

СЕКВЕНИРОВАНИЕ И АННОТАЦИЯ ХЛОРОПЛАСТНОГО ГЕНОМА *TRITICUM TIMONOVUM* HESLOT ET FERRARY

© 2024 г. А. Р. Кулуев¹, Р. Т. Матниязов¹, Б. Р. Кулуев¹*, Л. Ю. Привалов¹, А. В. Чемерис¹

¹Институт биохимии и генетики — обособленное структурное подразделение Федерального государственного бюджетного научного учреждения Уфимского федерального исследовательского центра Российской академии наук, Уфа, 450054 Россия

*e-mail: kuluev@bk.ru

Поступила в редакцию 14.12.2023 г.

После доработки 10.01.2024 г.

Принята к публикации 16.01.2024 г.

Впервые был секвенирован хлоропластный геном синтетической октаплоидной пшеницы *Triticum timonovum* Heslot et Ferrary (линия к-43065, Франция). Секвенирование пластома *T. timonovum* проводилось на секвенаторе Genolab M (GeneMind, Китай). Сборка кольца хлоропластного генома осуществлялась с помощью программы NOVOwarp. Размер хлоропластного генома *T. timonovum* составил 136158 пн. Длина области инвертированных повторов составила 21552 пн, области SSC — 12795 пн и области LSC — 80257 пн. Были сравнены хлоропластные геномы *T. timonovum* и различных линий *T. timopheevii* из GenBank. Хлоропластный геном *T. timonovum* линии к-43065 оказался наиболее близок к пластома *T. timopheevii* с номером доступа AB976560.1 и отличался от него лишь наличием одной вставки А в позиции 47891.

Ключевые слова: пшеница, октаплоид, синтетические виды, *Triticum timopheevii*, *T. timonovum*, хлоропластный геном.

DOI: 10.31857/S0016675824070124 **EDN:** BGHYZA

Triticum timonovum Heslot et Ferrary (пшеница тимоновум) — октаплоидная пшеница, искусственно созданная путем удвоения числа хромосом *T. timopheevii* Zhuk. с использованием колхичина. Авторами [1] это растение было обозначено как абберантный тетраплоид. Известно, что пшеница тимоновум — это синтетический октаплоид ($2n = 56$), который характеризуется такими полезными признаками, как высокое содержание белка в зерне и хорошие хлебопекарные качества, а также способностью вызывать у межвидовых гибридов цитоплазматическую мужскую стерильность. Между тем эта пшеница характеризуется медленным темпом развития и меньшей урожайностью чем *T. timopheevii*, что, по-видимому, связано с высокой плоидностью *T. timonovum* [2].

Принимая во внимание генотип исходной пшеницы, геном *T. timonovum* должен быть обозначен как GGA¹A¹GGA¹A¹, но по данным Е. Бадаевой и др. [3, 4] следует обозначать как GGA¹A¹BBA¹A¹, так как несколько хромосом генома А и одна хромосома генома G были заменены гомеологами геномов А и В от некоей неидентифицированной полбы. В дополнение к этому, в исследовании кариотипа *T. fungicidum* Zhuk. геномы А, В, А¹ и G этой пшеницы были идентичны родительским видам *T.*

timopheevii и *T. carthlicum* Nevsk., но в то же время геномы А и В отличались от соответствующих геномов *T. timonovum*. Авторы исследования предположили, что геномы А и В в геномы *T. fungicidum* и *T. timonovum* были переданы разными видами пшениц. Синтетический октаплоид *T. timonovum* — малоизученное растение, и его геном пока не исследован, к тому же не были изучены возможные изменения хлоропластного генома этой пшеницы под действием такого сильного мутагена как колхичин при создании *T. timonovum* путем полиплоидизации *T. timopheevii*.

Цель настоящей работы — секвенирование и аннотация хлоропластного генома *T. timonovum*.

Семена *T. timonovum* были предоставлены Федеральным исследовательским центром Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР, Санкт-Петербург). Растения проращивались в теплице при 21 °С в течение трех недель, затем было собрано около 20 г зеленых листьев и помещено на 48 ч в холод (+4 °С) и темноту для уменьшения содержания крахмала. Гомогенизацию листьев проводили при помощи фарфоровой ступки и пестика в жидком азоте. Хлоропластная ДНК была выделена из

зеленых листьев с помощью метода сахарозного градиента [5] с некоторыми модификациями.

Подготовка библиотек ДНК для секвенирования проводилась методом *shotgun* при помощи набора SG GM Plus (Raissol, Россия). Сначала проводилась ферментативная фрагментация хлоропластной ДНК до размера фрагментов 350 пн с одновременным восстановлением концов. Далее проводили лигирование адаптеров и очистку продукта реакции на магнитных частицах Smart beads (Raissol, Россия). На заключительном этапе проводилась индексная ПЦР и очистка продукта реакции на магнитных частицах Smart beads (Raissol, Россия). Библиотеки ДНК пулировали и проводили секвенирование при помощи набора Genolab M Kit V1.0 100 M reads/flow cell на секвенаторе Genolab M (GeneMind, Китай). Режим секвенирования — 2' 150 пн, файлы .base были демультиплексированы на приборе Genolab M с получением файлов с расширением fastq. В общей сложности было получено 13 млн ридов. Удаление адаптеров приготовленной библиотеки из файлов формата fastq проводили при помощи программы Trimmomatic v. 0.22 [6]. Обрезанные риды были собраны в континги, которые объединялись и записывались в файл формата fasta, содержащий информацию о полной кольцевой молекуле хлоропластной ДНК, которая была сгенерирована при помощи программы NOVOwrap [7].

В качестве референсного генома был использован хлоропластный геном *T. timopheevii* с номером доступа KJ614408.1. Полный хлоропластный геном *T. timonovum* был аннотирован при помощи ресурса Chloroplast Genome Annotation, Visualization, Analysis, and GenBank Submission 2 (CPGAVAS2) (<http://47.96.249.172:16019/analyzer/home>) [8]. Кольцевую карту хлоропластного генома визуализировали при помощи ресурса Chloroplast (<https://irscope.shinyapps.io/Chloroplast>) [9]. Выравнивание нуклеотидных последовательностей хлоропластных геномов проводили при помощи MAFFT v.7 [10]. Филогенетическое дерево строили и визуализировали при помощи Archaeopteryx (bootstrap = 1000) [11]. В качестве внешнего вида для филогенетического анализа был взят *Secale cereale* subsp. *segetale* (MZ507427.1) из GenBank.

Нуклеотидная последовательность хлоропластного генома *T. timonovum* была депонирована в GenBank (номер доступа OR936056). Аннотацию хлоропластного генома проводили при помощи ресурса CPGAVAS 2, при этом размер хлоропластного генома *T. timonovum* составил 136158 пн. Длина области инвертированных повторов 21552 пн, области SSC — 12795 пн, а области LSC — 80257 пн. Содержание GC-пар во всем пластидном геноме — 38.3%. В области SSC содержание GC составляет 32.17%, а в области LSC — 36.25%. В области IR *T. timonovum* GC-содержание доходит до 43.92%.

У пшеницы тимоновум аннотировали 132 структурных гена, из которых 85 — белок-кодирующие гены, 31 ген тРНК и 4 гена рРНК. В том числе 7 генов, кодирующих белки (*rps19*, *rpl2*, *rpl23*, *ndhB*, *rps7*, *rps12*, *rps15*), 8 генов тРНК (*trnH-GUG*, *trnM-CAU*, *trnL-CAA*, *trnV-GAC*, *trnT-CGU*, *trnA-UGC*, *trnR-ACG*, *trnN-GUU*), 4 гена рРНК (*rRNA4.5*, *rRNA23*, *rRNA16* и *rRNA5*) были дублированы из-за того, что находятся в области повтора IR. При этом в SSC-области обнаружено 10 белок-кодирующих генов и 1 ген тРНК, в LSC-области — 68 белок-кодирующих генов и 22 гена тРНК. Кроме того, из 132 генов 11 имеют по одному интрону (*atpF*, *ndhB*, *petB*, *petD*, *rpl2*, *trnI-GAU*, *ndhA*, *rpl16*, *rps12A*, *rps16*, *trnG-UCC*) и 9 генов — по 2 интрона (*trnK-UUU*, *trnS-CGA*, *trnL-UAA*, *trnV-UAC*, *trnT-CGU*, *trnA-UGC*, *ycf3*). Самый крупный интрон (2559 пн) находится в гене *trnK-UUU*, внутри которого располагается другой ген — *matK*. Все полученные результаты представлены на рис. 1, а в виде кольцевой структуры, полученной при помощи онлайн-ресурса Chloroplast. Различные цветовые блоки на внешнем круге кольца отображают принадлежность генов к тем или иным функциональным группам.

На основе анализа нуклеотидных последовательностей полных хлоропластных геномов различных образцов *T. timopheevii* из GenBank, *T. turgidum* MG958546.1 и *T. timonovum* k-43065 (секвенированный нами) было построено филогенетическое дерево, которое показало, что *T. timonovum* k-43065 наиболее близок к *T. timopheevii* (AB976560.1) (рис. 1, б), при этом другие образцы *T. timopheevii* расположились на древе немного дальше. *T. turgidum* MG958546.1 и *Secale cereale* subsp. *segetale* (MZ507427.1), как и ожидалось, оказались внешними видами по отношению к исследованным пшеницам. Пластомы *T. timonovum* и *T. timopheevii* (AB976560.1) отличались только одной вставкой A в положении 47891, при этом другие образцы *T. timopheevii* отличались от этих близких по хлоропластному геному образцов пшениц различными делециями и вставками.

Ранее японскими авторами [12] на основе микросателлитного анализа хлоропластной ДНК большого числа образцов *T. timopheevii* из дикой флоры были выявлены три подгруппы, которые можно видеть также на построенном нами древе (рис. 1, б). Лишь образец с номером KJ614407.1 расположился отдельно от этих трех групп.

К сожалению, французские авторы [1] в своей статье не указали, какой конкретно образец тетраплоидной пшеницы *T. timopheevii* с геномом GA они использовали при создании октаплоидного вида *T. timonovum* с геномом GGAA. Однако по нуклеотидной последовательности пластидного генома этот октаплоид оказался наиболее близок к образцу пшеницы Тимофеева (AB976560.1), секвенированного грузинскими авторами [13]; однако

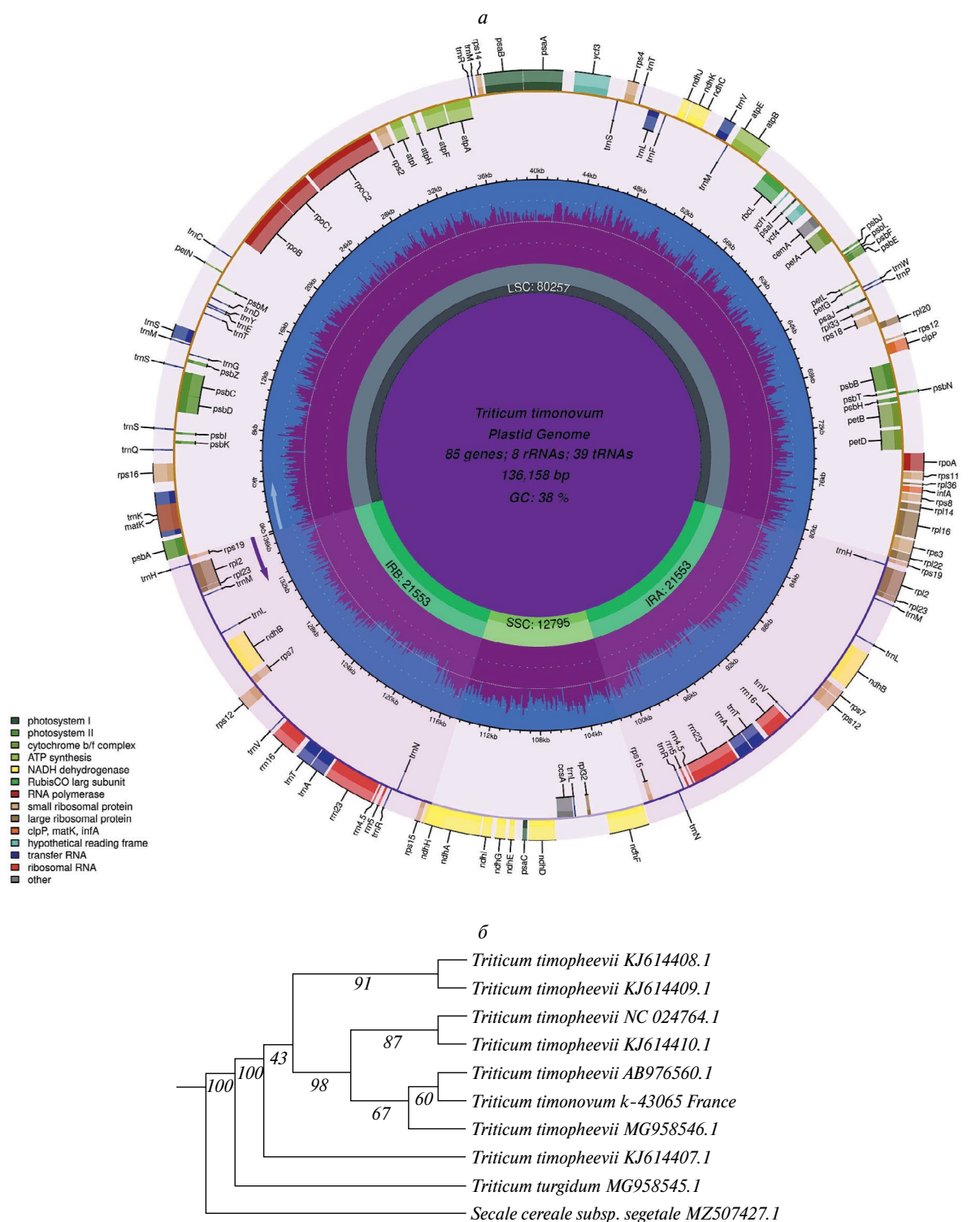


Рис. 1. *а* – визуальное представление в виде кольца секвенированного хлоропластного генома *T. timonovum* к-43065. Разными цветами отображены гены, синий круг посередине отображает уровень GC. IRA – инвертированная область повтора А, IRB – инвертированная область повтора В. Гены, расположенные за пределами внешнего круга, транскрибируются по часовой стрелке, а расположенные внутри гены – против часовой стрелки. *б* – филогенетическое древо, построенное на основе выравнивания нуклеотидных последовательностей хлоропластных геномов различных образцов *T. timopheevii* из GenBank, *T. turgidum* MG958546.1 и *T. timonovum* к-43065. В качестве внешнего вида представлен *Secale cereale* subsp. *segetale* (MZ507427.1).

в их статье также не сообщается ни каталожный номер, ни точное географическое происхождение секвенированного образца данной пшеницы.

Тем не менее, исходя из того, что от других представителей *T. timopheevii* хлоропластный геном пшеницы тимоновом отличается наличием ряда инделов, можно предположить, что именно образец *T. timopheevii* (AB976560.1) или близкий к нему был исходным при получении *T. timonovum*, при этом хлоропластный геном полученного октаплоида под действием колхицина сильным изменениям не подвергся.

Авторы выражают искреннюю признательность коллективу Отдела генетических ресурсов пшениц Всероссийского института генетических ресурсов растений им. Н.И. Вавилова (ВИР, Санкт-Петербург) за предоставление семенного материала для исследований.

Работа выполнена в рамках Государственного контракта по проекту РНФ № 23-24-00275 “Фило-генетические взаимоотношения отдельных видов пшенично-эгилопсного комплекса разных уровней плоидности через призму их полных хлоропластных геномов с прицелом на происхождение В- и G-субгеномов полиплоидных форм пшениц линий *turgidum-aestivum* и *timopheevii-zhukovskyi*”.

Настоящая статья не содержит каких-либо исследований с участием животных и людей в качестве объектов исследований.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Heslot H., Raymond R. Obtention experimentale d'un autotetraploide aberrant (*Triticum timonovum*) a partir de *Triticum timopheevi* Zhuk. // Compt. Rend. Hebd. Séances Acad. Sci. 1959. V. 248. P. 452–455.
2. Мурашëв В.В., Морозова З.А. Сравнительный морфогенез *Triticum timopheevii* (Zhuk.) и синтезированного октоплоидного вида *T. timonovum* Heslot et Ferrary // Вест. Моск. у-та. Серия 16. 2008. Т. 63. № 3. С. 127–133.
3. Badaeva E.D., Badaev N.S., Filatenko A.A. et al. Cytological investigation of cereal, hexa- and octoploid species containing G genome // Genetika (Mos.). 1990. V. 26. № 4. P. 708–716.
4. Badaeva E.D., Filatenko A.A., Badaev N.S. Cytogenetic investigation of *Triticum timopheevii* (Zhuk.) Zhuk. and related species using the C-banding technique // Theor. Appl. Genet. 1994. V. 89. P. 622–628.
5. Shi C., Hu N., Huang H. et al. An improved chloroplast DNA extraction procedure for whole plastid genome sequencing // PLoS One. 2012. V. 7. № 2. <https://doi.org/10.1371/journal.pone.0031468>
6. Bolger A.M., Lohse M., Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data // Bioinformatics. 2014. V. 30. P. 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
7. Wu P., Xu C., Chen H. et al. NOVOWrap: An automated solution for plastid genome assembly and structure standardization // Mol. Ecol. Resour. 2021. V. 21. № 6. P. 2177–2186. <https://doi.org/10.1111/1755-0998.13410>
8. Shi L., Chen H., Jiang M. et al. CPGAVAS2, an integrated plastome sequence annotator and analyzer // Nucl. Ac. Res. 2019. V. 47. P. W65–W73. <https://doi.org/10.1093/nar/gkz345>
9. Zheng S., Poczar P., Hyvönen J. et al. Chloroplast: An online program for the versatile plotting of organelle genomes // Front Genet. 2020. V. 11. <https://doi.org/10.3389/fgene.2020.576124>
10. Katoh K., Standley D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability // Mol. Biol. Evol. 2013. V. 30. № 4. P. 772–780. <https://doi.org/10.1093/molbev/mst010>
11. Han M.V., Zmasek C.M. phyloXML: XML for evolutionary biology and comparative genomics // BMC Bioinformatics. 2009. V. 10. P. 1–6. <https://doi.org/10.1186/1471-2105-10-356>
12. Mori N., Kondo Y., Ishii T. et al. Genetic diversity and origin of timopheevi wheat inferred by chloroplast DNA fingerprinting // Bred. Sci. 2009. V. 59. P. 571–578. <https://doi.org/10.1270/jsbbs.59.571>
13. Gogniashvili M., Naskidashvili P., Bedoshvili D. et al. Complete chloroplast DNA sequences of Zanduri wheat (*Triticum* spp.) // Genet. Resour. Crop. Evol. 2015. V. 62. P. 1269–1277. <https://doi.org/10.1007/s10722-015-0230-x>

Sequencing And Annotation Of The Chloroplast Genome Of *Triticum timonovum* Heslot et Ferrary

A. R. Kuluev¹, R. T. Matniyazov¹, B. R. Kuluev^{1, *}, L. Yu. Privalov¹, A. V. Chemeris¹

¹*Institute of Biochemistry and Genetics - Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, 450054 Russia*

**e-mail: kuluev@bk.ru*

The chloroplast genome of the synthetic octaploid *Triticum timonovum* Heslot et Ferrary k-43065 (France) was sequenced for the first time. Plastome sequencing was carried out on a Genolab M sequencer (GeneMind, China). The genome assembly was carried out using the NOVOwrap program. The size of the chloroplast genome of *T. timonovum* was 136158 bp. Meanwhile, the length of the inverted repeat region was 21552 bp, the SSC region was 12795 bp. and LSC – 80257 bp. The chloroplast genomes of *T. timonovum* and different *T. timopheevii* accessions from the GenBank database were compared. As for the chloroplast genome, *T. timonovum* was closer to *T. timopheevii* (AB976560.1), but differed from it by the presence of one insert A at position 47891.

Keywords: wheat, octaploid, synthetic species, *Triticum timopheevii*, chloroplast genome.