Szczegółowe parametry narzędzi wykorzystanych do analizy na platformie Galaxy

ETAP 3 - Dalsza analiza danych za pośrednictwem platformy Galaxy

2.5 Odfiltrowanie komórek o niskiej liczbie odczytów

Parameter	Value
Annotated data matrix	processed_anndata.h5ad
Method used for filtering	`pp.filter_cells`
Filter	`min_counts`
Minimum number of counts required	3
Output Log?	false

2.6 Parametry funkcji `pp.calculate_qc_metrics`

Input Parameter	Value
Annotated data matrix	Scanpy filter (pp.filter_cells) on data 1: Annotated data matrix
Method used for inspecting	`pp.calculate_qc_metrics`
Name of kind of values in X	`counts`
The kind of thing the variables are	`genes`
Keys for boolean columns in `.var`	`mito`
Proportions of top genes to cover	(empty)
Use expression from `adata.layers[layer]`	(empty)
Use `raw` attribute if present	false
Compute log1p transformed annotations	true
Output Log?	false

2.7 Filtrowanie komórek na podstawie liczby genów

Input Parameter	Value
Annotated data matrix	Scanpy Inspect and manipulate (pp.calculate_qc_metrics) on data 112: Annotated data matrix
Method used for filtering	`pp.filter_cells`
Filter	`min_genes`
Minimum number of genes required	200
Output Log?	false

2.8 Filtrowanie komórek o wysokim udziale mitochondrialnego RNA

Input Parameter	Value
Annotated data matrix	Scanpy filter (pp.filter_cells) on data 113: Annotated data matrix
Function to manipulate the object	`filter`
What to filter?	Observations (`obs`)
Type of filtering	`key`
Key to filter	`pct_counts_mito`
Type of value to filter	number
Filter	less than
Value	20

2.9 Normalizacja danych

Input Parameter	Value
Annotated data matrix	Manipulate AnnData (filter) on data 117
Method used for normalization	`pp.normalize_total`
Target sum	$10,000$
Exclude highly expressed genes from normalization factor	false
Field in `adata.obs` for normalization factor	`norm`
Layer to normalize instead of `X`	(empty)
Output Log?	false

2.10 Logarytmizacja danych

Input Parameter	Value
Annotated data matrix	Scanpy normalize (pp.normalize_total) on data 118: Annotated data matrix
Method used for inspecting	`pp.log1p`
Base of the logarithm	Not available
Entry of `layers` to transform	(empty)
Entry of `obsm` to transform	(empty)
Output Log?	false

2.11 Oznaczenie genów o największej zmienności (HVG)

Input Parameter	Value
Annotated data matrix	Scanpy Inspect and manipulate (pp.log1p) on data 120: Annotated data matrix
Method used for filtering	`pp.highly_variable_genes`
Flavor	`seurat`
Minimal mean cutoff	0.0125
Maximal mean cutoff	3
Minimal normalized dispersion cutoff	0.5
Maximal normalized dispersion cutoff	Not available
Number of bins for mean expression	20
Inplace subset to highly-variable genes?	false
Expression source (`adata.layers[layer]`)	(empty)
Batch key	(empty)
Output Log?	false

2.12 Regresja zmiennych technicznych

Input Parameter	Value
Annotated data matrix	Scanpy filter (pp.highly_variable_genes) on data 122: Annotated data matrix
Method used for plotting	`pp.regress_out`
Element of `layers` to regress on	(empty)
Keys for `obs` to regress on	`total_counts`, `pct_counts_mito`
Output Log?	false

2.13 Skalowanie

Input Parameter	Value
Annotated data matrix	Scanpy remove confounders (`pp.regress_out`) on data 123
Method used for inspecting	`pp.scale`
Zero center?	true
Maximum value	10
Element of `layers` to scale	(empty)
Element of `obsm` to scale	(empty)
Subset of observations to scale	(empty)
Output Log?	false

2.14 Analiza głównych składowych (PCA)

Input Parameter	Value
Annotated data matrix	Scanpy Inspect and manipulate (`pp.scale`) on data 124
Method used	`pp.pca`
Number of principal components	50
Element of `layers` to use for PCA	(empty)
Numpy data type of result	`float32`
Type of PCA?	false
Compute standard PCA from covariance matrix?	true
SVD solver	(nothing selected)
Random state initialization	0
Subset of genes for PCA	(empty)
Output Log?	false

2.15 Korekcja efektów batch metodą BBKNN

Input Parameter	Value
Annotated data matrix	Scanpy cluster, embed (pp.pca) on data 126: Annotated data matrix
Method used for plotting	`external.pp.bbknn`
Batch key	`batch`
Dimensionality reduction to use	`X_pca`
Approximate neighbour finding	`annoy`
Number of trees in annoy forest	10
Distance metric	`euclidean`
Top neighbours per batch	3
Number of dimensions used	28
Trim top connectivities	Not available
UMAP connectivity parameter	1
Fully connected nearest neighbors	1
Output Log?	false

2.16 Redukcja wymiarowości metodą UMAP

Input Parameter	Value
Annotated data matrix	Scanpy remove confounders (`external.pp.bbknn`) on data 127
Method used	`tl.umap`
Minimum distance between points	0.5
Scale of embedded points	1
Embedding dimensionality	2
Number of optimization iterations	Not available
Initial learning rate	1
Negative sample weighting	1
Number of negative samples per positive	5
Embedding initialization	Spectral embedding of the graph
Random seed	0
Neighbor/connectivity key	(empty)
Output Log?	false

2.17 Identyfikacja genów różnicujących klastry

Input Parameter	Value
Annotated data matrix	Scanpy cluster, embed (`tl.umap`) on data 132
Method used for inspecting	`tl.rank_genes_groups`
Export ranked genes as table?	true
Adjusted p-value cutoff	0.05
Minimum log fold change	Not available
Maximum log fold change	Not available
Gene symbol column in `.var`	(empty)
Observation grouping key	`louvain`
Use `raw` attribute?	false
Subset of groups for comparison	(empty)
Layer for testing	(empty)
Comparison	`rest`
Number of top genes to return	100
Statistical method	`wilcoxon`
P-value correction	Benjamini-Hochberg
Tie correction for Wilcoxon	false
Key in `adata.uns` for saving result	(empty)
Output Log?	false