

15-16 February 2021

COMETH Training course

From omics data

to tumor heterogeneity quantification

EIT Health is supported by the EIT,
a body of the European Union

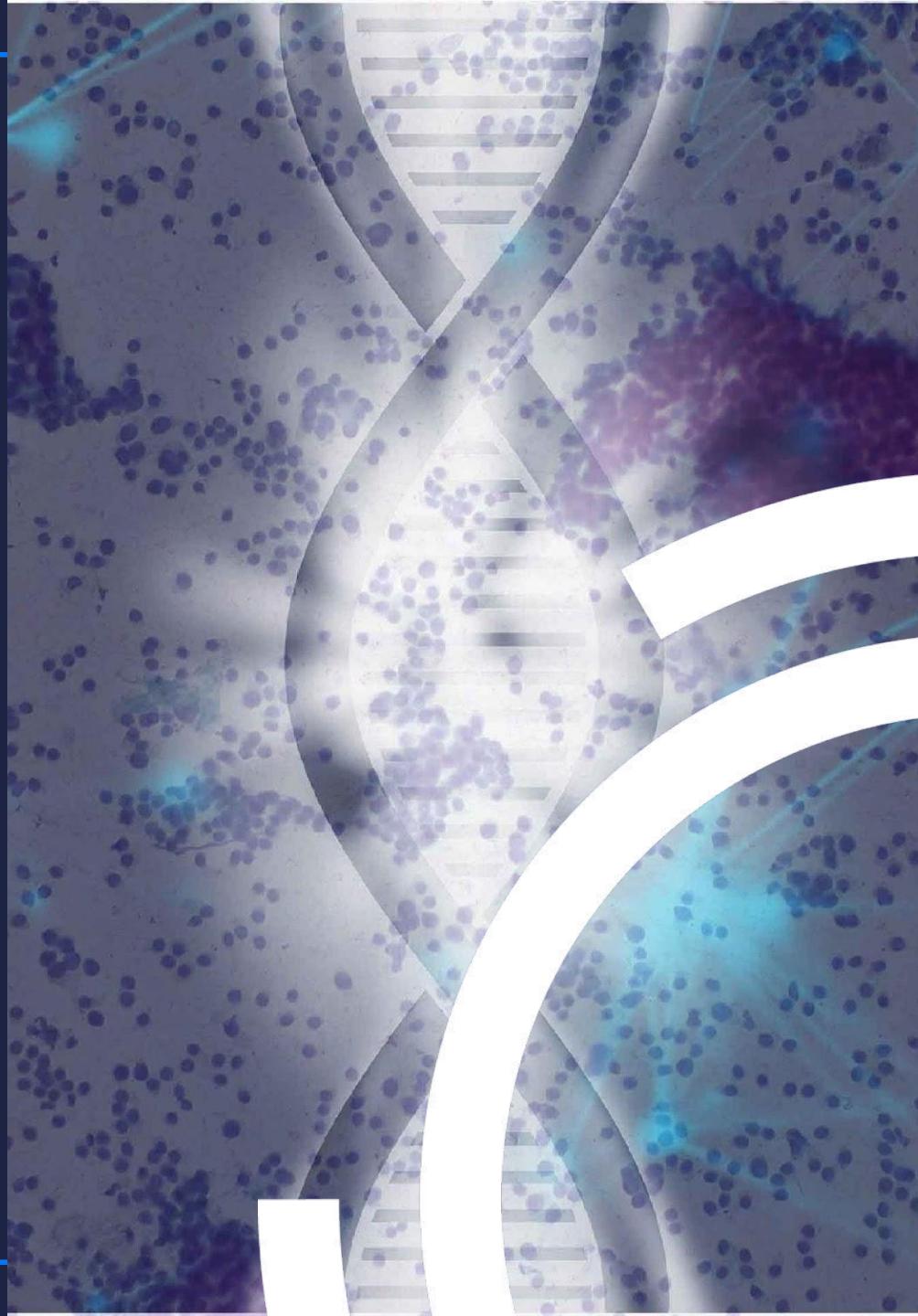




16 February 2021

Biological interpretation

Yuna Blum and Ashwini Sharma



compExplore Shiny app



Help you in the analysis, interpretation and visualization of the results

compExplore - Components explorer is a visualization tool to guide the user in the analysis and interpretation of the results from *Supervised* and *Unsupervised* gene expression deconvolution algorithms.

Different modules

The diagram illustrates the decomposition of a gene expression matrix into three components:

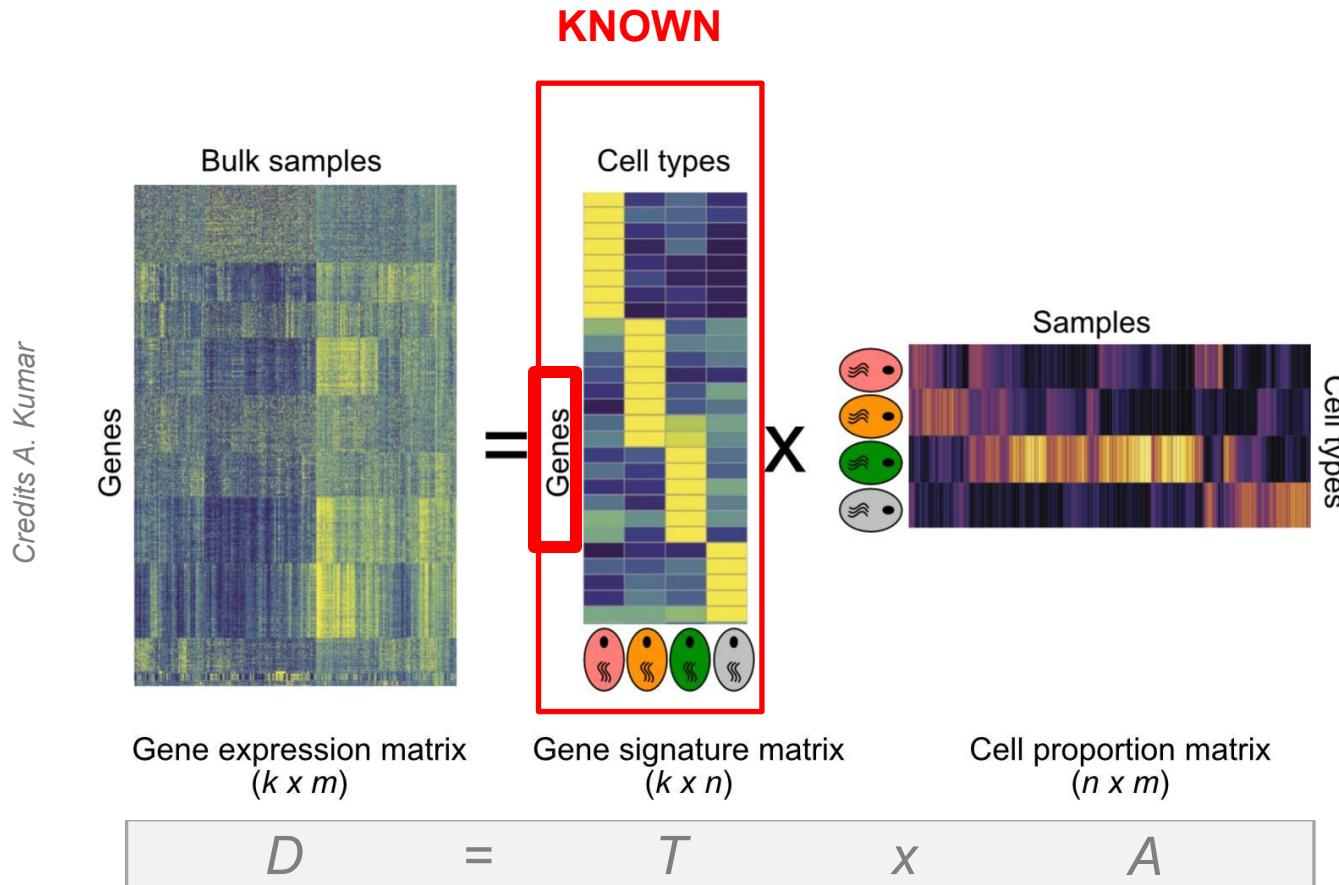
- Complex tissue with multiple cell types:** Shows a population of cells with different mRNA levels. A "Cell type" is highlighted.
- Bulk samples:** The result of "Bulk RNAseq" analysis, represented as a heatmap of gene expression levels across samples.
- Gene expression matrix ($k \times m$):** The total expression data.
- Gene signature matrix ($k \times n$):** A matrix where rows represent genes and columns represent cell types. It shows the contribution of each gene to each cell type.
- Cell proportion matrix ($n \times m$):** A matrix where rows represent cell types and columns represent samples. It shows the proportion of each cell type in each sample.

Terminology

1. Gene expression matrix - it is a $k \times m$ matrix with k rows of genes and m columns of samples. Each data point in this matrix represents the expression of a given gene in a given sample
2. Gene signature matrix - it is a $k \times n$ matrix with n rows of genes and m columns of cell fraction. Each data point in this matrix represents the contribution of a gene towards a cell type
3. Cell proportion matrix - it is a $n \times m$ matrix with n rows of cell types and m columns of samples. Each data point in this matrix represents the proportion of a given cell type in a given sample

<https://app.gebican.fr/compExplore/>

(Semi) Supervised methods

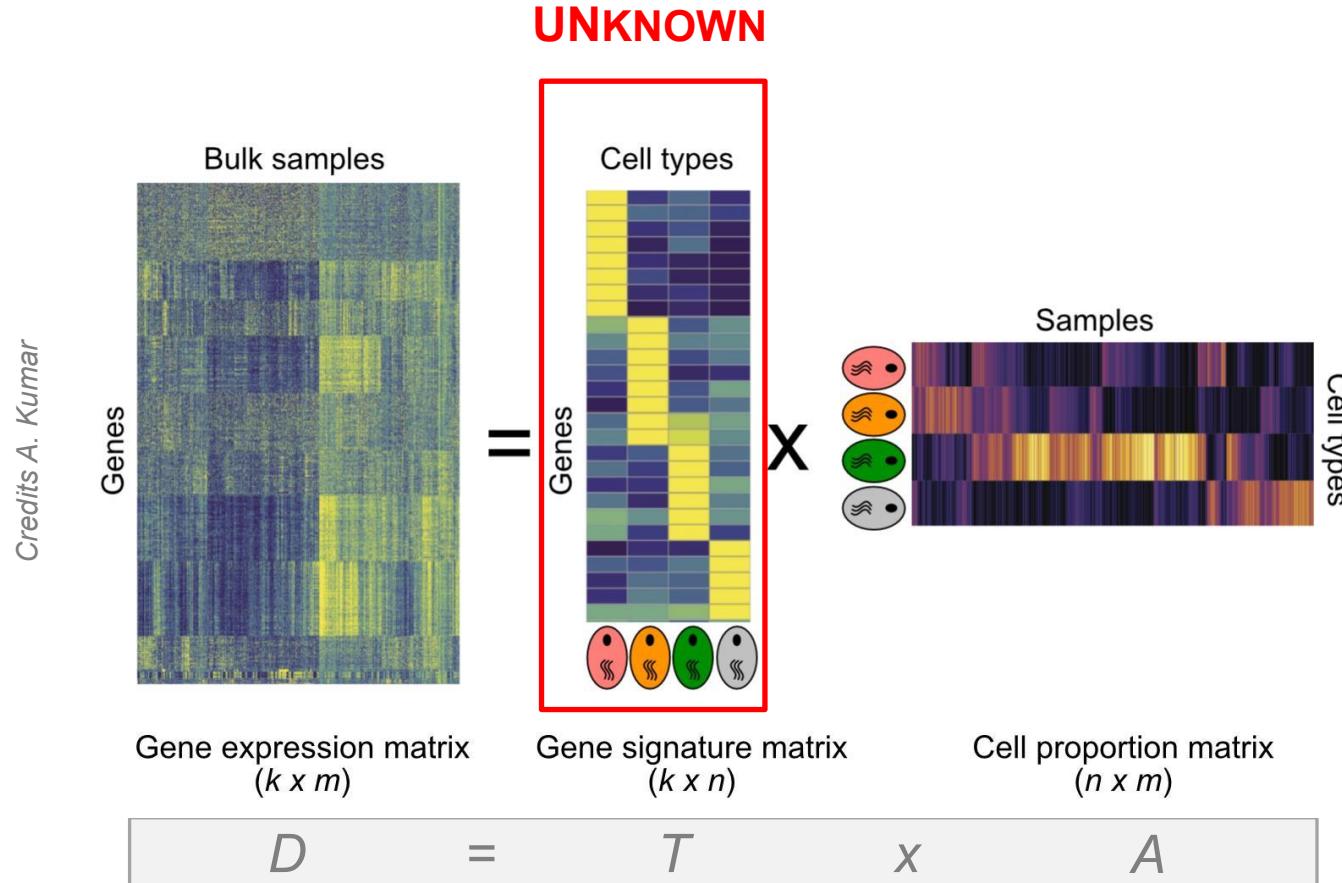


Methods available in the cometh web-app (#6)

Supervised: Cibersort (MT8), EPIC (MT9), quantiseq (MT11)

Semi-supervised: 3 different versions of **cellmix** (MT16, 17, 18) using cellMatch gene markers

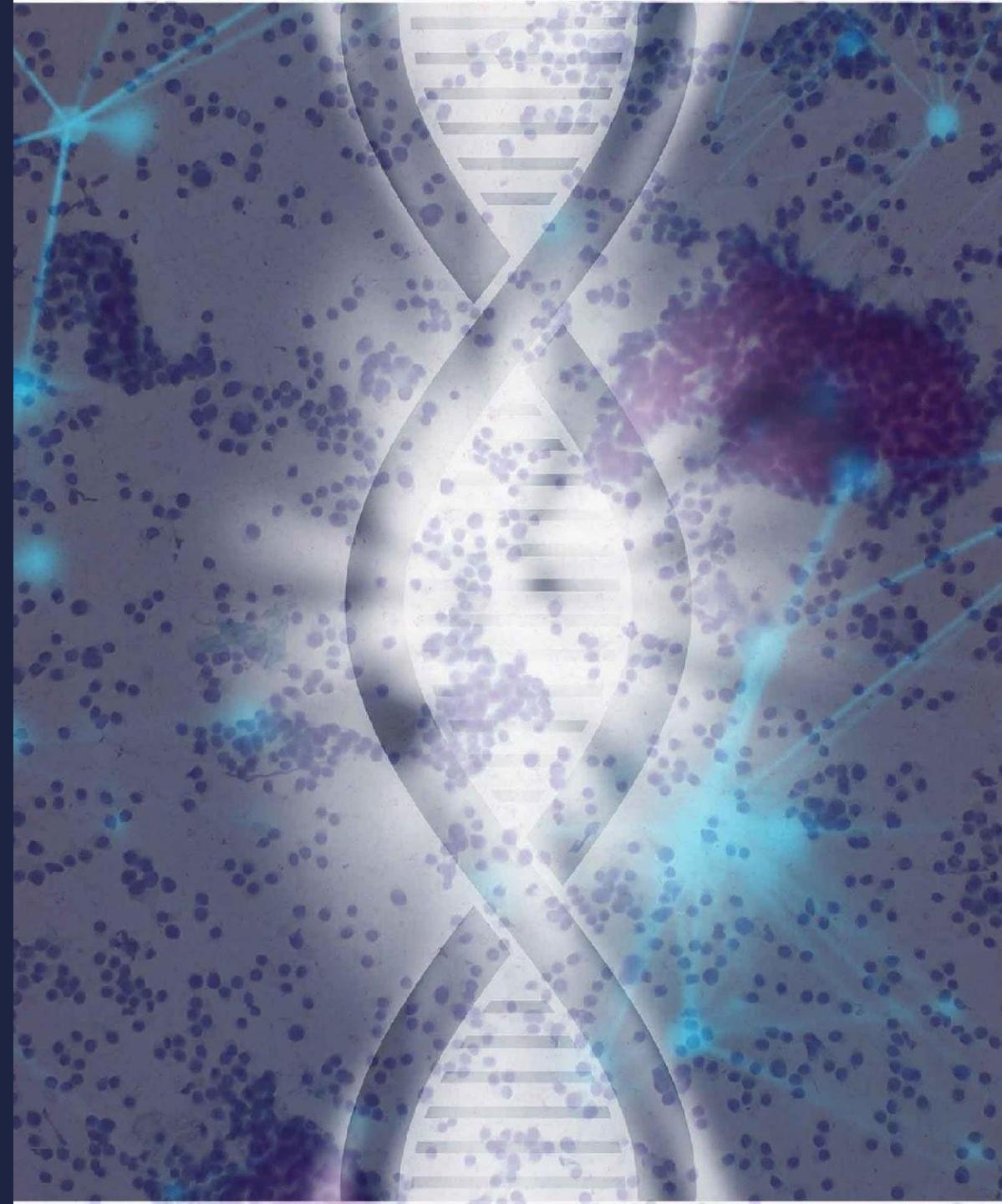
Unsupervised methods



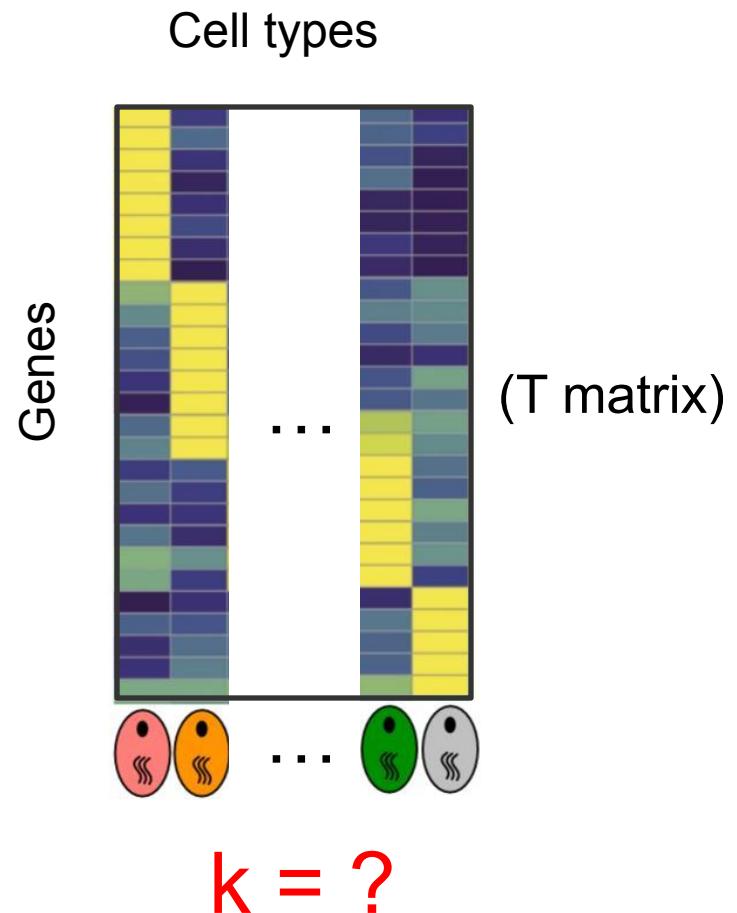
Methods available in the cometh web-app (#5)

ICA with ICA-based feature selection (MT1_ICA_fs, transcriptome), **NMF with ICA-based feature selection** (MT2_NMF_fs , transcriptome), **Edec method** (MT3_edec , methylome), **ICA without feature selection** (MT14_ICA transcriptome and methylome), **NMF without feature selection** (MT19_NMF, transcriptome and methylome)

Unsupervised methods: finding the number of k of cell types



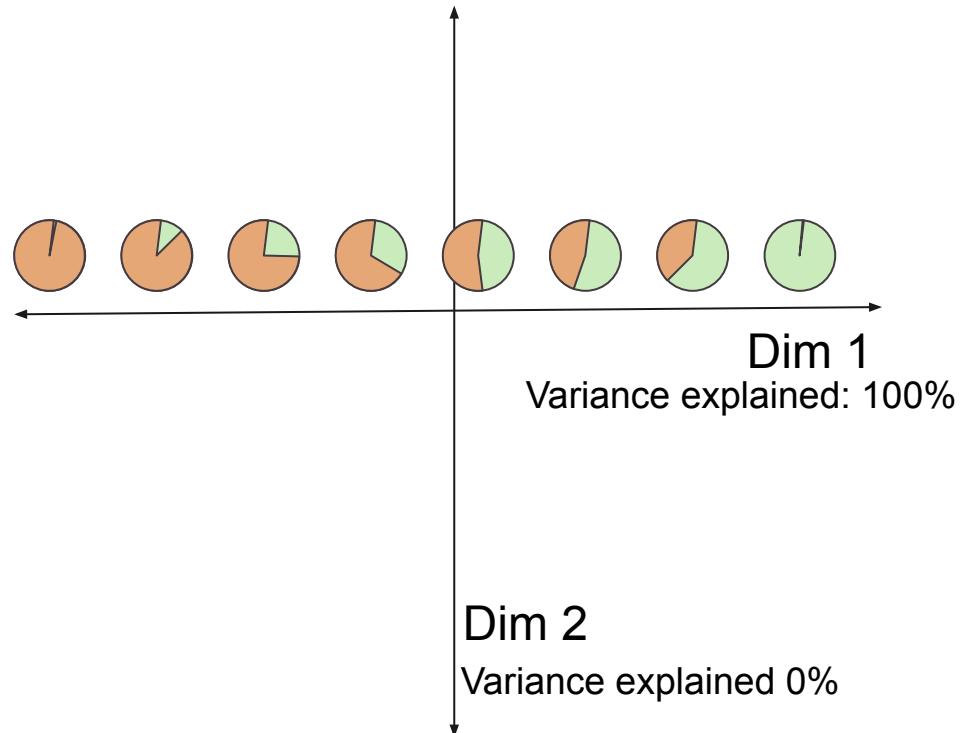
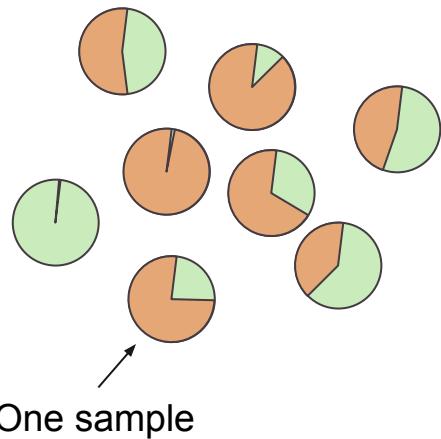
Finding the number of cell types k



Finding the number of cell types k

Principal Component Analysis (PCA)

Samples
mixtures of 2 cell types

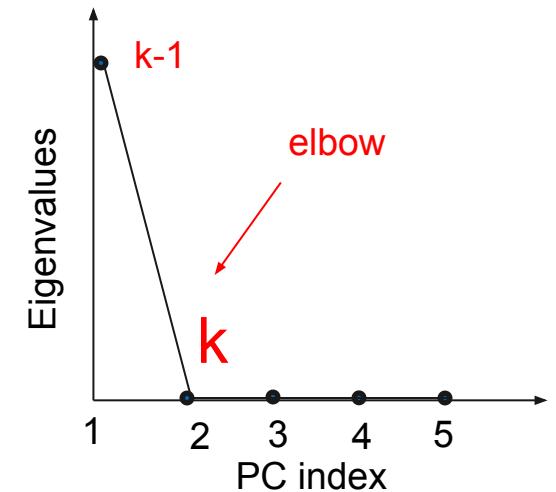


Reminder

Find the axes that maximized the explained variance (inertia)

Principal components are orthogonal

Plot of eigenvalues (=scree plot)



Eigenvalues represent the variance explained

Cattel's rule: $k = \text{PCs} + 1$

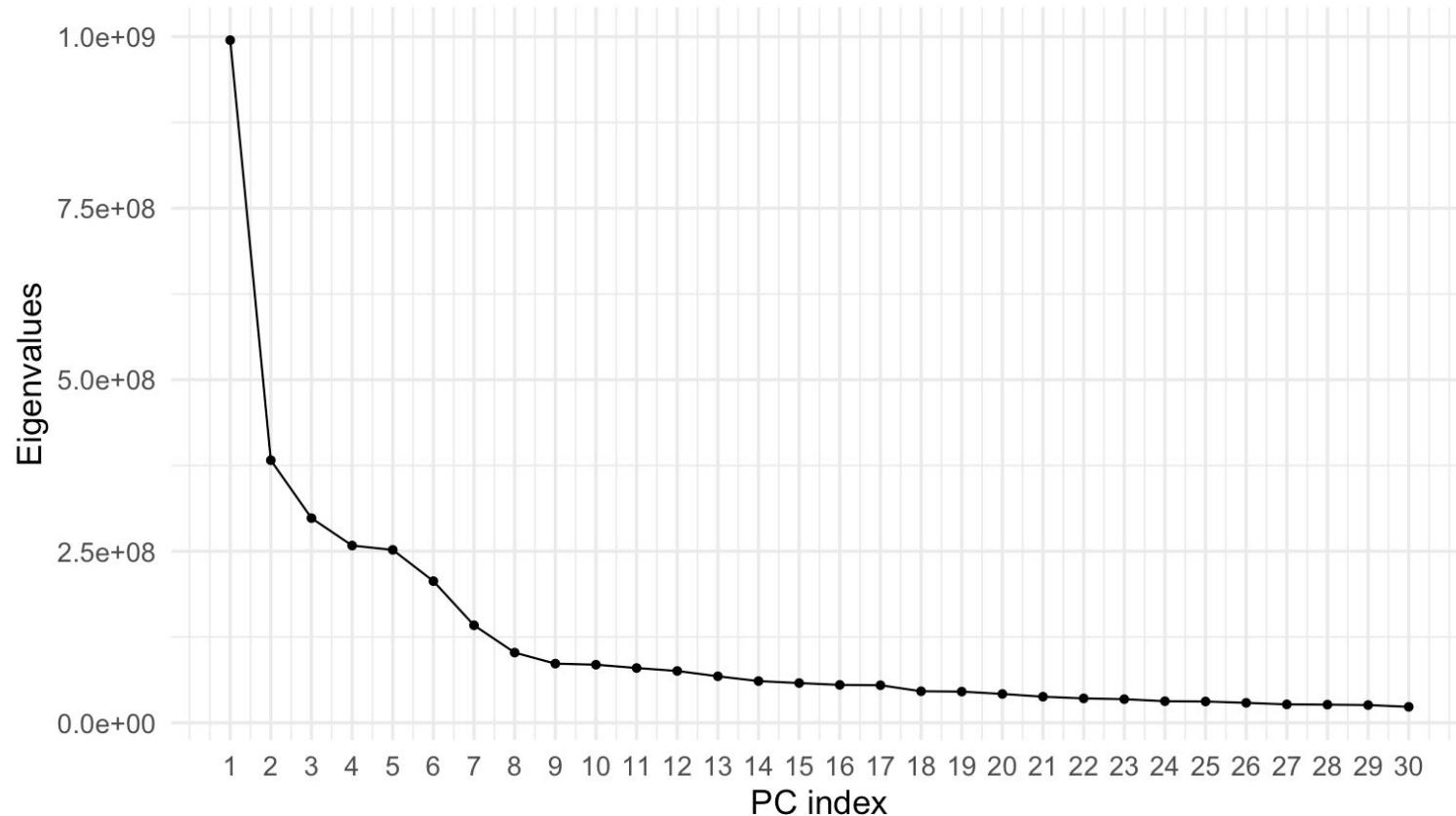
Number of relevant PCs

Finding the number of cell types k

Real life

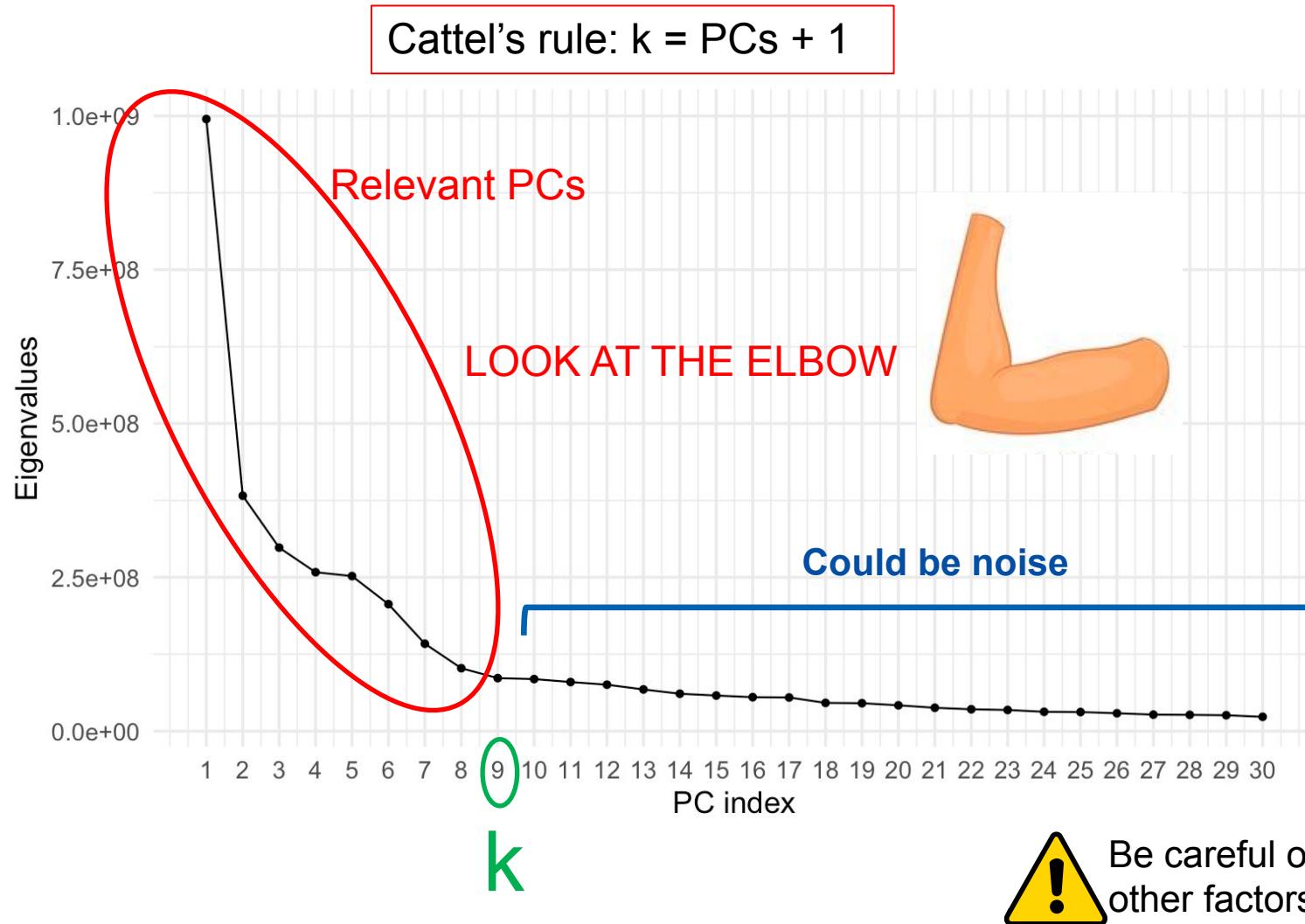


Cattel's rule: $k = \text{PCs} + 1$

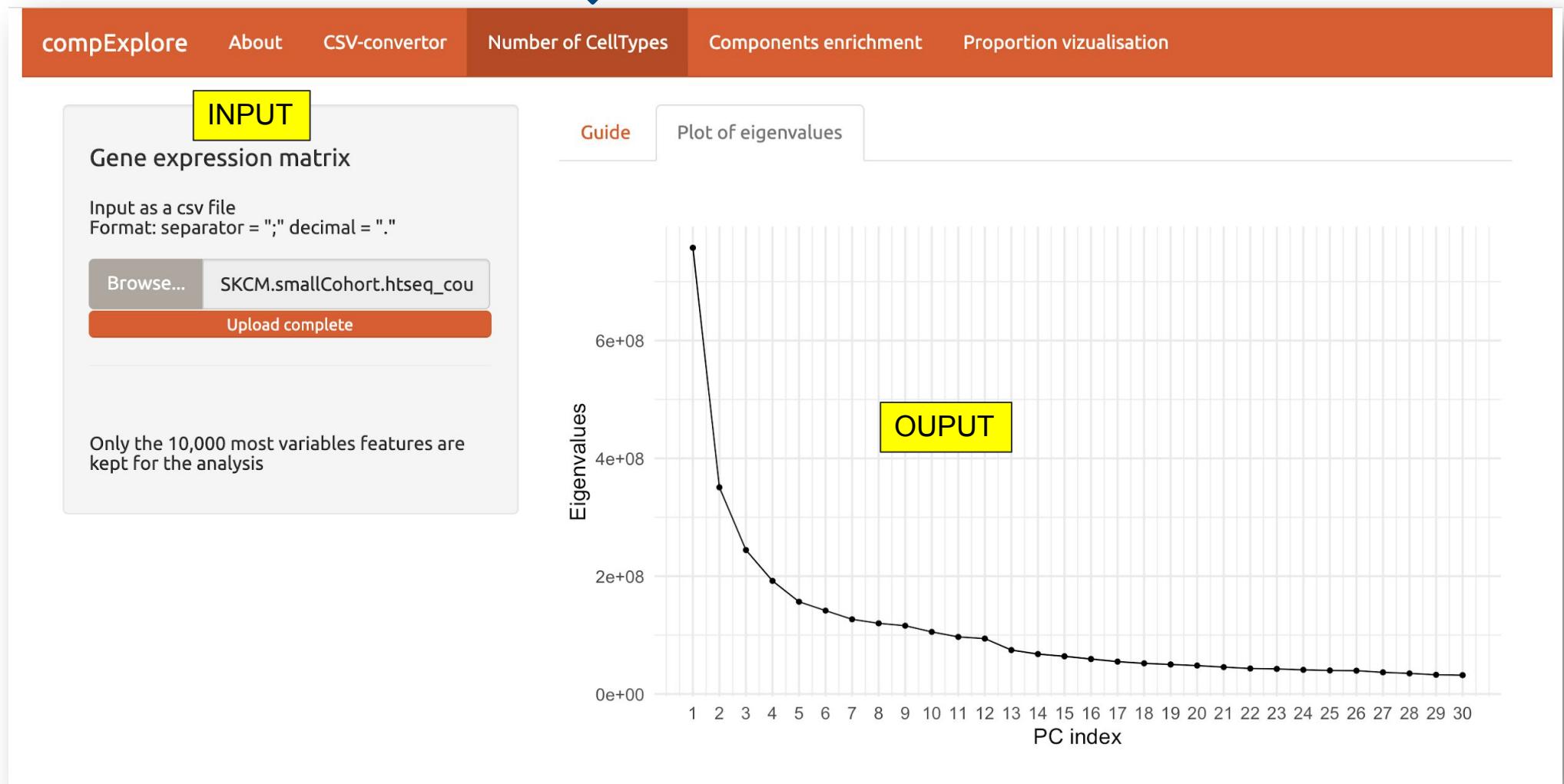


Finding the number of cell types k

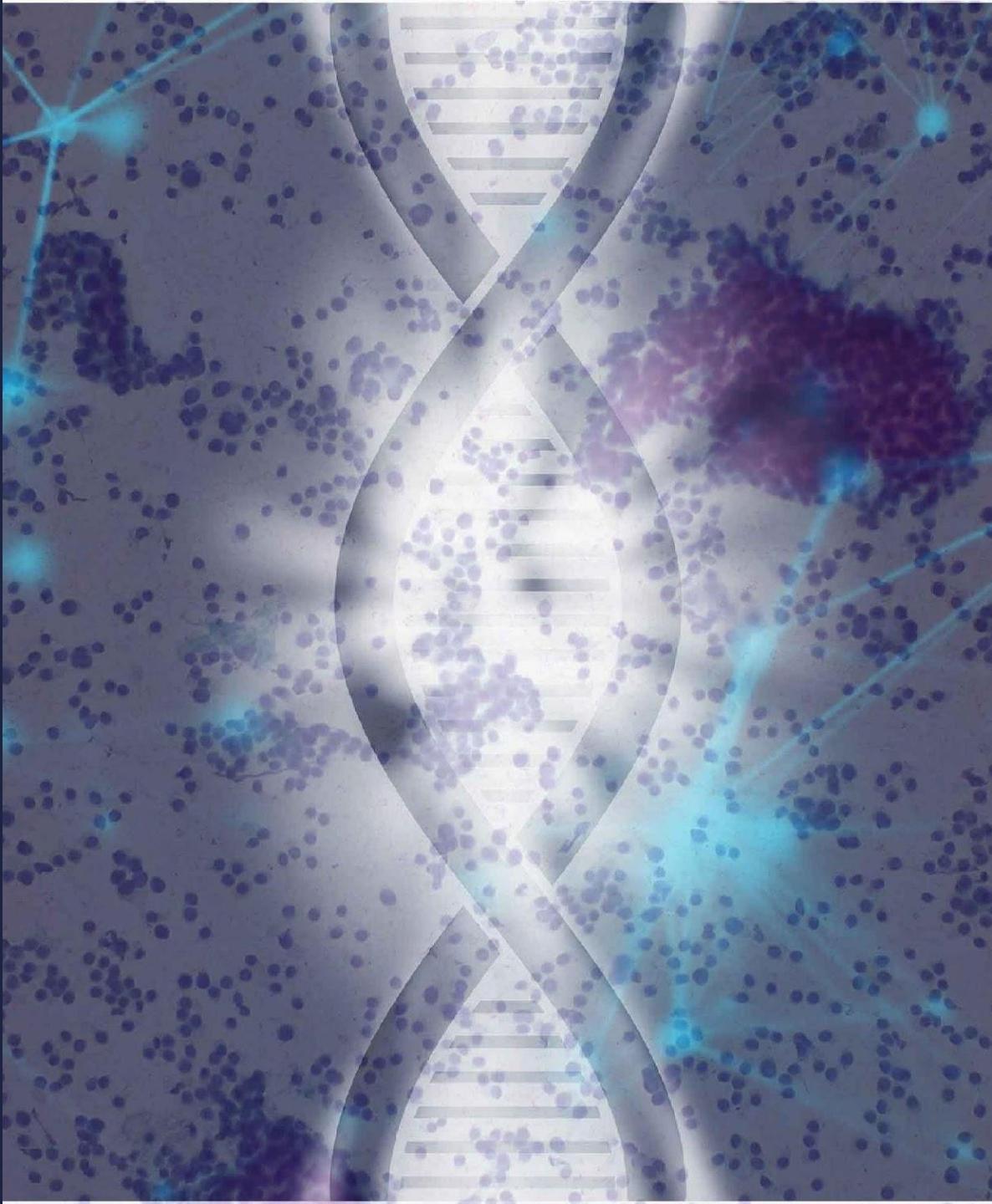
Real life



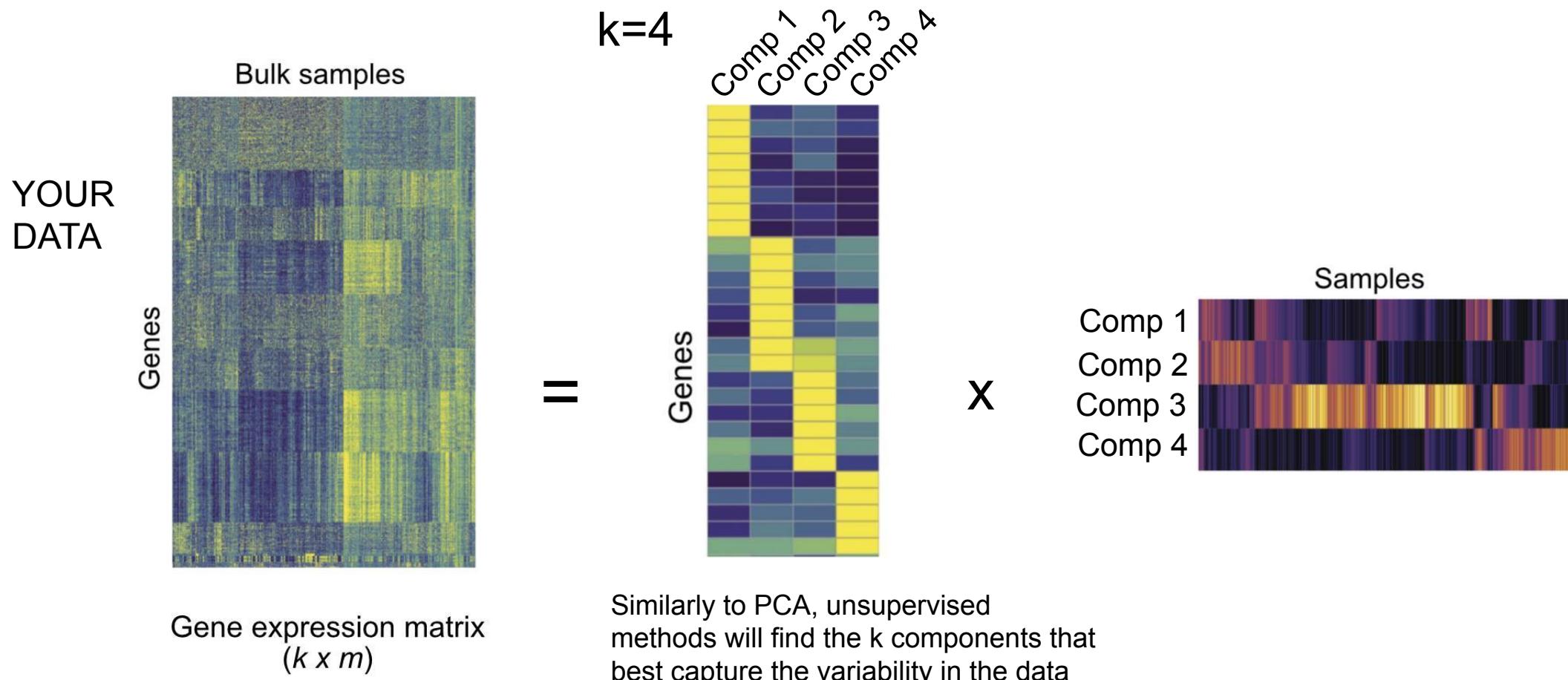
compExplore Shiny app



Unsupervised methods: Interpret the components identified

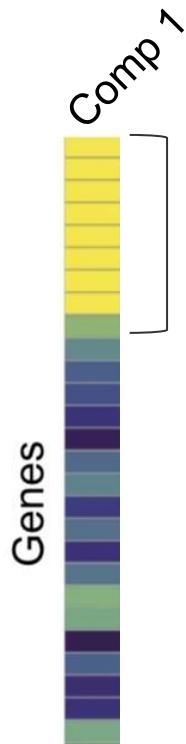


Interpret the components identified



To which cell type(s) corresponds each of the components identified by unsupervised methods?

Interpret the components identified



Genes with high scores
on the component

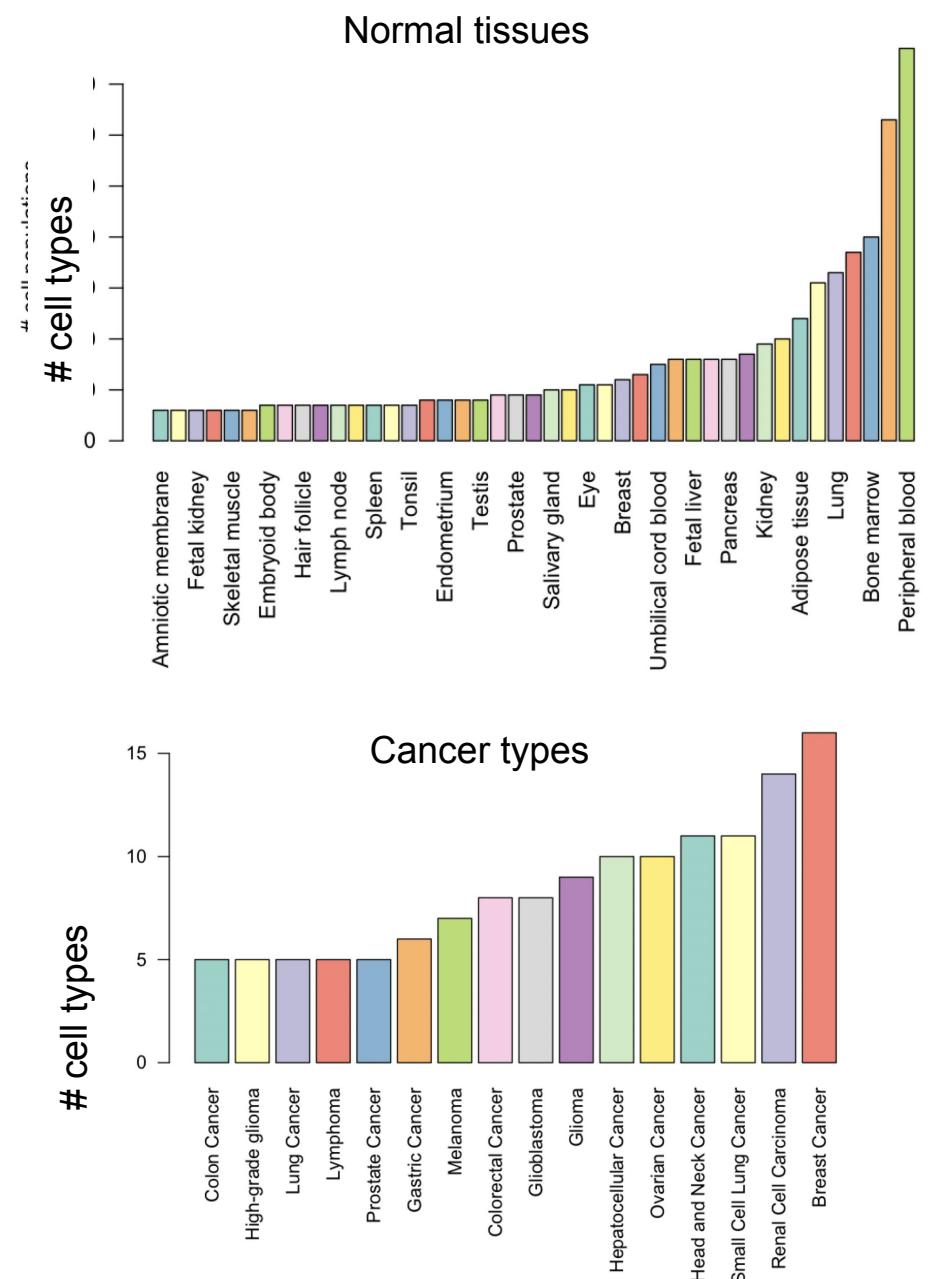
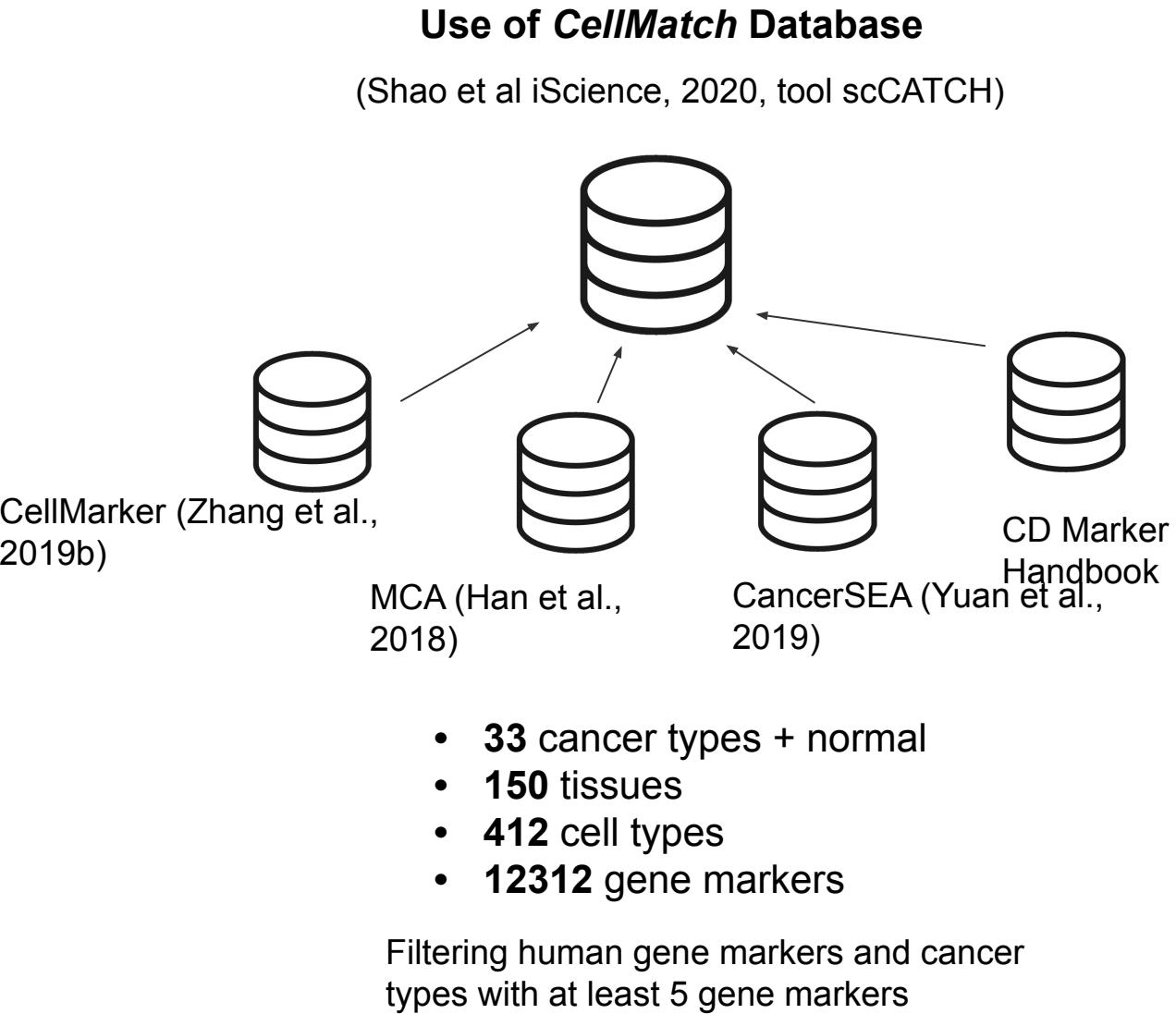


Markers of a particular cell
types?

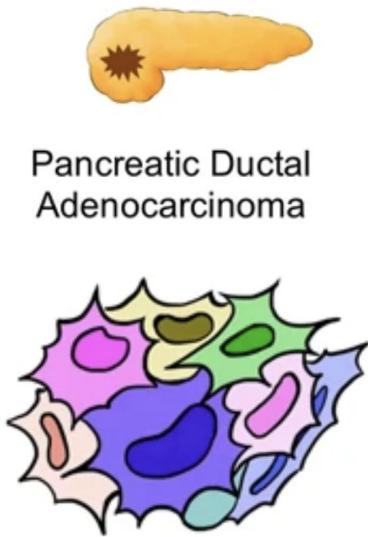
Use of *CellMatch* Database



Interpret the components identified



Interpret the components identified

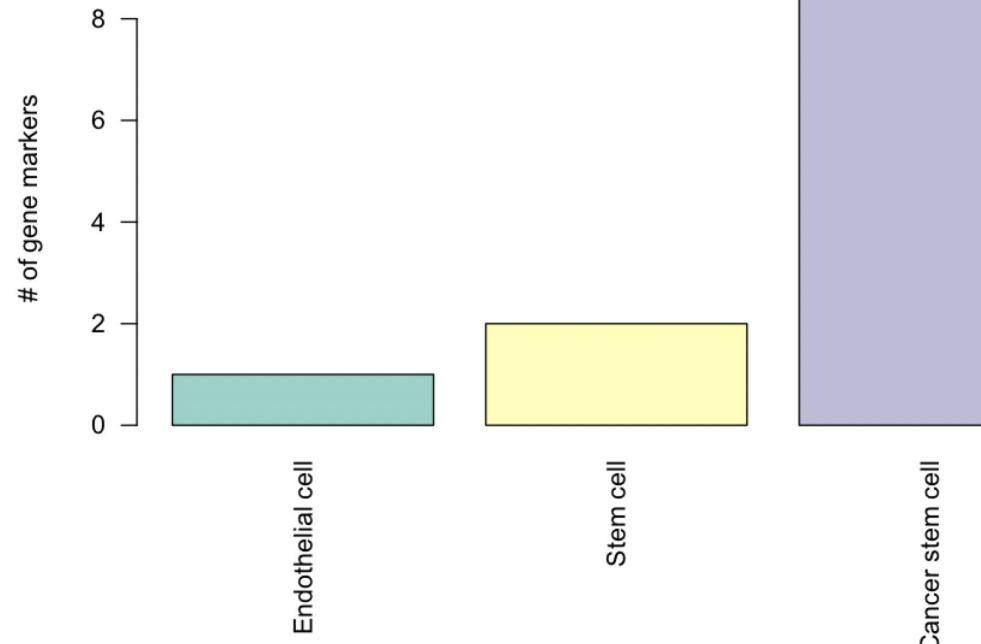


Cell Type
Ductal cell 1
Ductal cell 2
Acinar cell
Endocrine cell
Endothelial cell
Fibroblast
Stellate cell
Macrophage
T cell
B cell

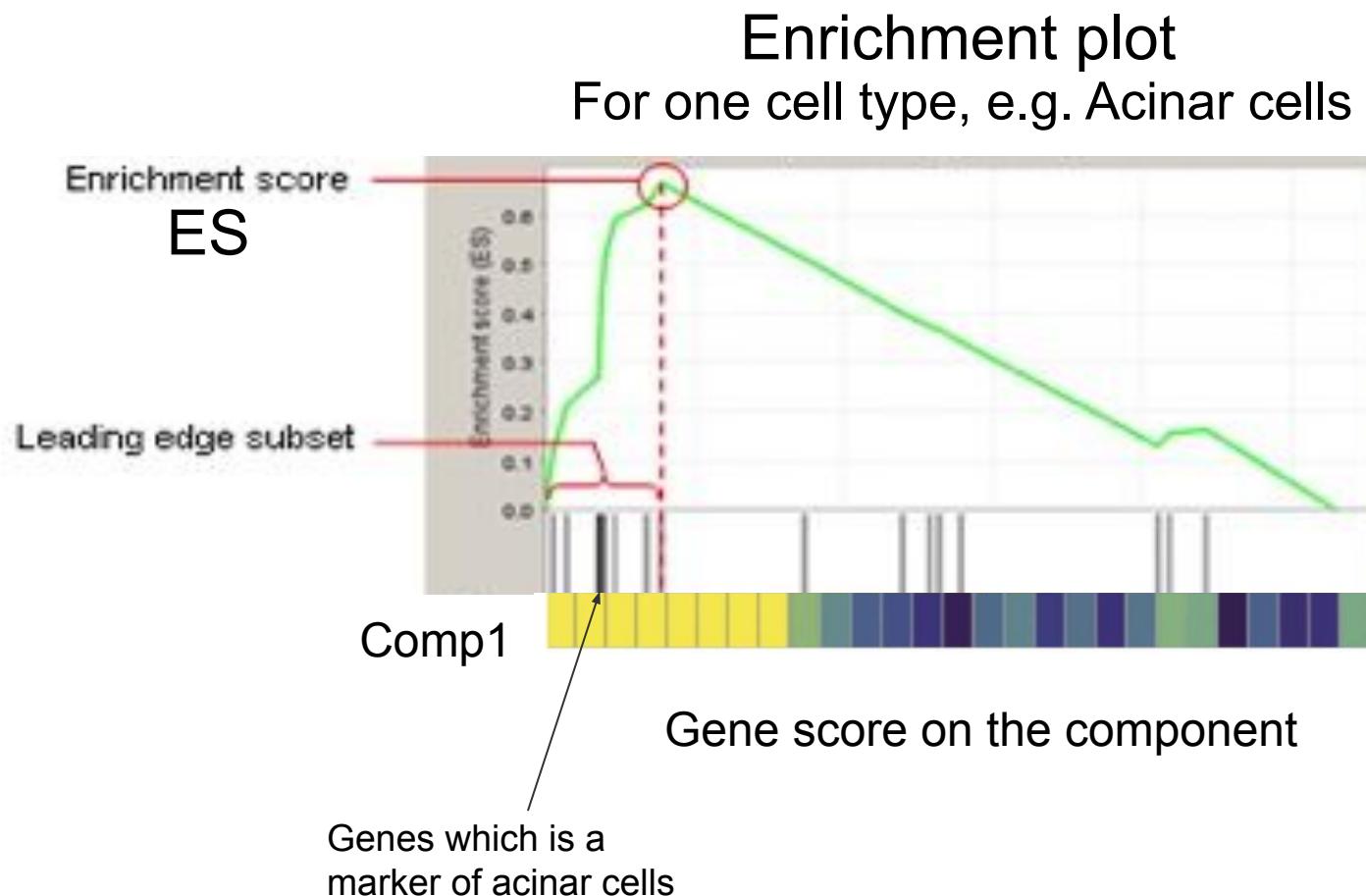
Peng et al 2019 Nature

CellMatch Database

PDAC cell types



Gene Set Enrichment Analysis (use of the fgsea R package)



- 1- Order the list of genes (test statistic, p-value, here component loadings...)
- 2- Calculation of the Enrichment Score (ES)

The algorithm scans the list: the score increases when the gene is part of the set (=cell type) and decreases otherwise. The increase and decrease values is weighted by the gene rank (for a gene set overexpressed, the increase will be higher at the beginning of the list).
The ES corresponds to the max score (absolute value).

- 3- Comparison of ES to a distribution of ES obtained on from random data (gene permutations) . → Calculation of a p-value

compExplore Shiny app



INPUT

Gene signature matrix

Input as a csv file
Format: separator = ";" decimal = ","

results_T_1.csv

The Gene signature matrix corresponds to the output results_T_1.csv in the comet web-app which is already in the requested format (separator = ";" decimal = ",").

Deconvolution method

ICA-based

NMF-based

Cancer type

ALL

Download top markers

Top100 gene markers for each component

Guide **Enrichment analysis** **OUTPUT**

Gene set enrichment analysis results using CellMatch DB

V1

pathway	ES	-log10(pval)
Non_Small_Cell_Lung_Cancer.Fibroblast	0.975	4.5
Normal.Fibroblast	0.950	3.8
Colorectal_Cancer.CancerAssociated_Fibroblast	0.935	3.5
Head_and_Neck_Cancer.Fibroblast	0.925	3.2
Oligodendroglioma.Microglial_Cell	0.915	3.0
Glioma.Astrocyte	0.905	2.8
B_Cell.Lymphoma.B_Cell	0.900	2.6
Normal.Circulating_Fetal_Cell	0.900	2.4
Melanoma.CancerAssociated_Fibroblast	0.900	2.2
Normal.Stromal_Cell	0.900	2.0

V2

pathway	ES	-log10(pval)
Normal.Primitive_Vesicle_Cell	0.80	3.8
Head_and_Neck_Cancer.Mycocyte	0.75	3.5
Renal_Cell_Carcinoma.Erythroblast	0.70	3.2
Melanoma.B_Cell	0.65	3.0
Normal.Photoreceptor_Cell	0.60	2.8
Normal.Enterocrinology_Cell	0.55	2.6
Normal.Streak_Cell	0.50	2.4
Normal.Lake_Et_Al.science.in2	0.45	2.2

V3

pathway	ES	-log10(pval)
Normal.1Cell_Stage_Cell_Blastomere	0.975	4.5
Normal.Secretory_Cell	0.950	3.8
Ovarian_Cancer.Cancer_Cell	0.935	3.5
Normal.Alpha_Cell	0.925	3.2
Astrocytoma.Astrocyte	0.915	3.0
Melanoma.Macrophage	0.905	2.8
Colon_Cancer.Stem_Cell	0.900	2.6
Glioma.Astrocyte	0.900	2.4
Normal.Mast_Cell	0.900	2.2
Renal_Cell_Carcinoma.Neutrophil	0.900	2.0

V4

pathway	ES	-log10(pval)
Normal.Idiopathic_Pulmonary_Fibrosis_Cell	0.95	4.5
Melanoma.CancerAssociated_Fibroblast	0.90	3.8
Head_and_Neck_Cancer.Fibroblast	0.85	3.5
Colorectal_Cancer.CancerAssociated_Fibroblast	0.80	3.2
Non_Small_Cell_Lung_Cancer.Myeloid_Cell	0.75	3.0
Normal.Bile_Duct_Cell	0.70	2.8
Normal.Mesangial_Cell	0.65	2.6
Head_and_Neck_Cancer.Cancer_Cell	0.60	2.4

V5

pathway	ES	-log10(pval)
Colorectal_Cancer.CancerAssociated_Fibroblast	0.98	6.0
Non_Small_Cell_Lung_Cancer.Fibroblast	0.975	5.8
Normal.Myofibroblast	0.970	5.5
Ovarian_Cancer.Mesenchymal_Cell	0.965	5.3
Normal.Mesangial_Cell	0.960	5.0
Normal.Bile_Duct_Cell	0.955	4.8
Head_and_Neck_Cancer.Fibroblast	0.950	4.5
Normal.Fibroblast	0.945	4.2
Normal.Idiopathic_Pulmonary_Fibrosis_Cell	0.940	4.0
Normal.Pneumocyte	0.940	3.8

compExplore Shiny app



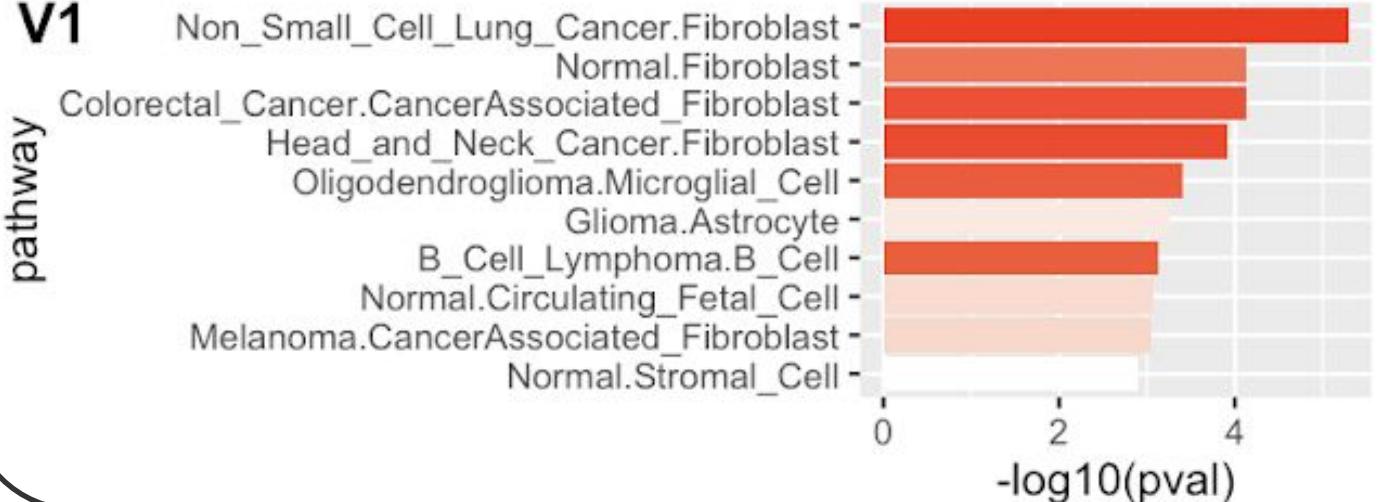
Example for component 1

(LUAD dataset, ICA method, k=5)

Enrichment Score

Gene set enrichment analysis results using CellMatch DB

V1



ES

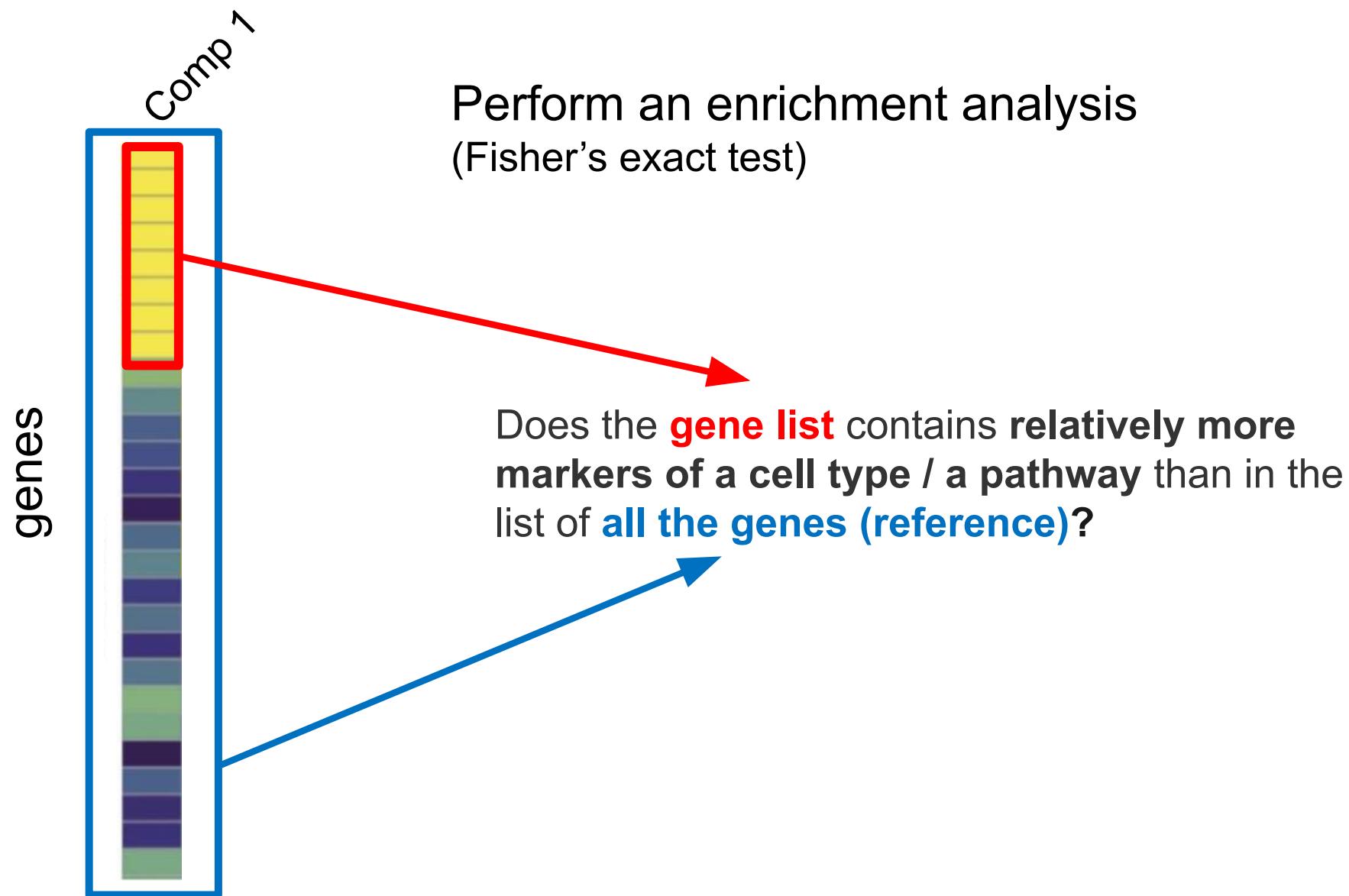
0.975
0.950
0.925
0.900

-log10(pval)

P-value of the GSEA test

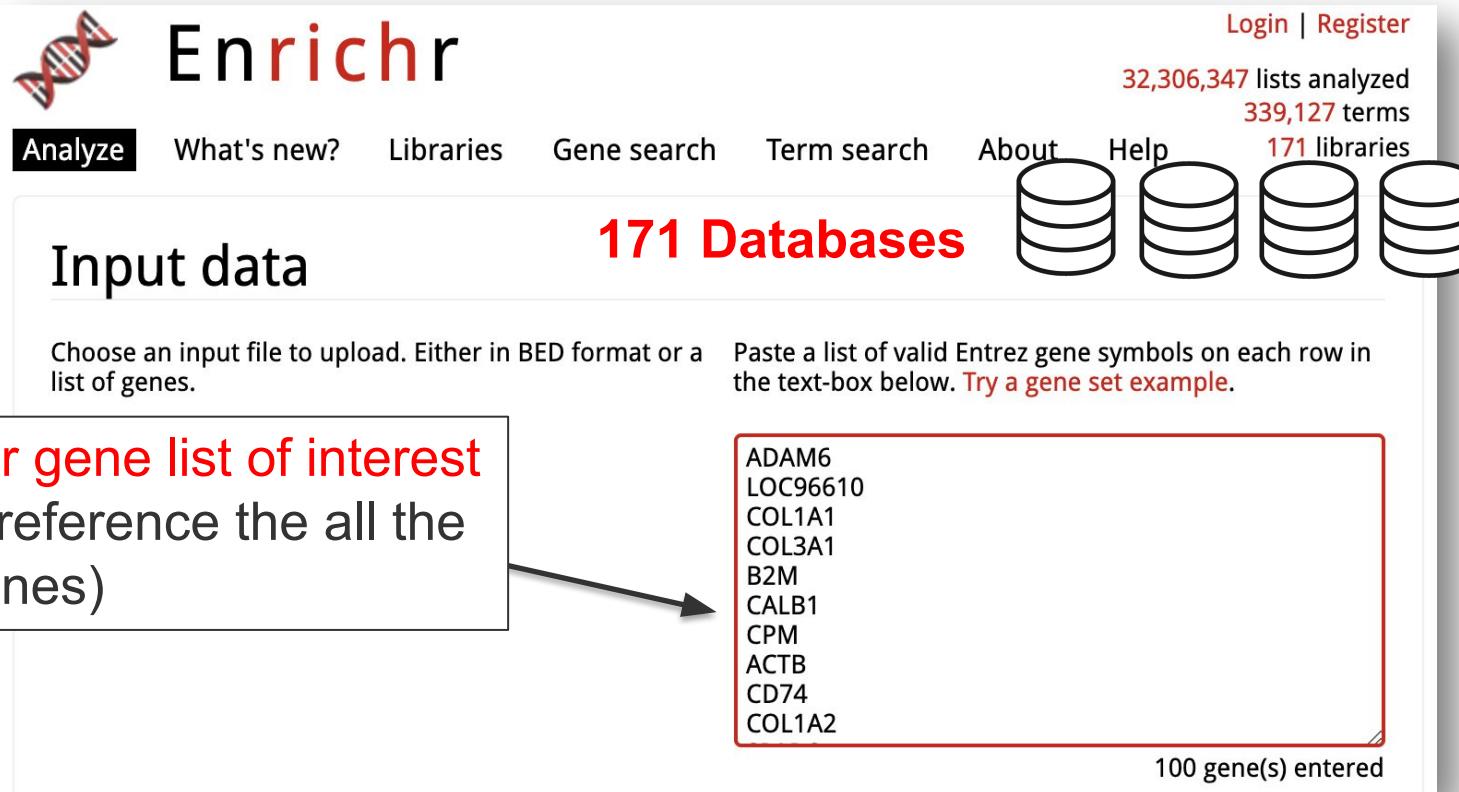
Stromal component, Fibroblast?

Cell types from the
CellMatch DataBase

Other option

Output examples

Enrichment analysis with Enrichr



The screenshot shows the Enrichr homepage. At the top right, it displays "32,306,347 lists analyzed" and "339,127 terms". Below this, there are four icons representing databases, labeled "171 libraries". The main area features a large red banner with the text "171 Databases". To the left, a box contains the text "Input data" and "Choose an input file to upload. Either in BED format or a list of genes." To the right, another box contains the text "Paste a list of valid Entrez gene symbols on each row in the text-box below. Try a gene set example." Below these boxes is a text area containing a list of genes: ADAM6, LOC96610, COL1A1, COL3A1, B2M, CALB1, CPM, ACTB, CD74, and COL1A2. A red border surrounds this list, and the text "100 gene(s) entered" is at the bottom.

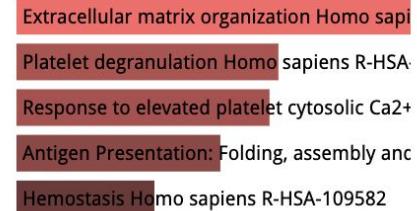
Paste your gene list of interest
(takes as reference the all the human genes)

171 Databases

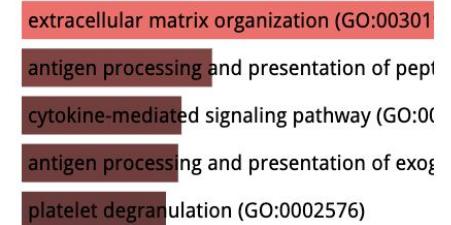
ADAM6
LOC96610
COL1A1
COL3A1
B2M
CALB1
CPM
ACTB
CD74
COL1A2

100 gene(s) entered

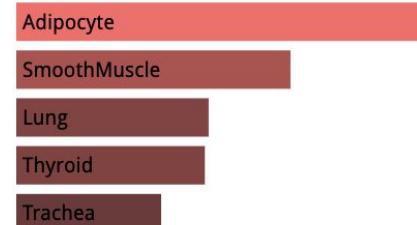
Reactome 2016



GO Biological Process 2018



Human Gene Atlas



Perform enrichment analyses with other external tools

Feature/Tool	DAVID	Enrichr	ToppGene	g:profiler	clusterProfiler	Goplot	BACA	FunMappOne
KEGG pathways	✓	✓	✓	✓	✓		✓	✓
Reactome pathways	✓	✓	✓	✓	✓			✓
Gene Ontology	✓	✓	✓	✓	✓	✓	✓	✓
Graphic representation		✓	✓		✓	✓	✓	✓
Graphic user interface	✓	✓		✓				✓

compExplore Shiny app



INPUT

Gene signature matrix

Input as a csv file
Format: separator = ";" decimal = ","

Browse... results_T_1.csv

Upload complete

The Gene signature matrix corresponds to the output results_T_1.csv in the comet web-app which is already in the requested format (separator = ";" decimal = ",").

Deconvolution method

ICA-based

NMF-based

Cancer type

ALL

Download top markers

Top100 gene markers for each component

Download

OUTPUT

Enrichment analysis

Gene set enrichment analysis results using CellMa

V1

pathway

Non_Small_Cell_Lung_Cancer.Fibroblast
Normal_Fibroblast
Colorectal_Cancer.CancerAssociated_Fibroblast
Head_and_Neck_Cancer.Fibroblast
Oligodendroglioma.Microglial_Cell
Glioma.Astrocyte
B_Cell.Lymphoma.B_Cell
Normal.Circulating_Fetal_Cell
Melanoma.CancerAssociated_Fibroblast
Normal.Stromal_Cell

-log10(pval)

V3

pathway

Normal_1Cell_Stage_Cell_Blastomere
Normal_Secretory_Cell
Ovarian_Cancer.Cancer_Cell
Normal_Alpha_Cell
Astrocytoma.Astrocyte
Melanoma.Macrophage
Colon_Cancer.Stem_Cell
Glioma.Astrocyte
Normal_Mast_Cell
Renal_Cell_Carcinoma.Neutrophil

-log10(pval)

V5

pathway

Colorectal_Cancer.CancerAssociated_Fibroblast
Non_Small_Cell_Lung_Cancer.Fibroblast
Normal_Myofibroblast
Ovarian_Cancer.Mesenchymal_Cell
Normal_Mesangial_Cell
Normal_Bile_Duct_Cell
Head_and_Neck_Cancer.Fibroblast
Normal_Fibroblast
Normal_Idiopathic_Pulmonary_Fibrosis_Cell
Normal_Pneumocyte

-log10(pval)

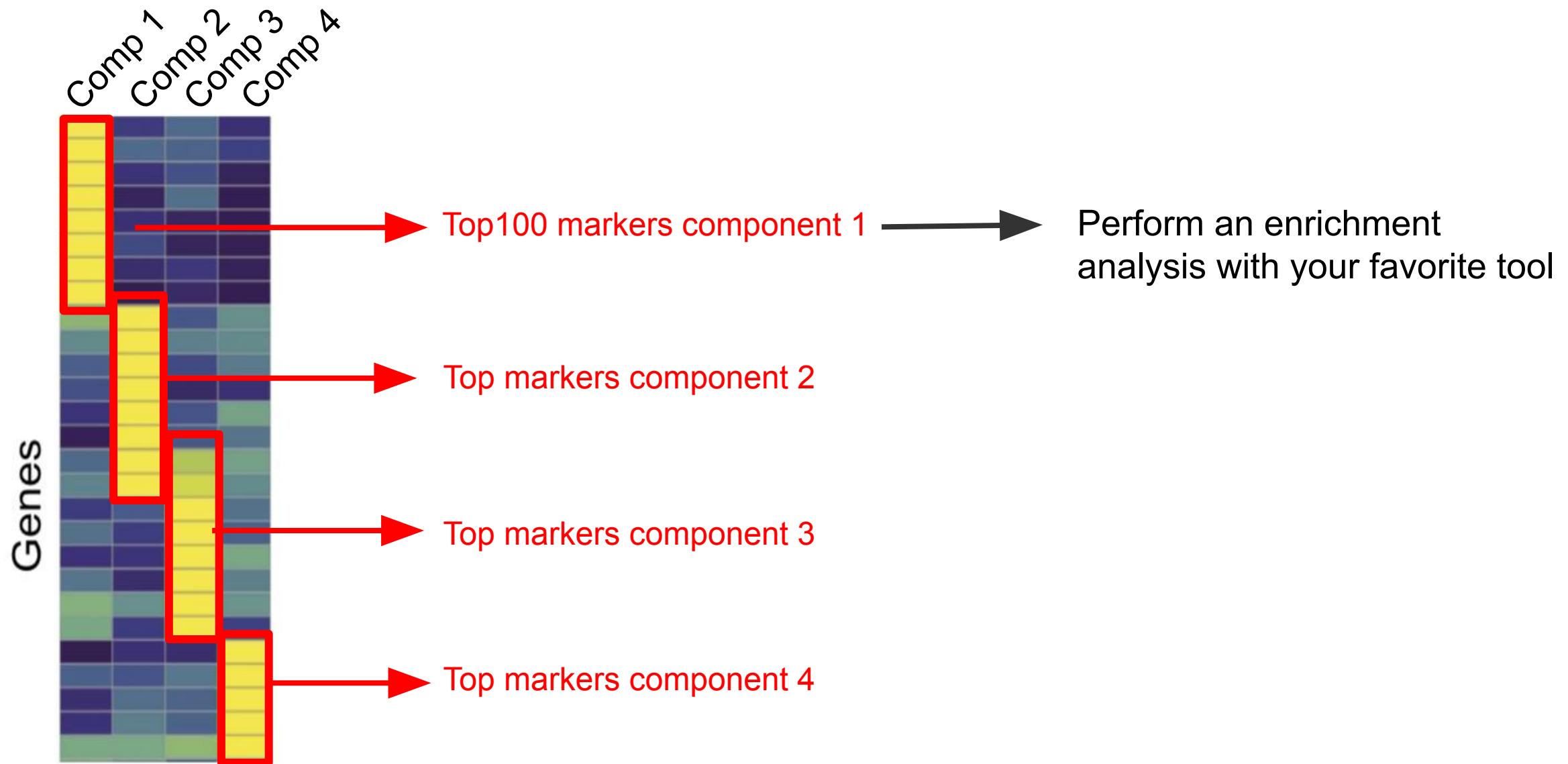
Components enrichment

Proportion vizualisation

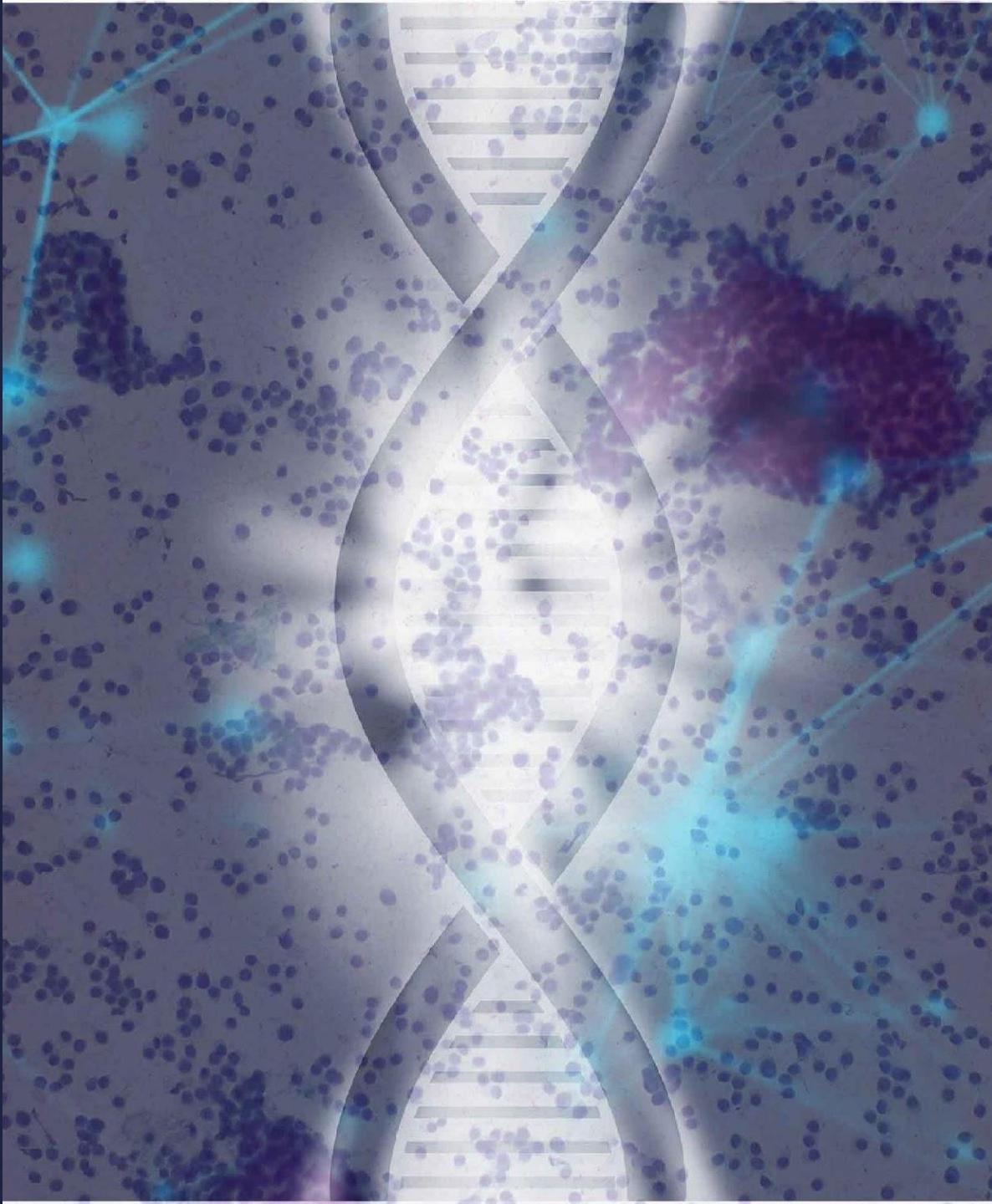
Csv file

	A	B	C	D	E	F
1	V1	V2	V3	V4	V5	
2	1 ADAM6	CLU	SFTPB	LOC96610	H19	
3	2 LOC96610	CALB1	FTL	CLU	FN1	
4	3 COL1A1	PCSK2	CLU	PLUNC	COL1A1	
5	4 COL3A1	PGC	SFTPA2	COL1A1	COL3A1	
6	5 B2M	SCGB3A2	TPT1	FGA	COL1A2	
7	6 CALB1	MYCL1	NAPSA	CALCA	PLUNC	
8	7 CPM	SCN3A	CEACAM6	FGG	IGF2	
9	8 ACTB	BAI1	CTSD	COL3A1	SPARC	
10	9 CD74	C16orf89	PLUNC	FGB	SFTPB	
11	10 COL1A2	GKN2	EEF1A1	FN1	ADAM6	
12	11 SPARC	GP2	CALCA	CEACAM6	CHGB	
13	12 TMSL3	AMBP	MSLN	GAPDH	CEACAM5	
14	13 HSP90B1	TMEM59L	FTH1	COL1A2	CALCA	
15	14 HLA-B	HPCAL4	AKR1C1	MUC5B	CEACAM6	
16	15 IGJ	GRIK1	MUC5B	SFTPC	COL6A3	
17	16 PABPC1	OBP2A	P4HB	SFTPA2	SLC34A2	
18	17 ACTG1	C1orf95	ACTG1	FTL	TMSL3	
19	18 PSAP	CHRDL2	PCSK2	CTSD	SFTPA2	
20	19 GAPDH	ADHFE1	SFTPC	CALB1	FLNA	
21	20 HLA-A	PCDHGA4	MUC1	SLC34A2	TIMP3	
22	21 HLA-DRA	GLYATL3	PABPC1	CPM	CPM	
23	22 CEACAM6	KRT40	LGALS3BP	SPARC	HMGB3	
24	23 LUM	ADRB1	EEF2	PCSK2	ODC1	
25	24 CCT2	LCN15	SFTPA1	MSLN	ATP1A1	
26	25 UBC	PLA2G10	FGB	CEACAM5	S100A6	
27	26 HLA-C	CBLN2	ACTB	HP	NDRG1	
28	27 KRT7	SCN2A	SCGB3A2	SFTPA1	GNAS	
29	28 BGN	LIMS3-LOC44	FGA	ENO1	VIM	
30	29 CALR	ITLN2	RPL8	PCSK1	SFTPC	
31	30 DHHR	STAG2	CDF2	HSDBOR1	RGN	

Interpret the components identified

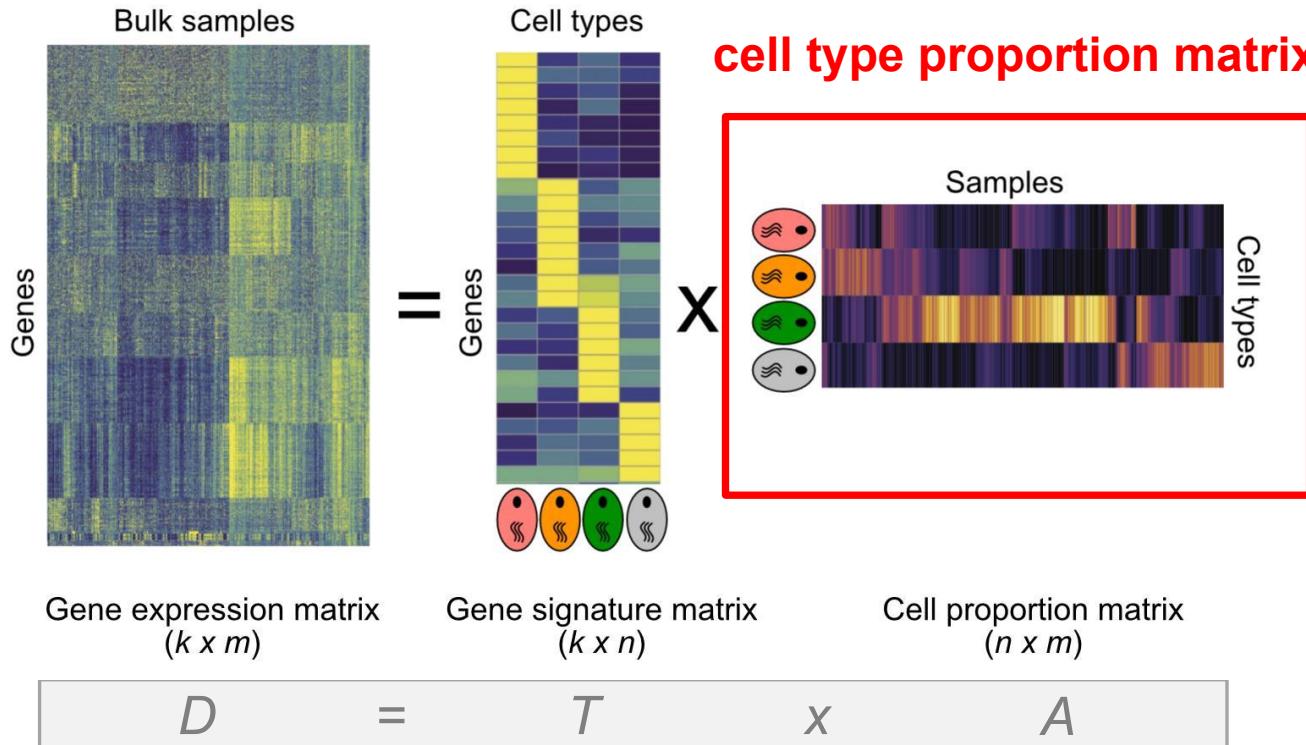


Visualize the cell type proportion matrix



Visualize the cell type proportion matrix

Credits A. Kumar



compExplore Shiny app



compExplore [About](#) [CSV-convertor](#) [Number of CellTypes](#) [Components enrichment](#) [Proportion vizualisation](#)

Cell proportion matrix

Input as a csv file
Format: separator = ";" decimal = ","

PUT

[Browse...](#) No file selected

The Gene proportion matrix corresponds to the output results_A_1.csv in the cometh web-app which is already in the requested format (separator = ";" decimal = ",").

Select samples

No choices here yet !!

Select cell types

No choices here yet !!

Guide **Vizualisation plots**

OUPUT

Heatmap of the proportion matrix

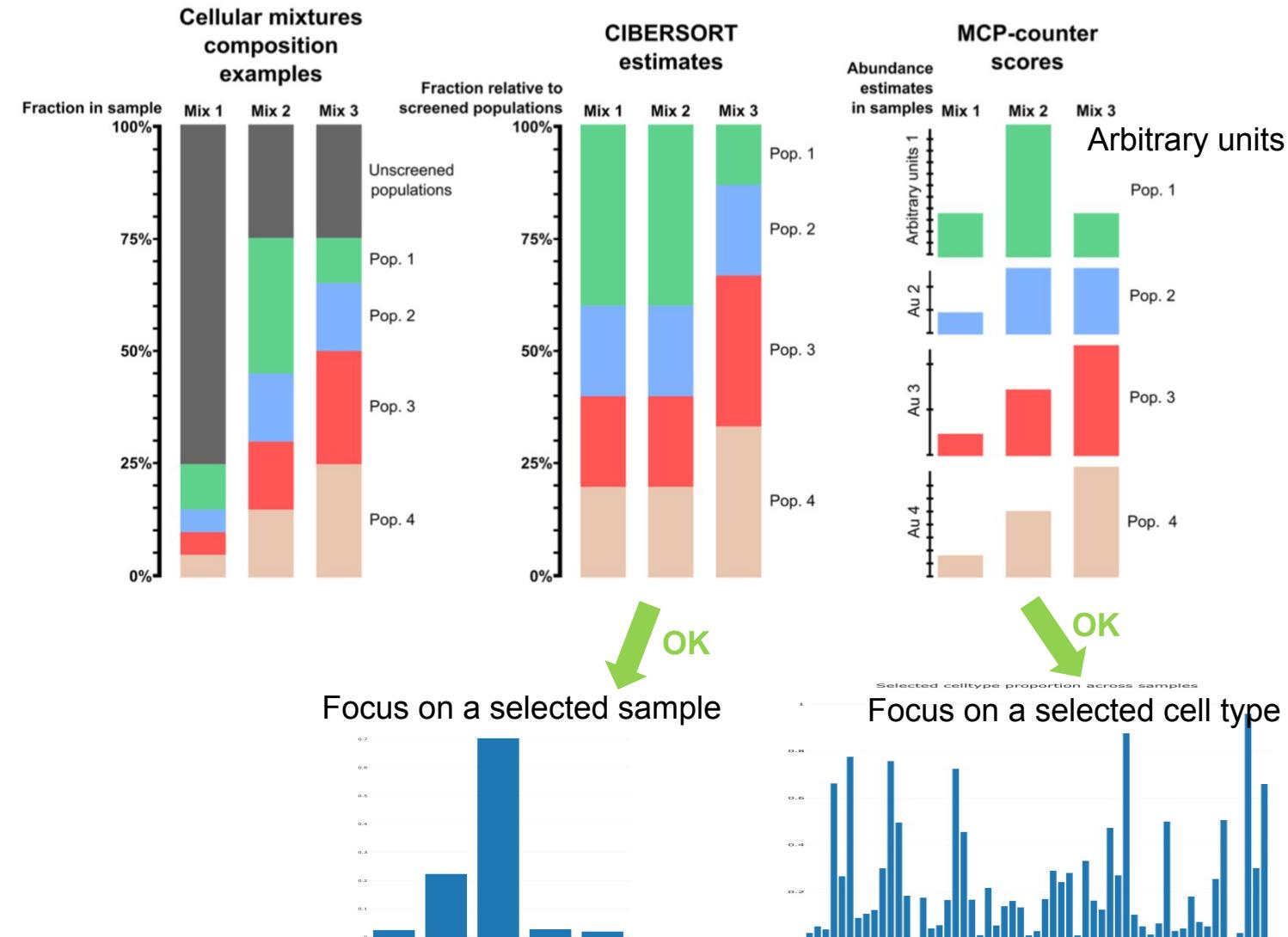
A heatmap visualization titled "Heatmap of the proportion matrix". The y-axis is labeled with sample names: IC2, IC1, IC5, IC4, and IC3. The x-axis lists various cell types: T0, 94, 91A, 05A, 05B, 06A, 06B, 07A, 07B, 08A, 08B, 09A, 09B, 10A, 10B, 11A, 11B, 12A, 12B, 13A, 13B, 14A, 14B, 15A, 15B, 16A, 16B, 17A, 17B, 18A, 18B, 19A, 19B, 20A, 20B, 21A, 21B, 22A, 22B, 23A, 23B, 24A, 24B, 25A, 25B, 26A, 26B, 27A, 27B, 28A, 28B, 29A, 29B, 30A, 30B, 31A, 31B, 32A, 32B, 33A, 33B, 34A, 34B, 35A, 35B, 36A, 36B, 37A, 37B, 38A, 38B, 39A, 39B, 40A, 40B, 41A, 41B, 42A, 42B, 43A, 43B, 44A, 44B, 45A, 45B, 46A, 46B, 47A, 47B, 48A, 48B, 49A, 49B, 50A, 50B, 51A, 51B, 52A, 52B, 53A, 53B, 54A, 54B, 55A, 55B, 56A, 56B, 57A, 57B, 58A, 58B, 59A, 59B, 60A, 60B, 61A, 61B, 62A, 62B, 63A, 63B, 64A, 64B, 65A, 65B, 66A, 66B, 67A, 67B, 68A, 68B, 69A, 69B, 70A, 70B, 71A, 71B, 72A, 72B, 73A, 73B, 74A, 74B, 75A, 75B, 76A, 76B, 77A, 77B, 78A, 78B, 79A, 79B, 80A, 80B, 81A, 81B, 82A, 82B, 83A, 83B, 84A, 84B, 85A, 85B, 86A, 86B, 87A, 87B, 88A, 88B, 89A, 89B, 90A, 90B, 91A, 91B, 92A, 92B, 93A, 93B, 94A, 94B, 95A, 95B, 96A, 96B, 97A, 97B, 98A, 98B, 99A, 99B, 100A, 100B, 101A, 101B, 102A, 102B, 103A, 103B, 104A, 104B, 105A, 105B, 106A, 106B, 107A, 107B, 108A, 108B, 109A, 109B, 110A, 110B, 111A, 111B, 112A, 112B, 113A, 113B, 114A, 114B, 115A, 115B, 116A, 116B, 117A, 117B, 118A, 118B, 119A, 119B, 120A, 120B, 121A, 121B, 122A, 122B, 123A, 123B, 124A, 124B, 125A, 125B, 126A, 126B, 127A, 127B, 128A, 128B, 129A, 129B, 130A, 130B, 131A, 131B, 132A, 132B, 133A, 133B, 134A, 134B, 135A, 135B, 136A, 136B, 137A, 137B, 138A, 138B, 139A, 139B, 140A, 140B, 141A, 141B, 142A, 142B, 143A, 143B, 144A, 144B, 145A, 145B, 146A, 146B, 147A, 147B, 148A, 148B, 149A, 149B, 150A, 150B, 151A, 151B, 152A, 152B, 153A, 153B, 154A, 154B, 155A, 155B, 156A, 156B, 157A, 157B, 158A, 158B, 159A, 159B, 160A, 160B, 161A, 161B, 162A, 162B, 163A, 163B, 164A, 164B, 165A, 165B, 166A, 166B, 167A, 167B, 168A, 168B, 169A, 169B, 170A, 170B, 171A, 171B, 172A, 172B, 173A, 173B, 174A, 174B, 175A, 175B, 176A, 176B, 177A, 177B, 178A, 178B, 179A, 179B, 180A, 180B, 181A, 181B, 182A, 182B, 183A, 183B, 184A, 184B, 185A, 185B, 186A, 186B, 187A, 187B, 188A, 188B, 189A, 189B, 190A, 190B, 191A, 191B, 192A, 192B, 193A, 193B, 194A, 194B, 195A, 195B, 196A, 196B, 197A, 197B, 198A, 198B, 199A, 199B, 200A, 200B, 201A, 201B, 202A, 202B, 203A, 203B, 204A, 204B, 205A, 205B, 206A, 206B, 207A, 207B, 208A, 208B, 209A, 209B, 210A, 210B, 211A, 211B, 212A, 212B, 213A, 213B, 214A, 214B, 215A, 215B, 216A, 216B, 217A, 217B, 218A, 218B, 219A, 219B, 220A, 220B, 221A, 221B, 222A, 222B, 223A, 223B, 224A, 224B, 225A, 225B, 226A, 226B, 227A, 227B, 228A, 228B, 229A, 229B, 230A, 230B, 231A, 231B, 232A, 232B, 233A, 233B, 234A, 234B, 235A, 235B, 236A, 236B, 237A, 237B, 238A, 238B, 239A, 239B, 240A, 240B, 241A, 241B, 242A, 242B, 243A, 243B, 244A, 244B, 245A, 245B, 246A, 246B, 247A, 247B, 248A, 248B, 249A, 249B, 250A, 250B, 251A, 251B, 252A, 252B, 253A, 253B, 254A, 254B, 255A, 255B, 256A, 256B, 257A, 257B, 258A, 258B, 259A, 259B, 260A, 260B, 261A, 261B, 262A, 262B, 263A, 263B, 264A, 264B, 265A, 265B, 266A, 266B, 267A, 267B, 268A, 268B, 269A, 269B, 270A, 270B, 271A, 271B, 272A, 272B, 273A, 273B, 274A, 274B, 275A, 275B, 276A, 276B, 277A, 277B, 278A, 278B, 279A, 279B, 280A, 280B, 281A, 281B, 282A, 282B, 283A, 283B, 284A, 284B, 285A, 285B, 286A, 286B, 287A, 287B, 288A, 288B, 289A, 289B, 290A, 290B, 291A, 291B, 292A, 292B, 293A, 293B, 294A, 294B, 295A, 295B, 296A, 296B, 297A, 297B, 298A, 298B, 299A, 299B, 300A, 300B, 301A, 301B, 302A, 302B, 303A, 303B, 304A, 304B, 305A, 305B, 306A, 306B, 307A, 307B, 308A, 308B, 309A, 309B, 310A, 310B, 311A, 311B, 312A, 312B, 313A, 313B, 314A, 314B, 315A, 315B, 316A, 316B, 317A, 317B, 318A, 318B, 319A, 319B, 320A, 320B, 321A, 321B, 322A, 322B, 323A, 323B, 324A, 324B, 325A, 325B, 326A, 326B, 327A, 327B, 328A, 328B, 329A, 329B, 330A, 330B, 331A, 331B, 332A, 332B, 333A, 333B, 334A, 334B, 335A, 335B, 336A, 336B, 337A, 337B, 338A, 338B, 339A, 339B, 340A, 340B, 341A, 341B, 342A, 342B, 343A, 343B, 344A, 344B, 345A, 345B, 346A, 346B, 347A, 347B, 348A, 348B, 349A, 349B, 350A, 350B, 351A, 351B, 352A, 352B, 353A, 353B, 354A, 354B, 355A, 355B, 356A, 356B, 357A, 357B, 358A, 358B, 359A, 359B, 360A, 360B, 361A, 361B, 362A, 362B, 363A, 363B, 364A, 364B, 365A, 365B, 366A, 366B, 367A, 367B, 368A, 368B, 369A, 369B, 370A, 370B, 371A, 371B, 372A, 372B, 373A, 373B, 374A, 374B, 375A, 375B, 376A, 376B, 377A, 377B, 378A, 378B, 379A, 379B, 380A, 380B, 381A, 381B, 382A, 382B, 383A, 383B, 384A, 384B, 385A, 385B, 386A, 386B, 387A, 387B, 388A, 388B, 389A, 389B, 390A, 390B, 391A, 391B, 392A, 392B, 393A, 393B, 394A, 394B, 395A, 395B, 396A, 396B, 397A, 397B, 398A, 398B, 399A, 399B, 400A, 400B, 401A, 401B, 402A, 402B, 403A, 403B, 404A, 404B, 405A, 405B, 406A, 406B, 407A, 407B, 408A, 408B, 409A, 409B, 410A, 410B, 411A, 411B, 412A, 412B, 413A, 413B, 414A, 414B, 415A, 415B, 416A, 416B, 417A, 417B, 418A, 418B, 419A, 419B, 420A, 420B, 421A, 421B, 422A, 422B, 423A, 423B, 424A, 424B, 425A, 425B, 426A, 426B, 427A, 427B, 428A, 428B, 429A, 429B, 430A, 430B, 431A, 431B, 432A, 432B, 433A, 433B, 434A, 434B, 435A, 435B, 436A, 436B, 437A, 437B, 438A, 438B, 439A, 439B, 440A, 440B, 441A, 441B, 442A, 442B, 443A, 443B, 444A, 444B, 445A, 445B, 446A, 446B, 447A, 447B, 448A, 448B, 449A, 449B, 450A, 450B, 451A, 451B, 452A, 452B, 453A, 453B, 454A, 454B, 455A, 455B, 456A, 456B, 457A, 457B, 458A, 458B, 459A, 459B, 460A, 460B, 461A, 461B, 462A, 462B, 463A, 463B, 464A, 464B, 465A, 465B, 466A, 466B, 467A, 467B, 468A, 468B, 469A, 469B, 470A, 470B, 471A, 471B, 472A, 472B, 473A, 473B, 474A, 474B, 475A, 475B, 476A, 476B, 477A, 477B, 478A, 478B, 479A, 479B, 480A, 480B, 481A, 481B, 482A, 482B, 483A, 483B, 484A, 484B, 485A, 485B, 486A, 486B, 487A, 487B, 488A, 488B, 489A, 489B, 490A, 490B, 491A, 491B, 492A, 492B, 493A, 493B, 494A, 494B, 495A, 495B, 496A, 496B, 497A, 497B, 498A, 498B, 499A, 499B, 500A, 500B, 501A, 501B, 502A, 502B, 503A, 503B, 504A, 504B, 505A, 505B, 506A, 506B, 507A, 507B, 508A, 508B, 509A, 509B, 510A, 510B, 511A, 511B, 512A, 512B, 513A, 513B, 514A, 514B, 515A, 515B, 516A, 516B, 517A, 517B, 518A, 518B, 519A, 519B, 520A, 520B, 521A, 521B, 522A, 522B, 523A, 523B, 524A, 524B, 525A, 525B, 526A, 526B, 527A, 527B, 528A, 528B, 529A, 529B, 530A, 530B, 531A, 531B, 532A, 532B, 533A, 533B, 534A, 534B, 535A, 535B, 536A, 536B, 537A, 537B, 538A, 538B, 539A, 539B, 540A, 540B, 541A, 541B, 542A, 542B, 543A, 543B, 544A, 544B, 545A, 545B, 546A, 546B, 547A, 547B, 548A, 548B, 549A, 549B, 550A, 550B, 551A, 551B, 552A, 552B, 553A, 553B, 554A, 554B, 555A, 555B, 556A, 556B, 557A, 557B, 558A, 558B, 559A, 559B, 560A, 560B, 561A, 561B, 562A, 562B, 563A, 563B, 564A, 564B, 565A, 565B, 566A, 566B, 567A, 567B, 568A, 568B, 569A, 569B, 570A, 570B, 571A, 571B, 572A, 572B, 573A, 573B, 574A, 574B, 575A, 575B, 576A, 576B, 577A, 577B, 578A, 578B, 579A, 579B, 580A, 580B, 581A, 581B, 582A, 582B, 583A, 583B, 584A, 584B, 585A, 585B, 586A, 586B, 587A, 587B, 588A, 588B, 589A, 589B, 590A, 590B, 591A, 591B, 592A, 592B, 593A, 593B, 594A, 594B, 595A, 595B, 596A, 596B, 597A, 597B, 598A, 598B, 599A, 599B, 600A, 600B, 601A, 601B, 602A, 602B, 603A, 603B, 604A, 604B, 605A, 605B, 606A, 606B, 607A, 607B, 608A, 608B, 609A, 609B, 610A, 610B, 611A, 611B, 612A, 612B, 613A, 613B, 614A, 614B, 615A, 615B, 616A, 616B, 617A, 617B, 618A, 618B, 619A, 619B, 620A, 620B, 621A, 621B, 622A, 622B, 623A, 623B, 624A, 624B, 625A, 625B, 626A, 626B, 627A, 627B, 628A, 628B, 629A, 629B, 630A, 630B, 631A, 631B, 632A, 632B, 633A, 633B, 634A, 634B, 635A, 635B, 636A, 636B, 637A, 637B, 638A, 638B, 639A, 639B, 640A, 640B, 641A, 641B, 642A, 642B, 643A, 643B, 644A, 644B, 645A, 645B, 646A, 646B, 647A, 647B, 648A, 648B, 649A, 649B, 650A, 650B, 651A, 651B, 652A, 652B, 653A, 653B, 654A, 654B, 655A, 655B, 656A, 656B, 657A, 657B, 658A, 658B, 659A, 659B, 660A, 660B, 661A, 661B, 662A, 662B, 663A, 663B, 664A, 664B, 665A, 665B, 666A, 666B, 667A, 667B, 668A, 668B, 669A, 669B, 670A, 670B, 671A, 671B, 672A, 672B, 673A, 673B, 674A, 674B, 675A, 675B, 676A, 676B, 677A, 677B, 678A, 678B, 679A, 679B, 680A, 680B, 681A, 681B, 682A, 682B, 683A, 683B, 684A, 684B, 685A, 685B, 686A, 686B, 687A, 687B, 688A, 688B, 689A, 689B, 690A, 690B, 691A, 691B, 692A, 692B, 693A, 693B, 694A, 694B, 695A, 695B, 696A, 696B, 697A, 697B, 698A, 698B, 699A, 699B, 700A, 700B, 701A, 701B, 702A, 702B, 703A, 703B, 704A, 704B, 705A, 705B, 706A, 706B, 707A, 707B, 708A, 708B, 709A, 709B, 710A, 710B, 711A, 711B, 712A, 712B, 713A, 713B, 714A, 714B, 715A, 715B, 716A, 716B, 717A, 717B, 718A, 718B, 719A, 719B, 720A, 720B, 721A, 721B, 722A, 722B, 723A, 723B, 724A, 724B, 725A, 725B, 726A, 726B, 727A, 727B, 728A, 728B, 729A, 729B, 730A, 730B, 731A, 731B, 732A, 732B, 733A, 733B, 734A, 734B, 735A, 735B, 736A, 736B, 737A, 737B, 738A, 738B, 739A, 739B, 740A, 740B, 741A, 741B, 742A, 742B, 743A, 743B, 744A, 744B, 745A, 745B, 746A, 746B, 747A, 747B, 748A, 748B, 749A, 749B, 750A, 750B, 751A, 751B, 752A, 752B, 753A, 753B, 754A, 754B, 755A, 755B, 756A, 756B, 757A, 757B, 758A, 758B, 759A, 759B, 760A, 760B, 761A, 761B, 762A, 762B, 763A, 763B, 764A, 764B, 765A, 765B, 766A, 766B, 767A, 767B, 768A, 768B, 769A, 769B, 770A, 770B, 771A, 771B, 772A, 772B, 773A, 773B, 774A, 774B, 775A, 775B, 776A, 776B, 777A, 777B, 778A, 778B, 779A, 779B, 780A, 780B, 781A, 781B, 782A, 782B, 783A, 783B, 784A, 784B, 785A, 785B, 786A, 786B, 787A, 787B, 788A, 788B, 789A, 789B, 790A, 790B, 791A, 791B, 792A, 792B, 793A, 793B, 794A, 794B, 795A, 795B, 796A, 796B, 797A, 797B, 798A, 798B, 799A, 799B, 800A, 800B, 801A, 801B, 802A, 802B, 803A, 803B, 804A, 804B, 805A, 805B, 806A, 806B, 807A, 807B, 808A, 808B, 809A, 809B, 810A, 810B, 811A, 811B, 812A, 812B, 813A, 813B, 814A, 814B, 815A, 815B, 816A, 816B, 817A, 817B, 818A, 818B, 819A, 819B, 820A, 820B, 821A, 821B, 822A, 822B, 823A, 823B, 824A, 824B, 825A, 825B, 826A, 826B, 827A, 827B, 828A, 828B, 829A, 829B, 830A, 830B, 831A, 831B, 832A, 832B, 833A, 833B, 834A, 834B, 835A, 835B, 836A, 836B, 837A, 837B, 838A, 838B, 839A, 839B, 840A, 840B, 841A, 841B, 842A, 842B, 843A, 843B, 844A, 844B, 845A, 845B, 846A, 846B, 847A, 847B, 848A, 848B, 849A, 849B, 850A, 850B, 851A, 851B, 852A, 852B, 853A, 853B, 854A, 854B, 855A, 855B, 856A, 856B, 857A, 857B, 858A, 858B, 859A, 859B, 860A, 860B, 861A, 861B, 862A, 862B, 863A, 863B, 864A, 864B, 865A, 865B, 866A, 866B, 867A, 867B, 868A, 868B, 869A, 869B, 870A, 870B, 871A, 871B, 872A, 872B, 873A, 873B, 874A, 874B, 875A, 875B, 876A, 876B, 877A, 877B, 878A, 878B, 879A, 879B, 880A, 880B, 881A, 881B, 882A, 882B, 883A, 883B, 884A, 884B, 885A, 885B, 886A, 886B, 887A, 887B, 888A, 888B, 889A, 889B, 890A, 890B, 891A, 891B, 892A, 892B, 893A, 893B, 894A, 894B, 895A, 895B, 896A, 896B, 897A, 897B, 898A, 898B, 899A, 899B, 900A, 900B, 901A, 901B, 902A, 902B, 903A, 903B, 904A, 904B, 905A, 905B, 906A, 906B, 907A, 907B, 908A, 908B, 909A, 909B, 910A, 910B, 911A, 911B, 912A, 912B, 913A, 913B, 914A, 914B, 915A, 915B, 916A, 916B, 917A, 917B, 918A, 918B, 919A, 919B, 920A, 920B, 921A, 921B, 922A, 922B, 923A, 923B, 924A, 924B, 925A, 925B, 926A, 926B, 927A, 927B, 928A, 928B, 929A, 929B, 930A, 930B, 931A, 931B, 932A, 932B, 933A, 933B, 934A, 934B, 935A, 935B, 936A, 936B, 937A, 937B, 938A, 938B, 939A, 939B, 940A, 940B, 941A, 941B, 942A, 942B, 943A, 943B, 944A, 944B, 945A, 945B, 946A, 946B, 947A, 947B, 948A, 948B, 949A, 949B, 950A, 950B, 951A, 951B, 952A, 952B, 953A, 953B, 954A, 954B, 955A, 955B, 956A, 956B, 957A, 957B, 958A, 958B, 959A, 959B, 960A, 960B, 961A, 961B, 962A, 962B, 963A, 963B, 964A, 964B, 965A, 965B, 966A, 966B, 967A, 967B, 968A, 968B, 969A, 969B, 970A, 970B, 971A, 971B, 972A, 972B, 973A, 973B, 974A, 974B, 975A, 975B, 976A, 976B, 977A, 977B, 978A, 978B, 979A, 979B, 980A, 980B, 981A, 981B, 982A, 982B, 983A, 983B, 984A, 984B, 985A, 985B, 986A, 986B, 987A, 987B, 988A, 988B, 989A, 989B, 990A, 990B, 991A, 991B, 992A, 992B, 993A, 993B, 994A, 994B, 995A, 995B, 996A, 996B, 997A, 997B, 998A, 998B, 999A, 999B, 1000A, 1000B, 1001A, 1001B, 1002A, 1002B, 1003A, 1003B, 1004A, 1004B, 1005A, 1005B, 1006A, 1006B, 1007A, 1007B, 1008A, 1008B, 1009A, 1009B, 1010A, 1010B, 1011A, 1011B, 1012A, 1012B, 1013A, 1013B, 1014A, 1014B, 1015A, 1015B, 1016A, 1016B, 1017A, 1017B, 1018A, 1018B, 1019A, 1019B, 1020A, 1020

Visualize the cell type proportion matrix

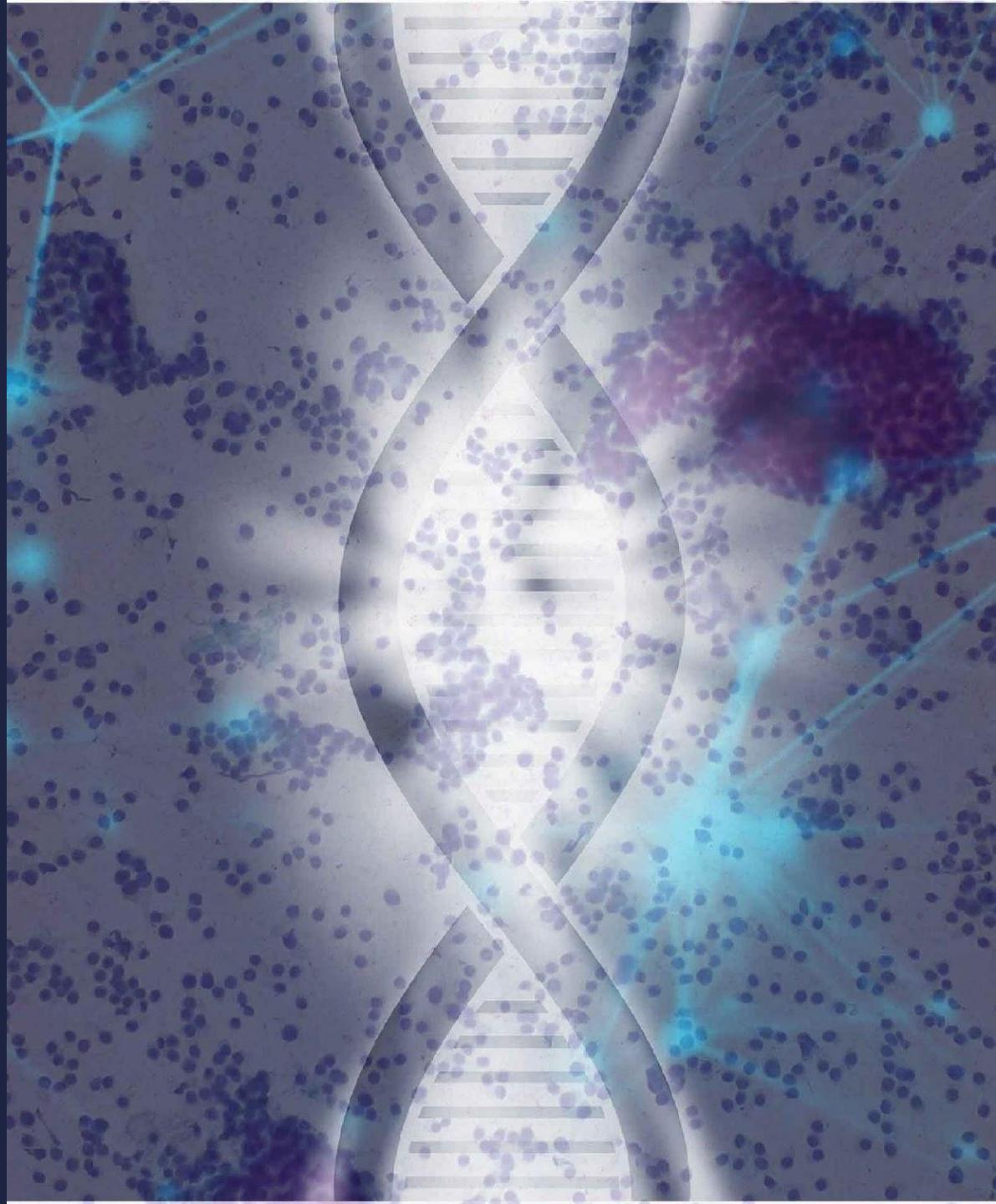


- (1) CIBERSORT-ABS, EPIC and quanTlseq can be used for both **inter- and intra- sample comparisons** i.e. comparing one cell-type within one sample and across samples is possible
- (2) CIBERSORT can be used only for **intra-sample comparisons** i.e. comparing different cell-types within each sample
- (3) MCP-Counter, TIMER and xCell (not provided yet in the comet web app) can be used only for **inter- sample comparisons** i.e. to compare one cell-type across multiple samples

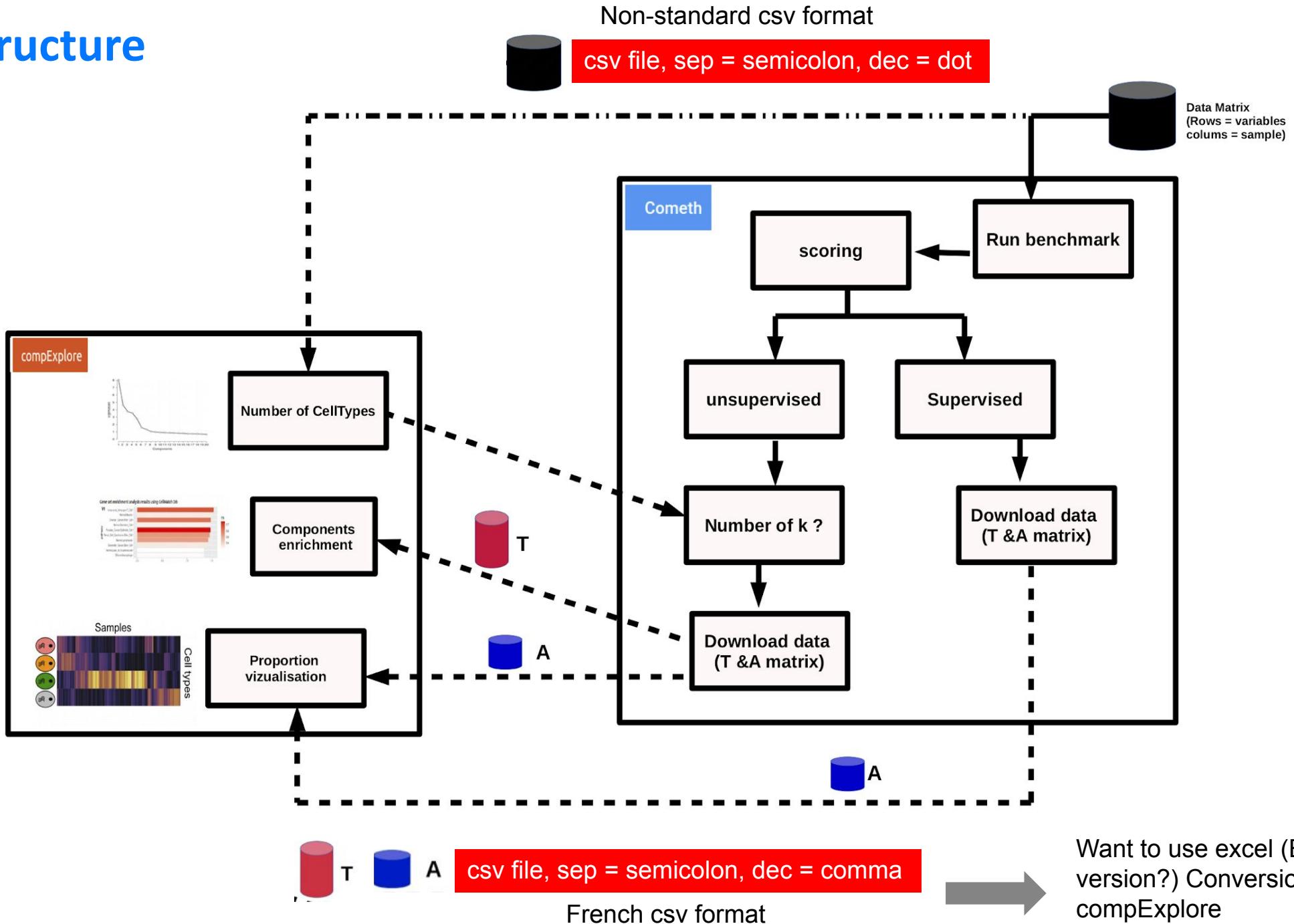
Petitprez et al., 2018 Cancer Immunol Immunother



Overall structure input/output format



Overall structure



compExplore Shiny app

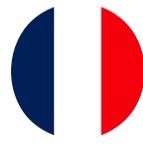


The screenshot shows the compExplore Shiny app interface. At the top, there is a navigation bar with tabs: compExplore, About, CSV-convertisor (which is highlighted in orange), Number of CellTypes, Components enrichment, and Proportion vizualisation.

The main area has two main sections:

- Your csv file**:
 - A file input field with "Browse..." button and "No file selected" message.
 - A note: "Note that output from the cometh web-app are in the french-format (Separator = ";" Decimal = ",")"
 - Separator settings:
 - Semicolon
 - Comma
 - Tab
 - Space
 - Decimal settings:
 - Comma
 - Dot
 - Convert you csv file into:
 - Filename (without csv extension):
 - Format:
 - English
 - French
 - Notes at the bottom: "English - Separator = ;, Decimal = ." and "French - Separator = ;, Decimal = ,,"
- Download your converted csv file**:
 - A "Convert" button with a download icon.
 - A note: "This module can be useful if you are, for instance, using an english version of excel: output of the cometh app are csv files in the french format. Convert them into the english format will allow you to open them directly in excel."

Cometh web app output



CSV

compExplore



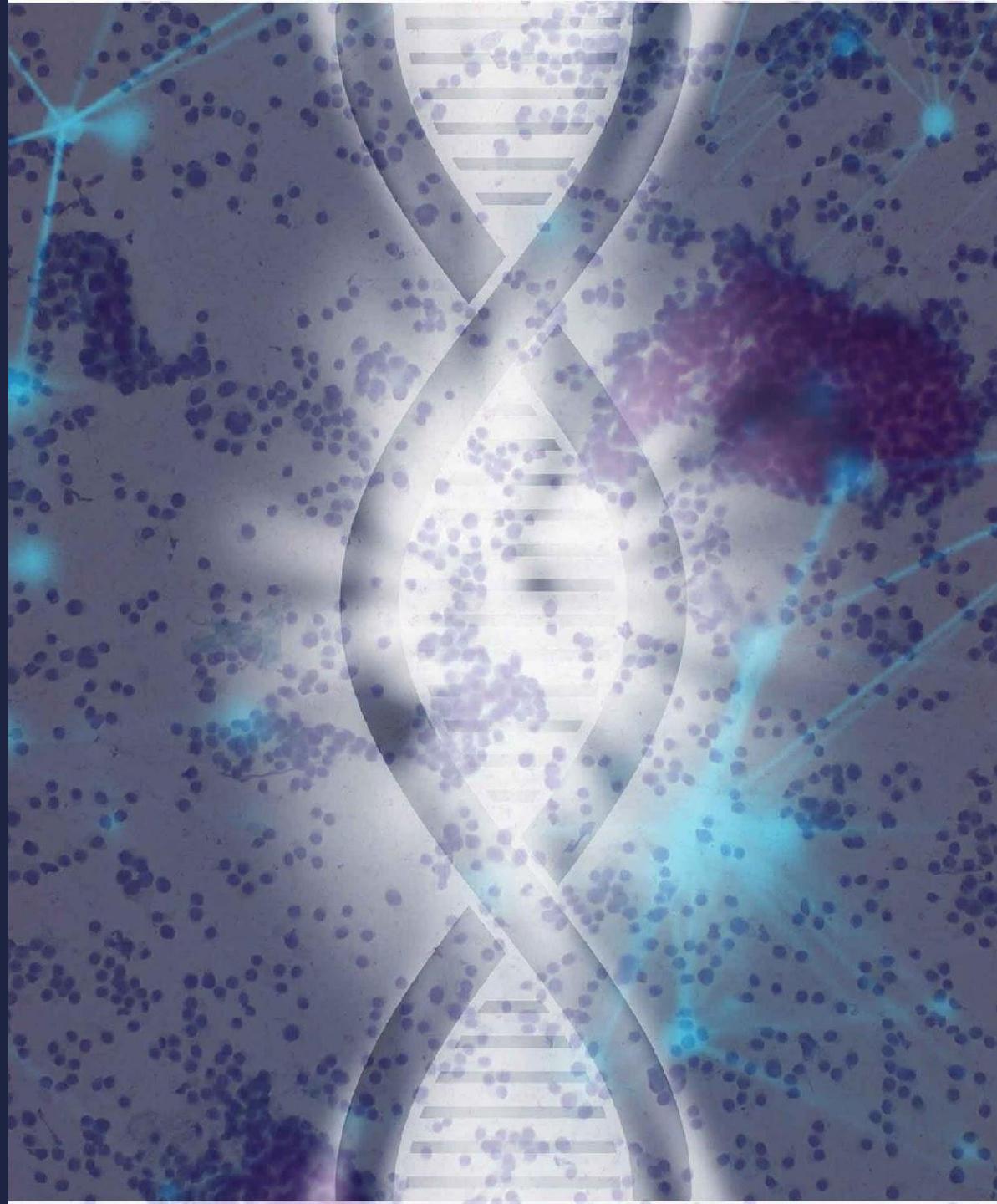
CSV

Separator = Semicolon
Decimal = comma

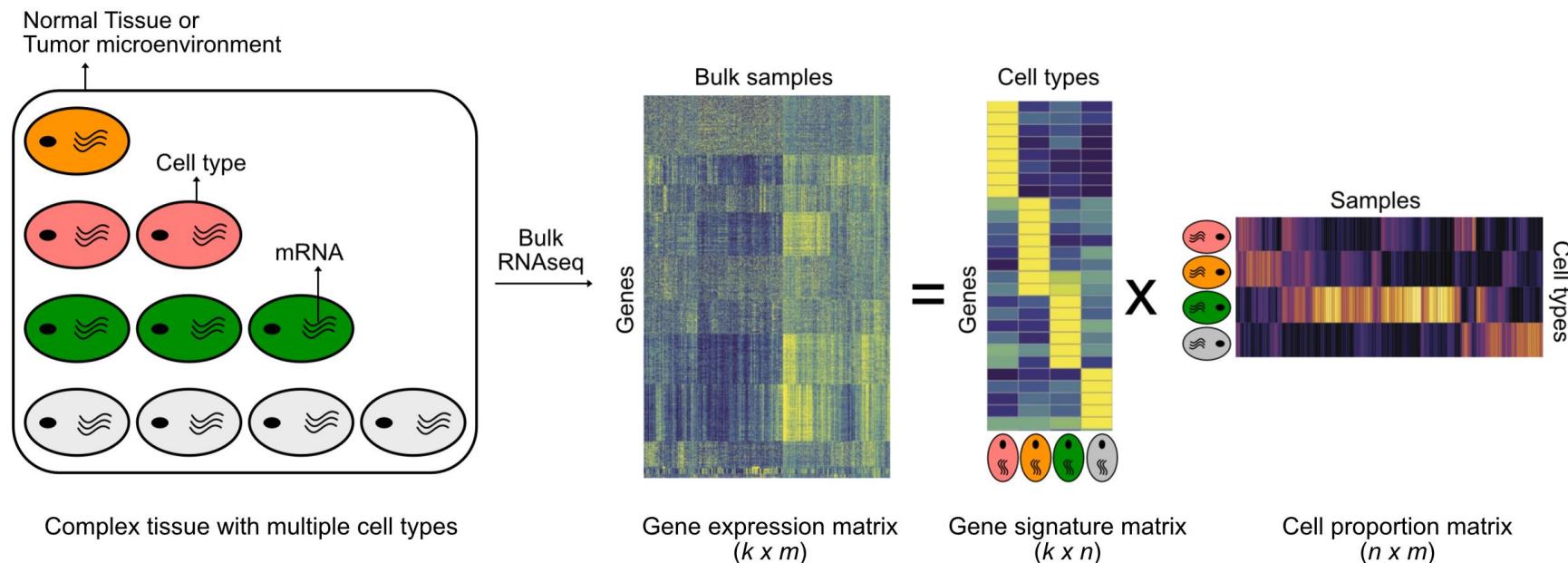
Separator = comma
Decimal = dot



Examples of success stories



Resolving cell types from complex tissue genomic data : RECAP



- **DeconRNAseq** (~160 | Apr, 2013 | <https://doi.org/10.1093/bioinformatics/btt090>)
- **CellMix** (~180 | Sep, 2013 | <https://doi.org/10.1093/bioinformatics/btt351>)
- **CIBERSORT** (~2000 | Mar, 2015 | <https://doi.org/10.1038/nmeth.3337>)
- **MCP-Counter** (~350 | Oct, 2016 | <https://doi.org/10.1186/s13059-016-1070-5>)
- **TIMER2.0** (~500 | Jul, 2017 | <https://doi.org/10.1093/nar/gkaa407>)
- **Xcell** (~400 | Nov, 2017 | <https://doi.org/10.1186/s13059-017-1349-1>)
- **EPIC** (~100 | Nov, 2017 | <https://doi.org/10.7554/eLife.26476>)
- **QuantiSeq** (~50 | May, 2019 | <https://doi.org/10.1186/s13073-019-0638-6>)

Neoantigen-directed immune escape in lung cancer evolution

Rachel Rosenthal, Elizabeth Larose Cadieux, Roberto Salgado, Maise Al Bakir, David A. Moore, Crispin T. Hiley, Tom Lund, Miljana Tanić, James L. Reading, Kroopa Joshi, Jake Y. Henry, Ehsan Ghorani, Gareth A. Wilson, Nicolai J. Birkbak, Mariam Jamal-Hanjani, Selvaraju Veeriah, Zoltan Szallasi, Sherene Loi, Matthew D. Hellmann, Andrew Feber, Benny Chain, Javier Herrero, Sergio A. Quezada, Jonas Demeulemeester, Peter Van Loo, Stephan Beck, Nicholas McGranahan , Charles Swanton  & The TRACERx consortium -Show fewer authors

Nature 567, 479–485(2019) | Cite this article

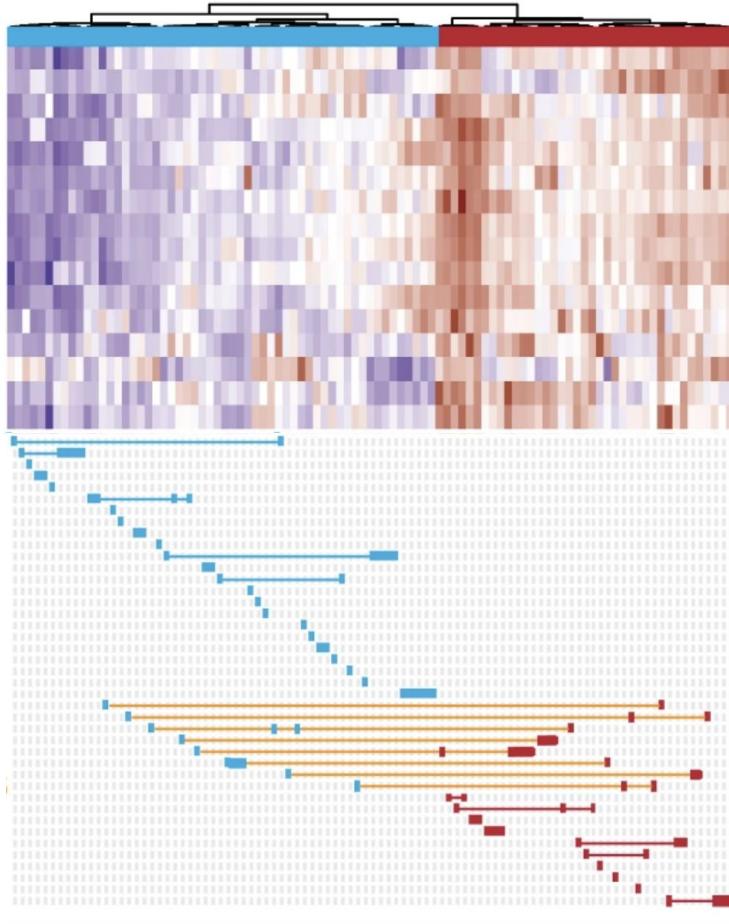
47k Accesses | 163 Citations | 359 Altmetric | Metrics

Cancer Research UK Lung Cancer Centre of Excellence, University College London
Cancer Institute, University College London, London, UK

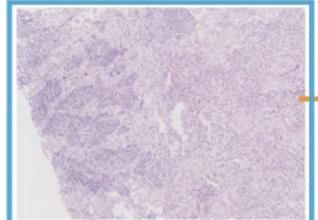
Cancer Genome Evolution Research Group, University College London Cancer
Institute, University College London, London, UK

<https://doi.org/10.1038/s41586-019-1032-7>

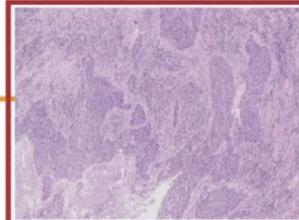
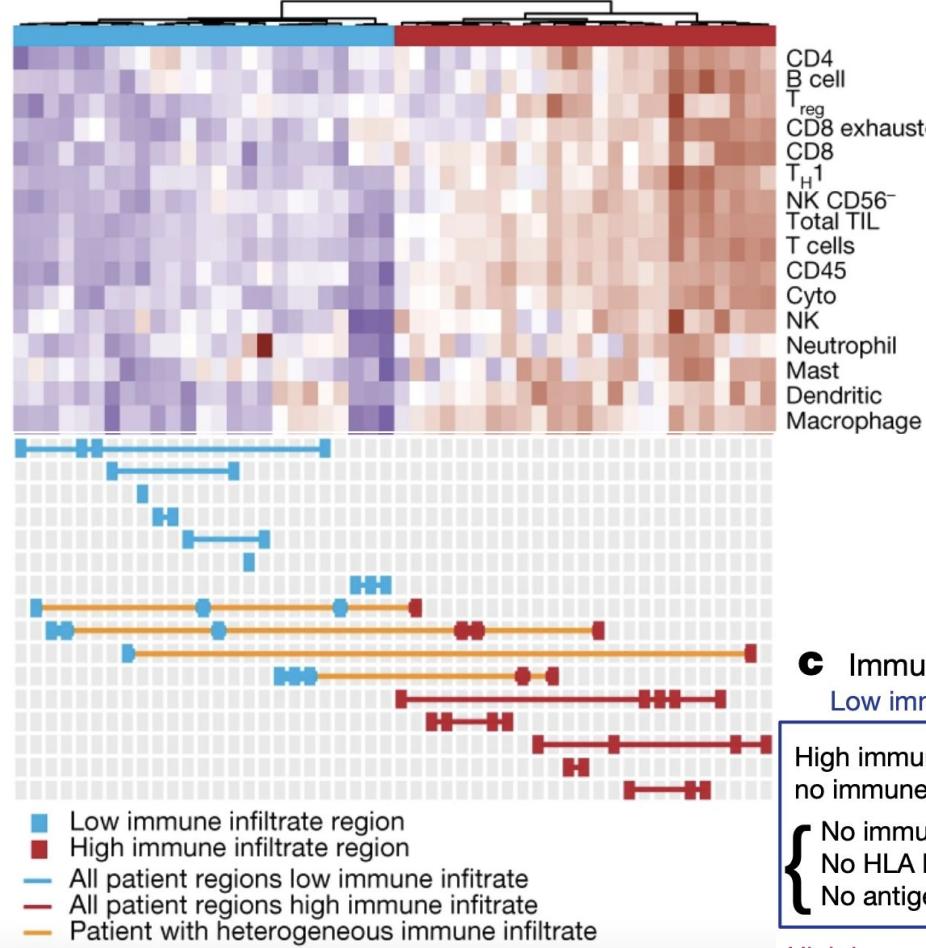
nature

a Lung adenocarcinoma

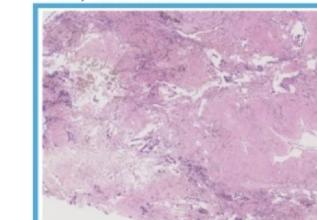
R3, low immune



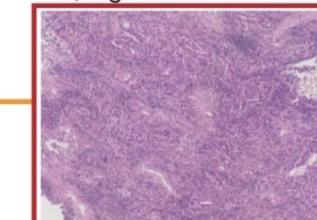
R2, high immune

**b** Lung squamous cell carcinoma

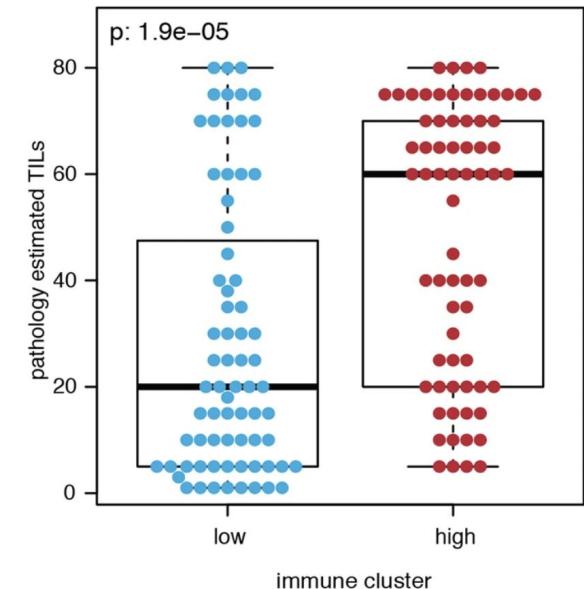
R2, low immune



R1, high immune



CD4
B cell
 T_{reg}
CD8 exhausted
CD8
 T_{H1}
NK CD56⁻
Total TIL
T cells
CD45
Cyto
NK
Neutrophil
Mast
Dendritic
Macrophage



c Immune-evasion capacity

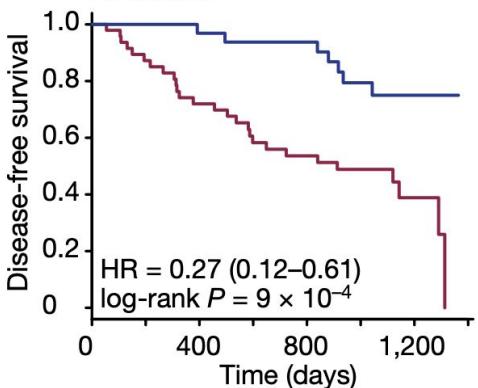
Low immune evasion

High immune infiltration or no immune escape
 { No immune editing
 No HLA LOH
 No antigen-processing defect

High immune evasion

Low/mixed immune infiltration and immune escape
 { Immune editing /
 HLA LOH /
 Antigen-processing defect

Immune-evasion capacity All tumors



HR = 0.27 ($0.12-0.61$)
 log-rank $P = 9 \times 10^{-4}$

34 32 31 30 28 19 7 0
 49 40 33 25 23 17 7 0

— Low immune-evasion capacity
 — High immune-evasion capacity

Super enhancers define regulatory subtypes and cell identity in neuroblastoma

Moritz Gartlgruber, Ashwini Kumar Sharma, Andrés Quintero, Daniel Dreidax, Selina Jansky, Young-Gyu Park, Sina Kreth, Johanna Meder, Daria Doncevic, Paul Saary, Umut H. Toprak, Naveed Ishaque, Elena Afanasyeva, Elisa Wecht, Jan Koster, Rogier Versteeg, Thomas G. P. Grünewald, David T. W. Jones, Stefan M. Pfister, Kai-Oliver Henrich, Johan van Nes, Carl Herrmann  & Frank Westermann 

Nature Cancer **2**, 114–128(2021) | Cite this article

1651 Accesses | **1** Citations | **38** Altmetric | Metrics

Health Data Science Unit, Medical Faculty Heidelberg and BioQuant, Heidelberg,

Germany

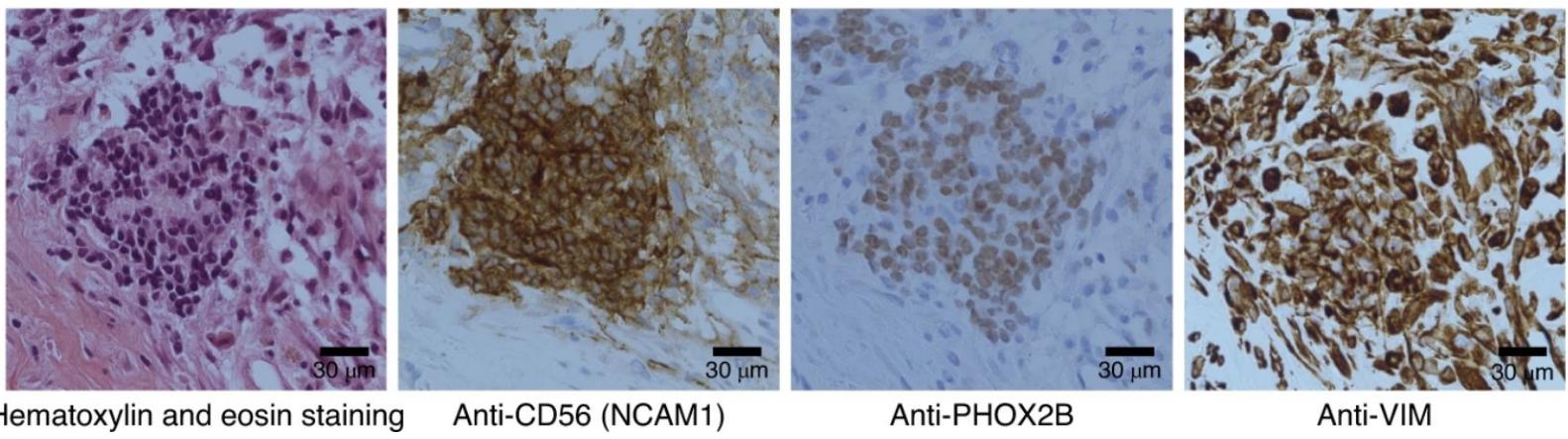
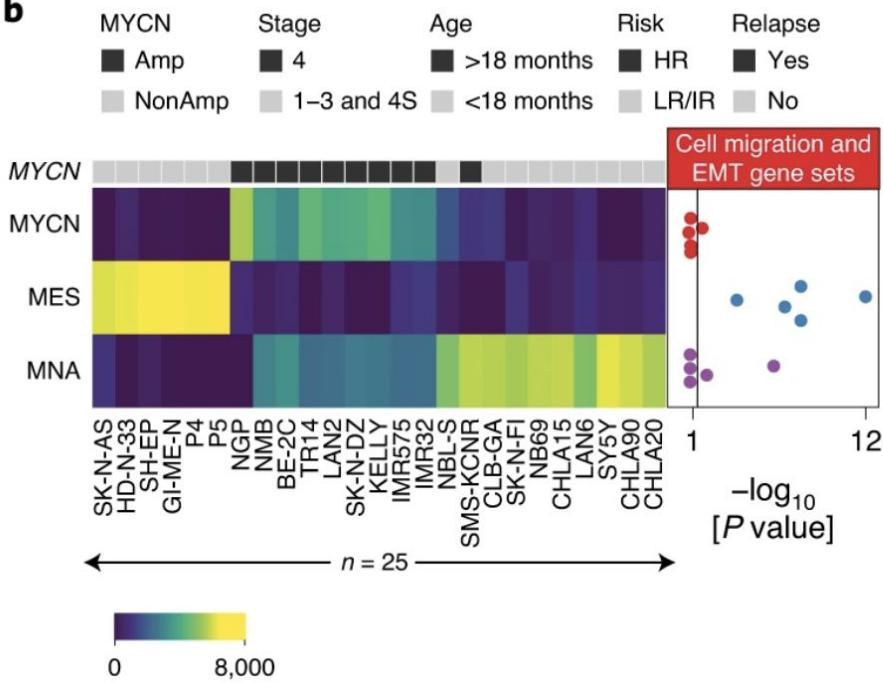
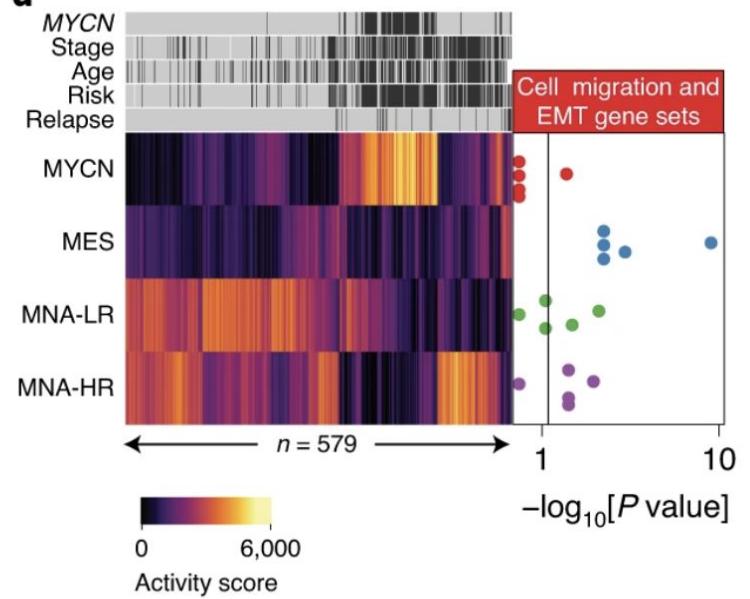
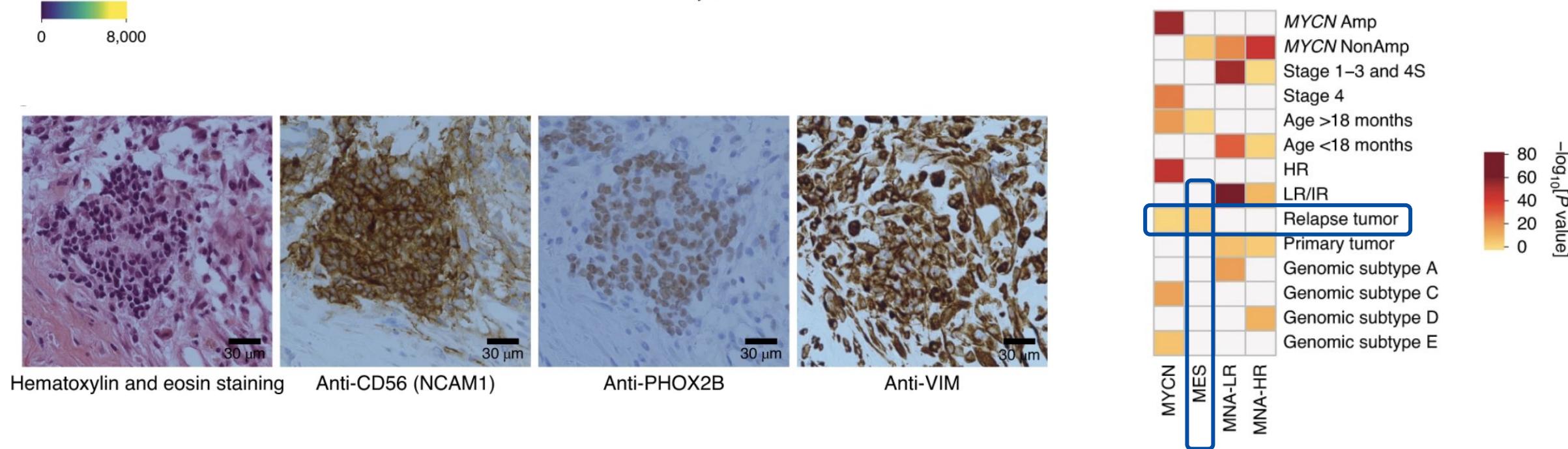
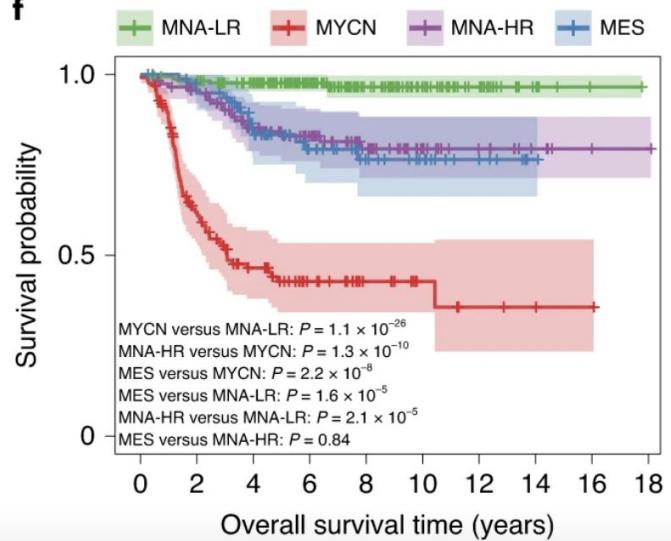
Hopp Children's Cancer Center Heidelberg (KiTZ), Heidelberg, Germany

Division of Neuroblastoma Genomics, German Cancer Research Center, Heidelberg,

Germany

<https://doi.org/10.1038/s43018-020-00145-w>

nature cancer

b**d****f**

Single cell guided deconvolution

CIBERSORTx (CSx)

Determining cell type abundance and expression from bulk tissues with digital cytometry

Aaron M. Newman , Chloé B. Steen, Chih Long Liu, Andrew J. Gentles, Aadel A. Chaudhuri, Florian Scherer, Michael S. Khodadoust, Mohammad S. Esfahani, Bogdan A. Luca, David Steiner, Maximilian Diehn & Ash A. Alizadeh 

Nature Biotechnology **37**, 773–782(2019) | [Cite this article](#)

39k Accesses | **160** Citations | **140** Altmetric | [Metrics](#)

Cell Population Mapping (CPM)

Article | Published: 18 March 2019

Cell composition analysis of bulk genomics using single-cell data

Amit Frishberg, Naama Peshes-Yaloz, Ofir Cohn, Diana Rosenthal, Yael Steuerman, Liran Valadarsky, Gal Yankovitz, Michal Mandelboim, Fuad A. Iraqi, Ido Amit, Lior Mayo, Eran Bacharach  & Irit Gat-Viks 

Nature Methods **16**, 327–332(2019) | [Cite this article](#)

12k Accesses | **22** Citations | **69** Altmetric | [Metrics](#)

Multi-subject Single Cell deconvolution (MuSiC)

Article | [Open Access](#) | Published: 22 January 2019

Bulk tissue cell type deconvolution with multi-subject single-cell expression reference

Xuran Wang, Jihwan Park, Katalin Susztak, Nancy R. Zhang  & Mingyao Li 

Nature Communications **10**, Article number: 380 (2019) | [Cite this article](#)

39k Accesses | **77** Citations | **81** Altmetric | [Metrics](#)

Single cell-assisted deconvolutional DNN (Scaden)

Deep learning-based cell composition analysis from tissue expression profiles

 Kevin Menden^{1,*},  Mohamed Marouf²,  Sergio Oller², Anupriya Dalmia¹,  Daniel Sumner Magruder^{2,3}, Karin Kloiber²,  Peter Heutink¹ and  Stefan Bonn^{1,2,*}

¹German Center for Neurodegenerative Diseases, Tuebingen, Germany.

²Institute of Medical Systems Biology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

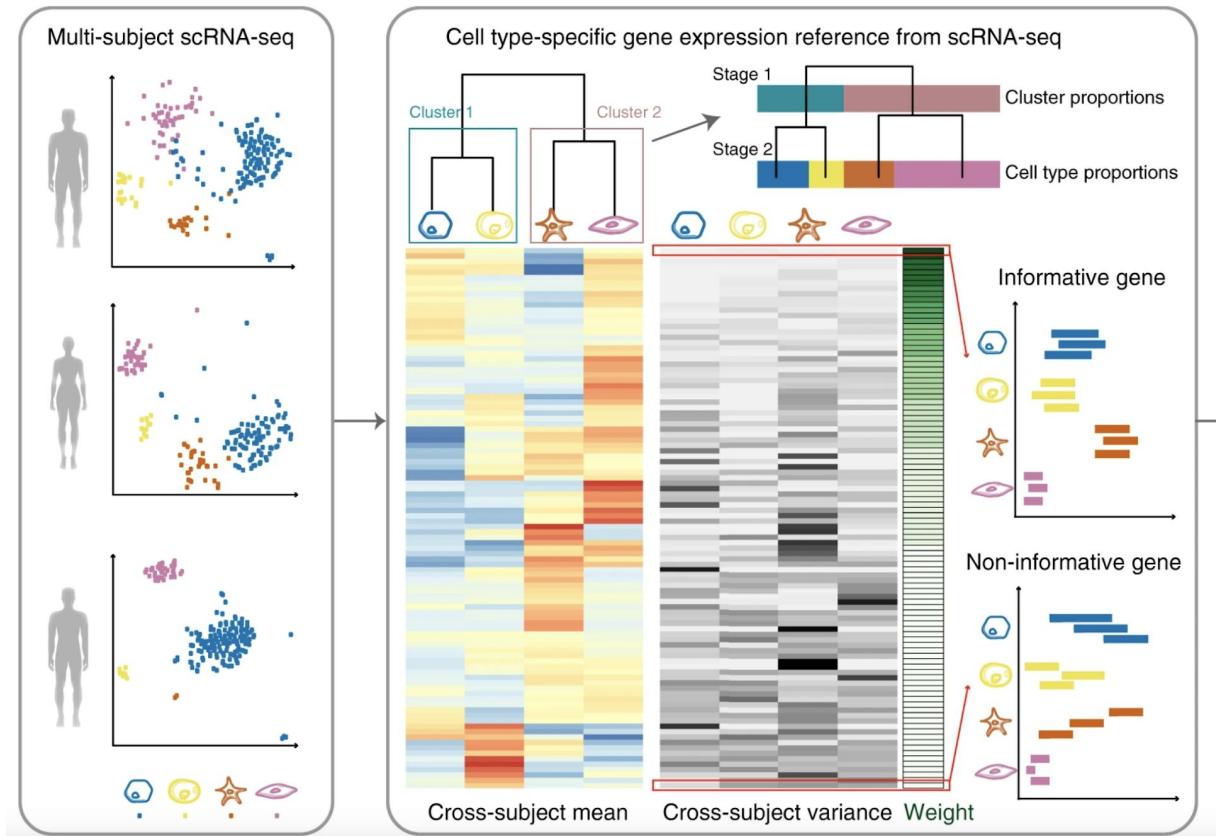
³Genevention GmbH, Goettingen, Germany.

*Corresponding author. Email: sbonn@uke.de (S.B.); kevin.menden@dzne.de (K.M.)

- Hide authors and affiliations

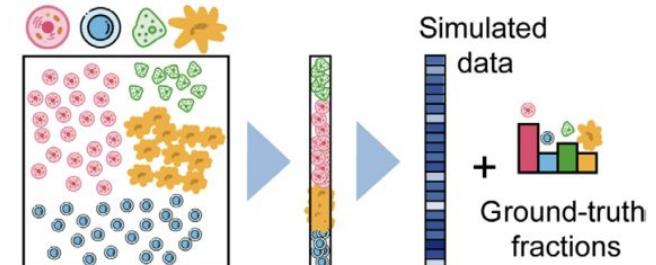
Science Advances 22 Jul 2020:
Vol. 6, no. 30, eaba2619
DOI: 10.1126/sciadv.eaba2619

MuSiC

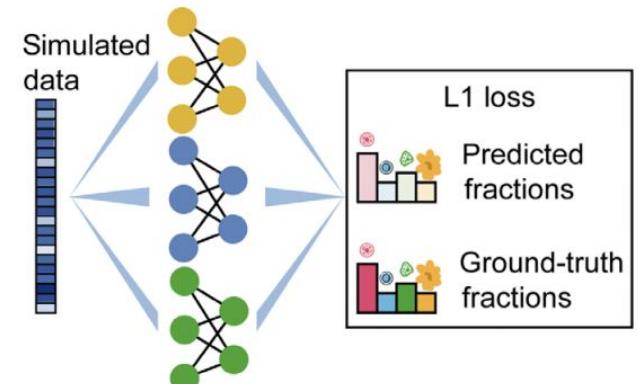


Scaden

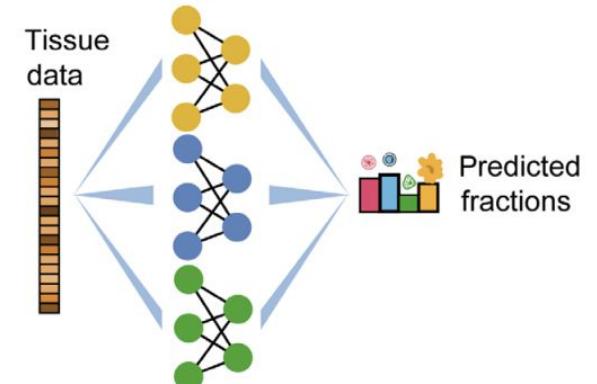
Simulated training data



B Scaden training

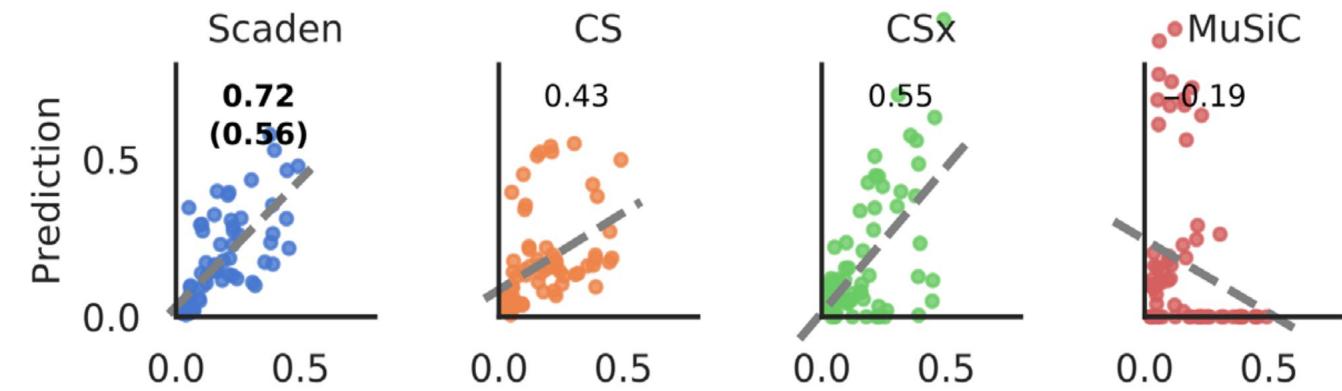
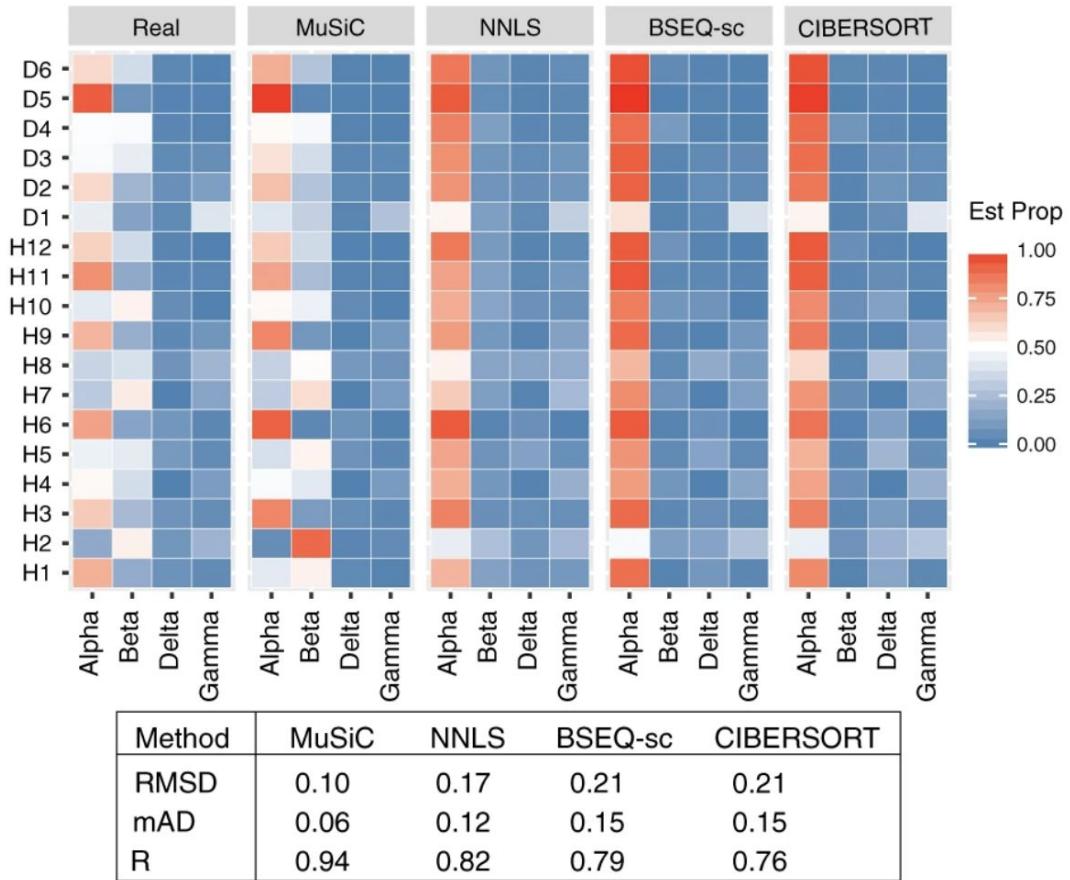


C Scaden predictions

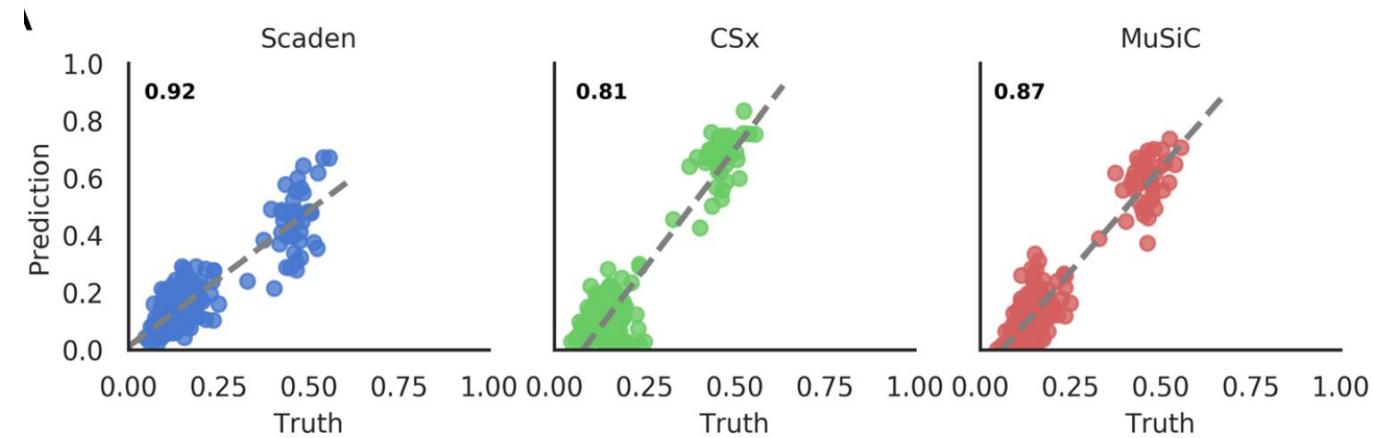


Peripheral Blood mononuclear cells

b



Brain cells





UNIVERSITAT DE
BARCELONA

Yuna Blum, Ligue contre le Cancer

Jérôme Cros, APHP

Clémentine Decamps, Uni Grenoble Alpes

Carl Herrmann, Medical Faculty Heidelberg

Slim Karkar, Uni Grenoble Alpes

Yasmina Kermezli, Uni Grenoble Alpes

Magali Richard, Uni Grenoble Alpes

Ashwini Sharma, University Hospital Heidelberg

https://cancer-heterogeneity.github.io/cometh_training.html

www.eithealth.eu | info@eithealth.eu