

ГЛАВА 1.1.7.

СТАНДАРТЫ ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ, БИОИНФОРМАТИКИ И ВЫЧИСЛИТЕЛЬНОЙ ГЕНОМИКИ

ВВЕДЕНИЕ

Высокопроизводительное секвенирование, биоинформатика и вычислительная геномика (HTS-BCG) должны использоваться в ветеринарии и исследованиях в сфере пищевой безопасности в соответствии со стандартами лабораторного тестирования, равно как и любая другая лабораторная процедура или любой другой лабораторный инструмент. HTS-BCG является относительно новой процедурой. Цель данной главы помочь лабораториям в ее использовании, определив стандарты, которые позволят включить данную методику в спектр методов, используемых в лаборатории, так, чтобы те, кто пользуется результатами данных исследований, относились к ним с доверием.

А. ОБЩИЕ ПОЛОЖЕНИЯ

Сведения, полученные в результате секвенирования, играют все более важную роль в диагностике и контроле микробных инфекций, также позволяют более подробно охарактеризовать инфекционные агенты, описать их потенциальные фенотипические свойства и эпидемиологию.

Более того, для полной идентификации и характеристики микроорганизма необходимо описание основных признаков генома. Для вирусов речь может идти о полном геноме, а для бактерий и паразитов могут быть описаны частичные последовательности. Хотя технологии секвенирования и развиваются довольно быстро, однако после разработки подходящих биоинформационных процедур в рабочем порядке и за довольно короткое время можно создать последовательности целого генома для более крупных микроорганизмов.

Описанные здесь стандарты применяют при генерировании сведений по геномным последовательностям во время расследований случаев заражения отдельных животных, в популяциях и в окружающей среде. Эти стандарты применяют для генерирования сведений, управления ими и использования в рамках принятых процедур ветеринарных расследований и в рамках лабораторной системы обеспечения качества.

В. ПРОВЕДЕНИЕ ВЕТЕРИНАРНЫХ РАССЛЕДОВАНИЙ

С ПРИВЛЕЧЕНИЕМ МЕТОДИКИ HTS-BCG

Данные о последовательностях микроорганизмов, генерированные методом HTS или методами метагеномики, являются лишь одним из инструментов, хотя и довольно мощным, способным принести пользу при проведении исследований в области ветеринарии и пищевой безопасности. Эксперты в указанной области должны проводить анализ данных последовательности. Интерпретировать сведения, относящиеся к изучению

болезни, должны ветеринары с надлежащей квалификацией в соответствии со стандартными требованиями к диагностике болезней животных.

Данные о последовательностях и данные анализа последовательностей для инфекционных болезней, ассоциированные со случаями, вспышками и расследованиями болезней животных и с расследованиями в области пищевой безопасности, которые проводит лаборатория, следует регистрировать и анализировать вместе с другой информацией, касающейся этих случаев и вспышек. Указанные сведения считают необходимой частью отчетности.

HTS-BCG можно применять для решения многих задач, связанных с выявлением и характеристикой инфекционных возбудителей, как в биологических материалах (образцы для диагностики или в рамках надзора), так и размноженных в культурах или взятых в качестве изолятов.

При проведении первичной диагностики пользователи методики должны сопоставить цели своего тестирования со стандартными целями тестирования, указанными в Главе 1.1.6. *Принципы и методы валидации диагностических тестов для инфекционных болезней*. HTS-BCG можно использовать, в том числе, в качестве подтверждающего теста для организмов, выявленных в ходе другого первичного анализа или для получения дополнительной характеристики таких организмов.

Кроме общих целей тестирования HTS-BCG предлагает и дополнительные возможности для:

- i) Выявления, идентификации и характеристики ранее неидентифицированных микроорганизмов;
- ii) Усовершенствования диагностики известных болезней;
- iii) Усовершенствования диагностики и повторного выявления эмерджентных болезней известной и неизвестной этиологии;
- iv) Разработки единых «универсальных» диагностических методов, способных идентифицировать любые потенциальные патогены;
- v) Одновременного и быстрого выявления многочисленных возбудителей болезней с многофакторной этиологией;
- vi) Активного изучения эволюционной динамики патогенов на фермах, на местном, государственном и международном уровнях;
- vii) Более глубокого понимания эпизоотологии инфекционных болезней и филогеографии инфекционных агентов;
- viii) Усиленной отслеживаемости инфекционных болезней и способов передачи патогенов, включая экспертно-криминалистическую эпидемиологию;
- ix) Более подробной характеристики «популяций» известных патогенов (например, релевантные штаммы неосновных подвидов, «ускользнувшие мутанты»), которая, в свою очередь, помогает совершенствовать состав вакцин, противовирусных препаратов и пр.;

- х) Обеспечения более крепких связей между генотипом и фенотипом патогена за счет полногеномной последовательности многочисленных штаммов (включая эталонные штаммы) одного возбудителя.

С. СТАНДАРТЫ ИСПОЛЬЗОВАНИЯ HTS-BCG

1. Выбор технологической платформы или услуги

Лаборатории могут установить HTS-BCG комплекс на своей территории, заключить договор на оказание услуг со сторонними организациями или подать образцы в специальные референтные центры.

Если лаборатория располагает своими собственными мощностями, то в ее распоряжение попадает несколько коммерчески доступных платформ для секвенирования с целью генерирования информации о последовательностях из тестовых образцов. Выбирать платформу следует исходя из поставленной цели или комбинации целей, перечисленных в Разделе В.

Особенно важно, чтобы выбранная технология соответствовала поставленной цели, подходила для получения информации о последовательностях по типам геномов, предназначенных для изучения. Следует также учитывать время, необходимое для проведения цикла секвенирования, включая подготовку пробы; вспомогательное оборудование в дополнение к основному устройству для секвенирования; капитальные расходы на закупку и установку необходимого оборудования, и стоимость ежегодных лицензий или договоров на оказание услуг, включая рекомендованный производителем график технического обслуживания; доступную экспертную помощь от поставщика; стоимость реagenтов для одного цикла секвенирования и вероятную доступность reagenтов в соответствующей стране; требования к персоналу и тренинги по использованию оборудования и проведению ассоциированных биоинформационных анализов; сведения о требованиях к управлению.

Уже изучены имеющиеся в настоящее время системы (Belak et al., 2013; Granberg et al., 2016; Marston et al., 2013), но следует ожидать появления новых моделей и технологий.

Если лаборатория или ветеринарные службы заключают договор о проведении HTS-BCG со сторонней организацией, то необходимо обеспечить, чтобы провайдер услуг выполнял требования, указанные в этой главе.

2. Отбор проб и отчетность

HTS-BCG – новая технология контроля болезней животных и использовать ее следует в контексте опробованных и принятых технологий контроля здоровья животных и пищевой безопасности, включая клинические и эпизоотологические полевые испытания и отбор проб у животных, в популяциях животных; или в контексте других эпизоотологически значимых ситуаций.

Применение технологии должно подходить целям расследования, стратегия отбора проб и сами пробы должны соответствовать расследованию с точки зрения патогенеза и эпизоотологии изучаемой болезни или вероятного патогенеза и эпизоотологии любого нового возбудителя инфекционной болезни. Такие расследования следует проводить под надзором ветеринаров, имеющих необходимую квалификацию.

В лабораториях, где используется HTS-BCG, управлять данной технологией следует в рамках лабораторной системы обеспечения качества. Следовательно, результаты HTS-

BCG следует интерпретировать в контексте патогенеза и эпизоотологии инфекционной болезни у изучаемых видов животных. Результаты тестирования сообщает ветеринарный специалист с соответствующей квалификацией, наделенный полномочиями по проведению диагностики в рамках лабораторной системы обеспечения качества и по месту проведения расследования.

3. Образцы и подготовка проб

Образцы следует собирать и поставлять в испытательную лабораторию в соответствии со стандартами, изложенными в Главе 1.1.2 *Сбор, предоставление и хранение диагностических образцов*. При оформлении образца, подаваемого в лабораторию, следует указывать информацию по конкретному животному, случаям болезни или указать основания для отбора проб вместе с соответствующей эпизоотологической информацией.

Критически важными являются лабораторные процессы, обеспечивающие целостность проб и образцов, подлежащих тестированию. Из образцов следует выделить нуклеиновые кислоты, либо ДНК, либо РНК. В некоторых случаях можно воспользоваться стратегиями обогащения для увеличения соотношения патогена к нуклеиновым кислотам хозяина, что позволит максимально повысить чувствительность метода. Однако следует избегать системных погрешностей в результатах с учетом предполагаемой области применения. Необходимо соблюдать меры предосторожности для обеспечения целостности и качества нуклеиновых кислот по аналогии с другими молекулярными методиками (например, полимеразная цепная реакция (ПЦР)), как было ранее описано в Руководстве 3.2. *Биотехнологии в диагностике инфекционных болезней*¹. После выделения нуклеиновых кислот из образцов их следует обработать (например, обратная транскрипция РНК в комплементарном ДНК) таким образом, чтобы использовать их в HTS. Различные технологические платформы требуют специфических наборов реагентов для создания конечного материала (библиотеки) для секвенирования.

Для этих целей существуют коммерческие наборы.

HTS является чрезвычайно чувствительной технологией, способной выявить даже незначительное количество нуклеиновых кислот. Поэтому следует принять меры предосторожности для недопущения перекрестной контаминации, как и в случае с другими молекулярными методами, используемыми для выявления нуклеиновых кислот (например, ПЦР). Важным требованием является исключение перекрестной контаминации в рабочих зонах. Кроме того, HTS довольно часто предполагает «мультиплексирование» нескольких проб за одну реакцию. Отдельные пробы «маркируют биркой» на одной из стадий пробоподготовки, используя для этого краткую последовательность индексов, связанных с молекулами нуклеиновой кислоты.

Последовательности индексов должны иметь соответствующие качество и вид, на которые можно сослаться, как на характерные символы последовательностей из библиотеки с маркированными образцами для использования в HTS. Это позволит избежать помех во время биоинформационного анализа сведений, полученных после секвенирования.

При каждом цикле HTS-BCG необходимо использовать положительные и отрицательные контроли, подходящие для расследования и включенные во все этапы процесса пробоподготовки цикла секвенирования, и для фактического цикла на технологической платформе. Необходимо использовать подходящие контроли для верификации каждого

¹ См. http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.2_BIOTECH_DIAG_INF_DIS.pdf

этапа процедуры, включая качество нуклеиновой кислоты, подготовку библиотеки, перекрестную контаминацию (включая мультиплексирование), чувствительность и воспроизводимость.

Как и для других диагностических методов, для подтверждения результатов необходимо повторно отобрать пробы от оригинального образца, который следует хранить в надлежащих условиях, не допускающих перекрестную контаминацию.

4. Генерирование данных последовательности

Поскольку платформы HTS существенно отличаются по ряду своих характеристик, то необходимо следовать основным принципам контроля технологии и разработать универсальные рекомендации для принятой системы контроля качества. Надлежащие меры контроля включают: использование циклов положительных, отрицательных контролей и контроля без матрицы в параллельных опытах; использование балльной системы оценки качества.

Метрические показатели качества секвенирования представляют собой подходящие параметры для валидации и мониторинга производительности платформы. Большинство платформ располагает возможностью обогащать контроли в реагентах и использовать метрические показатели контроля качества для мониторинга производительности платформы и реагента. Можно использовать дополнительные метрические показатели производительности специфической технологии для мониторинга производительности платформы и идентификации ошибочных циклов секвенирования.

Метрические показатели качества для оценки аналитической производительности тестов на основе HTS:

- i) Глубина покрытия. Показывает количество считываний последовательностей, предоставляя информацию об указанном нуклеотиде. Когда текущий мониторинг качества показывает, что глубина покрытия данного нуклеотида ниже минимального валидированного покрытия, необходимо предоставить подтверждение с использованием альтернативных методов (например, последовательность Сэнгера) или дополнительного секвенирования.
- ii) Равномерность покрытия. Данный параметр показывает, насколько распределена глубина покрытия по целевым участкам теста. Отклонения равномерности покрытия от валидированного диапазона потенциально указывает на погрешности в процессе тестирования.
- iii) GC-сдвиг. Содержимое GC (относительный избыток G и C нуклеотидов) целевого участка влияет на эффективность секвенирования и в итоге повлияет на равномерность покрытия. При возможности степень сдвига GC в целевых участках теста следует определять во время валидации и контролировать для оценки производительности теста.
- iv) Балльная оценка качества базового вызова. Отражение отношений сигнал/помехи на платформе, которое показывает вероятность правильности базового вызова. Допустимый порог качества необработанного базового вызова следует установить во время валидации и включить в биоинформационные фильтры для недопущения получения недостаточных сведений о качестве во время анализа.

- v) Снижение интенсивности сигнала или считанная длина. В зависимости от конкретного применения, платформы HTS и химических компонентов, считывания последовательностей имеют типичное распределение считанной длины и интенсивность сигнала. Ожидаемую интенсивность сигнала на протяжении считывания (или распределения считанной длины) следует устанавливать во время валидации и контролировать в каждом цикле. Отклонения в распределении или считанных длинах могут указывать на проблемные наборы данных.
- vi) Качество картирования. Мера неопределенности, показывающая, что считывание отображено на карте надлежащим образом относительно геномного положения в пределах целевого региона. Во время валидации биоинформационных операций следует установить допустимые показатели (а именно пропорцию считываний, отображающую целевой участок); а в рамках каждого цикла можно контролировать пропорцию считываний, не отображающих целевой участок.
- vii) Внутренние проверки. Большинство платформ предлагает возможность обогащать внутренний контроль на чрезвычайно низкой частоте во время цикла секвенирования. Метрические показатели качества этих считываний можно сравнить с ранее указанными метрическими показателями качества.

5. Биоинформатика

Любая лаборатория, намеревающаяся установить HTS-BCG комплекс, обязана применять специализированные методы биоинформатики. Использование указанных пакетов программ не исключает ответственности лаборатории при проведении компетентного анализа своих собственных данных, даже если в наличии должны были появиться платформы со вспомогательным ПО для специфических анализов в определенных клинических ситуациях.

Критически важными элементами в HTS-BCG являются биоинформационный анализ, предполагающий сбор геномной последовательности патогена из совокупности первоначальных данных, и последующий вторичный анализ. Следовательно, используемые подходы должны быть прозрачными, а описание используемых пакетов программного обеспечения, версий программного обеспечения и референтных баз данных или последовательностей должно быть неотъемлемой частью каждого отчета по секвенированию. Используемое для данных анализов программное обеспечение должно быть доступно (с соблюдением коммерческих интересов или находиться в открытом доступе), чтобы международное сообщество имело возможность провести его оценку.

Как и в отношении других лабораторных процедур, особое внимание следует уделять обеспечению качества. Метод тестирования должен включать критерии для принятия или отклонения каждого цикла на основе удовлетворительных результатов анализов контрольных образцов. Данные по секвенированию должны быть внесены в документы, позволяющие отобразить минимальные показатели качества и информацию по каждому нуклеотиду собранной, окончательной консенсусной последовательности.

Соответствие выбранного ПО для биоинформационных методов исследования конкретным анализам можно оценить путем сопоставления показателей его производительности стандартным наборам данных, содержащих сведения, касающиеся агентов, наличие которых предполагается в тестируемых образцах.

6. Управление данными

Созданные в ходе HTS-BCG операций сведения представляют важность для диагностических или иных целей расследований, таких как характеристика возбудителя, и являются неотъемлемым компонентом указанного процесса. По этой причине одним из важнейших обязательств лабораторий является наличие стратегий, процессов и поддерживающих систем для отбора, управления и хранения созданных данных.

Различные технологические платформы HTS производят первичные данные в различных форматах и на стадии предварительного анализа, поэтому важно, чтобы в лабораториях были стратегии и процессы работы с данными в зависимости от используемой технологической платформы.

Системы управления данными должны включать позиции, по которым необходимо сохранять сведения, и в этих системах должно быть указано время, в течение которого данные следует хранить. Необходимо наличие стратегий резервного хранения, позволяющих не допустить внезапные потери данных или намеренное их уничтожение. Необходимым условием является наличие метаданных, описывающих создание и анализ последовательностей данных, что делает возможным проанализировать и повторить сам процесс.

Если результаты анализа последовательности имеют значение для защиты здоровья животных, особенно если это касается вопросов торговли или имеет международное значение, то сведения, использованные при проведении анализа, должны в обязательном порядке быть доступны при проверке или при проведении подтверждающих анализов.

Сведения должны храниться в течение времени, которое сопоставимо по важности с результатами, касающимися здоровья животных. Неспособность произвести требуемые данные для независимого анализа, может быть основанием для признания полученных результатов недействительными.

Данные по секвенированию следует хранить таким образом, чтобы была выстроена четкая связь с метаданными, касающимися тестируемого образца. Исходя из стандартной практики проведения лабораторных последовательностей, к метаданным относят информацию о животном, от которого взят образец, его владельце и месте нахождения; также сопутствующую клиническую и эпидемиологическую информацию о популяции животных.

7. Валидация тестовых систем для конкретных целей

Концепции валидации тестов, как указано в Главе 1.1.6, широко применимы к HTS-BCG комплексу. Все процедуры, включая обработку образцов (выделение нуклеиновой кислоты, подготовку библиотеки, прикрепление бирки, целевое обогащение), секвенирование, биоинформатику и составление отчетности, должны быть изложены в виде СОПов еще до начала процедуры валидации. Необходимо сформировать данные на первом этапе валидации для подтверждения аналитической чувствительности (Se), специфичности (Sp) и воспроизводимости методики. Что касается анализов на основе секвенирования, то аналитическую чувствительность можно определить как вероятность того, что анализ позволит выявить разновидности целевых последовательностей, если таковые есть при заданном уровне вероятности (например, 95% доверительный уровень). Аналитическую специфичность можно определить как вероятность того, что анализ не позволит выявить вариацию последовательности, если таковая отсутствует при заданном

уровне вероятности. Более того, у каждого образца есть свои особенные характеристики, которые также необходимо учитывать. Например, особенности связанные с типами проб: носовой смыв, сыворотки или фекалии.

Для оценки производительности анализа можно воспользоваться хорошо описанными пробами целевого и нецелевого аналита и матричных компонентов известной концентрации. Как минимум сюда относят серийные разведения каждого вида образца, содержащего определенные организмы, чтобы зарегистрировать пределы выявления обозначенных цельных геномов или генетических последовательностей, характерных для типа, в отношении которого в лаборатории будут использоваться мощности HTS-BCG.

При расследовании в отношении вирусных болезней тестовые пробы можно подготовить так, чтобы они содержали репрезентативные вирусы из всего диапазона вирусных семей, представители которых могут присутствовать в тестовых пробах, подлежащих расследованиям в ходе рутинных операций. Будет создана документация по HTS-BCG системе для выявления этих вирусов. Аналогичные принципы применяют для генетических маркеров, бактерий или иных организмов, в отношении которых применяют HTS-BCG систему в рамках спланированных рутинных операций. В ходе всех указанных циклов, рассчитанных на выявление чувствительности и специфичности, частью каждой оценки должны быть этапы подготовки проб, поскольку они могут быть критически важными для всех аспектов производительности теста.

Ряд факторов осложняет валидацию NGS тестов, как первичных диагностических тестов, а именно:

- i) Весовой коэффициент валидации аналитических и диагностических методик (Глава 1.1.6);
- ii) Производственные издержки, связанные с применением технологии;
- iii) Сложности валидации схемы анализа данных;
- iv) Высокая необходимость инвестирования в техническое оснащение и повышение профессиональной компетенции сотрудников
- v) Время, необходимое для получения результата (в сопоставление дни-часы для специфических методов молекулярной диагностики, например, для ПЦР в реальном времени).

Дополнительные подтверждающие или вторичные диагностические тесты необходимо валидировать только по показателям их аналитической производительности, например, чувствительность и специфичность, повторяемость и первичная воспроизводимость (этап 1). Валидировать по всему спектру диагностических показателей не следует (диагностическая чувствительность и специфичность, этап 2).

Следует учитывать, что с практической точки зрения не всегда целесообразно производить большие объемы данных по показателям производительности тестов, которые обычно позволяют проводить расчеты чувствительности и специфичности диагностического теста, но при этом необходимо принять к рассмотрению другие аспекты валидации, такие как демонстрация воспроизводимости теста среди лабораторий, проводящих схожие расследования.

8. Обеспечение качества

Использовать комплекс HTS-BCG с целью расследования случаев болезней животных и в ситуациях, связанных с пищевой безопасностью, следует в соответствии с применяемой в лаборатории системой обеспечения качества, отличительные характеристики которой соответствуют стандартам, перечисленным в Главе 1.1.5. *Управление качеством в*

ветеринарных лабораториях. Если лаборатория аккредитована, то тестирование должно стать частью сферы аккредитации лаборатории.

Были разработаны наборы стандартных данных, на основе которых можно проверить практическую значимость программного обеспечения для биоинформационных методов. Лаборатории, пользующиеся HTS-BCG, должны гарантировать, что применяемые ими пакеты ПО для биоинформационных методов, соответствуют критериям производительности в сопоставлении со стандартами данных.

Если разработаны стратегии квалификационных испытаний, то лабораториям, применяющим HTS-BCG, следует участвовать в них.

9. Интерпретация результатов

Комплекс HTS-BCG можно использовать для различных целей, начиная с выявления патогенов, заканчивая диагностикой или подробной характеристикой возбудителей инфекционных болезней. Следовательно, интерпретировать полученные результаты необходимо в контексте особой клинической и эпизоотологической ситуации, при условии удовлетворительных показателей производительности, полученных по всем специфическим контролям и параметрам обеспечения качества. Как и в случае с другими лабораторными тестами, упомянутые параметры являются одними из множества тех параметров, которые необходимо учитывать.

Список литературы

BELÁK S., KARLSSON O.E., LEIJON M. & GRANBERG F. (2013). High-throughput sequencing in veterinary infection biology and diagnostics. *Rev. sci. tech. Off. int. Epiz.*, 32 (3), 893–915.

GRANBERG F., BÁLINT, A. & BELÁK, S. (2016). Novel technologies applied for the nucleotide sequencing and comparative sequence analysis of the genomes of infectious agents in veterinary medicine. *OIE Sci. Tech. Rev.*, 35 (1), (in press).

MARSTON D.A., MCELHINNEY L.M., ELLIS R.J., HORTON D.L., WISE E.L., LEECH S.L., DAVID D., DE LAMBALLERIE X. & FOOKS A.R. (2013). Next generation sequencing of viral RNA genomes. *BMC Genomics*, 14, 444.

VAN BORM S., WANG J., GRANBERG F. & COLLING A. (2016). Towards validation and quality assurance of next-generation sequencing workflows in veterinary infection biology. *OIE Sci. Tech. Rev.*, 35 (1), April 2016 (in press).

*

* *