Human B-Cell and Progenitor Stages as Determined by Probability State Modeling of Multidimensional Cytometry Data

Bagwell, CB[1], Hill, BL[1], Wood BL[2], Wallace, PK[3], Kelliher, AS[4], Preffer, FI[4]

[1] Verity Software House, [2] University of Washington, [3]Roswell Park Cancer Institute, [4] Massachusetts General Hospital

Introduction

The development of B lymphocytes in the bone marrow is well-studied due to the importance of understanding the dysregulation that occurs in leukemias and other diseases of the hematopoetic system. This development occurs in a well-ordered progression that can be defined by the sequential up-regulation and down-regulation of markers which can be subdivided into distinct stages [1-3]. These stages remain fairly consistent throughout a normal lifetime, but vary with the percentage of cells in a particular stage due to age, exposure to antigens and individual genetics[4].

Historically in the scientific literature there have been discrepancies and inconsistencies in the timing of the up-regulation or down-regulation of some markers in relationship to other markers [5-8]. Since it is important to have a clear picture of these relationships in normal bone marrow in order to diagnose and treat disrupted marrows, we undertook a statistical study of normal B-Cell progression using Probability State Modeling and data files from sixteen normal samples.

Probability State Modeling, or PSM, employs algorithms that allow one to model any number of flow cytometry data files and calculate the average expression of markers as well as the correlations between samples and directionality of the markers [9]. In addition, with PSM we can perform an unattended analysis of each data file, thus reducing the subjectivity that arises during traditional gating analysis.

Using typical B-Cell lineage markers, CD45, CD19, CD10, CD34, CD38, and CD20, we were able to clearly define four distinct stages of development of B cells and generate correlation statistics to support the transitions between these stages. We were also able to determine the expression profiles of several other markers, CD81, TdT, CD22, CD44, CD9, that contribute to the development of B-Cells in the marrow.

Furthermore, using the same markers and a subset of the data files, we were able to statistically define three very early preceding stages of lymphocyte development and the coordinated expression of markers defining these stages.

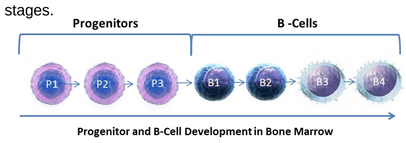

Figure 1. Poster Summary. Developmental Overview of B–Cell lineage progression.

Materials and Methods

Data analysis was performed using GemStone™ Probability State Modeling and Graphics Software program version 1.0.115.

Flow Cytometry listmode data files for the study were selected from our library of data files based on criteria summarized in Table 1 below. Briefly, files must include the markers CD19, CD45, CD10, CD34, and CD38 for staging, have an adequate number of events, and unequivocal separation between positive and negative peaks.

A normal B-cell developmental model was constructed using GemStone and Probability State Modeling by selecting for the CD19 positive to dim events and low SSC. Stratification for staging was done with CD10, CD34, CD38 and CD45. Each file was analyzed unattended with statistics collected during the analysis. Correlation coefficients were calculated to measure the modulation of each marker to determine coordinated changes in expression patterns. Based on these statistics an averaged model was created.

A normal progenitor model was constructed in a similar manner; however, progenitors were selected as CD34 positive, CD45 low to intermediate intensity and SSC low to intermediate intensity. Stratification was determined with CD38, CD10 and CD19 to identify the stages of very early B-cell development. Fewer files were used in this section of the study due to inadequate numbers of events in the omitted files.

File Name	# of Events	#CD19 dim/+	#Prog. Events	Date Collected	Instrument	Markers Present										
						CD19	CD10	CD34	CD38	CD45	CD20	TdT	CD81	CD22	CD44	CD9
B_BM1.fcs*	250000	16811	996	6-May-13	Facs Cantoll	x	x	x	x	x				x	x	
B_BM2.fcs	100000	2589		23-Apr-14	Facs Cantoll	x	x	x	x	x	x	x				
B_BM3.fcs	100000	4722		23-Apr-14	Facs Cantoll	x	x	x	x	x	x	x				
B_BM4.fcs	100000	3430		1-May-14	Facs Cantoll	x	x	x	x	x	x	x				
B_BM5.fcs	100000	8559		1-May-14	Facs Cantoll	x	x	x	x	x	x	x				
B_BM6.fcs	100000	5583		1-May-14	Facs Cantoll	x	x	x	x	x	x	x				
B_BM7.fcs*	250000	22012	15063	11-Jan-12	Facs Cantoll	x	x	x	x	x				x	x	
B_BM8.fcs*	337097	34849	5450	25-Aug-09	LSRII	x	x	x	x	x	x		x			
B_BM9.fcs*	417911	7446	7448	3-Feb-10	LSRII	x	x	x	x	x	x					
B_BM10.fcs*	422432	11898	2927	3-Feb-10	LSRII	x	x	x	x	x	x					
B_BM11.fcs*	742320	6622	7179	4-Feb-10	LSRII	x	x	x	x	x	x					
B_BM12.fcs*	522838	23228	4603	4-Feb-10	LSRII	x	x	x	x	x	x					
B_BM13.fcs	171218	32332		10-Feb-10	LSRII	x	x	x	x	x	x					x
B_BM14.fcs*	359633	30549	2476	17-Mar-10	LSRII	x	x	x	x	x	x					x
B_BM15.fcs*	321185	6343	7385	17-Mar-10	LSRII	x	x	x	x	x	x					x
B_BM16.fcs*	415403	10994	4737	20-Mar-10	LSRII	x	x	x	x	x	x					x

Table 1. Summary of all files used in study. To be included in this study files had to at least have staining for CD19, CD34, CD45, CD10 and CD38, have an adequate number of events (>50 events per state) and exhibit good separation between negative and positive staining. Files used for the progenitor study are starred.

	B1		B2		B3 and B4	
	a	b	c	d	e	f
Files	CD34	CD45	CD45	CD20	CD10	CD38
B_BM1	7.3	9.0	58.0		67.5	67.1
B_BM2	28.0	27.0	95.8	100.0	95.8	94.6
B_BM3	15.7	21.9	71.9	67.1	78.2	77.8
B_BM4	18.7	26.2	63.8	69.6	69.3	71.8
B_BM5	5.4	9.2	78.7	76.1	92.1	92.7
B_BM6	3.4	4.0	18.0	18.9	22.6	23.1
B_BM7	12.3	11.0	48.7	47.2	61.2	59.3
B_BM8	9.9	10.0	71.2	71.1	86.3	89.2
B_BM9	48.7	50.6	98.5	100.0	95.4	95.7
B_BM10	6.0	8.6	36.8	36.9	41.6	40.7
B_BM11	11.4	12.0	83.3	79.6	81.9	82.2
B_BM12	3.8	4.4	23.5	22.9	25.9	26.2
B_BM13	1.7	7.1	95.3	95.9	98.1	95.6
B_BM14	5.5	5.1	37.7	36.8	39.9	39.9
B_BM15	41.3	67.0	80.7	80.2	78.5	81.7
B_BM16	35.9	37.1	91.4	92.6	91.9	91.5
μ	15.9	19.4	65.8	66.3	70.4	70.6
σ	14.7	18.3	26.2	27.5	25.3	25.3
r	0.95		0.99		0.99	
H_0: r=0	<0.00001		<0.00001		<0.00001	
t	2.13		0.04		0.48	
H_0: t=0	N.S.		N.S.		N.S.	

Table 2. B-Cell Stage Results. All files were analyzed unattended with a B-Cell model to obtain critical stage locations for markers CD34, CD45, CD20, CD10, and CD38. The units for the locations are cumulative percentages. The recorded results for all sixteen study files are shown above. Markers were deemed to be on the same stage boundary if their correlation coefficients, r, were significant and their standard differences, t, were not (see bottom of Table 1).

The down-regulation of CD34 (see a and **Figure 2**) and the initial up-regulation of CD45 (b) were found to represent the end of the first stage boundary, B1 (red). The second up-regulation of CD45 (c) and the up-regulation of CD20 (d) defined the end of the second stage, B2 (green). The down-regulation of CD10 and CD38 (e and f) determined the end of B3 (blue) and the beginning of stage B4 (purple, not shown).

	P1	P2 and P3	
	a	b	c
Files	CD38	CD19	CD10
B_BM1.fcs	67.4	68.0	69.3
B_BM7.fcs	3.9	10.0	11.5
B_BM8.fcs	7.7	37.8	36.0
B_BM9.fcs	2.5	33.5	27.2
B_BM10.fcs	43.0	85.8	86.0
B_BM11.fcs	33.5	72.0	71.4
B_BM12.fcs	48.5	86.1	83.1
B_BM14.fcs	46.0	47.5	46.8
B_BM15.fcs	46.9	100.0	99.9
B_BM16.fcs	2.3	18.3	22.3
μ	30.2	55.9	55.4
σ	24.0	30.9	30.6
r		0.99	
H_0: r=0		<0.00001	
t		0.62	
H_0: t=0		N.S.	

Table 3. Progenitor Stage Results. Files with enough events to analyze were modeled to obtain critical stage locations for markers CD38, CD19 and CD10. The units for the locations are cumulative percent. The recorded results for ten study files are shown above. Markers were deemed to be on the same stage boundary if their correlation coefficients, r, were significant and their standardized differences, t, were not (see bottom of table).

The up-regulation of CD38 (see a and **Figure 5**) was found to represent the end of the first stage boundary, P1 (red). The up-regulation of CD19 (b) and CD10 (c) were determined to occur together forming the end of stage P2 (green) and the beginning of stage P3.

B-Cell Staging Results

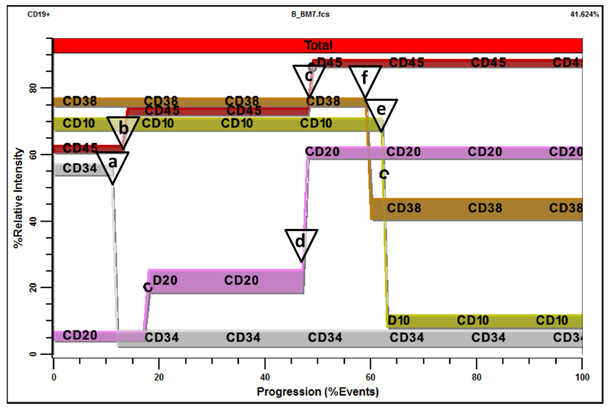

Figure 2. B-Cell Stage Analysis. Representative summary plot of an unattended analysis for one study file. This PSM overlay plot summarizes the modulation of CD34, CD45, CD20, CD10 and CD38 during B-cell ontogeny. Modeling reduces the listmode data to a set of critical control points that can be used to quantify the relative order of marker changes. The listmode events have been selected for CD19 dim to positive and SSC lymphocytes (see Materials and Methods for details). When CD34 down-regulates, usually CD45 up-regulates slightly (see open triangles for control points a and b). When CD45 up-regulates for the second time, CD20 up-regulates (control points c and d). Finally, when CD10 down-regulates, CD38 also down-regulates (control points e and f). These locations are quantified by the PSM model as cumulative percentages and summarized for all files in **Table 2**.

Normal B-Cell Model

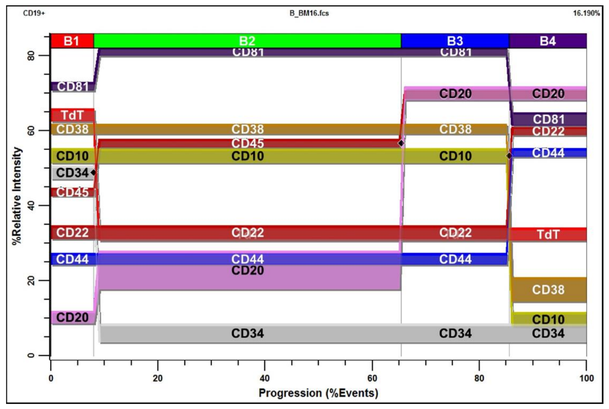

Figure 3. Normal B-Cell Model. Based on the B-Cell Staging results, a "normal" B-cell model was constructed that constrained appropriate marker modulations to the three critical control points shown in **Figure 2** and **Table 2** (a, c, and e). Markers that were not widely represented in this study (TdT, CD81, CD22 and CD44) were placed on the nearest "normal" stage boundary. The PSM overlay plot shown above is the average of all files in the unattended analysis. The end of the B1 stage is defined by the down-regulation of CD34 along with TdT, while CD45, CD81, and CD20 up-regulate slightly. Not shown here is a slight down-regulation of CD22 at this stage. At the end of B2, CD45 and CD20 up-regulate. The transition of B3 to B4 is marked by the down-regulation of CD10, CD38 and CD81, and the up-regulation of CD22 and CD44.

Normal B-Cell Expression Patterns

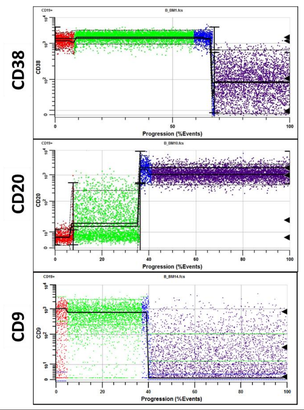

Figure 4. Panel A: A typical expression pattern for CD38. CD38 usually up-regulates slightly in B1 and is very heterogeneous upon down-regulation.

Panel B: When CD20 up-regulates at the end of B1 it becomes very heterogeneous, acting much like an activation marker. Upon entering the B4 stage, it slightly down-regulates.

Panel C: CD9 staining is very heterogeneous in stages B1 and B4.

Progenitor Staging Results

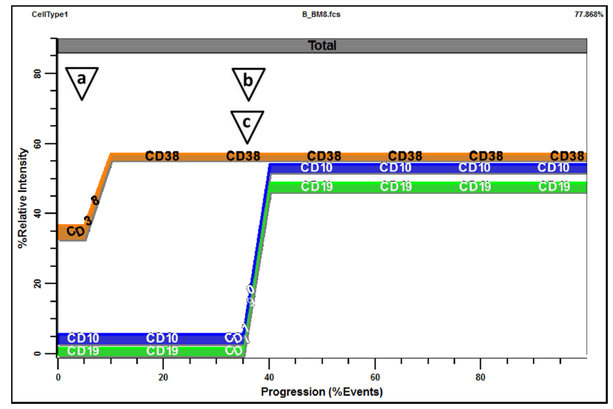

Figure 5. Progenitor Stage Analysis. Representative summary plot of one study file. The PSM overlay plot above summarizes the modulation of CD38, CD19 and CD10 during very early bone marrow ontogeny. The listmode events have been selected for CD45 low to intermediate intensity, SSC low to intermediate intensity, and CD34 positive. (see Materials and Methods for details). In our analysis CD38 up-regulates first (see a), then CD19 and CD10 up-regulate close together (b and c). These locations are quantified by the PSM model as cumulative percentages and summarized for all suitable files in **Table 3**.

Normal Progenitor Model

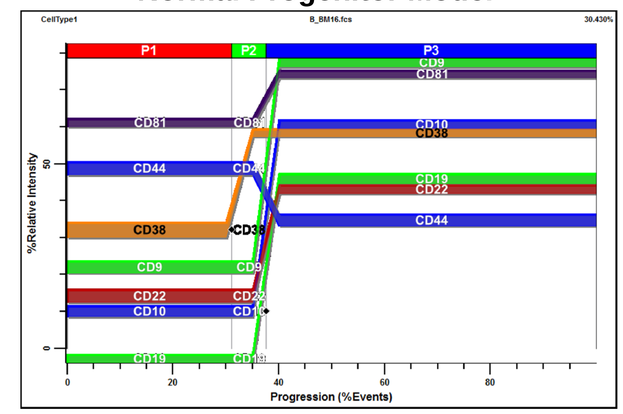

Figure 6. Normal Progenitor Model. Based on the Progenitor Staging results, a "normal" progenitor model was constructed that constrained appropriate marker modulations to the two critical control points shown in **Figure 5** and **Table 3** (see a and b). Markers that were not widely represented in this study (CD81, CD22, CD9 and CD44) were place on the nearest "normal" stage boundary. The PSM overlay plot shown above is the average model of all files analyzed for the progenitor population. The end of P1 is marked by the up-regulation of CD38. P2 ends when CD19, CD10, CD81, CD22 and CD9 up-regulate, while CD44 slightly down-regulates.

Normal Progenitor Expression Patterns

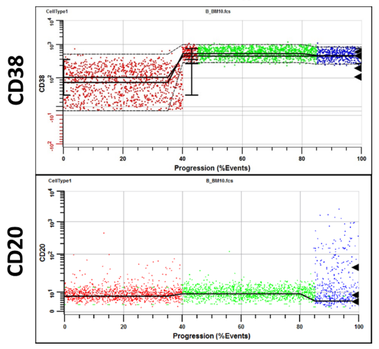

Figure 7. Panel A: A typical expression pattern for CD38 in progenitor cells. CD38 exhibits a heterogeneous staining pattern until it up-regulates.

Panel B. When CD20 starts to up-regulate at the end of P2, it becomes very heterogeneous, acting much like an activation marker.

Click for link to Website: www.vsh.com

References

1. LeBien TW, Womann B, Villablanca JG, et al. Multiparameter flow cytometric analysis of human fetal bone marrow B cells. *Leukemia*. 1990. 4: 354-358.
2. Loken MR, Shah VO, Dattilio KL, Civin CL. Flow cytometric analysis of human bone marrow II: normal B lymphocyte development. *Blood*. 1987. 70: 1316-1324.
3. Brashen CJ, Kersey JH, Bollum FJ, LeBien TW. Ontogenic studies of lymphoid progenitors in human bone marrow. *Exp. Hematol*. 1982. 10: 886-892.
4. Loken MR and Well DA. Normal antigen expression in hematopoiesis. *Immunophenotyping*. Stewart and Nicholson Editors. 2000. Wiley-Liss Inc. 133-160.
5. Beradi AC, Meffre E, Pflumio F, et al. Individual CD34+CD38lowCD19-CD10- progenitor cells from human cord blood generate B lymphocytes and granulocuytes. *Blood*. 1997. 89: 3553-3564.
6. Ghia P, tem Boekel E, Sanz E, de la Hera A, Rolink A, Melchers F. Ordering of human bone marrow B lymphocyte precursors by single-cell polymerase chain reaction analyses of the rearrangement status of the immunoglobulin H and L chain gene loci. *J Exp Med*. 1997. 194: 2217-2229.
7. Sanz E, Munoz AN, Monserrat J, Van-Den-Rym A, Escoll P, Ranz I, Alvarez-Mon M, de la Hera A. Ordering human CD34+CD10-CD19+ pre/pro B-cell and CD19- common lymphoid progenitor stages in two pro-B-cell development pathways. *Proc Natl Acad Sci USA*. 2010, Mar 30. 107; 13: 5925-5930.
8. LeBien TW and Tedder TF. B lymphocytes: how they develop and function. *Blood*. 2008. 112; 5: 1570-1580.
9. Inokuma MS, Maino VC, Bagwell CB. Probability state modeling of memory CD8+ T-cell differentiation. *JIM*. 2013. 397; 1-2: 8-17.

Conclusion

Using Probability State Modeling we have mathematically determined the normal developmental stages of B-Cells in human bone marrow as early progenitor cells as well as when they are committed to the B-Cell lineage. This study demonstrates the utility of PSM to statistically analyze any number of high-dimensional cytometry data files and present the results in a clear and concise manner. With an averaged model we are able to generate correlation coefficients that provide confirmation to the coordinated expression of markers.

This approach is a valuable step in evaluating the dysregulation of maturation patterns that occur in disease states, marrow regeneration, and in detection of minimal residual disease. With this information in hand we will be better able to model and enumerate disrupted changes in antigen expression in hopes to define more clearly diagnostic and prognostic outcomes in the near future.

This work has been submitted to *Cytometry A* for publication in an upcoming special issue on computational data analysis.