So... I didn't find a version of the pre-created Allen Brain Atlas gene-gene correlation network (yet), but I found something almost as good (or perhaps better?).

1) The ability to extract the experimental ids for all of the genes in the ABA in situ hybridization data as a .csv file and narrow them down by various characteristics (e.g., not deemed to have "Failed").

2) The ability to query for the correlations of a gene (experimental id) with all other genes (experimental Ids) in the ABA, filtering for things like whether to only include data from coronal sections.

... so if we wanted to play with this data in bulk, we would still need:

1) A wrapper function of some sort (something that uses the experiment ids to automatically run the http query and save it as a text file)

2) If we wanted to go further into building a network, we would need to actually covert the coexpression R values into an adjacency matrix and then use the adjacency matrix to construct a network.

... or we could just ask simpler questions like:

      A) Which of our top differentially-expressed genes are known to be spatially-coexpressed? (and where?)

      B) Is there greater spatial co-expression amongst our top genes than in our dataset as a whole? (to make sure this isn't silly and circular, we may need to filter our data first by some indication that the genes included in the analysis are definitely expressed/detectable in the hippocampus.

So... why this is potentially exciting:

Similar code can be used to query the ABA developmental datasets and human datasets.

With the developmental datasets, we might be able to ask the question not only regarding spatial correlation but temporal correlation.

... but just finding a list of the pre-identified modules in the network would still certainly be easier.

Cheers,

Megan

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Sample code & other information:

1) The ABA "Correlative Search" GUI:

How to (point & click) search for correlating genes for any single gene in ABA in situ hybridization data (or other ABA data):

<http://help.brain-map.org/display/mousebrain/In+Situ+Hybridization+%28ISH%29+Data#InSituHybridization(ISH)Data-CorrelativeSearch>

Or... to find genes specifically enriched in particular areas (e.g., layers of the hippocampus):

<http://help.brain-map.org/pages/viewpage.action?pageId=2424836>

2) How to automate it:

<http://help.brain-map.org/display/mousebrain/API>

A) First, we need to extract a list of "dataset IDs" identifying the in situ hybridization experiments ("SectionDataSet") characterizing our genes of interest (or genes in our dataset)

These can be found by hand via their GUI when doing a gene search or bulk search on their website (the number on the left hand side - labeled "Experiment"):

<http://mouse.brain-map.org/search/show?page_num=0&page_size=20&no_paging=false&search_term=Per2$Per1$Arntl$Fos$Dbp$Tef$$&search_type=bulk>

... but here is some example code for downloading the experiment IDs in bulk:

<http://help.brain-map.org/display/api/Example+Queries+for+Experiment+Metadata#ExampleQueriesforExperimentMetadata-AllMouseBrainISHexperimentsincommadelimitedformat>

<http://api.brain-map.org/api/v2/data/query.csv?criteria=>

model::SectionDataSet,

rma::criteria,[failed$eqfalse],products[abbreviation$eq'Mouse'],treatments[name$eq'ISH'],genes,plane\_of\_section,

rma::options,

[tabular$eq'[plane\_of\_sections.name](http://plane_of_sections.name/)+as+plane','genes.acronym+as+gene','[data\_sets.id](http://data_sets.id/)+as+section\_data\_set\_id'],

[order$eq'[plane\_of\_sections.name](http://plane_of_sections.name/),genes.acronym,[data\_sets.id](http://data_sets.id/)']

&num\_rows=50&start\_row=0

The num\_rows argument will control how many experiments are included in the output. We can also change the output to only include coronal experiments:

plane\_of\_section[name$eq'coronal']

E.g.: Like this:

<http://api.brain-map.org/api/v2/data/query.csv?criteria=>

model::SectionDataSet,

rma::criteria,[failed$eqfalse],products[abbreviation$eq'Mouse'],treatments[name$eq'ISH'],genes,plane\_of\_section[name$eq'coronal'],  
rma::options,

[tabular$eq'[plane\_of\_sections.name](http://plane_of_sections.name/)+as+plane','genes.acronym+as+gene','[data\_sets.id](http://data_sets.id/)+as+section\_data\_set\_id'],

[order$eq'[plane\_of\_sections.name](http://plane_of_sections.name/),genes.acronym,[data\_sets.id](http://data_sets.id/)']

&num\_rows=5000&start\_row=0

B) Then we could extract the correlation coefficients for each gene vs. other genes in the ABA in situ hybridization dataset:

"To perform a **Correlation search**, a user selects a seed SectionDataSet and a domain over which the similarity comparison is to be made. All voxels belonging to any of the domain structures form the domain voxel set. Pearson's correlation coefficient is computed between the domain voxel set from the seed SectionDataSet and every other SectionDataSet in the Product. The return list is sorted by descending correlation coefficient.

**Example:** Correlation search for genes with similar expression to the sagittal Pdyn SectionDataSet

* [Pipe: Set up the seed SectionDataSet by finding the sagittial SectionDataSet for gene Pdyn](http://api.brain-map.org/api/v2/data/query.xml?criteria=model::SectionDataSet,rma::criteria,genes%5Bacronym$eq%27Pdyn%27%5D,products%5Babbreviation$eq%27Mouse%27%5D,plane_of_section%5Bname$eq%27sagittal%27%5D)
* [Connect the pipe to **service::mouse\_correlation** to perform the correlation search](http://api.brain-map.org/api/v2/data/query.xml?criteria=model::SectionDataSet,rma::criteria,genes%5Bacronym$eq%27Pdyn%27%5D,products%5Babbreviation$eq%27Mouse%27%5D,plane_of_section%5Bname$eq%27sagittal%27%5D,pipe::list%5Bseed$eq%27id%27%5D,service::mouse_correlation%5Bset$eqmouse%5D%5Brow$eq$seed%5D%5Bstart_row$eq0%5D%5Bnum_rows$eq50%5D%5Bsort_order$eqdesc%5D)
* [Visualize the same search result in the Web application](http://mouse.brain-map.org/search/show?page_num=0&page_size=32&no_paging=false&domain=8&rows=69782969&image_set=P56&search_type=neuroblast)

See the [connected service](http://help.brain-map.org/display/api/Connected+Services+and+Pipes#ConnectedServicesandPipes-service::mouse_correlation) page for definitions of service::mouse\_correlation parameters.

Parameters:

<http://help.brain-map.org/display/api/Connected+Services+and+Pipes#ConnectedServicesandPipes-service::mouse_correlation>

service::mouse\_correlation

**Parameters**

|  |  |  |
| --- | --- | --- |
| set | String | Data set name: mouse or mouse\_coronal. |
| structures | Array (optional) | Comma delimited list of Structure.id or Structure.acronym. |
| row | Integer | DataSet.id to correlate against. |
| sort\_order | String (optional) | 'asc' or 'desc'. Defaults to 'desc'. |
| start\_row | Integer (optional) | Starting row of data for paging purposes. Defaults to 0. |
| num\_rows | Integer (optional) | Number of rows of data for paging purposes. Defaults to 2000. |

**Example**

<http://api.brain-map.org/api/v2/data/query.xml?criteria=>

service::mouse\_correlation[set$eqmouse][row$eq68918934]

<http://api.brain-map.org/api/v2/data/query.xml?criteria=>

service::mouse\_correlation[set$eqmouse\_coronal][row$eq357096]

Megan's Notes:

1) The "row" is the experiment id.

2) If you make the example say query.csv instead of query.xml you get .csv output.

E.g.

<http://api.brain-map.org/api/v2/data/query.csv?criteria=>

service::mouse\_correlation[set$eqmouse\_coronal][row$eq357096]