A functional variant in the miR-142 promoter modulating its expression and conferring risk of Alzheimer's disease

M. Ghanbari[1,2], S.T. Munshi[3], B. Ma[4], B. Lendemeijer[3], S. Bansal[3], H.H. Adams[1], W. Wang[4], K.Goth[5], D.E. Slump[3], M. C.G.N van den Hout[6], W.F.J. van IJcken[6], S.Bellusci[5], Q.Pan[4], S.J. Erkeland[7], F.M.S. de Vrij[3], S.A. Kushner[3], M. A.Ikram[1]

1. *Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, 3000 CA Rotterdam, the Netherlands.*
2. *Department of Genetics, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.*
3. *Department of Psychiatry, Erasmus MC, University Medical Center Rotterdam, 3000 CA Rotterdam, the Netherlands.*
4. *Department of Gastroenterology, Erasmus MC, University Medical Center Rotterdam, 3000 CA Rotterdam, the Netherlands.*
5. *Excellence Cluster Cardio-Pulmonary System (ECCPS), University Justus Liebig Giessen, 35392 Giessen, Germany.*
6. *Erasmus MC Center for Biomics, Erasmus MC, University Medical Center Rotterdam, 3000 CA Rotterdam, the Netherlands.*
7. *Department of Immunology, Erasmus MC, University Medical Center Rotterdam, 3000 CA Rotterdam, the Netherlands.*

Published in Human Mutation (2019), in press

ABSTRACT

Non-coding RNAs have been widely recognized as essential mediators of gene regulation. However, in contrast to protein-coding genes, much less is known about the influence of non-coding RNAs on human diseases. Here we examined the association of genetic variants located in primary microRNA sequences and long non-coding RNAs (lncRNAs) with Alzheimer's disease (AD) by leveraging data from the largest genome-wide association meta-analysis of late-onset AD. Variants annotated to five miRNAs and ten lncRNAs (in 7 distinct loci) exceeded the Bonferroni-corrected significance threshold (p-value $<1.02\times10^{-6}$). Among these, a leading variant (rs2526377:A>G) at the 17q22 locus annotated to two non-coding RNAs (*MIR142* and *BZRAP1-AS*) was significantly associated with a reduced risk of AD and fulfilled predefined criteria for being functional. Our functional genomic analyses revealed that rs2526377 affects the promoter activity and decreases the expression of miR-142. Moreover, differential expression analysis by RNA-Seq in human iPSC-derived neural progenitor cells and the hippocampus of miR-142 knockout mice demonstrated multiple target genes of miR-142 in the brain that are likely to be involved in the inflammatory and neurodegenerative manifestations of AD. These include *TGFBR1* and *PICALM*, of which their derepression in the brain due to reduced expression levels of miR-142-3p may reduce risk of AD.

INTRODUCTION

Alzheimer's disease (AD) is the most common age-related neurodegenerative disease worldwide manifested by the progressive loss of memory and cognitive decline[1] . Enormous efforts have been made over the past decades to discover risk factors for developing AD and to identify biomarkers for early diagnosis of the disease[2-4]. The determinants of early-onset AD have been primarily associated with mutations in one of three genes: *APP*, *PSEN1* and *PSEN2*[5]. In contrast, late-onset AD (after 65 years of age), the most common form of AD with a heritability of 60-80%, is a genetically heterogeneous disease[6] . In addition to apolipoprotein E (*APOE*) polymorphisms that explain ~25% of the heritability, more than 30 genetic loci have so far been established as contributing to late-onset AD risk[7,8]. However, they explain only a fraction of the estimated heritability and the genetics of AD are yet to be fully understood[9]. To fully grasp the contribution of genetic factors to AD, we must go beyond classical genetics, and explore the multiple interacting layers that regulate the genome. This includes the analysis of not only the protein-coding sequences, but the vast non-coding regions as well.

Recent developments in omics technologies have revealed the complexity of the human genome, displaying that protein-coding RNAs constitute only ~2% of the human transcriptome, highlighting the distinct possibility that non-coding RNAs (ncRNAs) might meaningfully contribute to human disease[10,11]. Non-coding RNAs are functional RNA molecules that are transcribed from DNA but not translated into proteins. They are frequently categorized, on the basis of transcript size, as small (less than 200 nucleotides (nt)) or long non-coding RNAs (over 200 nt). Among these, microRNAs (miRNAs), with approximately 21-23 nt in length, are currently the best-characterized ncRNAs. Many studies have shown the crucial role of miRNAs in neurodevelopmental regulation and disease-related neuropathology including AD[12,13]. Long non-coding RNAs (lncRNAs) comprise a large and diverse class of transcribed RNA molecules that are classified into different subtypes (e.g., antisense and intergenic) according to the position and direction of transcription with regard to other genes[14]. It has become increasingly evident that lncRNAs impact disease pathogenesis primarily through post-transcriptional regulation of gene expression[15]. Despite constituting the majority of non-coding transcriptome, few lncRNAs most notably *BACE1-AS* and *BC200* have been so far characterized to play a role in the pathogenesis of AD to date[16,17].

In the present study, we conducted a genome-wide scan to identify miRNAs and lncRNAs associated with AD by leveraging data from the largest available GWAS of late-onset AD[7]. We found several ncRNA loci significantly associated with AD, including a newly identified susceptibility locus on 17q22. We performed various *in silico* and *in vitro* studies to determine the functionality of ncRNA variant in this locus and to gain insight into the role of associated ncRNA in AD pathogenesis.

MATERIALS AND METHODS

Genome-wide association study on AD

Summary statistics data were retrieved from a recent large-scale GWAS meta-analysis of late-onset AD including 455,258 individuals of European ancestry, meta-analyzed in three phases[7]. Phase 1 consisted of 24,087 clinically diagnosed late-onset AD cases and 55,058 controls of European ancestry, which are collected by 3 independent consortia (Alzheimer disease working group of the Psychiatric Genomics Consortium (PGC-ALZ), the International Genomics of Alzheimer Project (IGAP), and the Alzheimer Disease Sequencing Project (ADSP)), and investigating 9,862,738 genetic variants. Phase 2 consisted of 376,113 individuals of European ancestry from the UK Biobank with parental AD status available (N proxy cases = 47,793; N proxy controls = 328,320). Phase 3 was the meta-analysis of phase 1 and 2, including 71,880 (proxy) AD cases and 383,378 (proxy) controls. More details about the consortia and participants are described elsewhere[7]. All participating studies in the AD GWAS had provided informed consent for participation in genetics studies and were approved by their local ethical committees.

Genetic variants in non-coding RNAs

Genetic variants in human lncRNA transcripts were extracted using lncRNASNP, a comprehensive database including 495,729 SNPs in 32,108 lncRNA transcripts of 17,436 lncRNAs[18]. Moreover, as primary transcript of miRNAs has been suggested to be 3-4kb in length[19], we used dbSNP database (https://www.ncbi.nlm.nih.gov/SNP/) to extract 16,178 SNPs located in +/-2kb of 1,318 mature miRNA sequences reported in miRBase v21 (http://www.mirbase.org/). We excluded SNPs with minor allele frequency (MAF) < 0.01. Of the remaining SNPs, we analyzed the association with AD of 96,950 SNPs in 14,790 lncRNA transcripts and 12,404 SNPs in 1,237 primary miRNA transcripts that were present in the GWAS summary statistics data[7]. To obtain the number of independent SNPs, we used the LD based SNP pruning in PLINK (http://pngu.mgh.harvard. edu/~purcell/plink/), where we excluded the SNPs with $R^2 > 0.7$. The Bonferroni correction was used to adjust *p*-value for the number of tests (0.05/49,323 independent SNPs) and the significance threshold was set at 1.02×0^{-6}. Regional plots showing the association of ncRNA SNPs and flanking variants in the corresponding loci with AD were generated by the LocusZoom web tool[20].

Assessing biological functionality of non-coding RNA variants

For the ncRNA SNPs associated with AD, the LD region ($R^2 > 0.7$) was determined using the 1000 Genomes Phase 3. We investigated whether known protein-coding variants were in strong LD with the associated ncRNA SNPs. Further, we examined whether the associated SNPs in ncRNAs are annotated to regulatory features, including promoter and enhancer regulatory motifs, DNase footprinting sites and conserved sequences using HaploReg (v4.1)[21].

For each set of variants in strong LD with a given ncRNA SNP, we also investigated whether the SNP was located in a potential regulatory region using the Roadmap consortium reference epigenomes dataset[22]. To test the association of ncRNA SNPs with gene expression, we used expression quantitative trait loci (eQTL) data provided by GTEx (https://www.gtexportal.org/home/) and BBMRI-NL atlas (http://atlas.bbmrirp3-lumc. surf-hosted.nl/#query)[23]. The UCSC genome browser was used for visualization of the ncRNA SNP location in the genome. The ncRNA secondary structure and the effect of a SNP on the minimum free energy (MFE) of the predicted ncRNA structure were investigated using the Vienna RNA Package 2.0[24].

Plasmids, miRNA promoter constructs and cell culture transfection

To compare the activity of miR-142 promoter containing either rs2526377 alleles, the full-length 589-bp fragment corresponding to the upstream region of the pri-miR-142 transcript was synthesized by Integrated DNA Technologies (IDT) and cloned into pGreenFire-CMV-EF1-puro (System Biosciences) digested with *EcoR I* and *Spe I*. DNA sequencing verified all constructs. HEK293 cells were then used to generate the lentivirus with co-transfection of reporter gene vectors, HIV gag-pol and VSV-G in a ratio of 1:0.8:0.2. For transduction assay, cells were seeded into 24-well plates and transduced with lentiviral particles. With selection by puromycin at a concentration of 2 ug/ml, cells were calculated and seeded into 96-well. After incubation for 24-hours, the cell supernatant was harvested and the luciferase activity was then measured on a luminometer (LB960; Berthold) using the Dual-Luciferase Reporter Assay System (Promega). The ratio of firefly luciferase to Renilla luciferase was calculated for each well. The experiments were performed five times.

Quantitative RT-PCR

Total RNA from human induced pluripotent stem cell (iPSC)-derived neural progenitor cells (NPCs) and human brain cryopreserved sections was isolated using Trizol LS reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocols. The concentration of total RNA was determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop, Wilmington, DE). TaqMan qPCR Assays were performed according to the manufacturer's protocols (Applied Biosystems, Foster City, CA, USA) to determine the expression levels of miR-142-3p, miR-142-5p, and *BZRAP1-AS1*. The assays were run using Applied Biosystems 7900HT Real-Time PCR system. RNU6B was used as an internal control for miRNA expression analysis. All the experiments were performed in triplicates. The human frozen brain tissues (n=3 gray matter and n=3 white matter) were obtained from the Netherlands Brain Bank (Amsterdam, the Netherlands). All samples were free of neurological disease.

Putative target genes of miR-142

TargetScan V7.1 (http://www.targetscan.org/)[25] was used to identify the putative targets of miR-142-3p and -5p in human and mouse. This program predicts biological targets of miR-

NAs by searching for conserved 7/8-mer sites that match the miRNA seed region. The predictions are ranked based on the putative efficacy of targeting as calculated using context scores of the sites, the higher context score, the greater the probability that a miRNA could target a particular gene. For our analysis, we used the predicted targets that had a recommended context score < -0.01. Further, we retrieved the list of putative target genes of miR-142 (3p and 5p) from two other widely used online miRNA target prediction databases, miRtarget2[26] and DIANA-microT[27]. Then, RNA-Seq data from the Human Body Map 2.0[28] was used to check which of the miR-142 putative target genes are expressed in the human brain (Fragments Per Kilobase Million, FPKM ≥ 1), target genes not expressed in the brain were excluded.

Pathway analysis was performed using KEGG and IPA databases. KEGG incorporates knowledge of known gene networks and identifies significantly enrichment of miRNA putative targets in these networks according to a t-test[29]. IPA is a knowledge database generated from peer-reviewed scientific publications that enables the discovery of highly represented biological mechanisms, pathways or functions most relevant to the genes of interest from large, quantitative datasets. We uploaded the list of the miR-142 target genes and performed a core analysis with the default settings in IPA. We mapped the miRNA target genes to biological functions or canonical pathways to see whether they are enriched in specific networks. The p-values are calculated using the right-tailed Fisher Exact Test and a p-value of less than 0.05 indicates a statistically significant, nonrandom association.

RNA-Seq analysis in human iPSC-derived neural progenitor cells (NPCs)

NPCs derived from human control iPSCs (Sigma-Aldrich line iPSC0028) were cultured to 70% confluency in 6-well plates (Corning) according to standard protocols[30]. NPCs were transfected with 10 nM miRNA mimics (mirVanaTM Mimics, Thermo Fisher Scientific) including miR-142-3p, miR-142-5p, and the standard negative control #1 (Catalog nr. 4464060), or without any mimic (untreated). Transfections were performed using X-treme GENE^{TM} transfection reagent (Merck) according to manufacturer's instructions. The experiment was run in triplicate. Total RNA was isolated 72 hours after transfection from the four groups of NPC samples using the RNeasy mini kit (Qiagen, 74104). The RNA quality was checked by Agilent's 2100 Bioanalyzer (using Eukaryote Total RNA Nano kit). RNA-Seq analysis was performed at Erasmus MC Center for Biomics to test the changes in gene expression pattern in NPCs after overexpression of either miR-142-3p or -5p compared to controls.

RNA-Seq analysis in the hippocampus of miR-142 KO mice and Wt littermates

MiR-142 in mice is located on chromosome 11 and in the vicinity of the second exon belonging to Mir142hg (ENSMUSG00000084796). The miR-142$^{-/-}$ knockout mouse is a model with complete deletion of miR-142, with a significant decrease in the expression levels of both miR-142-3p and -5p isoforms[31]. The expression of *Bzrap1*, a gene immediately flanking miR-142 is not altered in the miR-142-null mice, while the expression of a long non-coding RNA

(Mir142hg) embedded within the miR-142 gene is decreased. We confirmed the homo- and heterozygosity of miR-142 KO mice by qPCR. The hippocampi of miR-142$^{-/-}$, miR-142$^{-/+}$ and miR-142$^{+/+}$ littermates (n=4 for each group, age 13-14 weeks, gender balanced in all groups) were collected and total RNA was extracted using the RNeasy lipid tissue kit (Qiagen, 74804). RNA quality was checked by Agilent's 2100 Bioanalyzer (using Eukaryote Total RNA Nano kit). RNA-Seq analysis was performed at Erasmus MC Center for Biomics to identify target genes of mmu-miR-142a-3p and mmu-miR-142a-5p that are differentially expressed in the hippocampus samples of mice in the different groups. Animal experiments were approved by the Federal Authorities of Animal Research of the Regierungspräsidium Giessen, Hessen, Germany (Approved Protocol No. 613_M).

Data analysis and statistics

The nonparametric Mann–Whitney test was used to compare miR-142 expression between the genotype groups, and an unpaired/independent t-test was used to compare reporter gene activities. *In vitro* experiments were repeated at least three times and histograms represent mean ±S.D. Statistical differences were measured using unpaired two-sided Student's *t*-test. $P < 0.05$ was considered as statistically significant. Data analysis was performed using Excel Software Version 14.4.5.

RNA-Seq was performed with the Illumina TruSeq Stranded mRNA Library prep kit. The resulting DNA libraries were sequenced on the HiSeq2500, for single-end reads of 50bp length. Reads were generated of 50 base-pairs in length. Reads were mapped against the GRCm38 reference genome using HiSat2 (version 2.0.4)[32]. We called gene expression values (reads per gene) using htseq-count (version 0.6.1)[33]. We took only expressed genes into account, genes with at least 5 reads in at least 7 samples (half of the samples plus one). This filtering in iPSC-derived NPCs resulted in 17,181 genes. Differential expression analysis of the RNA–Seq data in human iPSC-derived NPCs was performed using R (version 3.3.2) and DESeq2 (version 1.14.1)[34-36]. Briefly, DESeq2 generated three values for each gene that were used for subsequent analysis: 1. Log2 fold change (Log2FC), 2. *p*-value based on the Wald test, and 3. Corrected *p*-value controlling the false discovery rate to 5%. Genes were considered differentially expressed if the corrected *p*-value was lower than 0.05. For the biological interpretation of the results, we placed an additional cut-off of fold change ≥ 1.2.

RESULTS

Non-coding RNAs associated with AD

In total, we examined the associations of 108,862 unique SNPs in primary miRNA sequences and lncRNAs with AD. Of these, SNPs annotated to 5 miRNAs and 10 lncRNAs, located in 7 distinct loci (each locus defined as 1 Mb), exceeded the significance threshold (*p*-value < 1.02

$\times 10^{-6}$) (**Table 1**). We assessed whether the associated ncRNA SNPs are likely to be functional in their corresponding loci based on a set of criteria recommended by previous studies to assess the potential functionality of ncRNA SNPs in GWAS results[37,38]. These criteria include an established association between SNP and the trait, the correlation of SNP with expression of the host ncRNA, the localization of SNP in the ncRNA regulatory regions and the potential of SNP for structural perturbations in the host ncRNA.

Table 1. Top variants in 15 non-coding RNAs significantly associated with Alzheimer's disease

SNP ID	ncRNA ID	Annotated gene	Chr:position	Locus	A1>A2	MAF	Beta	*p*-value
rs7384878	miR-6840	*PMS2P1*	7:100334426	1	C>T	0.32	-0.018	3.98×10^{-15}
rs611418	miR-6503	*MS4A4E*	11:60243540	2	C>T	0.35	-0.016	2.46×10^{-13}
rs10792264	lnc-MS4A4A-1	*NA*	11:60318017	2	A>C	0.36	0.012	2.29×10^{-8}
rs636355	lnc-CCDC83-1	*PICALM*	11:86013618	3	T>A	0.44	-0.019	1.25×10^{-17}
rs77162419	lnc-SLTM-2	*SLTM*	15:58926990	4	C>A	0.07	-0.021	8.06×10^{-7}
rs850520	lnc-ABI3-2:5	*FLJ40194*	17:49255705	5	A>G	0.46	0.010	9.25×10^{-7}
rs56229705	lnc-USP6-1	*LOC101928000*	17:5111494	5	G>A	0.12	0.018	1.09×10^{-7}
rs75511804	lnc-USP6-2	*LOC100130950*	17:5235009	5	C>T	0.12	0.020	1.68×10^{-9}
rs2632516	lnc-BZRAP1-1	*TSPOAP1-AS1/ MIR142*	17:58331728	6	G>A	0.47	-0.010	9.66×10^{-7}
rs2526377	miR-142	*TSPOAP1-AS1/ MIR142*	17:58332680	6	A>G	0.46	-0.011	9.13×10^{-7}
rs203709	miR-4531	*LOC107985305*	19:44658298	7	T>A	0.49	-0.020	4.15×10^{-17}
rs12459810	lnc-ZNF180-2	*BCL3*	19:44746404	7	C>T	0.27	0.083	4.08×10^{-44}
rs2965169	miR-8085	*BLC3*	19:44747899	7	A>C	0.47	-0.034	3.13×10^{-57}
rs3760628	lnc-ZNF296-1	*CLPTM1*	19:44953968	7	G>A	0.46	0.012	9.31×10^{-10}
rs1114831	lnc-NKPD1-1	*PPP1R37*	19:45133061	7	C>A	0.10	0.044	9.04×10^{-37}

Shown are the top variants in 5 miRNAs and 10 lncRNAs (located in 7 distinct loci), exceeding the significance threshold (p-value < 1.02×10^{-6}) to be associated with AD. The associations are based on the data from meta-analysis of phase 1 and 2 of the recent AD GWAS.[7] The table is sorted based on Chr and position (GRCh38.p12). Annotated gene, reported in dbSNP database; Chr, Chromosome; A1, Reference allele; A2, Alternative allele; MAF, Minor allele frequency; Beta, Effect estimate.

As shown in the regional association plots (Supp.**Figure S1**), in 6 of the 7 identified non-coding RNA loci were coding variants in strong LD ($R^2 > 0.7$) with the ncRNA SNPs more significantly associated with AD (Supp.**Table S1**), which complicating interpretations of the coding versus non-coding variants in these loci. In contrast, in the 17q22 locus, the top-associated variant (rs2526377:A>G, chr17:58332680) localized in two ncRNAs, *MIR142* and *BZRAP1-AS1*, and exhibited with no proxy variants in high LD in coding regions. We thus focused our further investigations on the 17q22 locus (**Figure 1**). Evaluation of the LD pattern in 17q22 revealed four SNPs in LD with an $R^2 > 0.7$. Of these, rs2632516 and rs2526377, in very high LD ($R^2 = 1.0$), are annotated to *MIR142* and, on the reverse DNA strand, located in

the last intron of *BZRAP1-AS1*. The other two SNPs, rs2526378 and rs2526380, are located in the first and last introns of *BZRAP1*, a coding gene ~3kb away from the top-associated SNP rs2526377(**Figure 2**). Regulome DB and HaploReg showed that three of the SNPs (rs2526378, rs2526380 and rs2632516) are intronic and without any predicted functions. Conversely, the top SNP rs2526377 is located in a highly conserved promoter region upstream of miR-142, which could control the expression of miR-142[39] (**Figure 2**). Using the UCSC genome browser and ENCODE data, we further found that rs2526377 overlaps with the well-conserved binding sites of multiple transcription factors (Supp.**Table S2** and Supp.**Figure S2**), which their bindings to the miR-142 promoter might be perturbed by the SNP. Moreover, the eQTL data from the BBMRI-NL consortium showed that the rs2632577 minor allele is associated with lower miR-142 expression levels in blood (p-value = 4.84× 10^{-11} , Z-score = -6.58).

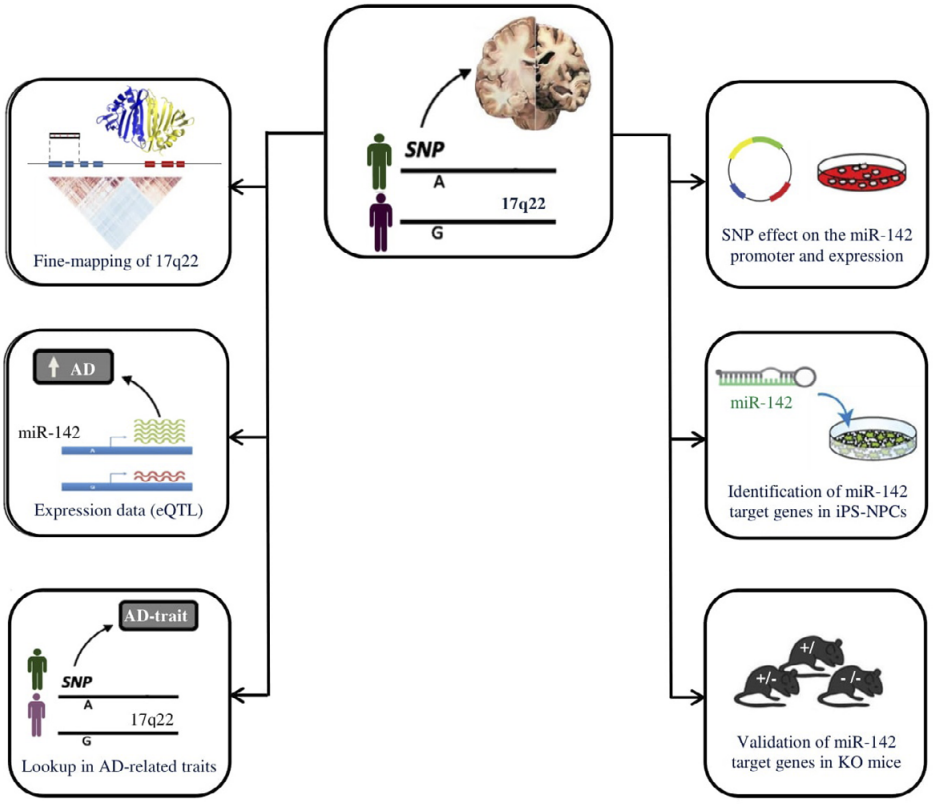

Figure 1. *In silico* and *in vitro* studies to elucidate the role of SNPs in 17q22 and miR-142 in AD. The figure summarizes our analyses to test the functionality of non-coding RNA variants at the 17q22 locus associated with AD and the role of miR-142 in AD pathogenesis. GWAS, Genome-wide association study; SNP, Single-nucleotide polymorphism; KO, Knock-out (miR-142$^{-/-}$); NPCs, human iPS-derived neural progenitor cells; eQTL, expression quantitative trait loci; AD-related traits, Look-up in GWAS of cognitive ability and educational attainment.

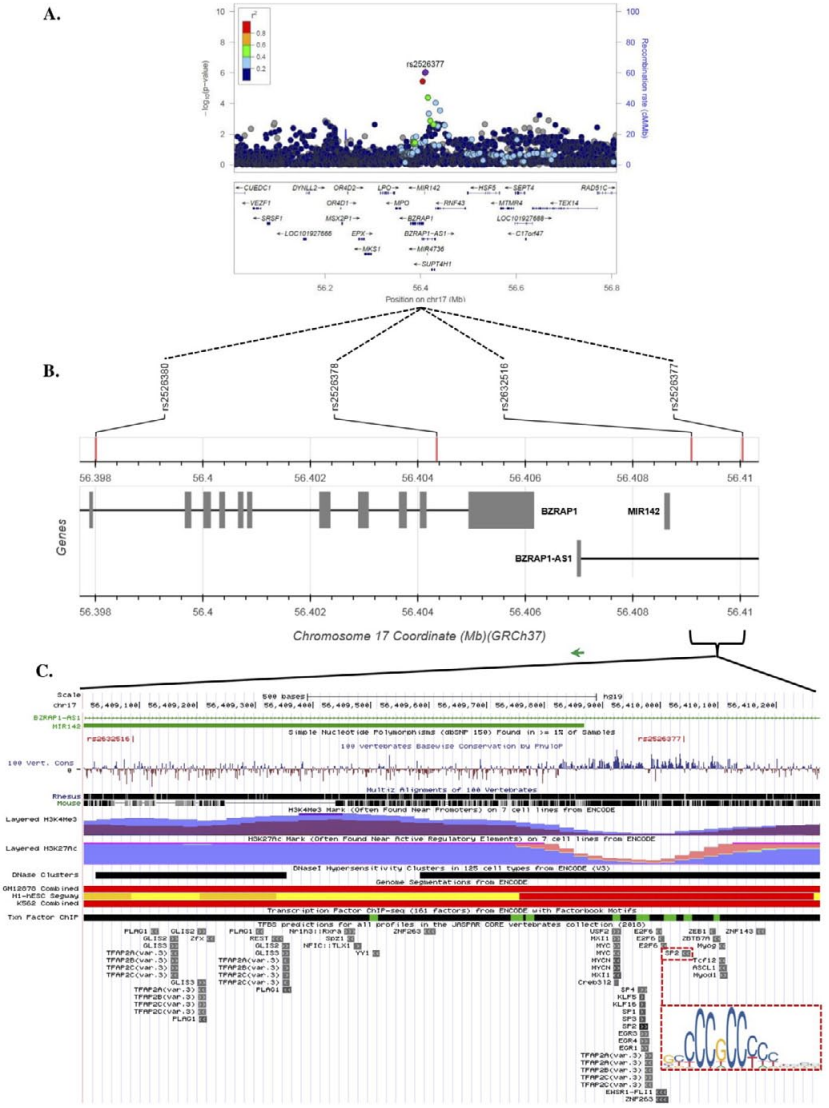

Figure 2. Fine-mapping of SNPs in the 17q22 locus associated with AD. A) The regional association plot shows the association of rs2526377 and its proxies at the 17q22 locus with AD based on the meta-analysis of phase 1 and 2 of the Jansen et al. paper. The *p*-values of variants are plotted (as –log10 of the p-value) against their physical position on the locus. A purple diamond represents the *p*-value for the top associated SNP rs2526377. Estimated recombination rates from the 1000 Genomes project (European population) show the local LD structure. The variant's colors indicate LD with the top SNP according to a scale from r^2= 0 to r^2= 1 based on pairwise r^2 values from the 1000 Genomes project. **B)** The figure displays the localization of the four SNPs in strong LD ($r^2 > 0.7$) at the 17q22 locus. These include two SNPs (rs2632516 and rs2526377) upstream of *MIR142*, and in the last intron of *BZRAP1-AS1*, and two SNPs (rs2526378 and rs2526380) in the first and last introns of *BZRAP1* gene. **C)** The functional mapping of the locus and gene annotations from the UCSC genome browser are also shown below.

As an additional analysis, we tested whether rs2526377 is connected to cognitive functioning prior to the clinical manifestations of AD, so that the variant can be used as early marker of disease. To this end, we examined the association of rs2526377 with cognitive ability and educational attainment using the publicly available GWAS data[40,41]. The SNP minor allele (G) was positively associated with cognitive function (*P*-value = 0.046, Beta = 0.011) and educational attainment (*P*-value = 0.005, Beta = 0.01), which is consistent with the protective effect of the G allele for AD risk.

Together, these data indicate that rs2526377 fulfilled predefined criteria for being a functional variant in the 17q22 locus, which possibly function by altering the expression of miR-142. We performed various *in vitro* and *in silico* studies to functionally show the impact of rs2526377 on the expression levels of miR-142 and gain insight into the function of miR-142 and its targets in the pathogenesis of AD.

The impact of rs2526377 on the promoter activity of miR-142

The promoter region and transcription start site (TSS) of miR-142 have previously been characterized[39], indicating that transcription of *MIR142* is initiated 1205bp upstream of the pre-miR-142 sequence and the promoter region is located between 1305 and 1570bp upstream of miR-142 (**Figure 3a**). Rs2526377 resides 1362bp upstream of the pre-miR-142 sequence and therefore within the miRNA promoter. To demonstrate whether rs2526377 alters the promoter activity of miR-142, we performed luciferase reporter assays in HEK293 cells. We generated reporter constructs containing either alleles of the SNP rs2526377 and transfected cells with the reporter plasmids, so that the effect of each allele on the promoter activity was evaluated. The construct carrying the rs2526377 major (A) allele exhibited 20% higher basal activity than the construct carrying minor (G) allele (*P*-value = 0.037) (**Figure 3b**). These data are consistent with the eQTL data from the BBMRI-NL consortium that show rs2526377 minor allele carriers have lower miR-142 expression levels in blood.

Potential miR-142 target genes implicated in AD

A miRNA and its target genes should be expressed in the target tissue for any biological function to be exerted. Thus, we first measured the expression of miR-142 in human brain. Both strands of miR-142 were expressed in the brain, miR-142-3p with an average Ct-value of 26 and miR-142-5p with an average Ct-value of 30, relative to the endogenous control RNU6B with an average Ct-value of 21.5 (Supp.**Table S3**). To identify target genes that could mediate the function of miR-142 in the brain, we compiled a list of all putative target genes of miR-142-3p and -5p from three miRNA target prediction databases (TargetsScan, miRDB and DIANA-microT). This resulted in 885 putative target genes for miR-142-3p and 1,541 putative target genes for miR-142-5p (**Figure 4**). We filtered these target genes on the basis of human brain expression, using the Illumina's Human Body Map RNA-Seq dataset. We focused our analysis on the 814 brain-expressed putative target genes of miR-142-3p and 1,393 of miR-142-5p. To

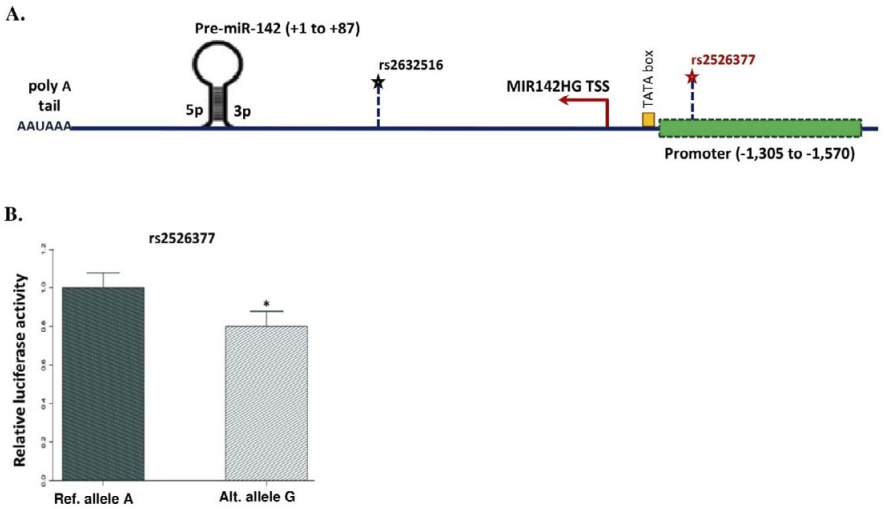

Figure 3. The genomic location of rs2526377 and its effect on the promoter activity and expression of miR-142. A) A schematic showing the position of rs2526377 upstream of pre-miR-142 sequence. The position of the miR-142 promoter region, transcription start site (TSS) and regulatory elements adapted from *Skarn et al.*, *PLosOne* 2013.[39] **B)** Luciferase reporter assay was performed to determine the effect of rs2526377 on the miR-142 promoter activity. The reporter gene constructs containing either the SNP alleles were generated and HEK293 cells were transfected with the reporter plasmids. The construct carrying the major allele A of rs2526377 had 20% higher basal activity in HEK293 cells than the construct carrying the minor allele G (*P*-value = 0.037). Error bars represent standard deviation (SD). $^*P < 0.05$ compared with the control group (Student's *t*-test). NS, non-significant.

examine the regulatory effect of miR-142-3p and -5p on the expression levels of their putative target genes, we used human iPSC-derived neural progenitor cells (NPCs). We overexpressed either miR-142-3p or -5p in NPCs using mirVana™ miRNA Mimics and performed differential expression analysis by RNA-Seq. To elucidate miR-142 target genes implicated in the AD pathogenesis, we applied the two most commonly used methods for detecting miRNA targets.

First, we conducted a hypothesis-free differential expression analysis considering all brain-expressed target genes of miR-142. We sought to identify target genes that were significantly down-regulated after overexpression of the mature miRNA (3p or 5p) with FDR-adjusted *P*-value < 0.05 and fold change ≥ 1.2. Of the 814 brain-expressed putative target genes of miR-142-3p, 280 genes were significantly down-regulated by the miR-142-3p mimic *vs* untreated, and 74 genes were significantly down-regulated in NPCs transfected with miR-142-3p mimic *vs* negative control (Supp.**Table S4** and **Table 2**). We performed KEGG pathway analysis for the 74 identified target genes of miR-142-3p and observed significant enrichment in Regulation of actin cytoskeleton (*WASL*, *ITGB8*, *APC*, *GNG12*, *CFL2*, *GNG12* and *ENAH*), Adherence junction (*TGFBR1*, *WASL*, *RAC1* and *YES1*) and Axon guidance (*CFL2*, *RAC1* and *SEMA3D*) (Supp.**Table S5** and Supp.**Figure S3**). Of the 1,393 brain-expressed putative target

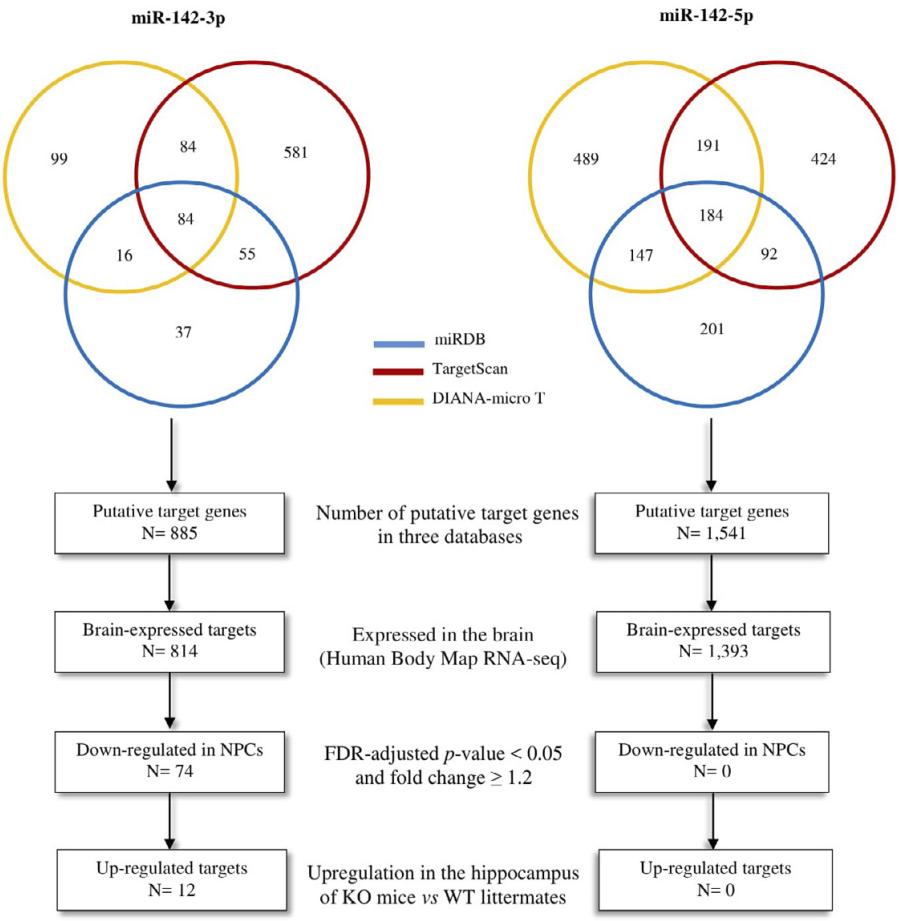

Figure 4. Identification of potential target genes of miR-142 in the brain implicated in AD. We compiled a list of all putative targets of miR-142-3p and -5p from three widely used miRNA target prediction databases (miRDb, TargetScan, and DIANA-micro T). The target genes found to be expressed in the human brain were included. We overexpressed miR-142 (3p or 5p) in human iPS-derived NPCs and performed RNA-Seq to examine the changes in gene expression pattern. Target genes significantly down-regulated (FDR-adjusted $P < 0.05$ and fold change ≥ 1.2) in NPCs transfected with mature miR-142 mimic (3p or 5p) *vs* untreated and negative control were retrieved. Subsequently, RNA-Seq was performed on hippocampus of miR-142 KO mice and their wildtype littermates to confirm miR-142-mediated regulation of the identified target genes in human iPS-derived NPCs.

genes of miR-142-5p, none of them were significantly down-regulated (FDR-adjusted *P*-value < 0.05 and fold change ≥ 1.2) in NPCs transfected with miR-142-5p mimic *vs* negative control.

Second, we examined the association of the 814 putative target genes of miR-142-3p and 1,393 putative target genes of miR-142-5p with AD using a candidate gene approach. To this end, we extracted genetic variants located in these target genes and tested their associations

Table 2. The top 20 target genes of miR-142-3p significantly down-regulated in human iPS-derived NPCs

Gene Name	miR-142-3p mimic *vs* untreated		miR-142-3p mimic *vs* negative control	
	Fold change	FDR-adj *p*-value	Fold change	FDR-adj *p*-value
WASL	1.7	7.14×10^{-21}	1.6	2.83×10^{-16}
YES1	1.6	3.60×10^{-21}	1.4	4.64×10^{-11}
BOD1	1.6	4.78×10^{-26}	1.4	4.64×10^{-11}
VAMP3	1.6	2.17×10^{-22}	1.4	9.39×10^{-11}
IL6ST	1.6	2.33×10^{-22}	1.4	1.03×10^{-10}
CFL2	1.5	1.69×10^{-11}	1.5	1.02×10^{-09}
SUCO	1.4	7.46×10^{-15}	1.4	4.24×10^{-09}
CASK	1.2	3.05×10^{-06}	1.3	8.12×10^{-09}
TWF1	2.0	5.20×10^{-30}	1.5	1.08×10^{-08}
TNFRSF12A	1.2	2.58×10^{-01}	2.0	1.12×10^{-08}
CLIC4	1.6	3.06×10^{-22}	1.3	1.13×10^{-08}
TGFBR1	1.4	1.43×10^{-11}	1.3	1.56×10^{-07}
ITGB8	1.6	3.17×10^{-16}	1.4	6.53×10^{-07}
MANBAL	1.3	3.44×10^{-05}	1.4	6.53×10^{-07}
FAM127B	1.3	7.88×10^{-05}	1.4	1.51×10^{-06}
RHOBTB3	1.8	7.20×10^{-33}	1.3	1.67×10^{-06}
HEATR5A	1.2	16.0×10^{-02}	1.3	1.74×10^{-05}
RAB2A	1.3	9.41×10^{-09}	1.3	2.00×10^{-05}
HSPA1B	1.8	2.69×10^{-20}	1.4	3.48×10^{-05}
DIRC2	1.2	1.01×10^{-02}	1.4	3.74×10^{-05}

The table shows the 20 most significantly down-regulated target genes in miR-142-3p overexpressing iPSC-derived neural progenitor cells (NPCs). Out of 814 predicted target genes of miR-142-3p, 280 were down-regulated in NPCs transfected with miR-142-3p mimic *vs* untreated and 74 were down-regulated in NPCs transfected with miR-142-3p mimic *vs* negative control (FDR-adjusted *p*-value < 0.05 and Fold change ≥ 1.2).

with AD using the GWAS data[7]. After Bonferroni correction for the number of tested variants in all target genes of miR-142-3p (0.05 /62,515 = 8.0×10^{-7}), four target genes passed the significance threshold (Supp.**Table S6**). Among these, *PICALM* (rs867611, *P*-value = 2.19×10^{-18}) was the only target gene significantly down-regulated in NPCs transfected with miR-142-3p mimic compared to both untreated (*P*-value = 1.6×10^{-8}, fold change = 1.2) and negative control (*P*-value = 2.7×10^{-2}, fold change = 1.1) conditions. We additionally confirmed the down-regulation of *PICALM* in NPCs transfected with miR-142-3p mimic by qPCR, which demonstrated ~30% reduction of the *PICALM* expression compared to untreated NPCs. A similar analysis was performed for 1,393 target genes of miR-142-5p. One target gene (*FAM63B*) passed the significance threshold (0.05/128,444 = 3.9×10^{-7}) (Supp.**Table S6**), however, the gene was not significantly down-regulated in NPCs transfected with miR-142-5p mimic compared to negative controls.

Validation of the identified miR-142-3p target genes in the hippocampus of miR-142$^{-/-}$ mice

To confirm regulation of the 74 identified target genes of miR-142-3p in the brain, we performed RNA-Seq of the hippocampus from miR-142 homozygous KO mice (miR-142$^{-/-}$) and compared it with heterozygous KO (miR-142$^{+/-}$) and wild-type (Wt) littermates (Supp.**Table S7**). Twelve (n=12) of the 74 identified target genes of miR-142-3p were up-regulated in the hippocampus of both homozygous and heterozygous KO mice *vs* Wt littermates (P-value < 0.05) (**Table 3**). IPA (Ingenuity Pathway Analysis) revealed that five of the twelve target genes (*TGFBR1*, *CFL2*, *SEMA3D*, *ALCAM* and *RHOQ*) are over-represented in Nervous System Development and Function (P-value = 4.98×10^{-2}-5.54×10^{-4}) (Supp.**Table S8**). Among the twelve target genes, miR-142-3p-mediated regulation of *TGFBR1* and *CFL2* have also been validated experimentally in the previous studies.[42,43]

Table 3. Twelve target genes of miR-142-3p up-regulated in the hippocampus of miR-142 KO mice

Gene Name	Homozygous KO *vs* WT		Hetrozygous *vs* WT	
	Fold change	*p*-value	Fold change	*p*-value
Tgfbr1	1.22	2.43×10^{-03}	1.20	6.68×10^{-03}
Rhoq	1.18	3.46×10^{-03}	1.15	1.50×10^{-02}
Slc39a10	1.15	5.62×10^{-03}	1.15	6.56×10^{-03}
Ppp1r2	1.67	5.69×10^{-03}	1.20	1.04×10^{-03}
Cfl2	1.18	7.43×10^{-03}	1.15	2.04×10^{-02}
Pafah1b2	1.11	7.51×10^{-03}	1.12	4.49×10^{-03}
Rab1a	1.10	8.11×10^{-03}	1.12	1.30×10^{-03}
Rab18	1.15	1.49×10^{-02}	1.20	2.25×10^{-03}
Alcam	1.13	1.64×10^{-02}	1.09	1.64×10^{-02}
Hspa4l	1.13	2.33×10^{-02}	1.14	1.81×10^{-02}
Rab2a	1.12	2.55×10^{-02}	1.13	1.70×10^{-02}
Sema3d	1.17	2.79×10^{-02}	1.09	2.42×10^{-01}

The table shows 12 target genes of miR-142-3p that are up-regulated in the hippocampus of KO mice (miR142$^{-/-}$). Out of the 74 identified target genes of miR-142-3p, which were significantly down-regulated in iPS-derived NPCs, twelve were confirmed to be up-regulated in the hippocampus of KO mice (miR-142$^{-/-}$) compared to their WT (miR-142$^{+/+}$) littermates (P < 0.05).

Moreover, *PICALM*, the target gene found to be significantly associated with AD in the GWAS data, was up-regulated in the hippocampus of miR-142 KO mice *vs* Wt littermates (P-value =1.5×10^{-2}, fold change = 1.13) and heterozygous KO mice *vs* Wt littermates (P-value = 7.2×10^{-4}, fold change = 1.2) (**Figure 5**).

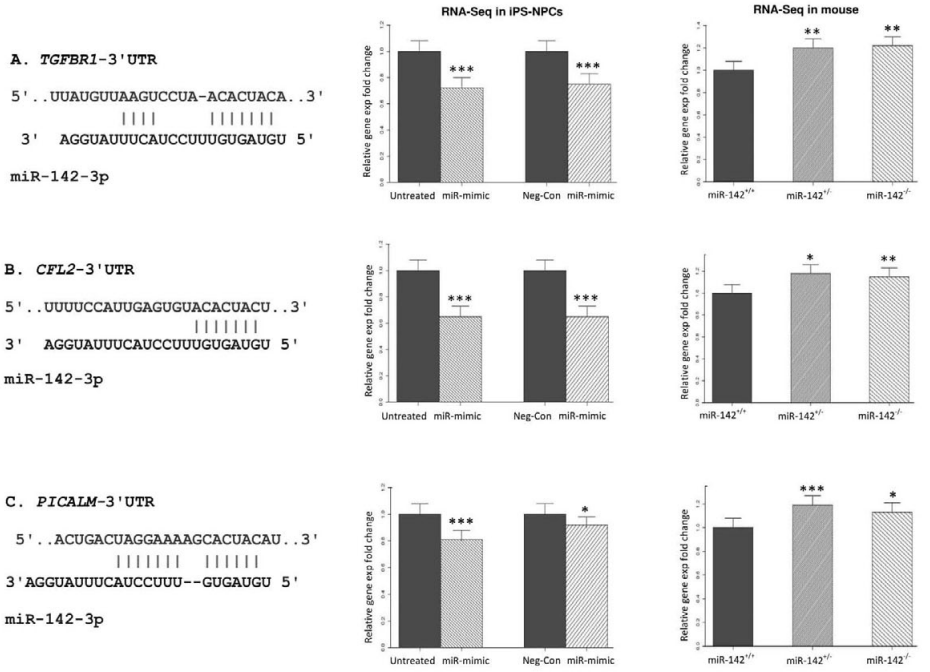

Figure 5. The interaction and regulatory effect between miR-142-3p and its three target genes. The figure illustrates the binding of miR-142-3p to its three highlighted target genes (*TGFBR1*, *CLF2*, and *PICALM*). The expression of these target genes were significantly down-regulated in human iPS-NPCs transfected with miR-142-3p mimic *vs* untreated, and in iPS-NPCs transfected with miR-142-3p mimic *vs* negative control. In contrast, the expression of these target genes were up-regulated in the hippocampus of miR-142 KO mice *vs* Wt littermates. Error bars represent standard deviation (SD). $^{*}P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$ compared with the control group (Wald-test).

DISCUSSION

Despite increasing interest in the biology of non-coding RNAs, relatively few genome-wide studies have thus far demonstrated associations with human disease. In this study, we performed a genome-wide scan to systematically investigate the association of miRNAs and lncRNAs with AD by leveraging publicly available GWAS summary statistics[7]. We found seven distinct ncRNA loci significantly associated with AD including a newly identified susceptibility locus on 17q22, in which the ncRNA variant leads the signal and fulfills predefined criteria for being functional. The locus has not been reported as significant in the original GWAS, because the p-value of the top SNP in the meta-analysis of phase 1 (AD case/control) and phase 2 (AD-by-proxy) was above the GWAS threshold[7]. However, the SNP exceeds the GWAS threshold in the phase 1 of this GWAS meta-analysis (P-value = 1.42×10^{-9}), combining data from the two large-scale AD case/control consortia, IGAP and PGC-ALZ. In the phase 2, using the AD-by-proxy phenotype from the UK biobank cohort, the association between

rs2632516 and AD is less significant (*P*-value = 5.0×10^{-3}), but still in the same direction. The lower association signal for the 17q22 locus in the UK biobank cohort could be explained by differences in case ascertainment of AD. In the UK biobank, Alzheimer dementia is ascertained via self-report information from family history (parent or first-degree relative with AD or dementia) as a proxy-phenotype for the participants[44]. This method relies on people to provide accurate information about whether their parents developed AD, for which misclassification of case status is of greater concern than consortia relying upon clinician reported diagnoses. In addition, a trans-ethnic GWAS, by adding more samples to the IGAP GWAS data, recently reported the significant association of 17q22 with AD[45]. In this trans-ethnic GWAS, however, the leading ncRNA variant in the 17q22 locus was annotated to the closest protein-coding gene (*BZRAP1*), and the potential impact of miR-142 has been overlooked. In contrast a more recent GWAS, investigating the association of rare coding variants with AD, with an even larger sample size did not find any significant association between rare variants in *BZRAP1* gene and AD[46]. In this line, our results demonstrated that miR-142 is the most likely functional target in the 17q22 locus implicated in AD pathogenesis.

Genetic variants in miRNA-encoding sequences have been shown previously to affect miRNAs expression and subsequently influence gene regulation in complex diseases[47-49]. Moreover, the functional impact of variants on the promoter activity of miRNAs has been revealed, most notably for rs57095329 located in miR-146a, by altering the miRNA processing and expression level[50]. Here, we demonstrated that rs2526377 affects the promoter activity and reduces the expression levels of miR-142. Previously, Skarn et al. characterized the miR-142 promoter region and demonstrated that DNA methylation of specific CpG sites in the region represses the promoter activity and reduces the expression level of miR-142 in mesenchymal stem cells[39]. Moreover, an independent study by Mor et al. revealed that hypomethylation of the CpGs in the miR-142 promoter region increases the miRNA expression level in the prefrontal cortex of autism patients[51]. These data may indicate that rs2526377 attenuates the risk of AD via reducing the miR-142 expression levels in the brain.

MiR-142 is a highly conserved miRNA amongst multiple invertebrate and vertebrate species. The role of miR-142 has extensively been studied in the hematopoietic system, lung development and cardiac hypertrophy[52]. Convergent evidence from multiple investigations also indicates the expression of miR-142 in the brain, suggesting that dysregulation or malfunction of miR-142 contribute to the pathogenesis of brain disorders. For instance, Junker et al. reported miR-142 among the 10 miRNAs that are more abundant in active multiple sclerosis (MS) brain lesions than normal white matter, and suggested miR-142 to be involved in the brain inflammatory and degenerative diseases[53]. Similarly, Mandolesi et al. observed that miR-142 is increased in the CSF of patients with active MS[54]. Moreover, Sorensen et al. performed miRNA expression profiles in CSF and blood of patients with AD and found a number of differentially expressed miRNAs, in which miR-142 is one of the significantly up-regulated miRNAs in AD patients compared to controls[55]. Two independent studies also revealed that

the expression of miR-142 is increased by age[56,57]. Here, our expression data confirmed that both mature miR-142-3p and -5p are expressed at relatively high levels in the brain; though, our RNA-Seq analysis proposed miR-142-3p, which is the guide strand of miR-142, to be more active on the regulation of its target genes in the brain. Consistent with this notion, Lau et al. have performed miRNA expression profiling of the hippocampus of a cohort of 41 AD patients and 23 age-matched controls and found miR-142-3p among the 15 significantly up-regulated miRNAs in the AD group[58]. Moreover, miR-142-3p has been reported as one of the eight miRNAs up-regulated in synaptoneurosomes from forebrains and hippocampus of mice during prion disease[59]. Together, these data endorse that alterations in the expression of miR-142 in the brain could confer AD risk, where higher levels of miR-142-3p increase a person's risk of developing Alzheimer's.

Up-regulation of miR-142 in the brain may influence AD risk through different mechanisms. Gene ontology analysis on the putative target genes of miR-142-3p and -5p has shown enrichment in categories related to synaptic transmission (dopaminergic synapse, neurotrophin signaling, axon guidance) and signal transduction (TGF-β signaling, MAPK signaling, ErbB signaling)[51,60]. Mandolesi et al. proposed miR-142 to be related to neuro-inflammatory changes in the brain occurring during MS by regulating the expression of *IL-1β*[54]. Further, Chaudhuri et al. suggested the involvement of miR-142 in autoimmune and neuro-inflammation in the brain, via miR-142-mediated repression of *SIRT1* in primary human neurons[61]. In an independent study, Chaudhuri et al. verified that miR-142 indirectly reduces MAOA protein level via regulating *SIRT1* expression[62]. Since *MAOA* is a neurotransmitter-metabolizing enzyme and delaminates serotonin, melanin, epinephrine and norepinephrine, they postulated that miR-142 up-regulation might contribute to change the dopaminergic neurotransmission by lowering *MAOA* expression and activity. In this study, we further demonstrated miR-142-3p-mediated regulation of multiple target genes in the brain that are involved in the pathways underlying AD. *TGFBR1* and *PICALM*, among others, are of particular interest (**Figure 6**). *TGFBR1* has been shown in several studies to be implicated in AD pathogenesis[63-68]. The regulation of *TGFBR1* expression by miR-142-3p has been experimentally confirmed at mRNA and protein levels in previous studies[43,69]. Our differential expression analysis for all miR-142-3p target genes demonstrated that *TGFBR1* was significantly down-regulated in miR-142-3p overexpressing human iPS-derived NPCs and the top target gene up-regulated in the hippocampus of miR-142 KO mice. Locating at the intersection of anti-inflammatory, anti-aging and neuroprotective pathways, *TGFBR1* makes a promising molecule for mediating the function of miR-142-3p in AD.

PICALM is ubiquitously expressed in all tissue types with prominent expression in neurons and is non-selectively distributed in pre- and postsynaptic terminals, where it plays an essential role in the fusion of synaptic vesicles to the presynaptic membrane in neurotransmitter release[70]. Several GWA studies have independently confirmed the association of *PICALM* with AD[8,46,71]. Recent studies have also shown that *PICALM* level is reduced in the AD brain

endothelium and postulated that it can potentially lead to Aβ accumulation in the brain by hindering LRP1-mediated Aβ transport[72,73]. These data strongly suggest that derepression of *PICALM* in response to the reduced miR-142 expression may decrease AD risk that deserve further and more deep investigation in future experimental work.

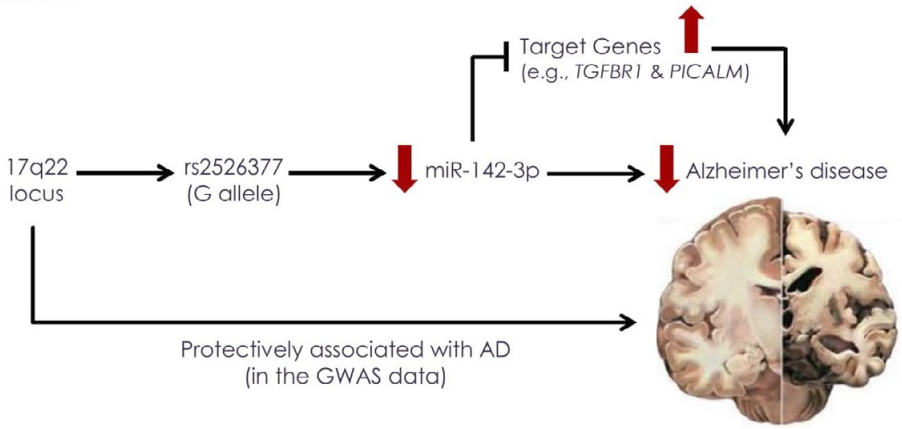

Figure 6. Rs2526377 in the promoter of miR-142 modulating its expression and conferring risk of AD. The SNP rs2526377 occurring within the promoter region of miR-142 alters the promoter activity and reduces the expression level of miR-142. Downregulation of miR-142-3p in the brain results in derepression of multiple target genes (e.g., *TGFBR1* and *PICALM*) that contribute to the pathogenesis of AD.

Conclusions

In this study, we endorse 17q22 as a susceptibility locus for AD and provide evidence demonstrating that miR-142 is the most likely functional target in the locus involved in AD pathogenesis. Furthermore, we revealed miR-142-3p-mediated regulation of multiple target genes in the brain that are implicated in the inflammatory and neurodegenerative manifestations of AD. These include two well-validated AD-associated genes, *TGFBR1* and *PICALM*, of which their derepression in the brain due to reduced expression levels of miR-142-3p may decrease risk of AD. Our findings may also suggest the therapeutic potential of miR-142 inhibition for AD, which warrants further investigations in future.

ABBREVIATIONS

AD, Alzheimer's disease; GWAS, Genome-wide association studies; SNP, Single-nucleotide polymorphism; lncRNA, long non-coding RNA; miRNA, microRNA; eQTL, expression quantitative trait loci; mRNA, messenger RNA; MFE, Minimum free energy; MAF, Minor allele frequency; LD, Linkage disequilibrium; GFP, green fluorescent protein; MSCV-BC,

Murine Stem Cell Virus-Bar Coded; TSS, Transcription start site; PCR, Polymerase change reaction; iPSC, induced pluripotent stem cell; FPKM, Fragments Per Kilobase Million; FDR, False discovery rate; Wt, wild-type; KO, Knock-out.

REFERENCES

1. Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. 2007. Forecasting the global burden of Alzheimer's disease. Alzheimers Dement 3(3):186-91.
2. Crews L, Masliah E. 2010. Molecular mechanisms of neurodegeneration in Alzheimer's disease. Hum Mol Genet 19(R1):R12-20.
3. Hampel H, Frank R, Broich K, Teipel SJ, Katz RG, Hardy J, Herholz K, Bokde AL, Jessen F, Hoessler YC and others. 2010. Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives. Nat Rev Drug Discov 9(7):560-74.
4. Schmechel DE, Saunders AM, Strittmatter WJ, Crain BJ, Hulette CM, Joo SH, Pericak-Vance MA, Goldgaber D, Roses AD. 1993. Increased amyloid beta-peptide deposition in cerebral cortex as a consequence of apolipoprotein E genotype in late-onset Alzheimer disease. Proc Natl Acad Sci U S A 90(20):9649-53.
5. Ertekin-Taner N. 2007. Genetics of Alzheimer's disease: a centennial review. Neurol Clin 25(3):611-67, v.
6. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL. 2006. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry 63(2):168-74.
7. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, Sealock J, Karlsson IK, Hagg S, Athanasiu L and others. 2019. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet 51(3):404-413.
8. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B and others. 2013. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 45(12):1452-8.
9. Zhou X, Chen Y, Mok KY, Zhao Q, Chen K, Chen Y, Hardy J, Li Y, Fu AKY, Guo Q and others. 2018. Identification of genetic risk factors in the Chinese population implicates a role of immune system in Alzheimer's disease pathogenesis. Proc Natl Acad Sci U S A 115(8):1697-1706.
10. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C and others. 2005. The transcriptional landscape of the mammalian genome. Science 309(5740):1559-63.
11. Consortium EP. 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489(7414):57-74.
12. Meza-Sosa KF, Valle-Garcia D, Pedraza-Alva G, Perez-Martinez L. 2012. Role of microRNAs in central nervous system development and pathology. J Neurosci Res 90(1):1-12.
13. O'Carroll D, Schaefer A. 2013. General principals of miRNA biogenesis and regulation in the brain. Neuropsychopharmacology 38(1):39-54.
14. Ma L, Bajic VB, Zhang Z. 2013. On the classification of long non-coding RNAs. RNA Biol 10(6):925-33.
15. Rinn JL, Chang HY. 2012. Genome regulation by long noncoding RNAs. Annu Rev Biochem 81:145-66.
16. Modarresi F, Faghihi MA, Patel NS, Sahagan BG, Wahlestedt C, Lopez-Toledano MA. 2011. Knockdown of BACE1-AS Nonprotein-Coding Transcript Modulates Beta-Amyloid-Related Hippocampal Neurogenesis. Int J Alzheimers Dis 2011:929042.
17. Mus E, Hof PR, Tiedge H. 2007. Dendritic BC200 RNA in aging and in Alzheimer's disease. Proc Natl Acad Sci U S A 104(25):10679-84.
18. Gong J, Wu Y, Zhang X, Liao Y, Sibanda VL, Liu W, Guo AY. 2014. Comprehensive analysis of human small RNA sequencing data provides insights into expression profiles and miRNA editing. RNA Biol 11(11):1375-85.

19. Saini HK, Griffiths-Jones S, Enright AJ. 2007. Genomic analysis of human microRNA transcripts. Proc Natl Acad Sci U S A 104(45):17719-24.
20. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. 2010. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26(18):2336-7.
21. Ward LD, Kellis M. 2016. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. Nucleic Acids Res 44(D1):D877-81.
22. Romanoski CE, Glass CK, Stunnenberg HG, Wilson L, Almouzni G. 2015. Epigenomics: Roadmap for regulation. Nature 518(7539):314-6.
23. Bonder MJ, Luijk R, Zhernakova DV, Moed M, Deelen P, Vermaat M, van Iterson M, van Dijk F, van Galen M, Bot J and others. 2017. Disease variants alter transcription factor levels and methylation of their binding sites. Nat Genet 49(1):131-138.
24. Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. Algorithms Mol Biol 6:26.
25. Agarwal V, Bell GW, Nam JW, Bartel DP. 2015. Predicting effective microRNA target sites in mammalian mRNAs. Elife 4.
26. Wang X, El Naqa IM. 2008. Prediction of both conserved and nonconserved microRNA targets in animals. Bioinformatics 24(3):325-32.
27. Paraskevopoulou MD, Georgakilas G, Kostoulas N, Vlachos IS, Vergoulis T, Reczko M, Filippidis C, Dalamagas T, Hatzigeorgiou AG. 2013. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. Nucleic Acids Res 41(Web Server issue):W169-73.
28. Kawamoto S, Yoshii J, Mizuno K, Ito K, Miyamoto Y, Ohnishi T, Matoba R, Hori N, Matsumoto Y, Okumura T and others. 2000. BodyMap: a collection of 3' ESTs for analysis of human gene expression information. Genome Res 10(11):1817-27.
29. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and others. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102(43):15545-50.
30. Shi Y, Kirwan P, Livesey FJ. 2012. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. Nat Protoc 7(10):1836-46.
31. Shrestha A, Carraro G, El Agha E, Mukhametshina R, Chao CM, Rizvanov A, Barreto G, Bellusci S. 2015. Generation and Validation of miR-142 Knock Out Mice. PLoS One 10(9):e0136913.
32. Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12(4):357-60.
33. Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31(2):166-9.
34. Cuypers B, Domagalska MA, Meysman P, Muylder G, Vanaerschot M, Imamura H, Dumetz F, Verdonckt TW, Myler PJ, Ramasamy G and others. 2017. Multiplexed Spliced-Leader Sequencing: A high-throughput, selective method for RNA-seq in Trypanosomatids. Sci Rep 7(1):3725.
35. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550.
36. Tian Y, Liao IH, Zhan X, Gunther JR, Ander BP, Liu D, Lit L, Jickling GC, Corbett BA, Bos-Veneman NG and others. 2011. Exon expression and alternatively spliced genes in Tourette Syndrome. Am J Med Genet B Neuropsychiatr Genet 156B(1):72-8.
37. Ghanbari M, Peters MJ, de Vries PS, Boer CG, van Rooij JGJ, Lee YC, Kumar V, Uitterlinden AG, Ikram MA, Wijmenga C and others. 2018. A systematic analysis highlights multiple long non-coding RNAs associated with cardiometabolic disorders. J Hum Genet 63(4):431-446.

38. Ryan BM, Robles AI, Harris CC. 2010. Genetic variation in microRNA networks: the implications for cancer research. Nat Rev Cancer 10(6):389-402.
39. Skarn M, Baroy T, Stratford EW, Myklebost O. 2013. Epigenetic regulation and functional characterization of microRNA-142 in mesenchymal cells. PLoS One 8(11):e79231.
40. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, Hagenaars SP, Ritchie SJ, Marioni RE, Fawns-Ritchie C and others. 2018. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun 9(1):2098.
41. Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, Turley P, Chen GB, Emilsson V, Meddens SF and others. 2016. Genome-wide association study identifies 74 loci associated with educational attainment. Nature 533(7604):539-42.
42. Schwickert A, Weghake E, Bruggemann K, Engbers A, Brinkmann BF, Kemper B, Seggewiss J, Stock C, Ebnet K, Kiesel L and others. 2015. microRNA miR-142-3p Inhibits Breast Cancer Cell Invasiveness by Synchronous Targeting of WASL, Integrin Alpha V, and Additional Cytoskeletal Elements. PLoS One 10(12):e0143993.
43. Yang X, Dan X, Men R, Ma L, Wen M, Peng Y, Yang L. 2017. MiR-142-3p blocks TGF-beta-induced activation of hepatic stellate cells through targeting TGFbetaRI. Life Sci 187:22-30.
44. Marioni RE, Harris SE, Zhang Q, McRae AF, Hagenaars SP, Hill WD, Davies G, Ritchie CW, Gale CR, Starr JM and others. 2018. GWAS on family history of Alzheimer's disease. Transl Psychiatry 8(1):99.
45. Jun GR, Chung J, Mez J, Barber R, Beecham GW, Bennett DA, Buxbaum JD, Byrd GS, Carrasquillo MM, Crane PK and others. 2017. Transethnic genome-wide scan identifies novel Alzheimer's disease loci. Alzheimers Dement 13(7):727-738.
46. Sims R, van der Lee SJ, Naj AC, Bellenguez C, Badarinarayan N, Jakobsdottir J, Kunkle BW, Boland A, Raybould R, Bis JC and others. 2017. Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. Nat Genet 49(9):1373-1384.
47. Dorn GW, 2nd, Matkovich SJ, Eschenbacher WH, Zhang Y. 2012. A human 3' miR-499 mutation alters cardiac mRNA targeting and function. Circ Res 110(7):958-67.
48. Ghanbari M, Darweesh SK, de Looper HW, van Luijn MM, Hofman A, Ikram MA, Franco OH, Erkeland SJ, Dehghan A. 2015. Genetic Variants in MicroRNAs and their Binding Sites are Associated with the Risk of Parkinson Disease. Hum Mutat 37(3):292-300.
49. Ghanbari M, de Vries PS, de Looper H, Peters MJ, Schurmann C, Yaghootkar H, Dorr M, Frayling TM, Uitterlinden AG, Hofman A and others. 2014. A genetic variant in the seed region of miR-4513 shows pleiotropic effects on lipid and glucose homeostasis, blood pressure, and coronary artery disease. Hum Mutat 35(12):1524-31.
50. Luo X, Yang W, Ye DQ, Cui H, Zhang Y, Hirankarn N, Qian X, Tang Y, Lau YL, de Vries N and others. 2011. A functional variant in microRNA-146a promoter modulates its expression and confers disease risk for systemic lupus erythematosus. PLoS Genet 7(6):e1002128.
51. Mor M, Nardone S, Sams DS, Elliott E. 2015. Hypomethylation of miR-142 promoter and upregulation of microRNAs that target the oxytocin receptor gene in the autism prefrontal cortex. Mol Autism 6:46.
52. Shrestha A, Mukhametshina RT, Taghizadeh S, Vasquez-Pacheco E, Cabrera-Fuentes H, Rizvanov A, Mari B, Carraro G, Bellusci S. 2016. MicroRNA-142 is a multifaceted regulator in organogenesis, homeostasis, and disease. Dev Dyn 246(4):285-290.
53. Junker A, Krumbholz M, Eisele S, Mohan H, Augstein F, Bittner R, Lassmann H, Wekerle H, Hohlfeld R, Meinl E. 2009. MicroRNA profiling of multiple sclerosis lesions identifies modulators of the regulatory protein CD47. Brain 132(Pt 12):3342-52.

54. Mandolesi G, De Vito F, Musella A, Gentile A, Bullitta S, Fresegna D, Sepman H, Di Sanza C, Haji N, Mori F and others. 2017. miR-142-3p Is a Key Regulator of IL-1beta-Dependent Synaptopathy in Neuroinflammation. J Neurosci 37(3):546-561.
55. Sorensen SS, Nygaard AB, Christensen T. 2016. miRNA expression profiles in cerebrospinal fluid and blood of patients with Alzheimer's disease and other types of dementia - an exploratory study. Transl Neurodegener 5:6.
56. Huan T, Chen G, Liu C, Bhattacharya A, Rong J, Chen BH, Seshadri S, Tanriverdi K, Freedman JE, Larson MG and others. 2018. Age-associated microRNA expression in human peripheral blood is associated with all-cause mortality and age-related traits. Aging Cell 17(1).
57. Zhang X, Azhar G, Wei JY. 2012. The expression of microRNA and microRNA clusters in the aging heart. PLoS One 7(4):e34688.
58. Lau P, Bossers K, Janky R, Salta E, Frigerio CS, Barbash S, Rothman R, Sierksma AS, Thathiah A, Greenberg D and others. 2013. Alteration of the microRNA network during the progression of Alzheimer's disease. EMBO Mol Med 5(10):1613-34.
59. Boese AS, Saba R, Campbell K, Majer A, Medina S, Burton L, Booth TF, Chong P, Westmacott G, Dutta SM and others. 2016. MicroRNA abundance is altered in synaptoneurosomes during prion disease. Mol Cell Neurosci 71:13-24.
60. Shrestha A, Mukhametshina RT, Taghizadeh S, Vasquez-Pacheco E, Cabrera-Fuentes H, Rizvanov A, Mari B, Carraro G, Bellusci S. 2017. MicroRNA-142 is a multifaceted regulator in organogenesis, homeostasis, and disease. Dev Dyn 246(4):285-290.
61. Chaudhuri AD, Yelamanchili SV, Marcondes MC, Fox HS. 2013b. Up-regulation of microRNA-142 in simian immunodeficiency virus encephalitis leads to repression of sirtuin1. FASEB J 27(9):3720-9.
62. Chaudhuri AD, Yelamanchili SV, Fox HS. 2013a. MicroRNA-142 reduces monoamine oxidase A expression and activity in neuronal cells by downregulating SIRT1. PLoS One 8(11):e79579.
63. Caraci F, Spampinato S, Sortino MA, Bosco P, Battaglia G, Bruno V, Drago F, Nicoletti F, Copani A. 2012. Dysfunction of TGF-beta1 signaling in Alzheimer's disease: perspectives for neuroprotection. Cell Tissue Res 347(1):291-301.
64. Chen JH, Ke KF, Lu JH, Qiu YH, Peng YP. 2015. Protection of TGF-beta1 against neuroinflammation and neurodegeneration in Abeta1-42-induced Alzheimer's disease model rats. PLoS One 10(2):e0116549.
65. Flanders KC, Lippa CF, Smith TW, Pollen DA, Sporn MB. 1995. Altered expression of transforming growth factor-beta in Alzheimer's disease. Neurology 45(8):1561-9.
66. Lippa CF, Flanders KC, Kim ES, Croul S. 1998. TGF-beta receptors-I and -II immunoexpression in Alzheimer's disease: a comparison with aging and progressive supranuclear palsy. Neurobiol Aging 19(6):527-33.
67. Tesseur I, Zou K, Esposito L, Bard F, Berber E, Can JV, Lin AH, Crews L, Tremblay P, Mathews P and others. 2006. Deficiency in neuronal TGF-beta signaling promotes neurodegeneration and Alzheimer's pathology. J Clin Invest 116(11):3060-9.
68. Wyss-Coray T, Yan F, Lin AH, Lambris JD, Alexander JJ, Quigg RJ, Masliah E. 2002. Prominent neurodegeneration and increased plaque formation in complement-inhibited Alzheimer's mice. Proc Natl Acad Sci U S A 99(16):10837-42.
69. Talebi F, Ghorbani S, Chan WF, Boghozian R, Masoumi F, Ghasemi S, Vojgani M, Power C, Noorbakhsh F. 2017. MicroRNA-142 regulates inflammation and T cell differentiation in an animal model of multiple sclerosis. J Neuroinflammation 14(1):55.
70. Ando K, Brion JP, Stygelbout V, Suain V, Authelet M, Dedecker R, Chanut A, Lacor P, Lavaur J, Sazdovitch V and others. 2013. Clathrin adaptor CALM/PICALM is associated with neurofibrillary tangles and is cleaved in Alzheimer's brains. Acta Neuropathol 125(6):861-78.

71. Naj AC, Jun G, Reitz C, Kunkle BW, Perry W, Park YS, Beecham GW, Rajbhandary RA, Hamilton-Nelson KL, Wang LS and others. 2014. Effects of multiple genetic loci on age at onset in late-onset Alzheimer disease: a genome-wide association study. JAMA Neurol 71(11):1394-404.
72. Parikh I, Fardo DW, Estus S. 2014. Genetics of PICALM expression and Alzheimer's disease. PLoS One 9(3):e91242.
73. Zhao Z, Sagare AP, Ma Q, Halliday MR, Kong P, Kisler K, Winkler EA, Ramanathan A, Kanekiyo T, Bu G and others. 2015. Central role for PICALM in amyloid-beta blood-brain barrier transcytosis and clearance. Nat Neurosci 18(7):978-87.

SUPPLEMENTARY DATA

Table S1. Functional annotation of the 15 AD-associated ncRNA SNPs and their proxies in high LD ($R^2 > 0.7$) using HaploReg v4 (Excel file)
Table S2. Transcription factor binding sites (TFBSs) might be perturbing by rs2526377
Table S3. Expression of miR-142-3p and -5p in the human brain regions
Table S4. Downregulated target genes of miR-142-3p in human iPSC-derived NPCs
Table S5. Putative miR-142-3p target genes that are involved in AD-relevant pathways
Table S6. Target genes of miR-142 with the most significant association with AD
Table S7. Differentially expressed target genes of miR-142-3p in the hippocampus of KO mice vs Wt littermates
Table S8. IPA pathway analysis for 12 target genes of miR-142-3p upregulated in the hippocampus of KO mice
Figure S1. Regional association plots showing the association of 15 ncRNA SNPs with AD using the phase 3 of AD GWAS
Figure S2. Transcription factor binding sites overlap with rs2526377 (using UCSC browser)
Figure S3. KEGG pathways analysis for some of the putative miR-142-3p target genes (orange boxes) involved in AD

All supplementary data can be found online: doi 10.1002/humu.23872.