

**БИОЛОГИЯ**  
**BIOLOGY**

УДК 577.214:004.9

Поступило в редакцию 03.02.2016  
Received 03.02.2016**Т. А. Подвицкий, Д. В. Галиновский, Н. В. Анисимова, академик Л. В. Хотылева,  
член-корреспондент А. В. Кильчевский***Институт генетики и цитологии НАН Беларуси, Минск, Республика Беларусь***ТРАНСКРИПТОМНЫЙ АНАЛИЗ ГЕНОВ ЦЕЛЛЮЛОЗОСИНТАЗ ВЫСШИХ  
РАСТЕНИЙ ПО ДАННЫМ РНК-СЕКВЕНИРОВАНИЯ**

В работе проанализированы транскриптомы различных органов и тканей четырех видов растений – тополя, эвкалипта, сои и фасоли. Установлены существенные различия в уровне экспрессии генов целлюлозосинтаз между растениями исследуемых видов, а также различия в экспрессии данных генов в органах и тканях одного растения. Для транскриптомов фасоли и эвкалипта характерно наличие «доминирующих» генов, которые дают более 50 % пула экспрессии генов целлюлозосинтаз. В случае тополя и сои экспрессионный пул распределен между большим числом генов. Возможно, стратегия экспрессии *CesA*-генов отражает процесс филогенеза исследуемых геномов.

*Ключевые слова:* целлюлозосинтаза, клеточная стенка, транскриптом, РНК-секвенирование.

**Ts. A. Padvitski, D. V. Galinousky, N. V. Anisimova, Academician L. V. Khotyleva,  
Corresponding Member A. V. Kilchevsky***Institute of Genetics and Cytology of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus***TRANSCRIPTOME ANALYZIS OF HIGHER PLANTS CELLULOSE SYNTHASE GENES  
USING RNA-SEQUENCING DATA**

In the given work the transcriptomes of different organs and tissues of western poplar, eucalyptus, soybean and common bean were studied. The expression level of cellulose synthase genes was notably different in the studied plant species and in different types of tissues and organs within plants. For common bean and eucalyptus transcriptome the domination of certain cellulose synthase genes was typical. These prevailing genes made up more than 50 % of the total expression pull of cellulose synthases. On the contrary, cellulose synthase expression pulls of western poplar and soybean were distributed between multiple genes. The different expression strategies of *CesA*-genes may reflect the phylogenetic process that occurred in the genomes studied.

*Keywords:* cellulose synthase, cell wall, transcriptome, RNA-sequencing.

**Введение.** В результате применения методов секвенирования нового поколения (NGS) были получены полногеномные нуклеотидные последовательности ряда высших растений – огурца (*Cucumis sativus* L.), томата (*Solanum lycopersicum* L.), кукурузы (*Zea mays* L.), льна (*Linum usitatissimum* L.) и мн. др. [1]. Следует отметить, что применение высокопроизводительных методов секвенирования не ограничивается определением нуклеотидной последовательности генома живых организмов. Одной из модификаций этого метода является РНК-секвенирование (RNA-seq), которое позволяет изучать транскриптом организма, ткани или клетки [2]. Изучение представленности транскриптов посредством анализа данных RNA-seq дополняет, а иногда и составляет альтернативу таким методам изучения транскриптома, как количественная ПЦР, Саузерн-гибридизация, микрочипы [3]. По оценкам специалистов, уже в ближайшие годы методы RNA-seq

станут рутинной процедурой, главным образом, за счет совершенствования и удешевления аппаратного обеспечения и упрощения пробоподготовки. Это, в свою очередь, многократно увеличит эмпирические данные и приведет к необходимости их качественного анализа и интерпретации [4].

С помощью RNA-seq и последующего анализа полученных данных уже сейчас решаются многие как фундаментальные, так и прикладные задачи. Так, объяснены и детализированы молекулярные механизмы, лежащие в основе ответа растения на стресс [5], а также открыты гены, ответственные за устойчивость растений к конкретным видам стресса [6]. Данные гены могут быть использованы для повышения устойчивости сельскохозяйственных культур к стрессовым факторам.

Методы генетического анализа, геномики и транскриптомики находят широкое применение при исследовании биосинтеза полисахаридов растительной клеточной стенки. Биосинтез целлюлозы (основной полисахарид клеточной стенки растений) – сложный биохимический процесс, который не удается воспроизвести *in vitro* [7]. Для его изучения перспективным представляется применение современных методов NGS, а именно, транскриптомного анализа генов целлюлозосинтаз (*CesA*-генов), который может дать косвенную, но ценную информацию о биогенезе растительной клеточной стенки. Доступность в on-line базах результатов RNA-seq различных культур [8] позволила нам провести биоинформационный анализ транскриптов *CesA*-генов ряда высших растений.

Цель работы – качественно и количественно охарактеризовать экспрессию генов целлюлозосинтаз в различных органах и тканях высших растений, используя данные RNA-seq.

**Материалы и методы исследования.** Для анализа экспрессии генов целлюлозосинтаз высших растений мы анализировали данные RNA-seq, к которым имеется открытый доступ на генетическом ресурсе Phytozome [8]. На время проведения анализа (10.10.2015) были доступны данные RNA-seq для шести серий экспериментов с разными видами и межвидовым гибридом растений, четыре из которых относятся к покрытосеменным, а именно – тополь (*Populus trichocarpa* Torr., A. Gray, 7 экспериментов), эвкалипт (гибрид *Eucalyptus grandis* × *Eucalyptus urophylla*, 6 экспери-

Т а б л и ц а 1. Вегетативные органы и ткани растений, которые использованы для RNA-seq [8]

Table 1. Vegetative organs and plant tissues used for RNA-seq [8]

Орган растения Plant organ	Название растения Plant			
	Тополь Poplar	Эвкалипт Eucalyptus	Соя Soybean	Фасоль Common bean
Корень Root	Верхушка корня Root tip Корень Root		Корень Root Корневые волоски Roots hairs nodules	Корень (10 дней) Root (10 days) Корень (19 дней) Root (19 days)
Стебель Stem	Узел Stem node Стебель Stem internode	Апикальная меристема стебля Shoot tips Флоэма Phloem Незрелая ксилема Immature xylem Зрелая ксилема Xylem	Апикальная меристема стебля Stem apical meristem Узел Node Стебель Stem	Узел Node Стебель (10 дней) Stem (10 days) Стебель (19 дней) Stem (19 days)
Лист Leaf	Зрелые листья Mature leaves Молодые листья Young leaves	Зрелые листья Mature leaves Молодые листья Young leaves	Листья Leaves	Листья Leaves Молодой трилистник Young trifoliates
Цветок Flower	Женская цветочная почка (ранняя стадия) Leaf female flower buds (early stage)		Цветок Flower	Цветочная почка Flower buds Цветок Flowers
Плод Fruit			Перикарпий плода Pod Семя Seed	Молодой боб Young pods Боб зеленой спелости Green mature pods

ментов), соя (*Glycine max* (L.) Merr., 9 экспериментов) и фасоль (*Phaseolus vulgaris* L., 11 экспериментов). Описание образцов, данные RNA-seq которых использованы в работе, приведено в табл. 1.

Для идентификации в геноме данных растений, а также в геноме *A. thaliana*, транскриптов, которые кодируют целлюлозосинтазы, использовали программу Biomart [9]. С помощью данной программы отбирали последовательности, в описании которых аннотирован белковый домен Pfam 03552 [10], характерный для суперсемейства целлюлозосинтаз, включающего *CesA*-гены и целлюлозосинтазоподобные гены. Последовательности генов с данным доменом из генома модельного организма *A. thaliana* использовались как референсные при реконструкции филогенетических отношений. Для выравнивания аминокислотных транслированных последовательностей использовали MUSCLE [11]. Для реконструкции филогенетических отношений всех найденных аминокислотных последовательностей использовали программу ClustalW2 Phylogeny [12] и метод кластеризации Neighbour-joining; филогению группы генов целлюлозосинтаз восстанавливали, используя программу PhyML с JTT-моделью аминокислотных замен [13].

Для анализа представленности транскриптов генов целлюлозосинтаз их Fpkm-значения были извлечены из базы данных Phytozome 10.3 (на 12.10.2015) с помощью интерфейса PhytoMine. Значение Fpkm является численным показателем количества «фрагментов на тысячу оснований на миллион картированных прочтений». Fpkm – относительная единица измерения экспрессии генов, которая используется при анализе результатов секвенирования парных прочтений и позволяет сравнивать результаты различных экспериментов [14]. Анализ данных и построение гистограмм осуществляли в R-среде.

Для оценки функциональной активности отдельных *CesA*-генов рассчитывали долю экспрессии индивидуальных генов от общего уровня экспрессии всего пула целлюлозосинтаз. Показатель «доля экспрессии» генов целлюлозосинтаз рассчитывали как частное между значением Fpkm целевого гена и Fpkm всего пула генов целлюлозосинтаз.

**Результаты и их обсуждение.** При анализе генома четырех исследуемых видов растений и генома резуховидки (*Arabidopsis thaliana* (L.) Heynh.) были идентифицированы гены, содержащие домен Pfam 03552. Всего в геномах данных растений идентифицировали 199 генов и, соответственно, 311 вариантов их транскриптов в составе транскриптомов. При помощи филогенетического анализа дифференцировали гены *CesA*-суперсемейства. Гены, транскрипты которых составляли единый кластер с целлюлозосинтазами *A. thaliana*, были отнесены к *CesA*-генам и далее проанализированы. В результате в геноме тополя обнаружили 17 генов целлюлозосинтаз, эвкалипта – 20, сои – 28 и в геноме фасоли – 15 *CesA*-генов (табл. 2).

RNA-seq эксперименты, анализируемые в данной работе, проводились разными исследователями в разное время и с разными целями, поэтому в них нет единого плана дифференциации тканей и органов растений.

По значению Fpkm для генов целлюлозосинтаз можно сделать вывод, что по уровню экспрессии данных генов исследованные растения существенно различаются между собой. Наиболее сильно *CesA*-гены экспрессируются в гибриде *E. grandis* × *E. urophylla*, слабее всего в растениях тополя. Общий уровень экспрессии генов целлюлозосинтаз в стебле эвкалипта (данные для ксилемы) в 8,2 раза выше, чем в стебле тополя, и в 2,98 раза выше, чем в стебле сои. В листьях эвкалипта – в 3,96 раза выше, чем в листьях тополя, и в 2,47 раз выше, чем в листьях сои.

Общий уровень экспрессии генов целлюлозосинтаз существенно различается в разных органах одного растения, хотя в пуле генов целлюлозосинтаз есть гены, экспрессия которых достаточно стабильна. У исследованных растений в стебле и корне гены целлюлозосинтаз экспрессируются мощнее, чем в листьях. Если сравнивать корень и стебель между собой, то у растений сои и фасоли показатель Fpkm, соответственно, в 1,24 и 1,49 раз выше для стебля, а у тополя в 1,25 раз выше для корня. Для эвкалипта такое сравнение провести не представляется возможным, поскольку отсутствуют данные для корня. При сравнимом общем уровне экспрессии *CesA*-генов в стеблях и корнях следует отметить разный профиль экспрессии *CesA*-генов в данных органах. Если в корне доминируют гены целлюлозосинтаз, ответственные за синтез первичной клеточной стенки (75 % и более от общего пула *CesA*-генов), а гены вторичной клеточной стенки в общем пуле транскриптов составляют не более 25 %, то в стебле доля генов вторичной клеточ-

Т а б л и ц а 2. Гены целлюлозосинтаз, аннотированные в исследуемых геномах

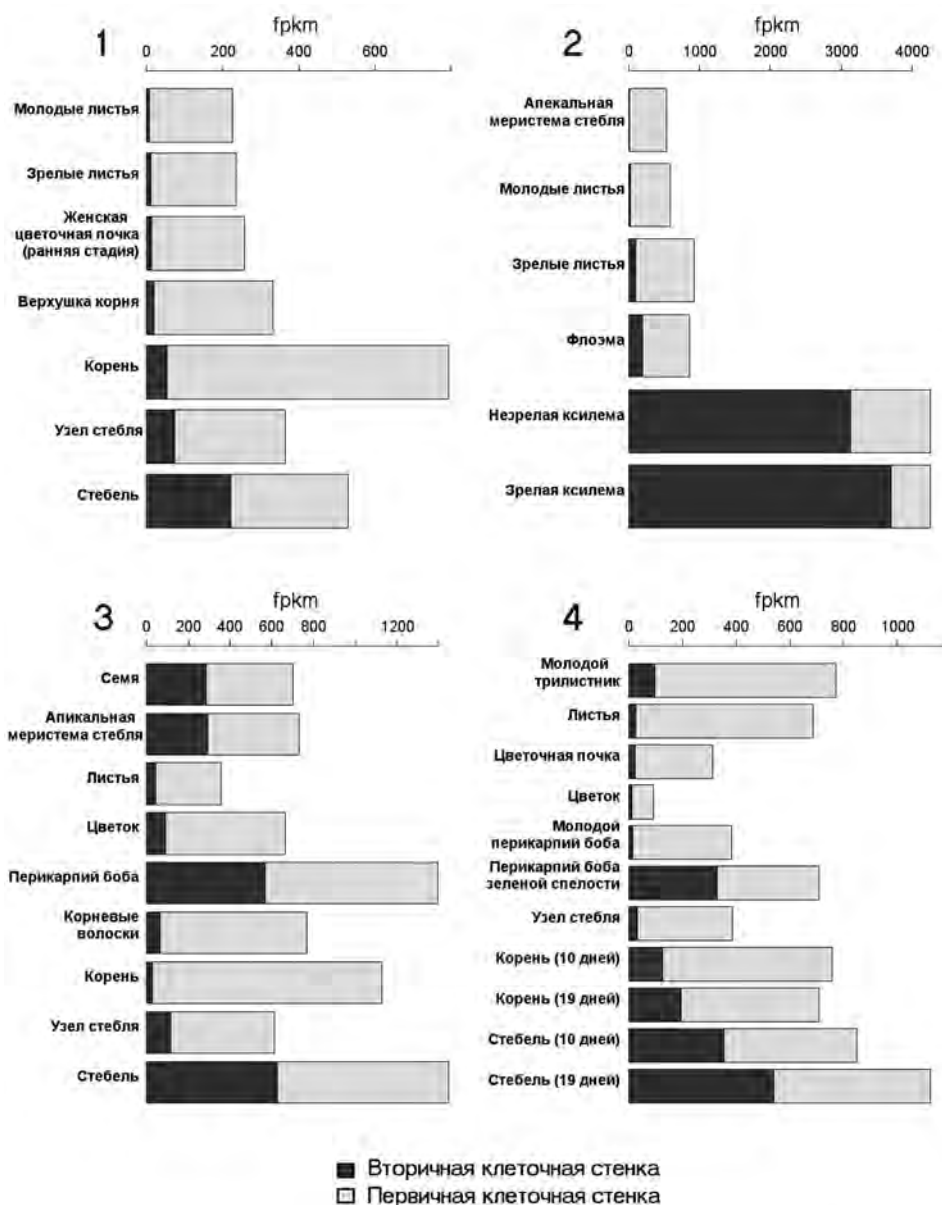
T a b l e 2. Cellulose synthase genes annotated in the genomes

Ген Gene	Код (ID) локуса генов целлюлозосинтаз Locus code (ID) of cellulose synthase genes				
	Резуховидка <i>Arabidopsis</i>	Тополь Poplar	Эвкалипт Eucalyptus	Соя Soybean	Фасоль Common bean
<i>Первичная клеточная стенка</i> <i>Primary cell wall</i>					
<i>CesA1/10</i>	AT2G25540 AT4G32410	Potri.018G029400 Potri.006G251900	Eucgr.C02801 Eucgr.C01769 Eucgr.H00939 Eucgr.J01639 Eucgr.L02402	Glyma.06G069600 Glyma.04G067900	Phvul.009G094200
<i>CesA2/5/6/9</i>	AT5G64740 AT5G09870 AT4G39350 AT2G21770	Potri.005G087500 Potri.007G076500 Potri.005G194200 Potri.002G066600 Potri.005G027600 Potri.013G019800	Eucgr.I00286 Eucgr.F04212 Eucgr.F04216 Eucgr.H02200 Eucgr.H00646 Eucgr.F03635 Eucgr.B01532 Eucgr.B01562 Eucgr.B03971	Glyma.05G187300 Glyma.08G145600 Glyma.02G080900 Glyma.16G165900 Glyma.10G223500 Glyma.16G081000 Glyma.06G316700	Phvul.002G268200 Phvul.004G093300 Phvul.007G081700 Phvul.005G010400
<i>CesA3</i>	AT5G05170	Potri.016G054900 Potri.006G052600 Potri.001G266400 Potri.009G060800	Eucgr.A02372 Eucgr.J01278 Eucgr.G03380	Glyma.09G103000 Glyma.15G275000 Glyma.13G202500 Glyma.12G237000	Phvul.005G022100 Phvul.011G211500
<i>Вторичная клеточная стенка</i> <i>Secondary cell wall</i>					
<i>CesA4</i>	AT5G44030	Potri.002G257900	Eucgr.A01324	Glyma.10G039600 Glyma.13G126000 Glyma.15G157100 Glyma.09G051100 Glyma.08G088400	Phvul.007G190300 Phvul.009G242700 Phvul.002G188600
<i>CesA7</i>	AT5G17420	Potri.018G103900 Potri.006G181900	Eucgr.D00476	Glyma.04G142700 Glyma.06G225400 Glyma.04G153700 Glyma.06G225500 Glyma.02G205800 Glyma.17G072200	Phvul.009G205200 Phvul.009G205100 Phvul.003G154600
<i>CesA8</i>	AT4G18780	Potri.011G069600 Potri.004G059600	Eucgr.C00246	Glyma.06G065000 Glyma.04G063800 Glyma.05G160000 Glyma.08G117500	Phvul.009G090100 Phvul.002G240200

ной стенки выше и составляет 42–47 %. Такое соотношение транскриптов наблюдали в структуре транскриптомов всех проанализированных растений. Даже в апикальной меристеме стебля наблюдали примерно такое же соотношение (61 % экспрессионного пула – гены целлюлозосинтаз первичной клеточной стенки и 39 % – вторичной, рисунок, 3). Следует отметить, что в узлах стебля наблюдали картину, отличную от остальной части стебля. Во-первых, общий уровень транскриптов *CesA*-генов ниже по сравнению с остальной частью стебля, а во-вторых, в пуле транскриптов *CesA*-генов низкая доля генов целлюлозосинтаз вторичной клеточной стенки – 11–18 % (рисунок, 1, 3, 4).

Самая высокая доля генов вторичной клеточной стенки из всех проанализированных экспериментов в ксилеме стебля у гибрида *E. grandis* × *E. urophylla*, здесь она составила 87 %, а уровень экспрессии генов целлюлозосинтаз в данных образцах был самым высоким из всех имеющихся.

Гены целлюлозосинтаз в листьях экспрессируются слабее среди всех вегетативных органов, в пуле *CesA*-генов доминируют гены целлюлозосинтаз первичной клеточной стенки, гены вторичной клеточной стенки составляют не более 14 % (рисунок).



Уровень экспрессии генов целлюлозосинтаз, отвечающих за биогенез первичной и вторичной клеточной стенки тополя (1), эвкалипта (2), сои (3) и фасоли (4)

The expression level of cellulose synthase genes responsible for the biogenesis of the primary and secondary cell wall of poplar (1), eucalyptus (2), soybean (3) and common bean (4)

Из полученных данных можно заключить, что в большинстве случаев экспрессия генов целлюлозосинтаз первичной клеточной стенки доминирует над генами вторичной. Есть образцы, где представленность транскриптов целлюлозосинтаз первичной и вторичной клеточной стенки примерно равна. Такое соотношение установлено для образцов стеблей растений и перикарпия бобов (для фасоли и сои). Только в одном случае был обнаружен паттерн экспрессии *CesA*-генов, в котором гены целлюлозосинтаз вторичной клеточной стенки доминируют над первичной – это образцы ксилемы (зрелой и незрелой) эвкалипта (рисунок, 2).

Видимо, в большинстве клеток растительного организма функционируют гены целлюлозосинтаз, связанные с биогенезом первичной клеточной стенки. *CesA*-гены вторичной клеточной стенки экспрессируются в клетках стебля, а в случае бобовых – в клетках семян и околоплодника. Усиление экспрессии генов вторичной клеточной стенки проявляется, возможно, в тех случаях, где необходимо обеспечение механической функции. Доминирование экспрессии генов вто-



ричной клеточной стенки в общем пуле генов целлюлозосинтаз, видимо, свидетельствует о специализации ткани.

Для каждого из растений имеются гены целлюлозосинтаз с низкой функциональной активностью, доля транскриптов которых в любом из вариантов транскриптомов не превышала 5 %. В случае гибрида *E. grandis* × *E. urophylla* было обнаружено 11 таких генов (Eucgr.H00939, Eucgr.J01639, Eucgr.L02402, Eucgr.F04212, Eucgr.F04216, Eucgr.H02200, Eucgr.H00646, Eucgr.B01532, Eucgr.B01562, Eucgr.B03971 и Eucgr.A02372), для *Phaseolus vulgaris* – 6 (Phvul.007G081700, Phvul.005G01040, Phvul.007G190300, Phvul.002G188600, Phvul.009G205200 и Phvul.002G240200), для *Glycine max* – 7 (Glyma.16G081000, Glyma.06G316700, Glyma.10G039600, Glyma.13G126000, Glyma.08G088400, Glyma.04G142700 и Glyma.05G160000) и один для *Populus trichocarpa* (Potri.006G052600). Транскрипты данных генов не учитывали в последующем анализе. Значения Fpkm для остальных генов целлюлозосинтаз приведены в табл. 3–6.

Среди генов целлюлозосинтаз первичной клеточной стенки тополя невозможно выделить однозначных «лидеров экспрессии». В большинстве случаев доля каждого гена составляла от 5 до 14 %, за исключением последовательности Potri.007G076500 (ген *CesA2/5/6/9*), которая в двух пробах несколько превышала 15 %, и последовательностей Potri.002G066600 и Potri.013G019800, доля которых (кроме транскрипта корня) составляла менее 5 % (табл. 3).

Ряд генов целлюлозосинтаз первичной клеточной стенки отличался достаточно стабильной экспрессией в надземных частях растений. Уровень экспрессии генов Potri.007G076500 (*CesA2/5/6/9*), Potri.005G194200 (*CesA2/5/6/9*), Potri.005G027600 (*CesA2/5/6/9*), Potri.018G029400 (*CesA1/10*), Potri.006G251900 (*CesA1/10*) и Potri.009G060800 (*CesA3*) в таких пробах различался в 1,5–1,8 раза (табл. 3).

Как уже указывалось выше, гены целлюлозосинтаз вторичной клеточной стенки экспрессировались в стебле, и среди них более 15 % пула составляли транскрипты гена Potri.002G257900, который экспрессировался в 1,7–2,2 раз сильнее, чем остальные *CesA*-гены вторичной клеточной стенки.

Т а б л и ц а 3. Экспрессия (Fpkm) отдельных генов целлюлозосинтаз тополя  
T a b l e 3. Expression (Fpkm) of individual cellulose synthase genes of poplar

Название гена Gene	ID гена Gene ID	Молодые листья Young leaves	Листья Leaves	Женская цветочная почка (ранняя стадия) Female flower bud (early stage)	Верхушка корня Root tip	Корень Root	Узел стебля Stem node	Стебель Stem internode
<i>Гены первичной клеточной стенки</i> <i>Primary cell wall genes</i>								
<i>CesA2/5/6/9</i>	Potri.005G087500	17	16	15	23	14	27	35
<i>CesA2/5/6/9</i>	Potri.007G076500	34	33	41	49	64	40	42
<i>CesA2/5/6/9</i>	Potri.005G194200	27	23	26	41	88	35	40
<i>CesA2/5/6/9</i>	Potri.002G066600	9	9	9	14	79	18	16
<i>CesA2/5/6/9</i>	Potri.005G027600	13	15	17	19	80	16	13
<i>CesA2/5/6/9</i>	Potri.013G019800	8	7	9	10	59	13	15
<i>CesA1/10</i>	Potri.018G029400	21	26	29	33	103	24	26
<i>CesA1/10</i>	Potri.006G251900	28	30	31	41	63	33	34
<i>CesA3</i>	Potri.016G054900	16	15	11	20	37	24	22
<i>CesA3</i>	Potri.001G266400	13	13	15	19	27	16	17
<i>CesA3</i>	Potri.009G060800	22	29	35	35	105	29	32
<i>Гены вторичной клеточной стенки</i> <i>Secondary cell wall genes</i>								
<i>CesA4</i>	Potri.002G257900	3	4	4	7	18	28	82
<i>CesA8</i>	Potri.004G059600	1	1	2	2	9	13	39
<i>CesA7</i>	Potri.018G103900	2	3	3	5	11	12	37
<i>CesA7</i>	Potri.006G181900	2	3	3	3	10	15	48

П р и м е ч а н и я: ■ – значения Fpkm более 15 % от общего пула *CesA*-генов; ■ – значения Fpkm от 5 до 14 % от общего пула *CesA*-генов; □ – значения Fpkm менее 5 % от общего пула *CesA*-генов.

N o t e s: ■ – Fpkm values are more than 15 % of the total pull of *CesA*-genes; ■ – Fpkm values range from 5 to 14 % of the total pull of *CesA*-genes; □ – Fpkm values are less than 5 % of the total pull of *CesA*-genes.

Среди генов целлюлозосинтаз эвкалипта, отвечающих за синтез первичной клеточной стенки, функционируют четыре гена – Eucgr.I00286 (*CesA2/5/6/9*), Eucgr.F03635 (*CesA2/5/6/9*), Eucgr.C02801 (*CesA1/10*) и Eucgr.G03380 (*CesA3*), причем данные гены экспрессируются достаточно стабильно в разных типах органов и тканей (разница в уровне экспрессии 2,2–3,8 раза). Наиболее стабильно экспрессируется ген Eucgr.I00286, разница экспрессии которого между минимальным значением (листья) и максимальным (незрелая ксилема) составила 2,2 раза (табл. 4). Возможно, такие результаты связаны с отсутствием данных для корня, цветков и плодов, где может меняться профиль экспрессии генов целлюлозосинтаз.

Гены целлюлозосинтаз вторичной клеточной стенки наиболее активно экспрессируются в ксилеме, где функционирует три гена – Eucgr.A01324 (*CesA4*), Eucgr.D00476 (*CesA7*) и Eucgr.C00246 (*CesA8*), последний из которых доминирует (его экспрессия в 3,9 раза выше по сравнению с Eucgr.A01324 и в 3,1 раза – с Eucgr.D00476). Соотношение транскриптов генов *CesA4* : *CesA7* : *CesA8* во флоэме и ксилеме стебля примерно 1 : 1 : 4, что может указывать на соотношение целлюлозосинтазных субъединиц в трансмембранном целлюлозосинтезирующем комплексе. Мы полагаем, что растения эвкалипта могут быть удобной моделью для изучения экспрессии генов биосинтеза целлюлозы (а, возможно, и других полимеров клеточной стенки) ввиду малого количества функционирующих генов и достаточно четких тканеспецифичных изменений в паттерне экспрессии генов.

Т а б л и ц а 4. Экспрессия (Fpkm) отдельных генов целлюлозосинтаз эвкалипта

Table 4. Expression (Fpkm) of individual cellulose synthase genes of eucalyptus

Название гена Gene	ID гена Gene ID	Апикальная меристема стебля Shoot tips	Молодые листья Young leaves	Зрелые листья Mature leaves	Флоэма Phloema	Незрелая ксилема Immature xylem	Ксилема Xylem
<i>Гены первичной клеточной стенки</i> <i>Primary cell wall genes</i>							
<i>CesA2/5/6/9</i>	Eucgr.I00286	134	145	116	187	254	163
<i>CesA2/5/6/9</i>	Eucgr.F03635	75	96	178	80	178	47
<i>CesA1/10</i>	Eucgr.C02801	138	174	309	134	270	104
<i>CesA3</i>	Eucgr.G03380	132	122	163	183	347	156
<i>Гены вторичной клеточной стенки</i> <i>Secondary cell wall genes</i>							
<i>CesA4</i>	Eucgr.A01324	7	9	24	34	506	596
<i>CesA7</i>	Eucgr.D00476	3	4	22	40	641	757
<i>CesA8</i>	Eucgr.C00246	4	12	52	122	1983	2352

Пр и м е ч а н и я: ■ – значения Fpkm более 15 % от общего пула *CesA*-генов; ■ – значения Fpkm от 5 до 14 % от общего пула *CesA*-генов; □ – значения Fpkm менее 5 % от общего пула *CesA*-генов.

Notes: ■ – Fpkm values are more than 15 % of the total pull of *CesA*-genes; ■ – Fpkm values range from 5 to 14 % of the total pull of *CesA*-genes; □ – Fpkm values are less than 5 % of the total pull of *CesA*-genes.

Экспрессионный профиль генов целлюлозосинтаз сои оказался наиболее сложным для интерпретации (табл. 5). Из его особенностей можно отметить органоспецифичный ген Glyma.10G223500 (*CesA2/5/6/9*), который функционировал в цветке (его транскрипты составляли 49 % от общего пула *CesA*-генов) и более нигде не проявлялся (табл. 5).

Среди генов целлюлозосинтаз первичной клеточной стенки только для генов Glyma.06G069600 (*CesA1/10*), Glyma.12G237000 (*CesA3*) и Glyma.16G165900 (*CesA2/5/6/9*) были установлены случаи, когда доля их экспрессии превышала 15 % в общем пуле генов. Для гена Glyma.06G069600 это были транскрипты листьев, корня, корневых волосков и узлов стебля, для гена Glyma.12G237000 – листьев и семян, для Glyma.16G165900 – корня. Большинство же генов вносили не более 14 % экспрессионного пула. Возможно, такое обстоятельство связано с эволюционно недавним удвоением генома сои [15], что привело к удвоению функционирующих генов и, как следствие, снижению доли транскриптов каждого из них.

Во всех рассмотренных экспериментах каждый из *CesA*-генов вторичной клеточной стенки вносил менее 15 % в общий экспрессионный пул. Максимальную долю (12 %) в общий пул целлюлозосинтаз вносил ген Glyma.09G051100 (*CesA4*) при экспрессии в перикарпиях бобов. Среди

Таблица 5. Экспрессия (Fpkm) отдельных генов целлюлозысинтазы сои  
 Table 5. Expression (Fpkm) of individual cellulose synthase genes of soybean

Название гена Gene	ID гена Gene ID	Семя Seed	Апикальная меристема стебля Stem apical meristem	Листья Leaves	Цветок Flower	Перикарпий боба Pod	Корневые волоски Root hairs	Корень Root	Узел стебля Stem modules	Стебель Stem
<i>Гены первичной клеточной стенки</i> <i>Primary cell wall genes</i>										
<i>CesA2/5/6/9</i>	Glyma.05G187300	45	56	34	25	88	92	98	66	101
<i>CesA2/5/6/9</i>	Glyma.08G145600	29	33	21	10	56	50	52	26	65
<i>CesA2/5/6/9</i>	Glyma.02G080900	21	20	24	5	52	36	139	45	84
<i>CesA2/5/6/9</i>	Glyma.16G165900	48	32	36	5	98	62	228	50	88
<i>CesA2/5/6/9</i>	Glyma.10G2233500	12	33	2	327	28	18	32	7	44
<i>CesA1/10</i>	Glyma.06G069600	51	79	56	37	164	159	179	99	157
<i>CesA1/10</i>	Glyma.04G067900	34	42	41	29	82	65	123	49	72
<i>CesA3</i>	Glyma.09G103000	15	26	6	6	47	107	151	54	64
<i>CesA3</i>	Glyma.15G275000	14	25	6	14	40	70	78	49	55
<i>CesA3</i>	Glyma.13G202500	40	32	21	52	44	12	6	12	23
<i>CesA3</i>	Glyma.12G237000	107	59	63	65	127	32	11	37	65
<i>Гены вторичной клеточной стенки</i> <i>Secondary cell wall genes</i>										
<i>CesA4</i>	Glyma.09G051100	70	59	9	6	169	15	4	24	133
<i>CesA8</i>	Glyma.06G065000	34	16	7	2	77	7	2	15	66
<i>CesA8</i>	Glyma.08G117500	17	69	0	2	1	3	5	4	37
<i>CesA7</i>	Glyma.06G225500	82	83	8	43	90	9	8	20	107

Примечания: ■ – значения Fpkm более 15 % от общего пула *CesA*-генов; ■ – значения Fpkm от 5 до 14 % от общего пула *CesA*-генов; □ – значения Fpkm менее 5 % от общего пула *CesA*-генов.

Notes: ■ – Fpkm values are more than 15 % of the total pull of *CesA*-genes; ■ – Fpkm values range from 5 to 14 % of the total pull of *CesA*-genes; □ – Fpkm values are less than 5 % of the total pull of *CesA*-genes.



Таблица 6. Экспрессия (Fpkm) отдельных генов целлюлозосинтаз фасоли  
 Table 6. Expression (Fpkm) of individual cellulose synthase genes of common bean

Название гена Gene	ID гена Gene ID	Молодой трилистник Young trifoliates	Листья Leaves	Цветочная почка Flower buds	Цветок Flowers	Перикарпий молодого боба Young pods	Перикарпий боба зеленой спелости Pericarp of green mature pods	Узел стебля Stem nodules	Корень (10 дн.) Root (10 days)	Корень (19 дн.) Root (19 days)	Стебель (10 дн.) Stem (10 days)	Стебель (19 дн.) Stem (19 days)
<i>Гены первичной клеточной стенки Primary cell wall genes</i>												
CesA2/5/6/9	Phvul.002G268200	197	191	58	11	81	61	75	151	106	102	113
CesA2/5/6/9	Phvul.004G093300	72	48	30	16	42	32	38	107	88	80	119
CesA1/10	Phvul.009G094200	204	213	115	25	128	151	111	166	150	156	185
CesA3	Phvul.005G022100	124	105	60	15	82	62	84	181	132	121	118
CesA3	Phvul.011G211500	66	88	24	13	29	68	44	22	30	28	38
<i>Гены вторичной клеточной стенки Secondary cell wall genes</i>												
CesA4	Phvul.009G242700	49	12	10	6	5	138	13	58	82	160	231
CesA8	Phvul.009G090100	18	6	5	2	2	45	5	22	34	66	96
CesA7	Phvul.009G205100	24	5	5	3	2	81	5	16	25	46	76
CesA7	Phvul.003G154600	3	2	3	1	5	64	5	20	30	73	90

Примечания: ■ — значения Fpkm более 15 % от общего пула CesA-генов; ■ — значения Fpkm от 5 до 14 % от общего пула CesA-генов; □ — значения Fpkm менее 5 % от общего пула CesA-генов.

Notes: ■ — Fpkm values are more than 15 % of the total pull of CesA-genes; ■ — Fpkm values range from 5 to 14 % of the total pull of CesA-genes; □ — Fpkm values are less than 5 % of the total pull of CesA-genes.

*CesA*-генов вторичной клеточной стенки этот же ген наиболее интенсивно экспрессировался в стебле сои (9 %) наряду с геном Glyma.06G225500 (*CesA7*) – 7 %.

При анализе данных RNA-seq *Phaseolus vulgaris* установили, что из 15 генов целлюлозосинтаз, имеющих в геноме, функционирует 9, из которых 5 – *CesA*-гены первичной клеточной стенки и 4 – вторичной (табл. 2, 6).

Среди генов целлюлозосинтаз первичной клеточной стенки наиболее сильно экспрессировались три гена – Phvul.002G268200 (*CesA2/5/6/9*), Phvul.009G094200 (*CesA1/10*) и Phvul.005G022100 (*CesA3*). Экспрессия данных генов достаточно стабильная во всех вегетативных органах (включая и корень) и различается в разных образцах в 1,9–2,6 раза. В цветках снижается общий уровень экспрессии генов целлюлозосинтаз и перераспределяются доли отдельных генов. При созревании плодов, в перикарпиях бобов, а также в стебле увеличивается экспрессия генов целлюлозосинтаз вторичной клеточной стенки, среди которых лидирует ген Phvul.009G242700 (*CesA4*). Данный ген также функционирует и в корнях, где представленность его транскриптов доходит до 12 % от общего пула генов целлюлозосинтаз (табл. 6).

При анализе профилей экспрессии *CesA*-генов можно уловить две альтернативные модели их функционирования, которые наиболее четко проявляются при экспрессии генов целлюлозосинтаз первичной клеточной стенки. Первая модель характеризуется наличием доминирующих генов целлюлозосинтаз (эвкалипт и фасоль), доля экспрессии каждого из которых может достигать 15 % и более. В таком случае экспрессия трех генов дает более 50 % пула экспрессии *CesA*-генов и можно говорить о «лидерах экспрессии». Во второй альтернативной модели нет «лидеров экспрессии», экспрессионный пул *CesA*-генов разделен между 8–10 генами, каждый из которых имеет менее 15 % в общем пуле. Возможно, этот факт отражает филогенетическую историю вида, а именно – события удвоения генома с сохранением функциональной активности удвоенных копий генов. В таком случае модель экспрессии *CesA*-генов связана с наличием или отсутствием событий полиплоидизации в процессе генезиса данного вида растений.

**Заключение.** Таким образом, в работе проанализированы транскриптомы различных органов и тканей четырех растений – тополя (*Populus trichocarpa*), эвкалипта (гибрид *E. grandis* × *E. urophylla*), сои (*Glycine max*) и фасоли (*Phaseolus vulgaris*). Установлено, что по уровню экспрессии генов целлюлозосинтаз исследованные растения существенно различаются между собой. Наиболее сильно *CesA*-гены экспрессируются в растениях эвкалипта, слабее всего – тополя. Также уровень экспрессии данных генов различается в органах и тканях одного растения, в стебле и корне гены целлюлозосинтаз экспрессируются сильнее, чем в листьях. Кроме количественных в органах растений отмечены и качественные различия в профиле экспрессии *CesA*-генов. В листьях и корне доля генов целлюлозосинтаз, ответственных за синтез вторичной клеточной стенки, составляет не более 25 % от общего пула *CesA*-генов, а в стебле – 42 % и более.

Транскрипционные профили *CesA*-генов растительных органов и тканей можно разделить на три группы. В первую группу включены образцы, в которых транскрипты генов целлюлозосинтаз, отвечающих за синтез первичной клеточной стенки, существенно преобладают (75 % и более) над уровнем транскриптов вторичной клеточной стенки. Такие транскрипционные профили характерны для листьев, корней, цветков. Вторую группу составляют образцы с примерно равными уровнями экспрессии генов вторичной (40–50 %) и первичной (50–60 %) клеточной стенки. Это соотношение установлено для образцов стеблей растений и перикарпия бобов (для фасоли и сои). В третью группу входят образцы, в которых транскрипты генов вторичной клеточной стенки преобладали (73 % и более) над первичными (ксилема эвкалипта).

Мы полагаем, что растения эвкалипта могут быть удобной моделью для изучения экспрессии генов биосинтеза целлюлозы (а, возможно, и других полимеров клеточной стенки) ввиду малого количества функционирующих генов и достаточно четких тканеспецифичных изменений в паттерне экспрессии генов.

При оценке функциональной активности генов целлюлозосинтаз установили альтернативные стратегии функционирования генов целлюлозосинтаз. Для транскриптомов фасоли и эвкалипта характерно наличие «доминирующих» генов, которые дают более 50 % пула экспрессии генов целлюлозосинтаз. В случае тополя и сои экспрессионный пул распределен между боль-

шим числом генов, на каждый из которых приходится менее 15 % от общего числа транскриптов. Возможно, стратегия экспрессии *CesA*-генов отражает процесс филогенеза исследуемых геномов, а именно – сравнительно недавнюю их полиплоидизацию, которая привела к увеличению копий функционирующих генов и снижению доли транскриптов каждого из них.

**Благодарности.** Работа выполнена при финансовой поддержке НАН Беларуси (договор № Б15УК/А-041) и Белорусского республиканского фонда фундаментальных исследований (проект Б15М-101).

**Acknowledgement.** The work was sponsored by the National Academy of Sciences of Belarus (Agreement no. Б15УК/А-041) and Belarusian Republican Foundation for Fundamental Research (Agreement no. Б15М-101).

### СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. JGI Phytozome 10.3 [Electronic resource]. – Mode of access: <http://phytozome.jgi.doe.gov>. – Date of access: 12.10.2015.
2. Wang, Z. RNA-Seq: a revolutionary tool for transcriptomics / Z. Wang, M. Gerstein, M. Snyder // *Nature reviews genetics*. – 2009. – Vol. 10, N 1. – P. 57–63.
3. Giorgi, F. M. Comparative study of RNA-seq- and microarray-derived coexpression networks in *Arabidopsis thaliana* / F. M. Giorgi, C. Del Fabbro, F. Licausi // *Bioinformatics*. – 2013. – Vol. 29, N 6. – P. 717–724.
4. Chu, Y. RNA Sequencing: Platform Selection, Experimental Design, and Data Interpretation / Y. Chu, D. R. Corey // *Nucleic Acid Therapeutics*. – 2012. – Vol. 22, N 4. – P. 271–274.
5. Wakasa, Y. RNA sequencing-mediated transcriptome analysis of rice plants in endoplasmic reticulum stress conditions / Y. Wakasa, Y. Oono, Sh. Yazawa // *BMC Plant Biology*. – 2014. – Vol. 14. – P. 101.
6. RNA-Seq Analysis of the Response of the Halophyte, *Mesembryanthemum crystallinum* (Ice Plant) to High Salinity / H. Tsukagoshi [et al.] // *PLoS ONE*. – 2015. – Vol. 10, N 2. – e0118339. doi.org/10.1371/journal.pone.0118339.
7. Горшкова, Т. А. Растительная клеточная стенка как динамическая система / Т. А. Горшкова. – Москва, 2007. – 429 с.
8. RNA-Seq data from Phytozome [Electronic resource]. – Mode of access: <http://phytozome.jgi.doe.gov/phytomine/aspect.do?name=Expression>. – Date of access: 12.10.2015.
9. Biomart [Electronic resource]. – Mode of access: <http://www.biomart.org/>. – Date of access: 12.10.2015.
10. EMBL-EBL: Family: Cellulose\_synt (PF03552) [Electronic resource]. – Mode of access: [http://pfam.xfam.org/family/Cellulose\\_synt#tabview=tab0](http://pfam.xfam.org/family/Cellulose_synt#tabview=tab0). – Date of access: 12.10.2015.
11. EMBL-EBL: MUSCLE. Multiple Sequence Alignment [Electronic resource]. – Mode of access: <https://www.ebi.ac.uk/Tools/msa/muscle/>. – Date of access: 12.10.2015.
12. EMBL-EBL: ClustalW2 – Phylogeny [Electronic resource]. – Mode of access: [http://www.ebi.ac.uk/Tools/phylogeny/clustalw2\\_phylogeny/](http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/). – Date of access: 12.10.2015.
13. ATGC: South of France bioinformatics platform [Electronic resource]. – Mode of access: <http://www.atgc-montpellier.fr/phyml/>. – Date of access: 12.10.2015.
14. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation / C. Trapnell [et al.] // *Nat. Biotechnol.* – 2010. – Vol. 28, N 5. – P. 511–515. doi.org/10.1038/nbt.1621.
15. Genome sequence of the palaeopolyploid soybean / J. Schmutz [et al.] // *Nature*. – 2010. – Vol. 463. – P. 178–183. doi.org/10.1038/nature08670.

### References

1. JGI Phytozome 10.3. Available at: <http://phytozome.jgi.doe.gov> (accessed 12 October 2015).
2. Wang Z., Gerstein M., Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews genetics*, 2009, vol. 10, no. 1, pp. 57–63. doi: 10.1038/nrg2484.
3. Giorgi F. M., Del Fabbro C., Licausi F. Comparative study of RNA-seq- and microarray-derived coexpression networks in *Arabidopsis thaliana*. *Bioinformatics*, 2013, vol. 29, no. 6, pp. 717–724. doi: 10.1093/bioinformatics/btt053.
4. Chu Y., Corey D. R. RNA Sequencing: Platform Selection, Experimental Design, and Data Interpretation. *Nucleic Acid Therapeutics*, 2012, vol. 22, no. 4, pp. 271–274. doi: 10.1089/nat.2012.0367.
5. Wakasa Y., Oono Y., Yazawa Sh. RNA sequencing-mediated transcriptome analysis of rice plants in endoplasmic reticulum stress conditions. *BMC Plant Biology*, 2014, vol. 14, no. 1, pp. 101. doi: 10.1186/1471-2229-14-101.
6. Tsukagoshi H., Suzuki T., Nishikawa K., Agarie S., Ishiguro S., Higashiyama T. RNA-Seq Analysis of the Response of the Halophyte, *Mesembryanthemum crystallinum* (Ice Plant) to High Salinity. *PLoS ONE*, 2015, vol. 10, no. 2, e0118339. doi:10.1371/journal.pone.0118339.
7. Gorshkova T. A. *Plant cell wall as a dynamic system*. Moscow, 2007. 429 p. (in Russian)
8. RNA-Seq data from Phytozome. Available at: <http://phytozome.jgi.doe.gov/phytomine/aspect.do?name=Expression> (accessed 12 October 2015).
9. Biomart. Available at: <http://www.biomart.org/> (accessed 12 October 2015).
10. EMBL-EBL: Family: Cellulose\_synt (PF03552). Available at: [http://pfam.xfam.org/family/Cellulose\\_synt#tabview=tab0](http://pfam.xfam.org/family/Cellulose_synt#tabview=tab0) (accessed 12 October 2015).
11. EMBL-EBL: MUSCLE. Multiple Sequence Alignment. Available at: <https://www.ebi.ac.uk/Tools/msa/muscle/> (accessed 12 October 2015).

12. EMBL-EBL: ClustalW2 – Phylogeny. Available at: [http://www.ebi.ac.uk/Tools/phylogeny/clustalw2\\_phylogeny/](http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/) (accessed 12 October 2015).

13. ATGC: South of France bioinformatics platform. Available at: <http://www.atgc-montpellier.fr/phyml/> (accessed 12 October 2015).

14. Trapnell C., Williams B. A., Pertea G., Mortazavi A., Kwan G., van Baren M. J., Salzberg S. L., Wold B. J., Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, 2010, vol. 28, no. 5, pp. 511–515. doi: 10.1038/nbt.1621.

15. Schmutz J., Cannon S. B., Schlueter J., Ma J., Mitros T., Nelson W., Hyten D. L., Song Q., Thelen J. J., Cheng J., Xu D., Hellsten U., May G. D., Yu Y., Sakurai T., Umezawa T., Bhattacharyya M. K., Sandhu D., Valliyodan B., Lindquist E., Peto M., Grant D., Shu S., Goodstein D., Barry K., Futrell-Griggs M., Abernathy B., Du J., Tian Z., Zhu L., Gill N., Joshi T., Libault M., Sethuraman A., Zhang X.-C., Shinozaki K., Nguyen H. T., Wing R. A., Cregan P., Specht J., Grimwood J., Rokhsar D., Stacey G., Shoemaker R. C., Jackson S. A. Genome sequence of the palaeopolyploid soybean. *Nature*, 2010, vol. 463, pp. 178–183. doi:10.1038/nature08670.

### Информация об авторах

*Подвицкий Тимофей Александрович* – магистр биоинформатики, магистр биологических наук, мл. науч. сотрудник, Институт генетики и цитологии НАН Беларуси (ул. Академическая, 27, 220072, Минск, Республика Беларусь). E-mail: tsimafeipadvitski@gmail.com.

*Галиновский Дмитрий Валентинович* – канд. биол. наук, науч. сотрудник, Институт генетики и цитологии НАН Беларуси (ул. Академическая, 27, 220072, Минск, Республика Беларусь). E-mail: D.Galinousky@igc.by.

*Анисимова Наталья Владимировна* – канд. биол. наук, ст. науч. сотрудник, Институт генетики и цитологии НАН Беларуси (ул. Академическая, 27, 220072, Минск, Республика Беларусь). E-mail: N.Anisimova@igc.by.

*Хотылева Любовь Владимировна* – академик, д-р биол. наук, профессор, гл. науч. сотрудник, Институт генетики и цитологии НАН Беларуси (ул. Академическая, 27, 220072, Минск, Республика Беларусь). E-mail: L.Khotyleva@igc.by.

*Кильчевский Александр Владимирович* – член-корреспондент, д-р биол. наук, профессор, заведующий лабораторией, Институт генетики и цитологии НАН Беларуси (ул. Академическая, 27, 220072, Минск, Республика Беларусь). E-mail: kilchev@presidium.bas-net.by.

### Для цитирования

Транскриптомный анализ генов целлюлозосинтазы высших растений по данным РНК-секвенирования / Т. А. Подвицкий [и др.] // Докл. НАН Беларуси. – 2016. – Т. 60, № 6. – С. 72–83.

### Information about the author

*Padvitski Tsimafei Alyksandravich* – Master of Bioinformatics, Master of Biology, Junior researcher, Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Akademicheskaya Str., 220072, Minsk, Republic of Belarus). E-mail: tsimafeipadvitski@gmail.com.

*Galinousky Dmitry Valentinovich* – Ph. D. (Biology), Researcher, Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Akademicheskaya Str., 220072, Minsk, Republic of Belarus). E-mail: D.Galinousky@igc.by.

*Anisimova Natallia Vladimirovna* – Ph. D. (Biology), Senior researcher, Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Akademicheskaya Str., 220072, Minsk, Republic of Belarus). E-mail: N.Anisimova@igc.by.

*Khotyleva Liubov Vladimirovna* – Academician, D. Sc. (Biology), Professor, Chief researcher, Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Akademicheskaya Str., 220072, Minsk, Republic of Belarus). E-mail: L.Khotyleva@igc.by.

*Kilchevsky Alexander Vladimirovich* – Corresponding Member, D. Sc. (Biology), Professor, Head of Laboratory, Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Akademicheskaya Str., 220072, Minsk, Republic of Belarus). E-mail: kilchev@presidium.bas-net.by.

### For citation

Padvitski Ts. A., Galinousky D. V., Anisimova N. V., Khotyleva L. V., Kilchevsky A. V. Transcriptome analysis of higher plants cellulose synthase genes using RNA-sequencing data. *Doklady Natsional'noi akademii nauk Belarusi* [Doklady of the National Academy of Sciences of Belarus], 2016, vol. 60, no. 6, pp. 72–83. (in Russian)