

ITMO UNIVERSITY

# Поиск активных модулей в сетях белок-белковых взаимодействий

Газизова Адель, Золотарев Андрей

*Руководители: Никита Алексеев,  
Алексей Сергушичев*

# Дифференциальная экспрессия генов

- Анализ сложный, много данных
- Существует большая вероятность что-то упустить или наоборот посчитать значимым случайную разницу экспрессии
- Необходимы новые автоматизированные методы оценки дифференциальной экспрессии генов
- Один из подходов – совмещение данных дифференциальной экспрессии и белок-белковых взаимодействий

# Задачи

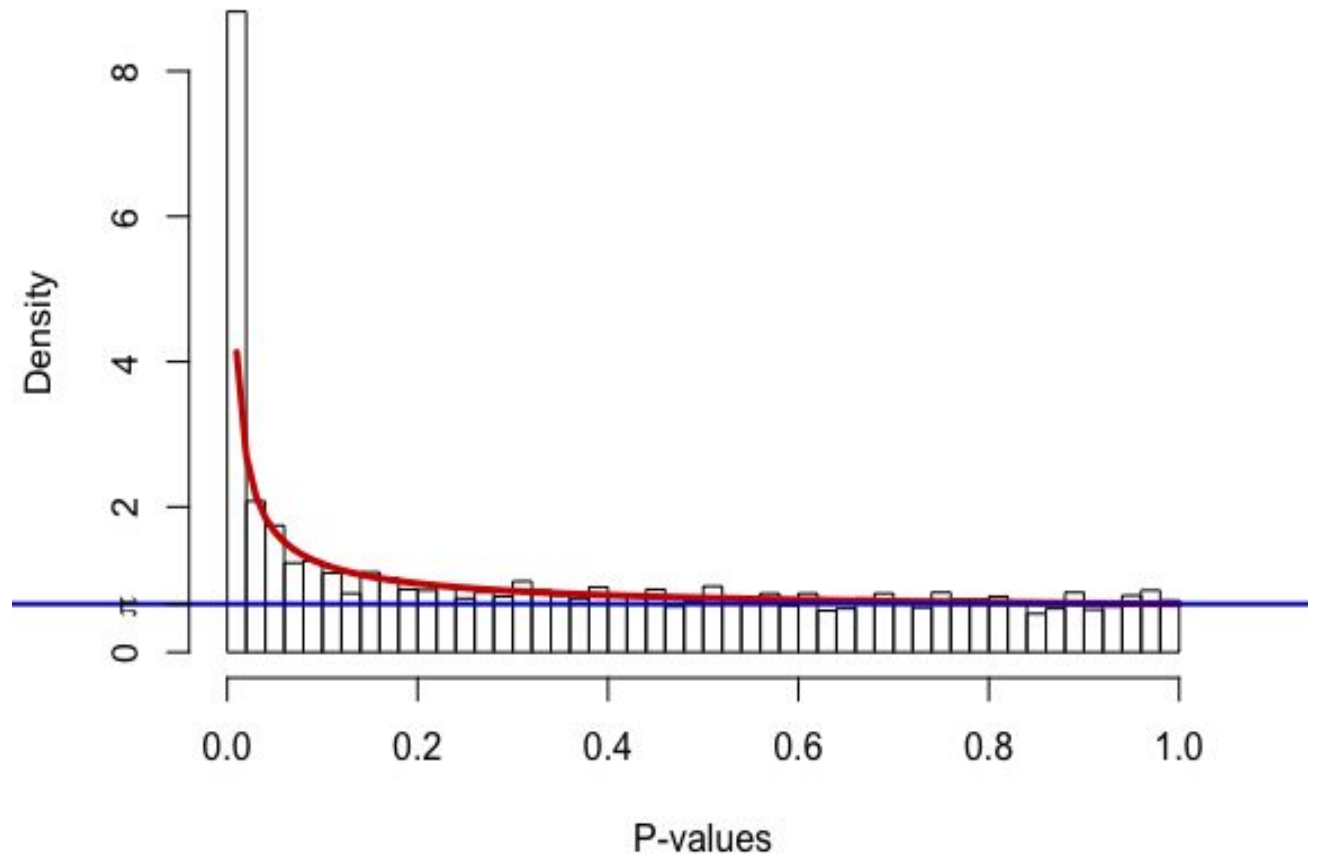
- Научиться генерировать данные дифференциальной экспрессии генов с заранее известным сигналом
- Применить алгоритм на реальных данных и с точки зрения биологической релевантности оценить результаты

# Распределение p-value в реальных данных

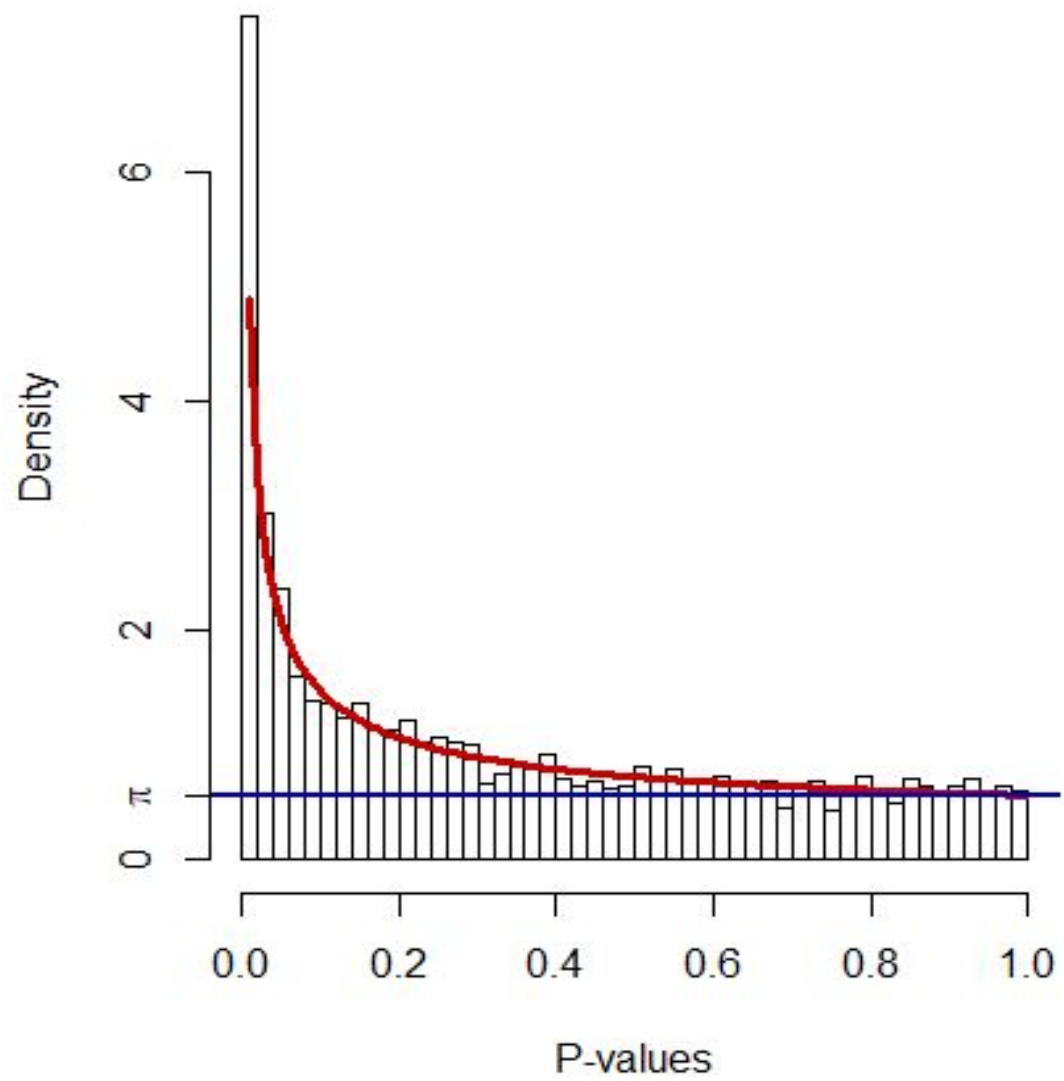
Beta-uniform distribution:

$$\lambda \cdot 1 + (1 - \lambda) \cdot a \cdot X^{a-1}$$

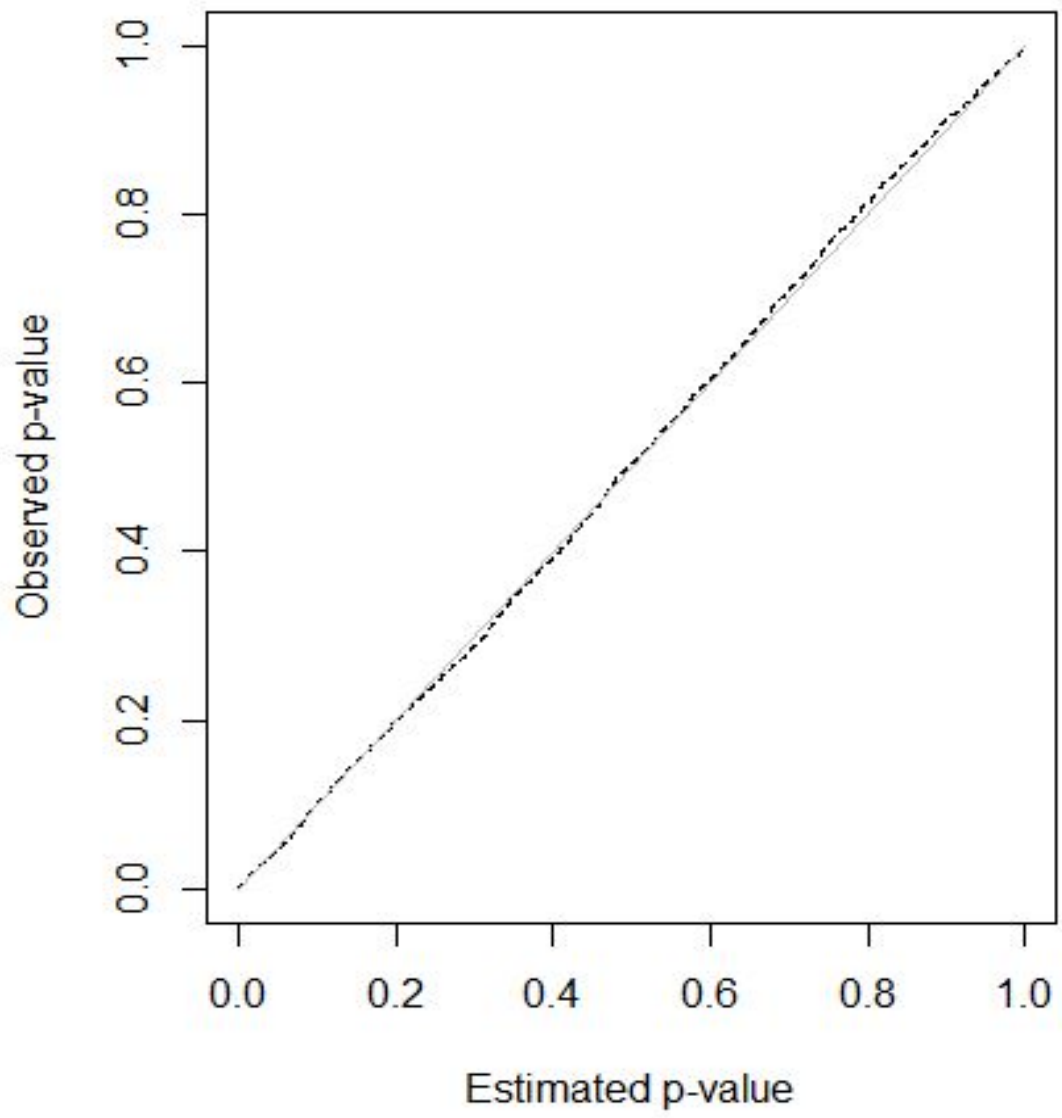
Histogram of p-values



### Histogram of p-values



### QQ-Plot



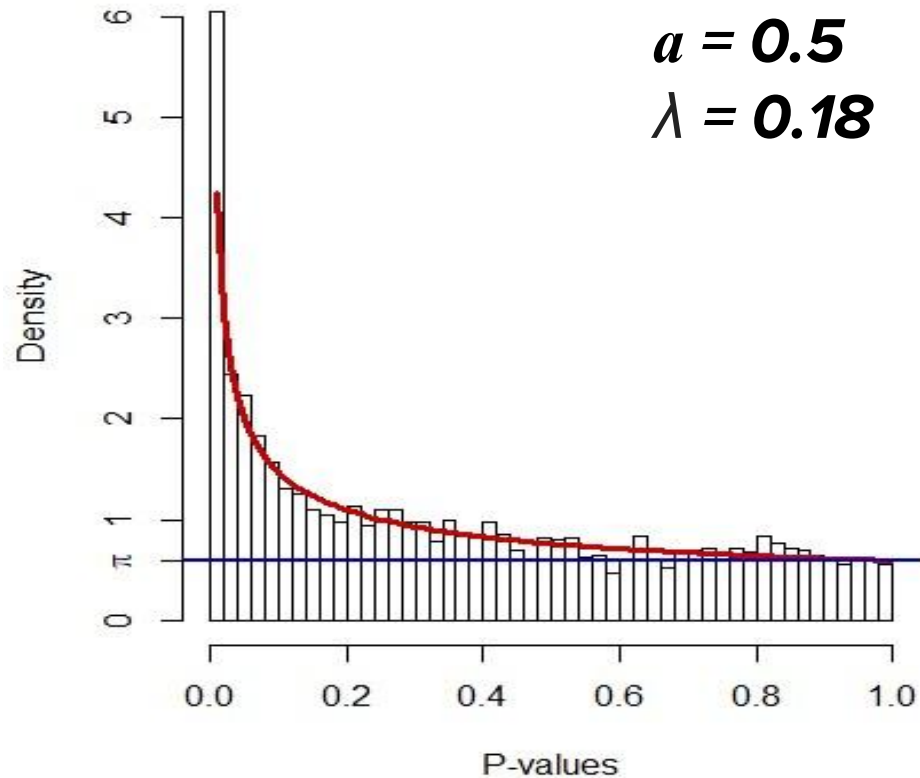
Задача - научиться симулировать сами экспрессии, а не p-value данных

Необходимо понять:

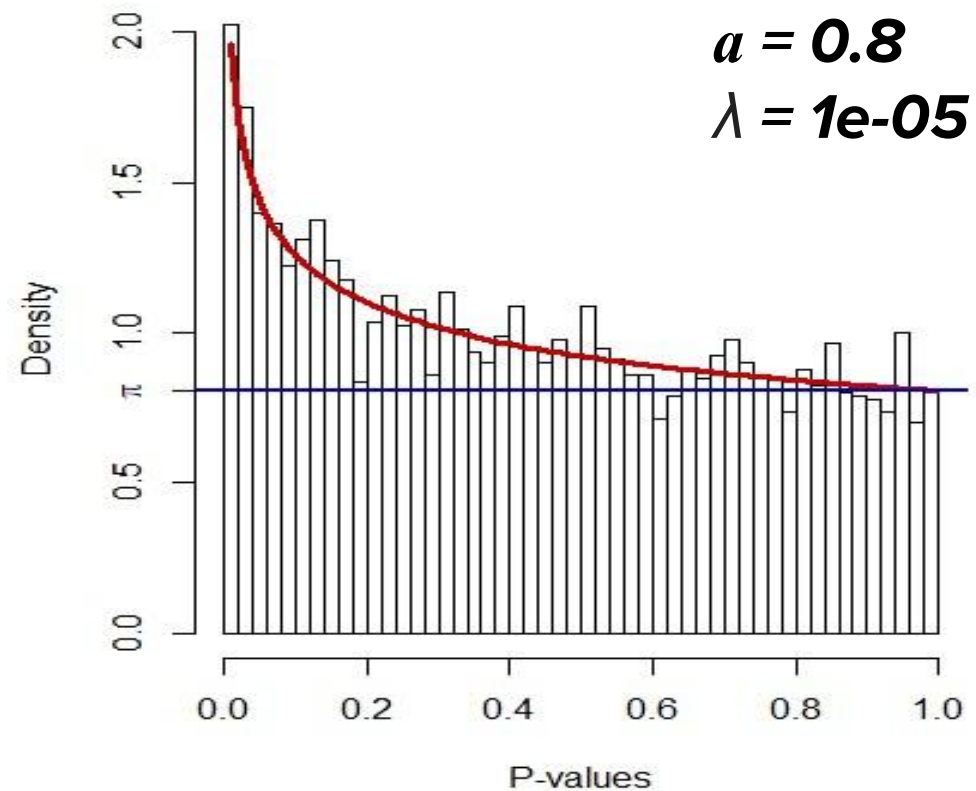
- Как распределена разница экспрессий
- Как устроены зависимости, и как они влияют на данные

# Расчет распределения сдвигов средних для любого значения параметра $a$

Histogram of p-values



Histogram of p-values



# Исследование зависимостей

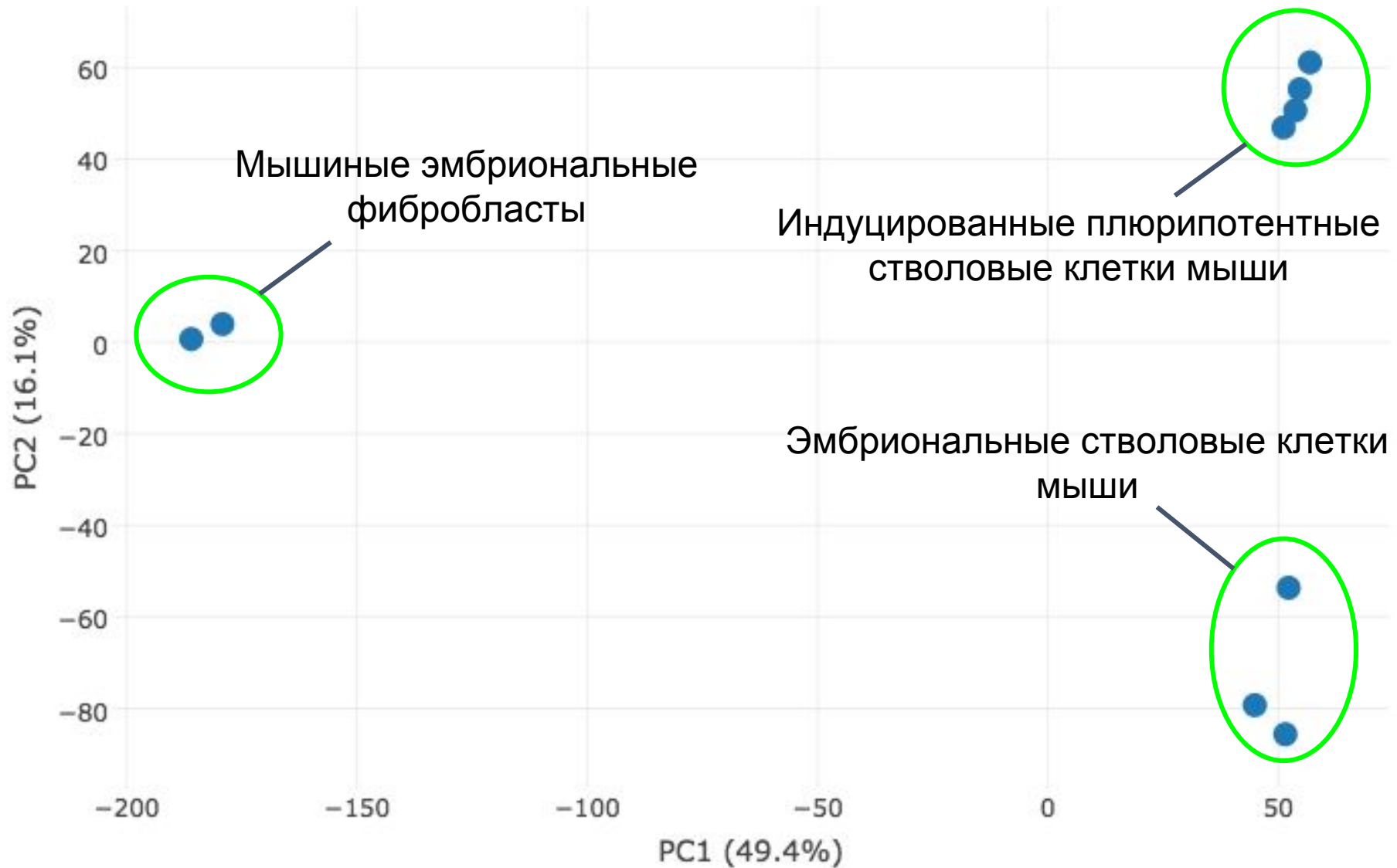
- Монотонные
  - Исследовали методом корреляционного анализа
  - Значимо не влияют на данные
- Немонотонные
  - Для исследования использовали критерий хи-квадрат
  - Также значимо не влияют на данные



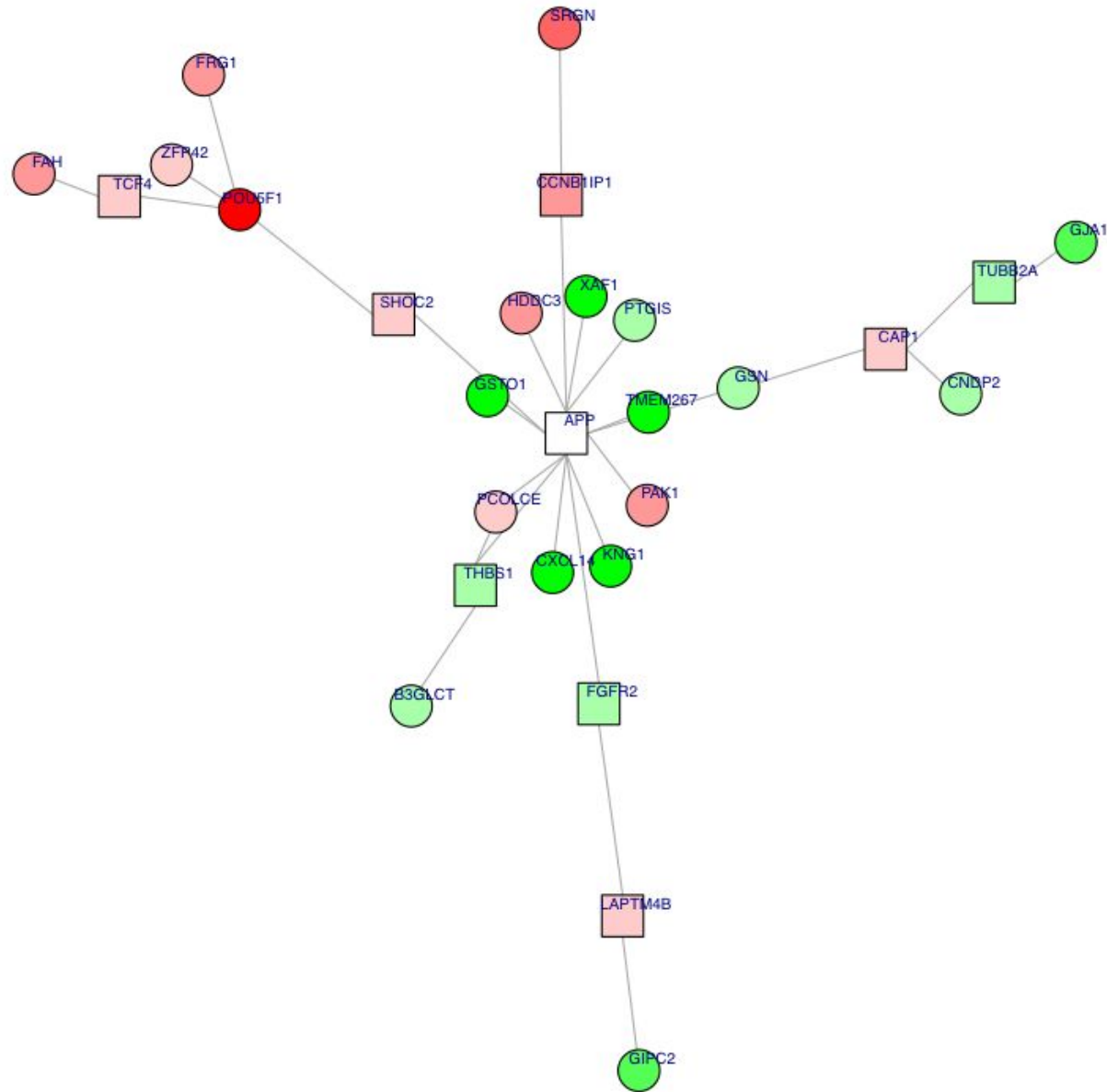
# Данные для работы

- Данные дифференциальной экспрессии генов в эмбриональных фибробластах (ЭФ), эмбриональных стволовых клетках (ЭСК) и индуцированных плюрипотентных стволовых клетках (иПСК) мыши
- Интересно сравнить экспрессию генов в ЭСК и иПСК, т.к. обсуждается использование иПСК в качестве материала для заместительной терапии при трансплантации, а также прямая конвертация клеток *in vivo*

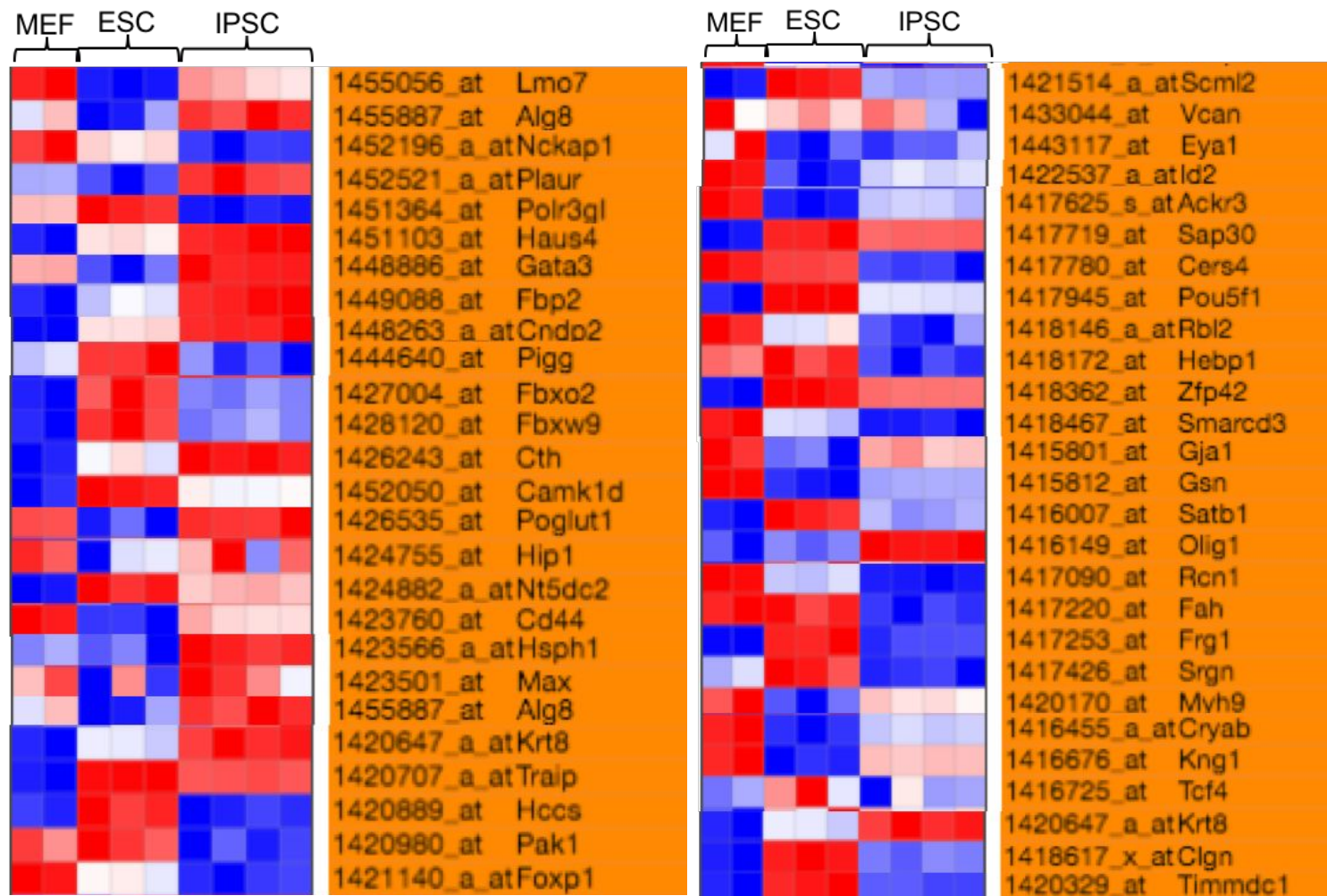
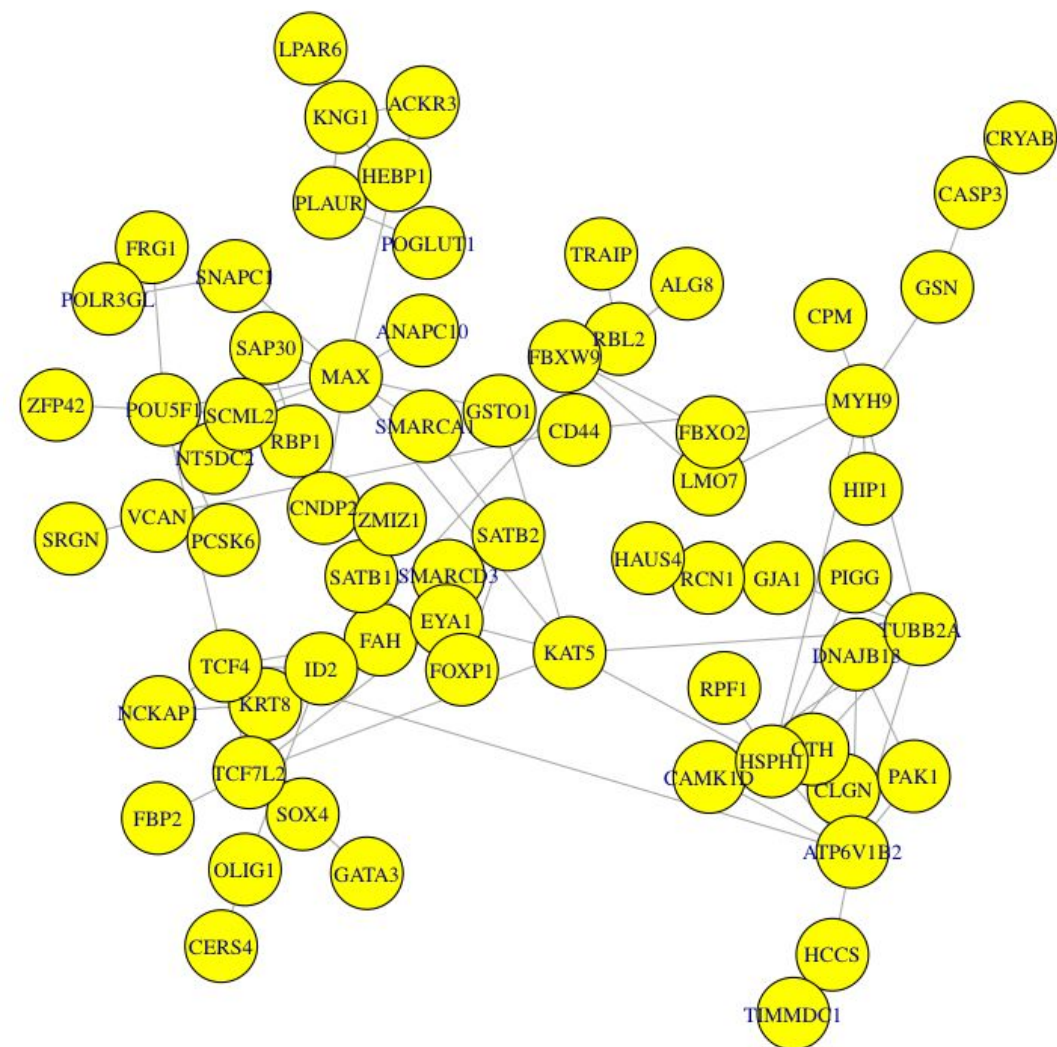
# Данные для работы



# BioNet



# MCMC-ranking



Спасибо за внимание!

GitHub: [https://github.com/flowdel/diff\\_gene\\_expression](https://github.com/flowdel/diff_gene_expression)