# Demo of EPP on OMIP-047 (HiD Phenotypic Characterization of B Cells) samples

Version 1, November 2021

This demo teaches how to compare AutoGate's fully automatic gating to the cell types defined in OMIP's published gating strategy - <u>http://cgworkspace.cytogenie.org/Tutorials/omipB.png</u> For reference, full publication is available here - <u>https://onlinelibrary.wiley.com/doi/full/10.1002/cyto.a.23488</u>

This document describes how to start in any position of a conventional gating hierarchy and run unsupervised automatic gating (EPP) and then compare the results between it and any other gating method AutoGate's HiD matching tools and highlighting tools.

# 1 Set up the experiment







9 available demos Select a demo and click the <b>Ok</b> bu	utton			
(To see supporting documentation click buttons Publicat	ion, Gating or Tutorial)	_		
				8
Basophil Activation Test 6 stains by Yael Gamez, Richard B. Moss et al	Publication		Tutorial	
Macrophages and B cells 10 stains by Eliver Ghosn, Leonore Herzenberg et al	Publication	Gating	Tutorial	
CyTOF Panorama data set 49 stains by Nikolay Samusik, Garry P Nolan et al	Publication	Gating	Tutorial	
OMIP-069: Forty-Color Full Spectrum Flow 40 stains by Lily M. Park et al@Cytek Biosciences Inc.	Publication	Gating	Tutorial	
CyTOF Quantitative Comparison 50 stains by William E. O'Gorman et al@Genentech Inc.	Publication	Gating		
<ul> <li>OMIP-047: Hi-D Phenotypic Characterization of B Cells</li> <li>16 stains by Thomas Liechti, et al.</li> </ul>	Publication	Gating		
OMIP-044: 28-Color Immunophenotyping Human Dendritic 28 stains by Florian Mair & Martin Prlic	Publication	Gating		
Comparison of CyTOF assays across sites 22 stains by Michael D. Leipold, Holden T. Maecker et al	Publication	Gating		
OMIP-044 with EPP: 28-Color Immunophenotyping Human Dendritic 28 stains by Florian Mair & Martin Prlic	Publication	Gating		

## 1.2 Pick OMIP-44: 28 Color Immunophenotyping Human Dendritic

AutoGate responds by opening up the GatingTree window on the publication's 4 full stained samples. The bundled demo has already completed the setup plus the replication of the published manual gating hierarchy for one sample. This starting point allows you to

- Begin a fully automated analysis playing with AutoGate's fully automatic tools or
- Continue with a conventional analysis by dragging and dropping the provided gating **hierarchy** to other samples

- 2 Run a fully automated unsupervised analysis with exhaustive projection pursuit (EPP)
- 2.1 Select **B cells** in the gating tree and then click the toolbar's Epp button.



2.2 AutoGate responds by bringing up the following EPP configuration screen.



## 2.3 Specify the parameter for EPP to gate on

To effectively compare EPP results to the published manual results for a specific sub hierarchy we need to identify both the parameters that

- 1. Were used for the manual analysis
- 2. Those that look informative for the same cell types
- 2.4 Point EPP at parameters used manually for B cells
- 2.4.1 Click on the tree icon under the 2 parameter lists



## 2.4.2 Pick **B cells** in the gating tree



AutoGate populates the left list with the parameters that the OMIP investigators used in this sub hierarchy of their gating tree.

$\circ$	Conf	igure EPP					
2	Fully HC13, , 150416_	/e projec	<b>tion pursuit (EPP)</b> starting at �Manual Singlets ◆B cells				
	Choose PHENOTYPIN	G reagents (	(stains/scatter)				
	Use these		Do not use these				
11/25	CD10:Brilliant Violet 650 (AKA: CALLA, NEP, gp100) > CD19:Brilliant Violet 785 (AKA: B4, Leu12, MGC12802) CD21:Brilliant Violet 705 (AKA: B2, C3DR, CR2, EBV-R) CD27:PE-Texas Red (AKA: S152, T14, TMFRSF7) CD38:Alexa Fluor 700 (AKA: Lau17, T10) > Dump:APC-Cy7 IgA:APC (AKA: A, Immunoglobulin) IgD:PE-Cy5 (AKA: D, Immunoglobulin) IgG3:FIL (AKA: G3, Immunoglobulin) IgG3:Fluorescein (FITC) (AKA: G3, Immunoglobulin)		CCR7:Brilliant Violet 605 (AKA: CD197, CMKBR7) CXCR3:PE-Cy7 (AKA: CD183, CKR-L2, CMKAR3, GPR CXCR4:PE-Cy5.5 (AKA: CD184, LAP3, LESTR, NPYR, W CXCR5:AmCyan (AKA: BLR1, CD185, MDR15) IL-21R:Cascade Blue Ki67:PerCP-Cy5.5 FSC-A (logicle) FSC-H (linear) FSC-H (linear) FSC-H (logicle) SSC-A (logicle) SSC-A (logicle) SSC-A (logicle) SSC-A (logicle) SSC-4 (logicle)				
Pick reas	gents using 👫 or 🧮 or 🔍		Hide advanced options				
Poly	gons via modal clustering 📀 W:	0.01 Sigma	: 3 Best balance				
r#ofcel	s um:75 □ Apply to Trim branches? gate	leaf s?					
KLD lev 1D:	els (0=normal distributi 0.16 2D: 0.16						
	See uncompensated? Export gates?	📌 <u>Highlight (</u>	back gate) Repeat on other samples?				
			Cancel				

- 3 Point EPP at other parameters that appear informative for B cells
- 4 Click on the table icon under the 2 parameter lists



AutoGate shows a table of reagents with those currently used by EPP selected in a gold color.

## 5 Select parameters that have high KLD but are not already selected

Ensure that the reagents are sorted by highest to lowest KLD (Kullback Leibler divergence). KLD scores from 0 to 1 measure who closely a data distribution resembles the normal distribution. 0 is most resemblance and thus least interesting/informative structure.

Marker	KLD (< 1 is <info)< th=""><th>Measurement distribution</th><th>Median</th><th>Stain/fluorophor</th><th>Mean</th><th>Standard deviation</th><th>Median abs. dev.</th><th>Bright- ness</th><th>Logicle W</th><th>Max scale</th><th>Chan #</th></info)<>	Measurement distribution	Median	Stain/fluorophor	Mean	Standard deviation	Median abs. dev.	Bright- ness	Logicle W	Max scale	Chan #
lgG1	0.63		51	PE	751	5k	58	0.16	0.705	262,144	18
lgG3	0.23		106	Fluorescein (FITC)	146	440	33	0.33	0.557	262,144	7
CCR7	0.19		3k	Brilliant Violet 605	3k	4k	2k	0.78	0.726	262,144	11
CXCR4	0.17		4k	PE-Cy5.5	5k	5k	2k	0.91	0.834	262,144	21
lgD	0.15		4k	PE-Cy5	7k	10k	3k	0.74	0.977	262,144	20
CD21	0.14		1k	Brilliant Violet 705	2k	1k	660	1.46	0.684	262,144	13
CD27	0.11		868	PE-Texas Red	2k	4k	960	0.61	0.727	262,144	19
CXCR5	0.11		2k	AmCyan	2k	834	537	2.18	0.557	262,144	10
CXCR3	0.09		193	PE-Cy7	369	819	174	0.45	0.883	262,144	22
IL-21R	0.09		183	Cascade Blue	376	490	205	0.77	0.635	262,144	9
Ki67	0.09		-479	PerCP-Cy5.5	-253	5k	620	-0.05	1.159	262,144	8
CD19	0.06		49k	Brilliant Violet 785	51k	21k	13k	2.41	1.206	262,144	14
SSC-W	0.03		69k	N/A	70k	2k	1k	30.71	0.739	262,144	6
CD10	0.03		74	Brilliant Violet 650	192	1,005	305	0.19	0.988	262,144	12
SSC-A	0.02		38k	N/A	41k	12k	7k	3.45	0.627	262,144	4
FSC-H	0.02		58k	N/A	57k	11k	7k	5.16	0.739	262,144	2
SSC-H	0.02		35k	N/A	37k	10k	6k	3.66	0.739	262,144	5
CD38	0.02		107	Alexa Fluor 700	198	650	118	0.3	0.625	262,144	16
lgA	0.01		143	APC	367	2k	282	0.19	0.855	262,144	15
FSC-A	0.01		71k	N/A	71k	12k	8k	5.97	0.739	262,144	1
FSC-W	0.01		79k	N/A	79k	4k	2k	20.27	0.739	262,144	3
Dump	0.01		127	APC-Cy7	141	124	71	1.14	0.72	262,144	17

All (11/22)

X Cancel V Pick (11)

Add (mac key + select / Ctrl key + select) the 6 reagents with the highest KLD that are not already selected: BV605, PE-Cy5.5, AmCyan, PE=Cy7, Cascade Blue and PerCP+Cy5.5

Marker	KLD (< 1 is <info)< th=""><th>Measurement distribution</th><th>Median</th><th>Stain/fluorophor</th><th>Mean</th><th>Standard deviation</th><th>Median abs. dev.</th><th>Bright- ness</th><th>Logicle W</th><th>Max scale</th><th>Chan #</th></info)<>	Measurement distribution	Median	Stain/fluorophor	Mean	Standard deviation	Median abs. dev.	Bright- ness	Logicle W	Max scale	Chan #
lgG1	0.63		51	PE	751	5k	58	0.16	0.705	262,144	18
lgG3	0.23		106	Fluorescein (FITC)	146	440	33	0.33	0.557	262,144	7
CCR7	0.19		3k	Brilliant Violet 605	3k	4k	2k	0.78	0.726	262,144	11
CXCR4	0.17		4k	PE-Cy5.5	5k	5k	2k	0.91	0.834	262,144	21
gD	0.15		4k	PE-Cy5	7k	10k	3k	0.74	0.977	262,144	20
CD21	0.14		1k	Brilliant Violet 705	2k	1k	660	1.46	0.684	262,144	13
CD27	0.11		868	PE-Texas Red	2k	4k	960	0.61	0.727	262,144	19
CXCR5	0.11		2k	AmCyan	2k	834	537	2.18	0.557	262,144	10
CXCR3	0.09		193	PE-Cy7	369	819	174	0.45	0.883	262,144	22
L-21R	0.09		183	Cascade Blue	376	490	205	0.77	0.635	262,144	9
<b>Ki67</b>	0.09		-479	PerCP-Cy5.5	-253	5k	620	-0.05	1.159	262,144	8
CD19	0.06		49k	Brilliant Violet 785	51k	21k	13k	2.41	1.206	262,144	14
SSC-W	0.03		69k	N/A	70k	2k	1k	30.71	0.739	262,144	6
CD10	0.03		74	Brilliant Violet 650	192	1,005	305	0.19	0.988	262,144	12
SSC-A	0.02		38k	N/A	41k	12k	7k	3.45	0.627	262,144	4
SC-H	0.02		58k	N/A	57k	11k	7k	5.16	0.739	262,144	2
SSC-H	0.02		35k	N/A	37k	10k	6k	3.66	0.739	262,144	5
CD38	0.02		107	Alexa Fluor 700	198	650	118	0.3	0.625	262,144	16
gA	0.01		143	APC	367	2k	282	0.19	0.855	262,144	15
SC-A	0.01		71k	N/A	71k	12k	8k	5.97	0.739	262,144	1
SC-W	0.01		79k	N/A	79k	4k	2k	20.27	0.739	262,144	3
Dump	0.01		127	APC-Cy7	141	124	71	1.14	0.72	262,144	17

All (17/22)

X Cancel V Pick (17)

## 6 Run EPP

If your EPP configuration settings match the screenshot below then click the Ok button to run EPP!!

	Configure EPP
	Fully gate         HC13, 150416_HC13_037.fcs* starting at          Manual Singlets         Image: Control of the control
	Choose PHENOTYPING reagents (stains/scatter)
	Use these Do not use these
17/36	CCR7:Brilliant Violet 605 (AKA: CD197, CMKBR7)         CD103:Brilliant Violet 650 (AKA: CALLA, NEP, gp100)         > CD193:Brilliant Violet 650 (AKA: 64, Lew12, MGC12802)         CD13:Brilliant Violet 765 (AKA: 82, C3DR, CR2, E8V-R)         CD21:Brilliant Violet 705 (AKA: 82, C3DR, CR2, E8V-R)         CD3:Brilliant Violet 705 (AKA: 82, C14, TMFSF7)         CD38:Alexa Fluor 700 (AKA: 1053, CKR4.2, CMKAR3, GP CXCR3:PE-Cy7 (AKA: CD164, LAP3, LESTR, NPYR,         VAX
Pick re Separ	aragents using 🚠 or 🔛 or 🔍 🖤 < Hide advanced options. atrix options ng method Best balance 💿
DE	BM cluster selections 📀 W: 0.01 Sigma: 3
f # of c Mini	rels imum: 50 Apply to 7 Trim leaf branches 7 2 Trim leaf gates ? evels (0=normal distribution) Cluster brightness levels
Test	0.25 STOP: 0 -10 Never (c) Line color (c)
Ve	s selection, avoid IF Cluster detail (low recommended) Cluster detail (low recommended) Clusters clusters clusters clusters clusters cluster = cl
C	See uncompensated? Export gates? 🎤 <u>Highlight (back gate)</u> Repeat on other samples?
	Cancel Ok

On a 6 core MacBook it takes EPP around 10 minutes to

- 1. Go through all X/Y pairings and find the top pairing where the data splits into two parts that have the best separation and are closest to each other's size
- 2. Narrow each of the two parts and repeat step 1 finding the next best data split to narrow into
- 3. Stop repeating steps 1 and 2 when good 2 way splits are no longer available

#### EPP in progress



#### **EPP** completed



## 7 Compare EPP to OMIP investigator's manual gates

When EPP completes its run it then shows the menu for running AutoGate's matching between your fresh new EP gates and the OMIP investigator's manual gates for the same data

### 7.1 Select Configure MDS & HiD match from H->2 toolbar



#### 7.2 Respond to overlap prompt

![](_page_13_Picture_5.jpeg)

# 7.3 Ensure your settings match as below

	Configure MI	DS & Hi-D I	match
e	Multidimensio	nal scal s C <sub>Epp balanc</sub>	ing & Hi-D matching
	Select which stair	ns/scatter to	o match on
	Use these		Do NOT use these
17/22	CCR7 CD10 CD19 CD21 CD27 CD38 CXCR3 CXCR4 CXCR5 Dump IgA IgG1 IgG1 IgG3 IL-21R Ki67 FSC-A		FSC-H FSC-W SSC-A SSC-H SSC-W
Pick reag	gents using 👬 or 🗮 or 🔍		C <<< <u>Hide advanced options</u>
Matching Mass	) method +distance similarity (QFMatch)	٢	Separate plots for 2+ groups? (e.g. EPP and non EPP)?
Merge ca Unlim	andidates ited matches per subset ise if the above limit is exceeded?	0	Best + top 1.5 * N matches 📀
- Match re Disallow that are	strictions r match IF the number of parameters too far apart is more than	0	Deviation unit thresholds Scatter 2 Stain 4
Logic	le data	0	standard deviation
			Cancel Ok

#### 7.4 Evaluate matching

AutoGate shows you a resulting MDS (multi-dimensional scaling) display. This offers many features for comparing the EPP gates to the OMIP investigator's gates.

![](_page_15_Picture_2.jpeg)

The quickest quantitative summary can be found by clicking on the table button of the toolbar which then shows best to least match with supporting statistics.

9	💉 💷 🐂 🏭 🐂	Reading the second s	ngs:  🚺 QF	Match (Hi-D ma	atch)		$\odot$	MATCH sets: training=9/10 (************************************					
#	Name	Mat- ches	Similarity (QFMatch)	Overlap (F-measure)	Plot	2 3 Freq.	# of events	Orphans	Gate ID	Parent orphans	Non Epp orphans	Rank <sup>1</sup>	
1	IgA- CXCR4+ CCR7+ Ep	5	98.1%	90.8%	1	81		N/A	185	N/A	N/A	1	
18	Naive B cells	5	98.1%	90.8%	2	37	:	N/A	80	N/A	N/A	1	
17	MZ B cells	5	98.1%	90.8%	2	31		N/A	78	N/A	N/A	1	
16	IgA-	5	98.1%	90.8%	2	7.1		N/A	75	N/A	N/A	1	
19	Transitional	5	98.1%	90.8%	2	3.:		N/A	83	N/A	N/A	1	
10	IgG1+ <sup>Ep</sup> P	2	96.5%	84.4%	1	3.1		N/A	905	N/A	N/A	2	
13	lgG1+	2	96.5%	84.4%	2	3.		N/A	68	N/A	N/A	2	
8	IgG3+ CXCR5 Ki67 Ep	3	96.1%	84.7%	1	0.;		N/A	808	N/A	N/A	3	
9	IgG3+ CD38- Ki67 Ep	3	96.1%	84.7%	1	0.;		N/A	893	N/A	N/A	3	
14	lgG3+	3	96.1%	84.7%	2	0.0	;	N/A	70	N/A	N/A	3	
2	RM <sup>E</sup> P	2	90.3%	67.7%	1	10		N/A	242	N/A	N/A	4	
12	RM	2	90.3%	67.7%	2	11		N/A	66	N/A	N/A	4	
5	PB CD27+ CD10- Ep	4	88.6%	45.3%	1	0.:		N/A	694	N/A	N/A	5	
7	PB CD27+ CD21- Ep	4	88.6%	45.3%	1	0.		N/A	709	N/A	N/A	5	
6	PB CXCR4- IgA Er	4	88.6%	45.3%	1	0.		N/A	701	N/A	N/A	5	
11	РВ	4	88.6%	45.3%	2	0.;		N/A	62	N/A	N/A	5	
3	IgA+ <sup>E</sup> P	2	44.7%	3%	1	0.:		N/A	423	N/A	N/A	6	
15	lgA+	2	44.7%	3%	2	4.		N/A	73	N/A	N/A	6	
4	CXCR4+ CCR7+	0	N/A	N/A	1	0.:		N/A	535	N/A	N/A	7	
20	RM	0	N/A	N/A	2	0.1	;	N/A	87	N/A	N/A	7	
21	lgG1+	0	N/A	N/A	2	0.:		N/A	89	N/A	N/A	7	
22	lgG3+	0	N/A	N/A	2	(		N/A	91	N/A	N/A	7	
23	la A	0	NI/A	NI/A	2			NI/A	04	NI/A	NI/A	7	

#### 7.5 Mass + Distance similarity and overlap (F- measure)

Click the highlighted buttons from the toolbar to view the graphs

![](_page_17_Figure_2.jpeg)

#### 7.6 Prediction Adjudicator

This helps determine how well one classification's subsets predict another's. PA reorganizes the predicting classifier's subsets into predicting subsets: true positive, false positive and false negative subsets. PA determines whether the false positives or false negatives have more QFMatch based similarity to the predicted subset.

# Click the green +- button available in the top right corner of the above table

•		Hi-D match (AG v4.670.20	a): 17D, bes	t + top 1.5 *	N "QFxF n	natches"	, 27 bins): 1	150416_B ce	l immuno	phenotypir	ng_Healthy	Donor		
۵.	8 🕥	🗲 💷 📼 🕂 📅 🧱 🔛 🌨 🚠 💿 👫	F-m Reading	: 🧕 QFN	latch (Hi-D mai	tch)	0	MATCH sets: tra	ining=9/13 <sup>(</sup>	UNMATCHED!),	est=9/10 <sup>(1 NEV</sup>	N1)	🛬 🔛 🚠	
	#	Name	Mat- ches	Similarity (QFMatch)	Overlap (F-measure)	Plot 2	2 3 # of -req. events	Orphans	Gate ID	Parent orphans	Non Epp orphans	Rank 1	See predic	tion accuracy in PredictionAdjudicator
	• 8	Naive B cells	5	98.1%	90.8%	1	37	2 N/A	80	N/A	N/A	1		
	• 7	MZ B cells	5	98.1%	90.8%	1	31	· N/A	78	N/A	N/A	1		
	6													

#### PA results below

Subset (class) name	1 Similarity (QFMatch)	# of events	Freq.	3 Overlap (F-measure)	Subset type	2 Subset ID	#	Mat- ches		Rai
gA+ Predicted	N/A	2k	4%	N/A	predicted	73	5	N/A	•	N/A
IgA+ false +	43.6%	101	0.2%	0%	false +	73.2	25	N/A		N/A
IgA+ false -	99.1%	2k	3.5%	99.2%	false -	73.3	26	N/A	•	N/A
IgA+ true +	45.9%	37	0.1%	3.2%	true +	73.1	24	N/A		N/A
gA- Predicted	N/A	4k	7.8%	N/A	predicted	75	6	N/A	•	N/A
IgA- false +	76.1%	453	0.7%	0%	false +	75.2	28	N/A		N/A
IgA- false -	93.6%	3k	5%	83.1%	false -	75.3	29	N/A	•	N/A
IgA- true +	84.2%	1k	2%	44.8%	true +	75.1	27	N/A		N/A
gG1+ Predicted	N/A	2k	3.1%	N/A	predicted	68	3	N/A	•	N/A
gG1+ Predicted	N/A	114	0.2%	N/A	predicted	89	11	N/A	•	N/A
lgG1+ <sup>false +</sup>	89%	390	0.6%	0%	false +	68.2	19	N/A		N/A
lgG1+ <sup>false -</sup>	70%	191	0.3%	19.5%	false -	68.3	20	N/A		N/A
lgG1+ false -	100%	114	0.2%	100%	false -	89.3	40	N/A	•	N/A
lgG1+ true +	96.4%	2k	2.5%	94.3%	true +	68.1	18	N/A	•	N/A
gG3+ Predicted	N/A	328	0.6%	N/A	predicted	70	4	N/A		N/A
lgG3+ false +	83.2%	68	0.1%	0%	false +	70.2	22	N/A		N/A
lgG3+ false -	79.9%	37	0.1%	20.3%	false -	70.3	23	N/A		N/A
lgG3+ true +	97.4%	291	0.5%	94%	true +	70.1	21	N/A	•	N/A
MZ B cells Predicted	N/A	18k	31.6%	N/A	predicted	78	7	N/A	•	N/A
MZ B cells false +	79.7%	2k	2.9%	0%	false +	78.2	31	N/A		N/A
MZ B cells false -	65.5%	545	0.9%	5.9%	false -	78.3	32	N/A		N/A
MZ B cells true +	98.9%	17k	27.3%	98.5%	true +	78.1	30	N/A	•	N/A
Naive B cells Predicted	N/A	21k	37.3%	N/A	predicted	80	8	N/A	•	N/A
Naive B cells false +	66.4%	2k	3.4%	0%	false +	80.2	34	N/A		N/A

### 7.7 Get False positive/negative rate

Click 'See false positive/negative results' blue button available on the top right corner of the toolbar

![](_page_19_Figure_2.jpeg)