

# Cytometry dATa anALYSIS Tools

&

**The solution**

Bead based compensation to correct for channel crosstalk in mass cytometry

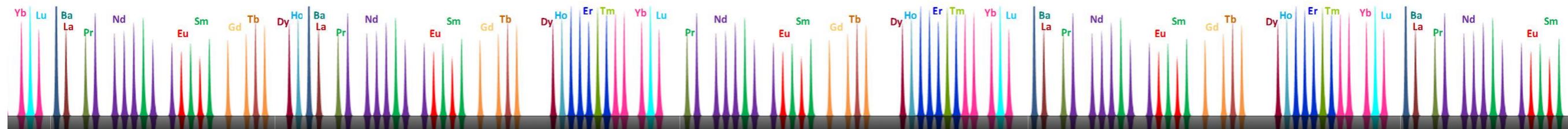
**The problem**

**The data**

Stéphane Chevrier\*, Helena L Crowell\*, Vito Zanotelli, Mark D Robinson, Bernd Bodenmiller

Institute of Molecular Life Sciences, University of Zurich, Switzerland

SIB Swiss Institute of Bioinformatics, University of Zurich, Zurich, Switzerland

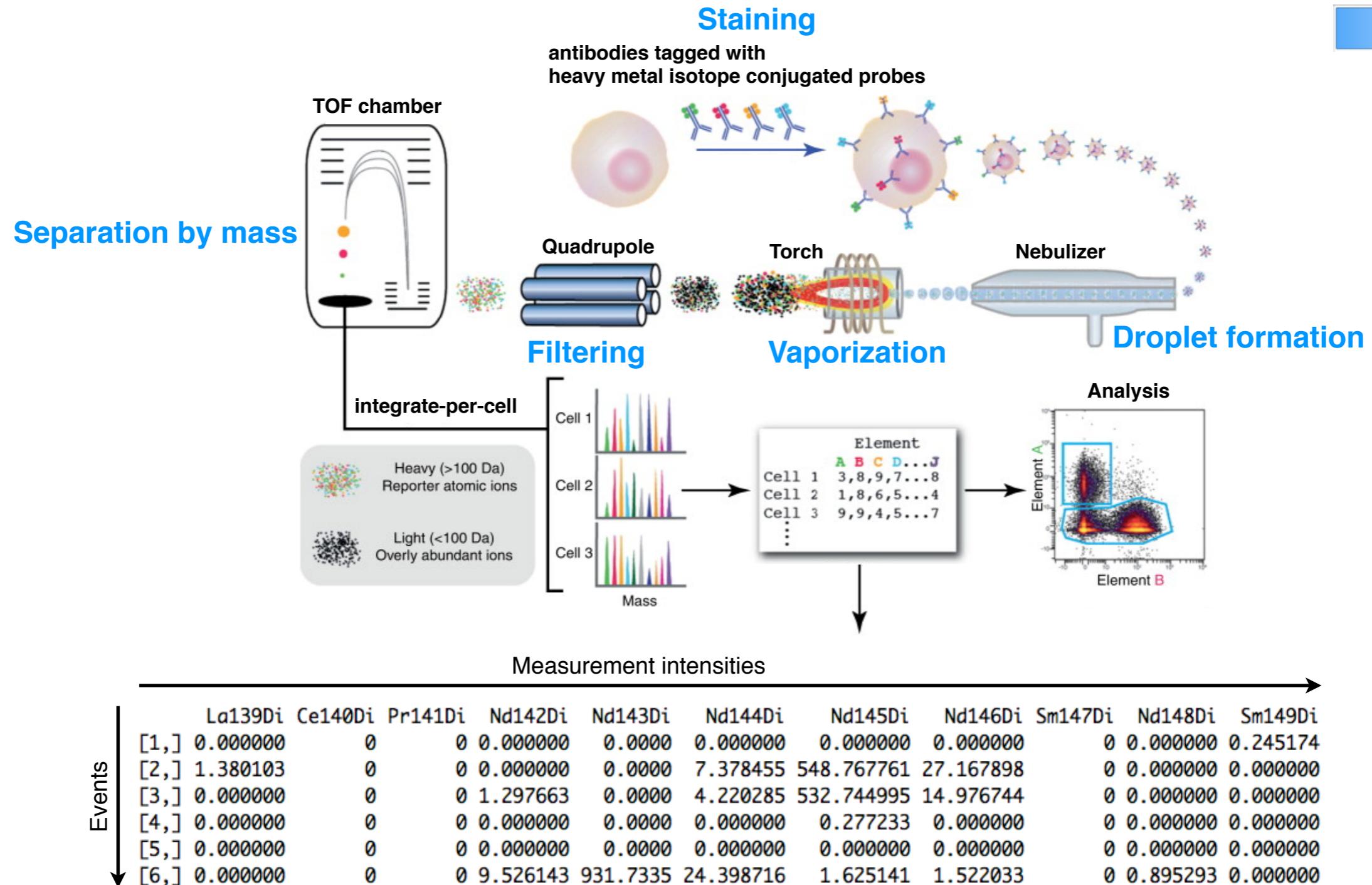


**Helena L Crowell, BSc**

Research Assistant, Robinson group UZH

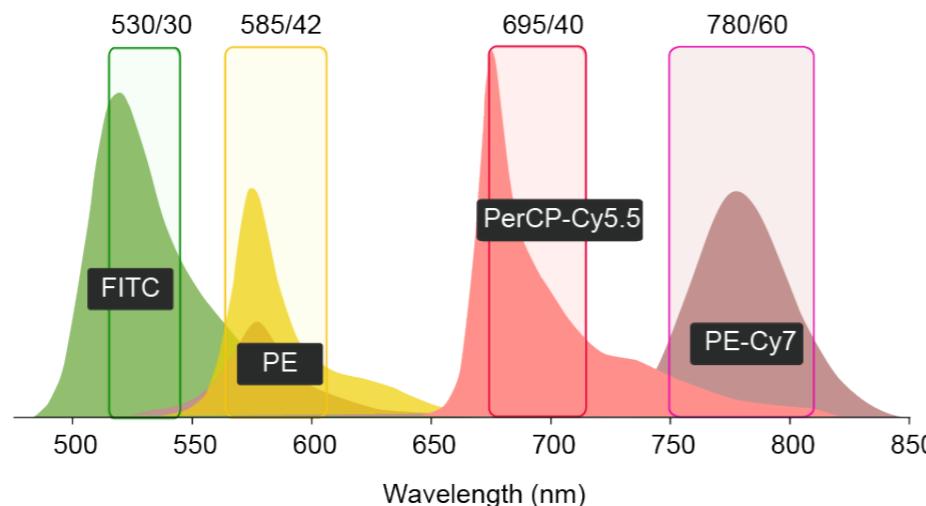
Master's Student Computational Biology and Bioinformatics, ETH

# Cytometry by Time of Flight (CyTOF)

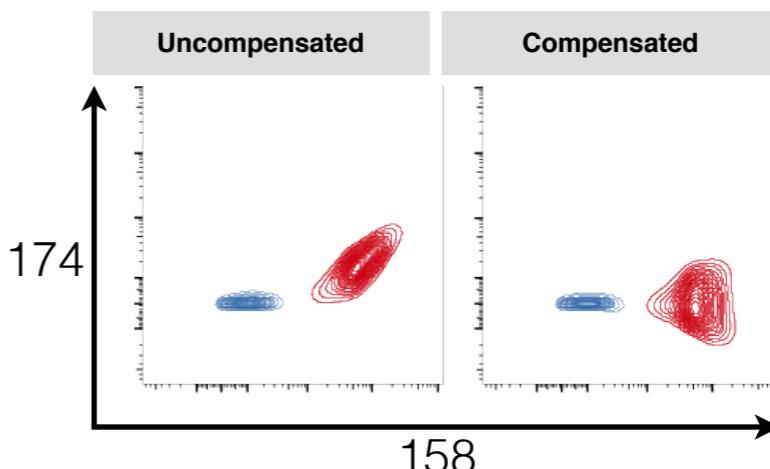
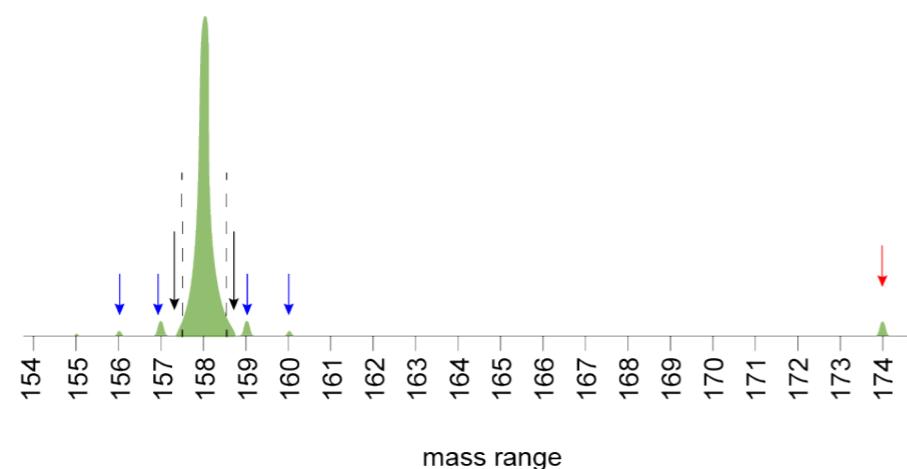


# Spectral overlap vs. spillover in mass cytometry

FACS (fluorescence-activated cell sorting)



Mass cytometry



**isotopic impurities:**

contamination after isotopic enrichment

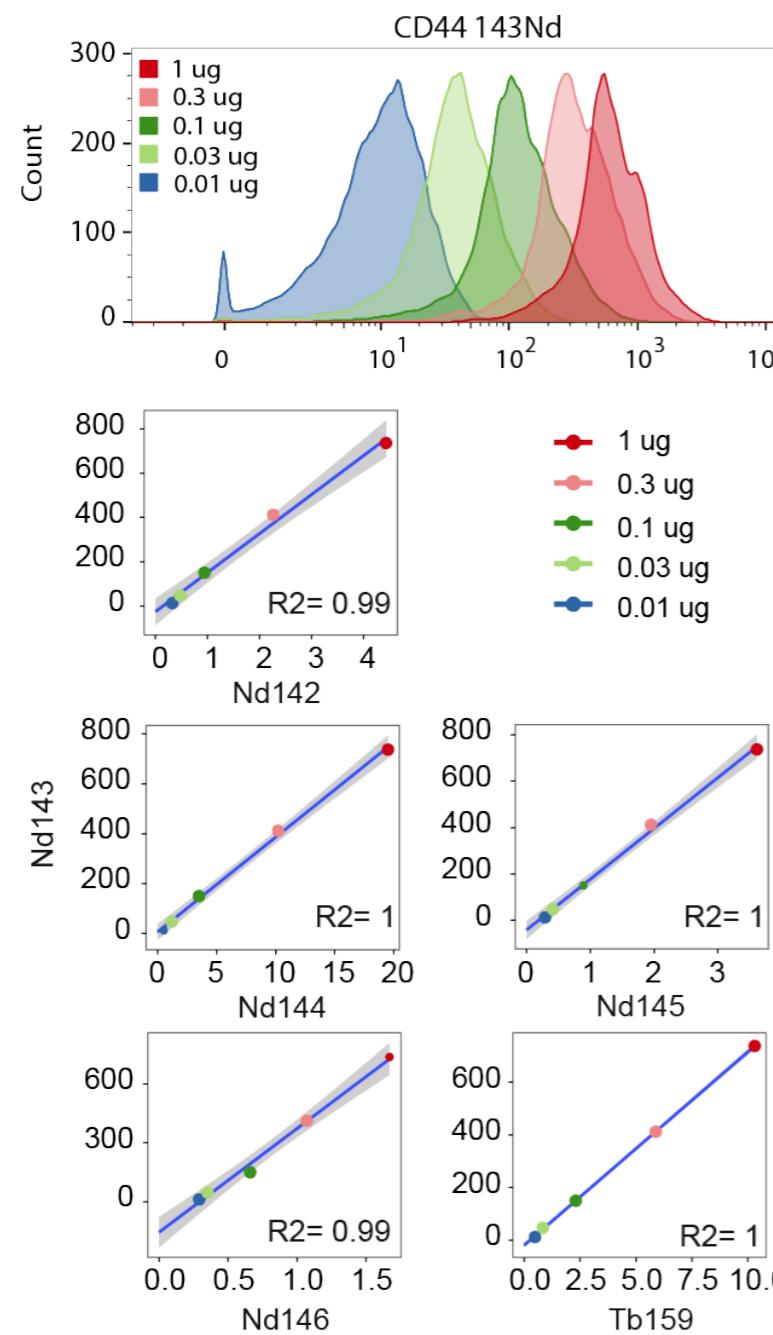
**abundance sensitivity:**

imprecision in ion detection ( $\pm 1M$ )

**oxide formation:**

dependent on plasma temperature & ion-oxygen bond strength ( $+16M$ )

# PoC: Spillover is a linear phenomenon



PBMC stained with increasing [CD44–Nd143]



linear model fitted with  $R > 0.99$

*predictor* = channel stained for, *response* = spill-effected channel



as in fluorescent flow cytometry,  
the linear regression slope may be used to estimate spill

# Compensation via single-staining controls

$$I_{j,observed} = I_{j,real} + \sum_{i \in w_j} s_{ij}$$

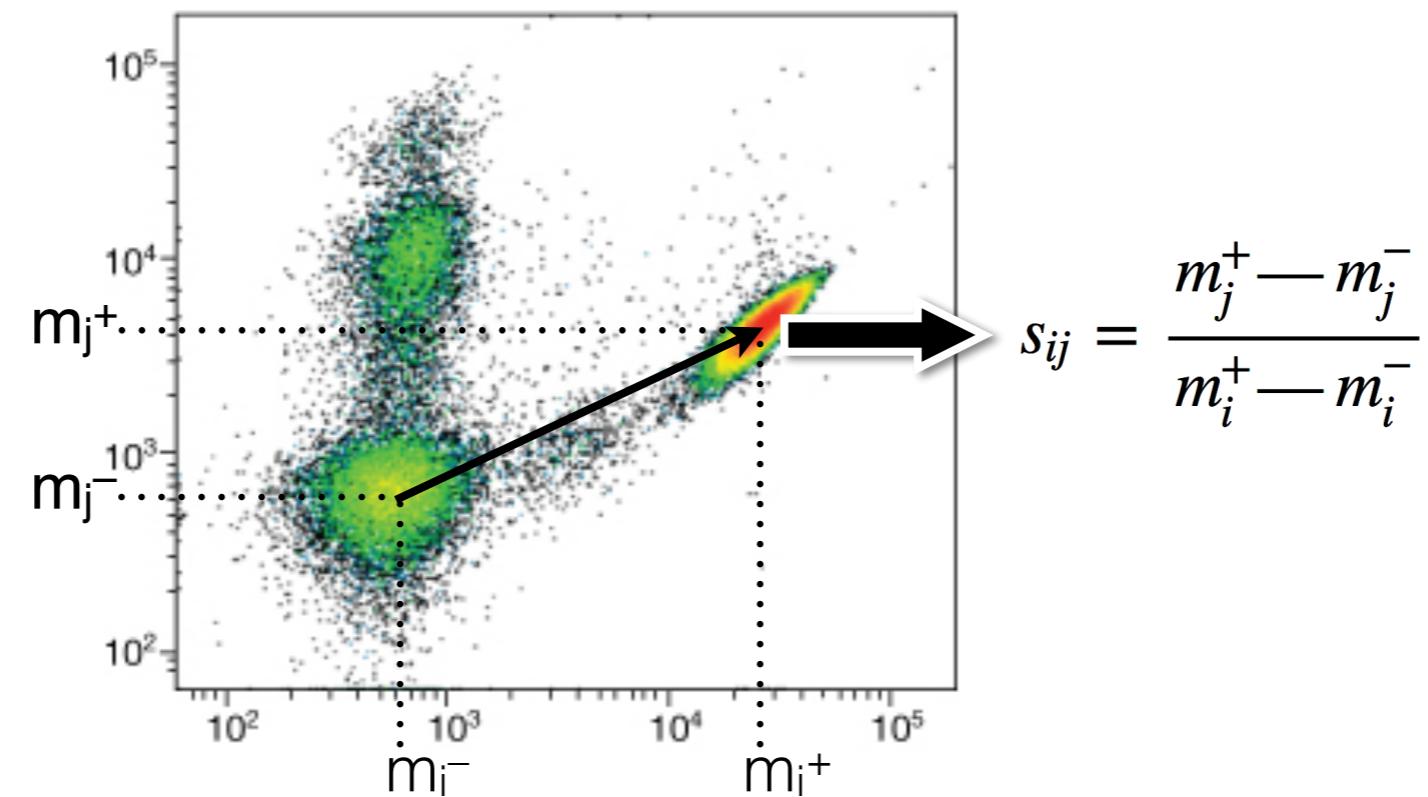
proportion of signal in channel i  
that is due to channel j signal

set of channels that spill into channel j

$$I_{observed} = I_{real} \cdot SM \longrightarrow \text{spillover matrix}$$

$$I_{real} = I_{observed} \cdot SM^{-1}$$

$$= I_{observed} \cdot CM \longrightarrow \text{compensation matrix}$$



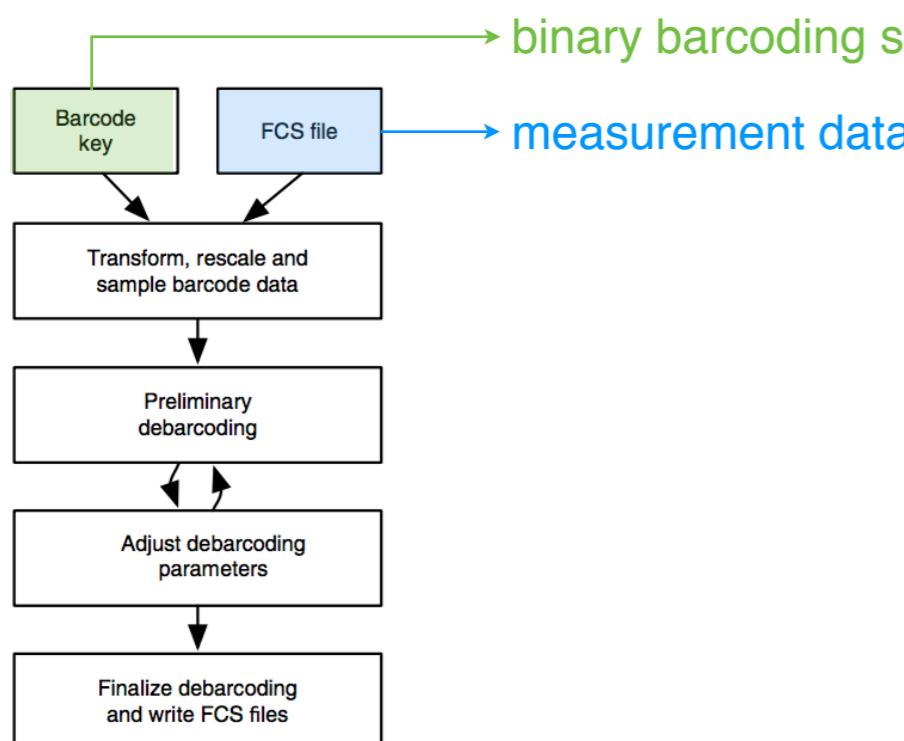
# Single-cell deconvolution

Nature Protocols 2015

## Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm

Eli R Zunder<sup>1,5</sup>, Rachel Finck<sup>1,5</sup>, Gregory K Behbehani<sup>1,2</sup>, El-ad D Amir<sup>3</sup>, Smita Krishnaswamy<sup>3</sup>, Veronica D Gonzalez<sup>1</sup>, Cynthia G Lorang<sup>1</sup>, Zach Bjornson<sup>1</sup>, Matthew H Spitzer<sup>1</sup>, Bernd Bodenmiller<sup>1,4</sup>, Wendy J Fantl<sup>1</sup>, Dana Pe'er<sup>3</sup> & Garry P Nolan<sup>1</sup>

### 1. preliminary assignment



##	139	141	142	143	144	145	146	147	148	149
## 139	1	0	0	0	0	0	0	0	0	0
## 141	0	1	0	0	0	0	0	0	0	0
## 142	0	0	1	0	0	0	0	0	0	0
## 143	0	0	0	1	0	0	0	0	0	0
## 144	0	0	0	0	1	0	0	0	0	0

Events	Measurement intensities										
	La139Di	Ce140Di	Pr141Di	Nd142Di	Nd143Di	Nd144Di	Nd145Di	Nd146Di	Sm147Di	Nd148Di	Sm149Di
[1,]	0.000000	0	0	0.000000	0.0000	0.000000	0.000000	0.000000	0.000000	0 0.000000	0.245174
[2,]	1.380103	0	0	0.000000	0.0000	7.378455	548.767761	27.167898	0 0.000000	0.000000	
[3,]	0.000000	0	0	1.297663	0.0000	4.220285	532.744995	14.976744	0 0.000000	0.000000	
[4,]	0.000000	0	0	0.000000	0.0000	0.000000	0.277233	0.000000	0 0.000000	0.000000	
[5,]	0.000000	0	0	0.000000	0.0000	0.000000	0.000000	0.000000	0 0.000000	0.000000	
[6,]	0.000000	0	0	9.526143	931.7335	24.398716	1.625141	1.522033	0 0.895293	0.000000	

143

145

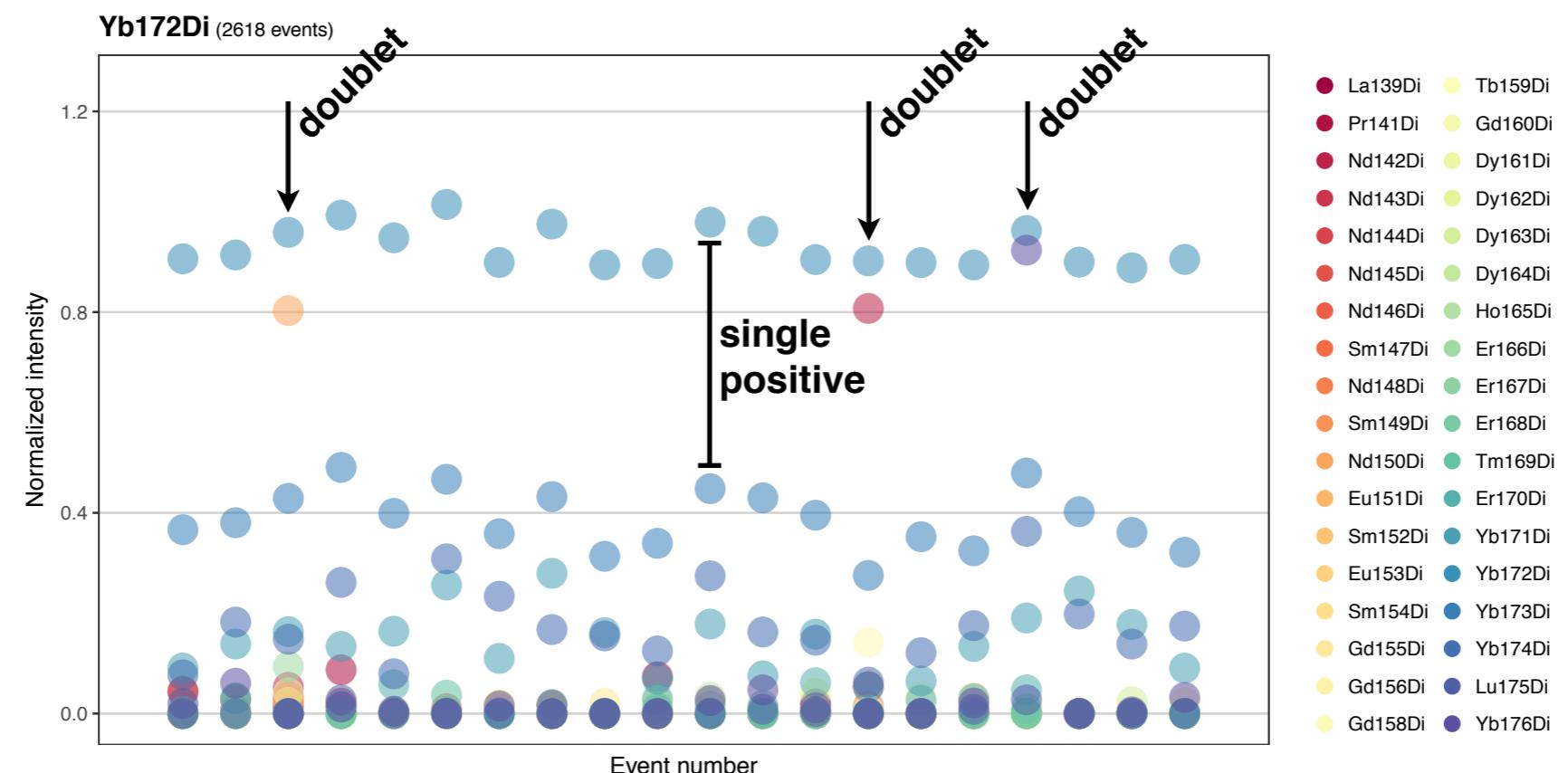
# Single-cell deconvolution

## Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm

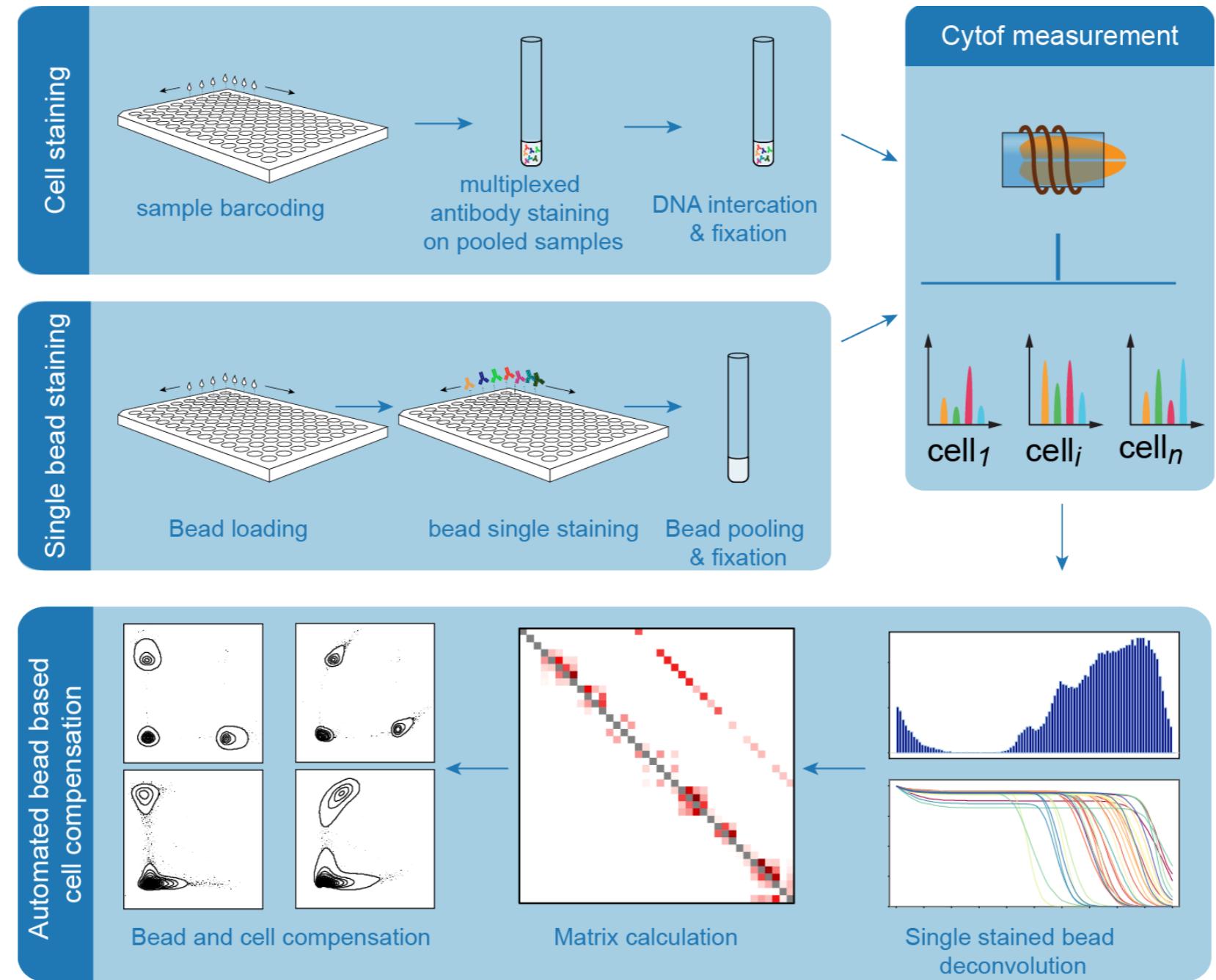
Eli R Zunder<sup>1,5</sup>, Rachel Finck<sup>1,5</sup>, Gregory K Behbehani<sup>1,2</sup>, El-ad D Amir<sup>3</sup>, Smita Krishnaswamy<sup>3</sup>, Veronica D Gonzalez<sup>1</sup>, Cynthia G Lorang<sup>1</sup>, Zach Bjornson<sup>1</sup>, Matthew H Spitzer<sup>1</sup>, Bernd Bodenmiller<sup>1,4</sup>, Wendy J Fantl<sup>1</sup>, Dana Pe'er<sup>3</sup> & Garry P Nolan<sup>1</sup>

1. preliminary assignment

2. doublet-removal



# SCD-based compensation work flow



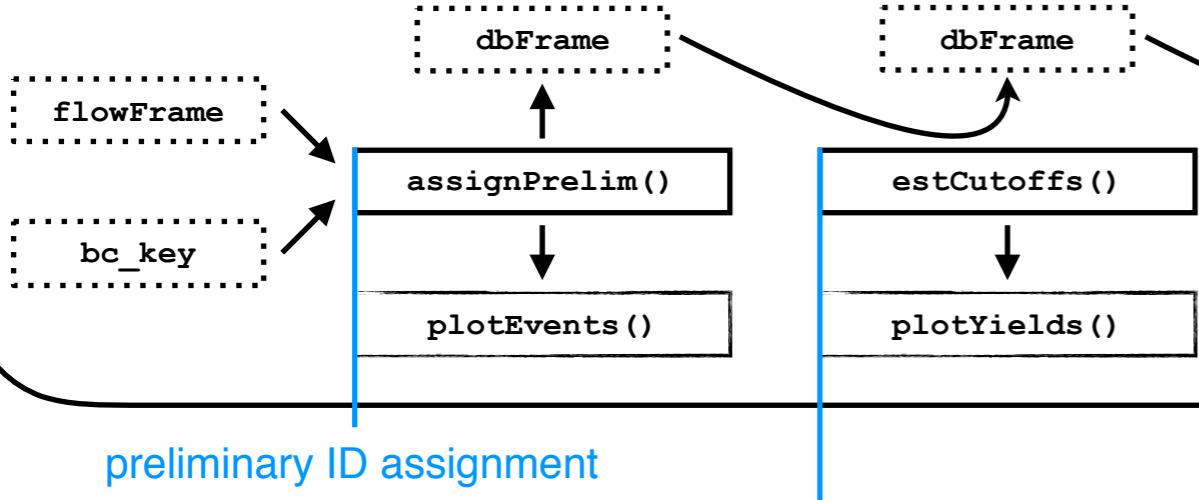
# R Implementation

```

bc_ms <- c(139, 141:156, 158:176)      # specify bc_key
re <- assignPrelim(x = ss_beads, y = bc_ms) # assign preliminary IDs
re <- estCutoffs(x = re)                   # estimate distance cutoffs
re <- applyCutoffs(x = re)                 # apply cutoffs
compMat <- computeCompmat(x = re)          # estimate CM

```

## DEBARCODING

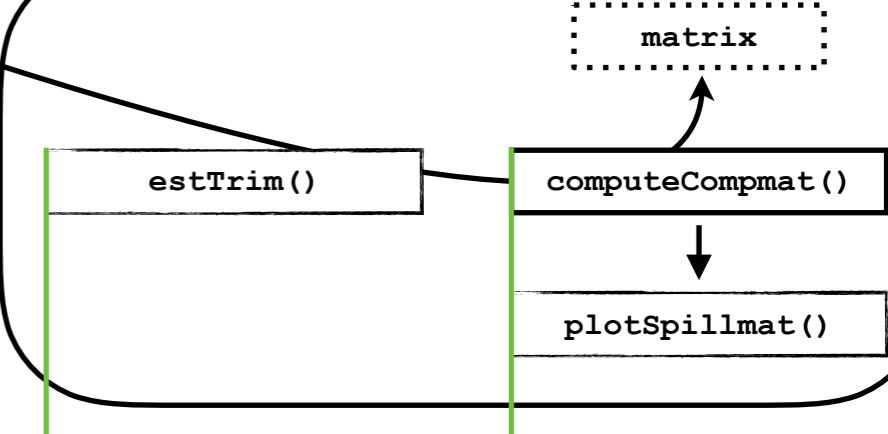


preliminary ID assignment

estimation of separation cutoffs

application of thresholds  
to arrive at final assignment

## COMPENSATION



estimation of optimal trim value

estimation of spill values from  
single positive populations

# The dbFrame class

## slots

exprs

bc\_key

bc\_ids

deltas

normed\_bcs

sep\_cutoffs

mhl\_cutoff

counts

yields

measurement data  
 binary barcoding scheme  
 event assignments  
 barcode separations  
 normalized intensities  
 barcode separation thresholds  
 Mahalanobis distance threshold  
 distribution of barcode separations  
 yields upon debarcoding

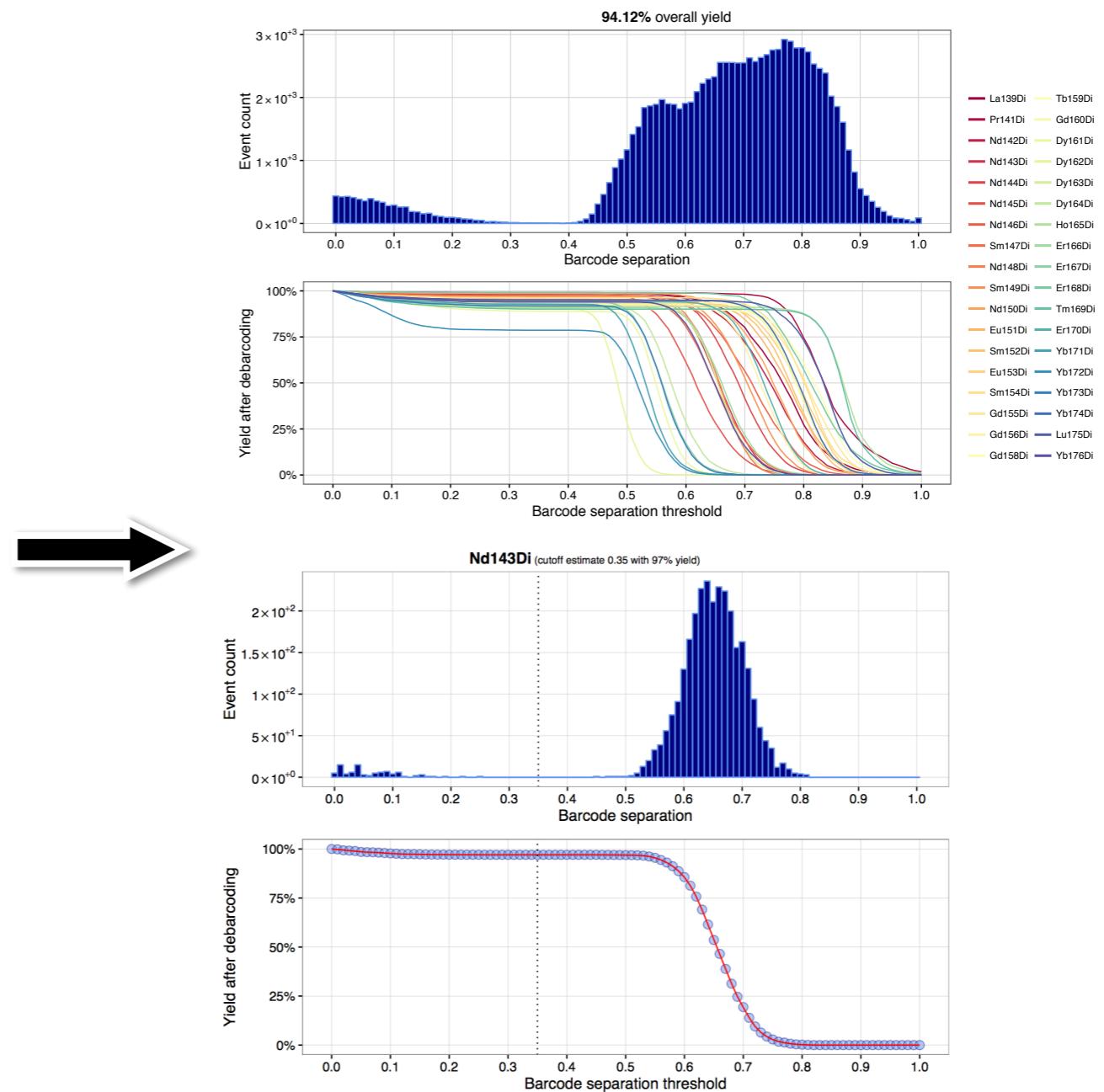
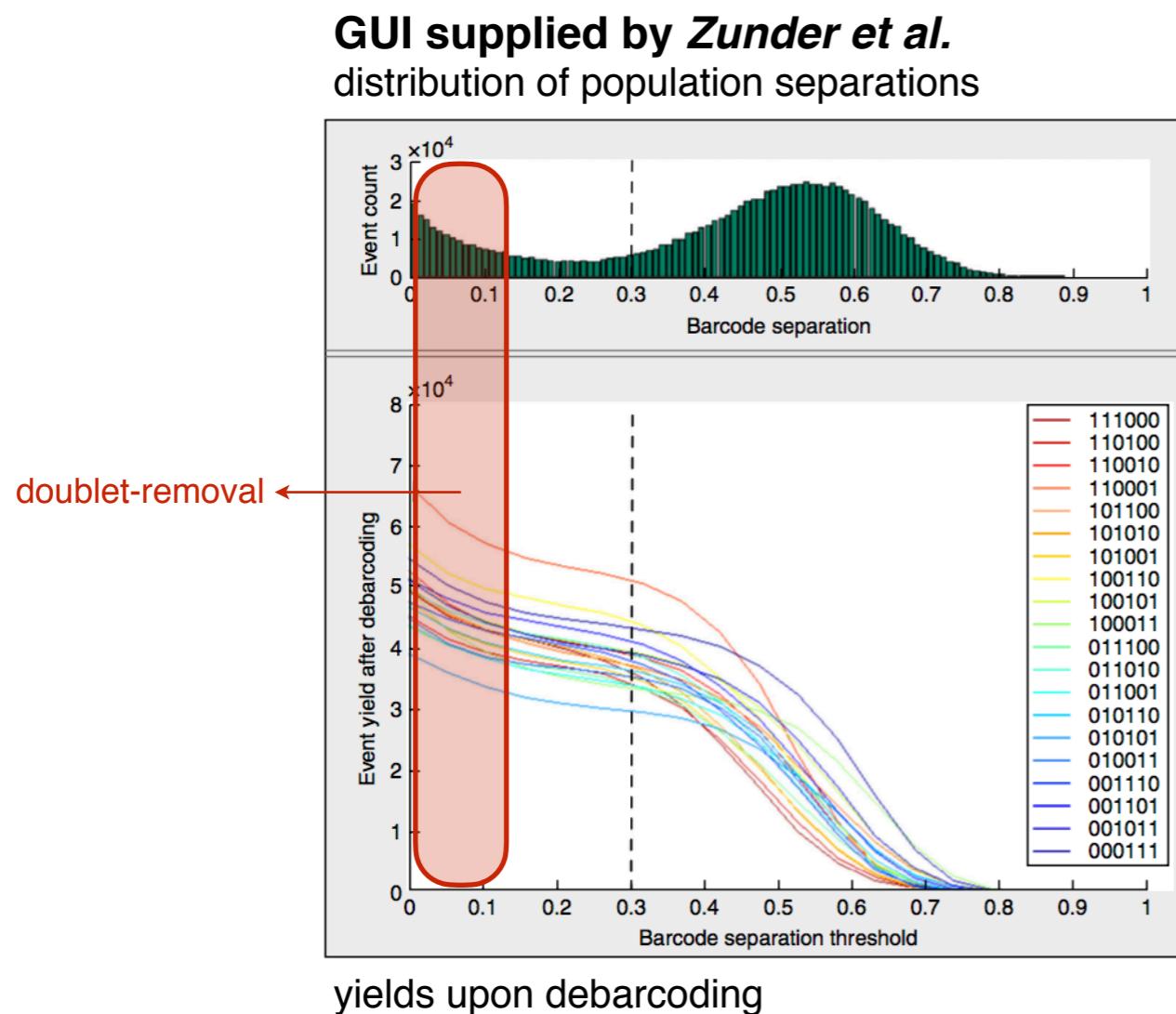
```
assignPrelim()
applyCutoffs()
```

```
estCutoffs()
```

```
## dbFrame objectect with
## 10000 events, 61 observables and 36 barcodes:
##
## Current assignments:
##   0 events unassigned
##   ID    162 147 169 151 172 174 146 159 144 152 161 171 165 166 173 143 176
##   Count 382 358 356 337 332 327 326 325 319 317 308 302 301 301 298 298
##
##   ID    148 149 145 153 175 150 164 160 167 163 155 158 141 156 170 154 142
##   Count 292 287 280 277 275 268 267 264 263 260 252 236 228 225 216 215 180
##
##   ID    139 168
##   Count 128 99
##
## Separation cutoffs:
##   ID    139 141 142 143 144 145 146 147 148 149 150 151 152
##   Yield 0.12 0.14 0.19 0.14 0.19 0.16 0.17 0.23 0.09 0.25 0.08 0.38 0.21
##
##   ID    153 154 155 156 158 159 160 161 162 163 164 165 166
##   Yield 0.31 0.14 0.19 0.26 0.25 0.40 0.09 0.20 0.35 0.09 0.19 0.28 0.20
##
##   ID    167 168 169 170 171 172 173 174 175 176
##   Yield 0.13 0.55 0.17 0.14 0.10 0.23 0.18 0.28 0.10 0.38
##
## Yields upon debarcoding:
##   94.66% overall yield
##   ID    139    141    142    143    144    145    146    147    148
##   Yield 97.66% 99.12% 98.89% 96.98% 99.37% 99.29% 97.24% 96.65% 97.26%
##
##   ID    149    150    151    152    153    154    155    156    158
##   Yield 96.17% 98.51% 93.18% 94.64% 93.14% 97.21% 94.84% 92.44% 96.19%
##
##   ID    159    160    161    162    163    164    165    166    167    168
##   Yield 88.92% 96.59% 87.34% 89.27% 96.92% 94.01% 91.03% 91.69% 95.06% 100%
##
##   ID    169    170    171    172    173    174    175    176
##   Yield 92.42% 93.98% 95.03% 76.81% 93.36% 94.5% 98.91% 93.29%
```

# plotYields

## Optimizing distance separation cutoffs

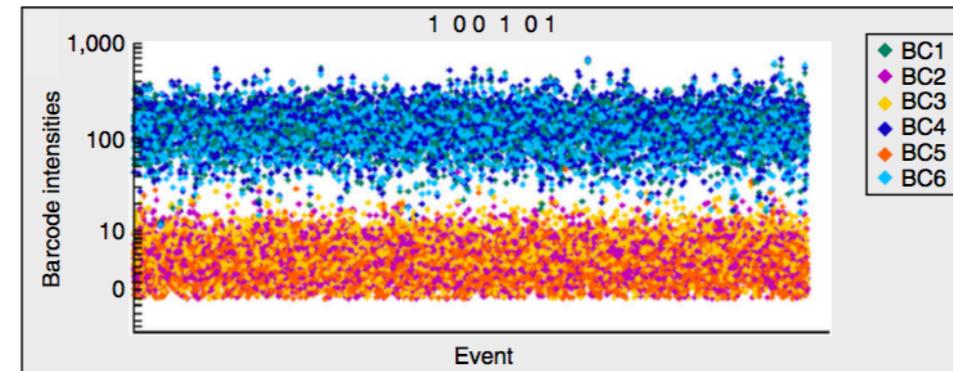


# plotEvents

## Barcode intensities normalized by population

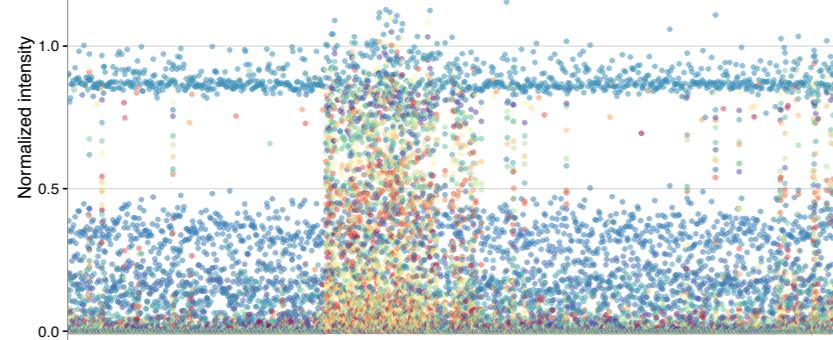
**GUI supplied by Zunder et al.**

Normalized barcode intensities



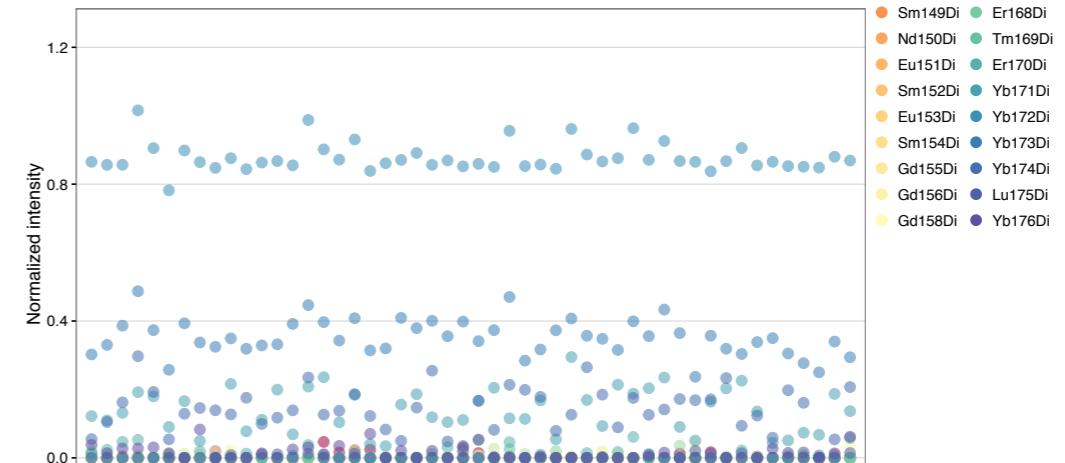
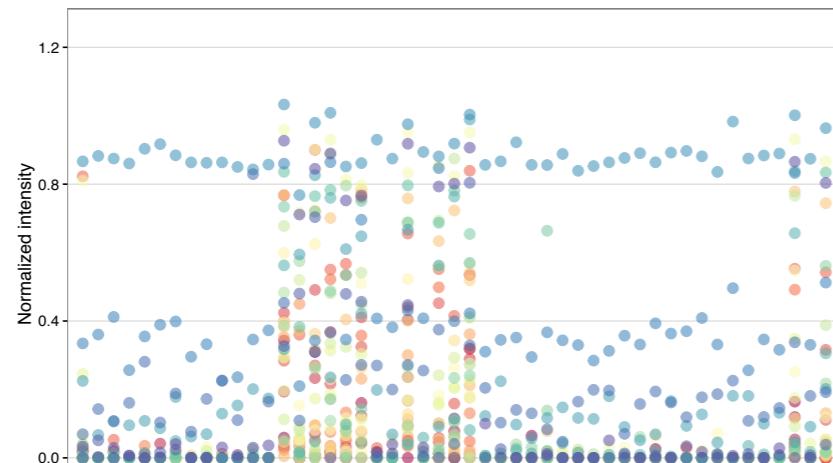
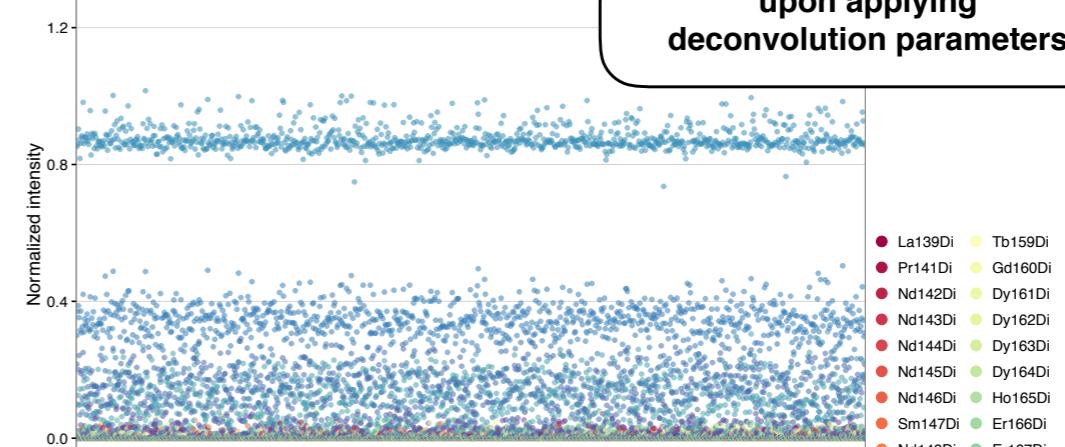
assignments returned by  
assignPrelim()

Yb172Di (3215 events)



Yb172Di (1870 events)

upon applying  
deconvolution parameters



# plotSpillmat

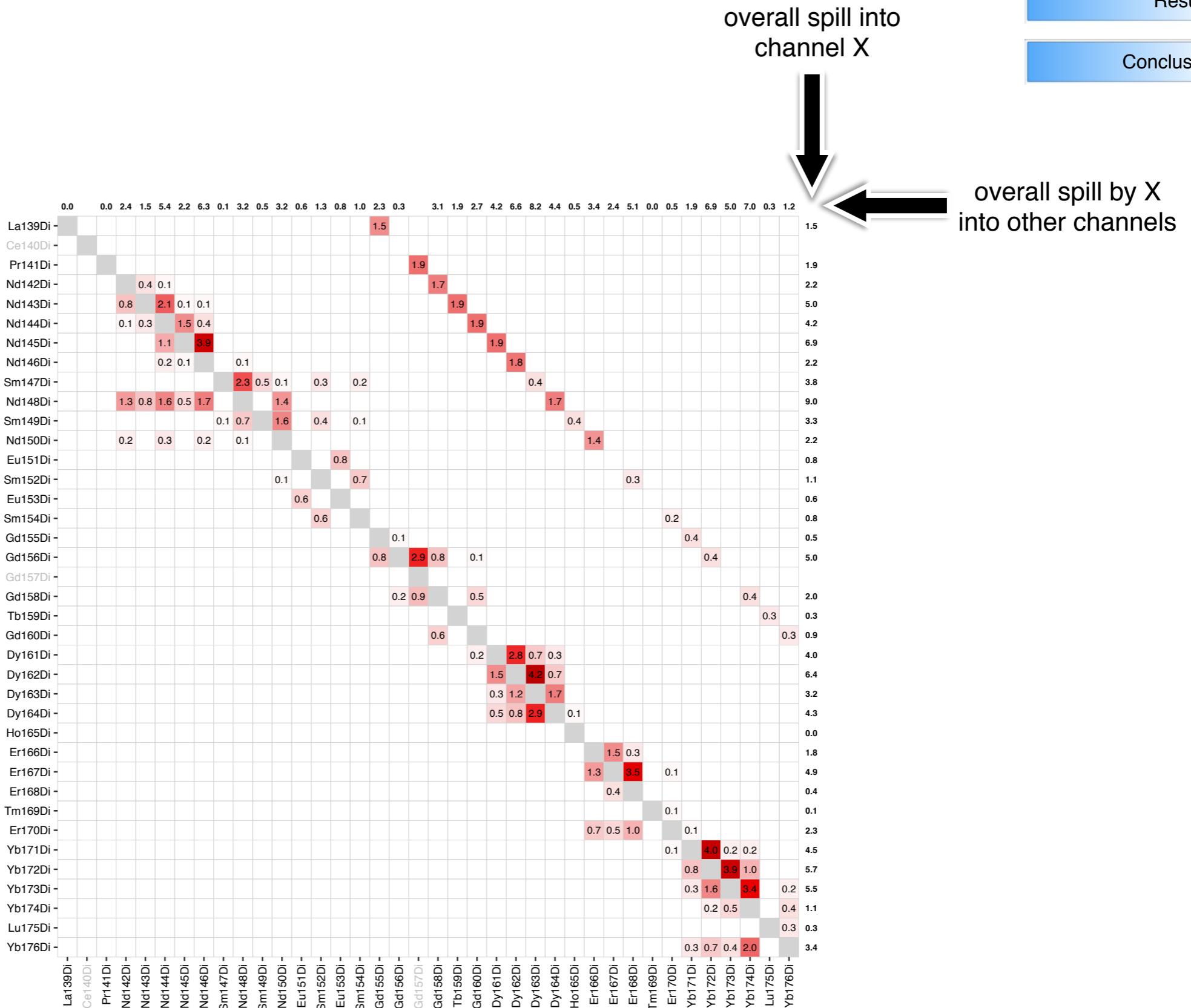
## Spillover matrix heat map representation

## Introduction

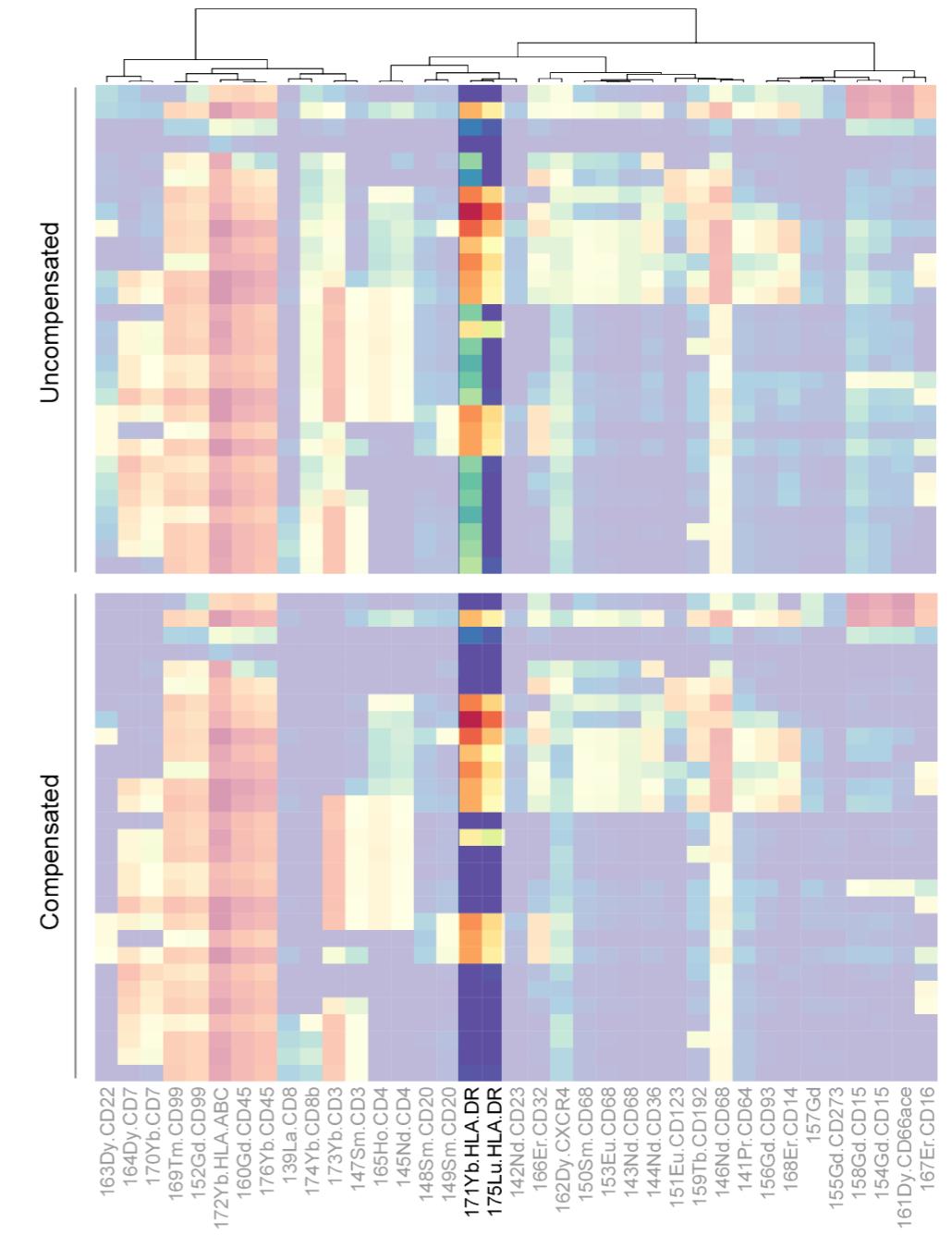
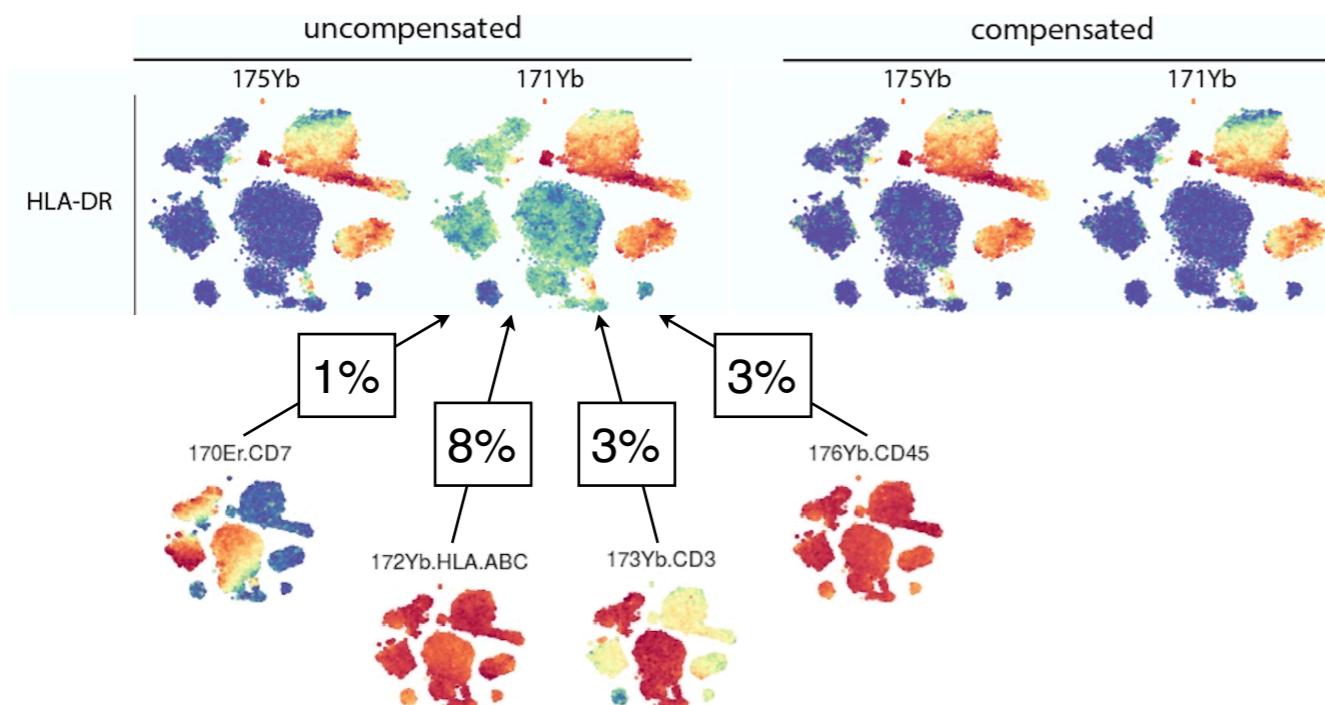
# CATALYST

## Results

## Conclusion



# Correction of spill artefacts on a 36ab panel



# In conclusion...

- R implementation of *Zunder et al.*'s SCD-algorithm including **automated estimation of population-specific separation cutoffs** ([Shiny app in the making!](#))
- as in fluorescent flow cytometry, **spillover is linear** and **can be estimated using single-color controls**
- removal of spill-artefacts will effect downstream analysis (e.g. increase correlation between the same antibody measured in different channels)

# Acknowledgement



**Prof Dr Bernd Bodenmiller**

Assistant Professor of  
Quantitative Biology



**Prof Dr Mark D Robinson**

Assistant Professor of  
Statistical Genomics



**Stéphane Chevrier, PhD**  
Postdoc Bodenmiller group



**Vito Zanotelli, MSc**  
PhD Student Bodenmiller group