Ответственный исполнитель: Климова Анна

Дата последнего редактирования: 27.11.2023 11.00

2 Материалы и методы  
2.1 Эпидемиологические методы исследования

Для определения генотипического пейзажа ВИЧ-1 в Челябинской области в период с января 2018 года по март 2022 было обследовано 38 ВИЧ-инфицированных пациентов. Возраст в среднем составлял 39 лет (среднее квадратичной отклонение σ = 7, коэффициент Шапиро-Вилка W = 0,917 при p = 0,008). Среди обследованных 20 человек составили мужчины (52,63%, 95% доверительный интервал (ДИ): [37,3 - 67,5]), 18 – женщины (47,37%, 95% ДИ: [32,5 - 62,7]). Все пациенты находились на стадии 4В ВИЧ-инфекции.

## 2.2 Лабораторные методы исследования

Определение уровня вирусной нагрузки проводили с использованием набора реагентов «АмплиСенс® ВИЧ-Монитор-FRT» (производитель ФБУН ЦНИИЭ Роспотребнадзора) методом количественного определения РНК вируса иммунодефицита человека типа 1 (ВИЧ-1) в клиническом материале методом полимеразной цепной реакции (ПЦР) с гибридизационно-флуоресцентной детекцией на амплификаторе Rotor-Gene Q («QUIAGEN», Германия).

Уровень вирусной нагрузки ВИЧ в плазме крови составлял по медиане 4,67 (МКИ: 4,19-5,40) lg копий/мл и статистически значимо превышал аналогичный показатель в ликворе, составляющий по медиане 3,87 (МКИ: 2,73 - 4,66) lg копий/мл, на 0,8 lg или в 6,27 раза (U=442, z=2,904, p=0,004).

Мутации устойчивости ВИЧ-1 к АРВП выявляли методом секвенирования амплифицированных фрагментов гена pol, кодирующего протеазу и часть обратной транскриптазы ВИЧ-1, с использованием набора реагентов «АмплиСенс® HIV-Resist-Seq» (производитель ФБУН ЦНИИЭ Роспотребнадзора) в соответствии с инструкцией производителя. Электрофорез высокого разрешения очищенных фрагментов с флюоресцирующими терминаторами проводили с помощью генетического анализатора Applied Biosystems 3500 Genetic Analyzer (Life Technologies, США).

Всего было получено 57 нуклеотидных последовательностей гена pol ВИЧ-1. Все нуклеотидные последовательности депонированы в международную базу данных GenBank (OR260484-OR260511).

В случае вирусной нагрузки менее 1000 копий/мл и/или отсутствия на элекрофореграмме специфических полос амплифицированных фрагментов ДНК, проводили предварительное ультрацентрифугирование 1,0 мл плазмы в течение 1 ч при скорости центрифуги не менее 24000 g при температуре от 2 до 80°С.

Стандартная методика для выявления резистентности ВИЧ к АРВП на основе генотипирования включала следующие стадии:

1) экстракция (выделение) РНК из клинического материала ‒ плазмы крови,

2) обратная транскрипция,

3) амплификация участков генома ВИЧ-1, кодирующих обратную транскриптазу и протеазу,

4) подготовка и проведение секвенирования включает следующие стадии:

a) проведение и очистка продуктов амплификации от невключившихся нуклеотидов и праймеров;

б) детекция и оценка концентрации очищенных ПЦР- продуктов в агарозном геле.

в) проведение реакции циклического секвенирования;

г) очистка продуктов реакции секвенирования от не включившихся терминаторов;

д) подготовка продуктов реакции секвенирования для проведения автоматической детекции нуклеотидной последовательности;

е) автоматическая детекция нуклеотидной последовательности с помощью генетического анализатора Applied Biosystems 3500 Genetic Analyzer (Life Technologies, США);

5) Анализ и интерпретация результатов: обработку электрофореграмм и получение консенсусной последовательности осуществляли с помощью программного обеспечение «Деона 1.7.0» («МАГ»), для дальнейшего анализа использовали сервисы Стэндфордского университета «HIVdb Program Genotypic Resistance Interpretation Algorithm».

Измерение количества CD4-клеток проводилось с помощью проточного цитометра BD FACSCanto™ II и набора реагентов BD Tritest CD4/CD8/CD3 (производитель Becton Dickinson).

## 2.3 Статистические методы исследования

Расчёт доверительных интервалов осуществляли по методу Уилсона для уровня ошибки 1-го типа 0,05 [33]. Для подтверждения статистически значимого различия использовали критерии непараметрической статистики (Хи-квадрат, точный критерий Фишера, критерий Манна-Уитни).

Статистическую обработку данных проводили с помощью программного продукта «Statistica v.12» (StatSoft Russia).

## 2.4 Биоинформационные методы исследования

Формирование консенсусных последовательностей из данных, полученных в результате секвенирования исследуемых образцов с помощью генетического анализатора Applied Biosystems 3500 Genetic Analyzer (Life Technologies, США), использовалось программное обеспечение Unipro UGENE (Okonechnikov K, Golosova O, Fursov M, the UGENE team. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 2012 28: 1166-1167. doi:10.1093/bioinformatics/bts091) и специальное прикладное программное обеспечение собственной разработки (на языке программирования Python 3.7) для анализа результатов секвенирования.

Субтипирование и определение мутаций полиморфизма и МЛУ выполняли на сервисе Стэнфордского университета HIV drug resistance database (HIVdb Program: Mutations Analysis, версия программы 3.4.3; версия алгоритма 9.4) [29].

Для определения филогенетического положения исследуемых штаммов у каждой из 57 последовательностей с помощью сервиса NCBI BLAST [27] найдены геномы ВИЧ, имеющие сходство 95% и выше. После исключения повторяющихся записей, синтетических последовательностей, а также записей, в которых страну происхождения штамма невозможно установить, осталось 2872 нуклеотидных последовательностей, имеющих среднюю идентичность по медиане 95,97% (МКИ: 95,52% - 96,51%).

В целях автоматизации процесса обращения к сервису NCBI BLAST была написана процедура на языке программирования Python 3.7.5 с использованием модулей Bio.Blast.NCBIWWW и Bio.SeqIO пакета Biopython 1.75 (https://biopython.org).

Для проведения филогенетического анализа множественное выравнивание выполняли с помощью алгоритма ClustalOmega на онлайн-сервисе EMBL [28]. Подбор наиболее подходящей для имеющихся данных модели замены нуклеотидов осуществляли с использованием онлайн сервиса FindModel [30]. Построение филогенетических деревьев 2929 фрагментов гена pol ВИЧ-1, кодирующего протеазу и часть ревертазы, выполняли методом максимального правдоподобия с моделью замены нуклеотидов GTR+G (General Time Reversible + gamma) с использованием программного обеспечения MEGA X [31]. В качестве меры статистической поддержки использовалось 500 бутстреп репликаций.

Трехмерные модели вирусных белков (протеазы и обратной транскриптазы) были получены методом гомологического построения. Для протеазы была использована модель кристаллической структуры протеазы ВИЧ-1, субтипа А (PDB ID: 3ixo). Данная модель была выбрана, так как она имела наибольшее сходство с нашими аминокислотными последовательностями, среди моделей протеаз ВИЧ-1, не связанных с ингибиторами и не имеющих МЛУ, а также относилась к наиболее близкому субтипу А.

Для обратной транскриптазы были использованы модель субъединицы p51 обратной транскриптазы ВИЧ-1 (PDB ID: 3kjv) и модель обратной транскриптазы/РНКазы Н ВИЧ-1 (PDB ID: 4icl). Данные модели были выбраны, так как они имели наибольшее сходство по первичной структуре с нашими последовательностями среди моделей, не находящихся в комплексе с ингибиторами и не имеющих МЛУ.

Построение трехмерных моделей белков и сравнение их структуры проводилось с использованием онлайн сервиса SWISS-MODEL [34,35].

# 3 Процесс теоретических исследований

## 3.3 Генетические особенности штаммов ВИЧ, выделенных от АРТ-наивных пациентов Челябинской области

### 3.3.1 Мониторинг эпидемического процесса ВИЧ-инфекции в Челябинской области

По данным анамнеза пациентов установлено, что продолжительность инфицирования ВИЧ с момента установления диагноза до госпитализации составляла в среднем 81 месяц (σ = 64, W = 0,924, p = 0,01). Большинство пациентов – 30 из 38 (79,0%, 95 % ДИ [63,7 – 88,9]), не имели опыта антиретровирусной терапии, остальные находились на терапии по медиане до 29 месяцев (МКИ: от 9 до 35), но при этом с низкой приверженностью, самостоятельно прерывая назначенные курсы лечения. Учитывая, что все пациенты были госпитализированы на 4В стадии ВИЧ-инфекции, их заражение произошло задолго до установления диагноза. Таким образом, мы наблюдали многолетнее течение ВИЧ-инфекции без АРТ.

Превышение уровня вирусной нагрузки ВИЧ в ликворе по сравнению с плазмой достаточно распространённое явление среди пациентов, не получающих эффективной АРТ, и среди АРТ-наивных пациентов, что было также показано в перекрестном многоцентровом исследовании, проведённом в крупных городах Европы в период с 1982 по 2017 год [38]. При этом различие в уровне вирусной нагрузки составляло 1,0 lg копий/мл и наблюдалась корреляция между данными показателями в плазме крове и ликворе. В нашем исследовании имелась слабая прямая корреляционная связь между уровнями вирусной нагрузки ВИЧ в плазме крови и ликворе: коэффициент Спирмена Rs = 0,38, p = 0,019.

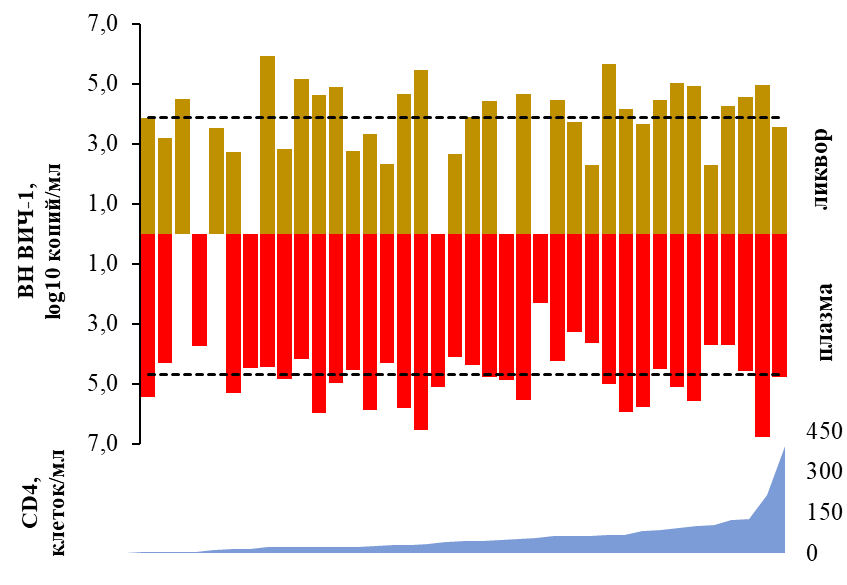


Рисунок 1. Уровень вирусной нагрузки ВИЧ в плазме крови и ликворе и количество CD4-клеток в выборке обследованных пациентов (n=19). Пунктиром обозначена медиана.

### 3.3.2 Филогенетический анализ штаммов ВИЧ, циркулировавших на территории Челябинской области

Вместе с исследуемыми образцами генотипическая структура была представлена 2508 штаммами субсубтипа А6 (85,5%, 95 % ДИ [84,1 – 86,7]), 397 CRF63\_02A6 (13,8%, 95 % ДИ [12,6 – 15,1]), 14 субтипов B (0,5%, 95 % ДИ [0,3 – 0,8]) и 7 CRF02\_A6G (0,2%, 95 % ДИ [0,1 – 0,5]) среди 2929 анализируемых штаммов.

В результате филогенетического анализа сформировалось 26 кластеров, в 11 из которых вошли 57 штаммов ВИЧ, выделенных от исследуемых пациентов (рисунок 2). Среди изолятов от пациентов из стран ближнего зарубежья в кластерах с исследуемыми образцами чаще всего оказывались изоляты, циркулирующие на Украине и в Киргизии (по 6 из 11 кластеров), в Белоруссии, Таджикистане, Казахстане, Армении (5 из 11 кластеров), среди государств из дальнего зарубежья – Польша и Германия (5 из 11 кластеров).

Среди 57 изолятов ВИЧ, выделенных от пациентов из исследуемой выборки, 11 (19,3%, 95 % ДИ [11,1 – 31,3]) сформировали отдельный кластер №3, где на их долю пришлось 73,3% (95 % ДИ [48 – 89,1]). Данные изоляты были выделены от 8 пациентов, из которых 6 женщин и 1 мужчина из города Челябинск и 1 женщина из г. Сатка Челябинской области. Из 8 пациентов только 2 заразились при употреблении наркотиков внутривенно, остальные – половым путём.

В самом крупном на филогенетическом дереве кластере №13, содержащем 873 генома ВИЧ (29,8%, 95 % ДИ [28,2 – 31,5]) оказалось 16 изолятов от пациентов из исследуемой группы, а также изоляты от пациентов из 19 иностранных государств. Наибольшие доли среди изолятов ВИЧ из иностранных государств в данном кластере пришлись на изоляты, циркулирующие в Киргизии (29,3%), Белоруссии (23,6%), Таджикистане (12,0%).

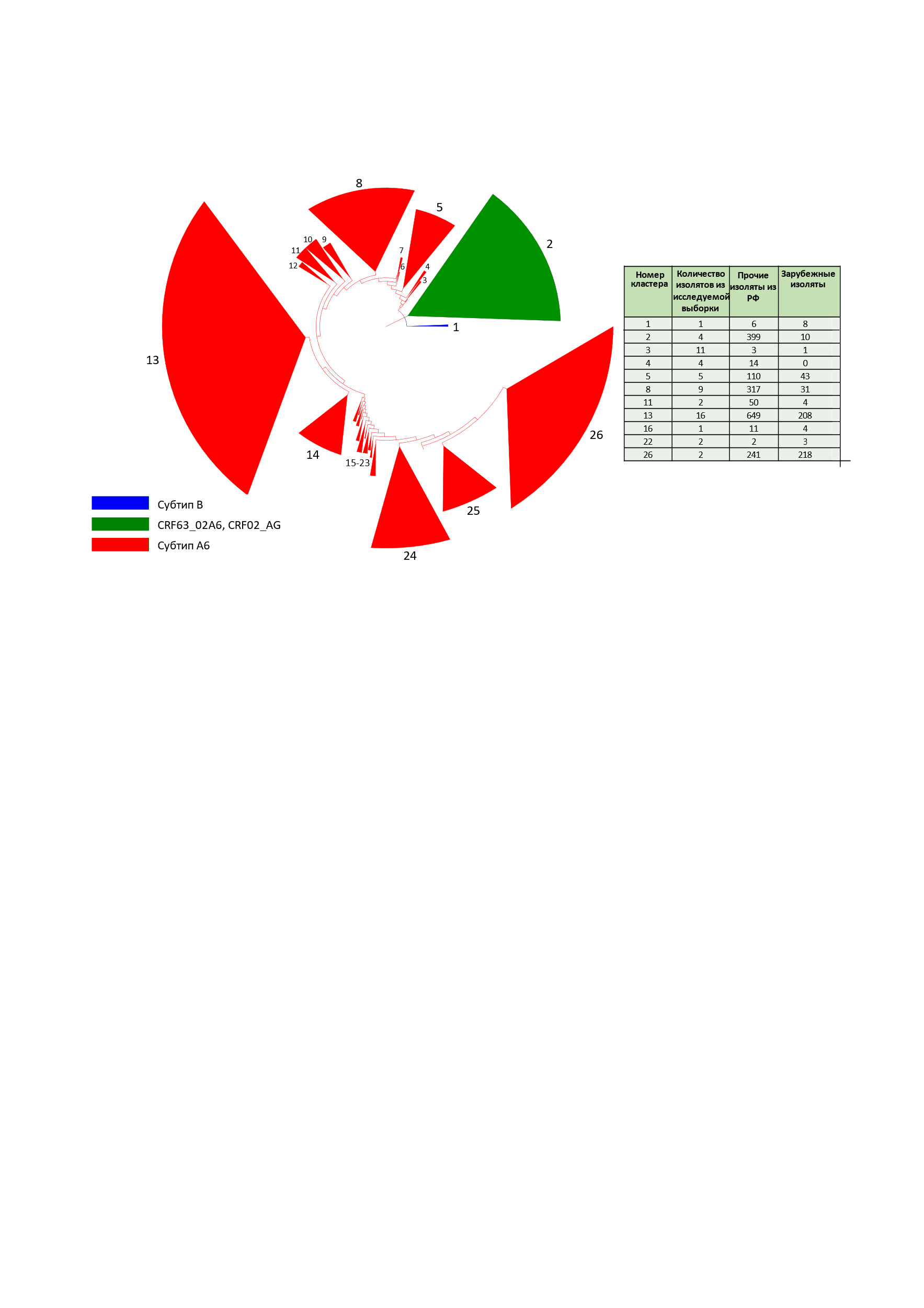


Рисунок 2. Филогенетическое дерево, построенное методом максимального правдоподобия на основе 2929 фрагментов гена pol ВИЧ-1 (кодирующего протеазу и часть обратной транскриптазы) с не менее чем 95% сходством со штаммами исследуемой выборки.

### 3.3.3 Анализ третичной структуры ферментов штаммов ВИЧ, циркулировавших на территории Челябинской области

Для анализа аминокислотных замен и третичной структуры протеазы и ревертазы ВИЧ были отобраны изоляты из парных проб от 13 пациентов из 19. Из сравнения исключены 5 пациентов с высокой гетерогенностью штаммов и подозрением на суперинфицирование и один пациент, у которого в изолятах из плазмы и из ликвора аминокислотный состав протеазы и ревертазы совпал полностью.

Были обнаружены высоковариабельные участки в структуре субъединицы p51 обратной транскриптазы ВИЧ-1 в аминокислотных позициях 16-20 и 210-235. У 5 пациентов структура фермента у изолятов из плазмы крови и из ликвора отличалась в обеих позициях, еще у 4 – только в позициях 210-235. У 3 пациентов третичная структура субъединицы p66 обратной транскриптазы ВИЧ-1 различалась в аминокислотных позициях 187-190.

Третичная структура субъединицы p51 обратной транскриптазы ВИЧ-1 оказалась наиболее вариабельной, при этом для возникновения конформационных различий между белками изолятов из плазмы крови и ликвора требовалось не менее двух эволюционных событий, связанных с заменой аминокислот, по одному в каждом из изолятов. Субъединица p51 каталитически неактивна и играет лишь структурную роль в комплексе обратной транскрипции. Аналогичные аминокислотные замены в субъединице p66 ревертазы, которая и выполняет каталитические функции, образуя участок связывания с ННИОТ, не приводили к изменениям в третичной структуре. В трех случаях различия в третичной структуре субъединицы p66 присутствовали, но были минимальны и охватывали участок, длиной в 3 аминокислоты. Различия третичной структуры субъединицы p66 ревертазы были связаны как минимум с 3 эволюционными событиями в одном из изолятов или с пятью событиями в изолятах из ликвора и плазмы.

Анализ аминокислотных замен в субъединицах p51 и p66 ревертазы продемонстрировал их связь с различиями в третичной структуре (Таблица 1) и позволяет предположить, что аминокислотные замены, приводящие к изменению третичной структуры одной субъединицы ОТ не обязательно ведут к изменению третичной структуры другой субъединицы. Так в 12 из 18 образцов различия в третичной структуре p51 не привели к изменениям в p66, в 2 образцах – изменения в p66 не коррелировали с изменением в p51.

Таблица 1. Аминокислотные замены, связанные с различием в третичной структуре обратной транскриптазы изолятов ВИЧ-1 из ликвора и плазмы крови

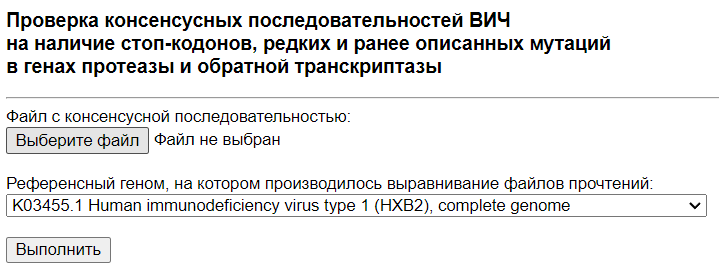
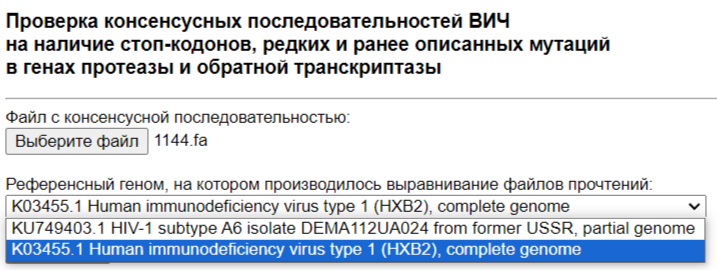
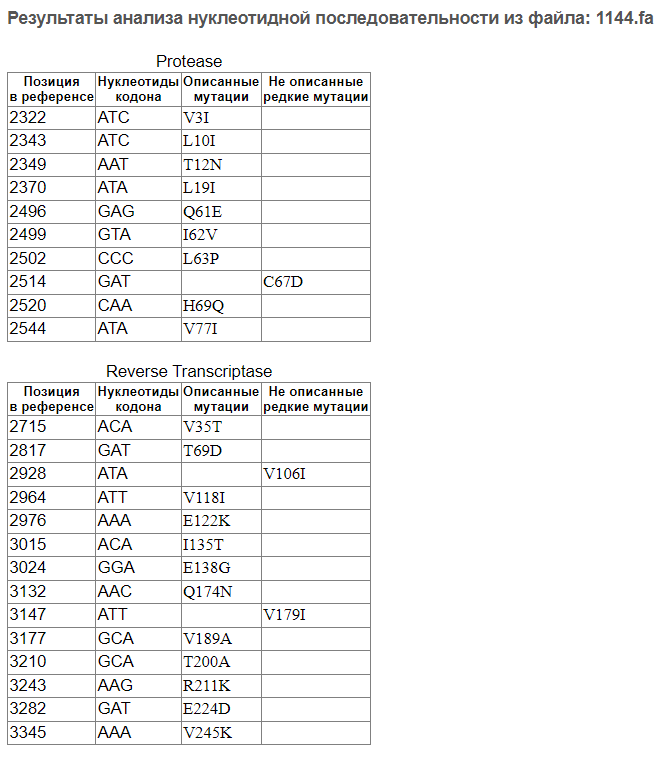
| ID пациента | Patient ID | Номер в GenBank | Accession number | Локус | Locus | Аминокислотные замены | Amino acid substitutions [[1]](#footnote-1) | Различия в третичной структуре субъединицы p51 ОТ в позициях 16-20 | Differences in the tertiary structure of RT p51 subunit in postions 16-20 | Различия в третичной структуре субъединицы p51 ОТ в позициях 210-235 | Differences in the tertiary structure of RT p51 subunit in postions 210-235 | Различия в третичной структуре субъединицы p66 ОТ в позициях 187-190 | Differences in the tertiary structure of RT p66 subunit in postions 187-190 |
| --- | --- | --- | --- | --- | --- | --- |
| p8 | OR260486 | плазма крови | T69S, Q242K |  | + | + |
| OR260531 | ликвор | E6D, K20E, E28K |
| p11 | OR260523 | плазма крови | T39D | + | + |  |
| OR260482 | ликвор | T39N |
| p26 | OR260512 | плазма крови | \_ | + | + |  |
| OR260533 | ликвор | E28K, K64R |
| p27 | OR260521 | плазма крови | K11A, T39K, V60I | + | + |  |
| OR260480 | ликвор | K11T, E36D, T39E, K64R |
| p53 | OR260516 | плазма крови | \_ |  | + |  |
| OR260535 | ликвор | K64R, A158S |
| p59 | OR260517 | плазма крови | V35K, T39K, I47M | + | + |  |
| OR260528 | ликвор | T39R, K64R, D67N, T200A |
| p95 | OR260507 | плазма крови | E40D, D86N, L214F | + | + | + |
| OR260508 | ликвор | \_ |
| p96 | OR260509 | плазма крови | V35T, T39M, S162H |  | + |  |
| OR260510 | ликвор | V35I, T39K, V60I, S162Y |
| p46 | OR260513 | плазма крови | V35K, F116Y |  |  | + |
| OR260534 | ликвор | T27P, V35T, S162C, D177N |

## 3.8 Web-сервис контроля сборки контигов фрагмента гена pol ВИЧ-1, кодирующих протеазу и обратную транскриптазу

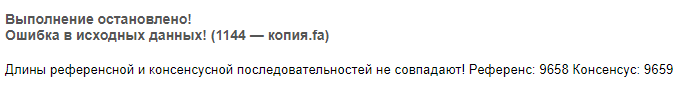
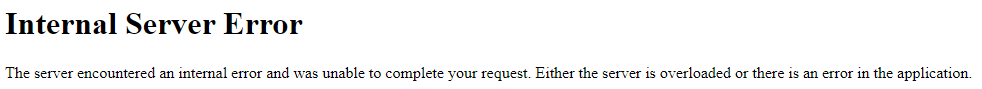
Ранее в отчете о научно-исследовательской работе «Риск-ориентированный подход к профилактике ВИЧ-инфекции в отдельных группах населения» по теме «Разработка подходов к выявлению групп риска первичной лекарственной устойчивости ВИЧ» от 2022 года сообщалось, что для анализа и обработки результатов секвенирования были разработаны базы данных и прикладное программное обеспечение: программа для ЭВМ «Программа для анализа первичных данных секвенирования региона pol ВИЧ-1» (при использовании Python 3.7, Biopython версии 1.75, свободно распространяемого ПО «Clustal W» версии 2.1).

К настоящему времени, на основе ранее разработанного ПО создан и зарегистрирован пользовательский web-сервис «Сервис анализа консенсусных последовательностей гена pol ВИЧ-1» (свидетельство о государственной регистрации для программы ЭВМ № 2023680626 от 03.10.2023 г.).

Программа предназначена для выявления мутаций во фрагментах гена pol ВИЧ-1, кодирующих протеазу и обратную транскриптазу, относительно референса ВИЧ-1 HXB2, с разделением на ранее описанные и нетипичные мутации и внедрена для практического использования в работе лаборатории УОЦС ФБУН ФНИИВИ «Виром» Роспотребнадзора. На постоянной основе ею пользуются различные сотрудники: врачи, научные сотрудники, эпидемиологи для оценки качества собранных нуклеотидных последовательностей и поиска ранее неописанных в литературе мутаций с целью их открытия. ПО зарекомендовало себя надежным сервисом и ежедневно помогает принимать решения в спорных моментах во время сборки геномов.

* ПО реализовано в виде web-сервиса, доступ к которому осуществляется по ссылке (<http://hivseqassist.niivirom.ru/>). На вход принимается собранная из отдельных ридов консенсусная последовательность, выравненная на референс с сохранением пробелов. Длина последовательности вместе с пробелами должна быть равна длине выбираемой из выпадающего списка референсной последовательности (9719 п.н для HXB2, 9658 п.н. для KU749403). Данные вводятся в формате fasta (fa, FASTA).  
  
* После загрузки исследуемой последовательности алгоритм производит ее выравнивание на один из ранее выбранных референсов, попарное сравнение каждого нуклеотида, определение состава кодонов и трансляцию их в аминокислотную последовательность. Некоторые из данных механизмов осуществляются посредством работы библиотеки Biopython, специально разработанной для работы с нуклеотидными последовательностями. Отчет работы библиотеки содержит 2 принимаемые на вход нуклеотидных последовательности, 2 соответственно транслированные аминокислотных последовательности. Определение состава кодонов задано вручную в коде программы и вычисляется арифметически.  
  
* После формируется подробный отчет об описанных и редких и ранее не встречавшихся в литературе мутациях отдельно для каждого фермента: протеазы и обратной транскриптазы ВИЧ.   
  Попарное сравнение нуклеотидного состава кодонов двух последовательностей осуществляется с использованием специально созданной базы данных, которая содержит информацию о порядковом номере кодона, порядковых номерах нуклеотида внутри всей последовательности и внутри отдельно взятого кодона, трансляции кодона относительно референса, ранее описанных мутациях для каждой позиции.   
  Алгоритм последовательно идет по каждому нуклеотиду, ищет несовпадения и сравнивает их с описанной базой. Если замена имеется в базе – она классифицируется как описанная мутация, если нет – мутации присваивается статус необычной или редкой.  
  Для каждой нуклеотидной замены на экран выводится ее позиция в референсе, благодаря этим данным возможна быстрая навигация по редактируемой последовательности; нуклеотиды кодона, для точного понимания границ интересующей области; описание аминокислотной замены (АК референса, порядковый номер АК, АК консенсуса).  
    
  

Список возможных ошибок:

1. Наблюдается при несовпадении количества символов в загружаемой последовательности и выбранной в качестве референса  
   
2. Наблюдается, если в последовательности не содержится данных для протеазы или ревертазы ВИЧ, т.е. не может быть произведено сравнение со специально созданной базой данных и формирование отчета  
   

Аналогом разработанного сервиса является HIVdb Program: Sequence Analysis (<https://hivdb.stanford.edu/hivdb/by-sequences/>) Стэндфордского университета на основе Hiv drug resistance database того же источника.   
Вышеописанное ПО так же опирается на использованную в аналоге базу данных и имеет дополнительный источник: The HIV mutation browser (https://www.hivmut.org/) - база данных информации о мутагенезе и мутациях вируса иммунодефицита человека (ВИЧ), собранная из научной литературы с использованием инструментов компьютерного анализа текста. Преимуществом сервиса так же является:

* автообновляемость указанных баз данных,
* возможность выбора референсных последовательностей среди HXB2 и специфичного для Российского региона ВИЧ субтипа А6,
* подробная форма отчета позволяет легко и быстро редактировать последовательность, основываясь на попадании мутации в раздел ранее не описанных. Для этого на экран выводится позиция нуклеотидной замены в референсе, что не предоставляет аналоговый сервис, и состав кодона.

Для разработки сервиса были использованы языки Python 3.7, SQL, HTML, объём программы: 1354 Kб. ПО совместимо со следующими ОС: Linux, MS Windows 7/8/8.1/10/11. Для доступа необходимы ПЭВМ или мобильное устройство с браузером (тестирование проводилось на браузерах на основе Google Chrome: Chrome, Yandex) и доступом в интернет.

Путями для развития сервиса без вмешательства в программный код являются:

* создание более наглядного приятного интерфейса,
* реализация возможности скачивания отчета в формате pdf, xlsx, txt, png\jpeg, xml,
* перевод сервиса на базу сайта ФБУН ФНИИВИ «Виром» Роспотребнадзора (<http://eniivi.niivirom.ru/>),
* осуществление массового доступа к программе.

# Заключение

Филогенетический анализ фрагмента гена pol ВИЧ-1, кодирующего протеазу и часть обратной транскриптазы, демонстрирует высокую степень гетерогенности, часть геномов кластеризуются с близкородственными штаммами, циркулирующими как в странах ближнего зарубежья на Украине и в Киргизии, так и в Центральной Европе: Польше и Германии. Это может свидетельствовать о высокой частоте заносов ВИЧ-инфекции в регион из-за рубежа.

Cравнение участков гена pol ВИЧ-1 (кодирующих протеазу и часть обратной транскриптазы) между изолятами из плазмы крови и из ликвора выявило у изолятов от 5 пациентов значительные генетические дистанции между геномами ВИЧ-1 (в одном случае штаммы принадлежали разным субтипам – А6 в плазме и B в ликворе).

Количество независимых аминокислотных замен в участке, кодирующем вирусную протеазу, у изолятов из плазмы крови составляло от 1 до 3, у изолятов из ликвора – от 1 до 2. Количество таких замен в участке, кодирующем фрагмент обратной транскриптазы ВИЧ-1 в изолятах из плазмы крови составляло от 1 до 6, а в изолятах из ликвора от 1 до 7.

Были обнаружены высоковариабельные участки в структуре субъединицы p51 обратной транскриптазы ВИЧ-1 в аминокислотных позициях 16-20 и 210-235. У 5 пациентов структура фермента у изолятов из плазмы крови и из ликвора отличалась в обеих позициях, еще у 4 – только в позициях 210-235. У 3 пациентов третичная структура субъединицы p66 обратной транскриптазы ВИЧ-1 различалась в аминокислотных позициях 187-190. Данная область входит в участок связывания с ННИОТ. Значительно большая разница в третичной структуре наблюдалась в субъединице p51, которая, как известно, не обладает каталитической активностью, однако играет важную структурную роль при формировании комплекса обратной транскрипции.

В третичной структуре вирусной протеазы различий между изолятами из плазмы крови и из ликвора выявлено не было.

Описанные наблюдения подтверждают наличие микроэволюционного процесса вируса, проявляющегося в изменениях как первичной, так и третичной структур ревертазы ВИЧ, идущего параллельно и независимо в организме одного пациента в разных компартментах, разделенных ГЭБ.

1. Аминокислотные замены относительно референсной последовательности HXB2 | Amino acid substitutions relative to HXB2 reference sequence [↑](#footnote-ref-1)