
Секвенирование длинных прочтений по технологии Oxford Nanopore – эффективный инструмент для изучения организации бактериальных геномов с варибельной структурой

Делеган Я.А.¹, Кочаровская Ю.Н.^{1,2}, Французова Е.Э.^{1,2}, Соляникова И.П.^{1,3}

¹ ФИЦ «Пушинский научный центр биологических исследований РАН»,
(Институт биохимии и физиологии микроорганизмов им. Г.К. Скрыбина РАН)

² ФГБОУ ВО Пушинский государственный естественно-научный институт

³ ФГАОУ ВО Белгородский государственный национальный исследовательский университет, mewgia@ya.ru

DOI: 10.34756/GEOS.2021.17.37953

В настоящее время секвенирование полных геномов является одним из основных подходов для определения генетической организации исследуемых организмов. Информация о последовательности и структуре полного генома позволяет с высокой точностью планировать молекулярно-биологические и генетические эксперименты.

Результатом секвенирования коротких прочтений (в основном по технологии Illumina) и последующей сборки являются протяжённые фрагменты – контиги, которые дают общее представление о геноме изучаемого организма. Тем не менее, воссоздать полные последовательности репликонов, особенно у штаммов с большим количеством повторов, мобильных генетических элементов (МГЭ) и коротких консервативных областей, с помощью секвенирования коротких прочтений практически невозможно даже при наличии референса – полностью собранного генома родственного организма.

Результатом секвенирования геномной ДНК по технологии Oxford Nanopore являются длинные прочтения, сборка которых позволяет с высокой точностью воссоздавать структуры полных репликонов даже в том случае, когда объектом исследования являются бактериальные геномы с большим количеством МГЭ и высококонсервативных повторяющихся областей. Коллективом авторов на оборудовании MinION (ONT) на базе лаборатории физиологии микроорганизмов выполняется секвенирование и сборка геномов различных актинобактерий (*Gordonia*, *Rhodococcus*), протеобактерий (*Pseudomonas*), а также дрожжей.

Выполнено секвенирование штамма *Rhodococcus (erythropolis)* VT6. Собранные последовательности всех репликонов были сравнены с ранее полученными коллективом авторов сборками штаммов *R. qingshengii* 7B [1] и *R. erythropolis* X5 [2]. Интересно отметить, что, несмотря на родство протяжённых участков ДНК, расположены они в геномах штаммов по-разному. При секвенировании штамма VT6 получено 1.716.067.812 нуклеотидов, распределённых по 171.434 прочтениям, из которых 130.800 (76.3%) имели качество $Q > 10$. Геном этого штамма имеет сложную структуру, включающую многочисленные МГЭ, а также внехромосомные генетические элементы. Геном состоит из хромосомы (6.464.421 п.н.) и 4 плазмид: 498 т.п.н., 187 т.п.н., 133 т.п.н. (кольц.), 100 т.п.н. (кольц.). Среднее покрытие – 120. Для коррекции мелких ошибок мы используем данные Illumina. Концы кольцевых элементов достоверно перекрываются.

Описанный в данной работе подход, включающий секвенирование длинных прочтений, сборку и анализ полученных данных, позволяет с достаточной точностью собирать не контиги, а полные репликоны, структура которых достоверно подтверждена. Коллектив авторов планирует расширять таксономическое разнообразие объектов, секвенируемых указанным методом, что позволит внести значительный вклад в изучение структур полных геномов микроорганизмов.

Выполнено в рамках гранта Министерства науки и высшего образования РФ по соглашению № 075-15-2021-973 (28.09.2021-31.12.2021)