Dedifferentiation orchestrated through remodeling of the chromatin landscape defines *PSEN1* mutation-induced Alzheimer's Disease

Andrew B. Caldwell[1], Qing Liu[2], Gary P. Schroth[3], Rudolph E. Tanzi[4], Douglas R. Galasko[2], Shauna H. Yuan[2], Steven L. Wagner[2,5], & Shankar Subramaniam[1,6,7,8*]

[1]Department of Bioengineering, University of California, San Diego, La Jolla, California, USA.
[2]Department of Neurosciences, University of California, San Diego, La Jolla, California, USA.
[3]Illumina, Inc., San Diego, California, USA.
[4]Department of Neurology, Massachusetts General Hospital, Charlestown, Massachusetts, USA.
[5]VA San Diego Healthcare System, La Jolla, California, USA.
[6]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, California, USA.
[7]Department of Nanoengineering, University of California, San Diego, La Jolla, California, USA.
[8]Department of Computer Science and Engineering, University of California, San Diego, La Jolla, California, USA.

Abstract

Early-Onset Familial Alzheimer's Disease (EOFAD) is a dominantly inherited neurodegenerative disorder elicited by mutations in the *PSEN1*, *PSEN2*, and *APP* genes[1]. Hallmark pathological changes and symptoms observed, namely the accumulation of misfolded Amyloid-β (Aβ) in plaques and Tau aggregates in neurofibrillary tangles associated with memory loss and cognitive decline, are understood to be temporally accelerated manifestations of the more common sporadic Late-Onset Alzheimer's Disease. The complete penetrance of EOFAD-causing mutations has allowed for experimental models which have proven integral to the overall understanding of AD[2]. However, the failure of pathology-targeting therapeutic development suggests that the formation of plaques and tangles may be symptomatic and not describe the etiology of the disease[3,4]. Here, we use an integrative, multi-omics approach and systems-level analysis in hiPSC-derived neurons to generate a mechanistic disease model for EOFAD. Using patient-specific cells from donors harboring mutations in *PSEN1* differentiated into neurons, we characterize the disease-related gene expression and chromatin accessibility changes by RNA-Seq, ATAC-Seq, and histone methylation ChIP-Seq. We show that the defining disease-causing mechanism of EOFAD is dedifferentiation, primarily through the REST-mediated repression of neuronal lineage specification gene programs and the activation of non-specific germ layer precursor gene programs concomitant with modifications in chromatin accessibility. These gene signature profiles and changes in chromatin topology illustrate that EOFAD neurons traverse the chromatin landscape from an ectodermal origin to a mixed germ lineage state. Further, a reanalysis of existing transcriptomic data from *PSEN1* patient brain samples demonstrates that the mechanisms identified in our experimental system recapitulate EOFAD in the human brain. Our results comprise a disease model which describes the mechanisms culminating in dedifferentiation that precede amyloid and tau pathology formation and engender neurodegeneration.

Results

Non-Demented Control (NDC), *PSEN1*^{M146L}, and *PSEN1*^{A246E} hiPSCs were generated and differentiated into neurons as previously described[5,6]. Subsequent RNA-Seq of FACS-purified neurons identified 5699 downregulated and 4321 upregulated genes in *PSEN1*^{M146L} neurons relative to NDC following differential expression analysis (Figure 1a). To ascertain the transcriptional regulators controlling these gene expression changes, we performed ISMARA[7] analysis to characterize the transcription factors (TF) and miRNAs (miR) for which the corresponding motif activity most significantly changed between the *PSEN1*^{M146L} and NDC

conditions (Figure 1b). The top transcriptional regulators coalesced into five general categories of gene programs, three upregulated (Cell Cycle, Pluripotency, and Dedifferentiation) and two downregulated (Neuron Lineage Specification and Function, Proliferation/Differentiation) (Figure 1c).

We confirmed this sequencing analysis in the *PSEN1*^{A246E} mutation, which exhibited similar dysregulation of key TFs and miRNAs related to these core disease-associated gene programs (Supplementary Figure 1a-b). As two key miRNAs related to neuron lineage specification, miR-124 and miR-9[8], were identified by ISMARA in the *PSEN1*^{M146L} and *PSEN1*^{A246E} neurons as having decreased activity and are known repressors of the most significantly activated TF identified, REST[9], we used the miRTarBase database to perform pre-ranked GSEA[10] and Enrichr[11] enrichment in the *PSEN1*^{M146L} (Figure 1d, Supplementary Figure 2a) and *PSEN1*^{A246E} (Supplementary Figure 1c-d) conditions; these results illustrate that the targets of these miRNAs are significantly upregulated in both conditions. Next, to determine whether this is due to loss of miR-124 and miR-9, we performed miRNA-qPCR. This revealed that miR-124 and miR-9 levels are significantly decreased in *PSEN1*^{M146L} and *PSEN1*^{A246E} neurons relative to NDC (Figure 1e).

To further expand the characterization of altered transcriptional regulation and gene programs, we carried out Enrichr TF-target enrichment using the ENCODE/ChEA consensus database, Panther[12] gene set enrichment analysis using the Reactome Pathway and Gene Ontology (GO) databases, and pre-ranked GSEA for Hallmark gene sets for the differentially regulated genes in the *PSEN1*^{M146L} (Figure 1f-g, Supplemental Figure 2b-c) and *PSEN1*^{A246E} (Supplemental Figure 1d-h) conditions; this revealed consistent upregulated enrichment of TFs and gene sets related to Cell Cycle[13], Pluripotency[14,15], and non-ectoderm lineage dedifferentiation[16,17] (EMT, ECM organization, and mesendoderm lineage specification captured by cardiac system-related development gene sets) and consistent downregulation of TFs related to neuron lineage specification (REST) and neuron function (REST, NRF1, CREB1)[18–20].

In order to validate these results, we performed qPCR in *PSEN1*^{M146L}, *PSEN1*^{A246E}, and *PSEN1*^{H136R} hiPSC-derived neurons to assess key cell cycle related and REST-repressed (*GAD1*) genes, recapitulating expression patterns observed in the *PSEN1*^{M146L} and *PSEN1*^{A246E} RNA-Seq data. Furthermore, as terminally differentiated neurons are generally thought to exist in the quiescent G_0 phase, yet *PSEN1*^{M146L} and *PSEN1*^{A246E} neurons exhibit significant gene set enrichment for cell cycle-related processes, we sought to determine the neuron cell cycle state. These analyses demonstrated that *PSEN1*^{M146L} and *PSEN1*^{A246E} neuron populations have a significantly lower percentage of cells in the G_0 and S phases and a higher percentage in the G2/M phases, suggesting they have re-entered the cell cycle (Figure 1i).

Next, we evaluated whether these observed transcriptional changes were due to modulation of the chromatin state. To this end, we performed ATAC-Seq on NDC and *PSEN1*^{M146L} hiPSC-derived neurons, identifying 14867 regions of chromatin with differential accessibility between the two conditions, 5451 of which occurred within differentially expressed genes (Figure 2a-b). Enrichr TF and Panther Reactome Pathway enrichment of upregulated and downregulated genes with differential chromatin accessibility identified TFs related to cell cycle (E2F4, FOXM1), pluripotency (SOX2, MYC), dedifferentiation (SALL4, ZNF217)[21,22], and chromatin remodeling (SUZ12, ZBTB7A)[23,24] among the upregulated genes and TFs related to neuron lineage specification[24,25] (REST, PRC2 components) and neuron mitochondrial function (NRF1, CREB1) among the downregulated genes (Figure 2c, Supplemental Figure 2a-c). Expanded enrichment revealed that NRF1 and CREB1 target genes were enriched among the downregulated genes with differential chromatin accessibility occurring within the promoter, whereas REST and PRC2 target genes were enriched among the downregulated genes with differential chromatin accessibility occurring outside the promoter. Interestingly, HINT TF footprinting analysis of DEGs with differential promoter accessibility identified NRF1 as the top TF with loss of binding motif accessibility in the

PSEN1^{M146L} condition (Figure 2d). Panther enrichment using the GO biological process and cellular component gene sets revealed that transcriptional upregulation of non-ectoderm lineage gene programs and downregulation of neuronal function programs is driven by changes in chromatin accessibility (Supplemental Figure 2g-h). Further, this gain or loss of promoter accessibility can be observed across all differentially accessible genes upregulated and downregulated at the gene expression level (Figure 2f).

As chromatin accessibility can be modulated through changes in histone methylation[26], and TF enrichment revealed REST and the histone-modulating PRC2 complex target genes as being significantly downregulated, we performed ChIP-Seq on NDC and *PSEN1*^{M146L} hiPSC-derived neurons for the activating mark H3K4Me3 and the repressive mark H3K27Me3. This revealed 7309 regions of differential H3K4 trimethylation and 21230 regions of differential H3K27 trimethylation (Supplemental Figure 4a); of these differentially methylated regions, 1583 upregulated genes intersected with increased H3K4Me3 and 667 with decreased H3K27Me3, whereas 733 downregulated genes intersected with decreased H3K4Me3 and 1147 intersected with increased H3K27Me3 (Figure 3a-b). Interestingly, Enrichr TF enrichment of these differentially expressed genes with differential methylation revealed that the upregulated genes were enriched for control by pluripotency-related TFs (SOX2, NANOG, KLF4), the dedifferentiation TF SALL4, and the PRC2-complex factor SUZ12, while the downregulated genes are enriched primarily for REST and PRC2-complex factors SUZ12 and EZH2 (Figure 3c-d). Panther enrichment of GO biological process gene sets revealed similar upregulation of non-ectoderm lineage development and downregulation of synaptic signaling processes in genes with differential H3K4Me3 or H3K27Me3, whereas enrichment of Reactome Pathway genes sets revealed upregulation of cell cycle programs unique to the differential H3K4Me3 genes (Supplemental Figure 4b-d).

Next, we integrated the ATAC-Seq data with the histone methylation ChIP-Seq data, leading to the discovery that genes with increased expression and chromatin accessibility showed increased H3K4Me3 density around the regions of maximum accessibility increase**,** while genes with decreased expression and accessibility showed increased H3K27Me3 around the regions of maximum accessibility decrease (Figure 3e). Further, the intersection of differentially expressed genes with differential methylation (Figure 3f-g) concomitant with differential chromatin accessibility (Figure 3h-i) revealed a substantial subset of genes. Read profiles of previously identified AD-associated genes[27,28] (*LRP10*, *VGF*) as well as non-ectoderm lineage (*TEAD2*, *TCF7L2*) or neuroectoderm lineage[29] (*BDNF*, *MYT1L*) specifying genes show chromatin accessibility and histone methylation changes which correspond directionally to their increased (Figure 3j) or decreased (Figure 3k) gene expression. As expected, the locus around the two key neuron lineage miRNAs downregulated in both *PSEN1*^{M146L} and *PSEN1*^{A246E} conditions, miR-9 and miR-124, show decreased chromatin accessibility and H3K4Me3 as well as increased H3K27Me3 (Figure 3l). A heat map corresponding to directionally correlated RNA expression, promoter accessibility, and H3K trimethylation for an expanded list of genes belonging to the core gene programs identified here or previously identified as AD marker genes illustrates key gene expression changes occurring due modifications of the chromatin landscape (Figure 3m).

In order to further delineate the mechanistic gene changes occurring due to modifications at the chromatin level, we separated the genes into categories of those with a change in chromatin accessibility concomitant with differential H3K4 or H3K27 trimethylation or those with a change in promoter chromatin accessibility alone; Enrichr TF enrichment and ToppGene[30] GO Biological Process gene overrepresentation analysis revealed that genes in the former group are primarily involved in pluripotency, dedifferentiation, and neuron lineage specification gene programs (Supplemental Figure 5a-b) while those in the latter group are involved in cell cycle, chromatin landscape modification, mitochondrial, and vesicle transport gene programs (Supplemental

Figure 4c-e). To gain insight into what types of histone modifications may be causing differential accessibility and expression in genes with no change in H3K4Me3 or H3K27Me3 status, we carried out Enrichr enrichment using the ENCODE and Roadmap Epigenomics histone modification gene sets. Interestingly, this approach revealed that genes with increased accessibility and expression were enriched for signatures of H3 acetylation at K56, K27, and K9 residues associated with less-differentiated, precursor cell types (Supplemental Figure 5f). Moreover, genes with decreased accessibility and expression were enriched for signatures of H3 acetylation and H3K79Me2 associated with precursor and non-ectoderm lineage cells. These results suggest the differentially accessible and expressed genes lacking a change in H3K4Me3 or H3K27Me3 state may primarily be modulated by histone acetylation, and the expression pattern due to this modification indicates a precursor-like, non-ectoderm cell lineage profile. Overall, this multi-omics, integrative analysis illustrates that REST and PRC2 neuronal target genes are transcriptionally repressed due to loss of chromatin accessibility driven by gain of H3K27Me3 and loss of H3K4Me3, whereas NRF1 and CREB1 neuronal target genes are repressed due to loss of chromatin accessibility in the promoter without change in the H3K4Me3 or H3K27Me3 methylation state.

Patient-specific iPSC-derived neurons are a powerful model system to investigate EOFAD-causing mutations in a mechanistic, systems-level fashion; to confirm whether the mechanisms identified in *PSEN1*^{M146L} neurons faithfully capture the disease-related processes occurring in AD patients, we reanalyzed a previous transcriptomic study[31] (GSE39420) of posterior cingulate cortex brain tissue from NDC and *PSEN1*^{M139T} EOFAD donors. Expression analysis revealed a substantial number of differentially expressed genes between the NDC and *PSEN1*^{M139T} brain overlapping with *PSEN1*^{M146L} and *PSEN1*^{A246E} neurons, particularly apparent for downregulated genes (Figure 4a-b). ISMARA motif analysis and Enrichr TF target enrichment showed similar TFs and miRNAs related to neuronal differentiation (REST, PRC2 complex, miR-124), non-ectoderm lineage specification and dedifferentiation (PRRX2, SMAD4, miR-124), pluripotency (SOX2, GATA2, LEF1), and cell cycle (E2F1/7, miR-9) as observed in *PSEN1*^{M146L} and *PSEN1*^{A246E} neurons, with REST identified as the top enriched/activated TF in *PSEN1*^{M139T} brain by both approaches (Supplemental Figure 6a-c).

Further comparison of ISMARA and pre-ranked GSEA results for TF motifs and TF targets, respectively, for these five core disease-related gene programs identified here revealed consistent TF modulation between *PSEN1* neurons and *PSEN1*^{M139T} brain (Figure 4c). As the transcriptional regulator MYT1L was identified at the intersection of our multi-omics approach as a key downregulated gene and recently was shown to promote neuronal differentiation and repress non-ectoderm lineage[32], we performed pre-ranked GSEA using a previously-identified list of MYT1L-regulated genes. Markedly, the normalized enrichment scores for MYT1L were similar across all disease conditions, suggesting that loss of MYT1L, a REST-repressed and miR-124 promoted gene, contributes to loss of the neuronal state (Supplemental Figure 6c). This consistent enrichment similarity between *PSEN1* neurons and *PSEN1*^{M139T} brain is also seen at the gene set level by Panther enrichment and pre-ranked GSEA approaches, particularly for the downregulated synaptic signaling and plasticity gene sets (Figure 4d, Supplemental Figure 6d-e). A selection of key differentially expressed genes common between all three conditions previously identified as AD marker genes[28] (*LRP10*, *HOMER1*, *VGF*) or those related to non-ectoderm dedifferentiation[33] (*YAP1*, *TEAD2*, *ZEB2*), neuronal gene splicing and activation[34,35] (*CELF3*, *ELAVL2*, *FOXP2*), and synaptic signaling[36] (*KCNB1*, *SNAP25, SYT1*) demonstrates that the core gene programs identified by our multi-omics approach in *PSEN1*^{M146L} neurons are analogously modulated by alternative mutations as well as in EOFAD patients (Figure 4e).

Our expanded enrichment analysis of *PSEN1* neuronal transcriptomics, combined with chromatin dynamics analysis in *PSEN1*^{M146L} neurons, has revealed five modulated gene programs integral

to the understanding of the overall disease mechanism: pluripotency, dedifferentiation, cell cycle re-entry, lineage miRNA, and neuronal specification encompassing lineage definition and synaptic function. Strikingly, the key aberrant gene programs in *PSEN1*^{M146L} and *PSEN1*^{A246E} neurons driving neurodegeneration are analogously dysregulated in *PSEN1*^{M139T} human patient brains. The comparative enrichment scores for the top transcriptional regulators belonging to these programs across each EOFAD condition gives insight into the documented severity of each mutation, as *PSEN1*^{M146L} causes a particularly early onset and accelerated progression of the disease.

The integration of ATAC-Seq and histone methylation ChIP-Seq with TF target and binding site enrichment analysis of RNA-Seq data in *PSEN1*^{M146L} neurons revealed that transcriptional modulation of the core gene programs have distinct mechanisms underlying the chromatin accessibility changes. Whereas upregulation of pluripotency genes and REST-mediated repression of neuronal genes occur due to promoter-proximal changes in H3K4Me3 and H3K27Me3 states leading to changes in accessibility, the upregulation of cell cycle and downregulation of NRF1 and CREB1-mediate neuronal maturation and mitochondrial function genes primarily occur due chromatin accessibility changes within promoter regions driven by alternative histone modifications combined with the loss of the pioneering TF capabilities of NRF1 and CREB1[37].

However, these core gene programs and the transcriptional regulators orchestrating their modulation do not exist in isolation from one another, and in certain key aspects are intertwined in how they reshape the chromatin landscape (Figure 4f). For example, commitment to a neuronal lineage involves the coordinated repression of REST, repression of the REST-recruited, histone-modifying complex PRC2, and switching of the chromatin-modulating SWI-SNF complex components from a npBAF (neuronal precursor) to a nBAF (neuronal) state, which predominantly occurs through increased expression of miR-124 and miR-9[8,38]. Furthermore, miR-124 and MYT1L[32] have been shown to restrict commitment to an neuroectoderm lineage through their repression of non-ectoderm lineage genes, while miR-9 plays an integral role in the cell cycle exit required for neuronal differentiation through its repression of key cell cycle regulators upon activation[39]. These multifaceted functions of REST, miR-124, miR-9, and MYT1L highlight their overarching identity as key neuron lineage determination factors; therefore, the combined reactivation of REST and downregulation of miR-124, miR-9, and MYT1L cause EOFAD neurons to dedifferentiate along an alternative axis through the repression of neuronal differentiation concomitant with activation of trans-dedifferentiation programs (Figure 4g). Consequently, the resulting EOFAD neurons exist in a precursor-like state with transcriptional signatures of ectoderm and non-ectoderm lineage alike.

This reversal to a mixed-lineage, germ layer state can be further inferred from the inflammatory expression signature observed in *PSEN1*^{M139T} brains as well as *PSEN1*^{M146L} and *PSEN1*^{A246E} neurons. The upregulation of key microglia lineage-specifying TFs and end state marker genes, such as *IRF8*, *STAT1*, and *TMEM119*[40], demonstrate the acquisition of non-ectoderm features contributing to the pronounced inflammatory signatures enriched in both *PSEN1* mutations (Supplemental Figure 6e-f). Additionally, the transcriptional regulators, which orchestrate dedifferentiation, particularly through EMT programs, enriched in both *PSEN1* neurons and brain samples, have alternative roles in shaping the overall inflammatory phenotype. Furthermore, transcriptional changes of factors shown to be indicative of the accumulation of tau and amyloid pathologies were observed in all three disease models. Key regulators involved in tau dephosphorylation such as *PP1A*, *PP2A*, and *GSK3B*[41], are transcriptionally downregulated, whereas *NOTCH1*, a gamma secretase substrate which becomes a transcriptional regulator upon its APP-like processing[42], is upregulated along with its direct target, *HES1* (Supplemental Figure 6g). These findings illustrate that the fundamental disease mechanism for EOFAD,

dedifferentiation, precedes the hallmark inflammatory and pathological consequences of the disease.

While EOFAD accounts for less than 5% of all AD cases, it undoubtedly shares the hallmark pathological signatures and symptoms of the late-onset, sporadic form of the disease. However, the similarities between the etiologies of these two forms of AD with respect to the early-stage neurodegenerative mechanisms preceding and driving the onset of pathology and cognitive decline are less well known. A previous analysis of over 600 sporadic AD and non-demented, age control patient brains identified EMT, a non-ectoderm dedifferentiation process significantly enriched in our *PSEN1*^{M146L} and *PSEN1*^{A246E} neurons as well as the *PSEN1*^{M139T} brains, as the top key upregulated gene program which facilities the progression of healthy aging into neurodegenerative AD[43]. Interestingly, miR-124, identified in our study, is the top enriched transcriptional regulator for the disease-contributing gene programs identified in this previous study; we found substantial gene overlap between the EMT genes dysregulated in sporadic AD patients and *PSEN1* neurons presented here (Supplemental Figure 6h). Moreover, the NIH AMP-AD program's 'Wall of Targets' list of sporadic AD susceptibility genes identified by multiple omics approaches show an overlap with EOFAD at both the gene and gene program level, with key drivers and markers of neuronal dedifferentiation, including *REST*, *HDAC1*, *ELAV4*, *GABRA1*, *SCN2A*, and *VGF,* similarly identified (Supplemental Figure 6i).

In summary, our model for EOFAD, based on an integrative multi-omics approach, shows that *PSEN1* neurons enter into a dedifferentiated state triggered by alterations in chromatin topology. This cellular state transformation is brought about primarily through the modulation of the lineage-defining factors REST, miR-9, and miR-124. The transcriptional regulatory reactivation of REST, concomitant with the decreased expression of miR-9 and miR-124, leads to neuronal program repression, increased precursor maintenance and non-ectoderm gene expression, and re-entry into the cell cycle, resulting in G2M phase arrest. Consequentially, these neurons undergo cellular remodeling into a more primitive, mixed germ-layer state. The similarity of these disease-defining mechanisms between our study and those based on transcriptional data from both EOFAD and LOAD human brain samples, illustrated by the common dysregulation of key genes and transcriptional programs across different mutations and analysis approaches, demonstrates their validity in vivo and provides a rational basis for novel therapeutic intervention at an earlier stage in the progression of the disease.

References

1. Tanzi, R. E. The Genetics of Alzheimer Disease. *Cold Spring Harb. Perspect. Med.* **2**, a006296 (2012).
2. Dawson, T. M., Golde, T. E. & Lagier-Tourenne, C. Animal models of neurodegenerative diseases. *Nat. Neurosci.* **21**, 1370 (2018).
3. Chételat, G. Alzheimer disease: Aβ-independent processes—rethinking preclinical AD. *Nat. Rev. Neurol.* **9**, 123–124 (2013).
4. Van Cauwenberghe, C., Van Broeckhoven, C. & Sleegers, K. The genetic landscape of Alzheimer disease: clinical implications and perspectives. *Genet. Med.* **18**, 421–430 (2016).
5. Israel, M. A. *et al.* Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature* **482**, 216–220 (2012).
6. Yuan, S. H. *et al.* Cell-Surface Marker Signatures for the Isolation of Neural Stem Cells, Glia and Neurons Derived from Human Pluripotent Stem Cells. *PLOS ONE* **6**, e17540 (2011).
7. Balwierz, P. J. *et al.* ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. *Genome Res.* **24**, 869–884 (2014).
8. Yoo, A. S., Staahl, B. T., Chen, L. & Crabtree, G. R. MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. *Nature* **460**, 642–646 (2009).
9. Wu, J. & Xie, X. Comparative sequence analysis reveals an intricate network among REST, CREBand miRNA in mediating neuronal gene expression. *Genome Biol.* **7**, R85 (2006).
10. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).
11. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
12. Mi, H. *et al.* PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* **45**, D183–D189 (2017).
13. Sadasivam, S. & DeCaprio, J. A. The DREAM complex: Master coordinator of cell cycle dependent gene expression. *Nat. Rev. Cancer* **13**, 585–595 (2013).
14. Filipczyk, A. *et al.* Network plasticity of pluripotency transcription factors in embryonic stem cells. *Nat. Cell Biol.* **17**, 1235–1246 (2015).
15. Gomes, A. M. *et al.* Cooperative Transcription Factor Induction Mediates Hemogenic Reprogramming. *Cell Rep.* **25**, 2821-2835.e7 (2018).
16. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **15**, 178–196 (2014).
17. Higuchi, M. *et al.* PRRX1- and PRRX2-positive mesenchymal stem/progenitor cells are involved in vasculogenesis during rat embryonic pituitary development. *Cell Tissue Res.* **361**, 557–565 (2015).
18. Ballas, N., Grunseich, C., Lu, D. D., Speh, J. C. & Mandel, G. REST and Its Corepressors Mediate Plasticity of Neuronal Gene Chromatin throughout Neurogenesis. *Cell* **121**, 645–657 (2005).
19. Dhar, S. S. & Wong-Riley, M. T. T. Coupling of Energy Metabolism and Synaptic Transmission at the Transcriptional Level: Role of Nuclear Respiratory Factor 1 in Regulating both Cytochrome c Oxidase and NMDA Glutamate Receptor Subunit Genes. *J. Neurosci.* **29**, 483–492 (2009).
20. Lonze, B. E. & Ginty, D. D. Function and Regulation of CREB Family Transcription Factors in the Nervous System. *Neuron* **35**, 605–623 (2002).
21. Pieraccioli, M. *et al.* ZNF281 inhibits neuronal differentiation and is a prognostic marker for neuroblastoma. *Proc. Natl. Acad. Sci.* 201801435 (2018). doi:10.1073/pnas.1801435115
22. Lee, D.-F., Walsh, M. J. & Aguilo, F. ZNF217/ZFP217 meets chromatin and RNA. *Trends Biochem. Sci.* **41**, 986–988 (2016).

23. Pittol, J. M. R., Oruba, A., Mittler, G., Saccani, S. & Essen, D. van. Zbtb7a is a transducer for the control of promoter accessibility by NF-kappa B and multiple other transcription factors. *PLOS Biol.* **16**, e2004526 (2018).
24. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343–349 (2011).
25. Dietrich, N. *et al.* REST–Mediated Recruitment of Polycomb Repressor Complexes in Mammalian Cells. *PLOS Genet.* **8**, e1002494 (2012).
26. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395 (2011).
27. Brodeur, J. *et al.* LDLR-related protein 10 (LRP10) regulates amyloid precursor protein (APP) trafficking and processing: evidence for a role in Alzheimer's disease. *Mol. Neurodegener.* **7**, 31 (2012).
28. Seyfried, N. T. *et al.* A Multi-network Approach Identifies Protein-Specific Co-expression in Asymptomatic and Symptomatic Alzheimer's Disease. *Cell Syst.* **4**, 60-72.e4 (2017).
29. Chen, B.-Y. *et al.* Brain-derived neurotrophic factor stimulates proliferation and differentiation of neural stem cells, possibly by triggering the Wnt/β-catenin signaling pathway. *J. Neurosci. Res.* **91**, 30–41 (2013).
30. Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**, W305–W311 (2009).
31. Antonell, A. *et al.* A preliminary study of the whole-genome expression profile of sporadic and monogenic early-onset Alzheimer's disease. *Neurobiol. Aging* **34**, 1772–1778 (2013).
32. Mall, M. *et al.* Myt1l safeguards neuronal identity by actively repressing many non-neuronal fates. *Nature* **544**, 245–249 (2017).
33. Lian, I. *et al.* The role of YAP transcription coactivator in regulating stem cell self-renewal and differentiation. *Genes Dev.* **24**, 1106–1118 (2010).
34. Darnell, R. B. RNA Protein Interaction in Neurons. *Annu. Rev. Neurosci.* **36**, 243–270 (2013).
35. Chiu, Y.-C. *et al.* Foxp2 regulates neuronal differentiation and neuronal subtype specification. *Dev. Neurobiol.* **74**, 723–738 (2014).
36. Saura, C. A., Parra-Damas, A. & Enriquez-Barreto, L. Gene expression parallels synaptic excitability and plasticity changes in Alzheimer's disease. *Front. Cell. Neurosci.* **9**, (2015).
37. Sherwood, R. I. *et al.* Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat. Biotechnol.* **32**, 171–178 (2014).
38. Lee, S. W., Oh, Y. M., Lu, Y.-L., Kim, W. K. & Yoo, A. S. MicroRNAs Overcome Cell Fate Barrier by Reducing EZH2-Controlled REST Stability during Neuronal Conversion of Human Adult Fibroblasts. *Dev. Cell* **46**, 73-84.e7 (2018).
39. Shenoy, A. & Blelloch, R. H. Regulation of microRNA function in somatic stem cell proliferation and differentiation. *Nat. Rev. Mol. Cell Biol.* **15**, 565–576 (2014).
40. Holtman, I. R., Skola, D. & Glass, C. K. Transcriptional control of microglia phenotypes in health and disease. *J. Clin. Invest.* **127**, 3220–3229 (2017).
41. Martin, L. *et al.* Tau protein phosphatases in Alzheimer's disease: The leading role of PP2A. *Ageing Res. Rev.* **12**, 39–49 (2013).
42. Marathe, S., Liu, S., Brai, E., Kaczarowski, M. & Alberi, L. Notch signaling in response to excitotoxicity induces neurodegeneration via erroneous cell cycle reentry. *Cell Death Differ.* **22**, 1775–1784 (2015).
43. Podtelezhnikov, A. A. *et al.* Molecular Insights into the Pathogenesis of Alzheimer's Disease and Its Relationship to Normal Aging. *PLOS ONE* **6**, e29610 (2011).
44. Liu, G.-H. *et al.* Recapitulation of premature ageing with iPSCs from Hutchinson-Gilford progeria syndrome. *Nature* **472**, 221–225 (2011).

45. Brian Bushnell. *BBMap short-read aligner, and other bioinformatics tools*. (2015).
46. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
47. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2016).
48. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
49. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
50. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
51. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
52. Tarbell, E. D. & Liu, T. HMMRATAC, the Hidden Markov ModeleR for ATAC-seq. *bioRxiv* 306621 (2018). doi:10.1101/306621
53. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
54. Zhu, L. J. *et al.* ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* **11**, 237 (2010).
55. Team B.C., Maintainer B.P. *TxDb.Hsapiens.UCSC.hg38.knownGene: Annotation package for TxDb object(s). R package version 3.4.0.* (2016).
56. Gusmao, E. G., Allhoff, M., Zenke, M. & Costa, I. G. Analysis of computational footprinting methods for DNase sequencing experiments. *Nat. Methods* **13**, 303–309 (2016).
57. Li, Z., Schulz, M. H., Zenke, M. & Filho, I. G. C. Identification of Transcription Factor Binding Sites using ATAC-seq. *bioRxiv* 362863 (2018). doi:10.1101/362863
58. Kulakovskiy, I. V. *et al.* HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **46**, D252–D259 (2018).
59. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
60. Bolstad, B. M., Irizarry, R. A., Åstrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
61. Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).
62. Irizarry, R. A. *et al.* Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31**, e15 (2003).
63. Carvalho, B. S. & Irizarry, R. A. A framework for oligonucleotide microarray preprocessing. *Bioinformatics* **26**, 2363–2367 (2010).
64. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).
65. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90-97 (2016).
66. Lachmann, A. *et al.* ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* **26**, 2438–2444 (2010).
67. Sloan, C. A. *et al.* ENCODE data at the ENCODE portal. *Nucleic Acids Res.* **44**, D726–D732 (2016).
68. Bernstein, B. E. *et al.* The NIH Roadmap Epigenomics Mapping Consortium. *Nat. Biotechnol.* **28**, 1045–1048 (2010).

69. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
70. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
71. Croft, D. *et al.* Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* **39**, D691–D697 (2011).
72. Mootha, V. K. *et al.* PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
73. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).
74. Hsu, S.-D. *et al.* miRTarBase: a database curates experimentally validated microRNA–target interactions. *Nucleic Acids Res.* **39**, D163–D169 (2011).

Acknowledgements

This study was supported by the Alzheimer's Association New Investigator Research Award (NIRG-14-322164) to S.H.Y.; NIH grants P50 AGO5131 (D.R.G.), R34 AG049646 (D.R.G. and S.L.W), U01 NS 074501-05 (S.L.W.), VA RR&D 1I01RX002259 (S.L.W.) and Cure Alzheimer's Fund (CAF) grants to S.L.W and R.E.T.; R01 LM012595 (S.S.), U01 CA198941 (S.S.), U01 DK097430 (S.S.), R01 HD084633 (S.S.), and R01 HL106579-07 (S.S.); NSF grant STC CCF-0939370 (S.S.). The authors thank Dr. Kristen Jepsen and the UC San Diego Institute for Genomics Medicine for their assistance with the sequencing performed in this research and Dr. Randy Bateman and the Dominantly Inherited Alzheimer's Network (DIAN) for certain EOFAD patient fibroblast lines.

Author Contributions

A.B.C. performed the ChIP-Seq and ATAC-Seq experiments, carried out the data processing and systems analysis of the RNA-Seq, ChIP-Seq, and ATAC-Seq experiments, developed the mechanistic framework, and wrote the manuscript with contributions to the Methods section from other authors; A.B.C. and S.S. were involved in the characterization of the mechanisms and networks associated with EOFAD; S.H.Y. was involved in the design of hiPSC differentiation into neurons and the validation experiments along with S.L.W. and S.S.; Q.L. carried out the hiPSC-neuron preparation, RNA isolation, qPCR validation, and cell cycle assay experiments; G.P.S. contributed to the RNA-, ChIP-, and ATAC-Sequencing; D.R.G. was involved in part of the study design; R.E.T. contributed to the revisions of the manuscript; S.S. supervised the overall study design, analysis, and revisions of the manuscript.

Author Information

The authors declare no competing financial interests. The contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government. Correspondence and requests for materials should be addressed to S.S. (shankar@ucsd.edu).

Data Availability

All sequencing presented in this work is available at the NCBI GEO under the SuperSeries accession GSE123998, encompassing RNA-Seq (GSE123891), ATAC-Seq (GSE123997), and ChIP-Seq (GSE123884) datasets.

Methods

hiPSC line generation

Non-Demented Control (NDC), *PSEN1*^{M146L}, and *PSEN1*^{H163R} fibroblasts were derived from skin biopsies in accordance with UC San Diego IRB approval, whereas *PSEN1*^{A246E} fibroblasts were obtained commercially (Coriell Cat. AG06840). The generation and characterization of the hiPSC lines were carried out as previously reported [5,44] from fibroblasts by retroviral transduction using the reprogramming factors *OCT4*, *SOX2*, *KLF4*, and *C-MYC*. These four iPSC lines were used for all downstream neuron differentiation applications.

Human neuron preparation

The protocol used for neuron preparation was previously described[6]. Briefly, neural stem cells (NSCs) were expanded in NSC growth medium (DMEM /F12, Glutamax™ (Thermo Fisher Cat. 10565018), 1x B-27 (Thermo Fisher Cat. 17504044), 1x N-2 (Thermo Fisher Cat. 17502001), 1x Penicillin-Streptomycin (Thermo Fisher Cat. 15070063), and 20ng/mL human bFGF-2 (BioPioneer Cat. HRP-0011)). At 80% confluence, the medium was changed to neural differentiation medium (DMEM /F12, Glutamax™, 1x B-27, 1x N-2, and 1x Penicillin-Streptomycin) for 3 weeks of differentiation. After differentiation, the cultures were dissociated with Accutase (Sigma Cat. A6964). Cells were resuspended in 200µL of IMag buffer (1x neural differentiation medium, 0.5µM EDTA, 0.5% Bovine Serum Albumin) followed by incubation with PE Mouse Anti-Human CD184 and CD44 antibodies (BD Biosciences Cat. 561733 and 561858, respectively) for 15min on ice in the dark. The mixture was washed with IMag buffer and subsequently incubated with magnetic beads (BD Biosciences) for 30min at room temperature. Magnetic bead separation was carried out for 8min according to the manufacturer's protocol (BD Biosciences). The supernatant containing purified CD184$^-$/CD44$^-$ neurons was then removed and spun down for downstream applications.

RNA Sequencing and Data Processing

For RNA-Seq, total RNA from magnetically purified human NDC, *PSEN1*^{M146L}, and *PSEN1*^{A246E} (replicates, n = 4) hiPSC-derived neurons was prepared using RNeasy Plus Micro Kit (Qiagen Cat. 74034) according to the manufacturer's protocol. On-column DNAse digestion was subsequently performed on total RNA extracts according to the manufacturer's recommendations to remove any genomic contamination (Qiagen Cat. 79254). Libraries were prepared for RNA-Seq using the TruSeq Stranded Total RNA Library prep kit (Illumina, Cat. RS-122-2303) by the Ribo-Zero ribosomal RNA reduction method (Illumina, Cat. MRZG12324). Samples were sequenced at the UC San Diego Institute for Genomics Medicine (IGM) sequencing core on an Illumina HiSeq4000 generating Paired-End, 100bp reads with an average of 25 million reads per sample (Illumina, Cat. FC-410-1001).

RNA-Seq data preprocessing was performed using the BBDuk package (v37.95) in the BBTools suite[45], removing sequencing adaptors and selecting for all paired-end reads above a quality score threshold (Phred Q>10). Trimmed RNA-Seq reads were mapped to the GRCh38.p12 human transcriptome using Salmon v0.9.12[46] with the options --gcBias --seqBias --biasSpeedSamp 5, followed by transcript level summation to the gene level using the R package tximport v1.8.0[47]. Differential expression analysis was performed for *PSEN1*^{M146L} and *PSEN1*^{A246E} iPSC-derived neurons relative to NDC using the R package DESeq2 v1.20.0[48], selecting for genes with a q-value (FDR-adjusted p-value) of <0.01 and a $\log_2$ fold-change of >0.5. These lists of differentially expressed genes (DEGs) for the *PSEN1*^{M146L} and *PSEN1*^{A246E} conditions were used for all subsequent analysis.

ATAC-Sequencing and Data Processing

ATAC-Seq transposition experiments were performed as previously reported[49] on 50,000 NDC and *PSEN1*^{M146L} (replicates, *n* = 2) hiPSC-derived neurons, using the Illumina Nextera DNA Sample Preperation kit (Illumina, Cat. 15028523) and Qiagen MinElute PCR Purification kit (Qiagen, Cat. 28004). ATAC-Seq libraries were generated from transposed DNA using the Kapa Biosystems Real-Time Library Amplification kit (Kapa Biosystems, Cat. 07959028001) according to the manufacturer's recommendations, monitoring amplification by qPCR and stopping the reaction when all samples reached a fluorescence amplification intensity between standards 1 and 3. ATAC-Seq libraries were further purified using the Qiagen MinElute PCR Purification kit and sequenced at the UC San Diego IGM sequencing core on an Illumina HiSeq4000 platform generating Paired-End, 50bp reads with an average of 25 million reads per sample.

ATAC-Seq data preprocessing was performed using the BBDuk package (v37.95) in the BBTools suite, removing sequencing adaptors and selecting for all paired-end reads above a quality score threshold (Phred Q>10). Trimmed reads were aligned to the GRCh38 human genome (GCA_000001405.15 with no alternative analysis) using BBMap v37.95 in the BBtools suite with the options maxindel=20 ambig=random, followed by sorting and indexing of bam files using Samtools v1.3[50], and annotation of PCR duplicates using the Picard v2.3.0 markDuplicates function with the option VALIDATION_STRINGENCY=LENIENT. All duplicates and Mitochondrial, Chromosome X, Chromosome Y, and EBV reads were removed using Samtools v1.3 command view with the options -b -h -f 3 -F 4 -F 8 -F 256 -F 1024 -F 2048. To determine regions of open chromatin, i.e. those accessible by the Tn5 transposase, bam files were first converted to bedGraph format using bedtools v2.25.0[51] functions bamtobed and genomecov, and then finally converted to bigWig format using UCSC bedGraphToBigWig. HMMRATAC v1.0[52] was used to call peaks on the ATAC-Seq data and determine regions of open chromatin with the options –m 50,200,400,600 --score all. These specifically called open regions of chromatin were then passed to the R package *Diffbind* v2.8.0[53] to determine regions of differential accessibility between the NDC and *PSEN1*^{M146L} conditions using the edgeR method. Differentially accessible regions of DNA were annotated using the R packages *ChIPpeakAnno* v3.14.0[54] and *TxDb.Hsapiens.UCSC.hg38.knownGene* v3.4.0[55], defining the promoter region -1000 to 1000 bp from the TSS.

In order to determine the transcription factor footprints associated with the gain or loss of chromatin accessibility, we used HINT v0.11.1[56,57] for the following steps; rgt-hint function footprinting with the options --atac-seq --paired-end --organism=hg38 to identify TF footprints in each sample; rgt-motifanalysis function matching with the options --organism=hg38 --motif-dbs ~/hocomoco to match footprints to known TFs in the HOCOMOCO v11[58] human pwm database; and rgt-hint function differential with the options --organism=hg38 --output-profiles and --window-size 200. Differential HINT footprinting was performed for all differentially accessible regions with corresponding differential gene expression over the entire HMMRATAC-called open chromatin region. In order to generate histograms of ATAC-Seq read density and H3K4Me3 or H3K27Me3 ChIP-Seq read density at the promoter-located summits of differential chromatin accessibility intersecting with differential gene expression, the Homer v4.10 function annotatePeaks was used with the options -size 2000 -hist 20 for ATAC-Seq read density, -size 5000 -hist 50 for H3K4Me3 read density, or -size 10000 -hist 100 for H3K27Me3 read density.

ChIP-Sequencing and Data Processing

For H3K4Me3 and H3K27Me3 ChIP-Seq experiments, NDC and *PSEN1*^{M146L} (replicates, *n* = 1) hiPSC-derived neurons were processed using the Active Motif ChIP-IT High Sensitivity kit according to the manufacturer's recommendations (Active Motif, Cat. 53040). Briefly, cells were washed with cold 1x PBS and collected in a 1.5 mL tube in PBS, followed by formaldehyde fixation for 10 min at room temperature. After addition of stop solution, cells were incubated at room temperature for 5 min and subsequently washed with 1x PBS Wash Buffer and resuspended in

Chromatin Prep Buffer. Following incubation on ice for 10 min, cells were lysed by performing 30 plunges in a chilled dounce homogenizer. Samples were then resuspended in Active Motif ChIP Buffer and chromatin sonication performed using a probe tip sonicator on ice for a total of 10 min per sample (30x 20 sec intervals). Optimal sonicated chromatin fragment lengths were confirmed by agarose gel and used for IP reactions. 500 ng of purified, sonicated chromatin was incubated with 5 µL of H3K4Me3 (Active Motif, Cat. 39159) or H3K27Me3 (Active Motif, Cat. 39155) Rabbit polyclonal antibody overnight at 4C. After de-crosslinking, ChIP DNA was purified using the Qiagen MinElute PCR Purification Kit (Qiagen, Cat. 28004); libraries for H3K4Me3, H3K27Me3, and 1% Input were generated for NDC and *PSEN1*^{M146L} conditions at the UC San Diego IGM sequencing core using Illumina TruSeq LT adaptors (Illumina Cat. FC-121-2001). Sequencing was performed on an Illumina HiSeq4000 generating Single-Read, 75bp reads with an average of 20 million reads per sample.

ChIP-Seq data preprocessing was performed using the BBDuk package (v37.95) in the BBTools suite, removing sequencing adaptors and selecting for all -end reads above a quality score threshold (Phred Q>10). Trimmed ChIP-Seq reads were aligned to the GRCh38 human genome (GCA_000001405.15 with no alternative analysis) using BBMap v37.95 in the BBtools suite with the options maxindel=20 ambig=random, followed by sorting and indexing of bam files using Samtools v1.3, annotation of PCR duplicates using Picard v2.3.0 markDuplicates with the option VALIDATION_STRINGENCY=LENIENT, and removal of PCR duplicates using Samtools v1.3[50] command view with the options -b -h -F 1024. ChIP-Seq peak calling and differential histone methylation was performed using Homer v4.10[59]; briefly, tags were generated from preprocessed bam files using the function makeTagDirectory, peaks called at a FDR of 0.05 using the function findPeaks with the options -style histone -gsize 2.7e9 -fdr 0.05 -region -size 500 for H3K4Me3 and -style histone -gsize 2.7e9 -fdr 0.05 -region -size 2000 for H3K27Me3, and differential histone methylation calculated using the function getDifferentialPeaks with the options -F 2 and -size 600 for H3K4Me3 and -F 2 and -size 2500 for H3K27Me3. The Homer function annotatePeaks was used to annotate regions of differential histone methylation to the GRCh38 genome as well as generate histograms of ChIP read density at histone methylation peak centers for genes with differential expression and differential histone methylation with the options -size 5000 -hist 50 for H3K4Me3 read density or -size 10000 -hist 100 for H3K27Me3 read density.

Human Brain Transcriptomic Analysis

Microarray data corresponding to NDC and *PSEN1*^{M139T} EOFAD human brain samples from a previously published study[31] were retrieved from the NCBI Gene Expression Omnibus (GEO) database (GSE39420) as raw CEL intensity files. 10 samples from this dataset, all obtained from the posterior cingulate cortex region, were selected for analysis: 6 non-demented control brains samples and 4 EOFAD brain samples from patients harboring the *PSEN1*^{M139T} mutation. Samples were background corrected and $\log_2$ normalized using the 'RMA' function[60–62] in the *oligo* package v1.44.0[63] in R v3.5.0. Differentially expressed genes were determined using the *limma* v3.36.2[64] package, selecting for genes with a FDR-adjusted p-value (q-value) < 0.05.

ISMARA Data Analysis

Quality and adaptor trimmed fastq.gz RNA-Seq files for NDC and *PSEN1*^{M146L} hiPSC-derived neurons as well as CEL microarray files for NDC and *PSEN1*^{M139T} EOFAD human brain samples were uploaded to ismara.unibas.ch[7] for automated processing. For RNA-Seq and microarray experiments, sample averaging was performed following motif enrichment. In order to determine a directional z-score for each enriched motif identified, the z-score for each given motif was multiplied by the sign of the Pearson correlation for the TF associated with the motif and the direction of change in expression for the genes associated with each motif (-1 for downregulated genes, +1 for upregulated genes). For motifs associated with miRNAs, qPCR expression (miR-9,

miR-124) or literature evidence was used to determine a positive of negative correlation with gene targets of the given motif.

Transcription Factor Target and Histone Modification Enrichment

Transcription Factor (TF) target enrichment was carried out on Enrichr[11,65] to investigate the TFs potentially controlling genes found to be differentially expressed[11,65]. Here, we selected three TF gene set libraries for use in target enrichment: ChEA, which uses the Chip-X Enrichment Analysis (ChEA) database for 645 TFs[66], consisting of genes differentially expressed by TFs in previously published work; ENCODE, which contains the 816 TF data sets from the Encyclopedia of DNA Elements (ENCODE) project[67] processed to generate target gene-TF relationships; and the ENCODE/ChEA Consensus database, incorporating the consensus genes targets identified for 104 TFs from all ENCODE and ChEA datasets for each given TF. Differentially up- or downregulated genes were enriched in an unweighted manner for each gene set library, generating an Enrichr Combined Score ranked list of TFs. Significantly ranked TFs resulting from the enrichment for each gene set library are listed; when redundant TFs were identified, the ranking with the higher Combined Score was presented. To determine candidate histone modifications of genes with differential expression and chromatin accessibility, we performed enrichment the ENCODE Histone Modification and Roadmap Epigenomics[68,69] gene set library in Enrichr.

Gene Set Enrichment and Pathway Analysis

Gene Set Enrichment and Pathway Analysis was performed primarily by two complementary approaches: 1) Gene Ontology (GO)[70] and Reactome Pathway[71] enrichment of differentially expressed genes using PANTHER v13.1[12] and 2) pre-ranked Gene Set Enrichment Analysis using the GSEAPreranked function in GSEA v3.0[10,72]. For all enrichments using PANTHER, a list of differentially expressed genes with the corresponding $\log_2$ fold RNA expression was created for each comparison in question and a statistical enrichment test performed using the default PANTHER settings. Enrichments were performed using the expanded GO v1.2 gene sets (Biological Process, Cellular Compartment, Molecular Function), and Reactome Pathway v58 gene sets; enriched gene sets and pathways were ranked by the $-\log_{10}$ of the Bonferroni-adjusted p-value (q-value). For all enrichments using the GSEAPreranked function, a pre-ranked list of genes was generated for each condition (*PSEN1*^{M146L} and *PSEN1*^{A246E} hiPSC-derived neurons, *PSEN1*^{M146L} human brain samples) by taking the sign of the $\log^2$ fold expression change relative to NDC multiplied by the $-\log_{10}$ of the p-value for all genes sequenced or probes in the microarray. This ranked list was used for all GSEAPreranked enrichments, using the options -rnd_seed timestamp -scoring_scheme classic -set_max 5000 -set_min 15. GMT files were downloaded from the MSigDB collections for all Hallmark v6.2, Gene Ontology, KEGG[73], and Reactome Pathway gene sets; custom GMT files were generated for ENCODE v2015, ChEA v2016, ENCODE/ChEA Consensus, and miRTarBase v2017[74] by downloading each database from Enrichr and converting into the GSEA GMT format. For MYT1L preranked GSEA analysis, we generated a GMT file from the list of MYT1L-controlled genes previously reported[32]. These GMT files were used for all GSEA Preranked gene set and TF enrichments. For gene set overrepresentation analysis, the ToppFun function of the ToppGene Suite[30] was used with the Gene Ontology Biological Process gene sets.

miRNA qPCR assay

Total RNA from magnetically purified NDC, *PSEN1*^{M146L}, and *PSEN1*^{A246E} (replicates, n = 3) human iPSC-derived neurons was prepared using miRNeasy Micro Kit (Qiagen Cat. 217084), based on manufacturer's procedures. cDNA was produced using TaqMan Advanced miRNA cDNA Synthesis Kit (Life Technologies, Cat. A28007) according to the manufacturer's instruction. Briefly, miRNA was polyadenylated at 37C for 45min using 20ng total RNA in a 5µL reaction. After

heat-stop of the polyadenylation reaction at 60C for 10min, the reaction was held on ice for 2min. The adaptor was then ligated to the miRNAs at 16C for 1hr using RNA ligase in a 15µl reaction. Next, cDNA was synthesized from the adapter-ligated miRNAs. qRT-PCR for miRNAs were carried out using advanced TaqMan Advanced miRNA assay (Life Technologies, Cat. A25576), with TaqMan Fast Advanced Master Mix (Life Technologies, Cat. 4444557) and TaqMan prime probe sets for miR-124 (Hsa-miR-124-3P, 477879_mir) and miR-9 (Hsa-miR-9-5P, 478214_mir). The qRT-PCR program was as follows: 95C for 10min, leading to 40 cycles at 95C for 10min, followed by 60C for 20sec. The $\log_2$ expression fold change for each gene was calculated by the ΔΔCt method.

qPCR validation

Total RNA from magnetically purified NDC, *PSEN1*^{M146L}, *PSEN1*^{A246E}, and *PSEN1*^{H163R} (replicates, *n* = 3) human iPSC-derived neurons was prepared using RNeasy Plus Micro Kit (Qiagen, Cat. 74034) according to the manufacturer's protocol. cDNA was synthesized from total RNA using random primer with iScript cDNA synthesis kit (BioRad, Cat. 1708890) according to the manufacturer's protocol. qRT-PCR reactions were conducted using a CFX96 thermocycler (BioRad) with cDNA template, using TaqMan Fast Advanced Master Mix (Life Technologies, Cat. 4444557) and TaqMan primer probe sets (ThermoFisher). TaqMan assay ID numbers used: *GAD1* (Hs01065893_m1), *HDAC1* (Hs02621185_s1), *E2F1* (Hs00153451_m1), *CDK1* (Hs00938777_m1), *hes1* (Hs00172878_m1), *tp53* (Hs01034249_m1), *ccl2* (Hs00234140_m1), *CDKN1A* (Hs00355782_m1), and *PGK1* (Hs00943178_g1). The qRT-PCR program was as follows: 95C for 10min, leading to 40 cycles of 95C for 20sec, followed by 60C for 20sec**.** The $\log_2$ expression fold change for each gene was calculated by the ΔΔCt method.

Cell Cycle Assay

Magnetically purified NDC, *PSEN1*^{M146L}, and *PSEN1*^{A246E} (biological replicates, *n* = 3) human iPSC-derived neurons were plated in 96 well plates. Cells were fixed in 70% ethanol and stored at 4C until the day of assay. Cells were stained with FxCycle™ Far Red Stain (Thermo Fisher, Cat. F10348) according to the manufacturer's protocol. Stained cells were analyzed by flow cytometry using an Accuri C6 flow cytometer (BD Biosciences).

Statistical analysis:

GraphPad Prism v7.0e software was used for statistical analysis to calculate either student's t-test or 1-way ANOVA with post-test analysis for Cell Cycle and qPCR assays.

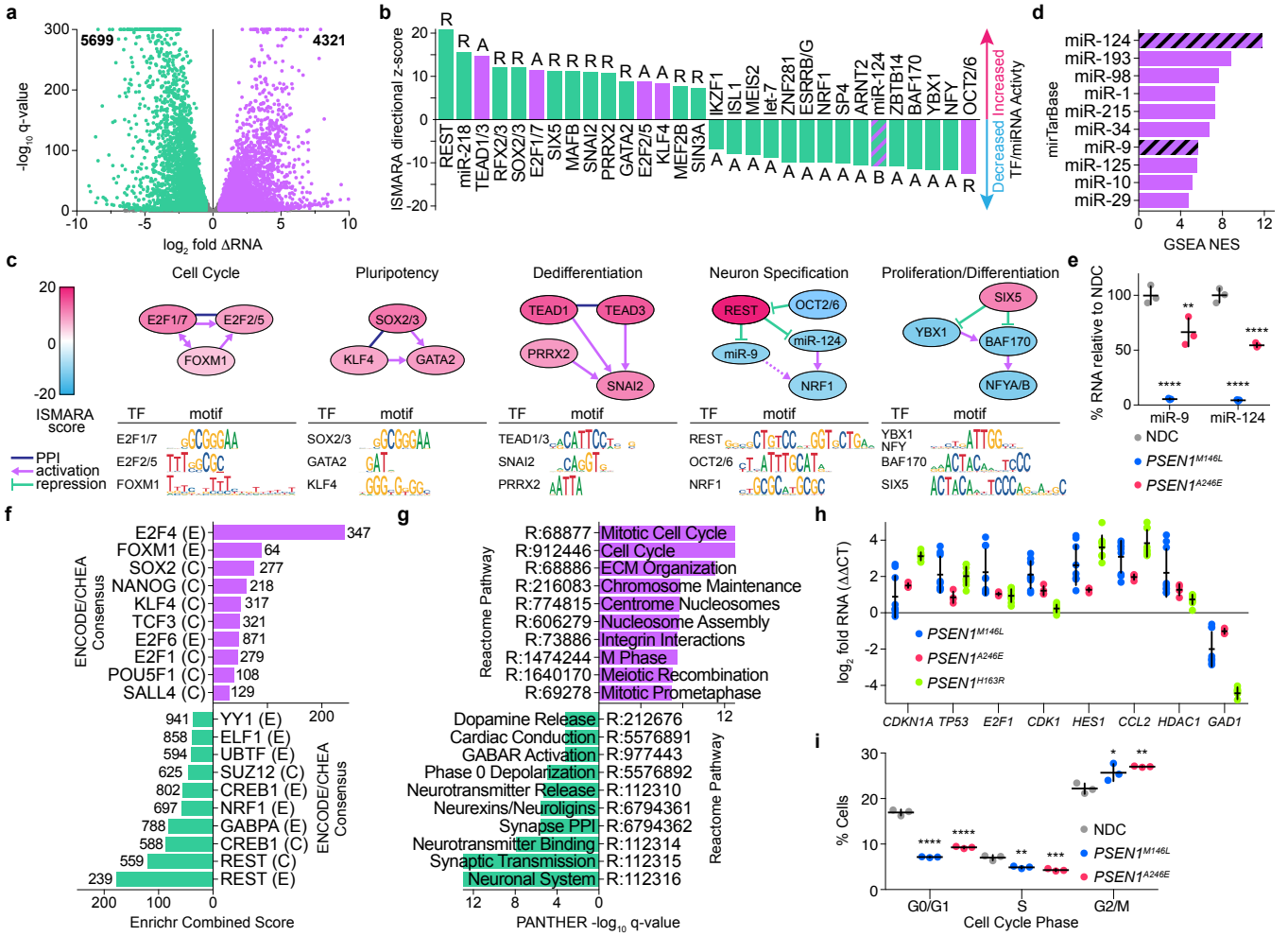


Figure 1

PSEN1^{M146L} hiPSC-derived neurons dedifferentiate through cell cycle re-entry, dedifferentiation, and repression of neuronal lineage programs

a $\log_2$ fold change of differentially expressed genes in *PSEN1^{M146L}* hiPSC-derived neurons relative to non-demented control (NDC) with a False Discovery Rate (FDR) adjusted p-value (q-value) < 0.01. (n = 4) **b** ISMARA motif analysis of RNA-Seq data ranking the TFs and miRNAs whose motif-associated activity is changing most significantly in *PSEN1^{M146L}* hiPSC-derived neurons relative to NDC based on ISMARA z-score and Pearson correlation; corresponding letter indicates each regulator's primary mode of function as an activator (A), repressor (R), or both (B). **c** Curation of TFs and miRNAs identified by ISMARA motif analysis into 5 key modulated gene programs in the *PSEN1^{M146L}* condition. **d** Ranked miRNAs with targets positively enriched in the *PSEN1^{M146L}* condition by preranked Gene Set Enrichment Analysis (GSEA) using the miRTarBase v7.0 miRNA database. **e** Total miRNA levels of miR-9 and miR-124 in *PSEN1^{M146L}* and *PSEN1^{A246E}* hiPSC-derived neurons relative to NDC. p-value: ** = < 0.01, **** = < 0.0001; mean ± SD. (n = 3) **f** Enrichr enrichment of upregulated (purple, top) and downregulated (green, bottom) genes in *PSEN1^{M146L}* hiPSC-derived neurons using the ENCODE/ChEA Consensus TF database. **g** Panther enrichment of differentially expressed genes in hiPSC-derived neurons using the Reactome Pathway database. **h** Confirmation of pathway-relevant differentially expressed genes by qPCR in *PSEN1^{M146L}* , *PSEN1^{A246E}*, or *PSEN1^{H163R}* hiPSC-derived neurons. (n=3) **i** Percentage of NDC, *PSEN1^{A246E}*, or *PSEN1^{M146L}* hiPSC-derived neurons in each cell cycle phase; cells stained with Propium Iodide and assayed by flow cytometry; p-value: * = < 0.05, ** = < 0.01, *** = < 0.001, **** = < 0.0001; mean ± SD. (n = 3)

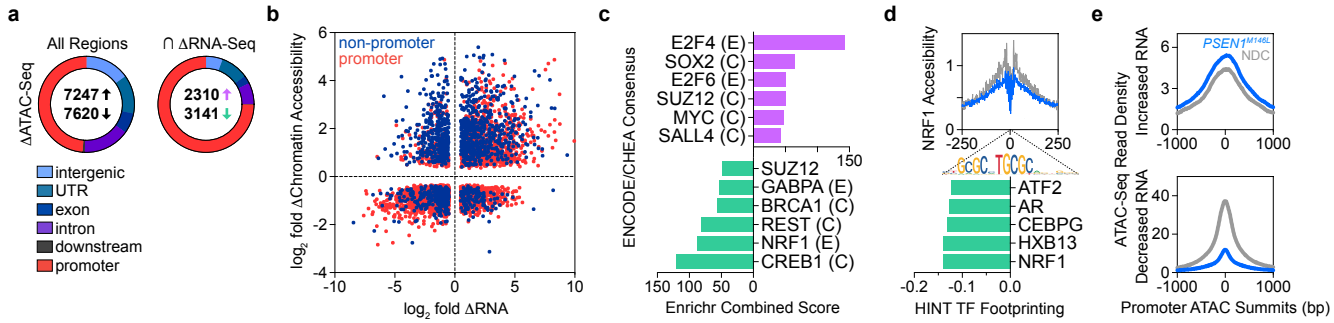


Figure 2

Modulation of chromatin accessibility drives differential gene expression in PSEN1^{M146L} hiPSC-derived neurons

a Differentially accessible regions of chromatin in *PSEN1^{M146L}* hiPSC-derived neurons relative to NDC as measured by ATAC-Seq; left, all differentially accessible regions; right, differentially expressed genes with differentially accessible regions of chromatin. (n = 2) **b** $\log_2$ fold change in chromatin accessibility and RNA expression in genes with differential chromatin accessibility and gene expression. **c** Enrichr enrichment of upregulated (purple, top) and downregulated (green, bottom) genes with differentially chromatin accessibility using the ENCODE/ChEA Consensus TF database. **d** Bottom, transcription factors with decreased binding motif activity in regions of differentially accessible chromatin by HINT TF footprint analysis; top, ATAC-Seq read profiles around NRF1 motifs in NDC and *PSEN1^{M146L}* hiPSC-derived neurons. **e** ATAC-Seq read profiles around the summits of differential chromatin accessibility for genes with upregulated (top) or downregulated (bottom) gene expression.

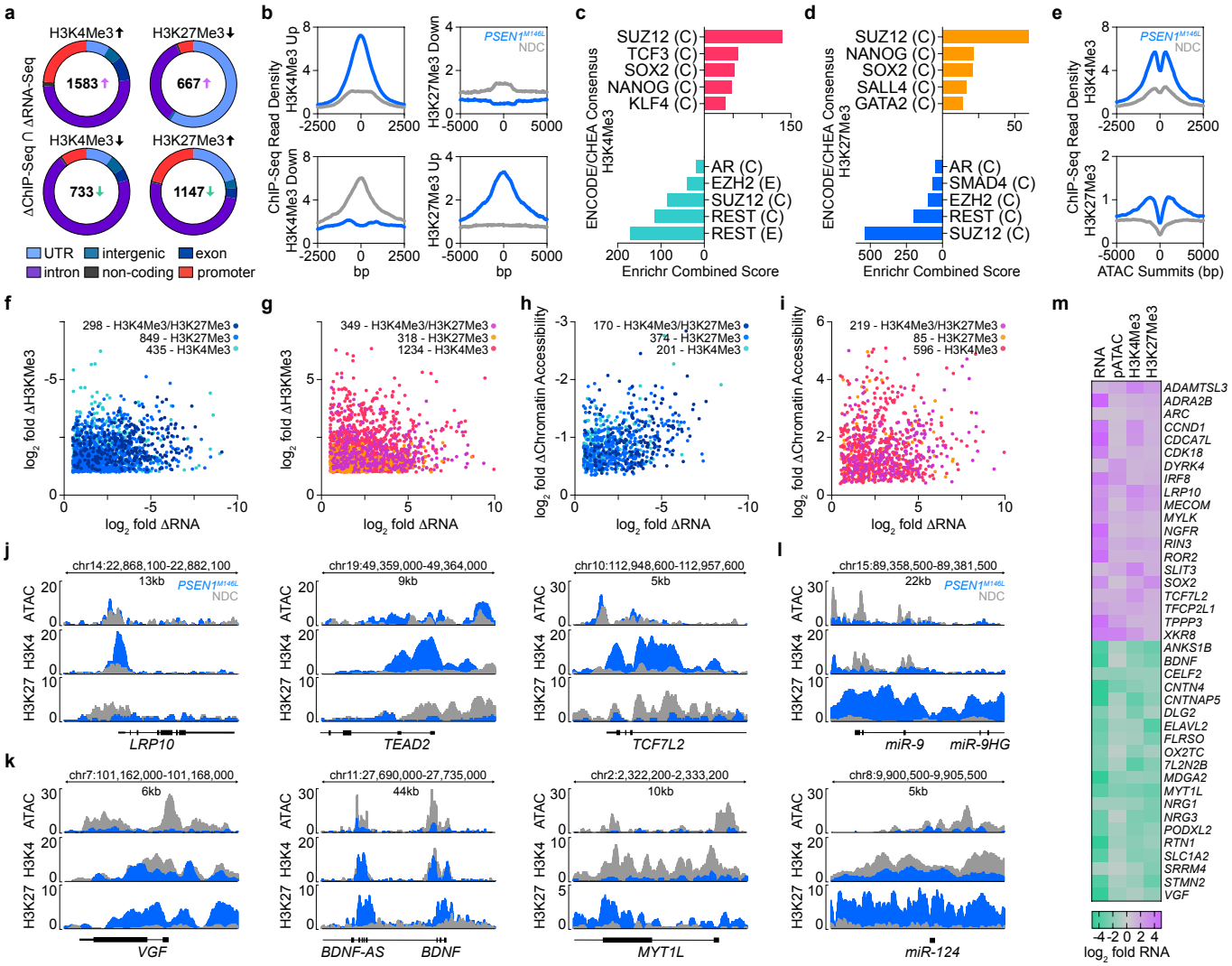


Figure 3

Dedifferentiation in PSEN1^{M146L} hiPSC-derived neurons is caused by changes in histone methylation leading to modulation of chromatin topology

a Differentially expressed genes with corresponding loss or gain of H3K4Me3 or H3K27Me3. **b** ChIP-Seq read profiles around the summits of differential H3K4Me3 or H3K27Me3 methylation for differentially expressed genes. **c-d** Enrichr enrichment of differentially expressed genes with **c** increased H3K4Me3 (top) or decreased H3K4Me3 (bottom) and **d** decreased H3K27Me3 (top) or increased H3K27Me3 (bottom) using the ENCODE/ChEA Consesus TF database. **e** ChIP-Seq read profiles for H3K4Me3 at summits of increased chromatin accessibility for genes with increased expression (top) or for H3K27Me3 at summits of decreased chromatin accessibility for genes with decreased expression (bottom). **f-g** $\log_2$ fold change in H3K trimethylation and RNA expression in genes with **f** decreased gene expression or **g** increased gene expression. **h-i** $\log_2$ fold change in chromatin accessibility and RNA expression in genes with **h** decreased gene expression and corresponding change in H3K trimethylation state or **i** increased gene expression and corresponding change in H3K trimethylation state. **j-l** ATAC-Seq, H3K4Me3, and H3K27Me3 read profiles around the promoter regions of **j** genes with increased expression, **k** genes with decreased expression, or **l** miRNAs with decreased expression. **m** Heatmap of $\log_2$ fold change in RNA expression (RNA), chromatin accessibility within the promoter (pATAC), H3K4Me3 status, or H3K27Me3 (projected in negative $\log_2$ space to correspond to expression) status for key upregulated and downregulated genes.

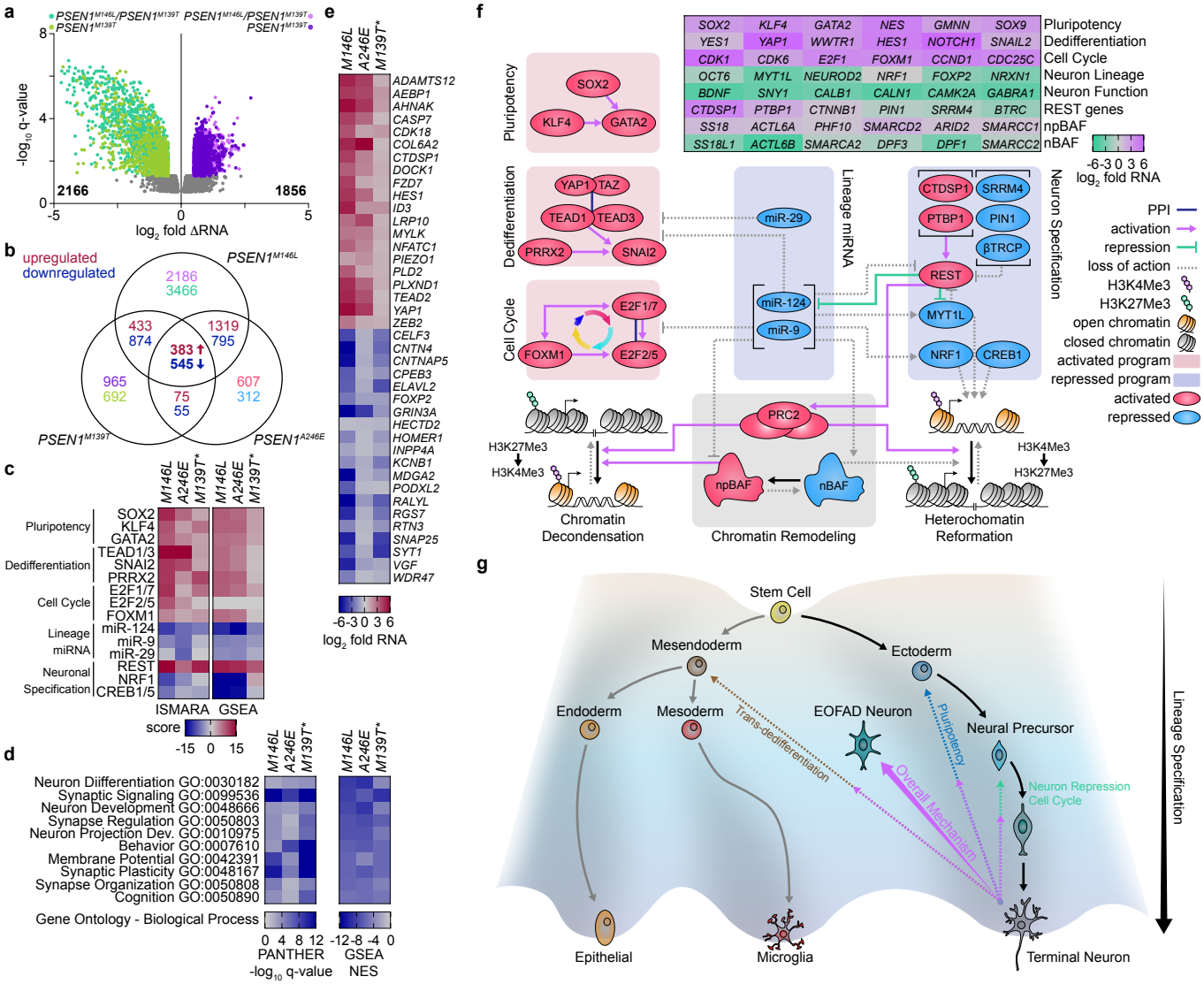
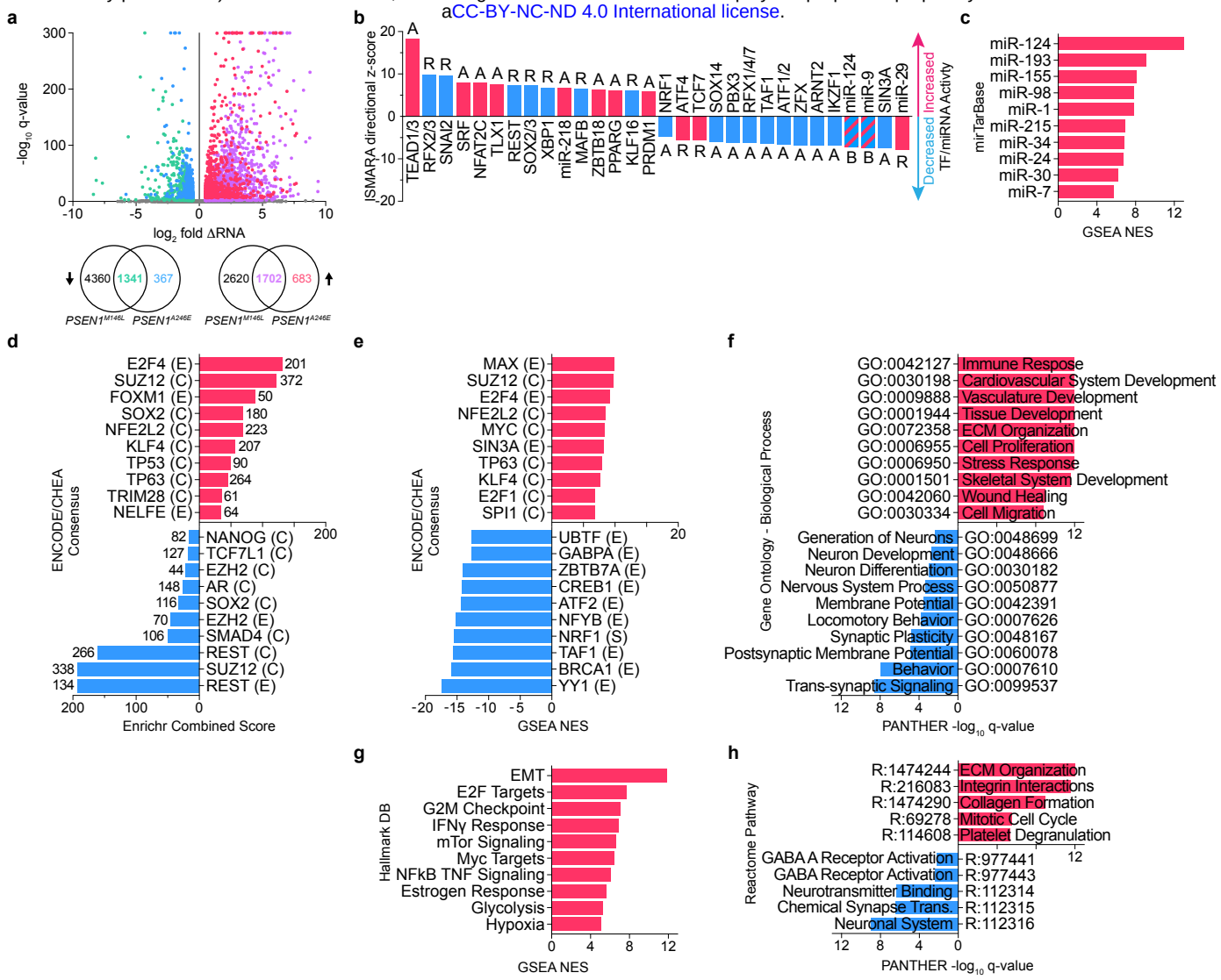


Figure 4

Mechanisms of neuronal dedifferentiation occurring in PSEN1 hiPSC-derived neurons are observed in EOFAD patient brains

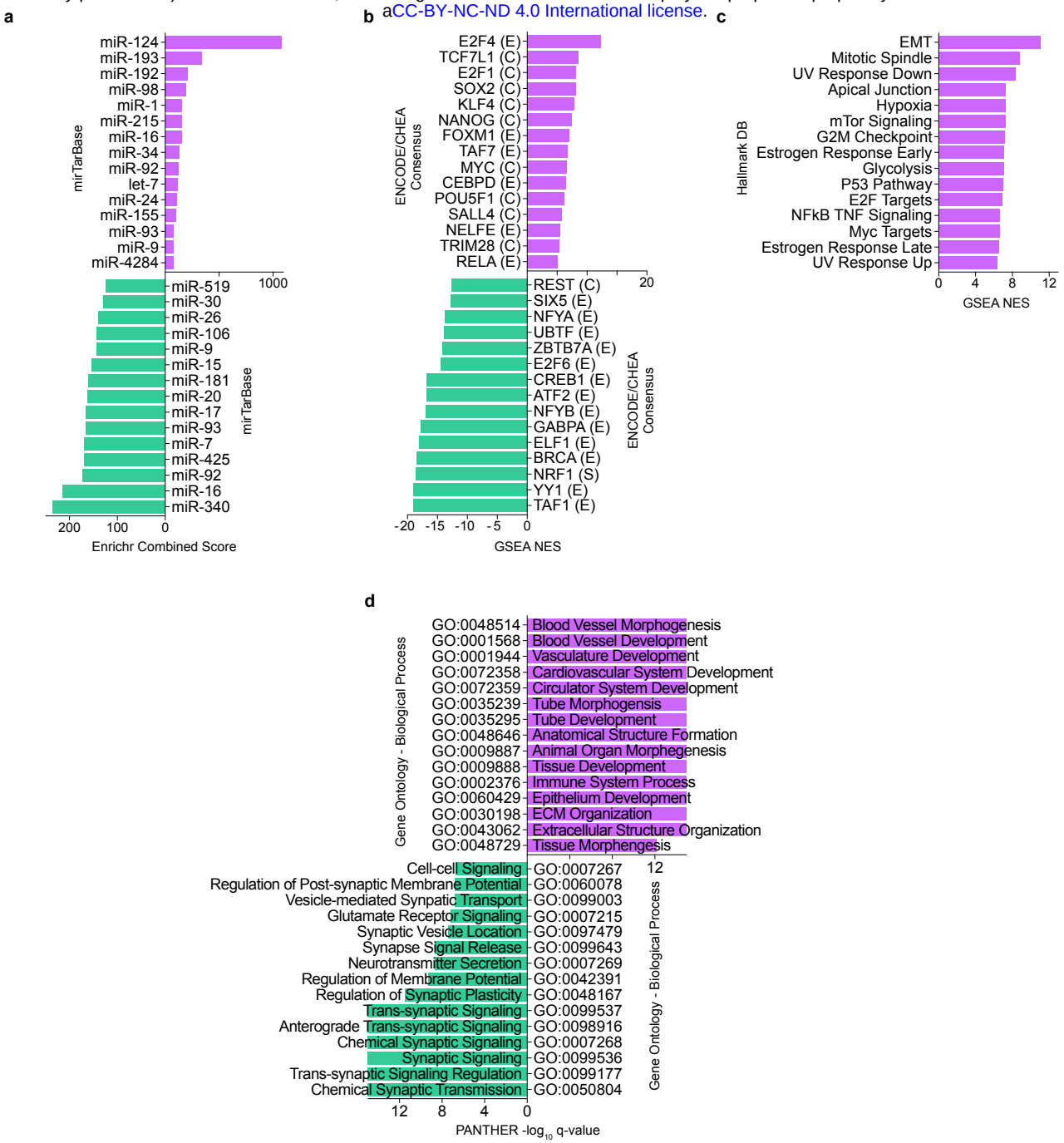
a $\log_2$ fold change of differentially expressed genes in *PSEN1*^{M139T} human brain samples relative to non-demented control (NDC) with a False Discovery Rate (FDR) adjusted p-value (q-value) < 0.05. (n = 4) **b** Venn diagram of differentially expressed genes in common between *PSEN1*^{M139T} human brain and *PSEN1*^{M146L} or *PSEN1*^{A246E} hiPSC-derived neurons. **c** Heatmap of ISMARA motif analysis (left) or pre-ranked GSEA TF enrichment (right) directional scores indicating gain or loss of activity for key TFs and miRNAs involved in neuronal dedifferentiation for *PSEN1*^{M139T} human brain and *PSEN1*^{M146L} or *PSEN1*^{A246E} hiPSC-derived neurons. d Top 10 downregulated Gene Ontology – Biological Process gene sets by Panther (left) or pre-ranked GSEA (right) enrichment in *PSEN1*^{M139T} human brain and *PSEN1*^{M146L} or *PSEN1*^{A246E} hiPSC-derived neurons. **e** Key commonly differentially expressed genes involved in neuronal dedifferentiation or previously identified as Alzheimer's disease markers in *PSEN1*^{M139T} human brain and *PSEN1*^{M146L} or *PSEN1*^{A246E} hiPSC-derived neurons. **f** Mechanistic integration of modulated genes and gene programs leading to remodeling of the chromatin state and dedifferentiation in the EOFAD disease state. Top right: heatmap of differentially expressed genes contributing to these mechanisms in the *PSEN1*^{M146L} condition. **g** Repression of neuron specification and neuron lineage miRNAs concomitant with activation of cell cycle, pluripotency, and dedifferentiation programs reverts EOFAD neuron along an alternative axis into a less differentiated, precursor state with non-ectoderm lineage characteristics.



Supplemental Figure 1

Mechanisms of neuronal dedifferentiation are similarly observed in PSEN1^{A246E} hiPSC-derived neurons

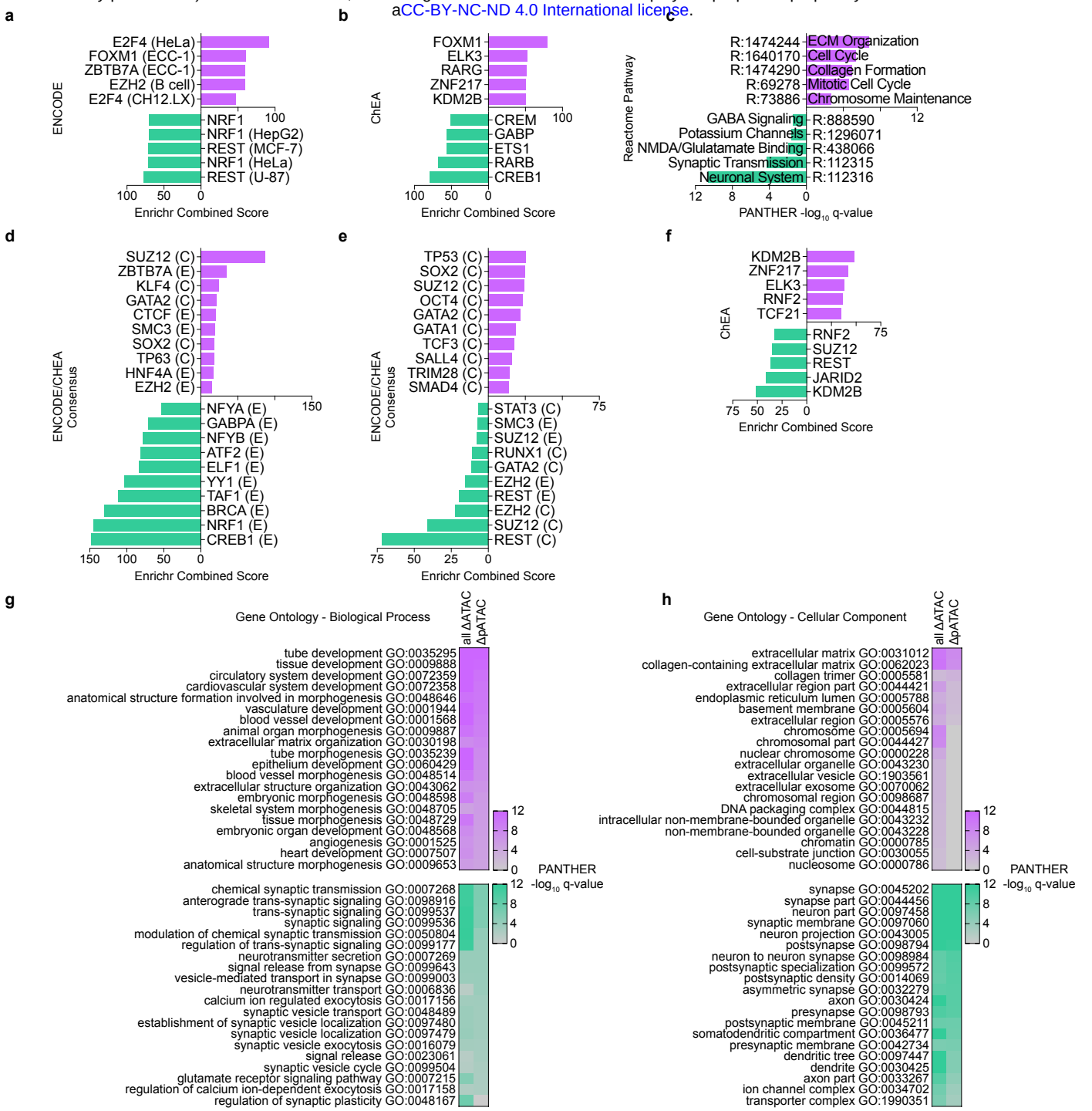
a Top, $\log_2$ fold change of differentially expressed genes in *PSEN1^{A246E}* hiPSC-derived neurons relative to non-demented control (NDC) with a False Discovery Rate (FDR) adjusted p-value (q-value) < 0.01. (n = 4); bottom, intersection of differentially upregulated and downregulated genes between *PSEN1^{A246E}* and *PSEN1^{M146L}* hiPSC-derived neurons. **b** ISMARA motif analysis of RNA-Seq ranking the TFs and miRNAs whose motif-associated activity is changing most significantly in *PSEN1^{A246E}* hiPSC-derived neurons relative to NDC based on ISMARA z-score and Pearson correlation. **c** Ranked miRNAs with targets positively enriched in the *PSEN1^{A246E}* condition by preranked Gene Set Enrichment Analysis (GSEA) using the miRTarBase v7.0 miRNA database. **d** Enrichr enrichment of upregulated (red, top) and downregulated (blue, bottom) genes in *PSEN1^{A246E}* hiPSC-derived neurons using the ENCODE/ChEA Consensus TF database. **e** Ranked TFs with targets positively enriched (red, top) or negatively enriched (blue, bottom) in the *PSEN1^{A246E}* condition by preranked Gene Set Enrichment Analysis (GSEA) using the ENCODE/ChEA Consensus TF database. **f** Panther enrichment of upregulated (red, top) and downregulated (blue, bottom) genes in *PSEN1^{A246E}* hiPSC-derived neurons using the Gene Ontology – Biological Process gene set database. **g** Ranked Hallmark gene sets with members positively enriched in the *PSEN1^{A246E}* condition by preranked Gene Set Enrichment Analysis (GSEA) using the Hallmark database. **h** Panther enrichment of differentially expressed genes in *PSEN1^{A246E}* hiPSC-derived neurons using the Reactome Pathway database.



Supplemental Figure 2

PSEN1^{M146L} hiPSC-derived neurons dedifferentiate through cell cycle re-entry, dedifferentiation, and repression of neuronal lineage programs

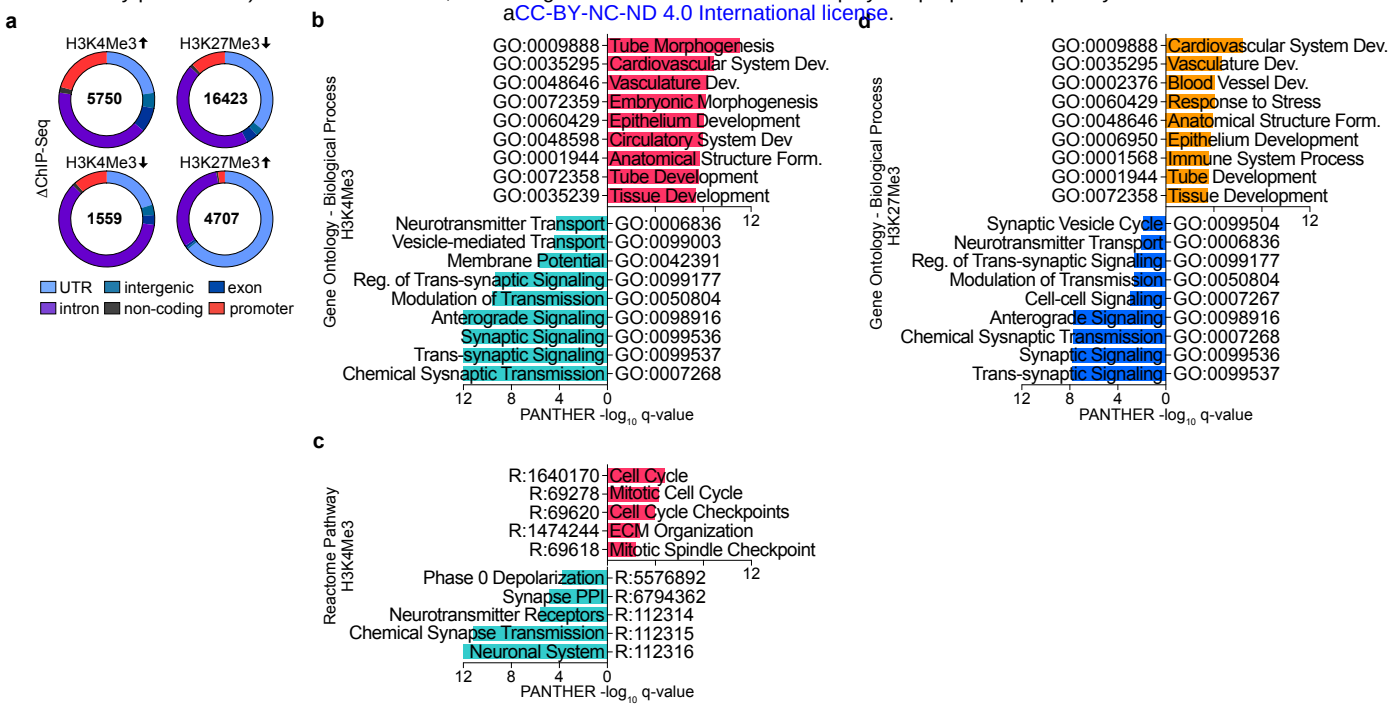
a Enrichr enrichment of upregulated (purple, top) and downregulated (green, bottom) genes in *PSEN1^{M146L}* hiPSC-derived neurons using the miRTarBase v7.0 miRNA database. **b** Ranked TFs with targets positively enriched (purple, top) or negatively enriched (green, bottom) in the *PSEN1^{M146L}* condition by preranked Gene Set Enrichment Analysis (GSEA) using the ENCODE/ChEA Consensus TF database. **c** Ranked Hallmark gene sets with members positively enriched in the *PSEN1^{M146L}* condition by preranked Gene Set Enrichment Analysis (GSEA) using the Hallmark database. **d** Panther enrichment of upregulated (purple, top) and downregulated (green, bottom) genes in *PSEN1^{M146L}* hiPSC-derived neurons using the Gene Ontology – Biological Process gene set database.



Supplemental Figure 3

REST- and NRF1-regulated genes are enriched among the downregulated genes with differential chromatin accessibility

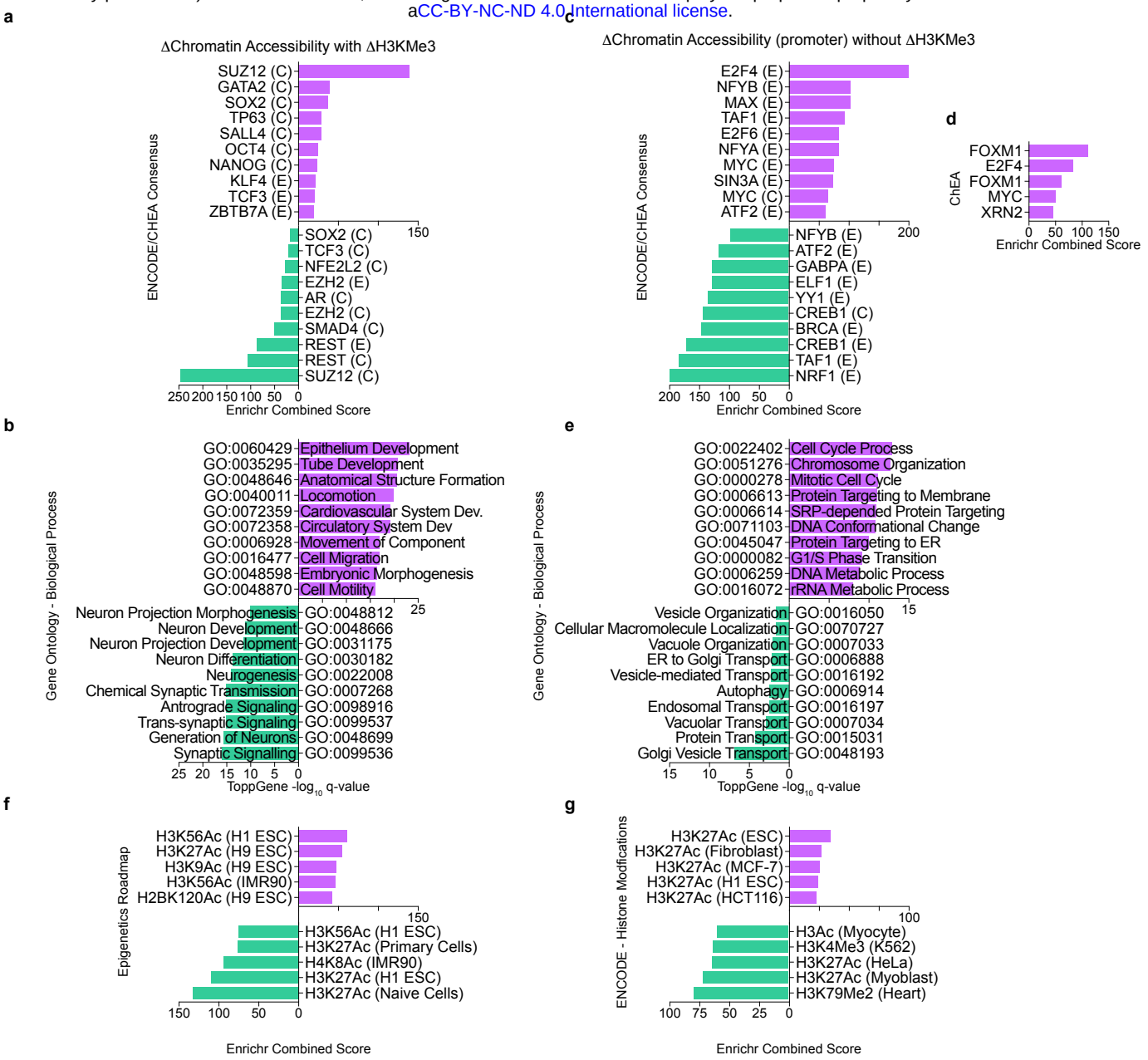
a-b Enrichr enrichment of upregulated (purple, top) and downregulated (green, bottom) genes with differential chromatin accessibility by ATAC-Seq in *PSEN1*^{M146L} hiPSC-derived neurons using the **a** ENCODE or **b** ChEA TF database. **c** Panther enrichment of differentially expressed genes with differential chromatin accessibility by ATAC-Seq in *PSEN1*^{M146L} hiPSC-derived neurons using the Reactome Pathway database. **d-e** Enrichr enrichment of upregulated (purple, top) and downregulated (green, bottom) genes with differential chromatin accessibility by ATAC-Seq in *PSEN1*^{M146L} hiPSC-derived neurons occurring **d** within the promoter region or **e** outside the promoter region using the ENCODE/ChEA Consensus TF database. **f** Enrichr enrichment of upregulated (purple, top) and downregulated (green, bottom) genes with differential chromatin accessibility by ATAC-Seq in *PSEN1*^{M146L} hiPSC-derived neurons occurring outside the promoter region using the ChEA TF database. **g-h** Panther enrichment of differentially expressed genes with differential chromatin accessibility by ATAC-Seq in *PSEN1*^{M146L} hiPSC-derived neurons for all regions (ΔATAC) or just gene promoter regions (Δ pATAC) using the Gene Ontology **g** Biological Process or **h** Cellular Component gene set database.



Supplemental Figure 4

Change in H3K4Me3 and H3K27Me3 methylation status are associated with dedifferentiation and neuronal repression

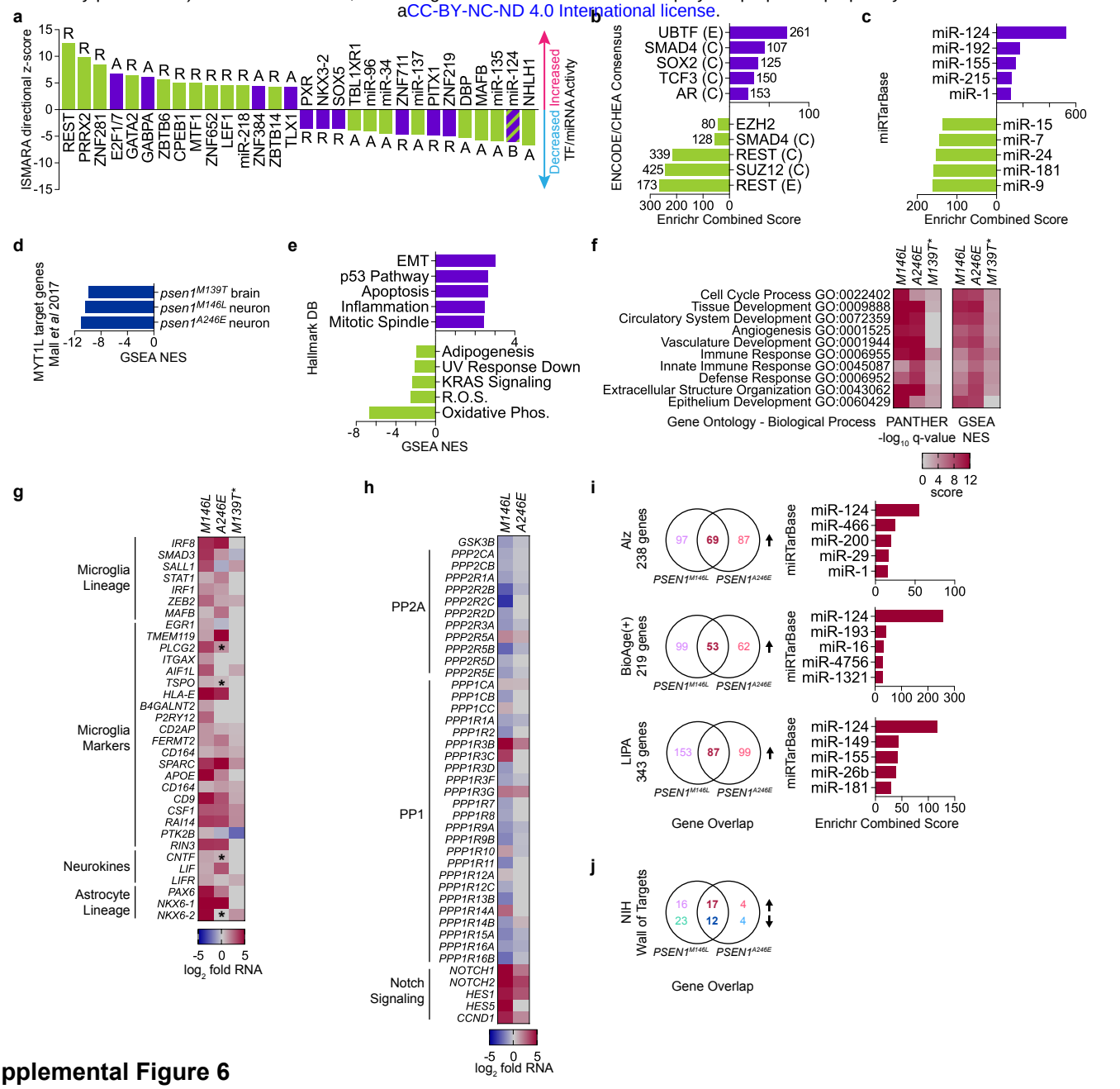
a Differentially methylated histone regions in *PSEN1*^{M146L} hiPSC-derived neurons relative to NDC as measured by ChIP-Seq for H3K4Me3 or H3K27Me3. (n = 1) **b-c** Panther enrichment of differentially expressed genes with differential H3K4Me3 methylation by ChIP-Seq in *PSEN1*^{M146L} hiPSC-derived neurons using the b Gene Ontology – Biological Process or **c** Reactome Pathway database. **d** Panther enrichment of differentially expressed genes with differential H3K27Me3 methylation by ChIP-Seq in *PSEN1*^{M146L} hiPSC-derived neurons using the Gene Ontology – Biological Process database.



Supplemental Figure 5

Change in H3K4Me3 and H3K27Me3 methylation status are associated with the neuronal repressor REST, PRC2, and pluripotent factors while change in chromatin accessibility are associated with neuronal mitochondrial TFs NRF1 and CREB1 (loss) and cell cycle TFs E2Fs and FOXM1 (gain)

a-b Enrichment of TFs and gene ontologies using **a** Enrichr enrichment and **b** ToppGene overrepresentation enrichment of upregulated (purple, top) and downregulated (green, bottom) genes in *PSEN1*^{M146L} hiPSC-derived neurons using the ENCODE/ChEA Consensus TF database for genes with differential chromatin accessibility and differential histone methylation (H3K4Me3 and/or H3K27Me3). **c-e** Enrichment of TFs and gene ontologies using **c-d** Enrichr enrichment and **e** ToppGene overrepresentation enrichment of upregulated (purple, top) and downregulated (green, bottom) genes in *PSEN1*^{M146L} hiPSC-derived neurons using the Gene Ontology – Biological Process database for for genes differential chromatin accessibility in the promoter region alone (no change in H3KMe3 status). **f-g** Enrichr enrichment of upregulated (purple, top) and downregulated (green, bottom) genes in *PSEN1*^{M146L} hiPSC-derived neurons using the **f** Epigenetics Roadmap database or **g** ENCODE Histone Modifications database for genes with differential chromatin accessibility in the promoter region alone (no change in H3KMe3 status).



Supplemental Figure 6

PSEN1^{M139T} human brain samples exhibit similar mechanistic signatures of neuron dedifferentiation

a ISMARA motif analysis of RNA-Seq ranking the TFs and miRNAs whose motif-associated activity is changing most significantly in *PSEN1^{M139T}* human brain to NDC based on ISMARA z-score and Pearson correlation. **b-c** Enrichr enrichment of upregulated (violet, top) and downregulated (green, bottom) genes in *PSEN1^{M139T}* human brain using the **b** ENCODE/ChEA Consensus TF database or **c** miRTarBase v7.0 miRNA database. **d** Negative enrichment of MYT1L-controlled genes from Mall et al 2017 for *PSEN1^{M146L}* neurons, *PSEN1^{A246E}* neurons, or *PSEN1^{M139T}* brain by preranked Gene Set Enrichment Analysis (GSEA). **e** Ranked Hallmark gene sets with members positively (violet, top) or negatively (green, bottom) enriched in the *PSEN1^{M139T}* condition by preranked Gene Set Enrichment Analysis (GSEA) using the Hallmark database. **f** Top 10 upregulated Gene Ontology – Biological Process gene sets by Panther DB (left) or pre-ranked GSEA (right) enrichment in *PSEN1^{M139T}* human brain and *PSEN1^{M146L}* or *PSEN1^{A246E}* hiPSC-derived neurons. **g** Expression of key genes involved in microglia lineage, microglia markers, neurokines, or astrocyte lineage in PSEN1M139T human brain and *PSEN1^{M146L}* or *PSEN1^{A246E}* hiPSC-derived neurons (* indicates *PSEN1^{M146L}* qval < 0.05). **h** Expression of factors involved in and indicative of tau and amyloid pathologies in *PSEN1^{M139T}* or *PSEN1^{A246E}* hiPSC-derived neurons. **i** Overlap with genes upregulated in *PSEN1^{M146L}* or *PSEN1^{A246E}* hiPSC-derived neurons (left) and Enrichr miRNA enrichment (right) for Alz, BioAge(+), and LIPA sporadic Alzheimer's Disease-associated metagene networks identified in Podtelezhnikov et al 2011. **j** Overlap between the 95 NIH Wall of Targets genes identified from sporadic AD cases and upregulated or downregulated genes in *PSEN1* hiPSC-derived neurons.