EXPERT REVIEW   OPEN

Step by step: towards a better understanding of the genetic architecture of Alzheimer's disease

Jean-Charles Lambert [ID][1 ✉], Alfredo Ramirez [ID][2,3,4,5,6], Benjamin Grenier-Boley[1] and Céline Bellenguez [ID][1]

© The Author(s) 2023

Alzheimer's disease (AD) is considered to have a large genetic component. Our knowledge of this component has progressed over the last 10 years, thanks notably to the advent of genome-wide association studies and the establishment of large consortia that make it possible to analyze hundreds of thousands of cases and controls. The characterization of dozens of chromosomal regions associated with the risk of developing AD and (in some loci) the causal genes responsible for the observed disease signal has confirmed the involvement of major pathophysiological pathways (such as amyloid precursor protein metabolism) and opened up new perspectives (such as the central role of microglia and inflammation). Furthermore, large-scale sequencing projects are starting to reveal the major impact of rare variants – even in genes like *APOE* – on the AD risk. This increasingly comprehensive knowledge is now being disseminated through translational research; in particular, the development of genetic risk/polygenic risk scores is helping to identify the subpopulations more at risk or less at risk of developing AD. Although it is difficult to assess the efforts still needed to comprehensively characterize the genetic component of AD, several lines of research can be improved or initiated. Ultimately, genetics (in combination with other biomarkers) might help to redefine the boundaries and relationships between various neurodegenerative diseases.

INTRODUCTION

Understanding the genetic component of Alzheimer's disease (AD) has been and still is a major research challenge. The main goal is to understand the pathophysiological mechanisms involved, and the characterization of the mutations responsible for monogenic forms of AD illustrates the scale of this challenge perfectly. The discovery of pathogenic mutations in the *APP*, *PSEN1* and *PSEN2* genes in the 1990s [1–3] led to the amyloid cascade hypothesis, which has greatly influenced the AD research field for more than three decades [4]. However, the amyloid cascade hypothesis was prompted by studies of a specific, small subset of patients (representing less than 1% of cases) and is now being called partly or wholly into question by the failure of most of the therapeutic approaches developed on this basis [5]. Even though the amyloid cascade hypothesis has been regularly modified to take account of developments in our knowledge of AD [6, 7], it now appears to be too simplistic and does not encompass the complexity and diversity of the pathophysiological processes involved in the common forms of the disease. Defining the genetic component of these common forms is one way of gaining a better understanding of the fundamental disease processes.

In 1993, the first genetic risk factor for common forms of AD was discovered: the ε4 allele of the apolipoprotein E (*APOE*) gene was found to be associated with a 3- to 4-fold increase in the AD risk [8]. A year later, it was reported that the *APOE* ε2 allele was associated with a two-fold decrease in that risk [9] – confirming the major role of the *APOE* gene in AD. Despite numerous efforts, our knowledge of the genetic component of common forms of AD did not extend much beyond *APOE* between 1993 to 2009, due to methodological and technological problems [10]. Eventually, the advent of genomic approaches (including genome-wide association studies (GWASs) and next-generation sequencing) boosted our characterization of the genetics of AD.

Here, we review the latest advances in our knowledge of the genetic landscape of AD, discuss the limitations and issues we are facing, and consider the potential consequences of these genetic findings for research on AD and related forms of dementia.

The last decade has uncovered a new genetic landscape for AD

Following the discovery of *APOE* as a major genetic risk factor for AD, more than 350 genes were selected from 1993 to 2009 on the basis of their potential implication in pathophysiological processes, and were tested in small-scale case-control association studies [11]. *SORL1* and *ACE* are the only well-established genetic risk factors for AD [12, 13] to have emerged from this candidate gene phase. This strategy was thus mostly unsuccessful (due to methodological and technological problems) and generated highly untrustworthy and confusing information for the AD research community [10].

Fortunately, AD genetics research has since greatly benefited from the advent of GWAS techniques. The results of first two

[1]Univ. Lille, Inserm, CHU Lille, Institut Pasteur de Lille, U1167-RID-AGE Facteurs de risque et déterminants moléculaires des maladies liées au vieillissement, Lille, France. [2]Division of Neurogenetics and Molecular Psychiatry, Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany. [3]Department of Neurodegenerative diseases and Geriatric Psychiatry, University Hospital Bonn, Medical Faculty, Bonn, Germany. [4]Department of Psychiatry & Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, San Antonio, TX, USA. [5]German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany. [6]Cluster of Excellence Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Cologne, Germany. ✉email: jean-charles.lambert@pasteur-lille.fr
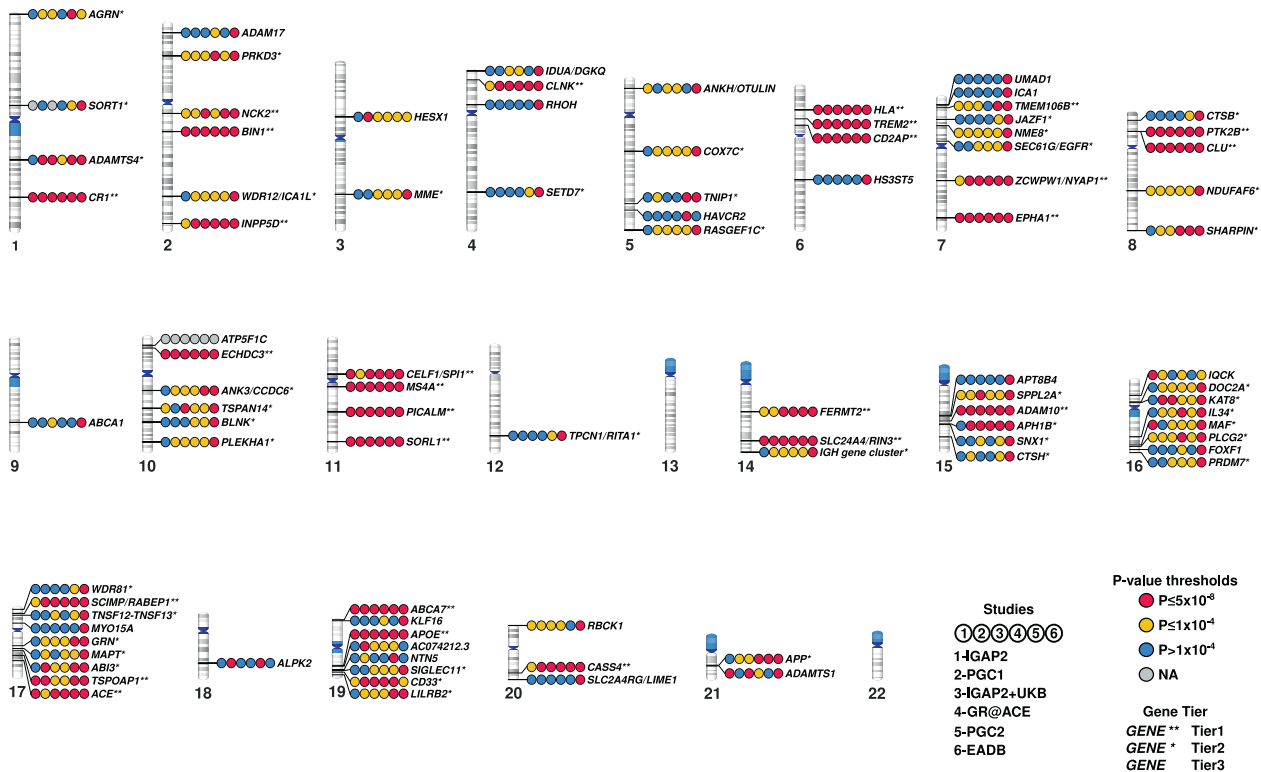
Fig. 1 Ideogram of the 90 loci with genome-wide significance extracted from the GWAS catalog and classified as tier 1, tier 2 or tier 3. For each locus, the figure shows the P-value categories for the association with AD in the six main GWASs published since 2019 (IGAP2, PGC1, IGAP2 + UKB, GR@ACE, PGC2, and EADB, [20–23, 26, 27]): $P \leq 5 \times 10^{-8}$, $P \leq 1 \times 10^{-4}$, $P > 1 \times 10^{-4}$, or NA (not available). See the additional note for details of the methods.

seminal GWAS papers were published in 2009, with the identification of three new loci close to or within the *CLU*, *CR1* and *PICALM* genes [14, 15]. As with many multifactorial diseases, the AD GWAS field has since incorporated (i) meta-analysis methodologies for the facilitated merger of independent GWAS results [16]; (ii) larger numbers of samples (with the International Genomics of Alzheimer Project (IGAP) meta-analysis in 2013 as a milestone [17]; (iii) increasingly powerful reference population panels, which considerably improve the imputation quality of common/rare variants as well as their number for analyses; and (iv) proxy-AD cases in the UK Biobank (UKB, based on self-reports of a family history of AD-related dementia) that has increased the statistical power of the AD GWASs since the end of the 2010s [16, 18–26]. The combination of these various advances ultimately led to a recent publication by the European Alzheimer & Dementia Biobank (EADB) consortium, which reported in addition to *APOE* the association of 75 loci with AD risk, of which 42 were newly identified at that time [27]. In the GWAS catalog, there are currently 97 entries for a locus having a significant genome-wide association with the AD risk in populations of European ancestry (Supplementary Table 1). We have classified these loci as tier 1, tier 2, tier 3, or "not validated", depending on the strength of the association in the six main GWASs performed in populations of European ancestry and published since 2019 [20–23, 26, 27] (see the additional note for a complete description of the criteria used). We classified 27, 41, 22 and 7 loci as tier 1, tier 2, tier 3 and "not validated", respectively (Supplementary Table 1 and Fig. 1). Some of the tier 3 signals (especially those just below the genome-wide significance threshold) might be false positives, and so further validation will be required. However, it is important to bear in mind that the significance threshold is arbitrary – even though it initially corresponded to a Bonferroni correction. Hence, a number of signals of potential interest located just below this threshold have received little attention.

GWASs in populations of non-European ancestry

The great majority of published GWASs in the field of AD were performed in populations of European ancestry, and the few performed in populations of non-European ancestry comprised limited numbers of cases and controls. Hence, comparisons are complicated by (i) the imbalance in statistical power between studies of populations of European ancestry and studies of populations of non-European-ancestry, and (ii) the risk of false positives and false negatives in small studies.

The results of the first (small) GWAS of a population of African-American ancestry were published in 2011 [28], and a GWAS comparable in size to the European studies published in 2009 was performed in 2013 [29]. The latter study found a significant, genome-wide association between a common *ABCA7* variant (which is rare in populations of European ancestry) and the AD risk. This finding emphasized the value of studying genetic diversity in populations with different ancestries, in order to capture additional information. The results of an updated GWAS with a 37% increase in the sample size (but only 826 new AD cases, in fact) were published in 2021. The study confirmed the association between the *ABCA7* variant and the AD risk, found a new signal with genome-wide significance (Supplementary Table 1), and detected nominal associations for six GWAS loci previously identified in populations of European ancestry [30].

In populations of Caribbean ancestry, few GWASs have been conducted and a small case-control study gave negative results [31]. In fact, most studies of these populations are based on linkage analyses of AD families [32]. A few case-control studies are available in central and South American countries with admixed populations [33–35]. However, many initiatives are being developed.

Lastly, although many small case-control studies of populations of East Asian ancestry have been conducted with the attempt to replicate known loci, the results of the first large GWASs of

populations of East-Asian ancestry were only published in 2021: respectively four, two and two new potential loci were identified in Chinese, Japanese and Korean populations (Supplementary Table 1)[36–38].

Next-generation sequencing: a new horizon?

Human genetic variability is primarily due to rare and very rare variants. It was soon suggested that most of the heritability of multifactorial diseases was associated with rare variants. In contrast, GWASs facilitate the analysis of variants that are frequent in the general population (those with a minor allele frequency >1%). Some genotyping chips have been specifically developed to study rare variants [39], and improved imputation panels have also facilitated this research. However, imputation tools are not effective for the vast majority of very rare variants in general and singleton variants in particular. Fortunately, the advent of next-generation sequencing and the marked fall in its cost over the last ten years have enabled the application of this technique to study rare variants in multifactorial diseases, for which large numbers of samples are required.

In 2012, the first applications of whole-exome sequencing (WES) and whole-genome sequencing (WGS)) to AD revealed that very rare/singleton mutations in the SORL1 gene were associated with familial early-onset forms of the disease [40]. The combination of WES/WGS with imputed GWAS data identified an association between the AD risk and the R47H variant in TREM2 (in 2013) [41, 42] and rare variants in ABCA7 (in 2015) [43]. Since then, the associations with these rare variants in SORL1, TREM2 and ABCA7 have been systematically validated and extended as the number of individual sequences has grown [44]. In 2019, the Alzheimer's Disease Sequencing Project published results of WES for 5740 cases of late-onset AD and 5096 cognitively normal controls [45]. Rare variants in several genes were found to be potentially associated with the AD risk. However, the signals observed in this study were not replicated by the Alzheimer Disease European Sequencing (ADES) consortium in the latest and most recent WES study in AD to date (with 12,652 AD cases and 8693 controls in stage one of the study) [46]. Along with ABCA7, TREM2 and SORL1, the ADES study found AD risk signals for rare variants in two new genes (ABCA1 and ATP8B4). It is noteworthy that a suggestive signal was also observed for ADAM10. Several important conclusions can be drawn from this work: (i) rare, damaging variants in these genes have a large effect on the AD risk; (ii) unsurprisingly, these variants are enriched in early-onset AD; (iii) all these genes with rare variants have been associated with the AD risk in GWASs of common variants. Furthermore, suggestive evidence of association was also reported with rare, damaging variants in the GWAS-identified genes RIN3, CLU, ZCWPW1, and ACE [46]. This is a remarkable convergence between signals from common variants and signals from rare variants, and so the genes concerned are likely to be of importance in understanding pathophysiological processes in AD; (iv) many of the genes with significant or suggestive signals present very rare loss-of-function mutations associated with a particular high risk of developing AD that is of particular interest at the biological level. For instance, the association observed with loss-of-function mutations in TREM2 suggests that the missense TREM2 R47H variant associated with an increased AD risk negatively impairs the protein's biological function.

We are still learning about APOE

As mentioned above, the association between the APOE gene and the AD risk was first reported in 1993 [8]. It has been estimated that at the age of 85, the lifetime risk of AD is 51% for APOE44 male carriers, 60% for APOE44 female carriers, 23% for APOE34 male carriers, and 30% for APOE34 female carriers. At the same age and without reference to the APOE genotype, this lifetime risk is 11% in males and 14% in females [47].

Many studies have been designed to determine which genetic factors (if any) can protect against or accentuate the risk linked to carriage of the APOE ε4 allele [48, 49]. WES studies have identified many novel potential associations of burden of rare variants with AD risk in individuals carrying APOE e4 alleles [50]. Some researchers reported that polygenic risk scores and genetic risk scores based on common variants associated with AD modify the disease risk and the age at onset of AD in APOE ε4 carriers [22, 51, 52].

In-depth analyses of the APOE locus itself have also revealed complex genetic patterns. However, these analyses are complicated by the strength of the association between APOE and the AD risk and the complex linkage disequilibrium in the APOE region. In the late 1990s, it had already been suggested that the APOE locus association was more complicated than a simple association between the ε2/ε3/ε4 alleles and the risk of developing AD. Several common variants in this locus have been proposed to modify APOE expression and promote an imbalance between APOE3 expression and APOE4 expression [53–55]. Similarly, differential regulation of APOE expression related to different ancestral genomic background around the locus might account for the differences in risk between populations of various ancestries [56–59]. More recently, sequencing studies have characterized two rare variants (V236E and R251G, in complete linkage disequilibrium with the ε3 and ε4 alleles, respectively) associated with a substantially lower risk of developing AD [60–62]. The V236E APOE3 mutation has been shown to reduce APOE aggregation, enhance APOE lipidation in human brains, and reduce amyloid pathology and neuritic dystrophy in an AD-like mouse model [63]. Lastly, two copies of the APOE3 R136S mutation were suggested to have delayed the development of AD by several decades in an individual carrying a PSEN1 mutation [64].

How have the results of the genetic studies influenced our knowledge of the pathophysiological processes in AD?

Integrating biological and GWAS data is essential for a better understanding of the pathophysiological processes involved. To this end, the first (and still the most popular) tool was enrichment pathway analysis. This is based on the postulate whereby a relevant pathway must be enriched in genetic risk factors for the disease in question. The results of the first attempts to apply this approach in AD were published in 2010; two studies of independent GWAS datasets highlighted the involvement of the immune system in AD for the first time [65, 66]. This observation was always confirmed as increasingly large GWASs were conducted. Enrichments in pathways involved in lipid metabolism and endocytosis also gave consistent results from one GWAS to another, while pathways directly involved in APP metabolism and Tau-related proteins became the most strongly associated pathways in the latest GWAS analyses [21, 27]. These latest findings are reassuring and underpin the major roles of the two main hallmarks of AD in the brain notably by establishing that the genetic factors implicated in common forms of AD also point to APP metabolism as a culprit. Importantly, APP and ADAM10 (whose corresponding protein is responsible for α-secretase activity in the brain) are both genetic risk factors for common forms of AD [22, 27]. However, it should be borne in mind that new genetic risk factors are often first evaluated in the context of known pathways. This approach may lead to circular reasoning and thus to an artificial enrichment in specific processes.

Giving biological meaning to data from GWASs and enrichment pathway analyses has some intrinsic limitations. Firstly, a large number of human genes have never been studied in a biological context; in some cases, their biological function is still deduced simply by sequence homology. Even for genes that have been studied, information on pleiotropy and/or function in the brain is not always available. Secondly, and even though the functional variant responsible for the GWAS signal can be assigned to a non-synonymous, deleterious variant in a few cases, the functional variant is usually located in an intergenic region and probably modulates the expression of the disease-causing gene. This causal gene is usually

close to the sentinel/functional variants [67]. However, the presence of several genes in a locus and/or complex linkage disequilibrium patterns that include the sentinel variant make it difficult to determine which gene is responsible for the observed association.

In order to characterize functional variants and genes, several statistical approaches have been developed. The objective is to prioritize disease-associated variants or genes in complex loci by (i) the combination of fine-mapping with methods that leverage enrichments in functional genomic annotations [26, 68]; (ii) co-localization analyses, based on the postulate whereby a GWAS signal that colocalizes with a quantitative trait locus is more likely to be functional [69, 70], and (iii) transcriptome-wide-association studies that identify gene–trait associations by integrating datasets from GWASs and gene expression studies [71]. These approaches were initially based on transcriptomic data but have been now extended to data on splicing, methylation, and protein quantitative trait loci [72–74]. Given the growing number of databases (for an entire organ or by cell type), scores of varying complexity can be used to prioritize genes of interest. However, to calculate these scores, weights have to be assigned subjectively to each level of information. In the latest AD GWAS, this approach prioritized 31 genes in the 42 novel loci associated with the AD risk [27].

Generating biological information to prioritize genes in complex GWAS loci and/or to define the genes pathophysiological roles is also challenging. As reported for other diseases in the post-GWAS era [75], our mechanistic understanding lags far behind the discovery of new AD risk loci in GWASs. Naturally, research efforts tend to shift towards "star" genes that feature damaging, non-synonymous, causal variants because they are easier to study. However, as mentioned above, most of the variants responsible for GWAS signals probably modulate the expression of the genes of interest in subtle ways; this complicates the analysis. Furthermore, the definition of phenotypes of interest in specific cell types in order to functionally characterize a gene of interest is a prerequisite that can lead to the development of hypothesis-driven approaches. Thus, the vast majority of post-GWAS functional studies have been based on aspects of the amyloid cascade hypothesis (the metabolism of APP, and the production and/or the toxicity of amyloid peptides) for a specific gene [76–78] or in systematic screens [79–82] (for a review, see ref. [83]). To a lesser extent, similar approaches have assessed the impact of GWAS-defined genes on Tau toxicity/accumulation [84–87]. Here, it is worth noting the remarkable convergence between BIN1 and Tau: (i) BIN1 modulates Tau toxicity in *Drosophila* and in mouse models [88, 89]; (ii) BIN1 interacts directly with Tau in a phosphorylation-dependent manner [90]; (iii) *BIN1* AD risk variants are associated with increased neurofibrillary tangles and higher Braak stages [88, 91, 92]; and (iv) *BIN1* AD risk variants are associated with levels of Tau/p-Tau (but not of Ab_{1-42}) in cerebrospinal fluid (CSF) [93] and Tau-PET results (but not amyloid-PET results) in the brain [94, 95].

In addition to the specific case of *BIN1*, the most notable success in the genomic/post-GWAS era is undoubtedly the identification of microglia as a cornerstone in the pathophysiology of AD. In 2013, it was found that rare, non-synonymous variants in *TREM2* (a gene almost only expressed in microglia) were associated with a significant elevation of the AD risk [41, 42]. On the genetic level, the importance of the microglia has since been reinforced by the discovery of rare, non-synonymous variants in *PLCG2* and *ABI3* – both of which are particularly expressed in microglia [39]. According to the GWAS results, AD risk alleles are specifically enriched in active enhancers of monocytes, macrophages and especially microglia [96, 97]. Consequently, many AD risk loci may be mediated through gene expression or splicing in microglia [98]. Many researchers have sought to link microglia, genetic risk factors (mainly *TREM2*) and Aβ peptides together through their impact on amyloid plaque formation/compaction [99–101], toxicity, and synapse pruning [102]. In animal models, GWAS-defined genes appear to determine the microglial response to Aβ

but not to Tau pathology [103, 104]. However, Aβ-activated microglia might control the seeding/spreading and accumulation of Tau pathology [105–107]; hence, the AD genetic risk would be downstream of the amyloid pathway but upstream of the Tau pathology pathway [104]. Accordingly, it has been suggested that the P522R *PLCG2* variant reduces AD progression in patients with mild cognitive impairment by mitigating Tau pathology in the presence of amyloid pathology [106].

In conclusion, the biological data produced by genomic studies have not only reinforced the roles of APP metabolism and Tau pathology in the etiology of AD but have also opened up a new field of investigation concerning the immune system in general and the microglia in particular. AD is no longer seen as a linear process defined by the amyloid cascade; in fact, it appears to be an increasingly complex phenomenon resulting from pathophysiological processes with many entry points that can trigger the disease or interact to speed up or slow-down disease progression. In view of this new paradigm for AD and the complexity suggested by the new genetic data, several new, non-exclusive hypotheses have been put forward. For example, the "cellular phase" hypothesis postulates feedback and anticipation reactions between all the various cell types in the brain [108]. The "genetically driven synaptic failure" model is based on changes in the focal adhesion pathway and the related cell signaling [83].

Although much progress has been made, how many genetic factors have yet to be identified?

Between 2009 and 2022, the number of risk loci for AD (apart from *APOE*) rose from 3 to more than 75. These discoveries have impacted our knowledge of pathophysiological processes, which is starting to be used in translational research on potential diagnostic/prognostic tools. One can legitimately wonder (i) how many genetic risk factors remain to be identified and (ii) how valuable will be a polygenic risk score (PRS). To that end, some researchers have tried to estimate the "missing heritability" in AD.

In fact, the estimates of AD's heritability vary greatly from one study to another. The highest values have been provided by twin studies (from 48% to 79% on the liability scale, denoted h^2_{twins}). Intermediate values have been obtained in analyses of individual-level GWAS data from unrelated individuals (from 24% to 55%, denoted h^2_{SNP}). Lastly, the lowest estimates were based on summary statistics from GWASs (from 2% to 42%, denoted $h^2_{summary}$)(Supplementary table 2). These differences are not unexpected, since heritability is a population-dependent measure. Differences between twin studies can also be explained by the rather small sample sizes considered, which lead to large confidence intervals for the estimates. However, all the h^2_{twins} estimates are quite broad, and a meta-analysis of twin studies estimated the heritability of dementia in AD to be 62.7% [109]. As has been observed for many complex diseases or traits, intermediate heritability values have been estimated in analyses of individual-level GWAS data from unrelated individuals. There are several explanations for this observation, including poor tagging of causal variants in GWAS data (for example rare causal variants), or the over-estimation of heritability in twin studies due to a common environment, gene-gene interactions, or gene-environment interactions [110–112]. More generally, the genomic heritabilities h^2_{SNP} and $h^2_{summary}$ are specific to the variant set considered. Filtering on the minor allele frequency is usual, sex chromosomes are excluded, and $h^2_{summary}$ is computed after removing the major histocompatibility complex region (which contains a genetic risk factor for AD) and loci with large effects (such as *APOE*); this leads to underestimation of the heritability. Furthermore, the restricted maximum likelihood method (used to compute h^2_{SNP} for AD) reportedly underestimates the heritability of binary traits in case-control studies, and, like the linkage disequilibrium score regression (LDSC) approach (used to compute $h^2_{summary}$) [113], can provide biased estimates in the

presence of strong non-genetic risk factors (such as age, for AD) [114, 115]. Variability in h^2_{SNP} and $h^2_{summary}$ is due in part to the difference in the prevalence values considered (from 2% to 33%). The prevalence varies greatly with age but also with sex and has a large impact on the estimation of heritability. Hence, an increase in the prevalence from 2% to 33% means that h^2_{SNP} rises from 24% to 55% [116, 117]. The $h^2_{summary}$ values are lower than the h^2_{SNP} values and range from 2% (with a prevalence of 5%) to 25% (with a prevalence of 17%), although one of the studies considered had a small sample size and estimated an outlying value of $h^2_{summary}$ of 42%. However, $h^2_{summary}$ is commonly computed with the LDSC approach, and considering linkage disequilibrium in an external reference panel rather than in the study sample, which biases the estimates [118, 119]. Furthermore, LDSC reportedly underestimates heritability when a major mutation (such as *APOE*) explains a high proportion of the heritability [120]. $h^2_{summary}$ is most often computed from GWAS meta-analysis results and tends to decrease with the size of the meta-analysis sample. This phenomenon has been observed for other diseases and traits and might be due to inter-study heterogeneity with regard to LD and characteristics of the study populations (such as age and *APOE* status), the accuracy of diagnoses, and the use of unscreened controls in the largest AD meta-analyses [119, 121–123]. In particular, the largest AD meta-analyses included some proxy-AD cases and thus assessed the genetics of AD and related dementias rather than AD alone. Furthermore, some studies do not take account of proxy-cases correctly when computing heritability, which leads to underestimated values [117]. Lastly, varying levels of ascertainment in meta-analyzed studies can also result in underestimation if it is not appropriately accounted for [124].

Overall, estimating the heritability of AD remains methodologically challenging. From a conceptual point of view, many of the inherent assumptions in the statistical models used to estimate heritability are not appropriate for a disease as complex as AD. This raises the question of whether estimates of the "missing heritability" are meaningful [111, 125].

How can we improve our knowledge of the genetics of AD?
The difficulty of understanding the genetics of AD is clearly illustrated by the fact that some researchers suggest that AD is an oligogenic disease involving around 100 common causal variants [120], while others favor a polygenic model with up to 11,000 common causative variants [126]. Consequently, it is reasonable to consider prudently that new loci have yet to be identified and that known loci require further characterization in AD.

Of course, it will always be necessary to increase the sample size (whether of European ancestry or multiple ancestries) in classical GWASs and sequencing analyses, in order to capture the genetic information carried respectively by common and rare variants. It will also be necessary to analyze all the types of variations in the genome. Structural variants (SVs i.e. changes larger than 50 bp) have been poorly studied in AD. SVs and (especially) copy number variations are major sources of genomic variation; although two individuals may differ genetically by 0.1% when considering single nucleotide variations, the difference increases to 1.5% when SVs are also taken into account [127–129]. However, apart from gene duplications leading to monogenic forms of neurodegenerative diseases [130–132], this field that has been poorly studied in common AD and the few available studies of GWAS addressing SV association with AD data lack consistency [133–137]; this might be attributable to various technical biases (e.g. different genotyping and sequencing platforms), batch effects, and a lack of statistical power. Nevertheless, methodological and technical progress (especially long-read sequencing) will probably make SVs easier to detect [138, 139].

Along with the ability to fully capture the variability in the human genome, it is also important to exploit this knowledge through a wide range of approaches to characterize the genetic component of AD. Fine-mapping approaches are needed to better understand the real importance of a given gene/locus: the risk conferred by a gene/locus can be underestimated if one is unaware of the existence of several independent causal variants, as observed for *BIN1* [26]. Heterogeneous or even hidden genetic signals (e.g. those that depend on the *APOE* genotype, sex, or early versus late onset) can also be assessed in interaction and stratification analyses. At another level of complexity, pangenomic searches for gene-gene or gene-environment interactions can be developed [140–142]. These approaches nevertheless require significant computing power and can generate false positives through multiple statistical testing. For gene-environment interactions more specifically, the question of statistical power arises because the longitudinal studies performed to date (e.g. the CHARGE consortium) included some tens of thousands of individuals. However, this limitation is being lifted by the creation of large biobanks (such as the UKB) in which the number of cases diagnosed will increase as the study population ages. Lastly, the implementation of increasingly large GWASs of many AD-related endophenotypes [93, 143–146] and the GWAS datasets' integration with other "omics" databases (e.g. systems biology) should make it possible to characterize key elements of AD genetics and related pathways. Importantly, the move towards systematic, detailed integration of the available data will make it difficult to validate results obtained independently. It will therefore become essential to demonstrate the biological relevance of genetic results (some of which will be generated by artificial intelligence) in appropriate cellular and/or animal models.

Lastly, many methodological issues related to heterogeneity in the generation of summary statistics can impact the results of GWASs and complicate comparisons and subsequent analyses. For example, analyses of proxy-AD cases require a correction factor, which can differ across studies, different covariates are used to adjust statistical models, and different imputation panels are used.

Since the IGAP's results were published, AD GWASs have been carried out by the meta-analysis of shared and (potential heterogenous) summary statistics and many findings might be driven by the IGAP summary statistics. Further, the number of controls has increased more quickly than the number of patients, following access to the very large population-based biobanks; this has led to stagnation in the number of novel loci characterized (Fig. 2).

The EADB was set up against this background. The biobank's main objective was to double the number of new, clinically diagnosed cases of AD available for analysis; this probably explains why the number of detected loci increased markedly (Fig. 2). Furthermore, the EADB shared raw data from the existing European GWASs, which enables the data to be checked for potential sample overlaps. It also facilitates the use of the same imputation panel (TOPMed), the application of the same quality control procedures, and the implementation of similar statistical models for homogeneous summary statistics prior to meta-analysis. Overall, sharing raw data improves and speeds up GWAS meta-analysis. Unfortunately, although the USA has a remarkable data sharing policy to be commended and encouraged, the European Union's General Data Protection Regulation (a legitimate effort to protect individual data) restricts data sharing outside Europe (and even into Europe).

Will genetics impact the definition of AD and related dementias?
In addition to the inherent methodological issues of GWASs, the diagnosis of AD can also be questioned and debated. The definition of around 40,000 to 50,000 proxy-AD cases in the UKB can be criticized because it is based on self-reports of a family history of AD dementia. Although the large number of proxy-AD cases increases the statistical power of the AD GWAS, and the genetic correlation between AD and proxy-AD is high (above 0.65) [20, 23, 25, 117], this "virtual diagnosis" might lead to the inclusion
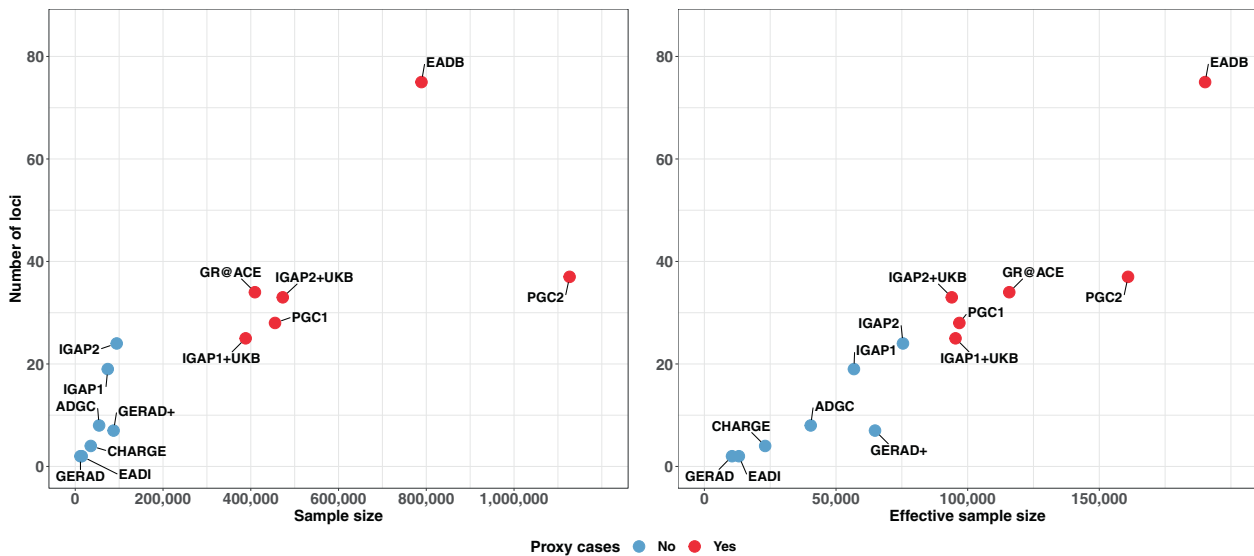
Fig. 2 Number of loci identified in the different GWASs. Number of loci identified as a function of the sample size (left) or the effective sample size (right) (cases and controls) in the main GWASs published since 2009 (EADI, GERAD, CHARGE, GERAD+, ADGC, IGAP1, IGAP1 + UKB (2017), IGAP1 + UKB (2018), PGC1, IGAP2, IGAP2 + UKB, GR@ACE, PGC2, and EADB [14–27]. The colors indicate the presence or absence of proxy cases in the GWAS. The effective sample size was computed per study included in the meta-analyses, and then summed across studies [124]. The effective sample size for the proxy UKB study was computed by dividing the raw number of proxy cases and proxy controls by four [24, 117].

of misdiagnosed individuals. In addition, even when studying clinically diagnosed cases, it is estimated that around 10–20% of diagnoses are incorrect. These points have prompted some experts to even consider that the detection of the same genetic signals for several neurodegenerative diseases could be artefactual; because of these misdiagnoses and the size of the AD GWASs, these GWASs were also statistically powerful enough to detect signals corresponding to the pathologies associated with these misdiagnoses.

To circumvent this problem, the selection of individuals on the basis of post-mortem pathology assessments is obviously the best option for establishing a definitive diagnosis. However, this approach is still strongly limited by the lack of access to large collections of brain samples – especially for age-matched controls [91, 147, 148]. Another option would be to use biomarker profiles to distinguish between AD cases, controls, and individuals with other neurodegenerative diseases. However, generalization of the recently developed amyloid/Tau/neurodegeneration (A/T/N) classification is limited by the cost of CSF biomarker assays and imaging [149]. Furthermore, the biomarker field is evolving rapidly, with the discovery of novel targets (e.g. p-Tau 217 and p-Tau 231) with potentially high diagnostic/prognostic value [150, 151]. These biomarkers might lead to a refinement of the diagnostic profile for AD and related disorders and therefore the corresponding genetic studies. However, these approaches might also lead to the over-selection of cases that are not representative of AD and the latter's complex relationship with the neurodegenerative landscape in real life.

Indeed, the postulate whereby detection of the same genetic signal in two different neurodegenerative diseases results from misdiagnosis is probably too simplistic and does not fully take account of important data from genetics and, above all, brain pathology markers. Firstly, in GWASs of clinical frontotemporal dementia, TMEM106B variants achieved only modest p-values and odds ratios for the behavioral subtype of frontotemporal dementia [152], and the associations were dependent on GRN mutations [153]. It is unlikely that this clinical subpopulation alone could drive a genome-wide, significant signal in an AD GWAS. Secondly, and even though a locus can be common to several diseases, it

appears that signals (and thus functional variants) can differ – as observed for the IDUA locus in AD and in Parkinson's disease. This may indicate that a common locus does not necessarily have the same pathological consequences in different neurodegenerative diseases. Thirdly, patients with AD often have other (concomitant) neurological diseases, which are potential "partners in crime" [154]. For instance, GRN and TMEM106B are reportedly involved in defective endosome/lysosome trafficking/function – a defect that is also observed in AD [155, 156]. GRN protects against amyloid-β deposition and toxicity in AD mouse models [157]. This is also illustrated by the association between BIN1 and the risk of developing Lewy body dementia, as observed in a large sample of autopsy-confirmed and clinically probable cases [92]. This BIN1 signal (like that observed in AD) was significantly associated with increased neurofibrillary tangles, as also observed in AD brains [90, 91]. Given the many observations linking BIN1 to Tau-related endophenotypes [93, 94] and pathological processes [89, 90], the identical genetic signal for BIN1 in two diseases pathologies suggest that similar Tau-related pathological processes are operating.

In general, one can argue that the initial and/or subsequent localization and development of concomitant diseases can trigger or favor AD. In other words, the interplay between pathophysiological processes in different pathologies might prompt the detection of a common genetic locus. The presence of common causal variants in the same gene might indicate a shared pathological role, whereas the presence of different causal variants in the said gene might indicate specific mechanisms or cell-type-specific expression. Common genetic risk factors in neurodegenerative diseases might have thus several important implications and do not necessarily reflect misdiagnosis in GWASs. In combination with other biomarkers, genetic markers could perhaps redefine the boundaries and relationships between neurodegenerative diseases (i.e. a hypothetical continuum). Hence, the inclusion of this information may be particularly useful for better defining AD and concomitant diseases in real life. For example, the development of PRSs might help to integrate this heterogeneity into pathophysiological processes for a given case – even when this case has been diagnosed as AD. This may be of

particular importance when developing drugs potentially targeting shared genes/pathways and the development of an efficient precision medicine.

Development of a tool for the diagnosis and prognosis of the common forms of AD

Translating new genetic knowledge into clinical practice (e.g. to identify individuals at risk of progression to dementia) is challenging, for many reasons. Researchers have designed PRSs that reflect an individual's overall genetic burden for a given disease [158–161]. The PRS provides a cumulative effect score summarizing the small information distributed across the individual susceptibility variants. To construct a PRS, an effect estimate has to be assigned to each variant included in the PRS; this estimate is obtained from the summary statistics derived from large GWASs and meta-analyses thereof [162]. This large GWAS dataset is considered to be the "training" dataset from which the effect estimators are obtained and used to build an informative PRS using an independent dataset called validation or test dataset. In this regard, the constant increase in sample size in GWASs has contributed significantly to the identification of genuine risk variants for diseases and has improved the effect estimates for each risk variant. All this information helps to better calibrate PRSs [163] for the assessment of at-risk individuals and the scores' translation into strategies for better diagnosis and (eventually) early intervention and prevention. PRS methodologies have been improved by the availability of genomic information on linkage disequilibrium dependencies between variants; this enables better definition of a single tagging variant for a defined locus and avoids redundant information and falsely inflated results [160, 164].

The potential value of PRSs in the context of AD has been emphasized by studies in which an AD-derived PRS was associated with the clinical diagnosis [165], cortical thickness [166], memory, hippocampal volume, cognitive decline [167], disease progression [168], and post-mortem confirmed cases [169]. However not everything has been "hunky-dory" with PRS because results from studies using an AD-derived PRS, incorporating all or part of the GWAS signals, have been inconclusive with regard to disease progression and clinical diagnosis [170, 171]. Thus, these contradictory results cast some doubts on the strategy used to compute the PRS. Consequently, several studies have focused on determining the best overall way to compute a PRS (for a more detailed review see [172]).

The term "PRS" is used liberally to refer to scores that include various number of SNPs. Current methods for PRS computation are designed to reduce the signal-to-noise ratio by selecting a small number of the most informative SNPs [173]. This specific selection of SNPs in the final model follows different selection criteria such as significance thresholds or a priori selection by researchers. This selection generally includes only common SNPs (minor allele frequency 5%) using a model that assumes only additive effects for these variants, and do not account for additional components like SNPxSNP interactions and dominance [174, 175]. Besides, rare variants (minor allele frequency <1%) or copy number variations are normally left out of a PRS as they are not included in the genetic data derived from GWAS. Thus, these inclusion criteria may restrict the genomic information included in the model leading to the inability of PRS to completely capture the genomic landscape of the selected trait [176, 177]. Other important aspects of PRS calculation are subject to debate: the best way to model *APOE* within a PRS, the p-value threshold for including SNPs, and the comparison of PRSs for independent cohorts [173, 178, 179]. As mentioned above, and although it is clear that several genes are involved in the AD process, there is still debate as to whether AD is polygenic or oligogenic [120, 165]. Resolving these issues will have a major impact on the strategy used to build a PRS in AD. Interestingly, a recent GWAS of human height (a strongly polygenic and heterogeneous trait) in over 5 million participants appeared to reach a plateau for the identification of genetic variants [180]. The investigators concluded that substantially increasing the sample size of GWAS might largely resolve the heritability attributed to common variation by identifying a finite set of SNPs [180].

While most of the debate concerning PRS has focused on selecting SNPs from a single summary statistic dataset, less attention has been paid to the fact that AD is a complex, heterogeneous phenotype [181]. As previously indicated, the pathophysiological processes in AD are diverse, and not all will be present and operating at the same time in a given patient. Hence, two patients with AD will present different subsets of the pathophysiological processes. Moreover, the various pathogenic processes and endophenotypes associated with AD might be driven by different genetic variants – ones that are not necessarily involved in the susceptibility to AD identified by case-control GWAS. In support of this hypothesis, GWASs using CSF levels of the hallmark biomarkers of AD (i.e. amyloid beta and tau) have identified a handful of genetic variants involved also in the susceptibility to AD [93, 182]. These studies also suggest that the genetics driving each of the biomarker's levels in CSF are independent of each other – except for the *APOE* locus. On the same lines, GWASs of cognitive performance in more than 300,000 cognitively healthy individuals also reported a handful of genome-wide significant signals in regions overlapping with loci responsible for susceptibility to AD [183]. This finding suggests that the genetic factors identified in case-control GWASs will only partly explain the variance observed in the AD phenotype. Hence, a simple PRS derived from case-control summary statistics might not explain the total genetic risk in an AD patient. As discussed above, some of the genetic drivers of AD are probably also involved in other types of dementia. Accordingly, a PRSs encompassing the genome-wide significant SNPs identified by the EADB consortium was consistently associated with progression to dementia in various cohorts of patients with mild cognitive impairment (prodromal dementia) and cognitively healthy individuals [27]. These findings underscore the contribution of the identified genetic variants to the progression of cognitive decline along the continuum of AD and may be also to related diseases, though this latter suggestion needs further research.

In conclusion, while PRS has been seen as a path to personalized medicine, the use of PRS in medicine is not exempt from limitations. PRS informs about the probability and not a destiny of a person based on the genetic burden of a particular disease. Hence, whether or not a PRS translates into pathology will not be reported by the PRS itself but from dynamic markers showing the ongoing pathology. This, however, also states the question of whether using PRS for age-related diseases in young populations may have any impact on prevention. However, our PRSs will become increasingly accurate as our knowledge of the genetics of susceptibility to AD improves. A PRS that provides information on the genetics of susceptibility to AD and on the related underlying endophenotypes would be especially useful. Likewise, strategies based on hazard ratios (rather than risk scores) might also improve our predictive ability and could ultimately lead to the "holy grail", i.e., a score for identifying cognitively healthy young individuals at risk of developing dementia. Identification of the pathways affected in each individual at risk would open up many opportunities, such as the development of personalized prevention strategies.

CONCLUSION

The AD genetic community has a clear responsibility for accelerating and fostering the generation of biological and clinically relevant results. The characterization of the genetics of multifactorial diseases like AD has a major impact on the research

in the fields and therefore the development of treatments. By way of an example, human genetic evidence supported two-thirds of the drugs approved by the US Food and Drug Administration in 2021 [184]. In view of this major responsibility, geneticists must ask themselves essential questions about their research practices and the resulting publications. This implies an appropriate understanding of the various biases in the results, how these biases should be taken into account, and how the results should be published. Geneticists must also be prepared to correct their mistakes, given the inherent risk of generating false positives in genetic studies.

Over the last ten years, we have made remarkable progress in understanding the genetics of AD but considerable additional efforts are still needed. Nevertheless, one can assume that the most obvious genetic factors/loci have already been identified; the remaining research efforts and investments will have to define new factors/loci with restricted impact – even though rare variants can still have a moderate impact on the AD risk. Regardless of how much of the genetic variability in AD remains to be discovered, knowledge of AD genetics has already had a considerable impact on (i) our understanding of the pathophysiology of AD (e.g. the central role of microglia), (ii) the definition of sub-populations at risk, and (iii) the interplay between AD and other neurodegenerative diseases. These genetics-related advances are especially important because few genetic risk factors have been studied in detail after their initial identification as being associated with the AD risk. We can therefore legitimately be optimistic and assume that the growth in post-GWAS research (such as the MODEL-AD project [185]) will enable to make rapid progress in the years to come.

REFERENCES

1. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature. 1991;349:704–6.
2. Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature. 1995;375:754–60.
3. Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. Science. 1995;269:973–7.
4. Hardy J, Higgins G. Alzheimer's disease: the amyloid cascade hypothesis. Science. 1992;256:184–5.
5. Herrup K. The case for rejecting the amyloid cascade hypothesis. Nat Neurosci. 2015;18:794–9.
6. Selkoe DJ, Hardy J. The amyloid hypothesis of Alzheimer's disease at 25 years. EMBO Mol Med. 2016;8:595–608.
7. Frisoni GB, Altomare D, Thal DR, Ribaldi F, van der Kant R, Ossenkoppele R, et al. The probabilistic model of Alzheimer disease: the amyloid hypothesis revised. Nat Rev Neurosci. 2022;23:53–66.
8. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, 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. 1993;90:1977–81.
9. Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmechel DE, Gaskell PC, et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. Nat Genet. 1994;7:180–4.
10. Lambert J-C, Amouyel P. Genetic heterogeneity of Alzheimer's disease: complexity and advances. Psychoneuroendocrinology. 2007;32:S62–70.
11. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet. 2007;39:17–23.
12. Farrer LA, Sherbatich T, Keryanov SA, Korovaitseva GI, Rogaeva EA, Petruk S, et al. Association between angiotensin-converting enzyme and Alzheimer disease. Arch Neurol. 2000;57:210–4.
13. Rogaeva E, Meng Y, Lee JH, Gu Y, Kawarai T, Zou F, et al. The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. Nat Genet. 2007;39:168–77.
14. Lambert J-C, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat Genet. 2009;41:1094–9.

15. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. 2009;41:1088–93.
16. Seshadri S. Genome-wide analysis of genetic loci associated with Alzheimer Disease. JAMA. 2010;303:1832.
17. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45:1452–8.
18. Hollingworth P, Harold D, Sims R, Gerrish A, Lambert J-C, Carrasquillo MM, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat Genet. 2011;43:429–35.
19. Naj AC, Jun G, Beecham GW, Wang L-S, Vardarajan BN, Buros J, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet. 2011;43:436–41.
20. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet. 2019;51:404–13.
21. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. Nat Genet. 2019;51:414–30.
22. de Rojas I, Moreno-Grau S, Tesi N, Grenier-Boley B, Andrade V, Jansen IE, et al. Common variants in Alzheimer's disease and risk stratification by polygenic risk scores. Nat Commun. 2021;12:3417.
23. Wightman DP, Jansen IE, Savage JE, Shadrin AA, Bahrami S, Holland D, et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. Nat Genet. 2021;53:1276–82.
24. Liu JZ, Erlich Y, Pickrell JK. Case-control association mapping by proxy using family history of disease. Nat Genet. 2017;49:325–31.
25. Marioni RE, Harris SE, Zhang Q, McRae AF, Hagenaars SP, Hill WD, et al. GWAS on family history of Alzheimer's disease. Transl Psychiatry. 2018;8:99.
26. Schwartzentruber J, Cooper S, Liu JZ, Barrio-Hernandez I, Bello E, Kumasaka N, et al. Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer's disease risk genes. Nat Genet. 2021;53:392–402.
27. Bellenguez C, Küçükali F, Jansen IE, Kleineidam L, Moreno-Grau S, Amin N, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat Genet. 2022;54:412–36.
28. Logue MW, Schu M, Vardarajan BN, Buros J, Green RC, Go RCP, et al. A comprehensive genetic association study of Alzheimer disease in African Americans. Arch Neurol. 2011;68:1569–79.
29. Reitz C, Jun G, Naj A, Rajbhandary R, Vardarajan BN, Wang L-S, et al. Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E ε4,and the risk of late-onset Alzheimer disease in African Americans. JAMA. 2013;309:1483–92.
30. Kunkle BW, Schmidt M, Klein H-U, Naj AC, Hamilton-Nelson KL, Larson EB, et al. Novel Alzheimer Disease risk loci and pathways in African American individuals using the African genome resources panel: a meta-analysis. JAMA Neurol. 2021;78:102–13.
31. Horimoto ARVR, Xue D, Thornton TA, Blue EE. Admixture mapping reveals the association between Native American ancestry at 3q13.11 and reduced risk of Alzheimer's disease in Caribbean Hispanics. Alzheimers Res Ther. 2021;13:122.
32. Rajabli F, Feliciano-Astacio BE, Cukier HN, Wang L, Griswold AJ, Hamilton-Nelson KL, et al. Linkage of Alzheimer disease families with Puerto Rican ancestry identifies a chromosome 9 locus. Neurobiol Aging. 2021;104:115.e1–115.e7.
33. Dalmasso MC, Brusco LI, Olivar N, Muchnik C, Hanses C, Milz E, et al. Transethnic meta-analysis of rare coding variants in PLCG2, ABI3, and TREM2 supports their general contribution to Alzheimer's disease. Transl Psychiatry. 2019;9:55.
34. Kretzschmar GC, Alencar NM, da Silva SSL, Sulzbach CD, Meissner CG, Petzl-Erler ML, et al. GWAS-top polymorphisms associated with late-onset Alzheimer Disease in Brazil: pointing out possible new culprits among non-coding RNAs. Front Mol Biosci. 2021;8:632314.
35. Marca-Ysabel MV, Rajabli F, Cornejo-Olivas M, Whitehead PG, Hofmann NK, Illanes Manrique MZ, et al. Dissecting the role of Amerindian genetic ancestry and the ApoE ε4 allele on Alzheimer disease in an admixed Peruvian population. Neurobiol Aging. 2021;101:298.e11–298.e15.
36. Kang S, Gim J, Lee J, Gunasekaran TI, Choi KY, Lee JJ, et al. Potential novel genes for late-onset Alzheimer's Disease in East-Asian descent identified by APOE-stratified genome-wide association study. J Alzheimers Dis JAD. 2021;82:1451–60.
37. Jia L, Li F, Wei C, Zhu M, Qu Q, Qin W, et al. Prediction of Alzheimer's disease using multi-variants from a Chinese genome-wide association study. Brain J Neurol. 2021;144:924–37.
38. Shigemizu D, Mitsumori R, Akiyama S, Miyashita A, Morizono T, Higaki S, et al. Ethnic and trans-ethnic genome-wide association studies identify new loci influencing Japanese Alzheimer's disease risk. Transl Psychiatry. 2021;11:151.
39. Sims R, van der Lee SJ, Naj AC, Bellenguez C, Badarinarayan N, Jakobsdottir J, et al. Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. Nat Genet. 2017;49:1373–84.

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

41. Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. N Engl J Med. 2013;368:107–16.

42. Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. TREM2 variants in Alzheimer's disease. N Engl J Med. 2013;368:117–27.

43. Steinberg S, Stefansson H, Jonsson T, Johannsdottir H, Ingason A, Helgason H, et al. Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. Nat Genet. 2015;47:445–7.

44. Bellenguez C, Charbonnier C, Grenier-Boley B, Quenez O, Le Guennec K, Nicolas G, et al. Contribution to Alzheimer's disease risk of rare variants in TREM2, SORL1, and ABCA7 in 1779 cases and 1273 controls. Neurobiol Aging. 2017;59:220.e1–220.e9.

45. Bis JC, Jian X, Kunkle BW, Chen Y, Hamilton-Nelson KL, Bush WS, et al. Whole exome sequencing study identifies novel rare and common Alzheimer's-Associated variants involved in immune response and transcriptional regulation. Mol Psychiatry. 2020;25:1859–75.

46. Holstege H, Hulsman M, Charbonnier C, Grenier-Boley B, Quenez O, Grozeva D, et al. Exome sequencing identifies novel AD-associated genes. MedRxiv. 2020. https://doi.org/10.1101/2020.07.22.20159251.

47. Genin E, Hannequin D, Wallon D, Sleegers K, Hiltunen M, Combarros O, et al. APOE and Alzheimer disease: a major gene with semi-dominant inheritance. Mol Psychiatry. 2011;16:903–7.

48. Huq AJ, Fransquet P, Laws SM, Ryan J, Sebra R, Masters CL, et al. Genetic resilience to Alzheimer's disease in APOE ε4 homozygotes: a systematic review. Alzheimers Dement J Alzheimers Assoc. 2019;15:1612–23.

49. Belloy ME, Napolioni V, Han SS, Le Guen Y, Greicius MD. Alzheimer's Disease Neuroimaging Initiative. Association of Klotho-VS heterozygosity with risk of Alzheimer disease in individuals who carry APOE4. JAMA Neurol. 2020;77:849–62.

50. Ma Y, Jun GR, Zhang X, Chung J, Naj AC, Chen Y, et al. Analysis of whole-exome sequencing data for Alzheimer disease stratified by APOE genotype. JAMA Neurol. 2019;76:1099–108.

51. van der Lee SJ, Wolters FJ, Ikram MK, Hofman A, Ikram MA, Amin N, et al. The effect of APOE and other common genetic variants on the onset of Alzheimer's disease and dementia: a community-based cohort study. Lancet Neurol. 2018;17:434–44.

52. Juul Rasmussen I, Rasmussen KL, Nordestgaard BG, Tybjærg-Hansen A, Frikke-Schmidt R. Impact of cardiovascular risk factors and genetics on 10-year absolute risk of dementia: risk charts for targeted prevention. Eur Heart J. 2020;41:4024–33.

53. Bullido MJ, Artiga MJ, Recuero M, Sastre I, García MA, Aldudo J, et al. A polymorphism in the regulatory region of APOE associated with risk for Alzheimer's dementia. Nat Genet. 1998;18:69–71.

54. Lambert JC, Berr C, Pasquier F, Delacourte A, Frigard B, Cottel D, et al. Pronounced impact of Th1/E47cs mutation compared with -491 AT mutation on neural APOE gene expression and risk of developing Alzheimer's disease. Hum Mol Genet. 1998;7:1511–6.

55. Lambert J-C, Araria-Goumidi L, Myllykangas L, Ellis C, Wang JC, Bullido MJ, et al. Contribution of APOE promoter polymorphisms to Alzheimer's disease risk. Neurology. 2002;59:59–66.

56. Rajabli F, Feliciano BE, Celis K, Hamilton-Nelson KL, Whitehead PL, Adams LD, et al. Ancestral origin of ApoE ε4 Alzheimer disease risk in Puerto Rican and African American populations. PLoS Genet. 2018;14:e1007791.

57. Bussies PL, Rajabli F, Griswold A, Dorfsman DA, Whitehead P, Adams LD, et al. Use of local genetic ancestry to assess TOMM40-523' and risk for Alzheimer disease. Neurol Genet. 2020;6:e404.

58. Nuytemans K, Lipkin Vasquez M, Wang L, Van Booven D, Griswold AJ, Rajabli F, et al. Identifying differential regulatory control of APOE ε4 on African versus European haplotypes as potential therapeutic targets. Alzheimers Dement J Alzheimers Assoc. 2022;18:1930–42.

59. Rajabli F, Beecham GW, Hendrie HC, Baiyewu O, Ogunniyi A, Gao S, et al. A locus at 19q13.31 significantly reduces the ApoE ε4 risk for Alzheimer's Disease in African Ancestry. PLoS Genet. 2022;18:e1009977.

60. Medway CW, Abdul-Hay S, Mims T, Ma L, Bisceglio G, Zou F, et al. ApoE variant p.V236E is associated with markedly reduced risk of Alzheimer's disease. Mol Neurodegener. 2014;9:11.

61. Le Guen Y, Belloy ME, Grenier-Boley B, de Rojas I, Castillo-Morales A, Jansen I, et al. Association of rare APOE missense variants V236E and R251G with risk of Alzheimer disease. JAMA Neurol. 2022;79:652–63.

62. Rasmussen KL, Tybjaerg-Hansen A, Nordestgaard BG, Frikke-Schmidt R. APOE and dementia - resequencing and genotyping in 105,597 individuals. Alzheimers Dement J Alzheimers Assoc. 2020;16:1624–37.

63. Liu C-C, Murray ME, Li X, Zhao N, Wang N, Heckman MG, et al. APOE3-Jacksonville (V236E) variant reduces self-aggregation and risk of dementia. Sci Transl Med. 2021;13:eabc9375.

64. Arboleda-Velasquez JF, Lopera F, O'Hare M, Delgado-Tirado S, Marino C, Chmielewska N, et al. Resistance to autosomal dominant Alzheimer's disease in an APOE3 Christchurch homozygote: a case report. Nat Med. 2019;25:1680–3.

65. Jones L, Holmans PA, Hamshere ML, Harold D, Moskvina V, Ivanov D, et al. Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. PLoS One. 2010;5:e13950.

66. Lambert J-C, Grenier-Boley B, Chouraki V, Heath S, Zelenika D, Fievet N, et al. Implication of the immune system in Alzheimer's disease: evidence from genome-wide pathway analysis. J Alzheimers Dis JAD. 2010;20:1107–18.

67. Mountjoy E, Schmidt EM, Carmona M, Schwartzentruber J, Peat G, Miranda A, et al. An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. Nat Genet. 2021;53:1527–33.

68. Kichaev G, Yang W-Y, Lindstrom S, Hormozdiari F, Eskin E, Price AL, et al. Integrating functional data to prioritize causal variants in statistical fine-mapping studies. PLoS Genet. 2014;10:e1004722.

69. Wallace C, Rotival M, Cooper JD, Rice CM, Yang JHM, McNeill M, et al. Statistical colocalization of monocyte gene expression and genetic risk variants for type 1 diabetes. Hum Mol Genet. 2012;21:2815–24.

70. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014;10:e1004383.

71. Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, et al. Opportunities and challenges for transcriptome-wide association studies. Nat Genet. 2019;51:592–9.

72. De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. Nat Neurosci. 2014;17:1156–63.

73. Lunnon K, Smith R, Hannon E, De Jager PL, Srivastava G, Volta M, et al. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. Nat Neurosci. 2014;17:1164–70.

74. Wingo AP, Liu Y, Gerasimov ES, Gockley J, Logsdon BA, Duong DM, et al. Integrating human brain proteomes with genome-wide association data implicates new proteins in Alzheimer's disease pathogenesis. Nat Genet. 2021;53:143–6.

75. Gallagher MD, Chen-Plotkin AS. The post-GWAS era: from association to function. Am J Hum Genet. 2018;102:717–30.

76. Tian Y, Chang JC, Fan EY, Flajolet M, Greengard P. Adaptor complex AP2/PICALM, through interaction with LC3, targets Alzheimer's APP-CTF for terminal degradation via autophagy. Proc Natl Acad Sci USA. 2013;110:17071–6.

77. Ubelmann F, Burrinha T, Salavessa L, Gomes R, Ferreira C, Moreno N, et al. Bin1 and CD2AP polarise the endocytic generation of beta-amyloid. EMBO Rep. 2017;18:102–22.

78. Sakae N, Liu C-C, Shinohara M, Frisch-Daiello J, Ma L, Yamazaki Y, et al. ABCA7 deficiency accelerates amyloid-β generation and Alzheimer's neuronal pathology. J Neurosci. 2016;36:3848–59.

79. Mukherjee S, Russell JC, Carr DT, Burgess JD, Allen M, Serie DJ, et al. Systems biology approach to late-onset Alzheimer's disease genome-wide association study identifies novel candidate genes validated using brain expression data and Caenorhabditis elegans experiments. Alzheimers Dement J Alzheimers Assoc. 2017;13:1133–42.

80. Chapuis J, Flaig A, Grenier-Boley B, Eysert F, Pottiez V, Deloison G, et al. Genome-wide, high-content siRNA screening identifies the Alzheimer's genetic risk factor FERMT2 as a major modulator of APP metabolism. Acta Neuropathol. 2017;133:955–66.

81. Camargo LM, Zhang XD, Loerch P, Caceres RM, Marine SD, Uva P, et al. Pathway-based analysis of genome-wide siRNA screens reveals the regulatory landscape of APP processing. PLoS One. 2015;10:e0115369.

82. Bali J, Gheinani AH, Zurbriggen S, Rajendran L. Role of genes linked to sporadic Alzheimer's disease risk in the production of β-amyloid peptides. Proc Natl Acad Sci USA. 2012;109:15307–11.

83. Dourlen P, Kilinc D, Malmanche N, Chapuis J, Lambert J-C. The new genetic landscape of Alzheimer's disease: from amyloid cascade to genetically driven synaptic failure hypothesis? Acta Neuropathol. 2019;138:221–36.

84. Shulman JM, Imboywa S, Giagtzoglou N, Powers MP, Hu Y, Devenport D, et al. Functional screening in Drosophila identifies Alzheimer's disease susceptibility genes and implicates Tau-mediated mechanisms. Hum Mol Genet. 2014;23:870–7.

85. Shulman JM, Chipendo P, Chibnik LB, Aubin C, Tran D, Keenan BT, et al. Functional screening of Alzheimer pathology genome-wide association signals in Drosophila. Am J Hum Genet. 2011;88:232–8.

86. Dourlen P, Fernandez-Gomez FJ, Dupont C, Grenier-Boley B, Bellenguez C, Obriot H, et al. Functional screening of Alzheimer risk loci identifies PTK2B as an

in vivo modulator and early marker of Tau pathology. Mol Psychiatry. 2017;22:874–83.

87. Ando K, De Decker R, Vergara C, Yilmaz Z, Mansour S, Suain V, et al. Picalm reduction exacerbates tau pathology in a murine tauopathy model. Acta Neuropathol. 2020;139:773–89.

88. Chapuis J, Hansmannel F, Gistelinck M, Mounier A, Van Cauwenberghe C, Kolen KV, et al. Increased expression of BIN1 mediates Alzheimer genetic risk by modulating tau pathology. Mol Psychiatry. 2013;18:1225–34.

89. Sartori M, Mendes T, Desai S, Lasorsa A, Herledan A, Malmanche N, et al. BIN1 recovers tauopathy-induced long-term memory deficits in mice and interacts with Tau through Thr348 phosphorylation. Acta Neuropathol. 2019;138:631–52.

90. Sottejeau Y, Bretteville A, Cantrelle F-X, Malmanche N, Demiaute F, Mendes T, et al. Tau phosphorylation regulates the interaction between BIN1's SH3 domain and Tau's proline-rich domain. Acta Neuropathol Commun. 2015;3:58.

91. Beecham GW, Hamilton K, Naj AC, Martin ER, Huentelman M, Myers AJ, et al. Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. PLoS Genet. 2014;10:e1004606.

92. Chia R, Sabir MS, Bandres-Ciga S, Saez-Atienzar S, Reynolds RH, Gustavsson E, et al. Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture. Nat Genet. 2021;53:294–303.

93. Jansen IE, van der Lee SJ, Gomez-Fonseca D, de Rojas I, Dalmasso MC, Grenier-Boley B, et al. Genome-wide meta-analysis for Alzheimer's disease cerebrospinal fluid biomarkers. Acta Neuropathol. 2022;144:821–42.

94. Franzmeier N, Rubinski A, Neitzel J, Ewers M. Alzheimer's Disease Neuroimaging Initiative (ADNI). The BIN1 rs744373 SNP is associated with increased tau-PET levels and impaired memory. Nat Commun. 2019;10:1766.

95. Franzmeier N, Ossenkoppele R, Brendel M, Rubinski A, Smith R, Kumar A, et al. The BIN1 rs744373 Alzheimer's disease risk SNP is associated with faster Aβ-associated tau accumulation and cognitive decline. Alzheimers Dement J Alzheimers Assoc. 2022;18:103–15.

96. Nott A, Holtman IR, Coufal NG, Schlachetzki JCM, Yu M, Hu R, et al. Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. Science. 2019;366:1134–9.

97. Novikova G, Kapoor M, Tcw J, Abud EM, Efthymiou AG, Chen SX, et al. Integration of Alzheimer's disease genetics and myeloid genomics identifies disease risk regulatory elements and genes. Nat Commun. 2021;12:1610.

98. Lopes K, de P, Snijders GJL, Humphrey J, Allan A, Sneeboer MAM, et al. Genetic analysis of the human microglial transcriptome across brain regions, aging and disease pathologies. Nat Genet. 2022;54:4–17.

99. Wang Y, Ulland TK, Ulrich JD, Song W, Tzaferis JA, Hole JT, et al. TREM2-mediated early microglial response limits diffusion and toxicity of amyloid plaques. J Exp Med. 2016;213:667–75.

100. Yuan P, Condello C, Keene CD, Wang Y, Bird TD, Paul SM, et al. TREM2 haplodeficiency in mice and humans impairs the microglia barrier function leading to decreased amyloid compaction and severe axonal dystrophy. Neuron. 2016;90:724–39.

101. Joshi P, Riffel F, Kumar S, Villacampa N, Theil S, Parhizkar S, et al. TREM2 modulates differential deposition of modified and non-modified Aβ species in extracellular plaques and intraneuronal deposits. Acta Neuropathol Commun. 2021;9:168.

102. Filipello F, Morini R, Corradini I, Zerbi V, Canzi A, Michalski B, et al. The microglial innate immune receptor TREM2 is required for synapse elimination and normal brain connectivity. Immunity. 2018;48:979–.e8.

103. Hong S, Dissing-Olesen L, Stevens B. New insights on the role of microglia in synaptic pruning in health and disease. Curr Opin Neurobiol. 2016;36:128–34.

104. Sierksma A, Lu A, Mancuso R, Fattorelli N, Thrupp N, Salta E, et al. Novel Alzheimer risk genes determine the microglia response to amyloid-β but not to TAU pathology. EMBO Mol Med. 2020;12:e10606.

105. Gratuze M, Chen Y, Parhizkar S, Jain N, Strickland MR, Serrano JR, et al. Activated microglia mitigate Aβ-associated tau seeding and spreading. J Exp Med. 2021;218:e20210542.

106. Kleineidam L, Chouraki V, Próchnicki T, van der Lee SJ, Madrid-Márquez L, Wagner-Thelen H, et al. PLCG2 protective variant p.P522R modulates tau pathology and disease progression in patients with mild cognitive impairment. Acta Neuropathol. 2020;139:1025–44.

107. Vautheny A, Duwat C, Aurégan G, Joséphine C, Hérard A-S, Jan C, et al. THY-Tau22 mouse model accumulates more tauopathy at late stage of the disease in response to microglia deactivation through TREM2 deficiency. Neurobiol Dis. 2021;155:105398.

108. De Strooper B, Karran E. The cellular phase of Alzheimer's disease. Cell. 2016;164:603–15.

109. Polderman TJC, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nat Genet. 2015;47:702–9.

110. Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM. Concepts, estimation and interpretation of SNP-based heritability. Nat Genet. 2017;49:1304–10.

111. Génin E. Missing heritability of complex diseases: case solved? Hum Genet. 2020;139:103–13.

112. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: genetic interactions create phantom heritability. Proc Natl Acad Sci. 2012;109:1193–8.

113. Bulik-Sullivan B, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47:291–5.

114. Golan D, Lander ES, Rosset S. Measuring missing heritability: Inferring the contribution of common variants. Proc Natl Acad Sci. 2014;111:E5272–E5281.

115. Weissbrod O, Flint J, Rosset S. Estimating SNP-based heritability and genetic correlation in case-control studies directly and with summary statistics. Am J Hum Genet. 2018;103:89–99.

116. 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. 2017;49:214.e7–214.e11.

117. Fuente J, de la, Grotzinger AD, Marioni RE, Nivard MG, Tucker-Drob EM. Integrated analysis of direct and proxy genome wide association studies highlights polygenicity of Alzheimer's disease outside of the APOE region. PLOS Genet. 2022;18:e1010208.

118. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh P-R, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet. 2015;47:1228–35.

119. Ni G, Moser G, Ripke S, Neale BM, Corvin A, Walters JTR, et al. Estimation of genetic correlation via linkage disequilibrium score regression and genomic restricted maximum likelihood. Am J Hum Genet. 2018;102:1185–94.

120. Zhang Q, Sidorenko J, Couvy-Duchesne B, Marioni RE, Wright MJ, Goate AM, et al. Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture. Nat Commun. 2020;11:4799.

121. Zhang Y, Qi G, Park J-H, Chatterjee N. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. Nat Genet. 2018;50:1318–26.

122. THE BRAINSTORM CONSORTIUM, Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, et al. Analysis of shared heritability in common disorders of the brain. Science. 2018;360:eaap8757.

123. Escott-Price V, Hardy J. Genome-wide association studies for Alzheimer's disease: bigger is not always better. Brain Commun. 2022;4:fcac125.

124. Grotzinger AD, Fuente J, Privé F, Nivard MG, Tucker-Drob EM. Pervasive downward bias in estimates of liability-scale heritability in GWAS Meta-Analysis: A Simple Solution. Biol Psychiatry. 2022. https://doi.org/10.1016/j.biopsych.2022.05.029.

125. Robette N, Génin E, Clerget-Darpoux F. Heritability: what's the point? What is it not for? A human genetics perspective. Genetica. 2022. https://doi.org/10.1007/s10709-022-00149-7.

126. Holland D, Frei O, Desikan R, Fan C-C, Shadrin AA, Smeland OB, et al. The genetic architecture of human complex phenotypes is modulated by linkage disequilibrium and heterozygosity. Genetics. 2021;217:iyaa046.

127. Levy-Sakin M, Pastor S, Mostovoy Y, Li L, Leung AKY, McCaffrey J, et al. Genome maps across 26 human populations reveal population-specific patterns of structural variation. Nat Commun. 2019;10:1025.

128. Mahmoud M, Gobet N, Cruz-Dávalos DI, Mounier N, Dessimoz C, Sedlazeck FJ. Structural variant calling: the long and the short of it. Genome Biol. 2019;20:246.

129. Collins RL, Brand H, Karczewski KJ, Zhao X, Alföldi J, Francioli LC, et al. A structural variation reference for medical and population genetics. Nature. 2020;581:444–51.

130. Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, et al. alpha-Synuclein locus triplication causes Parkinson's disease. Science. 2003;302:841.

131. Rovelet-Lecrux A, Hannequin D, Raux G, Le Meur N, Laquerrière A, Vital A, et al. APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. Nat Genet. 2006;38:24–26.

132. Le Guennec K, Quenez O, Nicolas G, Wallon D, Rousseau S, Richard A-C, et al. 17q21.31 duplication causes prominent tau-related dementia with increased MAPT expression. Mol Psychiatry. 2017;22:1119–25.

133. Li Y, Shaw CA, Sheffer I, Sule N, Powell SZ, Dawson B, et al. Integrated copy number and gene expression analysis detects a CREB1 association with Alzheimer's disease. Transl Psychiatry. 2012;2:e192.

134. Heinzen EL, Need AC, Hayden KM, Chiba-Falek O, Roses AD, Strittmatter WJ, et al. Genome-wide scan of copy number variation in late-onset Alzheimer's disease. J Alzheimers Dis JAD. 2010;9:69–77.

135. Swaminathan S, Shen L, Kim S, Inlow M, West JD, Faber KM, et al. Analysis of copy number variation in Alzheimer's disease: the NIALOAD/ NCRAD Family Study. Curr Alzheimer Res. 2012;9:801–14.

136. Chapman J, Rees E, Harold D, Ivanov D, Gerrish A, Sims R, et al. A genome-wide study shows a limited contribution of rare copy number variants to Alzheimer's disease risk. Hum Mol Genet. 2013;22:816–24.

137. Cuccaro D, De Marco EV, Cittadella R, Cavallaro S. Copy number variants in Alzheimer's disease. J Alzheimers Dis JAD. 2017;55:37–52.

138. De Roeck A, De Coster W, Bossaerts L, Cacace R, De Pooter T, Van, et al. NanoSatellite: accurate characterization of expanded tandem repeat length and sequence through whole genome long-read sequencing on PromethION. Genome Biol. 2019;20:239.

139. Hollox EJ, Zuccherato LW, Tucci S. Genome structural variation in human evolution. Trends Genet TIG. 2022;38:45–58.

140. Gusareva ES, Twizere J-C, Sleegers K, Dourlen P, Abisambra JF, Meier S, et al. Male-specific epistasis between WWC1 and TLN2 genes is associated with Alzheimer's disease. Neurobiol Aging. 2018;72:188.e3–188.e12.

141. Chang Y-C, Wu J-T, Hong M-Y, Tung Y-A, Hsieh P-H, Yee SW, et al. GenEpi: gene-based epistasis discovery using machine learning. BMC Bioinforma. 2020;21:68.

142. Sung YJ, Winkler TW, de Las Fuentes L, Bentley AR, Brown MR, Kraja AT, et al. A large-scale multi-ancestry genome-wide study accounting for smoking behavior identifies multiple significant loci for blood pressure. Am J Hum Genet. 2018;102:375–400.

143. Sarnowski C, Ghanbari M, Bis JC, Logue M, Fornage M, Mishra A, et al. Meta-analysis of genome-wide association studies identifies ancestry-specific associations underlying circulating total tau levels. Commun Biol. 2022;5:336.

144. Damotte V, van der Lee SJ, Chouraki V, Grenier-Boley B, Simino J, Adams H, et al. Plasma amyloid β levels are driven by genetic variants near APOE, BACE1, APP, PSEN2: a genome-wide association study in over 12,000 non-demented participants. Alzheimers Dement J Alzheimers Assoc. 2021;17:1663–74.

145. Kim H-R, Jung S-H, Kim J, Jang H, Kang SH, Hwangbo S, et al. Identifying novel genetic variants for brain amyloid deposition: a genome-wide association study in the Korean population. Alzheimers Res Ther. 2021;13:117.

146. Smith SM, Douaud G, Chen W, Hanayik T, Alfaro-Almagro F, Sharp K, et al. An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. Nat Neurosci. 2021;24:737–45.

147. Farfel JM, Yu L, Buchman AS, Schneider JA, De Jager PL, Bennett DA. Relation of genomic variants for Alzheimer disease dementia to common neuropathologies. Neurology. 2016;87:489–96.

148. Young J, Gallagher E, Koska K, Guetta-Baranes T, Morgan K, Thomas A, et al. Genome-wide association findings from the brains for dementia research cohort. Neurobiol Aging. 2021;107:159–67.

149. Jack CR, Bennett DA, Blennow K, Carrillo MC, Feldman HH, Frisoni GB, et al. A/T/N: an unbiased descriptive classification scheme for Alzheimer disease biomarkers. Neurology. 2016;87:539–47.

150. Palmqvist S, Janelidze S, Quiroz YT, Zetterberg H, Lopera F, Stomrud E, et al. Discriminative accuracy of plasma phospho-tau217 for Alzheimer disease vs other neurodegenerative disorders. JAMA. 2020;324:772–81.

151. Ashton NJ, Pascoal TA, Karikari TK, Benedet AL, Lantero-Rodriguez J, Brinkmalm G, et al. Plasma p-tau231: a new biomarker for incipient Alzheimer's disease pathology. Acta Neuropathol. 2021;141:709–24.

152. Ferrari R, Hernandez DG, Nalls MA, Rohrer JD, Ramasamy A, Kwok JBJ, et al. Frontotemporal dementia and its subtypes: a genome-wide association study. Lancet Neurol. 2014;13:686–99.

153. Finch N, Carrasquillo MM, Baker M, Rutherford NJ, Coppola G, Dejesus-Hernandez M, et al. TMEM106B regulates progranulin levels and the penetrance of FTLD in GRN mutation carriers. Neurology. 2011;76:467–74.

154. Tomé SO, Thal DR. Co-pathologies in Alzheimer's disease: just multiple pathologies or partners in crime? Brain. J Neurol. 2021;144:706–8.

155. Paushter DH, Du H, Feng T, Hu F. The lysosomal function of progranulin, a guardian against neurodegeneration. Acta Neuropathol. 2018;136:1–17.

156. Feng T, Lacrampe A, Hu F. Physiological and pathological functions of TMEM106B: a gene associated with brain aging and multiple brain disorders. Acta Neuropathol. 2021;141:327–39.

157. Minami SS, Min S-W, Krabbe G, Wang C, Zhou Y, Asgarov R, et al. Progranulin protects against amyloid β deposition and toxicity in Alzheimer's disease mouse models. Nat Med. 2014;20:1157–64.

158. International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009;460:748–52.

159. Speed D, Balding DJ. SumHer better estimates the SNP heritability of complex traits from summary statistics. Nat Genet. 2019;51:277–84.

160. Vilhjálmsson BJ, Yang J, Finucane HK, Gusev A, Lindström S, Ripke S, et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. Am J Hum Genet. 2015;97:576–92.

161. Lloyd-Jones LR, Zeng J, Sidorenko J, Yengo L, Moser G, Kemper KE, et al. Improved polygenic prediction by Bayesian multiple regression on summary statistics. Nat Commun. 2019;10:5086.

162. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc. 2020;15:2759–72.

163. Dudbridge F. Power and predictive accuracy of polygenic risk scores. PLoS Genet. 2013;9:e1003348.

164. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun. 2019;10:1776.

165. Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, et al. Common polygenic variation enhances risk prediction for Alzheimer's disease. Brain J Neurol. 2015;138:3673–84.

166. Sabuncu MR, Buckner RL, Smoller JW, Lee PH, Fischl B, Sperling RA, et al. The association between a polygenic Alzheimer score and cortical thickness in clinically normal subjects. Cereb Cortex. 2012;22:2653–61.

167. Mormino EC, Sperling RA, Holmes AJ, Buckner RL, De Jager PL, Smoller JW, et al. Polygenic risk of Alzheimer disease is associated with early- and late-life processes. Neurology. 2016;87:481–8.

168. Scelsi MA, Khan RR, Lorenzi M, Christopher L, Greicius MD, Schott JM, et al. Genetic study of multimodal imaging Alzheimer's disease progression score implicates novel loci. Brain. J Neurol. 2018;141:2167–80.

169. Escott-Price V, Myers AJ, Huentelman M, Hardy J. Polygenic risk score analysis of pathologically confirmed Alzheimer disease. Ann Neurol. 2017;82:311–4.

170. Lacour A, Espinosa A, Louwersheimer E, Heilmann S, Hernández I, Wolfsgruber S, et al. Genome-wide significant risk factors for Alzheimer's disease: role in progression to dementia due to Alzheimer's disease among subjects with mild cognitive impairment. Mol Psychiatry. 2017;22:153–60.

171. Louwersheimer E, Wolfsgruber S, Espinosa A, Lacour A, Heilmann-Heimbach S, Alegret M, et al. Alzheimer's disease risk variants modulate endophenotypes in mild cognitive impairment. Alzheimers Dement J Alzheimers Assoc. 2016;12:872–81.

172. Clark K, Leung YY, Lee W-P, Voight B, Wang L-S. Polygenic risk scores in Alzheimer's disease genetics: methodology, applications, inclusion, and diversity. J Alzheimers Dis JAD. 2022;89:1–12.

173. Leonenko G, Baker E, Stevenson-Hoare J, Sierksma A, Fiers M, Williams J, et al. Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores. Nat Commun. 2021;12:4506.

174. Bellou E, Stevenson-Hoare J, Escott-Price V. Polygenic risk and pleiotropy in neurodegenerative diseases. Neurobiol Dis. 2020;142:104953.

175. Mayhew AJ, Meyre D. Assessing the heritability of complex traits in humans: methodological challenges and opportunities. Curr Genom. 2017;18:332–40.

176. Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. Genome Med. 2020;12:44.

177. Nguyen TV, Eisman JA. Post-GWAS polygenic risk score: utility and challenges. JBMR. 2020;4:e10411.

178. Fulton-Howard B, Goate AM, Adelson RP, Koppel J, Gordon ML. Alzheimer's Disease Genetics Consortium, et al. Greater effect of polygenic risk score for Alzheimer's disease among younger cases who are apolipoprotein E-ε4 carriers. Neurobiol Aging. 2021;99:101.e1–101.e9.

179. Altmann A, Scelsi MA, Shoai M, de Silva E, Aksman LM, Cash DM, et al. A comprehensive analysis of methods for assessing polygenic burden on Alzheimer's disease pathology and risk beyond APOE. Brain Commun. 2020;2:fcz047.

180. Yengo L, Vedantam S, Marouli E, Sidorenko J, Bartell E, Sakaue S, et al. A saturated map of common genetic variants associated with human height. Nature. 2022;610:704–12.

181. Long JM, Holtzman DM. Alzheimer disease: an update on pathobiology and treatment strategies. Cell. 2019;179:312–39.

182. Deming Y, Li Z, Kapoor M, Harari O, Del-Aguila JL, Black K, et al. Genome-wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers. Acta Neuropathol. 2017;133:839–56.

183. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun. 2018;9:2098.

184. Ochoa D, Karim M, Ghoussaini M, Hulcoop DG, McDonagh EM, Dunham I. Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs. Nat Rev Drug Disco. 2022;21:551.

185. Oblak AL, Forner S, Territo PR, Sasner M, Carter GW, Howell GR, et al. Model organism development and evaluation for late-onset Alzheimer's disease: MODEL-AD. Alzheimers Dement N.Y. 2020;6:e12110.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

Initial conception of the manuscript (JCL), Preparation of initial version of the manuscript (JCL, AR, BGB, CB). Figure preparation (BGB, CB).

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION