The Genetic Research in Alzheimer Disease (GERALD) Initiative Finds rs9320913 as a Neural eQTL of lincRNA AL589740.1

1. Introduction

Alzheimer’s disease (AD) diagnosis involves a significant decline of intellectual abilities in one or more cognitive domains including learning and memory, language, and executive functions [1, 2]. A study conducted by Stern et al. [3] indicated that individuals with a low level of school-
previous genetic correlations suggest protective effects of the cognitive reserve for Alzheimer’s disease [5]. Different elements are associated with a greater cognitive reserve including the educational level and occupational and physical activities [6, 7]. Intelligence is a highly heritable trait and a net contributor to individual cognitive reserve [8, 9]. Recently, a Genome-Wide Association analysis comprising 78,308 subjects linked rs9320913 with fluid intelligence [10]. A previous work also associated rs9320913-A allele with a higher educational level and academic success [11]. The initial characterization of rs9320913 revealed that it lies within its own linkage disequilibrium block in a 1.6 Mb gene desert, 859 kb and 700 kb away of the nearest protein-coding genes. Further analysis of the genetic region showed that the polymorphism was located within AL589740.1 (ENCODE reference ENSG00000271860), a long intergenic noncoding RNA (lincRNA) whose longer transcript spans along 570 kb and 16 exons with 51 mature splice variants ranging from 0.3 to 3 kb. Unfortunately, their physiological role, as for many lincRNAs, is still unknown. This fact brings special relevance to the study of rs9320913, since lincRNAs have been associated to the molecular etiology of several neurodegenerative disorders including AD ([12] and references therein). In this context, we launched the Genetic Research in Alzheimer Disease (GERALD) initiative, a retrospective and prospective longitudinal project aimed at recruiting AD patients from Málaga (Southern Spain). The present work shows the pipeline of the project and the first results, derived from studying the rs9320913 variant in the course of the AD.

2. Materials and Methods

2.1. Study Population. The Genetic Research in Alzheimer Disease initiative is led by the Instituto Andaluz de Neurociencia, Málaga (https://ianec.es/), from Spain. Figure 1 summarizes the recruitment pipeline. Additional recruiting centers include the Asociación de Familiares de Alzheimer, Málaga; the Hospital Hermanas Hospitalarias del Sagrado Corazón, Málaga; the Centro Residencial Almudena, Málaga; the Centro Residencial Élite, Córdoba, Spain) following the XVIII World Medical Assembly Declaration of Helsinki and subsequent revisions.

2.2. DNA Extraction and Genotyping. DNA was extracted from each donor buccal swap using conventional digestion followed by ethanol precipitation. Blood DNA was extracted with a Blood Genomic DNA Extraction Kit (Canvax SL, Córdoba, Spain) following the manufacturer’s instructions. SNP rs9320913 was genotyped by PCR and subsequent enzymatic digestion using the following primers: forward 5′-CTCACATTTTCTGCATACCG-3′ and reverse 5′-AAA CCTGGATAAAAGCAATCAA-3′. Products were amplified in an Applied Biosystems 2700 thermocycler, using 30 cycles of 94°C for 30s, 60°C for 30s, and 72°C for 30s. Amplicons sizing 242 bp were digested with Apo I (Thermo Fisher Scientific, Massachusetts, USA) at 50°C for 1.5 hour and further run in a 3% agarose gel. The presence of the rs9320913-A allele generated two fragments of 172 and 70 bp, while the rs9320913-C allele retained the original length (Supplementary Figure 1). For the determination of the APOE genotype, the region comprising rs429358 and rs7412 was amplified by PCR with the oligonucleotides 5′-AGGCTGTCCAAAGGACGTGC-3′ and 5′-CGGCTGCC ATCTCCTCCATCC-3′. PCRs were performed at a 20 μl volume using 10 μg of each primer and 1x TB premix Ex Takara Bio Inc, Japan) with 5% dimethyl sulfoxide (DMSO), in an Applied Biosystems 2700 thermocycler with the following conditions: 95°C for 5 min followed by 35 cycles of

![Diagram showing GERALD project pipeline. De novo diagnoses undergo a full neuropsychological evaluation together with CNS magnetic resonance imaging (MRI). Peripheral blood extraction is used for both biochemical (BQ) analysis and DNA extraction for ApoE genotyping. Relevant data from the medical records, together with the different tests results, are integrated in a database. Subsequent follow-up results are also digitalized. Previously diagnosed patients undergo the neuropsychological evaluation which is longitudinally revised.](image-url)
94°C for 30 sec, 65°C for 30 sec, and 72°C for 30 sec. Next, PCRs were split in two independent LightCycler® Capillaries, each of them with 9 μl of the PCR product and 2 pg of the corresponding probe mix (rs429358 Anchor: 5′-CTGCAG GCGGCAGGGCCCCGCTGGGCCG[fluorescein]-3′ and sensor: 5′-[Cy5]ACATGGAGGACGTGCGCGG[phosphate]-3′; rs7412 Anchor: 5′-GCTGCGTAAGCGGCTCCTCCG CGATGCCG[fluorescein]-3′ and sensor: 5′-[Cy5]GACCTG CAGAGGCGCTGGG[phosphate]-3′). The fluorescence measurement was monitored using a LightCycler® (Roche Applied Science, Germany). For rs429358 (codon 112) determination, allele C exhibited a melting temperature (Tm) of 63°C, in contrast to the Tm = 55°C displayed by allele T. For rs7412 (codon 176), the C allele exhibited a Tm of 65°C while allele T had a Tm = 58°C (Supplementary Figure 2).

### 2.3. Statistical Analysis

The Hardy–Weinberg equilibrium was assayed using the corresponding online web tool from the Institute of Human Genetics of Munich (https://ihg.gsfc.de/ihg/snpcheck.html). Statistical analysis was performed using IBM SPSS Statistics v22. Power studies were calculated with Rothman’s Episheet [13]. The Kolmogorov–Smirnov test was used to determine the normality of the quantitative data series. For bivariate correlation studies, Pearson’s correlation coefficient and Spearman’s Rho were calculated. In order to evaluate cognitive decline, the slope between two Mini Mental State Examination (MMSE) scores obtained at AD diagnosis and at the first follow-up was obtained. Only data from the Q2 to Q4 quartiles of the MMSE scores at the time of diagnosis were taken into account for a better assessment of cognitive decline. The level of significance was defined at 0.05.

### 3. Results

#### 3.1. GERALD Series

Subjects comprised 121 women and 39 men, whose age at the time of diagnosis was from 39 to 91 years (mean age 77.21 ± 7.97 years), of which 76.7% (82.0% women and 18.0% men) presented a low educational level. The mean duration of symptoms determined through interviews with caregivers was 45.22 ± 31.79 months. MMSE scores determined at recruitment were 19.5 ± 5.9 points (average ± standard deviation). Global deterioration scores at recruitment ranged between 0.5 and 3, with an average of 1.98. Cognitive decline measured with MMSE varied between 0.78 and −2.22 points/month. When ApoE genotypes were determined, a single E4 allele was found in 55 (36.2%) patients, while homozygosity was found in 14 (9.2%). Deep white matter intensities ranged between 0 and 24, with an average of 5.1. Hippocampal atrophy ranged between 0 and 8, with an average of 3.37.

#### 3.2. Study of rs9320913 in AD Etiology

First reports mapped rs9320913 on a genetic desert on chromosome 6 flanked by POU3F2 and MMS22L, with no evident candidate gene potentially associated to the intelligence phenotype (Figure 2(a)). Our search was realized using the updated hg38 version of the genome browser, which revealed that rs9320913 now maps a long intervening noncoding RNA referenced AL589740.1 (Figure 2(b)). Moreover, when the AL589740.1 expression profile was explored in the GTEx portal, we found that its expression patterns were mainly constrained to the central nervous system (CNS) (Figure 3). Next, we examined whether rs9320913 genotype may be associated to the expression levels of AL589740.1 (expression quantitative trait locus or eQTL). Results showed that rs9320913 correlated with the brain-spinal cord...
the Munich University web tool.

check whether the e

value = 0.057) (Supplementary Table 1). In order to

portal (https://gtexportal.org). Signi

beta = 0

gene expression (indicate a correlation between rs9320913 and brain-amygdala

particular trend towards association, results for MMS22L

Although eQTL analysis for POU3F2 did not show any

expression pattern; in the telomeric side, POU3F2 showed a

centromeric border, MMS22L showed a virtually ubiquitous

genes might also exhibit a CNS-centered pattern. While in the

–

order, both parameters were negatively correlated

(beta = −0.18, p value = 0.028, Pearson’s test). Given its

higher dynamic range, only MMSE decay was included in

further analysis. Interestingly, MMSE decay showed a trend

towards less cognitive decline in homozygous rs9320913-A

individuals, although results were not statistically significant

probably due to the limited sample size. Figure 4 presents

study power was post hoc checked, we found that we were

80% powered to detect odds ratios ≥ 0.47. This suggests that

the effect size of rs9320913 in AD risk is, if any, below

OR = 0.47.

When the role of rs9320913 was examined according to

the MMSE at diagnostics, no significant differences were

observed between genotypes. The educational level and the

score obtained in the MMSE test at the time of diagnosis

showed a positive correlation (Spearman’s rho correlation

= 0.308, p value < 0.001) (Supplementary Figure 4). The

age of onset showed correlation with the MMSE test score

at the time of diagnosis (Spearman’s rho correlation of

-0.405, p value < 0.001) as well as an association with the

educational level (Spearman’s rho correlation of -0.189,

p value = 0.017). The inverse association between age of

onset and the MMSE score obtained indicates a trend

towards lower MMSE scores as patient age increases

(β = −0.345, p value < 0.001), while the effect of educational

level was proven to be positive, as higher educational levels

are related to higher MMSE scores at diagnosis (β = 0.255,

p value = 0.001). No significant associations were found

between any of the clinical variables measured and the

rs9320913 genotype. Thus, the previously reported e

first follow-up. Equiva-

lently, these calculations were performed for GDS. As

expected, both parameters were negatively correlated

(beta = −0.18, p value = 0.028, Pearson’s test). Given its

higher dynamic range, only MMSE decay was included in

further analysis. Interestingly, MMSE decay showed a trend

towards less cognitive decline in homozygous rs9320913-A

individuals, although results were not statistically significant

probably due to the limited sample size. Figure 4 presents

Figure 3: Violin plots representing AL589740.1 gene expression measured in transcripts per kilobase million (TPM) according to the GTEx portal (https://gtexportal.org). Significant expression is detected in different central nervous system areas, pituitary, and testis.

Table 1: rs9320913 genotypes in the case and control series.

<table>
<thead>
<tr>
<th></th>
<th>Controls (n = 254)</th>
<th>Cases (n = 153)</th>
<th>OR</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>88</td>
<td>61</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>112</td>
<td>69</td>
<td>0.89</td>
<td>0.60</td>
</tr>
<tr>
<td>AA</td>
<td>54</td>
<td>23</td>
<td>0.61</td>
<td>0.10</td>
</tr>
</tbody>
</table>

HWE p value 0.10 0.63

A: Dominant model (CC vs. GA); B: recessive model (C- vs. AA), according to the Munich University web tool.

(β = −0.23, p value = 0.037) and showed a trend towards association with the brain-frontal cortex (BA9) (β = −0.15, p value = 0.057) (Supplementary Table 1). In order to check whether the effect of rs9320913 might be extended to the closest protein-coding genes, we analyzed whether these genes might also exhibit a CNS-centered pattern. While in the centromeric border, MMS22L showed a virtually ubiquitous expression pattern; in the telomeric side, POU3F2 showed a CNS-associated expression pattern (Supplementary Figure 3). Although eQTL analysis for POU3F2 did not show any particular trend towards association, results for MMS22L indicate a correlation between rs9320913 and brain-amygdala gene expression (beta = 0.15, p value = 0.042) (Supplementary Table 1). These results suggest that rs9320913 acts as eQTL for AL589740.1 and could also act on MMS22L, but in different regions of the CNS.

When rs9320913 was genotyped in both case and control series, we observed that distributions fit the Hardy–Weinberg equilibrium. Allele frequencies were slightly different between cases and controls, since AA individuals were underrepresented in the group of cases: 15% versus the 21.2% found among the control group (Table 1). However, these results did not reach statistical difference. When the
Here, we present the creation of a novel longitudinal Alzheimer’s disease cohort devoted to contribute to revealing new genetic factors affecting AD cognitive decline. In this first work, we have focused on the characterization of the polymorphism rs9320913 given its potential role in cognitive reserve. Previous studies have reported the association between genetic variants and educational attainment and intelligence as well as an inverse relationship with the appearance of Alzheimer’s disease and depressive symptoms [10, 11, 14]. A protective effect between intelligence and disease development has been proposed [14]. As rs9320913 has been related to academic success [10, 11], our starting hypothesis was that the variant, being involved in processes that contribute to cognitive reserve, might influence the development of the disease. Thus, the effect of rs9320913 on the progression of AD is measured in cognitive decline using serial MMSE examinations. After genotyping the patient and control samples and performing the corresponding Hardy–Weinberg balance study, a statistical study was carried out in search of the relationships that could be established between the genotype of the variant and the cognitive decline due to AD.

Reviewing the genomic landscape of rs9320913, we found that it is located in a lncRNA, AL589740.1, of unknown function, but with a high expression in the central nervous system. This may suggest roles in neurodevelopmental and other processes specific to the central nervous system. In this regard, several lncRNAs have been linked to regulation of neuronal apoptosis, providing a new perspective on neurodegeneration [15]. lincRNA MALAT1, for instance, has been shown to inhibit neuron apoptosis and neuroinflammation [16], and lncRNA H19 has been found to initiate microglial pyroptosis [17].

The polymorphism rs9320913 turned out to be eQTL from AL589740.1, with significant association in the brain–spinal cord and brain–frontal cortex (BA9). A negative relation between the variant and the expression level was revealed. The implication of a lncRNA was of special interest as it has been reported that these molecules are expressed aberrantly during neurodegeneration, although their involvement in the process remains unclear. However, their possible effect on gene expression could establish a greater difference between patients and controls [18]. Nevertheless, a correlation between the rs9320913 genotype and the MMS22L expression has also been established, so we cannot discard a possible effect of the variant on an enhancer involved in the control of the MMS22L expression. However, the ubiquitous expression pattern of this gene diminishes the possible importance that its alteration can have on the development of intelligence.

Previous studies show the relationship between this allele and greater intelligence [10, 11]. Based on the theory of cognitive reserve, which argues that this ability optimizes cognitive functioning, it is expected that any effect on this capacity would affect the progression of the cognitive impairment associated with the disease. Despite this, no effect was found either on the parameter measuring decline or on the age of onset of cognitive impairment. Both the MMSE test score and age of onset were related to the educational level. Individuals with a higher educational level obtain higher scores in the MMSE test at the time of diagnosis, as was expected in the basis of the cognitive reserve theory [4]. However, results for age of onset were conflicting, as this variable showed a negative correlation with the educational level and this would imply that people with a higher cognitive reserve would present the symptoms earlier than those with a low cognitive reserve. An alternative explanation for this result is the greater sensitivity of highly educated people to their cognitive care, so that these people would detect
symptoms earlier, while those with lower educational levels would contact a professional with a greater deterioration. On the other side, specifically in the cohort of patients studied, the scale of the educational level used must be taken into account. The category of high educational level corresponds to patients with a bachelor’s degree and higher education levels, so there will be no differentiation among university levels. Other researches evaluate the effect of education based on years of schooling, which is a more precise scale to discern the effect of education on the development of the disease through cognitive reserve.

The statistical study carried out with rs9320913 did not reveal any significant association. However, the cognitive decline tended to be less pronounced in the presence of the A-allele, which would mean that this variant is associated to a less severe fall. Concerning the educational attainment, a trend towards a higher educational level was found related to the A-allele, which could be pointing out the effect of this variant on intelligence development. Both variables, MMSE decay slope and educational level, were found to be inversely related, and from this, we could extract either an effect of the polymorphism on the educational level, which in turn affects the lower cognitive decline, or a direct effect of the variant on both parameters. Nevertheless, the lack of statistical power of the study makes it impossible to validate these results and extract a clear conclusion. An extension of the sample size would be convenient to verify the results obtained, as well as the application of the parameter over a wider range.

Besides, we have focused on the MMSE test as a proxy for cognitive impairment, as it is the most widely used given its rapid clinical application and the proper information it offers in patients with typical Alzheimer’s disease [19].

Data Availability

Bioinformatic analysis is explained in the text and therefore can be independently reproduced. Raw personal data of Alzheimer patients is available upon the favorable pronunciation of the regional ethics committee.

Conflicts of Interest

The authors declare that there is no conflict of interest.

Authors’ Contributions

JLR and JMG-A designed the project; SM, EG, MPDG, JMM-C, AL, LC, OO, JT, MDN, CU, NP, ML, MGP, and RC performed psychological examinations; LL-G, EA-M, and JMC-G performed DNA analysis; LL-G and JLR performed statistical analysis; and LL-G, JLR, and JMG-A wrote the manuscript.

Acknowledgments

We wish to thank Irene Gonzalez for her valuable support in the laboratory routine and to all the patients, relatives, caregivers, and healthy controls who voluntarily enrolled in the study.

Supplementary Materials

Supplementary Figure 1: rs9320913 restriction fragment length polymorphism using Apo I. Supplementary Figure 2: ApoE genotyping plots obtained from a double heterozygote. Panel A shows the results for rs429358 while panel B refers to rs7412. Supplementary Figure 3: violin plots representing MMS22L (A) and POU3F2 (B) gene expression measured in transcripts per kilobase million (TPM) according to GTEx portal (https://gtexportal.org). MMS22L shows ubiquitous expression while POU3F2 present significant expression is detected in different central nervous system areas and the nerve. Supplementary Figure 4: boxplot of distribution of MMSE scores at the time of diagnosis according to educational level. Supplementary Table 1: eQTL between rs9320913 and AL589740.1, POU3F2, and MMS22L in the different tissues of the central nervous system. (Supplementary Materials)

References

[2] Committee on Decreasing the Risk of Alzheimer’s-Type Dementia Mild Cognitive Impairment and Age-Related Cognitive Impairment, Board on Health Science Policy, Institute of Medicine, and National Academies of Sciences Engineering and Medicine, Considerations for the design of a systemic review of interventions for preventing clinical Alzheimer’s-type dementia, mild cognitive impairment, and age-related cognitive decline, National Academies Press, Washington, DC, 2015.
[10] S. Sniekers, S. Stringer, K. Watanabe et al., “Genome-wide association meta-analysis of 78,308 individuals identifies new...


