alzheimer's
research&therapy

RESEARCH

Open Access

CrossMark

Role of ABCA7 loss-of-function variant in Alzheimer's disease: a replication study in European–Americans

Jorge L. Del-Aguila[1,2†], Maria Victoria Fernández[1,2†], Jessica Jimenez[1,2], Kathleen Black[1,2], Shengmei Ma[1,2], Yuetiva Deming[1,2], David Carrell[1], Ben Saef[1,2], Alzheimer's Disease Neuroimaging Initiative, Bill Howells[1,2], John Budde[1,2] and Carlos Cruchaga[1,2*]

Abstract

Introduction: A recent study found a significant increase of *ABCA7* loss-of-function variants in Alzheimer's disease (AD) cases compared to controls. Some variants were located on noncoding regions, but it was demonstrated that they affect splicing. Here, we try to replicate the association between AD risk and ABCA7 loss-of-function variants at both the single-variant and gene level in a large and well-characterized European American dataset.

Methods: We genotyped the *GWAS* common variant and four rare variants previously reported for *ABCA7* in 3476 European–Americans.

Results: We were not able to replicate the association at the single-variant level, likely due to a lower effect size on the European American population which led to limited statistical power. However, we did replicate the association at the gene level; we found a significant enrichment of *ABCA7* loss-of-function variants in AD cases compared to controls ($P = 0.0388$; odds ratio =1.54). We also confirmed that the association of the loss-of-function variants is independent of the previously reported genome-wide association study signal.

Conclusions: Although the effect size for the association of *ABCA7* loss-of-function variants with AD risk is lower in our study (odds ratio = 1.54) compared to the original report (odds ratio = 2.2), the replication of the findings of the original report provides a stronger foundation for future functional applications. The data indicate that different independent signals that modify risk for complex traits may exist on the same locus. Additionally, our results suggest that replication of rare-variant studies should be performed at the gene level rather than focusing on a single variant.

Introduction

A recent study found that loss-of-function variants in *ABCA7* (ATP-binding cassette transporter A7) confer greater risk for Alzheimer's disease (AD) [1]. Steinberg et al. [1] analyzed sequence, genome-wide association study (GWAS), and linkage data from 3419 individuals with AD and 151,805 controls from Iceland. Gene-based analyses, including nonsense, missense, frameshift splice-site variants and canonical splice-site variants

('loss-of-function'), identified *ABCA7* as the most significant gene (odds ratio (OR) = 2.12, $P = 2.2 \times 10^{-13}$) for AD. This association was mainly driven by a single splice-site variant, rs200538373 (OR = 4.47, $P = 3.4 \times 10^{-7}$), although other coding variants and splice variants were also found. This association was replicated at the gene level by genotyping the loss-of-function variants in more than 6500 AD cases and controls from four independent datasets (OR = 1.73, $P = 0.0056$). Interestingly, the OR for the variant (rs200538373) that led the association on the discovery series was in the opposite direction in the replication dataset (OR = 0.93). Additionally, none of these loss-of-function variants were in linkage disequilibrium (LD) with the *ABCA7* common variant

* Correspondence: ccruchaga@wustl.edu
†Equal contributors
[1]Department of Psychiatry, Washington University School of Medicine, 660 South Euclid Avenue B8134, St. Louis, MO 63110, USA
[2]The Hope Center for Neurological Disorders, BJC Institute of Health, 425 S. Euclid Ave, St. Louis, MO 63110, USA

Del-Aguila *et al. Alzheimer's Research & Therapy* (2015) 7:73

Page 2 of 4

identified by GWAS [2], suggesting that there are multiple and independent mechanisms throughout the *ABCA7* region that increase risk for AD. Based on these results, we tried to replicate the association of the *ABCA7* loss-of-function variants in a large cohort of European–Americans.

Methods

A total of 1776 AD cases and 1700 controls were selected from the Alzheimer's Disease Neuroimaging Initiative (ADNI), the Charles F. and Joanne Knight Alzheimer's Disease Research Center (Knight-ADRC) and the National Institute on Aging Genetics Initiative for Late-Onset Alzheimer's Disease (NIALOAD) [3, 4]. The Institutional Review Board at the Washington University School of Medicine in Saint Louis approved the study. Research was carried out in accordance with the approved protocol. Written informed consent was obtained from participants and their family members by the Clinical Core of the Knight-ADRC. The approval number for the Knight-ADRC Genetics Core family studies is 93-0006.

In the original study, Steinberg et al. [1] genotyped the *ABCA7* variants on 6681 European non-Icelandic individuals and only four variants were polymorphic (Supplementary Table 5 from Steinberg et al. [1]). Therefore, we decided to genotype the variants that were found to be polymorphic on the non-Icelandic population together with the GWAS common variant. The *ABCA7* GWAS common variant (rs4147929) as well as the loss-of-function variants (p.Glu709Alafs*86, p.Leu1403Argfs*7, rs200538373, and rs113809142) reported by Steinberg et al. [1] were genotyped in our dataset using competitive allele-specific polymerase chain reaction KASPar genotyping technologies as described previously [3, 4].

Allelic frequencies, Hardy-Weinberg equilibrium and single-variant association tests were performed with PLINK-1.9 [5]. For gene level analysis, burden analyses were performed using a combined multivariate and collapsing (CMC) test [6]. Age, gender, and principal component factors (PCs) were included in all association tests as covariates.

Results and discussion

In our dataset, the common variant rs4147929[A] was significantly associated with AD risk (Table 1; OR = 1.162, $P = 0.022$), suggesting that we have enough power to replicate known associations for common variants. All the genotyped loss-of-function variants were polymorphic in our dataset with similar frequencies to those reported in the European non-Icelandic datasets. As initially reported, none of the genotyped variants were in LD with the common variant rs4147929, allowing for independent analysis. In our dataset, all rare variants were more frequent in AD

Table 1 Association of *ABCA7* variants with Alzheimer's disease in studied groups based on Fisher's association test, including age, gender, and PC as covariates. Gene-based analysis was conducted with CMC collapsing method

rsID/position[a]	MAF		OR	P
	Cases	Controls	(95 % CI)	
rs4147929	0.179	0.1581	**1.162**	**0.022**
(GWAS SNP)			**(1.02–1.31)**	
19:998507[a]	0.006	0.0037	1.735	0.121
p.Glu709AlafsX86			(0.54–3.04)	
19:1006907[a]	0.003	0.0027	1.285	0.569
p.Leu1403ArgfsX7			(0.76–3.21)	
rs113809142	0.002	0.0012	1.695	0.451
c.4416+2T>G			(0.34–5.91)	
rs200538373	0.009	0.0076	1.231	0.476
c.5570+5G>C			(0.68–2.61)	
Loss-of-function	0.016	0.0107	**1.549**	**0.038**
(All low frequency variants)			**(1.02–2.34)**	

[a]NCBI Build 36
MAF
Values shown in bold are significant at the $P < 0.05$ level
CI Confidence interval, *CMC* Combined multivariate and collapsing, *GWAS* Genome-wide association study, *MAF* Minor allele frequency, *OR* Odds ratio, *PC* Principal component factors, *SNP* Single Nucleotide Polymorphism

cases than in controls with an OR ranging from 1.2 to 1.7; however, none of these values were statistically significant on their own (Table 1; $P > 0.1$).

Although we had statistical power to replicate the association of rs113809142 with AD risk based on the effect size reported in the Icelandic population (OR = 4.42, power = 0.887), we did not have enough power to replicate the single-variant analyses based on the minor allele frequency (MAF) and effect sizes reported on the European non-Icelandic dataset. There is some debate concerning what is the best approach to replicate the association of rare variants with complex traits [7–11]. It is clear that the MAF for these low-frequency variants varies widely, not only among populations, but also within populations [10–12]. If the cases and controls are not very well matched for local genetic background, the study can produce false-positive or false-negative results [7, 11, 13]. Additionally, it is possible that a specific variant is not found, or found in an extremely low frequency in a specific population; as with the case of the *TREM2* R47H variant in Asian [14, 15] or African–American populations. Although the association of the R47H variant with AD risk has been widely replicated in European–Americans, no significant association is found in Asian or African–American populations, because the MAF for this variant (R47H) is extremely low. However, other variants in the same gene could increase risk for diseases in these populations. Steinberg et al. [1] failed to replicate the

Del-Aguila *et al. Alzheimer's Research & Therapy* (2015) 7:73

Page 3 of 4

association of the rs113809142 in the European non-Icelandic population, but they were able to replicate the association at the gene level. These results support the notion that different (local) populations have varying genetic make-ups, and therefore single-variant analyses may not be the best approach for replicating these studies. This hypothesis is also supported by recent studies from Jin et al. [3, 15] in which deep resequencing of *TREM2* was performed on European–Americans and African–Americans; different variants were found in each population, and the variants in common presented very different MAFs and ORs. However, in both cases, the gene-based analyses supported the association of *TREM2* with AD risk.

For this reason, we decided to perform a gene-based analysis for all the reported polymorphic *ABCA7* loss-of-function variants. Since all variants presented the same direction of effect, we performed a CMC test. In our dataset, we found a significant enrichment of *ABCA7* loss-of-function variants in AD cases compared to controls ($P = 0.0388$; $OR = 1.54$). Therefore, despite none of the individual loss-of-function variants reported a significant association with *ABCA7* in this study, we were able to replicate two independent signals of the correlation of the *ABCA7* gene with AD: the common variant and the aggregation of the loss-of-function variants.

The point estimate for the OR in this study for the gene-based analysis is slightly lower than the reported OR for the Icelandic population ($OR = 2.12$) or the European non-Icelandic replication datasets ($OR = 1.73$), although the 95 % confidence interval in this study (1.02–2.34) includes both the Icelandic OR and the European non-Icelandic OR published by Steinberg et al. [1]. The current findings and those of the European non-Icelandic population support a possible "winner's curse" for the Icelandic discovery. Our sample size was smaller than both of the discovery series. Additionally, a proper gene-based replication would entail resequencing the candidate region to identify novel functional variants, and not just genotyping the reported variants; therefore, the real OR for the *ABCA7* loss-of-function variants remains to be determined. Despite these limitations, we were able to replicate the original report. Our data also indicate that the gene-based association of these loss-of-function variants is independent of the GWAS variant, and that the aggregate effect of these variants is larger than that of the common variant alone.

Our study validates the role of noncoding loss-of-function *ABCA7* variants in AD risk. Other population-specific independent variants with similar loss-of-function effects may contribute to AD risk or other complex traits. Supporting this hypothesis, a recent study has reported on an additional intronic low-frequency variant of *ABCA7* (rs78117248; $OR = 2.07$, $P = 0.0016$) that increases risk for AD, also independently of the common variant. Together, these results suggest that different and independent variants modify risk for complex diseases by different mechanisms existing on the same locus [16, 17]. Other genes will also harbor rare variants increasing risk for AD, independently of the GWAS hits [17].

Conclusions

In summary, our study replicates the association of *ABCA7* loss-of-function variants with AD risk, and highlights the necessity of performing gene-based, rather than single-variant analyses to replicate the association in this type of studies. Our study also confirms that there is high variability in the MAF of low-frequency variants within a population, so matching cases and controls for genetic background is a key step to avoiding false negatives or positives.

Abbreviations
ABCA7: ATC-binding cassette, subfamily A (ABC1), member 7; AD: Alzheimer's disease; ADNI: Alzheimer's Disease Neuroimaging Initiative; ADRC: Alzheimer's Disease Research Center; CMC: Combined multivariate and collapsing; GWAS: Genome-wide association study; LD: Linkage disequilibrium; MAF: Minor allele frequency; NIALOAD: National Institute on Aging Genetics Initiative for Late-Onset Alzheimer's Disease; OR: Odds ratio; PC: Principal component factors.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
Study design was developed by CC, MVF and JDA. DC, KB, JJ, SM, YD, JB, BH, BS, CC, MVF and JDA were involved in designing the current secondary data analysis study. Sample collection and DNA extraction were conducted by DC. Genotyping was conducted by KB and JJ. Phenotype data was collected by SM and principal component factors for population stratification were calculated by BH, KB, BS and BH. MVF, JDA and CC carried out the statistical analysis. All authors had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors participated in drafting, and read, revised and approved the final manuscript.

Del-Aguila *et al. Alzheimer's Research & Therapy* (2015) 7:73

Page 4 of 4

References

1. Steinberg S, Stefansson H, Jonsson T, Johannsdottir H, Ingason A, Helgason H, et al. Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. Nat Genet. 2015;47:445–7. doi:10.1038/ng.3246.
2. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45(12):1452–8. doi:10.1038/ng.2802.
3. Jin SC, Benitez BA, Karch CM, Cooper B, Skorupa T, Carrell D, et al. Coding variants in TREM2 increase risk for Alzheimer's disease. Hum Mol Genet. 2014;23(21):5838–46. doi:10.1093/hmg/ddu277.
4. Benitez BA, Jin SC, Guerreiro R, Graham R, Lord J, Harold D, et al. Missense variant in TREML2 protects against Alzheimer's disease. Neurobiol Aging. 2014;35:1510. doi:10.1016/j.neurobiolaging.2013.12.010. e19-26.
5. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7. doi:10.1186/s13742-015-0047-8.
6. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am J Hum Genet. 2008;83:311–21. doi:10.1016/j.ajhg.2008.06.024.
7. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. Am J Hum Genet. 2014;95:5–23. doi:10.1016/j.ajhg.2014.06.009.
8. Lambert JC, Grenier-Boley B, Bellengez C, Pasquier F, Campion D, Dartigues JF, et al. PLD3 and the risk of sporadic Alzheimer' disease. Nature. 2015;520(7545): E1. doi:10.1038/nature14036; In Press.
9. Heilmann S, Drichel D, Clarimon J, Fernandez V, Lacour A, Ramirez A. PLD3 in non-familial Alzheimer's disease. Nature. 2015;520(7545):E3–5 doi:10.1038/ nature14039.
10. Cruchaga C, Goate A. Reply to PLD3-variants in population studies. Nature. 2015;520(7545):E5–6. doi:10.1038/nature14037; In press.
11. Van der Lee SJ, Wong TY, Jakobsdottir M, Bis JC, Chouraki V, Van Duijn C. PLD3-variants in population studies. Nature. 2015;520(7545):E2–3. doi:10. 1038/nature14038 In Press.
12. Benitez BA, Cooper B, Pastor P, Jin SC, Lorenzo E, Cervantes S, et al. TREM2 is associated with the risk of Alzheimer's disease in Spanish population. Neurobiol Aging. 2013;34:1711. doi:10.1016/j.neurobiolaging.2012.12.018. e15-7.
13. Mathieson I, McVean G. Differential confounding of rare and common variants in spatially structured populations. Nat Genet. 2012;44:243–6. doi:10.1038/ng.1074.
14. Miyashita A, Wen Y, Kitamura N, Matsubara E, Kawarabayashi T, Shoji M, et al. Lack of genetic association between TREM2 and late-onset Alzheimer's disease in a Japanese population. J Alzheimers Dis. 2014;41:1031–8. doi:10.3233/JAD-140225.
15. Huang M, Wang D, Xu Z, Xu Y, Xu X, Ma Y, et al. Lack of genetic association between TREM2 and Alzheimer's disease in East Asian population: a systematic review and meta-analysis. Am J Alzheimers Dis Other Demen. 2015; 30(6):541-6. doi:10.1177/1533317515577128.
16. Cruchaga C, Haller G, Chakraverty S, Mayo K, Vallania FL, Mitra RD, et al. Rare variants in APP, PSEN1 and PSEN2 increase risk for AD in late-onset Alzheimer's disease families. PLoS One. 2012;7, e31039. doi:10.1371/journal.pone.0031039.
17. Vardarajan BN, Ghani M, Kahn A, Sheikh S, Sato C, Barral S, et al. Rare coding mutations identified by sequencing of Alzheimer disease genome-wide association studies loci. Ann Neurol. 2015;78:487–98. doi:10.1002/ana.24466.