



University of  
Zurich <sup>UZH</sup>

# Mass cytometry (CyTOF) workshop

University of Zurich (UZH) Mass Cytometry Facility

13-15 March 2017

Instructors: Vinko Tosevski, Tess Brodie,  
Mark Robinson, Lukas Weber

# Overview

- Workshop structure
- Day 1 and 2: “Experimental” and “Data analysis” (one day each; split into two groups and alternate)
- Day 3: “Tutorials and further data analysis”
- more details at:  
<http://www.cytometry.uzh.ch/en/mcf/training.html>

# Overview

- Data analysis sessions
- Overall structure: 1 day, 4x 1.5 hour sessions

Time	Content
9:00 – 10:30	Introduction
	<i>(short break)</i>
11:00 – 12:30	R programming
	<i>(lunch break)</i>
13:30 – 15:00	CyTOF workflows and plotting in R
	<i>(short break)</i>
15:30 – 17:00	CyTOF workflows and plotting in R (continued)

# Overview

- Data analysis sessions
- Information and materials available on GitHub:  
<https://github.com/lmweber/CyTOF-workshop-2017-03-13>
- *Please download data set (just in case problems later when we need it...)*
- Code of conduct

# Interactive workshop

- Please ask questions!
- Suggestions for examples?
- Interactive demonstrations
  - we will type R commands on screen
  - follow along on your own computer
  - R scripts containing all demonstration examples will be provided at end of workshop
- Etherpad: <https://beta.etherpad.org/p/VaaeWno43s>

# First exercise (5 mins)

- In Etherpad, please write down:
  - (1) What you would like to get out of today's data analysis sessions
  - (2) Any examples of types of data analyses (in R) that you would like to learn about
- Link to Etherpad (also in GitHub repository):  
<https://beta.etherpad.org/p/VaaeWno43s>
- (During workshop: Use Etherpad for notes, questions, code snippets, ...)

# R/RStudio

- R: programming language
  - popular for data analysis, statistics, plotting
  - freely available
  - download from “CRAN”: <https://cran.r-project.org/>
- RStudio
  - “integrated development environment (IDE)” or “editor”
  - to interact more easily with R
  - freely available (laptop/desktop version)
  - download from: <https://www.rstudio.com/>

# RStudio: panel layout

The screenshot displays the RStudio application window with four main panels. The top-left panel is the script editor, showing an R script with comments and package installation code. The bottom-left panel is the console, displaying the output of the script execution. The top-right panel is the Environment and History pane, showing the current environment. The bottom-right panel is the Plots pane, displaying a scatter plot of data points.

**scripts (.R files containing saved commands)**

```
1 #-----
2 # Script started by Felix J. Hartmann (Inst Experimental
3 # Immunology @UZH), heavily modified by Mark Robinson
4 # and Lukas Weber, IMLS @ UZH
5 #-----
6
7
8
9 #-----
10 # install packages if needed (only in a fresh R session)
11 #-----
12 source("http://bioconductor.org/biocLite.R")
13 biocLite(c("FlowCore", "ggplot2", "Rtsne", "Rphenograph",
14           "FlowSOM", "mclust", "mclustStats"))
15
16 biocLite("FlowSOM") # FlowSOM clustering
17 biocLite("plyr")     # for data handling
18 biocLite("xlsx")
19 biocLite("flowWorkspace")
20 biocLite("ggcyto")   # plotting flowSets
21 biocLite("devtools")
22 devtools::install_github("JinmiaoChenLab/Rphenograph")
23
24
25
26 #-----
27 # load packages, set working directory
28 #-----
29 setwd("~/projects/cytof/teaching")
30 library(ggplot2)
31 library(Rphenograph)
32
33
```

**command history, names of variables**

Environment is empty

**plots, help files, file explorer**

1:10

Index

Console

~/Dropbox/teaching/Cytof\_workshops/Cytof\_workshop\_March\_2017/draft\_script\_from\_Mark/

You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.

```
> 1 + 2
[1] 3
> ?seq
> plot(1:10)
>
```



# RStudio: key functionality

- Save commands in .R script file (top-left panel)
- Run/evaluate each line by clicking “Run”
  - or typing Cmd-Enter (Mac) or Ctrl-Enter (Windows)
- Lines beginning with # are “comments” (not evaluated)
  - very useful! always try to include lots of comments (for others as well as yourself in the future)
- R Notebooks / R Markdown
  - files containing R commands, text, and output/plots all together
- “Projects”: collect all files from a project

# CyTOF data analysis workflows and R

- Cytobank: cloud-based analysis platform:  
<https://www.cytobank.org/>
- Cytobank is very good at what it does, but during detailed analyses you will often reach limits, e.g:
  - data requires additional pre-processing? (quality control, filtering, transformations, ...)
  - want to try a different clustering algorithm?
  - adjustments to plots, different types of plots, heatmaps, ...
- Use whichever tools meet your data analysis needs, and allow you to perform rigorous, reproducible data analyses
- We want to teach you how to integrate R into your data analysis workflows: in particular, become comfortable to use R for steps where Cytobank is not enough

# CyTOF data analysis workflows and R

- Full workflow example
  - sessions 3 and 4 (afternoon)

# First example: getting data into and out of R

- See R script: “Session1\_data\_import\_export.R”

# Bioconductor

- Free, open source software project for analyzing high-throughput biology data using R
- Curated “package repository”
  - minimum standards regarding code quality and making sure packages do not “break”
  - easy to install packages
  - excellent documentation (“vignettes”)
  - inter-operability between packages (useful for building analysis pipelines)
- <http://bioconductor.org/>

## About *Bioconductor*

Bioconductor provides tools for the analysis and comprehension of high-throughput genomic data.

Bioconductor uses the R statistical programming language, and is open source and open development. It has two releases each year, [1296 software packages](#), and an active user community. Bioconductor is also available as an [AMI](#) (Amazon Machine Image) and a series of [Docker](#) images.

## News

- Bioconductor [3.4](#) is available.
- Bioconductor [F1000 Research Channel](#) launched.
- Orchestrating high-throughput genomic analysis with *Bioconductor* ([abstract](#)) and other [recent literature](#).
- Read our latest [newsletter](#) and [course material](#).
- Use the [support site](#) to get help installing, learning and using Bioconductor.

### Install »

Get started with *Bioconductor*

- [Install Bioconductor](#)
- [Explore packages](#)
- [Get support](#)
- [Latest newsletter](#)
- [Follow us on twitter](#)
- [Install R](#)

### Learn »

Master *Bioconductor* tools

- [Courses](#)
- [Support site](#)
- [Package vignettes](#)
- [Literature citations](#)
- [Common work flows](#)
- [FAQ](#)
- [Community resources](#)
- [Videos](#)

### Use »

Create bioinformatic solutions with *Bioconductor*

- [Software](#), [Annotation](#), and [Experiment packages](#)
- [Amazon Machine Image](#)
- [Latest release announcement](#)
- [Support site](#)

### Develop »

Contribute to *Bioconductor*

- [Developer resources](#)
- [Use Bioc 'devel'](#)
- 'Devel' [Software](#), [Annotation](#) and [Experiment](#) packages
- [Package guidelines](#)
- [New package submission](#)
- [Build reports](#)

# Installing Bioconductor

- See R script: “Session1\_Bioconductor.R”

# Note: R packages

- One of the strengths of R
  - an enormous range of “packages” is available, created by many different users, implementing a wide range of statistical methods
- Three main sources:
  - CRAN (all fields): <https://cran.r-project.org/>
  - Bioconductor (high-throughput biology): <http://bioconductor.org/>
  - GitHub (packages under development): <https://github.com/>



# Note: R packages

- Packages need to be “installed” and “loaded”
- Installation
  - from CRAN: `install.packages(“package_name”)`
  - from Bioconductor: using `biocLite()`; see example in R script
  - from GitHub: using `devtools` and `install_github`
  - note that `biocLite()` can also install packages from CRAN, but not the other way around
- Loading
  - `library(“package_name”)`

# Note: R packages

- Installation and loading
- Packages need to be:
  - installed once (save the package files on your local computer)
  - loaded at the start of every R session (load the package functions into your R session so you can use them)
- Packages may also need to be re-installed when you upgrade to a new version of R

# flowCore

- Bioconductor website
  - click on “Software”, then “flowCore”
- “Core” package containing fundamental “data structures” and functions to work with flow and mass cytometry (CyTOF) data
- .fcs files (“Flow Cytometry Standard”)

# .fcs files

## Markers

Cells

	FSC-A	FSC-H	SSC-A	B515-A	R780-A	R710-A	R660-A	V800-A	V655-A	V585-A	V450-A	G780-A	G710-A	G660-A	G610-A	G560-A
[1,]	27700.75	27291.75	177.52585	1984.485	625.0796	1232.1008	748.5101	1553.0295	1350.2565	3175.717	2338.1038	2286.1709	1758.4244	2550.914	1862.9843	1972.4854
[2,]	41264.25	39764.25	320.12296	3639.620	539.7032	1433.3112	1470.2659	2217.6750	2305.3516	5683.995	4767.4976	2617.9441	2063.0974	3401.507	2471.2236	2268.8318
[3,]	65054.75	57606.25	203.01607	2191.861	198.6541	726.9798	766.2198	802.2521	809.9579	1763.534	2870.2039	721.3581	750.2025	1156.660	879.6395	802.7821
[4,]	30584.00	31664.50	130.68690	1873.409	1304.0895	2528.7083	784.6980	1702.3671	1185.8608	2063.256	1839.7687	3945.7107	2480.2170	2585.438	1850.5399	2768.8843
[5,]	39505.75	39626.00	203.25166	2540.620	323.2625	857.1525	715.0004	1117.4775	1746.5798	3810.514	3442.1965	1766.7188	1595.8044	2746.546	2118.4902	1560.5364
[6,]	33171.50	34794.00	333.64246	2192.864	1408.8563	2573.5095	1604.2236	2128.1748	1727.5891	3734.910	2551.4509	5498.1421	2989.3887	3295.179	2938.3291	4663.7070
[7,]	63711.00	54475.50	1122.48340	3879.044	1730.8085	3573.5652	1691.8744	5106.0596	3578.0332	9183.305	7264.9624	7268.8550	4263.6313	4767.073	4025.7312	5758.2100
[8,]	40000.75	40213.50	236.54262	2545.858	1081.6753	2313.5962	1411.0983	2989.7524	1920.4047	4386.833	3081.6401	3615.0461	2552.7622	3304.552	2666.3101	3532.8064
[9,]	49286.00	49182.50	78.61845	1601.092	123.2834	493.6364	242.0255	633.3533	759.2227	1920.717	2082.5884	996.4465	805.6166	1197.744	957.8304	937.6376
[10,]	32209.75	33368.25	203.29897	2387.361	1056.0723	1769.4005	939.7758	1693.8635	1579.7000	3308.004	2106.1416	2835.7053	2057.3567	2653.473	1879.0623	2306.1431
[11,]	35937.25	36212.50	220.66580	2901.591	1218.1395	3202.3853	1059.7604	2443.0205	2253.0146	4684.219	2864.9727	5582.5425	3703.6475	4023.313	3111.4429	4476.1655
[12,]	32905.50	33897.50	233.98033	2726.240	1952.0721	3405.7139	2726.1091	2988.6882	2011.0159	4644.849	2661.8582	7080.1904	3540.6592	3915.066	3389.4014	6423.2324
[13,]	36028.50	35845.50	219.18674	3221.668	2542.3389	3895.0371	2283.0444	3331.8298	2479.6580	5253.382	2692.4919	9087.0859	4306.8105	4652.118	4095.7131	7948.1416
[14,]	38616.00	38775.00	218.46669	3218.305	582.6801	1022.7971	1255.5858	2150.4185	1993.2681	4657.159	1881.8761	2058.4124	1753.6392	2576.016	1914.5718	2386.2805
[15,]	45282.25	42223.25	1173.74487	6941.545	705.4651	1649.9570	1615.0811	4287.2036	3778.2302	11158.718	8247.2236	3138.5728	2382.1287	3487.720	2436.5886	3280.1677
[16,]	36246.25	36207.75	189.15569	3049.417	1736.7826	2823.7266	1031.0308	2824.6582	2053.6843	4689.628	2930.9685	6049.8901	3479.7673	3642.755	2885.2795	4767.1309
[17,]	29282.75	29884.00	209.64102	1836.197	612.2673	1149.7164	870.3303	1720.2170	1525.6914	2782.762	2335.3994	2180.0403	1825.0010	2702.035	1921.6849	1730.4067
[18,]	57757.25	54448.25	1999.17517	12972.877	4364.5908	11298.7070	6745.5039	20934.3457	17057.1934	30293.186	7082.0664	14765.4785	13275.3828	18369.047	14707.2549	9409.2637
[19,]	33301.00	33093.50	208.47151	2146.622	429.5022	855.5981	845.9418	1207.8969	1297.2683	2751.205	1081.5209	1716.1721	1479.2837	2110.357	1566.1290	1371.2155
[20,]	34478.25	35390.75	211.26921	3060.585	2016.3651	3442.5408	1348.4852	2673.9729	2259.8494	3781.046	1290.9741	7142.1543	3906.1033	4101.971	3586.8269	5697.7954
[21,]	29406.25	28219.50	231.55798	3008.380	997.8875	2319.5779	1514.2091	1757.2463	1675.9983	3019.569	1565.1250	4008.3364	2470.9639	2938.669	2289.2273	3631.8906
[22,]	49978.50	48517.75	537.04224	3122.343	981.1232	2252.1189	1861.3472	2518.4731	2230.5327	5010.112	3615.4170	4111.9141	2611.4641	3379.042	2634.6746	3722.1040
[23,]	39872.50	37620.75	198.75706	2719.222	1657.0939	2945.5713	1025.1293	2203.0527	1670.1367	3072.683	1879.3065	6735.8999	3531.4177	3581.812	2901.9954	4790.6758
[24,]	33395.00	35331.75	220.46056	2664.632	690.1926	1483.0898	1736.9537	1397.0316	1982.9124	3473.224	1987.4087	2142.7542	1793.1940	2844.251	2057.5803	2655.7776
[25,]	46976.00	47355.25	231.33037	2530.461	537.1376	1194.0681	1072.7083	1531.7494	1766.5841	3532.332	3953.0066	3967.8032	1918.4053	3249.490	2566.3657	3888.4016
[26,]	56663.75	51458.25	223.06416	3217.866	398.6222	1279.4880	1207.4561	1268.9905	1553.6884	3397.891	2623.9312	1686.8226	1583.9100	2461.450	1874.0762	2013.9438
[27,]	50818.75	48556.25	305.77182	3714.351	577.0732	1364.4095	1064.0983	1633.2513	2077.0466	3332.829	2654.2678	2074.0388	1966.8965	2876.737	1987.8170	2010.1973
[28,]	36225.25	36196.75	180.30524	2636.466	946.7570	2138.4143	1695.0502	1807.8429	2057.7292	3673.466	2611.5857	3530.6289	2630.3755	3608.507	2727.3037	2884.0103
[29,]	28509.25	30715.50	230.27397	1072.201	1867.2009	1643.1423	882.4811	1201.5806	688.1475	1810.280	616.8154	3604.4263	1530.6011	1353.209	1167.5325	2577.3455
[30,]	37198.75	36200.50	237.67776	3046.719	1376.3452	2580.9287	1326.2197	2599.6101	2196.7258	5525.008	3665.0623	4946.4224	3421.0840	3996.508	3045.4102	4413.0908

# .fcs files

- similar to a spreadsheet
- larger, faster to load, more efficient storage
- contains standard meta-data for flow cytometry experiments
- (our CyTOF workflow example will use data saved in .fcs files)

# .fcs files

- Importing and exporting data using .fcs files in R
- see R script: “Session1\_fcs\_files.R”