MENDELIAN ADULT-ONSET LEUKODYSTROPHY GENES IN ALZHEIMER’S DISEASE. CRITICAL INFLUENCE OF CSF1R AND NOTCH3

Sassi C.a, b, c, Nalls M.A.b, Ridge P.G*, Gibbs R.b, Lupton M.K.f, Troakes C.f, Lunnon K.f, Al-Sarraj S.f, Brown K.S.f, Medway C.f, Lord J., Turton J.f, Bras J.m, ARUK Consortium*, Blumenau S.c, Thielke M.c, Josties C.c, Freyer D.c, Dietrich A.d, Hammer M.b, Baier M.l, Dirnagl U.c, Morgan K., Powell J.F.l, Kauwe J.S.q, Cruchaga C.f, Goate AM.e, Singleton A.B.b, Guerreiro R.m, Angela Hodgesl, Hardy J.a

*The Alzheimer’s Research UK (ARUK) Consortium: Peter Passmore, David Craig, Janet Johnston, Bernadette McGuinness, Stephen Todd, Queen’s University Belfast, UK; Reinhard Heun, Royal Derby Hospital, UK; Heike Kölsch, University of Bonn, Germany; Patrick G. Kehoe, University of Bristol, UK; Emma R.L.C. Vardy, Salford Royal NHS Foundation Trust, UK; Nigel M. Hooper, David M. Mann, Stuart Pickering-Brown, University of Manchester, UK; Kristelle Brown, James Lowe, Kevin Morgan, University of Nottingham, UK; A. David Smith, Gordon Wilcock, Donald Warden, University of Oxford (OPTIMA), UK; Clive Holmes, University of Southampton, UK.

a Reta Lila, Weston Research Laboratories, Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK;

b Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA;

c Department of Experimental Neurology, Center for Stroke Research Berlin (CSB), Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health Berlin, Germany

d Departments of Biology, Neuroscience, Brigham Young University, Provo, UT, USA

f King’s College London Institute of Psychiatry, London, UK

g QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

h Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, Devon, UK

i Translation Cell Sciences-Human Genetics, School of Life Sciences, Queens Medical Centre, University of Nottingham, Nottingham, UK

j Neurodegenerative Diseases, Robert-Koch-Institut, Berlin

m Department of Molecular Neuroscience, Institute of Neurology, University College London, London, UK; Department of Medical Sciences, Institute of Biomedicine-iBiMED, University of Aveiro, Aveiro, Portugal; UK Dementia Research Institute at UCL (UK DRI), London, UK

q Department of Neuroscience, Brigham Young University, Provo, UT, USA

r Division of Biology and Biomedical Sciences, Washington University, St. Louis, MO, USA

s Icahn School of Medicine at Mount Sinai, Icahn Medical Institute, New York, NY, USA

**Corresponding author**

Celeste Sassi, celeste.sassi@charite.de
ABSTRACT

Mendelian adult-onset leukodystrophies are a spectrum of rare inherited progressive neurodegenerative disorders affecting the white matter of the central nervous system. Among these, Cerebral Autosomal Dominant and Recessive Arteriopathy with Subcortical Infarcts and Leukoencephalopathy (CADASIL and CARASIL), Cerebroretinal vasculopathy (CRV), Metachromatic leukodystrophy (MLD), Hereditary diffuse Leukoencephalopathy with spheroids (HDLS), Vanishing white matter disease (VWM) present with rapidly progressive dementia as dominant feature and are caused by mutations in NOTCH3, HTRA1, TREX1, ARSA, CSF1R, EIF2B1, EIF2B2, EIF2B3, EIF2B4, EIF2B5, respectively. Given the rare incidence of these disorders and the lack of unequivocally diagnostic features, leukodystrophies are frequently misdiagnosed with common sporadic dementing diseases such as Alzheimer’s disease (AD), raising the question of whether these overlapping phenotypes may be explained by shared genetic risk factors. To investigate this intriguing hypothesis, we have combined gene expression analysis 1) in 6 different AD mouse strains (APPPS1, HOTASTPM, HETASTPM, TPM, TAS10 and TAU), at 5 different developmental stages (Embryo [E15], 2 months, 4 months, 8 months and 18 months), 2) in APPPS1 primary cortical neurons under stress conditions (oxygen-glucose deprivation) and single-variant and single-gene (c-alpha and SKAT tests) based genetic screening in a cohort composed of 332 Caucasian late-onset AD patients and 676 Caucasian elderly controls. Csf1r was significantly overexpressed (Log2FC>1, adj p-val<0.05) in the cortex and hippocampus of aged HOTASTPM mice with extensive Aβ core dense plaque pathology. We identified 3 likely pathogenic mutations in CSF1R TK domain (p.L868R, p.Q691H and p.H703Y) in our discovery and validation cohort, composed of 465 AD and MCI Caucasian patients from the UK. Moreover, NOTCH3 was a significant hit in the c-alpha test (adj p-val = 0.01). Adult onset Mendelian leukodystrophy genes are not common factors implicated in AD. Nevertheless, our study suggests a potential pathogenic link between NOTCH3, CSF1R and sporadic LOAD, that warrants further investigation.
**KEY WORDS**

Alzheimer’s disease, Mendelian Leukodystrophies, *CSF1R, NOTCH3*

1. INTRODUCTION

Mendelian adult-onset leukodystrophies are a spectrum of rare chronic progressive disorders affecting the white matter of the central nervous system. Although a growing body of literature is reporting newly discovered forms, the most characterized adult-onset leukodystrophies are Cerebral Autosomal Dominant and Recessive Arteriopathy with Subcortical Infarcts and Leukoencephalopathy (CADASIL and CARASIL), Cerebroretinal vasculopathy (CRV), Metachromatic Leukodystrophy (MLD), Hereditary diffuse Leukoencephalopathy with spheroids (HDLS), Vanishing white matter disease (VWM), caused by mutations in *NOTCH3, HTRA1, TREX1, ARSA, CSF1R, EIF2B1, EIF2B2, EIF2B3, EIF2B4, EIF2B5*, respectively (Joutel et al., 1996), (Hara et al., 2009), (Richards et al., 2007), (Fluharty et al., 1991), (Rademakers et al., 2011), (Scali et al., 2006). Given the rare incidence of these disorders (5/100 000 to only few cases reported), the lack of peculiar and distinctive 1) clinical features, generally represented by rapidly progressive dementia, behavioural changes, pyramidal and extrapyramidal signs and, less commonly, ischemic strokes and epileptic seizures; 2) MRI lesion patterns, normally characterized by T2-weighted periventricular and subcortical, patchy and later confluent white matter hyperintensities with prominent frontal involvement, and 3) neuropathological features, frequently a combination of diverse neurodegenerative hallmarks, these rare Mendelian disorders are most frequently underrecognized and misdiagnosed with common sporadic dementias such as Alzheimer’s disease (AD). On the other hand, motor features like ataxia and spasticity may appear in the course of AD progression, particularly in the cases caused by or associated to *PSEN1* mutations (Rossor et al., 2010) and AD patients may display MRI patterns and neuropathological features typical of adult-onset leukodystrophies (Smith et al., 2000) (Marnane et al., 2016) (Barber et al., 1999)(Guerreiro et al., 2013), suggesting a potential common pathogenic ground.

In the past 10 years, Next Generation Sequencing (NGS) paved the way for groundbreaking discoveries in AD, showing that Mendelian rare disorders offer a unique window into the
sporadic complex traits and, particularly, that rare alleles in *TREM2*, *TYROBP* and *NOTCH3*, causative for adult-onset leukodystrophies, significantly influence the susceptibility for AD (Guerreiro et al., 2013) (Ma et al., 2015) (Guerreiro et al., 2012). Moreover, the sequencing of different mouse strains showed extensive similarities between mouse and human genome and validated the importance of using mouse models to illuminate the genetics of human diseases (Cheng et al., 2014) (Yue et al., 2014). Nevertheless, NGS still presents 2 main challenges: 1) the huge amount of data generated is difficult to mine and 2) the investigation of rare coding variants requires several thousands of samples. Consequently, the need for experimental methods that accurately identify critical genes and strategies to empower association studies became priorities. Therefore, we have applied a combination of cortical and hippocampal gene expression analysis in 6 diverse AD mouse strains (APPPS1, HOTASTPM, HETASTPM, TPM, TAS10 and TAU), at 5 different developmental stages (embryo [E15], 2 months, 4 months, 8 months and 18 months) to comprehensively study leukodystrophy gene expression pattern in relation to the progression of AD neuropathology and under stress conditions such as oxygen-glucose deprivation (OGD), which represents an *in vitro* model of ischemic stroke, a common feature in several adult-onset leukodystrophies and frequent comorbidity in AD. We then used exome and genome sequencing data in a cohort composed of 332 Caucasian late-onset AD patients and 676 Caucasian elderly controls to investigate rare coding variability in these main adult-onset Mendelian leukodystrophy genes. Among the studied genes, *Csf1r* was the only gene significantly overexpressed (Log2FC>1, p-val<0.05) in AD mouse models and its expression tightly correlated with the severity of core dense plaque deposition. Moreover, we identified a total of 3 rare variants in *CSF1R* tyrosine kinase (TK) domain and TK flanking regions (p.L868R and p.D565N, p.G957R, respectively) present only in cases and very likely pathogenic. We then screened *CSF1R* in an independent cohort composed of 465 MCI and AD cases, identifying 2 additional mutations in *CSF1R* TK domain (p.Q691H and p.H703Y). Finally, *NOTCH3* was a significant hit in the gene-based analysis (adj p-value=0.01), suggesting a potential role as disease modifier. We conclude that rare coding variability in adult-onset Mendelian leukodystrophy genes is not a common risk factor for AD. However, *CSF1R* coding variants clustering in the TK domain and *NOTCH3* may influence AD susceptibility.
2. MATERIALS AND METHODS

Adult-onset Leukodystrophy gene selection

The selected genes are all Mendelian leukodystrophy causative genes with a core clinical hallmark represented by adult-onset subacute dementia with frontal predominance revealed by T2 weighted magneto resonance imaging (MRI) (Table S1). Moreover, all of these candidate genes present more than one of the following features: 1) previously reported misdiagnosis with Alzheimer’s disease (CADASIL, HDLS, RVCL, MLD) (Guerreiro et al., 2012) (Rademakers et al., 2011) (Richards et al., 2007) (Johannsen et al., 2001); 2) molecular interaction with other genes playing a key role in AD (NOTCH3, CSF1R) (Thijs et al., 2003) (Otero et al., 2009); 3) genes taking part to APP-Aβ metabolism (NOTCH3, HTRA1)(Grau et al., 2005), 4) co-presence of AD neuropathological hallmarks reported (NOTCH3 and CSF1R)(Paquet et al., 2010) (Baba et al., 2006) and 5) most frequently mutated genes in adults with leukoencephalopathies (NOTCH3, EIF2B4, EIF2B5, CSF1R) (Lynch et al., 2017).

The pipeline followed in this study is described in Fig 1.

Gene expression analysis

We have used microarray data publicly available (MouseAC database [http://www.mouseac.org/])(Matarin et al., 2015) and real-time PCR data to analyze Arsa, Csf1r, Eif2b1, Eif2b2, Eif2b3, Eif2b4, Eif2b5, Htra1, Notch3 and Trex1 gene expression 1) in the hippocampus and cortex of 6 different AD mouse strains (APPPS1, HOTASTPM, HETASTPM, TPM, TAS10 and TAU), 2) at 5 different time points (embryonic development [E15], 2 months, 4 months, 8 months and 18 months), to comprehensively follow expression changes related to Aβ plaques density (HOTASTPM, HETASTPM and TAS10), neurofibrillary tangles (TAU) and absence of pathology (E15, TPM). Adult APPPS1 data for hippocampus were available only for 2 months of age, where no plaques were reported but only rare Aβ oligomers in cortex and surrounding cortical vessels (Fig. S1). Finally, considering that ischemic stroke is a common feature in several leukodystrophies and frequent comorbidity in AD, we used an in vitro model of ischemic stroke and performed OGD experiments in
APPSPS1 primary cortical neurons to test whether leukodystrophy gene expression pattern may have significantly differed between APPPS1 and WT mice under stress conditions.

Genetic screening

The discovery cohort

The discovery cohort was composed of 332 apparently sporadic AD cases and 676 elderly controls, neuropathologically and clinically confirmed, originating from the UK and North America. The mean age at disease onset was 71.66 years (range 41-94 years) for cases and the mean age of ascertainment was 78.15 years (range 60-102 years) for controls. The majority of the cases (77%) were late-onset (> 65 years at onset). Among the cases and controls, 42% and 51% were female, respectively. 58% and 47% of the cases and controls carried the APOE ε4 allele, respectively. The APOE ε4 allele was significantly associated to the disease status in the NIH and ADNI series (p-value= 0.02 and 1.19x10^{-9}, respectively). This cohort has already been described elsewhere (Sassi et al., 2016). The threshold call rate for inclusion of the subject in analysis was 95%. On this cohort we performed 1) gene-based analysis (SKAT and c-alpha tests) and 2) single-variant association analysis. Finally, we followed-up, in an independent Caucasian dataset, CSF1R, the only gene significantly overexpressed during AD most severe pathology (Figure 1, Table S2).

The follow-up dataset

The follow-up dataset was composed of 296 AD and 169 MCI late-onset cases (mean age at onset >75y) from UK (Table S2). Written informed consent was obtained for each clinically assessed individual and the study was approved by the appropriate institutional review boards. All samples had fully informed consent for retrieval and were authorized for ethically approved scientific investigation (UCLH Research Ethics Committee number 10/H0716/3, BYU IRB, Cardiff REC for Wales 08/MRE09/38+5, REC Reference 04/Q2404/130, National Research Ethics Service [NRES]).

Exome and genome sequencing

DNA was extracted from blood or brain for cases and brain only for controls using standard protocols. Library preparation for next generation sequencing was based on Roche Nimblegen.
Inc. or TruSeq, Illumina protocols and has been described elsewhere (Sassi et al., 2016). Genome sequencing was performed in 199 controls, from the Cache County Study on Memory in Aging. All samples were sequenced with the use of Illumina HiSeq technology. Sequence alignment and variant calling were performed against the reference human genome (UCSC hg19) and has been described in the eMaterials.

Initial analysis excluded pathogenic mutations in APP, PSEN1, PSEN2, MAPT, GRN and TREM2. All variants within the coding regions of the 10 adult-onset leukodystrophy candidate genes (ARSA [NM_000487]; CSF1R [NM_005211]; EIF2B1 [NM_001414]; EIF2B2 [NM_014239]; EIF2B3 [NM_001261418]; EIF2B4 [NM_001034116]; EIF2B5 [NM_003907]; HTRA1 [NM_002775]; NOTCH3 [NM_000435] and TREX1 [NM_016381] have been collected and analysed, including 20.8 Megabase pairs (Mbs) of coding sequence.

**Sanger sequencing**

Mutations in CSF1R TK domain and flanking regions were validated with Sanger Sequencing. CSF1R was screened in an additional follow-up cohort composed of 296 AD and 169 MCI cases (Supplementary Materials and Methods).

**Statistical Analysis**

In the single-variant analysis, allele frequencies were calculated for each low frequency and rare coding variants in cases and controls and Fisher’s exact test on allelic association was performed. The threshold call rate for inclusion of the subjects and genetic variants in analysis was 95%. MouseAC data have been analyzed and FDR correction was applied.

The Supplement provides a more detailed description of the methods used (mouse and human gene expression analysis, oxygen-glucose deprivation experiments, sanger sequencing, statistical analysis and bioinformatics).

3. RESULTS

**Gene expression analysis**

We do not report any significant differential expression in Arsa, Csf1r, Eif2b1, Eif2b2, Eif2b3, Eif2b4, Eif2b5, Htra1, Notch3 and Trex1 until the development of severe AD pathology, markedly pronounced in the most aggressive AD strain studied, HOTASTPM, homozygous for
the Swedish mutation APP p.K670N/M671L and PSEN1 p.M146V, 8 months of age (Fig. 2a-d, Table S3 and S4). Here, Csf1r was up to 2 folds significantly overexpressed both in the hippocampus and cortex (Log2FC=1.2 and 1.1; adj p-value= 2.5e-07 and 8.7E-05, respectively) and presented a trend at 18 months both in hippocampus and cortex (Log2FC=0.75 and 0.98; adj p-value= 2.7e-04 and 3e-04, respectively), in linear correlation with the most rapid and severe core dense plaque deposition (0.8 core-dense plaque/month and 0.5 core-dense plaque/month between 4m-8m and 8m-18m of age, respectively)(http://www.mouseac.org/)(Fig2a-d, Table S3 and S4). Moreover, Csf1r overexpression positively correlated also with tau pathology, suggesting that Csf1r upregulation is not Aβ plaque specific. By contrast, Csf1r was downregulated when plaque deposition was minimal (HETTAS10TPM, TAS10, TPM and TAU, 2m; TAS10 and TAU, 4m) (Table S3 and S4). Importantly, Csf1r upregulation relied on microglia infiltration and was co-expressed with other microglia markers such as Aif1, CD68, Trem2, Tyrobp and Grn. Particularly, Csf1r and Grn displayed the same pattern of overexpression, which was between one third to one fourth of Tyrobp and Trem2 overall upregulation (Fig. 3, Table S3 and S4).

Embryonal hippocampi and Primary neuronal cortical cultures oxygen glucose deprivation (OGD) experiments

We do not report any significant differential expression in the studied genes between APPPS1 and WT embryonal hippocampi and in APPPS1 and WT primary cortical neurons after OGD experiments (Table S5). This is likely due to the fact that most of the leukodystrophy genes are expressed on microglia, only moderately present in E15 hippocampal neurons and in primary neuronal cortical cultures. In line with this observation, Csf1r and its ligands (Csf1 and Il34), Grn, Trem2 and Aif2 were significantly overexpressed in both APPPS1 and WT adult hippocampi compared to the embryonal ones (Log2FC=4, 2.45, 7.9, 1.5, 2.24, 3.4 and 4.2, 2.7, 7.9, 1.77, 3.1 and 3.8, respectively)(Table S5). By contrast, Notch3 was up to 2 folds upregulated in both APPPS1 and WT embryonal hippocampi compared to adult hippocampi. Moreover, we noticed that TREAT 5’UTR displays the typical features of many transcripts, like BACE1, that are translationally controlled by cellular stress (O’Connor et al., 2008): TREAT 5’UTR is indeed particularly long (628nts), GC-rich (65%) and predicted to contain 6 upstream open reading frames (uORFs)(http://www.ncbi.nlm.nih.gov/orffinder/) (Fig. S2a,b), suggesting TREAT transcript might be a target of translation control by one or more stress-
activated pathway. Therefore, we have investigated TREX1 protein levels in both APPPS1 and WT adult brains and we do not report any macroscopically significant difference (Fig. S2c-IIV). This may be due to the fact that APPPS1 mice used for these experiments, being 2 months of age, did not display a severe pathology (Fig. S1).

Genetic screening
The study population consisted of a total of 332 sporadic and mainly late-onset AD cases and 676 elderly controls of British and North American ancestry. We do not report any pathogenic mutation in APP, PSEN1 and PSEN2 in our cohort. However, one of the controls was a heterozygous carrier of the protective variant APP p.A673T (MAF 7x10E-4 in our cohort and MAF 5x10E-4 among the European non-Finnish, ExAC database, released 13 January 2015).

We performed a single-variant and a single-gene association analysis in a pre-defined set of adult-onset Mendelian Leukodystrophy genes (ARSA, CSF1R, EIF2B1, EIF2B2, EIF2B3, EIF2B4, EIF2B5, HTRA1, NOTCH3, TREX1) including 20.8 Megabase pairs (Mbs) of coding sequence.

A total of 215 single nucleotide variants (SNVs) has been identified. Among these, 77(35.8%) were nonsynonymous, 59 (27.4%) were synonymous, 13 (6%) UTR variants. Among the missense variants, 192 (95%) were very rare (MAF<1%), 16 (7.9%) were low frequency (1%<MAF<5%) and 12(5.9%) were common (MAF>5%). In addition, we report 4 novel coding variants (NOTCH3, p.A2146E, CSF1R p.G957R and p.D565N and ARSA p.H425Y). Variant minor allele frequency and novel variants were based on ExAC database, European non-Finnish panel and EVS European-American panel, released March 14, 2016, or dbSNP 137 (Table S6).

The overall variant frequency in our cohort was in line with the variant frequency reported in the American-European cohort in the Exome Variant server database (Table S7).

Single-gene based analysis

We carried out gene-wide analysis to combine the joint signal from multiple variants (coding variants and flanking UTRs) within a gene and to provide greater statistical power than that for single-marker tests. All the variants (nonsynonymous, synonymous, UTRs, singletons)
located within the studied genes and their exon-intron flanking regions were collapsed together and their combined effect was studied.

*NOTCH3* is the only significant hit in the c-alpha test (adj p-value = 0.01) (Table S8a). The signal is driven by a common coding synonymous variant (p.P1521P) of moderate effect size (OR = 1.755, CI = 1.31-2.33), significant after Bonferroni correction (adj p-value = 0.02) (Table S9). *TREX1* was another hit in the c-alpha test, although nominally significant (adj p-value = 0.56) and the signal is mainly driven by a 5'UTR and synonymous (p.Y232Y) variants (Table S8a and Table S9). None of these variants were predicted to affect the splicing site (http://www.umd.be/HSF/) or a miRNA binding site (http://www.microrna.org/microrna/home.do).

**Single-variant based analysis**

A total of 69 rare and low frequency coding missense mutations was considered in the single-variant based analysis in the studied genes. Among these, the majority (62.8%) were singletons (Table S10).

Moreover, 41 missense variants (59.4%) were described as damaging variants by at least 2 out of 3 *in silico* prediction softwares (SIFT, Polyphen and Mutation Taster).

The study possessed relatively low power to detect a significant association between cases and controls for low frequency and rare variants, however we analyzed these variants because we could not preclude the possibility that high effect risk alleles were present. *EIF2B4* and *CSF1R* harbour the lowest and highest relative frequency of low frequency and rare coding variants (mean = 1.27 and 5.13 low freq-rare variant per kb of coding sequence, respectively), with 81.25% of the rare and low frequency coding variability in *CSF1R* clustering in the Ig-like domain (Table S11-S10).

The main hits, although not significant, are rare variants with moderate to strong effect sizes (0.6<OR<2.73) clustering to *EIF2B4, NOTCH3, TREX1* and *CSF1R* (Table S10).

None of the missense mutations leads to a premature stop codon.

**Singletons in CSF1R TK domain**

We report 2 heterozygous missense mutations in the *CSF1R* TK domain (exons 12-22, aa 582-910) in the discovery cohort (p.L868R and p.E694K), detected in one case and one control,
respectively. Moreover, we found 4 likely pathogenic variants in the TK domain flanking regions (aa 538-581 and 911-972): CSF1R p.D565N, p.E916K, p.E920D, p.G957R. We then screened CSF1R in an independent follow-up cohort of late-onset AD and MCI patients and we identified 2 additional mutations in CSF1R TK domain (p.Q691H and p.H703Y) (Table 1).

CSF1R TK mutation carriers (Patient E, F and H) presented a rather homogeneous phenotype (Table 2). All these carriers were LOAD cases displaying memory impairment at onset. Behavioural and motor signs eventually appeared. In 2/3 patients cardiovascular problems and strokes preceded the dementia. The neuropathology examination, available for Patient H and I, showed aggressive and diffuse neurodegeneration (Braak 6 and Cerad C). Two out of three carriers were heterozygous for APOE ε4 allele and do not report any familial history for dementia. By contrast, Patient H was homozygous for APOE ε4 allele, had a family history for dementia (4 brothers) and plausibly the combination of these risk factors, likely coupled with a pre-existent cerebrovascular disorder, may explain the earlier age at onset compared to the other patients (64 years). Patient I carried a missense mutation in CSF1R TK flanking region (p. G957R) and displayed a different clinical picture, dominated by early-onset dementia (49y) and language problems at onset. Although the small sample size, we do not report any association between age at onset, severity of the disease progression and disease duration.

Detailed clinical description was available for 4 patients. The clinical, neuroimaging and neuropathological features of the carriers are summarized in Table 2.

**Patient H (c.2603T>G; L868R)**

This male patient was born in 1932 and deceased aged 75 years. He was one of five siblings who all survived to old age, of which four experienced memory problems or received a diagnosis of dementia or AD. The informant reported he experienced sudden decline following a stroke at 65 years. He had obvious short-term memory problems and dysphasia. At this stage, he was considered to have probable vascular dementia. Pathological examination of the brain concluded this patient had a high probability of AD. Neurofibrillary tangle stage was consistent with Braak stage VI while plaque pathology met CERAD criteria for score C. In addition, there was evidence of amyloid angiopathy, focal TDP-43 positivity and occasional glial inclusions.
**Patient E (p.Q691H)**

This patient was born in 1922 and deceased aged 89 years. She complained of memory problems aged 82 and 2 years later underwent a MRI scan, which showed symmetric patchy periventricular hyperintensities, mainly pronounced in the frontal lobe (Figure 4a). Following annual visits involving neuropsychiatric testing, she received a diagnosis of AD aged 86 years. In the three years following her diagnosis, her symptoms were quite stable. She experienced a rapid deterioration in the last seven-eight months before her death.

**Patient F (p.H703Y)**

This patient had a strong history of cardiovascular disease and reported memory symptoms, aged 79, followed by irritability and anxiety and 2 years later received a clinical diagnosis of probable AD. Computed tomography of the brain showed supratentorial atrophy, temporal lobe atrophy, and slight vascular changes. The patient also experienced intermittent motor symptoms which included mild rigidity, tremor and slowness of movement. The MRI scans showed central and cortical atrophy and mild to moderate medial temporal lobe atrophy as well as a small old haemorrhage, ischemic lesions and bilateral lesions localized to the centrum semiovale (Figure 4b).

**Patient I**’s detailed description is in the supplement.

**Tissue expression of CSF1R**

We followed-up our findings checking CSF1R expression in late onset AD and control brain samples. We selected entorhinal cortex (EC) and BA9 pre-association cortex (BA9) because the brain regions primarily affected by AD spreading pathology (Khan et al., 2014). CSF1R was overexpressed in AD entorhinal cortex compared to AD BA9 pre-association cortex and compared to control brains (Fig. 5a).

It was not possible to quantitatively compare levels of CSF1R in all the 3 CSF1R TK mutation carriers due to a lack of available brain tissue. However, cDNA Sanger Sequencing revealed a possible allelic imbalance, with the WT allele normally expressed and the mutated one only
moderately both for Patient F (p.H703Y) and Patient H (p.L868R), suggesting a functional role of these mutations (Fig. 4g-h). RNA from entorhinal cortex and the BA9 pre-association cortex was available for Patient H and showed significantly lower expression of CSF1R in entorhinal cortex compared to BA9 pre-association cortex and 2) in Patient H’s entorhinal cortex compared to other AD patients and controls for all CSF1R primers tested (Fig. 5a-b).

4. DISCUSSION

Mendelian adult-onset leukodystrophies clinically resemble common dementias such as AD, potentially implying they may be influenced by shared genetic risk factors.

To comprehensively investigate this hypothesis we applied a combination of gene expression analysis in different AD mouse strains at diverse developmental stages (http://mouseac.org/) and single-variant and single-gene based genetic screening in a cohort composed of 332 LOAD cases and 676 elderly controls from the UK and USA (Fig. 1). Divergent gene expression between AD and WT mouse strains was only detected in aged mice with severe core-dense plaque deposition (Fig. 2a-d, Table S3 and S4).

Csf1r was the only gene displaying a significant differential expression between AD and WT mouse strains. It was up to 2 folds significantly overexpressed both in the hippocampus and cortex in HOTASTPM mice 8 months of age (Log2FC=1.2 and 1.1; adj p-value= 2.5E-07 and 8.7E-05, respectively) and its overexpression linearly correlated with the rapidity of core-dense plaque deposition rather than with their overall load (Fig2a-d, Table S3 and S4). By contrast, Csf1r was downregulated when the pathology was minimal or absent (Table S3 and S4), suggesting that Csf1r upregulation was tightly driven by and consequential to dense-core plaque formation and, to a lesser extent, neurofibrillary tau tangles.

Importantly, Csf1r expression pattern relied on microglia infiltration as overexpression of Aif1 suggested (Fig. 2a-b, Table S3 and S4). Csf1r was co-expressed and shared the same expression pattern with Trem2, Tyrobp and Grn, critical genes expressed on microglia, whose loss of function mutation in homozygosity is causative for adult-onset leukodystrophies such as Polycystic lipomembranous osteodysplasia with sclerosing leukoencephalopathy (PLOSL) (TREM2 and TYROBP) (Paloneva et al., 1993), and in heterozygosity causes FTLD-TDP
or is a significant risk factor for sporadic AD (TREM2) (Guerreiro et al., 2013) and whose overexpression is protective and limits AD neuropathology through a very effective clearance of Aβ plaques (GRN, TREM2) (Minami et al., 2014). Notably, in all the strains, Csf1r and Grn degree of expression correlated and this was generally one third of Tyrobp and Trem2 overall upregulation (Fig. 3, Table S3 and S4). This effect was not simply due to the aging process: we do not report significant differential expression in Csf1r, Grn, Trem2 and Tyrobp in the cortex and hippocampus of WT mice between 2 and 18 months of age (-0.04<Log2FC<0.2 and 0.04<Log2FC<0.4 in hippocampus and cortex, respectively) (Table S3 and S4). Importantly, CSF1R, TREM2 and TYROBP have been already shown to co-interact (Otero et al., 2009) (https://string-db.org/). Here, we report GRN as an additional potential key player and their co-expression on microglia strengthens their synergic function. Therefore, this may imply that CSF1R, in concert with TREM2, TYROBP and GRN plays a key role in Aβ plaque removal, hypothesis supported by previous literature, reporting CSF1R overexpressed in AD patients particularly around senile plaques and taking part to Aβ removal (Akiyama et al., 1994) (Murphy et al., 2000) (Mizuno et al., 2011) (Boissonneault et al., 2009) (Boissonneault et al., 2009)

By contrast, no differential expression of any adult-onset leukodystrophy gene was observed in the E15 hippocampi and in primary cortical neurons after OGD experiments between APPPS1 and WT strains, likely given the fact that these genes, although expressed in neurons, mainly exert a critical role on microglia (TREX1, CSF1R), astrocytes (NOTCH3, HTRA1) and endothelial cells (NOTCH3, HTRA1, TREX1, EIF2B2, EIF2B3, EIF2B5)(http://web.stanford.edu/group/barres_lab/brain_rnaseq.html), that are minimally present in E15 hippocampi and primary cortical neuronal cultures.

In our discovery and validation cohorts we detected 3 rare coding variants in the TK domain of CSF1R in 3 late-onset AD cases, one of these neuropathologically confirmed. Moreover, we report 2 cases harbouring rare mutations in the TK flanking regions (aa 538-580 and aa 911-972, encoded by exon 12 and 22, respectively), where an additional causative mutation for HDLS has been recently described (c.1736G>A, p.R579Q) (Ghadiri et al., 2014). These variants are very likely pathogenic: cluster to highly conserved domain among homologous proteins and different species (average PhyloP and PhastCons scores= 2.2 and 0.6, respectively)(Fig. 4i-m)and have been detected only in cases.
Moreover, *CSF1R* p.L868R is a functional mutation as it was associated with a reduced expression of the mutated allele (Fig. 4h) and decreased *CSF1R* expression in entorhinal cortex, a region primarily affected in AD and generally displaying *CSF1R* upregulation in AD patients (Fig 5a and b). Importantly, a different amino acid change, clustering within the same codon (p.L868P) has been reported as *de novo* mutation, causative for HDLS (Rademakers et al., 2011). Remarkably, to date, any missense mutation in Mendelian gene domains harbouring heterozygous causative mutations for autosomal dominant disorders such as familial AD and FTD (*APP, PSEN1, PSEN2* and *MAPT*) has always been reported as pathogenic (Table S12, http://www.molgen.ua.ac.be/ADMutations/). The only exception is represented by *APP* p.A673T, which is a very rare protective factor for AD (Jonsson et al., 2012). In addition, an intronic SNP in *CSF1R*, rs1010101, displayed a trend towards association (adj p-val=2E-4), in a Genome-wide association study (GWAS) performed in Caucasian LOAD patients (Wijsman et al., 2011), supporting *CSF1R* possible role in LOAD progression.

*CSF1R* mutation carriers presented a homogeneous phenotype, closely resembling HDLS. First, the symptom at onset was a memory deterioration followed by behavioural changes in 3/3 carriers. Second, the T2-weighted MRI, available for 2 patients, showed symmetric patchy periventricular hyperintensities, mainly pronounced in the frontal lobe (Patient E) and bilateral lesions localized to the centrum semiovale (Patient F) (Fig. 4b), that represent common MRI findings in HDLS patients (Ghadiri et al., 2014), (Rademakers et al., 2011) (Boissé et al., 2010) (Mateen et al., 2010). Finally, senile plaques, amyloid angiopathy and tau tangles have been reported also in the cortex and hippocampus of 2 familial and one sporadic HDLS patients (Baba et al., 2006) (Browne et al., 2003). Importantly, these 3 HDLS patients displaying AD neuropathological hallmarks presented an overall early age at onset (average 54 years [range 78-32y]), developed Parkinsonism, atypical Parkinsonism and motor impairment with increasing rigidity (Table S13). By contrast, only Patient F (20%) displayed intermittent mild rigidity, tremor and bradykinesia, arguing for Parkinsonism, however within a neurological picture already dominated by cerebrovascular disorders, and Patient H reported no sign of motor impairment beside 2 falls in few months. Nevertheless, although an earlier age at onset, Parkinsonism and distinctive motor features may be more common in HDLS patients presenting AD neuropathology than AD cases carrying *CSF1R* TK mutations, the average disease duration for both these HDLS and AD patients was 7 years. Therefore, in the
absence of accurate differential diagnostic criteria, this combination of clinical, neuroimaging and neuropathological features strikingly overlapping makes the definitive neurological diagnosis a real conundrum. The fact that potentially pathogenic mutations in the TK domain in heterozygosity may be detected either in databases and apparently healthy controls or give rise to both HDLS or being rare risk factors for AD may be due to different factors modifying the mutation penetrance (Karle et al., 2013). Analogously to GRN missense mutations in AD and FTD, CSF1R mutation penetrance may be influenced by APOE genotype, aging, disease duration or comorbidities such as cerebrovascular accidents, for which CSF1R has been already shown to play a critical protective role (Luo et al., 2013). In addition, It may be plausible that the majority of HDLS patients may not display AD neuropathology due to the rapid progression of the disease (Baba et al., 2006).

Finally, NOTCH3 was a significant hit in the gene-based analysis (c-alpha test, adj p-val= 0.01). The signal was driven both by a common synonymous variant (p.P1521P)(Table S9), that may influence gene expression (Sauna and Kimchi-Sarfaty, 2011) and 3 rare coding variants with large effect size (p.V1952M, p.V1183M, p.H170R, 2.73<OR<1.63), whose carrier frequency was between 2 and 3 times higher in cases compared to controls, although not significant (Table S6 and S10). Importantly, these rare variants (p.V1952M, p.V1183M, p.H170R) have been already reported to be significantly associated with severity of white matter lesions in elderly with hypertension (Schmidt et al., 2011), suggesting a potential role as disease modifier in LOAD. In addition, we report a heterozygous pathogenic gain of cysteine mutation in NOTCH3 (p.R578C) detected in one control and already reported in a Korean patient with clinical suspicious CADASIL, implying that the penetrance of NOTCH3 mutations is variable (Kim et al., 2014).

Therefore, although canonical NOTCH3 mutations causative for CADASIL are highly stereotyped: 1) cluster in epidermal growth factor-like repeat domains (EGFR), 2) in exon 3 and 4 and 3) consist in the gain or loss of cysteine, nevertheless, our study reports a possible synergetic effect of common and rare variants in NOTCH3 potentially influencing AD susceptibility through an increased risk for small vessel disease or white matter lesions. Our hypotheses are supported by a growing body of evidence showing that 1) NOTCH3 common variants (rs1043994, rs10404382, rs10423702, rs1043997) are significantly associated with white matter lesion in elderly with hypertension (Schmidt et al., 2011) and 2) rare non-
cysteine mutations may be pathogenic as they have been reported in Korean and Japanese CADASIL patients, in a French case with small vessel disease and have been associated to severe white matter lesions in elderly patients (Mizuno et al., 2008) (Fouillade et al., 2008) (Schmidt et al., 2011). Importantly, our findings add evidence to the pathogenic link between AD and CADASIL, displaying clinical shared features and rarely, as only few cases have been reported, neuropathological hallmarks characterized by Aβ plaques, amyloid angiopathy and neurofibrillary tangles (Thijs et al., 2003) (Guerreiro et al., 2012) (Gray et al., 1994) (Paquet et al., 2010). Biologically, presenilins cleaving both APP and NOTCH3 may bridge the gap between AD and CADASIL. However, whether NOTCH3 mutations or differential expression may accelerate a pre-existing AD or AD may contribute to CADASIL exacerbation remains to be elucidated.

In summary, adult-onset Mendelian leukodystrophy genes are not common and critical factors in AD, therefore the genetic screening plays a pivotal role in the differential diagnosis. However, genetically diagnosed HDLS and CADASIL patients may display clinical, neuroimaging and neuropathological features meeting the diagnostic criteria for AD, leaving the definitive diagnosis a significant challenge. Here we report neuropathologically confirmed AD patients carrying likely pathogenic mutations in CSF1R TK domain and a potential association between AD and NOTCH3. Our study provides compelling evidence that HDLS, CADASIL and AD may represent shades of the same disease spectrum. Moreover, we support previous studies, suggesting that CSF1R, in concert with TREM2, TYROBP and GRN, may play a critical role in Aβ plaque clearance and therefore may represent a pivotal, although rare, genetic factor influencing AD susceptibility. Given the very rare frequency of CSF1R TK pathogenic mutations detected in the screened patients (0.3% LOAD carriers), our hypotheses should foster genetic screening in larger cohorts of both early-onset and late onset AD cases and functional studies.

Acknowledgements
This study was supported by the Alzheimer's Research UK, the Medical Research Council (MRC), the Wellcome Trust/MRC Joint Call in Neurodegeneration Award (WT089698) to the UK Parkinson’s Disease Consortium (whose members are from the University College London Institute of Neurology, the University of Sheffield, and the MRC Protein Phosphorylation Unit
at the University of Dundee), grants (P50 AG016574, U01 AG006786, and R01 AG18023), the National Institute for Health Research Biomedical Research Unit in Dementia at University College London Hospitals, University College London; the Big Lottery (to Dr. Morgan); Jose Bras and Rita Guerreiro’s work is funded by Fellowships from the Alzheimer’s Society; Humboldt Fellowship (to Celeste Sassi) and the Intramural Research Programs of the National Institute on Aging and the National Institute of Neurological Disease and Stroke, National Institutes of Health (Department of Health and Human Services Project number, Z01 AG000950-10). Mike A. Nalls’ participation is supported by a consulting contract between Data Tecnica international and the National Institute on Aging NIH, Bethesda, MD, USA. The MRC London Neurodegenerative Diseases Brain Bank and the Manchester Brain Bank from Brains for Dementia Research are jointly funded from ARUK and AS. The Alzheimer’s Research UK (ARUK) Consortium: Peter Passmore, David Craig, Janet Johnston, Bernadette McGuinness, Stephen Todd, Queen’s University Belfast, UK; Reinhard Heun, Royal Derby Hospital, UK; Heike Kölsch, University of Bonn, Germany; Patrick G. Kehoe, University of Bristol, UK; Nigel M. Hooper, University of Leeds, UK; Emma R.L.C. Vardy, University of Newcastle, UK; David M. Mann, Stuart Pickering-Brown, University of Manchester, UK; Kristelle Brown, James Lowe, Kevin Morgan, University of Nottingham, UK; A. David Smith, Gordon Wilcock, Donald Warden, University of Oxford (OPTIMA), UK; Clive Holmes, University of Southampton, UK.

Tissue samples were supplied by The London Neurodegenerative Diseases Brain Bank, which receives funding from the MRC and as part of the Brains for Dementia Research programme, jointly funded by Alzheimer’s Research UK and Alzheimer’s Society.

CONFLICT OF INTEREST STATEMENT

All the authors declare no competing financial or personal interests that can influence the presented work. However, Mike A. Nalls’ participation is supported by a consulting contract between Data Tecnica international and the National Institute on Aging NIH, Bethesda, MD, USA, as a possible conflict of interest, Dr. Nalls consult also Illumina Inc, the Michael J. Fox Foundation and the University of California Healthcare among others.

References


