Herrera-Espejo, S., Santos-Zorrozua, B., Alvarez-Gonzalez, P. et al. *A Genome-Wide Study of Single-Nucleotide Polymorphisms in MicroRNAs and Further In Silico Analysis Reveals Their Putative Role in Susceptibility to Late-Onset Alzheimer's Disease*. **Mol Neurobiol** 58, 55–64 (2021). This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/s12035-020-02103-0

A genome-wide study of Single Nucleotide Polymorphisms in microRNAs and further *in silico* analysis reveals their putative role in susceptibility to Late-Onset Alzheimer's Disease.

Soraya Herrera-Espejo[1*], Borja Santos-Zorrozua[1*], Paula Alvarez-Gonzalez[1], Idoia Martin-Guerrero[1,2], Marian M. de Pancorbo[3], Africa Garcia-Orad[1,2**], and Elixabet Lopez-Lopez[1,2**]

[1]Department of Genetics, Physical Anthropology and Animal Physiology, Faculty of Medicine and Nursing, University of the Basque Country (UPV/EHU), Leioa, Spain
[2]BioCruces Bizkaia Health Research Institute, Barakaldo, Spain
[3]BIOMICs Research Group. Centro de Investigación "Lascaray" Ikergunea. University of the Basque Country (UPV/EHU).
[*]These authors share first authorship
[**]These authors share last authorship

Correspondence
Elixabet Lopez-Lopez
Department of Genetics, Physical Anthropology and Animal Physiology Faculty of Medicine and Nursing-University of the Basque Country (UPV/EHU) Barrio Sarriena s/n, 48940 Leioa, Spain.
Phone: international +34.946012951.
Fax: international +34.946013400 E-mail: elixabet.lopez@ehu.eus

Abstract

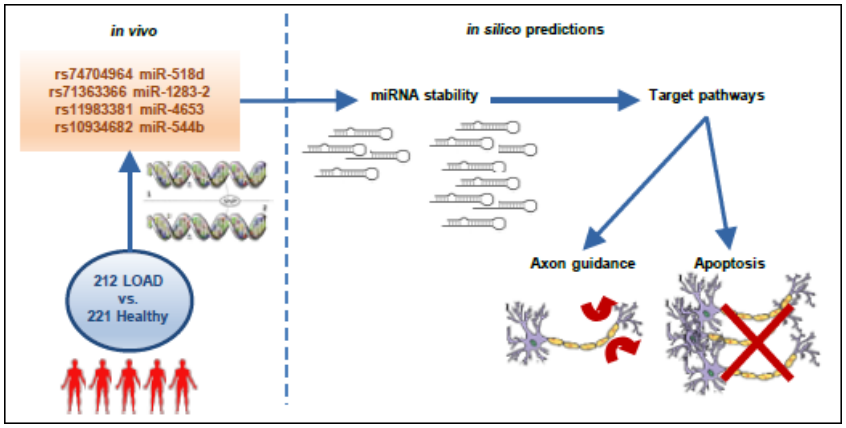
Aim: Late-onset Alzheimer's disease (LOAD) is a neurodegenerative disorder of growing relevance in aging societies for which predictive biomarkers are needed. Many genes involved in LOAD are tightly controlled by microRNAs (miRNAs), which can be modulated by Single Nucleotide Polymorphisms (SNPs). Our aim was to determine the association between SNPs in miRNAs and LOAD.

Methods: We selected all SNPs in pre-miRNAs with a minor allele frequency (MAF)>1% and genotyped them in a cohort of 229 individuals diagnosed with LOAD and 237 unrelated healthy controls. *In silico* analyses were performed to predict the effect of SNPs on miRNA stability and detect downstream pathways.

Results: Four SNPs were associated with LOAD risk with a p-value < 0.01 (rs74704964 in hsa-miR-518d, rs71363366 in hsa-miR-1283-2, rs11983381 in hsa-miR-4653, and rs10934682 in hsa-miR-544b). *In silico* analyses support a possible functional effect of those SNPs in miRNA levels and in the regulation of pathways of relevance for the development of LOAD.

Conclusion: Although the results are promising, additional studies are needed to validate the association between SNPs in miRNAs and the risk of developing LOAD.

Graphical abstract



Keywords

Single nucleotide polymorphisms, microRNAs, Late-Onset Alzheimer's Disease, susceptibility

1. Introduction

Alzheimer's disease (AD) is an irreversible progressive neurodegenerative disease that affects the central nervous system and encompasses > 80 % of dementia cases in people older than 65 [1] , harming more than 25 million people around the world [2].

Since AD has a preclinical stage that starts 15 to 20 years prior to diagnosis [3] and postmortem studies have revealed that the brains of people without diagnosis of cognitive impairment or with mild cognitive impairment often present similar histopathological hallmarks to AD patients [4-8], developing biomarkers to predict the disease in the earliest phases is very important [3].

Two types of AD can be distinguished with regard to patients' age: Early-Onset Alzheimer Disease (EOAD) and Late-Onset Alzheimer Disease (LOAD) [9, 10]. The former appears in people younger than 65 and LOAD is diagnosed in older people [9]. Regarding EOAD, several genetic markers have been established, this type of AD being a rare, autosomal, dominant, familial disease [11]. Conversely, diagnosis of LOAD is more difficult because of its complex pattern [11] and identifying biomarkers has proven to be much harder [12].

In order to identify genetic biomarkers in LOAD, most studies have focused on genetic variants located in coding regions [13-15]. However, in the last years, non-coding variants have also been related to this disease [1, 16]. One of the best known non-coding RNAs are microRNAs (miRNAs). MiRNAs are small molecules of single-stranded RNA of 21-25 nucleotides which are produced from hairpin-shaped precursors [17] . MiRNAs have the ability to regulate gene expression at post- transcriptional level through the binding of miRNA seed region to the 3'-untranslated region (3'-UTR) of their target mRNAs [18]. This way, they regulate more than 60 % of human genes, including those involved in each of the cellular changes that take place in LOAD. Thus, deregulation of miRNAs could be directly or indirectly involved in the pathogenesis of LOAD [19].

Single nucleotide polymorphisms (SNPs) are loci with alleles that differ in a single base, which are distributed in both coding and non-coding regions[20]. SNPs in miRNAs can modify their structure and stability, leading to alterations in miRNA levels if they are located in the precursor region, or changing their target mRNAs if they are located in the miRNA-mRNA binding region [21]. As a result, SNPs can interfere with miRNA function, which could be of relevance for the development of LOAD [22, 23].

So far, only a few studies have explored the effect of SNPs in miRNAs on the development of LOAD [10, 24-26]. In those studies, eight SNPs were associated with an increased risk of developing the disease, while 14 SNPs were described to have a protective effect against LOAD. However, none of those associations with LOAD were confirmed in other populations, pointing to a potential effect of miRNA-related SNPs in LOAD but warranting further research.

Therefore, the aim of this study was to analyze the association between SNPs located in miRNAs and the risk of developing LOAD in a Spanish cohort.

2. Materials and methods

2.1. Patients

A total of 433 individuals of European ancestry, including 229 individuals diagnosed with Late Onset Alzheimer Disease (LOAD) and 237 unrelated healthy controls from the collection

C.0001171 registered in the Institute of Health Carlos III, were included in this study (Table 1). LOAD patients were prospectively recruited by the Spanish National DNA Bank. The diagnosis of LOAD was based on a broad battery of neuropsychological tests: Minimental State Examination (MMSE), Clinical Dementia Rating scale (CDR), Consortium to Establish a Registry for Alzheimer's Disease (CERAD) protocol, Stroop test, unilateral and bilateral motor praxis, 7-minute test, trial making part A and B; and Neuropsychiatric Inventory (NPI). Clinical criteria for dementia and AD (Diagnostic and Statistical Manual of Mental Disorders, DSM-IV [27] and National Institute of Neurological and Communicative Disorders and the Alzheimer`s Disease and Related Disorders Association, NINCDS-ADRDA) [28] were used. Patients with a total score of less than 3 (1 and 2) on the CDR scale (mild to moderate dementia) were included. The exclusion criteria were the presence of previous cerebrovascular disorders (transient ischemic attacks, stroke or intracranial hemorrhage), other neurodegenerative diseases, severe comorbidity that made patient monitoring unlikely, manifestations of acute psychiatric disorders and the absence of a reliable informant. The figures obtained were corrected for age and education level. Sex and age data were systematically recorded blinded to genotypes from the patients' clinical records. Written informed consent was obtained from patients and/or their legally authorized representatives, as appropriate, and from healthy controls prior to sample collection. The study was approved by the Ethics Committee at the Spanish National DNA Bank (FO-24.01.1, 2012.07.02) and was carried out according to the Declaration of Helsinki.

2.2. Selection of miRNAs and polymorphisms

Since miRNAs are able to regulate a wide range of genes, which at present are not entirely defined, we selected all the SNPs described in pre-miRNAs with a MAF >0.01 in European/Caucasian populations until May 2014. SNP selection was performed using miRNA SNiPer (www.integratomics-time.com/miRNA-SNiPer/) . Of a total of 1910 SNPs that were identified in 969 miRNAs, we selected 213 SNPs in 201 pre-miRNAs which met the requirement of MAF >0.01 (Supplementary Table 1).

2.3. Genotyping

Genomic DNA was extracted from patients' and healthy controls' peripheral blood using standard phenol/chloroform extraction method. DNA was quantified using PicoGreen (Invitrogen Corp., Carlsbad, California, USA). For each sample, 400 ng of DNA were genotyped using the GoldenGate Genotyping Assay (Illumina Inc., San Diego, California, USA) with Veracode technology according to the published Illumina protocol. Data were analyzed using GenomeStudio (Illumina Inc.) software for genotype clustering and calling. Duplicate samples and CEPH trios (Coriell Cell Repository, Camden, New Jersey, USA) were genotyped across the plates. SNPs showing Mendelian allele transmission errors, discordant genotypes, or genotyping failure (no Polymerase Chain Reaction amplification, insufficient intensity for cluster separation, poor or no cluster definition, or genotyping obtention in <80 % of samples) were excluded from further analyses.

2.4. Statistical analysis

In order to detect any deviation from Hardy-Weinberg equilibrium, a chi-square test was carried out in the control population. To account for the possible confounding effect of age and sex, a t-test and a chi-square test were performed, respectively, to account for differences between patients and controls using GraphPad® Prism version 7.00 for Windows (GraphPad Software, La Jolla, California USA, www.graphpad.com). Association between LOAD and genetic polymorphisms was evaluated by logistic regression models and the effect sizes of the associations were estimated by the odds ratios (OR) with a 95 % confidence interval (95 % CI).

For each SNP, the most significant test among codominant (major allele homozygotes vs. heterozygotes and major allele homozygotes vs. minor allele homozygotes), dominant (major allele homozygotes vs. heterozygotes + minor allele homozygotes), recessive (major allele homozygotes + heterozygotes vs. minor allele homozygotes), and additive (doses-dependent effect: major allele homozygotes vs. heterozygotes vs. minor allele homozygotes) genetic models was selected. In all cases, the significance level was set at 5 %. The results were adjusted for multiple comparisons using the false discovery rate (FDR) correction, assuming 5 % level of significance for all tests [22]. Analyses were performed using the R 3.4.3 software [29] with the SNPassoc library [30].

2.5. Bioinformatics analysis

2.5.1. Prediction of miRNAs secondary structures

RNAfold web tool (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) [31] was used to estimate the impact of SNPs associated with LOAD on minimum free energy secondary structures and energy change (ΔΔG) of the hairpin structure of the miRNA.

2.5.2. Selection of target genes and pathways analysis

Predicted target genes for each of the miRNAs associated with LOAD were identified based on the miRWalk database (http://zmf.umm.uniheidelberg.de/apps/zmf/mirwalk2/index.html) [32]. Only those genes confirmed by six or more of the 12 prediction algorithms available in miRWalk were included.

Pathway enrichment analyses were performed with the over-representation analysis module of the ConsensusPathDB web tool (CPdB) [33, 34], assuming a conservative p-value cutoff of 0.0001.

3. Results

3.1. Genotyping results

Successful genotyping was achieved for 433 of 466 DNA samples (92.9 %), 212 out of 229 LOAD patients and 221 out of 237 controls (Table 1). A total of 130 out of 213 SNPs (61.03 %) were successfully genotyped and included in the association study after eliminating 60 SNPs showing genotyping failure, 11 SNPs that were monomorphic in our population, and 12 SNPs with deviations from HWE in the subsample of healthy controls (Supplementary Table 1).

3.2. Association study

To investigate whether genetic variation in miRNAs was linked to LOAD risk, the association between 130 SNPs in 126 miRNAs and this disease was analyzed. Since there were not statistically significant differences in age or sex distribution between LOAD patients and healthy controls in the study group (Table 1), these variables were not included as covariates. A total of 16 SNPs were associated with LOAD with a p-value < 0.05 (Supplementary Table 2). However, none of the SNPs reached statistical significance when FDR correction was applied. Thus, we focused on the four SNPs showing a p-value < 0.01 (Table 2).

The most significant result was found for rs74704964 in hsa-miR-518d (Table 2). CT genotype

displayed a 2.78-fold increased risk of LOAD (95 % CI: 1.30-5.95; p= 0.0056), in comparison with CC genotype, under the codominant model. No individual with the TT genotype was detected.

The second most significant result was found for rs71363366 in hsa-miR-1283-2. CG genotype was associated with a 2.64-fold increased risk of LOAD (95 % CI: 1.27 – 5.5; p = 0.0066), in comparison with CC genotype, under the codominant model. No individual with the GG genotype was detected.

The third most significant polymorphism was rs11983381 in hsa-miR-4653. The G allele was associated with a 1.66-fold increased risk of LOAD (95 % CI: 1.15 – 2.4; p = 0.0066) under the Log-additive model.

Finally, for rs10934682 located in hsa-miR-544b, the G allele presented a 1.64-fold increased risk of LOAD (95 % CI: 1.14 – 2.37; p = 0.007399) under the Log-additive model.

3.3. Bioinformatics analysis

3.3.1. Effect of genetic variants on the secondary structure of the pre-miRNAs

We analyzed *in silico* the energy changes (|ΔΔG|) and the modifications in secondary structures occurring as a consequence of the four SNPs most significantly associated with LOAD (p<0.01).

The four SNPs were located in the pre-miRNA sequences of their corresponding miRNAs. Although none of them produced major modifications in the secondary structure of the pre-miRNAs, all of them were associated with modest energy changes of the hairpin structure (Figure 1). Regarding rs74704964 in hsa-miR-518d, the change from the reference C allele to T allele caused a positive energy change (ΔΔG) of 2.5 kcal/mol (from -41.4 kcal/mol to -38.9 kcal/mol). Considering rs71363366 in hsa-miR-1283-2, the substitution of the C allele for a G allele induced a ΔΔG of 4 kcal/mol (from -37.7 kcal/mol to -33.7 kcal/mol). In rs11983381 in hsa-miR-4653, the substitution of the A allele for G allele induced an energy change of -4.7 kcal/mol (from -38.2 kcal/mol to -42.9 kcal/mol). Finally, in rs10934682 in hsa-miR-544b the change of the T allele for a G allele caused a ΔΔG of -3.3 kcal/mol (from -22.8 kcal/mol to -26.1 kcal/mol).

3.3.2. Target prediction and pathway analysis

We identified the predicted target genes for the miRNAs harboring the four most significant SNPs and, within those targets, we evaluated the overrepresented pathways with a plausible role in LOAD.

Regarding the predicted target genes of hsa-miR-518d, we found that axon guidance (Reactome; $p = 3.86 \cdot 10^{-06}$) (Supplementary Table 3) was the most statistically significantly overrepresented pathway.

In the case of hsa-miR-1283, we identified FoxO signaling pathway (KEGG; $p = 9.51 \cdot 10^{-06}$) among the ten most significant pathways (Supplementary Table 4).

Regarding hsa-miR-4653, we found that Axon guidance (Reactome; $p = 6.13 \cdot 10^{-08}$), BDNF signaling pathway (Wikipathways; $p = 1.33 \cdot 10^{-06}$) and Axon guidance (KEGG; $p = 2.48 \cdot 10^{-06}$) were among the five most significant pathways (Supplementary Table 5).

Finally, concerning hsa-miR-544b, apoptotic signaling pathway (Wikipathways; $p = 3.51 \cdot 10^{-07}$) and apoptosis (Wikipathways; $p = 5.34 \cdot 10^{-07}$) were among the five most significant pathways (Supplementary Table 6).

4. Discussion

In this study, we evaluated the role of SNPs in miRNAs in the risk of developing LOAD. In order to achieve this aim, we analyzed 130 SNPs in 129 miRNAs in a population of 212 LOAD patients and 221 healthy controls. Four SNPs were associated with the risk of developing LOAD with a p-value < 0.01. Although none of them remained statistically significant after FDR correction, *in silico* analyses support a plausible functional effect of those SNPs in miRNA levels and the regulation of pathways of relevance in the development of LOAD.

The SNP that showed the most significant association with LOAD was rs74704964, located in the pre-miRNA of hsa-miR-518d. For this SNP, CT genotype presented a 2.78-fold increased risk of developing LOAD under a codominant model. This SNP had not been previously studied in LOAD. Interestingly, the presence of the T allele in rs74704964 is predicted to induce an energy change of 2.50 kcal/mol in the secondary structure of the pre-miRNA. Positive energy changes have been associated with decreased stability and, therefore, the presence of the T allele could lead to a decrease in mature hsa-miR-518d levels [18] which would result in overexpression of its target genes (Figure 2a). Therefore, we performed a deep analysis of the target genes of this miRNA in order to search for enriched pathways with a plausible role in Alzheimer development, which would explain a possible role of hsa-miR-518d depletion in LOAD risk.

On the one hand, hsa-miR-518d has been previously shown to target peroxisome proliferator‑activated receptor-α (PPARα), PPARs being decisive in AD related processes, such as inflammation and diabetic phenotype [35, 36].This could be pointing to a relevant role of this miRNA in the disease. On the other hand, axon guidance pathway is overrepresented among the predicted target genes of hsa-miR- 518d (Supplementary Fig. 1). Among the seven processes controlled by this pathway, the predicted targets of hsa-miR-518d are mainly involved in L1 cell adhesion molecule (L1CAM) interactions and ephrin-ephrin (EPH-EPHRIN) signaling. On the one hand, L1CAM, is implicated in cell adhesion in the central nervous system [37], being involved in neuronal growth [38] and stimulating neuronal survival [39]. Therefore, a deregulation genes in this sub-pathway could imply an increase in apoptosis or aberrant neural development. On the other hand, EPH-EPHRIN signaling is involved in synaptic development [40] and in inflammatory and apoptotic processes [41]. Thus, the deregulation of target genes in this sub- pathway could interfere with the number and/or quality of synapses and neurons. As a result, all of this could help understand a putative role of changes in hsa-miR-518d levels in the development of LOAD.

Rs71363366 CG genotype in hsa-miR-1283-2 was paired with a 2.64-fold increase in the risk of LOAD development under a codominant model. This SNP, which had not been previously studied in LOAD, is located in the pre-miRNA sequence and, when the G allele is present, an energy change of the hairpin structure of the pre-miRNA of 4 kcal/mol is predicted. As mentioned above, such positive energy changes have been associated with decreased stability in the pre-miRNA structure and, therefore, the mature miRNA product would be diminished [21], which could result in an overexpression of its target genes (Figure 2b).

Inhibition of hsa-miR-1283 has been previously shown to promote endoplasmic reticulum stress and apoptosis through its role in the regulation of Activating Transcription Factor 4 (*ATF4*) expression in endothelial cells. This low expression of hsa-miR-1283 produces an overexpression of ATF4 which lead to vascular damage [42, 43]. Thus, depletion in hsa-miR-1283 expression associated with the presence of the G allele in rs7163366 could promote vascular damage, which

could explain its role in LOAD risk. In addition, the analysis *in silico* showed that Forkhead box O (FOXO) signaling pathway is overrepresented among target genes of hsa-miR-1283-2 (Supplementary Figure 2). Thus, genes such as *FOXO1* and *FOXO3* could be upregulated due to the presence of rs71363366 G allele in hsa-miR-1283. Although FOXO family has been related with a protective role in the age of AD development [44], once the disease has initiated its development, those genes have been shown to have an apoptotic function [45, 46]. Therefore, an upregulation of the FOXO family genes could trigger neuronal death and loss of neurons, some of the common signs of AD [47], which could be an alternative mechanism of action.

Regarding rs11983381, located in hsa-miR-4653, which presented the third most significant association with LOAD, the G allele showed a 1.64-fold increased risk of LOAD under a log-additive model. However, this SNP was not associated with LOAD in a previous study [25]. In this case, there is a predicted negative energy change in the pre-miRNA structure (ΔΔG = -4.7 kcal/mol), which would lead to a higher stability of hsa-miR-4653. Thus, there could be higher levels of mature miRNA [21] and its target genes could present a decreased expression (Figure 2c). Axon guidance pathway is overrepresented among the predicted target genes of hsa-miR-4653 (Supplementary Figure 3). The main function of this pathway is the regulation of three basic processes: axonal growth, axonal repulsion and axonal attraction. Therefore, upregulated hsa-miR-4653 could inhibit axonal growth or promote the presence of axons in inaccurate places. This could be linked to the existence of degenerated dendrites and axons in the Aβ plaques [48], a key characteristic of AD affected neurons [49].

Furthermore, *in silico* analysis determined that hsa-miR-4653 also targeted 27 genes located in the Brain-Derived Neurotrophic Factor (BDNF) signaling pathway (Supplementary Figure 4), which could be downregulated when the rs11983381 G allele is present. Among them, Neurotrophic Receptor Tyrosine Kinase 2 (*NTRK2*) gene plays an important role in the development of disease-controlling neuronal apoptosis [50]. In addition, downregulation of *NTRK2* and *NTRK3* expression has been observed in the neurons of the *Nucleus Basalis* during the progression of AD [51, 52]. Moreover, there is a selective loss of catalytic *NTRKB* immunoreactivity in both the temporal and frontal cortex of AD patients when compared to normal brains [53]. Therefore, the inhibition of this pathway could be a plausible mechanism linking rs11983381 and LOAD development.

Finally, rs10934682 was located in hsa-miR-544b and the G allele showed a 1.64-fold increased risk of LOAD according to a log-additive model. However, this SNP was not associated with LOAD in a previous study [25]. This allele induced an energy change of (ΔΔG) of -3.3 kcal/mol, in the pre-miRNA hairpin structure. This means that the pre-miRNA could be more stable, as mentioned above (Figure 2d). Consequently, mature miRNA product of hsa-miR-544b could be increased and the expression of its target genes would be decreased. Interestingly, hsa-miR-544 has been associated with the inhibition of tumor proliferation in breast cancer and glioma through the targeting of apoptosis-related proteins [54, 55]. In fact, hsa-miR-544b has numerous predicted target genes located in the apoptosis pathway (Supplementary Figure 5). Among them, B-cell lymphoma 2 (*BCL2*), which has an anti-apoptotic function [56], has been shown to be downregulated in brains with AD [57], which would match an upregulation of hsa-miR-544b. Likewise, *BCL2L1* induces apoptosis when it is underexpressed [58]. Additionally, the X-linked inhibitor of apoptosis protein (*XIAP*) has also been described as an antiapoptotic protein [59, 60] that suppresses the neuronal cell death process [61]. Therefore, overexpression of hsa-miR-544b seems to have a relevant role in apoptosis promotion, which could be a plausible explanation for the involvement of rs10934682 in LOAD risk.

Previously, a total of only five studies had been performed analyzing the role of SNPs in miRNAs in LOAD risk [10, 24-26, 62]. A total of 22 SNPs were associated with the development of LOAD.

Two of them were identified in specific studies on those SNPs in particular [24, 62] while the remaining 20 were identified in a more comprehensive study in which 237 SNPs were studied obtaining data from a large meta-analysis [25]. Thus, we present a review of the most significant results of those studies below.

The first of them, rs2910164, is a polymorphism located in hsa-miR-146a precursor in which the minor allele was associated with an increased risk of developing LOAD, performing both allelic and codominant inheritance genotypic analyses [62]. Supporting this association, functional studies revealed that the presence of the C risk allele led to a decreased expression of hsa-miR-146a *in vitro* and *in vivo*. Furthermore, Toll Like Receptor 2 (*TLR2*), a gene involved in neuroinflammatory activation and AD pathogenesis, was identified as a target of hsa-miR-146a *in vitro* providing a plausible functional explanation for the role of the miRNA in LOAD [62]. However, the association between rs2019164 and LOAD was not replicated in other two studies in which it was analyzed [10, 25] and it was not significantly associated with LOAD in the present study either. As a result, the implication of rs2910164 in the development of the disease cannot be confirmed.

Rs11014002 is a SNP located in the hsa-miR-603 precursor. It has been published that carriers of the T allele present a decreased risk of developing LOAD [24]. In addition, the T protective allele was shown to increase the expression of mature hsa-miR-603 *in vitro*, which could be an explanation for a possible role on LOAD. In fact, hsa-miR-603 has been shown to regulate genes of relevance for LOAD development, such as Low density lipoprotein Receptor-related Protein-Associated Protein 1 (*LRPAP1*), Low density lipoprotein Receptor-related Protein 1 (*LRP1*) and E2 Promoter Binding Factor 1 (*E2F1*) *[24]*. However, once again, the association of the SNP with LOAD was not replicated in the other study in which it was evaluated [25]. In this case, it should be noted that different inheritance models were analyzed, dominant and additive respectively, which may be contributing to the discrepancies in results. In our study, this SNP was not analyzed due to genotyping failure,.

Finally, the study performed by Ghanbari *et al* is the largest study on genetic variants in miRNAs, studying 237 SNPs in using data from a genome-wide meta-analysis of 37,154 controls and 17,008 LOAD patients [25]. Among them, 20 SNPs were found to be significantly associated with the development of LOAD. Rs2291418, in hsa-miR-1229, was the most relevant, remaining significantly associated with LOAD after Bonferroni correction, and the G allele being associated with an increased risk of developing the disease. In fact, the presence of the G allele was shown to increase the levels of mature hsa-miR-1229 *in vitro*. Several predicted hsa-miR-1229 target genes were reported to be involved in neurological pathways and disease. Among them, Sortilin Related Receptor 1 (*SORL1*) was demonstrated to be directly regulated by hsa-miR-1229- 3p *in vitro* [22]. Since both *SORL1* and hsa-miR-1229-3p are highly expressed in the brain, that regulatory effect is also expected in that tissue *in vivo*. Interestingly, decreased *SORL1* expression has been reported to be mechanistically involved in AD [63]. Therefore, increased levels of hsa-miR-1229-3p in rs2291418 G allele carriers could decrease *SORL1* levels and thus explain an increased risk of developing LOAD [25]. Nevertheless, this SNP had not been included in additional studies to further confirm this result and it is not associated with LOAD in the present study.

All of these results together warrant the need for additional studies regarding the role of miRNA-related SNPs in LOAD.

The current study presents some limitations that have to be addressed. For example, the relatively high failure rate of the genotyping technique. Nevertheless, this chance of failure was assumed from the beginning of the study because there was a low predicted score for

genotyping a number of the SNPs but no other design option to amplify the polymorphisms was possible using the GoldenGate Genotyping Assay. Another possible limitation is that these results did not reach a significant *p*-value when FDR correction was applied and further replication in additional populations would be required. This could be due to the low frequency of SNPs in miRNAs, because of their location in conserved regions [64]. In order to mitigate this, only those SNPs that showed the strongest association with the disease were selected ($p<0.01$) and we would like to emphasize that *in silico* analyses and previous literature provided plausible functional mechanisms through which those SNPs may be involved in LOAD development. In addition, we have to keep in mind the inaccuracy of the prediction algorithms of the databases used to define miRNA target genes [65, 66], but, unfortunately, nowadays this constraint has to be assumed. Finally, it is also important to highlight that it is possible that the association of the SNPs with LOAD might be independent of their possible role on the miRNAs and may be mediated by other cis or trans-regulatory mechanisms.

In spite of these limitations, our results, in combination with those of the other few studies performed so far, are very interesting since they open a new field of study with implications in the establishment of new biomarkers for LOAD. It would be of great interest to perform additional studies to validate these results and search for additional SNPs in miRNAs involved in LOAD. Besides, deep molecular studies would be of interest in order to demonstrate the biological effect of these SNPs in LOAD with the final aim to use these SNPs as high-fidelity biological identifiers of LOAD.

5. Conclusions

In conclusion, in this study we found four SNPs rs74704964, rs71363366, rs11983381 and rs10934682 located in the pre-miRNA structure of hsa-miR-518d, hsa-miR-1283-2, hsa-miR-4653 and hsa-miR-544b respectively, which could be associated with the risk of developing LOAD. These SNPs could change the expression levels of the miRNAs and, consequently, affect the expression of genes implicated in several biological processes associated with LOAD, as shown by *in silico* analyses. These results could indicate a potential involvement of these four SNPs in miRNAs with the development of LOAD. It would be interesting to further explore these results in other populations of LOAD patients.

6. Acknowledgements

This study was funded by the Basque Government (IT989-16). The authors would like to thank Estibaliz Cuesta Durán for her contribution regarding sample preparation. Support by the Spanish National Genotyping Center (CeGen) is also gratefully acknowledged.

7. Disclosure Statement

The authors declare that they have no conflict of interest.

8. References

1. Maoz R, Garfinkel BP, Soreq H. Alzheimer's Disease and ncRNAs. Adv Exp Med Biol. 2017;978:337-61; doi: 10.1007/978-3-319-53889-1_18.
2. Basavaraju M, de Lencastre A. Alzheimer's disease: presence and role of microRNAs. Biomol Concepts. 2016;7 4:241-52; doi: 10.1515/bmc-2016-0014.
3. Counts SE, Ikonomovic MD, Mercado N, Vega IE, Mufson EJ. Biomarkers for the Early Detection and Progression of Alzheimer's Disease. Neurotherapeutics. 2017;14 1:35-53; doi: 10.1007/s13311-016-0481-z.
4. Morris JC, Storandt M, Miller JP, McKeel DW, Price JL, Rubin EH, et al. Mild cognitive impairment represents early-stage Alzheimer disease. Arch Neurol. 2001;58 3:397-405.
5. Price JL, Morris JC. Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. Ann Neurol. 1999;45 3:358-68.
6. Bennett DA, Schneider JA, Bienias JL, Evans DA, Wilson RS. Mild cognitive impairment is related to Alzheimer disease pathology and cerebral infarctions. Neurology. 2005;64 5:834-41; doi: 10.1212/01.WNL.0000152982.47274.9E.
7. Markesbery WR, Schmitt FA, Kryscio RJ, Davis DG, Smith CD, Wekstein DR. Neuropathologic substrate of mild cognitive impairment. Arch Neurol. 2006;63 1:38-46; doi: 10.1001/archneur.63.1.38.
8. Mufson EJ, Chen EY, Cochran EJ, Beckett LA, Bennett DA, Kordower JH. Entorhinal cortex beta-amyloid load in individuals with mild cognitive impairment. Exp Neurol. 1999;158 2:469-90; doi: 10.1006/exnr.1999.7086.
9. Tellechea P, Pujol N, Esteve-Belloch P, Echeveste B, García-Eulate MR, Arbizu J, et al. Early- and late-onset Alzheimer disease: Are they the same entity? Neurologia. 2018;33 4:244-53; doi: 10.1016/j.nrl.2015.08.002.
10. Cui L, Li Y, Ma G, Wang Y, Cai Y, Liu S, et al. A functional polymorphism in the promoter region of microRNA-146a is associated with the risk of Alzheimer disease and the rate of cognitive decline in patients. PLoS One. 2014;9 2:e89019; doi: 10.1371/journal.pone.0089019.
11. Kim DH, Yeo SH, Park JM, Choi JY, Lee TH, Park SY, et al. Genetic markers for diagnosis and pathogenesis of Alzheimer's disease. Gene. 2014;545 2:185-93; doi: 10.1016/j.gene.2014.05.031.
12. Roy J, Mallick B. Altered gene expression in late-onset Alzheimer's disease due to SNPs within 3'UTR microRNA response elements. Genomics. 2017;109 3-4:177-85; doi: 10.1016/j.ygeno.2017.02.006.
13. Rezazadeh M, Khorrami A, Yeghaneh T, Talebi M, Kiani SJ, Heshmati Y, et al. Genetic Factors Affecting Late-Onset Alzheimer's Disease Susceptibility. Neuromolecular Med. 2016;18 1:37-49; doi: 10.1007/s12017-015-8376-4.
14. De Luca V, Spalletta G, Souza RP, Graff A, Bastos-Rodrigues L, Camargos Bicalho MA. Definition of Late Onset Alzheimer's Disease and Anticipation Effect of Genome-Wide Significant Risk Variants: Pilot Study of the APOE e4 Allele. Neuropsychobiology. 2019;77 1:8-12; doi: 10.1159/000490739.
15. Liu G, Sun JY, Xu M, Yang XY, Sun BL. SORL1 Variants Show Different Association with Early-Onset and Late-Onset Alzheimer's Disease Risk. J Alzheimers Dis. 2017;58 4:1121-8; doi: 10.3233/JAD-170005.
16. Martinez B, Peplow PV. MicroRNAs as diagnostic and therapeutic tools for Alzheimer's disease: advances and limitations. Neural Regen Res. 2019;14 2:242-55; doi: 10.4103/1673-5374.244784.
17. Wahid F, Shehzad A, Khan T, Kim YY. MicroRNAs: synthesis, mechanism, function, and recent clinical trials. Biochim Biophys Acta. 2010;1803 11:1231-43; doi: 10.1016/j.bbamcr.2010.06.013.
18. Guo H, Ingolia NT, Weissman JS, Bartel DP. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature. 2010;466 7308:835-40; doi: 10.1038/nature09267.
19. Shah SZA, Zhao D, Hussain T, Sabir N, Yang L. Regulation of MicroRNAs-Mediated Autophagic Flux: A New Regulatory Avenue for Neurodegenerative Diseases With Focus on Prion Diseases. Front Aging Neurosci. 2018;10:139.
20. Chakravarti A. Population genetics--making sense out of sequence. Nat Genet. 1999;21 1

Suppl:56-60; doi: 10.1038/4482.
21. Gong J, Tong Y, Zhang HM, Wang K, Hu T, Shan G, et al. Genome-wide identification of SNPs in microRNA genes and the SNP effects on microRNA target binding and biogenesis. Hum Mutat. 2012;33 1:254-63; doi: 10.1002/humu.21641.
22. Han Z, Huang H, Gao Y, Huang Q. Functional annotation of Alzheimer's disease associated loci revealed by GWASs. PLoS One. 2017;12 6:e0179677; doi: 10.1371/journal.pone.0179677.
23. Lau P, Bossers K, Janky Rs, Salta E, Frigerio CS, Barbash S, et al. Alteration of the microRNA network during the progression of Alzheimer's disease. Embo Molecular Medicine. 2013;5 10:1613-34.
24. Zhang C, Lu J, Liu B, Cui Q, Wang Y. Primate-specific miR-603 is implicated in the risk and pathogenesis of Alzheimer's disease. Aging (Albany NY). 2016;8 2:272-90; doi: 10.18632/aging.100887.
25. Ghanbari M, Ikram MA, de Looper HWJ, Hofman A, Erkeland SJ, Franco OH, et al. Genome-wide identification of microRNA-related variants associated with risk of Alzheimer's disease. Sci Rep. 2016;6:28387; doi: 10.1038/srep28387.
26. Qi L, Hu Y, Zhan Y, Wang J, Wang BB, Xia HF, et al. A SNP site in pri-miR-124 changes mature miR-124 expression but no contribution to Alzheimer's disease in a Mongolian population. Neurosci Lett. 2012;515 1:1-6; doi: 10.1016/j.neulet.2012.02.061.
27. Cooper JE. On the publication of the Diagnostic and Statistical Manual of Mental Disorders: Fourth Edition (DSM-IV). Br J Psychiatry. 1995;166 1:4-8; doi: 10.1192/bjp.166.1.4.
28. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984;34 7:939-44; doi: 10.1212/wnl.34.7.939.
29. Team RC: R: A language and environment for statistical computing. In.; 2015.
30. González JR, Armengol L, Solé X, Guinó E, Mercader JM, Estivill X, et al. SNPassoc: an R package to perform whole genome association studies. Bioinformatics. 2007;23 5:644-5; doi: 10.1093/bioinformatics/btm025.
31. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA websuite. Nucleic Acids Res. 2008;36 Web Server issue:W70-4; doi: 10.1093/nar/gkn188.
32. Dweep H, Sticht C, Pandey P, Gretz N. miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform. 2011;44 5:839-47; doi: 10.1016/j.jbi.2011.05.002.
33. Kamburov A, Wierling C, Lehrach H, Herwig R. ConsensusPathDB--a database for integrating human functional interaction networks. Nucleic Acids Res. 2009;37 Database issue:D623-8; doi: 10.1093/nar/gkn698.
34. Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R. ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Res. 2011;39 Database issue:D712-7; doi: 10.1093/nar/gkq1156.
35. Yilmaz-Aydogan H, Kurnaz O, Kucukhuseyin O, Akadam-Teker B, Kurt O, Eronat AP, et al. Different effects of PPARA, PPARG and ApoE SNPs on serum lipids in patients with coronary heart disease based on the presence of diabetes. Gene. 2013;523 1:20-6; doi: 10.1016/j.gene.2013.03.136.
36. Zhao C, Zhang T, Shi Z, Ding H, Ling X. MicroRNA-518d regulates PPARα protein expression in the placentas of females with gestational diabetes mellitus. Mol Med Rep. 2014;9 6:2085-90; doi: 10.3892/mmr.2014.2058.
37. Rathjen FG, Schachner M. Immunocytological and biochemical characterization of a new neuronal cell surface component (L1 antigen) which is involved in cell adhesion. EMBO J. 1984;3 1:1-10.
38. Chang S, Rathjen FG, Raper JA. Extension of neurites on axons is impaired by antibodies against specific neural cell surface glycoproteins. J Cell Biol. 1987;104 2:355-62.
39. Conacci-Sorrell M, Kaplan A, Raveh S, Gavert N, Sakurai T, Ben-Ze'ev A. The shed ectodomain of Nr-CAM stimulates cell proliferation and motility, and confers cell transformation. Cancer Res. 2005;65 24:11605-12; doi: 10.1158/0008-5472.CAN-05-2647.
40. Pitulescu ME, Adams RH. Eph/ephrin molecules--a hub for signaling and endocytosis. Genes Dev. 2010;24 22:2480-92; doi: 10.1101/gad.1973910.

41. Smith JA, Das A, Ray SK, Banik NL. Role of pro-inflammatory cytokines released from microglia in neurodegenerative diseases. Brain Res Bull. 2012;87 1:10-20; doi: 10.1016/j.brainresbull.2011.10.004.
42. He L, Yuan J, Xu Q, Chen R, Chen L, Fang M. miRNA-1283 Regulates the PERK/ATF4 Pathway in Vascular Injury by Targeting ATF4. PLoS One. 2016;11 8:e0159171; doi: 10.1371/journal.pone.0159171.
43. He L, Fang M, Chen L, Zhou J, Yuan J, Xu J, et al. Transcriptome analysis of blood stasis syndrome in subjects with hypertension. J Tradit Chin Med. 2016;36 2:173-80; doi: 10.1016/s0254-6272(16)30024-3.
44. Guo J, Cheng J, North BJ, Wei W. Functional analyses of major cancer-related signaling pathways in Alzheimer's disease etiology. Biochim Biophys Acta Rev Cancer. 2017;1868 2:341-58; doi: 10.1016/j.bbcan.2017.07.001.
45. Murtaza G, Khan AK, Rashid R, Muneer S, Hasan SMF, Chen J. FOXO Transcriptional Factors and Long-Term Living. Oxid Med Cell Longev. 2017;2017:3494289; doi: 10.1155/2017/3494289.
46. Wang Y, Zhou Y, Graves DT. FOXO transcription factors: their clinical significance and regulation. Biomed Res Int. 2014;2014:925350; doi: 10.1155/2014/925350.
47. Storz P. Forkhead homeobox type O transcription factors in the responses to oxidative stress. Antioxid Redox Signal. 2011;14 4:593-605; doi: 10.1089/ars.2010.3405.
48. Nelson PT, Alafuzoff I, Bigio EH, Bouras C, Braak H, Cairns NJ, et al. Correlation of Alzheimer disease neuropathologic changes with cognitive status: a review of the literature. J Neuropathol Exp Neurol. 2012;71 5:362-81; doi: 10.1097/NEN.0b013e31825018f7.
49. Nunomura A, Castellani RJ, Zhu X, Moreira PI, Perry G, Smith MA. Involvement of oxidative stress in Alzheimer disease. J Neuropathol Exp Neurol. 2006;65 7:631-41.
50. Chen Z, Simmons MS, Perry RT, Wiener HW, Harrell LE, Go RC. Genetic association of neurotrophic tyrosine kinase receptor type 2 (NTRK2) With Alzheimer's disease. Am J Med Genet B Neuropsychiatr Genet. 2008;147 3:363-9; doi: 10.1002/ajmg.b.30607.
51. Connor B, Young D, Lawlor P, Gai W, Waldvogel H, Faull RL, et al. Trk receptor alterations in Alzheimer's disease. Brain Res Mol Brain Res. 1996;42 1:1-17.
52. Ginsberg SD, Che S, Wuu J, Counts SE, Mufson EJ. Down regulation of trk but not p75NTR gene expression in single cholinergic basal forebrain neurons mark the progression of Alzheimer's disease. J Neurochem. 2006;97 2:475-87; doi: 10.1111/j.1471-4159.2006.03764.x.
53. Allen SJ, Wilcock GK, Dawbarn D. Profound and selective loss of catalytic TrkB immunoreactivity in Alzheimer's disease. Biochem Biophys Res Commun. 1999;264 3:648-51; doi: 10.1006/bbrc.1999.1561.
54. Zhu Z, Wang S, Zhu J, Yang Q, Dong H, Huang J. MicroRNA-544 down-regulates both Bcl6 and Stat3 to inhibit tumor growth of human triple negative breast cancer. Biol Chem. 2016;397 10:1087-95; doi: 10.1515/hsz-2016-0104.
55. Jin S, Dai Y, Li C, Fang X, Han H, Wang D. MicroRNA-544 inhibits glioma proliferation, invasion and migration but induces cell apoptosis by targeting PARK7. Am J Transl Res. 2016;8 4:1826-37.
56. Chipuk JE, Moldoveanu T, Llambi F, Parsons MJ, Green DR. The BCL-2 family reunion. Mol Cell. 2010;37 3:299-310; doi: 10.1016/j.molcel.2010.01.025.
57. Zhao S, Zhao J, Zhang T, Guo C. Increased apoptosis in the platelets of patients with Alzheimer's disease and amnestic mild cognitive impairment. Clin Neurol Neurosurg. 2016;143:46-50; doi: 10.1016/j.clineuro.2016.02.015.
58. Castora FJ, Liu TY, Gershon B, Shugoll J, Vora JS, Coleman RA. Brains of Alzheimer's patients show expression changes in five genes affecting mitochondrial morphology, membrane and neuronal apoptosis that directly or indirectly interact with APP. 2017. doi: https://doi.org/10.1016/j.jalz.2017.06.798.
59. Morizane Y, Honda R, Fukami K, Yasuda H. X-linked inhibitor of apoptosis functions as ubiquitin ligase toward mature caspase-9 and cytosolic Smac/DIABLO. J Biochem. 2005;137 2:125-32; doi: 10.1093/jb/mvi029.
60. Saraei R, Soleimani M, Movassaghpour Akbari AA, Farshdousti Hagh M, Hassanzadeh A, Solali S. The role of XIAP in resistance to TNF-related apoptosis-inducing ligand (TRAIL) in Leukemia. Biomed Pharmacother. 2018;107:1010-9; doi: 10.1016/j.biopha.2018.08.065.
61. Kato K, Tanaka T, Sadik G, Baba M, Maruyama D, Yanagida K, et al. Protein kinase C stabilizes X-

linked inhibitor of apoptosis protein (XIAP) through phosphorylation at Ser(87) to suppress apoptotic cell death. Psychogeriatrics. 2011;11 2:90-7; doi: 10.1111/j.1479-8301.2011.00355.x.
62. Zhang B, Wang A, Xia C, Lin Q, Chen C. A single nucleotide polymorphism in primary-microRNA-146a reduces the expression of mature microRNA-146a in patients with Alzheimer's disease and is associated with the pathogenesis of Alzheimer's disease. Mol Med Rep. 2015;12 3:4037-42; doi: 10.3892/mmr.2015.3968.
63. Vardarajan BN, Faber KM, Bird TD, Bennett DA, Rosenberg R, Boeve BF, et al. Age-specific incidence rates for dementia and Alzheimer disease in NIA-LOAD/NCRAD and EFIGA families: National Institute on Aging Genetics Initiative for Late-Onset Alzheimer Disease/National Cell Repository for Alzheimer Disease (NIA-LOAD/NCRAD) and Estudio Familiar de Influencia Genetica en Alzheimer (EFIGA). JAMA Neurol. 2014;71 3:315-23; doi: 10.1001/jamaneurol.2013.5570.
64. Mishra PJ, Banerjee D, Bertino JR. MiRSNPs or MiR-polymorphisms, new players in microRNA mediated regulation of the cell: Introducing microRNA pharmacogenomics. Cell Cycle. 2008;7 7:853-8; doi: 10.4161/cc.7.7.5666.
65. Lee YJ, Kim V, Muth DC, Witwer KW. Validated MicroRNA Target Databases: An Evaluation. Drug Dev Res. 2015;76 7:389-96; doi: 10.1002/ddr.21278.
66. Akhtar MM, Micolucci L, Islam MS, Olivieri F, Procopio AD. Bioinformatic tools for microRNA dissection. Nucleic Acids Res. 2016;44 1:24-44; doi: 10.1093/nar/gkv1221.

9. Figure legends

Fig. 1 Energy changes (ΔΔG) and minimum free-energy structures of the four most significant miRNA SNPs. a. hsa-miR-518d, **b.** hsa-miR-1283, **c.**hsa-miR-4653 and **d.** hsa-miR-544b. For each miRNA, the SNP alleles are marked with a circle and the arrow points out the variant allele. miRNA structure bordered with a rectangle was the most stable

Fig. 2 Proposed mechanism for LOAD risk for each of the SNPs in miRNAs. a. rs74704964, **b.** rs71363366, **c.** rs11983381, **d.** rs10934682

10. Table legends

Table 1. Characteristics of the study population.

Table 2. SNPs in miRNAs significantly associated with LOAD.

Figures

Figure 1

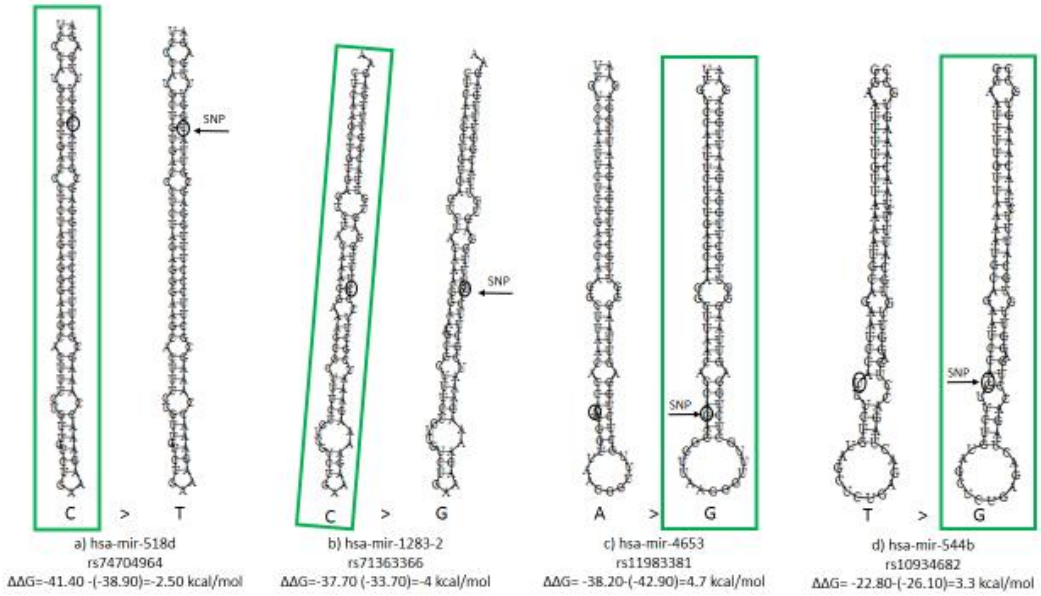
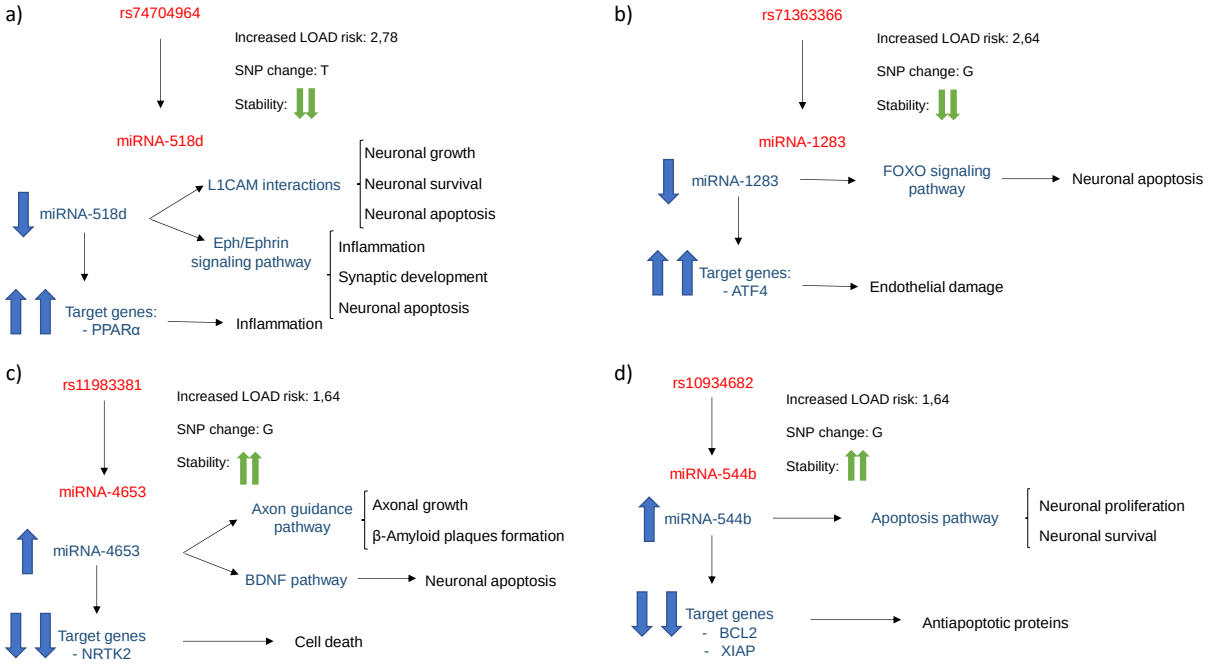


Figure 2



Tables

Table 1. Characteristics of the study population.

	LOAD patients	Healthy controls	p-value
N	212	221	
Sex (Male / Female)	67 (31,61 %) / 145 (68,39 %)	89 (40,27 %) / 132 (59,73 %)	0.0603
Age (mean ± SD)	78,99 ± 6,40	78,21 ± 18,49	0.5611

Table 2. SNPs in miRNAs significantly associated with LOAD (p<0.01)

SNP	Location	Genotype	N (%) Controls (N=221)	N (%) Cases (N=212)	OR (IC 95 %)	P-value
rs74704964	hsa-miR-518d	C/C	207 (95,4)	186 (88,2)	Codominant	0,005584
	19q13.42					
	pre-miRNA	C/T	10 (4,6)	25 (11,8)	2,78 (1,30-5,95)	
rs71363366	hsa-miR-1283-2	C/C	207 (95,0)	185 (87,7)	Codominant	0,006592
	19q13.42					
	pre-miRNA	C/G	11 (5,0)	26 (12,3)	2,64 (1,27-5,5)	
rs11983381	hsa-miR-4653	A/A	163 (73,8)	132 (63,5)	Log-additive	0,006604
	7q22.1	A/G	56 (25,3)	67 (32,2)	1,66 (1,15 - 2,4)	
	pre-miRNA	G/G	2 (0,9)	9 (4,3)		
rs10934682	hsa-miR-544b	T/T	162(73,3)	130 (62,2)	Log-additive	0,007399
	3q21.2	G/T	56 (25,3)	71 (34)	1,64 (1,14 - 2,37)	
	pre-miRNA	G/G	3 (1,4)	8 (3,8)		