

Проект по биоинформатике

Группа ZDNA_H3K4me1

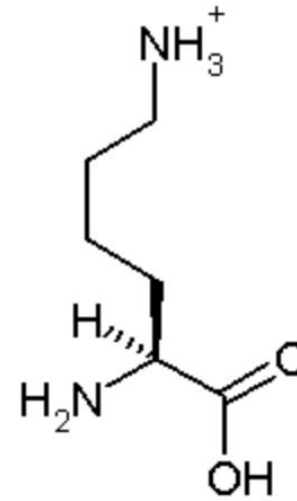
Состав команды: Эмилия Шаймуратова, Тагир Сафин, Андрей Рябинин, Дарья Барановская, Александр Залялов, Семен Кудрявцев, Михаил Фридман

Введение

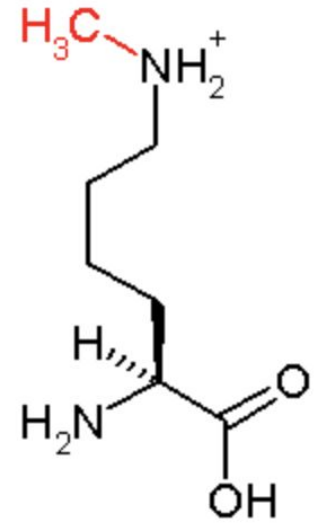
H3K4me1 представляет собой эпигенетическую модификацию белка упаковки ДНК гистона H3. Это метка, указывает на метилирование по лизину 4 и часто связана с энхансерами генов.

Функции: H3K4me1 обогащен активными и примированными энхансерами. Энхансеры транскрипции контролируют экспрессию гена идентичности клетки.

Известно о связи метки и комплекса BAF.



Lysine



Mono-methyl
lysine

Обзор исходных файлов

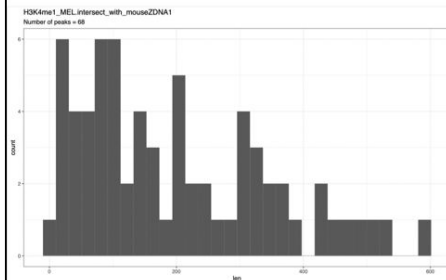
| Фамилия | Организм | Тип клетки | файл 1 | файл 2 |
|-------------|--------------|------------|-------------|-------------|
| Шаймуратова | mouse (mm10) | MEL | ENCFF267GXF | ENCFF086CHI |
| Сафин | mouse (mm10) | ES-E14 | ENCFF158GBZ | ENCFF147SYC |
| Рябинин | human (hg19) | HeLa-S3 | ENCFF572NWJ | ENCFF952QRP |
| Барановская | mouse (mm10) | MEL | ENCFF563AUO | ENCFF687RAP |
| Залялов | human (hg19) | K562 | ENCFF540NGG | ENCFF759NWD |
| Кудрявцев | mouse (mm10) | CH12.LX | ENCFF596ORE | ENCFF791SEO |
| Фридман | human (hg19) | H1 | ENCSR000ANA | ENCSR631RJR |

Обзор исходных файлов

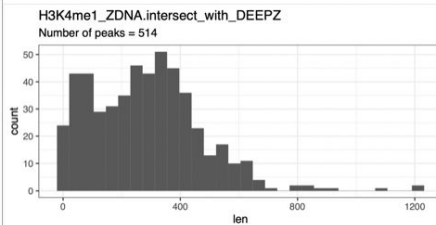
| Фамилия | # пиков в 1 файле | # пиков во 2 файле | # пиков в 1 файле после фильтрации | # пиков во 2 файле после фильтрации | # пиков после объединения | # участков вторичной структуры | #пересечений |
|--------------------------|-------------------|--------------------|------------------------------------|-------------------------------------|---------------------------|--------------------------------|--------------|
| Шаймуратова mouse (mm10) | 88102 | 60235 | 87623 | 60194 | 102857 | 3258 | 95 |
| Сафин mouse (mm10) | 131687 | 48495 | 129427 | 48266 | 153999 | 3258 | 197 |
| Рябинин human (hg19) | 97270 | 91178 | 96123 | 90031 | 103555 | 19394 | 2216 |
| Барановская mouse (mm10) | 62370 | 50333 | 62370 | 50333 | 84640 | 3258 | 68 |
| Залялов human (hg19) | 112834 | 116407 | 112145 | 115855 | 167644 | 412600 | 34381 |
| Кудрявцев mouse (mm10) | 170147 | 96956 | 168577 | 96955 | 189840 | 3258 | 514 |
| Фридман human (hg19) | 68550 | 135323 | 68379 | 134867 | 141063 | 19394 | 1953 |

Гистограммы пересечений hg19

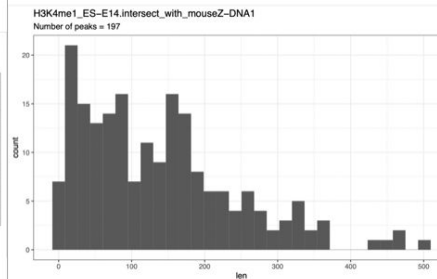
mouse (mm10)
MEL



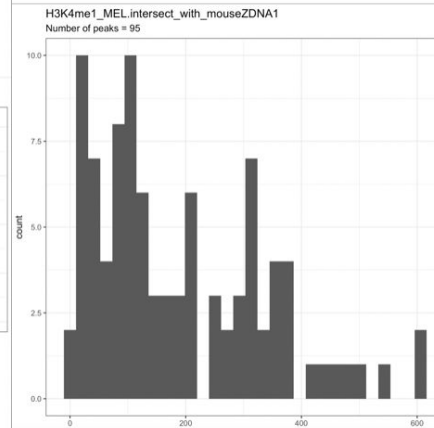
mouse (mm10) CH12.LX



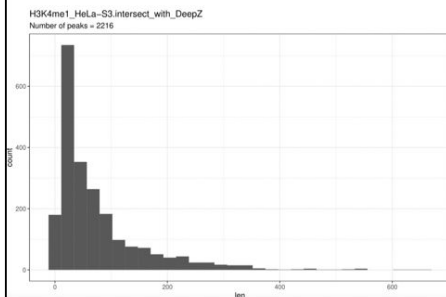
mouse (mm10) ES-E14



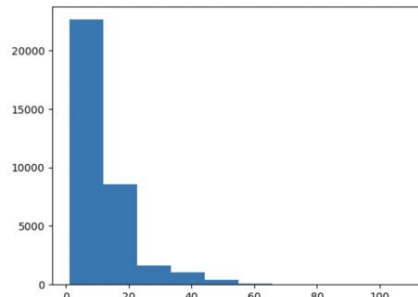
mouse (mm10)
MEL



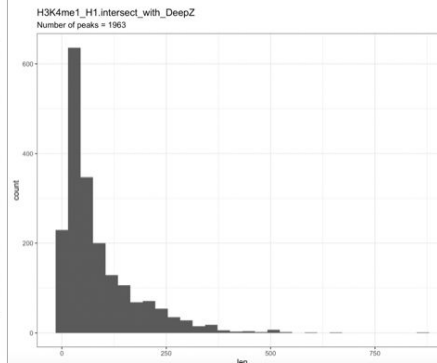
human (hg19) HeLa-S3



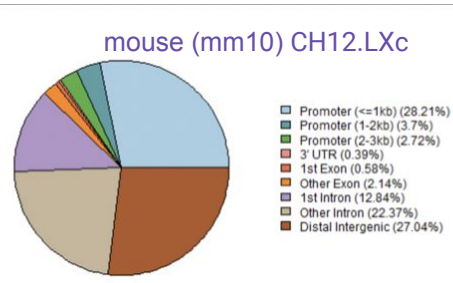
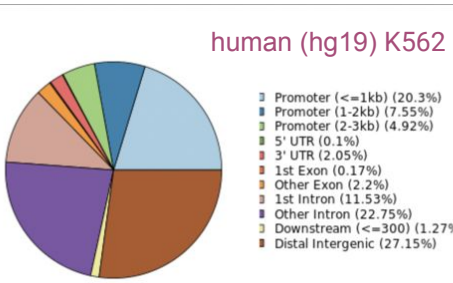
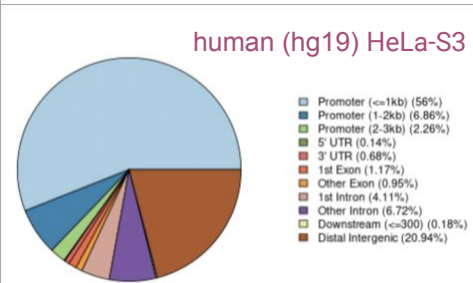
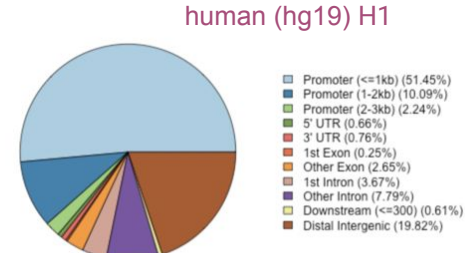
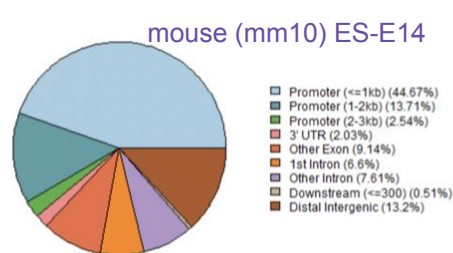
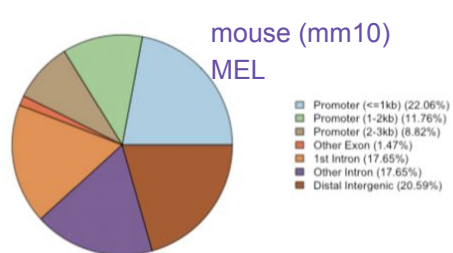
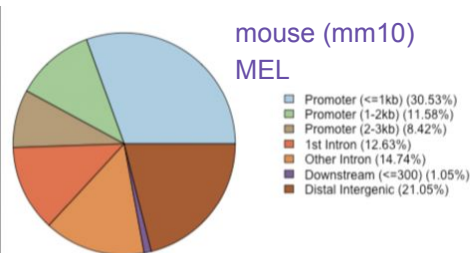
human (hg19) K562



human (hg19) H1



Круговые диаграммы пересечений



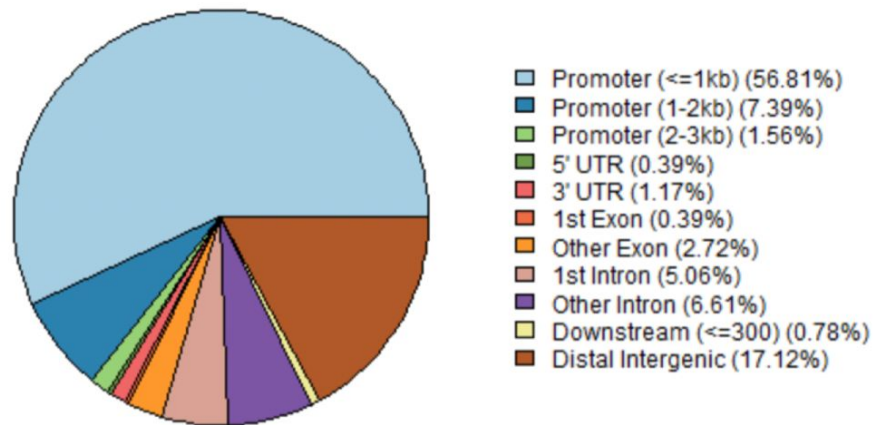
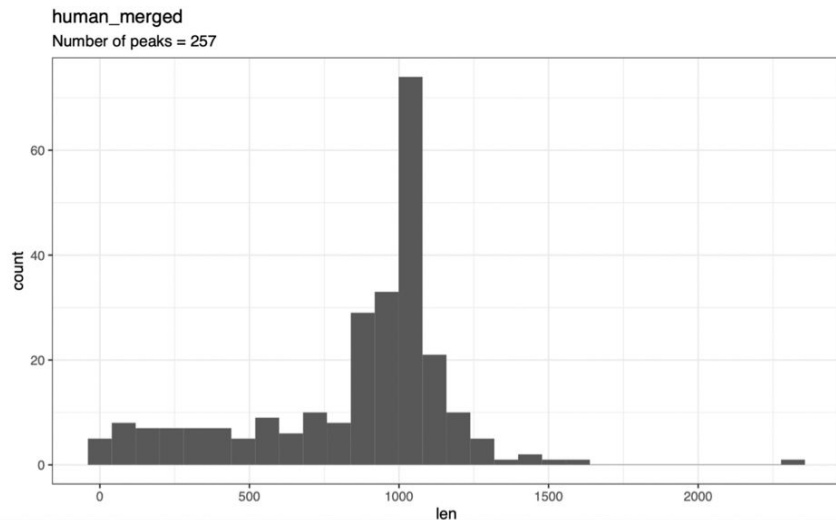
Пересечения intersect-файлов

| Организм и тип клетки | #пересечений у мыши | #пересечений после перевода |
|-----------------------------------|------------------------|-----------------------------------|
| mouse (mm10) MEL (Шаймуратова) | 95 | 22 |
| mouse (mm10) ES-E14 | 197 | 115 |
| mouse (mm10) MEL (Барановская) | 68 | 17 |
| mouse (mm10) CH12.LX | 514 | 122 |
| human (hg19) HeLa-S3 | - | 2216 |
| human (hg19) K562 | - | 34381 |
| human (hg19) H1 | - | 1953 |

| | mouse (mm10) MEL (Барановская) | mouse (mm10) MEL (Шаймуратова) | mouse (mm10) CH12.LX | mouse (mm10) ES-E14 |
|--|-----------------------------------|-----------------------------------|-------------------------|---------------------|
| пересечение всех людей human (hg19) HeLa-S3, K562, H1 | 0 | 3 | 3 | 0 |

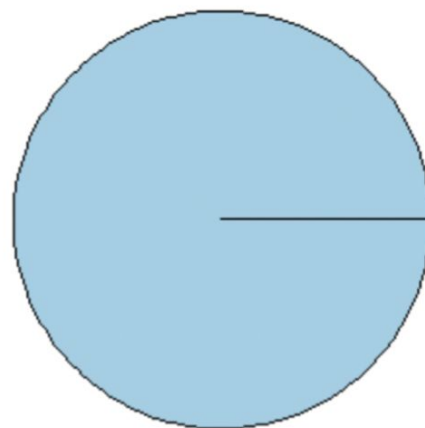
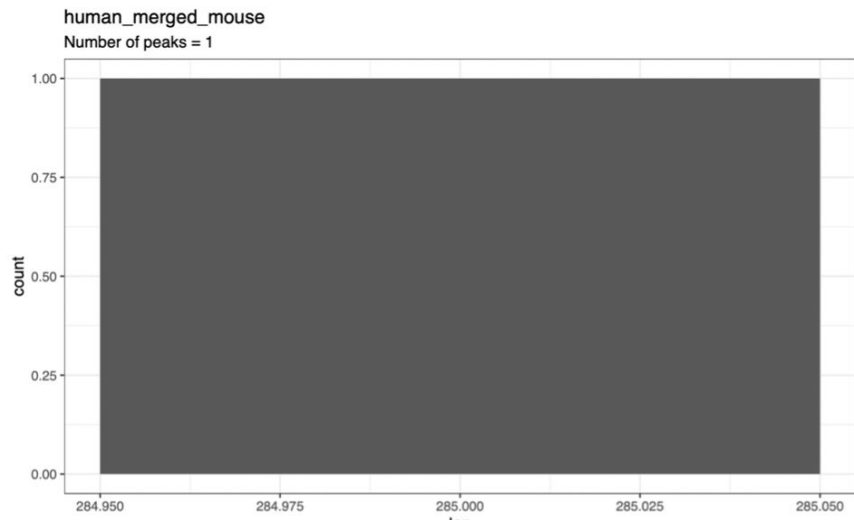
Гистограмма длин участков и круговая диаграмма после пересечения

пересечение всех людей human (hg19) HeLa-S3, K562, H1



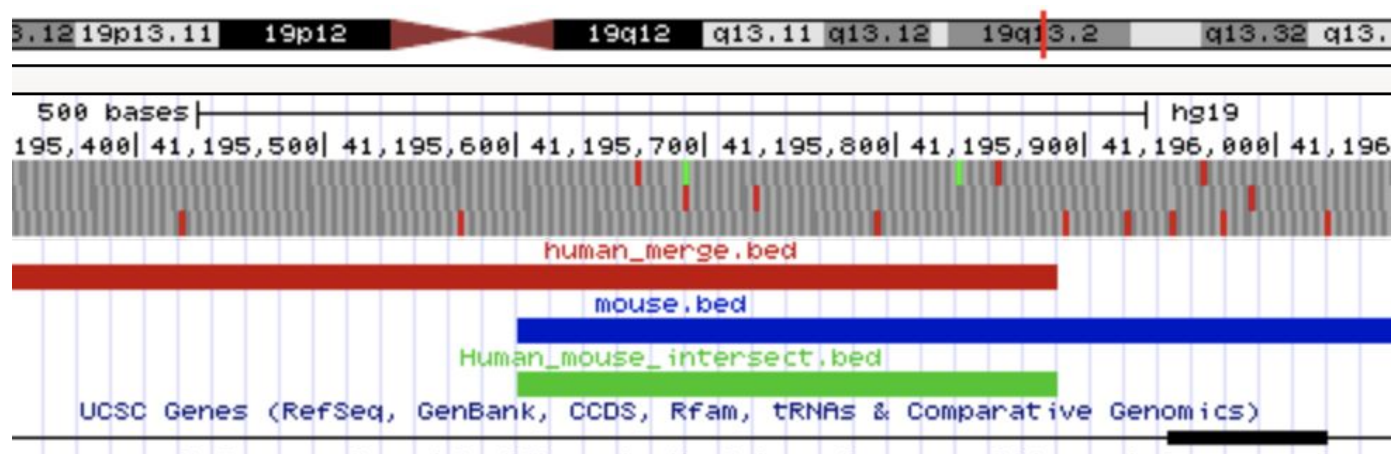
Гистограмма длин участков и круговая диаграмма после пересечения

пересечение всех людей с мышью



■ Promoter ($\leq 1\text{kb}$) (100%)

Локусы пересечения пиков



Результаты GO анализа

Results

| | Reference list | human_merged.genes_uniq.txt |
|-------------------------------|------------------------------------|--------------------------------|
| Uniquely Mapped IDS: | 20595 out of 20595 | 106 out of 107 |
| Unmapped IDs: | 0 | 6 |
| Multiple mapping information: | 0 | 1 |

Export [Table](#) [XML with user input ids](#) [JSON with user input ids](#)

No statistically significant results. [Click to see all results.](#)

Выводы

- Многие пики попадают в межгенное пространство
- Для мыши слишком мало пересечений после перевода в геном человека
- Нельзя сказать о прочной ассоциации гистоновой метки и вторичной структуры ДНК
- Пересечения пиков не лежат в значимых относительно FDR анализа регионах

GitHub: https://github.com/zilzilok/hse21_ZDNA_H3K4me1_group

Статьи

- (Identification of H3K4me1-Associated Proteins at Mammalian Enhancers)
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6007000/>
- Is H3K4me1 at enhancers correlative or causative?(<https://www.nature.com/articles/s41588-017-0018-3>)