Mutations in human DNA methyltransferase *DNMT1* induce specific genome-wide epigenomic and transcriptomic changes in neurodevelopment

Kasey N. Davis[1,2,†], Ping-Ping Qu[1,2,†], Shining Ma[3], Ling Lin[1,4], Melanie Plastini[1,2], Niklas Dahl[5], Giuseppe Plazzi[6,7], Fabio Pizza[6,8],

Ruth O'Hara[1], Wing Hung Wong[3,9], Joachim Hallmayer[1], Emmanuel Mignot[1,4], Xianglong Zhang [1,2,†,*] and Alexander E. Urban[1,2,*]

[1]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Palo Alto, CA 94304, USA
[2]Department of Genetics, Stanford University School of Medicine, Palo Alto CA 94304, USA
[3]Department of Statistics, Stanford University, Stanford, CA 94305, USA
[4]Center for Narcolepsy, Stanford University School of Medicine, Palo Alto, CA 94304, USA
[5]Department of Immunology, Genetics and Pathology Sciences for Life Laboratory, Uppsala University BMC, Uppsala 75122, Sweden
[6]IRCCS—Istituto delle Scienze Neurologiche di Bologna, Bologna 40139, Italy
[7]Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena 41125, Italy
[8]Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna 40126, Italy
[9]Department of Biomedical Data Science, Stanford University School of Medicine, Palo Alto, CA 94304, USA

*To whom correspondence should be addressed. Tel.: (650) 736-9528; Fax: (650) 725-4913; Email: aeurban@stanford.edu; Tel.: (650)334-9145; Emailxianglongruoying@gmail.com

†These authors contributed equally to this work.

Abstract

DNA methyltransferase type 1 (DNMT1) is a major enzyme involved in maintaining the methylation pattern after DNA replication. Mutations in *DNMT1* have been associated with autosomal dominant cerebellar ataxia, deafness and narcolepsy (ADCA-DN). We used fibroblasts, induced pluripotent stem cells (iPSCs) and induced neurons (iNs) generated from patients with ADCA-DN and controls, to explore the epigenomic and transcriptomic effects of mutations in *DNMT1*. We show cell type–specific changes in gene expression and DNA methylation patterns. DNA methylation and gene expression changes were negatively correlated in iPSCs and iNs. In addition, we identified a group of genes associated with clinical phenotypes of ADCA-DN, including *PDGFB* and *PRDM8* for cerebellar ataxia, psychosis and dementia and *NR2F1* for deafness and optic atrophy. Furthermore, *ZFP57*, which is required to maintain gene imprinting through DNA methylation during early development, was hypomethylated in promoters and exhibited upregulated expression in patients with ADCA-DN in both iPSC and iNs. Our results provide insight into the functions of *DNMT1* and the molecular changes associated with ADCA-DN, with potential implications for genes associated with related phenotypes.

Introduction

DNMT1 maintains DNA methylation patterns during DNA replication and DNA repair, functions that are essential for maintaining proper cellular epigenetic inheritance. It is essential for the proper regulation of gene transcription, genomic imprinting and survival of various tissues. DNMT1 is particularly important in neuronal survival, differentiation and migration and, in general, in the formation of central nervous system connectivity. DNMT1 deficiency in neural precursor cells results in DNA hypomethylation of daughter cells (1). These hypomethylated neurons come with maturation defects such as increased dendritic arborization and impaired neuronal excitability (2–4). An early mouse model study found that transgenic mice lacking DNMT1 exhibit embryonic lethality before midgestation (5). Recent research using *Dnmt1-M1* and *Dnmt1-M2* knock-in mouse models of mutations different from the ones studied here by us (leading to HSAN1E, see below), but also affecting the RFTS domain of the DNMT1 protein, revealed that neurodegenerative phenotypes, including impaired learning and memory, manifested in mutant heterozygous mice via having reduced DNMT1 proteins while

the homozygous mutants were not viable past the embryonic stage (6).

During organismal development, the genome undergoes multiple series of demethylation and methylation changes, resulting in cell type–specific and tissue-specific DNA methylation patterns. Across most tissues, *DNMT1* expression levels diminish greatly when no cell division is occurring. However, in both embryonic and adult postmitotic neurons, *DNMT1* is highly expressed, further underlining the importance of *DNMT1* in the nervous system (7). For example, a recent study discovered a correlation of DNMT1-activity and synaptic transmission in human induced pluripotent stem cell (iPSC)–derived glutamatergic cortical neurons, suggesting the importance of DNMT1 in the epigenetic regulation of synaptic function in human excitatory neurons (8).

DNA methylation is required in adult neurogenesis, and its misregulation has been reported to be involved in the pathophysiology of neurodegenerative disorders (9) such as autosomal dominant cerebellar ataxia, deafness and narcolepsy (ADCA-DN), the disorder investigated by this study. ADCA-DN was first reported in a Swedish pedigree in which five affected individuals were

described (10). This neurodegenerative disease is characterized by an adult-onset phenotype, with narcolepsy occurring years before the manifestation of deafness and cerebellar ataxia. Patients may also develop optic atrophy, sensorimotor polyneuropathy, psychosis, dementia and diabetes mellitus. Exome sequencing of individuals with ADCA-DN identified point mutations in exon 21 of the DNA methyltransferase type 1 (*DNMT1*) gene, affecting the regulator domain Replication Foci Targeting sequence (RFTS), as the causative type of mutation (11).

Interestingly, *DNMT1* mutations in a different exon than those seen in ADCA-DN also lead to a somewhat similar adult-onset neurological disorder: hereditary sensory and autonomic neuropathy with dementia and hearing loss type 1E (HSAN1E), in which point mutations in exon 20 have been identified (12). In blood cells of patients with HSAN1E and ADCA-DN, global hypomethylation has been reported, while Sun *et al.* observed patients with HSAN1E site-specific hypermethylation in peripheral blood cells (12–14). Furthermore, in *DNMT1* homozygous knockout human embryonic stem cells (hESCs), global demethylation was reported (15). Using the *Dnmt1* knock-in mouse models of these mutations in patients with HSAN1E, the mechanisms by which this mutation of the DNMT1 RFTS domain leads to neurodegenerative disease were explored. It was found that the functional defect of mutant DNMT1 in these cases may potentially be a mutation-induced RFTS cleavage of DNMT1 (6).

How ADCA-DN mutations in *DNMT1* manifest in relevant cell types, notably neurons, and generate the various clinical phenotype of *DNMT1* is unknown. To gain insights into this question and further explore the role of *DNMT1* in neurons, we investigated the presence and nature of cell type–specific changes in DNA methylation and gene expression patterns in ADCA-DN, in induced neurons (iNs) derived from induced pluripotent stem cells (iPSCs). To achieve this objective, we obtained fibroblast cultures from ADCA-DN patients and controls, generated iPSCs and reprogrammed these iPSCs into iNs.

These three cell types are commonly used in studies on neurodegenerative disorders (16–18). The utilization of the fibroblast cells in neurological studies takes advantage of their properties of comprising the chronological and biological aging of the donors according to their genetic and environmental background, in addition to the readily availability and robustness of this cell type (16). Furthermore, use of a non-neuronal cell type may make it possible to put signals from a neuronal cell type with the same mutation into sharper relief. Reprogrammed iPSCs represent the donor's genotype and genetically encoded molecular and cellular phenotypes from early development (19,20) while iPSC-derived early-stage iNs preserve patient-specific genetic architecture (21) while allowing a first view of neuron-specific effects of the mutations of interest. For all three cell types, genome-wide capture-based bisulfite sequencing was used to uncover DNA methylation changes associated with ADCA-DN. In parallel with this, RNA-Seq analysis was conducted to characterize associated gene expression changes.

Materials and Methods
Subjects

Written informed consent was obtained from all patients, three probands and two controls, for the collection of fibroblasts. Blood samples were obtained from the same probands and three controls (one from the two controls for the fibroblast collection above). Phenotypic details for the two Italian probands and Swedish proband have been described previously (10,11). For

Table 1. Symptoms in the three probands with approximate age of onset when documented

Phenotypes	Italian proband 1	Italian proband 2	Swedish proband
DNMT1 mutation	Ala570Val G > A	Gly605Ala C > G	Val606Phe C > A
Age at time of study	57	47	32
Gender	M	M	M
Cerebellar ataxia	46	47	12
Hearing deficit	43	43	4
Narcolepsy/Cataplexy	42	43	36
Psychosis	55	No	No
Dementia	43	No	No
Optic atrophy	55	47	No

additional details on the phenotype of all probands, see Table 1. For sample information of controls, see Supplementary Material, Table S1.

Mutation analysis by Sanger sequencing

Genomic DNA from fibroblasts and iPSCs was extracted with the Blood & Cell Culture DNA mini kit (Qiagen) according to the manufacturer's instructions. Primer pairs were designed to amplify the mutation region using Taq DNA Polymerase with ThermoPol Buffer (New England Biolabs). The PCR conditions were 95°C for 30 s, 30 cycles of 95°C for 30 s, 55°C for 55 s, 68°C for 1 min and 68°C for 5 min after the last cycle. PCR results were confirmed by Sanger sequencing.

iPSC culture

iPSCs were generated with CytoTune sendai reprogramming vectors (ThermoFisher Scientific) from dermal fibroblast from three individuals with ADCA-DN and two unaffected controls. iPSCs were maintained in feeder-free conditions using Matrigel (Corning)-coated 6-well plates in mTeSR1 media (STEMCELL Technologies) at 37°C incubators (5% CO_2) and passaged every 3–4 days with EDTA (Lonza).

Plasmid transformation

Plasmid DNA for pTet-O-Ngn2-puro, pTet-O-FUW-EGFP and FUW-rtTA were gifted from Marius Wernig (22). DNA was transformed into One Shot TOP10 Chemically Competent *E. coli* (Thermo Fisher). Each transformation was spread onto LB agar plates with 100 μg/ml Amplicillin (Teknova) and incubated at 37°C overnight. One clone from each sample was collected and allowed to amplify overnight, shaking at 37°C in LB/amp (100 μg/ml) media. Plasmids were purified using the Qiagen Plasmid Maxi Kit (Qiagen). After plasmid verification, plasmid DNAs were submitted to the Gene Vector and Virus Core (Stanford) for lentiviral production.

Generation of iNs from human iPSCs

iPSCs were maintained under feeder-free conditions in mTeSR1 media (STEMCELL Technologies). Media was changed every day. When cell density reached 75–80% confluence, clones were dissociated with Accutase (ThermoFisher Scientific) and plated onto Matrigel (Corning) –coated 6-well plates at a 1:3 dilution in mTeSR1 containing 10 μM of Y-27632 dihydrochloride (ROCK) (Tocris). On day 1, cells were infected with 2 μl of each lentivirus (RTTA, GFP and Ngn2) in mTeSR1 media containing 10 μM of ROCK. On day 2, mTeSR1 media was replaced with N2 media (DMEM/F12 [ThermoFisher Scientific], N2 [ThermoFisher

Scientific], penicillin/streptomycin [ThermoFisher Scientific]) containing doxycycline (2 μg/ml). Puromycin selection continued from day 3 to 6, with N2 media changes every day. On day 7, DNA and RNA were extracted from all iNs and the treated iNs were stained for neuronal cell markers MAP2 and TUJ1.

Immunocytochemistry

Cells were fixed with 4% paraformaldehyde for 15 min at room temperature. After fixation, cells were washed twice with PBS and incubated in blocking solution (BS) (PBS + 0.02% Triton X-100 + 5% goat serum) for 1 h at room temperature and then with desired primary antibodies at 4°C overnight. Following three washes in PBS, secondary antibodies were applied in the BS for 1 h at room temperature and finally washed with PBS. The following primary antibodies were used: mouse anti-Map2 (1500; Sigma-Aldrich) and rabbit anti-Tubulin β-3 (1100, BioLegend). Secondary antibodies used were goat anti-mouse Alexa Fluor 555 (1:1000; Invitrogen) and goat anti-rabbit Alexa Fluor 555 (11000, Invitrogen). Nuclear counterstaining was performed with 4′,6-diamidino-2-phenylindole (DAPI) from Sigma (11000).

DNA methylation analysis by genome-wide target-capture sequencing

Genomic DNA was extracted from fibroblasts, iPSCs and iNs using the Blood & Cell Culture DNA mini kit (Qiagen) and the Quick-DNA Miniprep Plus Kit (Zymo Research). Library construction was performed using the SeqCap Epi Enrichment System (Roche), as previously described (23). In brief, 1 μg of genomic DNA was sonicated (Covaris) to generate fragments 180–220 bp in size. Following fragmentation, DNA was used to construct library as described in the KAPA library Preparation Kit Illumina platforms (KAPA biosystems). After library construction, DNA libraries were bisulfite-converted and purified using the EZ DNA Methylation Lightning Kit (Zymo Research). Bisulfite converted samples libraries were amplified using LM-PCR. Amplified bisulfite converted samples and SeqCap Epi probes were hybridized for 64–72 h at 47°C. Captured DNA samples were washed and recovered, followed by amplification of the capture DNA samples using LM-PCR. Amplified samples were purified using Agencourt AMPure XP Beads (Beckman Coulter). Final libraries were checked for quality on a Bioanalyzer DNA 1000 chip (Agilent).

Bisulfite converted libraries captured by SeqCap Epi CpGiant Probes kit (the size of target regions is 80.5 Mb with >5.5 million CpGs) of all the samples were sequenced at the Functional Genomic Facility (Stanford University) on Illumina HiSeq 4000 platform by 2 × 150 paired-end sequencing with an average of 70 million reads generated for each sample. The SeqCap Epi CpGiant system guarantees a high sequencing coverage of CpG islands and is very effective for interrogating CpGs in CpG islands, CpG island shores, promoters and known DMRs (24,25). After trimming the adapters and low-quality ends by Cutadapt, the reads were mapped to human RefSeq genome (GRCh38.p10) using Bismark (Version 0.16.3) (26) with an average unique mapping rate of 73.8% of all the samples. Duplicates were removed by the deduplicate_bismark script in Bismark. Only one copy of the overlapping parts in the middle of paired-end reads was retained after clipping the read with the lower average quality in the overlap region by the 'clipOverlap' tool in bamUtil (Version 1.0.14). On-target read rate and coverage were calculated by Qualimap (Version 2.1) (27). The average on-target read rate of all the samples was 65.3%, and average on-target coverage was 34.3X.

The DNA methylation ratio for each CpG was extracted by the bismark_methylation_extractor script in Bismark. For each sample, only CpGs with at least 10 reads covering them were included in the downstream analysis (2 796 542 CpGs). Differentially methylated regions (DMRs) were identified between cell types as well as between patients and controls within each cell type by metilene (Version 0.2–6) (28) with ≥3 CpGs and a mean methylation difference between the two compared groups of ≥ 0.2. DMRs with FDR-corrected *P*-value <0.05 were considered significant. All the DMRs were assigned to genes whose transcription start sites (TSSs) were closet to them. The related genes were referred to as 'genes associated with DMRs'.

To determine bisulfite conversion efficiency of each captured library by the SeqCap Epi CpGiant Probes kit, unmethylated genomic DNA of Enterobacteria phage lambda was added and processed together with the sample DNA in the same tube throughout the library preparation and sequencing. Each kit contains probes to capture the lambda genomic region from base 4500 to 6500. The sequencing reads were aligned to the phage lambda genome (GenBank Accession NC_001416) by Bismark and the bismark_methylation_extractor script was used to extract the methylation ratio for CpGs of the lambda genome. The reads aligned to the captured region of the lambda genome can be used to calculate the overall conversion efficiency as follows: conversion rate = 1 − (sum(C_count)/sum(CT_count)). The average of bisulfite conversion rate of all the samples was 99.3%.

DNA methylation level distribution

All the CpGs with coverage ≥10 (1 433 461 CpGs) in all samples were used to estimate the distribution of the methylation level by the 'density' function in R for all cell lines. Methylation level distribution of each cell type was also estimated using all the CpGs with coverage ≥ 10 in all samples of the cell type (2 461 696 CpGs for blood samples, 1 660 669 for fibroblasts, 19 157 594 for iPSCs and 1 859 066 for iNs). DNA methylation levels were compared between patients with ADCA-DN and controls within each tissue (fibroblast, iPSCs, iNs and blood) and between any two different tissues using the Wilcoxon rank-sum test. Significance testing was done at the *P*-value <0.05 level.

Strand-specific DNA methylation analysis

Considering the function of DNMT1 in the faithful copying of methylation patterns from parental to progeny strand at replication forks (29–31), we examined whether the *DNMT1* mutations lead to strand-specific changes in DNA methylation levels. DMRs on top and bottom strands were called by metilene separately using the CpGs with coverage ≥ 10 for the corresponding strand. DMRs with ≥3 CpGs, a mean methylation difference of ≥0.2 and FDR < 0.05 were considered significant.

To identify regions with discordant strand-specific methylations, delta methylation values was calculated for each CpG for each sample as follows: delta methylation value = methylation ratio of top strand − methylation ratio of bottom strand. The Wilcoxon rank-sum test was performed on delta methylation values for each CpG between patients and controls within each cell type. Then, the whole genome was tiled into windows of 300 bp with step-size of 100 bp. Fisher's exact test was conducted to find windows enriched with CpGs showing significant *P*-values (< 0.05) in the Wilcoxon rank-sum test. Windows overlapping with the target region with a delta methylation value >0.1 or < −0.1 and with FDR < 0.05 were considered significant. Only 2 370 095 CpGs with coverage ≥ 10 on both strands were used in these analyses.

To investigate if the number of CpGs with the discordant methylation level between two strands were enriched in the patients, we assigned a binary methylation state (methylation ratio > 0.7 as methylated or < 0.3 as unmethylated) to each CpG on each strand for each sample. Then, the ratio of CpGs with discordant methylation states were calculated for each sample and the Student's t test showed no difference between patients and controls. Different cutoffs were also used for the binary methylation state (>0.8 or >0.9 as methylated, <0.2 or <0.1 as unmethylated), and no difference was observed between patients and controls. Consistently, patients were not enriched with the number of CpGs with a large delta methylation value (±0.1, ±0.2 or ±0.3).

Bisulfite pyrosequencing

For validation of DMRs, primers were designed to amplify bisulfite-converted DNA using PyroMark Assay Design SW 2.0 (Qiagen). In brief, 1 μg of genomic DNA was bisulfite-converted and purified using the EZ DNA Methylation Lightning Kit (Zymo Research). After purification of bisulfite-converted DNA, samples were amplified using the PyroMark PCR Kit (Qiagen). All samples were submitted to the Protein and Nucleic Acid Facility (Stanford University) for pyrosequencing sequencing.

RNA-Seq

Total RNA was extracted from fibroblasts, iPSCs and iNs using the RNeasy mini kit and RNeasy micro kit (Qiagen). Library construction was performed using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs) following the manufacturer's instructions. Briefly, 1 μg of total RNA was converted to mRNA using polyA purification. The mRNA was then fragmented and primed for cDNA synthesis. The cDNA was adenylated and adapters were ligated onto the cDNA. Samples were amplified using PCR; the amplified library was then purified using Agencount AMPure XP beads (Beckman Coulter). Final libraries were checked for quality on a Bioanalyzer DNA High Sensitivity chip (Agilent). Stand-specific RNA libraries of all the samples were sequenced on Illumina NextSeq 550 platform using 2 × 150 paired-end sequencing with 52–86 million reads generated for each sample. Cutadapt was used to trim Illumina TruSeq adapters and low-quality ends from the raw reads. Bowtie2 (Version 2.3.1) (32) was used to align the trimmed reads to the RefSeq transcriptome (GRCh38.p10), and RSEM (Version 1.2.30) (33) was used to quantify gene expression in a strand-specific manner by setting parameter '—forward-prob 0'. One sample (F03-3iN) was removed due to the extremely low mapping rate (7%), while the mapping rates for other samples ranged from 65% to 86%. DESeq2 (Version 1.12.4) (34) was used to perform differential expression analysis between cell types as well as between patients and controls within each cell type. Genes with FDR-adjusted *P*-value <0.05 were considered to be significant.

Co-expression network analysis

To identify modules of tissue-specific expressed genes, we performed weighted correlation network analysis (WGCNA, Version 1.46) (35) using all the samples by constructing signed network. Normalized count data obtained by the 'varianceStabilizingTransformation' function in DESeq2 were used.

Pathway enrichment analysis

To identify the biological processes that were significantly enriched with differentially expressed genes (DEGs; upregulated and downregulated genes considered separately) or enriched with genes associated with significant DMRs in each comparison, DAVID (Version 6.8) (36) was used to conduct pathway enrichment analysis using all the genes with an adjusted *P*-value reported by DESeq2 or all the genes with DMRs assigned as background. GO biological processes or KEGG pathways with FDR-corrected *P*-value <0.01 were considered significant.

qRT-PCR

Total RNA was isolated from iNs using the RNeasy Mini Kit (Qiagen). cDNA was prepared from 1 μg of total RNA by reverse transcription using the QuantiTect reverse transcript kit (Qigen). Gene expression was quantified using real-time quantitative PCR in combination with gene-specific primers and the Kapa SYBR FAST qPCR Kit (Kapa Biosystems). The reactions were performed with QuantStudio 6 Real-Time PCR system (Thermo Fisher Scientific). All samples were run in duplicates, and the 2-$^{\Delta\Delta CT}$ method was used to calculate relative expression levels normalized to UBC and HRPT.

Transcription factor binding site enrichment analysis

We first built an in-house database of TFs and their target genes. Briefly, TFs with known motifs were collected from existing databases and previous publications including Homer (37), Jaspar (38), ENCODE (39) and Jolma *et al.* (40). Due to the sample size restriction of iNs in relevant databases, target genes of these TFs were identified through correlation analysis between enhancer/promoter openness and gene expression in ENCODE blood cell lines as a compromise. Enrichment of TF binding motifs in the promoters and enhancers of the DEGs in iNs was performed using Homer.

Results
Subject description

Fibroblast cell cultures from three ADCA-DN patients with previously described ADCA-DN mutations (10,11) were obtained from Italy and Sweden. Point mutations for these patients were confirmed by Sanger sequencing of exon 21 of *DNMT1*. Sanger sequencing detected the mutation p.Ala554Val (RefSeq NM_001130823.1: c.1709G > A) in Italian proband 1, p.Gly589Ala (RefSeq NM_001130823.1 c.1814C > G) in Italian proband 2 and p.Val590Phe (RefSeq NM_001130823.1 c.1816C > A) in the Swedish proband (Table 1). As previously described, these three mutations of *DNMT1* are located in the replication foci targeting sequence motifs (RFTS) domain, which plays an inhibitory role in DNMT1-mediated DNA methylation (11,41) (Fig. 1A). The RFTS domain is located within the N-terminal region of the protein, which plays a role in recruitment and enzymatic activity of DNMT1 (41). The N-terminal region is composed of the RFTS domain, CXXC zinc finger domain and a pair of BAH (bromo-adjacent homology) domains (Fig. 1B). It is notable that residues A554 and V590 are involved in the formation of a hydrophobic cluster in the RFTS domain. It is possible that these mutations in this region cause altered stability and conformation of DNMT1.

Generation of induced glutamatergic neurons from iPS cells

In order to generate iNs derived from iPSCs, we used lentiviral delivery for vectors expressing rtTA, the Ngn2/puromycin resistance gene (which allowed us to select for cells expressing only Ngn2) and EGFP (which allowed us to identify cells that were transduced by the lentivirus) (Fig. 1C). All of the iNs were
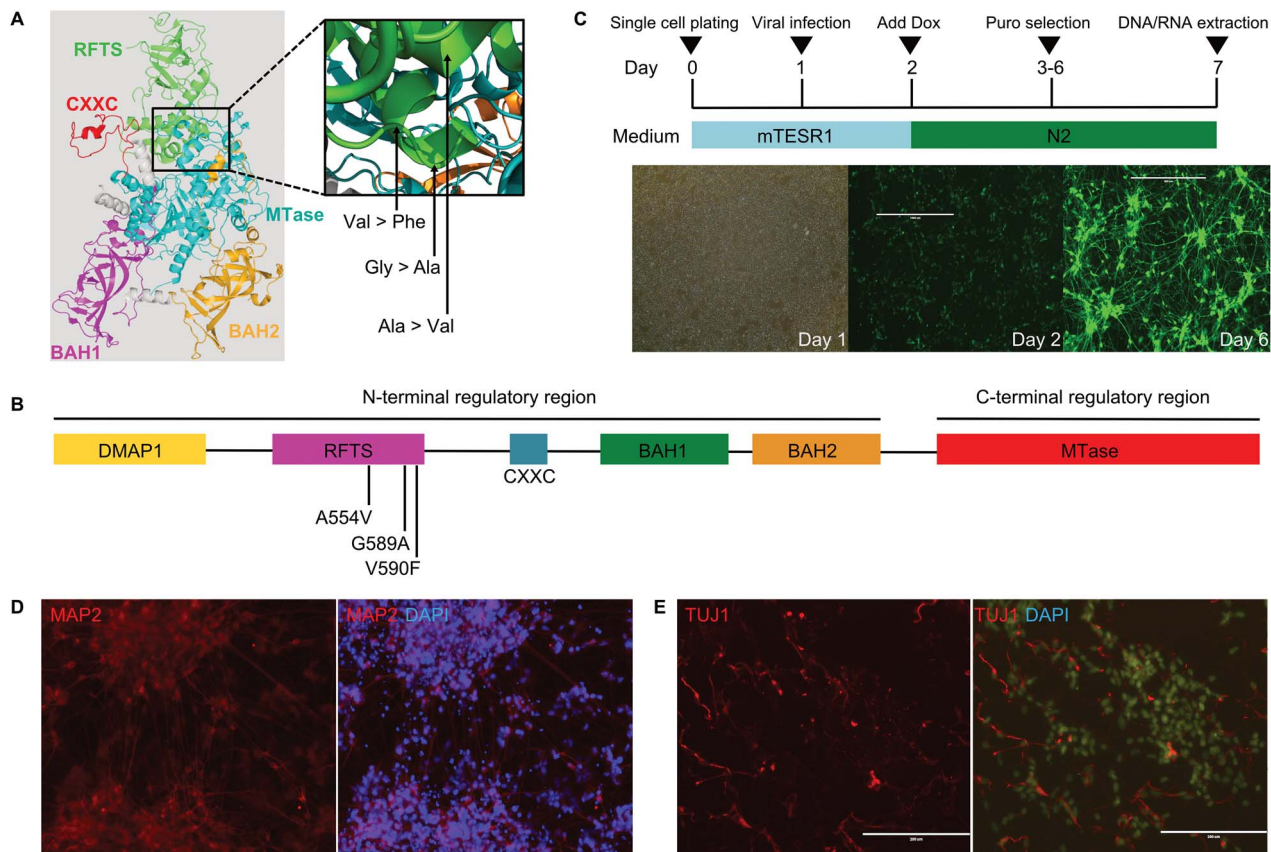
Figure 1. Protein structure of human DNMT1 (PDB accession number 4WXX, image generated with PyMol2). (**A**) Cartoon representation of human DNMT1 with domains RFTS, CXXC, BAH1, BAH2 and MTase. Sites of ADCA-DN mutations are (denoted by arrows) in the RFTS domain. Magnification structure was tilted in order to show the mutations. (**B**) Domain architecture of human DNMT1. Locations of the DNMT1 mutations are located in the RFTS domain. (**C**) Time-course of differentiation showing the generation of iNs including infection with lentivirus at day 1, dox treatment at day 2 and finally day 6 during puromycin selection. (**D**, **E**) Representative immunostaining images of iNs 7 days after infection stained with MAP2 and DAPI and TUJ1 and DAPI.

stained positive for neuronal cell markers MAP2 (Fig. 1D) and TUJ1 (Fig. 1E), indicating that all of the lentivirus treated iPSCs differentiated into iNs.

DNA methylation spectrum across various tissues

To examine the effects of *DNMT1* mutations on methylation profiles, we performed genome-wide capture-based bisulfite sequencing of genomic DNA from blood, fibroblasts, iPSCs and iNs, in both patients with ADCA-DN and controls. We observed distinct methylation profiles in different cell types or tissues (Fig. 2A). The global DNA methylation level increased from fibroblasts to iPSCs to blood to iNs (Fig. 2A and B). The difference of DNA methylation levels was significant for all comparisons between any two different tissues (Wilcoxon rank-sum test, *P*-value <0.05; Fig. 2B), while no significant difference was found between patients with ADCA-DN and controls within each tissue (P-value >0.05).

Differentially methylated genome regions in ADCA-DN

To investigate the effects of *DNMT1* mutations on DNA methylation patterns, we first identified differentially methylated regions (DMRs) between patients with ADCA-DN and controls for each tissue. DMRs were defined as genomic regions with ≥3 CpGs and a mean methylation difference ≥ 0.2 between the two compared

groups at FDR-corrected *P*-value <0.05. Using this threshold, we detected 140 DMRs in blood (Fig. 3A and B, Supplementary Material, Dataset S1), 349 in fibroblasts (Fig. 3C and D, Supplementary Material, Dataset S2), 385 in iPSCs (Fig. 3E and F, Supplementary Material, Dataset S3) and 344 in iNs (Fig. 3G and H, Supplementary Material, Dataset S4). To check accuracy of DMRs, we performed individual bisulfite pyrosequencing of the DMR in the promoter of *ZFP57* in iNs with a negative control CpG region and validated the findings (Supplementary Material, Table S2), which confirmed the reliability of our DNA methylation data.

We observed more hypermethylated DMRs in patients with ADCA-DN in blood and fibroblasts (Fig. 3A–D) but more hypomethylated DMRs in ADCA-DN patients in iPSCs and iNs (Fig. 3E–H). After assigning these DMRs to their closest transcriptional start site (TSS), we performed pathway enrichment analysis of the associated genes. Interestingly, the hypermethylated genes in patients in blood (Supplementary Material, Fig. S1A) and fibroblasts (Supplementary Material, Fig. S1C) were enriched in transcription related pathways, while it was the hypomethylated genes in iPSCs (Supplementary Material, Fig. S1D) and iNs (Supplementary Material, Fig. S1E) that were enriched in these pathways. Almost no enrichment of GO biological process pathways was seen within hypomethylated genes in blood and fibroblasts (Supplementary Material, Fig. S1B) as well as within hypermethylated genes in iPSCs and iNs. The hypomethylated genes were also enriched
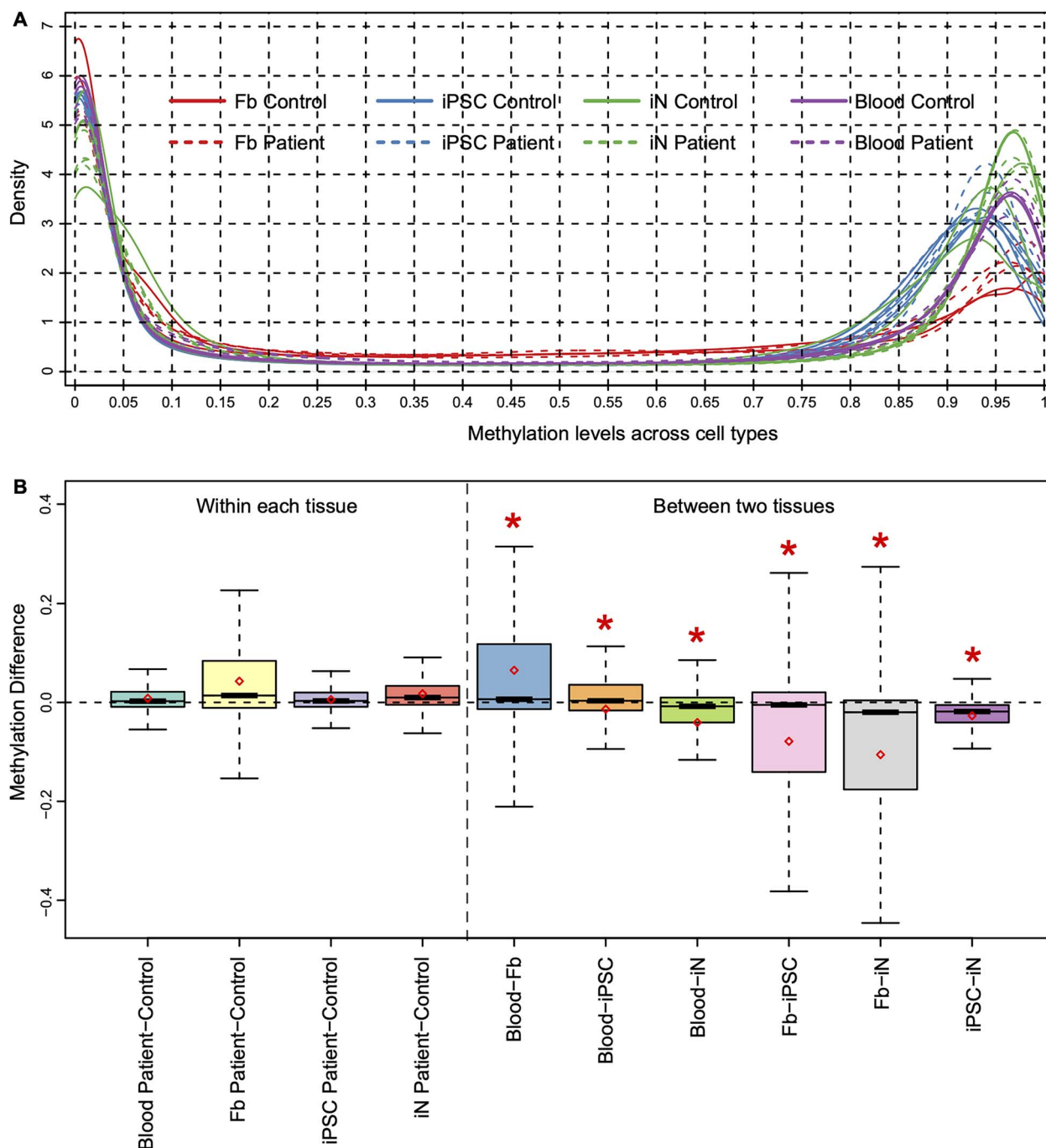
Figure 2. Methylation levels of CpGs across tissues (Fb: fibroblast; iPSC: induced pluripotent stem cell; iN: induced neuron; Blood: blood). (**A**) Density plot for methylation levels of CpGs for each sample across tissues. Patients with ADCA- DN and controls are denoted by solid and dashed lines, respectively. (**B**) Distribution of methylation level difference between patients with ADCA-DN and controls within each tissue and between any two different tissues. Diamonds and blocks in each box represent the mean and median values respectively. Comparisons on methylation levels between patients with ADCA-DN and controls within each tissue and between any two different tissues were conducted using the Wilcoxon rank-sum test. The red asterisks (∗) on the boxplots denote *P*-value <0.05 in the corresponding comparisons.

in pathways including nervous system development, synapse assembly and chemical synaptic transmission in both iPSCs (Supplementary Material, Fig. S1D) and iNs (Supplementary Material, Fig. S1E).

To find genes assigned with DMRs associated with ADCA-DN, we carried out an extensive literature research. This led us to 11 genes, based on the analysis at the iN stage in our study, that are known to be associated with various phenotypes of ADCA-DN (Fig. 4A): *KCNC3* (42), *KCNN3* (43), and *CNTN4*

(44) for cerebellar ataxia; *MYH14* (45), *NF2* (46), *GATA2* (47), and *FGF3* (48) for deafness; *KCNN3* (49), *CHL1* (50) and *GRM7* (51) for psychosis; *SORCS1* (52) and *TRIB2* (53) for narcolepsy. Four of these genes—*KCNC3*, *MYH14*, *NF2* and *SORCS1*—were also nominally differentially expressed between patients and controls (see below). *KCNC3*, *MYH14* and *SORCS1* were hypomethylated and with upregulated expression in patients, while *NF2* was hypermethylated and with downregulated expression in patients.
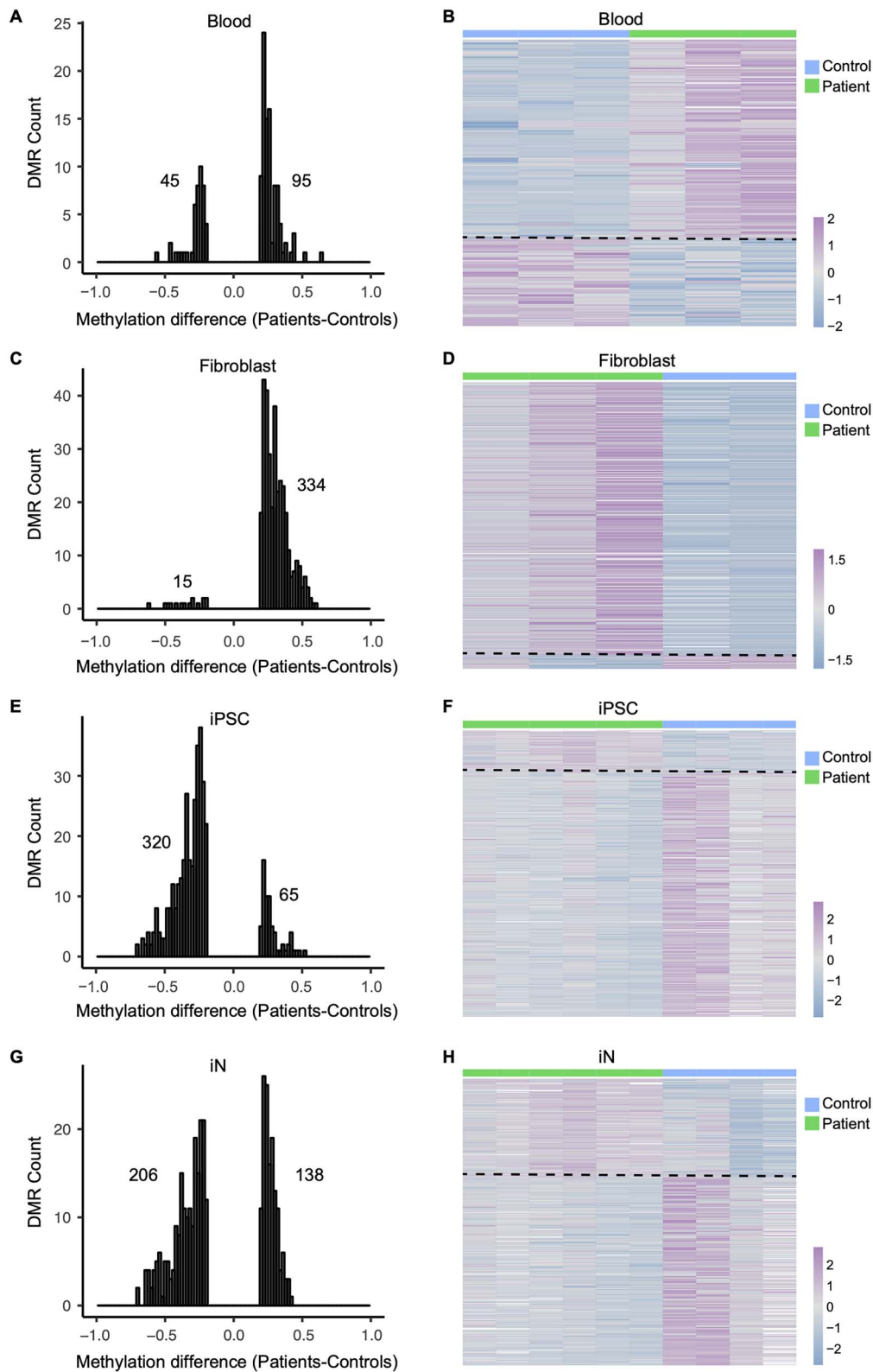
Figure 3. DMRs between patients with ADCA-DN and controls for each tissue. Distribution of DMRs in blood, fibroblasts, iPSCs and iNs are shown in (**A**), (**C**), (**E**) and (**G**), respectively. The scaled methylation levels of CpGs in the DMRs of blood, fibroblasts, iPSCs and iNs are shown in (**B**), (**D**), (**F**) and (**H**), respectively. In (A), (C), (E) and (G), x axis is the methylation difference between patients and controls with >0 meaning hypermethylation and < 0 meaning hypomethylation in patients, while the y axis is the number of DMRs at a specific interval of methylation level difference. The total number of DMRs is shown separately for hypermethylation and hypomethylation. In (B), (D), (F) and (H), the dashed black line separates hypermethylated and hypomethylated CpGs in patients. DMR calling cutoff: methylation difference > = 0.2, q value <0.05, number of CpG ≥ 3.
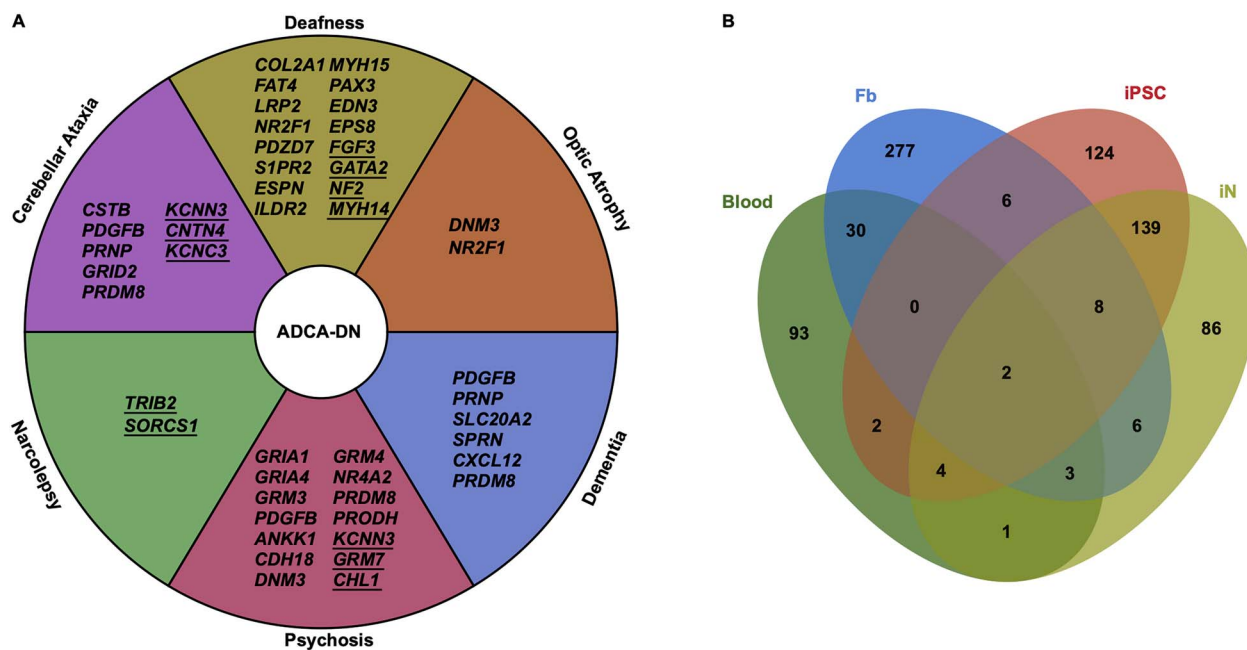
Figure 4. (**A**) Genes associated with differentially methylated regions (DMRs) and differentially expressed genes associated with phenotypes of ADCA-DN. The clinical phenotypes of ADCA-DN are classified into six categories. Shown is the genes associated with DMRs (underlined italic) and the differentially expressed genes (italic) between patients and controls in iNs, which have been associated with the corresponding phenotypes of ADCA-DN. (**B**) Overlap of genes associated with DMRs across tissues. DMRs called between patients with ADCA-DN and control within each tissue were assigned to their closest transcriptional start site (TSS), and the related genes were referred to as 'genes associated with DMRs'. Shown is the number of the overlapping genes associated with DMRs across tissues.

Shared genes associated with DMRs across cell types

Next, we investigated the overlap of genes with DMRs between patients and controls across tissues. The number of genes associated with DMRs overlapping between any two tissues was rather small, except between iPSCs and iNs (Fig. 4B). Our results suggested that the genes associated with DMRs between patients and controls can be distinct among tissues, which mirrored the lack of overlapping DEGs across various cell types (see below). iPSCs and iNs shared more than half of their genes associated with DMRs (Fig. 4B), indicating that a considerable proportion of DNA methylation changes seen in iNs might have occurred at the iPSC stage.

Strand-specific DNA methylation analysis

We next searched for strand-specific DNA methylation changes between patients with ADCA-DN and controls. In iNs, we only found one gene, *PCDHB12*, with a significant DNA methylation level difference between plus and minus strand in patients (Fisher's exact test *P*-value <2.38E−07). There were three genes, *PCBD2*, *CMC1* and *PCDHGB6*, with significantly smaller methylation level differences (Fisher's exact test *P*-value <8.68E−07) between strands in patients with ADCA-DN. In iPSCs, only three genes, *RIN1*, *C16orf95* and *MTRNR2L2*, exhibited significantly smaller strand-biased methylations in patients with ADCA-DN (Fisher's exact test *P*-value <2.35E−06) while no genes had larger strand-biased methylation differentials in patients with ADCA-DN. We did not observe evidence of a generally different discordance rate of the DNA methylation level of the two strands between patients and controls. Taken together, the methylation disturbance appeared to occur to both strands instead of being constrained to one strand.

Global gene expression changes across the cell types

We next characterized the expression profiles of fibroblasts, iPSCs and iNs in both patients with ADCA-DN and controls. Hierarchical clustering and principal component analysis (PCA) of RNA-Seq data demonstrated distinct gene expression patterns across cell types (Supplementary Material, Fig. S2). To identify the cell type–specific gene expression changes, we performed weighted gene co-expression network analysis (WGCNA), which reported six large gene co-expression modules (Fig. 6A). These modules ranged in size from 1233 to 3681 genes. Genes in module M1 were exclusively highly expressed in fibroblasts and were mainly involved in pathways such as extracellular matrix organization and the collagen metabolic process (Supplementary Material, Fig. S3A). Genes with exclusively high expression in iNs were observed in module M2 and M3. As expected, these genes were enriched in neuron- or synapse-related pathways such as synaptic transmission and nervous system development (Fig. 6B and C). Highly expressed genes in module M5 were specific to iPSCs, and these genes were enriched in pathways including RNA processing and DNA replication (Supplementary Material, Fig. S3B). Genes with similarly high expression levels in both fibroblasts and iPSCs or in both fibroblasts and iNs were seen in modules M4 (Supplementary Material, Fig. S3C) and M6 (Supplementary Material, Fig. S3D), respectively.

Differential gene expression in patients with ADCA-DN

To detect genes associated with ADCA-DN, we conducted differential expression analysis between patients with ADCA-DN and controls within each cell type. We detected 81 differentially expressed genes (DEGs) in fibroblasts (40 upregulated and 41

down-regulated in patients) (Fig. 5A–C, Supplementary Material, Dataset S5), 97 DEGs in iPSCs (53 upregulated and 44 downregulated in patients) (Fig. 5D–F, Supplementary Material, Dataset S6) and 394 DEGs in iNs (190 upregulated and 204 downregulated in patients) (Fig. 5G–I, Supplementary Material, Dataset S7). The fold changes of gene expression for a subset of the DEGs in iNs were confirmed by real-time quantitative PCR (qRT-PCR), in biological duplicates in all iN cell lines (Supplementary Material, Table S3). Pathway enrichment analysis of the DEGs identified only one KEGG pathway (p53 signaling pathway) within upregulated DEGs in patients in fibroblasts (Supplementary Material, Fig. S4A) and two in iPSCs (metabolic pathways and pathways in cancer, respectively) (Supplementary Material, Fig. S4B), while no pathway was enriched in the downregulated DEGs in patients with ADCA-DN in these two cell types. In iNs, KEGG pathways such as neuroactive ligand–receptor interaction were enriched in upregulated DEGs in patients (Supplementary Material, Fig. S4C) while synapse-related pathways including glutamatergic synapse and various signaling pathways including the ErbB signaling pathway were seen in downregulated DEGs (Supplementary Material, Fig. S4D). The larger numbers of DEGs and pathways identified in iNs demonstrate the necessity of studying the disease in the relevant tissues. By applying more stringent cutoff of FDR-corrected *P*-value <4.36E−03, expressions of 57 DEGs in iNs show complete separation between patients and controls (all patients showed higher or lower expression than any controls). However, no pathways were overrepresented in either upregulated or downregulated DEGs.

Overlapping DEGs across cell types

We next asked how much overlap there was among DEGs between the various cell types. Only 15 DEGs such as *ZFP57* were shared among at least two cell types, while the other DEGs were only observed in one cell type (Fig. 6D and E). Most of these shared DEGs (11 out of 15) displayed consistent expression changes across cell types, while four of these genes (*FAM167A, LRP2, FAT4* and *LIMCH1*) showed inconsistent expression changes (Fig. 6E). Previous studies have demonstrated that mutations and deletions in gene *GRID2* lead to cerebellar ataxia (54). Mutations in *LRP2* have been reported to cause Donnai–Barrow syndrome, which is associated with sensorineural hearing loss (55), while mutations in *FAT4* cause Van Maldergem syndrome (56), whose phenotypes include atresia of the external auditory meatus resulting in hearing loss (57). Furthermore, the majority (380 out of 394) of the DEGs in iNs were exclusively differentially expressed only in iNs (Fig. 6D), indicating that the genes associated with ADCA-DN may not exhibit expression disturbance until the mutation background is placed into the relevant cell type. This underlines the importance of studying the relevant tissues associated with the phenotypes of ADCA-DN.

Differentially expressed genes associated with phenotypes of ADCA-DN

Given that ADCA-DN is a disorder that mainly affects the central nervous system, we carried out an extensive review of the literature and were able to identify 30 genes out of the 394 DEGs in iNs, which are associated with subphenotypes present in ADCA-DN. We classified these DEGs into different phenotypic categories, including cerebellar ataxia, deafness, narcolepsy, psychosis, dementia and optic atrophy (Fig. 4A). Five were genes associated with cerebellar ataxia—*PRNP* (58), *PRDM8* (59), *GRID2* (54), *CSTB* (60) and *PDGFB* (61)—and 12 with deafness—*ESPN* (62),

PAX3 (63), *PDZD7* (64), *ESP8* (65), *S1PR2* (66), *LRP2* (55), *FAT4* (57), *MYH15* (45), *ILDR2* (67), *COL2A1* (68), *EDN3* (68) and *NR2F1* (69). *DNM3* (70) and *NR2F1* (71) were associated with optic atrophy. Eleven psychosis genes were identified—*PDGFB* (61), *NR4A2* (72), *PRMD8* (59), *GRIA1* (73), *GRM4* (51), *GRM3* (74), *GRIA4* (75), *CDH18* (76), *ANKK1* (77), *DNM3* (78) and *PRODH* (79). We also found six genes associated with this dementia—*PDGFB* (61), *SPRN* (80), *PRMD8* (59), *PRNP* (58), *CXCL12* (81) and *SLC20A2* (82). None of the DEGs showed genetic association with narcolepsy and cataplexy. We noticed that four genes were associated with two or three phenotypes: *PDGFB* and *PRDM8* with cerebellar ataxia, dementia and psychosis; *PRNP* with cerebellar ataxia and dementia; and *NR2F1* with deafness and optic atrophy.

Differentially expressed transcription factor genes

To uncover potential key regulators of these expression networks, we performed enrichment analysis of transcription factor (TF) binding motifs in the promoters and enhancers of the DEGs in iNs. We observed that the binding motifs of 91 TFs were significantly enriched in the promoters or enhancers of DEGs in iNs (raw *P*-value 0.01). Among them, four TFs—*ZFP57, NANOG, ETV1* and *KLF10*—were also DEGs themselves in iNs. The promoters of the DEGs in iNs were enriched with the binding motifs of *ZFP57* and *NANOG*, whereas the enhancers were enriched with the binding motifs of *ETV1* and *KLF10*.

Genes showing both differential methylation and differential expression in iNs

To examine whether DMRs are enriched for DEGs, we assigned DMRs to their closest TSS. Among the 221 expressed genes with DMRs assigned in iNs, 10 genes were DEGs (*CHL1, ZNF99, ZFP57, UNC13C, TCERG1L, HOOK2, PCDHB15, PCDHA3, DIO3* and *PDE10A*) and 55 genes were nominally differentially expressed (raw *P*-value <0.05, Supplementary Material, Dataset S8), which lead to significant enrichment (Fisher's exact test *P*-value = 8.82E−03) compared to genes expressed at background levels. The majority of these genes (47 of 55) exhibited consistent changes between DNA methylation level and gene expression, meaning either hypermethylation and downregulated expression or hypomethylation and upregulated expression in patients with ADCA-DN. We saw the same enrichment in iPSCs (Fisher's exact test *P*-value = 1.57E−04) but not in fibroblasts (Fisher's exact test *P*-value = 8.36E−02).

Of the 55 nominally differentially expressed genes with changes of DNA methylation levels in iNs, 9 are TFs. The most significantly differentially expressed TF is *ZFP57* (raw *P*-value = 6.01E−06, FDR-corrected *P*-value = 2.39E−03), which was hypomethylated in promoters and exhibited upregulated expression in patients. *ZFP57* encodes a transcription factor required to maintain maternal and paternal gene imprinting and acts by controlling DNA methylation during early development (83). Interestingly, 10 of the 221 expressed genes with DMRs assigned were imprinted genes, which was a significant enrichment compared to genome wide (Fisher's exact test *P*-value = 1.27E−04). The same enrichment was also seen in both iPSCs (Fisher's exact test *P*-value = 7.46E−05) and fibroblasts (Fisher's exact test *P*-value = 7.02E−03). It is noteworthy that *ZFP57* was differentially expressed and differentially methylated in both iPSCs and iNs, but was not differentially methylated in either fibroblasts or blood and not expressed in fibroblasts. However, the binding motif of *ZFP57* was only enriched in DEGs of iNs.
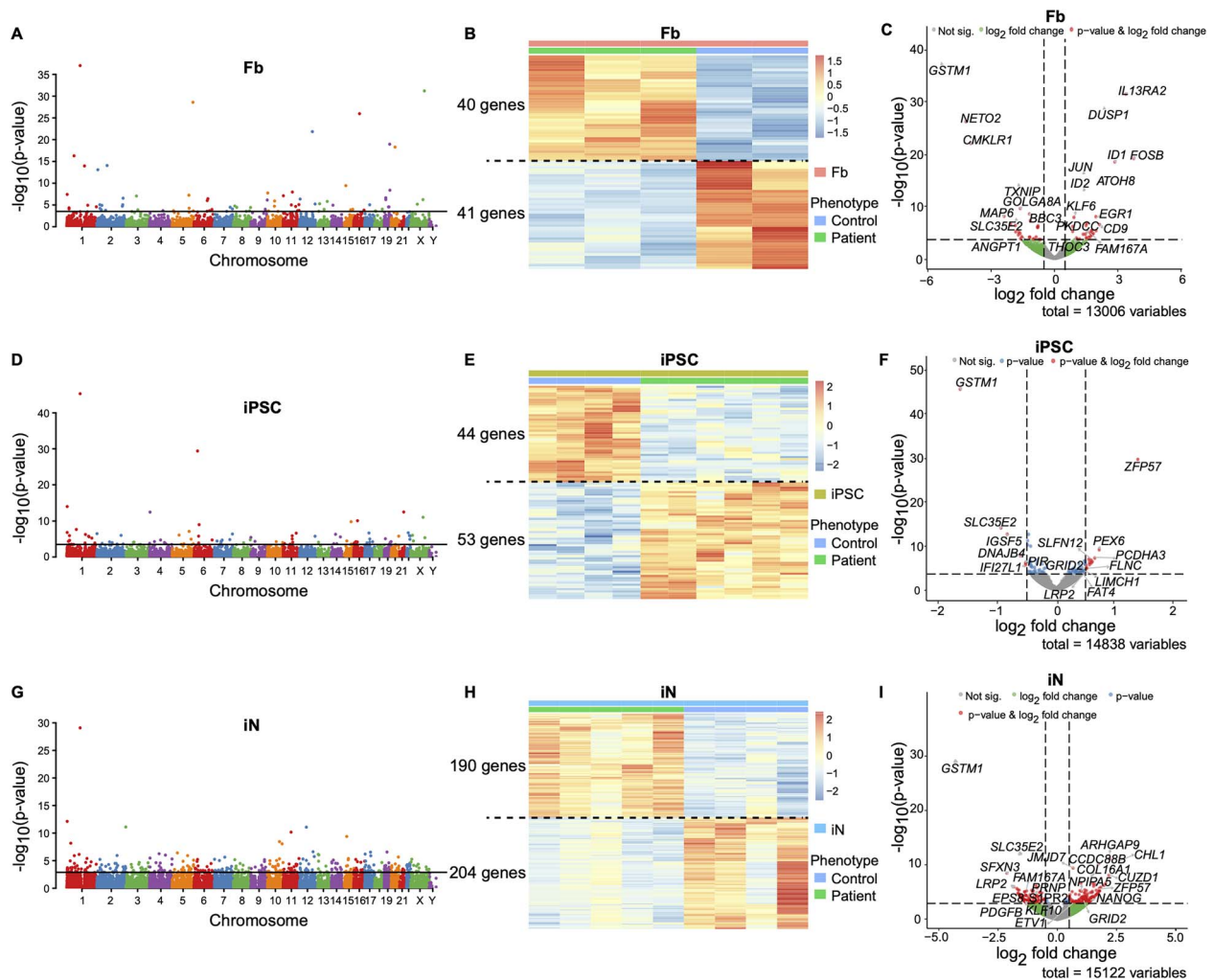
Figure 5. Differential expression analyses between patients with ADCA-DN and controls in three cell types. Genome-wide distribution of DEGs between patients and controls are shown in Manhattan plots for (**A**) fibroblasts, (**D**) iPSCs and (**G**) iNs. The x-axis denotes the −log10(P-values) of genes across the genome (y-axis) in the Manhattan plots. Genome-wide significance is based on FDR < 0.05 indicated by the horizontal black lines. Heatmaps show the scaled normalized expression for the DEGs in fibroblasts (**B**), iPSCs (**E**) and iNs (**H**), where the color scale goes from low expression to high expression. Volcano plots present the up- and downregulation of DEGs between patients and controls in fibroblasts (**C**), iPSCs (**F**) and iNs (**I**). Two vertical dashed lines in each volcano plot indicates a range of log₂ fold change values from −0.5 to 0.5, and the horizontal dashed line denotes the adjusted P-value cutoff of 0.05. A portion of important DEGs are labeled with gene names.

Discussion

Our present study provides the first analysis of global and cell type–specific changes in gene expression patterns and DNA methylation profiles induced by *DNMT1* mutations that are causal to ADCA-DN, in three different cell types, fibroblasts, iPSCs and iNs. We observed tissue-specific changes in gene expression and DNA methylation between patients with ADCA-DN and controls, underlining the importance of interrogating DNA methylation in relevant tissues and cell types. Furthermore, we were able to identify a group of genes that may play a key role in the molecular etiology of ADCA-DN (Fig. 4A).

Although ACDA-DN is caused by mutations in *DNMT1*, we did not observe evidence of differential expression of *DNMT1* between patients and controls in any cell type. Also, no evidence of allele-specific expression of *DNMT1* was seen in any cell type. The DNA methylation level in *DNMT1*'s promoter and its gene body appeared to remain unchanged in patients in all tissues of this study. Moreover, we did not see evidence of changes in

gene expression and DNA methylation of other DNA methyltransferases including *DNMT3a*, *DNMT3b*, *DNMT3L* and *DNMT2*. This indicates that it is not the expression level but the functionally effective mutations in the gene that lead to downstream consequences.

However, for another DNA methylation maintenance factor, *ZFP57*, we did observe changes of both DNA methylation and gene expression in patients in both iPSCs and iNs. *ZFP57*, a transcription factor, affects DNA methylation of imprinting control regions (84), maintains parent-of-origin-specific expression of corresponding imprinted genes and differentially affects non-imprinted targets in mouse embryonic stem cells (85). Furthermore, mutations of *ZFP57* result in hypomethylation at many imprinted regions in the human genome (83). Intriguingly, we also found that imprinted genes were enriched in the genes with DMRs assigned in all cell types and the binding motif of *ZFP57* was overrepresented in the promoters of the DEGs in iNs. Our results indicate that *ZFP57* might play an important role in reshaping the expression and DNA methylation spectrum of relevant tissues such as iNs and might
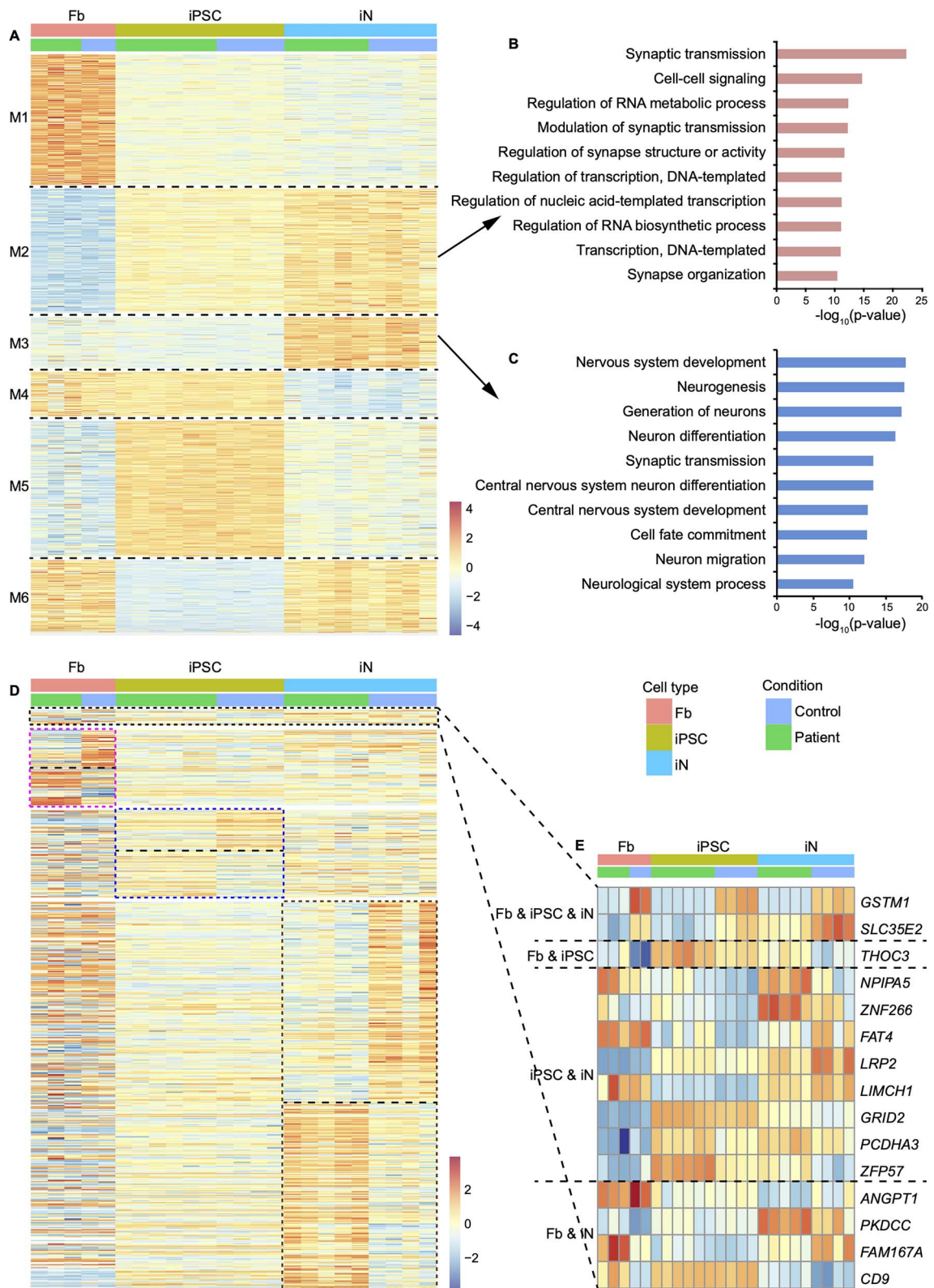
Figure 6. Tissue-specific expression and differential expression between patients with ADCA-DN and controls. (**A**) Distinct expression patterns across tissues. Heatmap shows normalized expression of genes in modules clustered by weighted correlation network analysis (WGCNA). Different modules are separated by black dashed lines. The color scale goes from low to high expression. (**B**, **C**) Pathways enriched in genes with exclusive high expression in iNs. The difference between the two gene modules is medium expression in iPSCs versus low expression in iPSCs. (**D**) DEGs in each tissue. Heatmap shows normalized expression of DEGs between patients and controls exclusively in fibroblasts, iPSCs and iNs as well as the shared DEGs among tissues. The color scale goes from low to high expression. The black dashed line separates the upregulated and downregulated DEGs within each box. (E) Share DEGs among tissues. Heatmap shows the zoomed-in view of the top box with dashed lines in (D).

be a key regulator toward the development of the symptoms of ADCA-DN. Of note, *ZFP57* was hypomethylated in promoters and its expression was upregulated in patients with ADCA-DN. Whether *ZFP57*'s overexpression is compensatory for correcting the disturbed DNA methylation that occurs with *DNMT1* mutation during early development remains to be investigated further.

We identified four differentially expressed TFs—*ZFP57, NANOG, ETV1* and *KLF10*—in iNs, whose target genes were also enriched in the DEGs in iNs. Another TF *NR2F2*, which has a binding motif of *ZFP57*, showed borderline significance of differential expression (raw *P*-value = 1.47E−03, FDR-corrected *P*-value = 5.24E−02) in iNs. *NR2F2* was differentially hypermethylated in one of its enhancers (FDR-corrected *P*-value = 5.10E−05) and showed downregulated expression in patients with ADCA-DN in iNs. Intriguingly, of the aforementioned genes associated with ADCA-DN phenotypes, five—*NR2F1, EPS8, S1PR2, PRNP* and *PDGFB*—were target genes of *NR2F2*. Among them, deletion of *NR2F1* has been reported to be associated with deafness (69) and optic atrophy (71). A nonsense mutation of *EPS8* has been associated with autosomal recessive profound deafness (65), and Eps8 knockout mice are profoundly deaf (86). Rare missense variants within *S1PR2* have been shown to cause autosomal-recessive hearing impairment (66) and knockout mice for *S1PR2* have been previously reported to have hearing loss (87). Nonsense and missense mutations of *PRNP* have been associated with neurodegenerative diseases, such as Creutzfeldt–Jakob disease, Gerstmann–Sträussler–Scheinker disease and fatal familial insomnia, which may all have clinical features of ataxia and dementia, although in these cases, spreading of misfolded *PRNP* protein is the known mechanism, an effect unlikely here (58). Loss-of-function mutations in the *PDGFB* gene have been found to be causative of idiopathic basal ganglia calcification (61), whose clinical features include ataxia, psychosis and dementia (88). Interestingly, these five genes were all downregulated in patients in iNs only, indicating the importance, again, of carrying out analyses in physiologically relevant cell types. The roles of these genes together with *NR2F2* in the development of ADCA-DN, especially for features such as deafness and cerebella ataxia, warrant further investigation.

While our study unveiled genes associated with most of the phenotypes of ADCA-DN, we did not observe differential expressions of any genes known to be associated with narcolepsy in the GWAS Catalog (89). Only one DEG—*NCKAP5*—in iNs has been associated with essential hypersomnia (90). However, we did find two genes with DMRs assigned to them—*SORCS1* and *TRIB2*—that are also associated with narcolepsy. Only *SORCS1* was nominally differentially expressed (upregulated in patients with its assigned DMRs hypomethylated). The lack of concordance is not surprising as the most frequent cause of narcolepsy is a loss of hypocretin cells secondary to an autoimmune process, thus leading to tight association of autoimmune loci with the pathology (91). In ADCA-DN, the likely origin of the narcolepsy phenotype may be a preferential vulnerability of hypocretin cells to DNMT1-mediated neurodegeneration. A recent study also linked ADCA-DN abnormalities to mitochondrial deficits (92), an interesting hypothesis as hypocretin neurons, with their long-range projections are likely very energetically demanding.

As *DNMT1* is responsible for maintaining DNA methylation patterns following DNA replication, we examined if mutations of *DNMT1* can cause differential methylation between the two DNA strands in patients with ADCA-DN. However, the vast majority (99.91%) of the CpGs exhibited either high methylation (>0.7) or low methylation (<0.3) on both strands in both patients and controls in all the four tissues studied (Supplementary Material, Table S4). Between the patients and the controls, there is no evidence of different rates of DNA methylation discordance between the two strands (Supplementary Material, Table S5). Our results indicate that the *DNMT1* mutations analyzed here do not impact the methylation level in a strand-specific way.

We observed significant enrichment of nominal DEGs in the genes with DMRs assigned to them in both iPSCs and iNs. Most of these genes (39 of 46 in iPSCs and 47 of 55 in iNs) were either hypermethylated at their assigned DMRs and downregulated in expression or hypomethylated and upregulated in patients with ADCA-DN. The negative correlation between methylation and gene expression changes indicates that the differential expression of these genes may at least to some extent be premised on the differential DNA methylation in patients, which was caused by the *DNMT1* mutations.

There are several limitations in the present study that merit consideration. First, this study employed the fibroblasts–iPSCs–iNs succession of cell types in culture. We employed this model system in order to enable us to address for the first time, in a direct and efficient manner, the general question whether the ADCA-DN mutations induce global changes in DNA methylation patterns and gene expression levels and whether such changes would be cell type specific. We chose to analyze iNs that were still at an early stage of maturation, as further maturation would have required to grow the neuronal cells on mouse glia, which could have introduced considerable confounding effects into the functional genomics analyses. Now that we have answered the question in the positive regarding whether there are neuronal cell–specific effects of the ADCA-DN mutations, follow-up studies would be expected to address the limitations of such early iNs. Namely, iNs could now be grown to maturation on mouse glia, a protocol separating the neuronal and the glial cells will have to be established to allow for functional genomics analyses in such a system. Single-cell assays should be considered in order to resolve sub-populations of neuronal cells, as even early iNs, and certainly iNs at a more mature stage, may be not entirely homogeneous regarding their specific neuronal identity (93). In such a system, additional later time points can also be considered for analysis, as well as the separate analysis of a variety of different neuronal cell types that have been purposely induced. For our present study, since the ADCA-DN mutations are autosomal dominant, we considered as justifiable the use of mutation-carrying iPSCs from ADCA-DN patients together with control iPSCs derived from unrelated donors. Once the analyses become more fine grained, one may reconsider whether iPSC lines with engineered ADCA-DN mutations, and matching isogenic control lines, would be advantageous to remove effects of genomic background (94,95). Knock-in or transgenic mouse models would be expected to be powerful complementary systems during these next stages of investigation.

The assay employed in the present study did not distinguish DNA hydroxymethylation from DNA methylation, and methylation levels at non-CpG sites were not examined. The DNA hydroxymethylation level in iPSCs and iNs and DNA methylation patterns at non-CpG sites in neurons warrant further investigation to understand the complexity of epigenetic regulation in neuronal cells. Transcription factor-binding enrichment analysis was done here in a purely computational manner. Now that our analyses have indicated that there may be relevant changes at the current level of analysis, this can also be expanded by experimental means in the next round of investigation. Specifically, in order to infer TF-binding by TF foot printing, ATAC-seq data, e.g. in

the more mature iNs described above, would provide further mechanistic insight based on experimental data (96).

Our present study, by showing that indeed the ADCA-DN mutations induce global, and cell type–specific, changes to patterns of DNA methylation and gene expression, have opened the door to such further study, and we expect that this will eventually not only lead to an understanding to the molecular etiology of ADCA-DN but also to useful insights into its component symptoms and overall into the molecular functions of the important gene *DNMT1*.

Supplementary Material

Supplementary Material is available at *HMG* online.

Conflict of Interest statement. None declared.

Funding

Web Resources

The URLs for data presented herein are as follows:

bamUtil, https://github.com/statgen/bamUtil

Cutadapt, https://journal.embnet.org/index.php/embnetjournal/article/view/200

DAVID, https://david.ncifcrf.gov/

KEGG, http://www.genome.jp/kegg/

RefSeq, http://www.ncbi.nlm.nih.gov/RefSeq

WGCNA, https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/

Data Availability

Raw sequencing data for this project have been deposited into GEO under accession number GSE126890.

Author Contributions

K.N.D. produced experimental data and contributed to writing the paper; X.Z. carried out data analysis and wrote the manuscript; P.Q. carried out data analysis and edited the manuscript; S.M. analyzed data; L.L. and M.P. produced experimental data; N.D. identified the Sweden patient; G.P. and F.P. identified the two Italian patients and deeply characterized them; R.O., W.H.W., J.H. and E.M. contributed to study design, contributed to directing data production and contributed to data interpretation; A.E.U. and X.Z. conceived of and designed the study, directed data production and interpretation.

References

1. Feng, J. and Fan, G. (2009) The role of DNA methylation in the central nervous system and neuropsychiatric disorders. *Int. Rev. Neurobiol.*, **89**, 67–84.

2. Golshani, P., Hutnick, L., Schweizer, F. and Fan, G. (2005) Conditional Dnmt1 deletion in dorsal forebrain disrupts development of somatosensory barrel cortex and thalamocortical long-term potentiation. *Thalamus Relat. Syst.*, **3**, 227–233.

3. Hutnick, L.K., Golshani, P., Namihira, M., Xue, Z., Matynia, A., Yang, X.W., Silva, A.J., Schweizer, F.E. and Fan, G. (2009) DNA hypomethylation restricted to the murine forebrain induces cortical degeneration and impairs postnatal neuronal maturation. *Hum. Mol. Genet.*, **18**, 2875–2888.

4. Feng, J., Zhou, Y., Campbell, S.L., Le, T., Li, E., Sweatt, J.D., Silva, A.J. and Fan, G. (2010) Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat. Neurosci.*, **13**, 423–430.

5. Li, E., Bestor, T.H. and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915–926.

6. Wang, W., Zhao, X., Shao, Y., Duan, X., Wang, Y., Li, J., Li, D., Li, X. and Wong, J. (2021) Mutation-induced DNMT1 cleavage drives neurodegenerative disease. *Sci. Adv.*, **7**, eabe8511.

7. Inano, K., Suetake, I., Ueda, T., Miyake, Y., Nakamura, M., Okada, M. and Tajima, S. (2000) Maintenance-type DNA methyltransferase is highly expressed in post-mitotic neurons and localized in the cytoplasmic compartment. *J. Biochem.*, **128**, 315–321.

8. Bachmann, S., Linde, J., Bell, M., Spehr, M., Zempel, H. and Zimmer-Bensch, G. (2021) DNA Methyltransferase 1 (DNMT1) shapes neuronal activity of human iPSC-derived glutamatergic cortical neurons. *Int. J. Mol. Sci.*, **22**, 2034.

9. Zhao, C., Deng, W. and Gage, F.H. (2008) Mechanisms and functional implications of adult neurogenesis. *Cell*, **132**, 645–660.

10. Melberg, A., Hetta, J., Dahl, N., Nennesmo, I., Bengtsson, M., Wibom, R., Grant, C., Gustavson, K.H. and Lundberg, P.O. (1995) Autosomal dominant cerebellar ataxia deafness and narcolepsy. *J. Neurol. Sci.*, **134**, 119–129.

11. Winkelmann, J., Lin, L., Schormair, B., Kornum, B.R., Faraco, J., Plazzi, G., Melberg, A., Cornelio, F., Urban, A.E., Pizza, F. *et al.* (2012) Mutations in DNMT1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Hum. Mol. Genet.*, **21**, 2205–2210.

12. Klein, C.J., Botuyan, M.V., Wu, Y., Ward, C.J., Nicholson, G.A., Hammans, S., Hojo, K., Yamanishi, H., Karpf, A.R., Wallace, D.C. *et al.* (2011) Mutations in *DNMT1* cause hereditary sensory neuropathy with dementia and hearing loss. *Nat. Genet.*, **43**, 595–600.

13. Sun, Z., Wu, Y., Ordog, T., Baheti, S., Nie, J., Duan, X., Hojo, K., Kocher, J.P., Dyck, P.J. and Klein, C.J. (2014) Aberrant signature methylome by DNMT1 hot spot mutation in hereditary sensory and autonomic neuropathy 1E. *Epigenetics*, **9**, 1184–1193.

14. Kernohan, K.D., Cigana Schenkel, L., Huang, L., Smith, A., Pare, G., Ainsworth, P., Boycott, K.M., Warman-Chardon, J., Sadikovic, B. and Consortium, C.R.C (2016) Identification of a methylation profile for *DNMT1*-associated autosomal dominant cerebellar ataxia, deafness, and narcolepsy. *Clin. Epigenetics*, **8**, 91.

15. Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C. *et al.* (2015) Targeted disruption of *DNMT1*, *DNMT3A* and *DNMT3B* in human embryonic stem cells. *Nat. Genet.*, **47**, 469–478.

16. Auburger, G., Klinkenberg, M., Drost, J., Marcus, K., Morales-Gordo, B., Kunz, W.S., Brandt, U., Broccoli, V., Reichmann, H., Gispert, S. and Jendrach, M. (2012) Primary skin fibroblasts as a model of Parkinson's disease. *Mol. Neurobiol.*, **46**, 20–27.

17. Kálmán, S., Garbett, K.A., Janka, Z. and Mirnics, K. (2016) Human dermal fibroblasts in psychiatry research. *Neuroscience*, **320**, 105–121.

18. TCW, J. (2019) Human iPSC application in Alzheimer's disease and tau-related neurodegenerative diseases. *Neurosci. Lett.*, **699**, 31–40.

19. Brennand, K.J. and Gage, F.H. (2011) Concise review: the promise of human induced pluripotent stem cell-based studies of schizophrenia. *Stem Cells*, **29**, 1915–1922.

20. Nagy, A. and Turksen, K. (eds) (2022) *Induced Pluripotent Stem (iPS) Cells: Methods and Protocols*. Springer Nature, New York, NY, USA.

21. Hoffmann, A., Ziller, M. and Spengler, D. (2020) Focus on causality in ESC/iPSC-based modeling of psychiatric disorders. *Cell*, **9**, 366.

22. Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., Marro, S., Patzke, C., Acuna, C., Covy, J. *et al.* (2013) Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron*, **78**, 785–798.

23. Wendt, J., Rosenbaum, H., Richmond, T.A., Jeddeloh, J.A. and Burgess, D.L. (2018) Targeted bisulfite sequencing using the Seq-Cap epi enrichment system. *Methods Mol. Biol.*, **1708**, 383–405.

24. Sun, Z., Cunningham, J., Slager, S. and Kocher, J.P. (2015) Base resolution methylome profiling: considerations in platform selection, data preprocessing and analysis. *Epigenomics*, **7**, 813–828.

25. Walker, D.L., Bhagwate, A.V., Baheti, S., Smalley, R.L., Hilker, C.A., Sun, Z. and Cunningham, J.M. (2015) DNA methylation profiling: comparison of genome-wide sequencing methods and the Infinium human methylation 450 bead Chip. *Epigenomics*, **7**, 1287–1302.

26. Krueger, F. and Andrews, S.R. (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*, **27**, 1571–1572.

27. García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Götz, S., Tarazona, S., Dopazo, J., Meyer, T.F. and Conesa, A. (2012) Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics*, **28**, 2678–2679.

28. Jühling, F., Kretzmer, H., Bernhart, S.H., Otto, C., Stadler, P.F. and Hoffmann, S. (2016) Metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res.*, **26**, 256–262.

29. Hermann, A., Goyal, R. and Jeltsch, A. (2004) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J. Biol. Chem.*, **279**, 48350–48359.

30. Valinluck, V. and Sowers, L.C. (2007) Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res.*, **67**, 946–950.

31. Probst, A.V., Dunleavy, E. and Almouzni, G. (2009) Epigenetic inheritance during the cell cycle. *Nat. Rev. Mol. Cell Biol.*, **10**, 192–206.

32. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with bowtie 2. *Nat. Methods*, **9**, 357–359.

33. Li, B. and Dewey, C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.*, **12**, 323.

34. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.

35. Langfelder, P. and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.*, **9**, 559.

36. Huang, d.W., Sherman, B.T. and Lempicki, R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.*, **37**, 1–13.

37. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H. and Glass, C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.

38. Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.Y., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R. *et al.* (2016) JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.*, **44**, D110–D115.

39. Kheradpour, P. and Kellis, M. (2014) Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Res.*, **42**, 2976–2987.

40. Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G. *et al.* (2013) DNA-binding specificities of human transcription factors. *Cell*, **152**, 327–339.

41. Zhang, Z.M., Liu, S., Lin, K., Luo, Y., Perry, J.J., Wang, Y. and Song, J. (2015) Crystal structure of human DNA methyltransferase 1. *J. Mol. Biol.*, **427**, 2520–2531.

42. Coutelier, M., Coarelli, G., Monin, M.L., Konop, J., Davoine, C.S., Tesson, C., Valter, R., Anheim, M., Behin, A., Castelnovo, G. *et al.* (2017) A panel study on patients with dominant cerebellar ataxia highlights the frequency of channelopathies. *Brain*, **140**, 1579–1594.

43. Figueroa, K.P., Chan, P., Schöls, L., Tanner, C., Riess, O., Perlman, S.L., Geschwind, D.H. and Pulst, S.M. (2001) Association of moderate polyglutamine tract expansions in the slow calcium-activated potassium channel type 3 with ataxia. *Arch. Neurol.*, **58**, 1649–1653.

44. Tanaka, E., Maruyama, H., Morino, H., Nakajima, E. and Kawakami, H. (2008) The *CNTN4* c.4256C>T mutation is rare in Japanese with inherited spinocerebellar ataxia. *J. Neurol. Sci.*, **266**, 180–181.

45. Wang, Y., Ru, Y., Liu, W., Wang, D., Zhou, J., Jiang, Y., Shi, S. and Qin, L. (2017) Morphological characterization and *HSP70*-, *IGS*-based phylogenetic analysis of two microsporidian parasites isolated from *Antheraea pernyi*. *Parasitol. Res.*, **116**, 971–977.

46. Asthagiri, A.R., Vasquez, R.A., Butman, J.A., Wu, T., Morgan, K., Brewer, C.C., King, K., Zalewski, C., Kim, H.J. and Lonser, R.R. (2012) Mechanisms of hearing loss in neurofibromatosis type 2. *PLoS One*, **7**, e46132.

47. Mutsaers, P.G., van de Loosdrecht, A.A., Tawana, K., Bödör, C., Fitzgibbon, J. and Menko, F.H. (2013) Highly variable clinical manifestations in a large family with a novel *GATA2* mutation. *Leukemia*, **27**, 2247–2248.

48. Tekin, M., Oztürkmen Akay, H., Fitoz, S., Birnbaum, S., Cengiz, F.B., Sennaroğlu, L., Incesulu, A., Yüksel Konuk, E.B., Hasanefendioğlu Bayrak, A., Sentürk, S. *et al.* (2008) Homozygous *FGF3* mutations result in congenital deafness with inner ear agenesis, microtia, and microdontia. *Clin. Genet.*, **73**, 554–565.

49. Tomita, H., Shakkottai, V.G., Gutman, G.A., Sun, G., Bunney, W.E., Cahalan, M.D., Chandy, K.G. and Gargus, J.J. (2003) Novel truncated isoform of SK3 potassium channel is a potent dominant-negative regulator of SK currents: implications in schizophrenia. *Mol. Psychiatry*, **8**(524–535), 460.

50. Chen, Q.Y., Chen, Q., Feng, G.Y., Lindpaintner, K., Chen, Y., Sun, X., Chen, Z., Gao, Z., Tang, J. and He, L. (2005) Case-control association study of the close homologue of L1 (*CHL1*) gene and schizophrenia in the Chinese population. *Schizophr. Res.*, **73**, 269–274.

51. Shibata, H., Tani, A., Chikuhara, T., Kikuta, R., Sakai, M., Ninomiya, H., Tashiro, N., Iwata, N., Ozaki, N. and Fukumaki, Y. (2009) Association study of polymorphisms in the group III metabotropic glutamate receptor genes, *GRM4* and *GRM7*, with schizophrenia. *Psychiatry Res.*, **167**, 88–96.

52. Shimada, M., Miyagawa, T., Kawashima, M., Tanaka, S., Honda, Y., Honda, M. and Tokunaga, K. (2010) An approach based on a genome-wide association study reveals candidate loci for narcolepsy. *Hum. Genet.*, **128**, 433–441.

53. Cvetkovic-Lopes, V., Bayer, L., Dorsaz, S., Maret, S., Pradervand, S., Dauvilliers, Y., Lecendreux, M., Lammers, G.J., Donjacour, C.E., Du Pasquier, R.A. *et al.* (2010) Elevated Tribbles homolog 2-specific antibody levels in narcolepsy patients. *J. Clin. Invest.*, **120**, 713–719.

54. Coutelier, M., Burglen, L., Mundwiller, E., Abada-Bendib, M., Rodriguez, D., Chantot-Bastaraud, S., Rougeot, C., Cournelle, M.A., Milh, M., Toutain, A. *et al.* (2015) *GRID2* mutations span from congenital to mild adult-onset cerebellar ataxia. *Neurology*, **84**, 1751–1759.

55. Kantarci, S., Al-Gazali, L., Hill, R.S., Donnai, D., Black, G.C., Bieth, E., Chassaing, N., Lacombe, D., Devriendt, K., Teebi, A. *et al.* (2007) Mutations in *LRP2*, which encodes the multiligand receptor megalin, cause Donnai-barrow and facio-oculo-acoustico-renal syndromes. *Nat. Genet.*, **39**, 957–959.

56. Cappello, S., Gray, M.J., Badouel, C., Lange, S., Einsiedler, M., Srour, M., Chitayat, D., Hamdan, F.F., Jenkins, Z.A., Morgan, T. *et al.* (2013) Mutations in genes encoding the cadherin receptor-ligand pair DCHS1 and FAT4 disrupt cerebral cortical development. *Nat. Genet.*, **45**, 1300–1308.

57. Mansour, S., Swinkels, M., Terhal, P.A., Wilson, L.C., Rich, P., Van Maldergem, L., Zwijnenburg, P.J., Hall, C.M., Robertson, S.P. and Newbury-Ecob, R. (2012) Van Maldergem syndrome: further characterisation and evidence for neuronal migration abnormalities and autosomal recessive inheritance. *Eur. J. Hum. Genet.*, **20**, 1024–1031.

58. Bagyinszky, E., Giau, V.V., Youn, Y.C., An, S.S.A. and Kim, S. (2018) Characterization of mutations in. *Neuropsychiatr. Dis. Treat.*, **14**, 2067–2085.

59. Turnbull, J., Girard, J.M., Lohi, H., Chan, E.M., Wang, P., Tiberia, E., Omer, S., Ahmed, M., Bennett, C., Chakrabarty, A. *et al.* (2012) Early-onset Lafora body disease. *Brain*, **135**, 2684–2698.

60. Pennacchio, L.A., Bouley, D.M., Higgins, K.M., Scott, M.P., Noebels, J.L. and Myers, R.M. (1998) Progressive ataxia, myoclonic epilepsy and cerebellar apoptosis in cystatin B-deficient mice. *Nat. Genet.*, **20**, 251–258.

61. Keller, A., Westenberger, A., Sobrido, M.J., García-Murias, M., Domingo, A., Sears, R.L., Lemos, R.R., Ordoñez-Ugalde, A., Nicolas, G., da Cunha, J.E. *et al.* (2013) Mutations in the gene encoding PDGF-B cause brain calcifications in humans and mice. *Nat. Genet.*, **45**, 1077–1082.

62. Naz, S., Griffith, A.J., Riazuddin, S., Hampton, L.L., Battey, J.F., Khan, S.N., Wilcox, E.R. and Friedman, T.B. (2004) Mutations of *ESPN* cause autosomal recessive deafness and vestibular dysfunction. *J. Med. Genet.*, **41**, 591–595.

63. Cesca, F., Bettella, E., Polli, R., Cama, E., Scimemi, P., Santarelli, R. and Murgia, A. (2018) A novel mutation of the *EYA4* gene associated with post-lingual hearing loss in a proband is co-segregating with a novel *PAX3* mutation in two congenitally deaf family members. *Int. J. Pediatr. Otorhinolaryngol.*, **104**, 88–93.

64. Du, H., Ren, R., Chen, P., Xu, Z. and Wang, Y. (2018) Identification of binding partners of deafness-related protein PDZD7. *Neural Plast*, **2018**, 1–10.

65. Behlouli, A., Bonnet, C., Abdi, S., Bouaita, A., Lelli, A., Hardelin, J.P., Schietroma, C., Rous, Y., Louha, M., Cheknane, A. *et al.* (2014) *EPS8*, encoding an actin-binding protein of cochlear hair cell stereocilia, is a new causal gene for autosomal recessive profound deafness. *Orphanet. J. Rare Dis.*, **9**, 55.

66. Santos-Cortez, R.L., Faridi, R., Rehman, A.U., Lee, K., Ansar, M., Wang, X., Morell, R.J., Isaacson, R., Belyantseva, I.A., Dai, H. *et al.* (2016) Autosomal-recessive hearing impairment due to rare missense variants within S1PR2. *Am. J. Hum. Genet.*, **98**, 331–338.

67. Borck, G., Ur Rehman, A., Lee, K., Pogoda, H.M., Kakar, N., von Ameln, S., Grillet, N., Hildebrand, M.S., Ahmed, Z.M., Nürnberg, G. *et al.* (2011) Loss-of-function mutations of *ILDR1* cause autosomal-recessive hearing impairment DFNB42. *Am. J. Hum. Genet.*, **88**, 127–137.

68. Azaiez, H., Booth, K.T., Ephraim, S.S., Crone, B., Black-Ziegelbein, E.A., Marini, R.J., Shearer, A.E., Sloan-Heggen, C.M., Kolbe, D., Casavant, T. *et al.* (2018) Genomic landscape and mutational signatures of deafness-associated genes. *Am. J. Hum. Genet.*, **103**, 484–497.

69. Brown, K.K., Alkuraya, F.S., Matos, M., Robertson, R.L., Kimonis, V.E. and Morton, C.C. (2009) *NR2F1* deletion in a patient with a de novo paracentric inversion, inv(5)(q15q33.2), and syndromic deafness. *Am. J. Med. Genet. A*, **149A**, 931–938.

70. Olichon, A., Landes, T., Arnauné-Pelloquin, L., Emorine, L.J., Mils, V., Guichet, A., Delettre, C., Hamel, C., Amati-Bonneau, P., Bonneau, D. *et al.* (2007) Effects of OPA1 mutations on mitochondrial morphology and apoptosis: relevance to ADOA pathogenesis. *J. Cell. Physiol.*, **211**, 423–430.

71. Al-Kateb, H., Shimony, J.S., Vineyard, M., Manwaring, L., Kulkarni, S. and Shinawi, M. (2013) *NR2F1* haploinsufficiency is associated with optic atrophy, dysmorphism and global developmental delay. *Am. J. Med. Genet. A*, **161A**, 377–381.

72. Liu, H., Fu, Y., Ren, J., Yu, S., Jiang, P., Dong, Y. and Li, H. (2015) Association between *NR4A2* genetic variation and schizophrenia: A comprehensive systematic review and meta-analysis. *Neurosci. Lett.*, **598**, 85–90.

73. Kang, W.S., Park, J.K., Kim, S.K., Park, H.J., Lee, S.M., Song, J.Y., Chung, J.H. and Kim, J.W. (2012) Genetic variants of *GRIA1* are associated with susceptibility to schizophrenia in Korean population. *Mol. Biol. Rep.*, **39**, 10697–10703.

74. Lencer, R., Bishop, J.R., Harris, M.S., Reilly, J.L., Patel, S., Kittles, R., Prasad, K.M., Nimgaonkar, V.L., Keshavan, M.S. and Sweeney, J.A. (2014) Association of variants in *DRD2* and *GRM3* with motor and cognitive function in first-episode psychosis. *Eur. Arch. Psychiatry Clin. Neurosci.*, **264**, 345–355.

75. Magri, C., Gardella, R., Valsecchi, P., Barlati, S.D., Guizzetti, L., Imperadori, L., Bonvicini, C., Tura, G.B., Gennarelli, M., Sacchetti, E. and Barlati, S. (2008) Study on *GRIA2*, *GRIA3* and *GRIA4* genes highlights a positive association between schizophrenia and *GRIA3* in female patients. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **147B**, 745–753.

76. Chen, X., Long, F., Cai, B., Chen, G. and Chen, G. (2017) A novel relationship for schizophrenia, bipolar and major depressive disorder part 5: a hint from chromosome 5 high density association screen. *Am. J. Transl. Res.*, **9**, 2473–2491.

77. Arab, A.H. and Elhawary, N.A. (2015) Association between *ANKK1* (rs1800497) and *LTA* (rs909253) genetic variants and risk of schizophrenia. *Biomed. Res. Int.*, **2015**, 821827.

78. Sakai, M., Watanabe, Y., Someya, T., Araki, K., Shibuya, M., Niizato, K., Oshima, K., Kunii, Y., Yabe, H., Matsumoto, J. *et al.* (2015) Assessment of copy number variations in the brain genome of schizophrenia patients. *Mol. Cytogenet.*, **8**, 46.

79. Liu, H., Heath, S.C., Sobin, C., Roos, J.L., Galke, B.L., Blundell, M.L., Lenane, M., Robertson, B., Wijsman, E.M., Rapoport, J.L., Gogos, J.A. and Karayiorgou, M. (2002) Genetic variation at the 22q11 *PRODH2/DGCR6* locus presents an unusual pattern and increases susceptibility to schizophrenia. *Proc. Natl. Acad. Sci. U. S. A.*, **99**, 3717–3722.

80. Beck, J.A., Campbell, T.A., Adamson, G., Poulter, M., Uphill, J.B., Molou, E., Collinge, J. and Mead, S. (2008) Association of a null allele of *SPRN* with variant Creutzfeldt-Jakob disease. *J. Med. Genet.*, **45**, 813–817.

81. Parachikova, A. and Cotman, C.W. (2007) Reduced CXCL12/CXCR4 results in impaired learning and is downregulated in a mouse model of Alzheimer disease. *Neurobiol. Dis.*, **28**, 143–153.

82. Wang, C., Li, Y., Shi, L., Ren, J., Patti, M., Wang, T., de Oliveira, J.R., Sobrido, M.J., Quintáns, B., Baquero, M. *et al.* (2012) Mutations in *SLC20A2* link familial idiopathic basal ganglia calcification with phosphate homeostasis. *Nat. Genet.*, **44**, 254–256.

83. Mackay, D.J., Callaway, J.L., Marks, S.M., White, H.E., Acerini, C.L., Boonen, S.E., Dayanikli, P., Firth, H.V., Goodship, J.A., Haemers, A.P. *et al.* (2008) Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in *ZFP57*. *Nat. Genet.*, **40**, 949–951.

84. Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P.V., Grimaldi, G., Riccio, A. and Trono, D. (2011) In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol. Cell*, **44**, 361–372.

85. Riso, V., Cammisa, M., Kukreja, H., Anvar, Z., Verde, G., Sparago, A., Acurzio, B., Lad, S., Lonardo, E., Sankar, A. *et al.* (2016) ZFP57 maintains the parent-of-origin-specific expression of the imprinted genes and differentially affects non-imprinted targets in mouse embryonic stem cells. *Nucleic Acids Res.*, **44**, 8165–8178.

86. Zampini, V., Rüttiger, L., Johnson, S.L., Franz, C., Furness, D.N., Waldhaus, J., Xiong, H., Hackney, C.M., Holley, M.C., Offenhauser, N. *et al.* (2011) Eps8 regulates hair bundle length and functional maturation of mammalian auditory hair cells. *PLoS Biol.*, **9**, e1001048.

87. MacLennan, A.J., Benner, S.J., Andringa, A., Chaves, A.H., Rosing, J.L., Vesey, R., Karpman, A.M., Cronier, S.A., Lee, N., Erway, L.C. and Miller, M.L. (2006) The S1P2 sphingosine 1-phosphate receptor is essential for auditory and vestibular function. *Hear. Res.*, **220**, 38–48.

88. Manyam, B.V. (2005) What is and what is not 'Fahr's disease'. *Parkinsonism Relat. Disord.*, **11**, 73–80.

89. MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Junkins, H., McMahon, A., Milano, A., Morales, J. *et al.* (2017) The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.*, **45**, D896–D901.

90. Khor, S.S., Miyagawa, T., Toyoda, H., Yamasaki, M., Kawamura, Y., Tanii, H., Okazaki, Y., Sasaki, T., Lin, L., Faraco, J. *et al.* (2013) Genome-wide association study of *HLA-DQB1∗06:02* negative essential hypersomnia. *PeerJ.*, **1**, e66.

91. Bonvalet, M., Ollila, H.M., Ambati, A. and Mignot, E. (2017) Autoimmunity in narcolepsy. *Curr. Opin. Pulm. Med.*, **23**, 522–529.

92. Maresca, A., Del Dotto, V., Capristo, M., Scimonelli, E., Tagliavini, F., Morandi, L., Tropeano, C.V., Caporali, L., Mohamed, S., Roberti, M. *et al.* (2020) *DNMT1* mutations leading to neurodegeneration paradoxically reflect on mitochondrial metabolism. *Hum. Mol. Genet.*, **29**, 1864–1881.

93. Lin, H.C., He, Z., Ebert, S., Schörnig, M., Santel, M., Nikolova, M.T., Weigert, A., Hevers, W., Kasri, N.N., Taverna, E., Camp, J.G. and Treutlein, B. (2021) *NGN2* induces diverse neuron types from human pluripotency. *Stem Cell Reports*, **16**, 2118–2127.

94. Soldner, F., Laganière, J., Cheng, A.W., Hockemeyer, D., Gao, Q., Alagappan, R., Khurana, V., Golbe, L.I., Myers, R.H., Lindquist, S. *et al.* (2011) Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell*, **146**, 318–331.

95. Singh Dolt, K., Hammachi, F. and Kunath, T. (2017) Modeling Parkinson's disease with induced pluripotent stem cells harboring *α*-synuclein mutations. *Brain Pathol.*, **27**, 545–551.

96. Zhang, S., Zhang, H., Zhou, Y., Qiao, M., Zhao, S., Kozlova, A., Shi, J., Sanders, A.R., Wang, G., Luo, K. *et al.* (2020) Allele-specific open chromatin in human iPSC neurons elucidates functional disease variants. *Science*, **369**, 561–565.