WEEK 6 ASSIGNMENT

Introduction to Computational Biology – BIOL 509000 | Fall 1 2020

Christina Morgenstern

Question #1: What member lab(s) have been credited with generating the data available for αβ T cells?

The Goldrath lab, the Kronenberg lab and the CBDM (Department of Immunology at Harvard Medical School) lab are responsible for generating data for αβ T cells (see Fig. 1)
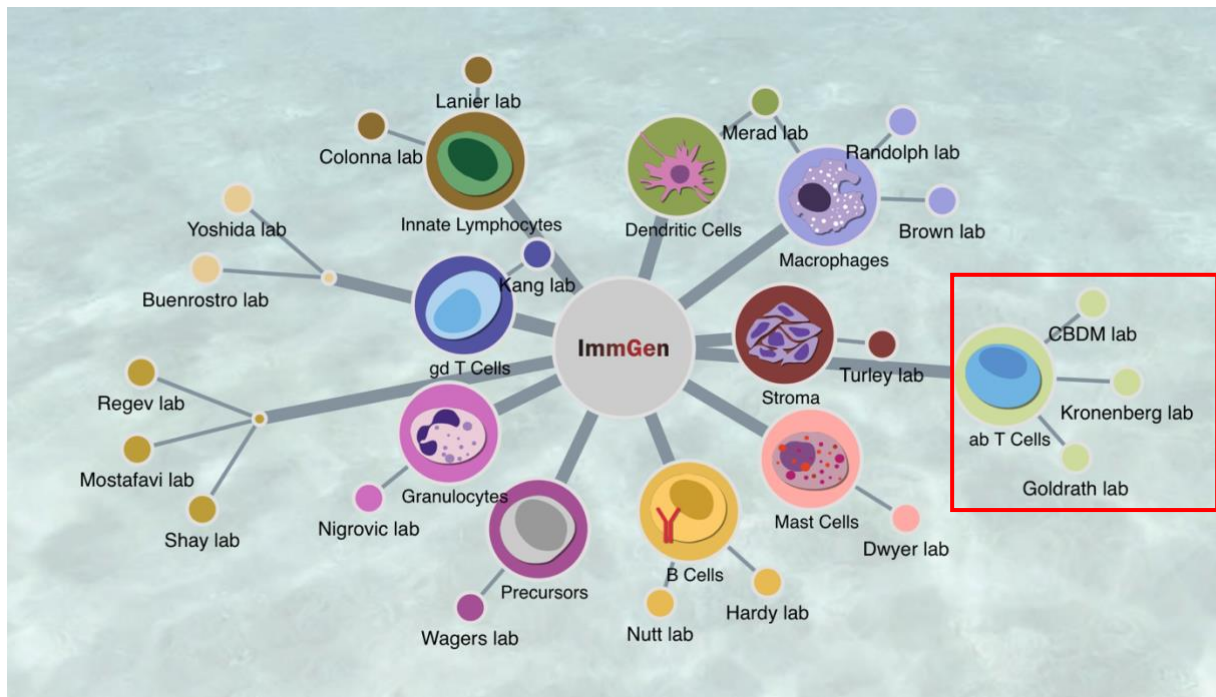

Figure 1. Members of the ImmGen consortium.

Question #2: In dot plot #4, what percentage of cells fall in the gate that is isolating cKit positive but Sca-1 negative cells?

51.1% of cells are cKit positive and Sca-1 negative.

Question #3: Which cell population expresses the highest level of Ccnd3? Which expresses the lowest level?

preB.FrC.BM cells express the highest level of *Ccnd3* (expression value: 1883.02).
B.MZ.Sp cells express the lowest level of *Ccnd3* (expression value: 229.75). See figure 2.
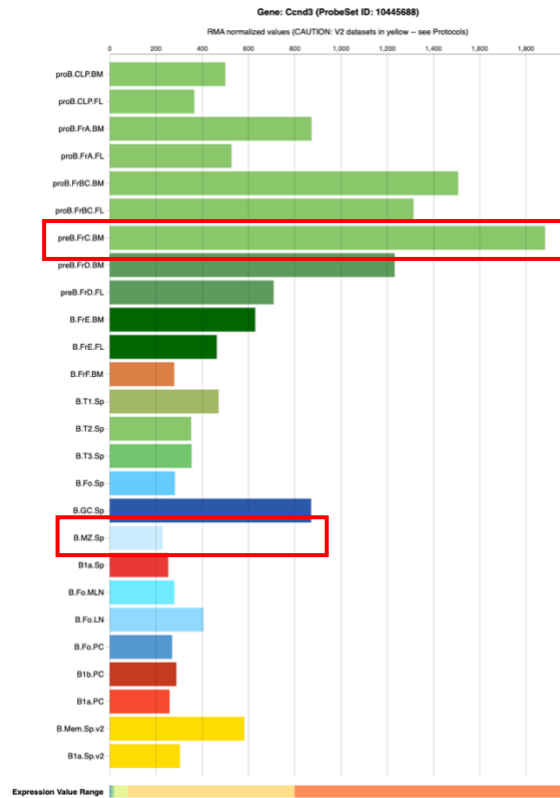
Figure 2. Expression of *Ccnd3* in different B cell populations.

Question #4: What are the four cell populations with the highest level of Ccnd3 expression?

1. T.DPsm.Th (expression value: 2187.67)
2. GN.BM (expression value: 1502.12)
3. NK.MCMV1.Sp (expression value: 1445.21)
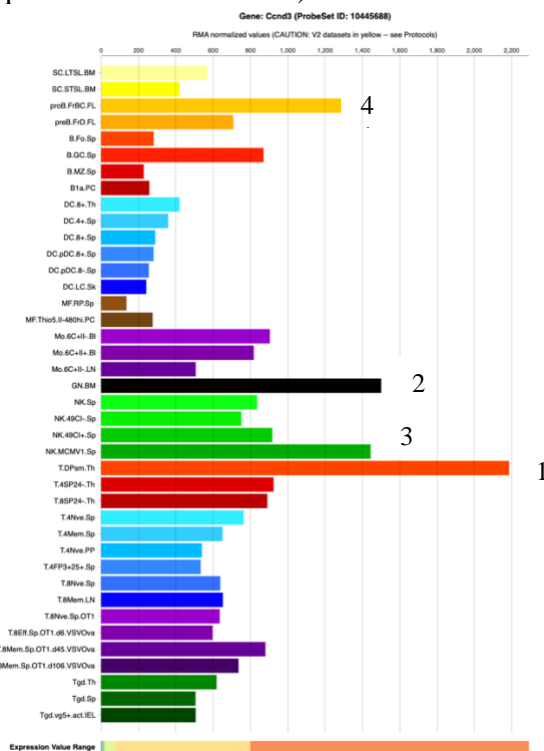4. proB.FrBC.FL (expression value: 1314.66)

Figure 3. Expression of *Ccnd3* in reference populations.

Question #5: Once the correlated factors are grouped by chromosome, determine how many and which chromosomes are represented.
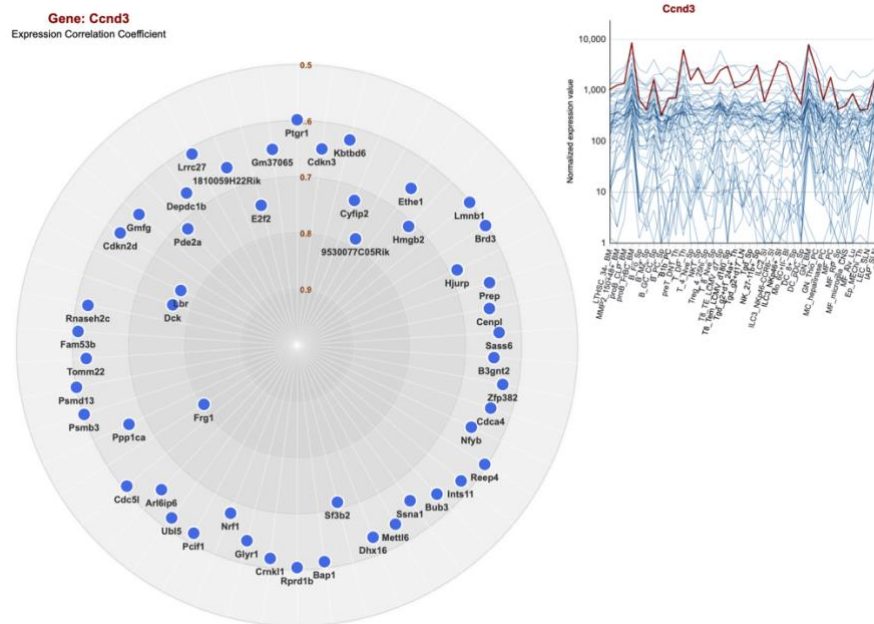

Figure 4. Expression of *Ccnd3* in reference populations.

The following 7 chromosomes are represented (see Fig. 5)
- chromosome 1 (Sgk3)
- chromosome 4 (Nadk)
- chromosome 7 (Xylt1, Pnkp, Klf13)
- chromosome 13 (Cap2)
- chromosome 15 (Nfam1)
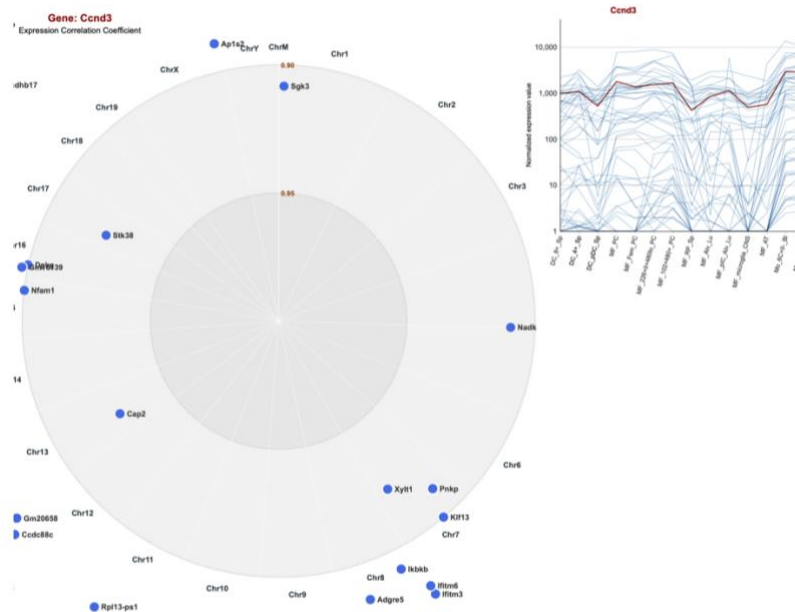- chromosome 16 (Dgkg)
- chromosome 17 (Stk38)

Figure 5. Expression of *Ccnd3* in myeloid cells arranged by gene location.

Question #6: As stated under the Gene Ontology section, what processes are Pnkp involved in?

Pnkp is involved in processes around DNA damage detection, response and repair. Further processes include (nucleotide) phosphorylation and dephosphorylation as well as negative regulation of protein ADP-ribosylation, metabolic processes and response to oxidative stress. Positive regulation of telomerase activity, telomere capping and telomer maintenance are also processes where Pnkp is involved in. See Fig. 6 for a detail list of processes for Pnkp.

Process	Evidence Code	Pubs
DNA damage response, detection of DNA damage	ISO	
DNA repair	IBA	PubMed
DNA repair	ISO	
cellular response to DNA damage stimulus	IEA	
dephosphorylation	ISO	
metabolic process	IEA	
negative regulation of protein ADP-ribosylation	ISO	
nucleotide phosphorylation	IBA	PubMed
nucleotide phosphorylation	ISO	
phosphorylation	IEA	
positive regulation of double-strand break repair via nonhomologous end joining	ISO	
positive regulation of telomerase activity	ISO	
positive regulation of telomere capping	ISO	
positive regulation of telomere maintenance via telomerase	ISO	
response to oxidative stress	ISO	

Figure 6. Involved processes for *Pnkp*.

Question #7: What are the four cell populations with the highest level of Ccnd3 expression? For each population, what is the reported expression value, and how many samples were included in the study?

1. proB.FrBC.BM (expression value: 8430.02, 2 samples)
2. GN.BM (expression value: 7708.01, 2 samples)
3. T.DP.Th (expression value: 6155.75, 2 samples)
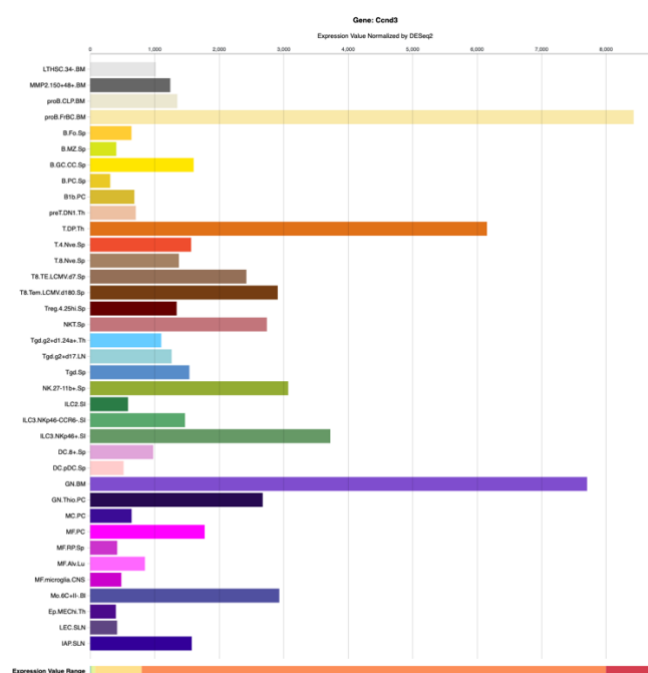4. ILC3.NKp46+.SI (expression value: 3725.42, 2 samples)

Figure 7. Expression values for *Ccnd3*.

Question #8: Looking at the Lck skyline, which populations have expression level greater than 2551?

The following cell populations show expression levels greater than 2551 for *Lck*:
- T.DPsm.Th (expression value: 5437.81)
- T.4SP24-.Th (expression value: 3862.93)
- T.8SP24-.Th (expression value: 3714.25)
- T.4Nve.Sp (expression value: 3349.96)
- T.4Mem.Sp (expression value: 3374.3)
- T.4Nve.PP (expression value: 3349.96)
- T.8Nve.Sp (expression value: 2929.93)
- T.8Mem.LN (expression value: 3156.65)
- T.8Nve.Sp.OT1 (expression value: 3485.87)
- T.8Eff.Sp.OT1.d6.VSVOva (expression value: 3627.6)
- T.8Mem.Sp.OT1.d45.VSVOva (expression value: 4002.12)
- T.8Mem.Sp.OT1.d106.VSVOva (expression value: 3260.34)
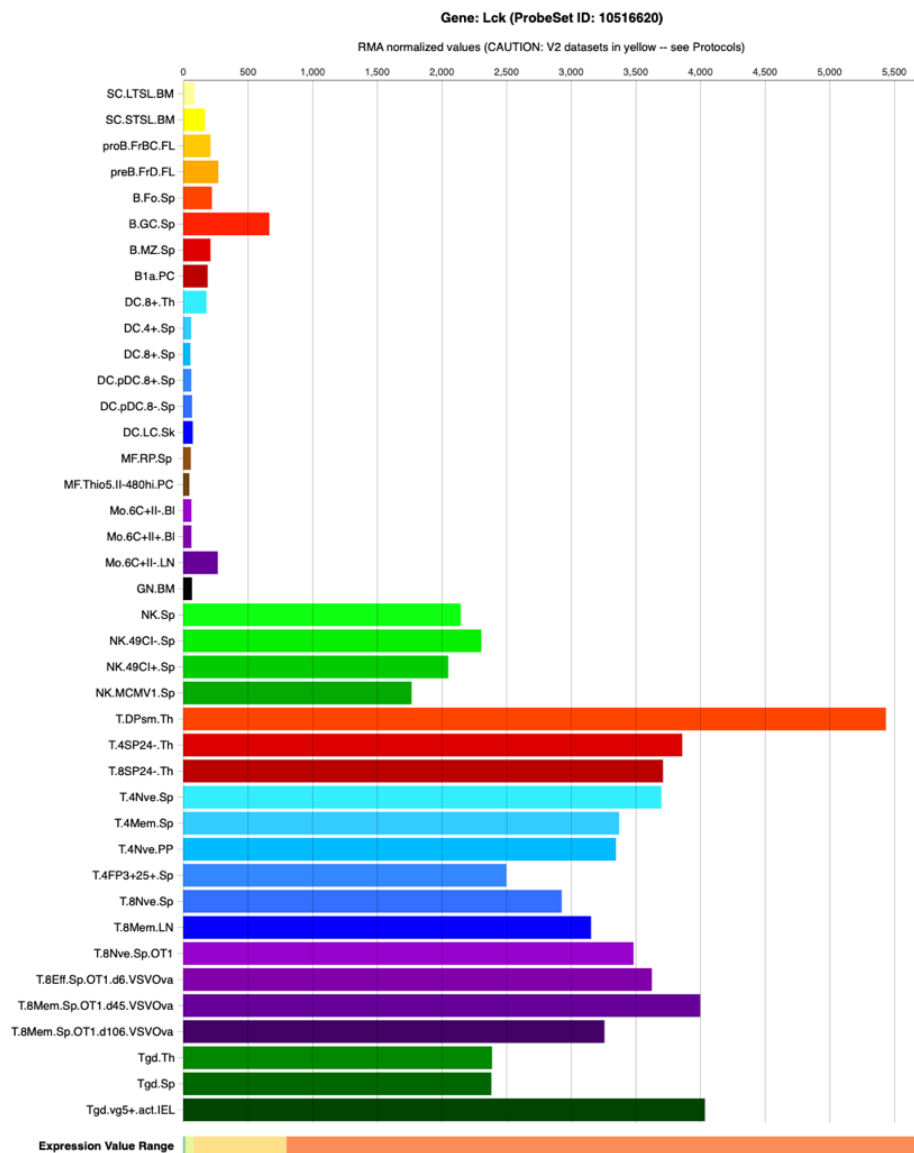- Tgd.vg5+.act.IEL (expression value: 4038.24)

Figure 8. Gene Skyline plot for *Lck*.

Question #9: Hover over the "B_Fo_Sp" name in the Group A box. What criteria are listed in the description for how these cells were sorted (give the factor name and whether +/-). Hover over the "B_GC_Sp" name in the Group B box. What criteria are listed in the description for how these cells were sorted?

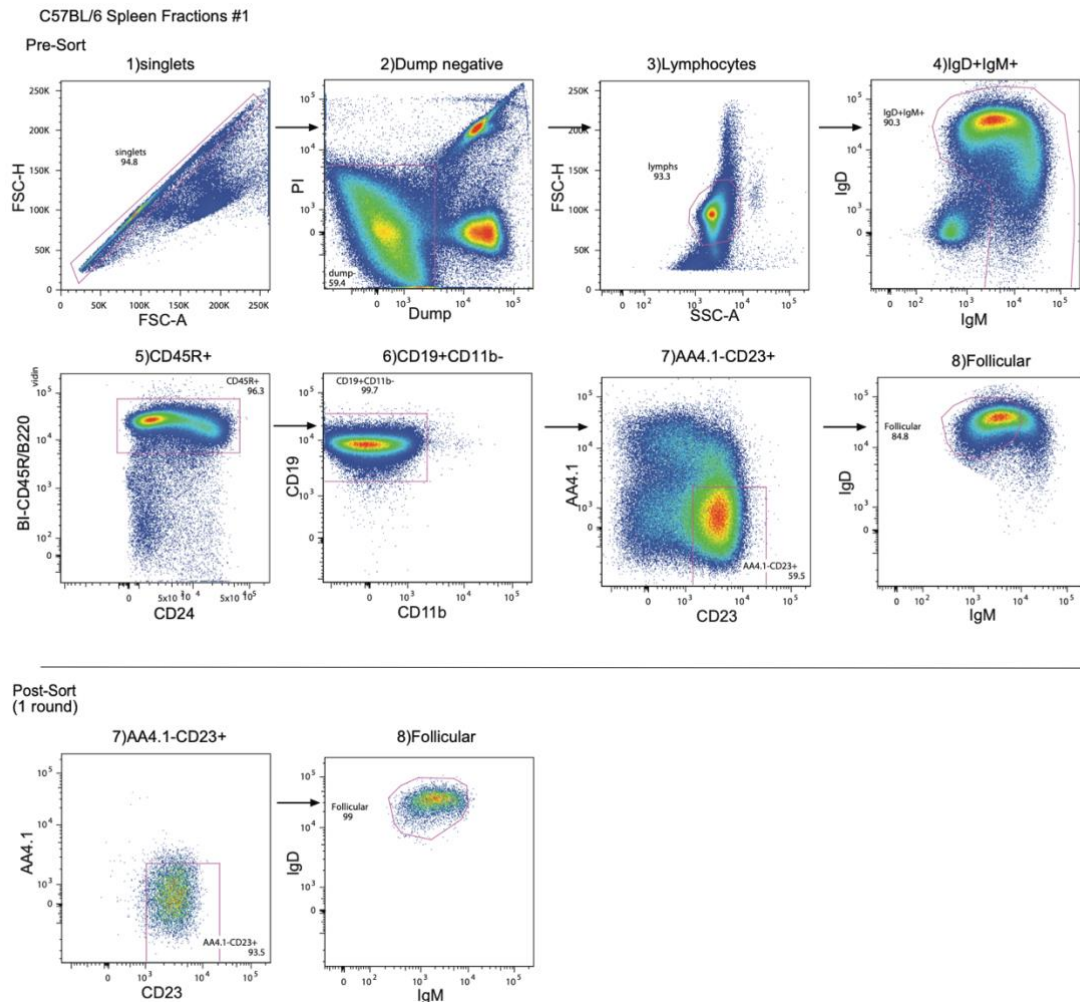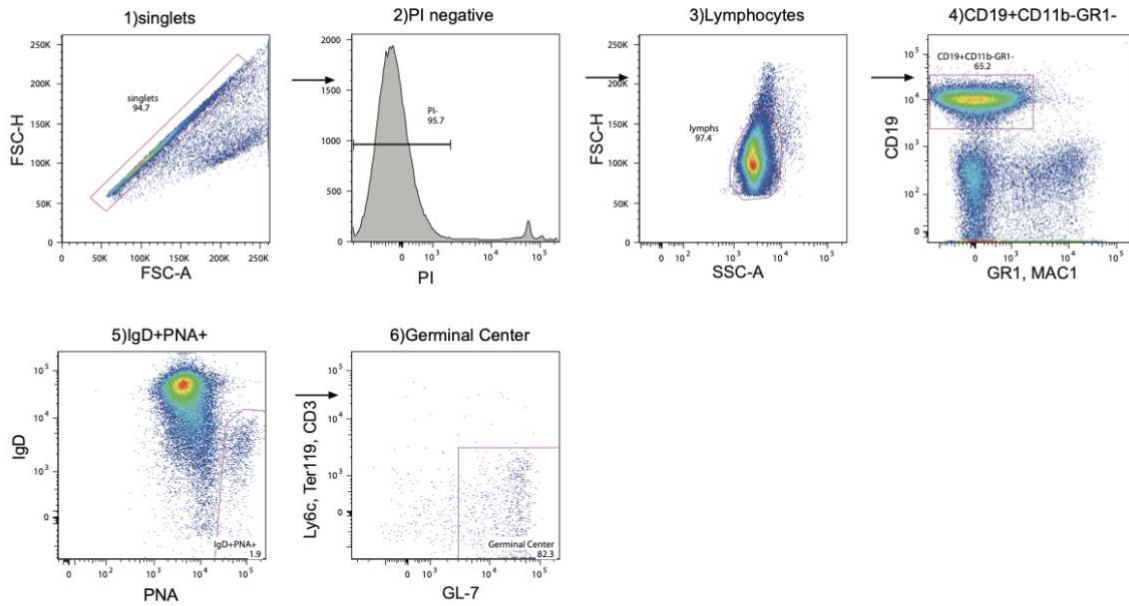After the pre-sorting, the B_Fo_Sp B cells were sorted for AA4.1- and CD23+ (see Fig. 9, bottom panel).

Figure 9. Sorting scheme for B_Fo_Sp B cell population.

After the pre-sorting, the B_BC_Sp B cells were sorted for IgD+ and PNA+ (see Fig. 10, bottom panel).
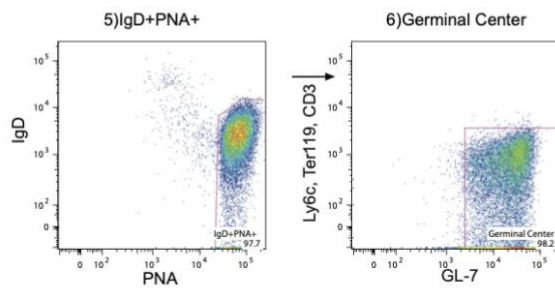
Pre-Sort

Post-Sort
(1 round)

Figure 10. Sorting scheme for B_GC_Sp B cell population.

Question #10: Once sorted by fold change, what are the top five genes that are differentially expressed between these two cell populations? For each, also list the fold change value and which cell population expresses the gene more.

The top five differentially expressed genes between B_Fo_Sp and B_GC_Sp B cell populations are *Sell, Cd55, Zfp318, Ifi203* and *Gm4759* (see Fig. 11).

Gene Symbol	SkyLine	Gene Description	ProbeSetID	Fold Change	P value	FDR	mean_A	mean_B
Sell	🔗	selectin lymphocyte	10351197	61.01	5.1e-13	6.3e-09	5073	83
Cd55	🔗	CD55 antigen	10357488	35.24	3.6e-11	2.4e-08	2303	65
Zfp318	🔗	zinc finger protein 318	10445531	23.48	3.1e-12	8.7e-09	2202	94
Ifi203	🔗	interferon activated gene 203	10360391	23.44	1.2e-11	1.5e-08	1994	85
Gm4759	🔗	GTPase very large interferon inducible 1 pseudogene	10566580	23.33	1.7e-11	1.5e-08	919	39

Figure 11. Top five differentially expressed genes between F_Fo_Sp and F_GC_Sp.