

ПОИСК ГЕННЫХ СИГНАТУР В SINGLE-CELL NAVIGATOR ПОЗВОЛЯЕТ НАХОДИТЬ РЕДКИЕ КЛЕТОЧНЫЕ ПОПУЛЯЦИИ В ПУБЛИЧНЫХ ДАННЫХ СЕКВЕНИРОВАНИЯ РНК ОДИНОЧНЫХ КЛЕТОК

Зайцев К.В. (Университет ИТМО),

Научный руководитель – к.т.н. Сергушичев А.А. (Университет ИТМО)

В данной работе разрабатывается онлайн-инструмент для анализа и визуализации данных секвенирования РНК одиночных клеток. Разрабатываемый инструмент также предоставляет другим исследователям доступ к публичным данным секвенирования РНК одиночных клеток.

В рамках данного инструмента был реализован метод поиска: по заданному набору генов (или по одиночным генам) метод находит публичные наборы данных, где экспрессия заданных генов повышена в одной или нескольких популяциях. Мы иллюстрируем применимость данного метода на примере поиска редких популяций в публичных данных секвенирования РНК одиночных клеток.

Введение. С растущим интересом в секвенированию РНК одиночных клеток исследователи производят всё больше наборов данных секвенирования РНК одиночных клеток. Однако некоторые клеточные популяции сложнее исследовать из-за их малого количества внутри биологического образца. Эти редкие клеточные популяции, однако, представлены в большом количестве публичных наборов данных. В данный момент отсутствуют легкодоступные инструменты поиска клеточных популяций по всем публичным данным. В данной работе мы показываем, как разработка такого инструмента может позволить исследователям собрать редкие популяции из разных наборов данных в один «мета»-набор данных. Анализ таких «мета»-наборов данных может позволить исследовать редкие клеточные популяции в большем масштабе.

Основная часть. Мы разработали Single-Cell Navigator - онлайн-инструмент для анализа, визуализации и поиска на основе генов в общедоступных наборах данных scRNA-seq. Используя стандартизированный конвейер обработки, мы повторно проанализировали большинство наборов данных scRNA-seq, доступных в NCBI GEO, чтобы обеспечить визуальный осмотр и базовый кластерный анализ. Наиболее важно то, что мы реализовали специальный алгоритм поиска сигнатур генов: по списку генов, scNavigator может найти набор данных, в котором экспрессия этих генов будет повышена в одном из кластеров / субпопуляций клеток в более чем 4000 наборах данных scRNA-seq в базе данных в разных биологических контекстах. Мы собрали эти наборы данных в общедоступный онлайн-ресурс (<https://artyomovlab.wustl.edu/scn/>), который может свободно использоваться научным сообществом для изучения отдельных генов и наборов генов, представляющих интерес, которые не были охвачены в этом исследовании.

Мы иллюстрируем эффективность подхода поиска редких популяций на нескольких примерах редких популяций – а именно лимфоидные эндотелиальные клетки и редкие популяции иммунных Т-клеток. Такой подход позволяет не только находить редкие популяции и тем самым подтверждать их наличие в других образцах, но и проводить сравнение этих редких популяций в разных тканях.

Выводы. Был разработан Single-Cell Navigator – база данных публичных данных секвенирования РНК одиночных клеток, с возможностью поиска по одиночным генам и по наборам генов. Была проиллюстрирована эффективность такого подхода: мы показываем, как использовать поиск для идентификации редких клеточных популяций в публичных данных

секвенирования РНК одиночных клеток. Мы считаем, что наш онлайн-инструмент может быть полезен не только в рамках поиска редких популяций, а также и для перекрестного сравнения популяций между собой, сравнения генных сигнатур, полученных любым способом с данными секвенирования РНК одиночных клеток, и для много другого – поэтому мы сделали наш онлайн-инструмент публично доступным по адресу <https://artyomovlab.wustl.edu/scn/>.

Зайцев. К. В. (автор)

Подпись

Сергушичев А.А. (научный руководитель)

Подпись