OXFORD

Human Molecular Genetics, 2016, Vol. 25, No. 22 4939–4950

doi: 10.1093/hmg/ddw317
Advance Access Publication Date: 16 September 2016
Original Article

ORIGINAL ARTICLE

Cis-acting single nucleotide polymorphisms alter MicroRNA-mediated regulation of human brain-expressed transcripts

Shyam Ramachandran[1], Stephanie L. Coffin[1], Tin-Yun Tang[2,†], Chintan D. Jobaliya[1,3], Ryan M. Spengler[4] and Beverly L. Davidson[1,5,*]

[1]Raymond G. Perelman Center for Cellular and Molecular Therapeutics, The Children's Hospital of Philadelphia, Philadelphia, USA, [2]Howard Hughes Medical Institute Medical Research Fellow, Howard Hughes Medical Institute, Chevy Chase, MD, USA, [3]Human Pluripotent Stem Cell Core, Raymond G. Perelman Center for Cellular and Molecular Therapeutics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA, [4]Department of Internal Medicine, Division of Hematology/Oncology, University of Michigan, Ann Arbor, MI, USA and [5]The Department of Pathology & Laboratory Medicine, The Children's Hospital of Philadelphia and The University of Pennsylvania, Philadelphia, PA, USA

*To whom correspondence should be addressed at: Beverly L Davidson, The Children's Hospital of Philadelphia, Philadelphia, 5060 Colket Translational Research Building, Pennsylvania 19104, USA. Tel: +1 267-426-0929; Fax: +1 215-590-3660; E-mail: davidsonbl@email.chop.edu

Abstract

Substantial variability exists in the presentation of complex neurological disorders, and the study of single nucleotide polymorphisms (SNPs) has shed light on disease mechanisms and pathophysiological variability in some cases. However, the vast majority of disease-linked SNPs have unidentified pathophysiological relevance. Here, we tested the hypothesis that SNPs within the miRNA recognition element (MRE; the region of the target transcript to which the miRNA binds) can impart changes in the expression of those genes, either by enhancing or reducing transcript and protein levels. To test this, we cross-referenced 7,153 miRNA-MRE brain interactions with the SNP database (dbSNP) to identify candidates, and functionally assessed 24 SNPs located in the 3'UTR or the coding sequence (CDS) of targets. For over half of the candidates tested, SNPs either enhanced (4 genes) or disrupted (10 genes) miRNA binding and target regulation. Additionally, SNPs causing a shift from a common to rare codon within the CDS facilitated miRNA binding downstream of the SNP, dramatically repressing target gene expression. The biological activity of the SNPs on miRNA regulation was also confirmed in induced pluripotent stem cell (iPSC) lines. These studies strongly support the notion that SNPs in the 3'UTR or the coding sequence of disease-relevant genes may be important in disease pathogenesis and should be reconsidered as candidate modifiers.

†Present address: The Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA 19107, USA.
Received: July 5, 2016. **Revised:** September 6, 2016. **Accepted:** September 12, 2016

Introduction

Genome wide association studies (GWAS) (1,2) and work defining quantitative trait loci (QTLs) (3,4) have shown that some single nucleotide polymorphisms (SNPs) associate with the onset or susceptibility of complex diseases (5). Although disease-associated SNPs commonly alter the amino acid sequence of the translated protein, SNPs located within a disease allele may not cause obvious protein misfolding or impair transcript stability, and may even present as conservative or synonymous mutations. While many SNPs are in linkage disequilibrium, there is limited information connecting the actual SNP to disease pathophysiology (6).

The majority of human genes are under post-transcriptional regulation by miRNAs (7) and single base changes can alter their effectiveness in repressing transcript levels. We postulate that an SNP-based effects on protein expression or stability could occur if they reside within a microRNA recognition element (MRE). In this manner, SNPs could create, destroy, or modify microRNA (miRNA) interactions with the mRNA. This phenomenon was demonstrated earlier when the SNP rs7646 was shown to alter miR-197 binding to the *MTHFD1L* 3'UTR, causing more efficient repression in the presence of the allele associated with increased risk of neural tube disorders (8). Similarly, SNPs alter miR-330-3p mediated regulation of *MAP2K5*, potentially contributing to mood and anxiety-related traits (9). Given the relevance of miRNAs in regulating neural patterning (10), neural stem cell differentiation (11), cell type specification (12), synaptic plasticity (13), and neuropsychiatric disorders (14–16) we looked at a role for SNPs in miRNA-based regulation on a broad scale.

Bioinformatics driven approaches predict a number of possible miRNA-MRE interactions in the human brain (17). A transcriptome-wide map of Argonaute (Ago2) CLIP-seq data has refined our understanding of those miRNAs engaged in gene silencing in mouse (18) and human brain (19). Of the 7,153 miRNA-MRE reported interactions constituting stringent Ago2 binding sites for the most abundant brain miRNAs, functional regulation of >3,000 genes are likely. The mapped interactions occur nearly equally in the coding sequence (CDS) and the 3' untranslated region (3' UTR) of mRNAs (18,19). While 3' UTR binding is the classical mechanism that prediction algorithms focus on, CDS interactions can occur, as MREs in transcript coding regions have been shown to impart gene regulation (20,21).

In this study, we cross-referenced SNP locations from dbSNP (22) against the 7,153 miRNA-MRE reported interactions in the human brain (19). The pipeline for prioritization of SNPs was to i) require that the SNPs reside within the seed sequence of the MRE, and ii) exclude genes with additional distinct MRE targets to more abundant miRNAs (by Ago2-CLIP). We selected 24 candidate genes for validation. In 13 the MRE resides in the CDS and in 11 the MRE is in the 3'UTR of the target gene. We also tested the impact of synonymous SNPs in the CDS on gene expression. In all, we demonstrate that SNP-resident MREs can alter mRNA and protein levels of brain-expressed transcripts.

Results

We used a systematic approach to prioritize gene candidates from reported miRNA-MRE brain interactions for further study (Fig. 1A) . We intersected the 7,153 Ago2-CLIP human brain data (19) with dbSNP (22) and identified 426 miRNA-MRE interactions containing SNPs within the 'seed sequence' of the associated MRE. Only genes with an alternative SNP (referred to as altSNP from this point forward) having a minor allele frequency (MAF) greater than 1% were considered for further analysis. Next, we used the biological abundance of the individual miRNA (only the top 50 expressed miRNAs were considered) (19) and the overlying gene's role in neurological processes (23) (Supplementary Material, Table S1) as additional selection criteria to narrow the candidate pool to 38. In anticipation of the possibility that the altSNP could facilitate binding of another abundant miRNA, we used the computational tool RNA22 V2 (24) to predict if any of the top 50 brain expressed miRNAs could bind from +21nt to -21nt of the altSNP site. None of the 38 candidate altSNPs generated novel binding sites for these miRNAs. Finally, genes with other distinct MREs (by Ago2-CLIP) that paired with more abundant miRNAs (compared to the miRNA binding the SNP-resident MRE) were excluded, resulting in a final list of 24 genes where the altSNP is predicted to alter binding to the most highly expressed regulatory miRNA (Fig. 1A). We deemed this filtering step essential as Ago2-CLIP studies suggest that the majority of Ago2 peaks could be explained by the most abundant miRNAs (18,19). We observed a fairly equal distribution of candidate genes; in 13 genes the SNP-resident MRE resided in the CDS and in 11 the SNP-resident MRE was in the 3'UTR of the target gene (Supplementary Material, Table S2).

Seventeen unique miRNAs were represented across the 24 candidates. We cloned each miRNA into a vector whereby expression was driven by Pol III U6. Their biological activity was confirmed by assessing their suppressive action on *Renilla* luciferase with each miRNA showing a dose dependent repression of their respective perfect targets (Supplementary Material, Fig. S1). All miRNAs under study were also expressed and interacted with target mRNAs by PAR-CLIP in HEK293 cells (25).

SNPs in 3'UTR MREs alter miRNA-mediated regulation

Next, we cloned the entire 3'UTR (containing the ancestral SNP, referred to as reference allele) of the 11 target genes with a 3'UTR-resident MRE (Supplementary Material, Table S2) downstream of a *Renilla* luciferase reporter (referred to as 'Ref' for reference). We generated the altSNPs by site-directed mutagenesis (referred to as 'Alt' for alternative) and scrambled the MRE at seed location 2-5 to abrogate putative miRNA binding (referred to as 'Scr'). Fig. 1B depicts how a C->G SNP could disrupt, and how a G->C SNP could enhance miRNA-mediated repression of a target gene.

When HEK293 cells were transfected with *RAPGEF1*-Alt or *SOD1*-Alt, a significant decrease in *Renilla* mRNA and protein levels was observed by RT-qPCR and luciferase assay, relative to the Ref construct (Fig. 1C and D). In contrast, there was a significant increase in target mRNA and protein levels in cells transfected with the Scr plasmid, indicating regulation by endogenous miR-124-3p (targets *RAPGEF1*) and miR-139-5p (targets *SOD1*) is lost (Fig. 1C and D). The effect of the miRNA on regulation of the altSNP-resident MRE was also confirmed by varying miRNA doses using miRNA overexpression constructs. Increasing miR-124-3p expression plasmid levels resulted in increased knockdown of RAFGEF1-Alt (compared to Ref) at every ratio measured, while scrambling the seed prevented this effect (Supplementary Material, Fig. S2A). SOD1-Alt showed similar results on overexpressing miR-139-5p (Supplementary Material, Fig. S2B). Thus SNPs rs3739497 (*RAPGEF1*) and rs41391245 (*SOD1*) improve miRNA-mediated repression.

Introducing the altSNP in *CARF* (rs62183749, Fig. 1E), *TUBB4A* (rs8113500, Fig. 1F), *FBXO45* (rs78195212, Fig. 1G) and *RPRD1A* (rs78544189, Fig. 1H) disrupted miRNA-mediated repression,

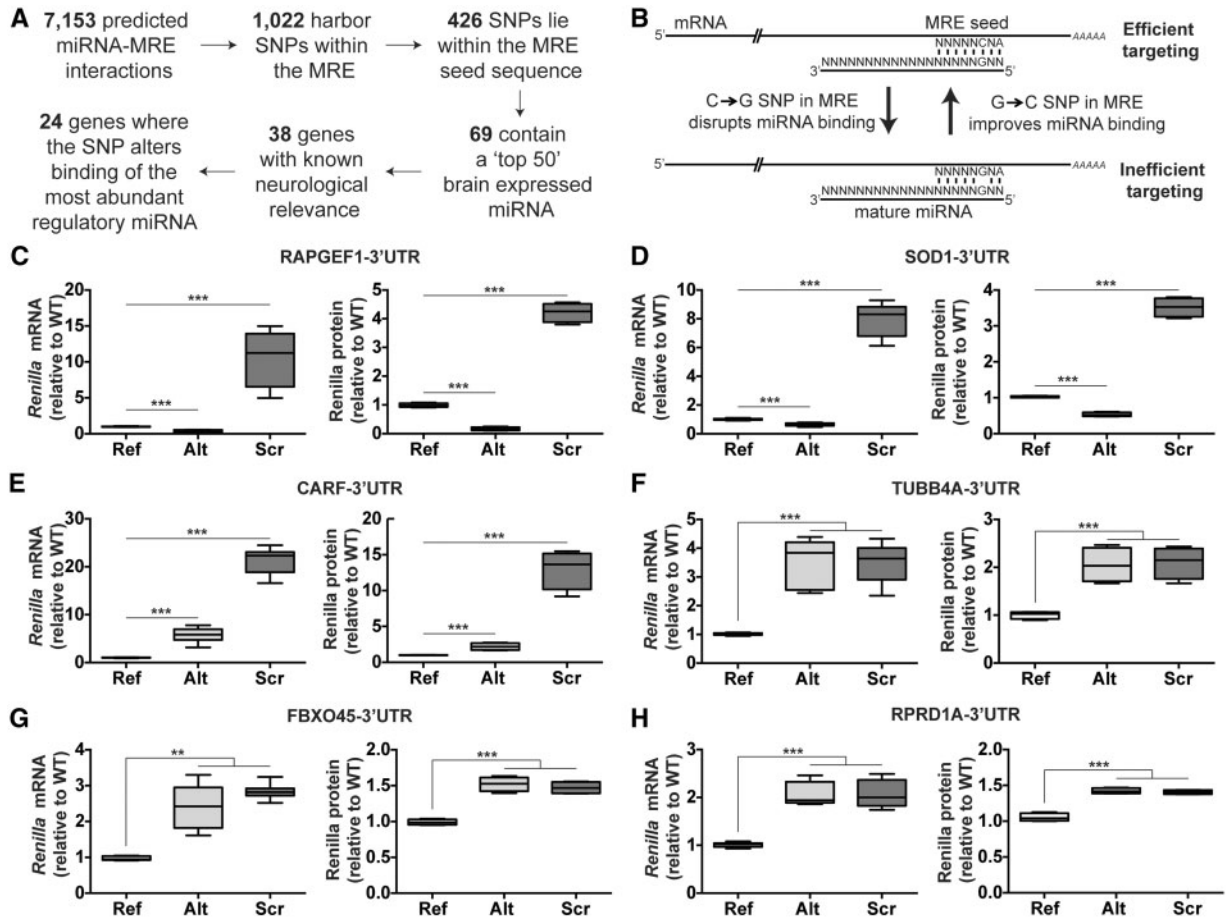

Figure 1. SNPs in the 3'UTR MRE modify miRNA-mediated regulation of target genes. **(A)** Flowchart describing steps utilized to prioritize candidate miRNA-MRE interactions for further study. Of the 24 candidates selected, 11 SNP-resident MREs are in the 3'UTR and 13 SNP-resident MREs are in the CDS of target genes. **(B)** Schematic describing how a SNP in a MRE could enhance or disrupt miRNA-mediated target gene regulation. **(C–H)** HEK293 cells were transfected with constructs containing the reference SNP (Ref), the alternative 'candidate' SNP (Alt), or a scrambled MRE (Scr) for the following genes **(C)** RAPGEF1, **(D)** SOD1, **(E)** CARF, **(F)** TUBB4A, **(G)** FBXO45 and **(H)** RPRD1A and analysed for RNA and protein expression. RNA and protein was harvested 48 h post-transfection. Left panel: RT-qPCR for *Renilla* (normalized to *Firefly* mRNA), presented as fold change relative to WT ($n = 9$). Right panel: Relative luminescence units (RLU) normalized (*Renilla*/*Firefly*) and presented as fold change relative to WT ($n = 12$). All panels: data presented as box & whiskers plot (min to max) and analysed by one-way ANOVA with Sidak's multiple comparison test, $^{**}P < 0.01$, $^{***}P < 0.001$.

resulting in increased *Renilla* mRNA and protein levels compared to WT, and often to the same extent as the Scr controls. Similarly, while increasing the dose of miR-181-5p (targets *CARF*, Supplementary Material, Fig. S2C), miR-1301-3p (targets *TUBB4A*, Supplementary Material, Fig. S2D), miR-381-3p (targets *FBXO45*, Supplementary Material, Fig. S2E) and miR-23a-3p (targets *RPRD1A*, Supplementary Material, Fig. S2F) boosted repression of the Ref constructs, it was less effective in reducing the levels of 3'UTRs harbouring the relevant altSNPs.

Other genes with altSNPs residing in 3'UTR MREs show a similar pattern of regulation, namely *PDE8B* (Supplementary Material, Fig. S2G) and *RAB5A* (Supplementary Material, Fig. S2I). In both cases, the Ref gene product showed significant repression by endogenous miRNAs, while the altSNP significantly diminished miRNA-mediated repression (*PDE8B*-rs1580, *RAB5A*-rs8682). Overexpressing the targeting miRNA boosted PDE8B-Ref (Supplementary Material, Fig. S2H) and RAB5A-Ref (Supplementary Material, Fig. S2J) repression, and introducing the altSNP dampened it, at least at low miRNA expression plasmid doses. We also found instances where the altSNP had no effect. For example, for PPIP5K2-Alt (rs8488, Supplementary Material, Fig. S2K-L) and UNC80-Alt (rs73080885, Supplementary Material, Fig. S2M and N) we found similar levels of miRNA-mediated repression to the Ref. SSB is probably not a real target (Supplementary Material, Fig. S2O and P), as significant repression occurred only with miR-7-5p overexpression. Thus, 8 of 11 candidate genes with 3'UTR MREs show strong altSNP-dependent effects on miRNA-mediated repression.

SNPs in CDS MREs alter miRNA-mediated regulation

We identified 13 genes where the SNP-resident MRE is located within the CDS of the target mRNA (Supplementary Material, Table S2). Their cDNAs, derived from human brain RNA to ensure that the appropriate isoform is studied, along with a Flag- and His-tag epitope, were cloned downstream of a CMV promoter (referred to as 'Ref') for expression. The relevant altSNP (referred to as 'Alt') or a scrambled MRE (seed nt 2-5; referred to as 'Scr') was introduced by site-directed mutagenesis. Of note, we made synonymous mutations with similar codon frequencies to the ancestral codon to generate the Scr constructs.

Introducing the corresponding altSNP significantly enhanced endogenous miRNA-mediated repression of *SREBF1* (rs2229591, Fig. 2A, Supplementary Material, Fig. S3A), *TUBA1A* (rs137853044,

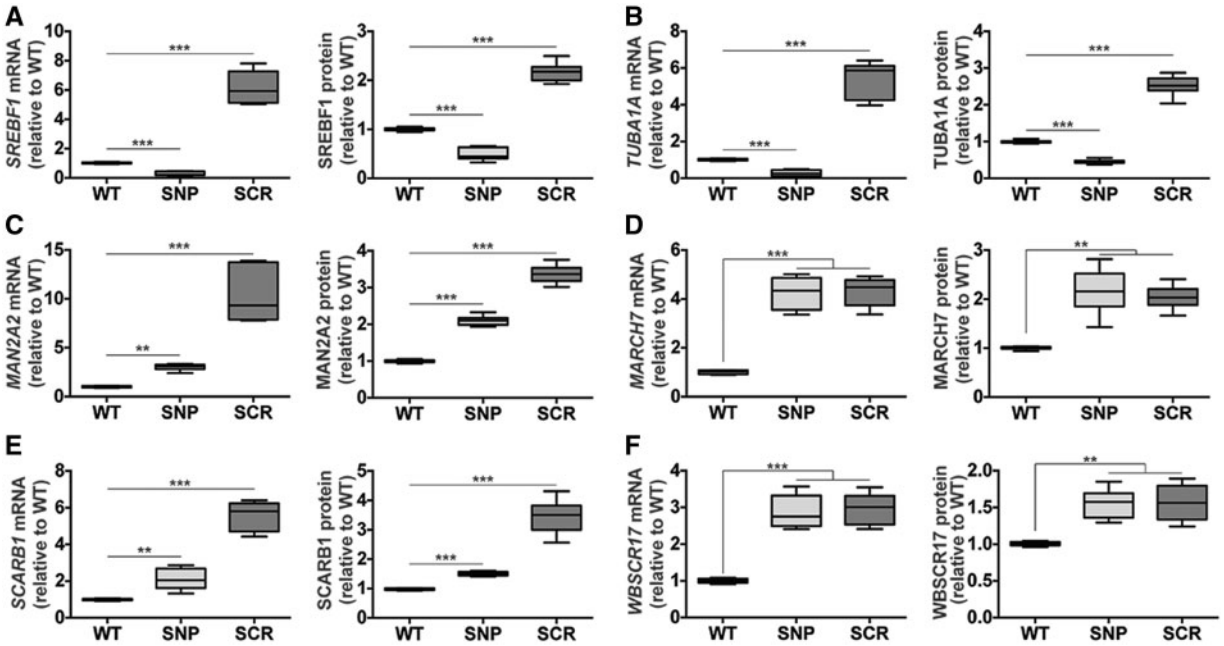

Figure 2. SNPs in the CDS MRE modify miRNA-mediated regulation of target genes. **(A–F)** HEK293 cells were transfected with Ref, Alt or Scr vectors for the following genes (A) SREBF1, (B) TUBA1A, (C) MAN2A2, (D) MARCH7, (E) SCARB1, and (F) WBSCR17. RNA and protein was harvested 48 h post-transfection. Left panel: RT-qPCR for respective genes (normalized to HPRT), presented as fold change relative to WT ($n = 9$). Right panel: ELISA (anti-Flag Ab) for the respective genes, presented as fold change relative to WT ($n = 10$). All panels: data presented as box & whiskers plot (min to max) and analysed by one-way ANOVA with Sidak's multiple comparison test, $^{**}P < 0.01$, $^{***}P < 0.001$.

Fig. 2B, Supplementary Material, Fig. S3B), and *HSPD1* (rs1050347, Supplementary Material, Figs S3C and S4C). Overexpression of miR-29b-3p (targets *SREBF1*, Supplementary Material, Fig. S4A) and miR-124-3p (targets *TUBA1A*, Supplementary Material, Fig. S4B) resulted in significantly greater repression of altSNP genes compared to Ref. Overexpression of miR-9-5p (targets *HSPD1*, Supplementary Material, Fig. S4D) resulted in higher levels of HSPD1-Ref. We speculate this is a consequence of miRNA overexpression and perturbation of other pathways regulating HSPD1 expression.

AltSNPs disrupted miRNA binding in *MAN2A2* (rs2677743, Fig. 2C, Supplementary Material, Fig. S3D), *MARCH7* (rs17813964, Fig. 2D, Supplementary Material, Fig. S3E), *SCARB1* (rs5889, Fig. 2E, Supplementary Material, Fig. S3F), and *WBSCR17* (rs17143818, Fig. 2F, Supplementary Material, Fig. S3G). AltSNPs in the CDS of *MARCH7* and *WBSCR17* abrogated miRNA-mediated repression to levels similar to Scr. However, altSNPs in the CDS of *MAN2A2* and *SCARB1*, while not as effective as the Scr, still caused significant de-repression of both genes. The significant de-repression in the setting of the Scr for these genes supports regulation by the endogenous corresponding miRNAs.

Overexpression of let-7a-5p (targets *MAN2A2* and *MARCH7*, Supplementary Material, Fig. S4E and F) and miR-124-3p (targets *SCARB1*, Supplementary Material, Fig. S4G) caused the increased repression of the Ref construct, and in each case, the altSNP construct showed less repression compared to Ref. Overexpression of miR-30b-5p (targets *WBSCR17*, Supplementary Material, Fig. S4H) caused increased repression of WBSCR17-Alt, divergent to the regulation observed under endogenous miRNA levels (Fig. 2F). Of the remaining 6 genes in the CDS screen, *PPP3R1*, *SYN1*, and *SYT* showed regulation by endogenous miRNAs (Supplementary Materials, Figs S3H–J, S4I–N), but the respective altSNPs had no impact. The remaining 3 genes were not regulated by the corresponding candidate miRNAs (Supplementary Materials, Figs. S3K–M, S4O–T). Thus, 7 of 13 CDS-resident MRE candidates show strong SNP-dependent effects on miRNA-mediated repression.

Rare codon-generating altSNPs trigger miRNA-mediated repression

Many groups have explored the biological basis for the evolutionary conservation of functional MREs in the 3'UTR of target mRNAs. It has been suggested that ribosomal traffic would minimize miRNA-mediated regulation in the CDS region (26–29). Gu and colleagues showed that the introduction of a rare codon in a reporter gene can cause ribosomal pausing, providing a platform for miRNA-mediated recognition and subsequent degradation of the target CDS (26). To test the impact of SNPs introducing synonymous codons on miRNA-mediated regulation, we intersected dbSNP (22) with the previously reported Ago2-CLIP data. This was done to ensure that candidate mRNAs were being regulated by the miRNA machinery. Of the resultant 2,314 genes, we ranked the codon (within which the SNP lay) based on human codon frequencies, and a log10 ratio of reference/altSNP codon for every SNP was generated. Those genes with a log10 ratio >1 of their altSNP-codon (24 genes) were then further refined for known neurological relevance (6 genes) (Fig. 3A) (Supplementary Material, Tables S1 and S3). We cloned the cDNAs (derived from human brain RNA) of the 6 genes along with a Flag- or His-tag as before (denoted as Ref), and tested the effects of the altSNPs (Alt) on miRNA mediated repression (Fig. 3B).

ABCA7-Alt (rs4147914) mRNA and protein levels were significantly reduced relative to the Ref construct (Fig. 3C, Supplementary Material, S5A and B). To test if the altSNP-resident

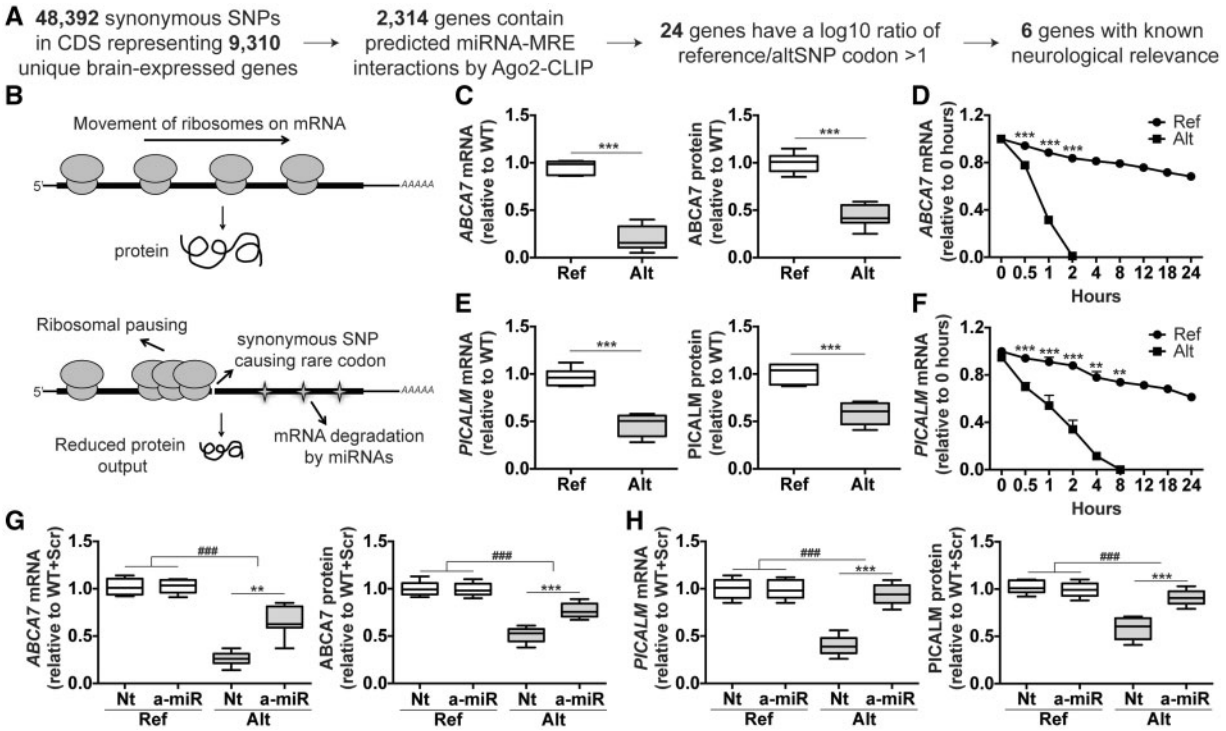

Figure 3. Synonymous codon changes generate active MREs. **(A)** Flowchart describes steps utilized to prioritize genes. **(B)** Schematic describing how a synonymous codon change via a SNP in the CDS could cause ribosomal pausing, resulting in miRNA-mediated silencing. HEK293 cells were transfected with the Ref or Alt plasmids for **(C, D)** ABCA7 or **(E–F)** PICALM. **(C, E)** Left panel: RT-qPCR for respective genes (normalized to HPRT), presented as fold change relative to WT ($n = 9$). Right panel: ELISA (anti-Flag Ab) for respective genes, presented as fold change relative to WT ($n = 10$). **(D, F)** Actinomycin-D was added 48 h post-transfection and chase performed (RNA harvest) at noted time points. RT-qPCR for the respective genes (normalized to HPRT), presented as fold change relative to time 0 ($n = 3$). **(G, H)** HEK293 cells were co-transfected with Ref or Alt plasmids along with a non-targeting oligonucleotide control (Nt) or **(G)** anti-miR-34a-5p or **(H)** anti-miR-34a-5p + anti-miR-103a-3p. RNA and protein was harvested 48 h post-transfection. Left panel: RT-qPCR for respective genes (normalized to HPRT), presented as fold change relative to Ref + Nt/anti-miR. $n = 9$. Right panel: ELISA (anti-Flag Ab) for respective genes, presented as fold change relative to Ref + Nt/anti-miR. $n = 10$. Panels D, F: data presented as mean with standard error; remaining panels: data presented as box & whiskers plot (min to max) and analysed by one-way ANOVA with Sidak's multiple comparison test, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Two-way ANOVA was used to test for interaction, $^{###}P < 0.001$.

mRNA was less stable, we performed an Actinomycin-D pulse chase. These data show that the half-life of the mRNA went from greater than 24 h to ~2 h with the introduction of the altSNP (Fig. 3D). Similar results were obtained for *PICALM* (rs592297) (Fig. 3E and F, Supplementary Material, Fig. S5C and D) and HMGCS1 (rs56257144) (Supplementary Material, Fig. S5E–G), wherein the altSNP significantly decreased gene expression. SNPs in the CDS of *CNTN6*, *GRIN2A* and *SORL1* (Supplementary Material, Fig. S5H–M) had no effect on mRNA or protein levels when tested in HEK293 cells.

To test which miRNAs mediated the observed reduction in mRNA levels, we searched for predicted MREs (of the top 50 brain expressed miRNAs) on the CDS of *ABCA7*, *PICALM*, and *HMGCS1* downstream of the SNP site using the computational tool RNA22 V2 (24) (Supplementary Material, Table S4). There are 46 MREs on *ABCA7*, comprising binding sites for 15 of the top 50 brain expressed miRNAs. *PICALM* has 8 predicted MREs (6 miRNAs) and *HMGCS1* has 1 predicted MRE. miR-34a-5p has 12 MREs on *ABCA7*. Although not the most abundant miRNA targeting *ABCA7*, based on the site type and folding energies for each MRE, miR-34a-5p is the top predicted regulatory miRNA. *PICALM* has two MREs each for miR-34a-5p and miR-103a-3p. To test their relevance, we co-transfected HEK293 cells with Ref or Alt plasmids for each gene along with the anti-miRNAs for miR-34a-5p or miR-103a-3p. We observed significant de-repression of both *ABCA7* (Fig. 3G, Supplementary Material, Fig. S5B on transfecting anti-miR-34a-5p) and *PICALM* (Fig. 3H, Supplementary Material, Fig. S5D on transfecting anti-miR-34a-5p + anti-miR-103a-3p). These results suggest that the activity of miR-34a-5p on *ABCA7* and the activity of miR-34a-5p + miR-103a-3p on *PICALM* are directly responsible for target repression when the altSNP is present.

SNPs in MREs alter miRNA binding

The data thus far demonstrate the effect of the altSNP on miRNA-mediated repression, either by enhancing that repression or reducing it. Whether this occurs through altSNP-induced enhanced or disrupted miRNA binding is unknown. To test this, we co-transfected HEK293 cells with the Ref + Alt expression constructs for the 15 candidate genes (described in Figs 1–3) and assessed target mRNA levels after argonaute 2 (Ago2) immunoprecipitation. Allele-specific primers allowed determination of the Ref to Alt product ratios by RT-qPCR. For *RAPGEF1*, *SREBF1* and *TUBA1A* we observed significantly increased association of the altSNP-resident mRNA to Ago2. (Fig. 4). All 3 genes had previously displayed significantly decreased mRNA and protein expression of the altSNP-containing construct. *CARF*, *TUBB4A*, *FBXO45*, *RPRD1A*, *MAN2A2*, *MARCH7* and *WBSCR17* showed significantly reduced association of the altSNP-resident mRNA to Ago2 (Fig. 4). Again, these 7 genes had previously displayed significantly increased mRNA and protein expression of

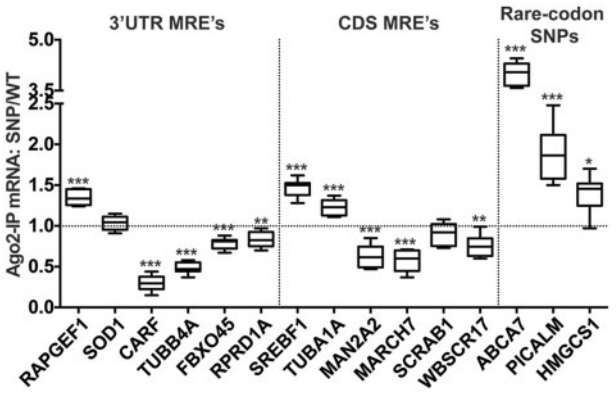

Figure 4. SNPs in MREs alter miRNA binding. HEK293 cells were co-transfected with Ref + Alt plasmids of the noted genes. Ago2 immunoprecipitation (anti-Ago2 Ab) and associated RNA was harvested 24 h later. RT-qPCR using allele-specific primers (Ref vs Alt) for the respective genes (normalized to HPRT and total gene expression Ref + Alt) was performed. Data are presented as the ratio of Alt/Ref ($n = 12$). Data are presented as box & whiskers plot (min to max) and analysed by one-way ANOVA with Sidak's multiple comparison test (SNP vs WT), $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

the altSNP-containing construct. The altSNP-containing mRNA of *ABCA7*, *PICALM* and *HMGCS1* showed significantly increased association with Ago2, lending validation to the hypothesis that altSNP-induced ribosomal pausing potentiated miRNA-binding downstream. These findings indicate that SNPs in the CDS or 3'UTR alter miRNA binding to target mRNAs.

SNPs modify miRNA-mediated regulation in induced pluripotent stem cells (iPSCs)

To test the effects of the altSNPs on miRNA-mediated regulation *in situ*, we screened a library of 14 iPSC lines generated from population control donors for heterozygosity. iPSC lines were classified as altSNP$^{-/-}$ (homozygous for ancestral/reference allele), altSNP$^{+/+}$ (homozygous for the alternative SNP), and altSNP$^{+/-}$ (alternative SNP/reference allele). As expected from genetic variation between individuals (30,31), the donors show a range of gene expression levels (Supplementary Material, Fig. S6).

The impact of altSNPs on miRNA-mediated regulation of *RAPGEF1*, *SOD1*, *SREBF1*, and *ABCA7* were assessed in iPSC lines from 3 altSNP$^{-/-}$ and 3 altSNP$^{+/-}$ lines, while 3 altSNP$^{+/+}$ and 3 altSNP$^{+/-}$ lines were used for *CARF* and *PICALM* because of the low incidence of the ancestral allele (Supplementary Material, Fig. S6). Changes in mRNA or protein levels in the altSNP$^{+/+}$, altSNP$^{+/-}$ or altSNP$^{-/-}$ lines in response to either blocking the miRNA with an anti-miR (a-miR), or transfecting in a miR-mimic (m-miR; depending on the target), was determined by RT-qPCR and immunoblot. We found that blocking miR-124-3p activity resulted in a slight increase in RAPGEF1 mRNA and protein levels in altSNP$^{-/-}$ lines, with greater increases in the altSNP$^{+/-}$ lines (Fig. 5A, Supplementary Material, Fig. S7). Similar results were observed for SOD1 (Fig. 5B, Supplementary Material, Fig. S7), SREBF1 (Fig. 5D, Supplementary Material, Fig. S7), ABCA7 (Fig. 5E, Supplementary Material, Fig. S7) and PICALM (Fig. 5F, Supplementary Material, Fig. S7). With respect to CARF, transfecting the targeting miR-mimic decreased CARF expression in both altSNP$^{+/+}$ and altSNP$^{+/-}$ lines. The decrease observed in the altSNP$^{+/+}$ lines, was however more robust and significantly different from that observed in the altSNP$^{+/-}$ (Fig. 5C, Supplementary Material, Fig. S7).

We next confirmed the effects of the SNPs on miRNA binding in 3 altSNP$^{+/-}$ lines by Ago2 immunoprecipitation (Supplementary Material, Fig. S6). For *RAPGEF1* and *SREBF1* there was significantly increased association of Ago2 with the altSNP-resident mRNA, relative to the wild type mRNA. Blocking activity of the associated miRNA using anti-miRs abrogated the observed preferential binding with Ago2 (Fig. 6). The increased association of Ago2 with the altSNP-resident mRNA is higher in the case of *ABCA7* and *PICALM*. Blocking activity of the most abundant targeting miRNAs predicted by RNA22 to target *ABCA7* or *PICALM* (Supplementary Material, Table S4) caused a drop in preferential Ago2 association with the altSNP-resident mRNA (Fig. 6). In the case of *CARF*, we observed decreased association of Ago2 with the altSNP-resident mRNA and introducing the targeting miRNA-mimic caused a significant increase in Ago2 association with the altSNP-resident mRNA (Fig. 6). Together, the results obtained in the iPSC lines corroborate the observations made in HEK293 cells. More importantly, they indicate that our predictions made with regard to SNPs in MREs modifying miRNA binding and miRNA-mediated regulation hold true under endogenous conditions.

Discussion

While disease-associated SNPs have been identified and mapped, the pathophysiological relevance for many remains largely anonymous (32). A functional consequence of SNPs within the MRE of a miRNA is that they can impart changes in the expression of host genes, either by enhancing or reducing transcript and protein levels. Many groups have described strategies to reveal SNPs within MREs that cause altered gene expression and in some cases, pathophysiological effects. Currently available computational tools include MirSNP (33), a computational approach to identifying SNPs in MREs by combining TargetScan (17) predictions, miRNA seed information, and dbSNP data to compute and predict the occurrences of 'seed-altering' SNPs. However, because TargetScan is based upon MREs in the 3'UTR, and the model was trained with data from a single cell type, not all SNP-MRE interactions will be captured. The PolymiRTS (34) database identifies and annotates SNPs in MREs which have been biologically tested, but again focuses on miRNA binding regions in the 3'UTR. Finally, mrSNP (35) provides a web service for researchers working with RNA-seq data to predict the impact of an SNP in the 3'UTR on miRNA binding. While all these approaches are useful to some degree, we developed an approach to capture SNP-resident MREs in the brain and brain diseases.

The novelty in our approach is three fold. First, we captured >95% of all predicted miRNA-mediated gene regulation events in the brain by focusing exclusively on the top 50 expressed miRNAs. Second, not being constrained by a 3'UTR bias in identifying candidate miRNA-MRE interactions allowed us to investigate physiologically relevant CDS events. Third, we excluded genes with additional distinct MREs to more abundant brain expressed miRNAs, prioritizing the most significant SNP-resident MREs. This was done to maximize the possibility that the altSNP altered activity of the most abundant, and possibly the most relevant miRNA, regulating that gene. Overall, 24 genes were investigated; 14 genes showed significant differences in gene expression levels when the altSNP was introduced in the MRE. These changes corresponded to predicted enhanced or reduced repression. It is possible that those MRE-resident SNPs

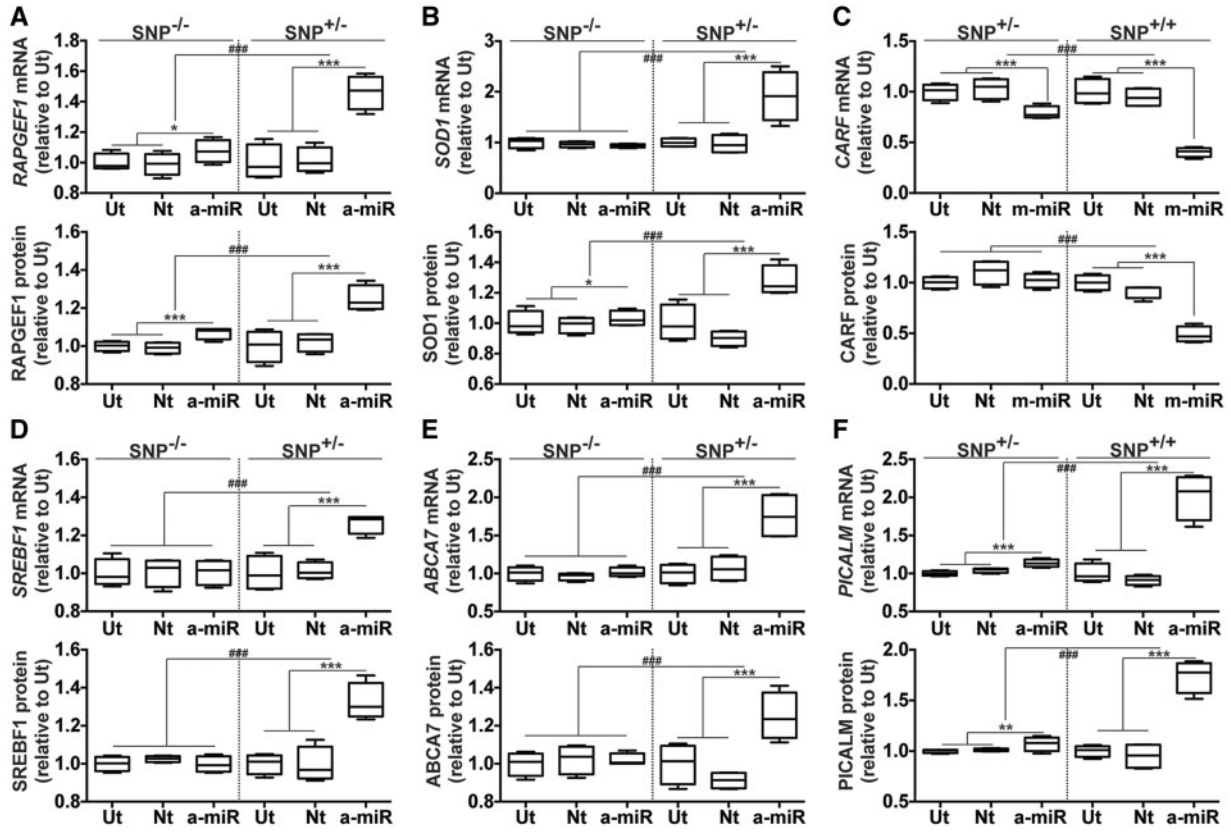

Figure 5. SNPs in MREs alter miRNA-mediated regulation in iPSCs. Three iPSC lines representing the following genotypes: altSNP$^{-/-}$ (homozygous for reference allele), altSNP$^{+/+}$ (homozygous for the alternative candidate SNP), and altSNP$^{+/-}$ (alternative SNP/reference allele) were selected for each gene. iPSC lines were left untreated (Ut) or transfected with the non-targeting (Nt) control or respective miRNA reagents: **(A)** RAPGEF1: anti-miR-124-3p; **(B)** SOD1: anti-miR-139-5p; **(C)** CARF: mimic of miR-181-5p; **(D)** SREBF1: anti-miR-29b-3p; **(E)** ABCA7: anti-miR-34a-5p; **(F)** PICALM: anti-miR-34a-5p + anti-miR-103a-3p. RNA and protein was harvested 48 h post-transfection. Top panel: RT-qPCR for the respective genes (normalized to HPRT), presented as fold change relative to WT ($n = 4$). Bottom panel: Densitometry of immunoblots performed for each gene (shown in Supplementary Material, Fig. S7) presented as fold change relative to Ut ($n = 3$). iPSC lines used (reference Supplementary Material, Fig. S6): *RAPGEF1* (1,4,7,2,13,14); *CARF* (1,6,12,13,2,9); *SOD1* (1,3,6,4,10,12); *SREBF1* (1,8,10,4,5,6); *ABCA7* (4,11,9,1,7,10); *PICALM* (3,7,11,1,5,10). Data are presented as box & whiskers plot (min to max) and analysed by one-way ANOVA with Sidak's multiple comparison test, $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

which did not affect gene expression (3 of 11 in the 3'UTR; 6 of 13 in the CDS) may represent false negatives caused by binding of multiple lower-abundant miRNAs involving MREs distinct from the one queried.

We also asked if synonymous altSNPs causing rare codons could result in significant mRNA repression. We identified three genes whose altSNP-resident mRNAs exhibited increased association with Ago2, higher mRNA decay rates, and significantly lower protein levels. Finally, while the use of exogenous expression systems in HEK293 cells is widely used to test effectiveness of miRNA-mediated target repression, using a non-native promoter to drive gene expression of the target could affect the impact of the altSNP. To validate our findings, we screened 14 iPSC lines for the altSNPs that altered miRNA binding, and show a similar impact.

Cis-acting regulatory SNPs can induce allelic imbalance in gene expression (36–40). While dramatic shifts in gene expression cause disease, as in the case of disease-associated mutations, allelic imbalances cause minor shifts in gene expression, whose consequences can accumulate with age. This can have a dramatic impact, especially on the pathogenesis and pathophysiology of neurologic diseases (41–44). This consequence was recently demonstrated with regard to the pathogenesis of Parkinson's disease (PD), where cis-acting SNPs either increased or decreased α-synuclein expression in human neurons, thereby increasing or decreasing PD risk respectively (45).

AltSNP-resident MREs could therefore result, over time, in the accumulation of allelic imbalances and alter disease pathophysiology for the better or worse. Calcium response factor (*CARF*) is a transcription factor that regulates BDNF expression in a calcium- and neuron-selective manner (46,47). Strong evidence point to deficits in BDNF signalling as a contributing factor to the pathogenesis of several major diseases and disorders such as Huntington's disease, Alzheimer's disease (AD), amyotrophic lateral sclerosis (ALS), and depression (48,49). The SNP rs62183749 increased CARF expression, possibly rendering a protective influence on disease pathogenesis. Superoxide dismutase 1 (*SOD1*) is important in apoptotic signalling and oxidative stress (50). Over 150 different toxic gain-of-function mutations in *SOD1* link it to familial ALS (51). Of interest however is that wild type *SOD1*, under conditions of cellular stress, is implicated in ~90% of sporadic ALS cases (52). While the exact mechanism of how aberrant wild type SOD1 levels contribute to ALS is unknown, it lends an interesting perspective on how the SNP rs41391245 could influence ALS pathogenesis by reducing SOD1 expression. SNPs in the ATP-binding cassette transporter A7 (*ABCA7*) were recently identified to increase AD risk (53). Decreased expression or loss of ABCA7 levels doubled insoluble

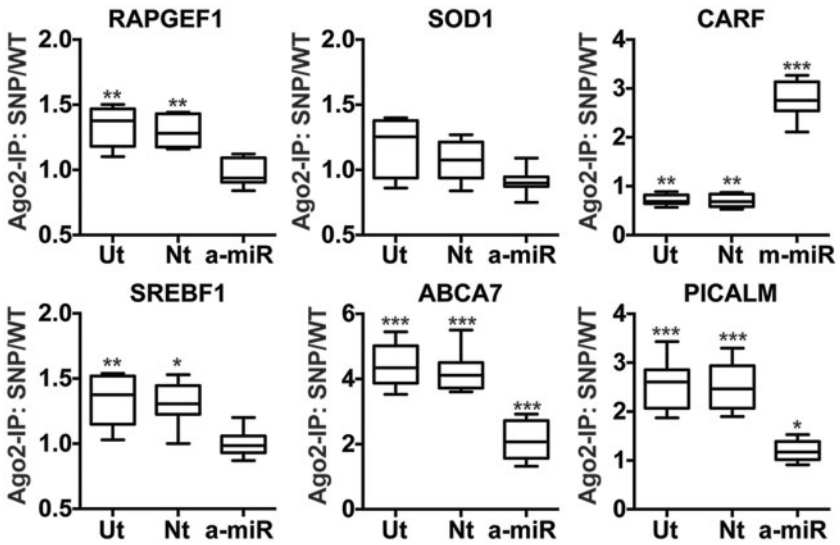

Figure 6. SNPs in the MRE alter miRNA binding in iPSCs. Three altSNP$^{+/-}$ iPSC lines for each of the 6 genes were left untreated (Ut) or transfected with the non-targeting (Nt) control or respective miRNA reagents (RAPGEF1:anti-miR-124-3p, SOD1:anti-miR-139-5p, CARF: mimic of miR-181-5p, SREBF1:anti-miR-29b-3p, ABCA7:anti-miR-34a-5p, PICALM: anti-miR-34a-5p + anti-miR-103a-3p). Ago2 immunoprecipitation (anti-Ago2 Ab) and associated RNA was harvested 24 h later. RT-qPCR using allele-specific primers (Ref vs Alt) for respective genes (normalized to HPRT and total gene expression Ref + Alt) was performed. iPSC lines used (reference **Figure S6**): *RAPGEF1* (2,13,14); *CARF* (1,6,12); *SOD1* (4,10,12); *SREBF1* (4,5,6); *ABCA7* (1,7,10); *PICALM* (3,7,11). Data are presented as the ratio of SNP/WT ($n = 12$). Data are presented as box & whiskers plot (min to max) and analysed by one-way ANOVA with Sidak's multiple comparison test, $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

Aβ levels and amyloid plaques in the brain (54). Furthermore, recent studies suggest that AD risk associated SNPs in *ABCA7* may also contribute to increased PD risk (55). Similarly, the phosphatidylinositol-binding clathrin assembly protein (*PICALM*) gene has also been identified as a susceptibility locus for late-onset AD (LOAD) incidence (56). *PICALM* plays a key role in mediating the clearance of Aβ at the blood-brain barrier, as well as mitigating Aβ toxicity in neurons (57). Early evidence suggests that increased PICALM expression may reduce AD risk (58). The SNPs rs4147914 and rs592297 dramatically reduce ABCA7 and PICALM expression respectively, possibly conferring a disease risk.

It is noteworthy that GWAS-SNPs were not prioritized in our search to identify functional SNP-resident MREs. However, of the 17 genes that showed altSNP-induced changes in miRNA-mediated regulation, 5 had GWAS associated hits. These include rs3739497 (*RAPGEF1*, associated with glycemic traits and epilepsy prognosis, GWAS Central Identifier: HGVM2257467), rs8113500 (*TUBB4A*, associated with Parkinson's disease and age-related macular degeneration, HGVM8374170), rs8682 (*RAB5A*, associated with Parkinson's disease, stroke and age-related macular degeneration, HGVM15436), rs17813964 (*MARCH7*, associated with psoriasis and movement-related adverse antipsychotic effects, HGVM11362713), and rs592297 (PICALM, associated with pulmonary function and epilepsy prognosis, HGVM379360).

Allele-specific differences in gene expression owing to cis-regulatory SNPs highlight the potential impact of SNPs on complex neurological disorders. We describe an experimental approach to identify functionally relevant SNP-resident MREs. Creating better predictors of SNPs in miRNA-MRE interactions will allow us to enhance our understanding of various disease mechanisms. Importantly, our experimental paradigm is applicable to any tissue of interest and can guide additional mechanistic studies to determine molecular consequences of risk alleles.

Materials and Methods

Expression constructs

Primary miRNAs were ordered as gBlocks® gene fragments from Integrated DNA Technologies (IDT), and cloned downstream of a U6 promoter by restriction digestion. Perfect targets and non-targeting controls for each miRNA were designed based on the mature miRNA sequence. Sequences were ordered as gBlocks® gene fragments from Integrated DNA Technologies (IDT), and cloned into a psiCHECK-2 vector by Gibson assembly. Complete 3'UTR and CDS of candidate genes were amplified by PCR from cDNA, generated from total RNA harvested from donor human brain samples. This was done to ensure brain expressed isoforms were being studied. The 3'UTR gene products were cloned into psiCHECK-2 vectors by Gibson assembly. The CDS genes were cloned downstream of a CMV promotor in a mammalian expression construct by Gibson assembly, along with a Flag and 6XHis tag on the C-terminal end. We used site-directed mutagenesis by overlap-extension PCR to generate altSNPs and Scr variants of each gene.

Transfection into HEK293 and induced pluripotent stem cells (iPSCs)

HEK293 cells (24 well plates) and iPSCs (6 well plates) were transfected with lipofectamine® 2000 (Thermo Fisher Scientific) using the manufacturers recommended protocols. Plasmid co-transfections were performed using molar concentrations of plasmids. Anti-miRs and miR-mimics were transfected RNAiMAX™ Reagent (Thermo Fisher) to a final concentration of 10nM using the manufacturers recommended protocols. Stock anti-miRs against miR-124-3p, miR-139-5p, miR-29b-3p, miR-34a-5p, and miR-103a-3p were procured from IDT. Stock non-targeting single stranded and double stranded control

oligonucleotides were ordered from IDT. A modified mimic of miR-181-5p was ordered from IDT as a duplex:

Mature Mimic Strand:/5PHOS/rArArUrArUrUrCrArArCrGrCrUrGrUrCrGrGrUrGmAmGrU

Complementary Mimic Strand: mUmCrAmCrCmGrAmCrAmGrCmGrUmUrGmArAmUrAmUT

Luciferase assay

Protein was harvested from HEK293 cells 48 h post-transfection. We used the Dual-Luciferase Reporter Assay (Promega) and followed the manufacture's recommended protocol.

RNA isolation

Total RNA from HEK293 cells and Embryoid bodies (generated from iPSCs) was isolated using the *mir*Vana™ miRNA isolation kit (ThermoFisher Scientific) or the TRIzol® Reagent (Life Technologies), according to the manufacturer's protocol. Plasmid and genomic DNA was degraded using DNase I treatment (New England Biolabs). Total RNA was tested for quality on an Agilent Model 2100 Bioanalyzer (Agilent Technologies).

Quantitative RT-PCR (RT-qPCR)

First-strand cDNA was synthesized using MultiScribe™ Reverse Transcriptase (Thermo Fisher Scientific) with oligo-dT or random-hexamer priming. Sequence and cDNA specific primers were designed for SYBR® green assays and were procured for the following human genes from IDT: *RAPGEF1, SOD1, CARF, TUBB4A, FBXO45, RPRD1A, PDE8B, RAB5A, PPIP5K2, UNC80, SSB, SREBF1, TUBA1A, HSPD1, MAN2A2, MARCH7, SCRAB1, WBSCR17, PPP3R1, SYN1, SYT1, FKBP4, GTDC1, NDUFV2, ABCA7, PICALM, HMGCS1, CNTN6, GRIN2A, SORL1, GAPDH, HPRT, RENILLA*, and *FIREFLY*. All reactions were setup using Power SYBR® Universal PCR Master Mix (Thermo Fisher Scientific) and run on the BioRad CFX96™ Real-Time PCR detection system. All experiments were performed in quadruplicate.

Allele-specific RT-qPCR

RNA was harvested, DNase treated, first-strand cDNA synthesis performed as described above. Primers were designed for each gene to capture total gene expression levels (Ref + altSNP). Allele-specific forward primers were made according to Gaudet *et al*. (59). A common reverse primer and the allele-specific forward primer were used for allele-specific quantitation (normalized to total). All reactions were in Power SYBR® Universal PCR Master Mix (Thermo Fisher Scientific) and run on the BioRad CFX96™ Real-Time PCR detection system. All experiments were performed in quadruplicate.

SDS-PAGE and Immunoblotting

HEK293 cells or Embryoid bodies (generated from iPSCs) were washed with PBS and lysed in freshly prepared lysis buffer (1% Triton, 25mM Tris pH 7.4, 150mM NaCl, protease inhibitors (cOmplete™, mini, EDTA-free, Roche)) for 30 min at 4°C. The lysates were centrifuged at 14,000 rpm for 20 min at 4°C, and the supernatant quantified by BCA Protein Assay kit (Pierce). Protein was denatured in 6X-Sample SDS buffer (375 mM Tris-HCl pH 6.8, 6% SDS, 48% glycerol, 9% 2-Mercaptoethanol, and 0.03% bromophenol blue). 20 μg of protein per lane was separated on a 4-15% SDS-PAGE gel for western blot analysis. Protein abundance was quantified by densitometry using an ChemiDoc MP Imaging System (Bio-Rad). Western blots were probed, stripped and re-probed as follows. PVDF membranes were first probed with the antibody against the gene of interest. After imaging, the PVDF membrane was stripped with Restore Western Blot Stripping Buffer (Thermo Scientific) for 15 min, washed in Tris Buffered Saline-Tween (TBS-T) and blocked in 5% Bovine Serum Albumin (BSA, Pierce) for 1 h. The membrane was washed in TBS-T and incubated with the goat anti-mouse or goat-anti-rabbit secondary antibody (1:10000, Sigma) for 1 h and imaged. If a signal was detected, the stripping procedure was repeated until no signal was observed. The membrane was washed in TBS-T, blocked for 1 h in 5% BSA and re-probed with the antibody against tubulin or GAPDH. Antibodies used in this study were procured from the following sources: Flag-M2 (F3165, Sigma); GAPDH (sc32233, SantaCruz); b-tubulin (sc5274, SantaCruz); RAPGEF1 (A301-965A, Bethyl Labs); SOD1 (A303-811A, Bethyl Labs); SREBF1 (ab3259, Abcam); ABCA7 (A304-427A, Bethyl Labs); CARF (A303-861A, Bethyl Labs); PICALM (ab172962, Abcam); His (ab9108, Abcam); Argonaute 2 (ab57113, Abcam).

Immunoprecipitation

Immunoprecipitation (IP) experiments were performed in HEK293 cells and $SNP^{+/-}$ iPS cells. To IP Argonaute 2, cells were lysed (as described above), and the supernatant (20–50 μg of protein) was incubated with anti-Ago2 Ab overnight at 4°C, followed by incubation with protein G-agarose (Invitrogen) for 1 h at 4°C. Beads were pelleted at 2,500 rpm for 30 sec, supernatant removed, and beads were resuspended in 500 mL RIP buffer. Three washes were repeated with RIP buffer, followed by one wash in PBS. The pellet was resuspended in TRIzol® Reagent (Life Technologies) (1 mL) and RNA harvested according to manufacturer's instructions. RNA was eluted with nuclease-free water (20 μL). RIP buffer: 150 mM KCl, 25 mM Tris pH 7.4, 5 mM EDTA, 0.5 mM DTT, 0.5% NP40, 100 U/ml RNAase inhibitor (added fresh each time), protease inhibitors (added fresh each time).

6XHis-trap anti-FLAG ELISA

Cells (grown in 24 well plates) were lysed in lysis buffer (50 mM Tris-Hcl, 1M NaCl, 10 mM TCEP, 25 mM Imidazole, protease inhibitors) and processed as described above. Total protein (50 ug) was added to a single well on the Pierce™ Nickel Coated Plates (Thermo Scientific) and incubated for 4–6 h at 4C. Each well was then gently washed 3 times with wash buffer (50 mM Tris-Hcl, 0.1M NaCl, 10 mM TCEP, 50 mM Imidazole). Each well was incubated with the anti-FLAG-M2-peroxidase antibody for 1 h at 4C. After washed each well 6 times with ice-cold PBS, the Amplex® UltraRed Reagent (Molecular Probes) assay solution was added (200 ul/well). Plates were incubated for 10-20 min at room temperature (in dark). 200 ul of the reaction was transferred to a 96well BLACK plate and read on the SpectraMax Minimax 300 Imaging Cytometer (Molecular Devices).

iPSC maintenance and orbital embryoid body generation

IPSC lines are stock control lines maintained by the Human Pluripotent Stem Cell Core at the Raymond G. Perelman Center for Cellular and Molecular Therapeutics, The Children's Hospital of Philadelphia. Protocols to maintain and expand iPSCs are described by Sullivan *et al*. (60). After transfection,

iPSC Line	Cell origin	Sex/Age (yrs)	iPSC Line	Cell origin	Sex/Age (yrs)
Line 1	Fibroblast	Female, 2	Line 8	PBMC	Male, Adult
Line 2	Fibroblast	Male, 3	Line 9	PBMC	Female, Adult
Line 3	Fibroblast	Female, 5	Line 10	PBMC	Male, Adult
Line 4	Fibroblast	Male, 2	Line 11	LCL	Male, 26
Line 5	Cord Blood	Female	Line 12	LCL	Male, Adult
Line 6	Bone Marrow	Male, Adult	Line 13	Fibroblast	Male, newborn
Line 7	PBMC	Male, Adult	Line 14	Fibroblast	Male, newborn

LCL: Lymphoblastoid cell line, PBMC: Primary peripheral blood mononuclear cells.

orbital embryoid bodies were generated from iPSCs to prevent contaminating DNA/RNA from mouse embryonic fibroblasts. Protocols to generate the embryoid bodies from iPSCs are described by Chen *et al.* (61).

Bioinformatics Methods

Experimental and analytical details for the human brain AGO2 HITS-CLIP data were reported previously (19). In brief, human brain motor cortex (BA4) and striatum tissue samples were obtained from 11 individuals and lysed. Lysates were UV-irradiated to crosslink RNA-protein complexes. AGO2 RNP complexes were enriched by IP, and Illumina libraries were prepared from the purified AGO2-associated RNAs (both miRNAs and bound target sites). Libraries were sequenced on the Illumina HiSeq 2500 platform. FASTQ files for each library were trimmed of adapter sequence and then mapped to the human genome (hg19) using Bowtie1, requiring unique mapping. To avoid false-positive peak calls resulting from PCR amplification error, alignments were collapsed to unique chromosome:start:end:strand coordinates. *P*-values were calculated based on a gene-wise zero truncated negative binomial (ZTNB) model. Fisher's method was used to combine the resulting positional p-values across all samples, and adjacent p-values < 0.05 were clustered. We manually examined each cluster to ensure the SNP resides within the MRE seed. In some cases, however, the interacting miRNA exhibited strong 3' complementarity with the MRE. In such an instance, we chose to include these interactions as the SNPs still held the potential to disrupt miRNA binding and miRNA mediated regulation.

SNP annotations were obtained from dbSNP138 filtered to contain brain expressed genes (9,310 genes)(62), and downloaded from the UCSC Genome Browser (hg19), using the Table Browser tool. Additional information for coding SNPs, including codon and amino acid changes and identities of altered proteins, was also obtained from the UCSC Table Browser, using the snp138CodingDbSnp linked table. Coding SNP overlap with AGO2 clusters was based on matching Gene IDs. We used the gene ontology tool DAVID (23) for functional annotation based clustering of genes, while the computational tool RNA22 V2 (24) was used to identify novel miRNA binding sites on transcripts.

Statistical Analysis

All statistical analysis was done using GraphPad PRISM V6.0h. Normality of data was tested using the D'Agostino & Pearson omnibus normality test. One-way ANOVA with Sidak's multiple comparison test was used to determine significance. Two-way ANOVA with Sidak's multiple comparison test was used to test for significant interaction.

The statistics used in individual experiments along with method of presenting data and error bars are described in the figure legends.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We thank Carolyn Yrigollen, Pavitra Ramachandran and Deborah French for critical comments on the manuscript. We would also like to thank Jean Ann Maguire, Liang 'Grace' Ge, Alejandro Mas Monteys, Yong Hong Chen, Luis Tecedor, Gregory Cajka and Kasey L Brida for providing essential reagents and for technical assistance.

Conflict of Interest statement. B.L.D. is a founder of Spark Therapeutics, Inc., a gene therapy company.

Funding

The authors also acknowledge the support of the Human Pluripotent Stem Cell Core, The Children's Hospital of Philadelphia. This work was supported by the National Institutes of Health [R01 NS076631 to B.L.D.]. Funding to pay the Open Access publication charges for this article was provided by Dr Davidson's grant funding.

References

1. Schaub, M.A., Boyle, A.P., Kundaje, A., Batzoglou, S. and Snyder, M. (2012) Linking disease associations with regulatory information in the human genome. *Genome Res.*, **22**, 1748–1759.
2. Johnson, A.D., Handsaker, R.E., Pulit, S.L., Nizzari, M.M., O'Donnell, C.J. and de Bakker, P.I. (2008) SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*, **24**, 2938–2939.
3. Kim, S., Cho, H., Lee, D. and Webster, M.J. (2012) Association between SNPs and gene expression in multiple regions of the human brain. *Transl Psychiatry*, **2**, e113.
4. Vosa, U., Esko, T., Kasela, S. and Annilo, T. (2015) Altered Gene Expression Associated with microRNA Binding Site Polymorphisms. *PLoS One*, **10**, e0141351.
5. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., *et al.* (2009) Finding the missing heritability of complex diseases. *Nature*, **461**, 747–753.

6. Goldstein, D.B. (2009) Common genetic variation and human traits. *N. Engl J. Med.*, **360**, 1696–1698.
7. Friedman, R.C., Farh, K.K., Burge, C.B. and Bartel, D.P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.*, **19**, 92–105.
8. Minguzzi, S., Selcuklu, S.D., Spillane, C. and Parle-McDermott, A. (2014) An NTD-associated polymorphism in the 3' UTR of MTHFD1L can affect disease risk by altering miRNA binding. *Hum. Mutat.*, **35**, 96–104.
9. Jensen, K.P., Kranzler, H.R., Stein, M.B. and Gelernter, J. (2014) The effects of a MAP2K5 microRNA target site SNP on risk for anxiety and depressive disorders. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **165B**, 175–183.
10. Ronshaugen, M., Biemar, F., Piel, J., Levine, M. and Lai, E.C. (2005) The Drosophila microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. *Genes Dev.*, **19**, 2947–2952.
11. Kuwabara, T., Hsieh, J., Nakashima, K., Taira, K. and Gage, F.H. (2004) A small modulatory dsRNA specifies the fate of adult neural stem cells. *Cell*, **116**, 779–793.
12. Poole, R.J. and Hobert, O. (2006) Early embryonic programming of neuronal left/right asymmetry in C. elegans. *Curr. Biol.*, **16**, 2279–2292.
13. Schratt, G.M., Tuebing, F., Nigh, E.A., Kane, C.G., Sabatini, M.E., Kiebler, M. and Greenberg, M.E. (2006) A brain-specific microRNA regulates dendritic spine development. *Nature*, **439**, 283–289.
14. Xu, Y., Li, F., Zhang, B., Zhang, K., Zhang, F., Huang, X., Sun, N., Ren, Y., Sui, M. and Liu, P. (2010) MicroRNAs and target site screening reveals a pre-microRNA-30e variant associated with schizophrenia. *Schizophr. Res.*, **119**, 219–227.
15. Xu, Y., Liu, H., Li, F., Sun, N., Ren, Y., Liu, Z., Cao, X., Wang, Y., Liu, P. and Zhang, K. (2010) A polymorphism in the microRNA-30e precursor associated with major depressive disorder risk and P300 waveform. *J. Affect. Disord.*, **127**, 332–336.
16. Shafi, G., Aliya, N. and Munshi, A. (2010) MicroRNA signatures in neurological disorders. *Can. J. Neurol. Sci.*, **37**, 177–185.
17. Agarwal, V., Bell, G.W., Nam, J.W. and Bartel, D.P. (2015) Predicting effective microRNA target sites in mammalian mRNAs. *Elife*, **4**.
18. Moore, M.J., Zhang, C., Gantman, E.C., Mele, A., Darnell, J.C.,. and Darnell, R.B. (2014) Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. *Nat. Protoc.*, **9**, 263–293.
19. Boudreau, R.L., Jiang, P., Gilmore, B.L., Spengler, R.M., Tirabassi, R., Nelson, J.A., Ross, C.A., Xing, Y. and Davidson, B.L. (2014) Transcriptome-wide discovery of microRNA binding sites in human brain. *Neuron*, **81**, 294–305.
20. Guojing, L., Zhang, R., Xu, J., Wu, C.I. and Lu, X. (2015) Functional Conservation of Both CDS- and 3'-UTR-Located MicroRNA Binding Sites between Species. *Mol. Biol. Evol.*, **32**, 3276.
21. Hausser, J., Syed, A.P., Bilen, B. and Zavolan, M. (2013) Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. *Genome Res.*, **23**, 604–615.
22. Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M. and Sirotkin, K. (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.*, **29**, 308–311.
23. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
24. Miranda, K.C., Huynh, T., Tay, Y., Ang, Y.S., Tam, W.L., Thomson, A.M., Lim, B. and Rigoutsos, I. (2006) A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell*, **126**, 1203–1217.
25. Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M., *et al.* (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*, **141**, 129–141.
26. Gu, S., Jin, L., Zhang, F., Sarnow, P. and Kay, M.A. (2009) Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs. *Nat. Struct. Mol. Biol.*, **16**, 144–150.
27. Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P. and Burge, C.B. (2003) Prediction of mammalian microRNA targets. *Cell*, **115**, 787–798.
28. Bartel, D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*, **136**, 215–233.
29. Shkumatava, A., Stark, A., Sive, H. and Bartel, D.P. (2009) Coherent but overlapping expression of microRNAs and their targets during vertebrate development. *Genes Dev.*, **23**, 466–481.
30. Storey, J.D., Madeoy, J., Strout, J.L., Wurfel, M., Ronald, J. and Akey, J.M. (2007) Gene-expression variation within and among human populations. *Am. J. Hum. Genet.*, **80**, 502–509.
31. Stranger, B.E., Nica, A.C., Forrest, M.S., Dimas, A., Bird, C.P., Beazley, C., Ingle, C.E., Dunning, M., Flicek, P., Koller, D., *et al.* (2007) Population genomics of human gene expression. *Nat. Genet.*, **39**, 1217–1224.
32. Issler, O. and Chen, A. (2015) Determining the role of microRNAs in psychiatric disorders. *Nat. Rev. Neurosci.*, **16**, 201–212.
33. Liu, C., Zhang, F., Li, T., Lu, M., Wang, L., Yue, W.,. and Zhang, D. (2012) MirSNP, a database of polymorphisms altering miRNA target sites, identifies miRNA-related SNPs in GWAS SNPs and eQTLs. *BMC Genomics*, **13**, 661.
34. Bhattacharya, A., Ziebarth, J.D. and Cui, Y. (2014) PolymiRTS Database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways. *Nucleic Acids Res.*, **42**, D86–D91.
35. Deveci, M., Catalyurek, U.V. and Toland, A.E. (2014) mrSNP: software to detect SNP effects on microRNA binding. *BMC Bioinformatics*, **15**, 73.
36. Almlof, J.C., Lundmark, P., Lundmark, A., Ge, B., Maouche, S., Goring, H.H., Liljedahl, U., Enstrom, C., Brocheton, J., Proust, C., *et al.* (2012) Powerful identification of cis-regulatory SNPs in human primary monocytes using allele-specific gene expression. *PLoS One*, **7**, e52260.
37. Bryois, J., Buil, A., Evans, D.M., Kemp, J.P., Montgomery, S.B., Conrad, D.F., Ho, K.M., Ring, S., Hurles, M., Deloukas, P., *et al.* (2014) Cis and trans effects of human genomic variants on gene expression. *PLoS Genet.*, **10**, e1004461.
38. GuhaThakurta, D., Xie, T., Anand, M., Edwards, S.W., Li, G., Wang, S.S. and Schadt, E.E. (2006) Cis-regulatory variations: a study of SNPs around genes showing cis-linkage in segregating mouse populations. *BMC Genomics*, **7**, 235.
39. Hu, P., Lan, H., Xu, W., Beyene, J. and Greenwood, C.M. (2007) Identifying cis- and trans-acting single-nucleotide polymorphisms controlling lymphocyte gene expression in humans. *BMC Proc.*, **1 Suppl 1**, S7.
40. Morley, M., Molony, C.M., Weber, T.M., Devlin, J.L., Ewens, K.G., Spielman, R.S. and Cheung, V.G. (2004) Genetic analysis

of genome-wide variation in human gene expression. *Nature*, **430**, 743–747.
41. Cooper-Knock, J., Kirby, J., Ferraiuolo, L., Heath, P.R., Rattray, M. and Shaw, P.J. (2012) Gene expression profiling in human neurodegenerative disease. *Nat. Rev. Neurol.*, **8**, 518–530.
42. Genetic Modifiers of Huntington's Disease, C. (2015) Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. *Cell*, **162**, 516–526.
43. Kearney, J.A. (2011) Genetic modifiers of neurological disease. *Curr. Opin. Genet. Dev.*, **21**, 349–353.
44. Li, J.L., Hayden, M.R., Warby, S.C., Durr, A., Morrison, P.J., Nance, M., Ross, C.A., Margolis, R.L., Rosenblatt, A., Squitieri, F., *et al.* (2006) Genome-wide significance for a modifier of age at neurological onset in Huntington's disease at 6q23-24: the HD MAPS study. *BMC Med. Genet.*, **7**, 71.
45. Soldner, F., Stelzer, Y., Shivalila, C.S., Abraham, B.J., Latourelle, J.C., Barrasa, M.I., Goldmann, J., Myers, R.H., Young, R.A. and Jaenisch, R. (2016) Parkinson-associated risk variant in distal enhancer of alpha-synuclein modulates target gene expression. *Nature*, **533**, 95–99.
46. Tao, X., West, A.E., Chen, W.G., Corfas, G. and Greenberg, M.E. (2002) A calcium-responsive transcription factor, CaRF, that regulates neuronal activity-dependent expression of BDNF. *Neuron*, **33**, 383–395.
47. McDowell, K.A., Hutchinson, A.N., Wong-Goodrich, S.J., Presby, M.M., Su, D., Rodriguiz, R.M., Law, K.C., Williams, C.L., Wetsel, W.C. and West, A.E. (2010) Reduced cortical BDNF expression and aberrant memory in Carf knock-out mice. *J. Neurosci.*, **30**, 7453–7465.
48. Wijesekera, L.C. and Leigh, P.N. (2009) Amyotrophic lateral sclerosis. *Orphanet. J. Rare Dis.*, **4**, 3.
49. Lu, B., Nagappan, G.,. and Lu, Y. (2014) BDNF and synaptic plasticity, cognitive function, and dysfunction. *Handb. Exp. Pharmacol.*, **220**, 223–250.
50. Danial, N.N. and Korsmeyer, S.J. (2004) Cell death: critical control points. *Cell*, **116**, 205–219.
51. Al-Chalabi, A. and Leigh, P.N. (2000) Recent advances in amyotrophic lateral sclerosis. *Curr. Opin. Neurol.*, **13**, 397–405.
52. Gagliardi, S., Cova, E., Davin, A., Guareschi, S., Abel, K., Alvisi, E., Laforenza, U., Ghidoni, R., Cashman, J.R., Ceroni, M., *et al.* (2010) SOD1 mRNA expression in sporadic amyotrophic lateral sclerosis. *Neurobiol. Dis.*, **39**, 198–203.
53. Steinberg, S., Stefansson, H., Jonsson, T., Johannsdottir, H., Ingason, A., Helgason, H., Sulem, P., Magnusson, O.T., Gudjonsson, S.A., Unnsteinsdottir, U., *et al.* (2015) Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. *Nat. Genet.*, **47**, 445–447.
54. Li, H., Karl, T. and Garner, B. (2015) Understanding the function of ABCA7 in Alzheimer's disease. *Biochem. Soc. Trans.*, **43**, 920–923.
55. Nuytemans, K., Maldonado, L., Ali, A., John-Williams, K., Beecham, G.W., Martin, E., Scott, W.K. and Vance, J.M. (2016) Overlap between Parkinson disease and Alzheimer disease in ABCA7 functional variants. *Neurol. Genet.*, **2**, e44.
56. Xu, W., Tan, L. and Yu, J.T. (2015) The Role of PICALM in Alzheimer's Disease. *Mol. Neurobiol.*, **52**, 399–413.
57. Zhao, Z., Sagare, A.P., Ma, Q., Halliday, M.R., Kong, P., Kisler, K., Winkler, E.A., Ramanathan, A., Kanekiyo, T., Bu, G., *et al.* (2015) Central role for PICALM in amyloid-beta blood-brain barrier transcytosis and clearance. *Nat. Neurosci.*, **18**, 978–987.
58. Parikh, I., Fardo, D.W. and Estus, S. (2014) Genetics of PICALM expression and Alzheimer's disease. *PLoS One*, **9**, e91242.
59. Gaudet, M., Fara, A.G., Beritognolo, I. and Sabatti, M. (2009) Allele-specific PCR in SNP genotyping. *Methods Mol. Biol.*, **578**, 415–424.
60. Sullivan, S.K., Mills, J.A., Koukouritaki, S.B., Vo, K.K., Lyde, R.B., Paluru, P., Zhao, G., Zhai, L., Sullivan, L.M., Wang, Y., *et al.* (2014) High-level transgene expression in induced pluripotent stem cell-derived megakaryocytes: correction of Glanzmann thrombasthenia. *Blood*, **123**, 753–757.
61. Chen, V.C., Couture, S.M., Ye, J., Lin, Z., Hua, G., Huang, H.I., Wu, J., Hsu, D., Carpenter, M.K.,. and Couture, L.A. (2012) Scalable GMP compliant suspension culture system for human ES cells. *Stem Cell Res.*, **8**, 388–402.
62. Lin, L., Park, J.W., Ramachandran, S., Zhang, Y., Tseng, Y.T., Shen, S., Waldvogel, H.J., Curtis, M.A., Faull, R.L., Troncoso, J.C., *et al.* (2016) Transcriptome sequencing reveals aberrant alternative splicing in Huntington's disease. *Hum. Mol. Genet.*, **25**, 3454–3466.