

## Комментарии к тестовому заданию

### Слайд №7:

Касательно strand. Походу, в задании вопрос с подвохом. Ссылка: <https://www.biostars.org/p/334485/> . Суть в том, что ChIP-Seq не стрэнд-специфичен. Цитата: There is no strand information, as ChIP-seq binding events are not strand-specific. DNA-binding factors and DNA-associated proteins, such as histones bind/are associated with double-stranded DNA.

Т.е. сделать с учетом цепи не получится в принципе. MACS2 не дает такую информацию.

По поводу qValue: как я понимаю, это pvalue после FDR correction, что делает отбор по нему даже лучше.

### Слайд №8:

Важно учесть, что qvalue в форме  $-\log(10)$ . Цитата отсюда: <https://www.biorxiv.org/content/biorxiv/suppl/2017/04/10/120113.DC2/120113-1.pdf> : A peak call with a larger signal value has a smaller p- and q-value, which indicates that it is more likely to reflect an actual protein-DNA binding event. Thus, a larger signal value will translate to a larger  $-\log_{10}(\text{p-value})$  and  $-\log_{10}(\text{q-value})$ .

Исправлено в новой версии. Также для отбора пиков я сортировал не только по qvalue, но и при равенстве qValue отбирал максимальные по signalValue, и при равенстве уже этого параметра отбирал по максимуму score.

### Слайд №10:

Мотивы здесь и ниже перечислены в том же порядке, что и в итоговом файле.

Первый же мотив однозначно указывает на YY1. Успех!

Второй мотив относится к Ets-related factors (о них ниже).

Следующий мотив (TFE3) это Transcription Factor For Immunoglobulin Heavy-Chain Enhancer 3 (подтверждается связь с регуляцией гуморального иммунитета) (<https://www.uniprot.org/uniprot/P19532>). В GO числится ДНК-связывающая активность, взаимодействие с РНК-полимеразой 2, что пересекается с предыдущими результатами.

Мотив №4 относится к homeo domain factors (связь с Homeobox, а значит с морфогенезом) (<http://jaspar.genereg.net/matrix/MA0063.2/?revcomp=1>)

Мотив №5 относится к гену MYBL2 (<https://www.genecards.org/cgi-bin/carddisp.pl?gene=MYBL2>). В описании значится регуляция клеточного деления и выживания.

Бегло просмотрев остальные мотивы, можно утверждать следующее: YY1 имеет ДНК-связывающую активность, взаимодействует с РНК-полимеразой 2, регулирует тем или иным образом В-клеточный иммунитет, клеточное деление, морфогенез.

Слайд №11:

Обнаружено совпадение de novo предсказанного мотива с опубликованными данными с factorbook. Это очень радует!

Слайд №12:

Насколько видно по результатам, найдено очень много взаимодействий с ETS семейством факторов. ETS = Erythroblast Transformation Specific; теперь результаты GREAT более осмысленны. Но вот в чем вопрос лично у меня: значит ли это, что YY1 активирует факторы семейства ETS, которые в свою очередь выполняют свои функции по контролю клеточного цикла/клеточной миграции/апоптоза(цитата с вики([https://en.wikipedia.org/wiki/ETS\\_transcription\\_factor\\_family](https://en.wikipedia.org/wiki/ETS_transcription_factor_family))): All ETS (Erythroblast Transformation Specific) family members are identified through a highly conserved [DNA binding domain](#), the **ETS domain**, which is a winged [helix-turn-helix](#) structure that binds to DNA sites with a central **GGA(A/T)** DNA sequence. As well as DNA-binding functions, evidence suggests that the ETS domain is also involved in [protein-protein interactions](#). There is limited similarity outside the ETS DNA binding domain.), или же мишени YY1 и ETS факторов похожи по структуре?

UPD: нашел древнюю статью про совместную работу ETS и YY1: <https://pubmed.ncbi.nlm.nih.gov/7607503/> Получается, они работают вместе?

UPD2: [https://www.cell.com/cell-stem-cell/pdfExtended/S1934-5909\(19\)30110-9](https://www.cell.com/cell-stem-cell/pdfExtended/S1934-5909(19)30110-9) статья свежая, детально не смотрел, цитата оттуда: Reconstruction of the YY1 transcriptional network in human melanoma revealed that E2F and ETS family members are putative YY1 co-factors. Значит, работают совместно.

Также еще нашлось несколько мотивов связи с Homeobox, что состыковывается с YY1 Uniprot (морфогенез); в принципе логично.

Слайды №14-17:

№14:

По биологическим процессам все совпадает с данными литературы. Например <https://www.nature.com/articles/nature06322> : we identified the transcription factor yin-yang 1 (YY1) as a common target of mTOR and PGC-1 $\alpha$ . Knockdown of YY1 caused a significant decrease in mitochondrial gene expression and in respiration, and YY1 was required for rapamycin-dependent repression of those genes. Moreover, mTOR and raptor interacted with YY1, and inhibition of mTOR resulted in a failure of YY1 to interact with and be coactivated by PGC-1 $\alpha$ .

По поводу связывания с мРНК: хорошая обзорная статья <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7554316/>.

№15:

Локализация совпадает с информацией из YY1 Uniprot: <https://www.uniprot.org/uniprot/P25490>

№16

Функции (гистондеацетилаза прямо как в википедии) тоже, Rho-GTP binding тоже логично (см. [https://en.wikipedia.org/wiki/Rho\\_family\\_of\\_GTPases](https://en.wikipedia.org/wiki/Rho_family_of_GTPases) ).

NADH активность тоже актуальна (<https://pubmed.ncbi.nlm.nih.gov/11955626/>).

О связи с иммунитетом: <https://www.pnas.org/content/pnas/113/27/E3911.full.pdf>

№17

В целом все верно и соответствует литературе. т.е. налицо связь действия YY1 и корректной работы митохондрий, морфогенеза, клеточного движения, развития цитоскелета, В-иммунного ответа. Здорово! Но не хватает времени для более полной интерпретации результатов.

Слайды №18-21

Полные результаты есть на Github. Радует, что результаты HOMER соотносятся с результатами bedtools и пересекаются с GREAT.