Introduction

To probe the depth to which ISH gene expression can be characterized on a genome-wide scale, comprehensive manual annotation of the mouse somatosensory cortex (SS) was performed. Image series data representing all available unique gene probes of the Allen Brain Atlas were identified, assessed for quality and categorized by general expression characteristics. Those experiments exhibiting laminar, sub-regional and/or interneuron-like patterning in SS were subsequently analyzed in greater detail by a team of trained annotators. A standardized density/intensity expression scoring system was employed for all layers and sub-regions of primary and secondary cortices. Sub-laminar and composite expression patterning was also noted when saliently apparent. A layer-collated database of SS metrics for the 3,300 scored genes can be found at http://www.brain-map.org/somatosensory.do.

Methods

I. Cortical Genome Survey of ABA Dataset

Representative ISH images for all available image series of the Allen Brain Atlas were retrieved and organized algorithmically for manual assessment of general expression characteristics. Sections from predetermined atlas levels representing a broad range of cortex were shown at a resolution sufficient for determining *a)* data quality *b)* yes/no expression in cortex and *c)* yes/no SS cortical patterning (discontinuous expression across layers and/or SS sub-regions or general scattering suggestive of interneuron cell types). ISH image sets exhibiting quality too poor for this analysis were excluded from the data set.

II. Generation of Somatosensory-Patterned Genes List for Deep Annotation

The initial list of image series positive for SS patterning was organized and condensed in an effort to minimize deep annotation redundancies. Because many genes were run multiple times in multiple planes, some with aliases, the SS-Patterned list was arranged by template ID, a unique probe sequence identifier from which just one image series would be subsequently analyzed. The choice of which series was made on the basis of *1)* plane of section and *2)* date of generation: coronals over sagittals, newest over older.

III. Deep Annotation of Somatosensory-Patterned Genes

Comprehensive scoring of somatosensory cortex was performed for 3,300 image series identified in I and II using the brainmap.org detailed viewer. A layer-by-region data matrix was devised to record observed expression and density values for downstream informatics analysis and manipulation. Defined layers and regions are:

```
Layers:  1                    Regions:    SS Primary
         2                                    -(General Region)
         3                                    -Trunk
         4                                    -Barrel Field
         5                                    -Lower Limb
         6                                    -Mouth
         6b                                   -Upper Limb
                                              -Nose
                                          SS Secondary
                                              -(General Region)
```

In order to observe seminal expression differences between the various fields as well as increase overall scoring consistency, multiple sections were considered during the deep annotation process. For sagittal image series, a complete analysis of primary subfields was not possible and the values for SSp sub-regions were not included. For each layer-region, a density and intensity expression metric was given using the following guideline:

Sublaminarity
    Uniform: pattern uniform across defined space
    Superficial: pattern preferentially dorsal
    Deep: pattern preferentially ventral
    Mid: pattern residing mid-space within a lamina
    Superficial and Deep: pattern except in mid-space

Density (for a given intensity)
    High: more than 80% of potential cells expressed
    Medium: 30-80% expressed
    Scattered: Fewer than 30% expressed
    Sparse: Far fewer than 10% expressed

Intensity
    Salient, but low expression: 1-2
    Moderate expression: 3-4
    Complete saturation: 5

Occasionally a layer-region exhibited concurrent, but distinct expression patterns. These were included in the final collated dataset as separate density and intensity values. "Primary" was given to the pattern exhibiting the higher density.