

Анализ вторичной структуры нетранслируемых регионов РНК флавиподобных вирусов с сегментированным геномом

Алхиреев Д.^{1,2}, Гладышева А.^{1*}

¹ Государственный научный центр вирусологии и биотехнологии «Вектор» Роспотребнадзора, р.п. Кольцово, Россия

² Новосибирский государственный университет, Новосибирск, Россия

* gladysheva_av@vector.nsc.ru

Ключевые слова: РНК-содержащие вирусы; *Flaviviridae*; Jingmen tick virus; 5'–3' UTR

Мотивация и цель: За последние пять лет с помощью высокопроизводительного секвенирования было обнаружено множество новых вирусов семейства *Flaviviridae*, обладающих атипичной структурой генома, что бросило вызов традиционным принципам классификации вирусов. Группа Jingmenvirus (JMV) состоит из новых РНК-содержащих вирусов, переносимых членистоногими, с неустановленной на данный момент патогенностью [1]. Kindia tick virus (KITV) – новый неклассифицированный флавиподобный вирус, относящийся к группе JMV семейства *Flaviviridae*, впервые обнаруженный в 2017 г. в иксодовых клещах на территории г. Киндия Гвинейской Республики [2, 3]. Он имеет сегментированный оцРНК(+) геном, состоящий из четырех сегментов, фланкированных 5' и 3' нетранслируемыми регионами (UTR). Два сегмента генетически и функционально связаны с генами, кодирующими неструктурные белки ортофлавиринов, роль других двух сегментов не установлена. Этот вирус имеет необычное устройство генома, в котором структура и роль 5' и 3' UTR вирусной РНК в реализации его генетической информации на сегодняшний день фактически полностью остаются неизученными. Поэтому целью данной работы стало изучение вторичной структуры 5' и 3' UTR РНК сегментированных флавиподобных вирусов на примере нового вируса KITV.

Методы и алгоритмы: Новые неразрешенные мотивы 5'–3' UTR (повторяющиеся шаблоны фиксированной длины) были найдены с использованием пакета программного обеспечения, предоставленного веб-сервером MEME Suite (версия 5.5.2). Прогнозирование структур 5'–3' UTR и выравнивание последовательностей выполнялись с использованием инструмента LocARNA. Полученные структурные выравнивания визуализировались как вторичные структуры РНК с помощью RNaPlot из пакета ViennaRNA с использованием режимов «наиболее информативной последовательности» и «аннотирования ковариации пар оснований» и алгоритма построения графиков RNApuzzler. Вторичные структуры 5' и 3' UTR РНК KITV и их комплексов совместно со шпилькой MS2 были предсказаны с использованием следующих независимых инструментов: ViennaRNA Fold, UNA MFOLD 3.6 и RNAstructure. Моделирование проводили при температуре 37 °C и в ионных условиях 1 М NaCl без двухвалентных ионов. Параметры «максимальное расстояние между парными основаниями» (MDBPB) и «процент субоптимальности» (%S) подбирались вручную. Были установлены MDBPB от 60 до 100 и %S до 50 %. Верхняя граница количества вычисленных складок и верхняя граница общего количества одноцепочечных оснований,

которые разрешены в выпуклости или внутренней петле, были установлены на уровне 25. Остальные параметры были установлены по умолчанию, а начальная свободная энергия ΔG – на минимальное значение. Для детального анализа вторичные структуры были линейаризованы с помощью программы VARNA 3.9 и перерисованы в графических редакторах. Моделирование белковых структур производили в AlphaFold 2.

Сборку ДНК копий 5' и 3' UTR РНК KITV, их комплексов совместно со шпилькой MS2 и гена, кодирующего белок оболочки фага MS2, осуществляли методом сборочной ПЦР с использованием высокоточной ДНК полимеразы Q5 (NEB, США) и олигонуклеотидов длиной 60 оснований и последующим клонированием продуктов амплификации в плазмидный вектор рJET. Для создания экспрессирующей генетической конструкции использовали плазмидный вектор рEASY.

Результаты: В настоящей работе проведены моделирование и функциональная аннотация вторичных структур 5' и 3' UTR вирусов группы JMV и KITV, а также получены синтетические ДНК копии 5' и 3' UTR РНК KITV и произведен дизайн комплексов РНК-белок для последующей кристаллизации и PCA.

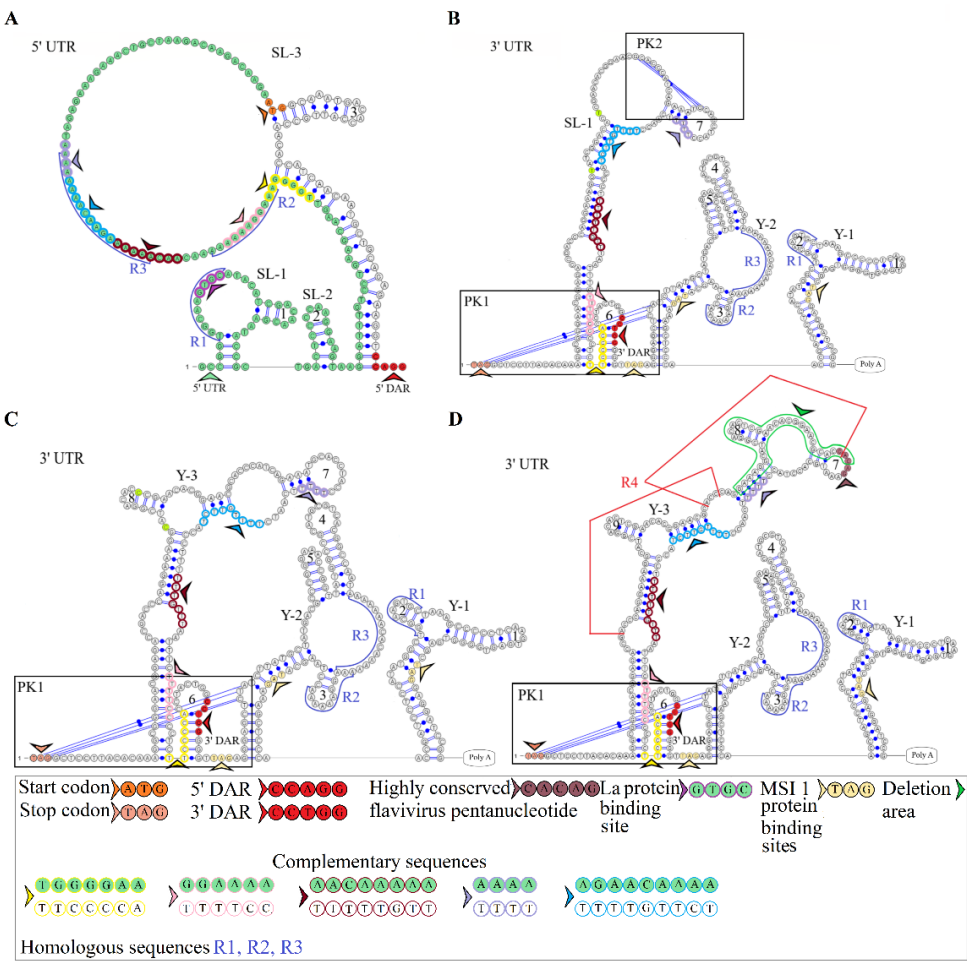


Рис. 1. Линейные модели вторичной структуры 5' UTR и 3' UTR сегмента 2 геномной РНК KITV. (A) 5' UTR изолята KITV/2018/1. (B) 3'UTR изолята KITV/2018/1. (C) 3' UTR изолята KITV/2018/2. (D) 3'UTR изолята KITV/2017/1

В ходе моделирования и анализа вторичных структур 5' и 3' UTR РНК KITV установлено, что для UTR различных сегментов KITV характерны: 1) полиаденированные 3' UTR; 2) мотивы 5' DAR и 3' DAR; 3) высококонсервативный пентануклеотид 5'-CACAG-3'; 4) регион-связывающий белок La; 5) множественные сайты UAG обеспечивающие взаимодействие с белком MSI1; 6) три гомологичные последовательности в 5' UTR и 3' UTR сегмента 2; 7) 3' UTR сегмента 2 изолята KITV/2017/1 содержит два последовательных повтора по 40 нуклеотидов, образующих Y-3 структуру; 8) делеция в 35 нуклеотидов во втором повторе 3' UTR 2-го сегмента изолятов KITV/2018/1 и KITV/2018/2, приводящая к модификации Y-3 структуры; 9) два псевдоузла PK1 и PK2 в 3' UTR сегмента 2; 10) 5' UTR и 3' UTR представлены паттернами консервативных мотивов; 11) расположение последовательности 5'-AAGUG-3' в первых шпильках UTR (рис. 1) [4]. Далее для получения экспериментальных данных о структуре 5' и 3' UTR были собраны полноразмерные ДНК копии 5' и 3' UTR РНК всех четырех сегментов KITV, содержащие T7 промоторную область для последующей транскрипции и область шпильки размером 18 нуклеотидов, взаимодействующей с белком оболочки фага MS2 для последующей кристаллизации комплексов РНК-белок. Также рекомбинантным путем в прокариотической системе клеток *E. coli* получен белок оболочки фага MS2, содержащий аффинную гексагистидиновую метку (6xHis-MS2). Моделированием структуры в AlphaFold 2 установлено, что химерный белок 6xHis-MS2 высокоструктурирован и аффинная метка не закрывает сайт связывания белка оболочки фага MS2 со специфичной 18-нуклеотидной шпилькой в составе 5' и 3' UTR РНК KITV.

Выводы: Таким образом, выявлены регуляторные элементы в структуре 5' и 3' UTR РНК KITV характерные для ортофлавириусов, что позволяет предположить их функциональную значимость для репликации вирусов группы JMV и эволюционную схожесть ортофлавириусов с флавигоподобными сегментированными вирусами. Далее планируется кристаллизация полученных комплексов белка оболочки фага MS2 с 5' и 3' UTR РНК KITV и их рентгеноструктурный анализ. Это позволит экспериментально подтвердить структуру 5' и 3' UTR РНК KITV и установить роль нетранслируемых регионов вирусной РНК в реализации генетической информации новых флавигоподобных вирусов с сегментированным геномом.

Финансирование: Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации (соглашение от 12.10.2021 № 075-15-2021-1355) в рамках реализации отдельных мероприятий Федеральной научно-технической программы развития синхротронных и нейтронных исследований и исследовательской инфраструктуры на 2019–2027 гг.

Analysis of the secondary structure of the untranslated RNA regions of flavi-like viruses with a segmented genome

Alkhireenko D.^{1,2}, Gladysheva A.^{1*}

¹ State Research Center of Virology and Biotechnology "Vector" Rospotrebnadzor, Koltsovo, Russia

² Novosibirsk State University, Novosibirsk, Russia

* gladysheva_av@vector.nsc.ru

Key words: RNA viruses; *Flaviviridae*; Jingmenvirus; 5'–3' UTR

Motivation and Aim: Over the past five years, high-throughput sequencing has discovered many new viruses in the family *Flaviviridae* that have atypical genome structures, challenging traditional principles of virus classification. The Jingmenvirus (JMV) group consists of novel arthropod-borne RNA viruses of currently unknown pathogenicity [1]. Kindia tick virus (KITV) is a new unclassified flavi-like virus belonging to the JMV group of the *Flaviviridae* family, first discovered in 2017 in ixodid ticks in the city of Kindia, Republic of Guinea. It has a segmented ssRNA(+) genome consisting of 4 segments flanked by 5' and 3' untranslated regions (UTR). Two segments are genetically and functionally related to genes encoding non-structural proteins of orthoflaviviruses; the role of the other two segments has not been established [2, 3]. This virus has an unusual genome structure, in which the structure and role of the 5' and 3' UTR of the viral RNA in the implementation of its genetic information remains virtually completely unexplored to date. Therefore, the goal of this work was to model and analyze the secondary structure of the 5'–3' UTR RNA of segmented flavi-like viruses using KITV as an example.

Methods and Algorithms: Novel, unresolved 5'–3' UTR motifs (recurring, fixed-length patterns) were searched using a software package provided by the MEME Suite web server (version: 5.5.2). Prediction of 5'–3' UTR structures and that of sequence alignment were performed using the LocARNA tool. The resulting structural alignments were visualized as RNA secondary structures with RNAplot from the ViennaRNA package using the “most informative sequence” and “annotate covariance of base pairs” modes and the RNApuzzler plotting layout algorithm. Secondary structures of the KITV RNA 5' and 3' UTR and their complexes with MS2 coat protein were predicted using three independent tools as follows: ViennaRNA Fold, UNA MFOLD 3.6, and RNAstructure. Simulation was performed at a folding temperature of 37°C and under ionic conditions of 1 M NaCl without divalent ions. The parameters of “maximal distance between paired bases” (MDBPB) and “percent suboptimality” (%S) were selected manually. An MDBPB of 60 to 100 and a %S of up to 50 % were established. The upper bound on the number of computed folds and the upper bound on the total number of single-stranded bases that are allowed in a bulge or interior loop were set at 25. Other parameters were set as default, and the initial free energy ΔG was set as the minimum value. For detailed analysis, secondary structures were linearized using the VARNA 3.9 software and redrawn using graphic editors. Modeling of protein structures was carried out in AlphaFold2.

Assembly of DNA copies of the 5'–3' UTR of KITV RNA, their complexes together with the MS2 hairpin and the gene encoding the MS2 coat protein was carried out by assembly PCR using high-precision Q5 DNA polymerase (NEB, USA) and oligonucleotides 60 bases long and subsequent cloning amplification products into the pJET plasmid. The pEASY plasmid was used to create an expression genetic construction.

Results: In this work, modeling and functional annotation of the 5' and 3' UTR secondary structures of the JMV group and KITV viruses were carried out, as well as synthetic DNA copies of the 5'–3' UTR RNA KITV were obtained and RNA-protein complexes were designed for crystallization and X-ray diffraction analysis.

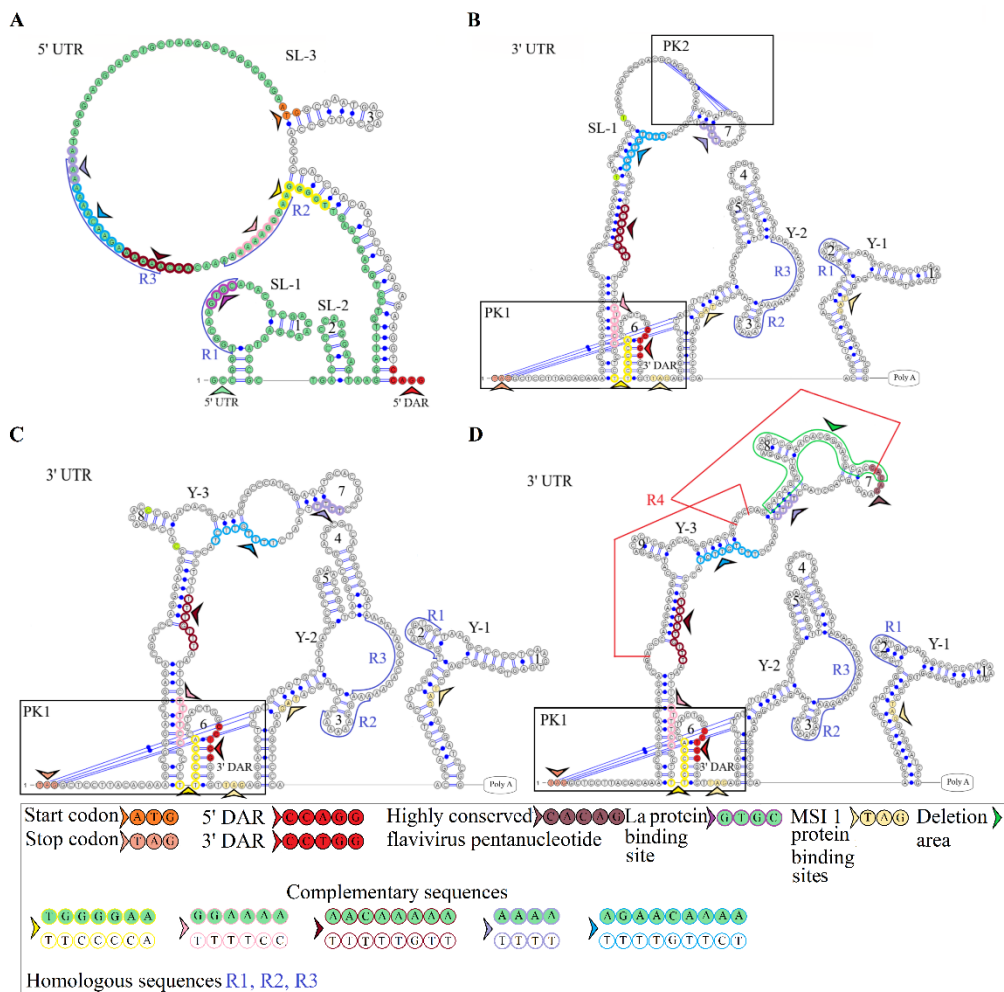


Fig. 1. Linear models of the secondary structure of 5' UTR and 3' UTR of genomic RNA segment 2 of KITV. (A) 5' UTR of the KITV/2018/1 isolate. (B) 3' UTR of the KITV/2018/1 isolate. (C) 3' UTR of the KITV/2018/2 isolate. (D) 3' UTR of the KITV/2017/1 isolate

UTRs of various KITV segments are characterized by the following points: (1) the polyadenylated 3' UTR; (2) 5' DAR and 3' DAR motifs; (3) a highly conserved 5'-CACAG-3' pentanucleotide; (4) a binding site of the La protein; (5) multiple UAG sites providing interactions with the MSI1 protein; (6) three homologous sequences in the 5' UTR and 3' UTR of segment 2; (7) the segment 2 3' UTR of a KITV/2017/1 isolate, which comprises two consecutive 40 nucleotide repeats forming a Y-3 structure; (8) a 35-nucleotide deletion in the second repeat of the segment 2 3' UTR of KITV/2018/1 and KITV/2018/2 isolates, leading to a modification of the Y-3 structure; (9) two pseudoknots in the segment 2 3' UTR; (10) the 5' UTR and 3' UTR being represented by

patterns of conserved motifs; (11) the 5'-CAAGUG-3' sequence occurring in early UTR hairpins (Fig. 1) [4].

Next, full-length DNA copies of the 5'-3' UTR RNA of all four KITV segments, containing the T7 promoter region for subsequent transcription and an 18-nucleotide hairpin region that interacts with the MS2 coat protein for subsequent crystallization of RNA-protein complexes were constructed to obtain experimental data of the 5'-3' UTR RNA KITV structure. Also, the MS2 phage envelope protein containing a hexahistidine affinity tag (6xHis-MS2) was obtained recombinantly in the *E. coli* cell. It was established that the chimeric protein 6xHis-MS2 is highly structured and the affinity tag does not overlap the binding site of the MS2 coat protein with a specific 18-nucleotide hairpin in the 5'-3' UTR RNA KITV.

Conclusion: Thus, we identified regulatory elements in the UTRs of KITV, which are characteristic of orthoflaviviruses. This suggests that they hold functional significance for the replication of JMV and the evolutionary similarity between orthoflaviviruses and segmented flavi-like viruses. Next, we plan to crystallize complexes of the MS2 phage coat protein with the 5' and 3' UTR RNA KITV and their X-ray diffraction analysis. This will allow us to experimentally confirm the structure of the 5' and 3' UTR RNA KITV and establish the role of the viral RNA untranslated regions in the implementation of the genetic information of new flavi-like viruses with a segmented genome.

Funding: The study was supported by the Ministry of Science and Higher Education of the Russian Federation (agreement No. 075-15-2021-1355 dated October 12, 2021) as a part of the implementation of certain activities of the Federal Scientific and Technical Program for the Development of Synchrotron and Neutron Research and Research Infrastructure for 2019–2027.

Список литературы/References

1. Zhang X., Wang N., Wang Z., Liu Q. The discovery of segmented flaviviruses: implications for viral emergence. *Curr Opin Virol.* 2020;40:11-18. doi 10.1016/j.coviro.2020.02.001
2. Gladysheva A.A., Gladysheva A.V., Ternovoi V.A., Loktev V.B. Structural Motifs and Spatial Structures of Helicase (NS3) and RNA-dependent RNA-polymerase (NS5) of a Flavi-like Kindia tick virus (unclassified Flaviviridae). *Vopr Virusol.* 2023;68(1):7-17. doi 10.36233/0507-4088-142 (in Russian)
3. Kartashov M.Y., Gladysheva A.V. et al. Molecular and genetic characteristics of the multicomponent flavi-like Kindia tick virus (Flaviviridae) found in ixodes ticks on the territory of the Republic of Guinea. *Vopr Virusol.* 2023;67(6):487-495. doi 10.36233/0507-4088-145 (in Russian)
4. Tsishevskaya A.A., Alkhireenko D.A., Bayandin R.B., Kartashov M.Y., Ternovoi V.A., Gladysheva A.V. Untranslated Regions of a Segmented Kindia Tick Virus Genome Are Highly Conserved and Contain Multiple Regulatory Elements for Viral Replication. *Microorganisms.* 2024;12(2):239. doi 10.3390/microorganisms12020239