HAL open science

Somatic variants in autosomal dominant genes are a rare cause of sporadic Alzheimer's disease

Gaël Nicolas, Rocio Acuna-Hidalgo, Michael Keogh, Olivier Quenez, Marloes Steehouwer, Stefan H. Lelieveld, Stéphane Rousseau, Anne-Claire Richard, Manon Oud, Florent Marguet, et al.

To cite this version:

Gaël Nicolas, Rocio Acuna-Hidalgo, Michael Keogh, Olivier Quenez, Marloes Steehouwer, et al.. Somatic variants in autosomal dominant genes are a rare cause of sporadic Alzheimer's disease. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 2018, 14 (12), pp.1632-1639. 10.1016/j.jalz.2018.06.3056 . hal-02540061

HAL Id: hal-02540061

https://normandie-univ.hal.science/hal-02540061v1

Submitted on 28 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Alzheimer's & Dementia 14 (2018) 1632-1639

Alzheimer's & Dementia

Featured Article

Somatic variants in autosomal dominant genes are a rare cause of sporadic Alzheimer's disease

Gaël Nicolas[a,b,*], Rocío Acuña-Hidalgo[a,c], Michael J. Keogh[d,e], Olivier Quenez[b], Marloes Steehouwer[a], Stefan Lelieveld[a], Stéphane Rousseau[b], Anne-Claire Richard[b], Manon S. Oud[a], Florent Marguet[f], Annie Laquerrière[f], Chris M. Morris[g], Johannes Attems[g], Colin Smith[h], Olaf Ansorge[i], Safa Al Sarraj[j], Thierry Frebourg[b], Dominique Campion[b,k], Didier Hannequin[l], David Wallon[m], Christian Gilissen[a], Patrick F. Chinnery[d,e], Joris A. Veltman[a,n], Alexander Hoischen[a,o]

[a]*Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands*
[b]*Normandie Univ, UNIROUEN, Inserm U1245 and Rouen University Hospital, Department of Genetics and CNR-MAJ, F 76000, Normandy Center for Genomic and Personalized Medicine, Rouen, France*
[c]*Max Planck Institute for Molecular Genetics, RG Development & Disease, Berlin, Germany*
[d]*Department of Clinical Neurosciences, University of Cambridge, Cambridge Biomedical Campus, Cambridge, CB2 0QQ, UK*
[e]*MRC Mitochondrial Biology Unit, University of Cambridge, Cambridge, CB2 0XY, UK*
[f]*Normandie Univ, UNIROUEN, Inserm U1245 and Rouen University Hospital, Department of Neuropathology, F 76000, Normandy Center for Genomic and Personalized Medicine, Rouen, France*
[g]*Institute of Neuroscience, Newcastle University, Newcastle Upon Tyne, NE4 5PL, UK*
[h]*National CJD Research & Surveillance Unit, Centre for Clinical Brain Sciences, University of Edinburgh, Western General Hospital, Edinburgh, EH4 2XU, UK*
[i]*Department of Neuropathology, John Radcliffe Hospital, Oxford, OX3 9DU, UK*
[j]*Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, SE5 8AF, UK*
[k]*Department of research, Rouvray Psychiatric Hospital, Sotteville-lès-Rouen, France*
[l]*Normandie Univ, UNIROUEN, Inserm U1245 and Rouen University Hospital, Department of Genetics, Department of Neurology and CNR-MAJ, F 76000, Normandy Center for Genomic and Personalized Medicine, Rouen, France*
[m]*Normandie Univ, UNIROUEN, Inserm U1245 and Rouen University Hospital, Department of Neurology and CNR-MAJ, F 76000, Normandy Center for Genomic and Personalized Medicine, Rouen, France*
[n]*Institute of Genetic Medicine, International Centre for Life, Newcastle University, Newcastle upon Tyne, United Kingdom*
[o]*Department of Internal Medicine and Radboud Center for Infectious Diseases (RCI), Radboud University Medical Center, Nijmegen, The Netherlands*

Abstract

Introduction: A minority of patients with sporadic early-onset Alzheimer's disease (AD) exhibit *de novo* germ line mutations in the autosomal dominant genes such as *APP*, *PSEN1*, or *PSEN2*. We hypothesized that negatively screened patients may harbor somatic variants in these genes.
Methods: We applied an ultrasensitive approach based on single-molecule molecular inversion probes followed by deep next generation sequencing of 11 genes to 100 brain and 355 blood samples from 445 sporadic patients with AD (>80% exhibited an early onset, <66 years).
Results: We identified and confirmed nine somatic variants (allele fractions: 0.2%–10.8%): two *APP*, five *SORL1*, one *NCSTN*, and one *MARK4* variants by independent amplicon-based deep sequencing.
Discussion: Two of the *SORL1* variant might have contributed to the disease, the two *APP* variants were interpreted as likely benign and the other variants remained of unknown significance. Somatic variants in the autosomal dominant AD genes may not be a common cause of sporadic AD, including early onset cases.
© 2018 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Mosaicism; Post-zygotic; Mutation; Alzheimer; Prion-like

Conflicts of interest: The authors report that they have no conflicts of interest to disclose.

*Corresponding author. Tel.: +33 2 35 14 82 80; Fax: +33 2 35 14 82 37.
E-mail address: gaelnicolas@hotmail.com

https://doi.org/10.1016/j.jalz.2018.06.3056
1552-5260/ © 2018 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Background

In the vast majority of the Alzheimer's disease (AD) cases, the disease is considered as a complex disorder with a high genetic component as part of a multifactorial determinism (for review, see [1]). However, AD can be inherited as an autosomal dominant trait in a few families, with highly penetrant pathogenic genetic variants in the *APP*, *PSEN1*, or *PSEN2* genes. These variants are sufficient to cause the disease, usually before the age of 66 years (early-onset Alzheimer's disease [EOAD]). *APP* encodes the precursor of the amyloid-β (Aβ) peptide, the aggregation of which triggers AD pathophysiology. Aβ is generated following the cleavage of APP by the β-secretase (encoded by *BACE1*) and the γ-secretase complex, the catalytic subunit of which is encoded by *PSEN1* or *PSEN2*. *APP*, *PSEN1*, and *PSEN2* pathogenic variants are typically identified in families with autosomal dominant EOAD, that is, at least two generations showing at least one relative affected by EOAD. However, patients with sporadic EOAD, that is, negative family history, have also been reported to carry a pathogenic variant in these genes. Recently, 18/129 (14%) patients with sporadic EOAD and an age of onset before 51 years were reported to present a pathogenic *PSEN1* variant or an *APP* duplication [2], although it was only 2/90 (2.2%) in patients with a relatively later onset (51–65) [3]. Importantly, the mutation occurred *de novo* in all 10 cases where DNA from the unaffected parents was available [2]. In addition, whole exome sequencing (WES) of EOAD patients and their unaffected parents revealed *de novo* germline mutations in two novel genes: *VPS35* and *MARK4* [4]. Overall, no pathogenic variants are found in a majority of the patients with sporadic AD undergoing screening for mutations in the known genes [3–5].

It has been hypothesized for decades that post-zygotic or even somatic, brain-specific, variants could cause the disease in a proportion of sporadic AD patients but remain undetected by standard sequencing techniques [6,7]. Recent advances in sequencing technologies currently allow the accurate assessment of this hypothesis for the first time (for review see [8]). For instance, deep sequencing of *APP*, *PSEN1*, *PSEN2*, and *MAPT* was recently applied to DNA isolated from the brain of 72 sporadic AD patients and 58 controls [9]. In another study, WES was performed in brain-blood paired samples of 17 sporadic AD patients (average depth of coverage: 60.8×) [10]. Although some somatic variants could be detected, no clear pathogenic variant was identified in these studies. Of note, the majority of the previously published patients exhibited a late onset of AD (after 65 years). One could hypothesize that, similar to inherited or *de novo* germline pathogenic variants, somatic variants with high penetrance could be associated with an early onset.

The first sequencing study of single neurons from nondiseased human brains recently revealed a high load of somatic genetic variations. The number of somatic single nucleotide variants could be as high as 1500 per neuronal genome [11,12]. Interestingly, most of the variants that were present in more than 5%–10% of the neurons were also detected in tissues originating from all three embryonic layers. This suggests that, if brain tissue is not available for sequencing, sequencing DNA isolated from other tissues including blood can allow the detection of post-zygotic variants. Whatever the tissue of detection and allelic ratios, assessing the pathogenicity of a given variant still requires accurate interpretation. Regarding AD, we found one example of a post-zygotic pathogenic *PSEN1* variant detected in 8% of the blood cells and 14% in the brain cells of an EOAD patient [13].

Given the knowledge on seeding and spreading of neuropathological lesions in AD brains [14], we hypothesized that patients without a germline pathogenic variant in AD autosomal dominant genes may harbor post-zygotic or somatic variants. The primary aim of this study was to assess the presence of post-zygotic or somatic variants in *APP*, *PSEN1*, and *PSEN2* in patients with sporadic AD using single-molecule molecular inversion probes (smMIPs). The smMIP technology uses molecular barcodes (unique molecular identifiers [UMI]) to allow for molecule-specific deep sequencing. This is therefore an ultrasensitive technique for the detection of low-level mosaics [15,16]. Our secondary aim was to assess the presence of post-zygotic or somatic variants in 8 additional genes, namely *BACE1*, *NCSTN*, *APH1A*, *APH1B*, *PSENEN*, *SORL1*, *VPS35*, and *MARK4*. We applied molecule-specific deep sequencing of this panel of 11 genes to DNA isolated from blood (355 samples) or from brain (100 samples) from 445 sporadic AD patients from France, the UK, and the Netherlands (Table 1).

2. Methods

We included 445 patients fulfilling the National Institute on Aging–Alzheimer's Association criteria for probable AD or a definite diagnosis of AD (i.e., high AD neuropathologic change according to National Institute on Aging–Alzheimer's Association criteria [17]) and a negative family history, one positive control carrying a pathogenic *PSEN1* variant, and 52 cognitively normal controls. All cases recruited by the French National Reference Center for Young Alzheimer Patients (CNRMAJ, Rouen, France) from multiple French centers exhibited an early onset (<66 years), the cases recruited by the Netherlands Brain Bank exhibited either an early onset or, when the age at onset was not available, age at death was before 76 years, and cases recruited by the Medical Research Council (MRC) brain bank were not selected in the light of ages of onset; nine of them had an early onset. Among cases, DNA was isolated from blood (n = 355 samples) and/or from brain tissue (n = 100 samples) (Table 1, Supplementary Tables S1–4). DNA was

Table 1
Inclusion of cases for ultrasensitive sequencing

Study	N patients (only blood)	N patients (only brain)	N patients (blood + brain)	Total N patients	Mean age at onset (range)	Mean age at death (range)
Rouen CNRMAJ, France	347	2*	2†	351	54.42 (44–65)	NA
MRC Brain Bank, UK	0	80	0	80	69.9 (53–82)‡	85 (71–99)
Netherlands Brain Bank	0	8	6	14	56.4 (48–63)§	66.9 (57–75)
Total	Total blood samples: 355 form 355 patients Total brain samples: 100 from 98 patients			445		

Abbreviation: MRC, Medical Research Council.
*One sample from cerebellum and one sample from frontal cortex for one patient, one sample from an unspecified region for the second patient.
†One sample from cerebellum and one sample from frontal cortex in one patient, one sample from frontal cortex for the second patient.
‡Among the 29/80 patients with available information.
§Among the 12/14 patients with available information.

isolated from blood for all 52 controls. All cases except those from the Netherlands Brain Bank (Netherlands Institute for Neuroscience, Amsterdam; open access: www.brainbank.nl) were previously negatively screened for germline pathogenic variants in *APP*, *PSEN1*, and *PSEN2*, either by whole exome or by Sanger sequencing [18,19]. All participants or their legal representatives provided written informed consent for genetic analyses and/or for a brain autopsy and for the use of the material and clinical information for research purposes. Ethical approval for the genetic analysis of postmortem brain tissue was obtained from the ethical review board of each participating center. For details on inclusion, see Supplementary Methods.

We designed and set up an ultrasensitive smMIP assay aiming at sequencing the coding regions of 11 genes including the three autosomal dominant AD genes (*APP*, *PSEN1*, and *PSEN2*), the genes recently identified in a trio-exome sequencing study in sporadic EOAD cases *VPS35* and *MARK4*, the risk factor gene *SORL1*, and, as an exploratory study, *BACE1* encoding the β-secretase, and the genes encoding the other members of the γ-secretase complex *NCSTN*, *PSENEN*, *APH1A*, and *APH1B*. After rebalancing the concentration of the smMIP pool following a first test run, we performed four independent runs of sequencing on an Illumina NextSeq sequencer (runs A-D, see Supplementary Methods). All cases, the positive control and 16 of the cognitively normal controls were assessed with two independent polymerase chain reaction (PCR) amplifications of the smMIP capture, while the remaining 32 controls were amplified once.

Raw bioinformatics data were processed following three distinct pipelines, all three contained a PCR duplicates removal step using the UMI information: BWA-GATK, Seqnext (JSI medical systems), and an in-house pipeline based on the pileup format as generated by SAMtools. Briefly, the latter pipeline consisted in the computation of a base-specific error rate per run as published previously [16], based on pileup formats, followed by the calling of candidate somatic variants significantly deviating from the base-specific error rate, for both PCR duplicates, followed by manual curation. Candidate somatic variants were confirmed by independent amplicon-based deep sequencing.

Detailed methods on smMIP assay design, library preparation, sequencing, bioinformatics analyses including DNA contamination assessment, and amplicon deep sequencing are provided in Supplementary Methods.

3. Results

3.1. Coverage statistics

After removal of PCR duplicates thanks to the UMI, the average single-molecule coverage was 1027× per smMIP (seven failed samples were excluded). Regarding the three autosomal dominant AD genes *APP*, *PSEN1*, and *PSEN2*, the single-molecule average coverage at all bases of interest (coding exons ± 2 bp; 3101 bp) was 2576×; 97.6% of the bases of interest were covered by at least 100 unique reads among more than 97% of the samples.

3.2. Identification and validation of somatic variants

We detected nine candidate somatic variants in nine patients (seven blood samples, two brain samples; Table 2). We performed an independent validation by amplicon deep sequencing, using PCR followed by Ion Torrent Personal Genome Machine sequencing (average depth of coverage of all nine amplicons: 60,104×) and validated all nine variants as true somatic events (Table 2, Supplementary Table S5). The variant allelic fractions (VAFs) ranged from 0.22% to 10.8% and were in similar ranges after amplicon deep sequencing. Six somatic variants were novel, and three were present in the gnomAD database with very low frequencies (3, 12, and 14 allele counts, respectively) [20].

Two of these variants were missense variants in exons 6 and 7 the *APP* gene, respectively. However, as all known pathogenic variants are located in the coding sequence of the Aβ peptide or its boundaries (exons 16-17), these variants were interpreted as likely benign regarding their

Table 2
Somatic variants identified in patients

Patient ID	Age at onset	Gender	APOE	Sample	Gene symbol	Chromosome position	cDNA nomenclature*	p. nomenclature	PolyPhen 2	SIFT	Mutation taster	gnomAD MAF	VAF PCR1 (%)	VAF PCR2 (%)	VAF average (%)	Validation VAF (ADS) (%)
ROU-1496-001	65	F	33	Blood	APP	chr21:27425605	c.415G>A	p.(Glu139Lys)	Prob. Dam.	Delet.	DC	0	0.29	0.26	0.28	0.37
C7	72	M	34	Brain	APP	chr21:27372399	c.964G>T	p.(Gly322Cys)	Prob. Dam.	Delet.	DC	0	0.29	0.15	0.22	0.24
EXT-0772-001	50	F	33	Blood	MARK4	chr19:45762351	c.156G>A	p.(Glu52 =)	NA	NA	NA	0.000012†	0.49	0.37	0.43	0.20
ROU-0085-001	53	F	33	Blood	NCSTN	chr1:160319955	c.497C>T	p.(Ser166Leu)	Benign	Delet.	Pol.	0.000049†	11.10	10.56	10.83	9.92
ROU-1347-001	51	M	33	Blood	SORL1	chr11:121421320	c.2207G>A	p.(Cys736Tyr)	Prob. Dam.	Delet.	DC	0	3.65	3.58	3.61	2.82
ROU-0778-001	60	M	34	Blood	SORL1	chr11:121425931	c.2475G>A	p.(Val825 =)	NA	NA	NA	0	0.49	0.24	0.36	0.74
ROU-0609-001	55	F	34	Blood	SORL1	chr11:121424794	c.2415C>T	p.(Ser805 =)	NA	NA	NA	0.000051†	8.41	7.41	7.91	8.63
EXT-0482-001	58	M	33	Blood	SORL1	chr11:121483508	c.5386T>C	p.(Leu1796 =)	NA	NA	NA	0	0.49	0.47	0.48	0.81
C10	NA‡	F	44	Brain	SORL1	chr11:121489480	c.5605-3C>T	p.?	NA	NA	NA	0	0.38	0.53	0.45	0.59

Abbreviations: ADS, amplicon deep sequencing; p. nomenclature, protein nomenclature; VAF, variant allele fraction; MAF, minor allele frequency; Prob. Dam., probably damaging; Delet., deleterious; DC, disease causing; Pol., polymorphism; NA, not available.

Reference genome: GRCh37/hg19.

*Reference transcripts: NM_000484.3 (APP), NM_001199867.1 (MARK4), NM_015331.2 (NCSTN), NM_003105.5 (SORL1).

†Allelic ratios are compatible with a germline origin in gnomAD.

‡Age at death: 82 years.

putative contribution to AD. No candidate somatic variant was detected in *PSEN1* and *PSEN2* across all samples.

The other somatic variants were located in *SORL1* (n = 5, including one missense, three synonymous, and one intronic variant), *NCSTN* (n = 1, missense) and *MARK4* (n = 1, synonymous). One of the *SORL1* somatic variants (NM_003105.5:c.2207G>A, VAF = 3.61% in blood) was annotated as missense and predicted damaging by 3/3 *in silico* prediction tools among Mutation Taster, PolyPhen 2, and SIFT (strictly damaging). The other *SORL1* somatic variants were synonymous (n = 3) or intronic (n = 1). Of note, one of the synonymous variants was predicted to strongly enhance a cryptic 5′ splicing site (NM_003105.5:c.2475G>A, VAF = 0.36% in blood, MaxEntScan score +202%) and hence might disrupt the *SORL1* coding sequence. The intronic *SORL1* mutation was close to a canonical splice site (c.5605-3C>T) although splicing prediction tools suggested a weak or absent effect.

In four samples, from one particular sequencing run, additional variants were identified with allelic fraction in the ranges of 1% to 3%. However, we considered these results as putative DNA contamination because the variants were known as common polymorphisms in variant databases (minor allele frequency > 1%), each putatively contaminated sample harbored at least two of these variants, and they were detected as germline heterozygous or homozygous in other samples from the same run, all samples initially belonged to a single plate, before capture. The presence of DNA contamination was further assessed using the same technique based on the pileup formats as for candidate somatic variants, in all four runs, taking into account nucleotide changes that correspond to known SNPs. No additional contamination was identified.

3.3. *Interpretation of probably germ line* APP, PSEN1, *and* PSEN2 *variants*

After variant calling by GATK and SeqNext, followed by annotation and variant interpretation, we accurately detected the probably germline heterozygous *PSEN1* variant included as a positive control in one brain sample from the MRC brain bank (Supplementary Table S6). No probable germline (allelic ratio 25%–100%) variant was rated as pathogenic or likely pathogenic in these genes in cases. Of note, we confirmed the presence of four known heterozygous missense variants of unknown significance (class 3 following the American College of Medical Genetics and Genomics and the Association for Molecular Pathology recommendations [21], one in *PSEN1*, and three in *PSEN2*) in two French patients (blood samples) and two patients from the MRC brain bank (brain samples) (Supplementary Table S6), including the p.V101M *PSEN2* variant that has been previously reported in the brain of a patient with sporadic AD [9], also as a probably germline variant. Additional variants were detected in all three genes, but they were

classified as benign or likely benign based on their predicted effect, variant frequencies in controls, and previous reports.

3.4. *Probable germline* VPS35 *and* MARK4 *variants*

VPS35 and *MARK4* are candidate genes for autosomal dominant EOAD given the observation of *de novo* germline mutations in two sporadic EOAD patients, and subsequent *in vitro* studies showing biochemical defects consistent with AD pathophysiology [4].

We identified one rare nonsynonymous variant in *VPS35* and five in *MARK4*, all with a VAF suggestive of a germline origin (Supplementary Table S7). Of note, the c.2320C>A, p.Leu774Met *VPS35* variant was found in both the brain and blood tissues in a patient from the Netherlands Brain Bank. This variant has already been detected in 2/863 cases (Austrian and German patients) with early onset Parkinson disease and 2/1014 controls [22]. It is located in the same C-terminal domain of the protein as the p.Asp620Asn Parkinson disease–causing mutation and the p.Leu625Pro deleterious *de novo* germline variant found in an EOAD patient [4,22]. Although they mapped very close to each other in the protein sequence, the latter two variants had distinct consequences on the retromer complex function *in vitro*, which is consistent with their association with distinct phenotypes. The p.Leu774Met variant mapped 3' from this region and was not predicted to have a strong impact on protein stability. Interestingly, the father of the proband was known to suffer from Parkinson disease although without dementia (no clinical details or DNA available). This variant remains of unknown significance.

In *MARK4*, one of the variants was predicted benign and was inherited from an unaffected parent (c.1553C>T, p.(Pro518Leu)). One variant was predicted damaging by the three assessed *in silico* prediction tools but was exclusively found in one of the unaffected parents (c.88G>T, p.(Gly30Cys)). The other three variants were found in cases only. Although they were predicted damaging by all three *in silico* prediction tools (c.1033C>T, p.(Arg345Trp), French patient, blood sample) or by 1 or 2/3 (c.1982G>A, p.(Ser661Asn); c.230G>A, p.(Arg77Gln); MRC patients, brain samples), no conclusion can be drawn due to the fact that they were most probably present in as heterozygous germline and that no segregation data are available.

3.5. *Probable germline* SORL1 *variants*

We found 15 protein-truncating or missense *SORL1* variants that were considered as strictly damaging (i.e., predicted damaging by the three *in silico* prediction tools PolyPhen 2, SIFT, and Mutation Taster), in 17 patients and no control (Supplementary Table S8). These categories of variants have been shown to increase the risk of EOAD [18,23]. All VAF were in ranges suggesting that they were present in the germline. Among them, 12 variants present in 14 patients were already reported in previous WES studies [18,23], and three variants (two novel) were found in novel patients, identified from brain tissues.

3.6. *Probable germline variants in* BACE1 *and genes encoding members of the* γ*-secretase complex*

We detected 11 rare nonsynonymous variants in 12 patients in *BACE1* (n = 2), *NCSTN* (n = 4), *APH1A* (n = 1), *APH1B* (n = 3), and *PSENEN* (n = 1) (Supplementary Table S9). These variants were detected in 10 blood samples and two brains samples and the VAF suggested their germline origin. All but one were missense. A frameshift variant was detected in the *APH1B* gene. However, this gene is not under strong constraint against loss of function, similar to *PSEN2*, judging by the probability of loss of function intolerance established from Exome Aggregation Consortium data [20]. All were detected in patients.

4. Discussion

In this study, we screened 11 genes for somatic mutations in 355 blood samples and 100 brain samples from 445 patients with AD, of which 372 (83.5%) exhibited an early onset (<66 years). In total, we identified nine somatic variants with variant fractions ranging from 0.2% to 10.8%. These variants were detected in multiple DNA copies and are more likely clonal than recurrent mutational events. The coverage statistics, together with the validation of all variants detected, including all six with an allelic ratio below than 1% (range 0.22%–0.48%), support the ultrasensitivity of our detection method. We did not find any candidate post-zygotic or low-level somatic variant in the three established autosomal dominant AD-causative genes *APP*, *PSEN1*, and *PSEN2* that could be interpreted as likely pathogenic. Given the high sensitivity of the assay, we consider our screen as negative regarding likely pathogenic variants in the coding region of these genes.

We could find only one example in the literature of an AD patient with a post-zygotic causative variant in *PSEN1* [13]. In this study, a patient with EOAD starting at the age of 27 years was found to have inherited a pathogenic mutation in *PSEN1* from her affected mother, who presented a disease onset at the age of 52. The mutation was present in 8% of the mother's blood cells and 14% of her brain cells, suggesting that the mutation occurred as a post-zygotic event in the mother and that it was present in variable proportions of cells in multiple tissues including the mother's oocytes [13]. Of note, the majority of our patients presented an early onset of sporadic AD (83.5%), and therefore this is the largest series of sporadic EOAD patients screened for pathogenic somatic variants causative for AD to date. The assessment of the somatic variant hypothesis in sporadic AD has been performed only recently, using deep sequencing [9] or brain-blood paired WES [10], in patients with a later onset on average than in our study. To our knowledge, our screen is the first to leverage UMIs to allow single-molecule tracing

and even better sensitivity. Taken together, we consider that somatic variants in *APP, PSEN1*, and *PSEN2* are not a common cause of sporadic AD, even in patients with an early onset. We acknowledge, however, that somatic variants might still be present as even more rare events in brain regions, which have not been assessed. Indeed, this and previous studies focused only on one or two brain regions per individual. The interpretation of putative region-specific variants may however be difficult. In addition, our assay did not allow the identification of mosaic copy number variations.

As part as our gene panel, we also sequenced the *VPS35* and *MARK4* genes. They were each previously hit by one *de novo* germline mutation in sporadic EOAD patients [4]. The effect of these variants was studied *in vitro*, and the location where the mutation occurred in the protein could be highly specific, given the results of functional assays. Despite the identification of a synonymous somatic variant in *MARK4* (VAF = 0.43%), we could not identify any putatively damaging variant in the corresponding exons as a germline or a somatic variant.

Germline protein truncating and rare missense predicted to be strictly damaging *SORL1* variants significantly increase the risk of EOAD [18,23]. We detected five *SORL1* somatic variants (VAF ranging from 0.63% to 7.91%). Among them, one was missense and classified as strictly damaging. It was detected in a blood sample of an EOAD patient and could, if present in the brain tissue, contribute to the genetic determinism of AD in this patient. Among the other *SORL1* somatic variants, one was predicted to enhance a cryptic 5′ splicing site and could disrupt SORL1 coding sequence. If so and if present in the brain tissue, it could also contribute to the genetic determinism of AD in this patient. *SORL1* rare damaging variants were originally identified in EOAD probands with a positive family history of EOAD, with no pathogenic *APP*, *PSEN1*, or *PSEN2* variant [24]. However, the paucity of segregation data still precludes the classification of *SORL1* as a putative Mendelian gene and association studies showed a role as a risk factor (for review, see [1]). Our results suggest that the other genes *TREM2* and *ABCA7*, the rare damaging variants of which having been shown to increase the risk of AD, should also be screened for post-zygotic and somatic variation.

We included in our assay the candidate genes *BACE1* encoding the β-secretase and the other genes encoding the other proteins from the γ-secretase complex (in addition to *PSEN1* and *PSEN2*). We detected one somatic variant in *NCSTN*, which was present in ∼22% of the sequenced cells from the blood of one EOAD patient (VAF = 10.8%). This variant introduced a missense that was predicted damaging by SIFT but not by PolyPhen2 and Mutation Taster. It has been observed in 12 individuals from the gnomAD database (minor allele frequency = 4.9×10^{-5}) [20]. The visualization of the BAM files of the three variant carriers available in the gnomAD website suggested that this variant was compatible with a heterozygous variant with germline origin, which is not consistent with the hypothesis of a damaging effect when carried as a post-zygotic event. Interestingly, we also detected 12 variants in 11 patients that were probably present in the germline. All were detected in patients. To our knowledge, there is no evidence of rare variants in these genes segregating in families further than by chance, or of a significant association of rare variants with AD. This study was not designed as an association study, and these genes were not reported among the latest large association studies including our own data from France [18]. By including these genes that play a key role in Aβ generation in the context of the γ-secretase complex, we made the hypothesis that the absence of damaging variants segregating in families in the literature could be explained by a putative intolerance (abnormal development, lethality, and other diseases). Post-zygotic damaging variants might be better tolerated and putatively increase the generation of Aβ through increased β or γ-cleavage or its regulation. We did not find such candidates somatic variants in our study. These genes remain biological candidates currently lacking genetic evidence.

In conclusion, we used single-molecule deep sequencing in brain and/or blood samples of 445 patients with sporadic AD and could detect nine somatic variants with allelic ratios as low as 0.2%. Although we detected a few putatively damaging *SORL1* somatic variants, we did not detect any candidate post-zygotic or somatic variant that could be interpreted as pathogenic or likely pathogenic in the three known autosomal dominant AD genes. Our results, together with a previous report [9], challenge the hypothesis that somatic mutations in key AD genes would cause a significant proportion of AD with a sporadic presentation. We conclude that somatic variation in these genes is most likely not a frequent cause of sporadic AD.

Acknowledgments

The authors thank the MRC brain bank, the Netherlands brain bank, and Rouen University Hospital Brain bank for providing brain samples or DNA isolated from brain tissue. G.N. acknowledges Fondation Bettencourt-Schueller, Fondation Philippe Chatrier, Fondation Charles Nicolle, and Association Cerveau Progrès. This work was funded by grants from the CNR-MAJ. R.A.H. was funded by a PhD grant from the Radboud University Medical Center. This work was in part financially supported by grants from the Netherlands Organization for Scientific Research (918-15-667 to JAV).

This work was funded by the UK Medical Research Council (13044). M.J.K. is a Wellcome Trust Clinical Research Training Fellow. P.F.C. is a Wellcome Trust Senior Fellow in Clinical Science (101876/Z/13/Z), and a UK NIHR Senior Investigator, who receives support from the Medical Research Council Mitochondrial Biology Unit (MC_UP_1501/2), the Evelyn Trust, the Medical Research Council (UK) Center for Translational Muscle Disease Research (G0601943), EU FP7 TIRCON, and the National Institute for Health Research (NIHR) Biomedical Research

Center based at Cambridge University Hospitals NHS Foundation Trust and the University of Cambridge. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.
Regarding the UK group, tissue for this study was provided by the Newcastle Brain Tissue Resource, which is funded in part by a grant from the UK Medical Research Council (G0400074), by Brains for Dementia research, a joint venture between Alzheimer's Society and Alzheimer's Research UK, and by the NIHR Newcastle Biomedical Research Center awarded to the Newcastle upon Tyne Hospitals NHS Foundation Trust and Newcastle University.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jalz.2018.06.3056.

RESEARCH IN CONTEXT

1. Systematic review: Attention toward the somatic variant hypothesis is growing. This hypothesis states that a proportion of patients with Alzheimer's disease could have developed the disease because of somatic mutations in the brain, leading to pathological lesions that would later spread into the brain. However, we could find only one published example. Advances in sequencing technologies allow the assessment of this hypothesis since very recently only.
2. Interpretation: We assessed this hypothesis using an ultrasensitive molecule-specific deep sequencing approach in young patients. Nine somatic variants were identified, and some of them could have contributed to the development of the disease. However, no pathogenic variant was found in the known autosomal dominant genes, thus challenging the hypothesis.
3. Future directions: Other techniques could be applied to detect other genomic variations such copy number variations. In addition, genetic variants in a small proportion of cells - not detectable by our technique - could be a future research direction.

References

[1] Nicolas G, Charbonnier C, Campion D. From common to rare variants: the genetic component of Alzheimer disease. Hum Hered 2016; 81:129–41.
[2] Lanoiselee HM, Nicolas G, Wallon D, Rovelet-Lecrux A, Lacour M, Rousseau S, et al. APP, PSEN1, and PSEN2 mutations in early-onset Alzheimer disease: a genetic screening study of familial and sporadic cases. PLoS Med 2017;14:e1002270.
[3] Nicolas G, Wallon D, Charbonnier C, Quenez O, Rousseau S, Richard AC, et al. Screening of dementia genes by whole-exome sequencing in early-onset Alzheimer disease: input and lessons. Eur J Hum Genet 2016;24:710–6.
[4] Rovelet-Lecrux A, Charbonnier C, Wallon D, Nicolas G, Seaman MN, Pottier C, et al. De novo deleterious genetic variations target a biological network centered on Abeta peptide in early-onset Alzheimer disease. Mol Psychiatry 2015;20:1046–56.
[5] Cruts M, van Duijn CM, Backhovens H, Van den Broeck M, Wehnert A, Serneels S, et al. Estimation of the genetic contribution of presenilin-1 and -2 mutations in a population-based study of presenile Alzheimer disease. Hum Mol Genet 1998;7:43–51.
[6] Vitek MP, Rasool CG, de Sauvage F, Vitek SM, Bartus RT, Beer B, et al. Absence of mutation in the beta-amyloid cDNAs cloned from the brains of three patients with sporadic Alzheimer's disease. Brain Res 1988;464:121–31.
[7] Reznik-Wolf H, Machado J, Haroutunian V, DeMarco L, Walter GF, Goldman B, et al. Somatic mutation analysis of the APP and Presenilin 1 and 2 genes in Alzheimer's disease brains. J Neurogenet 1998; 12:55–65.
[8] Acuna-Hidalgo R, Veltman JA, Hoischen A. New insights into the generation and role of de novo mutations in health and disease. Genome Biol 2016;17:241.
[9] Sala Frigerio C, Lau P, Troakes C, Deramecourt V, Gele P, Van Loo P, et al. On the identification of low allele frequency mosaic mutations in the brains of Alzheimer's disease patients. Alzheimers Dement 2015; 11:1265–76.
[10] Parcerisas A, Rubio SE, Muhaisen A, Gomez-Ramos A, Pujadas L, Puiggros M, et al. Somatic signature of brain-specific single nucleotide variations in sporadic Alzheimer's disease. J Alzheimers Dis 2014; 42:1357–82.
[11] Lodato MA, Woodworth MB, Lee S, Evrony GD, Mehta BK, Karger A, et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. Science 2015;350:94–8.
[12] Lodato MA, Rodin RE, Bohrson CL, Coulter ME, Barton AR, Kwon M, et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. Science 2018;359:555–9.
[13] Beck JA, Poulter M, Campbell TA, Uphill JB, Adamson G, Geddes JF, et al. Somatic and germline mosaicism in sporadic early-onset Alzheimer's disease. Hum Mol Genet 2004;13:1219–24.
[14] Brettschneider J, Del Tredici K, Lee VM, Trojanowski JQ. Spreading of pathology in neurodegenerative diseases: a focus on human studies. Nat Rev Neurosci 2015;16:109–20.
[15] Hiatt JB, Pritchard CC, Salipante SJ, O'Roak BJ, Shendure J. Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. Genome Res 2013;23:843–54.
[16] Acuna-Hidalgo R, Sengul H, Steehouwer M, van de Vorst M, Vermeulen SH, Kiemeney L, et al. Ultra-sensitive sequencing identifies high prevalence of clonal hematopoiesis-associated mutations throughout adult Life. Am J Hum Genet 2017;101:50–64.
[17] Montine TJ, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Dickson DW, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach. Acta Neuropathol 2012;123:1–11.
[18] Bellenguez C, Charbonnier C, Grenier-Boley B, Quenez O, Le Guennec K, Nicolas G, et al. Contribution to Alzheimer's disease risk of rare variants in TREM2, SORL1, and ABCA7 in 1779 cases and 1273 controls. Neurobiol Aging 2017;59:220.e1–220.e9.
[19] Keogh MJ, Wei W, Wilson I, Coxhead J, Ryan S, Rollinson S, et al. Genetic compendium of 1511 human brains available through the UK Medical Research Council Brain Banks Network Resource. Genome Res 2017;27:165–73.
[20] Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285–91.

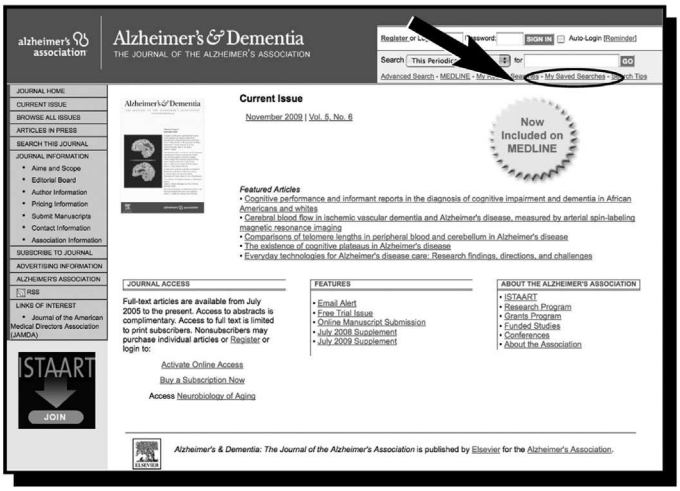
[21] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405–24.

[22] Zimprich A, Benet-Pages A, Struhal W, Graf E, Eck SH, Offman MN, et al. A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. Am J Hum Genet 2011;89:168–75.

[23] Nicolas G, Charbonnier C, Wallon D, Quenez O, Bellenguez C, Grenier-Boley B, et al. SORL1 rare variants: a major risk factor for familial early-onset Alzheimer's disease. Mol Psychiatry 2016;21:831–6.

[24] Pottier C, Hannequin D, Coutant S, Rovelet-Lecrux A, Wallon D, Rousseau S, et al. High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. Mol Psychiatry 2012;17:875–9.

Did you know?



You can save your online searches and get the results by email.

Visit www.alzheimersanddementia.org today!