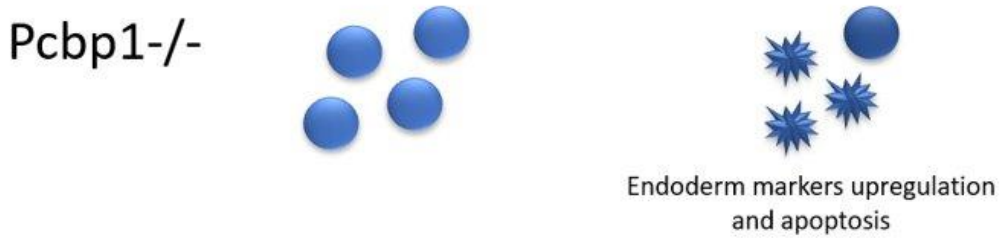
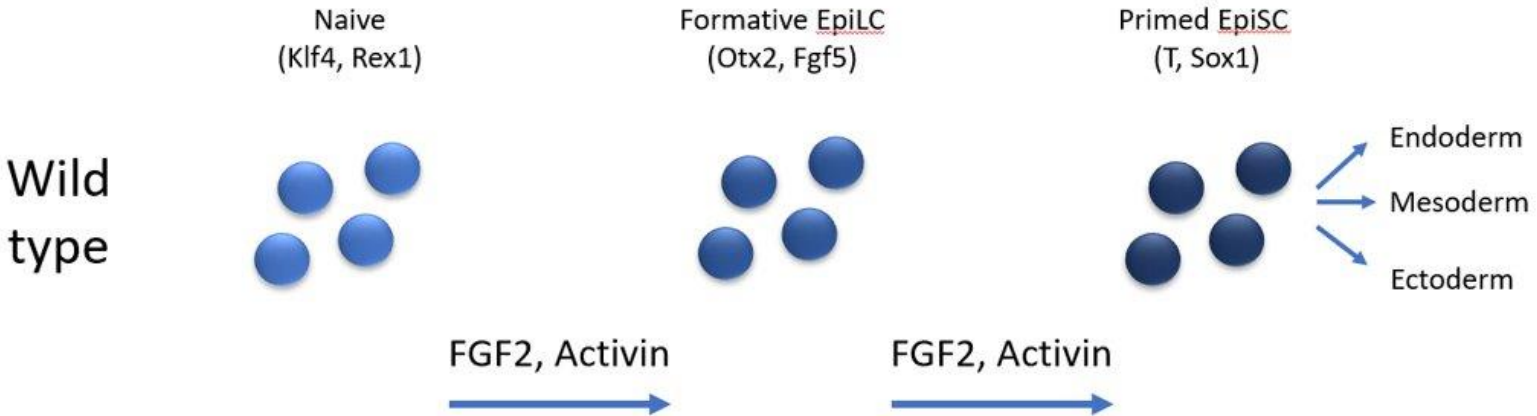


Обработка данных РНК-секвенирования с целью изучения функций белка Rcbp1

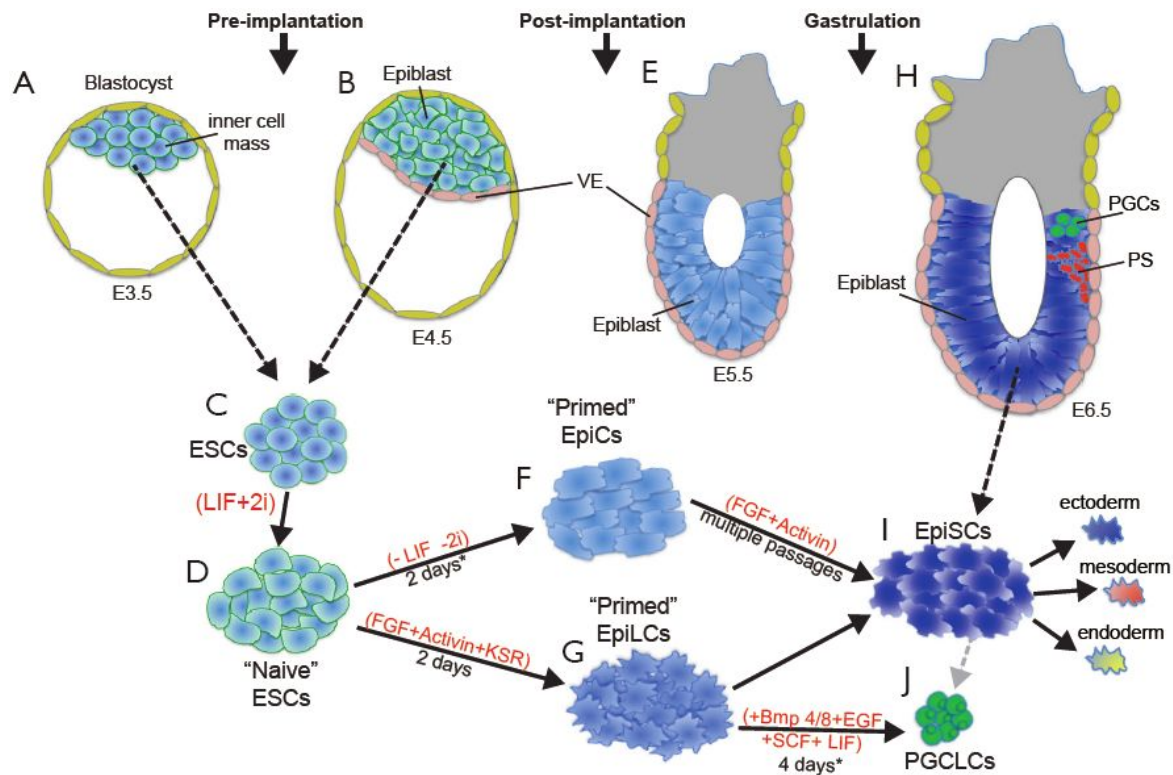
Потапенко Евгений

Оксана Станевич, Евгений Бакин
Институт биоинформатики

Проблема и суть задачи



Проблема и суть задачи



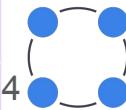
Объект / Данные

Объект исследования: эмбриональные стволовые клетки (ЭСК) мыши в исходном(Native) и праймированном (EpiLC) плюрипотентных состояниях с нокаутом по гену Pcbp1 и без.

Данные: 12 наборов SE ридов длиной ~ 100 п.н полученные на секвенаторе Illumina NovaSeq 6000 SP.

- ЭСК в исходном состоянии
- ЭСК в исходном состоянии с нокаутом Pcbp1
- ЭСК в праймированном состоянии
- ЭСК в праймированном состоянии с нокаутом Pcbp1

Sample	Work_number	Number of reads
S1 Native	1	42 380 298
S2 Native	2	41 535 143
S3 Native	3	50 376 749
P1-1 Native	4	45 108 795
P1-3 Native	5	47 062 337
P1-22Native	6	43 063 892
S1 Epi LC	7	50 987 437
S2 Epi LC	8	35 868 628
S3 Epi LC	9	39 986 664
P1-1 Epi LC	10	55 765 949
P1-3Epi LC	11	47 364 318
P1-22Epi LC	12	46 110 750



Цель проекта

Изучение профиля экспрессии генов эмбриональных стволовых клеток мыши при нокаутировании гена *Pcbp1*.



Задачи и методы

T O O L S

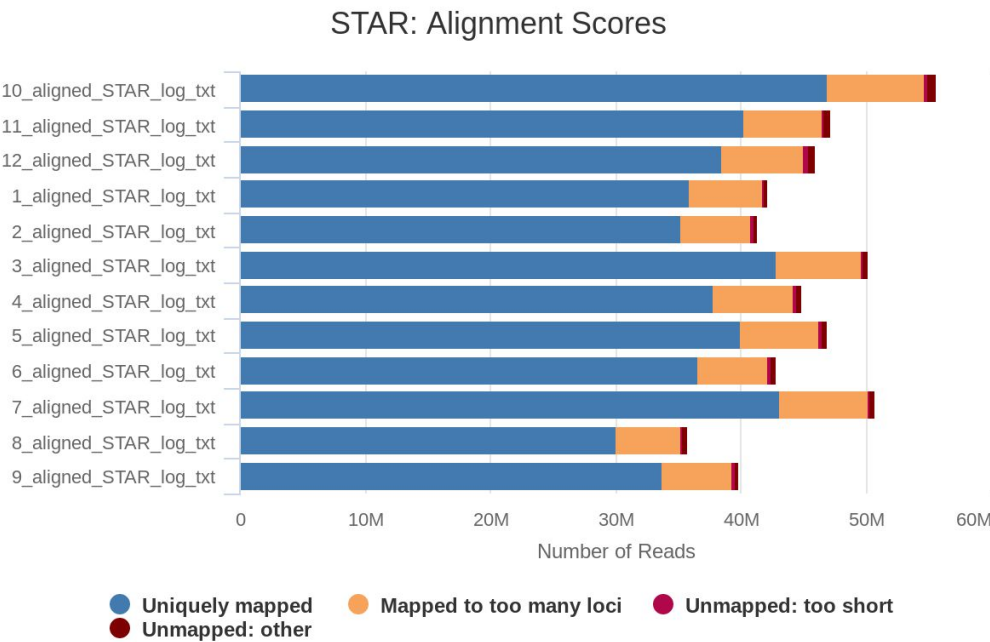
- 1) Анализ литературы по теме
- 2) Процессинг и оценка качества ридов
- 3) Выравнивание на референсный геном
- 4) Визуализация полученных данных
- 5) Дифференциальный анализ экспрессии генов
- 6) Анализ полученных результатов



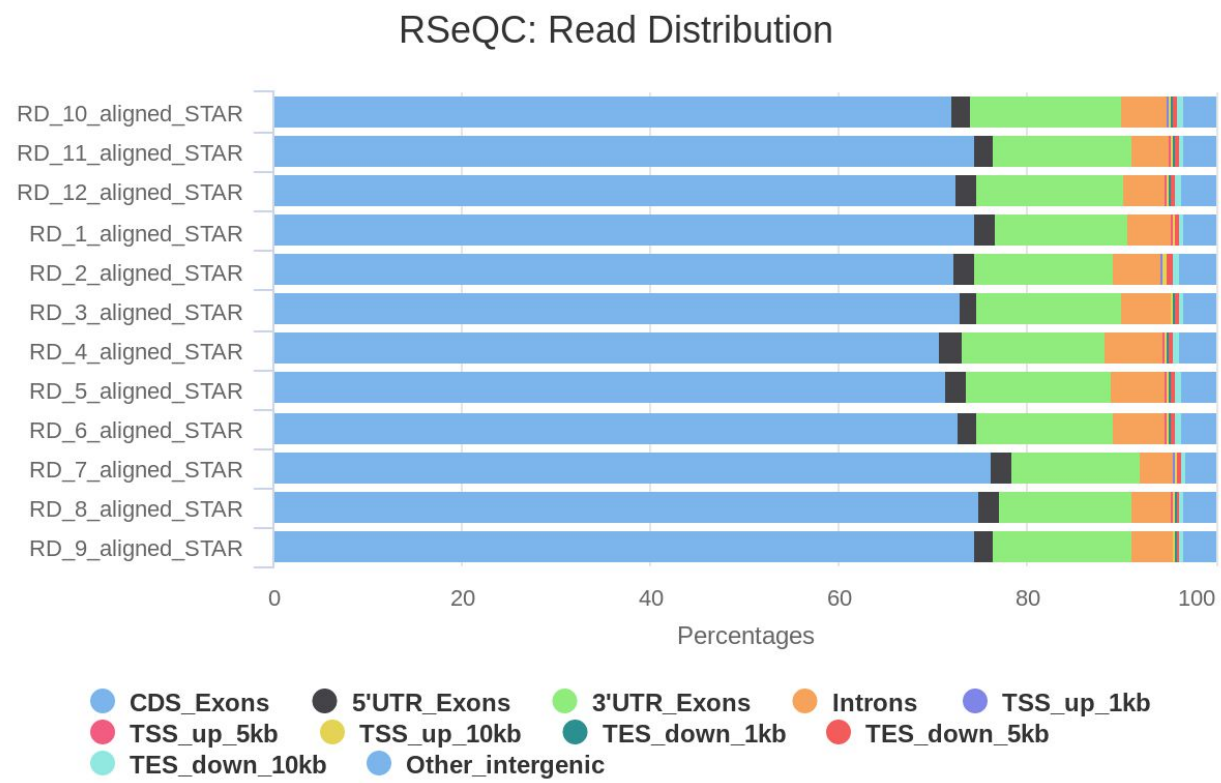
FASTQC
Trimmomatic
STAR
RseQC
featureCounts
DESeq2, edgeR, limma-voom
GO
KEGG

Reads and mapping quality

	% Dups	% GC	M Seqs
1	71.0%	50%	42.2
2	69.7%	50%	41.4
3	71.8%	49%	50.2
4	69.3%	49%	44.9
5	70.0%	49%	46.9
6	70.3%	50%	42.9
7	74.6%	51%	50.8
8	70.4%	50%	35.7
9	72.1%	50%	39.9
10	74.2%	49%	55.6
11	72.3%	50%	47.2
12	71.4%	49%	45.9

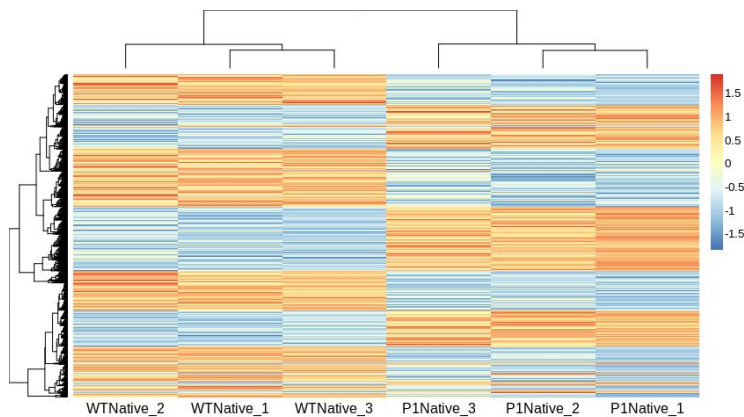
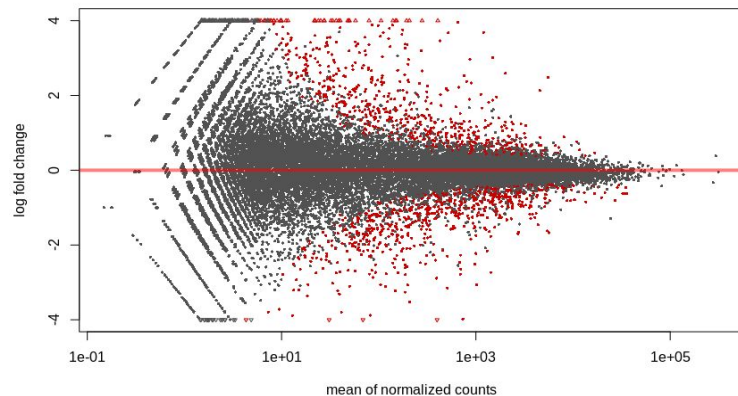


Distribution of reads

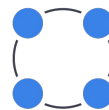
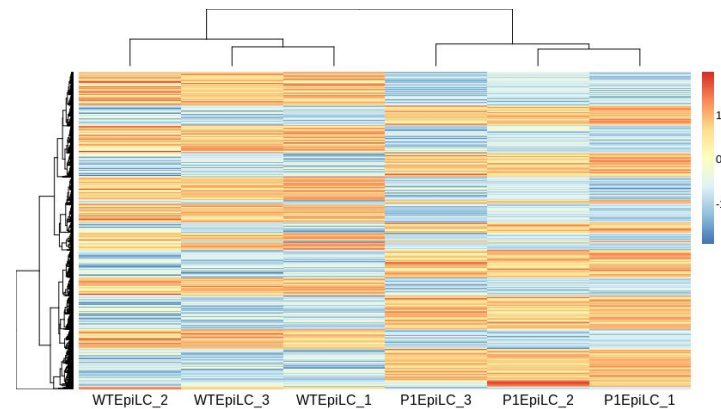
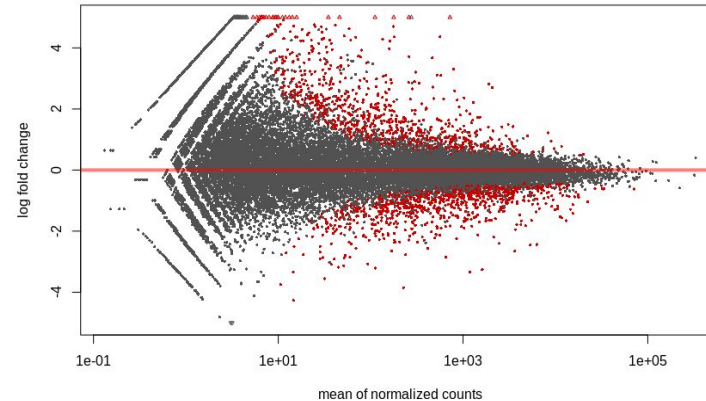


DGE analysis

Naive Pcbp1 KO

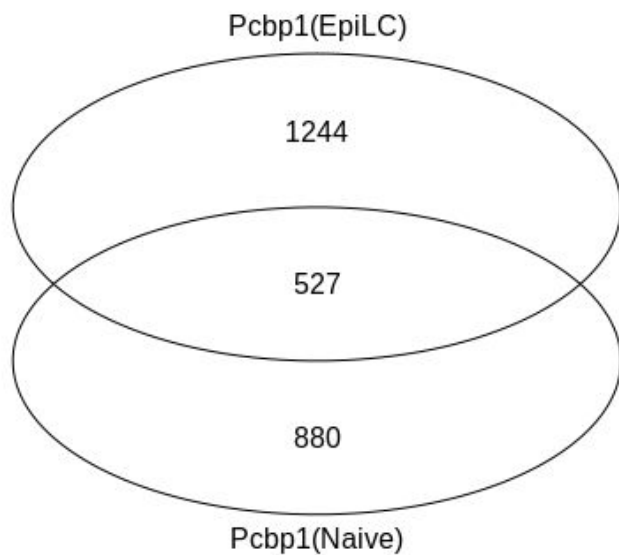


EpiLC Pcbp1 KO

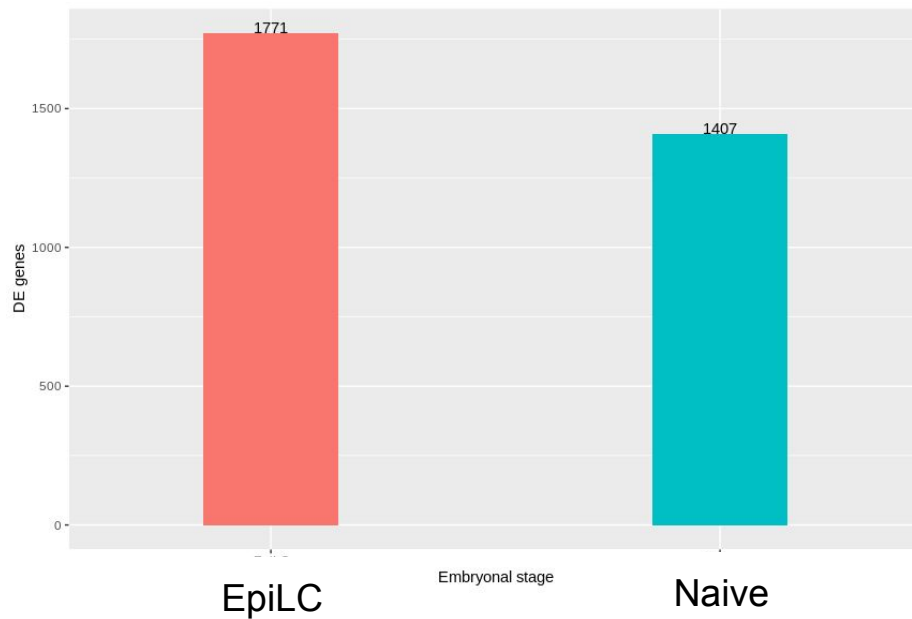


DGE analysis

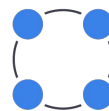
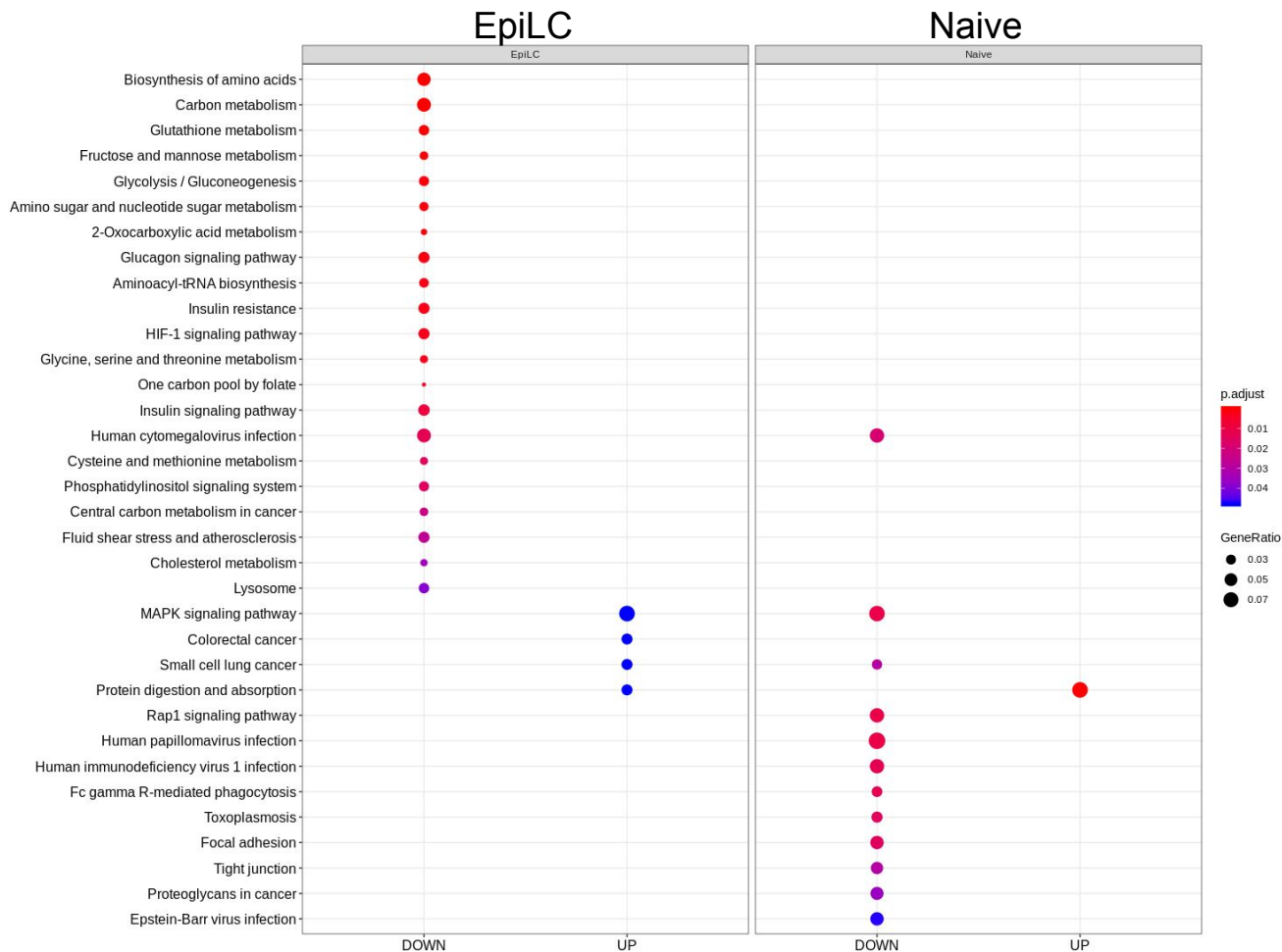
Overlapping of DE genes



Number of DE genes



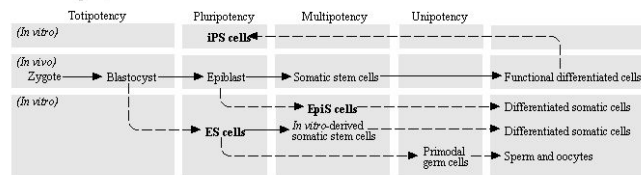
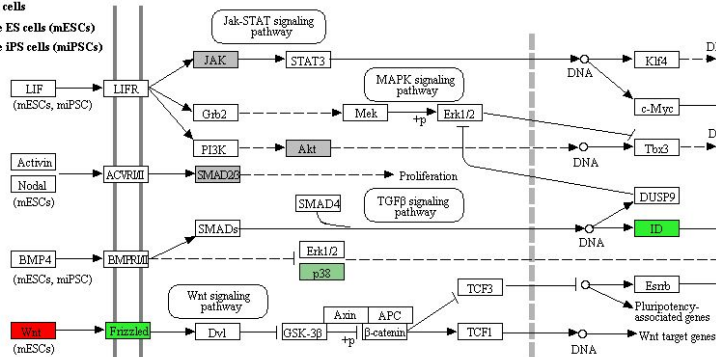
KEGG Pathway Enrichment Analysis of PCBP1



SIGNALING PATHWAYS REGULATING PLURIPOTENCY OF STEM CELLS

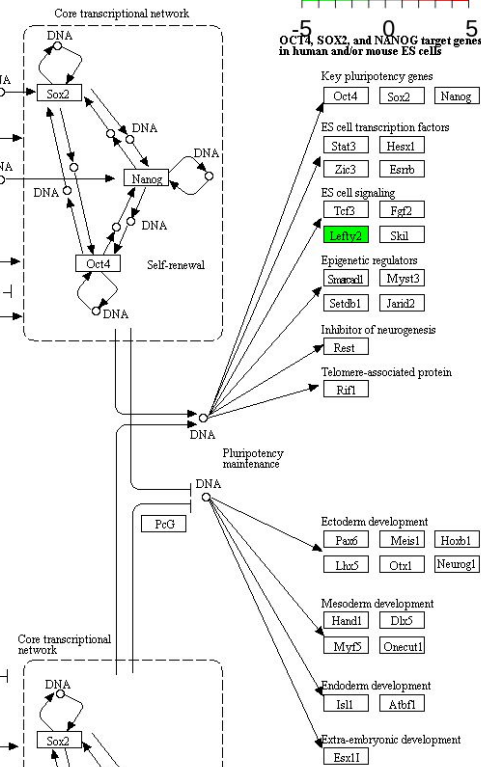
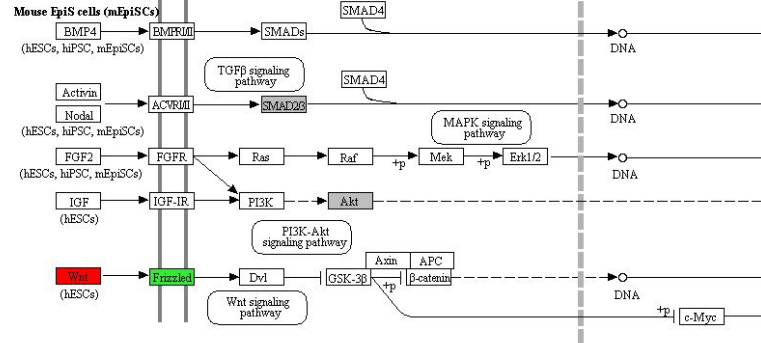
Naive stem cells

Mouse ES cells (mESCs)
Mouse iPS cells (miPSCs)

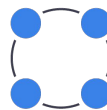


Primed stem cells

Human ES cells (hESCs)
Human iPS cells (hiPSCs)
Mouse EpiS cells (mEpiSCs)



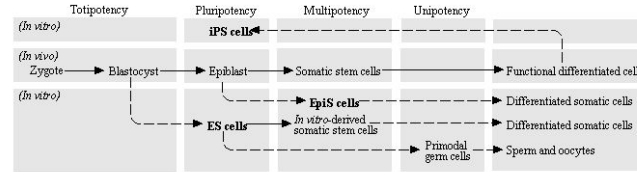
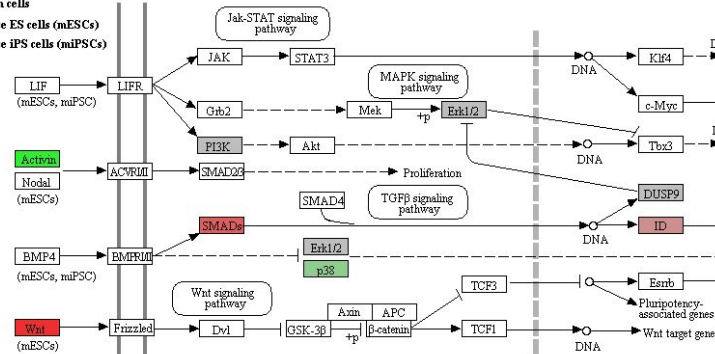
Stage: Naive
State: Pcbp1 KO



SIGNALING PATHWAYS REGULATING PLURIPOTENCY OF STEM CELLS

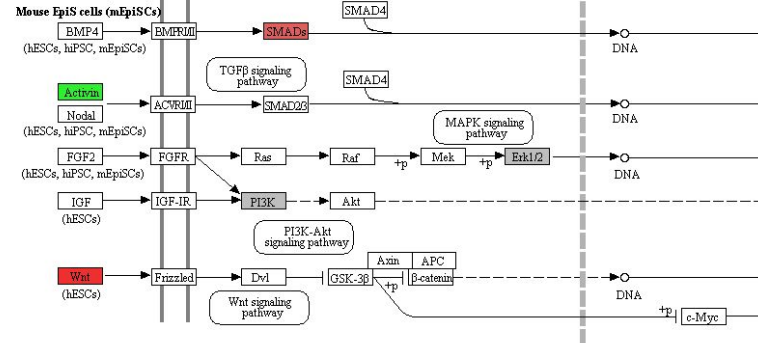
Naïve stem cells

Mouse ES cells (mESCs)
Mouse iPS cells (miPSCs)

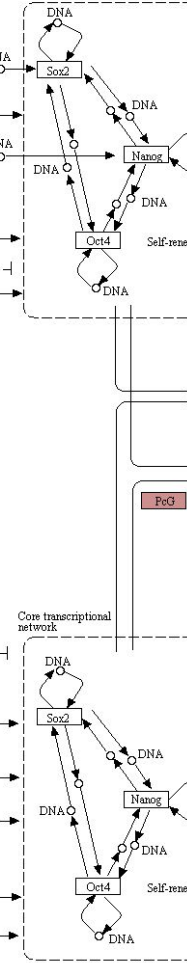


Primed stem cells

Human ES cells (hESCs)
Human iPS cells (hiPSCs)
Mouse EpiS cells (mEpiSCs)



Core transcriptional network

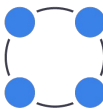


5 OCT4, SOX2, and NANOG target genes in human and/or mouse ES cells

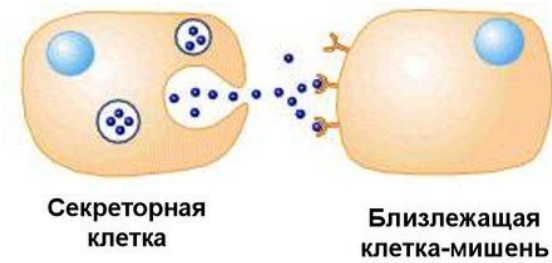
Key pluripotency genes
Oct4, Sox2, Nanog
ES cell transcription factors
Stat3, Hex1
Zic3, Esrrb
ES cell signaling
Tcf3, Fgf2
Lefty2, Ski
Epigenetic regulators
Smad3, Mst13
Setd1, Jarid2
Inhibitor of neurogenesis
Rest
Telomere-associated protein
Raf1

Phenotypic maintenance
PcG
Ectoderm development
Pax6, Meis1, Hoxb1
Lhx5, Otx1, Neurog1
Mesoderm development
Hand1, Dlx5
Myf5, Onecl1
Endoderm development
Isl1, Atf1
Extra-embryonic development
Esrr1

Stage: EpiLC
State: Pcbp1 KO



Defined genes of interest:



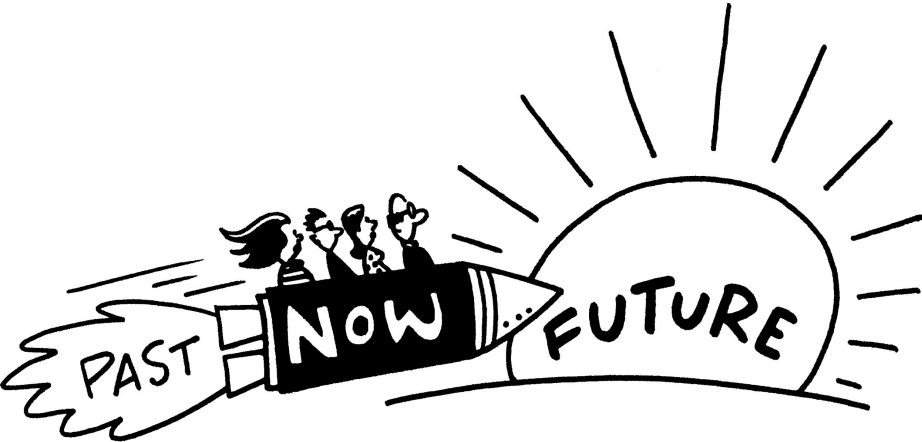
Паракринные влияния

SYMBOL	GENENAME	PATHWAY/FUNCTION
Wnt5a and Wnt9a	Ligands in Wnt-signaling pathway.	Wnt signaling
Inhbe	inhibin beta-E	TGF- β signaling
Cnpy1	canopy FGF signaling regulator 1	FGF signaling
Cnpy2	canopy FGF signaling regulator 2	



Дальнейшие планы:

- 1) Анализ имеющихся в общем доступе данных по интересующим сигнальным путям и генам интереса.
- 2) Сопоставления данных RNA-seq и ChIP-seq для определения генов потенциально регулируемых на транскрипционном уровне.



Github репозиторий проекта:

→ https://github.com/PotapenkoEugene/Bioinf_project_fall_2019



GitHub

**THANK YOU FOR YOUR
ATTENTION!**



ANY QUESTIONS?