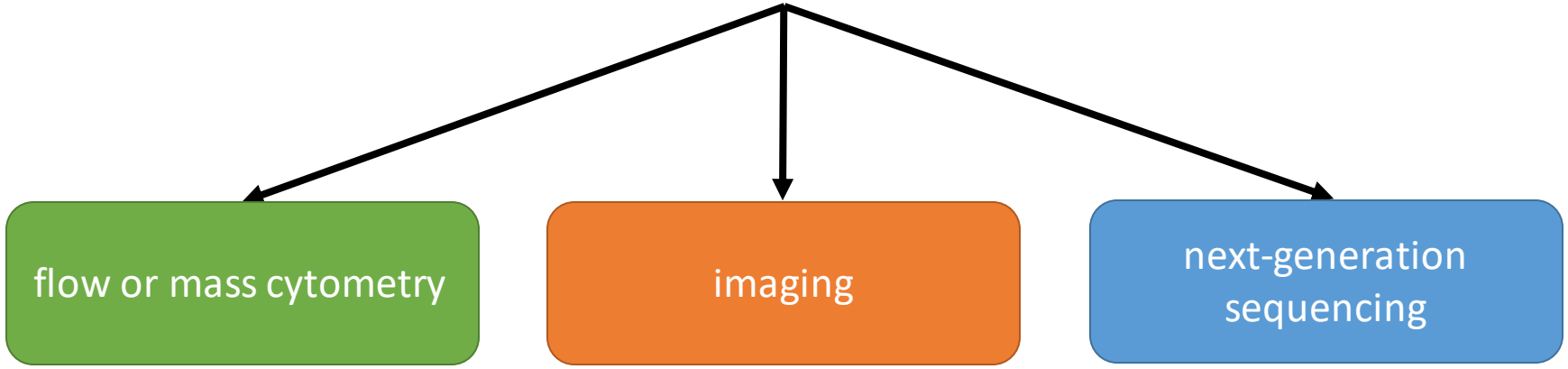
high-dimensional single-cell data and self-organizing maps

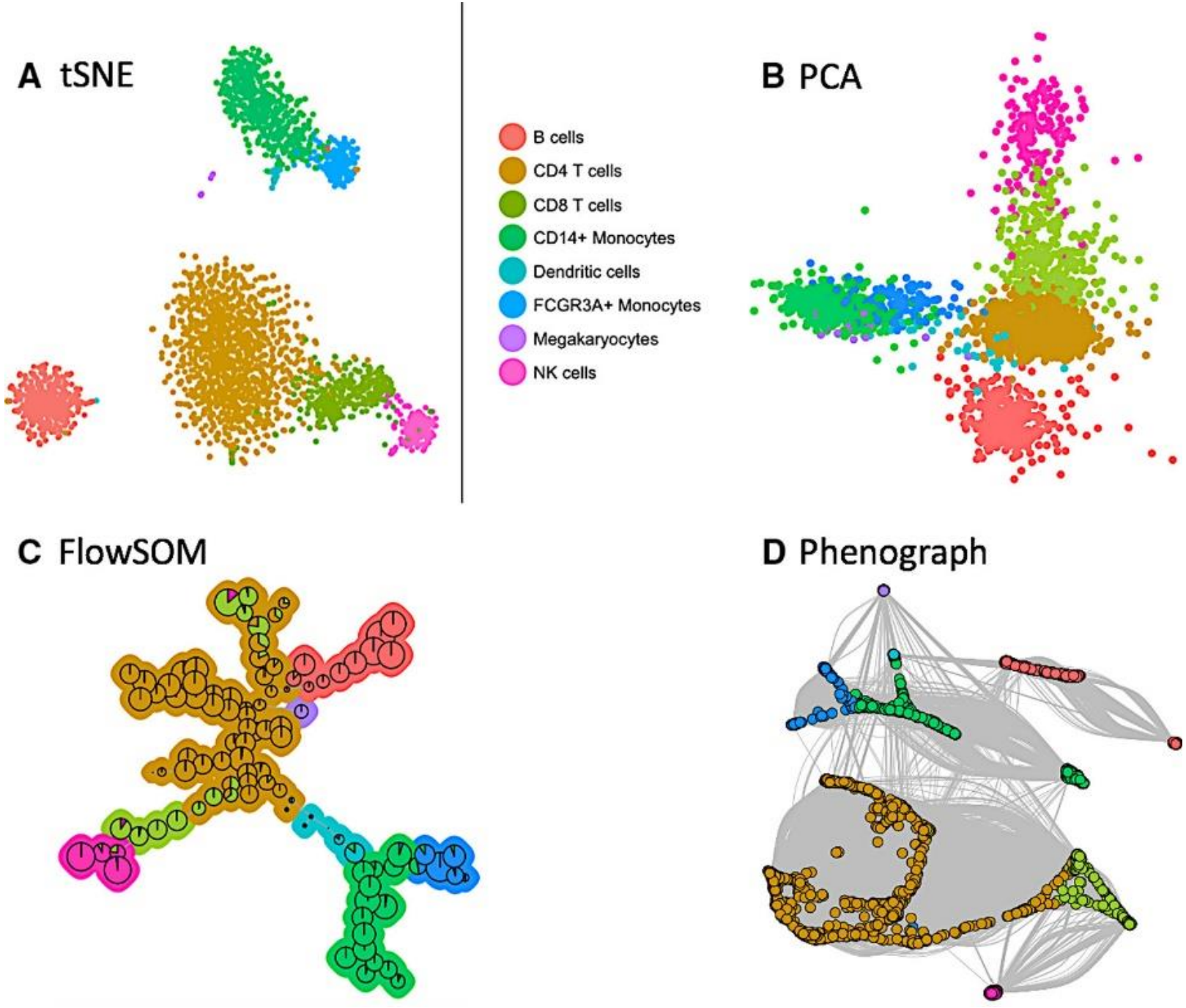
Journal Club

25.02.2020

Alexandra Bentrup

High-dimensional single-cell data technologies = HDcyto technologies

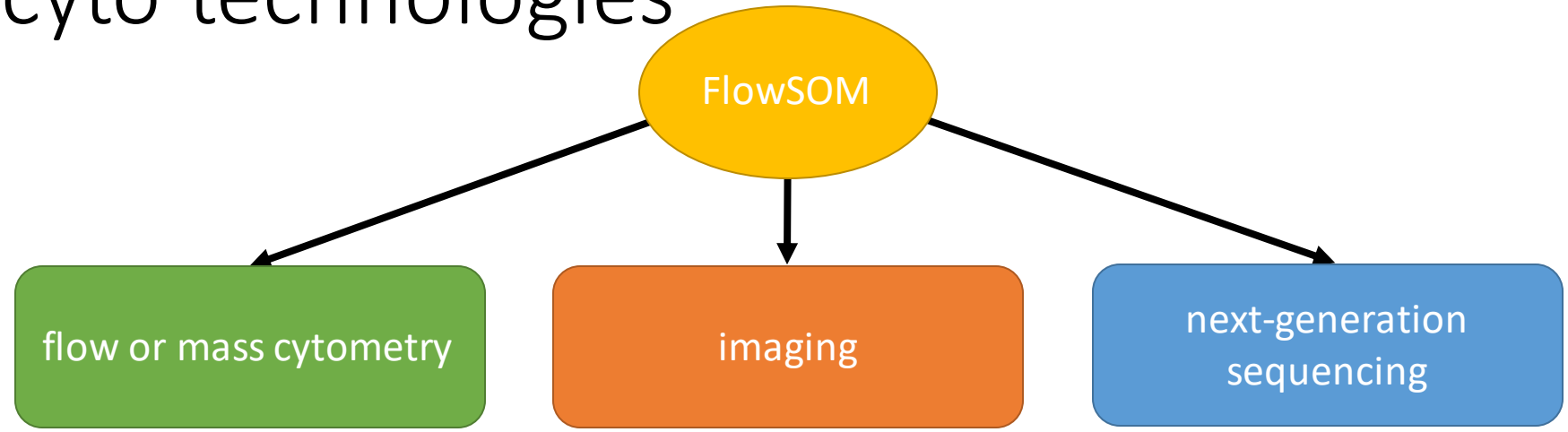




scRNA-Seq dataset of 3000 peripheral blood mononuclear cells (PBMCs) from the 10X Genomics platform (Zheng et al 2017 Nat Commun)

Todorov 2018
doi: 10.1111/febs.14613

High-dimensional single-cell data technologies = HDcyto technologies



A Systems-Level Study Reveals Regulators of Membrane-less Organelles in Human Cells

Doris Berchtold,[1] Nico Battich,[1] and Lucas Pelkmans[1,2,*]
[1]Institute of Molecular Life Sciences, University of Zurich, 8057 Zurich, Switzerland
[2]Lead Contact
*Correspondence: lucas.pelkmans@imls.uzh.ch
https://doi.org/10.1016/j.molcel.2018.10.036

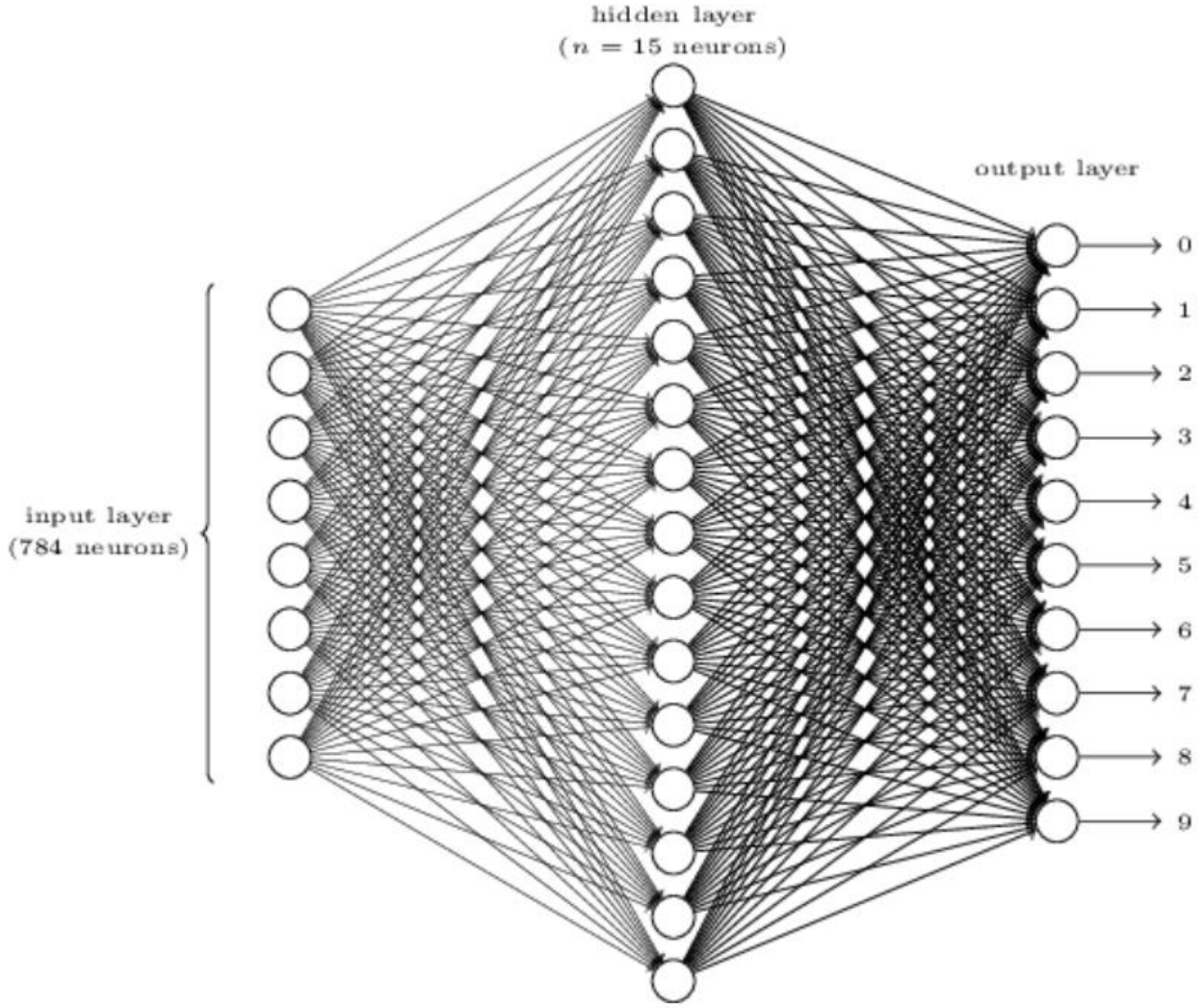
FlowSOM: Using Self-Organizing Maps for Visualization and Interpretation of Cytometry Data

Sofie Van Gassen,[1,2,3]* Britt Callebaut,[1] Mary J. Van Helden,[2,3] Bart N. Lambrecht,[2,3]
Piet Demeester,[1] Tom Dhaene,[1] Yvan Saeys[2,3]

A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging

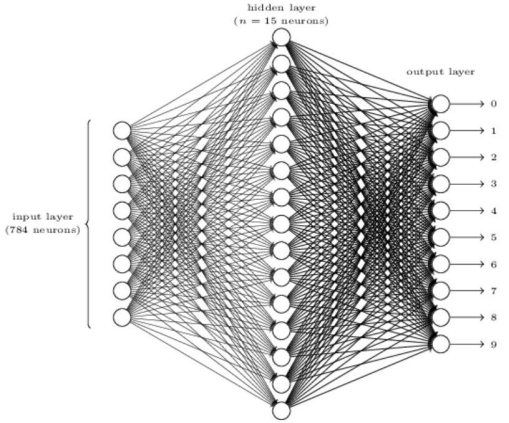
Leeat Keren,[1] Marc Bosse,[1] Diana Marquez,[1] Roshan Angoshtari,[1] Samir Jain,[1] Sushama Varma,[1] Soo-Ryum Yang,
Allison Kurian,[1] David Van Valen,[2,3] Robert West,[1] Sean C. Bendall,[1,*] and Michael Angelo[1,4,*]
[1]Department of Pathology, Stanford University, Stanford CA, 94305, USA
[2]Department of Biology, Caltech, Pasadena, CA 91125, USA
[3]Department of Bioengineering, Caltech, Pasadena, CA 91125, USA
[4]Lead Contact
*Correspondence: bendall@stanford.edu (S.C.B.), mangelo0@stanford.edu (M.A.)
https://doi.org/10.1016/j.cell.2018.08.039

Artificial neural networks



http://neuralnetworksanddeeplearning.com/chap1.html

Artificial neural networks



http://neuralnetworksanddeeplearning.com/chap1.html

1. Feedforward networks
   - Input-output transformation determined by external supervised adjustments

2. Feedback / recurrent networks
   - Neurons are not independent of each other, feedback loops &«memory» present

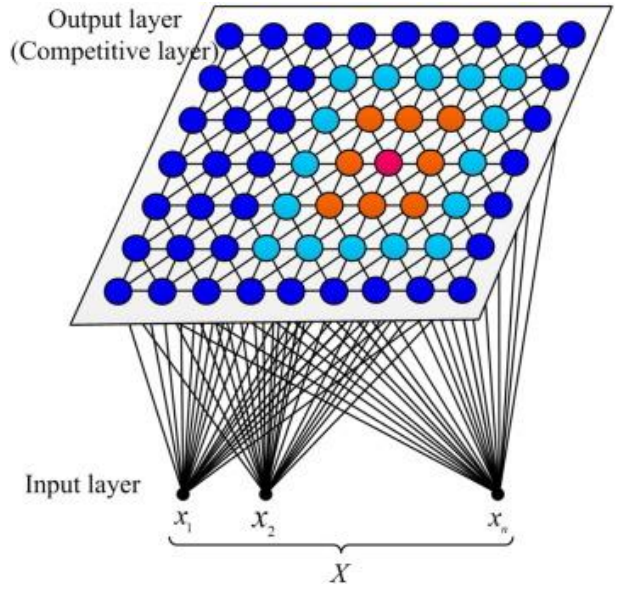
3. **Competitive / unsupervised / self-organizing**
   - **Competition in activity by mutual lateral interactions**
   - **Development of neurons into specific detectors of signal patterns**

Competitive ANN

- Competition between neurons to be activated
- Only one neuron is active at a time = winner neuron

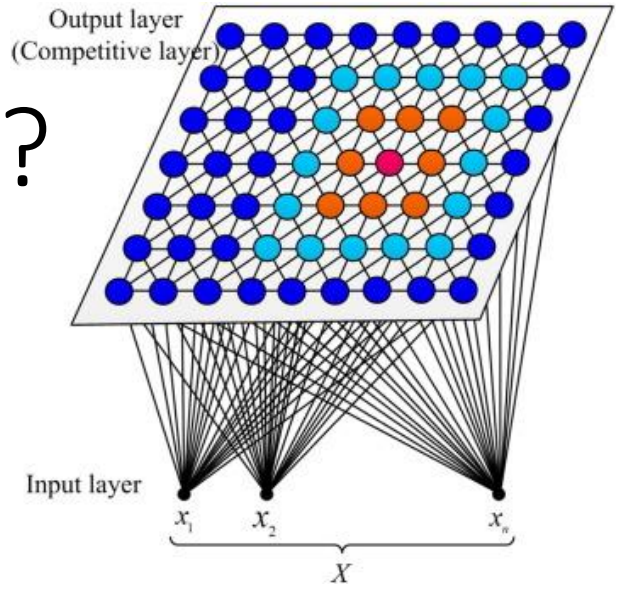
- Selective tuning of neuron to input pattern / class

- Ordering of neurons in meaningful coordinate system

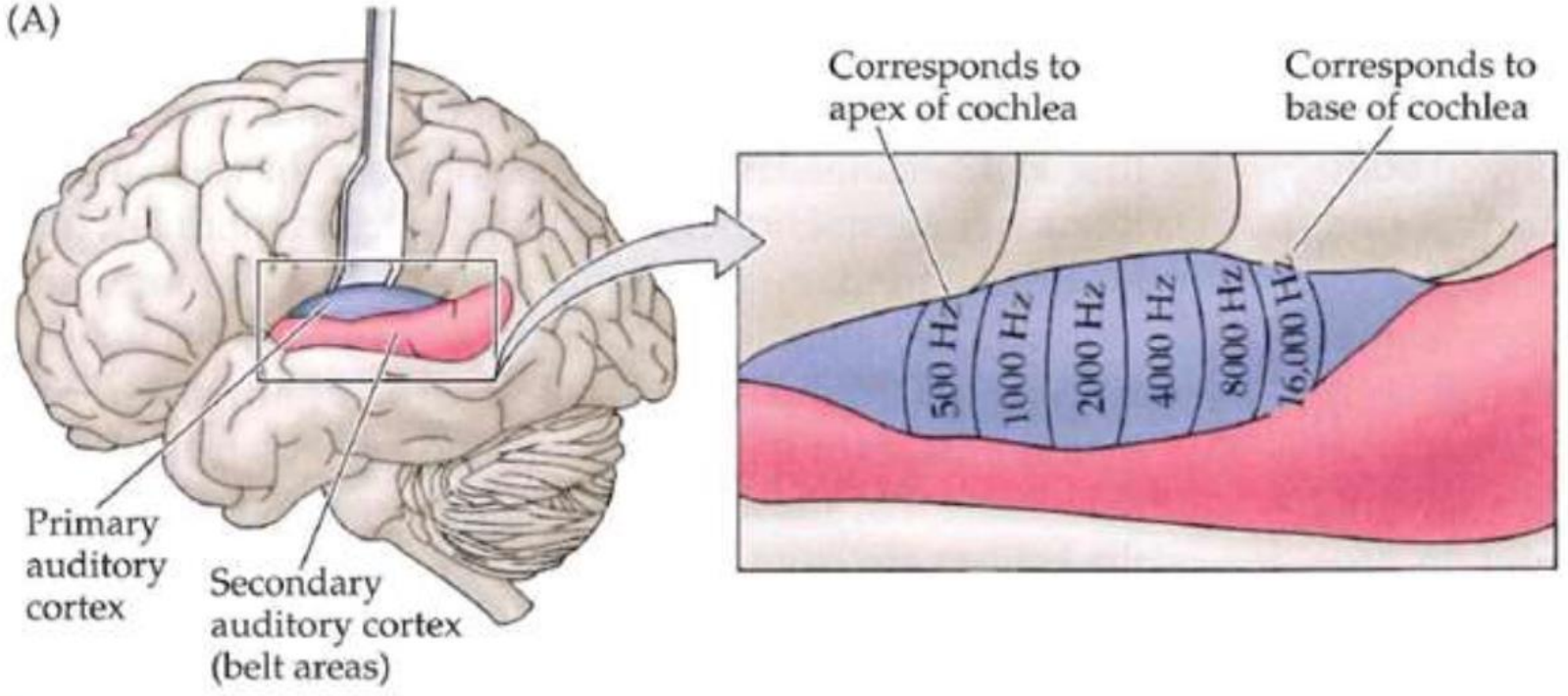


https://github.com/droidadroit/SOM/wiki/SOM

Competitive ANN = Neural Network?

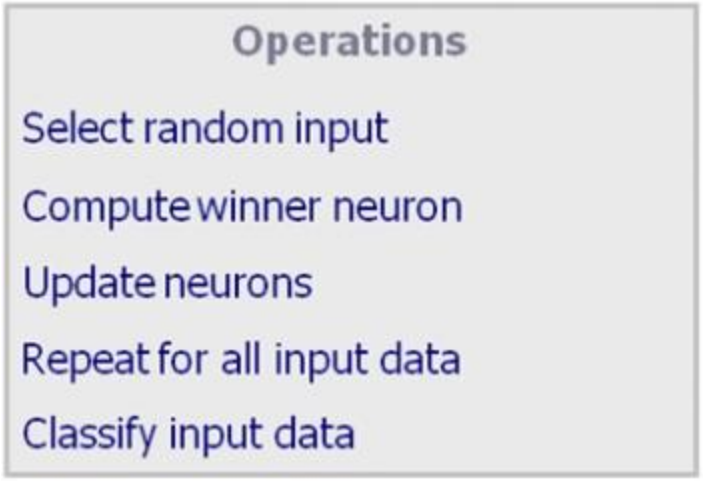


https://github.com/droidadroit/SOM/wiki/SOM



Kandel et al (2013). Principles of Neural Science

Self Organizing Maps = SOM (= Kohonen map)



https://www.youtube.com/watch?v=H9H6s-x-0YE

Self Organizing Maps = SOM (= Kohonen map)

Operations

Select random input

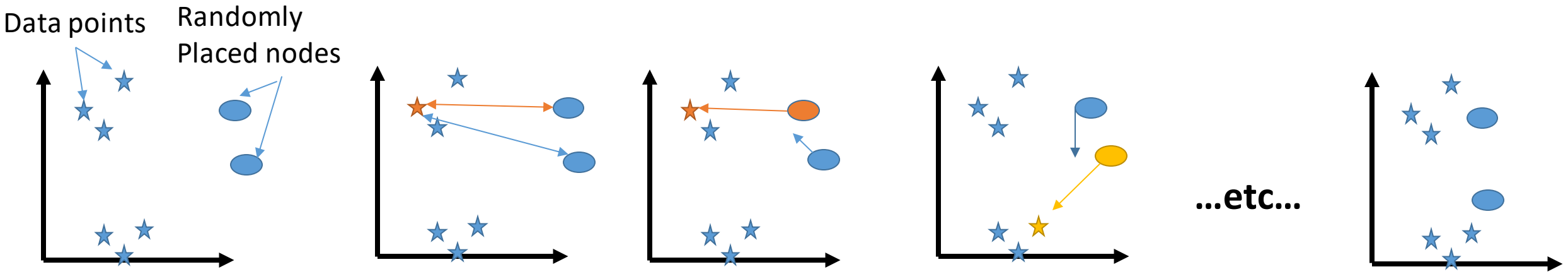
Compute winner neuron

Update neurons

Repeat for all input data

Classify input data

https://www.youtube.com/watch?v=H9H6s-x-0YE



The Self-Organizing Map

TEUVO KOHONEN, SENIOR MEMBER, IEEE

PROCEEDINGS OF THE IEEE, VOL. 78, NO. 9, SEPTEMBER 1990

Table 1 Input Data Matrix

	Item																													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6
Attribute																																
a_1	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_2	0	0	0	0	0	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_3	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	3	3	3	3	6	6	6	6	6	6	6	6	6	6
a_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	1	2	3	4	2	2	2	2	2	2
a_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6

The Self-Organizing Map

TEUVO KOHONEN, SENIOR MEMBER, IEEE

PROCEEDINGS OF THE IEEE, VOL. 78, NO. 9, SEPTEMBER 1990

Table 1 Input Data Matrix

	Item																													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6
Attribute																																
a_1	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_2	0	0	0	0	0	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_3	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	3	3	3	3	6	6	6	6	6	6	6	6	6	6
a_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	1	2	3	4	2	2	2	2	2	2
a_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6

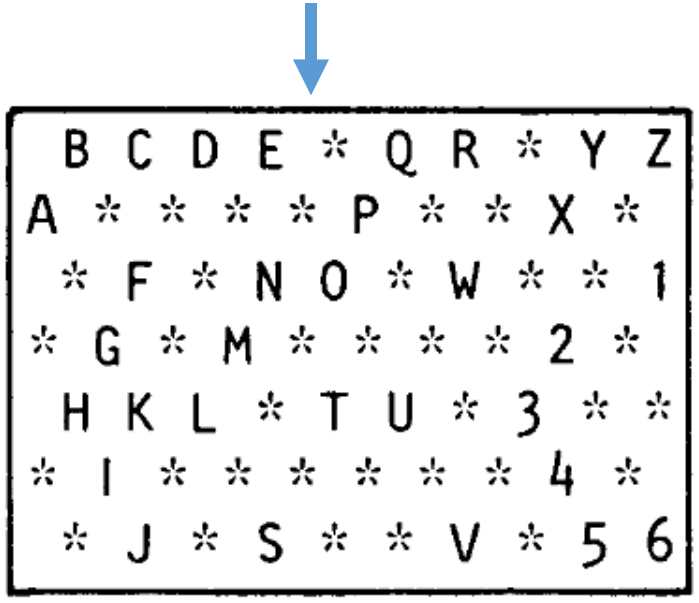


Fig. 6. Self-organized map of the data matrix of Table 1.

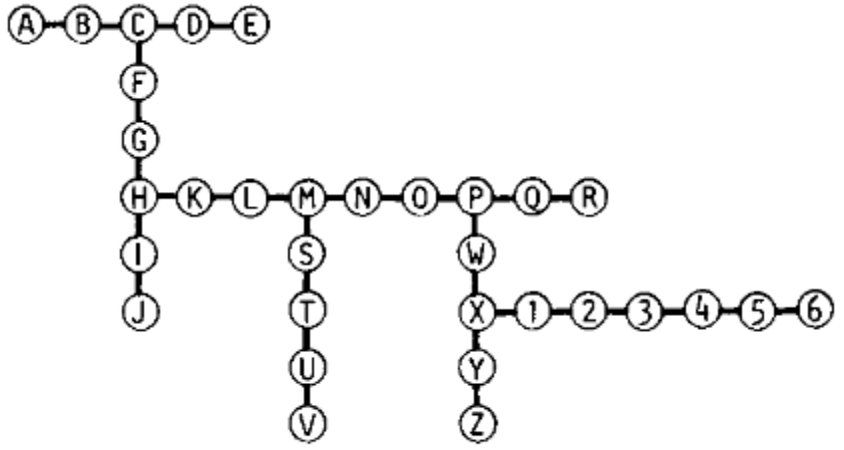


Fig. 7. Minimal spanning tree corresponding to Table 1.

The Self-Organizing Map

TEUVO KOHONEN, SENIOR MEMBER, IEEE

PROCEEDINGS OF THE IEEE, VOL. 78, NO. 9, SEPTEMBER 1990

Table 1 Input Data Matrix

	Item																												
Attribute	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6
a_1	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_2	0	0	0	0	0	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_3	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	3	3	3	3	6	6	6	6	6	6	6	6	6	6
a_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	1	2	3	4	2	2	2	2	2	2
a_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6

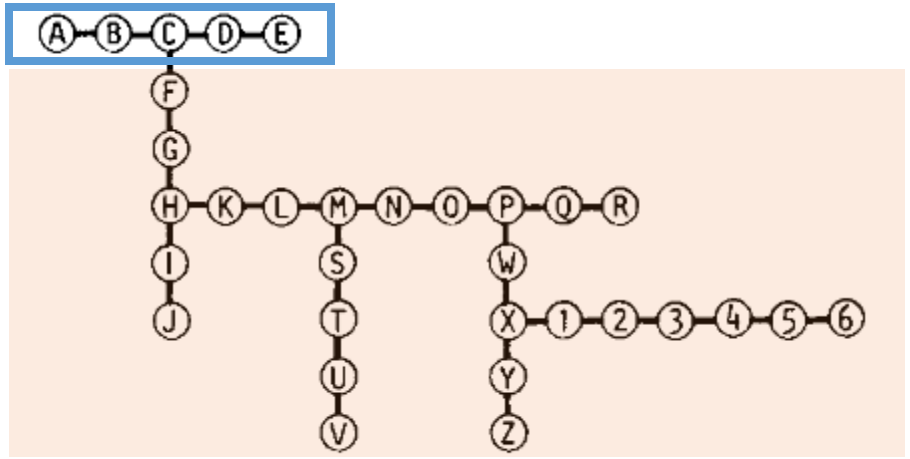


Fig. 7. Minimal spanning tree corresponding to Table 1.

The Self-Organizing Map

TEUVO KOHONEN, SENIOR MEMBER, IEEE

PROCEEDINGS OF THE IEEE, VOL. 78, NO. 9, SEPTEMBER 1990

Table 1 Input Data Matrix

	Item																													
Attribute	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6
a_1	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_2	0	0	0	0	0	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_3	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	3	3	3	3	6	6	6	6	6	6	6	6	6	6
a_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	1	2	3	4	2	2	2	2	2	2
a_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6

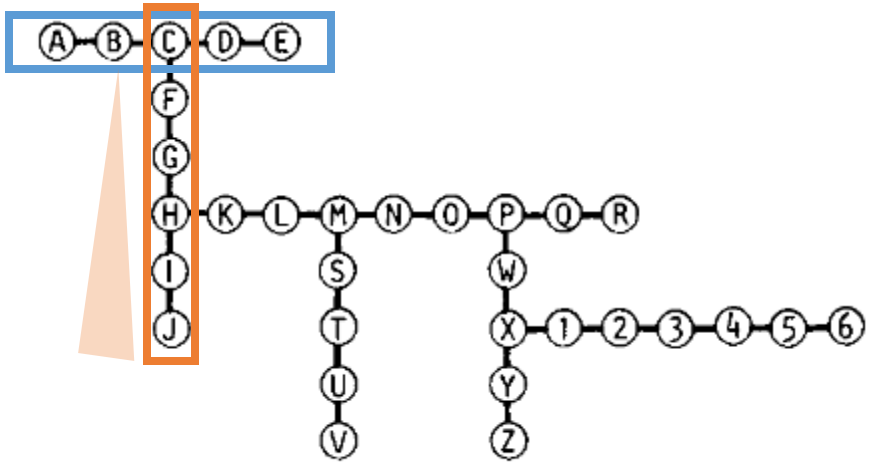


Fig. 7. Minimal spanning tree corresponding to Table 1.

The Self-Organizing Map

TEUVO KOHONEN, SENIOR MEMBER, IEEE

PROCEEDINGS OF THE IEEE, VOL. 78, NO. 9, SEPTEMBER 1990

Table 1 Input Data Matrix

	Item																												
Attribute	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6
a_1	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_2	0	0	0	0	0	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_3	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	3	3	3	3	6	6	6	6	6	6	6	6	6	6
a_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	1	2	3	4	2	2	2	2	2	2
a_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6

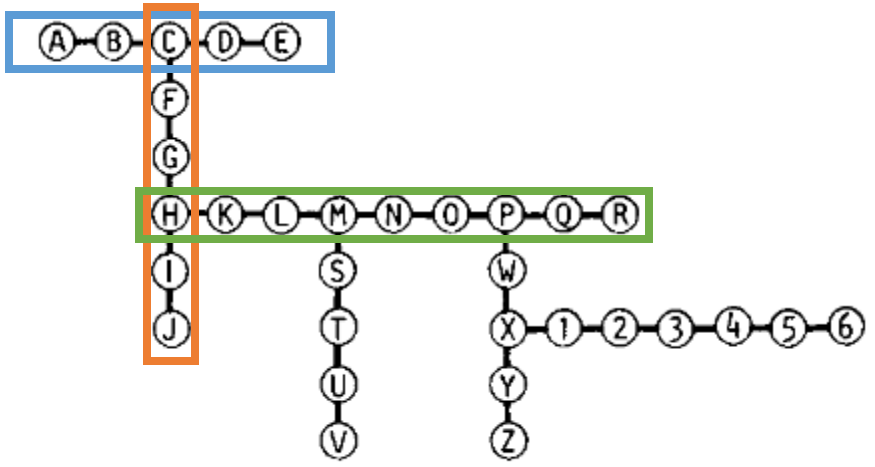


Fig. 7. Minimal spanning tree corresponding to Table 1.

The Self-Organizing Map

TEUVO KOHONEN, SENIOR MEMBER, IEEE

PROCEEDINGS OF THE IEEE, VOL. 78, NO. 9, SEPTEMBER 1990

Table 1 Input Data Matrix

	Item																											
Attribute	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6
a_1	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_2	0	0	0	0	0	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_3	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	3	3	3	3	6	6	6	6	6	6	6	6	6	6
a_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	1	2	3	4	2	2	2	2	2	2
a_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6

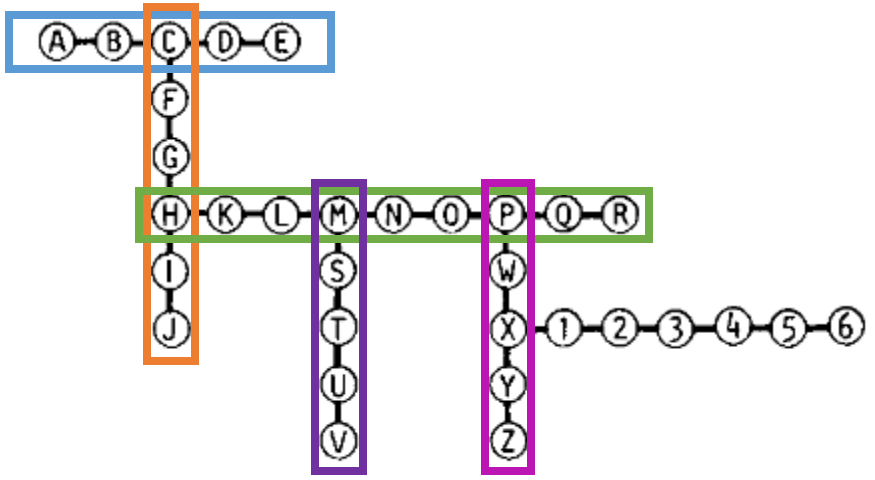


Fig. 7. Minimal spanning tree corresponding to Table 1.

The Self-Organizing Map

TEUVO KOHONEN, SENIOR MEMBER, IEEE

PROCEEDINGS OF THE IEEE, VOL. 78, NO. 9, SEPTEMBER 1990

Table 1 Input Data Matrix

	Item																											
Attribute	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6
a_1	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_2	0	0	0	0	0	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_3	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	3	3	3	3	6	6	6	6	6	6	6	6	6	6
a_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	1	2	3	4	2	2	2	2	2	2
a_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6

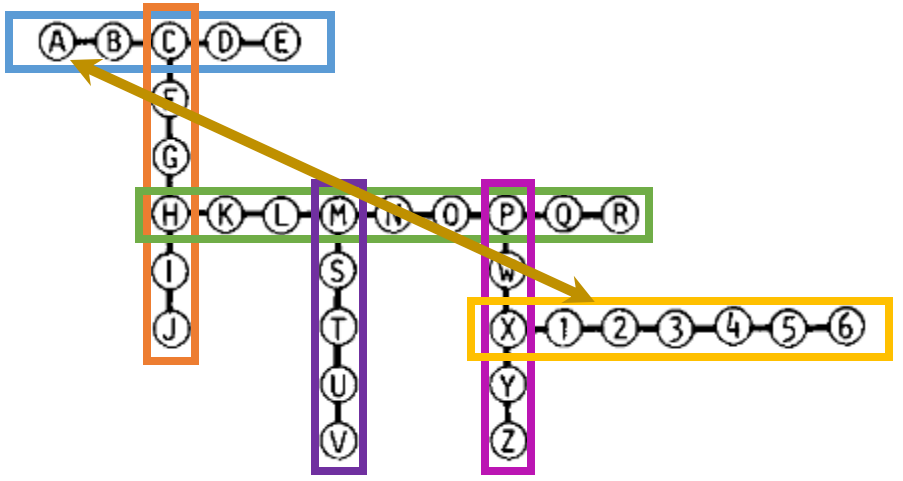


Fig. 7. Minimal spanning tree corresponding to Table 1.

The Self-Organizing Map

TEUVO KOHONEN, SENIOR MEMBER, IEEE

PROCEEDINGS OF THE IEEE, VOL. 78, NO. 9, SEPTEMBER 1990

Table 1 Input Data Matrix

	Item																											
Attribute	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6
a_1	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_2	0	0	0	0	0	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_3	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	3	3	3	3	6	6	6	6	6	6	6	6	6	6
a_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	1	2	3	4	2	2	2	2	2	2
a_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6

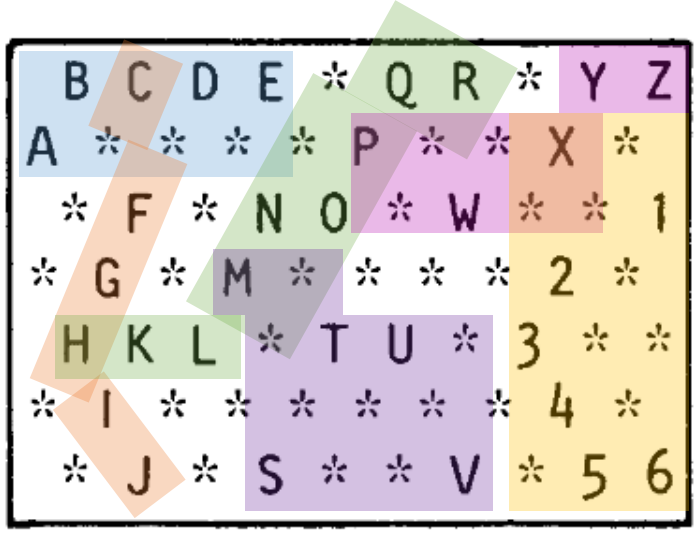


Fig. 6. Self-organized map of the data matrix of Table 1.

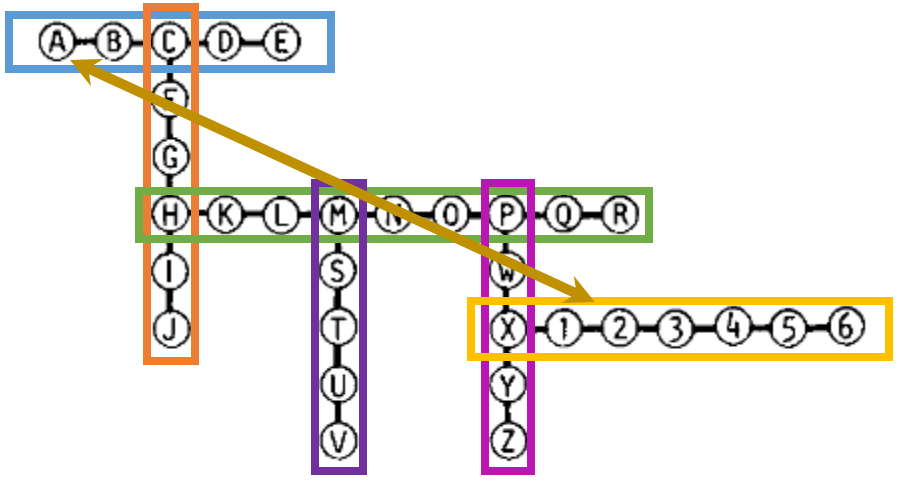


Fig. 7. Minimal spanning tree corresponding to Table 1.

ORIGINAL ARTICLE

Cytometry
PART A
Journal of the International Society for Advancement of Cytometry

MIFlowCyt

FlowSOM: Using Self-Organizing Maps for Visualization and Interpretation of Cytometry Data

Sofie Van Gassen,[1,2,3]* Britt Callebaut,[1] Mary J. Van Helden,[2,3] Bart N. Lambrecht,[2,3]
Piet Demeester,[1] Tom Dhaene,[1] Yvan Saeys[2,3]

Motivation:

- FACS experiments with many markers produce large dataset
- Create overview with minimal amount of graphs to be plotted
- Don’t loose information due to gating strategy

Read Data Tables

Clustering

Minimal Spanning Tree for Visualization

Second Clustering

Read Data Tables

Clustering

Minimal Spanning Tree for Visualization

Second Clustering

Table 1 Input Data Matrix

	Item																												
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6
Attribute																																
a_1	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_2	0	0	0	0	0	1	2	3	4	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
a_3	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	3	3	3	3	6	6	6	6	6	6	6	6	6	6
a_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	1	2	3	4	2	2	2	2	2	2
a_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6

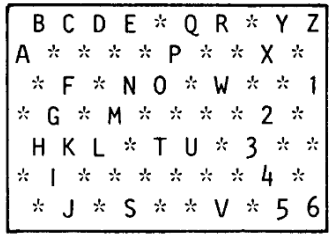


Fig. 6. Self-organized map of the data matrix of Table 1.

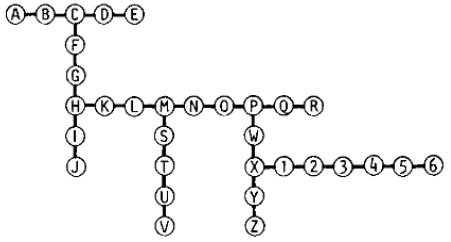


Fig. 7. Minimal spanning tree corresponding to Table 1.

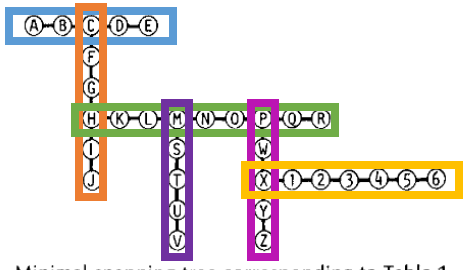
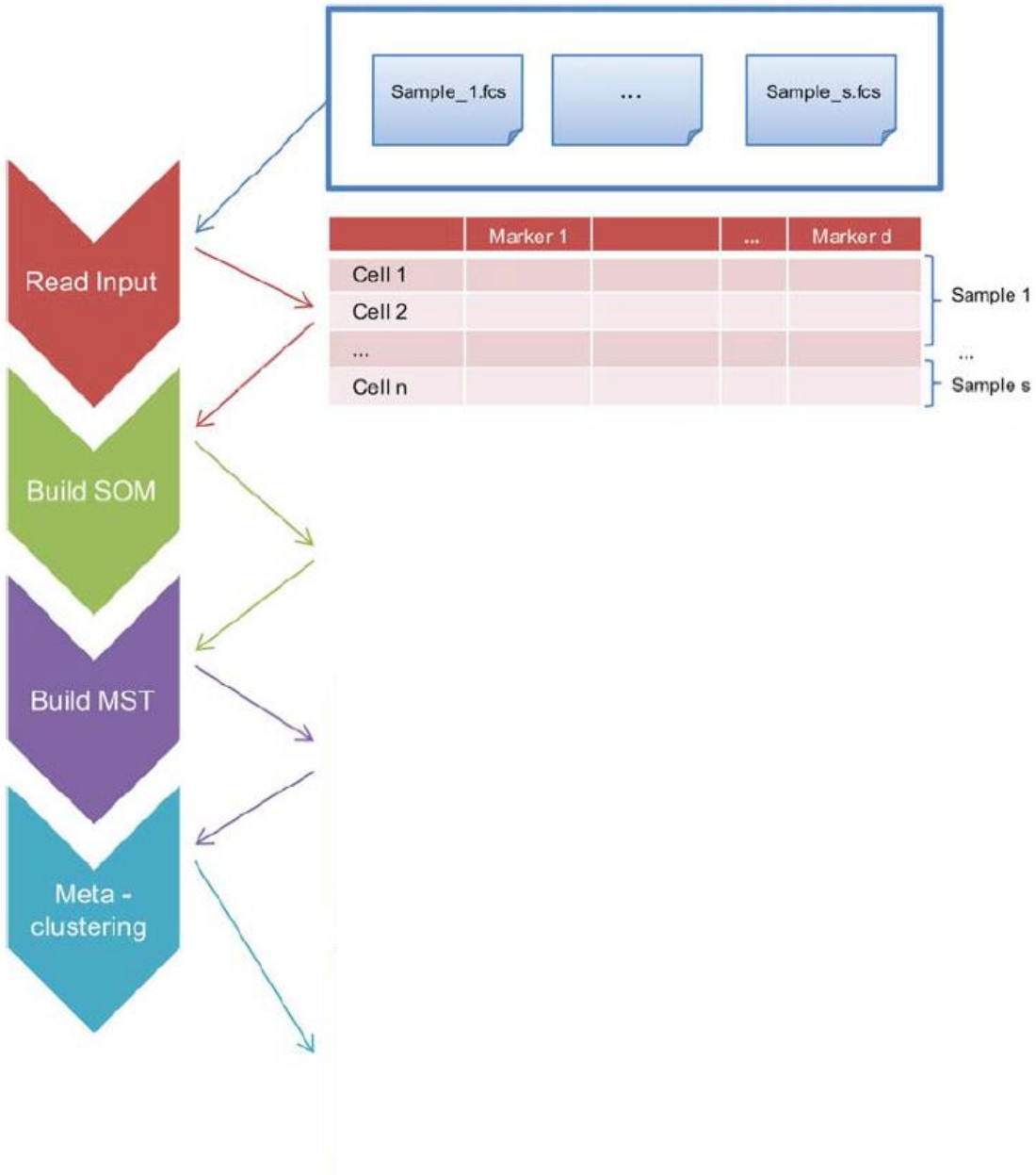
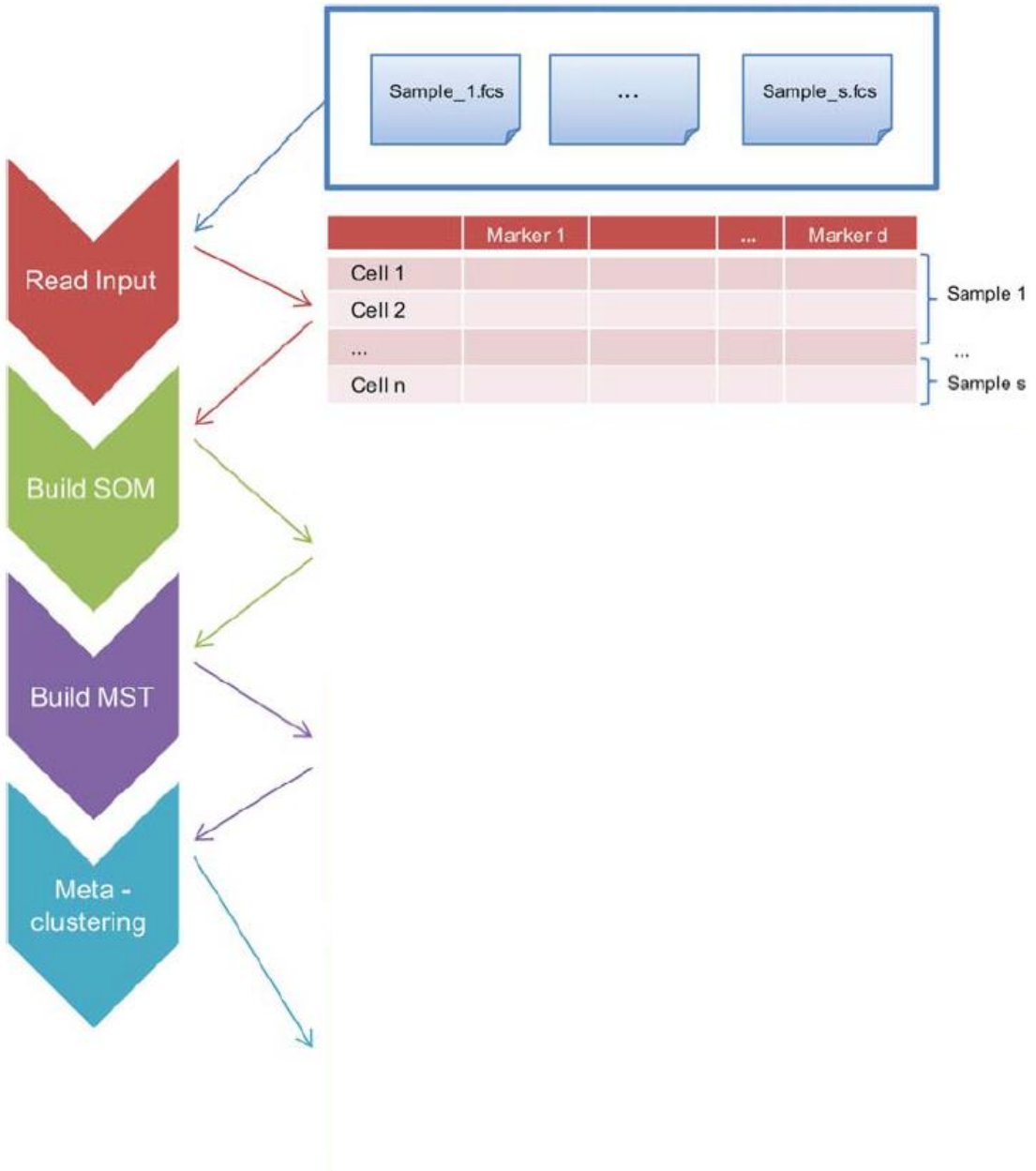


Fig. 7. Minimal spanning tree corresponding to Table 1.



Input:
Matrix of size «*cell number * marker number*»

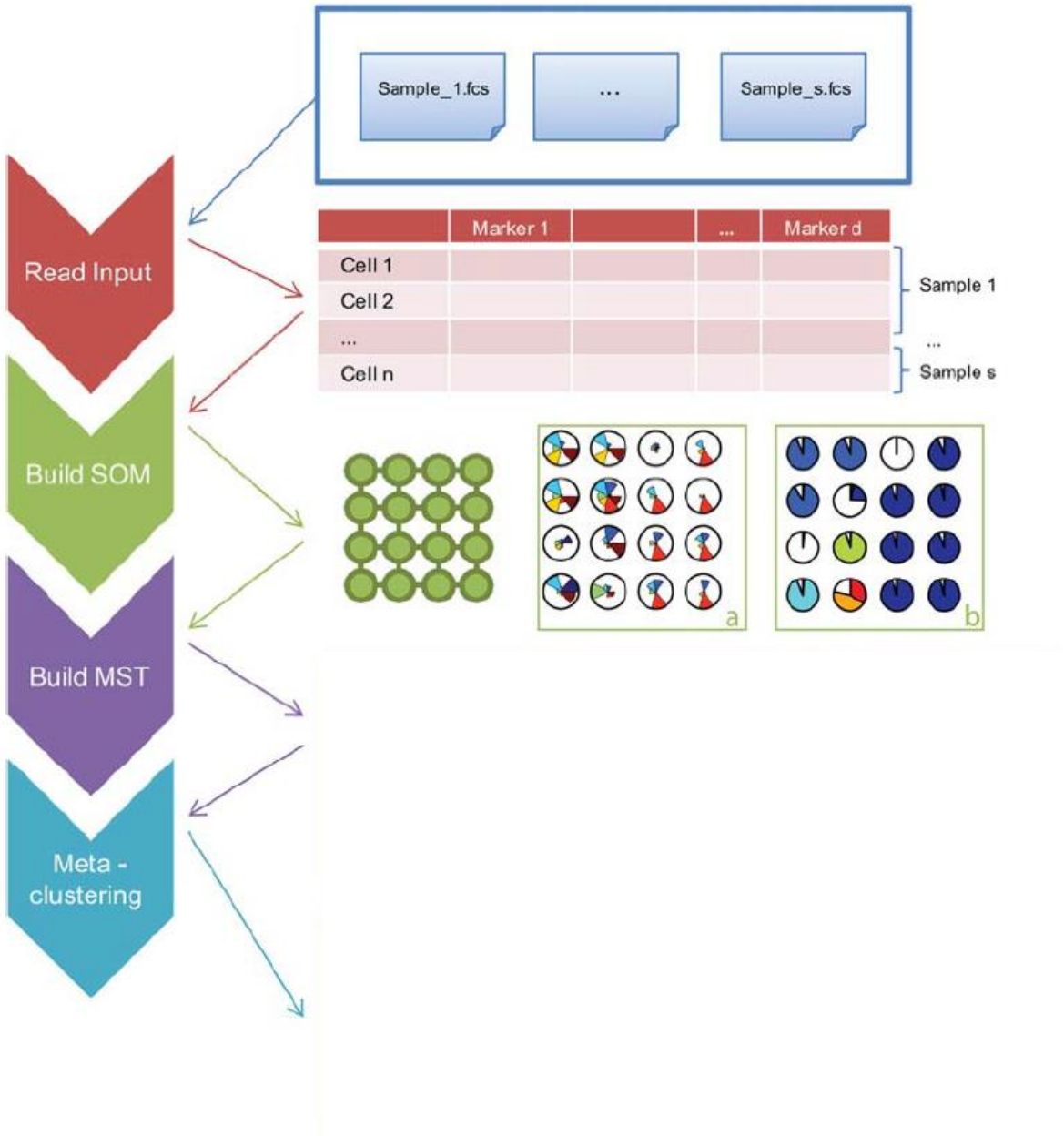
Output:
Each cell belongs to unknown cell class / type
Clusters correspond to *cell types*



Input:
Matrix of size «*cell number * marker number*»

Pre-processing:
Scaling of all features so that all columns/markers have a
Mean = 0
SD = 1
→ Equal importance of all markers

Output:
Each cell belongs to unknown cell class / type
Clusters correspond to cell types



Input:

Matrix of size «*cell number * marker number*»

Pre-processing

Scaling of all features so that all columns/markers have a

Mean = 0

SD = 1

→ Equal importance of all markers

Intermediate visualization:

More Clusters than cell types

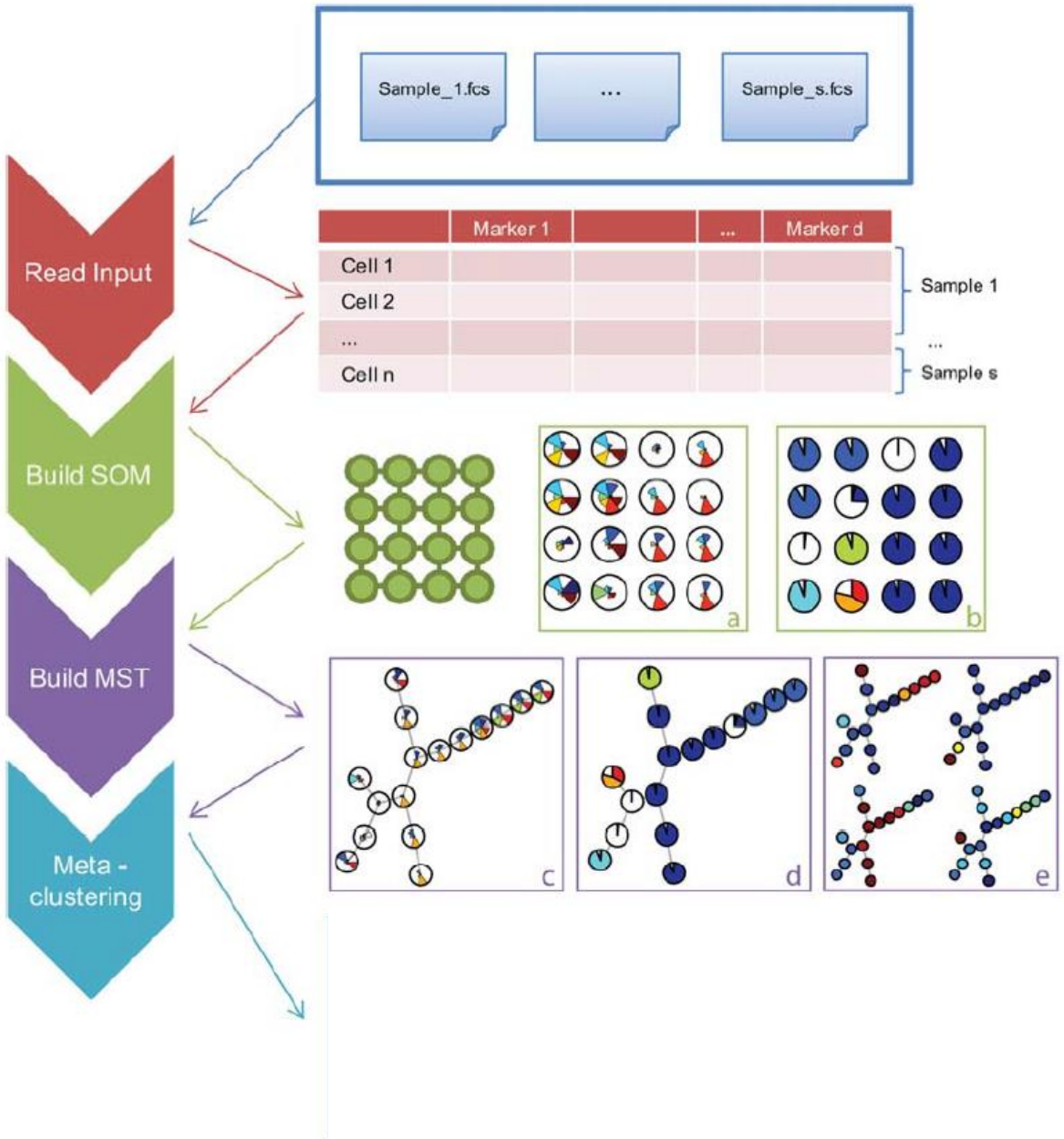
To have «pure» classes

→ reduce subsequently

Output:

Each cell belongs to unknown cell class / type

Clusters correspond to cell types



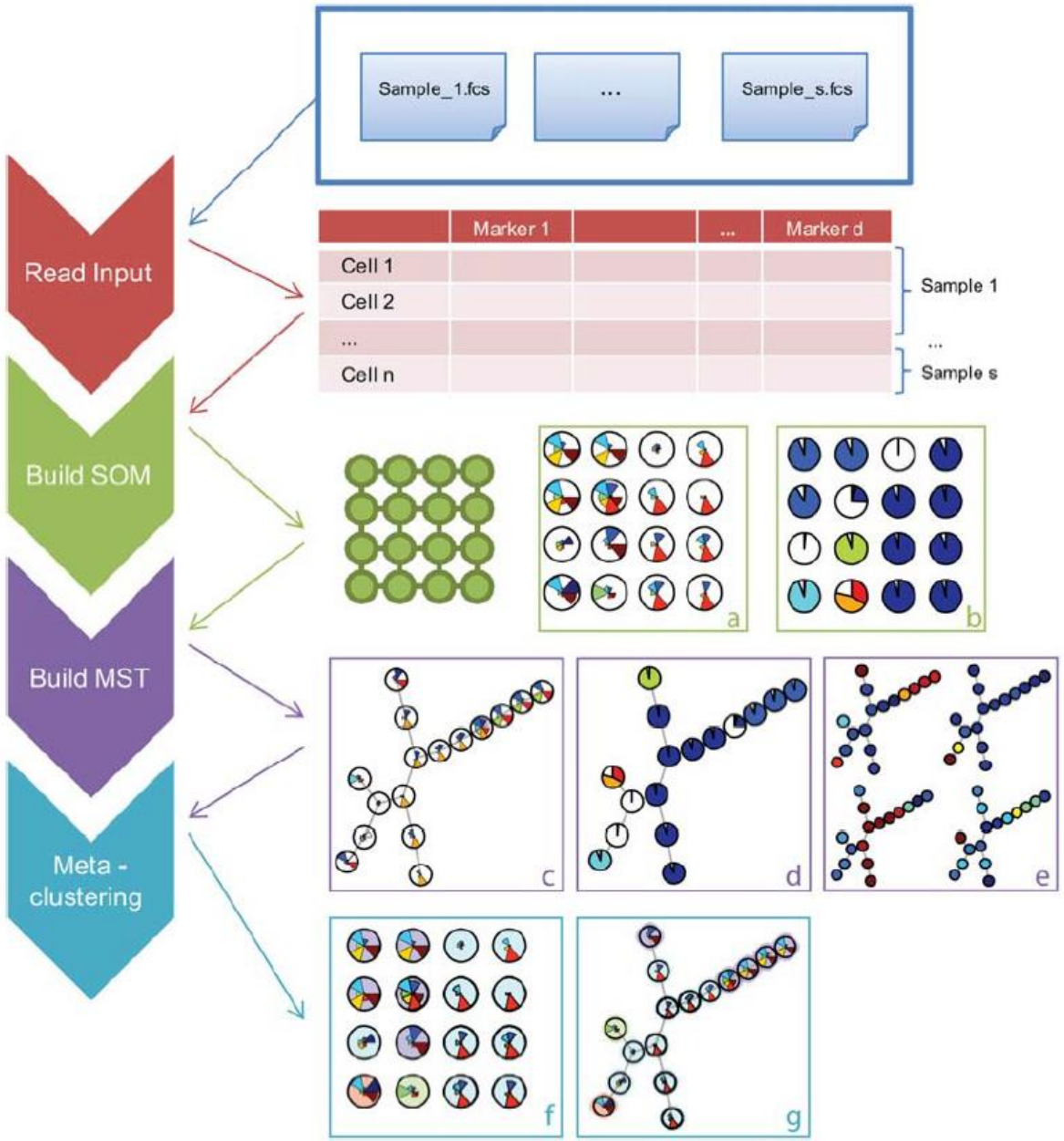
Input:
Matrix of size «*cell number * marker number*»

Pre-processing
Scaling of all features so that all columns/markers have a
Mean = 0
SD = 1
→ Equal importance of all markers

Intermediate visualization:
More Clusters than Cell types
To have «pure» classes
→ reduce subsequently

MST:
Nodes that are similar to each other
are closer connected
→ represents multidimensional topology

Output:
Each cell belongs to unknown cell class / type
Clusters correspond to cell types



Input:
Matrix of size «*cell number * marker number*»

Pre-processing
Scaling of all features so that all columns/markers have a
Mean = 0
SD = 1
→ Equal importance of all markers

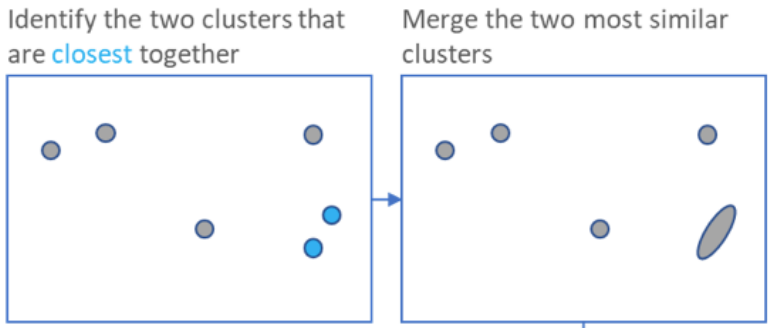
Intermediate visualization:
More Clusters than Cell types
To have «pure» classes
→ reduce subsequently

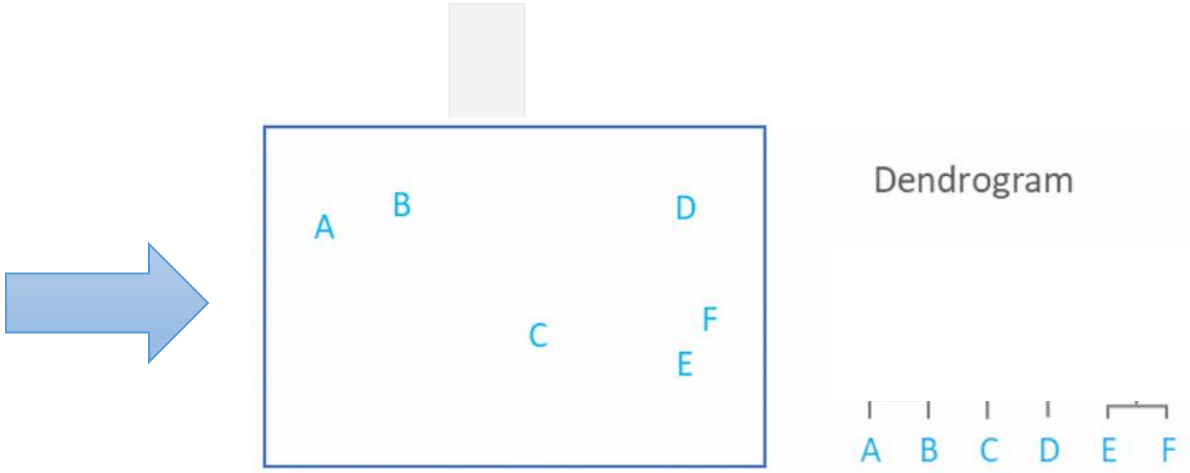
MST:
Nodes that are similar to each other
are closer connected
→ represents multidimensional topology

Metaclustering:
Clustering of Node-centres
to obtain final Cell Classes

Output:
Each cell belongs to unknown cell class / type
Clusters correspond to cell types

Hierarchical clustering



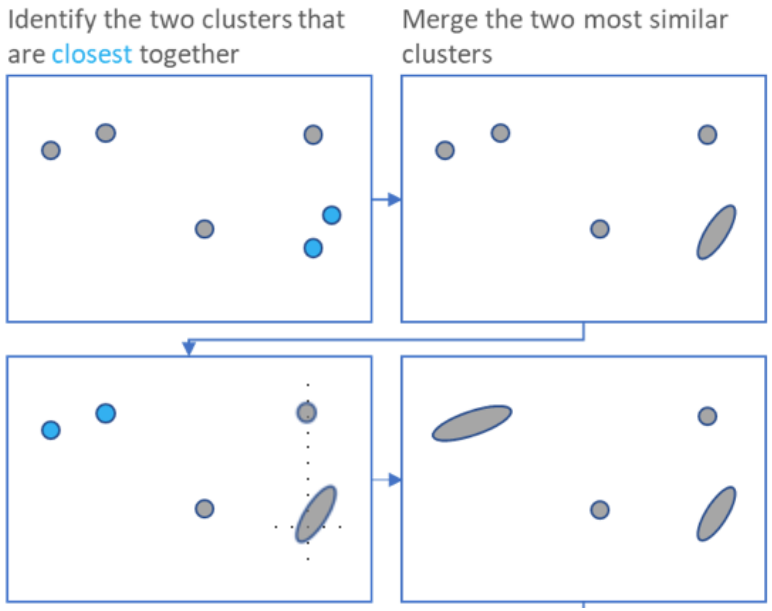


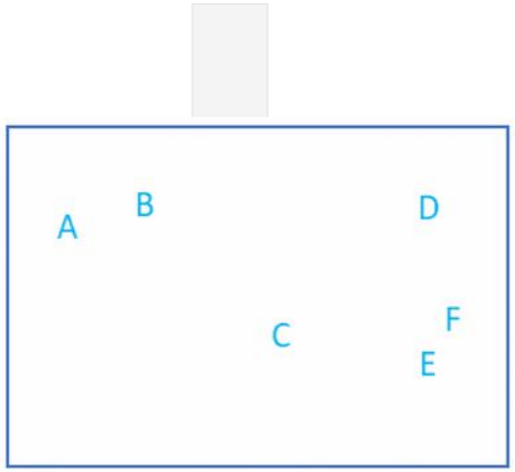
Metaclustering:
Clustering of Node-centres
to obtain final Cell Classes

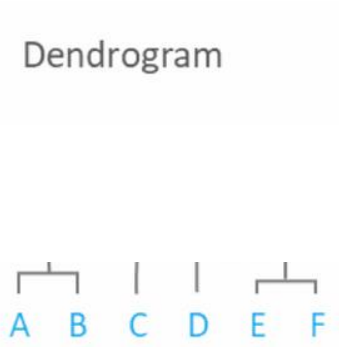
Output:
Each cell belongs to unknown cell class / type
Clusters correspond to cell types

https://www.displayr.com/what-is-hierarchical-clustering/

Hierarchical clustering





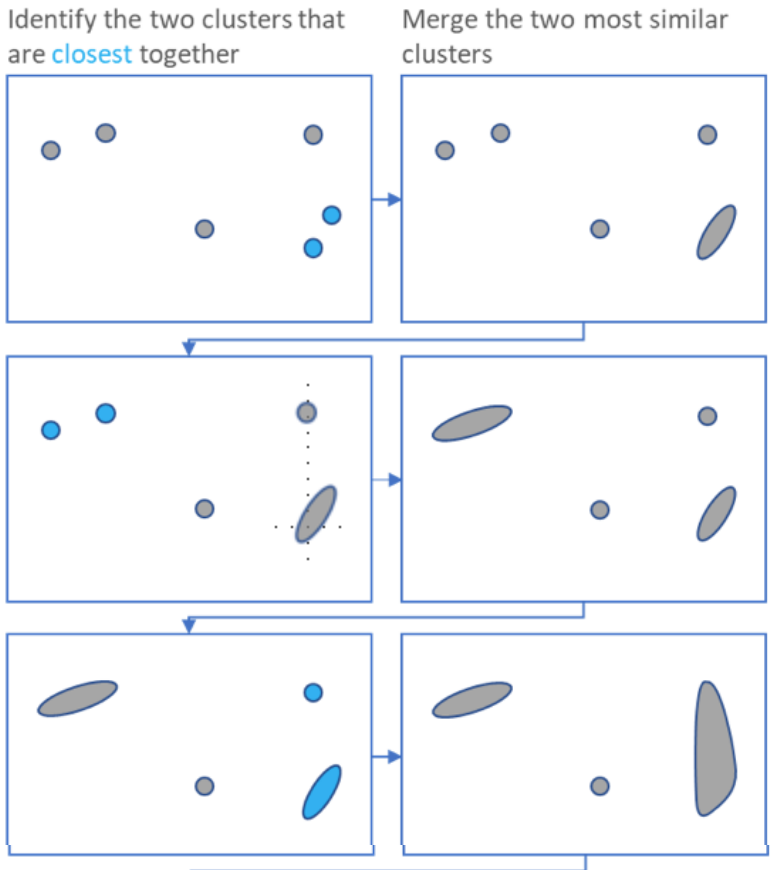


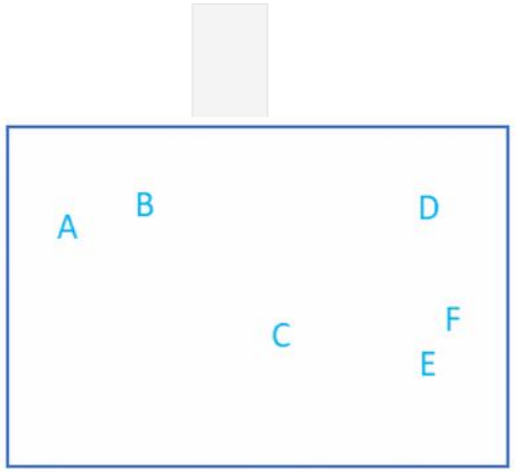
Metaclustering:
Clustering of Node-centres
to obtain final Cell Classes

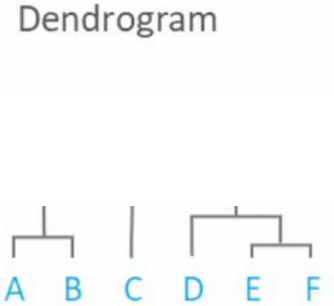
Output:
Each cell belongs to unknown cell class / type
Clusters correspond to cell types

https://www.displayr.com/what-is-hierarchical-clustering/

Hierarchical clustering





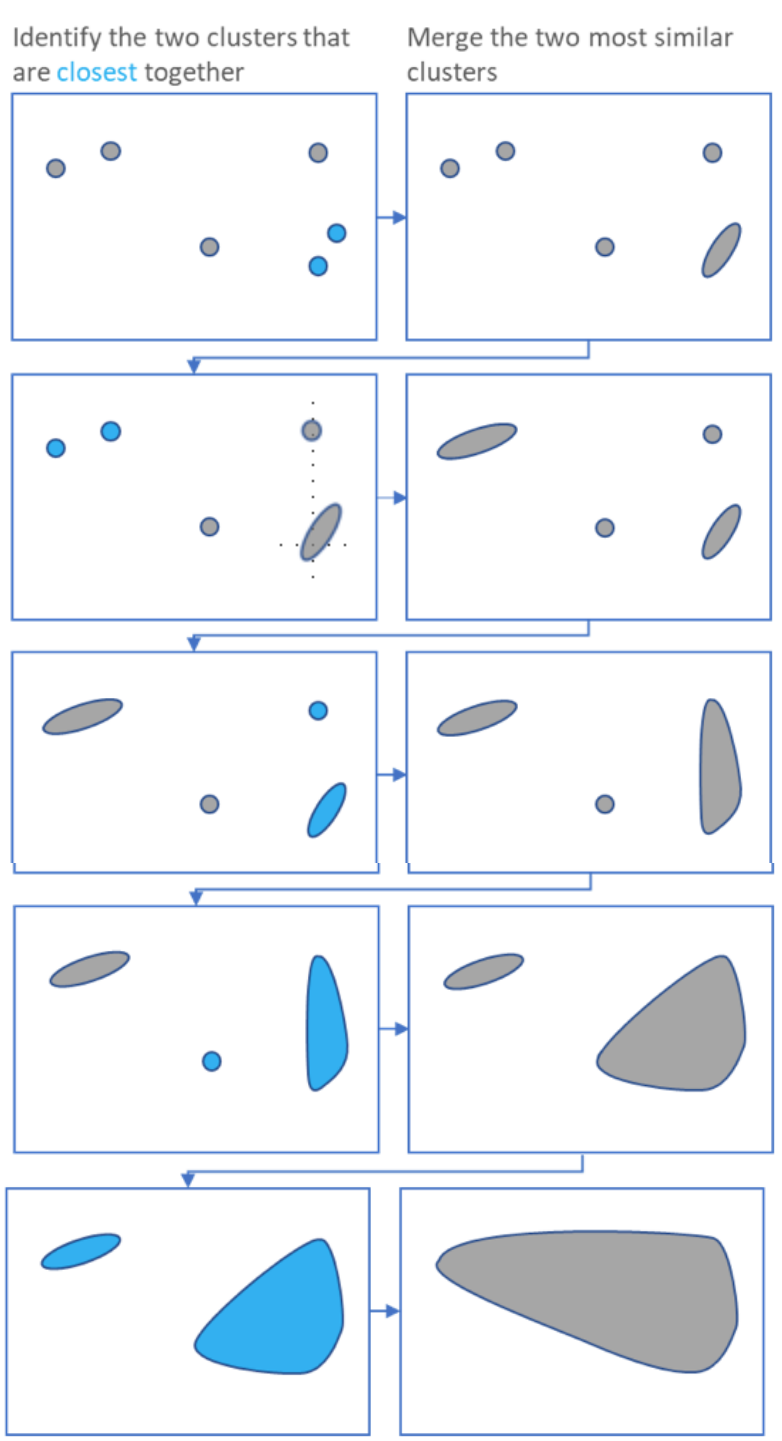


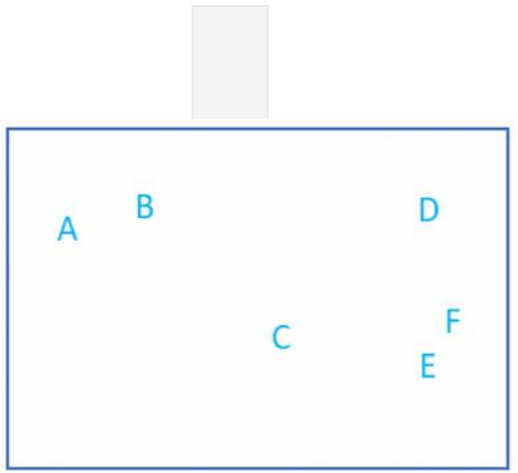
Metaclustering:
Clustering of Node-centres
to obtain final Cell Classes

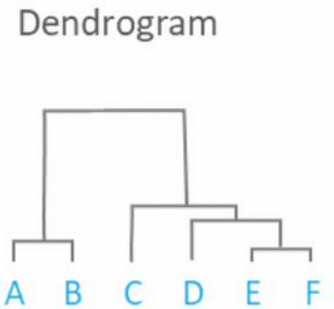
Output:
Each cell belongs to unknown cell class / type
Clusters correspond to cell types

https://www.displayr.com/what-is-hierarchical-clustering/

Hierarchical clustering







Metaclustering:
Clustering of Node-centres
to obtain final Cell Classes

Output:
Each cell belongs to unknown cell class / type
Clusters correspond to cell types

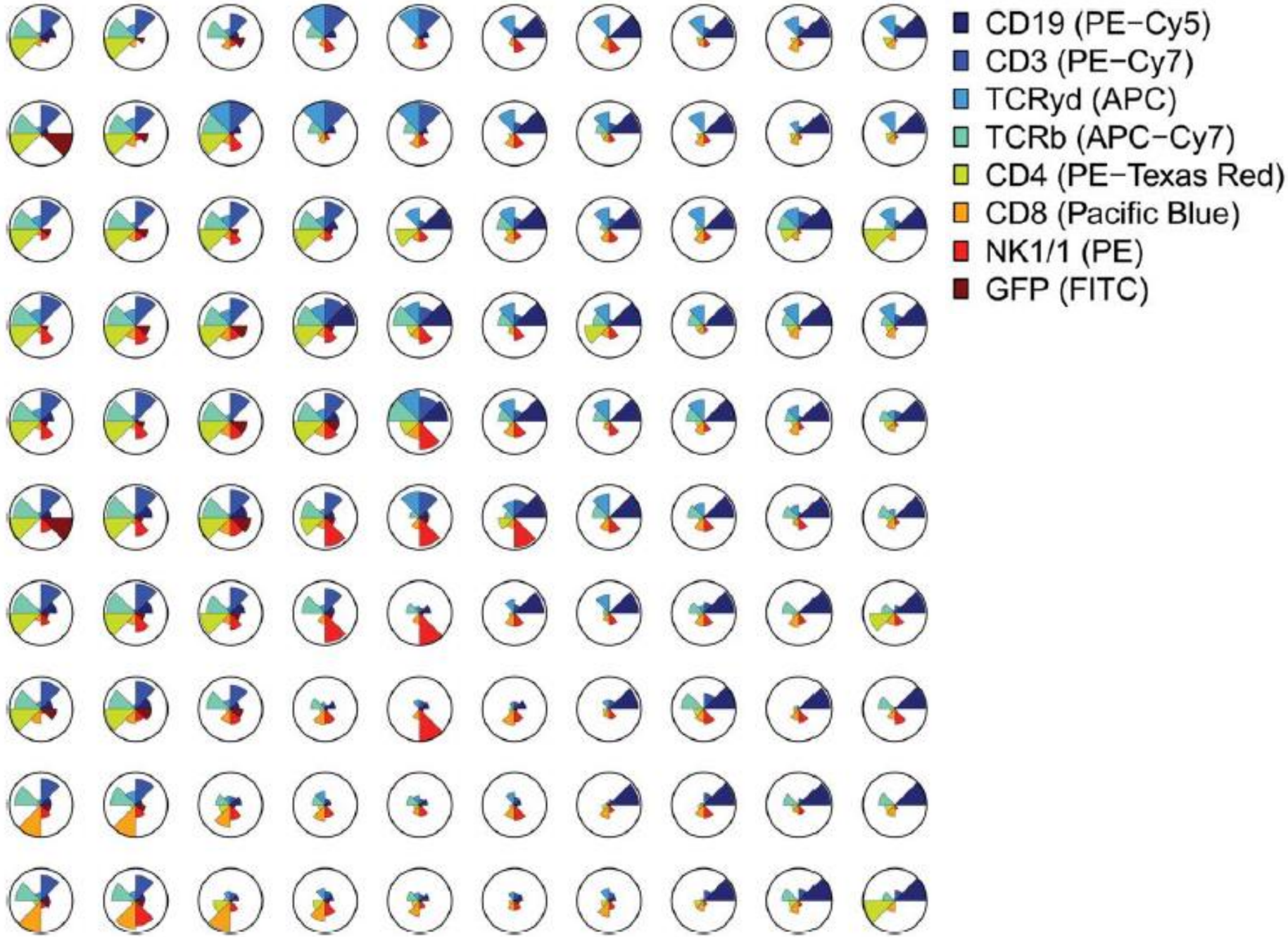
https://www.displayr.com/what-is-hierarchical-clustering/

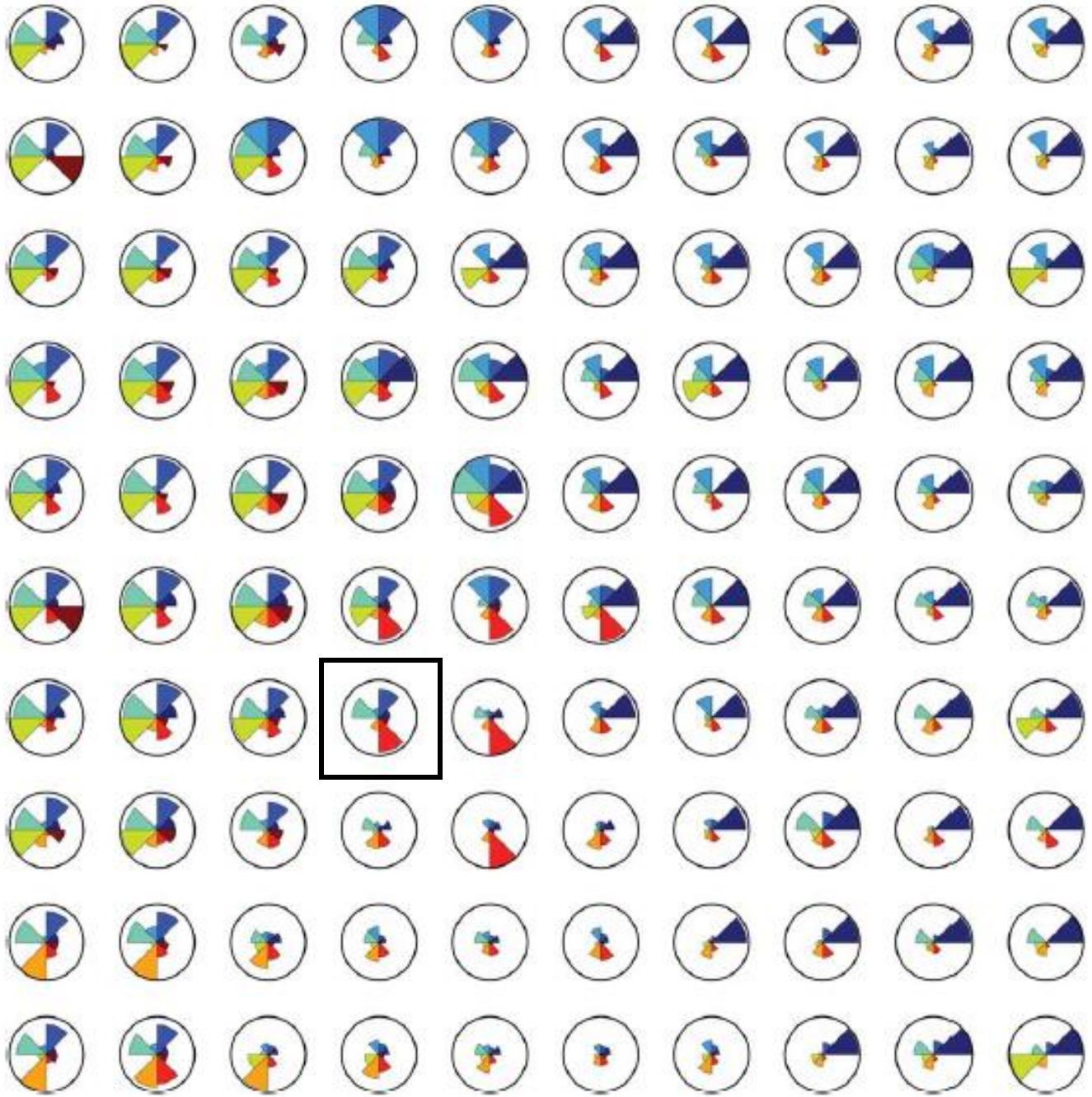
Published FACS datasets

- Comparison of FlowSOM with manual gating
- 7 surface markers + GFP

Table 1. Overview of datasets used in the results section

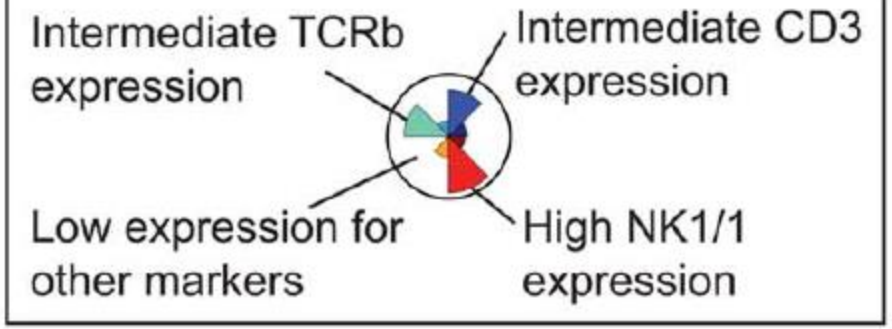
DATASET	MARKERS	SAMPLES	NUMBER OF EVENTS	NUMBER OF CELLS
FlowCAP I:				
Diffuse large B-cell lymphoma	3	30	–	308,676
Graft versus host disease	4	12	–	207,171
Hematopoietic stem cell transplant	4	30	–	278,005
Normal donor	10	30	–	1,778,883
Symptomatic West Nile virus	6	13	–	1,214,373
BAL staining	8	12	3,635,910	411,143
Bone marrow	31	3	1,264,755	660,084

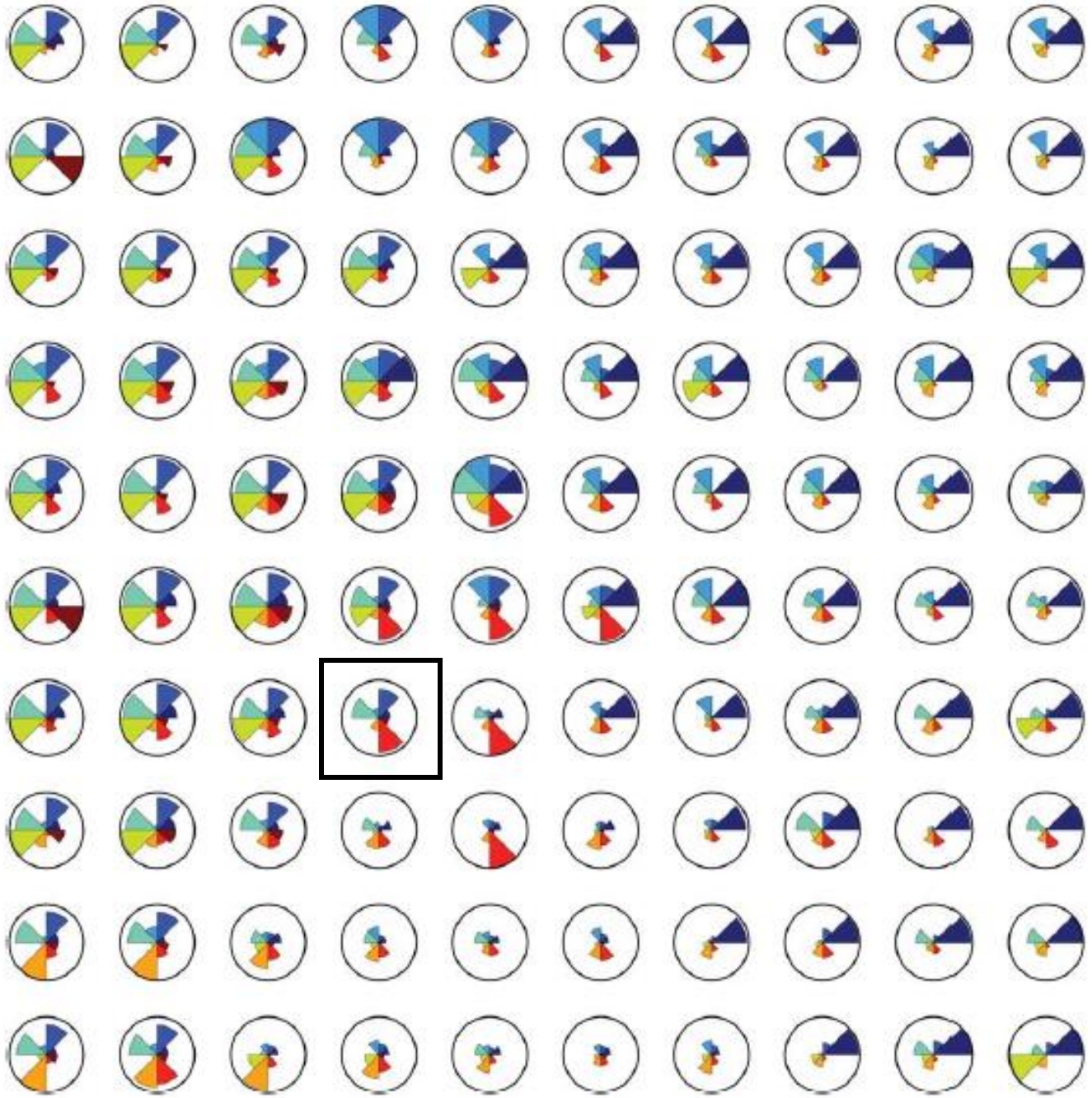

CD19 (PE-Cy5)
CD3 (PE-Cy7)
TCRyd (APC)
TCRb (APC-Cy7)
CD4 (PE-Texas Red)
CD8 (Pacific Blue)
NK1/1 (PE)
GFP (FITC)



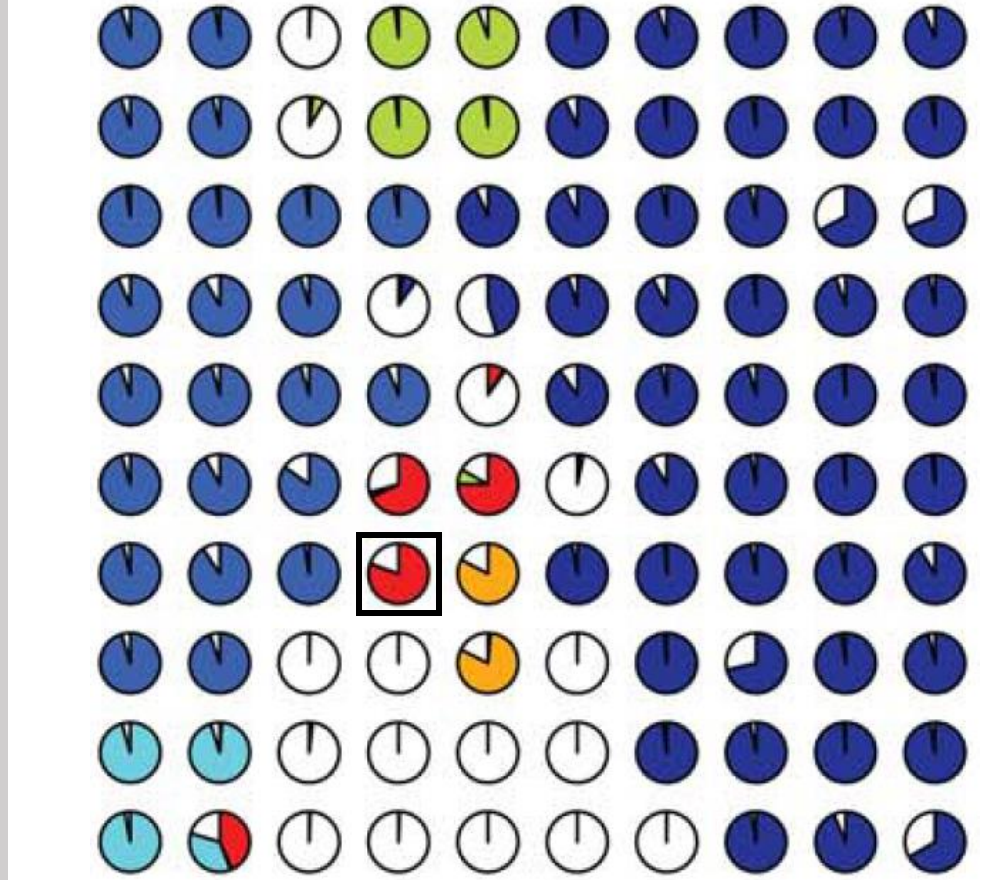
CD19 (PE-Cy5)
CD3 (PE-Cy7)
TCRyd (APC)
TCRb (APC-Cy7)
CD4 (PE-Texas Red)
CD8 (Pacific Blue)
NK1/1 (PE)
GFP (FITC)

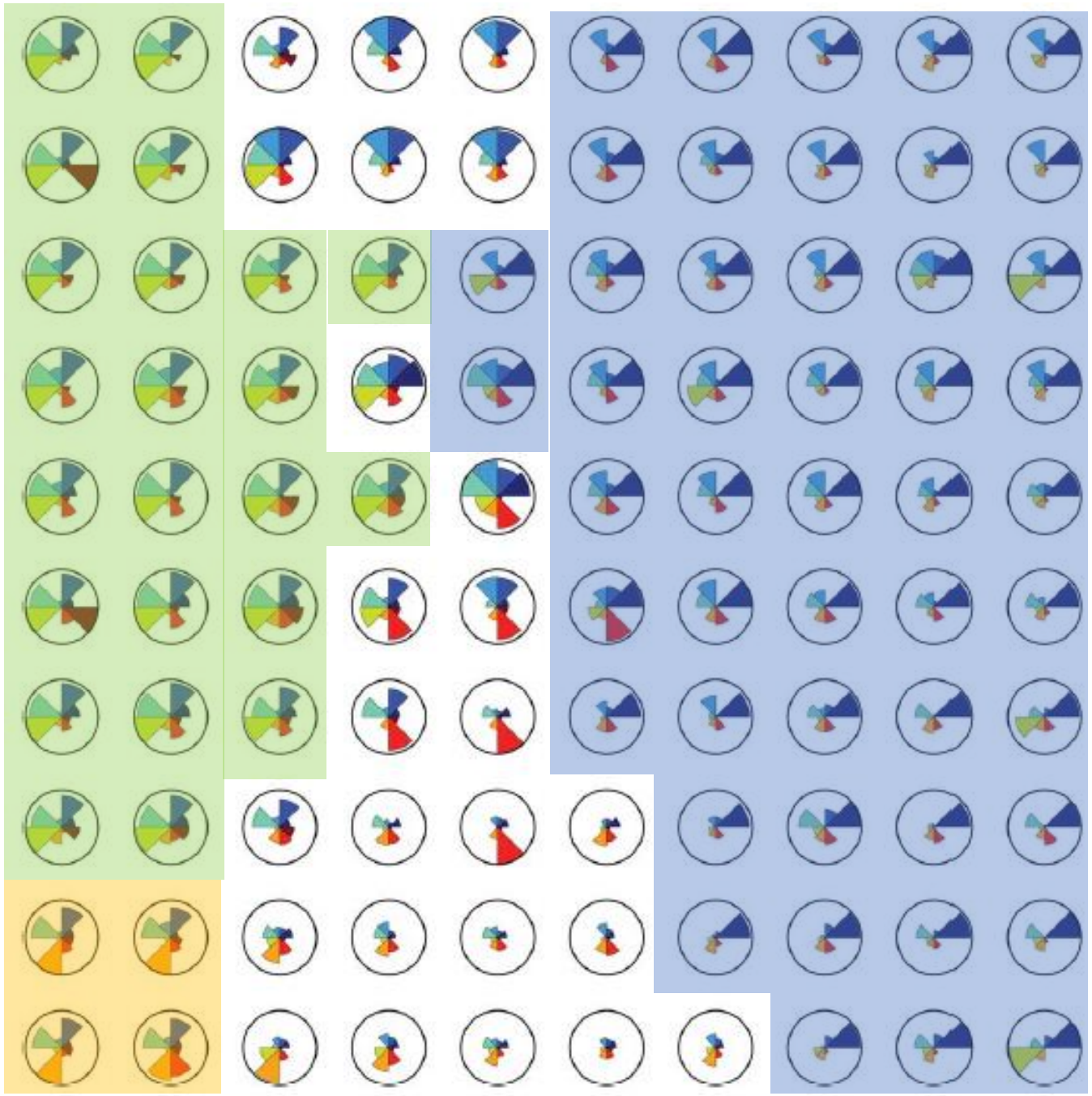
Example: a node with NKT cells





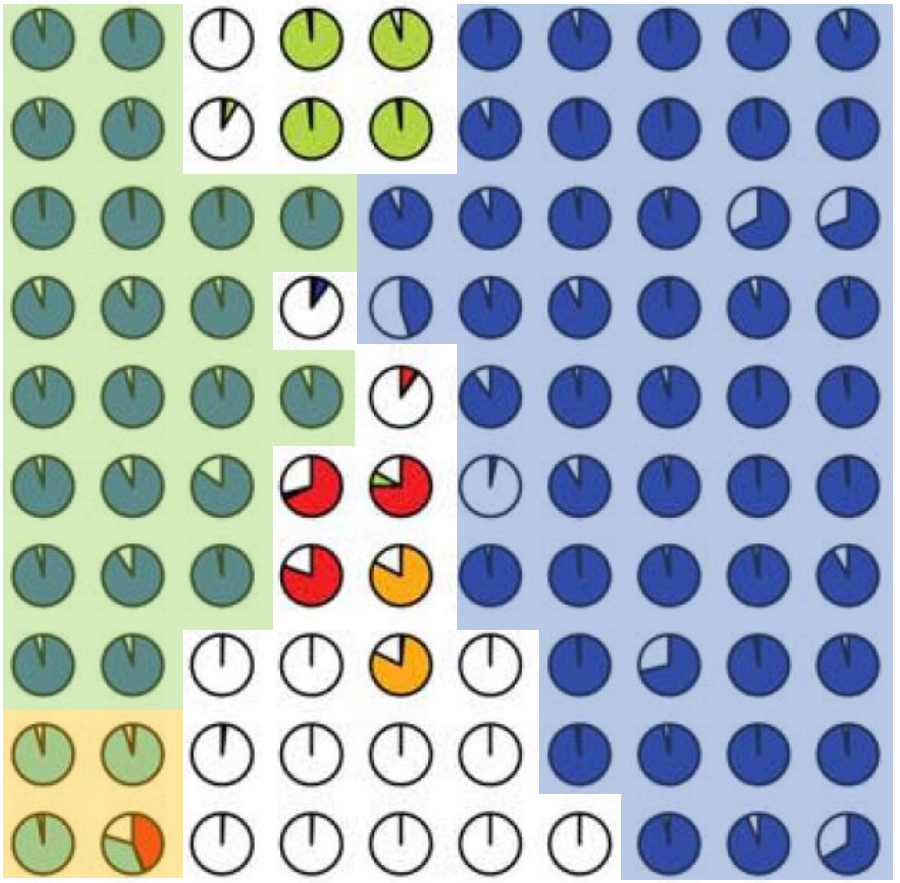
Manually gated cell types in SOM nodes:





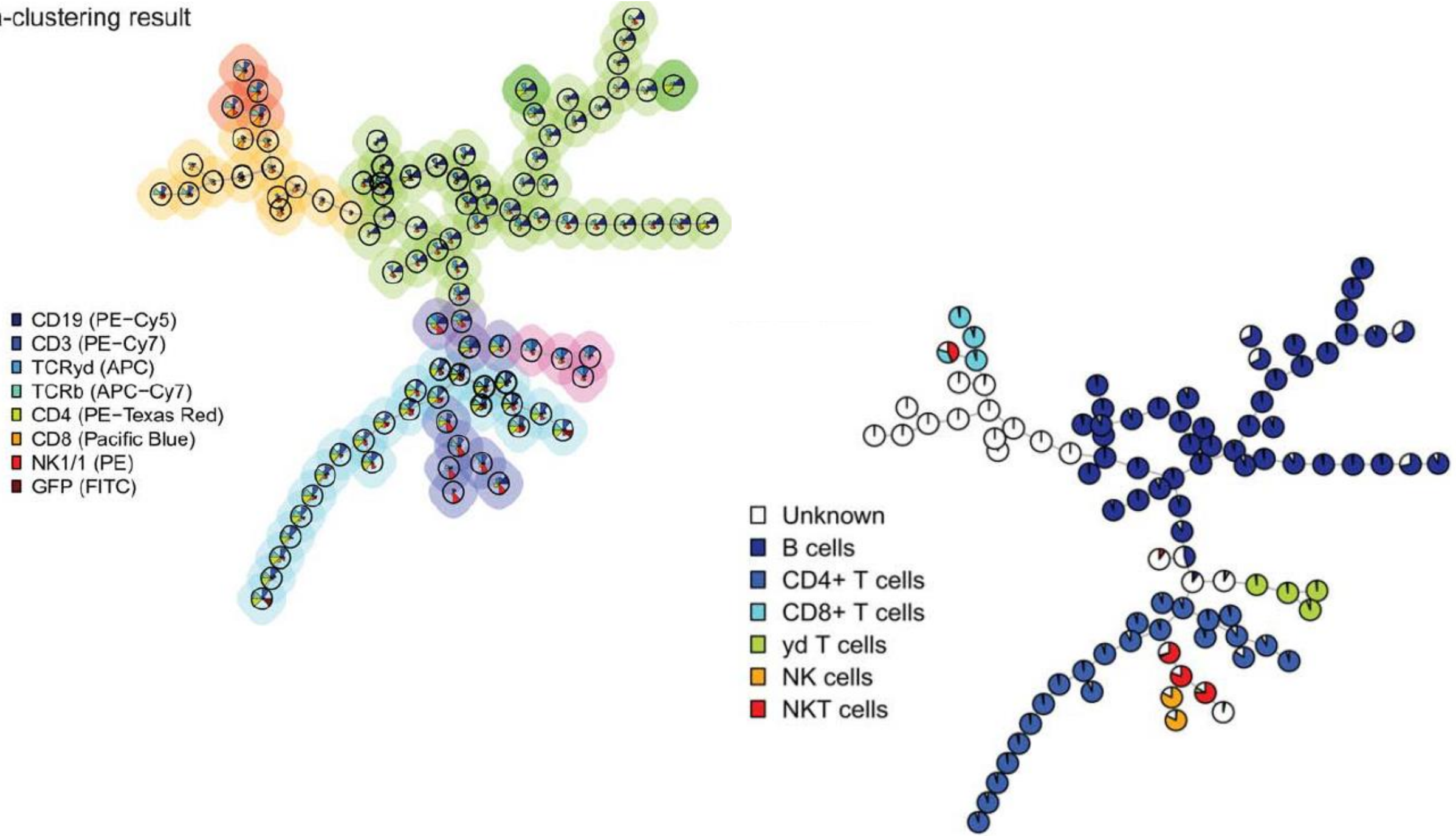
CD19 (PE–Cy5)
CD3 (PE–Cy7)
TCRyd (APC)
TCRb (APC–Cy7)
CD4 (PE–Texas Red)
CD8 (Pacific Blue)
NK1/1 (PE)
GFP (FITC)

Manually gated cell types in SOM nodes:

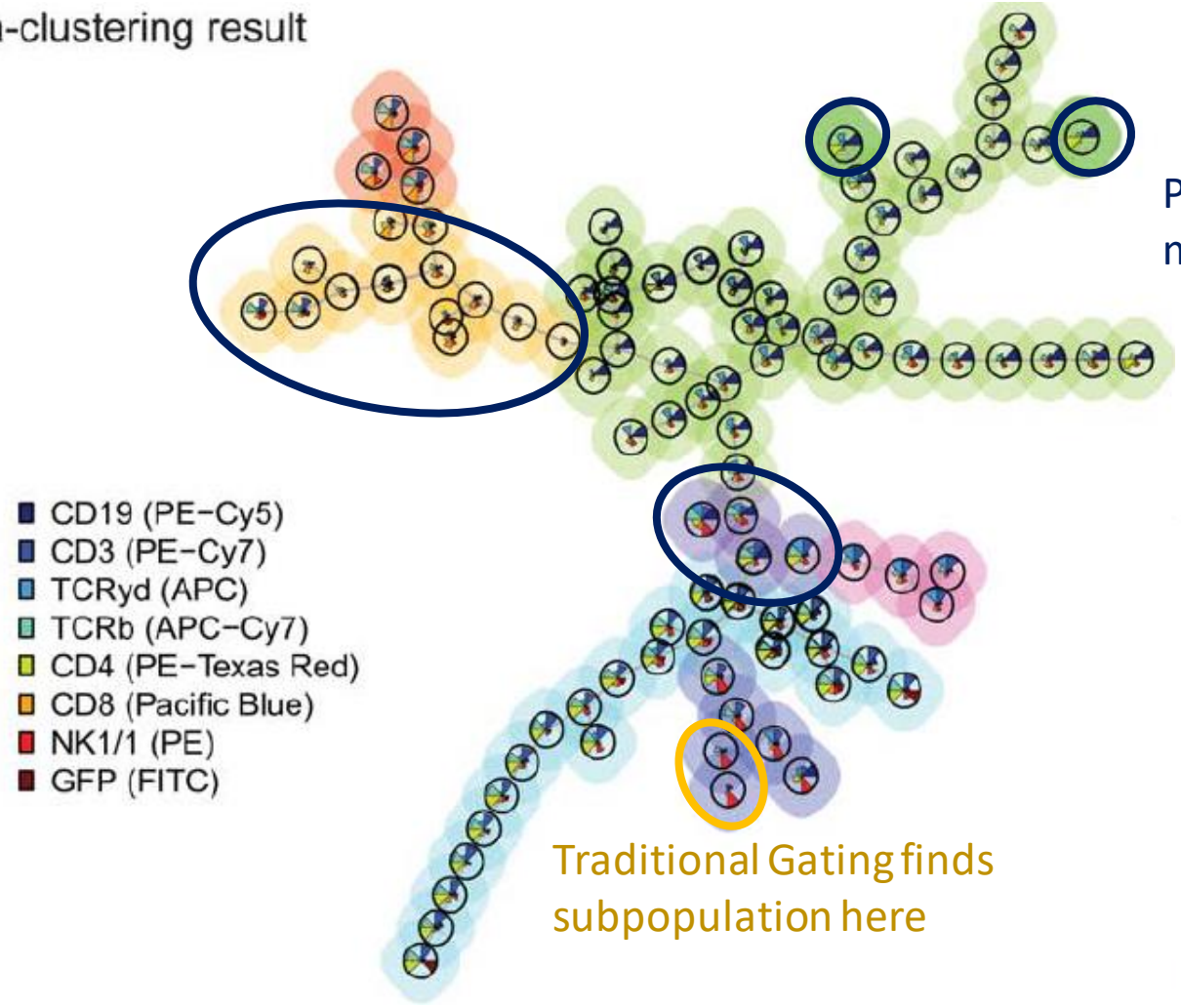


Unknown
B cells
CD4+ T cells
CD8+ T cells
γδ T cells
NK cells
NKT cells

C: Meta-clustering result

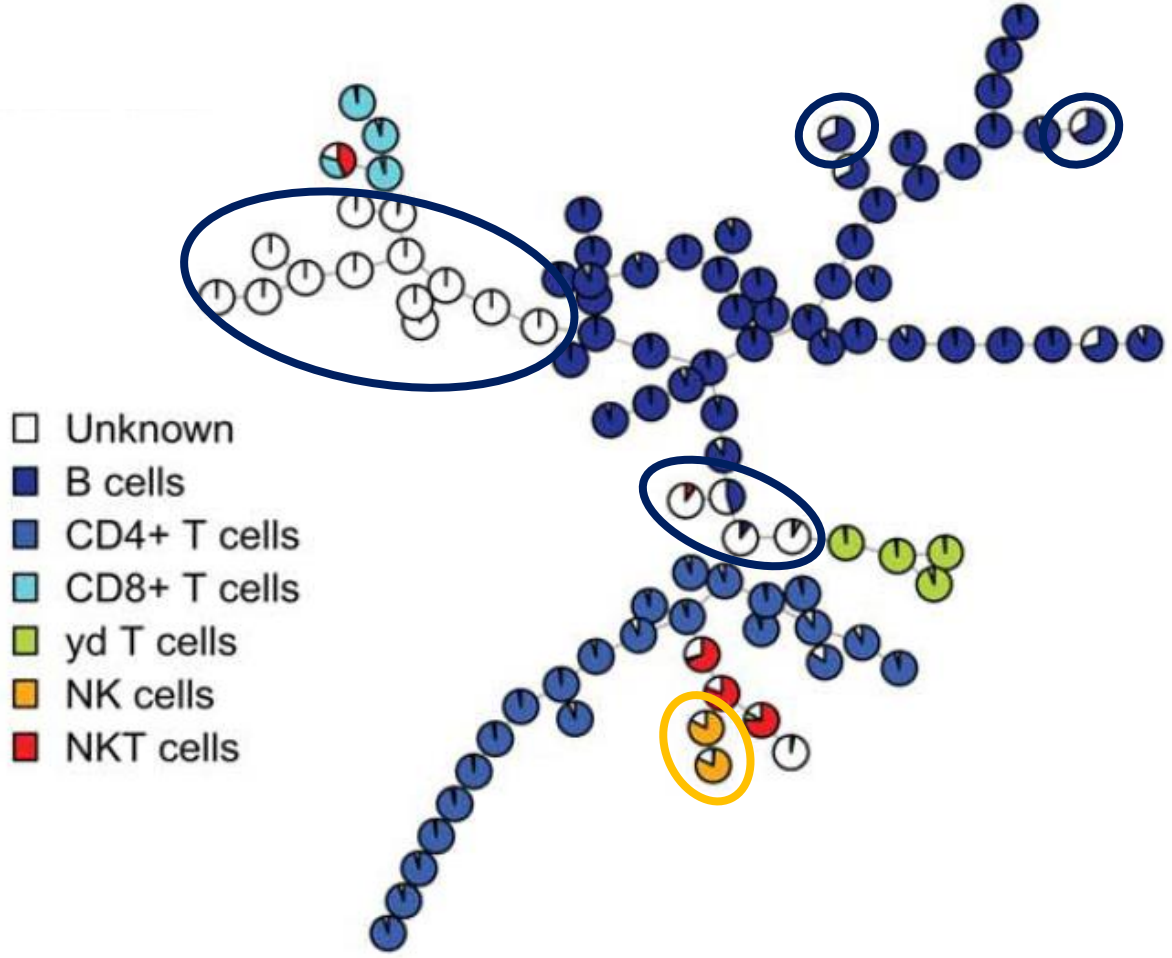


C: Meta-clustering result



Potentially previously missed subpopulations

Traditional Gating finds subpopulation here



ORIGINAL ARTICLE

Cytometry
PART A
Journal of the International Society for Advancement of Cytometry

MIFlowCyt

FlowSOM: Using Self-Organizing Maps for Visualization and Interpretation of Cytometry Data

Sofie Van Gassen,[1,2,3]* Britt Callebaut,[1] Mary J. Van Helden,[2,3] Bart N. Lambrecht,[2,3]
Piet Demeester,[1] Tom Dhaene,[1] Yvan Saeys[2,3]

- Overview over whole dataset
- No immediate exclusion of data by gating

- Specific 2D scatter plots can be created subsequently

- Unbiased discovery of small atypical, unknown or unexpected subsets
- Also some specific subsets might not be detected
  → Expert knowledge on cell types needed

Molecular Cell
Resource

A Systems-Level Study Reveals Regulators of Membrane-less Organelles in Human Cells

Doris Berchtold,[1] Nico Battich,[1] and Lucas Pelkmans[1,2,*]
[1]Institute of Molecular Life Sciences, University of Zurich, 8057 Zurich, Switzerland
[2]Lead Contact
*Correspondence: lucas.pelkmans@imls.uzh.ch
https://doi.org/10.1016/j.molcel.2018.10.036

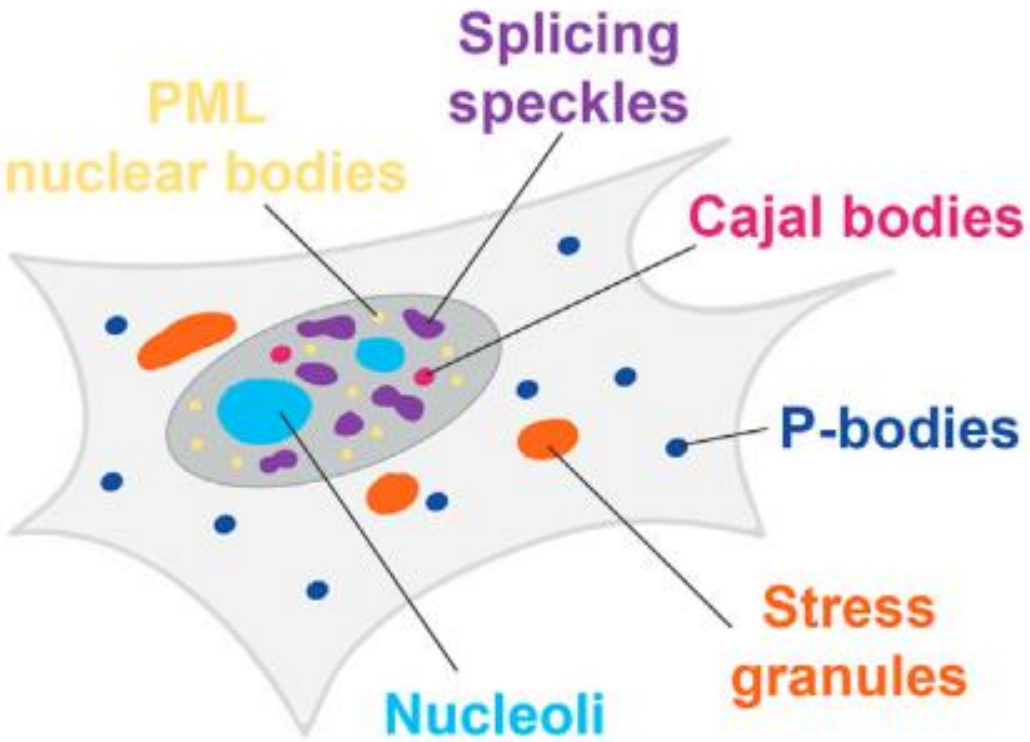
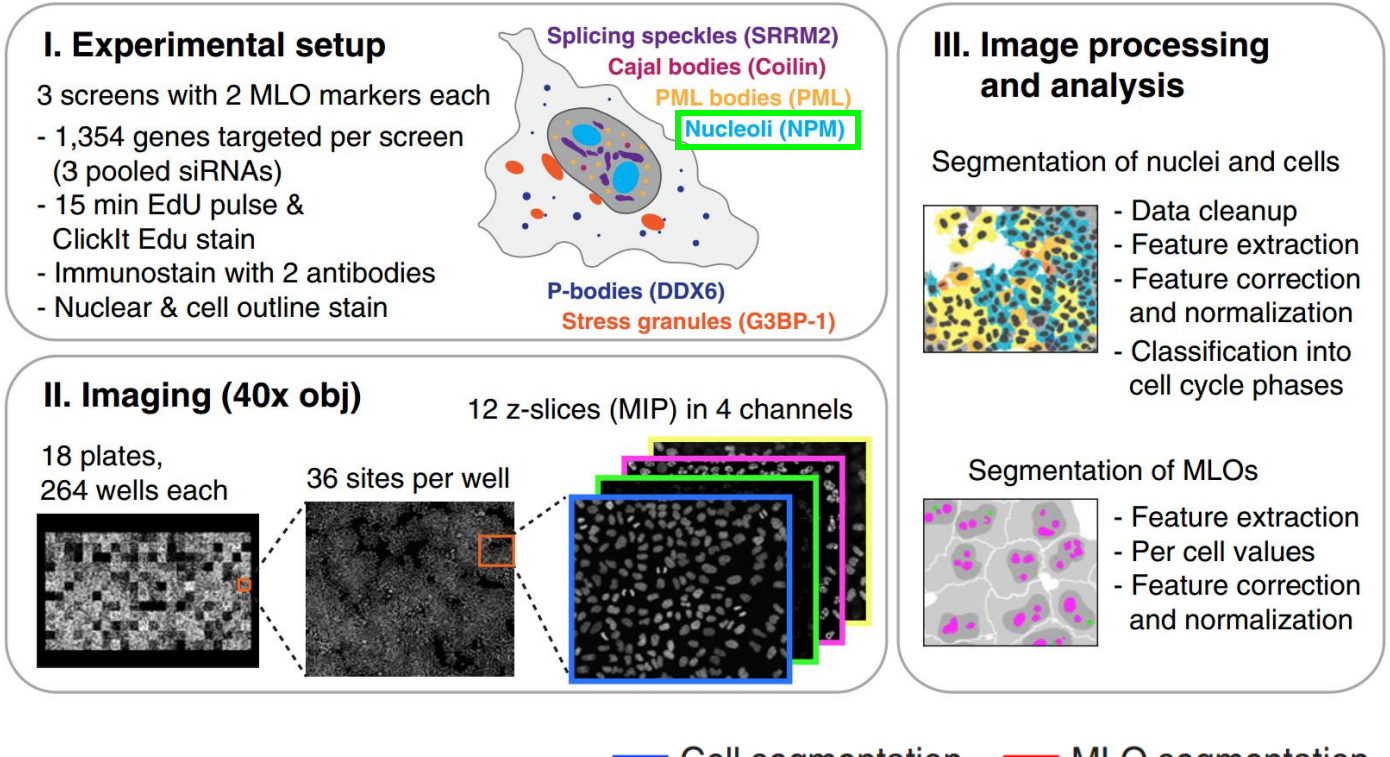
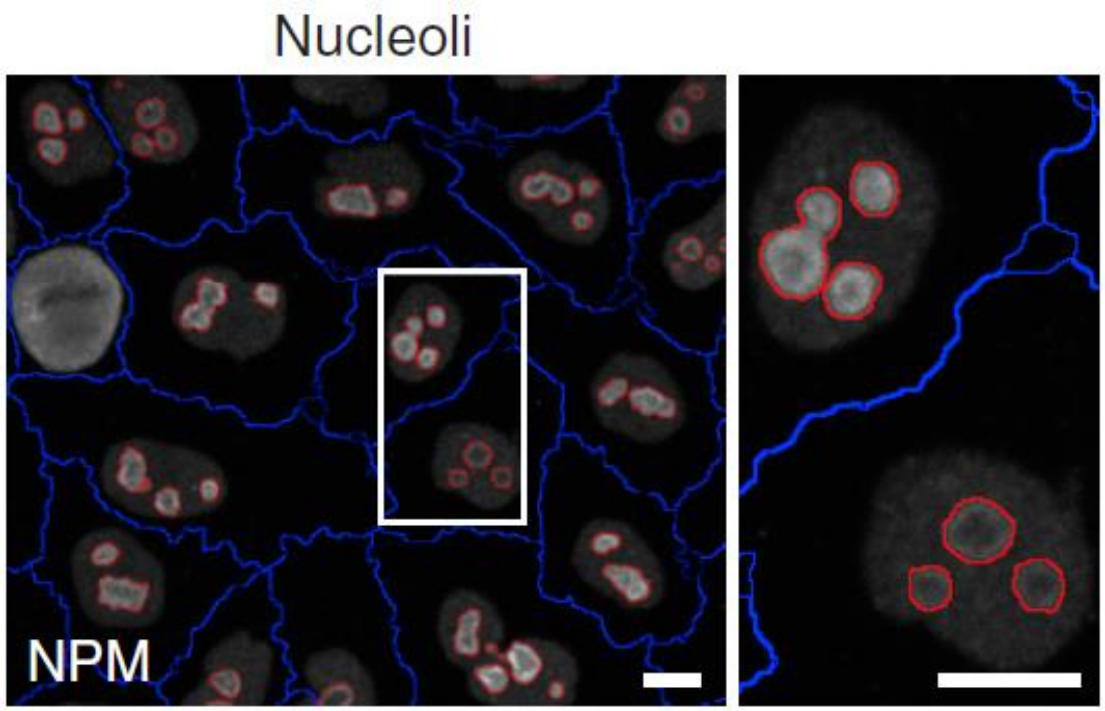
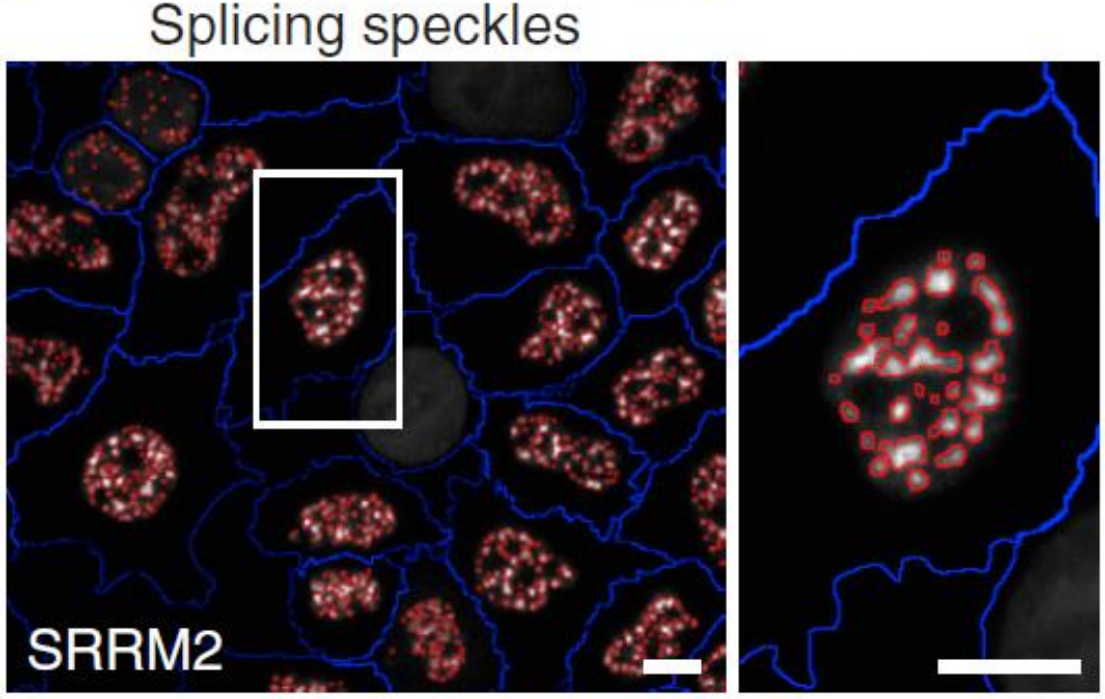


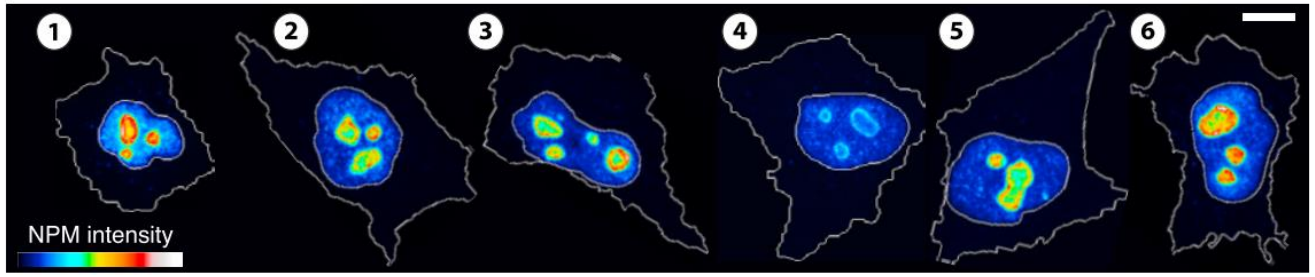
Image-based siRNA screen on membrale less organelles (MLO)





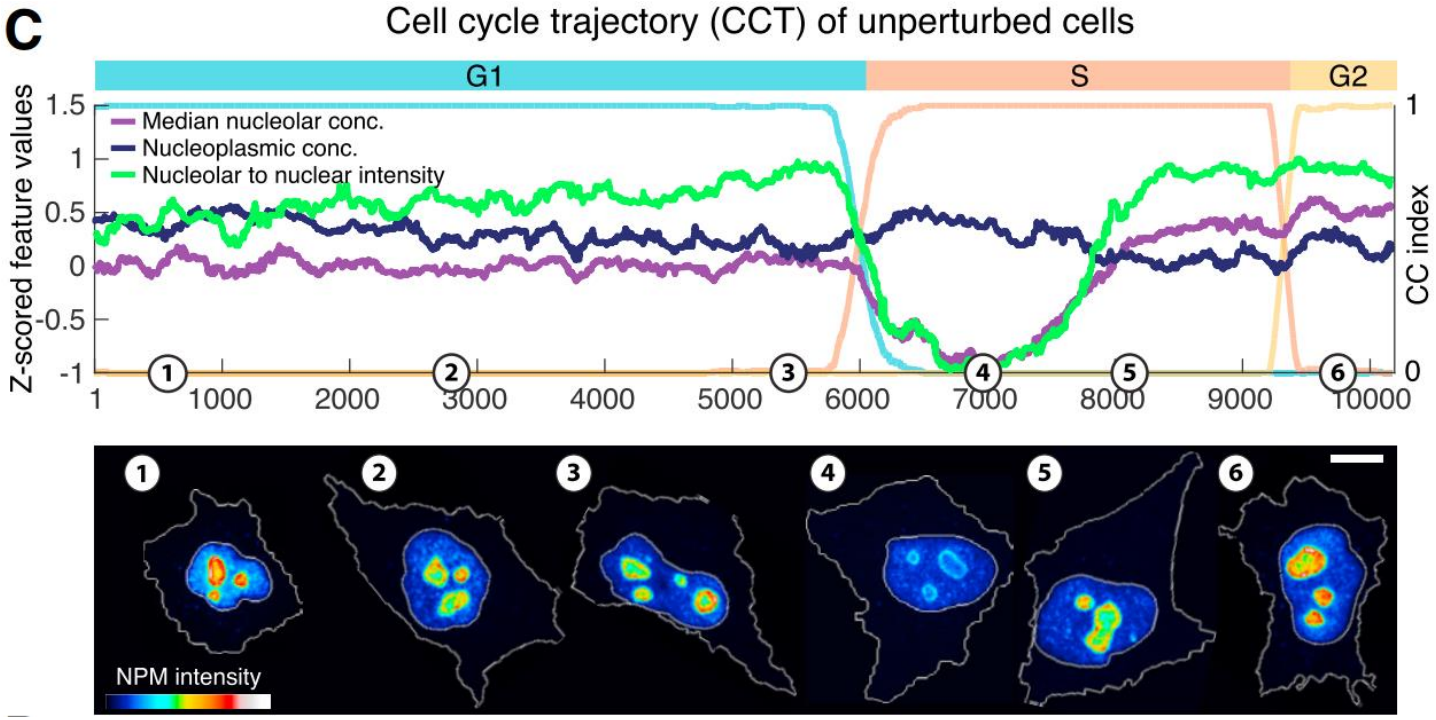


Problem:



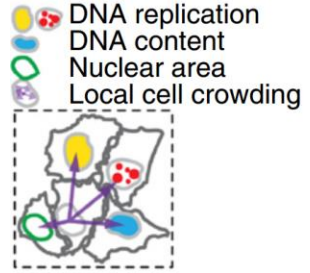
1. Look only at unperturbed control cells

Problem:



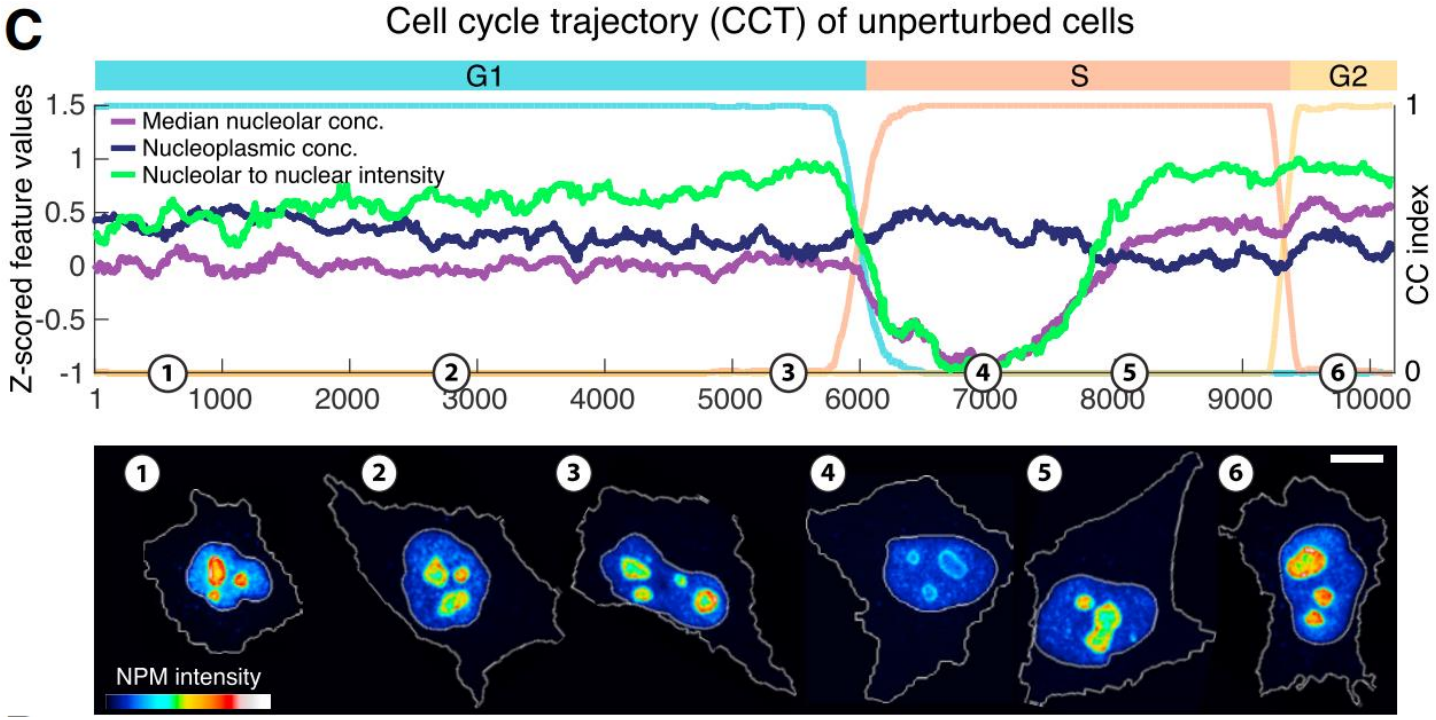
Trajectories of cell-cycle progression from fixed cell populations

Gabriele Gut[1,2,5], Michelle D Tadmor[3,5], Dana Pe'er[3,6], Lucas Pelkmans[1,6] & Prisca Liberali[1,4,6]



1. Look only at unperturbed control cells

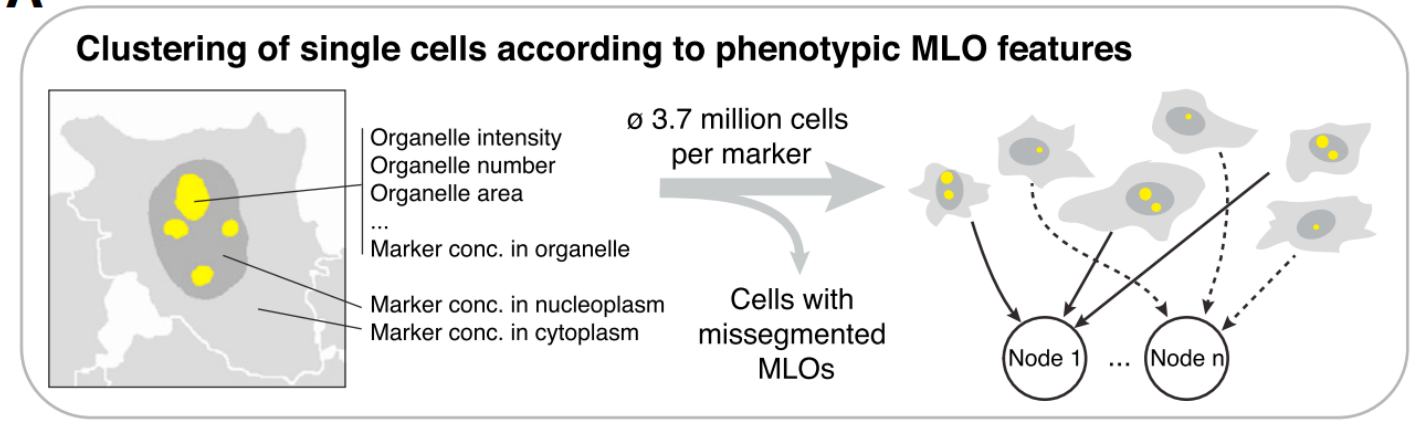
Problem:



Solution:

1. Look only at unperturbed control cells

→ Depending on cell cycle stage, cells are clustering to different nodes



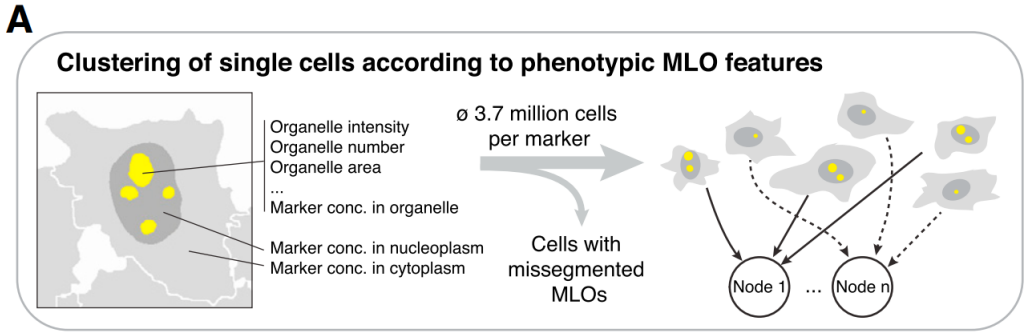
2. SOM Clustering of perturbed and unperturbed cells

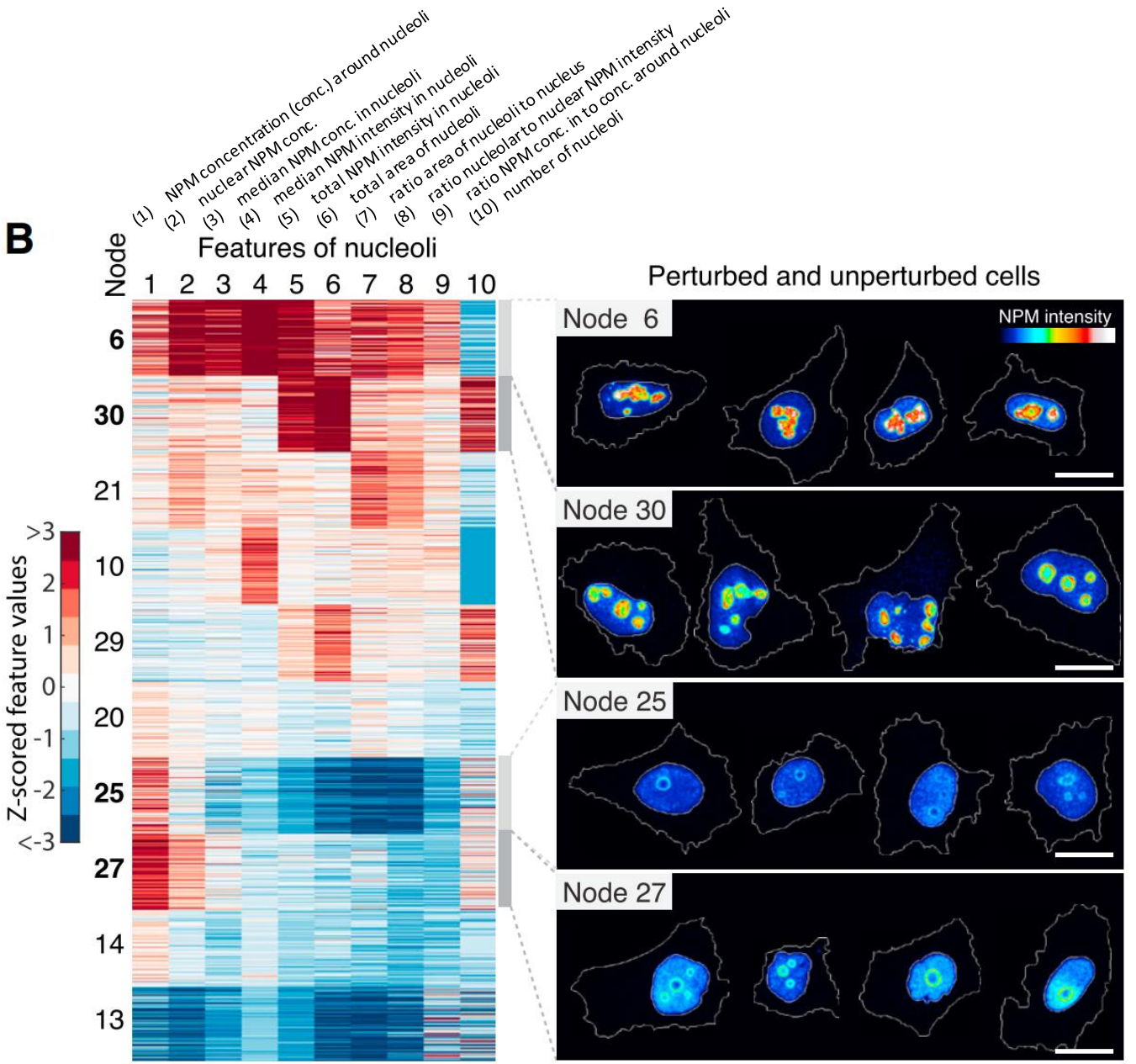
Strategy:

- Sequential clustering with high node numbers (e.g. 100)
- Cells in nodes with long distance to majority of nodes were missegmented (visual inspection)
  → excluded
- Do next clustering with remaining cells → … → …
- Node number reduced stepwise
  so that all cells in one node are highly similar: 15 - 30

2. SOM Clustering of perturbed and unperturbed cells
→ 30 nodes

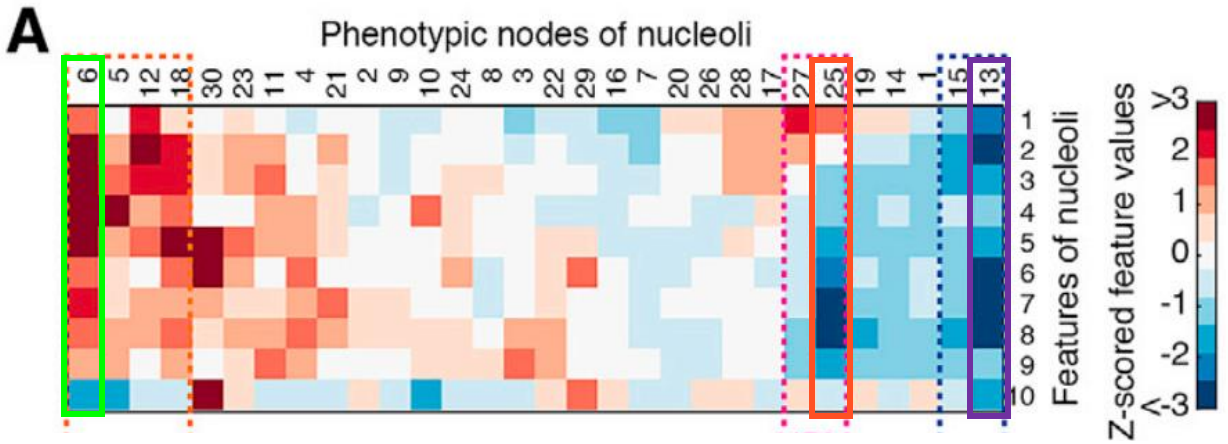
Solution:

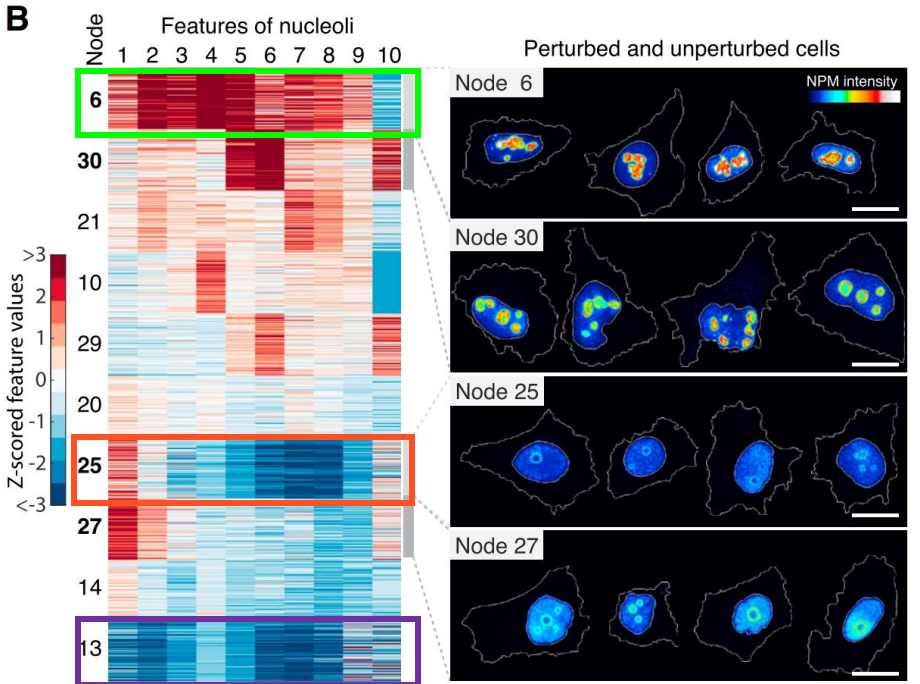




2. SOM Clustering of perturbed and unperturbed cells → 30 nodes

→ 3. **Median** of each feature in each node

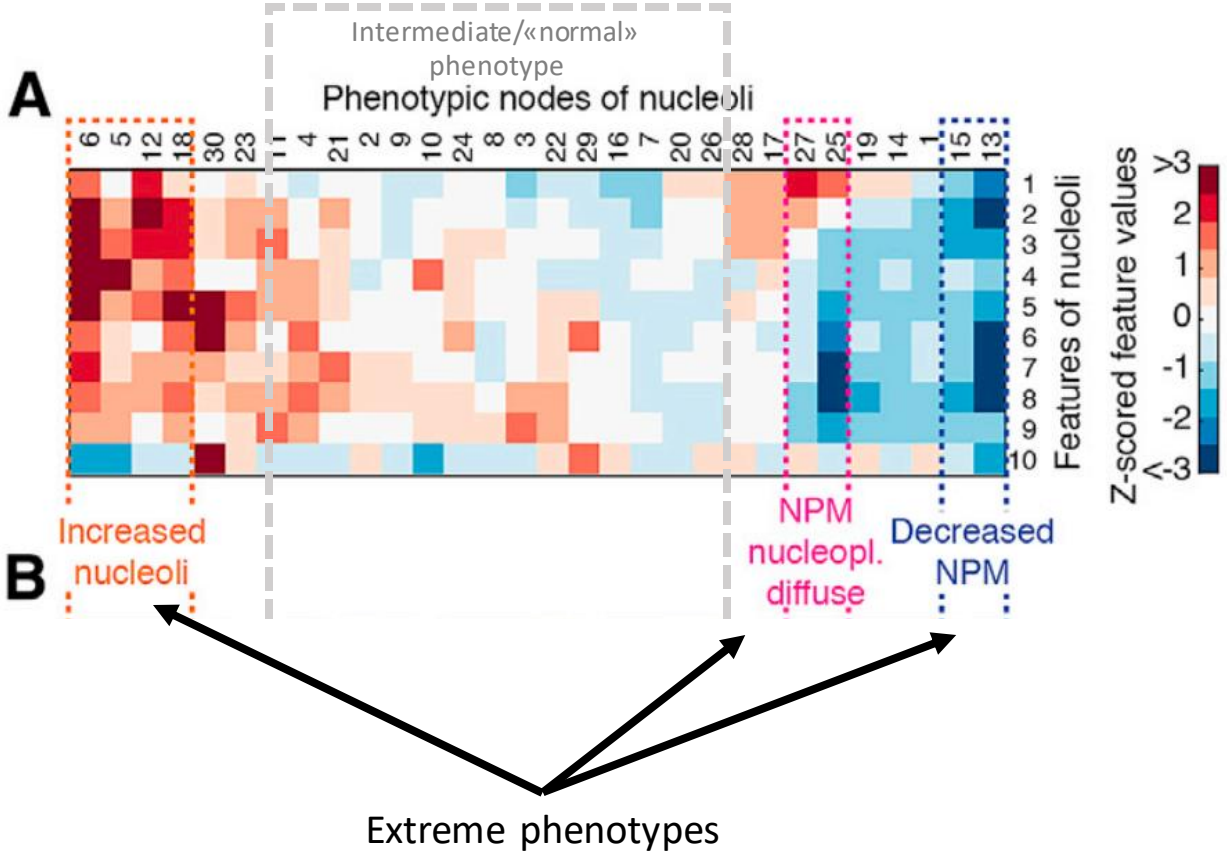




2. SOM Clustering of perturbed and unperturbed cells

→ 3. Median of each feature in each node

→ 4. Define **extreme or average** phenotype



Node selection:

- Median feature strength
- Absence of control cells
- Visual inspection

Increased MLO: high intensity of marker, high area

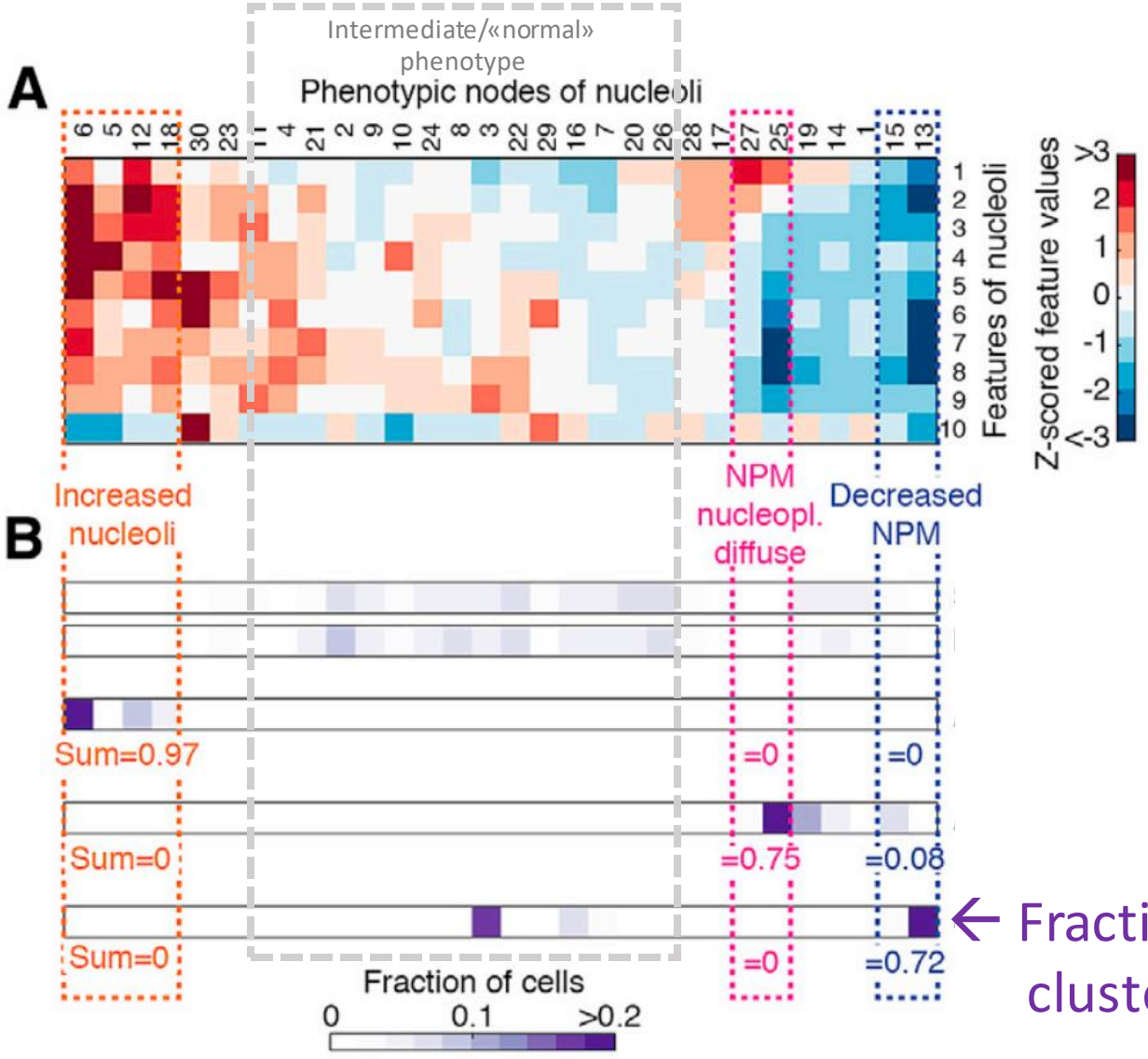
Decreased MLO: low intensity, small area or absent

Diffuse MLO: diffuse localization in nucleoplasm

2. SOM Clustering of perturbed and unperturbed cells

→ 3. Median of each feature in each node

→ 4. Define **extreme or average** phenotype



Node selection:

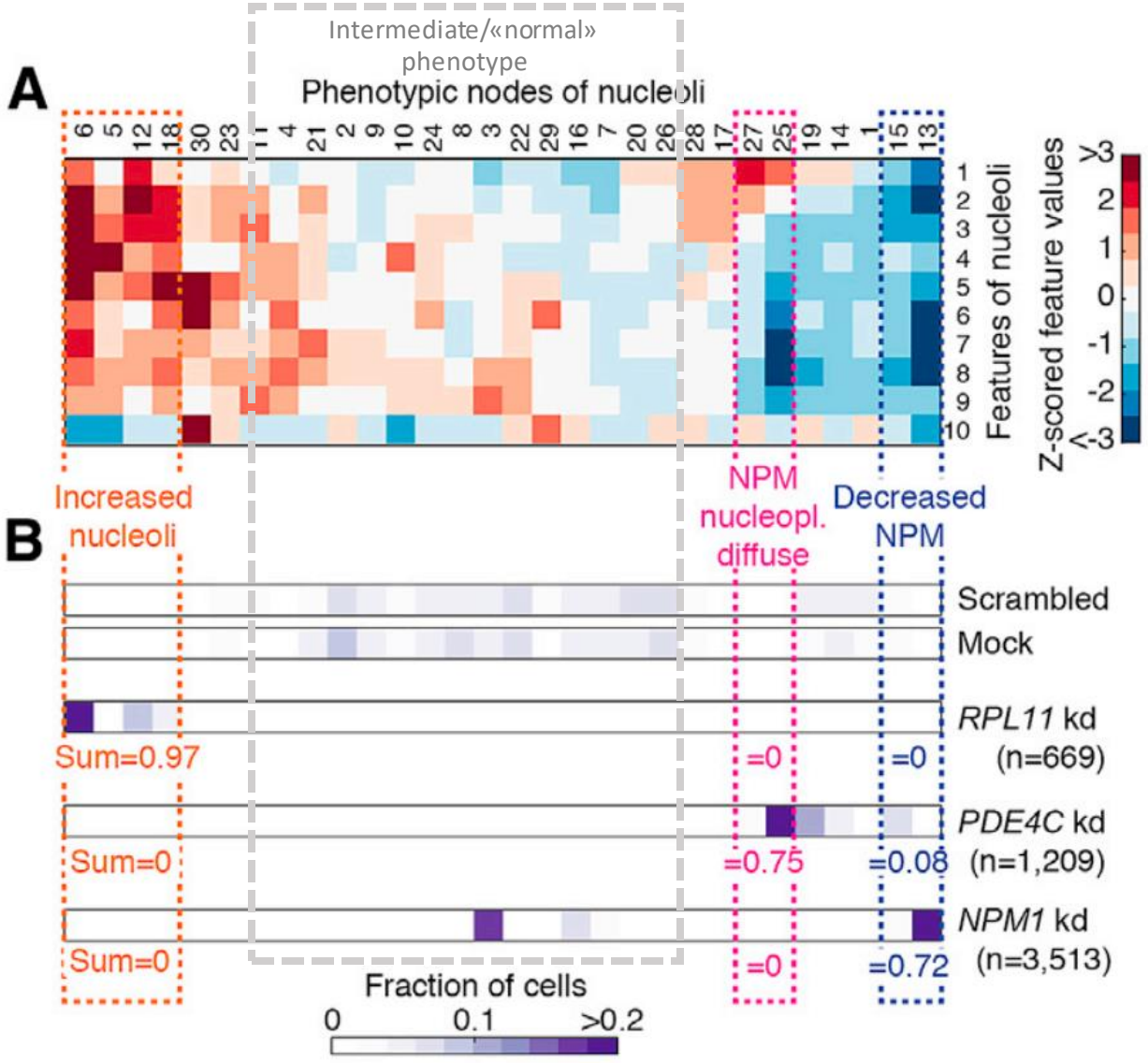
- Median feature strength
- **Absence of control cells**
- Visual inspection

← Fraction of cells per gene perturbation clustering to each node

2. SOM Clustering of perturbed and unperturbed cells

→ 3. Median of each feature in each node

→ 4. Define **extreme or average** phenotype



Node selection:

- Median feature strength
- **Absence of control cells**
- Visual inspection

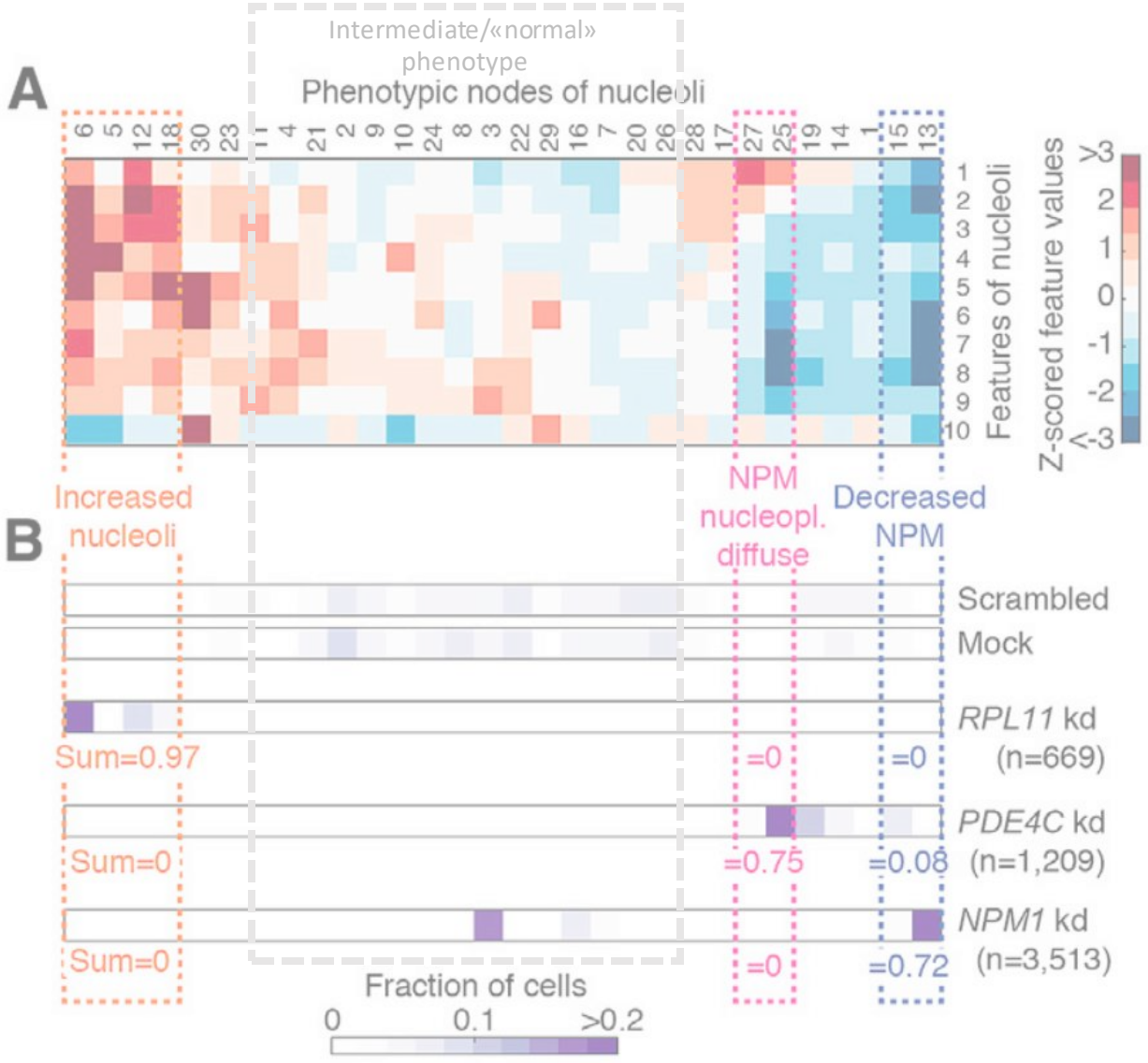
Unperturbed cells cluster to «**average**» nodes

Perturbed cells may cluster to nodes with **higher or lower** feature values

2. SOM Clustering of perturbed and unperturbed cells

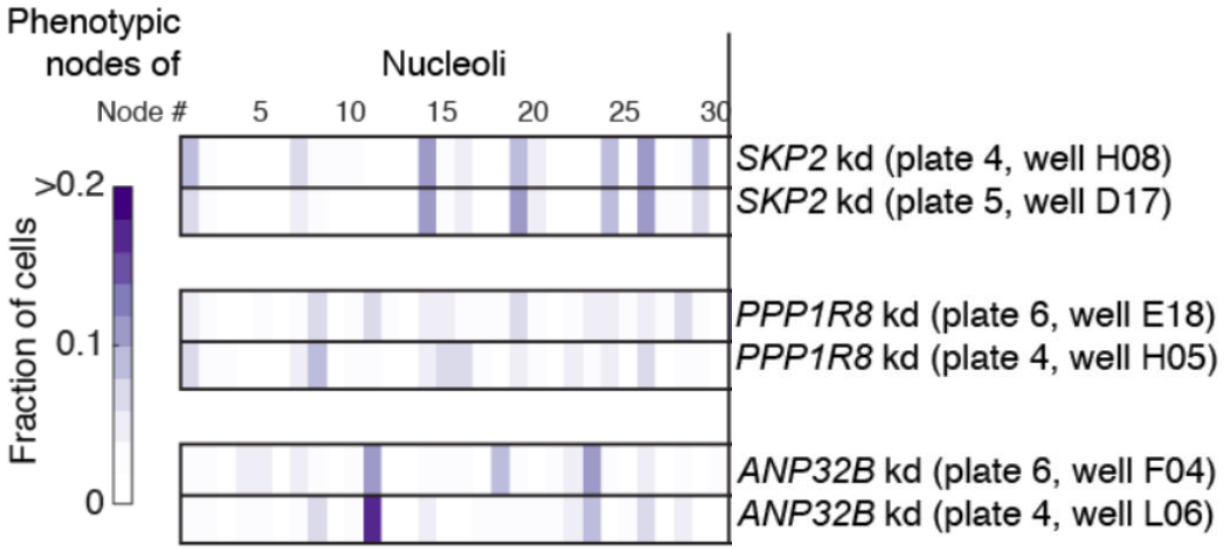
→ 3. Median of each feature in each node

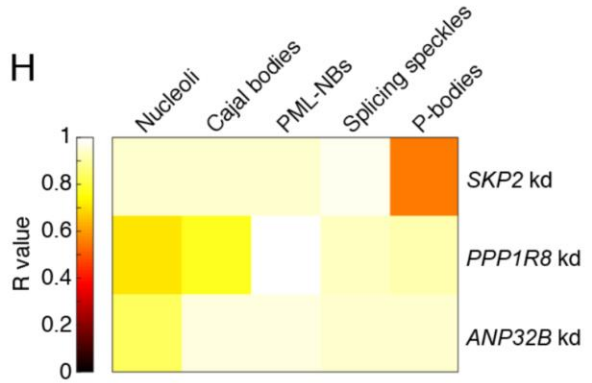
→ 4. Define extreme or average phenotype



Technical reproducibility of analysis?

→ How do biological replicates cluster?



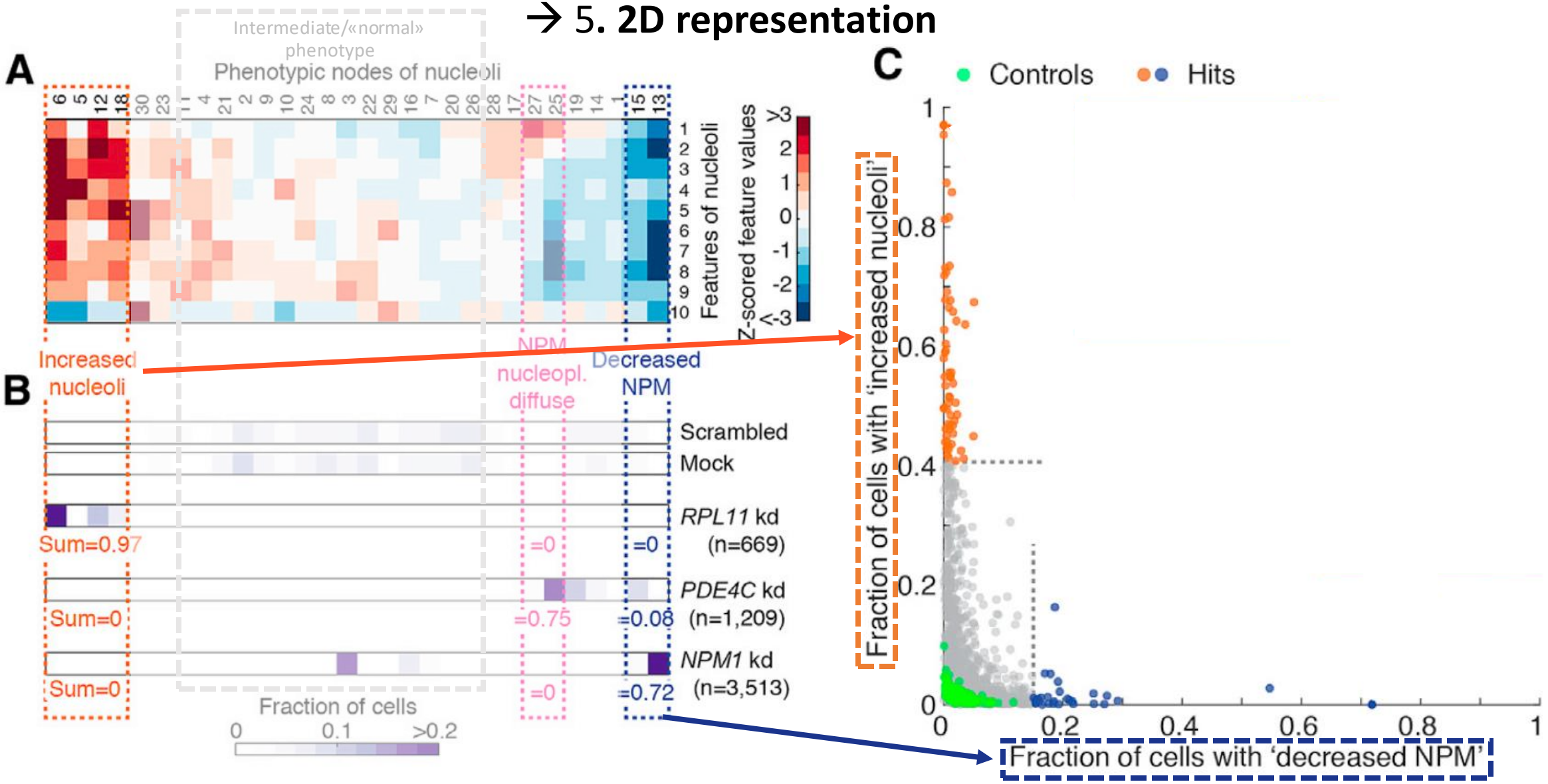


2. SOM Clustering of perturbed and unperturbed cells

→ 3. Median of each feature in each node

→ 4. Define extreme or average phenotype

→ **5. 2D representation**

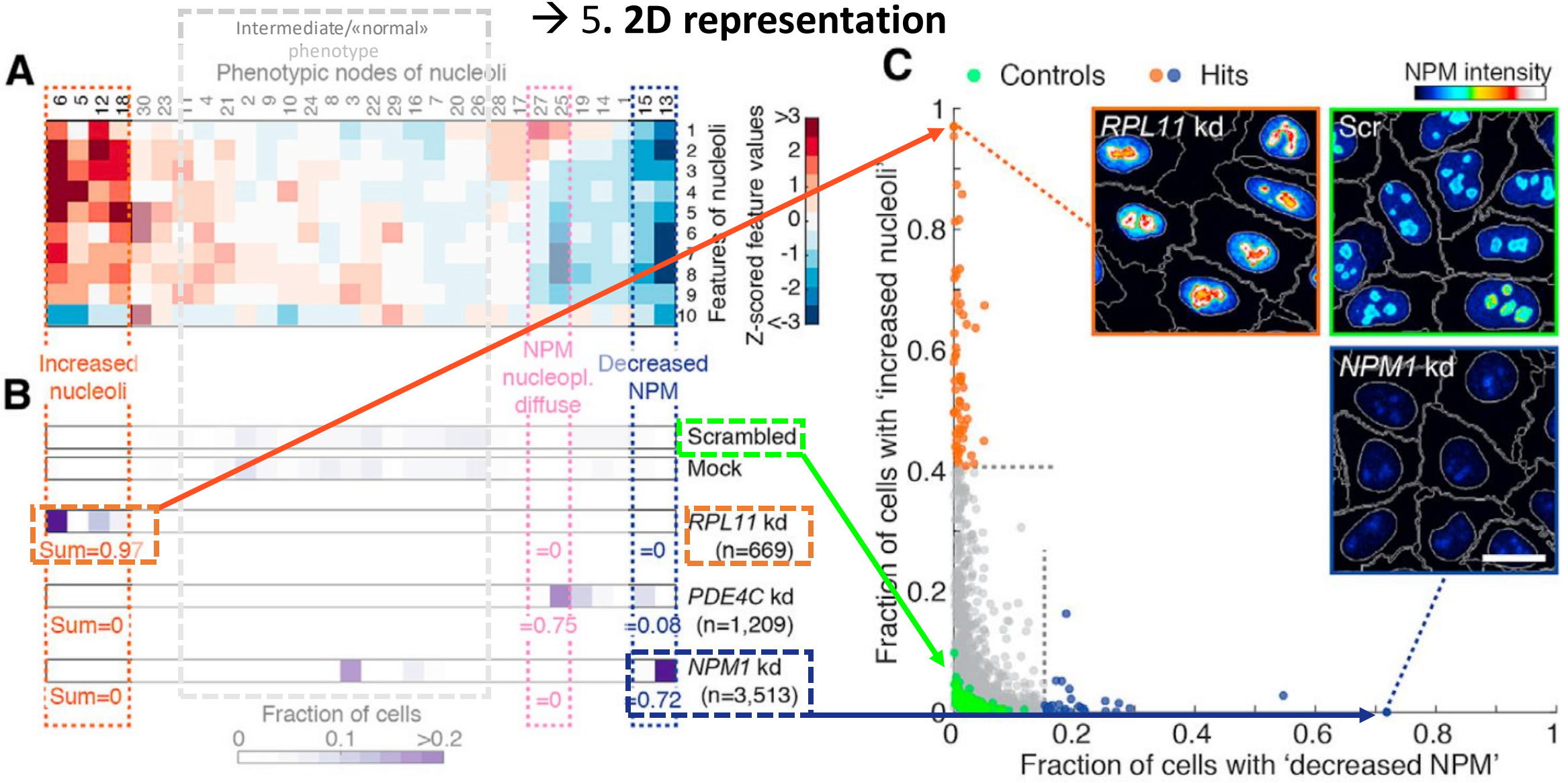


2. SOM Clustering of perturbed and unperturbed cells

→ 3. Median of each feature in each node

→ 4. Define extreme or average phenotype

→ **5. 2D representation**



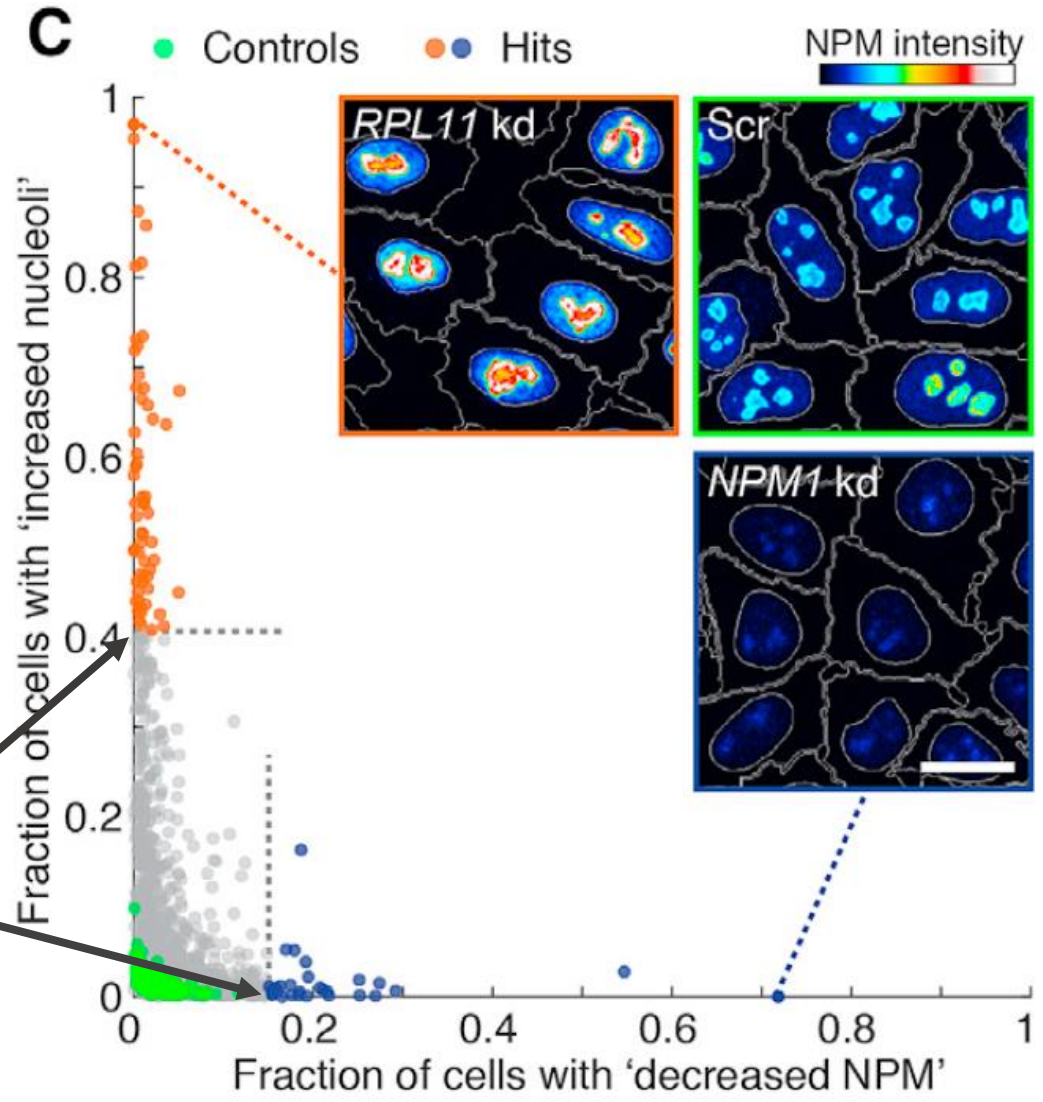
2. SOM Clustering of perturbed and unperturbed cells

→ 3. Median of each feature in each node

→ 4. Define extreme or average phenotype

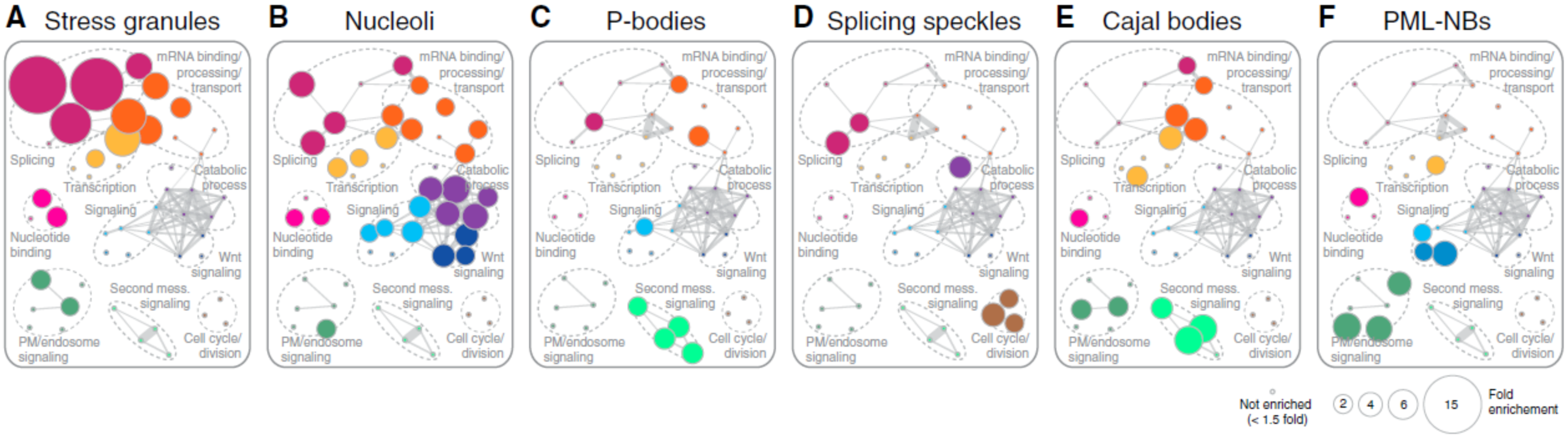
→ 5. 2D representation

→ 6. **Hit Calling**



> 2.5 SD from mean across the whole screen

GO Analysis with Hits



→ Follow up experiments

Molecular Cell
Resource

A Systems-Level Study Reveals Regulators of Membrane-less Organelles in Human Cells

Doris Berchtold,[1] Nico Battich,[1] and Lucas Pelkmans[1,2,*]
[1]Institute of Molecular Life Sciences, University of Zurich, 8057 Zurich, Switzerland
[2]Lead Contact
*Correspondence: lucas.pelkmans@imls.uzh.ch
https://doi.org/10.1016/j.molcel.2018.10.036

- Organelle changes can not always be described by one feature (size, intensity of marker, localization,…)
- Organelles may change physiologically during cell cycle

→ To find influence of genetic perturbation in HTS, one must consider and exclude these factors

Article

Cell

A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging

Leeat Keren,[1] Marc Bosse,[1] Diana Marquez,[1] Roshan Angoshtari,[1] Samir Jain,[1] Sushama Varma,[1] Soo-Ryum Yang,[1] Allison Kurian,[1] David Van Valen,[2,3] Robert West,[1] Sean C. Bendall,[1,*] and Michael Angelo[1,4,*]

[1]Department of Pathology, Stanford University, Stanford CA, 94305, USA
[2]Department of Biology, Caltech, Pasadena, CA 91125, USA
[3]Department of Bioengineering, Caltech, Pasadena, CA 91125, USA
[4]Lead Contact
*Correspondence: bendall@stanford.edu (S.C.B.), mangelo0@stanford.edu (M.A.)
https://doi.org/10.1016/j.cell.2018.08.039

- Multiplexing 36 marker proteins
- Retrospective TNBC cohort
- Identifying patterns linked to survival outcome

Article

Cell

A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging

Leeat Keren,[1] Marc Bosse,[1] Diana Marquez,[1] Roshan Angoshtari,[1] Samir Jain,[1] Sushama Varma,[1] Soo-Ryum Yang,[1] Allison Kurian,[1] David Van Valen,[2,3] Robert West,[1] Sean C. Bendall,[1,*] and Michael Angelo[1,4,*]

[1]Department of Pathology, Stanford University, Stanford CA, 94305, USA
[2]Department of Biology, Caltech, Pasadena, CA 91125, USA
[3]Department of Bioengineering, Caltech, Pasadena, CA 91125, USA
[4]Lead Contact
*Correspondence: bendall@stanford.edu (S.C.B.), mangelo0@stanford.edu (M.A.)
https://doi.org/10.1016/j.cell.2018.08.039

- TNBC = lack of therapeutic targets:
  - Estrogen receptor
  - Progesterone receptor
  - Her2

→ Radiation & chemotherapeutic neoadjuvant therapy

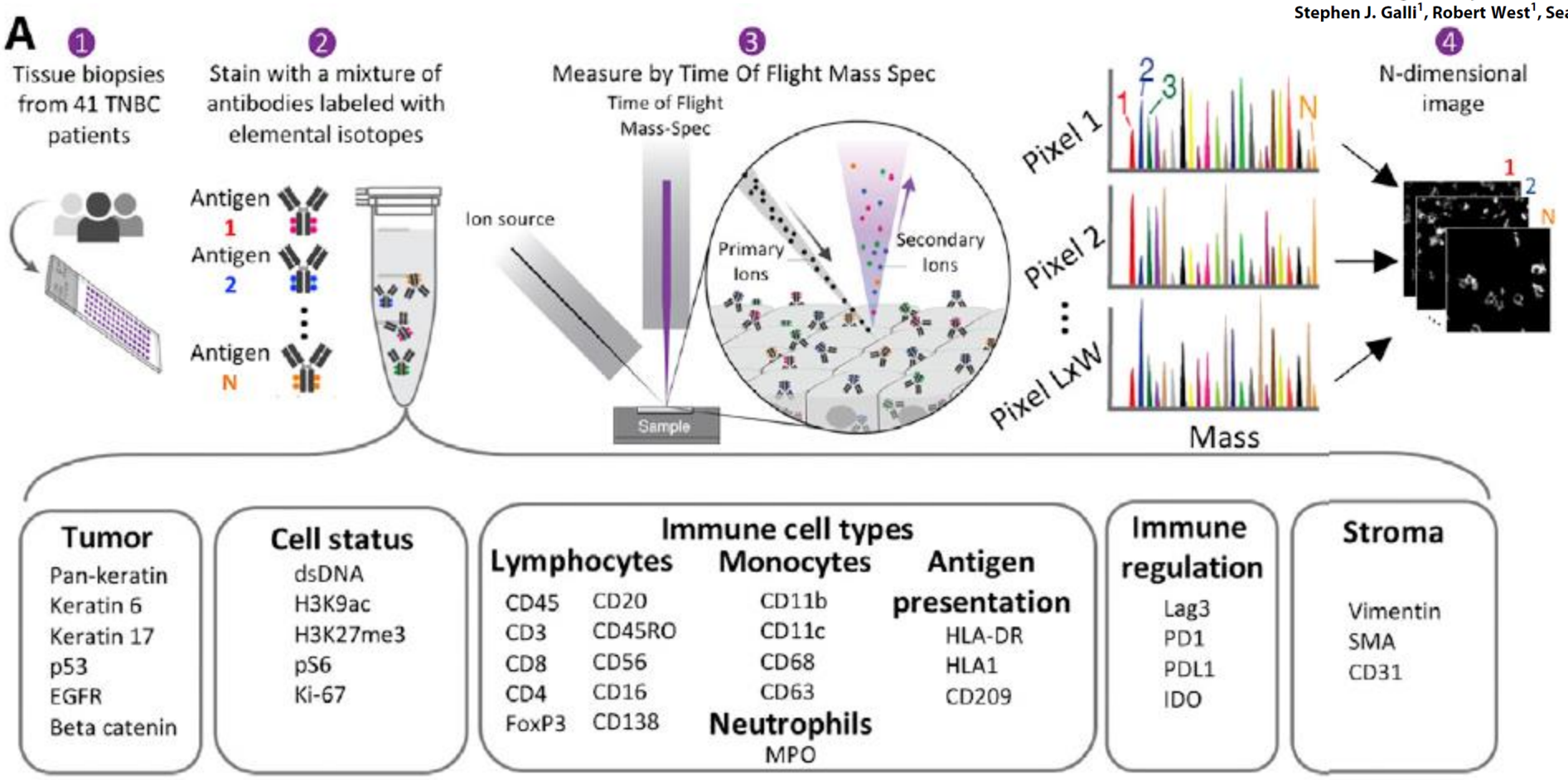
MIBI TOF Multiplexed imaging

SCIENCE ADVANCES | RESEARCH ARTICLE

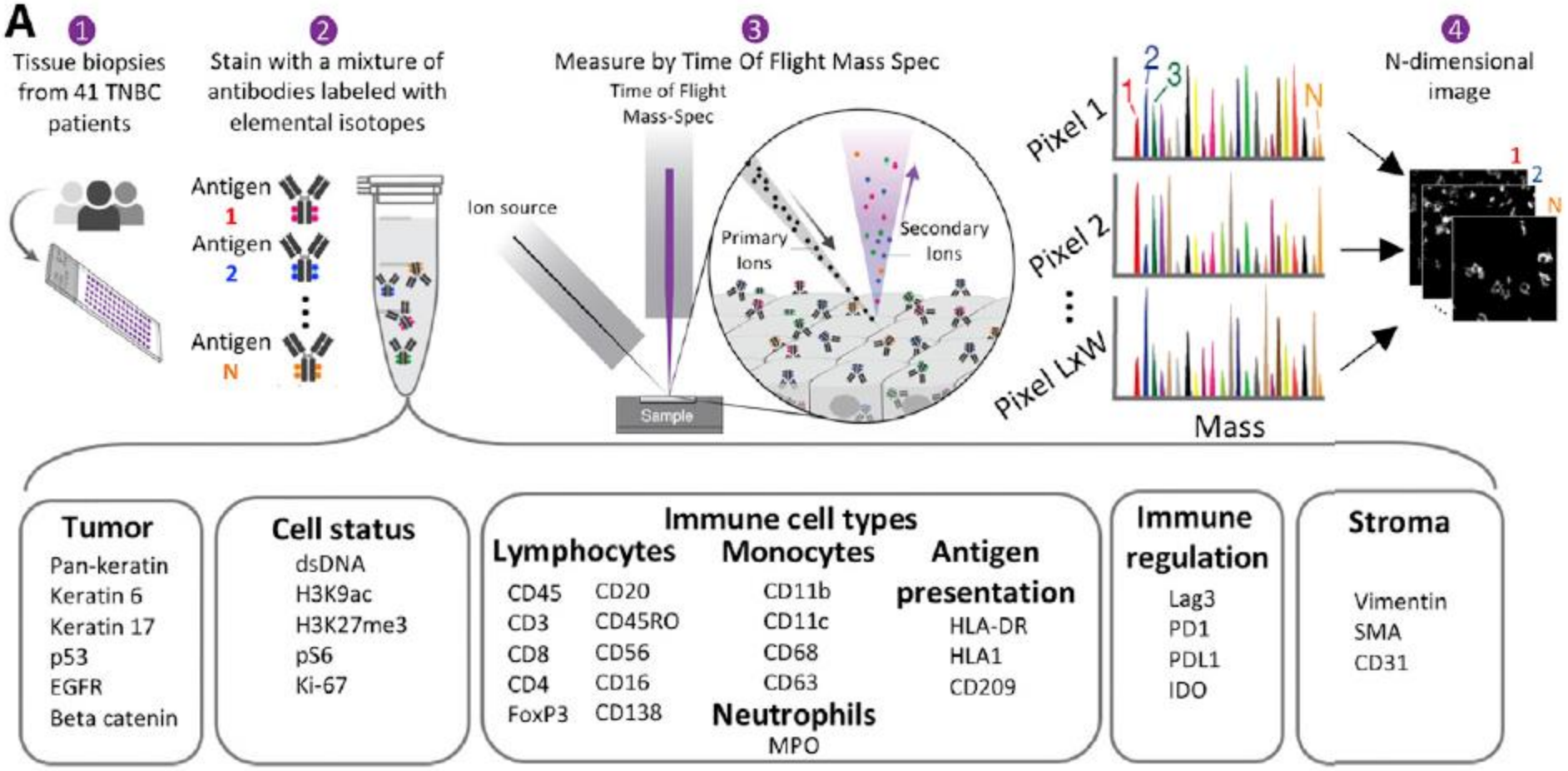
CANCER

MIBI-TOF: A multiplexed imaging platform relates cellular phenotypes and tissue structure

Leeat Keren[1]*, Marc Bosse[1]*, Steve Thompson[1], Tyler Risom[1], Kausalia Vijayaragavan[1], Erin McCaffrey[1,2], Diana Marquez[1], Roshan Angoshtari[1], Noah F. Greenwald[1,3], Harris Fienberg[1], Jennifer Wang[1], Neeraja Kambham[1], David Kirkwood[1], Garry Nolan[4], Thomas J. Montine[1], Stephen J. Galli[1], Robert West[1], Sean C. Bendall[1]*, Michael Angelo[1]*†



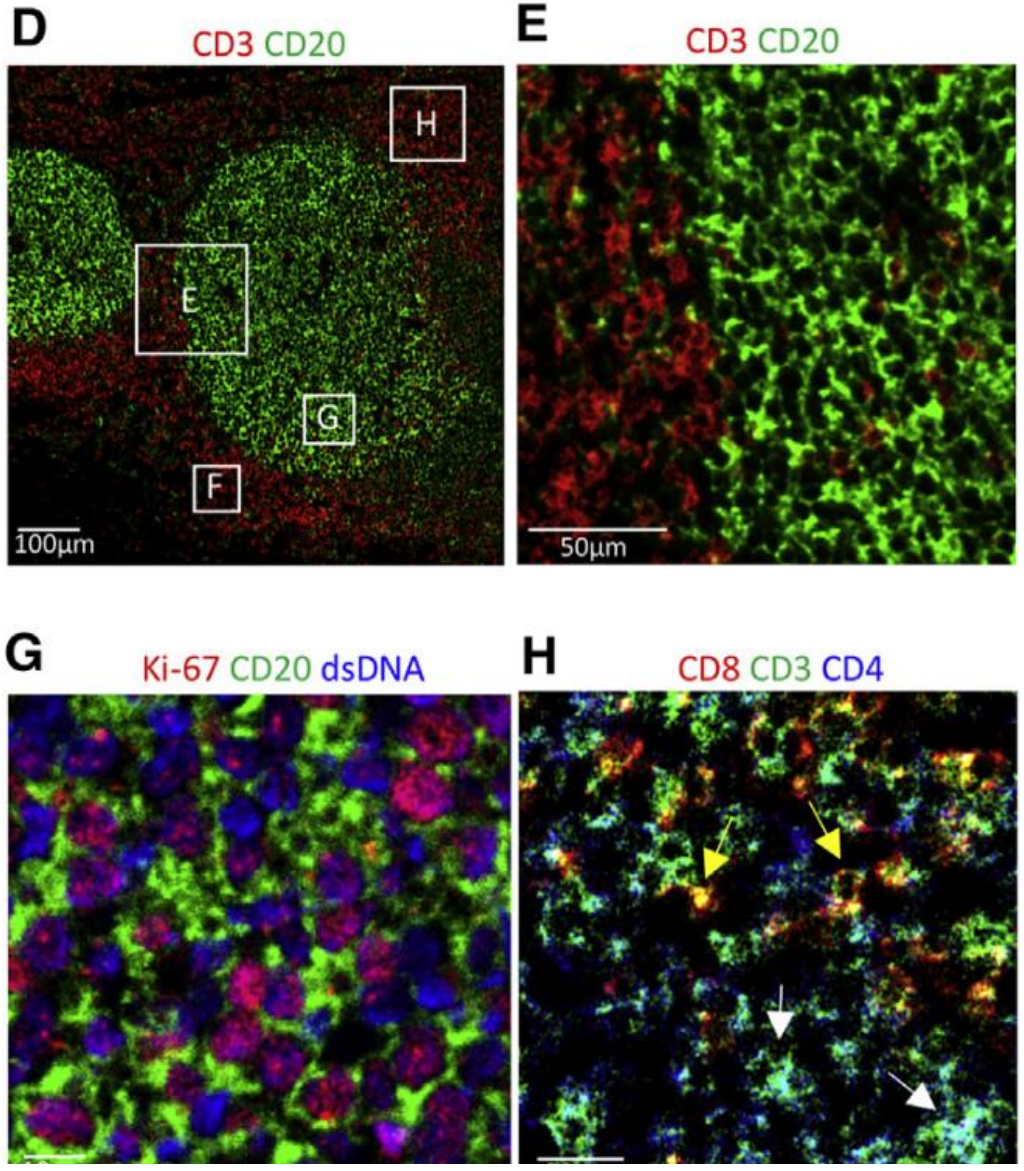
MIBI TOF Multiplexed imaging



Quality Controls:

- match w/ conventional IHC
- tissue specific localization
- subcellular localization
- co-expression of markers

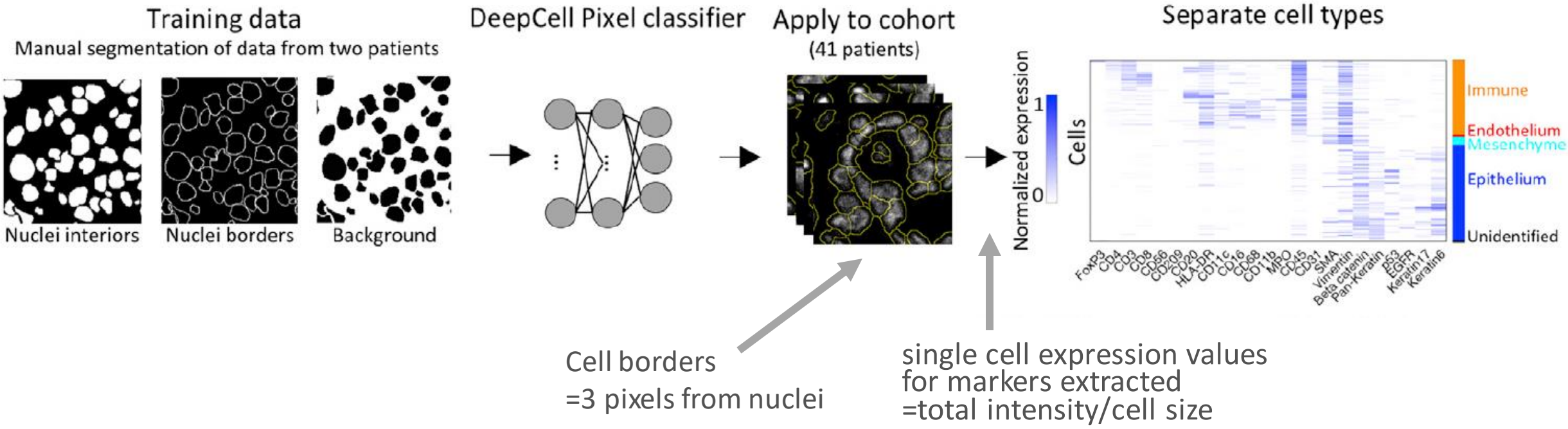
MIBI TOF Multiplexed imaging



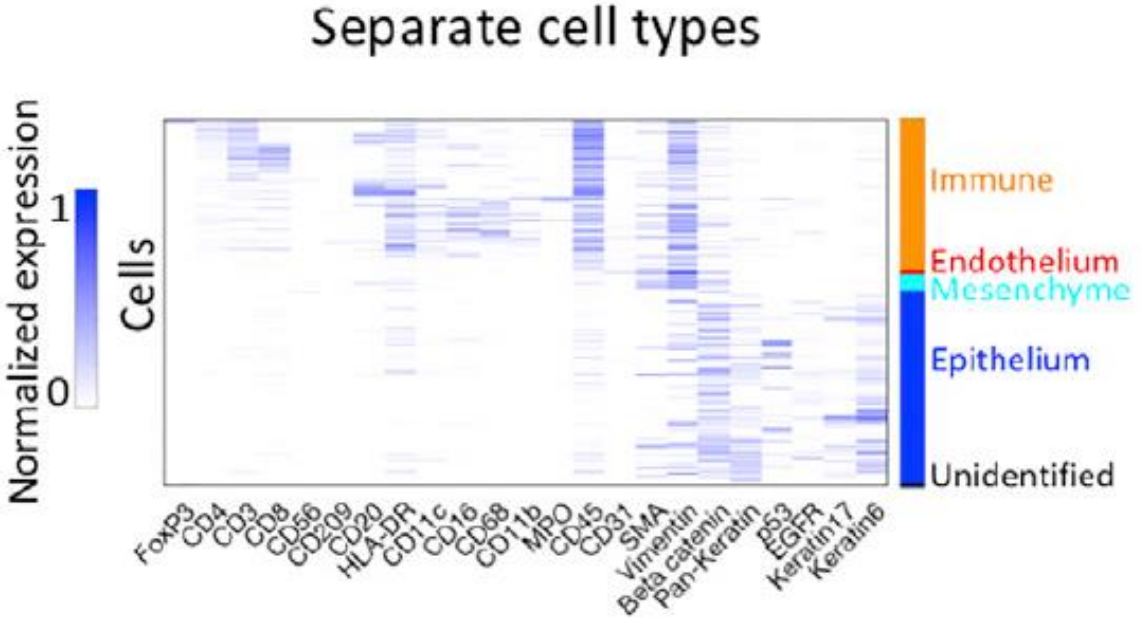
Quality Controls in e.g. tonsil:

- match w/ conventional IHC
- tissue specific localization
- subcellular localization
- co-expression of markers

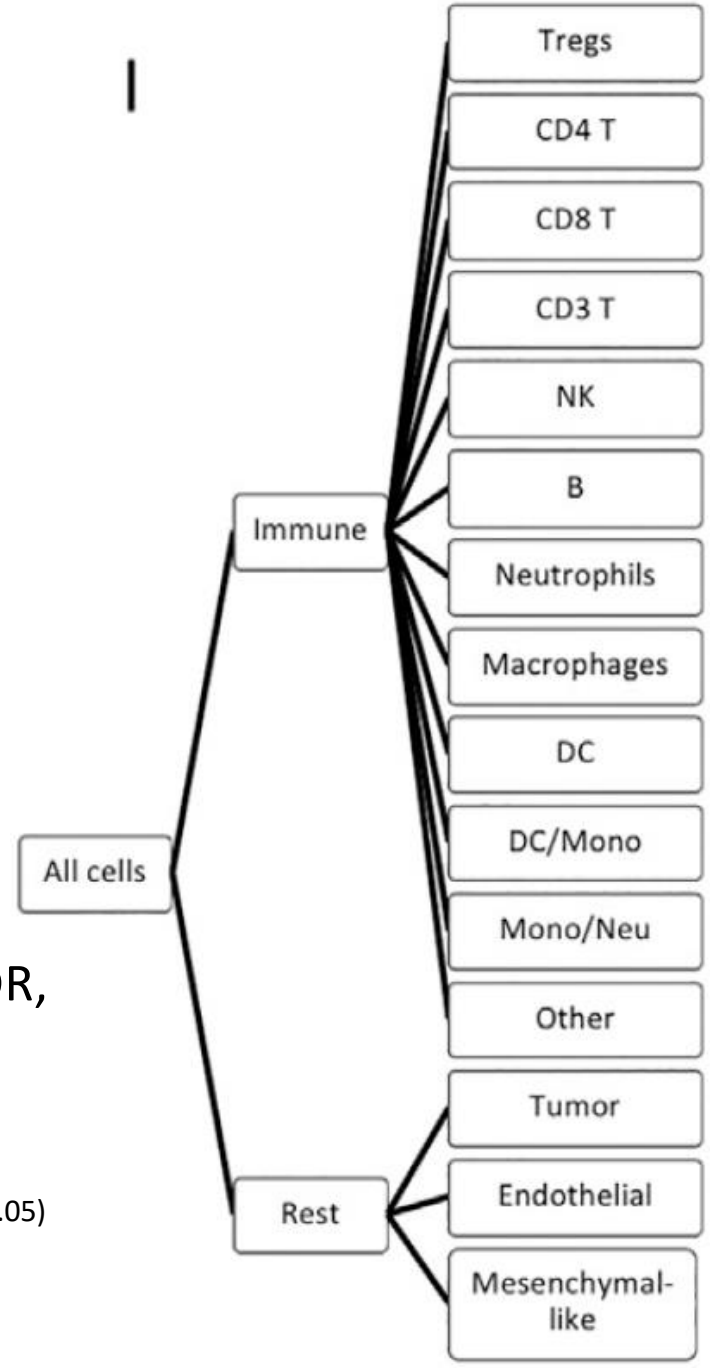
Automated Image Analysis



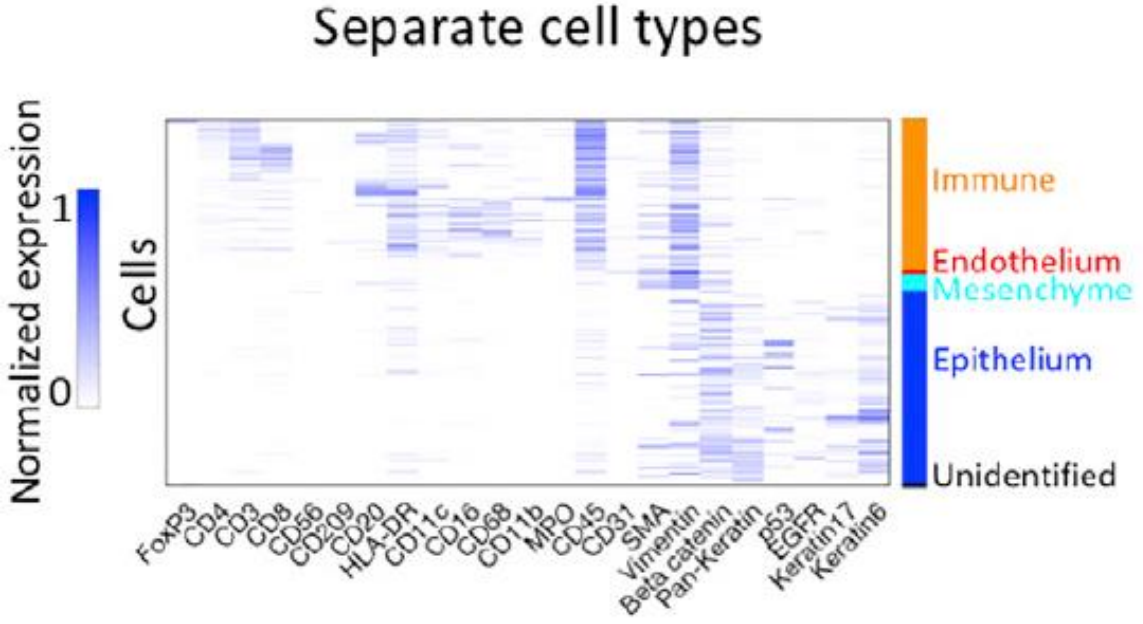
Automated Image Analysis

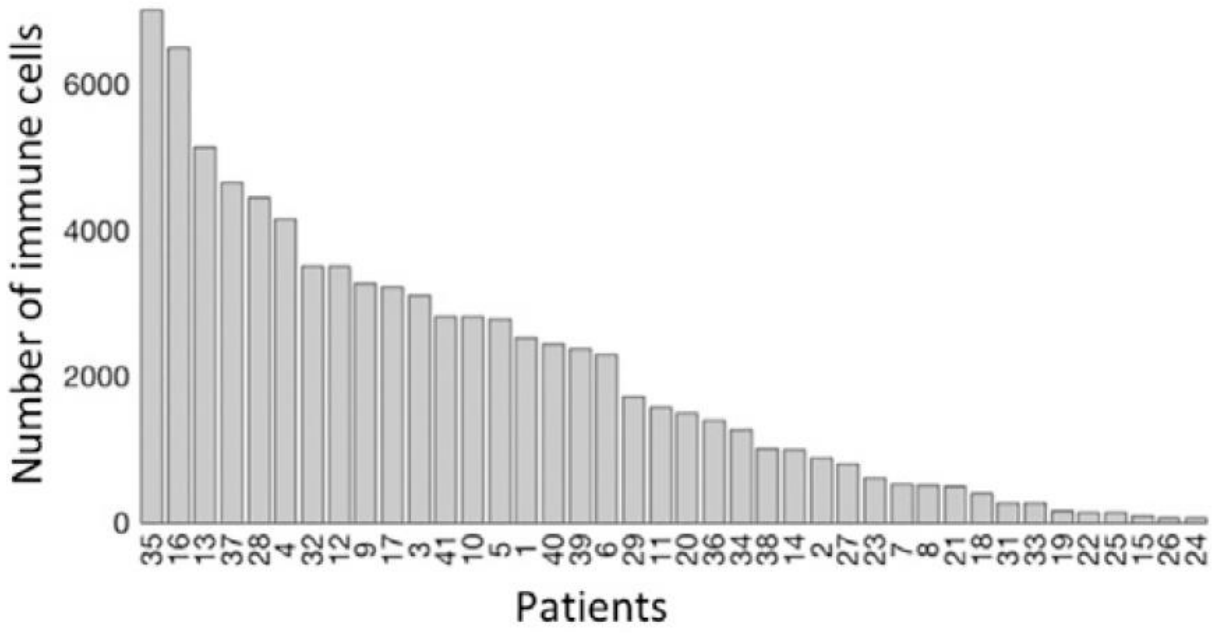


1. Clustering: Immune vs non-immune
   - 16 markers: CD45, FoxP3, CD4, CD8, CD3, CD20, CD16, CD68, MPO, HLA-DR, Pan-Keratin, Keratin17, Keratin6, p53, Beta catenin, EGFR
   - FlowSOM: 100 clusters
   - Merging of similar clusters by hierarchial clustering (cosine distance < 0.05)



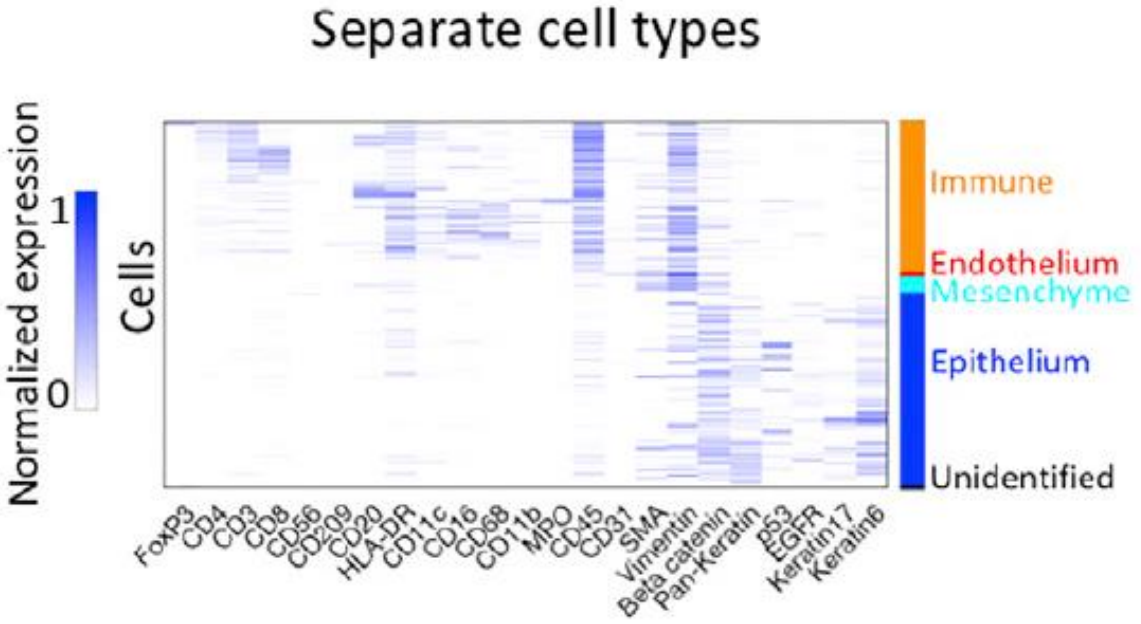
Automated Image Analysis





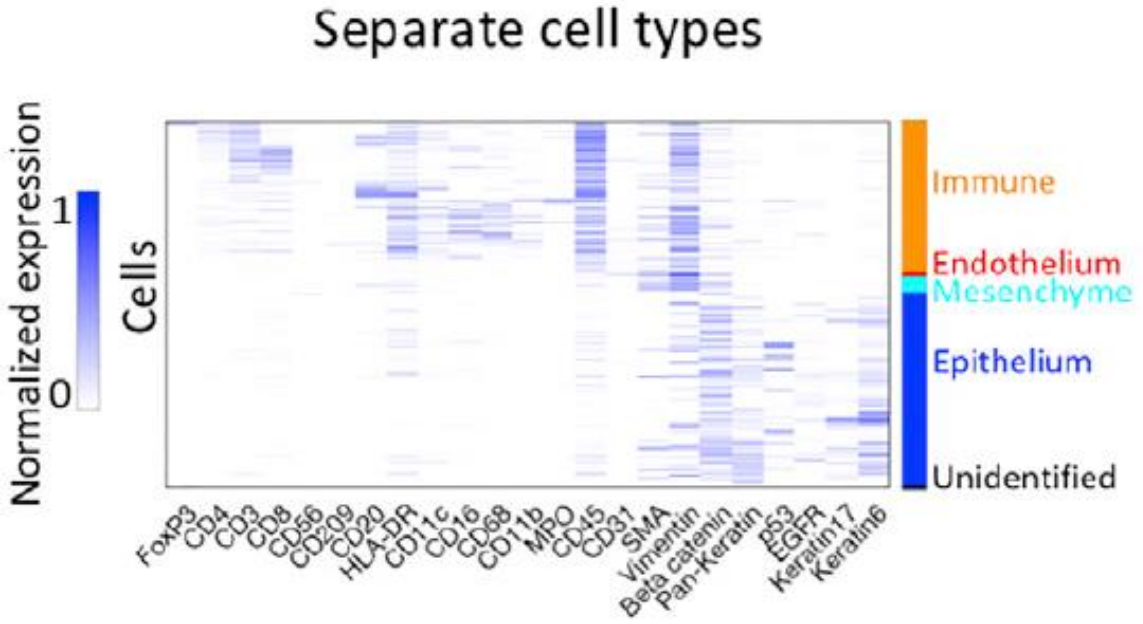
1. Clustering: Immune vs non-immune
   - 16 markers: CD45, FoxP3, CD4, CD8, CD3, CD20, CD16, CD68, MPO, HLA-DR, Pan-Keratin, Keratin17, Keratin6, p53, Beta catenin, EGFR
   - FlowSOM: 100 clusters
   - Merging of similar clusters by hierarchial clustering (cosine distance < 0.05)

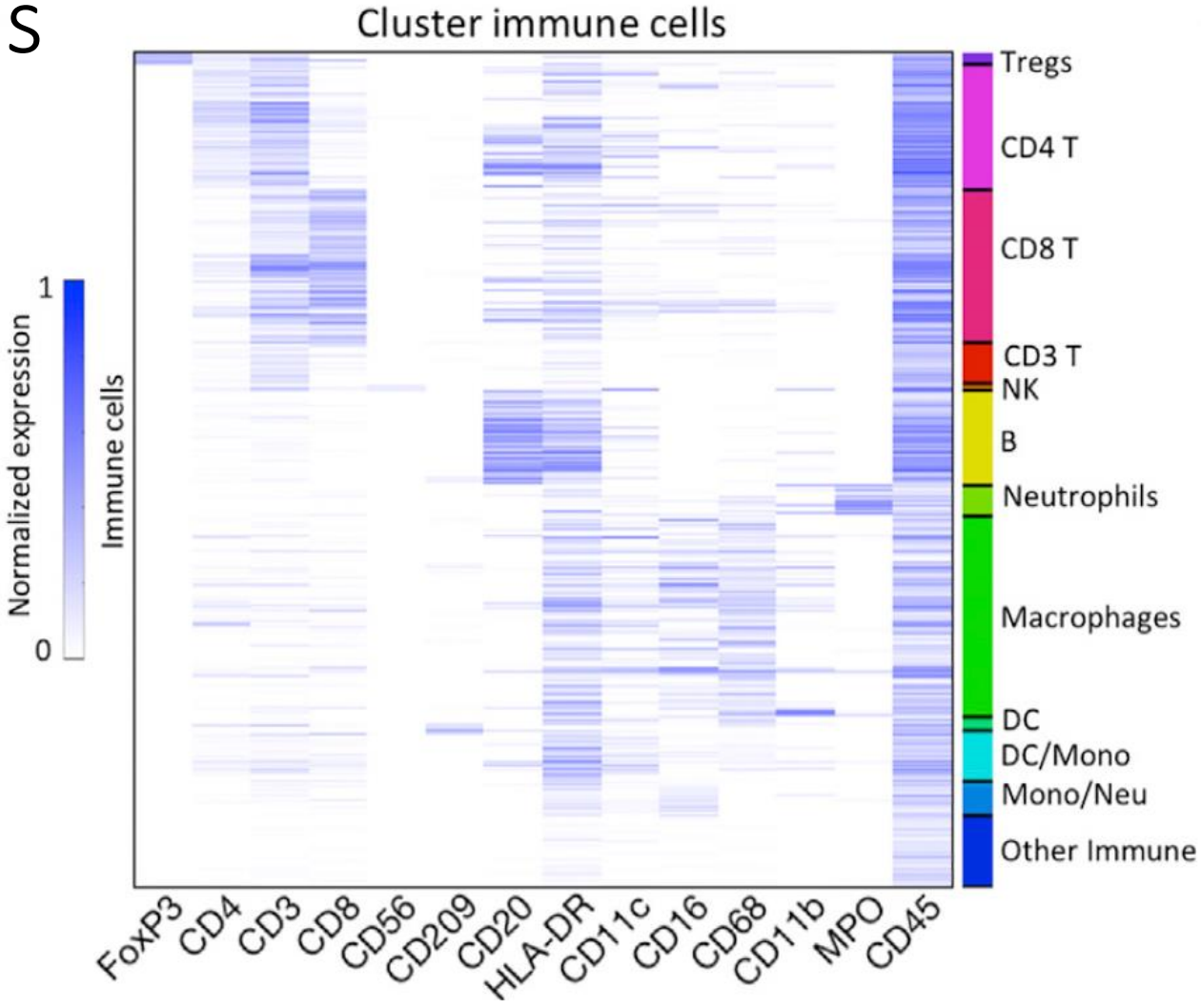
Automated Image Analysis



2. Clustering: Non-immune cells into Epithelial, Mesenchymal, Endothelial and Unidentified
   - 8 markers: Vimentin, SMA, CD31, Beta-catenin, EGFR, Keratin 17, Keratin 6, Pan-keratin
   - FlowSOM: 100 clusters
   - Merging of similar clusters by hierarchial clustering (cosine distance < 0.05)

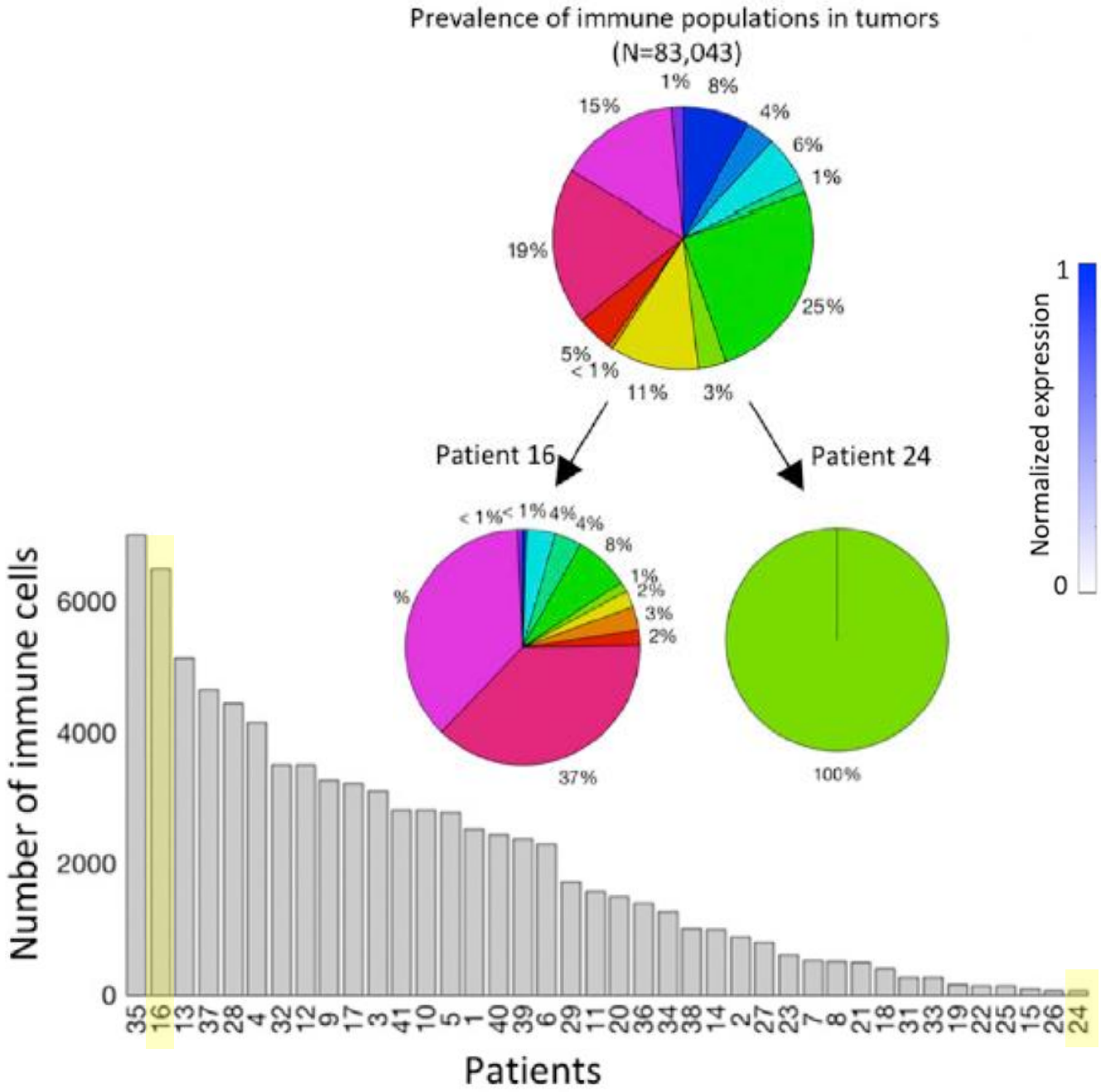
Automated Image Analysis

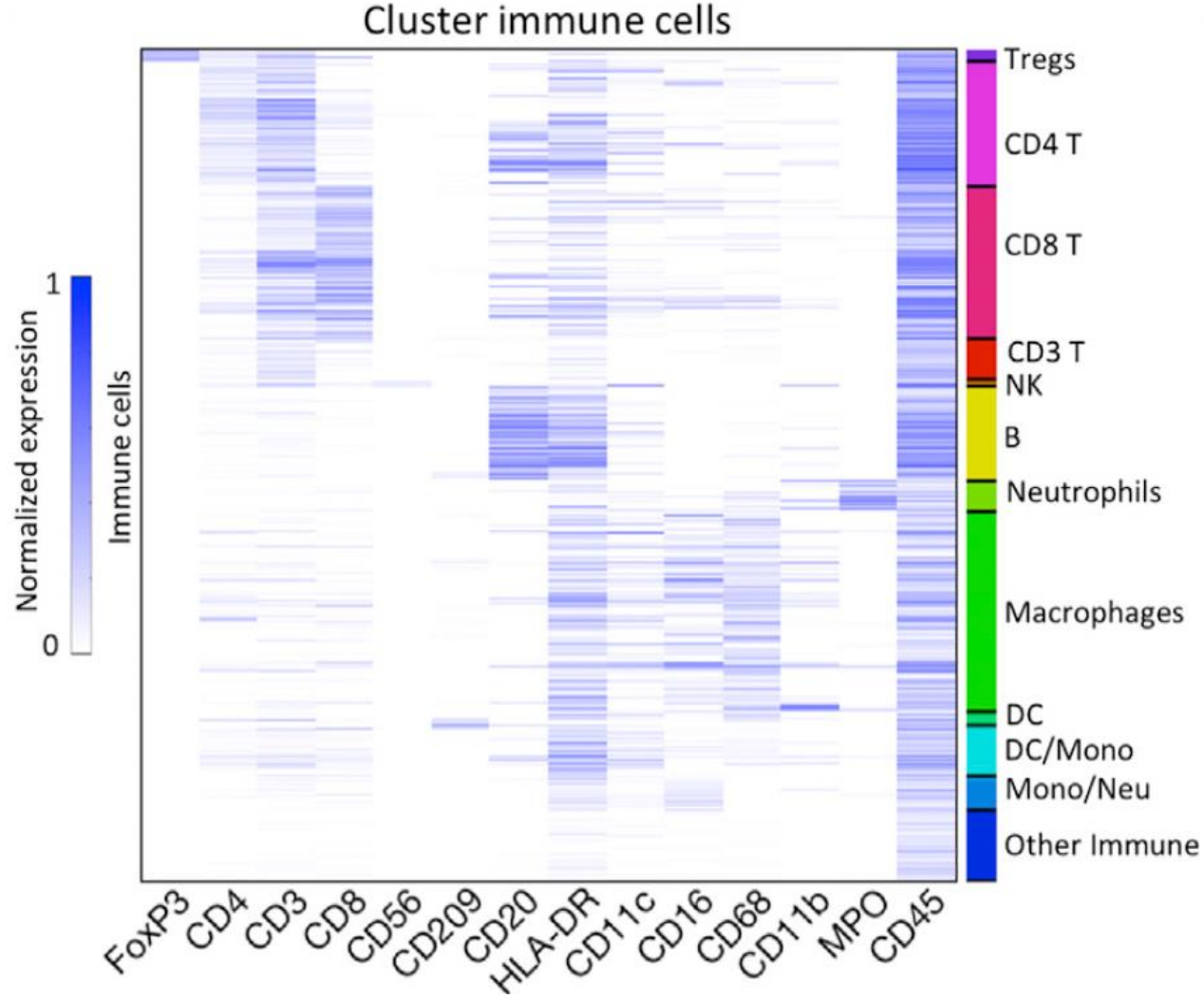


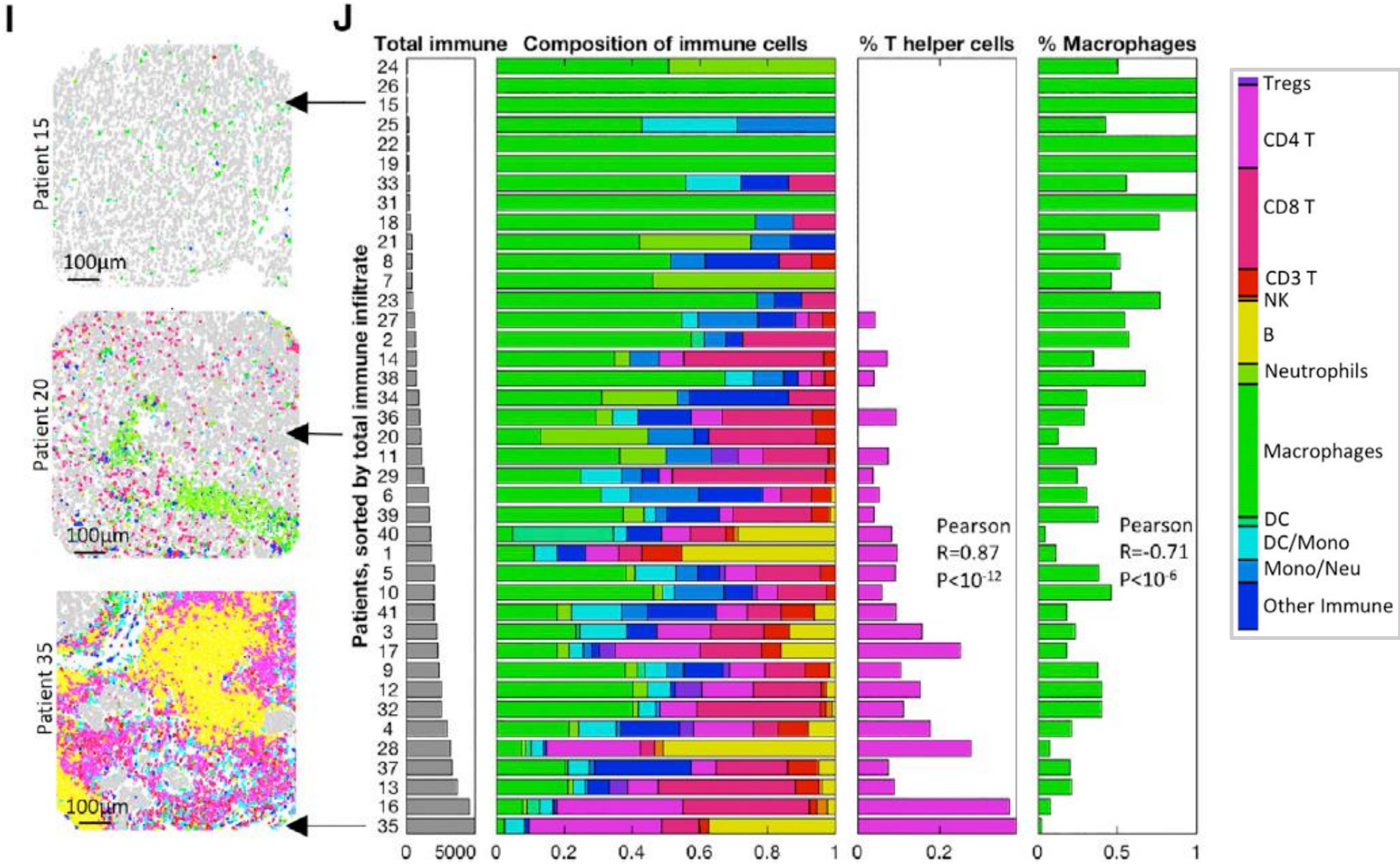


3. Clustering: Immune cells into 12 groups

- FlowSOM: 100 clusters
- Merging of similar clusters by hierarchial clustering (cosine distance < 0.05)

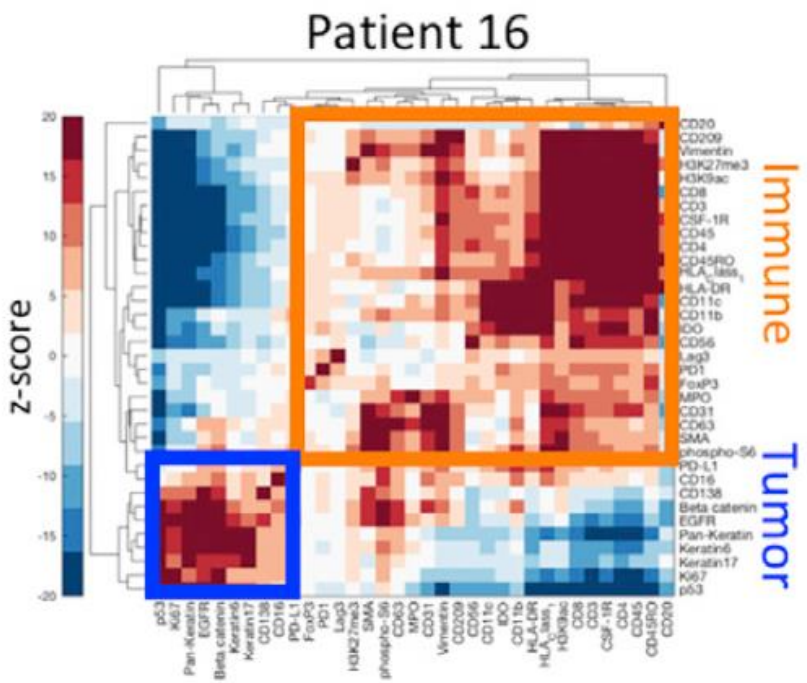

Prevalence of immune populations in tumors
(N=83,043)
1%
8%
4%
6%
1%
25%
3%
11%
5%
< 1%
19%
15%
Patient 16
Patient 24
< 1%
< 1%
4%
4%
8%
1%
2%
3%
2%
37%
100%
Number of immune cells
6000
4000
2000
0
35 16 13 37 28 4 32 12 9 17 3 41 10 5 1 40 39 6 29 11 20 36 34 38 14 2 27 23 7 8 21 18 31 33 19 22 25 15 26 24
Patients

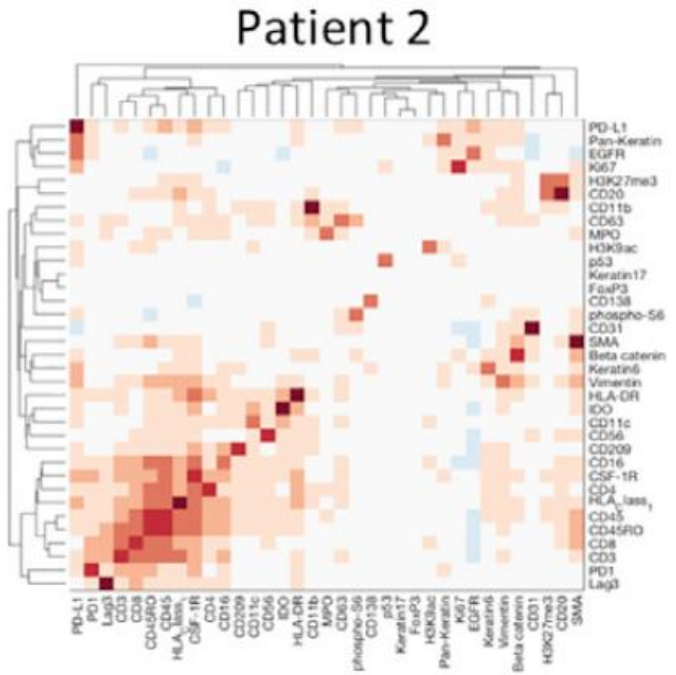

Cluster immune cells
Normalized expression
1
0
Immune cells
FoxP3
CD4
CD3
CD8
CD56
CD209
CD20
HLA-DR
CD11c
CD16
CD68
CD11b
MPO
CD45
Tregs
CD4 T
CD8 T
CD3 T
NK
B
Neutrophils
Macrophages
DC
DC/Mono
Mono/Neu
Other Immune


I
J
Patient 15
Patient 20
Patient 35
100µm
100µm
100µm
Total immune
Composition of immune cells
% T helper cells
% Macrophages
Patients, sorted by total immune infiltrate
24
26
15
25
22
19
33
31
18
21
8
7
23
27
2
14
38
34
36
20
11
29
6
39
40
1
5
10
41
3
17
9
12
32
4
28
37
13
16
35
0
5000
0
0.2
0.4
0.6
0.8
1
0
0.2
0
0.5
1
Pearson
R=0.87
P<10-12
Pearson
R=-0.71
P<10-6
Tregs
CD4 T
CD8 T
CD3 T
NK
B
Neutrophils
Macrophages
DC
DC/Mono
Mono/Neu
Other Immune

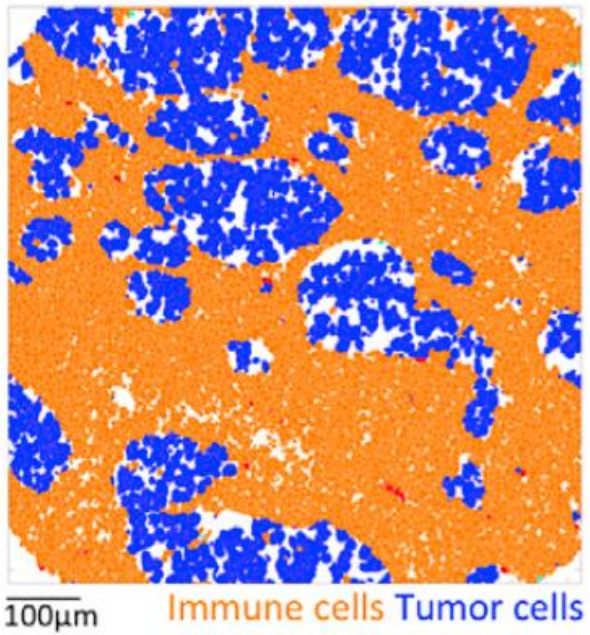
Immune markers are close to each other

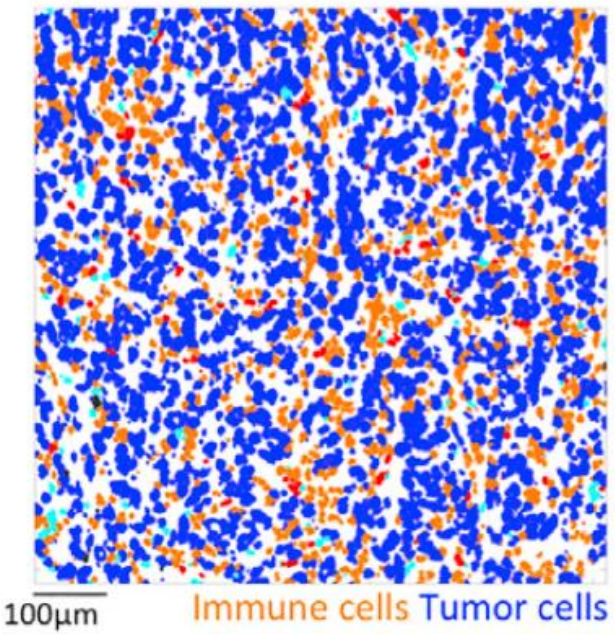
Tumor markers are close to each other



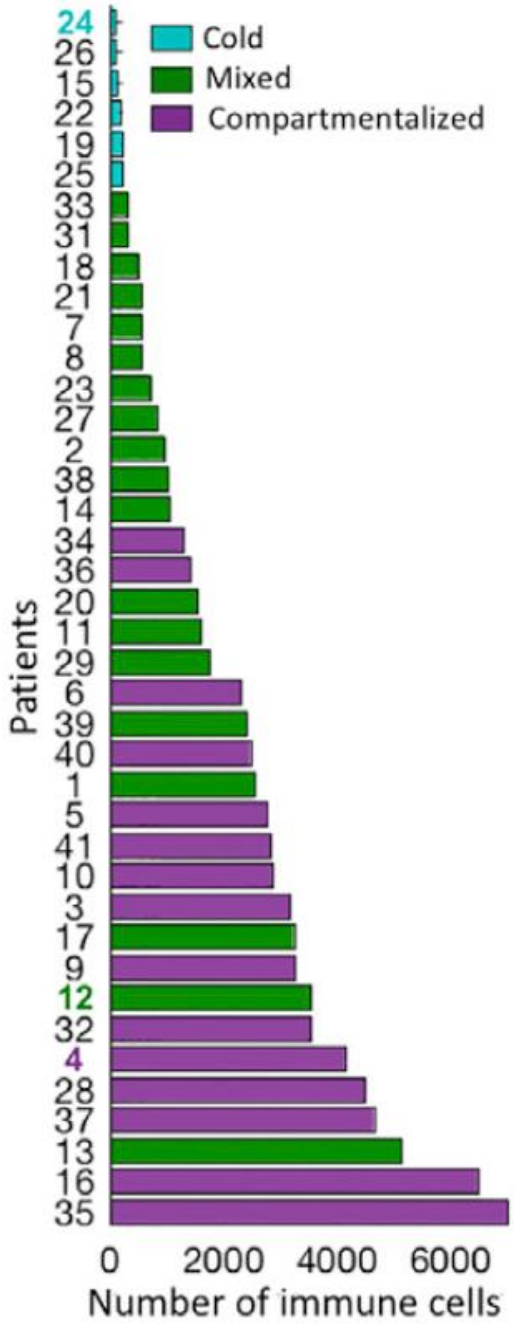


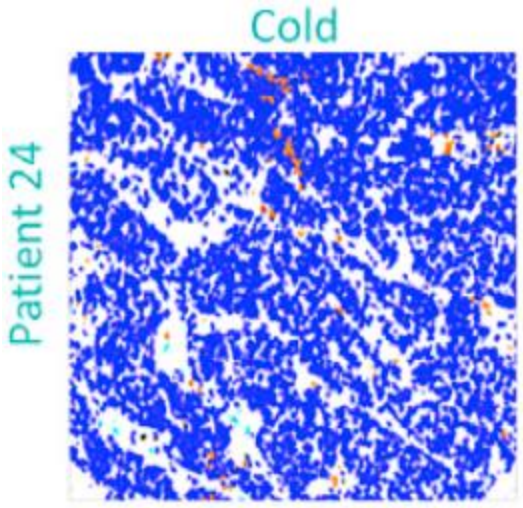
All mixed up

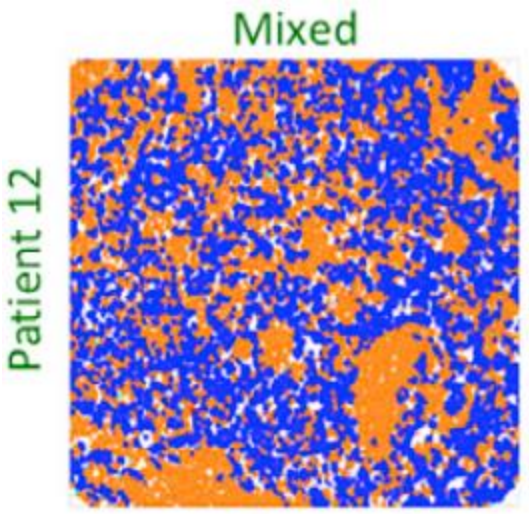


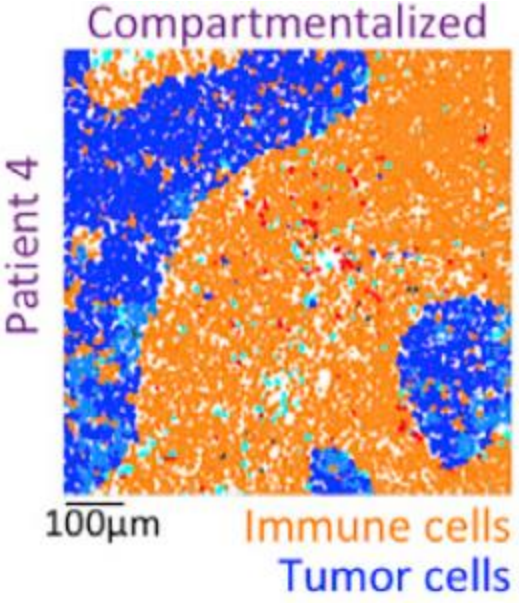


Definition of three «archetypical subtypes»

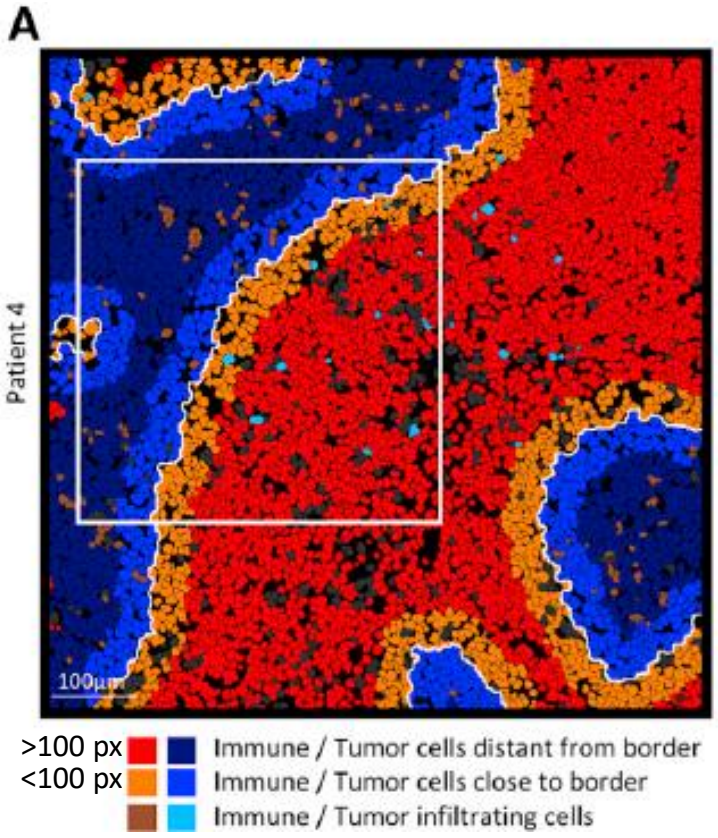


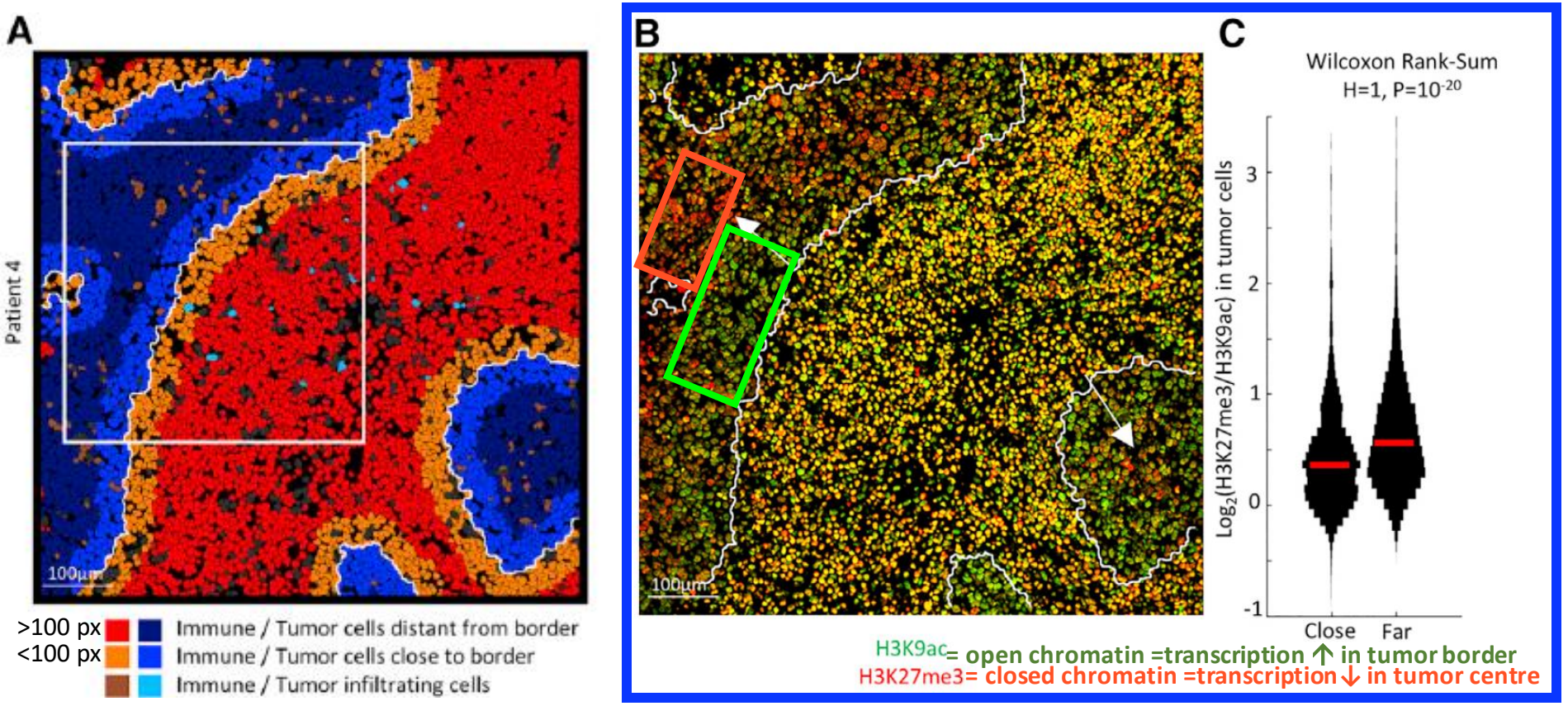






Compartmentalized Tumors

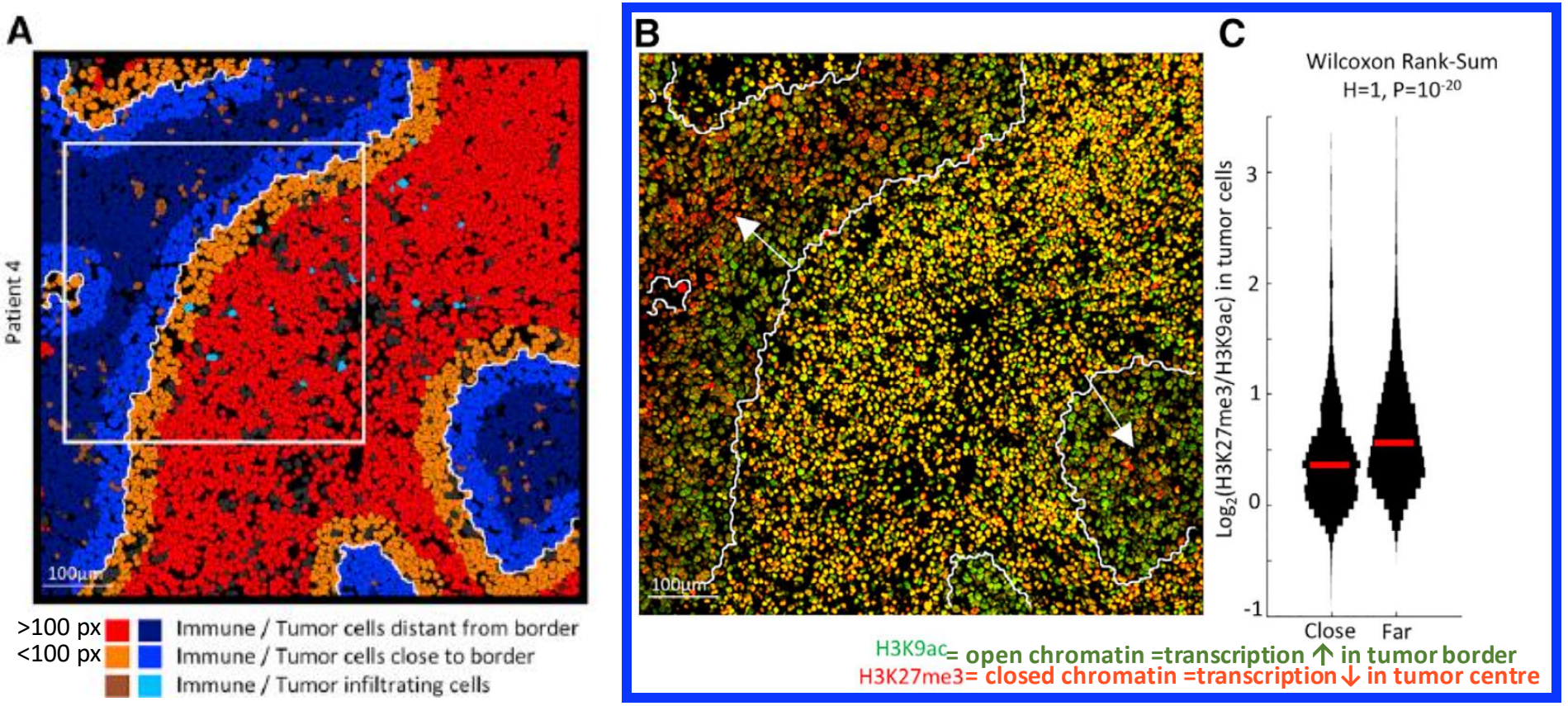


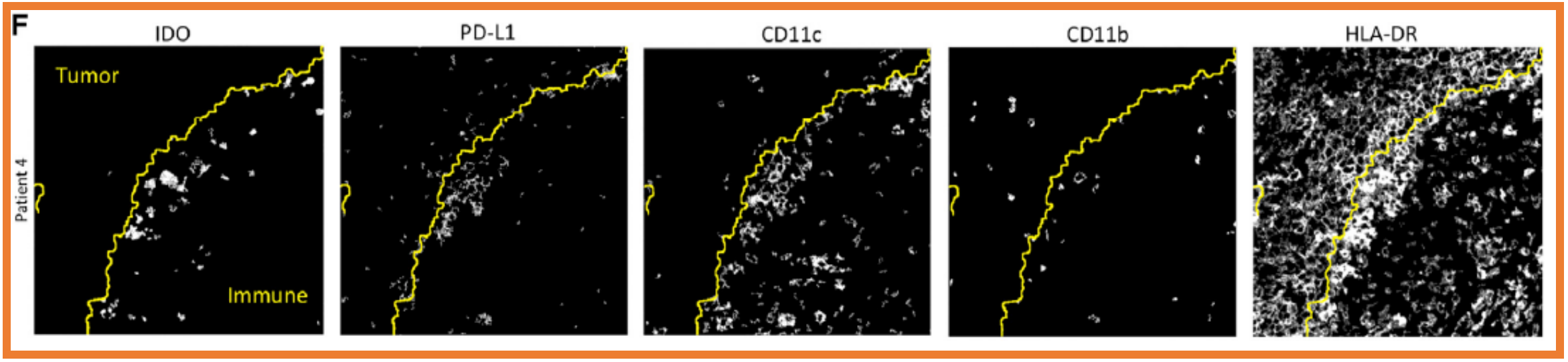


Compartmentalized Tumors

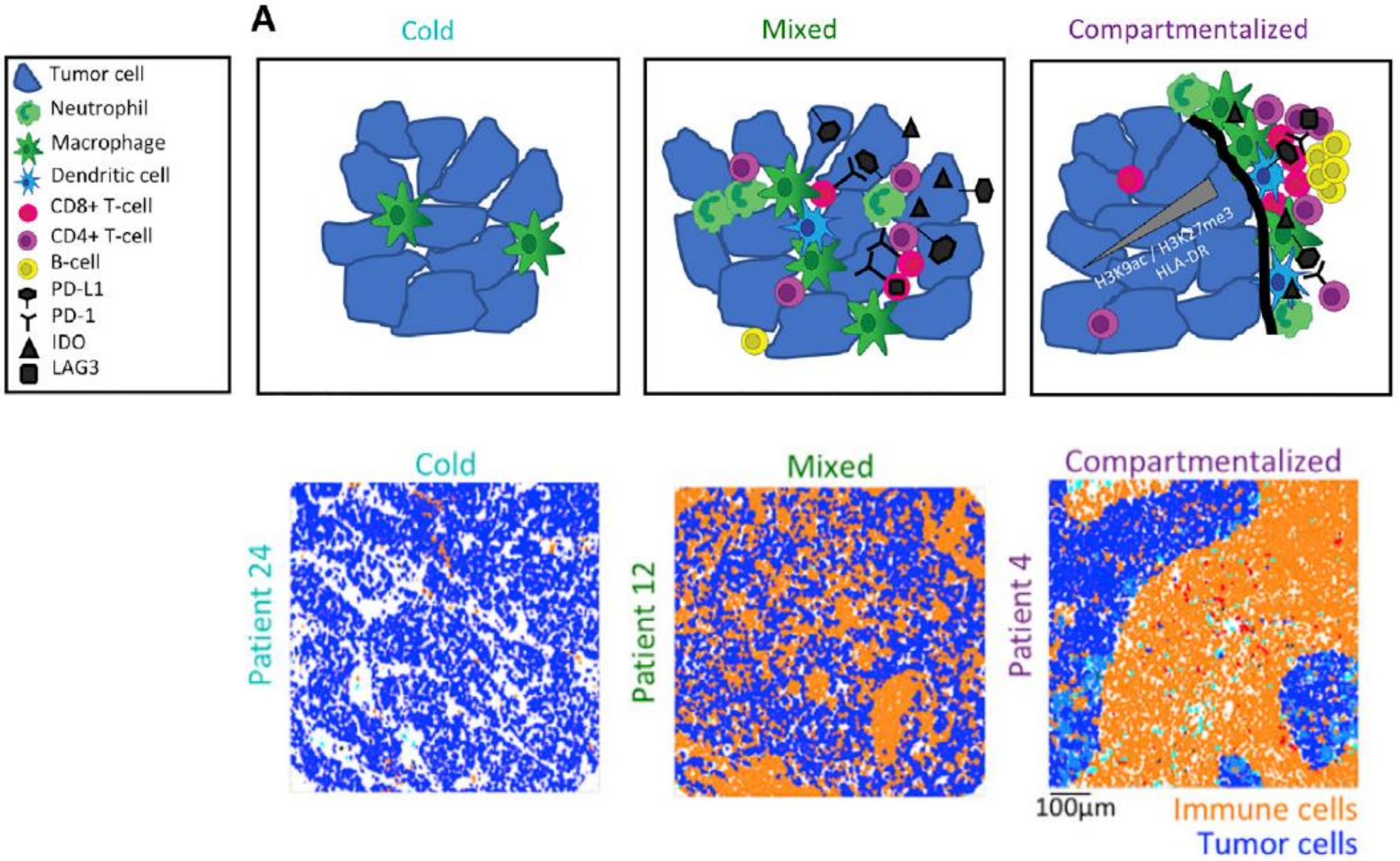
Compartmentalized Tumors

→ Tumor immune border as unique site of immune inhibition with altered expression profiles by both tumor and immune cells

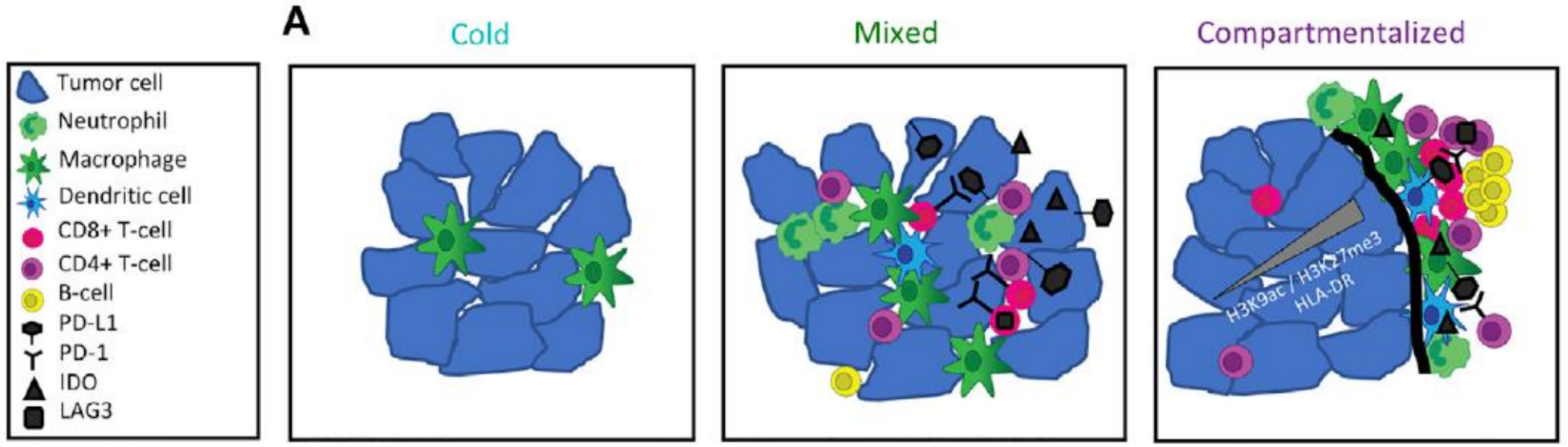


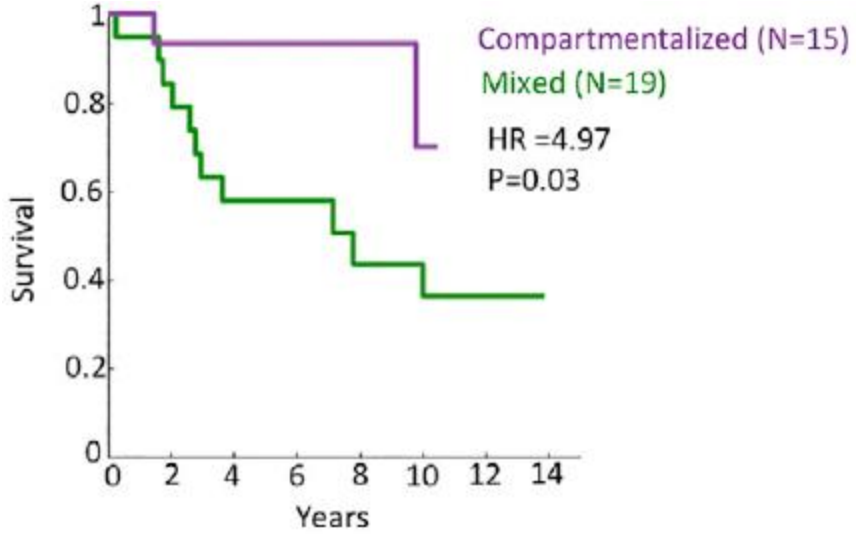


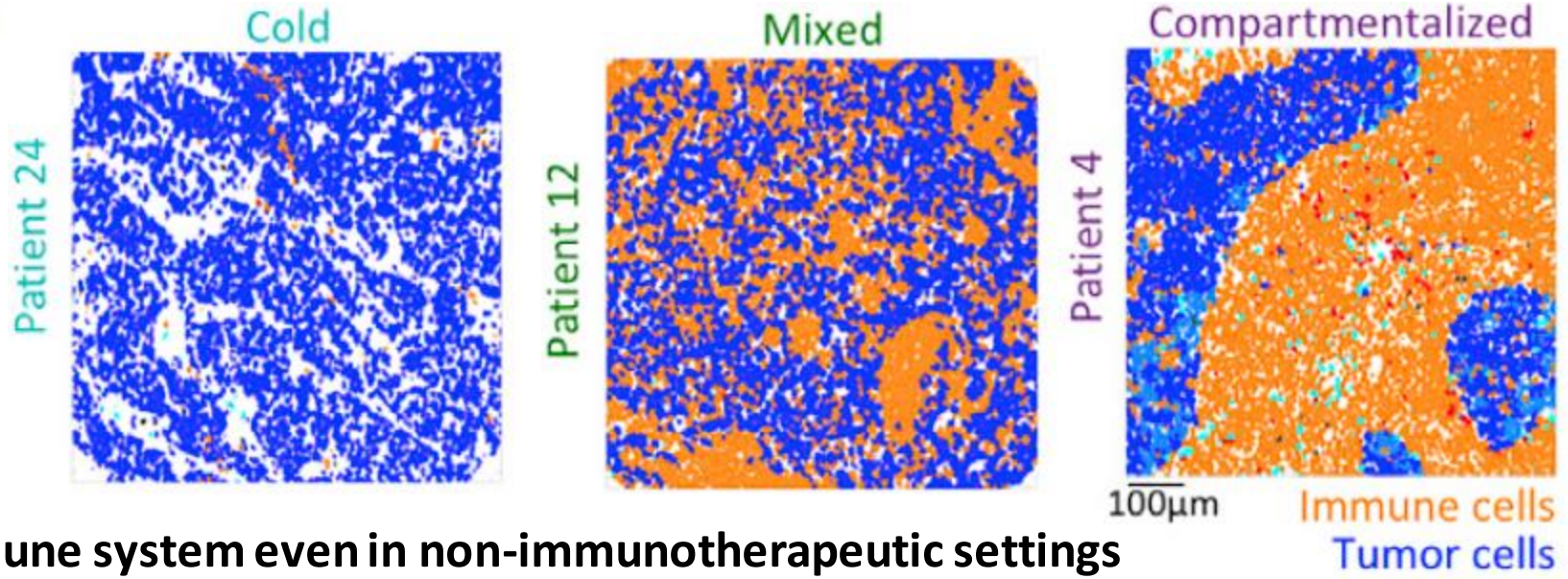
Definition & characterization of three «archetypical subtypes»



Definition & characterization of three «archetypical subtypes»







→ **clinical importance of the immune system even in non-immunotherapeutic settings**

Article

Cell

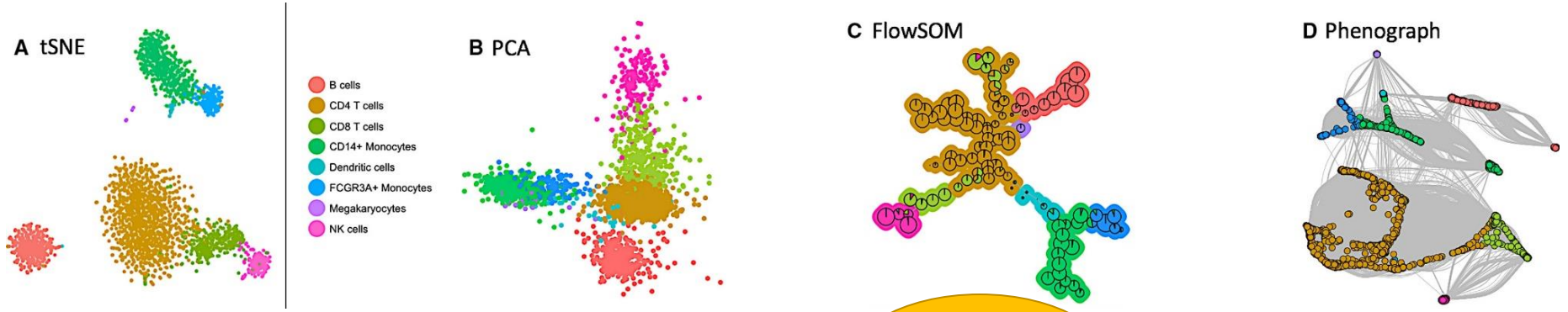
A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging

Leeat Keren,[1] Marc Bosse,[1] Diana Marquez,[1] Roshan Angoshtari,[1] Samir Jain,[1] Sushama Varma,[1] Soo-Ryum Yang,[1] Allison Kurian,[1] David Van Valen,[2,3] Robert West,[1] Sean C. Bendall,[1,*] and Michael Angelo[1,4,*]

[1]Department of Pathology, Stanford University, Stanford CA, 94305, USA
[2]Department of Biology, Caltech, Pasadena, CA 91125, USA
[3]Department of Bioengineering, Caltech, Pasadena, CA 91125, USA
[4]Lead Contact
*Correspondence: bendall@stanford.edu (S.C.B.), mangelo0@stanford.edu (M.A.)
https://doi.org/10.1016/j.cell.2018.08.039

- Automated image analysis on 36 markers
- Identification of 20 cell types
- Evaluation of their spatial distribution

→ Identification of prognosis-relevant tumor microenvironments



FlowSOM

flow or mass cytometry

imaging

next-generation sequencing

FlowSOM: Using Self-Organizing Maps for Visualization and Interpretation of Cytometry Data

Sofie Van Gassen,[1,2,3]* Britt Callebaut,[1] Mary J. Van Helden,[2,3] Bart N. Lambrecht,[2,3] Piet Demeester,[1] Tom Dhaene,[1] Yvan Saeys[2,3]

A Systems-Level Study Reveals Regulators of Membrane-less Organelles in Human Cells

Doris Berchtold,[1] Nico Battich,[1] and Lucas Pelkmans[1,2,*]
[1]Institute of Molecular Life Sciences, University of Zurich, 8057 Zurich, Switzerland
[2]Lead Contact
*Correspondence: lucas.pelkmans@imls.uzh.ch
https://doi.org/10.1016/j.molcel.2018.10.036

A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging

Leeat Keren,[1] Marc Bosse,[1] Diana Marquez,[1] Roshan Angoshtari,[1] Samir Jain,[1] Sushama Varma,[1] Soo-Ryum Yang, Allison Kurian,[1] David Van Valen,[2,3] Robert West,[1] Sean C. Bendall,[1,*] and Michael Angelo[1,4,*]
[1]Department of Pathology, Stanford University, Stanford CA, 94305, USA
[2]Department of Biology, Caltech, Pasadena, CA 91125, USA
[3]Department of Bioengineering, Caltech, Pasadena, CA 91125, USA
[4]Lead Contact
*Correspondence: bendall@stanford.edu (S.C.B.), mangelo0@stanford.edu (M.A.)
https://doi.org/10.1016/j.cell.2018.08.039

Todorov 2018
doi: 10.1111/febs.14613