



Анализ однонуклеотидных полиморфизмов риса *Oryza sativa* на основе транскриптомных данных

Выполнила: Баранова Я.В., гр.17410 ФЕН

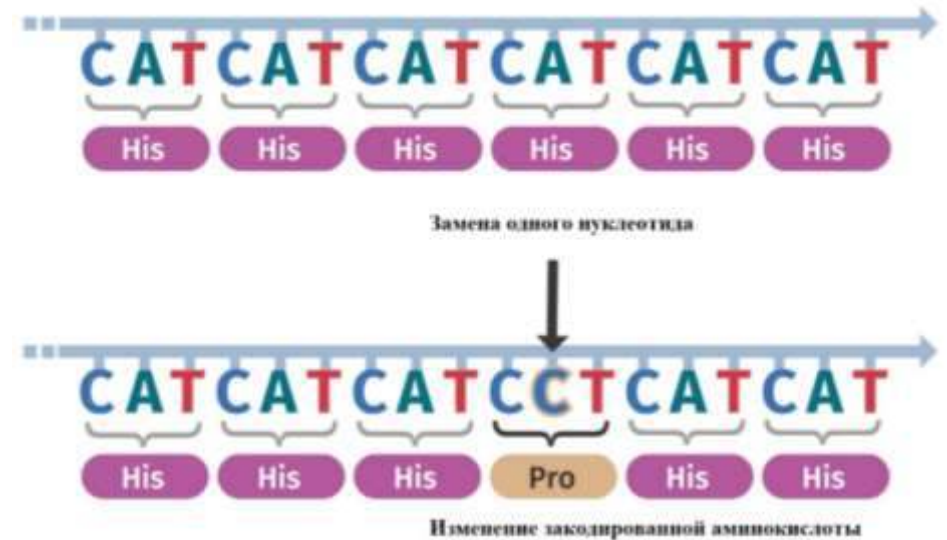
Научный руководитель: к.б.н. Генаев М.А.,
ст.н.с. лаб. эволюционной биоинформатики и
теоретической генетики ИЦиГ СО РАН

Новосибирск, 2021г.

Введение

Основной целью селекции растений является выведение более выгодных для человека форм сельскохозяйственных культур и их сортов. Поэтому изучение молекулярных причин формирования желаемого фенотипа, который проявляется не только в урожайности, но и в устойчивости к факторам среды, необходимо для дальнейшего развития сельского хозяйства и удовлетворения потребностей населения.

Однонуклеотидные полиморфизмы (ОНП) являются геномными молекулярными маркерами, характеризующимися высокой частотой встречаемости и эволюционной стабильностью, поэтому их использование эффективно для анализа функциональных полиморфизмов.



Объект исследования

В качестве модельного организма для исследования был выбран рис *Oryza sativa*. Это связано с разнообразием культивируемых и дикорастущих сортов растения, что позволяет сделать широкую выборку образцов и средовых факторов для анализа.

O. sativa имеет компактный диплоидный геном (375 млн п.н., $n=12$), поэтому является важным и удобным модельным видом среди однодольных и злаковых растений.

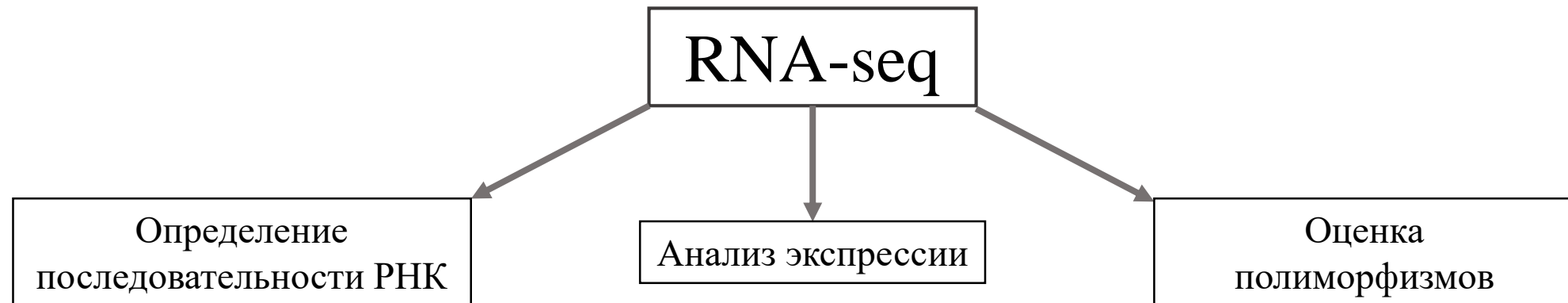
Также показано, что блоки функциональных сегментов на генетической карте риса позволяют предсказывать и выявлять подобные блоки на хромосомах кукурузы и пшеницы.



Секвенирование нового поколения

На сегодняшний день секвенирование нового поколения (NGS) является одним из наиболее эффективных, недорогих и массовых методов секвенирования. Эта технология позволяет прочитывать последовательности в масштабах полного генома, анализ же транскриптома дает информацию исключительно об экспрессирующихся генах.

Поэтому работа с данными транскриптомного секвенирования имеет определенные ограничения, например, отсутствует возможность оценить полиморфизмы в неэкспрессируемых частях генома.



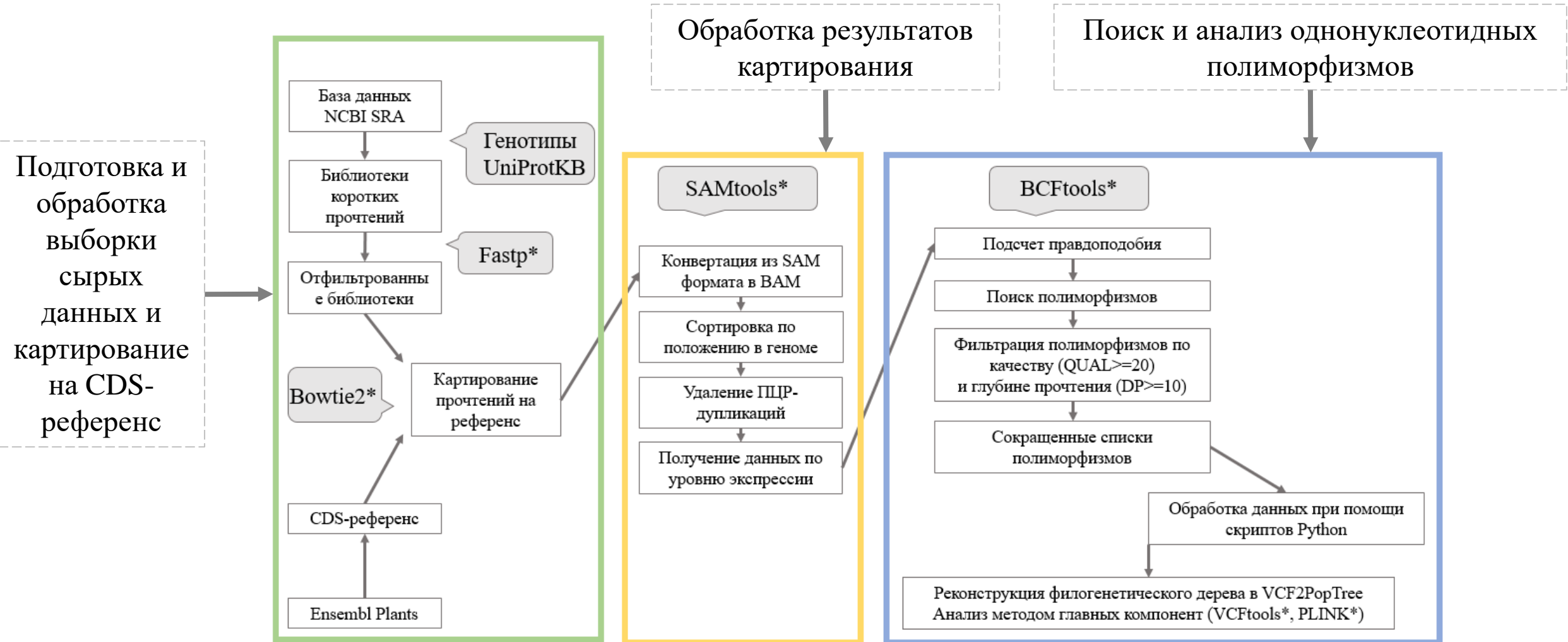
Цели и задачи исследования

Цель работы – сравнительный анализ генотипов риса *Oryza sativa* ssp. *Japonica* на основе данных об однонуклеотидных полиморфизмах.

Для достижения данной цели были поставлены следующие задачи:

1. Составить выборку генотипов риса, представленных в базе данных NCBI SRA;
2. Выявить однонуклеотидные полиморфизмы, встречающиеся в кодирующих последовательностях данных генотипов риса;
3. Сравнить соотношение синонимичных и несинонимичных полиморфизмов в данных генотипах риса;
4. Для подвыборки генотипов, транскриптомы которых были получены из семян, проанализировать связь наблюдаемых полиморфизмов с фенотипическими проявлениями.

Конвейер биоинформатической обработки



Работа с инструментами, отмеченными *, производилась на кластере ЦКП "Биоинформатика"

Обнаруженные однонуклеотидные полиморфизмы

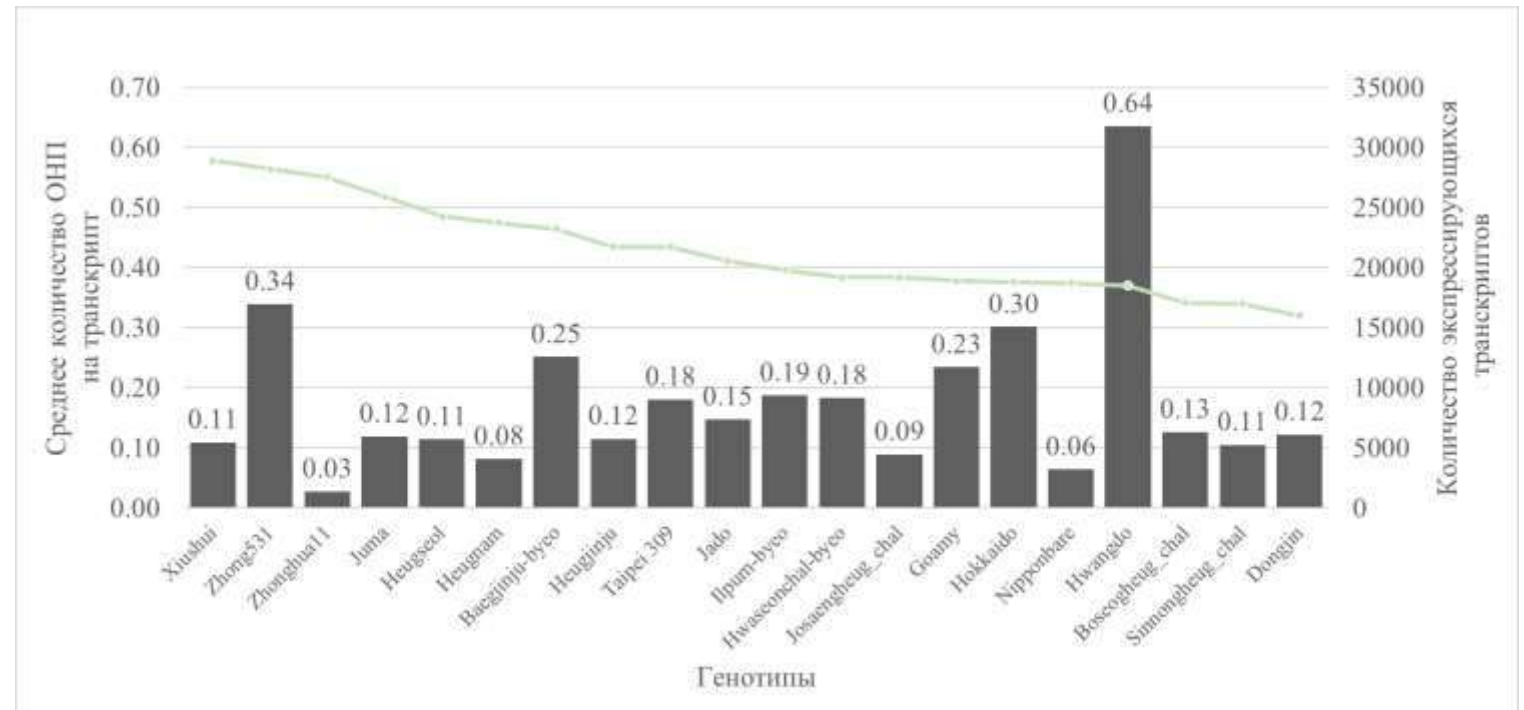
В данной работе было обработано 67 библиотек коротких прочтений, относящихся к 20 различным генотипам риса *O. sativa*.

Из них получен список всех однонуклеотидных полиморфизмов *O. sativa*, присутствующих как минимум в одном из генотипов. Список включает 35 572 ОНП в 9 510 транскриптах, из них 21 398 (60.15 %) являются уникальными, то есть характерны только для одного из генотипов, и 14 174 (39.85 %) встречаются сразу в нескольких генотипах.

Среднее количество ОНП в каждом из генотипов *O. sativa*

Генотипы Baegjinju-byeo, Goamy, Hokkaido, Hwangdo и Zhong531 характеризуются наибольшей эволюционной отдаленностью от референсного генотипа Nipponbare. У генотипов Heugnam, Josaengheug_chal и Zhonghua11, напротив, наблюдается схожесть с референсным генотипом.

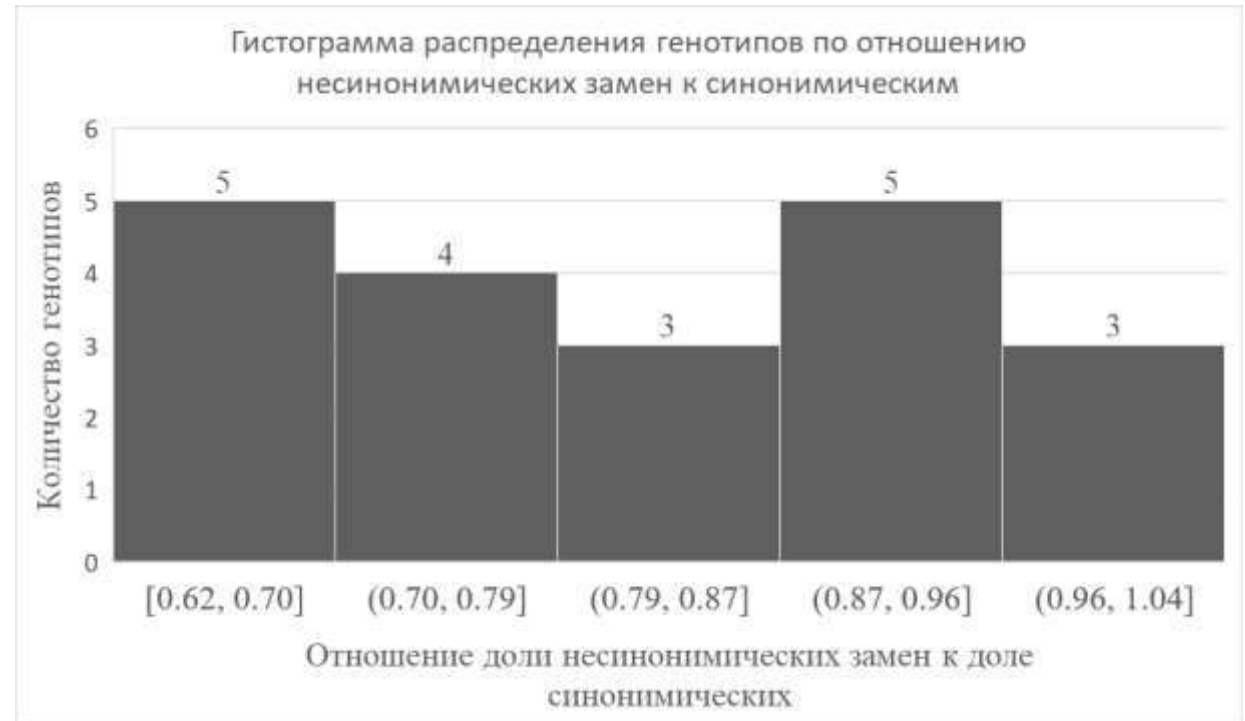
Интересно, что число ОНП относительно числа экспрессирующихся транскриптов у генотипа Hwangdo значительно превосходит показатели для остальных генотипов. Этот результат согласуется с полученными ранее данными о низкой гомологии последовательности генома Hwangdo и других генотипов *O. sativa* [1].



Анализ режима замен в кодонах

Наибольшим отношением несинонимических замен к синонимическим характеризуются генотип Hokkaido (1.04). Наименьшим – генотип Voseogheug_chal (0.62).

Исходя из гистограммы, можно предположить, что эволюция большинства генотипов *O. sativa* происходит в режиме слабого стабилизирующего отбора.



Плотность ОНП на хромосомах

Наиболее варьируемыми являются кодирующие последовательности 11 и 12 хромосом, а также кодирующие последовательности геномов органелл.

Эти данные согласуются с литературными [1-4].



[1] Yu J. et al. The Genomes of *Oryza sativa*: A History of Duplications // PLoS Biology. 2005. Vol. 3, № 2.

[2] Choise, N. et al. The sequence of rice chromosomes 11 and 12, rich in disease resistance genes and recent gene duplications // BMC Biology. 2005. Vol. 3, № 1. P. 20.

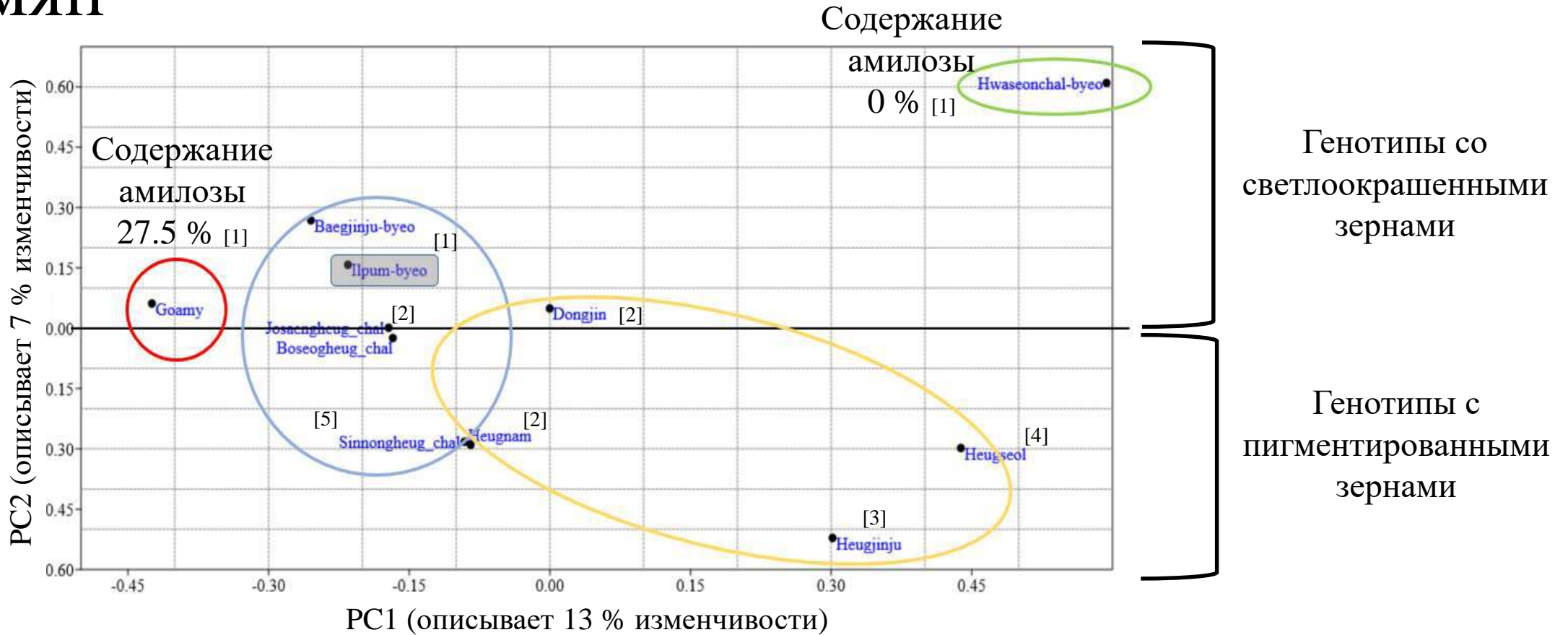
[3] Gao L.-Z. et al. Evolution of *Oryza* chloroplast genomes promoted adaptation to diverse ecological habitats // Communications Biology. 2019. Vol. 2, № 1.

[4] Cheng L., Kim K.-W., Park Y.-J. Evidence for selection events during domestication by extensive mitochondrial genome analysis between japonica and indica in cultivated rice // Scientific Reports. 2019. Vol. 9, № 1.

Гены риса с наибольшими числом замен

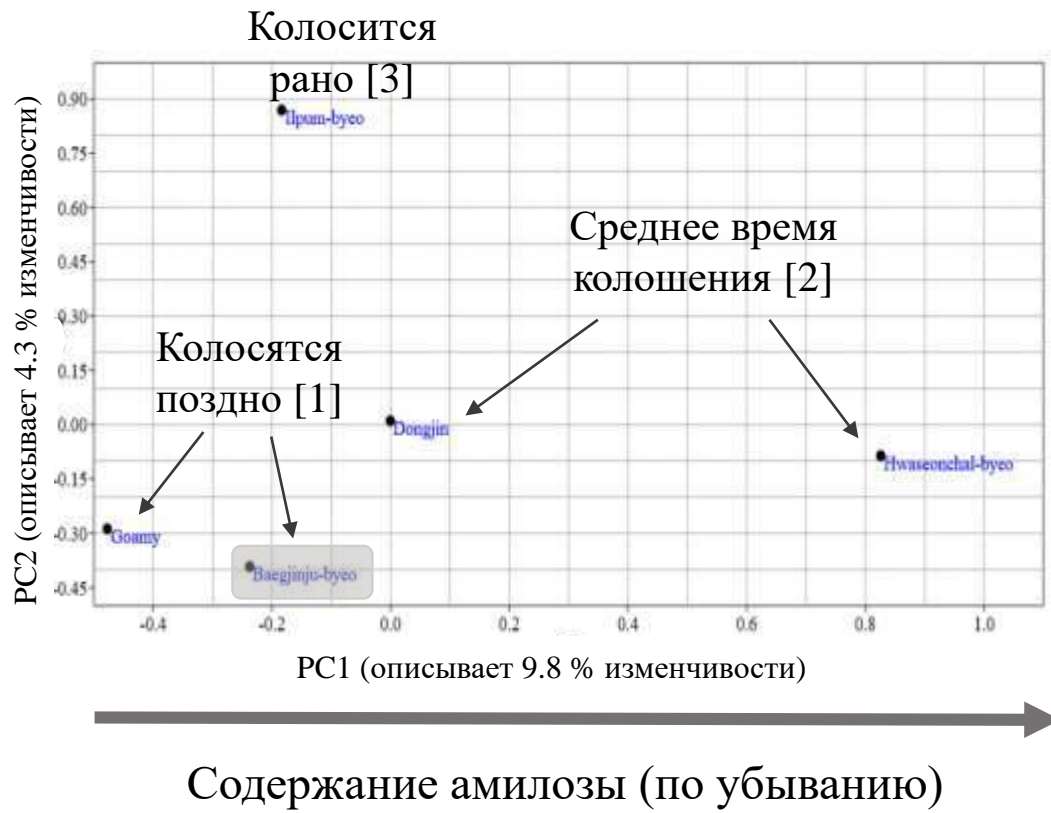
Ген	Среднее число замен на длину транскрипта	Описание белка
<i>Sdh3-2</i>	0.026	Заякоренная на внутренней мембране митохондрии субъединица комплекса II (SDH3-2) сукцинатдегидрогеназы (SDH). Участвует в окислительном фосфорилировании и цикле трикарбоновых кислот.
Os12g0632700	0.02	Глиоксисомальная малатдегидрогеназа. Участвует в глиоксилатном цикле, цикле трикарбоновых кислот и других реакциях с участием малата.
<i>Cgt</i>	0.017	UDP-глюкозозависимая гликозилтрансфераза, катализирующая С-гликозилирование дигидроксифлаванов, которые являются вторичными метаболитами растений, обладающими антиоксидантными, противомикробными и привлекающими насекомых свойствами.

Анализ методом главных компонент для ОНП генотипов, полученных при секвенировании семян

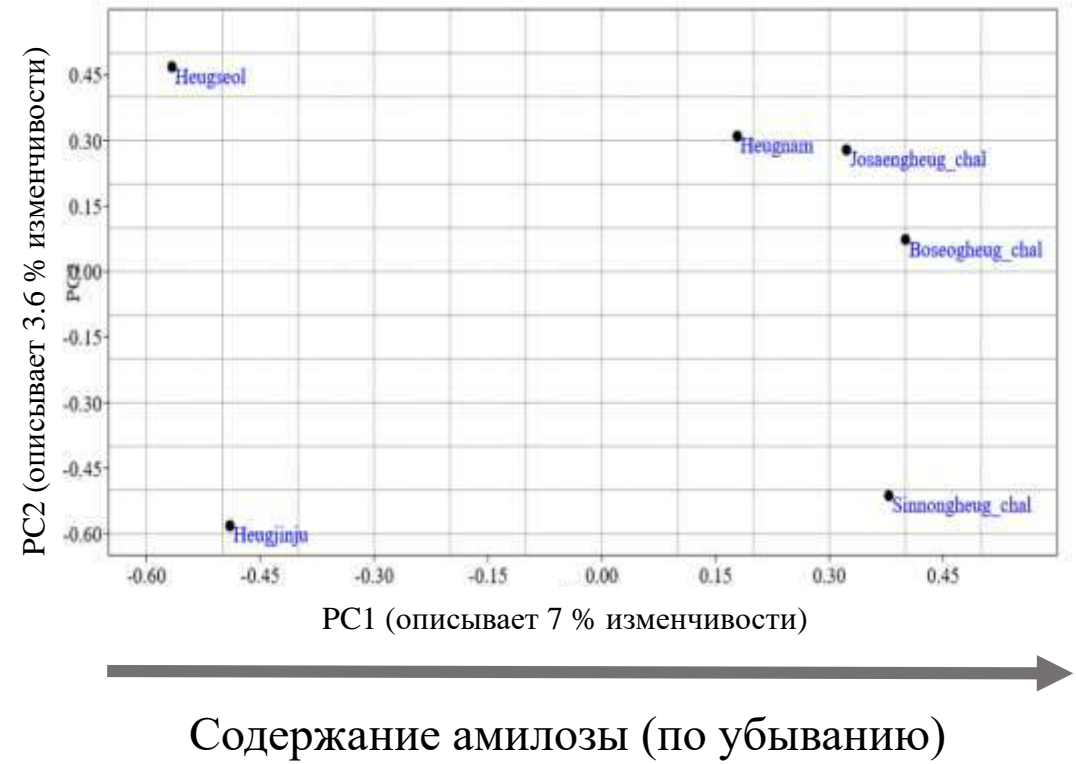


- [1] Lee J.-G. et al. The Chemical Quality Properties during Processing Treatment of Brown Rice Varieties with Different Amylose Content // Journal of Applied Biological Chemistry. 2010. Vol. 53, № 1.
- [2] Kim M.-S. et al. Breeding of High Cooking and Eating Quality in Rice by Marker-Assisted Backcrossing (MABc) Using KASP Markers // Plants. 2021. Vol. 10, № 4.
- [3] Ryu S.-N. Grain Quality of New Large-Grain Pigmented Rice Variety, “Daeripjamibyeo” // Korean Journal of Crop Science. 2013. Vol. 58, № 2.
- [4] Shin D.-S. et al. Quality of Tteokbokki tteok prepared by adding various concentration of brown rice // Korean Journal of Food Preservation. 2016. Vol. 23, № 2.
- [5] Surh J., Koh E. Effects of four different cooking methods on anthocyanins, total phenolics and antioxidant activity of black rice // Journal of the Science of Food and Agriculture. 2014. Vol. 94, № 15.

Анализ методом главных компонент для ОНП генотипов, характеризующихся различной окраской зерен



Время колошения (по
возрастанию)



[1] Surh J., Koh E. Effects of four different cooking methods on anthocyanins, total phenolics and antioxidant activity of black rice // Journal of the Science of Food and Agriculture. 2014. V. 94. № 15.

[2] Chun H., Choi. Current Status of Varietal Improvement in Specialty Rice Suitable for Food Processing and High-Functional Utility

[3] Improvement in Rice Technology and System in Korea [https://cropgenebank.sgrp.cgiar.org/] // The Crop Genebank Knowledge Base is an initiative of the System-wide Genetic Resources Programme (SGRP) of the Consultative Group on International Agricultural Research (CGIAR).

Выводы

1. Сформирована выборка из 67 библиотек коротких прочтений для 20 генотипов *O. sativa*;
2. Выявлены характерные для каждого из генотипов однонуклеотидные полиморфизмы: показано, что их число на транскрипт варьирует от 0.635 (генотип Hwangdo) до 0.027 (генотип Zhonghua 11);
3. Показано, что отношение количества несинонимических замен к синонимическим варьирует от 0.62 (генотип Boseogheug_chal) до 1.04 (генотип Hokkaido);
4. Анализ методом главных компонент позволил разделить исследованные генотипы риса *O. sativa* согласно содержанию амилозы в зернах, окраске зерен и временем колошения.

Спасибо за внимание!