

High-resolution gene expression atlases for adult and developing mouse brain and spinal cord

Alex M. Henry · John G. Hohmann

Received: 3 April 2012 / Accepted: 21 June 2012 / Published online: 26 July 2012
© Springer Science+Business Media, LLC 2012

Abstract Knowledge of the structure, genetics, circuits, and physiological properties of the mammalian brain in both normal and pathological states is ever increasing as research labs worldwide probe the various aspects of brain function. Until recently, however, comprehensive cataloging of gene expression across the central nervous system has been lacking. The Allen Institute for Brain Science, as part of its mission to propel neuroscience research, has completed several large gene-mapping projects in mouse, nonhuman primate, and human brain, producing informative online public resources and tools. Here we present the Allen Mouse Brain Atlas, covering ~20,000 genes throughout the adult mouse brain; the Allen Developing Mouse Brain Atlas, detailing expression of approximately 2,000 important developmental genes across seven embryonic and postnatal stages of brain growth; and the Allen Spinal Cord Atlas, revealing expression for ~20,000 genes in the adult and neonatal mouse spinal cords. Integrated data-mining tools, including reference atlases, informatics analyses, and 3-D viewers, are described. For these massive-scale projects, high-throughput industrial techniques were developed to standardize and reliably repeat experimental goals. To verify consistency and accuracy, a detailed analysis of the 1,000 most viewed genes for the adult mouse brain (according to website page views) was performed by comparing our data with peer-reviewed literature and other databases. We show that our data are highly consistent with independent sources and provide a comprehensive compendium of information and tools used by thousands of researchers each month. All data and tools are freely available via the Allen Brain Atlas portal (www.brain-map.org).

Introduction

Developing fundamental knowledge about the biology of the brain is essential if we are ever to understand how the most complex structure known to humanity enables us to breathe, think, and make sense of our world. While the rate of progress in neuroscience increases with each year, one stumbling block that remains is the lack of unifying information to provide foundational data for brain research. What is needed is comprehensive data on how the brain is constructed: the basic parts list. How does each gene function and where is it active in the brain? What are the functional units of information processing within the brain? How many cell types are there and how are they wired together? Fortunately, efforts to generate this “parts list” are underway in many large-scale investigations funded by both private and public sources. Genomes have been mapped for many species (Lander et al. 2001; Waterston et al. 2002), and microarray and RNAseq studies have revealed where genes are turned on and off across large brain areas (Kang et al. 2011; Lindblad-Toh et al. 2011; Shimogori et al. 2010). A truly massive effort is ongoing to map connectivity in the human brain using diffusion imaging and various MRI techniques (Van Essen et al. 2012), and the Allen Institute and others are mapping the connectome of the mouse brain with conventional and Cre-dependent viral tracers (Madisen et al. 2010). Since the inception of the Allen Institute, we have been forging ahead on numerous fronts to provide as much information as possible about the molecular neuroanatomy of the mammalian brain. We have produced detailed histological reference atlases of the developing and adult mouse brain (Dong 2008) and spinal cord (Watson et al. 2008) and are in the process of creating prenatal and adult human brain reference atlases. A gene expression map of approximately

A. M. Henry (✉) · J. G. Hohmann
The Allen Institute for Brain Science, Seattle, WA 98103, USA
e-mail: alexh@alleninstitute.org

500 regions per hemisphere of the adult human brain via microarray, and RNAseq data for 16 critical brain regions across several prenatal and postnatal ages are also available via the Allen Brain Atlas data portal (www.brain-map.org).

Our most longstanding projects are our high-resolution in situ hybridization (ISH) studies in mouse, human, and macaque monkey (Bernard et al. 2012; Lein et al. 2007; Zeng et al. 2012). Using industrial-scale methodologies, with careful and exhaustive quality control to ensure repeatability and accuracy, these studies provide a wealth of fundamental information about how the brain employs its ~20,000 available genes to construct and control a normal nervous system. In this review, we detail our efforts in the adult and developing mouse brain and spinal cord and describe the data, tools, and analyses of these projects (Box 1).

Allen mouse brain atlas

Since its original conception, the Allen Mouse Brain Atlas has grown from an ambitious idea into a prominent resource within the neuroscience community. The atlas uses gene-specific ISH transcripts to localize expression in the brains of adult C57BL/6J mice (for more detailed methods, see Lein et al. 2007). By employing a standardized, high-throughput production pipeline, the Allen Institute has been able to produce a vast catalog of comparable experiments. Since its first public data release in December 2004, the atlas has grown to contain 21,706 sagittal and 4,342 coronal ISH experiments for ~20,000 genes. For each of these experiments, brain sections were scanned at 10× in order to provide cellular resolution in all neuro-anatomical structures. Enhancing the value of the ISH experiments is an assortment of informatics tools designed to help a user visualize and sort the data in order to answer specific scientific questions. These tools include (1) an integrated reference atlas based on the Allen Reference Atlas (ARA) (Dong 2008); (2) informatics-based searches for genes with specific expression patterns; (3) the Anatomic Gene Expression Atlas (AGEA), an interactive viewer of brain-wide expression pattern correlations (Ng et al. 2009); and (4) the Brain Explorer 3-D viewer application (Lau et al. 2008) (Fig. 1). Including ISH experiments and analysis tools, approximately 17 terabytes of adult mouse brain data have been published online.

The integrated reference atlas is unquestionably the foundational framework within the Allen Mouse Brain Atlas. It is a high-resolution color atlas drawn over 21 sagittal and 132 coronal Nissl sections. It includes a systematic hierarchical ontology of brain structures. The ARA has been integrated into all other analysis tools in order to provide a consistent anatomic context for gene expression.

Box 1 Overview of Allen Brain Atlas Resources for the Mouse Central Nervous System. The following resources and tools are all freely available via the Allen Brain Atlas data portal at www.brain-map.org. Informational videos and tutorials demonstrating use of these resources are available via the data portal, and on YouTube (<http://www.youtube.com/user/AllenInstitute>)

Allen Mouse Brain Atlas

- Image-based in situ hybridization data
- Genome-wide coverage
- Comprehensive anatomic coverage
- Microscopic resolution down to the cellular level
- Powerful data search and visualization tools, including NeuroBlast and AGEA
- Interactive Brain Explorer® 3-D viewer
- Anatomic reference atlas, drawn by Hong-Wei Dong, MD, PhD

Allen Developing Mouse Brain Atlas

- Image-based in situ hybridization data for approximately 2,000 genes
- Anatomic coverage across the entire brain
- Coverage of multiple stages of development
- Microscopic resolution down to the cellular level
- Anatomic and temporal search
- NeuroBlast and AGEA
- Interactive Brain Explorer 3-D viewer
- Developmental reference atlases, drawn by Luis Puelles, MD, PhD

Allen Spinal Cord Atlas

- Image-based in situ hybridization data
- Genome-wide coverage
- Anatomic coverage spanning the full length of the spinal cord
- Data for neonate (P4) and adult (P56) stages
- Microscopic resolution down to the cellular level
- Nissl and acetyl cholinesterase-stained reference series
- Anatomic reference atlases, drawn by Charles Watson, MD and Gulgun Sengul, MD

The reference atlas can be viewed alongside ISH data. By using the “Synchronize with other viewers” tool, any point on an ISH experiment can be matched to the nearest point in the reference atlas (Fig. 1b). In order to make this possible, all ISH experiments have been virtually reconstructed in 3-D and registered to the 3-D volume of the reference atlas.

Another vital feature of the Allen Mouse Brain Atlas is its gene expression detection algorithm. For each ISH experiment, the algorithm identifies pixels that correspond with gene expression and assesses expression intensity. Combining these values with the reference atlas registration information, expression intensity scores can be calculated for any structure in the reference atlas ontology. The compilation of these scores is described as an experiment’s expression profile. Using this measure, genes can

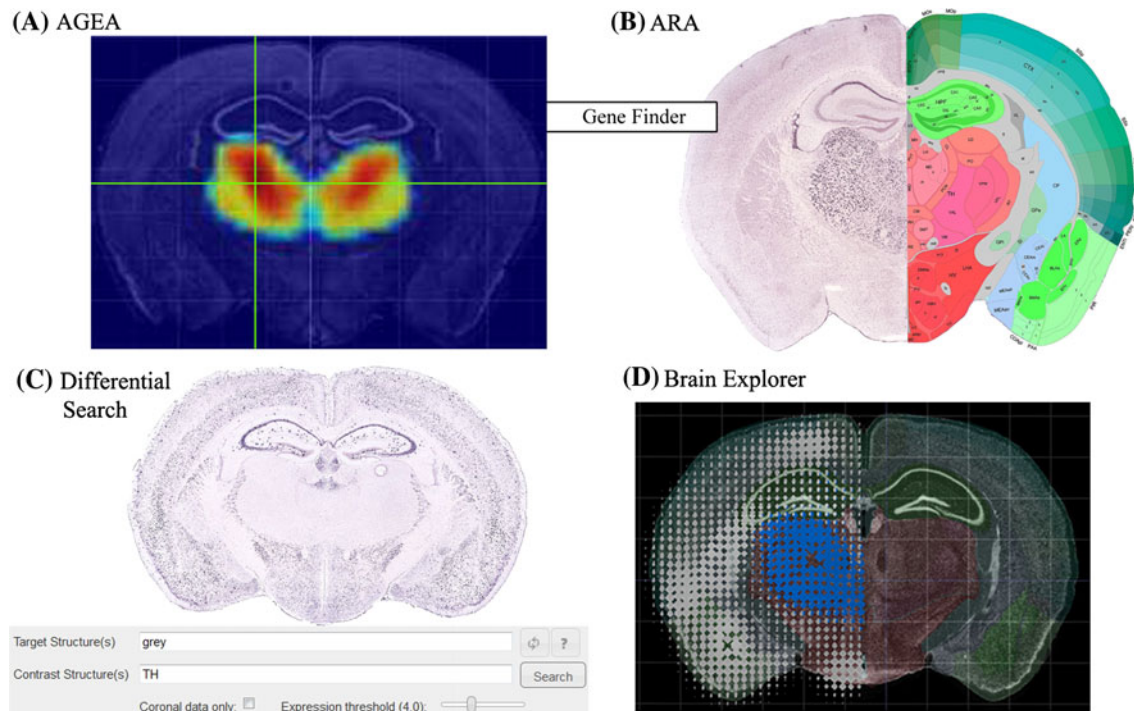


Fig. 1 Illustration of the assortment of Allen Mouse Brain Atlas informatics tools. **a** An AGEA correlation map with a seed voxel in the lateral group of the dorsal thalamus (*green crosshair*). **b** For this seed voxel, the gene *Plekhg1* (*left hemisphere*) is returned by AGEA’s “Gene Finder” feature. A matched ARA plate (*right hemisphere*)

shows that expression exists primarily in the thalamus. **c** The gene *Cnr1* is returned by the “differential search” feature when requesting expression in the whole brain, but NOT in the thalamus. **d** *Plekhg1* (*blue*) and *Cnr1* (*white*) expression visualized using Brain Explorer. The expression patterns of these genes are largely distinct

be classified and searched in different ways. The “NeuroBlast” search feature returns genes with an expression profile similar to that of a specified gene of interest. The “Differential” search returns a list of genes with higher expression in one structure compared to another (Fig. 1c).

Expanding upon the expression profile data calculated for each experiment is the AGEA analysis tool. Fundamentally, AGEA derives correlations between gene expression at single points in the brain and all other points. ISH experiments sectioned in the coronal plane were virtually reconstructed in 3-D and then divided into cubic voxels of 200 μm per side. Selecting a seed voxel will display a heat map showing the correlation of expression in other voxels (Fig. 1a). The reference atlas has been integrated into the AGEA tool such that each voxel has a neuroanatomical identity. However, the correlation maps are derived entirely from gene expression data. Utilizing the collection of correlation maps, a hierarchical binary tree of spatial clusters was generated. At each level of the tree, groups of voxels are dichotomized to achieve maximal dissimilarity of aggregate correlation values. Essentially, each group of voxels, beginning with the whole brain, is divided at each level by the most significant spatial difference in collective gene expression. AGEA’s “Gene Finder” tool returns a list of genes for which expression is specific to the highly correlated voxels in the heat map.

Once a list of genes has been accrued, the interactive Brain Explorer 3-D viewer is useful for visualizing expression. It displays voxel-level gene expression. Because the data have been registered into a canonical 3-D space, multiple experiments viewed simultaneously allow for a comparison of relative expression position (Fig. 1d), with anatomical context provided by 3-D brain structures derived from the reference atlas. Selecting any expressing voxel will show the user a thumbnail of the ISH data at that position and directly link to the high-resolution ISH images in the online Allen Mouse Brain Atlas.

The scope and utility of the Allen Mouse Brain Atlas has attracted a multitude of users. Since the publication of the majority of the data for the atlas in 2006, it has attracted an average of 10,352 unique visitors per month. In order to improve the user experience, regular updates to the atlas are made annually. These typically include replacement of poor-quality experiments and improvements to tools and viewers. According to PubMed, the primary platform paper for the atlas (Lein et al. 2007) has been cited 384 times as of May 2012, reflecting its extensive use as a validation of and catalyst for new research (Jones et al. 2009). For example, a recent project using RNAseq to generate a quantitative atlas of gene expression in the mouse neocortex (Belgard et al. 2011) uses ISH data from the Allen Mouse Brain Atlas to

visualize and verify their approach. Another recent publication informatically relates expression patterns from the Allen Mouse Brain Atlas with connectivity patterns from the Brain Architecture Management System, finding a statistically significant relationship between gene expression and connectivity (French and Pavlidis 2011).

Allen developing mouse brain atlas

To add a temporal dimension to the mouse brain ISH data set, the Allen Developing Mouse Brain Atlas was launched in 2008. Repeating the approach taken for the adult mouse brain atlas described above, the Allen Developing Mouse Brain Atlas details the expression patterns of ~2,000 unique genes at seven significant stages of mouse brain development. This survey comprises 15,066 sagittal and 714 coronal ISH experiments; all of which are displayed at 10× resolution. A set of tools similar to those of the Allen Mouse Brain Atlas has been deployed for this atlas, including (1) a series of anatomic reference atlases; (2) informatics-based searches for genes with specific expression patterns; (3) the Brain Explorer 3-D viewer application; and (4) AGEA. In all, approximately 12 terabytes of data and analysis tools have been made available online as part of this atlas. These data are accessed by an average of 2,903 unique users each month. As with the adult atlas, updates of data, tools and user interface are released each year.

Most genes were selected for inclusion in the atlas based on a known or hypothesized influence on brain development. Gene selection focused on transcription factors; neuropeptides, neurotransmitters, and their receptors; receptor tyrosine kinases and their ligands; axon guidance molecules; and anatomical and cell-type markers. Expression of these genes was assayed at embryonic days 11.5, 13.5, 15.5, and 18.5, and postnatal days 4, 14, and 28. By focusing on these types of genes at these specific ages, the progression of numerous neurodevelopmental events such as neuronal migration and differentiation has been captured. For example, the cerebral cortex is made up primarily of excitatory glutamatergic neurons and inhibitory GABAergic interneurons. Excitatory neurons mature in the ventricular zone of the dorsal pallium and radially migrate a short distance to form the cortical plate. Inhibitory interneurons are born primarily in the subpallium and tangentially migrate a relatively long distance to reach the cortical plate. The gene *Lhx6* expresses in the medial and central ganglionic eminences in the subpallium and marks a subset of interneurons migrating from these regions into cortex (Liodis et al. 2007). Data for *Lhx6* illustrate this key element of cortical development across time (Fig. 2).

A high-resolution, full-color anatomic reference atlas and ontology has been created for the Allen Developing Mouse Brain Atlas. This reference atlas was generated by Luis Puelles and includes annotations of the seven stages of mouse brain development utilized in the gene expression atlas. Instead of following the conventional model of brain neuroanatomy, which centers on defining functionally distinct nuclei and layers, Dr. Puelles has created an atlas set that illustrates the developmental Bauplan of the vertebrate brain. His “prosomeric model” holds that the developing brain is segmentally organized (Rubenstein et al. 1994). These segments, known as neuromeres, are delineated in the reference atlas.

Digital registration to the reference atlas allows gene expression data to be analyzed in many of the same ways described for the Allen Mouse Brain Atlas. A “NeuroBlast” search can be performed for any image series in order to find genes exhibiting similar expression patterns. The “Anatomic” search returns a list of genes that are enriched in a specified structure. The “Temporal” search finds genes that are enriched at a specified age in a specified structure. The Brain Explorer application can be used to view ISH data in 3-D. Within this application all seven developmental ages can be viewed simultaneously, giving an efficient overview of a gene’s expression pattern throughout development. The reference atlas has been integrated into this tool so expression can be visualized in a neuroanatomical context. There is also a version of AGEA derived from these data, which has most of the functionality of the Allen Mouse Brain Atlas version.

While the scope of the Allen Developing Mouse Brain Atlas is limited to approximately 2,000 genes instead of complete genomic coverage, its strategic gene selection, sampling at seven stages of development, and analytical tools make it an exceptional resource for visualizing neurodevelopmental events in the mouse brain.

Allen spinal cord atlas

The Allen Spinal Cord Atlas contains neonatal and adult mouse ISH data for about 20,000 genes spanning the length of the spinal cord. Coronal sections of spinal column at postnatal day 4 and spinal cord at postnatal day 56 were collected and processed with the ISH protocol described in Lein et al. (2007). These data are augmented by the inclusion of a detailed spinal cord reference atlas drawn by Charles Watson and Gulgun Sengul (Sengul et al. 2012; Watson et al. 2008). Over 80 structures in the neonate and adult cords have been delineated in the reference atlas. Each ISH section has an associated “nearest reference section,” which can be accessed by loading an ISH experiment and its age-appropriate reference atlas into the

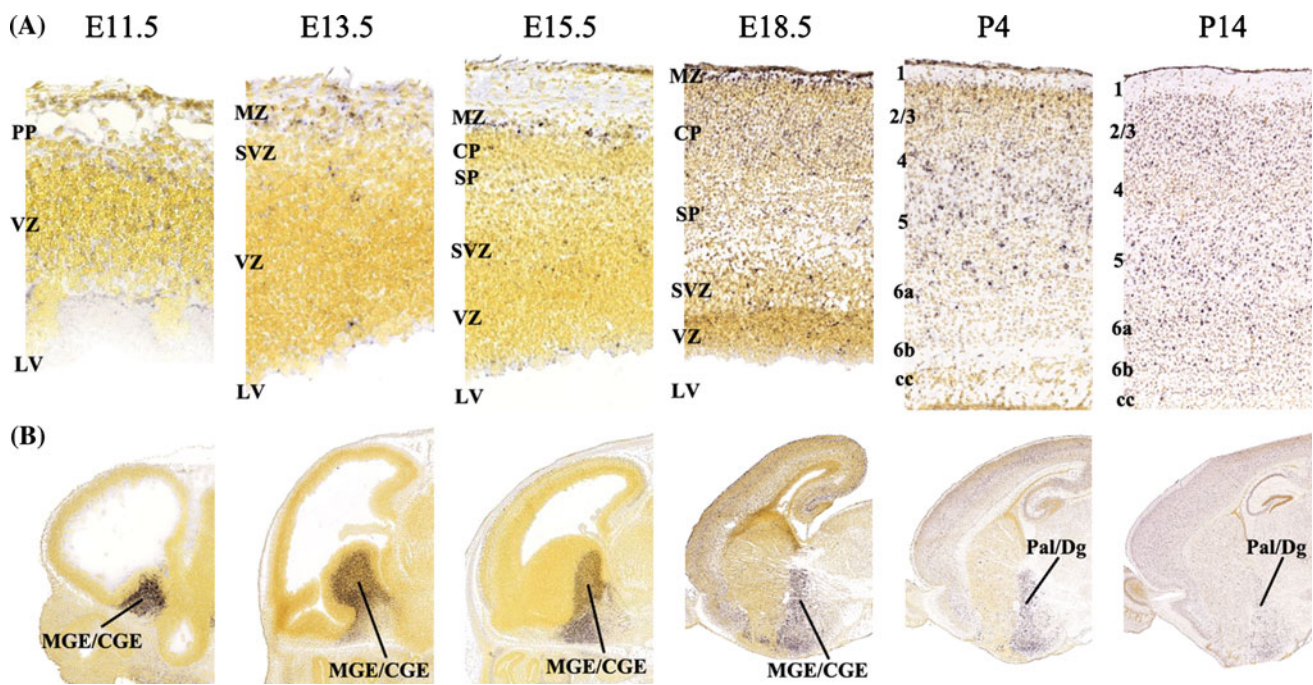


Fig. 2 ISH data for the interneuron marker gene *Lhx6* in the developing mouse brain. **a** Cross sections of pallium from embryonic day 11.5 to postnatal day 14. Migrating interneurons can be found in the marginal zone (MZ) as early as E13.5. By E18.5, *Lhx6*-positive interneurons are found throughout the depth of the pallium. This scattered expression pattern persists throughout adulthood. **b** Sagittal sections of the forebrain from E11.5 to P14. The medial and central ganglionic eminences (MGE and CGE), which are both sites of

cortical interneuron generation, are revealed by *Lhx6* expression. Over developmental time, expression in both areas is diminished as *Lhx6*-positive interneurons migrate to the cortex and cease to be generated. 1, 2/3, 4, 5, 6a, and 6b are distinct layers in the postnatal cortex; *cc* corpus callosum; *CP* cortical plate; *Dg* diagonal domain; *LV* lateral ventricle; *Pal* pallium; *PP* preplate; *SP* subplate; *SVZ* subventricular zone; *VZ* ventricular zone

viewer and using the “Synchronize with other viewers” tool. In addition, all adult spinal cord segments have been classified into cervical, thoracic, lumbar, sacral, or coccygeal groups. Nissl and AChE reference sets have been produced to further assist in the neuroanatomical localization of gene expression (Fig. 3).

The Allen Spinal Cord Atlas has received an average of 1,500 unique visitors each month since its first public data release in July 2008. In all, over 1.7 terabytes of data have been made available online. The atlas has been utilized by many researchers, but one of its most notable impacts was in 2011 when a paper titled “Adult Spinal Cord Radial Glia Display a Unique Progenitor Phenotype” (Petit et al. 2011) was published, describing a putative new cell type discovered based on the atlas data. The study was a collaboration effort with scientists from the University of British Columbia, The Montreal Neurological Institute, and the Allen Institute. It defined a previously unknown subtype of adult radial glia within the spinal cord that expresses a set of 122 genes, which are also expressed in classic neural stem cells. This discovery suggests the existence of previously overlooked stem cell-like populations that may one day assist in the repair of spinal cord injury or in the treatment of spinal cord disease.

Validation of the top 1,000 allen mouse brain atlas genes

While the scope of the Allen Mouse Brain Atlas is impressive, its true value depends on its biological accuracy. In order to characterize and improve data accuracy, the most frequently viewed genes in the atlas were compared to external sources (Allen Institute for Brain Science 2010). Cumulative webpage views from the Allen Mouse Brain Atlas were calculated for 17,825 unique genes from October 2005 to March 2010. A webpage hit was tallied whenever a page containing ISH data for a specific gene was accessed. The 1,000 genes receiving the greatest number of hits were selected for validation. We focused on the top 1,000 in order to preferentially improve accuracy among the most frequently viewed genes, which represented 49 % of all website hits (Fig. 4).

There were two ways in which a gene in the top 1,000 was validated: (1) comparison to peer-reviewed literature or (2) comparison to one of two online databases. The 100 highest-ranked genes were validated by published literature exclusively. These genes are highly studied within the neuroscience community so comparable data were readily available for all 100 (top 50 are shown in Table 1). Comparable data were defined as articles containing images of

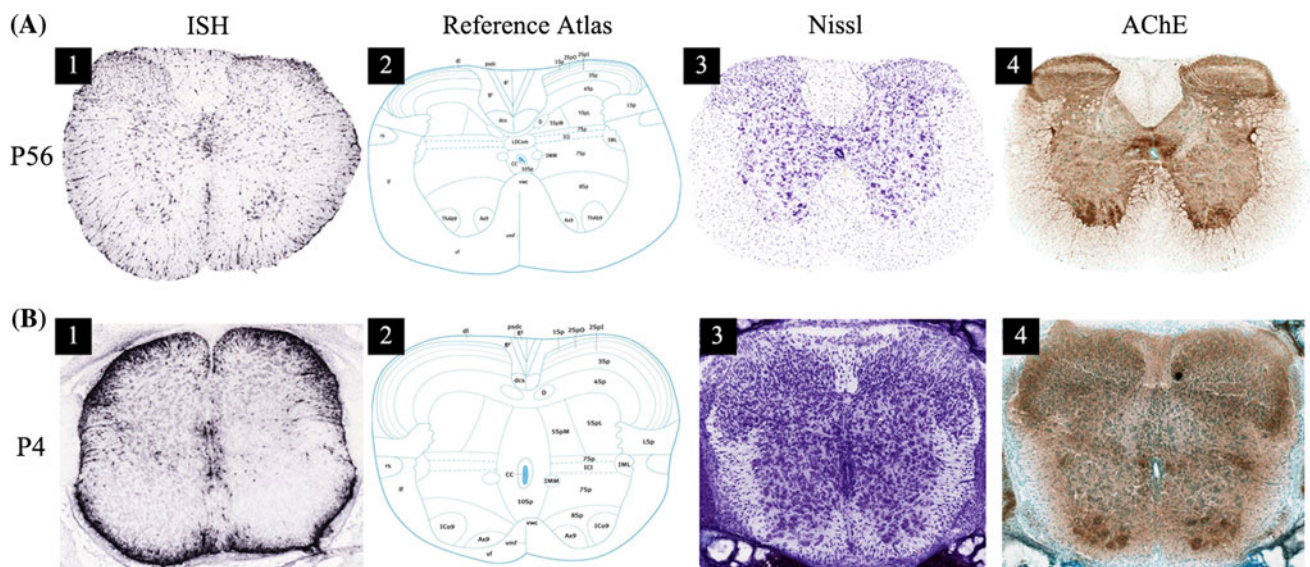


Fig. 3 ISH data and associated reference resources for the spinal cord radial glia marker *Gfap* in the adult (P56) and neonate (P4) mouse. **a1** Adult thoracic spinal cord section in the coronal plane showing *Gfap* expression radially arrayed in the white matter. **2** The nearest reference atlas section showing many divisions of white and gray matter. **3** The nearest reference Nissl section showing the pattern of cell bodies expected for the ISH section. **4** Nearest reference acetyl cholinesterase-stained (AChE) section showing cell bodies and

terminals involved in the cholinergic system. **b1** Thoracic spinal cord section from the P4 mouse showing *Gfap* expression radially arrayed in the white matter and surrounding the central canal. **2** The nearest reference atlas section showing detailed neuroanatomical annotations. **3** The nearest reference Nissl section showing the pattern of cell bodies expected for the ISH section. **4** The nearest reference AChE section showing cell bodies and terminals involved in the cholinergic system

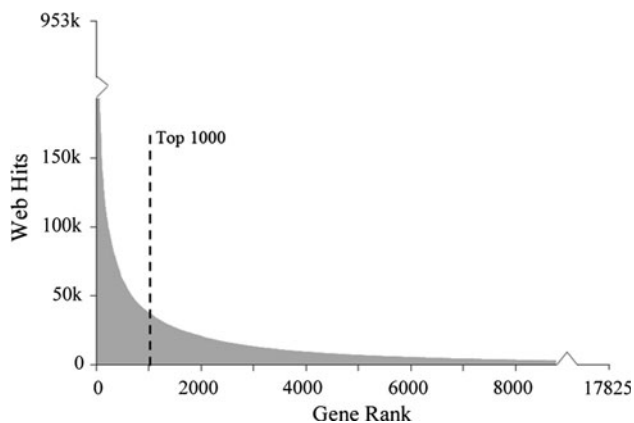


Fig. 4 Website hits shown by gene rank. The top 1,000 genes account for only 5 % of all genes but 49 % of the total website hits for the Allen Mouse Brain Atlas

mRNA or protein localization experiments from adult, wild-type mouse or rat. Images showing gene expression patterns were carefully compared to corresponding ISH images from the Allen Mouse Brain Atlas. Written descriptions of expression patterns were also utilized, but only as a complement to image comparison. Whole-brain comparisons were opted for whenever possible, but partial-brain comparison was accepted when necessary. The standard for validation was strict for comparisons to peer-reviewed literature sources. In order for a gene to be

validated, all structures available for comparison had to have equivalent expression patterns.

The remaining 900 genes from the top 1,000 initially were compared to two online databases: (1) the St. Jude Brain Gene Expression Map (BGEM) project (Magdaleno et al. 2006), and (2) the Gene Expression Nervous System Atlas (GENSAT) project (Gong et al. 2003). When a database comparison was not possible or did not validate our atlas data, peer-reviewed literature was consulted using the criteria described above. The BGEM database presents dark-field images of mRNA expression across the mouse brain using radioactive riboprobes for in situ hybridization. The GENSAT database contains gene expression data generated from bacterial artificial chromosome (BAC) transgenic mice, where endogenous protein-coding sequences have been replaced with sequences encoding an EGFP reporter gene in each transgenic vector. Because its method is more similar to that of the Allen Mouse Brain Atlas, BGEM was the preferred source for database comparisons. The standard for validation by database comparison was lower than that for literature comparison. For database comparison, gross agreement of expression pattern across major structures was the standard for validation (Fig. 5).

When no comparable database or literature source could be found for a gene in the top 1,000 genes, an attempt was made to validate gene expression using Allen Developing

Table 1 Comparison of top 50 Allen Mouse Brain Atlas genes to literature

Gene rank	Web hits	Gene symbol	Entrez gene ID	Comparison outcome	Reference
1	953,253	<i>Th</i>	21823	Validated	Yamada et al. (1992)
2	721,508	<i>Drd2</i>	13489	Validated	Chen et al. (1991)
3	705,345	<i>Gad1</i>	14415	Validated	Ferraguti et al. (1990)
4	703,364	<i>Slc17a6</i>	140919	Validated	Hisano et al. (2000)
5	593,001	<i>Pvalb</i>	19293	Validated	Seto-Ohshima et al. (1989)
6	586,270	<i>Htr2c</i>	15560	Validated	Wright et al. (1995)
7	584,511	<i>Bdnf</i>	12064	Validated	Hofer et al. (1990)
8	540,323	<i>Crh</i>	12918	Validated	Chen et al. (2000)
9	534,948	<i>Calb1</i>	12307	Validated	Frantz and Tobin (1994)
10	508,386	<i>Drd1a</i>	13488	Validated	Mansour et al. (1992)
11	490,434	<i>Gfap</i>	14580	Validated	Martin and O'Callaghan (1995)
12	445,989	<i>Chat</i>	12647	Validated	Lauterborn et al. (1993)
13	440,561	<i>Npy</i>	109648	Validated	Gehlert et al. (1987)
14	425,745	<i>Cnr1</i>	12801	Validated	Moldrich and Wenger (2000)
15	411,937	<i>Calb2</i>	12308	Validated	Resibois and Rogers (1992)
16	388,558	<i>Slc17a7</i>	72961	Validated	Danik et al. (2005)
17	345,749	<i>Camk2a</i>	12322	Validated	Ochiishi et al. (1994)
18	337,947	<i>Nos1</i>	18125	Validated	Gotti et al. (2005)
19	332,945	<i>Penk</i>	18619	Validated	Zagon et al. (1994)
20	330,884	<i>Trh</i>	22044	Validated	Merchenthaler et al. (1988)
21	313,915	<i>Arc</i>	11838	Validated	de Foubert et al. (2007)
22	311,763	<i>Lepr</i>	16847	Validated	Huang et al. (1996)
23	310,227	<i>Crhr1</i>	12921	Validated	Van Pett et al. (2000)
24	308,520	<i>Cartpt</i>	27220	Validated	Koylu et al. (1998)
25	307,104	<i>Gal</i>	14419	Validated	Cheung et al. (2001)
26	305,646	<i>Foxp2</i>	114142	Validated	Ferland et al. (2003)
27	297,326	<i>Etv1</i>	14009	Validated	Yoneshima et al. (2006)
28	289,949	<i>Nts</i>	67405	Validated	Jennes et al. (1982)
29	288,232	<i>Tph2</i>	216343	Validated	Clark et al. (2006)
30	287,239	<i>Pomc</i>	18976	Validated	Gee et al. (1983)
31	271,036	<i>Sst</i>	20604	Validated	Zavitsanou et al. (2002)
32	267,353	<i>Wfs1</i>	22393	Validated	Takeda et al. (2001)
33	264,385	<i>Avp</i>	11998	Validated	Castel and Morris (1988)
34	262,926	<i>Nes</i>	18008	Validated	Wei et al. (2002)
35	262,852	<i>Pdyn</i>	18610	Validated	Lin et al. (2006)
36	260,241	<i>Slc17a8</i>	216227	Validated	Herzog et al. (2004)
37	256,253	<i>Dlk1</i>	13386	Validated	Jensen et al. (2001)
38	255,497	<i>Cck</i>	12424	Validated	Zavitsanou et al. (2002)
39	254,001	<i>Nr3c1</i>	14815	Validated	Wei et al. (2004)
40	251,371	<i>Adcyap1</i>	11516	Validated	Kurrasch et al. (2007)
41	245,275	<i>Dcx</i>	13193	Validated	Yang et al. (2004)
42	239,046	<i>Prkcd</i>	18753	Validated	Garcia et al. (1993)
43	239,031	<i>Slc6a3</i>	13162	Validated	Shimada et al. (1992)
44	236,846	<i>Gabra6</i>	14399	Validated	Korpi et al. (1993)
45	231,388	<i>Grm5</i>	108071	Validated	Romano et al. (1995)
46	230,769	<i>Pax6</i>	18508	Validated	Stoykova and Gruss (1994)
47	228,745	<i>Grm1</i>	14816	Validated	Shigemoto et al. (1992)
48	227,398	<i>Gria1</i>	14799	Validated	Martin et al. (1993)

Table 1 continued

Gene rank	Web hits	Gene symbol	Entrez gene ID	Comparison outcome	Reference
49	227,379	<i>Tac1</i>	21333	Validated	Warden and Young (1988)
50	222,195	<i>Ntrk2</i>	18212	Validated	Klein et al. (1990)

Shown are the top 50 most accessed genes in the Allen Mouse Brain Atlas, their cumulative web hits, gene symbols, and Entrez gene identifiers. Citations of articles used to validate the data for these genes are included

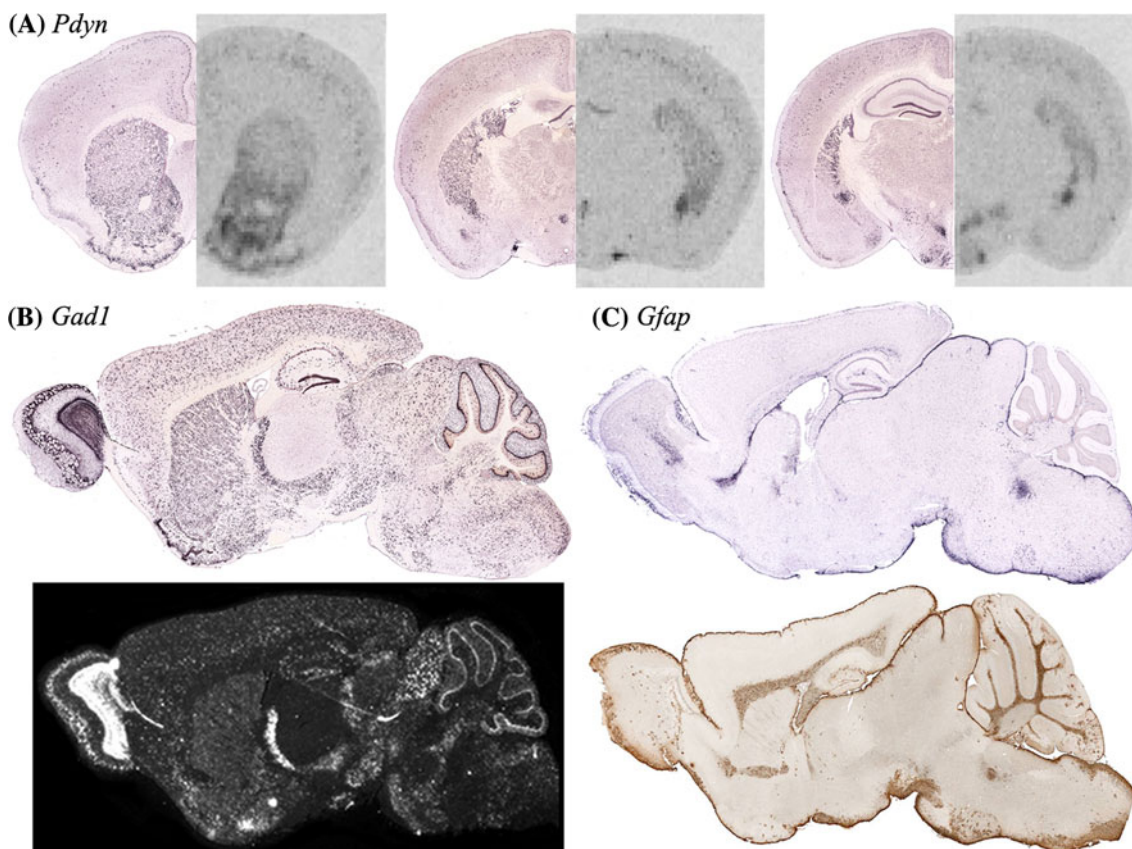


Fig. 5 Validation of Allen Mouse Brain Atlas experiments by external data sources. **a** *Pdyn* data from the Allen Mouse Brain Atlas (left hemisphere) compared with a radiolabeled in situ hybridization section from an adult mouse from Lin et al. (2006) (right hemisphere). **b** *Gad1* data from the Allen Mouse Brain Atlas (top)

compared with a radiolabeled in situ hybridization section from BGEM (bottom). **c** *Gfap* data from the Allen Mouse Brain Atlas (top) compared with a BAC transgenic adult mouse section from GENSAT (bottom). Regions where each gene is strongly expressed are similar between Allen Mouse Brain Atlas data and literature/database sources

Mouse Brain Atlas experiments. Data from the Allen Developing Mouse Brain Atlas using an identical ISH probe to the top 1,000 gene in question were compared to external sources. For example, no comparable adult data could be found for the gene *Lhx6* (data shown in Fig. 2). However, there were three developmental experiments available in the BGEM database (embryonic days 11 and 15 and postnatal day 7). Each of these experiments showed comparable expression to age-corresponding Allen Developing Mouse Brain Atlas experiments. Because the ISH probe could be shown to produce a validated expression

pattern earlier in development, the adult experiments using an identical probe sequence were also considered validated.

The results of the validation project corroborate the overall data accuracy of the top 1,000 genes from the Allen Mouse Brain Atlas (Table 2). We were able to find comparable external data for 782 of the top 1,000 genes. Of these, 96 % (752 genes) were validated by their comparison. The expression patterns of the remaining 4 % (30 genes) were found to contradict comparable external data. For these mismatches, a redo experiment was initiated in an attempt to produce a supported expression pattern.

Table 2 Overall results the of top 1,000 validation project

Top 100 genes		
Validated	100	
Compared to literature		100
Mismatch	0	
Unvalidated	0	
Next 900 genes		
Validated	661	
Compared to literature		261
Compared to BGEM1		359
Compared to GENSAT		94
Probe validation		47
Mismatch	21	
Unvalidated	218	

Subsequent experiments have produced validated data for 9 of these 30 genes. For the 21 remaining genes, further experiments are ongoing to resolve mismatches. The possibility remains that some external data used for comparison was either itself inaccurate or was not, in fact, biologically comparable.

Conclusions

The mapping resources described here allow researchers to compare gene expression across the entire brain and spinal cord in high spatial resolution. Addition of the temporal axis provides important new information about brain development and the localization of genes involved in neurogenesis, transcription, synaptic activity, and neurotransmission. The data for a single gene are easily obtained for all projects and can provide a revealing gestalt of that gene's activity over time and region. This is illustrated by the gene Somatostatin, which exhibits highly regionalized expression throughout development and in adulthood in both the brain and the spinal cord (Fig. 6).

Current efforts at the Allen Institute for Brain Science include nonhuman primate and human brain mapping studies that will allow seamless cross-species comparisons. The addition of human microarray and RNA sequencing data provides other genomic modalities to search across species and time. Future projects include detailed mapping of the connectome of the mouse brain and a complete cataloging of neuronal cell types in the cerebral cortex. Combined, these online public resources, which are made freely available to researchers around the world, will help to develop the essential parts list of molecular neuroanatomical tools needed to advance our understanding of how the human brain functions.

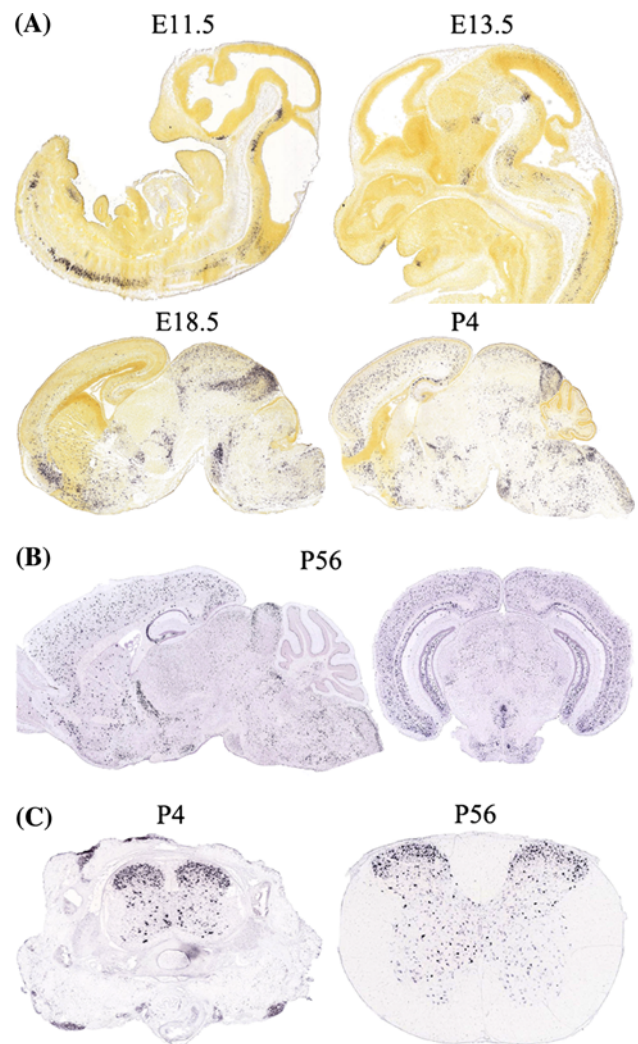


Fig. 6 Various ISH sections for a single mouse gene, Somatostatin (*Sst*). **a** Sagittal Allen Developing Mouse Brain Atlas sections showing the whole embryo at E11.5, the head at E13.5, and the brain at E18.5 and P4. *Sst* expression shifts from a few small regions at E11.5 to most major brain regions by P4. **b** Sagittal (*left*) and coronal (*right*) Allen Mouse Brain Atlas sections showing strong scattered expression in the adult cortex and caudoputamen. Enriched expression can be seen in the reticular thalamus and the superior colliculus. Many other sites also show regionalized *Sst* expression. **c** Cervical Allen Spinal Cord Atlas sections of neonate spinal column (*left*) and adult spinal cord (*right*) showing scattered expression throughout the spinal gray matter, and enrichment in laminae 1–3 compared to deeper layers. The enrichment in laminae 1–3 can be seen in the sagittal P4 and sagittal P56 brain sections near the base of the medulla. Undefined spinal cord expression can be seen in the sagittal E11.5 and E13.5 sections

Acknowledgments The authors thank the Allen Institute founders, Paul G. Allen and Jody Allen, for their vision, encouragement, and support. The Allen Spinal Cord Atlas was made possible by a unique funding consortium comprising public and private entities, including The ALS Association, PVA Research Foundation (grant No. 2525), Wyeth Research, National Multiple Sclerosis Society, International Spinal Research Trust, PEMCO Insurance, and philanthropist and

Allen Institute founder Paul G. Allen, as well as other anonymous donors.

References

- Allen institute for brain science (2010) Comparison of the top 1000 genes (October 2010), Allen mouse brain atlas. Available at <http://help.brain-map.org/display/mousebrain/Documentation>. Last accessed March 2012
- Belgard TG, Marques AC, Oliver PL, Abaan HO, Sirey TM et al (2011) A transcriptomic atlas of mouse neocortical layers. *Neuron* 71:605–616
- Bernard A, Lubbers LS, Tanis KQ, Luo R, Podtelezchnikov AA et al (2012) Transcriptional architecture of the primate neocortex. *Neuron* 73:1083–1099
- Castel M, Morris JF (1988) The neurophysin-containing innervation of the forebrain of the mouse. *Neuroscience* 24:937–966
- Chen JF, Qin ZH, Szele F, Bai G, Weiss B (1991) Neuronal localization and modulation of the D2 dopamine receptor mRNA in brain of normal mice and mice lesioned with 6-hydroxydopamine. *Neuropharmacology* 30:927–941
- Chen Y, Brunson KL, Muller MB, Cariaga W, Baram TZ (2000) Immunocytochemical distribution of corticotropin-releasing hormone receptor type-1 (CRF(1))-like immunoreactivity in the mouse brain: light microscopy analysis using an antibody directed against the C-terminus. *J Comp Neurol* 420:305–323
- Cheung CC, Hohmann JG, Clifton DK, Steiner RA (2001) Distribution of galanin messenger RNA-expressing cells in murine brain and their regulation by leptin in regions of the hypothalamus. *Neuroscience* 103:423–432
- Clark MS, McDevitt RA, Neumaier JF (2006) Quantitative mapping of tryptophan hydroxylase-2, 5-HT_{1A}, 5-HT_{1B}, and serotonin transporter expression across the anteroposterior axis of the rat dorsal and median raphe nuclei. *J Comp Neurol* 498:611–623
- Danik M, Cassoly E, Manseau F, Sotty F, Mouginot D et al (2005) Frequent coexpression of the vesicular glutamate transporter 1 and 2 genes, as well as coexpression with genes for choline acetyltransferase or glutamic acid decarboxylase in neurons of rat brain. *J Neurosci Res* 81:506–521
- de Foubert G, O'Neill MJ, Zetterstrom TS (2007) Acute onset by 5-HT₆-receptor activation on rat brain-derived neurotrophic factor and activity-regulated cytoskeletal-associated protein mRNA expression. *Neuroscience* 147:778–785
- Dong HW (2008) The Allen reference atlas: a digital color brain atlas of the C57BL/6J male mouse. Wiley, Hoboken
- Ferland RJ, Cherry TJ, Preware PO, Morrisey EE, Walsh CA (2003) Characterization of *Foxp2* and *Foxp1* mRNA and protein in the developing and mature brain. *J Comp Neurol* 460:266–279
- Ferraguti F, Zoli M, Aronsson M, Agnati LF, Goldstein M et al (1990) Distribution of glutamic acid decarboxylase messenger RNA-containing nerve cell populations of the male rat brain. *J Chem Neuroanat* 3:377–396
- Frantz GD, Tobin AJ (1994) Cellular distribution of calbindin D28 K mRNAs in the adult mouse brain. *J Neurosci Res* 37:287–302
- French L, Pavlidis P (2011) Relationships between gene expression and brain wiring in the adult rodent brain. *PLoS Comput Biol* 7:e1001049
- Garcia MM, Cusick CG, Harlan RE (1993) Protein kinase C-delta in rat brain: association with sensory neuronal hierarchies. *J Comp Neurol* 331:375–388
- Gee CE, Chen CL, Roberts JL, Thompson R, Watson SJ (1983) Identification of proopiomelanocortin neurones in rat hypothalamus by in situ cDNA-mRNA hybridization. *Nature* 306:374–376
- Gehlert DR, Chronwall BM, Schafer MP, O'Donohue TL (1987) Localization of neuropeptide Y messenger ribonucleic acid in rat and mouse brain by in situ hybridization. *Synapse* 1:25–31
- Gong S, Zheng C, Doughty ML, Losos K, Didkovsky N et al (2003) A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* 425:917–925
- Gotti S, Sica M, Viglietti-Panzica C, Panzica G (2005) Distribution of nitric oxide synthase immunoreactivity in the mouse brain. *Microsc Res Tech* 68:13–35
- Herzog E, Gilchrist J, Gras C, Muzerelle A, Ravassard P et al (2004) Localization of VGLUT3, the vesicular glutamate transporter type 3, in the rat brain. *Neuroscience* 123:983–1002
- Hisano S, Hoshi K, Ikeda Y, Maruyama D, Kanemoto M et al (2000) Regional expression of a gene encoding a neuron-specific Na(+)-dependent inorganic phosphate cotransporter (DNPI) in the rat forebrain. *Brain Res Mol Brain Res* 83:34–43
- Hofer M, Pagliusi SR, Hohn A, Leibrock J, Barde YA (1990) Regional distribution of brain-derived neurotrophic factor mRNA in the adult mouse brain. *EMBO J* 9:2459–2464
- Huang XF, Koutcherov I, Lin S, Wang HQ, Storlien L (1996) Localization of leptin receptor mRNA expression in mouse brain. *Neuroreport* 7:2635–2638
- Jennes L, Stumpf WE, Kalivas PW (1982) Neurotensin: topographical distribution in rat brain by immunohistochemistry. *J Comp Neurol* 210:211–224
- Jensen CH, Meyer M, Schroder HD, Kliem A, Zimmer J et al (2001) Neurons in the monoaminergic nuclei of the rat and human central nervous system express *FA1/dlk*. *Neuroreport* 12:3959–3963
- Jones AR, Overly CC, Sunkin SM (2009) The Allen brain atlas: 5 years and beyond. *Nat Rev Neurosci* 10:821–828
- Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X et al (2011) Spatio-temporal transcriptome of the human brain. *Nature* 478:483–489
- Klein R, Martin-Zanca D, Barbacid M, Parada LF (1990) Expression of the tyrosine kinase receptor gene *trkB* is confined to the murine embryonic and adult nervous system. *Development* 109:845–850
- Korpi ER, Kleingoor C, Kettenmann H, Seeburg PH (1993) Benzodiazepine-induced motor impairment linked to point mutation in cerebellar GABAA receptor. *Nature* 361:356–359
- Koylu EO, Couceyro PR, Lambert PD, Kuhar MJ (1998) Cocaine- and amphetamine-regulated transcript peptide immunohistochemical localization in the rat brain. *J Comp Neurol* 391:115–132
- Kurrasch DM, Cheung CC, Lee FY, Tran PV, Hata K et al (2007) The neonatal ventromedial hypothalamus transcriptome reveals novel markers with spatially distinct patterning. *J Neurosci* 27:13624–13634
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC et al (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921
- Lau C, Ng L, Thompson C, Pathak S, Kuan L et al (2008) Exploration and visualization of gene expression with neuroanatomy in the adult mouse brain. *BMC Bioinformatics* 9:153
- Lauterborn JC, Isackson PJ, Montalvo R, Gall CM (1993) In situ hybridization localization of choline acetyltransferase mRNA in adult rat brain and spinal cord. *Brain Res Mol Brain Res* 17:59–69
- Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A et al (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445:168–176
- Lin S, Boey D, Lee N, Schwarzer C, Sainsbury A et al (2006) Distribution of prodynorphin mRNA and its interaction with the NPY system in the mouse brain. *Neuropeptides* 40:115–123
- Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ et al (2011) A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478:476–482

- Liodis P, Denaxa M, Grigoriou M, Akufo-Addo C, Yanagawa Y et al (2007) Lhx6 activity is required for the normal migration and specification of cortical interneuron subtypes. *J Neurosci* 27:3078–3089
- Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA et al (2010) A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci* 13:133–140
- Magdalenos S, Jensen P, Brumwell CL, Seal A, Lehman K et al (2006) BGEM: an in situ hybridization database of gene expression in the embryonic and adult mouse nervous system. *PLoS Biol* 4:e86
- Mansour A, Meador-Woodruff JH, Zhou Q, Civelli O, Akil H et al (1992) A comparison of D1 receptor binding and mRNA in rat brain using receptor autoradiographic and in situ hybridization techniques. *Neuroscience* 46:959–971
- Martin PM, O'Callaghan JP (1995) A direct comparison of GFAP immunocytochemistry and GFAP concentration in various regions of ethanol-fixed Rat and mouse brain. *J Neurosci Methods* 58:181–192
- Martin LJ, Blackstone CD, Levey AI, Haganir RL, Price DL (1993) AMPA glutamate receptor subunits are differentially distributed in rat brain. *Neuroscience* 53:327–358
- Merchenthaler I, Csernus V, Csontos C, Petrusz P, Mess B (1988) New data on the immunocytochemical localization of thyrotropin-releasing hormone in the rat central nervous system. *Am J Anat* 181:359–376
- Moldrich G, Wenger T (2000) Localization of the CB1 cannabinoid receptor in the rat brain. An immunohistochemical study. *Peptides* 21:1735–1742
- Ng L, Bernard A, Lau C, Overly CC, Dong HW et al (2009) An anatomic gene expression atlas of the adult mouse brain. *Nat Neurosci* 12:356–362
- Ochiishi T, Terashima T, Yamauchi T (1994) Specific distribution of Ca²⁺/calmodulin-dependent protein kinase II alpha and beta isoforms in some structures of the rat forebrain. *Brain Res* 659:179–193
- Petit A, Sanders AD, Kennedy TE, Tetzlaff W, Glattfelder KJ et al (2011) Adult spinal cord radial glia display a unique progenitor phenotype. *PLoS One* 6:e24538
- Resibois A, Rogers JH (1992) Calretinin in rat brain: an immunohistochemical study. *Neuroscience* 46:101–134
- Romano C, Sesma MA, McDonald CT, O'Malley K, Van den Pol AN et al (1995) Distribution of metabotropic glutamate receptor mGluR5 immunoreactivity in rat brain. *J Comp Neurol* 355:455–469
- Rubenstein JL, Martinez S, Shimamura K, Puelles L (1994) The embryonic vertebrate forebrain: the prosomeric model. *Science* 266:578–580
- Sengul G, Puchalski RB, Watson C (2012) Cytoarchitecture of the spinal cord of the postnatal (P4) mouse. *Anat Rec (Hoboken)* 295(5):837–845
- Seto-Ohshima A, Emson PC, Berchtold MW, Heizmann CW (1989) Localization of parvalbumin mRNA in rat brain by in situ hybridization histochemistry. *Exp Brain Res* 75:653–658
- Shigemoto R, Nakanishi S, Mizuno N (1992) Distribution of the mRNA for a metabotropic glutamate receptor (mGluR1) in the central nervous system: an in situ hybridization study in adult and developing rat. *J Comp Neurol* 322:121–135
- Shimada S, Kitayama S, Walther D, Uhl G (1992) Dopamine transporter mRNA: dense expression in ventral midbrain neurons. *Brain Res Mol Brain Res* 13:359–362
- Shimogori T, Lee DA, Miranda-Angulo A, Yang Y, Wang H et al (2010) A genomic atlas of mouse hypothalamic development. *Nat Neurosci* 13:767–775
- Stoykova A, Gruss P (1994) Roles of Pax-genes in developing and adult brain as suggested by expression patterns. *J Neurosci* 14:1395–1412
- Takeda K, Inoue H, Tanizawa Y, Matsuzaki Y, Oba J et al (2001) WFS1 (Wolfram syndrome 1) gene product: predominant subcellular localization to endoplasmic reticulum in cultured cells and neuronal expression in rat brain. *Hum Mol Genet* 10:477–484
- Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TE et al (2012) The human connectome project: a data acquisition perspective. *NeuroImage*. doi:10.1016/j.neuroimage.2012.02.018
- Van Pett K, Viau V, Bittencourt JC, Chan RK, Li HY et al (2000) Distribution of mRNAs encoding CRF receptors in brain and pituitary of rat and mouse. *J Comp Neurol* 428:191–212
- Warden MK, Young WS (1988) Distribution of cells containing mRNAs encoding substance P and neurokinin B in the rat central nervous system. *J Comp Neurol* 272:90–113
- Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF et al (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520–562
- Watson C, Paxinos G, Kayalioglu G (2008) The spinal cord, 1st edn. Academic Press, Waltham
- Wei LC, Shi M, Chen LW, Cao R, Zhang P et al (2002) Nestin-containing cells express glial fibrillary acidic protein in the proliferative regions of central nervous system of postnatal developing and adult mice. *Brain Res Dev Brain Res* 139:9–17
- Wei Q, Lu XY, Liu L, Schafer G, Shieh KR et al (2004) Glucocorticoid receptor overexpression in forebrain: a mouse model of increased emotional lability. *Proc Natl Acad Sci USA* 101:11851–11856
- Wright DE, Seroogy KB, Lundgren KH, Davis BM, Jennes L (1995) Comparative localization of serotonin1A, 1C, and 2 receptor subtype mRNAs in rat brain. *J Comp Neurol* 351:357–373
- Yamada K, Sakai M, Okamura H, Ibata Y, Nagatsu I (1992) Detection of tyrosine hydroxylase and phenylethanolamine-N-methyltransferase messenger RNAs in the mouse adrenal gland and the brain by in situ hybridization. *Histochemistry* 97:201–206
- Yang HK, Sundholm-Peters NL, Goings GE, Walker AS, Hyland K et al (2004) Distribution of doublecortin expressing cells near the lateral ventricles in the adult mouse brain. *J Neurosci Res* 76:282–295
- Yoneshima H, Yamasaki S, Voelker CC, Molnar Z, Christophe E et al (2006) Er81 is expressed in a subpopulation of layer 5 neurons in rodent and primate neocortices. *Neuroscience* 137:401–412
- Zagon IS, Isayama T, McLaughlin PJ (1994) Preproenkephalin mRNA expression in the developing and adult rat brain. *Brain Res Mol Brain Res* 21:85–98
- Zavitsanou K, Triarhou LC, Kouvelas ED, Mitsacos A, Palacios JM et al (2002) Somatostatin, cholecystokinin and neuropeptide Y mRNAs in normal and weaver mouse brain. *J Neural Transm* 109:1337–1351
- Zeng H, Shen EH, Hohmann JG, Oh SW, Bernard A et al (2012) Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. *Cell* 149(2):483–496