Demo of EPP on OMIP-044 (28-color Immunophenotyping Human Dendritic) samples

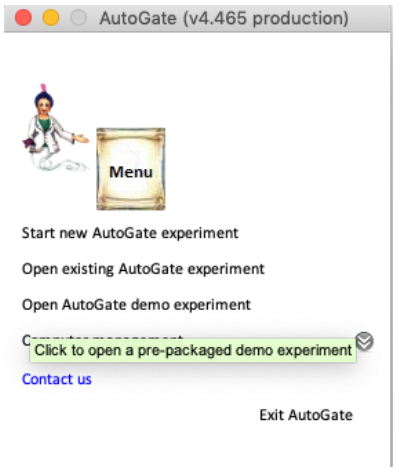
Version 1, November 2021

This demo teaches how to compare AutoGate's fully automatic gating to the cell types defined in OMIP's published gating strategy - http://cgworkspace.cytogenie.org/Tutorials/omip44.png .
For reference, full publication is available here - https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.23331

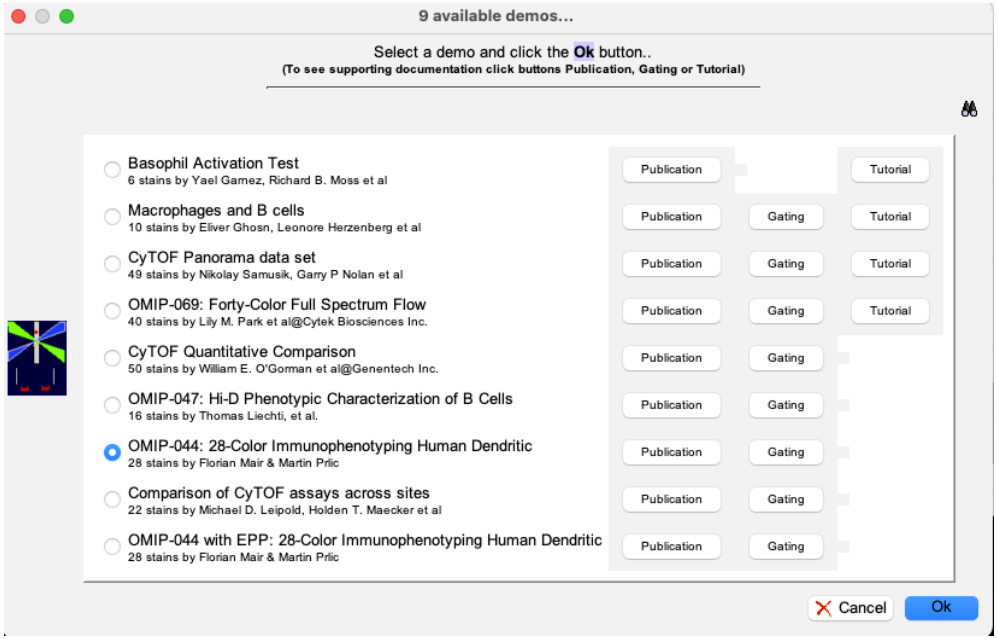
This document describes how to start in any position of a conventional gating hierarchy and run unsupervised automatic gating (EPP) and then compare the results between it and any other gating method AutoGate's HiD matching tools and highlighting tools.

1 Set up the experiment

1.1 Click "Open AutoGate demo experiment" on the main window



AutoGate responds by showing you the below list of demo experiments available as of November 2021



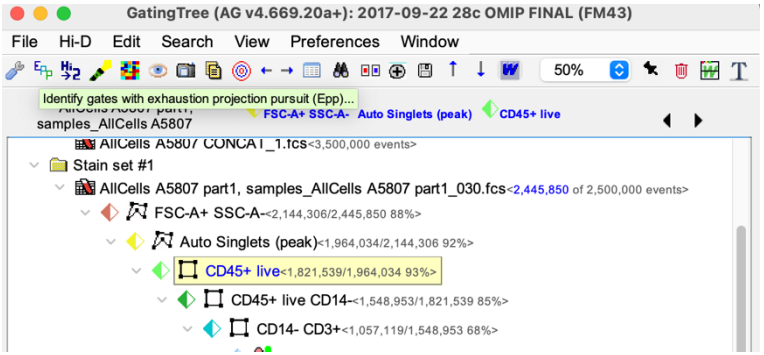
1.2 Pick OMIP-44: 28 Color Immunophenotyping Human Dendritic

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

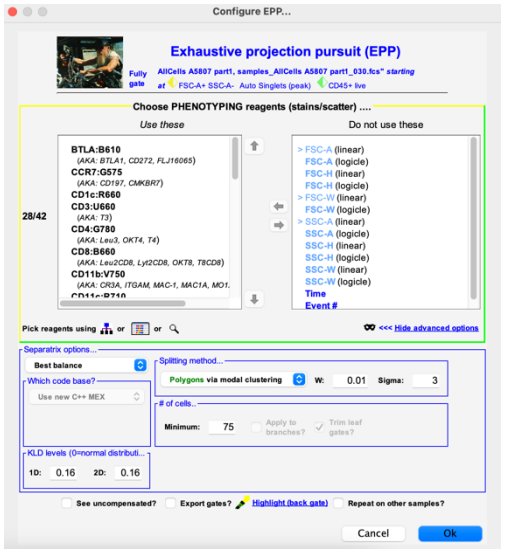
- Begin a fully automated analysis playing with AutoGate's fully automatic tools
  or
- Continue with a conventional analysis by dragging and dropping the provided gating **hierarchy** to other samples

2 Run a fully automated unsupervised analysis with exhaustive projection pursuit (EPP)

2.1 Select CD45+ live 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.



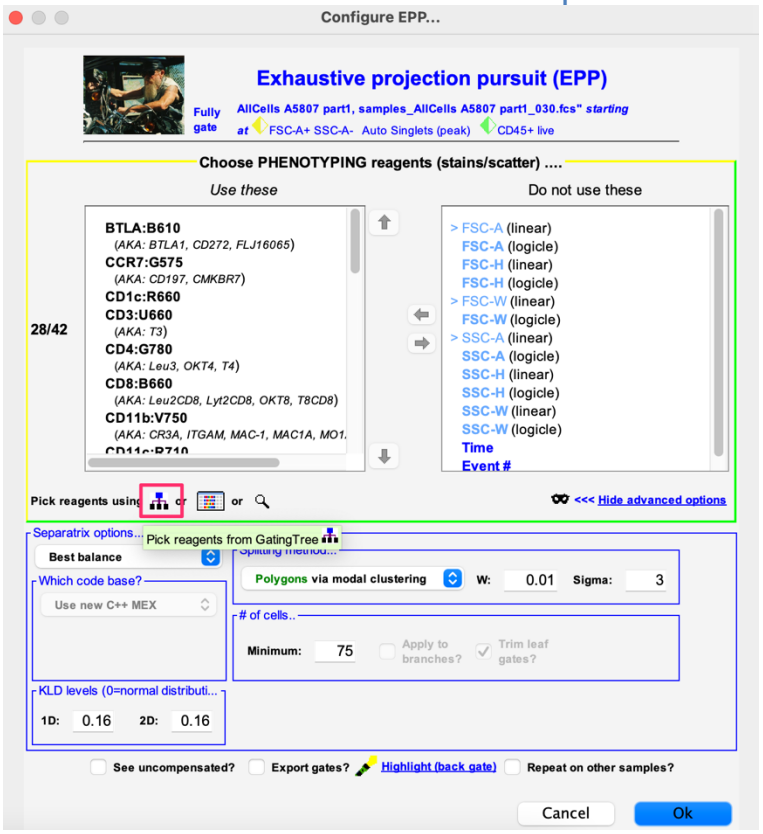
2.3 Specify the parameter for EPP to gate on

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

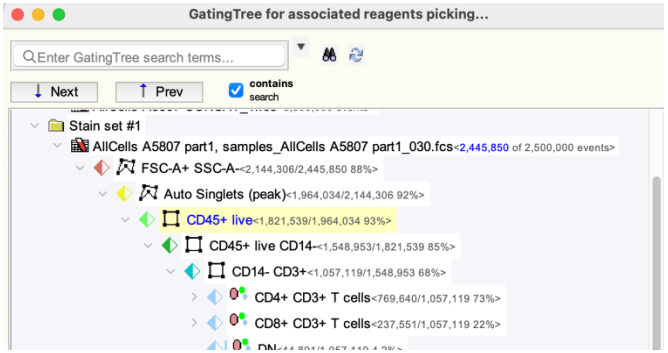
1. Were used for the manual analysis
2. Those that look informative for the same cell types

2.4 Point EPP at parameters used manually for **CD45+ live** cells

2.4.1 Click on the tree icon under the 2 parameter lists



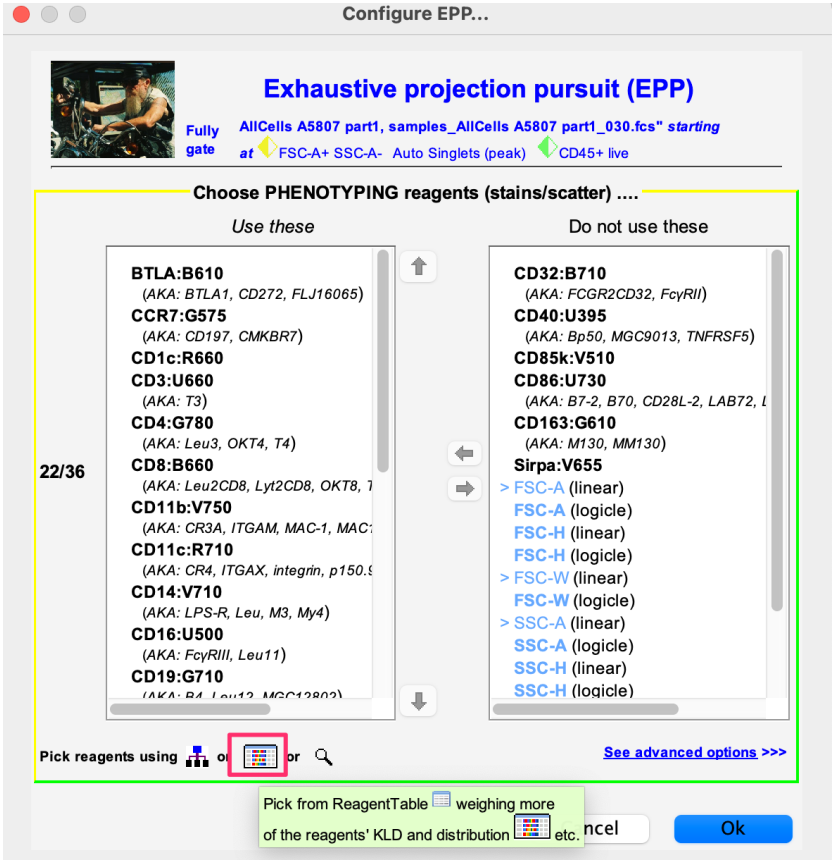
2.4.2 Select CD45+ live cells in the gating tree



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

3 Point EPP at other parameters that appear informative for CD45+ live cells

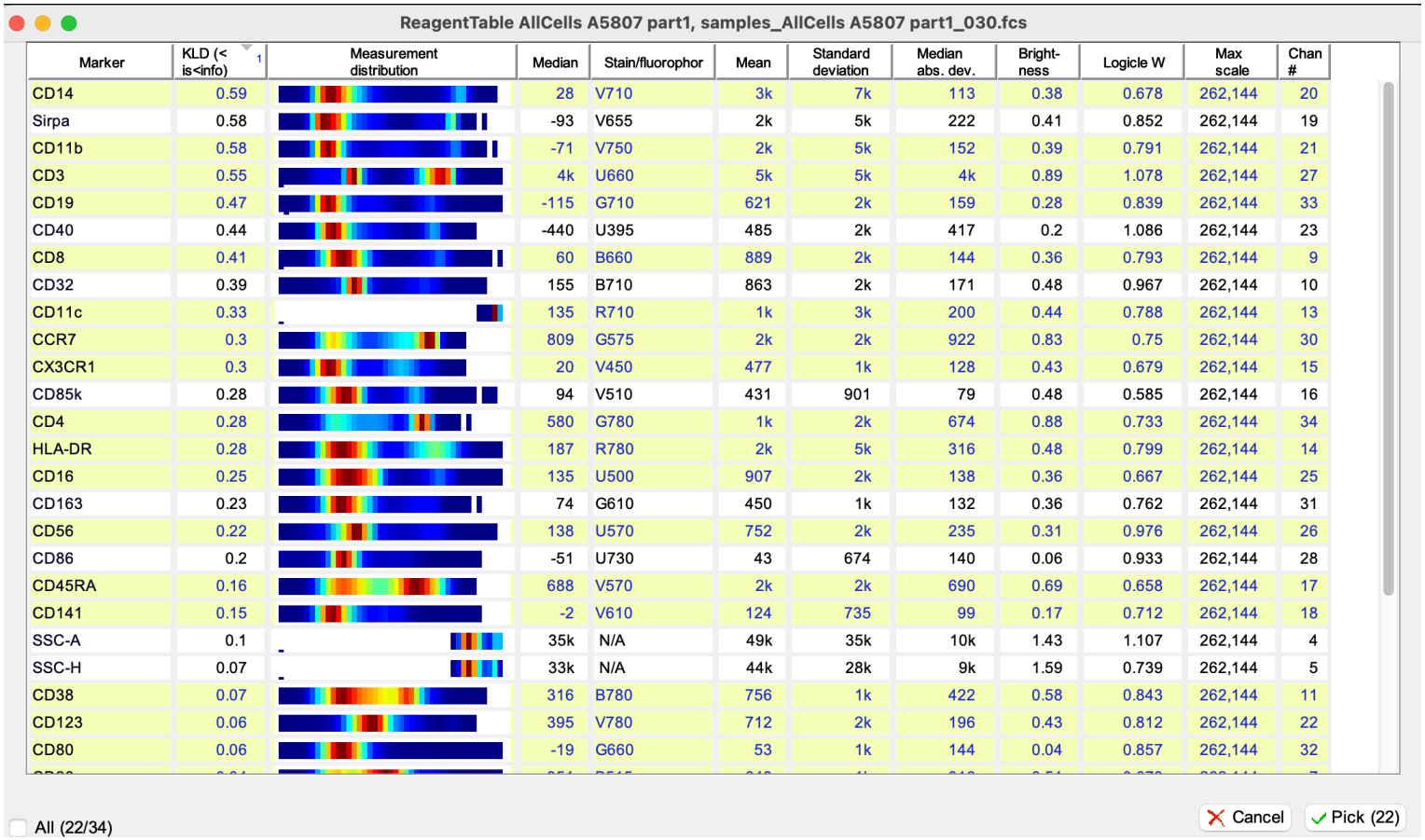
4 Click on the table icon under the 2 parameter lists



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

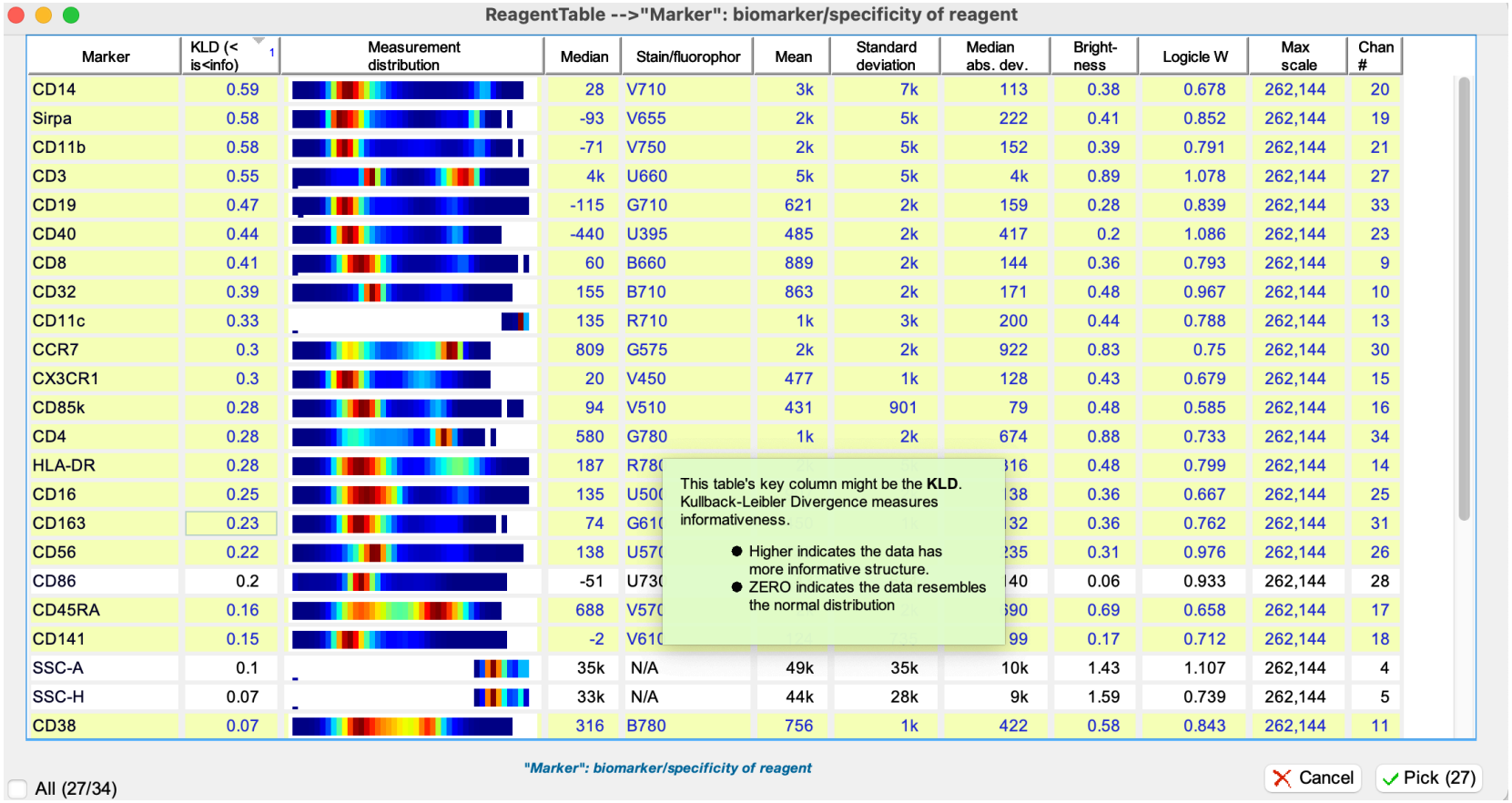
5 Select parameters that have high KLD but are not already selected

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 (< is<info)	Measurement distribution	Median	Stain/fluorophor	Mean	Standard deviation	Median abs. dev.	Bright-ness	Logicle W	Max scale	Chan #
CD14	0.59		28	V710	3k	7k	113	0.38	0.678	262,144	20
Sirpa	0.58		-93	V655	2k	5k	222	0.41	0.852	262,144	19
CD11b	0.58		-71	V750	2k	5k	152	0.39	0.791	262,144	21
CD3	0.55		4k	U660	5k	5k	4k	0.89	1.078	262,144	27
CD19	0.47		-115	G710	621	2k	159	0.28	0.839	262,144	33
CD40	0.44		-440	U395	485	2k	417	0.2	1.086	262,144	23
CD8	0.41		60	B660	889	2k	144	0.36	0.793	262,144	9
CD32	0.39		155	B710	863	2k	171	0.48	0.967	262,144	10
CD11c	0.33		135	R710	1k	3k	200	0.44	0.788	262,144	13
CCR7	0.3		809	G575	2k	2k	922	0.83	0.75	262,144	30
CX3CR1	0.3		20	V450	477	1k	128	0.43	0.679	262,144	15
CD85k	0.28		94	V510	431	901	79	0.48	0.585	262,144	16
CD4	0.28		580	G780	1k	2k	674	0.88	0.733	262,144	34
HLA-DR	0.28		187	R780	2k	5k	316	0.48	0.799	262,144	14
CD16	0.25		135	U500	907	2k	138	0.36	0.667	262,144	25
CD163	0.23		74	G610	450	1k	132	0.36	0.762	262,144	31
CD56	0.22		138	U570	752	2k	235	0.31	0.976	262,144	26
CD86	0.2		-51	U730	43	674	140	0.06	0.933	262,144	28
CD45RA	0.16		688	V570	2k	2k	690	0.69	0.658	262,144	17
CD141	0.15		-2	V610	124	735	99	0.17	0.712	262,144	18
SSC-A	0.1		35k	N/A	49k	35k	10k	1.43	1.107	262,144	4
SSC-H	0.07		33k	N/A	44k	28k	9k	1.59	0.739	262,144	5
CD38	0.07		316	B780	756	1k	422	0.58	0.843	262,144	11
CD123	0.06		395	V780	712	2k	196	0.43	0.812	262,144	22
CD80	0.06		-19	G660	53	1k	144	0.04	0.857	262,144	32

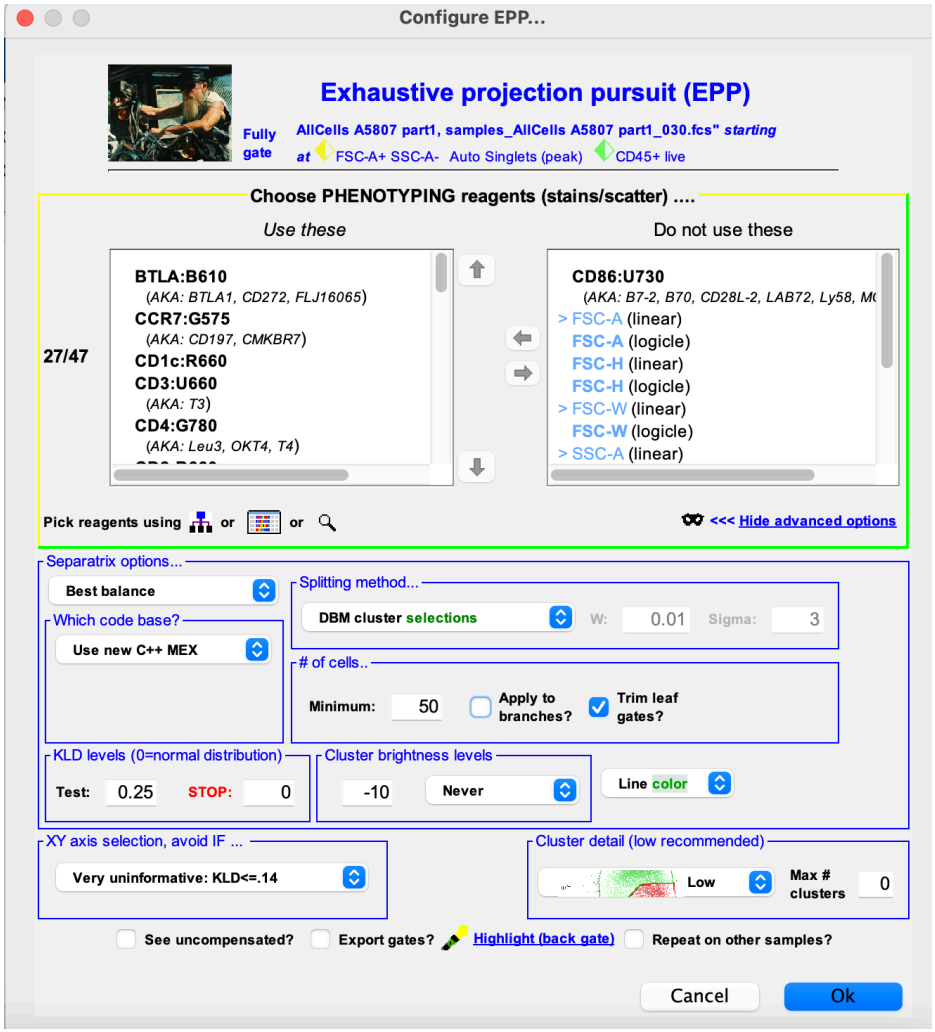
Add (mac key + select / Ctrl key + select) the 6 reagents with the highest KLD that are not already selected: Sirpa, CD40, CD95, CD32, CD85 and CD163



Marker	KLD (< is<info)	Measurement distribution	Median	Stain/fluorophor	Mean	Standard deviation	Median abs. dev.	Bright-ness	Logicle W	Max scale	Chan #
CD14	0.59		28	V710	3k	7k	113	0.38	0.678	262,144	20
Sirpa	0.58		-93	V655	2k	5k	222	0.41	0.852	262,144	19
CD11b	0.58		-71	V750	2k	5k	152	0.39	0.791	262,144	21
CD3	0.55		4k	U660	5k	5k	4k	0.89	1.078	262,144	27
CD19	0.47		-115	G710	621	2k	159	0.28	0.839	262,144	33
CD40	0.44		-440	U395	485	2k	417	0.2	1.086	262,144	23
CD8	0.41		60	B660	889	2k	144	0.36	0.793	262,144	9
CD32	0.39		155	B710	863	2k	171	0.48	0.967	262,144	10
CD11c	0.33		135	R710	1k	3k	200	0.44	0.788	262,144	13
CCR7	0.3		809	G575	2k	2k	922	0.83	0.75	262,144	30
CX3CR1	0.3		20	V450	477	1k	128	0.43	0.679	262,144	15
CD85k	0.28		94	V510	431	901	79	0.48	0.585	262,144	16
CD4	0.28		580	G780	1k	2k	674	0.88	0.733	262,144	34
HLA-DR	0.28		187	R78(			316	0.48	0.799	262,144	14
CD16	0.25		135	U50(			138	0.36	0.667	262,144	25
CD163	0.23		74	G61(			132	0.36	0.762	262,144	31
CD56	0.22		138	U57(			235	0.31	0.976	262,144	26
CD86	0.2		-51	U73(			140	0.06	0.933	262,144	28
CD45RA	0.16		688	V57(			590	0.69	0.658	262,144	17
CD141	0.15		-2	V61(			99	0.17	0.712	262,144	18
SSC-A	0.1		35k	N/A	49k	35k	10k	1.43	1.107	262,144	4
SSC-H	0.07		33k	N/A	44k	28k	9k	1.59	0.739	262,144	5
CD38	0.07		316	B780	756	1k	422	0.58	0.843	262,144	11

6 Run EPP

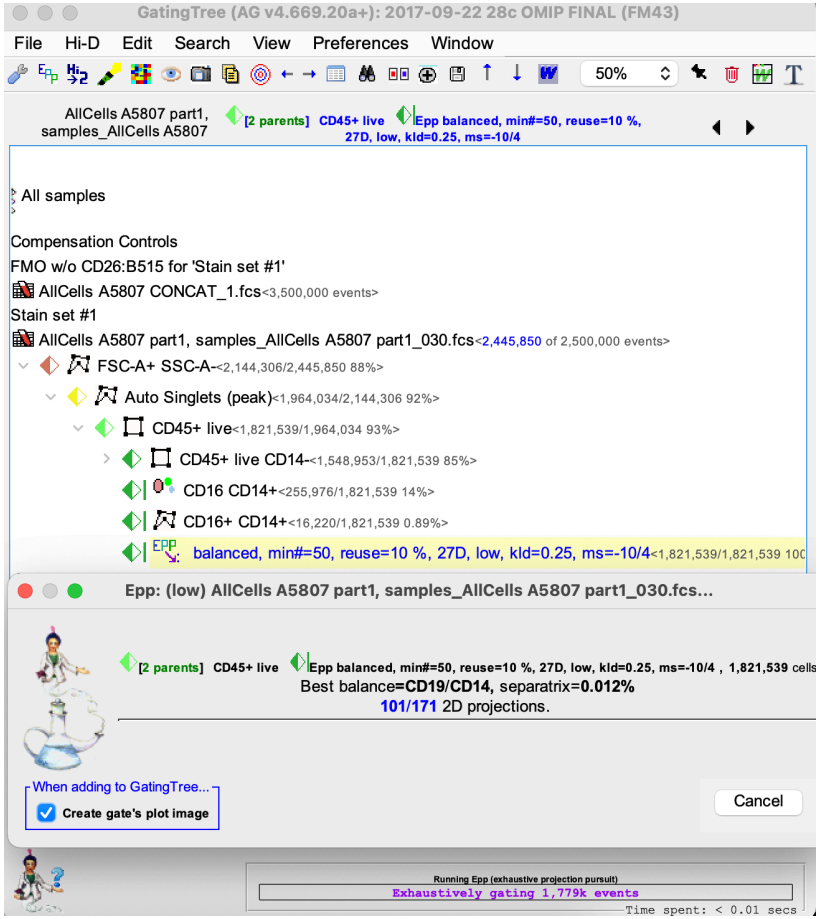
If your EPP configuration settings match the screenshot below then click the Ok button to run EPP!!



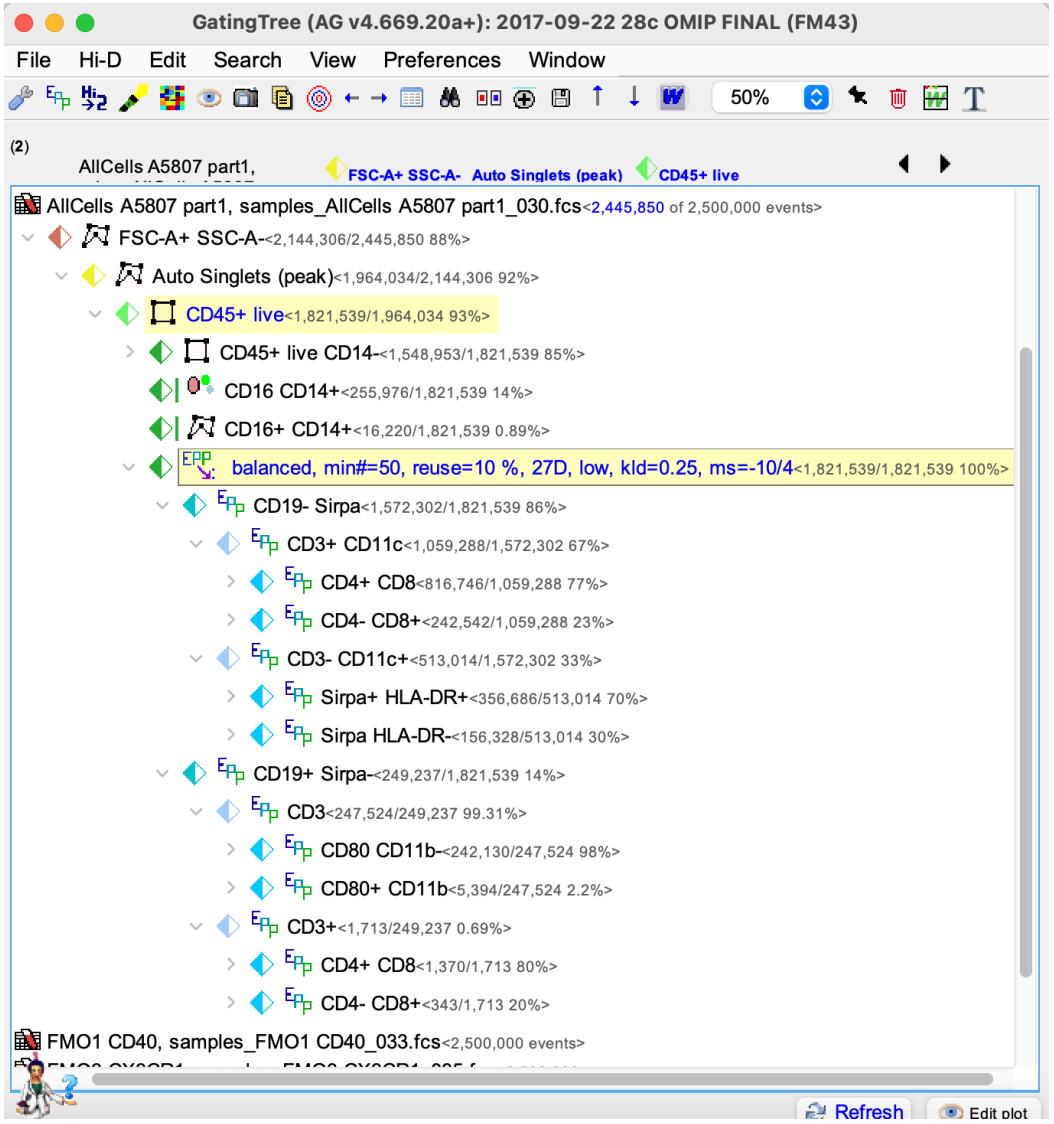
On a 6 core MacBook it takes EPP around 10 minutes to

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



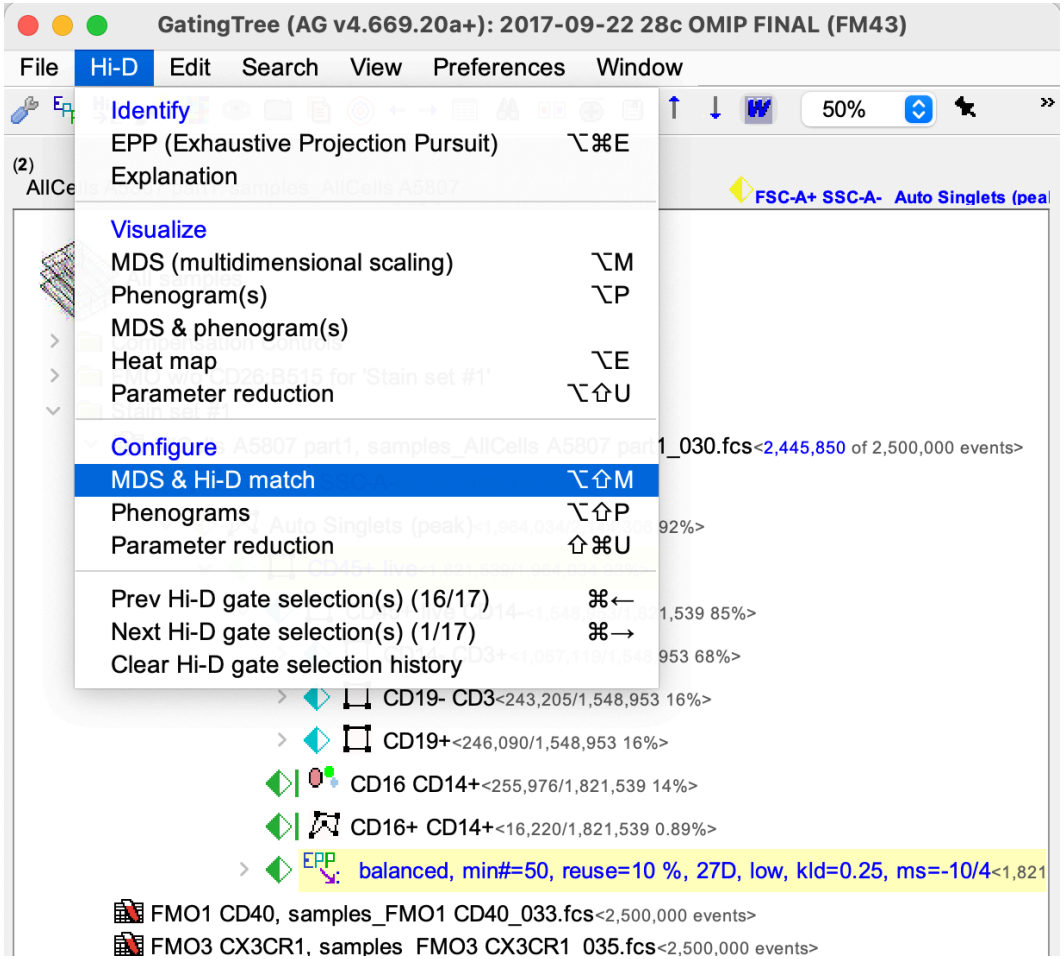
EPP completed



7 Compare EPP to OMIP investigator's manual gates

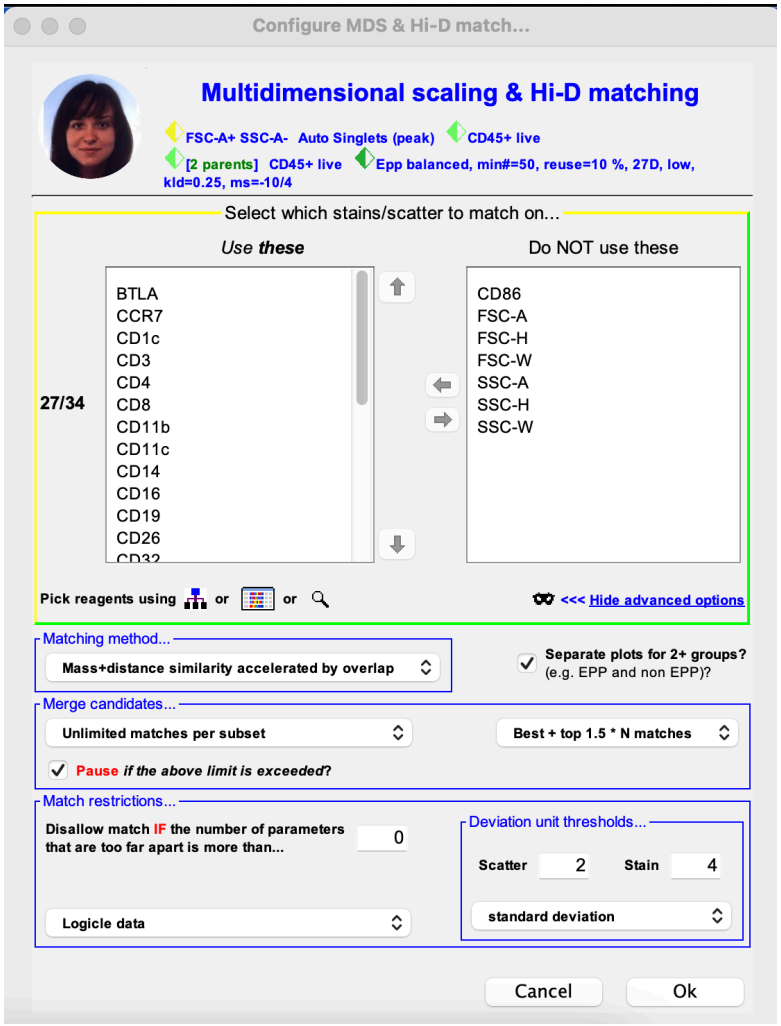
When EPP completes its run it then shows the menu for running AutoGate's matching between your fresh new EP gates and the OMIP investigator's manual gates for the same data

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



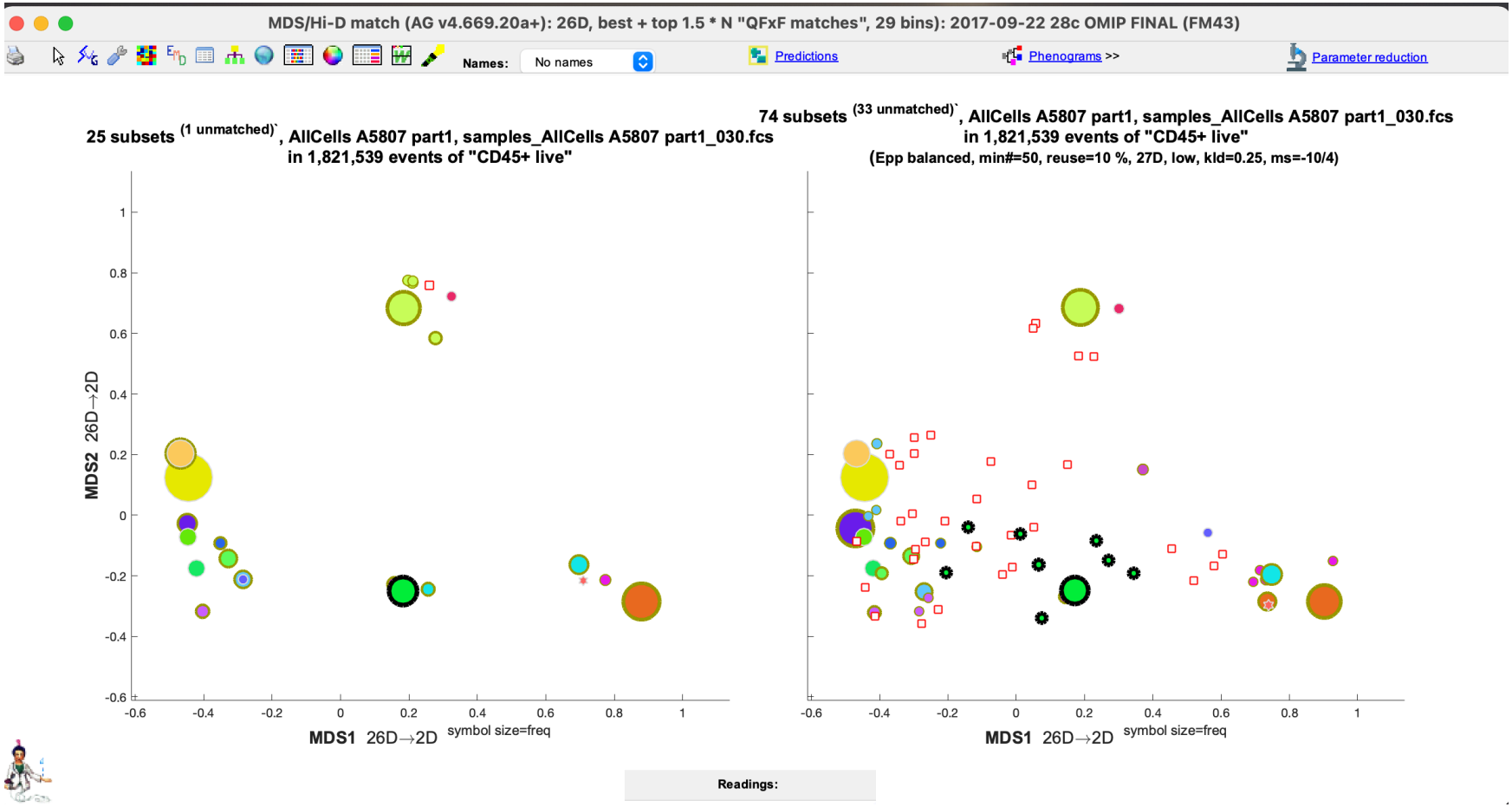
7.2 Start matching

Ensure your settings match as below

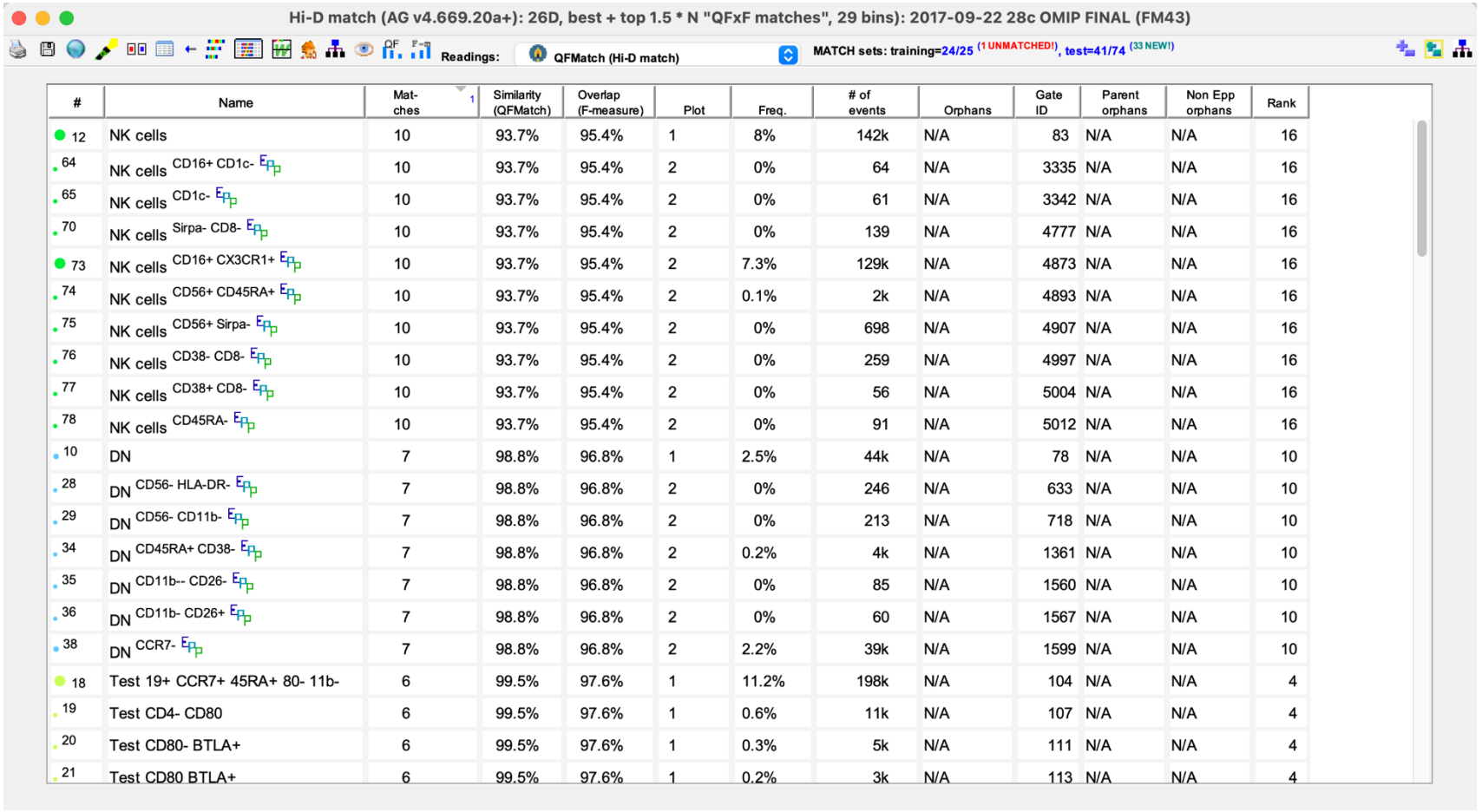


7.3 Evaluate matching

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.



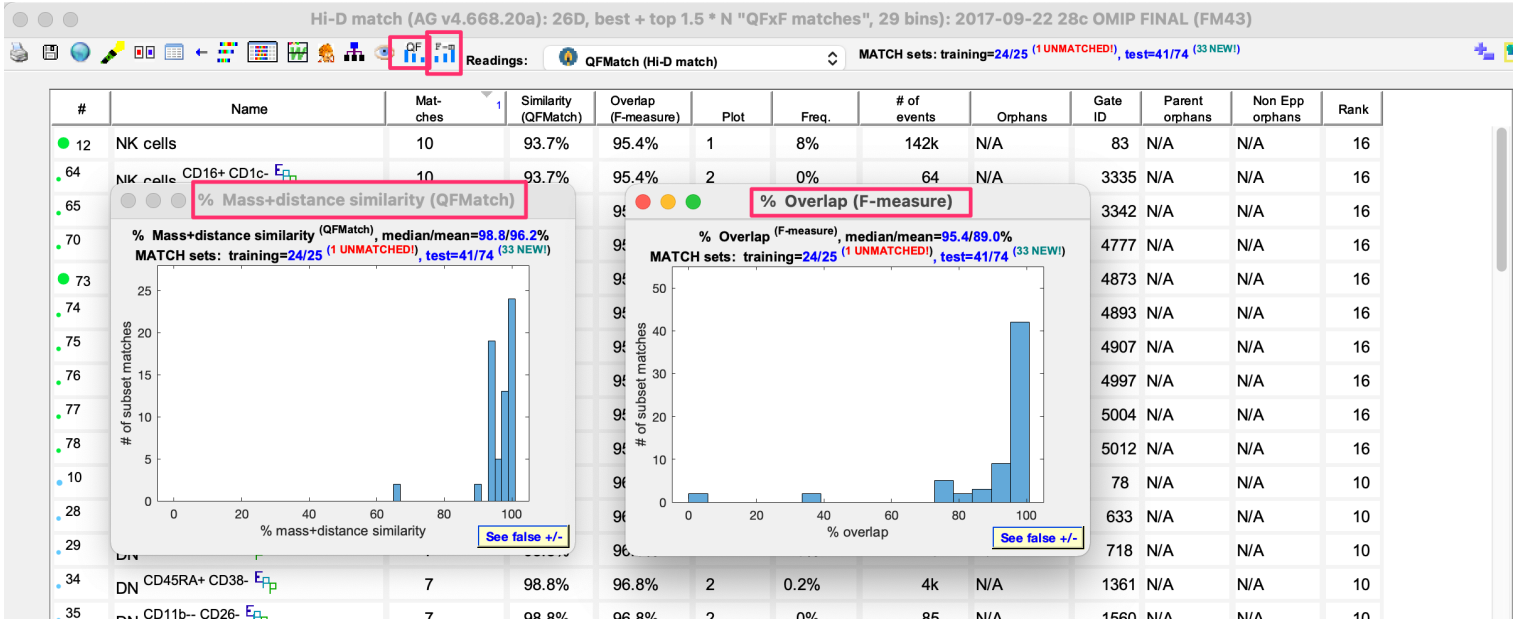
The quickest quantitative summary can be found by clicking on the table button of the toolbar which then shows best to least match with supporting statistics.



#	Name	Mat-ches	Similarity (QFMatch)	Overlap (F-measure)	Plot	Freq.	# of events	Orphans	Gate ID	Parent orphans	Non Epp orphans	Rank
12	NK cells	10	93.7%	95.4%	1	8%	142k	N/A	83	N/A	N/A	16
64	NK cells CD16+ CD1c-	10	93.7%	95.4%	2	0%	64	N/A	3335	N/A	N/A	16
65	NK cells CD1c-	10	93.7%	95.4%	2	0%	61	N/A	3342	N/A	N/A	16
70	NK cells Sirpa- CD8-	10	93.7%	95.4%	2	0%	139	N/A	4777	N/A	N/A	16
73	NK cells CD16+ CX3CR1+	10	93.7%	95.4%	2	7.3%	129k	N/A	4873	N/A	N/A	16
74	NK cells CD56+ CD45RA+	10	93.7%	95.4%	2	0.1%	2k	N/A	4893	N/A	N/A	16
75	NK cells CD56+ Sirpa-	10	93.7%	95.4%	2	0%	698	N/A	4907	N/A	N/A	16
76	NK cells CD38- CD8-	10	93.7%	95.4%	2	0%	259	N/A	4997	N/A	N/A	16
77	NK cells CD38+ CD8-	10	93.7%	95.4%	2	0%	56	N/A	5004	N/A	N/A	16
78	NK cells CD45RA-	10	93.7%	95.4%	2	0%	91	N/A	5012	N/A	N/A	16
10	DN	7	98.8%	96.8%	1	2.5%	44k	N/A	78	N/A	N/A	10
28	DN CD56- HLA-DR-	7	98.8%	96.8%	2	0%	246	N/A	633	N/A	N/A	10
29	DN CD56- CD11b-	7	98.8%	96.8%	2	0%	213	N/A	718	N/A	N/A	10
34	DN CD45RA+ CD38-	7	98.8%	96.8%	2	0.2%	4k	N/A	1361	N/A	N/A	10
35	DN CD11b-- CD26-	7	98.8%	96.8%	2	0%	85	N/A	1560	N/A	N/A	10
36	DN CD11b- CD26+	7	98.8%	96.8%	2	0%	60	N/A	1567	N/A	N/A	10
38	DN CCR7-	7	98.8%	96.8%	2	2.2%	39k	N/A	1599	N/A	N/A	10
18	Test 19+ CCR7+ 45RA+ 80- 11b-	6	99.5%	97.6%	1	11.2%	198k	N/A	104	N/A	N/A	4
19	Test CD4- CD80	6	99.5%	97.6%	1	0.6%	11k	N/A	107	N/A	N/A	4
20	Test CD80- BTLA+	6	99.5%	97.6%	1	0.3%	5k	N/A	111	N/A	N/A	4
21	Test CD80 BTLA+	6	99.5%	97.6%	1	0.2%	3k	N/A	113	N/A	N/A	4

7.4 Mass + Distance similarity and overlap (F- measure)

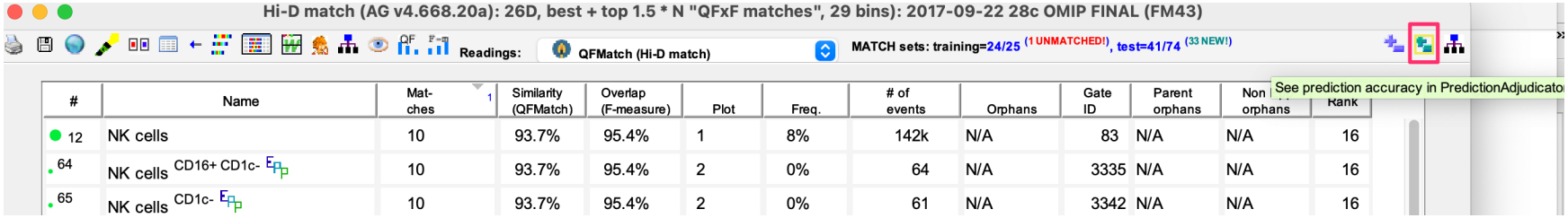
Click the highlighted buttons from the toolbar to view the graphs



7.5 Prediction Adjudicator

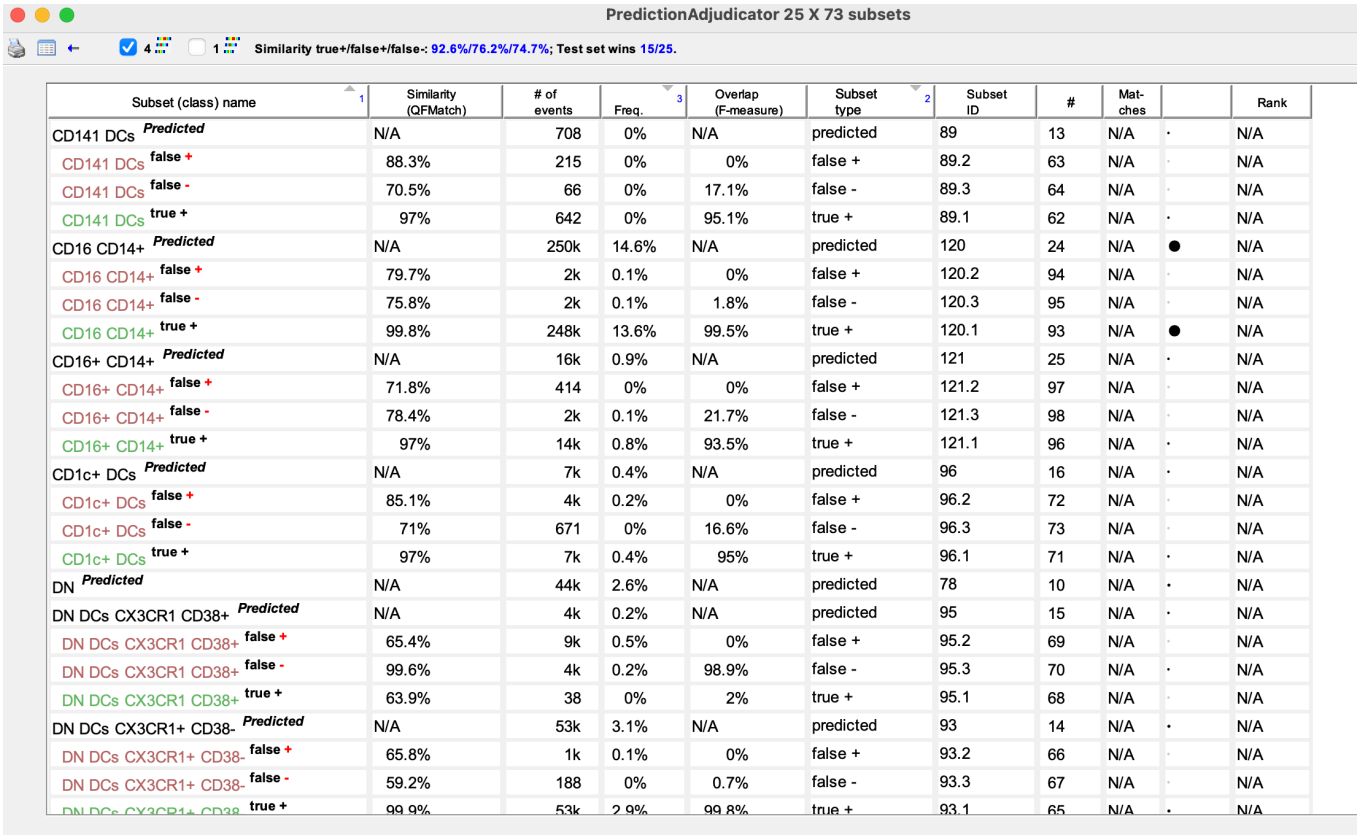
This helps determine how well one classification's subsets predict another's. PA reorganizes the predicting classifier's subsets into predicting subsets: true positive, false positive and false negative subsets. PA determines whether the false positives or false negatives have more QFMatch based similarity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

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



#	Name	Mat-ches	Similarity (QFMatch)	Overlap (F-measure)	Plot	Freq.	# of events	Orphans	Gate ID	Parent orphans	Non orphans	Rank
12	NK cells	10	93.7%	95.4%	1	8%	142k	N/A	83	N/A	N/A	16
64	NK cells CD16+ CD1c-	10	93.7%	95.4%	2	0%	64	N/A	3335	N/A	N/A	16
65	NK cells CD1c-	10	93.7%	95.4%	2	0%	61	N/A	3342	N/A	N/A	16

PA results below



Subset (class) name	Similarity (QFMatch)	# of events	Freq.	Overlap (F-measure)	Subset type	Subset ID	#	Mat-ches		Rank
CD141 DCs *Predicted*	N/A	708	0%	N/A	predicted	89	13	N/A	·	N/A
CD141 DCs false +	88.3%	215	0%	0%	false +	89.2	63	N/A		N/A
CD141 DCs false -	70.5%	66	0%	17.1%	false -	89.3	64	N/A		N/A
CD141 DCs true +	97%	642	0%	95.1%	true +	89.1	62	N/A	·	N/A
CD16 CD14+ *Predicted*	N/A	250k	14.6%	N/A	predicted	120	24	N/A	●	N/A
CD16 CD14+ false +	79.7%	2k	0.1%	0%	false +	120.2	94	N/A		N/A
CD16 CD14+ false -	75.8%	2k	0.1%	1.8%	false -	120.3	95	N/A		N/A
CD16 CD14+ true +	99.8%	248k	13.6%	99.5%	true +	120.1	93	N/A	●	N/A
CD16+ CD14+ *Predicted*	N/A	16k	0.9%	N/A	predicted	121	25	N/A	·	N/A
CD16+ CD14+ false +	71.8%	414	0%	0%	false +	121.2	97	N/A		N/A
CD16+ CD14+ false -	78.4%	2k	0.1%	21.7%	false -	121.3	98	N/A		N/A
CD16+ CD14+ true +	97%	14k	0.8%	93.5%	true +	121.1	96	N/A	·	N/A
CD1c+ DCs *Predicted*	N/A	7k	0.4%	N/A	predicted	96	16	N/A	·	N/A
CD1c+ DCs false +	85.1%	4k	0.2%	0%	false +	96.2	72	N/A		N/A
CD1c+ DCs false -	71%	671	0%	16.6%	false -	96.3	73	N/A		N/A
CD1c+ DCs true +	97%	7k	0.4%	95%	true +	96.1	71	N/A	·	N/A
DN *Predicted*	N/A	44k	2.6%	N/A	predicted	78	10	N/A	·	N/A
DN DCs CX3CR1 CD38+ *Predicted*	N/A	4k	0.2%	N/A	predicted	95	15	N/A	·	N/A
DN DCs CX3CR1 CD38+ false +	65.4%	9k	0.5%	0%	false +	95.2	69	N/A		N/A
DN DCs CX3CR1 CD38+ false -	99.6%	4k	0.2%	98.9%	false -	95.3	70	N/A	·	N/A
DN DCs CX3CR1 CD38+ true +	63.9%	38	0%	2%	true +	95.1	68	N/A		N/A
DN DCs CX3CR1+ CD38- *Predicted*	N/A	53k	3.1%	N/A	predicted	93	14	N/A	·	N/A
DN DCs CX3CR1+ CD38- false +	65.8%	1k	0.1%	0%	false +	93.2	66	N/A		N/A
DN DCs CX3CR1+ CD38- false -	59.2%	188	0%	0.7%	false -	93.3	67	N/A		N/A
DN DCs CX3CR1+ CD38- true +	99.9%	53k	2.9%	99.8%	true +	93.1	65	N/A	·	N/A

7.6 Get False positive/negative rate

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