ГЕНЕТИКА / GENETICS

DOI: https://doi.org/10.18454/jbg.2023.22.6

COMPLETE GENOME SEQUENCE OF THE PREDATORY MITE *NEOSEIULUS CALIFORNICUS (MESOSTIGMATA, PHYTOSEIIDAE) FROM LABORATORY STRAIN (BIODEFENCE)*

Research article

Uroshlev L.A.[1], Vasilenko O.V.[2], Meshkov Y.I.[3], Andrianov B.V.[4,*]

[1] ORCID : 0000-0003-2878-5703;
[2] ORCID : 0000-0002-3333-5558;
[3] ORCID : 0000-0001-5034-2533;
[4] ORCID : 0000-0002-0064-4696;

[1, 4] Vavilov Institute of General Genetics Russian Academy of Sciences, Moscow, Russian Federation
[2] Skryabin Institute of Biochemistry and Physiology of Microorganisms of the Pushchino Scientific Center for Biological Research of the Russian Academy of Sciences, Puschino, Russian Federation
[3] All-Russian Scientific Research Institute of a Phytopathology, Moscow, Russian Federation

* Corresponding author (andrianovb[at]mail.ru)

Abstract

Neoseiulus californicus is an effective predatory mite for spider mites biocontrol. We sequenced and assembled the genomes of two isofemale lines of *N. californicus*: BioDefence and BioDefence2, derived from a biocontrol mite population, using the Oxford Nanopore long-read sequencing method. A total of 3507265 reads were obtained for the genome of the BioDefence line and 1769217 reads for the BioDefence2 line. Canu v.2.3 software was used for contig assembly. The genome of the BioDefence line was assembled as 419 contigs of 191.4 Mb in length, with a G+C content of 50.33%. The N50 is 13147 and the average coverage is 26.798. The annotation identified – 12253 genes, including 11607 protein-coding genes, 28 microsatellites, 9 families of small RNA genes, 2 families of LTR retrotransposons, 3 families of DNA transposons, 12 rRNA genes and 634 tRNA genes. The genome of line BioDefence2 was obtained using a similar procedure. Based on the contigs obtained, we assembled two new complete mitochondrial genomes of *N. californicus* from the BioDefence and BioDefence2 lines. We identified hypervariable regions of the genome within the mitochondrial control region that show inter-lineage variability. This study adds a new complete genome of *N. californicus* assembled at the contig level to the two other complete *N. californicus* genomes registered in GenBank and assembled at the contig level.

Keywords: Amblyseiinae, Oxford Nanopore MinION, long reads, genome assembly.

ЗАВЕРШЁННАЯ НУКЛЕОТИДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ПОЛНОГО ГЕНОМА ХИЩНОГО КЛЕЩА *NEOSEIULUS CALIFORNICUS* (*MESOSTIGMATA, PHYTOSEIIDAE*) ЛАБОРАТОРНОЙ ЛИНИИ (BIODEFENCE)

Научная статья

Урошлев Л.А.[1], Василенко О.В.[2], Мешков Ю.И.[3], Андрианов Б.В.[4,*]

[1] ORCID : 0000-0003-2878-5703;
[2] ORCID : 0000-0002-3333-5558;
[3] ORCID : 0000-0001-5034-2533;
[4] ORCID : 0000-0002-0064-4696;

[1, 4] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Российская Федерация
[2] Институт биохимии и физиологии микроорганизмов им. Г.К. Скрябина Пущинского научного центра Российской Академии наук, Пущино, Российская Федерация
[3] Всероссийский научно-исследовательский институт фитопатологии, Москва, Российская Федерация

* Корреспондирующий автор (andrianovb[at]mail.ru)

Аннотация

Neoseiulus californicus – хищный клещ, используемый для контроля численности паутинных клещей. Мы секвенировали и собрали геномы двух изосамочных линий *N. californicus*: BioDefence и BioDefence2, полученных от биоконтрольной популяции клещей, используя медод секвенирования Oxford Nanopore с длинными прочтениями. Всего для линии клещей BioDefence было получено 3507265 чтений и 1769217 чтения в случае генома линии BioDefence2. Для сборки контигов использовалась программа Canu v.2.3. Геном линии BioDefence собран в виде 419 контигов длиной 191,4 Мб с G+C содержанием 50.33%. N50 равно 13147. Были выявлены 12253 гена, включая 11607 белок кодирующих гена, 28 микросателлитов, 9 генов малых РНК, 2 семейства ДКП ретротранспозонов, 3 семейства ДНК транспозонов, 12 генов рРНК, 634 гена тРНК. Геном линии BioDefence2 был получен по аналогичной процедуре. На основе полученных контигов мы собрали два новых полных митохондриальных генома *N. californicus* линий BioDefence и BioDefence2. Мы выявили гипервариабельные области генома в составе митохондриального контрольного района проявляющие изменчивость на межлинейном уровне. Это исследование добавляет новый полный геном *N. californicus* собранный на уровне Contig, к двум другим полными геномам *N. californicus*, зарегистрированных в GenBank и собранных на уровне Contig.

Ключевые слова: Amblyseiinae, Oxford Nanopore MinION, длинные прочтения, сборка генома.

Introduction

Neoseiulus californicus McGregor, 1954 was initially discovered and named *Typhlodromus californicus* McGregor, 1954 in California. Today, *N. californicus* can be found in tropical and subtropical climates worldwide. This global distribution is likely due to human activities, as these mites are introduced into natural environments from greenhouses where they are imported as biocontrol agents. The reason for the widespread use of *N. californicus* in agriculture is its effectiveness in controlling both the common two-spotted spider mite *Tetranychus urticae* Koch, 1836, and the more harmful red tomato spider mite *Tetranychus evansi* Baker &Pritchard, 1960, particularly on solanaceous crops like tomatoes. N. californicus, a predator with limited specialization, not only preys on spider mites but also effectively controls thrips, whiteflies, tarsonemid mites, and other small pests that affect plants. In addition, *N. californicus* can survive without foraging mites by feeding on pollen grains, making it adaptable to varying conditions. Another advantage of *N. californicus* is its ability to thrive in both greenhouse and field environments, even in low humidity [1], [2]. One interesting question to consider is how the type of food affects the growth parameters of the biocontrol population. The rate of development, frequency of egg-laying, and foraging and feeding activity in *N. californicus* can be influenced by various factors, including the diversity of their diet, which includes multiple mite species rather than just one [3]. Similar to other mites in the Phytoseiidae family, *N. californicus* reproduces through pseudo-arrhenotoky, with diploid females and haploid males [4]. For genome sequencing purposes, we acquired *N. californicus* from the Laboratory of Acarology and Entomology at the All-Russian Scientific Research Institute of Phytopathology maintains a biocontrol population of N. californicus [5]. This population has shown resistance to various acaricides, although the genetic basis of this resistance has not been investigated. The *N. californicus* mother culture is grown under specific conditions of temperature, humidity, and light, with the spider mite Tetranychus urticae serving as the source of food for the culture. In this announcement, we present the complete genome sequence of two sublines of N. californicus. The BioDefence and BioDefence2 lines were derived from a single fertilised female. In the BioDefence line, the sex ratio is close to 2:1. In the BioDefence2 line, the sex ratio was close to 1:1. We obtained the nucleotide sequences of the complete genomes of both lines. Since the predominance of females is an economically valuable trait due to their high nutritional activity, the BioDefence line was selected for detailed annotation, which we present in this communication. We extracted mitochondrial sequences from complete genomic information and assembled complete mitochondrial genomes.

Research methods and principles

Genomic DNA extraction. Adult predatory mites were transferred to a culture vial, where they were starved for two days to remove possible DNA contamination from food mites. The mites were then collected in a lysis solution. High molecular weight DNA for whole genome sequencing was isolated using the classic alkaline phenol-chloroform method, followed by treatment with Proteinase K enzyme. Most of the RNA was removed by T1 ribonuclease treatment. The quality and purity of the extracted DNA were assessed by employing a NanoDrop spectrophotometer (Thermo Scientific). The concentration of DNA was measured using a Qubit fluorometer (Invitrogen) and reagent kits (Qubit dsDNA HS Assay Kit, Invitrogen). For sequencing and basecalling purposes, a library for nanopore sequencing preparation was created using the SQK-LSK110 kit from Oxford Nanopore Technologies, following the instructions provided by the manufacturer. Additionally, a final step of long DNA selection was performed using the large fragment buffer (LFB). To prepare the library, the ends of the DNA molecules were repaired and then linked with adapters, utilizing the NEBNext Companion Module. Sequencing was performed in a MinION R9.4.1 streaming cell using MinKNOW v.5.1.0 without immediate base calling. Reads longer than 200 bp were retained. Base calling was performed on a server with two GeForce RTX 3090 graphics cards and 12 Intel Xeon processors running Guppy v6.0.1+652ffd179 software under the super-neat sup model [6].

Main results

After base calling, we obtained 3507265 reads for the BioDefence lineage and 1769217 reads for the BioDefence2 lineage. The complete genomic sequence was then assembled from the reads using the canu v programme [7]. The genomic assembly of *N. californicus* was verified using BUSCO version 5.4.2 software [8], which contains a database of arachnid genes, the most evolutionarily stable genes against which the assembly is compared. A total of 2821 complete BUSCO (C) – 96.1% – were obtained. Of these, 2626 were complete and single-copy BUSCOs (S): 89.5%, 195 complete and duplicate BUSCOs (D): 6.6%, 45 fragmented BUSCOs (F): 1.5%, 68 missing BUSCOs (M): 2.4%. Total number of BUSCO groups searched n: 2934. 96% of *N. californicus* genes were found in the assembly. The californicus genes were correctly assembled. All identified genes were translated into amino acid sequence using AUGUSTUS gene annotation software [9]. The results were filtered using OrthoFinder software [10], which assembles orthologous genes into groups. The identified genes were aligned using Blast software in the "nr" database. Out of a total of 3507265 reads from the BioDefence genome, 922936 reads remain after assembly and contig correction. The N50 is 13147 and the average coverage is 26.861. The assembly resulted in 419 contigs with a total length of 191.4 Mb and a G+C content of 50.33%. Annotation identified 12253 genes, including 11607 protein-coding genes, 28 microsatellites, 9 small RNA gene families, 2 LTR retrotransposon families, 3 DNA transposon families, 12 rRNA genes and 634 tRNA genes. Blast software was used to search for contigs assembled into the mitochondrial genome. Mitochondrial genes were identified based on their structural features and comparison with mitochondrial genes of other Phitoseiidae species previously published in GenBan The mitochondrial genomes obtained from the mite lines BioDefence and BioDefence2 have been recorded in GenBank with the accession numbers OQ026345 and OR195436, respectively. The mitochondrial genome of the BioDefence line of *N. californicus* is 21318 bp long and has a high A + T content of 78.4%. It contains the typical genes found in arthropods, including 13 protein-coding genes, 2 rRNA genes, 22 tRNA genes, and a large control region measuring 4846 bp. The length of the control region varies between the different mite lines. The mitochondrial genomes of BioDefence and BioDefence2 are almost identical, except for the control region, which is larger in BioDefence2, spanning 6645 bp. This lineage-specific DNA polymorphism can be used as a marker for that lineage.

Discussion

GenBank currently has a contig level genome of N. contig. californicus from China [GenBank ID: GCA_028455905.1] and a raw read level genome from Japan [BioProject Accession ID: PRJNA966143]. The genome presented in this article adds

a third genome of *N. californicus* from the Moscow region. In addition to the assembly of the nuclear genome, we assembled the complete mitochondrial genome of N. californicus, which has a unique mitochondrial gene sequence and an extended control region that varies in length between tick strains due to different numbers of direct incomplete repeats. Recently, the mitochondrial genome of *N. californicus* was obtained by Zu using Illumina sequencing technology and deposited in GenBank under accession no: NC_069213. This mitochondrial genome and two new mitochondrial genomes from *N. californicus* have the same gene order and very similar nucleotide sequence, but differ in the length of control regions due to different numbers of direct incomplete repeats. High variability in mitochondrial gene order is a characteristic feature of mites of the family Phytoseiidae. Mites belonging to the family Phytoseiidae possess a standard set of mitochondrial genes, with the exception of Metaseiulus occidentalis Nesbitt, 1951, which lacks the ND6 gene. However, these mites differ in the order of their mitochondrial genes [11]. Additionally, the mitochondrial genome of *N. californicus* has a unique species-specific gene order.

Conclusion

The genome sequence of the *N. californicus* strain named BioDefence has been submitted to the GenBank database under the accession number GCA_028659405.1 (version GCA_028659405.1). The corresponding BioProject accession number is PRJNA902297, and the BioSample accession number is SAMN32036229. The raw sequencing data are available from the SRA (Sequence Read Archive) under accession number SRP411367. The genome sequence of the strain N. californicus: BioDefence2 has been deposited in the GenBank database under accession number GCA_031001865.1 (version GCA_031001865.1). The corresponding BioProject accession number is PRJNA946627 and the BioSample accession number is SAMN33824953. The raw sequencing data are also available from the SRA (Sequence Read Archive) under accession number SRP428294. The *N. californicus* genome sequence presented here contains the genome sequences of the mite and its bacterial symbionts, obligate intracellular symbionts such as *Wolbachia* and *Spiroplasma*, and facultative intestinal symbionts that have not been removed by starvation for two days. The presented mitochondrial genome of *N. californicus* offers valuable insights for research on the genetic regulation of economically important characteristics of *N. californicus* and its hologenome, such as the makeup of beneficial and harmful bacteria that greatly influence the survival, reproductive capacity, and nutritional behavior of biocontrol mite populations belonging to the Phytoseiidae family.

Финансирование

This work was financially supported by the Russian Science Foundation within scientific project no. 22-24-00727.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Funding

Работа выполнена при поддержке гранта Российского научного фонда (РНФ) грант 22-24-00727.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Akyazi R. Biological Control of the Twospotted Spider Mite (Trombidiformes: Tetranychidae) with the Predatory Mite Neoseiulus californicus (Mesotigmata: Phytoseiidae) in Blackberries / R. Akyazi, O. E. Liburd // Florida Entomologist. — 2019. — 2. — p. 373-381. — DOI: 10.1653/024.102.0217.

2. Sanchez N.E. Biological Control by Neoseiulus californicus (McGregor) (Acari: Phytoseiidae) / N.E. Sanchez, N.M. Greco, C.V. Cedola; ed. by J.L. Capinera. — Dordrecht: Springer, 2008. — p. 493-495. — DOI: 10.1007/978-1-4020-6359-6_319.

3. Badrbani F.K. Oviposition, Development and Predation Rates of Neoseiulus Californicus Fed on Red and Green Forms of Tetranychus Urticae / F.K. Badrbani, A. Z. Golpayegani, A. Saboori et al. // Systematic & Applied Acarology. — 2015. — 20. — p. 603–611. — DOI: 10.11158/saa.20.6.3.

4. Nagelkerke C.J. Precise Control of Sex Allocation in Arrhenotokous Phytoseiid Mites / C.J. Nagelkerke, M.W. Sabelis // Journal of Evolutionary Biology. — 2002. — 11. — p. 649-684. — DOI: 10.1046/j.1420-9101.1998.11060649.x.

5. Глинушкин А.П. Влияние пестицидов, применяемых в защищенном грунте, на хищного клеща neoseiulus californicus (parasitiformes, phytoseiidae) / А.П. Глинушкин, И.Н. Яковлева, Ю.И. Мешков // Российская сельскохозяйственная наука. — 2019. — 3. — с. 32-34. — DOI: 10.31857/S2500-26272019332-34.

6. Benton M. Nanopore Guppy GPU Basecalling on Windows Using WSL2 / M. Benton. — 2021 — URL: https://hackmd.io/PrSp6UhqS2qxZ_rKOR18-g#Nanopore-Guppy-GPU-basecalling-on-Windows-using-WSL2 (accessed: 03.08.2022)

7. Koren S. Canu: Scalable and Accurate Long-read Assembly via Adaptive k-mer Weighting and Repeat Separation / S. Koren, B.P. Walenz, K. Berlin et al. // Genome Research. — 2017. — 27. — p. 722–736. — DOI: 10.1101/gr.215087.116.

8. Manni M. Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes / M. Manni, M.R. Berkeley, M. Seppey et al. // Molecular Biology and Evolution. — 2021. — 38. — p. 4647–4654. — DOI: 10.1093/molbev/msab199.

9. Stanke M. AUGUSTUS: ab initio Prediction of Alternative Transcripts / M. Stanke, O. Keller, I. Gunduz et al. // Nucleic Acids Research. — 2006. — 1. — DOI: 10.1093/nar/gkl200.

10. Emms D.M. OrthoFinder: Phylogenetic Orthology Inference for Ccomparative Genomics / D.M. Emms, S. Kelly // Genome Biology. — 2019. — 20. — p. 238. — DOI: 10.1186/s13059-019-1832-y.

11. Zhang B. Massive Gene Rearrangement in Mitogenomes of Phytoseiid Mites / B. Zhang, J. C. Havird, E. Wang et al. // International Journal of Biological Macromolecules. — 2021. — 186. — p. 33-39. — DOI: 10.1016/j.ijbiomac.2021.07.011.

Список литературы на английском языке / References in English

1. Akyazi R. Biological Control of the Twospotted Spider Mite (Trombidiformes: Tetranychidae) with the Predatory Mite Neoseiulus californicus (Mesotigmata: Phytoseiidae) in Blackberries / R. Akyazi, O. E. Liburd // Florida Entomologist. — 2019. — 2. — p. 373-381. — DOI: 10.1653/024.102.0217.

2. Sanchez N.E. Biological Control by Neoseiulus californicus (McGregor) (Acari: Phytoseiidae) / N.E. Sanchez, N.M. Greco, C.V. Cedola; ed. by J.L. Capinera. — Dordrecht: Springer, 2008. — p. 493-495. — DOI: 10.1007/978-1-4020-6359-6_319.

3. Badrbani F.K. Oviposition, Development and Predation Rates of Neoseiulus Californicus Fed on Red and Green Forms of Tetranychus Urticae / F.K. Badrbani, A. Z. Golpayegani, A. Saboori et al. // Systematic & Applied Acarology. — 2015. — 20. — p. 603–611. — DOI: 10.11158/saa.20.6.3.

4. Nagelkerke C.J. Precise Control of Sex Allocation in Arrhenotokous Phytoseiid Mites / C.J. Nagelkerke, M.W. Sabelis // Journal of Evolutionary Biology. — 2002. — 11. — p. 649-684. — DOI: 10.1046/j.1420-9101.1998.11060649.x.

5. Glinushkin A.P. Vlijanie pestitsidov, primenjaemyh v zaschischennom grunte, na hischnogo klescha neoseiulus californicus (parasitiformes, phytoseiidae) [The Impact of Pesticides Used in Greenhouses, on the Predatory Mite Neoseiulus Californicus (Parasitiformes, Phytoseiidae)] / A.P. Glinushkin, I.N. Jakovleva, Ju.I. Meshkov // Russian Agricultural Science. — 2019. — 3. — p. 32-34. — DOI: 10.31857/S2500-26272019332-34. [in Russian]

6. Benton M. Nanopore Guppy GPU Basecalling on Windows Using WSL2 / M. Benton. — 2021 — URL: https://hackmd.io/PrSp6UhqS2qxZ_rKOR18-g#Nanopore-Guppy-GPU-basecalling-on-Windows-using-WSL2 (accessed: 03.08.2022)

7. Koren S. Canu: Scalable and Accurate Long-read Assembly via Adaptive k-mer Weighting and Repeat Separation / S. Koren, B.P. Walenz, K. Berlin et al. // Genome Research. — 2017. — 27. — p. 722–736. — DOI: 10.1101/gr.215087.116.

8. Manni M. Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes / M. Manni, M.R. Berkeley, M. Seppey et al. // Molecular Biology and Evolution. — 2021. — 38. — p. 4647–4654. — DOI: 10.1093/molbev/msab199.

9. Stanke M. AUGUSTUS: ab initio Prediction of Alternative Transcripts / M. Stanke, O. Keller, I. Gunduz et al. // Nucleic Acids Research. — 2006. — 1. — DOI: 10.1093/nar/gkl200.

10. Emms D.M. OrthoFinder: Phylogenetic Orthology Inference for Ccomparative Genomics / D.M. Emms, S. Kelly // Genome Biology. — 2019. — 20. — p. 238. — DOI: 10.1186/s13059-019-1832-y.

11. Zhang B. Massive Gene Rearrangement in Mitogenomes of Phytoseiid Mites / B. Zhang, J. C. Havird, E. Wang et al. // International Journal of Biological Macromolecules. — 2021. — 186. — p. 33-39. — DOI: 10.1016/j.ijbiomac.2021.07.011.