

Алгоритм отбора минимального дискриминирующего набора SSR-локусов для целей генетической паспортизации растений: опыт с ячменём (*Hordeum vulgare* L.)

Научный руководитель – Дивашук Михаил Георгиевич

Стрембовский Илья Викторович

Аспирант

Всероссийский научно-исследовательский институт сельскохозяйственной

биотехнологии, Москва, Россия

E-mail: 0225vvs@gmail.com

Традиционно установление сортовой идентичности основано на оценке набора морфологических признаков, отличающих один сорт от другого. Однако внешне выравненный сорт на генетическом уровне может представлять собой сложную систему из внешне неразличимых биотипов. Длительное возделывание такого гетерогенного сорта в конкретных климатических условиях может привести к доминированию одного из биотипов и изменению общей архитектуры сорта [1]. В таком случае оценка по морфологическим признакам не является исчерпывающей и может быть дополнена данными о структуре полиморфных участков ДНК сорта – его генетическим паспортом. Методика паспортизации основана на генотипировании выборки растений сорта набором генетических маркеров и, в частности, маркеров, основанных на микросателлитных повторах (simple sequence repeats, SSR). Подходы по отбору наборов SSR-локусов, на которые в дальнейшем разрабатываются маркеры, крайне многочисленны и, в некоторых случаях, противоречивы, что приводит к формированию наборов с диаметрально противоположными характеристиками [2]. Поэтому для поддержания развития генетической паспортизации растений в России нами был разработан алгоритм, основанный на симуляции *in silico* ПЦР (полимеразной цепной реакции) и позволяющий отбирать наиболее информативные SSR-локусы. Модельной культурой был выбран ячмень (*Hordeum vulgare* L.), генетическая паспортизация которого широко применяется в производстве.

Концептуально предлагаемый алгоритм можно представить как набор последовательно исполняемых этапов: 1) сбор базы данных известных SSR-локусов и удаление дубликатов (наша база включала 634 уникальных SSR-локуса), 2) симуляция *in silico* ПЦР на пуле аннотированных сборок (27 шт. сборок генома ячменя) и отбор SSR-локусов, амплифицирующих специфичный ПЦР продукт оптимального для детекции размера в 800 п.н. (набор из 204 локусов), 3) сокращение массива отобранных SSR-локусов и отбор наиболее информативных из них с использованием алгоритма RFE (recursive feature elimination) с поддержкой методом бутстрапа с 10000 итерациями (7 SSR-локусов), 4) маскирование отобранного набора локусов и его искусственное расширение за счёт повторного применения алгоритма RFE (13 SSR-локусов) и 5) расширение набора и повышение его устойчивости за счёт добавления высокополиморфных SSR-локусов, отобранных алгоритмом RFE (18 SSR-локусов).

Конечный паспортизирующий набор, полученный по итогам работы алгоритма, состоял из 18 SSR-локусов (SSR1-SSR16), способный различать все доступные в настоящее время сборки генома ячменя, собранные до уровня хромосом и выше (27 шт.). Набор был намеренно сделан избыточным, чтобы различать генотипы, не представленные в публичных базах данных.

Оценка устойчивости набора к выпадению отдельных SSR-локусов набора, т.е. отсутствию амплификации по той или иной причине, проводимая методом симуляции, выявила,

что большинство локусов набора могут быть заменены в случае своего выпадения на аналогичный локус с вероятностью не ниже 45% (пороговое значение; Рисунок 1). Под вероятностью в данном случае понимается количество симуляций во время которых возможно выпадение одного из маркеров, но разрешающая способность всего набора (способность различать сорта) не снижается, т.к. выпавший маркер заменяется на другой. Исключение составляет локус SSR16, который может быть заменён на локус SSR1 или SSR15 с вероятностью 43,74%, т.е. ниже порогового значения в 45%. Из-за наличия подобных “малозаменимых” локусов конечный набор разрабатывается искусственно расширенным.

В дальнейшем на основании отобранных SSR-локусов были разработаны соответствующие генетические маркеры и эффективность паспортизирующего набора была подтверждена на практике.

Источники и литература

- 1) 1. Поморцев А. А., Лялина Е. В. Лабораторный контроль сортовой принадлежности и сортовой чистоты партий семян и товарного зерна пивоваренного ячменя // Зерновое хозяйство России. 2009. №. 6. С. 29-39.
- 2) 2. Development and genetic mapping of 127 new microsatellite markers in barley / J.Z. Li, T.G. Sjakste, M.S. Röder, M.W. Ganai // Theor. Appl. Genet. 2003. №6. С. 1021–1027. DOI : 10.1007/s00122-003-1345-6.

Иллюстрации

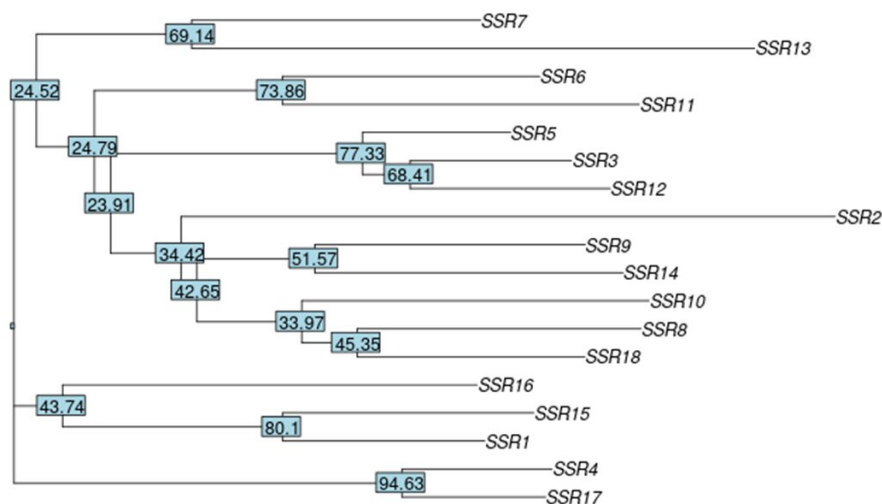


Рис. : Рисунок 1 – Дендрограмма подобия SSR-локусов с поддержкой бутстрапом. Числовые значения в узлах дерева показывают близость разрешающих способности локусов/кластера локусов