Хьюстон, у нас single cell RNA-seq

Елизавета Зезюля, Калининград

Дмитрий Сергеев, Москва

Ани Сахлян, Ереван

Оксана Иванова, Санкт - Петербург

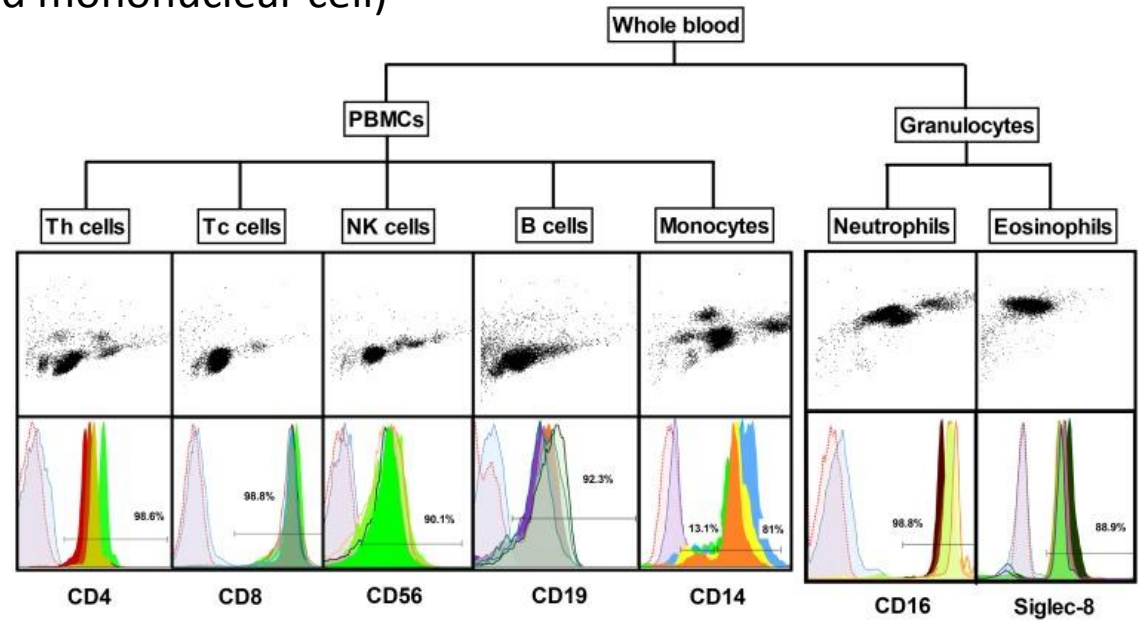
Ирина Овчинникова, Санкт - Петербург

ИНСТИТУТ
БИОИНФОРМАТИКИ

Выбор данных

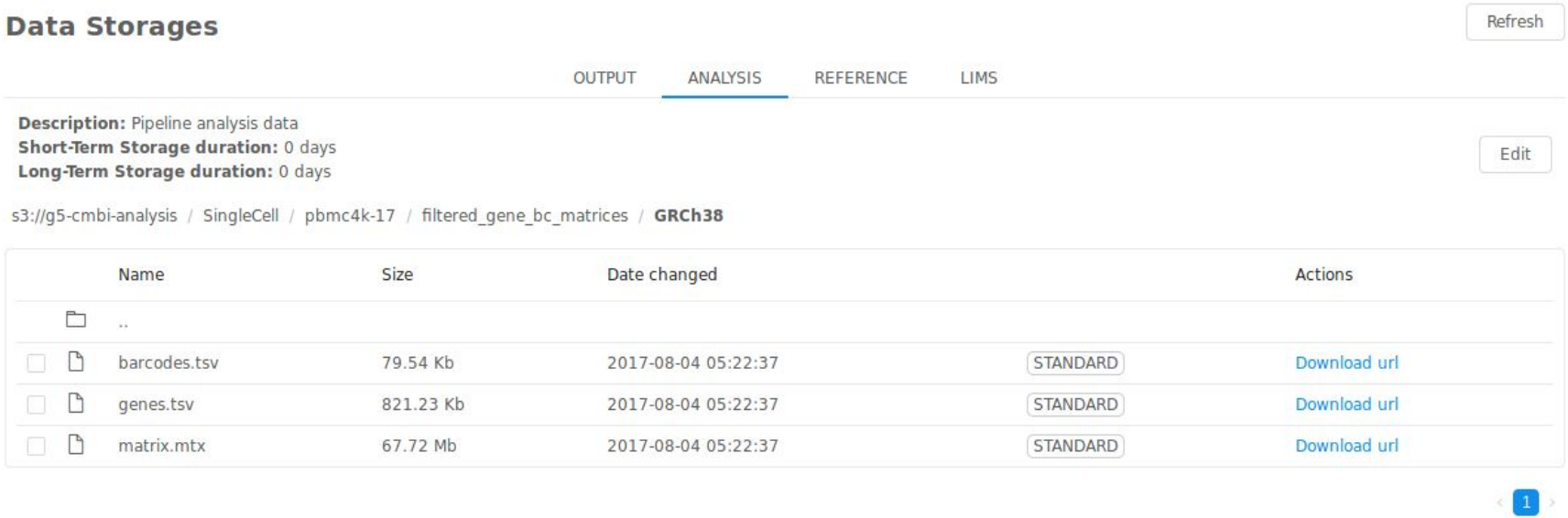
Был использован PBMC4k от здорового донора.

PBMC (peripheral blood mononuclear cell)



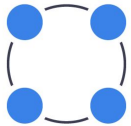
Ref.:Reinius, L. E., Acevedo, N., Joerink, M., Pershagen, G., Dahlén, S. E., Greco, D., ... & Kere, J. (2012). Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PloS one, 7(7), e41361.

Результаты CellRanger

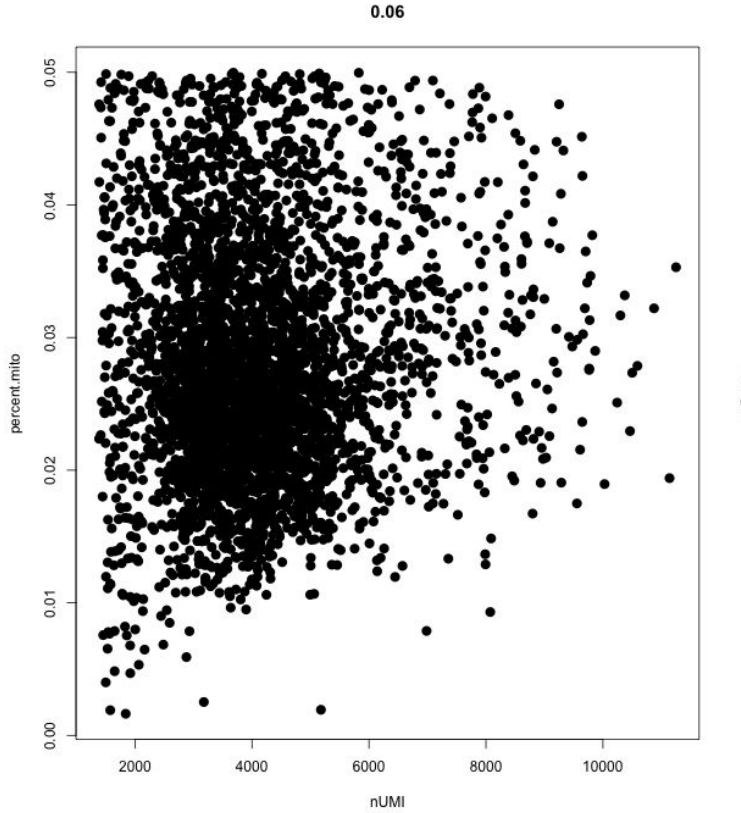


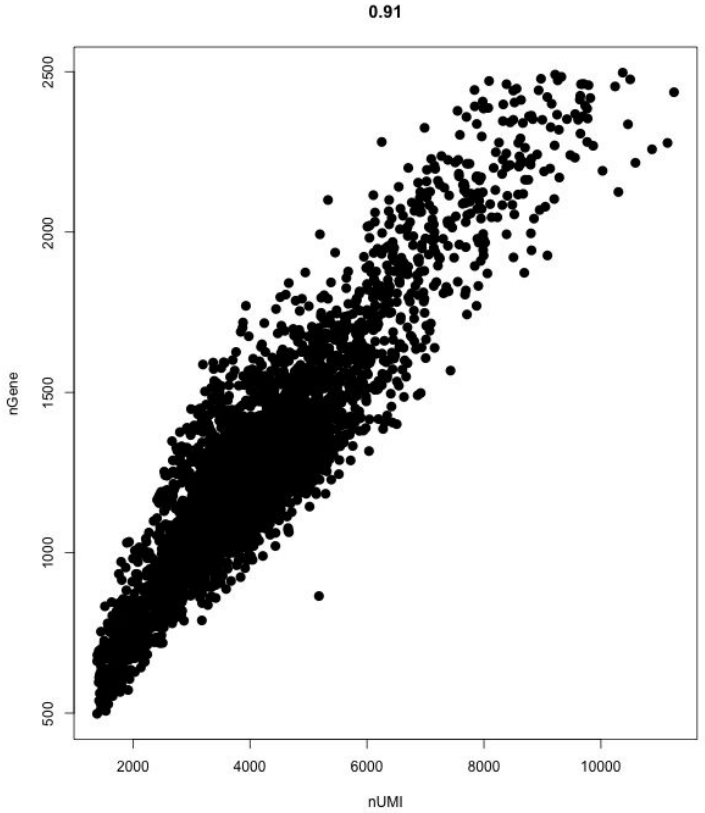
Статистическое описание

Estimated.Number.of.Cells	4,287
Mean.Reads.per.Cell	88,514
Median.Genes.per.Cell	1,239
Number.of.Reads	379,462,522
Valid.Barcodes	97.8%

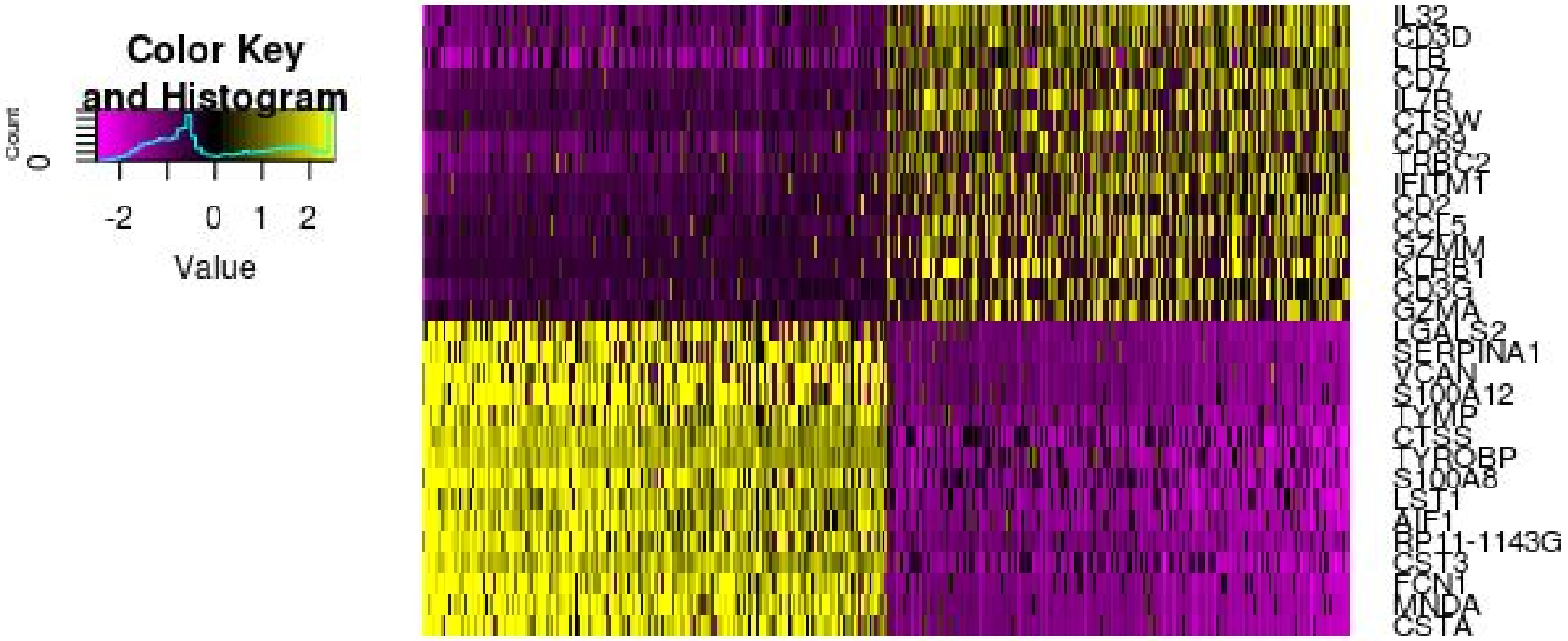


Фильтрация клеток

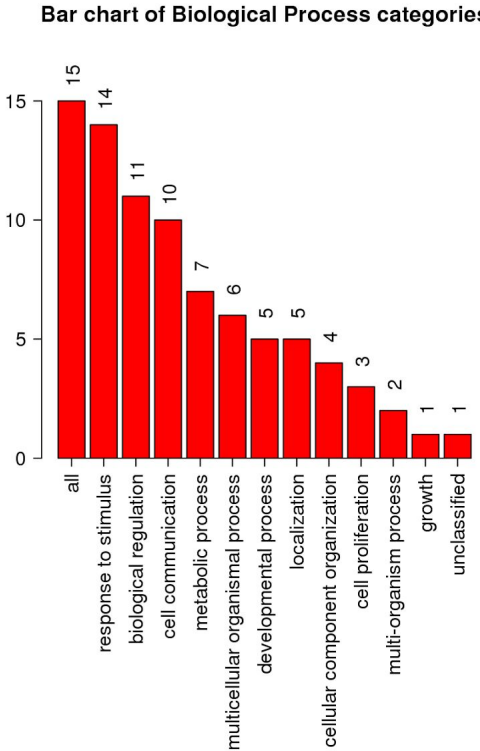




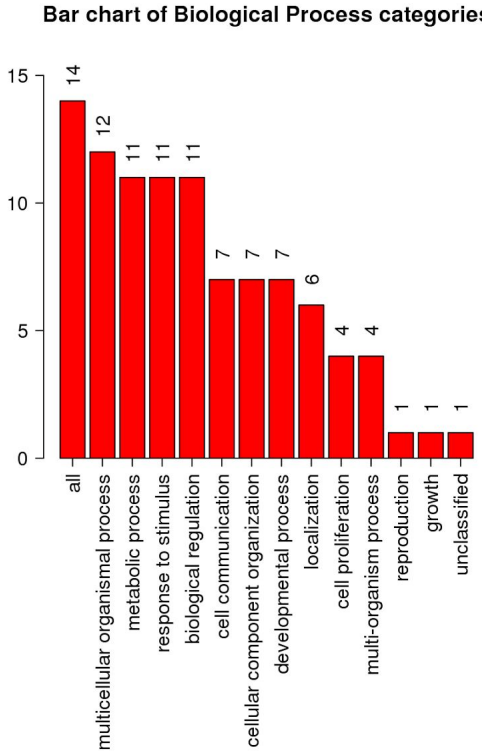
Анализ главных компонент (PCA)



Функциональный анализ генов



PCA1



PCA2

WEB-based GEne SeT AnaLysis Toolkit

Translating gene lists into biological insights...

Base ref.: Duncan, D.T., Shi, Z., Wang, J., & Zhang, B. (2013). WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. Nucleic Acids Research.

Функциональный анализ генов

GENEONTOLOGY Unifying Biology

Displaying only results with P<0.05; click here to display all results

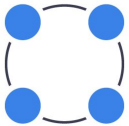
GO biological process complete	Homo sapiens (REF) #	upload_1 (▽ Hierarchy NEW!) #	expected	Fold Enrichment	+/-	P value
T cell activation	225	5	.16	31.11	+	3.33E-03
↳lymphocyte activation	355	5	.25	19.72	+	3.09E-02
↳immune system process	2533	13	1.81	7.19	+	8.14E-07
regulation of T cell activation	325	5	.23	21.54	+	2.01E-02
↳regulation of leukocyte cell-cell adhesion	340	5	.24	20.59	+	2.51E-02
↳regulation of lymphocyte activation	475	6	.34	17.69	+	4.83E-03
↳regulation of leukocyte activation	537	6	.38	15.64	+	9.86E-03
↳regulation of immune system process	1498	8	1.07	7.48	+	2.35E-02
↳regulation of cell activation	573	6	.41	14.66	+	1.44E-02

PCA1

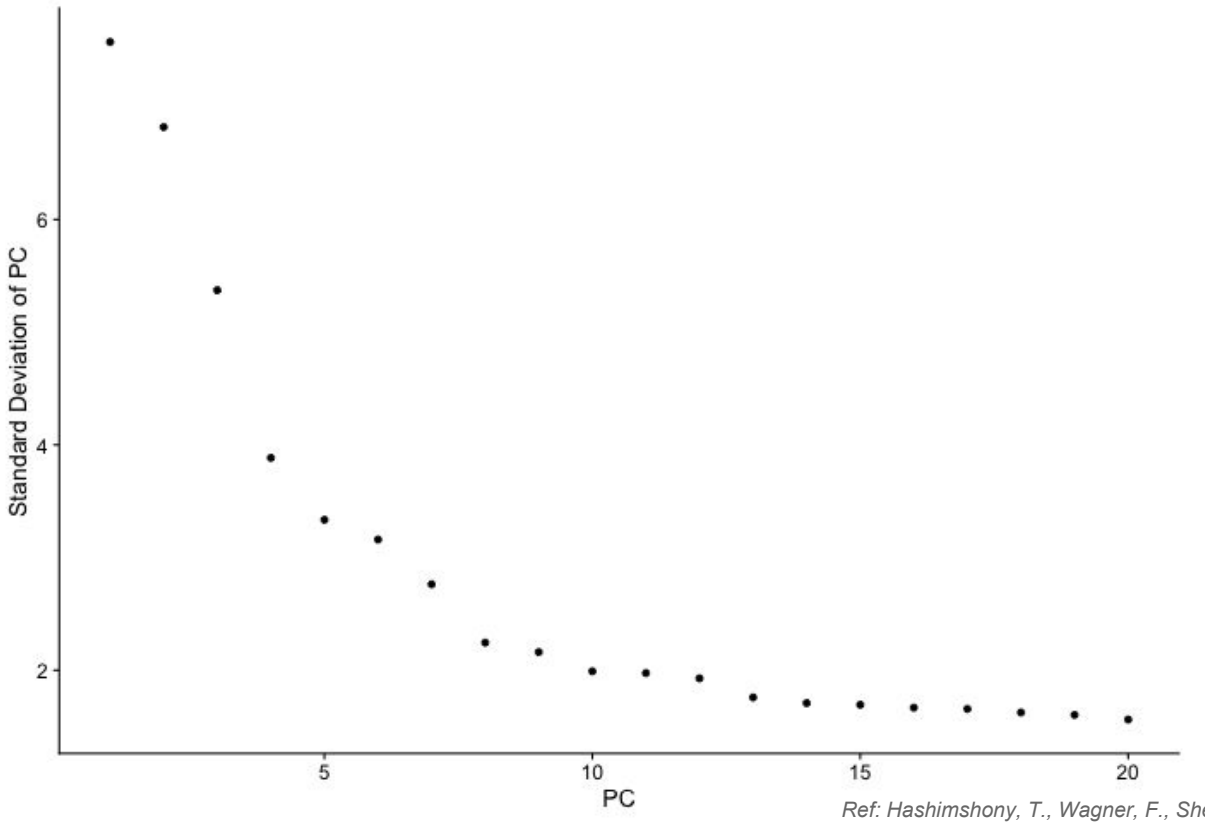
Displaying only results with P<0.05; click here to display all results

GO biological process complete	Homo sapiens (REF) #	upload_1 (▽ Hierarchy NEW!) #	expected	Fold Enrichment	+/-	P value
neutrophil degranulation	482	8	.34	23.24	+	3.69E-06
↳neutrophil mediated immunity	496	8	.35	22.58	+	4.62E-06
↳myeloid leukocyte mediated immunity	517	8	.37	21.67	+	6.39E-06
↳leukocyte mediated immunity	732	8	.52	15.30	+	9.67E-05
↳immune effector process	1017	8	.73	11.01	+	1.23E-03
↳immune system process	2533	10	1.81	5.53	+	9.42E-03

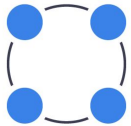
PCA2



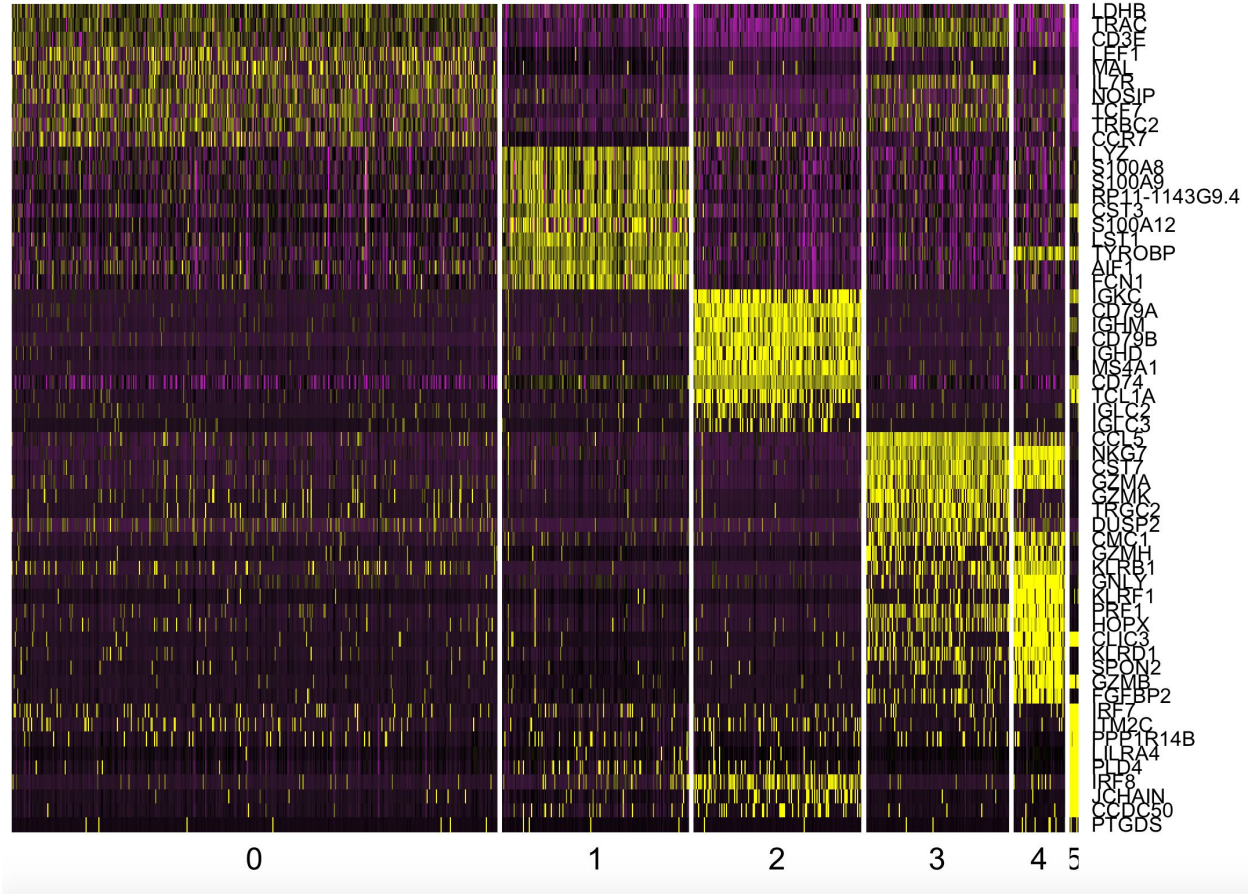
Выбор числа компонент

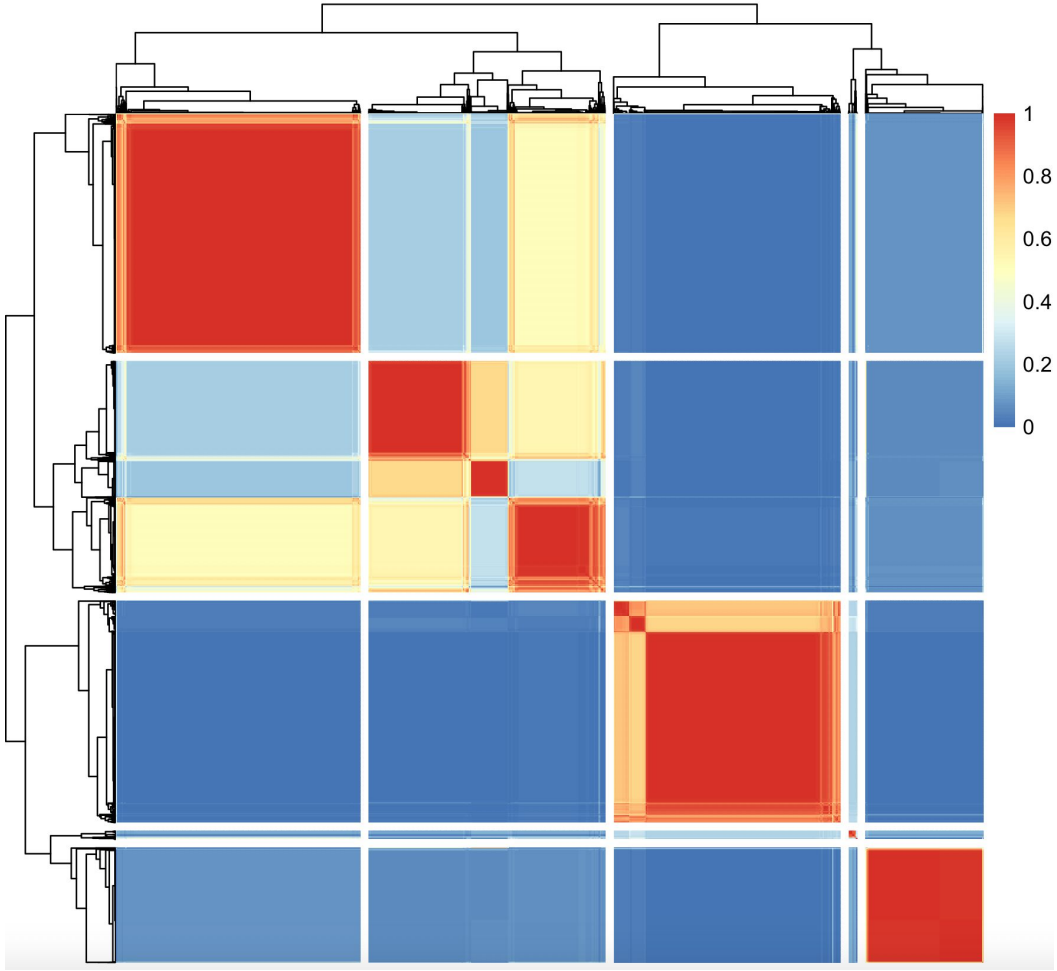


Ref: Hashimshony, T., Wagner, F., Sher, N., & Yanai, I. (2012). CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell reports, 2(3), 666-673.



Анализ главных компонент (PCA)

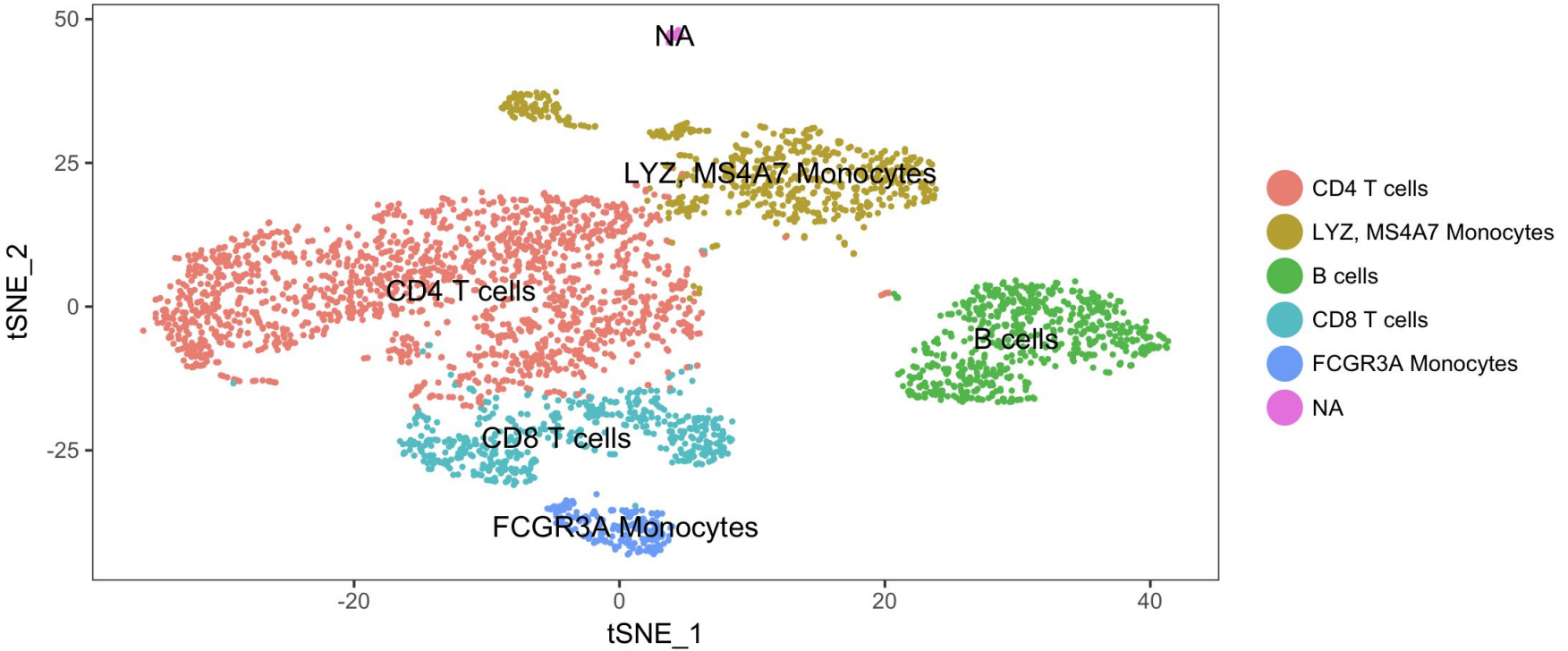




Матрица согласованности

SC3 package manual
https://bioconductor.org/packages/devel/bioc/vignettes/SC3/inst/doc/my-vignette.html

Shared nearest neighbor (SNN) + t-SNE с аннотациями кластеров



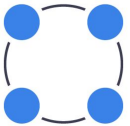
Hoek, K. L., Samir, P., Howard, L. M., Niu, X., Prasad, N., Galassie, A., ... & Shyr, Y. (2015). A cell-based systems biology assessment of human blood to monitor immune responses after influenza vaccination. PloS one, 10(2), e0118528.

Анализ функционального состава кластера B клеток

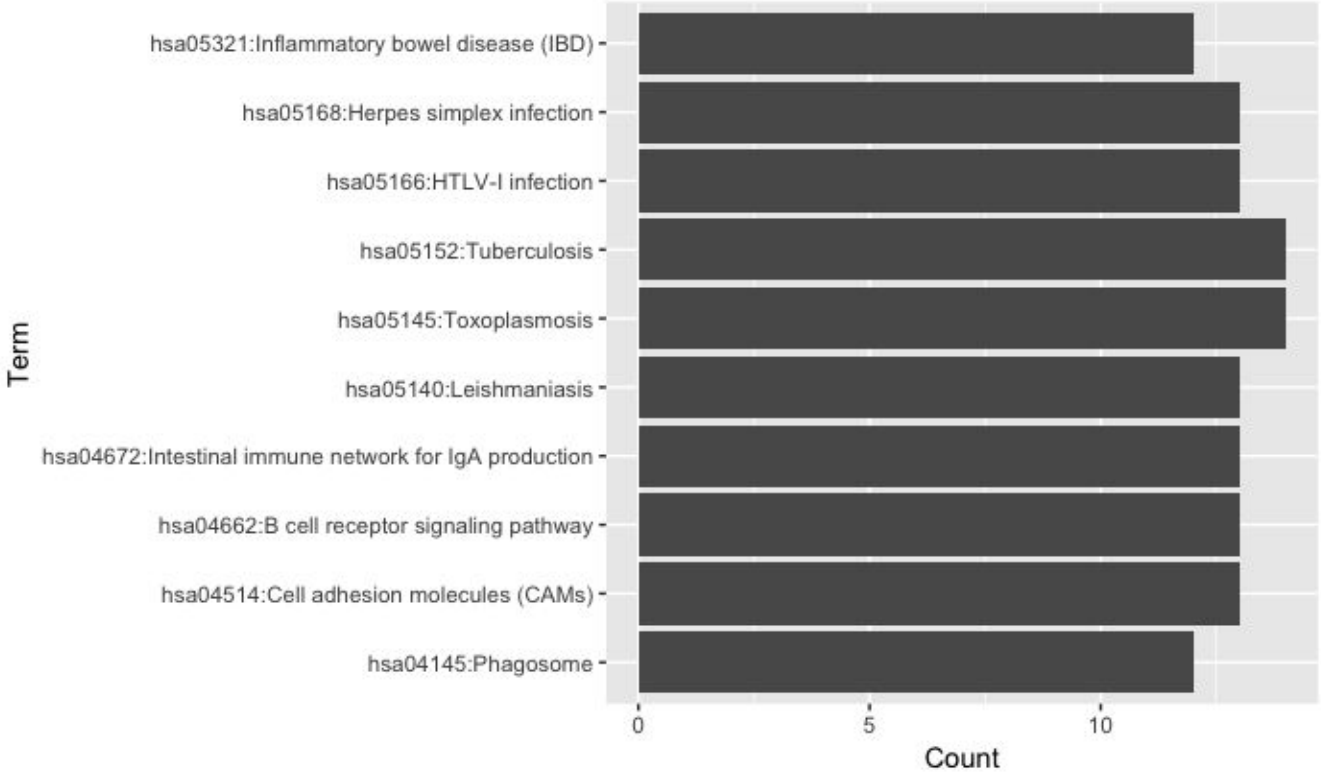
GENEONTOLOGY
Unifying Biology

Displaying only results with P<0.05; click here to display all results

GO cellular component complete	Homo sapiens (REF) #	upload_1 (▽ Hierarchy NEW!) #	expected	Fold Enrichment	+/-	P value
B cell receptor complex	3	3	.03	> 100	+	4.02E-03
↳plasma membrane part	2689	52	23.81	2.18	+	4.06E-05
↳plasma membrane	5364	79	47.51	1.66	+	4.96E-04
↳cell part	16712	172	148.01	1.16	+	1.49E-03
↳cell	16739	172	148.25	1.16	+	1.80E-03
↳membrane	9532	113	84.42	1.34	+	2.42E-02
↳cell periphery	5469	79	48.44	1.63	+	1.14E-03
↳immunoglobulin complex	64	7	.57	12.35	+	2.75E-03
↳protein complex	3211	61	28.44	2.15	+	2.90E-06
↳plasma membrane protein complex	245	18	2.17	8.30	+	1.57E-08
↳membrane protein complex	684	24	6.06	3.96	+	1.62E-05

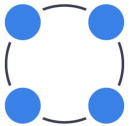


Определение ассоциированности с заболеваниями

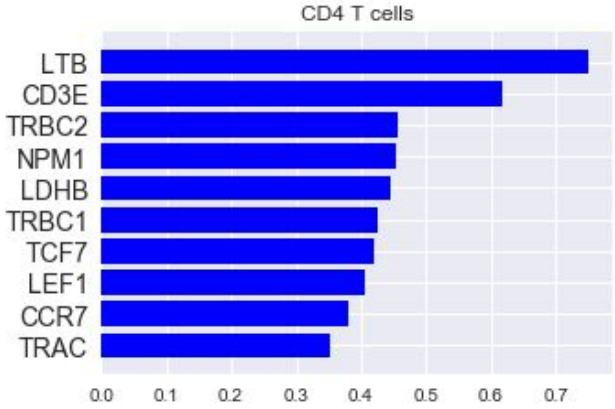


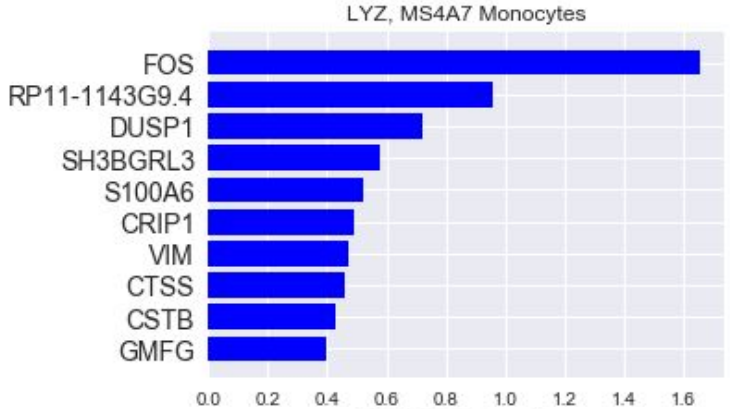
Результаты классификации

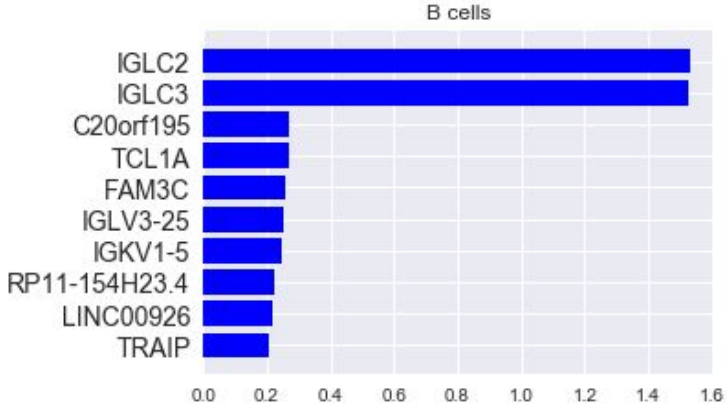
Cluster	precision	recall	f1-score	support
CD4 T cells	0.94	0.91	0.92	463
LYZ, MS4A7 Monocytes	0.84	0.91	0.87	509
B cells	0.71	0.62	0.66	52
FCGR3A Monocytes	0.78	0.73	0.75	278
avg	**0.86**	**0.86**	**0.86**	**1302**

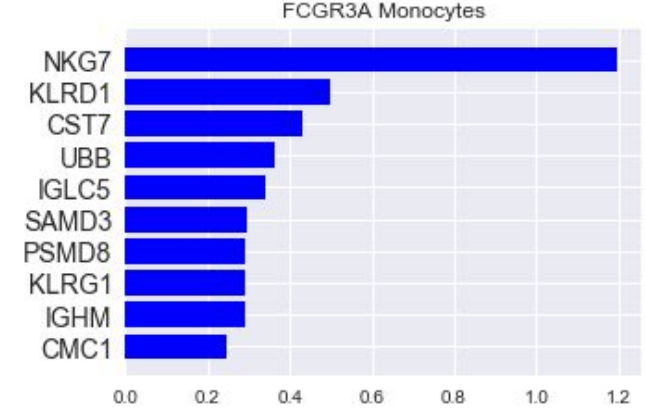


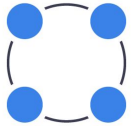
Оценка влияния признаков



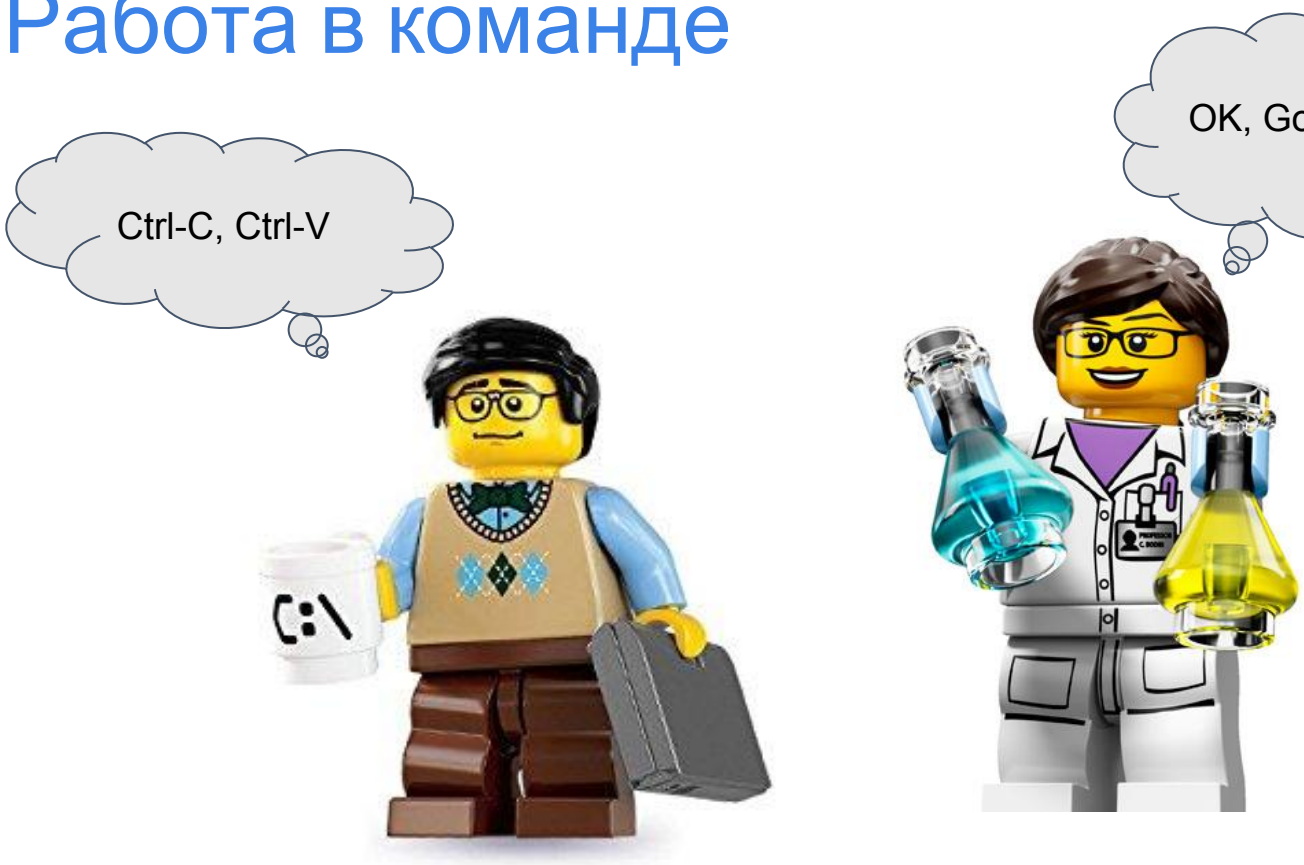








Работа в команде



Спасибо LEGO за наше счастливое детство и картинки на этом слайде.

Спасибо за внимание!

ИНСТИТУТ
БИОИНФОРМАТИКИ