NanopoReaTA: a user-friendly tool for nanopore-seq real-time transcriptional analysis

Anna Wierczeiko[1,#], Stefan Pastore[2,#], Vincent Dietrich[1], Stefan Mündnich[2], Mark Helm[2], Tamer Butto[2*] and Susanne Gerber[1,*]

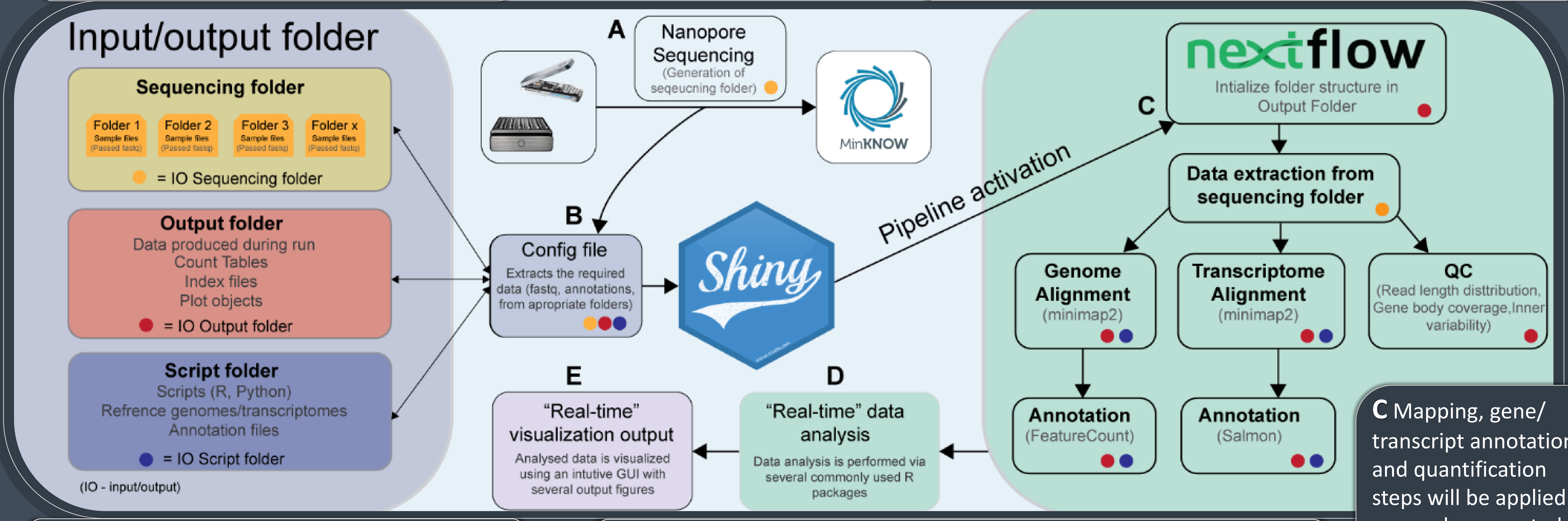
[1]Institute of Human Genetics, University Medical Center of the Johannes Gutenberg University Mainz, 55131 Mainz, Germany,
[2]Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, 55128 Mainz, Germany
[#] Authors contributed equally ; [*]corresponding authors

Background

Oxford Nanopore Technologies' (ONT) sequencing platform offers an excellent opportunity to perform real-time analysis during sequencing. This feature allows for early insights into experimental data and accelerates a potential decision-making process for further analysis. Here we introduce NanopoReaTA, a user-friendly real-time analysis toolbox for RNA sequencing data from ONT. Sequencing results from a running or finished experiment are processed through an R Shiny-based graphical user interface (GUI) with an integrated Nextflow pipeline for whole transcriptome and gene-specific analyses. NanopoReaTA provides visual snapshots of analysis results in progress, thus enabling interactive sequencing and rapid decision-making. The application can be installed via conda or docker and can be found on github: https://github.com/AnWiercze/NanopoReaTA.

A NanopoReaTA can direclty process actively generated ouput from MinKNOW.

B The user creates a configuration file within the GUI that gathers relevant data paths and parameters for preprocessing.

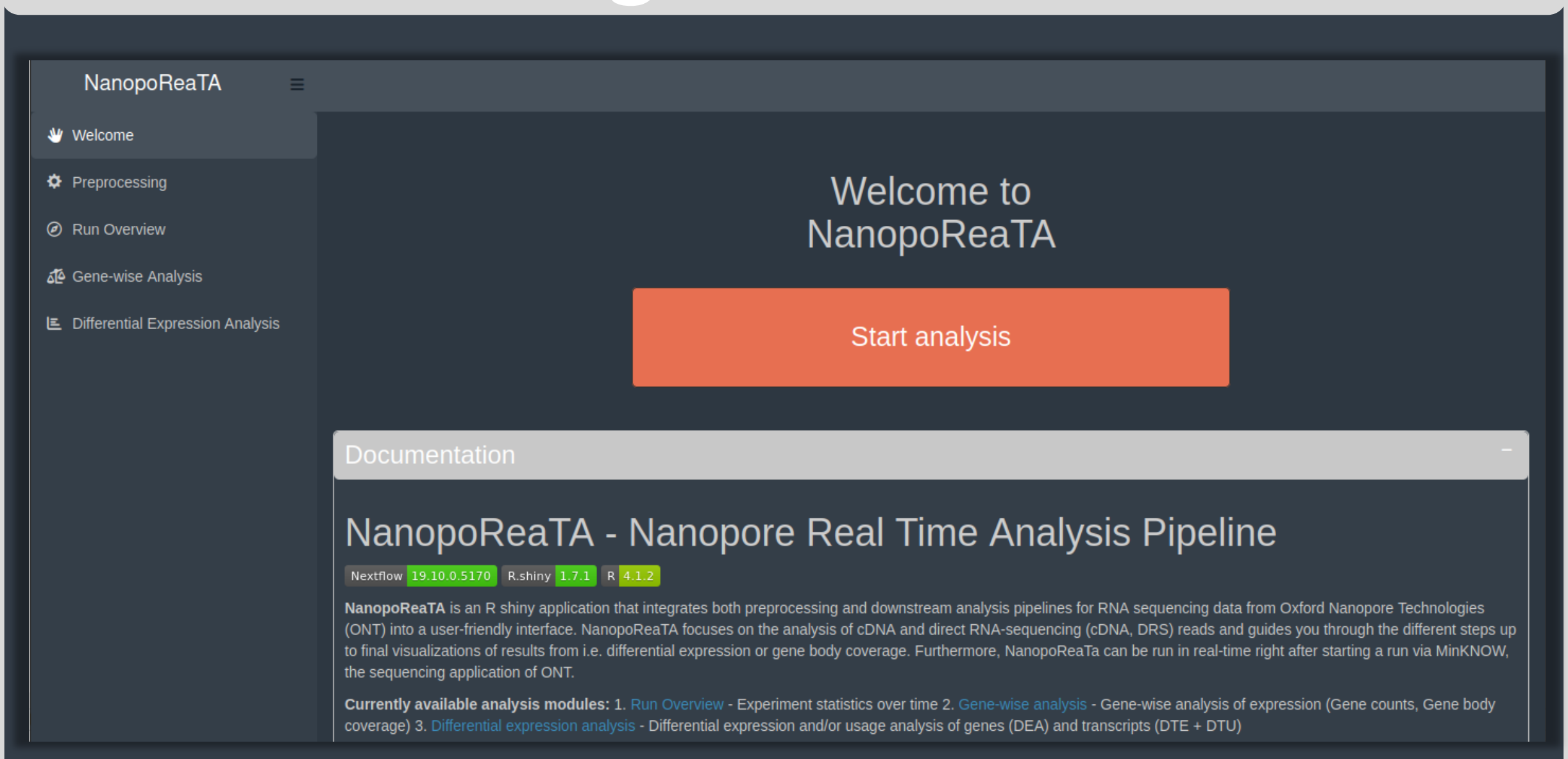


C Mapping, gene/ transcript annotation and quantification steps will be applied on newly generated data repetitively in batches (= iterations).

E Every generated plot can be downloaded within the GUI.

D The user can analyse the data summarized data after each preprocessing iteration using the stop and go functionality within the GUI.

Welcome Page



NanopoReaTA - Nanopore Real Time Analysis Pipeline

NanopoReaTA is an R shiny application that integrates both preprocessing and downstream analysis pipelines for RNA sequencing data from Oxford Nanopore Technologies (ONT) into a user-friendly interface. NanopoReaTA focuses on the analysis of cDNA and direct RNA-sequencing (cDNA, DRS) reads and guides you through the different steps up to final visualizations of results from i.e. differential expression or gene body coverage. Furthermore, NanopoReaTa can be run in real-time right after starting a run via MinKNOW, the sequencing application of ONT.

Currently available analysis modules: 1. Run Overview - Experiment statistics over time 2. Gene-wise analysis - Gene-wise analysis of expression (Gene counts, Gene body coverage) 3. Differential expression analysis - Differential expression and/or usage analysis of genes (DEA) and transcripts (DTE + DTU)

Input Data

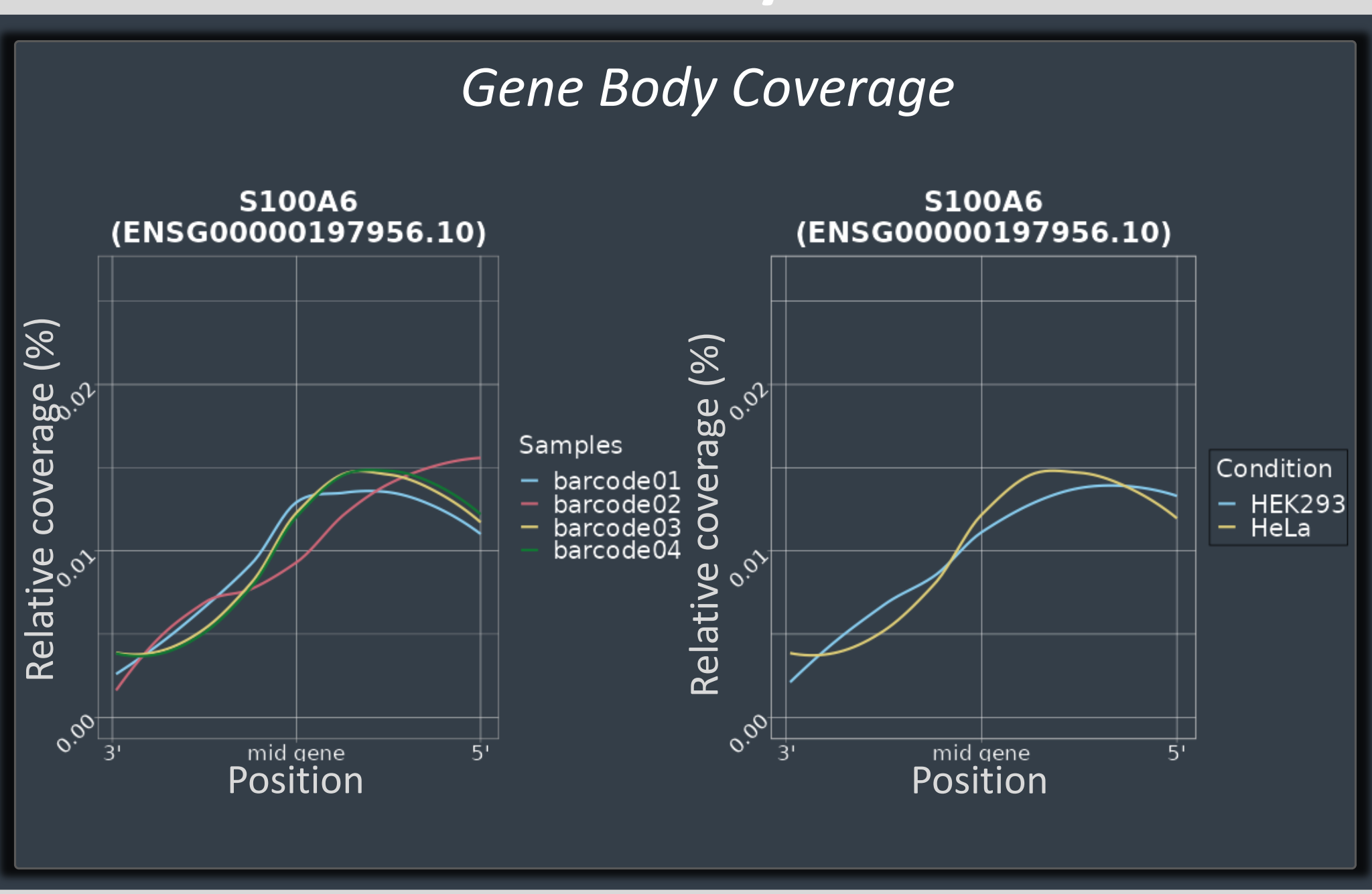
Samples:
- HEK293 cells
- HeLa cells

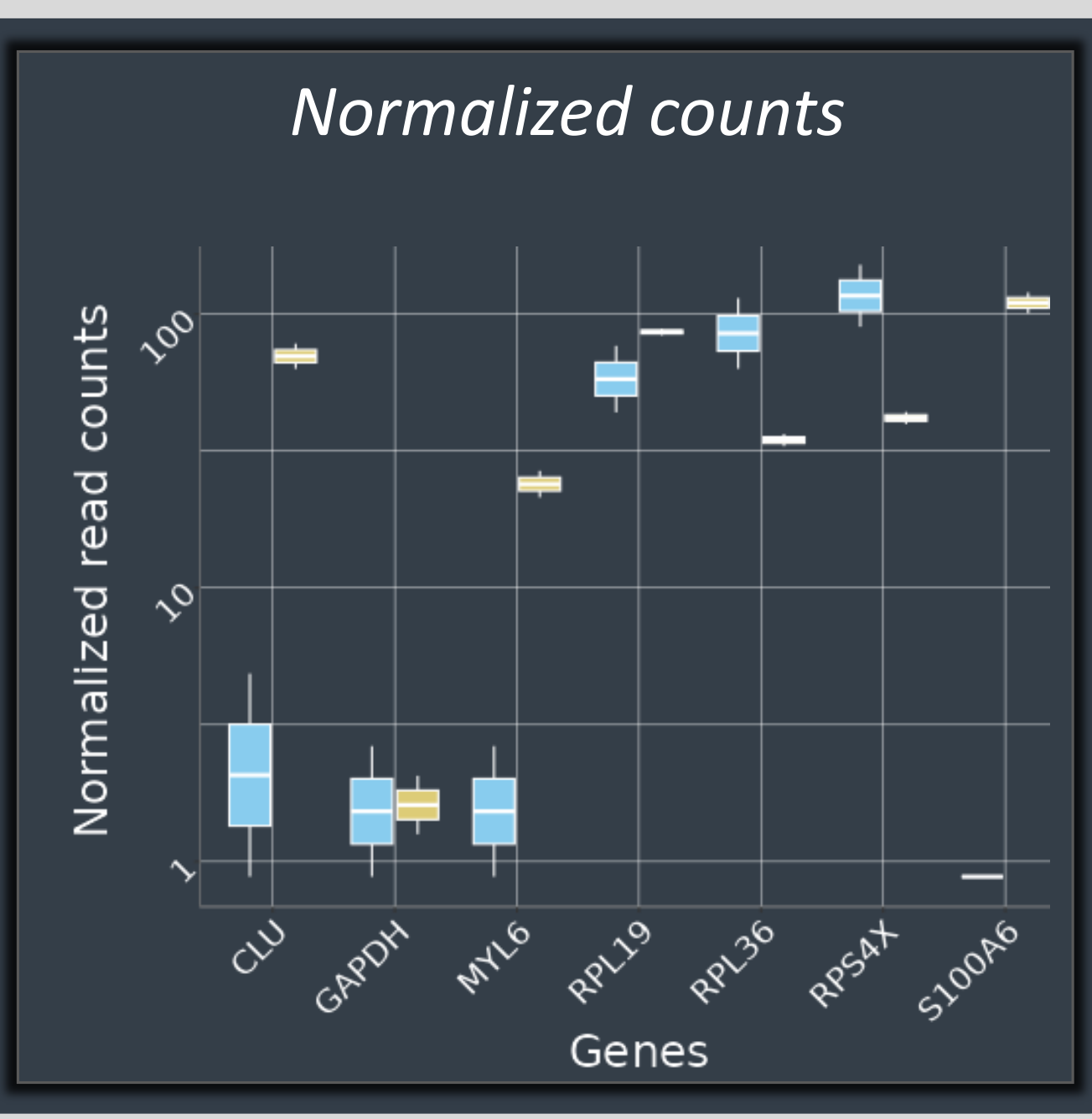
Device:
- MinION

Kit:
- Direct cDNA sequencing + native barcoding

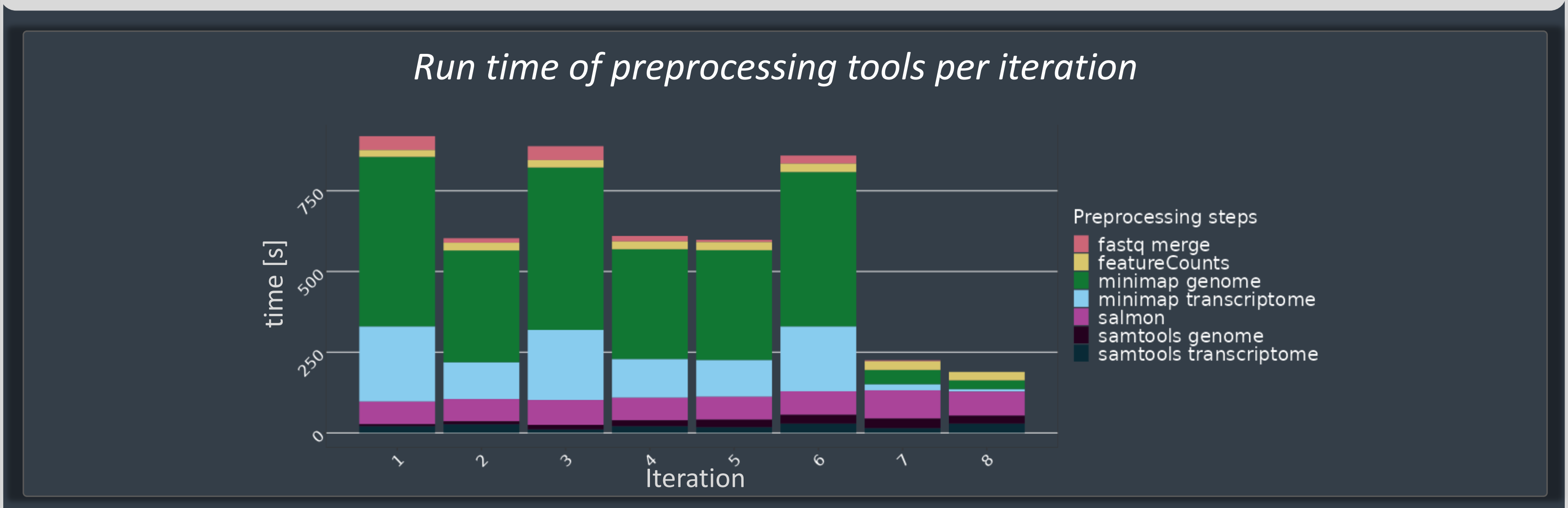
SQK-DCS109 & EXP-NBD104

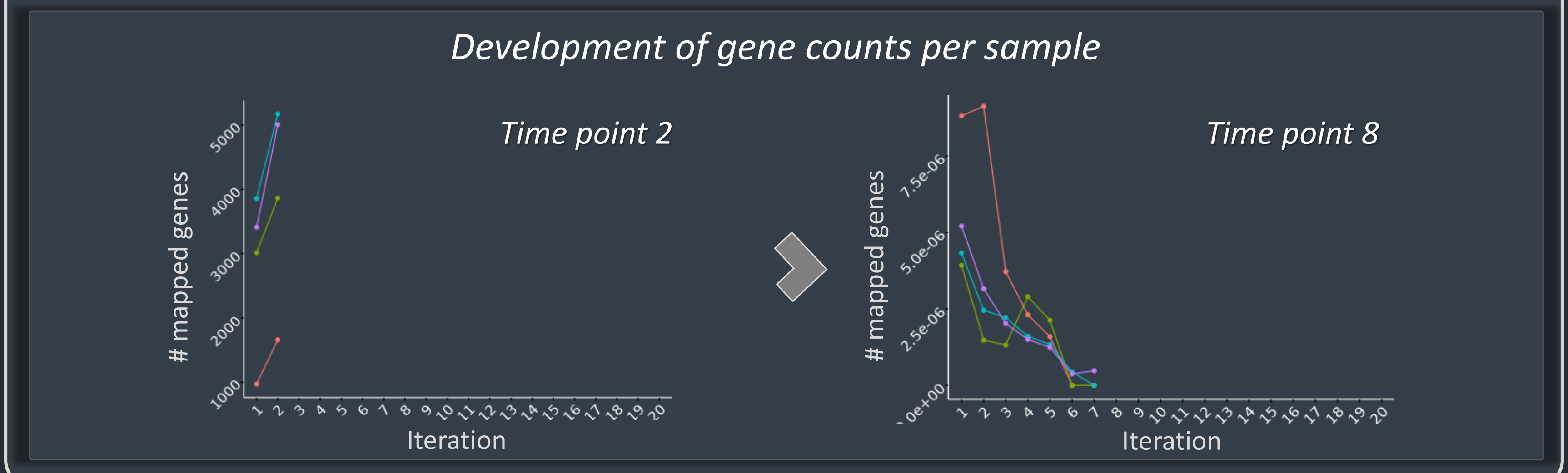
Gene-wise analysis


Gene Body Coverage


Normalized counts

Run overview


Run time of preprocessing tools per iteration

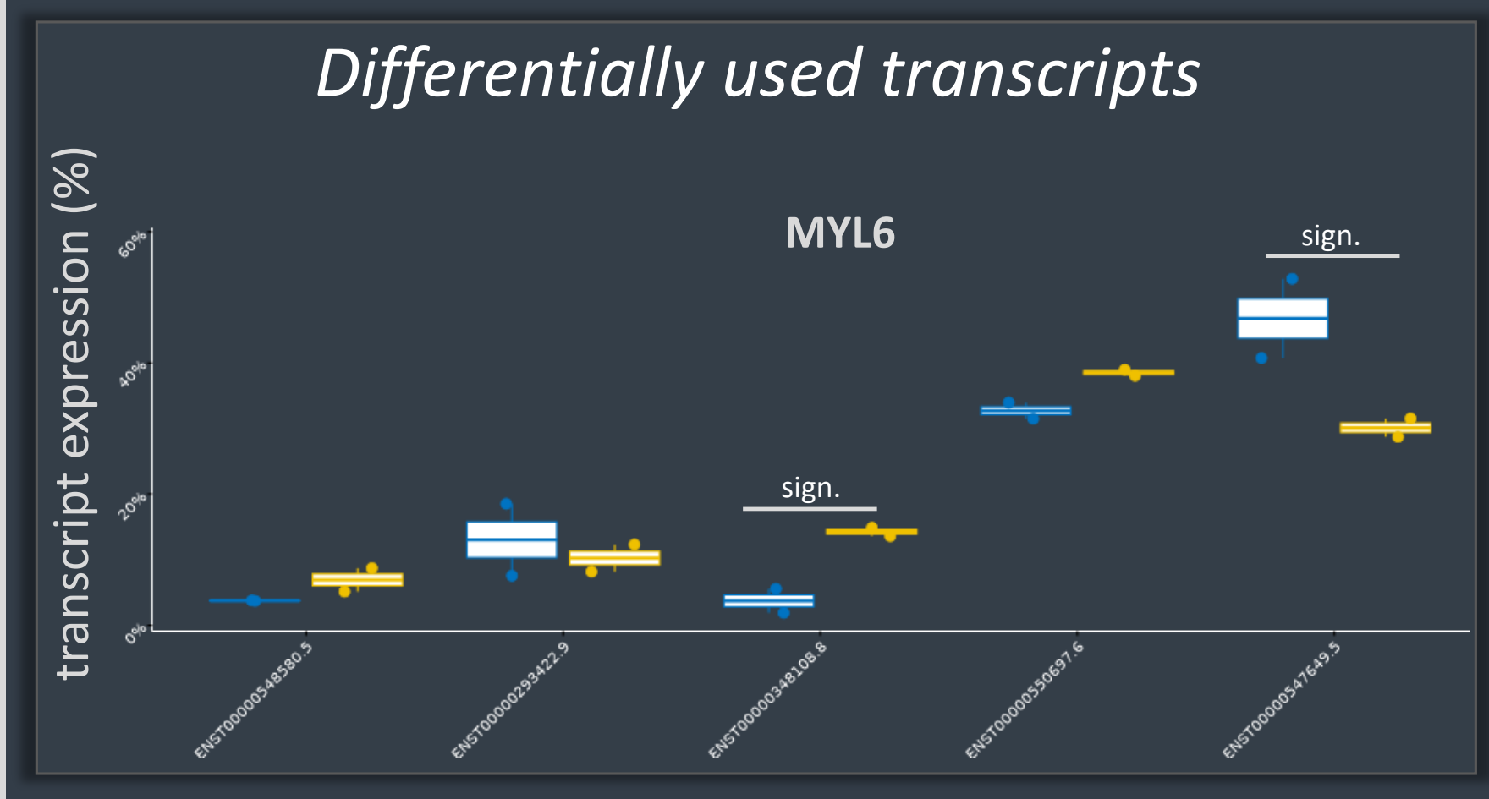

Development of gene counts per sample

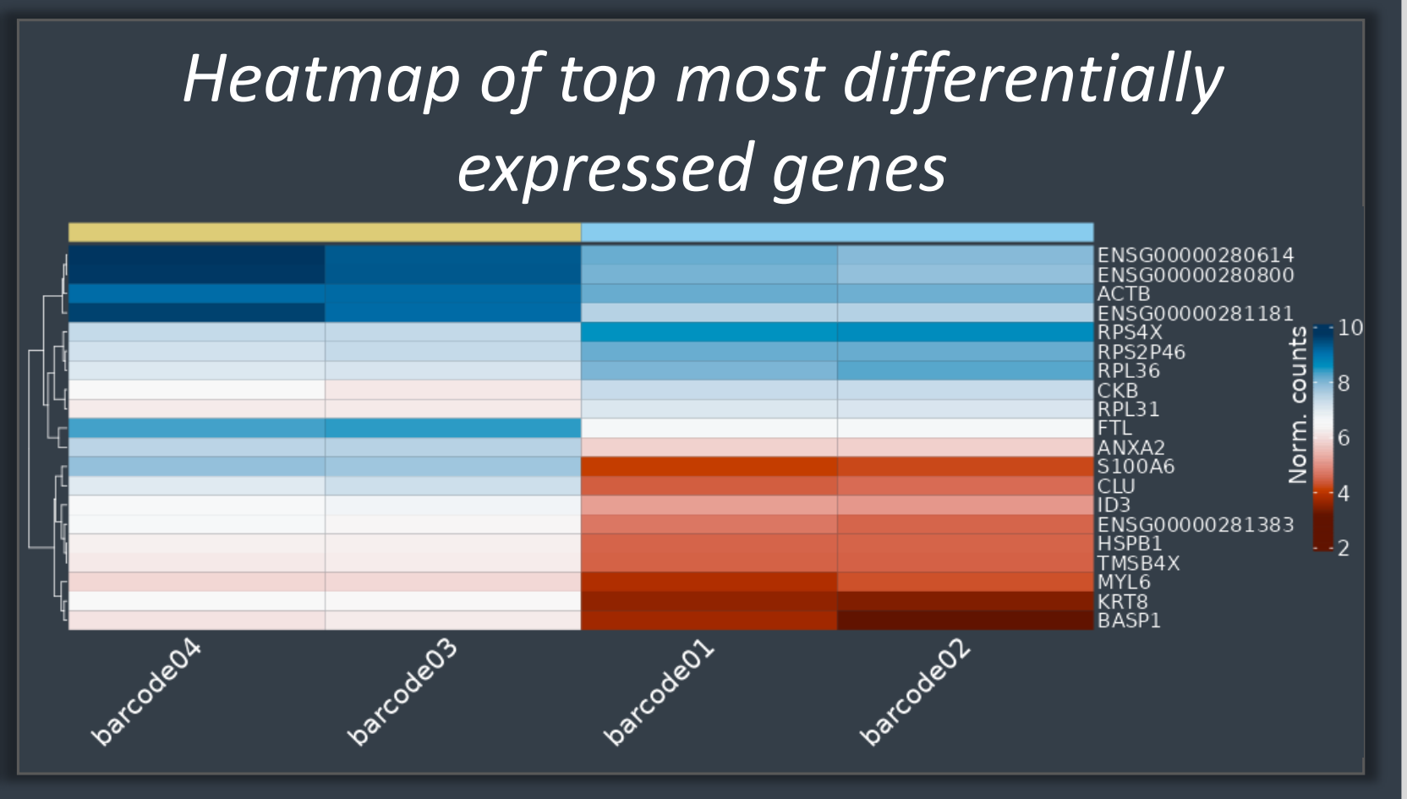
Differential Expression Analysis

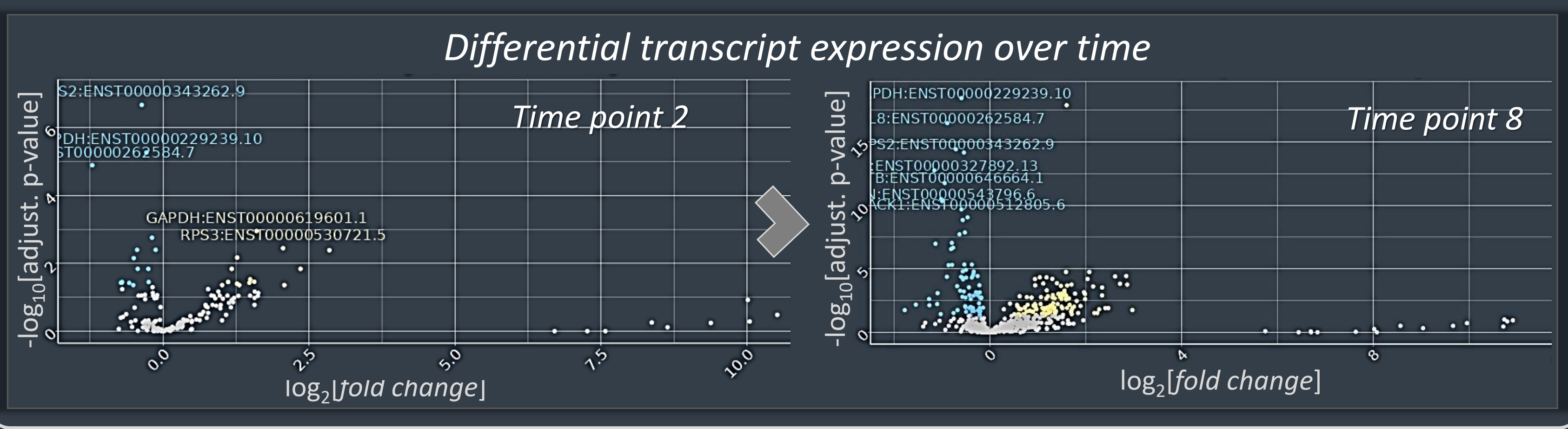
1. Differential gene expression

2. Differential transcript expression

3. Differential transcript usage


Differentially used transcripts


Heatmap of top most differentially expressed genes


Differential transcript expression over time