

Анализ открытости хроматина в процессе клеточного омоложения

Бушуев Сергей Евгеньевич

Студент (специалист)

Московский государственный университет имени М.В.Ломоносова, Москва, Россия

E-mail: bushsebzz@gmail.com

Введение Старение - фундаментальный процесс, имеющий огромное значение для жизни в целом. Многие поколения учёных занимались и продолжают заниматься этой темой, недавно, в 2006 было открыто клеточное репрограммирование, за что японский исследователь Синъя Яманака получил Нобелевскую премию по физиологии и медицине. Процесс позволяет индуцировать плюрипотентное состояние в соматических клетках, обернув историю клеточной дифференцировки.

Метод АТАС-seq позволяет количественно оценить открытость различных участков хроматина. В настоящее время представлено большое количество работ, связанных со старением и репрограммированием, которые выполнены с помощью АТАС-seq, что позволяет провести мета-анализ представленных работ.

Основной целью моей работы было нахождение взаимобратности этих процессов в омиксных данных с помощью мета-анализа.

Результаты

С помощью базы данных GEO NCBI было собрано 10 статей по тематике репрограммирования (для старения аналогичную работу выполнил студент нашей лаборатории - Кирилл Кузенков: 8 статей, дальнейшую работу по анализу данных по этой теме осуществлял он). Во всех собранных работах, связанных с репрограммированием, представлены данные АТАС-seq промежуточных стадий репрограммирования мышинных эмбриональных фибробластов. В других собранных статьях, связанных со старением, опубликованы аналогичные данные клеток тканей нескольких органов разных по возрасту представителей *Mus musculus*. Полученные данные были независимо обработаны с помощью программного конвейера `nf-core/atacseq`, написанного на `nexflow`. Одним из результатов работы конвейера является бинарная таблица с консенсусными пиками, содержащая информацию о том, открыт или закрыт конкретный интервал в геноме в каждом из образцов, имеющих временные метки. Для каждого пика в каждой таблице была построена линейная регрессия по времени, что позволило охарактеризовать пики соответствующими угловыми коэффициентами и их дисперсиями. В пределах тем (старения и репрограммирования) были пересечены пики, после чего проведен мета-регрессионный анализ для каждого пика. После отбора по поправленному *p-value* с порогом на уровень значимости - 0.1 в репрограммировании осталось 20000 пиков, а в старении - 810. После чего это пики были пересечены между таблицами двух тем. Их оказалось всего 168, из которых 153 были взаимно обратными (имели противоположные знаки наклонов регрессий).

С помощью аннотации, выполненной программой HOMER в составе конвейера, эти пики были ассоциированы с генами, список которых был отправлен в анализ обогащения. С помощью него было выявлено 2 группы генов: “TNF alpha сигнальный путь через NF-kB” и “отрицательная регуляция транспорта ионов калия”.