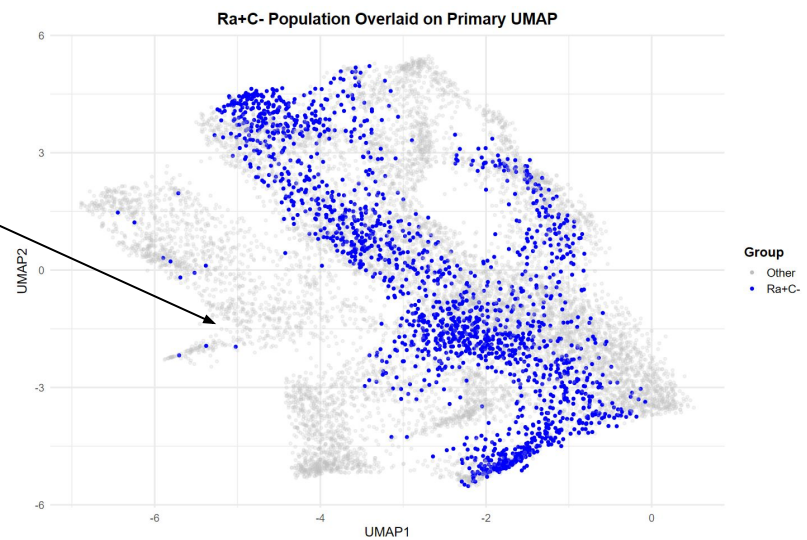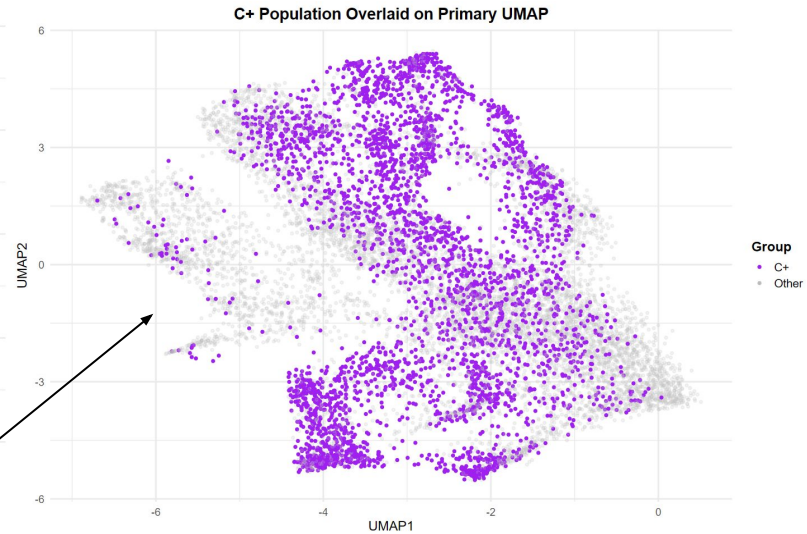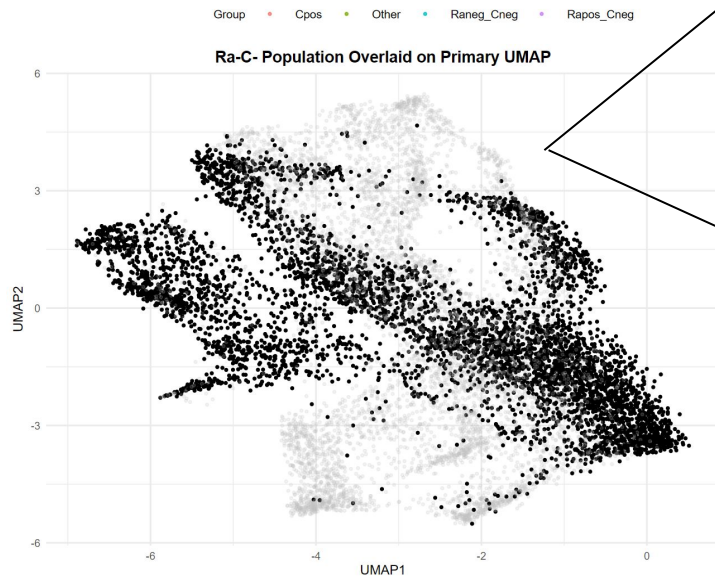# CITE-Seq Temporary Results - Aim 1

Last Update : March 27th

Aim1: Characterizing Dynamics of the 3 Populations

Pooled UMAP from Days 7,10,13

UMAP of Day 7 - Day 13 Gene Expression by Group

Group   Cpos   Other   Raneg_Cneg   Rapos_Cneg

C+ Population Overlaid on Primary UMAP

Group
C+
Other

Ra-C- Population Overlaid on Primary UMAP

Group
Other
Ra-C-

Ra+C- Population Overlaid on Primary UMAP

Group
Other
Ra+C-
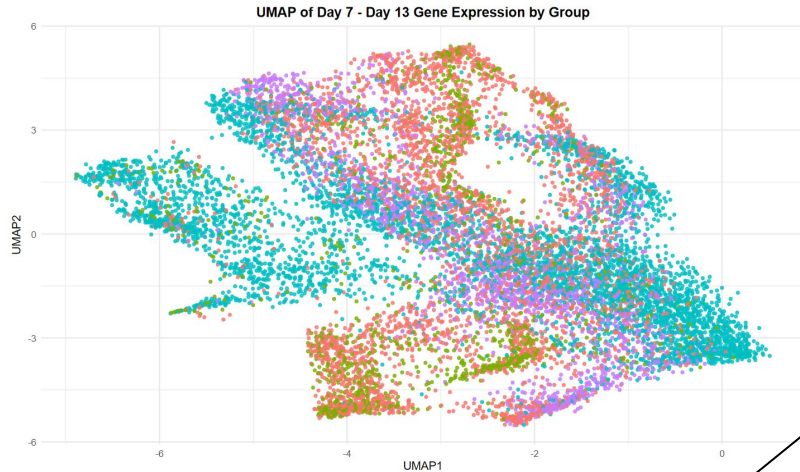
RQ1: Are there regions unique to each of the 3 populations in the Gene Space?

Unique Region for Each Population
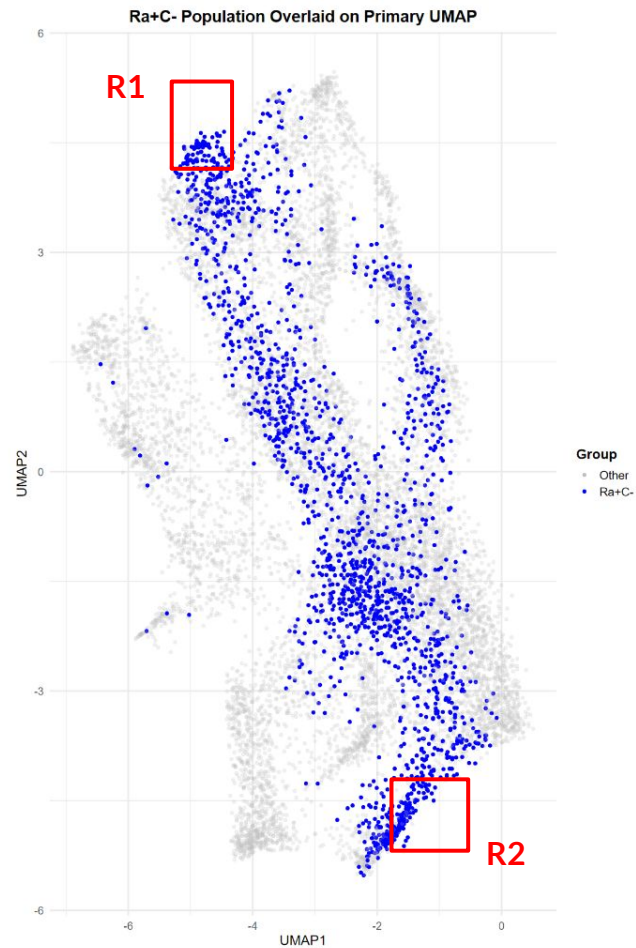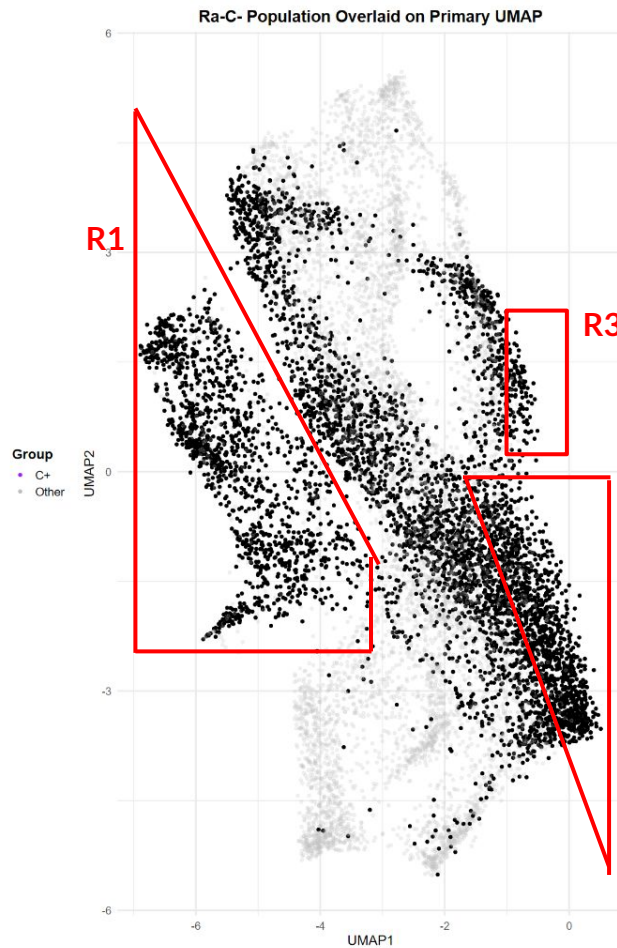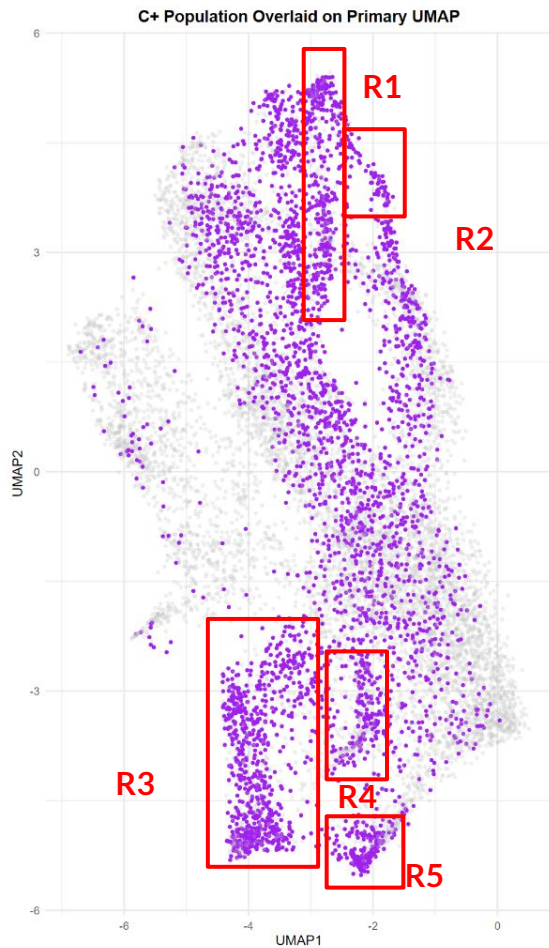
Actual Segmentation By Drawing Polygons

RQ2: What Genes are Differentially Expressed in these Unique Regions?

Characterizing R1 of the Ra+C- Population : Top Genes/CD Markers

Top 20 Genes

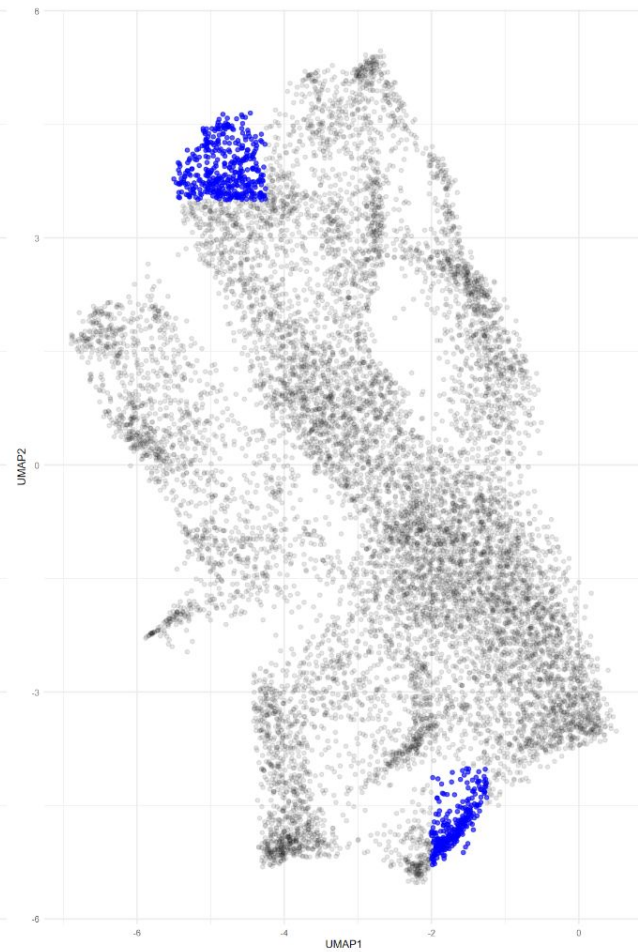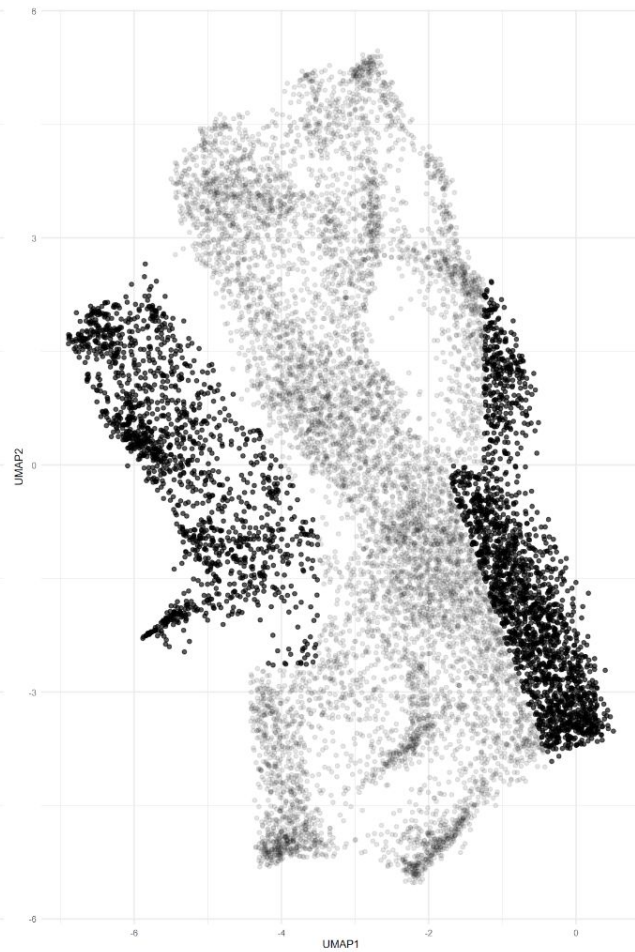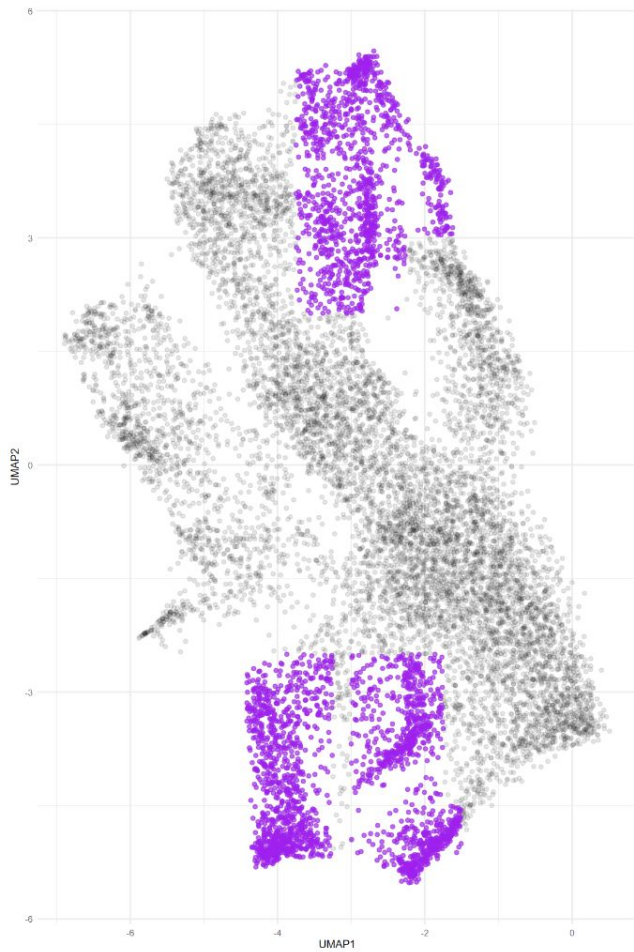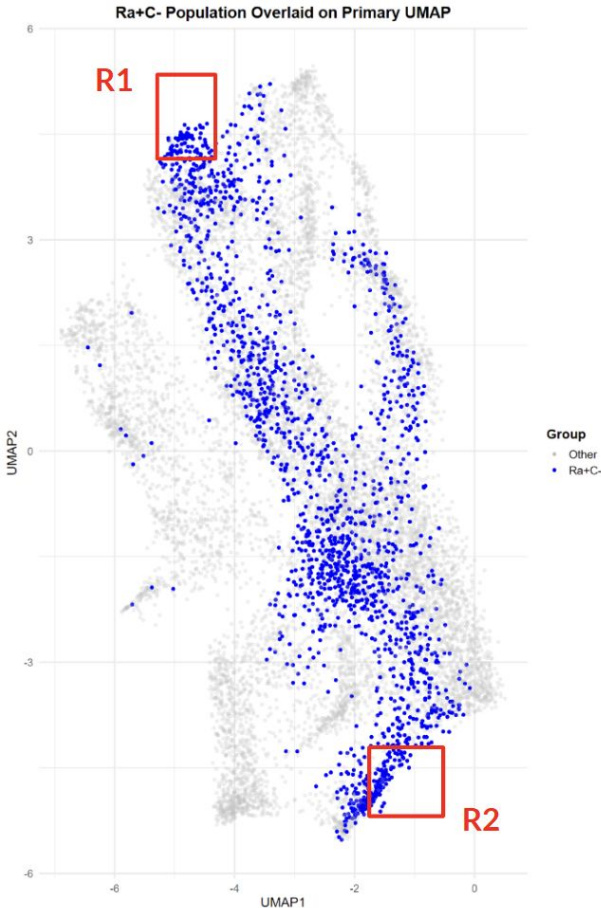| ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
HBD	2.600558	0.9332628
STXBP5	2.387527	1.9199339
ITGA2B	2.281450	0.7618509
LTBP1	2.165247	0.6594841
GP1BB	2.099730	0.6598842
RAP1B	1.979228	2.1936284
ABCC4	1.925287	1.0928290
SPINK2	-1.924872	1.7426931
PLCB1	-1.875570	1.8908794
RNF220	-1.828865	2.5707110
SLC24A3	1.793877	0.6780904
MED12L	1.747457	0.8803669
C1QTNF4	-1.742253	1.5501189
PLXDC2	1.705843	0.9520478
SH3BGRL3	1.699835	2.6590816
ATP8B4	-1.679106	1.6559922
NKAIN2	-1.664888	2.7157017
RAB27B	1.629171	0.9554767
UBE2C	1.595314	0.9767416
LAT	1.566811	0.4916719

Top 20 CD Markers

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
CD74	-1.3886235	2.2728535
CD99	-1.0149487	1.4580142
CD84	0.9984581	0.4955969
CD48	-0.9684666	0.8941089
CD63	0.9135788	2.2570242
CD36	0.8744668	0.3246596
CD34	-0.8677303	1.3150180
CD44	-0.8126885	1.2309836
CD55	0.7716636	0.5686104
CD52	-0.6601983	1.5512427
CD53	-0.6224523	0.8918330
CD164	-0.4921140	1.8707070
CD200	-0.4793880	0.4247843
BICD1	-0.4351051	1.0190811
CD69	0.3802578	0.6807524
CD37	-0.3784251	1.4375666
CD9	0.3453646	0.1117994
CD226	0.3390867	0.1304901
CD82	0.3335424	0.5675827
CD302	-0.3262979	0.5838855

Ra+C- Population Overlaid on Primary UMAP

Characterizing R1 of the Ra+C- Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
HBD	Subunit of hemoglobin A2. Highly specific to erythroid cells.	Strong indicator of **Erythroid commitment**
ITGA2B	Megakaryocyte/Platelet integrin	Canonical marker of **Megakaryocytes/Platelets**
STXBP5	Exocytosis regulator in platelets & neutrophils	Indicating **Megakaryocytes/Platelet** lineage
RAB27B	GTPase implicated in vesicle trafficking, highly expressed in platelets.	Strong indicator of **Megakaryocytes/Platelet** lineage
SPINK2	Serine protease inhibitor. Expressed in some RBC precursors.	Upregulation in **erythroid** contexts.
ABCC4	Transporter for cAMP efflux in platelets.	Indicating **Platelet** lineage
GP1BB	Subunit of receptor complex on platelets.	Indicating **Platelet/Megakaryocytic** lineage
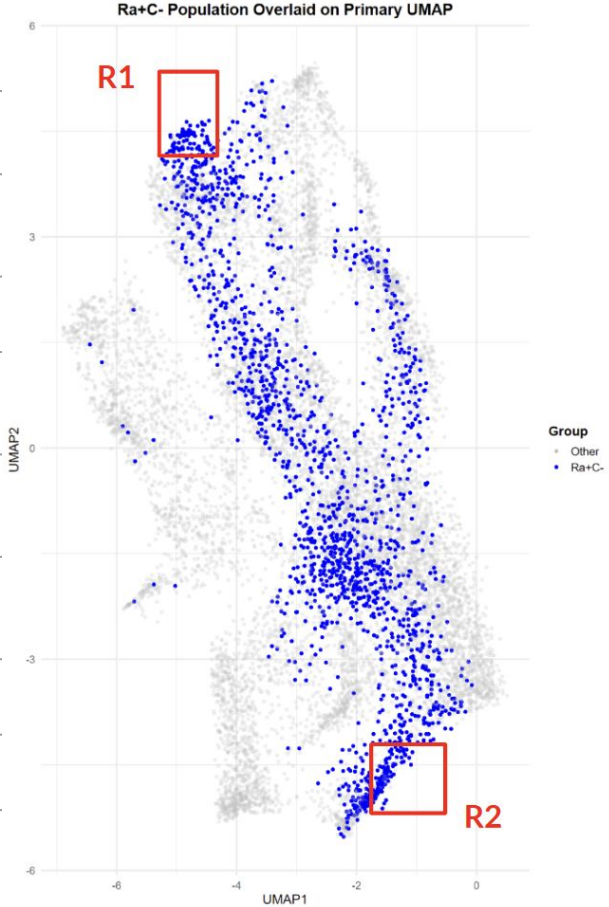
Ra+C- Population Overlaid on Primary UMAP

Characterizing R1 of the Ra+C- Population : Key Upregulated Genes and Their UniProt / Gene Context

Downregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
CD74	Components of BCR Complex.	Suppression of **B-Cell fate**.
CD99	Components of TCR complex, or T-cell surface activation.	Suppression of **T-Cell fate.**
CD48	SLAM family receptors on B/T/N cells.	Suppression of **lymphoid fate.**
CD44	Broad B/T/M marker involved in adhesion/homing.	Suppression of **lymphoid fate**.
CD302	C-type lectin, found in DCs and macrophages.	Suppression of **DC/Macrophage fate.**
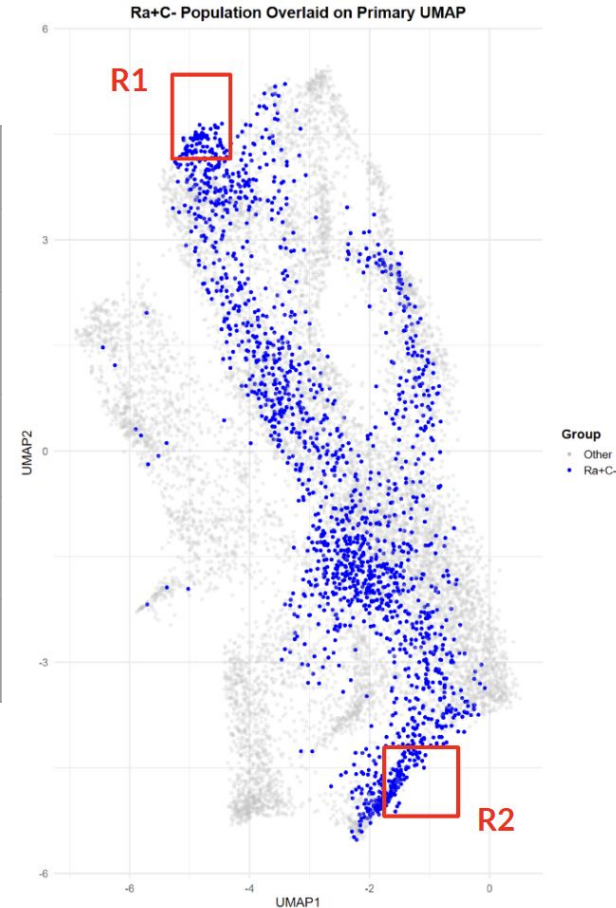
Ra+C- Population Overlaid on Primary UMAP

Characterizing R1 of the Ra+C- Population : Connection to FW/John/Yale's Papers + Characterization

In Line With Their Findings

> FW notes Ra+C- gating can have different fates ; some might be B-lymphoid biased, others RBC/platelet biased.

> John and Yale note that HPCs can spontaneously adopt RBC/Megakaryocyte identities when Notch or IL-7 Signals are insufficient.

Conclusion

1. Erythroid Platelet Signature
2. Exclusion of Lymphoid Fate
3. Given Notch or Hi IL-7 is not directing them to B/T lineages, this is not surprising

This region is characterized by erythroid/megakaryocytic fate.

Ra+C- Population Overlaid on Primary UMAP
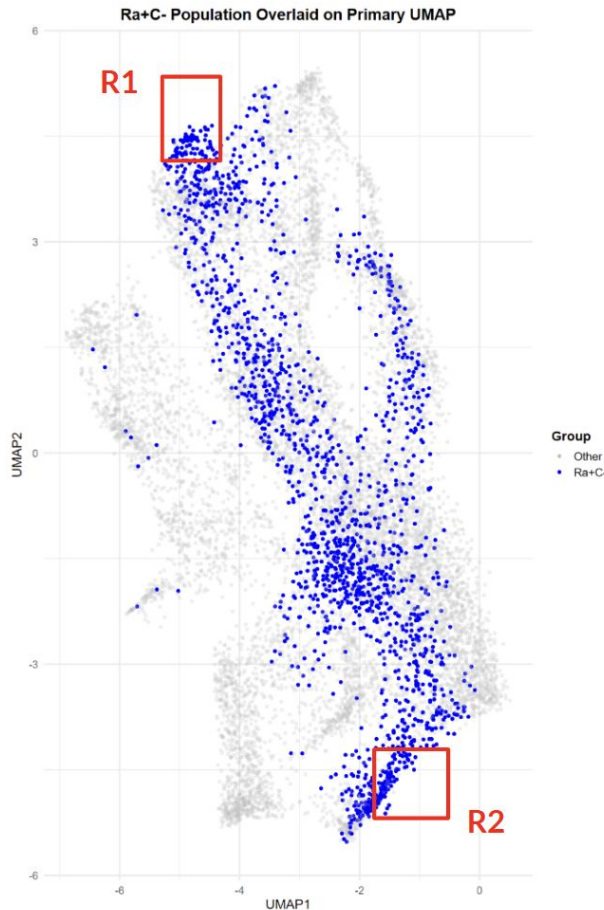
Characterizing R2 of the Ra+C- Population : Top Genes/CD Markers

Top 20 Genes

ID	logFC
<chr>	<dbl>
HIST1H4C	-1.2749375
HIST1H1B	-0.8467840
RBPMS	0.2835995
HDAC9	-0.8515274
HIST1H1C	-0.5566508
UBE2E2	-0.5896008
HIST1H1E	-0.4505406
RABGAP1L	-0.6140369
HIST1H1D	-0.6271537
HIST1H3D	-0.4628397
DIAPH3	-0.5253225
PTTG1	0.5344782
ALDH1A1	0.2691840
AFF3	-0.8891336
KIF18B	-0.3485664
CEP128	-0.4664184
FBXO5	-0.3487968
ADK	-0.4443190
RRM2	-0.4971882
DIAPH2	-0.5058354

Top 20 CD Markers

ID	logFC
<chr>	<dbl>
CD44	-0.28755418
CD34	0.25273801
CD59	0.17919633
CD99	0.23948973
CD52	0.27580557
CD38	-0.25596597
CD63	0.25708566
CDYL	-0.21229202
C2CD5	-0.13669178
CD81	0.16235847
CD84	-0.12110820
CDV3	0.15469354
CD79B	0.08226754
CD4	-0.11476129
BICD1	-0.13941746
CD58	0.09834087
CD99L2	-0.06261173
CD1D	0.09991974
TBCD	-0.07854332
CD248	0.00284589
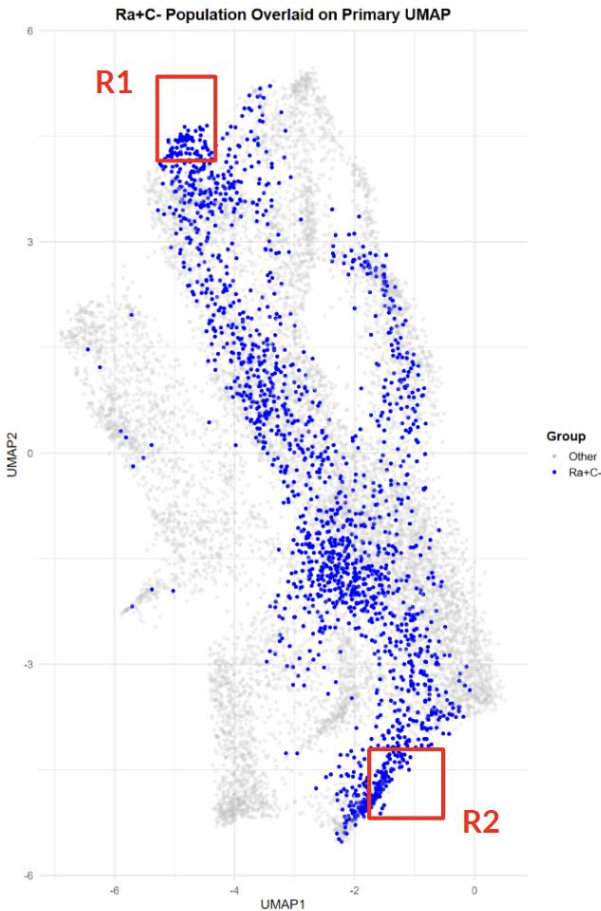
Ra+C- Population Overlaid on Primary UMAP

Characterizing R2 of the Ra+C- Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
TOP2A/NKAIN2	DNA Replication, chromosome segregation, ion transport.	Indicating **Actively Cycling** / Proliferative HSCs
HLA-DRA	MHC Class II Component	Indicator of **B-Cell** or **APC** Lineage
MSI2	Maintains HSC self-renewal	Early Progenitor / **Actively Cycling**
CD52	Widely expressed on M/B/T/NK ; Known target of alemtuzumab therapy	Indicator of of **B/T/NK** Lineage
CD74	Chaperone for MHC Class II Expression	Indicator of **B Cell** or **APC** Lineage
CD97/CD200	Adhesion GPCR / Immune Regulation Glycoprotein.	Indicator of **B/T** Lineage
CD48	SLAM family receptor	Indicator of **B/T/NK** Lineage
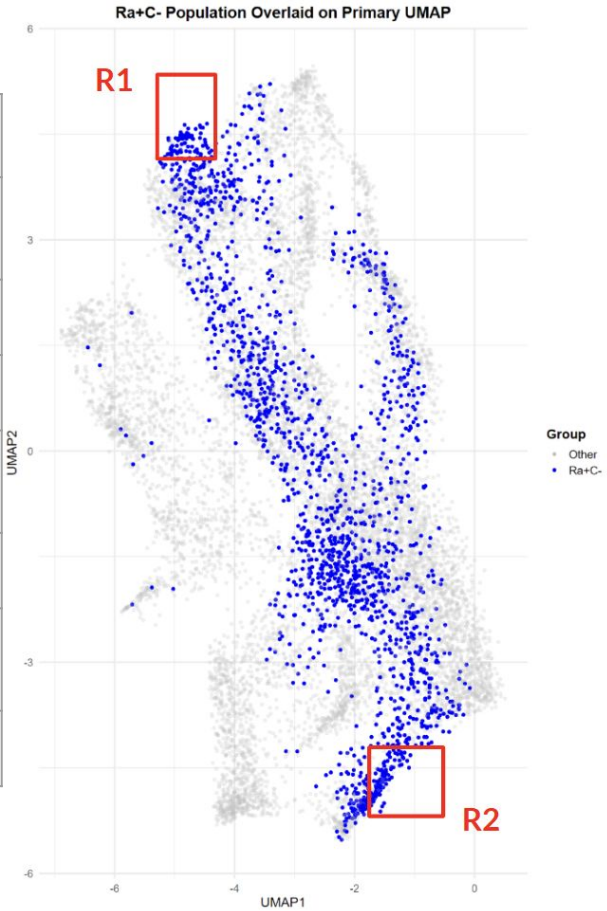
Ra+C- Population Overlaid on Primary UMAP

Characterizing R2 of the Ra+C- Population : Key Upregulated Genes and Their UniProt / Gene Context

Ra+C- Population Overlaid on Primary UMAP

Downregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
HIST1H4C, HMGB2, UBE2C, TUBB4B, UBE25, CENPF	A suite of cell cycle / chromatin regulators, typical in rapidly dividing cells.	Beginning to exit from the cell-cycle, but still in the cycling phase (**M-phase**).
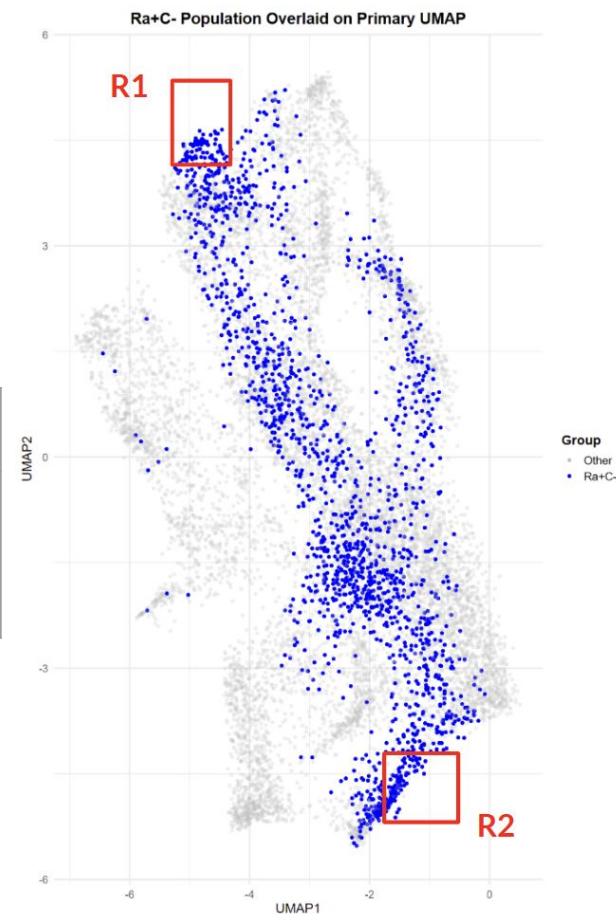
Characterizing R2 of the Ra+C- Population : Connection to FW/John/Yale's Papers + Characterization

In Line With Their Findings

> FW notes Ra+C- gating can have different fates ; some might be B-lymphoid biased, others RBC/platelet biased.

> Yale notes that B-lineage biased HPCs show robust MHC II expression and HPC markers once Erythroid/Myeloid genes are downregulated.

Conclusion

1. B-Lineage Oriented Progenitor
2. Actively Cycling, but beginning exit from active proliferation
3. Given Moderate IL-7 in 10% StemSpan, direction to B but not T lineages makes sense.

This region is characterized by actively cycling, B-cell fate biased cells.

Ra+C- Population Overlaid on Primary UMAP
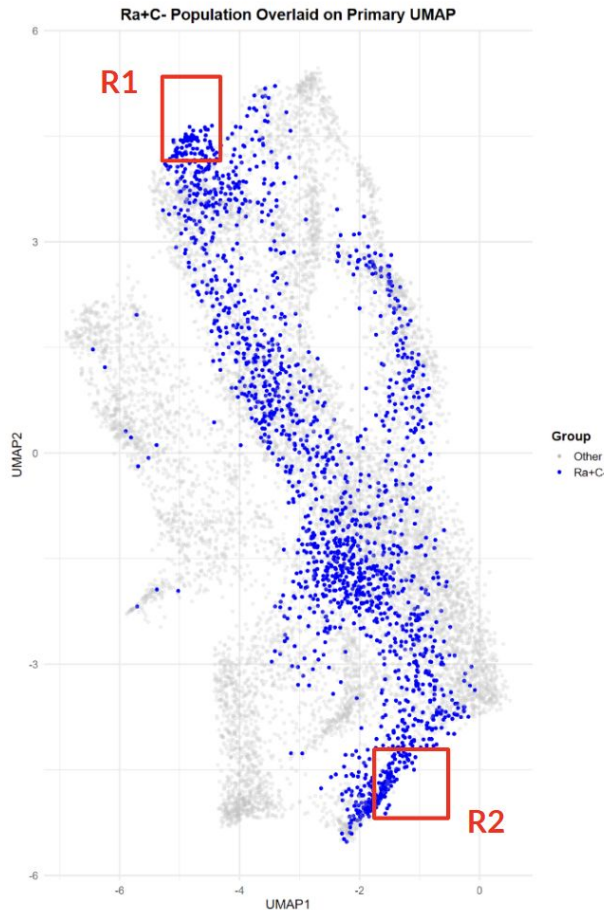
Slide Left Blank Intentionally

Characterizing R1 of the C+ Population : Top Genes/CD Markers

Top 20 Genes

| ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
TOP2A	1.943276	1.3769076
HIST1H4C	1.925317	2.3573200
MKI67	1.817081	1.2989187
UBE2C	1.698571	0.9020936
HMGB2	1.661068	2.7763606
CENPF	1.462420	1.3749132
ASPM	1.425954	1.0068074
NUSAP1	1.408244	1.0801089
UBE2S	1.311736	1.3815108
HIST1H1B	1.299407	1.0095044
CENPE	1.273867	0.8845728
HMMR	1.245465	0.7621866
TUBB4B	1.237407	1.5383101
CDK1	1.234474	0.7712793
KPNA2	1.178221	1.0672469
TUBB	1.146956	2.8305532
TPX2	1.140368	0.9513845
KIF11	1.110784	0.7894748
SMC4	1.108021	1.6302565
SAMHD1	1.067395	2.0192066

Top 20 CD Markers

ID	logFC
<chr>	<dbl>
CD34	-0.3903839
CD74	0.3669891
CD52	-0.2986400
CDYL	0.2070244
CD36	0.2044258
CD1D	0.1933482
CD2AP	0.1831994
CD37	-0.1793484
CD84	0.1769242
CD1C	0.1695980
CD53	0.1695372
SCD	0.1679191
CD96	-0.1498861
CD226	0.1455231
CD86	0.1390596
CD180	0.1379317
HACD4	0.1278411
CD9	0.1240892
CDV3	-0.1222131
CD79B	-0.1168537
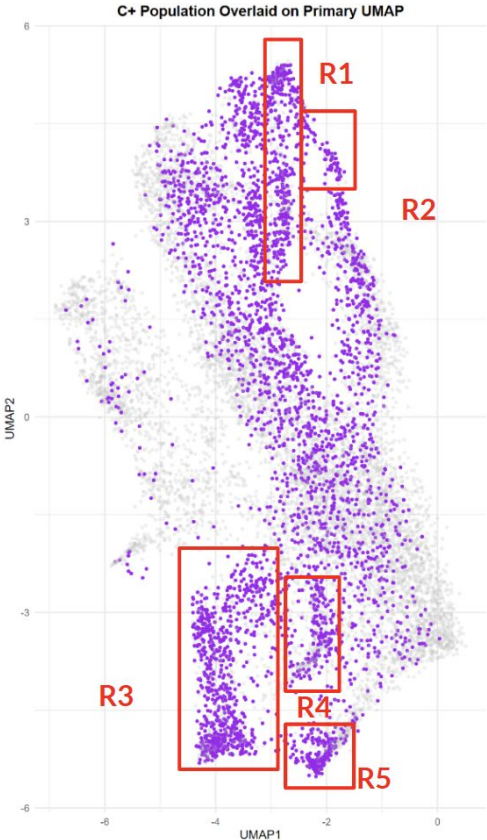
C+ Population Overlaid on Primary UMAP

Characterizing R1 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
TOP2A/ HIST1HXX/ CENPF/ HMGB2	DNA replication ; chromatin structure ; centromeres	Actively **Cell-Cycling**
CDK1	Master regulator of G2 -> M	Transition from **G2 to M Phase**
TUBB/ASPM/ TPX2	Microtubule, Spindle Assembly	Active **Cell Cycling**
CD16	Classical Myeloid marker	Strong indicator of **myeloid lineage**.
CD1D	MHC Class I molecule	Indicating **Monocyte/DC Lineage**
CD180	LPS Receptor cofactor	Indicating **B/M/DC Lineage**
CD74	MHC Class II Invariant Chain	Indicating **Antigen Presenting Capability (M/DC)**
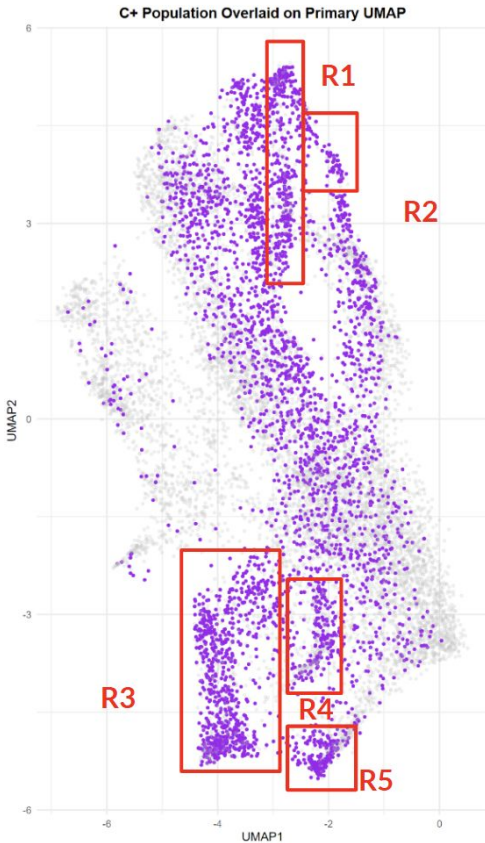
C+ Population Overlaid on Primary UMAP

Characterizing R1 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context
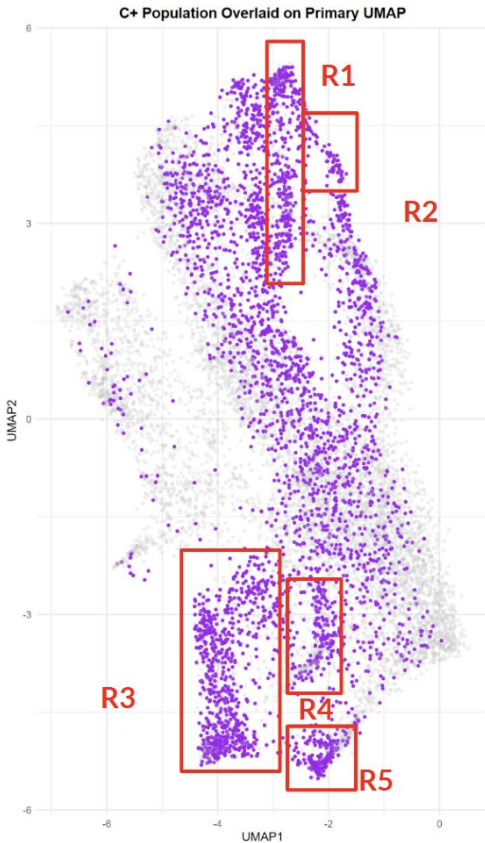
Downregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
CD34	Canonical HSC marker	**Losing stem cell state**, committing to lineage
CD52	GPI Anchored Glycoprotein expressed on lymphocytes and HSCs	Suppressing **Lymphoid and HSC lineage**
CD37	Tetraspanin on B/T/HSCs	Suppressing **Lymphoid lineage**
CD79B	BCR Component	Suppressing **B cell lineage**

C+ Population Overlaid on Primary UMAP

Characterizing R1 of the C+ Population : Connection to FW/John/Yale's Papers + Characterization

In Line With Their Findings
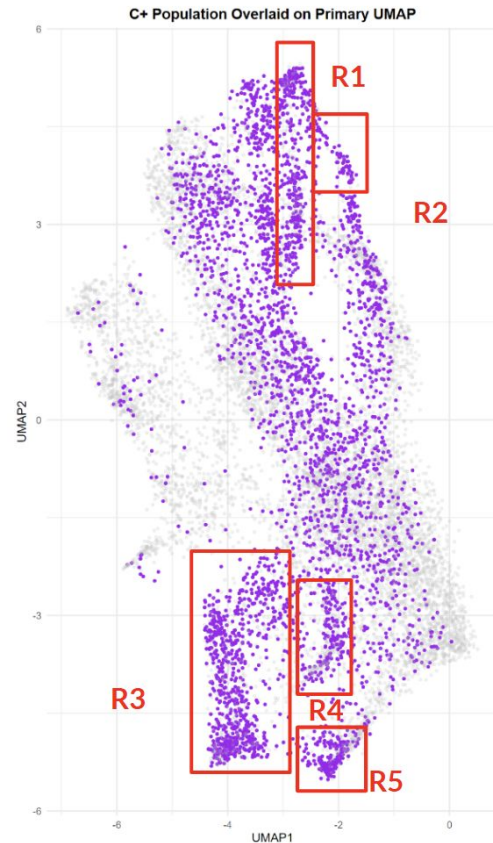
> Aligns with FW's C+ gating typically capturing NM progenitors.
> Aligns with John's Thesis, lack of Notch promoting myeloid fate (evidenced by CD16)

Conclusion

1. Myeloid/Monocyte-Neutrophil Identity
2. Highly Proliferative
3. Downregulation of HSC markers

*This region is characterized by mitosis undergoing, myeloid biased cells.*

C+ Population Overlaid on Primary UMAP

Characterizing R2 of the C+ Population : Top Genes/CD Markers

Top 20 Genes

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
HIST1H4C	-1.2749375	2.3573200
AFF3	-0.8891336	2.3235517
HDAC9	-0.8515274	2.6566573
HIST1H1B	-0.8467840	1.0095044
MPO	-0.7439102	1.9990494
HIST1H1D	-0.6271537	1.2377228
RABGAP1L	-0.6140369	1.8144553
RUNX2	-0.5982653	1.9477654
BCL2	-0.5901856	1.9437992
UBE2E2	-0.5896008	1.7408516
WWOX	-0.5687601	1.8820647
AUTS2	-0.5674274	1.8597347
HIST1H1C	-0.5566508	0.8933185
PRDX1	0.5466051	2.4794835
SFMBT2	-0.5393858	1.7277444
PTTG1	0.5344782	1.1813705
DIAPH3	-0.5253225	1.0074927
FCHSD2	-0.5165822	2.3827329
DIAPH2	-0.5058354	1.7573645
LINC01572	-0.5037277	1.1909193

Top 20 CD Markers

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
CD44	-0.28755418	1.66439588
CD52	0.27580557	0.73682286
CD63	0.25708566	1.89852791
CD38	-0.25596597	1.04718520
CD34	0.25273801	0.53207044
CD99	0.23948973	1.17319929
CDYL	-0.21229202	1.07175479
CD59	0.17919633	0.45740367
CD81	0.16235847	1.13598289
CDV3	0.15469354	1.35935331
BICD1	-0.13941746	0.94689170
C2CD5	-0.13669178	0.55794857
CD84	-0.12110820	0.39274125
CD4	-0.11476129	0.28359894
CD1D	0.09991974	0.35725796
CD58	0.09834087	0.46969995
CD79B	0.08226754	0.21083310
TBCD	-0.07854032	0.45369966
CD99L2	-0.06261173	0.21312940
CD200R1	-0.04557340	0.09994318
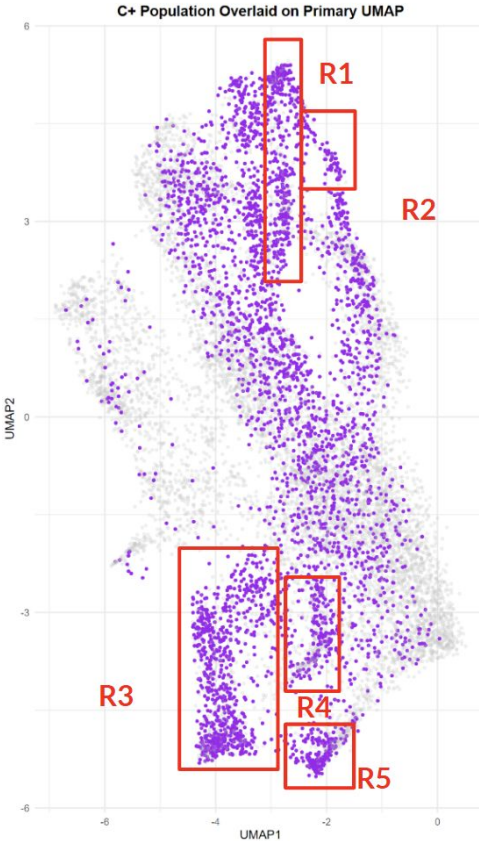
C+ Population Overlaid on Primary UMAP

Characterizing R2 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
PRDX1	Antioxidant enzyme highly expressed on HSCs	Cells coping with **ROS** ; **metabolic transition**
SFMBT2, PTTG1, DIAPH3, FCHSD2	Transcriptional repressors, Mitotic regulators	Mild proliferative or structural remodelling programs; **G0 state**
CD63	Tetraspanin involved in lysosomal trafficking and degranulation. Expressed in M/NK/DCs.	Early **myeloid priming** , **APC** vesicle secretion involvement
CD52	Surface Glycoprotein on B/T/M	Hematopoietic, but **uncommitted lineage**
CD34	Canonical HSC marker	Retaining some **stem cell features**
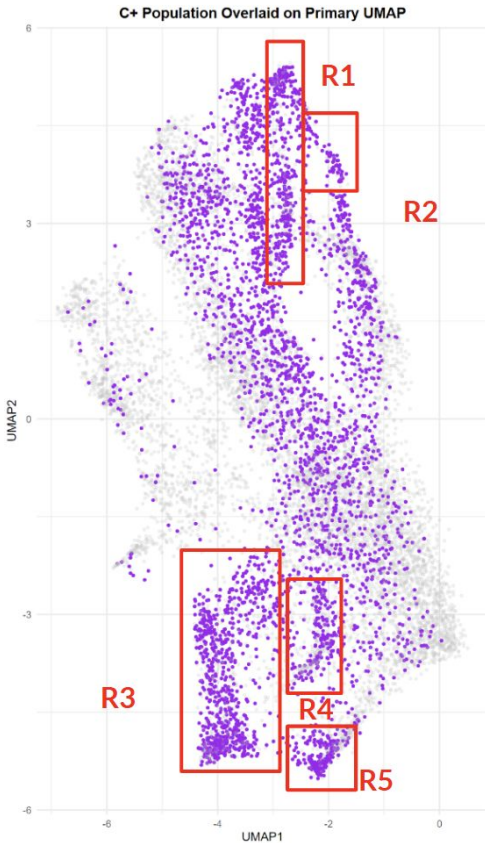
C+ Population Overlaid on Primary UMAP

Characterizing R2 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context

Downregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
HIST1H14C/ HIst1H1B/ HIST1H1D etc.	Core and linker histones; essential for chromatin compaction and proliferation	**Reduced cell cycle activity** ; Cells are exiting active proliferative state
HDAC9	Epigenetic regulator of stemness	Epigenetic relaxation; **Lineage Restriction**
MKI67	Classific proliferation marker	Reduced cell cycle activity; **quiescent state**
MPO	Enzyme in neutrophils and monocytes, involved in ROS production	Population not yet myeloid mature, or **diverging from granulocyte lineage**.
CD4 / CD79B	Classical TCR / B cell antigen receptor	Divergence from **Lymphoid lineage**
CD38	Found on activated B/T/NK cells	Divergence from **Lymphoid Lineage**
CD44	Adhesion molecule on HSCs	**Maturation** ; Divergence from migratory state
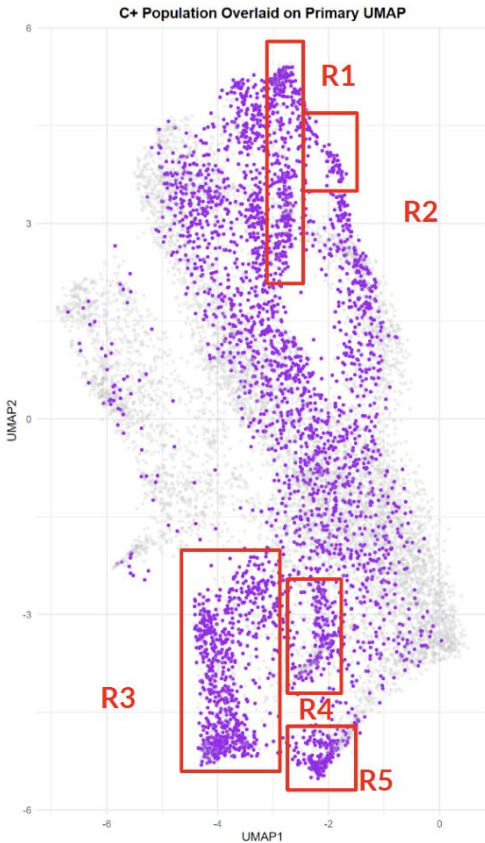
C+ Population Overlaid on Primary UMAP

Characterizing R2 of the C+ Population : Connection to FW/John/Yale's Papers + Characterization

In Line With Their Findings

> Aligns with FW's C+ gating typically capturing NM progenitors.
> Aligns with John's Thesis, loss of CD34/cell cycle genes accompanies NM differentiation
> Yale indicatest that MHCII and CD63 can increase in early monocyte/DC commitment

Conclusion

1. Reduced Cell Cycling
2. Retention of Stem Cell Markers
3. Monocyte/DC Primed path
4. Active suppression of Lymphoid/Adhesive programs

This region is characterized by quiescent, lineage primed NM progenitor.

C+ Population Overlaid on Primary UMAP
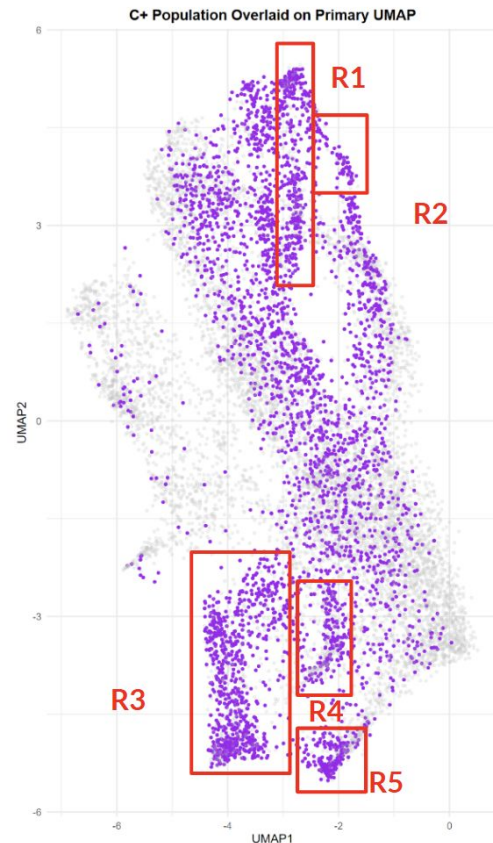
Characterizing R3 of the C+ Population : Top Genes/CD Markers

Top 20 Genes

| ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
CST3	2.844294	2.1637611
S100A6	2.082420	1.6373426
S100A10	2.045921	1.0808110
LYZ	2.007371	1.5582085
SAMHD1	1.901196	2.0192066
S100A4	1.889538	3.3488066
HLA-DRA	1.708710	3.2477191
VIM	1.691064	3.3620246
HLA-DPA1	1.674286	2.1864050
HLA-DPB1	1.604700	1.9964030
SLC8A1	1.603701	0.8390865
MPO	-1.565697	1.9990494
COTL1	1.554372	1.9108972
HLA-DRB1	1.507356	2.4250991
S100A11	1.502511	1.8005538
ANXA2	1.477705	1.0167411
CD74	1.467318	3.2601540
PLXDC2	1.466705	0.9928553
RTN1	1.419965	0.4684784
KCNQ5	-1.417320	2.3019023

Top 20 CD Markers

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
CD74	1.4673176	3.26015404
CD1C	1.1770754	0.47744823
CD36	0.9897158	0.56109847
CD48	0.8113073	1.29148444
CD86	0.7988978	0.41886951
CD1D	0.6891653	0.35725796
CD37	0.5373237	1.44754954
CD34	-0.5176654	0.53207044
CD4	0.5098941	0.28359894
CD302	0.4410703	1.10428854
CD63	0.3935811	1.89852791
CD38	-0.3743976	1.04718520
CD53	0.3648039	1.29404331
HACD3	-0.3519215	0.92054313
CD96	-0.2985389	0.31276717
HACD4	0.2869140	0.64434172
CD1E	0.2804092	0.10755063
CD200R1	0.2585260	0.09994318
CD68	0.2573989	0.18316388
CD99	-0.2208288	1.17319929
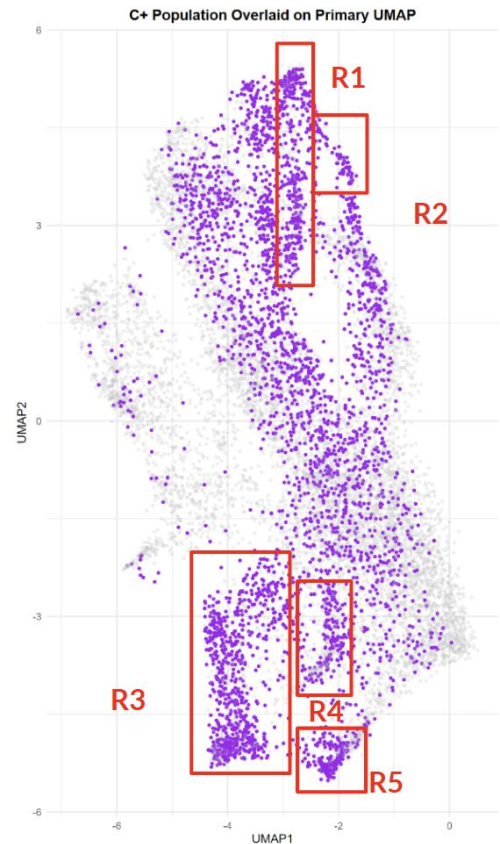
C+ Population Overlaid on Primary UMAP

Characterizing R3 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated Genes

Gene	Role / Pathway	Interpretation
CST3	Cysteine protease inhibitor, highly expressed in Mono, macro, DCs	Strong **monocyte / DC lineage** indicator
S100A6/A10/A4/A11	Calcium binding proteins; involved in inflammation, migration, myeloid activation	Commonly upregulated in **Mono,macro and DCs** - indicating **active myeloid lineage**
LYZ	Lysozyme; Key innate immunity enzyme	Classic **monocyte lineage** gene
HLA-DRA/DPA1/DPB1/DRB1	Antigen presentation MHC Class II components	Highly expressed in **mono,macro, DC** lineage, strongly indicating **APC lineage**
SAMHD1	dNTP pools and Viral restriction; expressed in myeloid cells, especially DCs	Indicator of **DC lineage**
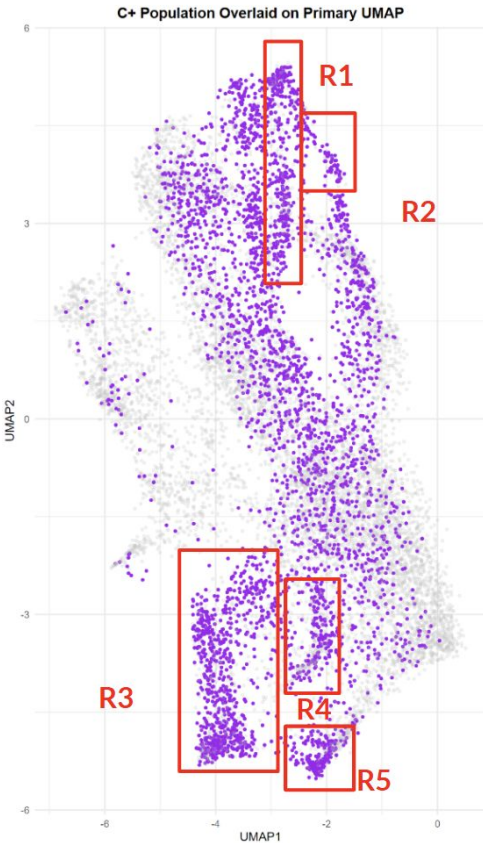
C+ Population Overlaid on Primary UMAP

Characterizing R3 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated CD Markers

| Gene	Role / Pathway	Interpretation
CD74	Invariant chain of MHC Class II complex	Critical **APC lineage** indicator ; aligns with high HLA Class II expression
CD1C	MHC-I like molecule; unique to DCs and some mono	Classic Indicator of **DCs**
CD36	Scavenger receptor for lipids ; high in mono/macro	Strong indicator of **phagocytic myeloid** function
CD86	Co-stimulatory molecule on APCs (mono, DCs)	Indicator of functional **APC identity** (DC, Mono)
CD1D	Lipid presenting molecule ; found on mono, B, DCs	Indicator of **APC Lineage** (DC, Mono)
CD48	SLAM family protein on Lymphoid/Myeloid cells	Indicator of **Early Progenitor / APCs**
CD302	C-type lectin receptor on Dendritic Cells	Indicator of **DC Lineage**
CD4	Modest expression ; In line with monocyte/DC fate	Supports **Monocyte/DC lineage**
CD63/53	Tetraspanins Common on Mono/Granulo/APC fate	Indicator of **Monocyte/Granulocyte/APC** Lineage
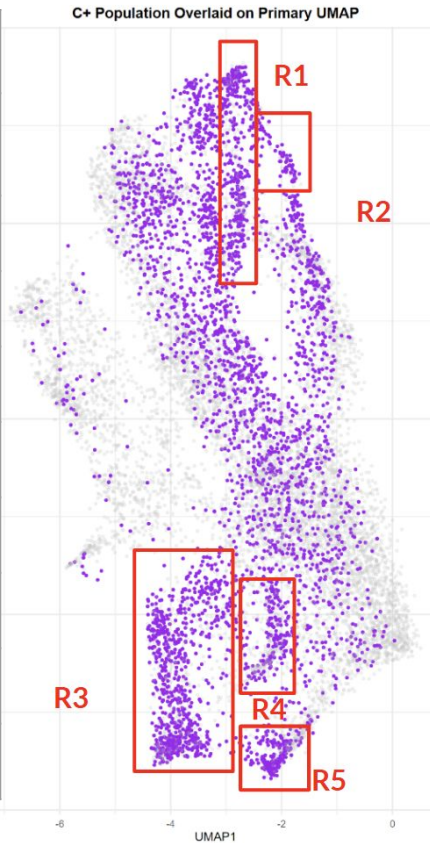
C+ Population Overlaid on Primary UMAP

Characterizing R3 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context

Downregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
CD34	Canonical Marker of Stem Cells	Diverging from Stem Cell Ness
CD96	Natural Killer associated programs	Divergence from **NK Fate**
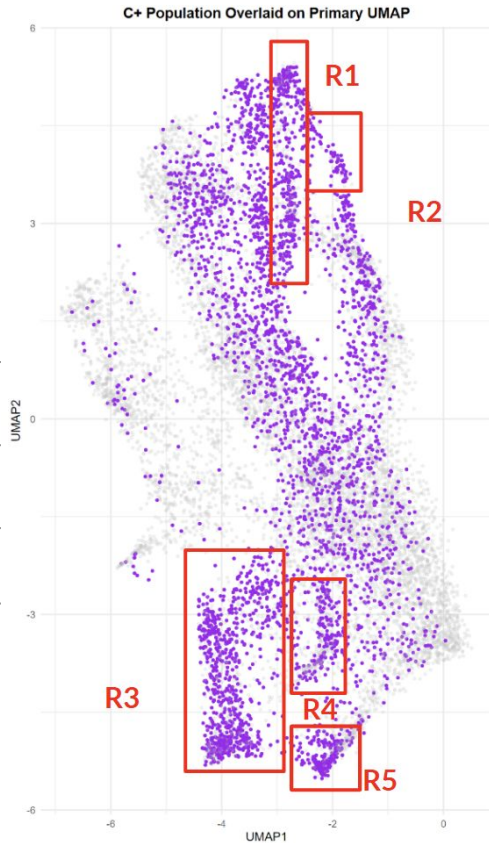
C+ Population Overlaid on Primary UMAP

Characterizing R3 of the C+ Population : Connection to FW/John/Yale's Papers + Characterization

In Line With Their Findings

> Aligns with FW's C+ gating typically capturing NM progenitors. In this case, R3 is strongly Monocyte oriented.
> John's Thesis notes Monocyte/DCs can emerge spontaneously under specific cytokines (SCF, Flt3L)

Alignment with Yale's Paper

Feature	FW C+ Monocyte/DC Population	Yale PSC-HPC Myeloid Output
High MHC Class II (HLA-DRA, -DPA1, -DPB1, -DRB1)	✔ Strong upregulation	✔ Highly expressed in DC-like cells
CD74 (MHC-II invariant chain)	✔ Strongly upregulated	✔ Marker of monocyte/DC identity
CD1C+ and CD86+ expression	✔ Both are upregulated in your C+ subset	✔ Used by Yale to define monocyte-derived DCs
Monocyte/innate immune gene signature	✔ CST3, LYZ, S100A4/6/10, ANXA2 all up	✔ Similar markers observed in PSC-derived CD14+/CD1C+ cells
Low CD34/CD38	✔ Downregulated — shows lineage commitment	✔ MHC II+ cells in Yale also downregulated CD34, indicating maturity
No Notch signals / low T/B output	✔ IL-7 present, but no DLL1/DLL4	✔ No Notch activation unless enforced experimentally
Culture-dependent bias toward myeloid/DC fates	✔ Observed even before clonal assays	✔ Intrinsic bias of PSC-HPCs under similar cytokines
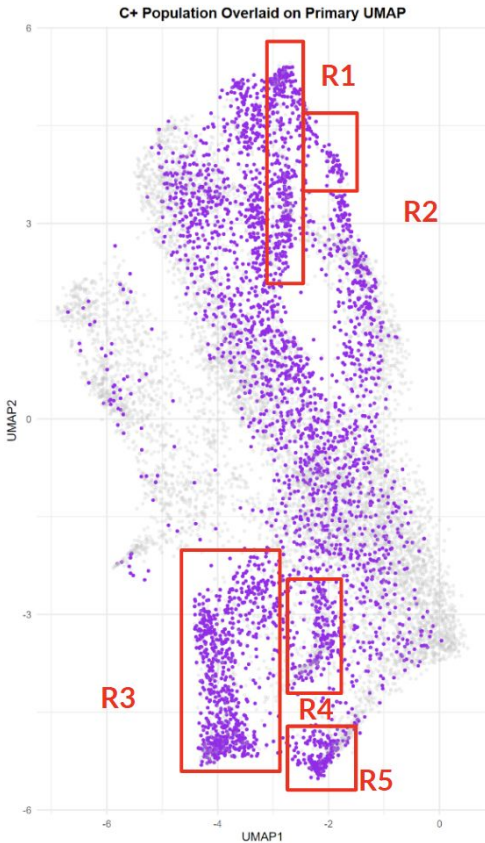
C+ Population Overlaid on Primary UMAP

Characterizing R3 of the C+ Population : Connection to FW/John/Yale's Papers + Characterization

Conclusion

R3 from FW's data shows good overlap with MHC-II[+] CD1C[+] Monocyte/DC-like cells described in Yale's PSC paper:
- Both populations emerge spontaneously under cytokine rich, Notch deficient environments
- Shared phenotype: MHCII^{hi}, CD74^{hi}, CD1C^{hi}, CD36^{hi}, CST3^{+}, LYZ^{+}
- Both lack CD34/CD38

Suggests that even through cellular origin is different (cord blood HSCs vs. PSC-derived HSCs), the environmental conditions in both studies converge on a myeloid antigen-presenting phenotype, with strong inclination of them being monocyte/dendritic cells.

Yale interpreted this as the default myeloid fate in the absence of strong Notch of lymphoid induced signals. R3's profile supports this model, suggesting that even fate-biased (via CLEC12 acquisition), freshly isolated cord blood progenitors will produce similiar monocyte/DC-like phenotypes when cultured in lymphoid cytokines without Notch activation.

R3 is characterized by a strong Monocyte/DC like population, with high antigen presentation, loss of stem cell markers, and consistent with FW's C+ compartment, but skewed towards Monocyte/DC axis instead of neutrophilic granulocyte fates.


C+ Population Overlaid on Primary UMAP
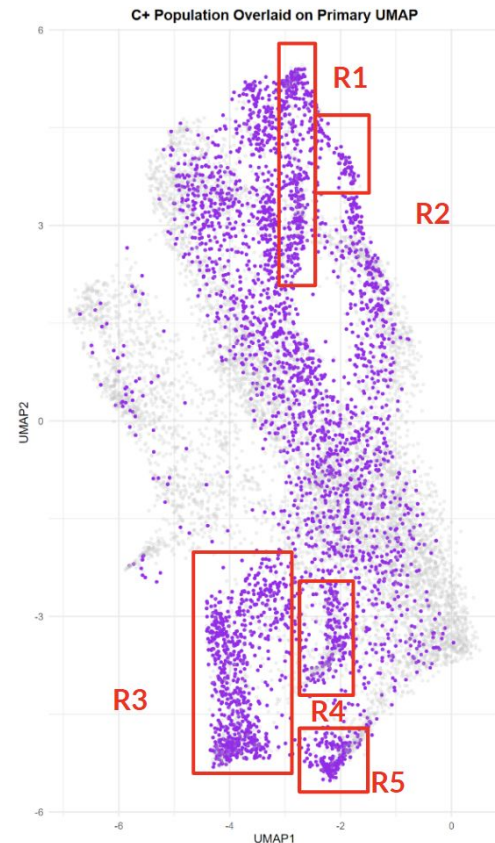
Characterizing R4 of the C+ Population : Top Genes/CD Markers

Top 20 Genes

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
MPO	2.8475984	1.9990494
PRTN3	2.3095339	0.4257153
AZU1	1.7662224	0.2779040
LRMDA	1.5977458	2.2763890
FNDC3B	1.3983028	1.2114140
SERPINB1	1.3156648	1.9682292
CST3	-1.2797978	2.1637611
HIST1H4C	-1.2525414	2.3573200
LYST	1.1831815	1.1446418
PRSS57	1.1804039	1.0461957
TUBB4B	-1.1790977	1.5383101
MKI67	-1.1584932	1.2989187
TOP2A	-1.1223199	1.3769076
TUBA1B	-1.0974248	3.9090800
S100A4	-1.0880403	3.3488066
HMGB2	-1.0364533	2.7763606
SRGN	0.9886033	1.9498024
HIST1H1B	-0.9691742	1.0095044
S100A6	-0.9655470	1.6373426
S100A10	-0.9496436	1.0808110

Top 20 CD Markers

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
CD74	-0.7150619	3.2601540
CD52	-0.4388130	0.7368229
CD164	-0.4225014	1.4962513
CD1C	-0.4094522	0.4774482
CD36	-0.4036869	0.5610985
CD53	-0.3805762	1.2940433
CD44	0.3305364	1.6643959
CD84	0.3025318	0.3927413
CD38	0.3009633	1.0471852
CD1D	-0.2890853	0.3572580
CD96	0.2637392	0.3127672
CD81	-0.2430534	1.1359829
HACD3	-0.2283692	0.9205431
CD63	-0.2279814	1.8985279
CD86	-0.2250331	0.4188695
CD48	-0.2156195	1.2914844
CD302	0.1958048	1.1042885
CD300A	-0.1778407	0.3300785
CD47	-0.1623991	1.1277816
CD34	-0.1471191	0.5320704
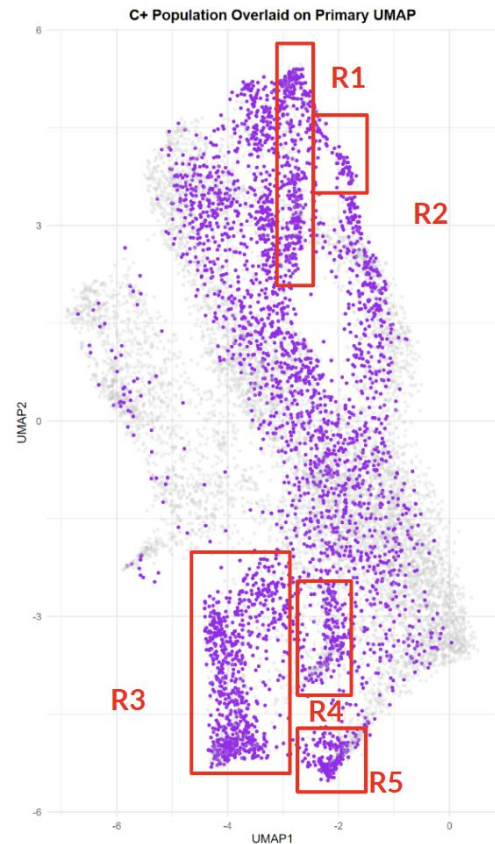
C+ Population Overlaid on Primary UMAP

Characterizing R4 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
MPO	Granule enzyme responsible for ROS generation	Indicating strong Neutrophil lineage.
PRTN3	Serine protease in azurophilic granules	Indicating Neutrophil/Granulocytic lineage.
AZU1	Antimicrobial protein in neutrophil granules	Indicating strong Neutrophil lineage.
SERPINB1	Serine protease inhibitor; protects Neutrophils from self inflicted damage	Indicating strong Neutrophil lineage.
LYST	Involved in granule formation and secretion	Indicating strong Granulocytic lineage.
CD44/84	Cell adhesion and migration / SLAM family receptor	Enhanced adhesive properties (N fate)
SRGN	Scaffold for granule proteins and secretion pathway	Indicating Neutrophil/Mast cells
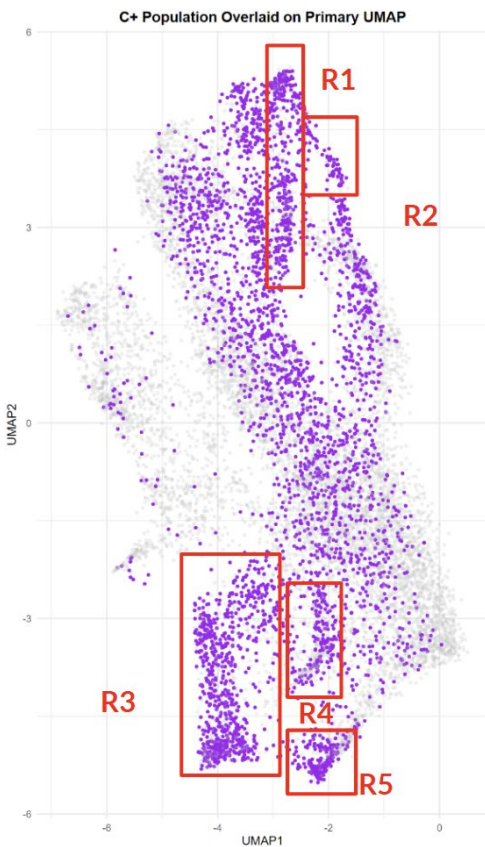
C+ Population Overlaid on Primary UMAP

Characterizing R4 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context

Downregulated Genes / CD Markers

| Gene	Role / Pathway	Interpretation
CD74	Invariant chain of MHC Class II complex	Divergence from APC (M/DC) fate
CD52	Typically expressed on B/T	Divergence from lymphoid fate
CD1C / 1D	Classical DC Cell Markers	Divergence from DC fate
CD36	Scavenger receptor on Mono/Macro	Divergence from Monocytes/Macrophages
CD86	Co-stimulatory molecule for antigen presentation	Divergence from DC/B-Cell fate
HIST1Hxx	Chromatin Structure and Cell Proliferation	Exiting Proliferative State
CST3	Cysteine Protease Inhibitor	Shift in protease regulation / differentiation
MKI67	Classic Marker of proliferation	Exiting Proliferative State
TOP2A	Classic Marker of proliferation	Exiting Proliferative State
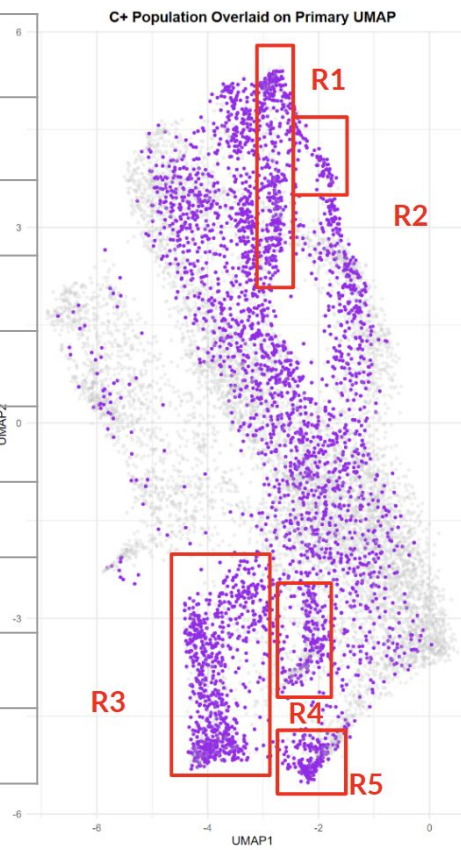
C+ Population Overlaid on Primary UMAP

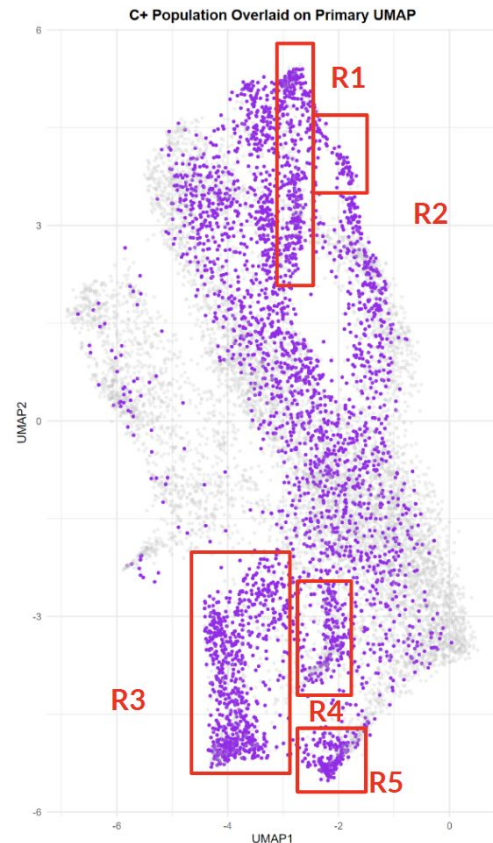# Characterizing R4 of the C+ Population : Connection to FW/John/Yale's Papers + Characterization

In Line With Their Findings

> Aligns with FW's C+ gating typically capturing NM progenitors.
> Aligns with John's Thesis, loss of CD34/cell cycle genes accompanies NM differentiation

Conclusion

1. Key upregulated genes are specific to neutrophil extracellular trap formation and phagosome maturation
2. Downregulation of proliferation markers
3. Downregulation of progenitor, lymphoid and DC programs, upregulation of adhesion and migration programs

*This region is characterized by a near terminally differentiated, lineage primed Neutrophil progenitor.*

C+ Population Overlaid on Primary UMAP

Characterizing R5 of the C+ Population : Top Genes/CD Markers

Top 20 Genes

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
LTB	3.106609	1.2740898
TUBA1B	-2.921061	3.9090800
TCF4	2.816225	1.7692575
JCHAIN	2.804626	0.4055437
H2AFZ	-2.686776	3.7948555
LINC01478	2.435756	0.2620662
HMGB2	-2.414741	2.7763606
CCDC50	2.359768	1.1358201
CUX2	2.315946	0.4353410
UGCG	2.248940	0.7868915
AC023590.1	2.203652	0.4591193
NIBAN3	2.148335	0.4199959
LINC01374	2.103218	0.3242030
RABGAP1L	2.059722	1.8144553
TPM2	2.043972	0.2870637
KCNQ5	-2.031652	2.3019023
CARD11	1.994982	0.6294967
IRF8	1.985258	1.2900046
FCHSD2	1.973939	2.3827329
LRMDA	-1.914075	2.2763890

Top 20 CD Markers

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
CD74	1.4602567	3.2601540
CD2AP	1.0363949	1.0172753
CD164	0.9516593	1.4962513
CD37	0.8932082	1.4475495
CD63	-0.8293675	1.8985279
CDYL	0.8201088	1.0717548
CD53	0.7633746	1.2940433
HACD3	-0.6845285	0.9205431
CD81	-0.5988106	1.1359829
CD38	-0.5814162	1.0471852
CDT1	-0.5783227	0.6152767
CD34	-0.5545999	0.5320704
CD4	0.4853419	0.2835989
SCD	-0.4639307	0.6539370
CD36	0.4080784	0.5610985
CD33	-0.4029567	0.5386984
CD48	-0.3963579	1.2914844
HACD1	-0.3655580	0.3836370
CD59	-0.3523438	0.4574037
CD180	0.3519795	0.3461945
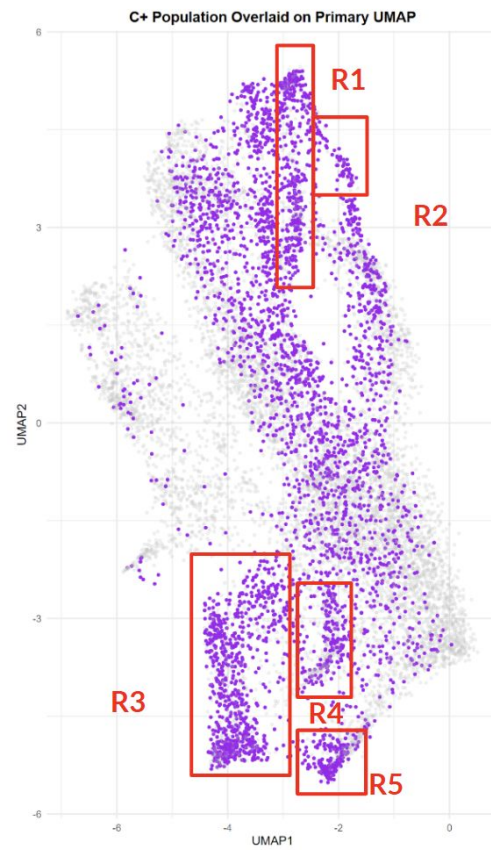
C+ Population Overlaid on Primary UMAP

Characterizing R4 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
LTB	Lymphotoxin Beta ; Lymphoid organogenesis	B/T/NK Lineage ; Heightened Inflammatory Signals
TCF4	Lymphoid Development ; GRN Activation	B Cells (Plasma) , T Cells
JCHAIN	Joining chain of multimeric immunoglobulins	B Plasma Cells ; Immunoglobulin Production
IRF8	TF critical for myeloid/DC/B decisions	DC / Macro / B Cells
CD74	Invariant chain of MHC II Antigen presentation	APCs - B cells / DCs / Monocytes
CD2AP	Adapter Protein in Cytoskeletal Reorganization	T/NK - Adhesive ; Enhanced Signalling Capacity
CD4	Co-receptor for MHC Class II Complex	Indicating a T/DC Lineage
CD37	Tetraspanin involved in cell signalling	B/T/NK/DCs ; Immune Signalling ; Maturation
TPM2/ RABGAP1L/ LINC01374	Cellular Programs for Morphological / Structural remodelling , lineage commitment	Population is becoming terminally committed
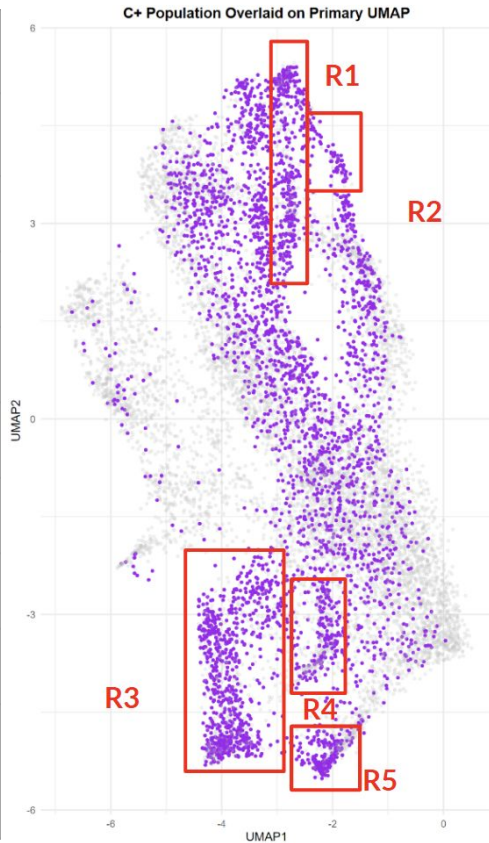
C+ Population Overlaid on Primary UMAP

Characterizing R4 of the C+ Population : Key Upregulated Genes and Their UniProt / Gene Context

Downregulated Genes / CD Markers

| Gene	Role / Pathway	Interpretation
H2AFZ/ HMGB2/ MKI67/ TOP2A	Ubiquitous in Actively Proliferating Cells	Cells are exiting proliferation, advanced differentiation stage
Broad Genes	Number of Housekeeping Tasks	Reducing metabolic flux to differentiate
CD63	Tetraspanin in vesicle/exosome forming	Suppressing platelet / monocyte / granulocyte fate
CD33	Classical myeloid lineage marker	Divergence from Mono/Macro lineages
CD48	Adhesion/ Costimulatory molecule	Suppression of B/T/NK/DC Lineages
CD34	Canonical HSC marker	Divergence from Stem Cell state
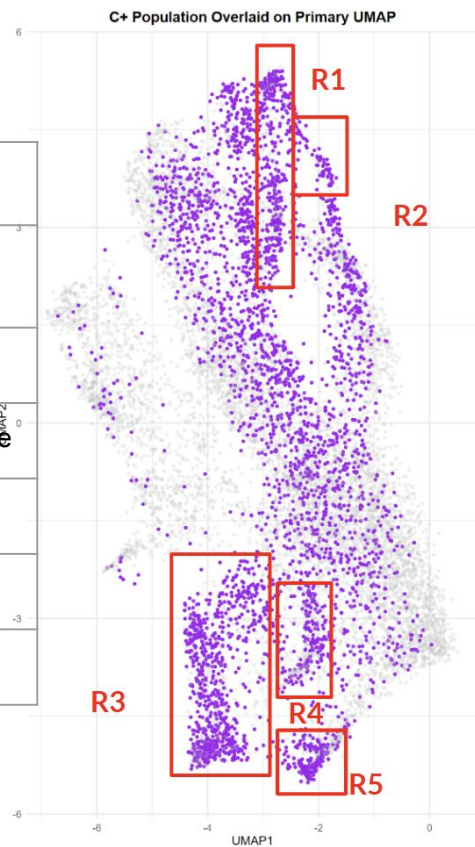
C+ Population Overlaid on Primary UMAP

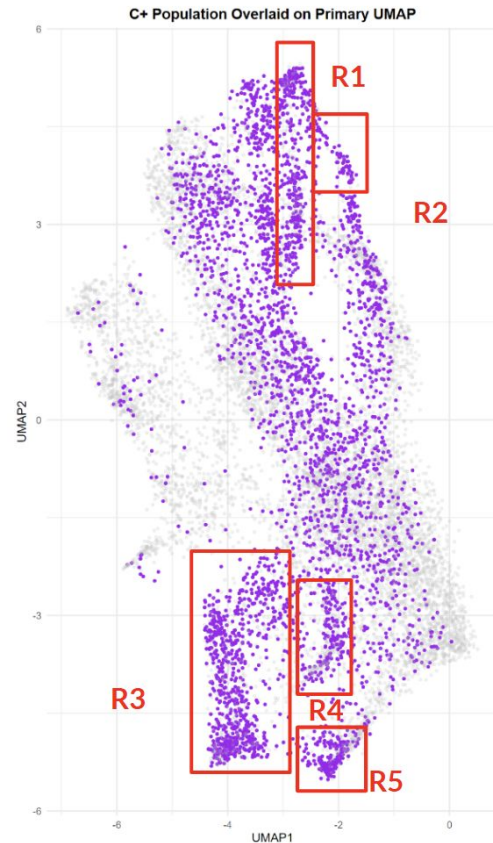# Characterizing R2 of the C+ Population : Connection to FW/John/Yale's Papers + Characterization

In Line With Their Findings

> Aligns with FW's C+ gating typically capturing NM progenitors, and downregulation of proliferative markers indicating shift toward specialized immune cell identity. Also mirror's John's thesis (decrease in licensing genes = cessation of proliferation / differentiation time)
> Yale mentions gene expression "bottleneck" when cells begin reprogramming and committing to final fate - evidenced here by broad downregulation of all sorts of genes.
> Similar upregulation of TCF4, IRF8 and downregulation of histones/cycling genes, as w/ Yale

Conclusion

R5 nearing terminal differentiation, but less mature than R3 and R4. Likely committing to neutrophil/granulocyte pathway, while turning off progenitor, lymphoid and proliferation programs.

This region is characterized by a actively differentiating, Neutrophil (R4) or Monoocyte/DC candidates (R3).
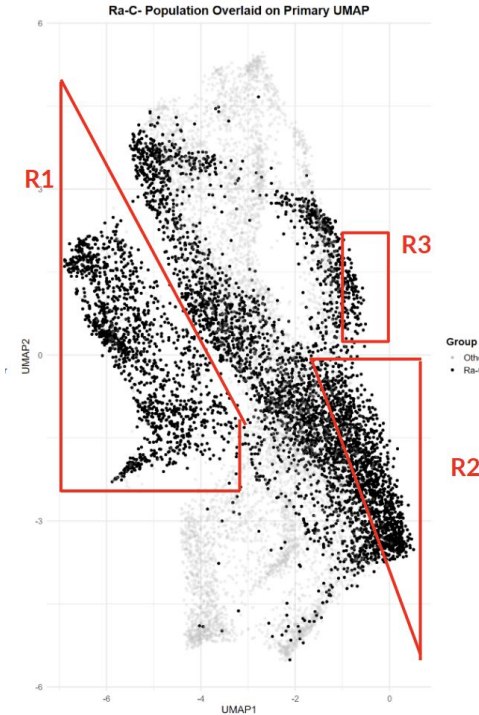

C+ Population Overlaid on Primary UMAP

Slide Intentionally Left Blank

Characterizing R1 of the Ra-C- Population : Top Genes/CD Markers

Top 20 Genes

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
HBD	2.600558	0.9332628
STXBP5	2.387527	1.9199339
ITGA2B	2.281450	0.7618509
LTBP1	2.165247	0.6594841
GP1BB	2.099730	0.6598842
RAP1B	1.979228	2.1936284
ABCC4	1.925287	1.0928290
SPINK2	-1.924872	1.7426931
PLCB1	-1.875570	1.8908794
RNF220	-1.828865	2.5707110
SLC24A3	1.793877	0.6780904
MED12L	1.747457	0.8803669
C1QTNF4	-1.742253	1.5501189
PLXDC2	1.705843	0.9520478
SH3BGRL3	1.699835	2.6590816
ATP8B4	-1.679106	1.6559922
NKAIN2	-1.664888	2.7157017
RAB27B	1.629171	0.9554767
UBE2C	1.595314	0.9767416
LAT	1.566811	0.4916719

Top 20 CD Markers

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
CD74	-1.3886235	2.2728535
CD99	-1.0149487	1.4580142
CD84	0.9984581	0.4955969
CD48	-0.9684666	0.8941089
CD63	0.9135788	2.2570242
CD36	0.8744668	0.3246596
CD34	-0.8677303	1.3150180
CD44	-0.8126885	1.2309836
CD55	0.7716636	0.5686104
CD52	-0.6601983	1.5512427
CD53	-0.6224523	0.8918330
CD164	-0.4921140	1.8707070
CD200	-0.4793880	0.4247843
BICD1	-0.4351051	1.0190811
CD69	0.3802578	0.6807524
CD37	-0.3784251	1.4375666
CD9	0.3453646	0.1117994
CD226	0.3390867	0.1304901
CD82	0.3335424	0.5675827
CD302	-0.3262979	0.5838855

Ra-C- Population Overlaid on Primary UMAP

Characterizing R1 of the Ra-C- Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
HBD	Hemoglobin subunit	Strong Erythroid / RBC Lineage Indicator
STXBP5 / CD41	Platelet Granule Trafficking/ Canonical Megakaryocyte marker	Major Platelet Surface Integrin / Definitive Megakaryocytic identity
GP1BB	Platelet adhesion receptor	Strong Platelet Indicator
RAP1B	Small GTPase; platelet activation/vascular interaction	Strong Platelet Indicator
ABCC4	Associated with Platelet dense granule function	Strong Platelet Indicator
CD36	Lipid Scavenger Receptor on Mono/Macro	Indicates Mono/Macro Bias
CD9	Exosome Tetraspanin present in Platelets	Platelet indicator
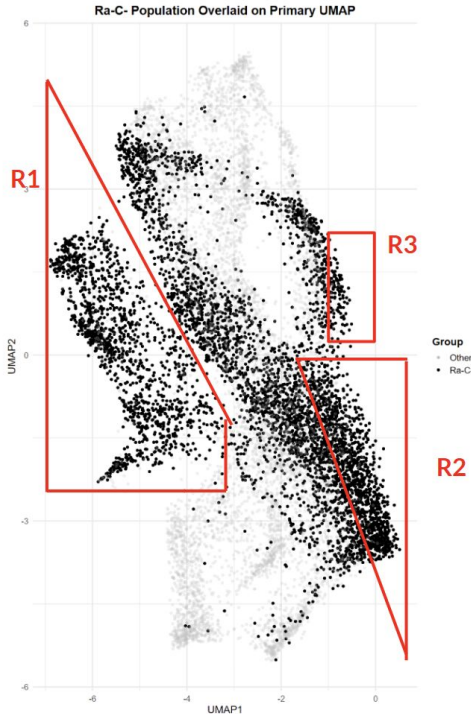
Ra-C- Population Overlaid on Primary UMAP

Characterizing R1 of the Ra-C- Population : Key Upregulated Genes and Their UniProt / Gene Context

Downregulated Genes / CD Markers

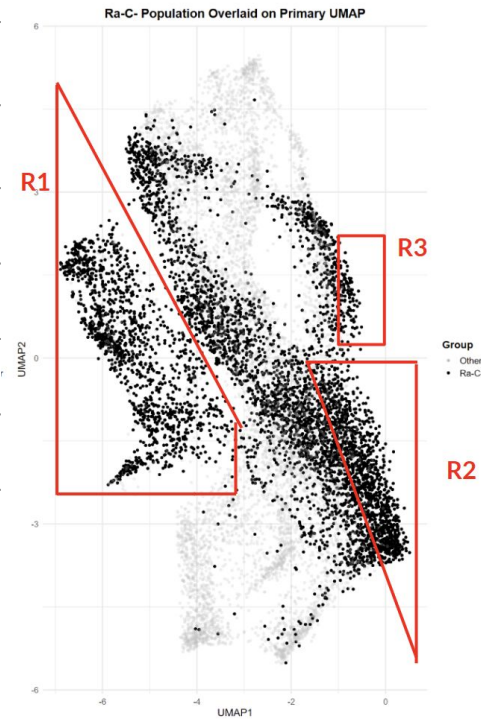| Gene	Role / Pathway	Interpretation
CD52/CD99/CD48	Lymphocyte Marker / Adhesion-Migration	Suppressing Lymphoid Lineage
CD74	MHC II Invariant Chain	Suppressing APC Lineage
CD34	Canonical HSC marker	Suppressing HSC multipotency
CD53	Tetraspanin in Lymphoid cells	Suppressing Lymphoid Lineage
CD37	Tetraspanin / B Cell Signalling	Suppressing B-Lineage

Ra-C- Population Overlaid on Primary UMAP

Characterizing R1 of the Ra-C- Population : Connection to FW/John/Yale's Papers + Characterization

In Line With Their Findings

FW: Identifies RA-C- as highly heterogeneous multipotent. R1 however demonstrates strong erythro–meg skewing, strong suppression of lympho-myelo signature. Potential for additional filtering here!

JE: Supports John's claim: in absence of lymphoid directing cues, HSCs trend towards default myelo-erythroid fate

Yale: Presence of residual lymphoid signals in clearly non lymphoid population validates claim: fate commitment is preceded by mixed transcriptional states.

Conclusion

- Upregulation of ITGA2B(CD41), GP1BB, HBD, CD36, STXB5 shows strong megakaryocyte and erythrocyte lineage priming.
- Downregulation of CD34, SPINK2, CD74 and CD48 indicating departure from HSCs
- Active suppression of both myeloid and lymphoid fates!

*This region is characterized by a Erythrocyte/Megakaryocyte primed population.*

Ra-C- Population Overlaid on Primary UMAP
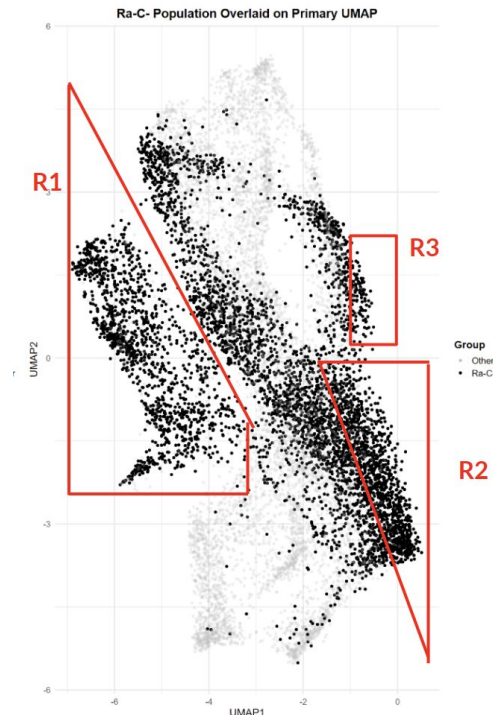
Characterizing R2 of the Ra-C- Population : Top Genes/CD Markers

Top 20 Genes

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
NKAIN2	1.6173566	2.7157017
HIST1H4C	-1.3112881	2.2538890
HMGB2	-1.2863952	2.6003865
TOP2A	-1.1748550	1.2552635
UBE2C	-1.1563485	0.9767416
TUBB4B	-1.1293153	1.5703254
HBD	-1.1049158	0.9332628
HLA-DRA	1.0408271	2.3423770
MSI2	1.0300816	2.5742865
RNF220	1.0277936	2.5707110
UBE2S	-1.0216306	1.3800974
SPINK2	1.0212230	1.7426931
HIST1H1B	-1.0188295	1.0765159
HOPX	1.0046959	1.1811817
MKI67	-0.9951843	1.0553881
CENPF	-0.9712544	1.3730378
CD74	0.9527469	2.2728535
PLCB1	0.9493143	1.8908794
RRM2	-0.9471306	0.8967019
NRIP1	0.9264414	2.4574898

Top 20 CD Markers

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
CD74	0.9527469	2.2728535
CD52	0.7753846	1.5512427
CD99	0.6737835	1.4580142
CD48	0.5969166	0.8941089
CD34	0.5490061	1.3150180
CD37	0.5443091	1.4375666
CD44	0.4687010	1.2309836
CD84	-0.4014853	0.4955969
CD53	0.4005713	0.8918330
CD63	-0.3767893	2.2570242
CD36	-0.3754074	0.3246596
CD200	0.3326701	0.4247843
CD109	0.3322049	0.8986139
CD55	-0.3204002	0.5686104
CD164	0.2829355	1.8707070
SCD	-0.2633034	0.6314002
CD79B	0.2341851	0.3376392
C2CD2	0.1882607	0.4018778
HACD3	-0.1857094	1.1536948
CD38	-0.1694150	0.7399566
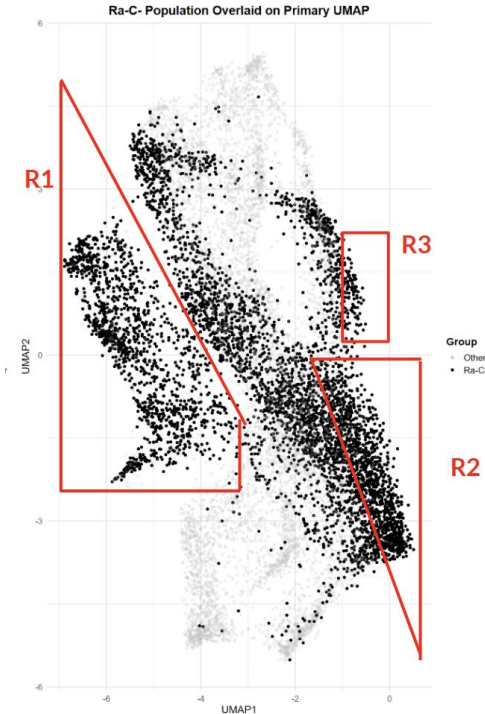
Ra-C- Population Overlaid on Primary UMAP

Characterizing R2 of the Ra-C- Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
HLA-DRA/ CD74/ CD44/ MSI2/ CD99/ CD48/ CD37/ RNF220/ MSI2	MHC II Antigen Presentation / RNA Regulation/ Notch modulation / Stemness / Immune Signalling	Immune Priming / Some APC Bias /
SPINK2/HOPX/ NKAIN2	Protease inhibition; Cell-fate decision; Na^+/K^+ transport	Maintaining Stem cell state / regulatory programs / Maintaining HSC State
RRM2/ PLCB1	DNA Synthesis / Phospholipid Signalling	Regulated metabolic readiness
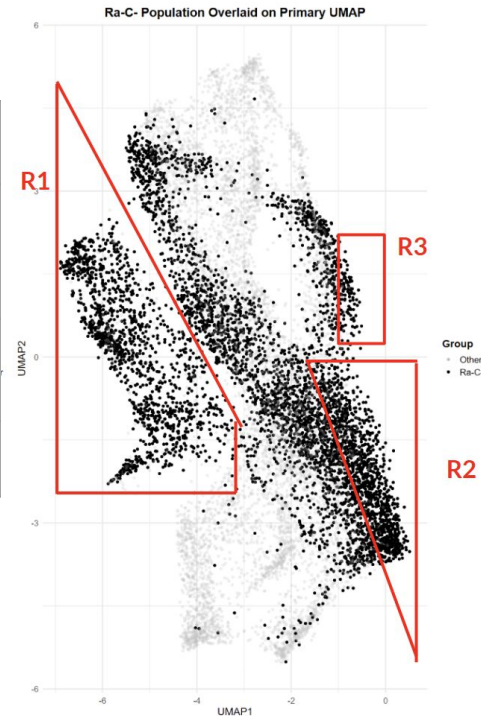
Ra-C- Population Overlaid on Primary UMAP

Characterizing R2 of the Ra-C- Population : Key Upregulated Genes and Their UniProt / Gene Context

Downregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
HMGB2/ TOP2A/ HIST1Hxx/ UBE2C/ MKI67/ CENPF	Chromatin Dynamics / Replication / Mitosis / Cell Cycling	Reduced proliferation and exit from cell cycle ; This population is not actively dividing
CD53	SLAM Family ; B/T Coreceptor	Suppression of lymphoid lymphoid fate
CD84	Tetraspanin Family ; Immune Activation	Suppression of myeloid/immune signalling
CD63	Exosome associated ; Immune Activation	Suppression of granulocyte lineage
CD79B	B-cell receptor component	Suppression of B/Lymphoid fate
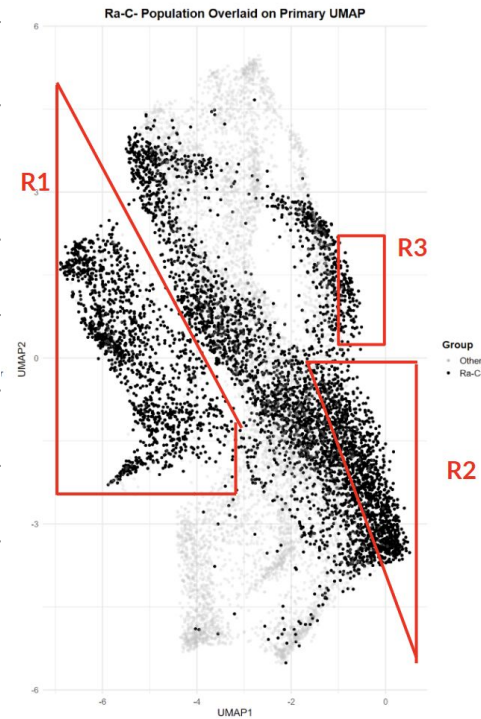
Ra-C- Population Overlaid on Primary UMAP

Characterizing R2 of the Ra-C- Population : Connection to FW/John/Yale's Papers + Characterization

In Line With Their Findings

> Aligns with FW's description of Ra-C- retaining multilineage potential. Claim: lineage restriction precedes proliferation, appears to be happening here

> Yale: "Pre-fate-resolution state": transcriptionally immune skewed, not yet clonally committed. Co-expression of stemness genes with APC/Lymphoid markers supports this

> John: Notch deficient lead to immune primed, undifferentiated state. Appears to be valid.

Conclusion

- MHCII and Co-receptor genes (HLA-DRA, CD74, CD44, CD37) upregulated suggesting APC potential
- Proliferation genes downregulated
- Stem-associated and regulatory factors maintained

*Non-committed, Immune Skewed HSCs, slight hints of Pre-DC or Pre-B/Myeloid Bifurcation Progenitors*

Ra-C- Population Overlaid on Primary UMAP
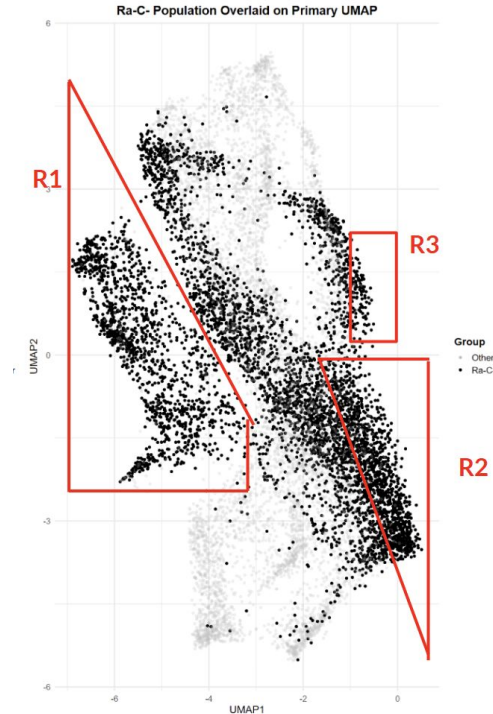
Characterizing R3 of the Ra-C- Population : Top Genes/CD Markers

Top 20 Genes

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
FHIT	-1.4255407	1.9786639
WWOX	-1.3788230	2.2088429
NKAIN2	-1.3215391	2.7157017
RAD51B	-1.3124129	1.8692422
IMMP2L	-1.2600372	1.9230808
LRMDA	-1.1954152	2.1615629
SMYD3	-1.1305355	2.3841554
MALAT1	-1.0991620	5.8601224
STXBP5	-1.0934012	1.9199339
HMGB2	1.0885756	2.6003865
PTTG1	1.0686504	1.1477022
ZBTB20	-1.0380624	1.4574735
LRBA	-1.0210120	2.6127848
MIR924HG	-0.9673563	1.6771880
INPP4B	-0.9477302	1.7127337
SNHG3	0.9329528	1.3746169
EIF4G3	-0.8830465	1.8522299
HIST1H4C	-0.8687174	2.2538890
HBD	-0.8567509	0.9332628
AC008014.1	-0.8518681	1.1810701

Top 20 CD Markers

ID	logFC	AveExpr
<chr>	<dbl>	<dbl>
CDV3	0.4613797	1.2229999
CD84	-0.3533260	0.4955969
CD164	0.3068021	1.8707070
CD36	-0.2896939	0.3246596
CD82	-0.2824216	0.5675827
CD63	-0.2598229	2.2570242
CD74	0.2591115	2.2728535
CD48	0.2480557	0.8941089
HACD3	0.2477402	1.1536948
CD37	-0.2100434	1.4375666
CD44	0.1974135	1.2309836
CDT1	-0.1961757	0.7378909
CD34	0.1893594	1.3150180
CD38	-0.1886832	0.7399566
CD99	0.1868761	1.4580142
CD55	-0.1723566	0.5686104
CD58	-0.1668748	0.5847240
BICD1	-0.1613899	1.0190811
CD2AP	0.1395765	0.9274637
CD47	-0.1384016	1.2640750
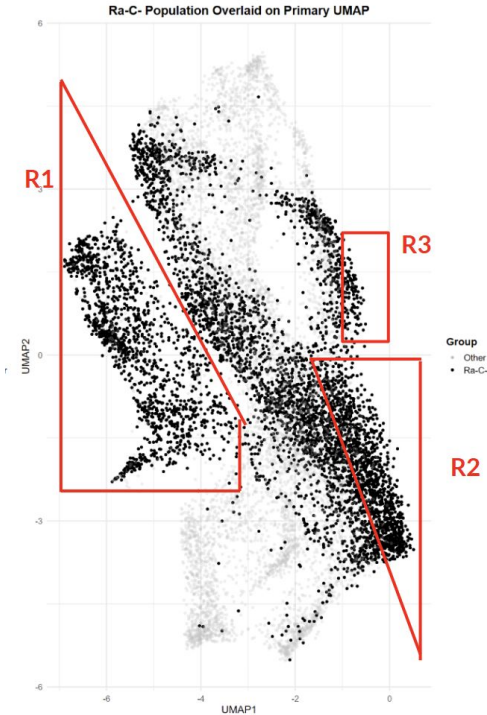
Ra-C- Population Overlaid on Primary UMAP

Characterizing R3 of the Ra-C- Population : Key Upregulated Genes and Their UniProt / Gene Context

Upregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
HMGB2	DNA Binding Chromatin Associated Protein	Proliferative State Commencing
PTTG1	Securin; Mitotic Checkpoint Regulation	Confirms Cell Cycle Activation
EIF4G3/ SNHG3/ RNF220	E3 Ubiquitin Ligase Module / Small nucleolar RNA host / Translation Initiation Scaffold Protein	Cell growth and RNA Biosynthesis / Broad pathways supporting cell cycle activation
CD3V	TCR Delta variable domain	Lymphoid Priming
CD74	MHC ClassII Chaperone	Lymphoid Priming
CD48	SLAM Family receptor on T/NK Cells	Indicating Lymphoid Priming
CD34	Canonical Stem cell gene	Maintaining HSC State
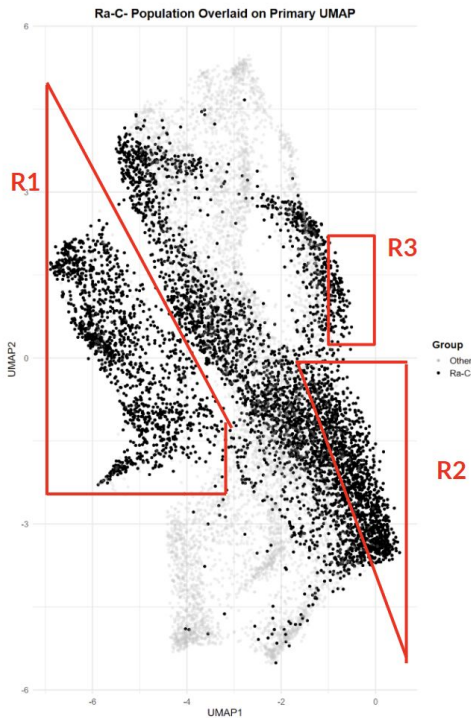
Ra-C- Population Overlaid on Primary UMAP

Characterizing R3 of the Ra-C- Population : Key Upregulated Genes and Their UniProt / Gene Context

Downregulated Genes / CD Markers

Gene	Role / Pathway	Interpretation
WWOX/FHIT	Tumour Suppressors/ DNA Repair	Loss linked to Cell Cycle Entry
STXBP5	Exocytosis Regulator ; Platelet Lineage	Divergence from Megakaryocytic fate
ZBTB20	Transcriptional Repressor, regulated B-cells	Divergence from B-lineage fate
LRBA	Vessel Trafficking in B-cells	Divergence from B-lineage fate
HBD	Haemoglobin Subunit	Divergence from Erythroid fate
CD36	Scavenger Receptor for Lipids	Divergence from Monocyte/Macrophage
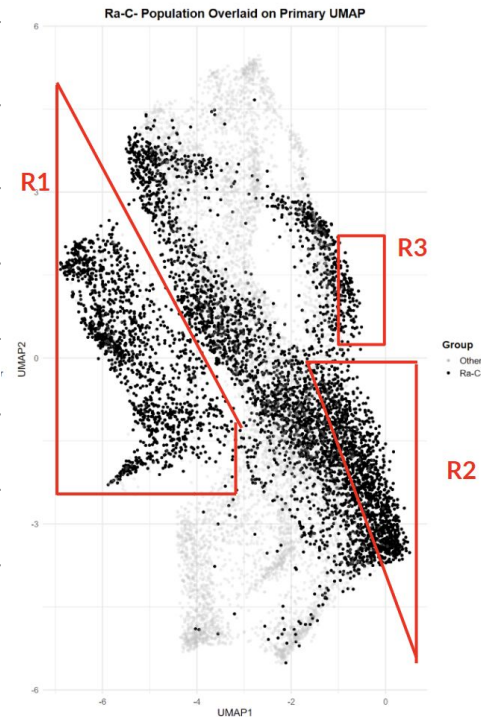
Ra-C- Population Overlaid on Primary UMAP

Characterizing R3 of the Ra-C- Population : Connection to FW/John/Yale's Papers + Characterization

In Line With Their Findings

FW: Fits FW's claim that lineage restriction is downstream of proliferation, not concurrent, since suppression of lineage specific genes, while upregulating MHC-II. Also, this is the most multipotent subset in the entire dataset by a long shot.

JE: Matches his description of niche-deprived progenitors in engineered environments : high proliferation, low specification

Yale: Aligns with his "Progenitor cloud" (high biosynthetic output, ambiguous lineage markers, no fate commitment Co-expression of CD34, CD74, HMGB2, CD38 matches his "pre-commitment" UMAP cluster signature.

Conclusion

- Upregulated mitotic/biosynthetic genes (HMGB2, PTTG1, SNHG3), indicating proliferation
- Mild upregulation of lymphoid priming
- Mild downregulation of erythroid (HBD), B-lineage (ZBTB20, LRBA) and immune regulators (CD84, CD55, C2AP) -> Non-committed, non-lineage skewed population

This region is characterized by cells beginning to proliferative, with almost no bias towards any particular fate.

Ra-C- Population Overlaid on Primary UMAP
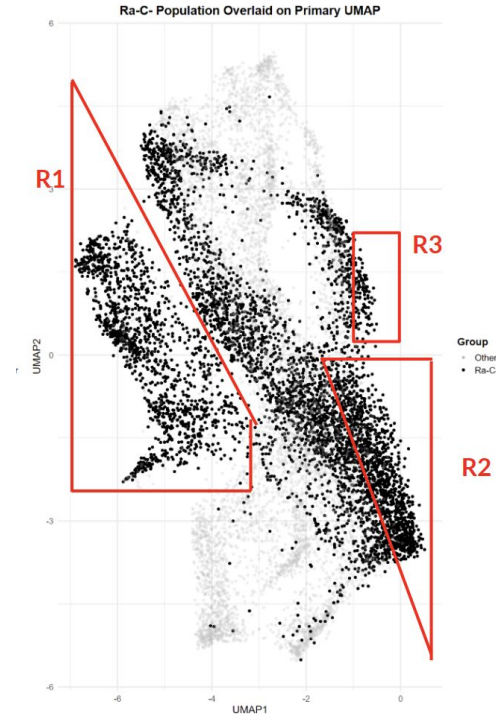
RQ3: Is There Quantifiable Evidence of Dendritic Cells in These Subsets?

RQ4: Is There Evidence of Mast Cell Fate in This Subset?

RQ5: Is There Statistically Significant Evidence of Heightened T-Cell Genes Among Any Population?