**Supplementary Figure 1**. Schematic of gating strategies for isolate brain cell types. Panels A,B,C are strategies used by the Complex Disease Epigenetic Group at Exeter, with panel D represents the strategy used to isolate cell types in the EPIGABA project. The blue boxes state the combination of antibodies used and the white boxes give the labels used to refer to that sample type in the manuscript.
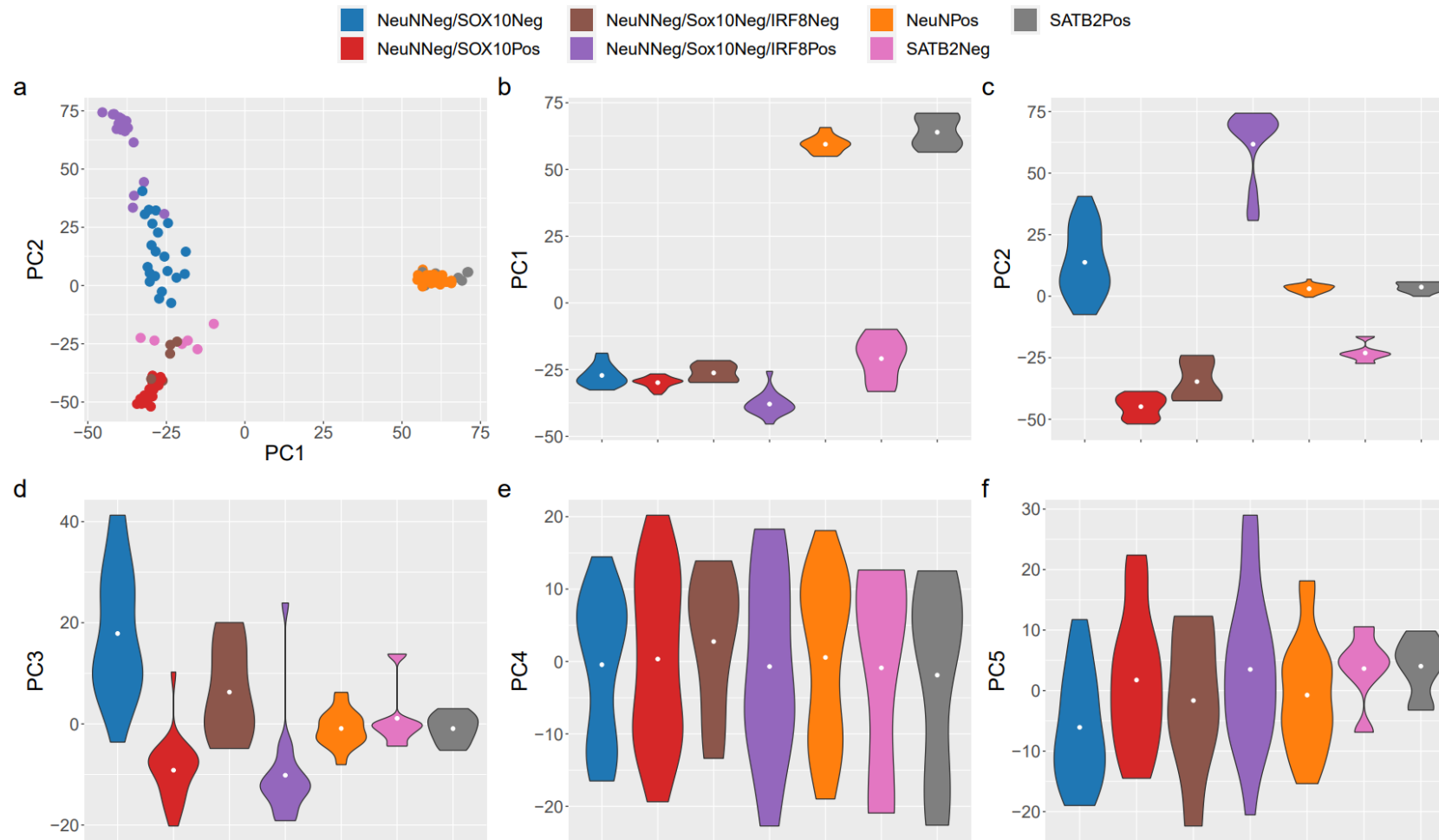
Supplementary Figure 2. Major axes of variation in DNA methylation data are driven by cell type. Scatterplot (a) of the first two principal components (first, second and sixth) that optimally separate different sample types (n = 112 samples from 43 donors). Each point represents a sample and the colour of the point indicates the nuclei fraction. Violin plots (b-f) for the first 5 principal components grouped by nuclei fraction.

Supplementary Figure 3. Schematic of reference panels form computational deconvolution of cellular composition in brain tissue. Each row represents a separate reference panel with a different combination of cell types for which deconvolution was performed.
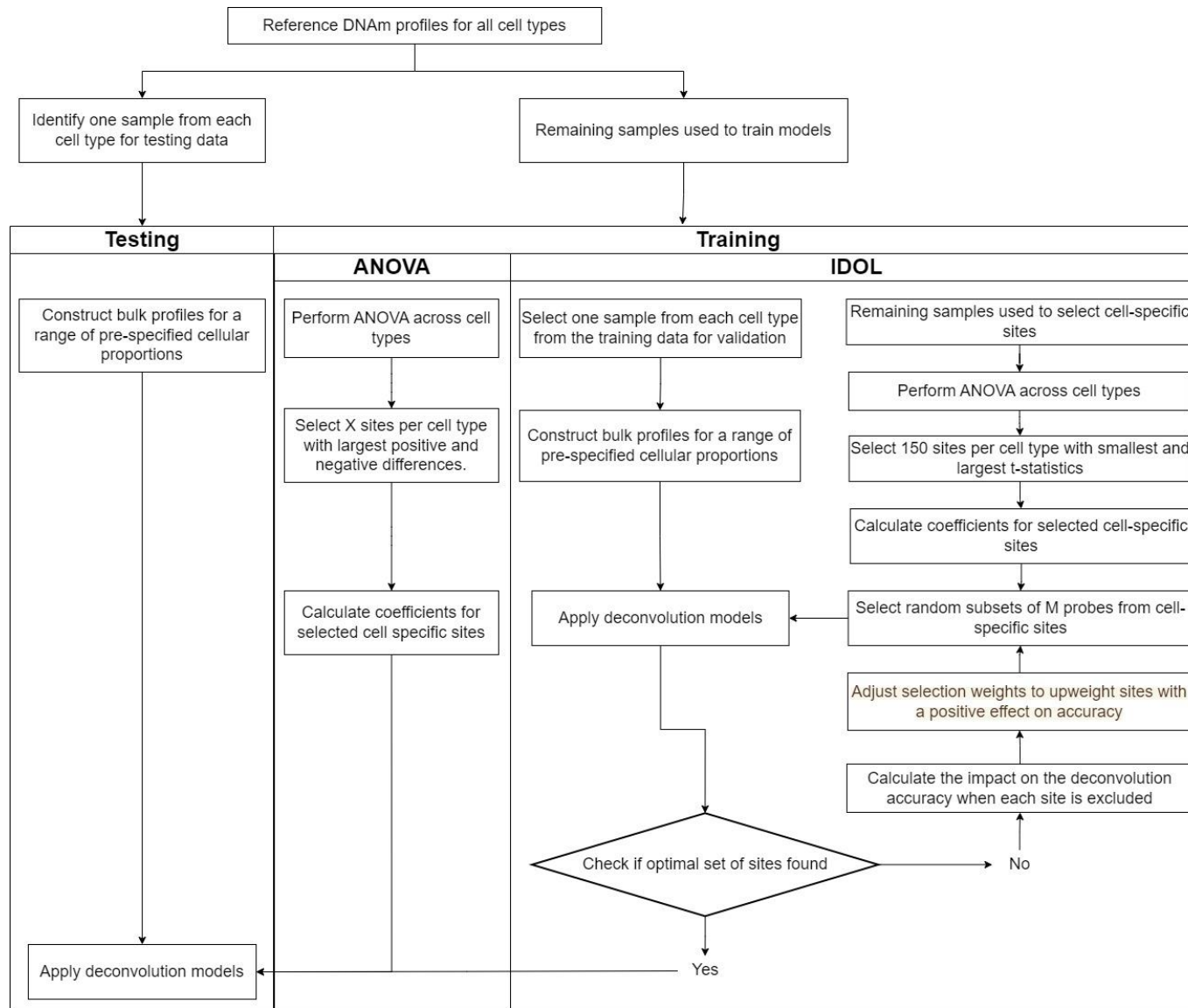
Supplementary Figure 4. Flowchart of methodology for simulation study to profile performance of different panels of brain cell types.
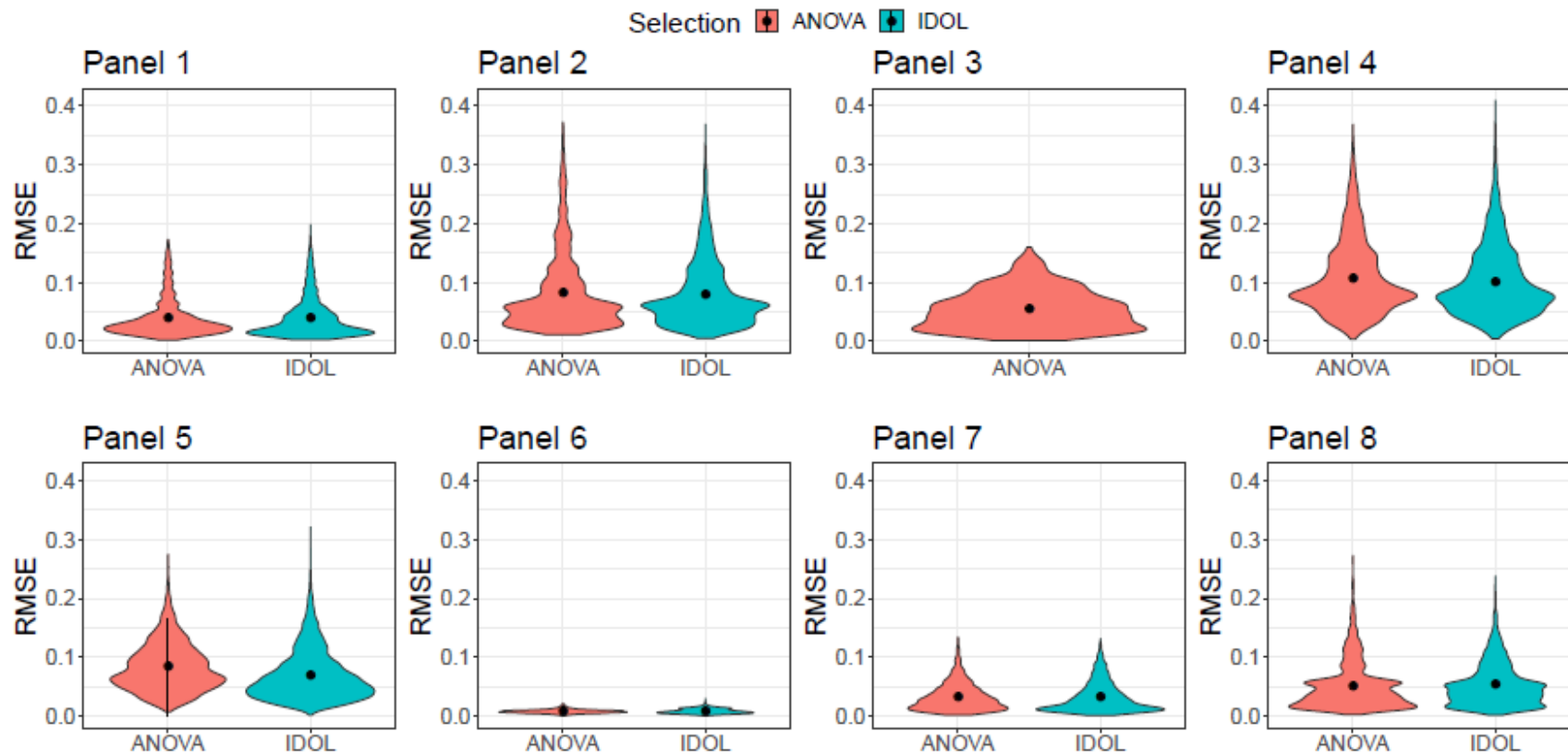
Supplementary Figure 5. Violin plots of the root mean square error (RMSE) between the fixed cellular proportions used to construct the bulk brain profiles (n = 360-1260) and the estimated proportions. Each figure panel represents a different panel of cell types (**Table 1/Additional File 1: Supplementary Figure 3**).
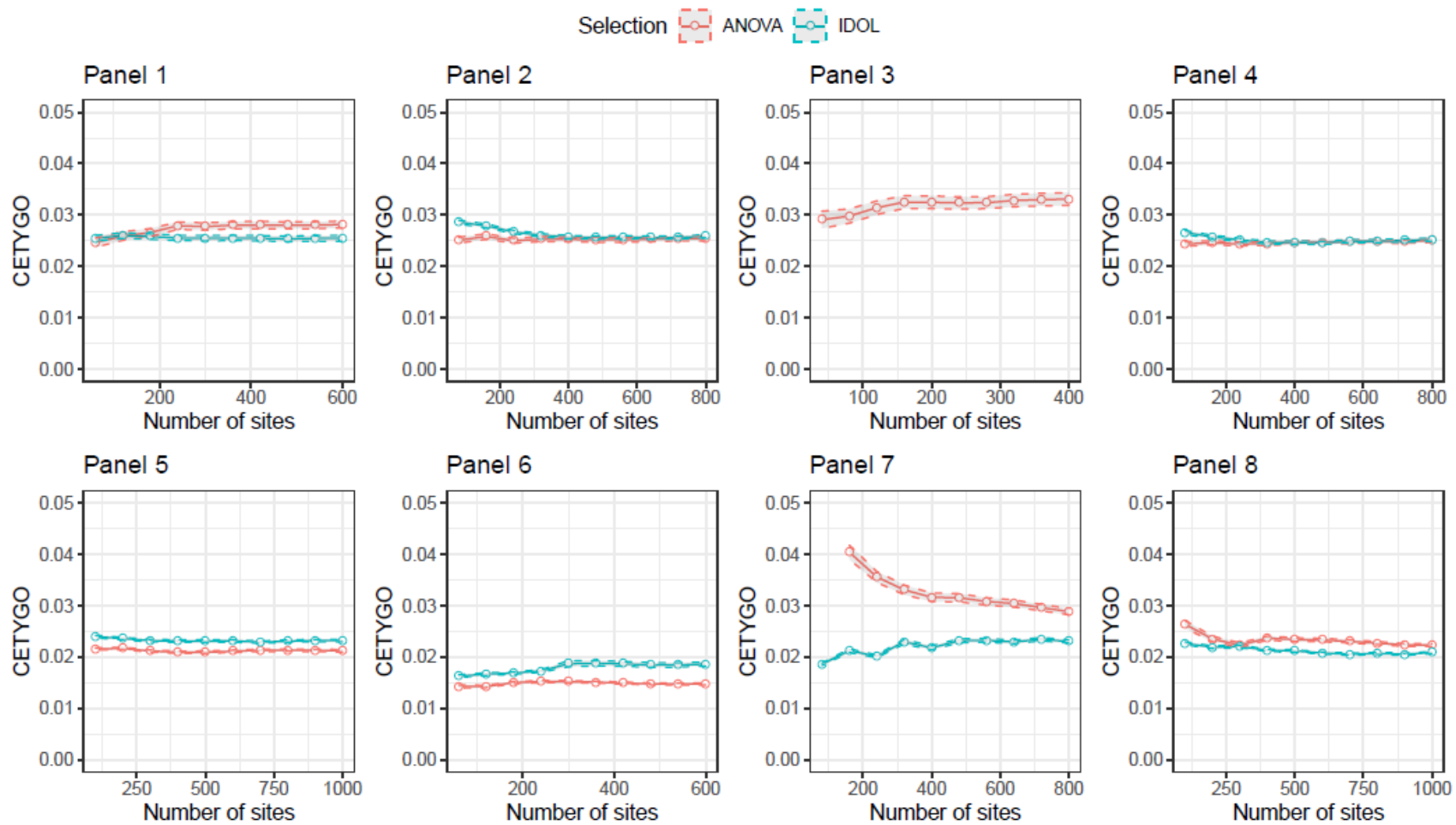
Supplementary Figure 6. Line graph of the effect of number of cell-specific probes on the accuracy of cellular deconvolution on reconstructed bulk brain profiles (n = 360-1260) measured by the CETYGO score. Each figure panel represents a different panel of cell types (**Table 1/Additional File 1: Supplementary Figure 3**).
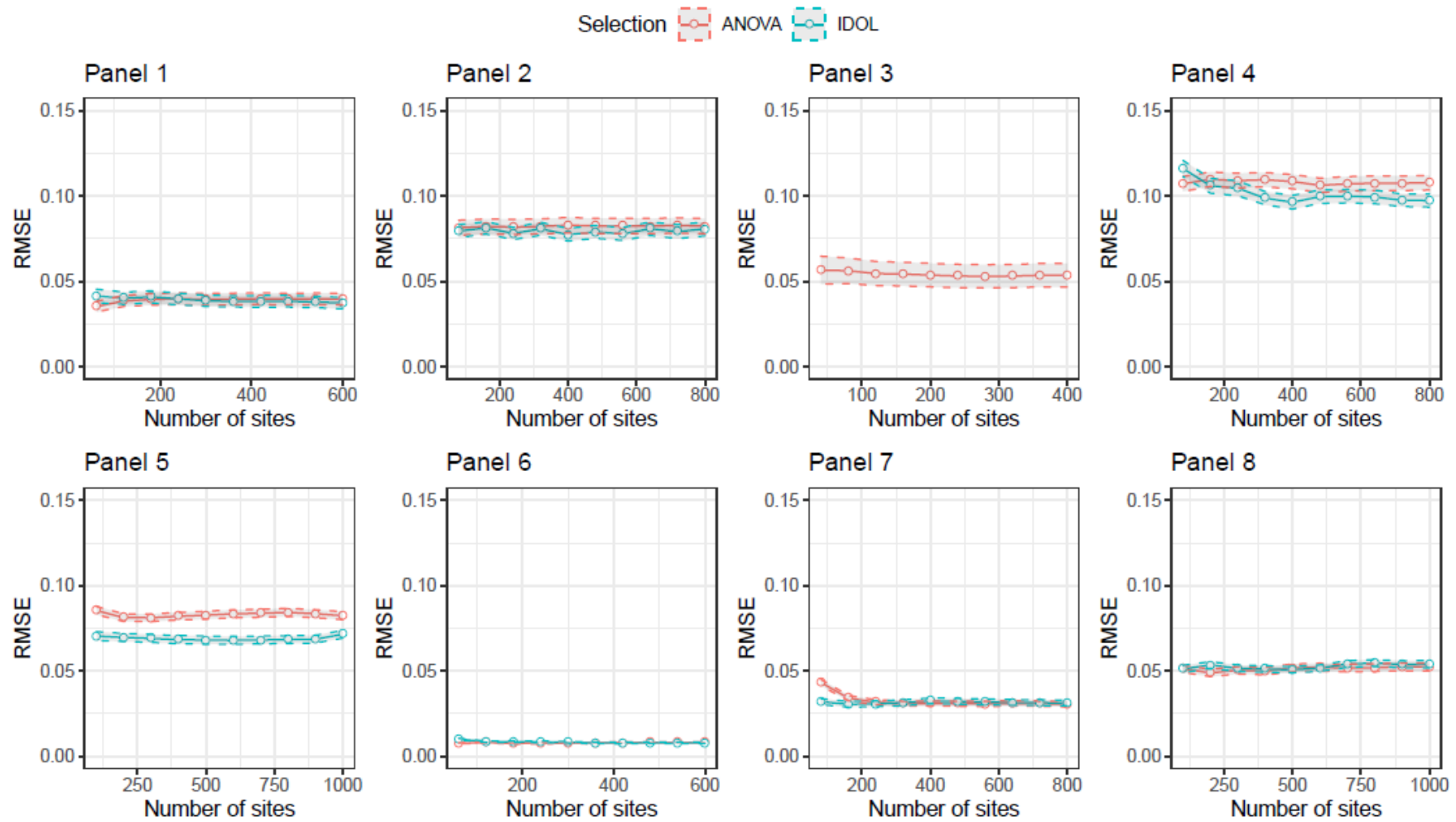
Supplementary Figure 7. Line graph of the effect of number of cell-specific probes on the accuracy of cellular deconvolution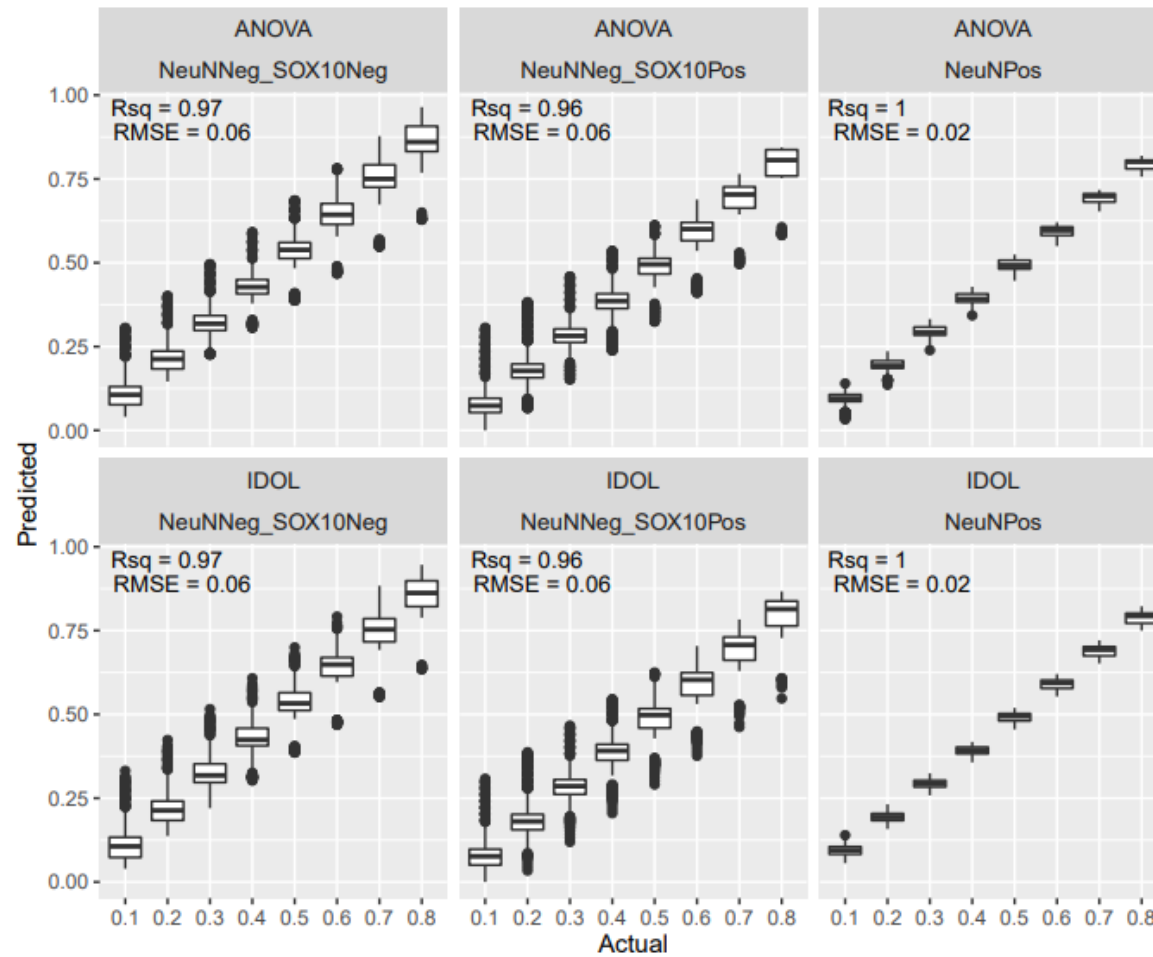 measured by the root mean square error (RMSE) between the fixed cellular proportions used to construct the bulk brain profiles (n = 360-1260) and the estimated proportions. Each figure panel represents a different panel of cell types (**Table 1/Additional File 1: Supplementary Figure 3**).
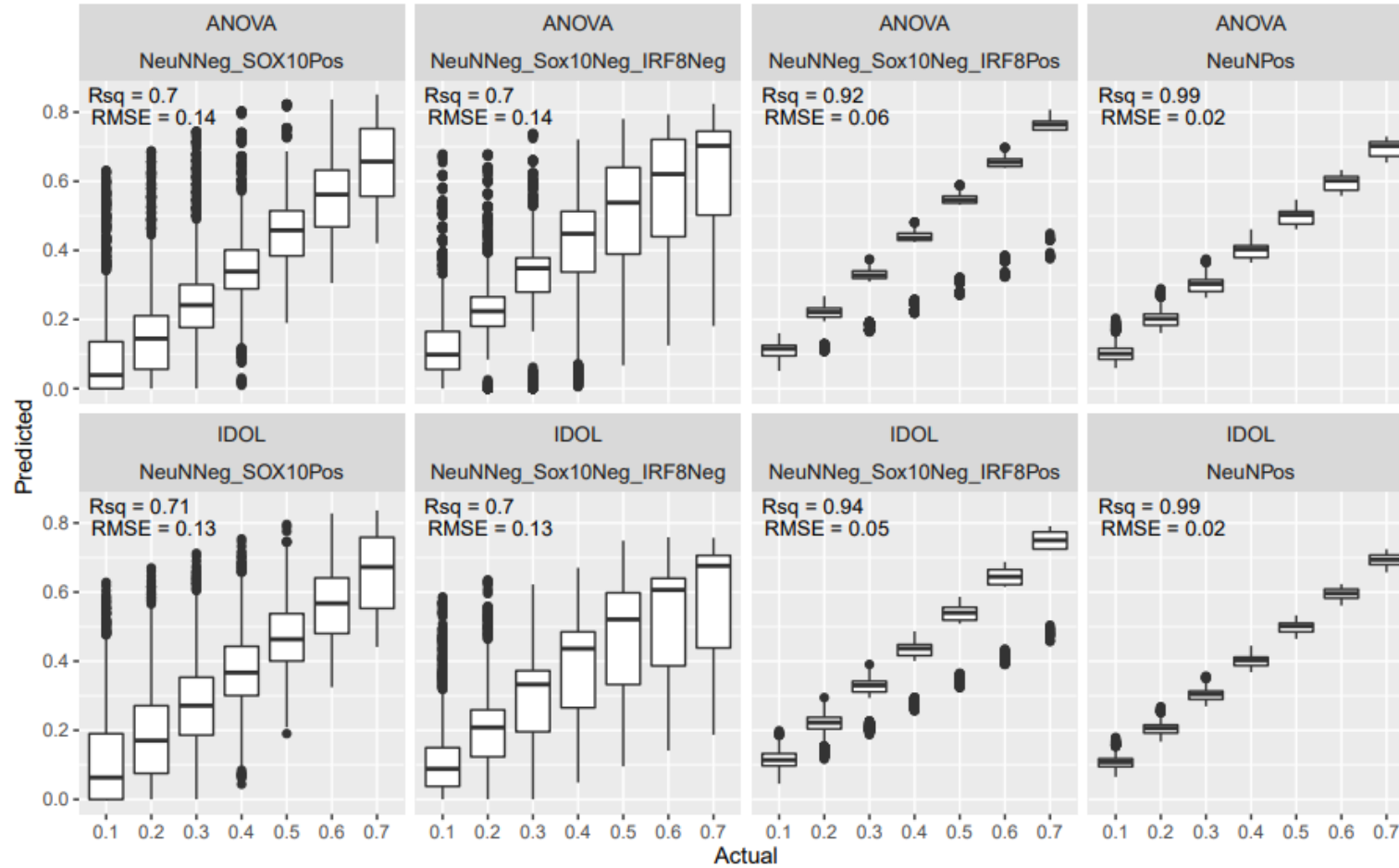
Supplementary Figure 8. Boxplots of distribution of predicted proportion (y-axis), grouped by actual proportion specified in the reconstruction of brain profiles (x-axis; n = 360). Shown here are the results for panel 1 of cell types (**Table 1**/**Additional File 1: Supplementary Figure 3**) using both the ANOVA and IDOL methods to select cell specific sites. Each panel in the figure represents a different cell type.

Supplementary Figure 9. Boxplots of distribution of predicted proportion (y-axis), grouped by actual proportion specified in the reconstruction of brain profiles (x-axis; n = 840). Shown here are the results for panel 2 of cell types (**Table 1**/**Additional File 1: Supplementary Figure 3**) using both the ANOVA and IDOL methods to select cell specific sites. Each panel in the figure represents a different cell type.
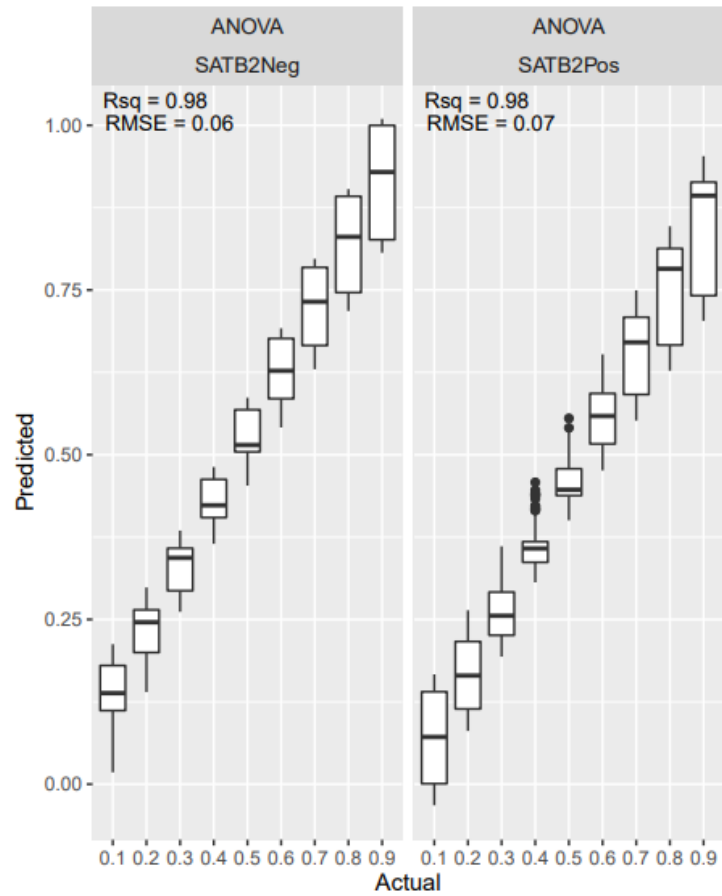
Supplementary Figure 10. Boxplots of distribution of predicted proportion (y-axis), grouped by actual proportion specified in the reconstruction of brain profiles (x-axis; n = 90). Shown here are the results for panel 3 of cell types (**Table 1/Additional File 1: Supplementary Figure 3**) using both the ANOVA method to select cell specific sites. Each panel in the figure represents a different cell type.
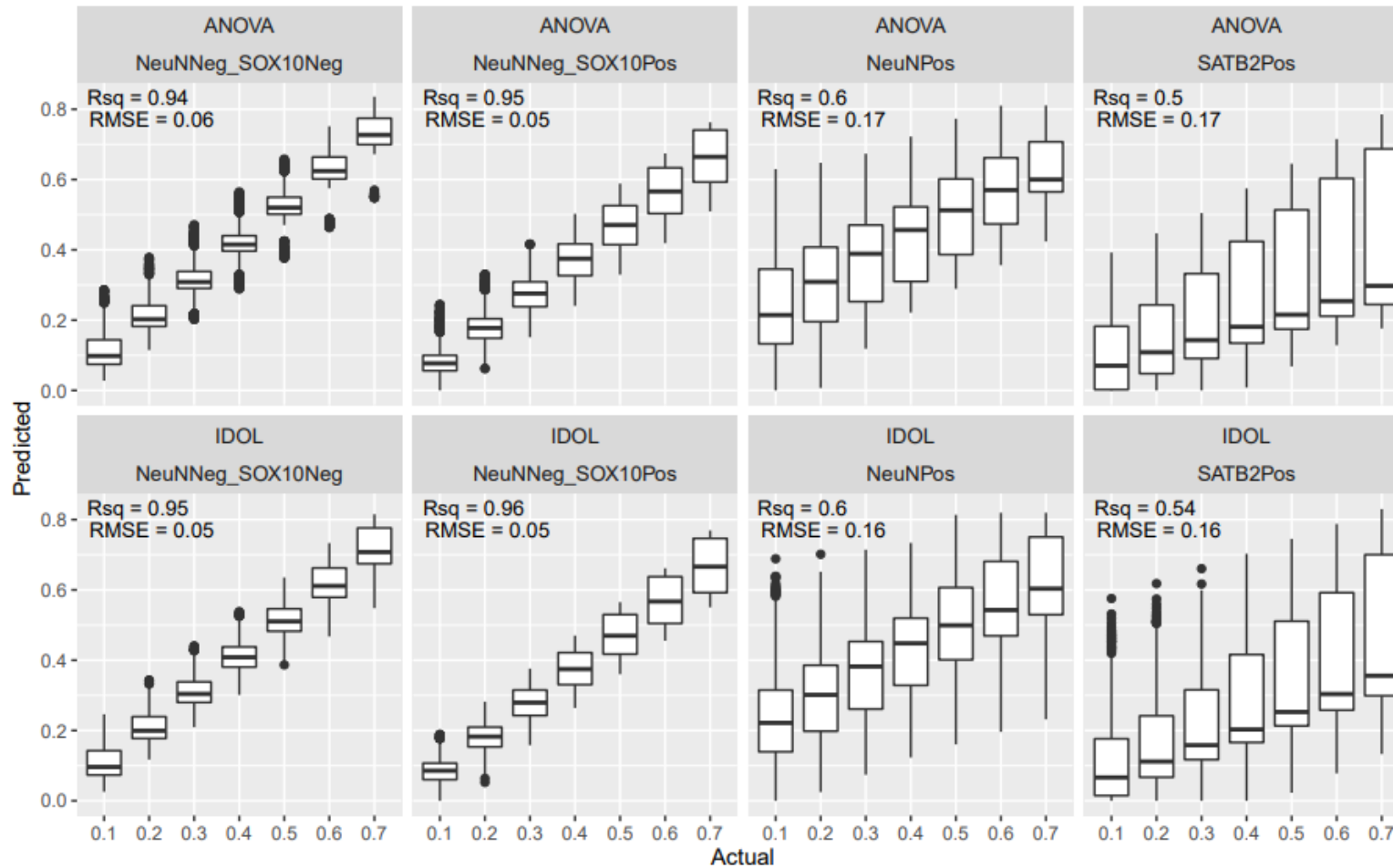
Supplementary Figure 11. Boxplots of distribution of predicted proportion (y-axis), grouped by actual proportion specified in the reconstruction of brain profiles (x-axis; n = 840). Shown here are the results for panel 4 of cell types (**Table 1/Additional File 1: Supplementary Figure 3**) using both the ANOVA and IDOL methods to select cell specific sites. Each panel in the figure represents a different cell type.
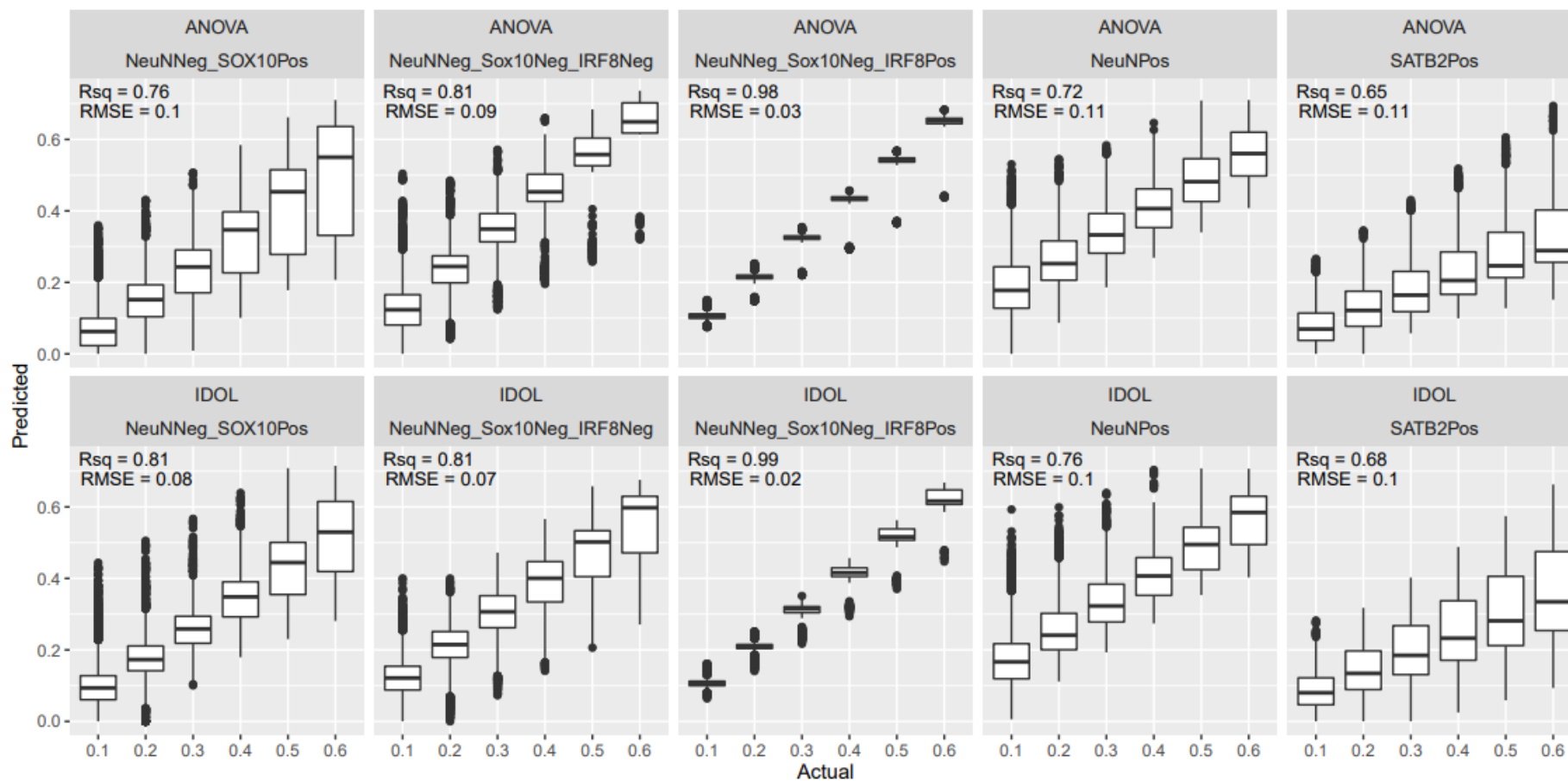
Supplementary Figure 12. Boxplots of distribution of predicted proportion (y-axis), grouped by actual proportion specified in the reconstruction of brain profiles (x-axis; n = 1260). Shown here are the results for panel 5 of cell types (**Table 1**/**Additional File 1: Supplementary Figure 3**) using both the ANOVA and IDOL methods to select cell specific sites. Each panel in the figure represents a different cell type.
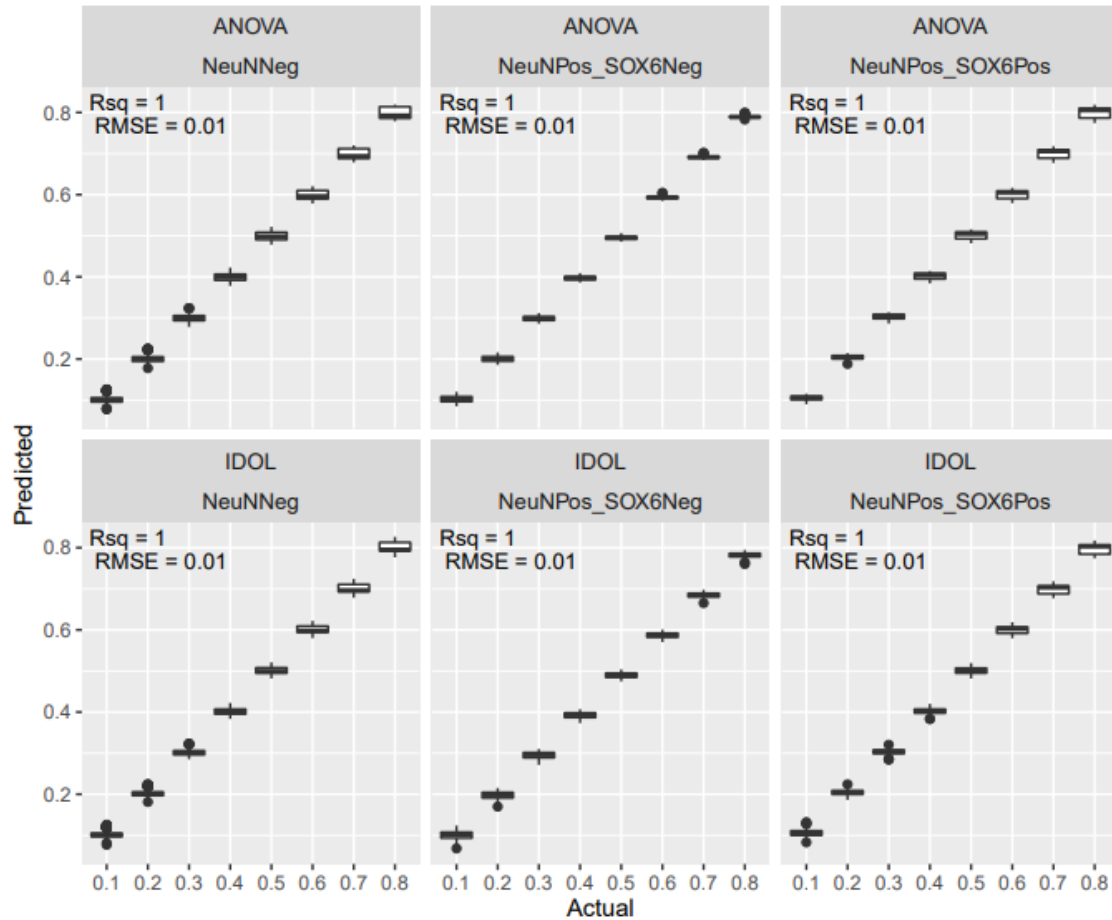
Supplementary Figure 13. Boxplots of distribution of predicted proportion (y-axis), grouped by actual proportion specified in the reconstruction of brain profiles (x-axis; n = 360). Shown here are the results for panel 6 of cell types (**Table 1/Additional File 1: Supplementary Figure 3**) using both the ANOVA and IDOL methods to select cell specific sites. Each panel in the figure represents a different cell type.
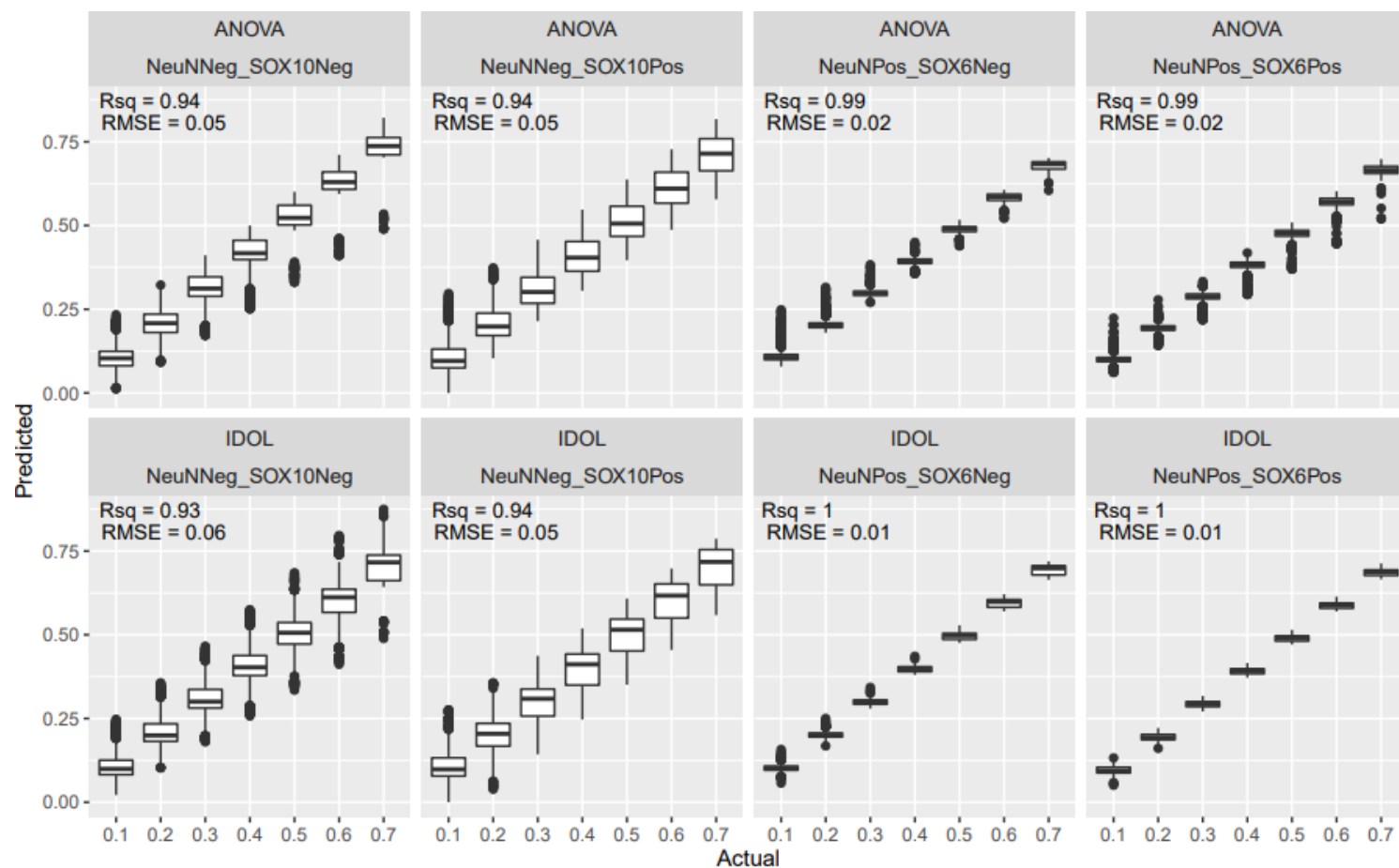
Supplementary Figure 14. Boxplots of distribution of predicted proportion (y-axis), grouped by actual proportion specified in the reconstruction of brain profiles (x-axis; n = 360). Shown here are the results for panel 7 of cell types (**Table 1/Additional File 1: Supplementary Figure 3**) using both the ANOVA and IDOL methods to select cell specific sites. Each panel in the figure represents a different cell type.

Supplementary Figure 15. Boxplots of distribution of predicted proportion (y-axis), grouped by actual proportion specified in the reconstruction of brain profiles (x-axis; n = 1260). Shown here are the results for panel 8 of cell types (**Table 1/Additional File 1: Supplementary Figure 3**) using both the ANOVA and IDOL methods to select cell specific sites. Each panel in the figure represents a different cell type.
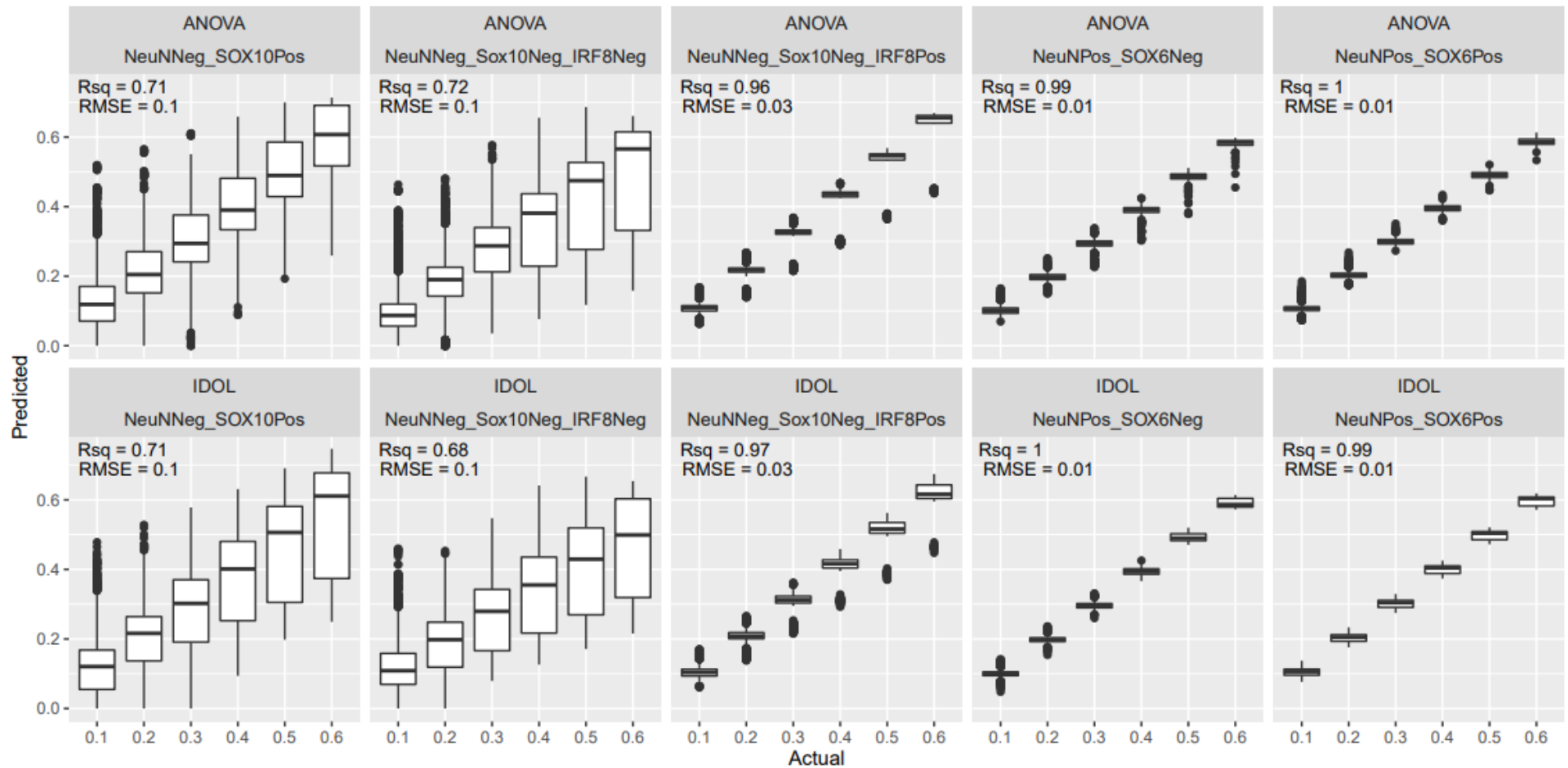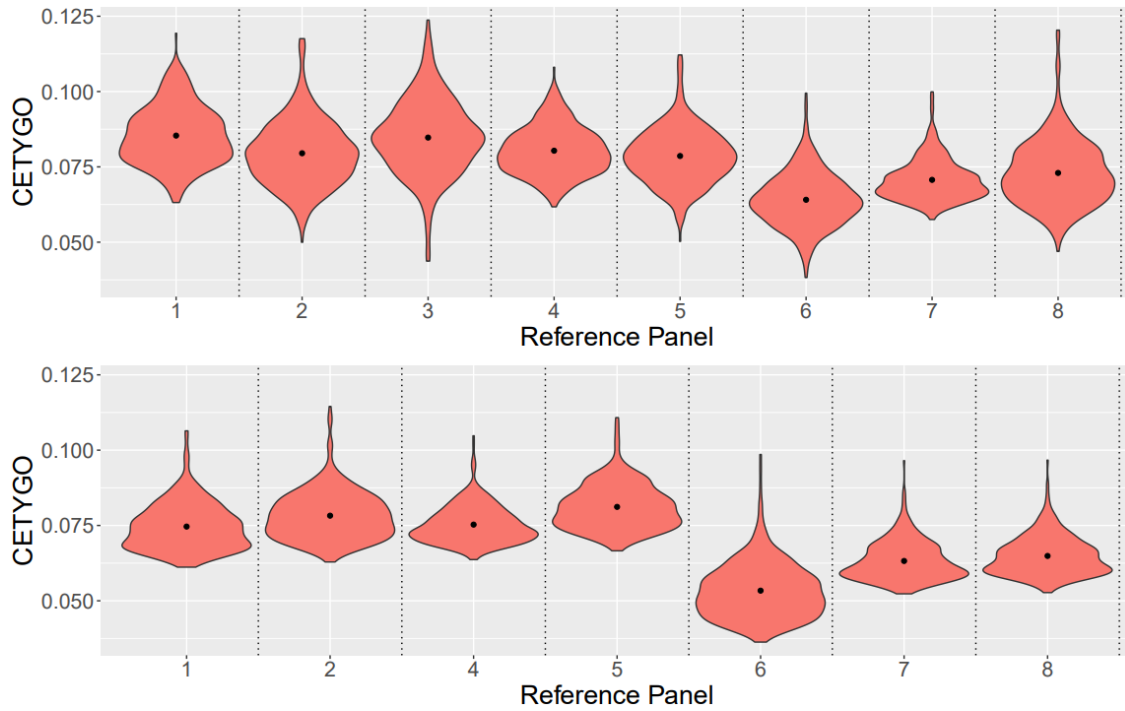
Supplementary Figure 16A. Violin plot of CETYGO score across 377 adult prefrontal cortex samples. Top panel, probes cell specific sites were selected with an ANOVA, bottom panel, cell specific sites were selected with the IDOL method.
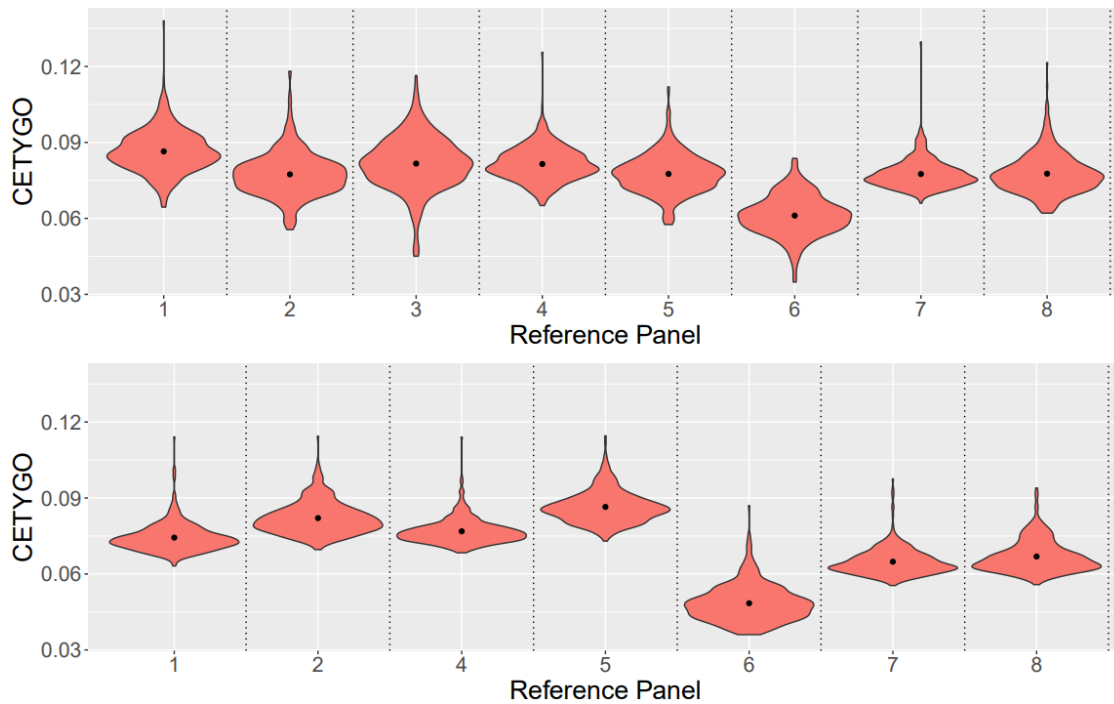
Supplementary Figure 16B. Violin plot of CETYGO score across 415 adult prefrontal cortex samples from the Jaffe dataset. Top panel, probes cell specific sites were selected with an ANOVA, bottom panel, cell specific sites were selected with the IDOL method.

Top heatmap (Exeter dataset):

|	ANOVA_M7	ANOVA_M1	ANOVA_M4	IDOL_M7	IDOL_M8	IDOL_M1	IDOL_M4	ANOVA_M3	ANOVA_M6	IDOL_M6	IDOL_M2	IDOL_M5	ANOVA_M8	ANOVA_M2	ANOVA_M5
ANOVA_M7															
ANOVA_M1	0.92														
ANOVA_M4	0.91	0.98													
IDOL_M7	0.85	0.82	0.82												
IDOL_M8	0.90	0.87	0.86	0.90											
IDOL_M1	0.91	0.89	0.88	0.91	0.92										
IDOL_M4	0.91	0.88	0.87	0.90	0.91	0.96									
ANOVA_M3	0.76	0.87	0.87	0.66	0.77	0.77	0.75								
ANOVA_M6	0.80	0.83	0.84	0.74	0.83	0.80	0.77	0.89							
IDOL_M6	0.79	0.75	0.77	0.77	0.87	0.85	0.82	0.77	0.84						
IDOL_M2	0.90	0.87	0.86	0.83	0.91	0.91	0.89	0.82	0.82	0.89					
IDOL_M5	0.89	0.89	0.89	0.84	0.91	0.92	0.91	0.85	0.85	0.89	0.96				
ANOVA_M8	0.90	0.86	0.86	0.78	0.89	0.89	0.86	0.86	0.89	0.89	0.95	0.94			
ANOVA_M2	0.90	0.91	0.91	0.78	0.89	0.89	0.86	0.91	0.91	0.87	0.93	0.94	0.98		
ANOVA_M5	0.86	0.90	0.91	0.74	0.85	0.86	0.83	0.93	0.91	0.86	0.91	0.93	0.96	0.99	

Bottom heatmap (Jaffe dataset):

|	ANOVA_M3	ANOVA_M6	ANOVA_M1	ANOVA_M4	ANOVA_M8	ANOVA_M2	ANOVA_M5	IDOL_M7	IDOL_M8	IDOL_M6	ANOVA_M7	IDOL_M2	IDOL_M5	IDOL_M1	IDOL_M4
ANOVA_M3															
ANOVA_M6	0.88														
ANOVA_M1	0.88	0.84													
ANOVA_M4	0.86	0.84	0.98												
ANOVA_M8	0.87	0.86	0.90	0.88											
ANOVA_M2	0.92	0.87	0.92	0.91	0.96										
ANOVA_M5	0.93	0.87	0.91	0.91	0.94	0.99									
IDOL_M7	0.40	0.58	0.64	0.63	0.60	0.52	0.48								
IDOL_M8	0.55	0.69	0.71	0.70	0.74	0.65	0.61	0.90							
IDOL_M6	0.64	0.79	0.73	0.73	0.80	0.75	0.73	0.72	0.82						
ANOVA_M7	0.70	0.73	0.88	0.88	0.87	0.82	0.78	0.75	0.81	0.75					
IDOL_M2	0.66	0.73	0.80	0.79	0.85	0.77	0.74	0.77	0.89	0.83	0.83				
IDOL_M5	0.73	0.76	0.85	0.83	0.89	0.83	0.80	0.74	0.86	0.84	0.84	0.96			
IDOL_M1	0.60	0.69	0.82	0.82	0.79	0.74	0.70	0.85	0.84	0.81	0.86	0.87	0.88		
IDOL_M4	0.60	0.67	0.81	0.80	0.78	0.72	0.68	0.84	0.85	0.78	0.88	0.86	0.88	0.93	

Supplementary Figure 17. Correlation heatmap of CETYGO scores from 15 different deconvolution models applied to adult prefrontal samples. Each row/column denotes a different model where the first word (ANOVA or IDOL) indicates how the cell-specific sites were selected and the number after the "M" represent which panel of cell types (as defined in **Additional File 1: Supplementary Figure 3**). The top heatmap is based on the Exeter dataset (n = 377), and the bottom heatmap the Jaffe dataset (n = 415).
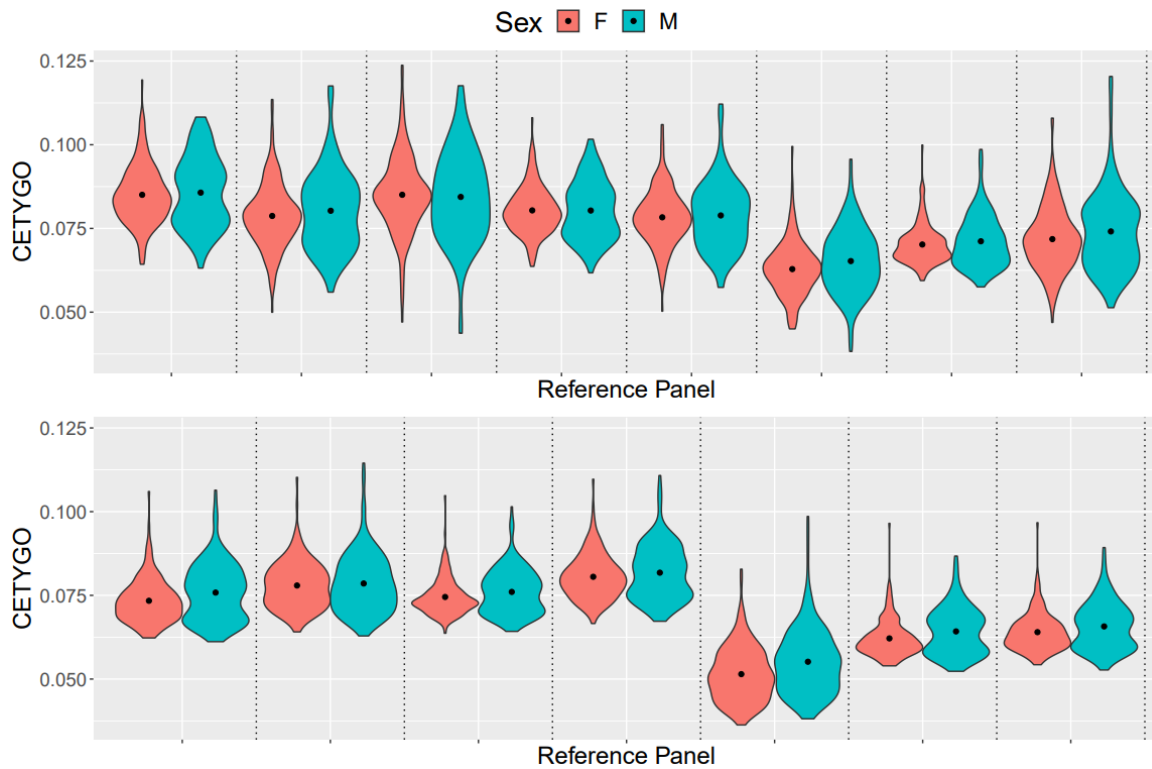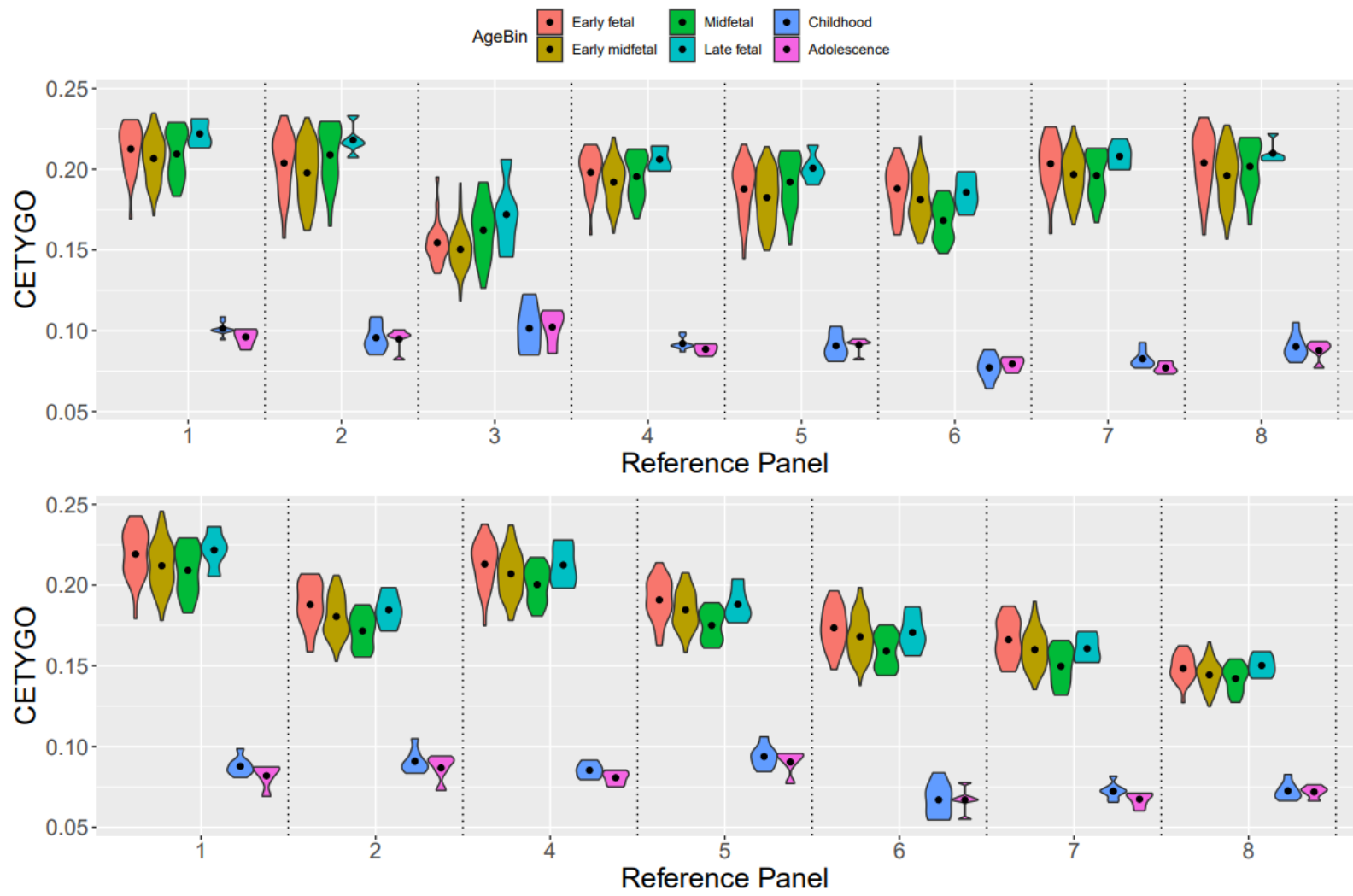
Supplementary Figure 18A. Violin plot of CETYGO score grouped by sex across 377 adult prefrontal cortex samples from the Exeter dataset. Top panel, probes cell specific sites were selected with an ANOVA, bottom panel, cell specific sites were selected with the IDOL method.
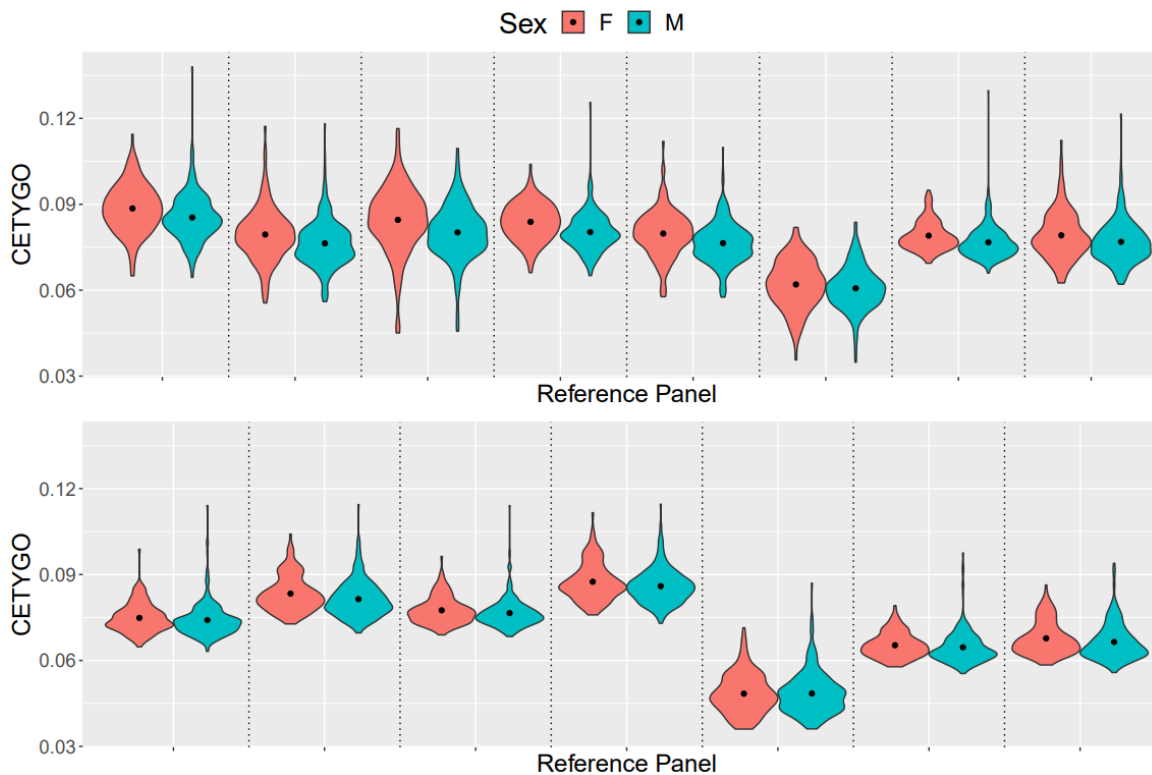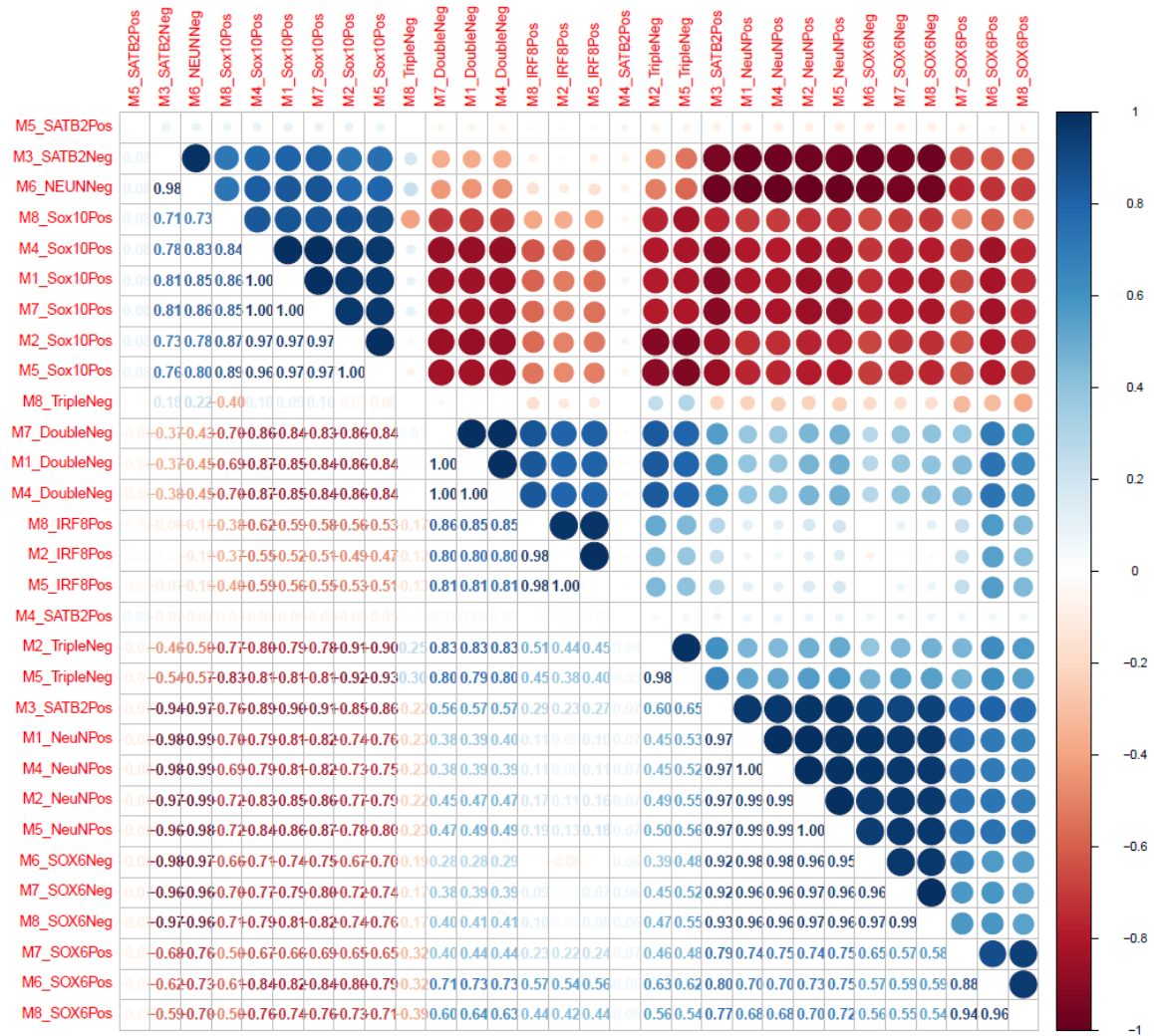

Supplementary Figure 18B. Violin plot of CETYGO score grouped by sex across 415 adult prefrontal cortex samples from the Jaffe dataset. Top panel, probes cell specific sites were selected with an ANOVA, bottom panel, cell specific sites were selected with the IDOL method.
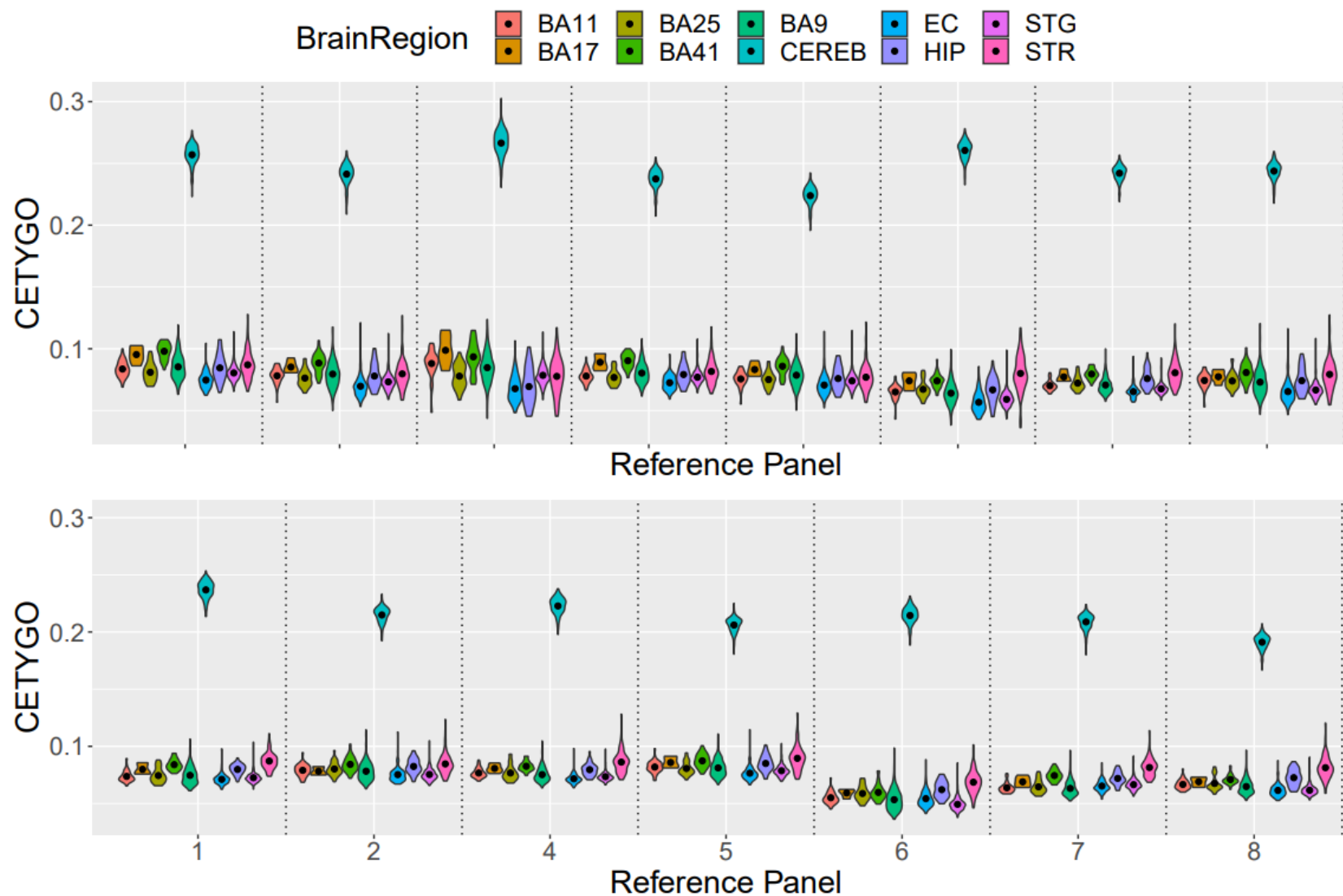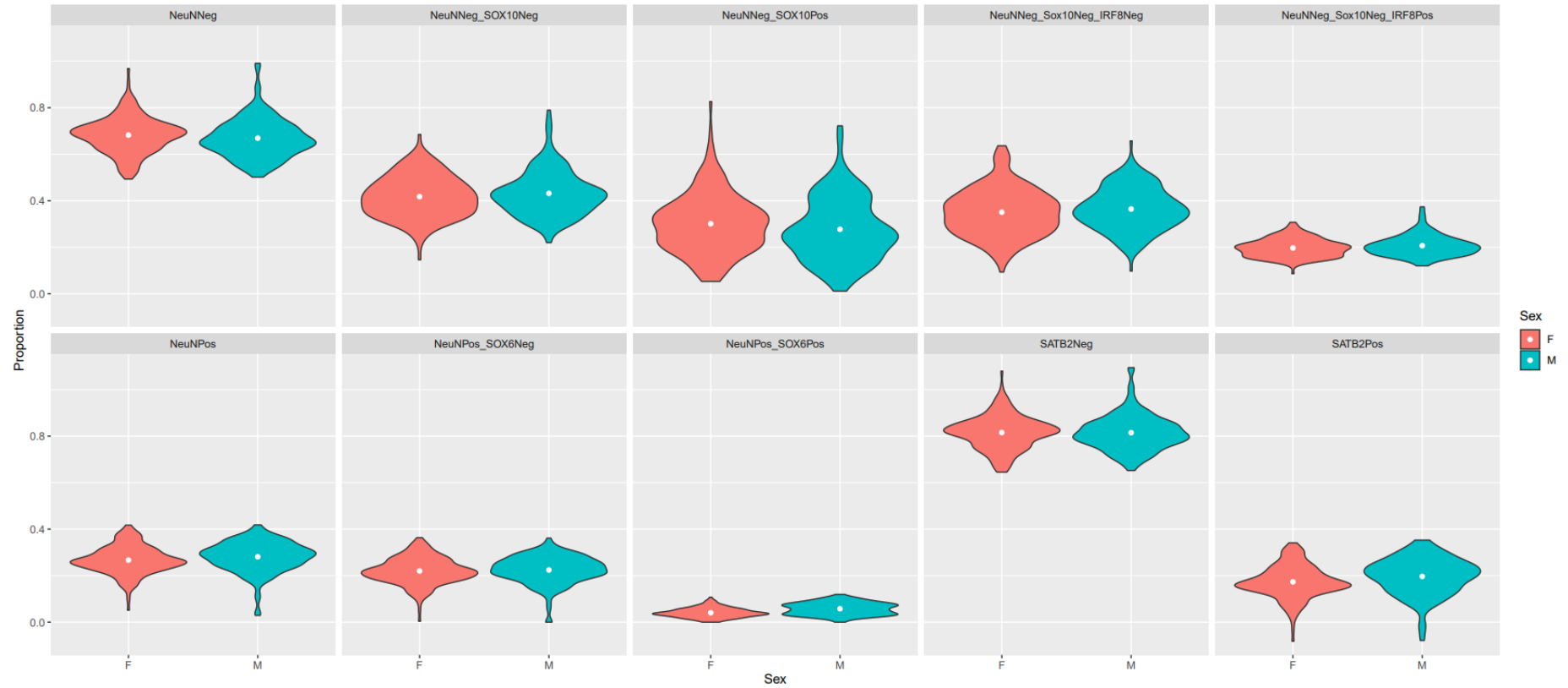
Supplementary Figure 19. Violin plot of CETYGO score across 851 adult tissue samples from different brain regions. Top panel, probes cell specific sites were selected with an ANOVA, bottom panel, cell specific sites were selected with the IDOL method.
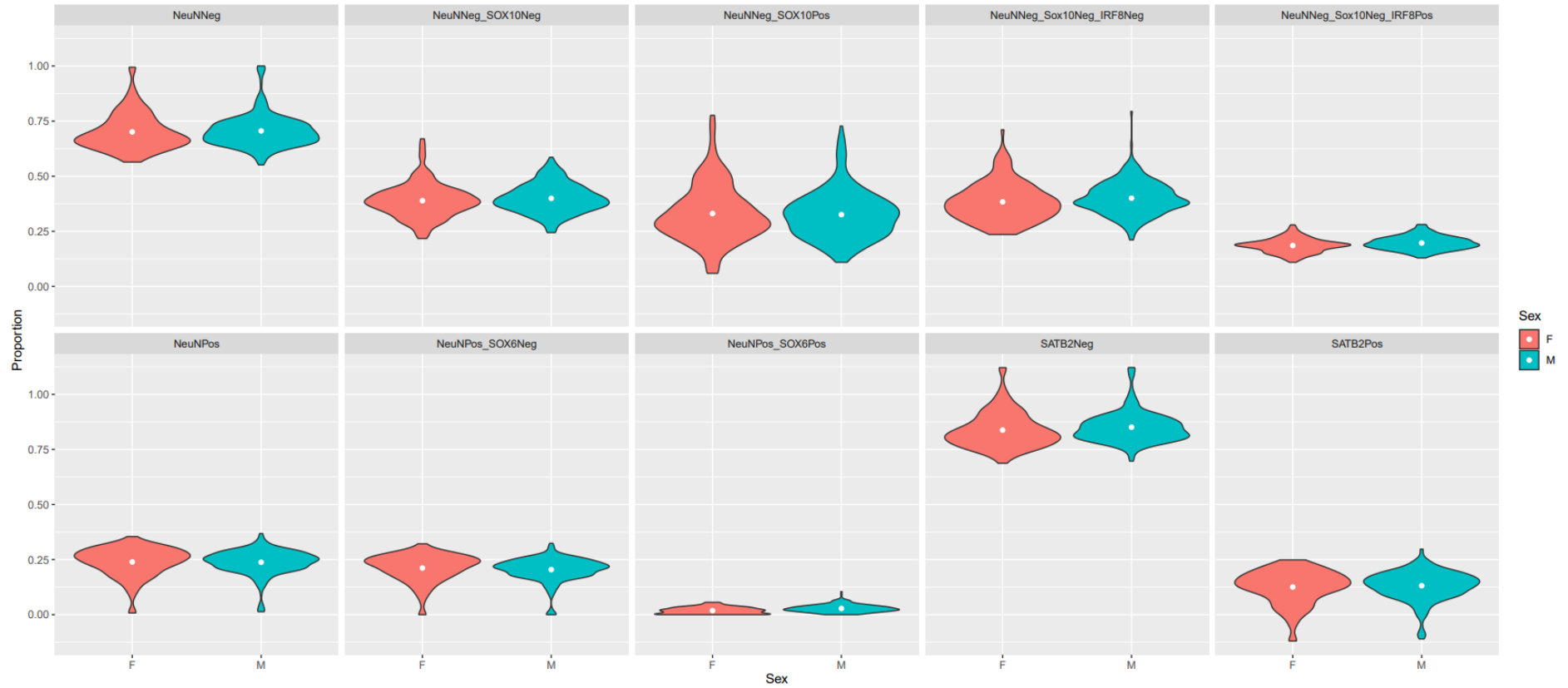
Supplementary Figure 20. Violin plot of CETYGO score across 167 fetal and childhood samples grouped into development stages. Top panel, probes cell specific sites were selected with an ANOVA, bottom panel, cell specific sites were selected with the IDOL method.

Supplementary Figure 21. Heatmap of correlations between predicted cellular proportions across all reference panels with ANOVA to select cell-specific sites in the Exeter dataset (n = 377).

Supplementary Figure 22. Violin plots of the distribution of proportion of brain cell types estimated from DNA methylation data generated from the prefrontal cortex grouped by sex across 377 adult prefrontal cortex samples from the Exeter dataset. Each panel represents a different cell type, estimated using the optimal reference panel for that cell type (**Additional File 2: Supplementary Table 8**).

Supplementary Figure 23. Violin plots of the distribution of proportion of brain cell types estimated from DNA methylation data generated from the prefrontal cortex grouped by sex across 415 adult prefrontal cortex samples from the Jaffe dataset. Each panel represents a different cell type, estimated using the optimal reference panel for that cell type (**Additional File 2: Supplementary Table 8**).
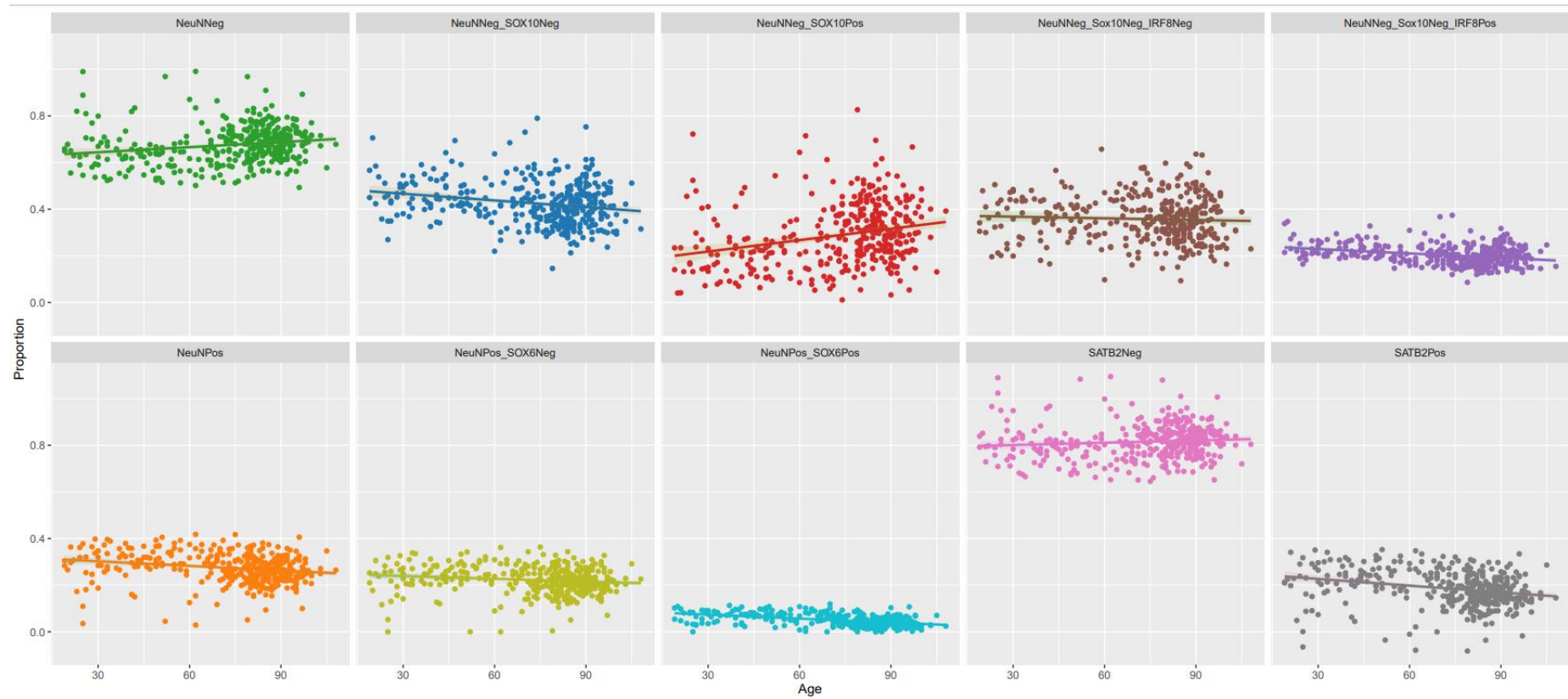
Supplementary Figure 24. Scatterplots plots of the of proportion of brain cell types estimated from DNA methylation data generated from the prefrontal cortex (y-axis) against age (years; x-axis) across 377 adult prefrontal cortex samples from the Exeter dataset. Each panel represents a different cell type, estimated using the optimal reference panel for that cell type (**Additional File 2: Supplementary Table 8**).

Supplementary Figure 25. Scatterplots plots of the of proportion of brain cell types estimated from DNA methylation data generated from the prefrontal cortex (y-axis) against age (years; x-axis) across 415 adult prefrontal cortex samples from the Jaffe dataset. Each panel represents a different cell type, estimated using the optimal reference panel for that cell type (**Additional File 2: Supplementary Table 8**).