ARTICLE
The Allen Brain Atlas as a Resource for Teaching Undergraduate Neuroscience

Terri L. Gilbert
GilHou Science Communication, Seattle, WA, 98106.

The open science movement has resulted in a growing field of data- and tool-sharing platforms that serve as a resource not only for sharing data and results in the field of brain science but has allowed students and researchers to learn neuroscientific skills and concepts. For over a decade, the Allen Institute for Brain Science has been meticulously collecting high quality data mapping gene expression, connectivity and, more recently, functional data from the brains of mice, macaques and humans. These open data have been paired with unique navigation and visualization tools such that the neuroscience researcher can explore, utilize and even incorporate these data into their publications. The tools created to explore and analyze the Allen Brain Atlas datasets have also been widely utilized to teach neuroscientific concepts to undergraduate and graduate students. This article aims to outline how to use the Allen Brain Atlas resources as teaching tools to impart neuroanatomic concepts to undergraduate and graduate neuroscience students.

Key words: Allen Brain Atlas, Gene Expression, In Situ Hybridization (ISH), Allen Mouse Brain Atlas.

The desire to understand how the brain works and how this fatty organ gives us our experience of ourselves and the world around us, makes neuroscience a fascinating and attractive field of study. Given that the brain is arguably the most complex structure in the known universe, studying the brain is also a non-trivial undertaking. The Allen Brain Atlas (www.brain-map.org) is an online freely available resource that includes high quality brain data merged with years of expertise in the field of neuroanatomy. This kind of resource lowers the barrier to learning and understanding brain structure and function and has been an educational resource to neuroanatomy bioinformatics professors for almost as long as it has been publishing data (Ramos et al., 2007; Jenks, 2009; Grisham et al., 2010; Chu et al., 2015; Grisham et al., 2017). Recent published data and tools (Allen Cell-type Database and the Allen Brain Observatory) are ideal resources to impart the basics of working with big data and include practice python notebooks to facilitate the analysis of these data (Gilbert and Ng, 2018). The current article describes how the unique informatics tools overlaid on the precise and systematically collected brain data available in the Allen Brain Atlas can be used as a tool not only for teaching neuroscience, but as a resource for hypothesis generation and to fuel discovery. Ultimately having the ability to use the available open science and data resources will forward people's ability to understand the brain especially when research funding is thin.

After a review of the resources available through the Allen Brain Atlas, a basic framework for using the online atlases as a teaching tool will be outlined. This framework will be demonstrated through a sample lesson plan using the Allen Mouse Brain Atlas. The intention of this article is for the reader to be able to create and adapt lesson plans based on the curriculum being taught paired with the resources available. This framework is designed to be effective with both the resources available from the Allen Institute website and as well as with other online open resources.

THE ALLEN BRAIN ATLAS

The Allen Institute for Brain Science began as a project to map gene expression in the mouse brain. Within three years, scientists had developed a systematic high-throughput method to complete a genome-wide survey of gene expression in the adult male mouse brain using colorimetric in situ hybridization (Lein et al., 2007). In order to make this vast amount of image data accessible, informatics and visualization tools were developed that, when integrated with the images, gave the online user the ability to navigate and explore the data without a prior understanding of function or structure (Ng et al., 2007). Subsequent projects mapped gene expression in the developing mouse brain (Thompson et al., 2014), the mouse spinal cord (Sengul et al., 2012), the non-human primate (Bernard et al., 2012), and in adult (Hawrylycz et al., 2012) and developing humans (Miller et al., 2014). Each of these resources laid the foundation to start asking more complex questions about the brain, such as "How are the regions of the brain connected?" and "How does the living brain function?". Current projects at the Allen Institute for Brain Science include resources addressing the questions of connectivity through the Allen Mouse Brain Connectivity Atlas (Oh et al., 2014), and function through the Allen Cell Types Database and the Allen Brain Observatory. What follows is a brief description of each of the Allen Institute resources. Familiarizing yourself with the kind of data available is the first step in utilizing resources like these in your lessons.

Allen Mouse Brain Atlas (mouse.brain-map.org). A genome-wide, three-dimensional map of gene expression throughout the adult mouse brain. Similar in scale to the Human Genome Project, the atlas comprises cellular resolution *in situ* hybridization (ISH) images with comprehensive anatomic coverage that reveal where each gene is expressed in the brain, as well as an integrated suite of powerful data search and visualization tools, including Correlative and Differential Searches, the

Anatomic Gene Expression Atlas (AGEA) and an anatomic reference atlas.

Allen Spinal Cord Atlas (mousespinal.brain-map.org). A genome-wide comprehensive resource detailing gene activity in the normal spinal cord. This atlas provides an essential baseline for understanding how gene expression in the spinal cord is altered in disease or injury. The Atlas includes image-based *in situ* hybridization data at cellular resolution from both juvenile (P4) and adult (P56) stages with anatomic reference atlases across the full length of the spinal cord.

Allen Developing Mouse Brain Atlas (developingmouse.brain-map.org). A detailed map of changes in gene expression across development of the mouse brain. This atlas provides a framework to explore both when and where genes are activated in the mouse brain from embryo through old age. Informatics data processing enables both spatial search and temporal search. Anatomic and temporal search locates enhanced gene expression in large structures, while manual data annotation allows for search of enhanced gene expression in small structures. Developmental reference atlases provide an additional framework for the data.

NIH Blueprint Non-Human Primate Atlas (blueprintnhpatlas.org). The National Institutes of Health Blueprint for Neuroscience Research awarded a contract to the Allen Institute for Brain Science, in partnership with researchers at the University of California at Davis, to generate an atlas of gene expression in the developing rhesus macaque brain. This atlas creates a developmental neuroanatomical framework for exploring the cellular and molecular architecture of the developing primate brain with direct relevance for human brain development.

Allen Human Brain Atlas (human.brain-map.org). A unique multi-modal atlas that maps gene expression across the adult human brain. This atlas integrates anatomic and genomic information, data modalities include magnetic resonance imaging (MRI), diffusion tensor imaging (DTI), histology, and gene expression data derived from both microarray and in situ hybridization (ISH) approaches. Key features include an "all genes, all structures" microarray survey spatially mapped to the MRI, ISH cellular resolution image data for selected genes in specific brain regions, and an annotated human brain atlas guide.

BrainSpan Atlas of the Developing Human Brain (brainspan.org). A unique resource for studying human brain development. The atlas provides a broad and detailed anatomical analysis of gene expression across human brain development, comprising *in situ* hybridization, RNA-sequencing and microarray approaches, along with supporting neuroanatomical reference content. The atlas was developed by a consortium of scientific partners from multiple organizations and was funded by awards from the National Institute of Mental Health.

Ivy Glioblastoma Atlas Project (glioblastoma.alleninstitute.org). A platform for exploring the anatomic and genetic basis of glioblastoma at the cellular and molecular levels. Data available via the Allen Brain Atlas data portal include cellular resolution in situ hybridization data mapping gene expression across the anatomic structures inherent in glioblastoma, genome-wide RNA-sequence profiling for anatomical structures identified in the ISH survey, as well as associated histological data suitable for neuropathological examination. A companion database, available at ivygap.org, linked by de-identified tumor specimen numbers and developed by project partners at the Ben and Catherine Ivy Center for Advanced Brain Tumor Treatment, provides additional clinical and genomic data. This project was made possible through the support of the Ben and Catherine Ivy Foundation.

Aging, Dementia and TBI Study (aging.brain-map.org). This resource is a detailed neuropathologic, molecular and transcriptomic characterization of brains with TBI exposure and control cases from a unique aged population-based cohort from the Adult Changes in Thought (ACT) study. Data available via the Allen Brain Atlas data portal include RNA sequencing data, high-resolution ISH image data, Luminex protein quantification, immunohistochemistry and histology, isoprostane quantification and specimen metadata (Miller et al., 2017). This project was developed by a consortium consisting of the University of Washington, Kaiser Permenente Research Institute and the Allen Institute for Brain Science and was made possible through the support of the Paul G. Allen Family Foundation.

Allen Mouse Brain Connectivity Atlas (connectivity.brain-map.org). A high-resolution map of long range neuronal projections in the mouse brain. Built on a suite of transgenic mice genetically engineered to visually target specific cell classes, the atlas comprises a unique compendium of projections from selected neuronal populations throughout the brain. Key features include tract tracing image data captured using serial two-photon tomography, transgenic characterization data detailing expression in Cre and other driver lines, and anatomic reference data. The data is presented within an interactive 3D viewer (Brain Explorer®) and can be searched by anatomical region, injection site or axonal trajectories, as well as with a "virtual" retrograde search.

Allen Cell Types Database (celltypes.brain-map.org). The initial step in characterizing the cellular components of the brain. The database is a multi-modal characterization of cells from the adult mouse visual cortex and surgical cortical samples from human donors based on their functional and structural characteristics. Users can investigate the brain's cell types using electrophysiological and morphological data from single cells and accompanying models of cellular behavior. Key features include whole-cell patch clamp recordings from targeted cells, images of biocytin-filled neurons, full 3D digital reconstructions of a subset of cells, and GLIF and biophysical models that can be downloaded via the Allen SDK (http://alleninstitute.github.io/AllenSDK/). As of 2017 single cell RNA sequencing had been recorded with the intention to build a data-driven cellular taxonomy in the future.

Allen Brain Observatory (observatory.brain-map.org). As of 2017, this inaugural dataset included a first step in characterizing the functional activity of individual neurons in the awake behaving mouse. This resource is a multi-

modal characterization of the adult mouse visual cortex, using widefield and *in vivo* calcium imaging of visually evoked responses from neurons sampled from selected brain areas, cortical layers, and Cre lines. This resource enables a quantitative exploration of the functional properties that underlie coding of sensory stimuli through the visual pathway, at both the single-cell and population level. Experimental data derived from standardized image acquisition have been generated using a robust, scalable laboratory methodology, and are available as raw files or summary graphical representations.

OUTLINE OF A LESSON PLAN

Understanding the data and tools available from the Allen Brain Atlas or other online resources is the first step in utilizing these resources in your lesson plans. These tools are ideal to study neuroanatomy at the cellular and structural level in mice, the non-human primate (rhesus macaque) and humans. The gene expression patterns can also be defined over development across these three species. These tools are also well suited to understand structural connectivity in the mouse brain. The latest projects underway at the Allen Institute for Brain Science lend themselves quite well to teaching neuroinformatics and using real data to learn how to work with big data.

Within these boundaries, the next step is to pose a question. Questions range from the very simple and straight-forward, “Where is dopa decarboxylase (or other gene of interest) expressed in the adult mouse brain?” to very complex, “Are axon guidance molecules expressed during early chick P1 development conserved in mammals?”

Once a question has been posed, the students need to be made aware of the data and tools available that will allow for inquiring into the question. Often there are limitations in the data that, if understood, will circumvent questions that may come up later, such as “Why, in the Allen Mouse Brain Atlas, is there only sagittal data for one hemisphere?” (Answer: Full brain or coronal sections were only performed on ~4000 genes that were chosen based on a genome-wide survey on one hemisphere in the sagittal plane). This will require making yourself aware of the data available in the resource by reading the methods papers in the Documentation section of the individual resource or by taking advantage of online videos that outline the data (https://www.terrigilbertphd.com/teaching-tools/).

The next step will be to train the students on the search and navigation features available in the resource you will be using. This is accomplished by reading the online help for each resource, viewing any tutorials that may be available (http://www.brain-map.org/tutorials/index.html) and also immersing yourself in the resource by asking yourself questions and seeing where your inquiry takes you.

The last step is to ask questions you have already sorted out for yourself first so that you build confidence in using the tools and provide a framework for your students to explore the data and start to discover the inner workings of the brain.

To summarize:

1. Familiarize yourself with the resource.
2. Pose a question.
3. Train on the basic and advanced search and navigation features available with that resource.
4. Practice on the resource so you have the skills to troubleshoot issues that arise from conducting research with real data.

EXAMPLE LESSON PLAN

This section aims to provide an example of how this plan can be used in the classroom. This example starts with posing a question and providing the steps necessary to prepare yourself to be able to impart this knowledge to your students.

Question: Where is dopa decarboxylase expressed and what does that tell you about its function?

To address this question in this example we will limit ourselves to the Allen Mouse Brain Atlas but the same framework should be used to address this question in the human resources.

ADULT MOUSE BRAIN ATLAS DATA

Familiarize yourself and your students with the data in this resource (see below). More detailed information regarding the methods used to create this resource, both the data itself as well as the web interface can be found in the Documentation tab located at mouse.brain-map.org.

The data for the Allen Mouse Brain Atlas was collected from C57BL/6J male mice (Jackson Laboratories, Farmington, CT) at post-natal day 56 from fresh frozen tissue (Fig. 1). An initial survey of ~20,000 genes was performed in the sagittal plane for approximately half the brain starting just prior to the midline. For ~4,000 genes, genes were probed in the coronal plane which ensured bilateral expression across the majority of the brain.

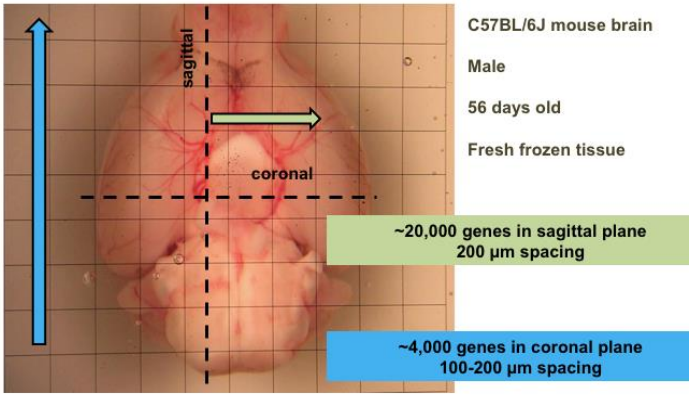

Figure 1. Adult Mouse Brain Atlas Data. A genome-wide survey of fresh frozen male C57BL/6J at P56 were sectioned at 200 µm spacing in the sagittal plane and a smaller number of genes were probed in the coronal plane.

Eight experimental assays were conducted on each mouse brain and were designed to capture the full extent of the brain. Each section was 25 µm thick and apportioning the sections across eight assays ensured 200

μm spacing between each section in a gene probe assay. Typically, six gene probe assays were performed per brain and paired with two Nissl histology assays for reference. Diagrams of the sectioning paradigm for the sagittal and coronal experiments are shown in Figures 2 and 3, respectively. This sectioning paradigm resulted in approximately 20 images per gene probe experiment in the sagittal plane and 50-60 images in the coronal plane.

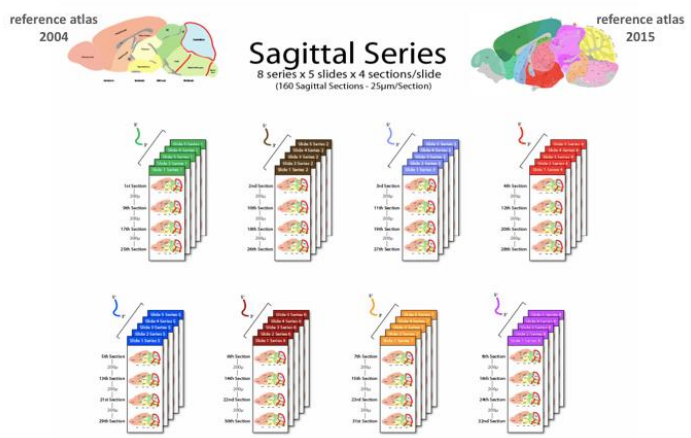

Figure 2. Sectioning paradigm for the sagittal series experiments in the Adult Mouse Brain Atlas. Each brain was sectioned into eight sagittal series of five slides that included four sections per slide. Sections were placed such that there was a spacing of 200 μm between sections. A single brain was sectioned so that six gene experiments and two Nissl histology experiments were conducted per brain. Each sagittal experiment includes approximately 20 images.

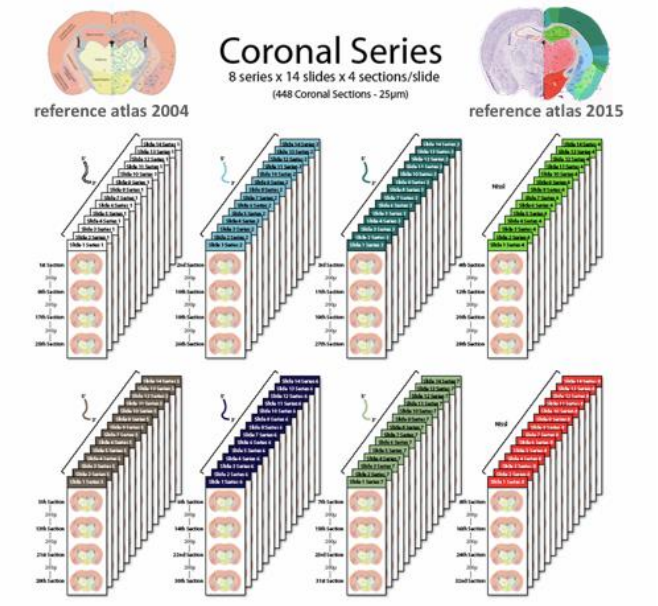

Figure 3. Sectioning paradigm for the coronal series experiments in the Adult Mouse Brain Atlas. Each brain was sectioned into eight sagittal series of fourteen slides that included four sections per slide. Sections were placed such that there was a spacing of 200 μm between sections. A single brain was sectioned so that six gene experiments and two Nissl histology experiments were conducted per brain. Each coronal experiment contains between 50-60 images.

Gene expression in the mouse brain atlas was assayed using colorimetric *in situ* hybridization which is effectively an antibody-based assay that qualitatively measures the presence of mRNA in a cell. Optimization of this platform was fully vetted and compared to equivalent autoradiography experiments to confirm sensitivity (see methods from the Documentation tab at mouse.brain-map.org for details). Because the assay is qualitative, gene levels can be approximated within an experiment, but comparing gene levels between experiments (even experiments using identical probes) is not the ideal manner in which to utilize this resource.

The Allen Institute developed a high-throughput method that enabled them to measure the expression of all the known genes within three years. First, slides were assembled into cassettes that formed a reservoir to enable reagent delivery to each section in an automated fashion (Fig. 4a). Slides were loaded into Tecan robots (Fig. 4b) which delivered reagents and altered temperatures in an automated fashion in accordance with the hybridization protocol. Probe dilution and aliquoting, cover-slipping post ISH (Fig. 4c), imaging and image processing were also entirely automated processes.

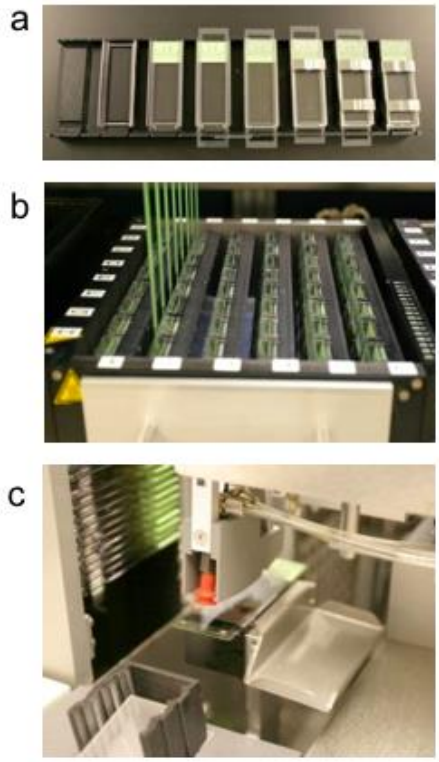

Figure 4. Automated high through-put *in situ* hybridization process. *a.* Each slide was assembled into a cassette which effectively created a reservoir in which the hybridization reagents could be washed over the sections using gravity. *b.* Tecan robots controlled temperature and reagent washes over the 23.5 hr process. *c.* Cover-slipping was performed automatedly to ensure reproducibility in this imaging-sensitive step.

The product of this considerable experimental undertaking is a vast number of two dimensional images that highlight positive expression of genes within discrete structures and cell populations (Fig. 5).

The meticulous nature of the data collection in this resource is only part of what makes the Allen Mouse Brain Atlas such a unique and useful resource. Informatics modules were developed to systematically process the images so that the data could be integrated both across experiments and across all the data so as to have the data start to inform the user about the structure and function of the brain (Fig. 6). Each experiment was registered into a 3-dimensional brain reference space referred to as the Common Coordinate Framework (CCF, see white paper in resource Documentation page) so as to overlay gene expression with the structures of the brain.

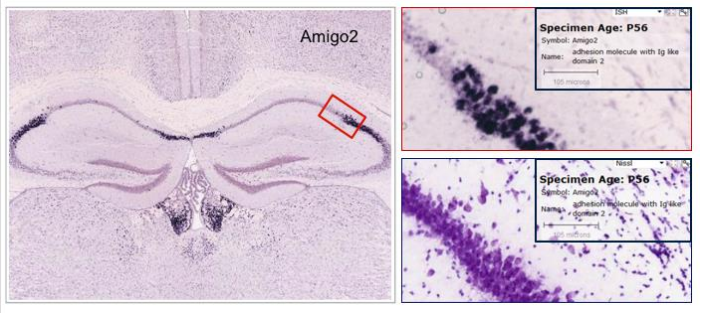

Figure 5. In situ hybridization image which shows the expression of Amigo2 in specific regions of the brain such as the CA2 region of the hippocampus (red box). The two images in the right panel show a magnified version of the image on the left, and a Nissl stain of the same region to illustrate that specific cells show the expression pattern.

Background signal, while not the best for human eyes inspecting the data, is ideal for algorithmic delineation of tissue edges (Preprocess, Fig. 6) as well as determining the context in which signal is detected. An adaptive thresholding algorithm determined from batch control samples subtracts background signal and colorizes expression patterns based on an intensity color map in which blue and green hues indicate low expression and the hotter reds and yellows, high expression (Expression Mask, Fig. 6).

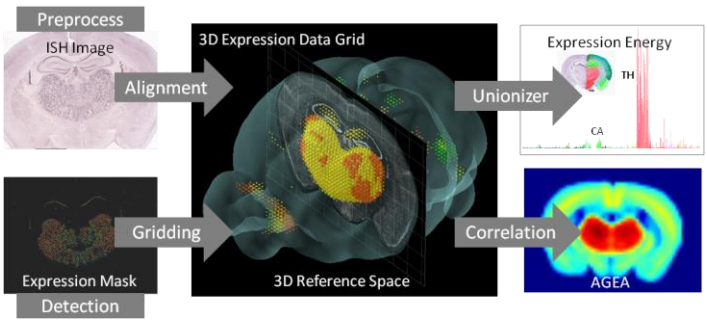

Figure 6. Mouse Brain Atlas Informatics. Modules enable registration of *in situ* hybridization image data into 3-dimensional space to allow comparison of gene expression patterns with brain structure.

Registration of the data into the CCF (Alignment and Gridding from Fig. 6) enables the quantification of gene expression across the brain within a single experiment (Unionizer in Fig. 6) as well as the generation of correlation maps to compare gene expression across structures (Correlation in Fig. 6).

Once the data and informatics modules overlaid on the data are understood, learning how to use the basic search and navigation is the next step.

BASIC SEARCH AND NAVIGATION

The data collected at the Allen Institute for Brain Science was meticulously collected so as to enable both understanding of the biology behind the experiments as well as to allow the data to be served up in a digital, user-friendly format. Over the years, feedback from users of the platform also informed upgrades and refinement to the tools overlaying the data so as to enable the user to better understand and interpret the data. As the resources and data-types expanded, attempts were made to standardize the navigation tools so as to allow for training on one of the projects to translate in a basic understanding of the search and visualization features of other projects. For the purposes of this article, the basic navigation features that need to be learned and then imparted to the student are the tools designed to view images (for the ISH and Connectivity data) and the tools that display genotyping data in heatmap form (microarray and RNA sequencing data).

The most ideal way to learn to use these tools is through hands-on practice, and as that is difficult to impart in print there were online tools created specifically to facilitate learning these skills. A companion webpage (https://www.terrigilbertphd.com/teaching-tools/) for this article was set-up to provide videos, PowerPoint files and links to enable this process.

This exploration began with a simple question: Where is dopa decarboxylase expressed and what does that tell you about its function? Once you have the background knowledge and practical skills to navigate the online resource, answering the original question simply requires following these steps:

1. **Perform a search for the gene of interest.**
2. **Using the Reference Atlas, find the structure(s) that expresses the gene of interest.**
3. **Use other resources to determine the function based on the structure or cell-type.**

Perform a search

From mouse.brain-map.org, type “dopa” into the search box. Notice that “dopa decarboxylase” is one of the suggested search terms. Clicking on this search term will return a list of experiments that fulfill the search criteria. Five experiments using probes against dopa decarboxylase (Ddc) are returned, which is rare since most genes only have a single experiment. Note the Probe Name column, which shows the riboprobe ID, and that three of the experiments were performed at Baylor, indicating they were part of the platform proof of concept. One of these experiments used a “sense” probe and serves as a negative control.

Determine the structure

Click on the experiment ID of the experiment conducted in the coronal plane. This takes you to the experiment page (Fig. 7) which includes all of the relevant information from

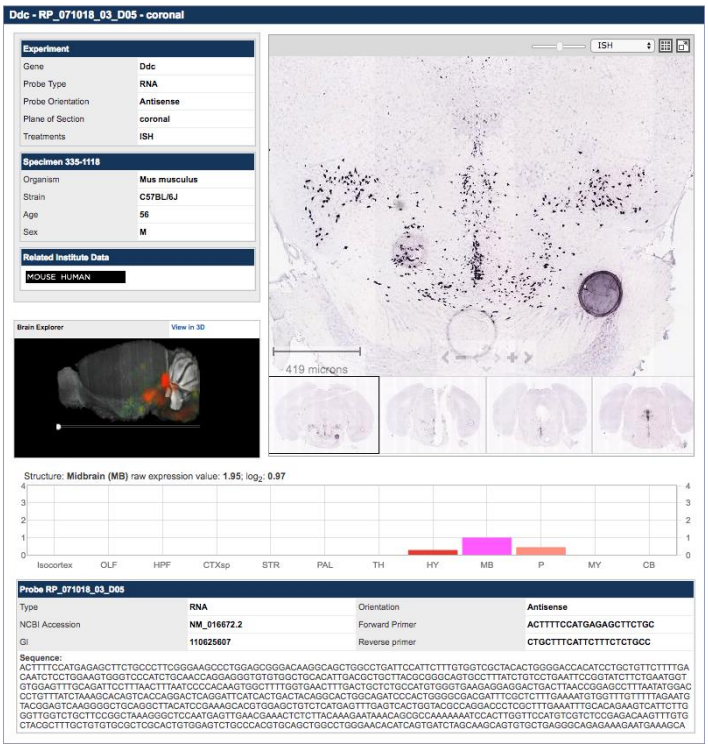

Figure 7. The experimental detail page from a coronal experiment probing for dopa decarboxylase (Ddc). This panel includes the image viewer (upper right), metadata about the experiment itself and a 3-D view of the gene expression (upper left), a histogram showing a semi-quantitative analysis of gene expression in coarse structures and metadata and sequence of the gene sequence used to create the riboprobe.

the experiment, including metadata, raw images and calculated expression values. Hovering over the columns in the histogram (i.e., MB) will bring up an image from the coarse structure that demonstrates gene expression; the midbrain in this case. Move through the thumbnails to find an image that shows discrete ISH positive expression. To further narrow down the structure(s) showing expression of Ddc, click the icon in the top right-hand corner of the image viewer to open a high-resolution image viewer. Clicking on the key icon in this window will open a side-by-side viewer of the ISH image with the Reference Atlas image.

Comparison of the gene expression pattern to the reference atlas suggests that Ddc shows relatively high expression in the ventral tegmental area and the substantia nigra compacta part (Fig. 8).

Research the Function of the Structures

Shallow review of the literature or of classic Neuroscience texts will describe the VTA and the SN as the only structures containing dopamine neurons and structures involved in the reward pathway. This exercise can stand alone or serve as a complement to lessons incorporating the reward pathway or the biochemistry of dopamine synthesis.

REFERENCES

Bernard A et al. (2012) Transcriptional architecture of the primate neocortex. Neuron 73:1083-1099.

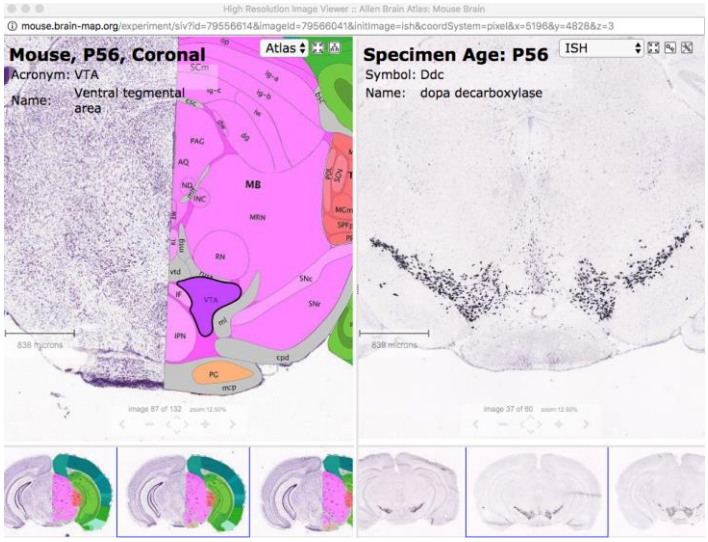

Figure 8. High Resolution Image Viewer showing expression of Ddc in the VTA and the SNc.

Chu P, Peck J, Brumberg JC (2015) Exercises in anatomy, connectivity, and morphology using Neuromorpho. org and the Allen Brain Atlas. J Undergrad Neurosci Educ 13:A95-A100.

Gilbert TL, Ng L (2018) The Allen Brain Atlas: toward understanding brain behavior and function through data acquisition, visualization, analysis, and integration. In: Molecular-genetic and statistical techniques for behavioral and neural research (Gerlai RT, ed) pp 51-72. Elsevier.

Grisham W, Schottler NA, Valli-Marill J, Beck L, Beatty J (2010) Teaching bioinformatics and neuroinformatics by using free web-based tools. CBE—Life Sci Educ 9:98-107.

Grisham W, Brumberg JC, Gilbert T, Lanyon L, Williams RW, Olivo R (2017) Teaching with big data: report from the 2016 Society for Neuroscience Teaching Workshop. J Undergrad Neurosci Educ 16:A68-A76.

Hawrylycz MJ et al. (2012) An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 489:391-399.

Jenks BG (2009) A self-study tutorial using the Allen Brain Explorer and Brain Atlas to teach concepts of mammalian neuroanatomy and brain function. J Undergrad Neurosci Educ 8:A21-A25.

Lein ES et al. (2007) Genome-wide atlas of gene expression in the adult mouse brain. Nature 445:168-176.

Miller JA et al. (2014) Transcriptional landscape of the prenatal human brain. Nature 508:199-206.

Miller JA, Guillozet-Bongaarts A, Gibbons LE, Postupna N, Renz A, Beller AE, Sunkin SM, Ng L, Rose SE, Smith KA (2017) Neuropathological and transcriptomic characteristics of the aged brain. Elife 6.

Ng L, Pathak SD, Kuan C, Lau C, Dong H, Sodt A, Dang C, Avants B, Yushkevich P, Gee JC, Haynor D, Lein E, Jones A, Hawrylycz M (2007) Neuroinformatics for genome-wide 3-D gene expression mapping in the mouse brain. IEEE/ACM Trans Comput Biol Bioinform (TCBB) 4:382-393.

Oh SW et al. (2014) A mesoscale connectome of the mouse brain. Nature 508:207-214.

Ramos RL, Smith PT, Brumberg JC (2007) Novel in silico method for teaching cytoarchitecture, cellular diversity, and gene expression in the mammalian brain. J Undergrad Neurosci Educ 6:A8-A13.

Sengul G, Puchalski RB, Watson C (2012) Cytoarchitecture of the spinal cord of the postnatal (P4) mouse. Anat Rec (Hoboken) 295:837-845.

Thompson CL et al. (2014) A high-resolution spatiotemporal atlas

of gene expression of the developing mouse brain. Neuron 83:309-323.

Received July 23, 2018; revised August 08, 2018; accepted August 08, 2018.

The Human Brain Reference Atlas was supported by Award Numbers RC2MH089921 (PIs: Ed Lein & Michael Hawrylycz, Allen Institute for Brain Science), RC2MH090047 (PI: James A. Knowles, University of Southern California) and RC2MH089929 (PI: Nenad Sestan, Yale University) from the National Institute of Mental Health.

The author would like to extend a most profound acknowledgement to Paul G. Allen for his vision and his generosity in funding such a worthy endeavor as the Allen Institute.

Address correspondence to: Dr. Terri L Gilbert, GilHou Science Communication, 1400 SW Thistle St, Seattle, WA, 98106. Email: terri@terrigilbertphd.com