ADAPTIVE TESTING OF SNP-BRAIN FUNCTIONAL CONNECTIVITY ASSOCIATION VIA A MODULAR NETWORK ANALYSIS

CHEN GAO, JUNGHI KIM and WEI PAN[*], for the Alzheimer's Disease Neuroimaging Initiative[*]

Division of Biostatistics, School of Public Health, University of Minnesota
[*]*E-mail: weip@biostat.umn.edu*

Due to its high dimensionality and high noise levels, analysis of a large brain functional network may not be powerful and easy to interpret; instead, decomposition of a large network into smaller subcomponents called modules may be more promising as suggested by some empirical evidence. For example, alteration of brain modularity is observed in patients suffering from various types of brain malfunctions. Although several methods exist for estimating brain functional networks, such as the sample correlation matrix or graphical lasso for a sparse precision matrix, it is still difficult to extract modules from such network estimates. Motivated by these considerations, we adapt a weighted gene co-expression network analysis (WGCNA) framework to resting-state fMRI (rs-fMRI) data to identify modular structures in brain functional networks. Modular structures are identified by using topological overlap matrix (TOM) elements in hierarchical clustering. We propose applying a new adaptive test built on the proportional odds model (POM) that can be applied to a high-dimensional setting, where the number of variables (p) can exceed the sample size (n) in addition to the usual $p < n$ setting. We applied our proposed methods to the ADNI data to test for associations between a genetic variant and either the whole brain functional network or its various subcomponents using various connectivity measures. We uncovered several modules based on the control cohort, and some of them were marginally associated with the APOE4 variant and several other SNPs; however, due to the small sample size of the ADNI data, larger studies are needed.

Keywords: aSPU test; brain functional connectivity; functional MRI; proportional odds model; single nucleotide polymorphism; weighted gene co-expression network analysis; WGCNA.

1. Introduction

Resting-state functional magnetic resonance imaging (rs-fMRI) is gaining popularity in studies of brain functional connectivity with applications to detection of subtle network reorganizations in Alzheimer's disease.[1] Disruption of connectivity in the brain functional network is related to many pathological conditions in the brain, such as Alzheimer's disease,[2] schizophrenia,[3] or autism.[4] This necessitates the development of methods for modelling the brain functional network its statistical inference.

A network is comprised of nodes and edges connecting the nodes. Based on functional MRI data, a popular choice of nodes are brain regions of interest (ROIs) while the edges are connectivities reflecting statistical dependencies between ROIs. An important network model, the scale-free network,[5] assumes that most nodes in a network are sparsely connected with the exception of a few "hub" nodes that are densely connected with other nodes. In the scale-free network model, new connections are more likely to occur for those hub nodes with already-high

connectivity. There has been empirical evidence supporting this model for brain functional networks,[6] though it is still debatable. In addition, the scale-free network model also admits a modular topological structure, which can be extracted for more efficient analyses for human brains.

Methods for drawing statistical inference to distinguish brain connectivity for different groups of subjects are still under development. The first question encountered is how to define brain functional connectivity. Ref. 7 discussed the choice between Pearson's marginal correlation coefficient and partial correlation coefficient as a network connectivity measure, though other measures are possible and it is yet unclear which one is best. To reduce dimensionality and to reach sparseness, graphical lasso is often used for estimating networks for different groups. Since an estimated network with the imposed sparsity penalty may not demonstrate modular structures, a better approach is to directly discover the modules in a network. A general framework for estimating scale-free networks and detecting modules is proposed in Ref. 8 for gene network analysis, which has gained tremendous popularity in genomics.[9] It starts by defining a similarity measure between two nodes in a network, called adjacency, using the marginal correlation coefficient. Soft-thresholding is then applied, leading to a weighted network. The soft-thresholded adjacency is further transformed to a topological overlap matrix (TOM) element, which is converted to a dissimilarity measure for hierarchical clustering, grouping closely connected nodes together as modules in the network. The above framework not only provides multiple network connectivity measures, but also carries out modular structure identification. The connectivity measures and identified modules in the brain functional network may help statistical inference and offer biological insights.[9]

In this paper, for the first time, we adapt the use of WGCNA for gene expression data to rs-fMRI data, constructing weighted brain functional networks and identifying their subnetworks or modules using the Alzheimer's Disease Neuroimaging Initiative (ADNI) data. We explored using the adjacency matrix element and TOM element, in addition to the marginal correlation or covariance, to characterize connectivity in brain functional networks. Taking advantages of detected network modules, we conduct association analysis of genetic variants with not only the whole brain functional network, but also its various subcomponents, including its modules, which aims to not only improve statistical power, but also offer better biological interpretation. We propose applying a new adaptive association test based on a proportional odds model (POM) accounting for the ordinal nature of the SNP genotype. We found evidence of associations between several network modules and the APOE4 variant, which is by far the most significant genetic risk factor for Alzheimer's disease.

This paper is organized as follows. We first review the method of WGCNA, including its module identification, then introduce the adaptive test based on a POM. We demonstrate the application of our methods to the ADNI data before summarizing our findings and future research directions in the discussion section.

2. Methods

2.1. *Module detection via weighted gene co-expression network analysis*

In this section, we briefly review the work in Ref. 8 on the weighted gene-coexpression network analysis (WGCNA) framework for network construction and module identification.

2.1.1. *Adjacency matrix*

The first step of the WGCNA framework is to define a similarity measure between gene expression profiles; in the current context, we use the BOLD signals in each of multiple ROIs from one or more subjects to calculate a similarity between any two ROIs. The similarity measure is required to take values between 0 and 1. A typical choice of this similarity measure is the absolute value of the Pearson correlation coefficient $s_{uv} = |cor(u, v)|$, for nodes u and v. Another choice, which preserves the sign of correlation, is defined as $s_{uv} = [1 + cor(u, v)]/2$. We refer the first one as unsigned similarity measure, and the second one as the signed similarity measure. From our experience of applications to the ADNI data, the identified modules have negligible differences using either unsigned or signed similarity measure. We used the unsigned similarity measure throughout this paper.

Once the similarity measure is computed, the next step is to transform the similarity matrix $S = [s_{uv}]$ into an adjacency matrix using an adjacency function. Hard thresholding is often used to yield a binary or unweighted network with a 0/1 adjacency indicating no-connection/connection and thus possible loss of information, though a more efficient multi-scale approach with multiple thresholds yielding a set of binary networks has been proposed.[10] Soft thresholding is a simple and popular alternative with more flexibilities. One choice is the power adjacency function

$$a_{uv} = power(s_{uv}, \delta) \equiv |s_{uv}|^{\delta} \tag{1}$$

with parameter δ, which is chosen as the smallest integer such that the scale-free network model fitting is above a certain threshold.

2.1.2. *Topological overlap matrix*

Instead of using only the adjacency matrix, Ref. 11 advocated a topological overlap matrix $\Omega = [\omega_{uv}]$ with its element as a potentially more useful measure that reflects the relative interconnectedness of two nodes u and v after accounting for their shared neighbors. The topological overlap matrix element is defined as

$$\omega_{uv} = \frac{l_{uv} + a_{uv}}{\min\{k_u, k_v\} + 1 - a_{uv}} \tag{2}$$

with $k_u = \sum_v a_{uv}$ and $l_{uv} = \sum_q a_{uq}a_{qv}$. For a binary network with $a_{uv} = 0$ or 1, k_u is the connectivity of node u representing the number of its direct neighbors, while l_{uv} equals the number of nodes that connect both nodes u and v; $\omega_{uv} = 0$ if the nodes u and v are not connected and they are not connected to the same neighbors; in contrast, $\omega_{uv} = 1$ if the nodes u and v are connected and the neighbors of the node with fewer edges are also connected to the one with more edges. For any network, $0 \leq a_{uv} \leq 1$ implies $0 \leq \omega_{uv} \leq 1$.

2.1.3. *Module identification*

To identify modules in a network, we need to have a dissimilarity or distance measure. An intuitive way is to convert a similarity measure. Based on the topological overlap matrix element ω_{uv}, we can simply define the dissimilarity measure as $d_{uv}^{\omega} = 1 - \omega_{uv}$. The TOM-based dissimilarity d_{uv}^{ω} is used as the input for average linkage hierarchical clustering. The output from hierarchical clustering is a dendrogram composed of branches and leaves. In a brain functional network, each leaf corresponds to a ROI. The hierarchical clustering algorithm groups the closest ROIs and forms the branches. By cutting the branches of the dendrogram, closely related ROIs are identified as a module. Among the several methods for cutting the branches of the dendrogram, the default used in the WGCNA framework is Dynamic Tree Cut from the R package `dynamicTreeCut`.

Once modules are identified, one can calculate an intramodular connectivity

$$\omega.in_u = \sum_{v \in M} \omega_{uv} \tag{3}$$

for each node u in its module M. Ref. 8 pointed out that intramodular connectivities $\omega.in$ may represent important features of the nodes (i.e. ROIs).

2.2. *An adaptive association test based on the proportional odds model*

Let $Y_i = 0, 1, 2$ denote the count of the minor allele for subject i for a given SNP of interest, then Y_i has $J = 3$ ordered categories. The logistic regression model cannot be applied in this situation, because it only allows the response variable to be binary. A popular choice for ordinal data is the proportional odds model (POM),[12] which we will briefly describe here.

Suppose subject i has p network connectivities denoted by $X_i = (x_{i1}, \ldots, x_{ip})$ and l covariates denoted by $Z_i = (z_{i1}, \ldots, z_{il})$. For the proportional odds model, we define the regression coefficients $\beta = (\beta_1, \ldots, \beta_p)'$ for the network connectivities and $\delta = (\delta_1, \ldots, \delta_l)'$, and a vector of intercepts $\alpha = (\alpha_0, \ldots, \alpha_{J-2})'$. The proportional odds model is

$$\text{logit}[Pr(Y_i \leq j)] = \alpha_j + Z_i\delta + X_i\beta, \qquad j = 0, 1. \tag{4}$$

The likelihood for equation Eq. 4 can be derived based on the multinomial distribution for the categorical variable Y_i, from which maximum likelihood estimates and statistical inference can be obtained as implemented in R package `MASS` or `VGAM`. However, numerical issues such as non-convergence arise when p, the dimension of β, is relatively large as compared to the sample size n.

Here we propose applying a class of tests that are applicable to the high-dimensional setting with $p > n$, from which an adaptive test is constructed to summarize information across the tests. No that most existing tests cannot be applied to the case $p > n$. To test the null hypothesis $H_0 : \beta = (\beta_1, \beta_2, \ldots, \beta_p)' = 0$, we can use the score vector derived in Ref. 13,

$$U_\beta = \sum_{i=1}^{n} \sum_{j=0}^{J-2} (1 - \hat{r}_{i(j-1)} - \hat{r}_{ij}) \cdot I(Y_i = j) \cdot X_i \tag{5}$$

where $\hat{r}_{ij} = \exp(\hat{\alpha} + Z_i\hat{\delta})/[1 + \exp(\hat{\alpha} + Z_i\hat{\delta})]$ comes from the fitted null model of Eq. 4 (i.e. with

$\beta = 0$); $\hat{\alpha}$ and $\hat{\delta}$ are estimated by the `polr` function in the R package `MASS`. Let U_k denote the kth component of the score vector $U_\beta = (U_1, \dots, U_p)'$. The SPU($\gamma$) test statistic is defined as

$$T_{SPU(\gamma)} = \sum_{k=1}^{p} U_k^\gamma, \tag{6}$$

where $\gamma \geq 1$ is an integer. As the parameter γ increases, a connectivity with a larger absolute value of the score gains a higher weight. In the extreme situation, when $\gamma \to \infty$ as an even integer, $SPU(\infty)$ takes only the maximum component of the score vector, i.e., $T_{SPU(\infty)} = \max_{k=1}^{p} |U_k|$.

The p-values of the SPU tests are computed by permuting the residuals from the null model B times, and the p-value can be calculated as

$$P_{SPU(\gamma)} = \frac{(\sum_{b=1}^{B} I[|T_{SPU(\gamma)}^{(b)}| \geq |T_{SPU(\gamma)}|] + 1)}{(B+1)}, \tag{7}$$

where $T_{SPU(\gamma)}^{(b)}$ is the SPU(γ) statistic based on the bth set of permuted residuals. Since the value of γ that yields highest power cannot be determined a priori, an adaptive SPU (aSPU) test is introduced to combine the evidence across multiple SPU tests,

$$T_{aSPU} = \min_{\gamma \in \Gamma} P_{SPU(\gamma)}, \tag{8}$$

where $P_{SPU(\gamma)}$ is the p-value of $SPU(\gamma)$ test statistics and Γ is a set of integers for the power of aSPU test. In the numerical examples throughout this paper, we chose γ from the set $\Gamma = \{1, 2, \dots, 8, \infty\}$. To calculate the p-value of T_{aSPU}, we can use the same permutation scheme as used for calculating the p-values of T_{SPU}'s. For each permuted residual set b, after calculating $T_{SPU(\gamma)}^{(b)}$ and its p-value $p_\gamma^{(b)} = (\sum_{b_1 \neq b} I[T_{SPU(\gamma)}^{(b_1)} \geq T_{SPU(\gamma)}^{(b)}] + 1)/B$. Then we can obtain $T_{aSPU}^{(b)} = \min_{\gamma \in \Gamma} p_\gamma^{(b)}$, and the p-value of T_{aSPU} is

$$P_{aSPU} = \frac{(\sum_{b=1}^{B} I[T_{aSPU}^{(b)} \leq T_{aSPU}] + 1)}{(B+1)}. \tag{9}$$

A step-wise procedure is used to gradually increase B if needed. We can start with $B = 10^3$ initially, then increase to $B = 10^5$ (or bigger) if a p-value is smaller than 5×10^{-3} (or smaller). The test is implemented in R package `POMaSPU` to be available on CRAN.

3. Results

3.1. *ADNI Data*

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (`adni.loni.usc.edu`). We included all subjects from the normal and Alzheimer's disease (AD) groups in the ADNI data. We applied motion correction and global signal regression to reduce noises.

Here we used the power adjacency function $a_{uv} = power(s_{uv}, \beta) = |s_{uv}|^\beta$ (equation (1)). β was selected as the smallest β such that the scale-free model fitting R^2 was above a pre-set threshold 0.85.

3.2. *Distinct modular structures in brain functional networks based on APOE4 SNP genotype scores*

For the ADNI data, we grouped the subjects based on the APOE4 SNP (rs429358) minor allele counts $(0, 1, 2)$. APOE4 plays a major role in the pathogenesis of Alzheimer's disease.[14,15] The APOE4 variant is a major risk factor for both early- and late-onset Alzheimer's disease.[14,15] We removed those subjects with a missing rs429358 value, resulting in a total of 162 subjects. Among them, 73 subjects have no minor allele at rs429358, whereas 67 subjects have one minor allele and 22 subjects have two. In order to establish possible modular structures in brain functional networks in the normal condition, we first applied the WGCNA framework to the rs-fMRI data of the control subjects only. Specifically, for each ROI, we concatenated the BOLD time series of all the control subjects, which were used to calculate the similarity between any two ROIs (i.e. the absolute value of Pearson's correlation between any two BOLD time series), then conducting the subsequent analyses in the WGCNA framework. At the end, we identified four modules based on the data from the control cohort (Figure 1).

Based on the modules identified, we continued to explore them for each APOE4 SNP genotype group. To measure the network connectivities, we used the correlation matrix, covariance matrix, and the topological overlap matrix (TOM). The rows and columns are ordered in the same way as in Figure 1. Distinct modular structures seem to be present in the correlation, covariance and TOM plots across the APOE4 genotype groups (Figure 2).

3.3. *Adaptive testing for SNP-module associations*

Using the APOE4 SNP (rs429358) minor allele counts as the response in a POM, we tested the association between the APOE4 SNP and the network connectivities. Covariates including age, gender and years of education were adjusted. Using the aSPU test, we found that the covariance matrix elements were marginally associated with the APOE4 SNP ($P = 0.033$, Table 1). We further decomposed the whole network connectivities into two exclusive subsets: connectivities within the four modules and those between the modules. Both the between-modular covariance and TOM were associated with the APOE4 SNP with $P < 0.05$.

Next we focused on the network connectivities in each individual module, and tested their association with the APOE4 SNP (Table 2). The network connectivities defined by the correlations in the yellow module showed evidence of association with the APOE4 SNP ($P = 0.017$). In addition, the network connectivities defined by covariance matrix elements in the blue and yellow modules were also associated with the APOE4 SNP ($P = 0.034, P = 0.011$).

Finally we tested for association between each module-specific intramodular connectivity $\omega.in$ and the APOE4 SNP. Only the yellow module showed a significant association with $P = 0.007$.

There are 30 and 19 ROIs in the blue and yellow modules, respectively. The ROIs identified in the yellow modules includes left/right sides of posterior cingulate cortex, angular gyrus, superior frontal cortex, middle frontal cortex, and inferior frontal cortex. For comparison, Ref. 13 identified 18 nodes related to the default mode network (DMN), including left/right sides of superior frontal cortex, medial prefrontal cortex, ventral anterior cingulate cortex, posterior cingulate cortex, parahipppocampal cortex, inferior parietal cortex, angular, middle
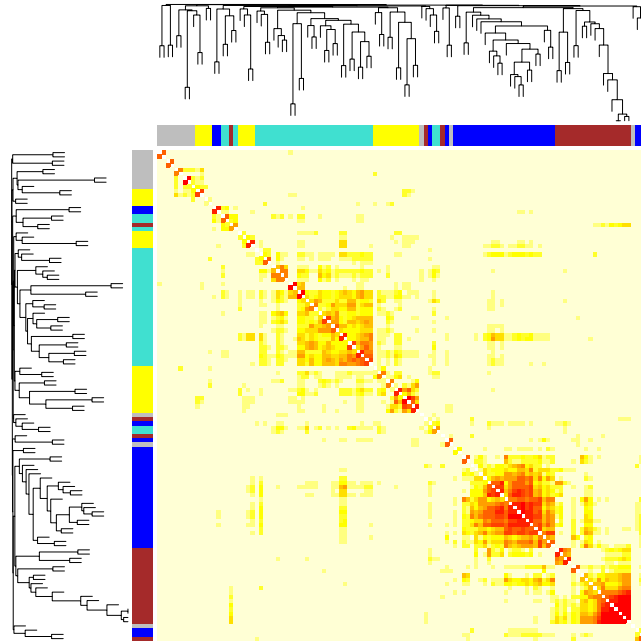
Fig. 1.   TOM plot of the whole brain functional network and its modules for normal subjects. The rows and columns are the ROIs, ordered by their distance in the tree.

temporal gyrus, and inferior temporal cortex.[16–18] We found that 15 ROIs in the yellow module are also related to the 18 nodes in the DMN. For example, the posterior cingulate cortex plays a pivotal role in the default mode network of the brain.[19,20] The posterior cingulate cortex is linked to cognitive functions such spatial memory, configural learning, and maintenance of discriminative avoidance learning and.[21,22] It is shown in the DMN that Alzheimer's disease affects the posterior cingulate cortex.[20] Angular gyrus is another region found in both DMN and the yellow module. Loss of grey matter volume in angular gyrus has been associated with dementia and progression to Alzheimer's disease.[23] The association between the APOE4 variant and the network connectivity measures in the yellow module also uncovers some key brain regions in DMN that were found to be affected in Alzheimer's disease.

The ROIs in the blue module includes the left/right sides of hippocampus, lingual gyrus, cuneus, calcarine fissure and superior occipital gyrus, cerebellum and vermis. Hippocampus is well known for its key role in memory.[24] Hippocampal neuronal loss and structural change have been connected with Alzheimer's disease.[25,26] Alzheimer's disease patients have also demonstrated neuronal and glial loss and structural changes in cerebellum and vermis.[27] Lingual gyrus, cuneus, calcarine fissure and superior occipital gyrus are located in the occipital lobe, which are mainly related to vision processing.[28] In addition, lingual gyrus plays an important role in the identification and recognition of words.[29] The association between the APOE4

SNP and the network connectivity measures may reflect the pathological changes of the brain functional network in Alzheimer's disease.

3.4. *GWAS scan with individual modules*

We tested for associations of the SNPs across the whole genome with the functional connectivity measures in the yellow and blue modules respectively. For genotype data, we included all SNPs with a minor allele frequency (MAF) ≥ 0.05, genotyping rate $\geq 90\%$, and passing the Hardy-Weinberg equilibrium test with a p-value > 0.001. After filtering with the above criteria, we obtained 579,382 SNPs.

The genome-wide scan showed that among the SNPs associated with the network connectivities (measured by Pearson's correlation) in the yellow module, rs17114690 on chromosome 14 was the only SNP that had a p-value smaller than 10^{-3}. Three SNPs were founded to be associated with the network connectivities (correlations) in the blue module, with p-values smaller than 10^{-3}. They are located on chromosome 1 (rs7536105, rs11265187) and chromosome 2 (rs17498117). rs7536105 is located in the chromatin interactive region, while rs11265187 is located in the enhancer region of gene olfactory receptor family 10 subfamily J member 9 pseudogene (OR10J9P).

The genome-wide scan also identified 5 SNPs associated with the intramodular network connectivity $\omega.in$ for the yellow module, with $P < 10^{-5}$. They are located on chromosome 1 (rs6656071, rs12043216), chromosome 7 (rs1178127, rs12674460), and chromosome 13 (rs2819239). SNP rs1178127 is a missense variant in gene histone deacetylase 9 (HDAC9),[30] an important gene with function in transcriptional regulation and cell cycle in the Wnt signalling pathway.

4. Discussion

In this paper we adapted WGCNA for network construction and module detection to rs-fMRI data. Based on the identified modules, we also proposed applying a new adaptive association test for single SNP association with the connectivities of the whole network or its components in a proportional odds model. While the whole network was not associated, some module-based connectivities were significantly associated with the APOE4 SNP rs429358. Given the major role of APOE4 in the pathogenesis of Alzheimer's disease, our finding seems plausible, suggesting its possible use for genome-wide scans to detect SNP variants associated with altered brain networks and AD. Although none of the associations was highly or genome-wide significant, it was perhaps due to a too small sample size; larger studies are needed. Our use of modules, with either various ROI-to-ROI connectivities (e.g. TOM in addition to standard correlations) or some module-based node measures (such as intramodular connectivity), not only may reduce the dimension and thus improve the statistical power, but also can enhance result interpretation, highlighting where is the association if any. In particular, we found that intramodular connectivities showed more significant associations with more SNPs, possibly due to their lower dimensions (i.e. p_1 in a module with p_1 ROIs as compared to $p_1(p_1 - 1)/2$ of ROI-to-ROI connectivities) and/or higher information contents.

The multiple traits used in this paper, including various network connectivity measures in

the whole network or its various subcomponents, differ from most of the previous neuroimaging studies,[31] in which the focus was on some direct measures on ROIs, not their connectivities as shown here. These phenotypes are often high dimensional with dimension exceeding the sample size. Many software packages cannot handle such a situation with $p > n$, which limits their use. The adaptive association test used in this paper can be applied to such high-dimensional traits. It can be a useful and powerful method for identifying associations between high-dimensional neuroimaging traits and SNPs. In this paper, we have focused on the study of the association between neuroimaging phenotypes and SNP genotype scores; however, other ordinal outcomes such as a disease status (e.g. normal, MCI and AD in the ADNI data) can be tested for their associations with neuroimaging and other endophenotypes.

Acknowledgment

References

1. Y. I. Sheline and M. E. Raichle, *Biological Psychiatry* **74**, 340 (2013).
2. K. Supekar, V. Menon, D. Rubin, M. Musen and M. D. Greicius, *PLoS Computational Biology* **4**, e1000100 (2008).
3. A. Zalesky, A. Fornito, M. L. Seal, L. Cocchi, C.-F. Westin, E. T. Bullmore, G. F. Egan and C. Pantelis, *Biological Psychiatry* **69**, 80 (2011).
4. M. K. Belmonte, G. Allen, A. Beckel-Mitchener, L. M. Boulanger, R. A. Carper and S. J. Webb, *The Journal of Neuroscience* **24**, 9228 (2004).
5. A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).

6. C. C. Hilgetag and A. Goulas, *Brain Structure and Function* , 1 (2015).
7. J. Kim, W. Pan and the Alzheimer's Disease Neuroimaging Initiative, *NeuroImage: Clinical* **9**, 625 (2015).
8. B. Zhang and S. Horvath, *Statistical Applications in Genetics and Molecular Biology* **4** (2005).
9. L. Zhu, J. Lei, B. Devlin and K. Roeder, *arXiv preprint arXiv:1606.00252* (2016).
10. H. Lee, H. Kang, M. K. Chung, B.-N. Kim and D. S. Lee, *IEEE transactions on medical imaging* **31**, 2267 (2012).
11. E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai and A.-L. Barabási, *Science* **297**, 1551 (2002).
12. P. McCullagh, *Journal of the Royal Statistical Society. Series B (Methodological)* , 109 (1980).
13. J. Kim, W. Pan and the Alzheimer's Disease Neuroimaging Initiative, *Unpublished* (2016).
14. J. Kim, J. M. Basak and D. M. Holtzman, *Neuron* **63**, 287 (2009).
15. E. Genin, D. Hannequin, D. Wallon, K. Sleegers, M. Hiltunen, O. Combarros, M. J. Bullido, S. Engelborghs, P. De Deyn, C. Berr *et al.*, *Molecular Psychiatry* **16**, 903 (2011).
16. M. D. Greicius, G. Srivastava, A. L. Reiss and V. Menon, *Proceedings of the National Academy of Sciences of the United States of America* **101**, 4637 (2004).
17. L. Q. Uddin, A. Clare Kelly, B. B. Biswal, F. Xavier Castellanos and M. P. Milham, *Human Brain Mapping* **30**, 625 (2009).
18. S. Passow, K. Specht, T. C. Adamsen, M. Biermann, N. Brekke, A. R. Craven, L. Ersland, R. Grüner, N. Kleven-Madsen, O.-H. Kvernenes *et al.*, *Human Brain Mapping* **36**, 2027 (2015).
19. P. Fransson and G. Marrelec, *Neuroimage* **42**, 1178 (2008).
20. R. L. Buckner, J. R. Andrews-Hanna and D. L. Schacter, *Annals of the New York Academy of Sciences* **1124**, 1 (2008).
21. R. J. Maddock, A. S. Garrett and M. H. Buonocore, *Neuroscience* **104**, 667 (2001).
22. R. Leech and D. J. Sharp, *Brain* **137**, 12 (2014).
23. G. Karas, J. Sluimer, R. Goekoop, W. Van Der Flier, S. Rombouts, H. Vrenken, P. Scheltens, N. Fox and F. Barkhof, *American Journal of Neuroradiology* **29**, 944 (2008).
24. L. R. Squire, *Psychological Review* **99**, 195 (1992).
25. B. T. Hyman, G. W. Van Hoesen, A. R. Damasio and C. L. Barnes, *Science* **225**, 1168 (1984).
26. M. J. West, P. D. Coleman, D. G. Flood and J. C. Troncoso, *The Lancet* **344**, 769 (1994).
27. M. Sjöbeck and E. Englund, *Dementia and Geriatric Cognitive Disorders* **12**, 211 (2001).
28. R. Malach, J. Reppas, R. Benson, K. Kwong, H. Jiang, W. Kennedy, P. Ledden, T. Brady, B. Rosen and R. Tootell, *Proceedings of the National Academy of Sciences* **92**, 8135 (1995).
29. A. Mechelli, G. W. Humphreys, K. Mayall, A. Olson and C. J. Price, *Proceedings of the Royal Society of London B: Biological Sciences* **267**, 1909 (2000).
30. C. Fernandez-Rozadilla, L. De Castro, J. Clofent, A. Brea-Fernandez, X. Bessa, A. Abuli, M. Andreu, R. Jover, R. Xicola, X. Llor *et al.*, *PLoS One* **5**, e12673 (2010).
31. L. Shen, S. Kim, S. L. Risacher, K. Nho, S. Swaminathan, J. D. West, T. Foroud, N. Pankratz, J. H. Moore, C. D. Sloan, M. J. Huentelman, D. W. Craig, B. M. DeChairo, S. G. Potkin, C. R. Jack Jr, M. W. Weiner, A. J. Saykin and the Alzheimer's Disease Neuroimaging Initiative, *Neuroimage* **53**, 1051 (2010).

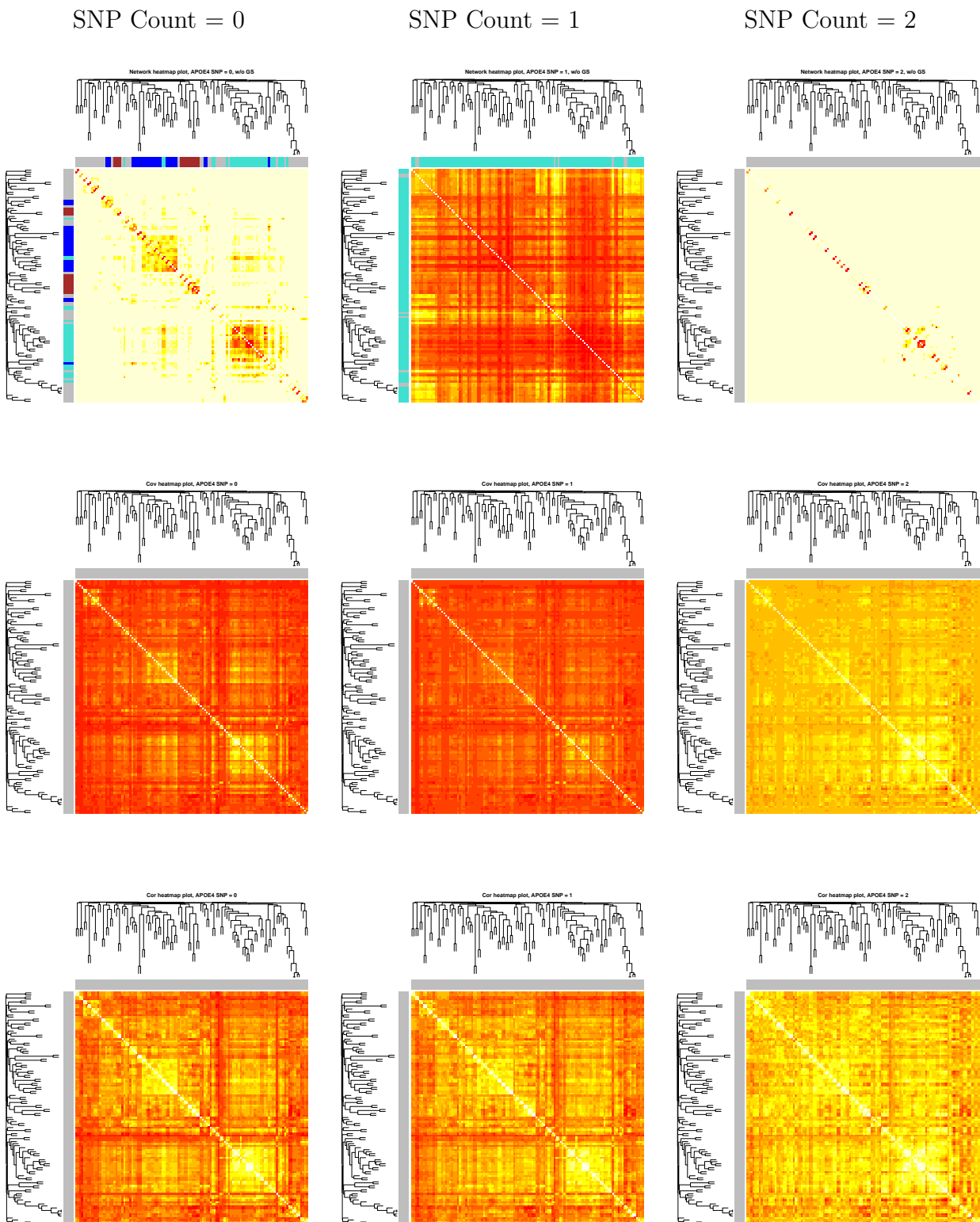SNP Count = 0   SNP Count = 1   SNP Count = 2

Fig. 2. TOM plot (top), covariance matrix plot (middle) and correlation matrix plot (bottom) of the brain functional networks for the three genotype groups based on APOE4 SNP (rs429358) (with its minor allele counts equal to 0, 1 or 2 from left to right).

Table 1. P-values of the tests for SNP-whole network associations using the correlation, covariance, TOM or adjacency matrix elements as the network connectivity measure respectively. W-mod and Btw-mod stand for within-modular and between-modular, respectively.

Test	Correlation			Covariance			TOM			Adjacency		
	All	W-mod	Btw-mod	All	W-mod	Btw-mod	All	W-mod	Btw-mod	All	W-mod	Btw-mod
SPU(1)	0.477	0.365	0.526	0.052	0.052	0.087	0.530	0.637	0.109	0.477	0.527	0.479
SPU(2)	0.161	0.154	0.207	0.012	0.016	0.008	0.099	0.250	0.014	0.477	0.515	0.487
SPU(3)	0.323	0.202	0.377	0.018	0.025	0.010	0.817	0.902	0.224	0.472	0.482	0.498
SPU(4)	0.150	0.122	0.197	0.019	0.009	0.004	0.325	0.424	0.172	0.463	0.434	0.516
SPU(5)	0.248	0.137	0.299	0.130	0.066	0.004	0.892	0.987	0.317	0.442	0.402	0.528
SPU(6)	0.141	0.101	0.209	0.120	0.008	0.003	0.444	0.533	0.330	0.416	0.365	0.554
SPU(7)	0.216	0.111	0.267	0.429	0.052	0.004	0.657	0.890	0.393	0.381	0.348	0.568
SPU(8)	0.137	0.089	0.225	0.188	0.009	0.004	0.498	0.603	0.410	0.348	0.334	0.583
SPU(∞)	0.122	0.079	0.356	0.210	0.009	0.007	0.463	0.706	0.655	0.263	0.239	0.460
aSPU	0.208	0.146	0.296	0.033	0.025	0.007	0.181	0.384	0.039	0.356	0.328	0.585

Table 2. P-values of the tests for SNP-individual network module associations using the correlation, covariance, TOM or adjacency matrix elements as the network connectivity measure.

Module	Test	Correlation		Covariance		TOM		Adjacency	
		W-mod	Btw-mod	W-mod	Btw-mod	W-mod	Btw-mod	W-mod	Btw-mod
Blue	SPU(1)	0.140	0.272	0.018	0.012	0.946	0.093	0.504	0.488
	SPU(2)	0.090	0.077	0.025	0.001	0.440	0.067	0.515	0.481
	SPU(3)	0.099	0.144	0.028	0.001	0.793	0.396	0.502	0.455
	SPU(4)	0.101	0.071	0.033	0.001	0.528	0.265	0.496	0.423
	SPU(8)	0.139	0.075	0.050	0.011	0.566	0.458	0.464	0.264
	SPU(∞)	0.267	0.070	0.069	0.015	0.571	0.554	0.458	0.180
	aSPU	0.160	0.119	0.034	0.003	0.654	0.141	0.581	0.244
Turquoise	aSPU	0.648	0.172	0.277	0.011	0.139	0.192	0.605	0.309
Brown	aSPU	0.260	0.182	0.016	0.015	0.459	0.219	0.249	0.327
Yellow	SPU(1)	0.040	0.500	0.005	0.084	0.083	0.172	0.357	0.510
	SPU(2)	0.024	0.219	0.005	0.012	0.060	0.155	0.323	0.509
	SPU(3)	0.020	0.350	0.005	0.015	0.147	0.458	0.297	0.514
	SPU(4)	0.016	0.220	0.010	0.006	0.188	0.445	0.262	0.522
	SPU(8)	0.008	0.271	0.100	0.004	0.317	0.690	0.177	0.500
	SPU(∞)	0.006	0.525	0.200	0.008	0.383	0.848	0.129	0.411
	aSPU	0.017	0.354	0.011	0.010	0.106	0.289	0.183	0.524