





Article

# The Impact of Complement Genes on the Risk of Late-Onset Alzheimer's Disease

Sarah M. Carpanini <sup>1,2,†</sup> , Janet C. Harwood <sup>3,†</sup> , Emily Baker <sup>1</sup>, Megan Torvell <sup>1,2</sup>, The GERAD1 Consortium <sup>‡</sup>, Rebecca Sims <sup>3</sup> , Julie Williams <sup>1</sup> and B. Paul Morgan <sup>1,2,\*</sup> 

<sup>1</sup> UK Dementia Research Institute at Cardiff University, School of Medicine, Cardiff, CF24 4HQ, UK; CarpaniniS@Cardiff.ac.uk (S.M.C.); BakerEA@cardiff.ac.uk (E.B.); TorvellM@cardiff.ac.uk (M.T.); WilliamsJ@cardiff.ac.uk (J.W.)

<sup>2</sup> Division of Infection and Immunity, School of Medicine, Systems Immunity Research Institute, Cardiff University, Cardiff, CF14 4XN, UK

<sup>3</sup> Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, CF24 4HQ, UK; HarwoodJC@cardiff.ac.uk (J.C.H.); SimsRC@cardiff.ac.uk (R.S.)

\* Correspondence: Morganbp@cardiff.ac.uk

† These authors contributed equally to this work.

‡ Data used in the preparation of this article were obtained from the Genetic and Environmental Risk for Alzheimer's disease (GERAD1) Consortium. As such, the investigators within the GERAD1 consortia contributed to the design and implementation of GERAD1 and/or provided data but did not participate in analysis or writing of this report. A full list of GERAD1 investigators and their affiliations is included in Supplementary File S1.



**Citation:** Carpanini, S.M.; Harwood, J.C.; Baker, E.; Torvell, M.; The GERAD1 Consortium; Sims, R.; Williams, J.; Morgan, B.P. The Impact of Complement Genes on the Risk of Late-Onset Alzheimer's Disease. *Genes* **2021**, *12*, 443. <https://doi.org/10.3390/genes12030443>

Academic Editors: Laura Ibanez and Justin Miller

Received: 27 February 2021

Accepted: 16 March 2021

Published: 20 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Late-onset Alzheimer's disease (LOAD), the most common cause of dementia, and a huge global health challenge, is a neurodegenerative disease of uncertain aetiology. To deliver effective diagnostics and therapeutics, understanding the molecular basis of the disease is essential. Contemporary large genome-wide association studies (GWAS) have identified over seventy novel genetic susceptibility loci for LOAD. Most are implicated in microglial or inflammatory pathways, bringing inflammation to the fore as a candidate pathological pathway. Among the most significant GWAS hits are three complement genes: *CLU*, encoding the fluid-phase complement inhibitor clusterin; *CR1* encoding complement receptor 1 (CR1); and recently, *C1S* encoding the complement enzyme C1s. Complement activation is a critical driver of inflammation; changes in complement genes may impact risk by altering the inflammatory status in the brain. To assess complement gene association with LOAD risk, we manually created a comprehensive complement gene list and tested these in gene-set analysis with LOAD summary statistics. We confirmed associations of *CLU* and *CR1* genes with LOAD but showed no significant associations for the complement gene-set when excluding *CLU* and *CR1*. No significant association with other complement genes, including *C1S*, was seen in the IGAP dataset; however, these may emerge from larger datasets.

**Keywords:** complement; complement receptor 1; clusterin; late-onset Alzheimer's disease; genetics; neuroinflammation

## 1. Introduction

Alzheimer's disease (AD) is the most common cause of dementia in the elderly. Pathologically, AD is a chronic neurodegenerative disease underpinned by neuronal and synaptic loss, the accumulation of amyloid- $\beta$  plaques, and neurofibrillary tangles composed of hyperphosphorylated tau. An important role for neuroinflammation has emerged in recent years. Evidence includes the presence of activated microglia in the brain innate immune cells, the presence of inflammatory markers, including complement proteins, in the brain, cerebrospinal fluid (CSF) and plasma, and the demonstration that chronic use of anti-inflammatory drugs may reduce disease incidence [1–3]. Perhaps the best evidence that inflammation may be involved in AD aetiology comes from genome-wide association

studies (GWAS); many of the genes most strongly associated with AD risk are involved in inflammation and immunity.

The first causative mutations for AD, identified over 25 years ago in the rare early-onset familial forms of AD, were in Amyloid precursor protein (*APP*), Presenilin 1 (*PSEN1*) and Presenilin 2 (*PSEN2*) genes [4–6]. *APP*, encoded by the *APP* gene, a broadly expressed transmembrane protein abundant in the brain, is sequentially cleaved by secretase enzymes. The precise cleavage patterns determine its propensity to seed A $\beta$  plaques. The presenilin proteins *PSEN1* and *PSEN2* are both components of the  $\gamma$ -secretase complex and important in the function of this enzyme; mutations in the genes encoding these proteins impact the *APP* cleavage pathway. The identification of early-onset AD-associated mutations in these three genes underpins the amyloid cascade hypothesis whereby abnormal *APP* processing leading to A $\beta$  plaque formation is considered the key underlying pathology associated with AD [7]. However, it is important to stress that these mutations are only relevant to early-onset familial AD which accounts for fewer than 1% of all AD cases. In late-onset Alzheimer's disease (LOAD), accounting for the large majority of AD cases, the strongest genetic risk factor is the presence of the  $\epsilon$ 4 allele of the gene encoding Apolipoprotein E (*ApoE*);  $\epsilon$ 4 confers increased risk, while the most common allele,  $\epsilon$ 3, is considered neutral for AD, and  $\epsilon$ 2 has a minor protective effect [8–11]. Homozygosity for *APOE*  $\epsilon$ 4 confers an ~11-fold increased risk of LOAD compared to  $\epsilon$ 3 homozygotes. Precisely how these variants in *APOE* impact disease risk remains a subject of ongoing research. *ApoE* is a lipoprotein present in biological fluids; therefore, roles in lipid transport and membrane repair in the brain have been proposed [12].

Over the past decade, large GWAS have identified variants in more than 70 genetic loci that are associated with LOAD, implicating multiple and diverse biological pathways [13–16]. Notably, ~20% of the genes in LOAD risk loci encode proteins with roles in inflammation and immunity [14,17,18]; many of these are predominantly expressed in microglia, notably *TREM2*, *ABI3* and *PLCG2* [15,19]. From GWAS, it has been shown that three complement system genes are significantly associated with LOAD: *CLU*, *CR1*, and recently, *C1S* encoding the classical pathway enzyme C1s was added to this list [13,16,20]. *CLU* encodes clusterin, a multifunctional plasma protein that regulates the complement terminal pathway, and *CR1* encodes complement receptor 1 (CR1), a receptor for complement fragments and regulator of activation. These are both regulators of the complement cascade and provide the impetus for this analysis of complement genetics in LOAD. To test whether complement genes beyond *CLU* and *CR1* (both genome-wide significant (GWS) in the International Genomics of Alzheimer's Project (IGAP) dataset) influence the risk of LOAD, we compiled a comprehensive complement gene-set containing only those genes that encoded proteins directly involved in complement activation, regulation, or recognition. Then, we undertook several methods of pathway analysis to test whether additional genes within the complement gene-set were associated with LOAD risk.

## 2. Materials and Methods

### 2.1. Complement Genes and Gene Exclusion Analyses in LOAD

In order to understand the genetics of the complement pathway in AD, we compiled a comprehensive gene-set comprising all complement genes and associated regulators and receptors. Genes were selected for inclusion based upon known biological relevance to the complement system rather than by using often inaccurate annotations in public databases. The resultant complement gene-set contained 56 genes, subdivided into their relevant functional groups (Table 1).

**Table 1.** Complement gene list including all complement genes and associated regulators and receptors. Genes are subdivided according to pathway; either classical, lectin, amplification loop or terminal and whether they are complement genes or associated regulators/receptors.

Pathway	HGNC Gene Name	Entrez Gene ID	HGNC Full Gene Name
Classical	<i>C1QA</i>	712	complement C1q A chain
Classical	<i>C1QB</i>	713	complement C1q B chain
Classical	<i>C1QC</i>	714	complement C1q C chain
Classical	<i>C1R</i>	715	complement C1r
Classical	<i>C1S</i>	716	complement C1s
Classical/Lectin	<i>C2</i>	717	complement C2
Classical/Lectin	<i>C4A</i>	720	complement C4A (Rodgers blood group)
Classical/Lectin	<i>C4B</i>	721	complement C4B (Chido blood group)
Lectin	<i>FCN1</i>	2219	ficolin 1
Lectin	<i>FCN2</i>	2220	ficolin 2
Lectin	<i>FCN3</i>	8547	ficolin 3
Lectin	<i>MASP1</i>	5648	mannan binding lectin serine peptidase 1
Lectin	<i>MASP2</i>	10747	mannan binding lectin serine peptidase 2
Lectin	<i>MBL2</i>	4153	mannose binding lectin 2
Amplification loop	<i>CFB</i>	629	complement factor B
Amplification loop	<i>CFD</i>	1675	complement factor D
Classical/Lectin/ Amplification loop	<i>C3</i>	718	complement C3
Terminal	<i>C5</i>	727	complement C5
Terminal	<i>C6</i>	729	complement C6
Terminal	<i>C7</i>	730	complement C7
Terminal	<i>C8A</i>	731	complement C8 $\alpha$ chain
Terminal	<i>C8B</i>	732	complement C8 $\beta$ chain
Terminal	<i>C8G</i>	733	complement C8 $\gamma$ chain
Terminal	<i>C9</i>	735	complement C9
Regulator/Receptor	<i>C1QBP</i>	708	complement C1q binding protein
Regulator/Receptor	<i>C3AR1</i>	719	complement C3a receptor 1
Regulator/Receptor	<i>C4BPA</i>	722	complement component 4 binding protein $\alpha$
Regulator/Receptor	<i>C4BPB</i>	725	complement component 4 binding protein $\beta$
Regulator/Receptor	<i>C5AR1</i>	728	complement C5a receptor 1
Regulator/Receptor	<i>C5AR2</i>	27202	complement component 5a receptor 2
Regulator/Receptor	<i>CD46</i>	4179	CD46 molecule
Regulator/Receptor	<i>CD55</i>	1604	CD55 molecule (Cromer blood group )
Regulator/Receptor	<i>CD59</i>	966	CD59 molecule
Regulator/Receptor	<i>CFH</i>	3075	complement factor H
Regulator/Receptor	<i>CFHR1</i>	3078	complement factor H related 1
Regulator/Receptor	<i>CFHR2</i>	3080	complement factor H related 2
Regulator/Receptor	<i>CFHR3</i>	10878	complement factor H related 3
Regulator/Receptor	<i>CFHR4</i>	10877	complement factor H related 4
Regulator/Receptor	<i>CFHR5</i>	81494	complement factor H related 5
Regulator/Receptor	<i>CFI</i>	3426	complement factor I
Regulator/Receptor	<i>CFP</i>	5199	complement factor properdin
Regulator/Receptor	<i>CLU</i>	1191	clusterin
Regulator/Receptor	<i>CR1</i>	1378	complement C3b/C4b receptor 1 (Knops blood group)
Regulator/Receptor	<i>CR2</i>	1380	complement C3d receptor 2
Regulator/Receptor	<i>CSMD1</i>	64478	CUB and Sushi multiple domains 1
Regulator/Receptor	<i>ITGAM</i>	3684	integrin subunit $\alpha$ M
Regulator/Receptor	<i>ITGAX</i>	3687	integrin subunit $\alpha$ X
Regulator/Receptor	<i>SERPING1</i>	710	serpin family G member 1
Regulator/Receptor	<i>VTN</i>	7448	Vitronectin
Regulator/Receptor	<i>CD93</i>	22918	C1q receptor phagocytosis
Complement-like	<i>C1QL1</i>	10882	complement C1q-like 1
Complement-like	<i>C1QL2</i>	165257	complement C1q-like 2
Complement-like	<i>C1QL3</i>	389941	complement C1q-like 3
Complement-like	<i>C1QL4</i>	338761	complement C1q-like 4
Complement-like	<i>C1RL</i>	51279	complement C1r subcomponent-like
Complement-like	<i>CR1L</i>	1379	complement C3b/C4b receptor 1-like

## 2.2. AD Summary Statistics

This study utilised summary statistics from the International Genomics of Alzheimer's Project (IGAP). IGAP is a large three-stage study based upon GWAS on individuals of European ancestry. In stage 1, IGAP used genotyped and imputed data on 11,480,632 single nucleotide polymorphisms (SNPs) to meta-analyse GWAS datasets consisting of 21,982 Alzheimer's disease cases and 41,944 cognitively normal controls from four consortia: the Alzheimer Disease Genetics Consortium (ADGC); the European Alzheimer's disease Initiative (EADI); the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium (CHARGE); and the Genetic and Environmental Risk in AD Consortium Genetic and Environmental Risk in AD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium (GERAD/PERADES). In stage 2, 11,632 SNPs were genotyped and tested for association in an independent set of 8362 Alzheimer's disease cases and 10,483 controls. Meta-analyses of variants selected for analysis in stage 3A ( $n = 11,666$ ) or stage 3B ( $n = 30,511$ ) samples brought the final sample to 35,274 clinical and autopsy-documented Alzheimer's disease cases and 59,163 controls.

Gene-set analysis was performed using the complement gene-set and stage 1 summary statistics from the International Genomics of Alzheimer's Project [14]. The individual and combined effects of the genome-wide significant (GWS) genes *CLU* and *CR1* within the complement gene-set were investigated by removing these genes individually and together. We utilised the most up-to-date publicly available GWAS dataset at the time of writing [14], and calculated the complement gene-set  $p$ -values when including and excluding those loci that reached genome-wide significance in the IGAP dataset. The recently identified LOAD-associated *C15* variant [13] does not show genome-wide statistical significance in the IGAP dataset; and therefore was not removed in the gene-set analysis. Complement gene-sets were tested for enrichment using the IGAP stage 1 summary statistics [14] in MAGMA version 1.06 [21]. Summary statistics were filtered for common variants ( $MAF \geq 0.01$ ) and all indels and merged deletions were removed; 8,608,484 SNPs were analysed. Genes were annotated using reference data files from the European population of Phase 3 of 1000 Genomes, human genome Build 37 using a window of 35 kb upstream and 10 kb downstream of each gene [22]. Ten thousand permutations were used to estimate  $p$ -values, corrected for multiple testing using the family-wise error rate (FWER). Gene-sets with a FWER-corrected  $p$ -value  $< 0.05$  under the "mean" model for estimating gene-level associations were reported as significant.

## 2.3. Complement Risk Score Analysis

A complement risk score combining the effects of all SNPs in the complement gene-set was produced. POLARIS [23] was used to compute risk scores in GERAD-genotyped data (3332 cases, 9832 controls) using SNP effect sizes from IGAP stage 1 summary statistics [14,16,20] (excluding GERAD subjects). Linkage disequilibrium (LD) was estimated from the GERAD data, and POLARIS was used to adjust the scores for LD between SNPs. The overall association of the complement gene-set with LOAD was determined using a logistic regression model, adjusting for population covariates, age, and sex. The logistic regression model included the baseline polygenic risk scores for all SNPs in the model, thereby testing for any association beyond the baseline polygenic effect.

Data used in the preparation of this article were obtained from the Genetic and Environmental Risk for Alzheimer's disease (GERAD) Consortium. The imputed GERAD sample comprised 3177 AD cases and 7277 controls with available age and gender data. Cases and elderly screened controls were recruited by the Medical Research Council (MRC) Genetic Resource for AD (Cardiff University; Institute of Psychiatry, London; Cambridge University; Trinity College Dublin), the Alzheimer's Research UK (ARUK) Collaboration (University of Nottingham; University of Manchester; University of Southampton; University of Bristol; Queen's University Belfast; the Oxford Project to Investigate Memory and Ageing (OPTIMA), Oxford University); Washington University, St Louis, United States; MRC PRION Unit, University College London; London and the South East Region

AD project (LASER-AD), University College London; Competence Network of Dementia (CND) and Department of Psychiatry, University of Bonn, Germany; the National Institute of Mental Health (NIMH) AD Genetics Initiative. A total of 6129 population controls were drawn from large existing cohorts with available GWAS data, including the 1958 British Birth Cohort (1958BC) (<http://www.b58cgene.sgu.ac.uk>, accessed on 15 March 2021), the KORA F4 Study, and the Heinz Nixdorf Recall Study. All AD cases met criteria for either probable (NINCDS-ADRDA, DSM-IV) or definite (CERAD) AD. All elderly controls were screened for dementia using the MMSE or ADAS-cog and were determined to be free from dementia at neuropathological examination or had a Braak score of 2.5 or lower. Genotypes from all cases and 4617 controls were previously included in the AD GWAS by Harold and colleagues (2009) [20]. Genotypes for the remaining population controls were obtained from WTCCC2. Imputation of the dataset was performed using IMPUTE2 and the 1000 genomes (<http://www.1000genomes.org/>, accessed on 15 March 2021) Dec2010 reference panel (NCBI build 37.1).

#### 2.4. Likelihood Ratio Analysis

A likelihood ratio test was used to estimate how much of the complement gene-set effect on LOAD risk was contributed by *CLU* and *CR1*, and to test whether there were residual polygenic effects of the remaining genes from the complement gene-set. The effects of *CLU* and *CR1* were estimated using a risk score combining all SNPs in the gene, produced using POLARIS in order to correct for LD. Likelihood ratio tests were used to compare individual models containing SNPs in *CLU* and *CR1* and models containing the combined risk conferred by SNPs in the rest of the complement gene-set.

### 3. Results

#### 3.1. MAGMA Analysis Reveals the Impact of Individual Complement Genes

From the MAGMA gene-set analysis, the complement gene-set comprising all 56 genes was significantly associated with LOAD ( $p = 0.011$ ) (Table 2). When the GWAS-significant genes *CLU* and *CR1* were excluded individually from the gene-set, the complement-minus-*CLU* gene-set was not significant ( $p = 0.057$ ), while the complement-minus-*CR1* gene-set was significant ( $p = 0.048$ ). As *CR1* and *CR1L* are located next to each other on chromosome 1, and linkage disequilibrium extends between the two genes, we excluded the *CR1/CR1L* locus from the gene-set. This gene-set was not significant ( $p = 0.082$ ). The gene set in which both *CLU* and *CR1* were excluded from the complement gene-set was not significantly associated with LOAD ( $p = 0.170$ ). The signal in the gene-set where *CR1L*, *CLU* and *CR1* were excluded was reduced compared with the signal derived from the gene-sets in which *CLU* and *CR1* were removed (Table 2). Taken together, these results suggest that the LOAD association signal in the complement gene-set is predominantly driven by *CLU* and *CR1*. Given the physical distance between *CR1* and *CR1L*, the use of extended gene boundaries and that linkage disequilibrium extends across both genes, we cannot resolve the signal between these two genes in the gene set analysis. Hence, we cannot confirm any independent contribution from *CR1L*.

**Table 2.** Complement gene-set analysis.

Gene-Set	Ngenes	OR	95% CI	$p$	$p$ FWER
Complement Genes	56	1.402	[1.068, 1.841]	0.008	0.011
Complement Genes Minus <i>CLU</i>	55	1.278	[0.969, 1.684]	0.041	0.057
Complement Genes Minus <i>CR1</i>	55	1.288	[0.981, 1.691]	0.034	0.048
Complement Genes Minus <i>CLU</i> , <i>CR1</i>	54	1.172	[0.891, 1.542]	0.129	0.170
Complement Genes Minus <i>CR1</i> , <i>CR1L</i>	54	1.244	[0.943, 1.639]	0.061	0.082
Complement Genes Minus <i>CLU</i> , <i>CR1</i> , <i>CR1L</i>	53	1.127	[0.854, 1.489]	0.199	0.246

Table 2 displays the results from the MAGMA analysis. Gene-sets were corrected for multiple testing using the family-wise error rate (FWER). The complement gene-set

is significant ( $p = 0.011$ ), but this effect is lost when *CLU* and *CR1* are excluded from the gene-set ( $p = 0.170$ ). *CLU* has the largest impact in the complement set, and the association with AD is predominantly driven by *CLU* and *CR1*.

### 3.2. Risk Score Analysis Supports the Impact of Complement Genes

To further explore the impact of complement genes on LOAD risk, we adopted a polygenic approach. We first applied risk score analysis to the dataset, then used logistic regression to explore the association between LOAD and complement gene-set risk scores in GERAD individuals (Table 3). The complement gene-set as a whole was strongly associated with AD in this analysis ( $p = 0.003$ ). Removal of *CLU* from the gene-set caused the largest reduction in significance ( $p = 0.003$  vs.  $p = 0.053$ ). Removal of *CR1*, or the *CR1/CR1L* locus had minimal impact on the significance of association in the gene-set, although when *CLU*, *CR1* and *CR1L* were eliminated, the significance was further reduced compared to the elimination of *CLU* alone ( $p = 0.148$  vs.  $p = 0.053$ ) (Table 3). These gene elimination analyses demonstrated that *CLU* and *CR1* were the major contributors to the risk of LOAD in the complement gene-set; however, the polygenic approach revealed that *CLU* was by far the more significant of these. In these data, the *CLU* gene shows a stronger association compared to *CR1* ( $p = 1.03 \times 10^{-5}$  and  $p = 1.5 \times 10^{-3}$ , respectively). The joint association of *CLU* and *CR1* is stronger still ( $p = 3.88 \times 10^{-7}$ ), showing that *CLU* and *CR1* are both independently associated with AD.

**Table 3.** Association between Alzheimer’s disease (AD) and complement gene-set risk score.

Gene-Set	Ngenes	OR	95% CI	<i>p</i>
Complement Genes	56	1.090	[1.028, 1.156]	0.003
Complement Genes Minus <i>CLU</i>	55	1.059	[0.998, 1.123]	0.053
Complement Genes Minus <i>CR1</i>	55	1.089	[1.027, 1.155]	0.004
Complement Genes Minus <i>CLU</i> , <i>CR1</i>	54	1.058	[0.997, 1.122]	0.059
Complement Genes Minus <i>CR1</i> , <i>CR1L</i>	54	1.077	[1.015, 1.142]	0.013
Complement Genes Minus <i>CLU</i> , <i>CR1</i> , <i>CR1L</i>	53	1.044	[0.984, 1.107]	0.148

Table 3 displays the results from the risk score analysis; the overall complement risk score shows an association with AD ( $p = 0.003$ ). *CLU* explains the majority of this signal.

### 3.3. Likelihood Ratio Analysis Confirms No Significant Impact of Other Complement Genes

We next tested complement gene-set effects using likelihood ratio analyses. Models in which *CLU*, *CR1* and *CR1/CR1L* were removed individually, showed significant residual impact in the gene-set ( $p = 0.0136$ ;  $p = 0.0091$ ;  $p = 0.0063$  respectively); after removal of *CR1* and *CLU* or *CR1*, *CLU* and *CR1L*, there was no significant residual impact in the gene-set, demonstrating that there was no significant polygenic effect of the remaining complement genes in the datasets used (Table 4). These results further support the conclusion that the complement gene-set association with LOAD is driven predominantly by *CLU* and *CR1*, but with no significant contribution from other complement gene-set members ( $p = 0.1457$ ; Table 4).

Table 4 shows the results from these likelihood ratio tests comparing models containing SNPs in *CLU*, *CR1* and *CR1/CR1L* only and models containing the combined risk in SNPs in the remaining complement genes. The *p*-values demonstrate whether the remaining genes in the complement explain any additional variation. These results further support the conclusion that the complement gene-set impact on LOAD risk is predominantly driven by *CLU* and *CR1*.

**Table 4.** Likelihood ratio test (LRT) comparing gene-set risk scores.

Models Compared	LRT <i>p</i> -Value
(1) <i>CLU</i> (2) <i>CLU</i> + Complement_minus_ <i>CLU</i>	0.0136
(1) <i>CR1</i> (2) <i>CR1</i> + Complement_minus_ <i>CR1</i>	0.0091
(1) <i>CLU</i> + <i>CR1</i> (2) <i>CLU</i> + <i>CR1</i> + Complement_minus_ <i>CLU</i> _ <i>CR1</i>	0.1457
(1) <i>CR1</i> + <i>CR1L</i> (2) <i>CR1</i> + <i>CR1L</i> + Complement_minus_ <i>CR1</i> _ <i>CR1L</i>	0.0063
(1) <i>CLU</i> + <i>CR1</i> + <i>CR1L</i> (2) <i>CLU</i> + <i>CR1</i> + <i>CR1L</i> + Complement_minus_ <i>CLU</i> _ <i>CR1</i> _ <i>CR1L</i>	0.1145

#### 4. Discussion

The first evidence implicating the complement system in LOAD came from immunostaining of post-mortem brain tissue. Complement components and activation products, notably C1q, C4b, C3b/iC3b and the membrane attack complex, were present and co-localised with amyloid plaques and neurofibrillary tangles in the AD brain [24–27]. C3 fragments were shown to opsonise amyloid for phagocytosis by microglia in the brain and facilitate transport on erythrocytes to the liver [28]. Complement activation is critically involved in synaptic pruning both in development and in diseases such as AD [29–32]. In AD mouse models, back-crossing to complement deficiencies has supported the critical role of complements in neuroinflammation and synapse loss [30,33]. The presence of complement activation biomarkers in CSF and/or plasma in LOAD suggested that complement dysregulation occurs early in the disease [2]. The demonstration that complement genes associated with LOAD provided compelling evidence that the complement was a driver of disease rather than a secondary event [13,14,16].

To further investigate the roles of complement genes in the risk of LOAD, we compiled a comprehensive complement gene-set and used a polygenic approach to identify genes contributing to AD risk. We have demonstrated that the signal for the association of the complement gene-set with LOAD is explained by the GWS genes *CLU* and *CR1*, and not by other complement genes tested here. This finding was unexpected. Based on knowledge from other chronic inflammatory diseases, we had hypothesised that many complement genes might influence LOAD risk. For example, in age-related macular degeneration (AMD), a retinal disease clinically and pathologically linked to LOAD, genes encoding complement components C2, C3, FB and C9, and regulators FH and FHR4, all contribute to risk [34]. Indeed, the demonstration that multiple complement genes can collaborate to cause dysregulation and disease informed the concept of the “complotype”, the set of complement gene variants inherited by an individual that dictates complement activity and disease risk [35]. The genetic associations in these other chronic inflammatory diseases influence systemic or local complement regulation and/or amplification of activation; these in turn cause complement dysregulation that drives inflammation. Our demonstration that the complement genetic signature in LOAD is restricted to the genes encoding clusterin and CR1 suggests that complement dysregulation is not critical in the disease process. However, it should be noted that this finding is dependent on the dataset being investigated. At the time of writing, we utilised the largest publicly available AD GWAS dataset [14]. A recent study by the European AD Biobank, currently a preprint, reported an LOAD GWAS-significant association with the complement gene *C1S* [13]; this suggests that larger datasets and different analytical methods may implicate other complement genes and further elucidate roles of the complement system in LOAD. Additionally, because of the highly repetitive nature of a number of the complement loci, for example, the regulators of complement activation (RCA) clusters on chromosome 1 [36], many complement genes may

be hidden from standard sequencing technologies; the application of emerging long-range sequencing methods may reveal additional genetic variation in complement genes linked to LOAD missed in current GWAS and whole exome/genome sequencing studies using short read sequencing technologies [37].

Of the complement genes tested here, *CLU* and *CR1* were significantly associated with LOAD through multiple analytical approaches. Clusterin is a multi-functional plasma protein; its role in the complement system is to restrict fluid-phase membrane attack pathway activation [38]; however, beyond the complement system, clusterin functions as an extracellular chaperone protein, is involved in oxidative stress and cell survival/cell death pathways, and functions as an apolipoprotein in lipid transport [38–41]. Any one or several of these functions might underpin the association with LOAD. Four SNPs in *CLU*, all intronic and in LD, have been associated with increased LOAD risk (rs11136000, rs2279590, rs9331888 and rs9331896) [16,20]; evidence to date suggests that these SNPs impact clusterin synthesis, and hence, plasma clusterin levels. *CR1* is a membrane-bound receptor for complement components (C1q, MBL) and fragments (C3b, C4b). The primary function of *CR1* is as a receptor for C3b/C4b-opsonised immune complexes. *CR1* on erythrocytes sequesters immune complexes and transports them to disposal sites, while *CR1* on phagocytic cells binds opsonised immune complexes and processes them for elimination via phagocytosis. This latter activity requires a second function of *CR1*, its cofactor activity for factor I cleavage of C3b to iC3b the ligand for the phagocytic receptor *CR3*. The biological relevance of the C1q/MBL binding functions of *CR1* are unclear. The human *CR1* gene is located in the *RCA* gene cluster on chromosome 1 (1q32); duplications and deletions in this highly repetitive gene generate multiple isoforms via copy number variation (CNV). The most common variant, *CR1*\*1 (allele frequency 0.87) comprises 30 tandem repeats of 60–70 amino acid units called short consensus repeats (SCRs), which are in turn grouped in four homologous sets of seven termed long homologous repeats (LHRs), each a separate C3b/C4b binding unit. The second most common variant *CR1*\*2 (allele frequency 0.11) is identical to *CR1*\*1 except for the acquisition of an additional LHR, a “gain-of-function”; this variant increases risk for LOAD by up to 30%, although precisely how is unclear [14,16,42–44]. It has been suggested that the *CR1*\*2 variant is associated with lower *CR1* expression on erythrocytes, reducing the efficiency of peripheral immune complex handling and impacting amyloid clearance from the brain [45,46].

Our original analysis suggested that some of the signal from the complement gene-set might be attributable to the *CR1L* gene. However, *CR1L* is immediately adjacent to *CR1* and the SNP signals cannot be resolved, so it is not possible to ascribe an independent signal to *CR1L* in this analysis. *CR1L* encodes a C4b-binding protein comprising 13 SCRs, expressed predominantly in haematopoietic tissues [47,48]. Its physiological role is unknown, and evidence mechanistically linking it to LOAD is absent.

## 5. Conclusions

Taken together, our findings confirm the strong genetic association of the complement genes *CR1* and *CLU* with LOAD and that there is no statistically significant association signal for other complement genes apparent in the dataset used for the analysis. *CR1* and clusterin are important regulators of the complement pathway, suggesting that its dysregulation is important in LOAD. The recent GWAS association of *C1S* with LOAD demonstrates the potential for missing associations in this complex gene-set and raises the possibility that other loci may be missed by current large-scale genotyping and short-read sequencing technologies. Application of long read sequencing technologies could significantly alter the current landscape of complement system genetics in relation to LOAD risk.



**Supplementary Materials:** The following are available online at <https://www.mdpi.com/2073-4425/12/3/443/s1>, Supplementary File 1: List of GERAD authors and affiliations.

**Author Contributions:** Conceptualization, B.P.M.; data curation, S.M.C., J.C.H., E.B., M.T. and The GERAD1 Consortium; formal analysis, J.C.H. and E.B.; funding acquisition, J.W. and B.P.M.; investigation, J.C.H. and E.B.; methodology, J.C.H. and E.B.; software, J.C.H. and E.B.; supervision, B.P.M.; writing—original draft preparation, S.M.C., J.C.H., E.B., M.T. and B.P.M.; writing—review and editing, S.M.C., J.C.H., E.B., M.T., R.S., J.W. and B.P.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the U.K. Dementia Research Institute which receives its funding from U.K. DRI Ltd., funded by the U.K. Medical Research Council, Alzheimer’s Society, and Alzheimer’s Research U.K.

**Data Availability Statement:** The GERAD data is available by request (GERADConsortium@cf.ac.uk) and Kunkle data is available to download (<https://www.niagads.org/datasets/ng00075> accessed on 15 March 2021).

**Acknowledgments:** Data used in the preparation of this article were obtained from the Genetic and Environmental Risk for Alzheimer’s disease (GERAD1) Consortium. Cardiff University was supported by the Wellcome Trust, Medical Research Council (MRC), Alzheimer’s Research U.K. (ARUK) and the Welsh Assembly Government. Cambridge University and Kings College London acknowledge support from the MRC. ARUK supported sample collections at the South West Dementia Bank and the Universities of Nottingham, Manchester, and Belfast. The Belfast group acknowledges support from the Alzheimer’s Society, Ulster Garden Villages, Northern Ireland R&D Office, and the Royal College of Physicians/Dunhill Medical Trust. The MRC and Mercer’s Institute for Research on Ageing supported the Trinity College group. The South West Dementia Brain Bank acknowledges support from Bristol Research into Alzheimer’s and Care of the Elderly. The Charles Wolfson Charitable Trust supported the OPTIMA group. Washington University was funded by NIH grants, Barnes Jewish Foundation and the Charles and Joanne Knight Alzheimer’s Research Initiative. Patient recruitment for the MRC Prion Unit/UCL Department of Neurodegenerative Disease collection was supported by the UCLH/UCL Biomedical Centre and NIHR Queen Square Dementia Biomedical Research Unit. LASER-AD was funded by Lundbeck SA. The Bonn group was supported by the German Federal Ministry of Education and Research (BMBF), Competence Network Dementia and Competence Network Degenerative Dementia, and by the Alfred Krupp von Bohlen und Halbach-Stiftung. The GERAD1 Consortium also used samples ascertained by the NIMH AD Genetics Initiative. The GERAD data used in this paper included 5770 additional population controls; including the 1958 British Birth Cohort (1958BC), the KORA F4 Study, Heinz Nixdorf Recall Study and controls from the National Blood Service genotyped as part of the Wellcome Trust Case Control Consortium. The KORA F4 studies were financed by Helmholtz Zentrum München; German Research Center for Environmental Health; BMBF; German National Genome Research Network and the Munich Center of Health Sciences. The Heinz Nixdorf Recall cohort was funded by the Heinz Nixdorf Foundation (jur. G. Schmidt, Chairman) and BMBF. Coriell Cell Repositories is supported by NINDS and the Intramural Research Program of the National Institute on Aging. We acknowledge the use of genotype data from the 1958 Birth Cohort collection, funded by the MRC and the Wellcome Trust, which was genotyped by the Wellcome Trust Case Control Consortium and the Type-1 Diabetes Genetics Consortium, sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Allergy and Infectious Diseases, National Human Genome Research Institute, National Institute of Child Health and Human Development and Juvenile Diabetes Research Foundation International. We thank the International Genomics of Alzheimer’s Project for providing summary data results for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in the analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i-Select chips were funded by the French National Foundation on Alzheimer’s disease and related disorders. EADI was supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant, Inserm, Institut Pasteur de Lille, Université de Lille 2, and the Lille University Hospital. GERAD/PERADES was supported by the Medical Research Council (grant no. 503480), Alzheimer’s Research U.K. (grant no. 503176), the Wellcome Trust (grant no. 082604/2/07/Z) and the German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant no. 01GI0102, 01GI0711, 01GI0420.

CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01-AG-12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the Alzheimer's Association grant ADGC-10-196728.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Etminan, M.; Gill, S.; Samii, A. Effect of Non-Steroidal Anti-Inflammatory Drugs on Risk of Alzheimer's Disease: Systematic Review and Meta-Analysis of Observational Studies. *BMJ* **2003**, *327*, 128. [[CrossRef](#)] [[PubMed](#)]
2. Hakobyan, S.; Harding, K.; Aiyaz, M.; Hye, A.; Dobson, R.; Baird, A.; Liu, B.; Harris, C.L.; Lovestone, S.; Morgan, B.P. Complement Biomarkers as Predictors of Disease Progression in Alzheimer's Disease. *J. Alzheimers Dis.* **2016**, *54*, 707–716. [[CrossRef](#)]
3. Morgan, A.R.; Touchard, S.; O'Hagan, C.; Sims, R.; Majounie, E.; Escott-Price, V.; Jones, L.; Williams, J.; Morgan, B.P. The Correlation between Inflammatory Biomarkers and Polygenic Risk Score in Alzheimer's Disease. *J. Alzheimers Dis.* **2017**, *56*, 25–36. [[CrossRef](#)]
4. Goate, A.; Chartier-Harlin, M.C.; Mullan, M.; Brown, J.; Crawford, F.; Fidani, L.; Giuffra, L.; Haynes, A.; Irving, N.; James, L.; et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* **1991**, *349*, 704–706. [[CrossRef](#)] [[PubMed](#)]
5. Sherrington, R.; Rogaev, E.I.; Liang, Y.; Rogaeva, E.A.; Levesque, G.; Ikeda, M.; Chi, H.; Lin, C.; Li, G.; Holman, K.; et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* **1995**, *375*, 754–760. [[CrossRef](#)] [[PubMed](#)]
6. Levy-Lahad, E.; Wasco, W.; Poorkaj, P.; Romano, D.M.; Oshima, J.; Pettingell, W.H.; Yu, C.E.; Jondro, P.D.; Schmidt, S.D.; Wang, K.; et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* **1995**, *269*, 973–977. [[CrossRef](#)]
7. Hardy, J. The discovery of Alzheimer-causing mutations in the APP gene and the formulation of the “amyloid cascade hypothesis”. *FEBS J.* **2017**, *284*, 1040–1044. [[CrossRef](#)]
8. Corder, E.H.; Saunders, A.M.; Risch, N.J.; Strittmatter, W.J.; Schmechel, D.E.; Gaskell, P.C., Jr.; Rimmler, J.B.; Locke, P.A.; Conneally, P.M.; Schmechel, K.E.; et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat. Genet.* **1994**, *7*, 180–184. [[CrossRef](#)]
9. Saunders, A.M.; Schmechel, K.; Breitner, J.C.; Benson, M.D.; Brown, W.T.; Goldfarb, L.; Goldgaber, D.; Manwaring, M.G.; Szymanski, M.H.; McCown, N.; et al. Apolipoprotein E epsilon 4 allele distributions in late-onset Alzheimer's disease and in other amyloid-forming diseases. *Lancet* **1993**, *342*, 710–711. [[CrossRef](#)]
10. Saunders, A.M.; Strittmatter, W.J.; Schmechel, D.; George-Hyslop, P.H.; Pericak-Vance, M.A.; Joo, S.H.; Rosi, B.L.; Gusella, J.F.; Crapper-MacLachlan, D.R.; Alberts, M.J.; et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* **1993**, *43*, 1467–1472. [[CrossRef](#)]
11. Strittmatter, W.J.; Weisgraber, K.H.; Huang, D.Y.; Dong, L.M.; Salvesen, G.S.; Pericak-Vance, M.; Schmechel, D.; Saunders, A.M.; Goldgaber, D.; Roses, A.D. Binding of human apolipoprotein E to synthetic amyloid beta peptide: Isoform-specific effects and implications for late-onset Alzheimer disease. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 8098–8102. [[CrossRef](#)]
12. Lanfranco, M.F.; Ng, C.A.; Rebeck, G.W. ApoE Lipidation as a Therapeutic Target in Alzheimer's Disease. *Int. J. Mol. Sci.* **2020**, *21*, 6336. [[CrossRef](#)]
13. Bellenguez, C.; Küçükali, F.; Jansen, I.; Andrade, V.; Moreno-Grau, S.; Amin, N.; Naj, A.C.; Grenier-Boley, B.; Campos-Martin, R.; Holmans, P.A.; et al. New insights on the genetic etiology of Alzheimer's and related dementia. *medRxiv* **2020**. [[CrossRef](#)]
14. Kunkle, B.W.; Grenier-Boley, B.; Sims, R.; Bis, J.C.; Damotte, V.; Naj, A.C.; Boland, A.; Vronskaya, M.; van der Lee, S.J.; Amlie-Wolf, A.; et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat. Genet.* **2019**, *51*, 414–430. [[CrossRef](#)]
15. Sims, R.; van der Lee, S.J.; Naj, A.C.; Bellenguez, C.; Badarinarayan, N.; Jakobsdottir, J.; Kunkle, B.W.; Boland, A.; Raybould, R.; Bis, J.C.; et al. Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. *Nat. Genet.* **2017**, *49*, 1373–1384. [[CrossRef](#)] [[PubMed](#)]
16. Lambert, J.C.; Heath, S.; Even, G.; Campion, D.; Sleegers, K.; Hiltunen, M.; Combarros, O.; Zelenika, D.; Bullido, M.J.; Tavernier, B.; et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat. Genet.* **2009**, *41*, 1094–1099. [[CrossRef](#)] [[PubMed](#)]
17. Jones, L.; Holmans, P.A.; Hamshere, M.L.; Harold, D.; Moskvina, V.; Ivanov, D.; Pocklington, A.; Abraham, R.; Hollingworth, P.; Sims, R.; et al. Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. *PLoS ONE* **2010**, *5*, e13950. [[CrossRef](#)]
18. Ahmad, S.; Bannister, C.; van der Lee, S.J.; Vojinovic, D.; Adams, H.H.H.; Ramirez, A.; Escott-Price, V.; Sims, R.; Baker, E.; Williams, J.; et al. Disentangling the biological pathways involved in early features of Alzheimer's disease in the Rotterdam Study. *Alzheimers Dement.* **2018**, *14*, 848–857. [[CrossRef](#)]

19. Guerreiro, R.; Wojtas, A.; Bras, J.; Carrasquillo, M.; Rogaeva, E.; Majounie, E.; Cruchaga, C.; Sassi, C.; Kauwe, J.S.; Younkin, S.; et al. TREM2 variants in Alzheimer's disease. *N. Engl. J. Med.* **2013**, *368*, 117–127. [[CrossRef](#)]
20. Harold, D.; Abraham, R.; Hollingworth, P.; Sims, R.; Gerrish, A.; Hamshere, M.L.; Pahwa, J.S.; Moskva, V.; Dowzell, K.; Williams, A.; et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat. Genet.* **2009**, *41*, 1088–1093. [[CrossRef](#)] [[PubMed](#)]
21. De Leeuw, C.A.; Mooij, J.M.; Heskes, T.; Posthuma, D. MAGMA: Generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **2015**, *11*, e1004219. [[CrossRef](#)]
22. Genomes Project, C.; Auton, A.; Brooks, L.D.; Durbin, R.M.; Garrison, E.P.; Kang, H.M.; Korbel, J.O.; Marchini, J.L.; McCarthy, S.; McVean, G.A.; et al. A global reference for human genetic variation. *Nature* **2015**, *526*, 68–74. [[CrossRef](#)]
23. Baker, E.; Schmidt, K.M.; Sims, R.; O'Donovan, M.C.; Williams, J.; Holmans, P.; Escott-Price, V.; Consortium, W.T.G. POLARIS: Polygenic LD-adjusted risk score approach for set-based analysis of GWAS data. *Genet. Epidemiol.* **2018**, *42*, 366–377. [[CrossRef](#)]
24. Veerhuis, R.; van der Valk, P.; Janssen, I.; Zhan, S.S.; Van Nostrand, W.E.; Eikelenboom, P. Complement activation in amyloid plaques in Alzheimer's disease brains does not proceed further than C3. *Virchows Arch.* **1995**, *426*, 603–610. [[CrossRef](#)]
25. Rogers, J.; Cooper, N.R.; Webster, S.; Schultz, J.; McGeer, P.L.; Styren, S.D.; Civin, W.H.; Brachova, L.; Bradt, B.; Ward, P.; et al. Complement activation by beta-amyloid in Alzheimer disease. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10016–10020. [[CrossRef](#)] [[PubMed](#)]
26. Webster, S.; Lue, L.F.; Brachova, L.; Tenner, A.J.; McGeer, P.L.; Terai, K.; Walker, D.G.; Bradt, B.; Cooper, N.R.; Rogers, J. Molecular and cellular characterization of the membrane attack complex, C5b-9, in Alzheimer's disease. *Neurobiol. Aging* **1997**, *18*, 415–421. [[CrossRef](#)]
27. Ishii, T.; Haga, S. Immuno-electron-microscopic localization of complements in amyloid fibrils of senile plaques. *Acta Neuropathol.* **1984**, *63*, 296–300. [[CrossRef](#)] [[PubMed](#)]
28. Bradt, B.M.; Kolb, W.P.; Cooper, N.R. Complement-dependent proinflammatory properties of the Alzheimer's disease beta-peptide. *J. Exp. Med.* **1998**, *188*, 431–438. [[CrossRef](#)]
29. Chu, Y.; Jin, X.; Parada, I.; Pesic, A.; Stevens, B.; Barres, B.; Prince, D.A. Enhanced synaptic connectivity and epilepsy in C1q knockout mice. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 7975–7980. [[CrossRef](#)]
30. Hong, S.; Beja-Glasser, V.F.; Nfonoyim, B.M.; Frouin, A.; Li, S.; Ramakrishnan, S.; Merry, K.M.; Shi, Q.; Rosenthal, A.; Barres, B.A.; et al. Complement and microglia mediate early synapse loss in Alzheimer mouse models. *Science* **2016**, *352*, 712–716. [[CrossRef](#)] [[PubMed](#)]
31. Sekar, A.; Bialas, A.R.; de Rivera, H.; Davis, A.; Hammond, T.R.; Kamitaki, N.; Tooley, K.; Presumey, J.; Baum, M.; Van Doren, V.; et al. Schizophrenia risk from complex variation of complement component 4. *Nature* **2016**, *530*, 177–183. [[CrossRef](#)]
32. Stevens, B.; Allen, N.J.; Vazquez, L.E.; Howell, G.R.; Christopherson, K.S.; Nouri, N.; Micheva, K.D.; Mehalow, A.K.; Huberman, A.D.; Stafford, B.; et al. The classical complement cascade mediates CNS synapse elimination. *Cell* **2007**, *131*, 1164–1178. [[CrossRef](#)] [[PubMed](#)]
33. Shi, Q.; Colodner, K.J.; Matousek, S.B.; Merry, K.; Hong, S.; Kenison, J.E.; Frost, J.L.; Le, K.X.; Li, S.; Dodart, J.C.; et al. Complement C3-Deficient Mice Fail to Display Age-Related Hippocampal Decline. *J. Neurosci.* **2015**, *35*, 13029–13042. [[CrossRef](#)]
34. Geerlings, M.J.; de Jong, E.K.; den Hollander, A.I. The complement system in age-related macular degeneration: A review of rare genetic variants and implications for personalized treatment. *Mol. Immunol.* **2017**, *84*, 65–76. [[CrossRef](#)] [[PubMed](#)]
35. Harris, C.L.; Heurich, M.; Rodriguez de Cordoba, S.; Morgan, B.P. The complement: Dictating risk for inflammation and infection. *Trends Immunol.* **2012**, *33*, 513–521. [[CrossRef](#)]
36. Hourcade, D.; Holers, V.M.; Atkinson, J.P. The regulators of complement activation (RCA) gene cluster. *Adv. Immunol.* **1989**, *45*, 381–416. [[CrossRef](#)]
37. Ebbert, M.T.W.; Jensen, T.D.; Jansen-West, K.; Sens, J.P.; Reddy, J.S.; Ridge, P.G.; Kauwe, J.S.K.; Belzil, V.; Prgent, L.; Carrasquillo, M.M.; et al. Systematic analysis of dark and camouflaged genes reveals disease-relevant genes hiding in plain sight. *Genome Biol.* **2019**, *20*, 97. [[CrossRef](#)]
38. Tschopp, J.; Chonn, A.; Hertig, S.; French, L.E. Clusterin, the human apolipoprotein and complement inhibitor, binds to complement C7, C8 beta, and the b domain of C9. *J. Immunol.* **1993**, *151*, 2159–2165. [[PubMed](#)]
39. Burkey, B.F.; Stuart, W.D.; Harmony, J.A. Hepatic apolipoprotein J is secreted as a lipoprotein. *J. Lipid Res.* **1992**, *33*, 1517–1526. [[CrossRef](#)]
40. Humphreys, D.T.; Carver, J.A.; Easterbrook-Smith, S.B.; Wilson, M.R. Clusterin has chaperone-like activity similar to that of small heat shock proteins. *J. Biol. Chem.* **1999**, *274*, 6875–6881. [[CrossRef](#)]
41. Jenne, D.E.; Tschopp, J. Molecular structure and functional characterization of a human complement cytolysis inhibitor found in blood and seminal plasma: Identity to sulfated glycoprotein 2, a constituent of rat testis fluid. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 7123–7127. [[CrossRef](#)] [[PubMed](#)]
42. Brouwers, N.; Van Cauwenberghe, C.; Engelborghs, S.; Lambert, J.C.; Bettens, K.; Le Bastard, N.; Pasquier, F.; Montoya, A.G.; Peeters, K.; Mattheijssens, M.; et al. Alzheimer risk associated with a copy number variation in the complement receptor 1 increasing C3b/C4b binding sites. *Mol. Psychiatry* **2012**, *17*, 223–233. [[CrossRef](#)]

43. Hazrati, L.N.; Van Cauwenberghe, C.; Brooks, P.L.; Brouwers, N.; Ghani, M.; Sato, C.; Cruts, M.; Sleegers, K.; St George-Hyslop, P.; Van Broeckhoven, C.; et al. Genetic association of CR1 with Alzheimer's disease: A tentative disease mechanism. *Neurobiol. Aging* **2012**, *33*, e2945–e2949. [[CrossRef](#)] [[PubMed](#)]
44. Kucukkilic, E.; Brookes, K.; Barber, I.; Guetta-Baranes, T.; Consortium, A.; Morgan, K.; Hollox, E.J. Complement receptor 1 gene (CR1) intragenic duplication and risk of Alzheimer's disease. *Hum. Genet.* **2018**, *137*, 305–314. [[CrossRef](#)] [[PubMed](#)]
45. Mahmoudi, R.; Feldman, S.; Kisserli, A.; Duret, V.; Tabary, T.; Bertholon, L.A.; Badr, S.; Nonnonhou, V.; Cesar, A.; Neuraz, A.; et al. Inherited and Acquired Decrease in Complement Receptor 1 (CR1) Density on Red Blood Cells Associated with High Levels of Soluble CR1 in Alzheimer's Disease. *Int. J. Mol. Sci.* **2018**, *19*, 2175. [[CrossRef](#)]
46. Mahmoudi, R.; Kisserli, A.; Novella, J.L.; Donvito, B.; Drame, M.; Reveil, B.; Duret, V.; Jolly, D.; Pham, B.N.; Cohen, J.H. Alzheimer's disease is associated with low density of the long CR1 isoform. *Neurobiol. Aging* **2015**, *36*, e1765–e1766. [[CrossRef](#)]
47. McLure, C.A.; Williamson, J.F.; Stewart, B.J.; Keating, P.J.; Dawkins, R.L. Indels and imperfect duplication have driven the evolution of human Complement Receptor 1 (CR1) and CR1-like from their precursor CR1 alpha: Importance of functional sets. *Hum. Immunol.* **2005**, *66*, 258–273. [[CrossRef](#)]
48. Logar, C.M.; Chen, W.; Schmitt, H.; Yu, C.Y.; Birmingham, D.J. A human CR1-like transcript containing sequence for a binding protein for iC4 is expressed in hematopoietic and fetal lymphoid tissue. *Mol. Immunol.* **2004**, *40*, 831–840. [[CrossRef](#)] [[PubMed](#)]