

DOI: 10.25205/978-5-4437-1691-6-38

**ИДЕНТИФИКАЦИЯ СЕГМЕНТОВ ИНТРОГРЕССИВНОЙ ГИБРИДИЗАЦИИ  
В ГЕНОМЕ ПШЕНИЦЫ С ИСПОЛЬЗОВАНИЕМ GBS\*****IDENTIFICATION OF INTROGRESSIVE HYBRIDIZATION SEGMENTS  
IN THE WHEAT GENOME USING GBS**Н. М. Съедина<sup>1</sup>, А. С. Ермолаев<sup>1</sup>, Д. С. Ульянов<sup>1</sup>, В. С. Воронежская<sup>1,2</sup>,  
А. В. Васильев<sup>1</sup>, С. В. Тощачков<sup>3</sup>, М. Г. Дивашук<sup>1</sup><sup>1</sup>Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии, Москва<sup>2</sup>Всероссийский научно-исследовательский институт радиологии и агроэкологии,  
НИЦ «Курчатовский институт», Обнинск<sup>3</sup>Национальный исследовательский центр «Курчатовский институт», МоскваN. M. Syedina<sup>1</sup>, A. S. Ermolaev<sup>1</sup>, D. S. Ulyanov<sup>1</sup>, V. S. Voronezhskaya<sup>1,2</sup>,  
A. V. Vasiliev<sup>1</sup>, S. V. Toshchakov<sup>3</sup>, M. G. Divashuk<sup>1</sup><sup>1</sup>All-Russia Research Institute of Agricultural Biotechnology, Moscow<sup>2</sup>All-Russia Research Institute of Radiology and Agroecology, National Research Centre “Kurchatov Institute”<sup>3</sup>National Research Centre “Kurchatov Institute”, Moscow

✉ nadi.sn@yandex.ru

**Аннотация**

В работе представлен подход для количественной оценки содержания геномного материала родителей в геноме потомков на основании данных секвенирования GBS. Кроме того, метод позволяет оценить распределение материала родителей в геноме потомков и идентифицировать точки рекомбинации и хромосомные перестройки. Количественная оценка позволяет сделать вывод о вкладе каждого родителя в геном потомка, что важно для селекционного процесса при бэккроссировании и интрогрессивной селекции для минимизации количества генетического материала одного из родителя в потомках.

**Abstract**

This study presents an approach for quantitative assessment of the content of parental genomic material in the offspring genome based on GBS sequencing data. In addition, the method allows for assessing the distribution of parental material in the offspring genome and identifying recombination points and chromosomal rearrangements. Quantitative assessment allows drawing a conclusion about the contribution of each parent to the offspring genome, which is important for the breeding process in backcrossing and introgression selection to minimize the amount of genetic material of one parent in the offspring.

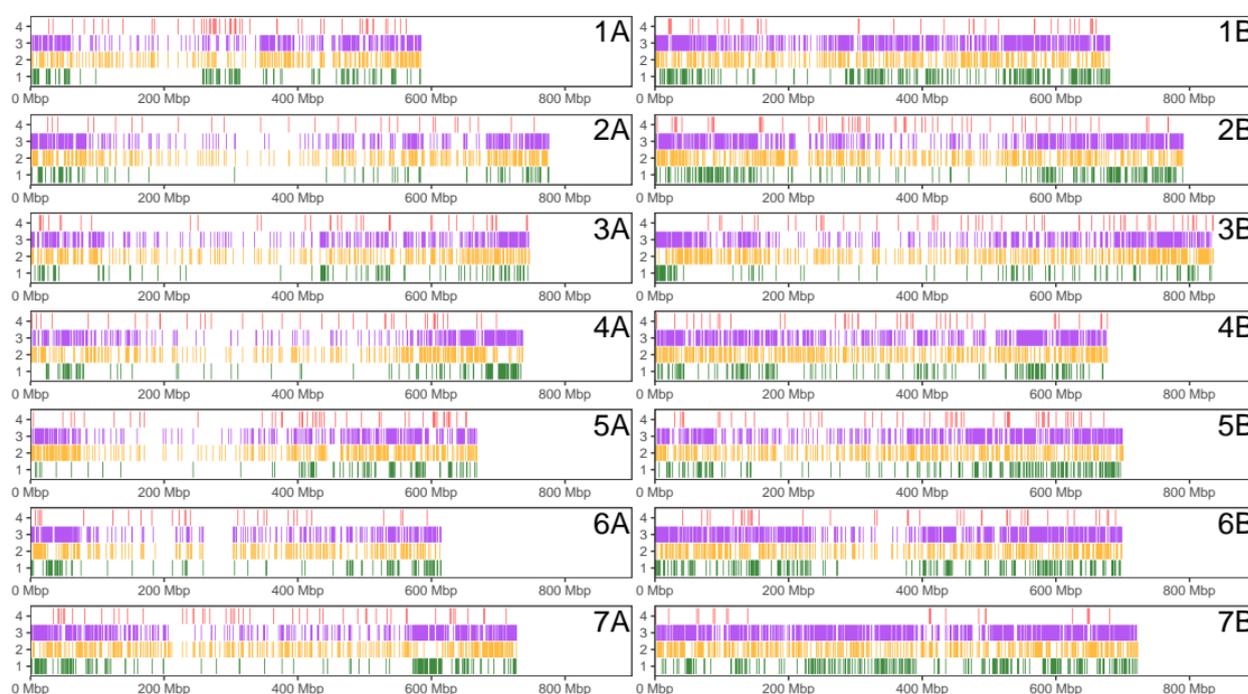
Интрогрессивная гибридизация играет важную роль в современной селекции, однако для создания новых сортов важно как можно полнее убирать генетический материал донора из генома реципиента (за исключением целевого участка). Количественная оценка остатков генетического материала донора в геноме реципиента является сложной задачей, особенно для мягкой пшеницы, геном которой составляет 16 Gbp. Еще более сложной задачей является оценка распределения генетического материала донора по геному из-за рекомбинации, которая уменьшает остаточные фрагменты генома донора из поколения в поколение. Одним из наиболее часто используемых методов для оценки генетического материала родителей в материале потомков является геномная *in situ* гибридизация (GISH), однако этот метод имеет ряд ограничений, таких как высокая стоимость, низкая производительность и разрешающая способность. Разработка подхода для оценки содержания и распределения генетического материала родителей в геноме потомков на основании GBS позволяет проводить одновременный анализ большого количества образцов, что значительно повышает производительность и значительно увеличивает разрешающую способность.

В проведенном исследовании для анализа принадлежности полиморфные SNP родителей и потомка в каждом сегменте генома мягкой и твердой пшеницы длиной 1 Mbp были использованы для построения филогенетического дерева для сегмента. Принадлежность сегмента в геноме потомка к одному из геномов родителей определялась на основе филогенетической дистанции между потомком и родителями. Подобный подход ранее никогда не проводился на растениях, а в литературе встречается лишь один пример его использования на малярийных

\* Исследование выполнено при поддержке РФФИ (проект № 24-16-00274).

комарах рода *Anopheles* [1]. Использование его на мягкой пшенице может быть осложнено из-за аллополиплоидной структуры генома, которая может привести к высокой гетерозиготности.

Для оценки возможности применения подхода на геноме мягкой и твердой пшеницы нами были секвенированы расщепляющиеся популяции — линии-потомки с поздних поколений самоопыления, а также их родители. Риды GBS образцов мягкой и твердой пшеницы выравнивались на референсные сборки IWGSC CS RefSeq v2.1 и Svevo.v1 соответственно. Картирование было проведено с использованием bowtie2 2.5.3 [2]. Для идентификации мутаций использовалась программа freebayes 1.3.7 [3]. Мультиаллельные SNP и инделы были удалены с помощью программы plink2 [4]. В результате после фильтров у мягкой пшеницы осталось 0,7–1,5 млн SNP, а у твердой пшеницы — 0,9–1,3 млн SNP. Для расчета деревьев в скользящем окне по геному был использован скрипт на R v4.3.1. Для построения дерева использовались только окна, содержащие не менее 10 SNPs. Каждое окно, принадлежность которого одному из родителей у потомков удалось определить однозначно, было окрашено в один цвет (см. рисунок).



Распределение генетического материала родителей в геноме потомков для двух образцов твердой пшеницы (1-2 и 3-4). Каждая вертикальная черта представляет из себя окно генома размером 1 Mbp. Цвет представляет принадлежность к одному из родителей. Белые зоны — области, непокрытые ридами при картировании, или окна, содержащие менее 10 SNPs

Результаты данной работы показали, что данные GBS могут быть использованы для определения принадлежности сегментов хромосом линий-потомков твердой и мягкой пшеницы к одному из линий-родителей. Данный алгоритм визуализации хромосом в дальнейшем может быть использован для поиска интрогрессий в других видах. Также планируется проверка данного метода на сниженном покрытии образцов потомков.

### Литература

1. Fontaine M.C. et al. Mosquito genomics. Extensive introgression in a malaria vector species complex revealed by phylogenomics // *Science* (N. Y.). 2015. Vol. 347, No. 6217. P. 1258524.
2. Langmead B. et al. Scaling read aligners to hundreds of threads on general-purpose processors // *Bioinformatics*. 2019. Vol. 35, No. 3. P. 421–432.
3. Garrison E., Marth G. Haplotype-based variant detection from short-read sequencing // arXiv. URL: <https://arxiv.org/abs/1207.3907>.
4. Chang C.C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets // *Gigascience*. 2015. Vol. 4, No. 1. P. s13742-015.