

УДК 571.27

<https://doi.org/10.31016/978-5-6053355-1-1.2025.26.301-306>

## СБОРКА *DE NOVO* И ОЦЕНКА КАЧЕСТВА СБОРКИ ТРАНСКРИПТОМА *DIBOTHRIOCEPHALUS DENDRITICUS*

Содномов Т. Ч. <sup>1</sup>,

аспирант кафедры биотехнологии,  
sodnomov.tamir@bk.ru

Кутырев И. А. <sup>1,2,3</sup>,

доктор биологических наук,  
доцент, ведущий научный сотрудник,  
ikutyrev@yandex.ru

Мазур О. Е. <sup>2</sup>,

кандидат биологических наук, научный сотрудник

Дугаров Ж. Н. <sup>2</sup>,

кандидат биологических наук, заведующий лабораторией  
паразитологии и экологии гидробионтов

### Аннотация

Цестоды, как паразитические черви, обладают уникальными механизмами взаимодействия с хозяином, что порождает интерес к их веществам как потенциальным источникам новых лекарственных средств, необходимых для лечения различных аутоиммунных заболеваний, аллергических заболеваний и воспалений. Нами впервые произведена *de novo* сборка транскриптома лентеца чаечного *D. dendriticus* на личиночной стадии развития. Оценка качества прочтений, полученных при секвенировании мРНК *Dibothriocephalus dendriticus*, проводили с помощью программы FastQC. Фильтрацию некачественных ридов проводили с помощью программы Trimmomatic. Сборку *de novo* транскриптома проводили программой Trinity. Оценка качества сборки проводили при помощи программы BUSCO. На первом этапе мы провели оценку качества сырых ридов трех образцов *D. dendriticus* и фильтрацию некачественных

<sup>1</sup> Федеральное государственное бюджетное образовательное учреждение высшего образования «Восточно-Сибирский государственный университет технологий и управления» (670013, Россия, г. Улан-Удэ, ул. Ключевская, д. 40В, стр. 1)

<sup>2</sup> Федеральное государственное бюджетное учреждение науки Институт общей и экспериментальной биологии Сибирского отделения Российской академии наук (670047, Россия, г. Улан-Удэ, ул. Сахьяновой, д. 6)

<sup>3</sup> Федеральное государственное бюджетное учреждение науки Институт проблем экологии и эволюции им А. Н. Северцова Российской академии наук (119071, Россия, г. Москва, Ленинский пр-т, д. 33)

прочтений. Далее подготовленные библиотеки были использованы для *de novo* сборки транскриптома. После сборки была проведена оценка качества сборки транскриптома. Показатель полностью собранных консервативных ортологов оказался достаточно большим (92,5%). Количество фрагментированных и несобранных ортологов невелико – 2,7 и 4,7 % соответственно. Это свидетельствует о хорошем качестве сборки, так как большинство транскриптов, продуцируемых в организме, было собрано. Однако большой показатель дублицированных последовательностей говорит об избыточности сборки транскриптов. В дальнейшем дублицированные транскрипты должны быть удалены соответствующими биоинформатическими методами.

**Ключевые слова:** *Dibothriocephalus dendriticus*, цестоды, мРНК, сборка *de novo*, транскриптом

## **DE NOVO DIBOTHRIOCEPHALUS DENDRITICUS TRANSCRIPTOME ASSEMBLY AND QUALITY ASSESSMENT OF THE ASSEMBLY**

**Sodnomov T. C.** <sup>1</sup>,

Postgraduate Student of the Department of Biotechnology,  
sodnomov.tamir@bk.ru

**Kutyrev I. A.** <sup>1, 2, 3</sup>,

Doctor of Biological Sciences,  
Associate Professor, Leading Researcher,  
ikutyrev@yandex.ru

**Mazur O. E.** <sup>2</sup>,

Candidate of Biological Sciences, Researcher

**Dugarov Z. N.** <sup>2</sup>,

Candidate of Biological Sciences, Head of the Laboratory  
of Parasitology and Ecology of Hydrobionts

### **Abstract**

Cestodes as parasitic worms have unique host interaction mechanisms, which generates interest in their substances as potential sources of new drugs needed to

---

<sup>1</sup> Federal State Budgetary Educational Institution of Higher Education "East Siberian State University of Technology and Management" (40B, Klyuchevskaya st., Bldg. 1, Ulan-Ude, 670013, Russia)

<sup>2</sup> The Institute of General and Experimental Biology of the Siberian Branch of the Russian Academy of Sciences (6, Sakhyanovoy st., Ulan-Ude, 670047, Russia)

<sup>3</sup> Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences (33, Leninsky Prospekt, Moscow, 119071, Russia)

treat various autoimmune diseases, allergic diseases, and inflammations. For the first time, we performed *de novo* assembly of the transcriptome of the gull-tapeworm *D. dendriticus* at the larval stage of development. The quality of reads obtained in sequencing a *Dibothriocephalus dendriticus* mRNA was assessed using FastQC. Low-quality reads were filtered using Trimmomatic. *De novo* transcriptome assembly was performed using Trinity. The assembly quality was assessed using BUSCO. At the first stage, we assessed quality of raw reads of three *D. dendriticus* samples and filtered out low-quality reads. Then, the prepared libraries were used for the *de novo* transcriptome assembly. After the assembly, quality of the transcriptome assembly was assessed. The rate of fully assembled conservative orthologs was quite high (92.5%). The number of fragmented and unassembled orthologs was low (2.7 and 4.7%, respectively). This indicates adequate quality of the assembly since most of the transcripts produced in the organism were assembled. However, the high rate of duplicated sequences indicates the redundancy of the transcript assembly. In the future, duplicated transcripts should be removed by appropriate bioinformatic methods.

**Keywords:** *Dibothriocephalus dendriticus*, cestodes, mRNA, *de novo* assembly, transcriptome

**Введение.** В ряде северных районов Сибири и в Прибайкалье основным возбудителем дифиллоботриоза человека и животных является лентец чаечный *Dibothriocephalus dendriticus* (син. *Diphyllobothrium dendriticum*). Цестоды представляют собой перспективный ресурс для разработки новых лекарственных средств, учитывая их уникальные биохимические свойства и взаимодействия с иммунной системой хозяев [4, 5]. Биологически активные вещества, выделяемые секретами и метаболитами цестод, представляют собой актуальную и многообещающую тему для научных исследований [3]. Иммунорегуляторные белки, выявленные в секрете *Dibothriocephalus dendriticus*, могут быть использованы в качестве мишеней для разработки новых лекарственных средств, направленных на лечение аутоиммунных заболеваний или для регуляции иммунного ответа в других патологиях. В работе рассматриваются методы сборки транскриптома цестод, необходимого для дальнейшей идентификации белков в их секреторно-экскреторных продуктах.

Целью настоящего исследования явилась сборка *de novo* и оценка качества сборки транскриптома *Dibothriocephalus dendriticus*.

Для достижения цели необходимо было выполнить следующие задачи:

- 1) Оценить качество фильтрации ридов *D. dendriticus* с помощью биоинформатических программ.

- 2) Собрать и оценить качество сборки транскриптома плероцеркоидов *D. dendriticus*.

**Материалы и методы.** Общую РНК из 0,5–1,0 г ткани выделяли с помощью реагента TRIzol (Ambion) и очищали с одновременной обработкой ДНКазой I на колонках PureLink RNA Mini (Invitrogen). Качество РНК определяли на биоанализаторе BA2100 с использованием набора RNA Nano Kit. Взяли прочтения мРНК трех образцов лентца *Dibothriocephalus dendriticus*. Каждый образец был отсековен в двух направлениях: прямом (forward) – R1 и обратном (reverse) – R2. Оценку качества прочтений проводили при помощи пакета программ FastQC [1]. Фильтрацию некачественных прочтений осуществляли в программе Trimmomatic [2]. Сборку транскриптома плероцеркоидов *D. dendriticus* производили с помощью пакета программ Trinity. В качестве входных файлов сначала собирали прямые фильтрованные прочтения, а затем, также в строгой последовательности, все обратные фильтрованные прочтения всех трех образцов. Оценку качества сборки транскриптома плероцеркоидов *D. dendriticus* проводили с помощью пакета программ BUSCO.

**Результаты исследований.** Нами впервые произведена *de novo* сборка транскриптома лентца чаечного *D. dendriticus* на личиночной стадии развития.

На первом этапе мы провели оценку качества сырых ридов трех образцов *D. dendriticus* и провели фильтрацию некачественных прочтений. В таблице представлены графические изображения оценки качества секвенирования первого образца лентца *Dibothriocephalus dendriticus* в прямом направлении (форвард, R1) до фильтрации (левый столбец) и после фильтрации (правый столбец).

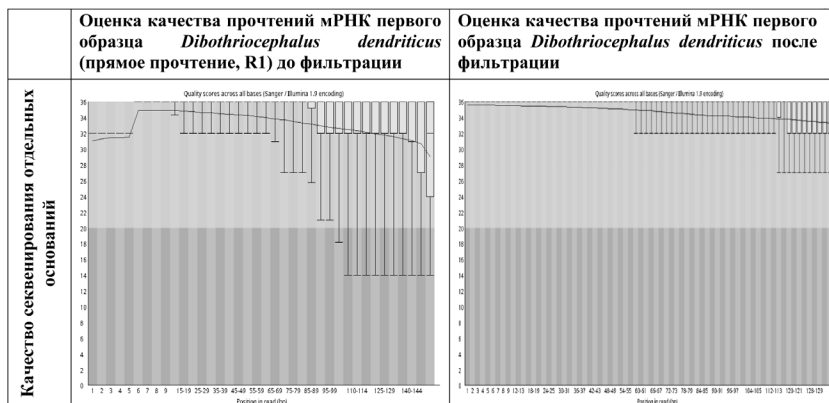
Далее подготовленные библиотеки были использованы для *de novo* сборки транскриптома с помощью пакета программ Trinity.

Общее количество собранных генов Trinity составляет 58 776, общее количество собранных транскриптов — 98 005, процент содержания GC — 47,19%.

Для оценки качества сборки полученного транскриптома использовали пакет программ BUSCO. Использовали базу данных однокопийных ортологов: eukaryota\_odb10 (Дата создания: 08.01.2024, количество генов: 70, количество BUSCO: 255). Для интерпретации результатов использовали следующие обозначения. Последовательность может присутствовать и быть собранной полностью (Complete, C) или частично

Таблица

**Оценка качества прочтений мРНК плероцеркоидов  
*Dibothriocephalus dendriticus***



(Fragmented, F). Может варьироваться число схожих транскриптов/изоформ — Single-copy, S (если в одной копии, что мы ожидаем в геноме) или Duplicated, D (последовательность встречается более 1 раза, что является нормой для транскриптовом в виду избыточности сборки). Если та или иная последовательность отсутствует в анализируемой сборке, что она указывается, как Missing, M.

После оценки качества сборки транскриптома плероцеркоидов *D. dendriticus* получили следующие результаты. Относительные показатели: C: 92.5% [S: 36.1%, D: 56.5%], F: 2.7%, M: 4.7%, n: 255. Абсолютные показатели: полностью собранные BUSCOs (C) — 236; собранные полностью и схожие транскрипты BUSCOs (S) — 92; собранные полностью и повторяющиеся последовательности BUSCOs (D) — 144; частично собранные последовательности BUSCOs (F) — 7; отсутствующие последовательности BUSCOs (M) — 12. Всего групп BUSCO найдено (n) — 255.

**Заключение.** Нами впервые произведена *de novo* сборка транскриптома лентца чаечного *D. dendriticus* на личиночной стадии развития. На первом этапе мы провели оценку качества сырых ридов трех образцов *D. dendriticus* и фильтрацию некачественных прочтений. Далее подготовленные библиотеки были использованы для *de novo* сборки транскриптома. После сборки была проведена оценка качества сбор-

ки транскриптома. Показатель полностью собранных консервативных ортологов оказался достаточно большим (92,5%) Количество фрагментированных и несобранных ортологов невелико – 2,7 и 4,7 % соответственно. Это свидетельствует о хорошем качестве сборки: большинство продуцируемых в организме транскриптов было собрано. Однако большой показатель дублицированных последовательностей говорит об избыточности сборки транскриптов. В дальнейшем дублицированные транскрипты должны быть удалены соответствующими биоинформатическими методами.

*Работа выполнена при финансовой поддержке гранта Министерства образования и науки Республики Бурятия № 502 от 22.11.2024 г. и в рамках темы госзадания «Биологическое разнообразие паразитов гидробионтов бассейна озера Байкал и Забайкалья» (FWSM-2021-0002).*

#### Список источников / References

1. Bhanu D., Ulaganathan K., Shanker A. K. Water Stress Responsive Differential Methylation of Organellar Genomes of *Zea mays* Z59. *American Journal of Plant Sciences*. 2020; 11(7): 1077–1100.
2. Bolger A. M., Lohse M., Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30(15): 2114–20.
3. Zhang T., Jiang J., Liu J. Effects of Parasitic Diseases on the Cardiovascular System. *Journal of Biosciences and Medicines*. 2022; 10(06): 90–102.
4. Do T., Nguyen B. X., Pham V., Tran T., Tjiputra E., Tran Q., Nguyen A. Reducing Training Time in Cross-Silo Federated Learning using Multigraph Topology. *International Conference on Computer Vision (ICCV)*. 2023: 19352–19362.
5. Li Z., Wen W., Qin M., He Y., Xu D., Li L. Biosynthetic Mechanisms of Secondary Metabolites Promoted by the Interaction Between Endophytes and Plant Hosts. *Frontiers in Microbiology*. 2022; 13: 928967.