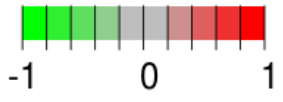


Поиск ключевых генов, инициирующих смену профиля экспрессии в клетке.

Студент:
Горбач Дарья

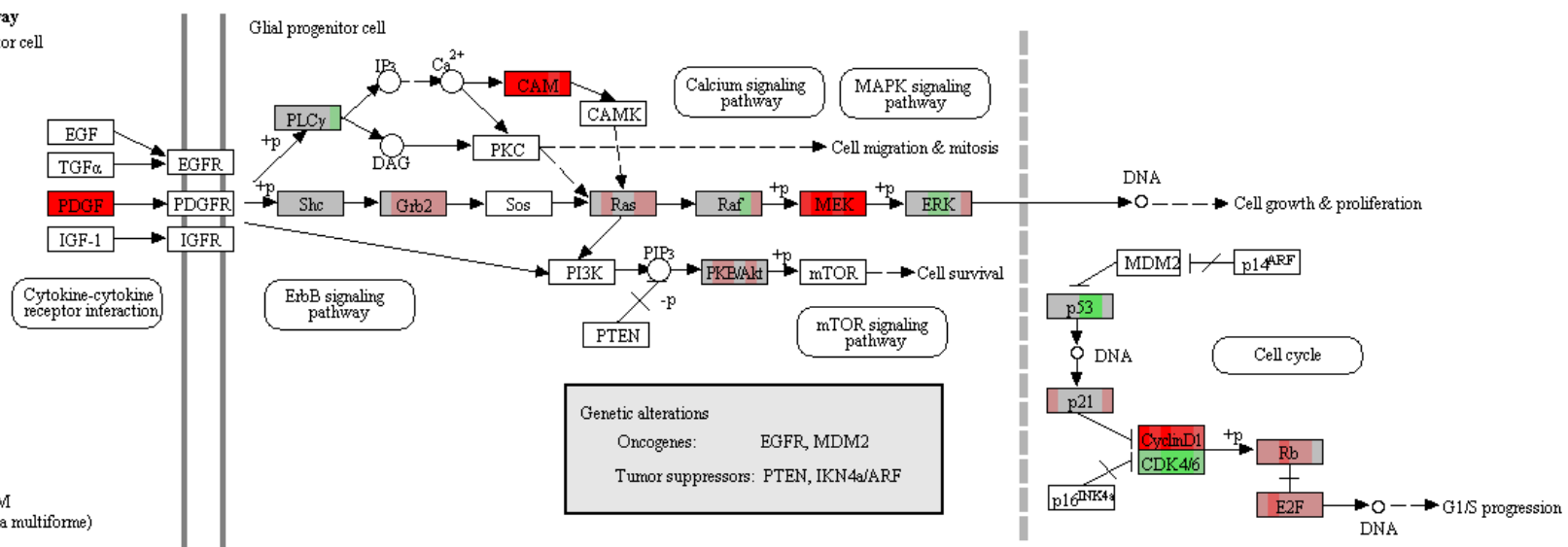
Руководители:
Оксана Станевич (ПСПбГМУ им. акад. И.П.Павлова)
Евгений Бакин (Центр геномной биоинформатики
им. Ф.Г. Добржанского)

Почему это важно

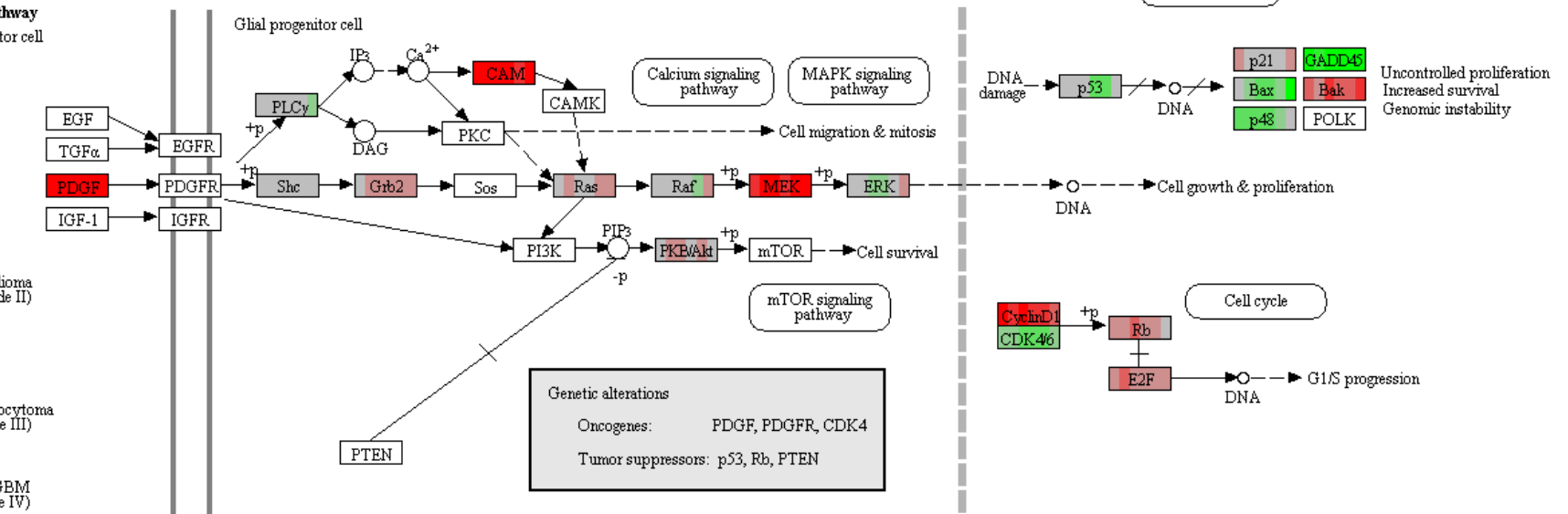


GLIOMA

De Novo pathway
Glial progenitor cell



Secondary pathway
Glial progenitor cell



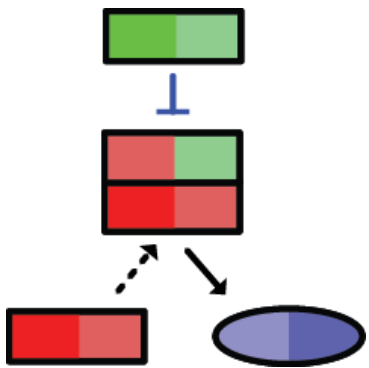
KEGG PATHWAY (Kyoto Encyclopedia of Genes and Genomes Pathway)



- Коллекция баз данных о генах, геномных продуктах и их взаимодействиях
- Также включает в себя ряд инструментов для анализа биологических и медицинских данных

KEGGgraph
(R package)

- Парсинг KGML файлов и перевод информации в графовое представление
- Основные операции над графами



PATHVIEW

Pathway based data integration and visualization

- Визуализация дифференциальной экспрессии

Цели и задачи

- Освоить основные методы работы с данными дифференциальной экспрессии на языке R (Limma, KEGGgraph, Pathview)
- Разработать графовый алгоритм поиска ключевых для смены профиля экспрессии генов
- Проанализировать биологическую релевантность полученных результатов

Получение значимо дифференциально экспрессирующихся генов с помощью **Limma (Phantasus)**

Загрузка соответствующих сигнальных путей из базы данных **Kegg**

Пайплайн

Разобраться со структурой данных **KEGGgraph**

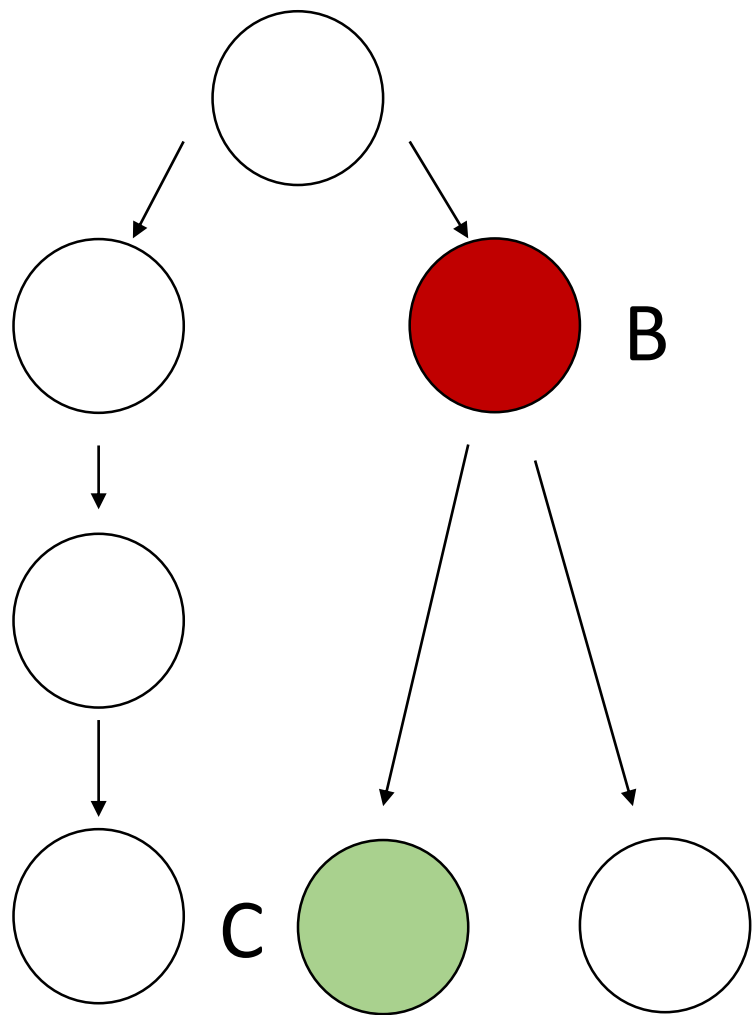
Поиск минимального значимого количества генов, объясняющих наибольшее количество сигнальных путей

Объединение нескольких сигнальных путей (в виде графов) в больший граф

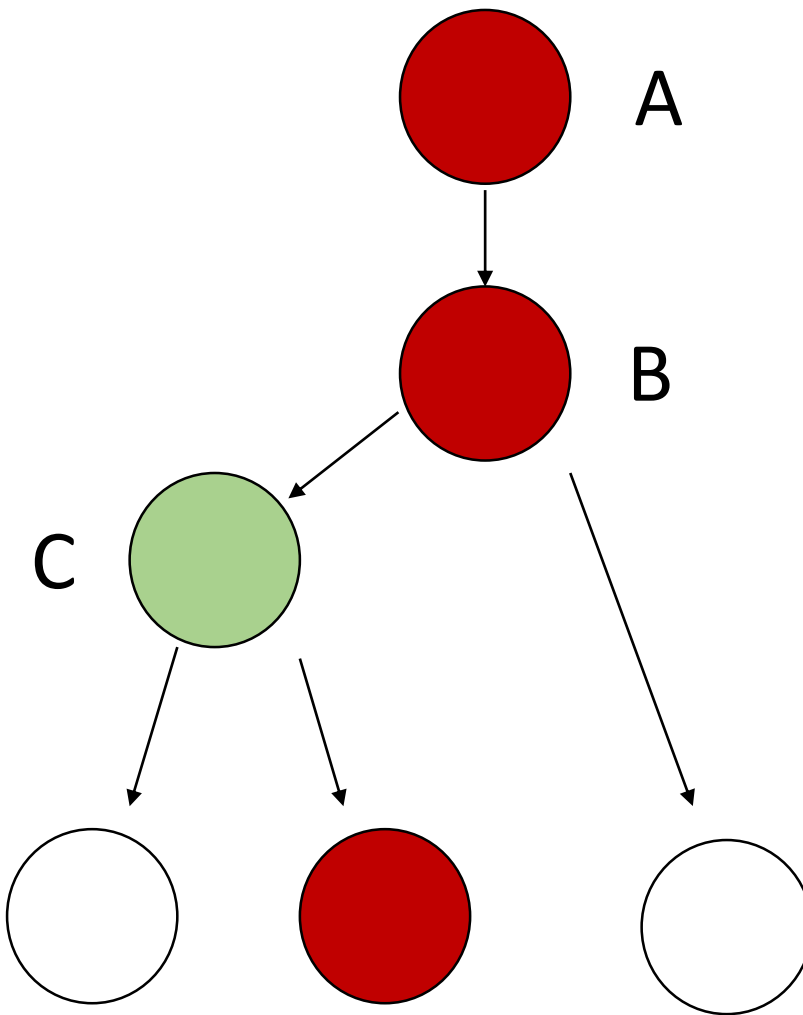
Принцип работы алгоритма

- Поиск такого подмножества генов, в котором содержатся все дифференциально экспрессирующиеся гены, а количество не дифференциально экспрессирующихся минимально

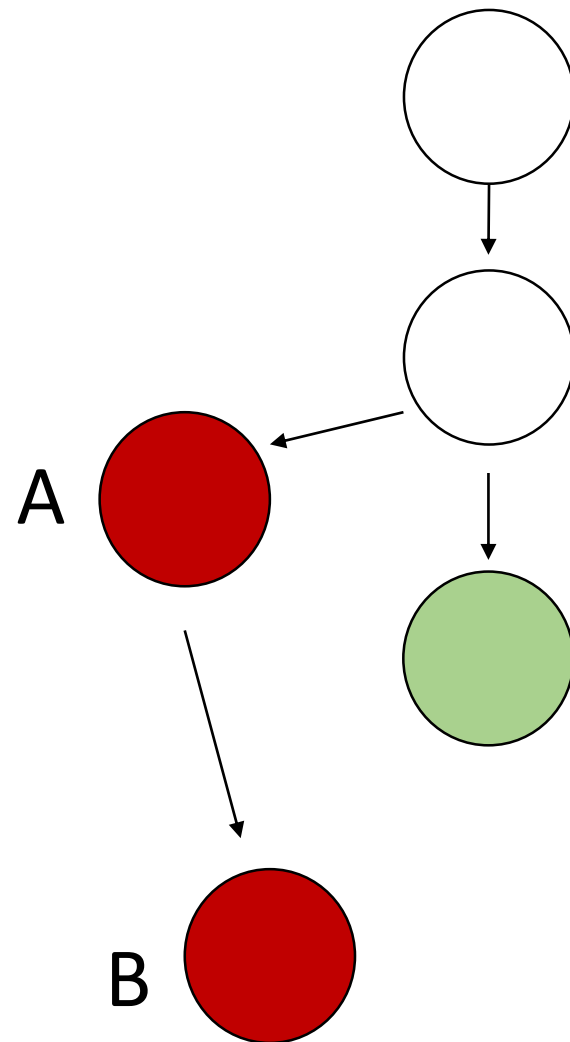
Сигнальный путь 1

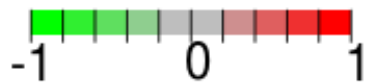


Сигнальный путь 3

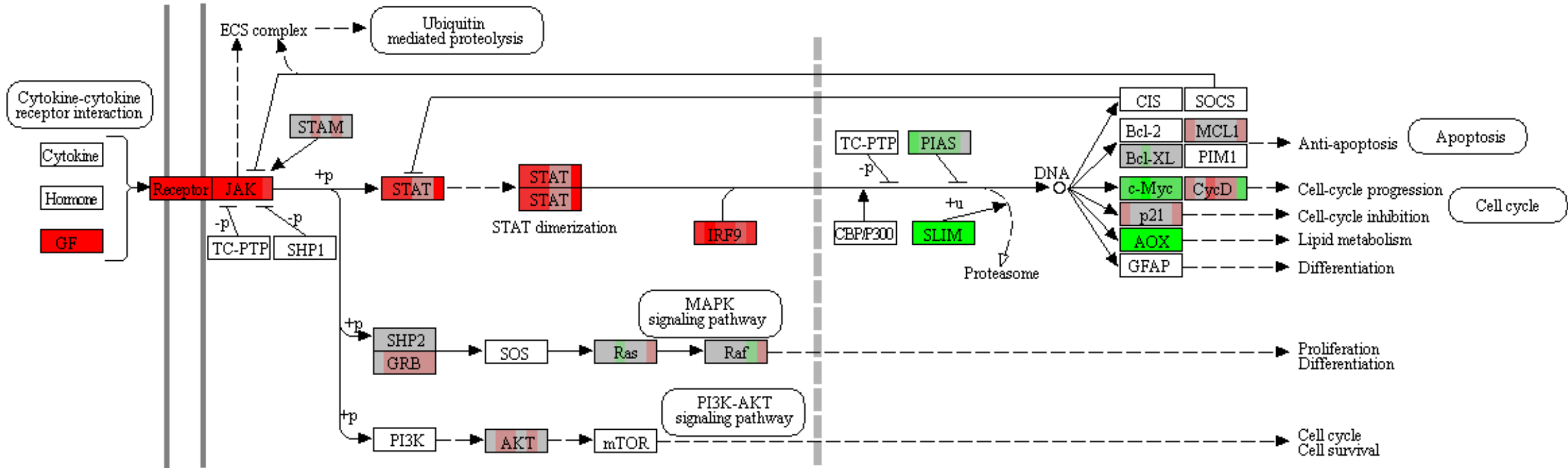


Сигнальный путь 2





JAK-STAT SIGNALING PATHWAY



Data on KEGG graph
Rendered by Pathview

KSHV LANA upregulates the expression of epidermal growth factor like domain 7 to promote angiogenesis

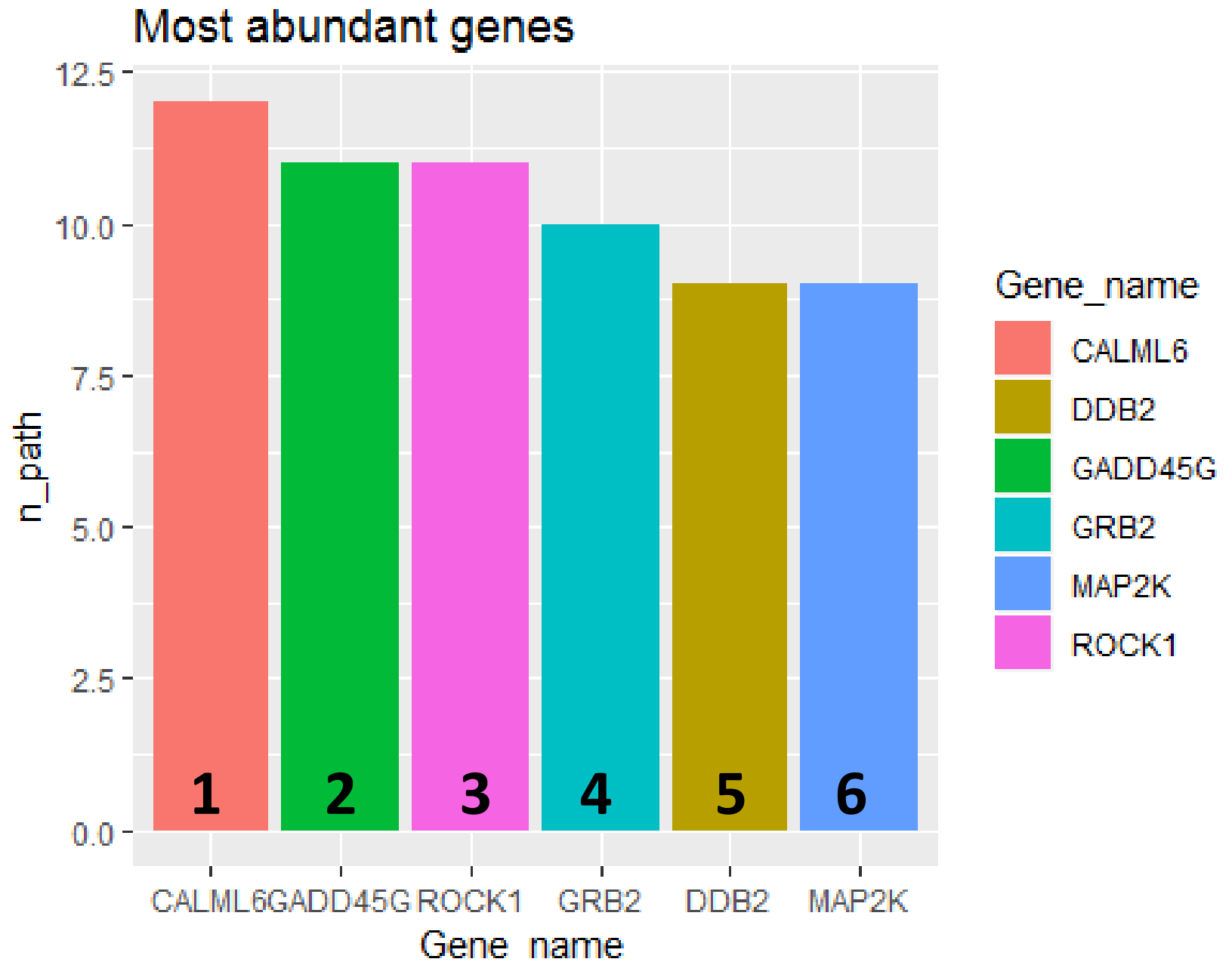
Suhani Thakker¹, Roxanne C. Strahan¹, Alexandra N. Scurry¹, Timsy Uppal¹ and Subhash C. Verma¹

¹Department of Microbiology and Immunology, School of Medicine, University of Nevada, Reno, NV 89557, USA

Принцип работы на нескольких сигнальных путях

- Объединение множества сигнальных путей в один мегаграф (кроме слишком разреженных: вершин больше, чем ребер)
- Пересечение вершин (генов) мегаграфа со списком дифференциально экспрессирующихся генов
- Подсчет частоты встречаемости каждого гена среди данных сигнальных путей

- 1 – Wang et al., 2019
- 2 – Liu, Happel et al., 2017;
Tamura et al., 2012
- 3 – Veettil et al., 2016
- 4 – Gallaher et al., 2013
- 5 – ?
- 6 – Pan et al., 2006



Планы на будущее

- Программная оптимизация подхода для поиска значимых генов в сигнальном пути
- Учет вторичных метаболитов
- Усовершенствование целевой функции

Спасибо за внимание!

GitHub:

[https://github.com/DariaGorbach/Dif_expression_profiles
project](https://github.com/DariaGorbach/Dif_expression_profiles_project)