API

ALLEN BRAIN ATLAS API

- Experimental Overview and Metadata
  - Supplemental RNA-Sequencing Data
- Downloading Expression Values
- Differential Search
- Correlative Search
- Magnetic Resonance Imaging

The **Allen Human Brain Atlas** is a multimodal atlas of the human brain that integrates anatomic and microarray-based gene expression information. Microarray sampling sites (~400-1000 sites per brain) were identified by expert anatomists using cytoarchitectural information from multiple histological stains. Sampling site delineations in the high resolution histological images were subsequently mapped into each individual's MR image space to provide 3-D anatomical context. All brains were also registered to MNI space to enable cross-individual comparisons.

From the API, you can:

Download expression values

Query the correlative and differential search services

Download MRI images

Experimental Overview and Metadata

RNA isolated from each sample area was hybridized to a custom Agilent 8x60k microarray chip to measure gene expression over the transcriptome. All least two different probes were available for 93% of genes. Probes were located on different exons as much as possible when multiple probes were available for a gene. For 60 genes, sets of tiling probes were designed.

Each sampling site was associated to a Structure by expert anatomists using cytoarchitectural information from multiple histological stains. Structures are organized hierarchically into a tree in which children structures are "parts of" their parent structure. Structures are assigned colors that visually emphasize the hierarchical relationships.

See the structure ontology page for more information.

Gene expression data for samples passing quality control are normalized to enable cross-comparison between batches of samples processed at different times or samples belonging to different donors. For more details on microarray data generation and processing see the Microarray whitepapers.

Normalized microarray expression values can be downloaded in several ways:

- From the web application Download page. Each zip file contains the normalized values for all probes and all samples for one donor)
- From the "Download this data" link below the heatmap in the web application.
- From the connected data service in the API

All experimental data from this study is associated with the "Human Brain Microarray" Product. All probe and sampling site information can be accessed through the API using RMA queries.

Example queries:

- All donors in the Product
- All sampling sites in the Product
- All microarray probes in the Product
- All microarray probes associated with gene prodynorhphin (PDYN)
- All samples associated with donor "H0351.2001" and the dentate gyrus
- Download an "raw" Aglient output file for one sample using the download-link (warning: large file)

Supplemental RNA-Sequencing Data

RNA-Sequencing (RNA-Seq) data were generated for a selected set of 240 samples (120 from each brain) representing matched cortical and sub-cortical regions across two brains (H0351.2001 and H0351.2002). The gene expression data (both raw and TPM counts) can be downloaded from the web application Download page.

Through a quantitative comparison of microarray and RNA-Seq data, a set of quality control metrics has been computed for each Agilent microarray probe which allows a user to filter out problematic probes or choose the most reliable probe for each gene. The probe metric table and metadata can be downloaded here.

Downloading Expression Values

Normalized expression values can be obtained by specifying:

- a list of probes
- a list of donors (optional), and
- a list of structures (optional)

See the connected service page for definitions of service::human_microarray_expression parameters.

Example:

Download expression values for donor "H0351.1015" in structure "locus ceruleus" for all probes associated with gene SLC6A.

- Find Donor ID for "H0351.1015" (id = 15496)

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=model::Donor,
rma::criteria,[name$eq'H0351.1015'],
rma::options[only$eq'donors.id']
```

- Find Structure ID for "locus ceruleus" (id = 9148)

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=model::Struct
ure,
rma::criteria,[name$il'locus ceruleus'],ontology[name$eq'Human Brain
Atlas'],
rma::options[only$eq'structures.id']
```

- Find Probes associated with gene SLC6A2 (id = 1023146,1023147)

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=model::Probe,
rma::criteria,[probe_type$eq'DNA'],products[abbreviation$eq'HumanMA']
,gene[acronym$eq'SLC6A2'],
rma::options[only$eq'probes.id']
```

- Use Donor, Structure and Probe IDs as parameters to the connected service

```
http://api.brain-map.org/api/v2/data/query.json?criteria=
service::human_microarray_expression[probes$eq1023146,1023147][donors
$eq15496][structures$eq9148]
```

The output of the service is two top level ordered arrays "probes" and "samples". For example:

```
"probes":[{
    "id":1023146,
    "name":"A_23_P358345",
    "gene-id":6494,
    "gene-symbol":"SLC6A2",
    "gene-name":"solute carrier family 6 (neurotransmitter transporter,
noradrenalin), member 2",

"entrez-id":6530,"chromosome":"16","start-position":"n/a","end-position":"
n/a",
    "expression_level":["13.2802","13.9603","13.9650"],
    "z-score":["9.3381","9.8663","9.8700"]},
    {
    "id":1023147,
    "name":"CUST_16472_PI416261804",
    "gene-id":6494,
    "gene-symbol":"SLC6A2",
    "gene-name":"solute carrier family 6 (neurotransmitter transporter,
noradrenalin), member 2",

"entrez-id":6530,"chromosome":"16","start-position":"n/a","end-position":"
n/a",
    "expression_level":["8.1878","8.5644","8.2310"],
    "z-score":["9.3201","9.8326","9.3790"]}
    ],
"samples":[
    {"donor":{"id":15496,"name":"H0351.1015","age":"49
years","color":"C2C200"},
    "sample":
        {"well":148955246,"polygon":127107914,"mri":[95,121,126]},
        "structure":{"id":9149,"name":"locus ceruleus,
Left","abbreviation":"LC","color":"00FFAA"},
        "top_level_structure":{"id":9135,"name":"Pontine
Tegmentum","abbreviation":"PTg","color":"00FFAA"}},
    {"donor":{"id":15496,"name":"H0351.1015","age":"49
years","color":"C2C200"},
    "sample":
        {"well":148955204,"polygon":126786164,"mri":[97,151,131]},
        "structure":{"id":9149,"name":"locus ceruleus,
Left","abbreviation":"LC","color":"00FFAA"},
        "top_level_structure":{"id":9135,"name":"Pontine
Tegmentum","abbreviation":"PTg","color":"00FFAA"}},
    {"donor":{"id":15496,"name":"H0351.1015","age":"49
years","color":"C2C200"},
    "sample":
        {"well":156435966,"polygon":126789834,"mri":[96,159,134]},
        "structure":{"id":9149,"name":"locus ceruleus,
Left","abbreviation":"LC","color":"00FFAA"},
        "top_level_structure":{"id":9135,"name":"Pontine
Tegmentum","abbreviation":"PTg","color":"00FFAA"}}
    ]
```

Each **probe** contains information about:

- the Probe(id, name), and
- the Gene (id, acronym, name, entrez-id), along with
- a vector of normalized expression values in the same order as the "samples" array.
- a vector of z-score values in the same order as the "samples" array. Note: z-score is computed independently for each probe over all donors and samples.

Each **sample** contains information about:

- the Donor (id, name, age),
- the Sample (well id and (x,y,z) coordinate in the MR volume in millimeters),
- the associated Structure (id, name, acronym and color), and
- the associated top (coarse) level Structure (id, name, acronym and color).

Differential Search

The differential search function finds probes that show the greatest difference between two sets (target and contrast) of user-defined structures. For each probe, a 2-sample t-test is performed followed by Benjamini and Hochberg false discovery rate correction. The null hypothesis is that the average expression level of samples in the contrast set of structures is greater than or equal to the average expression level of samples in the target set of structures. A statistically significant result (p-value less than user-defined threshold) allows us to reject the null hypothesis and conclude that the average expression level of samples in the target set of structures is greater than the average expression level of samples in the contrast set of structures. Resulting p-values are sorted in ascending order. Search results can also be sorted by fold-change (log ratio of expression) in descending order.

The differential search function can be accessed through the Web application or using the API.

See the connected service page for definitions of service::human_microarray_differential parameters.

Example:
Differential search for genes with higher expression in thalamus than the cerebral cortex

- Find Structure ID for "thalamus" (id = 4392)

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=model::Structure,
rma::criteria,[name$il'thalamus'],ontology[name$eq'Human Brain Atlas'],
rma::options[only$eq'structures.id']
```

- Find Structure ID for "cerebral cortex" (id = 4008)

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=model::Structure,
rma::criteria,[name$il'cerebral cortex'],ontology[name$eq'Human Brain Atlas'],
rma::options[only$eq'structures.id']
```

- Use Structure IDs as parameters to the connected service

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=service::human_microarray_differential
[structures1$eq4008][structures2$eq4392][sort_by$eq'fold-change']
```

- Visualize the same search result in the Web application

Figure: Screenshot of top returns of a differential search for genes with higher expression in the thalamus than in the cerebral cortex. The z-score heatmap shows enrichment (red) in the thalamus (structure color: light green) relative to other brain regions.

Usage of this service is demonstrated in the SPM example application.

Correlative Search

The correlative search function finds probes with expression profile similar to that of a selected seed probe over all samples within a user-specified structure. Pearson's correlation coefficient is computed for all probes and the results ranked in descending order.

This correlative search function can be access through the Web application or using the API.

See the connected service page for definitions of service::human_microarray_correlation parameters.

Example:
Correlative search for probes with similar expression to PVALB probe CUST_11451_PI416261804 over the whole Brain

- Find Probe ID for probe "CUST_11451_PI416261804" (id = 1052410)

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=model::Probe,
rma::criteria,[name$eq'CUST_11451_PI416261804'],
rma::options[only$eq'probes.id']
```

- Find Structure ID for "Brain" (id = 4005)

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=model::Struct
ure,
rma::criteria,[name$il'brain'],ontology[name$eq'Human Brain Atlas'],
rma::options[only$eq'structures.id']
```

- Use Probe and Structure IDs as parameters to the connected service

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=service::huma
n_microarray_correlation
[probes$eq1052410][structures$eq4005]
```

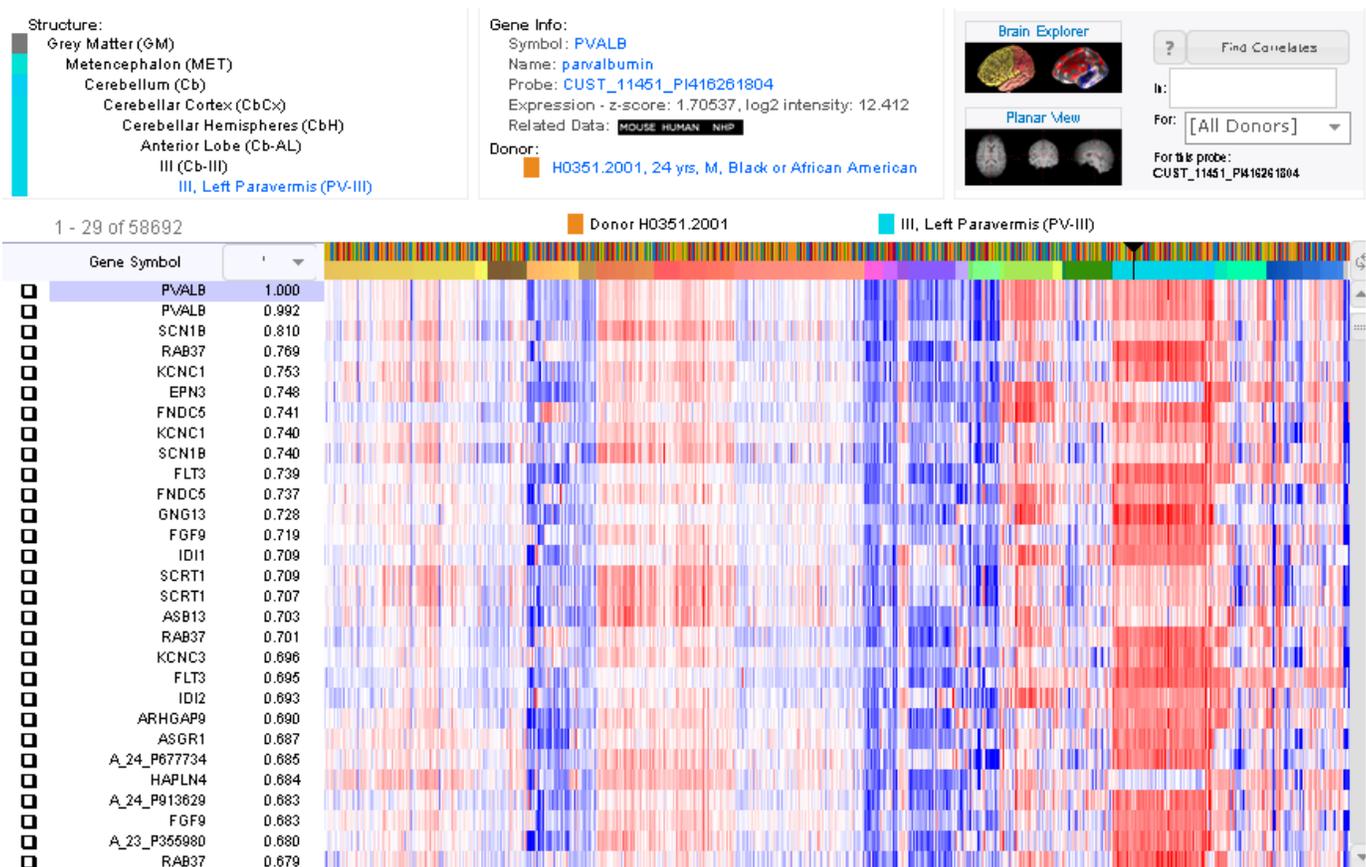- Visualize the same search result in the Web application

Figure: Screenshot of top returns of a correlative search for probes with similar expression to a PVALB probe with expression values displayed as a z-score heatmap.

Magnetic Resonance Imaging

T1-weighted MPRAGE scans were acquired for the postmortem brains using 3T Siemens Trio MR scanners (TI=900ms, TR=1900ms, TE=3.03ms, 9 degree flip angle, 1mm isotropic voxels). Scans were performed *in cranio* for some brains and *ex cranio* for others. See the Microarray whitepapers for more specific scan sequence details for each brain.

The T1, T2 and DTI (if available) volumetric data can be downloaded from the Web application or via the API.

All T1 images were registered to MNI space. FreeSurfer's affine registration was used for the *in cranio* scans. For *ex cranio* brains, the T1 was first rigidly aligned using FSL (Jenkinson, et. al, 2002) and then non-rigidly aligned using ANTS (Avants, et. al., 2011). The 3-D affine transform from a location in the MR volume to MNI space is encapsulated in the Alignment3d model.

Examples queries:

- Download links for all MR and DTI data available

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=model::Donor,
rma::criteria,products[abbreviation$in'HumanMA','HumanSZ','HumanCtx',
'HumanSubCtx'],organism[name$il'Homo Sapiens'],
rma::include,specimens(well_known_files(well_known_file_type[name$in'
T1-MRI','T2-MRI','DTI-MRI'])),
rma::options[only$eq'donors.name,products.name,well_known_files.downl
oad_link,specimens.id']
```

- Download TI MR scan for donor 'H0351.2002'

```
http://human.brain-map.org/api/v2/well_known_file_download/157723301
```

- Download the MR to MNI transform parameters

```
http://api.brain-map.org/api/v2/data/query.xml?criteria=model::Donor,
rma::criteria,products[abbreviation$eq'HumanMA'],
rma::include,specimens[parent_id$eqnull](alignment3d),
rma::options[only$eq'donors.id,donors.name,specimens.id']
```
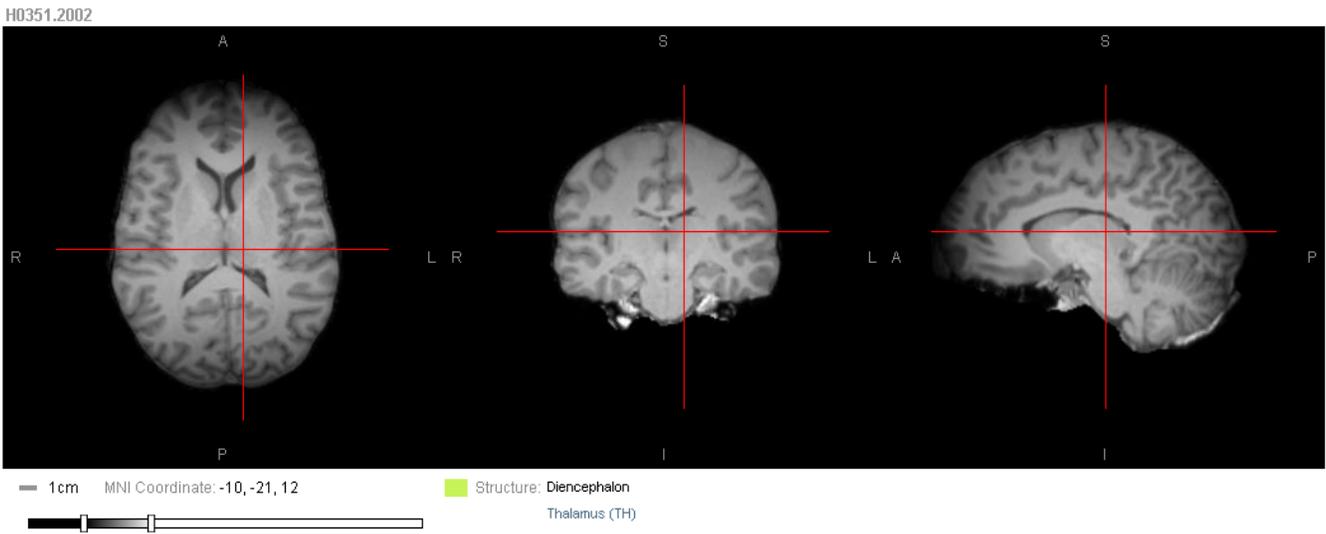

Figure: T1 MR scan for donor 'H0351.2002'

See example code on how to transform each microarray sample to MNI space.