

КОНТРОЛЬ КАЧЕСТВА ГЕНОМНЫХ БИБЛИОТЕК ДЛЯ СЕКВЕНАТОРОВ ВТОРОГО ПОКОЛЕНИЯ НАНОФОР СПС

Адельшина Э. В. (Университет ИТМО), Ващенко К. Д. (Университет ИТМО)

Научный руководитель – кандидат технических наук Петров А. И.
(Университет ИТМО, ИАП РАН)

Введение. НАНОФОР СПС – прибор для проведения геномного анализа методом массового параллельного секвенирования (NGS) [1]. Приборы такого класса предъявляют повышенное требование к качеству исследуемых образцов ДНК (геномных библиотек), поэтому при проведении исследований проводится контроль качества геномных библиотек. Целью данной работы является разработка математического аппарата, обеспечивающего проведение контроля качества геномных библиотек для секвенаторов второго поколения НАНОФОР СПС с помощью секвенатора первого поколения НАНОФОР 05 (по Сэнгеру).

Основная часть. На современном рынке представлен американский прибор Agilent 2100 Bioanalyzer, который даёт возможность осуществлять анализ качества геномных библиотек с помощью капиллярного электрофореза и имеет для этого необходимое программное обеспечение [2]. НАНОФОР 05 работает по методу капиллярного электрофореза и позволяет проводить фрагментный анализ, разновидностью которого может быть контроль качества геномных библиотек. Для того, чтобы решить эту задачу, необходимо разработать методику проведения неденатурирующего электрофореза, разработать алгоритм обработки данных, провести апробацию написанного алгоритма и на базе полученного алгоритма разработать программное обеспечение. Методика проведения неденатурирующего электрофореза была разработана сотрудниками НПК «Синтол». Методика предусматривает одновременный анализ геномной библиотеки и стандарта длин (лэддера). В ходе научно-исследовательской работы был разработан алгоритм в среде MATLAB, состоящий из трёх основных этапов: поиск пиков библиотеки, поиск пиков лэддера и получение итоговых результатов в виде оценок качества геномных библиотек:

1. Поиск пиков библиотеки.

- Очистка экспериментальных данных от шума с использованием фильтра Савицкого-Голея;
- Проверка соответствия пиков пороговому значению;
- Обнаружение всех локальных пиков с помощью второй производной;
- Получение координат времени выхода найденных пиков;
- Подсчёт площадей для каждого найденного пика.

По завершении этапа получаем таблицу со значениями времён выхода пиков геномной библиотеки и их площадями.

2. Поиск пиков лэддера

- Внесение значений длин фрагментов и концентрации лэддера;
- Обнаружение всех пиков лэддера и идентификация их согласно внесённому списку длин;
- Подсчёт площадей для каждого найденного пика.

По завершении этапа получаем таблицу со значениями времён выхода пиков лэддера и их площадями.

3. Получение итоговых результатов.

- Построение калибровочных кривых на основании данных лэддера;
- Расчёт длин фрагментов и концентрации для пиков библиотеки;
- Визуализация на экране графика электрофореграммы геномной библиотеки и таблицы, содержащей найденные значения длин фрагментов и их концентраций.

Качество геномной библиотеки оценивается на основе полученных данных и заданных значений длин фрагментов и концентрации из требований подготовки геномных библиотек для НАНОФОР СПС.

Выводы. В ходе научно-исследовательской работы был разработан алгоритм в среде MATLAB, который показал возможность проведения оценки качества геномных библиотек на секвенаторе первого поколения НАНОФОР 05. Созданный алгоритм будет добавлен в программное обеспечение прибора НАНОФОР 05, после чего секвенатор может быть использован для контроля качества геномных библиотек для секвенатора второго поколения НАНОФОР СПС.

Список использованных источников:

1. Различные типы секвенирования ДНК. — Текст : непосредственный // Молодой ученый. — 2018. — No 20 (206). — URL: <https://moluch.ru/archive/206/83008/> (дата обращения: 30.01.2024).
2. DNA Fragment Analysis by Capillary Electrophoresis User Guide. – 2014- No 4474504 - Rev. B – C. 220.