Rare Event Detection

Albert D. Donnenberg, Ph.D.

UNIVERSITY *of* PITTSBURGH CANCER INSTITUTE

donnenbergad@upmc.edu

Searching for a Needle in a Haystack

- Find the needle
- Determine that it really is a needle
- Make measurements to determine what kind of a needle it is

Key Elements

- Event frequency
  - Inherent property of sample
  - Enrichment possible
- Signal to noise ratio
  - Minimize noise
    - Nonspecific binding (1% mouse Ig)
    - Cellular autofluorescence (dump gate, green or red excitation, quenching dyes)
    - Doublets (ratio of peak height/integral or peak height/width)
    - Sporadic mechanical or electrical noise (time parameter)
    - Dead cells (vital dyes)
  - Maximize signal
    - Best fluorochrome for most critical determination
    - Optimal antibody concentration

Know Your Own Limit (of Detection)

- Limit of detection

  Frequency of false positives in appropriate negative control (FMO isotype control, FMO isoclonic control, TMer binding of MHC disparate cells, known negative sample)
- Calculate upper 95^{th} or 99^{th} percentile of the frequency false positive in a series of negative controls
- Caution: Rare events are log normally distributed. Use arithmetic means and you will get the wrong answer!

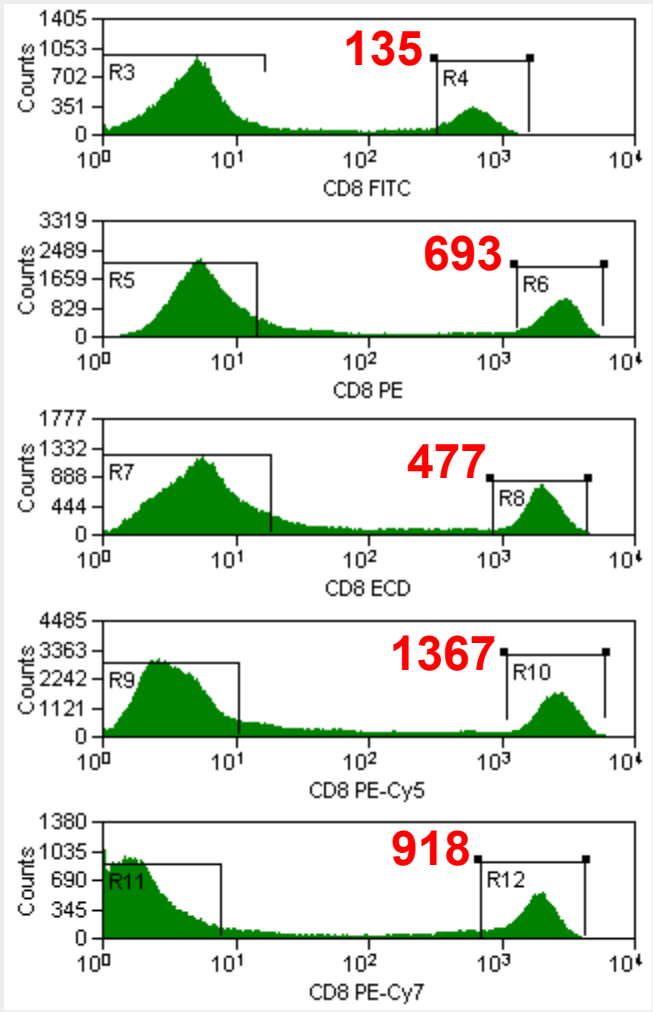
Pull the Noise From the Signal

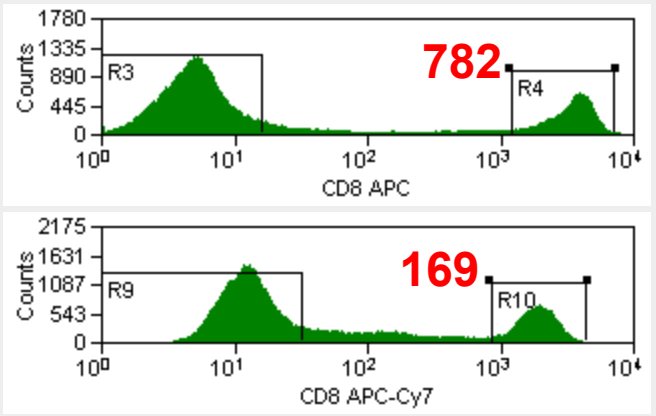
- Dump channel
- Unique location in multiparameter space
- Use the best fluorochrome for the most critical measurement

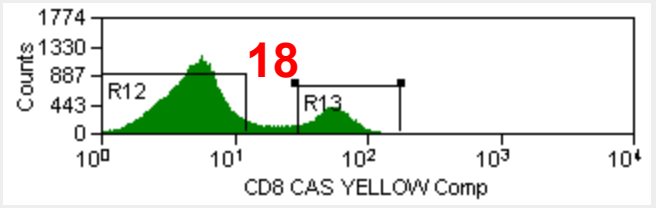
  PE has high quantum efficiency

  Red line used to excite APC and APC tandems excites less cellular autofluorescence

For a reagents available in several fluorochromes choose the one with the best signal to noise ratio for your critical measurement





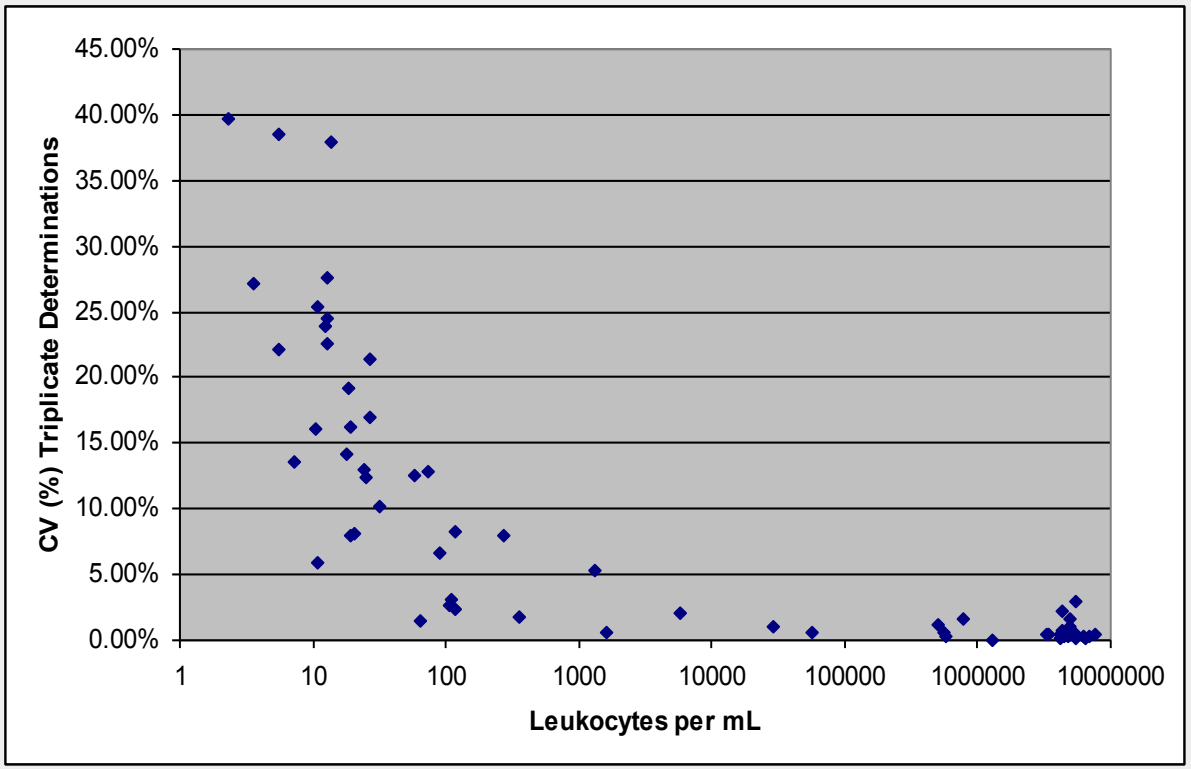


How Many Cells to Acquire

- Short answer: The rarer the event the more cells required
- Long answer: Depends on
  - Event frequency
  - Tightness of event cluster in multiparameter space
- You can determine the number empirically by determining the precision of replicate determinations
- No matter how many events you acquire, the limit of detection is governed by the signal to noise ratio
  -

Precision of Replicate Determinations

All events in three 5 mL aliquots of leukocyte depleted platelet product were acquired



Detection of leukocytes in filtered platelet components
Donnenberg et al Transfusion, 2000.

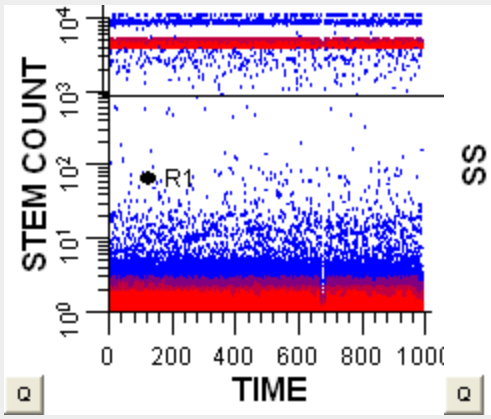
Autologous HSCT for SSc

- NIAMS-funded study Auto HSCT in rapidly progressing SSc
- CY (2g/m^2 over 24 hours), G-CSF (10 μg/kg/day sq) mobilization until WBC>2500
- Apheresis, T-depletion by Isolex 300i positive CD34 selection, negative CD3 depletion

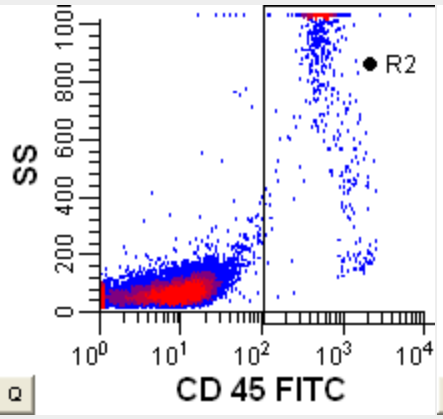
Patient	CD34 Dose/kg	CD34 Purity	CD3 Dose/kg	Log Depletion
1	1.18E+07	99.7%	9.76E+02	5.2
2	9.46E+06	96.4%	5.71E+01	6.3
3	1.30E+07	97.4%	4.54E+03	4.7
4	4.80E+06	98.9%	6.34E+02	4.8
5	1.07E+07	97.4%	4.85E+02	5.3
Mean	**9.94E+06**	**98.0%**	**1.34E+03**	**5.2**
SD	**3.16E+06**	**1.3%**	**1.82E+03**	**0.6**

Single Platform Absolute T-cell Count Day +3

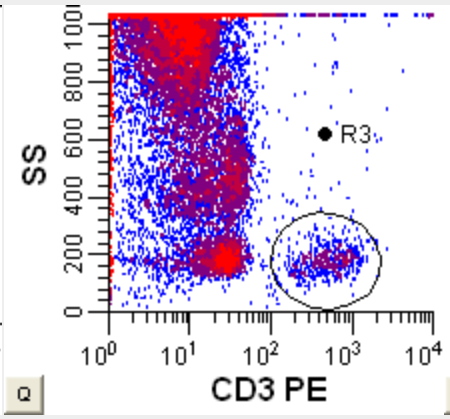
CAL BEADS (UNGATED)



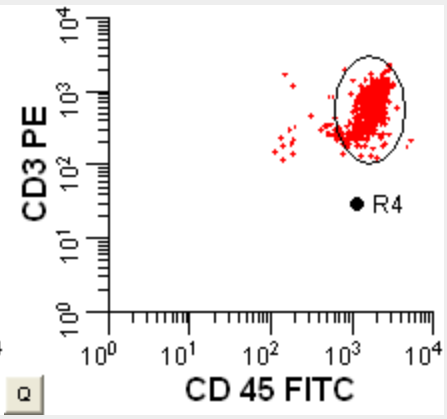
WBC (NOT R1)



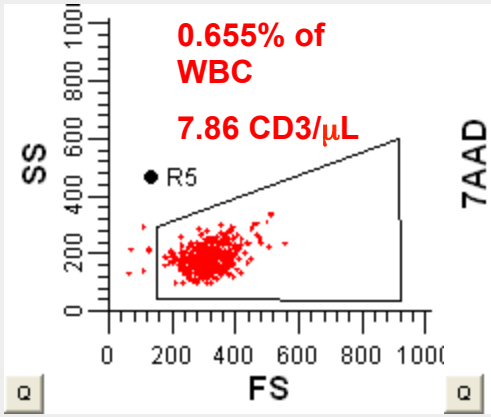
CD3+ (R2 NOT R1)



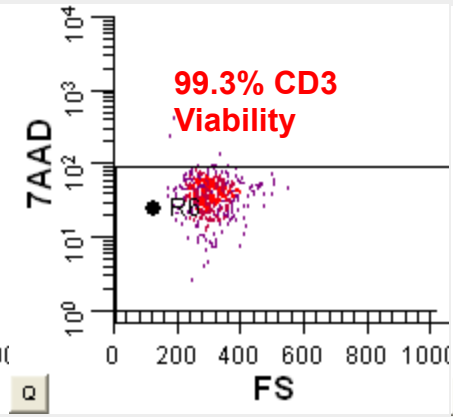
CD3+ CD45^{HI} (R2 R3 NOT R1)



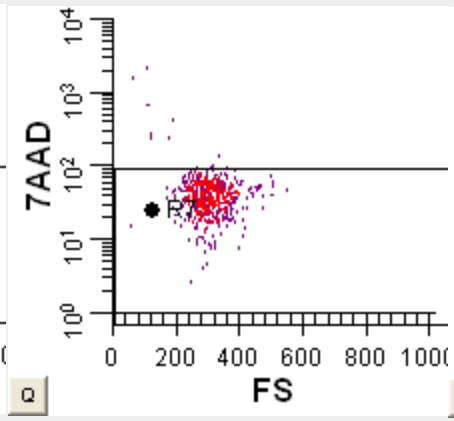
CD3+CD45^{HI} FS^{INT} (R2 R3 R4 NOT R1)



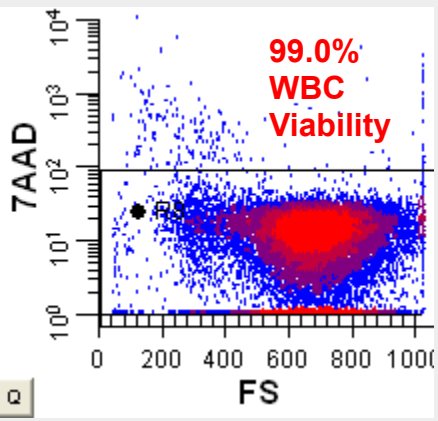
CD3+ VIABILITY (R2 R3 R4 R5 NOT R1)



TOTAL CD3+ VIABILITY (R2 R3 R4 NOT R1)



CD45 VIABILITY (R2 NOT R1)

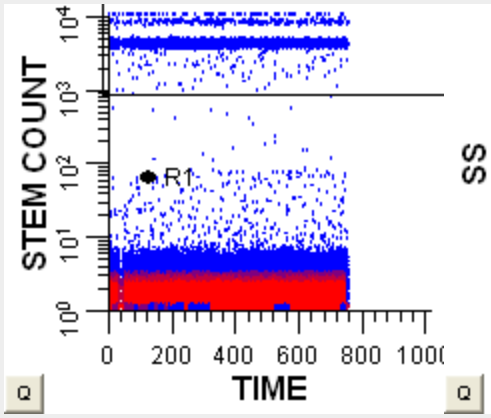


C5337.LMD

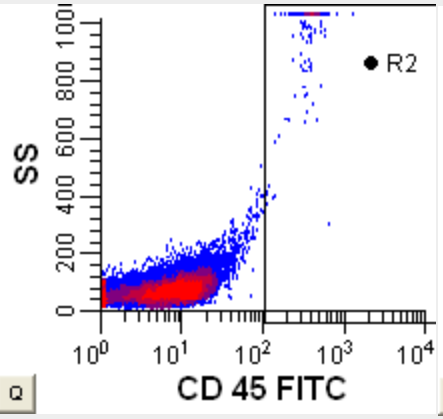
0.2 mL whole product x 6-plicate exhaustively acquired

Single Platform Absolute T-cell Count Day 0

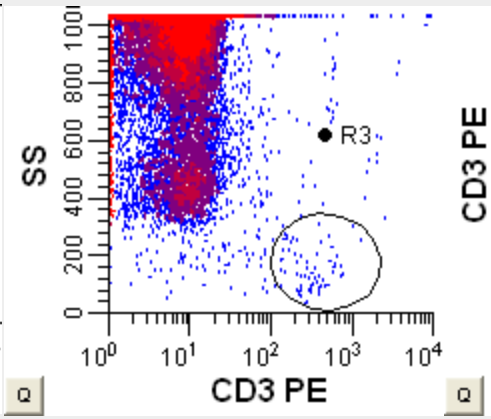
CAL BEADS (UNGATED)



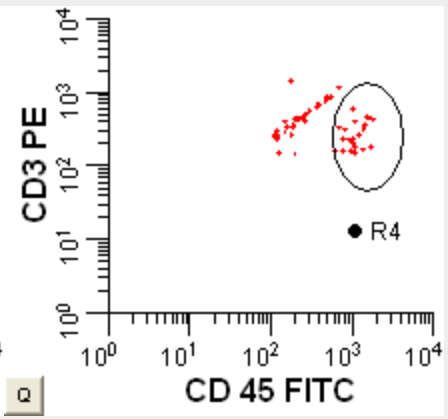
WBC (NOT R1)



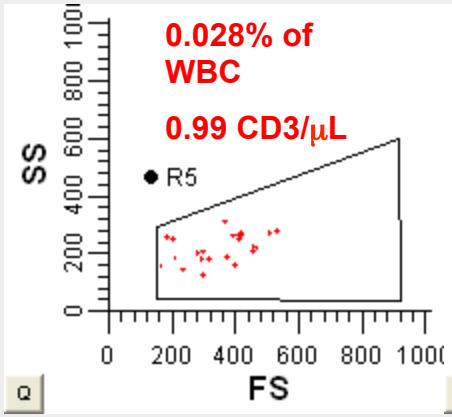
CD3+ (R2 NOT R1)



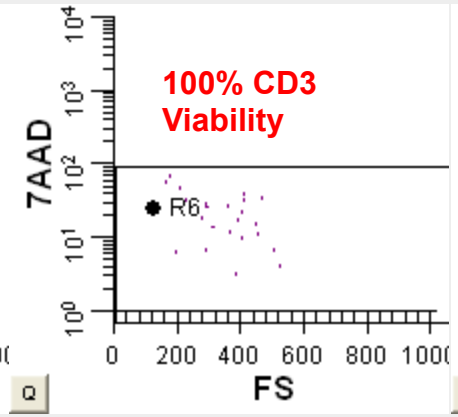
CD3+ CD45^{HI} (R2 R3 NOT R1)



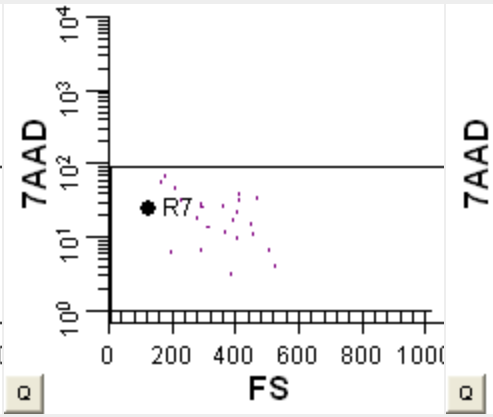
CD3+CD45^{HI} FS^{INT}
(R2 R3 R4 NOT R1)



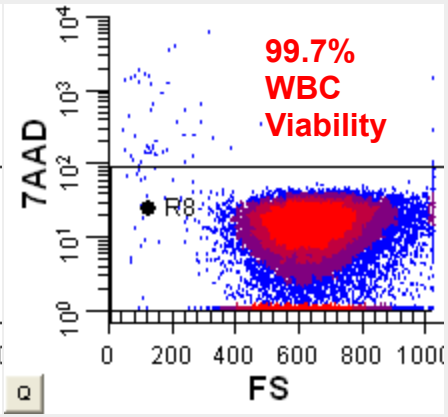
CD3+ VIABILITY
(R2 R3 R4 R5 NOT R1)



TOTAL CD3+ VIABILITY
(R2 R3 R4 NOT R1)

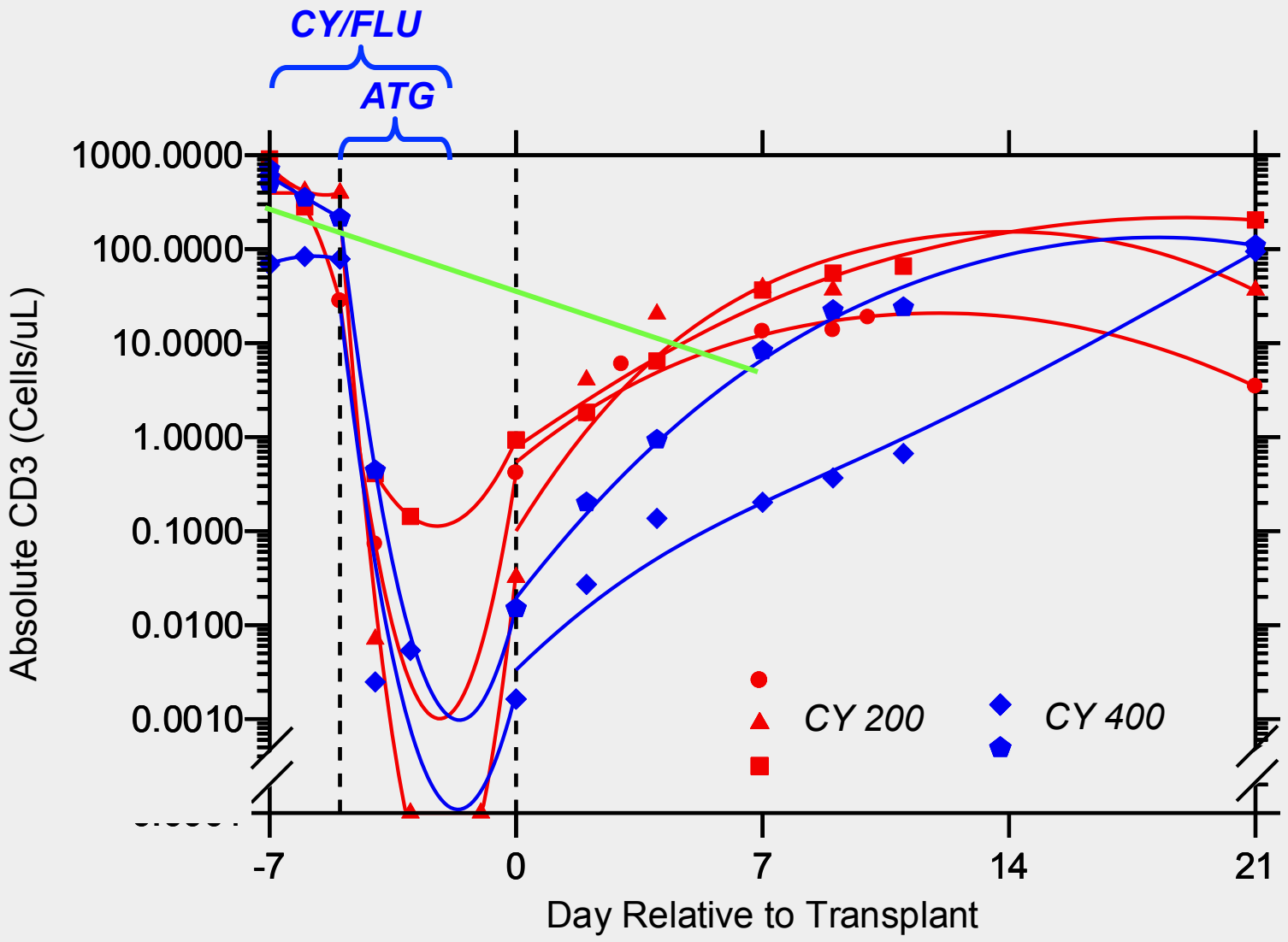


CD45 VIABILITY
(R2 NOT R1)



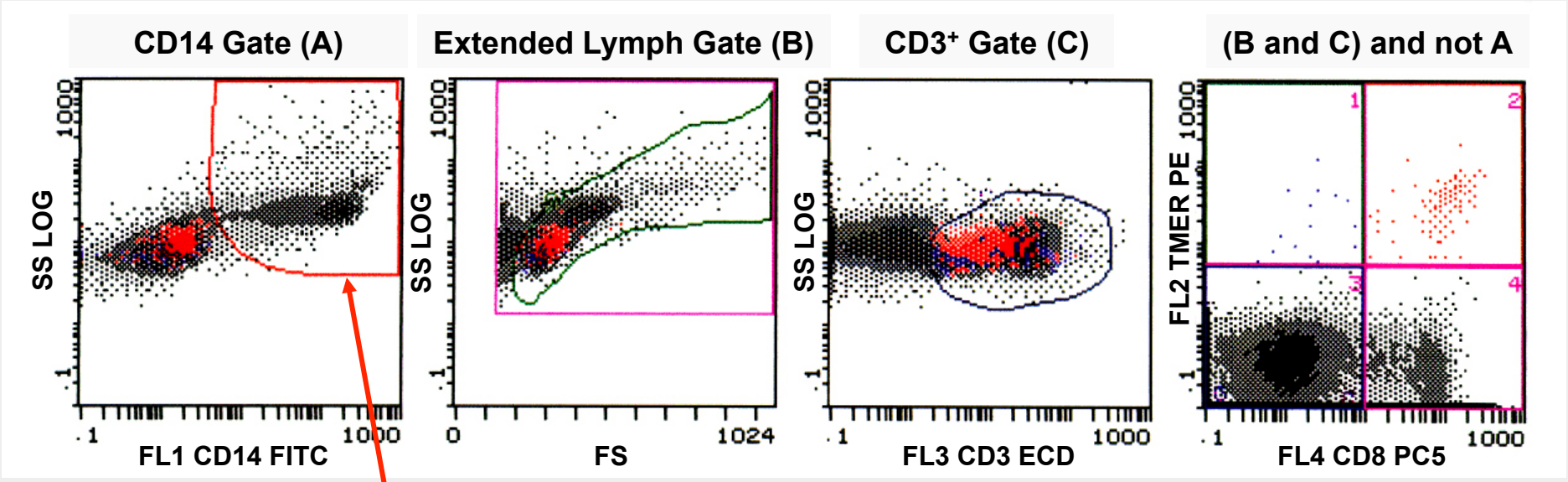
C5256.LMD

T-cell Kill Trajectory CY/Flu/ATG



Response of healthy $A2^+$ control subject to influenza tetramer (peptide GILGFVFTL)

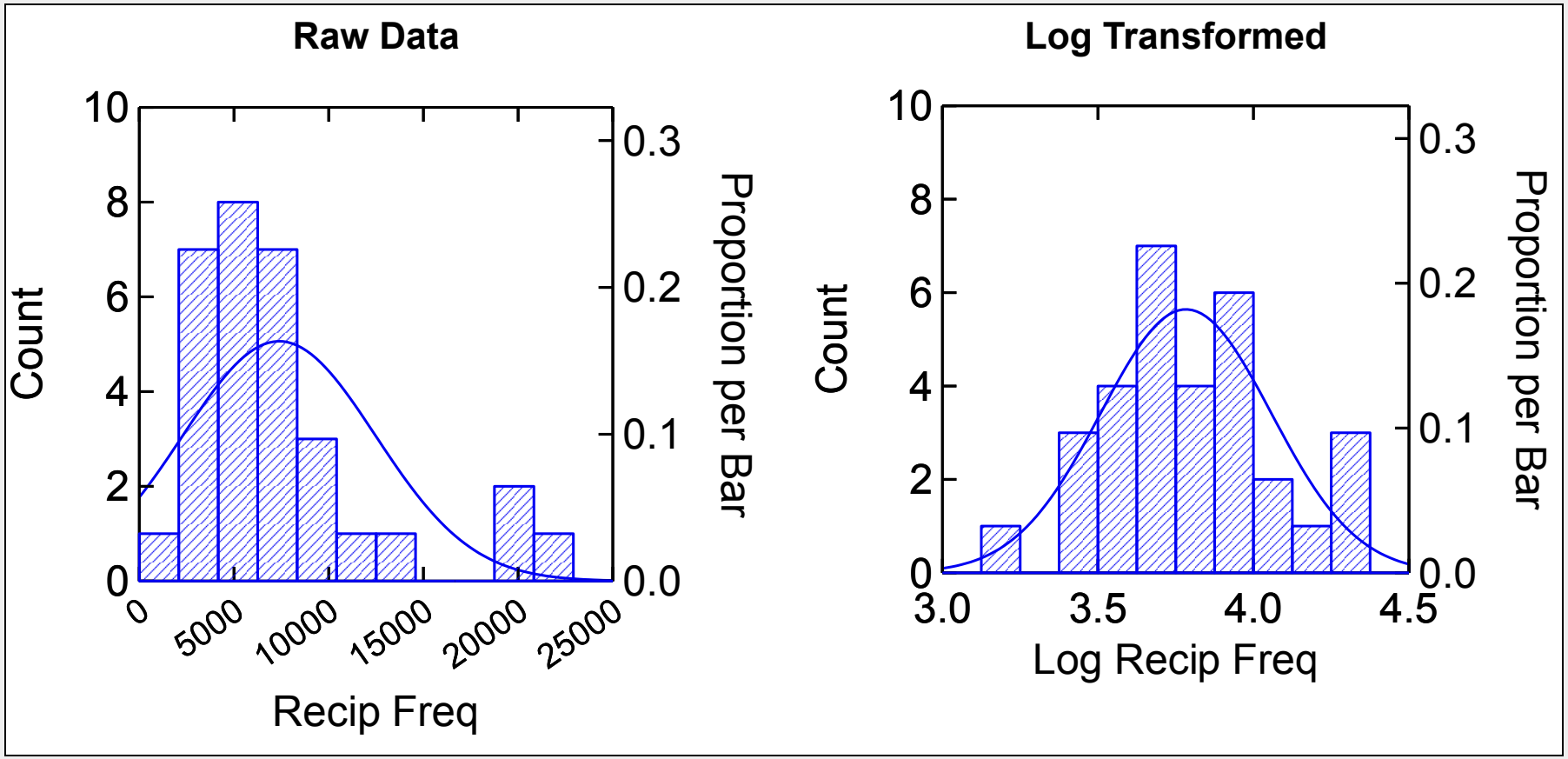
NOT CD14 NOT dead cells AND CD3

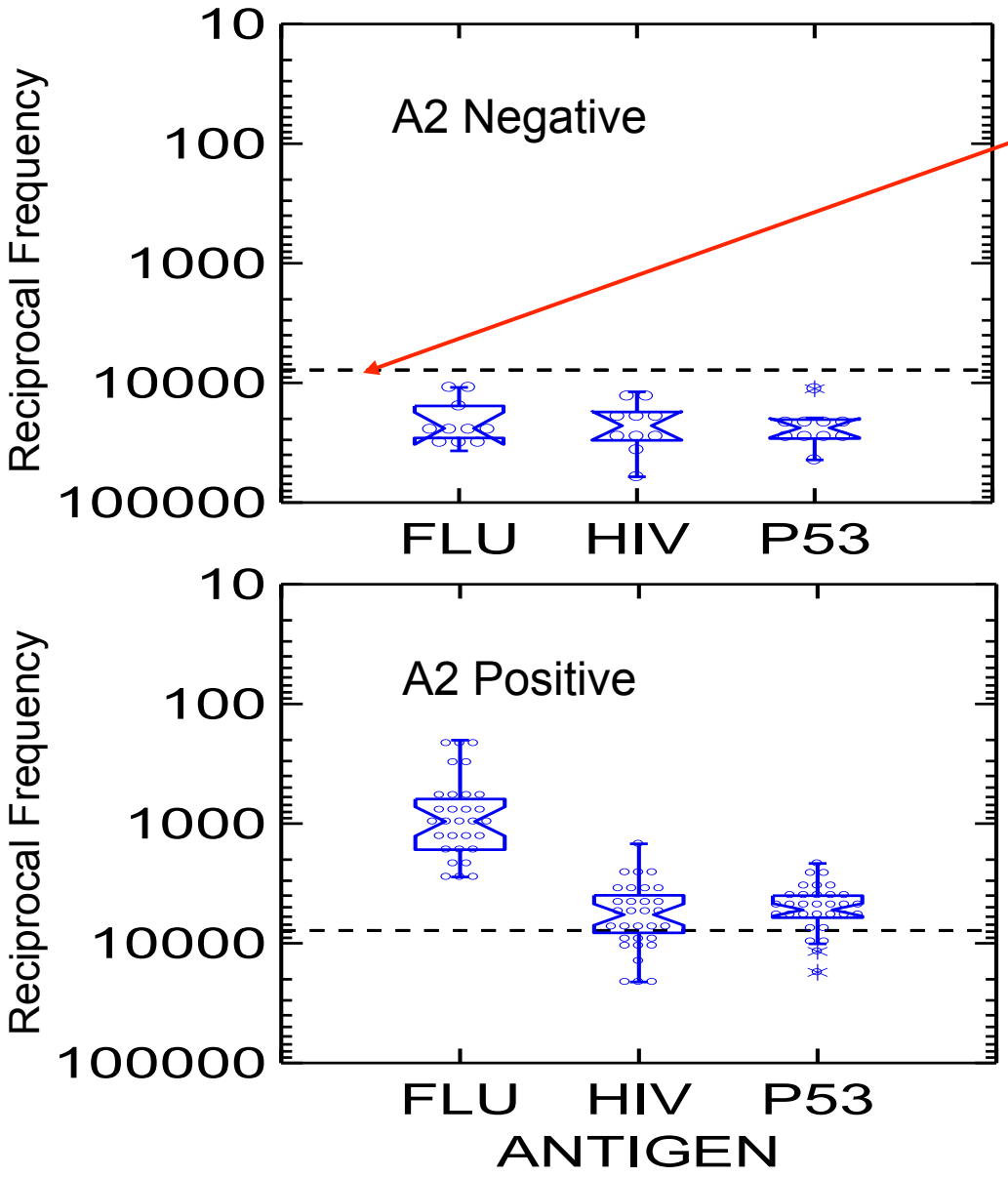


Dump gate

Cytometry, 41:321-328, 2000

Reciprocal frequencies (and % positive) are log normally distributed





Lower limit of detection defined as log mean of HLA-A2 negative subjects plus 2 SD = 1/7000

TCR Vβ Usage in T-cell subsets of HIV infected subjects

Healthy controls

N	**21 male subjects**
Mean Age	**42.5 years**

HIV-infected patients (on HAART) recruited from the Multicenter AIDS Cohort Study, Pittsburgh center.

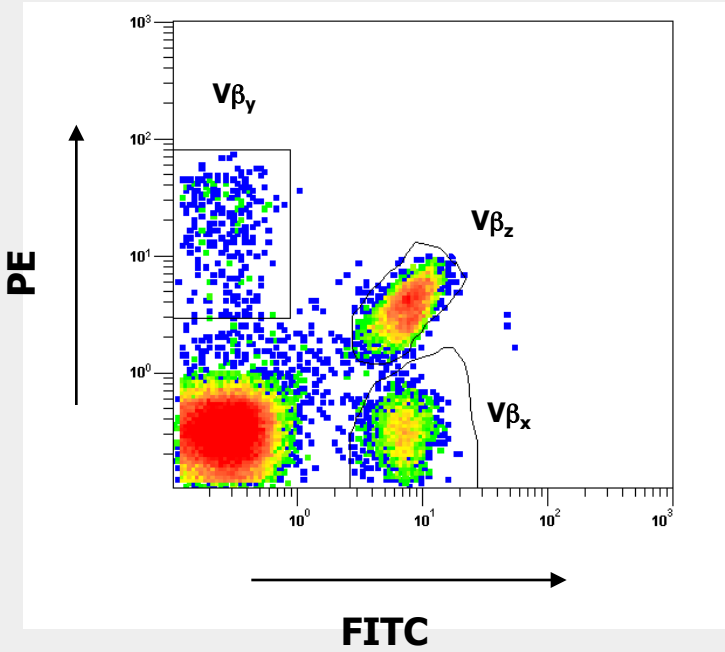
HIV+ patients

N	**39 male subjects**
Mean Age	**43.8 years**
CD4 > 500	**21**
CD4 (200-500)	**13**
CD4 < 200	**5**
Viral Load < 1000	**18**
Viral Load ≥ 1000	**21**

Tube	Vβ	Fluorochrome
A	Vb5.3 Vb7.1 Vb3	PE PE + FITC FITC
B	Vb9 Vb17 Vb16	PE PE + FITC FITC
C	Vb18 Vb5.1 Vb20	PE PE + FITC FITC
D	Vb13.1 Vb13.6 Vb8	PE PE + FITC FITC
E	Vb5.2 Vb2 Vb12	PE PE + FITC FITC
F	Vb23 Vb1 Vb21.3	PE PE + FITC FITC
G	Vb11 Vb22 Vb14	PE PE + FITC FITC
H	Vb13.2 Vb4 Vb7.2	PE PE + FITC FITC

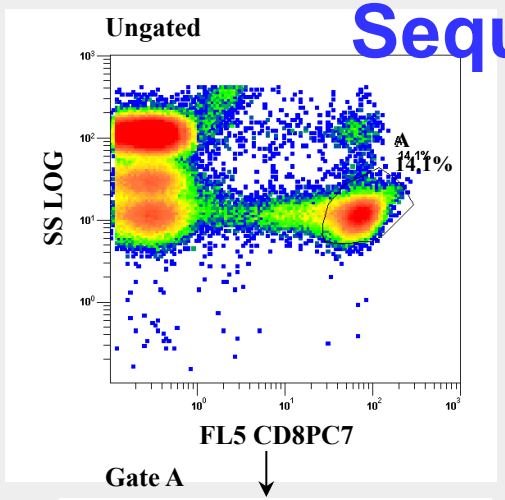
(24 different specificities, about 70% coverage of normal human TCR Vβ repertoire).

Multiplexed detection of TCR Vβ families.



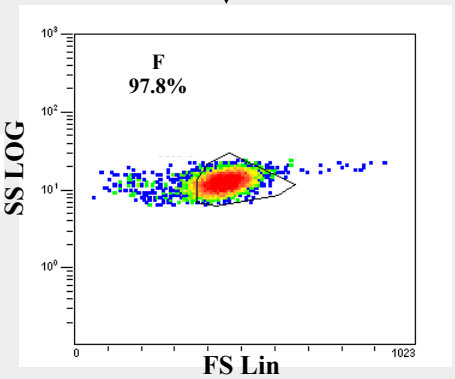
6 fluorescent parameters on a 5-color instrument

Sequential 5-color gating strategy

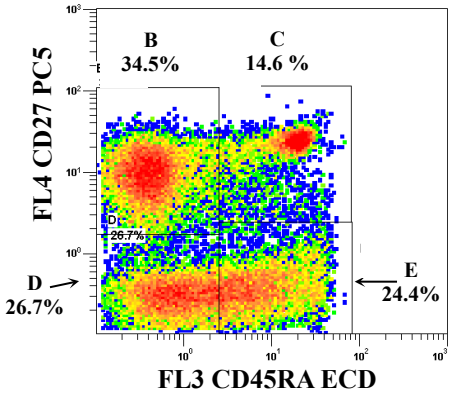


Classifiers

Gate A



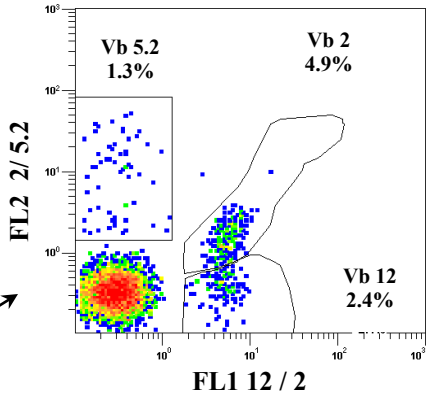
Gate A and F



Outcomes

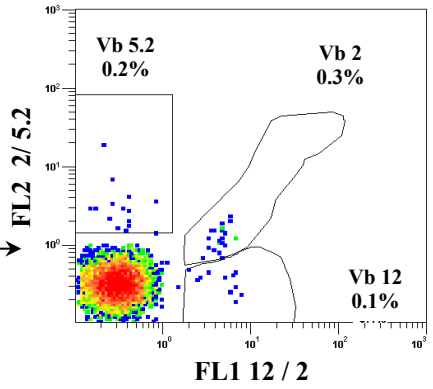
Naive

Gate A and F and C



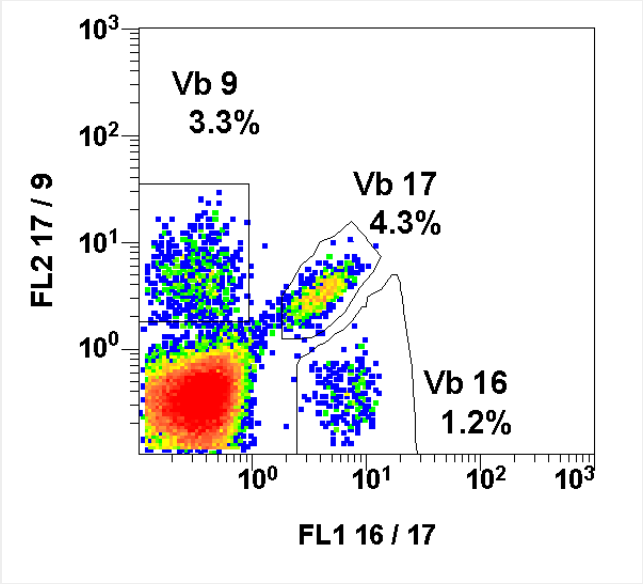
Effectors

Gate A and F and E

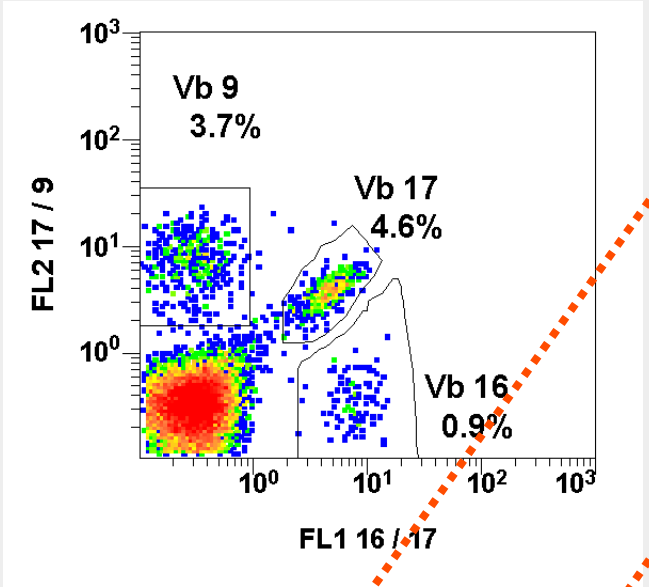


How to define expansions and deletions?

CD4 CD45RA- CD27+



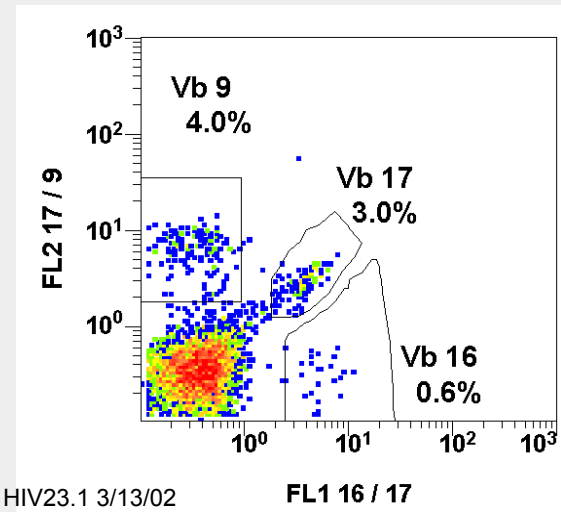
CD4 CD45RA+ CD27+



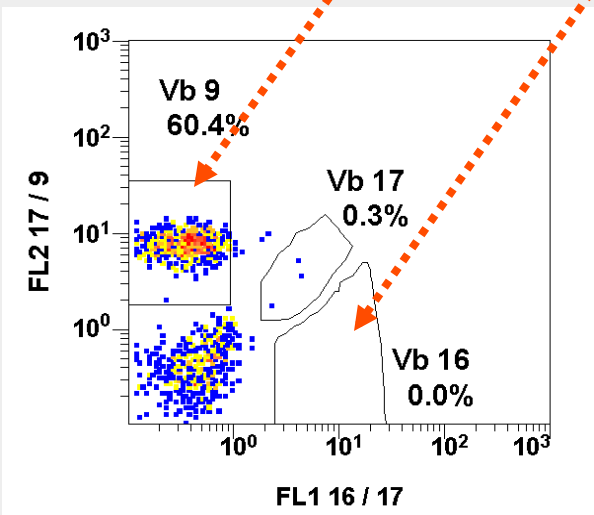
Vb9 Expansion ?

Vb16 Deletion ?

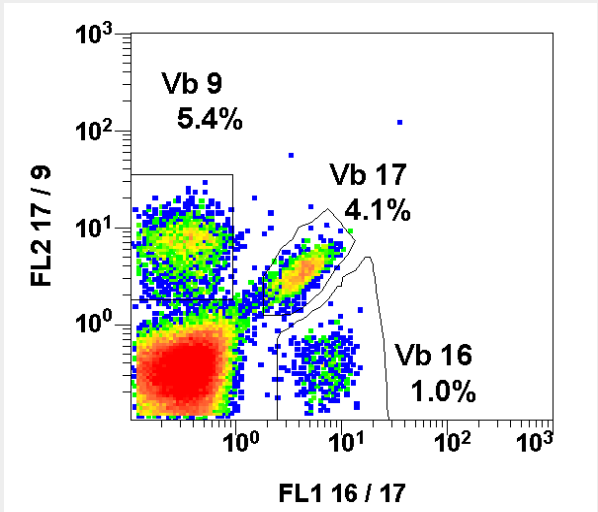
CD4 CD45RA- CD27-



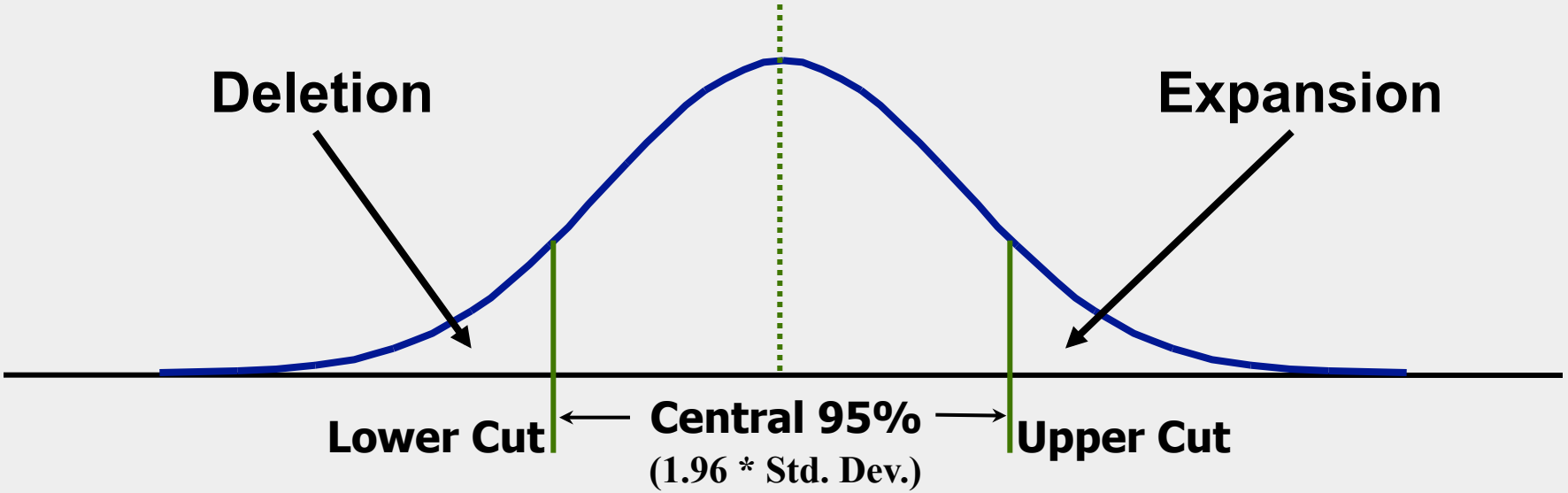
CD4 CD45RA+ CD27-



Total CD4

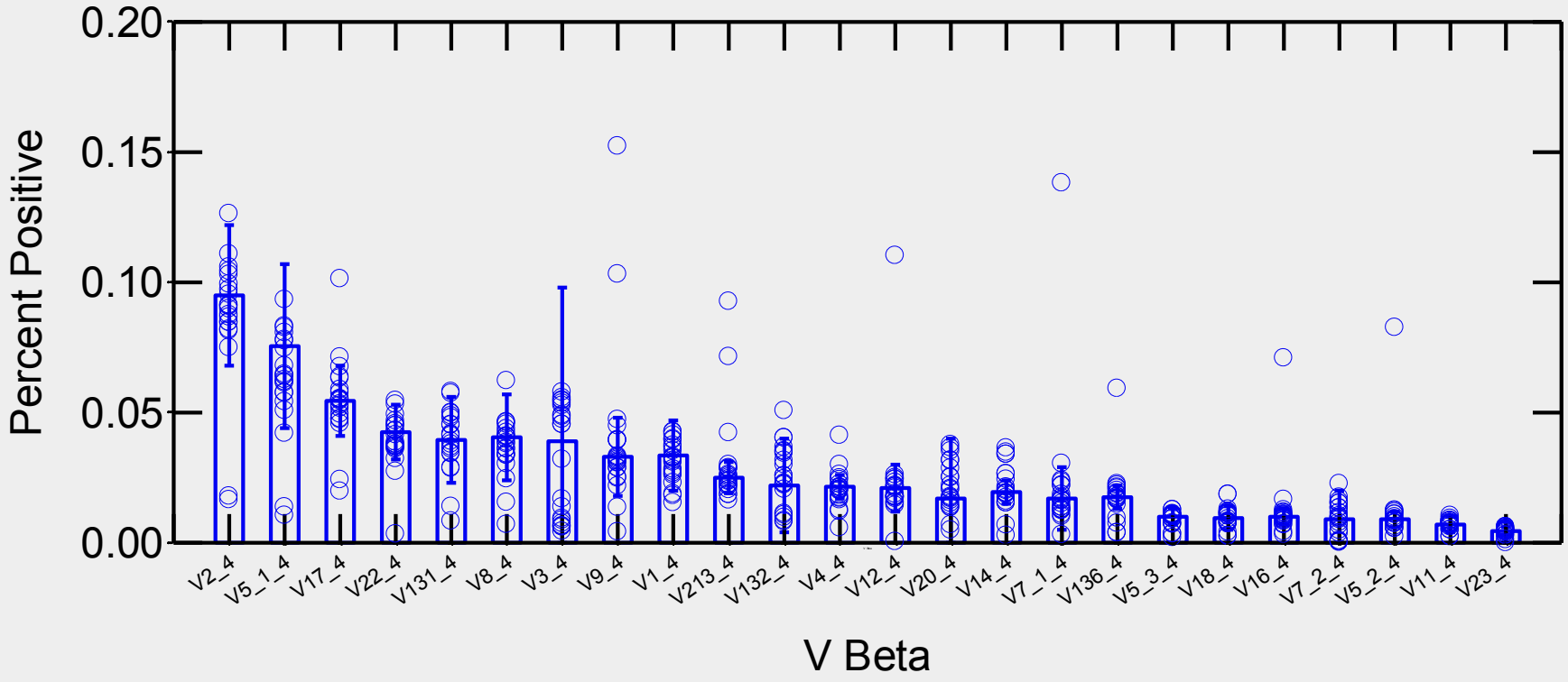


Predictions of the Normal Model



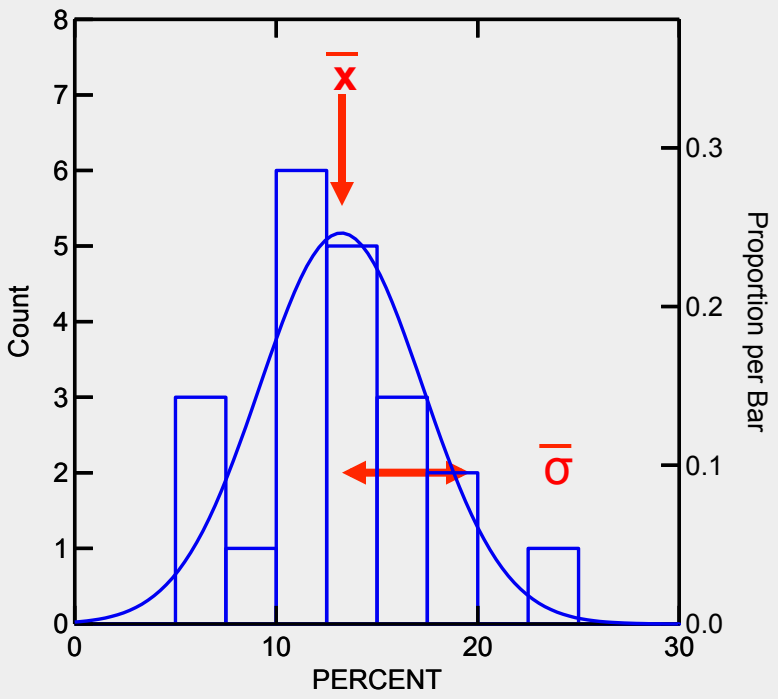
For our control dataset:
2.5% of 24 possible Vb * 21 control subjects = 12.6 expansions and 12.6 deletions

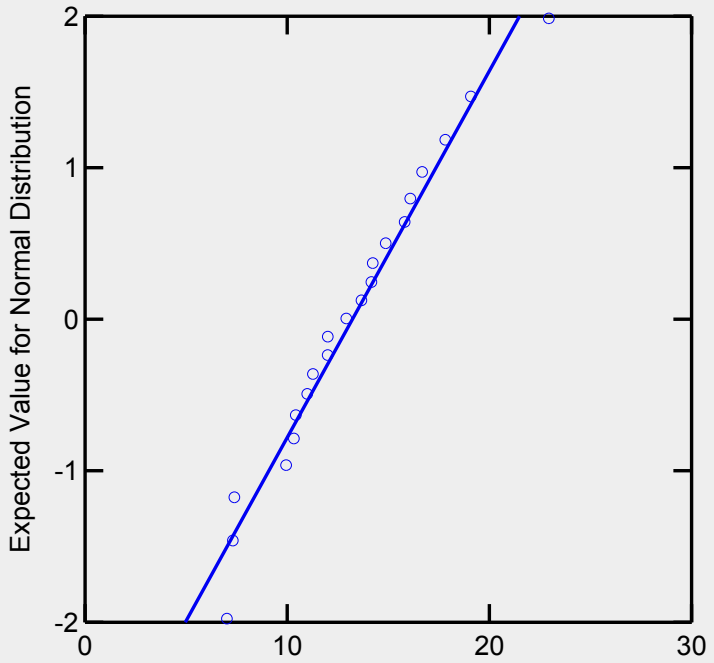
Control Subject VB Usage, Total CD4



25 expansions and 3 deletions!

CD4+ Percent is Normally Distributed

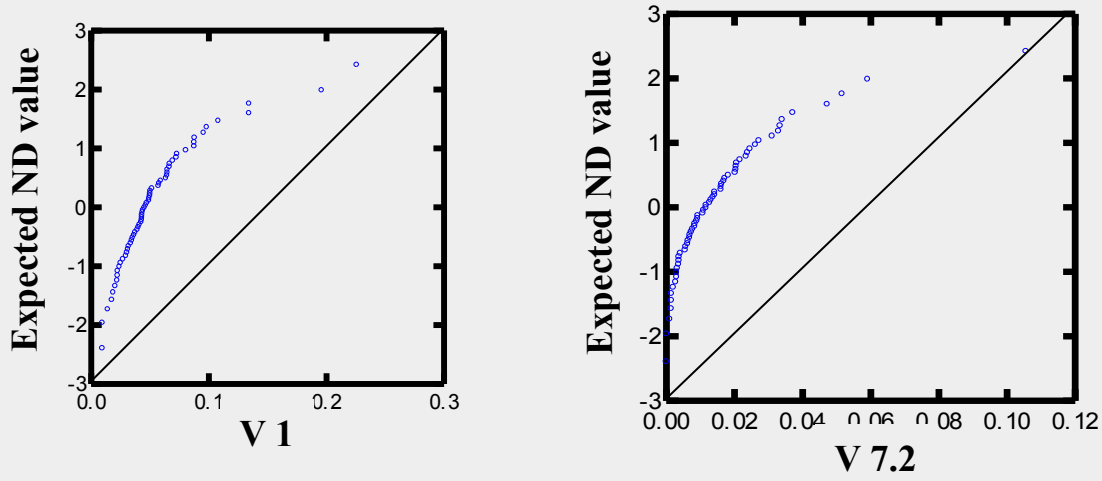




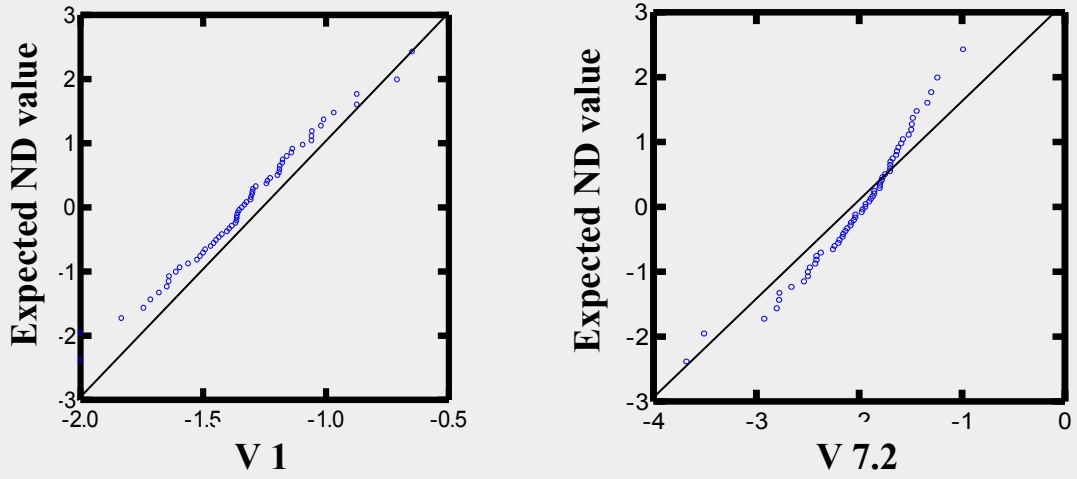
Percent CD4+ →

Vβ usage is log-normally distributed

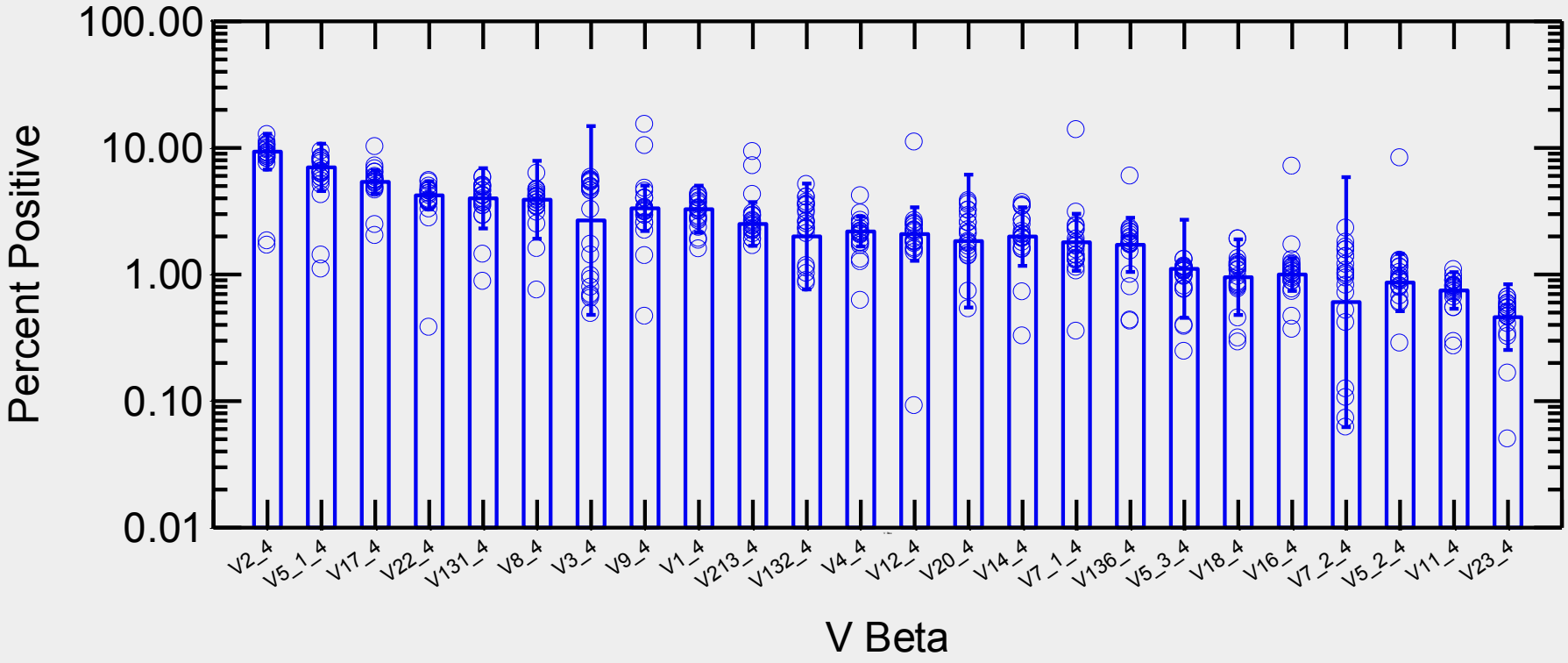
A. Linear Scale



B. Log Transformed



Control Subject VB usage, total CD4

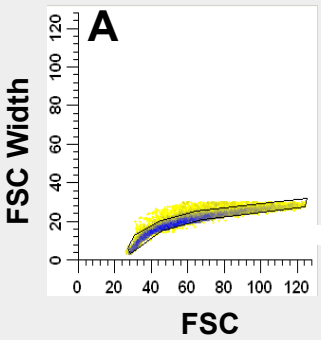


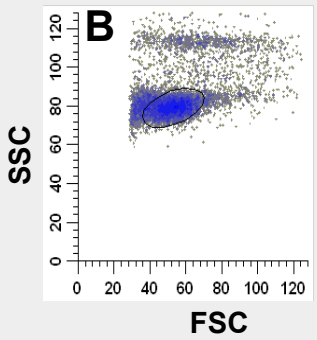
12 expansions and 14 deletions (expected 12.6)

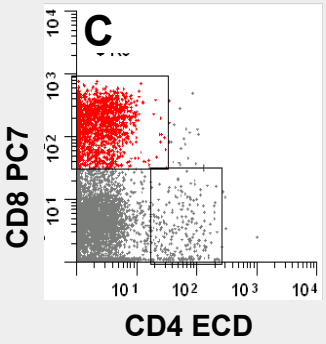
Conclusions

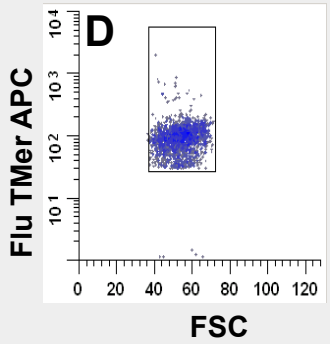
- Vβ usage is log-normally distributed.
- Failure to use log transformed data results in an overestimate of expansions, an underestimate of deletions, and loss of power to detect significant differences using parametric tests.

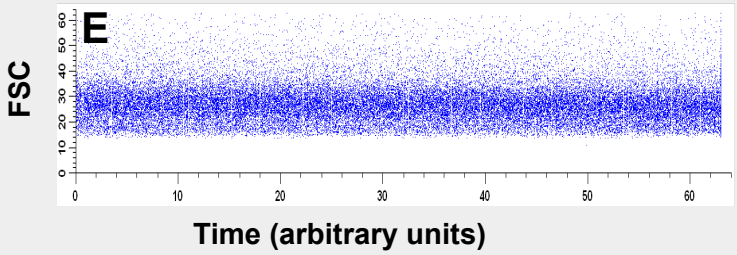
Measuring Vβ Repertoire on TMer+ T-cells

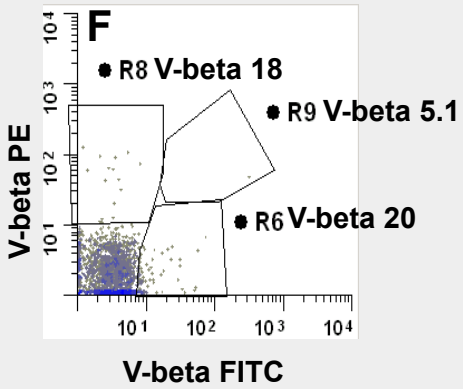


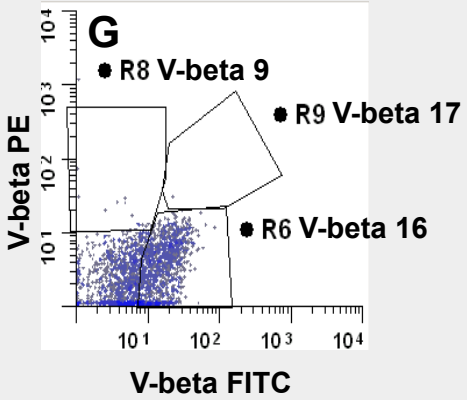


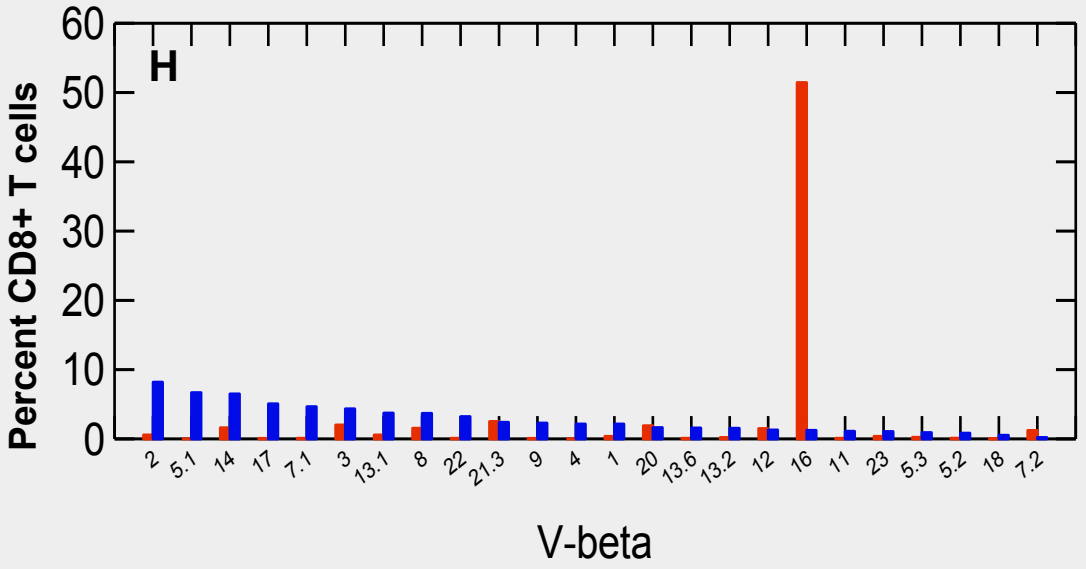








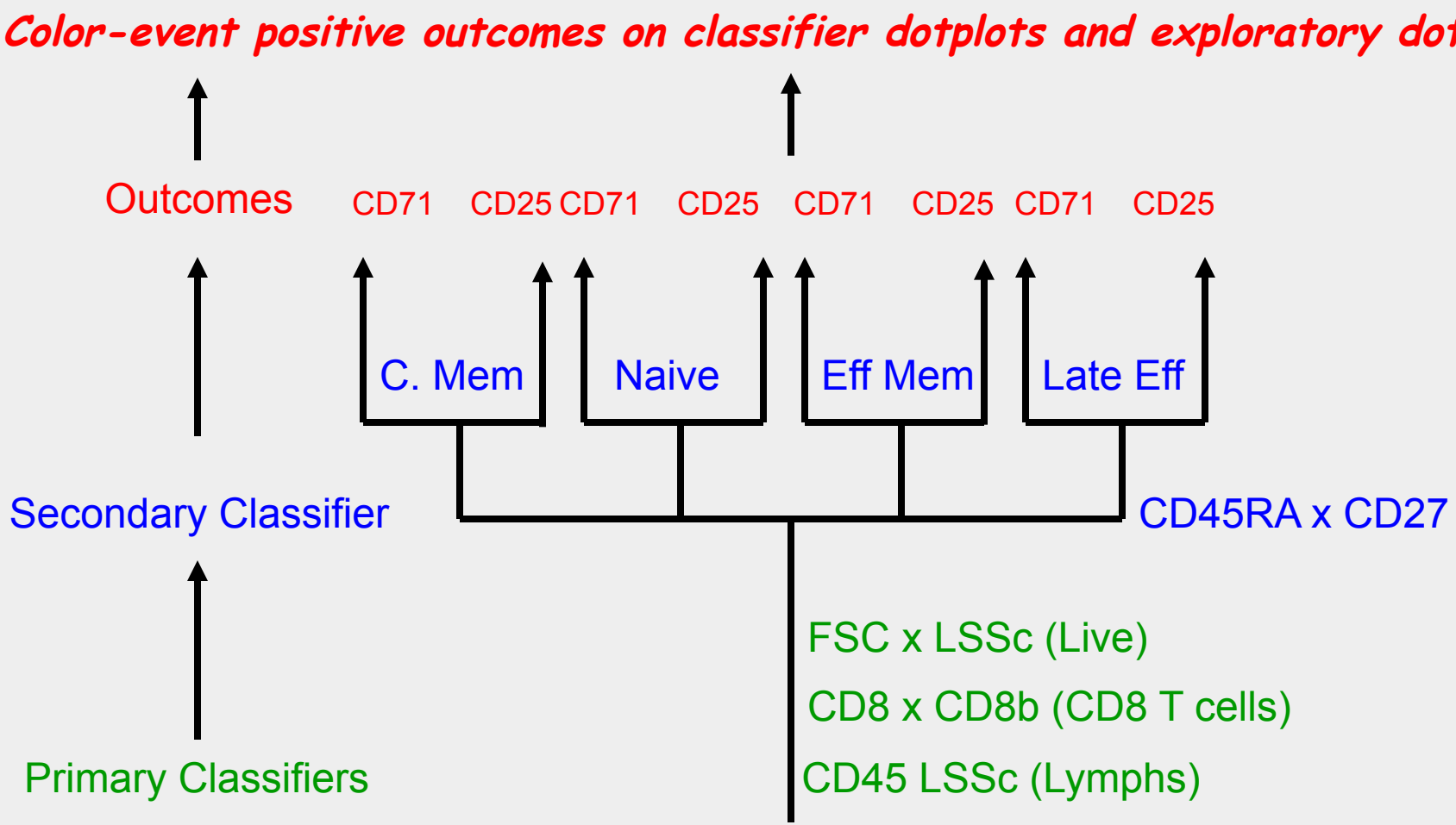


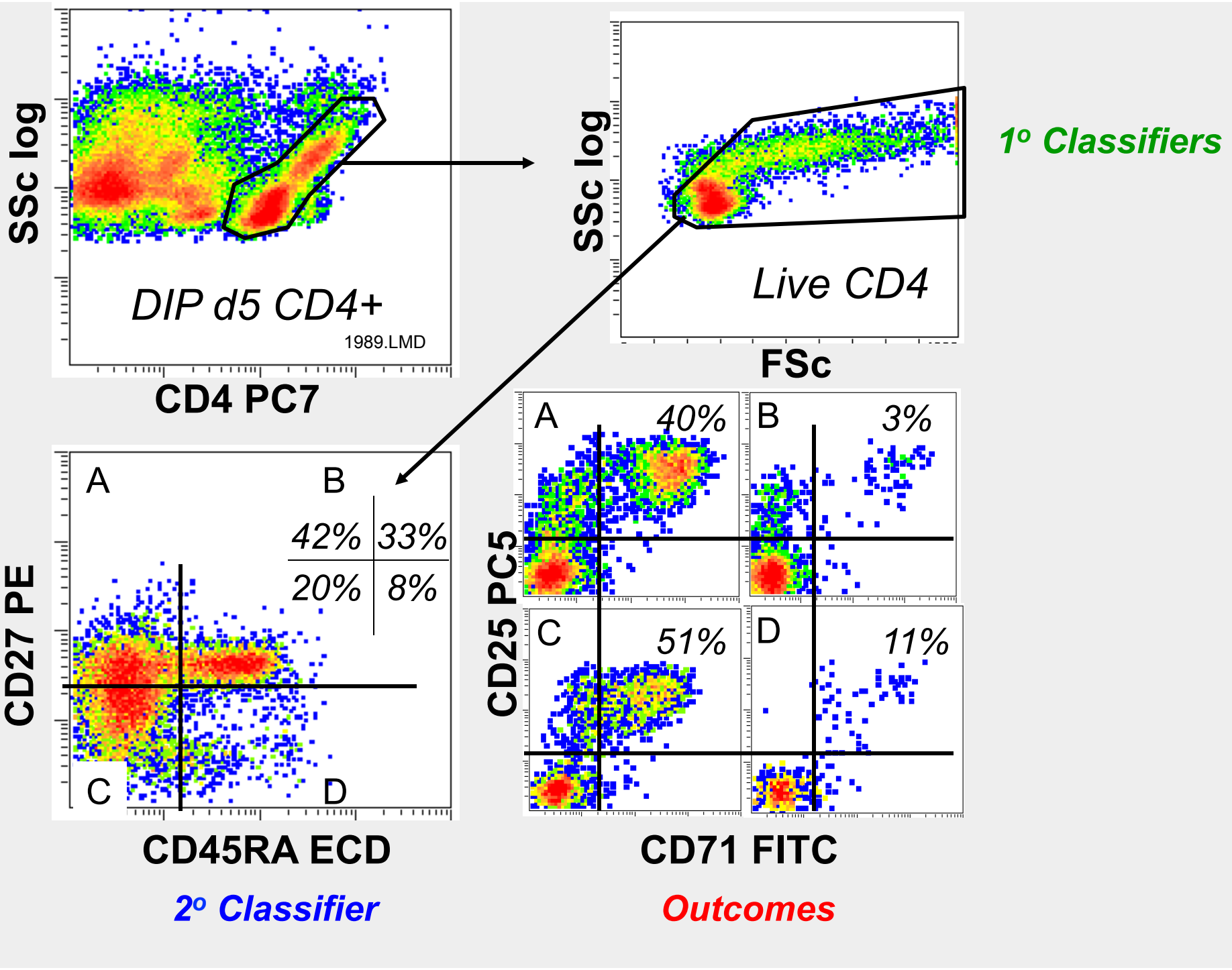


For Multicolor Analysis, Group Parameters as Classifiers or Outcomes

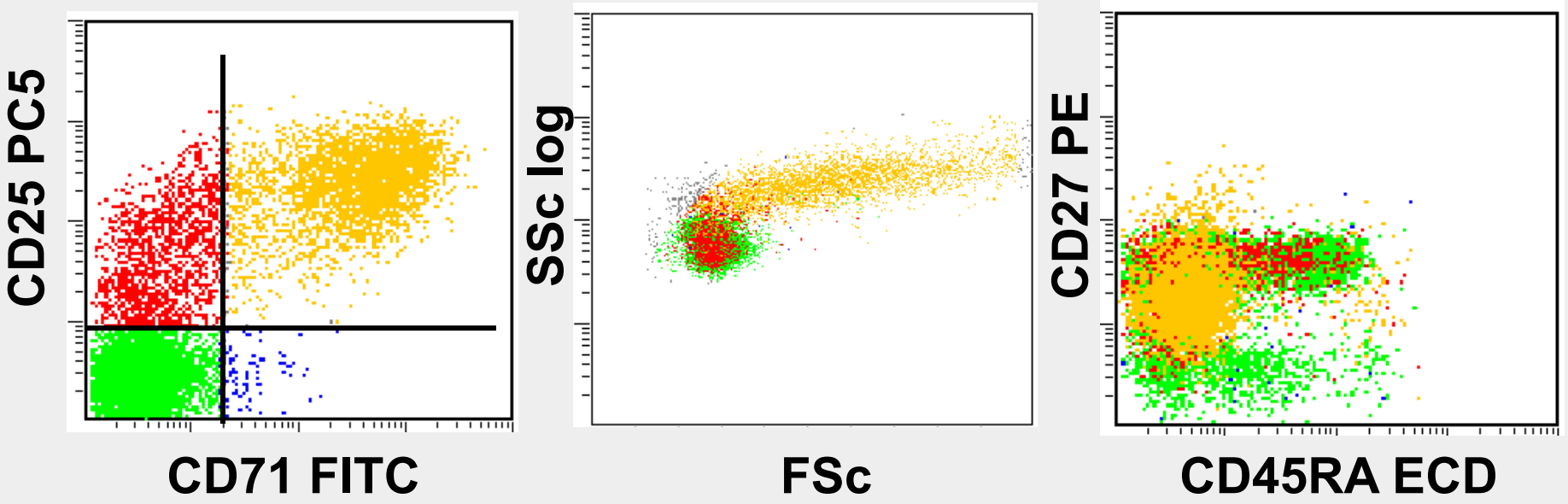
- Classifiers
  - Primary do not branch (e.g. singlets AND CD4+ low side scatter)
  - Secondary branch (e.g. memory, naïve and effector populations based on CD45RA and CD27 expression)
- Outcomes
  - Measured on each population defined by the classifiers
  - Color event and backgate to the classifiers

Hierarchical Gating Strategy for Multiparameter Functional Flow




SSc log
DIP d5 CD4+
1989.LMD
CD4 PC7
SSc log
Live CD4
FSc
1º Classifiers
A
B
42%
33%
20%
8%
C
D
CD27 PE
CD45RA ECD
2º Classifier
A
40%
B
3%
C
51%
D
11%
CD25 PC5
CD71 FITC
Outcomes

Gated on 1º Classifiers



Color Event the Outcomes ⟶ Project onto classifier dotplots

The Approach

- CMV high A2.01+ donor
- CMV pp65 tetramer added live during acquisition (kinetic measurements of tetramer binding)
- Kinetic measurement of calcium flux measured by Indo-1 ratio
- Sample maintained at 37° during 30 min acquisition
- High resolution immunophenotyping using a hierarchical gating strategy
- Beckman-Coulter NextGen prototype cytometer

The Panels

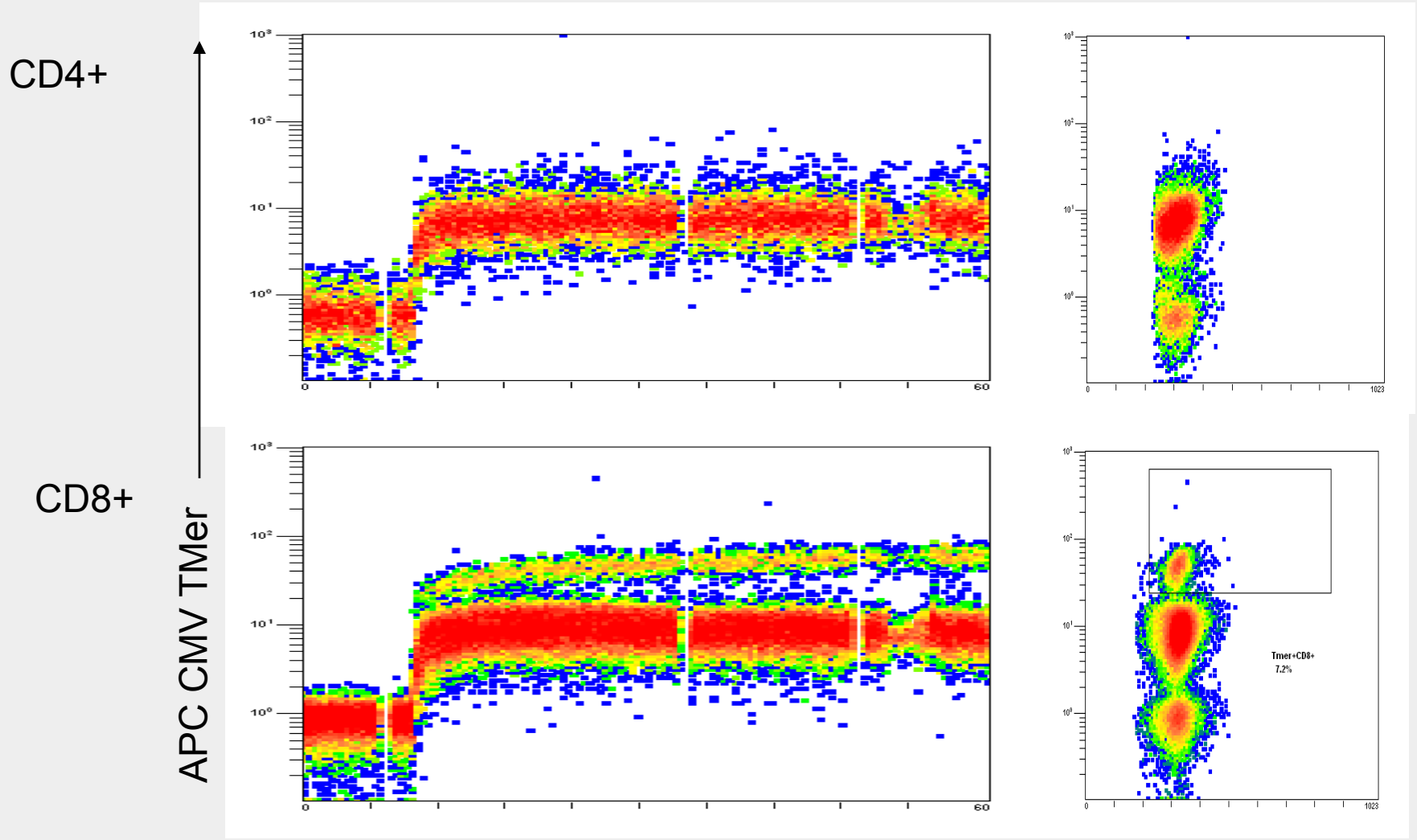
UV (355)	
1	2
Indo Short	Indo Long

Violet (408)	
3	4
Pacific Blue	Cascade Yellow
x	x
C1.7	CD45
C1.7	CD45

Blue (488)				
5	6	7	8	9
FITC	PE	ECD	PE-Cy5.5	PE-Cy7
CXCR4	x	CD8	x	CD4
CD7	CD95	CD45RA	CD8	CD8b
V-beta	V-beta	CD45RA	CD69	CD8

Red (635)		
10	11	12
APC	APC-Cy5	APC-Cy7
CMV TMER	x	x
CMV TMER	CD27	CD28
CMV TMER	CD27	CD28

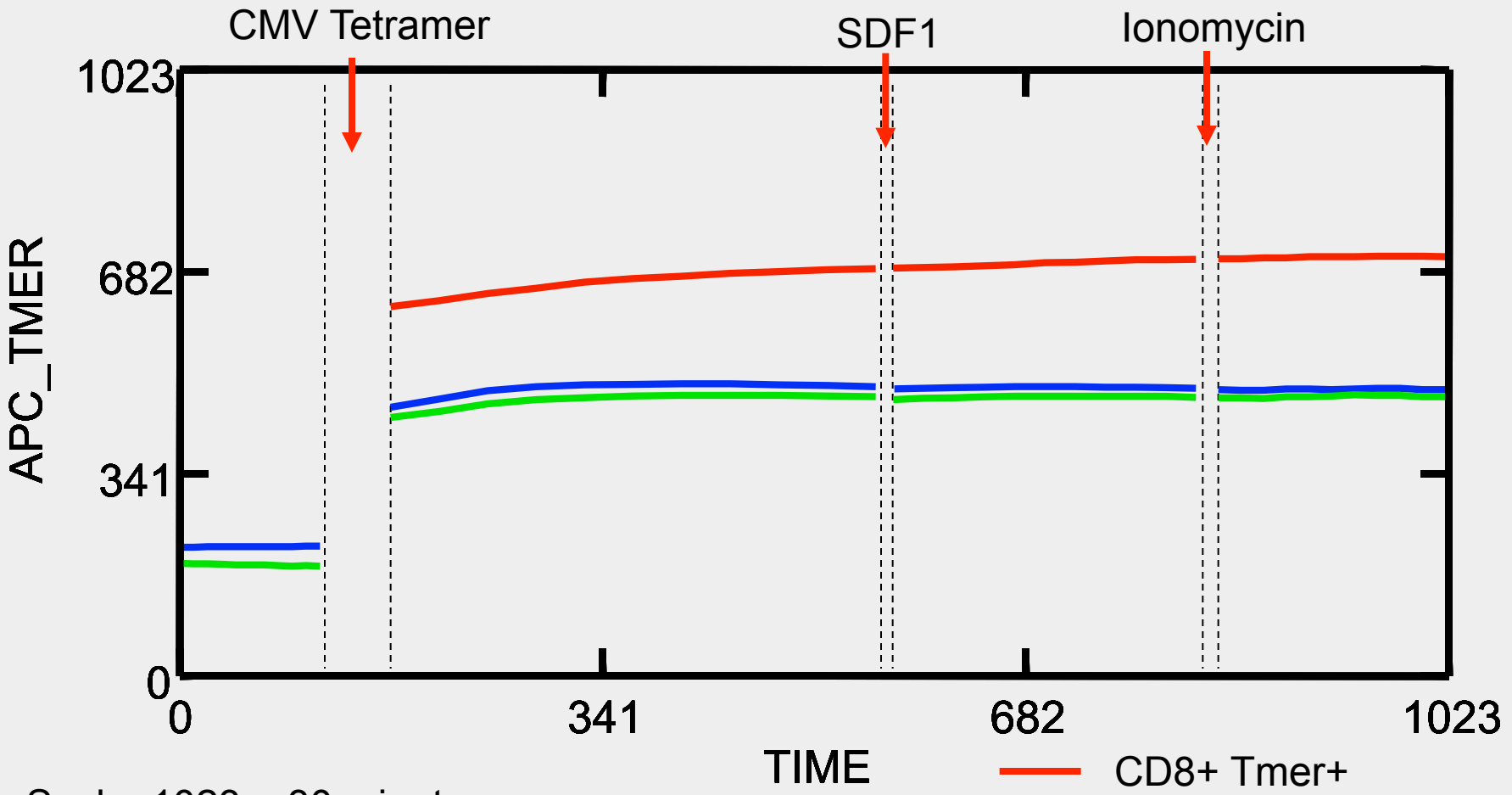
Kinetics of Tetramer Binding



CMV pp65 - NLVPMVATV

473.lmd

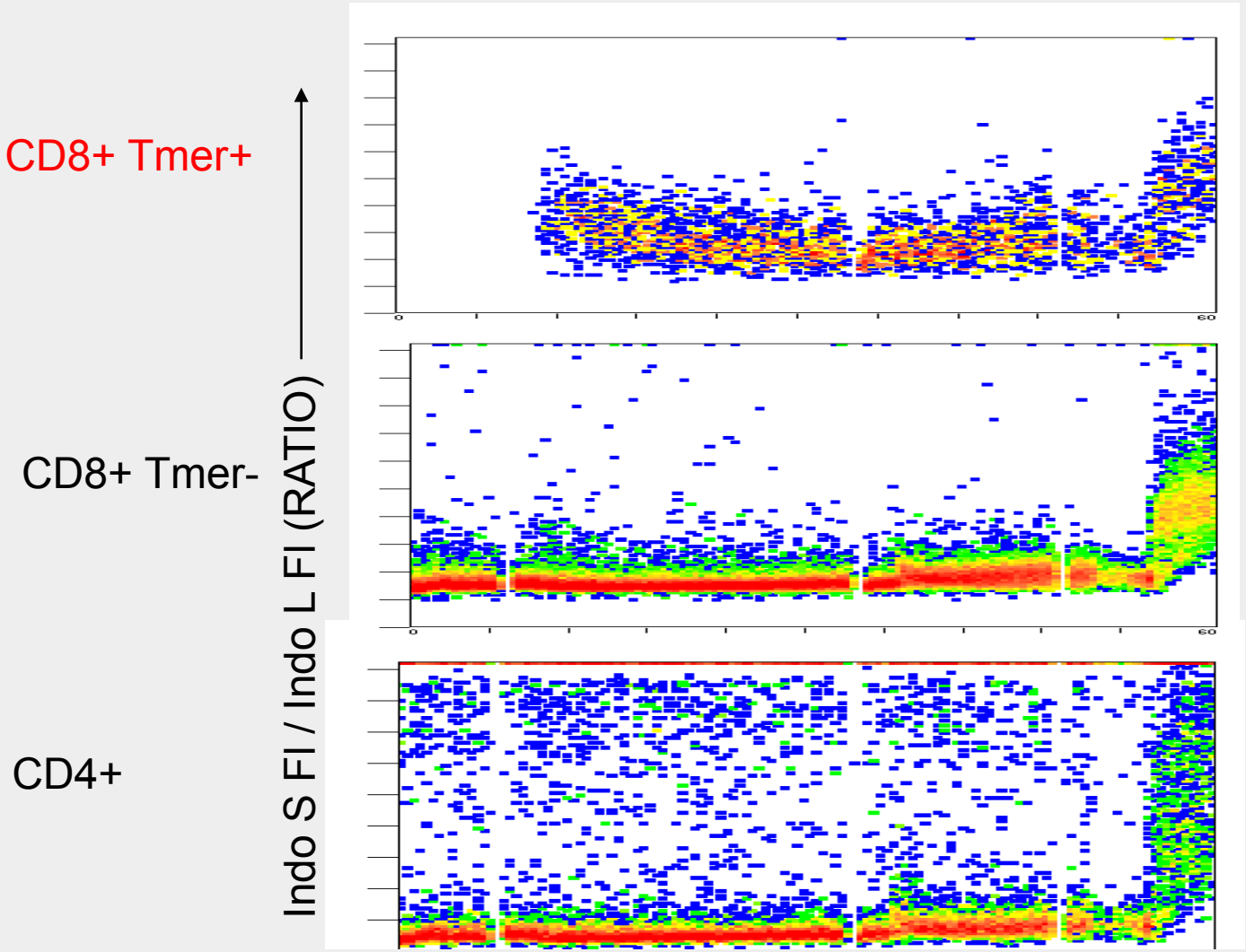
Kinetics of Tetramer Binding



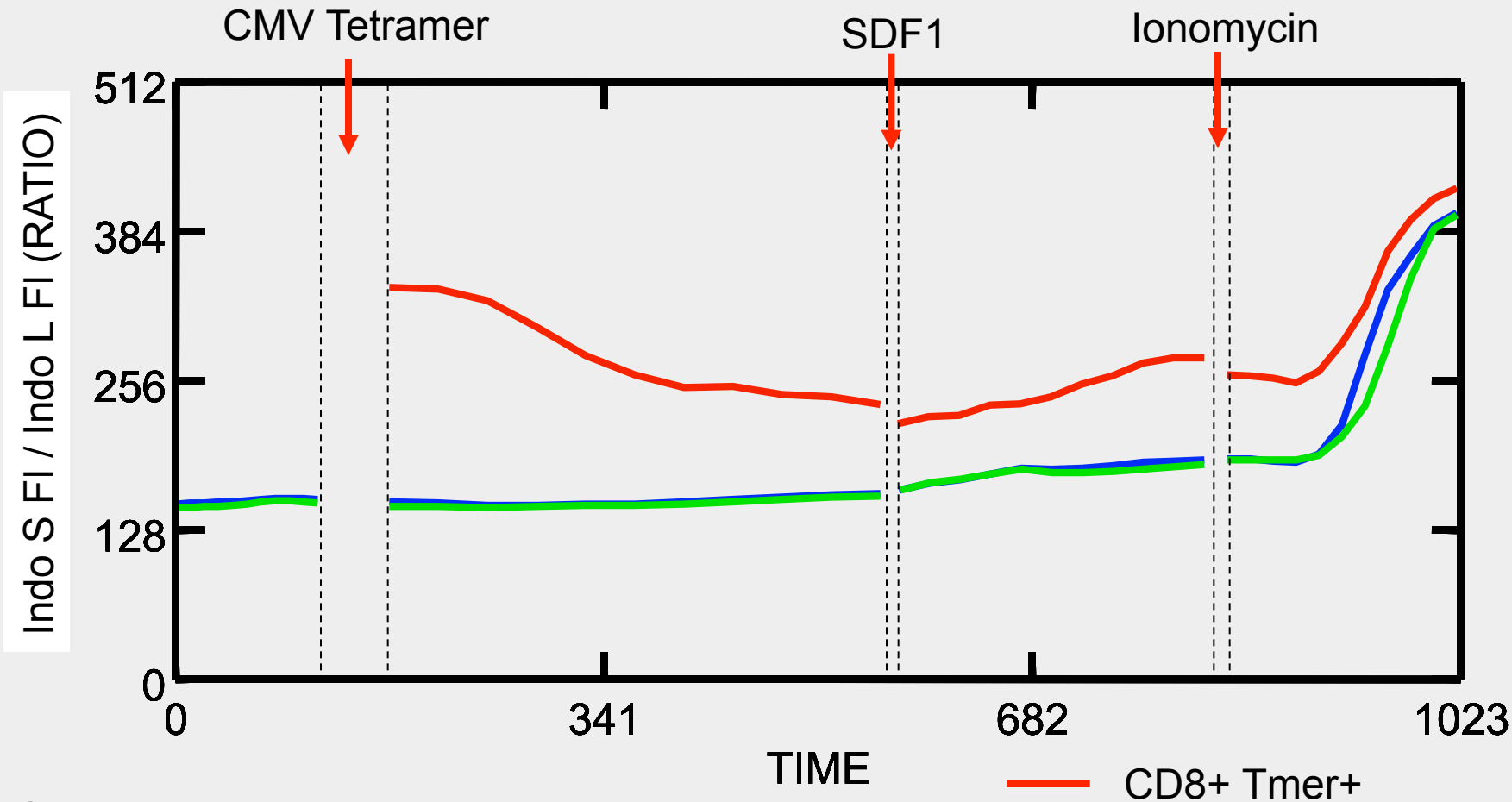
Scale: 1023 = 30 minutes
Running median smooth

473.lmd

Kinetics of Calcium Flux



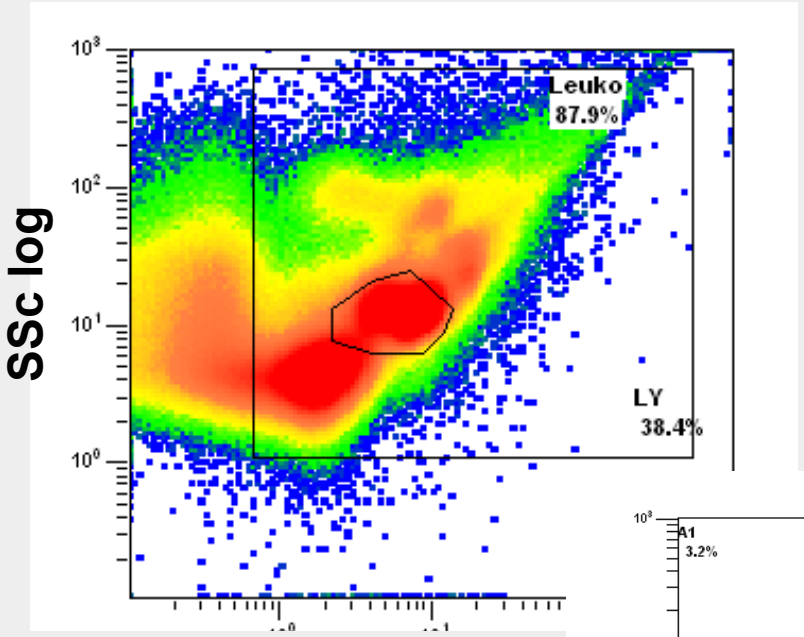
Kinetics of Calcium Flux

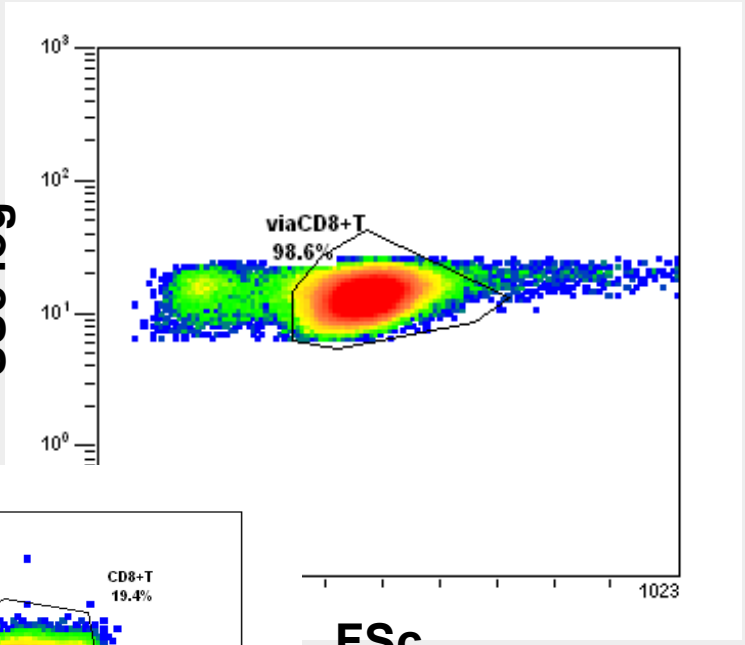


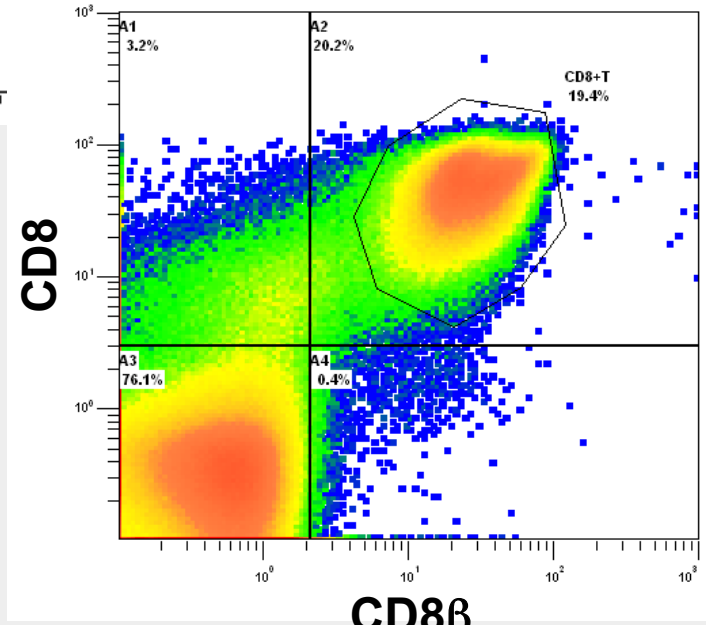
Scale: 1023 = 30 minutes

473.lmd

Primary Classifiers





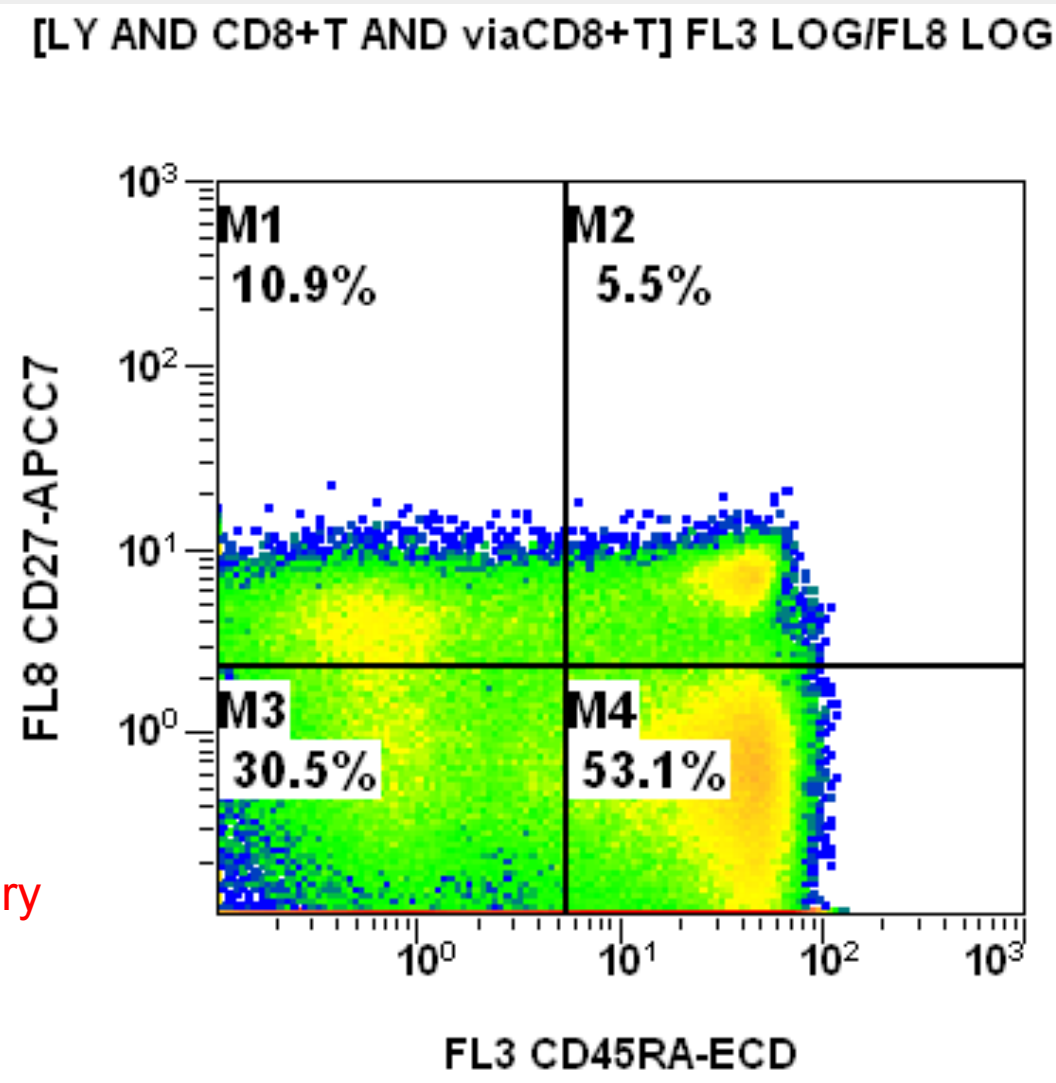


48680.lmd

Secondary Classifier

Central Memory

Naive



Effector Memory

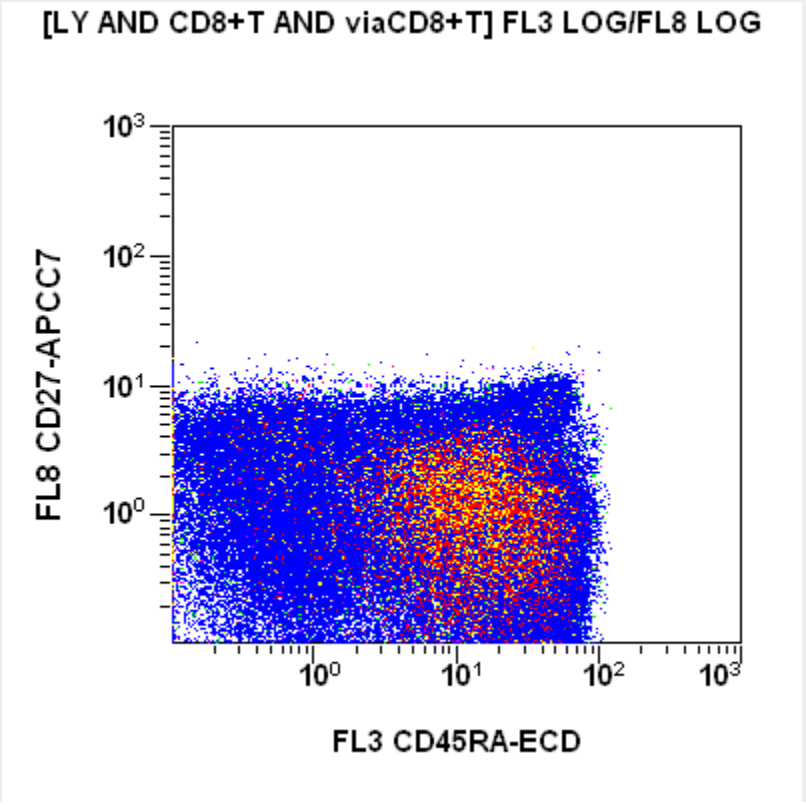
Effector

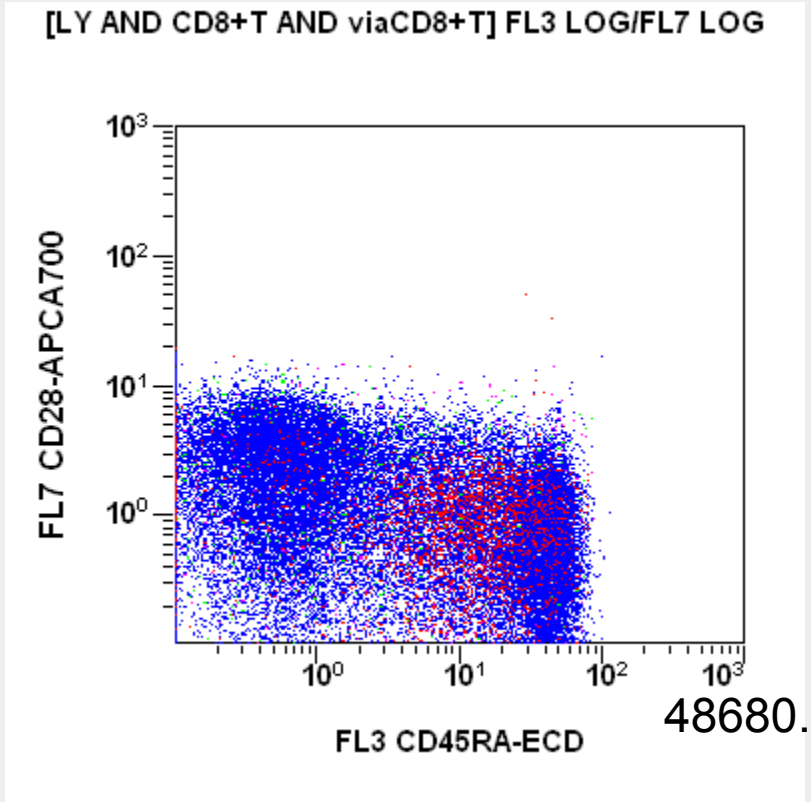
48680.lmd

Color Eventing Outcomes on Classifier and Exploratory Parameters

CMV Tetramer positive

High Calcium Flux



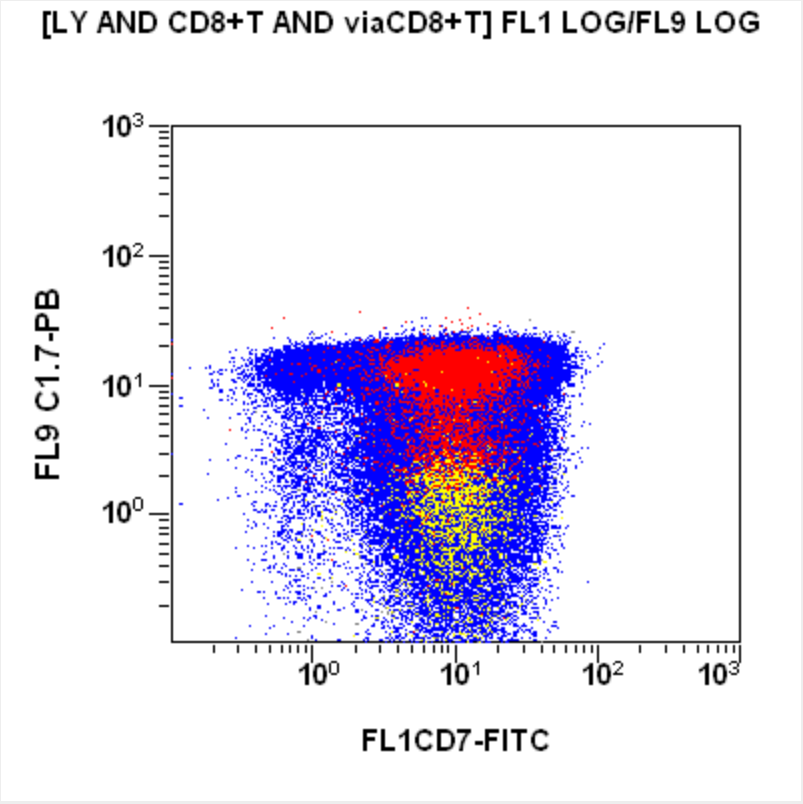


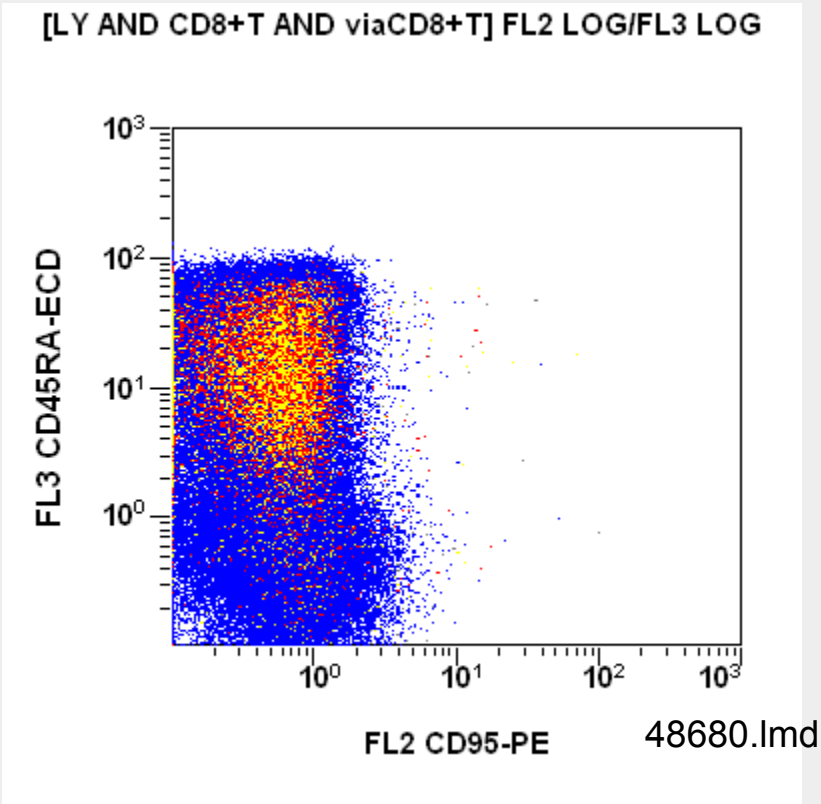
48680.lmd

Color Eventing Outcomes on Classifier and Exploratory Parameters

CMV Tetramer positive

High Calcium Flux

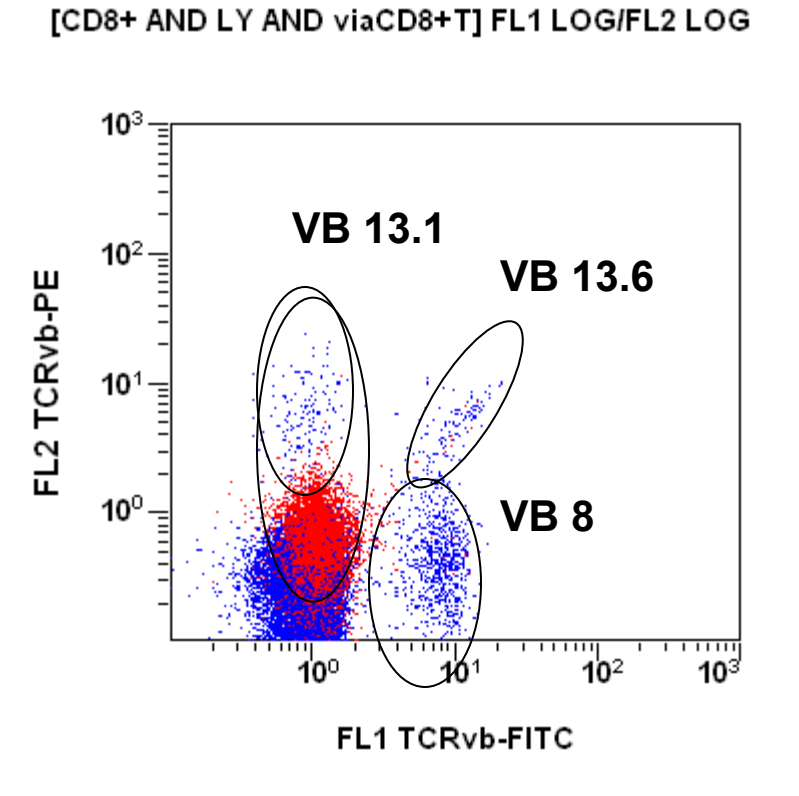


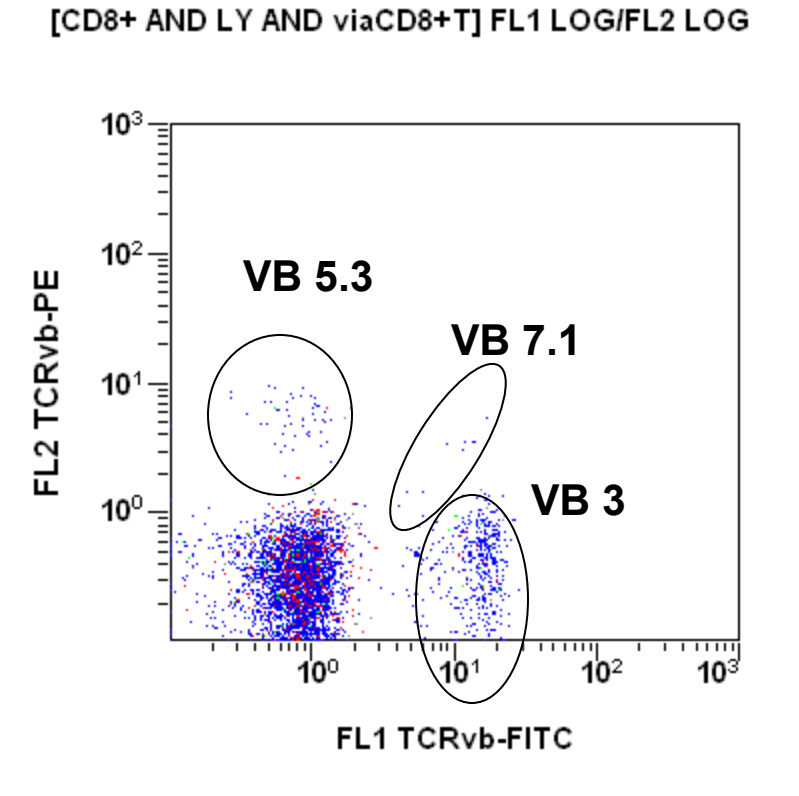


48680.lmd

Color Eventing Outcomes on Classifier and Exploratory Parameters

CMV Tetramer positive





506.lmd

Conclusions

- When dealing with multiple parameters it is useful to think of parameters as classifiers, outcomes, or exploratory
- This facilitates a hierarchical analysis that avoids the *all possible permutations* problem (*ie* 4096 distinct parameter combinations in this example)
- It is feasible and practical to perform functional, kinetic and high resolution immunophenotypic studies on populations that are considered rare events
- Hongmei Shen and Vera Donnenberg will show how these principles can be applied to detection of normal and malignant stem cells

Acknowledgements

- Vera Donnenberg
- Members of AVDLab past and present
  - E Michael Meyer
  - Debe Griffin
  - Hongmei Shen
  - Erin McClelland
  - Cassandra Singer
  - Thomas Hoffmann
  - Sumita Ganguly
  - Dawn Betters
  - Anita Popovic
- Kit Snow, Todd Lary, Meryl Forman and Tom Franks at Beckman Coulter