

DOI: 10.25205/978-5-4437-1691-6-15

**КЛАССИФИКАЦИЯ ОСНОВНЫХ ГЕТЕРОЗИСНЫХ ГРУПП КУКУРУЗЫ
НА ОСНОВЕ GENOTYPING-BY-SEQUENCING С ИСПОЛЬЗОВАНИЕМ WGS-ЯКОРЕЙ
ИЗ ОТКРЫТОГО ДОСТУПА***

**CLASSIFICATION OF MAJOR MAIZE HETEROTIC GROUPS BASED
ON GENOTYPING-BY-SEQUENCING USING PUBLICLY AVAILABLE
WGS REFERENCE SAMPLES**

В. С. Воронежская^{1,2}, А. С. Ермолаев¹, Д. С. Ульянов¹, А. Р. Дмитриева¹,
Н. М. Съедина¹, А. В. Васильев¹, Д. Ю. Сотченко^{1,3}, С. В. Тошчаков⁴

¹Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии, Москва

²Всероссийский научно-исследовательский институт радиологии и агроэкологии,
НИЦ «Курчатовский институт», Обнинск

³Всероссийский научно-исследовательский институт кукурузы, Пятигорск

⁴Национальный исследовательский центр «Курчатовский институт», Москва

V. S. Voronezhskaya^{1,2}, A. S. Ermolaev¹, D. S. Ulyanov¹, A. R. Dmitrieva¹,
N. M. Syedina¹, A. V. Vasiliev¹, D. Yu. Sotchenko^{1,3}, S. V. Toshchakov⁴

¹All-Russia Research Institute of Agricultural Biotechnology, Moscow

²All-Russia Research Institute of Radiology and Agroecology, National Research Centre “Kurchatov Institute”, Obninsk

³All-Russian Research Scientific Institute of Corn, Pyatigorsk

⁴National Research Centre “Kurchatov Institute”, Moscow

✉ voronezhskaya.vs@gmail.com

Аннотация

Предложен способ классификации гетерозисных групп кукурузы на основе результатов генотипирования образцов с помощью GBS. Преимуществом метода является отсутствие необходимости секвенировать эталонные образцы гетерозисных групп, поскольку в качестве якорей используются WGS основных гетерозисных групп из открытых баз данных. Для стратификации образцов кукурузы российской селекции на гетерозисные группы был использован анализ главных компонент (PCA) и оценка родства по мультилокусному набору SNP.

Abstract

A method for classifying heterotic groups of maize based on genotyping results from GBS (Genotyping-by-Sequencing) is proposed. The advantage of this method is the lack of necessity to sequence reference samples of heterotic groups, as the anchors are derived from whole-genome sequencing (WGS) of major heterotic groups available in public databases. To stratify samples of Russian-bred maize into heterotic groups, principal component analysis (PCA) and assessment of relatedness using a multilocus SNP panel were employed.

Явление гетерозиса заключается в увеличении агрономически важных показателей гетерозиготных гибридов F_1 по сравнению с гомозиготными инбредными родителями. Кукуруза была одной из первых культур, для которых гетерозис был использован в селекционном процессе. В современной селекции кукурузы используется несколько основных гетерозисных групп. Гетерозисная группа — это коллекция различных генотипов, дающих схожие по характеристикам гибриды при скрещивании с особями противоположных различных групп для получения лучшего потомства. При скрещивании растений, относящихся к одной группе, создаются новые инбредные линии, а в результате скрещивания разных групп создаются высокопродуктивные гибриды.

Определение гетерозисных групп крайне важно для селекции кукурузы и требует создания точных маркеров. Ранее для определения гетерозисных групп кукурузы преимущественно использовались ПЦР-маркеры (SSR, RAPD, AFLP и др.), однако с развитием технологий секвенирования наиболее активно стали использоваться SNP-маркеры, поскольку их легче идентифицировать в большом количестве и они обладают более высокой разрешающей способностью [1, 2]. В литературе уже описано SNP-генотипирование кукурузы с использова-

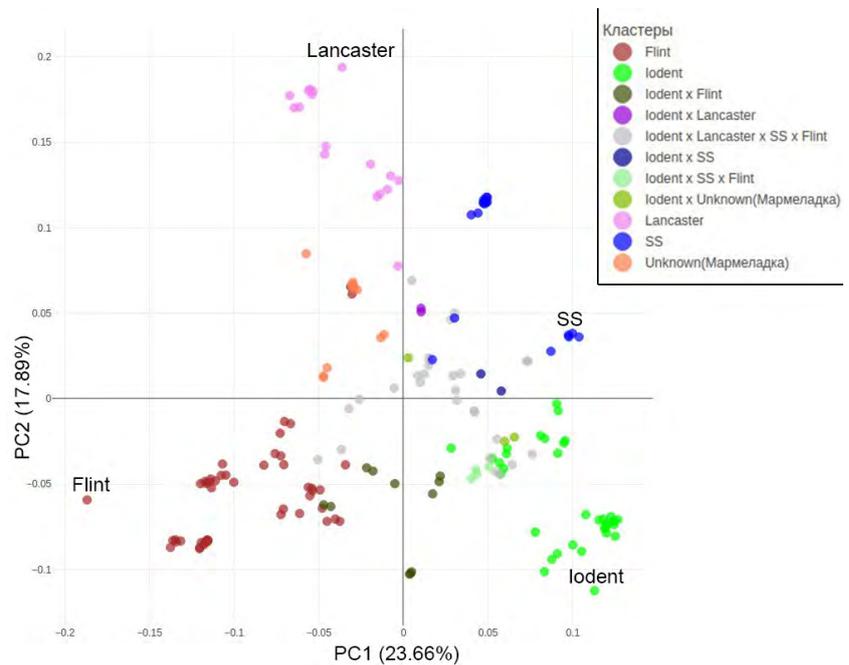
* Исследование выполнено при поддержке Министерства науки и высшего образования РФ (соглашение № 075-15-2019-1667).

© В. С. Воронежская, А. С. Ермолаев, Д. С. Ульянов, А. Р. Дмитриева, Н. М. Съедина, А. В. Васильев, Д. Ю. Сотченко, С. В. Тошчаков, 2024

нием транскриптомного [3] и полногеномного секвенирования с высоким покрытием [4]. GBS (Genotyping-by-Sequencing) также применялся для американских и китайских сортов [5].

В нашей работе представлен метод классификации образцов кукурузы по гетерозисным группам с использованием данных GBS, основанный на кластеризации образцов с референсными геномными сборками основных инбредных линий — B73 (Stiff Stalk), PH207 (Iodent), Mo17 (Lancaster), EP1 (Flint). В результате первичного SNP-коллинга было проанализировано 174 образца и выявлено 833 955 уникальных SNP. Дополнительно был проведен SNP-коллинг четырех референсных образцов (23523985 SNP). Результаты первичного SNP-коллинга для GBS и WGS-якорей были объединены с использованием программы Vcftools 1.20 [6]. Для повышения точности кластеризации пропущенные генотипы были предсказаны на основании генотипированных гапблоков с помощью программы Beagle 5.4 [7]. Всего было выбрано 1247 дискриминирующих SNP на основе анализа неравновесия сцепления локусов (linkage disequilibrium). Кластеризация проводилась в программе ADMIXTURE на основе оценки максимального правдоподобия мультилокусных наборов [8]. Результаты кластеризации показали, что выделенные группы образцов хорошо коррелируют с известными гетерозисными группами, что подтверждает эффективность нашего метода (см. рисунок).

Таким образом, в нашей работе был применен ранее не описанный метод кластеризации образцов GBS-секвенирования с WGS основных гетерозисных групп. Одним из главных преимуществ данного подхода является отсутствие необходимости в секвенировании геномов известных гетерозисных групп, что значительно сокращает временные и ресурсные затраты. Этот подход не только снижает затраты на исследование, но и делает его более доступным в условиях отсутствия коллекции якорных линий. Разработанный нами оптимизированный подход актуален для селекционных программ кукурузы, направленных на создание новых гибридов с повышенной продуктивностью.



PCA-график кластеризации образцов кукурузы по гетерозисным группам

Литература

1. Frascarelli E., Schrag T.A., Melchinger A. E. Genetic diversity analysis of elite European maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs // TAG. Theoretical and applied genetics. 2013. No. 1 (126). P. 133–141.
2. Yuan L. X. et al. Comparison of genetic diversity among maize inbred lines based on RFLPs, SSRs, AFLPs and RAPDs // Yi Chuan Xue Bao = Acta Genetica Sinica. 2000. No. 8 (27). P. 725–733.
3. Mazaheri M. et al. Genome-wide association analysis of stalk biomass and anatomical traits in maize // BMC Plant Biology. 2019. No. (19). P. 45.
4. Bornowski N. et al. Genomic variation within the maize stiff-stalk heterotic germplasm pool // The Plant Genome. 2021. No. 3 (14). P. e20114.
5. Shu G. et al. Genetic variation and population structure in China summer maize germplasm // Scientific Reports. 2021. No. 1 (11). P. 8012.
6. Danecsek P. et al. Twelve years of SAMtools and BCFtools // GigaScience. 2021. No. 2 (10). P. giab008.
7. Browning B. L., Zhou Y., Browning S. R. A One-Penny Imputed Genome from Next-Generation Reference Panels // The American Journal of Human Genetics. 2018. No. 3 (103). P. 338–348.
8. Mazaheri M. et al. Genome-wide association analysis of stalk biomass and anatomical traits in maize // BMC Plant Biology. 2019. No. (19). P. 45.