Figure 4

This notebook reproduces the main analyses from figure 4 of the manuscript. First we import the necessary packages.

```
library(voxhunt)
library(tidyverse)
library(Seurat)
```

Now we load the data. The loaded seurat object contains the neuronal popultions of the datasets shown in the manuscript. We further subset the ones shown in figure 2.

```
load_aba_data('voxhunt_data/')
neurons <- read_rds('combined_neurons_srt.rds')
neurons <- subset(neurons, orig.ident%in%c('cerebral', 'hCS', 'hSS', 'tanaka_thalamus'))
neurons <- subset(neurons,
    cluster%in%c('mesen_ex_cerebral', 'mesen_in_cerebral', 'ctx_ex_cerebral',
        'ge_in_cerebral', 'dien_ex_cerebral', 'ge_hss', 'ctx_hcs', 'dien_tho')
)
print(unique(neurons$cluster))
```

```
## [1] "mesen_ex_cerebral" "ctx_ex_cerebral"   "dien_ex_cerebral"
## [4] "ge_in_cerebral"    "mesen_in_cerebral" "ctx_hcs"
## [7] "ge_hss"            "dien_tho"
```

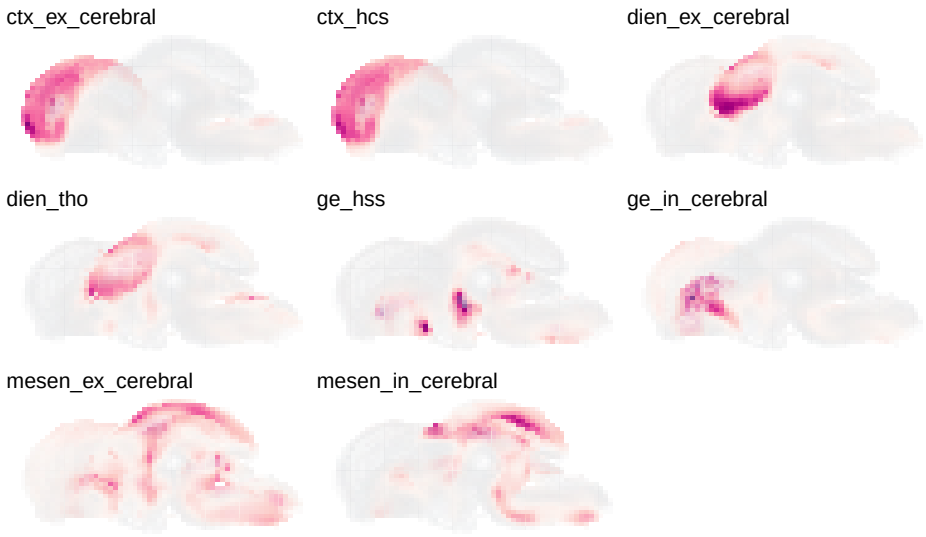
We can see that `cluster` already captures the different neuronal types we are interested in. Now we select some structure markers.

```
struct_markers <- structure_markers('E13', 'custom_3')
genes_use <- struct_markers %>%
    group_by(group) %>%
    top_n(10, auc) %>%
    pull(gene) %>% unique()
print(head(genes_use))
```

```
## [1] "OTP"    "MEST"   "DLK1"   "PEG10"  "CDH8"   "PLXNC1"
```

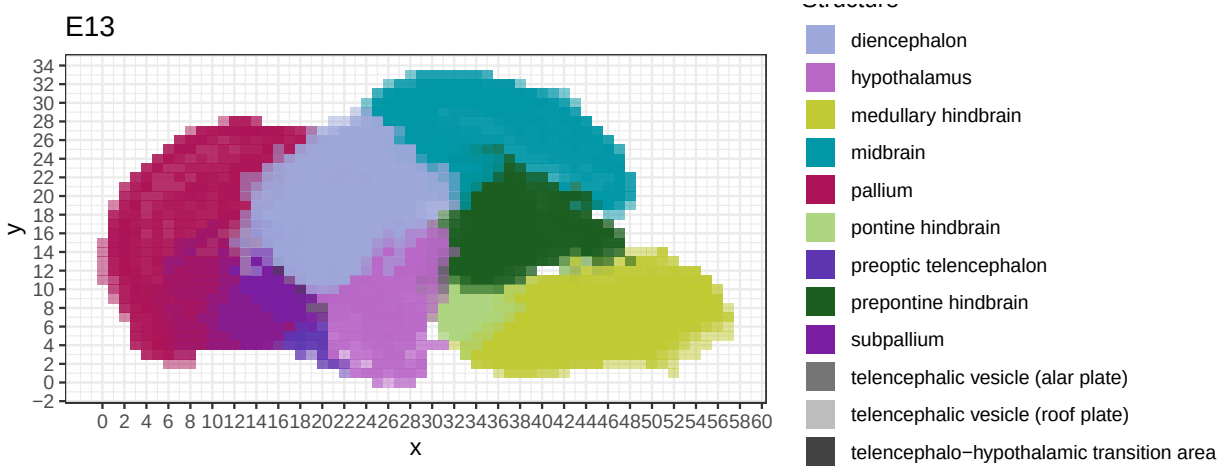
Now we run VoxHunt using these genes

```
neuron_voxmap <- voxel_map(
    neurons,
    group_name='cluster',
    genes_use=genes_use
)
plot_map(neuron_voxmap) & no_legend()
```



As shown in the figure, we can also plot coronal slices. We can first pick the slices from the annotated map.

```
voxhunt::plot_annotation('E13', show_coordinates = T, show_legend = T)
```



Now we plot slices 6, 11, 23 and 28

```
voxhunt::plot_map(neuron_voxmap, view='slice', slices=c(6, 11, 23, 28)) &
    no_legend()
```

Now we can also assign each cell to the highest correlating structure, similar as shown in figure 2.

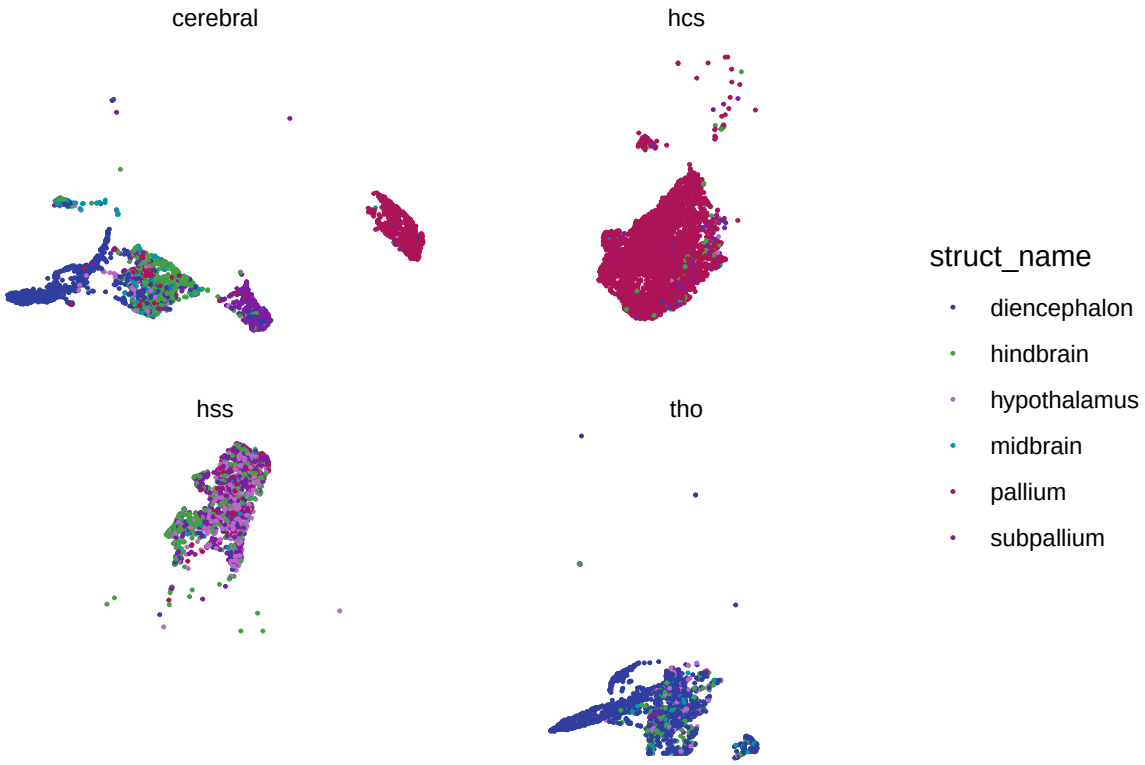
```
cell_assign <- assign_cells(neuron_voxmap)
cell_meta <- as_tibble(neurons@meta.data, rownames='cell') %>%
    dplyr::select(-stage) %>%
    dplyr::inner_join(cell_assign) %>%
    # dplyr::filter(custom_2!='medullary hindbrain') %>%
    dplyr::mutate(struct_name=case_when(
        str_detect(custom_2, 'hindbrain') ~ 'hindbrain',
        str_detect(custom_4, 'septum|subpall|striatum|amygda|telencephalic') ~ 'subpallium',
        str_detect(custom_2, 'telen') ~ 'pallium',
```

```
        TRUE ~ custom_2
    ))
```

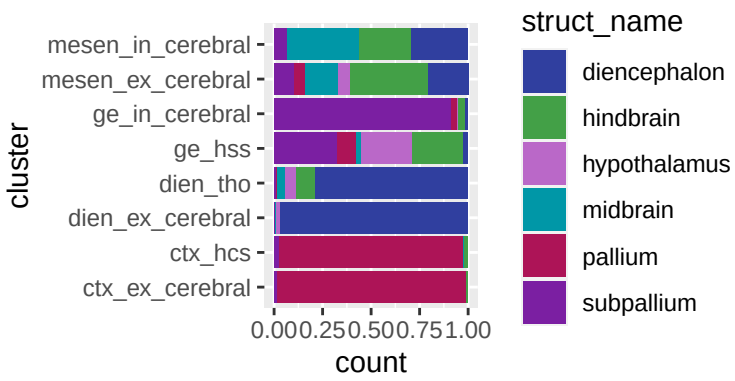
```
## Joining, by = "cell"
```

```
ggplot(cell_meta, aes(UMAP1, UMAP2, color=struct_name)) +
    geom_point(size=0.2) +
    facet_wrap(~dataset) +
    scale_color_manual(values=struct_colors) +
    theme_void()
```

```
## Warning: Removed 18655 rows containing missing values (geom_point).
```




```
ggplot(cell_meta, aes(cluster, fill=struct_name)) +
    geom_bar(position='fill') +
    coord_flip() +
    scale_fill_manual(values=struct_colors)
```


mesen_in_cerebral
mesen_ex_cerebral
ge_in_cerebral
ge_hss
dien_tho
dien_ex_cerebral
ctx_hcs
ctx_ex_cerebral
cluster
0.00
0.25
0.50
0.75
1.00
count
struct_name
diencephalon
hindbrain
hypothalamus
midbrain
pallium
subpallium