Digital Transcriptome Analysis in the Aging Cerebellum

MAGDALENA C. POPESCO, ADRIENNE FROSTHOLM, KATARZYNA REJNIAK, AND ANDREJ ROTTER

Department of Pharmacology and the Mathematical Biosciences Institute, The Ohio State University, Columbus, Ohio 43210, USA

ABSTRACT: Serial analysis of gene expression (SAGE) was used to identify and quantify all expressed cerebellar genes in the adult (P92) and aged (P810) C57BL/6J mouse cerebellum. A "closest-neighbor" algorithm was used to differentiate low abundance tags from possible sequencing errors in both libraries. Unique tags were categorized into four groups: (1) novel genes; (2) ESTs; (3) RIKEN, KIA, and hypothetical genes; and (4) known genes. Known genes were further subdivided into functional categories based on the gene ontology classification, using a web-based program developed in this laboratory (MmSAGEClass). Comparison of adult and aged cerebellar libraries revealed several genes that were differentially expressed, including growth hormone and prolactin, both of which were markedly decreased in the aged cerebellum. In addition, several tags showing differential expression were not identified in the Unigene database and are likely to represent novel genes. The present SAGE data on the aged cerebellar transcriptome may reveal candidate genes involved in the aging process.

KEYWORDS: digital transcriptome analysis; cerebellum; SAGE

The major function of the cerebellum is to integrate sensory input and motor output, thus modulating movement and balance. The cerebellar cortex of aged mice has a reduced number of Purkinje, basket, and stellate cells; granule cell numbers are retained relatively intact.[1] Senescent cerebellar cells also display dendritic atrophy,[2] reduction in somatic volume,[3] and alterations in electrophysiological properties[4] and microstructural features.[5] In recent studies, DNA microarrays were used to examine gene expression profiles in the aging mouse cerebellum;[6,7] the results suggested that senescence was accompanied by an inflammatory response, oxidative stress, and reduced neurotrophic support, all aspects of human neurodegenerative diseases.

Whereas DNA microarray technology is a "closed" system capable of detecting known genes, Serial Analysis of Gene Expression (SAGE) is an "open" system that detects and quantifies both known, and previously unknown, genes. The basic prin-

Address for correspondence: Department of Pharmacology and the Mathematical Biosciences Institute, The Ohio State University, Columbus, OH 43210. Voice: 614-292-7747.
rotter.1@osu.edu

Ann. N.Y. Acad. Sci. 1019: 58–63 (2004). **© 2004 New York Academy of Sciences.**
doi: 10.1196/annals.1297.013

ciple of SAGE is the isolation of a short cDNA sequence (tag) from a specific and invariable position within a given mRNA. This 14–15-bp tag is located immediately adjacent to the 3′ proximal *Nla*III restriction site; its sequence varies according to the particular mRNA from which it was derived. Each tag contains enough nucleotides to identify each transcript and thus is a unique marker for any expressed gene. The frequency with which any particular tag is detected is directly proportional to the number of mRNAs originally present in the cell or tissue being studied. Therefore, the number of identical tags detected is a measure of the abundance of the corresponding mRNAs in the original tissue. Each unique SAGE tag is matched with tags derived from the Unigene database. In the current study, serial analysis of gene expression was conducted in the adult and aged mouse cerebellum.

METHODS

SAGE libraries were constructed from the cerebellum of postnatal day (P) 92 and 810 mice. The aged mouse (C57BL/6Jnia) was purchased from the National Institute of Aging colony. The adult C57BL/6J animal was bred from Jackson Laboratories stocks. Animals were killed, cerebella were removed rapidly, and total RNA was prepared.[8] SAGE (I-SAGE kit) was performed as described by the manufacturer (Invitrogen). Sequence trace files were analyzed with the *Phred* base-calling software (Applied Biosystems). The SAGE2000 analysis software[8] was used to extract and analyze the primary sequence data from the electrophoretic trace files. The software extracts tag sequences from the sequence files, counts each tag, and provides a report containing the occurrence of each tag and its expression level and *P* value. Tag sequences were compared with the National Center for Biotechnology Information (NCBI) mouse SAGE tag-to-gene mapping reference database (ftp://ftp.ncbi.nih.gov/pub/sage/map/Mm/NlaIII/).

RESULTS AND DISCUSSION

Tag frequency in each SAGE library reflects the relative abundance of the corresponding cerebellar mRNA, a feature that allows digital comparisons between independently generated libraries. Both libraries were calculated to contain more than 100,000 tags, indicating that they were sufficiently large to obtain a good representation of expressed genes. The actual number of tags sequenced (total tags) was much smaller: 16,430 for the P92 library and 18,581 for the P810 library. This cutoff was arbitrary, based primarily on sequencing cost. Because the original libraries were retained, it is possible to obtain additional tags by sequencing more clones. Additional sequencing will permit the detection of additional low-frequency tags and more accurate quantitation of low-abundance tags. The tag sequences and counts were deposited in the SAGE database accessible via the NCBI Web site.

Although the majority of tags in both libraries were present in multiple copies, a significant number of tags were present once only. In most studies, these single tags are discounted because it is assumed that many could arise through errors due to base

substitution, deletion, or addition. However, this process tends to eliminate genuine tags corresponding to genes expressed at very low levels. To compromise between the two extremes of either retaining or discounting all single tags, we developed an algorithm ("closest neighbor") that substitutes, deletes, or adds bases to determine whether any tag in the SAGE library is related to another. In the event that such a tag is found, it is deleted from the list of total tags. This adjustment reduced the number of tags to 14,128 in the P92 library and 13,131 in the P810 library. All duplicate tags and genes were consolidated to obtain the number of unique tag sequences: the P92 library contained 4,842, and the P810 library had 5,094 unique tags.

Unique tags were further categorized into four groups: (1) novel genes, (2) ESTs, (3) RIKEN, KIA and hypothetical genes, and (4) known genes. The percentage of tags in each group was (1) 11 and 20, (2) 26 and 27, (3) 25 and 21, and (4) 38 and 32, for P92 and P810 mice, respectively. The high percentage of novel genes is surprising because they represent relatively abundant transcripts, which it might be assumed would have been detected previously. It is also notable that the percentage of novel tags is almost twice as high in the aged cerebellar library, suggesting that during the aging process novel genes become active. The known genes were further subdivided into functional categories based on the Gene Ontology classification (http://www.geneontology.org) using a Web-based program developed in our laboratory (http://mbi.osu.edu/~rejniak/MmSAGEClass.html). For both young adult and aged libraries, among the most numerous identifiers were enzyme activity (~6%), binding activity (~5%), transporter activity (~4%), signal transducer activity (~1%), and structural molecule activity (~1%). A notable feature of the classification scheme is that the sizes of the functional gene classes expressed in the two libraries were conserved. This indicates that there are no major shifts in the categories of abundantly expressed genes during aging, but changes do occur in the expression of individual genes within the larger functional categories.

The three hundred most abundant genes in the P92 and P810 libraries were compared. The results (TABLE 1) are listed in decreasing order of fold change (>5). This calculation takes into consideration the total tags in each group. The two libraries represent gene expression in neurons and glia of the entire cerebellar cortex and deep cerebellar nuclei, as well as in blood vessels, choroid plexus, and pia arachnoid membrane. Because granule cells compose the majority of all cerebellar cells, the SAGE transcriptome is presumably heavily weighted to this cell type. Genes demonstrating a large upregulation in the P810 cerebellum included testis-specific protein, dcd8; tumor rejection antigen, gp96; macrophage activation 2; olfactomedin 1 and several ESTs, RIKEN, and novel genes. Genes exhibiting a large downregulation included prolactin, Smt3h2, ribosomal protein S3a, glucose-regulated protein 58, growth hormone, a RIKEN gene, and a novel gene.

In this study, we have used the most stringent controls and data acceptance criteria to provide a conservative estimate of genes that may participate in cerebellar aging. We have demonstrated substantial changes in several transcripts in young adult and aged mouse cerebellum, the most striking of which was a strong downregulation in prolactin and growth hormone. Both are signal transduction molecules that interact with cognate receptors activating the JAK2/STAT5 signaling pathway.[9] Furthermore, growth hormone and prolactin influence the insulin growth factor–1 signaling pathway via IRS-1 and Shc,[9] respectively. Reduced activity of the IGF-1 pathway greatly increases life span in *C. elegans*.[10]

TABLE 1. Differentially expressed transcripts in a young adult (P92) and aged (P810) mouse cerebellum[a]

Tag Sequence	Unigene ID No.	Gene Name	Tag count P92	Tag count P810	Fold Change	*P* Chance
Increase in gene expression						
ATAAATACAT	41973	testis-specific protein, Ddc8	0	24	21.22	<2.3E-05
TATTAAATAC	34715; 131660	RIKEN cDNA 1700026N20 gene; ESTs	0	15	13.26	2.3E-05
ATAATACAAT		novel	0	15	13.26	2.3E-05
AATAATACAT	127042	EST	0	15	13.26	2.3E-05
TGTATAAAAA	4526; 227099	tumor rejection antigen gp96; tubulin, beta 2	1	14	12.38	5.6E-04
ATAATACATT	5057	macrophage activation 2	0	13	11.50	2.2E-04
ATAATACCAT	218873	RIKEN cDNA 3110054G10 gene	0	11	9.73	9.5E-04
ATAATAACAT	10735	calcium/calmodulin-dependent serine protein kinase	0	11	9.73	9.5E-04
AATACTGACA		novel	0	11	9.73	9.5E-04
ATTAATACAT		novel	0	10	8.84	0.002
TAACTTTAAG	43278	olfactomedin 1	2	12	5.31	0.012
CCCTTCTTCT	196110; 89136	hemoglobin alpha, adult chain 1; H3 histone, family 3A	9	52	5.11	<2.3E-05

TABLE 1. (*continued*) **Differentially expressed transcripts in a young adult (P92) and aged (P810) mouse cerebellum**[a]

Tag Sequence	Unigene ID No.	Gene Name	Tag count P92	Tag count P810	Fold Change	*P* Chance
Decrease in gene expression						
CTTGGGTGCA	1270	prolactin	29	2	−16.40	<2.3E-05
TAGGGCAATC	29923	SMT3 (suppressor of mif two, 3) homologue 2 (*S. cerevisiae*)	11	1	−12.44	4.0E-04
GGGAAGGCGG	6957	ribosomal protein S3a	11	1	−12.44	4.0E-04
GGCTGCATTC	29703; 709	DNA segment, Chr 7, expressed; glucose regulated protein	11	0	−12.44	9.0E-04
GGGTTGTTCA	74363	transient receptor cation channel, subfamily C, member 3	10	1	−11.31	9.0E-04
AAGTGTCGCC	1240	growth hormone	37	4	−10.46	<2.3E-05
TGTCATCTAG	4071; 5163	laminin receptor 1; Ras-like without CAAX 2	9	1	−10.18	0.004
GCTGCCCTAG	197515; 196396	tubulin, alpha 2; tubulin, alpha 1	8	1	−9.05	0.007
GAGCGTTTTG	5246	peptidylprolyl isomerase A	13	2	−7.35	5.2E-04
GACAAAGGGG	38576	DNA segment, Chr 11, Brigham	13	2	−7.35	5.2E-04
AGGAGGACTT	14534	RIKEN cDNA 9130221H12 gene	19	3	−7.16	2.0E-05
AACAATTTGG	14244	ribosomal protein L9	18	3	−6.79	2.3E-05
CCGCCCCTTT	29846	N-myc downstream regulated 4	6	1	−6.79	0.026
GATGTGGCTG	2718; 90587	eukaryotic translation elongation factor 1 beta 2; enolase 1, alpha	11	2	−6.22	0.002
GCCCCCCTCT	44101	hypothetical protein MGC27631	16	3	−6.03	1.5E-04
ATGACTGATA		novel	53	10	−5.99	<2.3E-05
GAGAGAAGAG	1287	microtubule-associated protein tau	15	3	−5.65	2.2E-04
TGTGTGAGGA	141230; 21086	1-acylglycerol-3-phos O-acyltransferase 3; elongation factor 1	10	2	−5.65	0.003

[a]Transcripts listed show greater than 5-fold increases.

ACKNOWLEDGMENTS

This work was supported by NIH Grants AG021698 and AA014422.

REFERENCES

1. STURROCK, R.R. 1989. Changes in neuron number in the cerebellar cortex of the ageing mouse. J. Hirnforsch. **30:** 499–503.
2. HADJ-SAHRAOUI, N., F. FREDERIC, H. ZANJANI, *et al*. 2001. Progressive atrophy of cerebellar Purkinje cell dendrites during aging of the heterozygous staggerer mouse (*Rora*$^{+/sg}$). Dev. Brain Res. **126:** 201–209.
3. MONTEIRO, R.A.F., R.M.F. HENRIQUE, E. ROCHA, *et al*. 1998. Age-related changes in the volume of somata and organelles of cerebellar granule cells. Neurobiol. Aging **19:** 325–332.
4. ROGERS, J., M.A. SILVER, W.J. SHOEMAKER & F.E. BLOOM. 1980. Senescent changes in a neurobiological model system: cerebellar Purkinje cell and correlative anatomy. Neurobiol. Aging **1:** 3–11.
5. ROGERS, J., S.F. ZORNETZER, F.E. BLOOM & R.E. MERVIS. 1984. Senescent microstructural changes in the rat cerebellum. Brain Res. **292:** 23–32.
6. LEE, C.-K., R. WEINDRUCH & T.A. PROLLA. 2000. Gene-expression profile of the ageing brain in mice. Nat. Genet. **25:** 294–297.
7. WEINDRUCH, R., T. KAYO, C.K. LEE & T.A. PROLLA. 2002. Gene expression profiling of aging using DNA microarrays. Mech. Aging Dev. **123:** 177–193.
8. VELCULESCU, V.E., L. ZHANG, B. VOGELSTEIN & K.W. KINZLER. 1995. Serial analysis of gene expression. Science **270:** 484–487.
9. CARTER-SU, C., A.P. KING, L.S. ARGETSINGER, *et al*. 1996. Signalling pathways of GH. Endocrinol. J. **43:** S65–S75.
10. MURPHY, C.T., S.A. MCCARROLL, C.I. BARGMANN, *et al*. 2003. Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. Nature **424:** 277–284.