Compare expression between cell types with multiple samples

Differential expression analysis can be used to compare cell types. Here, we will compare glioma and oligodendrocyte cells to identify genes differentially regulated in glioma cells from the oligodendroglioma subtype. Glioma cells in oligodendroglioma are thought to originate from oligodendrocytes, thus directly comparing the two cell types will identify genes that distinguish them.

Filter cells

To analyze only the oligodendroglioma subtype, we can filter the samples.

- Click the **Filtered counts** data node
- Expand **Filtering** in the task menu
- Click **Filter cells** (Figure 1)

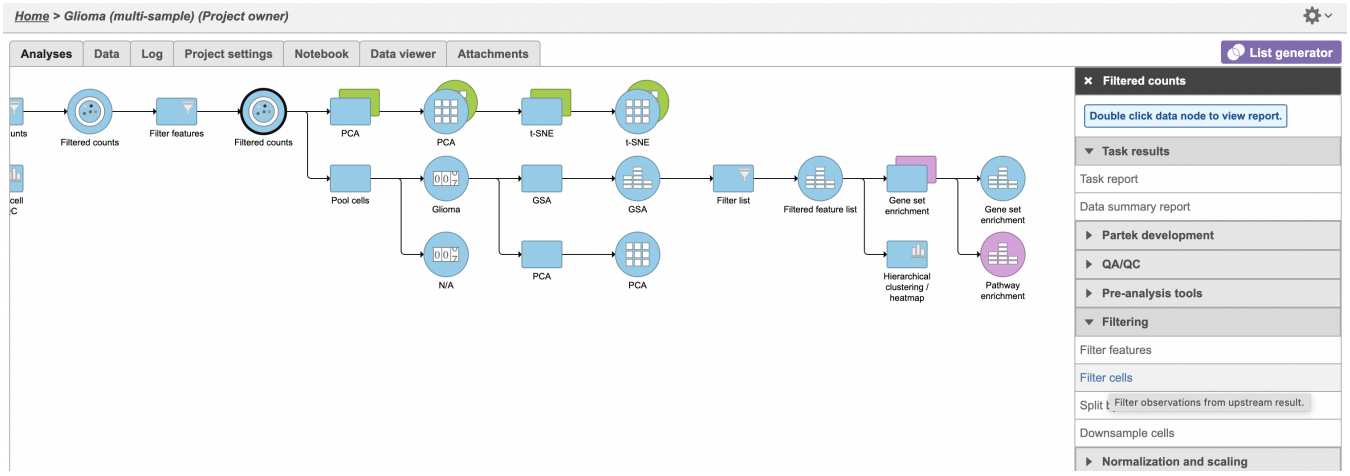


Figure 11. Invoking the sample filter

The filter lets us include or exclude samples based on sample ID and attribute.

- Set the filter to **Include** samples where **Subtype is Oligodendroglioma**
- Click **AND**
- Set the second filter to **exclude Cell type (multi-sample) is Microglia**
- Click **Finish** to apply the filter (Figure 2)

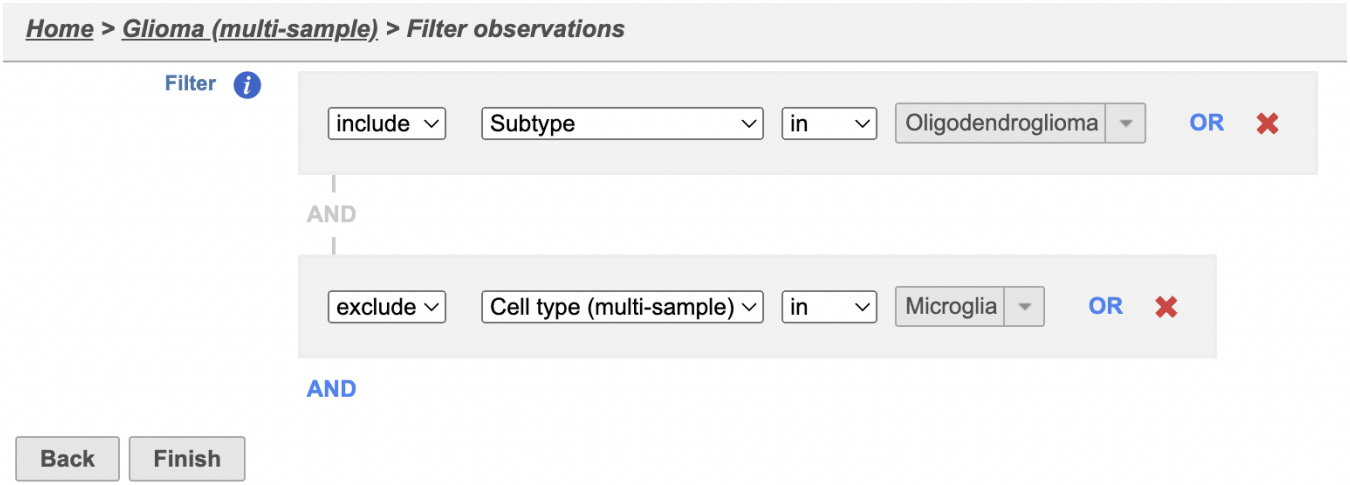


Figure 12. Configuring the group filter

A *Filtered counts* data node will be created with only cells that are from oligodendroglioma samples (Figure 3).

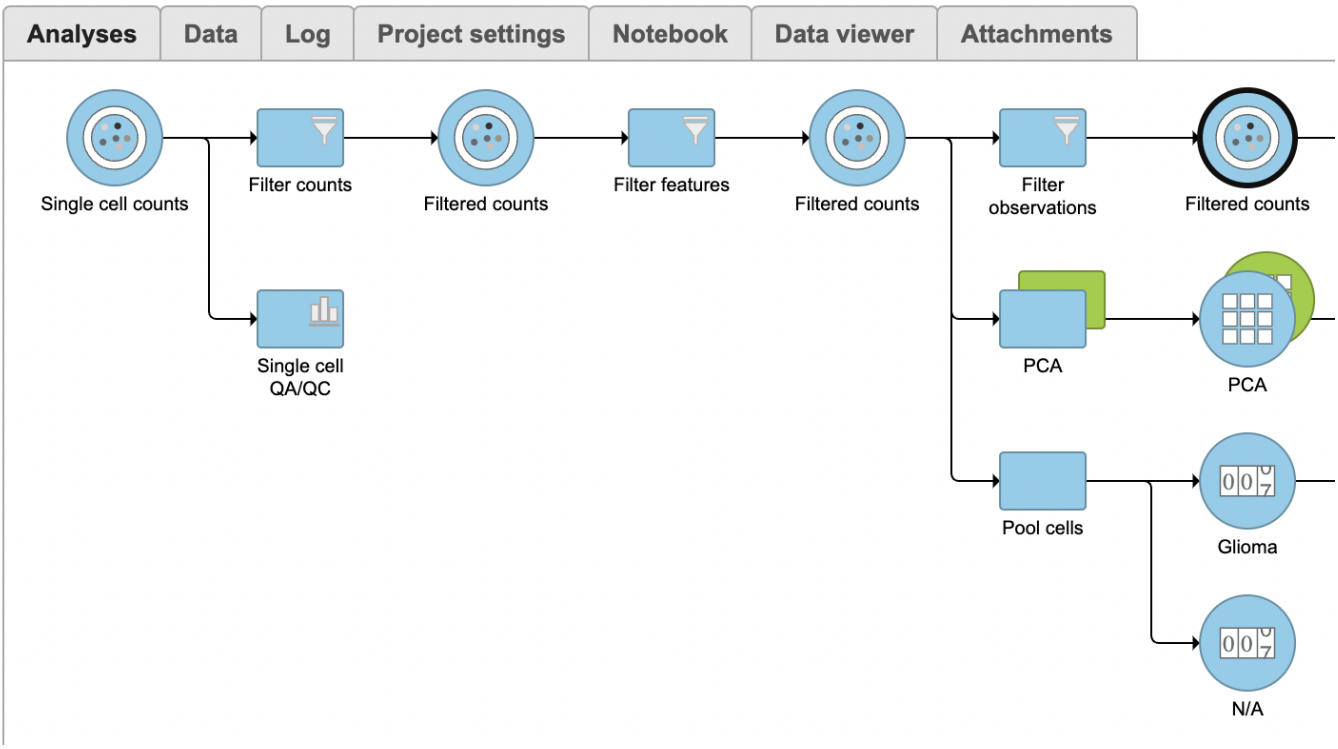


Figure 13. Filtering groups generates a Filtered counts data node

Identify differentially expressed genes

- Click the new **Filtered counts** data node
- Click **Statistics** > **Differential analysis** in the task menu
- Click **GSA**

The configuration options (Figure 4) includes sample and cell-level attributes. Here, we want to compare different cell types so we will include *Cell type (multi-sample)*.

- Click **Cell type (multi-sample)**
- Click **Next**

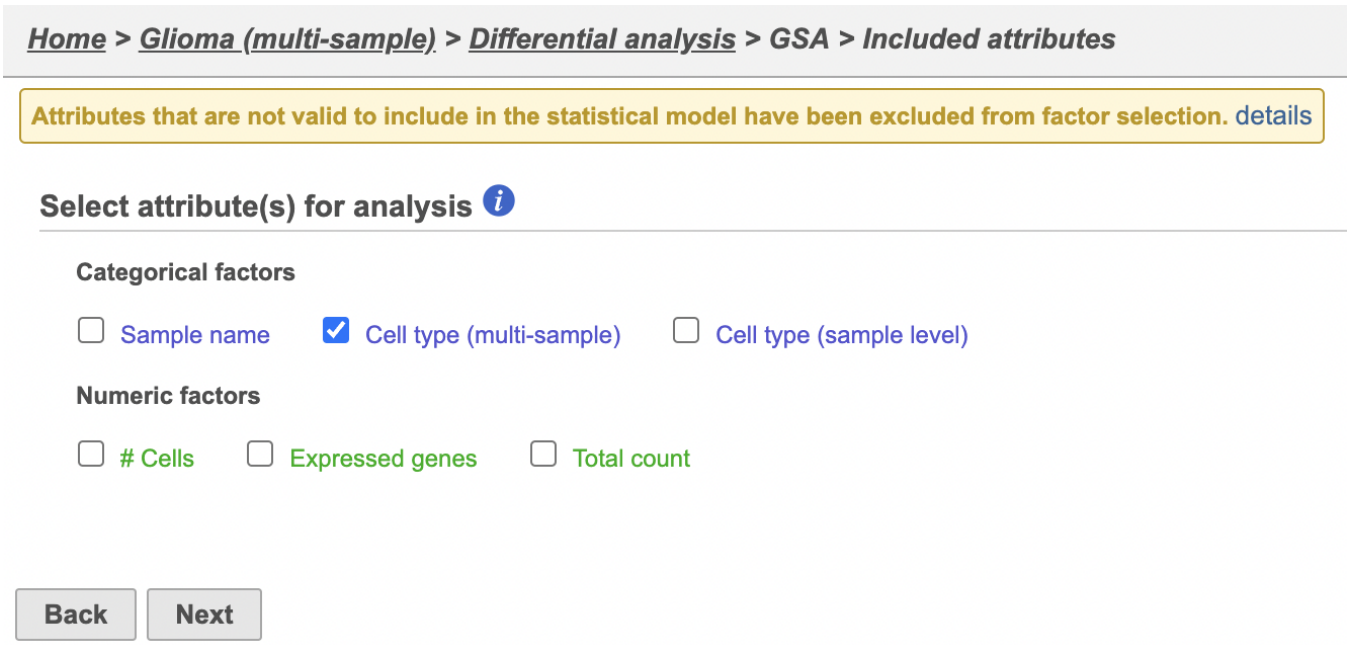


Figure 14. Choosing attributes to include in the statistical test

Next, we will set up a comparison between glioma and oligodendrocyte cells.

- Click **Glioma** in the top panel
- Click **Oligodendrocytes** in the bottom panel
- Click **Add comparison** (Figure 5)

This will set up fold calculations with glioma as the numerator and oligodendrocytes as the denominator.

Comparison selector

Cell type (multi-sample)

- [x] Glioma
- [ ] Oligodendrocytes

Glioma

vs.

Cell type (multi-sample)

- [ ] Glioma
- [x] Oligodendrocytes

Oligodendrocytes

Add comparison

	Cell type (multi-sample)		Cell type (multi-sample)	
1	Glioma	vs.	Oligodendrocytes	✖

Low value filter

Filter features by

- ○ Lowest average coverage — 1.00
- ○ Lowest maximum coverage — 1.00
- ○ Lowest total coverage — 10.00
- ◉ None

Advanced options

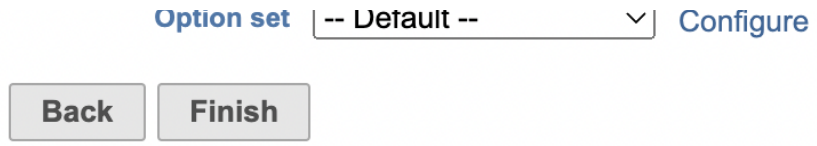


Figure 15. Defining the comparison between Glioma and Oligodendrocytes

- Click **Finish** to run the GSA

A green *GSA* data node will be generated containing the results of the GSA.

- Double-click the **green GSA** data node to open the GSA report

Because of the large number of cells and large differences between cell types, the p-values and FDR step up values are very low for highly significant genes. We can use the volcano plot to preview the effect of applying different significance thresholds.

- Click  to view the **Volcano plot**
- Open the **Style** icon on the left, change *Size* **point size to 6**
- Open the **Axes** icon on the left and change the Y-axis to **FDR step up (Glioma vs Oligodendrocytes)**
- Open the **Statistics** icon and change the *Significance* of **X threshold to -10 and 10** and the **Y threshold to 0.001**
- Open the **Select & Filter** icon, set the **Fold change thresholds to -10 and 10**
- In **Select & Filter**, click  to remove the **P-value (Glioma vs Oligodendrocytes)** selection rule. From the drop-down list, add **FDR step up (Glioma vs Oligodendrocytes)** as a selection rule and set the maximum to 0.001

Note these changes in the icon settings and volcano plot below (Figure 6).

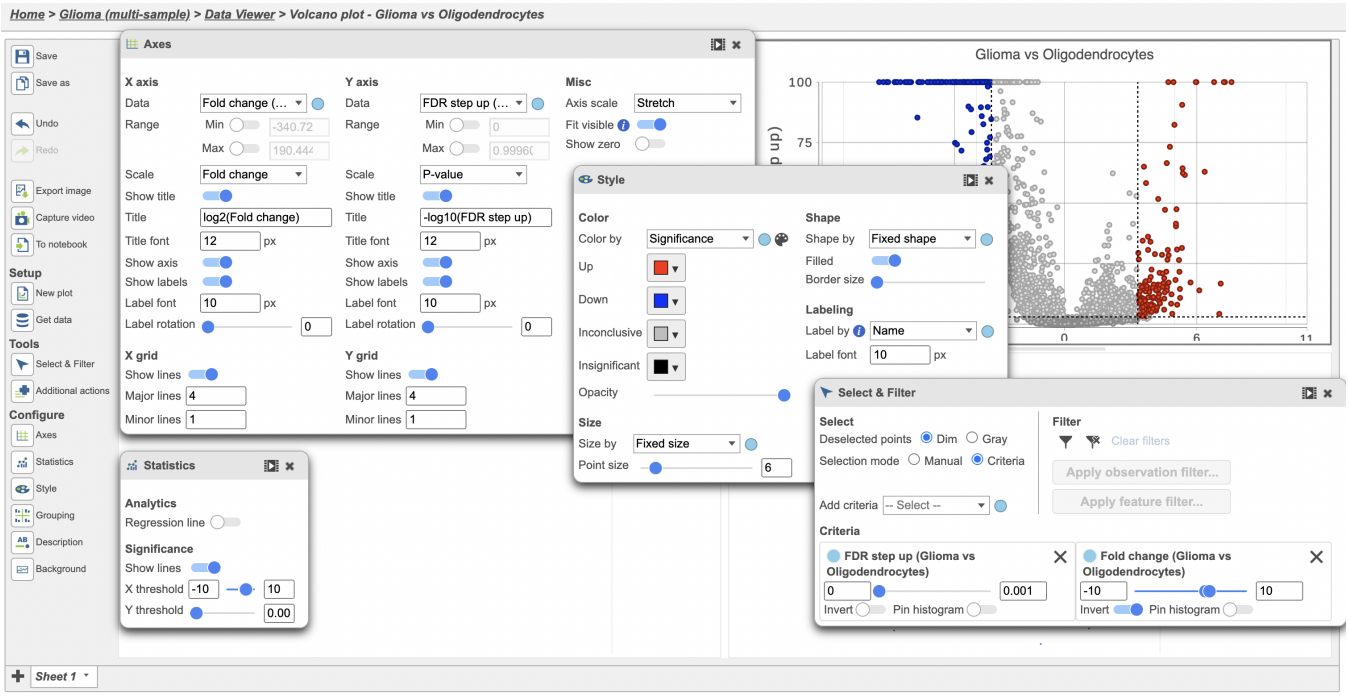
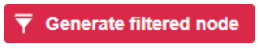


Figure 16. Previewing a filter by adjusting the size of the points, changing the Y-axis, adjusting the X & Y significance thresholds and changing the selection

We can now recreate these conditions in the GSA report filter.

- Click **GSA report** tab in your web browser to return to the GSA report
- Click **FDR step up**
- Set the **FDR step up** filter to **Less than or equal to 0.001**
- Press **Enter**
- Click **Fold change**
- Set the **Fold change** filter to **From -10 to 10**
- Press **Enter**

The filter should include 291 genes.

- Click  to apply the filter and generate a *Filtered Feature list* node

Exploring differentially expressed genes

To visualize the results, we can generate a hierarchical clustering heatmap.

- Click the **Filtered feature list** produced by the *Differential analysis filter* task
- Click **Exploratory analysis** in the task menu
- Click **Hierarchical clustering/heatmap**

Using the hierarchical clustering options we can choose to include only cells from certain samples. We can also choose the order of cells on the heatmap instead of clustering. Here, we will include only glioma cells and order the samples by sample name (Figure 7).

- Make sure **Cluster** is unchecked for *Cell order*
- Click **Filter cells** under *Filtering* and set the filter to **include Cell type (multi-sample) is Glioma**
- Choose **Sample name** from the *Cell order* drop-down menu in the *Assign order* section
- Click **Finish**

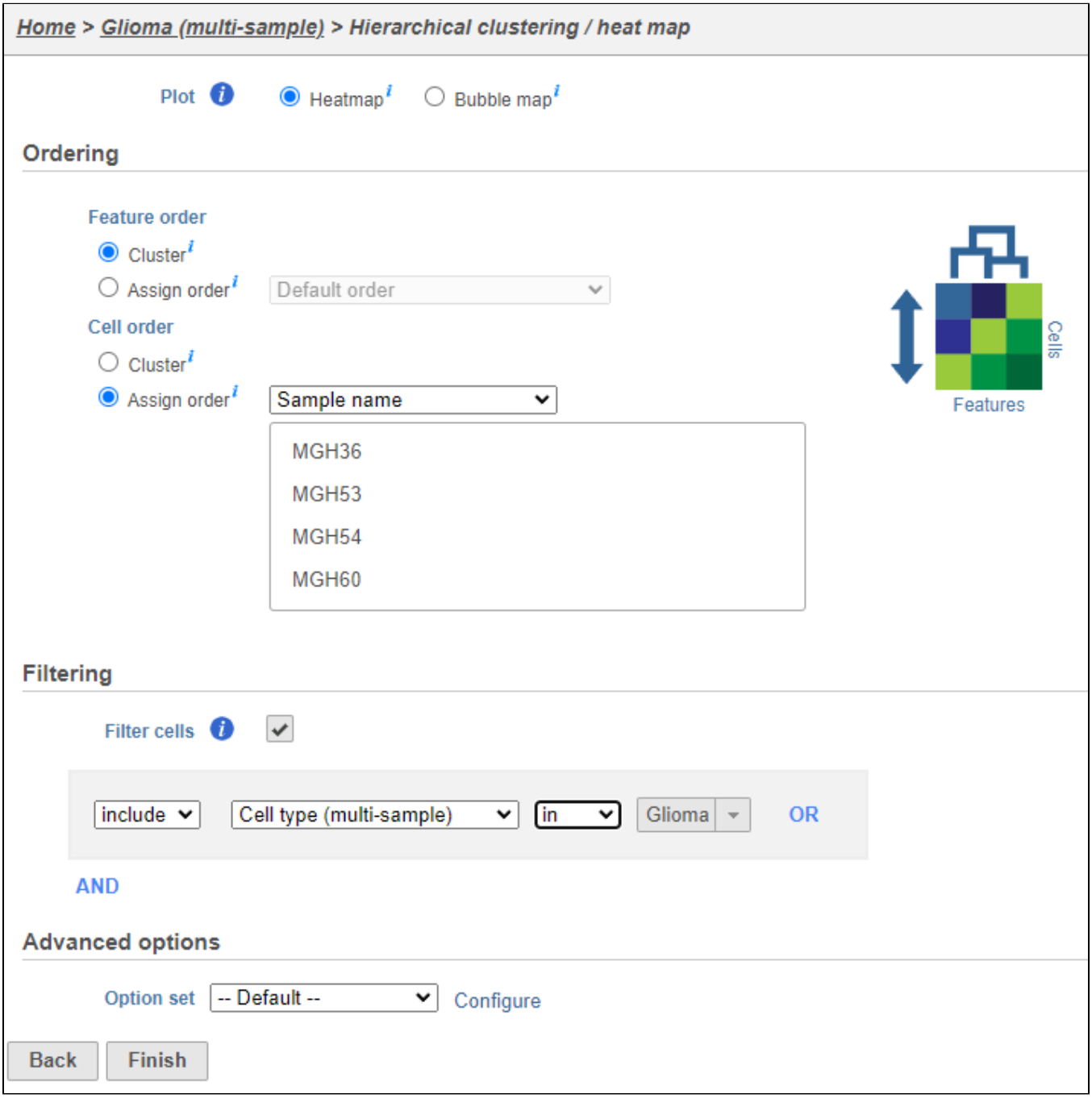


Figure 17. Configuring hierarchical clustering

- Double click the green **Hierarchical clustering** node to open the heatmap

The heatmap differences may be hard to distinguish at first; the range from red to blue with a white midpoint is set very wide because of a few outlier cells. We can adjust the range to make more subtle differences visible. We can also adjust the color.

- Set the **Range** toggle **Min** to **-1.5**
- Set the **Range** toggle **Max** to **1.5**

The heatmap now shows clear patterns of red and blue.

- Click **Axis titles** and deselect the **Row labels** and **Column labels** of the panel to hide sample and feature names, respectively.
- Select **Sample name** from the *Annotations* drop-down menu

Cells are now labeled with their sample name. Interestingly, samples show characteristic patterns of expression for these genes (Figure 8).

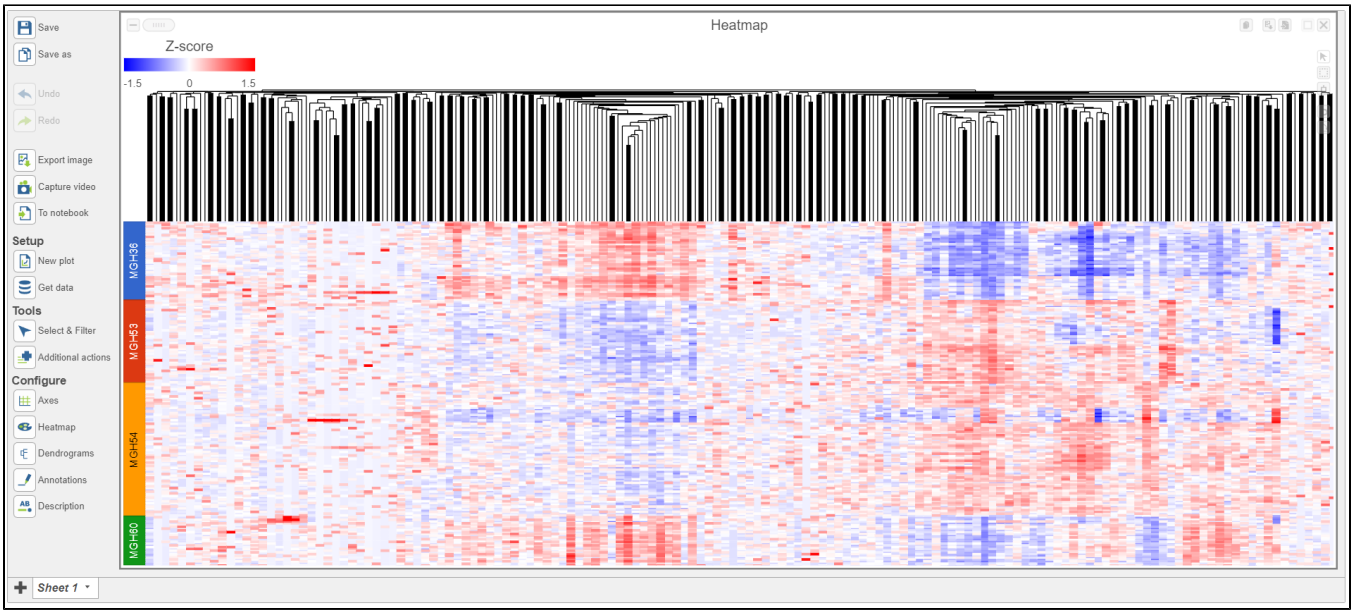


Figure 18. Hierarchical clustering heatmap with cells on rows (ordered by sample name) and genes on columns (clustered)

- Click **Glioma (multi-sample)** to return to the *Analyses* tab.

We can use gene set enrichment to further characterize the differences between glioma and oligodendrocyte cells.

- Click the **Filtered feature list** node
- Click **Biological interpretation** in the task menu
- Click **Gene set enrichment**
- Change *Database* to **Gene set database** and click Finish to continue with the most recent gene set (Figure 9)

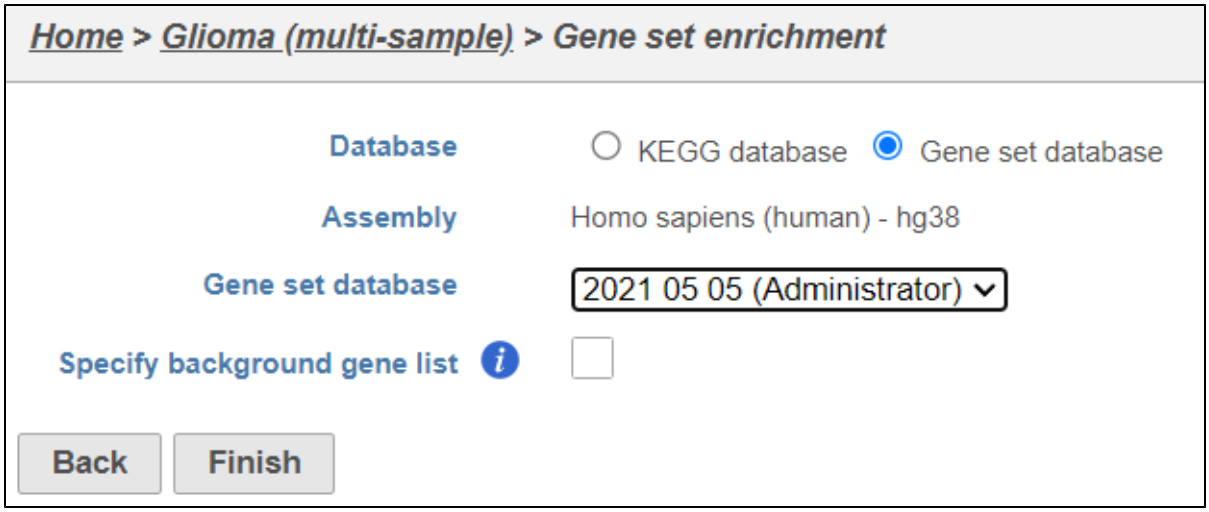


Figure 19. Gene set enrichment dialogue

A *Gene set enrichment* node will be added to the pipeline .

- Double-click the **Gene set enrichment** task node to open the task report

Top GO terms in the enrichment report include "ensheathment of neurons" and "axon ensheathment" (Figure 10), which corresponds well with the role of oligodendrocytes in creating the myelin sheath that supports and protect axons in the central nervous system.

Gene set	Description	Enrichment score	P-value	Genes in list	Genes not in list	
GO:0051960	regulation of nervous system development	30.93	3.7E-14	47	822	
GO:0050767	regulation of neurogenesis	27.72	9.18E-13	42	732	
GO:0048731	system development	27.17	1.58E-12	39	647	
GO:0045664	regulation of neuron differentiation	25.67	7.14E-12	36	585	
GO:0060284	regulation of cell development	24.62	2.03E-11	43	845	
GO:0051961	negative regulation of nervous system development	24.46	2.39E-11	24	267	
GO:0050768	negative regulation of neurogenesis	24.06	3.55E-11	23	248	
GO:0042552	myelination	22.09	2.55E-10	13	67	
GO:0007272	ensheathment of neurons	21.77	3.52E-10	13	69	
GO:0008366	axon ensheathment	21.77	3.52E-10	13	69	
GO:0007155	cell adhesion	21.45	4.86E-10	40	828	
GO:0010721	negative regulation of cell development	21.44	4.87E-10	23	286	
GO:0050793	regulation of developmental process	21.39	5.11E-10	76	2,358	
GO:0022610	biological adhesion	21.21	6.14E-10	40	835	
GO:0045665	negative regulation of neuron differentiation	20.91	8.26E-10	19	194	

Figure 20. GO enrichment task report

Additional Assistance

If you need additional assistance, please visit our support page to submit a help ticket or find phone numbers for regional support.

Your Rating:

Results: 31 rates