

ВНЕДРЕНИЕ ТЕХНОЛОГИИ 454 LIFE SCIENCES В ЛАБОРАТОРНУЮ ПРАКТИКУ

Н.Г. Зиняков¹, Е.В. Овчинникова², С.П. Лазарева³, А.А. Козлов⁴, И.А. Чвала⁵

¹ научный сотрудник, кандидат биологических наук, ФГБУ «ВНИИЗЖ», г. Владимир, e-mail: zinyakov@arriah.ru

² младший научный сотрудник, кандидат биологических наук, ФГБУ «ВНИИЗЖ», г. Владимир

³ ведущий биолог, ФГБУ «ВНИИЗЖ», г. Владимир

⁴ аспирант, ФГБУ «ВНИИЗЖ», г. Владимир

⁵ заведующий лабораторией, кандидат ветеринарных наук, ФГБУ «ВНИИЗЖ», г. Владимир

РЕЗЮМЕ

Показана возможность применения высокопроизводительного пирофосфатного секвенирования, также известного как технология 454 Life Sciences, для исследования высоковариабельных и малоизученных инфекционных агентов вирусной природы. Определены полногеномные нуклеотидные последовательности вирусных инфекционного бронхита кур, инфекционного ларинготрахеита, аденовируса птиц генотипа С и аденоассоциированного вируса птиц. Идентифицированы рекомбинации в геноме инфекционного бронхита кур, выявлены структурные и точечные мутации в геноме аденовируса птиц. Охарактеризована структура генома вируса инфекционного ларинготрахеита и аденоассоциированного вируса птиц.

Ключевые слова: инфекционный бронхит кур, инфекционный ларинготрахеит, аденовирус птиц, аденоассоциированный вирус птиц, геном, нуклеиновые кислоты, секвенирование.

ВВЕДЕНИЕ

В настоящее время для решения широкого круга научно-исследовательских задач активно используют методы высокопроизводительного секвенирования. Совокупность новых методов также называют секвенированием второго поколения или Next Generation Sequencing (NGS). Области применения технологий NGS охватывают работы, связанные с полногеномным секвенированием, таргетным секвенированием, анализом транскриптома и метагеномными исследованиями [7, 8]. NGS-технологии позволяют проводить работу без расчета праймеров за счет случайной фрагментации исходной ДНК и секвенирования полученной библиотеки, представляют огромные преимущества для работы с вариабельными и малоизученными инфекционными агентами. Первым методом, заложенным в основу высокопроизводительного секвенирования, было пирофосфатное секвенирование. Метод был основан на детекции флуоресцентного сигнала в результате разрушения пирофосфата (побочного продукта при синтезе ДНК) и являлся методом секвенирования путем синтеза. Впоследствии метод активно модифицировали и оптимизировали, что привело к созданию технологии 454 Life Sciences, также известной как «454 sequencing» (секвенирование 454). Отличительной особенностью

данной технологии среди других технологических решений в области высокопроизводительного секвенирования стала наибольшая длина единичных прочтений, что позволило применять ее для работ как с уже известными, так и с малоизученными организмами, поскольку, получая более длинные прочтения, выполнять сборку изучаемого генома можно при меньшем объеме данных и получать достоверные результаты при более низком покрытии. Однако использование NGS-технологий для расшифровки вирусных геномов сопряжено с рядом трудностей. Основная сложность заключается в очистке вируса от клеточного матрикса. В случае если очистка проведена некачественно, выход интересующих данных может оказаться незначительным и недостаточным для полноценного анализа. С целью увеличения информации об исследуемом объекте в генерируемом массиве данных проводят ряд дополнительных процедур по очистке и концентрации целевого объекта, например, иммунопреципитацию вирусных частиц, градиентное центрифугирование, выделение нуклеиновой кислоты с помощью вичипов со специальными зондами, а также дополнительную амплификацию со специфическими праймерами. Учитывая последующую сложную обработку полученного массива данных, наличие в геноме вирусов различных повторов и в целом более высокую вариабельность при анализе уже изученных вирусов, большая длина единичных прочтений упрощает дальнейшую биоинформатическую обработку.

Целью данной работы было изучение возможности применения технологии 454 Life Sciences для расшифровки генома возбудителей вирусных болезней птиц.

МАТЕРИАЛЫ И МЕТОДЫ

Выделение вируса инфекционного бронхита кур (ИБК) и инфекционного ларинготрахеита (ИЛТ) проводили с использованием СПФ-эмбрионов кур путем заражения суспензией патологического материала в аллантоисную полость и на хориоаллантоисную оболочку соответственно.

Аденовирус птиц выделяли с использованием периваемой культуры клеток гепатоцеллюлярной карциномы петуха (LMH). Для культивирования использо-

вали поддерживающую среду (рН 7,8), состоящую из питательной среды DMEM (Sigma, США) с добавлением фетальной сыворотки (Bioclot, Бразилия), в количестве 2% от общего объема, среды и раствора L-глутамина (Sigma, США) с конечной концентрацией 1% активного вещества.

Очистку вируса от клеточного матрикса проводили с использованием ультрацентрифугирования в градиенте сахарозы.

Для выделения РНК и ДНК из очищенного материала применяли набор для выделения нуклеиновых кислот RNeasy Mini Kit (Qiagen, Германия) в соответствии с инструкцией производителя.

Суммарную нуклеиновую кислоту вируса ИБК переводили в двухцепочечную кДНК, используя набор cDNA Synthesis System (Roche, Германия) согласно инструкции изготовителя.

Полногеномное секвенирование проводили методом пирофосфатного секвенирования 454 Life Sciences на автоматическом секвенаторе GS Junior (Roche, Германия) согласно протоколам изготовителя. Для сборки вирусного генома использовали программное обеспечение GS De Novo Assembler (454 Life Science Corp.) и GS Reference Mapper (454 Life Science Corp.). Для выявления рекомбинаций использовали программу Recombination Detection Program (RDP), версия 3.44.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В экспериментах по применению технологии 454 Life Sciences для расшифровки вирусных геномов было использовано несколько различных вирусов, имеющих определенные особенности в геномной организации. Первый эксперимент был проведен с возбудителем инфекционного бронхита кур (ИБК). Вирус ИБК относится к роду *Gamma coronavirus*, подсемейству *Coronavirinae*, семейству *Coronaviridae*, порядку *Nidovirales*. Особенностью данного возбудителя является генетическое разнообразие и высокая скорость изменчивости. Уровень нуклеотидных отличий между изолятами разных генетических групп составляет более 20%, кроме того, встречаются варианты изоляты вируса ИБК, отличающиеся от всех референтных штаммов и не относящиеся ни к одному известному генотипу. Изменения в геноме вируса ИБК происходят в результате накопления точечных мутаций, инсерций, делеций и рекомбинаций. Рекомбинация происходит при инфицировании одной клетки разными штаммами вируса. При этом дочерний вирус может иметь фрагменты генома родительских вирусов разных генотипов, включая вакцинные штаммы [3].

Для полногеномного секвенирования был выбран изолят IBV27-11 вируса ИБК, поскольку по результатам анализа нуклеотидной последовательности на участке гена S1 данного изолята была выявлена рекомбинация [3]. В полногеномном анализе было использовано 4542 единичных прочтения, из которых была собрана последовательность размером 27577 п.н. Анализ полученной нуклеотидной последовательности с помощью программы RDP позволил выявить мозаичную структуру генома с пятью рекомбинационными событиями (рис. 1).

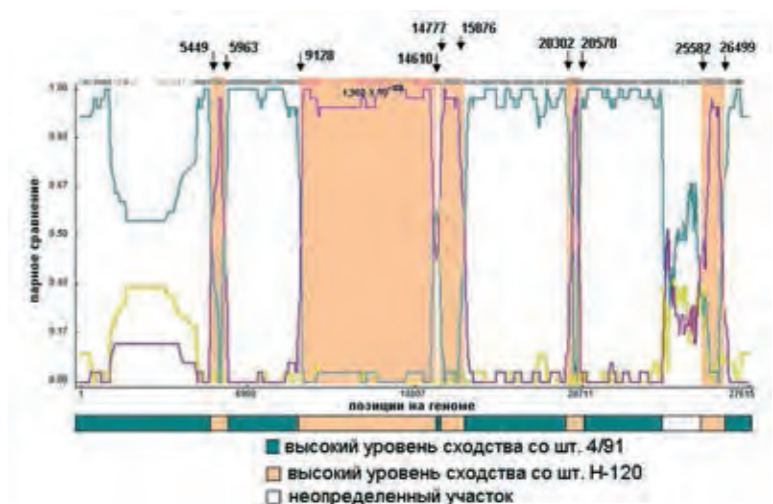
Попытки собрать геном вируса ИБК по алгоритму картирования на известную референтную последовательность с использованием программы GS Reference Mapper (454 Life Science Corp.) оказались безуспешны, при этом число собранных контигов варьировало от 5 до 14 в зависимости от использованной референтной

последовательности. Полученный результат демонстрирует необходимость использования алгоритмов сборки *de novo* для объектов с ожидаемой внутри- или межгеномной рекомбинацией.

Для изучения возможности применения технологии 454 Life Sciences для анализа вирусов, имеющих в геноме незначительные по размеру повторяющиеся мотивы, был исследован полевой изолят аденовируса птиц. Аденовирус птиц — широко распространенный инфекционный агент, относящийся к семейству *Adenoviridae*, роду *Aviadenovirus*. Наиболее известными болезнями птиц, вызываемыми аденовирусами и имеющими характерные патологоанатомические признаки, являются гепатит (IBH-inclusion body hepatitis) и синдром гидроперикардита (HNS-hepatitis-hydropericardium syndrome) [4]. Стоит отметить, что несмотря на то, что аденовирусы являются распространенными инфекционными агентами, на настоящий момент у аденовирусов птиц изучены и охарактеризованы лишь отдельные протеины. В данной работе был использован изолят «Краснодар 2009», выбор был обусловлен результатами предварительной работы [1]. В результате анализа нуклеотидных последовательностей генов фибер-1 и фибер-2 было показано отличие выделенного изолята от известных штаммов, в том числе от референтного штамма KR95. Отличия на генетическом уровне обусловлены как точечными заменами (1,4% отличия), так и структурными мутациями (1,2%). Из точечных замен 30% являются значимыми (0,36% от общего уровня отличий). Количество аминокислотных замен, выявленных у изолята «Краснодар 2009», в сравнении со штаммом KR95 по всем предполагаемым открытым рамкам считывания (ОРС) в результате точечных замен составило 168. Среди структурных мутаций были выявлены как незначительные по размеру делеции и вставки, так и крупные структурные мутации длиной до 147 п.н. Также были выявлены две вставки-дубликации. Генетические мутации типа инверсии и рекомбинации выявлены не были. Структурные мутации были отмечены как в кодирующей, так и в некодирующей части генома.

Незначительные по размеру структурные мутации (инсерции и делеции), затрагивающие предполагаемые рамки, были выявлены как среди гипотетических ранее предложенных рамок считывания, так и среди изучен-

Рис. 1. Мозаичная структура вируса ИБК, полученная с помощью программы RDP



	14720	14730	14740	14750	14760	14770	14780	14790	14800	14810
Fowl adenovirus C	K D R K									
Fowl adenovirus C	· H ·									
Fowl adenovirus I	· H ·									
Fowl adenovirus C	· H ·									
Krasnodar De Novo	R H ·									
Krasnodar Mapper	R H ·									

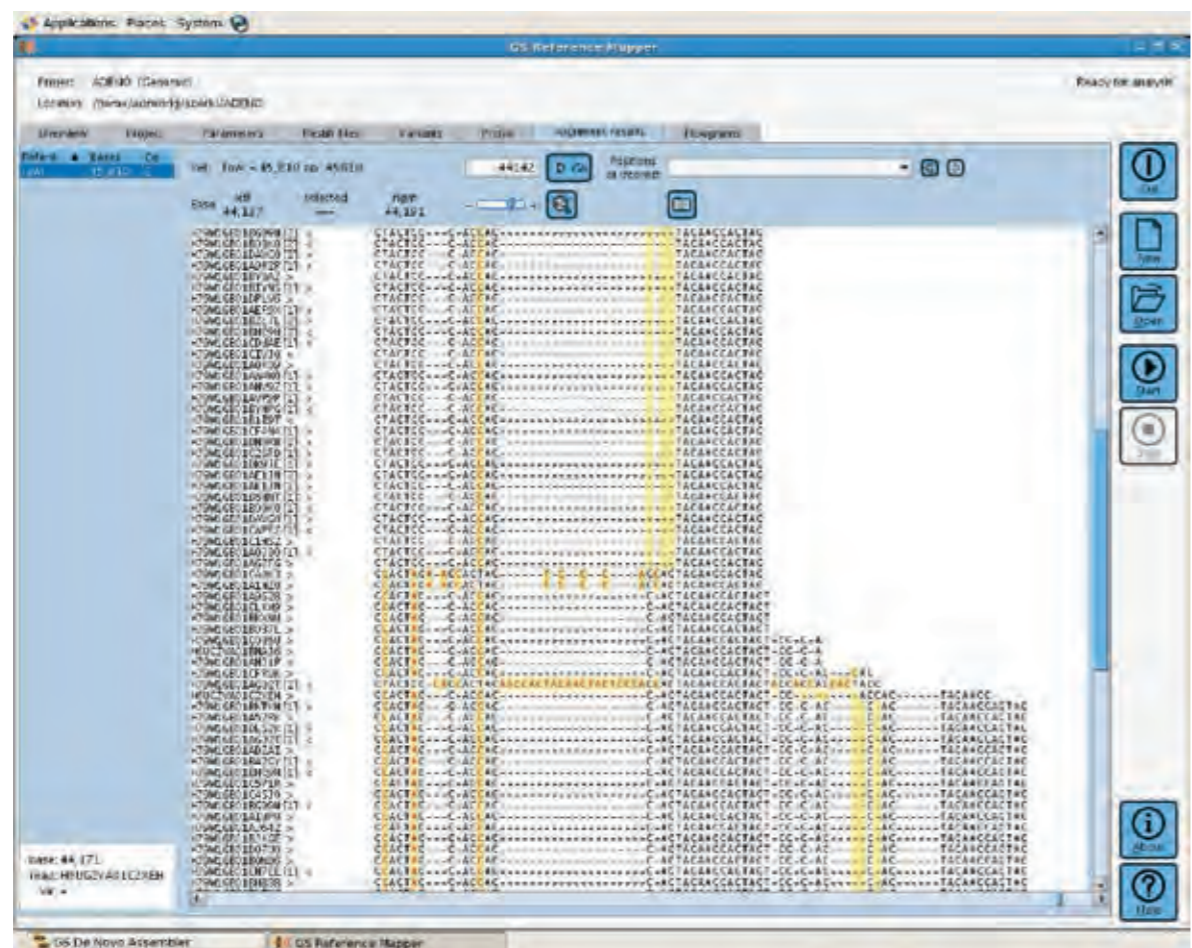
Рис. 2. Участок предсказанной аминокислотной последовательности белка 100 кДа аденовирусов птиц

ных протеинов (гены белков фибер-1, фибер-2, пентон, ДНК-полимераза).

При анализе полученного массива данных для участка генома, содержащего повторяющиеся генетические мотивы, были выявлены различия в полученных результатах. В частности, при анализе путем картирования на референтную последовательность в первичной нуклеотидной и предсказанной аминокислотной последовательности были выявлены различные одиночные замены, а также отсутствовала инсерция размером 60 п.н. (рис. 2).

Эти различия, по всей видимости, обусловлены иным алгоритмом обработки данных, используемым программой GS Reference Mapper. Здесь стоит отметить полезность функции визуализации полученных результатов в программах, используемых для анализа данных, поскольку выявленная гетерогенность исследованного участка статистически подкрепляется и могла бы быть расценена как истинная гетерогенность исследован-

Рис. 3. Скриншот результатов анализа массива данных с помощью программного обеспечения GS Reference Mapper



ного образца. Однако благодаря визуализации результатов можно отметить наличие характерной «обрезки» анализируемых прочтений по одной точке на всем массиве данных (рис. 3). Обнаружение подобной «обрезки» сборки указывает на потенциальные структурные изменения у исследуемого объекта (наличие инсерции, делеции, рекомбинации) и необходимость дополнительного анализа.

Кроме этого, при использовании алгоритма сборки *de novo* среди собранных контигов был выявлен фрагмент, который по результатам анализа был идентифицирован как аденоассоциированный вирус [2]. В результате сравнительного анализа собранного генома были выявлены характерные гены аденоассоциированных вирусов (ОРС неструктурного протеина «*rep protein*» и структурных белков «*cap protein*») [5]. Размер ОРС «*rep protein*» составил 1992 п.н. Процент сходства нуклеотидной последовательности ОРС «*rep protein*» составил 92, 95, 94% со штаммами VR-865 (AY629582), DA-1 (AY629583), YZ-1 (GQ368252) соответственно. Размер ОРС изолята ZN-1 «*cap protein*» составил 2232 п.н. Процент сходства нуклеотидной последовательности ОРС «*cap protein*» составил 87, 95, 95% со штаммами VR-865, DA-1, YZ-1 соответственно. Собранный геном изолята ZN-1 был депонирован в GenBank под номером KF937794. Таким образом, использование алгоритма сборки *de novo* является более целесообразным, по-

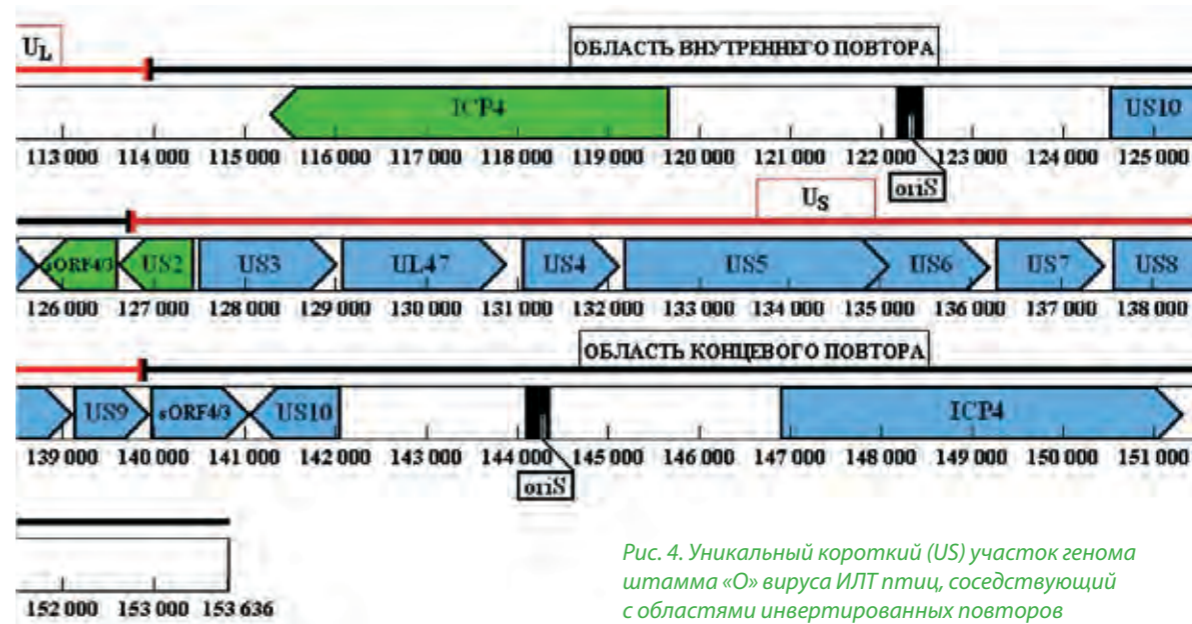


Рис. 4. Уникальный короткий (US) участок генома штамма «О» вируса ИЛТ птиц, соседствующий с областями инвертированных повторов

скольку позволяет исследовать вирусы со структурными мутациями, такими как рекомбинации, делеции, инсерции, а также позволяет выявлять вирусы-сателлиты.

Проведенные работы также показали возможность применения технологии 454 Life Sciences для анализа вирусов, в геноме которых содержатся инвертированные повторы. Вирус ИЛТ вызывает у кур острое респираторное заболевание. Возбудитель — ДНК-содержащий вирус семейства *Herpesviridae*, подсемейства *Alphaherpesvirinae*. Вирусные частицы имеют нуклеокапсиды диаметром 80–100 нм с икосаэдральной симметрией, полная вирусная частица диаметром 195–250 нм. Двухцепочечная линейная ДНК вируса размером 155 тысяч п.н. включает уникальные длинные и короткие участки, соседствующие с инвертированными повторами.

Для исследования был использован вакцинный штамм «О» вируса ИЛТ. Размер полученного генома составил 153636 п.н. Уникальный длинный (UL) и уникальный короткий (US) участки в ДНК составили 113921 п.н. и 13093 п.н. в длину соответственно. Уникальный короткий участок окружен внутренним (113922–126732 п.н.) и концевым (139826–153635 п.н.) инвертированными повторами в последовательности ДНК вируса ИЛТ птиц (рис. 4).

В областях инвертированных повторов находятся гены, кодирующие регуляторный белок ICP4, вирионный белок US10, а также *oriS* участки, которые имеют палиндромную организацию и содержат сайты инициации репликации вирусной ДНК [6]. При анализе результатов секвенирования вируса ИЛТ более приемлемым оказался алгоритм картирования полученных прочтений на референтную последовательность. Проведенная работа указывает на необходимость грамотного подбора алгоритмов последующего анализа данных, полученных с помощью технологии 454 Life Sciences, при осуществлении работ, связанных с расшифровкой вирусных геномов, имеющих свои специфические особенности.

ВЫВОДЫ

Полученные результаты показали возможность использования технологии 454 Life Sciences для исследования высоковариабельных в генетическом плане инфекционных агентов, проведения изучения вирусов, значительно отличающихся от референтных штаммов и имеющих межгеномные рекомбинации. Показано,

что для анализа полученных массивов данных целесообразно использование алгоритма сборки *de novo* при работе с вирусами, геном которых может содержать структурные мутации, повторяющиеся мотивы, геномные рекомбинации, тогда как для анализа вирусов с протяженными инвертированными геномными повторами более подходящим является алгоритм сборки картирования на референтную последовательность.

Таким образом, внедрение нового высокотехнологичного оборудования, использующего NGS-технологии, открывает дополнительные возможности в изучении генетических свойств вновь появляющихся и малоизученных микроорганизмов, а также возбудителей особо опасных болезней животных.

СПИСОК ЛИТЕРАТУРЫ

1. Генетический анализ изолята аденовируса птиц «Краснодар 2009» / О.С. Осипова, Н.Г. Зиняков, И.А. Чвала, В.В. Дрыгин // Вестник ветеринарии. — 2014. — № 3. — С. 33–37.
2. Зиняков Н.Г. Выявление аденоассоциированных вирусов с помощью NGS-технологии // Молекулярная диагностика. — 2014. — Т. 2. — С. 448–449.
3. Овчинникова Е.В. Молекулярно-биологические свойства изолятов вируса инфекционного бронхита кур, выявленных на территории России в период с 2005 по 2011 гг.: дис. ... канд. биол. наук. — Владимир, 2012. — 116 с.
4. Adair B.M., Fitzgerald S.D. Group 1 adenovirus infections // Diseases of Poultry / ed. Y.M. Saif, A.M. Fadly, J.R. Glisson [et al.]. — 12th ed. — Ames, Iowa, — 2008. — P. 252–266.
5. Bossis I., Chiorini J.A. Cloning of an Avian Adeno-Associated Virus (AAAV) and Generation of Recombinant AAAV Particles // J. Virol. — 2003. — Vol. 77. — P. 6799–6810.
6. Knopf C.W., Spies B., Kaerner H.C. The DNA replication origins of herpes simplex virus type 1 strain Angelotti // Nucleic Acids Research. — 1986. — Vol. 14, № 21. — P. 8655–8667.
7. Metagenomic analysis of the viromes of three North American bat species: viral diversity among different bat species that share a common habitat / E.F. Donaldson, A.N. Haskew, J.E. Gates [et al.] // J. Virol. — 2010. — Vol. 84. — P. 13004–13018.
8. Sequencing viral genomes from a single isolated plaque / J. DePew, B. Zhou, J.M. McCarrison [et al.] // J. Virol. — 2013. — Vol. 10 (181). — doi: 10.1186/1743-422X-10-181.