Secretory Landscape Along Hematopoiesis

Nathan Lewis Lab

Hsuan-lin Her (Charlene)

BISB year 1

Part I: Single Cell Analysis

Dataset and work flow

	Human hematopoietic stem cell (CD34+)	Human hematopoietic stem cell (LIN-)	Mouse stromal cell (wolock et. al)	Mouse stromal cell (Tihonova et.al)
FACs	CD34+ (haematopoietic stem cells)	LIN- (to later lineage)	long bones, followed by sorting viable CD45-Ter119-(non-hematopoietic) and CD31-(non-endothelial) cells.	CD45lowTER119low(non-hematopoietic)tdTomato+ cells(endothelial) were isolated
Technology	Drop-seq	Drop-seq	Drop-seq	
Summary stats				

	Cell 1	Cell 2
Gene 1	5	2
Gene 2	8	1

→

Filter cells

1. Transcript per cell (arbitrary threshold)
2. 2% mitochondrial transcript (25%)

→

Log normalize

1. Total transcript per cell (exclude highly expressed genes > 5%)
2. Log

* Pitfall: "compositional data"

→

Map to human homolog (Mouse Genome Informatics)

Visualization, clustering, trajectory inference and pseudotime

Highly variable gene (for each bin of mean expression, select dispersion z-score > 2)
Min_mean = 0.0125
Max_mean = 3
Min dispersion = 0.5

→ PCA (n=40) → Neighborhood graph →

1. UMAP
2. Force-directed layout
3. (PAGA)

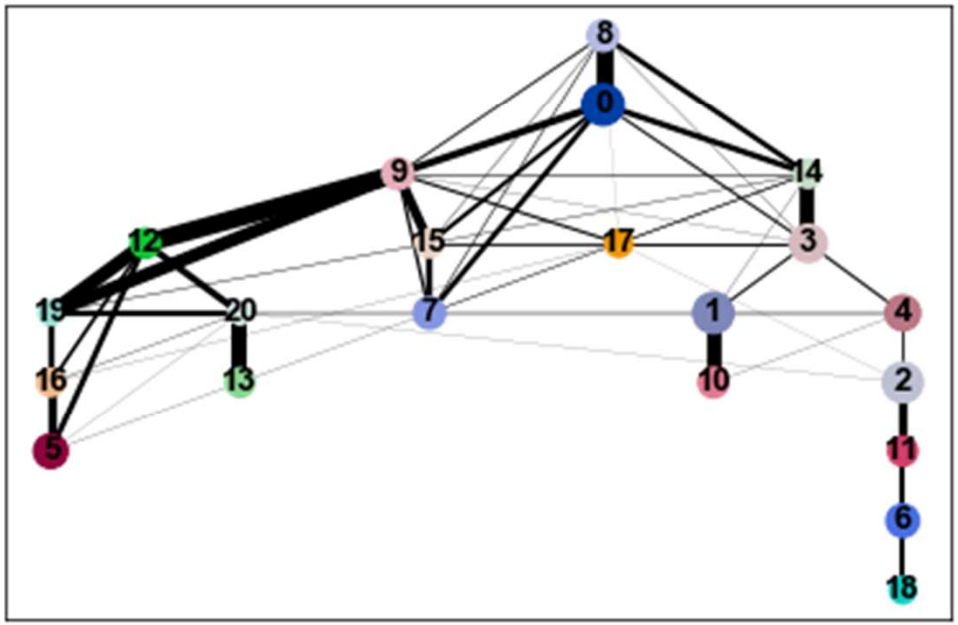
→

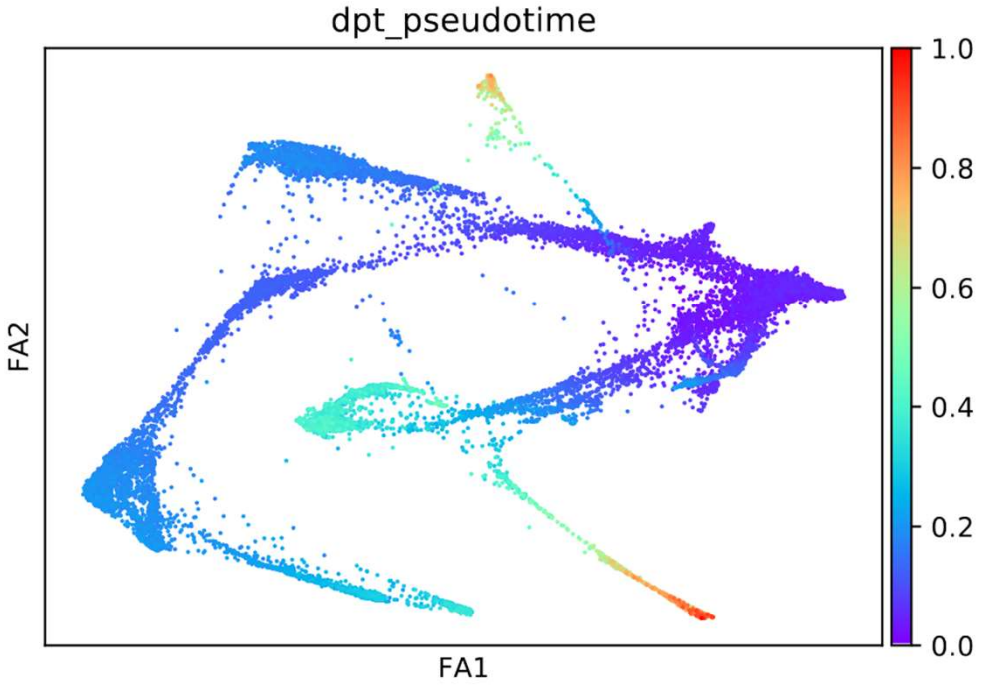
1. Louvain
2. Leiden

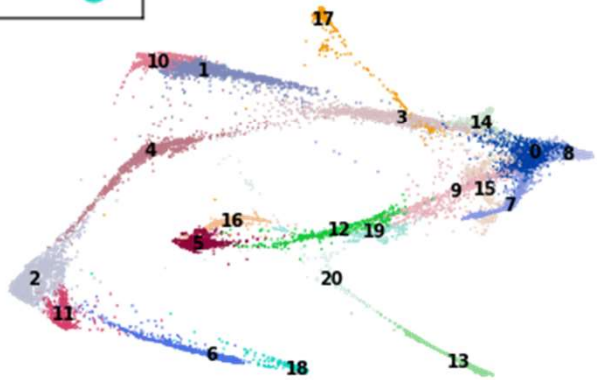
→

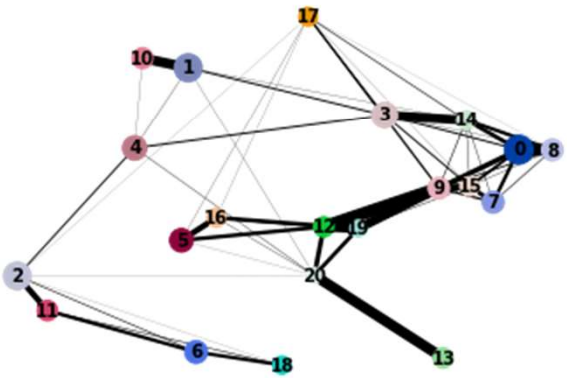
1. PAGA (branch detection)
2. Diffusion pseudotime
3. pdt

Trajectory Inference

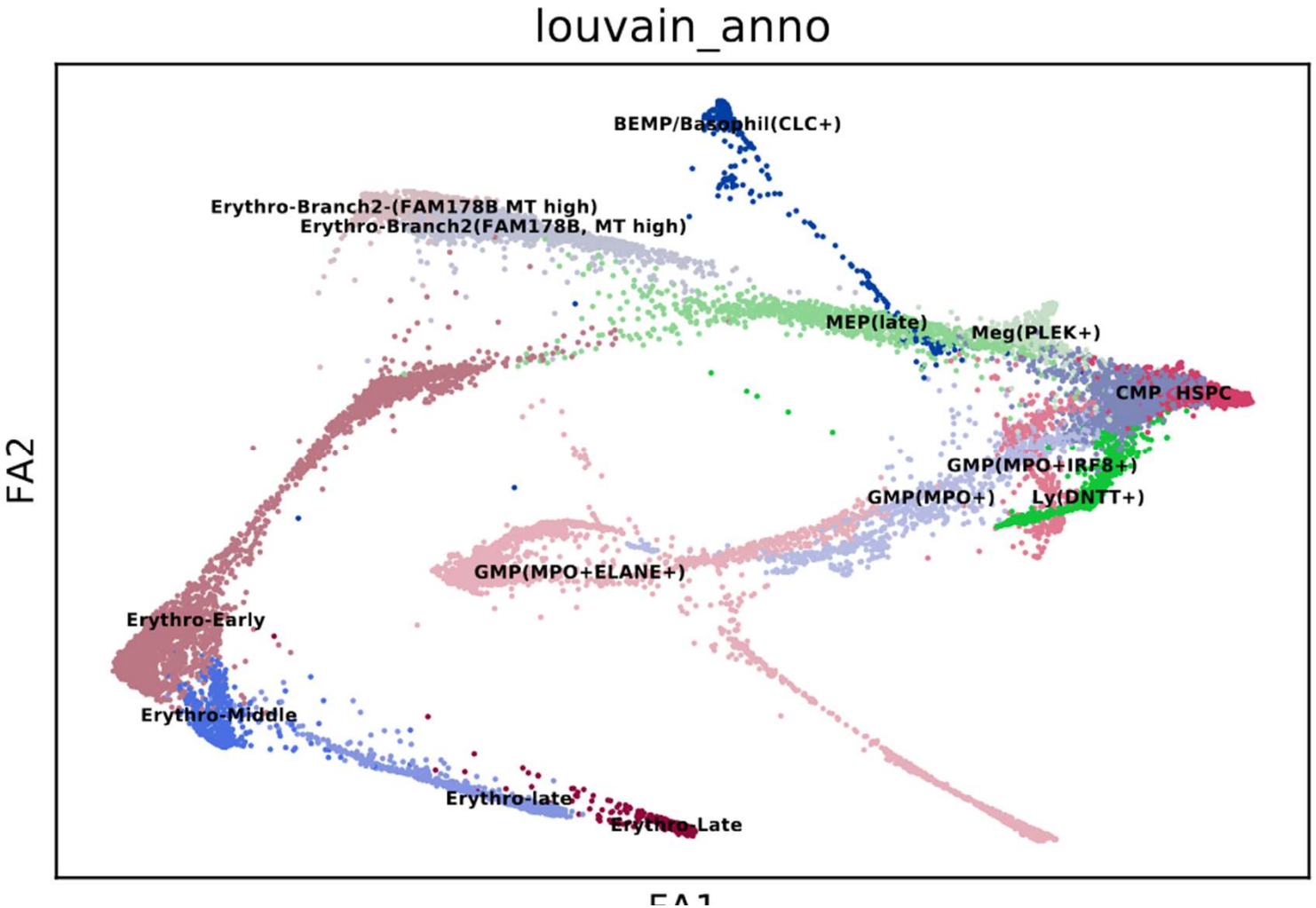






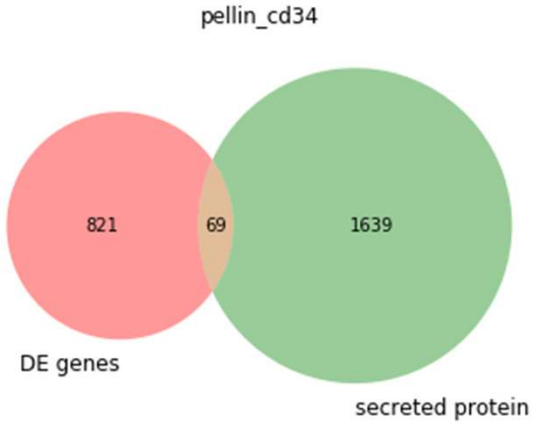


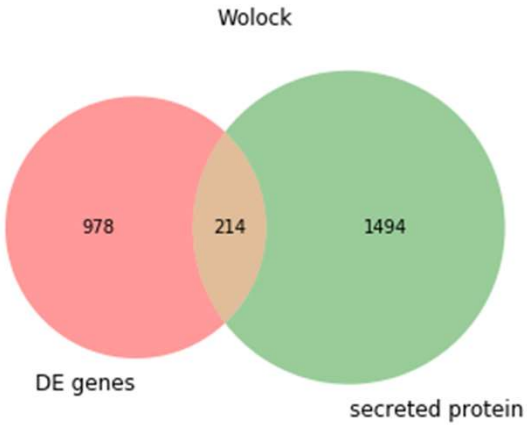
Labelling cluster: Hematopoietic cells

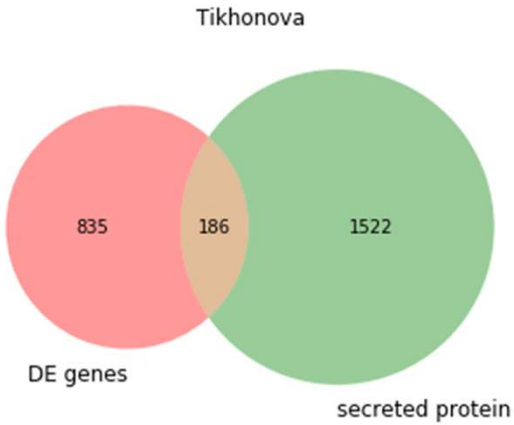


Part II: Secretome

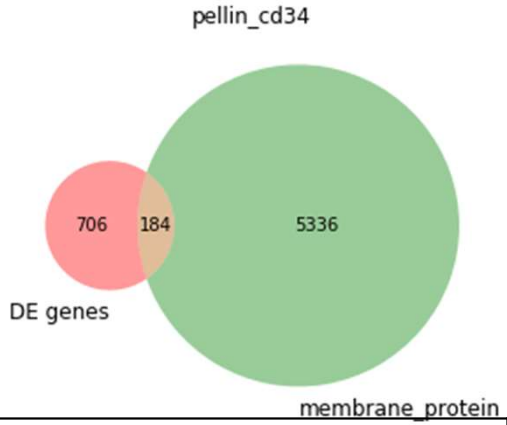
Fisher Exact test

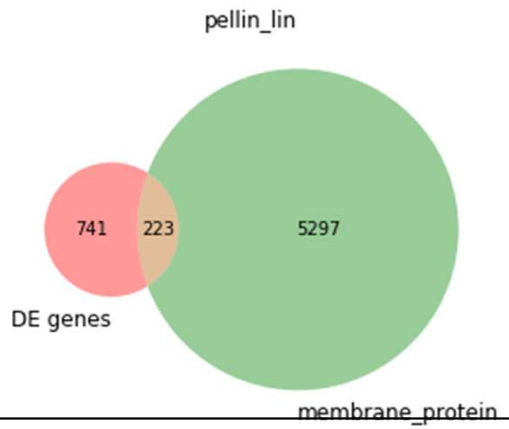


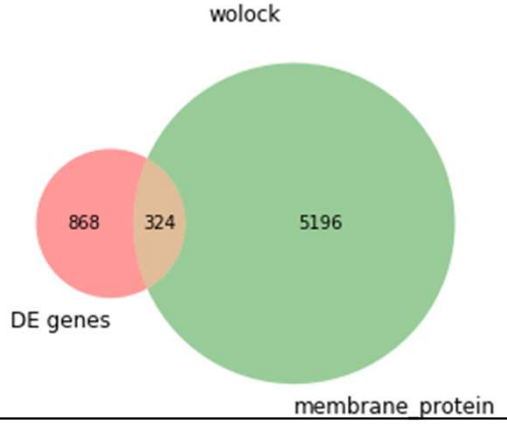


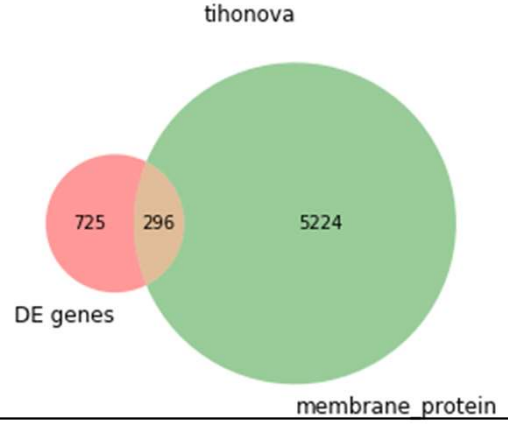


0.2189	0.1536	1.0191e-17	0.0002





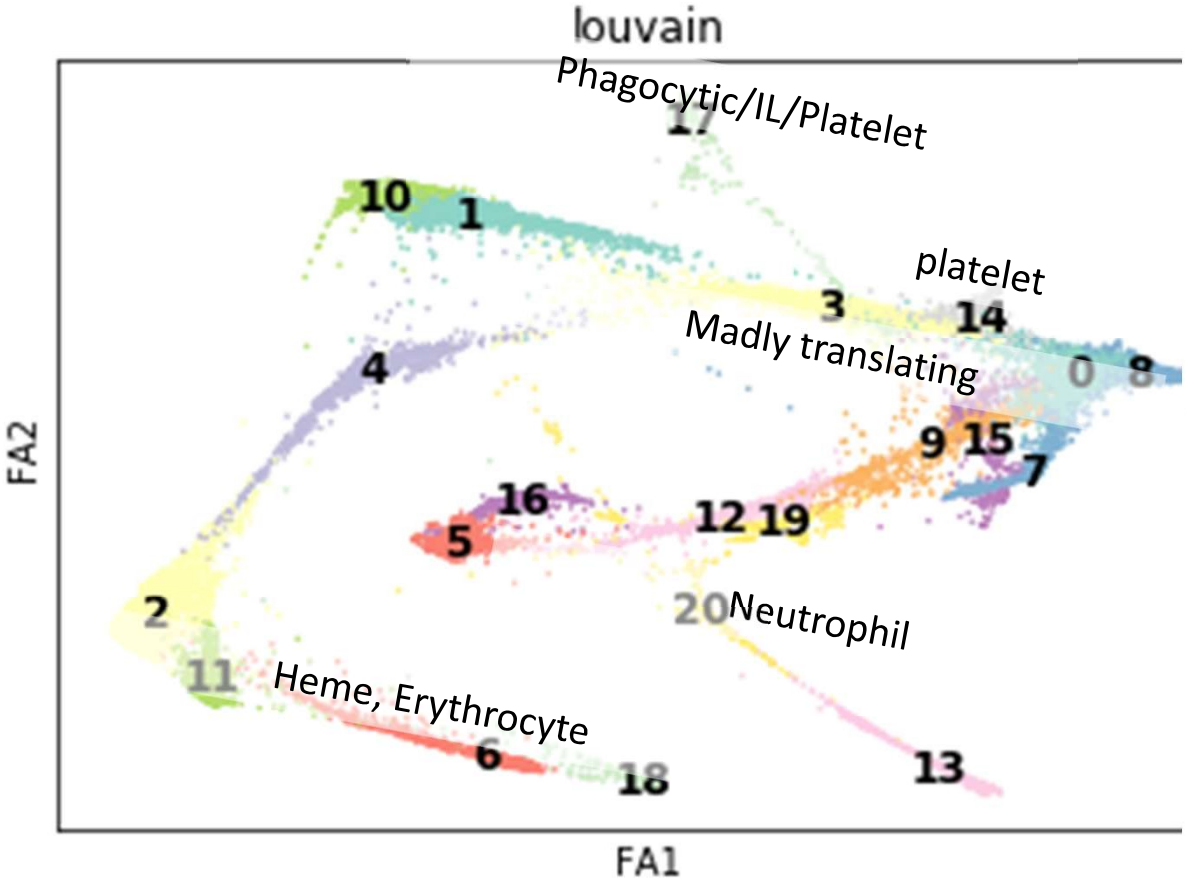


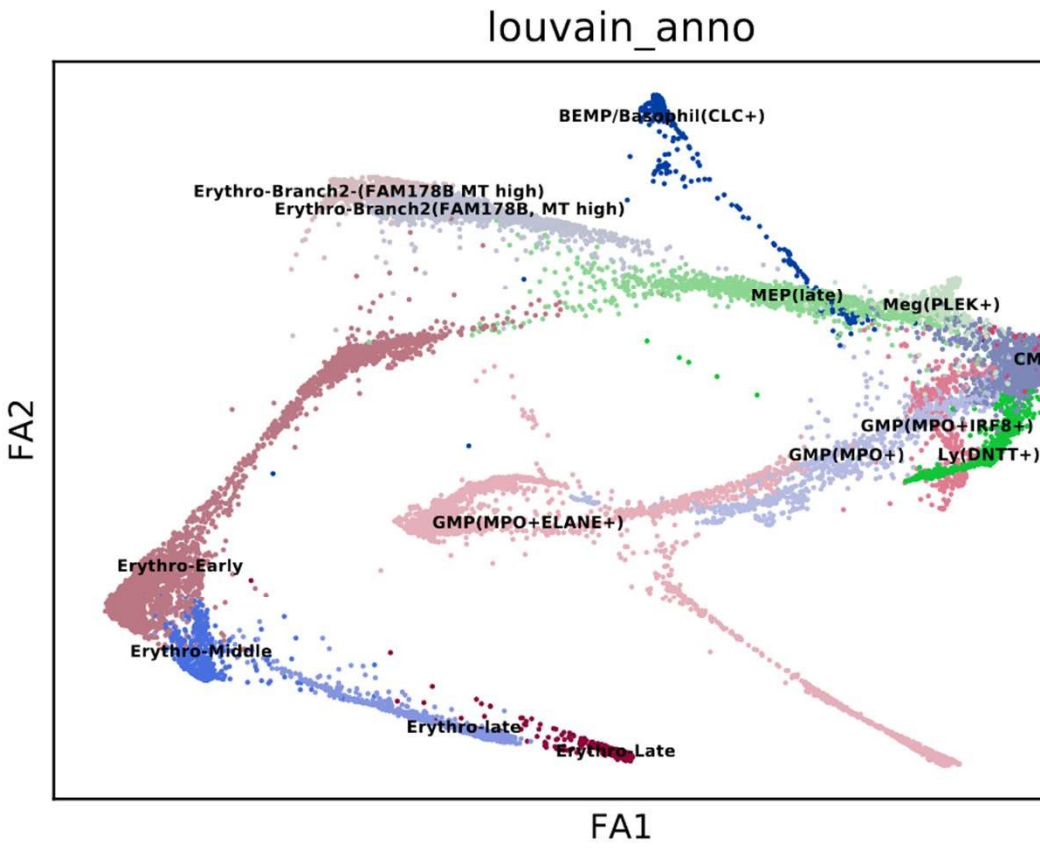


0.4815	0.5540	2.0242e-05	3.0663e-06

Are secretory protein important?

Are secretory proteins important





Cell2cell workflow

	Blood Cluster 1	Blood Cluster 2
Gene 1	23.4	3.5
Gene 2	32.5	9.62

	Stroma Cluster 1	Stroma Cluster 2
Gene 1	39.2	2.5
Gene 2	5.4	77.2

Gene percentile

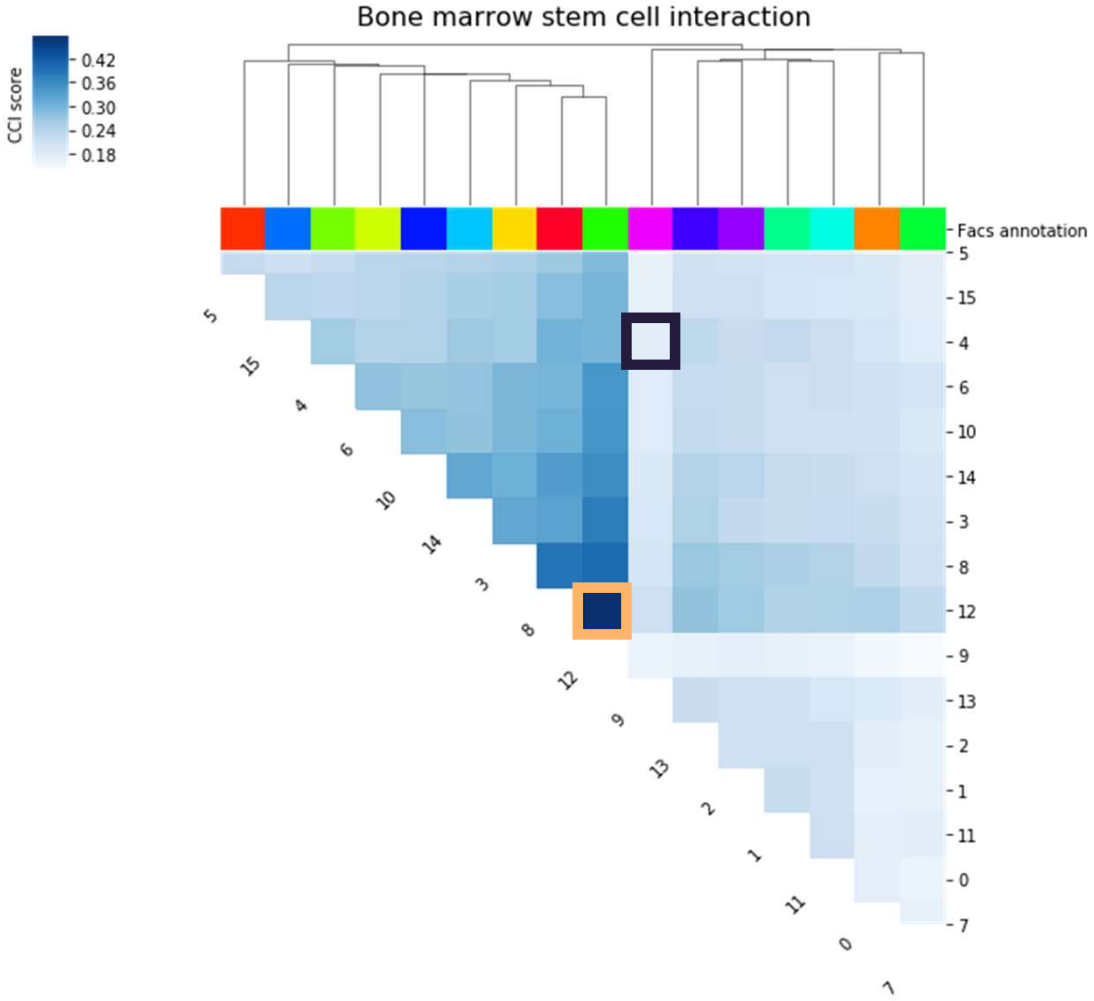
75%, 90% *

	b1	b2	s1	s2
g1	1	0	0	1
g2	1	0	0	1

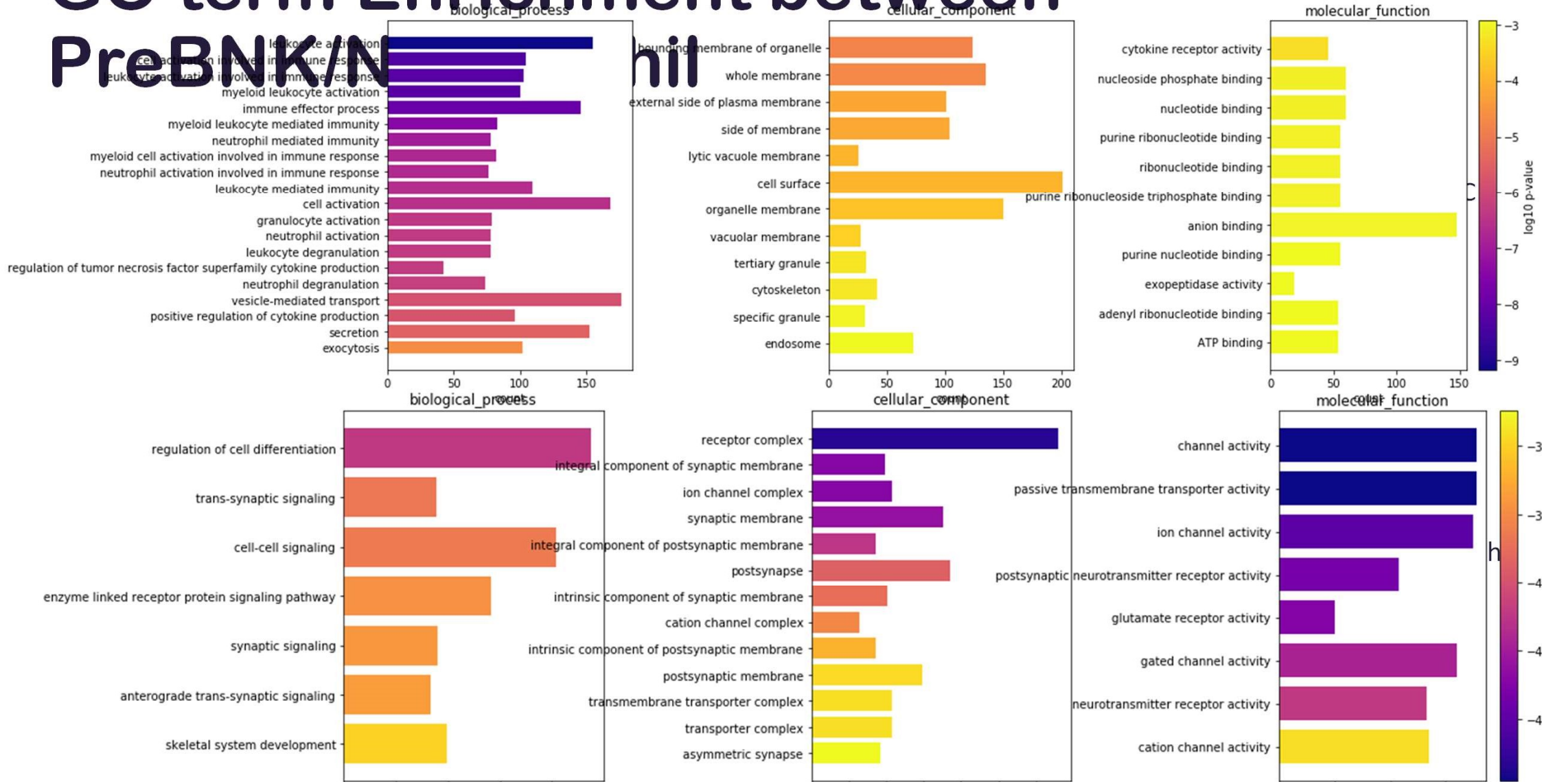
Average* expression profile per cluster
Cross dataset: normalized expression profile*

	Ligand	Receptor
Gene 1		
Gene 2		

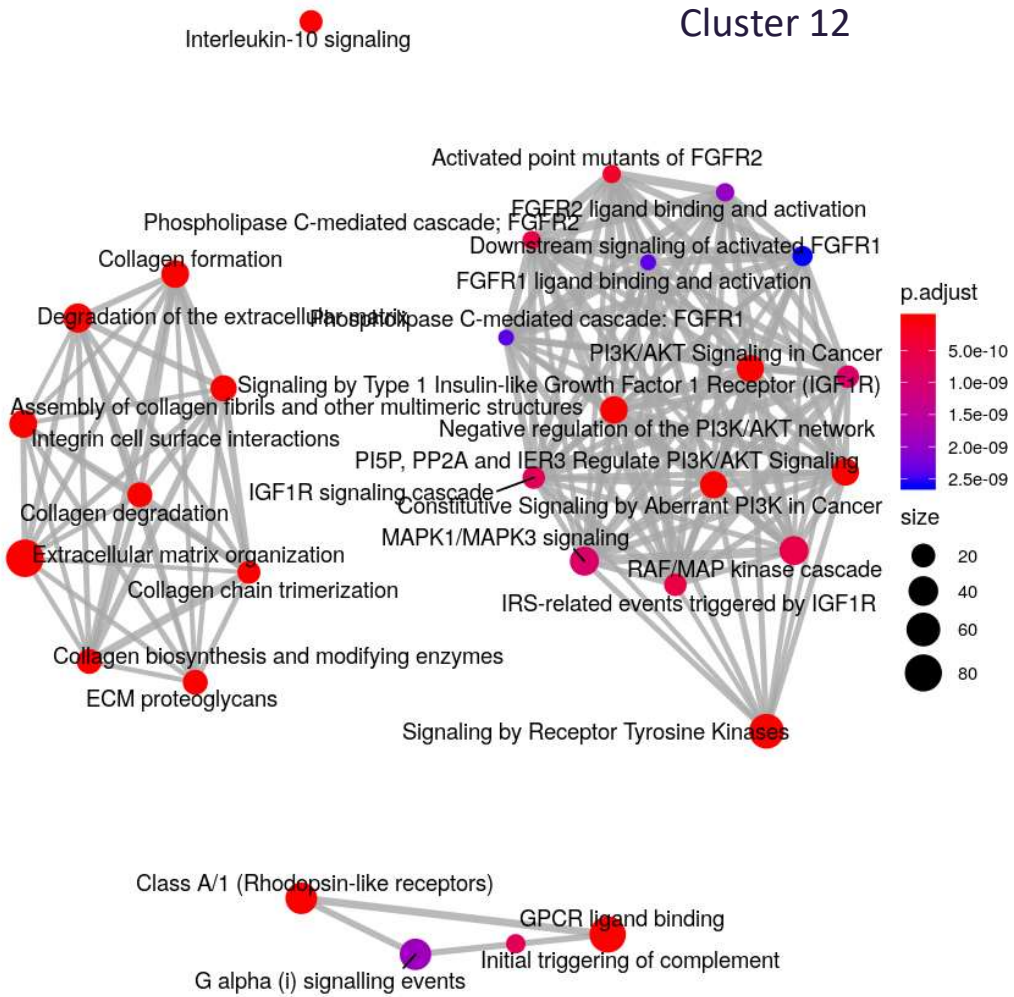
Interaction between blood cells

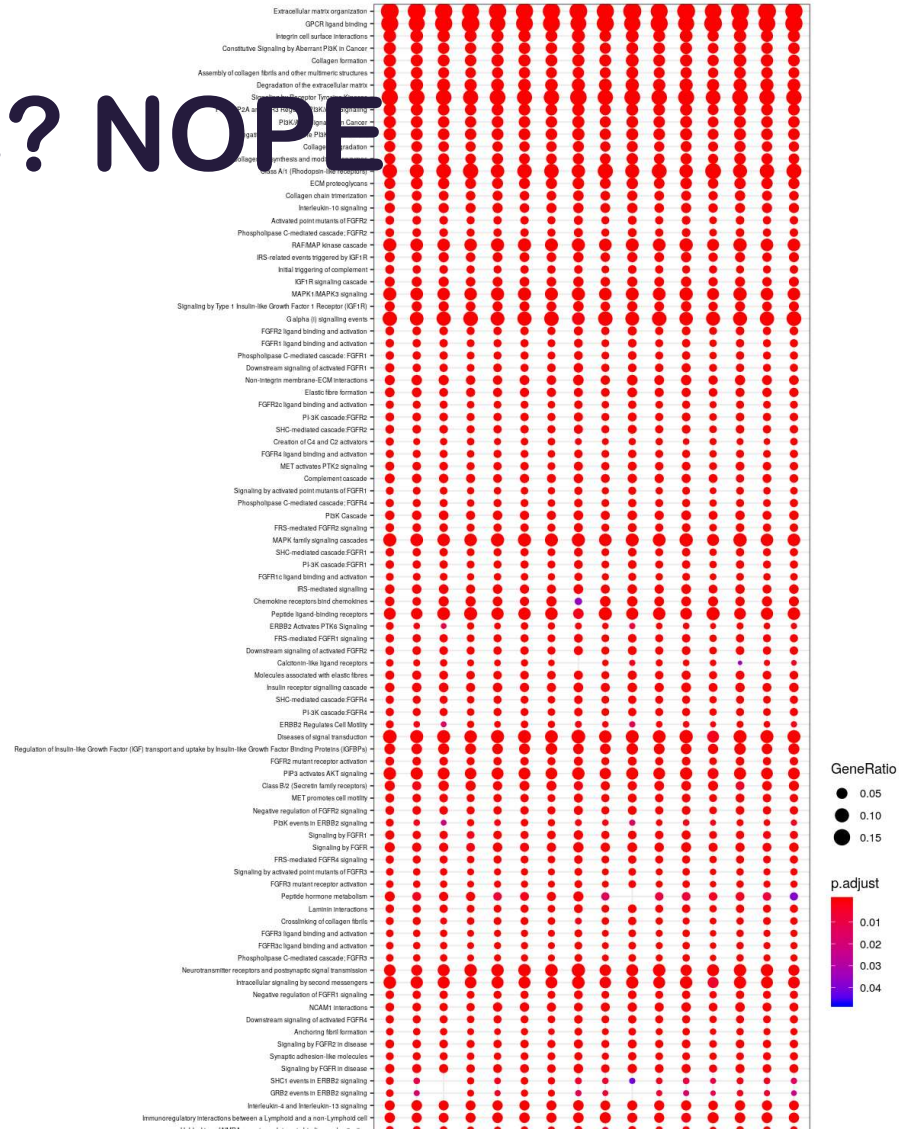


GO term Enrichment between PreBNK/N hil



Are interactions specific? NOPE





Cluster 12 all partners

Part III: Secretory Machinery

Method

Differential expression analysis
1. t-test
2. Not filter for fold

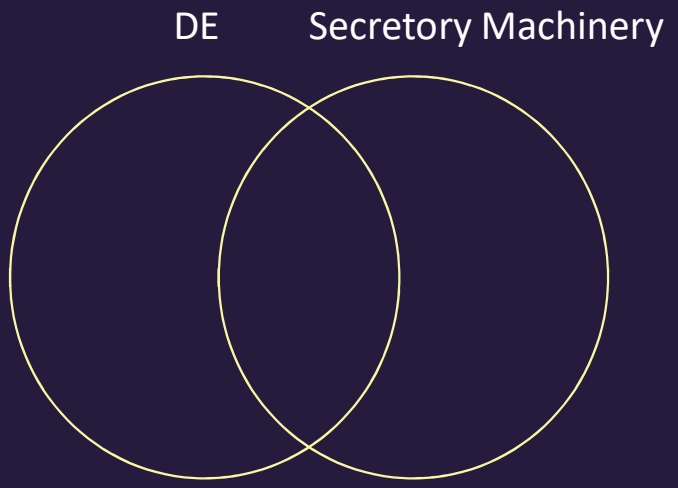
Module	Subsystem
Capacity Control	ERAD
	UPR
Folding	Protein folding
	Translocation
Glycosylation	ER glycosylation
	Golgi glycosylation
Trafficking	COPI
	COPII
	Post-Golgi trafficking
	Trafficking regulation

Feizi's reconstruction

```
gly['Process'].value_counts()

N-glycan                  42
O-glycan_notMucin         36
O-glycan_Mucin            35
lacto_neolacto series     27
heparan sulfate           24
GPI-anchor                23
additional                21
chondroitin sulfate       19
keratan sulfate           16
ganglio series            13
globo series              10
SLC_transporter           10
terminal modification      5
sphingolipid               4
Hyaluronic acid            3
Others                     2
unassigned                 1
WT                         1
Name: Process, dtype: int64
```

Lewis lab's glycosylation



Secretory Machinery differential analysis (Pe lin)