

АНАЛИЗ СТРАТЕГИЙ ЭКСПРЕССИИ ГЕНОВ ЦЕЛЛЮЛОЗОСИНТАЗ ВЫСШИМИ РАСТЕНИЯМИ НА ОСНОВЕ ДАННЫХ РНК-СЕКВЕНИРОВАНИЯ

Т.А. ПОДВИЦКИЙ¹, Д.В. ГАЛИНОВСКИЙ^{1*}, Н.В. АНИСИМОВА¹, Г.Я. БАЕР², Я.В. ПИРКО², А.И. ЕМЕЦ²,
Л.В. ХОТЫЛЕВА¹, Я.Б. БЛЮМ², А.В. КИЛЬЧЕВСКИЙ¹

¹ Институт генетики и цитологии НАН Беларуси, Минск

² Институт пищевой биотехнологии и геномики НАН Украины, Киев

* E-mail-galinousky@gmail.com

Проанализированы транскрипты различных органов и тканей четырех видов растений — тополя, эвкалипта, сои и фасоли. Установлены существенные различия уровня экспрессии генов целлюлозосинтаз 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. Вероятно, стратегия экспрессии CesaA-генов отражает процесс филогенеза исследуемых геномов.

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

Введение. В результате применения методов секвенирования нового поколения к настоящему времени получены полногеномные нуклеотидные последовательности огурца (*Cucumis sativus* L.), томата (*Solanum lycopersicum* L.), кукурузы (*Zea mays* L.), льна (*Linum usitatissimum* L.) и многих других видов высших растений (<http://phytozome.jgi.doe.gov>). Следует отметить, что применение высокопроизводительных методов секвенирования не ограничивается определением нуклеотидной последовательности генома живых организмов. Одной из модификаций этого метода является РНК-секвенирование, которое позволяет изучать транскриптом организма, ткани или клетки [1–3]. Изучение представленности транскриптов посредством анализа данных РНК-секвенирования дополняет, а иногда и составляет альтернативу таким методам изучения транскриптома, как количественная ПЦР, Саузерн-гибридизация, микрочипы [4]. По оценкам специалистов уже в ближайшие годы методы РНК-секвениро-

вания станут рутинной процедурой, главным образом, за счет совершенствования и удешевления аппаратного обеспечения и упрощения подготовки проб. Это в свою очередь многократно увеличит эмпирические данные и приведет к необходимости их качественного анализа и интерпретации [5, 6].

С помощью РНК-секвенирования и последующего анализа полученных данных уже сейчас решаются многие как фундаментальные, так и прикладные задачи. Например, использование такого подхода позволяет детализировать взаимодействие генов в полиплоидных геномах растений [7], заглянуть поглубже в молекулярно-генетические механизмы эволюции генома, архитектуру генома и доместификацию у бобовых [8]. С помощью РНК-секвенирования объяснены и детализированы молекулярные механизмы, лежащие в основе ответа растения на стресс [9], а также открыты гены, ответственные за устойчивость растений к конкретным видам стресса [10]. Такие 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* [11–13]. Для его изучения перспективным представляется применение современных методов секвенирования, а именно транскриптомного анализа генов целлюлозосинтаз (*CesaA*-генов), который позволит получить косвенную, но ценную информацию о биогенезе растительной клеточной стенки. Поэтому целью настоящего исследования стала

© Т.А. ПОДВИЦКИЙ, Д.В. ГАЛИНОВСКИЙ,
Н.В. АНИСИМОВА, Г.Я. БАЕР, Я.В. ПИРКО,
А.И. ЕМЕЦ, Л.В. ХОТЫЛЕВА, Я.Б. БЛЮМ,
А.В. КИЛЬЧЕВСКИЙ, 2016

качественная и количественная характеристика экспрессии генов целлюлозосинтаз в различных органах и тканях высших растений с использованием данных РНК-секвенирования. Доступность в on-line базах результатов РНК-секвенирования различных культур (<http://phytozome.jgi.doe.gov/phytozome/aspect.do?name=Expression>.) позволила нам провести биоинформационный анализ транскриптов *CesA*-генов ряда высших растений.

Материалы и методы. Для анализа экспрессии генов целлюлозосинтаз высших растений нами проанализированы данные РНК-секвенирования, к которым имеется открытый доступ на генетическом ресурсе Phytozome (<http://phytozome.jgi.doe.gov/phytozome/aspect.do?name=Expression>). На время проведения анализа (10.10.2015) были доступны данные РНК-секвенирования для шести серий экспериментов с различными 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* × *E. urophylla*, 6 экспериментов), соя (*Glycine max* (L.) Merr., 9 экспериментов) и фасоль (*Phaseolus vulgaris* L., 11 экспериментов). Описание образцов, дан-

ные РНК-секвенирования которых использованы в работе, приведено в табл. 1.

Для идентификации транскриптов, которые кодируют целлюлозосинтазы в геноме упомянутых растений, а также в геноме *A. thaliana*, использовали программу Biomart (<http://www.biomart.org/>). С помощью этой программы отбирали последовательности, в описании которых аннотирован белковый домен Pfam 03552 (http://pfam.xfam.org/family/Cellulose_synt#tabview=tab0), характерный для суперсемейства целлюлозосинтаз, включающего *CesA*-гены и целлюлозосинтазоподобные гены. Последовательности генов с упомянутым доменом из генома модельного организма *A. thaliana* использовали в качестве референтных при реконструкции филогенетических отношений. Выравнивание аминокислотных транслированных последовательностей осуществляли с помощью программы MUSCLE (<https://www.ebi.ac.uk/Tools/msa/muscle/>), реконструкцию филогенетических отношений всех найденных аминокислотных последовательностей — с помощью программы ClustalW2 Phylogeny (http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/) и метода кластеризации Neighbour-joining. Филогению группы генов целлюлозосинтаз восста-

Таблица 1. Вегетативные органы и ткани растений, использованные для РНК-секвенирования [8]

Орган растения	Растение			
	Тополь	Эвкалипт	Соя	Фасоль
Корень	Верхушка корня		Корень	Корень (10 дней)
	Корень		Корневые волоски	Корень (19 дней)
Стебель	Узел	Апикальная меристема	Апикальная меристема	Узел
	Стебель	стебля	стебля	Стебель (10 дней)
		Флоэма	Узел	Стебель (19 дней)
		Незрелая ксилема	Стебель	
	Зрелая ксилема			
Лист	Зрелые листья	Зрелые листья	Листья	Листья
	Молодые листья	Молодые листья		Молодой трилистник
Цветок	Женская цветочная почка (ранняя стадия)		Цветок	Цветочная почка Цветок
Плод			Перикарпий плода Семя	Молодой боб Боб зеленой зрелости

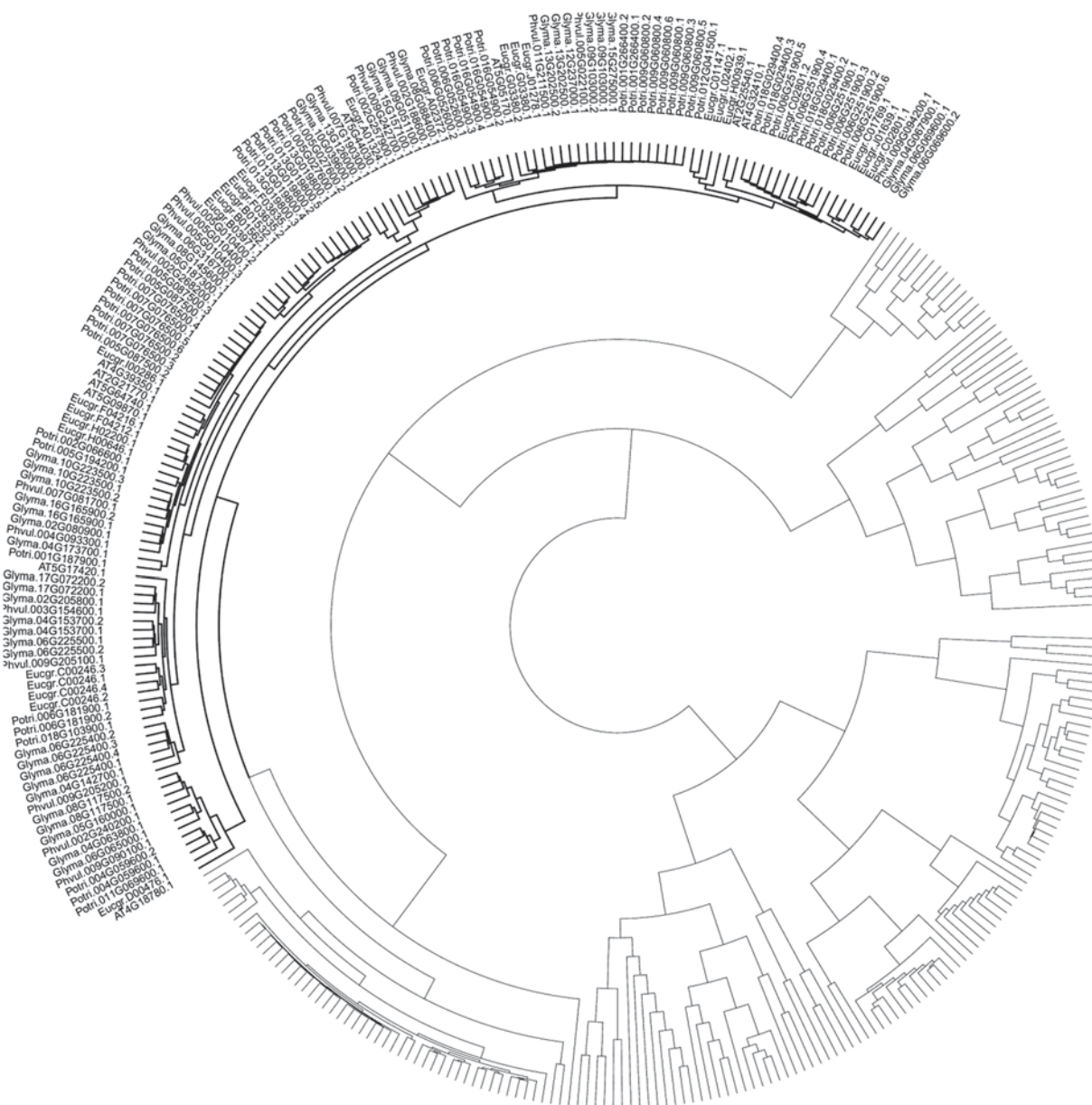


Рис. 1. Филогенетическое дерево транскриптов генов четырех исследуемых растений и *A. thaliana*, содержащих домен Pfam 03552. Кластер *CesaA*-генов выделен жирным начертанием

навливали, используя программу PhyML с JTT-моделью аминокислотных замен (<http://www.atgc-montpellier.fr/phyml/>).

Для анализа представленности транскриптов генов целлюлозосинтаз их Frkm-значения были извлечены из базы данных Phytozome10.3 с помощью интерфейса PhytoMine. Значение Frkm является числовым показателем коли-

чества «фрагментов на тысячу оснований на миллион картированных прочтений». Frkm — относительная единица измерения экспрессии генов, используемая при анализе результатов секвенирования парных прочтений и позволяющая сравнивать результаты различных экспериментов [14]. Анализ данных и построение гистограмм осуществляли в R-среде.

Анализ стратегий экспрессии генов целлюлозосинтаз высшими растениями

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 зна-

чением *F_{ркм}* целевого гена и *F_{ркм}* всего пула генов целлюлозосинтаз.

Результаты исследований и их обсуждение. В результате анализа генома четырех исследуемых видов растений и генома резуховидки (*A. thaliana* (L.) Heynh.) определили гены, со-

Таблица 2. Гены целлюлозосинтаз, аннотированные в исследуемых геномах

Ген	Код (ID) локуса генов целлюлозосинтаз						
	Резуховидка	Тополь	Эвкалипт	Соя	Фасоль		
<i>Первичная клеточная стенка</i>							
<i>CesA1/10</i>	AT2G25540	Potri.018G029400	Eucgr.C02801	Glyma.06G069600	Phvul.009G094200		
	AT4G32410	Potri.006G251900	Eucgr.C01769 Eucgr.H00939 Eucgr.J01639 Eucgr.L02402	Glyma.04G067900			
<i>CesA2/5/6/9</i>	AT5G64740	Potri.005G087500	Eucgr.I00286	Glyma.05G187300	Phvul.002G268200		
	AT5G09870	Potri.007G076500	Eucgr.F04212	Glyma.08G145600	Phvul.004G093300		
	AT4G39350	Potri.005G194200	Eucgr.F04216	Glyma.02G080900	Phvul.007G081700		
	AT2G21770	Potri.002G066600	Eucgr.H02200	Glyma.16G165900	Phvul.005G010400		
		Potri.005G027600	Eucgr.H00646	Glyma.10G223500			
		Potri.013G019800	Eucgr.F03635 Eucgr.B01532 Eucgr.B01562 Eucgr.B03971	Glyma.16G081000 Glyma.06G316700			
<i>CesA3</i>	AT5G05170	Potri.016G054900	Eucgr.A02372	Glyma.09G103000	Phvul.005G022100		
		Potri.006G052600	Eucgr.J01278	Glyma.15G275000	Phvul.011G211500		
		Potri.001G266400	Eucgr.G03380	Glyma.13G202500			
		Potri.009G060800		Glyma.12G237000			
<i>Вторичная клеточная стенка</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

держание домен Pfam 03552. Всего в исследованных нами геномах растений идентифицировано 199 генов и соответственно 311 вариантов их транскриптов в составе транскриптомов. При помощи филогенетического анализа дифференцировали гены *esA*-суперсемейства (рис. 1). Гены, транскрипты которых составили единый кластер с целлюлозосинтазами *A. thaliana*, были отнесены к *CesA*-генам и далее проанализированы. В геноме тополя обнаружили 17 генов целлюлозосинтаз, эвкалипта – 20, сои – 28 и в геноме фасоли – 15 *CesA*-генов (табл. 2). Следует отметить, что результаты экспериментов РНК-секвенирования, анализируемые в настоящей работе, получены разными исследователями в разное время и с разными целями, поэтому в них нет единого плана дифференциации тканей и органов растений. Выполненное нами соотнесение образцов из различных экспериментов приведено в табл. 1.

Общий уровень экспрессии генов целлюлозосинтаз существенно различается в разных органах одного растения, хотя в пуле генов целлюлозосинтаз есть гены, экспрессия которых достаточно стабильна. 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 % (рис. 2). 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 F_{rkm} соответственно в 1,24 и 1,49 раза выше для стебля, а у тополя в 1,25 раза 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 этих 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 клеточной стенки выше

и составляет не менее 42 %. Такое соотношение транскриптов наблюдали в структуре транскриптомов корней и стеблей всех проанализированных растений. Самая высокая доля генов вторичной клеточной стенки из всех проанализированных экспериментов выявлена в ксилеме стебля у гибрида *E. grandis* × *E. urophylla*, здесь она составила 87 %, а уровень экспрессии генов целлюлозосинтаз в образцах был самым высоким из всех имеющихся. Следует отметить, что в узлах стебля наблюдали картину, отличную от остальной части стебля. Во-первых, общий уровень транскриптов *CesA*-генов был ниже по сравнению с остальной частью стебля, а во-вторых, в пуле транскриптов *CesA*-генов обнаружена низкая доля генов целлюлозосинтаз вторичной клеточной стенки – 11–18 % (рис. 2, а, в, г).

Из полученных данных можно заключить, что в большинстве случаев экспрессия генов целлюлозосинтаз первичной клеточной стенки превалирует над экспрессией генов белков вторичной стенки. 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 обнаружен паттерн экспрессии *CesA*-генов, в котором гены целлюлозосинтаз вторичной клеточной стенки доминируют над первичной, – это образцы ксилемы (зрелой и незрелой) эвкалипта (рис. 2, б). 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 вторичной клеточной стенки в общем пуле генов целлюлозосинтаз, видимо, свидетельствует о специализации ткани. Это предположение согласуется с данными РНК-секвенирования растений джута (*Corchorus capsularis* L.), которые используются для получения растительного волокна. При биогенезе волокон джута показана экс-

Анализ стратегий экспрессии генов целлюлозосинтаз высшими растениями

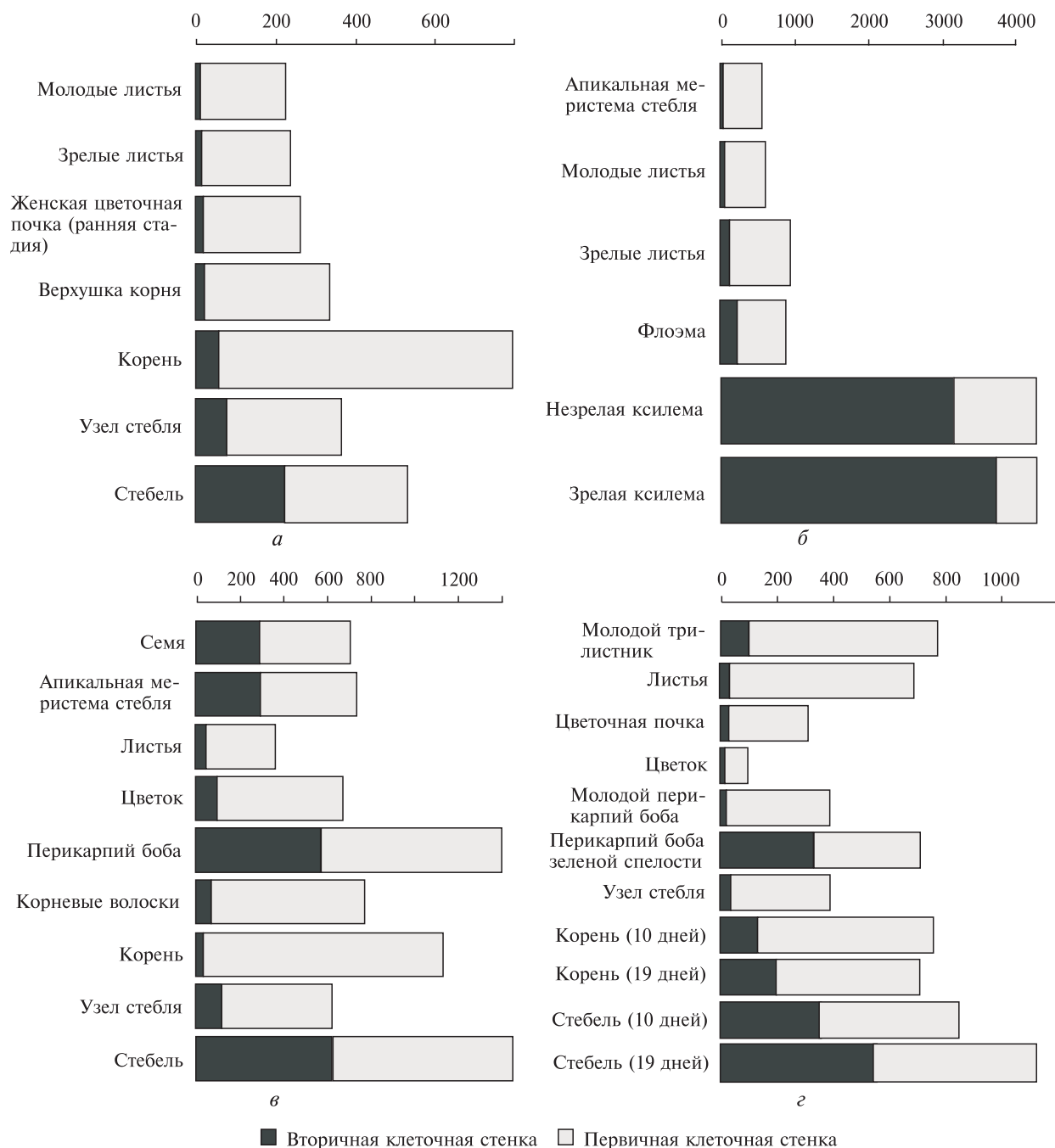


Рис. 2. Уровни экспрессии (Fpkm) генов целлюлозосинтаз, отвечающих за биогенез первичной и вторичной клеточной стенки тополя (а), эвкалипта (б), сои (в) и фасоли (г)

прессия девяти генов целлюлозосинтаз, из которых четыре экспрессировались наиболее сильно (comp11363_c0, comp11363_c1, comp24217_c0 и comp23531_c0) [15]. Для быстрорастущего тропического дерева кадамба (*Neolamarckia*

cadamba (Roxb.) Bosser) при формировании волокон в стебле показана преобладающая транскрипция четырех генов целлюлозосинтаз. Всего же в этих экспериментах зафиксирована транскрипция 18 генов целлюлозосинтаз. Транс-

крипты генов *NcCesA4* (comp78663_c0), *NcCesA7* (comp86337_c0), *NcCesA8* (comp82609_c0) и *NcCesA1* (comp86965_c0) в сумме составляли около 70 % общего числа транскриптов *CesA*-генов [16], причем в данном случае гены *NcCesA4*, *NcCesA7* и *NcCesA8* ассоциированы с биогенезом вторичной клеточной стенки, а ген *NcCesA1* – первичной [17]. В апикальном сегменте стебля дерева кадамба, где отсутствует активное депонирование растительных волокон, преобладают транскрипты генов *NcCesA1* (comp86965_c0), *NcCesA2* (comp86567_c1) и *NcCesA9* (comp86567_c0) [16]. Упомянутые гены являются ортологами соответствующих генов целлюлозосинтаз *A. thaliana* и ассоциированы с биосинтезом первичной клеточной стенки [17]. С помощью метода РНК-секвенирования у льна культурного (*Linum usitatissimum* L.), в стебле которого формируются флоэмные волокна, изучена транскрипция генов в апикальной и базальной частях стебля. Установлено, что экспрессия *CesA*-генов вторичной клеточной стенки *LusCesA4* (Lus10008225, Lus10008226) и

LusCesA8 (Lus10007296, Lus10029245) более чем в 100 раз усиливается в базальной части стебля по сравнению с апикальной [18]. Анализируя эти экспериментальные данные, а также результаты выполненной нами ранее аннотации генов целлюлозосинтаз льна [19], можно сделать вывод, что наибольший вес в общем пуле целлюлозосинтаз базальной части стеблей 14-дневных растений льна имеют гены *LusCesA4* (Lus10008226) – 23 %, *LusCesA3* (Lus10007538) – 13 % и *LusCesA8* (Lus10007296) – 10 %. Стоит отметить, что в работе [18] не проанализирована экспрессия гена *LusCesA7*, которая была показана рядом авторов при изучении биогенеза льноволокна [20–22].

В то же время для волокон рами (*Boehmeria nivea* L. Gaud), которые также ориентированы на масштабный синтез целлюлозы, характерна относительно сильная экспрессия 36 генов целлюлозосинтаз (всего идентифицирована экспрессия 51 *CesA*-гена) [23], что достаточно много по сравнению с другими рассмотренными культурами. Опираясь на эти результаты,

Таблица 3. Уровни экспрессии (Fpkм) отдельных генов целлюлозосинтаз тополя

Ген	ID гена	Молодые листья	Зрелые листья	Женская цветочная почка (ранняя стадия)	Верхушка корня	Корень	Узел стебля	Стебель
<i>Гены первичной клеточной стенки</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>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**

Примечание. Здесь и в табл. 4–6: *значения Fpkм более 15 % от общего пула *CesA*-генов; ** значения Fpkм от 5 до 14 % от общего пула *CesA*-генов. Остальные значения Fpkм менее 5 % от общего пула *CesA*-генов.

можно выделить семь *CesA*-генов (Unigene30783, Unigene23146, Unigene22477, Unigene4349, Unigene1078, Unigene1617, Unigene24841), экспрессия которых в коре стебля рами значительно превосходит экспрессию остальных упомянутых в статье генов целлюлозосинтаз [23].

Для каждого из исследованных в работе видов растений имеются гены целлюлозосинтаз с низкой функциональной активностью, доля транскриптов которых в любом из исследованных транскриптомов не превышала 5 %. В случае гибрида *E. grandis* × *E. urophylla* 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 перечисленных 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*), ко-

торая в двух пробах несколько превышала 15 %, и последовательностей Potri.002G066600 и Potri.013G019800, доля которых (кроме транскриптома корня) составляла менее 5 %. 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 образцах 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 раза сильнее, чем остальные *CesA*-гены вторичной клеточной стенки.

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*), экспрессирующиеся 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

Таблица 4. Уровни экспрессии (Fpkm) отдельных генов целлюлозосинтаз эвкалипта

Ген	ID гена	Апикальная меристема стебля	Молодые листья	Зрелые листья	Флоэма	Незрелая ксилема	Ксилема
<i>Гены первичной клеточной стенки</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>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*

данных для корня, цветков и плодов, где может меняться профиль экспрессии генов целлюлозосинтаз.

Гены целлюлозосинтаз вторичной клеточной стенки наиболее активно экспрессируются в ксилеме, где функционируют 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*). 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. Соотношение транскриптов генов *EuCesA4 : EuCesA7 : EuCesA8* во флоэме и ксилеме стебля примерно 1:1:4. Хотя не существует прямых экспериментальных доказательств, но считается, что различные три изоформы CESA собираются

в нативный целлюлозосинтезирующий комплекс в соотношении 1:1:1 [13]. Рассчитанное здесь соотношение транскриптов генов *EuCesA4*, *EuCesA7* и *EuCesA8* может указывать на соотношение целлюлозосинтазных субъединиц в трансмембранном целлюлозосинтезирующем комплексе либо на существование в одной ткани целлюлозосинтезирующих комплексов с разным составом субъединиц. Экспериментально установлено, что при отсутствии одной из изоформ CESA4, CESA7 или CESA8 пептида может осуществляться димеризация оставшихся двух. В случае изоформ CESA7 или CESA8 формируются гетеродимеры, а CESA4 способен формировать гомо- и гетеродимеры [13]. Эти данные согласуются с возможностью существования в одной ткани «популяций» целлюлозосинтезирующих комплексов, собранных из различных изоформ CESA белков.

Экспрессионный профиль генов целлюлозосинтаз сои оказался наиболее сложным для интерпретации (табл. 5). Из его особенностей можно отметить органоспецифичный ген

Таблица 5. Уровни экспрессии (Frkm) отдельных генов целлюлозосинтаз сои

Ген	ID гена	Семя	Апикальная меристема стебля	Листья	Цветок	Перикарпий боба	Корневые волоски	Корень	Узел стебля	Стебель
<i>Гены первичной клеточной стенки</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.10G223500	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>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**

Glyma.10G223500 (*CesA2/5/6/9*), функционирующий в цветке (его транскрипты составляют 49 % общего пула *CesA*-генов) и более нигде не проявляющийся. Среди генов целлюлозосинтаз первичной клеточной стенки только для генов Glyma.06G069600 (*CesA1/10*), Glyma.12G237000 (*CesA3*) и Glyma.16G165900 (*CesA2/5/6/9*) 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 [24], 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 *CesA*-генов вторичной клеточной стенки этот же ген наиболее интенсивно экспрессировался в стеб-

ле сои (9 %) наряду с геном Glyma.06G225500 (*CesA7*) – 7 %. Возможно, причину уже упоминавшегося большого количества *CesA*-генов, функционирующих при биогенезе волокна в стебле рами, также следует искать в филогенетической истории вида. В процессе эволюционного развития *Boehmeria nivea* могла произойти полиплоидизация его генома.

При анализе данных РНК-секвенирования *Phaseolus vulgaris* установлено, что из 15 генов целлюлозосинтаз, имеющих в геноме, функционируют 9, из которых 5 – *CesA*-гены первичной клеточной стенки и 4 – вторичной (табл. 2 и 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 этих генов достаточно стабильна 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 в перикарпиях бобов, а также в стебле увеличивается экспрессия генов целлюлозосинтаз вторичной клеточной стенки, среди которых лидирует ген

Таблица 6. Уровни экспрессии (Fpkм) отдельных генов целлюлозосинтаз фасоли

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

Phvul.009G242700 (*CesA4*), функционирующий также в корнях, где представленность его транскриптов доходит до 12 % 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 этого вида растений.

Выводы. Таким образом, в работе проанализированы транскрипты различных органов и тканей растений: тополя (*Populus trichocarpa*), эвкалипта (гибрид *E. grandis* × *E. urophylla*), сои (*Glycine max*) и фасоли (*Phaseolus vulgaris*). Уровень экспрессии генов целлюлозосинтаз существенно различался в органах и тканях одного конкретного растения. В стебле и корне *CesA*-гены экспрессировались 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, составляла не более 25 % общего пула *CesA*-генов, а в стебле – 42 % и более. При оценке функциональной активности генов целлюлозосинтаз выявили две альтернативные стратегии функционирования генов целлюлозосинтаз. Для изученных транскриптомов фасоли и эвкалипта характерно наличие «доминирующих» генов, которые давали более 50 % пула экспрессии генов целлюлозосинтаз. В случае

тополя и сои экспрессионный пул распределялся между большим числом генов, на каждый из которых приходилось менее 15 % общего числа транскриптов. Возможно, такой паттерн 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.

Работа выполнена при финансовой поддержке НАН Беларуси (договоры № Б15УК/А-041 и БРФФИ Б15М-101) и НАН Украины (проект № 0115U004560).

These sequence data used in this work were produced by the US Department of Energy Joint Genome Institute <http://www.jgi.doe.gov/> in collaboration with the user community.

ANALYSIS OF CELLULOSE SYNTHASE GENE EXPRESSION STRATEGIES IN HIGHER PLANTS USING RNA-SEQUENCING DATA

Ts.A. Padvitski, D.V. Galinously, N.V. Anisimova, G.Ya. Baer, Ya.V. Pirko, A.I. Yemets, L.V. Khotyleva, Ya.B. Blume, A.V. Kilchevsky

Institute of Genetic and Cytology of NAS of Belarus, Minsk

Institute of Food Biotechnology and Genomics of the NAS of Ukraine, Kyiv

E-mail: d.galinously@igc.by

The transcriptomes from 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 different types of tissues and organs in studied 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 a phylogenetic processes that occurred in genomes studied.

АНАЛІЗ СТРАТЕГІЙ ЕКСПРЕСІЇ ГЕНІВ ЦЕЛЮЛОЗОСИНТАЗ ВИЩИМИ РОСЛИНАМИ НА ОСНОВІ ДАНИХ РНК-СЕКВЕНУВАННЯ

Т.О. Подвицкий, Д.В. Галиновський, Н.В. Анісімова, Г.Я. Баєр, Я.В. Пірко, А.І. Ємець, Л.В. Хотильова, О.В. Кільчевський

Проаналізовано транскриптоми різних органів і тканин чотирьох видів рослин – тополі, евкалипта, сої та

квасолі. Встановлено істотні відмінності рівня експресії генів целлюлозосинтаз в органах і тканинах однієї рослини. Для транскриптомів квасолі та евкаліпта характерна наявність «домінуючих» генів, які дають понад 50 % пулу експресії генів целлюлозосинтаз. У тополі та сої експресійний пул розподілений між великим числом генів. Ймовірно, стратегія експресії *CesA*-генів відображає процес філогенезу досліджуваних геномів.

СПИСОК ЛИТЕРАТУРЫ

1. Wang, Z., Gerstein, M., and Snyder, M., RNA-Seq: a revolutionary tool for transcriptomics, *Nat. Rev. Genet.*, 2009, vol. 10, no. 1, pp. 57–63.
2. Egan, A.N., Schlueter, J., and Spooner, D.M., Applications of next-generation sequencing in plant biology, *Am. J. Bot.*, 2012, vol. 99, no. 2, pp. 175–185.
3. Strickler, S.R., Bombarely, A., and Mueller, L.A., Designing a transcriptome next-generation sequencing project for a nonmodel plant species, *Am. J. Bot.*, 2012, vol. 99, no. 2, pp. 257–266.
4. Giorgi, F.M., Del Fabbro, C., and Licausi, F., Comparative study of RNA-seq- and Microarray-derived coexpression networks in *Arabidopsis thaliana*, *Bioinformatics*, 2013, vol. 29, no. 6, pp. 717–724.
5. Chu, Y., and Corey, D.R., RNA Sequencing: Platform selection, experimental design, and data interpretation, *Nucl. Acid Ther.*, 2012, vol. 22, no. 4, pp. 271–274.
6. Finotello, F., and Di Camillo, B., Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis, *Brief. Funct. Genom.*, 2015, vol. 14, no. 2, pp. 130–142.
7. Pnut, D.C., Coate, J.E., Luciano, A.K., Owens, T.G., May, G.D., Farmer, A., and Doyle, J.J., A comparative transcriptomic study of an allotetraploid and its diploid progenitors illustrates the unique advantages and challenges of RNA-seq in plant species, *Am. J. Bot.*, 2012, vol. 99, no. 2, pp. 383–396.
8. O'Rourke, J.A., Bolon, Y.T., Bucciarelli, B., and Vance, C.P., Legume genomics: understanding biology through DNA and RNA sequencing, *Ann. Bot.*, 2014, vol. 113, no. 7, pp. 1107–1120.
9. Wakasa, Y., Oono, Y., Yazawa, S., Ozawa, K., Handa, H., Matsumoto, T., and Takaiwa, F., RNA sequencing-mediated transcriptome analysis of rice plants in endoplasmic reticulum stress conditions, *BMC Plant Biol.*, 2014, vol. 14, doi: 10.1186/1471-2229-14-101.
10. Tsukagoshi, H., Suzuki, T., Nishikawa, K., Agarie, S., Ishiguro, S., and 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.
11. Fry, S.C., Cell wall biosynthesis, *Handbook of Plant Science*, Roberts, K., Ed., Wiley, 2007, pp. 276–283.
12. Farrokhi, N., Burton, R.A., Brownfield, L., Hrmova, M., Wilson, S.M., Bacic, A., and Fincher, G.B., Plant cell wall biosynthesis: genetic, biochemical and functional genomics approaches to the identification of key genes, *Plant Biotechnol. J.*, 2006, vol. 4, no. 2, pp. 145–167.
13. Kumar, M., and Turner, S., Plant cellulose synthesis: CESA proteins crossing kingdoms, *Phytochemistry*, 2015, vol. 112, pp. 91–99.
14. Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L., Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms, *Nat. Biotechnol.*, 2010, vol. 28, no. 5, pp. 511–515.
15. Zhang, L., Ming, R., Zhang, J., Tao, A., Fang, P., and Qi, J., De novo transcriptome sequence and identification of major bast-related genes involved in cellulose biosynthesis in jute (*Corchorus capsularis* L.), *BMC Genom.*, 2015, vol. 16:1062, doi: 10.1186/s12864-015-2256-z.
16. Ouyang, K., Li, J., Zhao, X., Que, Q., Li, P., Huang, H., Deng, X., Singh, S.K., Wu, A.-M., and Chen, X., Transcriptomic analysis of multipurpose timber yielding tree *Neolamarckia cadambaduring* xylogenesis using RNA-Seq, *PLoS One*, 2016, vol. 11, no. 7, e0159407.
17. Maleki, S.S., Mohammadi, K., and Ji, K.-Sh., Characterization of cellulose synthesis in plant cells, *Sci. World J.*, 2016, Article ID 8641373, 8 p., <http://dx.doi.org/10.1155/2016/8641373>.
18. Zhang, N., and Deyholos, M.K., RNA Seq analysis of the shoot apex of flax (*Linum usitatissimum*) to identify phloem fiber specification genes, *Front. Plant Sci.*, 2016, vol. 7:950, doi: 10.3389/fpls.2016.00950.
19. Pydiura, N.A., Bayer, G.Ya., Galinouski, D.V., Yemets, A.I., Pirko, Ya.V., Padvitski, T.A., Anisimova, N.V., Khotyleva, L.V., Kilchevski, A.V., and Blume, Ya.B., Bioinformatic search for cellulose synthase genes in flax (*Linum usitatissimum*) and their phylogenetic analysis, *Cytol. Genet.*, 2015, vol. 49, no. 5, pp. 279–287.
20. Galinousky, D.V., Anisimova, N.V., Raiski, A.P., Leontiev, V.N., Titok, V.V., and Khotyleva, L.V., Cellulose synthase genes that control the fiber formation of flax (*Linum usitatissimum* L.), *Russ. J. Genet.*, 2014, vol. 50, no. 1, pp. 26–34.

21. Mokshina, N., Gorshkova, T., and Deyholos, M.K., Chitinase-like (*CTL*) and cellulose synthase (*CESA*) gene expression in gelatinous-type cellulosic walls of flax (*Linum usitatissimum* L.) bast fibers, *PLoS One*, 2014, vol. 9, no. 2, e97949.
22. Chantreau, M., Chabbert, B., Billiard, S., Hawkins, S., and Neutelings, G., Functional analyses of cellulose synthase genes in flax (*Linum usitatissimum*) by virus-induced gene silencing, *Plant Biotechnol. J.*, 2015, vol. 13, no. 9, pp. 1312–1324.
23. Liu, T., Zhu, S., Tang, Q., Chen, P., Yu, Y., and Tang, Sh., De novo assembly and characterization of transcriptome using Illumina paired-end sequencing and identification of *CesA* gene in ramie (*Boehmeria nivea* L. Gaud), *BMC Genom.*, 2013, vol. 14:125, doi: 10.1186/1471-2164-14-125.
24. Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, Th., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakura, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan B., Lindquist, E., Peto, M., Grant, D., Shu, Sh., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Zh., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.-Ch., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C., and Jackson, S.A., Genome sequence of the palaeopolyploid soybean, *Nature*, 2010, vol. 463, no. 7278, pp. 178–183.

Поступила 06.08.16