

Исследование транскриптома почти изогенных линий ячменя с альбинизмом и меланизмом колоса

Участники:

1. Шмаков Николай Александрович, аспирант ИЦиГ СО РАН
2. Афонников Дмитрий Аркадьевич, зав лаб ИЦиГ СО РАН, к.б.н.
3. Хлёткина Елена Константиновна, зав лаб ИЦиГ СО РАН, д.б.н.
4. Васильев Георгий Владимирович, с.н.с, к.б.н.
5. Глаголева Анастасия Юрьевна, магистрант НГУ

Гранты:

РФФИ 16-34-00924

РНФ 16-14-00086

Описание:

1. Постановка задачи
Изучить генетические и транскриптомные особенности почти изогенных линий ячменя *Hordeum vulgare* с целью выявить гены, ассоциированные с образованием тканеспецифичных аномальных типов окраски
2. Состояние проблемы
Ячмень *Hordeum vulgare* L. ($2n = 2x = 14$; размер генома - 5.5млн.п.н.) – один из основных культурных видов злаков. В 2012 году была опубликована сборка генома ячменя, включающая в себя, по меньшей мере, 95% последовательности генома ячменя. Это существенно облегчает возможность использования высокпроизводительных методов, таких как RNA-seq для решения задач, связанных с идентификацией генов и реконструкцией генных сетей, лежащих в основе фенотипического разнообразия по тому или иному признаку. Растительные пигменты – собирательное название разнородных по химической природе и происхождению соединений, придающих окраску органам и тканям растений. Бесспорно, наиболее известным, распространённым и важным растительным пигментом является хлорофилл. При этом, несмотря на подробное описание биосинтеза хлорофилла на биохимическом уровне, и известную структуру многих генов, участвующих в этом процессе, неизвестными остаются генетические механизмы, контролирующие тканеспецифичный синтез хлорофилла. Мутация *Alm*, описанная у ячменя, приводит к потере синтеза хлорофилла в перикарпе, цветковой чешуе, стебле и ушках. *Alm* – это ядерный ген, локализованный в коротком плече хромосомы 3Н. Его последовательность не известна, как и структура его белкового продукта. Не ясно и взаимодействие данного гена с другими генами, участвующими в синтезе хлорофилла. Другой класс растительных пигментов – фитомеланины, придающие чёрную окраску некоторым из органов растений, например – коже семен подсолнечника. Эти соединения устойчивы к действию большинства растворителей, поэтому их точная химическая природа не ясна, как и механизмы их биосинтеза и гены, участвующие в этом процессе. Мутация в гене *Blp* у ячменя приводит к накоплению фитомеланинов в лемме и перикарпе. Как и в случае с геном *Alm*, ген *Blp* локализован в геноме с точностью до длинного плеча хромосомы 1Н, но его структура не известна, как и его белковый продукт, механизм воздействия и взаимодействие с другими генами. Для выяснения этих вопросов наиболее современным и подходящим методом является секвенирование транскриптома. Хорошими моделями для выявления дифференциально экспрессирующихся генов методом RNAseq являются мутантные либо изогенные линии. Почти изогенные линии – подходящая модель

для выделения генов, лежащих в основе фенотипической вариации. Линия ячменя Bowman и её почти изогенные линии, несущие соответствующие мутации в генах *Alm* и *Blp*, были любезно предоставлены др. Бёрнером (IPK Gatersleben, Germany). Полученные результаты будут способствовать решению одной из актуальных задач молекулярной генетики растений – установлению взаимодействия ядерного и хлоропластного геномов злаковых растений на уровне тканеспецифичной регуляции генов.

3. Методы

Секвенирование транскриптомов перикарпа исследуемых почти изогенных линий ячменя, а также изогенной линии Bowman, использованной в обоих случаях в качестве контроля, проводилось на платформе IonProton. Библиотеки были обработаны программами FastQC и Prinseq с целью оценки распределения параметров качества и длин прочтений. Последовательности адаптеров были удалены программой CutAdapt. При помощи программы Prinseq также была проведена фильтрация фрагментов по длинам и значениям качества. Были удалены последовательности с длиной менее 50 нуклеотидов и средним значением качества Phred меньше 20.

Для картирования библиотек использовали геном ячменя версии 082214v1.28, аннотированный в базе данных Ensembl Plants. Картирование осуществлялось при помощи программ STAR и TopHat2. В целях подбора оптимальной стратегии картирование осуществлялось программой TopHat2 с различными значениями параметра «количество допустимых несоответствий» (allowed mismatches), от 2 до 20. Всего было проведено 10 картирований для каждой библиотеки. При помощи программы STAR каждая библиотека была картирована 4 раза, со значениями параметра "outFileMismatchMax" равным 4, 8, 12 и 18. Были оценены такие параметры, как доля картированных фрагментов из каждой библиотеки и количество покрытых участков генома. В результате, в дальнейшую обработку были взяты картирования, произведённые программой TopHat2 с допущенными 18 несоответствиями между картируемой последовательностью и референсом.

Оценка дифференциальной экспрессии проводилась при помощи пакета программ Cufflinks и пакета EdgeR для языка R. Гены, имеющие пониженный и повышенный уровни экспрессии в каждой из исследуемых линий по сравнению с контрольной линией, были проанализированы по отдельности на предмет участия в метаболических и регуляторных генных сетях при помощи онлайн-сервиса, предоставляемого базой данных PlantCyc. Анализ генной онтологии был проведён с помощью базы данных AgriGO

С помощью программы Trinity была осуществлена сборка транскриптома de novo линии Alm. Для проведения сборки транскриптома de novo библиотеки, соответствующие линии Alm и контрольной линии Bowman, были сведены в один файл. Далее каждый из файлов был обработан программой Trinity. Полученные сборки были выровнены на имеющуюся версию генома ячменя при помощи программы Blast.

4. Результаты

Около 13% прочтений были удалены из библиотек линии Alm и Bowman в процессе фильтрации. Из библиотек линии Blp было удалено в среднем 15% ридов.

Сборки de novo для линий Alm и Bowman состоят из 86,8 и 107 тысяч транскриптов, соответственно, параметр N50 составляет 1112 и 1150, соответственно. 564 транскрипта из сборки линии Alm и 604 транскрипта для сборки Bowman не показали гомологии к геному ячменя, из них 320 транскриптов встречаются в обеих сборках. Для линии Blp сборка de novo не проводилась.

Линейка программ Cufflinks выделила 119 транскриптов, имеющих повышенный уровень экспрессии в линии Alm и 783 транскрипта, имеющих пониженный

уровень экспрессии в этой линии. Для линии B1p было выделено 632 и 325 up- и down-регулируемых генов, соответственно. Пакет EdgeR выделил 79 транскриптов, имеющих повышенный уровень экспрессии в линии Alm и 802 транскрипта, имеющих пониженный уровень экспрессии в этой линии. В линии B1p пакет EdgeR выделил 1115 и 648 up- и down-регулируемых генов, соответственно. При этом были выделены 40 оперонов хлоропластного генома, имеющих пониженный уровень экспрессии в линии Alm, и не было выделено оперонов в геноме хлоропластов, имеющих повышенный уровень экспрессии в этой линии.

Было установлено, что гены, имеющие более низкий уровень экспрессии в линии Alm, участвуют в таких метаболических путях, как цикл Кальвина-Бенсона, цикл ксантинов, цикл хлорофилла и другие. Далее, анализ генной онтологии показал, что для генов, имеющих более высокий уровень экспрессии в линии Alm наиболее представленным термином генной онтологии является связь с цитоплазматическими везикулами, в то время как для генов с пониженной экспрессией в этой линии наиболее достоверными терминами генной онтологии являются «фотосинтез», «биосинтез порфиринов», а также локализация в клетке в пластидах и в мембранных органоидах. Для ДЭГ в линии B1p было обнаружено участие в метаболическом пути «биосинтез флавонолов» и «дефосфорилирование UTP и CTP», и представленность ГО терминов «биосинтез жирных кислот», «тилакоиды» и «активность транскрипционных факторов»

Использование ресурсов ИВЦ НГУ:

Анализ и фильтрация библиотек, картирование на геном, оценка уровней экспрессии, поиск дифференциальной экспрессии проводились на ресурсах ИВЦ НГУ.

Публикации:

Glagoleva AY, Shmakov NA, Shoeva OY, Vasiliev GV, Shatskaya NV, Börner A, Afonnikov DA, Khlestkina EK. Metabolic pathways and genes identified by RNA-seq analysis of barley near-isogenic lines differing by allelic state of the Black lemma and pericarp (B1p) gene // BMC Plant Biology. 2017; 17(Suppl 1):182. doi: 10.1186/s12870-017-1124-1

Shmakov NA, Vasyliiev GV, Shatskaya NV, Doroshkov AV, Gordeeva EI, Afonnikov DA, Khlestkina EK. Identification of nuclear genes controlling chlorophyll synthesis in barley by RNA-seq // BMC Plant Biology. 2016; 16 (Suppl 3) :245. doi: 10.1186/s12870-016-0926-x