

## Supplementary Figure Captions

**Figure S1:** Example scenarios for gene expression patterns with and without noise.

**Figure S2:** Clustering-based tree of cell types for plant root and mouse neuronal types. Heatmap clustergrams were generated in mouse and Arabidopsis by first filtering for significantly regulated genes using a one-way ANOVA on cell types with replicates ( $P < 0.0001$ ). This yielded 11,465 genes for Arabidopsis and 3,754 genes for mouse. For hierarchical clustering, the distance metric was Pearson correlation used with an average linkage method.

**Figure S3:** Histograms of maximally informative cellular domain sizes, for each cell type. The genes displaying maximally informative domains, identified as bright fingers in Figure 2, are histogrammed. For each gene belonging a maximally informative domain of size  $D$ , the  $D$  cell types for which it is maximally informative were identified. For each cell type, the histogram displays the number of genes whose maximally informative domains overlap the given type, at each value of  $D$ .

**Figure S4:** Genomic structure of hormone treatment specificity. This plot is identical in nature to Figure 2, but is shown for the hormone dataset.

**Figure S5:** Comparison of the bin-based *Spec* estimator with its true value in the Gaussian model (see Supplementary Text). (a) The plot shows a histogram of *Spec*( $y$ ) values computed in the plant dataset, using both the “true” Gaussian distribution, and the bin-based estimator, with  $n_{\text{bin}} = 3$ . Histogram is over all genes and all cell types. (b) The correlation value of true *Spec* versus the bin-based *Spec* estimator, as a function of the number of bins,  $n_{\text{bin}}$ .

## Supplementary Table Captions

**Table S1:** Results of Gene Ontology (GO) analysis of genes with high *Spec* values in both QC and lateral root meristem cell types.

**Table S2:** Results of GO analysis of genes with different cellular domain sizes (Plant).

**Table S3:** Results of GO analysis of genes with different cellular domain sizes (Mouse).

**Table S4:** List of known auxin-responsive genes.

**Table S5:** Top 20 ranked markers of auxin according to *Spec* and GenePattern, with known auxin markers highlighted.

**Table S6:** Raw gene expression data (Plant).

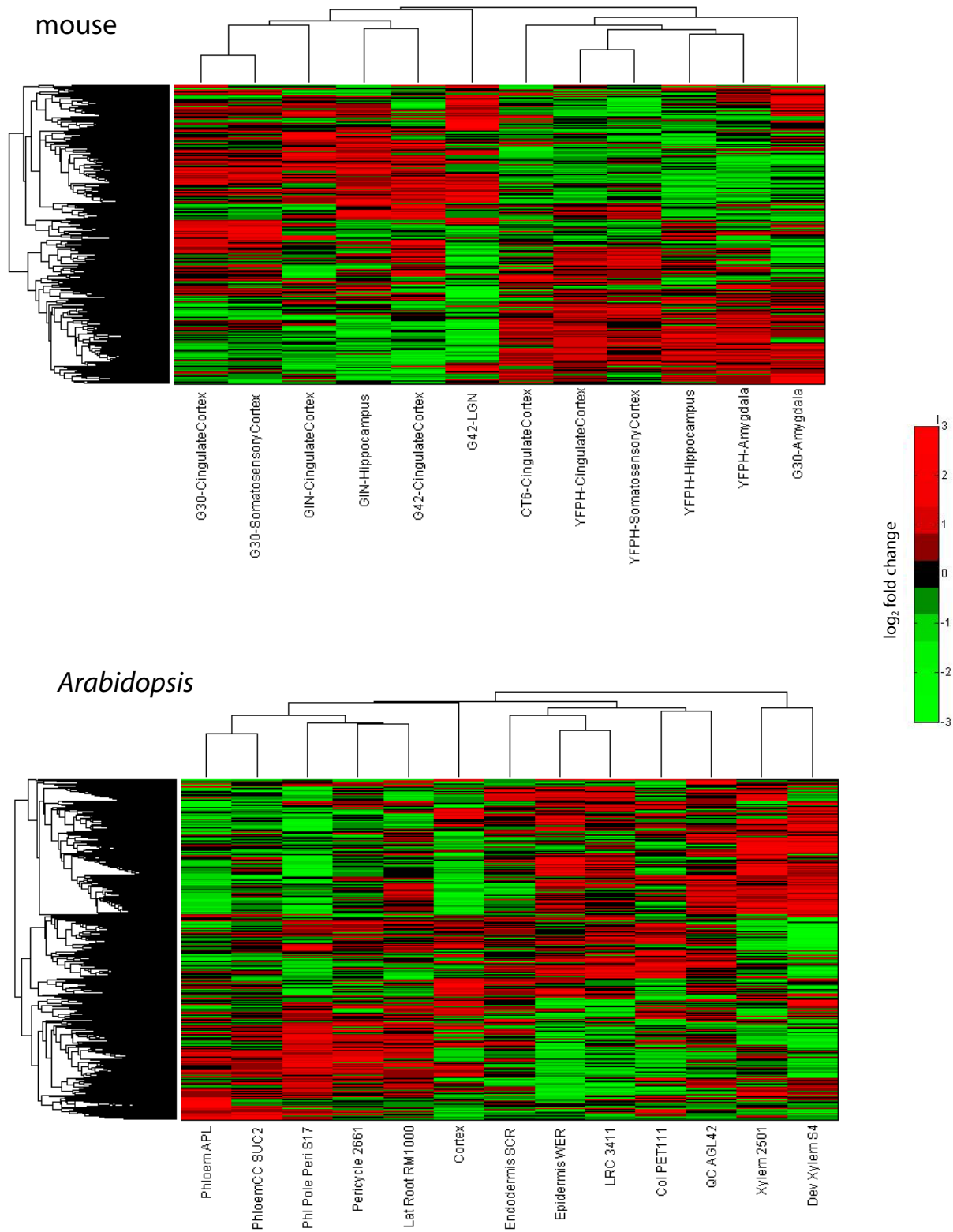
**Table S7:** Raw gene expression data (Mouse).

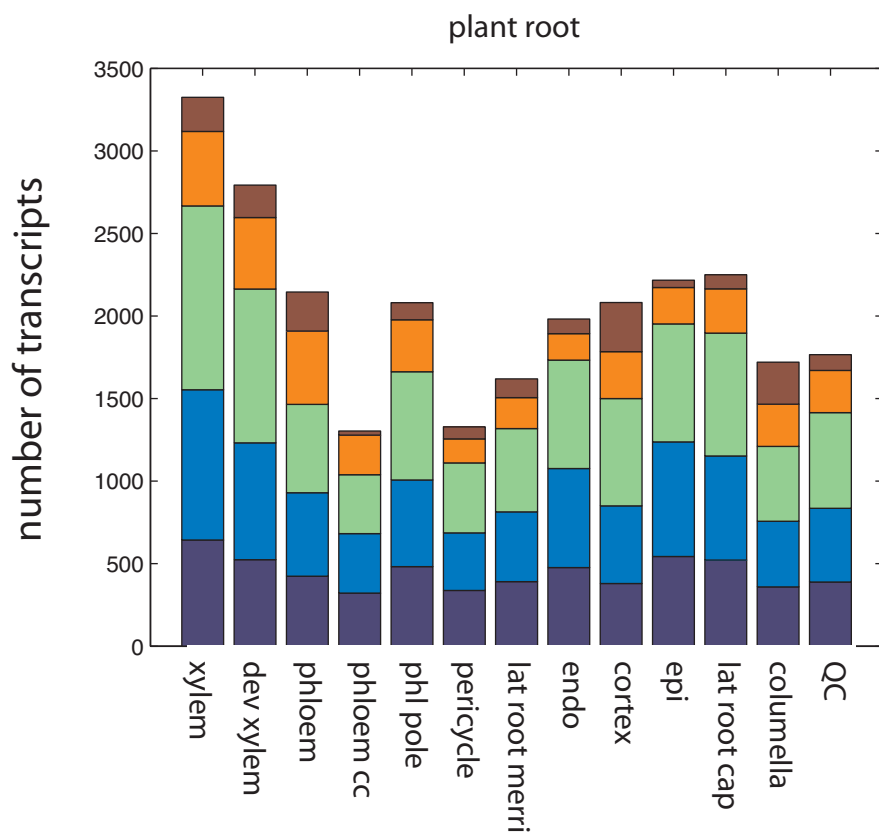
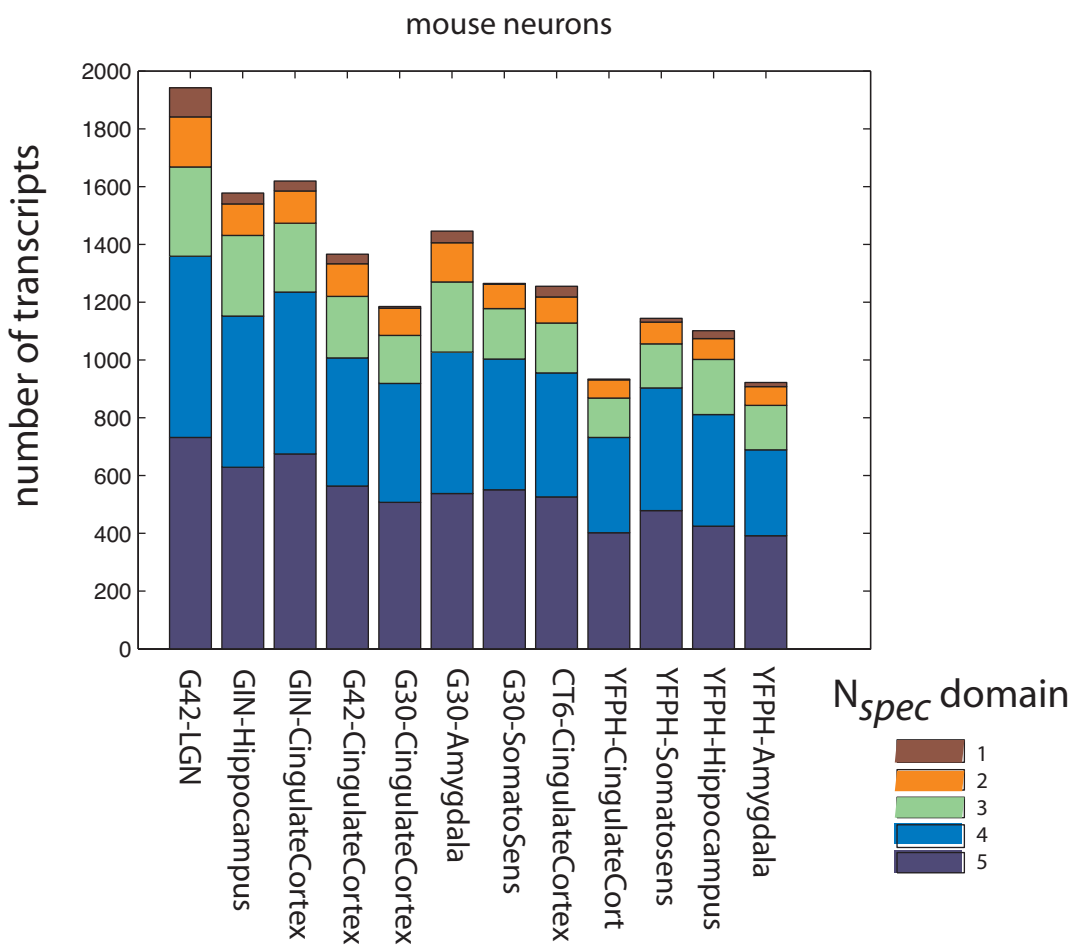
**Table S8:** Raw gene expression data for plant hormone treatments.

**Table S9:** *Spec* and *dSpec* output from analysis of hormone treatment data.

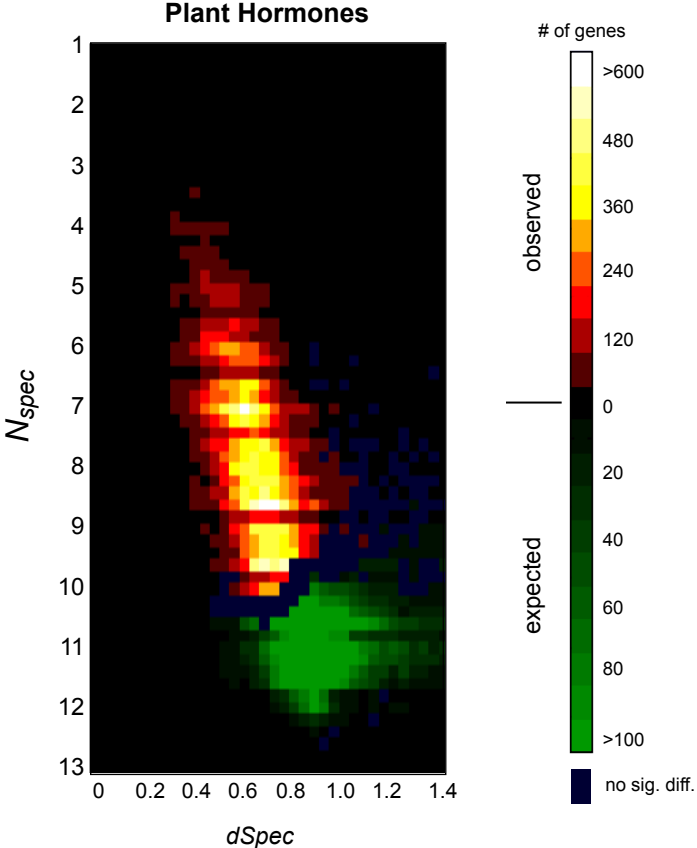
**Table S10:** GenePattern output from analysis of hormone treatment data.



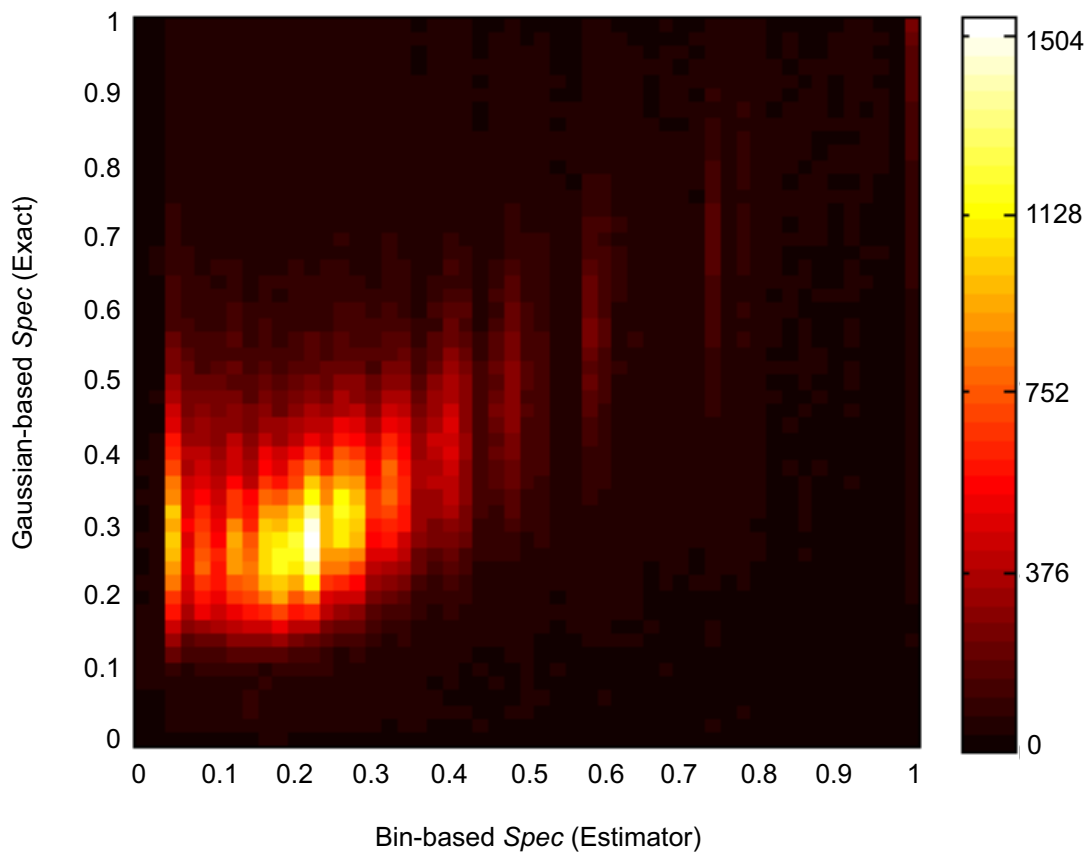




SUPPLEMENTARY FIGURE 4



**A**



**B**

