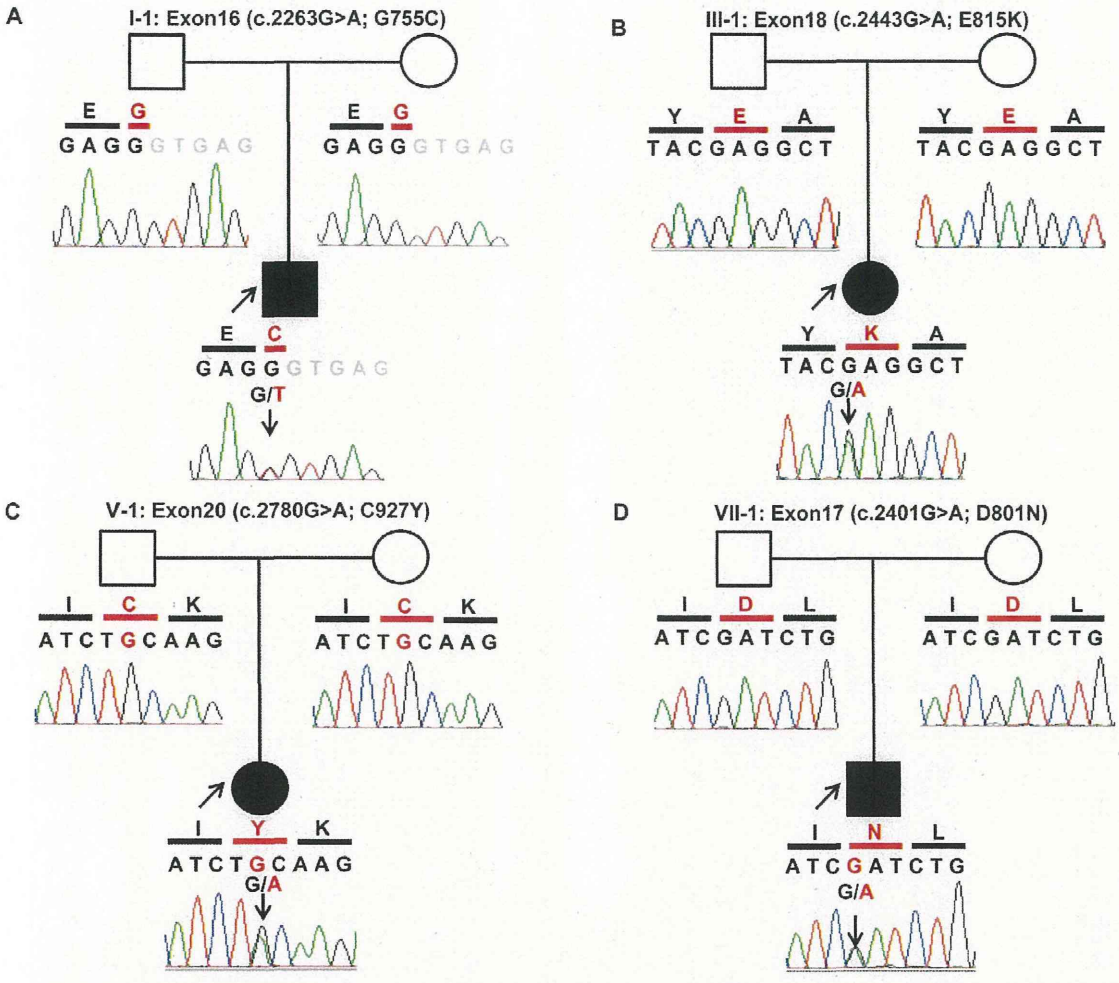

Figure 2. Chromatograms of four *de novo* mutations identified in *ATP1A3*. Data were obtained by Sanger sequencing during the confirmation process. In trio of each pedigree, black shadow represents the proband. In the chromatograms, *Black letters* show exonic nucleotide sequences, *gray letters* show intronic nucleotide sequences. Amino acids are shown in a single letter notation. Nucleotides and amino acids in red indicate mutations. (A) G755C was identified only in Patient I-1. (B) E815K was identified in Patients II-1, III-1, IV-1, IX-1 and X-1. (C) C927Y was identified in Patient V-1 only. (D) D801N was identified in Patients VI-1, VII-1 and VIII-1. None of the mutations was detected in the father or mother except for Patient IX-1, whose parents refused to undergo genetic analysis.
doi:10.1371/journal.pone.0056120.g002

function. In fact, *in silico* analysis of the mutations identified in the present study suggests a profound damage of the ATPase molecule and hence accord well with functional deficits of the ATPase encountered with the recently described mutations [10].

It is noteworthy that several mutations of *ATP1A3* have been reported to cause RDP [9]. RDP is an autosomal dominant disease characterized by abrupt onset of dystonia and Parkinsonism, developing within minutes to days of onset [13–16]. Recently reported were two infantile RDP patients with *ATP1A3* mutations (R756H and D923N); onset began for one of them at 11 months and for the other at 4 years of age. Major symptoms included motor delay, hypotonia, and ataxia [17,18]. Involuntary movements such as dystonia overlap with AHC, however, their clinical features and age of onset are different than those of AHC, which mainly shows repeated attacks of alternating hemiplegia and which begins with abnormal ocular movements by 3 months of age. Both typical and infantile RDP show different clinical features and processes than AHC, although *ATP1A3* seems to be pathologically involved in both disorders. In particular, D801N, one of the *ATP1A3* mutations identified in the present study, affected D801, where D801Y had been found in RDP. Thus, it seems that two substitutions in the same amino acid result in two distinguished phenotypes. Initially, we could not identify D801N in *ATP1A3* from novel variant. The reason for the erroneous results was the extraction of novel variants from all the variants using chromosome position only during the collation of databases. The position 42474557 of chromosome 19, where the G to A transversion resulted in D801N identified by our exome sequencing, had been registered as the nucleotide where the G to T transition is identified in rapid-onset dystonia-parkinsonism. Based on the backup plans involving reconfirmation of the gene identified with novel variants, using all variants, and to re-sequence the gene in our pipeline with the Sanger sequencer, D801N was not overlooked in the present study. These results suggest that confirmation by Sanger sequencer is useful in avoiding any oversight in the field of gene identification.

Table 3. Clinical data of 10 unrelated individuals with AHC.

Patient ID	I-1	II-1	III-1	IV-1	V-1	VI-1	VII-1	VIII-1	IX-1	X-1
Mutations	**G755C**	**E815K**	**E815K**	**E815K**	**C927Y**	**D801N**	**D801N**	**D801N**	**E815K**	**E815K**
Age (year)/sex	18/male	13/male	32/female	6/male	16/female	17/male	9/male	12/male	9/male	1/male
Age at onset (day)	60	17	2	1	60	1	120	0	Infant	Neonatal
Age at onset of paralysis (month)	6	10	12	4	12	4	9	9	Infant	9
Initial symptoms/signs	L versive movement of neck, monocular deviation of L eye to the left	Tonic fits	Tonic fits	Upward gaze, tonic fits	Nystagmus, ocular deviation to right	Nystagmus, focal clonic seizure	Clonic seizure	Nystagmus	Apnea	Nystagmus, downward gaze, tonic fits
Paralytic type	Flaccid	Flaccid	Flaccid	Flaccid	Rigid	Flaccid	Flaccid	Flaccid	Flaccid	Flaccid
Paralytic symptoms	Paralysis of unilateral arm or leg on R or L, or hemiparesis, sometimes continues with shift to opposite side. Rarely quadriplegia.	Paralysis of unilateral arm or leg on R or L, or hemiparesis, sometimes shifts to opposite side. Rarely quadriplegia.	Hemiparesis. Sometimes quadriplegia. No episodic paralysis since stabilizing of quadriplegia at 14 years.	Paralysis or hemiparesis of R arm.	Rigidity of R arm. Alternating flaccid hemiplegia since 1 year of age.	Alternating hemiparesis every 2–3 months	Alternating hemiplegia (R>L), only a few days every month.	R or L unilateral arm or leg paralysis, sometimes systemic paralysis. Tendency to occur following tonic fits.	Quadriplegia with/without bulbar palsy, for a few min to several hrs every day. Sometimes hemiplegia. Sometimes paralysis shifts to other parts.	Exterior ocular deviation on R side. Systemic cataplexy. Alternating paraparesis
Other neurological abnormalities	Choreoathetosis, aphonia	Choreoathetosis, facial dyskinesia	Dystonia, oral or facial dyskinesia	Aphonia	Spastic diplegia	None	Left hemidystonia	Dystonia	Dystonia	Head lag, nystagmus, ocular deviation
Motor development	walks alone	stands with support	walks with support	sits alone	walks alone	walks alone	walks alone	walks with support	Unable to sit	rolling over
Intellectual development	two words	only words	only words	no words	Normal	three phrases	three word phrases	only words	No words	delay
Regression	No	Yes	Yes	No	Yes	No	No	Yes	Yes	No
Epilepsy	4 years	2 years	4 years	None	None	None	4months	8 years	Yes	9months
Epileptic status	No	Yes	Yes	No	No	No	No	Yes	Yes	Yes
Headache	Yes	Yes	No	No	No	No	No	No	unknown	unknown
Head MRI	Normal	Cerebellar atrophy	Cerebellar atrophy	Normal	Mild enlargement of inferior horns bilaterally	Normal	Normal	High intensity in hippocampus	N/A	Normal
Respiratory status	Apnea	Normal	Use of ventilator	Apnea	Normal	Normal	Normal	Apnea	Apnea	Apnea
Effective drugs for paralysis	flunarizine	CZP	CZP, flunarizine	flunarizine	CZP	flunarizine	flunarizine	flunarizine	none (flunarizine not tried)	MDL
Family history	None	None	None	Headache, epilepsy	None	None	Migraine	Headache, epilepsy	Headache	None
Gestational age	40 weeks	34 weeks 3days	42 weeks	40 weeks	unknown	41 weeks 4 days	39 weeks 3 days	41 weeks	40 weeks	37 weeks 3 days
Birth weight (g)	3148	2218	3260	3392	unknown	3526	3200	3008	3550	2962
Asphyxia	None	No crying unless stimulated	Unknown	None	unknown	None	unknown	None	None	None

MDL: midazolam, CZP: clonazepam, L: left, R: right.

doi:10.1371/journal.pone.0056120.t003

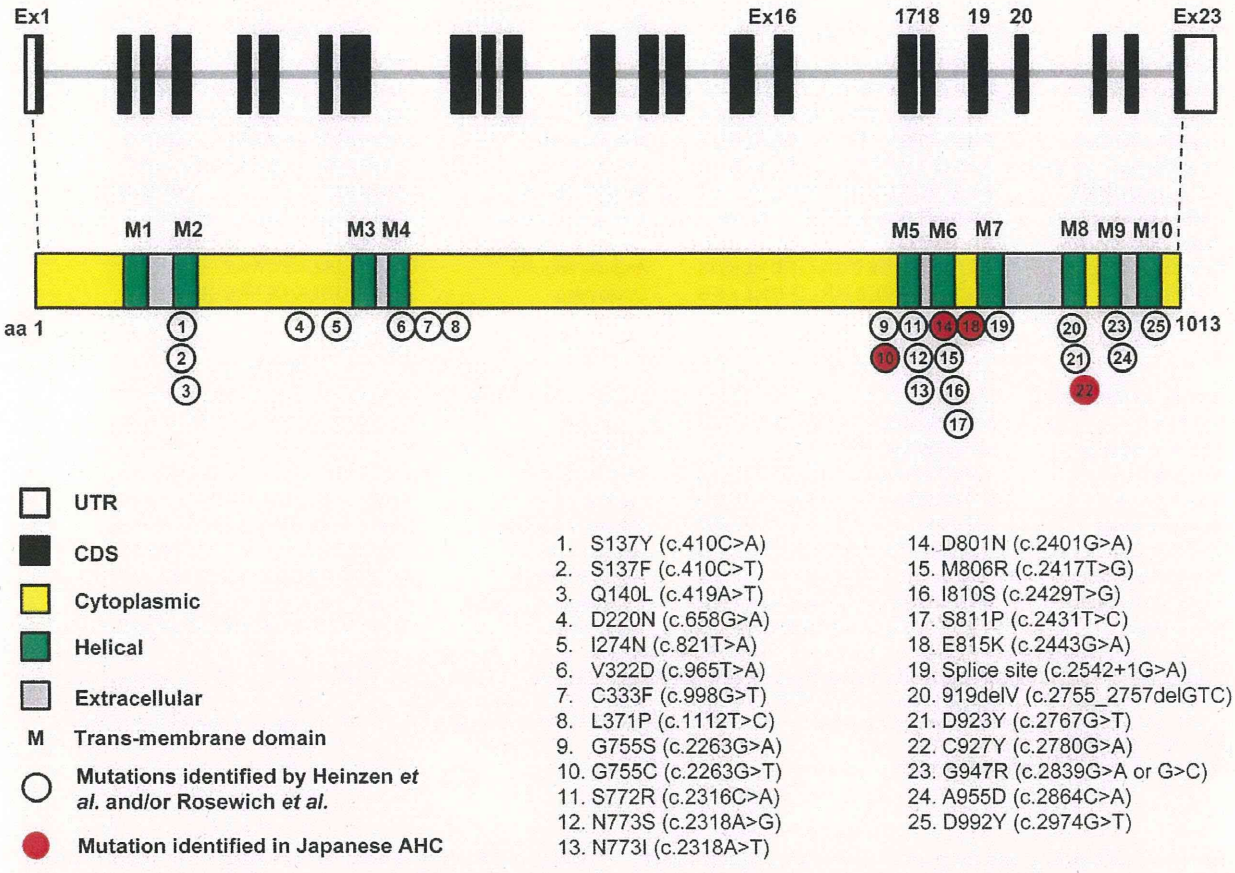

Figure 3. *ATP1A3* mutations and their protein domain structures. *Black lined circle:* Mutations reported recently [10,11]. *Red colored circle:* Mutations identified in the present study in a Japanese cohort with AHC. The *ATP1A3* gene consists of 23 exons that encode several domains in the ATP1A3 protein molecule, including 6 cytoplasmic, 10 helical and 5 extracellular domains. G755C and E815K were located in the cytoplasmic domains. Notably, E815K was resident of the transmembrane domain rather than the cytoplasmic domain. D801N and C927Y were located in the helical domains. C927Y was identified in this study only and hence considered novel.
doi:10.1371/journal.pone.0056120.g003

Functional analysis of *ATP1A3* mutations in RDP by haploinsufficiency demonstrated low protein levels of the corresponding ATPase [9]. In addition, Heinzen *et al.* demonstrated that none of the mutations causes AHC reduced protein levels, whereas both mutations of AHC and those of RDP reduced ATPase activity [10]. These studies suggested that mutations identified in AHC affect the Na^+/K^+ ATPase pump function due to inhibition of ion binding. This implies that D801N substitutions can cause pump dysfunction more than D801Y. Heterozygous knock-out mice and knock-in mice deficient in *ATP1A3* have been generated. The *ATP1A3* knock-out mice were found to have reduced NMDA receptors and exhibited neurological abnormalities such as hyperactivity, spatial learning and memory deficit [19]. The mice harboring mutation I810N of *ATP1A3*, which were neither RPD nor AHC, developed seizures [20]. While these phenotypes do not necessarily correspond with the typical clinical manifestations observed in either RDP or AHC, some similarities do exist.

In total, we identified four *ATP1A3* mutations in 10 Japanese AHC patients. All were heterozygous and *de novo*. Although the number of patients was small (10 individuals), E815K and D801N were observed in 5 (50%) and 3 (30%) of the 10 patients, respectively.

The exact mechanism of *de novo* mutation identified in this study is not clear at present. The nucleotides of both E815K and D801N are located in the GC-rich sequences of *ATP1A3*, and within 6-bp palindrome. These features may be related to the development of these *de novo* mutations.

Intriguingly, E815K mutation of *ATP1A3* found in half of our patients was associated with the presence of severe neurological symptoms, respiratory failure, status epilepticus and resistance to medications. The attending physicians consider, with hindsight clinical experience that flunarizine seems to be less effective in individuals with E815K mutation, compared to those with other mutations. However, the association between genotype and phenotype remains undefined due to the small number of the cohort. The present findings and those of other groups on AHC associated with *ATP1A3* mutations warrant further studies to understand the relation between genotype and phenotype in AHC and to develop new tools for the diagnosis and treatment of AHC.

Patients and Methods

Ethics statement

The present study was approved by the Ethics Review Committees of Fukuoka University and the University of Tokyo.

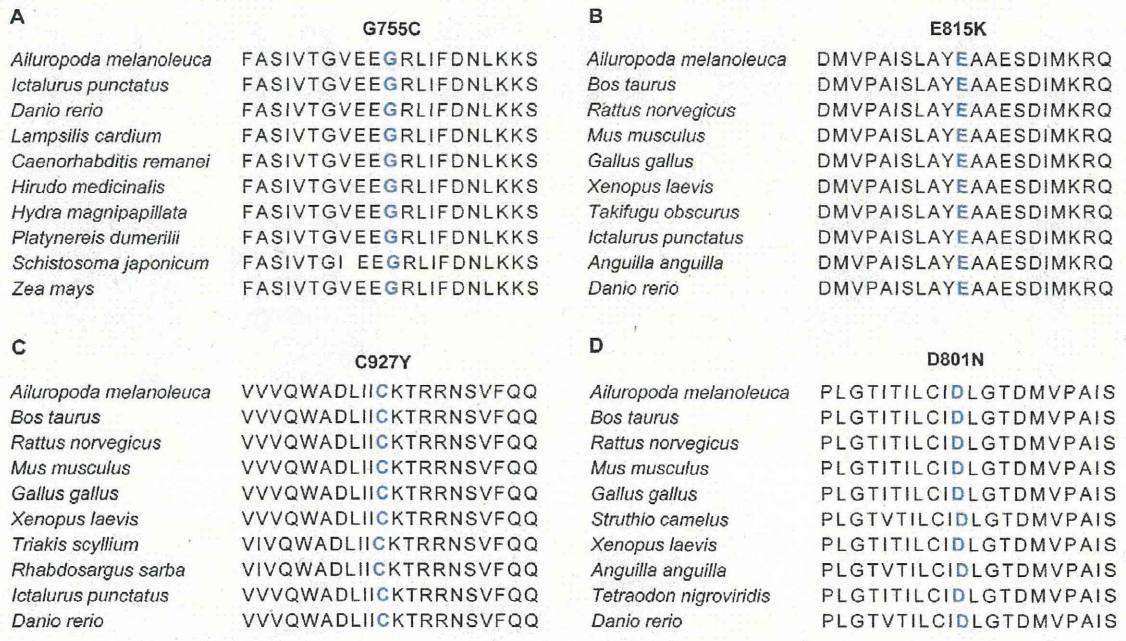

Figure 4. Homologous comparison of altering-protein. *Blue letters:* altering-protein by mutation, *red letters:* differential protein with human. (A) G755C changed by novel SNVs (c.2263G>T) of *ATP1A3* in Patient I-1. (B) E815K changed by novel SNVs (c.2443 G>A) of *ATP1A3* in Patients II-1, III-1, IV-1, IX-1 and X-1. (C) C927Y changed by novel SNVs (c.2780 G>A) of *ATP1A3* in Patient V-1. (D) D801N changed by novel SNVs (c.2401 G>A) of *ATP1A3* in Patient VI-1, VII-1 and VIII-1.
doi:10.1371/journal.pone.0056120.g004

Parents of each patient and the parents themselves provided signed informed consent before the study.

Patients

We initially recruited 10 unrelated Japanese individuals with clinical features of typical sporadic AHC. The diagnosis of AHC was based on the criteria of AHC [1,2]. The clinical presentations of these patients were typical but the neurological symptoms showed some variations, including aphonia, choreoathetosis, dyskinesia, epilepsy, and episodic apnea. Furthermore, variability in the response to different medications, such as flunarizine, was also noted among the patients (Table 3). Flunarizine was used for the treatment of 9 patients to control paralysis. The frequency of the paretic symptom decreased somewhat following the treatment, compared to that with other medications. However, the response to treatment, as evaluated subjectively by the attending physician, was not remarkable. Two patients (II-1 and V-1) showed a better response to clonazepam than to flunarizine.

The patients studied were 8 males and 2 females with similar clinical presentation, including infantile onset and psychomotor retardation. MRI images showed high-intensity hippocampal region in patient VIII-1 (Table 3), which was considered secondary to repeated episodes of epileptic convulsions. MRI images in patients II-1 and III-1 showed cerebellar atrophy, which was considered a primary lesion similar to FHM. The MRI findings in patient V-1 were considered non-specific.

Based on the availability of samples from the parents of the 9 patients, we selected 8 probands (subjects I-1 to VIII-1, Table 3) for exome sequencing analysis. After the identification of *de novo* heterozygous mutations in 8 patients, we also collected samples from the parents of patient IX-1 and also samples from patient X-1 and his parents. Parents of the patients with available genomic DNAs were also enrolled in this study. We also recruited 96 unrelated healthy Japanese volunteers as the control group who were free of seizures or history of epilepsy.

Genomic DNA was prepared from EDTA-Na_2-containing blood samples using the QIAamp DNA Blood Maxi Kit (Qiagen, Hilden, Germany), using the protocol provided by the manufacturer.

Exome sequencing

The exonic sequences were enriched using the Agilent SureSelect technology for targeted exon capture (213,383 exons, covering approximately 50 Mb of the CCDS database) (Agilent Technologies, Santa Clara, CA) from 3 μg of genomic DNA, using the protocol provided by the manufacturer. The captured DNAs were subjected to massively parallel sequencing (100 bp paired-end reads) on the Illumina Hiseq2000 (Illumina, San Diego, CA). The average of 1.3 billion bases of the sequence data was obtained for each individual. On average, 99.08% of the total bases were mapped to the reference genome with a mean coverage of 182.8x, which encompassed 92.99% of the targeted regions with coverage >10x. Burrows Wheeler Aligner [21] and Samtools [22] were used as default settings for alignment of raw reads and detection of variations. The variants were filtered against dbSNP (build 135). The aligned short reads were viewed using the University of Tokyo Genome Browser (UTGB) [23].

Sanger sequencing

Sanger sequencing was performed to validate the presence of each variant detected by exome sequencing in patients with AHC and the absence of each in the parental genomes. The entire exons and the intron-exon boundaries of *ATP1A3*, *CNTN4* (NM_175607) and *SYNE1* (NM_033071) were amplified by PCR using the

designed PCR primers (Table S1 lists the primer sequences and the PCR conditions). The PCR products were purified in ExoSAP-IT for PCR Product Clean-Up (Affymetrix, Santa Clara, CA) set at one cycle of 15 min at 37°C and 15 min at 80°C. The purified PCR products were sequenced using the ABI PRISM BigDye 3.1 terminator method (Applied Biosystems, Foster City, CA) and the ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems).

URLs

BLAST: http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD = Web&PAGE_TYPE = BlastHome Japanese Society of Alternating hemiplegia of childhood: http://www008.upp.so-net.ne.jp/ahc/

Accession numbers

Reference sequences are available from NCBI under the following accession codes: *CACNA1A*:NM_000068
ATP1A2:MN_000702
CNTN4: NM_175607
ATP1A3: NM_152296
SYNE1: NM_033071

Supporting Information

Figure S1 Rations of single nucleotide variations (SNVs) overlapping with known polymorphisms in various ethnic backgrounds.
(DOC)

Note S1 Brain-expressed genes.
(DOC)

Table S1 PCR primers and conditions designed for *ATP1A3*.
(DOC)

Acknowledgments

We thank all members of the family and also the "Society of AHC Japan" for their helpful cooperation in this study. The authors also thank Minako Yonetani and Akiyo Hamachi for the excellent technical assistance, and Takako Umemoto and Sumie Matsunaga for formatting and typing the manuscript.

Author Contributions

Confirmed the diagnosis in each patients participating in this study: MS YS. Conceived and designed the experiments: AI YS SM MS ST SH. Performed the experiments: AI JM HI. Analyzed the data: AI JY. Contributed reagents/materials/analysis tools: MS YS HA SY SK HO. Wrote the paper: AI ST SH.

References

1. Bourgeois M, Aicardi J, Goutieres F (1993) Alternating hemiplegia of childhood. J Pediatr 122: 673–679.
2. Sweney MT, Silver K, Gerard-Blanluet M, Pedespan JM, Renault F, et al. (2009) Alternating hemiplegia of childhood: early characteristics and evolution of a neurodevelopmental syndrome. Pediatrics 123: e534–541.
3. Bassi MT, Bresolin N, Tonelli A, Nazos K, Crippa F, et al. (2004) A novel mutation in the ATP1A2 gene causes alternating hemiplegia of childhood. J Med Genet 41: 621–628.
4. Neville BG, Ninan M (2007) The treatment and management of alternating hemiplegia of childhood. Dev Med Child Neurol 49: 777–780.
5. Ducros A, Denier C, Joutel A, Vahedi K, Michel A, et al. (1999) Recurrence of the T666M calcium channel CACNA1A gene mutation in familial hemiplegic migraine with progressive cerebellar ataxia. Am J Hum Genet 64: 89–98.
6. De Fusco M, Marconi R, Silvestri L, Atorino L, Rampoldi L, et al. (2003) Haploinsufficiency of ATP1A2 encoding the Na+/K+ pump alpha2 subunit associated with familial hemiplegic migraine type 2. Nat Genet 33: 192–196.
7. Vanmolkot KR, Kors EE, Hottenga JJ, Terwindt GM, Haan J, et al. (2003) Novel mutations in the Na+, K+-ATPase pump gene ATP1A2 associated with familial hemiplegic migraine and benign familial infantile convulsions. Ann Neurol 54: 360–366.
8. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, et al. (2011) Spatio-temporal transcriptome of the human brain. Nature 478: 483–489.
9. de Carvalho Aguiar P, Sweadner KJ, Penniston JT, Zaremba J, Liu L, et al. (2004) Mutations in the Na+/K+ -ATPase alpha3 gene ATP1A3 are associated with rapid-onset dystonia parkinsonism. Neuron 43: 169–175.
10. Heinzen EL, Swoboda KJ, Hitomi Y, Gurrieri F, Nicole S, et al. (2012) De novo mutations in ATP1A3 cause alternating hemiplegia of childhood. Nat Genet 44: 1030–1034.
11. Rosewich H, Thiele H, Ohlenbusch A, Maschke U, Altmuller J, et al. (2012) Heterozygous de-novo mutations in ATP1A3 in patients with alternating hemiplegia of childhood: a whole-exome sequencing gene-identification study. Lancet Neurol 11: 764–773.
12. McGrail KM, Phillips JM, Sweadner KJ (1991) Immunofluorescent localization of three Na,K-ATPase isozymes in the rat central nervous system: both neurons and glia can express more than one Na,K-ATPase. J Neurosci 11: 381–391.
13. Brashear A, DeLeon D, Bressman SB, Thyagarajan D, Farlow MR, et al. (1997) Rapid-onset dystonia-parkinsonism in a second family. Neurology 48: 1066–1069.
14. Dobyns WB, Ozelius LJ, Kramer PL, Brashear A, Farlow MR, et al. (1993) Rapid-onset dystonia-parkinsonism. Neurology 43: 2596–2602.
15. Linazasoro G, Indakoetxea B, Ruiz J, Van Blercom N, Lasa A (2002) Possible sporadic rapid-onset dystonia-parkinsonism. Mov Disord 17: 608–609.
16. Pittock SJ, Joyce C, O'Keane V, Hugle B, Hardiman MO, et al. (2000) Rapid-onset dystonia-parkinsonism: a clinical and genetic analysis of a new kindred. Neurology 55: 991–995.
17. Anselm IA, Sweadner KJ, Gollamudi S, Ozelius LJ, Darras BT (2009) Rapid-onset dystonia-parkinsonism in a child with a novel atp1a3 gene mutation. Neurology 73: 400–401.
18. Brashear A, Mink JW, Hill DF, Boggs N, McCall WV, et al. (2012) ATP1A3 mutations in infants: a new rapid-onset dystonia–Parkinsonism phenotype characterized by motor delay and ataxia. Dev Med Child Neurol 54: 1065–1067.
19. Moseley AE, Williams MT, Schaefer TL, Bohanan CS, Neumann JC, et al. (2007) Deficiency in Na,K-ATPase alpha isoform genes alters spatial learning, motor activity, and anxiety in mice. J Neurosci 27: 616–626.
20. Clapcote SJ, Duffy S, Xie G, Kirshenbaum G, Bechard AR, et al. (2009) Mutation I810N in the alpha3 isoform of Na+,K+-ATPase causes impairments in the sodium pump and hyperexcitability in the CNS. Proc Natl Acad Sci U S A 106: 14085–14090.
21. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.
22. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.
23. Saito TL, Yoshimura J, Sasaki S, Ahsan B, Sasaki A, et al. (2009) UTGB toolkit for personalized genome browsers. Bioinformatics 25: 1856–1861.

OPEN ACCESS Freely available online

PLOS | ONE

SORL1 Is Genetically Associated with Late-Onset Alzheimer's Disease in Japanese, Koreans and Caucasians

Akinori Miyashita[1☯], Asako Koike[2☯], Gyungah Jun[3☯], Li-San Wang[4], Satoshi Takahashi[5†], Etsuro Matsubara[6], Takeshi Kawarabayashi[6], Mikio Shoji[6], Naoki Tomita[7], Hiroyuki Arai[7], Takashi Asada[8], Yasuo Harigaya[9], Masaki Ikeda[10], Masakuni Amari[10], Haruo Hanyu[11], Susumu Higuchi[12], Takeshi Ikeuchi[13], Masatoyo Nishizawa[13], Masaichi Suga[14], Yasuhiro Kawase[15], Hiroyasu Akatsu[16], Kenji Kosaka[16], Takayuki Yamamoto[16], Masaki Imagawa[17], Tsuyoshi Hamaguchi[18], Masahito Yamada[18], Takashi Moriaha[19], Masatoshi Takeda[19], Takeo Takao[20], Kenji Nakata[21], Yoshikatsu Fujisawa[21†], Ken Sasaki[21], Ken Watanabe[22], Kenji Nakashima[23], Katsuya Urakami[24], Terumi Ooya[25], Mitsuo Takahashi[26], Takefumi Yuzuriha[27], Kayoko Serikawa[28], Seishi Yoshimoto[28], Ryuji Nakagawa[28], Jong-Won Kim[29], Chang-Seok Ki[29], Hong-Hee Won[29], Duk L. Na[30], Sang Won Seo[30], Inhee Mook-Jung[31], The Alzheimer Disease Genetics Consortium[‡], Peter St. George-Hyslop[32], Richard Mayeux[33], Jonathan L. Haines[34], Margaret A. Pericak-Vance[35], Makiko Yoshida[2], Nao Nishida[36], Katsushi Tokunaga[36], Ken Yamamoto[37], Shoji Tsuji[38], Ichiro Kanazawa[39], Yasuo Ihara[40], Gerard D. Schellenberg[4], Lindsay A. Farrer[3,41¶*], Ryozo Kuwano[1¶*]

1 Department of Molecular Genetics, Brain Research Institute, Niigata University, Niigata, Japan, **2** Central Research Laboratory, Hitachi Ltd, Tokyo, Japan, **3** Departments of Medicine (Biomedical Genetics), Ophthalmology and Biostatistics, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, United States of America, **4** Department of Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, United States of America, **5** Department of Neurology, Iwate Medical University, Morioka, Japan, **6** Department of Neurology, Hirosaki University Graduate School of Medicine, Hirosaki, Japan, **7** Department of Geriatric and Complementary Medicine, Tohoku University Graduate School of Medicine, Sendai, Japan, **8** Department of Psychiatry, University of Tsukuba, Tsukuba, Japan, **9** Department of Neurology, Maebashi Red Cross Hospital, Maebashi, Japan, **10** Department of Neurology, Gunma University Graduate School of Medicine, Maebashi, Japan, **11** Department of Geriatric Medicine, Tokyo Medical University, Tokyo, Japan, **12** Division of Clinical Research, Kurihama Alcoholism Center, Yokosuka, Japan, **13** Department of Neurology, Brain Research Institute, Niigata University, Niigata, Japan, **14** Higashi Niigata Hospital, Niigata, Japan, **15** Kawase Neurology Clinic, Sanjo, Japan, **16** Choju Medical Institute, Fukushimura Hospital, Toyohashi, Japan, **17** Imagawa Clinic, Osaka, Japan, **18** Department of Neurology and Neurobiology of Aging, Kanazawa University Graduate School of Medical Science, Kanazawa, Japan, **19** Department of Psychiatry, Osaka University Graduate School of Medicine, Osaka University, Osaka, Japan, **20** Kurashiki Heisei Hospital, Kurashiki, Japan, **21** Kinoko Espoir Hospital, Kasaoka, Japan, **22** Watanabe Hospital, Tottori, Japan, **23** Department of Neurology Tottori University, Yonago, Japan, **24** Department of Biological Regulation, Section of Environment and Health Science, Tottori University, Yonago, Japan, **25** Town Office, Onan, Japan, **26** Department of Clinical Pharmacology, Fukuoka University, Fukuoka, Japan, **27** Department of Psychiatry, National Hospital Organization, Hizen Psychiatric Center, Saga, Japan, **28** Ureshino-Onsen Hospital, Saga, Japan, **29** Department of Laboratory Medicine & Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea, **30** Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea, **31** Department of Biochemistry & Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea, **32** Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Canada, and the Department of Clinical Neurosciences, Cambridge Institute for Medical Research, Cambridge, United Kingdom, **33** Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, United States of America, **34** Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee, United States of America, **35** The John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida, United States of America, **36** Department of Human Genetics, University of Tokyo, Tokyo, Japan, **37** Department of Molecular Genetics, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan, **38** Department of Neurology, University of Tokyo, Tokyo, Japan, **39** National Center for Neurology and Psychiatry, Kodaira, Japan, **40** Department of Neuropathology, Doshisha University, Kyoto, Japan, **41** Departments of Neurology, Ophthalmology, Genetics & Genomics, and Epidemiology, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, United States of America

Abstract

To discover susceptibility genes of late-onset Alzheimer's disease (LOAD), we conducted a 3-stage genome-wide association study (GWAS) using three populations: Japanese from the Japanese Genetic Consortium for Alzheimer Disease (JGSCAD), Koreans, and Caucasians from the Alzheimer Disease Genetic Consortium (ADGC). In Stage 1, we evaluated data for 5,877,918 genotyped and imputed SNPs in Japanese cases (n = 1,008) and controls (n = 1,016). Genome-wide significance was observed with 12 SNPs in the *APOE* region. Seven SNPs from other distinct regions with p-values $<2\times10^{-5}$ were genotyped in a second Japanese sample (885 cases, 985 controls), and evidence of association was confirmed for one *SORL1* SNP (rs3781834, $P = 7.33\times10^{-7}$ in the combined sample). Subsequent analysis combining results for several SORL1 SNPs in the Japanese, Korean (339 cases, 1,129 controls) and Caucasians (11,840 AD cases, 10,931 controls) revealed genome wide significance with rs11218343 ($P = 1.77\times10^{-9}$) and rs3781834 ($P = 1.04\times10^{-8}$). SNPs in previously established AD loci in Caucasians showed strong evidence of association in Japanese including rs3851179 near *PICALM* ($P = 1.71\times10^{-5}$) and rs744373 near *BIN1* ($P = 1.39\times10^{-4}$). The associated allele for each of these SNPs was the same as in Caucasians. These data demonstrate for the first time genome-wide significance of LOAD with *SORL1* and confirm the role of other known loci for LOAD in Japanese. Our study highlights the importance of examining associations in multiple ethnic populations.

Citation: Miyashita A, Koike A, Jun G, Wang L-S, Takahashi S, et al. (2013) *SORL1* Is Genetically Associated with Late-Onset Alzheimer's Disease in Japanese, Koreans and Caucasians. PLoS ONE 8(4): e58618. doi:10.1371/journal.pone.0058618

Editor: Mathias Toft, Oslo University Hospital, Norway

Received November 2, 2012; **Accepted** February 5, 2013; **Published** April 2, 2013

Funding: This study was supported by a Grants-in-Aid for Scientific Research on Priority Areas, 'Advanced Brain Science Project' (Y.I.), 'Comprehensive Genomics' (R.K., K.T. and K.Y.), and 'Applied Genomics' (S.T.) and 'Integrated Database Project' (A.K.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, 'Japanese Alzheimer's Disease Neuroimaging Initiative' (R.K.) by New Energy and Industrial Technology Development Organization (NEDO) in Japan, and by a grant from the Korea Health Technology R&D Project, Ministry of Health & Welfare, Republic of Korea (A120030). The United States National Institutes of Health (NIH), National Institute on Aging (NIH-NIA) supported this work through the following grants: ADGC, U01 AG032984, RC2 AG036528; NACC, U01 AG016976; NCRAD, U24 AG021886; NIA LOAD, U24 AG026395, U24 AG026390; BU ADC P30; MIRAGE R01 AG025259; Banner Sun Health Research Institute P30 AG019610; Boston University, P30 AG013846, U01 AG10483, R01 CA129769, R01 MH080295, R01 AG017173, R01AG33193; Columbia University, P50 AG008702, R37 AG015473; Duke University, P30 AG028377, AG05128; Emory University, AG025688; Group Health Research Institute, UO1 AG06781, UO1 HG004610; Indiana University, P30 AG10133; Johns Hopkins University, P50 AG005146, R01 AG020688; Massachusetts General Hospital, P50 AG005134; Mayo Clinic, P50 AG016574; Mount Sinai School of Medicine, P50 AG005138, P01 AG002219; New York University, P30 AG08051, MO1RR00096, and UL1 RR029893; Northwestern University, P30 AG013854; Oregon Health & Science University, P30 AG008017, R01 AG026916; Rush University, P30 AG010161, R01 AG019085, R01 AG15819, R01 AG17917, R01 AG30146; TGen, R01 NS059873; University of Alabama at Birmingham, P50 AG016582, UL1RR02777; University of Arizona, R01 AG031581; University of California, Davis, P30 AG010129; University of California, Irvine, P50 AG016573, P50, P50 AG016575, P50 AG016576, P50 AG016577; University of California, Los Angeles, P50 AG016570; University of California, San Diego, P50 AG005131; University of California, San Francisco, P50 AG023501, P01 AG019724; University of Kentucky, P30 AG028383; University of Michigan, P50 AG008671; University of Pennsylvania, P30 AG010124; University of Pittsburgh, P50 AG005133, AG030653; University of Southern California, P50 AG005142; University of Texas Southwestern, P30 AG012300; University of Miami, R01 AG027944, AG010491, AG027944, AG021547, AG019757; University of Washington, P50 AG005136; Vanderbilt University, R01 AG019085; and Washington University, P50 AG005681, P01 AG03991. The Kathleen Price Bryan Brain Bank at Duke University Medical Center is funded by NINDS grant #NS39764, NIMH MH60451 and by Glaxo Smith Kline. Genotyping of the TGEN2 cohort was supported by Kronos Science. The TGen series was also funded by NIA grant AG034504 to AJM, The Banner Alzheimer's Foundation, The Johnnie B. Byrd Sr. Alzheimer's Institute, the Medical Research Council, and the state of Arizona and also includes samples from the following sites: Newcastle Brain Tissue Resource (funding via the Medical Research Council, local NHS trusts and Newcastle University), MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council), South West Dementia Brain Bank (funding via numerous sources including the Higher Education Funding Council for England (HEFCE), Alzheimer's Research Trust (ART), BRACE as well as North Bristol NHS Trust Research and Innovation Department and DeNDRoN), The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, International Parkinson Fonds, Internationale Stiching Alzheimer Onderzoek), Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona. Marcelle Morrison-Bogorad, PhD, Tony Phelps, PhD, and Walter Kukull, PhD, are thanked for helping to co-ordinate this collection. ADNI Funding for ADNI is through the Northern California Institute for Research and Education by grants from Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., Alzheimer's Association, Alzheimer's Drug Discovery Foundation, the Dana Foundation, and by the National Institute of Biomedical Imaging and Bioengineering and NIA grants U01 AG024904, RC2 AG036535, K01 AG030514. The authors thank Drs. D. Stephen Snyder and Marilyn Miller from NIA who are ex-officio ADGC members. Support was also from the Alzheimer's Association (LAF, IIRG-08-89720; MP-V, IIRG-05-14147) and the US Department of Veterans Affairs Administration, Office of Research and Development, Biomedical Laboratory Research Program. P.S.G.-H. is supported by the Wellcome Trust, Howard Hughes Medical Institute, and Canadian Institute of Health Research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: We have the following interests. The Kathleen Price Bryan Brain Bank at Duke University Medical Center is funded by NINDS grant #NS39764, NIMH MH60451 and by Glaxo Smith Kline. Genotyping of the TGEN2 cohort was supported by Kronos Science. Asako Koike and Makiko Yoshida are employed by Central Research Laboratory, Hitachi Ltd. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

* E-mail: ryosun@bri.niigata-u.ac.jp (RK); farrer@bu.edu (LAF)

❡ These authors contributed equally to this work.

¶ These authors also contributed equally to this work.

‡ Membership of The Alzheimer Disease Genetics Consortium is provided in the Acknowledgments.

† Deceased.

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder characterized by cognitive dysfunction and memory loss. Multiple rare mutations in *APP*, *PSEN1*, *PSEN2* and *SORL1* account for most cases of early-onset autosomal dominant AD [1,2]. Risk of late-onset AD (LOAD), the most common type of dementia in the elderly, is associated with complex interactions between genetic and environmental factors. Until recently, *APOE* was the only unequivocally recognized major susceptibility gene for LOAD [1,3]. Several genome-wide association studies (GWAS) each including more than 5,000 Caucasians identified genome-wide significant associations for LOAD with nine other loci including *ABCA7*, *BIN1*, *CD2AP*, *CD33*, *CLU*, *CR1*, *EPHA1*, *MS4A* gene cluster, and *PICALM* [4,5]. To our knowledge, no large GWAS for LOAD has been performed in any Asian population. Because there is a possibility that there exist ethnic-specific LOAD susceptibility variants, we carried out a large-scale GWAS to confirm associations at known loci and identify novel loci for LOAD using a three-stage design including a discovery Japanese cohort and replication cohorts of Japanese, Korean and Caucasian subjects.

Methods

Subjects

Japanese datasets. Clinically defined subjects were recruited by the Japanese Genetic Study Consortium of Alzheimer's Disease (JGSCAD: principal investigator, Y.I.) [6,7]. Probable AD cases were ascertained on the basis of the criteria of the National Institute of Neurological and Communicative Disorders, and Stroke-Alzheimer's Disease and Related Disorders (NINCDS/ADRDA) [8]. The Mini-Mental State Examination [9], Clinical Dementia Rating [10], and/or Function Assessment Staging [11] were primarily used for evaluation of cognitive impairment. Elders living in an unassisted manner in the local community with no signs of dementia were used as controls. DNA was extracted from peripheral blood leukocytes using standard protocols [6]. For the purpose of this study, the Stage 1 genome-wide association study (GWAS) dataset included 2024 subjects (1008 AD cases and 1016

controls) and the Stage 2 dataset included 1870 subjects (885 AD cases and 985 controls).

Korean dataset. A total of 339 subjects with AD were recruited at the Samsung Medical Center in Seoul, Korea. All AD subjects fulfilled NINCDS-ADRDA criteria for probable AD [8]. These subjects underwent a clinical interview and neurological examination that were previously described [12]. The absence of secondary causes of cognitive deficits was assessed by laboratory tests including complete blood count, blood chemistry, vitamin B12/folate, syphilis serology, and thyroid function tests. Conventional brain MRI scans (T1-weighted, T2-weighted, and FLAIR images) confirmed the absence of territorial cerebral infarctions, brain tumors, and other structural lesions. Healthy control subjects (n = 1,129) ages 55 to 85 years were recruited from routine health examination at the same location and showed no evidence of cognitive dysfunction.

Alzheimer Disease Genetics Consortium dataset. Summarized information from tests of genetic association of AD with SNPs located in the candidate gene regions was culled from a recent large GWAS conducted by the Alzheimer Disease Genetics Consortium (ADGC) [5]. Results were computed for SNPs throughout the genome in a sample composed of 11,840 AD cases and 10,931 cognitively normal elders from 15 independent Caucasian data sets. Details of the quality control and statistical analysis procedures and genetic models has been published elsewhere [5].

This study was approved by the Boston University Institutional Review Board, Institutional Review Board of Niigata University, and the Institutional Review Boards of all participating institutions. Written informed consent was obtained from all participants. Next of kin, carer takers or guardians consented on the behalf of participants whose capacity to consent was compromised. All subjects were anonymously genotyped. The basic demographics of the cases and controls before QC in each dataset are presented in Table 1.

Genotyping

GWAS genotyping was performed in the Stage 1 sample using Affymetrix GeneChip 6.0 microarrays containing 909,622 SNPs. Applied Biosystems' (ABI) TaqMan Assays were used to genotype individual SNPs in the Japanese and Korean replication cohorts. *APOE* genotypes in the Japanese and Korean samples were determined by haplotypes derived from rs7412 and rs429358 which were genotyped using TaqMan Assays. Details of *APOE* genotyping in each ADGC dataset were described previously [13].

Quality Control and Population Substructure

In the Stage 1 sample, SNPs with a genotype call rate (GCR) <95%, a minor allele frequency (MAF) <0.05, or significant deviation from the Hardy-Weinberg equilibrium (HWE) in controls ($P<10^{-6}$) were excluded. After excluding 83,673 low quality and 298,304 low frequency SNPs, we removed 196 subjects with a GCR <95% and 41 subjects whose gender as determined by analysis of X-chromosome data using the PLINK program (ver. 1.06) [14] was inconsistent with the reported gender. The same QC procedures were applied to the Japanese and Korean replication datasets. We examined population substructure in the GWAS dataset by analyzing tagging SNPs from the genome-wide panels using the *smartpca* module from EIGENSTRAT software [15] in a manner described previously [5]. Subsequently, we excluded three subjects who were cryptically related to other subjects in the dataset and 49 individuals who were population outliers. The strength of association of the top 10 principal components (PCs) was tested with AD status. The first

Table 1. Sample size and characteristics of the discovery and replication datasets.

Population (Stage)	Total	Alzheimer Disease Cases				Cognitively Normal Controls			
		N	Female (%)	Age at onset (mean ± SD)	Age at exam (mean ± SD)	*APOE* ε2/ε3/ε4 Frequency	N	Female (%)	Age at exam (mean ± SD)	*APOE* ε2/ε3/ε4 Frequency
Japanese Discovery (Stage 1)	2,024	1,008	723 (72%)	73.0 (4.28)	NA	0.02/0.65/0.33	1,016	583 (57%)	77.0 (5.89)	0.04/0.87/0.09
Japanese Replication (Stage 2)	1,870	885	574 (65%)	74.3 (6.98)	NA	0.02/0.69/0.29	985	618 (63%)	73.74 (5.84)	0.05/0.86/0.09
Korean (Stage 3)	1,468	339	245 (72%)	NA	73.67 (9.49)	0.03/0.70/0.27	1,129	550 (49%)	71.04 (4.86)	0.06/0.85/0.09
Caucasian (Stage 3)	22,771	11,840	7168 (61%)	76.37 (5.18)	80.59 (4.92)	0.04/0.61/0.36	10,931	6418 (59%)	76.77 (3.55)	0.08/0.78/0.14
TOTAL	**28,133**	**14,072**					**14,061**			

doi:10.1371/journal.pone.0058618.t001

three PCs were nominally associated with AD status. A total of 574,828 SNPs and 1,735 subjects comprising 891 cases and 844 controls passed the QC and were used for imputation and in further statistical analyses.

Genotype Imputation

Genotypes for all SNPs in Japanese and Caucasians were imputed with the Markov Chain haplotyping (MaCH) software [16] using reference haplotypes in the 1000 Genomes database (version released in February 2012 for Japanese datasets and version released in December 2010 for Caucasian datasets). This procedure also filled in missing data for the genotyped SNPs. Imputation quality was determined as R^2, which estimates the squared correlation between imputed and true genotypes. We applied threshold criteria for quality control assessment of imputed SNPs ($R^2 \geq 0.8$) as recommended for 1000 Genomes imputed data using the IMPUTE2 program [17]. Genotype probabilities for 5,877,918 genotyped and reliably imputed SNPs with a minor allele frequency (MAF) >0.02 were included in the Japanese GWAS.

Statistical analysis

Genotyped and imputed SNPs were tested for association with AD in the Stage 1 dataset using a logistic generalized linear model (GLM) controlling for age-at-onset (cases)/age-at-exam (controls), sex and the first three principal components from analysis of of population substructure. Stage 1 analyses were also performed based on a model adjusting for these covariates and the number of *APOE* ε4 alleles. SNPs in the *APOE* region (between map positions 45,000 kb and 45,800 kb on chromosome 19) were also tested for association in ε3/ε3 and ε3/ε4 subgroups. Genotyped SNPs were coded as 0, 1, or 2 according to the number of minor alleles under the additive genetic model. For imputed SNPs, a quantitative estimate between 0 and 2 for the dose of the minor allele were used to incorporate the uncertainty of the imputation estimates. All analyses were performed using PLINK. SNPs attaining a P value below 5×10^{-5} were considered for replication in Stage 2. Initially, only one SNP per region was tested in the replication sample to minimize the penalty for multiple testing. Additional SNPs from regions meeting the signifcance threshold in the replication sample were also evaluated. SNPs with a P value below 1×10^{-5} in the combined Stages 1 and 2 samples and nominally significant in Stage 2 ($P<0.05$) were advanced to Stage 3.

SNP association results obtained from individual datasets were combined by meta-analysis using the inverse variance method implemented in the software package METAL (http://www.sph.umich.edu/csg/abecasis/Metal/index.html) [18]. An additive model was assumed and the association results across datasets were combined by summing the regression coefficients weighted by the inverse variance of the coefficients. The meta-analysis P-value of the association was estimated by the summarized test statistic, after applying a genomic control within each individual study. Effect sizes were weighted by their inverse variance and a combined estimate was calculated by summing the weighted estimates and dividing by the summed weights.

Results

The quantile-quantile plot indicated limited genomic inflation ($\lambda = 1.04$in the Stage 1 GWAS results (Fig. S1). A total of 125 SNPs from seven distinct regions showed evidence of association with $P<10^{-4}$ (Table S1, Fig. S2). In addition to *APOE* SNP rs429358 ($P=2.46\times10^{-49}$, OR [95% CI] =5.5 [4.4–6.9]), 12 other SNPs in the *APOE* region were associated with LOAD at the genome-wide significance level of $P<5.0\times10^{-8}$. The two most significant results in this group of SNPs were rs12610605 (*PVRL2*: $P=1.38\times10^{-13}$, OR [95% CI] =1.8 [1.5–2.0]) and rs62117161 (between *CEACAM16* and *BCL3*: $P=3.46\times10^{-12}$, OR [95% CI] =0.47 [0.38–0.58]). Since imputation in the *APOE* region using the 1000 Genomes reference panel is unreliable [6], we genotyped nine SNPs from this region, spanning multiple linkage disequilibrium (LD) blocks (Fig. S3) and that were nominally significant in the *APOE* ε3/ε3 subgroup, in the Japanese discovery and replication samples using TaqMan assays (Table S2). Genome-wide significant results were obtained for five of these SNPs, but only the association with *PPP1R37* SNP rs 17643262 remained nominally significant after adjustment for the number of *APOE* ε4 alleles ($P=3.96\times10^{-4}$) or in analyses stratified by *APOE* genotype (ε3/ε3: $P=0.01$; ε3/ε4: P = 0.0016).

SNPs from six other distinct chromosomal regions met Stage 2 follow-up criteria ($P<5\times10^{-5}$) and the top SNP from each region was genotyped in an independent Japanese sample (Table 2). Two SNPs were nominally significant in the replication sample, however the effect direction for *KIAA0494* SNP rs7519866 differed from the discovery sample. Modest evidence for replication was observed only with *SORL1* SNP rs4598682 ($P\leq0.05$). Subsequently, we selected an additional four *SORL1* SNPs (rs3781834, rs2282647, rs17125523, and rs3737529) for testing in the Japanese replication sample that were among the most significant in the basic or extended models in the discovery sample (Table S1) and not in LD with rs4598682 ($r^2<0.2$, Figure S4). Two of these SNPs (rs3781834 and rs17125523) were chosen also because they were genotyped in the discovery sample and thus would minimize the effects of potential imputation artifacts in meta-analysis of the two Japanese samples. Highly significant results were obtained for *SORL1* SNPs rs4598682 ($P=9.51\times10^{-6}$), rs3781834 ($P=7.33\times10^{-7}$), rs17125523 ($P=5.51\times10^{-6}$), and rs3737529 ($P=4.14\times10^{-6}$) after combining results from the discovery and replication samples (Table S3).

These four *SORL1* SNPs showing significant association in the combined samples from Stages 1 and 2 were considered for further replication in Stage 3. We added rs11218343 to this stage of the analysis because it was the most significant *SORL1* SNP in the large Caucasian dataset ($P=1.0\times10^{-7}$), a result which emerged after pooling the Caucasian discovery GWAS sample and unpublished data in the replication sample from our previously published GWAS [5]. These five SNPs were subsequently evaluated in Stage 3 by meta analysis including the Stage 1 and 2 Japanese, Korean and ADGC Caucasian datasets. SNPs rs11218343 ($P=2.20\times10^{-9}$) and rs3781834 ($P=9.90\times10^{-9}$), attained genome-wide significance in the sample of datasets from all stages (Table 3, Fig. 1). There was modest evidence of replication for rs17125523 (meta $P=3.30\times10^{-6}$) and rs 3737529 (meta $P=5.10\times10^{-6}$). Although the allele frequencies for the top SNPs were very different between the Asian (MAF >0.2) and Caucasian (MAF <0.05) samples (Table 3), there was no evidence of heterogeneity in the magnitude of the odds ratios or effect direction among the population groups ($P>0.15$, Fig. 2). There was no apparent association in the comparably smaller Korean dataset; however, the direction of the effect for each SNP was the same as in the Japanese and Caucasian datsets.

Next, we investigated whether robust genetic associations for LOAD reported previously in Caucasians [4,5] generalize to Japanese. After correcting for 15 tests, SNPs rs3851179 located approximately 90 kb upstream from *PICALM* ($P=1.71\times10^{-5}$) and rs744373 located approximately 30 kb upstream from *BIN1* ($P=1.39\times10^{-4}$) were significantly associated with LOAD risk in the Japanese Stage 1 dataset (Table 4). Nominally significant

Table 2. Top-ranked genome-wide association results in the Japanese discovery (Stage 1) sample ($P<2.5\times10^{-5}$) and their replication in Japanese (Stage 2).

SNP	CH:MB	Nearest Gene	MA	MAF	# SNPs	Discovery (Stage 1)		Replication (Stage 2)		Meta-Analysis (Stages 1+2)	
						OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
rs7519866	1:47.0	*KIAA0494*	G	0.37	52	0.71 (0.61–0.83)	9.70×10^{-6}	1.15 (1.01–1.32)	0.04	0.90 (0.57–1.44)	0.67
rs913360	9:111.7	*PALM2*	G	0.28	20	1.56 (1.43–1.70)	1.83×10^{-7}	1.11 (0.96–1.29)	0.16	1.29 (1.15–1.44)	6.6×10^{-6}
rs1273007	10:9.0	*LOC338591*	T	0.27	39	0.68 (0.62–0.74)	3.08×10^{-6}	0.95 (0.81–1.10)	0.47	0.81 (0.73–0.91)	2.2×10^{-4}
rs10898417	11:85.2	*SYTL2*	G	0.15	2	0.59 (0.53–0.66)	1.17×10^{-6}	1.02 (0.85–1.22)	0.83	0.82 (0.71–0.93)	0.003
rs4598682	11:121.1	*SORL1*	G	0.23	11	0.68 (0.57–0.81)	2.25×10^{-5}	0.83 (0.68–1.00)	0.05	0.75 (0.66–0.85)	9.5×10^{-6}
rs11621843	14:92.2	*RIN3*	G	0.26	19	1.47 (1.35–1.60)	5.19×10^{-6}	1.03 (0.88–1.20)	0.72	1.21 (1.08–1.36)	8.1×10^{-4}

CH:MB, chromosome:position (in megabasepairs, build 19); MA, minor allele; MAF, minor allele frequency; # SNPs, the number of SNPs for which $P\leq1\times10^{-4}$ in the discovery (Stage 1) sample; OR, odds ratio; *P* P-value;
Selected SNPs represent the strongest association within each locus.
doi:10.1371/journal.pone.0058618.t002

associations were also observed for SNPs in *CR1*, *CLU*, and *ABCA7*. Of the eight SNPs tested in the small Korean sample, nominally signficant results ($P<0.05$) were obtained for one SNP in *CLU* and *PICALM*, each with the same pattern of association and comparable effect size as in Japanese.

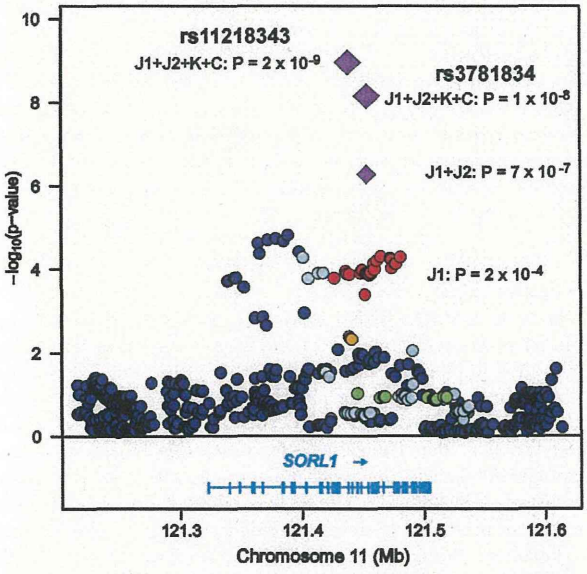

Figure 1. Regional association plot for the *SORL1* region on chromosome 11 in the three-stage design. For each SNP, the chromosomal location is shown on the x-axis and the significance level for association with LOAD is indicated by a $-\log_{10}$P value on the y-axis. P-values are expressed as $-\log_{10}$(P) (y-axis) for every tested SNP ordered by chromosomal location (x-axis). Genomic position was determined using the NCBI database (Build 37.1). Computed estimates of linkage disequilibrium (LD; r^2) between SNPs in this region with the top-ranked SNP (rs3781834) in the Japanese discovery (J1) dataset are shown as red circles for $r^2\geq0.8$, orange circles for $0.5\leq r^2<0.8$, light blue circles for $0.2\leq r^2<0.5$, and dark blue circles for $r^2<0.2$ using hg19/1000 Genomes of Asian populations (ASN; release on November 2010) combined from Han Chinese (CHB) and Japanese (JPT). Meta-analysis *P*-values are shown as purple diamonds for the Japanese datasets (J1+J2) and all datasets (J1+J2+K+C) including Japanese, Korean (K), and Caucasians (C). Two genome-wide significant SNPs in the final stage (rs3781834 and rs11218343) are presented. The gene structure and reading frame are shown below the plot. Exons are denoted with vertical bars. The LD between rs3781834 and rs11218343 is 0.57 in the ASN reference population.
doi:10.1371/journal.pone.0058618.g001

Discussion

Our multi-stage GWAS of LOAD identified for the first-time genome-wide significant association with *SORL1*. Genetic association with *SORL1* was first established in a study focused on genes encoding proteins involved in vacuolar protein sorting [19]. Most, but not all, subsequent studies in Caucsians replicated this finding (summarized in Alzgene database: http://www.alzgene.org/). Confirmatory evidence of association with *SORL1* SNPs has also been reported in comparatively small samples of Chinese and Japanese (reviewed in [20]). These findings are independent of previous candidate gene studies of *SORL1* in Japanese (two subjects in common) and with Caucasians in the Rogaeva et al. study [19] (less than 2% overlap).

The two genome-wide significant *SORL1* SNPs, rs11218343 and rs3781834 are located at chromosome positions 121,435,587 base pairs and 121,445,940 base pairs, respectively, and thus between the two previously reported strongly associated 3-marker haplotypes that extend upstream from rs641120 (121,380,965 base pairs) and downstream from rs1699102 (121,456,962 base pairs) [19]. A recent meta-analysis including more than 30,000 Caucasian and Asian subjects demonstrated that multiple *SORL1* SNPs in distinct regions are associated with AD [20], a finding substantiated in an association study of *SORL1* SNPs with brain MRI traits in LOAD families [21]. Further analysis of our large Caucasian sample suggests that the association peak at rs3781834 is independent of at least one of the two distinct haplotypes previously associated with AD in an independent sample of non-Hispanic Caucasians, Caribbean Hispanics and Israeli-Arabs (Fig. S5) [19], Since all of the SNPs at the association peaks reported in this study and previously are intronic, functional studies are required to determine the identity of pathogenic variants at these locations.

Remarkably, the less frequent alleles at rs11218343 and rs3781834 are protective in both Japanese and Caucasian datasets with very similar odds ratios (range 0.74 to 0.83) despite the fact that these alleles are much rarer in Caucasians (4% and 2%, respectively) than in Japanese (34% and 23%, respectively). The rarity of these SNPs in Caucasians, as well as allelic heterogeneity, may explain why *SORL1* did not previously emerged as a genome-