Program number 186

A Novel Modeling System for Analysis of High Dimensional data: Definition of T-cell Effector and Memory Subsets Using GemStone™ Software

Inokuma MS[1], Trotter J[2], Hunsberger BC[3], Munson ME[3], Herbert DJ[3], Bray CM[3], Ghanekar SA[1], Maino VC[1], and Bagwell CB[3]

[1]BD Biosciences, 2350 Qume Dr., San Jose, CA 95131; [2]BD Biosciences, 11077 N. Torrey Pines Rd. La Jolla, CA 92037; [3]Verity Software House, 45A Augusta Rd. Topsham, ME 04086

Abstract

The differentiation pathways of memory and effector T-cells are poorly defined in CD8 and CD4 T-cell subsets. These subsets, often defined by phenotypic and functional characteristics are usually determined using multiparameter flow cytometry. Analysis has been limited to arrays of 2-dimensional (2D) plots, resulting in lack of consensus in defining effector and memory subsets. These limitations in technology have hindered the development of relevant models in the field of T-cell mediated immunity. We introduce GemStone™ software, based on a Probability State Modeling system, paired with multicolor flow cytometry data, which allows for the interrogation and presentation of high dimensional flow cytometry data. The development of this system allows for a more sophisticated understanding of the phenotypic and functional correlates associated with T-cell maturation. Traditional methods of analysis of gating 2D dot plots were compared with the modeling system using GemStone. Additionally, multiple cocktails were combined to display a comprehensive model of phenotypes involved in differentiation of T-cells. From this, a simplified combination of CD45RA, CCR7, and CD28 was found to efficiently delineate the major effector and memory T cell subsets. Antigen specific cytokine responses were mapped to determine the types of CD8+ T-cell subsets. Knowledge of where a specific T-cell lies on these models can give information on its functional capacity and antigen exposure history.

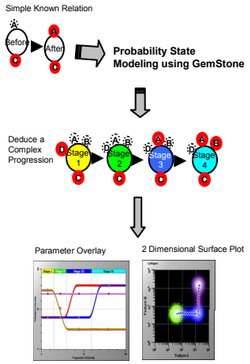
Introduction

The differentiation of human T-cells has been the focus of intense discussion in defining the lineage relationship of effector and memory cells, and more recently, in understanding the heterogeneity of these populations. The large pool of available surface and intracellular phenotyping markers, the inherent complexity of the *in vivo* systems, and the lack of comprehensive tools has resulted in much debate.

To approach this challenge from a new direction, we used GemStone software, a novel approach for the analysis of multiparameter flow data by using a Probability State Model (PSM) to identify and quantify subsets. A PSM is used to classify events into populations by probability, based on a model defined by some basic biological information. A model can be made up of one or several cell types. The phenotyping markers in the assays are used to create a set of parameter profiles for each cell type subset. A parameter profile then uses a set of control points to define how the subset transitions over the state index axis.

Using this method, subjective gating and associated errors are eliminated. Population overlaps in multidimensional data are accounted for. Data defined by numerous correlated parameters are presented in simple-to-understand graphics. Multiple samples may be combined into one cohesive analysis.

Probability State Modeling

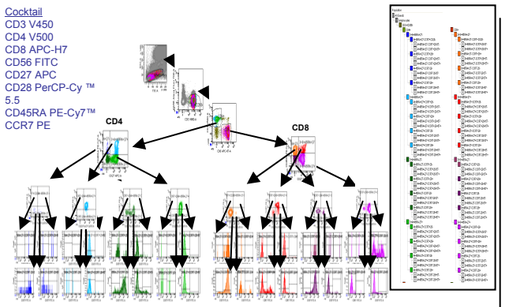


Probability State Modeling is a tool used to leverage a simple known relation (top) to deduce a much more complex progression (bottom). The simple relation in this example is the knowledge that cells selected by feature C initially are A- but later are A+. With this simple relation it is possible to unravel a very complex progression involving many features. Not only can the system determine order, but it can produce a single graph that shows the order and percentages of all the intermediate stages (Parameter Overlay) and two-dimensional contextual surface plots that appropriately blend colors from defined stages. For more details see poster P96, program number 182, "Modeling Sequential Cellular Processes Defined by Numerous Measurements" or visit www.vsh.com

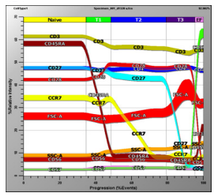
1. Simplification of multicolor data analysis: Dot plots versus GemStone models

1a Standard 2-dimensional dot plot analysis of 8-color data of both the CD4 and CD8 T-cells, resulting in 64 distinct subpopulations represented in 21 dotplots

Cocktail
CD3 V450
CD4 V500
CD8 APC-H7
CD56 FITC
CD27 APC
CD28 PerCP-Cy ™ 5.5
CD45RA PE-Cy7™
CCR7 PE



1b Progression of T-cell ontology is represented in 2 parameter overlay plots



1c Percent positive of T-cell stages are similar between GemStone model and gated dot plots

CD4

	%CD4	%CD4
CD45RA	CD27	CCR7	CD28	CD57	Gemstone	dotplot
Naive	+	+	+	+	+	55.15	48.44
T1	+	+	+	+	-	14.3	20.13
T2	+	+	+	-	-	21.15	20.17
T3	-	+	+	-	-	10.1	8.7
EF	+	-	-	-	-	0.3	4.04

CD8

	%CD8	%CD8
CD45RA	CD27	CCR7	CD28	CD57	Gemstone	dotplot
Naive	+	+	+	+	+	47.19	55.0
T1	+	+	+	+	-	24.66	26.7
T2	+	+	+	-	-	7.24	8.43
T3	-	+	+	-	-	5.64	7.26
EF	+	-	-	-	-	20.26	23.84

CD4

%CD4 CD45RA	CD27	CCR7	CD28	CD57
48.44	+	+	+	+
20.13	-	+	+	+
18.59	+	+	+	-
1.67	+	-	+	+
1.02	-	-	+	+
4.71	+	-	+	-
0.33	+	-	-	+
0.03	-	-	+	-
0.00	-	-	-	+
0.09	+	+	+	+
0.14	-	+	+	+
0.32	+	+	+	-
0.06	-	+	+	-
0.00	+	+	-	+
0.01	-	+	-	+
0.15	+	+	-	-
0.01	-	+	-	-
0.07	+	-	-	+
1.51	-	-	-	+
0.12	+	-	-	-
1.71	-	-	-	-
23.38	+	-	-	-
0.04	+	+	+	-
0.01	-	+	+	-
0.01	+	+	-	-
0.01	-	+	-	-
0.13	+	-	+	-
0.03	-	-	+	-
0.04	+	-	-	-
0.11	-	-	-	-
0.00	+	+	+	+
0.01	-	+	+	+

CD8

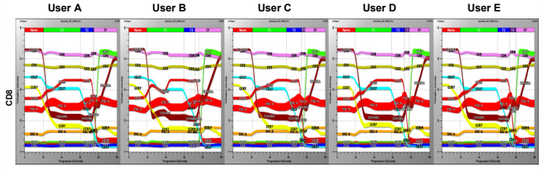
%CD8 CD45RA	CD27	CCR7	CD28	CD57
1.48	-	+	+	-
2.03	+	+	+	-
23.19	-	+	-	-
0.48	+	+	+	+
1.31	-	+	+	+
1.45	+	+	-	+
0.35	-	+	-	+
1.13	+	+	-	-
0.08	-	+	-	-
0.45	+	-	+	+
0.54	-	-	+	+
0.68	+	-	+	-
1.71	-	-	+	-
0.00	+	-	-	+
0.01	-	-	-	+
0.07	+	-	-	-
5.49	-	-	-	-
0.07	+	+	+	+
0.27	-	+	+	+
0.14	+	+	-	+
0.24	-	+	-	+
0.35	+	+	+	-
0.00	-	+	+	-
0.04	+	-	+	-
0.12	-	-	+	-
0.01	+	-	-	-
0.01	-	-	-	-

Healthy donor peripheral blood stained with a 10-color panel was acquired with a special order BD™ LSR II flow cytometer. The data was analyzed in two ways. Figure 1a shows the standard 2-dimensional dot plot breakdown of both the CD4 and CD8 T-cells, resulting in 64 subpopulations represented in 21 dot plots. The same set of data is represented in 2 GemStone parameter overlays (Figure 1b). Using GemStone, the populations of interest were selected using CD16, CD56, CD3, CD4, and CD8. There is prior knowledge that CD28 is expressed in naïve cells, and that expression of CD28 diminishes as T-cells mature[1]. Similar knowledge of expression patterns of CCR7 and CD45RA in T-cells was also used to determine progression of T-cell differentiation and is represented from left to right. T-cell differentiation stages were determined by observing number of general transition points and defined by CD45RA, CD28 and CCR7. The GemStone models were used to determine phenotype of T cell subsets: Naïve, T1, T2, T3, T effector (EF). The percent populations were compared between the GemStone model and the gated dot plots (Figure 1c).

- Positive and negative gates or analysis regions from dot plots are often subjective, resulting in differences in percent positive between dot plots and parameter profiles.
- When T-cell subsets are defined by 2D dot plots, illogical stages, (e.g. CD45RA+CD27+CCR7+CD28+CD57+) are created, making it much more difficult to map T-cell stages. See Figure 1c, orange rows.
- GemStone models present high-dimensional data in a biologically relevant, easily interpreted format.

2. GemStone models are reproducible between different users and between healthy donors

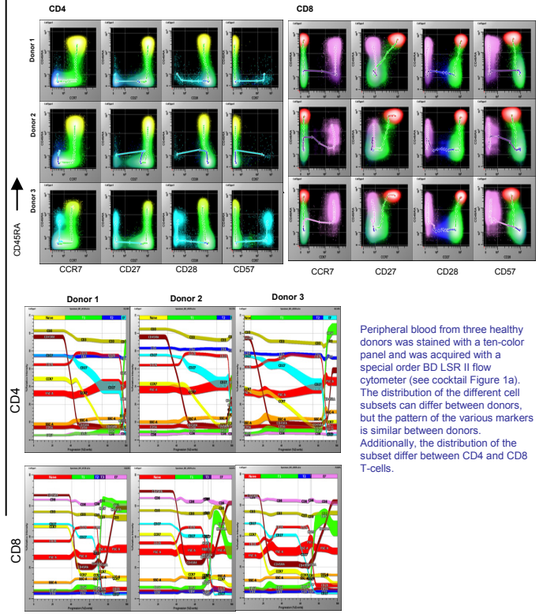
2a Different users generate similar GemStone models



Healthy donor peripheral blood stained with a ten-color panel was acquired with a special order BD LSR II flow cytometer (see cocktail Figure 1a). Each user was given minimal information about the markers in the file, then was asked to independently generate a model. Analysis of cell types is equivalent between the three different users.

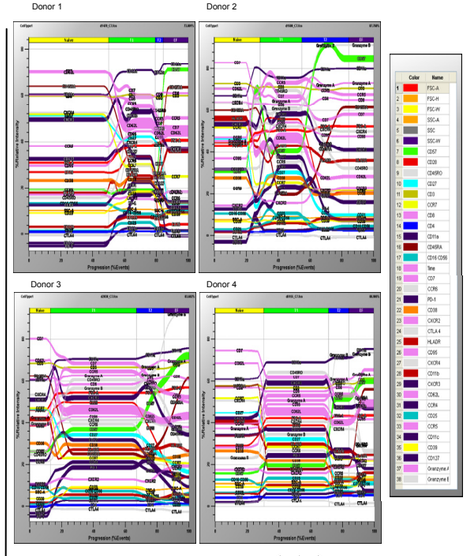
User	A	B	C	D	E
Naïve	21.44	20.37	21.64	20.4	21.04
T1	39.61	40.25	37.77	39.35	39.69
T2	15.03	15.28	16.36	14.35	13.11
T3	0.95	0	3.78	4.36	5.83
EF	22.97	27.56	20.45	21.56	20.32

2b Data from healthy donors result in similar models



Peripheral blood from three healthy donors was stained with a ten-color panel and was acquired with a special order BD LSR II flow cytometer (see cocktail Figure 1a). The distribution of the different cell subsets can differ between donors, but the pattern of the various markers is similar between donors. Additionally, the distribution of the subset differ between CD4 and CD8 T-cells.

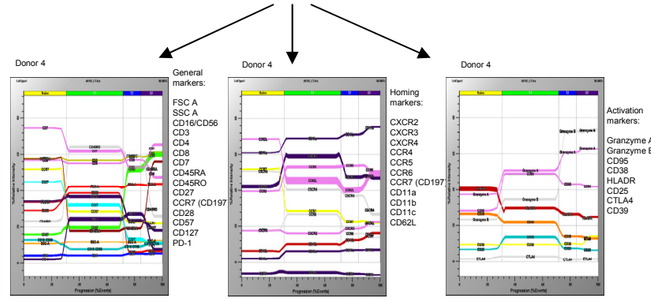
3. Analysis of 38 correlated phenotypic markers on CD8 T-cells



Peripheral blood of four healthy donors was stained with 7 different 9 to 11-color panels and was acquired with a special order BD LSR II flow cytometer.

Using GemStone, the populations of interest were selected using CD16/CD56, CD3, CD4, CD8. *A priori* knowledge of expression patterns of CD28, CCR7 (CD197) and CD45RA in T-cells was used to determine progression of T-cell differentiation in the model. Remaining markers were allowed fall in place in reference to the progression established by the three model markers. By having a scaffold of matching markers (in blue), the seven cocktails can be combined into one parameter profile so that all the markers can be correlated to one another.

The parameter profile of donor 4 was separated out into three groupings for easier analysis. The first grouping displays general markers commonly used in T-cell phenotyping, the second set shows homing markers, and the third set represents activation markers.



Antibody cocktails

	C1	C2	C3	C4	C5	C6	C7
PE-Cy5	CD16 CD56	CD16 CD56	CD16 CD56	CD16 CD56	CD16 CD56	CD16 CD56	CD16 CD56
APC-H7	CD3	CD3	CD3	CD3	CD3	CD3	CD3
Alexa 594	CD4	CD4	CD4	CD4	CD4	CD4	CD4
V500	CD8	CD8	CD8	CD8	CD8	CD8	CD8
PrCP/Cy5.5	CD28	CD28	CD28	CD28	CD28	CD28	CD28
PE-Cy7	CD45RA	CD45RA	CD45RA	CD45RA	CD45RA	CD45RA	CD45RA
V450	CCR7	CCR7	CCR7	CCR7	CCR7	CCR7	CCR7
FITC	CD57	CD7	CXCR2	CD95	CXCR3	CD25	Granzyme A
PE	CD11a	PD-1	CD127	CXCR5	CCR4	CD38	Granzyme B
APC	CD45RO	CD6	CTLA4	CXCR6	CD39	HLADR	
Alexa 700	CD27	CD38	HLADR	CD11b			

T cell subsets:

From the parameter profiles, four major CD8 T-cell subsets can be identified with the transition of the phenotypic markers: Naïve, T1, T2 and EF (terminal effectors).

Specific CD8+ T-cell phenotyping markers:

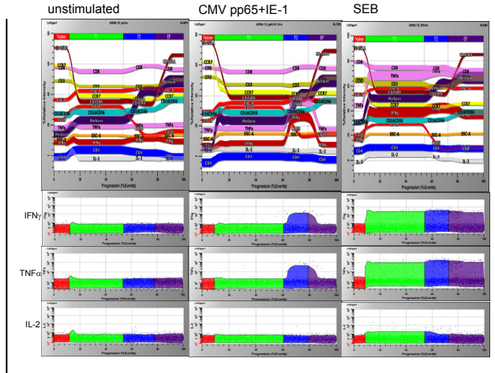
Contrary to previously published models[1,2], several differences can be seen in the data from four healthy donors:

-CCR7 and CD45RA start first transition at approximately the same stage

-CD11a expression increases with T-cell differentiation

-CD28 shows a significant (and reproducible) increase in T1 stage on T-cells

4. Cytokine expression in activated CD8+ T-cells can be associated with memory phenotypes



Peripheral blood from a known CMV responsive donor was stimulated with CMV peptide pools of pp65+IE-1 or with SEB for 6 hrs, then stained with a 10-color cocktail. Cells were selected for CD16/CD56^{neg}, CD3^{pos}, CD4^{neg}, CD8^{pos}, then *a priori* knowledge of expression patterns of CD28, CCR7 (CD197) and CD45RA in T-cells was used to determine progression of T-cell differentiation in the model. To show subtle activation patterns along the progression axis, Perforin, IFNγ, TNFα, and IL-2 were analyzed using the Hi-Lo Display and Analysis System in GemStone.

Antibody cocktail

PE-Cy5	CD16 CD56
APC-H7	CD3
Alexa 594	CD4
V500	CD8
PrCP/Cy5.5	CD28
PE-Cy7	CD45RA
V450	CCR7
FITC	TNFα
PE	perforin
APC	IL-2
Alexa 700	IFNγ

CMV pp65+IE-1 stimulation:

IL-2 expression peaks at the beginning of T1, then rapidly diminishes in expression through T2. Both TNFα and IFNγ are expressed during early terminal effector stage (EF), then demonstrate diminished expression in late EF.

SEB stimulation:

IL-2 expression peaks at beginning of T1, and plateaus through beginning of T2 stage.

TNFα and IFNγ also are expressed at T1, then slightly increase through T2 and at beginning of EF stage, then diminishes in expression.

Summary

- GemStone models present high dimensional data in a biologically relevant, easy to interpret format.
- Generation of models are consistent between multiple users and across healthy donors.
- When 38 phenotypic markers were combined into one model, four major CD8+ T-cell subsets can be seen with the transition of the markers: Naïve, T1, T2, and EF.
- These four major CD8+ T-cell subsets can be efficiently defined with three markers: CD45RA, CD28, and CCR7.
- By looking at data from healthy donors, several patterns were observed that were contrary to previously published models.
- Cytokine responses to CMV peptide pools and SEB can be mapped to specific T-cell stages.

References

1. Appay et al. Phenotype and Function of Human T Lymphocyte Subsets: Consensus and Issues. Cytometry Part A 2008; 73A:975-983.
2. Sallusto et al. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. Nature 1999; 401:708-712.

Acknowledgements

Special thanks to Holden Maecker and Jack Dunne for valuable insights, and to Larry Duckett and crew for the special order BD LSR II flow cytometers used in this study