

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ МНОГОМЕРНОГО АНАЛИЗА МИКРОЧИПОВЫХ ДАННЫХ: ИНТЕРФЕЙС, ВХОДНОЙ ЯЗЫК

И. А. Штайгер

Институт цитологии и генетики СО РАН
Новосибирский государственный университет

В молекулярной биологии и медицине широко используется технология, основанная на ДНК-микрочипах, позволяющая производить анализ экспрессии генов. На ДНК-микрочип с некоторой специфической последовательностью ДНК наносится образец, после чего, при определенных условиях, достигается гибридизация зонда и мишени. Затем, с помощью флюоресценции определяется количество ДНК с заданной последовательностью в образце. Набор таких данных называется микрочиповыми данными.

Существует большое количество специализированных программных средств, предназначенных для анализа микрочиповых данных [2]. Однако, они не слишком удобны в использовании и некоторые из них являются коммерческими. Основные проблемы таких пакетов – закрытая реализация и использование собственных форматов данных. Существуют и более универсальные средства, обладающие, однако, тяжеловесным интерфейсом (например, пакет R) [1]. Поэтому разработка простого и удобного программного комплекса для анализа микрочиповых данных представляется весьма актуальной.

При проектировании программного комплекса был разработан интуитивно понятный конечному пользователю язык для написания скриптов, близкий к естественному (русскому или английскому) языку.

Схема работы программного комплекса: программа-диспетчер получает скрипт, написанный пользователем, в формате csv, проверяет его корректность с точки зрения входного языка, и, если скрипт корректен, создает файл, который содержит список команд с параметрами, а также индекс-файл, который обеспечивает быстрый доступ к любой команде. Затем, диспетчер считывает очередную команду из файла и вызывает соответствующий модуль, который перед завершением вызывает диспетчер. Такой программный комплекс содержит все необходимые функции для удобной работы пользователя-биолога.

1. J. M. Curran, Introduction to Data Analysis with R for Forensic Scientists. CRC Press, Boca Raton, FL, (2011).

2. Биоинформатика. Инфраструктура. Электронный ресурс.
<http://www.bionet.nsc.ru/bioinf/files/bioinformatica.pdf>

Научный руководитель – д-р биол. наук, доцент. В. М. Ефимов