
SINGLE-CELL TRANSCRIPTOMIC PROFILING OF BETA CELLS IN HUMAN PANCREATIC ISLETS: COMPARATIVE ANALYSIS OF HEALTHY AND TYPE 2 DIABETIC STATES

Group - 3

Dinesh Sambhaji Pradhan
Asmitha Nagajothi Purushotam
Vedant Kulkarni
Rich Goodier

BINF6310 - Introduction to Computational Methods in Bioinformatics
Instructor: Andrey J. Savol, PhD

Agenda

- BACKGROUND
- MOTIVATION AND OBJECTIVE
- DATA SOURCE
- WORKFLOW
- KEY FINDINGS (Analysis)
- CONCLUSION & FUTURE WORK
- REFERENCES
- MISCELLANEOUS

Background

- **Type 2 Diabetes on the rise** - According to IDF (2025) 1 in 9 have diabetes and by 2050, it is expected to increase 46%.
- **Role of Islets of Langerhans** - endocrine powerhouse, dysfunction disrupts homeostasis
- **Cause of T2D** - insulin resistance, islet cell breakdown

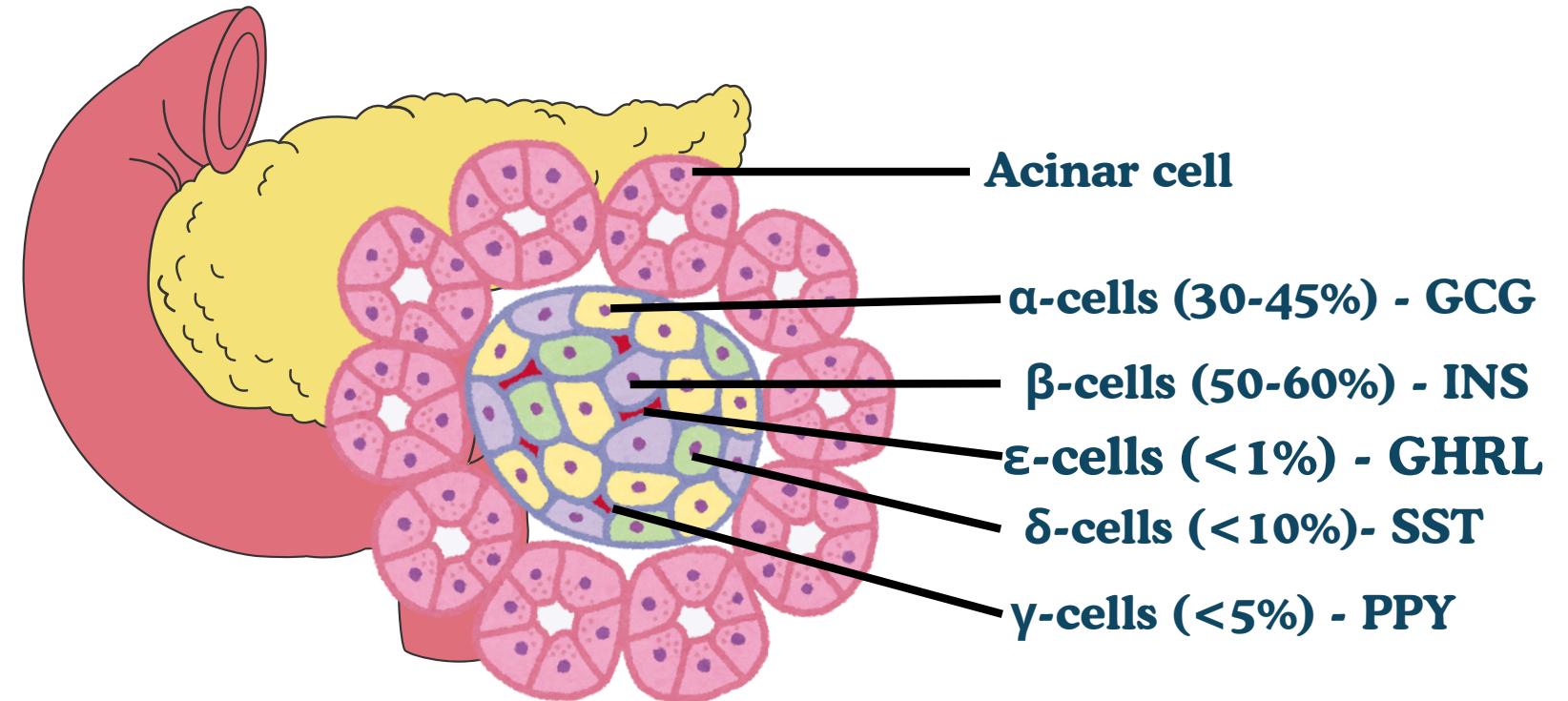


Fig. 1. Islets of Langerhans

Motivation and Objective



Motivation

- Bulk RNA-seq lacks cellular resolution; single-cell RNA-seq reveals beta-cell heterogeneity.
- Profiling gene expression in healthy vs. diabetic beta cells can uncover disease-related targets



Objective

- Build a pipeline to analyze single-cell RNA-seq data.
- Compare beta-cell transcriptomes in healthy and diabetic samples.
- Identify differentially expressed genes.

Dataset & Reference Details

Sample Overview

Condition	Number of individuals	Number of cells	Cell Type	Protocol
Healthy	6	1,492	Beta Cells	Smart-Seq2
T2D Patients	4	717		

Total Cells Analyzed: 2,209

Data Acquisition - [scRNA_human_pancreas.fastq](#)

Reference Genome - GRCh38 with Ensembl GTF annotation

Methods

Smartseq2 Protocol

Cell Isolation

Individual cells are isolated, often via micromanipulation, FACS, or microfluidics.



Cell Lysis & RNA Capture

Each cell is lysed, and mRNA is captured using oligo(dT) primers



Reverse Transcription

mRNA is reverse transcribed into cDNA using a template-switching mechanism, to improve coverage at the 5' end of transcripts



cDNA Amplification

Full-length cDNA is amplified via PCR, preserving transcript information for low-abundance genes



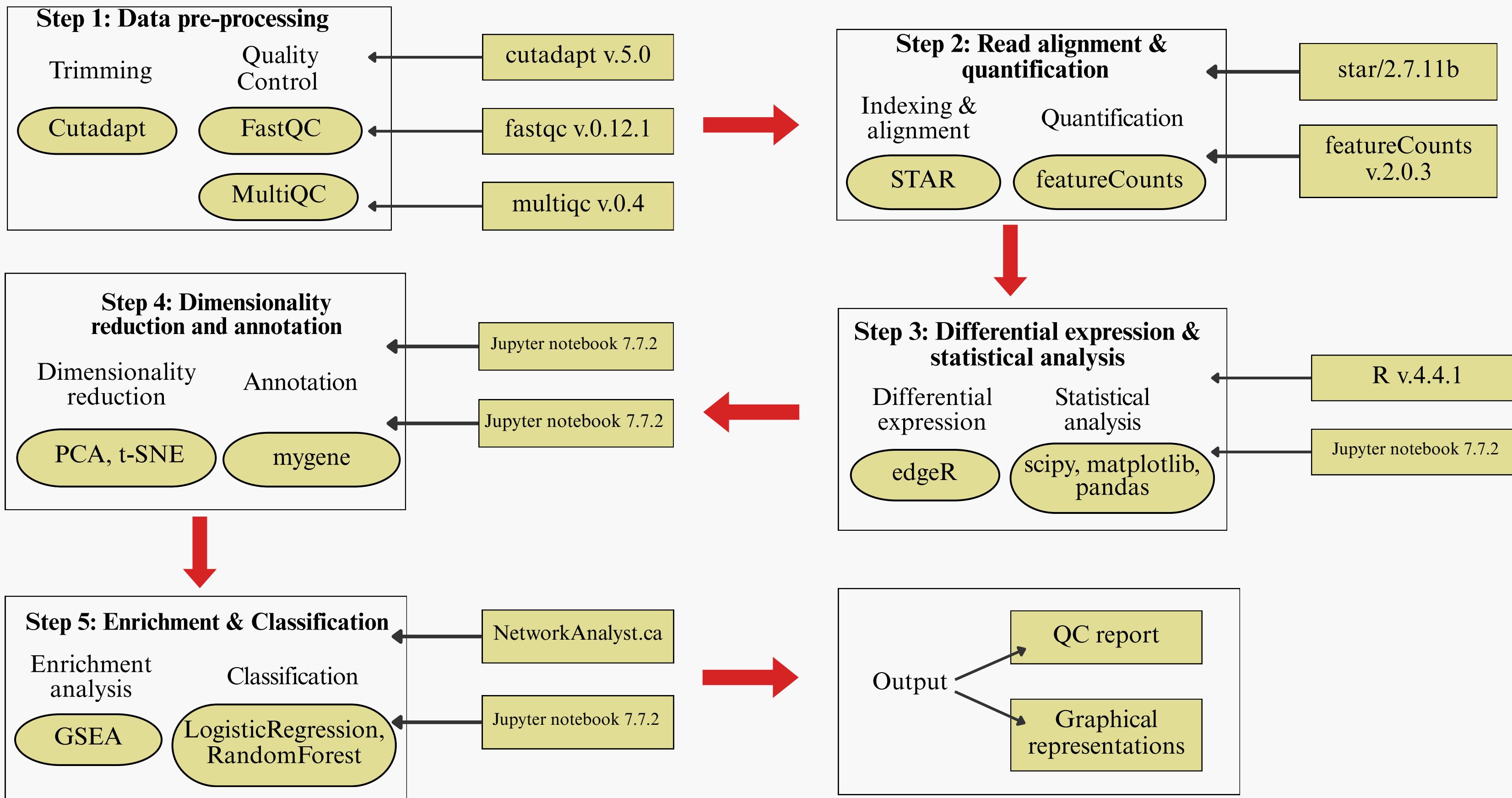
Library Preparation & Sequencing

Libraries sequenced on Illumina HiSeq2000

Pipeline - RNA-seq Analysis

Data Acquisition - .fastq

[*scrna_human_pancreas.fastq*](#)



Results

Multi QC Report

General Statistics

Showing 3274 samples.

Configure columns

Summarize plot Export...

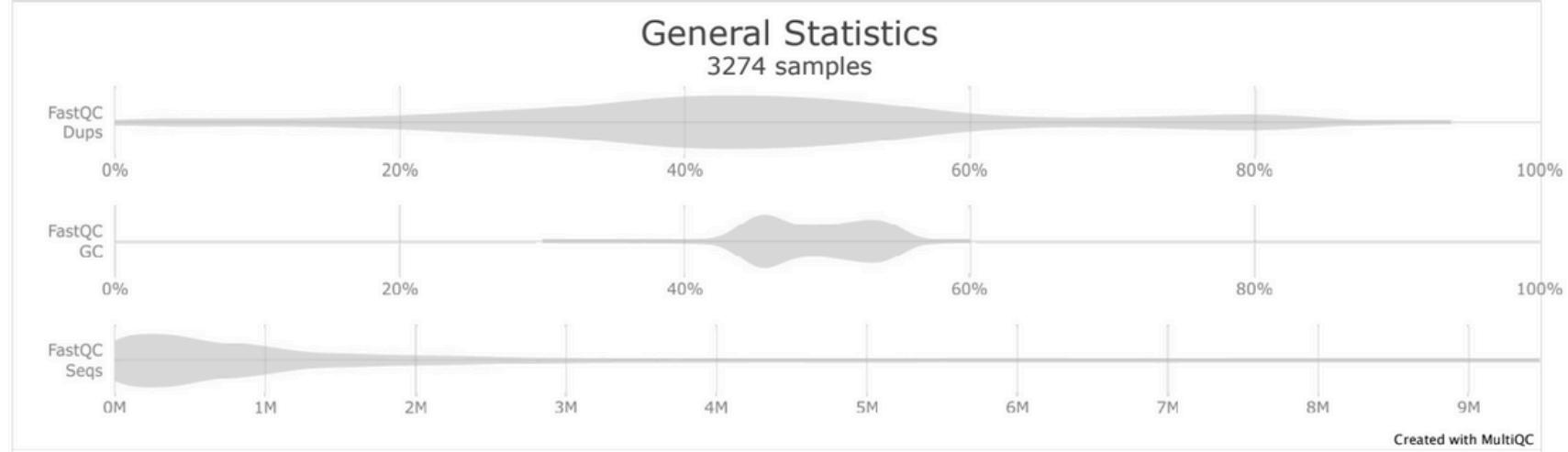


Fig. 2.1. General statistics

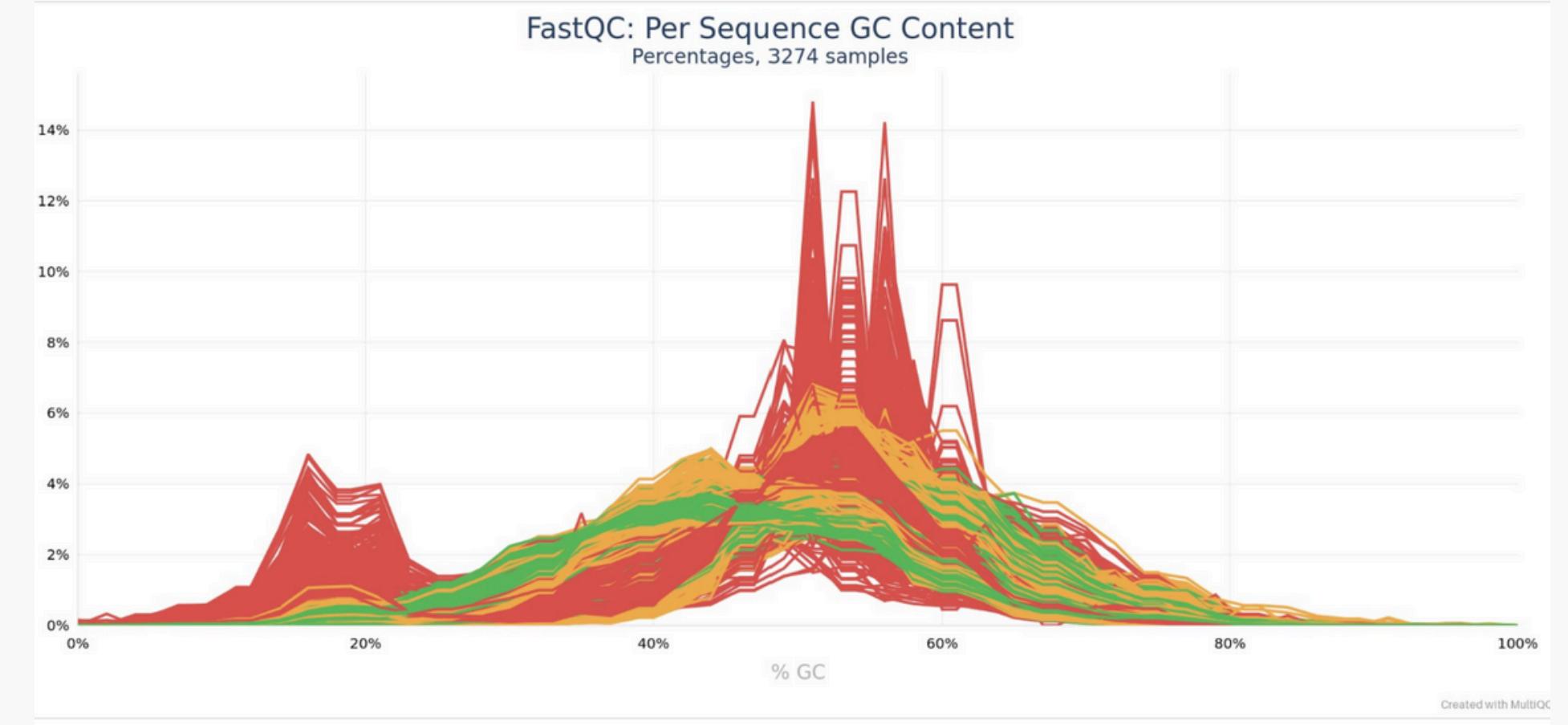


Fig. 2.2 Per Sequence GC Content

Multi QC Report

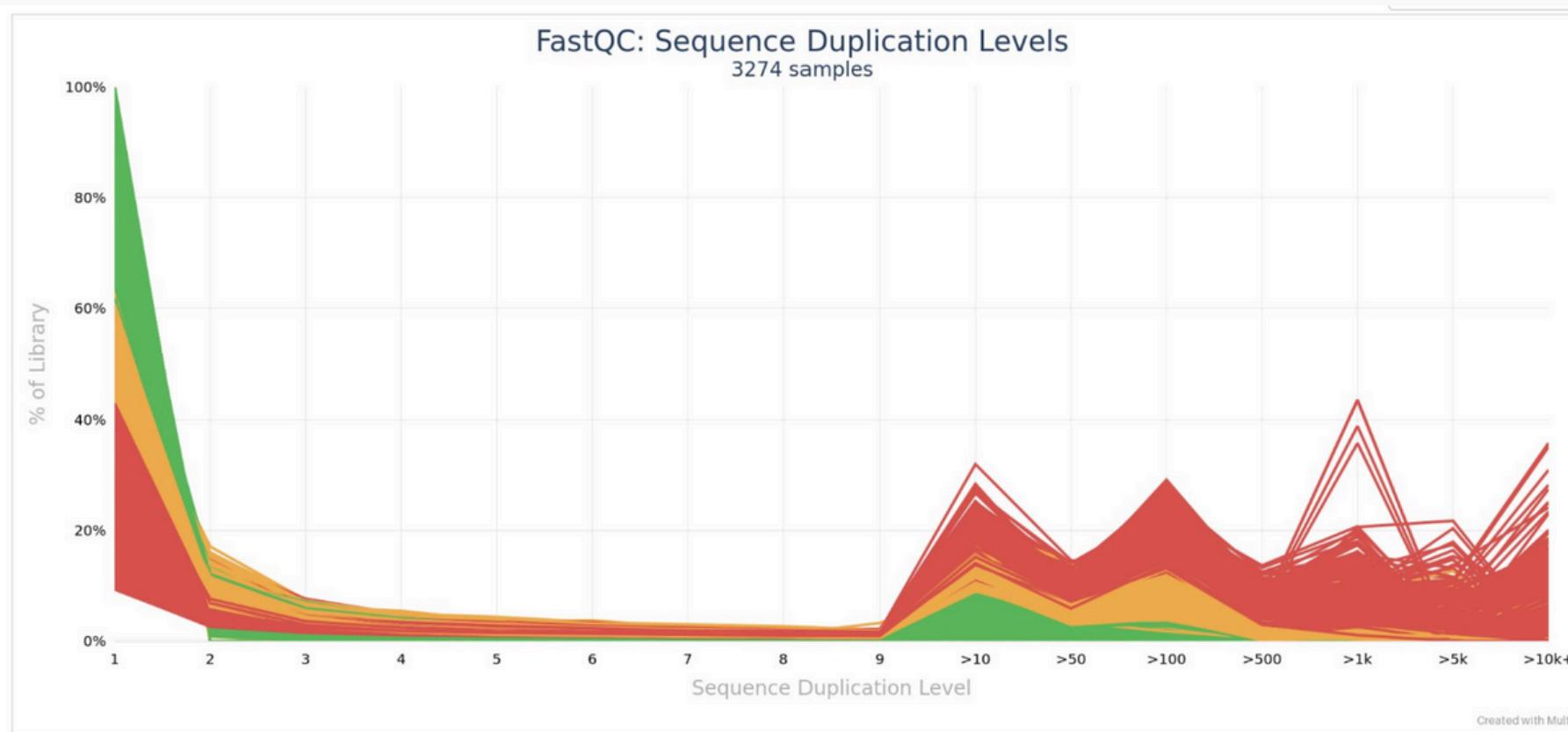


Fig. 2.3 Sequence Duplication Levels

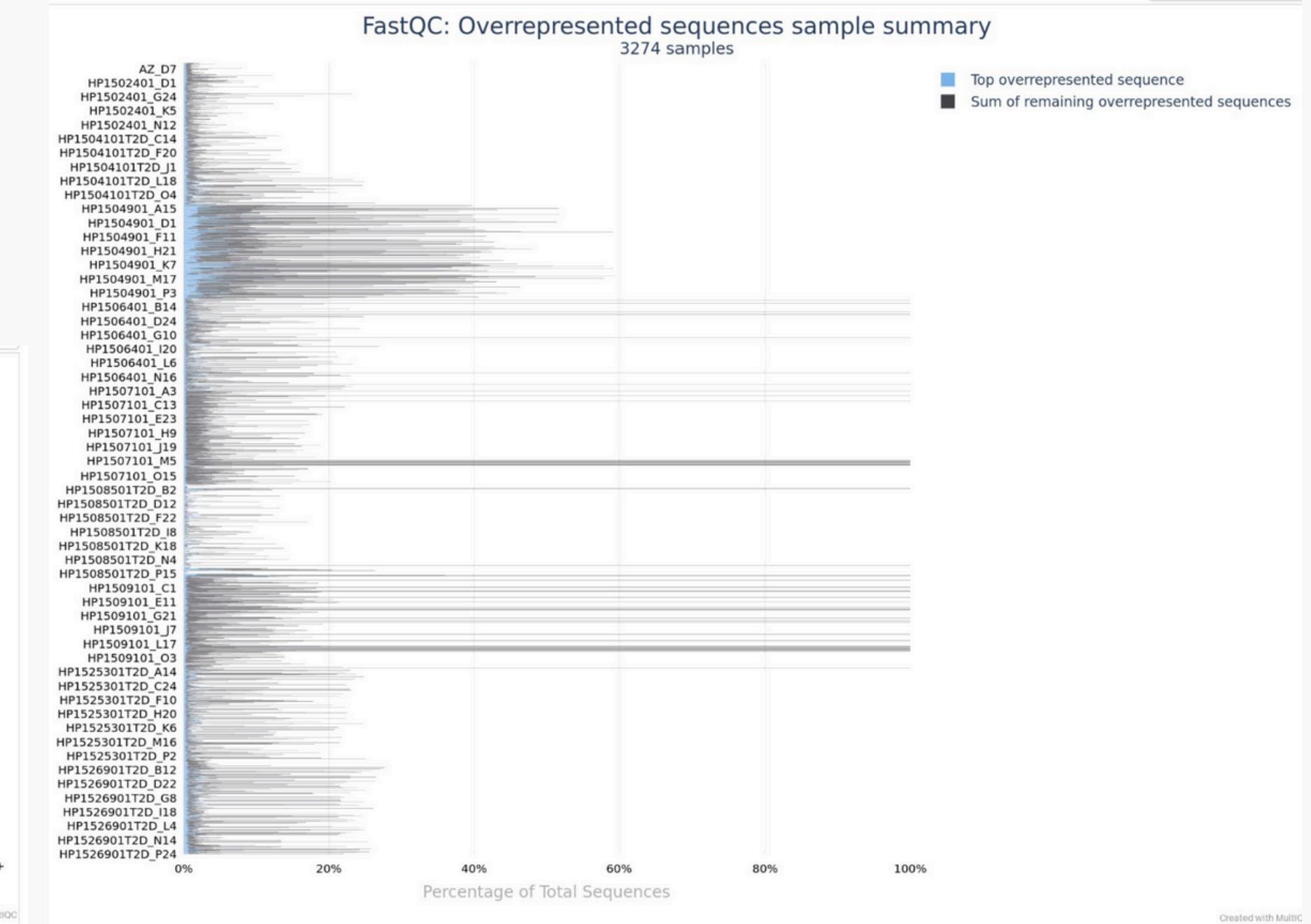


Fig. 2.4 Overrepresented sequences by sample

Comparison of β -cells by INS expression

Scatter plots display expression of endocrine hormones within β -cells, with emphasis on INS (insulin).

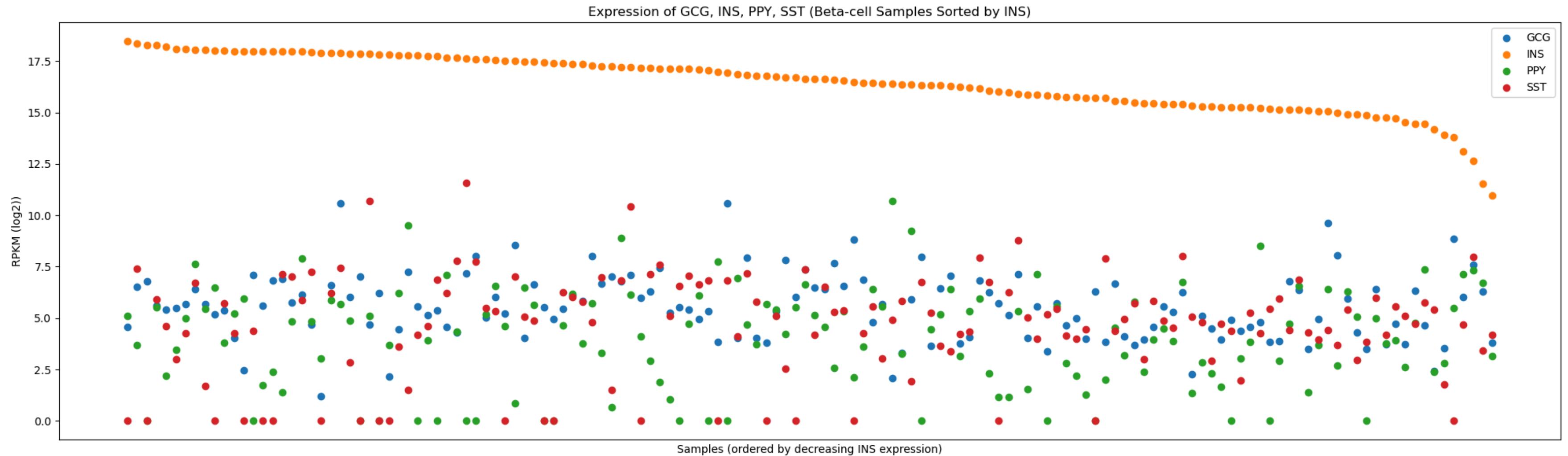


Fig. 3.1. Expression of GCG, INS, PPY, and SST in β -cells Sorted by INS Expression

Comparison of β -cells in Healthy vs T2D

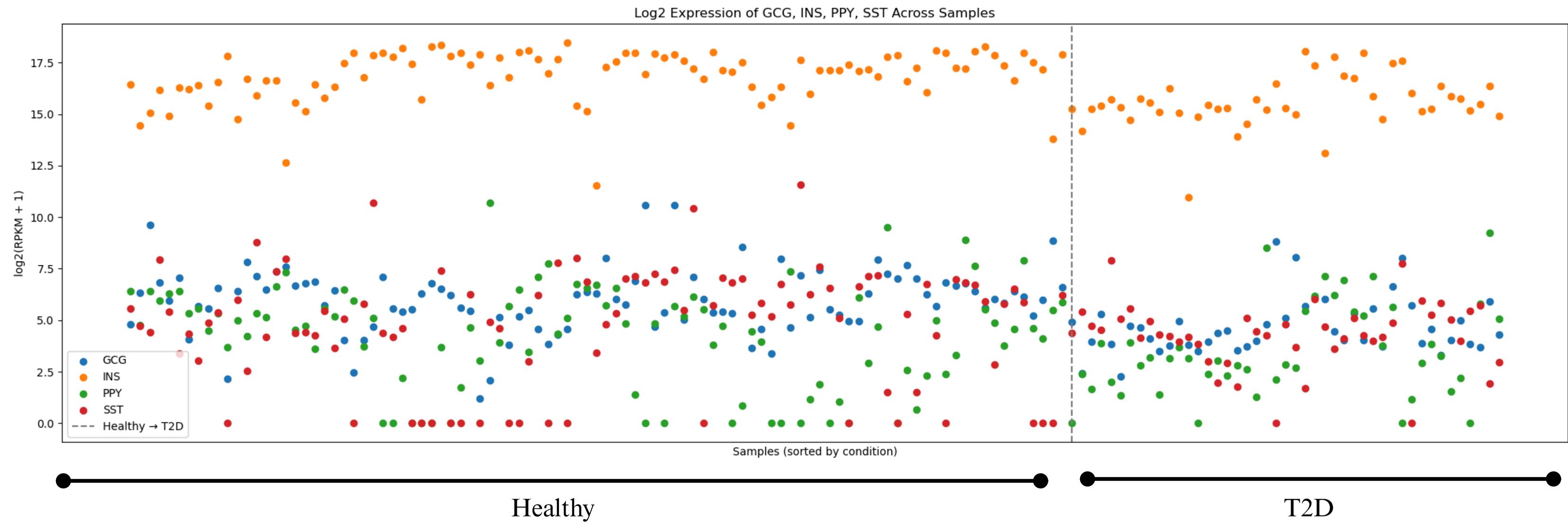


Fig. 3.2. Comparison of GCG, INS, PPY, and SST Expression in β -cells from Healthy and T2D Donors

Comparison of β -cells in Healthy vs T2D

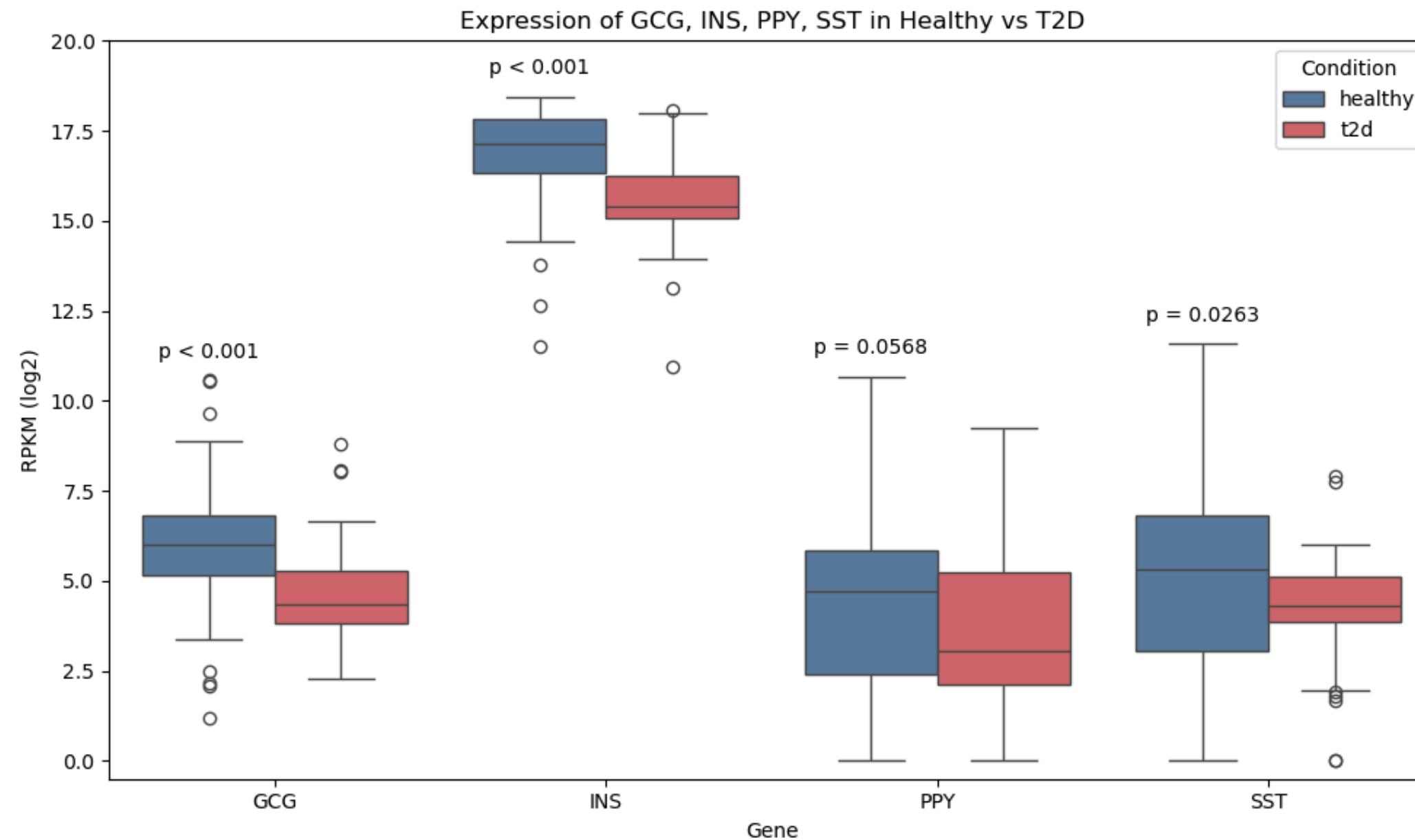


Fig. 4. Boxplot comparison of GCG, INS, PPY, and SST Expression in β -cells from Healthy and T2D Donors

Normalization Effect on Gene Expression Data (PCA Plot)

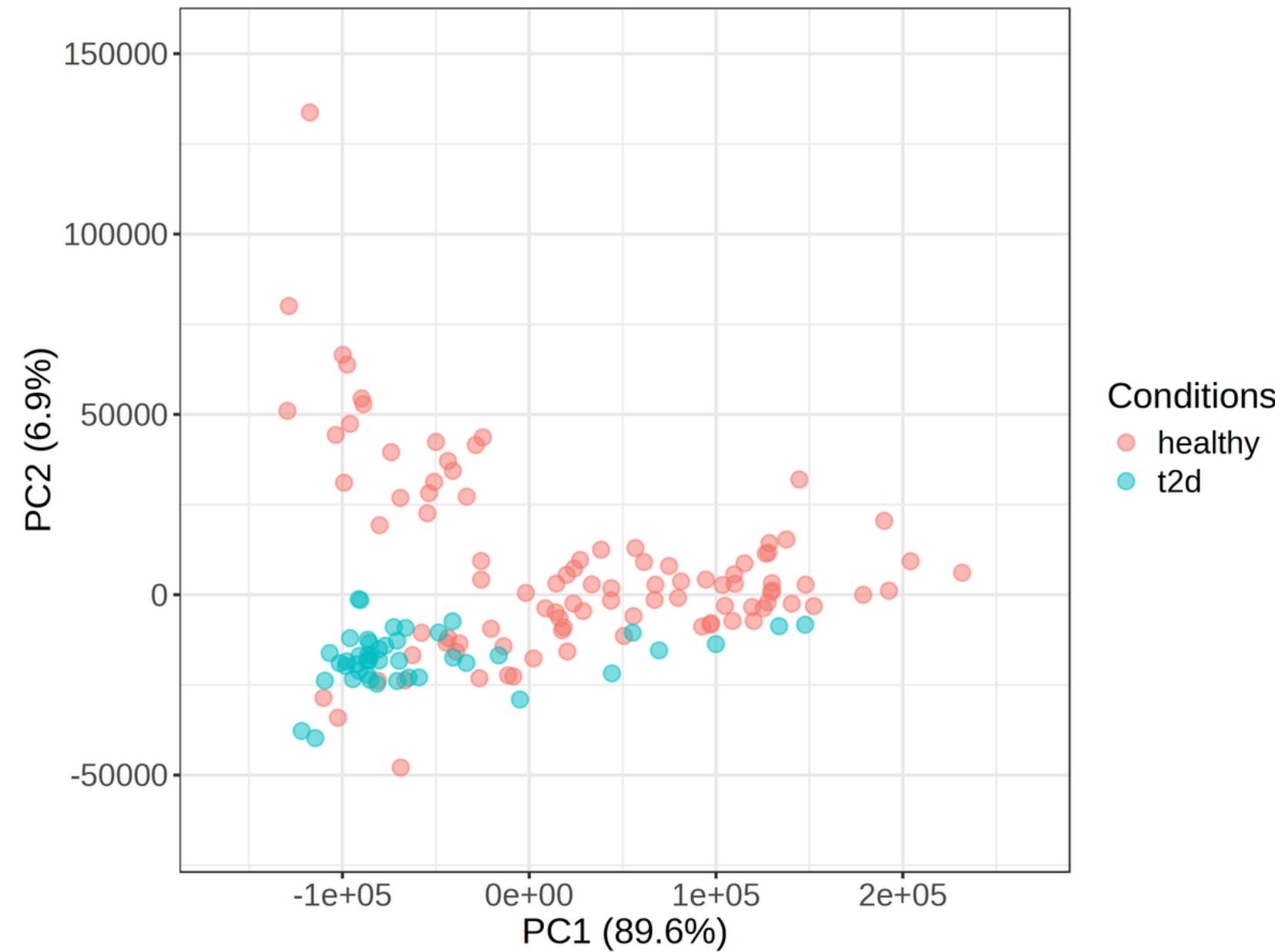


Fig. 7.1. Pre-Normalization

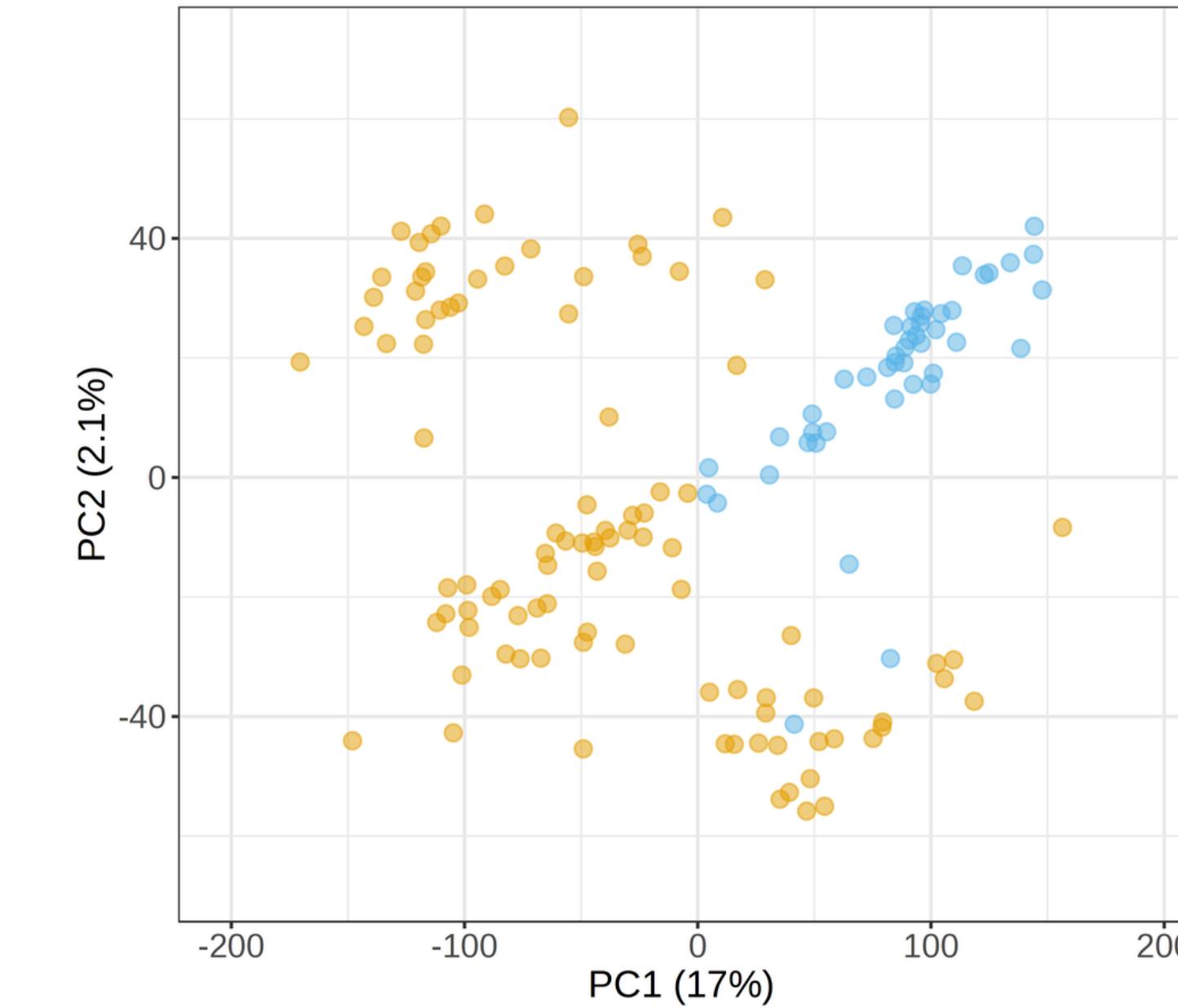


Fig. 7.2. Post-Normalization

t-SNE Visualization of β -Cells by Condition and Donor

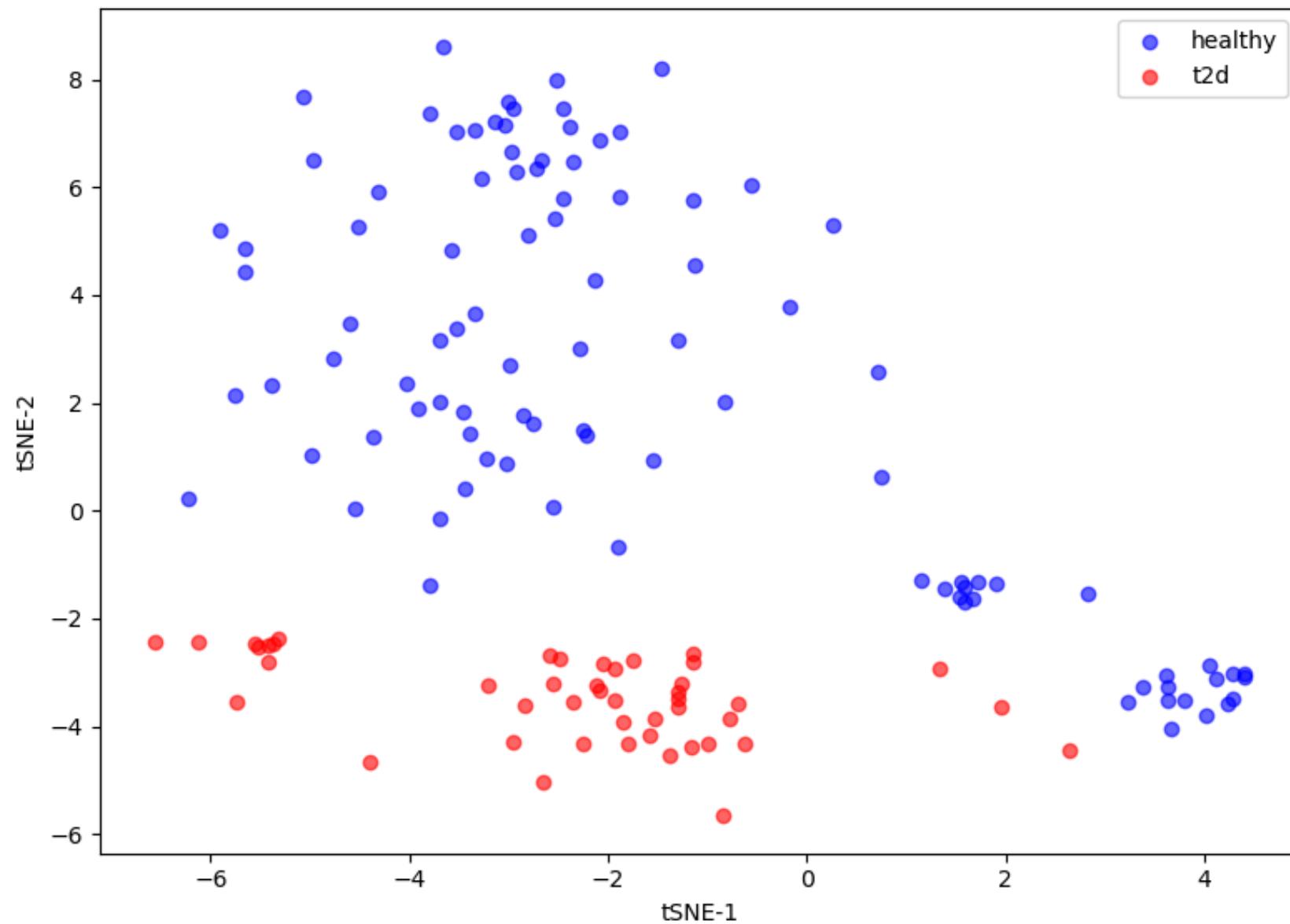


Fig. 6.1 Colored by condition

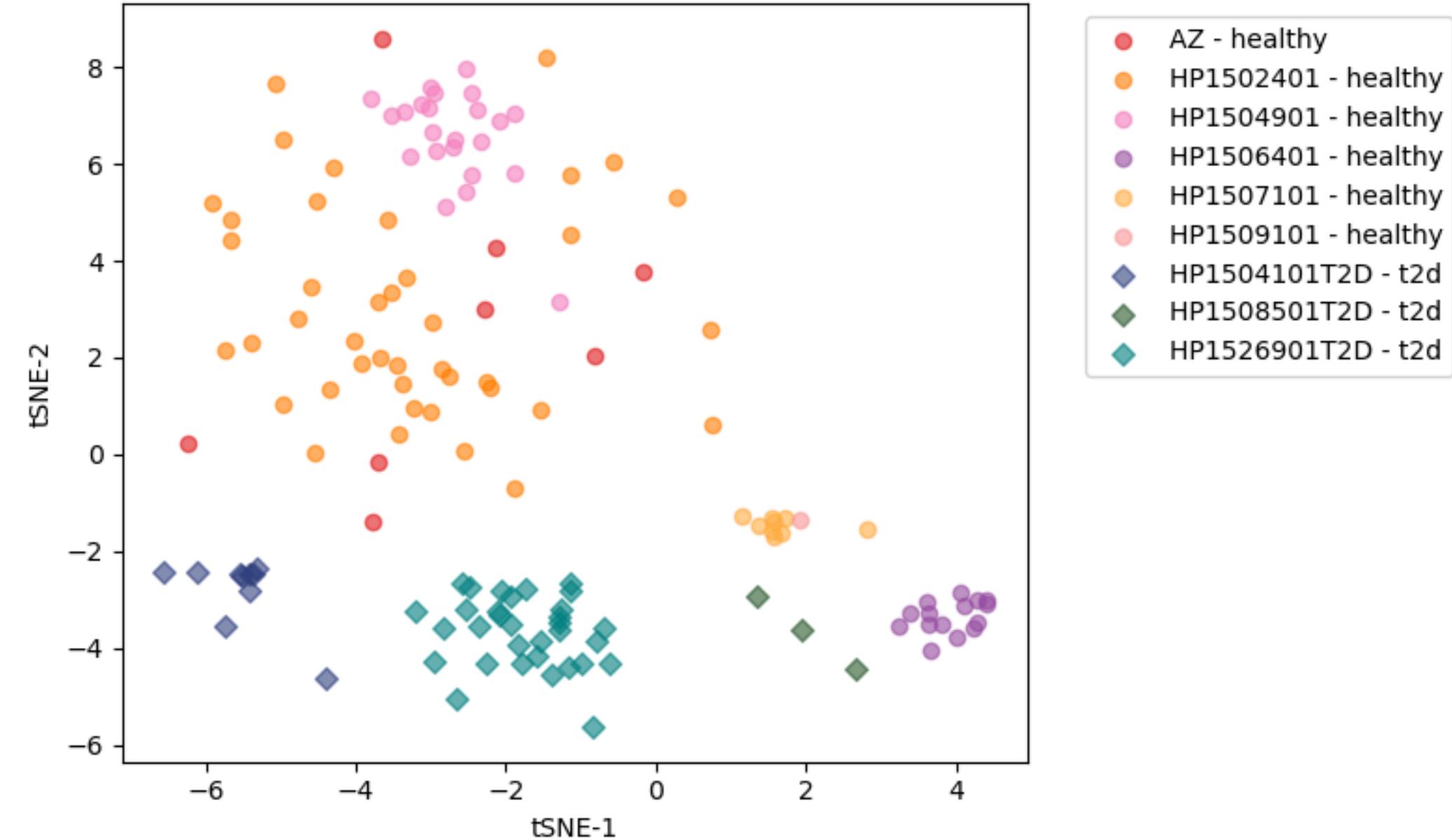


Fig. 6.2 Colored by Donor

t-SNE representation of differentially expressed genes

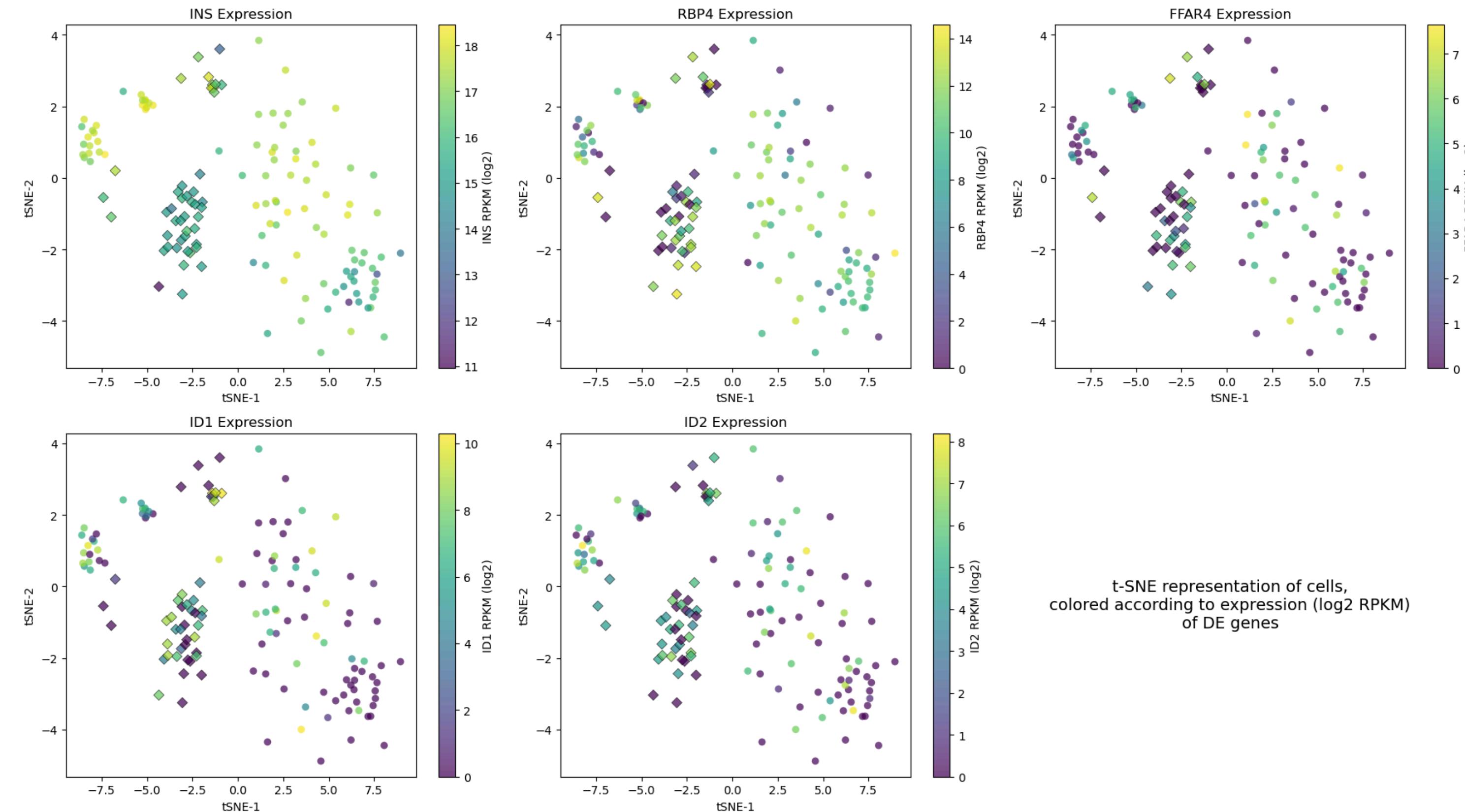
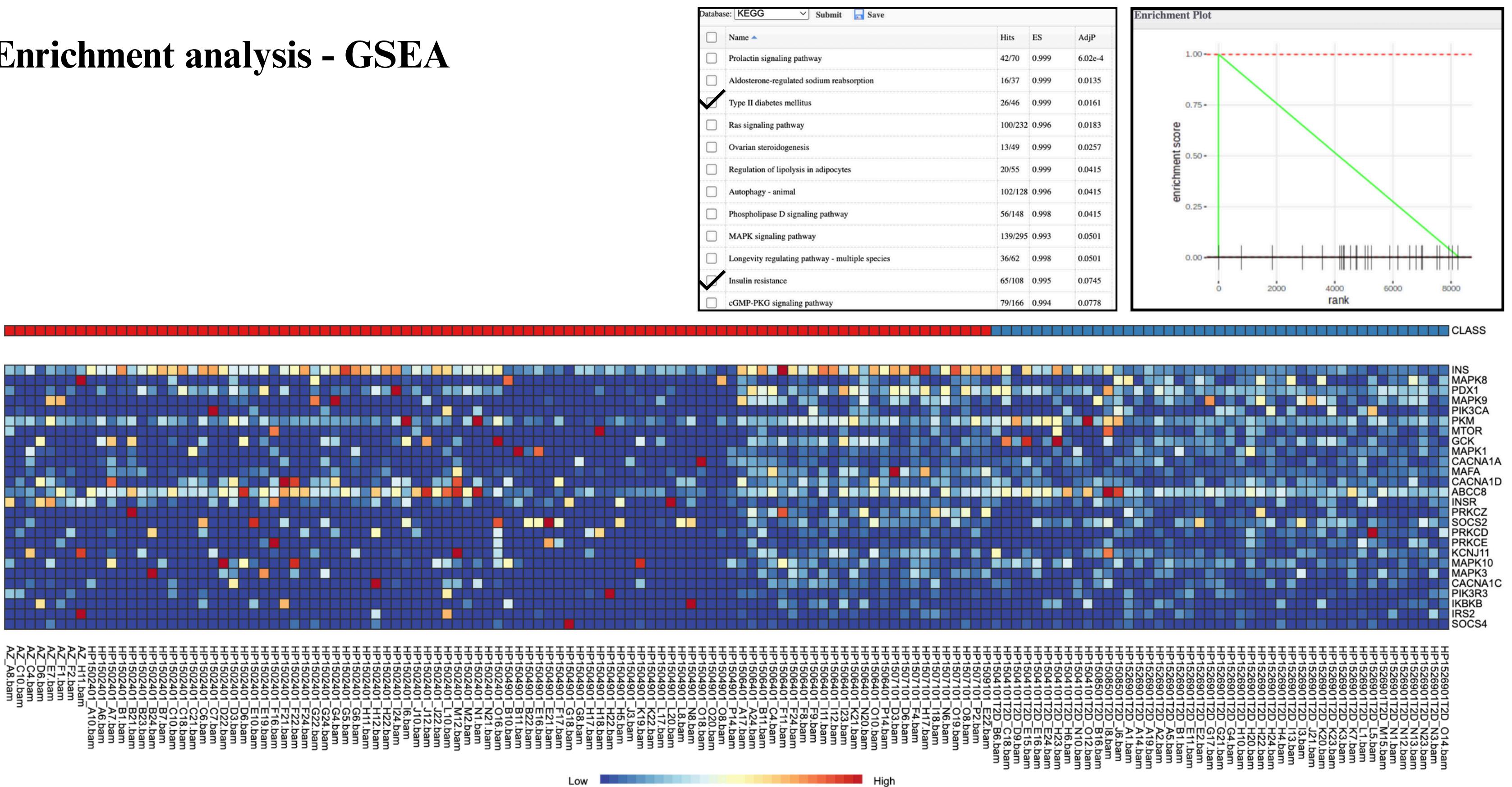


Fig. 8. t-SNE representation of beta cells among five clusters

Enrichment analysis - GSEA



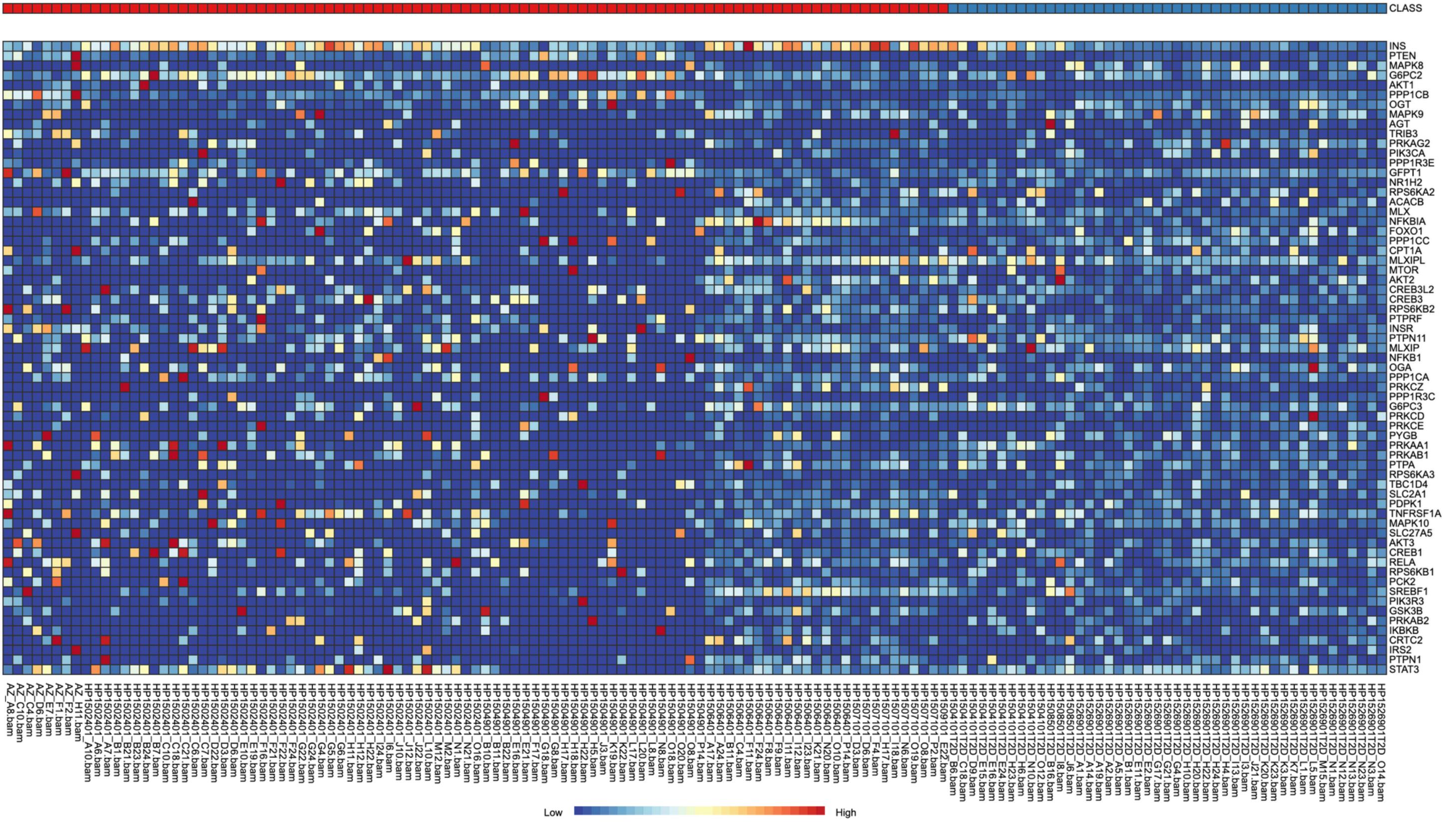


Fig. 9.2. GSEA enrichment analysis of insulin resistance β -cells

Predictive genes linked to T2D

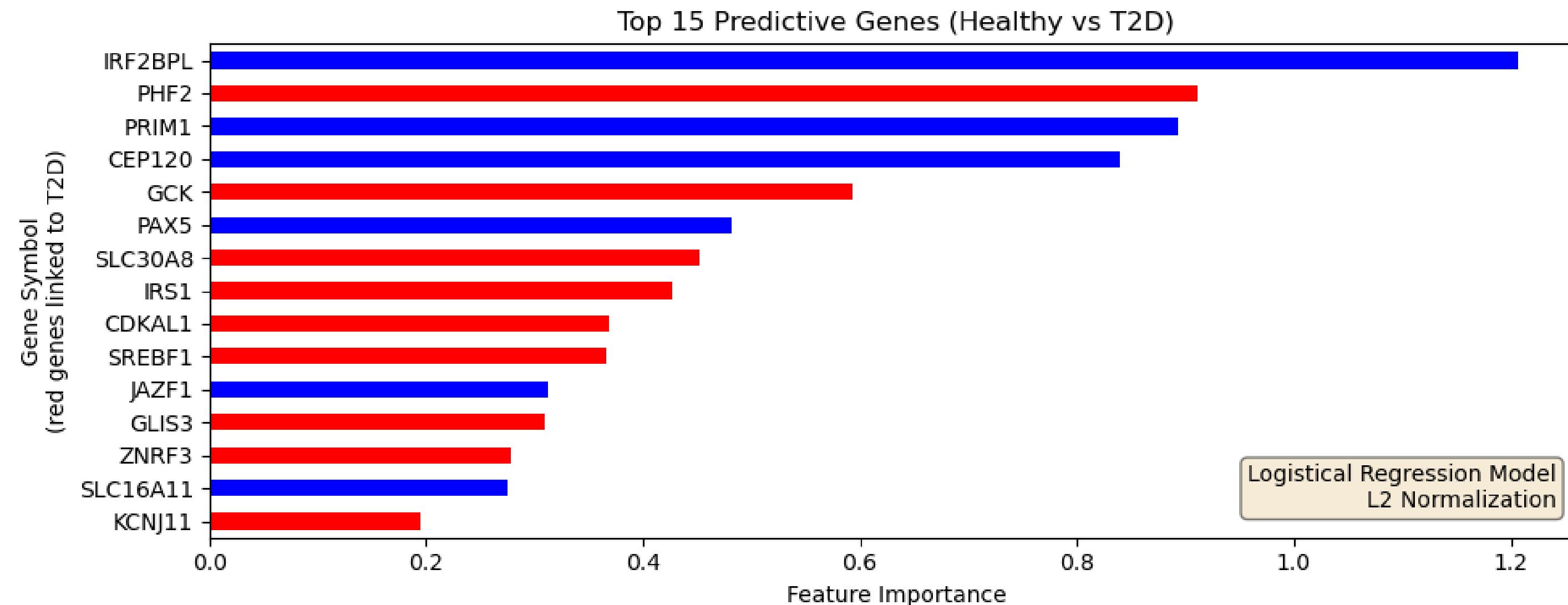


Fig. 10. Predictive genes linked to T2D

Key Findings

- scRNA-seq revealed **unique gene expression changes** in β -cells from T2D donors, compared to healthy individuals
 - t-SNE clustering combined with high **INS expression** was used to define β -cells, revealing subpopulations including RBP4-expressing clusters that suggest functional heterogeneity within the β -cell population.
 - Essential β -cell identity genes such as **PDX1 and INS** were **downregulated** in T2D donors, supporting the idea of β -cell dedifferentiation or loss of identity.
 - We also predicted genes related to T2D using logistic regression namely **PHF2, GCK**.
-

Retrospective analysis

What Went Well

Challenges Faced

Areas for
Improvement

Future Directions

CONCLUSION

- Built and executed a full scRNA-seq analysis pipeline.
- Found key gene changes in T2D β -cells: ↓ **INS, PDX1**; ↑ GPD2, LEPROT1.
- Detected **β -cell heterogeneity**, including RBP4+ subpopulations.
- GSEA showed enrichment of stress and metabolic pathways in T2D.
- Results support **β -cell identity loss and dysfunction** in T2D

References

- Segerstolpe, Å., Palasantza, A., Eliasson, P., Andersson, E., Andréasson, A., Sun, X., ... & Sandberg, R. (2016). Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. *Cell Metabolism*, 24(4), 593-607.
<https://doi.org/10.1016/j.cmet.2016.08.020>

Submissions

1. [MultiQC report](#)
2. [Pipeline](#)
3. [Genes linked to T2D](#)

Miscellaneous

1. Workflow of Segerstolpe, et al. 2016
2. Reproduced graphs

1. Workflow of Segerstolpe, et al. 2016

	Tools
Step 1: Tissue acquisition of healthy and T2D	Prodo Laboraties Inc.
Step 2: Cell dissociation and sorting	FACS (Flourescence-activated cell sorting)
Step 3: cDNA library generation	Smartseq2 protocol
Step 4: Sequencing	Illumina HiSeq 2000
Step 5: Read quantification	rpkmforgenes
Step 6: Gene expression analysis	t-SNE, ANOVA
Step 7: Dimesionality reduction	t-SNE
Step 8: Cell type identification	t-SNE
Step 9: Cell composition analysis	Immunohistochemistry, FACS, BioPix software

2. Reproduced graphs

Fig. 3.1. in slide 10 matches exactly with Figure S2(C) beta cells

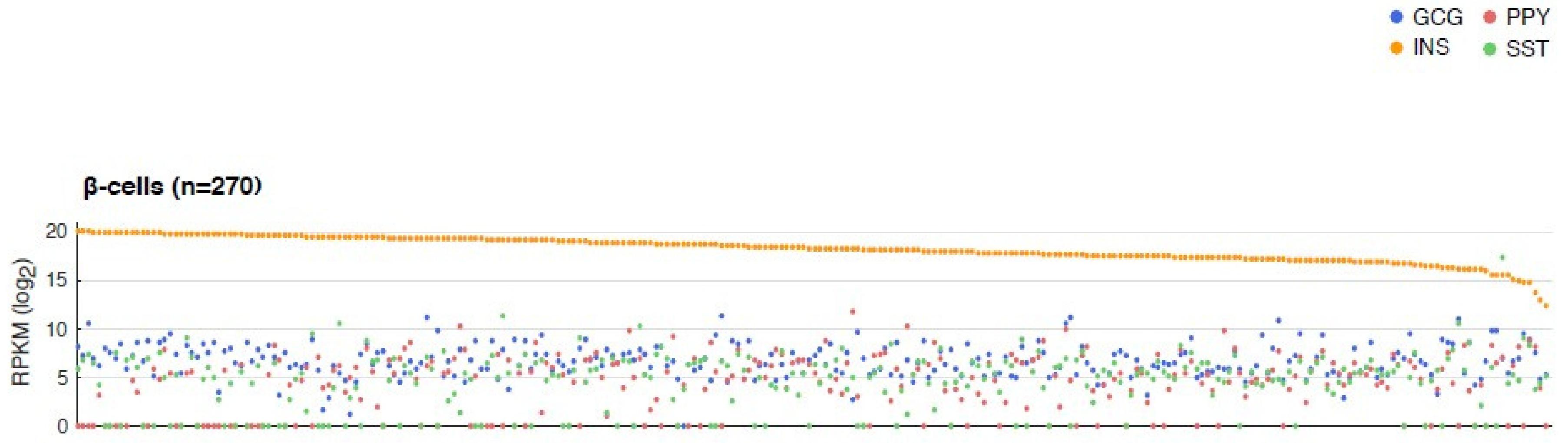


Fig 6.1., 6.2. can be matched with Figure 1(E) maybe we can tell since this is a more focused version of beta cells, we got a graph similar to 1(E) - the shape maybe.

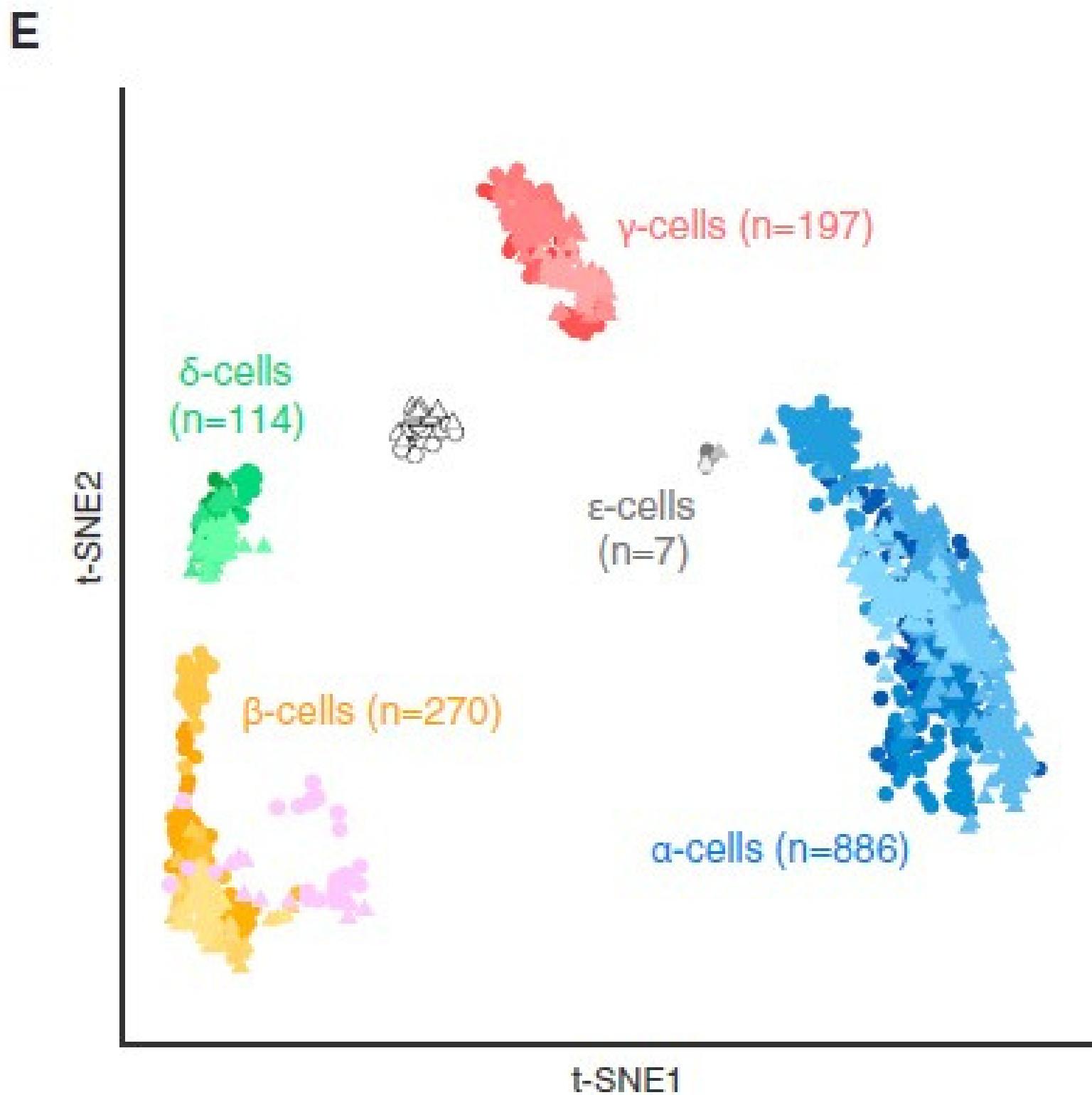
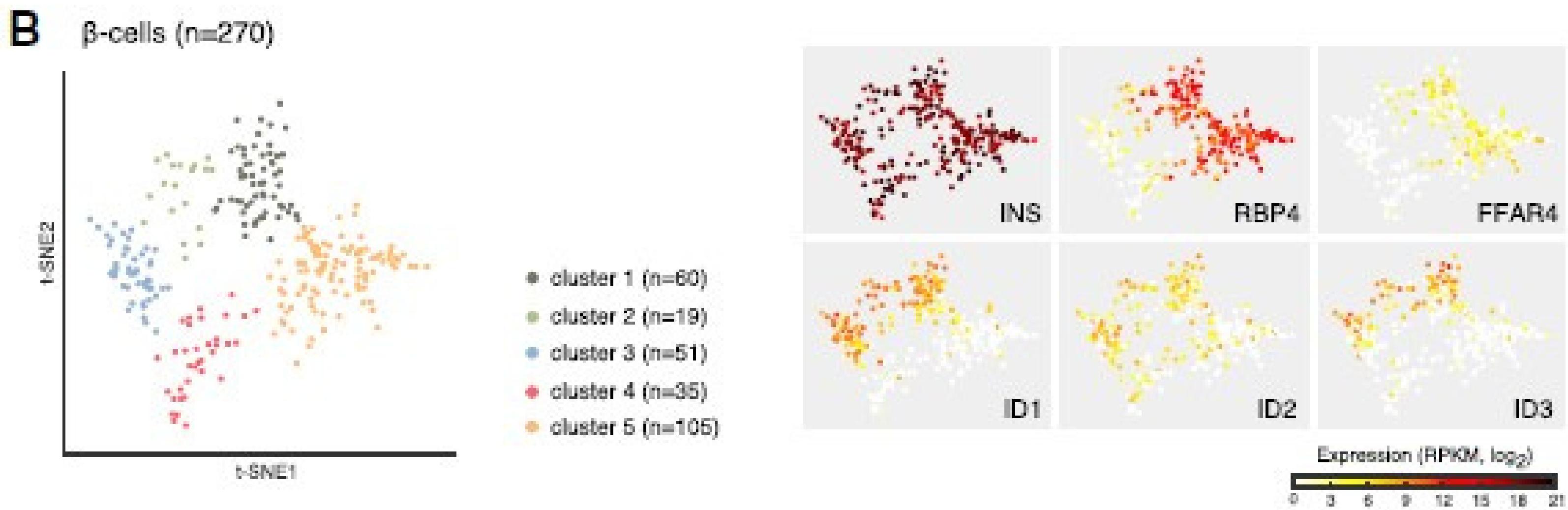


Fig 8. can be matched with Figure 4(B) - The cells are clustered irrespective of condition to demonstrate gene expression levels differing by cluster.



THANK YOU!
