (G) Genes in 2D UMAP space colored by cluster assignment (Gaussian mixture model; see Methods) with per-cluster pathway enrichments shown (GO BP, hypergeometric enrichment, $p < 0.01$). (H) Cell type-specific gene cluster scores ($SC = \text{mean}(S_i)$, for genes i in cluster c). * indicates permutation FDR-adjusted p-value < 0.01 and $|SC| > 0.25$.

Fig. 2. Transcriptional Perturbations in Excitatory Neurons in ABCA7 LoF and ABCA7 p.Ala1527Gly Variant Carriers. (A) 2D UMAP projection of individual cells colored by log(Exp), where Exp represents log-normalized ABCA7 expression values. (B) Kernighan-Lin (K/L) clustering on leading edge genes from pathways perturbed in ABCA7 LoF excitatory neurons, where $p < 0.05$. Colors indicate distinct K/L gene clusters, which are numbered from 0 to 7. (C) Gaussian kernel density estimate plots of gene scores S for genes belonging to a given gene cluster. $S > 0$ indicates upregulation in ABCA7 LoF. Solid lines indicate distribution means. (D) Representative pathways that annotate the largest number of genes within a cluster (i.e., with the highest intra-cluster connectivity) shown per-cluster. (E) Schematic indicating the genomic location of the p.Ala1527Gly codon change. A purple arrow indicates the location of the missense variant in the ABCA7 gene. Minor allele frequency shown to the right.

(F) Overview of snRNA-seq cohort of ABCA7 p.Ala1527Gly carriers (homozygous and heterozygous) vs. control non-carriers (minor allele frequency approx. 18%). (G) Perturbation of ABCA7 LoF-associated gene clusters from (B-D) in excitatory neurons of p.Ala1527Gly variant-carriers vs. non-carrier controls, computed by FGSEA. Top p-values ($p < 0.1$) are indicated. $S > 0$ indicates upregulation in carriers. (H) Distributions of gene scores S for genes belonging to a given gene cluster for ABCA7 p.Ala1527Gly (no fill) or ABCA7 LoF-variants (solid fill). $S > 0$ indicates upregulation in ABCA7 variant. * indicates FGSEA p-value<0.1 from (G). Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). (I) Closed conformation ABCA7 protein structure. ABCA7 domain between residues 1517 and 1756 used for simulations is shown in yellow. Lipid bilayer shown in orange. (J) Expanded yellow domain (inset from *I*), with A1527 variant (light grey) and G1527 variant (purple). (K) Expanded inset from *J* with residues of interest rendered. (L) Root mean squared deviations of closed conformation domains from *J* with A1527 (light grey) or G1527 (purple) under simulation. Structural deviations over time were computed with respect to reference closed conformation from *J*. Violin plot inset indicates average C_α atom positional fluctuations over time. (M) Projection of C_α atom positional fluctuations under simulation onto the first two principal components, for closed conformation domain from *J* with A1527 (top, light grey) or G1527 (bottom, purple).

Fig. 3. ABCA7 LoF Neurons Have Neutral Lipid Accumulation and Phospholipid Imbalances. (A) Volcano plot showing the distribution of lipid species by log-fold-change and log-p-value, where log-fold-change > 0 indicates up-regulation in ABCA7 LoF post-mortem PFC (N=8) vs. Control (N=8). A subset of lipid species is labeled ($p < 0.1$ & logFC>0.5 (red) or logFC<-0.5 (blue)) (p-values by independent sample t-test). (B) Overview of two iPSC-derived isogenic neuronal lines carrying ABCA7 LoF variants. ABCA7 gene map depicts exons (black rectangles) and introns (black lines). (C) CRISPR-Cas9 was used to generate an ABCA7 LoF isogenic iPSC line by introducing a premature termination codon into exon 3 or exon 15 (ABCA7 p.Glu50fs*3, blue or p.Tyr622*, orange, respectively). (D) Example MAP2 staining of 2-week-old p.Tyr622* iNs. (E) Representative sweeps show action potentials elicited by 800 ms of current injections in patched 4-week-old iNs. (F) Summary of action potential frequency (means ± SEM) elicited with different amounts of injected current in 4-week-old iNs. (G) Projection of control and ABCA7 p.Glu50fs*3 lipidomes (per-sample z-scaled peak areas; $N = 6$ per genotype) onto the first two principal components from lipid space. Fraction of explained variance shown along each axis. (H) Volcano plot showing perturbed lipid species by class (p-values by independent sample t-test). (I) Overview of perturbed lipid species by lipid subclass. $U =$ number of lipids in that subclass where $p < 5 \times 10^{-3}$ & fold-change > 0. $D =$ number of lipids in that subclass where $p < 5 \times 10^{-3}$ & fold-change < 0. $T = U + D$. % = (T/N)$\times 100$, where $N =$ total number of lipids in a given subclass (p-values by independent sample t-test; $N = 6$ per genotype). (J,K) Volcano plot showing the distribution of triglyceride

(TG) species (J) and phosphatidylcholine (PC) species (K) by log-fold-change and log-p-value, where log-fold-change > 0 indicates up-regulation in p.Glu50fs*3 ($N = 6$) vs. WT ($N = 6$) iPSC-derived neurons. A subset of lipid species is labeled ($p < 0.05$ & logFC>2 (red) or logFC<-2 (blue)) (p-values by independent sample t-test). (L) Schematic of metabolic link between TG, monoglyceride (MG), and diglyceride (DG). (M-Q) Relative abundance of select lipid species in p.Glu50fs*3 iN ($N = 6$) and WT iN ($N = 6$) (p-values by independent sample t-test). Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). (R) Schematic from (L) overlaid with abundance changes in p.Glu50fs*3 iNs vs. WT iNs. Schematics in (C, L, R) created with BioRender.com.

Fig. 4. ABCA7 LoF Impairs Regulation of Mitochondrial Uncoupling and Carbon Flux in Neurons. (A) Projection of WT and p.Glu50fs*3 metabolomes (per-sample z-scaled normalized peak areas; $N = 6$ per genotype) onto the first two principal components from metabolite space. Fraction of explained variance shown along each axis. The dotted red line separating the two genotypes indicates the median PC1 value. (B) Volcano plot indicating differentially regulated metabolites by log2-fold-change and log10(p-value), where log2(FC) > 0 indicates increased abundance in p.Glu50fs*3 vs. WT. Top up- and down-regulated metabolites are shown in red and blue, respectively (p-values by independent sample t-test). (C) Abundance of carnitine species in WT vs. p.Glu50fs*3 iNs. (D) Correlation of carnitine species abundance (normalized peak areas by LC-MS) and monoglyceride (MG) abundance (peak areas by LC-MS) from matched metabolomic-lipidomic samples. Grey points indicate WT iNs ($N = 6$). Blue points indicate p.Glu50fs*3 iNs ($N = 6$). The equation indicates the linear function fit. Grey error bar indicates 95% confidence interval. (E) Schematic indicating the relationship between oxygen consumption as a measure of proton current (I), which sustains the proton motive force (Δp; voltage (V)). Regulation of ATP synthase and uncoupling protein (UCP) activity modifies resistance (R) and depletes Δp. (F) Schematic indicating how relative uncoupling is computed from oxygen consumption rate (OCR). Remaining oxygen consumption after pharmacological inhibition of ATP synthase gives the proportion of basal oxygen consumption attributed to proton leak. (G) Relative uncoupling quantified by Seahorse oxygen consumption assay (see Methods) in 4-week-old WT vs. ABCA7 LoF iNs. P-values computed by independent sample t-test. N wells = 18 (WT), 17 (p.Tyr622*), 13 (p.Glu50fs*3) across two independent differentiation batches and Seahorse experiments (see Fig. S11E). (H) Left: Quantification of neuronal HCS MitoHealth dye fluorescence intensity as a measure of mitochondrial membrane potential. P-values computed by linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 8$ (WT), 11 (p.Tyr622*), 9 wells (p.Glu50fs*3) (3000 cells per condition) from three independent differentiation batches (see Fig. S11F). Individual data points indicate per-well averages of cell-level intensities. Right: Representative images per condition as mean-intensity projections of the entire image (NeuN+) and within NeuN+ volumes considered for quantification (MitoHealth, Hoechst). Representative images for the MitoHealth channel were processed with condition-wide percentile-based background subtraction and

thresholding. Representative images of cell soma underwent per-image percentile-based background subtraction and thresholding, reflecting the segmentation methodology. For (C, G, H) boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile).

Fig. 5. Supplementation with CDP-choline Reduces Neutral Lipid Accumulation, Restores Mitochondrial Function, and Reduces Amyloid Pathology in ABCA7 LoF Neurons. (A) Left: Representative images per condition as mean-intensity projections of the entire image (NeuN+) and within NeuN+ volumes considered for quantification (LipidSpot, PLIN2). Right: Quantification of neuronal LipidSpot dye fluorescence intensity. P-values computed by linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 16$ (p.Tyr622* + H2O), 21 wells (p.Tyr622* + CDP-choline) (5419 cells per condition) from three independent differentiation batches of 4-week-old iNs treated for 2 weeks (see Fig. S12B). (B) Relative mitochondrial uncoupling, quantified by Seahorse oxygen consumption assay for 4-week-old p.Tyr622* iNs treated with CDP-choline or vehicle control for 2 weeks. Relative uncoupling gives the proportion of basal oxygen consumption attributed to uncoupled proton leak. P-values computed by independent sample t-test. N wells = 6 (p.Tyr622* + H2O), 8 (p.Tyr622* + CDP-choline). (C) Left: Representative images per condition as mean-intensity projections of the entire image (NeuN+) and within NeuN+ volumes considered for quantification (MitoHealth, Hoechst). Right: Quantification of neuronal HCS MitoHealth dye fluorescence intensity as a measure of mitochondrial membrane potential. P-values computed by linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 11$ (p.Tyr622* + H2O; datapoints from Fig. 4H), 12 wells (p.Tyr622* + CDP-choline) (3929 cells per condition) from three independent differentiation batches of 4-week-old iNs treated for 2 weeks (see Fig. S12F). (D) Left: Representative images per condition as mean-intensity projections of the entire image (NeuN+) and projections within NeuN+ volumes considered for quantification (Aβ42). Right: Quantification of neuronal Aβ42 fluorescence intensity. P-values computed by linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 8$ (p.Tyr622* + H2O; datapointes from Fig. S9A) (1466 cells), 7 wells (p.Tyr622* + CDP-choline) (1102 cells) from 4-week-old iNs treated for 2 weeks. (E) Left: Representative images per condition as mean-intensity projections of the entire image (NeuN+) and within NeuN+ volumes considered for quantification (β-secretase). Right: Quantification of neuronal β-secretase fluorescence intensity. P-values computed by linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 4$ (p.Tyr622* + H2O) (3107 cells), 4 wells (p.Tyr622* + CDP-choline) (2829 cells) from 4-week-old iNs treated for 2 weeks. (F) Proposed model of ABCA7 dysfunction in neurons: ABCA7 loss-of-function (LoF) induces a compositional shift away from phosphatidylcholines (PCs) (1,2), which may directly increase triglycerides (TGs) by increasing precursor availability (3, 4) or indirectly by impairing lipid droplet processing (5) and mitochondrial uncoupling (6). Impaired uncoupling can lead to multiple detrimental cellular effects, including a decreased ability to meet the cell's

energy demand, impaired regulation of carbon breakdown via oxidative phosphorylation (OXPHOS), and increased oxidative stress (7). Abbreviations: DG = diglyceride, TG = triglyceride, MG = monoglyceride, PC = phosphatidylcholine, ROS = reactive oxygen species. For (A-E), + = treated with CDP-choline. - = treated with vehicle control. Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). For (A, C-E), individual data points indicate per-well averages of cell-level intensities. Representative images for the quantified channel were processed with condition-wide percentile-based background subtraction and thresholding. Representative images of cell soma underwent per-image percentile-based background subtraction and thresholding, reflecting the segmentation methodology. Schematic in (F) created with Biorender.com.

1. Alzheimer's Disease International, Guerchet, M., Prince, M. & Prina, M. Numbers of people with dementia worldwide: An update to the estimates in the World Alzheimer Report 2015. (2020).
2. Alzheimer's Association. 2016 Alzheimer's disease facts and figures. *Wiley Online Library* https://doi.org/10.1016/j.jalz.2016.03.001 (2016) doi:10.1016/j.jalz.2016.03.001.
3. Karlsson, I. K. *et al.* Measuring heritable contributions to Alzheimer's disease: polygenic risk score analysis with twins. *Brain Commun* **4**, fcab308 (2022).
4. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).
5. Marioni, R. E. *et al.* Correction: GWAS on family history of Alzheimer's disease. *Transl. Psychiatry* **9**, 161 (2019).
6. Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019).
7. Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019).
8. de Rojas, I. *et al.* Common variants in Alzheimer's disease and risk stratification by polygenic risk scores. *Nat. Commun.* **12**, 3417 (2021).
9. Wightman, D. P. *et al.* A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat. Genet.* **53**, 1276–1282 (2021).
10. Bellenguez, C. *et al.* New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat. Genet.* **54**, 412–436 (2022).
11. Belloy, M. E. *et al.* APOE Genotype and Alzheimer Disease Risk Across Age, Sex, and Population Ancestry. *JAMA neurology* vol. 80 1284–1294 (2023).
12. Steinberg, S. *et al.* Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. *Nat. Genet.* **47**, 445–447 (2015).
13. De Roeck, A., Van Broeckhoven, C. & Sleegers, K. The role of ABCA7 in Alzheimer's disease: evidence from genomics, transcriptomics and methylomics. *Acta Neuropathol.* **138**, 201–220 (2019).
14. Reitz, C. *et al.* Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E ϵ4,and the risk of late-onset Alzheimer disease in African Americans. *JAMA* **309**, 1483–1492 (2013).
15. Holstege, H. *et al.* Exome sequencing identifies rare damaging variants in ATP8B4 and ABCA1 as risk factors for Alzheimer's disease. *Nat. Genet.* **54**, 1786–1794 (2022).
16. Lyssenko, N. N. & Praticò, D. ABCA7 and the altered lipidostasis hypothesis of Alzheimer's disease. *Alzheimers. Dement.* **17**, 164–174 (2021).
17. Le Guennec, K. *et al.* ABCA7 rare variants and Alzheimer disease risk. *Neurology* **86**, 2134–2137 (2016).
18. Hollingworth, P. *et al.* Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat. Genet.* **43**, 429–435 (2011).
19. Naj, A. C. *et al.* Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat. Genet.* **43**, 436–441 (2011).
20. Kim, W. S., Weickert, C. S. & Garner, B. Role of ATP-binding cassette transporters in brain lipid transport and neurological disease. *J. Neurochem.* **104**, 1145–1166 (2008).
21. Koldamova, R., Fitz, N. F. & Lefterov, I. ATP-binding cassette transporter A1: from metabolism to neurodegeneration. *Neurobiol. Dis.* **72 Pt A**, 13–21 (2014).
22. Abe-Dohmae, S. *et al.* Human ABCA7 supports apolipoprotein-mediated release of cellular cholesterol and phospholipid to generate high density lipoprotein. *J. Biol. Chem.* **279**, 604–611 (2004).
23. Wang, N. *et al.* ATP-binding cassette transporter A7 (ABCA7) binds apolipoprotein A-I and mediates cellular phospholipid but not cholesterol efflux. *J. Biol. Chem.* **278**, 42906–42912 (2003).
24. Tomioka, M. *et al.* Lysophosphatidylcholine export by human ABCA7. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1862**, 658–665 (2017).
25. Picataggi, A. *et al.* Specificity of ABCA7-mediated cell lipid efflux. *Biochim. Biophys. Acta Mol.*

Cell Biol. Lipids **1867**, 159157 (2022).
26. Quazi, F. & Molday, R. S. Differential phospholipid substrates and directional transport by ATP-binding cassette proteins ABCA1, ABCA7, and ABCA4 and disease-causing mutants. *J. Biol. Chem.* **288**, 34414–34426 (2013).
27. Aikawa, T., Holm, M.-L. & Kanekiyo, T. ABCA7 and Pathogenic Pathways of Alzheimer's Disease. *Brain Sci* **8**, (2018).
28. Tanaka, N., Abe-Dohmae, S., Iwamoto, N. & Yokoyama, S. Roles of ATP-binding cassette transporter A7 in cholesterol homeostasis and host defense system. *J. Atheroscler. Thromb.* **18**, 274–281 (2011).
29. Duchateau, L. *et al.* CSF biomarker analysis of ABCA7 mutation carriers suggests altered APP processing and reduced inflammatory response. *Alzheimers. Res. Ther.* **15**, 195 (2023).
30. Kawatani, K. *et al.* ABCA7 deficiency causes neuronal dysregulation by altering mitochondrial lipid metabolism. *Mol. Psychiatry* 1–11 (2023).
31. Tayran, H. *et al.* ABCA7-dependent Neuropeptide-Y signalling is a resilience mechanism required for synaptic integrity in Alzheimer's disease. *bioRxiv* (2024) doi:10.1101/2024.01.02.573893.
32. Satoh, K., Abe-Dohmae, S., Yokoyama, S., St George-Hyslop, P. & Fraser, P. E. ATP-binding cassette transporter A7 (ABCA7) loss of function alters Alzheimer amyloid processing. *J. Biol. Chem.* **290**, 24152–24165 (2015).
33. Sakae, N. *et al.* ABCA7 Deficiency Accelerates Amyloid-β Generation and Alzheimer's Neuronal Pathology. *J. Neurosci.* **36**, 3848–3859 (2016).
34. Chan, S. L. *et al.* ATP-binding cassette transporter A7 regulates processing of amyloid precursor protein in vitro. *J. Neurochem.* **106**, 793–804 (2008).
35. Bamji-Mirza, M. *et al.* Genetic Variations in ABCA7 Can Increase Secreted Levels of Amyloid-β40 and Amyloid-β42 Peptides and ABCA7 Transcription in Cell Culture Models. *J. Alzheimers. Dis.* **66**, 853–854 (2018).
36. Kim, W. S. *et al.* Deletion of Abca7 increases cerebral amyloid-β accumulation in the J20 mouse model of Alzheimer's disease. *J. Neurosci.* **33**, 4387–4394 (2013).
37. Fu, Y., Hsiao, J.-H. T., Paxinos, G., Halliday, G. M. & Kim, W. S. ABCA7 Mediates Phagocytic Clearance of Amyloid-β in the Brain. *J. Alzheimers. Dis.* **54**, 569–584 (2016).
38. Aikawa, T. *et al.* ABCA7 haplodeficiency disturbs microglial immune responses in the mouse brain. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 23790–23796 (2019).
39. Aikawa, T. *et al.* ABCA7 Regulates Brain Fatty Acid Metabolism During LPS-Induced Acute Inflammation. *Front. Neurosci.* **15**, 647974 (2021).
40. Allen, M. *et al.* loss-of-function variants, expression, and neurologic disease risk. *Neurol Genet* **3**, e126 (2017).
41. Liu, Y. *et al.* Plasma lipidome is dysregulated in Alzheimer's disease and is associated with disease risk genes. *Transl. Psychiatry* **11**, 344 (2021).
42. Brase, L. *et al.* Single-nucleus RNA-sequencing of autosomal dominant Alzheimer disease and risk variant carriers. *Nat. Commun.* **14**, 2314 (2023).
43. Blanchard, J. W. *et al.* APOE4 impairs myelination via cholesterol dysregulation in oligodendrocytes. *Nature* **611**, 769–779 (2022).
44. Sayed, F. A. *et al.* AD-linked R47H-TREM2 mutation induces disease-enhancing microglial states via AKT hyperactivation. *Sci. Transl. Med.* (2021) doi:10.1126/scitranslmed.abe3947.
45. Wamsley, B. *et al.* Molecular cascades and cell type-specific signatures in ASD revealed by single-cell genomics. *Science* **384**, eadh2602 (2024).
46. Kamath, T. *et al.* Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in Parkinson's disease. *Nat. Neurosci.* **25**, 588–595 (2022).
47. Pauwels, K. *et al.* Structural basis for increased toxicity of pathological aβ42:aβ40 ratios in Alzheimer disease. *J. Biol. Chem.* **287**, 5650–5660 (2012).
48. Fraser, P. E., Nguyen, J. T., Surewicz, W. K. & Kirschner, D. A. pH-dependent structural transitions of Alzheimer amyloid peptides. *Biophys. J.* **60**, 1190–1201 (1991).
49. Pike, C. J., Burdick, D., Walencewicz, A. J., Glabe, C. G. & Cotman, C. W. Neurodegeneration

induced by beta-amyloid peptides in vitro: the role of peptide assembly state. *J. Neurosci.* **13**, 1676–1687 (1993).
50. Pettegrew, J. W., Panchalingam, K., Hamilton, R. L. & McClure, R. J. Brain membrane phospholipid alterations in Alzheimer's disease. *Neurochem. Res.* **26**, 771–782 (2001).
51. Phillips, J. C. Why Aβ42 Is Much More Toxic than Aβ40. *ACS Chem. Neurosci.* **10**, 2843–2847 (2019).
52. Duchateau, L., Wawrzyniak, N. & Sleegers, K. The ABC's of Alzheimer risk gene ABCA7. *Alzheimers. Dement.* (2024) doi:10.1002/alz.13805.
53. Johnson, E. C. B. *et al.* Large-scale proteomic analysis of Alzheimer's disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. *Nat. Med.* **26**, 769–780 (2020).
54. Welch, G. M. *et al.* Neurons burdened by DNA double-strand breaks incite microglia activation through antiviral-like signaling in neurodegeneration. *Sci Adv* **8**, eabo4662 (2022).
55. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).
56. Mathys, H. *et al.* Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology. *Cell* **186**, 4365–4385.e27 (2023).
57. Le Thi My, L. *et al.* Cryo-EM structures of human ABCA7 provide insights into its phospholipid translocation mechanisms. *EMBO J.* (2022) doi:10.15252/embj.2022111065.
58. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
59. Ho, S.-M. *et al.* Rapid Ngn2-induction of excitatory neurons from hiPSC-derived neural progenitor cells. *Methods* **101**, 113–124 (2016).
60. De Roeck, A. *et al.* An intronic VNTR affects splicing of ABCA7 and increases risk of Alzheimer's disease. *Acta Neuropathol.* **135**, (2018).
61. Krahmer, N. *et al.* Phosphatidylcholine synthesis for lipid droplet expansion is mediated by localized activation of CTP:phosphocholine cytidylyltransferase. *Cell Metab.* **14**, 504–515 (2011).
62. Guo, Y. *et al.* Functional genomic screen reveals genes involved in lipid-droplet formation and utilization. *Nature* **453**, 657–661 (2008).
63. Fei, W. *et al.* A role for phosphatidic acid in the formation of 'supersized' lipid droplets. *PLoS Genet.* **7**, e1002201 (2011).
64. Schott, M. B. *et al.* Lipid droplet size directs lipolysis and lipophagy catabolism in hepatocytes. *J. Cell Biol.* **218**, 3320–3335 (2019).
65. Virmani, M. A. & Cirulli, M. The Role of l-Carnitine in Mitochondria, Prevention of Metabolic Inflexibility and Disease Initiation. *Int. J. Mol. Sci.* **23**, 2717 (2022).
66. Yoo, H. C., Yu, Y. C., Sung, Y. & Han, J. M. Glutamine reliance in cell metabolism. *Exp. Mol. Med.* **52**, 1496–1516 (2020).
67. Glutamine metabolism in diseases associated with mitochondrial dysfunction. *Mol. Cell. Neurosci.* **126**, 103887 (2023).
68. Chatree, S., Thongmaen, N., Tantivejkul, K., Sitticharoon, C. & Vucenik, I. Role of Inositols and Inositol Phosphates in Energy Metabolism. *Molecules* **25**, 5079 (2020).
69. Basile, M. S., Bramanti, P. & Mazzon, E. Inosine in Neurodegenerative Diseases: From the Bench to the Bedside. *Molecules* **27**, 4644 (2022).
70. Furuhashi, M. *et al.* Differential regulation of hypoxanthine and xanthine by obesity in a general population. *J. Diabetes Investig.* **11**, 878–887 (2020).
71. Lee, J. S. *et al.* Hypoxanthine is a checkpoint stress metabolite in colonic epithelial energy modulation and barrier function. *J. Biol. Chem.* **293**, 6039–6051 (2018).
72. Morland, C. & Nordengen, K. -Acetyl-Aspartyl-Glutamate in Brain Health and Disease. *Int. J. Mol. Sci.* **23**, (2022).
73. Evolution of glutamatergic signaling and synapses. *Neuropharmacology* **199**, 108740 (2021).
74. The glutamatergic synapse – A key hub in neuronal metabolism, signalling and plasticity. *Neuropharmacology* **207**, 108945 (2022).

75. Inazu M. & Matsumiya T. [Physiological functions of carnitine and carnitine transporters in the central nervous system]. *Nihon Shinkei Seishin Yakurigaku Zasshi* **28**, 113–120 (2008).
76. Acetyl-l-carnitine restores synaptic transmission and enhances the inducibility of stable LTP after oxygen–glucose deprivation. *Neuroscience* **332**, 203–211 (2016).
77. Janiri, L., Falcone, M., Persico, A. & Tempesta, E. Activity of L-carnitine and L-acetylcarnitine on cholinoceptive neocortical neurons of the rat in vivo. *Journal of Neural Transmission / General Section JNT* **86**, 135–146 (1991).
78. Carnitine transport and fatty acid oxidation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1863**, 2422–2435 (2016).
79. Carnitine: transport and physiological functions in the brain. *Mol. Aspects Med.* **25**, 551–567 (2004).
80. Dienel, G. A. Brain Glucose Metabolism: Integration of Energetics with Function. *Physiol. Rev.* (2018) doi:10.1152/physrev.00062.2017.
81. Trigo, D., Avelar, C., Fernandes, M., Sá, J. & da Cruz e Silva, O. Mitochondria, energy, and metabolism in neuronal health and disease. *FEBS Lett.* **596**, 1095–1110 (2022).
82. Yellen, G. Fueling thought: Management of glycolysis and oxidative phosphorylation in neuronal metabolism. *J. Cell Biol.* **217**, 2235–2246 (2018).
83. Morant-Ferrando, B. *et al.* Fatty acid oxidation organizes mitochondrial supercomplexes to sustain astrocytic ROS and cognition. *Nat Metab* **5**, 1290–1302 (2023).
84. Divakaruni, A. S., Paradyse, A., Ferrick, D. A., Murphy, A. N. & Jastroch, M. Analysis and interpretation of microplate-based oxygen consumption and pH data. *Methods Enzymol.* **547**, 309–354 (2014).
85. Gu, X., Ma, Y., Liu, Y. & Wan, Q. Measurement of mitochondrial respiration in adherent cells by Seahorse XF96 Cell Mito Stress Test. *STAR Protoc* **2**, 100245 (2021).
86. Divakaruni, A. S. & Jastroch, M. A practical guide for the analysis, standardization and interpretation of oxygen consumption measurements. *Nature Metabolism* **4**, 978–994 (2022).
87. Divakaruni, A. S. & Brand, M. D. The regulation and physiology of mitochondrial proton leak. *Physiology* **26**, 192–205 (2011).
88. Jekabsons, M. B. & Nicholls, D. G. In Situ Respiration and Bioenergetic Status of Mitochondria in Primary Cerebellar Granule Neuronal Cultures Exposed Continuously to Glutamate *. *J. Biol. Chem.* **279**, 32989–33000 (2004).
89. Demine, S., Renard, P. & Arnould, T. Mitochondrial Uncoupling: A Key Controller of Biological Processes in Physiology and Diseases. *Cells* **8**, (2019).
90. Zorov, D. B. *et al.* Neuroprotective Potential of Mild Uncoupling in Mitochondria. Pros and Cons. *Brain Sci* **11**, (2021).
91. Mitochondrial proton leaks and uncoupling proteins. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1862**, 148428 (2021).
92. Uncoupling proteins in the mitochondrial defense against oxidative stress. *Prog. Retin. Eye Res.* **83**, 100941 (2021).
93. Suski, J. M. *et al.* Relation between mitochondrial membrane potential and ROS formation. *Methods Mol. Biol.* **810**, 183–205 (2012).
94. Hoang, T., Kuljanin, M., Smith, M. D. & Jelokhani-Niaraki, M. A biophysical study on molecular physiology of the uncoupling proteins of the central nervous system. *Biosci. Rep.* **35**, e00226 (2015).
95. Transport Function and Regulation of Mitochondrial Uncoupling Proteins 2 and 3. *J. Biol. Chem.* **274**, 26003–26007 (1999).
96. Ardalan, A., Smith, M. D. & Jelokhani-Niaraki, M. Uncoupling Proteins and Regulated Proton Leak in Mitochondria. *Int. J. Mol. Sci.* **23**, 1528 (2022).
97. Farese, R. V., Jr & Walther, T. C. Glycerolipid Synthesis and Lipid Droplet Formation in the Endoplasmic Reticulum. *Cold Spring Harb. Perspect. Biol.* **15**, (2023).
98. Schuler, M.-H., Di Bartolomeo, F., Mårtensson, C. U., Daum, G. & Becker, T. Phosphatidylcholine Affects Inner Membrane Protein Translocases of Mitochondria. *J. Biol. Chem.* **291**, 18718–18729 (2016).
99. Szymkowicz, L., Lento, C. & Wilson, D. J. Impact of Cardiolipin and Phosphatidylcholine

Interactions on the Conformational Ensemble of Cytochrome. *Biochemistry* **58**, 3617–3626 (2019).
100. Mejia, E. M. & Hatch, G. M. Mitochondrial phospholipids: role in mitochondrial function. *J. Bioenerg. Biomembr.* **48**, 99–112 (2015).
101. Prola, A. *et al.* Cardiolipin content controls mitochondrial coupling and energetic efficiency in muscle. *Sci Adv* **7**, (2021).
102. Zeisel, S. H. & da Costa, K.-A. Choline: an essential nutrient for public health. *Nutr. Rev.* **67**, 615–623 (2009).
103. Son, Y., Kenny, T. C., Khan, A., Birsoy, K. & Hite, R. K. Structural basis of lipid head group entry to the Kennedy pathway by FLVCR1. *Nature* 1–7 (2024).
104. Lipid Droplets and Neurodegeneration. *Neuroscience* **549**, 13–23 (2024).
105. Olzmann, J. A. & Carvalho, P. Dynamics and functions of lipid droplets. *Nat. Rev. Mol. Cell Biol.* **20**, 137–155 (2019).
106. van der Veen, J. N. *et al.* The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochim. Biophys. Acta Biomembr.* **1859**, 1558–1572 (2017).
107. Bharadwaj, P. *et al.* Role of the cell membrane interface in modulating production and uptake of Alzheimer's beta amyloid protein. *Biochim. Biophys. Acta Biomembr.* **1860**, 1639–1651 (2018).
108. Yang, X., Sun, G. Y., Eckert, G. P. & Lee, J. C.-M. Cellular membrane fluidity in amyloid precursor protein processing. *Mol. Neurobiol.* **50**, 119–129 (2014).
109. Walter, J. & van Echten-Deckert, G. Cross-talk of membrane lipids and Alzheimer-related proteins. *Mol. Neurodegener.* **8**, 34 (2013).
110. Takasugi, N. *et al.* BACE1 activity is modulated by cell-associated sphingosine-1-phosphate. *J. Neurosci.* **31**, 6850–6857 (2011).
111. Kalvodova, L. *et al.* Lipids as modulators of proteolytic activity of BACE: involvement of cholesterol, glycosphingolipids, and anionic phospholipids in vitro. *J. Biol. Chem.* **280**, 36815–36823 (2005).
112. Tseng, Y.-H., Cypess, A. M. & Kahn, C. R. Cellular bioenergetics as a target for obesity therapy. *Nat. Rev. Drug Discov.* **9**, 465–482 (2010).
113. Rossmeisl, M. *et al.* Triglyceride-lowering effect of respiratory uncoupling in white adipose tissue. *Obes. Res.* **13**, (2005).
114. Goedeke, L. & Shulman, G. I. Therapeutic potential of mitochondrial uncouplers for the treatment of metabolic associated fatty liver disease and NASH. *Molecular Metabolism* **46**, (2021).
115. Amorim, J. A. *et al.* Mitochondrial and metabolic dysfunction in ageing and age-related diseases. *Nat. Rev. Endocrinol.* **18**, 243–258 (2022).
116. Ralhan, I., Chang, C.-L., Lippincott-Schwartz, J. & Ioannou, M. S. Lipid droplets in the nervous system. *J. Cell Biol.* **220**, (2021).
117. Islimye, E., Girard, V. & Gould, A. P. Functions of Stress-Induced Lipid Droplets in the Nervous System. *Front Cell Dev Biol* **10**, 863907 (2022).
118. Ioannou, M. S. *et al.* Neuron-Astrocyte Metabolic Coupling Protects against Activity-Induced Fatty Acid Toxicity. *Cell* **177**, 1522–1535.e14 (2019).
119. Moulton, M. J. *et al.* Neuronal ROS-induced glial lipid droplet formation is altered by loss of Alzheimer's disease–associated genes. *Proceedings of the National Academy of Sciences* **118**, e2112095118 (2021).
120. McFadden, J. W. *et al.* Increasing fatty acid oxidation remodels the hypothalamic neurometabolome to mitigate stress and inflammation. *PLoS One* **9**, e115642 (2014).
121. C75, a Fatty Acid Synthase Inhibitor, Modulates AMP-activated Protein Kinase to Alter Neuronal Energy Metabolism. *J. Biol. Chem.* **279**, 3817–3827 (2004).
122. Jastroch, M. Uncoupling protein 1 controls reactive oxygen species in brown adipose tissue. *Proceedings of the National Academy of Sciences* **114**, 7744–7746 (2017).
123. Giralt, M. & Villarroya, F. Mitochondrial Uncoupling and the Regulation of Glucose Homeostasis. *Curr. Diabetes Rev.* **13**, 386–394.
124. Fedorenko, A., Lishko, P. V. & Kirichok, Y. Mechanism of fatty-acid-dependent UCP1 uncoupling in brown fat mitochondria. *Cell* **151**, 400–413 (2012).

125. CTP:phosphocholine cytidylyltransferase. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1348**, 79–90 (1997).
126. Tau interactome maps synaptic and mitochondrial processes associated with neurodegeneration. *Cell* **185**, 712–728.e14 (2022).
127. Bano, D., Ehninger, D. & Bagetta, G. Decoding metabolic signatures in Alzheimer's disease: a mitochondrial perspective. *Cell Death Discovery* **9**, 1–4 (2023).
128. Zong, Y. *et al.* Mitochondrial dysfunction: mechanisms and advances in therapy. *Signal Transduction and Targeted Therapy* **9**, 1–29 (2024).
129. Declining NAD+ Induces a Pseudohypoxic State Disrupting Nuclear-Mitochondrial Communication during Aging. *Cell* **155**, 1624–1638 (2013).
130. Picca, A., Faitg, J., Auwerx, J., Ferrucci, L. & D'Amico, D. Mitophagy in human health, ageing and disease. *Nature Metabolism* **5**, 2047–2061 (2023).
131. Park, A. *et al.* Mitochondrial matrix protein LETMD1 maintains thermogenic capacity of brown adipose tissue in male mice. *Nat. Commun.* **14**, 1–14 (2023).
132. Mitochondria controlled by UCP2 determine hypoxia-induced synaptic remodeling in the cortex and hippocampus. *Neurobiol. Dis.* **90**, 68–74 (2016).
133. Shadel, G. S. & Horvath, T. L. Mitochondrial ROS signaling in organismal homeostasis. *Cell* **163**, 560–569 (2015).
134. Korshunov, S. S., Skulachev, V. P. & Starkov, A. A. High protonic potential actuates a mechanism of production of reactive oxygen species in mitochondria. *FEBS Lett.* **416**, (1997).
135. Wisløff, U. *et al.* Cardiovascular Risk Factors Emerge After Artificial Selection for Low Aerobic Capacity. *Science* (2005) doi:10.1126/science.1108177.
136. Andrews, Z. B., Diano, S. & Horvath, T. L. Mitochondrial uncoupling proteins in the cns: in support of function and survival. *Nat. Rev. Neurosci.* **6**, 829–840 (2005).
137. Uncoupling proteins as a therapeutic target for the development of new era drugs against neurodegenerative disorder. *Biomed. Pharmacother.* **147**, 112656 (2022).
138. Robert, G. & Richard Wagner, J. ROS-Induced DNA Damage as an Underlying Cause of Aging. *Advances in Geriatric Medicine and Research* **4**, (2020).
139. Volanti, C., Matroule, J.-Y. & Piette, J. Involvement of oxidative stress in NF-kappaB activation in endothelial cells treated by photodynamic therapy. *Photochem. Photobiol.* **75**, 36–45 (2002).
140. Canty, T. G., Jr *et al.* Oxidative stress induces NF-kappaB nuclear translocation without degradation of IkappaBalpha. *Circulation* **100**, II361–4 (1999).
141. Schreck, R., Albermann, K. & Baeuerle, P. A. Nuclear factor kappa B: an oxidative stress-responsive transcription factor of eukaryotic cells (a review). *Free Radic. Res. Commun.* **17**, 221–237 (1992).
142. Byrns, C. N. *et al.* Senescent glia link mitochondrial dysfunction and lipid accumulation. *Nature* 1–9 (2024).
143. Gavrilova, S. I., Kolykhalov, I. V., Ponomareva, E. V., Fedorova, Y. B. & Selezneva, N. D. [Clinical efficacy and safety of choline alfoscerate in the treatment of late-onset cognitive impairment]. *Zh. Nevrol. Psikhiatr. Im. S S Korsakova* **118**, 45–53 (2018).
144. Blusztajn, J. K., Slack, B. E. & Mellott, T. J. Neuroprotective Actions of Dietary Choline. *Nutrients* **9**, 815 (2017).
145. Bhattacharyya, R., Fenn, R. H., Barren, C., Tanzi, R. E. & Kovacs, D. M. Palmitoylated APP Forms Dimers, Cleaved by BACE1. *PLoS One* **11**, e0166400 (2016).
146. Sienski, G. *et al.* disrupts intracellular lipid homeostasis in human iPSC-derived glia. *Sci. Transl. Med.* **13**, (2021).
147. Wang, H., Bennett, D. A., De Jager, P. L., Zhang, Q.-Y. & Zhang, H.-Y. Genome-wide epistasis analysis for Alzheimer's disease and implications for genetic risk prediction. *Alzheimers. Res. Ther.* **13**, 55 (2021).
148. Hemani, G., Knott, S. & Haley, C. An evolutionary perspective on epistasis and the missing heritability. *PLoS Genet.* **9**, e1003295 (2013).
149. Haig, D. Does heritability hide in epistasis between linked SNPs? *Eur. J. Hum. Genet.* **19**, 123 (2011).

150. Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 1193–1198 (2012).
151. Haney, M. S. *et al.* APOE4/4 is linked to damaging lipid droplets in Alzheimer's disease microglia. *Nature* **628**, 154–161 (2024).
152. Victor, M. B. *et al.* Lipid accumulation induced by APOE4 impairs microglial surveillance of neuronal-network activity. *Cell Stem Cell* **29**, 1197–1212.e8 (2022).

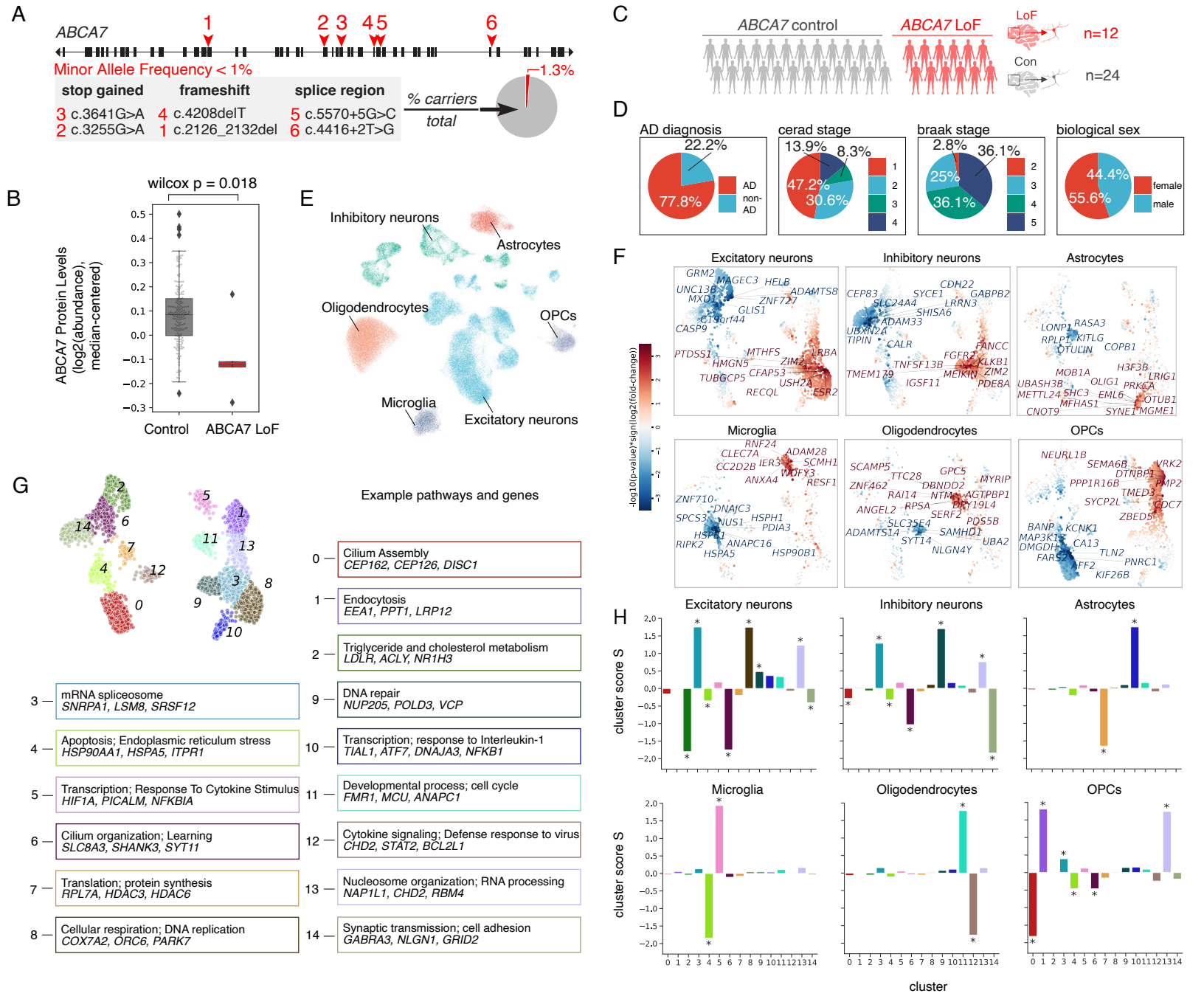


Fig. 1. Single-nuclear RNA-sequencing Atlas of Human Post-mortem Prefrontal Cortex Reveals Cell Type-specific Gene Changes in ABCA7 LoF. (A) Overview of ABCA7 gene structure with the location of variants represented in this study (average minor allele frequency for depicted variants is < 1%). Exons are depicted as black rectangles, and introns as black lines. The pie chart indicates the frequency of ABCA7 PTC-variant-carriers within the ROSMAP cohort. (B) ABCA7 protein levels (log2(abundance)) from post-mortem human prefrontal cortex in all available controls ($N = 180$) vs. ABCA7 LoF carriers ($N = 5$). P-value computed by Wilcoxon rank sum test. Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). (C) Overview of human cohort for snRNA-seq (Created with BioRender.com). (D) Overview of snRNA-seq cohort metadata for 32 individuals. (E) 2D UMAP projection of per-cell gene expression values and their transcriptionally defined cell type. (F) 2D UMAP projection of ABCA7 LoF gene perturbation scores ($S = -\log_{10}(\text{p-value}) \times \text{sign}(\log_2(\text{fold change}))$); Red = $S > 1.3$, Blue = $S < -1.3$; Point size indicates $|S|$). Up to top 20 genes by $|S|$ are labeled. (G) Genes in 2D UMAP space colored by cluster assignment (Gaussian mixture model; see Methods) with per-cluster pathway enrichments shown (GO BP, hypergeometric enrichment, $p < 0.01$). (H) Cell type-specific gene cluster scores ($SC = \text{mean}(S_i)$, for genes i in cluster c). * indicates permutation FDR-adjusted p-value < 0.01 and $|SC| > 0.25$.

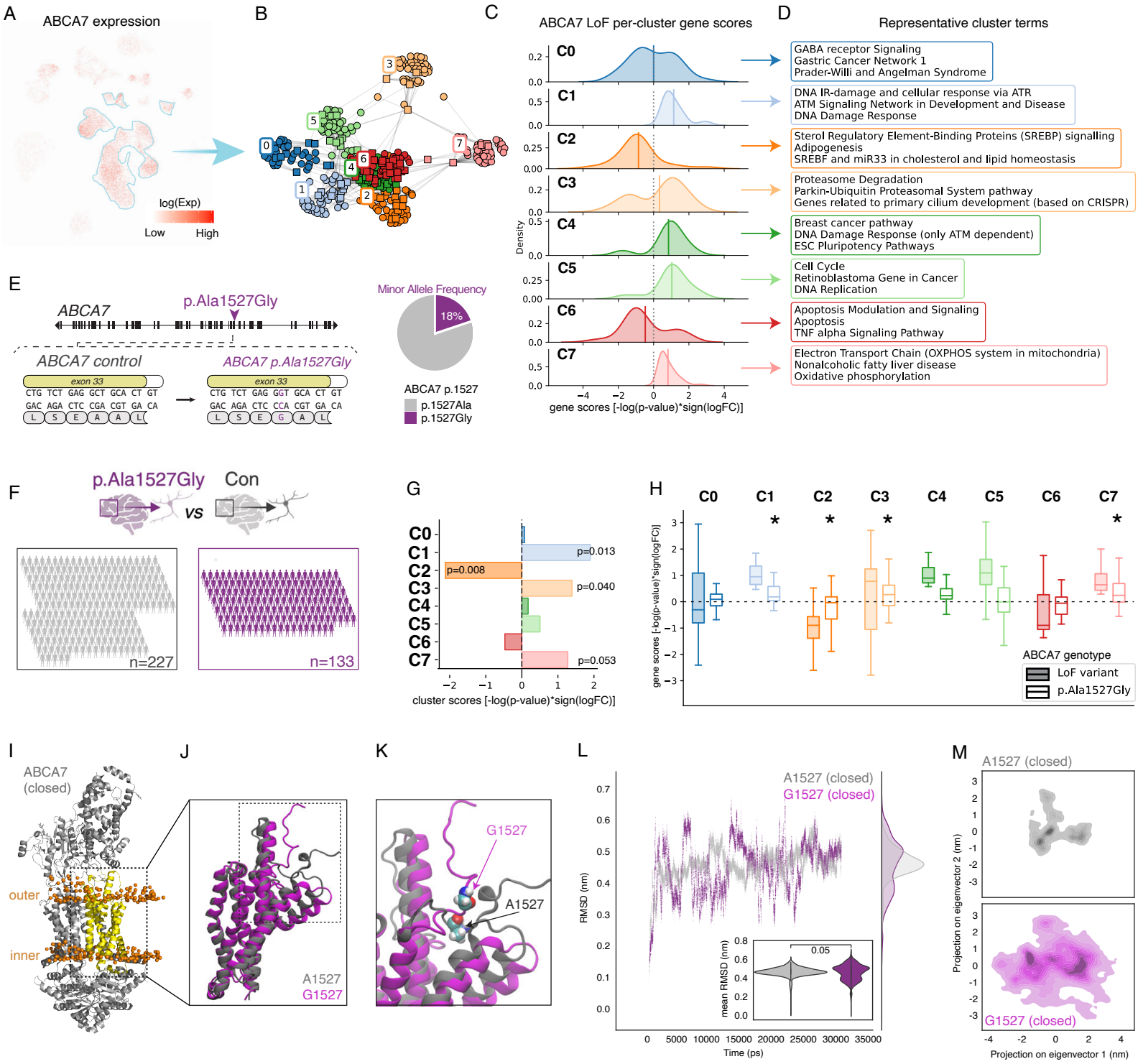
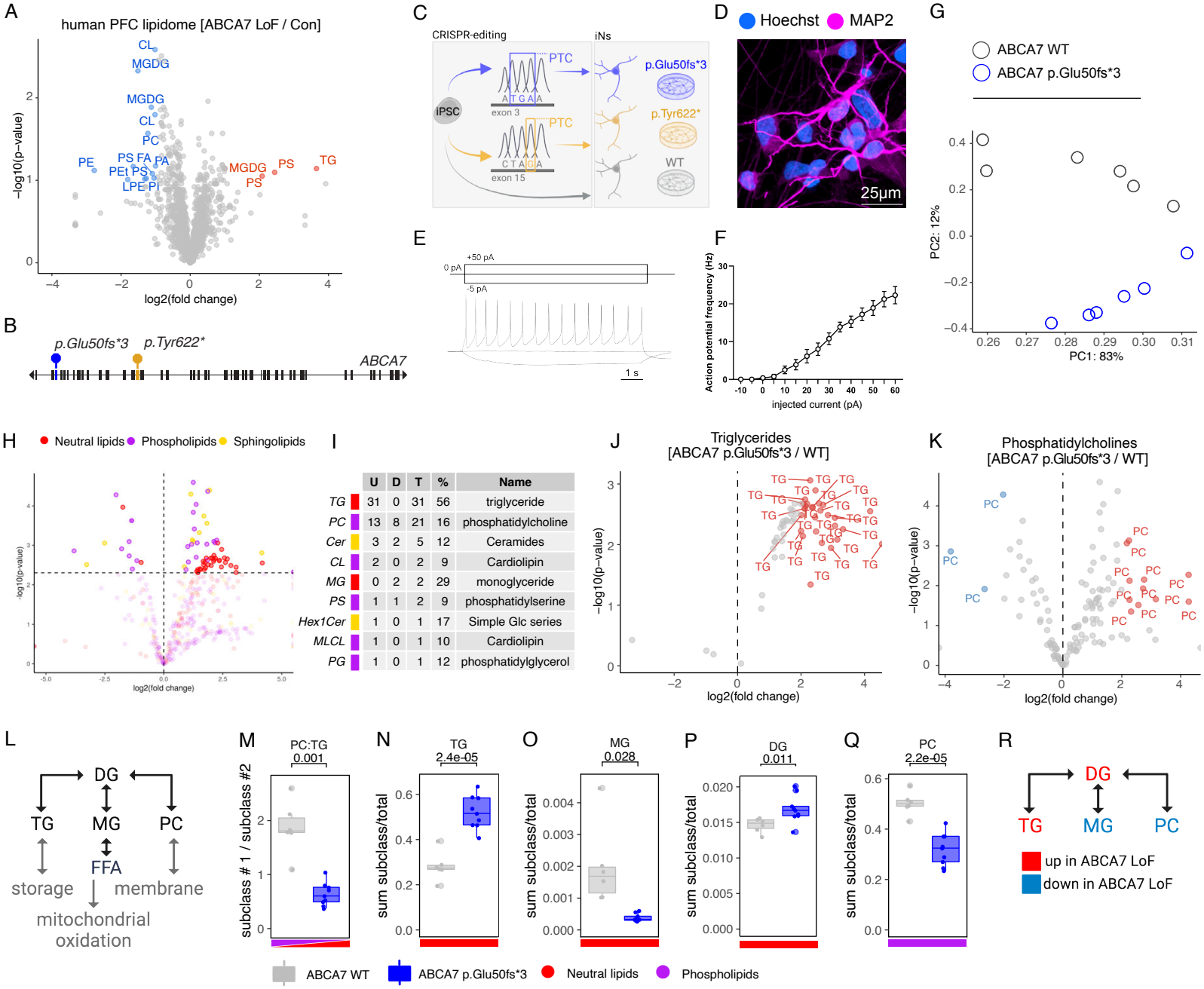


Fig. 2. Transcriptional Perturbations in Excitatory Neurons in ABCA7 LoF and ABCA7 p.Ala1527Gly Variant Carriers. (A) 2D UMAP projection of individual cells colored by log(Exp), where Exp represents log-normalized ABCA7 expression values. (B) Kernighan-Lin (K/L) clustering on leading edge genes from pathways perturbed in ABCA7 LoF excitatory neurons, where $p < 0.05$. Colors indicate distinct K/L gene clusters, which are numbered from 0 to 7. (C) Gaussian kernel density estimate plots of gene scores S for genes belonging to a given gene cluster. $S > 0$ indicates upregulation in ABCA7 LoF. Solid lines indicate distribution means. (D) Representative pathways that annotate the largest number of genes within a cluster (i.e., with the highest intra-cluster connectivity) shown per-cluster. (E) Schematic indicating the genomic location of the p.Ala1527Gly codon change. A purple arrow indicates the location of the missense variant in the ABCA7 gene. Minor allele frequency shown to the right. (F) Overview of snRNA-seq cohort of ABCA7 p.Ala1527Gly carriers (homozygous and heterozygous) vs. control non-carriers (minor allele frequency approx. 18%). (G) Perturbation of ABCA7 LoF-associated gene clusters from (B-D) in excitatory neurons of p.Ala1527Gly variant-carriers vs. non-carrier controls, computed by FGSEA. Top p-values ($p < 0.1$) are indicated. $S > 0$ indicates upregulation in carriers. (H) Distributions of gene scores S for genes belonging to a given gene cluster for ABCA7 p.Ala1527Gly (no fill) or ABCA7 LoF-variants (solid fill). $S > 0$ indicates upregulation in ABCA7 variant. * indicates FGSEA p-value<0.1 from (G). Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5

times the interquartile range from the first or third quartile). (I) Closed conformation ABCA7 protein structure. ABCA7 domain between residues 1517 and 1756 used for simulations is shown in yellow. Lipid bilayer shown in orange. (J) Expanded yellow domain (inset from *I*), with A1527 variant (light grey) and G1527 variant (purple). (K) Expanded inset from *J* with residues of interest rendered. (L) Root mean squared deviations of closed conformation domains from *J* with A1527 (light grey) or G1527 (purple) under simulation. Structural deviations over time were computed with respect to reference closed conformation from *J*. Violin plot inset indicates average C_α atom positional fluctuations over time. (M) Projection of C_α atom positional fluctuations under simulation onto the first two principal components, for closed conformation domain from *J* with A1527 (top, light grey) or G1527 (bottom, purple).



	U	D	T	%	Name
TG	31	0	31	56	triglyceride
PC	13	8	21	16	phosphatidylcholine
Cer	3	2	5	12	Ceramides
CL	2	0	2	9	Cardiolipin
MG	0	2	2	29	monoglyceride
PS	1	1	2	9	phosphatidylserine
Hex1Cer	1	0	1	17	Simple Glc series
MLCL	1	0	1	10	Cardiolipin
PG	1	0	1	12	phosphatidylglycerol

Fig. 3. ABCA7 LoF Neurons Have Neutral Lipid Accumulation and Phospholipid Imbalances. (A) Volcano plot showing the distribution of lipid species by log-fold-change and log-p-value, where log-fold-change > 0 indicates up-regulation in ABCA7 LoF post-mortem PFC (N=8) vs. Control (N=8). A subset of lipid species is labeled ($p < 0.1$ & logFC>0.5 (red) or logFC<-0.5 (blue)) (p-values by independent sample t-test). (B) Overview of two iPSC-derived isogenic neuronal lines carrying ABCA7 LoF variants. ABCA7 gene map depicts exons (black rectangles) and introns (black lines). (C) CRISPR-Cas9 was used to generate an ABCA7 LoF isogenic iPSC line by introducing a premature termination codon into exon 3 or exon 15 (ABCA7 p.Glu50fs*3, blue or p.Tyr622*, orange, respectively). (D) Example MAP2 staining of 2-week-old p.Tyr622* iNs. (E) Representative sweeps show action potentials elicited by 800 ms of current injections in patched 4-week-old iNs. (F) Summary of action potential frequency (means ± SEM) elicited with different amounts of injected current in 4-week-old iNs. (G) Projection of control and ABCA7 p.Glu50fs*3 lipidomes (per-sample z-scaled peak areas; $N = 6$ per genotype) onto the first two principal components from lipid space. Fraction of explained variance shown along each axis. (H) Volcano plot showing perturbed lipid species by class (p-values by independent sample t-test). (I) Overview of perturbed lipid species by lipid subclass. $U =$ number of lipids in that subclass where $p < 5 \times 10^{-3}$ & fold-change > 0. $D =$ number of lipids in that subclass where $p < 5 \times 10^{-3}$ & fold-change < 0. $T = U + D$. % = $(T/N) \times 100$, where $N =$ total number of lipids in a given subclass (p-values by independent sample t-test; $N = 6$ per genotype). (J,K) Volcano plot showing the distribution of triglyceride (TG) species (J) and phosphatidylcholine (PC) species (K) by log-fold-change and log-p-value, where log-fold-change > 0 indicates up-regulation in p.Glu50fs*3 ($N = 6$) vs. WT ($N = 6$) iPSC-derived neurons. A subset of lipid species is labeled ($p < 0.05$ & logFC>2 (red) or logFC<-2 (blue)) (p-values by independent sample t-test). (L) Schematic of metabolic link between TG, monoglyceride (MG), and diglyceride (DG). (M-Q) Relative abundance of select lipid species in p.Glu50fs*3 iN ($N = 6$) and WT iN ($N = 6$) (p-values by independent sample t-test). Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). (R) Schematic from (L) overlaid with abundance changes in p.Glu50fs*3 iNs vs. WT iNs. Schematics in (C, L, R) created with BioRender.com.

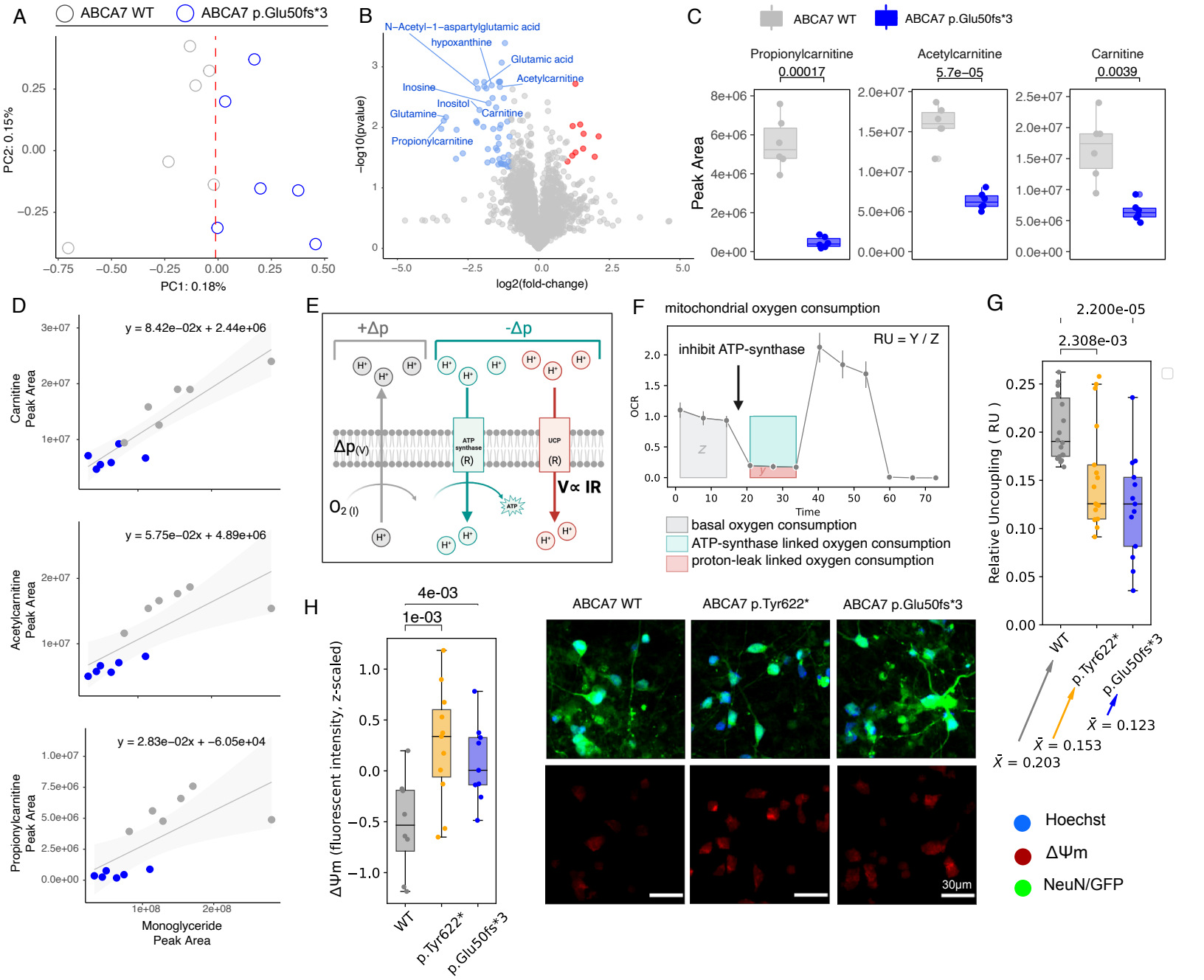


Fig. 4. ABCA7 LoF Impairs Regulation of Mitochondrial Uncoupling and Carbon Flux in Neurons. (A) Projection of WT and p.Glu50fs*3 metabolomes (per-sample z-scaled normalized peak areas; $N = 6$ per genotype) onto the first two principal components from metabolite space. Fraction of explained variance shown along each axis. The dotted red line separating the two genotypes indicates the median PC1 value. (B) Volcano plot indicating differentially regulated metabolites by log2-fold-change and log10(p-value), where log2(FC) > 0 indicates increased abundance in p.Glu50fs*3 vs. WT. Top up- and down-regulated metabolites are shown in red and blue, respectively (p-values by independent sample t-test). (C) Abundance of carnitine species in WT vs. p.Glu50fs*3 iNs. (D) Correlation of carnitine species abundance (normalized peak areas by LC-MS) and monoglyceride (MG) abundance (peak areas by LC-MS) from matched metabolomic-lipidomic samples. Grey points indicate WT iNs ($N = 6$). Blue points indicate p.Glu50fs*3 iNs ($N = 6$). The equation indicates the linear function fit. Grey error bar indicates 95% confidence interval. (E) Schematic indicating the relationship between oxygen consumption as a measure of proton current (I), which sustains the proton motive force (Δp; voltage (V)). Regulation of ATP synthase and uncoupling protein (UCP) activity modifies resistance (R) and depletes Δp. (F) Schematic indicating how relative uncoupling is computed from oxygen consumption rate (OCR). Remaining oxygen consumption after pharmacological inhibition of ATP synthase gives the proportion of basal oxygen consumption attributed to proton leak. (G) Relative uncoupling quantified by Seahorse oxygen consumption assay (see Methods) in 4-week-old WT vs. ABCA7 LoF iNs. P-values computed by independent sample t-test. N wells = 18 (WT), 17 (p.Tyr622*), 13 (p.Glu50fs*3) across two independent differentiation batches and Seahorse experiments (see Fig. S11E). (H) Left: Quantification of neuronal HCS MitoHealth dye fluorescence intensity as a measure of mitochondrial membrane potential. P-values computed by linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 8$ (WT), 11 (p.Tyr622*), 9 wells (p.Glu50fs*3) (3000 cells per condition) from three independent differentiation batches (see Fig. S11F). Individual data points indicate per-well averages of cell-level intensities. Right: Representative images per condition as mean-intensity projections of the entire image (NeuN+) and within NeuN+ volumes considered for quantification (MitoHealth, Hoechst). Representative images for the MitoHealth channel were processed with condition-wide percentile-based background subtraction and thresholding. Representative images of cell soma underwent per-image percentile-based background subtraction and thresholding, reflecting the segmentation methodology. For (C, G, H) boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile).

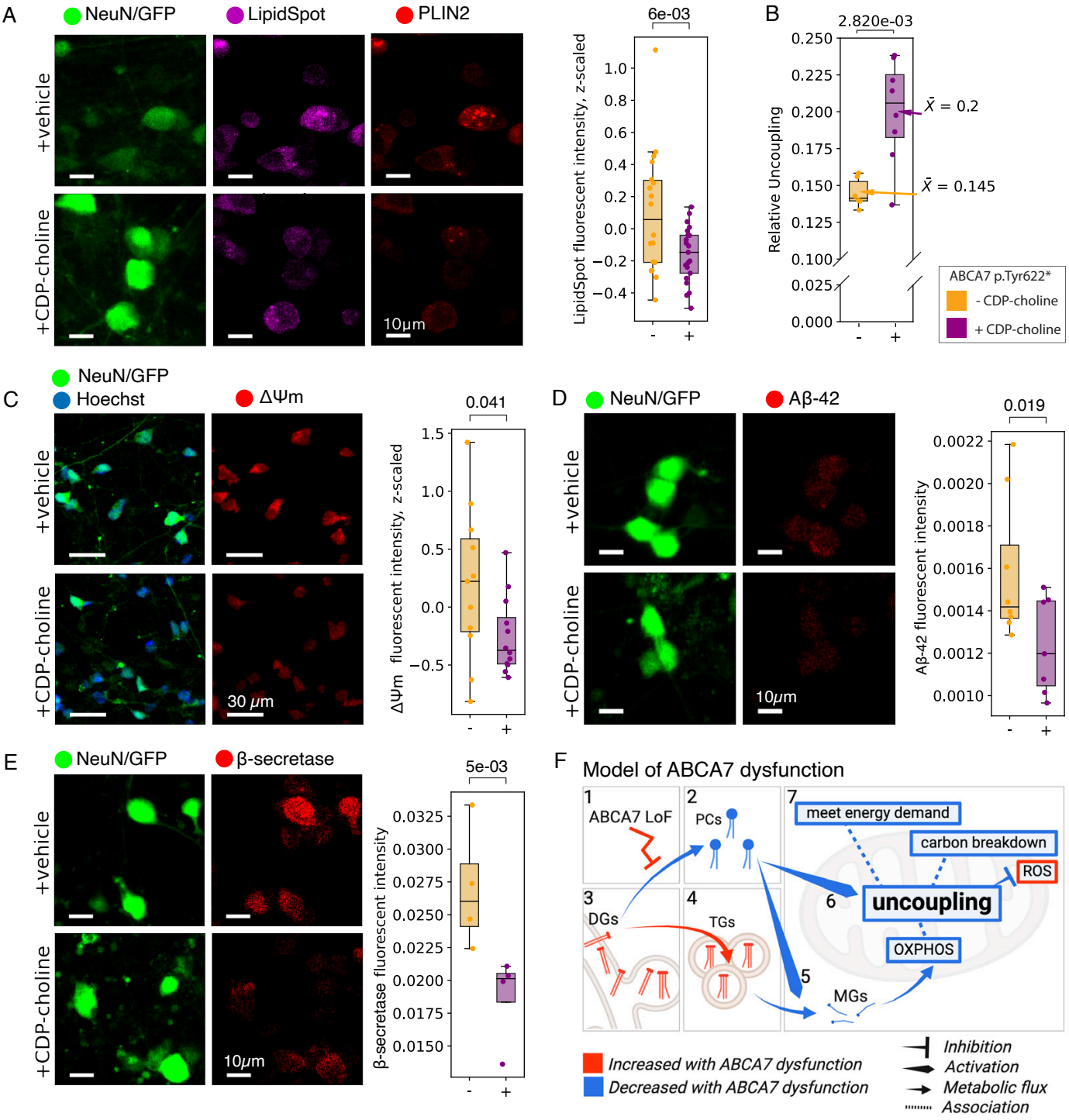


Fig. 5. Supplementation with CDP-choline Reduces Neutral Lipid Accumulation, Restores Mitochondrial Function, and Reduces Amyloid Pathology in ABCA7 LoF Neurons. (A) Left: Representative images per condition as mean-intensity projections of the entire image (NeuN+) and within NeuN+ volumes considered for quantification (LipidSpot, PLIN2). Right: Quantification of neuronal LipidSpot dye fluorescence intensity. P-values computed by linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 16$ (p.Tyr622* + H2O), 21 wells (p.Tyr622* + CDP-choline) (5419 cells per condition) from three independent differentiation batches of 4-week-old iNs treated for 2 weeks (see Fig. S12B). (B) Relative mitochondrial uncoupling, quantified by Seahorse oxygen consumption assay for 4-week-old p.Tyr622* iNs treated with CDP-choline or vehicle control for 2 weeks. Relative uncoupling gives the proportion of basal oxygen consumption attributed to uncoupled proton leak. P-values computed by independent sample t-test. N wells = 6 (p.Tyr622* + H2O), 8 (p.Tyr622* + CDP-choline). (C) Left: Representative images per condition as mean-intensity projections of the entire image (NeuN+) and within NeuN+ volumes considered for quantification (MitoHealth, Hoechst). Right: Quantification of neuronal HCS MitoHealth dye fluorescence intensity as a measure of mitochondrial membrane potential. P-values computed by linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 11$ (p.Tyr622* + H2O; datapoints from Fig. 4H), 12 wells (p.Tyr622* + CDP-choline) (3929 cells per condition) from three independent differentiation batches of 4-week-old iNs treated for 2 weeks (see Fig. S12F). (D) Left: Representative images per condition as mean-intensity projections of the entire image (NeuN+) and projections within NeuN+ volumes considered for quantification (Aβ42). Right: Quantification of neuronal Aβ42 fluorescence intensity. P-values

computed by linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 8$ (p.Tyr622* + H2O; datapointes from Fig. S9A) (1466 cells), 7 wells (p.Tyr622* + CDP-choline) (1102 cells) from 4-week-old iNs treated for 2 weeks. (E) Left: Representative images per condition as mean-intensity projections of the entire image (NeuN+) and within NeuN+ volumes considered for quantification (β-secretase). Right: Quantification of neuronal β-secretase fluorescence intensity. P-values computed by linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 4$ (p.Tyr622* + H2O) (3107 cells), 4 wells (p.Tyr622* + CDP-choline) (2829 cells) from 4-week-old iNs treated for 2 weeks. (F) Proposed model of ABCA7 dysfunction in neurons: ABCA7 loss-of-function (LoF) induces a compositional shift away from phosphatidylcholines (PCs) (1,2), which may directly increase triglycerides (TGs) by increasing precursor availability (3, 4) or indirectly by impairing lipid droplet processing (5) and mitochondrial uncoupling (6). Impaired uncoupling can lead to multiple detrimental cellular effects, including a decreased ability to meet the cell's energy demand, impaired regulation of carbon breakdown via oxidative phosphorylation (OXPHOS), and increased oxidative stress (7). Abbreviations: DG = diglyceride, TG = triglyceride, MG = monoglyceride, PC = phosphatidylcholine, ROS = reactive oxygen species. For (A-E), + = treated with CDP-choline. - = treated with vehicle control. Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). For (A, C-E), individual data points indicate per-well averages of cell-level intensities. Representative images for the quantified channel were processed with condition-wide percentile-based background subtraction and thresholding. Representative images of cell soma underwent per-image percentile-based background subtraction and thresholding, reflecting the segmentation methodology. Schematic in (F) created with Biorender.com.

Supplementary Materials for

Single-cell atlas of ABCA7 loss-of-function reveals impaired neuronal respiration via choline-dependent lipid imbalances

Materials and Methods **1**
- Experimental Methods 1
  - Human post-mortem 1
  - NGN2 3
- Computational Methods 8
  - Single-Cell 8
  - Image Processing 15
  - Molecular Dynamics Simulations 17
  - LC-MS data analysis 18
  - Other data analysis 19

Supplementary Text **20**
- Description of variables according to the Rush Alzheimer's Disease Center Codebook 20
- Choosing a partitioning heuristic for gene-pathway grouping 21
- Molecular Dynamics Simulations Results 23

Supplementary Figure Legends **25**

Supplementary Tables **29**

Extended Data Descriptions **32**

Supplementary References **33**

Materials and Methods

Experimental Methods

Human post-mortem

Isolation of nuclei from frozen *post-mortem* brain tissue

For batch #1: The protocol for the isolation of nuclei from frozen *post-mortem* brain tissue (region BA10) was adapted for smaller sample volumes from a previous study [1]. All procedures were carried out on ice or at 4°C. In brief, *post-mortem* brain tissue was homogenized in 700 µl Homogenization Buffer (320 mM sucrose, 5 mM CaCl2, 3 mM Mg(CH3COO)2, 10 mM Tris HCl pH 7.8, 0.1 mM EDTA pH 8.0, 0.1% IGEPAL CA-630, 1 mM β-mercaptoethanol, and 0.4 U µl−1 recombinant RNase inhibitor (Clontech)) using a Wheaton Dounce tissue grinder (15 strokes with the loose pestle). Homogenized tissue was filtered through a 40-µm cell strainer, mixed with an equal volume of Working Solution, which is prepared by mixing Diluent (30mM CaCl2, 18mM Mg(CH3COO)2, 60mM Tris pH 7.8, 0.6mM EDTA, 6mM β-mercaptoethanol) with Optiprep density gradient solution (Sigma-Aldrich D1556-250ML) in a 1:5 ratio. The sample mix was then loaded on top of an Optiprep density gradient consisting of 750 µl 30% OptiPrep solution (1.5:1 ratio of Working Solution:Homogenization Buffer) on top of 300 µl 40% OptiPrep solution (4:1 ratio of Working Solution:Homogenization Buffer). The nuclei were separated by centrifugation (5 min, 10,000 g, 4 °C). Approximately 100µl of nuclei were collected from the 30%/40% interface and washed twice with 1 ml of PBS containing 0.04% BSA, centrifuging 300g for 3 min (4 °C) in between, then resuspended in 100µl PBS containing 0.04% BSA. The nuclei were counted on C-Chip disposable hemocytometer and diluted to 1000 nuclei per µl in PBS containing 0.04% BSA. *For batch #2:* These samples (fresh post-mortem brain; PFC BA10) were prepared as part of and according to a previous study [1].

Informed consent and Anatomical Gift Act consent were obtained from each participant. The Religious Orders Study and Rush Memory and Aging Project were approved by the Institutional Review Board (IRB) of Rush University Medical Center. All participants signed a repository consent, allowing their data and biospecimens to be shared.

Droplet-based snRNA-seq

For batch #1: cDNA libraries were prepared in two batches using the Chromium Single Cell 3′ Reagent Kits v2 or v3 according to the manufacturer's protocol (10x Genomics). The generated scRNA-seq libraries from batch # 1 were sequenced using NovaSeq 6000 S2 (28 + 91 bases pair-end run with 8 nucleotide index). Samples were split over two lanes. To increase sequencing depth, samples were sequenced twice, on two flow cells. The generated scRNA-seq libraries for

batch # 2 were sequenced using NextSeq 500/550 High Output v2 kits (150 cycles) as part of a previous study, described here [1].

RNAscope in *post-mortem* human brain tissue

Fresh frozen PFC tissue (region BA10) was sectioned on a cryostat microtome (Leica CM3050 S) at 10µm. RNAscope was performed using RNAscope Multiplex Fluorescent Reagent Kit v2 with TSA Vivid Dyes (ACD Bio 323270) according to manufacturer's instructions with probes targeting human *ACLY* (Cat. No. 460391-C2), *SCP2* (Cat. No. 875961), *COX7A2* (Cat. No. 1288461-C2), or *TLK2* (Cat. No. 1288451-C2). All hybridizations were performed with an additional probe to label vGlut1 positive cells, human *SLC17A7* (Cat. No. 415611-C3). Images were acquired using a Zeiss LSM 880 confocal microscope at 40X for quantification. Images were analyzed blinded to genotype with a custom ImageJ macros. In brief, the macros first identified ROIs based on DAPI that were positive for *SLC17A7* signal, and within those ROIs, performed a particle analysis to count RNAscope probe dots per cell. Results were reported as dot/ROI and as H-score (defined by ACD Bio analysis guidelines). 3-4 images were acquired and scored per individual, n = 4 individuals per genotype.

Lipidomics of *post-mortem* human brain tissue

A biphasic extraction protocol was used to isolate a lipid fraction from frozen *post-mortem* prefrontal cortex (~100mg). Briefly, weighed tissue was homogenized using Bio-vortexer Homogenizer (Daigger) in 1 mL cold methanol (Sigma MX0486) and transferred to glass vials (VWR 66011-550) with an additional 1mL of cold methanol. Chloroform (Sigma 1.02444) (4 mL; cold) was added to each vial, and mixed by vortexing for 1 min. Water (Sigma WX0001) (2 mL; cold) was added to each vial, and mixed by vortexing for 1 min. Vials were placed in 50 mL conical tubes and centrifuged for 10 min at 3000 rcf for phase separation. The lower, chloroform phase was collected (3 mL from each sample) and transferred to new vials. A blank treated the same as samples was included throughout the biphasic extraction for lipidomic analysis. Lipidomics and analysis was performed in collaboration with the Harvard Center for Mass Spectrometry (HCMS).

For lipidomics, samples were dried under nitrogen flow and resuspended in 60 µL of chloroform. Each sample was split into two equal aliquots, one for each polarity analysis. LC–MS analyses were modified from [2] and were performed on an Orbitrap Exactive plus MS (Thermo Scientific) in line with an Ultimate 3000 LC (Thermo Scientific). Each sample was analyzed in positive and negative modes, in top 5 automatic data-dependent MS/MS mode. Column hardware consisted of a Biobond C4 column (4.6 × 50 mm, 5 µm, Dikma Technologies). Flow rate was set to 100 µL min^{-1} for 5 min with 0% mobile phase B (MB), then switched to 400 µL min^{-1} for 50 min, with a linear gradient of MB from 20% to 100%. The column was then washed at 500 µL min^{-1} for 8

min at 100% MB before being re-equilibrated for 7 min at 0% MB and 500 µL min^{-1}. For positive mode runs, buffers consisted for mobile phase A (MA) of 5mM ammonium formate, 0.1 % formic acid and 5% methanol in water, and for MB of 5 mM ammonium formate, 0.1% formic acid, 5% water, 35% methanol in isopropanol. For negative runs, buffers consisted for MA of 0.03% ammonium hydroxide, 5% methanol in water, and for MB of 0.03% ammonium hydroxide, 5% water, 35% methanol in isopropanol. Lipids were identified and their signal integrated using the Lipidsearch © software (version 4.2.27, Mitsui Knowledge Industry, University of Tokyo). Integrations and peak quality were curated manually before exporting.

NGN2

Culture and generation of human isogenic iPSCs

A control parental line was derived from a 75-year-old female (AG09173) with an APOE3/3 genotype by the Picower Institute for Learning and Memory iPSC Facility as first described [3]. Two ABCA7 LoF isogenic lines were derived from parental AG09173. ABCA7 p.Glu50fs*3, generated by Synthego (www.synthego.com), contains a premature termination codon in exon 3 (**Fig. S11A**), which to our knowledge has not been discovered in patients, but is functionally analogous to patient loss-of-function mutations.

ABCA7 p.Tyr622* contains a patient-derived mutation (Y622*) [4] and was generated in house by CRISPR-Cas9 genome editing. The CRISPR/Cas9-ABCA7-Y622* sgRNA plasmid was prepared followed by the published protocol [5]. In brief, a sgRNA sequence within 10 nucleotides from the target site was designed using the CRISPR/Cas9 Design Tool (http://crispr.mit.edu). The oligomer pairs (forward: 5'-CACCGCCCCTACAGCCACCCGGGCG-3' and reverse: 5'-AAACCGCCCGGGTGGCTGTAGGGGC-3') were annealed and cloned into pSpCas9-2A-GFP (PX458) plasmid (Addgene #48138). Plasmid DNA was submitted for Sanger sequencing to confirm correct ABCA7 sgRNA sequence (**Fig. S11B**).

AG09173 iPSCs were dissociated with Accutase (Thermo Fisher Scientific) supplemented with 10 µM ROCK inhibitor (Tocris) for electroporation using Amaxa and Human Stem Cell Nucleofector Kit I (Lonza). ~$5x10^6$ cells were resuspended in 100 µl of reaction buffer supplemented with 7.5 µg of CRISPR/Cas9-ABCA7 sgRNA plasmid and 15 µg of single-strand oligodeoxynucleotide (ssODN) template (5'-GGTGCGCGCCCCCAGGCCAATCCAGGAGCTGCACCCTAAGCTCCCGTTGCCTCTC ACAGCTGGGAGACATCCTCCCCTAGAGCCACCCGGGCGTCGTCTTCCTGTTCTTGGC AGCCTTCGCGGTGGCCACGGTGACCCAGAGCTTCCTGCTCAGCGCCTTCTTCTCCCG CGCCAACCTGG-3'). This reaction mixture was nucleofected with program A-23, resuspended with media supplemented with 10 µM ROCK inhibitor and seeded on MEF plates. Two days

after electroporation, cells were dissociated and filtered through Falcon polystyrene test tubes (Corning #352235), transferred to Falcon polypropylene test tubes (Corning #352063) and sorted by BD FACS Aria IIU in FACS Facility at the Whitehead Institute and seeded as single cells in media supplemented with 1X Penicillin-Streptomycin (P/S, Gemini Bio-products) and 10 µM ROCK inhibitor. After sufficient colony growth, each colony was transferred in part to a 12-well plate while the remainder was collected and used to extract genomic DNA (Qiagen DNeasy Blood & Tissue Kit, Cat. No. 69504) and screen for the Y622* mutation by sanger sequencing.

All lines used were confirmed to have normal karyotypes before use and periodically reviewed (Cell Line Genetics) (**Fig. S10C**). All human iPSCs were maintained at 37°C and 5% CO_2, in feeder-free conditions in mTeSR-1 medium (Cat #85850; STEMCELL Technologies) on Matrigel-coated plates (Cat # 354277; Corning; hESC-Qualified Matrix). iPSCs were passaged at 60–80% confluence using ReLeSR (Cat# 05872; STEMCELL Technologies) and reseeded between 1:6 and 1:24 (depending on desired density) onto Matrigel-coated plates.

<u>rTTA and NGN2 Virus production</u>

HEK293T cells (ATCC, Cat#CRL-3216) were maintained in DMEM/F-12, GlutaMAX (ThermoFisher, Cat#10565018), 10% fetal bovine serum (GeminiBio, SKU#100-106), 1% MEM Non-essential amino acids (Sigma, Cat#M7145), 1% sodium pyruvate (ThermoFisher, Cat#11360070), and 1% Penicillin-Streptomycin (GeminiBio, SKU#400-109). Cells were passaged for maintenance with TrypLE (ThermoFisher, Cat#12605010) at 70-80% confluence and reseeded 1:10 in 10 cm tissue culture plates.

For transfection, HEK293T cells were seeded at $5x10^6$ cells per 10 cm plate. Transfection mixtures containing the components required for 3rd generation lentiviral production (per 10 cm dish: 10 µg EF1a-rtTA-Hygro (Addgene #66810) or pLV-TetO-hNGN2-eGFP-Puro (Addgene #79823), 5 µg pMDL g/pRRE, 2.5 µg pRSV-Rev, 2.5 µg MD2.G, and 48 µL polyethyleneimine (1 mg/mL) in 600 uL OptiMEM (Fisher, Cat#51-985-034)). Mixtures were inverted 10X and incubated at RT for 20 min, then added dropwise to the dish. Transfection media was removed 16h later and replaced with 10 mL fresh media. Three days after transfection, media was collected and centrifuged at 3000 xg for 5 min at 4°C to pellet any contaminating cells. Supernatant was transferred to sterile Millex glass ultracentrifuge tubes and centrifuged at 25,000 rpm for 2 hours using a SW32Ti rotor in a Beckman Optima L-90K Ultracentrifuge. The pellets were resuspended in 1 mL PBS per 10 cm plate, and stored at -80°C until use.

<u>Lentivirus-mediated NGN2 induction in iPSCs and drug treatments</u>

iPSCs were dissociated into single cell suspension with Cell Dissociation Buffer (Life Technologies, Cat#13151-014), centrifuged at 300 xg for 5 min, and resuspended in mTeSR1

media with Rock inhibitor (Rockout; Abcam, ab285418). Single-cell suspension was plated in a 6-well plate coated with Matrigel for an optimized seeding density of 50-60% confluence 24 hours after plating. One day after plating, cells were co-transduced with 80 µL pLV-TetO-hNGN2-eGFP-Puro and 80 µL EF1a-rtTA-Hygro added in 1 mL fresh media per well and incubated overnight at 37°C. NGN2 expression was then induced 24 with addition of 2 mL fresh media supplemented with doxycycline (DOX, 1 µg/mL, final concentration) and Rock inhibitor. Puromycin (1 µg/mL) selection of non-NGN2 expressing cells was performed with media change 24 hours after induction, with continued DOX supplementation. After 24 hours of puromycin selection, immature neurons were re-plated onto PDL/laminin coated plates at $1x10^6$ cells/well on 6-well plates or $5x10^4$ cells/well on 96-well plates. Neurons were maintained in BrainPhys Neuronal Media (STEMCELL Technologies, Cat#05793) with Neurocult SM1 Neuronal Supplement (STEMCELL Technologies, Cat#05711), N2-supplement-A (STEMCELL Technologies, Cat#07152), and DOX (1 µg/mL) with half media changes every 3-4 days. Neuronal cultures were maintained for 28 days before experimentation.

iPSC-derived neurons were treated with cytidine 5'-diphosphocholine (CDP-choline; Millipore Sigma 30290) or DGAT1/2 inhibitors (PF-04620110; PF-06424439) to final concentrations of 100 µM and 1 µM, respectively, beginning at day 14 and repeated with each media change until 28 days matured.

<u>Lipidomics and Metabolomics of iPSC-derived neurons</u>

iPSC-derived neurons were washed once with cold PBS (Fisher; Cat#MT21040CM) and lifted off plate with a cell scraper in 1 mL cold PBS . Cells were centrifuged at 2000 xg for 5 min. PBS was removed, and cells were resuspended in 2 mL cold methanol for biphasic extraction as described above for *post-mortem* brain tissue. In addition to transferring the chloroform phase for lipidomics, the aqueous phase was also collected for metabolomics. Lipidomics and Metabolomics was performed in collaboration with the Harvard Center for Mass Spectrometry (HCMS).

Lipidomics was performed as described above for post-mortem samples. For metabolomics, samples were dried under nitrogen flow and resuspended in acetonitrile 50% in water. Resuspension volume was scaled to biomass, with the lowest biomass resuspended in 25 µL. 15 µL of each sample was transferred to glass microinserts for analysis. The remaining of the sample volumes were combined to create a pool sample used for MS2/MS3 data acquisition.

Samples were analyzed by LC-MS on a Vanquish LC coupled to an ID-X MS (Thermofisher Scientific). Five µL of sample was injected on a ZIC-pHILIC peek-coated column (150 mm x 2.1 mm, 5 micron particles, maintained at 40 °C, SigmaAldrich). Buffer A was 20 mM Ammonium Carbonate, 0.1% Ammonium hydroxide in water and Buffer B was Acetonitrile

97% in water. The LC program was as follows: starting at 93% B, to 40% B in 19 min, then to 0% B in 9 min, maintained at 0% B for 5 min, then back to 93% B in 3 min and re-equilibrated at 93% B for 9 min. The flow rate was maintained at 0.15 mL min-1, except for the first 30 seconds where the flow rate was uniformly ramped from 0.05 to 0.15 mL min-1. Data was acquired on the ID-X in switching polarities at 120000 resolution, with an AGC target of 1e5, and a m/z range of 65 to 1000. MS1 data is acquired in switching polarities for all samples. MS2 and MS3 data was acquired on the pool sample using the AquirX DeepScan function, with 5 reinjections, separately in positive and negative ion mode.

Electrophysiological recordings

Cells were placed in a recording chamber and perfused with oxygenated artificial cerebrospinal fluid (ACSF) contains (in mM) 125 NaCl, 2.5 KCl, 1.2 NaH_2PO_4•H_2O, 2.4 $CaCl_2$•$2H_2O$, 1.2 $MgCl_2$•$6H_2O$, 26 $NaHCO_3$ and 11 D-Glucose at a constant rate of 2 mL/min at ~32°C. Cells were visualized using infrared differential interference contrast (IR-DIC) imaging on an Olympus BX-50WI microscope.

Recordings were performed using Axon Multiclamp 700B and Clampex 11.2 (Molecular Devices). Action potentials were generated by injecting various steps of currents using current clamp configuration. Whole-cell currents were recorded from a holding potential of −80 mV by stepping to various voltages using voltage clamp configuration. Signals were filtered at 1 kHz using the amplifier's four-pole, low-pass Bessel filter, digitized at 10 kHz with a Digidata 1550B interface (Molecular Devices). Pipette solution contained (in mM) 120 K gluconate, 5 KCl, 2 $MgCl_2$•$6H_2O$, 10 HEPES, 4 ATP, 0.2 GTP. pClamp 11.2 (Molecular Devices) and GraphPad Prism 10 software suites were used for data acquisition and analysis. Data are presented as means ± standard errors of means (SEM).

Aβ Enzyme linked immunosorbent assays

Media was collected from 4 week old iNs and flash frozen. ELISAs were performed on thawed media according to manufacturer's instructions to measure Aβ40 (ThermoFisher Scientific, KHB3481) and Aβ42 (ThermoFisher Scientific, KHB3441) respectively. Amyloid levels were normalized to total protein content in media calculated using the Pierce BCA Protein assay (ThermoFisher Scientific, 23225) according to manufacturer's instructions.

Immunocytochemistry, LipidSpot stain, and Mitochondrial Health live cell stain

Neurons were fixed in 4% paraformaldehyde/4% sucrose in PBS at 4°C for 15 min at room temperature, washed 3X with PBS, then permeabilized with 0.1% Triton-X in PBS for 5 min at room temperature. Cells were blocked in 2% Bovine Serum Albumin (BSA, Fisher Bioreagents,

BP9703) in PBS for 20 min at room temperature, then incubated in primary antibodies diluted in blocking solution (1:500) overnight at 4°C. Cells were washed 3X for 5 min with PBS, then incubated in secondary antibodies diluted in blocking solution (1:1000) for 2 hours at room temperature. When used, LipidSpot (1:1000; Biotium 70069) was added along with secondary antibodies. Cells were washed 3X for 5 min with PBS, then incubated for 10 min with 1:2000 Hoechst 33342 (Invitrogen, H3570). Cells were washed 1X with PBS, and wells flooded with PBS for imaging.

HCS Mitochondrial Health kit (ThermoFisher, Cat#H10295) were used on live cells according to manufacturer's protocols. In brief, CellROX dye was added directly to cell media for a final concentration of 5µM, while 50 uL of media containing 1.5 µL MitoHealth dye was added to each well. Both stains were incubated on live cells for 30 min at 37°C. Cells were fixed in 4% paraformaldehyde in PBS, then co-stained according to immunocytochemistry procedures described above.

Seahorse Metabolic Assays

iPSCs were differentiated as described above directly on 96-well Agilent Seahorse XFe96/XF Pro cell culture microplates and matured for 28 days before assaying on a Seahorse XFe96 Analyzer. Seahorse XF Cell Mito Stress Test and Oxidation Stress Tests were performed according to manufacturer protocol with the following final drug concentration: Oligomycin, 2.5 µM; FCCP, 1 µM; Rotenone/Antimycin, 0.5 µM.

Confocal Image acquisition

All confocal images were acquired on a Zeiss LSM900. Acquisition settings were kept constant within each imaging batch (where conditions of interest were uniformly distributed across plates). The minimum and maximum z-plane was manually determined for each culture well, to accommodate differences in culture thickness. Cultures were imaged at 1 µm intervals along the z-axis.

Computational Methods

Single-Cell

Variant calling and ROSMAP subject selection

A total of 36 individuals were selected from the ROSMAP cohort, a longitudinal cohort study of aging and dementia in elderly nuns, priests, and brothers. Processed whole genome sequencing (WGS) variant call files for all ROSMAP samples, where available ($N = 1249$ sequencing samples), were downloaded from Synapse (syn11724057). Variant call data were downloaded for

chromosomes harboring SORL1, TREM2, ABCA7, ATP8B4, ABCA1, and ADAM10 (see Github Repository). When more than one WGS sample existed for a given subject, the sample with the higher Genomic Quality Score was chosen. Only samples that did not have sex mismatches and were consistent with previous array-based genotype data were considered (see syn12178037). Only variants that passed quality control ('FILTER_PASS') were considered.

Potential PTC (protein-truncating) variants in each of the aforementioned genes were flagged based on the following criteria: the variant had to be either a splice, missense, frameshift, nonsense, or premature start variant and be annotated as 'LOF' (loss of function). For ABCA7, this filtering captured known ABCA7 LoF risk variants from the literature, except for c.5570+5G>C, which was manually added to the filtered variants. Also see syn10901595 (https://www.synapse.org/#!Synapse:syn10901595) for information on WGS library preparation, quality control, variant annotations, and impact predictions. Annotated ABCA7 PTC variants are shown in Data S1.

The WGS data was used to identify 12 subjects who did not carry a known PTC variant in one of the aforementioned genes, other than in ABCA7, and for whom fresh-frozen post-mortem tissue was available for request from Rush University (termed 'LoF' samples). We also selected 24 individuals who do not carry a single ABCA7 PTC mutation or PTC variants in one of the aforementioned genes (termed 'control' samples). Control samples were matched on age, sex, and pathology.

Read Counting & Alignment

Library demultiplexing was performed using the BMC/BCC pipelines BioMicroCenter Software. Fast-q reads were aligned to the human genome GRCh38 and counted using the `cellranger count()` function from Cell Ranger version 6.1.2 (10x Genomics). Introns were included in counting to allow for the detection of unspliced transcripts, and the expected number of cells was set to 5000. Otherwise, Cell Ranger (v.6.1.2) default parameters were used. Counts across individual samples were then aggregated using a custom aggregation script (see GitHub Repository), resulting in a total of 150,456 cells.

Sample Swap Analysis

Sample swap analysis was performed using a previously established pipeline (MVV; QTLtools_1.1) [6], which compares allelic concordance between genomic and transcriptomic sequencing data. As input, we used the BAM files generated in the cellranger counting step and the chromosome 19 (the chromosome harboring ABCA7) variant call files (VCF). When comparing the concordance of BAM and VCF data for homozygous and heterozygous sites, the

expected WGS sample appeared as a clear outlier (more consistent along both dimensions than any of the other 1249 WGS ROSMAP samples) for all single cell samples (Fig. S1).

Cell filtering metrics

Prior to cell type annotation, we performed a series of quality control steps on the aggregated counts matrix. First, we filtered cells based on N_g, the number of genes for each cell where counts > 0, and kept cells for which $500 < N_g < 10000$. Next, we removed all cells with a high fraction of counts from mitochondrial-encoded genes. Mitochondrial fraction (M_f) is a commonly used per-cell metric to measure compromised nuclear integrity, with high fractions indicating low-quality nuclei, where C_{mt} is the total counts of mitochondrially-encoded genes for a cell, C_t is the total count of all genes for the same cell, and $M_f = \frac{C_{mt}}{C_t}$.

We fit a Gaussian mixture model (GMM) using sklearn's GaussianMixture implementation to $M'_f = \log_{10}(M_f + \epsilon)$, where ϵ is a small value added to M_f to avoid taking the logarithm of zero. A GMM models the data as independently sampled from a mixture of k Gaussian probability densities parameterized by a mean μ_k, a variance σ_k, and a mixture weight π_k indicating the proportion of the data derived from each component. The following log-likelihood function was maximized:

$$\ln[L(\theta|M'_f)] = \sum_{i=0}^{n} \ln\left(\sum_{j=0}^{k} \pi_j N(M'_{fi}|\mu_j, \sigma_j)\right)$$

The model with $k = 5$ components had the lowest Bayesian information criterion (BIC) score (GridSearchCV from sklearn.preprocessing), where:

$$\mathrm{BIC} = -2\ln(L) + k\ln(n)$$

where L is the maximized log-likelihood of the model, and n is the number of observations (cells). Finally, each cell i is assigned a component in k according to:

$$\mathrm{argmax}_k \left[\pi_k N(M'_{fi}|\mu_k, \sigma_k)\right]$$

Cells assigned to the component k with the highest mean M'_f scores were presumed to constitute a population of low-quality cells and were removed from further analysis. This initial filtering removed approximately 20,000 cells.

Considering all remaining cells in marker-gene expression space, where marker genes include only known cell type-specific genes for the major human PFC cell types, including astrocytes (159 markers), excitatory neurons (113 markers), inhibitory neurons (83 markers), microglia (97 markers), oligodendrocytes (179 markers), OPCs (143 markers), and vascular cells (124 markers) (Reference 1; Table S2), normalized to total library size $NC_m = \frac{C_m}{C_t}$, where NC_m and C_m are respectively the normalized and unnormalized count values for a given marker gene m and C_t is the total counts of all genes for the same cell. Next, we performed a memory-efficient implementation of singular value decomposition (IncrementalPCA from sklearn.decomposition) to transform cells from the marker-gene space (mean-centered and unit-variance) into a lower dimensional space (top 50 principal components sorted by variance). Visually, the cells formed a number of Gaussian-like clusters when projected onto the first two principal components. Under the assumption that each Gaussian cluster represented a different cell type in the brain, we again fit a GMM, as described above, except this time parameterized by a covariance matrix Σ_k instead of variance σ_k, to the projected data. The model with $k = 10$ and full covariance had the lowest BIC score. Each resulting cell cluster was enriched for a subset of major cell type markers in the brain, indicating clusters of astrocytes, microglia, OPCs, oligodendrocytes, excitatory neurons, inhibitory neurons, and a heterogeneous cluster of vascular cells.

To remove cells that were not well-explained by the GMM and likely represent low-quality cells, we next computed the per-cell log-probability given the model $L_i = \ln[P(x_i|\theta)]$, using Sklearn's GaussianMixture score_samples function, and removed cells with $L_i < -100$. We also removed two Gaussian clusters whose probability distributions constituted clear outliers compared to remaining clusters. The excluded cells had lower C_t and higher M_f compared to those that passed the log-likelihood filter, suggesting that the removed cells were indeed of low quality. As expected, when examining the data visually projected onto the first two principal components, this filtering removed many of the cells that were not visibly associated with a main Gaussian cluster. Together, this filtering removed an additional approximately 12,000 cells, leaving a total of 118,668 cells.

<u>Gene filtering metrics</u>

For the remaining downstream analysis we only considered genes that were both nuclear-encoded and protein-coding, which constituted a total of 19384 genes, based on annotation of ensembl GRCh38p12.

<u>Cell type annotations</u>

To remove variance explained by sequencing batch and individual-of-origin, we first applied the Python implementation of the Harmony algorithm [7] with individual-of-origin as an indicator

vector to the low-dimensional embedding of cells (first 50 principal components) remaining after the initial rounds of quality control described above. Next, we computed a neighborhood graph on the Harmony-corrected values in the PC embedding space, as implemented in the Scanpy Python package [8], using default parameters. Finally, we applied the Leiden graph-clustering algorithm to cluster this neighborhood graph of cells, using the Scanpy implementation of the Leiden algorithm[9].

We used the Scanpy `rank_genes_groups` function to compute top marker genes per Leiden cluster. Briefly, we assigned a major cell type label c to each Leiden cluster, where $c \in \{\text{'Ex', 'In', 'Ast', 'Mic', 'Oli', 'Opc', 'Vascular'}\}$, by computing the average cell-type-specific marker gene enrichment per Leiden cluster. Specifically, this is a vector of cell type signatures S_c for each Leiden cluster, where

$$S_c = \frac{1}{n} \sum_{i=0}^{n} \log_2 \left(\frac{I_i^{\text{In}}}{I_i^{\text{Out}}} \right)$$

where n is the total number of marker genes assigned to a cell type in c and I_i^{In} and I_i^{Out} indicate average gene expression values for a gene i for cells inside or outside a specific Leiden cluster, respectively. Then each Leiden cluster is assigned a label c by

$$\text{argmax}_c(S_c).$$

Finally, we sub-clustered cells from each major cell type using the Leiden clustering algorithm and examined distributions of mitochondrial fractions M_f and total counts C_t among subclusters s of the same cell type. Clusters were removed if

$$M_{f_s} > (2 \cdot \text{std}(M_f) + \overline{M_f})$$

or if

$$|C_{t_s}| > (2 \cdot \text{std}(C_t) + \overline{C_t}),$$

where M_{f_s} and C_{t_s} are respectively the mean M_f and C_t for all cells in a given Leiden cluster s and $\overline{M_f}$ and $\overline{C_t}$ are respectively the means of those values across all Leiden clusters, considering only clusters with the same cell type annotation c, because variance in M_f and C_t across major cell types can be biologically explained. Manual inspection of the removed clusters revealed that they tended to have fewer cells and low individual-level representations, and were not well-connected in the graph.

Individual-level filtering

After all rounds of quality control as described above, we noted a subset of individuals ($N = 6$) with very few cells (< 500). These subjects were removed from further analysis, resulting in 24 control individuals and 12 ABCA7 LoF individuals. None of these individuals carried ABCA7 PTC variants, and removing them did not substantially alter the distribution of clinical variables across genotypes.

Differential gene expression

Summed (pseudo-bulked) gene expression values were computed by matrix multiplication XI, where X is the gene x cell counts matrix and I is a cell x individual binary matrix indicating the individual-of-origin for each cell, resulting in 36 gene expression vectors for each of the six major cell types. For each cell type, only genes with a nonzero detection rate > 0.10 were considered for differential expression. Summed counts were normalized using the edgeR TMM method. The residual mean-variance trend not explained by the multivariate linear model (formalized below) was removed using Limma-Voom. Unknown sources of variance were captured in the model using surrogate variable analysis (SVA). Limma's *lmFit*, *eBayes*, and *topTable* functions were then used to estimate differential gene expression statistics, as reported in Data S3. The following model was fit for each cell type:

$$G_i = \beta_0 \times \text{ABCA7LoF} + \beta_1 \times \text{msex} + \beta_2 \times \text{nft} + \beta_3 \times \text{amyloid} + \beta_4 \times \text{age_death} + \beta_5 \times \text{PMI} + \beta_6 \times \text{batch} + \beta_7 \times \text{APOE4} + \beta_8 \times \text{SV0}$$

G_i refers to a vector of expression profiles of size 1×36 for a gene i in a given cell type. ABCA7LoF is a binary variable, encoding the presence of an ABCA7 variant predicted to cause loss of function (see Data S2). See Supplementary Text for descriptions of the remaining variables included in the model. SV0 refers to the first surrogate variable estimated from the data. The exact number of surrogate variables per cell type to include as additive terms in the model was estimated using *the num.sv()* function in R.

Gene-pathway projections

For each cell type, we computed a set of gene-wise scores quantifying the direction and statistical significance of gene expression changes (computed as part of the differential gene expression analysis) associated with ABCA7 LoF:

$$S = \text{sign}(\log_2 \text{FC}) \times -\log_{10}(\text{p-value})$$

where $\log_2 \mathrm{FC} > 0$ indicates up-regulation in ABCA7 LoF vs control. Top differentially expressed genes per cell type ($|S| > 1.3$) were projected from 6-dimensional score space, where each dimension captures ABCA7 LoF perturbation scores in one of the major cell types (Ex, In, Ast, Mic, Oli, OPC), into two dimensions, using the UMAP algorithm (using the `umap` Python package). Gene scores that were not detected in >10% of cells in a given cell type were set to 0.

We performed a grid search for Gaussian mixture parameters (parameter 1: number of components; parameter 2: covariance type) on the embedded cells (using the Python `sklearn` package) to assign genes to clusters in the 2D space. We proceeded with the model with the lowest BIC score, which had 15 components and a tied covariance matrix.

Each cluster was assigned representative pathway names by testing genes in that cluster for enrichment with Gene Ontology Biological Process pathways (Table S2) against the background of all genes in the embedding space, by hypergeometric enrichment (using the Python package `gseapy`). Pathways with an enrichment p-value < 0.01 were considered for cluster annotation.

Per-cell-type perturbation scores (S_c) for each cluster were computed as the average gene score S (for a given cell type) for all genes in that cluster. The statistical significance of each cell type-specific cluster score was assessed by permuting cluster assignments (100,000 permutations).

<u>Gene-set enrichment</u>

Genes were rank-ordered based on their scores S (see description in Gene-Pathway Projections). An R implementation of gene set enrichment analysis (GSEA) [10,11] (fast gene set enrichment analysis, fGSEA) was run with 10,000 permutations to estimate the statistical overrepresentation of gene sets in the WikiPathways databases (Table S2) within high-scoring ($|S|$), differentially expressed genes. Gene sets with a minimum size of 5 and a maximum size of 1000 were considered.

<u>Gene-pathway clustering using Kernighan-Lin heuristic</u>

To reduce the solution's computational search space, we reformulated the gene-pathway association problem as a bipartite graph G constructed from all the genes in the Leading Edge subset (LE) and their associated pathways. LE was defined as the set of 268 genes driving the enrichment signal for pathways that passed a significance threshold of $p < 0.05$ (fGSEA) in Con vs. ABCA7 LoF excitatory neurons. G was constructed from an $n \times m$ unweighted adjacency matrix, where n represented the number of LE genes and m represented the number of pathways associated with four or more LE genes, as specified in the WikiPathways database.

We chose to group gene-pathways into clusters of approximately equal size, making this a graph partitioning problem. We found that removing this constraint made the grouping results highly susceptible to outliers (Supplementary Text; Fig. S5C). Of the three graph partitioning algorithms tried, METIS and the Kernighan-Lin (K/L) algorithms had the lowest loss (Supplementary Text; Fig. S5B). Both METIS and K/L achieved very comparable losses (within 1.8% of each other, after 5.0×10^4 random initiations) and produced almost identical solutions (Rand index=0.98, after 5.0×10^4 random initiations) (Supplementary Text; Fig. S5B, D-F). We proceeded with the K/L algorithm for gene-pathway groupings as we found this algorithm to perform consistently better than METIS across a wider range of graph sizes (not shown).

The K/L algorithm was implemented in Python (see GitHub Repository) based on its original paper [12] and run with parameters set as $C = 0$, $KL_modified = True$, $random_labels = True$, $unweighted = True$, and $K = 50$ to partition G into 8 groups. We performed 5.0×10^4 random initiations on G and report the partitioning with the lowest loss among all initiations.

Gene-pathway graph layouts were computed using the `networkx` Python package with the spring layout algorithm, using 10,000 iterations. Layouts were visualized using the `matplotlib` `pyplot` package in Python.

Representative pathways for each cluster were inferred from the graph by averaging the ABCA7 LoF perturbation scores S for all genes in the cluster of interest sharing an edge with the pathway in question. Scores for pathways with intra-cluster degrees ≥ 5 were reported in the figures. Manually picked subsets of genes with the largest scores ($|S| > 1$) were reported in the figures. All gene statistics are reported in Data S3, and cluster assignments are reported in Data S6.

ABCA7 p.Ala1527Gly variant calling and gene-pathway clustering comparisons

We followed the same steps indicated in the section "Variant Calling and ROSMAP Subject Selection" to identify subjects who carried the p.Ala1527Gly variant and had snRNAseq performed on the PFC as part of a previous study. Differentially expressed genes were computed as described in the section "Differential Gene Expression," followed by gene set enrichment analysis of ABCA7 LoF-associated K/L clusters in excitatory neurons as described in the section "Gene Set Enrichment."

Image Processing

Confocal Image Processing

Image data were acquired at 8 or 16 bits, with voxel sizes of 1 µm x 0.62 µm x 0.62 µm (MitoHealth) or 1 µm x 0.31 µm x 0.31 µm (PLIN2, LipidSpot, Aβ, β-secretase). Image files extracted from the confocal Zeiss microscope (.czi format) were loaded into Python using the `aicsimageio` package and converted to floating-point format in the range $[0, 1]$. Confocal image acquisition settings were kept consistent within each imaging batch.

A pre-trained model ("cyto2" from Cellpose[13]) was applied to segment NeuN+ cell bodies per image. For images sampled at 0.62 µm along the xy-plane, segmentation on the NeuN channel in 3D produced the best results. For images sampled at 0.31 µm along the xy-plane, segmentation on the NeuN and Hoechst channels in 2D (xy), with subsequent stitching along the z-axis, produced the best results. Specific segmentation settings were determined for each imaging experiment. Segmentation quality was assessed manually, blinded to condition, and images with low-quality segmentations were discarded.

The model outputs per-voxel probabilities representing the Bernoulli probability that a given voxel lies within a cell (any cell) and per-voxel masks—recovered from flow vectors and from the pixel probabilities output by the model—representing regions of interest (cells). We leveraged these per-voxel probabilities to compute the expected fluorescence intensity $E(I_t)$ for our target channel t in each cell c. This is calculated as $E(I_t) = a \cdot b$, where a is a 1-dimensional vector containing the measured intensities for channel t across all n voxels annotated as part of the region of interest for cell c, and b is a 1-dimensional vector of the same length n. Each element b_i in b represents the normalized probability that the corresponding voxel i belongs to cell c, calculated as:

$$b_i = \frac{\Pr(v_i \in c)}{\sum_{j=0}^{n-1} \Pr(v_j \in c)}$$

This normalization ensures that the probabilities sum to 1, providing a weighted contribution of each voxel to the total expected fluorescence intensity for the cell.

A linear mixed-effects model was fit (using `mixedlm()` from the `statsmodels` package) to cell-level average fluorescence intensities, with treatment or genotype as a fixed effect and well-of-origin as a random effect; formalized as follows:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + u_j + \epsilon_{ij}$$

where:

- Y_{ij} is the observed fluorescence intensity for cell i in well j,
- β_0 is the intercept,
- β_1 is the coefficient for the fixed effect (treatment or genotype),
- X_{ij} is the fixed effect predictor (treatment or genotype) for cell i in well j,
- u_j is the random effect for well j, assumed to be normally distributed with mean 0 and variance σ_u^2,
- ϵ_{ij} is the residual error for cell i in well j, assumed to be normally distributed with mean 0 and variance σ^2.

Where indicated, measurements were combined over multiple differentiation batches (independent staining and imaging experiments). To this end, an equal number of cells from each experimental condition were sampled uniformly per batch, fluorescent values were z-scaled within that batch, and then combined. Indicator vectors for well-of-origin and batch-of-origin were included in the model. Before applying this linear transformation, per-cell per-image clipping was determined to be low (< 0.1%) and the response function of the confocal microscope was assumed to be linear.

For each condition, representative images were chosen from a single batch as the images closest to the mean fluorescence intensity for each condition. Voxels not belonging to a cell (i.e., not used in quantification) were masked prior to mean-projection for visualization.

For the confocal image analysis code, see GitHub Repository.

Molecular Dynamics Simulations

The initial structure of ABCA7 was obtained from the Protein Data Bank (PDB) in unbound-open and bound-closed conformations, PDB IDs 8EE6 and 8EOP, respectively. These experimentally solved structures harbor the G1527 variant. The A1527 structure was generated by mutating the glycine residue to alanine using pymol software.

The ABCA7 domain between residues 1517 and 1756 was embedded in a dipalmitoylphosphatidylcholine (DPPC) membrane using the CHARMM-GUI web server and oriented according to the Orientations of Proteins in Membranes (OPM) database. Four different simulations were performed using GROMACS 2022.3, as reported in Table S5. The CHARMM36M force field was used for all simulations.

The protein-membrane system was solvated in a cubic box with a minimum distance of 1.0 nm between the protein and the box edge, using the TIP3P water model. Energy minimization was performed using the steepest descent algorithm with a maximum force threshold of 1000 kJ/mol/nm to relieve any steric clashes or bad contacts. The system was equilibrated in six phases, each 125 ps long, to equilibrate volume (NVT) and pressure (NPT). The production run, 300 ns long, was performed in the NPT ensemble at 323 K using a v-rescale thermostat and 1 bar using the Parrinello-Rahman barostat. A 2 fs time step with h-bonds constraints was used with periodic boundary conditions applied in all directions. Long-range electrostatics were handled using the Particle Mesh Ewald (PME) method with a cutoff of 1.0 nm for non-bonded interactions.

RMSD was calculated to monitor the conformational stability of a given structure over the course of the simulation by comparing the position of C_α at time t under simulation to its reference position (in 8EOP or 8EE6).

Principal Component Analysis (PCA) was conducted to identify the major conformational changes during the simulation. The analysis involved the following steps:

1. A covariance matrix of the C_α
2. atom positional fluctuations was constructed using the `gmx covar` tool.
2. The covariance matrix was diagonalized to obtain the eigenvalues and eigenvectors, representing the principal components (PCs).
3. The trajectory was projected onto the first two principal components (PC1 and PC2) using the `gmx anaeig` tool to visualize the dominant motions.
4. A kernel density estimate (KDE) plot implemented in seaborn python3 was used to visualize the first two eigenvectors for each simulation, corresponding to 45%, 40%, 40% and 33% of variance in 8EOP-G1527, 8EE6-A1527, 8EE6-G1527 and 8EOP-A1527 respectively.

Visualization of the trajectories was carried out using VMD software.

LC-MS data analysis

<u>Lipidomic data from *postmortem* tissue and iNs</u>

Lipids were identified, and their signals integrated using the Lipidsearch © software (version 4.2.27, Mitsui Knowledge Industry, University of Tokyo). Integrations and peak quality were curated manually. For specific lipid species (see Figure 3A), statistical significance for differences in peak distributions between control and ABCA7 LoF were computed by two-sided unpaired t-test. For lipid classes (see Figure S9C), peaks were summed for individual lipid species of a given class, normalized to per-sample total lipid abundance, and compared by two-sided unpaired t-test.

Metabolomic data from iN

Data were analyzed using Compound Discoverer 3.2 (CD, Thermo Fisher Scientific). Identification was based on MS2/MS3 matching with a local mzVault library and corresponding retention time built with pure standards (level 1), or on mzCloud match (level 2). Each match was manually inspected. For specific metabolite comparisons, only metabolites with high-confidence annotation ('Level 1 ID', 'Level 2 ID', 'MasslistRT ID') that were not also detected in the background media were considered for analysis. Statistical significance for differences in normalized peak areas between control and ABCA7 LoF were computed by two-sided unpaired t-test.

Other data analysis

Oxygen consumption rate data analysis

The oxygen consumption rate (OCR) of cells was determined over time using a Seahorse XF Analyzer. Prior to analysis, OCR curves were visually inspected in a blinded manner to exclude wells that did not respond to drug injections. To calculate per-well total oxygen consumption for a given experimental period (e.g., under basal conditions prior to injections of uncouplers), integrals between specific experimental time points were computed from the OCR curve. The following measurements were made:

1. **Basal respiration** was computed as the total oxygen consumption prior to oligomycin injection.
2. **Proton leak** was computed as the total oxygen consumed after oligomycin injection and prior to FCCP injection.
3. **Maximal respiration** was computed as the total oxygen consumption after FCCP and prior to Rotenone + Antimycin injection.
4. **Relative uncoupling** was computed as the fraction of basal respiration attributed to proton leak.

723 5. **Spare respiratory capacity** was determined as the ratio of basal respiration to maximal
724 respiration.

Supplementary Text

Description of variables according to the Rush Alzheimer's Disease Center Codebook

1. **age_death.** Age at death
2. **amyloid.** Overall amyloid level - Mean of 8 brain regions. Amyloid beta protein identified by molecularly-specific immunohistochemistry and quantified by image analysis. Value is percent area of cortex occupied by amyloid beta. Mean of amyloid beta score in 8 regions (4 or more regions are needed to calculate). The 8 regions are hippocampus, entorhinal cortex, midfrontal cortex, inferior temporal gyrus, angular gyrus, calcarine cortex, anterior cingulate cortex, superior frontal cortex.
3. **braaksc.** Braak stage. Semi quantitative measure of neurofibrillary tangles. Braak Stage is a semi quantitative measure of severity of neurofibrillary tangle (NFT) pathology. Bielschowsky silver stain was used to visualize NFTs in the frontal, temporal, parietal, entorhinal cortex, and the hippocampus. Braak stages were based upon the distribution and severity of NFT pathology: Braak stages I and II indicate NFTs confined mainly to the entorhinal region of the brain; Braak stages III and IV indicate involvement of limbic regions such as the hippocampus; Braak stages V and VI indicate moderate to severe neocortical involvement.
4. **ceradsc - CERAD score.** Semiquantitative measure of neuritic plaques.CERAD score is a semiquantitative measure of neuritic plaques. A neuropathologic diagnosis was made of no AD (value 4), possible AD (value 3), probable AD (value 2), or definite AD (value 1) based on semiquantitative estimates of neuritic plaque density as recommended by the Consortium to Establish a Registry for Alzheimer's Disease (CERAD), modified to be implemented without adjustment for age and clinical diagnosis. A CERAD neuropathologic diagnosis of AD required moderate (probable AD) or frequent neuritic plaques (definite AD) in one or more neocortical regions. Diagnosis includes algorithm and neuropathologist's opinion, blinded to age and all clinical data. Value 1: definite AD, Value 2: probable AD, Value 3: possible AD, Value 4: no AD.
5. **cogdx - Final consensus cognitive diagnosis.** Clinical consensus diagnosis of cognitive status at time of death. At the time of death, all available clinical data were reviewed by a neurologist with expertise in dementia, and a summary diagnostic opinion was rendered regarding the most likely clinical diagnosis at the time of death. Summary diagnoses were made blinded to all post-mortem data. Case conferences including one or more neurologists and a neuropsychologist were used for consensus on selected cases. Value 1: NCI, No cognitive impairment (No impaired domains), Value 2: MCI, Mild cognitive impairment (One impaired domain) and NO other cause of CI, Value 3: MCI, Mild cognitive impairment (One impaired domain) AND another cause of CI, Value 4: AD, Alzheimer's disease and NO other cause of CI (NINCDS PROB AD), Value 5: AD,

Alzheimer's disease AND another cause of CI (NINCDS POSS AD), Value 6: Other dementia. Other primary cause of dementia

6. **msex – Sex.** Self-reported sex, with "1" indicating male sex. 1 = Male 0 = Female
7. **nft - Neurofibrillary tangle burden.** Neurofibrillary tangle summary based on 5 regions. Neurofibrillary tangle burden is determined by microscopic examination of silver-stained slides from 5 regions: midfrontal cortex, midtemporal cortex, inferior parietal cortex, entorhinal cortex, and hippocampus. The count of each region is scaled by dividing by the corresponding standard deviation. The 5 scaled regional measures are then averaged to obtain a summary measure for neurofibrillary tangle burden.
8. **pmi - Post-mortem interval.** Time interval in hours from time of death to autopsy. *Post-mortem* interval (PMI) refers to the interval between death and tissue preservation in hours.

Choosing a partitioning heuristic for gene-pathway grouping

Methods

The heatmap in Fig. S5A highlights how frequently pathways within a pathway database, such as WikiPathways, share gene members. On average, every pathway shown in Fig. S5A shares at least one gene with approximately 40% of the other pathways, highlighting that there is redundancy in this matrix that could be summarized in simpler terms.

To summarize redundant gene-pathway information into a limited number of non-redundant gene-pathway groups, we reformulated the gene-pathway association problem as a bipartite graph G constructed from all the genes in the Leading Edge subset (LE) and their associated pathways. LE was defined as the set of 268 genes driving the enrichment signal for pathways that passed a significance threshold of $p < 0.05$ (fGSEA) in Con vs. ABCA7 LoF excitatory neurons. G was constructed from an $n \times m$ unweighted adjacency matrix, where n represented the number of LE genes and m the number of pathways associated with four or more LE genes, as specified in the WikiPathways database.

Graph partitioning involves segmenting the vertices of a graph into equal-sized partitions, optimizing for the minimal number of interconnecting edges (i.e., "total cut size"). We tested three prominent graph partitioning techniques, as outlined by Elsner (1997)[14], to approximate optimal partitioning. These methods include:

1. **Recursive Spectral Bisection:** Implemented in Python using the numpy linear algebra package, this method was executed for $\log_2(N)$ iterations, yielding $N = 8$ partitions. A detailed description of the algorithm can be found in Elsner (1997)[14].

2. **Multilevel Graph Partitioning:** Leveraging the METIS software package[15] in Python using the following parameters: `nparts=8`, `tpwgts=None`, `ubvec=None`, `recursive=False`.
3. **Kernighan-Lin (K/L) Algorithm:** Based on its original paper[12], this algorithm was implemented in Python and run with parameters set as $C = 0$, `KL_modified=True`, `random_labels=True`, `unweighted=True`, and $K = 50$.

Additionally, the Spectral Clustering algorithm, a commonly used clustering method, was applied using the `SpectralClustering()` function from the `sklearn` Python package with default parameters, apart from `n_clusters=8` and `assign_labels='kmeans'`. We stipulated eight clusters for each algorithm, as qualitatively, this resolution seemed to strike a good balance to summarize main biological effects.

For benchmarking purposes, the three graph partitioning techniques and the spectral clustering algorithm were evaluated by segmenting graph G into eight gene-pathway clusters using the respective algorithms. Spectral clustering was run over 1,000 initiations, while K/L and METIS were run over 50,000 iterations because their solutions were slightly more variable across runs. The deterministic bisection method was run only once. A randomized graph partitioning benchmark was also computed by permuting the eight cluster labels of approximately equivalent size for 1,000 initiations. Average losses were computed per algorithm on all initiations. The benchmarking process and source code are available at: GitHub Repository.

Results

Spectral clustering performed significantly better than all other algorithms based on the loss (Fig. S5B). This was expected, as spectral clustering does not place a constraint on cluster size. Spectral clustering results were characterized by a single large cluster and many small clusters (Fig. S5C), indicating that this clustering algorithm was highly susceptible to outliers and suggesting that graph partitioning, which imposes the constraint of equal partitioning, was a better approach to the problem of grouping genes and pathways into biologically informative groups. Indeed, all three graph partitioning algorithms divided the graph into more uniformly-sized groups (Fig. S5C). Among the partitioning algorithms, K/L and METIS produced the most uniformly sized groups (Fig. S5C) and also had significantly lower losses compared to the spectral bisection algorithm (Fig. S5B). K/L and METIS solutions were very similar, with their respective best solutions (lowest loss) having an average Jaccard similarity index of 0.91 on the diagonal (Fig. S5D,E). K/L and METIS solutions were also consistent across pairwise random initiations, both when comparing within K/L or METIS solutions (Rand Index=0.87 and 0.91, respectively) and when comparing all pairwise K/L and METIS solutions (Rand Index=0.88) (Fig. S5F).

Overall, these results indicate the importance of non-redundant gene-pathway groupings to interpret biological effects. They also indicate that for some gene-pathway graphs, such as the one in this study, graph partitioning is a better approach than clustering.

Molecular Dynamics Simulations Results

The root mean square deviation (RMSD) analysis was performed to assess the conformational stability of ABCA7 in different states and mutations (**Fig. 2I**; **Fig. S6A**). The RMSD values of the C_α atoms were calculated over the 300 ns simulation period for both the closed and open conformations, with G1527 and A1527 mutations.

The RMSD plot for ABCA7 in the closed conformation with the G1527 mutation (**Fig. 2J, K**) exhibited large fluctuations throughout the simulation (**Fig. 2L**). The distribution of RMSD values was broad, indicating significant conformational variability (**Fig. 2L; Fig. S6E**). The high RMSD values suggest that the G1527 mutation in the closed state induces larger flexibility, leading to substantial deviations from the initial structure.

In contrast, the RMSD analysis for the closed conformation with the A1527 mutation (**Fig. 2J,K**) showed minor fluctuations (**Fig.2L**). The RMSD values were consistently lower than those observed for the G1527 mutation, and the distribution of RMSD values followed a normal distribution (**Fig. 2L; Fig. S6E**). This indicates that the A1527 mutation confers lower local flexibility and greater local stability to the closed conformation of ABCA7.

For the open conformation of ABCA7, both G1527 and A1527 mutations (**Fig. S6B,C**) exhibited minor RMSD fluctuations (**Fig. S6D**). The RMSD values for these mutations were significantly lower compared to the closed conformation with the G1527 mutation (**Fig. S6E**). The RMSD distributions for both mutations in the open conformation were narrow (**Fig. S6D; Fig. S6E**), suggesting local stable conformational behavior.

Principal component analysis (PCA) was performed to further investigate the conformational dynamics of ABCA7 in the different states and mutations. The PCA of the closed conformation with the G1527 mutation showed a widespread along the first principal component (PC1), reflecting the large conformational space sampled during the simulation (**Fig. 2M**). This is consistent with the broad RMSD distribution, indicating high flexibility and multiple conformational states.

In contrast, the closed conformation with the A1527 mutation showed a much more confined spread along PC1 and PC2 (**Fig. 2M**). This limited spread corresponds to the minor RMSD fluctuations observed, suggesting a locally more stable and less flexible structure.

For the open conformation, PCA indicated two distinct clusters for both G1527 and A1527 mutations (**Fig. S6F**). These clusters represent two closely related conformational states, indicating that both mutations stabilize the open conformation into two main conformational states with minimal fluctuations.

Supplementary Figure Legends

Fig. S1. Overview of Human snRNA-Sequencing Cohort. (A) Sanger sequencing of ABCA7 LoF variants in prefrontal cortex genomic DNA samples from 3 ABCA7 LoF carriers and 3 controls from the snRNA-seq cohort. Sequencing confirmed heterozygosity of the indicated variant in LoF samples, with variant location marked by a black box. (B) Example plots validating matches between whole genome sequencing (WGS) and snRNA-seq libraries. Each plot shows the concordance of homo- and heterozygous SNP calls between WGS and snRNA-seq data for a single individual. Matches between WGS SNP calls and snRNA-seq BAM inferred SNP calls are indicated by extreme outliers. Expected (i.e., correct) matches are indicated in blue/purple. (C) Protein levels from post-mortem human prefrontal cortex (see Table S2 for external dataset used) showing ABCA7 protein levels (left) and NeuN (RBFOX3) levels (middle) for a subset of individuals in the snRNA-seq cohort (N=6 control and N=4 ABCA7 LoF carriers). The right panel shows NeuN (RBFOX3) protein levels by genotype in all available control samples (N=180) vs. ABCA7 LoF carriers (N=5). (D) Distributions of continuous metadata variables (see Supplementary Text for descriptions) for control individuals (N=24) vs. ABCA7 LoF carriers (N=12). For panels C and D, boxes indicate dataset quartiles per condition, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). (E) Distributions of discrete metadata variables for control individuals (N=24) vs. ABCA7 LoF carriers (N=12). Con=control, LoF=ABCA7 loss-of-function. P-values in panels C and D were computed by two-sided Wilcoxon rank sum test. P-values in panel E were computed by two-sided Fisher's exact test.

Fig. S2. Overview of snRNA-sequencing Batch Correction, Data Quality, and Cell Type Annotations. (A,B) Distributions of continuous (A) and discrete (B) metadata variables by sequencing batch. (C) Two-dimensional UMAP projection of snRNA-seq single cells from gene expression space, colored by batch 1 (pink) or batch 2 (blue) after all rounds of quality control. (D) Two-dimensional UMAP projection of snRNA-seq single cells from gene expression space, colored by individuals of origin after all rounds of quality control. Each individual is indicated by a different color. (E) Two-dimensional UMAP projections of individual cells from gene expression space, colored by Leiden clusters. (F) Average marker gene expression (per-cluster mean log(fold-change)) for all marker genes for the cell type indicated along the x-axis. Log(fold-changes) are computed for the cluster of interest vs. all remaining clusters. Reference 1 (Table 2) marker genes were used. (G) Cladogram visualizing subcluster relationships based on pairwise distances between per-cluster gene expression profiles. (H) Average marker gene expression profiles (x-axis) per major cell type annotation (y-axis) for two marker gene references (Table 2). (I) Per-cell distribution of select marker gene expression by cell type. Y-axis indicates log counts. (J) Median number of cells per cell type per individual. (K) Cell type fraction by individual. (L,M) Individual-level gene expression correlations by cell type. For all panels, p-values for all continuous variables were computed by two-sided Wilcoxon rank sum test. P-values for all discrete variables were computed by two-sided Fisher's exact test. For A, I,

M boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile).

Fig. S3. Single-nuclear RNA-sequencing of Human Post-mortem Prefrontal Cortex Reveals Cell Type-specific Gene Changes in ABCA7 LoF Variant Carriers. (A) Cartoon overview of gene score analysis strategy (see Methods) (Created with BioRender.com). Genes from cell-type-specific ABCA7 LoF perturbation score space were projected onto the first two UMAP dimensions, followed by clustering and pathway enrichment to provide a unified visual overview of gene expression changes associated with ABCA7 LoF across all major cell types. (B, C) Projections of gene scores onto the top two dimensions computed by UMAP, showing up-regulated genes per cell type ($S > 1.3$) (B) or down-regulated genes per cell type ($S < -1.3$) (C). (D) Perturbation patterns of gene clusters from Fig. 1G in indicated cell types shown as histograms (astrocytes (Ast), red; excitatory neurons (Ex), light blue; inhibitory neurons (In), green; microglia (Mic), navy blue; oligodendrocytes (Oli), peach; oligodendrocyte precursor cells (Opc), grey). Gene scores S are plotted on the x-axis and frequency is plotted on the y-axis.

Fig. S4. Neuronal Expression of ABCA7 in the Post-mortem Human Brain. (A) Per cell type ABCA7 detection rate of major cell types in the post-mortem PFC as quantified by snRNA-seq. (B) Normalized expression of indicated gene in glial cells (per-individual mean expression profiles across Oli, Opc, Ast, Mic) vs. neuronal cells (per-individual mean expression profiles across Ex and In) from post-mortem snRNA-seq data. (C) Normalized expression of indicated genes in NeuN- vs. NeuN+ cells (N=6 individuals, from [16]; see Table 2). All p-values are computed by paired two-sided t-test. Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile).

Fig. S5. Benchmarking Partitioning and Clustering Algorithms for Gene-Pathway Grouping. (A) Jaccard indices quantifying overlap of genes for all 111 pathways in Fig. 2B (see Methods; Supplementary Text). (B) Average loss (total cut size; see Methods) associated with applying each algorithm (spectral clustering (SC), METIS, Kernighan-Lin (K/L), spectral bisection (SB), or random permutation) to G (with 379 vertices; see Methods) over 1000 initiations (SC, random permutation) or 5×10^5 initiations (METIS, K/L). The SB implementation is deterministic and was run only once. Error bars indicate the standard deviation. (C) Unweighted adjacency matrix for G sorted by labels assigned by the indicated algorithm. Red indicates the presence of an edge between two vertices. For each algorithm, labels corresponding to the best initiation (lowest loss) over 1000 initiations (SC, random permutation) or 5×10^5 initiations (METIS, K/L) are shown. (D) Pairwise labeling consistency for the best K/L initiation and the best METIS initiation. Cluster labels corresponding to each are shown on the X- and Y-axes, respectively. Each color entry indicates the fraction of shared vertices per cluster across two initiations. Consistency is quantified using the Jaccard Index (JI).

$\mathrm{JI} = \frac{|A \cap B|}{|A \cup B|}$, where A and B are two sets (i.e., cluster A from initiation #1 and cluster B from initiation #2). (E) Same as (D), but comparing the best K/L initiation against the best random permutation initiation. (F) Average Rand index (RI) for all pairwise initiations from (B). "METIS," "Kernighan-Lin," and "Permuted" labels on the Y-axis indicate the average RI (consistency across two sets of labels) for all combinations of initiations within the specified algorithm. "METIS-K/L" indicates the average RI for all combinations of initiations across the METIS and Kernighan-Lin algorithms. Error bars indicate standard deviations. ($\mathrm{RI} = \frac{\text{number of agreeing vertex pairs}}{\text{number of vertex pairs}}$).

Fig. S6. Molecular Dynamics Simulations of ABCA7 open conformations with p.Ala1527Gly substitution. (A) Open conformation ABCA7 protein structure. ABCA7 domain between residues 1517 and 1756 used for simulations is shown in yellow. (B) Expanded yellow domain (inset from *A*), with A1527 variant (light grey) and G1527 variant (purple). (C) Expanded inset from *A* with residues of interest rendered. (D) Root mean squared deviations of open conformation domains from *B* with A1527 (light grey) or G1527 (purple) under simulation. Structural deviations over time were computed with respect to reference open structures from *B*. (E) Violin plot indicating average C_α atom positional fluctuations over time. (F) Projection of C_α atom positional fluctuations under simulation onto the first two principal components, for open conformation domain from *B* with A1527 (top, light grey) or G1527 (bottom, purple).

Fig. S7. Lipid and Mitochondrial Transcriptional Perturbations in Excitatory Neurons from Human ABCA7 LoF Variant-Carriers vs. Non-Carriers. (A) Lipid synthesis and storage pathways perturbed in ABCA7 LoF excitatory neurons vs. control as measured by snRNA-seq on the post-mortem human PFC. Enrichments of biological processes were computed using FGSEA. Red = enrichment > 0, Blue = enrichment < 0. * = $p < 0.05$. (B) Schematic model showing anabolic processes feeding from the TCA cycle towards fatty acid (FA) and triglyceride (TG) synthesis. DG = diacylglyceride, PA = phosphatidic acid, PC = phosphatidylcholine, PE = phosphatidylethanolamine, PS = phosphatidylserine. * = differentially expressed in ABCA7 LoF vs. control excitatory neurons from post-mortem human brain at $p < 0.05$ and $\log \mathrm{FC} < 0$. (C) *β*-oxidation and TCA pathways perturbed in ABCA7 LoF excitatory neurons vs. control as measured by snRNA-seq on the post-mortem human PFC. Enrichments of biological processes were computed using FGSEA. Red = enrichment > 0, Blue = enrichment < 0. * = $p < 0.05$. (D) Schematic model showing catabolic processes feeding into the TCA cycle and oxidative phosphorylation with key genes from (C) highlighted in red or blue. * = $p < 0.05$. For (A, C,) boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). (E, F) Transcript levels of ACLY (E) and SCP2 (F) assessed in

post-mortem human PFC by RNAscope. Transcript counts per SLC17A7+ cell are reported in each bar chart. $N = 8$ individuals per genotype. Per-cell Wilcoxon rank-sum p-values are reported.

Fig. S8. Differentiating and Profiling iPSC-Derived Neurons Harboring ABCA7 PTC Variants. (A) Sanger sequencing chromatogram confirming single nucleotide insertion in ABCA7 exon 3 to introduce a premature termination codon into the isogenic iPSC line ABCA7 p.Glu50fs*3 using CRISPR-Cas9 gene editing. (B) Sanger sequencing chromatogram confirming patient single nucleotide polymorphism in ABCA7 exon 15 to introduce a premature termination codon into the isogenic iPSC line ABCA7 p.Tyr622* using CRISPR-Cas9 gene editing. (C) Normal karyotypes were observed for control, ABCA7 p.Glu50fs*3, and ABCA7 p.Tyr622* isogenic iPSC lines. (D) iPSCs were plated at low density for NGN2 viral transduction. Expression of NGN2 was driven by doxycycline (DOX) induction with puromycin (PURO) selection, then re-plated to match neuronal densities. Neurons were maintained for 4 weeks (DIV 28) before experimentation (Created with BioRender.com). (E) Neuronal marker gene expression in 2 and 4-week matured iNs. (F) Representative sweeps of whole-cell current flow of inward (upper panel) and outward (lower panel) current recordings from WT 4-week-old neurons. (G) Quantification of (F). (H) Resting membrane potential (mV) of 4-week-old WT, ABCA7 p.Tyr622*, and ABCA7 p.Glufs*3 neurons. (I) Rheobase (pA) of 4-week-old WT, ABCA7 p.Tyr622*, and ABCA7 p.Glufs*3 neurons. (J) Action potential frequency of 4-week-old WT, ABCA7 p.Tyr622*, and ABCA7 p.Glufs*3 neurons with indicated current injections. For panels F-J: WT: $n = 24$; Y622: $n = 13$; G2: $n = 23$. For all panels: $P* < 0.05$, $P** < 0.001$. Graphs are mean ± SEM.

Fig. S9. Quantification of Aβ42 in iPSC-Derived Neurons Harboring ABCA7 PTC Variants. (A) Quantification of neuronal Aβ42 fluorescence intensity. P-values were computed by a linear mixed-effects model on per-NeuN+ volume averages, including well-of-origin as a random effect. $N = 16$ (WT; 2261 cells), $N = 8$ (p.Tyr622*; 1466 cells), $N = 6$ wells (p.Glu50fs*3; 999 cells) from 4-week-old iNs. Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). Individual data points represent per-well averages of cell-level intensities. (B) Representative images per condition showing mean-intensity projections of the entire image (NeuN+) and projections within NeuN+ volumes considered for quantification (Aβ42). Representative images for the Aβ42 channel were processed with condition-wide percentile-based background subtraction and thresholding. Representative images of cell soma underwent per-image percentile-based background subtraction and thresholding, reflecting the segmentation methodology.

Fig. S10. Lipidomic Analysis in ABCA7 LoF vs. Control iNs. (A) Pairwise Pearson correlation of iN lipidomic profiles. (B) Correlation of lipidomic scores ($S = \text{sign}(\log(\text{fold-change})) \times -\log_{10}(\text{p-value})$, T-test) between WT and ABCA7 p.Glu50fs*3

iNs by batch. Top consistent triglyceride (TG), phosphatidylcholine (PC), and ceramide (Cer) lipid changes, where $|S| > 1.3$ in both batches, shown in green. Grey error bar indicates 95% confidence interval for simple linear model fit. (C) Relative TG and PC abundances in post-mortem human brain. Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). (D) Correlation of metabolomic scores (computed and plotted as in B) for both differentiation batches. (E) Average expression and nonzero detection rate of select lipid oxidation genes in major cell types in the human brain, assessed by snRNA-seq of post-mortem PFC.

Fig. S11. Analysis of Oxygen Consumption Rates in ABCA7 LoF vs. Control iNs. (A) Example oxygen consumption rate (OCR) curves from Batch 1 of the two differentiation batches used for analysis in Fig. 4G. The line plot indicates the per-condition mean estimator, and the error bars indicate the 95% confidence interval. (B) Representative per-well traces from (A). (C) Schematic indicating measurement of maximal and basal oxygen consumption to compute SRC, as shown in (D) for WT, ABCA7 p.Glu50fs*3, and ABCA7 p.Tyr622* iNs. P-values computed by independent sample t-test. N wells = 18 (WT), 17 (p.Tyr622*), 13 (p.Glu50fs*3) across two independent differentiation batches and Seahorse experiments. (E) Relative uncoupling measured for two independent iN differentiation batches and separate Seahorse experiments shown combined in Fig. 4G. P-values computed by independent sample t-test. Batch 1 (left); N wells = 10 (WT), 7 (p.Tyr622*), 7 (p.Glu50fs*3). Batch 2 (right); N wells = 8 (WT), 10 (p.Tyr622*), 6 (p.Glu50fs*3) shown per differentiation batch. For (D, E) boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). (F) Per-batch cell-level MioHealth fluorescence intensities (related to Fig. 4H).

Fig. S12. Lipid and Mitochondrial Effects of Treatment with CDP-choline. (A) Per-cell correlation of average PLIN2 and LipidSpot fluorescent intensities shown as a density plot. (B) Per-batch LipidSpot fluorescence intensities (related to Fig. 5A) in ABCA7 p.Tyr622* iNs treated with CDP-choline or H20 vehicle control. X-axis indicates z-scaled log-fluorescence intensity. (C) Example oxygen consumption rate (OCR) curves used for analysis in Fig. 5. The line plot indicates the per-condition mean estimator, and the error bars indicate the 95% confidence interval. (D) Representative per-well traces from (C). (E) Quantification of SRC from curves in (D). P-values computed by independent sample t-test. N wells = 6 (p.Tyr622* + H20), 8 (p.Tyr622* + CDP-choline). Boxes indicate per-condition dataset quartiles, and whiskers extend to the most extreme data points not considered outliers (i.e., within 1.5 times the interquartile range from the first or third quartile). (F) Per-batch cell-level MioHealth fluorescence intensities (related to Fig. 5C) in ABCA7 p.Tyr622* iNs treated with CDP-choline or H20 vehicle control. X-axis indicates z-scaled fluorescence intensity.

1061 # Supplementary Tables

1062 **Table S1.** Annotation of ABCA7 loss of function variants used in this study

rsID	HGVS.c	HGVS.p	annotation	AD association	N in cohort
rs113809142	c.4416+2T>G	NA	splice donor variant [17]	Steinberg et al (2015), Nature Genetics, Table 1 [18]	1
rs200538373	c.5570+5G>C	NA	splice region variant [17,18]	Steinberg et al (2015), Nature Genetics, Table 1 [18]	4
rs538591288	c.4208delT	p.Leu1403fs	frameshift variant [17]	Steinberg et al (2015), Nature Genetics, Table 1 [18]	1
rs547447016	c.2126_2132delAGCAGGG	p.Glu709fs	frameshift variant [17]	Steinberg et al (2015), Nature Genetics, Table 1 [18]	4
rs201060968	c.3641G>A	p.Trp1214*	stop gained	NA	1
19_1053362_G_A	c.3255G>A	p.Trp1085*	stop gained	NA	1

1063

1064 **Table S2.** External datasets used

Description	**Access**	**Reference**
Post-mortem human PFC proteomic data	https://www.synapse.org/#!Synapse:syn21449368	[19]
Reference 1: Cell type specific marker genes for human brain	https://osf.io/vn7w2/	[20]
Reference 2: Cell type specific marker genes for human brain	https://osf.io/vn7w2/	[21]
Gene Ontology Biological Process 2023	https://maayanlab.cloud/Enrichr/#libraries	NA
NeuN+/- bulk RNA-sequencing from *post-mortem* human brain	https://osf.io/vn7w2/	[16]
WikiPathways 2019 Human	https://maayanlab.cloud/Enrichr/#libraries	NA
snRNAseq from *postmortem* human PFC from p.Ala1527Gly variant-carriers and controls	https://www.synapse.org/#!Synapse:syn52293417	[22]

1065

1066 **Table S3.** Antibodies used

Antibody name	Company	Catalog No.
NeuN	Synaptic Systems	266004
Tuj1	BioLegend	MMS-435P
SM312 (pan-axonal marker)	Biolegend	837904
MAP2	Biolegend	822501
PLIN2	Proteintech	15294-1-AP

Aβ42 (12F4)	BioLegend	805503
BACE	Millipore	MAB5308

Table S4. PCR/Sanger sequencing (SS) primers

Primer	Sequence
rs547447016_FOR	5'-ACGCTGGCCTGGATCTACTC-3'
rs547447016_REV	5'-TGCATGCGTGTGCCAAGAAG-3'
chr19.1053362G>A_rs201060968_FOR	5'-CTGAAGCACCCCTTTGTCCAC-3'
chr19.1053362G>A_rs201060968_REV	5'-GAAAGCGCTTGAGAAGCAGGG-3'
chr19.1053362G>A_REV_SS	5'-GCTGCTCATAAACACGCTATTCATCCTTC-3'
rs201060968_FOR_SS	5'-CATTGCTGGCCTAGACGTAA-3'
ABCA7_p.Glu50fs*3_FOR	5'-GTGACGAAAGCGTTAAGCCC-3'
ABCA7_p.Glu50fs*3_REV	5'-GCAGTGGCTTGTTTGGGAAG-3'
ABCA7_p.Tyr622*_FOR	5'-CTGGTTCTGGTGCTCAAG-3'
ABCA7_p.Tyr622*_REV	5'-CCTACGGCAGACGTCTTCAG-3'

Table S5. Experimentally-determined 3D ABCA7 structures used in molecular dynamics simulations

System	**PDB ID**	**State**
CLOSE-G1527	8EOP	HOLO
CLOSE-A1527	8EOP	HOLO
OPEN-G1527	8EE6	APO
OPEN-A1527	8EE6	APO

Extended Data Descriptions

Data S1. ABCA7 PTC variants represented in snRNAseq cohort (related to **Figure 1**).

Data S2. ROSMAP cohort metadata for selected individuals in this study (related to **Figure 1**).

Data S3. Differential expression statistics for each cell type (related to **Figure 1**).

Data S4. Cluster pathway enrichments (related to **Figure 1**).

Data S5. Cluster cell type perturbation scores (related to **Figure 1**).

Data S6. Kernighan-Lin cluster assignments (related to **Figure 2**).

Data S7. *Post-mortem* PFC lipidome statistics (related to **Figure 3**).

Data S8. NGN2 lipidome statistics (related to **Figure 3**).

Data S9. NGN2 metabolome statistics (related to **Figure 4**).

Supplementary References

1. Mathys, H. *et al.* Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).
2. Miraldi, E. R. *et al.* Molecular network analysis of phosphotyrosine and lipid metabolism in hepatic PTP1b deletion mice. *Int Bio (Cam)* **5**, 940–963 (2013).
3. Lin, Y.-T. *et al.* APOE4 Causes Widespread Molecular and Cellular Alterations Associated with Alzheimer's Disease Phenotypes in Human iPSC-Derived Brain Cell Types. *Neuron* **98**, 1294 (2018).
4. De Roeck, A., Van Broeckhoven, C. & Sleegers, K. The role of ABCA7 in Alzheimer's disease: evidence from genomics, transcriptomics and methylomics. *Acta Neuropathol.* **138**, 201–220 (2019).
5. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
6. Fort, A. *et al.* MBV: a method to solve sample mislabeling and detect technical bias in large combined genotype and sequencing assay datasets. *Bioinformatics* **33**, (2017).
7. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
8. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 1–5 (2018).
9. Traag, V., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. (2018) doi:10.1038/s41598-019-41695-z.
10. Leek, J. T. & Storey, J. D. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet.* **3**, 1724–1735 (2007).
11. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).
12. Kernighan, B. W. & Lin, S. An efficient heuristic procedure for partitioning graphs. *Bell Syst. Tech. J.* **49**, 291–307 (1970).
13. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
14. Elsner, U. *Graph Partitioning: A Survey*. (1997).
15. Karypis, G. & Kumar, V. METIS: A Software Package for Partitioning Unstructured Graphs, Partitioning Meshes, and Computing Fill-Reducing Orderings of Sparse Matrices. (1997).
16. Welch, G. M. *et al.* Neurons burdened by DNA double-strand breaks incite microglia activation through antiviral-like signaling in neurodegeneration. *Sci Adv* **8**, eabo4662 (2022).
17. Allen, M. *et al.* loss-of-function variants, expression, and neurologic disease risk. *Neurol Genet* **3**, e126 (2017).
18. Steinberg, S. *et al.* Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. *Nat. Genet.* **47**, 445–447 (2015).
19. Johnson, E. C. B. *et al.* Large-scale proteomic analysis of Alzheimer's disease brain and

cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. *Nat. Med.* **26**, 769–780 (2020).

20. Wang, D. *et al.* Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, (2018).
21. Franzén, O., Gan, L.-M. & Björkegren, J. L. M. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database* **2019**, baz046 (2019).
22. Mathys, H. *et al.* Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology. *Cell* **186**, 4365–4385.e27 (2023).

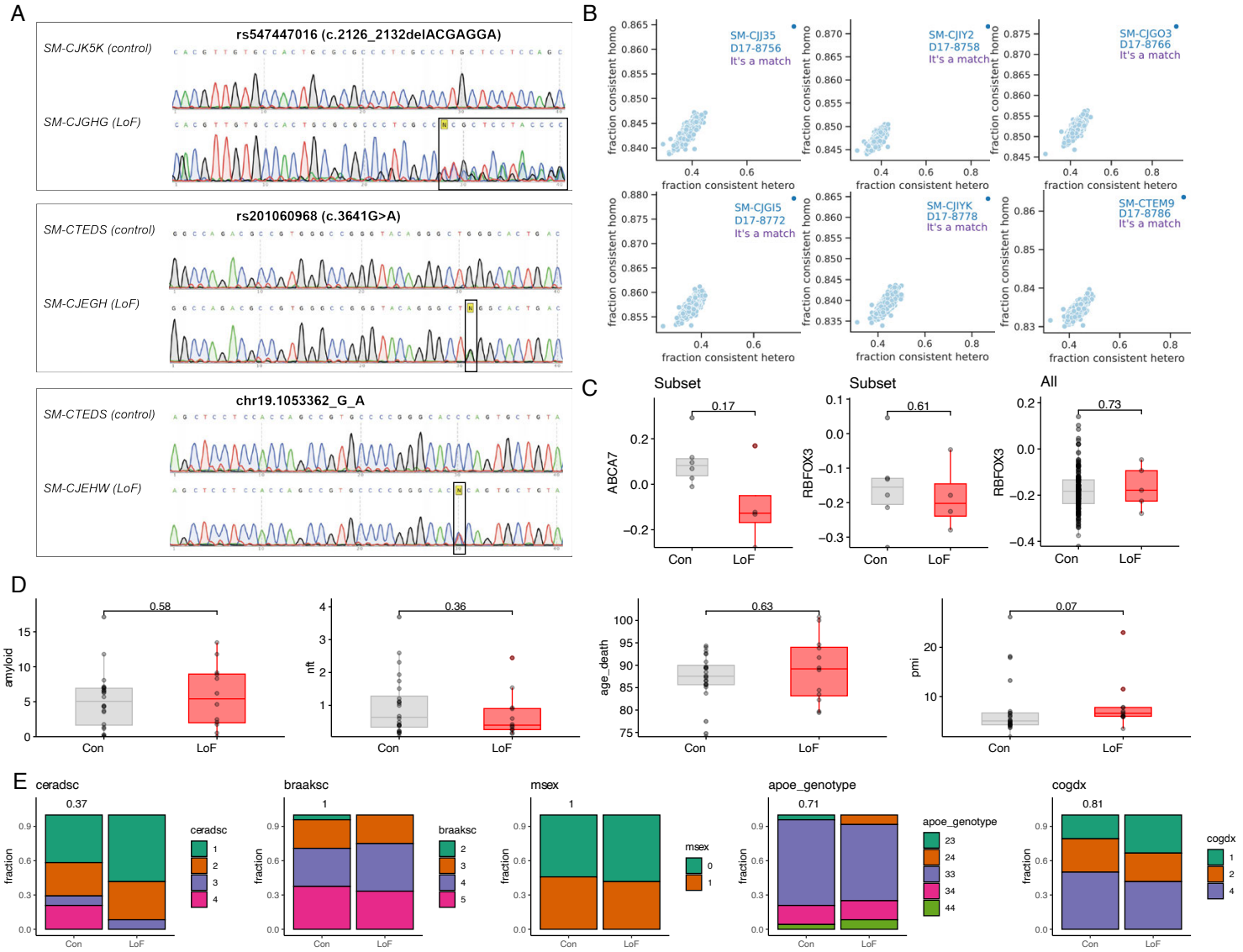

A
rs547447016 (c.2126_2132delACGAGGA)
SM-CJK5K (control)
SM-CJGHG (LoF)
rs201060968 (c.3641G>A)
SM-CTEDS (control)
SM-CJEGH (LoF)
chr19.1053362_G_A
SM-CTEDS (control)
SM-CJEHW (LoF)
B
SM-CJJ35
D17-8756
It's a match
SM-CJIY2
D17-8758
It's a match
SM-CJGO3
D17-8766
It's a match
SM-CJGI5
D17-8772
It's a match
SM-CJIYK
D17-8778
It's a match
SM-CTEM9
D17-8786
It's a match
fraction consistent homo
fraction consistent hetero
C
Subset
Subset
All
ABCA7
RBFOX3
RBFOX3
0.17
0.61
0.73
Con
LoF
D
amyloid
nft
age_death
pmi
0.58
0.36
0.63
0.07
Con
LoF
E
ceradsc
braaksc
msex
apoe_genotype
cogdx
0.37
1
1
0.71
0.81
fraction
Con
LoF

Fig. S1.

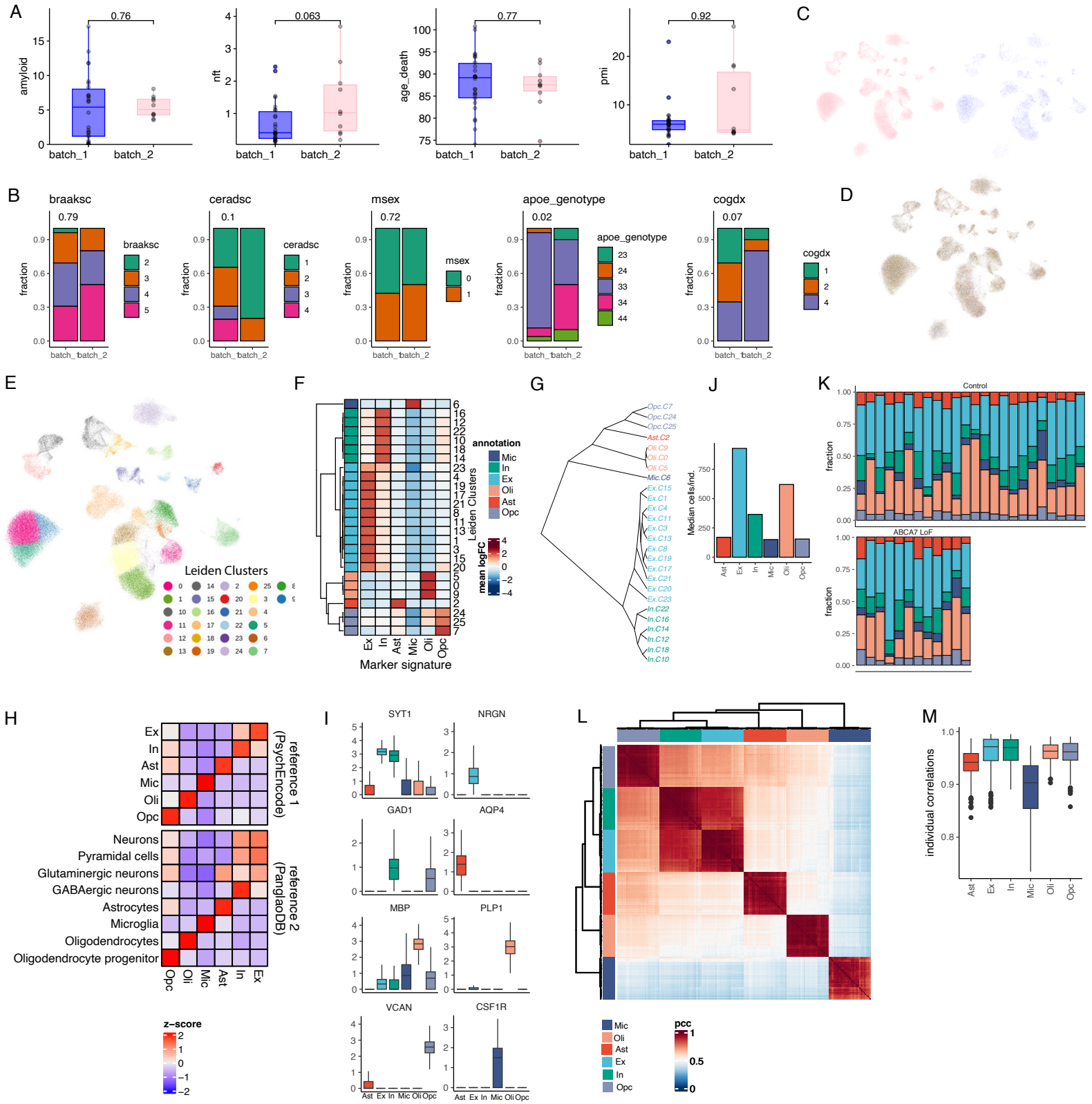

A
amyloid
nft
age_death
pmi
0.76
0.063
0.77
0.92
batch_1
batch_2
B
braaksc
ceradsc
msex
apoe_genotype
cogdx
0.79
0.1
0.72
0.02
0.07
fraction
C
D
E
Leiden Clusters
F
annotation
Mic
In
Ex
Oli
Ast
Opc
Leiden Clusters
mean logFC
Marker signature
G
J
Median cells/ind.
K
Control
ABCA7 LoF
H
reference 1 (PsychEncode)
reference 2 (PanglaoDB)
Neurons
Pyramidal cells
Glutaminergic neurons
GABAergic neurons
Astrocytes
Microglia
Oligodendrocytes
Oligodendrocyte progenitor
z-score
I
SYT1
NRGN
GAD1
AQP4
MBP
PLP1
VCAN
CSF1R
L
pcc
M
individual correlations

Fig. S2.

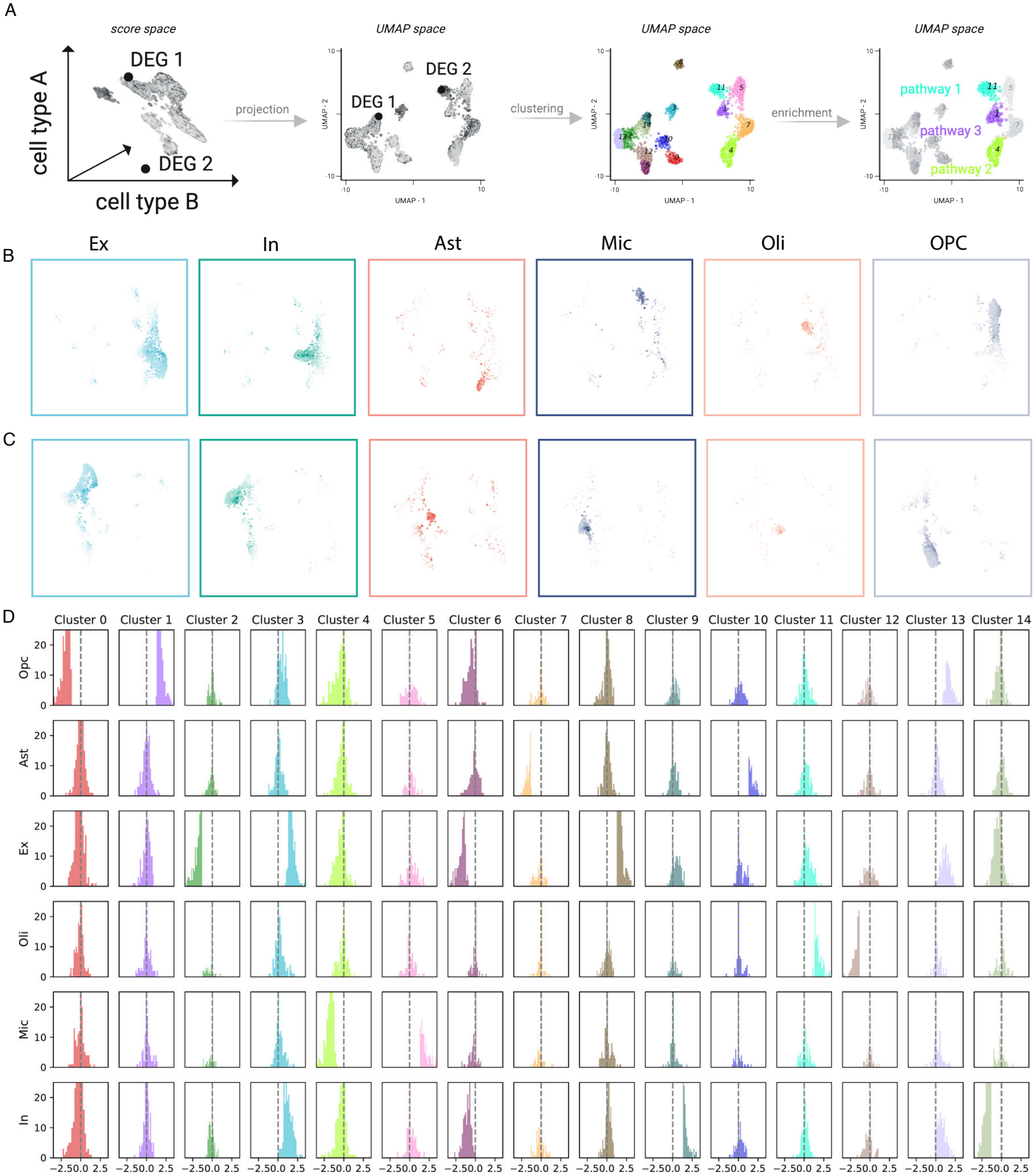

A
score space
UMAP space
UMAP space
UMAP space
cell type A
cell type B
DEG 1
DEG 2
projection
clustering
enrichment
pathway 1
pathway 2
pathway 3
UMAP - 1
UMAP - 2
B
C
D
Ex
In
Ast
Mic
Oli
OPC
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Cluster 11
Cluster 12
Cluster 13
Cluster 14
Opc
Ast
Ex
Oli
Mic
In

Fig. S3.

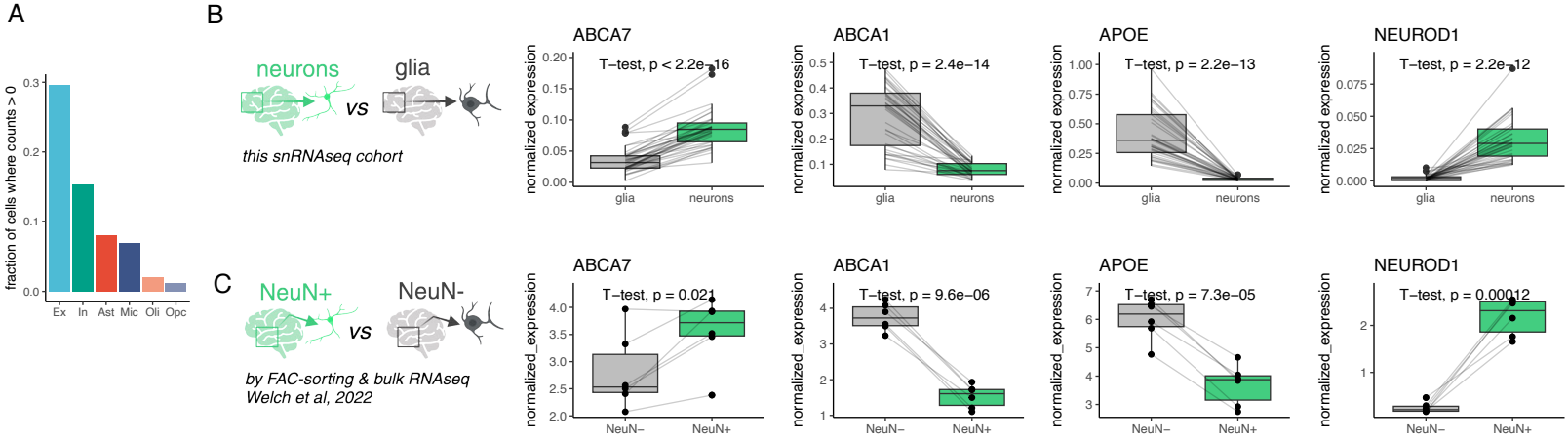

A
fraction of cells where counts > 0
Ex
In
Ast
Mic
Oli
Opc
B
neurons
glia
vs
this snRNAseq cohort
ABCA7
T−test, p < 2.2e−16
ABCA1
T−test, p = 2.4e−14
APOE
T−test, p = 2.2e−13
NEUROD1
T−test, p = 2.2e−12
normalized expression
glia
neurons
C
NeuN+
NeuN-
vs
by FAC-sorting & bulk RNAseq
Welch et al, 2022
ABCA7
T−test, p = 0.021
ABCA1
T−test, p = 9.6e−06
APOE
T−test, p = 7.3e−05
NEUROD1
T−test, p = 0.00012
normalized_expression
NeuN−
NeuN+

Fig. S4.

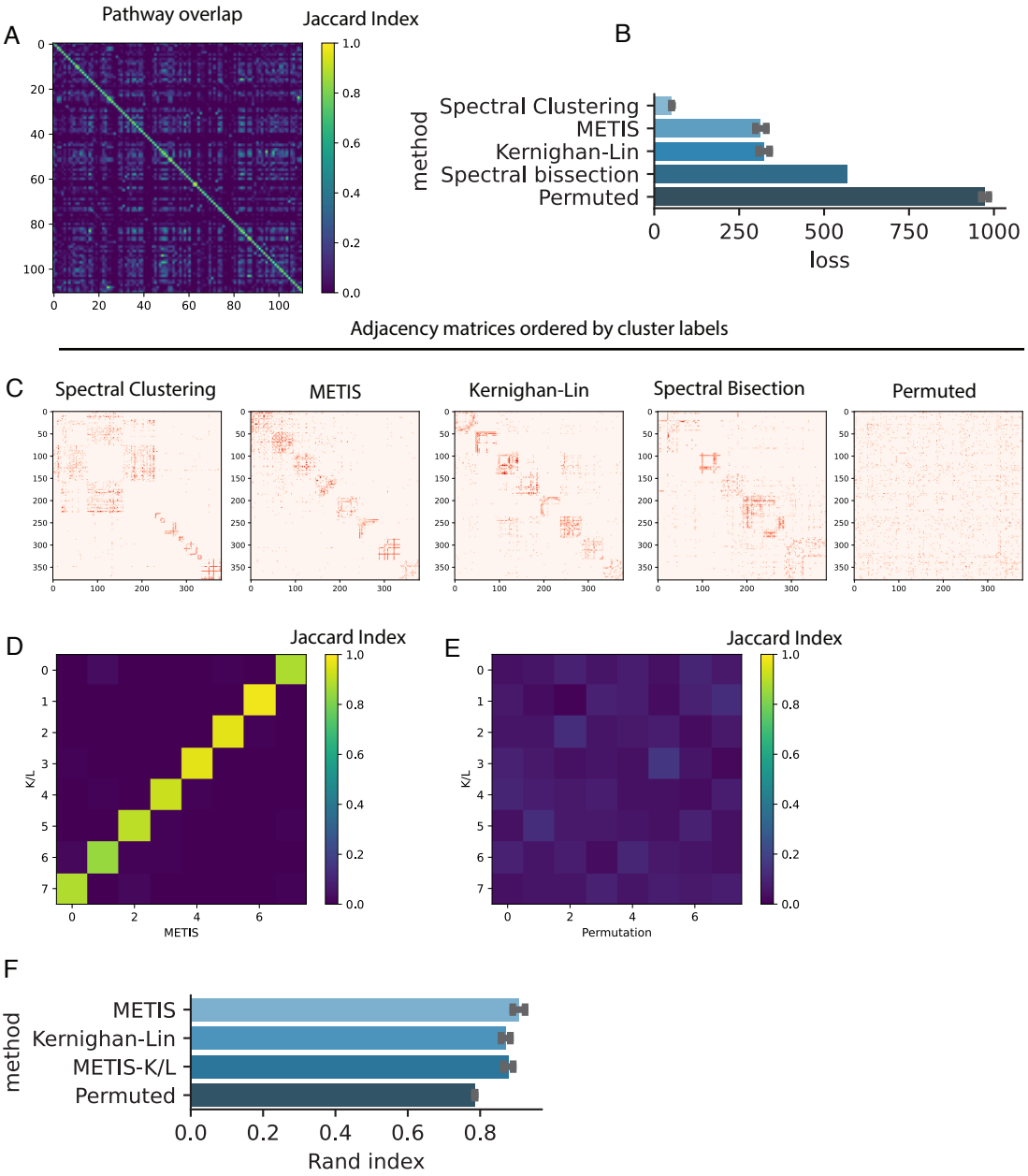

A
Pathway overlap
Jaccard Index
B
method
Spectral Clustering
METIS
Kernighan-Lin
Spectral bissection
Permuted
loss
Adjacency matrices ordered by cluster labels
C
Spectral Clustering
METIS
Kernighan-Lin
Spectral Bisection
Permuted
D
Jaccard Index
K/L
METIS
E
Jaccard Index
K/L
Permutation
F
method
METIS
Kernighan-Lin
METIS-K/L
Permuted
Rand index

Fig. S5.

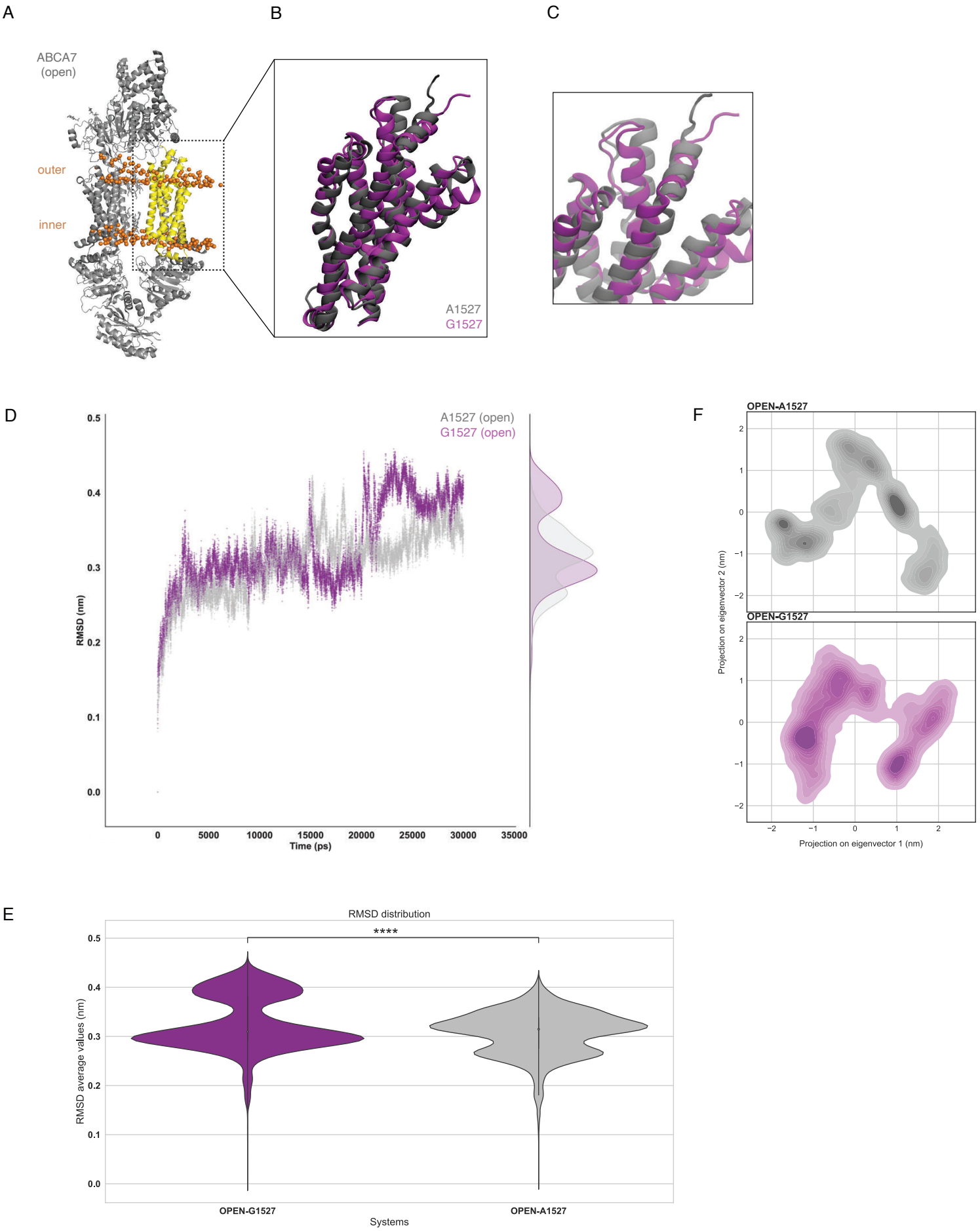

A
B
C
ABCA7
(open)
outer
inner
A1527
G1527
D
A1527 (open)
G1527 (open)
RMSD (nm)
Time (ps)
F
OPEN-A1527
OPEN-G1527
Projection on eigenvector 2 (nm)
Projection on eigenvector 1 (nm)
E
RMSD distribution
****
RMSD average values (nm)
OPEN-G1527
OPEN-A1527
Systems

Fig. S6.

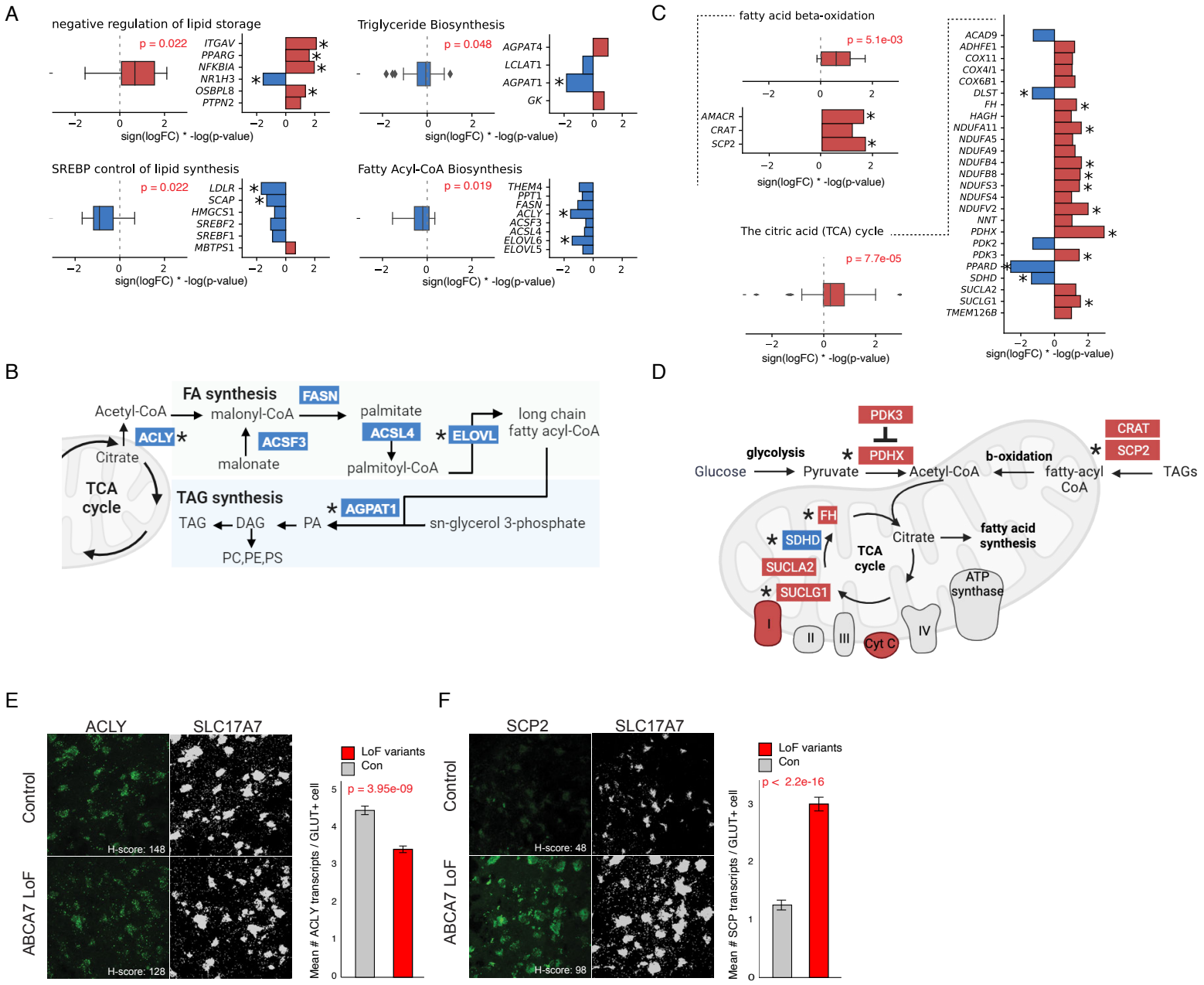

A
negative regulation of lipid storage
p = 0.022
ITGAV
PPARG
NFKBIA
NR1H3
OSBPL8
PTPN2
sign(logFC) * -log(p-value)
Triglyceride Biosynthesis
p = 0.048
AGPAT4
LCLAT1
AGPAT1
GK
SREBP control of lipid synthesis
p = 0.022
LDLR
SCAP
HMGCS1
SREBF2
SREBF1
MBTPS1
Fatty Acyl-CoA Biosynthesis
p = 0.019
THEM4
PPT1
FASN
ACLY
ACSF3
ACSL4
ELOVL6
ELOVL5
B
FA synthesis
FASN
Acetyl-CoA
malonyl-CoA
palmitate
ACLY
ACSF3
ACSL4
ELOVL
long chain fatty acyl-CoA
Citrate
malonate
palmitoyl-CoA
TCA cycle
TAG synthesis
AGPAT1
TAG
DAG
PA
sn-glycerol 3-phosphate
PC,PE,PS
C
fatty acid beta-oxidation
p = 5.1e-03
AMACR
CRAT
SCP2
The citric acid (TCA) cycle
p = 7.7e-05
ACAD9
ADHFE1
COX11
COX4I1
COX6B1
DLST
FH
HAGH
NDUFA11
NDUFA5
NDUFA9
NDUFB4
NDUFB8
NDUFS3
NDUFS4
NDUFV2
NNT
PDHX
PDK2
PDK3
PPARD
SDHD
SUCLA2
SUCLG1
TMEM126B
D
PDK3
glycolysis
PDHX
CRAT
SCP2
Glucose
Pyruvate
Acetyl-CoA
b-oxidation
fatty-acyl CoA
TAGs
FH
SDHD
SUCLA2
SUCLG1
Citrate
TCA cycle
fatty acid synthesis
ATP synthase
I
II
III
Cyt C
IV
E
ACLY
SLC17A7
Control
ABCA7 LoF
H-score: 148
H-score: 128
LoF variants
Con
p = 3.95e-09
Mean # ACLY transcripts / GLUT+ cell
F
SCP2
SLC17A7
Control
ABCA7 LoF
H-score: 48
H-score: 98
LoF variants
Con
p < 2.2e-16
Mean # SCP transcripts / GLUT+ cell

Fig. S7.

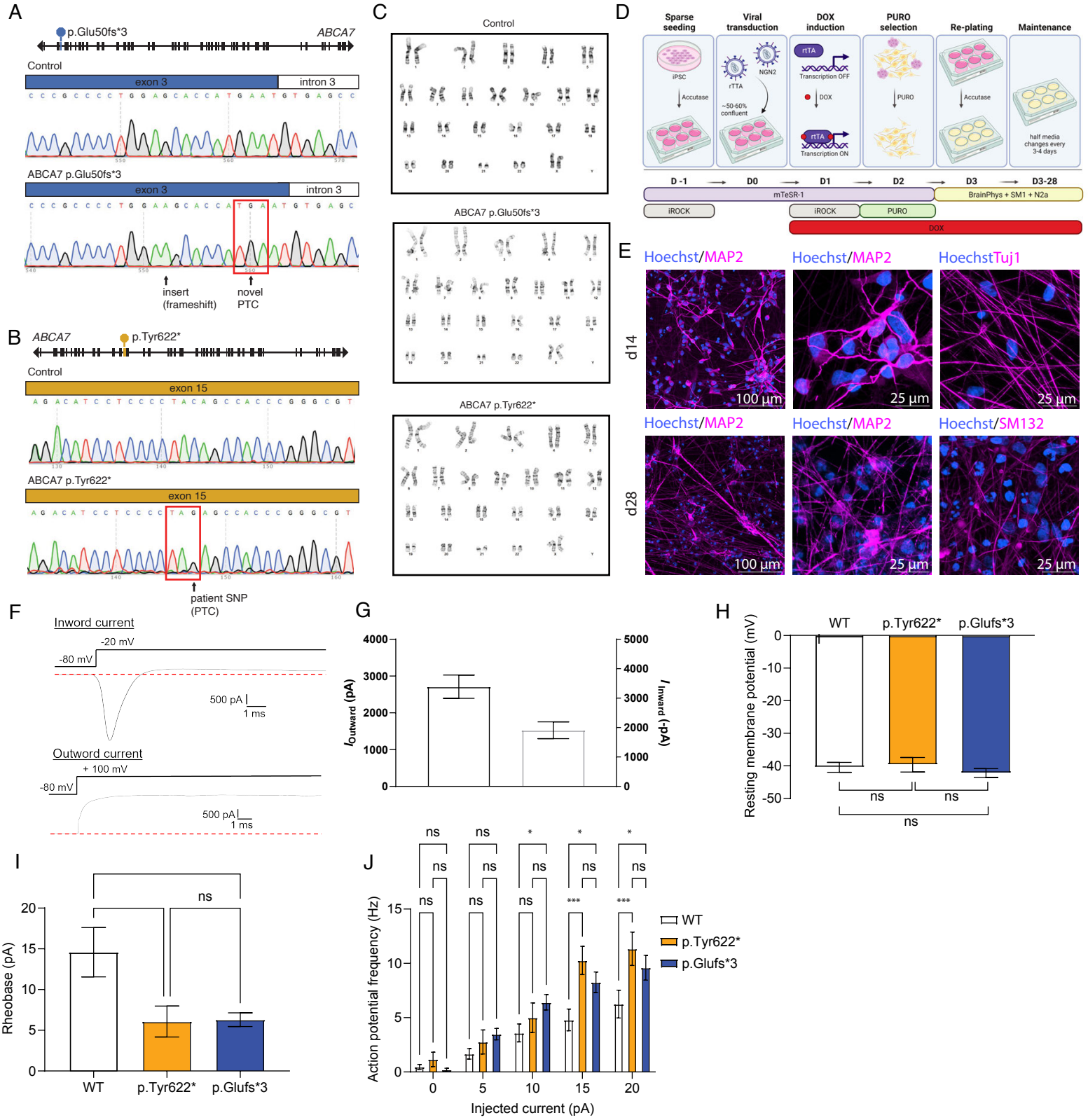

A
p.Glu50fs*3
ABCA7
Control
exon 3
intron 3
ABCA7 p.Glu50fs*3
insert (frameshift)
novel PTC
B
ABCA7
p.Tyr622*
Control
exon 15
ABCA7 p.Tyr622*
patient SNP (PTC)
C
Control
ABCA7 p.Glu50fs*3
ABCA7 p.Tyr622*
D
Sparse seeding
Viral transduction
DOX induction
PURO selection
Re-plating
Maintenance
iPSC
Accutase
rTTA
NGN2
~50-60% confluent
rtTA
Transcription OFF
DOX
Transcription ON
PURO
half media changes every 3-4 days
D -1
D0
D1
D2
D3
D3-28
mTeSR-1
BrainPhys + SM1 + N2a
iROCK
PURO
DOX
E
Hoechst/MAP2
Hoechst/MAP2
HoechstTuj1
d14
100 µm
25 µm
25 µm
Hoechst/MAP2
Hoechst/MAP2
Hoechst/SM132
d28
100 µm
25 µm
25 µm
F
Inward current
-20 mV
-80 mV
500 pA
1 ms
Outward current
+ 100 mV
-80 mV
500 pA
1 ms
G
$I_{outward}$ (pA)
I_{inward} (-pA)
H
WT
p.Tyr622*
p.Glufs*3
Resting membrane potential (mV)
ns
ns
ns
I
Rheobase (pA)
ns
WT
p.Tyr622*
p.Glufs*3
J
Action potential frequency (Hz)
Injected current (pA)
WT
p.Tyr622*
p.Glufs*3

Fig. S8.

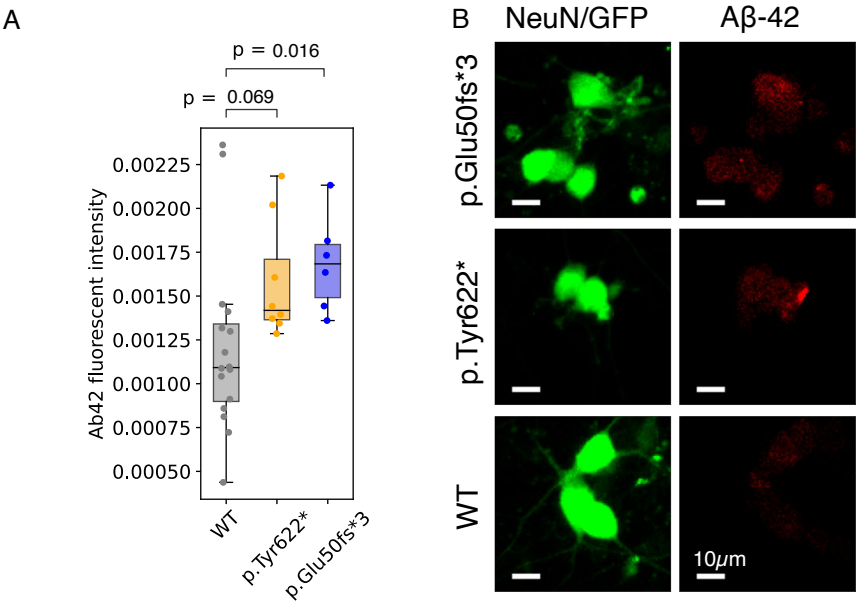

A
p = 0.016
p = 0.069
Ab42 fluorescent intensity
0.00225
0.00200
0.00175
0.00150
0.00125
0.00100
0.00075
0.00050
WT
p.Tyr622*
p.Glu50fs*3
B
NeuN/GFP
Aβ-42
p.Glu50fs*3
p.Tyr622*
WT
10μm

Fig. S9

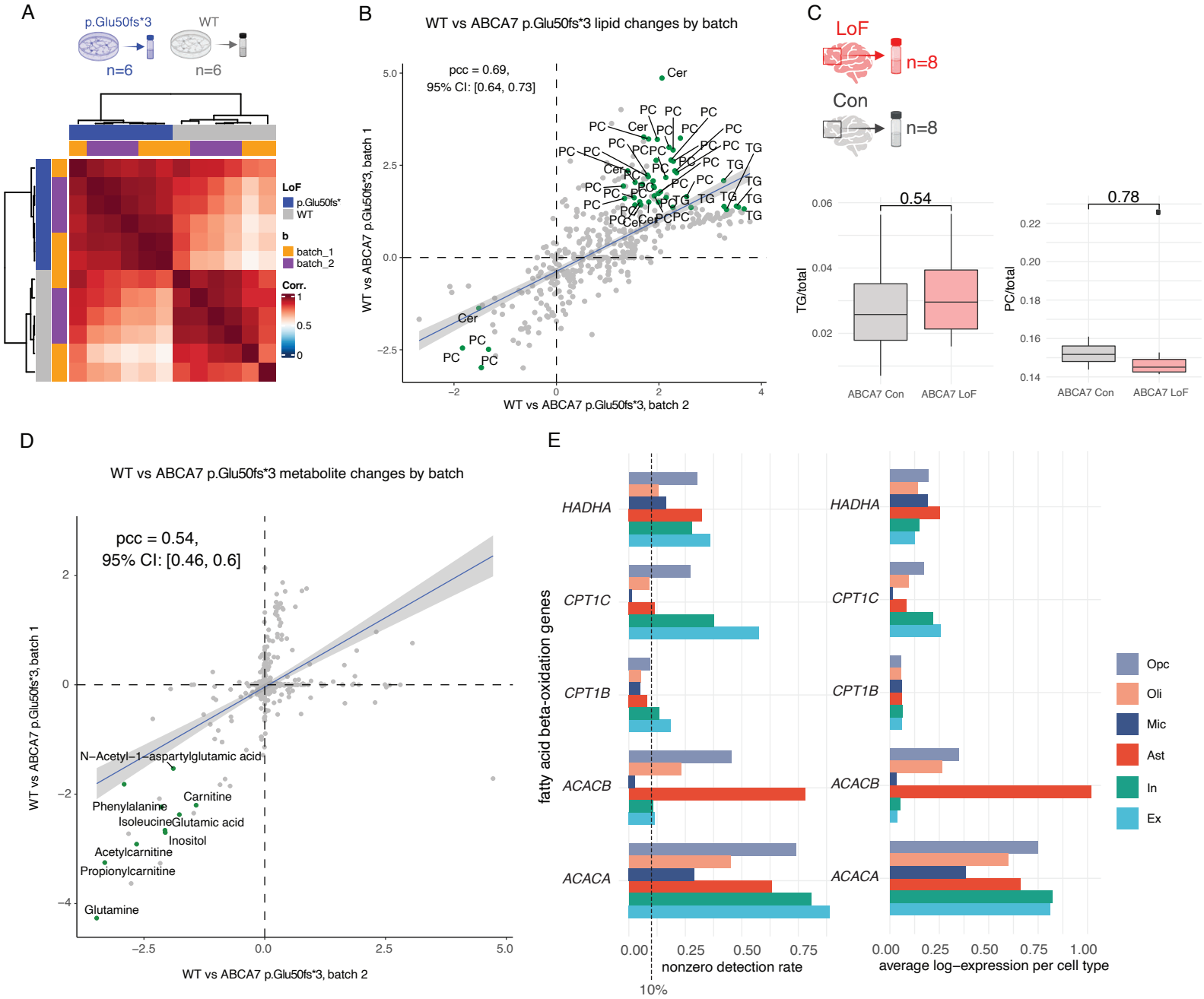

A
p.Glu50fs*3
WT
n=6
n=6
LoF
p.Glu50fs*
WT
b
batch_1
batch_2
Corr.
1
0.5
0
B
WT vs ABCA7 p.Glu50fs*3 lipid changes by batch
pcc = 0.69,
95% CI: [0.64, 0.73]
WT vs ABCA7 p.Glu50fs*3, batch 1
WT vs ABCA7 p.Glu50fs*3, batch 2
Cer
PC
TG
C
LoF
n=8
Con
n=8
0.54
0.78
TG/total
PC/total
ABCA7 Con
ABCA7 LoF
D
WT vs ABCA7 p.Glu50fs*3 metabolite changes by batch
pcc = 0.54,
95% CI: [0.46, 0.6]
WT vs ABCA7 p.Glu50fs*3, batch 1
WT vs ABCA7 p.Glu50fs*3, batch 2
N−Acetyl−1−aspartylglutamic acid
Carnitine
Phenylalanine
Isoleucine
Glutamic acid
Inositol
Acetylcarnitine
Propionylcarnitine
Glutamine
E
fatty acid beta-oxidation genes
HADHA
CPT1C
CPT1B
ACACB
ACACA
nonzero detection rate
10%
average log−expression per cell type
Opc
Oli
Mic
Ast
In
Ex

Fig. S10.

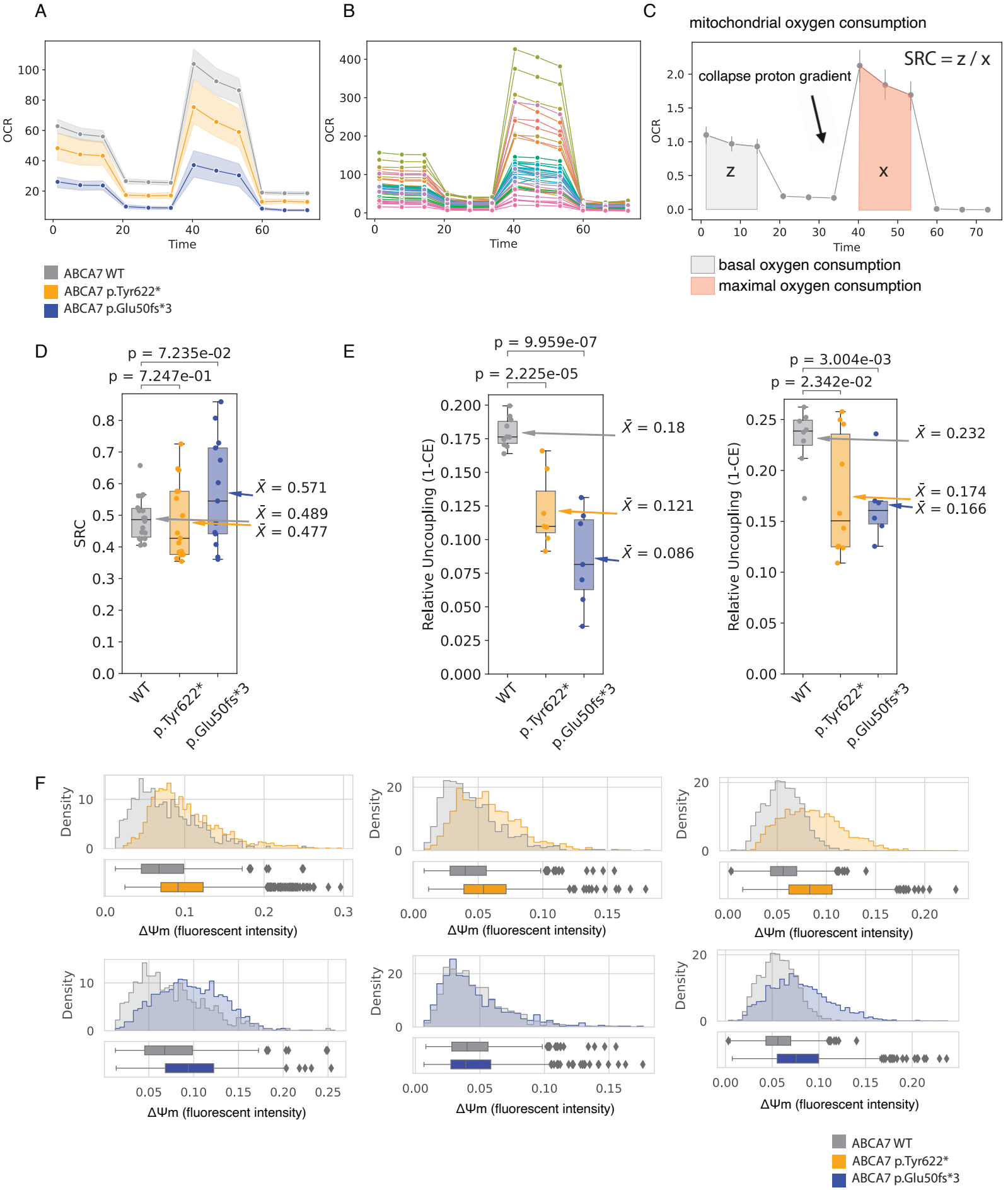

A
B
C
mitochondrial oxygen consumption
SRC = z / x
collapse proton gradient
OCR
Time
z
x
basal oxygen consumption
maximal oxygen consumption
ABCA7 WT
ABCA7 p.Tyr622*
ABCA7 p.Glu50fs*3
D
E
p = 7.235e-02
p = 7.247e-01
SRC
$\bar{X}$ = 0.571
$\bar{X}$ = 0.489
$\bar{X}$ = 0.477
p = 9.959e-07
p = 2.225e-05
Relative Uncoupling (1-CE)
$\bar{X}$ = 0.18
$\bar{X}$ = 0.121
$\bar{X}$ = 0.086
p = 3.004e-03
p = 2.342e-02
$\bar{X}$ = 0.232
$\bar{X}$ = 0.174
$\bar{X}$ = 0.166
WT
p.Tyr622*
p.Glu50fs*3
F
Density
ΔΨm (fluorescent intensity)

Fig. S11.

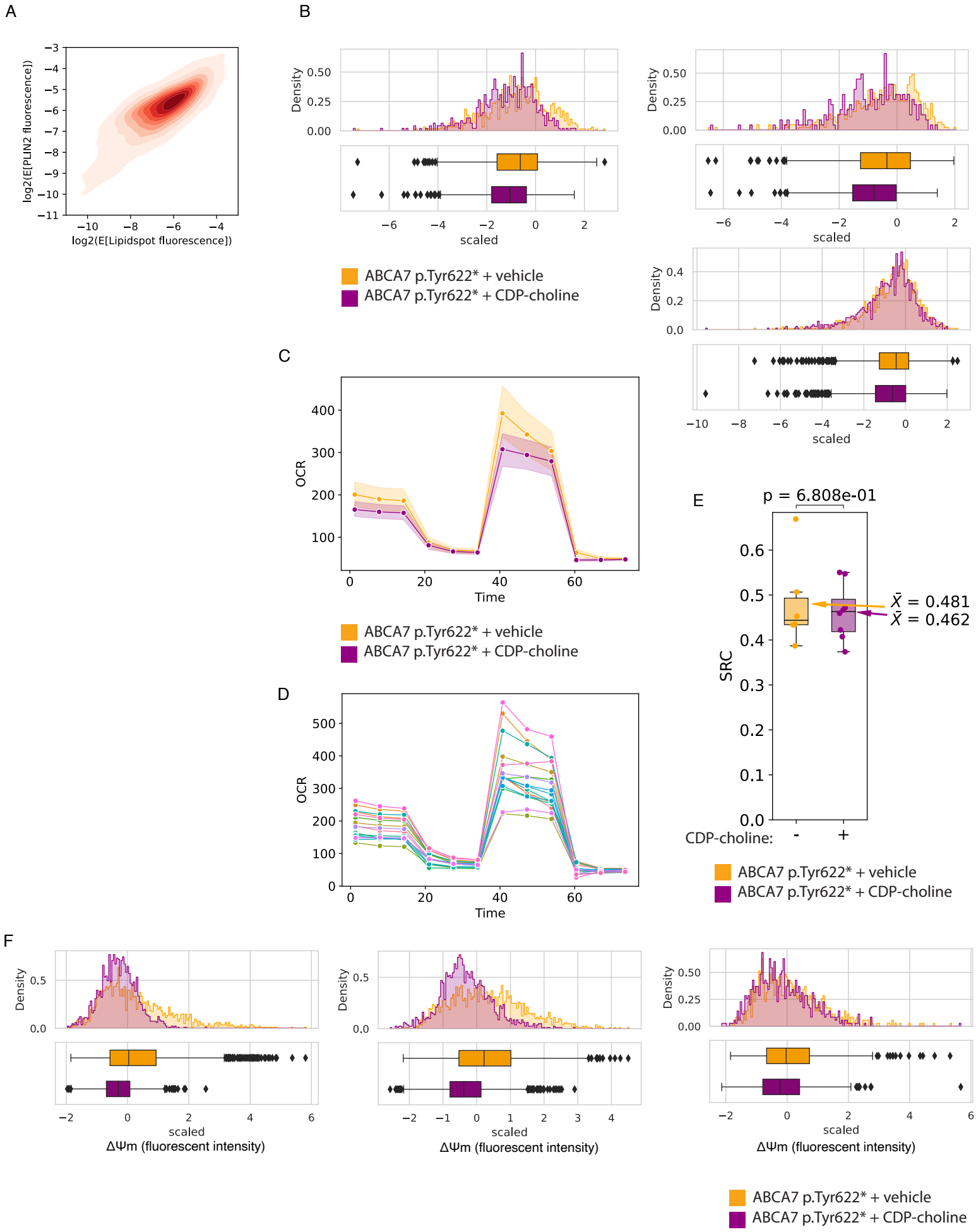

A
log2(E[PLIN2 fluorescence])
log2(E[Lipidspot fluorescence])
B
Density
scaled
ABCA7 p.Tyr622* + vehicle
ABCA7 p.Tyr622* + CDP-choline
C
OCR
Time
ABCA7 p.Tyr622* + vehicle
ABCA7 p.Tyr622* + CDP-choline
D
OCR
Time
E
p = 6.808e-01
SRC
$\bar{X}$ = 0.481
$\bar{X}$ = 0.462
CDP-choline: - +
ABCA7 p.Tyr622* + vehicle
ABCA7 p.Tyr622* + CDP-choline
F
Density
scaled
ΔΨm (fluorescent intensity)
ABCA7 p.Tyr622* + vehicle
ABCA7 p.Tyr622* + CDP-choline

Fig. S12.