

OCCUPANCY INDEX OF LIME LEAF MINER ON *TILIA CORDATA* ON
THE SECOND GENERATION IN THE CONDITIONS OF MINSK

Yu.O. Shibanova¹, A.A. Khackevich², M.A. Bobrova³

¹ *Belarusian state pedagogical university named after Maxim Tank,
Minsk, Belarus*

² *State Educational Establishment "High school №110", Minsk, Belarus*

³ *Branch of the Belarusian state economic university "Minsk Trade College",
Minsk, Belarus*

aleh.sinchuk@gmail.com; al.hatzkevich@yandex.ru; bobrova.082000@mail.ru

Occupancy index of lime leaf miner (*Phyllonorycter issikii*) on *Tilia cordata* at the end of the second generation varies from 2 % to 46 % which may be due to the density of plantations.

**ОПТИМИЗАЦИЯ ПАРАМЕТРА ДЛИНЫ К-
МЕРЫ ПРИ СБОРКЕ ПРОЧТЕНИЙ ГЕНОМА *APHIS FABAE*
MORDVILKOI BÖRNER & JANISCH, 1922
С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМА ABYSS2.0**

Р.С. Шулинский, П.Ю. Кветко, Ю.В. Бондаренко

*Белорусский государственный университет, Минск, Беларусь
shulinsky@mail.ru*

Сборка геномов *denovo* является сложной задачей, особенно при работе с крупными геномами, типичными для эукариот. Проблема сборки геномов *denovo* связана с реконструкцией хромосомной последовательности из коротких перекрывающихся прочтений, которые на несколько порядков короче, чем геном изучаемого организма [1]. На практике современное программное обеспечение для сборки геномов *denovo* не полностью восстанавливает хромосомную последовательность, а скорее создает более протяженные смежные консенсусные последовательности – контиги.

Одним из важнейших этапов при сборке геномов *denovo* является оптимизация длины k-меры (искусственно заданной длины, представляющей собой фрагмент анализируемого прочтения), что влияет на ключевые показатели качества сборки контиг. Следовательно, оптимизация параметров для сборки контиг, а, в частности, длины k-меры, является обязательным этапом в сборке генома *denovo* [2].

Сборка контиг производилась из данных полногеномного секвенирования *Aphisfabaemordvilkoic* помощью программы ABySS2.0. При выполнении данной работы были использованы такие значения длины k-меры, как 76, 80 и 86.

Нами были получены следующие результаты сборки контиг при различном значении k-меры: при значении длины k-меры 76 было получено 3 □ 174 □ 569 контиг со значением N50 равным 760 и суммарной длиной контиг $126,6 \times 10^6$; при длине k-меры равной 80 было получено 2 □ 756 □ 558 контиг со значением N50 равным 762 и суммарной длиной контиг $129,4 \times 10^6$; и при значении длины k-меры 86 было получено 2 □ 237 □ 855 контиг со значением N50 равным 753 и суммарной длиной контиг $127,6 \times 10^6$.

Пиковый объем оперативной памяти при сборке контиг со значениями длины k-меры 76, 80 и 86 был равен 116 Гб, 128 Гб и 138 Гб соответственно. Как можно видеть из выше представленных данных, оптимальной длиной k-меры для прочтений генома *A. fabaemordvilkoia* является 80, т.к. значение N50 и показатель суммарной длины контиг максимальны.

Таким образом, было оптимизировано значение k-меры и установлен минимальный объем оперативной памяти, необходимой для сборки контиг для данных полногеномного секвенирования *A. fabaemordvilkoia*.

1. ABySS: A parallel assembler for short read sequence data / J.T. Simpson [et al.] // Genome Research. – 2009. – Vol. 19, No □ 6. – P. □ 1117–1123.

2. ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter / S.D. Jackman [et al.] // Genome Research. – 2017. – Vol. 27, No 5. – P. □ 768.– 777.

OPTIMIZATION OF K-MER LENGTH PARAMETER FOR ASSEMBLY
READING OF THE *APHIS FABAE MORDVILKOI* BÖRNER & JANISCH,
1922 USING ABYSS2.0 ALGORITHM

R.S. Shulinsky, P.Y. Kviatko, Y.V. Bandarenka

Belarusian State University, Minsk, Belarus

shulinsky@mail.ru

Aptitude of ABySS2.0 *de novo* assembly algorithm to genome reads of *Aphis fabae mordvilkoia* was evaluated. Single-end genome reads were produced with IonTorrent technology. We evaluated genome assembly statistics produced with different values of k–mer length as well as memory usage statistics.

We conclude the optimal length of k-mer for *A. fabae* ssp reads data being 80 with memory usage of 128 Gb. Results can be applied in insect genomics and molecular genetics.
