nature neuroscience

REVIEW ARTICLE

https://doi.org/10.1038/s41593-020-0599-5

The multiplex model of the genetics of Alzheimer's disease

Rebecca Sims[1], Matthew Hill[1,2] and Julie Williams[1,2] ✉

Genes play a strong role in Alzheimer's disease (AD), with late-onset AD showing heritability of 58–79% and early-onset AD showing over 90%. Genetic association provides a robust platform to build our understanding of the etiology of this complex disease. Over 50 loci are now implicated for AD, suggesting that AD is a disease of multiple components, as supported by pathway analyses (immunity, endocytosis, cholesterol transport, ubiquitination, amyloid-β and tau processing). Over 50% of late-onset AD heritability has been captured, allowing researchers to calculate the accumulation of AD genetic risk through polygenic risk scores. A polygenic risk score predicts disease with up to 90% accuracy and is an exciting tool in our research armory that could allow selection of those with high polygenic risk scores for clinical trials and precision medicine. It could also allow cellular modelling of the combined risk. Here we propose the multiplex model as a new perspective from which to understand AD. The multiplex model reflects the combination of some, or all, of these model components (genetic and environmental), in a tissue-specific manner, to trigger or sustain a disease cascade, which ultimately results in the cell and synaptic loss observed in AD.

Alzheimer's disease genetics

Early-onset AD and the amyloid cascade hypothesis. Early findings of disease mutations in the amyloid precursor protein gene (*APP*) and presenilin genes[1–3] were pivotal to the development of the amyloid cascade hypothesis[4]. It posits that misprocessing of amyloid-β (Aβ) and its deposition are the primary causal event in AD pathogenesis. Although these mutations explain less than 1% of AD, there is no doubt that this hypothesis has shaped mechanistic research and drug development for AD over the last 25 years[5]. However, recent failures in clinical trials based on removing soluble and/or insoluble Aβ or targeting enzymes responsible for cleavage of APP have thrown doubt on the hypothesis[6,7]. Several possibilities may explain this lack of success. First, the hypothesis may only relate to rare forms of early-onset AD in which causal mutations are observed. Second, the drug treatments may only be effective in the early stages of AD and not when the disease has already caused extensive neurodegeneration[8,9]. Indeed, evidence suggests that the disease process begins up to 20 years before the first cognitive symptoms are observed[10]. The hope is that amyloid-based drug trials on mutation carriers, recruited and treated presymptomatically, will inform our understanding here[11]. Third, Aβ and the associated amyloid plaques may be correlates of disease mechanisms that have the primary influence on disease development[12].

Late-onset AD genetics: common variation. Looking beyond AD mutations, genetic research has now produced extensive evidence that other genetic factors contribute to disease. Common forms of late-onset AD (LOAD) have heritability estimates of 56–79%[13], and rarer forms with early-onset (5% of AD cases) with heritability of over 90%[14], are contributed to by multiple genetic risk factors.

Apolipoprotein E (*APOE*) on chromosome 19 was the first risk gene identified as associating with LOAD[15], as well as influencing familial and early forms of disease, and it remains the strongest genetic risk factor. The differential expression of the three major isoforms of ApoE (ε2, ε3 and ε4) is dependent on two point mutations (rs429358 and rs7412) within exon 4 of the gene. An increased risk of AD is found in carriers of the ε4 allele, whereas the ε2 allele confers a small protective effect[16,17]. Risk is dose-dependent, with a threefold increase in ε4 heterozygotes (ApoE ε3/ε4) and a 15-fold increase in ε4 homozygotes (ApoE ε4/ε4). Disease susceptibility is thought to result from a conformational change in ApoE that affects the protein's ability to bind ligands, including Aβ and TREM2[18]. ApoE ε4 is thought to be less efficient in mediating clearance of soluble and aggregated Aβ[19], but is also implicated in other cellular processes and tissues and certainly needs more study to define its full contribution to disease[20].

Perhaps the most successful approach to identifying the genetic architecture of AD is the genome-wide association study (GWAS). In 2009, the first novel genetic associations were identified using GWAS, showing genome-wide statistically significant association between AD and variants within the *CLU*, *PICALM* and *CR1* loci[21,22]. To date, over 50 risk loci (Fig. 1 and Table 1) with genome-wide significance (GWS; $P < 5 \times 10^{-8}$) are associated with AD. This success in identifying risk loci can be attributed to the extensive national and international collaboration seen within the field. The initial Genetic and Environmental Risk in AD (GERAD) and European AD Initiative (EADI) GWAS[21,22] were quickly followed by studies led by the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE)[23] and AD Genetics Consortium (ADGC)[24], as well as an additional GERAD study[25]. These four consortia subsequently joined together to form the International Genomics of Alzheimer's Project (IGAP) who, in 2013, identified a further 11 risk loci as novel genome-wide significant LOAD susceptibility loci[26]. The IGAP GWAS results summary dataset is freely available to researchers (individual-level data available upon request to the relevant consortia) and has been pivotal to multiple successive studies in a variety of research areas[27–29].

Building upon the IGAP[26] dataset, single-nucleotide[26,30–34], genome-wide[35,36], transethnic[32] and proxy design[31,37,38] (based on

[1]Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK. [2]UK Dementia Research Institute, School of Medicine, Cardiff University, Cardiff, UK. ✉e-mail: williamsj@cardiff.ac.uk

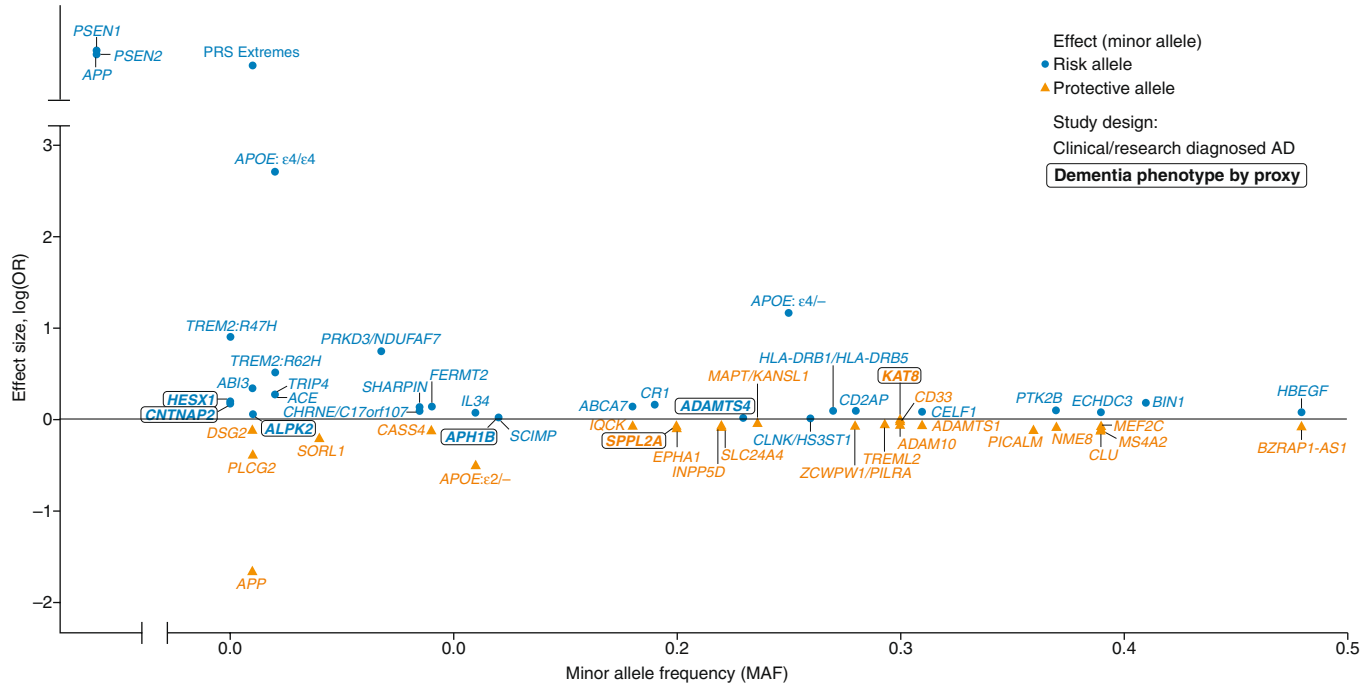

Fig. 1 | Schematic of Mendelian disease-causing genes and loci reaching GWS for single-variant (not gene-wide) association with sporadic AD. Blue circles and orange triangles represent risk and protective association, respectively. Associations identified in AD-diagnosed cohorts are not in boxes, while associations identified in meta-analysis of AD-diagnosed and proxy-diagnosed cohorts are indicated by black box outlines. Of note, AD case–control data is absent for the *CNTNAP2* and *HESX1* loci and is only weakly supportive for the *ALPK2*and *APH1B* loci ($P = 10^{-2}$). OR, odds ratio.

reports of parental history) studies have identified numerous novel GWS loci (Fig. 1 and Table 1).

It is estimated that a substantial proportion (up to 60%)[39,40] of the genetic variance of LOAD is not accounted for. Given the success in other diseases[41], there is no doubt that more powerful GWAS will identify additional associations. Currently, studies using research-based or clinically diagnosed AD have examined 33,692 cases and 56,077 controls[42], so more associations will be found with increasing sample size and greater power in the future. However, this 'missing heritability' may also be contributed to by rare or low-frequency susceptibility genes.

Late-onset AD genetics: rare variation. The primary technology for the detection of rare genetic variation (population frequency less than 1%) has been sequencing. Next-generation sequencing (NGS) technologies in the form of whole-exome and whole-genome sequencing have identified protein-coding changes associated with disease[43–46]. Interestingly, a number of rare disease-associated variants are identified in loci with common variants associated with LOAD[47–50], suggesting that these genes influence disease susceptibility in multiple ways. A number of additional loci have received attention as putative risk genes[51–53].

Sequencing has historically proven to be prohibitively expensive for broad use in the field. While the costs of such experiments are falling, future gene discovery may be increasingly tractable through enriching sequencing for the most heritable samples, including early-onset and familial AD. An alternate approach for rare-variant detection is the use of exome-wide microarrays with variants selected from whole-exome sequencing. This approach has limitations and can only test what is known. Sims and colleagues used this approach in a powerful genome-wide study and found novel coding variants that influence AD and also showed that improved imputation panels now make GWAS more amenable to detecting rare variants[54]. It is important to note that variants of small statistical effect can show substantive biological changes of disease relevance. For example, Sims et al.[54] identified *PLCG2* p.(Arg522), which has an effect size of 0.68 and has been shown to increase enzymatic activity in cell lines[55], human microglia derived from induced pluripotent stem cells (iPSC) and mouse microglia (Phillips, T. et al., unpublished data).

Systematic analysis of gene–gene interaction or epistasis in AD has been limited, largely due to insufficient power and the massive multiple-testing burden inherent in genome-wide epistasis screening. Initial studies have identified interaction between single-nucleotide polymorphisms (SNPs) that require replication[56]. It is noteworthy that a small number of individuals live well beyond 75 years of age without any symptoms of cognitive decline despite possessing a large number of risk factors for AD. These 'AD-resilient' individuals may harbor protective genetic variation[57].

Sub-phenotypes of disease. Genetic relationships have also been sought for disease phenotypes. Aside from the core cognitive symptoms of disease, individuals with AD can develop a range of behavioral symptoms. One area that has received attention in recent years is psychosis. Psychotic symptoms are significantly more common in AD than the general population, affecting ~40% of cases[58]. They are associated with decreased quality of life for caregivers and patients[59], more rapid cognitive[60] and functional decline[61], and premature institutionalization[59]. While no gene has, thus far, shown genome-wide significant association to psychosis in AD, evidence suggests that loci influencing psychosis in disease do so with a greater effect than generally seen in LOAD (excluding *APOE*)[62] and that the lack of a significant association may be accounted for by the small sample sizes tested to date. Another area of study is rate of decline. Early studies show that the genetic architectures for AD disease risk and rate of decline are distinct, with *APOE* showing no association with disease progression[60]. Recent work investigating the impact of both single-nucleotide AD-associated variation and polygenic risk score (PRS) (generated from the IGAP genome-wide significant hits) on rate of decline show association between both

Table 1 | Loci reaching genome-wide significance for association with sporadic AD

Locus	GWS locus or gene	Original SNP and publication	Dataset	Functional information
1	*APOE*	rs429358 p.(Cys112Arg); ref. [15]	Case–control	A multifactorial protein, known primarily for its role in lipid transport. Known to bind soluble Aβ.
		rs7412 p.(Cys158Arg); ref. [15]		
2	*EPHA1*	rs11767557; refs. [24,25]	Combined ADGC and GERAD+	Receptor tyrosine kinase. Role in immunity and endocytosis. Regulates cell morphology and motility, including permeability of the blood–brain barrier to leucocytes.
3	*CLU*	rs11136000; refs. [21,22]	GERAD EADI	Molecular chaperone. Role in immunity and cholesterol metabolism. Binds Aβ.
4	*INPP5D*	rs35349669; ref. [26]	IGAP	Inositol polyphosphate-5-phosphatase. Role in immunity and cholesterol metabolism. Mediates signaling of multiple myeloid-cell pathways including proliferation, survival and chemotaxis. Inhibits TREM2 signaling by association with DAP12.
5	*HLA-DRB5/HLA-DRB1*	rs9271192; ref. [26]	IGAP	HLA class II histocompatibility antigen. Role in immunity, including involvement in antigen presentation.
6	*CR1*	rs6656401; ref. [22]	EADI	Complement receptor. Role in immunity; functions include clearance of complement opsonized molecules and microglial phagocytosis.
7	*TREM2*	rs75932628 p.(Arg47His); refs. [42,44]	Mixed-cohorts	Receptor of the immunoglobulin superfamily, binds lipids and Aβ. Signals to affect multiple processes in myeloid cells including phagocytosis and cellular metabolism.
		rs143332484 p.(Arg62His); ref. [54]	IGAP	
8	*CD33*	rs3865444; refs. [24,25]	Combined ADGC and GERAD+	Myeloid-cell transmembrane receptor that binds sialic acids. Role in immunity.
9	*MS4A* gene cluster	rs4938933; ref. [24] rs610932; ref. [25]	ADGC GERAD+	Specific function unknown. Expressed predominately in immune cells.
10	*ABI3*	rs616338 p.(Ser209Phe); ref. [54]	IGAP	Component of Abi–WAVE complex, which regulates actin polymerization. Role in immunity.
11	*PLCG2*	rs72824905 p.(Pro522Arg); ref. [54]	IGAP	Phospholipase catalyzing the conversion of IP3 and DAG. Signal transducer of multiple pathways in immune cells.
12	*ZCWPW1* and *PILRA*	rs1476679; ref. [26]	IGAP	ZCWPW1: Unknown function. Possible reader of histone modifications. PILRA: Control of cell signaling via SHP-1.
13	*MEF2C*	rs190982; ref. [26]	IGAP	Transcription factor involved in development of multiple tissue types. Putative master regulator of microglia. In neurons, controls activity-dependent synapse number. Hub gene.
14	*CD2AP*	rs9349407; refs. [24,25]	Combined ADGC and GERAD+	Adaptor molecule involved in cytoskeletal dynamics. Involved in early endosome morphology.
15	*BIN1*	rs744373; ref. [23]	CHARGE	Involved in endocytic recycling and Aβ production. Also involved in membrane folding.
16	*PICALM*	rs3851179; ref. [21]	GERAD	Clathrin assembly protein involved in clathrin-mediated endocytosis and transcytosis.
17	*CASS4*	rs7274581; ref. [26]	IGAP	Regulates focal adhesion integrity and cell spreading. Roles in cytoskeleton and axon development and tau metabolism.
18	*CELF1/SPI1*	rs10838725; ref. [26]	IGAP	RNA-binding protein involved in pre-mRNA alternative splicing. Role in cytoskeleton and axon development.
19	*FERMT2*	rs17125944; ref. [26]	IGAP	Scaffolding protein, part of the extracellular matrix; controls cell shape.
20	*NME8*	rs2718058; ref. [26]	IGAP	Unknown function. Possibly involved in ciliary function with a role in cytoskeleton and/or axon development.
21	*SORL1*	rs11218343; ref. [26]	IGAP	Endocytic receptor involved in the uptake of lipoproteins, APP processing and lysosomal targeting of Aβ.
		Gene-wide[48]	ADES-FR	
22	*ABCA7*	rs3764650; ref. [25]	GERAD+	Transporter involved in cholesterol metabolism and phagocytic clearance of Aβ.
		Gene-wide[50]	IGAP	

Continued

Table 1 | Loci reaching genome-wide significance for association with sporadic AD (continued)

Locus	GWS locus or gene	Original SNP and publication	Dataset	Functional information
23	*SLC24A4–RIN3*	rs10498633; ref. [26]	IGAP	SLC24A4: Na^+Ca^{2+}, K^+ exchange. RIN3: Ras interaction–interference protein regulating endocytosis. Role in cholesterol metabolism.
24	*PTK2B*	rs28834970; ref. [26]	IGAP	Cytoplasmic protein tyrosine kinase sensitive to calcium. Regulation of ion channels in neurons, cell spreading and migration and immune cell function.
25	*ADAM10*	rs593742; refs. [33,37]	IGAP+	Metalloprotease responsible for proteolytic processing of APP.
			Combined UK Biobank and IGAP	
26	*IGHV1-67*	Gene-wide[35]	IGAP	Unknown function.
27	*PPARGC1A*	Gene-wide[36]	IGAP	Transcriptional coactivator regulation mitochondrial oxidative metabolism.
28	*TP53INP1*	Gene-wide[35]	IGAP	Tumor suppressor activity; regulates autophagy and transcription.
29	*ECHDC3*	rs7920721; ref. [32]	ADGC and IGAP	Unknown function.
			IGAP+	
30	*ACE*	rs138190086; ref. [33]	IGAP+	Catalyzes the conversion of angiotensin I into a physiologically active peptide, angiotensin II. Controls blood pressure and fluid–electrolyte balance.
		rs6504163; ref. [37]	Combined UK Biobank and IGAP	
31	*ADAMTS1*	rs2830500; ref. [33]	IGAP+	Metalloproteinase. Degrades extracellular matrix proteoglycans. Expression is induced by immune response.
32	*IQCK*	rs7185636; ref. [33]	IGAP+	Unknown function.
33	*TRIP4*	rs74615166; ref. [30]	Fundaciô ACE & IGAP	Transcriptional coactivator of nuclear receptors.
34	*RORA*	Gene-wide[36]	IGAP	Nuclear hormone receptor. Possible roles in circadian rhythm, cholesterol metabolism and inflammation.
35	*ZNF423*	Gene-wide[36]	IGAP	DNA-binding transcription factor. Involved in differentiation of adipocytes, neurons and leukemia.
36	*APP*	rs63750847, p.(Ala673Thr); ref. [46]	Icelandic, Finnish and Swedish	Amyloid precursor protein.
37	*IGHG3*	rs77307099; ref. [44]	ADSP	Immunoglobulin gene whose antibodies interact with Aβ.
38	*AC099552.4*	7:154988675:G:A; ref. [44]	ADSP	Non-coding RNA.
39	*ZNF655*	Gene-wide[44]	ADSP	Zinc-finger protein; transcriptional regulation.
40	*HBEGF/AFDN1*	rs11168036; ref. [32]	Transethnic ADGC and IGAP	Heparin-binding EGF-like growth factor. May be involved in macrophage-mediated cellular proliferation.
41	*BZRAP1-AS1*	rs2632516; ref. [32]	Transethnic ADGC and IGAP	Non-coding RNA.
42	*TPBG*	Gene-wide[32]	Transethnic ADGC and IGAP	Trophoblast glycoprotein encodes a leucine-rich transmembrane glycoprotein that may be involved in cell adhesion.
43	*DSG2*	rs8093731; refs. [26,31]	IGAP	Desmoglein 2, a cell-adhesion molecule. Desmogleins are calcium-binding transmembrane glycoprotein components of desmosomes, cell–cell junctions
			Combined ADSP, IGAP, PGC–ALZ and deCODE	
44	*CLNK* and *HS3ST1*	rs6448453; ref. [31]	Combined UK Biobank, ADSP, IGAP, PGC–ALZ and deCODE	CLNK: member of SLP-76 family of immune-cell-specific adaptors.
		rs4351014; ref. [34]	Combined GR@CE–DEGESCO, IGAP and UK Biobank	HS3ST1: Sulfotransferase that utilizes 3′-phospho-5′-adenylyl sulfate (PAPS) to catalyze the transfer of a sulfo group to position 3 of glucosamine residues in heparan.
45	*SCIMP*	rs113260531; ref. [31]	Combined UK Biobank, ADSP, IGAP, PGC–ALZ and deCODE	Transmembrane adaptor protein involved in MHC class II signaling transduction.
		rs77493189; ref. [38]	Combined UK Biobank and IGAP	

Continued

Table 1 | Loci reaching genome-wide significance for association with sporadic AD (continued)

Locus	GWS locus or gene	Original SNP and publication	Dataset	Functional information
46	*PRKD3* and *NDUFAF7*	rs876461; ref. [34]	Combined GR@CE-DEGESCO, IGAP and UK Biobank	PRKD3: Protein kinase D family of serine–threonine kinases, which bind diacylglycerol and phorbol esters.
				NDUFAF7: Assembly factor protein, assembly and stabilization of the mitochondrial respiratory chain complex I.
47	*TREML2*	rs9381040; ref. [34]	Combined GR@CE-DEGESCO, IGAP and UK Biobank	Cell surface receptor that may play a role in innate and adaptive immune response-enhancing T-cell activation.
48	*SHARPIN*	rs34674752 p.(Pro294Ser); ref. [34]	Combined GR@CE-DEGESCO, IGAP and UK Biobank	Component of the LUBAC complex; plays a key role in NF-κB activation and regulation of inflammation.
		rs34173062 p.(Ser17Phe); ref. [34]		
49	*MAPT* and *KANSL1*#	rs2732703; ref. [34]	Combined GR@CE-DEGESCO, IGAP and UK Biobank	MAPT: Transcripts are differentially expressed in the nervous system, depending on stage of neuronal maturation and neuron type.
				KANSL1: Subunit of histone acetylation complexes MLL1 and NSL1. The NSL complex may be involved in transcription regulation.
50	*CHRNE* and *C17orf107*	rs72835061; ref. [34]	Combined GR@CE-DEGESCO, IGAP and UK Biobank	CHRNE: Controls an ion-conducting channel across the plasma membrane. C17orf107: Unknown function.
51	*IL34*	rs4985556 p.(Tyr213Ter); refs. [34,37]	Combined UK Biobank and IGAP	Interleukin-34. Cytokine that promotes the proliferation, survival and differentiation of monocytes and macrophages.
			Combined GR@CE-DEGESCO, IGAP and UK Biobank	
Dementia in parental-proxy observation (not GWS in AD diagnosed).				
52	*CNTNAP2**	rs114360492; ref. [31]	Combined UK Biobank, ADSP, IGAP, PGC-ALZ and deCODE	Member of the neurexin family.
53	*ALPK2***	rs76726049; ref. [31]	Combined UK Biobank, ADSP, IGAP, PGC-ALZ and deCODE	Alpha kinase. Specific function is unknown.
54	*ADAMTS4*	rs4575098; ref. [31]	Combined UK Biobank, ADSP, IGAP, PGC-ALZ and deCODE	Extracellular matrix metalloproteinase (aggrecanase-1)
55	*APH1B***	rs117618017 p.(Thr27Ile); ref. [31]	Combined UK Biobank, ADSP, IGAP, PGC-ALZ and deCODE	Component of the gamma-secretase complex; assists in the cleavage of APP.
56	*KAT8*	rs59735493; ref. [31]	Combined UK Biobank, ADSP, IGAP, PGC-ALZ and deCODE	Histone acetyltransferase. Regulates the outcome of autophagy.
		rs889555; ref. [37]	Combined UK Biobank and IGAP	
57	*SPPL2A*	rs59685680; ref. [38]	Combined UK Biobank and IGAP	Signal peptide peptidase-like 2A, related to TNF signaling (REACTOME) and signaling by GPCR. May play a role in the regulation of innate and adaptive immunity.
58	*HESX1**	rs184384746; ref. [31]	Combined UK Biobank, ADSP, IGAP, PGC-ALZ and deCODE	Homeobox protein that is a transcriptional repressor.

Table differentiates between loci reaching GWS in AD diagnosed cohorts (loci numbered 1–43) and loci reaching GWS when AD diagnosed cohorts are meta-analyzed with UK Biobank proxy diagnosed cohorts (loci numbered 44–52). Datasets: Fundació ACE, a non-profit entity at the service of people with Alzheimer's disease or other dementias and their caregivers (https://www.fundacioace.com/en); Alzheimer's disease sequencing project (ADSP); Psychiatric Genomics Consortium Alzheimer's disease working group (PGC-ALZ) Genome Research at Fundació ACE (GR@CE); Dementia Genetics Spanish Consortium (DEGESCO); deCODE, a private corporation (https://www.decode.com); + indicates the sample consortium but with additional samples. # Genome-wide significant association only seen in *APOE* ε4 analysis. * Indicates SNPs and loci with missing data in AD case–control datasets. ** Indicates SNPs and loci with minimal support ($P = 10^{-2}$ genome-wide significance level) in AD case–control datasets.

the PRS and the rare *TREM2* p.(His47) variant[63]. In fact, *TREM2* p.(His47) carriers show a 23% faster rate of decline compared with non-variant carriers.

Comorbid traits. Epidemiological observations of shared comorbidity in twin and family studies have long provided evidence for genetic correlations among diseases[64], as has the co-occurrence of

multiple diseases in the same individual[65]. The advent of GWAS allowed, for the first-time, systematic cross-phenotype analyses, with a significant number of traits sharing genetic architecture[66]. Indeed, up to 4.6% of SNPs and 16.9% of genes have cross-phenotype associations[67]. In dementia, initial work shows that AD and Parkinson's disease (PD) are genetically distinct[68], but that dementia with Lewy bodies (DLB) is correlated to both AD and PD[69,70]. Work by the Brainstorm consortium attempted to quantify the degree of overlap in genetic risk factors of 25 common brain disorders including AD[29] and a range of behavioral–cognitive phenotypes. While AD shows no significant evidence of correlation with psychiatric or neurological traits, strong negative correlations with college attainment, years of education and intelligence are observed. AD and some aspects of cardiovascular disease also share common risk variants[71]. We are now in the era where sufficiently powered genome-wide datasets are available to extend these sophisticated analyses to a range of phenotypes and sub-phenotypes seen to overlap traditional diagnostic boundaries.

Functional genomics

The progression from genetic association to biological mechanism poses a significant challenge to exploiting the findings of GWAS in the development of new therapies. This is, in part, due to the location of the majority of risk variants in non-coding elements of the genome. Combined with the polygenic nature of many diseases, it is clear that analytical approaches that combine multiple data types are required to assist in this translation.

Pathway analysis. The identification of many risk genes suggests commonalities or convergence in function. As with studies of gene expression, 'pathway' analysis methods have been developed for genomic association data that aim to identify, in general, an excess of association signal in sets of genes based on independent annotations (for example, ALIGATOR[72], INRICH[73] and MAGMA[74]). They often incorporate risk loci that fall below the traditional genome-wide significant threshold, and they can therefore offer insights into risk mechanisms beyond select loci, capturing the maximum amount of genetic association information available. Application of these methods to AD GWAS results has been particularly powerful in identifying disease-relevant processes. Indeed, these approaches provide some of the first convincing genetic evidence that the immune system contributes to AD risk (Fig. 2)[75]. Other pathways implicated include endocytosis, cholesterol metabolism, ubiquitination and, more recently, Aβ clearance and tau biology (Fig. 2)[42].

Gene expression. In parallel with the identification of risk variants by GWAS, the genetic control of gene expression has been investigated using studies of expression quantitative trait loci (eQTLs)[76]. These studies aim to link specific variants with levels of gene expression, often across multiple tissue types and cellular contexts. As such, they are a powerful tool for investigating the relationship between genetic disease risk and gene expression and for linking non-coding variants to target genes. Analytical advances such as transcriptome-wide association studies[77] and PrediXcan[78] will also be useful for linking risk alleles to gene expression mechanisms, and have recently been applied to AD GWAS to identify genetically mediated changes in brain mRNA splicing[79]. Although many resources for brain tissue exist[80–82] and continue to be enhanced with increasing cellular and developmental resolution, a striking overlap of AD risk variants and eQTLs in monocytes from blood has also been reported[83]. eQTL studies represent the gold standard for linking variants to gene expression changes, but they require multiple donors with matched genotype and RNA-expression measurements. The sample sizes often range from hundreds to thousands, making them expensive and difficult to perform on hard-to-isolate cell types. In contrast, gene expression measurements from a small

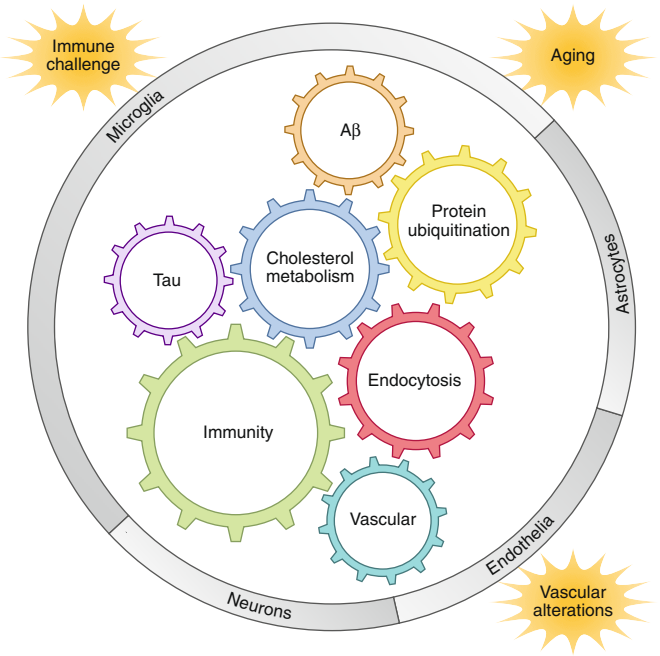

Fig. 2 | Multiplex model of Alzheimer's disease. Pathways associated in the etiology of LOAD by genetic analysis, represented as cogs within the cell, the cell types implicated in disease and the environmental cues thought to directly affect biology.

number of samples have also proven useful in linking putative risk genes to specific cell types. For AD, the integration of GWAS results with cell types identified from single-cell RNA-sequencing of brain tissue has highlighted microglia as the most enriched cell type[84,85]; although the causal relationship is less clear with these studies than with eQTLs, they again support the role of immune cells in AD.

Tissue specificity. In addition to identifying likely causal cell types, the application of single-cell technologies to heterogeneous tissues will help resolve different cellular states. This is particularly important for cells of the immune system, which are known to rapidly respond to environmental cues and to adopt long-lasting 'activation' states. Indeed, recent single-cell RNA-sequencing (scRNA-seq) profiling of microglia from mouse models of AD has identified a subset of 'disease-associated microglia'[86]. Distinct microglia subsets, based on scRNA-seq, have also been identified during normal development and in response to injury[87,88]. Identification of the molecular and environmental regulators of these cells states open up new opportunities for the manipulation of microglia function. Likewise, the influence of AD associated variants and genes on the transition between these states may have important consequences for understanding and treating the disease. Beyond measurements of gene expression, single-cell omics technologies are now capable of interrogating the chromatin landscape[89,90], DNA methylation[91] and targeted protein abundances[92,93]. The availability of increasingly high-resolution data on cell types of interest (for example, microglia) promises to refine these findings further[94]. Finally, convergence between genes at genetic risk loci and molecular system level changes in aged or diseased brains suggest that AD risk genes operate in pathways relevant to pathology[95,96], including those that change expression in response to Aβ accumulation[97].

Epigenome. The gene-regulatory mechanisms underlying eQTLs and non-coding risk variants are often poorly understood, but our knowledge of the gene-regulatory landscape (the epigenome) of cell types is rapidly expanding with the advent of genome-wide

sequencing applications such as chromatin immunoprecipitation and sequencing (ChIP–seq). These assays are able to provide genome-wide profiles of regulatory features based on histone modifications, the binding of individual transcription factors, and biophysical properties such as open chromatin. Integration of these data types with GWAS findings can provide insights into risk mechanisms at individual loci, as well as identify cell types in which multiple loci operate. For AD GWAS, integrative analyses with gene-regulatory elements have identified immune cell types, particularly monocytes, as likely effectors of risk at genome-wide significant loci[98–100] and are starting to identify functional variants underling risk-associated eQTLs[101]. These approaches have been extended with methods such as stratified linkage disequilibrium (LD) score regression[102] to partition the heritability by gene-regulatory elements from different cell types. Again, SNPs located in immune cell types, including microglia, are the most enriched[100,103,104]. Recently, these approaches have been used with gene-regulatory information from human microglia[105] to increase the resolution from cell type to transcription-factor cistromes. Tansey et al. identified an enrichment of genome-wide significant AD risk variants within a particular microglial–macrophage motif containing DNA elements[104]; these sites were also enriched for AD common variant heritability. Amongst these enriched cistromes were those targeted by PU.1 (encoded by *SPI1*) and MEF2 (encoded by *MEF2C*). Interestingly, both *SPI1* and *MEF2C* have been identified as AD risk or onset-modifying loci[26,35,103,106]. These findings suggest that common-variant AD risk operates through transcriptional networks controlled by other AD risk genes that act as 'hubs'. Such genes have also been referred to as 'peripheral master regulators'[107]. Through coordinated regulation of other risk genes, they could provide important avenues into trait biology.

It is noteworthy that the majority of human functional genomic data produced to date uses postmortem tissue and therefore poorly captures dynamic changes in gene regulation (for example, during development or response to an environmental challenge). To address this, collections of induced pluripotent stem cells from genotyped individuals are being generated to explore the genetic control of context-specific gene expression[108].

Somatic mutations

Single-cell technologies are also being used to probe heterogeneity in cellular DNA content and sequence. These postzygotic changes are known as somatic mutations. Studies of somatic mutation in the brain are in their infancy. Nevertheless, they do occur in healthy brain tissue, resulting in mosaicism[109,110]. Damaging mutations can therefore occur and accumulate in a subset of cells, resulting in restricted cell-type consequences, including vulnerability to cell death[111]. Whole-genome approaches to single-cell DNA content are largely restricted by the cost of obtaining sufficient sequencing coverage for reliable quantification. However, targeted approaches have identified changes in *APP* copy number in cells from AD brain samples compared to controls[112], as well as *APP* recombination events that result in the insertion of known disease-causing *APP* mutations into the genome of individual neurons[113]. The general importance of this type of mutation is still to be quantified, and it should be noted that they do not contribute to the observed heritability of the AD. They are therefore likely to operate in conjunction with common-variant risk factors.

Risk prediction

The finding that LOAD is the result of the combined influences of multiple genetic loci or polygenic effects, and that these effects can be captured in one algorithm, has enormous utility in the field. While *APOE* has an established role as the strongest single genetic risk factor for sporadic AD, it is neither necessary nor sufficient to cause disease. The effect estimates of the other associated risk loci range from an odds ratio of approximately 1.1 to 2.1 for each disease-associated allele, meaning their individual contribution to disease risk is relatively small. However, the cumulative effect of these susceptibility loci can be captured by PRS analysis. This takes advantage of all relevant association information and thus captures most of the variance of GWAS studies, including true genetic risk loci that are hypothesized to lie below the genome-wide significance threshold. This approach is supported by the observed increase in explained heritability when weak effect loci are also considered[114].

Early work showed that AD is a polygenic disease ($P = 4.9 \times 10^{-26}$)[115], an enrichment that remains significant after *APOE* and other genome-wide associated regions are excluded ($P = 3.4 \times 10^{-19}$). Escott-Price and colleagues created the Cardiff PRS (CPRS) from 17,008 AD cases and 37,154 controls taken from the IGAP dataset[26]. Using an association cut-off of $P < 0.5$ they produced an algorithm based upon over 87,000 variants, incorporating age and sex, which showed an area under the curve (AUC) of 0.78, indicating that this CPRS could correctly classify cases and controls 78% of the time. The predictive utility of CPRS has now been validated in a number of independent datasets[116], with the predictive accuracy of disease status reaching 84% in neuropathologically confirmed AD samples, 82% in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort (Fig. 3) and extending to over 90% sensitivity (correctly identifying AD cases) in those individuals showing the greatest genetic risk for disease.

Notably, those with a CPRS 2 s.d. above the mean (extreme PRS) showed highly accurate prediction of AD diagnosis (Fig. 3). Thus, the CPRS could identify those at high risk early in life. This facility could transform our understanding of the first stages of disease and also provide a means to develop high- and low-risk stem cell models to explore disease mechanisms in human systems. Interestingly, the ADNI dataset also showed that *APOE* was just as good as CPRS at detecting individuals with amyloid plaque deposition. However, only 62% of these went on to develop AD and CPRS; extreme CPRS still remained the best predictor of amyloid deposition within AD. The genetic heritability explained by *APOE* and the genome-wide significant loci is not high ($h^2 = 5.1\%$)[117], as compared to genome-wide estimates ($h^2 = 24–53\%$)[40,114,118]. The CPRS[115] shows prediction accuracy of AUC = 75–84% (compared to AUC = 66% for *APOE* and GWS loci[119]) in clinical and pathology-confirmed samples, respectively[115,120]. These AUC estimates are very close to the maximum prediction accuracy that can be achieved on the basis of SNP-based heritability captured by the whole genome[117] and can be used for AD risk prediction with more confidence. If used in the general population, the majority of people will gain little from CPRS, but those with extreme CPRS will have a high degree of confidence that they either will develop or never develop AD.

Current research is exploring the utility of using CPRS calculated for the biological pathways implicated in AD, enabling participant stratification for preventative and clinical trials of relevant targets and potentially for precision medicine. Initial work assessing the cumulative risk of 20 AD associated risk variants categorized by biological pathway suggests that the clinical model of early AD pathology is explained by different biological pathways[121]. In particular, the endocytosis pathway shows relevance in subjects with mild cognitive impairment[121]. Development of full PRS models for each AD-implicated pathway are now needed to improve the quality of pathway-specific genetic scores that could feed into future research, including clinical trials of drugs targeting relevant pathways. Targeted drugs will also need pathway-specific biomarkers and drug trials that possibly move away from full disease measures to define outcomes, with the consequence of a likely reduction in timescale and cost. Identifying individuals at high or low risk of developing AD will also allow better understanding of the earliest signs of disease, develop appropriate biomarkers through imaging and bio-sampling, and help test for relationships with environmental

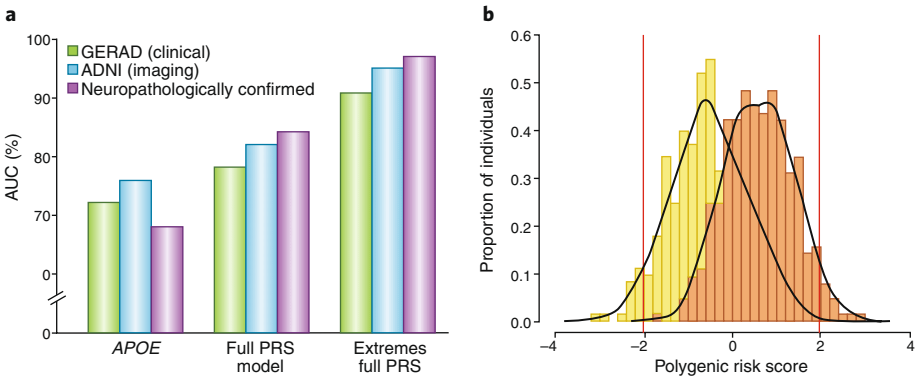

Fig. 3 | Polygenic risk score (PRS) in Alzheimer's disease. a, The accuracy of classifying AD and control status (area under the curve, AUC) using *APOE*, full, and extreme PRS classed as 2 s.d. above or below the mean across clinically diagnosed (GERAD)[21], neuroimaging-positive (ADNI; http://adni.loni.usc.edu/) and neuropathically confirmed cohorts[119]. **b**, The distribution of PRS, highlighting values 2 s.d. above or below the mean, indicated by the red lines. The yellow histogram bars represent controls and the orange histogram bars represent AD cases. The separate normal distributions for AD cases and controls are depicted by the black lines.

factors that may interact with genetics to delay or exasperate disease mechanisms.

Neuroimaging approaches offer insight into AD pathogenesis in vivo, demonstrating how the combined impact of AD risk genes are associated with altered brain physiology[122–125]. Accumulating evidence suggests that AD GWAS risk alleles influence brain structure and function in asymptomatic individuals. Early studies showcase the potential roles of individual GWAS AD SNPs on brain structure and function[126,127]; however, recent work now assesses the impact of the cumulative impact of AD-risk SNPs through CPRS. These studies have primarily focused on putatively AD-susceptible brain regions, such as medial temporal lobe macrostructure (hippocampal formation; amygdala) and other in vivo biomarkers of AD pathology such as $A\beta_{42}$ deposition[128–132]. Collectively, these observations suggest that an excessive burden of AD risk alleles may compromise brain health in individuals years before the onset of clinical symptoms. These hypotheses are further supported by large GWAS of neuroimaging data, demonstrating genetic correlations between AD and markers of brain health, such as subcortical brain volumes[133,134]. However, the extended impact of AD PRS on the brain remains relatively unknown. This is largely due to constraints such as the need to do multisequence, multimodal imaging in large sample sizes and constraints intrinsic to harmonization procedures[135]. Initial evidence from a middle-aged population cohort (UK Biobank) does suggest relationships between CPRS and surface areas of the frontal and cingulate cortex, specifically with the anterior cingulate for the microglia-mediated innate-immunity PRS[136].

Genetic modelling and disease mechanisms

Establishing animal and cellular models of AD mutations or functional coding variants is now routine. Several transgenic mouse models have been developed using AD mutations in *APP* and the presenilin genes, but none recapitulate the full profile of the disease as seen in humans[137]. While they do show accumulation of Aβ peptide in the brain and cognitive deficits, they rarely show AD associated cell death or tau dysfunction (unless tau mutations are also introduced). It is noteworthy that rodents do not naturally develop AD and that human-based manipulations are necessary to produce AD relevant changes. If, as the genetics of common forms of AD suggests, the disease requires multiple components to change to trigger AD, then it is not surprising that transgenic models of single AD components do not reflect full blown AD. Transgenic models of *APOE* are less numerous, but have shown interesting results when crossed with *APP* transgenic models[137]. *APOE* is shown to influence Aβ aggregation and clearance from the brain, although other outcomes are now the focus of new research. Indeed, as many more *APOE* models are being produced (MODEL-AD) we will soon see a much broader capture of the AD phenotype. *Drosophila* models of AD have also been the source of much research and benefit from the speed at which results and manipulations can be achieved. *Drosophila* have low redundancy, which simplifies the analysis of gene disruption. Early work focused on *APP* and tau models, but lately the models are facilitating the screening of GWAS susceptibility genes[138].

Stem-cell-derived models of AD genetic risk variants have, understandably, focused on rare variants, particularly those that cause familial AD. Lines derived from AD cases carrying *PSEN1* and *PSEN2* mutations were the first to be investigated[139], followed by *APP* duplications[140]. More recently, *TREM2* variant and null lines have been developed and used to investigate AD-related microglial function[141,142]. Common-variant stem cell models have generally lagged behind, with only a small number of target models developed, most notably *APOE*[143,144]. Only three loci identified exclusively by GWAS, *PICALM*[145], *CLU*[146] and *PLCG2*, have been used for stem cell models. Of these, only the *PICALM* and *PLCG2* models are based on a likely causal variant, highlighting the challenges of moving from GWAS association to a cell model. These models have begun to identify important AD-relevant biology. For example, neurons deficient in *CLU* protein are resistant to neurodegeneration in response to Aβ risk, and altered *PICLAM* expression manifests as disrupted transcytosis of Aβ by endothelial cells. Models of *APOE*-mediated risk have identified multiple dysregulated processes across different stem-cell-derived cell types, for example, diminished neurotrophic function of *APOE* ε4/ε4 astrocytes[143], differential activation of neuronal *APP* transcription and Aβ-synthesis by glial APOE isoforms[147], and altered Aβ aggregates and hyperphosphorylation of tau in organoids[144]. However, none, as they stand, recapitulate all aspects of the human AD[148].

The recent advances in identifying multiple genetic risk factors for AD, as described above, open new avenues for disease modelling. Specifically, they create the ability to construct induced pluripotent stem cells (iPSC) from individuals with high or low PRS for AD or its component pathways, thereby creating resources which capture multiple disease factors in the same cells. However, there are challenges with these approaches. Individuals selected for high PRS will vary in other ways that could influence outcomes. Accordingly, studies involving many different iPSC donors will be needed to overcome this natural variation and identify the disease

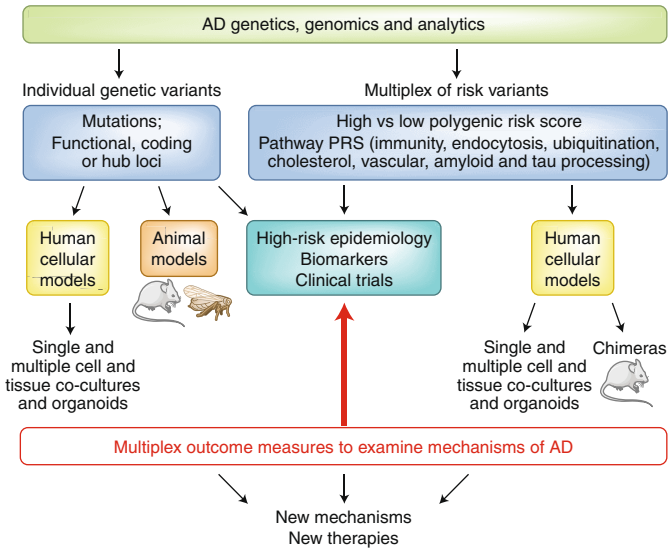

Fig. 4 | Schematic demonstrating the complexity and methods of discovering Alzheimer's disease mechanisms. Identifying and understanding the mechanisms involved in Alzheimer's disease etiology and pathology paves the way for new therapeutic targets and new therapies.

relevant consistencies. Thus, future research could combine information from models of specific known AD variants of *APP*, presenilin, *APOE*, *TREM2*, *PLCG2* and *ABI3* (for example) with the outcomes of high and low PRS models to triangulate disease relevant mechanisms (Fig. 4). The multiplex model of AD (see below) also has implications for what are tested as outcome measures. Recent advances in cellular approaches such as scRNA-seq, three-dimensional (3D) cultures, organoids of neurons, glia and epithelial cells, and the ability to transplant cells into the brains of mice to form chimeras will undoubtedly expand our knowledge of disease mechanisms underpinning the AD model.

The multiplex model of Alzheimer's disease

The multiplex model builds on evidence we observe from genetic, and for that matter, environmental studies of AD. Genetics show us that AD is a multicomponent disease and that deficits combine additively to trigger disease. There is strong evidence for changes in immunity and inflammation, Aβ production and clearance, endocytosis, ubiquitination, cholesterol, and tau processing. Epidemiological research (not reviewed here) also highlights a significant vascular component to AD development (Fig. 2)[149]. The multiplex model assumes that changes to some or all of these model components act together to trigger a disease cascade, which ultimately results in the cell and synaptic loss observed in AD. AD could be triggered by a number of different patterns of deficits that may differ between tissues and over the course of disease development. Indeed, in time we may characterize AD as several diseases. However, until we understand the specific biological mechanisms that underlie the model, it is beneficial to continue viewing AD as a single disease. As we learn more, we will refine the model. For example, we already have evidence that endocytosis could affect Aβ clearance[145]. However, with current knowledge there is simply not enough evidence to show that they pinpoint the same disease mechanism. It is also assumed that the liability threshold for disease could result from extreme loading on a limited number of components or indeed, moderate vulnerability across multiple components. Future treatments and preventative approaches may focus on one or multiple AD components, which may also change over the course of disease development. The multiplex model of AD encourages future research to focus on a broader range of outcome measures to understand disease mechanisms and identify several new targets for treatments, and it may ultimately change the way we diagnose AD.

Conclusions

It is now well-established that drug trials based on evidence with a genetic basis are more likely to succeed[150]. Thus, using this well-replicated robust biological evidence for future research into disease mechanisms and therapies seems the logical step. A variety of genome-wide approaches have already identified over 50 loci associated with AD at a genome-wide level of significance. Pathway and functional genomic analyses have shown strong patterns of susceptibility implicating immunity, endocytosis, cholesterol transport, ubiquitination, Aβ and tau processing and have highlighted several hub genes of significant influence. Using most of the information from GWAS data, accounting for up to 50% of heritability, PRS can be calculated which show around 80% accuracy in predicting AD in a variety of independent datasets. Moreover, selecting individuals at the polygenic extremes achieves sensitivities of over 90% for the detection of AD cases. Applications of overall and AD-pathway-specific CPRS to future research could include selection and enrichment for clinical trials and precision medicine, understanding of early disease development through risk related epidemiology, selective biomarkers and iPSC models of PRS risk for single-cell, multitissue, multi-organoid and whole-system chimeric analyses. Combine this with the growing multi-omic approaches now available and it is clear our understanding of this complex disease will advance at a considerable pace. Genetic studies have changed our perception of AD, highlighting its multifactorial complexity. Building on these findings, together with the role of vascular factors implicated by epidemiology, we propose the multiplex model as a way of integrating evidence from several domains to support our understanding of Alzheimer's disease.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-020-0599-5.

Received: 18 October 2018; Accepted: 24 January 2020;
Published online: 28 February 2020

References

1. Goate, A. et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* **349**, 704–706 (1991).
2. Sherrington, R. et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* **375**, 754–760 (1995).
3. Rogaev, E. I. et al. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature* **376**, 775–778 (1995).
4. Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297**, 353–356 (2002).
5. Ricciarelli, R. & Fedele, E. The amyloid cascade hypothesis in Alzheimer's disease: it's time to change our mind. *Curr. Neuropharmacol.* **15**, 926–935 (2017).
6. Doody, R. S., Farlow, M. & Aisen, P. S., Alzheimer's Disease Cooperative Study Data Analysis and Publication Committee. Phase 3 trials of solanezumab and bapineuzumab for Alzheimer's disease. *N. Engl. J. Med.* **370**, 1460 (2014).
7. Honig, L. S. et al. Trial of solanezumab for mild dementia due to Alzheimer's disease. *N. Engl. J. Med.* **378**, 321–330 (2018).
8. Galimberti, D. & Scarpini, E. Disease-modifying treatments for Alzheimer's disease. *Ther. Adv. Neurol. Disord.* **4**, 203–216 (2011).
9. Yiannopoulou, K. G. & Papageorgiou, S. G. Current and future treatments for Alzheimer's disease. *Ther. Adv. Neurol. Disord.* **6**, 19–33 (2013).
10. Jagust, W. Imaging the evolution and pathophysiology of Alzheimer disease. *Nat. Rev. Neurosci.* **19**, 687–700 (2018).

11. Rajan, K. B., Wilson, R. S., Weuve, J., Barnes, L. L. & Evans, D. A. Cognitive impairment 18 years before clinical diagnosis of Alzheimer disease dementia. *Neurology* **85**, 898–904 (2015).
12. Rosenblum, W. I. Why Alzheimer trials fail: removing soluble oligomeric beta amyloid is essential, inconsistent, and difficult. *Neurobiol. Aging* **35**, 969–974 (2014).
13. Gatz, M. et al. Role of genes and environments for explaining Alzheimer disease. *Arch. Gen. Psychiatry* **63**, 168–174 (2006).
14. Wingo, T. S., Lah, J. J., Levey, A. I. & Cutler, D. J. Autosomal recessive causes likely in early-onset Alzheimer disease. *Arch. Neurol.* **69**, 59–64 (2012).
15. Saunders, A. M. et al. *Apolipoprotein* E epsilon 4 allele distributions in late-onset Alzheimer's disease and in other amyloid-forming diseases. *Lancet* **342**, 710–711 (1993).
16. Strittmatter, W. J. et al. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc. Natl Acad. Sci. USA* **90**, 1977–1981 (1993).
17. Corder, E. H. et al. Protective effect of *apolipoprotein E* type 2 allele for late onset Alzheimer disease. *Nat. Genet.* **7**, 180–184 (1994).
18. Liao, F., Yoon, H. & Kim, J. Apolipoprotein E metabolism and functions in brain and its role in Alzheimer's disease. *Curr. Opin. Lipidol.* **28**, 60–67 (2017).
19. Deane, R. et al. ApoE isoform-specific disruption of amyloid beta peptide clearance from mouse brain. *J. Clin. Invest.* **118**, 4002–4013 (2008).
20. Verghese, P. B., Castellano, J. M. & Holtzman, D. M. Apolipoprotein E in Alzheimer's disease and other neurological disorders. *Lancet Neurol.* **10**, 241–252 (2011).
21. Harold, D. et al. Genome-wide association study identifies variants at *CLU* and *PICALM* associated with Alzheimer's disease. *Nat. Genet.* **41**, 1088–1093 (2009).
22. Lambert, J. C. et al. Genome-wide association study identifies variants at *CLU* and *CR1* associated with Alzheimer's disease. *Nat. Genet.* **41**, 1094–1099 (2009).
23. Seshadri, S. et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. *J. Am. Med. Assoc.* **303**, 1832–1840 (2010).
24. Naj, A. C. et al. Common variants at *MS4A4/MS4A6E*, *CD2AP*, *CD33* and *EPHA1* are associated with late-onset Alzheimer's disease. *Nat. Genet.* **43**, 436–441 (2011).
25. Hollingworth, P. et al. Common variants at *ABCA7*, *MS4A6A/MS4A4E*, *EPHA1*, *CD33* and *CD2AP* are associated with Alzheimer's disease. *Nat. Genet.* **43**, 429–435 (2011).
26. Lambert, J. C. et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).
27. Desikan, R. S. et al. Genetic assessment of age-associated Alzheimer disease risk: development and validation of a polygenic hazard score. *PLoS Med.* **14**, e1002258 (2017).
28. Johnson, E. C. B. et al. Deep proteomic network analysis of Alzheimer's disease brain reveals alterations in RNA binding proteins and RNA splicing associated with disease. *Mol. Neurodegener.* **13**, 52 (2018).
29. Anttila, V. et al. Analysis of shared heritability in common disorders of the brain. *Science* **360**, eaap8757 (2018).
30. Ruiz, A. et al. Follow-up of loci from the International Genomics of Alzheimer's Disease Project identifies *TRIP4* as a novel susceptibility gene. *Transl. Psychiatry* **4**, e358 (2014).
31. Jansen, I. E. et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019).
32. Jun, G. R. et al. Transethnic genome-wide scan identifies novel Alzheimer's disease loci. *Alzheimers Dement.* **13**, 727–738 (2017).
33. Kunkle, B. W. et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019).
34. de Rojas, I., et al. Common variants in Alzheimer's disease: novel association of six genetic variants with AD and risk stratification by polygenic risk scores. Preprint at *medRxiv* https://doi.org/10.1101/19012021 (2019).
35. Escott-Price, V. et al. Gene-wide analysis detects two new susceptibility genes for Alzheimer's disease. *PLoS One* **9**, e94661 (2014).
36. Baker, E. et al. Gene-based analysis in HRC imputed genome wide association data identifies three novel genes for Alzheimer's disease. *PLoS One* **14**, e0218111 (2019).
37. Marioni, R. E. et al. GWAS on family history of Alzheimer's disease. *Transl. Psychiatry* **8**, 99 (2018).
38. Liu, J. Z., Erlich, Y. & Pickrell, J. K. Case-control association mapping by proxy using family history of disease. *Nat. Genet.* **49**, 325–331 (2017).
39. So, H. C., Gui, A. H., Cherny, S. S. & Sham, P. C. Evaluating the heritability explained by known susceptibility variants: a survey of ten complex diseases. *Genet. Epidemiol.* **35**, 310–317 (2011).
40. Ridge, P. G., Mukherjee, S., Crane, P. K. & Kauwe, J. S. Alzheimer's Disease Genetics Consortium. Alzheimer's disease: analyzing the missing heritability. *PLoS One* **8**, e79771 (2013).
41. Meng, W. et al. A genome-wide association study finds genetic associations with broadly-defined headache in UK Biobank (N=223,773). *EBioMedicine* **28**, 180–186 (2018).
42. Kunkle, B. et al. Meta-analysis of genetic association with diagnosed Alzheimer's disease identifies novel risk loci and implicates Abeta, Tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019).
43. Guerreiro, R. et al. *TREM2* variants in Alzheimer's disease. *N. Engl. J. Med.* **368**, 117–127 (2013).
44. Bis, J.C. et al. Whole exome sequencing study identifies novel rare and common Alzheimer's-associated variants involved in immune response and transcriptional regulation. *Mol. Psychiatry* https://doi.org/10.1038/s41380-018-0112-7 (2018).
45. Jonsson, T. et al. Variant of *TREM2* associated with the risk of Alzheimer's disease. *N. Engl. J. Med.* **368**, 107–116 (2013).
46. Jonsson, T. et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature* **488**, 96–99 (2012).
47. Kunkle, B. W. et al. Early-onset Alzheimer disease and candidate risk genes involved in endolysosomal transport. *JAMA Neurol.* **74**, 1113–1122 (2017).
48. Bellenguez, C. et al. Contribution to Alzheimer's disease risk of rare variants in *TREM2*, *SORL1*, and ABCA7 in 1779 cases and 1273 controls. *Neurobiol. Aging* **59**, 220.e1–220.e9 (2017).
49. Louwersheimer, E. et al. Influence of genetic variants in *SORL1* gene on the manifestation of Alzheimer's disease. *Neurobiol. Aging* **36**, 1605.e13–1605.e20 (2015).
50. Steinberg, S. et al. Loss-of-function variants in *ABCA7* confer risk of Alzheimer's disease. *Nat. Genet.* **47**, 445–447 (2015).
51. Cruchaga, C. et al. Rare coding variants in the *phospholipase D3* gene confer risk for Alzheimer's disease. *Nature* **505**, 550–554 (2014).
52. Logue, M. W. et al. Two rare *AKAP9* variants are associated with Alzheimer's disease in African Americans. *Alzheimers Dement.* **10**, 609–618.e11 (2014).
53. Wetzel-Smith, M. K. et al. A rare mutation in *UNC5C* predisposes to late-onset Alzheimer's disease and increases neuronal cell death. *Nat. Med.* **20**, 1452–1457 (2014).
54. Sims, R. et al. Rare coding variants in *PLCG2*, *ABI3*, and *TREM2* implicate microglial-mediated innate immunity in Alzheimer's disease. *Nat. Genet.* **49**, 1373–1384 (2017).
55. Magno, L. et al. Alzheimer's disease phospholipase C-gamma-2 (PLCG2) protective variant is a functional hypermorph. *Alzheimers Res. Ther.* **11**, 16 (2019).
56. Gusareva, E. S. et al. Male-specific epistasis between *WWC1* and *TLN2* genes is associated with Alzheimer's disease. *Neurobiol. Aging* **72**, 188.e3–188.e12 (2018).
57. Ridge, P. G. et al. Linkage, whole genome sequence, and biological data implicate variants in *RAB10* in Alzheimer's disease resilience. *Genome Med.* **9**, 100 (2017).
58. Ropacki, S. A. & Jeste, D. V. Epidemiology of and risk factors for psychosis of Alzheimer's disease: a review of 55 studies published from 1990 to 2003. *Am. J. Psychiatry* **162**, 2022–2030 (2005).
59. Shin, I. S., Carter, M., Masterman, D., Fairbanks, L. & Cummings, J. L. Neuropsychiatric symptoms and quality of life in Alzheimer disease. *Am. J. Geriatr. Psychiatry* **13**, 469–474 (2005).
60. Wilkosz, P. A. et al. Trajectories of cognitive decline in Alzheimer's disease. *Int. Psychogeriatr.* **22**, 281–290 (2010).
61. Lopez, O. L. et al. Long-term effects of the concomitant use of memantine with cholinesterase inhibition in Alzheimer disease. *J. Neurol. Neurosurg. Psychiatry* **80**, 600–607 (2009).
62. Hollingworth, P. et al. Genome-wide association study of Alzheimer's disease with psychotic symptoms. *Mol. Psychiatry* **17**, 1316–1327 (2012).
63. Del-Aguila, J. L. et al. Assessment of the genetic architecture of Alzheimer's disease risk in rate of memory decline. *J. Alzheimers Dis.* **62**, 745–756 (2018).
64. Criswell, L. A. et al. Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the *PTPN22* 620W allele associates with multiple autoimmune phenotypes. *Am. J. Hum. Genet.* **76**, 561–571 (2005).
65. Eaton, W. W., Rose, N. R., Kalaydjian, A., Pedersen, M. G. & Mortensen, P. B. Epidemiology of autoimmune diseases in Denmark. *J. Autoimmun.* **29**, 1–9 (2007).
66. Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* **14**, 483–495 (2013).
67. Sivakumaran, S. et al. Abundant pleiotropy in human complex diseases and traits. *Am. J. Hum. Genet.* **89**, 607–618 (2011).
68. Moskvina, V. et al. Analysis of genome-wide association studies of Alzheimer disease and of Parkinson disease to determine if these 2 diseases share a common genetic risk. *JAMA Neurol.* **70**, 1268–1276 (2013).
69. Guerreiro, R. et al. Genome-wide analysis of genetic correlation in dementia with Lewy bodies, Parkinson's and Alzheimer's diseases. *Neurobiol. Aging* **38**, 214.e7–214.e10 (2016).

70. Geiger, J. T. et al. Next-generation sequencing reveals substantial genetic contribution to dementia with Lewy bodies. *Neurobiol. Dis.* **94**, 55–62 (2016).
71. Liu, G. et al. Cardiovascular disease contributes to Alzheimer's disease: evidence from large-scale genome-wide association studies. *Neurobiol. Aging* **35**, 786–792 (2014).
72. Holmans, P. et al. Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. *Am. J. Hum. Genet.* **85**, 13–24 (2009).
73. Lee, P. H., O'Dushlaine, C., Thomas, B. & Purcell, S. M. INRICH: interval-based enrichment analysis for genome-wide association studies. *Bioinformatics* **28**, 1797–1799 (2012).
74. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLOS Comput. Biol.* **11**, e1004219 (2015).
75. International Genomics of Alzheimer's Disease Consortium (IGAP). Convergent genetic and expression data implicate immunity in Alzheimer's disease. *Alzheimers Dement.* **11**, 658–671 (2015).
76. Yu, C. H., Pal, L. R. & Moult, J. Consensus genome-wide expression quantitative trait loci and their relationship with human complex trait disease. *OMICS* **20**, 400–414 (2016).
77. Gusev, A. et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
78. Gamazon, E. R. et al. A gene-based association method for mapping traits using reference transcriptome data. *Nat. Genet.* **47**, 1091–1098 (2015).
79. Raj, T. et al. Integrative transcriptome analyses of the aging brain implicate altered splicing in Alzheimer's disease susceptibility. *Nat. Genet.* **50**, 1584–1592 (2018).
80. Ramasamy, A. et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat. Neurosci.* **17**, 1418–1428 (2014).
81. Akbarian, S. et al. The PsychENCODE project. *Nat. Neurosci.* **18**, 1707–1712 (2015).
82. Fromer, M. et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.* **19**, 1442–1453 (2016).
83. Raj, T. et al. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. *Science* **344**, 519–523 (2014).
84. Skene, N. G. & Grant, S. G. Identification of vulnerable cell types in major brain disorders using single cell transcriptomes and expression weighted cell type enrichment. *Front. Neurosci.* **10**, 16 (2016).
85. Skene, N. G. et al. Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).
86. Keren-Shaul, H. et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* **169**, 1276–1290.e17 (2017).
87. Hammond, T. R. et al. Single-cell RNA sequencing of microglia throughout the mouse lifespan and in the injured brain reveals complex cell-state changes. *Immunity* **50**, 253–271.e6 (2019).
88. Li, Q. et al. Developmental heterogeneity of microglia and brain myeloid cells revealed by deep single-cell RNA sequencing. *Neuron* **101**, 207–223. e10 (2019).
89. Buenrostro, J. D. et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486–490 (2015).
90. Cusanovich, D. A. et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910–914 (2015).
91. Farlik, M. et al. Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. *Cell Rep.* **10**, 1386–1397 (2015).
92. Peterson, V. M. et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* **35**, 936–939 (2017).
93. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
94. Olah, M. et al. A single cell-based atlas of human microglial states reveals associations with neurological disorders and histopathological features of the aging brain. Preprint at *bioRxiv* https://doi.org/10.1101/343780 (2018).
95. Mostafavi, S. et al. A molecular network of the aging human brain provides insights into the pathology and cognitive decline of Alzheimer's disease. *Nat. Neurosci.* **21**, 811–819 (2018).
96. Zhang, B. et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* **153**, 707–720 (2013).
97. Salih, D.A. et al. Genetic variability in response to amyloid beta deposition influences Alzheimer's risk. *Brain Commun.* **1**, fcz022 (2019).
98. Gjoneska, E. et al. Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature* **518**, 365–369 (2015).
99. Lu, Q. et al. Systematic tissue-specific functional annotation of the human genome highlights immune-related DNA elements for late-onset Alzheimer's disease. *PLoS Genet.* **13**, e1006933 (2017).
100. Novikova, G. et al. Integration of Alzheimer's disease genetics and myeloid cell genomics identifies novel causal variants, regulatory elements, genes and pathways. Preprint at *bioRxiv* https://doi.org/10.1101/694281 (2019).
101. Amlie-Wolf, A. et al. Inferring the molecular mechanisms of noncoding Alzheimer's disease-associated genetic variants. *J. Alzheimers Dis.* **72**, 301–318 (2019).
102. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
103. Huang, K. L. et al. A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. *Nat. Neurosci.* **20**, 1052–1061 (2017).
104. Tansey, K. E., Cameron, D. & Hill, M. J. Genetic risk for Alzheimer's disease is concentrated in specific macrophage and microglial transcriptional networks. *Genome Med.* **10**, 14 (2018).
105. Gosselin, D. et al. An environment-dependent transcriptional network specifies human microglia identity. *Science* **356**, eaal3222 (2017).
106. Matarin, M. et al. A genome-wide gene-expression analysis and database in transgenic mice during development of amyloid or tau pathology. *Cell Rep.* **10**, 633–644 (2015).
107. Liu, X., Li, Y. I. & Pritchard, J. K. Trans effects on gene expression can drive omnigenic inheritance. *Cell* **177**, 1022–1034.e6 (2019).
108. Alasoo, K. et al. Shared genetic effects on chromatin and gene expression indicate a role for enhancer priming in immune response. *Nat. Genet.* **50**, 424–431 (2018).
109. Lodato, M. A. et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* **350**, 94–98 (2015).
110. McConnell, M. J. et al. Intersection of diverse neuronal genomes and neuropsychiatric disease: the brain somatic mosaicism network. *Science* **356**, eaal1641 (2017).
111. Arendt, T., Brückner, M. K., Mosch, B. & Lösche, A. Selective cell death of hyperploid neurons in Alzheimer's disease. *Am. J. Pathol.* **177**, 15–20 (2010).
112. Bushman, D. M. et al. Genomic mosaicism with increased amyloid precursor protein (*APP*) gene copy number in single neurons from sporadic Alzheimer's disease brains. *eLife* **4**, e05116 (2015).
113. Lee, M. H. et al. Somatic APP gene recombination in Alzheimer's disease and normal neurons. *Nature* **563**, 639–645 (2018).
114. Lee, S. H. et al. Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. *Hum. Mol. Genet.* **22**, 832–841 (2013).
115. Escott-Price, V. et al. Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain* **138**, 3673–3684 (2015).
116. Escott-Price, V. & Jones, L. Genomic profiling and diagnostic biomarkers in Alzheimer's disease. *Lancet Neurol.* **16**, 582–583 (2017).
117. Escott-Price, V., Shoai, M., Pither, R., Williams, J. & Hardy, J. Polygenic score prediction captures nearly all common genetic risk for Alzheimer's disease. *Neurobiol. Aging* **49**, 214.e7–214.e11 (2017).
118. Ridge, P. G. et al. Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiol. Aging* **41**, 200.e13–200.e20 (2016).
119. Lewis, C. M. & Vassos, E. Prospects for using risk scores in polygenic medicine. *Genome Med.* **9**, 96 (2017).
120. Escott-Price, V., Myers, A. J., Huentelman, M. & Hardy, J. Polygenic risk score analysis of pathologically confirmed Alzheimer disease. *Ann. Neurol.* **82**, 311–314 (2017).
121. Ahmad, S. et al. Disentangling the biological pathways involved in early features of Alzheimer's disease in the Rotterdam Study. *Alzheimers Dement.* **14**, 848–857 (2018).
122. Biffi, A. et al. Genetic variation and neuroimaging measures in Alzheimer disease. *Arch. Neurol.* **67**, 677–685 (2010).
123. Braskie, M. N., Ringman, J. M. & Thompson, P. M. Neuroimaging measures as endophenotypes in Alzheimer's disease. *Int. J. Alzheimers Dis.* **2011**, 490140 (2011).
124. Saykin, A. J. et al. Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: genetics core aims, progress, and plans. *Alzheimers Dement.* **6**, 265–273 (2010).
125. Mormino, E. C. et al. Polygenic risk of Alzheimer disease is associated with early- and late-life processes. *Neurology* **87**, 481–488 (2016).
126. Bralten, J. et al. CR1 genotype is associated with entorhinal cortex volume in young healthy adults. *Neurobiol. Aging* **32**, 2106.e7–2106.e11 (2011).
127. Lancaster, T. M. et al. Alzheimer's disease risk variant in *CLU* is associated with neural inefficiency in healthy individuals. *Alzheimers Dement.* **11**, 1144–1152 (2015).
128. Foley, S. F. et al. Multimodal brain imaging reveals structural differences in Alzheimer's disease polygenic risk carriers: a study in healthy young adults. *Biol. Psychiatry* **81**, 154–161 (2017).
129. Axelrud, L. K. et al. Polygenic risk score for Alzheimer's disease: implications for memory performance and hippocampal volumes in early life. *Am. J. Psychiatry* **175**, 555–563 (2018).
130. Chandler, H. L. et al. Polygenic impact of common genetic risk loci for Alzheimer's disease on cerebral blood flow in young individuals. *Sci. Rep.* **9**, 467 (2019).
131. Lupton, M. K. et al. The effect of increased genetic risk for Alzheimer's disease on hippocampal and amygdala volume. *Neurobiol. Aging* **40**, 68–77 (2016).

132. Chauhan, G. et al. Association of Alzheimer's disease GWAS loci with MRI markers of brain aging. *Neurobiol. Aging* **36**, 1765.e7–1765.e16 (2015).
133. Hibar, D. P. et al. Novel genetic loci associated with hippocampal volume. *Nat. Commun.* **8**, 13624 (2017).
134. Satizabal, C.L. et al. Genetic architecture of subcortical brain structures in 38,851 individuals. *Nat. Genet.* **51**, 1624–1636 (2019).
135. Thompson, P. M. et al. ENIGMA and the individual: predicting factors that affect the brain in 35 countries worldwide. *Neuroimage* **145**, 389–408 (2017). Pt B.
136. Lancaster, T. M., Hill, M. J., Sims, R. & Williams, J. Microglia - mediated immunity partly contributes to the genetic association between Alzheimer's disease and hippocampal volume. *Brain Behav. Immun.* **79**, 267–273 (2019).
137. Esquerda-Canals, G., Montoliu-Gaya, L., Güell-Bosch, J. & Villegas, S. Mouse models of Alzheimer's disease. *J. Alzheimers Dis.* **57**, 1171–1183 (2017).
138. Bouleau, S. & Tricoire, H. *Drosophila* models of Alzheimer's disease: advances, limits, and perspectives. *J. Alzheimers Dis.* **45**, 1015–1038 (2015).
139. Yagi, T. et al. [Modeling familial Alzheimer's disease with induced pluripotent stem cells]. *Rinsho Shinkeigaku* **52**, 1134–1136 (2012).
140. Israel, M. A. et al. Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature* **482**, 216–220 (2012).
141. Garcia-Reitboeck, P. et al. Human induced pluripotent stem cell-derived microglia-like cells harboring *TREM2* missense mutations show specific deficits in phagocytosis. *Cell Rep.* **24**, 2300–2311 (2018).
142. Claes, C. et al. Human stem cell-derived monocytes and microglia-like cells reveal impaired amyloid plaque clearance upon heterozygous or homozygous loss of *TREM2*. *Alzheimers Dement.* **15**, 453–464 (2019).
143. Zhao, J. et al. APOE ε4/ε4 diminishes neurotrophic function of human iPSC-derived astrocytes. *Hum. Mol. Genet.* **26**, 2690–2700 (2017).
144. Lin, Y. T. et al. APOE4 causes widespread molecular and cellular alterations associated with Alzheimer's disease phenotypes in human iPSC-derived brain cell types. *Neuron* **98**, 1141–1154.e7 (2018).
145. Zhao, Z. et al. Central role for *PICALM* in amyloid-β blood-brain barrier transcytosis and clearance. *Nat. Neurosci.* **18**, 978–987 (2015).
146. Robbins, J. P. et al. Clusterin is required for β-amyloid toxicity in human iPSC-derived neurons. *Front. Neurosci.* **12**, 504 (2018).
147. Huang, Y. A., Zhou, B., Wernig, M. & Südhof, T. C. ApoE2, ApoE3, and ApoE4 differentially stimulate *APP* transcription and Aβ secretion. *Cell* **168**, 427–441.e21 (2017).
148. Arber, C., Lovejoy, C. & Wray, S. Stem cell models of Alzheimer's disease: progress and challenges. *Alzheimers Res. Ther.* **9**, 42 (2017).
149. Rius-Pérez, S., Tormos, A. M., Pérez, S. & Taléns-Visconti, R. Vascular pathology: cause or effect in Alzheimer disease? *Neurologia* **33**, 112–120 (2018).
150. Nelson, M. R. et al. The support of human genetic evidence for approved drug indications. *Nat. Genet.* **47**, 856–860 (2015).

Acknowledgements

Cardiff University was supported by the UK Dementia Research Institute at Cardiff (MC_PC_17112), Medical Research Council MR/K013041/1), the Alzheimer's Society, Alzheimer's Research UK (ARUK-NC2018-WAL), Dementia Platform UK (HQR00720), the European Joint Programme for Neurodegenerative Disease (MR/N029402/1), the Welsh Assembly Government, Centre for Ageing & Dementia Research (SGR544:CADR) and a donation from the Moondance Charitable Foundation. We thank T. D. Cushion, G. Leonenko, J. Harwood and B. Lan-Leung for their support preparing this manuscript. Finally, we thank all patients and affected families for their continued generosity and willingness to participate in medical research; without their involvement, there would be no advancement in our knowledge of disease genetics.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to J.W.

Peer review information *Nature Neuroscience* thanks Mina Ryten and Bryan Traynor for their contribution to the peer review of this work.