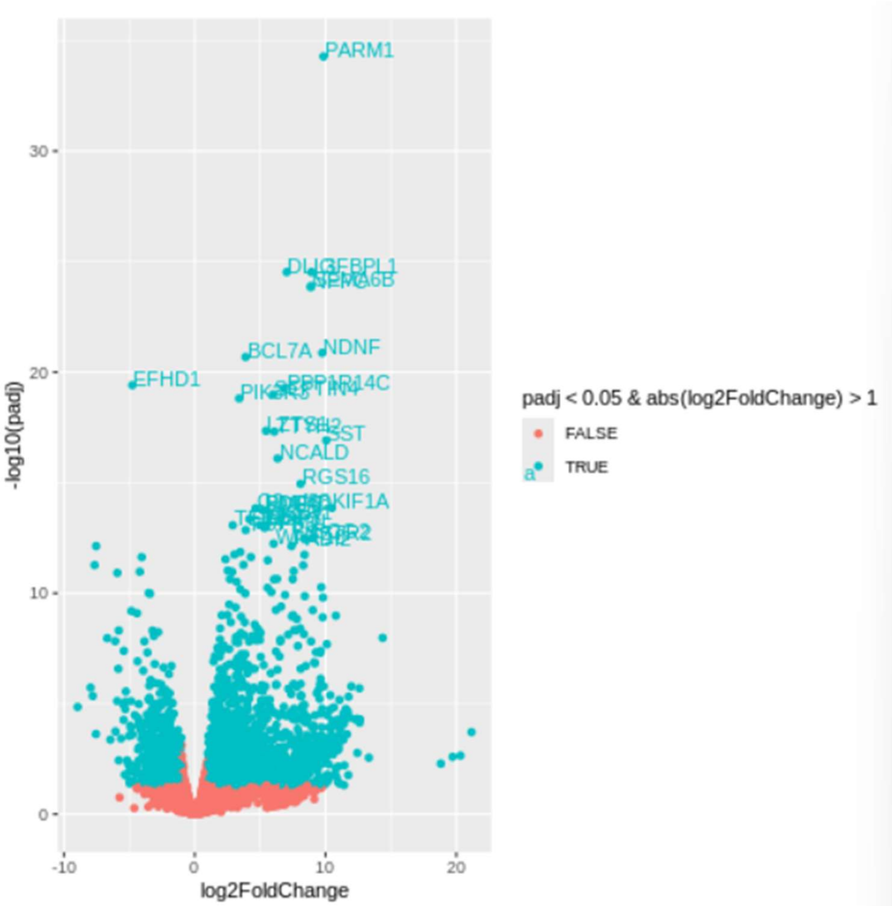
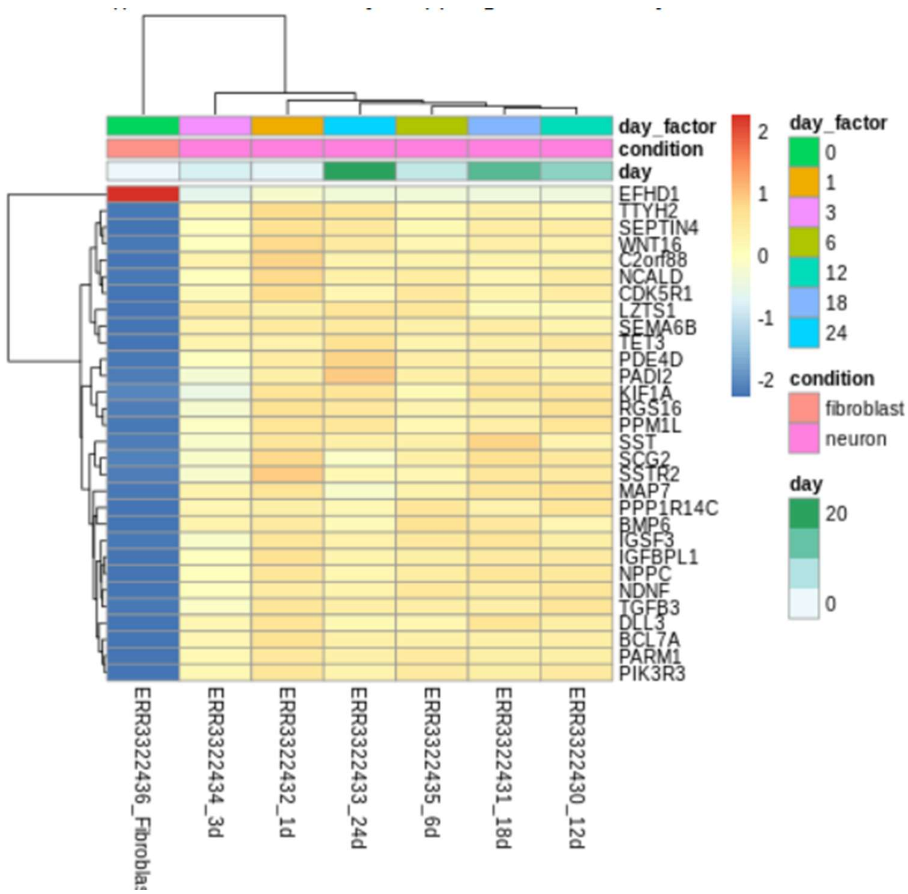
Week 7 – Further analysis of top 30 genes to GSEA

A volcano plot was constructed for the top 30 genes:



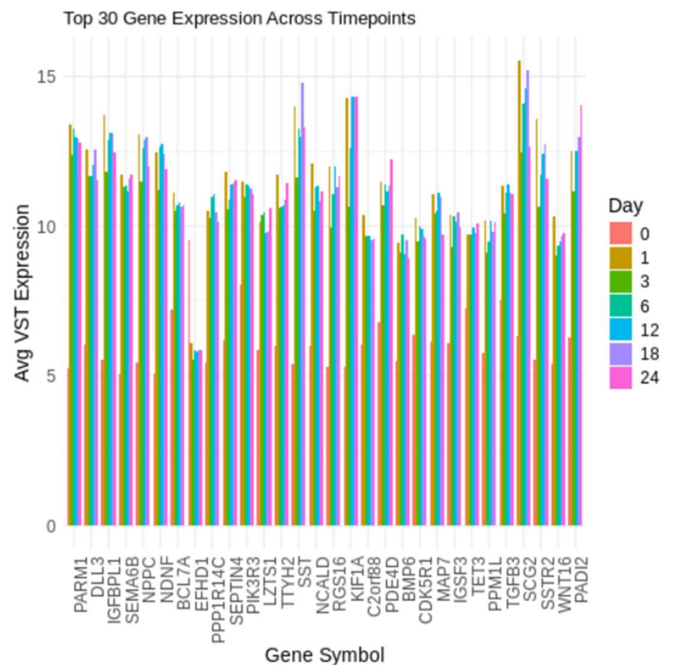
This volcano plot visualizes differential gene expression, with log2 fold change on the x-axis (indicating up- or downregulation) and -log10 adjusted p-value on the y-axis (indicating significance). Each point represents a gene, with turquoise points marking those that are statistically significant (padj < 0.05) and have a large effect size (|log2FoldChange| > 1). These highlighted genes are the most biologically relevant, with some top ones labeled, such as PARM1 and DLG3. The red points are either not significant or have small fold changes.

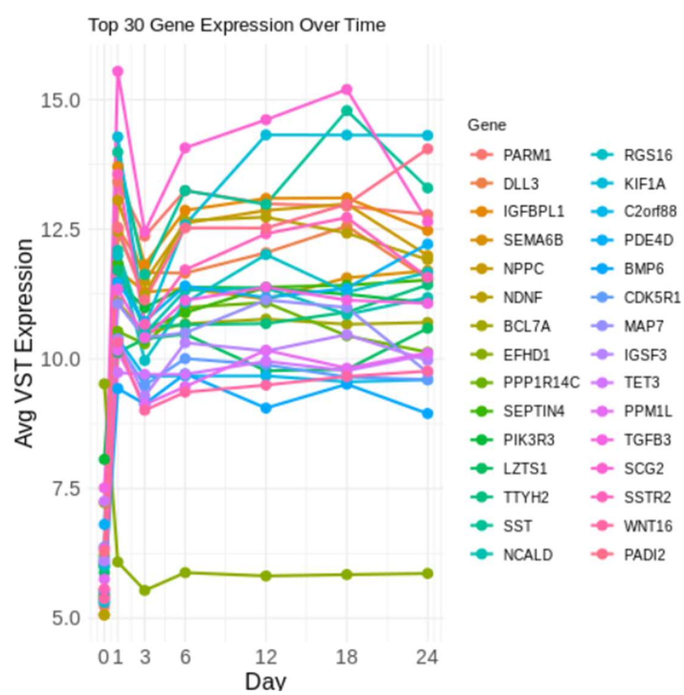
Then a heatmap was plotted to understand the relationship between the different genes



This heatmap shows the expression patterns of selected genes across different samples and time points during the differentiation of fibroblasts into neurons. Each row represents a gene, and each column a sample, with colors indicating relative expression levels (blue = low, yellow/red = high). Samples are grouped by condition (fibroblast vs neuron) and time point (from day 0 to day 24), with clustering revealing similar expression patterns over time. Early samples (e.g., fibroblasts at day 0) show distinct gene expression (blue, low), while neuronal samples progressively increase expression in key genes (toward yellow/red), reflecting dynamic transcriptional changes during differentiation.

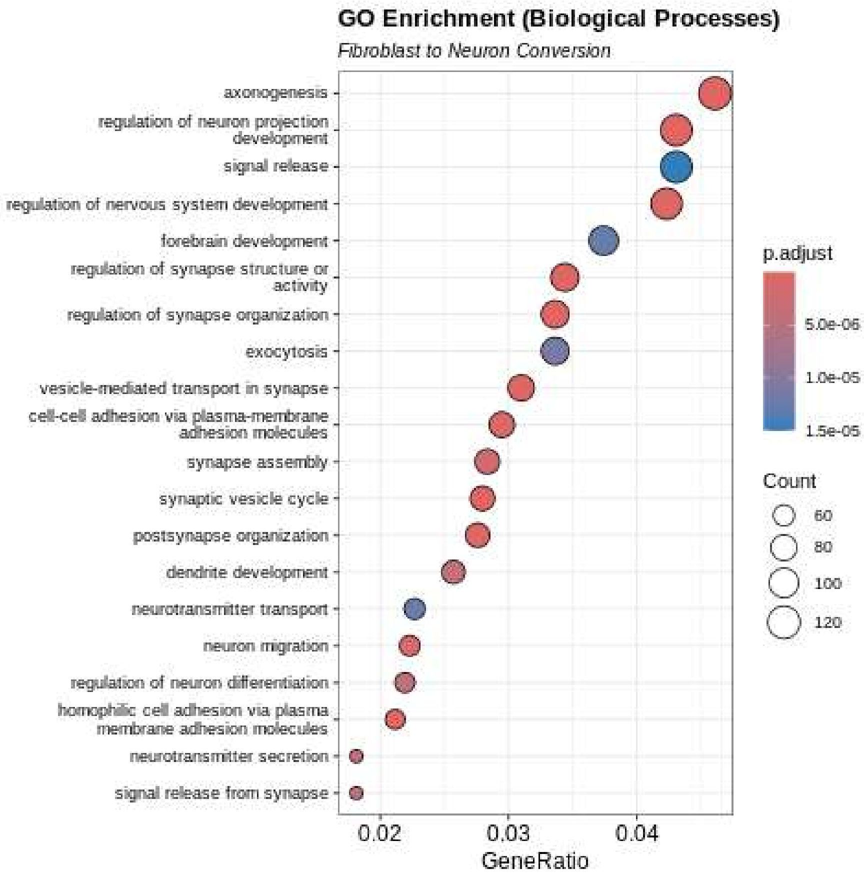
The obtained top 30 genes were then plotted, in two ways. The first, with genes on the x-axis and their days of expression on the y-axis and the second with genes on the y axis and no. of days on the x-axis to understand better their expression patterns.





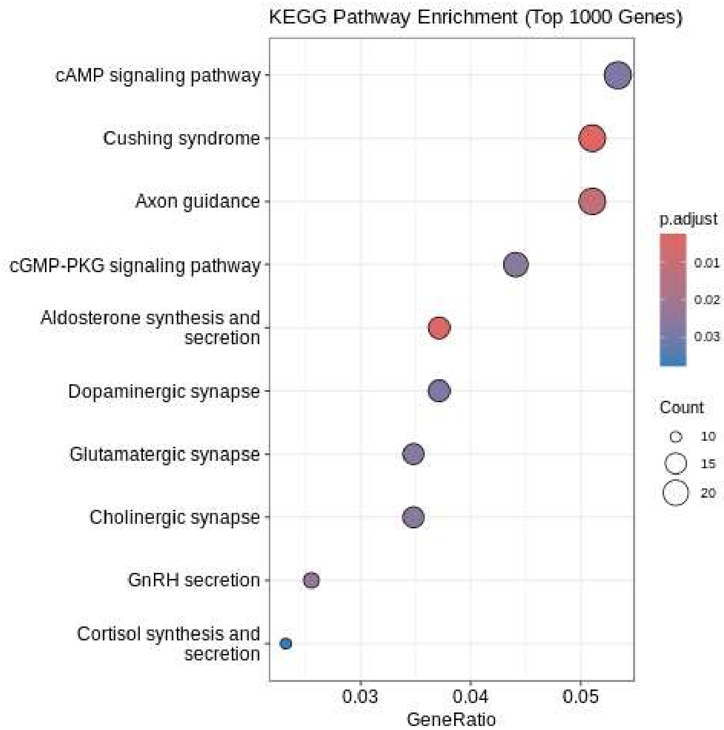
A sharp increase in expression for most genes between Day 0 and Day 1, followed by stabilization or gradual elevation through Day 24. Notably, genes such as **IGFBPL1, IGSF3, SEMA6B, NPPC, and TGFβ3** exhibit the highest average VST (Variance Stabilizing Transformation) expression levels across the time course, consistently maintaining values around or above 13–15. These genes show sustained high expression from early stages onward, suggesting their prominent involvement in neuronal development and maturation processes.

Next, GO enrichment analysis was performed.



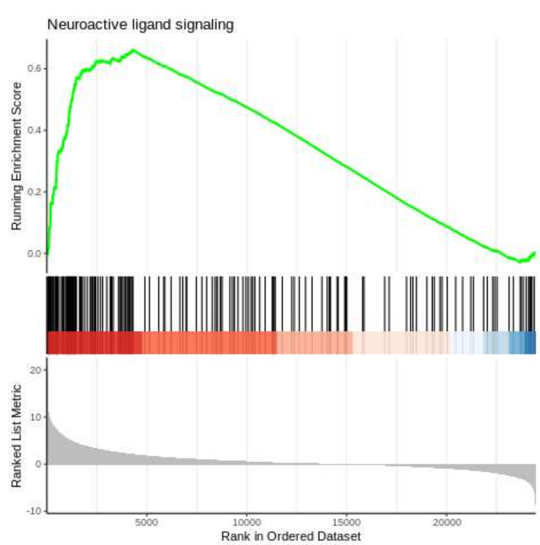
The Gene Ontology (GO) enrichment plot highlights the most significantly enriched biological processes during fibroblast-to-neuron conversion. The x-axis shows the GeneRatio (proportion of input genes associated with each GO term), while bubble size reflects gene count, and colour indicates the adjusted p-value (darker red = more significant). Key enriched processes include axonogenesis, regulation of neuron projection development, signal release, and nervous system development, suggesting that the reprogramming process strongly activates pathways critical for neuronal differentiation and synaptic function. Overall, the results indicate a robust shift toward neural identity, emphasizing processes involved in neuron structure, connectivity, and communication.

This was followed by KEGG Pathway enrichment analysis for top 1000 genes, to increase resolution, as significant results were not being shown for <500 genes.



The most significant pathways, including cAMP signaling and axon guidance, demonstrate the complex cellular reprogramming required to transform a fibroblast into a functional neuron. Multiple neurotransmitter synaptic pathways (dopaminergic, glutamatergic, cholinergic) are enriched, indicating the extensive molecular redesign necessary for neuronal differentiation. The prominent axon guidance pathway suggests critical architectural restructuring, while the cAMP signaling pathway points to fundamental changes in gene expression, synaptic plasticity, and neuronal development. This analysis underscores that converting fibroblasts to neurons is far more than a simple cell type switch—it's a comprehensive molecular transformation involving structural remodeling, signaling network reconfiguration, and deep epigenetic reprogramming.

Finally the Gene set enrichment analysis graph was plotted. GSEA (Gene Set Enrichment Analysis) is a computational method used to determine whether predefined sets of genes (e.g., pathways or functional categories) show statistically significant, coordinated differences between two biological states (such as treated vs. control). Unlike single-gene analysis, GSEA considers all genes ranked by expression changes, capturing subtle but coordinated shifts in gene groups.



This graph depicts the Neuroactive Ligand Signaling pathway during fibroblast-to-neuron conversion, revealing a dynamic molecular transformation. The green line shows a striking enrichment score that peak and then gradually declines, suggesting a complex progression of signaling mechanisms critical to neuronal reprogramming. The color-coded bar beneath indicates varied gene expression patterns, with red and blue regions potentially representing different stages of cellular conversion. The shifting enrichment score implies that neuroactive ligand signaling is not a static process but a nuanced journey of molecular reconfiguration, where certain signaling pathways are dramatically upregulated during the initial stages of fibroblast reprogramming and then modulate as the cells transition towards a more mature neuronal state. This visualization captures the intricate molecular choreography involved in converting fibroblasts into functional neurons, highlighting the dynamic nature of cellular reprogramming.