Trends in Genetics

CellPress
REVIEWS

Review

Understanding Alzheimer Disease at the Interface between Genetics and Transcriptomics

Jan Verheijen[1,2] and Kristel Sleegers[1,2,*]

Over 25 genes are known to affect the risk of developing Alzheimer disease (AD), the most common neurodegenerative dementia. However, mechanistic insights and improved disease management remains limited, due to difficulties in determining the functional consequences of genetic associations. Transcriptomics is increasingly being used to corroborate or enhance interpretation of genetic discoveries. These approaches, which include second and third generation sequencing, single-cell sequencing, and bioinformatics, reveal allele-specific events connecting AD risk genes to expression profiles, and provide converging evidence of pathophysiological pathways underlying AD. Simultaneously, they highlight brain region- and cell-type-specific expression patterns, and alternative splicing events that affect the straightforward relation between a genetic variant and AD, re-emphasizing the need for an integrated approach of genetics and transcriptomics in understanding AD.

Highlights

Due to risk gene pleiotropy, difficulty in finding functional variants, and poor reflection of physiological complexity in genetic analysis, translation of new genetic findings for Alzheimer disease (AD) into functional mechanisms has been difficult.

Transcriptomic analysis has provided additional support for previously identified risk genes while also identifying novel associated genes, helping to elucidate mechanisms of disease.

Refinement of transcriptomics through 2nd and 3rd generation sequencing, single-cell sequencing and bioinformatics is revealing mechanisms involved in AD in previously unattainable detail, including brain region- and cell-type-specific expression changes and molecular processes such as transcript rescue events, challenging the direct interpretation of an association between genetic variant and phenotype.

Transcriptome analysis in postmortem brain has uncovered central biological pathways and central regulator 'hub' genes in disease, for example, SPI.1 and TYROBP in the brain immune response.

Alzheimer Disease: Common and Complex

Alzheimer's disease (AD) is a genetically complex, multifactorial disease that leads to neurodegenerative dementia. There is no cure for AD yet, and due to a high prevalence and continuously increasing incidence it poses a major threat to personal health as well as to the health care system. Patients display a progressive decline of cognitive capabilities, with characteristic early loss of episodic memory, eventually resulting in complete dependency and death. The disease is preceded by a long prodromal phase [1,2]. Neuropathological changes in the AD brain include progressive hippocampal and cortical atrophy, visible upon neuroimaging and macroscopic examination. Characteristic microscopic features are intracellular neurofibrillary tangles (NFTs) of hyperphosphorylated tau protein and extracellular depositions of Amyloid-β $(Aβ)_{1–42}$ peptide, accompanied by neuronal and synapse loss and reactive gliosis [3].

Initial evidence of a genetic etiology of AD was presented by the observation of families with multiple generations affected by a rare early onset form of AD (EOAD, <65 years). Molecular genetic investigation of these pedigrees resulted in the identification of the amyloid precursor protein (*APP*), presenilin 1 (*PSEN1*) and presenilin 2 (*PSEN2*) genes as disease genes in AD. Pathogenic mutations in these genes converge to a general mechanism of increased $Aβ_{1-42}$ accumulation or increased $Aβ_{1–42}/Aβ_{1–40}$ ratio. Hundreds of dominantly inherited pathogenic mutations have since been described for these genes, mostly in EOAD patients, although only explaining up to 10% of EOAD (reviewed in [4]).

[1]Neurodegenerative Brain Diseases Group, Center for Molecular Neurology, VIB, Antwerp, B-2610, Belgium
[2]Laboratory of Neurogenetics, Institute Born-Bunge, University of Antwerp, Antwerp, B-2610, Belgium

*Correspondence:
kristel.sleegers@molgen.vib-ua.be
(K. Sleegers).

Most patients have late-onset AD (LOAD). While a pathogenic mutation in *APP*, *PSEN1* or *PSEN2* is infrequently identified in LOAD patients, LOAD is typically considered multifactorial, with a strong polygenic component and an estimated heritability of up to 80%. The most well-known genetic risk factor for AD is the ***APOE* ε4 allele** (see Glossary), explaining approximately 25% of the heritability of liability [5]. Over the past decade, complex genetic research on AD has been successful in identifying additional genetic risk factors for AD, both low-penetrant common risk factors (e.g., [6]) and rare alleles with intermediate to high penetrance (e.g., [7–9]). While the amyloid cascade hypothesis has long dominated the efforts towards development of diagnostics and therapeutics for AD, gene discovery studies and pathway-based analysis of genome-wide studies shed light on a range of additional biological processes contributing to AD, proposing new targets for therapy development. Translational impact of these findings is still limited however, owing to – amongst others – pleiotropy of risk genes, limited understanding which genetic variants in identified risk genes or loci actually affect AD risk and how, and the physiological complexity of tissues which is insufficiently represented in genetic analysis.

While the field of AD genetics is still in active pursuit of additional genetic risk loci through **genome-wide association studies (GWAS)** and next generation sequencing studies on increasingly large AD cohorts, a trend is emerging of simultaneous interrogation of transcriptomics data to study the effect of newly identified genetic risk factors at the level of the transcriptome (e.g., [10]). In parallel, refinement of transcriptomics through second and third generation sequencing methodology, single-cell sequencing and bioinformatics analysis allows investigating functional mechanisms involved in AD in previously unattainable detail. Here, we review the state-of-the-art of genetic discovery and transcriptome investigations in multifactorial AD, with emphasis on the insights emerging at their interface.

Genetic Risk Loci and Pathways in Polygenic AD

Common Genetic Risk Factors

The search for genetic risk factors for polygenic AD was initially dominated by studies querying common genetic variation, most successfully through GWAS. At least 42 genes/loci have been associated with LOAD at genome-wide significance in at least one GWAS [6,11–20], of which *ABCA7*, *BIN1*, *CASS4*, *CD2AP*, *CD33*, *CELF1*, *CLU*, *CR1*, *EPHA1*, *FERMT2*, HLA-cluster, *INPP5D*, *MEF2C*, *MS4A6A*, *NME8*, *PICALM*, *PTK2B*, *SLC24A4/RIN3*, *SORL1*, *CELF1*, *NME8*, *FERMT2*, *CASS4*, *DGS2*, and *ZCWPW1* have been identified or confirmed in GWAS meta-analysis, and are since regarded as established risk genes/loci for LOAD. *CD33* and *DSG2* did not show genome-wide significance in the replication phase of the largest AD GWAS meta-analysis published to date [6]. Family-based GWAS approaches reported significant loci overlapping with those in case-control GWAS, including *APOE*, *BIN1*, *CLU*, and *CD33*. In addition, *PLXNA4, CUGBP2, TRPC4AP, ATXN1,* and *APOC1* as well as uncharacterized chromosome 14 locus 14q31.2 were reported [18–20] but not replicated in case-control approaches. Alternative analytical approaches of case-control GWAS detected additional risk loci such as *FRMD4A* in a sliding window haplotype-based approach [21], and *TP53INP1* and *IGHV1-67* in gene-wide analysis [22]. Expanding analysis of genome-wide variant data beyond single-variant or single-gene level to pathway-based analysis revealed common pathways associated with AD, despite differences in analytical approach and pathway definition between studies. Common biological pathways emerging from these studies are immune response, lipid metabolism, endocytosis, and cell adhesion molecule (CAM) pathways [10,23–29] (Table 1). A study using the methodologically distinct approach of genome-wide heritability partitioning confirmed enrichment of the adaptive and innate immune response in AD [30].

Glossary

***APOE* ε4 allele:** *APOE* is the gene encoding Apolipoprotein E. The ε4 allele is one of three possible isoforms, determined by two single nucleotide polymorphisms resulting in arginine residues at positions 112 and 158.
Cell type marker genes: canonical genes whose expression is associated with a specific cell (sub) type, used to identify and quantify cell types within tissue samples.
ChIP-Seq: sequencing approach applying chromatin immunoprecipitation to elucidate genetic regions interacting with proteins, applied in the identification of DNA regulatory regions.
Epigenome-wide association studies (EWAS): a genome-wide analysis of the association between epigenetic marks, such as DNA methylation, and a phenotype.
Expression quantitative trait loci (eQTL): genetic loci associated with regulation of gene expression levels.
Gene coexpression networks: construction of networks of genes with correlating expression direction across samples, applied to identify gene modules and the underlying biological mechanisms associated with a phenotype.
Genome-wide association studies (GWAS): a statistical approach to identify single-nucleotide polymorphisms associated with a phenotype, typically investigating millions of variants across the genome in very large cohorts.
Hub genes: central regulator genes with high connectivity in gene coexpression networks.
Meta-analysis: an epidemiological approach to combine and analyze results from previous studies. Is often performed to increase statistical power.
Splice quantitative trait locus (sQTL): a genetic locus causing a specific splice event of a gene
Transcriptome-wide association studies (TWAS): statistical analysis to identify genes whose expression is associated with the disease or trait of interest at transcriptome-wide scale, but without gene expression quantification. Gene expression levels are imputed based on a reference dataset.

Table 1. Significantly Enriched Pathways in GWAS Meta-analysis Studies on AD[a]

GWAS data sets analyzed	Consulted database	Lipid metabolism	Immune response	Endocytosis	Synaptic transmission	Cell adhesion molecules	Miscellaneous
GERAD/EADI [23]	ALIGATOR/GSA: KEGG, GO databases	Sterol transport ($P = 0.0079$), cholesterol transport ($P = 0.0079$)	Immunoglobulin mediated immune response ($P = 4 \times 10^{-3}$), immune response ($P = 3 \times 10^{-3}$)		Synaptic transmission, cholinergic ($P = 5.0 \times 10^{-3}$)		
EADI [24]	KEGG, GO databases		RIG-I-like receptor signaling ($P = 3 \times 10^{-2}$) Antigen processing and presentation ($P = 2.0 \times 10^{-2}$)				Regulation of autophagy ($P = 0.007$)
EADI [25]	Gencodis/DAVID: GO database			Intracellular transmembrane protein transport ($P = 7.2 \times 10^{-6}$)			
Discovery data EADI, replication data ADNI [26]	IGSEA: KEGG database. WebGestalt/DAVID: GO database		RIG-I-like receptor signaling ($P = 7.00 \times 10^{-4}$), Natural killer cell mediated cytotoxicity ($P = 8.56 \times 10^{-5}$), Antigen processing and presentation ($P = 3.50 \times 10^{-7}$)			Cell adhesion molecules KEGG ($P = 1.84 \times 10^{-6}$)	Regulation of autophagy ($P = 6.22 \times 10^{-5}$)
ADNI, composite measure of memory as phenotype [27]	GSA-SNP: BioCarta, KEGG, GO, Reactome databases		Allograft rejection ($P = 3.9 \times 10^{-2}$)		Transmission across chemical synapses ($P = 1.77 \times 10^{-4}$)	Focal adhesion ($P = 0.006$), Cell adhesion molecules (CAMs) ($P = 2.9 \times 10^{-2}$)	Calcium signaling pathway ($P = 1.17 \times 10^{-4}$), Viral myocarditis ($P = 0.039$), Long-term depression/ potentiation ($P = 8.0 \times 10^{-3}$)/ ($P = 1.9 \times 10^{-2}$)
Combined TGen1, NIA-LOAD/NCRAD, ADNI replication data [28]	KEGG, GO databases	Lipid transport ($P = 8.12 \times 10^{-9}$)		Endocytosis ($P = 2.24 \times 10^{-6}$)	Glutamate receptor signaling pathway ($P = 1.86 \times 10^{-11}$), axon guidance ($P = 1.20 \times 10^{-9}$)	Focal adhesion ($P = 2.63 \times 10^{-10}$)	Protein autophosphorylation ($P = 2.30 \times 10^{-12}$), transmembrane receptor protein kinase activity ($P = 1.18 \times 10^{-12}$), Calcium signaling pathway ($P = 1.27 \times 10^{-11}$)

Table 1. (continued)

GWAS data sets analyzed	Consulted database	Lipid metabolism	Immune response	Endocytosis	Synaptic transmission	Cell adhesion molecules	Miscellaneous
IGAP [10]	ALIGATOR/GSEA: KEGG, GO databases	Cholesterol transport ($P = 2.96 \times 10^{-9}$), sterol transport ($P = 3.91 \times 10^{-9}$)	Humoral immune response mediated by circulating immunoglobulin ($P = 3.27 \times 10^{-12}$), regulation of immune response ($P = 3.27 \times 10^{-12}$)	Regulation of endocytosis ($P = 1.31 \times 10^{-11}$)			Clathrin adaptor complex ($P = 1.20 \times 10^{-3}$), protein folding ($P = 1.60 \times 10^{-3}$)
GERAD integrated with transcriptome profiling data of temporal cortex [29].	KEGG				Axon guidance ($P = 3.03 \times 10^{-3}$)	Cell adhesion molecules (CAMs) ($P = 1.04 \times 10^{-5}$)	Calcium signaling pathway ($P = 2.08 \times 10^{-3}$), Viral myocarditis ($P = 4.00 \times 10^{-4}$), Purine metabolism ($P = 4.00 \times 10^{-4}$)

[a]*P* values indicate multiple-testing corrected values according to respective studies for terms showing strongest interaction within each pathway. Abbreviations: ADGC, Alzheimer's Disease Genetics Consortium; CHARGE, Cohorts for Heart and Aging in Genomic Epidemiology; EADI, European Alzheimer's Disease Initiative; GERAD, Genetic and Environmental Risk in Alzheimer's Disease; IGAP, International Genomics of Alzheimer's Project.

Rare Genetic Risk Factors

The investigation of rare variants in susceptibility of AD has gained significant momentum after the identification through combined whole genome (WGS) and whole exome sequencing (WES) of a rare non-synonymous mutation, p.R47H, in *TREM2* that increases risk of AD [7,31], which has since been widely replicated. Numerous independent studies further reported an enrichment of rare coding variants among AD patients in *SORL1* and *ABCA7*, which were already known to be implicated in AD risk through common risk factors. In *SORL1*, both heterozygous missense and premature termination codon (PTC) mutations were found, most notably in patients with EOAD and/or familial disease [32]. For *ABCA7*, strongest evidence of association was found for heterozygous PTC mutations, with widely varying onset age among carriers but an increased proportion of positive family history [8,33]. Large-scale sequencing efforts, such as the Alzheimer Disease Sequencing Project (ADSP) which includes ~11 000 participants, are providing further evidence of rare risk variants in genes initially identified in GWAS [34]. These studies are likely to provide insight into novel risk genes for AD in the near future. An exome-wide association analysis on rare and low-frequency variants using exome chip genotyping proposed rare variants in two additional genes, that is, *ABI3* and *PLGC2*, the latter being protective for AD [9].

Understanding AD through Transcriptome Analysis

Both classical gene discovery and pathway-based analysis of AD GWAS shed light on the range of biological processes that contribute to AD in addition to amyloid pathology. The straightforward interpretation and functional investigation of the association between a genetic variant and AD remains challenging, however, due to difficulty in identifying disease variants, pleiotropy of risk genes, and pathophysiological complexity. Hundreds of molecular factors putatively interact in multiple networks at different time points and at distinct biological levels including subcellular, cellular, tissue, and organic level. Increasingly, studies address this complexity by combining genetic data with gene expression studies. This allows studying the effect of genetic risk factors at the level of the transcriptome in a tissue, topology, time and cell-type-dependent manner, to identify similarities between disease networks, and uncover molecular interactors that relate to hub nodes connecting pathways. Transcriptome profiling is commonly performed by either microarray hybridization or next-generation RNA sequencing. Although both methods enable large-scale investigation of gene expression, the underlying principles of expression profiling differ fundamentally (Box 1).

AD Risk Genes and Loci at the Transcript Level

In an endeavor to provide insight in the molecular mechanisms underlying the association between a risk gene or locus and AD, numerous studies have investigated differential expression of these genes, motivated in part by the observation that many GWAS signals were strongest in noncoding – thus potentially regulatory – regions of the genome. Targeted studies in postmortem cortical tissue observed differential expression for a range of AD risk genes, including genes within the *HLA*-cluster, *MS4A*-cluster genes, *FRMD4A*, *CLU*, *SORL1*, *ABCA7*, *PTK2B*, *BIN1*, *TREM2*, and *CD33* [35–37]. Transcriptome-wide microarray studies and **meta-analyses** of these studies on AD and control brain largely corroborate these findings (Table 2). However, inconsistencies exist between studies, which could in part be due to technological variances and small differences in expression between phenotypes, but also to variation in investigated brain regions and cellular composition between samples.

For several of these loci genetic risk variants have been found associated with or directly affecting gene expression, indicating that the observed differences in expression are not merely by-products of the disease process. For example, a 3′ risk haplotype in *SORL1* has been

Box 1. Common Methods for Transcriptome Profiling

Microarray transcriptome profiling provides a cost-efficient approach to quantification of thousands of transcripts in parallel. Complementary DNA (cDNA) libraries reverse transcribed from RNA samples are introduced to the array prior to further analysis such as differential gene expression. However, the necessity of correcting for nonbiological effects on signal output represents a drawback. Microarray profiling is further confounded by variations in probe hybridization sensitivity within and between array platforms. Additionally, quantification of very low or very highly expressed transcripts proves problematic [77]. As probes cannot be designed for unknown sequences, microarray profiling is unable to identify novel transcripts. Probe design is restricted to knowledge of the genome at the time of array design. Human reference genome builds have been updated multiple times over the last ten years [78], which should be kept in mind when interpreting meta-analyses of several independent datasets.

By contrast, RNA sequencing (RNA-Seq) represents a hybridization-free method allowing massive parallel sequencing of cDNA using next-generation sequencing platforms. Selection of subset of RNAs from a total RNA sample is performed prior to reverse transcription to obtain a cDNA library enriched for RNAs, such as miRNA or (poly-adenylated) mRNA. Common procedures for enrichment involve removal of highly abundant ribosomal RNA by depletion of ribosomal RNA or pull-down of poly-adenylated (poly-A) RNAs by oligo-dT beads. Of note, RNAs lacking a poly-A tail including small mRNAs and non-coding RNAs (ncRNAs) are not retained with poly-A pulldown, while these molecules have been shown relevant to AD. For example, the poly-A tail lacking ncRNA 51A maps to intron 1 of the *SORL1* gene in antisense direction. ncRNA 51A is upregulated in AD frontal cortex and regulates SORL1 expression and splicing [79]. Known degradation of poly-A sequences in postmortem tissue presents a limitation for use of poly-A selected RNA-Seq in the context of AD, which can be overcome by 3′ mRNA-sequencing, where oligo-dT annealing is performed on the 3′ UTR, circumventing loss of non-polyadenylated transcripts. A disadvantage of this method is the inability to discriminate between different transcript isoforms. Quantification of transcripts generated by RNA-Seq involves read alignment and quality control filtering, where the number of reads aligned to each transcript represents a measure of transcript abundance. The RNA-Seq approach enables directional sequencing of cDNA, generating reads spanning multiple exons while maintaining directional information of reads. This allows for identification of novel splice events, which are of specific interest in the understanding of complex diseases [80].

associated with downregulation of *SORL1* expression [48], and rare loss-of-function mutations in *SORL1* have been identified in EOAD patients [49]. A three-nucleotide insertion near *BIN1*, rs59335482, increases *BIN1* expression, as well as risk of AD and NFT pathology in brain [50]. The HLA locus harbors a large cluster of genes involved in immune response modulation which includes *HLA-DRA*, *HLA-DRB5* and *HLA-DRB1*. HLA locus genes are commonly upregulated in AD brains, specifically *HLA-DRA*, which was found to be upregulated by microarray meta-analysis and studies investigating specific AD risk genes [39]. A risk variant, rs9271192, in *HLA-DRB1* is associated with increased expression of this gene [51]. The immune response locus MS4A includes AD risk genes *MS4A4A* and *MS4A6A,* among others. Both *MS4A4A* and *MS4A6A* are upregulated in AD brain [37,43,44]. AD risk variants have been associated with increased expression of *MS4A4A* [35]. *PICALM* transcripts have been found to be both upregulated [45] as well as downregulated in AD brains [40]. However, the GWAS index SNP rs3851179 which is associated with decreased risk of AD has been associated with increased PICALM expression [52]. Similarly, expression studies on *ABCA7* have been inconsistent, yet rare loss-of-function mutations have been associated with increased disease risk, suggesting reduced dosage of *ABCA7* increases risk of AD [33]. An AD risk allele, rs7143400-T, in the 3′ untranslated region (3′ UTR) of *FERMT2* creates a canonical binding site for microRNA miR-4505, resulting in allele-specific downregulation of FERMT2 in the presence of miR-4505 *in vitro* [53].

In addition to allele-specific differences in expression, alternative splicing of AD risk genes has been associated with the disease. Alternative splicing results in the production of distinct mRNA isoforms from one genetic locus, and is a widespread phenomenon, with reported rates of up to 95% of human multiexon genes undergoing splicing [54]. It represents an important part of gene product diversity and acts as a mechanism of regulating transcript expression levels by priming transcripts for nonsense-mediated mRNA decay. Gene products are regulated differentially between tissues and cell types, with the number of tissue-specific splice isoforms reported highest in the brain [55].

Table 2. Expression Regulation of AD Risk Genes in Brain[a]

Gene	Primary associated pathway	Primary expressed brain cell type	Differential expression direction in targeted AD risk gene studies	Differential expression direction in microarray meta-analyses [43–46]
ABCA7	Immune response/ Lipid metabolism	Microglia	+/– [38]	+ + ? +
BIN1	Endocytosis/Synaptic transmission	Oligodendrocytes	+ [37]	? + ? ?
CD2AP	Endocytosis	Endothelia		+ ? ? ?
CD33	Immune response	Microglia	+ [37]	+ + ? ?
CLU	Immune response/Lipid metabolism	Astrocyte	+ [36,37]	+ + ? ?
CR1	Immune response	Microglia	+ [37]	? + ? ?
EPHA1	Endocytosis/Synaptic transmission	Endothelia		? + ? ?
HLA-locus	Immune response	Microglia	HLA-DRA: + [39]	HLA-DRA: + ? + ?
MEF2C	Immune response/ Synaptic transmission	Endothelia/Microglia		– – – –
MS4A-cluster	Immune response	Microglia	MS4A4A and *MS4A6A:* + [37]	MS4A 4A/6A: + + ? ?
PICALM	Endocytosis/Synaptic transmission	Endothelia	– [40] + [45]	? ? + ?
PTK2B	Immune response/ Synaptic transmission	Microglia		? – ? ?
SORL1	Endocytosis	Astrocyte/Neuron	– [41]	? – ? ?
TREM2	Immune response	Microglia	+ [36,42]	+ ? ? ?

[a]Significant differential expression results are provided for targeted AD risk gene studies and meta-analysis. Direction of expression symbols +/–/? indicate upregulated/downregulated/missing expression direction in disease versus control. Meta-analysis incorporated samples derived from: Frontal cortex [43], temporal pole [44], frontal and temporal cortical regions [45], and frontal, temporal and entorhinal cortical regions [46]. Primary expressing cell types were based on data from Darmanis *et al*. 2015 [47], and the Barres lab brain RNA-seq webpage[iv].

A recent RNA-Seq study reported up to 91% of differentially expressed isoforms in AD postmortem brain to be unannotated [56]. Differential exon usage was reported for ~25% of genes analyzed in two separate studies on postmortem AD brain [44]. A novel method applying local splice variation (LSV) as a model for alternative splicing detection reported around 200 splice events distinct between AD and non-demented control brains [54]. Exon-array profiling of brain transcriptome in AD reported 22 candidate alternatively spliced genes relevant to disease [57]. RNA-Seq of frontal cortex showed a higher proportion of reads mapping to intronic regions in AD patients compared to controls, indicating differential alternative splicing events such as intron retention in disease. In particular, splice efficiency reduction was reported for several known AD genes including *APOE*, *BACE1*, *CLU*, *BIN1*, *PICALM*, *PSEN1*, and *SORL1* [58]. Targeted gene approaches further elucidated alternative splicing events relevant to disease in a number of genes involved in tau and amyloid processing, and several AD risk genes identified through GWAS approaches [59] (Box 2).

Box 2. Alternative Splicing Events in AD Risk Genes

Alternative transcript isoform generation by alternative splicing arises through the inclusion or exclusion of exons during pre-mRNA processing to result in functionally distinct RNA and protein products. Several mechanisms of splicing exist, including exon skipping, alternative donor or acceptor splice site usage, mutually exclusive exon usage, and intron retention. Of these, exon skipping is most prominent and is estimated to underlie up to 30% of AS events in vertebrates [81]. For numerous AD risk genes, alternative splicing events have been associated with AD risk or disease progression. For example, splicing of *SORL1* is regulated by an intron-1 antisense noncoding RNA. AD brains show overexpression of the resulting *SORL1* 51A isoform and corresponding decrease in SORL1 protein [79]. *CD33* shows alternative splicing controlled by a variant in linkage disequilibrium with the GWAS index variant in this gene. The *CD33* splice variant, lacking exon 2, alters *CD33* function and is associated with decreased risk of AD [82]. Up to ten *BIN1* isoforms have been reported, along with numerous splice variants. Tissue-specific alternative splicing events regulate interaction of *BIN1* relevant to AD, with seven isoforms specific to brain [83]. Alternative splicing of *CLU* generates an intracellular isoform, which is more abundant in the presence of Aβ *in vitro,* and is associated with amyloid pathology [84]. Interaction between the intracellular *CLU* isoform and *BIN1* is regulated by one of the *BIN1* splice variants [85]. A coding SNP in the splice enhancer site of exon 5 in the *PICALM* gene regulates splicing of exon 5. This splice variant is in linkage disequilibrium with the observed GWAS signal in *PICALM* [86]. Expression of an isoform lacking exon 13 has further been associated with the same GWAS signal SNP in AD [52]. Third generation, long-read, cDNA sequencing of *ABCA7* revealed numerous previously unobserved alternative splicing events that rescued the truncating effect of PTC mutations in the same gene, which might have phenotype-modifying effects [87]. By shedding light on a greater extent of gene product diversity than previously considered, 2nd and 3rd generation RNA sequencing studies further challenge the straightforward interpretation of the effect of genetic risk variants on disease, warranting deeper investigation.

Transcriptome-wide Profiling of Postmortem AD Brain

Correlation between variants within GWAS loci and transcriptome regulation are typically investigated by **expression quantitative trait loci (eQTL)** and **splice QTL (sQTL)** analysis. QTLs have been reported for variants in *CR1*, *HLA-DRB1*, *ZCWPW1*, *SLC24A4*, *CLU* and *MS4A4A*. However, the identified QTLs are not always within the GWAS haplotype, hence may not be relevant for the disease process, and replication of QTL findings remains inconclusive. One drawback of some genome-wide expression profiling studies is the relatively limited number of samples included in analyses. This is partly due to sparse availability of suitable source tissue, such as clinically characterized postmortem human brain. Due to the huge multiple testing burden of genome-wide and transcriptome-wide analyses, and the presence of potentially high-impact low-frequency variants, integrative analysis of DNA and RNA sequencing requires large well-characterized cohorts. Investigating e- and s-QTL effects in AD risk genes could further benefit from increased resolution introduced by multiple-tissue sequencing. In recent years, the extent of data generated by RNA sequencing is rapidly increasing. Access to large scale transcriptome profiling studies, such as through the Synapse platform[i], the Brain eQTL Almanac web server[ii], and the Genotype-Tissue expression (GTEx) project[iii] greatly facilitates genome-transcriptome integrated analysis. Recently, GWAS methods have been extended to enable **transcriptome-wide association studies (TWAS)**, through imputation of eQTL and sQTL data from these reference datasets onto large-scale GWAS data. A TWAS on AD revealed 61 sQTLs in known genes including *CLU*, *PTK2B*, and *CR1*; it also proposed novel candidates, including *AP2A1* which is an interactor of *PICALM* [60]. With rapid increase in sequencing depth and cohort size in RNA sequencing, eQTL and sQTL analysis is anticipated to elucidate additional genetic regulation of expression of risk genes.

Beyond hypothesis-driven expression studies focusing on known AD risk genes, microarray and RNA-sequencing based transcriptome studies of postmortem brain in AD and healthy controls have identified a myriad of differentially expressed genes and associated functional pathways. This yielded a core set of differentially expressed pathways including immune response, apoptosis, cell proliferation, energy metabolism, and synaptic transmission [43,46]. RNA-Seq of human AD postmortem brains additionally showed a large fraction of novel isoforms deregulated in parietal cortex and enrichment of pathways associated with neurite differentiation, immune response, and lipid metabolism [59]. These analyses thus corroborate findings of GWAS pathway analyses (Table 1).

Gene Coexpression Networks and Hub Genes

Unlike manually curated pathways of biological mechanisms and disease networks, **gene coexpression networks** are not restricted by previous findings, thus reducing the effect of literature bias and theoretical framework. An estimated 20–50% of transcripts detectable though RNA sequencing cannot be functionally annotated using public databases. Gene coexpression networks enable the identification of novel gene interactions and networks unrelated to currently known gene network clusters. Gene–gene correlation alterations or differential gene coexpression can be identified in absence of significant differential expression association. This holds the potential for identifying new '**hub genes**' in disease susceptibility and progression, that is, genes with many connections in the gene network which are considered crucial in maintaining network functionality. Genetic variants in hub genes are predicted to have a more profound effect on network dysfunction and disease than other genes in the network. *TYROBP* was identified as a central regulator of the top-ranked immune/microglia networks, also including *TREM2*, *MS4A4A*, *MS4A6A*, and *CD33* [61]. TYROBP is a binding partner of TREM2, and a genetic variant (p.D50_L51ins14) in *TYROBP* is reported to regulate TREM2 expression levels [62]. Upregulation of TYROBP in AD brains and interaction with TREM2 expression has been validated by independent studies [63]. TYROBP has been suggested to function in neuronal pruning and stimulation of microglial phagocytosis of amyloid-β, although the exact function of TYROBP in AD susceptibility and progress remains unclear [63]. Additionally, *PTK2B* has been reported as a hub regulator identified through microarray meta-analysis in the frontal cortex [43]. Protein–protein interaction analysis proposed PTK2B as a regulator of several differentially expressed genes, most of which are upregulated in AD and associated with the immune response including *TREM2,* suggesting an overlap between regulatory functions for *TYROBP* and *PTK2B*. The *SPI1* gene located within the *CELF1* GWAS risk locus has recently been identified as a regulator of the microglial activity pathway including *ABCA7*, *TREM2* and *TYROBP* (Box 3).

Epigenetic Regulation of Gene Expression

Epigenetic mechanisms, such as DNA methylation, histone modification and chromatin regulation by long noncoding RNAs, are known to regulate dynamic changes in gene expression in a cell-type- and tissue-specific manner. In postmitotic neurons, for example, epigenetic transcriptional regulation is involved in synaptic activity, learning and memory (reviewed in [64]). Epigenetic mechanisms have also been implicated in AD. The most studied epigenetic

Box 3. *SPI1* as Hub Gene in the Microglial Activity Pathway in AD

The identification of the role of the *SPI1* gene in AD exemplifies the added value of integrating findings at the genetic level with emerging insights from transcriptome profiling to provide insight in the biological mechanisms associated with AD. The protective rs1057233-G allele within the GWAS haplotype at the *CELF1* locus is reported to decrease *SPI1* expression in monocytes and macrophages and delay AD onset age [88]. The proposed functional mechanisms by which rs1057233 regulates *SPI1* expression is through altering the binding region of microRNA miR-569. *SPI1* encodes the PU.1 protein, which acts as a transcription factor in myeloid cells, including microglia and monocytes. PU.1 regulates gene expression by interaction with purine-rich regions in promotor regions. Binding of *SPI1* has been reported for genes *ABCA7*, *CD33*, *MS4A4A*, *MS4A6A*, *TREM2* and *TYROBP,* among others [88]. **ChIP-Seq** on mouse microglial cells identified Trem2 and Tyrobp as spi1 target genes, with spi1 binding to intron 1 in Trem2 and the promotor region of Tyrobp [89]. Combined RNA-Seq in mouse cell lines and human hippocampal neurons indicated expression of *SPI1* in neurodegeneration. Integrated epigenetic analysis of promotor regions proposes *SPI1* as increased-level expression enhancer [90]. *SPI1* putatively serves as a central hub in myeloid immune response in neurodegeneration, and regulates the balance between protective and pathogenic microglial activation in AD. Follow-up investigation of the role *SPI1* in AD could entail single-cell RNA sequencing to uncover cell-type-specific responses to *SPI1* network activation in early stages of disease, and the eQTL effects of the protective allele on neuronal, endothelial and microglial expression profiles. This knowledge could be beneficial to development of potential markers of inflammation, and present strategies to dampen early microglial inflammatory activation in disease.

mechanism in AD involves methylation of the 5th position of the cytosine base (5-methylcytosine, 5mC) in CpG dinucleotides. **Epigenome-wide association studies (EWAS)** that allow the simultaneous interrogation of methylation state at many hundred thousands of CpG sites across the genome have proposed several differentially methylated genes in AD pathogenesis. Although these analyses come with limitations [65], they have resulted in compelling findings for *ANK1, RHBDF2*, *RPL13*, and *CDH23* in association with AD neuropathology, as they were identified in two independent large-scale studies [66,67]. In line with the anticipated transcriptional regulation of cytosine methylation, these genes also showed changes in expression in AD brains [66]. Differential methylation has also been reported for several known AD risk genes, of which *ABCA7* is most consistently found between studies [44,66,68,69]. In line with evolving possibilities in imputation of gene expression data using public repositories, epigenomic imputation using publicly accessible reference data into large GWAS cohorts enables larger scale investigations. These studies may suffer a loss of statistical power due to imputation, and may be sensitive to issues such as size and quality of the reference dataset, and limited cross-tissue or cell-type predictability due to the dynamic nature of the epigenome, warranting caution in interpretation [70,71].

Hurdles in Interpretation

As AD is a neurodegenerative brain disease, many AD transcriptome studies have investigated transcriptional regulation in the brain. However, studies on human brain tissue are generally restricted to postmortem material, which represents an end-stage reflection of AD. In addition, postmortem interval and storage conditions of brain material influence transcriptome stability and preservation. Moreover, differences in cellular composition between diseased and healthy brain tissue may confound the interpretation of differential expression of cell-type-specific genes, such as genes implicated in synapse function in neurons, or inflammatory response genes in microglia. Due to neuronal loss and reactive gliosis, the observed changes in gene expression could be reflective of differences in cellular composition between affected and healthy tissue, or between different disease stages, rather than being indicative of disease-causing differences in transcriptional regulation. Different approaches exist to control for different cellular composition, for example through controlling for expression levels of cell-type-specific genes in the statistical framework, although a validated consensus set of **cell type marker genes** remains absent. A digital deconvolution of different brain cell types has been proposed, using reference panels of cell-type specific expression profiles [72]. Meanwhile, it will not be straightforward to untangle biologically relevant effects (cf. the genetic association of genes implicated in the immune response and endocytosis with AD risk) from side effects of neuronal loss or reactive gliosis. Indeed, ongoing work suggests that the immune response may act early in the disease process [73]. Studies on model systems offer an opportunity to distinguish between early and late (probably secondary) changes in gene expression. A systematic transcriptome-wide investigation of different transgenic mouse models harboring human *APP*, *PSEN1*, and *MAPT* mutations throughout the development of amyloid or tau pathology demonstrated, amongst others, a strong correlation between increased expression of immune response genes and amyloid pathology [74]. This included numerous known AD risk genes, as well as genes that were only later identified to be risk genes in genetic studies (Abi3 and Plcg2) [9].

The dynamic character of the transcriptome and heterogeneity between tissues and cell types further stresses the complexity of elucidating molecular mechanisms contributing to AD. Brain-region specific expression profiles of AD have been highlighted by comparative analysis of brain regions. Weighted gene coexpression network analysis (WGCNA) on 19 brain regions in 125 AD patients identified primary vulnerability of the temporal gyrus [45]. Gene coexpression

analysis of laser captured neurons in hippocampus, posterior cingulate cortex, and middle temporal gyrus elucidated distinct differences in onset and extent of AD pathology between these regions, suggesting a direct relation between brain topology and disease severity effects [75]. Single-cell RNA sequencing of adult and fetal non-demented temporal cortex cells identified distinct gene expression patterns between age, glial and neuronal cells types and between neuronal excitatory and inhibitory subtypes [47]. Neuronal subtype profiling on sorted cortical neurons revealed distinct differences in gene expression between excitatory and inhibitory neuronal subpopulations, and topology-specific neuronal composition between cortical regions [76]. Relative presence of excitatory and inhibitory subtypes varied between cortical regions, elucidating distinct neuronal subtype composition profiles between Brodmann areas. Since pathways involving inflammation are commonly identified upregulated in AD transcriptome analysis, investigating cell-type-specific signatures in AD could shed light on confounding variables in AD transcriptome profiling.

Concluding Remarks and Future Perspectives

Genetic investigation on AD susceptibility is uncovering a growing number of risk genes. In addition to rare variants with large effect sizes, over 25 risk genes have been identified through GWAS approaches, commonly exerting only moderate risk of AD. Pathway analysis of GWAS data has derived valuable insights and uncovered central roles for immunity, lipid metabolism, synaptic function and cell adhesion molecules in AD. Integration of gene expression data enables the identification of hub genes from these networks that may provide targets for biomarker or therapy development. While transcriptome profiling of AD brain has detected converging evidence of pathways involved in AD, it also highlights various levels of complexity in translating genetic findings into mechanistic insights. This includes brain-region- and cell-type-specific expression patterns, and alternative splicing events that may affect the straightforward relation between a genetic variant and a phenotype. Conversely, identification of genetic drivers of disease-related differences in gene expression patterns and molecular networks may help to delineate subgroups of individuals who could benefit from targeted clinical approaches. These observations re-emphasize the need for an integrated approach of genetics and transcriptomics (see Outstanding Questions).

While postmortem human brain material is relatively sparse and has inherent limitations as discussed above, additional investigations could benefit from analysis of *in vitro* cellular models to corroborate findings in patient material. Generation of neuronal cultures derived from reprogrammed patient material will allow for integrated analysis of postmortem brain material and living cultures from the same individual or group of individuals, for example, carrying a similar genetic risk variant. This may reveal molecular mechanisms specific to a genetic subtype. Nevertheless, it will be challenging to mimic the effects of a senescent or degenerative central nervous system *in vitro*. Ongoing refinements in iPSC and organotypic culture and RNA sequencing methods could corroborate postmortem brain studies in AD by providing a living model of matching genetic background and cellular composition for in depth investigation of disease mechanisms in AD. Importantly, evaluation of potential therapeutic agents could be performed using *in vitro* cultures.

The topological complexity in neuropathological vulnerability and transcriptional regulation in AD in the brain proposes a need to characterize in more detail the molecular mechanisms underlying disease susceptibility and progression. Additional experiments are warranted to elucidate local and regional brain cell (sub) type diversity in further detail to inform functional specialization of brain regions and regional susceptibility to disease. A comprehensive overview of cell-type-specific expression profiles in disease, and topological composition of cell types holds the potential to shed light on the diversity in brain deterioration and spread of neuropathology observed between AD patients.

Outstanding Questions

Can interindividual variation in risk allele expression, for example, through unmapped alternative splicing events, contribute to reduced penetrance of AD in variant carriers?

Can expression analysis segregate patients based on affected disease pathways, for example, between amyloid-negative, limbic-dominant, and hippocampal-sparing subcohorts?

Should outcomes of early comparative transcriptomics analyses of AD be revisited to control for confounding due to disease-stage-related changes in cell type composition of the tissue from which RNA was extracted?

Is it possible to reliably distinguish between primary changes in gene expression of immune response genes and neuronal genes on the one hand, and secondary changes in cell type composition of brain samples due to neuronal and synaptic loss and reactive gliosis on the other hand, when controlling for this confounding factor in comparative transcriptomics analyses?

Can an enhanced knowledge of the temporal, topological and cell-type-specific changes in gene expression and splicing further our understanding of the genetic contributors to AD?

Can disease-relevant cell-type-specific changes in gene expression be pharmacologically modulated?

Acknowledgments

The research in the authors' team is funded in part by the European Commission Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 305299 (AgedBrainSYSBIO), the Flanders Impulse Program on Networks for Dementia Research, the Alzheimer Research Foundation (SAO-FRA), and the University of Antwerp Research Fund.

Resources

[i]www.synapse.org/
[ii]www.braineac.org/
[iii]www.gtexportal.org/
[iv]https://web.stanford.edu/group/barres_lab/brain_rnaseq.html

References

1. Jack, C.R., Jr *et al.* (2011) Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 257–262
2. McKhann, G.M. *et al.* (2011) The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 263–269
3. Braak, H. and Braak, E. (1991) Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol.* 82, 239–259
4. Cacace, R. *et al.* (2016) Molecular genetics of early-onset Alzheimer's disease revisited. *Alzheimers Dement.* 12, 733–748
5. Cuyvers, E. *et al.* (2016) Genetic variations underlying Alzheimer's disease: evidence from genome-wide association studies and beyond. *Lancet Neurol.* 15, 857–868
6. Lambert, J.C. *et al.* (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* 45, 1452–1458
7. Guerreiro, R. *et al.* (2013) TREM2 variants in Alzheimer's disease. *N. Engl. J. Med.* 368, 117–127
8. Steinberg, S. *et al.* (2015) Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. *Nat. Genet.* 47, 445–447
9. Sims, R. *et al.* (2017) Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. *Nat. Genet.* 49, 1373–1384
10. Jones, L. *et al.* (2015) Convergent genetic and expression data implicate immunity in Alzheimer's disease. *Alzheimers Dement.* 11, 658–671
11. Lambert, J.C. *et al.* (2009) Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat. Genet.* 41, 1094–1099
12. Harold, D. *et al.* (2009) Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat. Genet.* 41, 1088–1093
13. Seshadri, S. *et al.* (2010) Genome-wide analysis of genetic loci associated with Alzheimer disease. *J. Am. Med. Assoc.* 303, 1832–1840
14. Naj, A.C. *et al.* (2011) Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat. Genet.* 43, 436–441
15. Hollingworth, P. *et al.* (2011) Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat. Genet.* 43, 429–435
16. Lee, J.H. *et al.* (2011) Identification of novel loci for Alzheimer disease and replication of CLU, PICALM, and BIN1 in Caribbean Hispanic individuals. *Arch. Neurol.* 68, 320–328
17. Miyashita, A. *et al.* (2013) SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians. *PLoS One* 8, e58618
18. Bertram, L. *et al.* (2008) Genome-wide association analysis reveals putative Alzheimer's disease susceptibility loci in addition to APOE. *Am. J. Hum. Genet.* 83, 623–632
19. Jun, G. *et al.* (2014) PLXNA4 is associated with Alzheimer disease and modulates tau phosphorylation. *Ann. Neurol.* 76, 379–392
20. Wijsman, E.M. *et al.* (2011) Genome-wide association of familial late-onset Alzheimer's disease replicates BIN1 and CLU and nominates CUGBP2 in interaction with APOE. *PLoS Genet.* 7, e1001308
21. Lambert, J.C. *et al.* (2013) Genome-wide haplotype association study identifies the FRMD4A gene as a risk locus for Alzheimer's disease. *Mol. Psychiatry* 18, 461–470
22. Escott-Price, V. *et al.* (2014) Gene-wide analysis detects two new susceptibility genes for Alzheimer's disease. *PLoS One* 9, e94661
23. Jones, L. *et al.* (2010) Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. *PLoS One* 5, e13950
24. Lambert, J.C. *et al.* (2010) Implication of the immune system in Alzheimer's disease: evidence from genome-wide pathway analysis. *J. Alzheimers Dis.* 20, 1107–1118
25. Hong, M.G. *et al.* (2010) Genome-wide pathway analysis implicates intracellular transmembrane protein transport in Alzheimer disease. *J. Hum. Genet.* 55, 707–709
26. Liu, G. *et al.* (2012) Cell adhesion molecules contribute to Alzheimer's disease: multiple pathway analyses of two genome-wide association studies. *J. Neurochem.* 120, 190–198
27. Ramanan, V.K. *et al.* (2012) Genome-wide pathway analysis of memory impairment in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort implicates gene candidates, canonical pathways, and networks. *Brain. Imaging Behav.* 6, 634–648
28. Perez-Palma, E. *et al.* (2014) Overrepresentation of glutamate signaling in Alzheimer's disease: network-based pathway enrichment using meta-analysis of genome-wide association studies. *PLoS One* 9, e95413
29. Xiang, Z. *et al.* (2015) Integrating genome-wide association study and brain expression data highlights cell adhesion molecules and purine metabolism in Alzheimer's disease. *Mol. Neurobiol.* 52, 514–521
30. Gagliano, S.A. *et al.* (2016) Genomics implicates adaptive and innate immunity in Alzheimer's and Parkinson's diseases. *Ann. Clin. Transl. Neurol.* 3, 924–933
31. Jonsson, T. *et al.* (2013) Variant of TREM2 associated with the risk of Alzheimer's disease. *N. Engl. J. Med.* 368, 107–116
32. Pottier, C. *et al.* (2012) High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. *Mol. Psychiatry* 17, 875–879
33. Cuyvers, E. *et al.* (2015) Mutations in ABCA7 in a Belgian cohort of Alzheimer's disease patients: a targeted resequencing study. *Lancet Neurol.* 14, 814–822
34. Beecham, G. (2017) Whole-genome sequencing in familial late-onset Alzheimer's disease identifies rare variation in AD candidate genes, Alzheimer's Association International Conference | July 16-20, 2017, Alzheimer's & Dementia: The Journal of the Alzheimer's Association. London, England

35. Allen, M. *et al.* (2012) Novel late-onset Alzheimer disease loci variants associate with brain gene expression. *Neurology* 79, 221–228
36. Martiskainen, H. *et al.* (2015) Transcriptomics and mechanistic elucidation of Alzheimer's disease risk genes in the brain and *in vitro* models. *Neurobiol. Aging* 36, 1221.e15–28
37. Karch, C.M. *et al.* (2012) Expression of novel Alzheimer's disease risk genes in control and Alzheimer's disease brains. *PLoS One* 7, e50976
38. Zhao, Q.F. *et al.* (2015) ABCA7 in Alzheimer's disease. *Mol. Neurobiol.* 51, 1008–1016
39. Yokoyama, J.S. *et al.* (2016) Association Between Genetic Traits for Immune-Mediated Diseases and Alzheimer Disease. *JAMA Neurol.* 73, 691–697
40. Zhao, Z. *et al.* (2015) Central role for PICALM in amyloid-beta blood-brain barrier transcytosis and clearance. *Nat. Neurosci.* 18, 978–987
41. Scherzer, C.R. *et al.* (2004) Loss of apolipoprotein E receptor LR11 in Alzheimer disease. *Arch. Neurol.* 61, 1200–1205
42. Lue, L.F. *et al.* (2015) TREM2 protein expression changes correlate with Alzheimer's disease neurodegenerative pathologies in post-mortem temporal cortices. *Brain Pathol.* 25, 469–480
43. Li, X. *et al.* (2015) Integrated genomic approaches identify major pathways and upstream regulators in late onset Alzheimer's disease. *Sci. Rep.* 5, 12393
44. Humphries, C. *et al.* (2015) Alzheimer disease (AD) specific transcription, DNA methylation and splicing in twenty AD associated loci. *Mol. Cell. Neurosci.* 67, 37–45
45. Wang, M. *et al.* (2016) Integrative network analysis of nineteen brain regions identifies molecular signatures and networks underlying selective regional vulnerability to Alzheimer's disease. *Genome Med.* 8, 104
46. Ciryam, P. *et al.* (2016) A transcriptional signature of Alzheimer's disease is associated with a metastable subproteome at risk for aggregation. *Proc. Natl. Acad. Sci. U. S. A.* 113, 4753–4758
47. Darmanis, S. *et al.* (2015) A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. U. S. A.* 112, 7285–7290
48. Caglayan, S. *et al.* (2012) Identification of Alzheimer disease risk genotype that predicts efficiency of SORL1 expression in the brain. *Arch. Neurol.* 69, 373–379
49. Verheijen, J. *et al.* (2016) A comprehensive study of the genetic impact of rare variants in SORL1 in European early-onset Alzheimer's disease. *Acta Neuropathol.* 132, 213–224
50. Chapuis, J. *et al.* (2013) Increased expression of BIN1 mediates Alzheimer genetic risk by modulating tau pathology. *Mol. Psychiatry* 18, 1225–1234
51. Allen, M. *et al.* (2015) Late-onset Alzheimer disease risk variants mark brain regulatory loci. *Neurol. Genet.* 1, e15
52. Parikh, I. *et al.* (2014) Genetics of PICALM expression and Alzheimer's disease. *PLoS One* 9, e91242
53. Delay, C. *et al.* (2016) miRNA-dependent target regulation: functional characterization of single-nucleotide polymorphisms identified in genome-wide association studies of Alzheimer's disease. *Alzheimers Res. Ther.* 8, 20
54. Vaquero-Garcia, J. *et al.* (2016) A new view of transcriptome complexity and regulation through the lens of local splicing variations. *eLife* 5, e11752
55. Xu, Q. *et al.* (2002) Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res.* 30, 3754–3766
56. Mills, J.D. *et al.* (2012) Alternative splicing of mRNA in the molecular pathology of neurodegenerative diseases. *Neurobiol. Aging* 33, 1012.e11–24
57. Lai, M.K. *et al.* (2014) Genome-wide profiling of alternative splicing in Alzheimer's disease. *Genom. Data* 2, 290–292
58. Bai, B. *et al.* (2013) U1 small nuclear ribonucleoprotein complex and RNA splicing alterations in Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* 110, 16562–16567
59. Mills, J.D. *et al.* (2013) RNA-Seq analysis of the parietal cortex in Alzheimer's disease reveals alternatively spliced isoforms related to lipid metabolism. *Neurosci. Lett.* 536, 90–95
60. Raj, T. (2017) Genetically Regulated Transcriptomic Study of Alzheimer's Disease Yields Mechanistic Insights, Alzheimer's Association International Conference, July 16-20, 2017, Alzheimer's & Dementia: The Journal of the Alzheimer's Association. London, England
61. Zhang, B. *et al.* (2013) Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* 153, 707–720
62. Pottier, C. *et al.* (2016) TYROBP genetic variants in early-onset Alzheimer's disease. *Neurobiol. Aging* 48, 222 e9–222.e15
63. Ma, J. *et al.* (2015) TYROBP in Alzheimer's disease. *Mol. Neurobiol.* 51, 820–826
64. Christopher, M.A. *et al.* (2017) Neuroepigenetic mechanisms in disease. *Epigenetics Chromatin* 10, 47
65. Lord, J. and Cruchaga, C. (2014) The epigenetic landscape of Alzheimer's disease. *Nat. Neurosci.* 17, 1138–1140
66. De Jager, P.L. *et al.* (2014) Alzheimer's disease pathology is associated with early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat. Neurosci.* 17, 1156–1163
67. Lunnon, K. *et al.* (2014) Cross-tissue methylomic profiling strongly implicates a role for cortex-specific deregulation of ANK1 in Alzheimer's disease neuropathology. *Nat. Neurosci.* 17, 1164–1170
68. Chibnik, L.B. *et al.* (2015) Alzheimer's loci: epigenetic associations and interaction with genetic factors. *Ann. Clin. Transl. Neurol.* 2, 636–647
69. Yu, L. *et al.* (2015) Association of Brain DNA methylation in SORL1, ABCA7, HLA-DRB5, SLC24A4, and BIN1 with pathological diagnosis of Alzheimer disease. *JAMA Neurol.* 72, 15–24
70. Rawlik, K. *et al.* (2016) Imputation of DNA methylation levels in the brain implicates a risk factor for Parkinson's disease. *Genetics* 204, 771–781
71. Ernst, J. and Kellis, M. (2015) Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nat. Biotechnol.* 33, 364–376
72. Harari, O. (2017) Cell-Type Profiling to Identify the Transcriptomic Downstream Events Triggered By Early-Onset Autosomal Dominant AD Mutations Alzheimer's Association International Conference | July 16-20, 2017, Alzheimer's & Dementia: The Journal of the Alzheimer's Association, London, England
73. Hong, S. *et al.* (2016) Complement and microglia mediate early synapse loss in Alzheimer mouse models. *Science* 352, 712–716
74. Matarin, M. *et al.* (2015) A genome-wide gene-expression analysis and database in transgenic mice during development of amyloid or tau pathology. *Cell Rep.* 10, 633–644
75. Ray, M. *et al.* (2010) Analysis of Alzheimer's disease severity across brain regions by topological analysis of gene co-expression networks. *BMC Syst. Biol.* 4, 136
76. Lake, B.B. *et al.* (2016) Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* 352, 1586–1590
77. Shendure, J. (2008) The beginning of the end for microarrays? *Nat. Methods* 5, 585–587
78. Tyner, C. *et al.* (2017) The UCSC Genome Browser database: 2017 update. *Nucleic Acids Res.* 45, D626–D634
79. Ciarlo, E. *et al.* (2013) An intronic ncRNA-dependent regulation of SORL1 expression affecting Aβ formation is upregulated in post-mortem Alzheimer's disease brain samples. *Dis. Model Mech.* 6, 424–433
80. Sutherland, G.T. *et al.* (2011) Understanding the pathogenesis of Alzheimer's disease: will RNA-Seq realize the promise of transcriptomics? *J. Neurochem.* 116, 937–946
81. Blencowe, B.J. (2006) Alternative splicing: new insights from global analyses. *Cell* 126, 37–47

82. Malik, M. *et al.* (2013) CD33 Alzheimer's risk-altering polymorphism, CD33 expression, and exon 2 splicing. *J. Neurosci.* 33, 13320–13325

83. Tan, M.S. *et al.* (2013) Bridging integrator 1 (BIN1): form, function, and Alzheimer's disease. *Trends Mol. Med.* 19, 594–603

84. Killick, R. *et al.* (2014) Clusterin regulates beta-amyloid toxicity via Dickkopf-1-driven induction of the wnt-PCP-JNK pathway. *Mol. Psychiatry* 19, 88–98

85. Zhou, Y. *et al.* (2014) Intracellular clusterin interacts with brain isoforms of the bridging integrator 1 and with the microtubule-associated protein tau in Alzheimer's disease. *PLoS One* 9, e103187

86. Schnetz-Boutaud, N.C. *et al.* (2012) Identification and confirmation of an exonic splicing enhancer variation in exon 5 of the Alzheimer disease associated PICALM gene. *Ann. Hum. Genet.* 76, 448–453

87. De Roeck, A. *et al.* (2017) Deleterious ABCA7 mutations and transcript rescue mechanisms in early onset Alzheimer's disease. *Acta Neuropathol.* 134, 475–487

88. Huang, K.L. *et al.* (2017) A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. *Nat. Neurosci.* 20, 1052–1061

89. Satoh, J. *et al.* (2014) A comprehensive profile of ChIP-Seq-based PU.1/Spi1 target genes in microglia. *Gene Regul. Syst. Biol.* 8, 127–139

90. Gjoneska, E. *et al.* (2015) Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature* 518, 365–369