

Я.М.Краснов¹, В.А.Сафронов¹, Н.Ю.Носов¹, В.В.Кутырев¹, А.Ю.Попова^{2,3}**СРАВНИТЕЛЬНЫЙ АНАЛИЗ СЕКВЕНИРОВАННЫХ ОБРАЗЦОВ ВИРУСА ZAIRE EBOLAVIRUS ИЗ ГВИНЕЙСКОЙ РЕСПУБЛИКИ**¹ФКУЗ «Российский научно-исследовательский противочумный институт «Микроб», Саратов, Российская Федерация; ²Федеральная служба по надзору в сфере защиты прав потребителей и благополучия человека, Москва, Российская Федерация; ³Российская медицинская академия последипломного образования, Москва, Российская Федерация

В 2013–2015 гг. в Западной Африке зарегистрирована крупнейшая в истории эпидемия лихорадки Эбола, вызванная вирусом вида *Zaire ebolavirus*. Из Гвинейской Республики вирус быстро распространился на соседние страны. В статье представлены результаты сравнительного анализа трех геномов *Zaire ebolavirus*, кДНК которых получена из клинического материала от больных лихорадкой Эбола с лабораторно подтвержденным диагнозом на территории Гвинейской Республики в 2014 г. Полученные нуклеотидные последовательности геномов позволили установить их филогенетическую близость к определенным вариантам, представленным в базе данных NCBI GenBank. Общий филогенетический анализ с использованием репрезентативной выборки показал два направления эволюционного развития генома от одного исходного предка. Первая ветвь эволюции представлена относительно небольшим числом изолятов *Zaire ebolavirus*, получивших распространение в Гвинее. Второе, наиболее многочисленное направление эволюции, вызвано 4 вариантами генома *Zaire ebolavirus*, которые первоначально начали выделяться в конце мая – начале июня 2014 г. в Сьерра-Леоне. За короткий период времени они охватили территорию Сьерра-Леоне, Гвинеи и Либерии.

Ключевые слова: геном возбудителя БВВЭ, лихорадка Эбола, секвенирование, филогенетический анализ, Гвинейская Республика, Западная Африка.

Ya.M.Krasnov¹, V.A.Safronov¹, N.Yu.Nosov¹, V.V.Kutyrev¹, A.Yu.Popova^{2,3}**Comparative Analysis of the Sequenced *Zaire ebolavirus* Samples from the Republic of Guinea**¹Russian Research Anti-Plague Institute "Microbe", Saratov, Russian Federation; ²Federal Service for Surveillance in the Sphere of Consumers Rights Protection and Human Welfare, Moscow, Russian Federation; ³Russian Medical Academy for Post-Graduate Training, Moscow, Russian Federation

An unprecedented on its scale Ebola fever epidemic, caused by *Zaire ebolavirus*, has been registered in the territory of West Africa in 2014–2015. The virus quickly spread from the Republic of Guinea into the neighboring countries. Presented are the results of comparative analysis between three *Zaire ebolavirus* genomes, the complementary DNA of which was obtained from the samples from Ebola fever patients in the Republic of Guinea in 2014. The nucleotide sequences of the genomes received have allowed for identification of phylogenetic affinity to certain references, displayed in the database of NCBI GenBank. General phylogenetic analysis by means of representative sampling has revealed two trends of genome evolutionary development from one and the same ancestral form. The first branch of evolution is represented by a relatively small number of *Zaire ebolavirus* isolates, disseminated in Guinea. The second one, the most numerous trend, is generated by four variants of *Zaire ebolavirus* genome, which have been identified starting from May – early June, 2014 in Sierra-Leone, but within a short period of time they have won the whole territory of the country and the neighboring Guinea and Liberia.

Key words: EVD agent genome, Ebola fever, sequencing, phylogenetic analysis, Republic of Guinea, West Africa.

Текущая эпидемия лихорадки Эбола в Западной Африке, о начале которой было сообщено в марте 2014 г., является на данный момент крупнейшей из зафиксированных для данной инфекционной болезни. На 9 августа 2015 г. известно о 27965 случаях болезни, 11298 из них закончились летально [4]. В отдельные временные периоды смертность от болезни, вызванной вирусом Эбола (БВВЭ), составляла порядка 90 %. Чрезвычайная ситуация в области общественного здравоохранения, вызванная текущей эпидемией БВВЭ, осложняется отсутствием средств специфической профилактики и лечения.

Этиологическим агентом данной эпидемии является вирус вида *Zaire ebolavirus*, по видимому, однократно попавший в удаленном районе Лесной Гвинеи из неизвестного природного резервуара в че-

ловеческую популяцию [2, 5].

Вирус вида *Zaire ebolavirus* (ZE) принадлежит роду *Ebolavirus* (семейство *Filoviridae*), вызывающему лихорадку Эбола у высших приматов. К роду *Ebolavirus* принадлежат еще 4 вида: *Sudan ebolavirus*, *Bundibugyo ebolavirus* (для которых отмечены неоднократные эпидемические осложнения, локализованные на Африканском континенте), *Reston ebolavirus* (отмеченный только на Филиппинах) и *TaiForest ebolavirus*. Морфологические признаки *Ebolavirus* схожи с вирусом *Marburgvirus*, также принадлежащим семейству *Filoviridae* и вызывающим подобную болезнь [7].

В этой связи важен своевременный мониторинг эволюции вирусного генома, который позволяет выявлять мутации, происходящие в геноме вируса, и

устанавливать филогенетические связи с его изолятами, выделенными ранее. Нами проведен сравнительный анализ трех геномов ZE, кДНК которых для секвенирования получена из клинического материала от больных лихорадкой Эбола в Гвинейской Республике в 2014 г.

Материалы и методы

Исходный клинический материал для диагностического анализа в виде сыворотки крови от лиц с подозрением на БВВЭ на территории Гвинейской Республики направлялся в мобильную лабораторию СПЭБ Роспотребнадзора. Проводилась инактивация материала в соответствии с МУ 1.3.2569-09 «Организация работы лабораторий, использующих методы амплификации нуклеиновых кислот при работе с материалом, содержащим микроорганизмы I–IV групп патогенности» [1]. Диагностический тест на наличие вирусной РНК ZE определяли с помощью технологии ПЦР-РВ и набора «АмплиСенс EBOV Zaire-FL» (ЦНИИ Эпидемиологии). Из образцов материала с положительным ответом проводили выделение вирусной РНК с помощью набора «QIAamp Viral RNA Mini Kit» (Qiagen) и последующий синтез кДНК с помощью набора «Реверта-Л» (ЦНИИ Эпидемиологии).

В исследовании нами использовались три образца кДНК:

- 211 (13.10.2014, Конакри), житель района Ратома, больной «И.», мужчина, 34 года, торговый агент. Поступил с лихорадкой, астенией, головной болью, миалгией и артралгией. Болезнь закончилась выздоровлением;

- 223 (14.10.2014, Конакри), житель района Колома, больной «Д.», женского пола, 14 лет, учащаяся. Поступила с лихорадкой, астенией и артралгией. Болезнь закончилась выздоровлением;

- 802 (18.12.2014, Киндиа), больной «К.», мужчина, 45 лет, водитель. Поступил в общемедицинскую сеть в тяжелом состоянии с лихорадкой и подозрением на малярию. Болезнь закончилась летальным исходом.

Получение необходимых фрагментов ДНК для секвенирования проводили в два этапа, используя гнездовую ПЦР. С помощью рассчитанных нами двух комплектов перекрывающихся пар праймеров сначала получали более протяженные фрагменты ДНК (1000–2900 п.н.), которые в дальнейшем использовали для амплификации более мелких частей ДНК (500–750 п.н.). Учет результатов ПЦР проводили методом гель-электрофореза в 1,5 % агарозном геле. Секвенирование полученных специфичных фрагментов ДНК проводили с использованием генетического анализатора «ABI 3500xl» (LifeTechnologies), согласно стандартному протоколу производителя. Часть фрагментов генома визуально не детектируемых при учете результатов ПЦР в агарозном геле секвенировали после целевого обогащения исходной

кДНК, используя систему высокопроизводительного секвенирования PGM (IonTorrent), согласно стандартному протоколу производителя. Фрагменты генома после секвенирования собирали в единую последовательность с помощью выравнивая на референс – изолят Makona-Kissidougou-C15 (KJ660346, 17.03.2014, Гвинея) с использованием программ SeqMan Pro 11 (DNASTAR Lasergene, <http://www.dnastar.com/t-seqmanpro.aspx>) и MEGA 6.0. (<http://www.megasoftware.net>).

Для проведения сравнительного анализа полученных данных нами использована репрезентативная выборка из международной базы данных NCBI GenBank, содержащая 645 геномов ZE, полученных из клинического материала от больных лихорадкой Эбола на территории стран Западной Африки (Гвинейская Республика, Сьерра-Леоне, Либерия и Мали), взятого с марта 2014 по март 2015 года.

Поиск единичных нуклеотидных полиморфизмов (SNP) для всей группы из 648 геномов ZE проводили с помощью программы Wombac 2.0.

Для построения филогенетических дендрограмм и дальнейшего анализа использовали программный пакет Bionumerics 7.5.

Результаты и обсуждение

Нами было собрано 18867 (99,5 %) нуклеотидов первичной структуры полного генома вируса ZE для образца 223; 18747 (98,9 %) – для образца 802 и 16679 (88 %) – для образца 211.

Полный геном РНК вируса ZE составляет от 18956 до 18961 нуклеотидов [9].

Полученные нуклеотидные последовательности геномов этих образцов депонированы в международную базу данных NCBI GenBank под следующими номерами доступа: 802 – KT633509, 223 – KT633510 и 211 – KT633511.

С помощью Интернет-ресурса BLAST найдены ближайшие к ним по подобию геномы вируса ZE из базы данных NCBI.

Образец 223 (14.10.14, Конакри, Республика Гвинея – здесь и далее по тексту указывается дата и место забора клинического материала) максимально близок с изолятами Conakry-1481 (KR534575, 15.10.14, Гвинейская Республика) и 076192 (KR817142, 18.10.14, Гвинейская Республика), отличаясь от них только 1 SNP. Так же они близки как по времени, так и по месту получения материала от больных.

Ближайшими по подобию к образцу 211 (13.10.14, Конакри, Гвинейская Республика), является изолят Conakry-1651 (KR534570, 24.10.14, Конакри, Гвинейская Республика), отличие составляет 1 SNP, что также соотносится с информацией о месте и времени выделения материала от больных.

Ближайшими по подобию к 802 (18.12.14, Киндиа, Гвинейская Республика) являются 3 изолята ZE, выделенные в середине октября 2014 г. в Гвинею

(Коя) под номерами доступа в NCBI GenBank: KR534591, KR534590 и KR534571. Образец 802 отличается от них на 5–6 SNP. Забор материала от больного для образца 802 сделан на 2 месяца позже, чем от ближайших вариантов из базы данных NCBI, а территориально районы Коя и Киндиа находятся по соседству.

Следует отметить, что секвенированные нами три генома *ZE* и другие варианты первичной структуры геномов *ZE* из стран Западной Африки отличаются более чем на 600 SNP от вариантов изолированных летом 2014 г. при вспышке БВВЭ в Центральной Африке (Демократическая Республика Конго). Такое значительное отличие (3 % и более) говорит о не зависимости друг от друга возникновения данных эпидемических ситуаций.

С помощью выравнивания относительно референса (Makona-Kissidougou-C15, KJ660346) нуклеотидных последовательностей трех полученных нами геномов *ZE*, установлены SNP, отличающие от референса исследуемые геномы (таблица).

Как видно из таблицы, образцы 802, 211 и 223 имеют 14, 11 и 12 SNP соответственно относительно референса. При этом 211 и 223 часто имеют одинаковые SNP, а 802 отличен от них в большинстве позиций. Сравнение геномов попарно показало, что 223 и 211 отличаются по 8 SNP, 223 и 802 – по 27 SNP, 211 и 802 – по 23 SNP (принимая во внимания, что анализ был сделан только по 88 % полученного генома 211 и 98,9 % для 802). Помимо этого в геноме образца 802 в некодирующей части гена VP24 обнаружена вставка единичного нуклеотида А (после позиции 10342), где расположен гомополимерный фрагмент из 6 адениловых остатков подряд.

В генетическом банке данных NCBI представлены всего 4 изолята вируса *ZE*, из более чем 830, имеющих аналогичную мутацию: 13625 Kikwit (KC242796, выделен в 1995 г. от больного в г. Киквит, Демократическая Республика Конго); Mayinga (AF499101, адаптированный на мышах вариант из Центральной Африки, депонирован в 2002 г.); Makona-20140910 (KR653269, выделен 18.09.14 от больного из Сьерра-Леоне, провинция Кенема); Makona-20142127 (KR653234, выделен 12.10.14 от больного из Сьерра-Леоне, провинция Бомбали).

Ближайшие по времени и территории выделения два изолята из Сьерра-Леоне, имеющие вставку +А после позиции 10342 существенно отличаются от 802 по другим позициям в геноме (21 и 24 SNP). Согласно литературным данным, скорость накопления мутаций в геноме *ZE* составляет $1-2 \cdot 10^{-3}$ на один нуклеотид в год [5, 8], соответственно для всего генома *ZE* нужно не менее 7 месяцев, чтобы приобрести 21 SNP. Вероятно, 802, Makona-20140910 и Makona-20142127 имели общего предка, получившего эту мутацию в начале 2014 г., на территории Гвинейской Республики, затем попавшего в Сьерра-Леоне.

Используя данные работы Н.Еbihara *et al.* [3] по адаптации генома вируса *ZE* на мышах, приво-

дящей к появлению уникальной вставки +А в некодирующей области гена VP24, которая может повлиять на репликацию и эффективность транскрипции, мы провели сравнение для 802, Makona-20140910 и Makona-20142127 по другим 12 SNP, которыми сопровождается адаптация на мышах генома вируса *ZE* и выявили несовпадение по 11 из 12 SNP. В данном случае вариант приобретения вставки +А в геноме *ZE* при многократной передаче вируса от человека к мышевидным грызунам и обратно маловероятен.

Мы провели филогенетический анализ полученных геномов с геномами *ZE* из международной базы данных NCBI GenBank, для этого сделана репрезентативная выборка 645 геномов *ZE*, нуклеотидная последовательность каждого из которых содержала не менее 99 % полного генома вируса. Эти геномы получены из клинического материала, выделенного с марта 2014 по март 2015 года от больных лихорадкой Эбола на территории стран Западной Африки (Гвинейская Республика, Сьерра-Леоне, Либерия и Мали).

Сначала мы провели выравнивание и поиск SNP в группе из 646 последовательностей геномов (645 геномов из GenBank и образец 223). В качестве референсной последовательности использовался геном изолята Makona-Kissidougou-C15. В результате для этой группы на протяжении всего генома получен профиль из 931 SNP по каждому из геномов. Большинство из этих 931 SNP, а именно 587 точечных мутаций, являются уникальными (каждая из них встречалась однократно в сравниваемой группе из 646 геномов).

Затем в группу из 646 геномов добавили нуклеотидные последовательности образцов 211 и 802, провели выравнивание и поиск SNP. Для данной группы из 648 геномов поиск SNP был ограничен длиной последовательности наименьшего в ней генома (88 % от полного генома у образца 211), поэтому полученный профиль составил 825 SNP.

Полученные профили из 931 и 825 SNP мы использовали для построения не укорененных филогенетических дендрограмм (рисунок, А, В).

Филогенетическое дерево А (рисунок) делится на 5 кластеров. Кластер 1 состоит из 77 геномов *ZE* (март–октябрь 2014 г., Гвинейская Республика, 76 геномов), из Либерии – 1 геном, 01.04.14. Геном референс Makona-Kissidougou-C15 (17.03.14, Гвинейская Республика) расположен у основания данного кластера (отмечен фиолетовой точкой), то есть является его родоначальником.

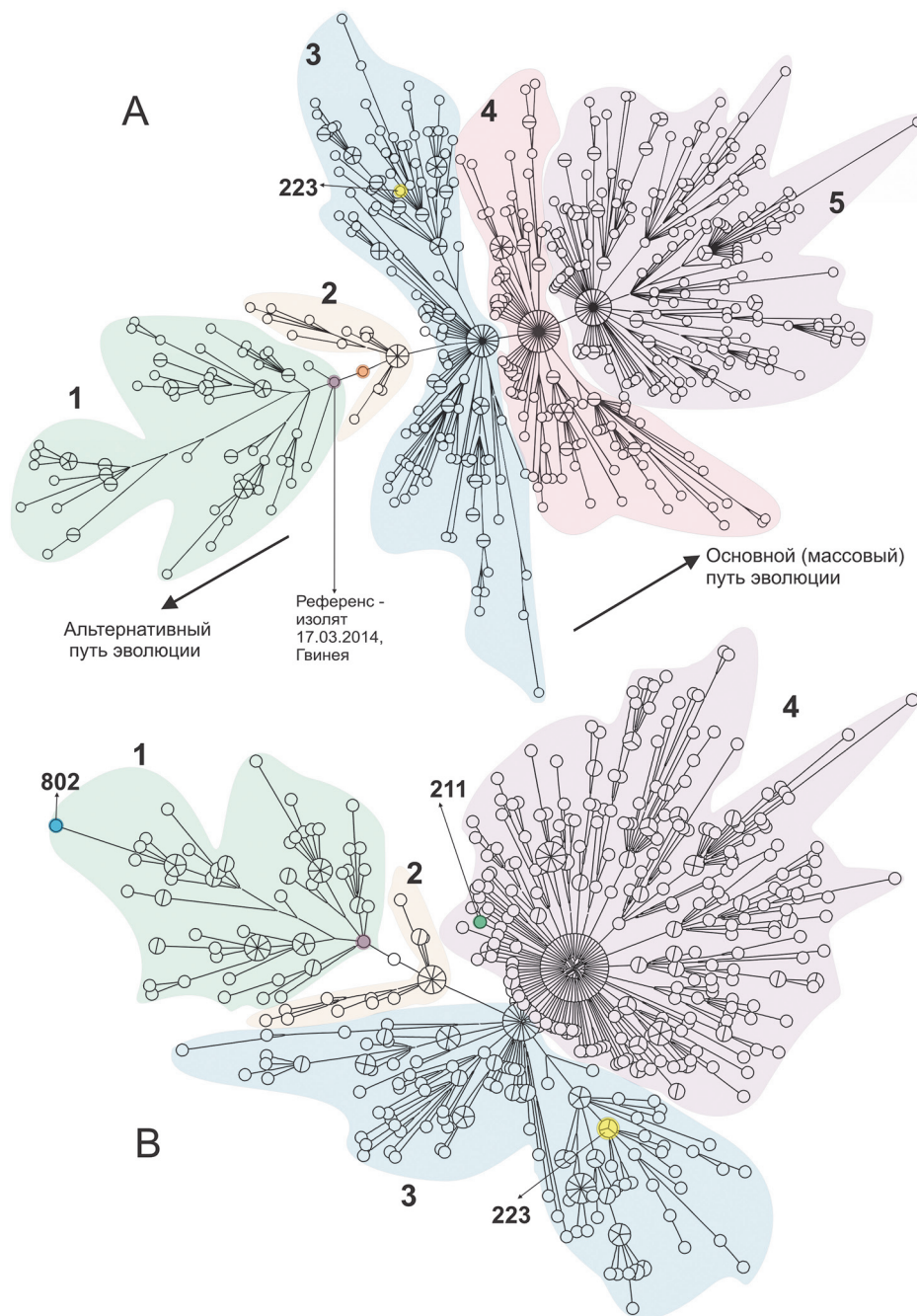
Кластер 2 образуют в общей сложности 22 генома *ZE*, на дендрограмме он отходит в правую сторону от референса и имеет выраженный центр из 8 одинаковых геномов (7 из Сьерра-Леоне, май–июнь 2014 г.) и 1 изолят из Гвинеи (06.06.14). Из центра расходятся ответвления – варианты дальнейшей эволюции генома с более поздними датами выделения от исходного единого варианта в центре, распространившиеся по Сьерра-Леоне, Гвинее и Либерии.

Кластер 3 представлен 176 геномами *ZE*, среди

Единичные нуклеотидные замены (выделены жирным шрифтом), отличающие анализируемые геномы образцов 211, 223 и 802 от референса (KJ660346, 17.03.14, Гвинейская Республика)

Позиция по референсу	Референс	Образец 802	Образец 211	Образец 223	Свойство замены
800 ген NP	C	C	T	T	Несиноним. Аргинин – Цистеин
1602 ген NP	A	A	G	A	Несиноним. Аспарагин – Серин
1777 ген NP	C	T	C	C	Синоним.
1849 ген NP	T	T	C	C	Синоним.
2373 ген NP	G	G	A	G	Несиноним. Аргинин – Лизин
4116 ген NP	C	T	C	C	Несиноним. Лейцин – Фенилаланин
5312 VP40	C	T	C	C	Синоним.
6056 ген GP	A	A	C	C	Синоним.
6283 ген GP	C	C	T	T	Несиноним. Аланин – Валин
6357 ген GP	A	G	A	A	Несиноним. Аспарагин – Аспарагиновая кислота
7476 ген GP	G	A	G	G	Несиноним. Глицин – Аспарагиновая кислота
7777 ген GP	A	G	A	A	Синоним.
8263 НОГ, ген GP	C	T	C	C	-
8928 ген VP30	A	A	C	C	Синоним.
9264 ген VP30	A	A	A	G	Синоним.
10087 НОГ, ген VP24	A	T	A	A	-
10218 НОГ, ген VP24	G	G	A	G	-
10315 НОГ, ген VP24	C	C	C	T	-
После 10342 НОГ, ген VP24	-	вставка A	-	-	-
13794 ген L	T	C	н.д.	T	Синоним.
13856 ген L	G	A	н.д.	G	Несиноним. Глицин – Аспарагиновая кислота
14253 ген L	C	C	C	T	Синоним.
14613 ген L	C	T	C	C	Синоним.
14724 ген L	A	A	н.д.	G	Синоним.
14943 ген L	T	C	н.д.	T	Синоним.
15660 ген L	C	T	C	C	Синоним.
15963 ген L	G	G	A	A	Синоним.
16446 ген L	C	T	н.д.	C	Синоним.
16514 ген L	G	G	н.д.	A	Несиноним. Серин – Аспарагин
17142 ген L	T	T	C	C	Синоним.
18813 5'-трейлерная нетранслируемая область	T	T	C	T	-

Примечания: А – аденин, G – гуанин, T – тимин, C – цитозин, НОГ – некодирующая область гена.



Не укорененное филогенетическое дерево, построенное по алгоритму Maximum parsimony на основании профиля из 931 SNP (A) и 825 SNP (B) от группы из 646 и 648 геномов *ZE* соответственно

которых 110 геномов из Гвинеи (август 2014 – январь 2015 г.), 37 – из Сьерра-Леоне (май–сентябрь 2014 г.), 25 – из Либерии (июнь 2014 – январь 2015 г.) и 4 – из Мали (октябрь–ноябрь 2014 г.). KP260799 (23.10.14) из Мали полностью совпадает с KR817135 (10.10.14) из Гвинеи, другие 3 изолята из Мали (ноябрь 2014 г.) имеют ближайшие по подобию 4 изолята из Гвинеи (август 2014 г.), что подтверждает Гвинейское происхождение вспышки БВВЭ осенью 2014 г. в Мали. Центр кластера 3 состоит из 17 одинаковых геномов *ZE* из Сьерра-Леоне (май–сентябрь 2014 г.), от которых происходят все остальные варианты эволюции геномов *ZE* данного кластера.

На верхней ветви кластера 3, вместе с другими изолятами из Гвинеи (64 генома, август 2014 – январь 2015), расположен образец 223 (отмечен желтой

точкой). Ближайшие по расположению к нему оказались геномы Conakry-1481 (KR534575, 15.10.14, Гвинейская Республика) и 076192 (KR817142, 18.10.14, Гвинейская Республика), что полностью соответствует полученным нами данным при сравнении образца 223 с базой данных NCBI с помощью BLAST. На этой же ветви находятся 4 генома *ZE* из Мали и 4 из Либерии (август 2014 г.).

Кластер 4 состоит из 148 геномов *ZE*, причем 144 из Сьерра-Леоне (май–ноябрь 2014 г.) и 4 генома из Гвинеи (август и октябрь 2014 г.). Центр кластера состоит из 30 одинаковых геномов *ZE* из Сьерра-Леоне (май–июль 2014 г.).

Кластер 5 состоит из 221 генома *ZE*, из них 191 получен из материала, взятого с июня по ноябрь 2014 г. от больных лихорадкой Эбола на территории

Сьерра-Леоне и 30 геномов *ZE* из Гвинеи (материал взят в июле–октябре 2014 г.). Центр кластера состоит из 22 одинаковых геномов, среди которых 13 геномов *ZE* из Сьерра-Леоне (июнь–сентябрь 2014 г.) и 9 геномов из Гвинеи (июль–сентябрь 2014 г.).

Общей для ответвлений, расходящихся от центров (оснований) всех 5 кластеров, является связь с определенным периодом взятия материала от больных, часто имеется территориальная обобщенность с тенденцией расположения геномов, полученных от материала с более поздним временем сбора к концу ветвей.

Мы провели сравнительный анализ между группами геномов, составляющих центры кластеров 2–5, которые являются исходными вариантами для отходящих от каждого из центров десятков и сотен их эволюционных потомков. При сравнении с референсом Makona-Kissidougou-C15 геномов *ZE* из центра кластеров 2–5 были найдены 7 ключевых SNP, образующих эти центры кластеров. Все геномы, составляющие центр кластера 2, отличаются от генома референса только появлением 2 SNP – первая мутация в позиции 1849 (синонимичная замена Т на С в гене NP), вторая – 6283 (несинонимичная замена С на Т, приводящая к замене аминокислоты Аланина на Валин в позиции 82 от начала кодирующей последовательности в гене GP). Гликопротеин – белок гена GP – имеет поверхностное расположение, осуществляет связывание вирусной частицы с мембранными рецепторами клетки-мишени и опосредует проникновение вируса внутрь клетки [11]. Геномы *ZE* в центре кластера 3 имеют помимо 2 SNP как у геномов из центра кластера 2, еще 3 SNP: в позиции 800 (несинонимичная замена С на Т, приводящая к замене аминокислоты Аргинин на Цистеин в позиции 111 от начала кодирующей последовательности в гене NP). Белок нуклеопротеин гена NP важен для образования стабильного рибонуклеопротеинового комплекса в процессе сборки вирусной частицы *ZE* [6]. Следующие две синонимичные замены в позиции 15963 (G на A в гене L) и в позиции 17142 (T на C в гене L). Геномы *ZE* в центре кластера 4 отличаются от геномов *ZE* в центре кластера 3 только появлением у них еще 1 SNP – в позиции 10218 (замена G на A в некодирующей части гена VP24).

Геномы в центре кластера 5 имеют на 1 SNP больше в позиции 14019 (синонимичная замена Т на С в гене L) относительно геномов в центре кластера 4.

Полученные нами результаты согласуются с работой D.J.Park *et al.* [8], в которой исследователи анализировали эволюционные изменения среди 232 геномов *ZE*, полученных из материала, выделяемого в течение семи месяцев от больных лихорадкой Эбола на территории Сьерра-Леоне, и выявили подобные 4 кластера (соответствуют кластерам 2–5, рисунок, А). В работе E.Simon-Lorriere *et al.* [10] провели сравнительный анализ нуклеотидных последовательностей 85 геномов *ZE*, полученных от пациентов, инфицированных на территории Гвинеи с июля по ноябрь 2014 г., со 110 геномами *ZE*, представленными в

базе данных NCBI (пробы взяты от больных на территории Гвинеи, Мали и Сьерра-Леоне). Авторами получены 3 большие генетические линии *ZE*, первая из которых соответствует кластеру 1 (рисунок, А) – вариантам распространенным только в Гвинеи с марта по октябрь 2014 года. Две другие генетические линии *ZE* изначально представлены изолятами из Сьерра-Леоне (май–июнь 2014 г.) и в дальнейшем распространились на территорию Гвинеи и Мали. При детальном рассмотрении каждая из этих двух больших генетических линий *ZE* может быть разбита на 2 подкластера, соответствующих полученным нами кластерам 2–5 (рисунок, А).

Проведенный нами анализ доступных данных в NCBI GenBank показывает, что из вариантов, выделенных еще в конце марта 2014 г. на территории Гвинеи, в базе данных NCBI представлен изолят KR817182 (31.03.14, Гвинейская Республика), который содержит одну (в позиции 6283) из описанных 7 SNP и, возможно, является исходным для линии изолятов *ZE* из Сьерра-Леоне, представленных в банке данных NCBI с 25 мая 2014 г. На рисунке (А) данный изолят отмечен оранжевой точкой и располагается на прямой между референсом и центром кластера 2. От генома референса (KJ660346, 17.03.14, Гвинейская Республика) изолят KR817182 отличается только одной заменой в позиции 6283.

Первый по дате материал, выделенный от больных лихорадкой Эбола на территории Сьерра-Леоне, из которого получены геномы KM034549 и KM034550 (NCBI GenBank), находящиеся в центре кластера 2, относится к 25 мая 2014 г. Уже 26 мая 2014 г. на территории Сьерра-Леоне от больных выделен материал, из которого получены геномы KM034552 и KM034556 из центра кластера 3, отличающиеся 4 SNP от генома KR817182. В центре кластера 4 уже 31 мая 2014 г. на территории Сьерра-Леоне от больных выделен материал, из которого секвенирован геном KM233049, отличающийся 5 SNP от генома KR817182. В центре кластера 5 расположен геном KM233048, который был секвенирован из материала, выделенного 9 июня 2014 г. от больных на территории Сьерра-Леоне, данный геном отличается 6 SNP от генома KR817182.

Таким образом, если исходить от даты выделения предполагаемого предшественника (KR817182, 31.03.14, Республика Гвинея), то все 4 геноварианта с 1, 4, 5 и 6 SNP из Сьерра-Леоне, которые отличают их от генома KR817182, могли приобрести данное количество единичных мутаций в течение 2 месяцев. В случае появления 5 и 6 SNP за этот период, скорость приобретения единичных мутаций высока и близка к $2 \cdot 10^{-3}$ на один нуклеотид в год. В то же время, исходя из представленных в базе NCBI данных, рассматриваемые 4 геноварианта не имеют тенденции к последовательному накоплению описанных 7 ключевых SNP, образующих центры кластеров с переходом к своей конечной форме – геномам из центра кластера 5. Они эволюционируют независимо, приобретая

мутации в других точках генома, на фоне сохраняющихся в течение второй половины 2014 и начала 2015 года исходных 1, 4, 5 и 6 SNP.

В базе данных NCBI GenBank не представлено ни одного из описываемых 4 геновариантов *ZE* до 25 мая 2014 г., или промежуточных вариантов, отличающихся от KR817182 (31.03.14, Республика Гвинея), например, на 2 или 3 SNP. Появление данных 4 геновариантов совпадает со стремительным обострением эпидемиологической ситуации в Сьерра-Леоне, после чего происходит их распространение в Гвинею и Либерию. Причина столь интенсивного дальнейшего распространения лихорадки Эбола на основе 4 геновариантов *ZE* заключается в их генетических особенностях и/или в результате сложившейся эпидемической ситуации.

Нами найдены всего 2 изолята, которые представляют дополнительные варианты генома для описанных выше 4 геновариантов *ZE*, но не получивших существенного распространения, – изолят KR105215 (06.07.2014, Сьерра-Леоне, Кенема), который имеет 2 SNP из описанных 7 относительно референса, но в другой комбинации точек – 10218 и 17142, вместо 1849 и 6283; изолят KP342330 (октябрь 2014 г., Республика Гвинея) имеет 5 SNP относительно референса, но тоже в другой комбинации – позиции 800, 6283, 10218, 15963 и 17142, вместо 800, 1849, 6283, 15963 и 17142.

В этой связи кластер 1 (рисунок, А) представляется как альтернативный вариант эволюции *ZE*. Ни один из входящих в его состав 77 геномов (все получены из материала, взятого от больных на территории Гвинеи, за исключением одного из Либерии), начиная от референса, выделенного в середине марта 2014 г. и заканчивая вариантами генома *ZE*, полученными из материала, взятого от пациентов в октябре 2014 г., не имеют ни одной из 7 ключевых SNP, образующих центры кластеров 2–5, сформированных первоначально изолятами *ZE* из Сьерра-Леоне. Отличие геномов, расположенных максимально далеко в этом кластере от исходного варианта (референса), достигает 12 SNP, но в других точках на геноме. Геномы *ZE*, относящиеся к кластеру 1, составляют не более 10 % от всех представленных в базе NCBI GenBank геномов этого вируса, а по дате изоляции последний из подобных вариантов относится к 25 октября 2014 г. (Гвинейская Республика).

Филогенетическое дерево, построенное на основании профиля из 825 SNP (рисунок, В), во многом подобно варианту А. Присутствуют кластеры с 1 по 4. Но последний кластер 5 объединен с кластером 4 из-за отсутствия данных по некоторым участкам генома у образца 211, в частности, по фрагменту, содержащему 7-ю кластерообразующую SNP (позиция 14019), исключенную из анализа по всему профилю у всех остальных 648 геномов группы сравнения. Так же из-за уменьшения количества точек различия (SNP) часто происходит объединение ранее отдельных ближайших геномов в общие круги.

Геном образца 211 расположился в кластере 4 недалеко от центрального круга (выделен зеленым цветом), геном образца 802 (18.12.2014) оказался продолжением эволюции вируса *ZE* в альтернативном направлении и располагается в самой отдаленной точке от референса в кластере 1 (рисунок, В). Номера ближайших по расположению в кластерах геномов (соответственно и подобию) из NCBI GenBank для образцов 802 и 211 полностью совпали с найденными для них ранее с помощью Интернет-ресурса BLAST NCBI.

Таким образом, полученные нуклеотидные последовательности геномов для исследуемых образцов 211, 223 и 802 позволили установить их филогенетическую близость к Западно-Африканским вариантам *ZE*, представленным в базе данных NCBI GenBank. Отмечена высокая степень совпадения по времени и месту выделения сходных с ними изолятов *ZE*.

Первоначально выделяемые в конце мая – начале июня 2014 г. в Сьерра-Леоне 4 варианта генома *ZE*, в силу своих генетических особенностей и/или в результате сложившейся эпидемической ситуации, послужили мощным толчком для развития эпидемии лихорадки Эбола в Западно-Африканском регионе.

Общий филогенетический анализ с использованием репрезентативной выборки генов *ZE* из базы данных NCBI GenBank показал два направления эволюционного развития генома *ZE* от одного исходного предка. Первая ветвь эволюции (кластер 1, рисунок, А) представлена относительно небольшим числом изолятов (77), получивших распространение только в Гвинею, и характеризуется постепенным характером приобретения SNP. Второе, наиболее многочисленное направление эволюции (568 геномов, рисунок, А, кластеры 2–5), получило свое развитие одновременно со стремительным появлением в конце мая – начале июня 2014 г. на территории Сьерра-Леоне 4 вариантов одного исходного генома *ZE* (возможно, изолята KR817182, 31.03.14, Гвинейская Республика), которые за короткий период времени охватили не только территорию Сьерра-Леоне, но и соседние с ней страны – Гвинею и Либерию.

Авторы подтверждают отсутствие конфликта финансовых/нефинансовых интересов, связанных с написанием статьи.

СПИСОК ЛИТЕРАТУРЫ

1. Организация работы лабораторий, использующих методы амплификации нуклеиновых кислот при работе с материалом, содержащим микроорганизмы I–IV групп патогенности: МУ 1.3.2569-09. М.: Федеральный центр гигиены и эпидемиологии Роспотребнадзора; 2010. 51 с.
2. Baize S., Pannetier D., Oestereich L., Rieger T., Koivogui L., Magassouba N'Faly, Soropogui B., Sow M.S., Keita S., Clerck H., Tiffany A., Dominguez G., Loua M., Traoré A., Kolié M., Malano E.R., Heleze E., Bocquin A., Mély S., Raoul H., Caro V., Cadar D., Gabriel M., Pahlmann M., Tappe D., Schmidt-Chanasit J., Impouma B., Diallo A.K., Formenty P., Van Herp M., Günther S. Emergence of Zaire Ebola virus disease in Guinea. *N. Engl. J. Med.* 2014; 371:1418–25.
3. Ebihara H., Takada A., Kobasa D., Jones S., Neumann G., Theriault S., Bray M., Feldmann H., Kawaoka Y. Molecular determinants of Ebola virus virulence in mice. *PLoS Pathog.* 2006; 2(7):e73.
4. Ebola response roadmap – Situation report. <http://www>

who.int/csr/disease/ebola/situation-reports/en. (дата обращения 12.08.2015).

5. Gire S.K., Goba A., Andersen K.G., Sealfon R.S., Park D.J., Kanneh L., Jalloh S., Momoh M., Fullah M., Dudas G., Wohl S., Moses L.M., Yozwiak N.L., Winnicki S., Matranga C.B., Malboeuf C.M., Qu J., Gladden A.D., Schaffner S.F., Yang X., Jiang P.P., Nekoui M., Colubri A., Coomber M.R., Fonnine M., Moigboi A., Gbakie M., Kamara F.K., Tucker V., Konuwa E., Saffa S., Sellu J., Jalloh A.A., Kovoma A., Koninga J., Mustapha I., Kargbo K., Foday M., Yillah M., Kanneh F., Robert W., Massally J.L., Chapman S.B., Bochicchio J., Murphy C., Nusbaum C., Young S., Birren B.W., Grant D.S., Scheffelin J.S., Lander E.S., Hapji C., Gevao S.M., Gnirke A., Rambaut A., Garry R.F., Khan S.H., Sabeti P.C. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science*. 2014; 345(6202):1369–72.

6. Huang Y., Xu L., Sun Y., Nabel G.J. The assembly of Ebola virus nucleocapsid requires virion-associated proteins 35 and 24 and posttranslational modification of nucleoprotein. *Mol. Cell*. 2002; 10:307–16.

7. Kuhn J.H., Andersen K.G., Baize S., Bao Y., Bavari S., Berthet N., Blinkova O., Brister J.R., Clawson A.N., Fair J., Gabriel M., Garry R.F., Gire S.K., Goba A., Gonzalez J.P., Günther S., Hapji C.T., Jahrling P.B., Kapetski J., Kobinger G., Kugelman J.R., Leroy E.M., Maganga G.D., Mbala P.K., Moses L.M., Muyembe-Tamfum J.J., N'Faly M., Nichol S.T., Omilabu S.A., Palacios G., Park D.J., Paweska J.T., Radoshitzky S.R., Rossi C.A., Sabeti P.C., Schieffelin J.S., Schoepp R.J., Sealfon R., Swanepoel R., Towner J.S., Wada J., Wauquier N., Yozwiak N.L., Formenty P. Nomenclature- and database-compatible names for the two Ebola virus variants that emerged in Guinea and the Democratic Republic of the Congo in 2014. *Viruses*. 2014; 6(11):4760–99.

8. Park D.J., Dudas G., Wohl S., Goba A., Whitmer S.L., Andersen K.G., Sealfon R.S., Ladner J.T., Kugelman J.R., Matranga C.B., Winnicki S.M., Qu J., Gire S.K., Gladden-Young A., Jalloh S., Nosamiefan D., Yozwiak N.L., Moses L.M., Jiang P.P., Lin A.E., Schaffner S.F., Bird B., Towner J., Mamoh M., Gbakie M., Kanneh L., Kargbo D., Massally J.L., Kamara F.K., Konuwa E., Sellu J., Jalloh A.A., Mustapha I., Foday M., Yillah M., Erickson B.R., Sealy T., Blau D., Paddock C., Brault A., Amman B., Basile J., Bearden S., Belsler J., Bergeron E., Campbell S., Chakrabarti A., Dodd K., Flint M., Gibbons A., Goodman C., Klena J., McMullan L., Morgan L., Russell B., Salzer J., Sanchez A., Wang D., Jungreis I., Tomkins-Tinch C., Kislyuk A., Lin M.F., Chapman S., Maclinnis B., Matthews A., Bochicchio J., Hensley L.E., Kuhn J.H., Nusbaum C., Schieffelin J.S., Birren B.W., Forget M., Nichol S.T., Palacios G.F., Ndiaye D., Hapji C., Gevao S.M., Vandi M.A., Kargbo B., Holmes E.C., Bedford T., Gnirke A., Stroher U., Rambaut A., Garry R.F., Sabeti P.C. Ebola virus Epidemiology, Transmission, and Evolution during Seven Months in Sierra Leone. *Cell*. 2015; 161(7):1516–26.

9. Sanchez A., Kiley M.P., Holloway B.P., Auperin D.D. Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Res*. 1993; 29(3):215–40.

10. Simon-Loriere E., Faye O., Faye O., Koivogui L., Magassouba N., Keita S., Thiberge J.M., Diancourt L., Bouchier C., Vandenbogaert M., Caro V., Fall G., Buchmann J.P., Matranga C.B., Sabeti P.C., Manuguerra J.C., Holmes E.C., Sall A.A. Distinct lineages of Ebola virus in Guinea during the 2014 West African epidemic. *Nature*. 2015; 524(7563):102–4.

11. Volchkov V.E., Becker S., Volchkova V.A., Ternovoj V.A., Kotov A.N., Netesov S.V., Klenk H.-D. GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and Vaccinia virus polymerases. *Virology*. 1995; 214:421–30.

References

- [Organization of works of laboratories utilizing nucleic acids amplification methods while handling materials containing microorganisms of I–IV pathogenicity groups]. MU.1.3.2569-09. M.; 2010. 51 p.
- Baize S., Pannetier D., Oestereich L., Rieger T., Koivogui L., Magassouba N'Faly, Soropogui B., Sow M.S., Keita S., Clerck H., Tiffany A., Dominguez G., Loua M., Traoré A., Kolié M., Malano E.R., Heleze E., Bocquin A., Mély S., Raoul H., Caro V., Cadar D., Gabriel M., Pahlmann M., Tappe D., Schmidt-Chanasit J., Impouma B., Diallo A.K., Formenty P., Van Herp M., Günther S. Emergence of Zaire Ebola virus disease in Guinea. *N. Engl. J. Med*. 2014; 371:1418–25.
- Ebihara H., Takada A., Kobasa D., Jones S., Neumann G., Theriault

S., Bray M., Feldmann H., Kawaokaet Y. Molecular determinants of Ebola virus virulence in mice. *PLoS Pathog*. 2006; 2(7):e73.

4. Ebola response roadmap – Situation report. [cited 12 Aug 2015]. Available from: <http://www.who.int/csr/disease/ebola/situation-reports/en>.

5. Gire S.K., Goba A., Andersen K.G., Sealfon R.S., Park D.J., Kanneh L., Jalloh S., Momoh M., Fullah M., Dudas G., Wohl S., Moses L.M., Yozwiak N.L., Winnicki S., Matranga C.B., Malboeuf C.M., Qu J., Gladden A.D., Schaffner S.F., Yang X., Jiang P.P., Nekoui M., Colubri A., Coomber M.R., Fonnine M., Moigboi A., Gbakie M., Kamara F.K., Tucker V., Konuwa E., Saffa S., Sellu J., Jalloh A.A., Kovoma A., Koninga J., Mustapha I., Kargbo K., Foday M., Yillah M., Kanneh F., Robert W., Massally J.L., Chapman S.B., Bochicchio J., Murphy C., Nusbaum C., Young S., Birren B.W., Grant D.S., Scheffelin J.S., Lander E.S., Hapji C., Gevao S.M., Gnirke A., Rambaut A., Garry R.F., Khan S.H., Sabeti P.C. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science*. 2014; 345(6202):1369–72.

6. Huang Y., Xu L., Sun Y., Nabel G.J. The assembly of Ebola virus nucleocapsid requires virion-associated proteins 35 and 24 and posttranslational modification of nucleoprotein. *Mol. Cell*. 2002; 10:307–16.

7. Kuhn J.H., Andersen K.G., Baize S., Bao Y., Bavari S., Berthet N., Blinkova O., Brister J.R., Clawson A.N., Fair J., Gabriel M., Garry R.F., Gire S.K., Goba A., Gonzalez J.P., Günther S., Hapji C.T., Jahrling P.B., Kapetski J., Kobinger G., Kugelman J.R., Leroy E.M., Maganga G.D., Mbala P.K., Palacios G., Park D.J., Paweska J.T., Radoshitzky S.R., Rossi C.A., Sabeti P.C., Schieffelin J.S., Schoepp R.J., Sealfon R., Swanepoel R., Towner J.S., Wada J., Wauquier N., Yozwiak N.L., Formenty P. Nomenclature- and database-compatible names for the two Ebola virus variants that emerged in Guinea and the Democratic Republic of the Congo in 2014. *Viruses*. 2014; 6(11):4760–99.

8. Park D.J., Dudas G., Wohl S., Goba A., Whitmer S.L., Andersen K.G., Sealfon R.S., Ladner J.T., Kugelman J.R., Matranga C.B., Winnicki S.M., Qu J., Gire S.K., Gladden-Young A., Jalloh S., Nosamiefan D., Yozwiak N.L., Moses L.M., Jiang P.P., Lin A.E., Schaffner S.F., Bird B., Towner J., Mamoh M., Gbakie M., Kanneh L., Kargbo D., Massally J.L., Kamara F.K., Konuwa E., Sellu J., Jalloh A.A., Mustapha I., Foday M., Yillah M., Erickson B.R., Sealy T., Blau D., Paddock C., Brault A., Amman B., Basile J., Bearden S., Belsler J., Bergeron E., Campbell S., Chakrabarti A., Dodd K., Flint M., Gibbons A., Goodman C., Klena J., McMullan L., Morgan L., Russell B., Salzer J., Sanchez A., Wang D., Jungreis I., Tomkins-Tinch C., Kislyuk A., Lin M.F., Chapman S., Maclinnis B., Matthews A., Bochicchio J., Hensley L.E., Kuhn J.H., Nusbaum C., Schieffelin J.S., Birren B.W., Forget M., Nichol S.T., Palacios G.F., Ndiaye D., Hapji C., Gevao S.M., Vandi M.A., Kargbo B., Holmes E.C., Bedford T., Gnirke A., Stroher U., Rambaut A., Garry R.F., Sabeti P.C. Ebola virus Epidemiology, Transmission, and Evolution during Seven Months in Sierra Leone. *Cell*. 2015; 161(7):1516–26.

9. Sanchez A., Kiley M.P., Holloway B.P., Auperin D.D. Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Res*. 1993; 29(3):215–40.

10. Simon-Loriere E., Faye O., Faye O., Koivogui L., Magassouba N., Keita S., Thiberge J.M., Diancourt L., Bouchier C., Vandenbogaert M., Caro V., Fall G., Buchmann J.P., Matranga C.B., Sabeti P.C., Manuguerra J.C., Holmes E.C., Sall A.A. Distinct lineages of Ebola virus in Guinea during the 2014 West African epidemic. *Nature*. 2015; 524(7563):102–4.

11. Volchkov V.E., Becker S., Volchkova V.A., Ternovoj V.A., Kotov A.N., Netesov S.V., Klenk H.-D. GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and Vaccinia virus polymerases. *Virology*. 1995; 214:421–30.

Authors:

Krasnov Ya.M., Safronov V.A., Nosov N.Yu., Kutyrer V.V. Russian Research Anti-Plague Institute “Microbe”. 46, Universitetskaya St., Saratov, 410005, Russian Federation. E-mail: rusrapi@microbe.ru

Popova A.Yu. Federal Service for Surveillance in the Sphere of Consumers Rights Protection and Human Welfare; 18, Bld. 5 and 7, Vadkovsky Pereulok, Moscow, 127994, Russian Federation. Russian Medical Academy for Post-Graduate Training; 2/1, Barrikadnaya St., Moscow, 125993, Russian Federation.

Об авторах:

Краснов Я.М., Сафронов В.А., Носов Н.Ю., Кутырев В.В. Российский научно-исследовательский противочумный институт «Микроб». Российская Федерация, 410005, Саратов, ул. Университетская, 46. E-mail: rusrapi@microbe.ru

Попова А.Ю. Федеральная служба по надзору в сфере защиты прав потребителей и благополучия человека; Российская Федерация, 127994, Москва, Вадковский переулок, дом 18, строение 5 и 7. Российская медицинская академия последилового образования; Российская Федерация, 125993, Москва, ул. Баррикадная, 2/1.

Поступила 19.08.15.