Biological Research

SHORT REPORT

CrossMark

Disease association and inter-connectivity analysis of human brain specific co-expressed functional modules

Kimin Oh[1†], Taeho Hwang[2†], Kihoon Cha[1] and Gwan-Su Yi[1*]

Abstract

Background: In the recent studies, it is suggested that the analysis of transcriptomic change of functional modules instead of individual genes would be more effective for system-wide identification of cellular functions. This could also provide a new possibility for the better understanding of difference between human and chimpanzee.

Results: In this study, we analyzed to find molecular characteristics of human brain functions from the difference of transcriptome between human and chimpanzee's brain using the functional module-centric co-expression analysis. We performed analysis of brain disease association and systems-level connectivity of species-specific co-expressed functional modules.

Conclusions: Throughout the analyses, we found human-specific functional modules and significant overlap between their genes in known brain disease genes, suggesting that human brain disorder could be mediated by the perturbation of modular activities emerged in human brain specialization. In addition, the human-specific modules having neurobiological functions exhibited higher networking than other functional modules. This finding suggests that the expression of neural functions are more connected than other functions, and the resulting high-order brain functions could be identified as a result of consolidated inter-modular gene activities. Our result also showed that the functional module based transcriptome analysis has a potential to expand molecular understanding of high-order complex functions like cognitive abilities and brain disorders.

Keywords: Co-expressed functional module, Disease association, Inter-connectivity analysis

Background

To characterize the molecular bases associated with human's remarkably advanced high-order brain functions and vulnerability to various brain disorders, a comparative analysis between human and chimpanzee brain transcriptome is considered as an important way [1]. Until now, several previous studies compared the transcriptome data of human and chimpanzee brains. Despite some successes, little has been understood to account for unique features of human brain. In the previous study, we showed that a co-expression analysis of functional modules has shown increased sensitivity for identifying implication of more diverse genes and cellular functions that were previously undetected [2]. Recent studies have shown the existence of inter-module co-expression networks in the human brain, suggesting that systems-level relationships might also uniquely constitute human brain specificity [3, 4]. However, the implication of functional modules and their network properties have not been investigated in depth to further extract functional meanings of each module and their interplay with respect to human brain specialization. In this study, we analysed to find molecular characteristics of human brain functions using the human-specific co-expressed functional modules (HS-COMODs) and the chimpanzee-specific co-expressed functional modules (CS-COMODs). Our new approach using integrated analysis of gene expression

*Correspondence: gsyi@kaist.ac.kr
†Kimin Oh and Taeho Hwang contributed equally to this work
[1] Department of Bio and Brain Engineering, KAIST, Daejeon, Korea
Full list of author information is available at the end of the article

BioMed Central

Oh *et al. Biol Res* (2015) 48:67

Page 2 of 6

data should be an aid in molecular interpretation of other complex biological functions too.

Results

Association of HS-COMODs and brain disorders

We questioned if the human vulnerability to neurodegenerative and psychiatric diseases might be relevant to the emergent co-expression of the functional modules in the human brain. To examine the possibility, we compared the HS-COMODs and the CS-COMODs with the brain disease modules. Functional modules associated to neuro/psychological diseases were defined using the Fisher's exact test. The association between the species-specific co-expression and brain disease relevance of the functional modules was measured using the one-sided Fisher's exact test. Our analysis showed that there was a statistically significant overlap between the HS-COMODs and the brain disease modules (Fig. 1). The association was observed more strongly with a more stringent cut-off used for defining the brain disease modules. On the other hand, such association was not observed for the CP-COMODs. This data suggests that the functional modules that were emergently functionalized in the human brain might comprise the cellular functions sensitive to the genetic perturbations and the etiology.

Systems-level characteristics of HS-COMODs

To extract the functional meanings, we analysed the inter-module co-expression network based on the network connectivity. Since the number of non-overlapping functional modules was different for each functional module, we normalized the connectivity by dividing by the number of non-overlapping HS-COMODs. The HS-COMODs ranked with the top 10 highest normalized

Fig. 1 Overlap between species-specific modules and brain disease modules. The association between the species-specific co-expression and brain disease relevance of the functional modules was measured using the one-sided Fisher's exact test

connectivity (one-sided Fisher's exact test, $P = 1.52E{-}03$) (Fig. 2a) includes pathways for 'Alzheimer's disease', 'Huntington's disease', 'GRIN1 network', 'Integration of energy metabolism', and protein interaction networks pivoted by GRIN1 and CDK5, all of which show implication to the learning and memory.

There are 25 neurobiological modules in total 248 non-overlapping HS-COMODs in human-specific functional module networks. In top 5 % (13 functional modules) HS-COMODs which have highest normalized connectivity, there are 5 neurobiological modules (one-sided Fisher's exact test, $P = 4.53E{-}03$). In addition, average connectivity score of neurobiological modules is 0.308 compare to average connectivity of total functional modules and non-neurobiological module is 0.246 and 0.239, respectively. These results showed that the neurobiological modules had significantly higher normalized connectivity than did the other modules. In addition, the neurobiological modules had higher normalized connectivity than the rest, particularly in the 'denser' network (Fig. 2b). Our data suggests the central importance of the neurobiological modules, several of which are influential to the high-order brain functions such as learning or memory, and the prevalent functional cooperation of the neurobiological modules with a wide variety of functional modules in human brain.

We identified the HS-COMODs with the largest normalized connectivity to the neurobiological modules. More than half of the HS-COMODs with normalized connectivity over than 0.5 were involved in protein modification which was overrepresented functional category in our results (Table 1); especially, they were associated with protein ubiquitination. The concerted activity of the individual players such as the E3 ubiquitin ligase and proteasome upon the intra and extracellular stimuli might be crucial, as it has been shown that ubiquitin-dependent protein degradation delicately controls the life cycle of synaptic proteins for synapse regulation and organization, and in turn controls learning and memory [5, 6]. Implications of ubiquitin system have been also reported in previous analyses [7, 8] but their unique characteristics in terms of functional interplay have not been noted. Several E3s, such as AMFR, RNF5, and PARK2, involved in 'protein ubiquitination', were shown to specifically target the synaptic proteins involved in, for example, the postsynaptic density (GO:0014069) and nerve–nerve synaptic transmission (GO:0007270). The other modules, those that are highly connected to the neurobiological modules such as 'HIV Infection', and 'generation of precursor
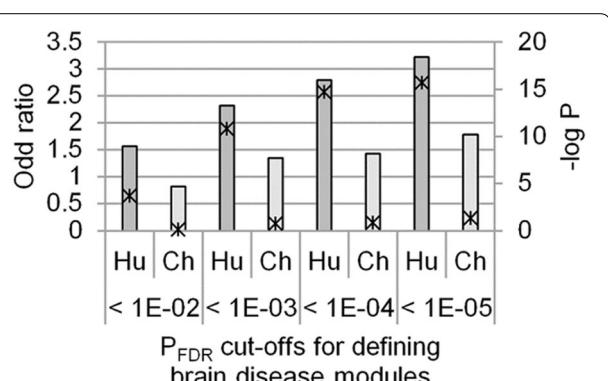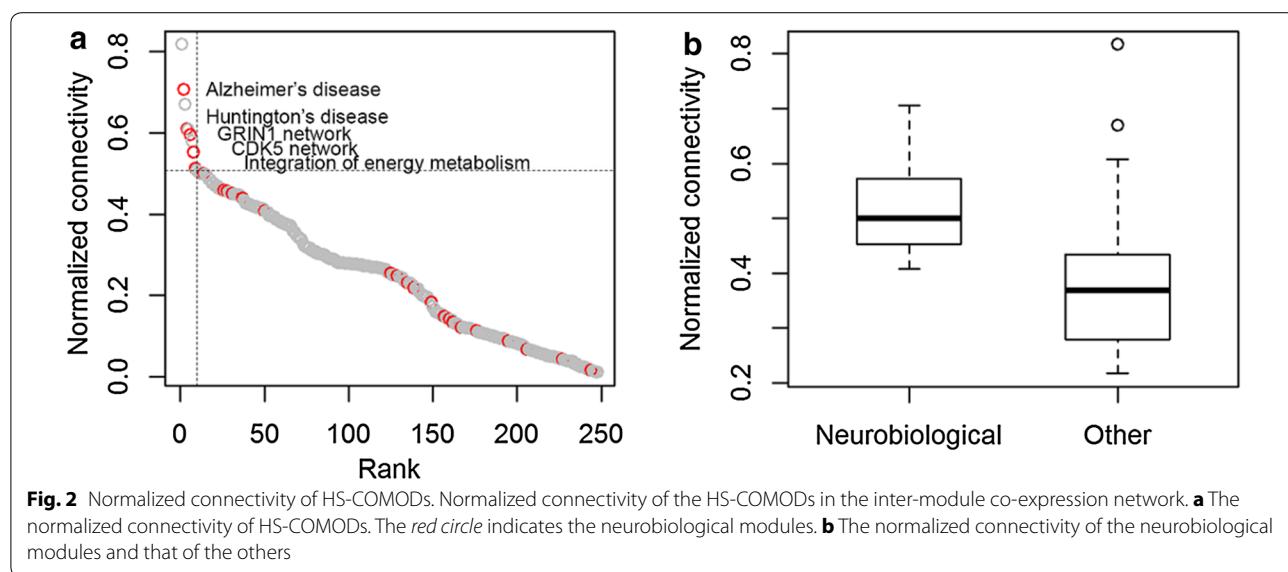
Oh *et al. Biol Res* (2015) 48:67

Page 3 of 6

Fig. 2 Normalized connectivity of HS-COMODs. Normalized connectivity of the HS-COMODs in the inter-module co-expression network. **a** The normalized connectivity of HS-COMODs. The *red circle* indicates the neurobiological modules. **b** The normalized connectivity of the neurobiological modules and that of the others

Table 1 Top 10 HS-COMODs showing highest connectivity to neurobiological modules

ID	Type	Name	Normalized connectivity
86	Biological process	Negative regulation of protein modification process	0.78
250	Pathway	HIV Infection	0.67
45	Biological process	Generation of precursor metabolites and energy	0.64
200	Hub network	RNF5 hub protein network	0.58
226	Pathway	Ubiquitin mediated proteolysis (KEGG)	0.56
284	Pathway	Protein ubiquitination (UniPathway)	0.56
120	Biological process	Protein ubiquitination	0.55
118	Biological process	Protein modification by small protein conjugation	0.55
38	Biological process	Energy derivation by oxidation of organic compounds	0.54
117	Biological process	Proteasomal ubiquitin-dependent protein catabolic process	0.53

metabolites and energy' are also intriguing as it has been shown that positive manipulation of the energy metabolism may be effective in preventing or reversing cognitive impairments [9] and that the impaired immune function in HIV patients may lead to dementia, manifesting the cognitive dysfunction [10]. In summary, inter-module co-expression analysis among the HS-COMODs revealed functional interplay between specific and distinct functional modules. The neurobiological HS-COMODs tend to show co-expression with other diverse HS-COMODs, implying that the high-order human brain function such as cognition has been emerged by the systemic and parallel change of multiple functional modules in a concerted manner.

Discussion

It is important to unravel the molecular mechanisms underlying the high-order human brain functions for the drug development for the cognition enhancement. Despite tremendous researches, only few drugs have been approved [11] and the efficacy of those drugs for both healthy people and patients with cognitive dysfunctions seems to be modest and even harmful in some ways [12]. Recently it has been suggested that targeting multiple genes or complementary mechanisms by multiple drugs would be more desirable approach for the improved therapeutics [13, 14]. Moreover, it has been shown that evolutionarily emerging genes are likely to be the targets of successful drugs [15]. In these regards, this

Oh *et al. Biol Res* (2015) 48:67

Page 4 of 6

study might provide insight to the drug development. To make it more plausible, it would be further necessary to elucidate the specificity of the transcription factors or the epigenetic regulators on those key functional modules.

Conclusion

We showed the significant overlap between the HS-COMODs and the brain disease modules, suggesting that the emergent functionalization of modular activities in human brain might be sensitive to the perturbations. In a system-level analysis into the HS-COMODs, it was shown that the functional modules show complex inter-module co-expression in human brain. Of particular note, the functional modules implicated to the neurobiological processes showed significantly higher connectivity than the others. From this point of view, neurobiological modules might have the potential contribution to cooperate with a wide variety of functional modules to drive cognitive functions at the systems level. Therefore, our findings showed that a systems approach adopted in the interpretation of transcriptomic change between human and chimpanzee brains has a potential to improve our molecular understanding of high-order complex functions like cognitive abilities and brain disorders.

Methods

Identification of brain disease modules

The genes associated with human brain disorders was compiled from public databases including the OMIM, the Genetic Association Database, the Cancer Gene Census, the KEGG Disease, and the HugeNavigator. The heterogeneous disease names were unified into a controlled vocabulary in the Unified Medical Language System (UMLS) using MetaMap. The individual diseases converted to the UMLS terms were further categorized into the 298 groups of higher classifications available in the International Classification of Diseases 10 provided by the World Health Organization. Among these 298 disease groups, we manually inspected and selected 15 categories (Table 2) as representative brain disease. To identify brain disease module, we used the one-sided Fisher's exact test for the function enrichment analysis. The brain disease module was defined to be any functional module [16, 17] significantly enriched (FDR adjusted P <0.01) for

Table 2 Disease categories used to compile brain disease modules

Disease categories
Alzheimer's disease
Cerebral infarction
Cerebral palsy and other paralytic syndromes
Dementia
Epilepsy
Mental and behavioural disorders due to other psycho-active substance use
Mental and behavioural disorders due to use of alcohol
Mental retardation
Mood (affective) disorders
Nerve, nerve root and plexus disorders
Neurotic, stress-related and somatoform disorders
Other congenital malformations of the nervous system
Other diseases of the nervous system
Parkinson's disease
Schizophrenia, schizotypal and delusional disorders

the known genes associated with any brain disease category among the 15 selected categories.

Construction of co-expression network and characterization of inter-modular correlation

We constructed the each species-specific functional modules networks. Each of the HS-COMODs and the CP-COMODs was represented by the module eigengene which correspond the first principal component extracted from the expression levels. We used the ModuleColor and WGCNA R packages to determine the module eigengene. Next, the co-expression between every pair of module eigengenes was calculated with the Pearson's correlation coefficient and the significance of the co-expression was assigned to each pair among the 277 distinct HS-COMODs and the 49 distinct CP-COMODs, respectively [18]. The p values were adjusted using the Benjamini and Hochberg method. However, we restricted our analysis to pairs of each species-specific co-expressed functional modules with no overlapped component genes on the microarray data, since the overlapping genes significantly affect the co-expression measure. Using the significantly coexpressed pairs of HS-COMODs and CP-COMODs (FDR adjusted P <0.01),
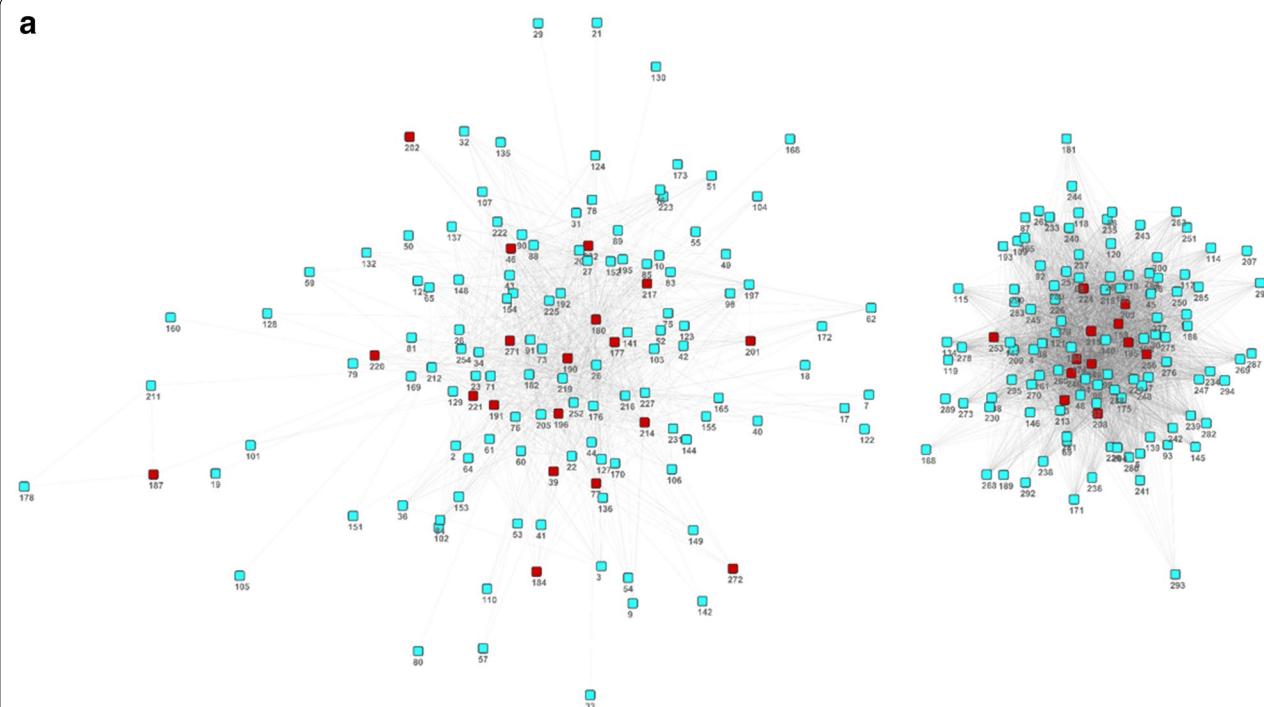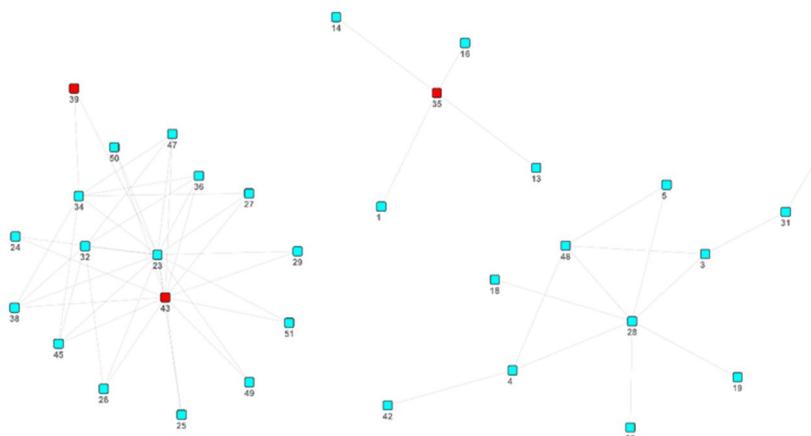
Oh *et al. Biol Res (2015) 48:67*

Page 5 of 6

Fig. 3 Co-expression network of HS-COMODs and CS-COMODs. **a** Co-expression network of HS-COMODs. **b** Co-expression network of CS-COMODs. The neurobiological modules were colored with *red*

we produced the two human-specific functional module networks, one with higher density (Fig. 3a) and the three small and sparse chimpanzee-specific functional module networks (Fig. 3b).

Authors' contribution
KO and TH designed and conducted the experiments and wrote the manuscript. KC helped preparation of data from public databases. GSY designed and supervised the experiments and wrote the manuscript. All authors discussed the results and implications and commented on the manuscript at all stages. All authors read and approved the final manuscript.

Author details
[1] Department of Bio and Brain Engineering, KAIST, Daejeon, Korea. [2] Department of Information and Communications Engineering, KAIST, Daejeon, Korea.

Oh *et al. Biol Res (2015) 48:67*

Page 6 of 6

References
1. Khaitovich P, Muetzel B, She X, Lachmann M, Hellmann I, Dietzsch J, Steigele S, Do HH, Weiss G, Enard W, et al. Regional patterns of gene expression in human and chimpanzee brains. Genome Res. 2004;14(8):1462–73.
2. Kimin Oh TH, Kihoon Cha, Gwan-Su Yi. Functional module-centric interpretationi of transcriptomic change between human and chimpanzee cerebral cortex. In: Bioinformatics and Biomedicine (BIBM), 2014 IEEE International Conference; Belfast. IEEE 2014. pp. 162–167.
3. Konopka G, Friedrich T, Davis-Turak J, Winden K, Oldham MC, Gao F, Chen L, Wang GZ, Luo R, Preuss TM, et al. Human-specific transcriptional networks in the brain. Neuron. 2012;75(4):601–17.
4. Oldham MC, Konopka G, Iwamoto K, Langfelder P, Kato T, Horvath S, Geschwind DH. Functional organization of the transcriptome in human brain. Nat Neurosci. 2008;11(11):1271–82.
5. Mabb AM, Ehlers MD. Ubiquitination in postsynaptic function and plasticity. Annu Rev Cell Dev Biol. 2010;26:179–210.
6. Han Y, Lee H, Park JC, Yi GS. E3Net: a system for exploring E3-mediated regulatory networks of cellular functions. Mol Cell Proteom MCP. 2012;11(4):O111 014076.
7. Nowick K, Gernat T, Almaas E, Stubbs L. Differences in human and chimpanzee gene expression patterns define an evolving network of transcription factors in brain. Proc Natl Acad Sci USA. 2009;106(52):22358–63.
8. Oldham MC, Horvath S, Geschwind DH. Conservation and evolution of gene co-expression networks in human and chimpanzee brains. Proc Natl Acad Sci USA. 2006;103(47):17973–8.
9. Kapogiannis D, Mattson MP. Disrupted energy metabolism and neuronal circuit dysfunction in cognitive impairment and Alzheimer's disease. Lancet Neurol. 2011;10(2):187–98.
10. Kipnis J, Derecki NC, Yang C, Scrable H. Immunity and cognition: what do age-related dementia, HIV-dementia and 'chemo-brain' have in common? Trends Immunol. 2008;29(10):455–63.
11. Wallace TL, Ballard TM, Pouzet B, Riedel WJ, Wettstein JG. Drug targets for cognitive enhancement in neuropsychiatric disorders. Pharmacol Biochem Behav. 2011;99(2):130–45.
12. Husain M, Mehta MA. Cognitive enhancement by drugs in health and disease. Trends in cognitive sciences. 2011;15(1):28–36.
13. Millan MJ, Agid Y, Brune M, Bullmore ET, Carter CS, Clayton NS, Connor R, Davis S, Deakin B, DeRubeis RJ, et al. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. Nat Rev Drug Discovery. 2012;11(2):141–68.
14. Cha K, Hwang T, Oh K, Yi GS. Discovering transnosological molecular basis of human brain diseases using biclustering analysis of integrated gene expression data. BMC Med Inform Decis Mak. 2015;15(Suppl 1):S7.
15. Wang ZY, Fu LY, Zhang HY. Can medical genetics and evolutionary biology inspire drug target identification? Trends Mol Med. 2012;18(2):69–71.
16. Sun CH, Kim MS, Han Y, Yi GS. COFECO: composite function annotation enriched by protein complex data. Nucleic Acids Res. 2009;37(Web Server issue):W350–5.
17. Sun CH, Hwang T, Oh K, Yi GS. DynaMod: dynamic functional modularity analysis. Nucleic Acids Res. 2010;38(Web Server issue):W103–8.
18. Langfelder P, Horvath S. Eigengene networks for studying the relationships between co-expression modules. BMC Syst Biol. 2007;1:54.