Формирование биологического банка данных на основании проведения молекулярно-генетического мониторинга возбудителя ретровирусной инфекции

Creation of a biological data bank on the basis of molecular genetic monitoring of the pathogen of retroviral infection

Козырева Н.Г.
г. Москва

DOI CrossRef: 10.30917/ATT-VK-1814-9588-2018-7-14
УДК 619:616-07; 577.21

Ключевые слова: база данных нуклеотидных последовательностей, вирус лейкоза крупного рогатого скота (ВЛКРС), крупный рогатый скот (КРС), ПЦР, секвенирование, филогенетический анализ.

Key words: nucleotide sequence database, cattle, bovine leukaemia virus (BLV), PCR, sequence analysis, phylogenetic analysis.

Резюме. Охарактеризована база данных ВИЭВ (БД ВИЭВ): обработаны массивы генетической информации (186 нуклеотидных последовательностей, полученных на основании результатов филогенетического анализа env, pol генов вируса лейкоза крупного рогатого скота (ВЛКРС)), данных по животным, проведено наполнение информацией и формирование основных объектов БД. Определена гетерогенность исследуемой популяции ВЛКРС по гену env: установлена циркуляция 5 генотипов вируса лейкоза крупного рогатого скота (ВЛКРС) – IG, IIG, IVG, VIIG, VIIIG с доминирующим IVG вариантом; выявлены внутригрупповые эволюционные расстояния. Степень внутригрупповой дивергенции по гену env находится в диапазоне от 0,005 до 0,012 в зависимости от генетического варианта ВЛКРС; по гену pol – в среднем 0,018. В международный генетический банк депонировано 86 нуклеотидных последовательностей.

Summary. The VIEV database was characterized: arrays of genetic information were processed (186 nucleotide sequences obtained on the basis of phylogenetic analysis of the env, pol genes of the bovine leukemia virus (BLV)), processed data files for animals, information filling and formation of the main objects of the database. Conducted phylogenetic analysis based on the env gene made it possible to evaluate the heterogeneity – circulation of 5 genotypes of bovine leukemia virus (BLV) – IG, IIG, IVG, VIIG, VIIIG with the dominant IVG variant was established and based on the pol gene allowed to estimate the homogeneity of the studied BLV population; revealed intragroup evolutionary distances. The degree of intragroup divergence for the env gene is in the range from 0,005 to 0,012, depending on the genetic variant of BLV, for the pol gene – an

average of 0,018. 86 nucleotide sequences were deposited in the international genetic bank GenBank.

Введение

Анализ генетического разнообразия популяции ВЛКРС, циркулирующей на территории РФ с последующим выявлением родственных связей между различными вариантами возбудителя и создание БД охарактеризованных нуклеотидных последовательностей изолятов возбудителя в современной ситуации внедрения компьютерных технологий в рамках молекулярной эпизоотологии приобретает все более важное и актуальное значение [1,2,3,4,5]. ВЛКРС является представителем дельтаретровирусов (наряду с Т-лимфотропными вирусами приматов (PTLV) – человека и обезьян (HTLV и STLV)) семейства Retroviridae и индуцирует злокачественное лимфопролиферативное заболевание – энзоотический лейкоз – медленную ретровирусную инфекцию, широко распространенную по всему миру [7].

Для молекулярно-эпизоотологического анализа исследуются генетические последовательности области поверхностного гликопротеида gp51 ВЛКРС, что связано со значительной изменчивостью этого участка генома вируса. На сегодняшний день в мире выявлено 10 генетических вариантов ВЛКРС на основе гена env [7]. Ген pol является наиболее консервативным участком в составе генома ВЛКРС. Результаты наших экспериментов показали, что вариативность изученного ранее целевого участка гена pol ВЛКРС не превышала 3% [4].

Объектом исследования является генетическая информация – нуклеотидные последовательности изолятов ВЛКРС.

Цели работы: наполнение, формирование, анализ собственной БД ВИЭВ по нуклеотидным последовательностям изолятов ВЛКРС.

Материалы и методы

Для статистической обработки данных, формирования БД использовали алгоритмы компьютерного приложения Microsoft Office Excel 2003. Возможен экспорт массивов данных таблиц в Access и использование такого алгоритма в дальнейшем для построения ГИС (геоинформационной системы) с применением программного обеспечения ArcGIS [3]. Работа с генетической информацией проводилась с помощью программы BioEdit v.7, алгоритмов сервиса BLAST, филогенетический анализ выполняли с помощью программы Mega 5. Дендрограммы построены дистанционными методом присоединения соседей NJ (neighbor-joining method; Saitou, Nei, 1987) с использованием модели р-дистанций. Статистическую достоверность топологии деревьев оценивали с помощью метода бутстрэп-анализа при количестве случайных выборок 1000. Расчет эволюционных дистанций произведен с использованием модели Джукса-Кантора (Jukes, Cantor, 1969).

Для сравнения использовали некоторые референсные нуклеотидные последовательности целевых генов ВЛКРС, представленные в международном БД GenBank, ресурса NCBI (<http://www.ncbi.nlm.nih.gov>). Генотипы ВЛКРС дифференцировали согласно классификации, предложенной S.M. Rodriguez et al., 2009 с последующими дополнениями.

Результаты и обсуждения

Наполнение основных объектов БД нуклеотидных последовательностей ВЛКРС осуществлено на основании полученных молекулярных данных по результатам ранее проведенного филогенетического анализа целевых генов (env, pol: 148 и 38 изолятов, соответственно), которые в дальнейшем были систематизированы, структурированы и преобразованы в актуальную версию. БД представляет упорядо-

Авторы / Authors

Козырева Н.Г., кандидат биологических наук, ведущий научный сотрудник лаборатории лейкозологии, nk07-73@mail.ru
Иванова Л.А., кандидат биологических наук, ведущий научный сотрудник лаборатории лейкозологии,
Kozyreva N.G., Ivanova L.A.
ФГБНУ ФНЦ "Всероссийский научно-исследовательский институт экспериментальной ветеринарии имени К.И. Скрябина и Я.Р. Коваленко РАН" (ФГБНУ ФНЦ ВИЭВ РАН), Москва, Россия.

ченную определенным образом информацию в рамках заданной структуры (Табл. 1). Общий массив БД ВИЭВ составляет 186 нуклеотидных последовательностей с присвоенными каждой уникальным названием и номером.

Мониторинговые исследования полиморфизма ВЛКРС на основе гена *env*, циркулирующего в 14 регионах РФ, проводились на крупном рогатом скоте в возрасте от 7 месяцев до 11 лет 7 разных пород [2,3], что позволило оценить гетерогенность исследуемой популяции ВЛКРС. Так, на территориях Пензенской, Ростовской, Рязанской, Новосибирской, Московской областей обнаружена циркуляция более одного (2–4) генотипа ВЛКРС вируса. Для Пензенской области характерно присутствие 2-х (IV, VII), для Ростовской 3-х (I, II, IV) и для Рязанской 2-х (II, IV) генотипов ВЛКРС. В остальных регионах выявлено по одному генотипу вируса. Степень дивергенции между последовательностями исследуемых изолятов из БД ВИЭВ и известных мировых референтных штаммов составляет от 0,005 до 0,012 (0,5–1,2%) в зависимости от генетического варианта ВЛКРС. Минимальные значения внутригрупповых дистанций наблюдается для II генотипа, максимальные – для I, что коррелирует с данными других исследователей [9] и VIII генотипов ВЛКРС. При изучении генетического разнообразия ВЛКРС выявлено 5 генотипов вируса – IG (5%), IIG (7%), IVG (75%), VIIG (11%), VIIIG (1%) – основная доля циркулирующих провирусов приходится на доминирующий IVG генотип, так называемый "европейский кластер".

По результатам филогенетического анализа на основе гена *pol* была установлена гомогенность исследуемой популяции ВЛКРС. Исследуемые изоляты из регионов РФ кластеризуются с изолятами, к которым относятся также штаммы из Аргентины, Австралии, Японии, США, Бразилии. Тем не менее, исследуемые изоляты отличаются от штаммов из США (M10987 BLV GAGA), Северной Италии (S83529), Уругвая (HE967303), а также США (NC001414, AF033818), России и Украины, Уругвая (HE967301), составляющих отдельные клады. Степень дивергенции внутри группы в среднем составляет 0,018 (1,8%), что соотносится с данными других исследователей [8].

На сегодняшний день в международный генетический банк данных депонированы охарактеризованные нуклеотидные последовательности генов ВЛКРС: *env* (как частичные участки, так и последовательности полного гена) – 77 и локусов гена *pol* – 9.

Преобладание IVG ("европейский кластер") является отражением длительной циркуляции на территории России варианта ВЛКРС, изначально возможно попавшего с территории Восточной Пруссии (где лейкоз КРС был выявлен еще в 1916 году) при вхождении районов Мемеля (Клайпеда) и Кёнигсберг (с 1946 г. Калининград) в состав СССР [6], а также ввезенного из стран Европы с инфицированным крупным рогатым скотом после второй мировой войны в результате проведения торговых операций [7].

К сожалению, не имея информации о движении скота, сегодня нельзя определить причину появления необычных вариантов ВЛКРС и дифференцировать это как результат мутагенеза давно циркулирующего в данном регионе возбудителя или завоза с импортным скотом. В большинстве случаев для выявленных на территории РФ генетических вариантов ВЛКРС, как правило, отсутствуют данные по происхождению и перемещению инфицированного скота. Обладая информацией о движении скота как внутри страны, так и в случае покупки скота из неблагополучных по лейкозу стран нам удавалось устанавливать принадлеж-

ность полученных изолятов вируса. Интересно, что результаты филогенетических анализов на основе двух целевых генов в некоторых случаях коррелировали между собой [1,3].

Создание с 2009 г. собственной компьютерной БД ВИЭВ по нуклеотидным последовательностям изолятов ВЛКРС имеет важное научное и практическое значение. Возможность быстрого получения и анализа большого количества генетической информации при изучении распространения инфекционных заболеваний с последующей разработкой мер по их предотвращению и ограничению позволяет решать традиционные задачи эпизоотологии на молекулярном уровне.

Выводы

Проведено наполнение собственной БД ВИЭВ нуклеотидных последовательностей изолятов ВЛКРС, распространенных на территории РФ с определением для каждого генетического статуса на основе филогенетического анализа по генам *env*, *pol* – общий массив составляет 186 изолятов ВЛКРС. Сформирована собственная БД ВИЭВ, которая упорядочена определенным образом в рамках заданной структуры (Табл. 1).

Проанализировано распределение охарактеризованных генетических вариантов ВЛКРС на обследованных территориях некоторых регионов РФ – выявлено 5 генотипов вируса – IG (5%), IIG (7%), IVG (75%), VIIG (11%), VIIIG (1%) с доминирующим IVG вариантом. Определены генетические гетерогенность по гену *env* и гомогенность по гену *pol* анализируемой популяции возбудителя. В международный банк данных депонировано 86 охарактеризованных последовательностей генов *env*, *pol* изолятов ВЛКРС.

Финансирование: данная работа выполнена в рамках государственного задания по теме 0578-2014-0004 "Сформировать базу данных генетического полиморфизма вариантов вируса лейкоза крупного рогатого скота, распространенных на территории России". **Конфликтные интересы:** авторы статьи заявляют об отсутствии конфликтов интересов.

Список литературы

1. Гулюкин М.И., Иванова Л.А., Генджиева О.Б., Козырева Н.Г. Генотипирование изолятов ВЛКРС, распространенных на территории республики Калмыкия. Ветеринария Кубани, 2012, 4: 4-7.
2. Гулюкин М.И., Козырева Н.Г., Иванова Л.А. [и др.]. Генетический полиморфизм вируса лейкоза крупного рогатого скота на территории РФ. Российская сельскохозяйственная наука, 2016, 5:56-59.
3. Козырева Н.Г., Гулюкин М.И. Распространение лейкоза крупного рогатого скота и генетические варианты возбудителя на территории животноводческих хозяйств Центрального федерального округа Российской Федерации. Ветеринария Кубани, 2017, 6: 4-9.
4. Козырева Н.Г., Гулюкин М.И., Иванова Л.А. [и др.]. Филогенетический анализ участка гена *pol* изолятов провируса лейкоза КРС, обнаруженных у животных из различных хозяйств регионов Российской Федерации. Доклады Российской Академии с.- х. наук, 2011, 6: 48-51.
5. Козырева Н.Г., Иванова Л.А., [и др.] Определение генетического статуса изолятов ВЛКРС, распространенных на территории Московской области ЦФО РФ. Сб. тр. IX Всерос. науч.-практ. конф. с междунар. участием "Молекулярная диагностика - 2017". М., 2017, Т.2: 394-395.
6. Лейкозы и злокачественные опухоли животных [Текст] / Под ред. В.П. Шишкова, Л.Г. Бурбы. М., Колос, 1977, 375 с.
7. Polat M., Takeshima S.N., Aida Y. Epidemiology and genetic diversity of bovine leukemia virus. J. Virol., 2017, 14:209 (doi: 10.1186/s12985-017-0876-4).
8. Polat M., Takeshima S.-n., Hosomichi K. et al. A new genotype of bovine leukemia virus in South America identified by NGS-based whole genome sequencing and molecular evolutionary genetic analysis. J. Retrovirology, 2016, 13:4: 23p. URL: <http://retrovirology.biomedcentral.com/articles/10.1186/s12977-016-0239-z.pdf>.
9. Rola-Luszczak M., Pluta A., Olech M. et al. The molecular characterization of bovine leukaemia virus isolates from Eastern Europe and Siberia and its impact on phylogeny. PLoS ONE: electronic resource, 2013, 8(3). DOI: 10.1371/journal.pone.0058705.

Таблица 1. Структура базы данных нуклеотидных последовательностей ВЛКРС ВИЭВ

N n/p	Информация об изоляте										Информация о животном, от которого выделен изолят					Упоминание информации об изоляте						
	Название изолята	per N в GenBank (NCBI)	Год регистрации в GenBank (NCBI)	Год взятия материала	Год выделения изолята	Ген провируса	Генетический статус	Источник выделения (биоматериал)	Страна	Область, регион	Шифр	Имя, N/клика животного	Вид животного	Возраст, г/р	Порода	Условия содержания	Хозяйство	Примечание	в публикации		в отчетности	
																			да/нет	год	да/нет	год