

# User Guide: ALS-ST.NYGENOME.ORG

# Index

- Introduction
- General Interface
  - Selecting a view
  - Zoom and Pan
  - Gene symbol entry
- Views
  - Arrays
    - Description of view
    - Array selection
    - Color scale
  - Overlay
    - Description of view
    - Color scale
    - Adjusting dynamic range
  - Coexpression
    - Description of view
    - Color manipulation
  - Coefficient
    - Description of view
    - Adjusting expression scale
  - Expression Trajectory
    - Description of view
    - Adjusting cutoffs

# Introduction

ALS-ST.NYGENOME.ORG is a suite of interactive visualization tools for exploring the data presented in

**Spatiotemporal Dynamics of Molecular Pathology in Amyotrophic Lateral Sclerosis**

Silas Maniatis, Tarmo Aijo, Sanja Vickovic, Catherine Braine, Kristy Kang, Annelie Mollbrink, Zaneta Andrusivova, Sami Saarenpaa, Gonzalo Saiz-Castro, Miguel Cuevas, Aaron Watters, Joakim Lundeberg, Richard Bonneau, Hemali Phatnani

**doi:** <https://doi.org/10.1101/389270>

The intent of this suite of tools is to make our data and analysis easily accessible to non-specialists, and to users without computational expertise.

For a detailed description of the samples and mathematical parameters used, please refer to the supplementary methods presented with the manuscript.

# General Interface

The suite of tools is divided into views, each of which is intended to visualize data in a way that is best suited for particular types of analyses. These views are selected from the home page of [als-st.nygenome.org](http://als-st.nygenome.org).

**The available views are:**

- Mouse arrays
  - Visualization of a single gene's expression in the context of histology images for individual ST arrays for the mouse data set.
- Human arrays
  - Visualization of a single gene's expression in the context of histology images for individual ST arrays for the human data set.
- Mouse overlay
  - Visualization of a single gene's expression for all ST spots from the mouse dataset. ST spots are spatially registered using a common coordinate system, and data from all arrays for each condition are overlaid.
- Human overlay
  - Visualization of a single gene's expression for all ST spots from the mouse dataset. ST spots are spatially registered using a common coordinate system, and data from all arrays for each condition are overlaid.
- Mouse coexpression
  - Visualization of expression for several genes simultaneously for all ST spots from the mouse dataset. ST spots are spatially registered using a common coordinate system, and data from all arrays for each condition are overlaid.
- Coefficients
  - Visualization of a single gene's expression from our novel Bayesian generative model. Expression is reported for each anatomical region, time point, and genotype distinctly and simultaneously, allowing for cross-condition comparisons.
- Trajectories
  - Visualization of differences in a single gene's expression between SOD1-WT and SOD1-G93A across time within each anatomical region.

# General Interface

Each view has its own set of parameters that can be adjusted. In the following sections, you will find descriptions of these. However, several interface elements are shared across all tools.



- Gene symbol entry
  - Requires official NCBI gene symbols (case sensitive)
    - *Autofill feature has known issues*
- Pan
  - When selected, allows the field of view to be adjusted by mouse click and drag
- Zoom
  - When selected, allows the zoom to be adjusted by mouse wheel
- Reset
  - When clicked, resets all parameter adjustments and view settings to default

# Views: Arrays

## Description:

The “Mouse Arrays” and “Human Arrays” views represent expression (rate  $\lambda$ ) for a single gene on a per-spot basis, in the context of the histology image acquired for a given array. Circles with 100micron diameter are drawn at the coordinate for each ST spot, and each spot is assigned the color corresponding to the expression rate measured for the selected gene for that ST spot. The actual area covered by the array feature may differ slightly from this representation, as the array features are not always exactly circular or 100microns in diameter. Expression rate as visualized can be compared across spots, arrays, and genes. However, this measure does not represent the data as fully modeled by our method. As such, substantial inter-array variability will be present. This view is therefore best for obtaining a general impression of expression in the context of the information conveyed by the histology image for a single array. For a view of spatial expression across the entire dataset, the overlay or coefficients views are a better choice.

# Views: Arrays

## Array selection

Arrays are selected through several dropdowns at the top of the page. The rightmost dropdown selects individual arrays within the condition selected through filtering through the dropdowns to the left.

## Color scale

The color scale represents expression rate, with low/no expression (rate  $\lambda = 0$ ) colored dark purple, and the highest expression rate measured for the selected gene in the entire dataset colored bright yellow.

# Views: Overlay

## Description:

The “Mouse Overlay” and “Human Overlay” views represent expression (rate  $\lambda$ ) for a single gene on a per-spot basis, for all spots in either dataset. Spatial coordinates for spots mapping to each tissue section are transformed by a registration procedure described in the manuscript. Circles are drawn at the registered coordinate for each ST spot, and each spot is assigned the color corresponding to the expression rate measured for the selected gene for that ST spot. All spots within a given condition are drawn on the same plot, resulting in an “overlaid” view using a common coordinate system for all spots from all tissue sections in each condition. Expression rates can be compared across spots, conditions, and genes. This view is useful for obtaining a general impression of spatio-temporal dynamics of gene expression for a single gene.

# Views: Overlay

## Color scale

The color scale represents expression rate, with low/no expression (rate  $\lambda = 0$ ) colored black, and the highest expression rate measured for the selected gene in the entire dataset colored bright yellow.

## Adjusting dynamic range

The limits of the color scale can be adjusted using the “Colormap Limits” slider. This is particularly useful for resolving temporal changes for poorly expressed (maximum rate  $\lambda \leq .15$ ) or very highly expressed (maximum rate  $\lambda \geq 2$ ) genes.

# Views: Mouse Coexpression

## Description:

This view represents expression (rate  $\lambda$ ) for up to 5 genes on a per-spot basis, for all spots in either dataset. Spatial coordinates for spots mapping to each tissue section are transformed by a registration procedure described in the manuscript. Circles are drawn at the registered coordinate for each ST spot, and each spot is assigned the color corresponding to the expression rate measured for the selected gene for that ST spot. The color scales are on a per-gene basis and represent expression from rate  $\lambda = 0$  (black) to the highest rate  $\lambda$  observed for that gene for the entire dataset (encoded by the maximum luminosity for the selected color). All spots within a given condition are drawn on the same plot, resulting in an “overlaid” view using a common coordinate system for all spots from all tissue sections in each condition. The plotted color represents the mixture resulting from blending the color values for all plotted genes at each spot. This view is useful for obtaining an *“in silico FISH”* impression of spatio-temporal dynamics of gene expression for up to 5 genes.

# Views: Mouse Coexpression

## Color manipulation

The color scales are on a per-gene basis and represent expression rate from  $\lambda = 0$  (black) to the highest rate  $\lambda$  observed for that gene in the entire dataset (encoded by the maximum luminosity for the selected color). Look up tables (colors) can be assigned through the dropdown menus. The Gamma value will adjust the histogram of the lookup table (Lower values result in more intense color).

# Views: Coefficient

## Description:

This view depicts the modeled expression (The posterior distributions of coefficient parameters  $\beta$ ) for a single gene, for all spots in the mouse data set. The coefficient parameters  $\beta$  capture offsets of expression (in natural logarithmic space) in distinct anatomical annotation areas across all tissue sections of a given condition. Another way to understand these plots is that the curves represent probabilities that expression in a given condition is at a given level. The position of the peak on the horizontal axis gives the most likely expression level, while the height and width of the curve can be understood as the model's confidence in calling expression at that level. Tall, narrow peaks represent high confidence that expression of the selected gene is at the level indicated by the peak, while broad peaks represent low confidence. Offsets between the curves for different conditions are indicative of differential expression. That is to say, substantial overlap between the areas under curves for different conditions indicates low probability that expression differs between the conditions, while small or no overlap indicates higher probability of the gene being differentially expressed between the conditions. Expression can be compared across conditions and genes on the same scale. The curves can be organized either by anatomical annotation area, or by genotype by selecting the tabs at the top of the page. This view is useful for quickly understanding what the model says about spatiotemporal dynamics of gene expression across conditions, and for quickly assessing differential expression.

# Views: Coefficient

## Adjusting expression scale

The horizontal axis represents modeled expression (The posterior distributions of coefficient parameters  $\beta$ ) for a single gene. To adjust the limits of this scale, one can manipulate the “X-axis limits” sliders.

# Views: Expression Trajectory

## Description:

This view depicts the difference in expression between SOD1-WT and SOD1-G93A, within each anatomical annotation area, at each time point. Positive values represent higher expression in SOD1-G93A, while negative values represent higher expression in SOD1-WT. Only genes that pass the user defined statistical cutoffs at the selected time point are shown. This view is useful for understanding the temporal dynamics of differentially expressed genes.

# Views: Expression Trajectory

## Adjusting cutoffs

The gene set plotted can be filtered along several parameters by selecting from the drop down menus at the top of the window. Only genes that pass the selected cutoffs in each anatomical annotation area will be plotted. The “Timepoint” drop down selects at which time point differential expression is to be considered. The “Minimum Bayes factor” dropdown selects how stringent the test for differential expression should be. Higher values represent more confidence in differential expression. A minimal expression value can also be selected using the “Minimum expression” slider. Generally, genes with a minimum expression level of  $\geq -3$ , and a Bayes factor of  $\geq 5$  are true positives.

