Basic Data Analysis, Gating, and Statistics in Flow Cytometry



Alex Henkel

Associate Instrumentation Specialist

amhenkel@wisc.edu

uwflow@uwcarbone.wisc.edu

<u>Outline</u>

- What is an FCS File?
- Visualization and Scaling
- Gating and Controls
- Basic Statistics in Flow Data
- The Process

<u>The FCS File – Flow Cytometry Standard File</u>

2	3	4	5	6	7	8	9	10
FSC-H	FSC-W	SSC-A	SSC-H	SSC-W	488 A 695/40	561 D 610/20	561 A 780/60	405 D 525/50
262144	262144	262144	262144	262144	262144	262144	262144	262144
32	32	32	32	32	32	32	32	32
136813.00	72175.84	100686.88	88489.00	74569.90	9000.32	15158.64	3904.32	15887.52
136392.00	72013.31	98798.56	86540.00	74819.30	8873.76	15250.76	3290.84	16505.28
136761.00	72370.88	99523.20	86557.00	75353.27	8064.00	14991.06	3386.88	16967.61
136288.00	72599.80	102604.32	90900.00	73974.45	8835.68	13644.54	2762.62	16658.73
136493.00	72679.64	102473.28	90521.00	74189.29	8384.32	14253.12	3108.56	15756.84
137809.00	71919.16	102944.80	91413.00	73803.40	8487.36	14612.78	3410.40	17198.28
136451.00	72557.54	97821.92	85559.00	74929.08	8755.04	14392.28	3188.92	15426.18
131217.00	73191.84	87923.36	77317.00	74526.24	7935.20	13974.80	3613.26	14421.33
137763.00	72796.29	98178.08	86391.00	74477.66	8622.88	14378.56	2479.40	15633.09
138296.00	71846.40	101661.28	90380.00	73716.24	8266.72	15017.52	3336.90	17202.24
138825.00	72279.32	108109.12	94829.00	74713.85	9089.92	14271.74	3181.08	16338.96
137084.00	71431.94	97625.92	86136.00	74278.02	9569.28	15296.82	3063.48	17455.68
136099.00	72974.48	94170.72	81575.00	75655.19	8542.24	16242.52	3597.58	17812.08
138561.00	72564.37	100914.24	87989.00	75162.98	8611.68	15122.38	3232.04	17161.65
120599.00	73751.43	128493.12	112851.00	74619.85	13147.68	22317.54	4793.18	23744.16
132646.00	73071.72	91556.64	79586.00	75393.36	7881.44	14176.68	3687.74	14792.58
	FSC-H 262144 32 136813.00 136813.00 136392.00 136761.00 136761.00 136761.00 136761.00 136761.00 136761.00 136761.00 136761.00 136288.00 136763.00 137763.00 138296.00 138296.00 137084.00 136561.00 120599.00	FSC-H FSC-W 262144 262144 32 32 136813.00 72175.84 136392.00 72013.31 136761.00 72370.88 136288.00 72599.80 136493.00 72679.64 137809.00 7191.16 136451.00 7257.54 138296.00 72193.21 138296.00 71846.40 138296.00 71431.94 137084.00 71431.94 136099.00 72974.48 138561.00 73751.43	FSC-H FSC-W SSC-A 262144 262144 262144 32 32 32 136813.00 72175.84 100686.88 136392.00 72013.31 98798.56 136761.00 72370.88 99523.20 136493.00 72679.64 102473.28 136493.00 7257.54 97821.92 136493.00 7257.54 97821.92 136493.00 7257.54 97821.92 136493.00 7257.54 97821.92 13769.00 7191.84 87923.36 13625.00 7279.29 98178.08 138296.00 71846.40 101661.28 138025.00 71431.94 97625.92 137084.00 71431.94 97625.92 136099.00 72974.48 94170.72 138561.00 72564.37 100914.24 120599.00 73751.43 128493.12	FSC-H FSC-W SSC-A SSC-H 262144 262144 262144 262144 32 32 32 32 136813.00 72175.84 100686.88 88489.00 136392.00 72013.31 98798.56 86540.00 136761.00 72370.88 99523.20 86557.00 136288.00 72599.80 102604.32 9090.00 136493.00 72679.64 102473.28 90521.00 136451.00 72557.54 97821.32 85559.00 131217.00 73191.84 87923.36 77317.00 138296.00 71431.94 97625.92 86136.00 138025.00 72279.32 108109.12 94829.00 137084.00 71431.94 97625.92 86136.00 136099.00 72974.48 94170.72 81575.00 136099.00 72974.89 94170.72 81575.00 136099.00 72974.88 94170.72 81575.00 136099.00 72974.89 94170.72 8157	FSC-H FSC-W SSC-A SSC-H SSC-W 262144 262144 262144 262144 262144 262144 32 32 32 32 32 32 136813.00 72175.84 100686.88 88489.00 74569.90 136871.00 72013.31 98798.56 86540.00 74819.30 136761.00 72370.88 99523.20 86557.00 75353.27 136288.00 72599.80 102604.32 90900.00 73974.45 136493.00 72679.64 102473.28 90521.00 74189.29 137809.00 71919.16 102944.80 91413.00 7390.40 136451.00 7257.54 97821.92 85559.00 74292.08 137673.00 72796.29 98178.08 86391.00 74477.66 138296.00 71846.40 101661.28 90380.00 73716.24 138295.00 72279.32 108109.12 94829.00 74713.85 137084.00 71431.94 97625.92 8613	FSC-H FSC-W SSC-A SSC-H SSC-W 488 A 695/40 262144	FSC-H FSC-W SSC-A SSC-H SSC-W 488 A 695/40 561 D 610/20 Z62144 Z62144 <thz6214< th=""> <thz614< th=""> <thz6214< th=""></thz6214<></thz614<></thz6214<>	FSC-H FSC-W SSC-A SSC-H SSC-W 488 A 695/40 561 D 610/20 561 A 780/60 262144 <th< td=""></th<>

- Comprised of a text segment and data segment.
- FCS Files are in a list mode data format.
- Rows = Events
- Columns = Parameters
 - H,A,W each get their own column.
- Annotate data before recording.

The FCS File - Header

Keyword	Value					
\$ENDSTEXT	0					
\$BEGINDATA	3015					
\$ENDDATA	969014					
\$FIL	Rainbow_Beads_001.fcs					
\$SYS	Windows 7 6.1					
\$TOT	17250					
\$PAR	14					
\$MODE	C C					
\$BYTEORD	4,3,2,1					
\$DATATYPE	F					
\$NEXTDATA	0					
CREATOR	BD FACSDiva Software Version 8.0					
TUBE NAME	Beads					
\$SRC	Rainbow					

	Name	
€	Mouse_LN CD25 FMO_017.fcs	
OH	Mouse_LN Control 1_011.fcs (Control)	
OH	Mouse_LN Control 2_012.fcs (Control)	
OH	Mouse_LN Control 3_013.fcs (Control)	
OI	Mouse_LN FoxP3 FMO_018.fcs	
OH	Mouse_LN Naive_010.fcs	
OH	Mouse_LN UV 1_014.fcs	
OI	Mouse_LN UV 2_015.fcs	
OI	Mouse_LN UV 3_016.fcs	
OH	Mouse_SP CD25 FMO_026.fcs	
OI	Mouse_SP Control 1_020.fcs (Control)	
OH	Mouse_SP Control 2_021.fcs (Control)	
OI	Mouse_SP Control 3_022.fcs (Control)	
OH	Mouse_SP FoxP3 FMO_027.fcs	
OI	Mouse_SP Naive_019.fcs	
OH	Mouse_SP UV 1_023.fcs	
	Mouse SP UV 2 024.fcs	

Help								•		0	ок
Date: 28-JAN-2016		C	tion Matrix								
System:Windows 7 6.1		Compens	non wants								_
Cytometer:LSRII			405 D	405 E	488		31 A	561 D		42 A	Г
File:Mouse_LN Control 1_011.fcs			525 50-A					10 20-A		60-A	7
F	1	405 D	_		-					-	F
\$BEGINANALYSIS: 0		525_50-A	100	15.708	1.19	88 0	.716	0.717		0.239	
SEGINDATA: 3694	=	105.5		1					- <u>-</u>	_	F
\$BEGINSTEXT: 0	-	405 E 450 50-A	13.868	100	0.0	13 0	.063	0.658		0.013	
\$BTIM: 11:44:57										_	F
\$BYTEORD: 4,3,2,1		488 A	-0.013	-0.052	10	0 6	.679	0	1	1.827	
\$CYT: LSRII		695_40-A									
\$DATATYPE: F		RR1 A	1	11					11.		Ι.
\$DATE: 28-JAN-2016					- 110						_
\$ENDANALYSIS: 0		Parame	ters and Sta	ains							
\$ENDDATA: 10347173						Decade					_
\$ENDSTEXT: 0		Paramet (\$PnN)		Range (\$PnR)					Itage PnV)	Deriv	
\$ETIM: 11:46:27		FSC-A	(31113)			0.0	10	544		Fron	
\$FIL: Mouse_LN Control 1_011.fcs		FSC-H				0.0	1.0	544			-
\$INST:		FSC-W		262144	2	0.0	1.0	544		-	_
\$MODE: L		SSC-A		282144 3	12	0,0	1.0	399	-		-
\$NEXTDATA: 0		SSC-H		262144 3	12	0,0	1.0	399			_
\$OP:		SSC-W		262144	12	0,0	1.0	399			
\$PAR: 14 \$SRC: Mouse		488 A 6	D11a P	262144 3	12 1	0,0	1.0	546			
SSYS: Windows 7.6.1		561 D 6	D3 PCF	262144 3	12	0,0	1.0	495			_
STS: WHOWS 7 6.1 STIMESTEP: 0.01		561 A 7	oxp3 PC7	262144 3	12	0,0	1.0	631			
STOT: 184705		405 D 5			12	0,0	1.0	429			_
APPLY COMPENSATION: TRUE		405 E 4				0,0	1.0	416			
AUTOBS: TRUE		642 B 7	CD8 rF710	262144	12	0,0	1.0	600			_
CREATOR: BD FACSDiva Software Version 8.0		Median	of Parame	eters vs. Tim	e						
CST BASELINE DATE: 2014-11-21T15:48:51-06:00											
CST BEADS EXPIRED: False											-
CST BEADS LOT ID: 24628											
CST PERFORMANCE EXPIRED: 2014-12-11T09:00:10-06:00		1A	Mq _	5-1	bre		~	-14	1-	7_	~
CST REGULATORY STATUS: RUO Performance Check			V	· · · ·			·	· M	1	V	
CST SETUP DATE: 2014-12-10T09:00:10-06:00											
		FSC-A		FSC-H		FSC-W		I I St	BC-A-		
CST SETUP STATUS: SUCCESS WITH WARNING										-	_

- Header shows information about the data collection.
- Stores all the metadata from the instrument, including any added labels.

 Double clicking on the diamond next to a file in FlowJo will open the file information, including header information.

<u>Keywords</u>

- All the header information can be extracted as keywords in data analysis software, allowing you to group, or sort files by keyword.
- Keywords can be added to tables with sample statistics.
- Keywords can be metadata.

Name 🛆	Value	Explanation
SBEGINANALYSIS	0	
SBEGINDATA	3685	
SBEGINSTEXT	0	
SBTIM	12:11:44	The time at the beginning of data collection, format hh:mm:ss.
SBYTEORD	4,3,2,1	Byte Order
SCYT	LSRII	The cytometer used to acquire the data.
\$DATATYPE	F	Data format.
SDATE	28-JAN-2016	The date the file was created, format dd-mmm-yy.
SENDANALYSIS	0	
SENDDATA	10704836	
SENDSTEXT	0	
SETIM	12:12:13	The time at the end of data collection, format hh:mm:ss.
SFIL	Mouse_SP Naive_0	The name of the data file when it originally was created.
SINST		The institution where the data were collected.
SMODE	L	Data Mode. U = single-parameter histograms, C = correlated multi-parameter histog
SNEXTDATA	0	The byte offset of additional data set included in the file.
SOP	DeLuca	The name of the operator.
SP10B	32	Number of bits in parameter 10.
SP10E	0,0	The type of amplification for parameter 10 (logarithmic or linear).
SP10G	1.0	
SP10N	405 D 525/50-A	The name of the parameter 10.
SP10R	262144	The range of parameter 10.
SP105	LD GV 510	The name of the fluorescent stain or probe used with parameter 10.
SP10V	429	
SP11B	32	Number of bits in parameter 11.
SP11E	0.0	The type of amplification for parameter 11 (logarithmic or linear).
SP11G	1.0	
SP11N	405 E 450/50-A	The name of the parameter 11.
SP11R	262144	The range of parameter 11.
SP11S	CD4 vF450	The name of the fluorescent stain or probe used with parameter 11.
SP11V	416	
SP12B	32	Number of bits in parameter 12.
SP12E	0,0	The type of amplification for parameter 12 (logarithmic or linear).
\$P12G	1.0	1 2356 vit 1920 vit 177 1920.
SP12N	642 B 730/45-A	The name of the parameter 12.
SP12R	262144	The range of parameter 12.

How does the cytometer generate values?

Event

1

2

3

6

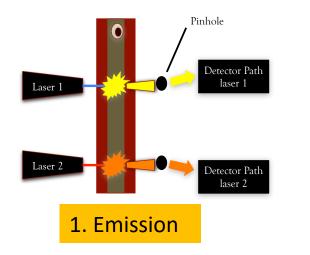
7

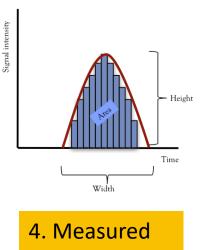
8

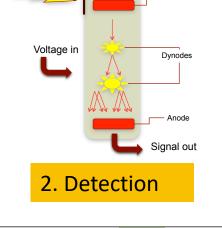
9

10

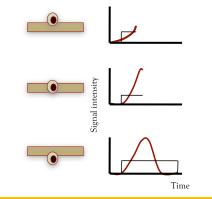
- Detection of light from scatter or fluorescence in cells is converted into a voltage pulse.
- Each pulse generated by an object passing by the laser has a height, area and width measurement.
- Values for H, A, and W from each parameter stored in a listmode data format in FCS file.



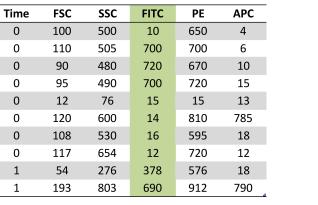




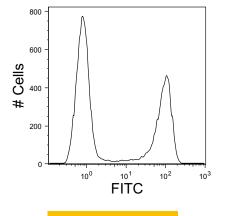
Photocathode



3. Converted to Voltage



5. File Generated



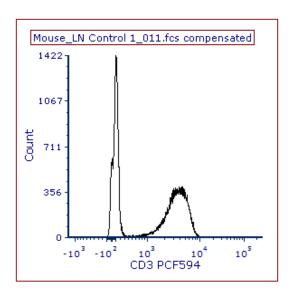
6. Plotted

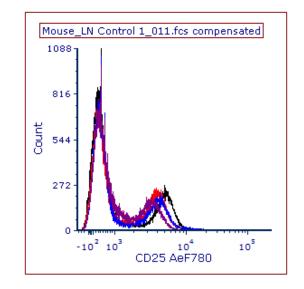
From Excyte Expert Cytometry

Visualization

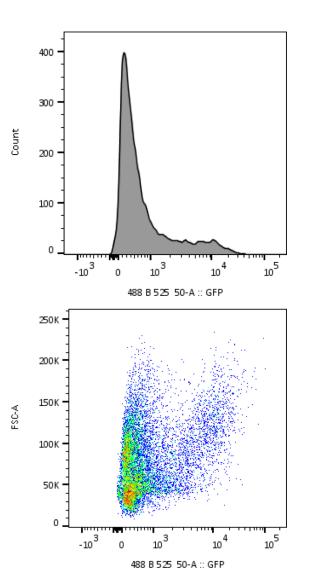
- We can create several types of plots with the generated data.
- Histograms, dot plots, density plots, contour plots.
- We want to display the data in a way that relates to our hypothesis.

<u>Histograms</u>

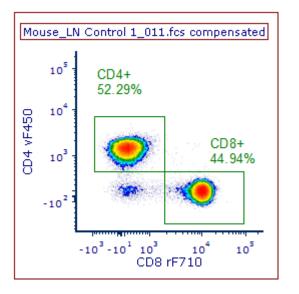


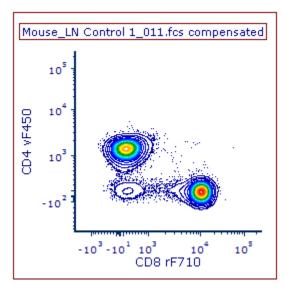


- Shows the distribution of values for a specific parameter.
- Cannot see the relationship between two populations.
- Can miss sub-populations that have similar values in one parameter.

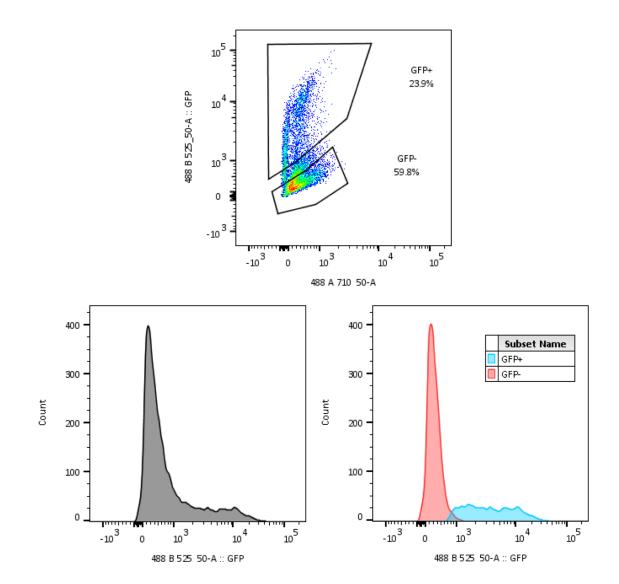


Bi-variate Dot Plots



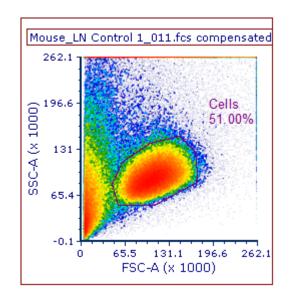


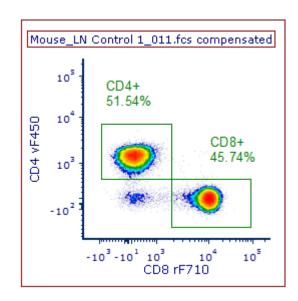
- Can see relationships between markers.
- Can see sub-populations from distinctions in two dimensions instead of one.



Scaling Considerations

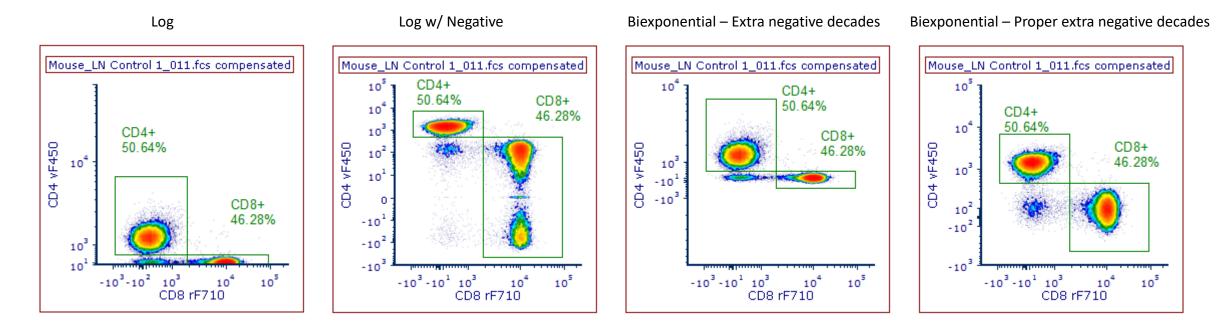
- Forward and Side scatter <u>almost</u> always displayed in <u>linear scale</u>.
 - Exception for very small things like bacteria, extracellular vesicles, and nuclei.
 - Use linear for small dynamic range.
- Fluorescence parameters <u>almost</u> always in <u>log scale</u>.
 - Exception for low signal increase. Ex. Cell cycle assay will only have a two-fold increase in fluorescent intensity.
 - Use log for large dynamic range.



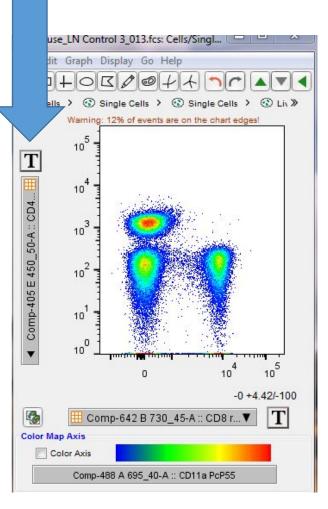


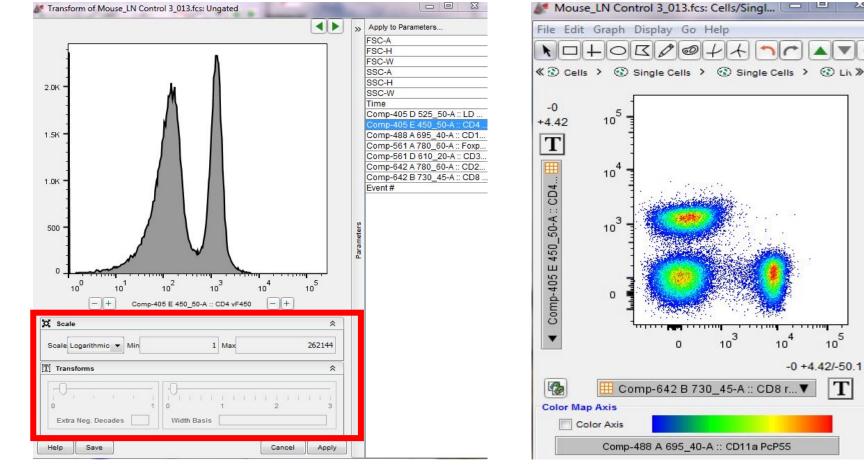
Scaling Plots to Display Data

- Sometimes plots don't display the data in the best way.
- Changing the scaling does not change the values, just the display of the data.
- Linear, Log, Biexponential, Hyperlog.



Adjusting Scaling in FlowJo





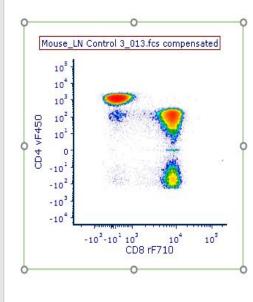
105

T

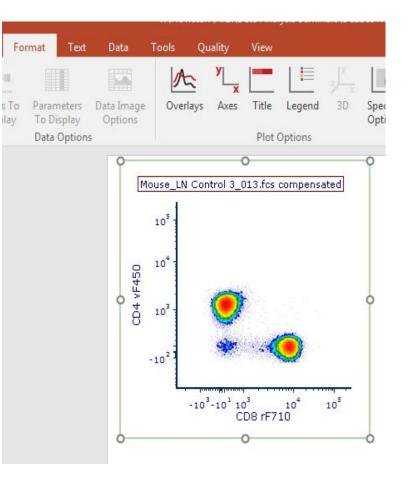
Adjusting Scaling in FCS Express

X



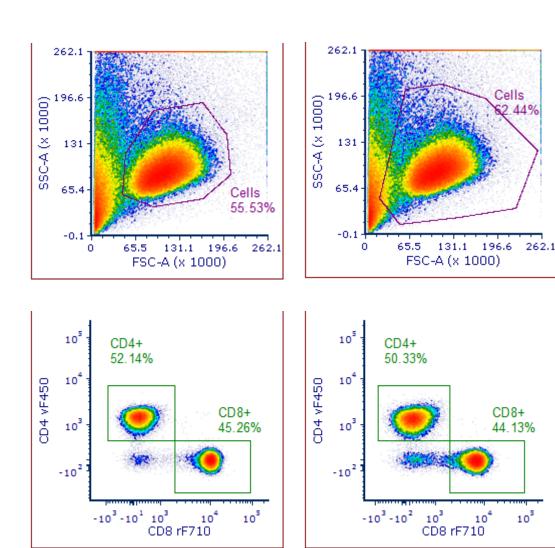


Derive Axis Titles	From		
Parameter Nar	me		
🔵 Parameter lab	el set		
The specified	text below		
Select Axis			
		Y Axis	X Axis
Scale			
Automatic	.og with Negat	ive	•
Range			Title
🔽 Automatic			Visible
Minimum	0		Text
Maximum	10000		Angle 90
# Ticks	4		Font
Axis Line			Labels
Width	2.00 🌲	Visible	Visible
Style	-		Angle 0
Color			Font
Grid Lines			Divide by Factor
Width	1.00	Visible	V Divide Axis Labels
Style			Factor 1000
Color			Show Division Factor in Title
Ticks			Other
Width	1.00 🌲	Visible	Display Axis at Zero

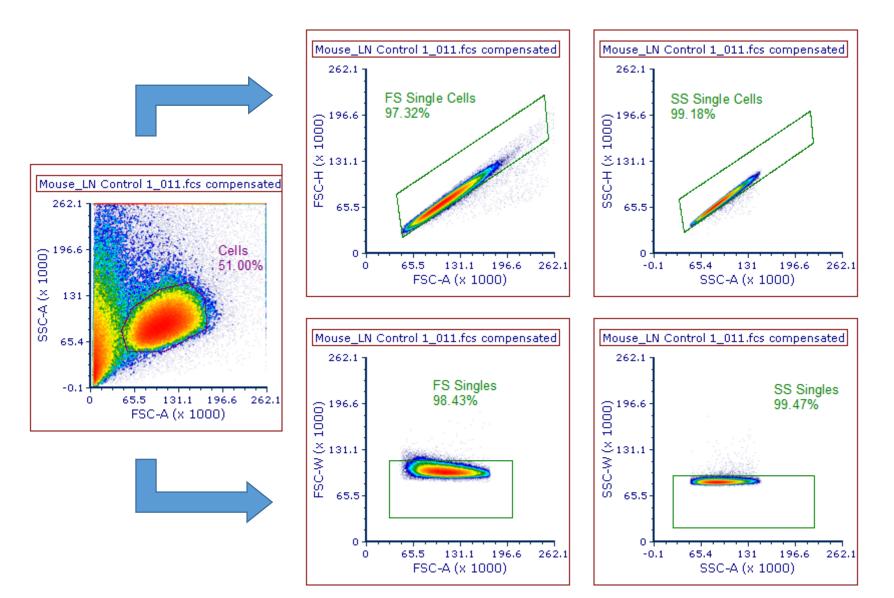


Basic Gating Considerations

- Gating on cells only
 - Exclude debris.
- Doublet Discrimination
 - Removes events that are two cells stuck together.
- Live/Dead gating
 - Dead cells soak up antibody.
- Each of these things lead to double positive or strange events!

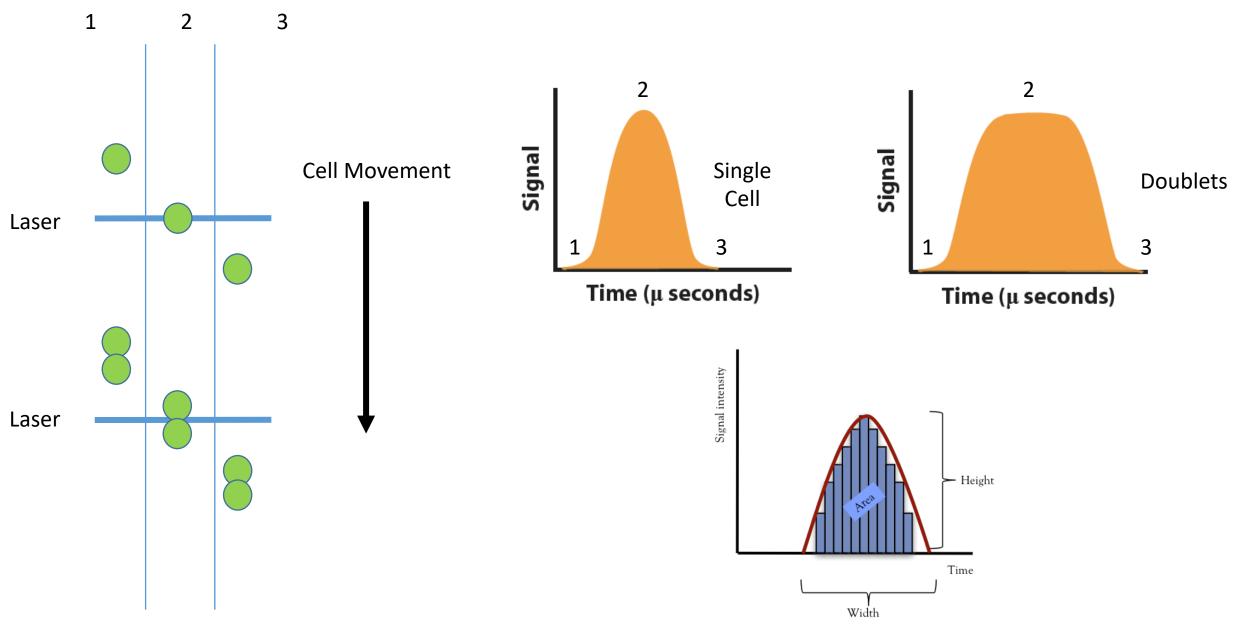


Doublet Discrimination

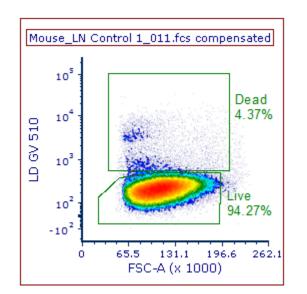


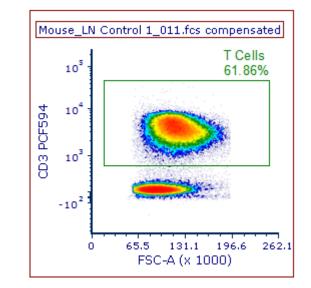
- Helps remove events that are two or more cells stuck together.
- Reduces contribution to false positive or double positive events.

Why Doublet Discrimination Works



Live/Dead Gating



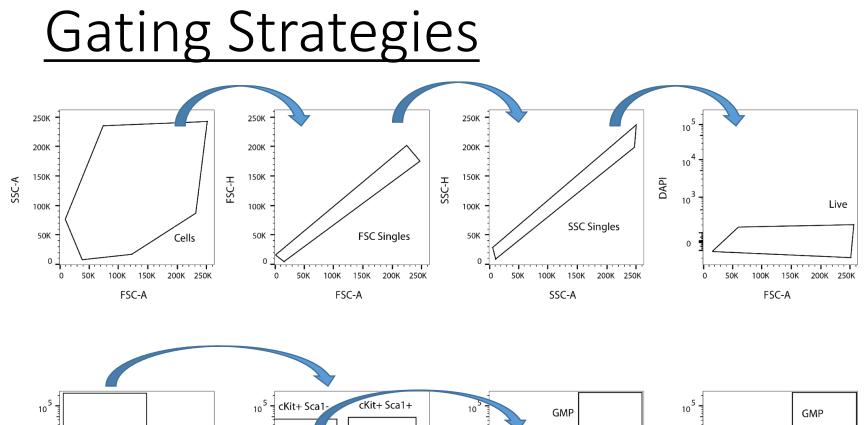


Tcells 30.43%

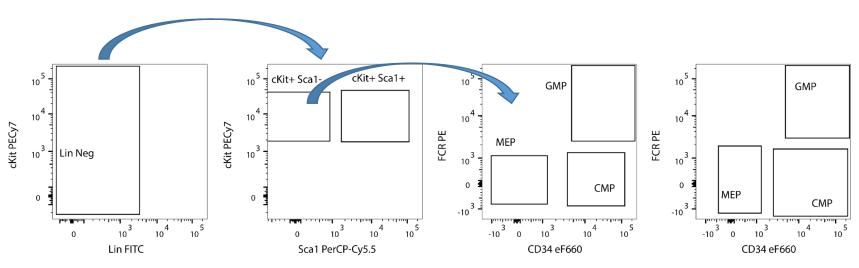
Mouse_LN Control 1_011.fcs compensated

From Live gate

From Dead gate



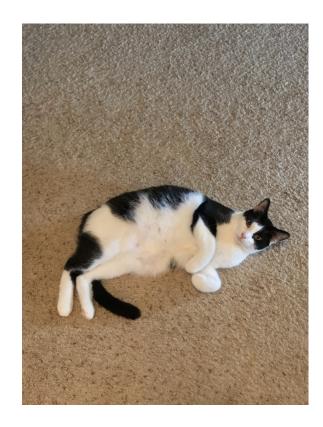
- General gating strategy
 - Doesn't have to be in this order.



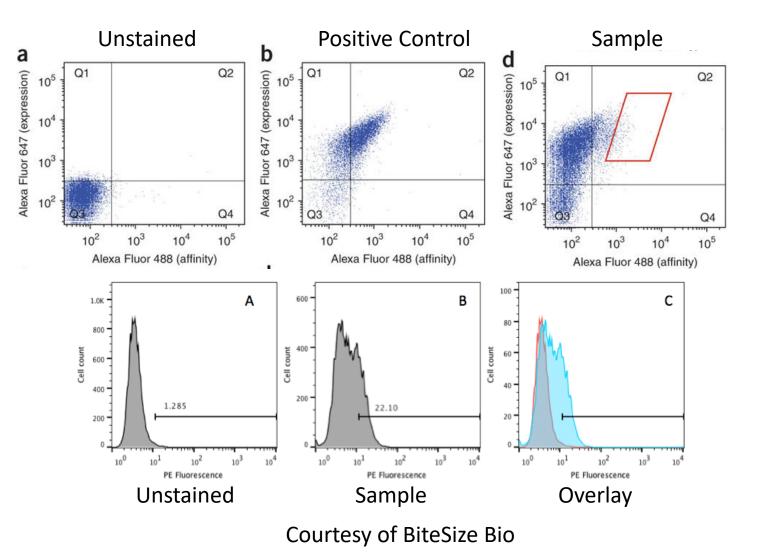
UWCCC Flow Lab for Kirby Johnson, PhD

Common Gating Controls

- Fluorescence minus one (FMO) Control
 - Gating control, shows the background and contributions from neighboring fluorescence spillover.
- Positive Control
 - Standardize gating procedure and observe staining profile.
 - Treated to induce positivity.
- Biological Controls
 - Stim vs. Unstim, T0 vs. Time course, Treated vs. Untreated.
 - Any control you need to prove your hypothesis.
- Unstained control
 - To evaluate inherent background and autofluorescence.

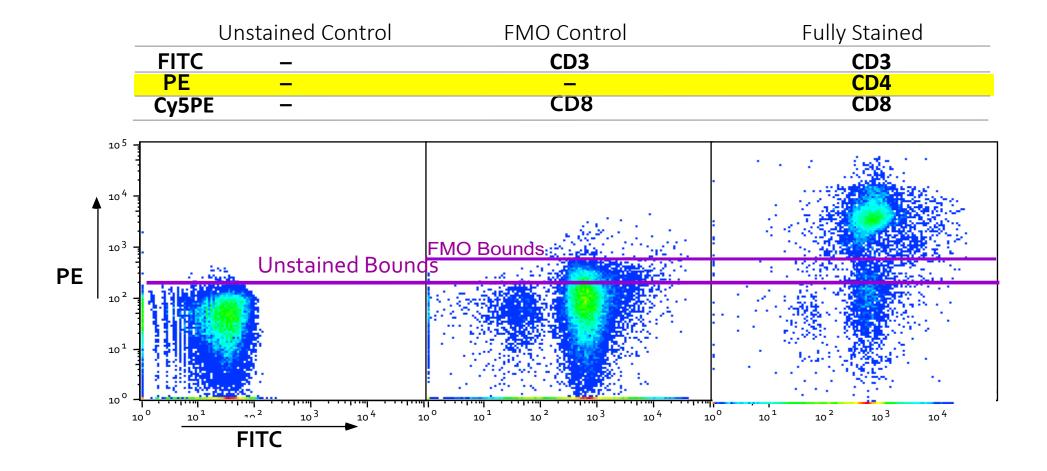


Negative and Positive controls slide



- Positive controls
 - Biological control to assess what the signal looks like when the antigen of interest is present.
 - Useful for rare positive populations or when antigen expression is variable between samples.
- Unstained/Negative control
 - Helps make gating decisions.
 - Visualizing autofluorescence.

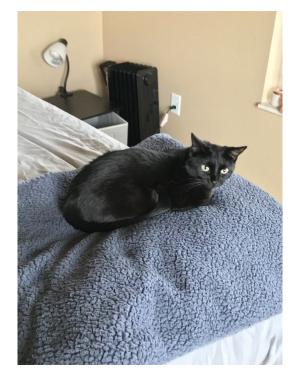
<u>FMOs</u>



From Excyte Expert Cytometry (Courtesy of M. Roederer, Ph.D, NIH Vaccine Center)

Why NOT Isotype Controls?

- Nearly impossible to determine if the isotype antibody has the same number of average fluorophores attached per Ab as experimental Ab.
- Different antibody than test sample, different binding properties.
- Maecker HT, and Trotter J. <u>Flow cytometry controls, instrument setup</u> <u>and determination of positivity.</u> Cytometry Part A 2006; 69A:1037– 1042
 - Flow Server (T drive) > Flowdata > Flow Resources > Flow References > Isotypes.
- Can use isotype control to test how well the blocking buffer worked.

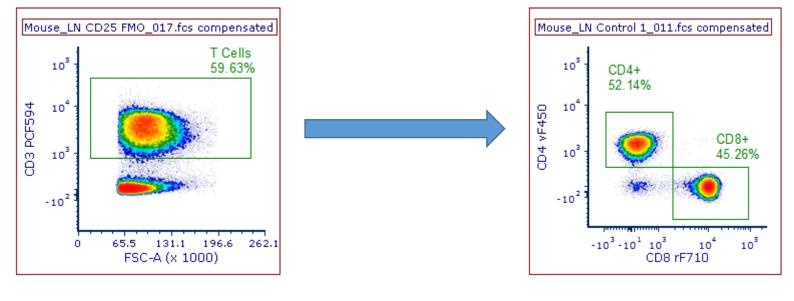


Basic Statistics in Flow Cytometry

- Typically described using frequencies and fluorescence intensity.
- Frequency
 - Number of events in the target population within a larger population.
- MFI (Median Fluorescence Intensity)
 - NOT mean. Mean is subject to outliers, median is less affected.
- Statistical modeling (Following Seminars)
 - DNA Cell Cycle analysis
 - Proliferation analysis
- Absolute counts
 - Volumetric based acquisition cytometer or counting beads spiked in sample at known concentration. (Counting beads tend to be problematic)

Frequency

- Ex. Number of CD4+ cells in a population of Live, single, CD3+ positive cells.
- Used to analyze presence of antigen/marker.



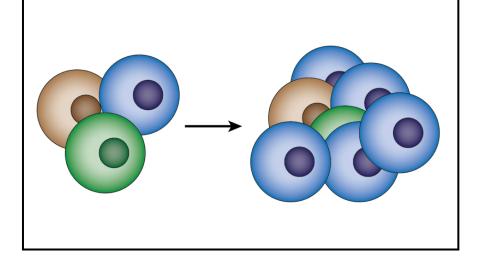
From Live Single Cells

49.81% of CD3+ cells are CD4+ 44.86% of CD3+ cells are CD8+

Frequency Hypotheses

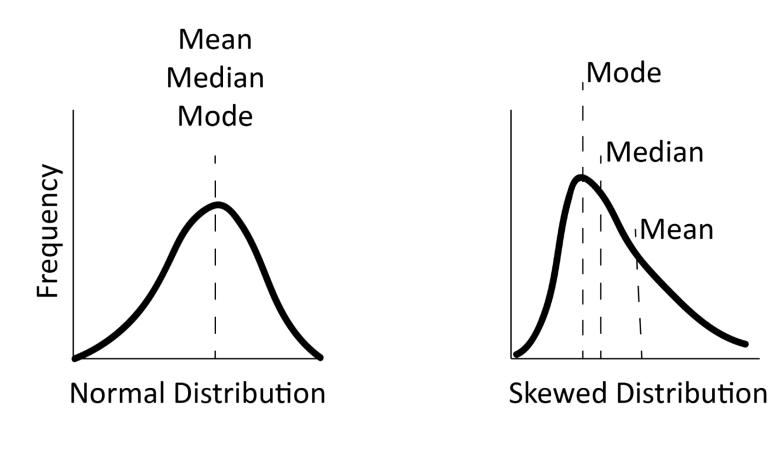
Treatment increases numbers of cell type Y

Cell type Y is more prevalent in disease state A



Report % positive to evaluate changes in composition of cell populations

Central Tendency

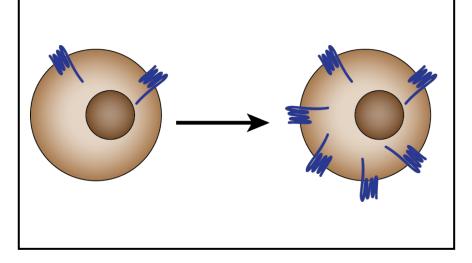


Most flow cytometry data is displayed on a Logarithmic scale – What looks symmetrical is actually skewed!

Median Fluorescence Intensity

Treatment increases expression of protein X

Protein X is upregulated in disease state Y



- Use the MFI to assess *levels* of target protein expression
 - Median for logarithmic data.
 - Mean is ok for linear only.
- Standardizing your assay is critical
 - Reference Standard for PMT sensitivity (Rainbow Beads).
 - Can compare samples run on different days.
 - <u>https://cancer.wisc.edu/research/wp-content/uploads/2017/03/Flow_TechNotes</u>
 <u>Rainbow-Standard-Tech-Note_20170918.pdf</u>
- Fold increase?

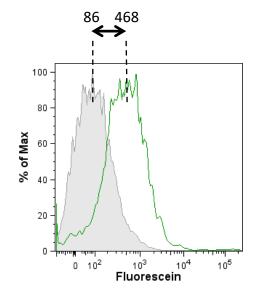
Fold-change in MFI

- Used in comparison of expression level of antigen/marker between samples.
 - Fold-change in MFI = MFI(sample)/MFI(control)
 - Can compare fold-change in MFI between

treatments/samples.

Caution:

- In order to use Fold-change in MFI, need to be aware of potential skewing of data due to log scale.
- Small changes in negative can translate into large changes in the fold.



Control MFI = 86 Experimental MFI = 468 Fold-change in MFI = 468/86 = 5.44

The Data Analysis Process

- 1. Have a specific Hypothesis. ASK A SPECIFIC QUESTION!
 - 1. Need to know which statistics you are after.
- 2. Gate on live, single cells and use controls to gate each fluorescent parameter.
- 3. Gather statistics from plots and gates.
- 4. Perform Analyses.

Acknowledgements

- Thank you to the DeLuca Lab for data used in this presentation.
- Excyte Expert Cytometry for graphics used in this presentation.

Mark your calendars for upcoming UWCCC Flow Lab Seminars!

Rigor and Reproducibility in Flow Cytometry

Friday, December 14, 2018 10am, WIMR 7001A

Overview of Computational Data Analysis Platforms for Flow Cytometry

Friday, January 11, 2019 10am, WIMR 7001A

Flow Cytometry – Compensation with Confidence

Friday February 1, 2019 10am, WIMR 7001A

Multicolor Panel Design for Flow Cytometry

Tuesday, March 5, 2019 2pm, WIMR 7001A

Flow Cytometry Current Best Practices for Pls

Thursday, February 14, 2019 7:30am, WIMR 7170

Data Analysis with Alex II

Tuesday, March 7, 2019 10am, WIMR 7170

Data Analysis with Alex III

Wednesday, May 16, 2019 10am, WIMR 7170

