

# Combining Neuroinformatics Databases for Multi-Level Analysis of Brain Disorders

Ha Sun Yu<sup>1</sup>, Joon Bang<sup>2</sup>, Yousang Jo<sup>1</sup> and Doheon Lee<sup>1,\*</sup>

<sup>1</sup>Department of Bio and Brain Engineering, KAIST, Daejeon, Korea

<sup>2</sup>Winchester College, College Street, Winchester, UK

**Subject areas;** Biological computation/ Database

**Author contribution;** H.S.Y. writing and design of the manuscript; J.B. writing part of 'brain imaging databases', 'electrophysiological databases', and 'databases of genetic information of brain'; Y.J. writing part of 'brain connectivity database' and 'multi-level analysis combining different types of databases'; D.L. study design, management, and advice of the study and manuscript writing.

\***Correspondence** and requests for materials should be addressed to D.L. (dhlee@kaist.ac.kr).

**Editor;** Keun Woo Lee, Gyeongsang National University, Korea

**Received** August 02, 2012

**Accepted** August 10, 2012

**Published** August 10, 2012

**Citation;** Yu, H.S., et al. Combining Neuroinformatics Databases for Multi-Level Analysis of Brain Disorders. IBC 2012, 4:7, 1-8. doi: 10.4051/ibc.2012.4.3.0007

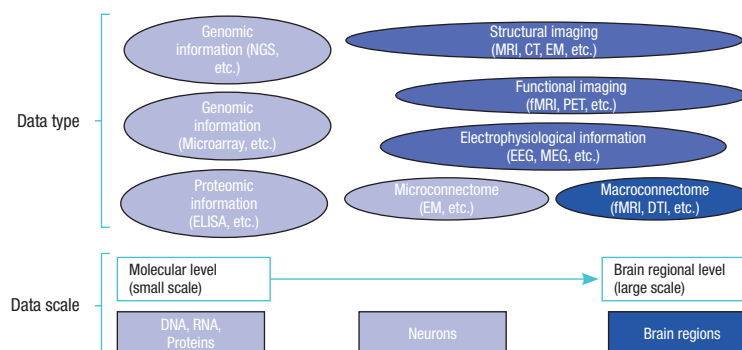
**Funding;** This work was supported by the National Research Foundation of Korea funded by the Korea Government, Ministry of Education, Science and Technology through the WCU (World Class University) program (R32-2008-000-10218-0) and Basic Research Laboratory grant (2009-0086964). It was also supported by the KISTI (Korea Institute of Science and Technology Information).

**Competing interest;** All authors declare no financial or personal conflict that could inappropriately bias their experiments or writing.

© Yu, H.S. et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## SYNOPSIS

With the development of many methods of studying the brain, the field of neuroscience has generated large amounts of information obtained from various techniques: imaging techniques, electrophysiological techniques, techniques for analyzing brain connectivity, techniques for getting molecular information of the brain, etc. A plenty of neuroinformatics databases have been made for storing and sharing this useful information and those databases can be publicly accessed by researchers as needed. However, since there are too many neuroinformatics databases, it is difficult to find the appropriate database depending on the needs of researcher. Moreover, many researchers in neuroscience fields are unfamiliar with using neuroinformatics databases for their studies because data is too diverse for neuroscientists to handle this and there is little precedent for using neuroinformatics databases for their research. Therefore, in this article, we review databases in the field of neuroscience according to both their methods for obtaining data and their objectives to help researchers to use databases properly. We also introduce major neuroinformatics databases for each type of information. In addition, to show examples of novel uses of neuroinformatics databases, we represent several studies that combine neuroinformatics databases of different information types and discover new findings. Finally, we conclude our paper with the discussion of potential applications of neuroinformatics databases.



**Key Words:** neuroscience; database; neuroinformatics databases; usages of neuroinformatics databases; application of neuroinformatics databases

## INTRODUCTION

Brain is one of the most complex organs in our body and it has been studied for a long time in the field of neuroscience. Researchers have studied the brain using various technologies (Figure 1): magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), and computerized tomography (CT) to study functions, connectivity, and structures of the brain; microarray, *in situ* hybridization (ISH), and next generation sequencing (NGS) to study the molecular state of the brain; electroencephalography (EEG) and magnetoencephalography (MEG) to study the electrophysiology of the brain. Each technology has its own pros and cons to study the brain, and researchers have used these technologies based on the objective of their research. A large amount of studies and data for neuroscience has been produced using these technologies and has been opened to the public with the establishment of databases. However, too many neuroinformatics databases have been generated. Furthermore, many neuroscientists are not familiar with using those databases for their research.

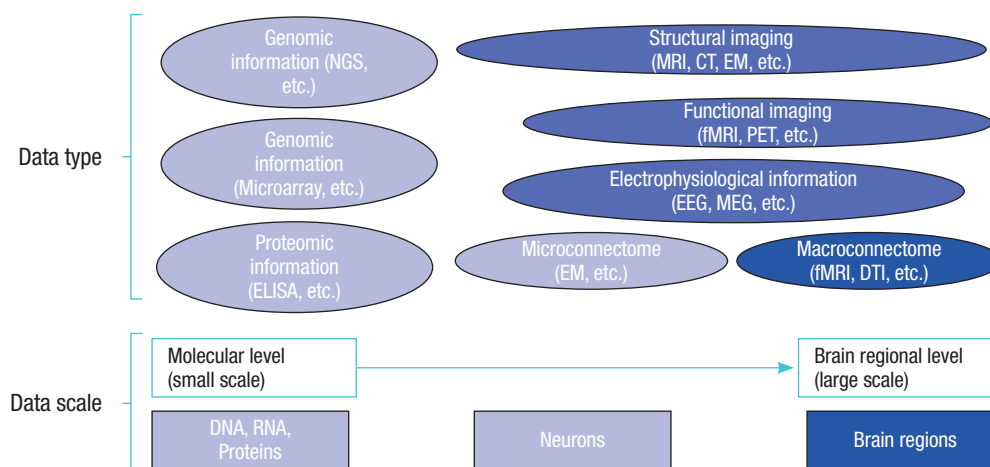
To give a guideline of databases related to neuroscience fields, called neuroinformatics databases, some researchers reviewed neuroinformatics databases<sup>1-4</sup>. For example, Leon French et al. recently reviewed the application of informatics to neuroscience and several useful neuroinformatics databases<sup>3</sup>. However, since much time has passed after their reviews, there have been a lot of new studies of neuroscience, and many databases have been newly established and updated. Furthermore, they did not review neuroinformatics databases related to diseases of brains and they did not represent research using neuroinformatics databases.

Therefore, in this review, we review neuroinformatics databases of various information types: imaging databases; electro-

physiological databases; brain connectivity databases; genetic databases and represent the latest useful neuroinformatics databases (Table 1). We also present databases related to brain disorders such as Alzheimer's diseases, Parkinson's disease, schizophrenia, depression, and bipolar disorder (Table 2). In addition, we introduce several studies integrating databases of multi-level information to show potential application of neuroinformatics databases. Our review will aid researchers in using neuroinformatics databases in their research.

**Table 1.** Brain imaging, electrophysiology, connectivity, and genetic databases

Database	URL	Data type
Brain-development.org	<a href="http://www.brain-development.org/">http://www.brain-development.org/</a>	Structural images, functional images
Brainmuseum.org	<a href="http://www.brainmuseum.org/">http://www.brainmuseum.org/</a>	Structural images
Harvard whole brain atlas	<a href="http://www.med.harvard.edu/AANLIB/home.html">http://www.med.harvard.edu/AANLIB/home.html</a>	Structural images, functional images
BrainMap.org	<a href="http://brainmap.org/">http://brainmap.org/</a>	Structural images, functional images
BrainMaps.org	<a href="http://brainmaps.org/">http://brainmaps.org/</a>	Structural images, functional images
NeuroMorpho.org	<a href="http://neuromorpho.org/neuroMorpho/index.jsp">http://neuromorpho.org/neuroMorpho/index.jsp</a>	3D model images
CARMEN Project	<a href="http://www.carmen.org.uk/">http://www.carmen.org.uk/</a>	Neurophysiological data
CRNCS	<a href="http://crcns.org/">http://crcns.org/</a>	Neurophysiological data
Open Connectome Project	<a href="http://openconnectome.me/">http://openconnectome.me/</a>	Microconnectome, macroconnectome
CoCoMac	<a href="http://www.cocomac.org/">http://www.cocomac.org/</a>	Macroconnectome
BAMS	<a href="http://brancusi.usc.edu/bkms/">http://brancusi.usc.edu/bkms/</a>	Macroconnectome
UCLA-MCD	<a href="http://umcd.humanconnectomeproject.org/">http://umcd.humanconnectomeproject.org/</a>	Macroconnectome
Gene Expression Omnibus (GEO)	<a href="http://www.ncbi.nlm.nih.gov/geo/">http://www.ncbi.nlm.nih.gov/geo/</a>	Gene expression
ArrayExpress	<a href="http://www.ebi.ac.uk/arrayexpress/">http://www.ebi.ac.uk/arrayexpress/</a>	Gene expression
Allen Brain Atlas	<a href="http://www.brain-map.org/">http://www.brain-map.org/</a>	Gene expression, structural images



**Figure 1.** Taxonomy of data in the field of neuroscience according to their types of information.

**Table 2.** Neuroinformatics databases related to brain disease

Database	URL	Data type	Disease
ADNI	<a href="http://www.adni-info.org/">http://www.adni-info.org/</a>	Structural images, functional images	Alzheimer's disease
MaND	<a href="http://www.depressiondatabase.org/">www.depressiondatabase.org/</a>	Study information (MRI, CT)	Major depressive disorder
OASIS	<a href="http://www.oasis-brains.org/">www.oasis-brains.org/</a>	Structural images	Dementia
SMRI	<a href="https://www.stanleygenomics.org/">https://www.stanleygenomics.org/</a>	Gene expression, DNA sequencing	Schizophrenia, bipolar disorder, depression
Parkinson's disease database	<a href="http://www2.cancer.ucl.ac.uk/Parkinson_Db2/">http://www2.cancer.ucl.ac.uk/Parkinson_Db2/</a>	Gene expression	Parkinson's disease
PDGene	<a href="http://www.pdgene.org/">http://www.pdgene.org/</a>	Genetic association studies	Parkinson's disease
SZGene	<a href="http://www.szgene.org/">http://www.szgene.org/</a>	Genetic association studies	Schizophrenia
ALSGene	<a href="http://www.alsgene.org/">http://www.alsgene.org/</a>	Genetic association studies	Amyotrophic lateral sclerosis
MSGene	<a href="http://www.msgene.org/">http://www.msgene.org/</a>	Genetic association studies	Multiple sclerosis
Alzgene	<a href="http://www.alzgene.org/">http://www.alzgene.org/</a>	Genetic association studies	Alzheimer's disease
MDPD	<a href="http://datam.i2r.a-star.edu.sg/mdpd/index.php">http://datam.i2r.a-star.edu.sg/mdpd/index.php</a>	Genetic association studies	Parkinson's disease
NHANES	<a href="http://www.cdc.gov/nchs/nhanes.htm/">http://www.cdc.gov/nchs/nhanes.htm/</a>	Environmental factors	Depression, schizophrenia, etc.

## NEUROINFORMATICS DATABASES

### Brain imaging databases

Imaging techniques for a brain allow us to study functions and structure of brains. Current techniques in brain imaging are divided into mainly two different areas: structural imaging technique for studying the anatomy of the brain, for example, CT and MRI; functional imaging techniques for studying its functions or the connectivity of the brain, for example, fMRI, positron emission tomography (PET), and diffusion tensor imaging (DTI). Researchers used techniques according to the purpose of their study. Several useful databases stored information about studies and their structural and functional images. These databases are public and can be used for various uses such as validation of their experiment.

Brain-development.org<sup>5</sup> is a part of the research project of Imperial College London containing datasets of over 600 MRI images of healthy people utilizing T1 T2 and PD-weighted images, MRA images, DTI images, and Diffusion weighted images of normal and healthy human brains. This project also provides demographic information of each subject. To analyze brain images, tools for image registration toolkit (IRTK) are also available.

Brainmuseum.org<sup>6</sup> is a public repository providing images and information for well-preserved, dissected, and stained mammal brains. This database contains brain images over 100 different species of mammals including chimpanzees, monkeys, Florida manatee, big brown bat, American badger, ro-

dents, cow, and humans. This database is closely related to MSU brain diversity bank which is a repository containing both cell stain images and Fiber Stain images for each Sagittal section of the brain.

Harvard whole brain atlas<sup>7</sup> contains CT and MRI images for structural imaging and SPECT/PET images for functional imaging. This database consists of normal brain images and brain images with diseases and disorders such as cerebrovascular diseases, neoplastic diseases, degenerative diseases, and inflammatory or infectious diseases.

BrainMap.org<sup>8</sup> is a database collecting published functional and structural neuroimaging experiments with coordinate-based results (x, y, z) in Talairach and MNI space. This database shares neuroimaging results and enables meta-analysis for anatomy and function of human brain in healthy and diseased subject.

BrainMaps.org<sup>9</sup> is an interactive multi-resolution brain atlas of various species both primate, e.g. *Homo sapiens*, and non-primate, e.g. *C. auratus*. BrainMaps.org is now based on over 20 million pixels of Nissl stained, annotated, and scanned images of both primate and non-primate brains. It provides software for analyzing this database. It also contains connectivity maps, Connectomes.org<sup>10</sup>, of some species such as *C. Elegans*.

The images of small scale such as neuronal images are also interesting subjects for researchers to study the brain, and there are several techniques for small scale images such as Electron Microscopy (EM) or computer simulation. NeuroMorpho database<sup>11</sup> is the largest public repository of 3-D neuronal reconstruction. This database aims for researchers to facilitate data of neuronal scale. It contains 7,986 digitally reconstructed neurons obtained from WWW databases and peer-reviewed publications. This database can be used for analysis, visualization, and modeling of neuronal data.

### Electrophysiological databases

Since biological information is converted into electrical signals for a brain to interpret, researchers are interested in electrophysiology of a brain with recording techniques, e.g. EEG. However, there are not yet firm standards how electrophysiological data to be described and integrated due to its complexity and the diversity of it. In spite of these difficulties, some useful databases and tools are developed and are available now.

The Code Analysis, Modeling and Repository for E-Neuroscience (CARMEN) Project<sup>12</sup>, established in 1<sup>st</sup> October 2006, is aiming to provide a virtual electrophysiological laboratory on the web which allows us to share their programs and primary data types such as the neural activity recording including signals and series of images. In order to share information, one must sign up which allows one to access, upload, and curate data.

Collaborative Research in Computational Neuroscience

(CRNCS)<sup>13</sup> offers a forum of discussion over electrophysiological datasets as well as the sharing data. Data sets of this forum include physiological recordings from the sensory, i.e. both visual and auditory, and memory systems. Also, they include data dealing with the movement of the eye. Their data will contain description of experimental conditions such as paradigms, as well as identifying the species, surgical procedures, and its specific recording techniques such whether if they used the electrode type, clamp method etc. In order to download and share information, one should register on the site and login.

### Brain connectivity databases

To transfer information between anatomical structures of a brain, the brain is wired with transferring electrical signals. Common natures of brain wiring are axonal connections and synapses. But there are numerous connections between brain components of various scales stretching from a neuronal scale to a brain regional scale, and different technologies are needed to analyze each scale of connection. Connectome databases are divided to microconnectome and macroconnectome by data scale of connections.

Microconnectome is the connection that its scale includes single neuron and synapse. So it focuses on axonal connection between neuron and each synaptic property. Open Connectome Project<sup>14</sup> is a connectome database opened to anybody to view, manipulate, analyze, and contribute. It provides high-resolution neuroanatomical images and its ultimate goal is to reconstruct the complete 3D model of brain. It is a collaborated project between many laboratories and is funded by NIH. Open Connectome Project essentially provides all levels of data from micro to macro, but this database is focused on microconnectome database such as EM images.

Macroconnectome is the connectome with its scale including brain regions. It thus focuses on the strength of connections and the number of neuron between two brain regions. Macroconnectome data is usually obtained from fMRI, Diffusion MRI, and DTI. Macroconnectome databases are relatively more and are better organized than microconnectome databases. It is well-quantized and interpreted as 2-dimension matrix form known as 'connectivity matrix'. So it is easily used for data mining and informatics and can be used to assist other research.

Collection of Connectivity data on the Macaque brain (CoCoMac) provides records of known wiring of Macaque brain<sup>15</sup>. It provides connectivity matrix data of primate brains. Since connectivity matrix is well-quantized such as each cell of matrix has a number of neurons composing each connection, its data can be easily interpreted and its use is facilitated by data mining technique. CoCoMac also provides large amount of data and 3D analysis tool with mandatory registration.

Brain Architecture Management System (BAMS) is an online

database for information about neural circuitry of a mouse brain<sup>16</sup>. It provides connectivity matrix data of rat brain. Special features of BAMS are its atlas and parsing engine. BAMS has more than 5 brain atlases so it can be easily related with other neuroscience database. Also, BAMS provides web parsing engine which makes connectivity map containing brain regions and connection strength.

UCLA Multimodal Connectivity Database (UCLA-MCD) is an open human connectome database project<sup>17</sup>. Users can upload their connectome data with some description such as status of sample, used atlas and etc. And also, any user can download human connectome data from it. It consists of connectivity matrix, atlas map and coordinate map. Since it has x, y, z coordinate map, this database can reconstruct 3D network structure of human brain from each data with including engine.

### Databases of genetic information of brain

Gene expressions in various brain regions and various cell types are important information for understanding a relationship between regulation of molecular level and physiological level of brains. With recent development of high throughput technologies such as microarray, *in situ* hybridization, and RNA sequencing, the volume of genetic data has grown exponentially and researches have studied gene expression of brains. Gene Expression Omnibus (GEO)<sup>18</sup> and ArrayExpress<sup>19</sup> are representative repositories of publicly available gene expression data from National Center for Biotechnology Information (NCBI) and European Bioinformatics Institute (EBI). These databases stored vast amount of gene expression profiles from brains of various species, various brain regions, and brains with some disorders. In this review, we will specially focus some of useful and novel datasets containing genetic information of brains and neurons.

Allen Brain Atlas ([www.brain-map.org](http://www.brain-map.org)) is a growing collection of public resources integrating gene expression and neuroanatomical data<sup>20</sup>. This database firstly contains whole genome-scale collection of cellular gene expression profiles using *in situ* hybridization method. In addition, ABA also has launched Allen Developing Mouse Brain Atlas, Allen Human Brain Atlas. Allen Developing Mouse Brain Atlas is gene expression data in the mouse brain beginning with mid-gestation through to juvenile/young adult. The Allen Developing Mouse Brain Atlas shows temporal and spatial regulation of gene expression of mouse brain. Allen Human Brain Atlas provides gene expression from human whole brain regions. It is the first and unique multi-modal gene expression atlas of the human brain. The ABA shows heterogeneity in brain regions and finds region specific enriched genes, brain housekeeping genes and cell type specific enriched genes. Bohland et al.<sup>21</sup> reported spatial gene expression patterns of mouse brain are closely related with classically defined neuroanatomical structures of mouse brain

using ABA. They suggested that analyzing spatial gene expression patterns can make a connection between molecular level and higher level information about brain organization.

Recently, Kang et al.<sup>22</sup> reported spatio-temporal transcriptomes of human brain. In this research, they discussed not only spatial gene expression patterns but also temporal gene expression patterns and it is the first transcriptome study of many regions of the developing human brain with large number of samples. They explored the transcriptomes of 16 human brain regions, neocortex, amygdala, hippocampus, striatum, thalamus, and cerebellar cortex, for 15-periods from embryonic development to late adulthood. In this research, they found that transcriptomes of the human brain across regions and times are more different than sexes, ethnicities, or individuals.

The brain has multiple cell types and gene expressions can differ across cell types. However, there was a lack of methods to separate neural cell types. John D. Cahoy et al. used Fluorescent-Activated Cell Sorting (FACS) and microarray to create a transcriptome database for three main CNS neural cell types, astrocyte, neurons, and oligodendrocytes<sup>23</sup>. This database provides global characterization of gene expression levels by acutely separated astrocyte, neurons, and oligodendrocytes at various postnatal ages from postnatal day 1 (P1) to P30. Using this database, they found that thousands of genes enriched in specific cell types and they can provide new information for understanding of interactions between different cells, neural development, function, and brain disease.

### Databases related to brain disorders

There are many studies about various brain disorders: Alzheimer's disease, Parkinson's disease, Huntington's disease, depression, bipolar disorder, schizophrenia, and so on. Researchers have studied these brain disorders with various approaches such as imaging, microarray, sequencing, and so on and they created some useful databases that other researchers can use in their research.

Some databases collected imaging studies about a specific disorder. Alzheimer's Disease Neuroimaging Initiative (ADNI) tried to study the progress of Mild cognitive impairment and Alzheimer's disease, provide a large databases<sup>24</sup>. ADNI contains MRI, fMRI, DTI, and PET images, cerebral spinal fluid, and blood biomarkers to study progression of Alzheimer's disease. Using this database, researchers can validate their biomarkers obtained from MRI/PET imaging, blood tests, and tests of cerebrospinal fluid for Alzheimer's disease (AD) clinical trials and diagnosis.

Major Depressive Disorder Neuroimaging Database (MaND) provides an Excel spreadsheet file containing database and meta-analysis<sup>25</sup>. It contains information of 225 imaging studies of brain structures (using MRI and CT scans) of major depres-

sive disorder patients and control groups. It also contains the meta-analysis result of those studies.

Open Access Series of Imaging Studies (OASIS) is a database collecting cross-sectional MRI in young, middle aged, nondemented and demented older adults<sup>26</sup>. Cross sectional MRI data contains images of 416 subjects aged 18 to 96 and longitudinal MRI data in nondemented and demented older adults. All subjects are right-handed and both male and female. Some of subjects were diagnosed as AD and not demented. Longitudinal MRI data contains longitudinal collection of 150 subjects aged 60 to 96. In Longitudinal MRI datasets, 72 subjects were not demented throughout the study, 64 subjects were characterized as demented at initial image, and 14 subjects were not demented at initial image and subsequently characterized as demented through a longitudinal study.

There are many studies using gene expression profiles of postmortem brains stored in GEO<sup>18</sup>, ArrayExpress<sup>19</sup> and other repositories. Online Genomics Databases of The Stanley Medical Research Institute is one of the largest databases of depression, schizophrenia and bipolar disorder<sup>27</sup>. This database provides gene expression profiles using microarray for postmortem brain tissues (prefrontal cortex, temporal cortex, hippocampus, thalamus, and cerebellum) of depression, bipolar disorder, and schizophrenia patients and control groups.

The Parkinson's Disease Database is a database for users to use data collected from gene expression data of Parkinson's diseases from ArrayExpress<sup>28</sup>. This database provides curated and re-annotated datasets which can be queried for individual gene expression profiles under diverse biological and clinical conditions and across various organisms.

To find causes and their effect of brains disorders, brain diseases have been studied by using gene expression and genomic variation with association study. Many of those studies could not find significant and reproducible biomarkers. However, there are some databases containing genetic information reliably related to a specific brain disorder: PDGene<sup>29</sup>, SZGene<sup>30</sup>, ALSGene<sup>31</sup>, MSGene<sup>32</sup>, and Alzgene<sup>33</sup> are databases for providing Parkinson's disease, schizophrenia, amyotrophic lateral sclerosis, multiple sclerosis, and Alzheimer's disease genetic association studies and meta-analyses available for all eligible polymorphism with sufficient data; Mutation Database for Parkinson's Disease (MDPD) provides not only an integrated genetic information resource for Parkinson's disease but also each genetic substitution and its resulting impact with its reference<sup>34</sup>.

Many brain disorders are caused by both genetic and environmental factor. Thus, Environment Wide Association Study (EWAS) using environmental factors with brain disorders can be conducted for brain disorders<sup>35</sup>. National Health and Nutrition Examination Survey (NHANES)<sup>36</sup>, which provides studies

designed to assess the health and nutritional status of adults and children in the United States, contains information of samples about not only environmental factors but also clinical status of mental illness (e.g. depression) and those information can be used for surveying brain disorders.

### Integration of neuroinformatics databases

With increasing amount of data of brain generated from various technologies, neuroscientists have made an effort to not only integrate and manage data efficiently but also make a “gold standard” of ontologies and nomenclatures for brains. The Neuroscience Information Framework (NIF)<sup>37</sup>, as an example of this, was launched in 2005. It is a project highly dependent upon the Neuroscience Lexicon (Neurolex)<sup>38</sup>, a dynamic lexicon of neuroscience terms. NIF provides a framework for researchers to discover and access to a large amount of public neuroscience data and tools easily.

Neuronames is comprehensive hierarchical nomenclatures specialized in Brain and Spinal cords of Human, Mouse, Rat and Macaque<sup>39</sup>. It displays suitable neuroanatomical ontology and vocabularies to refine neuroinformatics data-mining. It provides information more than 15,000 neuroanatomical terms, defines a brain in terms of 550 primary structures, and shows hierarchical relationship between each structures and neuroanatomical terms.

### MULTI-LEVEL ANALYSIS COMBINING DIFFERENT TYPES OF DATABASES

Studies using data obtained from only one technology or one domain have its own limitations. For example, different gene expressions from postmortem brain studies of mental illness can be caused by not only effects of mental illness but also by treatments of the brain disorder. Thus, some researchers have tried to study brains by combining different types of data for analyzing the brain.

Leon French et al.<sup>40</sup> studied the relationship between gene expression and brain connectivity of an adult rodent brain. They combined gene expression signatures of mouse brains from Allen Brain Atlas<sup>20</sup> with connectivity of mouse brain from the Brain Architecture Management System<sup>16</sup>. Their results showed that gene expression signature of mouse brain regions are significantly related to connectivity of mouse brain regions. They also found a set of genes that are closely correlated with neuroanatomical connectivity. This is the first research showing a possibility of combining gene expression data with macroconnectome data.

Lior Wolf et al. also studied the link between gene expressions of mouse brain regions and neural connectivity patterns of mouse brain regions<sup>41</sup>. They also used Allen Brain Atlas and

Brain Architecture Management System. They found gene expression levels of mouse brain regions can highly predict the connectivity of mouse brain regions-this result is closely related to the previous study<sup>40</sup>. They also identified a list of genes that may play an important role in the connectivity of mouse brain regions. A part of genes of the list are known as associated to axon guidance, axonogenesis, and several brain disorders such as schizophrenia, autism, major depression, and bipolar disorders based on literatures. This information can give us to novel inference in relationship between the gene expression changes of brain disorders and the connectivity of brain disorders.

Unlike most of studies of connectivity of Alzheimer’s disease using fMRI experiment, Byungkyu Park et al. combined gene expression data with diffusion tensor images of human brain to study Alzheimer’s disease<sup>42</sup>. They modeled interactions of Alzheimer-related genes from the fiber pathway by using microarray data of Allen Human Brain Atlas and diffusion tensor images of Allen Brain Atlas. Although their research have some limitations owing to data they used, their approach, integrating gene expression data with DTI data, is novel and can be used again when more suitable data is available.

### CONCLUSION AND PROSPECTS

In this review, we have reviewed neuroinformatics databases according to their information types, and we represent several useful neuroinformatics databases for each type of information. We also introduce several studies that use databases of different information types. These databases and related studies give a new way to study the brain.

Large amounts of data for neuroscience fields are already available, and additionally, we expect that more information about neuroscience, obtained from various technologies, will be available to researchers in the future. Therefore, researchers will have increased opportunities to study the brain from new perspectives using various databases. Researchers can use databases for their *in silico* experiments, validation of their results, integrating or combining different types of databases, and other related research tasks. Such novel uses of databases will give us new insights and novel findings in the field of neuroscience.

### ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea funded by the Korea Government, Ministry of Education, Science and Technology through the WCU (World Class University) program (R32-2008-000-10218-0) and Basic Research Laboratory grant (2009-0086964). It was also supported by the KISTI (Korea Institute of Science and Technology Information).

## REFERENCES

- Kotter, R. (2001). Neuroscience databases: tools for exploring brain structure-function relationships. *Philos Trans R Soc Lond B Biol Sci* 356, 1111-1120.
- Yarkoni, T., Poldrack, R.A., Van Essen, D.C., and Wager, T.D. (2010). Cognitive neuroscience 2.0: building a cumulative science of human brain function. *Trends Cogn Sci* 14, 489-496.
- French, L., and Pavlidis, P. (2007). Informatics in neuroscience. *Brief Bioinform* 8, 446-456.
- Akil, H., Martone, M.E., and Van Essen, D.C. (2011). Challenges and opportunities in mining neuroscience data. *Science* 331, 708-712.
- <http://www.brain-development.org/>.
- <http://www.brainmuseum.org/index.html>.
- <http://www.med.harvard.edu/AANLIB/home.html>.
- Laird, A.R., Lancaster, J.L., and Fox, P.T. (2005). BrainMap: the social evolution of a human brain mapping database. *Neuroinformatics* 3, 65-78.
- Mikula, S., Trotts, I., Stone, J.M., and Jones, E.G. (2007). Internet-enabled high-resolution brain mapping and virtual microscopy. *Neuroimage* 35, 9-15.
- <http://connectomes.org/>.
- Ascoli, G.A., Donohue, D.E., and Halavi, M. (2007). NeuroMorpho.Org: a central resource for neuronal morphologies. *J Neurosci* 27, 9247-9251.
- <http://www.carmen.org.uk/>.
- <http://crcns.org/>.
- <http://openconnectomeproject.org/>.
- Stephan, K.E., Kamper, L., Bozkurt, A., Burns, G.A., Young, M.P., and Kotter, R. (2001). Advanced database methodology for the Collation of Connectivity data on the Macaque brain (CoCoMac). *Philos Trans R Soc Lond B Biol Sci* 356, 1159-1186.
- Bota, M., Dong, H.W., and Swanson, L.W. (2005). Brain architecture management system. *Neuroinformatics* 3, 15-48.
- Jesse Brown, J.R., and Susan Bookheimer. (2011). Networks in the Cloud: Web-based Neuroimaging Brain Network Analysis and Data Sharing. Organization for Human Brain Mapping, 2011, Quebec City, Quebec.
- Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30, 207-210.
- Rocca-Serra, P., Brazma, A., Parkinson, H., Sarkans, U., Shojatalab, M., Contrino, S., Vilo, J., Abeygunawardena, N., Mukherjee, G., Holloway, E., et al. (2003). ArrayExpress: a public database of gene expression data at EBI. *C R Biol* 326, 1075-1078.
- Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168-176.
- Bohland, J.W., Bokil, H., Pathak, S.D., Lee, C.K., Ng, L., Lau, C., Kuan, C., Hawrylycz, M., and Mitra, P.P. (2010). Clustering of spatial gene expression patterns in the mouse brain and comparison with classical neuroanatomy. *Methods* 50, 105-112.
- Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al. (2011). Spatio-temporal transcriptome of the human brain. *Nature* 478, 483-489.
- Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A., et al. (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci* 28, 264-278.
- Jack, C.R. Jr., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., J, L.W., Ward, C., et al. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging* 27, 685-691.
- Kempton, M.J., Salvador, Z., Munafo, M.R., Geddes, J.R., Simmons, A., Frangou, S., and Williams, S.C. (2011). Structural neuroimaging studies in major depressive disorder. Meta-analysis and comparison with bipolar disorder. *Arch Gen Psychiatry* 68, 675-690.
- Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., and Buckner, R.L. (2007). Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J Cogn Neurosci* 19, 1498-1507.
- Torrey, E.F., Webster, M., Knable, M., Johnston, N., and Yolken, R.H. (2000). The stanley foundation brain collection and neuropathology consortium. *Schizophr Res* 44, 151-155.
- Taccioli, C., Tegner, J., Maselli, V., Gomez-Cabrero, D., Altobelli, G., Emmett, W., Lescai, F., Gustincich, S., and Stupka, E. (2011). ParkDB: a Parkinson's disease gene expression database. *Database (Oxford)* 2011, bar007.
- Lill, C.M., Roehr, J.T., McQueen, M.B., Kavvoura, F.K., Bagade, S., Schjeide, B.M., Schjeide, L.M., Meissner, E., Zauft, U., Allen, N.C., et al. (2012). Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. *PLoS Genet* 8, e1002548.
- Allen, N.C., Bagade, S., McQueen, M.B., Ioannidis, J.P., Kavvoura, F.K., Khoury, M.J., Tanzi, R.E., and Bertram, L. (2008). Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* 40, 827-834.
- Lill, C.M., Abel, O., Bertram, L., and Al-Chalabi, A. (2011). Keeping up with genetic discoveries in amyotrophic lateral sclerosis: the ALSod and ALSGene databases. *Amyotroph Lateral Scler* 12, 238-249.
- Lill, C.M., R.J., McQueen, M.B., Bagade, S., Schjeide, B.M., Zipp, F., and Bertram, L. The MSGene Database. Alzheimer Research Forum.
- Bertram, L., McQueen, M.B., Mullin, K., Blacker, D., and Tanzi, R.E. (2007). Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* 39, 17-23.
- Tang, S., Zhang, Z., Kavitha, G., Tan, E.K., and Ng, S.K. (2009). MDPD: an integrated genetic information resource for Parkinson's disease. *Nucleic Acids Res* 37, D858-862.
- Patel, C.J., Bhattacharya, J., and Butte, A.J. (2010). An Environment-Wide Association Study (EWAS) on type 2 diabetes mellitus. *PLoS One* 5, e10746.
- <http://www.cdc.gov/nchs/NHANES.htm>.
- Gardner, D., Akil, H., Ascoli, G.A., Bowden, D.M., Bug, W., Donohue, D.E., Goldberg, D.H., Grafstein, B., Grethe, J.S., Gupta, A., et al. (2008). The neuroscience information framework: a data and knowledge environment for neuroscience. *Neuroinformatics* 6, 149-160.
- [http://neurolex.org/wiki/Main\\_Page](http://neurolex.org/wiki/Main_Page).

39. Bowden, D.M., Song, E., Kosheleva, J., and Dubach, M.F. (2012). NeuroNames: an ontology for the BrainInfo portal to neuroscience on the web. *Neuroinformatics* 10, 97-114.
40. French, L., and Pavlidis, P. (2011). Relationships between gene expression and brain wiring in the adult rodent brain. *PLoS Comput Biol* 7, e1001049.
41. Wolf, L., Goldberg, C., Manor, N., Sharan, R., and Ruppin, E. (2011). Gene expression in the rodent brain is associated with its regional connectivity. *PLoS Comput Biol* 7, e1002040.
42. Park, B., Lee, W., and Han, K. (2012). Modeling the interactions of Alzheimer-related genes from the whole brain microarray data and diffusion tensor images of human brain. *BMC Bioinformatics* 13 Suppl 7, S10.