
Access from the University of Nottingham repository: http://eprints.nottingham.ac.uk/2749/1/alz.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the Creative Commons Attribution Non-commercial No Derivatives licence and may be reused according to the conditions of the licence. For more details see: http://creativecommons.org/licenses/by-nc-nd/2.5/

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk
Review: The genetics of Alzheimer’s disease; putting flesh on the bones

C. Medway and K. Morgan
Translation Cell Sciences – Human Genetics, School of Life Sciences, Queens Medical Centre, Nottingham, UK


The genetics of Alzheimer’s disease; putting flesh on the bones

For two decades the search for genes involved in Alzheimer’s disease brought little reward; it was not until the advent of genome-wide association studies (GWAS) that genetic associations started to be revealed. Since 2009 increasingly large GWAS have revealed 20 loci, which in itself is a substantial increase in our understanding, but perhaps the more important feature is that these studies have highlighted novel pathways that are potentially involved in the disease process. This commentary assembles our latest knowledge while acknowledging that the casual functional variants, and undoubtedly other genes are still yet to be discovered. This is the challenge that remains and the promise of next-generation sequencing is anticipated as there are a number of large initiatives which themselves should start to yield information before long.

Keywords: Alzheimer’s, GWAS, NGS

Fifteen years separated the discovery of APOE, the first genetic risk factor for sporadic, late-onset Alzheimer’s disease (LOAD, also referred to as sporadic Alzheimer’s disease – sAD), and the first replicable genome-wide associations in 2009. However with patience comes reward; a recent meta-analysis by the International Genomics of Alzheimer’s Project (IGAP) reported 11 new Alzheimer’s susceptibility loci (CASS4, CELF1, FERMT2, HLA-DRB5/HLA-DRB1, INPP5D, MEF2C, NME8, PTK2B, SLC24A4/RIN3, SORL1 and ZCWPW1), and confirmed eight (CR1, BIN1, CD2AP, EPHA1, CLU, MS4A6A, PICALM and ABOCA7) of the nine previously reported genome-wide associations in addition to APOE [1]; the exception being CD33 which failed to replicate. Consequently genetic discoveries within the last 5 years account for ~47% of the population attributable risk (PAR) of LOAD. This rises to ~61% with the established APOE haplotype (Table 1).

However there is still a substantial component of ‘missing heritability’ waiting to be detected, which may be accounted for by multiple variants imparting modest effect (polygenic model) and/or fewer rare mutations of larger effect [4].

Unlike the rarer, early onset familial form of Alzheimer’s disease (fAD), the result of deterministic mutations in one of three amyloid processing genes (APP, PSEN1 or PSEN2), the aetiology of LOAD remains unknown. Considering LOAD shares clinical and neuropathological features with the familial disease, it is attractive to speculate that abnormalities within the amyloid pathway are similarly culpable. However, despite years of research, there is no concrete evidence that the ‘amyloid cascade’ is causally linked to late-onset disease, or to ratify its therapeutic potential.

The era of the genome-wide association study (GWAS) heralded the arrival of new LOAD alleles, each common (minor allele frequency, MAF > 5%) and conveying modest genetic effects [3,5–8]. Enrichment of the first eight genome-wide associated genes within three pathways (cholesterol metabolism, immune system function and synaptic vesicle recycling/endocytosis) [9,10] nomi-
nated novel pathways for therapeutic intervention. While these pathways may modify amyloid aggregation and/or clearance, to what extent these they are implicated in amyloid homeostasis remains to be determined [11,12]. A summary of how these new genes link to pathways likely to be involved in sAD is presented (see Figure 1).

New sAD genes, new pathways?

The IGAP consortium recently reported 11 new genes for sAD (Table 1b) [1]. IGAP drew on a combined resource of 74 046 samples, the largest GWAS meta-analysis of LOAD to date. The ensuing power boost elevated new risk alleles previously masked due to low frequency, weak genetic effects or incomplete tagging of the causal mutation. This may explain why SORL1, which had long been suspected of harbouring genuine disease alleles using candidate gene approaches [13], attained genome-wide significance in this series but not earlier GWAS. The encoded sortilin-related receptor is a neuronal APOE receptor and is understood to direct APP trafficking to amyloidogenic endocytic pathways [14].

Similarly, immune system function is represented in new candidate genes. Most notable are the association signals within the HLA region (which encompasses a number of candidate genes, the strongest associations are found at \textit{DRB1} and \textit{DRB5}). These genes have been previously associated with Parkinson’s disease [15] and multiple sclerosis [16]. \textit{INPP5D} encodes a member of the inositol polyphosphate-5-phosphatase family of enzymes involved in second messenger signalling in myeloid cells. \textit{INPP5D} affects pathways associated with cell proliferation and the regulation of inflammatory responses [17].

The remaining genes (PTK2B, MEFC2, CASS4, FERMT2, CELF1, ZCWPW1, SLC24A4, NME8) cannot be readily attributed to pathways which are biologically relevant to LOAD. Among them are cell adhesion molecules and kinases which signal in pathways attributed to...
neuronal function. PTK2B is a cytoplasmic protein tyrosine kinase expressed in brain, where it is understood to regulate neuronal activity via the MAP-kinase signalling pathway. Specific functions in long-term potentiation (LTP) and memory have been proposed [18].

MEF2C is also expressed in the cortex and is involved in the MAP kinase signalling pathway, responding to calcium influx to activate survival genes [19]. MEF2C encodes a transcription factor of the MEF2 family, which has a widely reported role in muscle (cardiac and skeletal) and vascular development [20]. Like PTK2B, MEF2C has been implicated in hippocampal synaptic plasticity and LTP [21]. Haploinsufficiency of MEF2C manifests mental retardation and epilepsy [22].

PTK2B is a focal adhesion kinase (FAK2), which is known to interact with the CASS4 (also genetically implicated in the IGAP series) homologue NEDD9 and integrin to regulate apoptosis and cell adhesion [23]. While CASS4 encodes a poorly characterized scaffold protein, variants in other members of the CAS gene family (NEDD9) have been associated with dementia and Alzheimer’s disease [24,25]. Fermatin family member 2 [FERMT2 or kindling-2 (KIND2)] is a neuronally expressed adhesion molecule implicated in actin cytoskeleton organization [26]. FERMT2 also plays a role in signalling via integrin activation and the Wnt signalling pathway [27]. Abnormal Wnt signalling has previously attracted attention as a candidate pathway in Alzheimer’s disease and other dementing syndromes [28].

CELF1, ZCWPW1, SLC24A4 and NME8 are particularly poorly characterized. CELF1 (CUGBP1 – CUG triplet repeat-binding protein 1) encodes an RNA-binding protein involved in alternative mRNA splicing, editing and stability. A CUG triplet repeat expansion is responsible for myotonic dystrophy type-1. Given the size of the associated locus in the IGAP publication, the causal gene may lie further afield; a MAP-kinase activating death domain 130kb downstream has attracted attention although RAPSN, a receptor-associated protein found at the synapse, is also within this region. Similarly, the association signal attributed to ZCWPW1 (rs1476679), encoding an epigenetic regulatory protein [29], is part of a larger LD block, encompassing many candidate genes. SLC24A4 (NCKX4) is a sodium/potassium/calcium exchanger expressed in human brain [30]. Variation in SLC24A4 gene has been associated with hair, eye and skin pigmentation [31]. More recently variation in SLC24A4 has been associated with lipid metabolism [32]. It is important to note that the SLC24A4 association signal (rs10498633) is towards the 3’-end of the gene, which is neighboured by RIN3. RIN3 interacts with BIN1 (amphiphysin-2), and may be part of the endocytic machinery [33]. Finally NME8, encoding a Thioredoxin Domain-containing Protein, is associated with primary ciliary dyskinesia-6 [34].

Towards identifying the causal alleles: why we find genome-wide association signals

In each instance identification of the causal allele(s) will be instrumental towards conclusively pinpointing the affected gene, accurately attributing risk to each gene/loci, and to fully appreciate the mechanism of pathogenesis. However, given that the causal mutations underpinning earlier genome-wide associations remain largely unknown this will be a considerable challenge. This has led some to postulate that we need to dig deeper.
resequencing candidate loci to identify rare mutation(s) explaining GWAS associations [35]. Coined ‘synthetic associations’ [36], there is no evidence that any of the common Alzheimer’s disease alleles discovered to date are precipitated through correlation (linkage disequilibrium) to rarer causal mutations. Furthermore, it would be expected that GWAS ‘hits’ conforming to the synthetic model will have allele frequencies skewed towards the lower end of the allele spectrum, which is not the rule in complex disorders [37]. The more common the associated GWAS SNPs (single nucleotide polymorphisms), the larger the effect (odds ratio) at the rare position(s) would have to be to fully explain the signal – this makes it more likely that they would have already been detected through family-based linkage approaches [37]. IGAP reports three alleles with frequencies below 10%. It will remain to be seen if, in these instances, genome resequencing unearth rare alleles evoking these signals as they represent the best potential examples to date that are possibly compatible with this concept.

Genome resequencing has already contributed to our understanding of sAD genetics. Two independent groups have reported rare coding variants within TREM2 (exon 2) conferring increased risk for LOAD [2,38]. TREM2 encodes a receptor expressed on myeloid cells, and is understood to mediate anti-inflammatory responses. Interestingly, this gene has been previously associated with other dementing phenotypes; the rare recessive Nasu-Hakola disease [39], early-onset dementia [40] and frontotemporal-like dementia [41]. Although there is evidence of allelic heterogeneity, with greater burden in cases, carriers of one nonsynonymous change (exon 2) conferring increased risk for LOAD [2,38]. TREM2 encodes a receptor expressed on myeloid cells, and is understood to mediate anti-inflammatory responses. Interestingly, this gene has been previously associated with other dementing phenotypes; the rare recessive Nasu-Hakola disease [39], early-onset dementia [40] and frontotemporal-like dementia [41]. Although there is evidence of allelic heterogeneity, with greater burden in cases, carriers of one nonsynonymous change (exon 2) conferring increased risk for LOAD [2,38].

International Genomics of Alzheimer’s Project reports a common variant (rs9381040), 3.4 kb downstream of TREM2’s neighbour TREML2, which falls just short of genome-wide significance. The fact that the alleles (R47H and rs9381040) are in complete linkage disequilibrium (D’ = 1) (there are no recombination hotspots segregating these positions), coupled with the observation that the TREM2 risk allele is in phase with the major allele at the TREML2 SNP, is consistent with a synthetic model. This would explain why the TREML2 minor allele reports a protective effect while the TREM2 allele is associated with risk. Interestingly, there is also prior linkage evidence of the TREML2 locus in Alzheimer’s disease [42]. In order to fully evaluate to what extent rare mutations in the TREM2 locus account for the TREML2 signal, a greater understanding of the allelic heterogeneity will have to be sought. Interestingly this observation may represent the interface, or juxta-positioning, of GWAS and NGS technologies. As GWAS are conducted with ever increasing sample sizes they will undoubtedly start to uncover rarer disease associated variants as exemplified by the IGAP study. This bodes well for the concept that more in-depth NGS will be able to contribute to some of the remaining ‘missing heritability’.

One recent observation that disease causing, coding mutations for common disorders are relatively ‘young’ – in fact occurred within the last 10 000 years [43] adds impetus to the NGS drive. For a common variant (MAF > 5%) to become fixed, and hence readily detectable in a GWAS, the change needs to have occurred over 200 000 years ago. Previously we have been using modern technologies to investigate ancient DNA changes; we are on the verge of an era where we are now able to use these approaches to look at ‘modern’ DNA changes, that is, the contribution of rare variants with much lower MAFs.

**Putting flesh on the bones; osteoclasts, microglia and Alzheimer’s disease**

Defects in the TREM2-DAP12 signalling pathway in osteoclasts and microglia, both of which are derived from myeloid progenitor cells, may explain the dual bone and brain manifestations in Nasu-Hakola disease. DAP12 and TREM2 are both expressed in osteoclasts and microglia (and dendritic cells and oligodendrocytes). Osteoclasts oversee the demineralization, reabsorption and recycling (via endocytosis) of bone; a similar recycling/removal mechanism for microglia within the brain is plausible.

It is interesting to note that other LOAD candidate genes have been implicated in osteoclast function downstream of TREM2-DAP12 signalling. PTK2B (Pyk2) is one of the kinases acting downstream of DAP12 [44]. INPP5D negatively regulates osteoclast formation and function by directly binding to DAP12 and preventing PI3K signalling [45,46]. CASS4 family members (CASL and CASS1), a downstream target of PTK2B, are required for bone matrix remodelling [23] and expression is reduced in TREM2 knockout mice [40]. Siglec (CD33-related family) expressed in myeloid derived cells also signal via DAP12, suggesting a common signalling pathway [47]. A recent systems approach has championed microglia...
phagocytosis as a prominent functional group, where DAP12 is a key regulator [48]. Consequently, there may be a mechanistic overlap between TREM2/DAP12 and some of the GWAS derived genes. The possibility of a shared mechanism between the bone and neurological manifestations of Alzheimer’s disease has been raised [49].

The future

The heritable component of sAD which remains elusive may conform to a polygenic model (many more alleles of weak genetic effects), and will require larger cohorts to detect vanishingly smaller genetic effects. Polygenic risk scoring in other complex diseases suggests this is indeed the case [50]. Alternatively fewer, rarer variants conferring larger genetic effects, like TREM2, may be responsible [2]. While weak genetic effects (polygenic model) and rare allele (rare variants model) are criticized for their individually small contributions to risk, the insight they offer into disease biology is invaluable.

The time when GWAS and resequencing approaches occupy their own niche is coming to an end; GWAS as a tool to explore common variants (MAF > 5%) in sufficiently large numbers to detect weak genetic effects (OR > 1.3), and NGS as a tool to detect rare variants in modest sample numbers due to cost implications. As custom array technology begins to include rarer variants, discovered through 1000 Genomes and/or other resequencing initiatives, and whole-genome sequencing becomes cost-effective these boundaries will blur. With this will come a greater appreciation of how common and rare alleles summate to influence a phenotype at any given disease loci.

References


17 Ware MD, Rosten P, Dumen JE, Liu L, Humphries RK, Krystal G. Cloning and characterization of human SHIP, the 145-kD inositol 5-phosphatase that associates with SHC after cytokine stimulation. Blood 1996; 88: 2833–40


20 Zweier M, Gregor A, Zweier C, Engels H, Sticht H, Grasshoff U, Johnson DS, Robertson L, Firth H V, Ekici AB, Reis A, Rauch A. Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MEF2P and CDKL5 expression. Hum Mutat 2010; 31: 722–33


32 Wray NR, Purcell SM, Visscher PM. Synthetic associations created by rare variants do not explain most GWAS results. PLoS Biol 2011; 9: e1000579


46 Peng Q, Malhotra S, Torchia JA, Kerr WG, Coggeshall KM, Humphrey MB. TREM2- and DAP12-dependent activation of PI3K requires DAP10 and is inhibited by SHIP1. *Sci Signal* 2010; 3: ra38


Received 29 October 2013
Accepted after revision 21 November 2013
Published online Article Accepted on 3 December 2013