The *MS4A* gene cluster is a key modulator of soluble TREM2 and Alzheimer disease risk

Yuetiva Deming[1][*], Fabia Filipello[2][*], Francesca Cignarella[2][*], Claudia Cantoni[2], Simon Hsu[1], Robert Mikesell[2], Zeran Li[1], Jorge L Del-Aguila[1], Umber Dube[1], Fabiana Geraldo Farias[1], Joseph Bradley[1], John Budde[1], Laura Ibanez[1], Maria Victoria Fernandez[1], Alzheimer's Disease Neuroimaging Initiative (ADNI)[†], Dominantly Inherited Alzheimer Network (DIAN), Kaj Blennow[3,4,5,6], Henrik Zetterberg[3,4,5,6], Amanda Heslegrave[5,6], Per M Johansson[5,6,7], Johan Svensson[8], Bengt Nellgård[8], Alberto Lleo[9], Daniel Alcolea[9], Jordi Clarimon[9], Lorena Rami[10], José Luis Molinuevo[11,12], Marc Suárez-Calvet[12,13,14], Estrella Morenas-Rodríguez[13,15], Gernot Kleinberger[13,15], Michael Ewers[16], Oscar Harari[1], Christian Haass[13,14,15], Thomas J Brett[17], Bruno A. Benitez[1,#], Celeste M. Karch[1,18,#], Laura Piccio[2,18,#], Carlos Cruchaga[1,18,*,#]

Affiliations:
[1]Department of Psychiatry, Washington University School of Medicine, 660 S. Euclid Ave. B8134, St. Louis, MO 63110, USA.
[2]Department of Neurology, Washington University School of Medicine, St Louis, MO, USA
[3]Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, the Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden.
[4]Clinical Neurochemistry Laboratory, Department of Neuroscience and Physiology, University of Gothenburg, Sahlgrenska University Hospital, Mölndal, Sweden.
[5]Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, London, UK.
[6]UK Dementia Research Institute at UCL, London, UK.
[7]Department of Anesthesiology and Intensive Care Medicine, Sahlgrenska University. Hospital, Mölndal, Sweden and Institute of Clinical Sciences, the Sahlgrenska Academy at the University of Gothenburg, Sweden
[8]Department of Internal Medicine, Institute of Medicine, the Sahlgrenska Academy at the University of Gothenburg, Göteborg, Sweden.
[9]Neurology Dept. Hospital de la Santa Creu i Sant Pau. Sant Antoni Mª Claret 167. Barcelona 08025. Spain
[10]IDIBAPS. Alzheimer´s disease and other cognitive disorders unit, Neurology Service, ICN Hospital Clinic.
[11]Alzheimer´s disease and other cognitive disorders unit, Neurology Service, ICN Hospital Clinic i Universitari
[12]BarcelonaBeta Brain Research Center, Pasqual Maragall Foundation. Barcelona, Spain
[13]Biomedical Center (BMC), Biochemistry, Ludwig-Maximilians-Universität München, Munich, Germany.
[14]German Center for Neurodegenerative Diseases (DZNE), Munich, Germany
[15]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
[16]Institute for Stroke and Dementia Research, University Hospital, LMU Munich, Germany

[17]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Washington University School of Medicine, St. Louis, United States
[18]Hope Center for Neurological Disorders. Washington University School of Medicine, 660 S. Euclid Ave. B8111, St. Louis, MO 63110, USA.

#To whom correspondence should be addressed: Drs. Celeste Karch (karchc@wustl.edu), Laura Piccio (picciol@wustl.edu) or Carlos Cruchaga (ccruchaga@wustl.edu)

# Authors with equal contribution
*Authors with equal contribution

One Sentence Summary: Common variants in the microglia-specific *MS4A* gene cluster modify

risk for late-onset Alzheimer disease and modulate extracellular soluble TREM2.

Abstract: Soluble triggering receptor expressed on myeloid cells 2 (sTREM2) in the cerebrospinal fluid (CSF) has been associated with Alzheimer disease (AD). TREM2 plays a critical role in microglial activation, survival, and phagocytosis; however, the pathophysiological role of sTREM2 in AD is not well understood. Understanding the role of sTREM2 in AD may reveal novel pathological mechanisms underlying AD and identify distinct therapeutic targets. We performed a genome-wide association study (GWAS) to identify genetic modifiers of CSF sTREM2 obtained from the Alzheimer Disease Neuroimaging Initiative (ADNI, $N = 813$). Common variants in the membrane-spanning 4-domains subfamily A (*MS4A*) gene region were associated with CSF sTREM2 (rs1582763; $P = 1.15\times10^{-15}$) which replicated in independent datasets. The variants associated with increased CSF sTREM2 are associated with reduced AD risk and delayed age-at-onset. Rs1582763 modifies expression of *MS4A4A* and *MS4A6A* in multiple tissues, suggesting that one or both of these genes are important for modulating sTREM2. *MS4A* genes encode transmembrane proteins that are expressed in microglia. Using human macrophages to model the relationship between MS4A4A and TREM2, we found that MS4A4A and TREM2 colocalize on lipid rafts at the plasma membrane, that sTREM2 increases with *MS4A4A* overexpression, and that antibody-mediated targeting of MS4A4A reduces sTREM2. Thus, genetic, molecular, and cellular findings suggest that MS4A4A modulates sTREM2. These findings also provide a mechanistic explanation for the original GWAS signal in the *MS4A* locus for AD risk and indicate that TREM2 is involved in sporadic AD pathogenesis, not only in *TREM2* risk-variant carriers.

Introduction

In 2013, two groups independently identified a rare variant in *TREM2* (p.R47H) that increased risk for AD almost three-fold, making *TREM2* the strongest genetic risk factor for late-onset AD since the identification of *APOE* $\varepsilon4$ 30 years earlier (*1, 2*). *TREM2* p.R47H is also associated with clinical, imaging, and neuropathological AD phenotypes, including advanced behavioral symptoms, gray matter atrophy, and Braak staging (*3, 4*). Additional rare variants in *TREM2* have also been associated with AD risk including p.R62H, p.H157Y, p.R98W, p.D87N, p.T66M, p.Y38C, and p.Q33X (*1, 5-7*).

TREM2 encodes a protein that is part of a transmembrane receptor-signaling complex essential for the immune response of myeloid cells such as microglia. TREM2 is highly and specifically expressed in microglia in the central nervous system (CNS) where it plays a key role in microglial activation, survival, and phagocytosis (*8-11*). Recent evidence from AD animal models suggests that TREM2 has a complex relationship with the key proteins involved in AD pathogenesis (amyloid-beta (Aβ) and tau) (*12*). In studies suggesting that TREM2 may be protective in AD, TREM2 has been associated with phagocytosis in the presence of Aβ (*13*). The introduction of the human *TREM2* transgene reduced pathology and rescued cognitive function in amyloid-bearing mice (*14*). A common phenotype across mouse models of amyloidosis is that reduction or loss of *Trem2* leads to fewer amyloid plaque-associated microglia, resulting in more diffuse plaques and enhanced neuritic damage (*15-17*). In contrast, studies of tau pathology indicate that *Trem2* deficiency protects against neurodegeneration, suggesting a detrimental role for TREM2 in the presence of tau pathology (*18, 19*). Together, these studies demonstrate that TREM2 plays an important, yet complex, role in AD pathology.

In addition to the full-length TREM2 which makes up part of the transmembrane receptor, a soluble form of TREM2 (sTREM2) is produced by alternative splicing or proteolytic cleavage of the full-length TREM2 protein (*6, 13, 20*). Three distinct transcripts for *TREM2* are found in human brain tissue, including a short alternatively spliced transcript that excludes exon 4 (encoding the transmembrane domain) to produce a soluble isoform of TREM2 (*6*). Recently a cleavage site was identified in TREM2 that can be targeted by proteases such as ADAM10, ADAM17, and gamma-secretase to generate sTREM2 (*21, 22*). sTREM2 is released into the cerebrospinal fluid (CSF) where it can be quantified (*13, 23-28*). CSF sTREM2 is hypothesized to increase in response to microglial activation due to neurodegenerative processes (*24, 27*). Our group and others have demonstrated that CSF sTREM2 is elevated in AD (*23, 24, 26*). Changes in CSF sTREM2 appear to occur after amyloid accumulation, beginning approximately five years before clinical symptom onset in autosomal dominant forms of AD (*25*). Additionally, CSF sTREM2 positively correlates with CSF tau and phosphorylated tau (ptau), but not with CSF $A\beta_{42}$, suggesting that sTREM2 may be associated with pathological processes occurring after the accumulation of $A\beta$ (*23, 24, 26*). Thus, CSF sTREM2 has emerged as an important and dynamic biomarker of disease processes throughout AD pathogenesis.

Our group and others have shown that using CSF biomarkers of complex diseases as quantitative endophenotypes can help identify genes that contribute to key biological pathways involved in disease (*29-31*). In this study, we aimed to identify genetic modifiers of CSF sTREM2. We conducted a genome wide association study (GWAS) of CSF sTREM2 from more than 1,390 individuals and detected a locus within the *MS4A* gene region (11q12.2) that displays genome-wide significant association with CSF sTREM2. We then used bioinformatic, transcriptomic, and cellular approaches to examine correlations between *MS4A* genes and

TREM2. Based on these findings, we propose that the *MS4A* gene family contains one or more key modulators of sTREM2.

Results

The goal of this study was to identify genetic modifiers of CSF sTREM2 and to better understand the role of sTREM2 in biological pathways relevant to AD. To accomplish this goal, we analyzed genetic data and CSF sTREM2 from 813 individuals from the Alzheimer's Disease Neuroimaging Initiative (ADNI). We then sought to replicate our findings in independent datasets ($N = 580$).

CSF sTREM2 correlates with age, sex, and AD biomarkers

Our discovery cohort was obtained from the ADNI ($N = 813$), the largest dataset for CSF sTREM2 to date. The analyzed dataset included 172 AD, 169 cognitive normal,183 early Mild Cognitive impairment (eMCI), 221 late-MCI and 68 Significant Memory Concern (SMC) individuals. We excluded participants with MCI due to non-AD causes and participants carrying previously identified risk mutations in *TREM2* (Supplementary Table 1). CSF sTREM2 in our analyses approximated a normal distribution without transformation; therefore, we analyzed the raw values, unless otherwise stated.

We tested the correlation between age at the time of lumbar puncture (LP) and CSF sTREM2 and found a positive correlation, consistent with our previous reports (Pearson $r = 0.270$, $P = 4.66 \times 10^{-15}$; Fig. S1) (*24*). We also evaluated whether CSF sTREM2 was associated with sex, as previously reported (*24-26, 32*). There was no significant difference between CSF

sTREM2 in males (3,926 ± 1,969 pg/mL) versus females (3,873 ± 1,856 pg/mL, one-tailed $P = 0.694$; Fig. S1). CSF sTREM2 positively correlated with CSF tau ($r = 0.377$, $P = 9.59 \times 10^{-29}$) and ptau ($r = 0.348$, $P = 2.08 \times 10^{-24}$; Fig. S1). Consistent with our previous findings, there was no significant correlation between CSF sTREM2 and CSF $A\beta_{42}$ levels ($r = 0.074$, $P = 0.052$; Fig. S1). CSF sTREM2 was significantly correlated with tau/$A\beta_{42}$ and ptau/$A\beta_{42}$ ratio ($r = 0.171$, $P = 5.71 \times 10^{-6}$; $r = 0.176$, $P = 3.18 \times 10^{-6}$; respectively). However, these correlations were less than observed between CSF sTREM2 and tau or ptau, suggesting that the correlations are likely driven by the associations with CSF tau and ptau levels (Supplementary Fig. S1).

We also examined CSF sTREM2 in 38 individuals heterozygous for *TREM2* risk variants which were not included in the GWAS: p.D87N, p.L211P, p.R47H, p.R62H, and p.H157Y (Table 2). There was a significant difference in CSF sTREM2 between *TREM2* risk variant carriers ($P = 0.042$) as reported previously (*24*). We also observed variant specific effects on CSF sTREM2. CSF sTREM2 was significantly lower in p.L211P carriers (2,578 ± 1,444 pg/mL) compared to non-carriers (3,903 ± 1,919 pg/mL, $P = 0.026$; Supplementary Fig. S2). CSF sTREM2 was significantly elevated in p.R47H carriers (4,852 ± 1,172 pg/mL) compared to p.L211P carriers ($P = 0.029$) and p.D87N carriers (2,082 ± 239 pg/mL, $P = 0.041$), but not significantly different from non-carriers ($P = 0.223$; Fig. S2).

The MS4A gene region is associated with CSF sTREM2

To identify genetic variants that modify CSF sTREM2, we tested for association between CSF sTREM2 and single-nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) >0.02 using an additive linear regression model with age, sex, and two principal

components (PCs) as covariates. A total of 7,320,475 genotyped and imputed SNPs passed strict quality control as described in the methods. *TREM2* risk-variant carriers were excluded from the GWAS analyses.

A genome-wide significant genetic association with CSF sTREM2 was identified on chromosome 11 within the *MS4A* gene region (Fig. 1, and Supplementary Table S2 and Fig. S3). The top SNP was rs1582763, an intergenic variant nearest *MS4A4A* (located 26 Kb 5′of *MS4A4A*; 11q12.2). This common variant (A allele, MAF = 0.368, genotyped) produced a genome-wide significant association with CSF sTREM2 ($N = 807$, $\beta = 735.1$, $P = 1.15 \times 10^{-15}$; Fig. 1) and explained more than 6% of the variance in sTREM2 (adjusted $r^2 = 0.064$; see supplementary results). Results from case-control stratified analyses indicate that both clinically diagnosed cognitively impaired individuals ($N = 606$, $\beta = 675.8$, $P = 8.19 \times 10^{-10}$) and cognitively normal controls ($N = 207$, $\beta = 912.4$, $P = 5.20 \times 10^{-8}$) contributed to the association between the *MS4A* locus and CSF sTREM2 (Supplementary Fig. S4). No common variants were identified that modified CSF sTREM2 and reached genome-wide significant P values in the *APOE, TYROBP, TREM2, ADAM10* or *ADAM17* regions (Supplementary Fig. S5). Interestingly, common variants within the *TREM* family gene locus showed a marginal association with CSF sTREM2 (rs200443049, intronic indel, $N = 805$, $\beta = 416.5$, $P = 2.01 \times 10^{-5}$; Supplementary Fig. S5).

To determine whether there was more than one independent signal within the *MS4A* locus, we performed conditional analyses including rs1582763 in the regression model. After conditioning for rs1582763, the most significant variant was rs6591561 (G allele, MAF = 0.316, genotyped), a missense variant within *MS4A4A* (NP_076926.2:p.M159V), which is not in linkage disequilibrium (LD) with rs1582763 ($r^2 = 0.112$, D′ = 0.646). Rs6591561 (*MS4A4A*

p.M159V) was associated with CSF sTREM2 but in the opposite direction of the main signal

(rs1582763). The minor allele of rs6591561 was associated with reduced CSF sTREM2, and this

association was independent of rs1582763 (before conditioning: β = -783.7, P = 1.47×10^{-9}; after

conditioning: β = -378.4, P = 1.55×10^{-4}; Fig. 1). After conditioning for rs6591561 (*MS4A4A*

p.M159V), rs1582763 remained genome-wide significant (β = 625.7, P = 4.52×10^{-11}; Fig. 1),

providing further evidence that these are independent signals.

Replication in independent datasets and meta-analysis

To replicate the associations of rs1582763 and rs6591561(*MS4A4A* p.M159V), we

analyzed an additional 580 samples with CSF sTREM2 and genetic data that were obtained from

six different studies (Charles F. and Joanne Knight Alzheimer Disease Research Center (Knight

ADRC), the Dominantly Inherited Alzheimer Network (DIAN), two studies from the

Sahlgrenska Academy at the University of Gothenburg (GHDEM and GHPH), the Memory Unit

and Alzheimer's laboratory at the Hospital of Sant Pau in Barcelona (Sant Pau Initiative on

Neurodegeneration; SPIN), and the ICN Hospital Clinic-IDIBAPS in Barcelona. These datasets

have been described previously (*24-26, 32*) (Supplementary Table S3).

We replicated the association of rs1582763 with elevated CSF sTREM2 (z = 5.743, P =

9.28×10^{-9}; Fig. 2). A meta-analysis including ADNI and the replication datasets showed a strong

association between rs1582763 and elevated CSF sTREM2 (P_{meta} = 4.48×10^{-21}; Fig. 2).

Additionally, we replicated the association between rs6591561 (*MS4A4A* p.M159V) and reduced

CSF sTREM2 in the independent datasets (z = -3.715, P = 2.03×10^{-4}; Fig. 2). A meta-analysis of

rs6591561 including ADNI with the replication datasets also showed a genome-wide significant

association ($P_{meta} = 1.65 \times 10^{-11}$; Fig. 2). Individual results for each dataset are shown in Supplementary Table S4.

Comparisons between MS4A association and APOE

APOE genotype is the strongest genetic risk factor for AD (*33*) and *APOE* $\varepsilon 4$ is the strongest genetic association for CSF Aβ_{42} and tau (*31, 34*). To determine whether *APOE* genotype is associated with CSF sTREM2, we created a variable representing *APOE*-mediated AD risk by recoding *APOE* genotype for each individual as 0 for $\varepsilon 2/\varepsilon 2$, 1 for $\varepsilon 2/\varepsilon 3$, 2 for $\varepsilon 3/\varepsilon 3$, 3 for $\varepsilon 2/\varepsilon 4$, 4 for $\varepsilon 3/\varepsilon 4$, and 5 for $\varepsilon 4/\varepsilon 4$. We added this *APOE* genotype risk variable to a regression model, adjusting for age, sex, and two PCs for population stratification and found that *APOE* genotype was not significantly associated with CSF sTREM2 ($\beta = -26.5$, $P = 0.613$). The SNP commonly used as a proxy for *APOE* $\varepsilon 4$, rs769449, was also not associated with CSF sTREM2 ($\beta = -106.9$, $P = 0.308$).

The influence of *APOE* genotype on Aβ pathology, and thereby biomarkers for Aβ pathology including CSF Aβ_{42}, has been a consistent finding (*31, 35*). Of the 813 individuals in this study, we had data for CSF Aβ_{42} available in 695 individuals. To compare CSF Aβ_{42} and sTREM2 on the same scale, we converted Aβ_{42} and sTREM2 values to z-scores by subtracting the mean and dividing by the standard deviation for each respective protein. After conversion, we verified the association between *APOE* genotype and CSF Aβ_{42}. As expected, there was a strong negative effect of *APOE* genotype on CSF Aβ_{42} ($\beta = -0.361$, $P = 1.69 \times 10^{-34}$). We also stratified the CSF Aβ_{42} into quartiles and calculated the OR of *APOE* genotype for the first vs.

the last quartile. *APOE* genotype had an OR = 2.84 for lower CSF Aβ_{42} (95% CI = 2.30 – 3.56, *P* = 8.44×10^{-21}).

Within the same subset of individuals, the effect size of rs1582763 on CSF sTREM2 was similar (β = 0.366, *P* = 7.53×10^{-13}) to that of *APOE* on CSF Aβ_{42} (β = -0.361, *P* = 1.69×10^{-34}). When comparing the lower vs. upper quartiles of CSF sTREM2, rs1582763 had an OR = 3.34 for higher CSF sTREM2 (95% CI = 2.37 – 4.82, *P* = 2.50×10^{-11}), which is similar to the magnitude of effect of *APOE* on CSF Aβ_{42} (OR = 2.84, 95% CI 2.30 – 3.56). Thus, the impact of rs1582763, located in the *MS4A* gene region, on CSF sTREM2 is similar to that of *APOE*, the major modulator of AD risk, on CSF Aβ_{42}.

Additional genes associated with CSF sTREM2

To identify genes associated with CSF sTREM2, we used MAGMA, which maps every SNP to the nearest gene, takes into account LD structure, and uses multiple regression analyses to provide a P value for the association of each gene with the tested phenotype. There were four genes associated with CSF sTREM2 that passed multiple test correction (*P* < 2.75×10^{-6}). All of these genes belong to the *MS4A* gene family: *MS4A4A* (*P* = 3.15×10^{-11}), *MS4A4E* (*P* = 6.13×10^{-12}), *MS4A2* (*P* = 1.29×10^{-11}), and *MS4A6A* (*P* = 1.44×10^{-11}; Fig. 3). Interestingly, *TREML2* approached gene-wide significance with 89 mapped variants (*P* = 3.23×10^{-6}; Fig. 3), suggesting there may be additional genes contributing to CSF sTREM2.

Mendelian Randomization

We used Mendelian randomization (MR) to determine whether CSF sTREM2 is involved on AD pathogenesis or if it is just associated with disease due to some confounding factor or reverse causality. Mendelian randomization estimates a causal effect from observational data, in this case CSF sTREM2, in the presence of confounding factors. We used the two independent genome-wide SNPs in Chr. 11 (rs1582763 and rs6591561) as instrumental factors. We also used the top SNP in Chr. 6 (rs28385608), as this locus was marginally significant in both the single-variant and gene-based analyses. Summary statistics from the largest AD risk GWAS published to date (*36*) were used as input data. We ran multiple MR models to take into account any potential heterogeneity in the data and to obtain robust estimates. For example, the penalized MR-Egger model downweights the variants with heterogeneous causal estimates. In all MR models, the analyses yielded a significant association ($P \leq 2.14 \times 10^{-3}$; Fig. 4), suggesting that CSF sTREM2 may be involved in AD pathogenesis.

Functional annotation of genome-wide significant signals in the MS4A gene region

The genome-wide significant association with CSF sTREM2 is located in a gene-rich region including at least 15 genes, most of which are members of the *MS4A* gene family: *OOSP2, OOSP4B, OOSP1, MS4A2*, *MS4A6A*, *MS4A4E*, *MS4A4A*, *MS4A6E*, *MS4A7*, *MS4A14*, *MS4A5*, *MS4A1*, *MS4A12*, *MS4A13*, and *MS4A8*. To determine the functional variants and specific genes that modulate CSF sTREM2, we performed additional bioinformatic analyses.

We first examined whether any of the genome-wide significant variants were located within exonic regions. We identified two coding variants: the missense variant within *MS4A4A*

rs6591561 (p.M159V, MAF = 0.316, β = -593.6, P = 1.47×10⁻⁹), which is reported above, and a synonymous variant within *MS4A6A*, rs12453 (p.L137L; MAF = 0.392, β = 710.5, P = 1.77×10⁻¹⁵) which is in high LD with rs1582763 (r^2 = 0.782, D' = 0.924). Neither variant was predicted by SIFT or PolyPhen to be loss-of-function, pathogenic, or damaging.

Because the top signal, rs1582763, is independent of the *MS4A4A* coding variant, we analyzed if this top variant showed evidence of expression quantitative trait locus (eQTL) effects on *MS4A* genes. In whole blood, rs1582763 was associated with reduced expression of *MS4A6A* (GTEx: β = -0.089, P = 3.90×10⁻⁵; Westra: Z = -23.42, P = 2.95×10⁻¹²¹) and *MS4A4A* (GTEx: β = -0.123, P = 6.1×10⁻⁵; Westra: Z = -7.21, P = 5.52×10⁻¹³), and increased expression of *MS4A2* (Westra: Z = 8.06, P = 7.49×10⁻¹⁶). Using gene expression data from blood obtained from 365 individuals in ADNI, we verified that rs1582763 produces a *cis*-eQTL with reduced *MS4A4A* (β = -0.226, P = 1.04×10⁻⁴) and *MS4A6A* expression (β = -0.080, P = 6.02×10⁻⁶). The eQTL effect of rs1582763 on *MS4A4A* and *MS4A6A* was observed in both clinically diagnosed AD cases (N = 235; *MS4A4A*: β = -0.194, P = 6.05×10⁻³; *MS4A6A*: β = -0.053, P=1.24×10⁻²) and cognitively normal controls (N = 80; *MS4A4A*: β = -0.326, P=1.10×10⁻²; *MS4A6A*: β = -0.138, P = 3.32×10⁻⁴), suggesting that the *cis*-eQTL effect is independent of disease status.

We also analyzed whether rs1582763 produces an eQTL effect in human brain tissue using GTEx and Braineac data. Interestingly, rs1582763 was nominally associated with *MS4A4A* expression in cortex from both GTEx (β = 0.215, P = 8.0×10⁻³) and Braineac (β = 0.327, P = 6.3×10⁻³; Supplementary Tables S5-S6). However, the direction of the eQTL effect in brain was opposite of the effect in blood. Similarly, rs1582763 produced a nominal *cis*-eQTL effect on *MS4A6A* expression in GTEx (Frontal cortex: β = 0.183, P = 0.04 and Cortex: β = 0.165, P =

0.05) and Braineac (medulla; $\beta = 0.337$, $P = 4.50\times10^{-4}$). These data suggest there may be tissue specific effects of rs1582763 on the expression of *MS4A* genes. These results also suggest that *MS4A4A* and/or *MS4A6A* are the genes modulating CSF sTREM2.

To determine whether rs1582763 produced a *trans*-eQTL effect on *TREM2* in brain tissue, we analyzed genetic and expression data from the Knight ADRC and Mount Sinai Brain Bank (*37*). Rs1582763 was not associated with *TREM2* gene expression levels (Knight ADRC, parietal lobe, $P = 0.61$; Mount Sinai Brain Bank, BM10 $P = 0.251$; BM22 $P = 0.635$; BM36 $P = 0.269$ and BM44 $P = 0.999$). These findings suggest that the association between rs1582763 and CSF sTREM2 is driven by the *cis*-eQTL effect on *MS4A4A* and/or *MS4A6A* rather than a direct effect on *TREM2* expression.

MS4A genes and TREM2 gene expression are highly correlated

To begin to understand the relationship between *MS4A* genes and *TREM2*, we examined the correlation between expression of *TREM2* and genes within the *MS4A* cluster using brain RNA-seq data. Among the 16 genes tested from the *MS4A* gene cluster, the expression of three genes (*MS4A4A*, *MS4A6A*, and *MS4A7*) were consistently positively correlated with *TREM2* gene expression in human brain tissue. In autopsy-confirmed late-onset AD cases and controls (Knight ADRC), *TREM2* expression was significantly correlated with expression of *MS4A4A* ($N = 40$, $r = 0.41$, $P = 8.00\times10^{-3}$) and *MS4A6A* ($N = 41$, $r = 0.67$, $P = 1.60\times10^{-6}$; Supplementary Fig. S6 and Table S7). These findings were replicated in two independent RNA-seq datasets obtained from the Mayo Clinic Brain Bank ($N = 162$; *MS4A4A*: $r = 0.68$, $P = 1.00\times10^{-23}$; *MS4A6A*: $r = 0.76$, $P = 1.60\times10^{-31}$; Supplementary Fig. S7 and Table S8) and Mount Sinai Brain Bank ($N = 300$; *MS4A4A*: $r = 0.61$, $P = 1.60\times10^{-16}$; *MS4A6A*: $r = 0.52$, $P = 7.90\times10^{-8}$; Supplementary Fig.

S8 and Table S9). These data further support that *MS4A4A* and/or *MS4A6A* modulate CSF sTREM2.

Targeting MS4A4A on human macrophages decreases soluble TREM2 in vitro

MS4A4A and *MS4A6A* are both highly expressed in microglia (*38*). The mouse ortholog for *MS4A6A* is *Ms4a6e,* however there is no mouse ortholog for *MS4A4A* (*39, 40*). Thus, to understand the role of *MS4A* genes on TREM2 function, we utilized human models. To begin evaluating the functional connection between TREM2 and MS4A4A and MS4A6A, we used human macrophages, since they express TREM2, MS4A4A, and MS4A6A, and are frequently employed to model microglia (*41*).

To examine the therapeutic potential of targeting these MS4A proteins, we first used an antibody-based approach. Human macrophages were derived *in vitro* from blood monocytes, and macrophage culture media tested for the presence of sTREM2 by ELISA. Since IL-4-mediated macrophage stimulation results in upregulation of TREM2 and MS4A4A (*42, 43*), we measured sTREM2 in the presence or absence of IL-4 as previously described (*27*). We detected sTREM2 in the macrophage media beginning at day 3-post isolation which increased progressively with culture time, consistent with previous reports ($P < 0.0001$; Supplementary Fig. S9A) (*27*). By day 7, stimulation with IL-4 had significantly increased sTREM2 in the media compared to un-treated cells ($P = 0.015$; Supplementary Fig. S9B).

Next, we evaluated whether antibody-mediated targeting of MS4A4A or MS4A6A modulated extracellular sTREM2. Human macrophages were treated on day 7 with commercially available antibodies to MS4A4A, MS4A6A, and isotype controls, then 48 hours post-treatment

sTREM2 was measured in culture media. We found that an antibody directed against MS4A4A was sufficient to significantly reduce sTREM2 ($P < 0.0001$; Fig. 5A). However, treatment with an anti-MS4A6A antibody failed to alter sTREM2 (Fig. 5A). We observed similar results when macrophages were stimulated with IL-4 prior to antibody treatment ($P < 0.0001$; Supplementary Fig. S9C). The effect of the MS4A4A antibody on sTREM2 levels was dose-dependent. Treatment of macrophages with 1, 5, or 10 μg/mL of the MS4A4A antibody produced a dose-dependent decrease in sTREM2 ($P = 9.63 \times 10^{-8}$, Spearman $R^2 = -0.88$; Fig. 5). Extracellular sTREM2 was unchanged with increasing concentrations of the isotype-matched control antibody (IgG Mouse; Fig. 5B). The slope for the anti-MS4A4A antibody was significantly different from the control antibody ($P = 0.0018$; Fig. 5B). These results suggest that specifically targeting MS4A4A reduces sTREM2 produce by human macrophages.

To determine whether *MS4A4A* expression is sufficient to change the release of sTREM2 in culture media, we overexpressed human *MS4A4A* with lentivirus in human macrophages. Macrophages were transduced with lentiviral particles containing *MS4A4A* or GFP for 24 hours. Cells and media were collected 48 hours post-transduction. Overexpression of *MS4A4A* WT was sufficient to significantly increase *MS4A4A* mRNA ($P = 0.0043$; Fig. 5C) and MS4A4A protein levels (P <0.0001; Fig. 5D, and Supplementary Fig. S9D) compared to the GFP control-treated macrophages and protein levels. Macrophages overexpressing *MS4A4A* also produced significantly more sTREM2 compared to GFP control-treated cells ($P = 0.028$; Fig. 5E). These results were replicated in four independent experiments and demonstrate that targeting MS4A4A at the molecular or protein level is sufficient to alter sTREM2.

MS4A4A colocalizes with TREM2 in human macrophages

We performed immunocytochemistry studies in human macrophages to determine the cellular localization of MS4A4A and TREM2. MS4A4A has been reported to be expressed on the macrophage plasma membrane (*42*) and on the plasma membrane and intracellular organelles in human mast cells, where it may play a role in protein trafficking (*44*). Consistent with these reports, we found MS4A4A primarily expressed on the plasma membrane of human macrophages, and also intracellularly (Fig. 5F, see white arrows). Intense staining for TREM2 was detected in areas surrounding the nucleus and in discrete regions on the plasma membrane (Fig. 5F, see yellow arrows). To determine specific colocalization of MS4A4A and TREM2, we stained human macrophages with markers for the endoplasmic reticulum (KDEL), Golgi (Giantin), and lipid rafts (Caveolin-1) (Fig. 5G and Supplementary Fig. S9G). MS4A4A and TREM2 colocalized with Caveolin-1 positive lipid-enriched structures on the plasma membrane (Fig. 5G, see white arrowheads). We also detected partial colocalization of TREM2 with both endoplasmic reticulum (KDEL) and Golgi (Giantin). In contrast, MS4A4A did not colocalize with these intracellular markers (Supplementary Fig. S9G). IL-4-mediated stimulation of human macrophages increased expression of both MS4A4A and TREM2 (Supplementary Fig. S9E, F and H), which is consistent with previous reports (*41, 42*).

Discussion

In this study, we performed a genome-wide analysis for genetic modifiers of CSF sTREM2. We observed two independent signals in the *MS4A* gene region that passed genome-wide significance thresholds for association with CSF sTREM2: rs1582763 and rs6591561 (*MS4A4A* p.M159V). Importantly, rs1582763 is associated with elevated CSF sTREM2, reduced

AD risk (*36*), and delayed age-at-onset (*40*). Conversely, rs6591561 (*MS4A4A* p.M159V) is associated with reduced CSF sTREM2, increased AD risk (*36*), and accelerated age-at-onset (*40*). These SNPs modify expression of *MS4A4A* and *MS4A6A* in human blood and brain tissue. The expression of both these genes is highly correlated with *TREM2* gene expression. In human macrophages MS4A4A and TREM2 colocalize on lipid rafts, and pharmacologic or genetic modification of MS4A4A produced changes in sTREM2. Thus, we present genetic, molecular, and functional evidence suggesting that *MS4A* genes, particularly *MS4A4A*, are key modulators of sTREM2.

Our findings that variants in the *MS4A* gene region are associated with CSF sTREM2 provide a putative biological connection between the MS4A family, TREM2, and AD risk. Prior to this study, *TREM2* was directly implicated in AD through the identification of rare variants that increase AD risk (*1, 2*). Thus, TREM2 was thought to only impact a small percentage of AD cases. The putative role of TREM2 in sporadic disease was less clear. Common variants in the *MS4A* locus, including the top SNP we found associated with CSF sTREM2, were reported to be associated with AD risk (rs1582763, OR = 0.898, $P = 1.81 \times 10^{-15}$) (*36*) and age-at-onset (HR = 0.929, $P = 9.26 \times 10^{-9}$) (*40*), but the functional variant and mechanism of action was unknown. Here we show that variants within this AD risk locus modify CSF sTREM2. More specifically our data indicate that AD risk variants are associated with lower CSF sTREM2 and the protective AD variants are associated with elevated CSF sTREM2. This relationship between CSF sTREM2 and AD risk is supported by previous studies suggesting that AD-risk and Nasu-Hakola causal *TREM2* variants are loss-of-function and that increasing TREM2 or activating the TREM2 signaling pathway could offer a new therapeutic approach (*45*).

We present multiple lines of evidence that genes within the *MS4A* cluster modulate sTREM2. First, we observed a strong and consistent association across multiple independent datasets ($P_{meta} = 4.48 \times 10^{-21}$; Fig. 2). For comparison, the effect size of rs1582763 on CSF sTREM2 (OR = 3.34; 95% CI = 2.37 – 4.82) is similar to that of *APOE* genotype on CSF Aβ$_{42}$ (OR = 2.84; 95% CI = 2.30 – 3.50). Second, the *MS4A* association with CSF sTREM2 represents a *trans* effect: *TREM2* is located on chromosome 9 and this *MS4A* locus is on chromosome 11. As our group and others have reported, it is not common to find *trans*-protein quantitative trait loci (pQTLs) with large effect sizes and most *trans*-pQTLs occur in genes which encode receptors or that have other regulatory influences on the protein studied (*46-51*). Thus, we hypothesize that one or more genes within the *MS4A* locus may be key modulators of sTREM2. Third, our functional studies support a relationship between MS4A4A and sTREM2. Using pharmacologic (antibody) or genetic (lentivirus-mediated overexpression) approaches, we targeted MS4A4A and observed changes in sTREM2 produced by human macrophages. Overexpression of *MS4A4A* resulted in elevated extracellular sTREM2 compared to GFP expressing cells, and antibodies targeting MS4A4A led to a dose-dependent reduction of sTREM2 levels.

These results suggest that MS4A4A may be a therapeutic target for AD. MS4A4A colocalizes with TREM2 in human macrophages and both proteins are upregulated in response to IL-4-mediated stimulation. *TREM2* and *MS4A4A* are both highly and specifically expressed in human microglia (*38*). While the function of MS4A4A in the brain is poorly understood, it may be involved in protein trafficking and clathrin-dependent endocytosis (*44*). The finding that modifying MS4A4A affects sTREM2 in human macrophages, support a putative relationship

between these two proteins. Compounds that overexpress MS4A4A or increase MS4A4A activity could modify AD risk through regulation of sTREM2.

The origin and function of sTREM2 is still poorly understood. Most studies investigating the role of TREM2 in AD pathogenesis have focused on characterizing the impact of *TREM2* loss-of-function or rare *TREM2* risk variants. These studies evaluate the impact of TREM2 deficiency (*18, 52*) or haploinsufficiency (*15*) but are unable to draw any conclusions about a potential role for sTREM2 specifically. Some evidence suggests that sTREM2 binds to and sequesters TREM2 ligands, such as apolipoproteins, thus preventing these ligands from activating the TREM2 receptor (*53, 54*). Others propose that sTREM2 works independently of the TREM2 receptor. Treatment with sTREM2 enhanced microglial survival and cytokine release in both WT and *Trem2*$^{-/-}$ cells (*55*). In both WT and *Trem2*$^{-/-}$ primary microglia, the introduction of sTREM2 increased mRNA levels of IL-1β, IL-6, TNF, and IL-10 in a dose-dependent manner, further suggesting that sTREM2 may trigger microglial activation independently of the TREM2 receptor (*55*). These findings demonstrate that further research is necessary to determine the role of sTREM2 in AD. The correlations between CSF sTREM2 and CSF tau ($r = 0.377$, $P = 9.59 \times 10^{-29}$) and CSF ptau ($r = 0.348$, $P = 2.08 \times 10^{-24}$), which have been reported consistently in independent studies (*23, 24, 26*), may provide some clue. There is increasing evidence that microglia play a key role in tau-mediated neurodegeneration in AD and similar tauopathies (*56*). Considering that *TREM2*, *MS4A4A*, and *MS4A6A* are exclusively expressed in microglia in the CNS, our findings add further support to this putative role for microglia in tau-mediated pathogenic mechanisms.

Our Mendelian randomization analyses provide additional support that sTREM2 may be involved in AD pathogenesis. Mendelian randomization uses genetic variation as instrumental

variables to analyze whether a specific variable, which can be an exposure or protein (*57*), is associated with a specific outcome (AD risk in this case) independently of reverse causation or confounding effect. We used, as instrumental variables, the three independent SNPs that showed strong association with sTREM2 and found significant associations for CSF sTREM2 with AD risk, demonstrating for the first time that sTREM2 may actually be involved in AD and is not merely a product of the disease process.

There are some limitations to the present study. The first limitation is that this study only included common variants (MAF > 2%); therefore, we are unable to determine the effect of genes that only harbor low frequency or rare functional variants. Although some studies have linked APOE with TREM2 (*58*), we did not observe an association between *APOE* genotype and CSF sTREM2. Thus, *APOE* does not play a role in modulating sTREM2. However, we cannot yet conclude whether other genes such *TYROBP, ADAM10*, or *ADAM17* play a role in modulating CSF sTREM2. These genes were reported to be involved in *TREM2* cleavage (*21, 22*); therefore, we would expect to find a strong association with CSF sTREM2 in these gene regions. However, the risk variants in *TYROBP* are all rare (*59, 60*) and although there are variants in *ADAM10* known to affect the function or expression of ADAM10, these variants also present low frequency (*61*). By excluding low frequency and rare variants, this study was not suitable to identify genetic signals in these genes. Additional studies of sequence data in a larger sample size will be necessary to determine the role of *TYROBP, ADAM10* or *ADAM17* in CSF sTREM2 modulation.

A second limitation of this study is that we cannot conclude whether other genes in the *MS4A* locus also modulate sTREM2. Our data indicate that *MS4A4A* modulates CSF sTREM2. This is supported by the eQTL analyses, the correlation of *MS4A4A* and *TREM2* mRNA levels,

the identification of a non-synonymous *MS4A4A* variant associated with CSF sTREM2, and our functional studies. Our preliminary functional studies suggest that MS4A6A does not modify sTREM2, but due to limited availability of MS4A6A reagents we cannot definitively rule out a relationship between MS4A6A and sTREM2. Since the GWAS and bioinformatics analyses only consistently supported a putative functional role for *MS4A4A* and *MS4A6A*, we did not investigate the other *MS4A* gene family members; therefore, we cannot conclude that none of the other genes in the *MS4A* region modulate sTREM2.

Another limitation of this study is that we are only beginning to uncover the functional relationship between MS4A4A and sTREM2. The importance of this study resides in the identification and functional validation of the association of *MS4A4A* with sTREM2. This study also begins to provide a functional explanation for the original AD risk association identified in the *MS4A* gene region. Our functional data support the genetic association between *MS4A4A* and sTREM2 as we were able to modify extracellular sTREM2 by modulating *MS4A4A* expression or targeting MS4A4A with specific antibodies. MS4A4A may modify sTREM2 as part of the machinery needed for cleavage, through intracellular TREM2 trafficking, or by binding to extracellular sTREM2. Additional functional studies interrogating all the possible mechanisms will be necessary to address these important questions.

In summary, this study identified for the first time *MS4A4A* as an important player in TREM2 biology, providing a mechanistic explanation of the *MS4A* genetic association with AD risk, supporting the importance of TREM2 biology in AD risk in general, and demonstrating that sTREM2 may be involved in AD pathology.

Materials and Methods

Study Design

The goal of this study was to identify common genetic variants and genes associated with CSF sTREM2. To do this, we used a three-stage genome-wide association study: discovery, replication, and meta-analyses. The discovery phase included 813 individuals from ADNI, and the replication phase included 580 independent samples from six different studies. Meta-analyses were performed using a fixed effect model. Genetic loci that passed the multiple test correction for GWAS ($P < 5 \times 10^{-8}$) were functionally annotated using bioinformatics tools to identify variants and genes most likely driving the GWAS signal.

To validate the genetic findings, we performed cell-based studies modeling the most likely functional genes identified by the functional annotation. Our analyses identified *MS4A4A* and *MS4A6A* as putative functional genes. Our hypothesis is that changes in protein levels or function of the functional gene(s) should also result in changes of sTREM2. These genes, as well as TREM2, are primarily expressed in microglia in the brain. Additionally, *MS4A4A* does not have a mouse ortholog. For this reason, we decided to use primary human macrophages. Transfecting/transducing primary human macrophages is challenging; therefore, we decided to first use antibodies against MS4A4A and MS4A6A as a quick screening method. Primary human macrophages, were treated with the specific MS4A antibodies and sTREM2 was measured in the media by ELISA. Then we transduced primary macrophages using lentivirus to overexpress *MS4A4A*. ELISA was used to measure sTREM2 in the media.

Ethics statement

The Institutional Review Boards of all participating institutions approved the study and research was carried out in accordance with the approved protocols. Written informed consent was obtained from participants or their family members.

Cohort demographics

Cerebrospinal fluid was obtained from the Alzheimer Disease Neuroimaging Initiative (ADNI) and sTREM2 was measured with enzyme-linked immunosorbent assay (ELISA) by a group of investigators at Washington University in St Louis (WashU) and a group at Ludwig-Maximilians-Universität (LMU) München department of Neurology (Munich, Germany). A total of 172 AD, 169 cognitive normal, 183 early Mild Cognitive impairment (eMCI), 221 late-MCI and 68 Significant Memory Concern (SMC) individuals were studied in addition to 38 TREM2 mutation carriers (p.D87N, p.H157Y, p.L211P, p.R47H, and p.R62H). TREM2 risk-variant carriers were excluded from the GWAS analyses. Demographic characteristics of the datasets can be seen in Table 1, Table 2, and Supplementary tables S1 and S3. Additional datasets were used for replication of the top SNPs, demographic characteristics are shown in Supplementary Table S3. Data were obtained from The Charles F. and Joanne Knight Alzheimer's Disease Research Center (Knight ADRC), GHPH, GHDEM, SPIN, Clinic-IDIBAPS, and The Dominantly Inherited Alzheimer Network (DIAN) (*24-26, 32*).

RNA-seq data were obtained from four independent cohorts as described previously (*37*). The Knight ADRC provided data from 28 brains taken from neuropathologically-confirmed late-onset AD, an additional 10 AD cases with known *TREM2* mutations, and 14 cognitively healthy controls. DIAN provided data from 19 autosomal-dominantly inherited early-onset AD. RNA extraction and sequencing methods were described previously (*37*). RNA-seq data were obtained

from the Mayo Clinic Brain Bank via the AMP-AD Knowledge portal (https://www.synapse.org; synapse ID = 5550404; accessed January 2017), the RNA extraction and sequencing methods were described previously (*62*). The AMP-AD portal was also utilized to obtain RNA-seq data from the Mount Sinai Brain Bank (https://www.synapse.org; synapse ID = 3157743; accessed January 2017) which included data from 1,030 samples collected from four brain regions taken from 300 individuals, as described previously (*63*). QC and processing were performed as previously described and additional information is in the supplementary methods (*37*).

ELISA for CSF sTREM2

CSF samples were obtained and measured separately by investigators at WashU as described previously (24) and LMU. CSF sTREM2 measurements from LMU were done with an ELISA based on the MSD platform and is comprehensively described in previous publications (*13, 25, 26*). Measured CSF values were corrected by plate-specific correction factors obtained by each group and the corrected values were used for analyses. Pearson's correlation was used to compare the corrected sTREM2 values between the two different ELISA methods in overlapping samples ($N = 980$, $r = 0.834$, $P = 2.2 \times 10^{-254}$). Detailed description of the ELISA and QC can be found in the supplementary methods.

When more than one experiment was performed in vitro to measure sTREM2 in cell culture media results are expressed as normalized values on untreated controls.

Genotyping and imputation

Samples were genotyped with the Illumina 610 or Omniexpress chip. Stringent quality control (QC) criteria were applied to each genotyping array separately before combining genotype data. The minimum call rate for single nucleotide polymorphisms (SNPs) and individuals was 98% and autosomal SNPs not in Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$) were excluded. X-chromosome SNPs were analyzed to verify gender identification. Unanticipated duplicates and cryptic relatedness (Pihat ≥0.25) among samples were tested by pairwise genome-wide estimates of proportion identity-by-descent, and when a pair of identical or related samples was identified, the sample from Knight-ADRC or with a higher number of variants that passed QC was prioritized. EIGENSTRAT (*64*) was used to calculate principal components. *APOE* ε2, ε3, and ε4 isoforms were determined by genotyping rs7412 and rs429358 using Taqman genotyping technology as previously described (*65-67*). The 1000 Genomes Project Phase 3 data (October 2014), SHAPEIT v2.r837 (*68*), and IMPUTE2 v2.3.2 (*69*) were used for phasing and imputation. Individual genotypes imputed with probability < 0.90 were set to missing and imputed genotypes with probability ≥0.90 were analyzed as fully observed. Genotyped and imputed variants with MAF < 0.02 or IMPUTE2 information score < 0.30 were excluded, leaving 7,320,475 variants for analyses.

Statistical methods

Statistical analyses and data visualization were performed in R v3.4.0 (*70*), PLINK v1.9 (*71*), and LocusZoom v1.3 (*72*). Corrected raw values for CSF sTREM2 were normally distributed, so two-sample t-tests were used to compare CSF levels between males and females. Kruskal-Wallis with *post-hoc* Dunn's test was used for multiple group comparisons of the *TREM2* mutation carriers and non-carriers. Pearson's product moment correlation coefficient

was used for correlation analyses. Single-variant associations with CSF sTREM2 were tested

using the additive linear regression model in PLINK v1.9 (*71*). Covariates included age at time

of LP, sex, and the first two principal component factors to account for population structure. To

determine whether there were spurious associations due to differences in assay, we tested the

CSF sTREM2 measures from WashU and LMU and in both cases a genome-wide signal in the

MS4A cluster were found (See Figure 1 and Supplementary Figure S3). The genomic inflation

factor was $\lambda < 1.008$ for all of the genetic analyses. Statistical significance for single-variant

analyses was selected based on the commonly used threshold estimated from Bonferroni

correction of the likely number of independent tests in genome-wide analyses ($P < 5 \times 10^{-8}$).

Meta-analyses were performed using a fixed effect model in METAL (*73*).

To identify additional independent genetic signals, conditional analyses were conducted

by adding the SNP with the smallest P value as a covariate into the default regression model and

testing all remaining regional SNPs for association. The dataset was stratified by the most

recently reported case status and the genetic signals from the joint dataset were tested to

determine whether the genetic associations were driven by cases or controls. To determine

whether the identified genetic association within the joint dataset was sex-specific, the dataset

was stratified by sex and each sex was tested for the association. Results from the sex-specific

analyses were verified in a subset of age-matched samples ($N = 367$ each for males and females,

Table S1).

Gene expression was inferred using Salmon v0.7.2 (*74*) transcript expression

quantification of the coding transcripts of *Homo sapiens* included in the GENCODE reference

genome (GRCh37.75). Gene counts data were transformed to stabilize expression variances

along the range of mean values and normalized according to library size using the variance

stabilizing transformation (VST) function in DESeq2 (*75*). Correlations between expression of *TREM2* and each member of the *MS4A* gene family, with an average TPM > 1, were tested in each independent RNA-seq cohort.

Differences in mRNA levels, and sTREM2 levels for the cell-based studies were analyzing using the non-parametric Mann-Whitney test or Kruskal Wallis H test for multi-group comparisons.

Bioinformatics annotation

All variants below the threshold for suggestive significance ($P < 1 \times 10^{-5}$) were taken forward for functional annotation using ANNOVAR version 2015-06-17 (*76*) and examined for potential regulatory functions using HaploReg v4.1 (*77*) and RegulomeDB v1.1 (*78*). Search tools from publicly available databases, Genotype-Tissue Expression (GTEx) Analysis V7 (*79*), the Brain eQTL Almanac (Braineac) (*80*), and the Westra Blood eQTL browser (*48*), were utilized to determine if significant SNPs were reported eQTLs. The Encyclopedia of DNA elements (ENCODE, https://www.encodeproject.org/) was queried, filtering for brain tissue in *Homo sapiens*, to examine DNA elements affected by the associated SNPs. The Ensembl Variant Effect Predictor (VEP, http://grch37.ensembl.org/Homo_sapiens/Tools/VEP) (*81*) was queried for the predicted effects of identified coding variants, including any applicable SIFT and PolyPhen scores.

Mendelian Randomization

We used the R package MendelianRandomization (*82*) which includes three primary methods: inverse-variance weighted (IVW), median-based (simple or weighted), and MR-Egger. The MR-Egger method has an option to use robust regression instead of standard regression or a

penalized option that downweights the contribution of genetic variants with outlying (heterogeneous) causal estimates. We tested all three primary methods and the different options for the MR-Egger method.

The SNPs selected for the analysis were the two independent SNPs (rs1582763, and rs659156) from chromosome 11 as well as the top SNP in Chr 6 (rs28385608). This SNP was included because the locus was near significant in the gene-based analyses and this SNP had the smallest P value within the region in the single-variant analysis. The beta-coefficients and standard errors for the three selected SNPs from this study were used as input with summary statistics from a large GWAS for AD risk (*36*).

Human macrophages cultures

Peripheral blood mononuclear cells (PBMCs) were purified from human blood on Ficoll-Paque PLUS density gradient (Amersham Biosciences, Piscataway, NJ). To generate macrophages, PBMCs were cultured in 6-well culture plates (3×10^6 cell/well) in RPMI-1640 without fetal bovine serum. After 2 hours of culture PBMCs were washed twice with PBS 1X and cultured in RPMI supplemented with 50 ng/ml MCSF for 7 days at 37^0C with 5% CO_2. Supernatants from macrophage cultures were collected at different time points, centrifuged at 21000g for 15 min and filtered through a 0.22 µm filter to remove all cells and membrane debris. Macrophage supernatants were then frozen in aliquots at –80°C until used in the ELISA for human sTREM2.

In vitro experiments with anti MS4A antibodies

Human macrophages after 7 days in culture were incubated with 10, 5 and 1 µg/ml of anti-MS4A4A Biolegend (cat #372502 (clone 5C12)); or anti-MS4A6A (Invitrogen cat #PA5-

72732) antibodies in the presence or absence of IL-4 (Peprotech; 15 ng/ml) in culture. Mouse

IgG Isotype (Thermo Fisher cat: #10400C) was used as control. After 48 hours cell culture media

were collected, centrifuged at 21,000g for 15 min and filtered through a 0.22 μm filter to remove

all cells and membrane debris. Macrophage supernatants were then frozen in aliquots at –80°C

until used in the ELISA for human sTREM2.

Lentivirus vector preparation

Lentiviral vectors encoding human MS4A4A WT cDNA (NM_148975) Myc-DDK-

tagged (RC204646L1) and GFP (PS100071) were obtained from Origene and produced as

previously described (*83, 84*). All constructs were verified by Sanger sequencing. To generate

viral particles, HEK-293T cells were transfected using calcium phosphate with the packaging

construct and packaging plasmids Gag-Pol, Rev, and VSV-G. Supernatant containing lentiviral

particles were collected at 48 and 72 h after transfection. Viral supernatant was collected

according to previously published protocols (*84*). Transduction was carried out with different

volumes 1, 5, 10, 50, 100 and 500 uL of supernatant in HEK-293T cells for 24 hours.

Transduction efficiency, defined as percentage of transduced cells expressing GFP, was assessed

qualitatively using an Olympus IX81 fluorescence microscope (Center Valley, PA).

Transductions of Human macrophages

Cultured monocyte-derived macrophages were transduced with lentiviral vectors. The

highest transduction efficiency (60-70%) was reached with 400 uL of unconcentrated virus at

five days in culture for human macrophages. Macrophages were transduced with lentivirus

vector for 24 hours, rinsed and fresh media added. Cells and media were collected 48 hours later

for sTREM2 ELISA and qPCR analyses.

Quantitative PCR

The effect of overexpression was measured by real-time qPCR analysis using probe

specific for *MS4A4A*. Macrophages transduced with the lentiviral particles were collected in

TRIzol reagent and total RNA was extracted using the RNeasy Mini Kit (Qiagen). cDNA was

prepared from the total RNA using a High-Capacity cDNA Reverse Transcription Kit

(Thermo Fisher). Gene expression levels were analyzed by real-time PCR using TaqMan

assays for *MS4A4A* (Hs00254780_m1) and *GAPDH* (Hs02758991_g1) on a QuantStudio 12K

Flex Real-Time PCR System (Thermo Fisher). To avoid amplification interference,

expression assays were run in separate wells from the housekeeping gene *GAPDH*. Real-time

data were analyzed by the comparative C_T method. Average C_T values for each sample were

normalized to the average C_T values for the housekeeping gene *GAPDH*. The resulting value

was corrected for assay efficiency. Samples with a standard error of 20% or less were

analyzed.

Immunocytochemistry and cell imaging

Cells were fixed for 15-20 min at room temperature (RT) in 4 % (w/v) PFA, 4 % (w/v)

sucrose, 20 mM NaOH and 5 mM MgCl2 in PBS, pH 7.4. For intracellular staining, cells were

permeabilized and blocked for 60 min at RT in 5% Horse serum, 0.1% Saponin in PBS and

incubated at 4 °C overnight with human anti-TREM2 (R&D, AF1828) and human anti-MS4A4A

(Biolegend, clone 5C12) primary antibodies diluted in PBS, 5% Horse serum. Samples were

analyzed with an Olympus FV1200 scanning confocal microscope and a Zeiss LSM880 laser scanning confocal microscope (Carl Zeiss Inc, Thornwood, NY) equipped with 63X, 1.4 NA Zeiss Plan Apochromat oil objective. The Olympus FV1200 laser scanning confocal: (Olympus-America, Inc, Waltham, MA) equipped with five detectors (2 spectral and 1filter based) and two GaAsP PMTs. Images were captured using an OLYMPUS PlanApoN 60X, 1.4 NA super corrected oil objective [PlanApoN 60X OIL SC] objectives. The 405nm, 488nm, 559nm diode lasers and 635nm HeNe (helium neon) lasers were utilized with an optimal pinhole of 1 airy unit to acquire images. Images were finally processed with ImageJ software. DF ZEN 2.1 black edition software was used to obtain Z-stacks through the entire height of the cells with confocal Z-slices of 1.5µm (63X) and an interval of 0.5 µm. Some of the images were acquired using a Nikon Eclipse 90i fluorescent and bright field microscope and analyzed for quantitation with the Metamorph 7.7 software.

REFERENCES

1. R. Guerreiro, A. Wojtas, J. Bras, M. Carrasquillo, E. Rogaeva, E. Majounie, C. Cruchaga, C. Sassi, J. S. K. Kauwe, M. K. Lupton, M. Ryten, K. Brown, J. Lowe, P. G. Ridge, M. B. Hammer, Y. Wakutani, P. Proitsi, S. Newhouse, E. Lohmann, N. Erginel-Unaltuna, C. Medway, H. Hanagasi, C. Troakes, H. Gurvit, B. Bilgic, S. Al-Sarraj, B. Benitez, B. Cooper, D. Carrell, M. Emre, F. G. Zou, L. Ma, M. E. Murray, D. W. Dickson, S. Younkin, L. Hazrati, R. C. Petersen, C. D. Corcoran, Y. F. Cai, C. Oliveira, M. H. Ribeiro, I. Santana, J. T. Tschanz, J. R. Gibbs, M. C. Norton, I. Kloszewska, P. Mecocci, H. Soininen, M. Tsolaki, B. Vellas, R. G. Munger, D. M. A. Mann, S. Pickering-Brown, S. Lovestone, J. Beck, S. Mead, J. Collinge, L. Parsons, J. Pocock, J. C. Morris, T. Revesz, T. Lashley, N. C. Fox, M. N. Rossor, B. Grenier-Boley, C. Bellenguez, V. Moskvina, R. Sims, D. Harold, J. Williams, J. C. Lambert, P. Amouyel, N. Graff-Radford, A. Goate, R. Rademakers, K. Morgan, J. Powell, P. St George-Hyslop, A. Singleton, J. Hardy, E. Consortium, G. Consortium, U. Consortium, A. G. A. Grp, TREM2 Variants in Alzheimer's Disease. *New Engl J Med* **368**, 117-127 (2013).

2. T. Jonsson, H. Stefansson, S. Steinberg, I. Jonsdottir, P. V. Jonsson, J. Snaedal, S. Bjornsson, J. Huttenlocher, A. I. Levey, J. J. Lah, D. Rujescu, H. Hampel, I. Giegling, O. A. Andreassen, K. Engedal, I. Ulstein, S. Djurovic, C. Ibrahim-Verbaas, A. Hofman, M. A. Ikram, C. M. van Duijn, U. Thorsteinsdottir, A. Kong, K. Stefansson, Variant of TREM2 Associated with the Risk of Alzheimer's Disease. *New Engl J Med* **368**, 107-116 (2013).

3. E. O. Luis, S. Ortega-Cubero, I. Lamet, C. Razquin, C. Cruchaga, B. A. Benitez, E. Lorenzo, J. Irigoyen, I. Alzheimer's Disease Neuroimaging, M. A. Pastor, P. Pastor, Frontobasal gray matter loss is associated with the TREM2 p.R47H variant. *Neurobiol Aging* **35**, 2681-2690 (2014).

4. R. Guerreiro, B. Bilgic, G. Guven, J. Bras, J. Rohrer, E. Lohmann, H. Hanagasi, H. Gurvit, M. Emre, A novel compound heterozygous mutation in TREM2 found in a Turkish frontotemporal dementia-like family. *Neurobiology of Aging* **34**, (2013).

5. B. A. Benitez, B. Cooper, P. Pastor, S. C. Jin, E. Lorenzo, S. Cervantes, C. Cruchaga, TREM2 is associated with the risk of Alzheimer's disease in Spanish population. *Neurobiol Aging* **34**, 1711 e1715-1717 (2013).

6. S. C. Jin, B. A. Benitez, C. M. Karch, B. Cooper, T. Skorupa, D. Carrell, J. B. Norton, S. Hsu, O. Harari, Y. Cai, S. Bertelsen, A. M. Goate, C. Cruchaga, Coding variants in TREM2 increase risk for Alzheimer's disease. *Hum Mol Genet* **23**, 5838-5846 (2014).

7. S. C. Jin, M. M. Carrasquillo, B. A. Benitez, T. Skorupa, D. Carrell, D. Patel, S. Lincoln, S. Krishnan, M. Kachadoorian, C. Reitz, R. Mayeux, T. S. Wingo, J. J. Lah, A. I. Levey, J. Murrell, H. Hendrie, T. Foroud, N. R. Graff-Radford, A. M. Goate, C. Cruchaga, N. Ertekin-Taner, TREM2 is associated with increased risk for Alzheimer's disease in African Americans. *Mol Neurodegener* **10**, 19 (2015).

8. S. E. Hickman, J. El Khoury, TREM2 and the neuroimmunology of Alzheimer's disease. *Biochem Pharmacol* **88**, 495-498 (2014).

9. Y. Wang, M. Cella, K. Mallinson, J. D. Ulrich, K. L. Young, M. L. Robinette, S. Gilfillan, G. M. Krishnan, S. Sudhakar, B. H. Zinselmeyer, D. M. Holtzman, J. R. Cirrito, M. Colonna, TREM2 lipid sensing sustains the microglial response in an Alzheimer's disease model. *Cell* **160**, 1061-1071 (2015).

10. J. D. Ulrich, T. K. Ulland, M. Colonna, D. M. Holtzman, Elucidating the Role of TREM2 in Alzheimer's Disease. *Neuron* **94**, 237-248 (2017).

11. C. Cantoni, B. Bollman, D. Licastro, M. Xie, R. Mikesell, R. Schmidt, C. M. Yuede, D. Galimberti, G. Olivecrona, R. S. Klein, A. H. Cross, K. Otero, L. Piccio, TREM2 regulates microglial cell activation in response to demyelination in vivo. *Acta Neuropathol* **129**, 429-447 (2015).

12. Y. Deming, Z. Li, B. A. Benitez, C. Cruchaga, Triggering receptor expressed on myeloid cells 2 (TREM2): a potential therapeutic target for Alzheimer disease? *Expert Opin Ther Targets* **22**, 587-598 (2018).

13. G. Kleinberger, Y. Yamanishi, M. Suarez-Calvet, E. Czirr, E. Lohmann, E. Cuyvers, H. Struyfs, N. Pettkus, A. Wenninger-Weinzierl, F. Mazaheri, S. Tahirovic, A. Lleo, D. Alcolea, J. Fortea, M. Willem, S. Lammich, J. L. Molinuevo, R. Sanchez-Valle, A. Antonell, A. Ramirez, M. T. Heneka, K. Sleegers, J. van der Zee, J. J. Martin, S. Engelborghs, A. Demirtas-Tatlidede, H. Zetterberg, C. Van Broeckhoven, H. Gurvit, T. Wyss-Coray, J. Hardy, M. Colonna, C. Haass, TREM2 mutations implicated in neurodegeneration impair cell surface transport and phagocytosis. *Sci Transl Med* **6**, 243ra286 (2014).

14. C. Y. D. Lee, A. Daggett, X. Gu, L. L. Jiang, P. Langfelder, X. Li, N. Wang, Y. Zhao, C. S. Park, Y. Cooper, I. Ferando, I. Mody, G. Coppola, H. Xu, X. W. Yang, Elevated TREM2 Gene Dosage Reprograms Microglia Responsivity and Ameliorates Pathological Phenotypes in Alzheimer's Disease Models. *Neuron* **97**, 1032-1048 e1035 (2018).

15.   J. D. Ulrich, M. B. Finn, Y. Wang, A. Shen, T. E. Mahan, H. Jiang, F. R. Stewart, L. Piccio, M. Colonna, D. M. Holtzman, Altered microglial response to Abeta plaques in APPPS1-21 mice heterozygous for TREM2. *Mol Neurodegener* **9**, 20 (2014).

16.   Y. Wang, T. K. Ulland, J. D. Ulrich, W. Song, J. A. Tzaferis, J. T. Hole, P. Yuan, T. E. Mahan, Y. Shi, S. Gilfillan, M. Cella, J. Grutzendler, R. B. DeMattos, J. R. Cirrito, D. M. Holtzman, M. Colonna, TREM2-mediated early microglial response limits diffusion and toxicity of amyloid plaques. *J Exp Med* **213**, 667-675 (2016).

17.   P. Yuan, C. Condello, C. D. Keene, Y. Wang, T. D. Bird, S. M. Paul, W. Luo, M. Colonna, D. Baddeley, J. Grutzendler, TREM2 Haplodeficiency in Mice and Humans Impairs the Microglia Barrier Function Leading to Decreased Amyloid Compaction and Severe Axonal Dystrophy. *Neuron* **92**, 252-264 (2016).

18.   C. E. G. Leyns, J. D. Ulrich, M. B. Finn, F. R. Stewart, L. J. Koscal, J. Remolina Serrano, G. O. Robinson, E. Anderson, M. Colonna, D. M. Holtzman, TREM2 deficiency attenuates neuroinflammation and protects against neurodegeneration in a mouse model of tauopathy. *Proc Natl Acad Sci U S A* **114**, 11524-11529 (2017).

19.   S. M. Bemiller, T. J. McCray, K. Allan, S. V. Formica, G. Xu, G. Wilson, O. N. Kokiko-Cochran, S. D. Crish, C. A. Lasagna-Reeves, R. M. Ransohoff, G. E. Landreth, B. T. Lamb, TREM2 deficiency exacerbates tau pathology through dysregulated kinase signaling in a mouse model of tauopathy. *Mol Neurodegener* **12**, 74 (2017).

20.   P. Wunderlich, K. Glebov, N. Kemmerling, N. T. Tien, H. Neumann, J. Walter, Sequential proteolytic processing of the triggering receptor expressed on myeloid cells-2 (TREM2) protein by ectodomain shedding and gamma-secretase-dependent intramembranous cleavage. *J Biol Chem* **288**, 33027-33036 (2013).

21.   K. Schlepckow, G. Kleinberger, A. Fukumori, R. Feederle, S. F. Lichtenthaler, H. Steiner, C. Haass, An Alzheimer-associated TREM2 variant occurs at the ADAM cleavage site and affects shedding and phagocytic function. *EMBO Mol Med* **9**, 1356-1365 (2017).

22.   P. Thornton, J. Sevalle, M. J. Deery, G. Fraser, Y. Zhou, S. Stahl, E. H. Franssen, R. B. Dodd, S. Qamar, B. Gomez Perez-Nievas, L. S. Nicol, S. Eketjall, J. Revell, C. Jones, A. Billinton, P. H. St George-Hyslop, I. Chessell, D. C. Crowther, TREM2 shedding by cleavage at the H157-S158 bond is accelerated for the Alzheimer's disease-associated H157Y variant. *EMBO Mol Med* **9**, 1366-1378 (2017).

23.   A. Heslegrave, W. Heywood, R. Paterson, N. Magdalinou, J. Svensson, P. Johansson, A. Ohrfelt, K. Blennow, J. Hardy, J. Schott, K. Mills, H. Zetterberg, Increased cerebrospinal fluid soluble TREM2 concentration in Alzheimer's disease. *Mol Neurodegener* **11**, 3 (2016).

24.   L. Piccio, Y. Deming, J. L. Del-Aguila, L. Ghezzi, D. M. Holtzman, A. M. Fagan, C. Fenoglio, D. Galimberti, B. Borroni, C. Cruchaga, Cerebrospinal fluid soluble TREM2 is higher in Alzheimer disease and associated with mutation status. *Acta Neuropathol* **131**, 925-933 (2016).

25.   M. Suarez-Calvet, M. A. Araque Caballero, G. Kleinberger, R. J. Bateman, A. M. Fagan, J. C. Morris, J. Levin, A. Danek, M. Ewers, C. Haass, N. Dominantly Inherited Alzheimer, Early changes in CSF sTREM2 in dominantly inherited Alzheimer's disease occur after amyloid deposition and neuronal injury. *Sci Transl Med* **8**, 369ra178 (2016).

26.   M. Suarez-Calvet, G. Kleinberger, M. A. Araque Caballero, M. Brendel, A. Rominger, D. Alcolea, J. Fortea, A. Lleo, R. Blesa, J. D. Gispert, R. Sanchez-Valle, A. Antonell, L. Rami, J. L. Molinuevo, F. Brosseron, A. Traschutz, M. T. Heneka, H. Struyfs, S. Engelborghs, K. Sleegers, C. Van Broeckhoven, H. Zetterberg, B. Nellgard, K. Blennow, A. Crispin, M. Ewers, C. Haass, sTREM2 cerebrospinal fluid levels are a potential biomarker for microglia activity in early-stage Alzheimer's disease and associate with neuronal injury markers. *EMBO Mol Med* **8**, 466-476 (2016).

27.   L. Piccio, C. Buonsanti, M. Cella, I. Tassi, R. E. Schmidt, C. Fenoglio, J. Rinker, 2nd, R. T. Naismith, P. Panina-Bordignon, N. Passini, D. Galimberti, E. Scarpini, M. Colonna, A. H. Cross, Identification of soluble TREM-2 in the cerebrospinal fluid and its association with multiple sclerosis and CNS inflammation. *Brain* **131**, 3081-3091 (2008).

28.   K. Henjum, I. S. Almdahl, V. Arskog, L. Minthon, O. Hansson, T. Fladby, L. N. Nilsson, Cerebrospinal fluid soluble TREM2 in aging and Alzheimer's disease. *Alzheimers Res Ther* **8**, 17 (2016).

29.   B. Benyamin, A. E. Mcrae, G. Zhu, S. Gordon, A. K. Henders, A. Palotie, L. Peltonen, N. G. Martin, G. W. Montgomery, J. B. Whitfield, P. M. Visscher, Variants in TF and HFE Explain similar to 40% of Genetic Variation in Serum-Transferrin Levels. *Am J Hum Genet* **84**, 60-65 (2009).

30.   K. Panoutsopoulou, S. Thiagarajah, A. Day-Williams, L. Southam, K. Hatzikotoulas, A. Matchan, M. Doherty, J. M. Wilkinson, E. Zeggini, a. Consortium, Powerful Detection of Osteoarthritis Susceptibility

Loci by Comprehensive Examination of Clinically Important Endophenotypes. *Osteoarthr Cartilage* **22**, S232-S232 (2014).

31. Y. Deming, Z. Li, M. Kapoor, O. Harari, J. L. Del-Aguila, K. Black, D. Carrell, Y. Cai, M. V. Fernandez, J. Budde, S. Ma, B. Saef, B. Howells, K. L. Huang, S. Bertelsen, A. M. Fagan, D. M. Holtzman, J. C. Morris, S. Kim, A. J. Saykin, P. L. De Jager, M. Albert, A. Moghekar, R. O'Brien, M. Riemenschneider, R. C. Petersen, K. Blennow, H. Zetterberg, L. Minthon, V. M. Van Deerlin, V. M. Lee, L. M. Shaw, J. Q. Trojanowski, G. Schellenberg, J. L. Haines, R. Mayeux, M. A. Pericak-Vance, L. A. Farrer, E. R. Peskind, G. Li, A. F. Di Narzo, I. Alzheimer's Disease Neuroimaging, C. Alzheimer Disease Genetic, J. S. Kauwe, A. M. Goate, C. Cruchaga, Genome-wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers. *Acta Neuropathol* **133**, 839-856 (2017).

32. P. Johansson, N. Mattsson, O. Hansson, A. Wallin, J. O. Johansson, U. Andreasson, H. Zetterberg, K. Blennow, J. Svensson, Cerebrospinal fluid biomarkers for Alzheimer's disease: diagnostic performance in a homogeneous mono-center population. *J Alzheimers Dis* **24**, 537-546 (2011).

33. K. D. Coon, A. J. Myers, D. W. Craig, J. A. Webster, J. V. Pearson, D. H. Lince, V. L. Zismann, T. G. Beach, D. Leung, L. Bryden, R. F. Halperin, L. Marlowe, M. Kaleem, D. G. Walker, R. Ravid, C. B. Heward, J. Rogers, A. Papassotiropoulos, E. M. Reiman, J. Hardy, D. A. Stephan, A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. *J Clin Psychiatry* **68**, 613-618 (2007).

34. C. Cruchaga, J. S. Kauwe, O. Harari, S. C. Jin, Y. Cai, C. M. Karch, B. A. Benitez, A. T. Jeng, T. Skorupa, D. Carrell, S. Bertelsen, M. Bailey, D. McKean, J. M. Shulman, P. L. De Jager, L. Chibnik, D. A. Bennett, S. E. Arnold, D. Harold, R. Sims, A. Gerrish, J. Williams, V. M. Van Deerlin, V. M. Lee, L. M. Shaw, J. Q. Trojanowski, J. L. Haines, R. Mayeux, M. A. Pericak-Vance, L. A. Farrer, G. D. Schellenberg, E. R. Peskind, D. Galasko, A. M. Fagan, D. M. Holtzman, J. C. Morris, G. Consortium, I. Alzheimer's Disease Neuroimaging, C. Alzheimer Disease Genetic, A. M. Goate, GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. *Neuron* **78**, 256-268 (2013).

35. R. Lautner, S. Palmqvist, N. Mattsson, U. Andreasson, A. Wallin, E. Palsson, J. Jakobsson, S. K. Herukka, R. Owenius, B. Olsson, H. Hampel, D. Rujescu, M. Ewers, M. Landen, L. Minthon, K. Blennow, H. Zetterberg, O. Hansson, I. Alzheimer's Disease Neuroimaging, Apolipoprotein E genotype and the diagnostic accuracy of cerebrospinal fluid biomarkers for Alzheimer disease. *JAMA Psychiatry* **71**, 1183-1191 (2014).

36. J. C. Lambert, C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, A. L. DeStafano, J. C. Bis, G. W. Beecham, B. Grenier-Boley, G. Russo, T. A. Thorton-Wells, N. Jones, A. V. Smith, V. Chouraki, C. Thomas, M. A. Ikram, D. Zelenika, B. N. Vardarajan, Y. Kamatani, C. F. Lin, A. Gerrish, H. Schmidt, B. Kunkle, M. L. Dunstan, A. Ruiz, M. T. Bihoreau, S. H. Choi, C. Reitz, F. Pasquier, C. Cruchaga, D. Craig, N. Amin, C. Berr, O. L. Lopez, P. L. De Jager, V. Deramecourt, J. A. Johnston, D. Evans, S. Lovestone, L. Letenneur, F. J. Moron, D. C. Rubinsztein, G. Eiriksdottir, K. Sleegers, A. M. Goate, N. Fievet, M. W. Huentelman, M. Gill, K. Brown, M. I. Kamboh, L. Keller, P. Barberger-Gateau, B. McGuiness, E. B. Larson, R. Green, A. J. Myers, C. Dufouil, S. Todd, D. Wallon, S. Love, E. Rogaeva, J. Gallacher, P. St George-Hyslop, J. Clarimon, A. Lleo, A. Bayer, D. W. Tsuang, L. Yu, M. Tsolaki, P. Bossu, G. Spalletta, P. Proitsi, J. Collinge, S. Sorbi, F. Sanchez-Garcia, N. C. Fox, J. Hardy, M. C. Deniz Naranjo, P. Bosco, R. Clarke, C. Brayne, D. Galimberti, M. Mancuso, F. Matthews, I. European Alzheimer's Disease, Genetic, D. Environmental Risk in Alzheimer's, C. Alzheimer's Disease Genetic, H. Cohorts for, E. Aging Research in Genomic, S. Moebus, P. Mecocci, M. Del Zompo, W. Maier, H. Hampel, A. Pilotto, M. Bullido, F. Panza, P. Caffarra, B. Nacmias, J. R. Gilbert, M. Mayhaus, L. Lannefelt, H. Hakonarson, S. Pichler, M. M. Carrasquillo, M. Ingelsson, D. Beekly, V. Alvarez, F. Zou, O. Valladares, S. G. Younkin, E. Coto, K. L. Hamilton-Nelson, W. Gu, C. Razquin, P. Pastor, I. Mateo, M. J. Owen, K. M. Faber, P. V. Jonsson, O. Combarros, M. C. O'Donovan, L. B. Cantwell, H. Soininen, D. Blacker, S. Mead, T. H. Mosley, Jr., D. A. Bennett, T. B. Harris, L. Fratiglioni, C. Holmes, R. F. de Bruijn, P. Passmore, T. J. Montine, K. Bettens, J. I. Rotter, A. Brice, K. Morgan, T. M. Foroud, W. A. Kukull, D. Hannequin, J. F. Powell, M. A. Nalls, K. Ritchie, K. L. Lunetta, J. S. Kauwe, E. Boerwinkle, M. Riemenschneider, M. Boada, M. Hiltuenen, E. R. Martin, R. Schmidt, D. Rujescu, L. S. Wang, J. F. Dartigues, R. Mayeux, C. Tzourio, A. Hofman, M. M. Nothen, C. Graff, B. M. Psaty, L. Jones, J. L. Haines, P. A. Holmans, M. Lathrop, M. A. Pericak-Vance, L. J. Launer, L. A. Farrer, C. M. van Duijn, C. Van Broeckhoven, V. Moskvina, S. Seshadri, J. Williams, G. D. Schellenberg, P. Amouyel, Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-1458 (2013).

37.   Z. Li, J. L. Del Aguila, U. Dube, J. Budde, R. Martinez, K. Black, Q. Xiao, N. J. Cairns, J. D. Dougherty, J.-M. Lee, J. C. Morris, R. J. Bateman, C. M. Karch, C. Cruchaga, O. Harari, Genetic variants associated with Alzheimer's disease confer different cerebral cortex cell-type population structure. *bioRxiv*, (2018).

38.   Y. Zhang, K. Chen, S. A. Sloan, M. L. Bennett, A. R. Scholze, S. O'Keeffe, H. P. Phatnani, P. Guarnieri, C. Caneda, N. Ruderisch, S. Deng, S. A. Liddelow, C. Zhang, R. Daneman, T. Maniatis, B. A. Barres, J. Q. Wu, An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci* **34**, 11929-11947 (2014).

39.   Y. Zhang, S. A. Sloan, L. E. Clarke, C. Caneda, C. A. Plaza, P. D. Blumenthal, H. Vogel, G. K. Steinberg, M. S. Edwards, G. Li, J. A. Duncan, 3rd, S. H. Cheshier, L. M. Shuer, E. F. Chang, G. A. Grant, M. G. Gephart, B. A. Barres, Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron* **89**, 37-53 (2016).

40.   K. L. Huang, E. Marcora, A. A. Pimenova, A. F. Di Narzo, M. Kapoor, S. C. Jin, O. Harari, S. Bertelsen, B. P. Fairfax, J. Czajkowski, V. Chouraki, B. Grenier-Boley, C. Bellenguez, Y. Deming, A. McKenzie, T. Raj, A. E. Renton, J. Budde, A. Smith, A. Fitzpatrick, J. C. Bis, A. DeStefano, H. H. H. Adams, M. A. Ikram, S. van der Lee, J. L. Del-Aguila, M. V. Fernandez, L. Ibanez, P. International Genomics of Alzheimer's, I. Alzheimer's Disease Neuroimaging, R. Sims, V. Escott-Price, R. Mayeux, J. L. Haines, L. A. Farrer, M. A. Pericak-Vance, J. C. Lambert, C. van Duijn, L. Launer, S. Seshadri, J. Williams, P. Amouyel, G. D. Schellenberg, B. Zhang, I. Borecki, J. S. K. Kauwe, C. Cruchaga, K. Hao, A. M. Goate, A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. *Nat Neurosci* **20**, 1052-1061 (2017).

41.   M. Cella, C. Buonsanti, C. Strader, T. Kondo, A. Salmaggi, M. Colonna, Impaired differentiation of osteoclasts in TREM-2-deficient individuals. *J Exp Med* **198**, 645-651 (2003).

42.   R. Sanyal, M. J. Polyak, J. Zuccolo, M. Puri, L. Deng, L. Roberts, A. Zuba, J. Storek, J. M. Luider, E. M. Sundberg, A. Mansoor, E. Baigorri, M. P. Chu, A. R. Belch, L. M. Pilarski, J. P. Deans, MS4A4A: a novel cell surface marker for M2 macrophages and plasma cells. *Immunol Cell Biol* **95**, 611-619 (2017).

43.   I. R. Turnbull, S. Gilfillan, M. Cella, T. Aoshi, M. Miller, L. Piccio, M. Hernandez, M. Colonna, Cutting edge: TREM-2 attenuates macrophage activation. *J Immunol* **177**, 3520-3524 (2006).

44.   G. Cruse, M. A. Beaven, S. C. Music, P. Bradding, A. M. Gilfillan, D. D. Metcalfe, The CD20 homologue MS4A4 directs trafficking of KIT toward clathrin-independent endocytosis pathways and thus regulates receptor signaling and recycling. *Mol Biol Cell* **26**, 1711-1727 (2015).

45.   S. Carmona, K. Zahs, E. Wu, K. Dakin, J. Bras, R. Guerreiro, The role of TREM2 in Alzheimer's disease and other neurodegenerative disorders. *Lancet Neurol* **17**, 721-730 (2018).

46.   Y. Deming, J. Xia, Y. F. Cai, J. Lord, J. L. Del-Aguila, M. V. Fernandez, D. Carrell, K. Black, J. Budde, S. M. Ma, B. Saef, B. Howells, S. Bertelsen, M. Bailey, P. G. Ridge, D. Holtzman, J. C. Morris, K. Bales, E. H. Pickering, J. M. Lee, L. Heitsch, J. Kauwe, A. Goate, L. Piccio, C. Cruchaga, Adni, Genetic studies of plasma analytes identify novel potential biomarkers for several complex traits. *Sci Rep-Uk* **6**, 18092 (2016).

47.   N. Garge, H. Q. Pan, M. D. Rowland, B. J. Cargile, X. X. Zhang, P. C. Cooley, G. P. Page, M. K. Bunger, Identification of Quantitative Trait Loci Underlying Proteome Variation in Human Lymphoblastoid Cells. *Mol Cell Proteomics* **9**, 1383-1399 (2010).

48.   H. J. Westra, M. J. Peters, T. Esko, H. Yaghootkar, C. Schurmann, J. Kettunen, M. W. Christiansen, B. P. Fairfax, K. Schramm, J. E. Powell, A. Zhernakova, D. V. Zhernakova, J. H. Veldink, L. H. Van den Berg, J. Karjalainen, S. Withoff, A. G. Uitterlinden, A. Hofman, F. Rivadeneira, P. A. C. 't Hoen, E. Reinmaa, K. Fischer, M. Nelis, L. Milani, D. Melzer, L. Ferrucci, A. B. Singleton, D. G. Hernandez, M. A. Nalls, G. Homuth, M. Nauck, D. Radke, U. Volker, M. Perola, V. Salomaa, J. Brody, A. Suchy-Dicey, S. A. Gharib, D. A. Enquobahrie, T. Lumley, G. W. Montgomery, S. Makino, H. Prokisch, C. Herder, M. Roden, H. Grallert, T. Meitinger, K. Strauch, Y. Li, R. C. Jansen, P. M. Visscher, J. C. Knight, B. M. Psaty, S. Ripatti, A. Teumer, T. M. Frayling, A. Metspalu, J. B. J. van Meurs, L. Franke, Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature Genetics* **45**, 1238-U1195 (2013).

49.   K. Suhre, M. Arnold, A. M. Bhagwat, R. J. Cotton, R. Engelke, J. Raffler, H. Sarwath, G. Thareja, A. Wahl, R. K. DeLisle, L. Gold, M. Pezer, G. Lauc, M. A. El-Din Selim, D. O. Mook-Kanamori, E. K. Al-Dous, Y. A. Mohamoud, J. Malek, K. Strauch, H. Grallert, A. Peters, G. Kastenmuller, C. Gieger, J. Graumann, Connecting genetic risk to disease end points through the human blood plasma proteome. *Nat Commun* **8**, 14357 (2017).

50.   L. Folkersen, E. Fauman, M. Sabater-Lleal, R. J. Strawbridge, M. Franberg, B. Sennblad, D. Baldassarre, F. Veglia, S. E. Humphries, R. Rauramaa, U. de Faire, A. J. Smit, P. Giral, S. Kurl, E. Mannarino, S. Enroth, A. Johansson, S. B. Enroth, S. Gustafsson, L. Lind, C. Lindgren, A. P. Morris, V. Giedraitis, A. Silveira, A.

Franco-Cereceda, E. Tremoli, I. s. group, U. Gyllensten, E. Ingelsson, S. Brunak, P. Eriksson, D. Ziemek, A. Hamsten, A. Malarstig, Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. *PLoS Genet* **13**, e1006706 (2017).

51. T. T. Chang, J. W. Chen, Emerging role of chemokine CC motif ligand 4 related mechanisms in diabetes mellitus and cardiovascular disease: friends or foes? *Cardiovasc Diabetol* **15**, 117 (2016).

52. T. R. Jay, C. M. Miller, P. J. Cheng, L. C. Graham, S. Bemiller, M. L. Broihier, G. Xu, D. Margevicius, J. C. Karlo, G. L. Sousa, A. C. Cotleur, O. Butovsky, L. Bekris, S. M. Staugaitis, J. B. Leverenz, S. W. Pimplikar, G. E. Landreth, G. R. Howell, R. M. Ransohoff, B. T. Lamb, TREM2 deficiency eliminates TREM2+ inflammatory macrophages and ameliorates pathology in Alzheimer's disease mouse models. *J Exp Med* **212**, 287-295 (2015).

53. D. L. Kober, T. J. Brett, TREM2-Ligand Interactions in Health and Disease. *J Mol Biol* **429**, 1607-1629 (2017).

54. M. M. Painter, Y. Atagi, C. C. Liu, R. Rademakers, H. Xu, J. D. Fryer, G. Bu, TREM2 in CNS homeostasis and neurodegenerative disease. *Mol Neurodegener* **10**, 43 (2015).

55. L. Zhong, X. F. Chen, T. Wang, Z. Wang, C. Liao, Z. Wang, R. Huang, D. Wang, X. Li, L. Wu, L. Jia, H. Zheng, M. Painter, Y. Atagi, C. C. Liu, Y. W. Zhang, J. D. Fryer, H. Xu, G. Bu, Soluble TREM2 induces inflammatory responses and enhances microglial survival. *J Exp Med* **214**, 597-607 (2017).

56. C. E. G. Leyns, D. M. Holtzman, Glial contributions to neurodegeneration in tauopathies. *Mol Neurodegener* **12**, 50 (2017).

57. D. I. Swerdlow, K. B. Kuchenbaecker, S. Shah, R. Sofat, M. V. Holmes, J. White, J. S. Mindell, M. Kivimaki, E. J. Brunner, J. C. Whittaker, J. P. Casas, A. D. Hingorani, Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. *Int J Epidemiol* **45**, 1600-1616 (2016).

58. S. Krasemann, C. Madore, R. Cialic, C. Baufeld, N. Calcagno, R. El Fatimy, L. Beckers, E. O'Loughlin, Y. Xu, Z. Fanek, D. J. Greco, S. T. Smith, G. Tweet, Z. Humulock, T. Zrzavy, P. Conde-Sanroman, M. Gacias, Z. Weng, H. Chen, E. Tjon, F. Mazaheri, K. Hartmann, A. Madi, J. D. Ulrich, M. Glatzel, A. Worthmann, J. Heeren, B. Budnik, C. Lemere, T. Ikezu, F. L. Heppner, V. Litvak, D. M. Holtzman, H. Lassmann, H. L. Weiner, J. Ochando, C. Haass, O. Butovsky, The TREM2-APOE Pathway Drives the Transcriptional Phenotype of Dysfunctional Microglia in Neurodegenerative Diseases. *Immunity* **47**, 566-581 e569 (2017).

59. J. Paloneva, M. Kestila, J. Wu, A. Salminen, T. Bohling, V. Ruotsalainen, P. Hakola, A. B. Bakker, J. H. Phillips, P. Pekkarinen, L. L. Lanier, T. Timonen, L. Peltonen, Loss-of-function mutations in TYROBP (DAP12) result in a presenile dementia with bone cysts. *Nat Genet* **25**, 357-361 (2000).

60. J. Paloneva, T. Manninen, G. Christman, K. Hovanes, J. Mandelin, R. Adolfsson, M. Bianchin, T. Bird, R. Miranda, A. Salmaggi, L. Tranebjaerg, Y. Konttinen, L. Peltonen, Mutations in two genes encoding different subunits of a receptor signaling complex result in an identical disease phenotype. *Am J Hum Genet* **71**, 656-662 (2002).

61. L. Cui, Y. Gao, Y. Xie, Y. Wang, Y. Cai, X. Shao, X. Ma, Y. Li, G. Ma, G. Liu, W. Cheng, Y. Liu, T. Liu, Q. Pan, H. Tao, Z. Liu, B. Zhao, Y. Shao, K. Li, An ADAM10 promoter polymorphism is a functional variant in severe sepsis patients and confers susceptibility to the development of sepsis. *Crit Care* **19**, 73 (2015).

62. M. Allen, M. M. Carrasquillo, C. Funk, B. D. Heavner, F. Zou, C. S. Younkin, J. D. Burgess, H. S. Chai, J. Crook, J. A. Eddy, H. Li, B. Logsdon, M. A. Peters, K. K. Dang, X. Wang, D. Serie, C. Wang, T. Nguyen, S. Lincoln, K. Malphrus, G. Bisceglio, M. Li, T. E. Golde, L. M. Mangravite, Y. Asmann, N. D. Price, R. C. Petersen, N. R. Graff-Radford, D. W. Dickson, S. G. Younkin, N. Ertekin-Taner, Human whole genome genotype and transcriptome data for Alzheimer's and other neurodegenerative diseases. *Sci Data* **3**, 160089 (2016).

63. M. Wang, P. Roussos, A. McKenzie, X. Zhou, Y. Kajiwara, K. J. Brennand, G. C. De Luca, J. F. Crary, P. Casaccia, J. D. Buxbaum, M. Ehrlich, S. Gandy, A. Goate, P. Katsel, E. Schadt, V. Haroutunian, B. Zhang, Integrative network analysis of nineteen brain regions identifies molecular signatures and networks underlying selective regional vulnerability to Alzheimer's disease. *Genome Med* **8**, 104 (2016).

64. A. L. Price, N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, D. Reich, Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-909 (2006).

65. C. Cruchaga, J. S. Kauwe, K. Mayo, N. Spiegel, S. Bertelsen, P. Nowotny, A. R. Shah, R. Abraham, P. Hollingworth, D. Harold, M. M. Owen, J. Williams, S. Lovestone, E. R. Peskind, G. Li, J. B. Leverenz, D. Galasko, I. Alzheimer's Disease Neuroimaging, J. C. Morris, A. M. Fagan, D. M. Holtzman, A. M. Goate,

SNPs associated with cerebrospinal fluid phospho-tau levels influence rate of decline in Alzheimer's disease. *PLoS Genet* **6**, e1001101 (2010).

66. C. Cruchaga, J. S. Kauwe, P. Nowotny, K. Bales, E. H. Pickering, K. Mayo, S. Bertelsen, A. Hinrichs, I. Alzheimer's Disease Neuroimaging, A. M. Fagan, D. M. Holtzman, J. C. Morris, A. M. Goate, Cerebrospinal fluid APOE levels: an endophenotype for genetic studies for Alzheimer's disease. *Hum Mol Genet* **21**, 4558-4571 (2012).

67. J. S. Kauwe, C. Cruchaga, S. Bertelsen, K. Mayo, W. Latu, P. Nowotny, A. L. Hinrichs, A. M. Fagan, D. M. Holtzman, I. Alzheimer's Disease Neuroimaging, A. M. Goate, Validating predicted biological effects of Alzheimer's disease associated SNPs using CSF biomarker levels. *J Alzheimers Dis* **21**, 833-842 (2010).

68. O. Delaneau, J. Marchini, G. P. Consortium, Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nature Communications* **5**, (2014).

69. B. Howie, C. Fuchsberger, M. Stephens, J. Marchini, G. R. Abecasis, Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature Genetics* **44**, 955-+ (2012).

70. R Core Team. (R Foundation for Statistical Computing, Vienna, Austria, 2017).

71. C. C. Chang, C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, J. J. Lee, Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

72. R. J. Pruim, R. P. Welch, S. Sanna, T. M. Teslovich, P. S. Chines, T. P. Gliedt, M. Boehnke, G. R. Abecasis, C. J. Willer, LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-2337 (2010).

73. C. J. Willer, Y. Li, G. R. Abecasis, METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191 (2010).

74. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419 (2017).

75. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

76. K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).

77. L. D. Ward, M. Kellis, HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40**, D930-934 (2012).

78. A. P. Boyle, E. L. Hong, M. Hariharan, Y. Cheng, M. A. Schaub, M. Kasowski, K. J. Karczewski, J. Park, B. C. Hitz, S. Weng, J. M. Cherry, M. Snyder, Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* **22**, 1790-1797 (2012).

79. GTEx Consortium, Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-660 (2015).

80. D. Trabzuni, M. Ryten, R. Walker, C. Smith, S. Imran, A. Ramasamy, M. E. Weale, J. Hardy, Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. *J Neurochem* **119**, 275-282 (2011).

81. W. McLaren, L. Gil, S. E. Hunt, H. S. Riat, G. R. Ritchie, A. Thormann, P. Flicek, F. Cunningham, The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122 (2016).

82. O. O. Yavorska, S. Burgess, MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol* **46**, 1734-1739 (2017).

83. T. Dull, R. Zufferey, M. Kelly, R. J. Mandel, M. Nguyen, D. Trono, L. Naldini, A third-generation lentivirus vector with a conditional packaging system. *J Virol* **72**, 8463-8471 (1998).

84. B. A. Benitez, M. S. Sands, Primary fibroblasts from CSPalpha mutation carriers recapitulate hallmarks of the adult onset neuronal ceroid lipofuscinosis. *Sci Rep* **7**, 6332 (2017).

85. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

86. N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**, 525-527 (2016).

Christian Haass has a patent for TREM2 cleavage modulators and uses thereof

(PCT/EP2017/068684)

Supplementary Materials

Supplementary Results

Supplementary Material and Methods

Supplementary Figure S1-S9

Supplementary Tables S1-S9

Supplementary References (References number 85-86)

Table 1. ADNI CSF cohort characteristics.

	AD	Controls	EMCI	LMCI	SMC	*P*
N **(813)**	172	169	183	221	68	
Age	74.39 ± 8.56	74.47 ± 5.85	71.23 ± 7.39	73.06 ± 7.41	71.58 ± 5.57	< 0.001
females (%)	74 (43.0)	80 (47.3)	77 (42.1)	91 (41.2)	38 (55.9)	0.224
APOE ε4+ **(%)**	115 (66.9)	40 (23.7)	87 (47.5)	129 (58.4)	24 (35.3)	< 0.001
CDR at LP (%)						< 0.001
0	0 (0.0)	169 (100.0)	0 (0.0)	1 (0.5)	68 (100.0)	
0.5	74 (43.0)	0 (0.0)	183 (100.0)	219 (99.1)	0 (0.0)	
1	98 (57.0)	0 (0.0)	0 (0.0)	1 (0.5)	0 (0.0)	
WashU sTREM2	2433.99 ± 776.77	2436.56 ± 768.94	2391.85 ± 728.48	2401.97 ± 711.59	2444.99 ± 690.01	0.964
LMU sTREM2	4018.65 ± 1946.09	3988.02 ± 1923.06	3741.56 ± 2066.17	3918.40 ± 1828.83	3784.28 ± 1724.71	0.641
CSF Aβ$_{42}$	626.69 ± 247.96	1059.19 ± 358.69	941.90 ± 367.92	747.16 ± 304.79	1115.83 ± 366.22	< 0.001
CSF tau	367.80 ± 149.72	232.57 ± 81.04	262.10 ± 132.04	310.72 ± 126.60	240.33 ± 90.54	< 0.001
CSF ptau$_{121}$	36.88 ± 16.49	21.21 ± 8.01	24.90 ± 14.87	30.64 ± 14.40	22.04 ± 9.39	< 0.001

Disease status is based on diagnosis at lumbar puncture (LP):

AD = Alzheimer disease case

Controls = cognitively healthy individuals

EMCI = Early mild cognitive impairment (MCI)

LMCI = Late MCI

SMC = Subjective memory complaint

Age is age at LP. Values are reported in years, mean ± SD

CDR at LP is the Clinical Dementia Rating (CDR) at LP:

0 = no cognitive impairment, 0.5 = very mild cognitive impairment, 1 = mild cognitive impairment.

WashU sTREM2 represents the ADNI CSF sTREM2 values measured at Washington University

LMU sTREM2 represents the ADNI CSF sTREM2 values measured at Ludwig-Maximilians-Universität

All the CSF values are represented as pg/mL Mean ± SD

Table 2. ADNI CSF cohort characteristics by *TREM2* mutation carriers.

	p.R62H	p.R47H	p.L211P	p.D87N	p.H157Y	*P*
N	18	4	11	4	1	
Age	75.01 ± 6.48	73.24 ± 14.51	72.78 ± 4.36	71.14 ± 5.95	73.10 ± NA	0.843
Females (%)	9 (50.0)	2 (50.0)	6 (54.5)	0 (0)	1 (100.0)	0.393
APOE ε4+ (%)	7 (38.9)	3 (75.0)	2 (18.2)	2 (50.0)	0 (0)	0.194
CDR at LP (%)						0.815
0	4 (22.2)	1 (25.0)	5 (45.5)	1 (25.0)	1 (100.0)	
0.5	10 (55.6)	3 (75.0)	5 (45.5)	3 (75.0)	-	
1	4 (22.2)	-	-	-	-	
2	-	-	1 (9.1)	4 (0.7)	-	
Case status						0.871
AD	13 (72.2)	2 (50.0)	6 (54.5)	1 (25.0)	-	
CO	3 (16.7)	1 (25.0)	3 (27.3)	1 (25.0)	1 (100.0)	
WashU sTREM2	2153 ± 842.34	2573 ± 237.61	1636± 563.94	1570 ± 162.55	3658.23 ± NA	0.140
LMU sTREM2	3335 ± 1882	4791 ± 965.51	2386± 1389	1615 ± 757.86	5641.71 ± NA	0.055
CSF Aβ$_{42}$	976.76 ± 381.52	828.13 ± 351.90	1076 ± 507.09	685.67 ± 251.93	NA	0.443
CSF tau	303.46 ± 135.43	307.85 ± 97.93	246.24 ± 113.22	252.97 ± 122.51	213.70 ± NA	0.678
CSF ptau$_{121}$	29.31 ± 14.86	30.80 ± 13.93	22.18 ± 12.66	23.42 ± 12.99	18.05 ± NA	0.729

Each column is a *TREM2* variant that has been associated with AD risk and was present in the available ADNI data.

P is the P value for an ANOVA comparing all the variants

Age represents age at lumbar puncture (LP). Values represent years, mean ± SD

CDR at LP is the Clinical Dementia Rating at LP

WashU sTREM2 are CSF sTREM2 ADNI values measured at Washington University

LMU sTREM2 are CSF sTREM2 ADNI values measured at Ludwig-Maximilians-Universität.

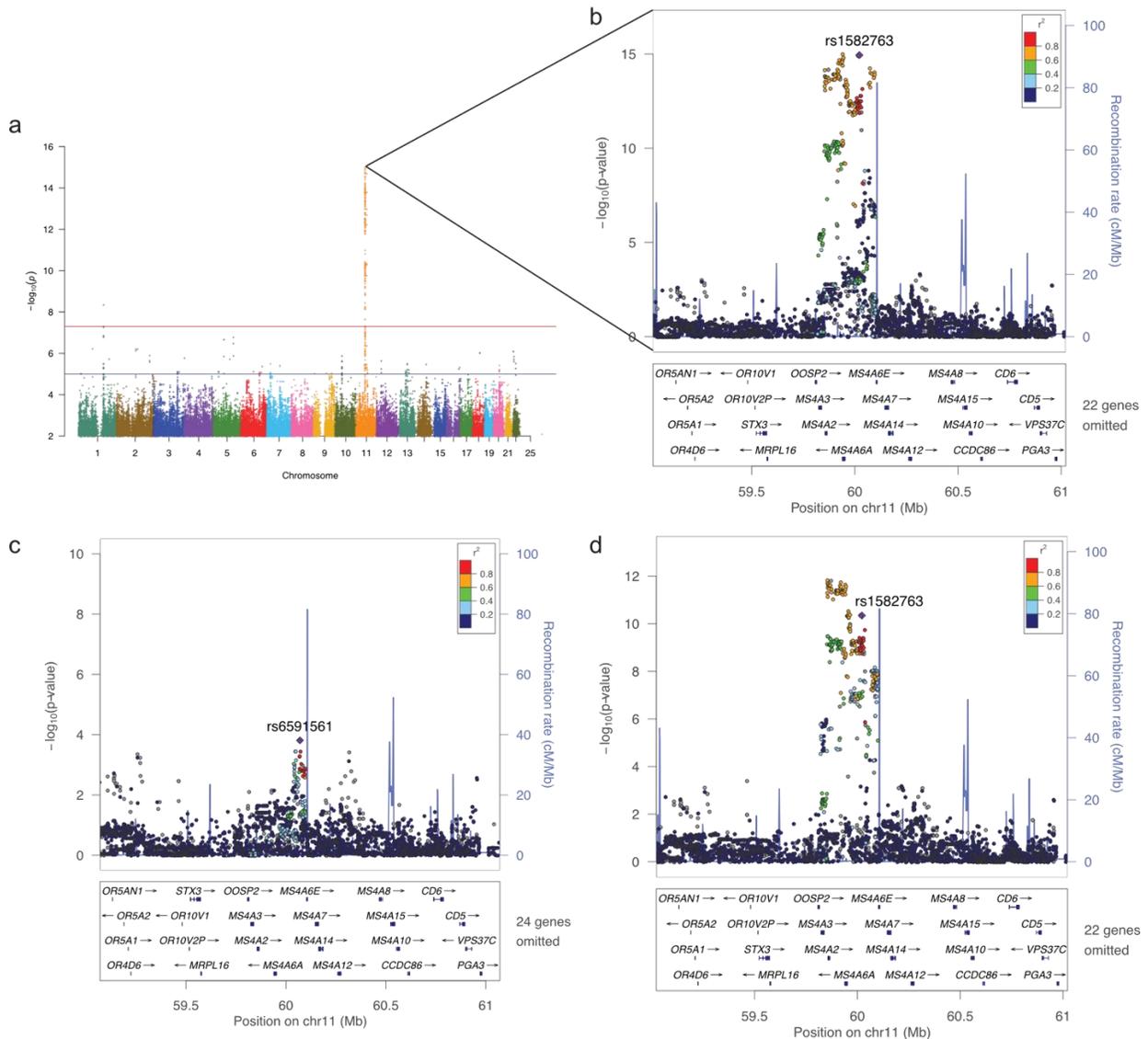All the CSF values are represented as pg/mL Mean ± SD

FIGURES

Figure 1 Association plots from single-variant analysis of CSF sTREM2 (A) Manhattan plot shows the negative log$_{10}$-transformed *P* values on the y-axis for CSF sTREM2 based on the LMU measures (see material and methods). Comparable results were found with the WU measures (see supplementary material). The horizontal lines represent the genome-wide significance threshold (red) and suggestive threshold (blue). **(B)** Regional association plot of the *MS4A* gene region in the single-variant analysis. **(C)** Regional association plot of the *MS4A* region after conditioning on the top SNP (rs1582763) and **(D)** after conditioning on rs6591561 (*MS4A4A* p.M159V). The SNPs labeled on each regional plot are represented by a purple diamond. Each dot represents individual SNPs and dot colors in the regional plots represent LD with the named SNP. Blue vertical lines in the regional plots show recombination rate as marked on the right-hand y-axis.
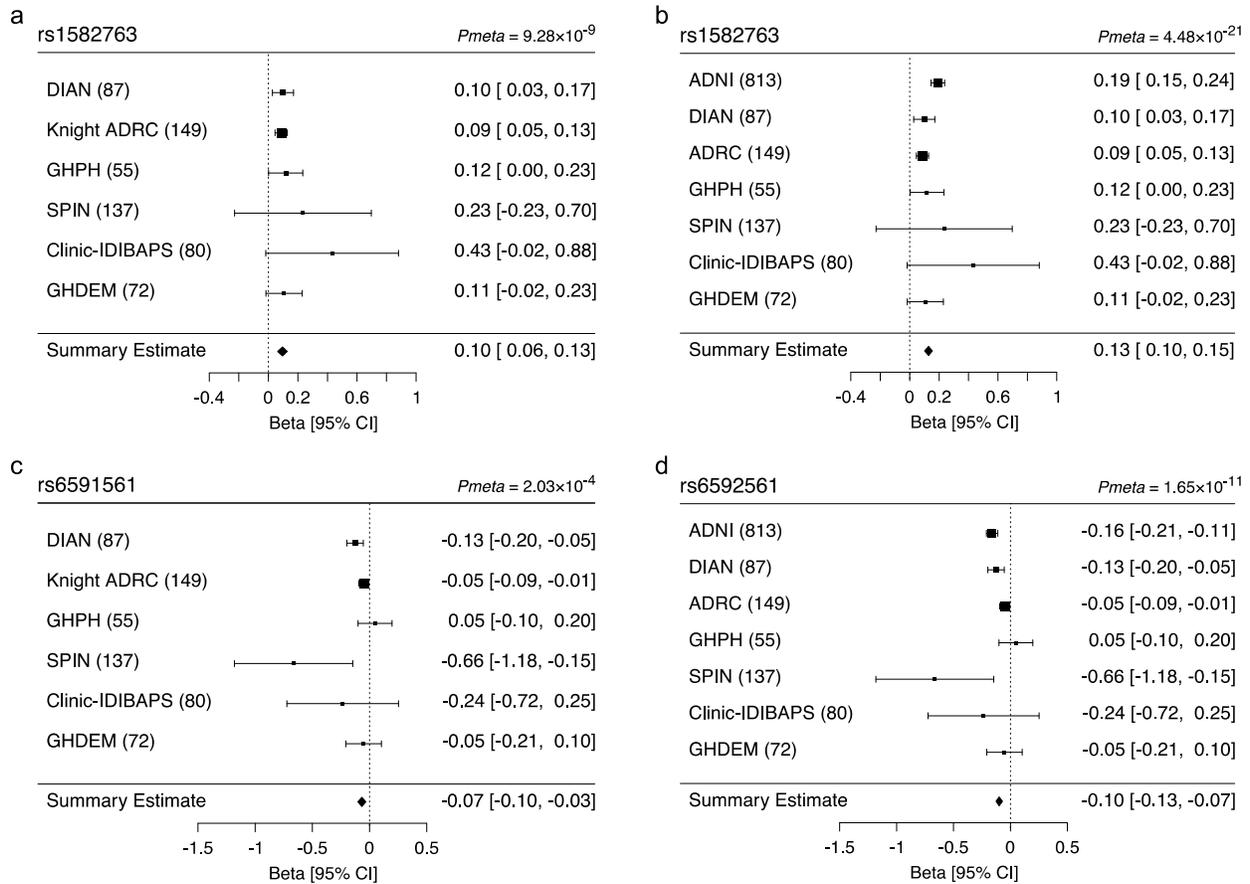
Figure 2 Meta-analysis of replication data sets (A) Forest plot of the meta-analysis of rs1582763 in the replication datasets from DIAN, Knight ADRC, GHPH, SPIN, Clinic-IDIBAPS, and GHDEM; **(B)** forest plot of the meta-analysis of rs1582763 including ADNI; **(C)** forest plot of *MS4A4A* p.M159V (rs6591561) in the meta-analysis of DIAN, Knight ADRC, GHPH, SPIN, Clinic-IDIBAPS, and GHDEM; and **(D)** including ADNI. The sample size (N) for each study is in parentheses next to the study name.
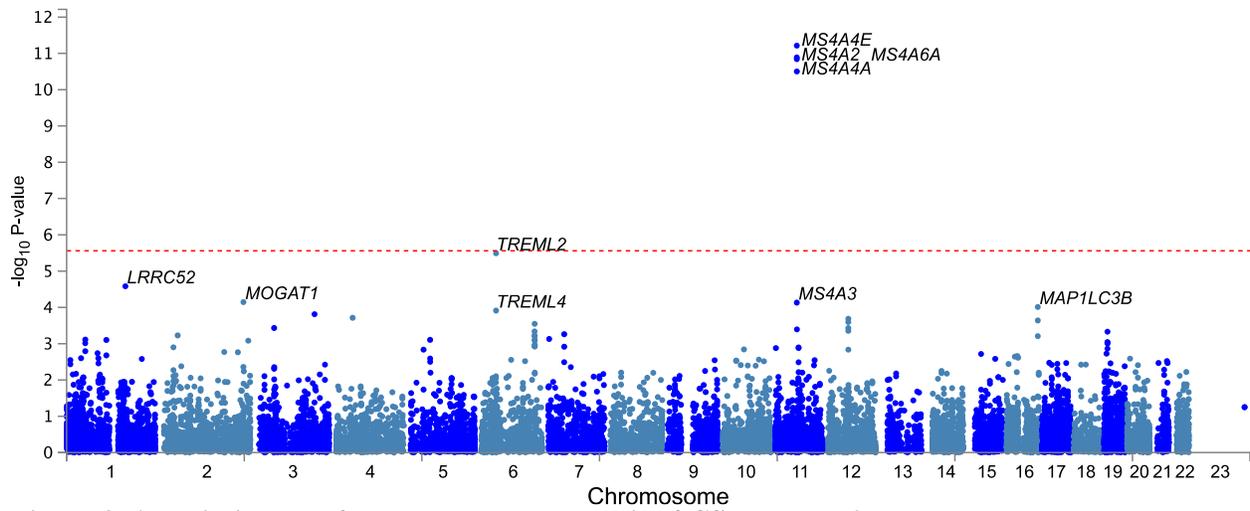
Figure 3. Association plot from gene-based analysis of CSF sTREM2. Manhattan plot shows the negative $\log_{10}$-transformed P values on the y-axis. Each dot represents one gene; the top ten genes in the MAGMA results are labeled. The red dotted line represents the gene-wide significance threshold ($P = 2.74\times10^{-6}$).

A

Mendelian Randomization analyses for CSF sTREM2 vs. AD risk

Method	Estimate	S.E	95% CI		P
MR-Egger	-3.35×10^{-04}	7.26×10^{-05}	-4.77×10^{-04}	-1.92×10^{-04}	3.97×10^{-06}
Penalized MR-Egger	-3.35×10^{-04}	7.26×10^{-05}	-4.77×10^{-04}	-1.92×10^{-04}	3.97×10^{-06}
Robust MR-Egger	-3.35×10^{-04}	1.93×10^{-05}	-3.73×10^{-04}	-2.97×10^{-04}	$<1.00\times10^{-06}$
Penalized robust MR-Egger	-3.35×10^{-04}	1.93×10^{-05}	-3.73×10^{-04}	-2.97×10^{-04}	$<1.00\times10^{-06}$

Residual Standard Error : 0.934
Residual standard error is set to 1 in calculation of confidence interval when its estimate is less than 1.
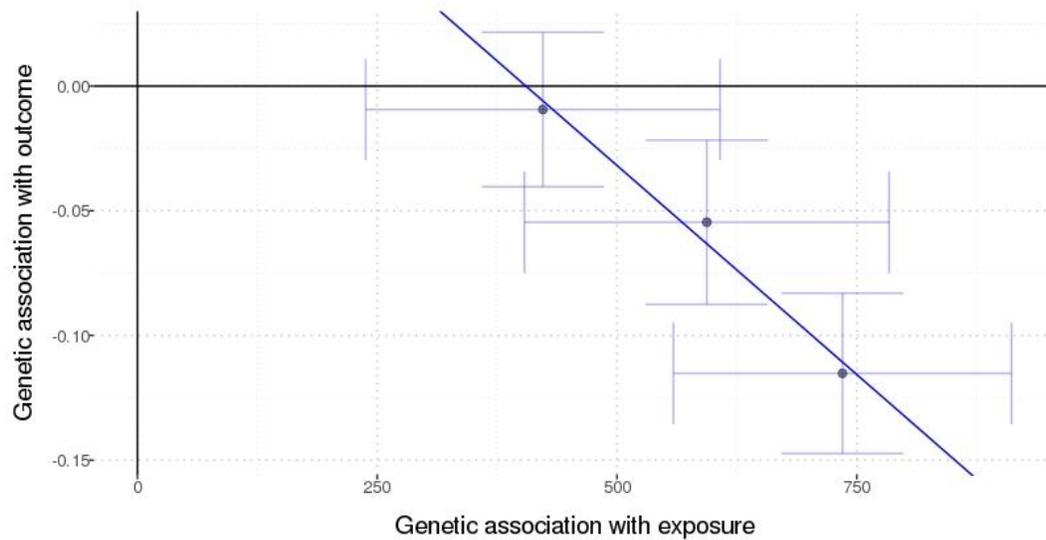Heterogeneity test statistic = 0.3884 on 1 degrees of freedom, ($P = 0.533$).
I^2_GX statistic: 60.0%

B

Figure 4. Mendelian randomization analyses for CSF sTREM2 vs AD risk. (A) Summary statistics for the different MR models analyzed. **(B)** Scatter plot illustrating genetic associations with AD risk (outcome) against genetic associations with CSF sTREM2 (exposure) with error bars representing 95% confidence intervals for the associations. The variants are all oriented to the CSF sTREM2 effect. The line represents the MR-Egger causal estimate.
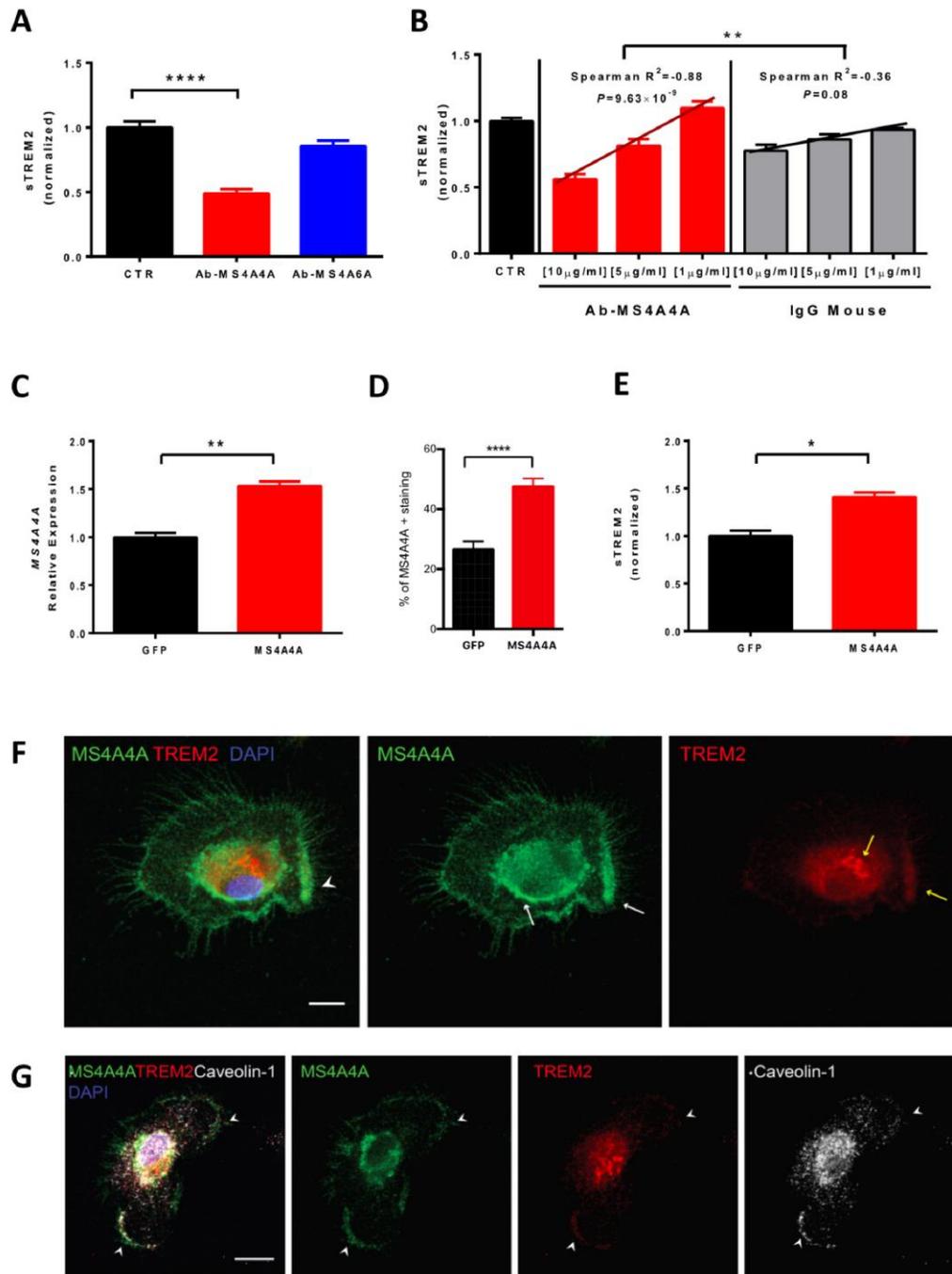
Figure 5: sTREM2 in cell culture supernatant is modified by overexpressing *MS4A4A* or by treating cells with anti-MS4A4A antibody (**A**) sTREM2 detected in cell culture media 48h after treatment with antibodies against MS4A4A and MS4A6A, normalized values are shown. Error bars represent means ± SEM. ****$P < 0.0001$. P values were calculated using the Mann-Whitney test. (**B**) Dose-response effect of MS4A4A antibody on sTREM2 or an isotypic control (IgG Mouse). Difference in the slope between the dose-response for MS4A4A and IgG-Mouse antibody was tested using ANOVA, **$P < 0.01$. Four experiments were performed and results for sTREM2 are expressed as normalized values on untreated controls. (**C**) *MS4A4A* was overexpressed in primary human macrophages using a lentiviral vector encoding

for the human *MS4A4A* cDNA (NM_148975). The plot represents mRNA levels. Error bars represent means ± SD, **$P < 0.01$. P values were calculated by Mann-Whitney test. (D) Quantification of MS4A4A expression by immunofluorescence in control human macrophages or overexpressing MS4A4A. Staining were analyzed as the percentage area of positive staining (number of positive pixels/1 mm2) within the area of interest. **(E)** sTREM2 was significantly higher in cells overexpressing *MS4A4A*, *$P < 0.05$. P values were calculated using the Mann-Whitney test. (**F**) Confocal images of human macrophages stained for TREM2 (red) and MS4A4A (green) antibodies. White arrows indicate MS4A4A positive signal while yellow arrows indicate TREM2 positive signal. Arrowheads (white) indicate colocalizing signals for TREM2 and MS4A4A on the cell membrane. Scale bar 10 µm. **(G)** Representative confocal images of macrophages stained for Caveolin-1 (grey), TREM2 (red), and MS4A4A antibodies (green). Arrowheads (white) indicate colocalization signals. Scale bar 10 µm.