BMC Medical Genomics

RESEARCH

Open Access

Rare variants in the splicing regulatory elements of EXOC3L4 are associated with brain glucose metabolism in Alzheimer's disease

Jason E. Miller[1,5], Manu K. Shivakumar[1], Younghee Lee[2], Seonggyun Han[2], Emrin Horgousluoglu[3], Shannon L. Risacher[3], Andrew J. Saykin[3], Kwangsik Nho[3*], Dokyoon Kim[1,4*] and for the Alzheimer's Disease Neuroimaging Initiative

Abstract

Background: Alzheimer's disease (AD) is one of the most common neurodegenerative diseases that causes problems related to brain function. To some extent it is understood on a molecular level how AD arises, however there are a lack of biomarkers that can be used for early diagnosis. Two popular methods to identify AD-related biomarkers use genetics and neuroimaging. Genes and neuroimaging phenotypes have provided some insights as to the potential for AD biomarkers. While the field of imaging-genomics has identified genetic features associated with structural and functional neuroimaging phenotypes, it remains unclear how variants that affect splicing could be important for understanding the genetic etiology of AD.

Methods: In this study, rare variants (minor allele frequency < 0.01) in splicing regulatory element (SRE) loci from whole genome sequencing (WGS) in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort, were used to identify genes that are associated with global brain cortical glucose metabolism in AD measured by FDG PET-scans. Gene-based associated analyses of rare variants were performed using the program BioBin and the optimal Sequence Kernel Association Test (SKAT-O).

Results: The gene, EXOC3L4, was identified as significantly associated with global cortical glucose metabolism (FDR (false discovery rate) corrected *p* < 0.05) using SRE coding variants only. Three loci that may affect splicing within EXOC3L4 contribute to the association.

Conclusion: Based on sequence homology, EXOC3L4 is likely a part of the exocyst complex. Our results suggest the possibility that variants which affect proper splicing of EXOC3L4 via SREs may impact vesicle transport, giving rise to AD related phenotypes. Overall, by utilizing WGS and functional neuroimaging we have identified a gene significantly associated with an AD related endophenotype, potentially through a mechanism that involves splicing.

Keywords: Alternative splicing, Imaging genomics, Alzheimer's disease, Whole genome sequencing, Rare variants

* Correspondence: knho@iupui.edu; dkim@geisinger.edu
[3]Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN, USA
[1]Biomedical and Translational Informatics Institute, Geisinger Health System, Danville, PA, USA
Full list of author information is available at the end of the article

Miller *et al. BMC Medical Genomics* 2018, **11**(Suppl 3):76
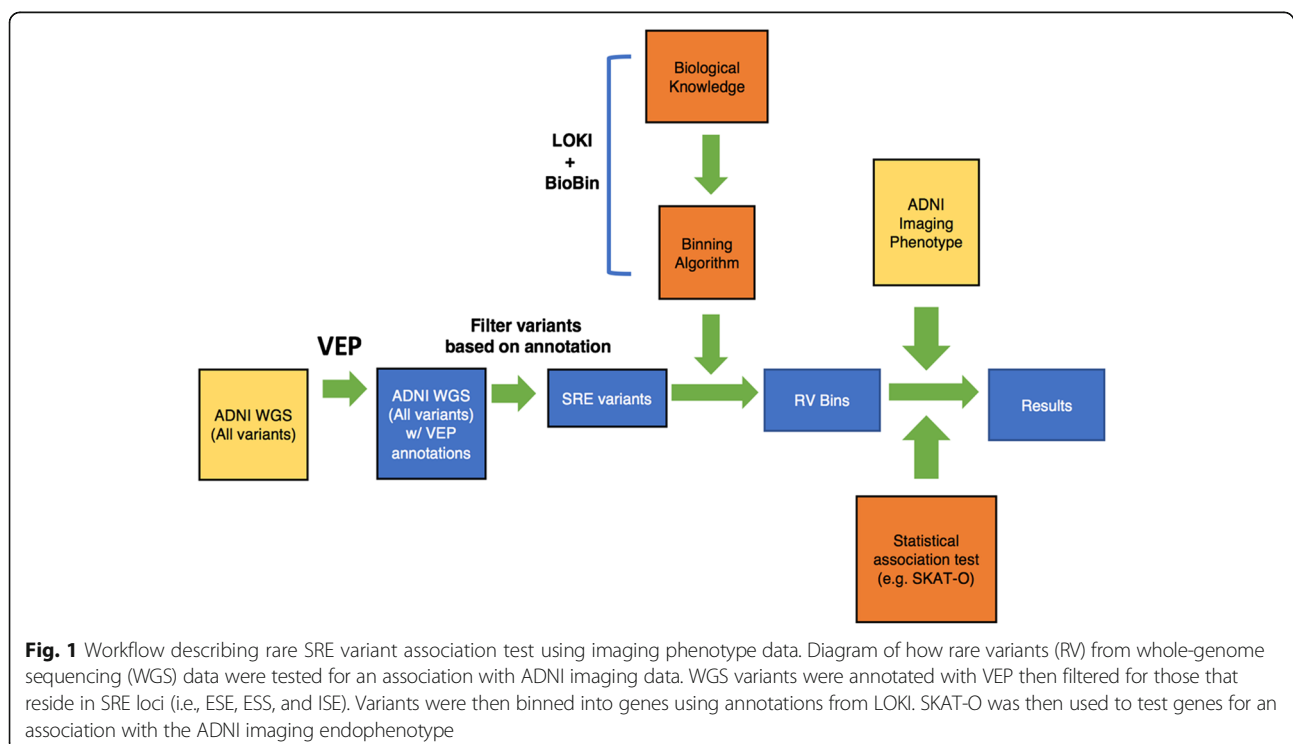
Page 46 of 71

Background

Late-onset Alzheimer's disease (LOAD) is a progressive common neurodegenerative disorder that causes problems with memory, thinking, and behavior and pathologically characterized by the presence of amyloid deposition and neurofibrillary tangles in the brain [1, 2]. 5.5 million Americans are estimated to have AD in 2017 and the number of Americans with AD is rapidly increasing because of the growing number of older adults [1]. Currently, there is no available cure for AD. As a result, without earlier diagnosis and early disease-modifying intervention, the total number of individuals with AD is predicted to quadruple by 2050, causing a great economic and social burden [1]. Furthermore, a biomarker for early diagnosis could benefit clinical trials for AD by precisely classifying prognosis, stage, and determining a clinical endpoint [3]. Thus, AD related research is increasingly important, especially as it relates to early diagnosis.

Genetic variation may play an essential role in AD pathogenesis [4]. Recently, a large-scale genome-wide association study (GWAS) identified and validated more than 22 susceptibility genes for LOAD [5]. After the success of GWAS for common SNPs, large-scale whole exome and genome sequencing studies have successfully identified several rare risk variants for LOAD [6–9]. Recently, the genetics of AD has been investigated in the context of imaging data. By combining information from genetic architecture, functional neuroimaging, and multi-omics data, genetic variation associated with AD-related imaging biomarkers

have been identified and thus the potential influence of genetic variation on brain structure and function related to AD pathophysiology [10, 11]. Furthermore, imaging endophenotypes can substantially increase statistical detection power of genetic association analysis through the use of quantitative traits as phenotypes [12].

Rare and low-frequency variants play an important role in the heritability of disease. However, the spurious nature of rare variants makes them difficult to run an association test. With so few occurrences of the variant at a given loci most tests will be underpowered [13]. In order to overcome this problem, variants can be grouped together by prior biological knowledge, such as genes, conserved loci, and pathways [14–16]. This strategy will accumulate effects of rare variants within a knowledge-driven region and reduce the number of statistical tests, thereby increasing the power to detect an association. Additionally, focusing on specific types of variants, such as those that lead to non-synonymous changes can also reduce the multiple testing burden and provide a potential explanation for the gene associated with the phenotype [17]. An attractive category of variant for studying AD are those that impact splicing.

Alternative splicing (AS) is an important gene regulatory mechanism underlying neurological function and development [18]. While motifs along splice junctions have been well studies, the effect of genetic variants in splicing regulatory elements (SREs) is less understood in the context of AD. There are multiple types of SREs that

Fig. 1 Workflow describing rare SRE variant association test using imaging phenotype data. Diagram of how rare variants (RV) from whole-genome sequencing (WGS) data were tested for an association with ADNI imaging data. WGS variants were annotated with VEP then filtered for those that reside in SRE loci (i.e., ESE, ESS, and ISE). Variants were then binned into genes using annotations from LOKI. SKAT-O was then used to test genes for an association with the ADNI imaging endophenotype

Miller *et al. BMC Medical Genomics* 2018, **11**(Suppl 3):76

Page 47 of 71

Table 1 Summary statistics of variables used as covariates in association study

Demographics and Covariates	Values ($N = 695$)
Sex (M/F)	391/304
Age in years (Mean/Std)	72.95 (+/− 7.05)

can impact splicing in different ways including intronic splicing enhancers (ISE), intronic splicing silencers (ISS), exonic splicing enhancers (ESE), and exonic splicing silencers (ESS) [19]. However, the significance of SRE in relation to AD related phenotypes remains unknown. In this study, an imaging genomics approach was taken to investigate AD by identifying rare variants within SREs associated with a neuroimaging phenotype using ADNI data. The Alzheimer's Disease Neuroimaging Initiative (ADNI) has provided publically available whole-genome sequencing (WGS) data, along with imaging phenotypes. BioBin, an open-source program that was developed to group or bin the variants using information from multiple databases, was employed to bin rare variants from ADNI WGS data [14]. Gene-based analyses were performed using the optimal Sequence Kernel Association Test (SKAT-O), which maximizes power by adaptively using the data to optimally combine the burden test and dispersion test.

Methods
Study sample
All whole-genome sequencing (WGS) and imaging data came from the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort. The cohort used here was made up of participants with cognitive normal (CN), early mild cognitive impairment (EMCI), late MCI (LMCI), and AD. The demographic data, along with sequencing and imaging data were downloaded from the ADNI data repository (http://www.loni.usc.edu/ADNI/). All participants provided written informed consent and study protocols were approved by participating sites' Institutional Review Board. WGS was performed using blood-derived genomic DNA samples from ADNI participants. Sequencing was performed using 100 bp paired-end reads on the Illumina HiSeq2000 platform (www.illumina.com). As previously described using Broad GATK and BWA-mem, reads were

mapped and aligned to the human genome (build 37), then variants were called [8, 20].

Neuroimaging analysis
Pre-processed [18F] FDG PET scans were downloaded from the LONI (http://loni.usc.edu). As previously described in detail, these FDG scans (Co-registered, Averaged, Standardized Image and Voxel Size, Uniform Resolution) were already averaged, aligned to a standard space, re-sampled to a standard image and voxel size (2 × 2 × 2 mm), smoothed to a uniform resolution, and intensity normalized [21]. The pre-processed images were aligned to each individual's MRI scan at the same visit and normalized to MNI space using SPM8 as previously described [22]. The intensity of the resulting scans was re-scaled to a pons reference region and then the final [18F] FDG standardized uptake value ratio (SUVR) images were created. A global cortical glucose metabolism measured from [18F] FDG-PET scans was used as an AD-related quantitative endophenotype with age at baseline and sex as covariates.

Variant annotation
The VCF file containing 695 non-Hispanic Caucasian participants with imaging phenotype, covariates and genomic data was annotated using the variant effect predictor (VEP) package [23]. The variant_effect_predictor.pl script was applied to the VCF file using cache, refseq, and pick flags. Variants were then selected if they were also annotated with an SRE element. Sequences and organization of SREs in humans were identified using previously developed method [19, 24] which required the use of dbSNP version 137 and hg19 reference genome [25, 26]. In brief, this method predicts hexamer motifs associated with the following types of SREs associated with exon skipping events: intronic splicing enhancer (ISE), exonic splicing enhancer (ESE), and exonic splicing silencer (ESS). The program *twoBitToFa* was used to find the genome sequences surrounding the SNPs of interest with hg19 reference [27]. 11-mer sequences were interrogated surrounding the SNPs (5 bp on each side) using the hexamer motifs. While ESE and ESS were coding SNPs, the ISE were defined as intronic SNPs. Using the published methods [19], an SRE was counted if there was a match between a hexamer motif and any part of the

Table 2 Top 5 genes associated with imaging phenotype using ISE variants only

Gene	Unique Loci	Variants across cohort	SKAT-O *p*-value	FDR corrected *p*-value
TNFAIP2	7	10	2.47E-05	0.123
STK35	56	148	3.09E-05	0.123
PWRN1	34	101	4.15E-05	0.123
EXOC3L4	8	21	6.11E-05	0.123
TMEM182	23	67	6.22E-05	0.123

Miller *et al. BMC Medical Genomics* 2018, **11**(Suppl 3):76

Page 48 of 71

Table 3 EXOC3L4 gene is associated with imaging phenotype using ESE/ESS variants only

| Gene	Unique Loci	Variants across cohort	SKAT-O *p*-value	FDR corrected *p*-value
EXOC3L4	4	16	7.48E-06	0.038

11-mer, and associated with exon skipping. ISE SNPs were included based on the exons bordering the intron in which the SNP was located.

Variant binning, association test, and analysis

BioBin was used to group rare variants by genic region (minor allele frequency (MAF) < 0.01). BioBin uses gene annotations from LOKI (the library of knowledge integration), which contains a number of widely used publically available databases such as NCBI Entrez, UCSC Genome Browser, Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, Genome Ontology (GO) and others. Association tests were performed using SKAT-O [28], adjusting for age and sex. The minimum bin size included in the association test was five variants across samples. The bins were tested for an association with global brain cortical glucose metabolism measured by FDG PET scans (often referred to as the "imaging phenotype"). For both annotation and BioBin, the GRCh37 assembly was specified. Finally, the false discovery rate for BioBin output *p*-values was then calculated in R using the p.adjust function, using the "FDR" method. Variant effect analysis using PROVEAN and SIFT was performed online at http://provean.jcvi.org [29–31]. The UCSC genome browser was used to visualize the *ECO3L4* gene along with splicing isoforms and protein domains [25].

Results

First, variants from the ADNI WGS study were selected that were located in the SRE coding and/or ISE loci (Fig. 1). Next BioBin was employed to bin variants with minor allele frequency (MAF) less than 0.01 into their respective genes. SKAT-O was used to test if these rare variants in each gene were associated with the phenotype derived from FDG PET scans from ADNI. These associations were adjusted for covariates such as age and sex to reduce the effect of confounding variables (Table 1). The *p*-values were adjusted for multiple testing, and genes with a false discovery rate (FDR) less than 0.05 were considered significant, while those with FDR < 0.1 were suggestive of statistical significance. Using ISE variants, there were no genes that had a significant association with the imaging phenotype (Table 2). However, using SRE coding variants (i.e., ESE and ESS), *EXOC3L4* was identified as having a genome-wide significant (FDR < 5%) association with the imaging phenotype (Table 3 and Fig. 2). PROVEAN and SIFT predictions for variants in *EXOC3L4* were considered neutral and tolerated, respectively, which suggests SRE annotations offer novel insight into the consequence of rare variants. Also, the sequence homology between *EXOC3L4* and exocyst complex components suggests SRE sites are within the Sec6 domain, indicating these splicing elements could have a functional impact on the protein (uniprot.org). Alternatively, when combining both sets of variants there were several genes which were only
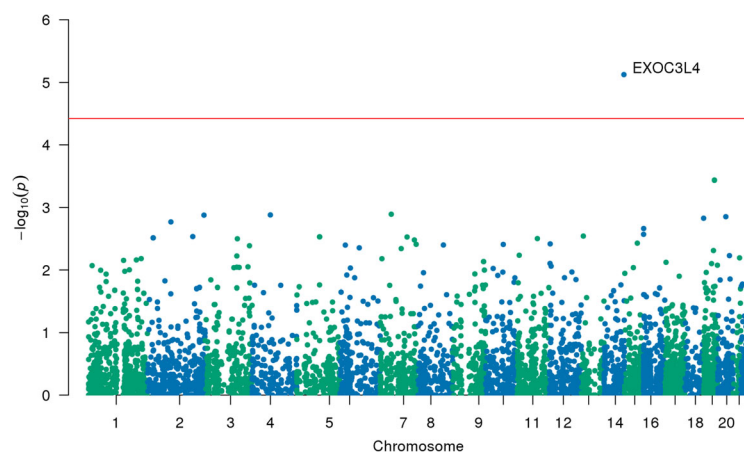
Fig. 2 Manhattan plot of *p*-values from association between genes and the imaging phenotype using SRE coding variants. Manhattan plot which shows the results from the association test between the imaging phenotype and each gene tested using SKAT-O. Only variants that fell into SRE coding loci were used. The blue and red lines represent 0.05 *p*-value and 0.05 FDR cutoffs, respectively

Miller *et al. BMC Medical Genomics* 2018, **11**(Suppl 3):76

Page 49 of 71

Table 4 Top 5 genes associated with imaging phenotype using ESE/ESS and ISE variants

Gene	Unique Loci	Variants across cohort	SKAT-O *p*-value	FDR corrected *p*-value
TNFAIP2	8	14	1.20E-05	0.094
EXOC3L4	10	35	1.99E-05	0.094
STK35	56	148	3.09E-05	0.094
STEAP4	7	13	3.23E-05	0.094
PWRN1	34	101	4.15E-05	0.097

suggestive of having a significant (FDR < 10%) association with the imaging phenotype (Table 4). In summary, SRE coding annotations provided increased power to detect *EXOC3L4* as having an association with the imaging phenotype.

To investigate the association with *EXOC3L4* further, each of the four loci in *EXOC3L4* with rare variants were interrogated to define their contribution to the association. This analysis was performed by removing each SNP individually then rerunning the association test to retrieve a *p*-value for only *EXOC3L4* (Table 5). After removing SNPs rs10142287, rs9324055, or rs148718670, *EXOC3L4* had a *p*-value that was less significant compared to the original *p*-value of *EXOC3L4* with 4 variants. These effects are unlikely to be caused purely by the number of variants at each locus removed, as only one or three variants were removed, suggesting there is something specific about the loci which leads to the association with the phenotype. As shown in the UCSC genome browser (Fig. 3), there is evidence that alternative splicing of *EXOC3L4* can lead to the existence of a transcript that skips the second exon which harbors two SNPs within ESE sites, rs10142287 and rs9324055. The skipped exon is part of a region encoding the Sec6 domain (Fig. 3). These results help explain why variability in SRE sites of *EXOC3L4* could impact the phenotype through a mechanism involving AS. On the other hand, removing the rs117708804 SNP resulted in an increase in significance as illustrated by the lower *p*-value, suggesting that *EXOC3L4* can absorb variation at this locus, and that variants here may be spurious or not important for the context of this association.

Discussion

Although very little is known about *EXOC3L4* or its orthologues, BLAST search results using its amino acid sequence suggests it is likely to be an exocyst complex component (uniprot.org). This information lends itself to a number of possible models for how *EXOC3L4* may be involved in AD. In mammals, the exocyst complex is an eight-subunit complex that is ubiquitously expressed [32]. The exocyst proteins are important for vesicle trafficking along with SNARE proteins [33, 34]. It has been suggested that SNARE proteins are important for glucose uptake in the context of proper neuronal function [35]. And the imaging phenotype from this study, FDG PET-scans, quantifies global brain cortical glucose metabolism in AD. Additionally, vesicle transport is used for lysosomal transport, such as seen in autophagy [36]. AD is defined by the accumulation of proteins like amyloid plaques, which can be removed via the autophagy-lysosome pathway [37]. There is evidence that defects in this autophagy process can lead to AD [38]. Thus, our results suggest a model where variants in *EXOC3L4* that are located in SRE coding loci may alter the function of the protein through exon skipping, which may inhibit proper vesicle transport. Moreover, these variants may lead to AD related phenotypes.
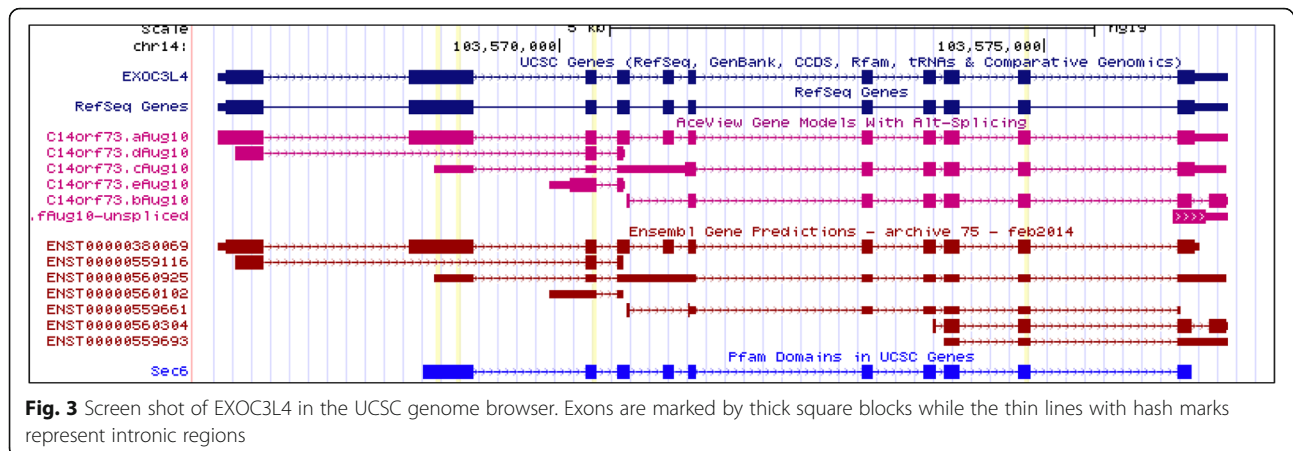
Evidence suggests the exocyst plays important roles in embryogenesis, neuronal cell polarity, and cell motility [32]. *EXOC3L4* shares high sequence similarity with M-seq (also known as *TNFaip2*), a protein that shares structural similarity to *Sec6* (uniprot.org). It has been suggested that *TNFaip2* has a role in filopedia development in neurons [32]. Thus, if *EXOC3L4* does not carry out its function through interactions with the exocyst

Table 5 Characterization of EXOC3L4 rare variant loci

rsID	Consequence	*p*-value[a]	SRE type	Variants across cohort
rs117708804	missense	4.32E-07	ESE, ESS	11
EXOC3L4		7.48E-06[b]		16
rs10142287	synonymous	1.58E-04	ESE	1
rs9324055	missense	1.59E-04	ESE	1
rs148718670	missense	1.68E-04	ESE	3

[a]SKAT-O p-value results after removing the variant from EXOC3L4
[b]SKAT-O *p*-value result using all variants from EXOC3L4

Miller *et al. BMC Medical Genomics* 2018, **11**(Suppl 3):76

Page 50 of 71

Fig. 3 Screen shot of EXOC3L4 in the UCSC genome browser. Exons are marked by thick square blocks while the thin lines with hash marks represent intronic regions

complex, there is evidence that exocyst-like proteins are important for neuronal cell function through an alternative mechanism.

One limitation of this study is the samples size. *TNFAIP2* was only suggestive of statistical significance, however it is another protein that is functionally relevant to SNARE proteins and thus vesicle transport (uniprot.org). Since only few samples contained the rare variants, it will be important for this study to be replicated in another independent cohort. Additionally, since these are associations it will be important to perform follow-up experiments to identify a causal link between *EXOC3L4* and AD. None the less, these results suggest more genes that contain rare variants in SRE loci, and are important for proper exocytosis and autophagy, are likely to be identified in studies with increased sample size. In summary, annotating variants as SRE provides novel insight as to how rare variants may be useful when finding an association between an imaging phenotype and genetic variants in the context of AD.

Conclusions

In this study, we set out to find associations between a neuroimaging phenotype and rare SRE variants from WGS in AD. While it is common to perform a gene-wise association test, we hypothesized that the power of this study could be increased by focusing on functionally relevant loci such as those that impact splicing. Furthermore, associations with annotated regions can lead to easier interpretation afterward. Thus, rare variants that fell into SRE loci from WGS from the ADNI cohort were collapsed into genes using BioBin. These rare variants were used in an association test for an imaging phenotype and *EXOC3L4* was identified as having a statistically significant association. And while intronic elements did not detect a statistically significant association, exonic splicing elements did. In summary, utilizing prior biological

knowledge in the form of splicing elements serves as an important means to identify genotype-phenotype relationships with respect to imaging data and AD.

Miller *et al. BMC Medical Genomics* 2018, **11**(Suppl 3):76

Page 51 of 71

Availability of data and materials

Demographic information, raw neuroimaging scan data, whole genome sequencing data, neuropsychological tests scores, and diagnostic information are available from the ADNI data repository at http://www.loni.usc.edu/ADNI/.

About this supplement

This article has been published as part of *BMC Medical Genomics* Volume 11 Supplement 3, 2018: Selected articles from the 7th Translational Bioinformatics Conference (TBC 2017): medical genomics. The full contents of the supplement are available online at https://bmcmedgenomics.biomedcentral.com/articles/supplements/volume-11-supplement-3.

Authors' contributions

All authors contributed substantively to this work. JEM, YL, KN, and DK were involved in study conception and design. JEM, MKS, YL, KN, EH, DK, SH, and AJS were involved in data organization, whole genome sequencing analysis and statistical analyses. AJS was involved in coordination and data collection and processing for ADNI. JEM, MKS, KN and DK drafted the report and prepared all figures and tables. All authors were involved in reviewing and editing of the manuscript and approved it.

Ethics approval and consent to participate

At the time of enrollment, a written informed consent was acquired for imaging and genetic sample collection along with protocols of consent forms were approved by each participating sites' Institutional Review Board (IRB).

Consent for publication

Not applicable.

Competing interests

The authors declare they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

[1]Biomedical and Translational Informatics Institute, Geisinger Health System, Danville, PA, USA. [2]Department of Biomedical Informatics, University of Utah School of Medicine, Salt Lake City, UT 84106, USA. [3]Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN, USA. [4]Huck Institute of the Life Sciences, Pennsylvania State University, University Park, PA, USA. [5]Present Address: Department of Genetics, Institute for Biomedical Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

Published: 14 September 2018

References

1. Alzheimer's A. 2015 Alzheimer's disease facts and figures. Alzheimers Dement. 2015;11(3):332–84.
2. Bloom GS. Amyloid-beta and tau: the trigger and bullet in Alzheimer disease pathogenesis. JAMA Neurol. 2014;71(4):505–8.
3. Beach TG. A review of biomarkers for neurodegenerative disease: will they swing us across the valley? Neurol Ther. 2017;6(Suppl 1):5–13.
4. Tanzi RE. The genetics of Alzheimer disease. Cold Spring Harb Perspect Med. 2012;2(10)
5. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45(12):1452–8.
6. Cruchaga C, Karch CM, Jin SC, Benitez BA, Cai Y, Guerreiro R, Harari O, Norton J, Budde J, Bertelsen S, et al. Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. Nature. 2014;505(7484):550–4.
7. Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, Cruchaga C, Sassi C, Kauwe JS, Younkin S, et al. TREM2 variants in Alzheimer's disease. N Engl J Med. 2013;368(2):117–27.
8. Nho K, Horgusluoglu E, Kim S, Risacher SL, Kim D, Foroud T, Aisen PS, Petersen RC, Jack CR Jr, Shaw LM, et al. Integration of bioinformatics and imaging informatics for identifying rare PSEN1 variants in Alzheimer's disease. BMC Med Genomics. 2016;9(Suppl 1):30.
9. Steinberg S, Stefansson H, Jonsson T, Johannsdottir H, Ingason A, Helgason H, Sulem P, Magnusson OT, Gudjonsson SA, Unnsteinsdottir U, et al. Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. Nat Genet. 2015;47(5):445–7.
10. Nho K, Kim S, Horgusluoglu E, Risacher SL, Shen L, Kim D, Lee S, Foroud T, Shaw LM, Trojanowski JQ, et al. Association analysis of rare variants near the APOE region with CSF and neuroimaging biomarkers of Alzheimer's disease. BMC Med Genet. 2017;10(Suppl 1):29.
11. Saykin AJ, Shen L, Yao X, Kim S, Nho K, Risacher SL, Ramanan VK, Foroud TM, Faber KM, Sarwar N, et al. Genetic studies of quantitative MCI and AD phenotypes in ADNI: progress, opportunities, and plans. Alzheimers Dement. 2015;11(7):792–814.
12. Shen L, Thompson PM, Potkin SG, Bertram L, Farrer LA, Foroud TM, Green RC, Hu X, Huentelman MJ, Kim S, et al. Genetic analysis of quantitative phenotypes in AD and MCI: imaging, cognition and biomarkers. Brain Imaging Behav. 2014;8(2):183–207.
13. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. Am J Hum Genet. 2014;95(1):5–23.
14. Moore CB, Wallace JR, Frase AT, Pendergrass SA, Ritchie MD. BioBin: a bioinformatics tool for automating the binning of rare variants using publicly available biological knowledge. BMC Med Genomics. 2013;6(Suppl 2):S6.
15. Moore CB, Wallace JR, Frase AT, Pendergrass SA, Ritchie MD. Using BioBin to explore rare variant population stratification. Pac Symp Biocomput. 2013:332–43.
16. Basile AO, Wallace JR, Peissig P, McCarty CA, Brilliant M, Ritchie MD. Knowledge driven binning and Phewas analysis in Marshfield personalized medicine research project using Biobin. Pac Symp Biocomput. 2016;21:249–60.
17. Kim D, Basile AO, Bang L, Horgusluoglu E, Lee S, Ritchie MD, Saykin AJ, Nho K. Knowledge-driven binning approach for rare variant association analysis: application to neuroimaging biomarkers in Alzheimer's disease. BMC Med Inform Decis Mak. 2017;17(Suppl 1):61.
18. Vuong CK, Black DL, Zheng S. The neurogenetics of alternative splicing. Nat Rev Neurosci. 2016;17(5):265–81.
19. Lee Y, Gamazon ER, Rebman E, Lee Y, Lee S, Dolan ME, Cox NJ, Lussier YA. Variants affecting exon skipping contribute to complex traits. PLoS Genet. 2012;8(10):e1002998.
20. Nho K, West JD, Li H, Henschel R, Bharthur A, Tavares MC, Saykin AJ. Comparison of multi-sample variant calling methods for whole genome sequencing. IEEE Int Conf Systems Biol. 2014;2014:59–62.
21. Jagust WJ, Bandy D, Chen K, Foster NL, Landau SM, Mathis CA, Price JC, Reiman EM, Skovronsky D, Koeppe RA, et al. The Alzheimer's disease neuroimaging initiative positron emission tomography core. Alzheimers Dement. 2010;6(3):221–9.
22. Risacher SL, Kim S, Nho K, Foroud T, Shen L, Petersen RC, Jack CR Jr, Beckett LA, Aisen PS, Koeppe RA, et al. APOE effect on Alzheimer's disease biomarkers in older adults with significant memory concern. Alzheimers Dement. 2015;11(12):1417–29.
23. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP effect predictor. Bioinformatics. 2010;26(16):2069–70.
24. Gamazon ER, Konkashbaev A, Derks EM, Cox NJ, Lee Y. Evidence of selection on splicing-associated loci in human populations and relevance to disease loci mapping. Sci Rep. 2017;7(1):5980.
25. Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, Goldman M, Barber GP, Clawson H, Coelho A, et al. The UCSC genome browser database: update. Nucleic Acids Res. 2011;39(Database issue):D876–82.
26. Yeo G, Hoon S, Venkatesh B, Burge CB. Variation in sequence and organization of splicing regulatory elements in vertebrate genes. Proc Natl Acad Sci U S A. 2004;101(44):15700–5.
27. Kent WJ. BLAT--the BLAST-like alignment tool. Genome Res. 2002;12(4):656–64.
28. Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, Team NGESP-ELP, Christiani DC, Wurfel MM, Lin X. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. Am J Hum Genet. 2012;91(2):224–37.
29. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. PLoS One. 2012;7(10):e46688.
30. Choi Y. A fast computation of pairwise sequence alignment scores between a protein and a set of single-locus variants of another protein. In: Proceedings of the ACM conference on bioinformatics, computational biology and biomedicine. Orlando, Florida: ACM; 2012. p. 414–7.
31. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res. 2001;11(5):863–74.

Miller *et al. BMC Medical Genomics* 2018, **11**(Suppl 3):76

Page 52 of 71

32. Martin-Urdiroz M, Deeks MJ, Horton CG, Dawe HR, Jourdain I. The exocyst complex in health and disease. Front Cell Dev Biol. 2016;4:24.
33. Wu B, Guo W. The exocyst at a glance. J Cell Sci. 2015;128(16):2957–64.
34. He B, Guo W. The exocyst complex in polarized exocytosis. Curr Opin Cell Biol. 2009;21(4):537–42.
35. Park SJ, Jung YJ, Kim YA, Lee-Kang JH, Lee KE. Glucose/oxygen deprivation and reperfusion upregulate SNAREs and complexin in organotypic hippocampal slice cultures. Neuropathology. 2008;28(6):612–20.
36. Moreau K, Renna M, Rubinsztein DC. Connections between SNAREs and autophagy. Trends Biochem Sci. 2013;38(2):57–63.
37. Zare-Shahabadi A, Masliah E, Johnson GV, Rezaei N. Autophagy in Alzheimer's disease. Rev Neurosci. 2015;26(4):385–95.
38. Nixon RA. The role of autophagy in neurodegenerative disease. Nat Med. 2013;19(8):983–97.