

Биоинформатический поиск и анализ генов, участвующих в формировании чёрной и белой окраски колоса ячменя *Hordeum vulgare***Научный руководитель – Афонников Дмитрий Аркадьевич*****Шмаков Николай Александрович****Аспирант*

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения РАН, Новосибирск, Россия

E-mail: shmakov@bionet.nsc.ru

Растительные пигменты - собирательное название соединений различной природы и происхождения, придающих окраску органам и тканям растений. Наиболее известные растительные пигменты - хлорофиллы, участвующие в процессе фотосинтеза и, следовательно, имеющие огромную важность для жизни растения и огромное сельскохозяйственное значение. Другой класс растительных пигментов - флавоноиды, производные флавонола, придающие частям растений различную окраску, от синей до чёрной. Флавоноиды участвуют в защите растений от патогенов и паразитов и обладают антиоксидантным эффектом.

RNA-seq - технология геномных и транскриптомных исследований, позволяющая одновременно оценить уровень экспрессии всех генов в исследуемых органах или тканях [6]. Наиболее часто технология RNA-seq применяется для поиска дифференциальной экспрессии генов в изучаемых образцах. Помимо этого технология предоставляет возможность обнаружения не аннотированных ранее генов, изоформ, замен и делеций/вставок в кодирующих частях генов.

Ячмень обыкновенный *Hordeum vulgare* L. - сельскохозяйственно-важный вид злаков. Геном ячменя диплоидный, состоит из семи пар хромосом, секвенирован и аннотирован. Текущая версия сборки содержит около 83 тысяч генов [5]. Кроме того, существуют почти изогенные линии ячменя i:BwAlm и BLP, выведенные на основе линии Bowman, отличающиеся отсутствием зелёной окраски леммы и чёрной окраской леммы и перикарпа, соответственно. Наличие высококачественной сборки генома и почти изогенных линий, имеющих нужные признаки, позволяет провести эксперимент RNA-seq для выявления особенностей транскриптома растений, имеющих аномальную окраску колоса.

В эксперименте была выделена РНК из колосковых чешуй растений линий i:BwAlm, BLP и Bowman, после чего было проведено секвенирование на платформе IonTorrent. После предобработки библиотеки были картированы на референсный геном *H. vulgare* программами TopHat2 [4] и STAR [2]. При помощи линейки программ Cufflinks [8] и пакета EdgeR [7] для языка R был осуществлён поиск дифференциальной экспрессии генов. Гены, имеющие дифференциальную экспрессию, были проанализированы на предмет перепредавленности терминов генной онтологии и участия в метаболических и регуляторных генных сетях. Программа Trinity [3] была использована для сборки транскриптома *de novo*, после чего программы gmap [9], blast [1] и вспомогательные утилиты Trinity использовались для аннотации полученных контигов.

Для генов, имеющих повышенную экспрессию в линии i:BwAlm, была показана функциональная связь с везикулярным транспортом. Для генов, имеющих повышенную экспрессию в линии BLP, было показано участие в пути синтеза суберина, кутина, фенилпропаноидов и фитоалексинов. При анализе *de novo* сборки транскриптома были обнаружены контиги, не имеющие гомологии с текущей сборкой генома ячменя, но имеющие

значительные гомологии с последовательностями, представленными в базе данных NCBI Nucleotide.

Работа была поддержана грантами РФФИ №16-34-00924, РНФ №14-14-00734. Вычисления проводились в Информационно-вычислительном центре НГУ

Источники и литература

- 1) 1. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications // BMC Bioinformatics. 2008. no. 10:421.
- 2) 2. Dobin A. et al. STAR: ultrafast universal RNA-seq aligner // Bioinformatics. 2013
- 3) 3. Grabherr M. et al. Full-length transcriptome assembly from RNA-seq data without a reference genome // Nature Biotechnology. 2011. no 29, 644-652
- 4) 4. Kim D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions // Genome Biology. 2013. 14:R3
- 5) 5. Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V, Dockter C, Hedley PE, Russell J et al. A chromosome conformation capture ordered sequence of the barley genome // Nature. 2017. no 544:427-433
- 6) 6. Mortazavi A., Williams B., McCue K., Schaeffer L., Wold B. Mapping and quantifying mammalian transcriptomes by RNA-seq // Nature Methods online. 2008. p 1-8
- 7) 7. Robinson M., McCarthy D. and Smyth G. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data // Bioinformatics. 2010. no. 26.
- 8) 8. Trapnell C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation // Nature Biotechnology. 2010. no. 28, 511–515
- 9) 9. Wu T.D., Watanabe C.K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences // Bioinformatics. 2005. V. 21-9, 1859–1875