Neural Networks 21 (2008) 1056–1069

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

2008 Special Issue

Cerebellar development transcriptome database (CDT-DB): Profiling of spatio-temporal gene expression during the postnatal development of mouse cerebellum

Akira Sato [a,b], Yukiko Sekine [a], Chihiro Saruta [a], Hirozumi Nishibe [a], Noriyuki Morita [a,c], Yumi Sato [a], Tetsushi Sadakata [a], Yo Shinoda [a], Toshio Kojima [b], Teiichi Furuichi [a,*]

[a] *Laboratory of Molecular Neurogenesis, RIKEN Brain Science Institute, Wako, Saitama 351-0198, Japan*

[b] *Synthetic Biology Team, RIKEN Advanced Science Institute, Yokohama 230-0045, Japan*

[c] *Mimasaka Junior College, Tsuyama, Okayama 708-8511, Japan*

ARTICLE INFO

Article history:
Received 31 October 2007
Received in revised form
8 April 2008
Accepted 28 May 2008

Keywords:
Cerebellum
Postnatal development
Gene expression
in situ hybridization
Microarray

ABSTRACT

A large amount of genetic information is devoted to brain development and functioning. The neural circuit of the mouse cerebellum develops through a series of cellular and morphological events (including neuronal proliferation and migration, axogenesis, dendritogenesis, synaptogenesis and myelination) all within three weeks of birth. All of these events are controlled by specific gene groups, whose temporal and spatial expression profiles must be encoded in the genome. To understand the genetic basis underlying cerebellar circuit development, we analyzed gene expression (transcriptome) during the developmental stages on a genome-wide basis. Spatio-temporal gene expression data were collected using *in situ* hybridization for spatial (cellular and regional) resolution and fluorescence differential display, GeneChip, microarray and RT-PCR for temporal (developmental time series) resolution, and were annotated using Gene Ontology (controlled terminology for genes and gene products) and anatomical context (cerebellar cell types and circuit structures). The annotated experimental data were integrated into a knowledge resource database, the Cerebellar Development Transcriptome Database (CDT-DB http://www.cdtdb.brain.riken.jp), with seamless links to the relevant information at various bioinformatics database websites. The CDT-DB not only provides a unique informatics tool for mining both spatial and temporal pattern information on gene expression in developing mouse brains, but also opens up opportunities to elucidate the transcriptome for cerebellar development.

1. Introduction

The brain is the ultimate genetic system to which a large number of genes are devoted (Sutcliffe, 1988). Data collection at the whole genome level is crucial for comprehensive understanding of the molecular mechanisms underpinning higher-order brain structures and functions of normal brain as well as those of brain diseases in the post-genomic sequencing era (Geschwind, 2000; Hatten & Heintz, 2005). The application of state-of-the-art technologies, including DNA microarrays, on a genome-wide basis, to various aspects of neuroscience, has led to the accumulation of huge amounts of gene expression data; for example, in developing brains (Diaz et al., 2002; Jensen et al., 2004; Kagami & Furuichi, 2001; Matsuki, Hori, & Furuichi, 2005; Mody et al., 2001; Pollard et al., 2006), in neurons in response to neuronal activities (Newton et al., 2003; Tropea et al., 2006), in neuron types (Lein, Zhao, & Gage, 2004; Sugino et al., 2006; Toledo-Rodriguez et al., 2004) and in ageing brains (Lee, Weidruch, & Prolla, 2000; Lu et al., 2004). Most recently, the advent of the neuroinformatics era has enabled us to systematize such large datasets of brain gene expression information in a database, and to create a platform for sharing and mining data; for example, the *in situ* hybridization (ISH) brain atlas databases GenePaint (Carson et al., 2005; Visel et al., 2004, 2007), Brain Gene Expression Map (BGEM) (Magdaleno et al., 2006) and the Allen Brain Atlas (ABA) (Lein et al., 2007), a database for transcription factor gene expression termed the Functional Genomic Atlas of the Mouse Brain (Gray et al., 2004) and a database for nuclear receptor gene expression termed MousePat (Gofflot et al., 2007).

We have focused on the neural circuit of the mouse cerebellum, which has a layered structure in which there are five major neuronal cell types (Purkinje cells, granule cells, stellate cells,

* Corresponding address: Laboratory for Molecular Neurogenesis, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. Tel.: +81 48 467 5906; fax: +81 48 467 6079.
E-mail address: tfuruichi@brain.riken.jp (T. Furuichi).

0893-6080/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2008.05.004

basket cells, and Golgi cells), two major inputs (mossy fibers and climbing fibers), and a single output (Purkinje cell axons) (Ito, 2006). The mouse cerebellar circuit develops through a series of cellular and morphological events (cell proliferation and migration, dendritogenesis and axogenesis, synaptogenesis, myelination, foliation and fissurization) that all occur within the first three weeks of life, and which appear to be genetically coded because of the presence of many cerebellar mutant mice available (Goldowitz & Hamre, 1998). Although the hereditary plan for cerebellar development remains incompletely understood, all of the transcription events (transcriptome) during this period are known to be more complex than those in other mouse organs (Kaplan, 1982, 1987). To understand the genetic basis of cerebellar development we carried out genome-wide profiling of spatial (cellular and regional) and temporal (developmental time series) gene expression patterns in developing mouse brains and systematized them in a database called the Cerebellar Development Transcriptome Database (CDT-DB) (Sato et al., 2004). The CDT-DB is a useful knowledge resource database to share and mine the annotated experimental data on the cerebellar transcriptome.

2. Materials and methods

2.1. Genome-wide analyses of spatio-temporal gene expression

ICR and C57B6/J mice (Nihon SLC, Hamamatsu, Japan) were used according to the RIKEN guidelines for animal research. We analyzed spatio-temporal gene expression patterns during the various developmental stages on a genome-wide basis (Sato et al., 2004). Briefly, to search for genes that are differentially expressed in the mouse cerebellum during eight developmental stages (embryonic day [E]18, postnatal day [P]0, P3, P7, P12. P15, P21 and P56), we employed a fluorescence differential display (FDD) approach. Among a total of approximately 12,000 bands displayed using 300 primer sets (100 arbitrary 10mers and 3 kinds of anchor primers), 83.1% exhibited almost constant patterns or only slight changes in their intensities, throughout the postnatal stages, whereas the remaining bands (16.9%, about 2,000 bands) had different intensities during different developmental stages. By cloning and sequencing these bands, we succeeded in obtaining 2194 non-redundant clones derived from transcripts expressed during cerebellar development. The temporal (developmental time series) expression profiles of cerebellar genes were verified by semi-quantitative RT-PCR or custom-made cDNA microarray (Sato et al., 2004). FDD and RT-PCR gel images were digitized using a fluorescence image scanner (Molecular Imager FX, Rio-Rad Laboratories). We also performed genome-wide analysis of temporal gene expression at E18, P7, P14, P21 and P56 by utilizing Affymetrix GeneChip Mu11K arrays (12,654 transcripts) (Kagami & Furuichi, 2001) and Mouse Genome 430 2.0 arrays (39,000 transcripts). We identified about 7,100 cerebellar development (CD) genes by statistical analyses (ANOVA) as described previously (Matsuki et al., 2005). In total, we identified ∼9000 CD genes showing differential expression patterns during the postnatal development of mouse cerebellum (Table 1).

Spatial (cellular and regional) expression profile information was collected by *in situ* hybridization (ISH) histochemistry using riboprobes labeled with digoxigenin (DIG)-UTP (Roche Diagnostics) and parasagittal paraffin sections (6 μm in thickness) of mouse brains at P7 and P21 as described previously (Sadakata, Washida, & Furuichi, 2007; Shiraishi et al., 1999). ISH reactions were carried out manually or semi-automatically (Freedom EVO GenePaint, TECAN). Images of hybridized brain sections were digitized using charge-coupled device camera-equipped microscopes or a digital slide scanner (NanoZoomer Digital Pathology, Hamamatsu Photonics K.K.).

Table 1
CDT-DB datasets

Experiments	Developmental stages analyzed	Amount of data (as of April, 2008)
FDD/RT-PCR	8 stages (E18, P0, P3, P7, P12, P15, P21, P56)	1769 gel images
Microarray	8 stages (E18, P0, P3, P7, P12, P15, P21, P56)	209 graphs
GeneChip	5 stages (E18, P7, P14, P21, P56)	7096 graphs
ISH	2 stages (P7, P21)	10,182 images (1697 genes × 2 stages × 3 images)
Tissue RT-PCR	1 stage (either P7 or P21)	1755 gel images

The CDT-DB includes the gene and hyperlink information on 8807 genes (as of April, 2008).

The tissue specificity of CD gene expression was determined by RT-PCR analysis of eight different tissues (brain, thymus, lung, heart, liver, spleen, kidney, and testis) obtained at either P7 or P21, as described previously (Matsuki et al., 2005).

2.2. Generation of the cerebellar development transcriptome database (CDT-DB)

We developed the CDT-DB (version 3.1) (http://www.cdtdb.brain.riken.jp) by systematizing all gene expression data incorporating temporal information throughout the developmental stages and spatial information defined by brain regions (cerebellar cell types and brain structures). The CD genes registered in the CDT-DB were given gene names according to the nomenclatures used in the National Center for Biotechnology Information (NCBI)-Entrez Gene (http://www.ncbi.nlm.nih.gov/) and the Jackson laboratory (JAX)-Mouse Genome Informatics (MGI) (http://www.informatics.jax.org/). The CD genes were also annotated using common terminology for biological processes, cellular components and molecular functions defined by Open Biomedical Ontologies (OBO)-Gene Ontology (http://www.geneontology.org/) (The Gene Ontology Consortium, 2000). The annotated gene expression datasets accompanying these metadata are registered in the CDT-DB in a searchable format with which multiple queries about annotated gene expression patterns (developmental time series specificity, cerebellar cell type and layer specificity, brain region specificity and tissue distribution specificity) and keywords (for example, gene names, protein and cellular properties) can be executed. For easy access to additional information, each gene was directly linked to the corresponding websites of the relevant public bioinformatics databases: NCBI-Nucleotide, -Entrez -Gene, -UniGene, -GEO, -OMIM, -PubMed; JAX-MGI; Sanger Institute/EMBL European Bioinformatics Institute-Ensembl (http://www.ensembl.org/Mus_musculus/index.html); Kyoto University-KEGG (http://www.genome.jp/kegg/).

3. Results

3.1. Acquisition and systematization of spatial and temporal gene expression profiles in the CDT-DB

To understand the genetic basis of postnatal development of the mouse cerebellum, we investigated the spatial (cellular and regional) and temporal (developmental time series) expression patterns of cerebellar development (CD) genes during the developmental stages using genome-wide approaches. Spatio-temporal gene expression profile data were systematized into the Cerebellar Development Transcriptome Database (CDT-DB).