

Полногеномное секвенирование микроспоридий рода *Nosema* – внутриклеточных паразитов чешуекрылых насекомых

Малыш С.М.^{1*}, Румянцева А.С.¹, Сахабеев Р.Г.², Ефейкин Б.Д.³, Данилов Л.Г.⁴, Токарев Ю.С.¹

¹ Всероссийский институт защиты растений, Санкт-Петербург, Россия

² Санкт-Петербургский государственный технологический институт (технический университет), Санкт-Петербург, Россия

³ Институт проблем экологии и эволюции им. А.Н. Северцова РАН, Москва, Россия

⁴ Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

* malyshshvetlana@gmail.com

Ключевые слова: микроспоридии; эволюция генома; мультигенная филогения; мультилокусное генотипирование; генодиагностика; близкородственные виды

Мотивация и цель: Микроспоридии – родственные грибам протисты, наиболее широко распространенные в качестве паразитов членистоногих и других животных [1, 2]. В связи с адаптацией к внутриклеточному паразитизму, они характеризуются высоким уровнем редукции клеточной организации и компактизации генома и служат важной моделью исследований молекулярной и клеточной биологии. Практическое значение микроспоридий определяется способностью вызывать заболевания экономически значимых видов беспозвоночных и позвоночных, а также человека [3–5]. Описано свыше 200 родов микроспоридий, содержащих около 1300 видов, для более 50 из которых получены полногеномные сиквенсы [6]. В связи с таким маленьким количеством геномных последовательностей и все большим количеством новых изолятов возникают трудности с проведением видовой идентификации близкородственных видов микроспоридий. Цель работы – получение исходных данных и первичная оценка результатов полногеномного секвенирования микроспоридий *Nosema pyrausta* из кукурузного мотылька *Ostrinia nubilalis* и *Nosema* sp. NspHA22 из хлопковой совки *Helicoverpa armigera* в сравнении с референтными последовательностями *N. bombycis* из тутового шелкопряда *Bombyx mori* и *N. antheraeae* из китайского дубового шелкопряда *Antheraea pernyi*. Это необходимо для изучения внутривидового полиморфизма, поиска молекулярных маркеров для разграничения близкородственных видов, выявления генетических детерминант факторов вирулентности, филогеномного анализа и других исследований молекулярной генетики энтомопатогенных микроспоридий. В частности, известно, что по общепринятому диагностическому фрагменту гена рРНК вышеуказанные изоляты из тутового шелкопряда и хлопковой совки неотличимы друг от друга, однако их биологические свойства существенно различаются, что не позволяет относить их к одному виду [7].

Методы и алгоритмы: Для накопления спор использовали гусениц первого лабораторного поколения сибирского шелкопряда *Dendrolimus sibiricus*, любезно предоставленных Александром А. Агеевым (филиал ВНИИЛМ «Центр лесной пирологии», Красноярск). Гусениц второго возраста заражали в дозировке 1 млн

спор на гусеницу и содержали в чашках Петри на хвое пихты сибирской при +24 °С. Через 2 месяца гусениц вскрывали, препарировали ткани внутренних органов в дистиллированной воде и гомогенизировали вручную в стеклянном гомогенизаторе с тефлоновым пестиком, центрифугировали при 1000 g в течение 5 мин, осадок промывали в дистиллированной воде, удаляя верхние слои и повторно центрифугируя. Количество спор подсчитывали в камере Горяева. Образец, содержащий от 1 до 1.5 млрд чистых спор, растирали в фарфоровой ступке с добавлением жидкого азота и использовали для экстракции ДНК фенольно-хлороформным методом [8]. Качество геномной ДНК определяли с помощью гель-электрофореза и наноспектрофотометра Implen, а качество ридов – утилитой FastQC. Полногеномное секвенирование осуществляли на платформе Illumina Nextseq в центре коллективного пользования Сколковского института науки и технологий (Сколтех, Москва). Сборка генома осуществлялась программой MEGAHIT. Множественное выравнивание нуклеотидных последовательностей выполнялось с использованием алгоритма ClustalW, построение филогенетических деревьев осуществлялось методом максимального правдоподобия.

Результаты: Геном *N. pyrausta* депонирован в Генбанк под номером доступа GCA_001924031.1. Приблизительный размер генома составил 15.9 Мб, количество предсказанных генов 4468, что очень близко к соответствующим показателям близкородственного вида *N. bombycis*. При этом для другого вида этого рода, а именно *N. antheraeae*, размер генома примерно в четыре раза меньше (табл. 1).

Таблица 1. Основные характеристики генома трех близкородственных видов микроспоридий рода *Nosema*

Характеристика	<i>N. bombycis</i>	<i>N. antheraeae</i>	<i>N. pyrausta</i>
Номер доступа в Генбанке	NA30919	NA183977	GCA_001924031.1
Название изолята	NbCQ1	YY	Npyr2020
Размер генома, Мб	15.7	6.6	15.9
N50	57394	1883	8050
Количество генов	4458	3413	4468
GC состав, %	31	28	34.5
Количество контигов	3558	?	4115
Количество скэффолдов	1607	6215	1880
Ссылка	Pan et al., 2013	Pan et al., 2013	Собственные данные

Для генома NspHA22 на момент подачи тезисов сборка и аннотирование не завершены из-за более высокой примеси геномной ДНК хозяина, что требует дополнительной работы по фильтрации неспецифических сигналов. Тем не менее первичные полногеномные данные обоих изолятов микроспоридий использованы для извлечения нуклеотидных последовательностей отдельных генов, наиболее часто используемых для мультилокусного генотипирования и мультигенных филогенетических реконструкций: белка теплового шока HSP70 и РНК-полимеразы RPB1, а также гексокиназы, дополнительно предложенной для меж- и внутривидовой дифференциации микроспоридий [9]. Анализ соответствующих нуклеотидных последовательностей, изучаемых микроспоридий, показал их разделение на две сестринские филогенетические линии внутри рода *Nosema*. В

одну группу вошли *N. bombycis* и NspHA22, в другую – *N. pyrausta* и *N. antheraeae*. Сходство по каждому из трех генов внутри группы составило около 98–99 %, а между группами – 92–93 %. Это позволяет достаточно четко разграничить изучаемые изоляты, однако требует секвенирования соответствующих локусов для наиболее близкородственных форм, так как низкая частота нуклеотидных замен не дает возможности сконструировать праймеры, специфичные для конкретного вида (изолята).

Выводы: Считается, что такой относительно большой, по сравнению с другими видами рода *Nosema*, размер генома *N. bombycis* обусловлен недавним событием геномной дупликации [10]. Следовательно, данное событие затронуло как *N. bombycis*, так и *N. pyrausta*, а поскольку круг их восприимчивых хозяев существенно различается [11, 12], полученная информация может послужить дальнейшему поиску генетических детерминант вирулентности энтомопатогенных микроспоридий, вызывающих опасные заболевания культивируемых насекомых и требующих разработки подходов для профилактики и лечения [13], а также перспективных с точки зрения защиты растений [14]. Геномные данные послужат также для поиска вариабельных участков генома для экспресс-диагностики, в частности, для разработки видоспецифичных праймеров.

Финансирование: Исследование выполнено в рамках проекта РНФ № 23-16-00262 и проекта Госзадания лаборатории молекулярной защиты растений ФГБНУ ВИЗР № FGEU-2022-0019 (№ госрегистрации 122032900144-2).

Whole genome sequencing of microsporidia of the genus *Nosema* – intracellular parasites of lepidopteran insects

Malysh S.^{1*}, Rumiantseva A.¹, Sakhabeev R.², Efeikin B.³, Danilov L.⁴, Tokarev Y.¹

¹ All-Russian Institute of Plant Protection, Saint Petersburg, Russia

² St. Petersburg State Institute of Technology, Saint Petersburg, Russia

³ A.N. Severtsov Institute of Ecology and Evolution, Moscow, Russia

⁴ Saint Petersburg State University, Saint Petersburg, Russia

* malyshsvetlana@gmail.com

Key words: microsporidia; genome evolution; multigene phylogeny; multilocus sequence typing; molecular diagnostics; closely related species

Motivation and Aim: Microsporidia are protists related to fungi, most widely distributed as parasites of arthropods and other animals [1, 2]. Due to adaptation to intracellular parasitism, they are characterized by a high level of reduction of cellular organization and genome compaction and serve as an important model for research in molecular and cellular biology. The practical importance of microsporidia is determined by the ability to cause diseases of economically important species of invertebrates and vertebrates, as well as humans [3–5]. Over 200 microsporidia genera have been described, containing about 1300 species, for more than 50 of which full genome sequences have been obtained [6]. And because of such a low number of whole genome sequences and an increasing number of newly discovered isolates, difficulties began to arise in identifying the species in closely related microsporidian taxa. The aim of the work was to perform whole-genome sequencing of microsporidian isolates *Nosema pyrausta* from the European corn

borer *Ostrinia nubilalis* and *Nosema* sp. NspHA22 from the cotton bollworm *Helicoverpa armigera*. The reference sequences were those of *N. bombycis* from the silkworm *Bombyx mori* and *N. antheraeae* from the Chinese oak silkworm *Antheraea pernyi* to study intraspecific polymorphism, search for molecular markers to distinguish closely related species, identify genetic determinants of virulence factors, phylogenomic analysis and other studies of the molecular genetics of entomopathogenic microsporidia. In particular, it is known that according to the generally accepted diagnostic fragment of the rRNA gene, the above isolates from the silkworm and the cotton bollworm are indistinguishable from each other, but their biological properties differ significantly, which does not allow them to be classified as the same species [7].

Methods and Algorithms: To accumulate spores, we used caterpillars of the first laboratory generation of the Siberian silkworm *Dendrolimus sibiricus*, kindly provided by Alexander A. Ageev (branch of the VNIILM Center for Forest Pyrology, Krasnoyarsk). Second instar caterpillars were infected at a dosage of 1 million spores per caterpillar and kept in Petri dishes on Siberian fir needles at +24 °C. After 2 months, the caterpillars were dissected, the tissues of the internal organs were dissected in distilled water and homogenized manually in a glass homogenizer with a Teflon pestle, centrifuged at 1000 g for 5 min, the sediment was washed in distilled water, removing the upper layers and centrifuging again. The number of spores was counted in Goryaev's chamber. A sample containing 1 to 1.5 billion purified spores was ground in a porcelain mortar with liquid nitrogen and used for DNA extraction using the phenol-chloroform method [8]. The quality of genomic DNA was determined using gel electrophoresis and Implen nanospectrophotometer, and the quality of reads was determined using the FastQC utility. Whole-genome sequencing was carried out on the Illumina Nextseq platform at the shared use center of the Skolkovo Institute of Science and Technology (Skoltech, Moscow) using the MEGAHIT program. Multiple alignment of nucleotide sequences was performed using the ClustalW algorithm, and the phylogenetic trees were constructed using the maximum likelihood method.

Results: *Nosema pyrausta* was deposited in the GenBank under accession number GCA_001924031.1. The approximate genome size was 15.9 Mb, the number of predicted genes was 4468, which is very close to the corresponding indicators of the closely related species *N. bombycis* (Table 1).

Table 1. Main characteristics of the genome of three closely related species of microsporidia of the genus *Nosema*

Metrics	<i>N. bombycis</i>	<i>N. antheraeae</i>	<i>N. pyrausta</i>
GenBank accession number	NA30919	NA183977	GCA_001924031.1
Isolate name	NbCQ1	YY	Npyr2020
Genome size, Mb	15.7	6.6	15.9
N50	57394	1883	8050
Number of genes	4458	3413	4468
GC, %	31	28	34.5
Number of contigs	3558	?	4115
Number of scaffolds	1607	6215	1880
Link	Pan et al., 2013	Pan et al., 2013	Own data

The *Nosema* sp. NspHA22 genome assembly and annotation are not finished by the time of submission of the present work due to higher admixture of host genomic DNA, which requires additional work to filter the nonspecific signals out. Anyway, the raw whole genome data from both microsporidian isolates were used to extract nucleotide sequences of individual genes most often used for multilocus genotyping and multigene phylogenetic reconstructions: heat shock protein HSP70 and RNA polymerase RPB1, as well as hexokinase, additionally proposed for inter- and intraspecific differentiation of microsporidia [9]. Analysis of the corresponding nucleotide sequences of the studied microsporidia showed their division into two sister phylogenetic lineages within the genus *Nosema*. One group included *N. bombycis* and NspHA22, the other included *N. pyrausta* and *N. antheraeae*. The similarity for each of the three genes within the group was about 98–99 %, and between groups – 92–93 %. This makes it possible to fairly clearly distinguish between the studied isolates, but requires sequencing of the corresponding loci for the most closely related forms, since the low frequency of nucleotide substitutions excludes the ability to design primers specific to a particular species (isolate).

Conclusion: The relatively large genome size of *N. pyrausta* compared to other *Nosema* species is believed to be such due to a recent genomic duplication event [10]. Consequently, this event affected both species under consideration, and since their range of susceptible hosts differs significantly [11, 12], the information obtained can serve in further search for genetic determinants of virulence of entomopathogenic microsporidia that cause dangerous diseases of cultivated insects and requiring the development of approaches for prevention and treatment [13], as well as promising from the point of view of plant protection [14]. Genomic data will also facilitate the search for variable regions of the genome for express diagnostics, in particular, for the development of species-specific primers.

Funding: The study is supported the Russian Science Foundation project No. 23-16-00262 and State Assignment project No. FGEU-2022-0019 (state registration No. 122032900144-2).

Список литературы/References

1. Becnel J.J., Andreadis T.G. Microsporidia in Insects. In: Microsporidia: Pathogens of Opportunity. 2014;521-570. doi 10.1002/9781118395264.ch21
2. Han B., Weiss L.M. Microsporidia: obligate intracellular pathogens within the fungal kingdom. *Microbiol Spectr.* 2017;5(2). doi 10.1128/microbiolspec.funk-0018-2016
3. Исси И.В. Развитие микроспоридиологии в России. *Вестник защиты растений.* 2020;103(3):161-176
4. Токарев Ю.С. Значение микроспоридий в борьбе с вредными членистоногими. *Защита и карантин растений.* 2007;(12):14-16
5. Wittner M. Historic Perspective on the Microsporidia: Expanding Horizons. In: The microsporidia and microsporidiosis. Washington: ASM Press, 1999;1-6. doi 10.1128/9781555818227.ch1
6. Bojko J. Microsporidia: a new taxonomic, evolutionary, and ecological synthesis. *Trends Parasitol.* 2022;38(8):642-659
7. Malysh S.M. et al. Susceptibility of Bombyx mori larvae to the microsporidium *Nosema bombycis* from the silkworm and *Nosema* sp. from the cotton bollworm. *Plant Protect News.* 2023;4:210-214. doi 10.31993/2308-6459-2023-106-4-16148
8. Sambrook J., Fritsch E.R., Maniatis T. Molecular Cloning: A Laboratory Manual. NY: Cold Spring Harbor Laboratory Press, 1989
9. Tokarev Y.S. et al. Ecological vs physiological host specificity: the case of the microsporidium *Nosema pyrausta* (Paillot) Weiser, 1961. *Acta Biol Sibirica.* 2022;8:297-316. doi 10.14258/abs.v8.e19
10. Pan G. et al. Comparative genomics of parasitic silkworm microsporidia reveal an association between genome expansion and host adaptation. *BMC Genomics.* 2013;14:186

11. Tokarev Y.S. et al. Hexokinase as a versatile molecular genetic marker for Microsporidia. *Parasitology*. 2019;146(4):472-478. doi 10.1017/S0031182018001737
12. Weiser J. Die Mikrosporidien als Parasiten der Insekten. Beihefte zur Angewandten Entomologie. Hamburg: P.P. Parey, 1961
13. Dolgikh V.V. Construction and heterologous overexpression of two chimeric proteins carrying outer hydrophilic loops of *Vairimorpha ceranae* and *Nosema bombycis* ATP/ADP carriers. *J Invertebr Pathol*. 2020;171:107337
14. Malysh J.M., Chertkova E.A., Tokarev Y.S. The microsporidium *Nosema pyrausta* as a potent microbial control agent of the beet webworm *Loxostege sticticalis*. *J Invertebr Pathol*. 2021;186:107675. doi 10.1016/j.jip.2021.107675