Loss-of-function variants in *ABCA7* confer risk of Alzheimer's disease

Stacy Steinberg[1,26], Hreinn Stefansson[1,26], Thorlakur Jonsson[1,26], Hrefna Johannsdottir[1], Andres Ingason[1], Hannes Helgason[1], Patrick Sulem[1], Olafur Th Magnusson[1], Sigurjon A Gudjonsson[1], Unnur Unnsteinsdottir[1], Augustine Kong[1], Seppo Helisalmi[2,3], Hilkka Soininen[2,3], James J Lah[4], DemGene[5], Dag Aarsland[6–8], Tormod Fladby[8,9], Ingun D Ulstein[10,11], Srdjan Djurovic[12,13], Sigrid B Sando[14,15], Linda R White[14,15], Gun-Peggy Knudsen[16], Lars T Westlye[17,18], Geir Selbæk[19], Ina Giegling[20], Harald Hampel[21], Mikko Hiltunen[2,3,22], Allan I Levey[4], Ole A Andreassen[17,23], Dan Rujescu[20], Palmi V Jonsson[24,25], Sigurbjorn Bjornsson[25], Jon Snaedal[25] & Kari Stefansson[1,24]

We conducted a search for rare, functional variants altering susceptibility to Alzheimer's disease that exploited knowledge of common variants associated with the same disease. We found that loss-of-function variants in *ABCA7* confer risk of Alzheimer's disease in Icelanders (odds ratio (OR) = 2.12, $P = 2.2 \times 10^{-13}$) and discovered that the association replicated in study groups from Europe and the United States (combined OR = 2.03, $P = 6.8 \times 10^{-15}$).

Alzheimer's disease is a common neurodegenerative disorder with high heritability. The ε4 allele of *APOE* was found to associate with Alzheimer's disease about 20 years ago[1,2], and rare variants in three genes (*APP*, *PSEN1* and *PSEN2*) were implicated in familial, early-onset forms of the disease around the same time[3]. More recently, rare variants in *APP* and *TREM2* were discovered to affect susceptibility to the sporadic, late-onset form of Alzheimer's disease[4–6], and the latest meta-analysis of common variants demonstrated that 19 loci, in addition to the *APOE* locus, harbor variants showing genome-wide significant association with Alzheimer's disease[7]. The new common variant association regions include the major histocompatibility complex (MHC) locus and 18 other regions in which the association signal extends over 104 genes (summarized in **Supplementary Table 1**)[7].

We used our imputation of the whole-genome sequences of 2,636 Icelanders into 104,220 long-range phased individuals and their close relatives[8] to investigate whether any of the genes located in the regions showing common variant association with Alzheimer's disease (excluding the MHC) also harbored rare variants conferring higher risk. We considered two classes of variants: nonsense, frameshift or canonical splice-site variants ('loss of function'), and missense or splice-site variants ('missense'). Because loss-of-function mutations are likely to have the same biological effect—no protein is produced as a result of nonsense-mediated decay of mRNA containing premature stop codons—we collapsed these variants into a single allele and estimated one effect. Missense variants may have different biological consequences; thus, for these variants, we ran analyses using SKAT[9], a gene test that allows different effects, in addition to a collapsing test.

We carried out association testing using information from 3,419 individuals with Alzheimer's disease and 151,805 population controls. For loss-of-function variants, collapsing tests using variants with minor allele frequency (MAF) <0.5% and information >0.8 resulted in examination of 20 genes in 10 common variant association regions. The most significant association result was for *ABCA7* (OR = 2.08, $P = 3.8 \times 10^{-5}$) (**Supplementary Table 2**). For missense variants, we carried out collapsing tests for variants with MAF <0.5% and information >0.8 (79 genes in 17 regions), and SKAT scans using both all variants with MAF <1% and information >0.8 (82 genes in 17 regions) and the subset of those variants predicted to be damaging or possibly damaging by PolyPhen (63 genes in 17 regions). The most significant result for the missense variant tests was for *ABCA7*, obtained using SKAT with all variants ($P = 0.00020$; **Supplementary Tables 3** and **4**).

As the most significant result for both the loss-of-function and missense variant scans was for *ABCA7*, we searched for functional commonality among the variants producing significance in the two scans. We found that the *ABCA7* missense variant test was significant due to a single splice-site variant, c.5570+5G>C ($P = 0.46$ without this variant). Examination of transcripts from several c.5570+5G>C carriers using RNA sequencing (RNA-seq) showed that this variant led to mRNA containing intronic sequences (**Supplementary Fig. 1**) that eventually included a stop codon. Thus, c.5570+5G>C

[1]deCODE Genetics, Reykjavik, Iceland. [2]Institute of Clinical Medicine–Neurology, University of Eastern Finland, Kuopio, Finland. [3]NeuroCenter, Kuopio University Hospital, Kuopio, Finland. [4]School of Medicine, Emory University, Atlanta, Georgia, USA. [5]A list of members and affiliations appears at the end of the paper. [6]Alzheimer's Disease Research Centre, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden. [7]Center for Age-Related Diseases, Stavanger University Hospital, Stavanger, Norway. [8]Institute of Clinical Medicine, Division of Medicine and Laboratory Sciences, University of Oslo, Oslo, Norway. [9]Department of Neurology, Akershus University Hospital, Lørenskog, Norway. [10]Department of Psychiatry of Old Age, Oslo University Hospital, Oslo, Norway. [11]Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [12]Department of Medical Genetics, Oslo University Hospital, Oslo, Norway. [13]NORMENT–K.G. Jebsen Centre, Department of Clinical Science, University of Bergen, Bergen, Norway. [14]Department of Neuroscience, Norwegian University of Science and Technology, Trondheim, Norway. [15]Department of Neurology, St. Olav's Hospital, Trondheim University Hospital, Trondheim, Norway. [16]Division of Mental Health, Norwegian Institute of Public Health, Oslo, Norway. [17]NORMENT–K.G. Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. [18]Department of Psychology, University of Oslo, Oslo, Norway. [19]Ageing and Health, Norwegian Centre for Research, Education and Service Development, Vestfold Hospital Trust, Tønsberg, Norway. [20]Department of Psychiatry, University of Halle, Halle, Germany. [21]Sorbonne Universités, Université Pierre et Marie Curie, Département de Neurologie, Hôpital Pitié Salpêtrière, Paris, France. [22]Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland. [23]NORMENT–K.G. Jebsen Centre, Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [24]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [25]Department of Geriatric Medicine, Landspitali University Hospital, Reykjavik, Iceland. [26]These authors contributed equally to this work. Correspondence should be addressed to K.S. (kstefans@decode.is).

Received 8 June 2014; accepted 13 February 2015; published online 25 March 2015; doi:10.1038/ng.3246

Table 1 Association of loss-of-function variants in *ABCA7* with Alzheimer's disease in Iceland

Variant	Position[a]	Rs ID	MAF (%)	OR	*P*	Info
p.Tyr622*	Chr19:998,176	–	0.041	4.42	0.0034	0.98
p.Glu709Alafs*86	Chr19:998,507	–	0.274	2.14	0.0014	0.98
p.Arg1118*	Chr19:100,4459	–	0.085	0.96	0.95	0.99
p.Leu1403Argfs*7	Chr19:1,006,907	–	0.085	1.78	0.25	0.99
c.4416+2T>G	Chr19:1,007,244	rs113809142	0.129	4.47	3.4×10^{-7}	0.99
p.Arg1561*	Chr19:1,008,000	–	0.045	1.30	0.69	0.99
c.5570+5G>C	Chr19:1,012,892	rs200538373	0.800	1.91	3.8×10^{-6}	0.99
c.6044+1G>A	Chr19:1,015,253	–	0.028	1.76	0.53	0.99
All	–	–	1.487	2.12	2.2×10^{-13}	0.99

Variant names are based on RefSeq records NM_019112.3 and NP_061985.2. MAF, minor allele frequency in controls; OR, odds ratio.

[a]NCBI Build 36.

could be included in the loss-of-function test, which then had an OR of 1.97 and a *P* value of 5.3×10^{-10}.

We followed up the *ABCA7* loss-of-function association result further by carrying out Sanger sequencing of predicted variant carriers to validate the variants and to create a larger reference set for reimputation. Five of the six variants included in the original loss-of-function test as well as c.5570+5G>C and two variants not included in the initial scan because of low information scores were validated. All pairs of variants were uncorrelated ($r^2 < 0.001$). Each of the variants was predicted to cause the insertion of a stop codon into the transcript before the last two exons. After reimputation, *ABCA7* loss-of-function variants were associated with Alzheimer's disease with an OR of 2.12 and a *P* value of 2.2×10^{-13} (**Table 1**).

Using our long-range phased haplotypes, we found that the *ABCA7* loss-of-function variants were never located on the background of rs4147929[A], the common variant previously associated with Alzheimer's disease[7]. Thus, the loss-of-function variant association described here could not account for the common variant association, nor vice versa. We did not find evidence of interaction between the rare *ABCA7* loss-of-function alleles and rs4147929[A] ($P = 0.35$).

To investigate the effect of these variants on Alzheimer's disease outside Iceland, we genotyped the *ABCA7* loss-of-function variants in study groups from Finland, Germany, Norway and the United States. We found six of the eight variants in at least one of the study groups (**Supplementary Table 5**). In the combined replication study groups, loss-of-function variants were associated with Alzheimer's disease with an OR of 1.73 and a *P* value of 0.0056 (**Table 2** and **Supplementary Table 5**). Meta-analysis with the Icelandic data yielded an OR of 2.03 and a *P* value of 6.8×10^{-15}.

ABCA7, or ATP-binding cassette transporter A7, is a member of the A subfamily of ABC transporters that move primarily lipids across membranes[10]. It is strongly expressed in brain, with the highest levels found in microglia[10]. ABCA7 promotes the efflux of phospholipids and cholesterol to apolipoprotein A-I (apoA-I) and apolipoprotein E (apoE) in cell culture[11,12]. *ABCA7* is also an ortholog of the *Caenorhabditis elegans* gene *ced-7*, the product of which has a crucial role in the engulfment of apoptotic cells. In line with this function, ABCA7 is involved in the promotion of phagocytosis in several human cell lines[13,14].

Recent results from mouse models suggest that the primary function of ABCA7 is in the regulation of phagocytosis rather than in cholesterol metabolism. *Abca7*-null mice have only modestly altered serum cholesterol levels, and cholesterol and phospholipid efflux in macrophages isolated from these mice does not differ from that of wild-type macrophages[15]. In addition, although J20 *Abca7*-null mice have greater insoluble amyloid β (Aβ) levels than J20 mice, they do not have elevated apoE concentrations, suggesting that the increase in Aβ levels is not a result of increased efflux to apoE[16].

To investigate whether genetic data provide any insights into the pathway through which ABCA7 acts, we tested for statistical interaction of the *ABCA7* loss-of-function variants with either *APOE* ε4 or the Alzheimer's disease–associated variant in *TREM2*, which encodes a protein involved in microglial activation[17,18]. We found no evidence of interaction of the *ABCA7* loss-of-function variants with the Alzheimer's disease–associated alleles of either *APOE* or *TREM2* ($P = 0.72$ and 0.38, respectively).

This study reinforces the claim that *ABCA7*, rather than neighboring genes, is involved in the pathogenesis of Alzheimer's disease. In addition, although it is currently unknown how the causal variant tagged by rs4147929[A] affects *ABCA7* expression in the brain—both increased and decreased mRNA levels have been reported[19,20]—our results for the loss-of-function variants suggest that rs4147929[A] is likely to lead to reduced levels of the protein. Moreover, because of the high OR associated with *ABCA7* loss-of-function variants, carriers of these variants should be useful participants in clinical studies seeking to improve the understanding and treatment of Alzheimer's disease.

Table 2 Association of loss-of-function variants in *ABCA7* with Alzheimer's disease in four non-Icelandic data sets

Study	*n* cases	*n* controls	MAF (%)	OR	*P*
United Sates (Emory)	409	406	0.49	2.50	0.18
Finland	515	588	0.60	2.13	0.12
Germany	522	1,917	0.94	1.33	0.38
Norway	919	1,405	0.50	1.86	0.097
All	2,365	4,316	–	1.73	0.0056

MAF, minor allele frequency in controls; OR, odds ratio.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was done in collaboration with Illumina. The authors are grateful to the participants and their relatives. We also thank the staff and researchers at the Krókháls recruitment center, the Memory Clinic at the National Hospital of Iceland and the deCODE Genetics core facilities. This work was carried out with support from the Academy of Finland; EVO grant 5772708 of Kuopio University Hospital; the Strategic Funding of the University on Eastern Finland (University of Eastern Finland–Brain); the Sigrid Juselius Foundation; Framework Programme 7, grant agreement 601055; VPH Dementia Research Enabled by IT VPH-DARE@IT; the Research Council of Norway (223273, 225989, 237250/EU JPND); the K.G. Jebsen Foundation; the Norwegian Health Association; the AXA Research Fund, the Fondation Université Pierre et Marie Curie and the Fondation pour la Recherche sur Alzheimer, Paris, France; and the program 'Investissements d'Avenir' ANR-10-IAIHU-06.

AUTHOR CONTRIBUTIONS

S.S., H. Stefansson, T.J. and K.S. designed the study. S.S., H.J., A.I., H. Helgason, P.S., S.A.G., A.K. and U.U. carried out data analysis. O.T.M. performed experiments. S.H., H. Soininen, M.H., DemGene, D.A., T.F., I.D.U., S.D., S.B.S., L.R.W., G.-P.K., L.T.W., G.S., O.A.A., J.J.L., A.I.L., I.G., H. Hampel, D.R., P.V.J., S.B. and J.S. diagnosed and recruited patients. S.S., H. Stefansson, T.J. and K.S. wrote the manuscript with the support of the remaining authors.

COMPETING FINANCIAL INTERESTS
The authors declare competing financial interests: details are available in the online version of the paper.

1. Saunders, A.M. *et al. Neurology* **43**, 1467–1472 (1993).
2. Strittmatter, W.J. *et al. Proc. Natl. Acad. Sci. USA* **90**, 1977–1981 (1993).
3. Cruts, M., Theuns, J. & Van Broeckhoven, C. *Hum. Mutat.* **33**, 1340–1344 (2012).
4. Guerreiro, R. *et al. N. Engl. J. Med.* **368**, 117–127 (2013).
5. Jonsson, T. *et al. Nature* **488**, 96–99 (2012).
6. Jonsson, T. *et al. N. Engl. J. Med.* **368**, 107–116 (2013).
7. Lambert, J.C. *et al. Nat. Genet.* **45**, 1452–1458 (2013).
8. Kong, A. *et al. Nat. Genet.* **40**, 1068–1075 (2008).
9. Wu, M.C. *et al. Am. J. Hum. Genet.* **89**, 82–93 (2011).
10. Kim, W.S., Weickert, C.S. & Garner, B. *J. Neurochem.* **104**, 1145–1166 (2008).
11. Abe-Dohmae, S. *et al. J. Biol. Chem.* **279**, 604–611 (2004).
12. Wang, N. *et al. J. Biol. Chem.* **278**, 42906–42912 (2003).
13. Iwamoto, N., Abe-Dohmae, S., Sato, R. & Yokoyama, S. *J. Lipid Res.* **47**, 1915–1927 (2006).
14. Jehle, A.W. *et al. J. Cell Biol.* **174**, 547–556 (2006).
15. Kim, W.S. *et al. J. Biol. Chem.* **280**, 3989–3995 (2005).
16. Kim, W.S. *et al. J. Neurosci.* **33**, 4387–4394 (2013).
17. Griciuc, A. *et al. Neuron* **78**, 631–643 (2013).
18. Takahashi, K., Rochford, C.D. & Neumann, H. *J. Exp. Med.* **201**, 647–657 (2005).
19. Allen, M. *et al. Neurology* **79**, 221–228 (2012).
20. Vasquez, J.B., Fardo, D.W. & Estus, S. *Neurosci. Lett.* **556**, 58–62 (2013).

Norwegian Dementia Genetics Consortium (DemGene):

Ina Selseth Almdahl[9], Fred Andersen[27], Nenad Bogdanovic[28], Anne Brækhus[28], Knut Engedal[19], Arvid Rongve[29,30], Ingvild Saltvedt[14], Eystein Stordal[14,31] & Aree Witoelar[11,17]

[27]Department of Community Medicine, University of Tromsø, Tromsø, Norway. [28]Geriatric Department, University Hospital Oslo and University of Oslo, Oslo, Norway. [29]Division of Mental Health, Helse Fonna, Haugesund, Norway. [30]Department of Clinical Medicine, University of Bergen, Bergen, Norway. [31]Department of Psychiatry, Namsos Hospital, Namsos, Norway.

ONLINE METHODS

Subjects. The subjects from Iceland, the United States (Emory University) and Germany have previously been described[6]. Briefly, patients from Iceland were diagnosed with definite, probable or possible Alzheimer's disease on the basis of the NINCDS-ADRDA criteria[21] (n = 1,035) or according to guidelines for ICD-10 F00 (n = 2,384) and were compared to population controls. The patients from Emory and Germany were diagnosed with probable Alzheimer's disease according to the NINCDS-ADRDA criteria[21] and were compared to either screened (Emory) or population (Germany) controls. The Norwegian sample included cases from 5 studies: the AHUS study (n = 75), the Dementia Study of Western Norway (DemVest; n = 82), the Dementia Study in Rural Northern Norway (n = 74), the Norwegian Memory Register study (n = 243) and the TrønderBrain study (n = 445). These cases were diagnosed according to the ICD-10 research criteria[22] (the Memory Register study and the Dementia Study in Rural Northern Norway), the recommendations from the National Institute on Aging–Alzheimer's Association (NIA/AA)[23] (AHUS) or the NINCDS-ADRDA criteria[21] (DemVest and TrønderBrain). The controls from Norway included screened controls obtained through the AHUS study (n = 89), the Dementia Study in Rural Northern Norway (n = 136) and the TrønderBrain study (n = 651) as well as population controls from the Norwegian Mother and Child Cohort Study (MoBa) conducted by the Norwegian Institute of Public Health[24] (n = 381) and a lifespan study (n = 148). See the **Supplementary Note** for further information on the Norwegian studies. The Finnish sample consisted of 515 cases diagnosed with probable Alzheimer's disease according to NINCDS-ADRDA clinical criteria[21] and 588 age-matched controls; this sample has been described previously[25]. The Icelandic study was approved by the Data Protection Authority and the National Bioethics Committee, and the non-Icelandic studies were approved by local ethics boards. Participants giving samples also gave written, informed consent.

Genotyping and sequencing. Genotyping on Illumina arrays, whole-genome sequencing and variant calling of the Icelandic samples was carried out as previously described[26]. Validation of the sequencing and imputation results was performed using Sanger sequencing as detailed earlier[26]. Sample preparation, sequencing and analysis for RNA-seq were carried out as described in the **Supplementary Note**. Genotyping of the Emory, Finland, Germany and Norway samples was performed using Centaurus assays (p.Tyr622*, p.Leu1403Argfs*7, c.4416+2T>G, c.5570+5G>C and c.6044+1G>A) or Sanger sequencing (p.Glu709Alafs*86, p.Arg1118* and p.Arg1561*).

Long-range phasing, imputation and association testing. Long-range phasing and imputation of the Icelandic sample were carried out as previously described[26]. The Alzheimer's disease common variant association regions were defined as regions in which SNPs correlated ($r^2 > 0.2$) with the most significant SNP at the locus showing association above baseline (Fig. 2 and Supplementary Figs. 2–11 in ref. 7). For the collapsing tests, appropriate variants in each gene were combined into a single allele, and analysis was conducted using logistic regression assuming a log-additive model for the two alleles and including sex, age, age^2 and county of birth as covariates. The SKAT[9] scan included the same covariates as the collapsing tests. Annotation was performed using VEP[27] on the basis of RefSeq transcripts. Variants were defined as 'probably damaging' or 'possibly damaging' using the PolyPhen-2 HDIV database[28]. For the collapsing and individual-variant tests, we included information both from typed individuals (2,476 cases and 65,347 controls) and untyped, related individuals (943 cases and 86,458 controls) as previously described[26]. For the SKAT scans, only information from typed individuals was used. Genomic control[29] was applied to all Icelandic tests to account for relatedness. Association testing in the replication study groups was carried out using Fisher's exact test, and the results from the different replication study groups were combined using the exact Mantel-Haenzel test.

21. McKhann, G. *et al. Neurology* **34**, 939–944 (1984).
22. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines* (World Health Organization, 1992).
23. McKhann, G.M. *et al. Alzheimers Dement.* **7**, 263–269 (2011).
24. Magnus, P. *et al. Int. J. Epidemiol.* **35**, 1146–1150 (2006).
25. Natunen, T. *et al. J. Alzheimers Dis.* **37**, 217–232 (2013).
26. Styrkarsdottir, U. *et al. Nature* **497**, 517–520 (2013).
27. McLaren, W. *et al. Bioinformatics* **26**, 2069–2070 (2010).
28. Adzhubei, I.A. *et al. Nat. Methods* **7**, 248–249 (2010).
29. Devlin, B. & Roeder, K. *Biometrics* **55**, 997–1004 (1999).

doi:10.1038/ng.3246