

МОЛЕКУЛЯРНАЯ БИОЛОГИЯ

УДК 577.21

СЕКВЕНИРОВАНИЕ ГЕНОМА И ТРАНСКРИПТОМА МИНИАТЮРНОГО ПЕРЕПОНЧАТОКРЫЛОГО *MEGAPHRAGMA AMALPHITANUM* (HYMENOPTERA: TRICHOGRAMMATIDAE)Е.Б. Прохорчук^{1,3}, А.В. Недолужко^{2,*}, Ф.С. Шарко³, С.В. Цыганкова², Е.С. Булыгина², С.М. Расторгуев², А.С. Соколов³, А.М. Мазур³, А.А. Полилов⁴, К.Г. Скрябин^{1,2,3}¹ Кафедра биотехнологии, биологический факультет, Московский государственный университет имени М.В. Ломоносова; Россия, 119234, г. Москва, Ленинские горы, д. 1, стр. 12;² Национальный исследовательский центр “Курчатовский институт”; Россия, 123182, г. Москва, пл. Академика Курчатова, д. 1;³ Институт биоинженерии, Федеральный исследовательский центр “Фундаментальные основы биотехнологии” Российской академии наук; Россия, 119071, г. Москва, Ленинский пр-т, д. 33, корп. 2;⁴ Кафедра энтомологии, биологический факультет, Московский государственный университет имени М.В. Ломоносова; Россия, 119234, г. Москва, Ленинские горы, д. 1, стр. 12
*e-mail: nedoluzhko@gmail.com

Миниатюризация организмов — одно из наиболее интересных эволюционных явлений. Это весьма распространенный процесс, широко представленный у многоклеточных организмов: беспозвоночных (немертины, круглые черви, брахиоподы, моллюски, паукообразные и насекомые) и позвоночных (рыбы, амфибии, птицы и даже млекопитающие). Миниатюризация приводит к значительному сокращению размеров организма и упрощению ряда систем и органов без потери жизнеспособности. Как правило, эта комплексная эволюционная модификация связана с занятием особой экологической ниши, свободной от конкурентов. В ходе этого процесса не только изменяется общий размер организма, но и уменьшаются размеры клеток, составляющих его, а также меняется структура генома и отдельных генов. Паразитический наездник *Megaphragma amalphitanum* (Hymenoptera: Trichogrammatidae) считается одним из мельчайших насекомых, описанных к настоящему времени, и потому представляет значительный интерес в качестве объекта для изучения генетических аспектов уменьшения размеров организмов в процессе эволюции. Секвенирование полного генома *M. amalphitanum* и его последующая аннотация, а также сборка транскриптома позволяют понять, насколько меняется генетический материал таких организмов при миниатюризации по сравнению с другими (более крупными) представителями отряда Hymenoptera. Предварительный анализ показал, что размер генома *M. amalphitanum* составляет 346 миллионов пар оснований, а размер транскриптома — 27,5 миллионов пар оснований.

Ключевые слова: Hymenoptera, геном, транскриптом, *Megaphragma amalphitanum*, паразитический наездник, миниатюризация.

Понимание процесса миниатюризации многоклеточных организмов имеет колоссальное фундаментальное значение. Для позвоночных показано, что миниатюризация приводит к изменениям в строении тела, у беспозвоночных изменения, связанные с уменьшением размеров тела, также затрагивают многие органы и системы органов [1–3].

Процесс экстремального уменьшения размеров приводит к постепенной миниатюризации организма посредством последовательной редукции предкового фенотипа и появления новых, модифицированных, органов и систем. В ходе этого процесса не только изменяется общий размер организма, но и сокращаются размеры клеток, составляющих его.

Миниатюризация иногда может приводить к изменению структуры и размеров генома и, как следствие, изменению экспрессии определенных генов, связанных с ростом и развитием организма, хотя, к сожалению, к настоящему времени подобных исследований проведено мало [4–6].

В данной работе представлены результаты геномного и транскриптомного секвенирования одного из мельчайших насекомых, описанных к настоящему времени, — паразитического наездника *Megaphragma amalphitanum* (Hymenoptera: Trichogrammatidae). Показано, что размер генома и транскриптома этого насекомого сопоставим с этими показателями для других, более крупных, представителей Hymenoptera.

Материалы и методы

Выделение ДНК и РНК, приготовление библиотек и секвенирование. В работе использовали взрослых особей *M. amalphitanum*, которые были выведены из яиц трипсов *Heliothrips haemorrhoidalis* (Thysanoptera: Thripidae), собранных в окрестностях города Санта-Маргарита (Северная Италия). Выделение ДНК и создание ДНК-библиотек проводили по методике, описанной ранее [7]; тотальную РНК выделяли с использованием реагента Тризол (Thermo

Fisher Scientific, США) по стандартному протоколу, кДНК-библиотеки конструировали, используя коммерческий набор Ovation RNA-Seq System V2 kit (NuGEN, США).

Созданные ДНК-библиотеки секвенировали на геномном анализаторе Illumina HiSeq 1500 (Illumina, США).

Биоинформатический анализ данных. В качестве исходного материала для сборки генома нами были использованы данные, полученные при секвенировании (исходный формат *.fastq) ДНК-библиотек, приготовленных из геномной ДНК. Для сборки полной последовательности генома паразитического наездника были использованы 102188833 парных чтений длиной 150 нуклеотидов. Для реконструкции генома *M. amalphantum* проводилась сборка *de novo* с использованием нескольких общедоступных программ (ABYSS, Velvet, CLC, SOAP и SPAdes). На первом этапе сборки мы объединили исходные чтения с помощью программы Pear [8] и удалили ПЦР-дубликаты.

Программный пакет SPAdes показал лучшие результаты сборки генома, поэтому в дальнейшем мы использовали контиги, полученные с его помощью. Сборку скаффолдов проводили с использованием утилиты SOAPdenovo Data Preparation Module [9].

В качестве исходного материала для сборки транскриптома нами были использованы данные, полученные при секвенировании (исходный формат *.fastq) кДНК-библиотек, приготовленных из тотальной РНК. Для сборки полной последовательности генома паразитического наездника были использованы 59 790 973 парных чтений длиной 125 нуклеотидов. Для сборки *de novo* транскриптома *M. amalphantum* использовали программный пакет Trinity со стандартными параметрами [10]. Аннотацию генома проводили с помощью программы AUGUSTUS [11].

Результаты и обсуждение

В результате проведенного биоинформатического анализа была получена геномная сборка паразитического наездника *M. amalphantum* (PRJNA344956) — её суммарный размер составляет 346 миллионов пар оснований (п.о.) при N50 контига, равном 4275 п.о.

Комбинирование двух *de novo*-сборщиков SOAP and SPAdes позволило собрать 94 687 скаффолдов (N50 = 10 296 п.о.). Программный пакет SPAdes продемонстрировал лучшие результаты сборки (табл. 1).

Используя полученные чтения кДНК-библиотеки, полученных из тотальной ДНК и отсекаемых, собрали 46 841 контигов с N50, равным 633 п.о. Суммарный размер транскриптома составил 27 407 647 миллионов п.о. Полученные геномные и транскриптомные данные были использованы при аннотации генома, которая позволила выявить ряд белок-кодирующих и некодирующих последовательностей (доступны по ссылке: www.treefrog.ru/megaphragma.gff). Эти последовательности в дальнейшем будут использованы при сравнении полученных результатов с данными, касающимися более крупных представителей Hymenoptera. Это позволит выявить гены/генные кластеры/метаболические пути, вероятно, связанные с миниатюризацией *M. amalphantum*.

К настоящему времени не существует единого мнения насчет корреляции между размером тела и размером генома. В частности, в исследованиях плоских червей и веслоногих рачков (копепод) продемонстрировали наличие достоверной связи между этими двумя признаками [4], в то время как в работах, посвященных размерам генома у муравьев, не просматривается подобной корреляции [12]. В случае с *M. amalphantum* (предварительный размер генома — 346 миллионов п.о.) мы также не просматриваем корреляции между размером организма и его геномом — согласно данным, депонированным в NCBI (Национальный центр биотехнологической информации США), размер генома более крупного паразитического наездника *Nasonia vitripennis* составляет 295 миллионов п.о., а наиболее качественно собранного генома медоносной пчелы (*Apis mellifera*) — 247 миллионов п.о.

Анализ транскриптома *M. amalphantum*, который был использован нами для аннотации, показал, что он также имеет стандартный размер для перепончатокрылых — 27,5 миллионов п.о.

Работа выполнена при финансовой поддержке Российского научного фонда (проект № 14-24-00175).

Таблица 1

Статистика сборки *de novo* генома паразитического наездника *M. amalphantum* с использованием ассемблеров ABYSS, Velvet, CLC и SPAdes, SOAP

Количество собранных контигов	Минимальный размер контига	N50	Максимальный размер контига	Суммарный размер сборки	Использованный ассемблер
541 950	300	869	19 313	2,91×10 ⁸	Velvet
484 255	300	1425	33 321	3,07×10 ⁸	CLC
553 157	300	4285	56 202	3,58×10 ⁸	SPAdes
4,58E+06	300	974	21 336	2,20×10 ⁸	ABYSS
94 687	300	10 296	895 906	3,46×10 ⁸	SPAdes+SOAP

СПИСОК ЛИТЕРАТУРЫ

1. Grebennikov V.V., Beutel R.G. Morphology of the minute larva of *Ptinella tenella*, with special reference to effects of miniaturisation and the systematic position of Ptiliidae (Coleoptera: Staphylinioidea) // *Arthropod Struct. Dev.* 2002. Vol. 31. N 2. P. 157–172.
2. Polilov A.A. The smallest insects evolve anucleate neurons // *Arthropod Struct. Dev.* 2012. Vol. 41. N 1 P. 29–34.
3. Polilov A.A. Consequences of miniaturization in insect morphology // *Moscow Univ. Biol. Sci. Bull.* 2015. Vol. 70. N 3. P. 37–43.
4. Gregory T.R., Hebert P.D., Kolasa J. Evolutionary implications of the relationship between genome size and body size in flatworms and copepods // *Heredity (Edinb.)*. 2000. Vol. 84. Pt. 2. P. 201–208.
5. Liu S., Hui T.H., Tan S.L., Hong Y. Chromosome evolution and genome miniaturization in minifish // *PLoS One*. 2012. Vol. 7. N 5. e37305.
6. Harris R.A., Tardif S.D., Vinar T., Wildman D.E., Rutherford J.N., Rogers J., Worley K.C., Aagaard K.M. Evolutionary genetics and implications of small size and twinning in callitrichine primates // *Proc. Natl. Acad. Sci. U.S.A.* 2014. Vol. 111. N 4. P. 1467–1472.
7. Nedoluzhko A.V., Sharko F.S., Boulygina E.S., Tsygankova S.V., Sokolov A.S., Mazur A.M., Polilov A.A., Prokhortchouk E.B., Skryabin K.G. Mitochondrial genome of *Megaphragma amalphitanum* (Hymenoptera: Trichogrammatidae) // *Mitochondrial DNA Part A*. 2016. Vol. 27. N 6. P. 4526–4527.
8. Zhang J., Kobert K., Flouri T., Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR // *Bioinformatics*. 2014. Vol. 30. N 5. P. 614–620.
9. Luo R., Liu B., Xie Y., et al. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler // *Gigascience*. 2012. Vol. 1. N 1. P. 18.
10. Haas B.J., Papanicolaou A., Yassour M., et al. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis // *Nat. Protoc.* 2013. Vol. 8. N 8. P. 1494–1512.
11. Stanke M., Steinkamp R., Waack S., Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes // *Nucleic Acids Res.* 2004. Vol. 32 (Web Server issue). W309–W312.
12. Tsutsui N.D., Suarez A.V., Spagna J.C., Johnston J.S. The evolution of genome size in ants // *BMC Evol. Biol.* 2008. Vol. 8 : 64.

Поступила в редакцию
09.11.2016
Принята в печать
06.12.2016

MOLECULAR BIOLOGY

THE GENOME SEQUENCING AND TRANSCRIPTOMIC ASSEMBLY OF THE PARASITOID WASP *MEGAPHRAGMA AMALPHITANUM* (HYMENOPTERA: TRICHOGRAMMATIDAE)

E.B. Prokhortchouk^{1,3}, A.V. Nedoluzhko^{2,*}, F.S. Sharko³, S.V. Tsygankova², E.S. Boulygina²,
S.M. Rastorguev², A.S. Sokolov³, A.M. Mazur³, A.A. Polilov⁴, K.G. Skryabin^{1,2,3}

¹ Department of Biotechnology, Faculty of Biology, Lomonosov Moscow State University,
Leninskiye gory 1-12, Moscow, 119234, Russia;

² National Research Centre “Kurchatov Institute”, Kurchatov sq. 1, Moscow, 123182, Russia;

³ Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of Sciences,
Leninsky prospect 33-2, Moscow, 119071, Russia

⁴ Department of Entomology, Faculty of Biology, Lomonosov Moscow State University,
Leninskiye gory 1-12, Moscow, 119234, Russia

*e-mail: nedoluzhko@gmail.com

Miniaturization of organisms is one of the most interesting evolutionary phenomena. This is a very common process, widely represented in multicellular organisms: invertebrates (nemertean, nematodes, brachiopods, mollusca, arachnids, and insects) and vertebrates (fishes, amphibians, birds, and even mammals). Miniaturization leads to a significant reduction in the body size and simplifying of some systems and organs without loss of species viability. The cell size reduce as well as genome and gene structure change during miniaturization. Typically, this complex evolutionary modification is associated with the occupation of ecological niches. Parasitoid wasp *Megaphragma amalphitanum* (Hymenoptera: Trichogrammatidae) is considered one of the smallest insects described to date and therefore is attractive as an object for studying of genetic aspects of body size reduction during evolution. The whole genome sequencing of *M. amalphitanum* and subsequent genome annotation and transcriptome assembly allow us to understand how the genetic material changes during miniaturization in relation to the corresponding characteristics of the larger Hymenoptera species. Preliminary analysis showed that the genome size of the *M. amalphitanum* is 346 Mbp; the transcriptome size is 27.5 Mbp.

Keywords: *Hymenoptera*, *genome*, *transcriptome*, *Megaphragma amalphitanum*, *parasitoid wasp*, *miniaturization*.

Сведения об авторах

Прохорчук Егор Борисович — докт. биол. наук, проф. кафедры биотехнологии биологического факультета МГУ, заведующий лабораторией Института биоинженерии, ФИЦ “Фундаментальные основы биотехнологии” РАН. Тел.: 8-499-135-53-37; e-mail: prokhortchouk@gmail.com

Недолужко Артём Валерьевич — канд. биол. наук, начальник лаборатории НИЦ “Курчатовский институт”. Тел.: 8-499-196-79-37; e-mail: nedoluzhko@gmail.com

Шарко Фёдор Сергеевич — аспирант, мл. науч. сотр. Института биоинженерии ФИЦ “Фундаментальные основы биотехнологии” РАН. Тел.: 8-499-135-53-37; e-mail: fedosic@gmail.com

Цыганкова Светлана Валерьевна — канд. биол. наук, ст. науч. сотр. НИЦ “Курчатовский институт”. Тел.: 8-499-196-79-37; e-mail: svetlana.tsygankova@gmail.com

Булыгина Евгения Станиславовна — канд. биол. наук, вед. науч. сотр. НИЦ “Курчатовский институт”. Тел.: 8-499-196-79-37; e-mail: eugenia.bulygina@gmail.com

Расторгуев Сергей Михайлович — канд. биол. наук, ст. науч. сотр. НИЦ “Курчатовский институт”. Тел.: 8-499-196-79-37; e-mail: rastorgueff@gmail.com

Соколов Алексей Сергеевич — мл. науч. сотр. Института биоинженерии, ФИЦ “Фундаментальные основы биотехнологии” РАН. Тел.: 8-499-135-53-37; e-mail: sokolovbiotech@gmail.com

Мазур Александр Михайлович — канд. физ.-мат. наук, ст. науч. сотр. Института биоинженерии, ФИЦ “Фундаментальные основы биотехнологии” РАН. Тел.: 8-499-135-53-37; e-mail: mazur.am@gmail.com

Полилов Алексей Алексеевич — докт. биол. наук, зав. кафедрой биологического факультета МГУ. Тел.: 8-495-939-35-32; e-mail: polilov@gmail.com

Скрябин Константин Георгиевич — докт. биол. наук, проф., академик РАН, заведующий кафедрой биотехнологии биологического факультета МГУ, научный руководитель ФИЦ “Фундаментальные основы биотехнологии” РАН. Тел.: 8-499-135-73-19; e-mail: office@biengi.ac.ru