scLTdb tutorial

Junyao Jiang

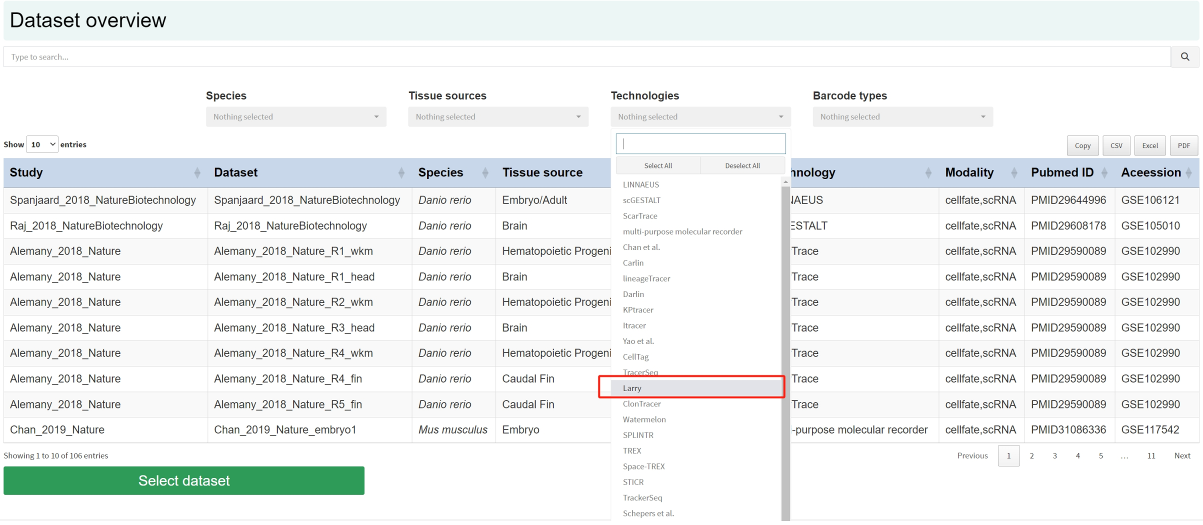
2024-06-26

Introduction

This tutorial provides users with information on how to effectively **search** for datasets, **re-analyze** and **visualize** datasets, and **download** datasets within the scLTdb. We will focusIf you have any suggestions or find bugs, please feel free to contact us.

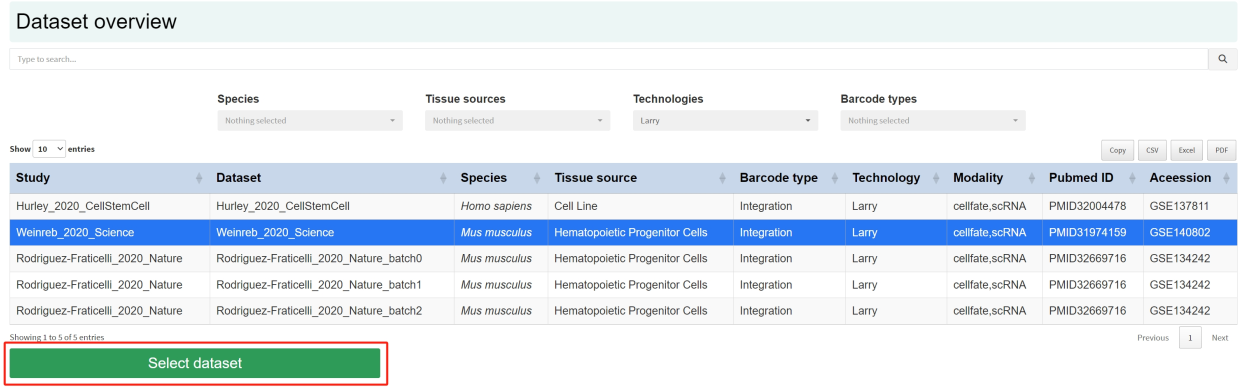
Search & querry datasets

The search page on scLTdb offers the functionality to search and query datasets based on several features, namely: 'Species', 'Tissue sources', 'Technologies', and 'Barcode types'. To illustrate, when you click on the drop-down box located under the **Technologies** feature and choose **Larry**, the interactive table will present all datasets that have **Larry** as their designated **Technologies**.



Study	Dataset	Species	Tissue source	nology	Modality	Pubmed ID	Aceession
Spanjaard_2018_NatureBiotechnology	Spanjaard_2018_NatureBiotechnology	*Danio rerio*	Embryo/Adult	AEUS	cellfate,scRNA	PMID29644996	GSE106121
Raj_2018_NatureBiotechnology	Raj_2018_NatureBiotechnology	*Danio rerio*	Brain	ESTALT	cellfate,scRNA	PMID29608178	GSE105010
Alemany_2018_Nature	Alemany_2018_Nature_R1_wkm	*Danio rerio*	Hematopoietic Progeni	Trace	cellfate,scRNA	PMID29590089	GSE102990
Alemany_2018_Nature	Alemany_2018_Nature_R1_head	*Danio rerio*	Brain	Trace	cellfate,scRNA	PMID29590089	GSE102990
Alemany_2018_Nature	Alemany_2018_Nature_R2_wkm	*Danio rerio*	Hematopoietic Progeni	Trace	cellfate,scRNA	PMID29590089	GSE102990
Alemany_2018_Nature	Alemany_2018_Nature_R3_head	*Danio rerio*	Brain	Trace	cellfate,scRNA	PMID29590089	GSE102990
Alemany_2018_Nature	Alemany_2018_Nature_R4_wkm	*Danio rerio*	Hematopoietic Progeni	Trace	cellfate,scRNA	PMID29590089	GSE102990
Alemany_2018_Nature	Alemany_2018_Nature_R4_fin	*Danio rerio*	Caudal Fin	Trace	cellfate,scRNA	PMID29590089	GSE102990
Alemany_2018_Nature	Alemany_2018_Nature_R5_fin	*Danio rerio*	Caudal Fin	Trace	cellfate,scRNA	PMID29590089	GSE102990
Chan_2019_Nature	Chan_2019_Nature_embryo1	*Mus musculus*	Embryo	-purpose molecular recorder	cellfate,scRNA	PMID31086336	GSE117542

We begin by selecting the **Weinreb_2020_Science** dataset and then click the **Select dataset** button. Subsequently, scLTdb will redirect us to the analysis result page specific to the **Weinreb_2020_Science** dataset.



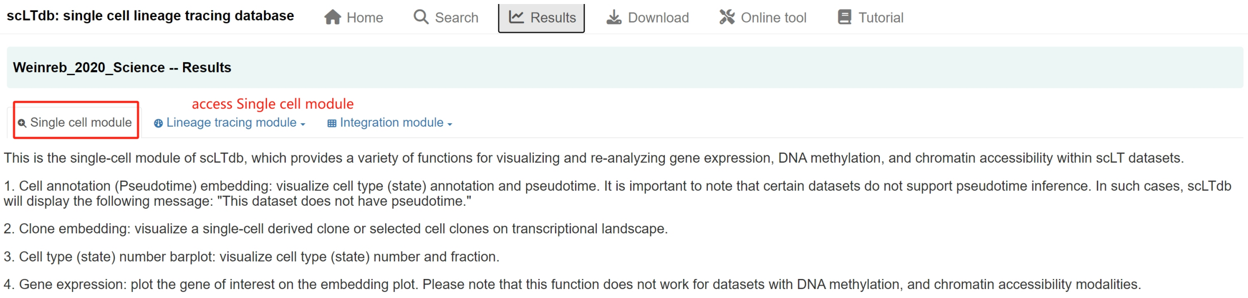
Study	Dataset	Species	Tissue source	Barcode type	Technology	Modality	Pubmed ID	Aceession
Hurley_2020_CellStemCell	Hurley_2020_CellStemCell	*Homo sapiens*	Cell Line	Integration	Larry	cellfate,scRNA	PMID32004478	GSE137811
Weinreb_2020_Science	Weinreb_2020_Science	*Mus musculus*	Hematopoietic Progenitor Cells	Integration	Larry	cellfate,scRNA	PMID31974159	GSE140802
Rodriguez-Fraticelli_2020_Nature	Rodriguez-Fraticelli_2020_Nature_batch0	*Mus musculus*	Hematopoietic Progenitor Cells	Integration	Larry	cellfate,scRNA	PMID32669716	GSE134242
Rodriguez-Fraticelli_2020_Nature	Rodriguez-Fraticelli_2020_Nature_batch1	*Mus musculus*	Hematopoietic Progenitor Cells	Integration	Larry	cellfate,scRNA	PMID32669716	GSE134242
Rodriguez-Fraticelli_2020_Nature	Rodriguez-Fraticelli_2020_Nature_batch2	*Mus musculus*	Hematopoietic Progenitor Cells	Integration	Larry	cellfate,scRNA	PMID32669716	GSE134242

Visualize & re-analyze datasets

After clicking on the **Select dataset** option, the user will be directed to the result page of the selected dataset. This result page provides three interactive modules for visualizing and analyzing single-cell lineage tracing datasets.

single cell module

The single cell module of scLTdb provides a variety of functions for visualizing and re-analyzing gene expression, DNA methylation, and chromatin accessibility within scLT datasets. Users can click the **Single cell module** button to access this module.



This is the single-cell module of scLTdb, which provides a variety of functions for visualizing and re-analyzing gene expression, DNA methylation, and chromatin accessibility within scLT datasets.

1. Cell annotation (Pseudotime) embedding: visualize cell type (state) annotation and pseudotime. It is important to note that certain datasets do not support pseudotime inference. In such cases, scLTdb will display the following message: "This dataset does not have pseudotime."

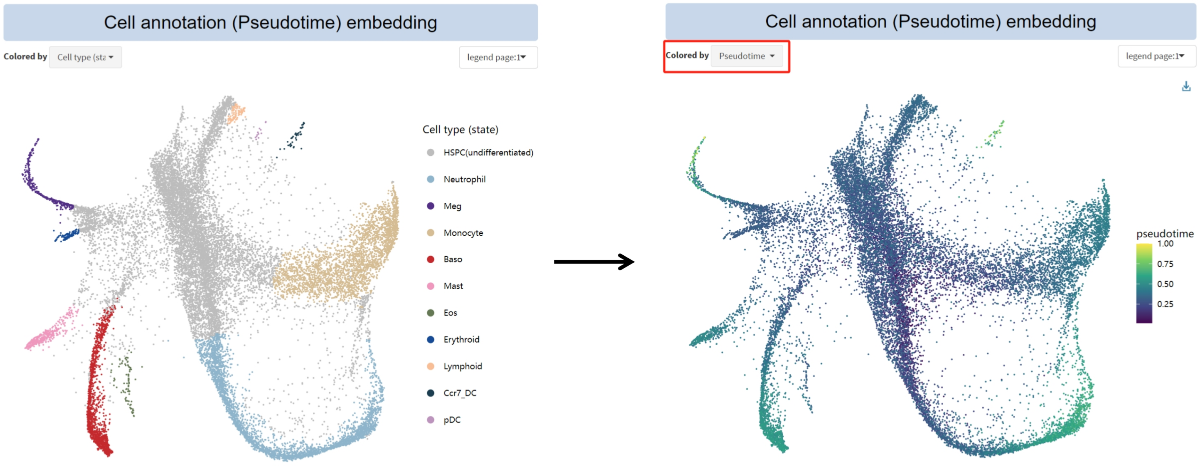
2. Clone embedding: visualize a single-cell derived clone or selected cell clones on transcriptional landscape.

3. Cell type (state) number barplot: visualize cell type (state) number and fraction.

4. Gene expression: plot the gene of interest on the embedding plot. Please note that this function does not work for datasets with DNA methylation, and chromatin accessibility modalities.

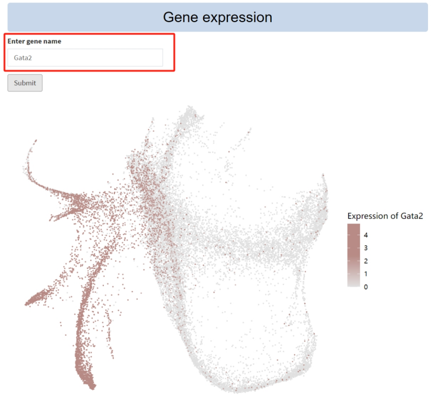
Visualize cell annotation and pseudotime

You can click **colored by** button in the **Cell annotation (Pseudotime) embedding** section to change cell color (colored by cell type or pseudotime).



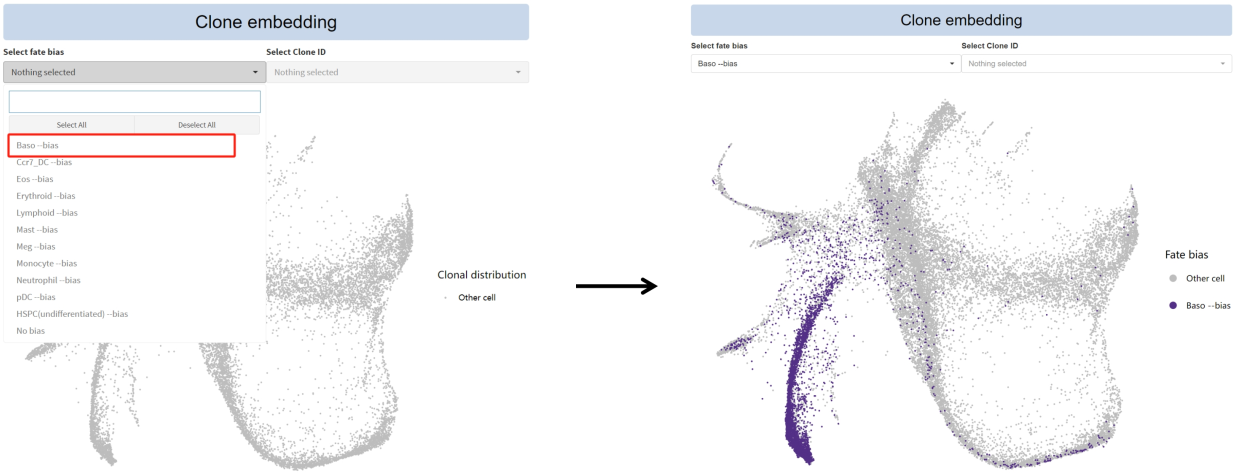
Visualize gene expression

To visualize gene expression on the embedding plot, you can input interested gene name in the **Enter gene name** box

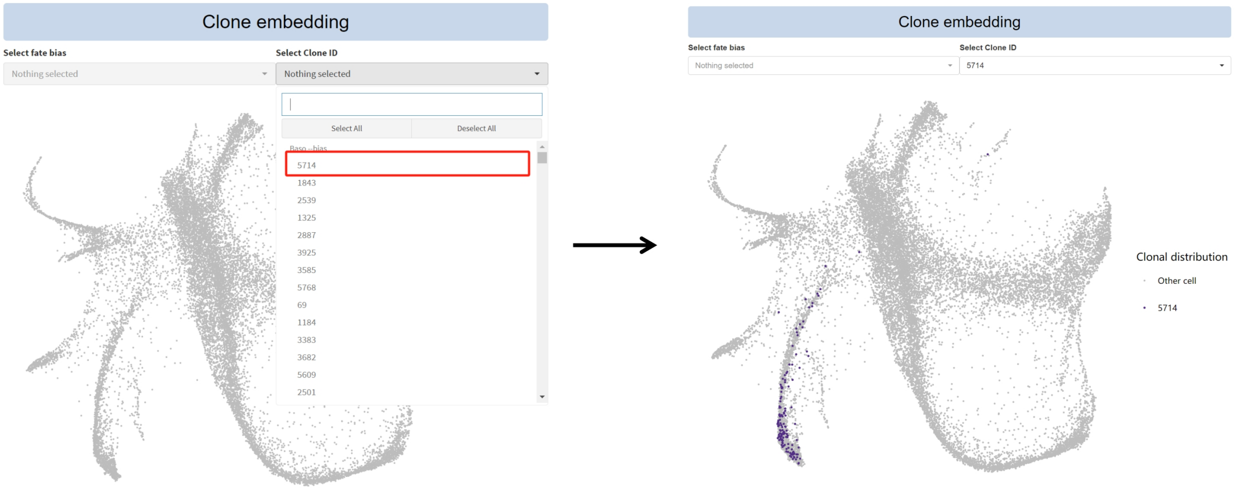


Visualize clone embedding

To visualize a set of clones with the same fate bias, you can click the **Select fate bias** button in the clone embedding section.

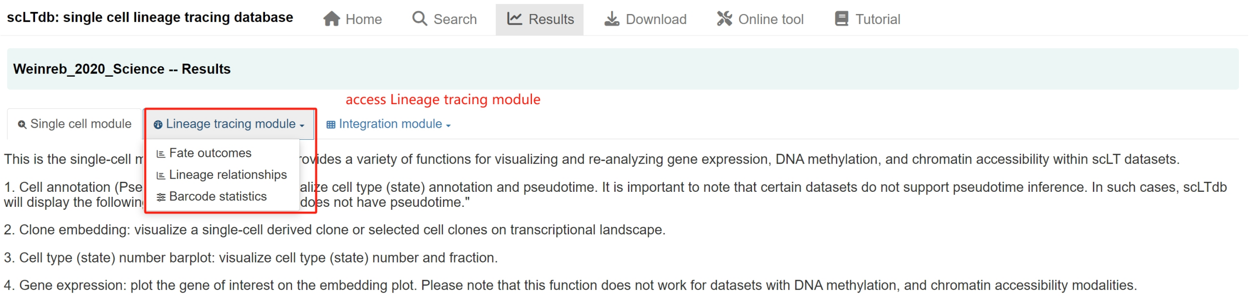


In a similar way, you can visualize single clone by clicking **Select Clone ID** button.



Lineage tracing module

The lineage tracing module of the scLTdb offers a range of functions to visualize and re-analyze lineage tracing (clone) data modalities within scLT datasets. Users can click the **Lineage tracing module** button to access this module.



Weinreb_2020_Science -- Results

This is the single-cell m rovides a variety of functions for visualizing and re-analyzing gene expression, DNA methylation, and chromatin accessibility within scLT datasets.

1. Cell annotation (Pse alize cell type (state) annotation and pseudotime. It is important to note that certain datasets do not support pseudotime inference. In such cases, scLTdb will display the followin does not have pseudotime."

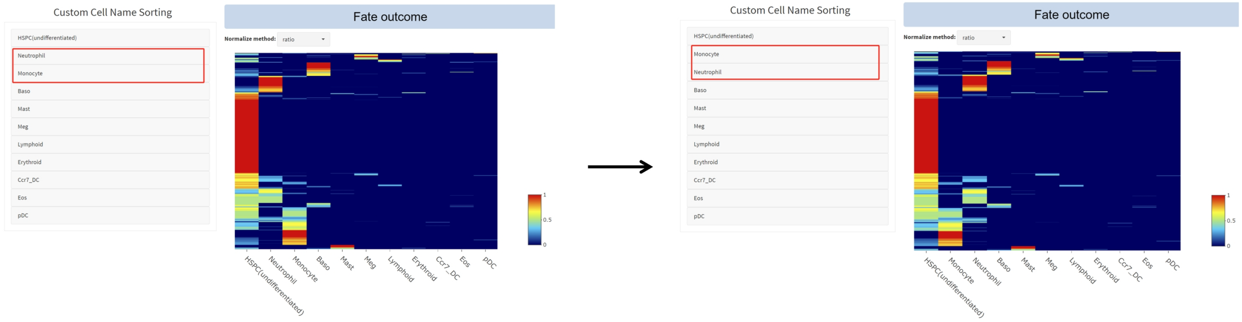
2. Clone embedding: visualize a single-cell derived clone or selected cell clones on transcriptional landscape.

3. Cell type (state) number barplot: visualize cell type (state) number and fraction.

4. Gene expression: plot the gene of interest on the embedding plot. Please note that this function does not work for datasets with DNA methylation, and chromatin accessibility modalities.

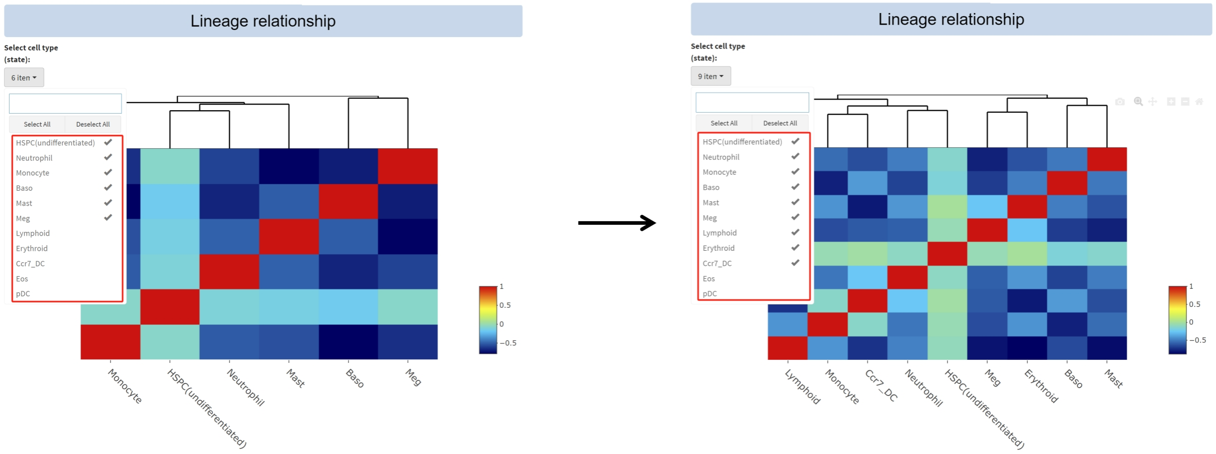
Fate outcomes

This function can assist users to re-map cell fate of targeted populations by visualizing the barcode propagation across various cell types or states. Users can adjust the column order (cell type order) of the heatmap. This can be done by simply rolling the cell types (states) bars located on the left panel of the heatmap.



Lineage relationships

This function is used to compare the lineage similarity between cell types (states) based on the number and frequency of barcodes present in each cell type (state). If two cell types share many barcodes at similar frequencies, they are likely to have arisen from a common developmental pathway; if not, they probably developed more independently. Users can select specific cell types to reconstruct lineage relationships by clicking on **Select cell type (state)**.

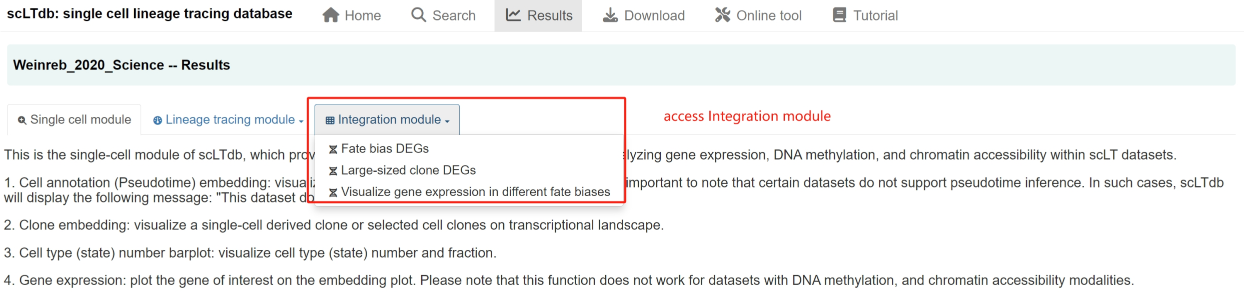


Barcode statistics

This section help users to visualize the number of cells carry the same clonal barcode.

Integration module

The integration module offers a range of functions to perform integrative analysis and visualization of gene expression and lineage tracing clones. Users can click the **Integration module** button to access this module.



This is the single-cell module of scLTdb, which prov lyzing gene expression, DNA methylation, and chromatin accessibility within scLT datasets.

1. Cell annotation (Pseudotime) embedding: visuali important to note that certain datasets do not support pseudotime inference. In such cases, scLTdb will display the following message: "This dataset d

2. Clone embedding: visualize a single-cell derived clone or selected cell clones on transcriptional landscape.

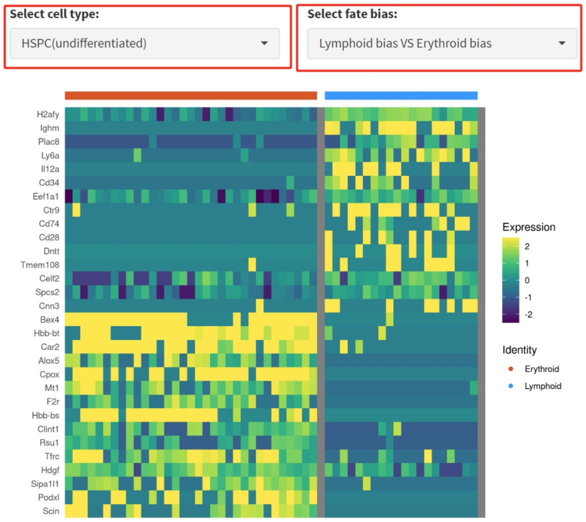
3. Cell type (state) number barplot: visualize cell type (state) number and fraction.

4. Gene expression: plot the gene of interest on the embedding plot. Please note that this function does not work for datasets with DNA methylation, and chromatin accessibility modalities.

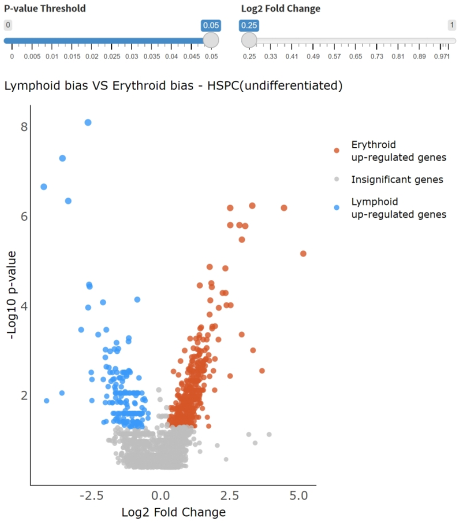
Fate bias DEGs

This section provides Differentially Expressed Genes (DEGs) analysis for cells with different fate biases. For example, we want to analyze the transcriptome heterogeneity between **Lymphroid biased HSPC** and **Erythroid biased HSPC**, we select cell type as **HSPC**, and fate bias as **Lymphroid bias vs Erythroid bias**.

For the volcano plot, you can use the slide windows of **P-value Threshold** and **Log2 Fold Change** to adjust the significant genes.


Select cell type: HSPC(undifferentiated)

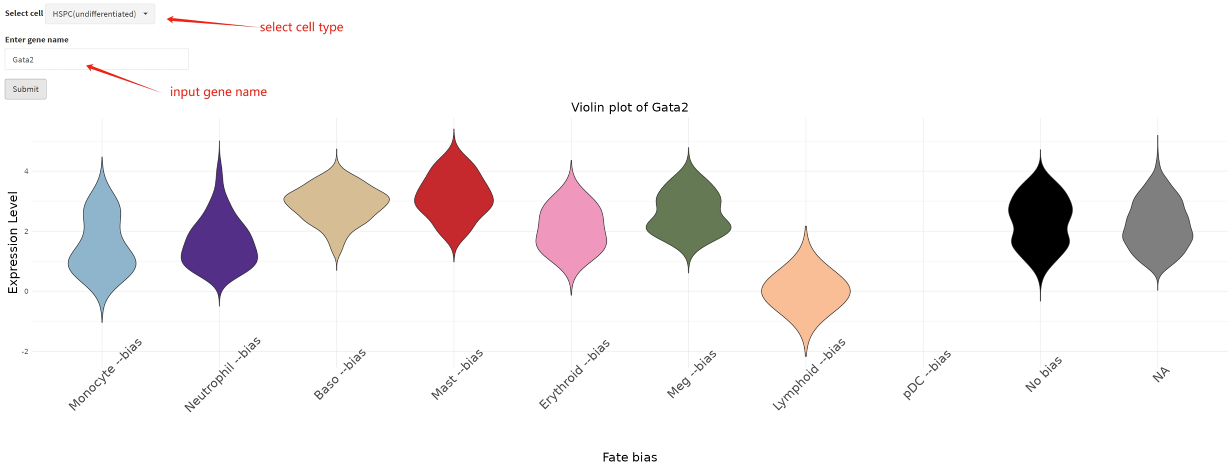
Select fate bias: Lymphoid bias VS Erythroid bias


P-value Threshold: 0.05

Log2 Fold Change: 0.25

Violin plot of gene with different fate bias

scLTdb also provides function for users to demonstrate the variation in gene expression across different fate biases within a particular cell type. For example, we want to visualize the expression of Gata2 among different **HSPC fate biases**. We firstly need to select cell type **HSPC**, then type **Gata2** at the **Enter gene name** box.


Select cell: HSPC(undifferentiated)

Enter gene name: Gata2

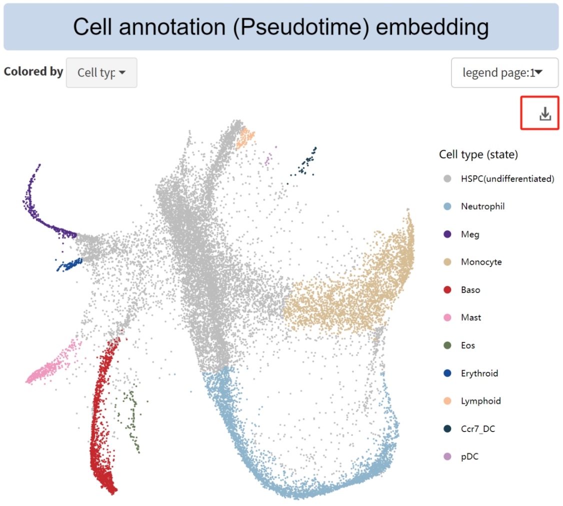
select cell type

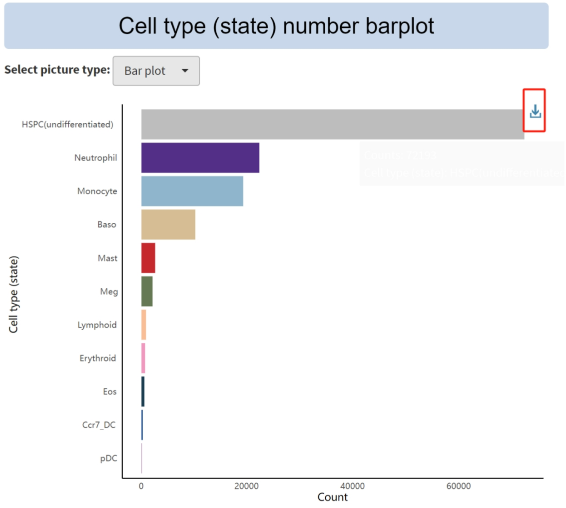
input gene name

Download

Download plots of online analyzed result.

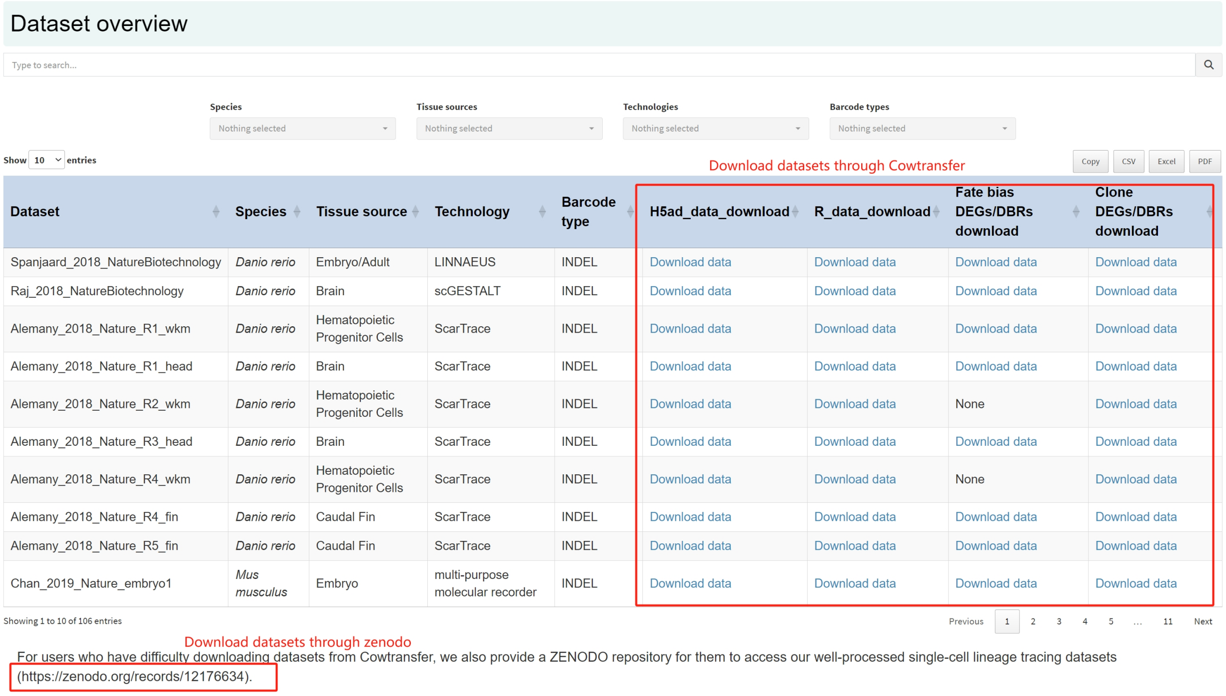
When users hover their mouse over the figures in scLTdb, a download icon will be displayed on the left side of the plot. Users can click on this icon to download the figures they are interested in.





Download well-processed datasets and analysis result

scLTdb also enables users to download both well-processed single-cell lineage tracing datasets and the corresponding analysis results. These download links are conveniently accessible through an interactive table. Furthermore, we have created a backup repository in ZENODO where users can also access and download our datasets.



Download datasets through Cowtransfer

Dataset	Species	Tissue source	Technology	Barcode type	H5ad_data_download	R_data_download	Fate bias DEGs/DBRs download	Clone DEGs/DBRs download
Spanjaard_2018_NatureBiotechnology	*Danio rerio*	Embryo/Adult	LINNAEUS	INDEL	Download data	Download data	Download data	Download data
Raj_2018_NatureBiotechnology	*Danio rerio*	Brain	scGESTALT	INDEL	Download data	Download data	Download data	Download data
Alemany_2018_Nature_R1_wkm	*Danio rerio*	Hematopoietic Progenitor Cells	ScarTrace	INDEL	Download data	Download data	Download data	Download data
Alemany_2018_Nature_R1_head	*Danio rerio*	Brain	ScarTrace	INDEL	Download data	Download data	Download data	Download data
Alemany_2018_Nature_R2_wkm	*Danio rerio*	Hematopoietic Progenitor Cells	ScarTrace	INDEL	Download data	Download data	None	Download data
Alemany_2018_Nature_R3_head	*Danio rerio*	Brain	ScarTrace	INDEL	Download data	Download data	Download data	Download data
Alemany_2018_Nature_R4_wkm	*Danio rerio*	Hematopoietic Progenitor Cells	ScarTrace	INDEL	Download data	Download data	None	Download data
Alemany_2018_Nature_R4_fin	*Danio rerio*	Caudal Fin	ScarTrace	INDEL	Download data	Download data	Download data	Download data
Alemany_2018_Nature_R5_fin	*Danio rerio*	Caudal Fin	ScarTrace	INDEL	Download data	Download data	Download data	Download data
Chan_2019_Nature_embryo1	*Mus musculus*	Embryo	multi-purpose molecular recorder	INDEL	Download data	Download data	Download data	Download data

Showing 1 to 10 of 106 entries

Download datasets through zenodo

For users who have difficulty downloading datasets from Cowtransfer, we also provide a ZENODO repository for them to access our well-processed single-cell lineage tracing datasets (https://zenodo.org/records/12176634).