Gene expression profiling of cerebellar development with high-throughput functional analysis

Sakae Saito,1 Kimi Honma,2 Hiroko Kita-Matsuo, Takahiro Ochiya, and Kikuya Kato. Gene expression profiling of cerebellar development with high-throughput functional analysis. Physiol Genomics 22: 8–13, 2005. First published March 29, 2005; 10.1152/physiolgenomics.00142.2004.—We measured the expression levels of 450 genes during mouse postnatal cerebellar development by quantitative PCR using RNA purified from layers of the cerebellar cortex. Principal component analysis of the data matrix demonstrated that the first and second components corresponded to general levels of gene expression and gene expression patterns, respectively. We introduced 288 of the 450 genes into PC12 cells using a high-throughput transfection assay based on atelocollagen and determined the ability of each gene to promote neurite outgrowth or cell proliferation. Five genes induced neurite outgrowth, and seven genes enhanced proliferation. Evaluation of the functional data and gene expression patterns showed that none of these genes exhibited elevated expression at maturation, suggesting that genes characteristic of mature neurons are not likely to participate in neuronal development. These results demonstrate that functional data can facilitate interpretation of expression profiles and identification of new molecules that participate in biological processes.

Adaptor-tagged competitive polymerase chain reaction; cell transfection array

DEVELOPMENTAL PROCESSES are executed by intrinsic programs encoded in the genome with modulation by extrinsic factors. Because the execution of such programs is, for the most part, dependent on gene expression, it should be possible to directly elucidate the programs by analyzing gene expression. Conventional approaches such as genetic studies based on mutant analysis have so far identified genes responsible for particular developmental processes. Many of these genes are transcription factors such as homeotic genes; however, many of the molecular events downstream of the transcription factors still remain to be elucidated.

Although gene expression profiling is expected to shed light upon downstream molecular events, the lack of valid information regarding gene functions makes it difficult to interpret the expression data. Complementary methods of analysis include bioinformatics of the scientific literature (24) and proteomics (7, 10). However, direct analysis of gene function is more desirable. Recently, two techniques for high-throughput cell transfection have been described: reverse transfection (33) and a cell transfection array using atelocollagen (16). These two techniques enable large-scale experiments to study the overexpression or suppression of genes.

In this report, we measured the expression levels of 450 cerebellar genes using RNA purified from each layer of the cerebellar cortex. In parallel, we analyzed 288 of the 450 genes by cell transfection array using PC12 cells. The cerebellar cortex is a classic model system in developmental biology (2). The postnatal development of the cerebellum includes most of the events characteristic of nervous system development including axon elongation, cell proliferation, cell migration, synapse formation, and apoptosis. PC12 cells are also a classic model system of neuronal differentiation (13). This cell line is derived from the rat pheochromocytoma, and the cells extend neurites under appropriate stimulation, such as treatment with nerve growth factor (NGF). Combined use of these classical systems has allowed the description of both the transcriptional and functional aspects of nervous system development, demonstrating the correlation between gene expression and function.

MATERIALS AND METHODS

Genes subjected to analysis. We have previously performed gene expression profiling of postnatal cerebellar development using the whole cerebellum (22). This experiment analyzed 1,869 genes selected by descending order of abundance from our expressed sequence tag collection (21). Of these 1,869 genes, 768 genes, including 384 genes whose expression changed during development and 384 genes that were previously reported to be involved in brain development, were selected for the present study (22). Because RNA quantitation using samples recovered by laser capture microdissection (LCM) is demanding, we succeeded in analyzing only 450 of the 768 genes. The full-length cDNAs of 288 of these genes were subcloned into expression vectors. It should be noted that the selected genes may be biased toward abundance but should not be biased to other factors such as gene function or annotation.

LCM. With the use of a PixCell LCM microscope (Arcturus Engineering), the granule cell layers corresponding to the distinct areas described in the main text were dissected from hematoxylin-eosin-stained 12-μm frozen tissue sections of 4-, 8-, 12-, and 21-day-old mice. Three mice were used for LCM at each age. Experimental details are described in protocols for microdissection and RNA preparation on the NIH LCM web page (http://dir.nichd.nih.gov/lcm/lcm.htm). Approximately 100 ng of total RNA were obtained from 1,000 transfer shots of the external granular layer (EGL) and the internal granular layer (IGL).

Adaptor-tagged competitive PCR and data analysis. The adaptor-tagged competitive PCR (ATAC-PCR) protocol is essentially the same as described previously (18, 22, 30). Total RNA from the
cerebrum at postnatal week 6 was used as a standard. The relative expression levels versus the 6-wk cerebrum were calculated. Among the 768 genes subjected to ATAC-PCR, 450 genes had no missing values. This data matrix was normalized as previously described (25). In brief, the data were first divided by the median of the samples and converted into a logarithmic scale. The cutoff values were set at 20 and 0.05 before the logarithmic conversion. Principal component analysis (PCA) was performed using GeneMath 2.0 (AppliedMath).

The chance of biased distribution of functional genes (P) was calculated using the hypergeometric distribution, described as follows:

\[
P = 1 - \sum_{x=0}^{n-1} \left( \frac{mC_x}{mC_{n-x}} \right) \left( \frac{N-mC_x}{N-mC_{n-x}} \right)^m \]

where N is the number of genes in the total population, M is the number of functional genes in the total population, n is the number of genes in a particular group, and m is the number of functional genes in the group.

Culture and differentiation of PC12 cells. PC12 cells, obtained from the American Type Culture Collection, were grown in DMEM with 10% charcoal-stripped horse serum (Biofluid), 5% diaz Yield FLABTR), 10 mM HEPES (pH 7.4), 50 mg/ml streptomycin, and 50 U/ml penicillin. To differentiate PC12 cells, 10% charcoal-stripped horse serum (Biofluid), 5% dialyzed FBS from the American Type Culture Collection, were grown in DMEM Medium was replaced every 3 days. On day 7, neurite extension was clearly observed in ≥90% cells. Passage 24 clones were used in all the experiments.

Atelocollagen complex and gene transfer. The full-length open reading frames of 288 genes were amplified from 8-day-old mouse cerebellar cDNA. Because the 8-day-old mouse cerebellum contains both the EGL and the IGL, the selected genes were well amplified and subcloned into pDEST26 with Gateway Technology (Invitrogen). Each purified pDEST26 plasmid and a pCMV-EGFP plasmid were mixed with Atelocollagen (Koken) and dissolved in PBS buffer (Sigma)-coated 25-cm² culture flasks. The following day, cells were incubated in PC12 medium containing 100 ng/ml of 5.5% NGF (Collaborative Biomedical Products) for 1 wk. Medium was replaced every 3 days. On day 7, neurite extension was clearly observed in ≥90% cells. Passage 24 clones were used in all the experiments.

Cell transfection array. Postnatal cerebellar development includes cell proliferation and axon elongation. Genes involved in these processes are expected to induce phenotypic changes in PC12 cells; 288 of the 450 cerebellar genes that were subjected to gene expression analysis were also transfected into PC12 cells. To enable high-throughput functional analysis, we employed a cell transfection array using atelocollagen. Atelocollagen is a biocompatible natural polymer that is free of telopeptides and is obtained by pepsin treatment of highly purified type I collagen. Atelocollagen can be used to condense and deliver DNA, antisense oligodeoxynucleotides, virus vectors, and siRNAs into cells (27, 28).

Cells transfected with five genes, including neuronatin (Nnat), retinoblastoma 1 (Rb1), potassium channel subfamily K member 2 (TREK), ring finger protein 13 (Rnf13), and eukaryotic translation initiation factor 4A1 (eIF4A1), exhibited typical cell proliferation and axon elongation during the second week (2). The developing granule cell layer typically consists of two sublayers: the EGL and the IGL. Furthermore, the EGL and IGL typically consist of two areas. The proliferating granule cells are located in the superficial area of the EGL (represented as EGLa in Fig. 1), whereas postmitotic and migrating granule cells are located in the internal area of the EGL (EGLb). These granule cells migrate to the IGL, extending parallel fibers. The morphological changes in these cells are complete by the third week (2) and regarded as the mature cerebellum. The 21-day cerebellum therefore only consists of IGLb and has the structure as the same as that of an adult. Using LCM, we dissected the granule cell layers from 4-, 8-, 12-, and 21-day-old mice, collecting a total of 12 samples (4-day EGLa, 4-day EGLb, 4-day IGL, 8-day EGLa, 8-day EGLb, 8-day IGLa, 8-day IGLb, 12-day EGLa, 12-day IGLa, 21-day IGLa, and 21-day IGLb). It should be noted that we were not able to discriminate among the four layers in some samples because the fraction of EGL decreases and the fraction of IGL increases as development proceeds. The relative expression levels of the 450 genes were obtained using ATAC-PCR. ATAC-PCR is an advanced version of quantitative PCR and has been established as a mature expression profiling technique in several studies (30). The data matrix consisting of 450 genes × 12 samples is supplied as supplementary data (http://physiolgenomics.physiology.org/cgi/content/full/00142.2004/DC1).

We applied several unsupervised feature extraction methods to the data matrix. PCA, a statistical technique to summarize a multivariate data set using a few components, was successful. This analysis demonstrated that the first and second components corresponded to general levels of gene expression and to gene expression patterns, respectively (Fig. 1, A and B). Sorting of genes by the factor scores of the second component ordered the genes from those elevated during development to those elevated at maturation. In contrast to the prominent temporal changes, differences between the layers were obscure.

**RESULTS**

Gene expression profiling using RNA purified from granule cell layers. In rodents, the development of the cerebellar cortex begins just after birth; the peak of granule cell proliferation occurs during the first week, and the peak of cell migration and axon elongation occurs during the second week (2). The developing granule cell layer typically consists of two sublayers.
derived growth factor (Hdgf), mitogen-activated protein kinase 6 (Mapk6/Erk3), and Finkel-Biskis-Reilly murine sarcoma virus (FBR-MuSV) ubiquitously expressed gene (Fau/MNSF/H9252).

**Correlation between gene expression patterns and functions.** To explore the relationship between gene expression and function, we aligned genes in the functional categories identified by the cell transfection array to the gene expression patterns sorted by the second component (Fig. 1B). The genes that functioned in neurite outgrowth and cell proliferation had characteristic distribution; none of these genes exhibited elevated expression at maturation. Figure 1B includes genes that were not analyzed in transfection experiments. In Fig. 3, the expression patterns of the 288 genes analyzed by transfection are shown in a more quantitative manner: the quantitative differences between the expression levels in 8-day and 12-day/3-week cerebella are plotted by the order of factor scores of the second component. There are no genes that functioned either in neurite outgrowth or cell proliferation distributed in the region where the indicator exhibited negative values.

We calculated the chance probability of the biased gene distribution. Because Nnat-3 is located nearest to the right end, we calculated the probability that all genes of a category are located on the left side of the Nnat-3 position. The P values for all genes, cell proliferation genes, and neurite outgrowth genes are 0.002, 0.028, and 0.267, respectively.

**DISCUSSION**

Identification of transcriptional activation or deactivation of functional gene classes would facilitate the understanding of molecular mechanisms underlying biological processes. Exploration of the correlations between gene expression patterns and functions has therefore been a central issue of transcriptome analysis. However, these analyses have usually been performed by subjective interpretation of individual scientists without valid evidence. Such arguments have often been based on previously published scientific literature rather than on information derived from the analysis of experimental systems. Experimental data regarding gene functions should provide more valid information. In this report, we described the simultaneous analysis of gene expression and function to test whether an expression-function correlation could be experimentally identified. One important finding herein is that none of the genes in the functional groups exhibited elevated expression at maturation. This result agrees with the gene expression-function correlation previously identified from the literature (20, 22). Taken together, these data imply that molecules that are functionally active in mature neurons are not likely to be active participants in the developmental process. Although confirmation with a larger number of genes is needed, our approach will provide a rigid basis for future studies of gene expression-function correlations.

The rat pheochromocytoma cell line PC12 is a well-characterized model for the study of neuron-like differentiation and signal transduction. There are many cases where genes involved in cerebellar development are induced phenotypic changes in PC12 cells. In response to expression of Wnt-1, a secreted signaling factor required for development of mammalian cerebellum, PC12 cells adopt a flat morphology reminiscent of epithelial cells, form extensive cell-cell contacts, and become refractory to neuronal differentiation induced by NGF (31). Moreover, NCAM, a member of the immunoglobulin...
superfamily implicated in cerebellar granule cell migration, has
stimulated neurite outgrowth from PC12 cells (8). Nnat protein
is expressed much more highly in fetal and neonatal brains than
in the adult brain (9) and is expressed in a segment-specific
pattern during early hindbrain development (32), suggesting
involvement in nervous system development. Rb1 is known to
be a tumor suppressor gene and is hypophosphorylated in PC12
cells by NGF through the Ras signaling pathway (19). Hypo-
phosphorylated Rb1 is essential for neurite outgrowth, as was
demonstrated by inhibition with a monoclonal antibody against
hypophosphorylated Rb (19). The related protein Rb2/p130
was reported to induce neurite outgrowth in the absence of
NGF (29). Including our results, there is enough evidence to
support the involvement of Rb1 in differentiation of PC12
cells. However, the role of Rb1 in cerebellar development is
unknown. There have been very few reports of other genes that
induced neurite outgrowth. Our findings should be a starting
point for elucidating the roles of these genes in this process.

Most of the genes we identified as promoting cell growth
have been described in the literature as having similar func-
tions. Hdgf was found to have activity in hepatocytes (26) but
may also be involved in nervous system development (1).

Fig. 2. Cell transfection array. PC12 cells transfected with enhanced green fluorescent protein alone (A and B) and PC12 cells differentiated into neurons by
treatment with 100 ng/ml neuron growth factor (NGF; C). Fluorescent micrographs show neurite outgrowth of cells that overexpressed Nnat (D), Rb1 (E),
TREK-1 (F), Rnf13 (G), and eIF4A1 (H) in the absence of NGF.

Fig. 3. The characteristics of the expression-
function correlation. Vertical axis, the log-
transformed ratio of the mean value of ex-
pression levels in 8-day cerebella to the
mean value of expression levels in 12-day
and 3-wk cerebella. Horizontal axis, the 288
genes subjected to transfection were sorted
by the factor score of the second component
of PCA. Crossbars designate genes whose
overexpression resulted in phenotypic
changes. From left to right, blue bars (neurite
outgrowth) indicate TREK-1, Rb1, Rnf13,
eIF4A1, and Nnat, and red bars (cell prolif-
eration) indicate Hdgf, FIN14/Ddx3x, Ier2/
pip92/ETR109, Mapk6/Erk3, Epha3/Hek/
Mek4, Sipa1/Spa1, and Fau/MNSFβ.
Fin14/Ddx3x was initially identified as a FGF-inducible gene (14). Ier2/pip92/ETR101 was identified in FGF-stimulated hippocampal cells during neuronal differentiation (5) and in PC12 cells treated with cell proliferation agents, neuronal differentiation agents, and membrane depolarization agents (3). Sipa1/Spa1 is transcriptionally induced in murine lymphoid cells after mitogenic stimulation (15). The expression pattern of Epha3/Hek/Mek4 suggests its participation in guidance of cell migration and axonal growth cones during development (6, 12). Epha3/Hek/Mek4 displayed a lobe-dependent variability in expression pattern in the developing chick cerebellum (17), although its role in cerebellar development is poorly understood. Mapk6/Erk3 is localized in the nucleus in exponentially growing, quiescent, and growth factor-stimulated cells (4). Fau/MNSFB, a lymphokine product of a T cell hybridoma, was originally isolated from a radiation-induced mouse osteosarcoma (11) and possibly acts as a tumor suppressor gene (23). Among the genes selected by the cell transfection assay, only a few genes, i.e., Nnat, Ier2/pip92/ETR101, and Epha3/Hek/Mek4, have direct evidence of gene expression in neuronal development. However, this does not exclude the possibility of other genes involvement in neuronal development.

There may be criticism that functionality in PC12 cells and cerebellar development are different biological processes. Thus activity in PC12 cells may not reflect processes occurring in cerebellar granule cells. However, another interpretation of our system is that this heterogeneity implies stringency of our experimental conditions: the genes we have identified are active enough to induce phenotypic changes in different types of cells, i.e., both PC12 and granule cells. The identification of functional genes depends on the assay system, and our current procedure is the most stringent as far as using this heterogeneous system. More sensitive assays such as transfection under optimal conditions. Despite several reservations, we stress that increase pseudopositive hits, and it will be important to choose other genes. It should be noted that sensitive assays may weak stimulation by growth factors, e.g., NGF, may identify neurous system. More sensitive assays such as transfection under procedure is the most stringent as far as using this heteroge-

Among the genes selected by the cell transfection assay, only originally isolated from a radiation-induced mouse osteosar-

Spa1 is transcriptionally induced in murine lymphoid cells

effects in the developing chick cerebellum (17). The expression pattern of Epha3/Hek/Mek4 suggests its participation in guidance of cell migration and axonal growth cones during development (6, 12). Epha3/Hek/Mek4 displayed a lobe-dependent variability in expression pattern in the developing chick cerebellum (17), although its role in cerebellar development is poorly understood. Mapk6/Erk3 is localized in the nucleus in exponentially growing, quiescent, and growth factor-stimulated cells (4). Fau/MNSFB, a lymphokine product of a T cell hybridoma, was originally isolated from a radiation-induced mouse osteosarcoma (11) and possibly acts as a tumor suppressor gene (23). Among the genes selected by the cell transfection assay, only a few genes, i.e., Nnat, Ier2/pip92/ETR101, and Epha3/Hek/Mek4, have direct evidence of gene expression in neuronal development. However, this does not exclude the possibility of other genes involvement in neuronal development.

There may be criticism that functionality in PC12 cells and cerebellar development are different biological processes. Thus activity in PC12 cells may not reflect processes occurring in cerebellar granule cells. However, another interpretation of our system is that this heterogeneity implies stringency of our experimental conditions: the genes we have identified are active enough to induce phenotypic changes in different types of cells, i.e., both PC12 and granule cells. The identification of functional genes depends on the assay system, and our current procedure is the most stringent as far as using this heterogeneous system. More sensitive assays such as transfection under weak stimulation by growth factors, e.g., NGF, may identify other genes. It should be noted that sensitive assays may increase pseudopositive hits, and it will be important to choose optimal conditions. Despite several reservations, we stress that the cell transfection array or comparable large-scale functional analysis is per se a strong approach to finding new molecules involved in a particular biological process.

GRANTS

This study was supported by the Program for Promotion of Fundamental Studies in Health Sciences of the Organization for Pharmaceuticals and Medical Devices Agency.


