


RESEARCH

Open Access



# Multivariate genome wide association and network analysis of subcortical imaging phenotypes in Alzheimer's disease

Xianglian Meng<sup>1</sup>, Jin Li<sup>2</sup>, Qiushi Zhang<sup>3</sup>, Feng Chen<sup>2</sup>, Chenyuan Bian<sup>2</sup>, Xiaohui Yao<sup>4</sup>, Jingwen Yan<sup>5,6</sup>, Zhe Xu<sup>1</sup>, Shannon L. Risacher<sup>5</sup>, Andrew J. Saykin<sup>5</sup>, Hong Liang<sup>2\*</sup> , Li Shen<sup>4\*</sup> and for the Alzheimer's Disease Neuroimaging Initiative

From The International Conference on Intelligent Biology and Medicine (ICIBM) 2020 Virtual. 9-10 August 2020

## Abstract

**Background:** Genome-wide association studies (GWAS) have identified many individual genes associated with brain imaging quantitative traits (QTs) in Alzheimer's disease (AD). However single marker level association discovery may not be able to address the underlying biological interactions with disease mechanism.

**Results:** In this paper, we used the MGAS (Multivariate Gene-based Association test by extended Simes procedure) tool to perform multivariate GWAS on eight AD-relevant subcortical imaging measures. We conducted multiple iPINBPA (integrative Protein-Interaction-Network-Based Pathway Analysis) network analyses on MGAS findings using protein-protein interaction (PPI) data, and identified five Consensus Modules (CMs) from the PPI network. Functional annotation and network analysis were performed on the identified CMs. The MGAS yielded significant hits within *APOE*, *TOMM40* and *APOC1* genes, which were known AD risk factors, as well as a few new genes such as *LAMA1*, *XYLB*, *HSD17B7P2*, and *NPEPL1*. The identified five CMs were enriched by biological processes related to disorders such as Alzheimer's disease, Legionellosis, Pertussis, and Serotonergic synapse.

**Conclusions:** The statistical power of coupling MGAS with iPINBPA was higher than traditional GWAS method, and yielded new findings that were missed by GWAS. This study provides novel insights into the molecular mechanism of Alzheimer's Disease and will be of value to novel gene discovery and functional genomic studies.

**Keywords:** Brain imaging, Multivariate gene-based genome-wide analysis, iPINBPA network analysis, Consensus modules

\* Correspondence: [lh@hrbeu.edu.cn](mailto:lh@hrbeu.edu.cn); [lishen@penncmedicine.upenn.edu](mailto:lishen@penncmedicine.upenn.edu)

<sup>2</sup>College of Automation, Harbin Engineering University, Harbin 150001, China

<sup>4</sup>Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Alzheimer's disease (AD) is a debilitating and highly heritable disease with great complexity in its genetic contributors [1]. Genome-wide association studies (GWAS) of AD or AD biomarkers have been performed at the single-nucleotide polymorphism (SNP) level [2–4] as well as at the higher level (e.g., gene, pathway and/or network) [5–8]. It is widely recognized that AD has a complicated genetic mechanism involving multiple genes. Different combinations of functionally related variants in genes and pathways may interact to produce the phenotypic outcomes in AD, single SNP-level and gene-level GWAS results are unlikely to completely reveal the underlying genetic mechanism in AD. GWAS have greatly facilitated the identification of genetic markers (e.g., single nucleotide polymorphisms or SNPs) associated with brain imaging quantitative traits (QTs) in AD [9, 10]. As a complex disease, it is highly likely that AD is influenced by multiple genetic variants [11, 12]. The identified single-SNP-single-QT associations typically have small effect sizes. To bridge this gap, exploring single-SNP-multi-QT associations may have the potential to increase statistical power and identify meaningful imaging genetic associations. With this observation, we employ the MGAS (Multivariate Gene-based Association test by extended Simes procedure) tool [13] to perform multivariate GWAS on eight AD-relevant subcortical imaging measures.

In addition, biological interactions may be important in contributing to intermediate imaging QTs and overall disease outcomes [14]. Network-based analysis guided by biologically relevant connections from public databases provides a powerful tool for improved mechanistic understanding of complex disorders [15–18]. Considering that the etiology of AD might depend on functional protein-protein interaction (PPI) network, we conduct multiple iPINBPA (integrative protein-interaction-network-based pathway analysis) [19] network analyses on MGAS findings using the PPI data, and identify Consensus Modules (CMs) based on the iPINBPA discoveries. Functional annotation and network analysis are subsequently performed on the identified CMs.

In order to enhance the ability to recognize the aggregation effect of multiple SNPs, it may be desirable to perform association analysis at the SNP set (or gene) level rather than at a single SNP level. This paper aims to reveal the relationship between genetic markers and multiple phenotypes, improve statistical power, and find GWAS missing results by MGAS. Network analysis could provide meaningful biological relationships to help interpret GWAS data to further study the genetic mechanism of AD. A schematic framework of our analysis is shown in Fig. 1.

## Results

### Participant characteristics

The subjects ( $N = 866$ ) consisted of 467 males (53.9%) and 399 females (46.1%) aged 48–91 years. Shown in Table 1 are the demographic and clinical characteristics of these subjects stratified by five diagnostic groups. There is no significant difference on the *APOE* e4 status in the five diagnostic groups. Significant differences are observed in gender ( $p = 0.035$ ) and education ( $p = 0.037$ ). Age is significantly different across the five groups ( $p < 0.001$ ). Furthermore, eight neuroimaging phenotypes (L AmygVol, R AmygVol, L HippVol, R HippVol, L AccumVol, R AccumVol, L PutamVol, R PutamVol; see Table 2) show the significant difference across the five diagnostic groups ( $p < 0.001$ ). Shown in Fig. 2 is the correlation matrix of these eight phenotypes. The correlation between L HippVol and R HippVol ( $r = 0.83$ ) and that between L PutamVol and R PutamVol ( $r = 0.90$ ) are among the highest.

### Multivariate genome wide association study

In multivariate genome wide association study (MGAS) [13, 20], the top SNP hit is rs769449 from the *APOE* and *TOMM40* region ( $p = 1.19E-09$ ) (Table 3). According to the hypothesis that genes are the functional units in biology [15, 21], multivariate gene-level association  $p$ -values were also obtained by MGAS which combines  $p$ -value information in regressing univariate phenotypes on common SNPs. Figure 3 shows the Manhattan plot of the gene-based MGAS results. Using Bonferroni corrected  $p$ -value of 0.05 as the threshold, three genes (*APOE*, *TOMM40*, *APOC1*) were significantly associated with the studied eight subcortical measures. Table 4 shows that the top 10 gene-level findings identified by MGAS, where *APOE* ( $p = 2.77E-08$ ), *TOMM40* ( $p = 3.49E-08$ ), and *APOC1* ( $p = 2.09E-06$ ) are the well-known AD risk regions. *LAMA1* ( $p = 3.79E-05$ ) was reported to encode the laminin alpha subunit associated with late onset AD in the Amish [22]. *HSD17B7P2* ( $p = 8.40E-05$ ) was reported to play an important role in brain development [23]. The other five gene-level findings in top 10 are *XYLB*, *NPEPL1*, *CYP24A1*, *OR5B2* and *MIR7160*.

### Consensus modules

Consensus modules (CMs) were constructed based on our previous work [24]. To search for subnetworks in the multivariate GWAS finding, we ran iPINBPA ten times by varying the random seed value from 1 to 10. Table 5 shows the top 5 subnetworks identified in each run, including the Dice's coefficient value with the most similar modules in other runs. Compared with the standard iPINBPA method, our CM-based network strategy was designed to identify more reliable modules across multiple runs.



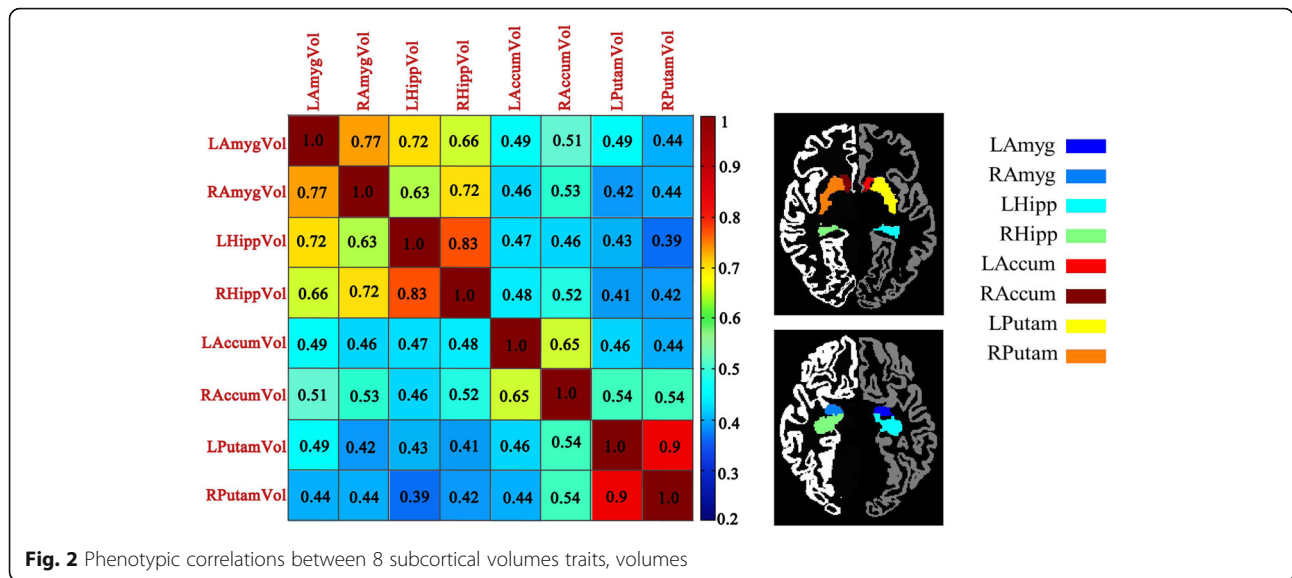
**Table 1** Demographic information and total number of participants involved in each analysis

	<b>CN (N = 183)</b>	<b>SMC (N = 95)</b>	<b>EMCI (N = 281)</b>	<b>LMCI (N = 177)</b>	<b>AD (N = 130)</b>	<b>p-values</b>
Age (years)	74.29(6.01)	72.20(5.67)	71.35(7.30)	71.87(7.98)	74.56(8.07)	$p < 0.001$
Gender(M/F)	92/91	39/56	159/122	99/78	78/52	0.035
Education (years)	15.78(2.69)	16.81(2.55)	16.07(2.66)	16.31(2.89)	15.78(2.69)	0.037
APOE e4 allele present	62(34%)	37(39%)	108 (38%)	64(36%)	45 (35%)	0.161
<b>LAmygVol (i.e., QT for GWAS)</b>	<b>1377.2 (230.40)</b>	<b>1434.48(202.03)</b>	<b>1387.06(257.16)</b>	<b>1258.07(286.96)</b>	<b>1081.93 (229.42)</b>	<b><math>p &lt; 0.001</math></b>
<b>RAmygVol (i.e., QT for GWAS)</b>	<b>1425.05 (221.11)</b>	<b>1494.02 (210.39)</b>	<b>1450.03 (255.77)</b>	<b>1327.93 (282.28)</b>	<b>1178.03 (240.54)</b>	<b><math>p &lt; 0.001</math></b>
<b>LHippVol (i.e., QT for GWAS)</b>	<b>3626.23 (497.63)</b>	<b>3730.98 (529.42)</b>	<b>3573.44 (558.02)</b>	<b>3243.94 (635.34)</b>	<b>2912.15 (518.08)</b>	<b><math>p &lt; 0.001</math></b>
<b>RHippVol (i.e., QT for GWAS)</b>	<b>3679.88 (500.54)</b>	<b>3824.23 (487.64)</b>	<b>3663.93 (535.43)</b>	<b>3319.32 (641.19)</b>	<b>2985.63 (540.11)</b>	<b><math>p &lt; 0.001</math></b>
<b>LAccumVol (i.e., QT for GWAS)</b>	<b>463.51 (100.88)</b>	<b>481.37 (92.49)</b>	<b>474.07 (93.96)</b>	<b>447.91 (101.37)</b>	<b>417.88 (96.74)</b>	<b><math>p &lt; 0.001</math></b>
<b>RAccumVol (i.e., QT for GWAS)</b>	<b>490.44 (95.69)</b>	<b>506.76 (93.38)</b>	<b>508.65 (107.70)</b>	<b>474.41 (105.41)</b>	<b>429.57 (96.56)</b>	<b><math>p &lt; 0.001</math></b>
LCaudVol	3442.57 (505.06)	3411.83 (518.23)	3477.88 (572.20)	3460.31 (518.58)	3429.95 (698.19)	0.0851
RCaudVol	3583.43 (528.08)	3539.56 (545.67)	3640.80 (623.54)	3588.41 (540.35)	3577.51 (697.23)	0.608
LPallVol	1602.48 (204.00)	1592.57 (196.97)	1633.63 (211.56)	1587.56 (220.94)	1584.27 (231.88)	0.101
RPallVol	1413.36 (189.35)	1434.79 (177.66)	1444.59 (192.61)	1433.18 (202.08)	1414.99 (217.48)	0.444
<b>LPutamVol (i.e., QT for GWAS)</b>	<b>4788.04 (636.56)</b>	<b>4788.21 (687.42)</b>	<b>4939.46 (750.31)</b>	<b>4733.41 (671.80)</b>	<b>4479.92 (688.84)</b>	<b><math>p &lt; 0.001</math></b>
<b>RPutamVol (i.e., QT for GWAS)</b>	<b>4586.97 (599.25)</b>	<b>4584.27 (604.70)</b>	<b>4708.87 (752.76)</b>	<b>4522.97 (669.07)</b>	<b>4327.07 (687.91)</b>	<b><math>p &lt; 0.001</math></b>
LThalVol	6060.0 (714.50)	6064.96 (868.69)	6072.55 (604.95)	6054.47 (745.24)	5953.05 (718.64)	0.592
RThalVol	6081.58 (684.84)	6056.69 (722.49)	6176.78 (663.50)	6128.47 (732.28)	5980.53 (750.51)	0.098

AD Alzheimer's disease, ADNI Alzheimer's Disease Neuroimaging Initiative, CDR-SOB clinical dementia rating-sum of boxes, CN cognitively normal, SMC significant memory concern, EMCI early mild cognitive impairment, LMCI late mild cognitive impairment  
Number (%) or mean (s.d.) is shown in each entry. P-values are computed using one-way ANOVA (\*except for gender using chi-square test)

**Table 2** 14 FreeSurfer subcortical ROIs

<b>Phenotype ID</b>	<b>Description</b>	<b>Region</b>
LAmygVol	the volume of the left amygdala	Subcortical (temporal)
RAmygVol	the volume of the right amygdala	
LHippVol	the volume of the left hippocampus	
RHippVol	the volume of the right hippocampus	
LAccumVol	the volume of the left accumbens	Subcortical (striatum/basal ganglia)
RAccumVol	the volume of the right accumbens	
LCaudVol	the volume of the left caudate	
RCaudVol	the volume of the right caudate	
LPallVol	the volume of the left pallidum	
RPallVol	the volume of the right pallidum	
LPutamVol	the volume of the left putamen	
RPutamVol	the volume of the right putamen	
LThalVol	the volume of the left thalamus	Subcortical (thalamus)
RThalVol	the volume of the right thalamus	



RAmygVol, LHippVol and RHippVol. *LAMA1* ( $p = 3.79E-05$ ,  $rs656734$ ) encodes one of the alpha 1 subunits of Laminin, which has been demonstrated to be expressed in the hippocampal neuronal cell layers [27]. *NPEPL1* was confirmed to be a potential direct target of miR-19a in a breast cancer study [28] and miR-19a was up-regulated in primary motor cortex and hippocampus in the brain of amyotrophic lateral sclerosis mice at late disease stage [29]. In our study, we found that *NPEPL1* was associated to LPutamVol ( $P_{LPutamVol} = 2.13E-05$ ) and RPutamVol ( $P_{RPutamVol} = 3.39E-03$ ). In Table 4, the hippocampus and amygdala volumes were associated with multiple genes. While *CYP24A1* was associated with none of eight studied QTs, it was identified in MGAS to have an overall association with all eight QTs. The statistical efficacy of MGAS of the detected gene associations appears to be more powerful than univariate phenotype models. Given that *OR5B2* and *MIR7160*

have not been reported to be related to AD or AD related biomarkers, it warrants further investigation to examine their roles on AD in independent cohorts.

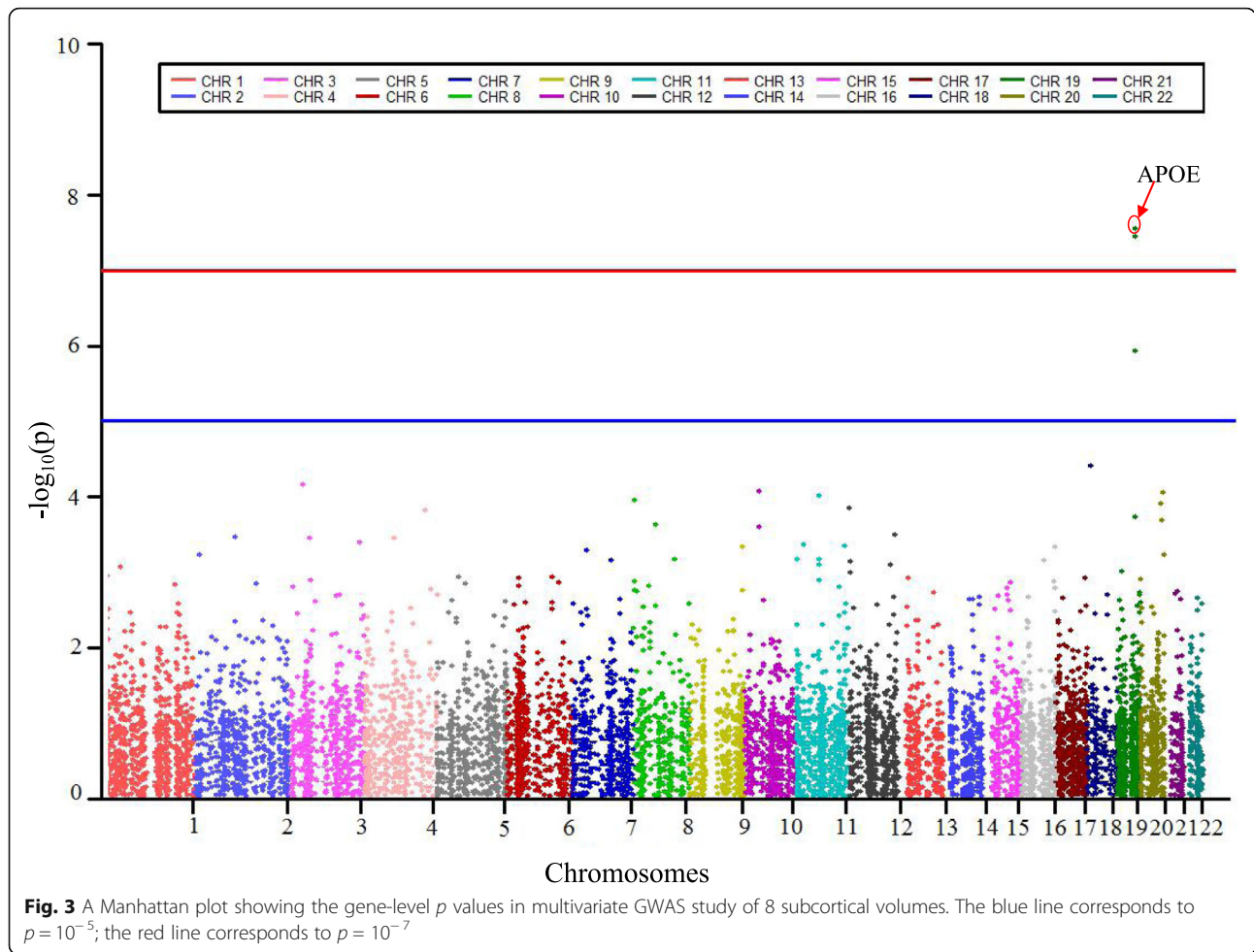
Because we found that different subnetworks could be identified by using different random seed values, we present the consensus modules discovered by an enhanced iPINBPA strategy. The genes for the CMs might not show a direct individual statistical significance but demonstrated a collected effect on the studied phenotypes. We assessed the significance of each identified consensus module (Table 6).  $CM_1$  ( $Score = 3.32$ ,  $p = 9.00E-04$ ) contains totally 8 genes, including KEGG AD genes *TNFRSF1A*.  $CM_2$  ( $Score = 1.41$ ,  $p = 0.16$ ) contains totally 4 genes without reaching the significance level.  $CM_3$  ( $Score = 4.37$ ,  $p = 1.24E-05$ ) contains 6 genes, including KEGG AD genes *APOE*, *APP*, and *CASP3*.  $CM_4$  ( $Score = 4.32$ ,  $p = 1.56E-05$ ) contains 5 genes, including KEGG AD genes *APOE*, *APP*, and *CASP3*.  $CM_5$  ( $Score = 1.89$ ,  $p = 5.88E-02$ ) contains 6 genes with a marginal significance. The genes in the significant CMs warrant further investigation. The consensus module strategy applied to the iPINBPA framework yielded more stable results than the standard iPINBPA.

The intersection of  $CM_3$  and  $CM_4$  yielded five genes, including *APP*, *APOE*, *CASP3*, *C3*, and *PLTP*. The *C3* gene was shown to contribute to the pathogenesis of demyelinating disease by directly or indirectly chemoattracting encephalitogenic cells to the CNS [30]. The *PLTP* gene was reported to play an important role in  $A\beta$  metabolism and it is an interesting topic to further elucidate functions of *PLTP* in AD susceptibility. Table 7 shows the top ten pathways enriched by the intersection genes. Among these genes, the *APP*, *APOE*, and *CASP3* genes are known AD risk factors. Several significant pathways were observed, including Alzheimer’s disease ( $p$ -

**Table 3** The top 10 SNPs identified by MGAS

SNP	Position	P <sub>SNP</sub>	Chr	Gene	P <sub>MGAS</sub>
rs769449	45,410,002	1.19E-09	19	APOE	2.77E-08
rs405509	45,408,836	2.01E-04	19	APOE	2.77E-08
rs439401	45,414,451	1.99E-01	19	APOE	2.77E-08
rs584007	45,416,478	2.25E-01	19	APOE	2.77E-08
rs445925	45,415,640	2.62E-01	19	APOE	2.77E-08
rs769449	45,410,002	1.19E-09	19	TOMM40	3.49E-08
rs2075650	45,395,619	2.19E-06	19	TOMM40	3.49E-08
rs157582	45,396,219	8.54E-05	19	TOMM40	3.49E-08
rs1160985	45,403,412	6.88E-06	19	TOMM40	3.49E-08
rs405509	45,408,836	2.01E-04	19	TOMM40	3.49E-08

P<sub>SNP</sub> is the p-value of top 10 snp identified by MGAS; Chr: Chromosome; P<sub>MGAS</sub> is the p-value of top 10 gene identified by MGAS



**Table 4** The top 10 FDR corrected genes identified by MGAS in 8 subcortical ROIs

NO.	Gene	$P_{MGAS}$	Chr	SNP	$P_{LAmgyVol}$	$P_{RAmgyVol}$	$P_{LHippVol}$	$P_{RHippVol}$	$P_{LAccumVol}$	$P_{RAccumVol}$	$P_{LPutamVol}$	$P_{RPutamVol}$
1	APOE	2.77E-08	19	rs769449	<b>3.98E-06</b>	1.38E-04	<b>5.77E-07</b>	<b>9.52E-09</b>	1.58E+00	1.48E+00	2.11E+00	2.44E+00
2	TOMM40	3.49E-08	19	rs769449	<b>3.98E-06</b>	1.38E-04	<b>5.77E-07</b>	<b>9.52E-09</b>	1.58E+00	1.48E+00	2.11E+00	2.44E+00
3	APOC1	1.18E-06	19	rs4420638	3.09E-05	3.46E-05	<b>7.70E-06</b>	<b>5.23E-07</b>	6.99E-01	1.34E+00	3.06E+00	2.71E+00
4	LAMA1	3.79E-05	18	rs656734	2.62E-05	<b>8.64E-07</b>	3.11E-04	5.53E-04	5.64E+00	2.45E+00	4.27E+00	5.26E+00
5	XYLB	6.72E-05	3	rs196376	1.80E-02	6.61E-04	8.64E-05	<b>6.63E-06</b>	2.06E-02	5.68E-01	9.22E-01	7.13E-02
6	HSD17B7P2	8.40E-05	10	rs12221164	3.42E+00	4.05E+00	5.92E+00	5.09E+00	1.16E-04	1.02E-01	7.86E+00	7.93E+00
7	NPEPL1	8.63E-05	20	rs2426778	5.96E-01	3.89E+00	7.38E-01	2.17E+00	7.32E+00	2.09E-02	2.13E-05	3.39E-03
8	OR5B2	9.40E-05	11	rs11229440	1.62E+00	8.20E-01	3.90E-01	1.05E+00	1.67E+00	8.30E-03	2.22E-04	9.52E-05
9	MIR7160	1.10E-04	8	rs6558595	8.22E-01	7.69E+00	3.07E-01	1.13E+00	8.50E-02	1.68E+00	3.49E-05	3.67E-04
10	CYP24A1	1.20E-04	20	rs3787555	2.10E-01	4.75E+00	3.64E+00	2.19E+00	7.80E+00	1.75E+00	1.71E-01	9.55E-01

$P_{MGAS}$  is the  $p$ -value of top 10 genes identified by MGAS; Chr: Chromosome;  $P_{LAmgyVol}$  is the  $p$ -value of top 10 genes associated to LAmgyVol;  $P_{RAmgyVol}$  is the  $p$ -value of top 10 genes associated to RAmgyVol;  $P_{LHippVol}$  is the  $p$ -value of top 10 genes associated to LHippVol;  $P_{RHippVol}$  is the  $p$ -value of top 10 genes associated to RHippVol;  $P_{LAccumVol}$  is the  $p$ -value of top 10 genes associated to LAccumVol;  $P_{RAccumVol}$  is the  $p$ -value of top 10 genes associated to RAccumVol;  $P_{LPutamVol}$  is the  $p$ -value of top 10 genes associated to LPutamVol;  $P_{RPutamVol}$  is the  $p$ -value of top 10 genes associated to RPutamVol, Bold font indicates  $p$ -value < 0.000001, Italic font indicates  $p$ -value < 0.00001

**Table 5** The characteristics of the identified consensus modules in 10 iPINBPA runs

CM	RunA	Ta: the top subnetwork in RunA. Sb: the most similar subnetwork to Ta in RunB										
		RunB	1	2	3	4	5	6	7	8	9	10
1	TN11	Rank of Sb in RunB	TN <sub>11</sub>	TN <sub>12</sub>	TN <sub>13</sub>	TN <sub>14</sub>	TN <sub>15</sub>	TN <sub>16</sub>	TN <sub>17</sub>	TN <sub>18</sub>	TN <sub>19</sub>	TN <sub>1,10</sub>
		DC(x, y)	1	0.87	0.63	0.38	0.87	0.53	0.44	0.57	0.92	0.51
2	TN21	Rank of Sb in RunB	TN <sub>21</sub>	TN <sub>22</sub>	TN <sub>23</sub>	TN <sub>24</sub>	TN <sub>25</sub>	TN <sub>26</sub>	TN <sub>27</sub>	TN <sub>28</sub>	TN <sub>29</sub>	TN <sub>2,10</sub>
		DC(x, y)	1	0.77	0.58	0.71	0.95	0.59	0.47	0.7	0.77	0.38
3	TN31	Rank of Sb in RunB	TN <sub>31</sub>	TN <sub>32</sub>	TN <sub>33</sub>	TN <sub>34</sub>	TN <sub>45</sub>	TN <sub>36</sub>	TN <sub>37</sub>	TN <sub>38</sub>	TN <sub>49</sub>	TN <sub>4,10</sub>
		DC(x,y)	1	0.89	0.89	1	0.74	0.74	0.77	0.89	0.89	0.44
4	TN41	Rank of Sb in RunB	TN <sub>41</sub>	TN <sub>32</sub>	TN <sub>33</sub>	TN <sub>34</sub>	TN <sub>45</sub>	TN <sub>36</sub>	TN <sub>37</sub>	TN <sub>38</sub>	TN <sub>49</sub>	TN <sub>3,10</sub>
		DC(x,y)	1	0.61	0.61	0.7	0.96	0.67	0.67	0.7	0.7	0.36
5	TN51	Rank of Sb in RunB	TN <sub>51</sub>	TN <sub>62</sub>	TN <sub>53</sub>	TN <sub>74</sub>	TN <sub>75</sub>	TN <sub>86</sub>	TN <sub>67</sub>	TN <sub>88</sub>	TN <sub>89</sub>	TN <sub>4,10</sub>
		DC(x, y)	1	1	1	1	1	1	1	1	0.96	0.52

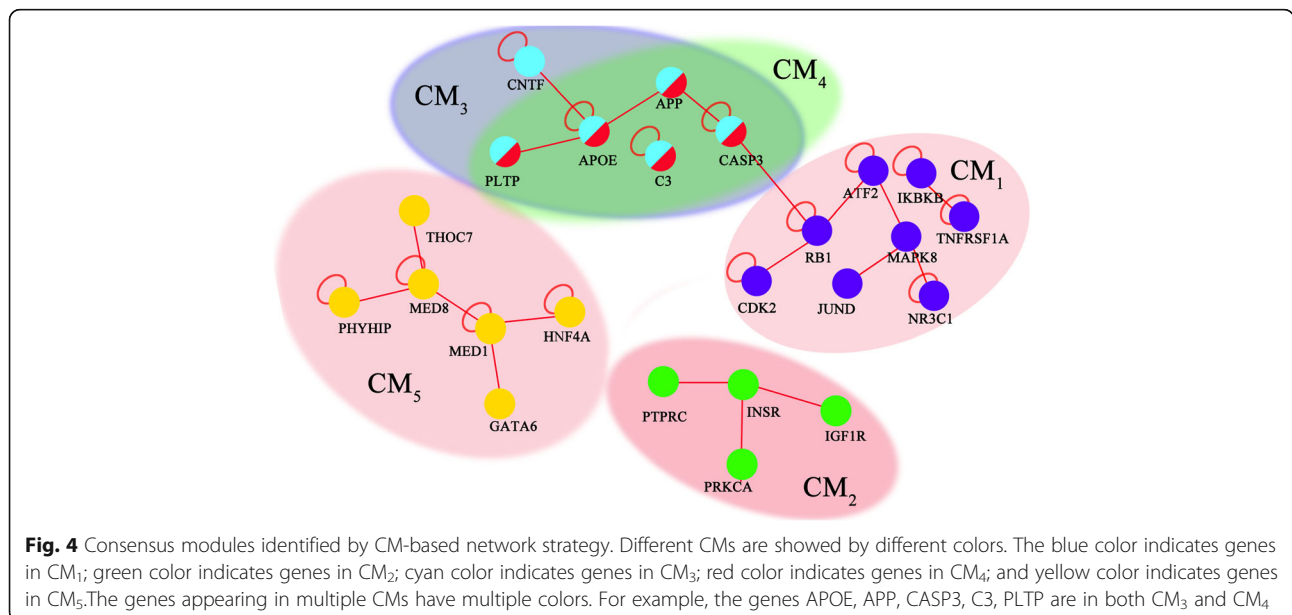
value = 2.50E-05, FDR = 1.54E-03), Legionellosis (*p*-value = 1.90E-04, FDR = 7.03E-03), Pertussis (*p*-value = 3.63E-04, FDR = 8.40E-03), Serotonergic synapse (*p*-value = 8.02E-04, FDR = 1.28E-02), Tuberculosis (*p*-value = 2.00E-03, FDR = 1.92E-02), Herpes simplex infection (*p*-value = 2.13E-03, FDR = 1.92E-02) and so on. It has been reported that *Legionella pneumophila*, one species of *Legionella*, is an intracellular microorganism that causes Legionellosis. This type of pulmonary infection is usually associated with neurological dysfunction [31]. Serotonergic neurotransmission and synapse activity are highlighted as primary pathological factors in neuropsychiatric symptoms [32, 33]. Pertussis toxin inhibits the apoptosis and DNA synthesis caused by FAD APP mutants which precedes FAD APP-mediated apoptosis in neurons and inhibition of neuronal entry into the cell cycle inhibits the apoptosis

[34]. Apoptotic pathways and DNA synthesis are activated in neurons in the brains of individuals with AD.

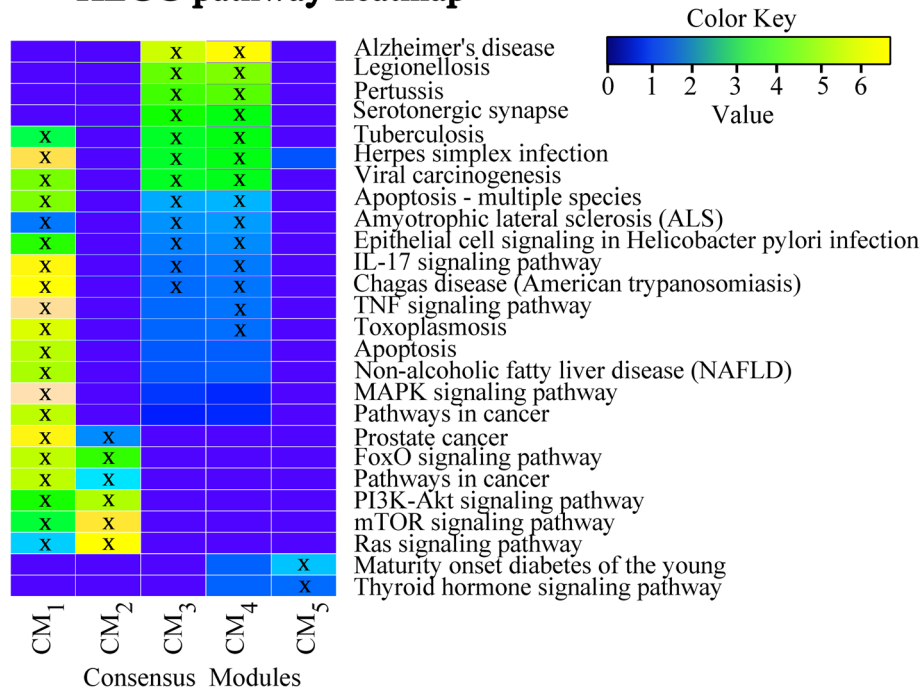
Due to the limited number of samples available to us, in this work we were only able to perform a discovery study. In the future, when more data become available replication studies in independent cohorts warrant investigation to validate the identified CMs.

**Conclusion**

In this study, we performed MGAS analysis to explore the multivariate imaging genetic association effects for a set of AD-related subcortical measures. In addition, we conducted the iPINBPA network analysis to discover consensus modules related to these imaging phenotypes from a protein-protein interaction network. The MGAS analysis identified several genes associated with the studied



### KEGG pathway heatmap



**Fig. 5** Functional annotation of the five identified consensus modules (CM<sub>1</sub>-CM<sub>5</sub>) using KEGG pathways. The five consensus modules were treated as five gene sets, and went through pathway enrichment analysis based on the KEGG pathway database. The enrichment results at a nominal statistical threshold of  $p < 0.05$  are shown.  $-\log_{10}(p)$  values are color-mapped and displayed in the heat map. Heat map blocks labeled with "x" reach the nominal significance level of  $p < 0.05$ . Only top enrichment findings are included in the heat map, and so each row (pathway) has at least one "x" block

imaging phenotypes, including *APOE*, *TOMM40*, *APOC1*, *LAMA1*, *XYLB*, *HSD17B7P2* and others. The statistical power of coupling MGAS with iPINBPA was higher than traditional GWAS method, and yielded findings missed by GWAS. In this work, we reported top five consensus modules based on MGAS results. Network-based analysis can take into account information on biological relationships to interpret GWAS data. Our results suggested several susceptible genes and network modules for further investigation and replication to better understand the genetic mechanism of Alzheimer's Disease.

### Methods

#### Subjects and data

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative

(ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see [www.adni-info.org](http://www.adni-info.org).

Baseline 3 T MRI scans, demographic information, and diagnosis for the ADNI-1 and ADNI-GO/2 cohorts were downloaded [35]. MRI scans were analyzed using FreeSurfer version 5.1 for brain segmentation. We examined the volume measures of 14 subcortical ROIs; see Tables 1-2. We performed analysis of variance (ANOVA) to evaluate

**Table 6** The properties of Consensus Modules identified from the PPI network

Consensus Module	Nodes	Score	P-value
CM <sub>1</sub>	MAPK8, ATF2, TNFRSF1A, JUND, NR3C1, RB1, IKKBK, CDK2	3.32	9.00E-04
CM <sub>2</sub>	IGF1R, PRKCA, INSR, PTPRC	1.41	1.59E-01
CM <sub>3</sub>	APP, APOE, CASP3, C3, PLTP, CNTF	4.37	1.24E-05
CM <sub>4</sub>	C3, PLTP, APP, APOE, CASP3	4.32	1.56E-05
CM <sub>5</sub>	MED8, GATA6, HNF4A, MED1, THOC7, PHYHIP	1.89	5.88E-02

Scores were computed by using the adjusted network score



**Table 7** The pathway of genes appearing in all five consensus modules

NO.	Pathway	p-value	FDR	Genes
1	Alzheimer's disease	2.50E-05	1.54E-03	CASP3, APOE, APP
2	Legionellosis	1.90E-04	7.03E-03	CASP3, C3
3	Pertussis	3.63E-04	8.40E-03	CASP3, C3
4	Serotonergic synapse	8.02E-04	1.28E-02	CASP3, APP
5	Tuberculosis	2.00E-03	1.92E-02	CASP3, C3
6	Herpes simplex infection	2.13E-03	1.92E-02	CASP3, C3
7	Viral carcinogenesis	2.51E-03	1.92E-02	CASP3, C3
8	Apoptosis - multiple species	1.32E-02	2.94E-02	CASP3
9	Amyotrophic lateral sclerosis (ALS)	2.03E-02	3.92E-02	CASP3
10	<i>Staphylococcus aureus</i> infection	2.23E-02	4.12E-02	C3

the diagnostic effect on 14 volume measures. Using the significance level of  $p < 0.05$ , we focused on the volume measures of eight subcortical ROIs (i.e., LAmygVol, RAmygVol, LHippVol, RHippVol, LAccumVol, LPutamVol, RPutamVol; see Table 2) in subsequent genetic association studies.

Genotyping data of both ADNI-1 and ADNI-GO/2 cohorts were downloaded, and then quality controlled and combined as described in [36]. A total of 866 non-Hispanic Caucasian participants with both complete subcortical imaging measurements and genotyping data were included in the study. The study sample ( $N = 866$ ) included 183 cognitively normal (CN), 95 significant memory concern (SMC), 281 early MCI (EMCI), 177 late MCI (LMCI) and 130 AD subjects. The demographic and clinical characteristics of participants, stratified by the diagnosis, are shown in Table 1.

#### Multivariate genome wide association study

GWAS was performed to examine the main effects of 563,980 SNPs on eight subcortical measures as quantitative traits (QTs). Linear regression model was performed using PLINK to examine the association between each SNP-QT pair (<https://www.cog-genomics.org/plink2>) [37]. An additive genetic model was tested with age, gender and brain volume as covariates. We computed the correlation matrix ( $8 \times 8$  matrix) for the QT data containing eight imaging phenotypes (Fig. 2). We applied the MGAS (Multivariate Gene-based Association test by extended Simes procedure) tool to all 563,980 SNPs and examined their multivariate gene-based associations with eight imaging QTs [13]. A Manhattan plot was generated using R (<http://www.r-project.org>) to visualize the gene-level MGAS results for our work (Fig. 3). We obtained one multivariate gene-based  $p$ -value  $P_{MGAS}$  as follows.

$$P_{MGAS} = \min \left( \frac{q_e p_j}{q_{ej}} \right)$$

Here,  $q_e$  represents the effective number of  $p$ -values within a gene,  $q_{ej}$  represents the effective number of  $p$ -values among the top  $j$   $p$ -values where  $j$  runs from 1 to  $8 \times 563,980$ , and  $p_j$  represents the  $j$ -th  $p$ -value in the list of ordered  $p$ -values.  $P_{MGAS}$  is the smallest weighted  $p$ -value within a gene associated with the null hypothesis that none of the eight phenotypes are related to the 563,980 SNPs within the gene, and the alternative hypothesis that at least one of the eight phenotypes is related to at least one of the 563,980 SNPs. We identified 1386 genes with  $p$ -value  $< 0.05$  [13, 38].

#### Identifying consensus models using iPINBPA

This study used the protein-protein interaction (PPI) data from the Human Protein Reference Database (HPRD, <http://www.hprd.org>) [39], containing 9617 proteins and 39,240 interactions. Gene-level  $p$ -values obtained from MGAS of subcortical imaging phenotypes were mapped to the PPI network. Then the network was followed by an iPINBPA (integrative protein-interaction-network-based pathway analysis) procedure [19] to identify enriched PPI network modules. Consensus modules (CMs) were identified using the following approach based on our prior study [24].

Briefly, building on our prior study, we focus on analyzing the top 5 subnetworks ( $TN_1, TN_2, TN_3, TN_4, TN_5$ ) in each iPINBPA run. Let  $TN_{ij}$  be the top  $i$ -th subnetwork identified in the  $j$ -th run, where  $i \in \{1, 2, \dots, 5\}$  and  $j \in \{1, 2, \dots, 10\}$ . We first find  $SN_n(TN_{ij})$ , which is the most similar subnetwork to  $TN_{ij}$  in the  $n$ -th run, where  $n \in \{1, 2, \dots, 10\} \setminus \{j\}$ . Clearly, we have.

$$SN_n(TN_{ij}) = \operatorname{argmax}_{sn} DC(TN_{ij}, sn),$$

where  $sn$  is any subnetwork enriched in Run  $n$ , and  $DC(x, y)$  indicates the dice coefficients between two subnetworks  $x$  and  $y$ . Consequently, for Run  $j$ , we define its  $i$ -th consensus module  $CM_{ij}$  as follows.

$$CM_{ij} = TN_{ij} \cap \left( \bigcap_{j=1,2,\dots,10}^{i=1,2,\dots,5} SN_n(TN_{ij}) \right), n \in \{1, 2, \dots, 10\} \setminus \{j\}$$

Namely,  $CM_{ij}$  is the intersection of  $TN_{ij}$  and its most similar subnetworks identified in all the other runs. In our empirical study, we will report the consensus modules based on Run 1, i.e.,  $CM_{i1}$  as the  $i$ -th consensus module.

### Functional analysis

Cytoscape 3.4 [40] was used to visualize the identified CMs. We used ToppGene online tool (<https://toppgene.cchmc.org/>) for functional enrichment analysis. The ToppGene suite is an advanced bioinformatics tool, it could detect and arrange candidate genes through a comprehensive assessment of a variety of factors, including gene ontology (GO) annotating, phenotype, signaling pathway and protein interactions from a specific list of genes [41]. In this case, the top 10 findings of our multivariate gene-based association analysis were analyzed for functional enrichment. For the identified CMs, we also performed functional enrichment analysis using the ToppGene Suite.

### Abbreviations

AD: Alzheimer's disease; QTs: Quantitative traits; GWAS: Genome-Wide Association Studies; QT: Quantitative trait; SNPs: Single Nucleotide Polymorphisms; AD: Alzheimer's Disease; MGAS: Multivariate Gene-based Association test by extended Simes procedure; iPINBPA: Integrative Protein-Interaction-Network-Based Pathway Analysis; PPI: Protein-protein interaction; CM: Consensus Module; CMs: Consensus Modules; SNP: Single nucleotide polymorphism; SNPs: Single nucleotide polymorphisms; HPRD: Human Protein Reference Database; ADNI: Alzheimer's Disease Neuroimaging Initiative; MRI: Magnetic resonance imaging; PET: Positron emission tomography; MCI: Mild cognitive impairment; CN: Cognitively normal; SMC: Significant memory concern; EMCI: Early mild cognitive impairment; LMCI: Late mild cognitive impairment; ANOVA: Analysis of variance; DC: Dice's coefficient; GNS: Groups of enriched subnetworks; TN: Top subnetwork.

### Acknowledgements

Alzheimer's Disease Neuroimaging Initiative. Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

### About this supplement

This article has been published as part of BMC Genomics Volume 21 Supplement 112,020: Bioinformatics methods for biomedical data science. The full contents of the supplement are available at <https://bmcbgenomics.biomedcentral.com/articles/supplements/volume-21-supplement-11>.

### Authors' contributions

X.M., J.L., H.L. and L.S. led and supervised the research. X.M., H.L., L.S., and J.L. designed the research and wrote the article. X.M. performed the gene analysis, MGAS analysis, network mapping, consensus modules analysis, and pathway analysis. C.B., Q.Z., F.C., and Z.X. prepared subcortical imaging measurements and genotyping data, and performed quality control. X.Y., J.Y., S.L.R., A.J.S., and L.S. provided guidance and consultation on the genotyping and biomarker details about ADNI data, data preprocessing, quality control, population stratification, GWAS protocol, gene analysis, network analysis and pathway analysis. All the authors reviewed, commented, edited and approved the manuscript.

### Funding

Data analysis, result interpretation and manuscript writing were supported in part by grants from National Natural Science Foundation of China (61901063, 61803117, 61773134) and Hei long jiang Provincial Natural Science Foundation of China (QC2018080), Humanities and Social Science Fund of Ministry of Education of China (19YJCZH120), Qinglan Project of Jiangsu Province(2020), Science and Technology Plan Project of Changzhou (CE20205042) and Natural Science Foundation of the Jiangsu Higher Education Institutions of China (19KJB520003). Data analysis, method development, and manuscript editing were supported in part by NIH R01 EB022574, U01 AG024904, R01 AG19771, and P30 AG10133 at U Penn and IU. Publication costs were funded by National Natural Science Foundation of China (61901063).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd. and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

### Availability of data and materials

The genotyping and subcortical imaging phenotypes data were downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu/>). Application for access to the ADNI data can be submitted by anyone at <http://adni.loni.usc.edu/data-samples/access-data/>. The process includes completion of an online application form and acceptance of Data Use Agreement. We have received administrative approval for access to the ADNI database. The Human PPI data were downloaded from the public Human Protein Reference Database (<http://www.hprd.org/>).

### Ethics approval and consent to participate

The study procedures were approved by the institutional review boards of all participating centers ([https://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)), and written informed consent was obtained from all participants or their authorized representatives. Ethics approval was obtained from the institutional review boards of each institution involved: Oregon Health and Science University; University of Southern California; University of California—San Diego; University of Michigan; Mayo Clinic, Rochester; Baylor College of Medicine; Columbia University Medical Center; Washington University, St. Louis; University of Alabama at Birmingham; Mount Sinai School of Medicine; Rush University Medical Center; Wien Center; Johns Hopkins University; New York University; Duke University Medical Center; University of Pennsylvania; University of Kentucky; University of Pittsburgh; University of Rochester Medical Center; University of California, Irvine; University of Texas Southwestern Medical School; Emory University; University of Kansas, Medical Center; University of California, Los Angeles; Mayo Clinic, Jacksonville; Indiana University; Yale University School of Medicine; McGill University, Montreal-Jewish General Hospital; Sunnybrook Health Sciences, Ontario; U.B.C.Clinic for AD & Related Disorders; Cognitive Neurology—St. Joseph's, Ontario; Cleveland Clinic Lou Ruvo Center for Brain Health; Northwestern University; Premiere Research Inst (Palm Beach Neurology); Georgetown University Medical Center; Brigham

and Women's Hospital; Stanford University; Banner Sun Health Research Institute; Boston University; Howard University; Case Western Reserve University; University of California, Davis—Sacramento; Neurological Care of CNY; Parkwood Hospital; University of Wisconsin; University of California, Irvine—BIC; Banner Alzheimer's Institute; Dent Neurologic Institute; Ohio State University; Albany Medical College; Hartford Hospital, Olin Neuropsychiatry Research Center; Dartmouth-Hitchcock Medical Center; Wake Forest University Health Sciences; Rhode Island Hospital; Butler Hospital; UC San Francisco; Medical University South Carolina; St. Joseph's Health Care Nathan Kline Institute; University of Iowa College of Medicine; Cornell University; and University of South Florida: USF Health Byrd Alzheimer's Institute.

#### Consent for publication

Not applicable.

#### Competing interests

The authors have no actual or potential conflicts of interest including any financial, personal, or other relationships with other people or organizations that could inappropriately influence (bias) our work.

#### Author details

<sup>1</sup>School of Computer Information & Engineering, Changzhou Institute of Technology, Changzhou 213032, China. <sup>2</sup>College of Automation, Harbin Engineering University, Harbin 150001, China. <sup>3</sup>School of Computer Science, Northeast Electric Power University, Jilin 132012, China. <sup>4</sup>Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA. <sup>5</sup>Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN 46202, USA. <sup>6</sup>Department of BioHealth Informatics, Indiana University School of Informatics and Computing, Indianapolis, IN 46202, USA.

Published: 29 December 2020

#### References

- Andrews SJ, Fulton-Howard B, Goate A. Interpretation of risk loci from genome-wide association studies of Alzheimer's disease. *Lancet Neurol.* 2020;19(4):326–35.
- Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, Boland A, Vronskaya M, van der Lee SJ, Amlie-Wolf A, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet.* 2019;51(3):414–30.
- Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, Sealock J, Karlsson IK, Hagg S, Athanasiu L, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet.* 2019;51(3):404–13.
- 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.
- Mukherjee S, Kim S, Ramanan VK, Gibbons LE, Nho K, Glymour MM, Ertekin-Taner N, Montine TJ, Saykin AJ, Crane PK, et al. Gene-based GWAS and biological pathway analysis of the resilience of executive functioning. *Brain Imaging Behav.* 2014;8(1):110–8.
- Ramanan VK, Kim S, Holohan K, Shen L, Nho K, Risacher SL, Foroud TM, Mukherjee S, Crane PK, Aisen PS, et al. 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.* 2012;6(4):634–48.
- Yao X, Yan J, Liu K, Kim S, Nho K, Risacher SL, Greene CS, Moore JH, Saykin AJ, Shen L, et al. Tissue-specific network-based genome wide study of amygdala imaging phenotypes to identify functional interaction modules. *Bioinformatics.* 2017;33(20):3250–7.
- Shen L, Thompson PM. Brain imaging genomics: integrated analysis and machine learning. *Proc IEEE Inst Electr Electron Eng.* 2020;108(1):125–62.
- Shen L, Kim S, Risacher SL, Nho K, Swaminathan S, West JD, Foroud T, Pankratz N, Moore JH, Sloan CD, et al. Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: a study of the ADNI cohort. *NeuroImage.* 2010;53(3):1051–63.
- 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.
- Lee S, Kerns S, Ostrer H, Rosenstein B, Deasy JO, Oh JH. Machine learning on a genome-wide association study to predict late genitourinary toxicity after prostate radiation therapy. *Int J Radiat Oncol Biol Phys.* 2018;101(1):128–35.
- Kim J, Zhang Y, Pan W. Alzheimer's disease neuroimaging I: powerful and adaptive testing for multi-trait and multi-SNP associations with GWAS and sequencing data. *Genetics.* 2016;203(2):715–31.
- Van der Sluis S, Dolan CV, Li J, Song Y, Sham P, Posthuma D, Li MX. MGAS: a powerful tool for multivariate gene-based genome-wide association analysis. *Bioinformatics.* 2015;31(7):1007–15.
- Liu Y, Maxwell S, Feng T, Zhu X, Elston RC, Koyuturk M, Chance MR. Gene, pathway and network frameworks to identify epistatic interactions of single nucleotide polymorphisms derived from GWAS data. *BMC Syst Biol.* 2012; 6(Suppl 3):S15.
- Marin M, Esteban FJ, Ramirez-Rodrigo H, Ros E, Saez-Lara MJ. An integrative methodology based on protein-protein interaction networks for identification and functional annotation of disease-relevant genes applied to channelopathies. *BMC Bioinformatics.* 2019;20(1):565.
- Gosak M, Markovic R, Dolensek J, Slak Rupnik M, Marhl M, Stozar A, Perc M. Network science of biological systems at different scales: a review. *Phys Life Rev.* 2018;24:118–35.
- Yan J, Risacher SL, Shen L, Saykin AJ. Network approaches to systems biology analysis of complex disease: integrative methods for multi-omics data. *Brief Bioinform.* 2018;19(6):1370–81.
- Li J, Chen F, Zhang Q, Meng X, Yao X, Risacher SL, Yan J, Saykin AJ, Liang H, Shen L, et al. Genome-wide network-assisted association and enrichment study of amyloid imaging phenotype in Alzheimer's disease. *Curr Alzheimer Res.* 2019;16(13):1163–74.
- Wang L, Mousavi P, Baranzini SE. iPINBPA: an integrative network-based functional module discovery tool for genome-wide association studies. *Pac Symp Biocomput.* 2015:255–66.
- Vroom CR, Posthuma D, Li MX, Dolan CV, van der Sluis S. Multivariate gene-based association test on family data in MGAS. *Behav Genet.* 2016;46(5): 718–25.
- Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet.* 2011;12(1):56–68.
- D'Aoust LN, Cummings AC, Laux R, Fuzzell D, Caywood L, Reinhart-Mercer L, Scott WK, Pericak-Vance MA, Haines JL. Examination of candidate Exonic variants for association to Alzheimer disease in the Amish. *PLoS One.* 2015; 10(2):e0118043.
- Shehu A, Mao J, Gibori GB, Halperin J, Le J, Devi YS, Merrill B, Kiyokawa H, Gibori G. Prolactin receptor-associated protein/17beta-hydroxysteroid dehydrogenase type 7 gene (Hsd17b7) plays a crucial role in embryonic development and fetal survival. *Mol Endocrinol.* 2008;22(10): 2268–77.
- Cong W, Meng X, Li J, Zhang Q, Chen F, Liu W, Wang Y, Cheng S, Yao X, Yan J, et al. Genome-wide network-based pathway analysis of CSF t-tau/Abeta1-42 ratio in the ADNI cohort. *BMC Genomics.* 2017; 18(1):421.
- Tamura K, Wakimoto H, Agarwal AS, Rabkin SD, Bhere D, Martuza RL, Kuroda T, Kasmieh R, Shah K. Multimechanistic tumor targeted oncolytic virus overcomes resistance in brain tumors. *Mol Ther.* 2013;21(1):68–77.
- Zhang YW, Tong YQ, Zhang Y, Ding H, Zhang H, Geng YJ, Zhang RL, Ke YB, Han JJ, Yan ZX, et al. Two novel susceptibility SNPs for ischemic stroke using exome sequencing in Chinese Han population. *Mol Neurobiol.* 2014; 49(2):852–62.
- Chen ZL, Strickland S. Neuronal death in the hippocampus is promoted by plasmin-catalyzed degradation of laminin. *Cell.* 1997;91(7):917–25.
- Ouchida M, Kanzaki H, Ito S, Hanafusa H, Jitsumori Y, Tamaru S, Shimizu K. Novel direct targets of miR-19a identified in breast cancer cells by a quantitative proteomic approach. *PLoS One.* 2012;7(8):e44095.
- Marcuzzo S, Bonanno S, Kapetis D, Barzago C, Cavalcante P, D'Alessandro S, Mantegazza R, Bernasconi P. Up-regulation of neural and cell cycle-related microRNAs in brain of amyotrophic lateral sclerosis mice at late disease stage. *Mol Brain.* 2015;8:5.
- Boos L, Campbell IL, Ames R, Wetsel RA, Barnum SR. Deletion of the complement anaphylatoxin C3a receptor attenuates, whereas ectopic expression of C3a in the brain exacerbates, experimental autoimmune encephalomyelitis. *J Immunol.* 2004;173(7):4708–14.

31. Lagana P, Soraci L, Gambuzza ME, Delia S. Innate immune surveillance in the central nervous system in *Legionella Pneumophila* infection. *CNS Neurol Disord Drug Targets*. 2017.
32. Ballard C, Aarsland D, Francis P, Corbett A. Neuropsychiatric symptoms in patients with dementias associated with cortical Lewy bodies: pathophysiology, clinical features, and pharmacological management. *Drugs Aging*. 2013;30(8):603–11.
33. Doraiswamy PM. Non-cholinergic strategies for treating and preventing Alzheimer's disease. *CNS Drugs*. 2002;16(12):811–24.
34. McPhie DL, Coopersmith R, Hines-Peralta A, Chen Y, Ivins KJ, Manly SP, Kozlowski MR, Neve KA, Neve RL. DNA synthesis and neuronal apoptosis caused by familial Alzheimer disease mutants of the amyloid precursor protein are mediated by the p21 activated kinase PAK3. *J Neurosci*. 2003; 23(17):6914–27.
35. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A*. 2000;97(20): 11050–5.
36. Li J, Zhang Q, Chen F, Yan J, Kim S, Wang L, Feng W, Saykin AJ, Liang H, Shen L. Genetic interactions explain variance in cingulate amyloid burden: an AV-45 PET genome-wide association and interaction study in the ADNI cohort. *Biomed Res Int*. 2015;2015:647389.
37. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81(3):559–75.
38. Li MX, Gui HS, Kwan JS, Sham PC. GATES: a rapid and powerful gene-based association test using extended Simes procedure. *Am J Hum Genet*. 2011; 88(3):283–93.
39. Goel R, Harsha HC, Pandey A, Prasad TS. Human protein reference database and human Proteinpedia as resources for phosphoproteome analysis. *Mol BioSyst*. 2012;8(2):453–63.
40. Demchak B, Hull T, Reich M, Liefeld T, Smoot M, Ideker T, Mesirov JP. Cytoscape: the network visualization tool for GenomeSpace workflows. *F1000Res*. 2014;3:151.
41. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res*. 2009;37(Web Server issue):W305–11.

## Publisher's Note

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

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

