

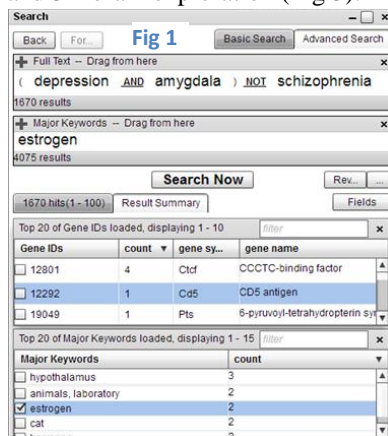
PubAnatomy 3D: Integrating Medline Exploration with the Allen Mouse Brain Atlas

Yang Gang¹, Manhong Dai¹, Jean Song², Barbara Mirel³, Fan Meng^{1,4,5}

¹Molecular and Behavioral Neuroscience Institute, ²Health Sciences Library, ³School of Education, ⁴Psychiatry Department and ⁵Center for Computational Medicine and Bioinformatics, University of Michigan

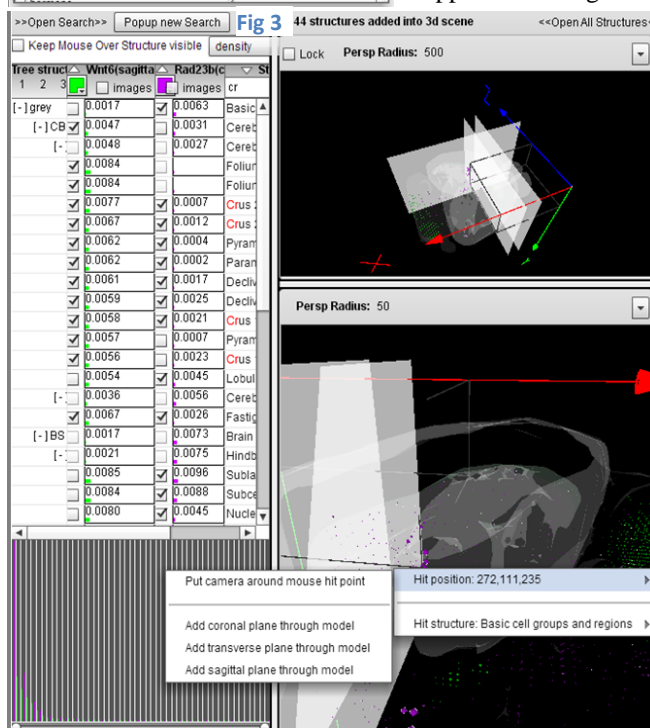
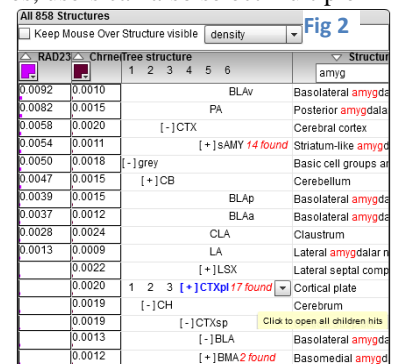
Neurobiology data generated by the big science approaches, such as the Allen brain atlases and the NIH Human Connectome Project, enables the understanding how brain works at both molecular and anatomical levels with unprecedented resolution and completeness. For example, researchers can now use the Brain Explorer and the NeuroBlast tools from the Allen Institute for Brain Science to explore genes and their spatial expression pattern across the whole brain for learning the potential biological implications of genes identified in their experiments and designing follow-up experiments, such as knockout/knockin mouse for modeling human diseases. However, since it is impossible for individual scientists to grasp all the known functional roles of genes, anatomical structures and the functional relationships among them, researchers frequently need to perform literature searches to guide big data set explorations. Similarly, literature exploration can also benefit significantly from the relationships among genes and anatomical structures presented in such data sets. The need to use two separate data exploration tools such as the Brain Explorer and PubMed leads to technical hurdles as well as gaps in the thinking process that can significantly constrain the hypothesis development process.

PubAnatomy 3D is designed to provide a seamless exploration environment across the Allen Mouse Atlas data and the Medline literature for iterative data- and literature-guided hypothesis development. We mapped genes and anatomical structures in the Allen Mouse Atlas data set to individual Medline records and developed a flexible web-based search interface for integrated Medline and mouse atlas data exploration. A typical use case is a researcher starting with Medline search for their interested topic, such as diseases and brain structures (Fig 1, upper part), to obtain a list of Medline records, which is annotated with different concept categories such as genes as well as summary statistics, e.g., number of Medline records associated with each gene in the search results (Fig 1, lower part). Researchers can use the summary statistics, various filtering criteria and the content of relevant Medline records to select genes and anatomical structures through simple drag-and-drop for evaluating gene expression levels and brain structures (Fig 2) and 3D brain exploration (Fig 3). Besides displaying the voxel level data for selected brain structures, users can also select multiple



arbitrary 2D intersections in the coronal and sagittal directions to view the raw *in situ* images (Fig 3). Two or more genes can be displayed side-by-side for detailed raw *in situ* data inspection together with brain structure annotations for each 2D intersection (Fig 4). Researchers can use new structures or additional genes identified in the brain data exploration process to filter Medline search results or to start new or modify existing Medline search through simple drag-and-drop, greatly facilitating the iterative literature and data exploration during hypothesis development process.

Pubanatomy3D is available at <http://brainarray.mbni.med.umich.edu/PubAnatomy3D/>. We plan to define and publish the Application Programming Interface for Pubanatomy 3D to enable third party developers to



access data and functions in PubAnatomy as well as passing their own data such as gene or SNP lists to PubAnatomy 3D. We will also integrate more data sets, such as connectome, pathway and protein interaction, Gene Ontology, etc. into PubAnatomy to further enhance its usefulness for hypothesis development in neurobiology.

