Neurogenetics (2006) 7: 67–80
DOI 10.1007/s10048-006-0032-6

ORIGINAL ARTICLE

Richard B. Roth · Peter Hevezi · Jerry Lee · Dorian Willhite · Sandra M. Lechner · Alan C. Foster · Albert Zlotnik

Gene expression analyses reveal molecular relationships among 20 regions of the human CNS

Received: 22 December 2005 / Accepted: 16 February 2006 / Published online: 30 March 2006

Abstract Transcriptional profiling was performed to survey the global expression patterns of 20 anatomically distinct sites of the human central nervous system (CNS). Forty-five non-CNS tissues were also profiled to allow for comparative analyses. Using principal component analysis and hierarchical clustering, we were able to show that the expression patterns of the 20 CNS sites profiled were significantly different from all non-CNS tissues and were also similar to one another, indicating an underlying common expression signature. By focusing our analyses on the 20 sites of the CNS, we were able to show that these 20 sites could be segregated into discrete groups with underlying similarities in anatomical structure and, in many cases, functional activity. These findings suggest that gene expression data can help define CNS function at the molecular level. We have identified subsets of genes with the following patterns of expression: (1) across the CNS, suggesting homeostatic/housekeeping function; (2) in subsets of functionally related sites of the CNS identified by our unsupervised learning analyses; and (3) in single sites within the CNS, indicating their participation in distinct site-specific functions. By performing network analyses on these gene sets, we identified many pathways that are upregulated in particular sites of the CNS, some of which were previously described in the literature, validating both our dataset and approach. In summary, we have generated a database of gene expression that can be used to gain valuable insight into the molecular characterization of functions carried out by different sites of the human CNS.

Keywords Central nervous system · Human · Transcriptional profiling · Microarray · Network analysis

Abbreviations *ADORA2A*: Adenosine A2a receptor · *AMPH*: Amphiphysin (Stiff–Man syndrome with breast cancer 128 kDa autoantigen) · *CACNA1A*: Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit · *CACNA1B*: Calcium channel, voltage-dependent, L type, alpha 1B subunit · CNS: Central nervous system · *CRTAM*: Class I MHC-restricted T cell-associated molecule · CTP: Cytosine triphosphate · *DLG2*: Discs, large homologue 2, chapsyn-110 (*Drosophila*) · *DLG4*: Discs, large homologue 4 (*Drosophila*) · *DLGAP1*: Discs, large (*Drosophila*) homologue-associated protein 1 · *DNM1*: Dynamin 1 · *DRD2*: Dopamine receptor D2*EDNRB*-endothelin receptor type B · *EPS15*: Epidermal growth factor receptor pathway substrate 15 · *FOSB*: FBJ murine osteosarcoma viral oncogene homologue B · *FYN*: FYN oncogene related to SRC, FGR, and YES · *GABRA6*: Gamma-aminobutyric acid (GABA) A receptor, alpha 6 · GPCR: G-protein coupled receptor · *GRIK2*: Glutamate receptor, ionotropic, kainate 2 · *GRIN1*: Glutamate receptor, ionotropic, *N*-methyl D-aspartate 1 · *GRM3*: Glutamate receptor, metabotropic 3 · IPKB: Ingenuity Pathways Knowledge Base · IVT: In vitro transcription · *MBP*: Myelin basic protein · PCA: Principal component analysis · PCR: Polymerase chain reaction · *PMCH*: Pro-melanin-concentrating hormone · PMI: Postmortem interval · PNS: Peripheral nervous system · *PTPN5*: Protein tyrosine phosphatase, nonreceptor type 5 · qPCR: Quantitative polymerase chain reaction · *RAB3A*: RAB3A, member RAS oncogene family · RMA: Robust multiarray analysis; robust multichip average · RNA: Ribonucleic acid · *SFRP4*: Secreted frizzled-related protein 4 · *SLC1A3*: Solute carrier family 1 (glial high affinity glutamate

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s10048-006-0032-6

R. B. Roth (✉) · P. Hevezi · J. Lee · D. Willhite · A. Zlotnik
Department of Molecular Medicine,
Neurocrine Biosciences, Incorporated,
12790 El Camino Real,
San Diego, CA 92130, USA
e-mail: rroth@neurocrine.com
Tel.: +1-858-6177204
Fax: +1-858-6177696

S. M. Lechner · A. C. Foster
Department of Neuroscience,
Neurocrine Biosciences, Incorporated,
12790 El Camino Real,
San Diego, CA 92130, USA

transporter), member 3 · *SNAP25*: Synaptosomal-associated protein, 25 kDa · *SYT1*: Synaptotagmin I · *SYT3*: Synaptotagmin III · *SYT4*: Synaptotagmin IV · *TH*: Tyrosine hydroxylase · TOI: Target of interest · UTP: Uridine triphosphate · *VAMP2*: Vesicle-associated membrane protein 2 (synaptobrevin 2)

Background

A considerable amount of research in neuroscience has focused on understanding the anatomical design, physiology, and function of the human central nervous system (CNS). This knowledge has led to a greater understanding of the pathology of several CNS-linked diseases. Unfortunately, a complete understanding of the molecular events that are altered during disease development and progression remains a significant challenge for many CNS diseases. In many areas of biological research, significant progress has been made in defining the molecular events underlying disease development and progression through the use of genomics-based tools that include genotyping, proteomics, and transcriptomics. For example, in cancer research, transcriptional profiling has enabled the differentiation of cancer subtypes from one another [1, 2], allowed evaluation of long-term prognoses [3–7], and identified novel potential prognostic [8–10] and therapeutic targets [10–12]. The use of transcriptional profiling in CNS research until now has been limited by the difficulty in obtaining high quality tissue samples from many regions of the CNS, particularly from healthy donors, making it difficult to perform comparative analyses between diseased and nonpathogenic samples. In addition, the inherent structural and functional complexity of the human CNS has limited the impact of transcriptional profiling at a regional level.

Recently, a few studies reported comparative analyses of gene expression across multiple healthy human tissues [13–19]. Using hierarchical clustering [20] and principal component analysis (PCA) [21, 22], several groups have shown that gene expression data can correctly segregate samples by tissue type. Furthermore, a number of genes showing tissue-specificity were also identified [13, 16, 17, 19]. These groups have successfully validated this profiling approach as a powerful tool to correlate gene expression with different tissues. As a result, they have provided a foundation upon which the genes responsible for a tissue's function and in the event of disease, dysfunction can be identified.

Although these studies often included CNS tissues, none of these studies have focused on the CNS, resulting in either a lack of sufficient biological replicate depth or a lack of gender representation across tissues. In the present study, we have focused on the human CNS and profiled samples from 20 different CNS regions totaling 169 samples with seven to nine biological replicates surveyed per tissue. For each CNS region, four male and four female biological replicates were profiled with the exception of the vestibular nuclei superior where only three male samples were profiled. We used the Affymetrix human U133 plus 2.0 array to measure expression levels of more than 47,000 unique transcripts, providing a genome-wide perspective of each tissue's expression signature. This represents the most comprehensive molecular profiling of the human CNS to date. Furthermore, to allow for comparative analyses with tissues outside of the CNS, we also profiled samples from 45 additional sites of the human body, generating a dataset of 353 samples representing 65 tissues. All data from this study were uploaded to National Center for Biotechnology Information, Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/projects/geo) with the GSE ID GSE3526. For a drug discovery and development program focused on neurological disorders, the utility of such a database can be very helpful. For example, by identifying genes that are differentially expressed between a set of neurological disease samples and nondisease samples such as those present in our CNS database, novel drug targets can be revealed. Furthermore, during the selection process of a target molecule for drug development, understanding the expression profile of that target across the CNS, in fact across the entire body, provides an idea of target specificity and can also reveal potential negative side effects resulting from expression in a nontarget tissue.

Using the dataset generated from this study, the gene expression signatures of the 20 CNS sites profiled were compared, providing insight into each tissue's anatomy and function at the molecular level. Using PCA and hierarchical clustering, we show that a common underlying molecular signature exists among the CNS samples, readily differentiating them from nonnervous tissues. Similarly, within the CNS, expression patterns are distinct enough to cluster the samples in a pattern that is consistent with known anatomical and functional differences. These results indicate that gene expression has the potential to be used to better understand CNS structure–function relationships. We therefore sought to identify genes that are (1) highly expressed only in the CNS, (2) coexpressed only in CNS sites of common anatomy and function, and (3) expressed in a single tissue within the human CNS. Furthermore, by performing network analyses, we have identified several potential networks, functions, and pathways active in these tissues. Our end result is the most comprehensive molecular profile of the human CNS to date that can be used as a starting point for more detailed molecular analyses of CNS structure and function.

Results

Transcriptional profiling of the human CNS

To implement our transcriptional profiling study, human postmortem tissue samples from ten donors were used (see "Materials and methods"). Donors were healthy and free of chronic disease at the time of death and cause of death in all cases was the result of a sudden event. To ensure sample quality and minimize RNA degradation, all tissue samples were flash frozen within 8.5 h postmortem [maximum

Table 1 CNS sites profiled

Tissue	*Number of samples*
Accumbens	9
Amygdala	8
Cerebellum	9
Cerebral cortex	9
Corpus callosum	9
Frontal lobe	9
Hippocampus	9
Hypothalamus	8
Medulla	9
Midbrain	9
Occipital Lobe	8
Parietal lobe	9
Putamen	9
Spinal cord	8
Substantia nigra	8
Subthalamic nucleus	8
Temporal lobe	8
Thalamus	8
Ventral tegmental area	8
Vestibular nuclei-superior	7
Total	169

Table 2 Non-CNS tissues profiled

Tissue	*Number of samples*
Adipose tissue	3
Adipose tissue, omental	4
Adipose tissue, subcutaneous	3
Adrenal gland, cortex	4
Bone marrow	5
Bronchus	3
Cervix	4
Colon cecum	3
Coronary artery	3
Dorsal root ganglia	8
Endometrium	4
Esophagus	4
Heart atrium	4
Heart ventricle	3
Kidney cortex	4
Kidney medulla	4
Liver	4
Lung	3
Lymph nodes	4
Mammary gland	3
Myometrium	5
Nipple, cross section	4
Nodose nucleus	8
Oral mucosa	4
Ovary	4
Pharyngeal mucosa	4
Pituitary gland	8
Prostate gland	3
Salivary gland	4
Saphenous vein	3
Skeletal muscle	5
Spleen	4
Stomach cardiac	3
Stomach fundus	4
Stomach pyloric	4
Testes	3
Thyroid gland	4
Tongue, main corpus	4
Tongue, superior w/ papillae	4
Tonsil	3
Trachea	3
Trigeminal ganglia	8
Urethra	3
Vagina	4
Vulva	4
Total	184

postmortem interval (PMI)=8.5 h]. Collectively, from each donor, samples from 20 regions of the CNS (Table 1) and 45 non-CNS tissues (Table 2) were procured. At least eight samples per CNS site (with the exception of the vestibular nuclei superior, *N*=7) and at least three samples per non-CNS site were profiled. In total, 353 samples were processed including 169 from the CNS and 184 from non-CNS sites. Together, these samples represent all major systems of the human body and include the nervous, circulatory, digestive, endocrine, excretory, immune, musculoskeletal, reproductive, respiratory, and skin systems. To monitor gene expression across the genome, Affymetrix human U133 plus 2.0 arrays that contain 54,675 probe sets per array, representing 21,974 unique UniGene clusters and more than 47,000 transcripts were used. Samples passing Affymetrix quality control metrics (3′/5′ glyceraldehyde-3-phosphate dehydrogenase ratio, percent present calls, etc.) were subsequently background corrected, normalized, and polished using robust multichip average (RMA) [23, 24] before analysis.

Genome-wide unsupervised analysis

To gain a global perspective of the transcriptional profile similarities and differences among the 65 tissues processed, PCA and hierarchical clustering were performed using all probe sets present on the array. Figure 1a shows the results of PCA. The transcriptional profiles of the 20 CNS regions are very different from nonnervous tissues and furthermore, the 20 CNS regions maintain some level of similarity at the level of gene expression as they form a single cluster. A second PCA was subsequently performed, this time using data from the 20 CNS regions only (Fig. 1b), and showed that even with the underlying similarity of transcript profiles among the CNS sites, differences in expression are distinct enough to separate the regions from one another. The cerebellum is the most distinct from the remaining 19 CNS sites, indicating a significantly different

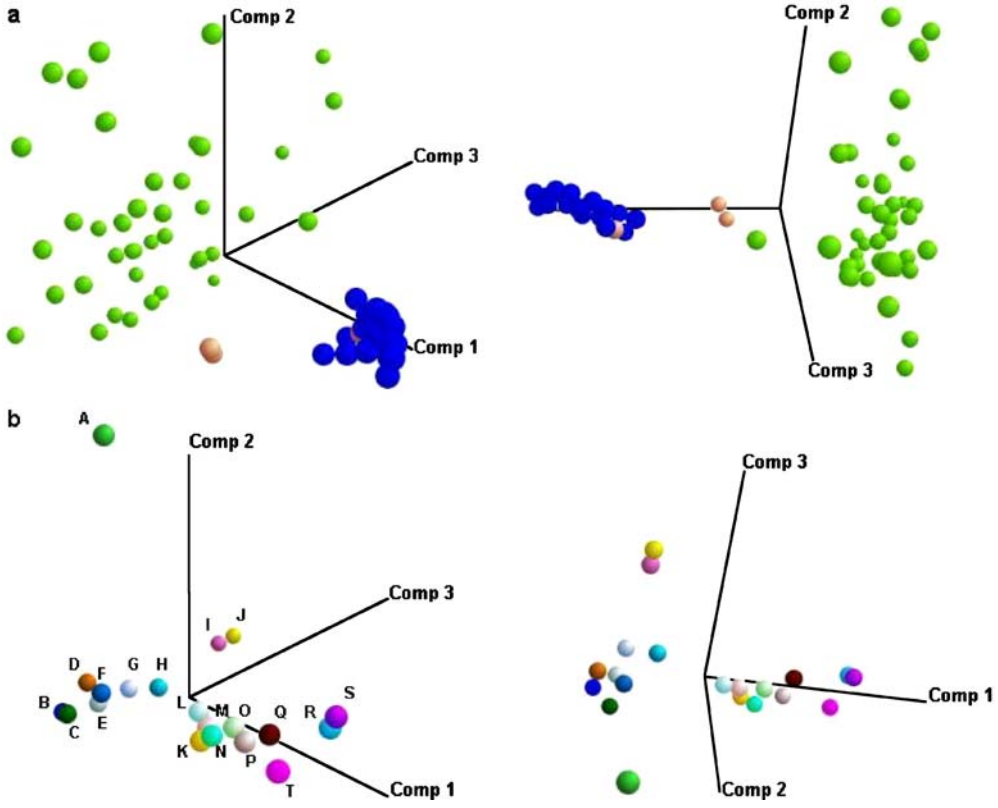

Fig. 1 PCA of tissues. **a** The expression profiles of 65 tissues were compared using all probe sets present on the Affymetrix Human Genome U133 plus 2.0 array (*N*=54,675) using PCA. The three most important principal components are shown (*Comp 1*, *Comp 2*, and *Comp 3*). Tissues are represented by single-colored spheres: *green spheres* represent nonnervous tissues, *brown spheres* represent peripheral nervous tissues, and *blue spheres* represent CNS tissues. **b** PCA was performed as in subpanel **a** but restricted to only CNS sites. *A* Cerebellum, *B* parietal lobe, *C* occipital lobe, *D* frontal lobe, *E* temporal lobe, *F* cerebral cortex, *G* amygdala, *H* hippocampus, *I* accumbens, *J* putamen, *K* vestibular nuclei superior, *L* thalamus, *M* hypothalamus, *N* midbrain, *O* ventral tegmental area, *P* medulla, *Q* substantia nigra, *R* subthalamic nuclei, *S* corpus callosum, and *T* spinal cord

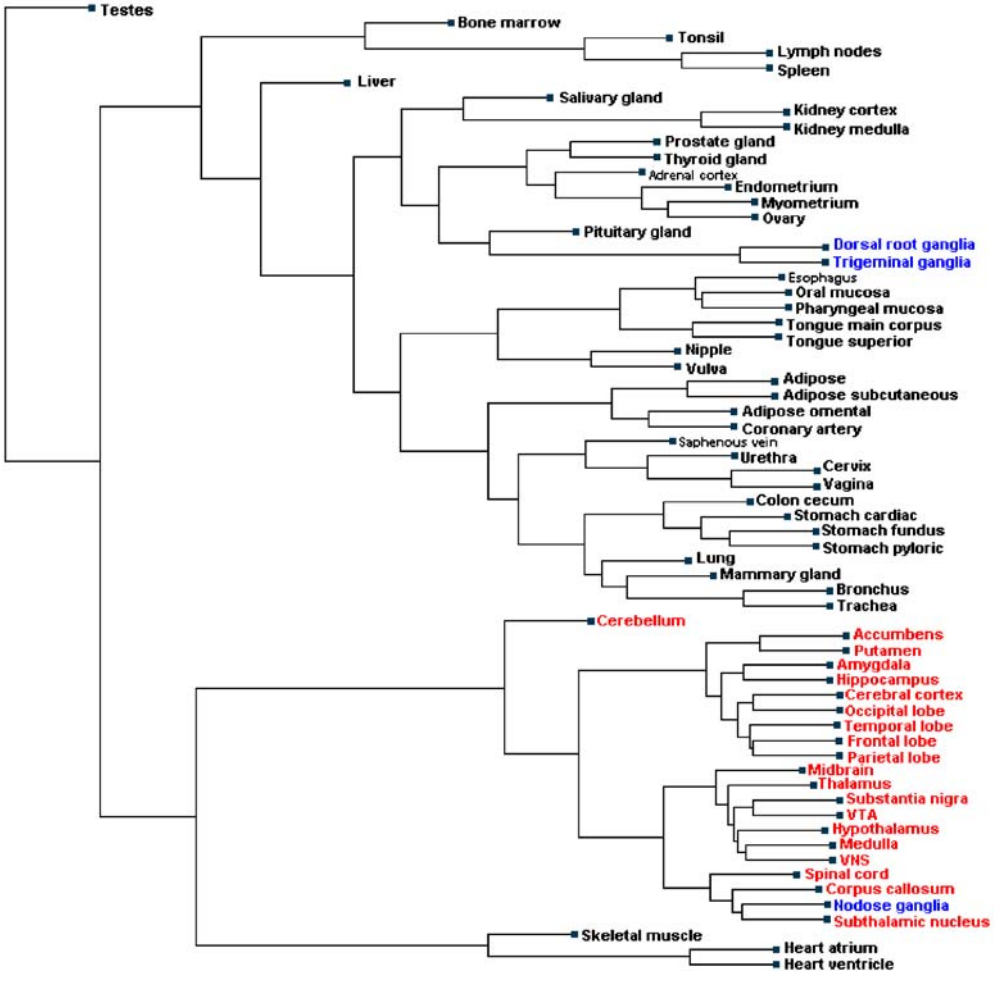

Fig. 2 Hierarchical clustering of all tissues. Using hierarchical clustering, the expression profiles of 65 tissues were compared using all probe sets present on the Affymetrix Human Genome U133 plus 2.0 array (*N*=54,675). The 20 CNS sites included in this analysis were all found in a single branch of the resulting dendrogram. CNS tissue in *red text*, PNS tissue in *blue text*, and nonnervous tissue in *black text*

gene expression profile. The putamen and accumbens, which form part of the basal ganglia, clustered together as did the five cortex sites: cerebral cortex, frontal lobe, parietal lobe, occipital lobe, and temporal lobe. These data suggest that the global expression profiles reflect the structural similarities, common functions, and ontogeny of the different CNS regions.

The result of hierarchical clustering of all 65 tissues is shown in Fig. 2a. Once again, the 20 CNS sites profiled cluster together and are found in a single branch of the dendrogram. Hierarchical clustering performed using data only from the 20 CNS sites revealed three CNS branches (Fig. 3). One branch contains only the cerebellum, again indicating a pattern of gene expression significantly different from the rest of the CNS. A second branch consists of CNS sites found only in the forebrain, including the telencephalon and diencephalon. This branch further separates into three clusters that correlate with the known functions of cluster members. For example, the putamen and accumbens form one cluster, which are both components of the basal ganglia; the second cluster contains the amygdala and hippocampus, otherwise known as the limbic system; and the third cluster consists of the five cortex samples in our dataset, all of which play a role in higher learning. The remaining CNS branch is populated with CNS sites representing the forebrain, midbrain, hindbrain, and spinal cord, yet, these sites also ultimately segregate into several clusters that reflect a common anatomy and/or function. The spinal cord shares the least similarity with the other CNS sites found in this branch of the CNS and with the exception of the cerebellum, exhibits the most distinct expression profile of all the CNS sites analyzed. Based on its location and function relative to the other CNS sites, this finding is not unexpected. Other examples from this branch include the clustering of the substantia nigra and the ventral tegmental area, which together form the dopaminergic center of the brain. Taken together, these results indicate that the global expression signatures of the human CNS reflect the anatomy and function of the CNS and raise the possibility of identifying genes associated with each of these regions that potentially mediate the functions of each of these CNS regions.

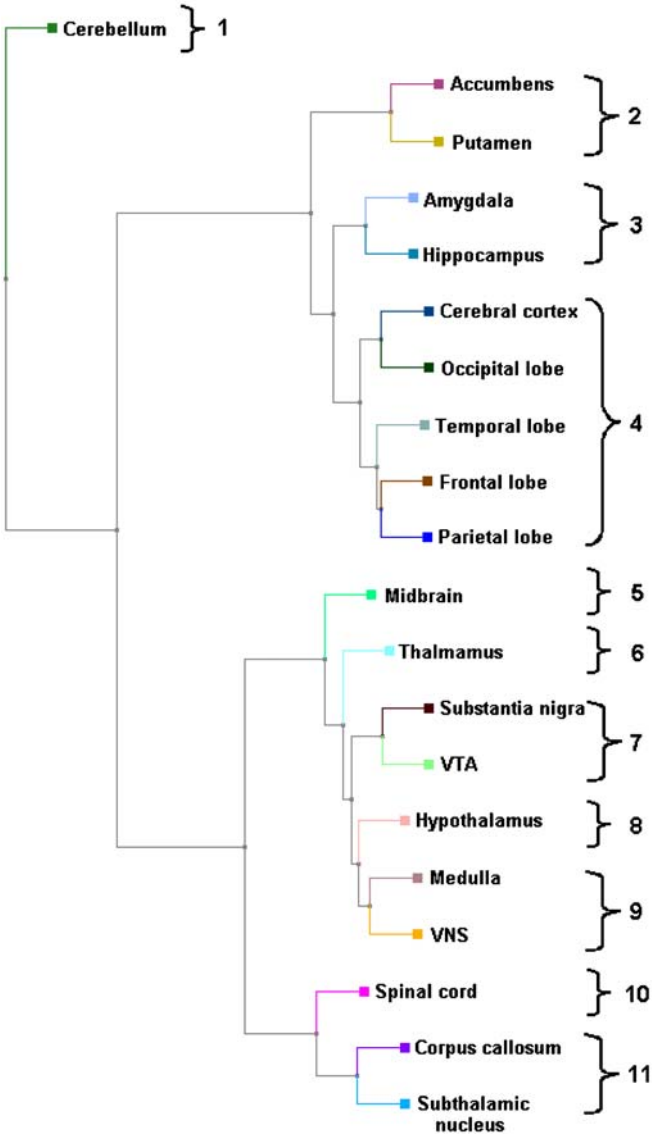

Fig. 3 Hierarchical clustering of CNS tissues. Using hierarchical clustering, the expression profiles of 20 CNS tissues were compared using all probe sets present on the Affymetrix Human Genome U133 plus 2.0 array (*N*=54,675). From this analysis, several clusters that correlate with known anatomical/functional groupings emerge. These clusters were used to find cluster-specific gene sets

CNS-specific, cluster-specific, and region-specific expression

Three approaches were used to find genes that (1) are expressed across the CNS as compared to non-CNS tissues (pan-CNS), (2) show expression correlating with each cluster as identified by hierarchical clustering (cluster-specific, Fig. 3), and (3) are expressed in a tissue-specific or highly specific fashion within the CNS (site-specific). An initial ANOVA filter was applied to find genes that were differentially expressed across the CNS, and genes meeting the *p* value cutoff criteria were used for all downstream analyses ($p<0.05$ with a Bonferroni correction for multiple testing). All in all, approximately 23% or 12,784 probe sets met these criteria.

We initially identified genes that showed highly elevated expression specifically across the CNS relative to non-CNS tissues (Pearson correlation score of 0.7 or greater). These pan-CNS genes are listed in the supplemental Table 1. By defining genes that are pan-CNS-specific, anatomical and functional attributes that are unique to the CNS may be potentially revealed. These genes likely include mostly CNS housekeeping or homeostatic genes—they may participate in the maintenance of the structure and function of neuronal tissues. Many genes identified as pan-CNS-specific are involved in well-defined nervous tissue-specific functions such as signal potentiation (*FYN*, *GRIK2*, and *PTPN5*), synaptic transmission (*DLG2*, *DLG4*, *DLGAP1*,

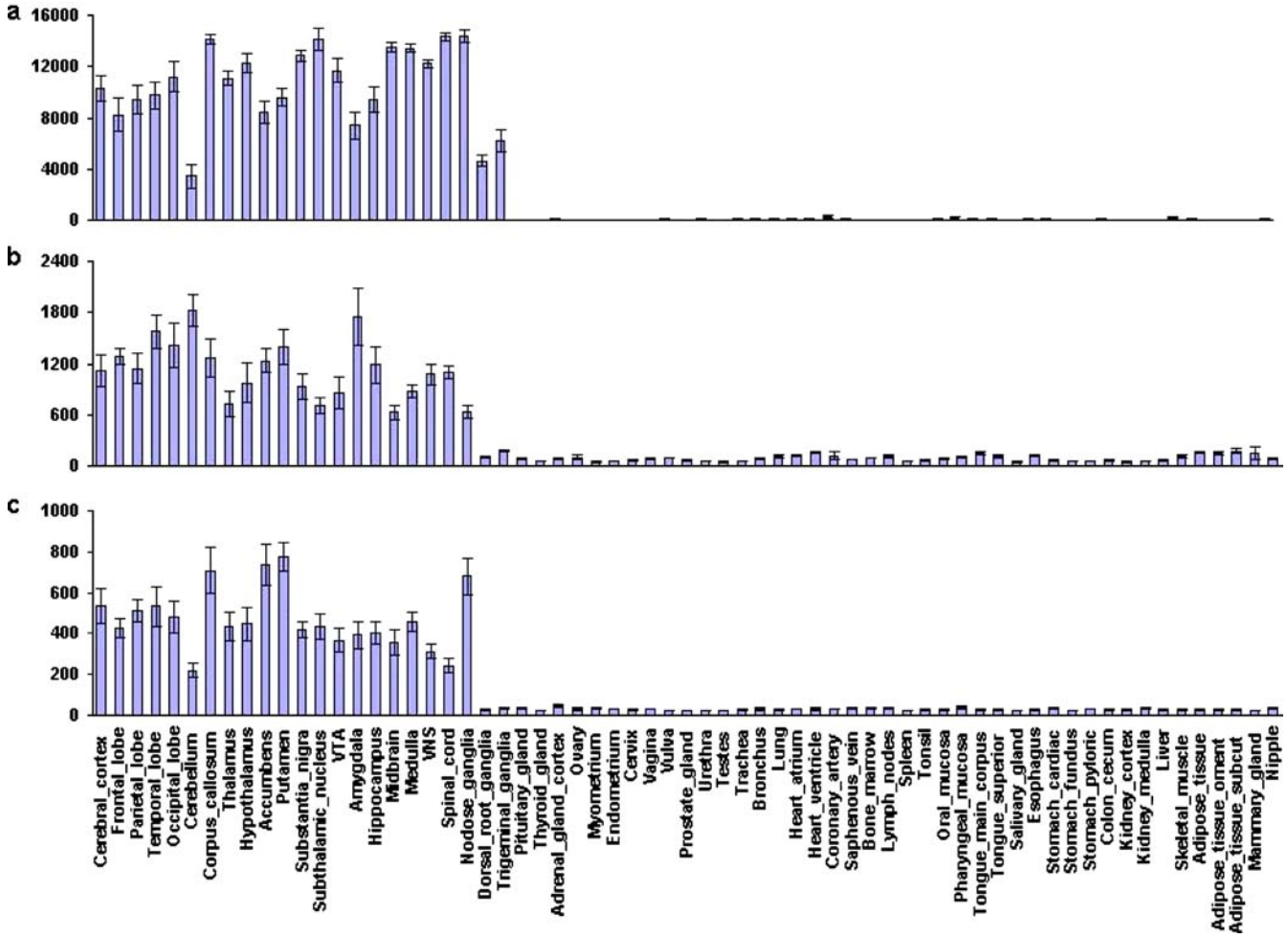

Fig. 4 pan-CNS-specific expression profiles. Depicted are representative expression profiles for several pan-CNS-specific probe sets across all 65 tissues. *Y-axis*, expression level (RMA values); *X-axis*, individual tissues; and *error bars* are SEMs. **a** 209072_at (*MBP* myelin basic protein), Pearson correlation=0.90. **b** 202800_at [*SLC1A3* solute carrier family 1 (glial high affinity glutamate transporter), member 3], Pearson correlation=0.94. **c** 205814_at (*GRM3* glutamate receptor, metabotropic 3), Pearson correlation=0.94. *VTA* Ventral tegmental area, *DRG* dorsal root ganglia, and *VNS* vestibular nuclei superior

and *GRIN1*), vesicle transport (*NAPG*, *RAB3A*, and *SYT3*), endocytosis of synaptic vesicles (*AMPH*, *DNM1*, and *EPS15*), and regulation of neurotransmitter levels (*CACNA1B*, *SNAP25*, *SYT1*, *SYT4*, and *VAMP2*). The expression profiles of three representative examples of pan-CNS genes are shown in Fig. 4.

Table 3 Summary of number of cluster-specific probe sets per cluster

Cluster	*Number of sets*
Accumbens and putamen	594
Amygdala and hippocampus	34
Cerebellum	1938
Corpus callosum and subthalamic nuclei	275
Cortex and lobes	766
Hypothalamus	8
Medulla and VNS	7
Midbrain	24
Spinal cord	201
Substantia nigra and VTA	11
Thalamus	39
Total	3897

VNS Vestibular nuclei superior and *VTA* ventral tegmental area

We next identified genes that exhibit up- or down-regulation in a single CNS cluster as defined in Fig. 3, relative to all other CNS clusters (Pearson correlation score of 0.7 or greater). Genes meeting these criteria are listed in the supplemental Table 2 in the and are summarized by cluster in Table 3. Cluster-specific genes likely participate in processes specific to each set of tissues found in that cluster and facilitate the molecular characterization of each cluster's distinct anatomy and function. Figure 5 depicts examples of expression profiles of genes representing several clusters.

Genes identified as tissue-specific or highly specific in a single tissue relative to all others can be used to molecularly characterize specific CNS sites and provide a starting point for further investigation into the functional activities occurring at each site. Supplemental Table 3 lists the genes meeting these criteria. It is interesting to note that

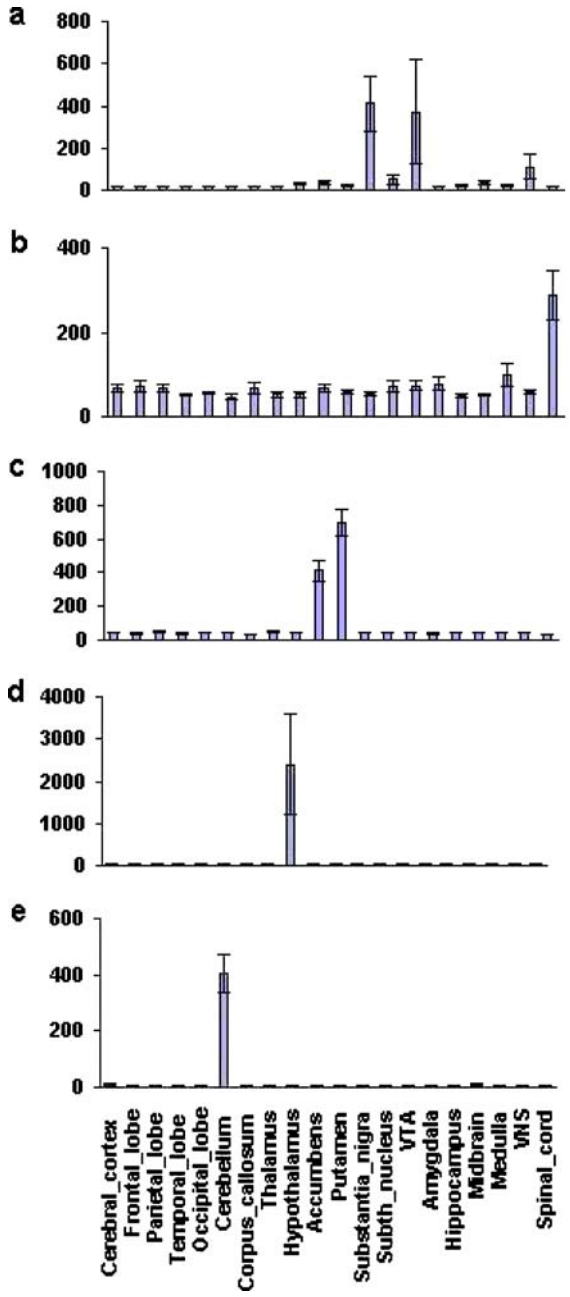

Fig. 5 Cluster-specific expression profiles. Shown are representative expression profiles for several CNS clusters as defined in Fig. 2b. *Y-axis*, expression level (RMA values); *X-axis*, individual tissues; and *error bars* are SEMs. **a** Substantia nigra and VTA: 208291_s_at (*TH* tyrosine hydroxylase), Pearson correlation=0.92. **b** Spinal cord: 202768_at (*FOSB* FBJ murine osteosarcoma viral oncogene homologue B), Pearson correlation=0.94. **c** Accumbens and putamen: 1554242_a_at (*ADORA2A* adenosine A2a receptor), Pearson correlation=0.99. **d** Hypothalamus: 206942_s_at (*PMCH* pro-melanin-concentrating hormone), Pearson correlation=0.99. **e** Cerebellum: 206914_at (*CRTAM* class-I MHC-restricted T cell-associated molecule), Pearson correlation=0.97. *VTA* Ventral tegmental area and *VNS* vestibular nuclei superior

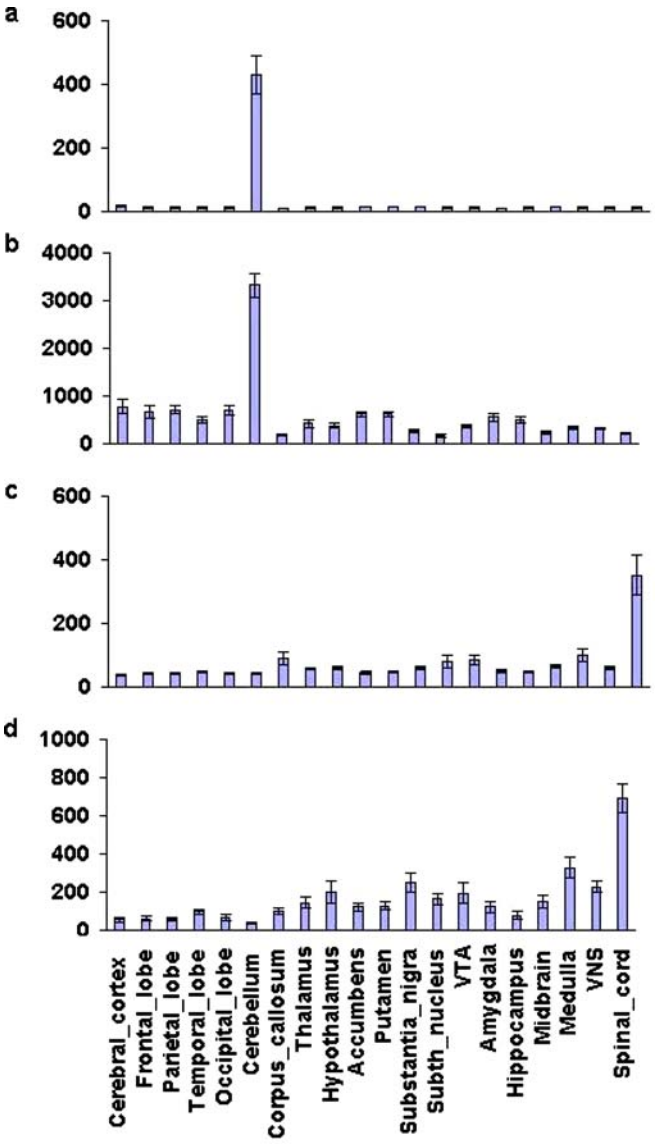

Fig. 6 Tissue-specific and highly specific expression profiles. Depicted across the CNS tissues are representative expression profiles of tissue-specific and highly specific probe sets. *Y-axis*, expression level (RMA values); *X-axis*, individual tissues; and *error bars* are SEMs. **a** 207182_at [*GABRA6* gamma-aminobutyric acid (GABA) A receptor, alpha 6], tissue-specific for the cerebellum. **b** 214933_at (*CACNA1A* calcium channel, voltage-dependent, P/Q type, alpha 1A subunit), highly specific for the cerebellum. **c** 204051_s_at (*SFRP4* secreted frizzle-related protein 4), tissue-specific for the spinal cord. **d** 204273_at (*EDNRB* endothelin receptor type B), highly specific for the spinal cord. *VTA* Ventral tegmental area and *VNS* vestibular nuclei superior

with the exception of the cerebellum and the spinal cord, very few CNS regions contain tissue-specific or highly specific genes as defined by these criteria (see "Materials and methods"). Examples of genes expressed in a tissue-specific fashion across the CNS are shown in Fig. 6. In particular, the cerebellum possesses a large number of genes expressed in a tissue-specific (N=156) or highly specific (N=170) pattern, reflecting a unique function and structure for this CNS region. A preview of these findings was revealed earlier in this study by both PCA and hierarchical clustering (Figs. 1 and 2), which showed that the gene expression profile of the cerebellum is very distinct from the other 19 CNS sites analyzed.

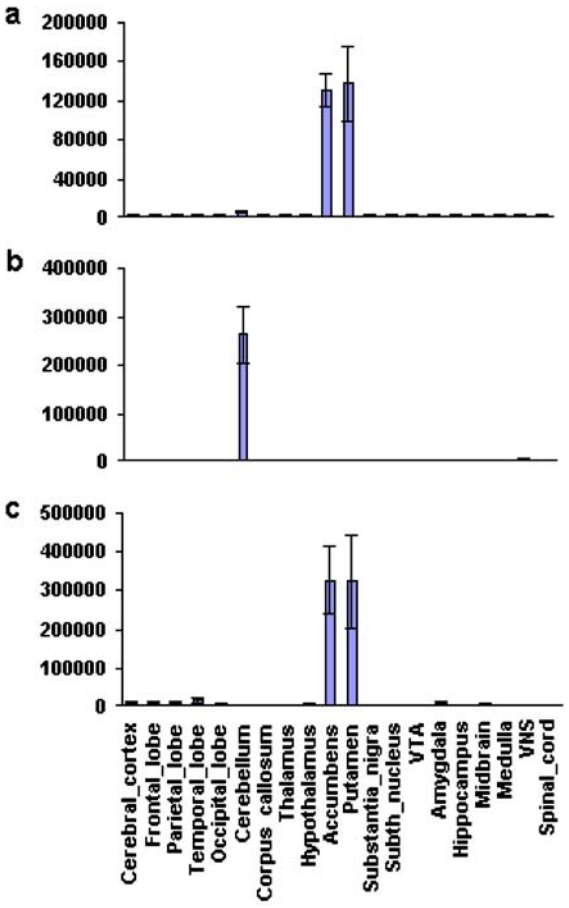

Fig. 7 Real time RT-PCR validation. Real time RT-PCR was used to validate the expression profiles of several genes identified as tissue-specific, highly specific, cluster-specific, and pan-CNS-specific by transcriptional profiling. *Y-axis*, relative expression as $2^{(40-\text{deltaCt})}$; *X-axis*, individual tissues; and *error bars* are SEMs. **a** *ADORA2A* Adenosine A2a receptor. **b** *CRTAM* Class-I MHC-restricted T cell-associated molecule. **c** *GPR6* G protein-coupled receptor 6. *VTA* Ventral tegmental area and *VNS* vestibular nuclei superior

To validate the identification of tissue-specific, highly specific, cluster-specific, and pan-CNS-specific genes first revealed through transcriptional profiling, several genes were selected for real time reverse transcriptase polymerase chain reaction (RT-PCR) (TaqMan) analysis. In all cases, the expression profiles as determined by real time RT-PCR, confirmed the initial microarray results. Figure 7 shows the real time RT-PCR results for the three genes *ADORA2A* (cluster-specific), *CRTAM* (tissue-specific), and *GPR6* (cluster-specific) profiled across the CNS.

Taken together, these data aid in the identification of genes that help define the unique attributes of the human CNS, enabling insight into the processes that determine neurological activity at a system, regional, and site-specific level. One approach to help validate both our methodology and hypothesis that CNS structure–function can be characterized at the molecular level using transcriptional profiling is to use network analysis to verify that specific sets of genes identified in our analyses as system-, regional-, or site-specific, populate pathways known to be present in the CNS, its subregions, and specific tissues, respectively.

Network analysis of the cerebellum

Using the pan-CNS, cluster-specific, and tissue/highly specific gene sets detailed in supplemental Tables 1, 2, and 3 we attempted to identify the networks, functions, and pathways that help define the attributes and features of the CNS, each CNS cluster (see Fig. 3), and each CNS tissue, respectively. To perform this analysis, we used the Ingenuity Pathways Knowledge Base (IPKB; Ingenuity, Mountain View, CA, USA), a database of molecular and functional interactions modeled into a global network, to reveal networks and pathways that are upregulated in either the CNS relative to non-CNS tissues, a single CNS cluster relative to all other CNS clusters, or in a single CNS site relative to all other CNS sites profiled. As an example, the set of 1,938 cerebellum cluster-specific probe sets was queried using IPKB to identify networks and pathways overrepresented in the cerebellum. It should be noted that the two most significantly scoring networks identified consisted entirely of genes found in our cerebellum cluster-specific set as no additional genes from the IPKB were required to generate each network (Fig. 8). This finding provides strong evidence that the pathways and functions that make up these networks are overrepresented in the cerebellum. The Fisher's exact test (see "Materials and methods") was used to show that several of the functions found in these networks as defined by IPKB are over-represented in the cerebellum, including neuronal cell differentiation ($p=7.52\times10^{-4}$), neuronal cell development ($p=7.58\times10^{-4}$), cell survival ($p=1.21\times10^{-5}$), tissue development ($p=1.59\times10^{-5}$), and apoptosis ($p=7.53\times10^{-5}$). These functions are not unique to the cerebellum but many of the genes that constitute these functions are over-represented in the cerebellum and as such, they help to define attributes of the cerebellum that distinguish it from other CNS tissues. Furthermore, several pathways that are important to the function of the cerebellum were also found to be overrepresented. Granule cells are the most frequently found neuron in the cerebellum and they are glutamatergic in nature. Both the nitrogen metabolism pathway ($p=1.71\times10^{-2}$), which generates L-glutamate, and the glutamate receptor-signaling pathway itself ($p=2.61\times10^{-2}$) were found to be overrepresented in the cerebellum.

Discussion

Using the Affymetrix human (genome-wide) U133 plus 2.0 array, we have profiled 169 samples from 20 regions of the human CNS (Table 1), by far, the most extensive CNS transcriptional profiling study to date. We have also profiled 184 samples from 45 non-CNS tissues (Table 2), allowing us to compare gene expression in the CNS relative to the rest of the human body. The complete database currently includes 353 samples covering 65

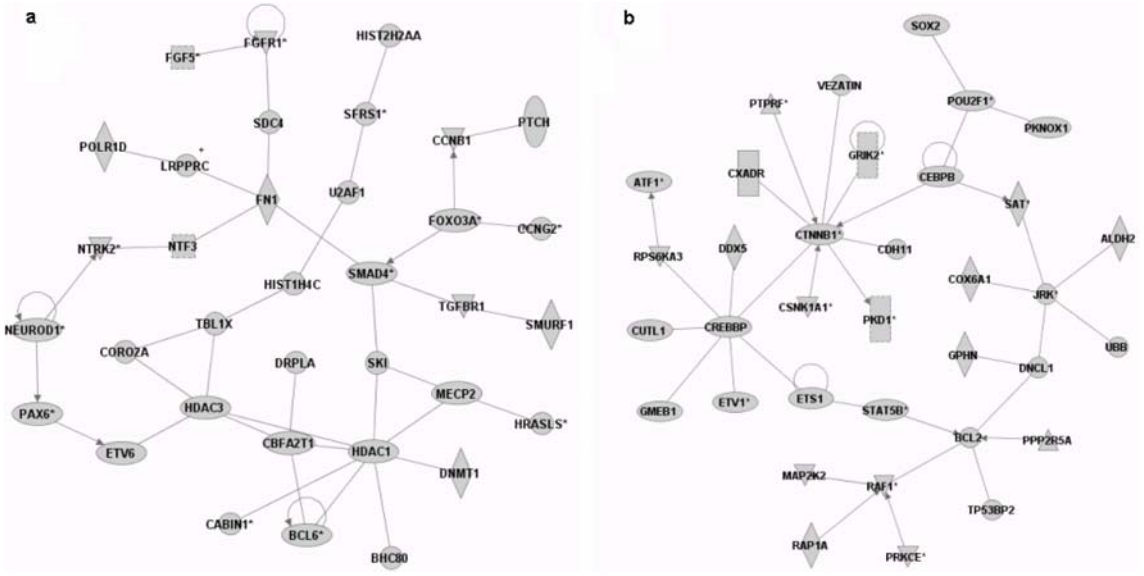

Fig. 8 Cerebellum cluster-specific network analysis. Several networks were found to be overrepresented in the cerebellum, most notably, the networks shown. Both *networks a* and *b* consist entirely of genes found in the cerebellum cluster-specific set and do not require the addition of supplementary genes found in the protein knowledge base to connect the network members together. These two networks are intimately involved in the development, growth, and maintenance of the cerebellum. Network annotation in terms of node shape: *circle*, other; *hashed square*, growth factor; *hashed rectangle*, ion channel; *diamond*, enzyme; *oval*, transcription factor; *triangle facing down*, kinase; and *triangle facing up*, phosphatase. Edge types: *line with arrow*, acts on and *line without arrow*, binds to. *Multiple probe sets in dataset for this gene

different tissues. This database is an important resource for the drug discovery and development process. By defining the baseline expression profiles of 20 distinct sites within the CNS, we can now use this database as a tool to find potential novel drug targets for neurological disorders by comparison to diseased samples and identify genes differentially expressed between normal and diseased tissues. Furthermore, drug target specificity can be assessed by comparing the expression profiles of all tissues represented in this database for each particular gene of interest. In the same way, potential adverse side effects resulting from expression of a target molecule in a nontarget tissue can be revealed before the development process.

Before performing gene expression analyses, we initially characterized the within-tissue variability among samples and the within-individual variability among tissues by calculating the percent coefficients of variation (%CV) for all probe sets present on the U133 plus 2.0 array and subsequently condensed these data into global median % CV values for each tissue and individual (data not shown). It is interesting to note that variability within tissues and within individuals was similarly small yet sufficient enough to make it possible to identify gene expression differences between tissues. Based on these data, we subsequently performed unsupervised analyses to compare the global expression signatures of the 65 tissues profiled.

It is important to note that we were able to show through unsupervised pattern-recognition algorithms (PCA and hierarchical clustering) that gene expression data from 20 different CNS regions are very distinct from nonnervous tissues and closely parallel known anatomical and/or functional interactions within the CNS [25]. This implies that it should be possible to identify genes expressed by each of these regions that should help define their functionality. It is well documented that genes specifically or highly expressed in a given organ are often closely associated with the functions of that particular organ [14, 16]. Thus, our database may allow the identification of sets of genes whose expression is associated with the function of each anatomical region of the CNS.

We have queried our data to identify genes that are (1) system-specific, expressed exclusively across most, if not all, tissues in the CNS; (2) region-specific, expressed in a subset of areas in the CNS that are functionally and/or anatomically linked, and (3) site-specific, expressed exclusively at a single site in the CNS.

CNS-specific genes

By performing a pan-CNS analysis, the anatomical and functional characteristics unique to the CNS were revealed by identifying genes that are expressed ubiquitously within the CNS but to a much lesser extent in nonnervous tissue. Many examples of pan-CNS genes were found. We hypothesize that genes in this category represent, for the most part, housekeeping and homeostatic genes of the CNS, possibly including genes expressed by glial cells that are found throughout the CNS. The pan-CNS set of genes consists of genes of known and unknown functions and by showing that their expression is CNS-wide, we can start to understand the roles of many of the genes with unknown functions. Of those pan-CNS genes with known functions,

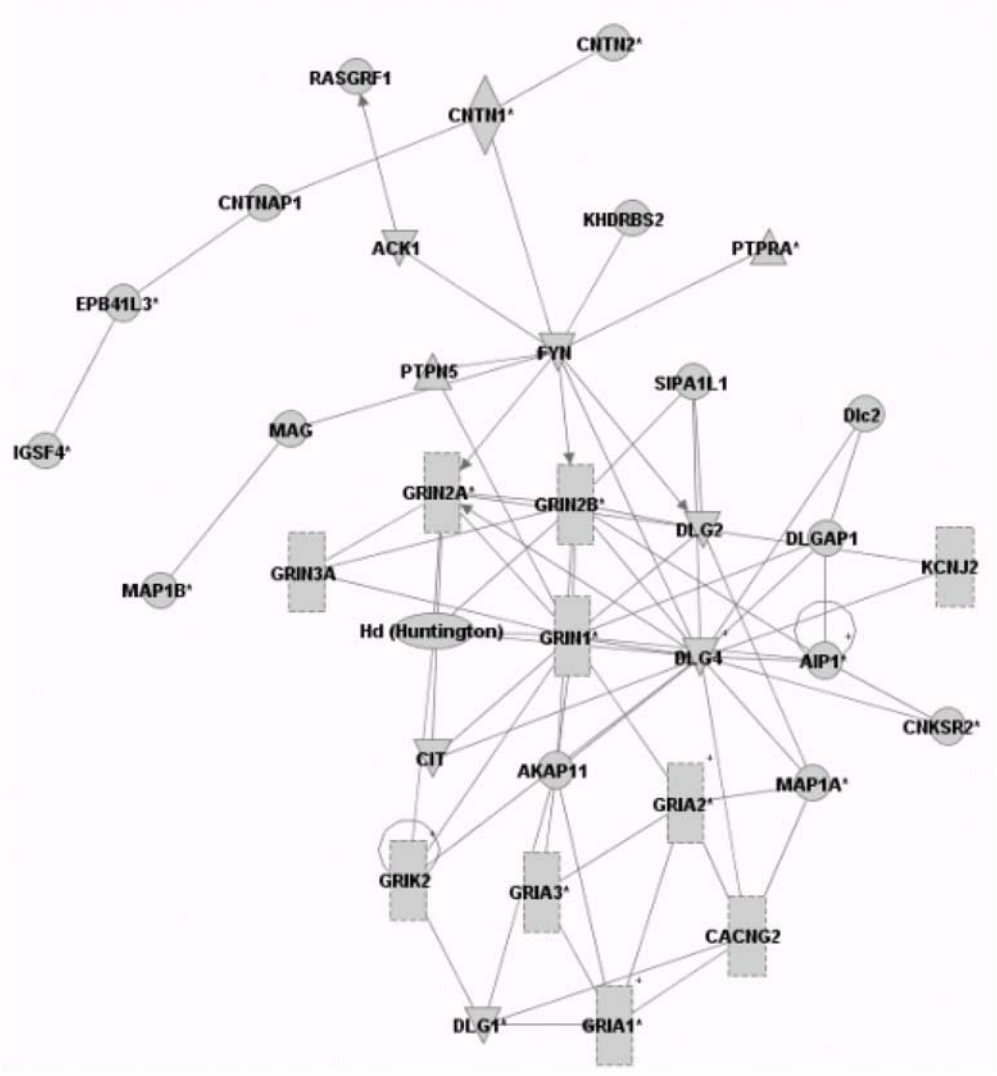

Fig. 9 pan-CNS-specific network analysis. The entire set of pan-CNS-specific probe sets was used to identify networks and pathways overrepresented in the CNS relative to non-CNS tissues using Ingenuity's Protein Knowledge Base. From this analysis, several networks were found to be overrepresented in the CNS, most notably, the network shown here. This network consists entirely of genes found in the pan-CNS-specific set and did not require the addition of supplementary genes found in the protein knowledge base to connect the network members together. Network annotation in terms of node shape: *circle*, other; *hashed square*, growth factor; *hashed rectangle*, ion channel; *diamond*, enzyme; *oval*, transcription factor; *triangle facing down*, kinase; and *triangle facing up*, phosphatase. Edge types: *line with arrow*, acts on and *line without arrow* binds to. *Multiple probe sets in dataset for this gene

there is an extensive array of receptors, ion channels, and solute carriers, as might be expected from the molecular processes used to relay information between different regions of the CNS. Using IPKB, we identified several networks and functions intimately involved in nervous tissue-specific functions from the pan-CNS gene set including signal potentiation, synaptic transmission, vesicle transport, and vesicle endocytosis, which are activities that provide the groundwork for CNS function. For example, the network shown in Fig. 9 consists entirely of pan-CNS genes, many of which are involved in synaptic transmission. These networks provide the link between the function of the CNS and the set of genes identified as pan-CNS-specific, validating this approach as a method to help define CNS function at the molecular level.

Region-specific genes

The accumbens and putamen are anatomically and functionally connected as components of the basal ganglia and under PCA, form a distinct cluster. The amygdala and hippocampus are components of the limbic system and also cluster together. A third cluster consists of the cerebral cortex, frontal lobe, occipital lobe, parietal lobe, and temporal lobe. The cerebellum and spinal cord did not cluster with any of the other 18 CNS sites, suggesting that their expression signatures are distinct within the CNS and reflecting that they functionally carry out unique sets of activities.

Using a Pearson correlation, the clusters defined in Fig. 3 were compared to one another to find genes specifically up- or downregulated within a single cluster relative to all others. These up- and downregulated gene sets help to define the function of each CNS cluster at the molecular level. Using the IPKB, we examined the genes found to be upregulated specifically in the accumbens and putamen cluster. Our data indicate that the expression levels of several genes involved in adenosine and dopamine signaling are significantly upregulated in the accumbens and putamen compared to the rest of the CNS, including the GPCRs *ADORA2A* and *DRD2* (p=7.11×10^{-10}, Fisher's exact test). This is consistent with previous studies that have shown a high level of expression and colocalization of

ADORA2A and DRD2 in these regions [26–28]. By performing this type of network analysis on all CNS cluster-specific gene sets identified in Fig. 3, several examples of overrepresented pathways were found, highlighting the potential for this approach as a tool that can be used to provide a molecular characterization of the known functions of distinct regions of the CNS (data not shown).

One caveat to this approach is that as neuronal processes such as axons and dendrites often overlap multiple regions of the CNS, our ability to find differences between regions is reduced. An alternate approach would be to dissect the CNS into individual ganglia and tracts or even individual cell types by laser capture microscopy before analysis, a process that presents significant technical challenges. We chose to identify the similarities and differences found among anatomically distinct regions of the CNS whose functions ultimately result from the interplay of the multiple cell types found within each region.

Site-specific genes

By restricting our analysis to find genes expressed in single sites of the CNS, we identified 362 probe sets that meet these criteria. Most striking was the finding that the cerebellum and to a significantly lesser extent, the spinal cord, were the only CNS sites that contained multiple genes showing tissue-specific or highly specific expression patterns.

The cerebellum and spinal cord were the two CNS sites that diverged earliest from the remaining CNS sites during hierarchical clustering as shown in Fig. 3. Network analysis (IPKB) showed that the cerebellum-specific gene set contained a preponderance of glutamate receptor-signaling pathway members (p=9.05×10^{-4}, Fisher's exact test). The most numerous neurons found in the cerebellum are granule cells, which are glutamatergic in nature. We also identified pathways involved in cellular maintenance, such as metabolism of cellular building blocks (nucleotide metabolism and amino acid metabolism), cell cycle control (G2/M and G1/S checkpoint regulation), apoptosis, and signal transduction as upregulated in this tissue (data not shown), providing insight into the features that distinguish the cerebellum from all other CNS regions. Further investigation is required to elucidate the role played by these potential networks in cerebellar function.

It is interesting to note that during our analysis, we identified a subset of genes that appeared to be cluster-specific or site-specific but were not included in the cluster-specific or site-specific gene sets because their variability was too high. The current dataset was generated using samples from five male and five female donors. This led us to surmise that gender-specific differences in expression across the genome might be responsible for increasing the variability of a subset of genes. There have been reports showing gender-specific differences in gene expression in CNS tissues and the prevalence of certain CNS diseases that are more common in one gender than in the other [29–33]. We have grouped the samples by gender and have begun an analysis similar to that described in this study and our preliminary data indicates that there are several genes displaying gender-specific gene expression.

In conclusion, gene expression in the CNS is clearly distinct from other sites in the body and reflects the unique make-up of CNS tissues. Furthermore, the CNS itself can be broadly segregated into known functional and anatomical groupings based on these data. By conducting this study of gene expression across 20 distinct regions of the human CNS, we have identified many of the genes that (1) distinguish the CNS from all other tissues, (2) are expressed distinctly among subsets of regions that share a common structure and function, and (3) provide the molecular basis for a particular site's distinct features. As such, our human CNS gene expression dataset reflects the differing functions of each region and is a starting point for a greater molecular understanding of the functional roles played by these regions.

Materials and methods

Tissue procurement

Tissue samples from ten postmortem donors (five female and five male) were procured from Zoion Diagnostics (Hawthorne, NY, USA). All donors were Caucasian and ranged in age from 23 to 53 (donor 1: 25-year-old male, donor 2: 38-year-old male, donor 3: 39-year-old female, donor 4: 30-year-old male, donor 5: 35-year-old male, donor 6: 52-year-old male, donor 7: 50-year-old female, donor 8: 48-year-old female, donor 9: 53-year-old female, and donor 10: 23-year-old female). All tissue samples were obtained after the approval of a local Institutional Review Board and appropriate informed consents were given. All donors were free of chronic disease and in all cases, death was the result of a sudden event. Autopsies were performed within 3 h postmortem and all tissue samples were flash frozen in liquid N_2 and stored at −80°C within 8.5 h postmortem (maximum PMI of 8.5 h). Hematoxylin and eosin-stained slides of tissues were used to confirm tissue type. From seven to nine biological replicates per CNS site and at least three biological replicates per non-CNS tissue were profiled to measure gene expression, totaling 169 CNS and 184 non-CNS samples, respectively. Tables 1 and 2 summarize the number of biological replicates profiled for each of the 65 tissues represented in this study.

RNA preparation, Affymetrix GeneChip hybridization, and real time RT-PCR

Total RNA was isolated using a standard Trizol protocol, checked for integrity by gel electrophoresis, and further purified using RNeasy columns (Qiagen, Valencia, CA, USA). Five micrograms of total RNA from each sample were used to direct first strand cDNA synthesis using a T 7-oligo (dT) 24 primer and PowerScript reverse transcriptase (Clontech, Palo Alto, CA, USA). After second strand

synthesis and clean-up with a Qiaquick spin column (Qiagen), the double-stranded cDNA was used in a MEGAscript T7 RNA polymerase in vitro transcription reaction (Ambion, Austin, TX, USA) containing biotin-labeled ribonucleotides CTP and UTP. The resulting labeled cRNAs were fragmented and hybridized in Affymetrix Human Genome U133 Plus 2.0 arrays (Affymetrix, Santa Clara, CA, USA) overnight, washed, and scanned according to the manufacturer's protocols using a GeneChip Scanner 3000 and GeneChip Operating Software (Affymetrix).

For real time RT-PCR, total RNAs were converted into single-stranded cDNAs using the High-Capacity cDNA Archive Kit (Applied Biosystems Group, Foster City, CA, USA) according to the manufacturer's instructions. PCR was performed on the ABI PRISM 7900HT Sequence Detection System in 384 well plates. TaqMan Universal PCR MasterMix and Assays-on-Demand Gene Expression probes (Applied Biosystems) were used for the PCR step according to the manufacturer's instructions. Expression values obtained were corrected for loading by measuring 18S RNA relative expression levels.

Data processing and analysis

Affymetrix Human Genome U133 Plus 2.0 arrays (Affymetrix) were used throughout this study. Each array contains 54,675 probe sets representing 47,000 unique transcripts and more than 21,000 UniGene clusters. Affymetrix quality control metrics were used to remove arrays from downstream analyses that did not meet the threshold criteria. All remaining Affymetrix .cel files were background corrected, quantile-normalized, polished by RMA [23, 24], and subsequently analyzed using Expressionist Pro2 (Genedata, Basel, Switzerland) software. All statistical analyses were carried out on $\log_2$-transformed data. Data were retransformed back to linear to show fold change values between tissues and to depict figures showing relative expression values.

Principal component and hierarchical clustering analyses were performed using the mean values of each tissue obtained by averaging the expression values of biological replicates. For PCA, a covariance matrix was used; hierarchical clustering was performed using a correlation metric with complete linkage. One-way ANOVA was performed across all samples grouped by tissue and was used to filter genes that did not show a significant difference in expression across the tissues profiled. A p value of <0.05 with a Bonferroni correction for multiple testing was used to define the filter threshold. From this resulting set of probe sets, pan-CNS-specific and cluster-specific genes were identified using a pattern matching method with the Pearson correlation. Genes that showed at least a twofold increase or decrease in expression across the CNS relative to all non-CNS tissues with a similarity coefficient score of greater than 0.7 were defined as pan-CNS-specific. Genes that showed at least a twofold increase or decrease in expression in a single cluster relative to all other CNS clusters with a similarity coefficient score of greater than 0.7 were defined as cluster-specific.

Table 4 Criteria for selection of tissue-specific and highly specific probe sets

Criteria	Tissue-specific	Highly specific
Expression rank (TOI)	1	1
Expression level (TOI)	>100	>100
Expression level of tissue ranked #2	<100	–
Affymetrix present call % (TOI)	≥50	≥50
Standard deviation (TOI)	<1/2 TOI mean	<1/2 TOI mean
Expression ratio, TOI/tissue ranked #2	≥2	≥2

TOI Tissue of interest

Probe sets were identified as either tissue-specific or highly specific according to the six criteria described in Table 4. Briefly, for each probe set, the tissue showing the highest expression level was identified and this tissue was denoted as the tissue of interest (TOI). The expression value of the probe set in the TOI had to be at least 100 and for the tissue with the second highest expression level, expression had to be less than 100 (assuring the remaining 18 CNS tissues also had expression levels of less than 100). Next, at least 50% of the TOI's biological replicates were required to be called present by Affymetrix's absent/present algorithm. In addition, the TOI's standard deviation had to be less than half of its mean expression value. Finally, the TOI's expression value had to be at least two times greater than the tissue with the second highest expression value. If these thresholds were met, the probe set was denoted as tissue-specific. A highly specific probe set had to meet all of the above criteria as well, except that the tissue with the second highest expression level could have an expression value that was greater than 100. Each of the probe sets on the Affymetrix Human Genome U133 plus 2.0 array were analyzed in this fashion.

Network analysis

To identify networks, functions and pathways over-represented in specific gene sets of interest, Ingenuity's Pathways Analysis and Pathways Knowledge Base were used. Gene sets defined as pan-CNS-specific, cluster-specific, and site-specific were uploaded to Ingenuity's Pathways Analysis web-based interface and subsequently analyzed using the IPKB. This knowledge base contains millions of interactions that were mostly hand-curated and combined to build global networks. The genes from each gene set of interest were used to populate the networks found in the IPKB. Networks resulting from this analysis for a particular gene set of interest are ranked according to how relevant they are to the genes in each set, i.e., the greater the number of genes from a gene set of interest found in a network, the more relevant the network is to the

gene set. A Fisher's exact *t* test was used in conjunction with a 2×2 contingency table to identify functions and pathways significantly overrepresented in gene set's of interest ($p<0.05$ cutoff). Fisher's exact *t* test *p* values were not corrected for multiple testing to minimize type II errors. For each Fisher's exact test performed, pathways analyzed were rank ordered by *p* value and pathways ranked highest (lowest *p* values) were considered overrepresented in each analysis group.

Acknowledgement We thank Dr. Richard A. Maki for helpful discussion and critical reading of the manuscript.

References

1. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A, Sampas N, Dougherty E, Wang E, Marincola F, Gooden C, Lueders J, Glatfelter A, Pollock P, Carpten J, Gillanders E, Leja D, Dietrich K, Beaudry C, Berens M, Alberts D, Sondak V (2000) Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 406:536–540
2. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D (2000) Molecular portraits of human breast tumours. Nature 406:747–752
3. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R (2002) A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347:1999–2009
4. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH (2002) Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530–536
5. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET (2003) Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci U S A 100:10393–10398
6. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98:10869–10874
7. Bertucci F, Finetti P, Rougemont J, Charafe-Jauffret E, Nasser V, Loriod B, Camerlo J, Tagett R, Tarpin C, Houvenaeghel G, Nguyen C, Maraninchi D, Jacquemier J, Houlgatte R, Birnbaum D, Viens P (2004) Gene expression profiling for molecular characterization of inflammatory breast cancer and prediction of response to chemotherapy. Cancer Res 64: 8558–8565
8. Brennan DJ, O'Brien SL, Fagan A, Culhane AC, Higgins DG, Duffy MJ, Gallagher WM (2005) Application of DNA microarray technology in determining breast cancer prognosis and therapeutic response. Expert Opin Biol Ther 5:1069–1083
9. Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, Pienta KJ, Rubin MA, Chinnaiyan AM (2001) Delineation of prognostic biomarkers in prostate cancer. Nature 412:822–826
10. Welsh JB, Sapinoso LM, Su AI, Kern SG, Wang-Rodriguez J, Moskaluk CA, Frierson HF Jr, Hampton GM (2001) Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. Cancer Res 61:5974–5978
11. Heighway J, Knapp T, Boyce L, Brennand S, Field JK, Betticher DC, Ratschiller D, Gugger M, Donovan M, Lasek A, Rickert P (2002) Expression profiling of primary non-small cell lung cancer for target identification. Oncogene 21:7749–7763
12. Velasco AM, Gillis KA, Li Y, Brown EL, Sadler TM, Achilleos M, Greenberger LM, Frost P, Bai W, Zhang Y (2004) Identification and validation of novel androgen-regulated genes in prostate cancer. Endocrinology 145:3913–3924
13. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A, Patapoutian A, Hampton GM, Schultz PG, Hogenesch JB (2002) Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A 99:4465–4470
14. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, Cooke MP, Walker JR, Hogenesch JB (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A 101:6062–6067
15. Haverty PM, Weng Z, Best NL, Auerbach KR, Hsiao LL, Jensen RV, Gullans SR (2002) HugeIndex: a database with visualization tools for high-density oligonucleotide array data from normal human tissues. Nucleic Acids Res 30:214–217
16. Son CG, Bilke S, Davis S, Greer BT, Wei JS, Whiteford CC, Chen QR, Cenacchi N, Khan J (2005) Database of mRNA gene expression profiles of multiple human organs. Genome Res 15:443–450
17. Shyamsundar R, Kim YH, Higgins JP, Montgomery K, Jorden M, Sethuraman A, van de Rijn M, Botstein D, Brown PO, Pollack JR (2005) A DNA microarray survey of gene expression in normal human tissues. Genome Biol 6:R22
18. Hsiao LL, Dangond F, Yoshida T, Hong R, Jensen RV, Misra J, Dillon W, Lee KF, Clark KE, Haverty P, Weng Z, Mutter GL, Frosch MP, Macdonald ME, Milford EL, Crum CP, Bueno R, Pratt RE, Mahadevappa M, Warrington JA, Stephanopoulos G, Gullans SR (2001) A compendium of gene expression in normal human tissues. Physiol Genomics 7:97–104
19. Saito-Hisaminato A, Katagiri T, Kakiuchi S, Nakamura T, Tsunoda T, Nakamura Y (2002) Genome-wide profiling of gene expression in 29 normal human tissues with a cDNA microarray. DNA Res 9:35–45
20. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863–14868
21. Misra J, Schmitt W, Hwang D, Hsiao LL, Gullans S, Stephanopoulos G (2002) Interactive exploration of microarray gene expression patterns in a reduced dimensional space. Genome Res 12:1112–1120
22. Raychaudhuri S, Stuart JM, Altman RB (2000) Principal components analysis to summarize microarray experiments: application to sporulation time series. Pac Symp Biocomput 455–466
23. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4:249–264
24. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP (2003) Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31:e15
25. Nolte J (2002) The human brain. An introduction to its functional anatomy, 5th edn. Mosby, St. Louis, Missouri
26. Standaert DG (2003) Adenosine A2A receptor modulation of motor systems for symptomatic therapy in Parkinson's disease. Neurology 61:S30–S31
27. Ferre S, Fredholm BB, Morelli M, Popoli P, Fuxe K (1997) Adenosine-dopamine receptor-receptor interactions as an integrative mechanism in the basal ganglia. Trends Neurosci 20:482–487

28. Fredholm BB, Svenningsson P (2003) Adenosine–dopamine interactions: development of a concept and some comments on therapeutic possibilities. Neurology 61:S5–S9
29. Baum LW (2005) Sex, hormones, and Alzheimer's disease. J Gerontol A Biol Sci Med Sci 60:736–743
30. Czlonkowska A, Ciesielska A, Gromadzka G, Kurkowska-Jastrzebska I (2005) Estrogen and cytokines production—the possible cause of gender differences in neurological diseases. Curr Pharm Des 11:1017–1030
31. Duquette P, Pleines J, Girard M, Charest L, Senecal-Quevillon M, Masse C (1992) The increased susceptibility of women to multiple sclerosis. Can J Neurol Sci 19:466–471
32. McGrath J, Saha S, Welham J, El Saadi O, MacCauley C, Chant D (2004) A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. BMC Med 2:13
33. Nelson LM (1995) Epidemiology of ALS. Clin Neurosci 3:327–331