

Анализ вариаций в геномах картофеля на основе данных полногеномного секвенирования

Д.И. Каретников^{1*}, Г.В. Васильев¹, С.В. Тощак², Н.А. Шмаков¹, М.А. Генаев¹, М.А. Нестеров¹, С.М. Ибрагимова¹, Д.А. Рыбаков³, Т.А. Гавриленко³, Е.А. Салина¹, М.В. Патрушев³, А.В. Кочетов¹, Д.А. Афонников¹

¹ Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² НИЦ Курчатовский институт, Москва, Россия

³ ФИЦ Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова, Санкт-Петербург, Россия

* karetnikovmit@bionet.nsc.ru

Цель: *Solanum tuberosum* L. – важнейшая сельскохозяйственная культура, выращиваемая практически во всем мире. Общий объем производства составляет более 370 млн тонн в год, а количество сортов превышает 4000. Картофель является автотетраплоидом, обладает высокой гетерозиготностью, обусловленной преимущественно вегетативным способом размножения. Полногеномное секвенирование открывает путь к изучению структуры генома картофеля, выявлению однонуклеотидных полиморфизмов, связанных с различными фенотипическими признаками. Одной из важных особенностей анализа генома является вариабельность числа копий (CNV), определяемая как участки генома размером от 1 Кб до нескольких мегабаз, которые демонстрируют разное количество копий в популяции. На функциональном уровне CNV связан с генами, участвующими в стрессовых реакциях и устойчивости к различным факторам.

Ранее было секвенировано 15 сортов картофеля *S. tuberosum*, выращиваемых в России. На основе секвенированных геномов был реконструирован пангеном и определены потенциальные CNV, ассоциированные с генами. В этой работе выборка картофеля, выращиваемого в России, была расширена.

Материалы и методы: Геномы *S. tuberosum* были секвенированы с использованием платформы Illumina, что позволило реконструировать геномы *de novo*. На основе коротких прочтений и референсного генома DM1-3 версии 6 с помощью CNVpytor были идентифицированы CNV в геномах. Кроме того, мы идентифицировали SNP, короткие инсерции и делеции во всех хромосомах с помощью программного комплекса GATK. Все выявленные варианты были функционально аннотированы, на основе чего были выявлены варианты, связанные с высоким функциональным вкладом в набор генов, ассоциированных с циркадными ритмами, созреванием и клубнеобразованием. Также на основе коротких прочтений были реконструированы пластымы отечественных сортов картофеля и выявлены их типы.

Результаты: Реконструированный пангеном картофеля *S. tuberosum* сортов отечественной селекции продемонстрировал увеличение вариабельной части, что говорит о том, что пангеном картофеля отечественных сортов является открытым. Анализ CNV выявил количественные различия между сортами, выращиваемыми в России, и южноамериканскими сортами, которые связаны с генами, ассоциированными с циркадными ритмами картофеля. Кроме того, короткие считывания позволили идентифицировать однонуклеотидные полиморфизмы во всех геномах и провести их дальнейший количественный анализ, а также сравнительный относительно различных сортов Европы и дикого картофеля. Дополнительно мы оценили встречаемость полиморфизмов с высоким функциональным вкладом в генах, ассоциированных с созреванием картофеля, а также проанализировали аллельное разнообразие гена StCDF1, являющегося главным регулятором созревания картофеля, влияя на сроки клубнеобразования.

Выводы: Таким образом, мы провели комплексный анализ геномов картофеля *S. tuberosum* отечественных сортов, оценив встречаемость CNV, частоты SNP и InDel, а также частоты аллелей гена StCDF1. Помимо этого мы выявили типы хлоропластной ДНК картофеля.

Финансирование: Работа выполнена за счет финансирования бюджетного проекта FWNR-2022-0020. Вычисления проводились с использованием ресурсов ЦКП «Биоинформатика».