High-dimensional single cell analysis by mass cytometry - introduction and call for collaboration

Vinko Tosevski, Ph.D.
Head of Mass Cytometry Facility
University of Zurich
Disclaimer

• I declare no conflict of interest
  – Products
  – Services
  – ...

University of Zurich

Mass Cytometry Facility
Hypothesis

- You do single cells analysis
- You are open to new technologies that can advance your research
- You didn’t come across mass cytometry before
Outline

• What is mass cytometry?
• What can be done with it?
• What does it take to start doing it?
• How can we assist you in doing it?
  – Who are “we”?
What is mass cytometry?

• Novel high-dimensional analysis technology that allows highly multiplexed measurement of protein and RNA content at the single cell level

• Novel
  – 1st prototype published in 2009
  – Currently around 150 instruments worldwide

• High-dimensional
  – 130 parameters (measurement channels) *possible*
What is mass cytometry?


Dmitry R. Bandura*, Vladimir I. Baranov†, Olga I. Ormatsky†, Alexei Antonov‡, Robert Kinach‡, Xudong Lou†, Sergei Pavlov*, Sergey Vorobiev*, John E. Dick§ and Scott D. Tanner‡
Department of Chemistry, University of Toronto, 80 St. George Street, Toronto, Ontario M5S 3H6, Canada, DVS Sciences, Inc., 70 Peninsula Crescent, Richmond Hill, Ontario L4S 1Z5, Canada, and University Health Network, Toronto Medical Discovery Tower, 101 College Street, Toronto, Ontario M5G 1L7, Canada

DOI: 10.1021/ac901049w
Publication Date (Web): July 14, 2009
Copyright © 2009 American Chemical Society

* Corresponding author; Phone: +1 416 946 8420. Fax: +1 416 978 4317. E-mail: Dmitry.Bandura@utoronto.ca, † University of Toronto, ‡ DVS Sciences, Inc. & University Health Network.
What is mass cytometry?

RESEARCH ARTICLE

Single-Cell Mass Cytometry of Differential Immune and Drug Responses Across a Human Hematopoietic Continuum

Sean C. Bendall1,*, Erin F. Simonds1,*, Peng Qiu2, El-ad D. Amir3, Peter O. Krutzik1, Rachel Finck1, Robert V. Bruggner1,7, Rachel Melamed3, Angelica Trejo1, Olga I. Ornatsky4,5, Robert S. Balderas6, Sylvia K. Plevritis2, Karen Sachs1, Dana Pe'er3, Scott D. Tanner4,5, Garry P. Nolan1,†

†To whom correspondence can be addressed. E-mail: gnolan@stanford.edu

* These authors contributed equally to this work.

Science. 06 May 2011; Vol. 332, Issue 6030, pp. 687-696
DOI: 10.1126/science.1193704
What is mass cytometry?

• Protein and RNA content
  – Antibody-based (proteins)
  – Nucleotide probe-based (RNA)
CyTOF = mass cytometry

• CyTOF - Cytometry by Time of Flight

• Currently, high end flow cytometers can measure up to 20 parameters simultaneously
  – FACSSymphony just came out with 30 detectors installed, but instrument performance and fluorochrome situation is unclear at the moment.

• CyTOF currently measures around 50 parameters!
CyTOF

• How was this leap in dimensionality achieved?
  – Instead of fluorochromes, antibodies are tagged with stable metal isotopes (Lanthanides)

Source: BD, Fluidigm
Possibly useful elements

1. Non-rare
2. Non-biological
3. Non-radioactive

Source: Fluidigm
Pros and cons

- A lot of limitations stems from the fact that ions are measured the way they are

- **Pros**
  - Many channels
  - Lower «autofluorescence –like» background

- **Cons**
  - Slower (cells/sec)
  - Less sample transmission efficiency
  - Less per-channel sensitivity (not really a problem)
Pros and cons

• Cons
  – …
  – Dependence on Abs (dominantly but not exclusively, i.e. cell cycle or proliferation analyses work beautifully with IdU directly).
## Side-by-side comparison


### Table 1. Comparison of utility and performance of state of the art commercial fluorescence flow cytometry and mass cytometry single-cell analysis platforms.

<table>
<thead>
<tr>
<th>Technology</th>
<th>Fluorescence flow cytometry</th>
<th>Mass cytometry</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Measurement basis</strong></td>
<td>Fluorescent probes</td>
<td>Stable mass isotope probes</td>
</tr>
<tr>
<td><strong>Experimental design</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max no. of measurements</td>
<td>20 (18 fluorescence)</td>
<td>37 (including DNA)</td>
</tr>
<tr>
<td>Theoretical no. of subsets(^a)</td>
<td>(2.6 \times 10^5)</td>
<td>(1.4 \times 10^{11})</td>
</tr>
<tr>
<td>Panel design complexity (no. of probes)</td>
<td>Easy</td>
<td>(&lt;8)</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Hard</td>
<td>8–12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12–18</td>
</tr>
<tr>
<td>Sensitivity range for different probes(^b)</td>
<td>Easy</td>
<td>0.1–10</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hard</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sample throughput</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sampling efficiency</td>
<td>&gt; 95%</td>
<td>&lt; 30%</td>
</tr>
<tr>
<td>Measured cells/s</td>
<td>25 000</td>
<td>500–1000</td>
</tr>
<tr>
<td>Cells/h</td>
<td>25–60 million</td>
<td>2 million</td>
</tr>
<tr>
<td><strong>Commercial reagent cost</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per probe per test(^c)</td>
<td>$2.00–$8.00</td>
<td>$1.50–$3.00</td>
</tr>
</tbody>
</table>

\(^a\)Theoretical number of subsets is the number of distinct cell types determinable, assuming only on or off for each marker; that is, \(2^{\text{colors}}\).

\(^b\)Sensitivity range is in arbitrary units, and compares the rough sensitivity for different probes (fluorescence or ICP-MS) to detect a given epitope on a cell by immunophenotyping.

\(^c\)Estimated based on the price of commercially conjugated reagents or unconjugated antibodies and commercial conjugation kits.
CyTOF

Source: http://pubs.acs.org/action/showimage?doi=10.1021%2Fac901049w&iName=master.img-000.png&type=master
Sample introduction and ionization

- A lot of cells get lost on the way to the detector in the tubing and the glassware.
- CyTOF2 had 30% to 40% cell transmission efficiency (i.e. you have 100K cells in the tube and end up with 30K to 40K cells on the FCS file).
- Helios (3rd generation CyTOF instrument) has 50% to 70%
- Caveat! 75% on beads translates into 50% on hPBMCs (in our hands)
Sample introduction

Source: Fluidigm
Sample introduction and ionization

Source: Fluidigm
Sample introduction and ionization

Source: Fluidigm
The vacuum interface includes the three nickel interface cones: sampler (red), skimmer (blue) and reducer (green).
Processing of the ion cloud

Source: Fluidigm
Ion separation and detection

Source: Fluidigm
Ion separation and detection

Push-out plate: ions pushed into TOF chamber at 13 µsec intervals (“pushes”)

Source: Fluidigm
Push

- **Pushes**: 13 µsec slice of the ion stream entering the ToF chamber.
- **TOF**: Separation of ions **within a push** by mass.

<table>
<thead>
<tr>
<th>TOF in ns →</th>
<th>9,000</th>
<th>9,500</th>
<th>10,000</th>
<th>10,500</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image" alt="Red Circle" /></td>
<td><img src="image" alt="Orange Circle" /></td>
<td><img src="image" alt="Green Circle" /></td>
<td><img src="image" alt="Blue Circle" /></td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Red Circle" /></td>
<td><img src="image" alt="Orange Circle" /></td>
<td><img src="image" alt="Green Circle" /></td>
<td><img src="image" alt="Blue Circle" /></td>
</tr>
</tbody>
</table>

Source: Fluidigm
From cell to ion cloud

well-spaced → diffusion → individual clouds

Metals + Markers

Source: Fluidigm

University of Zurich UZH
Mass Cytometry Facility
From cell to ion cloud

Source: Fluidigm

too close

diffusion

cloud fusion

Example: concentrated europium beads giving a larger cell length value
Event duration is the period of time, measured in pushes, that the ion signal intensity is above the threshold.

Any events outside the range of 10 to 150 pushes are excluded from conversion to the FCS file data set.
Raw data – rain plot

Source: Ornatsky et al., J Immunol Methods 2010
FCS file in the end

Application highlights
Aplications

- No sorting or calcium influx 😊
- Extensive phenotyping and functional profiling
  - Immunophenotype
  - Signaling state
  - Cytokine/chemokine expression
  - Health & Viability
  - Proliferation
  - Apoptosis
  - ...
Seminal work


- 13 „core“ lineage markers, 18 subset-specific surface markers, 18 intracellular epitopes, 13 ex-vivo stimulation conditions.

- Huge scope of the results (929 pages of SOM)
Closer to home

Becher et al.
Around the same time

The mass cytometry toolbox
The challenge?

• Let’s CyTOF!
• Most common concerns
  – Panel assembly (know-how and price)
  – Barcoding yes/no?
  – Data analysis
Price of a mass cytometry pilot experiment

• No stocks of metal-labeled antibodies around
• Off-the-shelf chemicals/buffers often weren’t good enough for elemental analysis
• Whole labs not suited for elemental analysis!
• Cyto 2014 – CyTOF UGM
  Jared K. Burks
  MD Anderson, TX, USA
• Reagent repository!
Reagent repository

• 100+ metal-conjugated antibodies (Ms, Hu)
• All (!) metal conjugation kits
• All the buffers required to perform the standard surface, ICS and nuclear staining.
• Additional reagents
  – Intercalator(s), Dead cell reagent(s)
    • Iridium, Rhodium, cisplatin
  – Normalization beads
    • EQ four element beads

• The aim – provide the means for «the quick test» if mass cytometry works for you without breaking the budget.
<table>
<thead>
<tr>
<th>Reagent repository</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ly6G</td>
</tr>
<tr>
<td>CD11c</td>
</tr>
<tr>
<td>CD115</td>
</tr>
<tr>
<td>CD69</td>
</tr>
<tr>
<td>CD25</td>
</tr>
<tr>
<td>CD3e</td>
</tr>
<tr>
<td>CD335, NKp46</td>
</tr>
<tr>
<td>CD62L</td>
</tr>
<tr>
<td>CCR7</td>
</tr>
<tr>
<td>CD8a</td>
</tr>
<tr>
<td>TCRβ</td>
</tr>
<tr>
<td>NK1.1</td>
</tr>
<tr>
<td>B220</td>
</tr>
<tr>
<td>CD45</td>
</tr>
<tr>
<td>CD11b (MAC1)</td>
</tr>
<tr>
<td>CD19</td>
</tr>
<tr>
<td>Ly6C</td>
</tr>
<tr>
<td>CD44</td>
</tr>
<tr>
<td>CD4</td>
</tr>
<tr>
<td>I-A/I-E</td>
</tr>
</tbody>
</table>

+ for 4 samples ≈ 150 CHF
Know-how

• By now, quite some support information out there (Fluidigm, peer-reviewed journals, UGMs,...)
• MaxPar reagents are a good start
• On average, conceiving and validating the panel not as hard as generally perceived (depends on the application!)
• Panel building takes from 0 to ∞
• Having deep experience with FC helps!
Instrumentation know-how

- Howard Shapiro, Cerberus
Barcoding

- Barcoding really found its place in mass cytometry – should be considered as well as means of to reduce the batch effects and improve throughput.
Barcoding

• Mass cytometry barcoding (MCB).
• Instead of NHS-functionalized fluorochromes (FCB), MCB used mDOTA (maleimido-mono-amide-DOTA).

Barcoding

# Barcoding

## Table

<table>
<thead>
<tr>
<th>Barcode number</th>
<th>BC1</th>
<th>BC2</th>
<th>BC3</th>
<th>BC4</th>
<th>BC5</th>
<th>BC6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Diagram

- **Barcode doublet**
- **Barcode singlet**

### Mass Cytometry Facility

**University of Zurich (UZH)**
Data analysis

• The biggest perceived challenge of all!

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Plots</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>36</td>
</tr>
<tr>
<td>32</td>
<td>496</td>
</tr>
</tbody>
</table>

University of Zurich

Mass Cytometry Facility
Data analysis

- “Unfortunately, the use of three or more independent fluorescent parameters complicates the analysis of the resulting data significantly.” *Murphy, Cytometry (1985)*
- Does one have to become a bioinformatician to analyze CyTOF data?
- «Hypothesis-driven analysis» vs. «Exploratory data analysis»
Tools for automated data analysis

- Bioconductor
- FlowJo
- Cytobank
- GemStone
- Cyt
Bioconductor

- http://bioconductor.org
- BioConductor provides R software modules for biological and clinical data analysis
- A scripted approach to high throughput data analysis
  - Non-interactive, reproducible
  - Breaks problem into smaller pieces (packages)
  - Modules can plug-in & swap-out
Why R?

• Big community
• A lot of available packages
• Well suited for developing new workflows
  – Interactive mode vs. batch mode
• Designed with data analysis in mind
• New algorithms will often be made available in R by the authors *right away*
FlowJo

- Historically, FlowJo offered some basic integration with few popular R packages
FlowJo Exchange

- [https://github.com/FlowJo](https://github.com/FlowJo)
- FlowJo Exchange houses scripts and plugins for FlowJo that anyone can contribute to or download from.
- New version 10.1 promised “Multiple improvements which allow any algorithm to be plugged into FlowJo.”
- At the beginning 5 scripts on the exchange - Indexed Sorting, Basic Looping, Collection Order Sort, Control Based Inclusion Gating, and a Relative Gate.
FlowJo v9

- [Link to FlowJo v9 documentation](http://docs.flowjo.com/v9/flowjo-v9-documentation-home/platforms/t-sne/)
Cytobank with SPADE, viSNE, CITRUS


SPADE

- Clustering algorithm

1. Cytometry data
2. Density-dependent down-sampling
3. Down-sampled data
4. Agglomerative clustering
5. Clustering result
6. Minimum spanning tree construction
7. SPADE tree
8. Up-sampling
9. Colored tree showing cellular heterogeneity
viSNE (t-SNE, bh SNE)

- Dimensionality reduction algorithm

viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia

El-ad David Amir¹, Kara L Davis²,³, Michelle D Tadmor¹,³, Erin F Simonds²,³, Jacob H Levine¹,³, Sean C Bendall²,³, Daniel K Shenfeld¹,³, Smita Krishnaswamy¹, Garry P Nolan²,⁴ & Dana Pe'er¹,⁴
viSNE (t-SNE, bh SNE)

1. Calculate high-dimensional similarity matrix

2. Reduce to 2D matrix, preserve multi-D relationships as distance

3. viSNE map of cell types
   A) 2D view of cells that shows multi-D similarity / difference using distance
   B) Can show protein expression as heat

- t-SNE: a unitless axis where distance represents multi-D phenotype difference
viSNE (t-SNE, bh SNE)

• Visualization of high-dimensional single-cell data in 2D
• The resulting map provides a visual representation of the single-cell data where the positions of cells reflects their proximity in high-dimensional space.
• Color can be utilized as a third dimension to interactively visualize features of these cells.
CITRUS


• “cluster identification, characterization, and regression”
CITRUS

● Cells from N samples are combined and clustered in a semi(un)supervised manner to automatically identify C clusters of related cells.
● Descriptive statistics characterizing various properties of each cluster (cluster features) are extracted on a per-sample basis.
● Extracted cluster features are used in conjunction with a user-specified endpoint of interest to train a supervised model.
● Internal cross-validation is used to evaluate model fit and select the appropriate regularization threshold for a final model.
● Model features are plotted as a function of endpoint of interest and cluster phenotypes are determined by density plots of markers used for clustering.
Rather than cutting the dendrogram at a fixed height to identify clusters, all clusters $C$ in the hierarchy of merged clusters larger than a user-specified size are retained for subsequent analysis.
CITRUS

GemStone

- Uses continuous expression patterns of various parameters and employs probability state modeling to organize and visualize cell populations relative to one another.
- While it still requires a priori knowledge of the relationship between at least some of the markers measured, it still visually summarizes all cells in a given sample and can reveal cell subsets and relationships that other tools may not.
Analysis of CD8 T cells showing progression of phenotypic markers including branching expression of markers like CD57.
CYT is an interactive visualization tool designed for the analysis of high-dimensional mass or flow cytometry data. The tool encompasses multiple computational features (viSNE, Wanderlust, PhenoGraph and more).
• viSNE: see before
• PhenoGraph: computing PhenoGraph clusters. The clusters can be visualized by a marker expression heatmap or the cluster centroids can be used to generate a tSNE map to be gated on or overlaid with other markers.
• Wanderlust: computing a wanderlust trajectory (a nonlinear principal component or ordering of the data). Cyt can visualize the average expression of markers as a function of the wanderlust trajectory (or any desired marker).
• Wishbone: to align single cells from differentiation systems with bifurcating branches.
• cyt also implements some basic data analysis techniques such as PCA, kMeans, EMGM, and more.
Current organization and scope of service

• We are able to support the entire hypothetical workflow based on mass cytometry!

• Project
  – Panel design
  – Sample preparation
  – Troubleshooting/optimization

• Reagents
  – Reagent repository (Abs, buffers, conjugation kits)

• Instrument
  – Sample acquisition
  – Maintenance
  – Repairs
  – QC

• Data analysis
  – Preprocessing
  – Exploratory data analysis
Current organization and scope of service

- Project
  - Panel design
  - Sample preparation
  - Troubleshooting/optimization
- Reagents
  - Reagent repository (Abs, buffers, conjugation kits)
- Instrument
  - Sample acquisition
  - Maintenance
  - Repairs
  - QC
- Data analysis
  - Preprocessing
  - Exploratory data analysis
Current organization and scope of service

- **Project**
  - Panel design
  - Sample preparation
  - Troubleshooting/optimization

- **Reagents**
  - Reagent repository (Abs, buffers, conjugation kits)

- **Instrument**
  - Sample acquisition
  - Maintenance
  - Repairs
  - QC

- **Data analysis**
  - Preprocessing
  - Exploratory data analysis
Cytobank

- uzh.cytobank.org
- Great tool for data storage, sharing and analysis
- Well suited for core facilities as it simplifies support
- Has SPADE, viSNE, CITRUS
- Limitations?
Hybrid workflow

• Cytobank – Bioconductor – Cytobank
• Cytobank – Bioconductor
• (Cytobank ↔ custom code through API)
Hybrid workflow

F. Hartmann

• We run clustering on the dataset and visualize the resulting clusters on the viSNE map

• Currently we use FlowSOM
  – SAMSpectral, Phenograph, densityCut...

• Good reference to check:

• Metaclustering step
  – Existing knowledge of the number of populations
  – «elbow» method
  – Guided by the tSNE map?
Hybrid workflow

F. Hartmann
Conclusion

• Since few years now there has been a massive excitement about mass cytometry
• Not every site that purchased the instrument ended up happy, though...
• Be well prepared
• Well designed shared resource labs can minimize the entry barriers
We’re open for collaboration

• Discussion never harms...
• http://www.cytometry.uzh.ch/mcf
• vinko.tosevski@uzh.ch
Thank you for your attention!