

# Journées Pratiques

## Analyses non supervisées en cytométrie



Du 5 au 7 Février 2020  
Sophia-Antipolis



*Présidées par :*

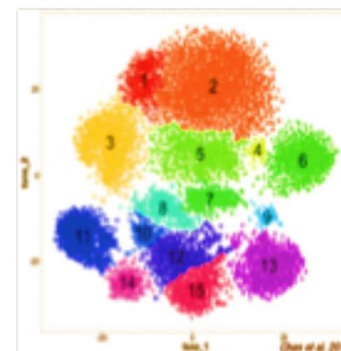
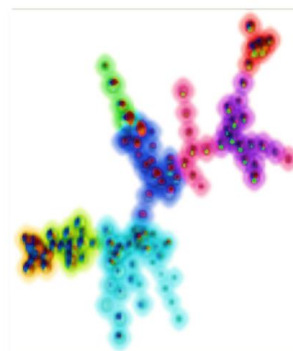
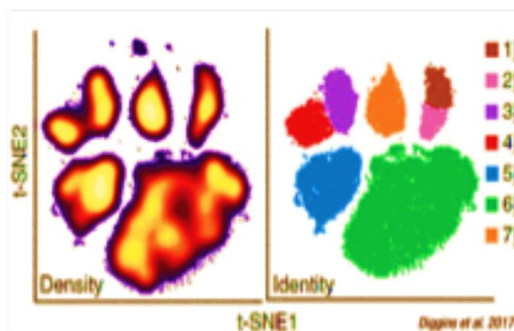
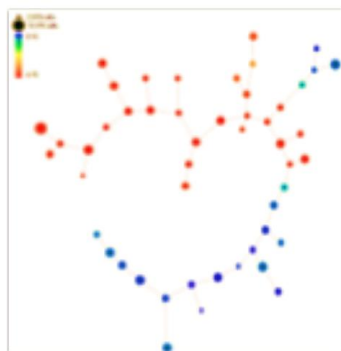
**Jonathan M. IRISH**

Mass Cytometry Center of Excellence (MCCE), Vanderbilt University, Nashville, (USA)

*Co-animées par :*

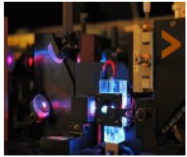
**Aïda Meghraoui-Kheddar** (IPMC, CNRS Valbonne) et **Samuel Granjeaud** (CRCM, INSERM Marseille)

**Julie Cazareth** (IPMC, CNRS Valbonne) et **Sierra Barone** (Vanderbilt University, Nashville USA)

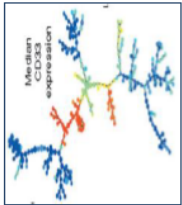


# Passeport cytométrie Marseille

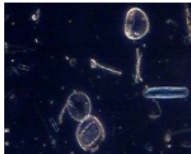
## Édition 2020



**Cytométrie multiparamétrique avancée : théorie & pratique**  
du 24 au 27 mars 2020

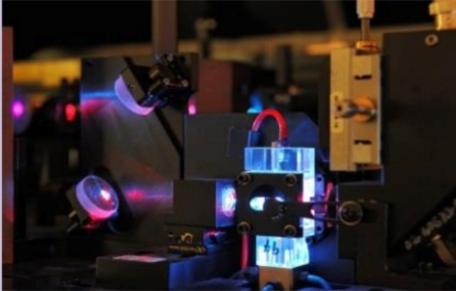


**Outils d'analyse de données avancées  
en cytométrie de flux et de masse : théorie & pratique**  
du 9 au 12 juin 2020



**Cytométrie en flux : du photon à la cellule : théorie & pratique**  
novembre 2020

# Passeport cytométrie Marseille



## Cytométrie multiparamétrique avancée

Du 24 au 27 mars 2020, à Marseille

### Public

Tout public intéressé par l'application de la cytométrie à son champ expérimental et désireux d'augmenter le nombre de paramètres étudiés simultanément sur ses cellules d'intérêt

### Prérequis

Connaissances de base en cytométrie de flux conventionnelle (4-6 paramètres)

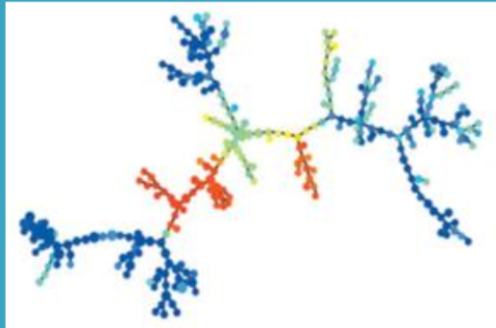
### Objectif

Mieux appréhender les approches multiparamétriques en cytométrie (Flux, Masse et Spectrale)

### Programme

- Rappels théoriques sur la cytométrie en flux multiparamétrique
- Réglages, optimisation et standardisation des cytomètres
- Notions sur les microparticules : optimisation des paramètres d'acquisitions
- Mise au point de panel à façon, marquage 15 couleurs
- Acquisition sur BD LSR2 et Fortessa
- Initiation à la cytométrie de Masse sur Helios
- Initiation à la cytométrie Spectrale sur Cyttek Aurora
- Notions de signalisation intracellulaire

# Passeport cytométrie Marseille



Outils d'analyse de données avancées en  
cytométrie de flux et de masse - Théorie & pratique

**Du 9 au 12 juin 2020, Marseille**

## Public

Chercheurs et ingénieurs effectuant des analyses en cytométrie multiparamétrique.  
Cette offre de formation s'adresse aux biologistes de préférence.

**Prérequis : notions d'anglais**

## Objectifs

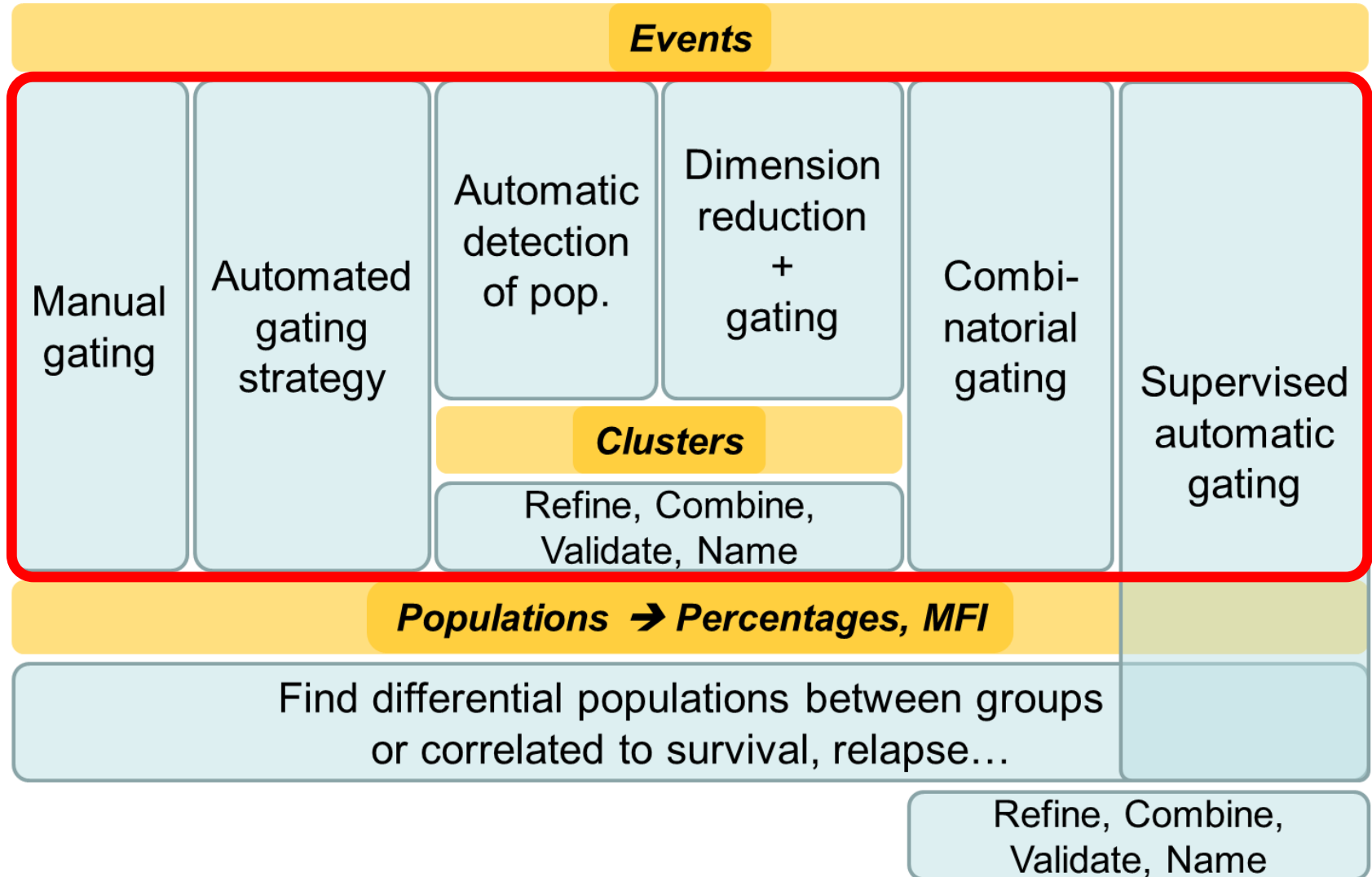
- Connaître les outils actuels d'analyse de données avancées, les mettre en œuvre sur des jeux de données tests afin d'apprendre à les maîtriser
- Permettre aux participants d'identifier la méthode d'analyse de choix appropriée pour une question définie et extraire le maximum d'informations à partir d'un set de données
- Présenter des solutions logicielles simples permettant de visualiser et de synthétiser les résultats autrement qu'en histogrammes ou en cytogrammes bivariants classiques
- Réaliser des analyses intégratives de données issues aussi bien de plusieurs analyses complexes en cytométrie de flux que d'autres types de tests (multiplex immuno-assay, formule sanguine...)

**Lieu :** Délégation régionale Inserm - 13009 Marseille

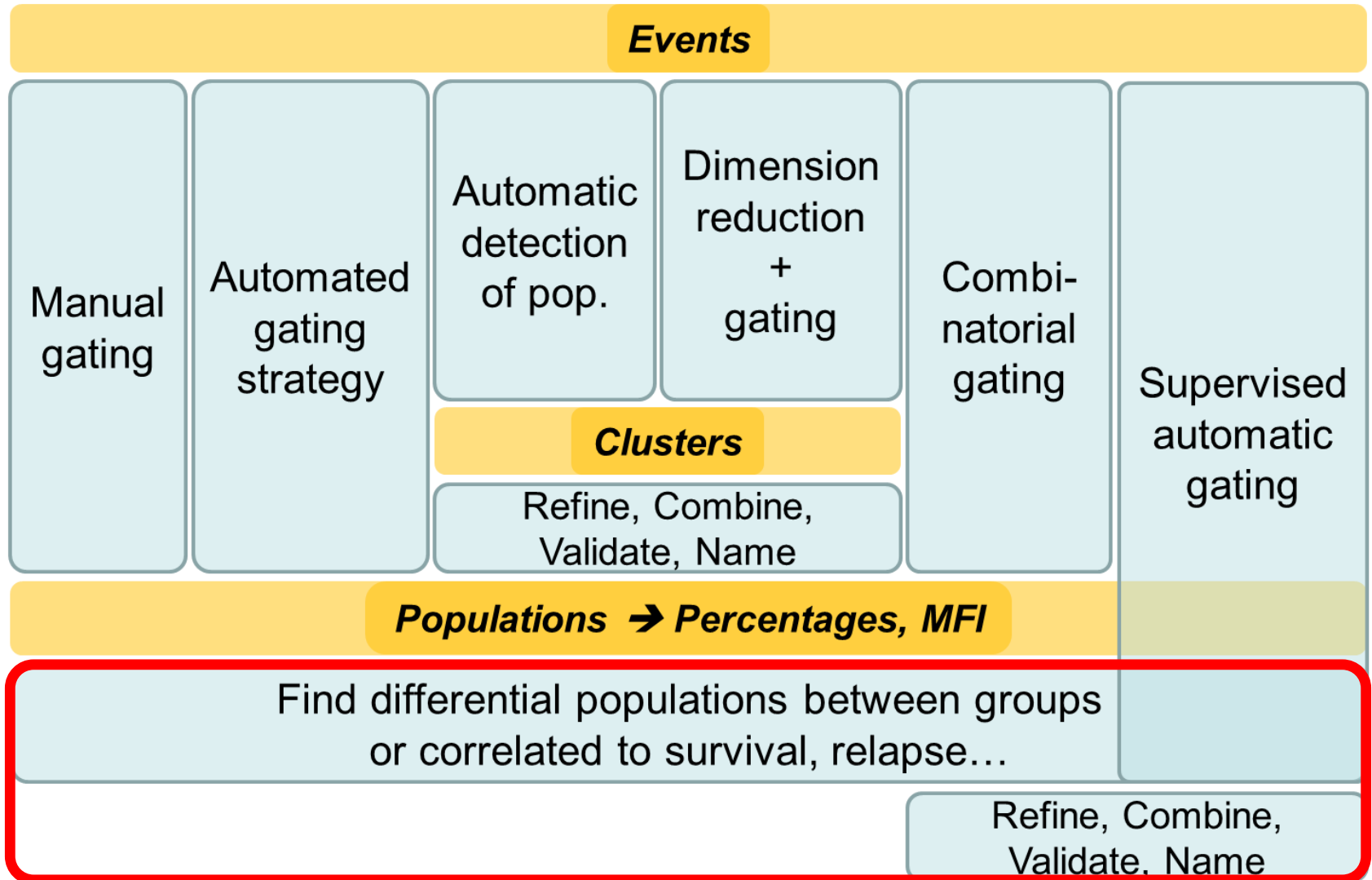
**Participants :** 8 personnes

# **PIPELINES OVERVIEW AND DETAILS**

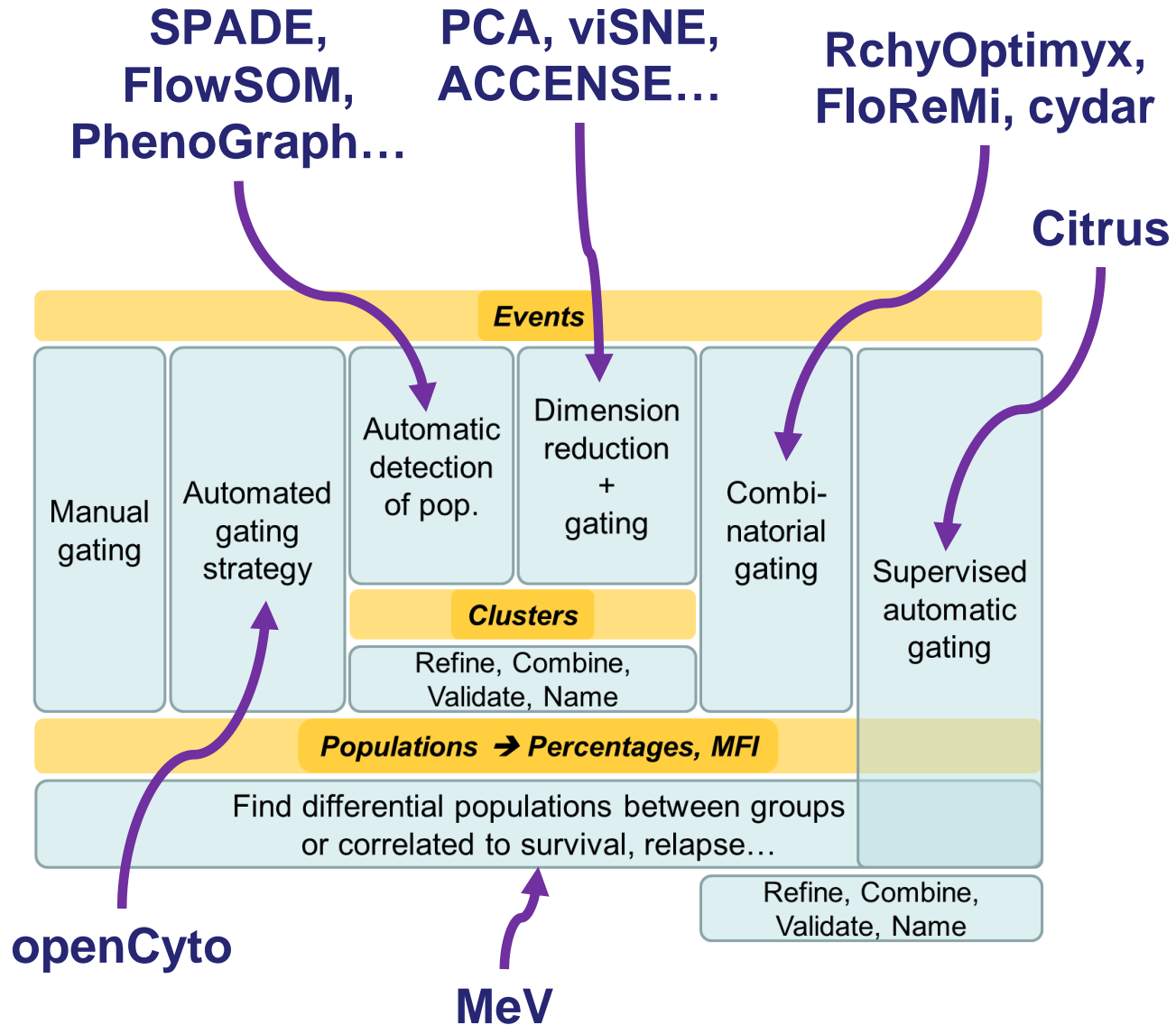
# Strategies



# Features analysis

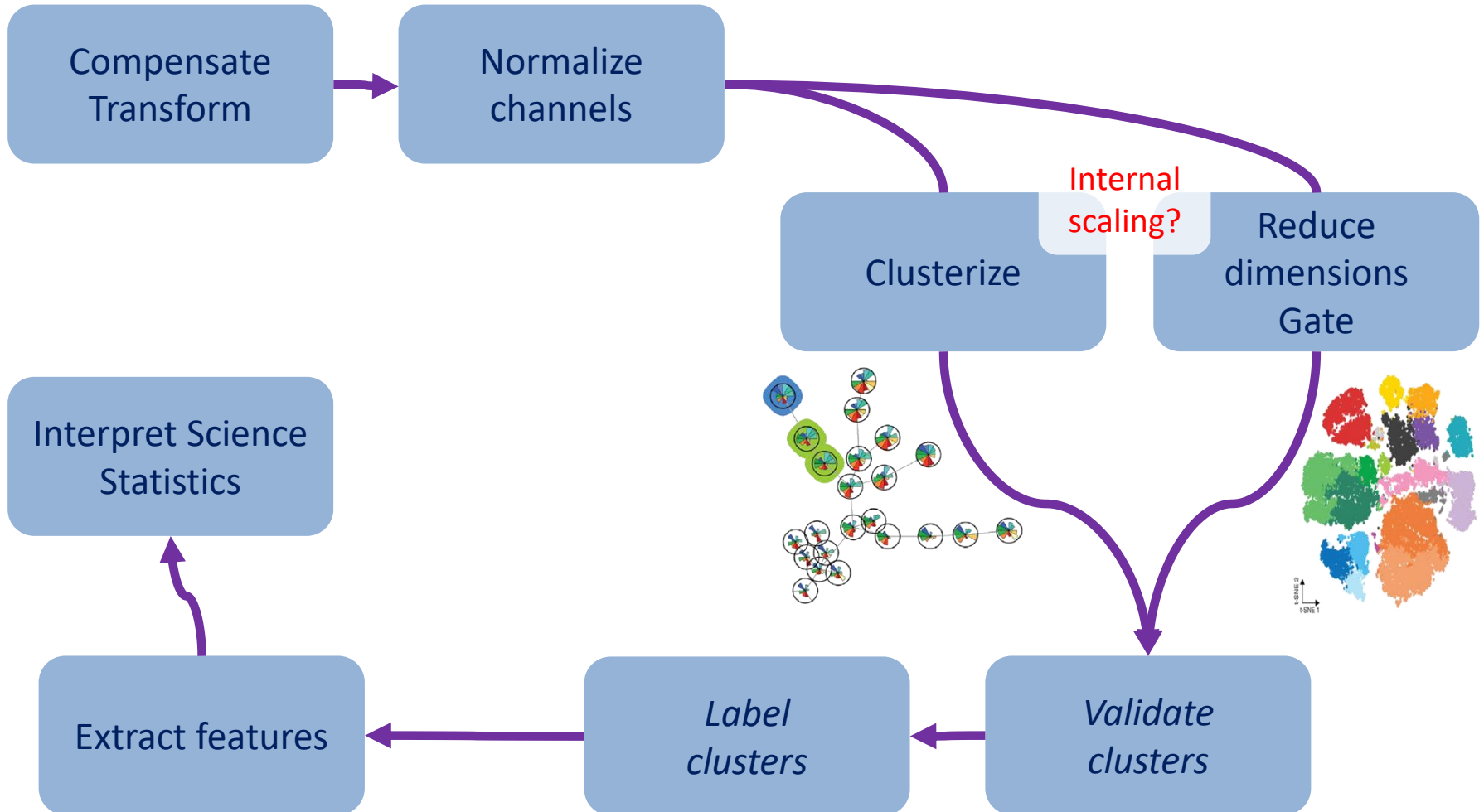


# Free tools



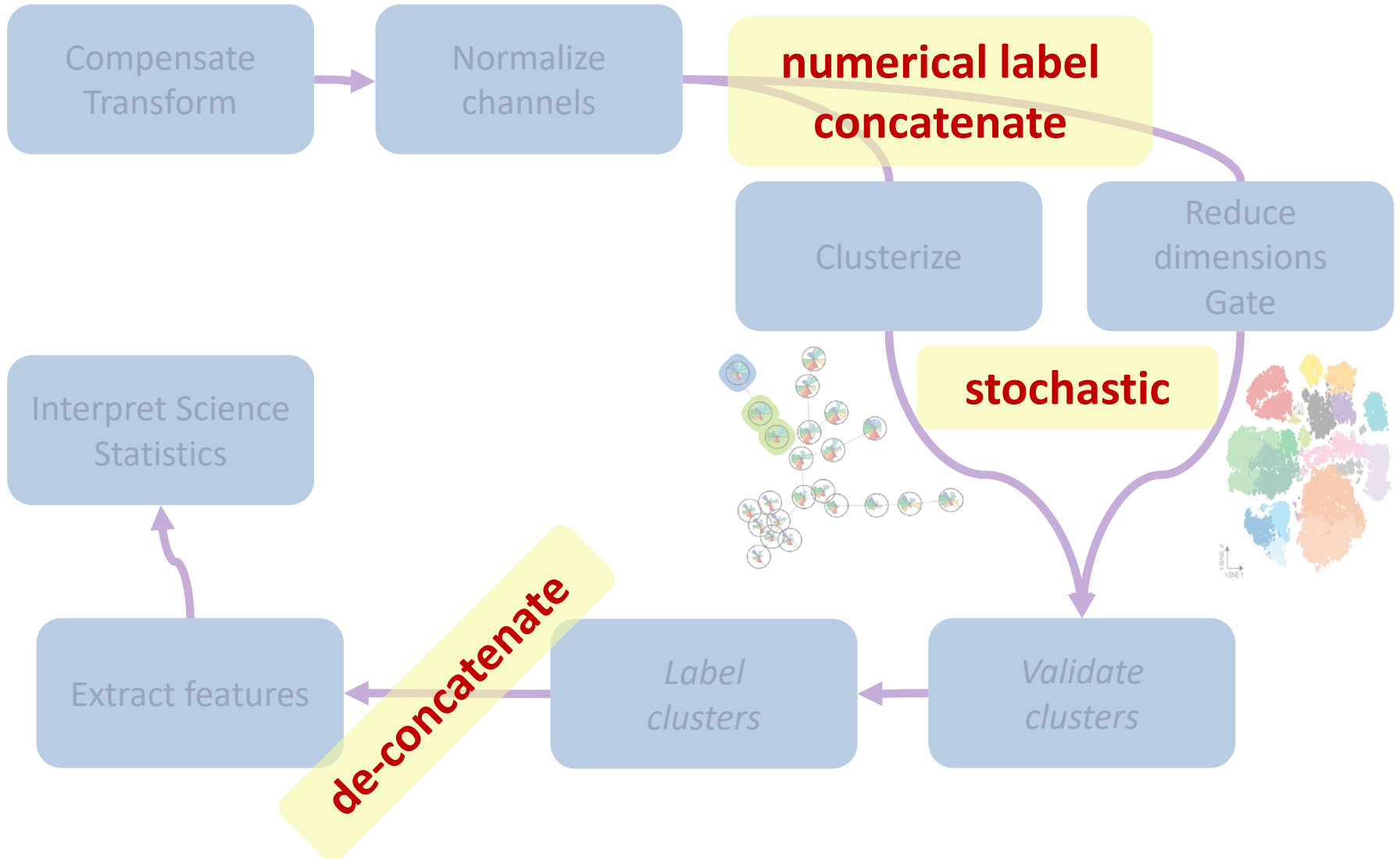


# Pipeline (one sample)

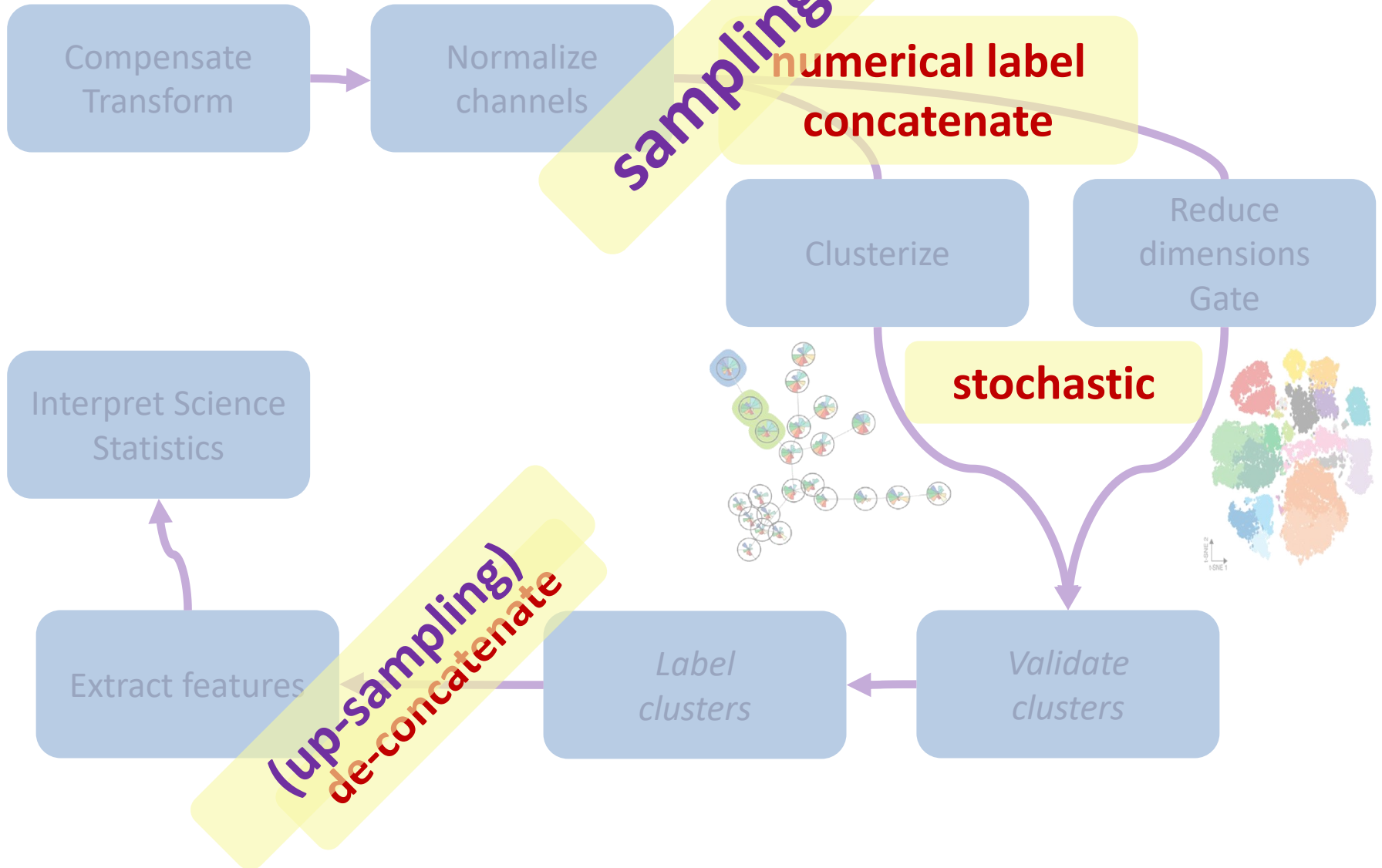


**strength = weakest step**

# Pipeline (multi-samples)



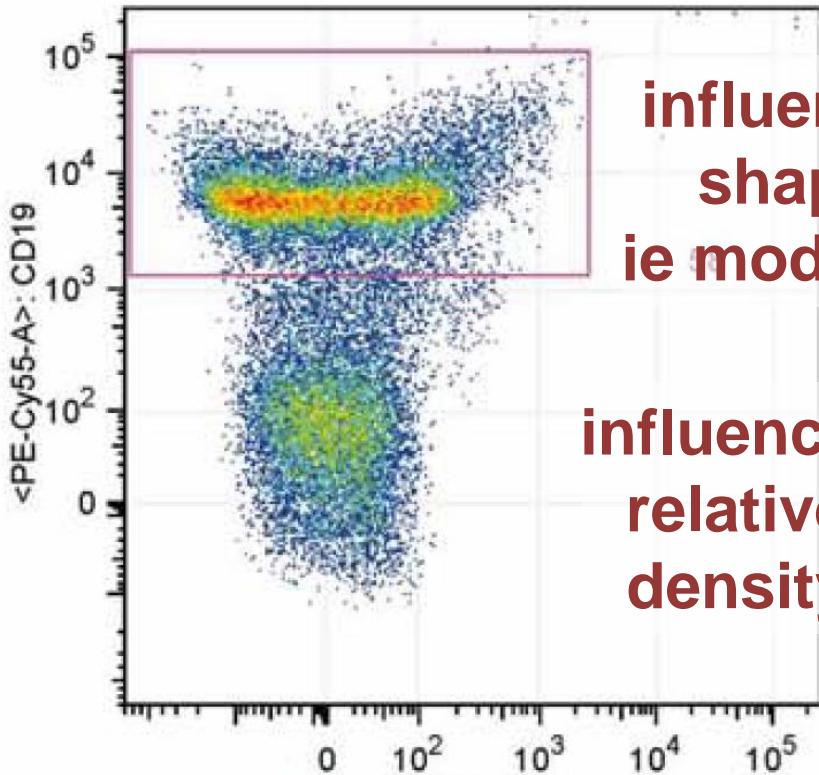
# Pipeline (multi-samples)



# **TRANSFORMATIONS: ASINH VS BI-EXP**

# Transformation parameters

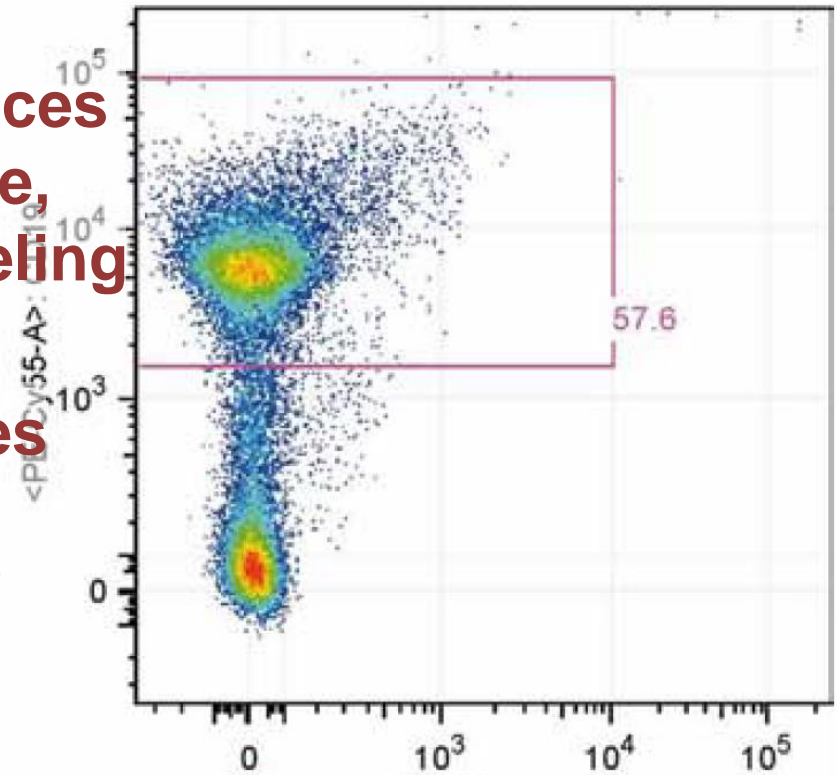
D. Compensated  
 $M=4.5, A=1.5, W=0.25$



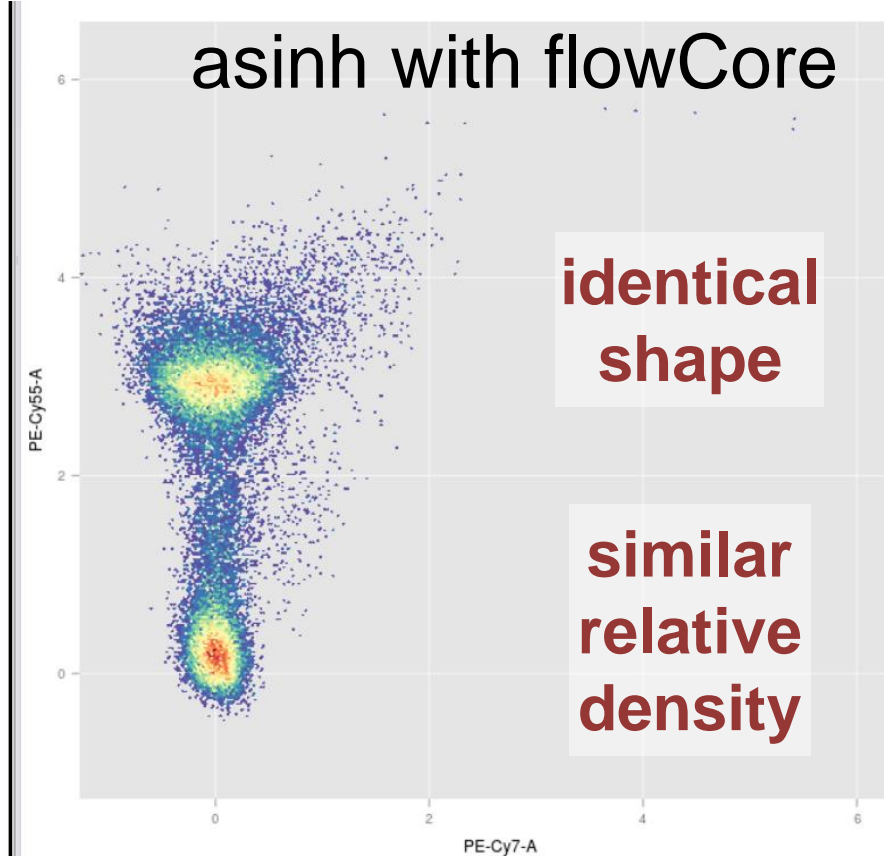
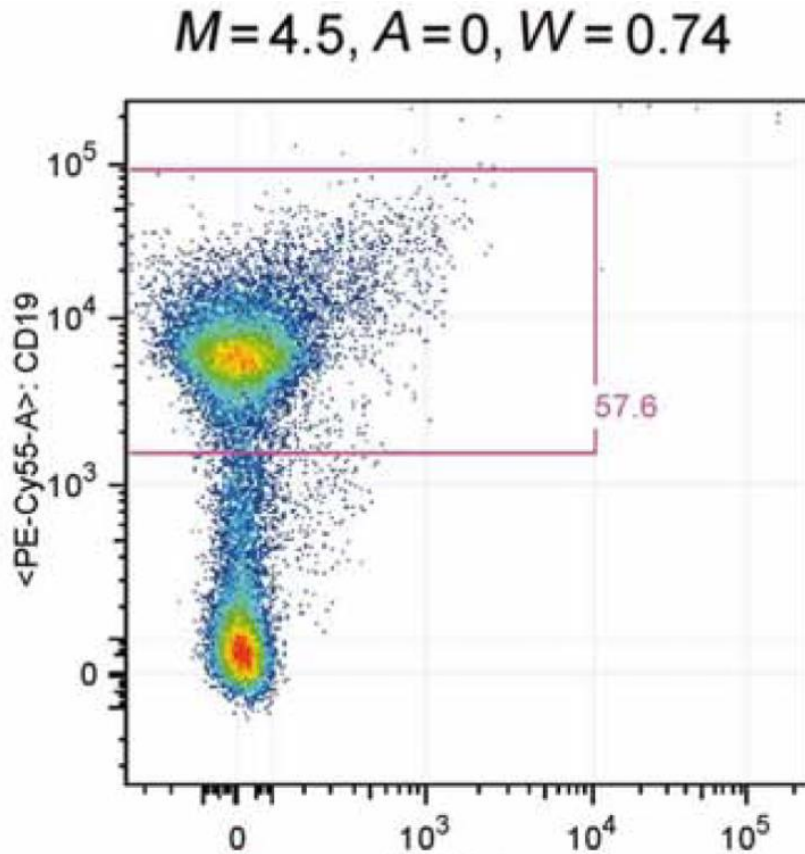
influences  
shape,  
ie modeling

influences  
relative  
density

E. Compensated  
 $M=4.5, A=0, W=0.74$



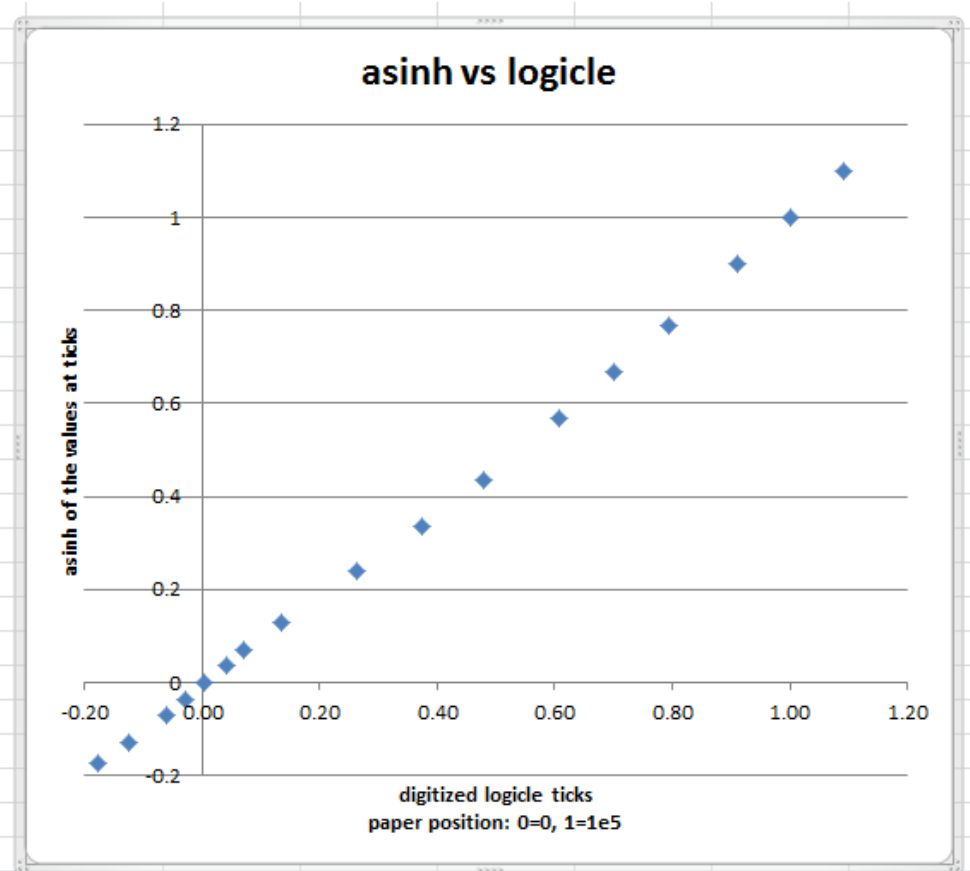
# logicle vs asinh



# logicle vs asinh

## one coefficient

		scaling	200		
log10	units	digitized	asinh	delta%	
1.06	2.00E+05	1.09	1.10	-1	
1.00	1.00E+05	1.00	1.00	0	
0.94	5.00E+04	0.91	0.90	1	
0.86	2.00E+04	0.79	0.77	4	
0.80	1.00E+04	0.70	0.67	5	
0.74	5.00E+03	0.61	0.57	7	
0.66	2.00E+03	0.48	0.43	10	
0.60	1.00E+03	0.37	0.33	12	
0.54	5.00E+02	0.26	0.24	10	
0.46	2.00E+02	0.13	0.13	6	
0.40	1.00E+02	0.07	0.07	2	
0.34	5.00E+01	0.04	0.04	17	
0.00	0.00E+00	0.00	0.00	#DIV/0!	
0.34	-5.00E+01	-0.03	-0.04	-20	
0.40	-1.00E+02	-0.06	-0.07	-11	
0.46	-2.00E+02	-0.13	-0.13	-1	
0.50	-3.00E+02	-0.18	-0.17	2	

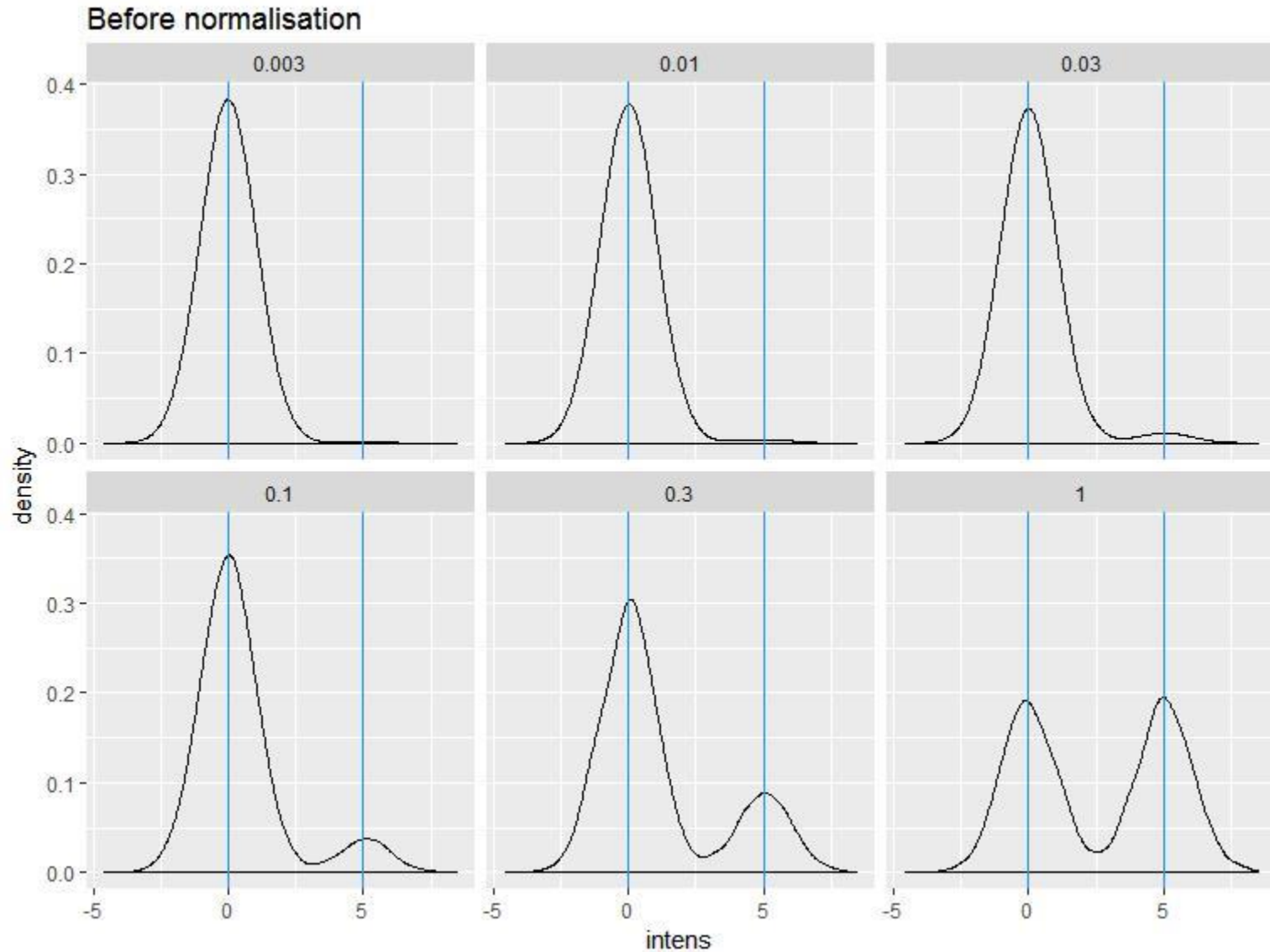


available in Excel ;-)

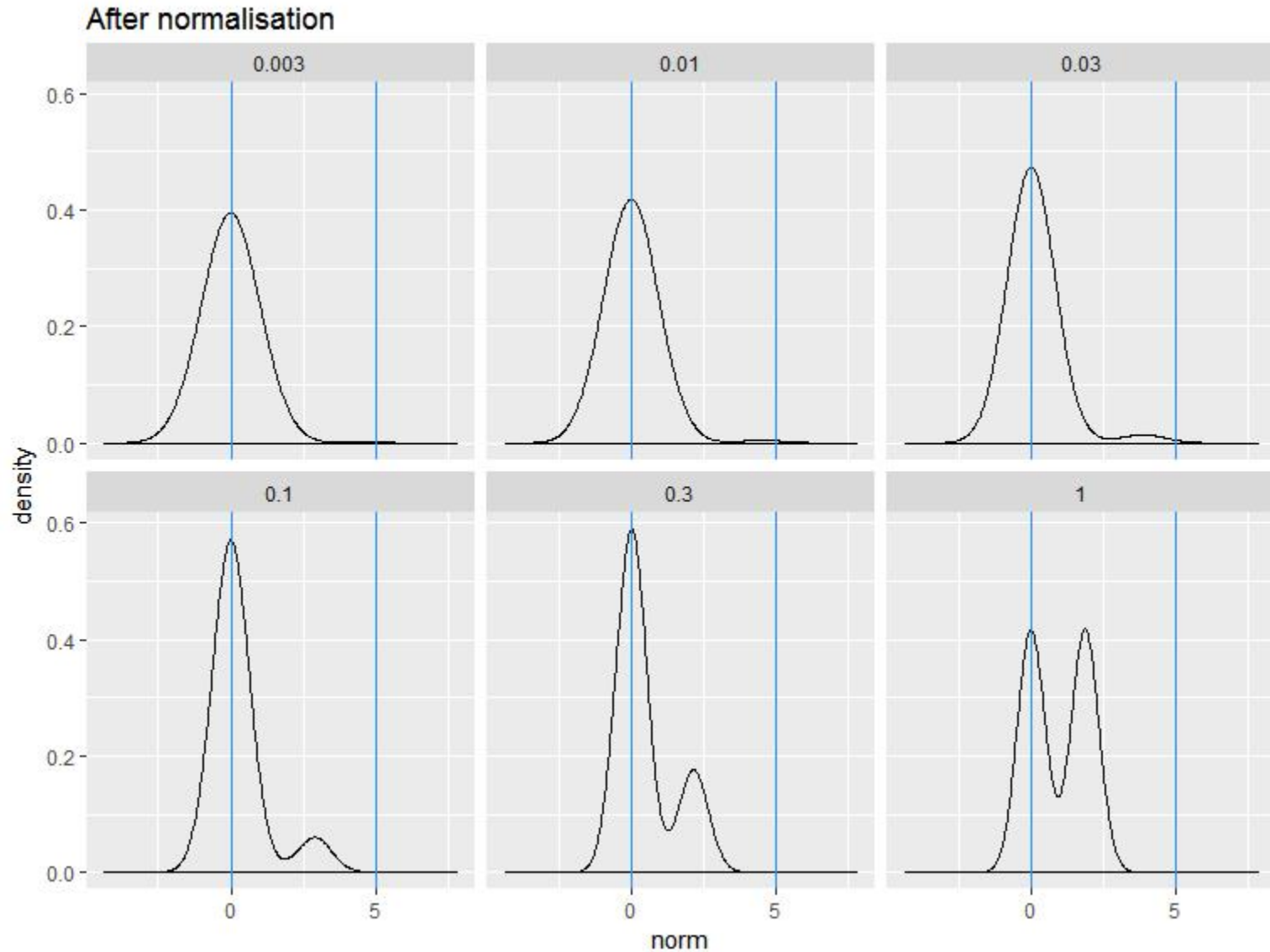
# **NORMALIZATION DIFFICULTIES**



# Normalisation, scaling...

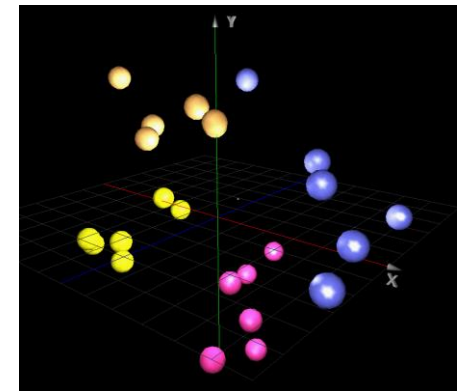
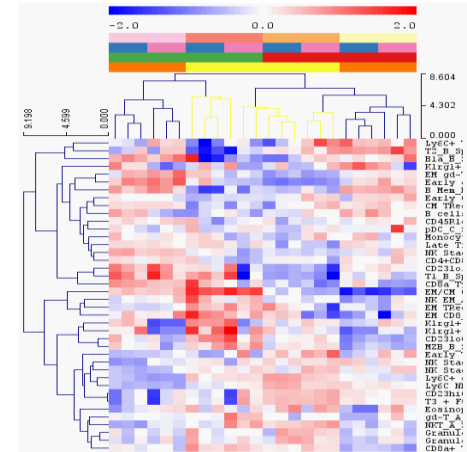
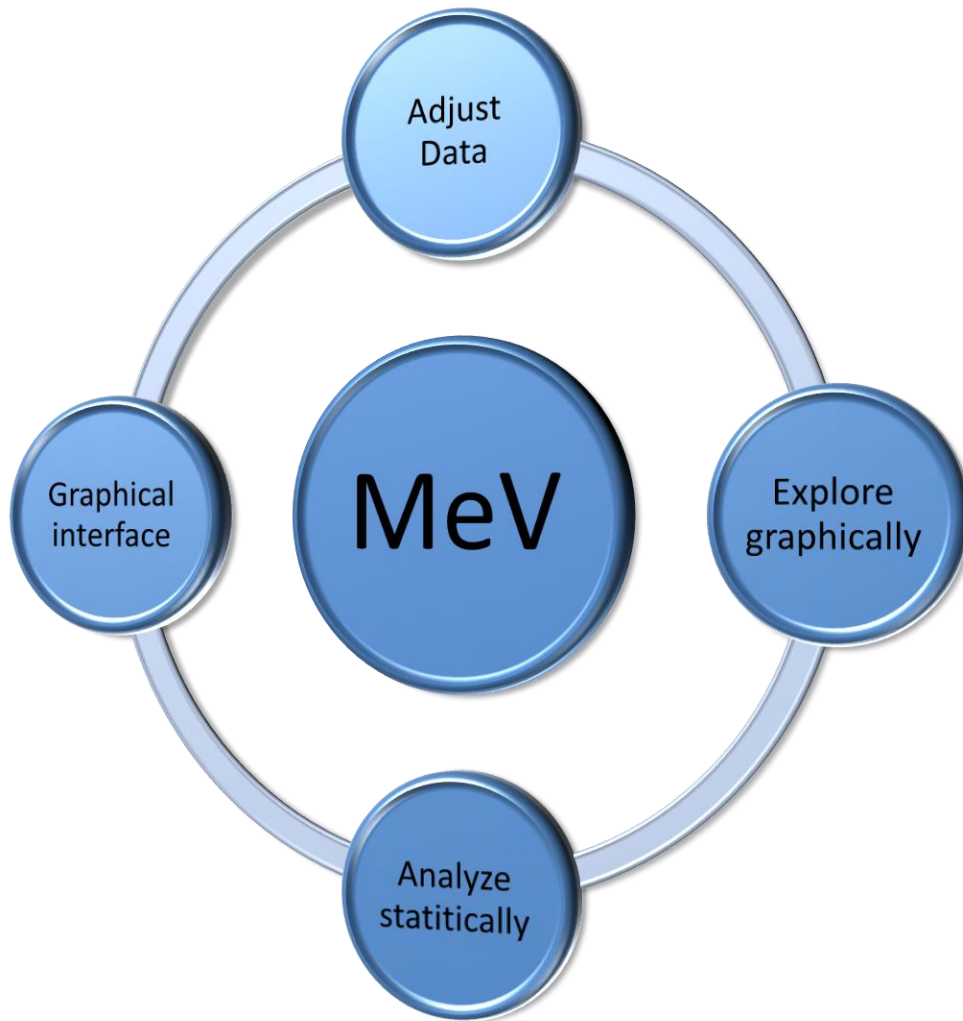


# Normalisation, scaling...



# **PERCENTAGES ANALYSIS**

# MeV capabilities



# Analysis Pipeline

## Adjust

- Log2 transform (or asinh)
- Center each row
- *Filter rows*

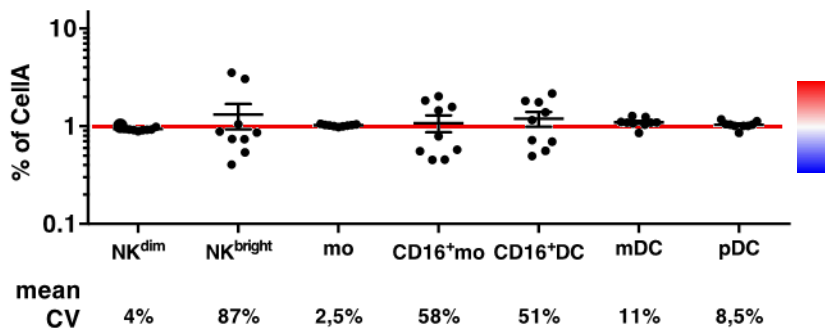
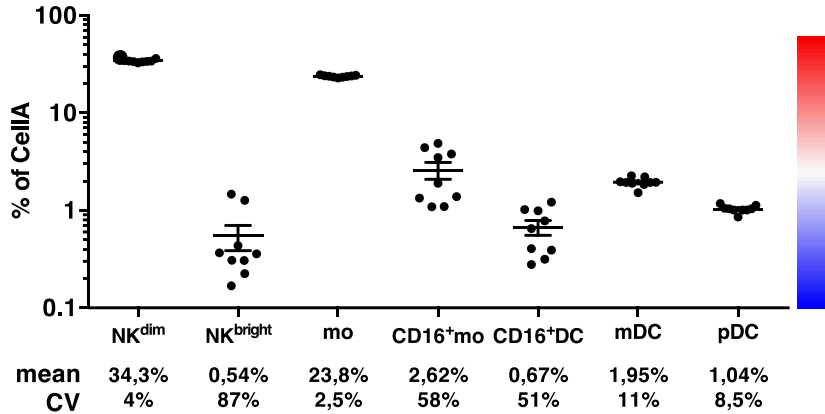
## Explore

- Unsupervised analysis: HClustering, PCA
- Identify sample outliers
- *Remove/mute sample outliers*

## Identify

- Supervised analysis: statistical methods
- Identify difference between groups
- *Ignore small differences*

# Adjust



- Percentages are usually displayed in a log scale
- some MFI also
- Luminex concentr.
- Centering focus on differences between groups within populations

# Log2 properties

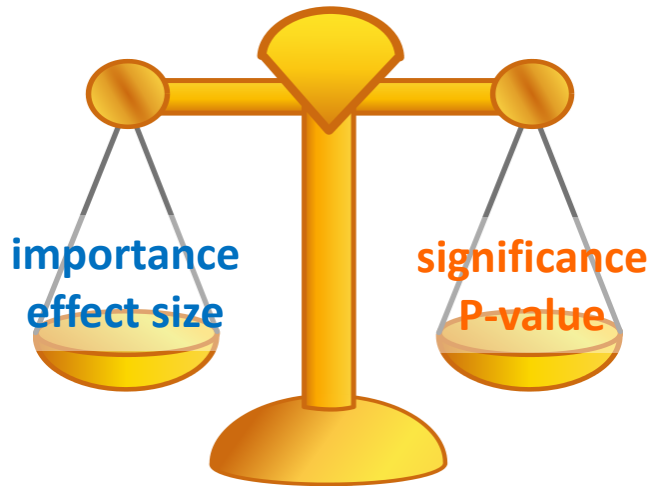
- $\log_2$  is proportional to  $\log_{10}$
- 1 qRT-PCR cycle  $\sim \times 2$
- ratio  $\Rightarrow$  addition  
     $\times 2 \Rightarrow + 1$
- is symmetric:  $+100\% = \times 2 = +1$   
                   $- 50\% = / 2 = -1$
- stabilizes the dispersion
- $\log_2( a / b ) = \log_2( a ) - \log_2( b )$

# Statistics

- Important, Unavoidable
- Abusive or abused
  - » Statistics does not tell us whether we are right. It tells us the chances of being wrong.
- Point Of Significance in Nat. Meth.
  - » <http://mkweb.bcgsc.ca/pointsofsignificance/>
  - » [http://blogs.nature.com/methagora/2013/08/giving\\_statistics\\_the\\_attention\\_it\\_deserves.html](http://blogs.nature.com/methagora/2013/08/giving_statistics_the_attention_it_deserves.html)

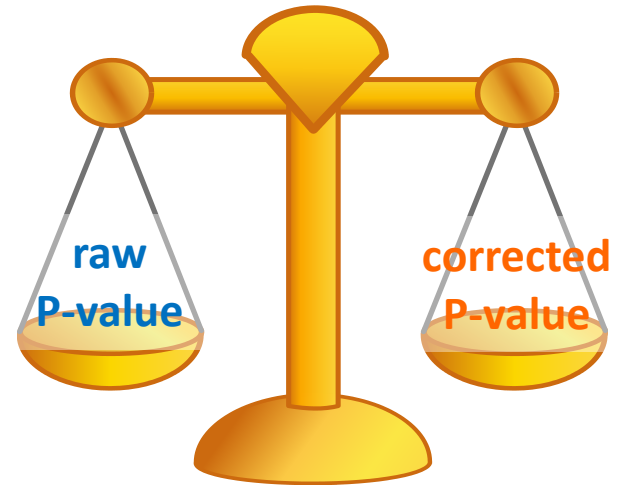


# Statistical trade-offs



**Volcano Plot**

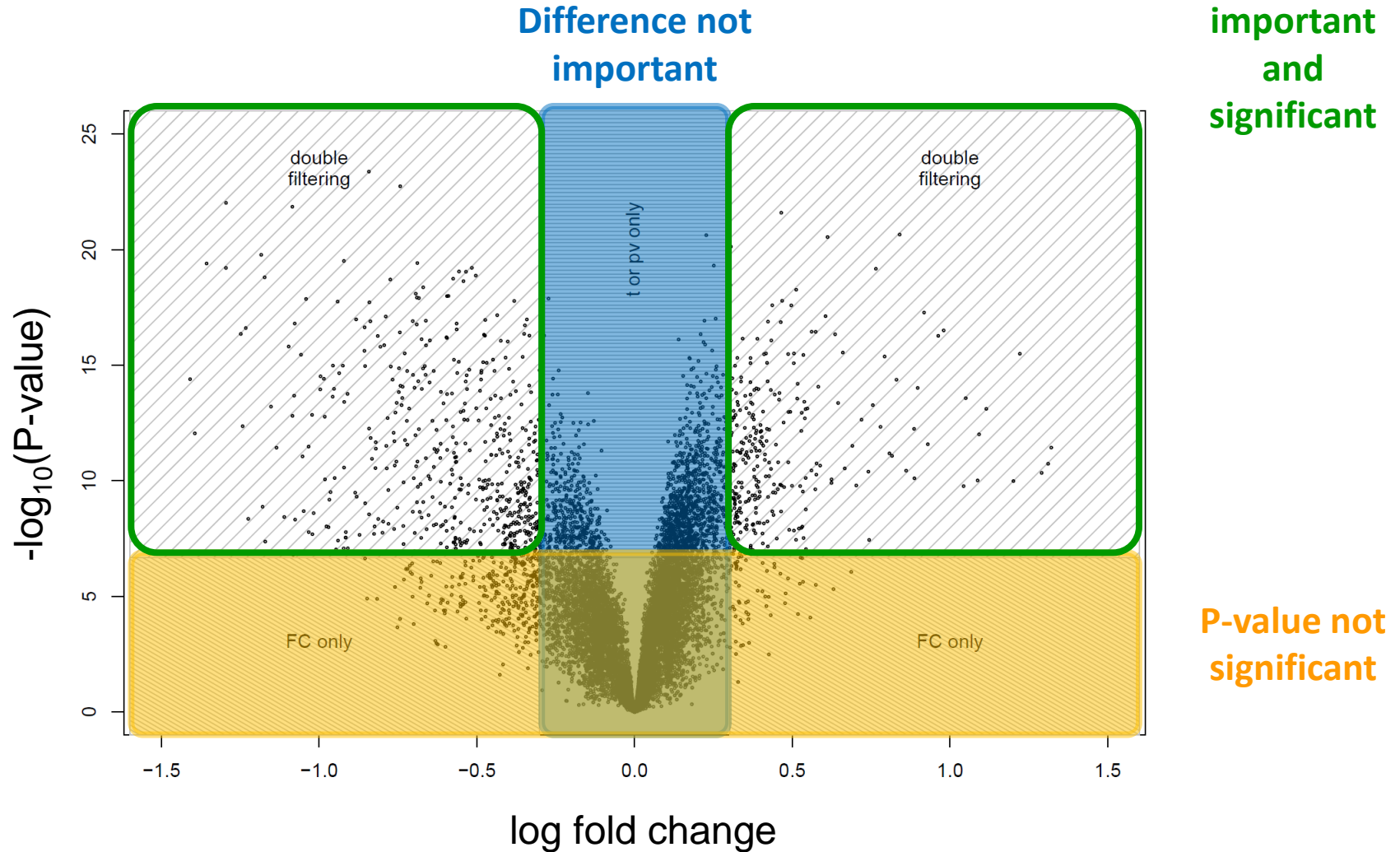
**report important  
and significant**



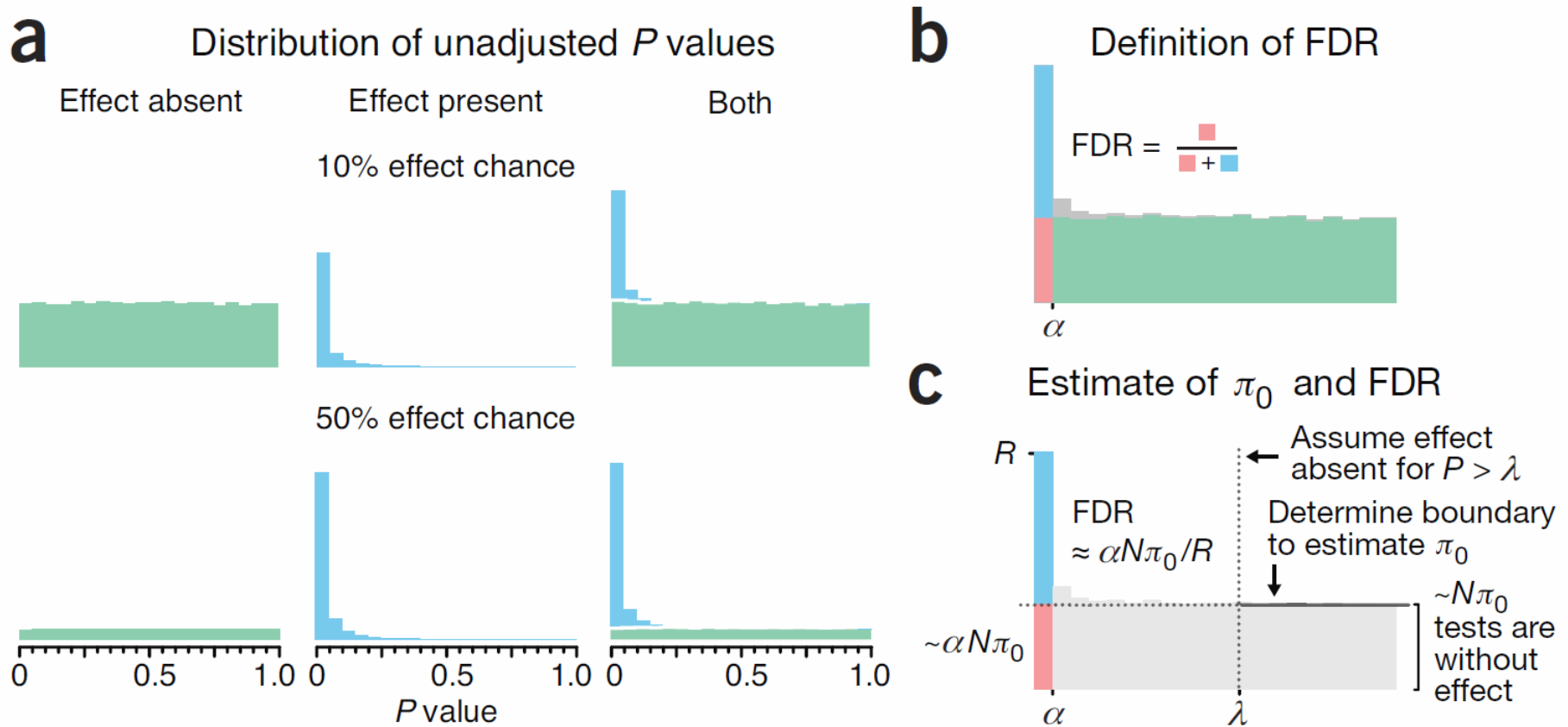
**False Discovery Rate**

**correct  
multiple tests**

# FC vs P: Volcano plot



# More about multiple testing



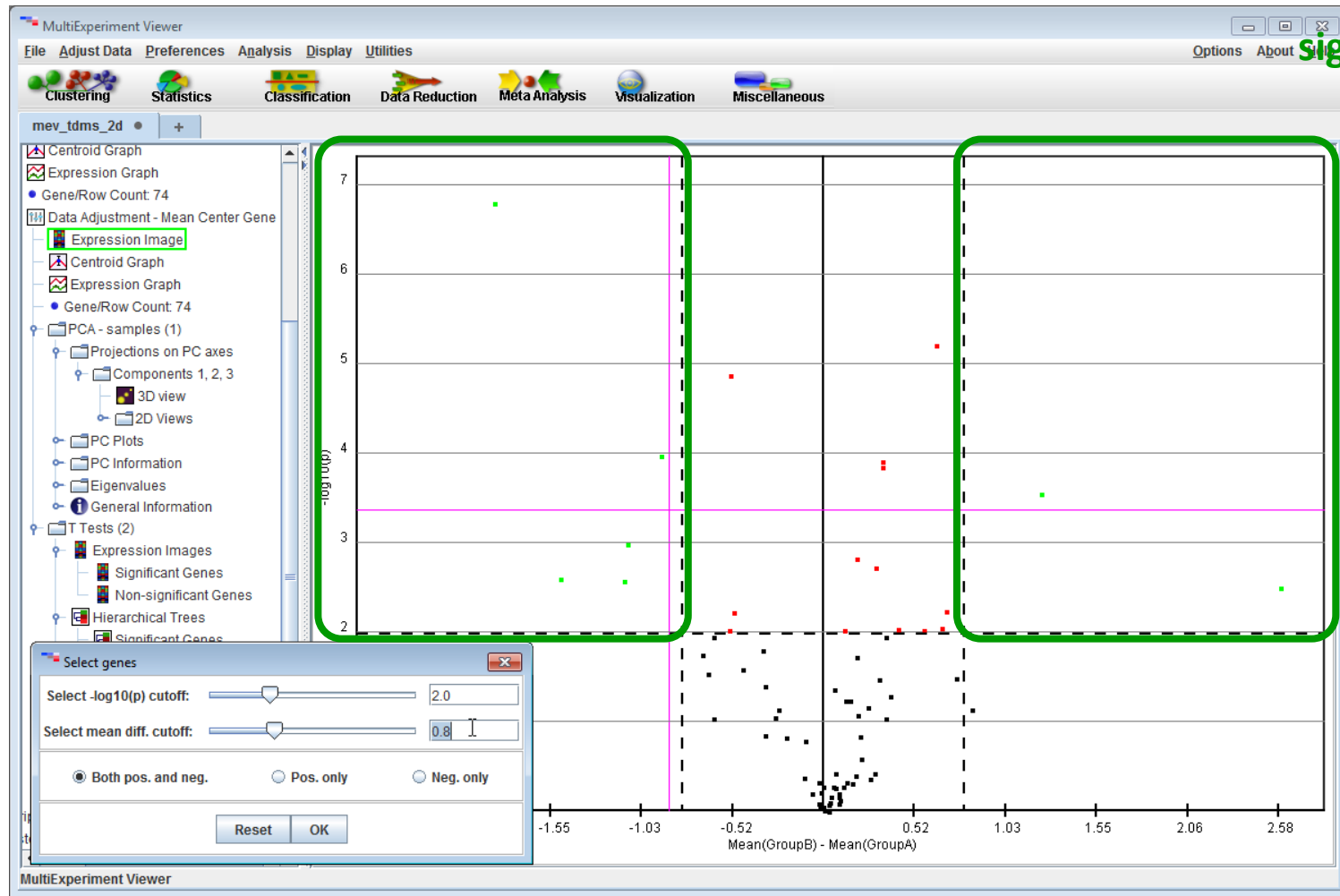
**Figure 3** | The shape of the distribution of unadjusted  $P$  values can be used to infer the fraction of hypotheses that are null and the false discovery rate (FDR).

# What is "significance"?

- Statistical significance is not the same as practical importance.
- P-value does not tell whether the result is of a **practical importance**.
- Statistics does not tell us whether we are right. It tells us the chances of being wrong.
- Any particular threshold for declaring significance is arbitrary.

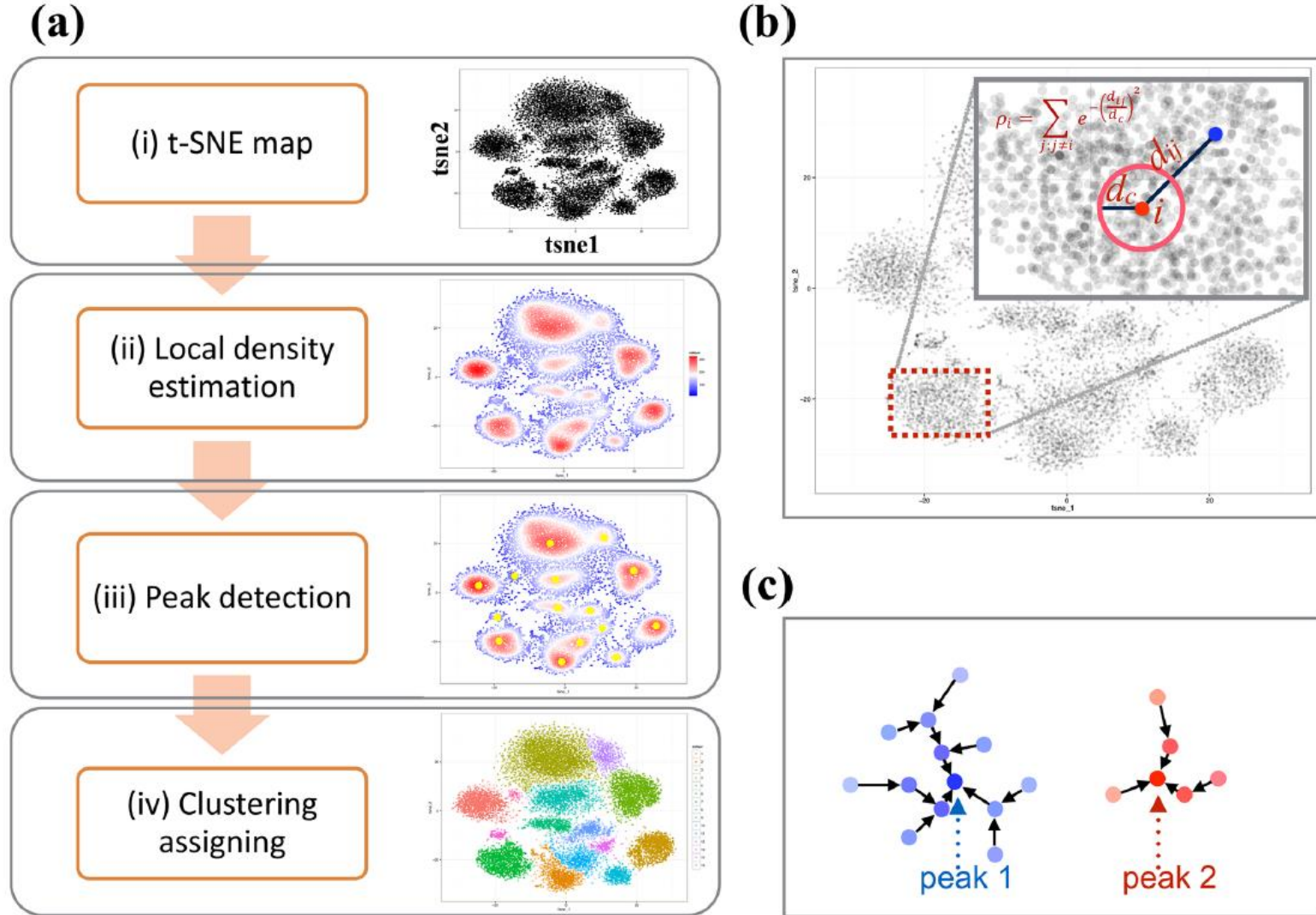
# Difference threshold

important  
and  
significant



# **CLUSTERING ALGORITHMS IN CYTOFKIT**

# ClusterX

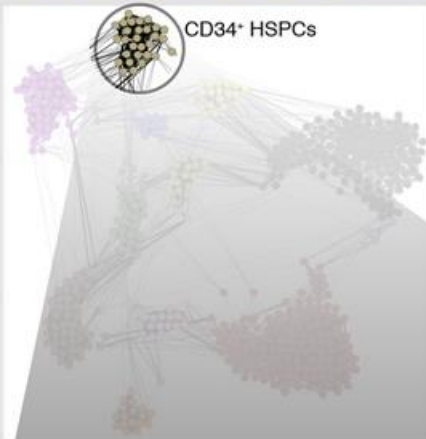


**Fig 2. Workflow of ClusterX for mass cytometry data clustering.** (a) depict the workflow of ClusterX for mass cytometry data clustering, which contains four steps: (i) t-SNE dimensionality reduction (ii) estimate the local density on the t-SNE map (iii) detect the density peaks represented as cluster centers and (iv) assign the remaining cells to clusters. (b) Explains the local density estimation method. (c) Illustrate the cluster assigning step using two peaks, peak1 and peak 2. Each point is a cell and the color intensity represents the local density of the cell. Then each cell is assigned to be the same cluster as its nearest neighbor cell which has higher density than it.

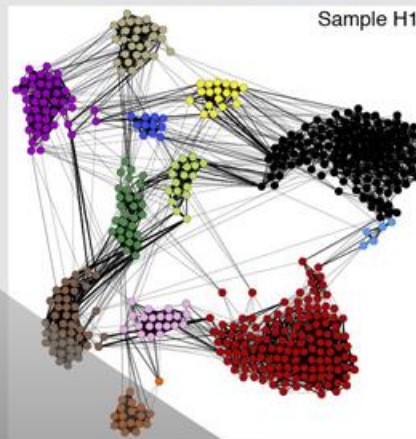
# Phenograph

PhenoGraph

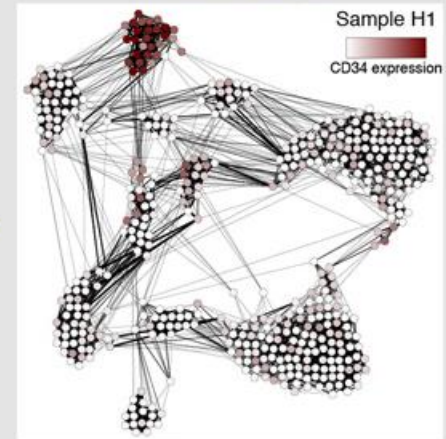
Extract surface and signaling features for each subpopulation



Partition each graph into distinct subpopulations

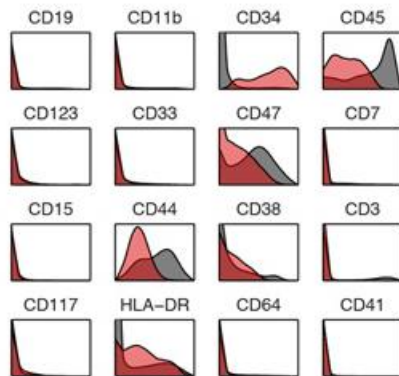


Build single-cell graph for each sample

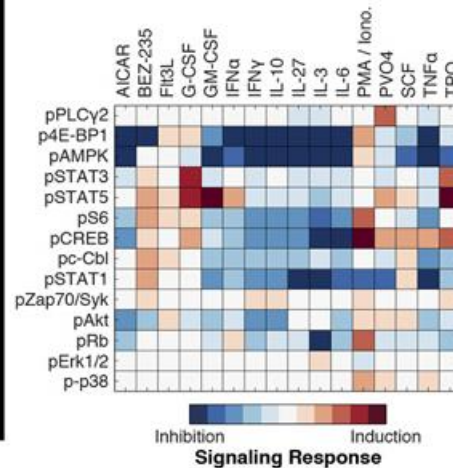


Repeat for each sample

## Surface Phenotype



## Intracellular Signaling





# Résumé

- ClusterX
  - recherche de sommets et de montagnes
  - [Rodriguez et Laio Science 2014](#)
- Phenograph
  - recherche de groupes dans les réseaux sociaux
  - [Levine et al. Cell 2015 Web](#)
- FlowSOM
  - déformation d'une grille pour l'adapter aux données
  - [van Gassen et al. Cyto A 2015 Web](#)