

**ОБНАРУЖЕНИЕ НУКЛЕОТИДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ,
КОДИРУЮЩИХ DCL-БЕЛКИ В РАСТЕНИЯХ
ШАЛОТА *Allium cepa* L. var. *aggregatum* G. Don**

А.В. АРХИПОВ, В.К. ВИШНИЧЕНКО

Впервые в мировой литературе представлены полученные с помощью методов полимеразной цепной реакции, секвенирования и компьютерного анализа доказательства экспрессии в растениях лука-шалота (*Allium cepa* L. var. *aggregatum* G. Don) нуклеотидных последовательностей, кодирующих полипептиды, гомологичные известным растительным DCL-белкам (Dicer-like). В составе исследуемых полипептидных цепей выявлены характерные для молекулы DCL-белка домены dsRNA bind, PAZ и RIBOc. Решение поставленной задачи оказалось возможным только при применении диагностических праймеров, дизайн которых базировался на использовании информации по EST-последовательностям (expressed sequence tags) у *Allium sativum* (база данных GarlicESTdb). Результаты позволяют сделать вывод о высокой степени дивергенции DCL-белков арабидопсиса, тополя и риса, с одной стороны, и растений рода *Allium* — с другой, что выявляется на уровне как нуклеотидных последовательностей мРНК, так и аминокислотных последовательностей исследуемых полипептидов.

Ключевые слова: РНК-интерференция, белки Dicer и Dicer-like (DCL), лук-шалот (*Allium cepa* L. var. *aggregatum*), X-вирус шалота.

Keywords: RNA-interference, Dicer&Dicer-like (DCL) proteins, shallot (*Allium cepa* L. var. *aggregatum*), Shallot virus X.

В растениях уровень экспрессии многих генов может модулироваться эпигенетически при участии механизмов РНК-интерференции, ключевыми элементами которых служат эндорибонуклеазы (RNase III) Dicer или Dicer-like (DCL), эндонуклеазы Argonautes (AGOs) и клеточная РНК-зависимая РНК-полимераза (RDR) (1). В результате действия этих ферментов образуется широкий спектр малых РНК (микроРНК и малых интерферирующих РНК — siRNA), принимающих участие в осуществлении разнообразных биологических процессов, связанных, в частности, с развитием растений (2). Мощными экзогенными индукторами РНК-интерференции в растениях служат фитовирусы, и вызванный вирусной инфекцией эффект выключения, или «замолкания», генов (virus-induced gene silencing, VIGS) (3) представляет собой результат специфического расщепления репликативных (двуцепочечных) форм вирусных РНК растительными DCL-белками. Деградация вирусной РНК эндорибонуклеазами этого класса, сопровождающаяся образованием вирус-специфических siRNA, — это первый этап антивирусной защитной реакции у растения, базирующейся на феномене РНК-интерференции (4).

В геномах растений присутствуют несколько гомологичных последовательностей, кодирующих различные DCL-белки (5); так, у *Arabidopsis thaliana* базовый набор из четырех паралогов кодирует белки AtDCL 1-4, функционально специализированные и характеризующиеся выраженной избирательностью в отношении субстрата (6). AtDCL1 генерирует малые РНК, используя в качестве субстрата одноцепочечные РНК, отдельные участки которых имеют несовершенную двуцепочечную структуру типа stem-loop; AtDCL2, AtDCL3 и AtDCL4 обладают большим сродством по отношению к молекулам РНК с совершенной двуцепочечной структурой. При вирусной инфекции все четыре типа DCL-белков арабидопсиса участвуют в расщеплении репликативных форм вирусных РНК, однако основную антивирусную функцию выполняет AtDCL4: в результате действия этого белка образуются вирус-специфические siRNA длиной 21 п.н., которые, во-первых, принимают активное участие в специфической деградации вирусной РНК эффекторным комплексом RISC (RNA-induced silencing complex), во-вторых, функционируют в качестве мобильного сигнала РНК-интерференции (7). Малые интер-

ферирующие РНК, генерируемые AtDCL2, играют в этих процессах подчиненную роль, а AtDCL1 и AtDCL3, хотя и способны разрезать вирусные двуцепочечные РНК на короткие (24 bp) фрагменты, антивирусной активностью не обладают. Не исключено, однако, что описанная модель не универсальна, и в растениях с другим, нежели у арабидопсиса, набором DCL-белков, функциональные взаимосвязи между ними могут быть иными.

Нашей задачей было обнаружение экспрессии и анализ нуклеотидных последовательностей, кодирующих DCL-белки в растениях *Allium sera* var. *aggregatum* — единственного восприимчивого хозяина X-вируса шалота (ХВШ, прототип рода аллексивирусов) (8). До проведения настоящего исследования какая-либо информация о DCL-белках *Allium sera* в мировой научной литературе отсутствовала.

Методика. Вегетативное размножение зараженных ХВШ растений шалота (*Allium sera* L. var. *aggregatum* G. Don) осуществляли в условиях моно-вирусной инфекции. В качестве неинфицированных образцов (контроль) использовали растения шалота сорта Red Sun («Crocus Co.», Великобритания).

Тотальную РНК экстрагировали из листьев через 2 нед после высадки луковиц с помощью тризола (TRIzol Reagent, «Life Technologies», США) в соответствии с протоколом производителя. Синтез кДНК выполняли следующим образом: на стадии отжига инкубационную смесь, содержащую 0,5–2,0 мкг тотальной РНК, 2 мкл олиго(dT)₁₈ (40 мкМ) или «случайных» (random hexamers, 40 мкМ) праймеров, 4 мкл смеси dNTP (каждый в концентрации 2,5 мМ, pH 7,0, Трис-HCl) и свободную от нуклеаз H₂O (общий объем смеси 16 мкл), прогревали при 70 °С в течение 5 мин, центрифугировали и немедленно помещали на лед; к смеси добавляли 2 мкл 10×RT буфера, 1 мкл (10 ед.) ингибитора рибонуклеаз (RNasin) и 1 мкл (50 ед.) ревертазы (M-MLV, Moloney) («Синтол», Россия). Обратную транскрипцию (в объеме 20 мкл) проводили при 42 °С в течение 1 ч, затем фермент инактивировали нагреванием при 90 °С в течение 10 мин. При дизайне ПЦР-праймеров использовали две их группы: диагностические праймеры, соответствующие консервативным участкам геномных последовательностей DCL-белков арабидопсиса и позволяющие идентифицировать DCL-белки 1–4 методом полимеразной цепной реакции (9), и праймеры, при конструировании которых в настоящей работе был применен альтернативный подход, основанный на использовании базы данных GarlicESTdb (10). Полимеразную цепную реакцию осуществляли в следующем режиме: денатурация при 94 °С 2 мин, затем 30 циклов (94 °С, 30 с; 60 °С, 30 с; 72 °С, 30 с) и финальная элонгация при 72 °С 2 мин. Ампликоны фракционировали в 0,8 % агарозных гелях, экстрагировали и очищали с применением набора реагентов DNA extraction kit # K0513 («Fermentas», Литва), нуклеотидную последовательность продукта определяли в секвенаторе 3130xl («Applied Biosystems», США). Компьютерный анализ полученных последовательностей выполняли с помощью соответствующих баз данных и пакета компьютерных программ NCBI (The National Center for Biotechnology Information, США).

Результаты. Для решения поставленной задачи предполагалось использовать диагностические праймеры, соответствующие консервативным участкам в геномных последовательностях DCL-белков арабидопсиса (9). Подобный экспериментальный подход основан на концепции, согласно которой для всех покрытосеменных и, возможно, для всех многоклеточных растений характерен базовый набор четырех типов DCL-белков, обладающих высокой степенью гомологии; имеющиеся экспериментальные данные свидетельствуют, что такие праймеры действительно позволяют идентифицировать гены DCL-белков 1–4 методом ПЦР не только у арабидопсиса, но также у ячменя, кукурузы, хлопчатника и люпина (9). В наших экспериментах с помощью указанных праймеров были надежно идентифицированы DCL-белки

1-4 в растениях арабидопсиса, однако их применение на ядерной ДНК или кДНК шалота оказалось безуспешным: ампликоны либо не синтезировались вовсе, либо продукты реакции не обладали сколько-нибудь выраженной гомологией с последовательностями, кодирующими DCL-белки 1-4 (данные не приведены). Мы интерпретировали это как свидетельство выраженной дивергенции генов DCL-белков *A. thaliana* и *Al. cepa*. В настоящее время прямое сравнение последовательностей ДНК арабидопсиса и репчатого лука невозможно из-за ограниченности соответствующей информации по *Al. cepa*: секвенирование генома у этого вида представляет собой трудную и до сих пор не решенную проблему; опубликованные недавно результаты многолетней работы по секвенированию генома репчатого лука (11) выявили лишь протяженные участки вырожденных ретровирусных элементов и транспозонов, однако никакой информации о генах DCL-белков получено не было. В связи с этим мы предприняли попытку применить базу данных GarlicESTdb (<http://garlicdb.kribb.re.kr>), включающую информацию о 21,595 EST-последовательностях (expressed sequence tags) у *Al. sativum* (10). Использование в аналогичных целях базы данных по *Al. cepa* (12), содержащей значительно меньшее число EST-последовательностей, оказалось безуспешным.

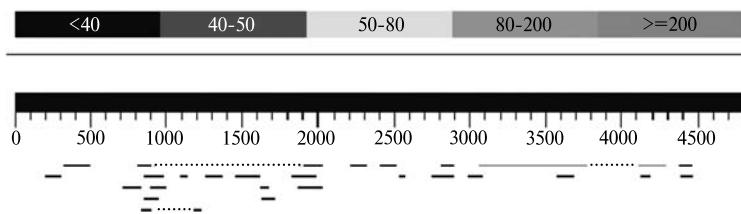


Рис. 1. Результаты скрининга EST-последовательностей, комплементарных мРНК AtDCL2, в базе данных GarlicESTdb. Обнаруженные EST (expressed sequence tags) соответствуют последовательностям нуклеотидов мРНК в позициях 3100-3750 и 4100-4250. Аналогичные данные были получены с участием мРНК, кодирующих все известные растительные DCL-белки.

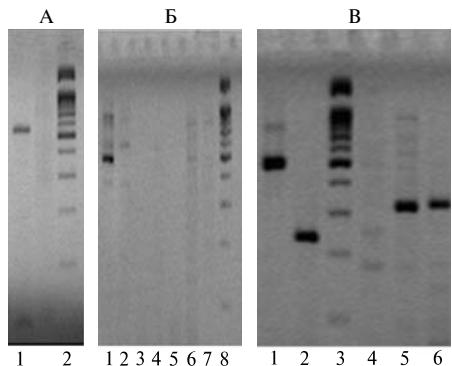
Конструирование праймеров на основе базы данных GarlicESTdb выполняли следующим образом. С помощью программы BLAST/tblastx, которая транслирует предложенную нуклеотидную последовательность и затем сравнивает полученные аминокислотные последовательности между собой, в базе данных GarlicESTdb был осуществлен поиск EST-последовательностей, кодирующих полипептиды с высокой степенью гомологии с различными DCL-белками. На рисунке 1 представлен результат такого скрининга с участием матричной мРНК AtDCL2; аналогичные данные получили для мРНК всех известных растительных DCL-белков. Подобным образом из базы данных GarlicESTdb выбрали три EST-последовательности — EPP005KGAA12S003959 (783 bp), EPP004KGAA12S004834 (679 bp) и EPP004KGAA12S005240 (675 bp); все они содержали по одной открытой рамке считывания, и степень гомологии транслируемых с этих последовательностей полипептидов с различными DCL-белками арабидопсиса, риса и тополя варьировала в пределах 60-70 %. При этом соответствующий компьютерный анализ выявил в составе исследуемых полипептидных цепей характерные для молекулы DCL-белка dsRNA bind, PAZ и RIBOc (RNase IIIa и RNase IIIb) домены (данные не приведены). На основе полученной информации были сконструированы праймеры, соответствующие различным участкам трех выбранных EST-последовательностей (табл.).

При проведении обратной транскрипции и полимеразной цепной реакции с участием тотальной РНК здоровых или инфицированных растений шалота в стандартных условиях наиболее эффективными и позволяющими получать стабильные результаты оказались пары праймеров L-106—R-644 (EPP005KGAA12S003959), L-133—R-632 (EPP004KGAA12S004834) и L-38—R-552 (EPP004KGAA12S005240) (рис. 2, А, Б, В). После экстракции из геля, очистки и секвенирования ампликонов компьютерный анализ не выявил в со-

Диагностические ПЦР-праймеры, использованные для амплификации нуклеотидных последовательностей, кодирующих DCL-белки в растениях шалота

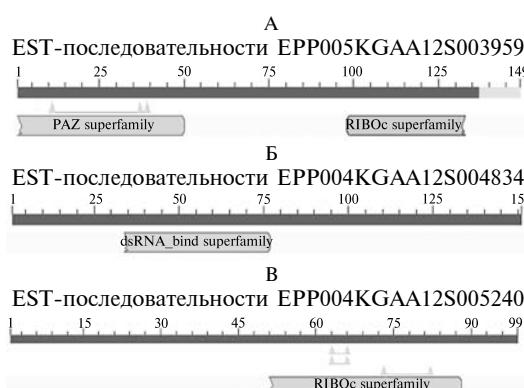
Обозначение праймера	Нуклеотидная последовательность праймера
	E S T - п о с л е д о в а т е л ь н о с т ь EPP004KGAA12S005240
L-8	CGCTCAGCCCTTTAACAA
L-38 ^a	TATGAGCTGGCCATTGCTAA
L-102	ACTTCTTGCAATTGGCAGA
L-378	ACGAGGAAGCCAAAAGGAT
R-552 ^b	GCATCCCCAAGATACTCCAA
R-397	ATCCCTTTGGCTTCCTCGT
R-120	CTGCCCAATGCAAAGAAGTT
R-628	AAATCTGAAATTGACCTGGTTTC
	E S T - п о с л е д о в а т е л ь н о с т ь EPP004KGAA12S004834
L-133 ^a	GTCCTCCGTTTGATTGAC
L-90	TTCTACTGCTCCGACTTCC
L-38	TGAATCAACTGGAGCAAACC
L-323	AGCAGAGCTTGCTTCGAGG
L-395	GGTTGATTCTGGCTCTCGT
R-567	CGCCCTGGTCTTCTCTATG
R-632 ^b	AGCCATCAAAACCTGCTCAC
R-415	ACGAGGACCAAGAATCAACC
R-4834	ATATTCTGACGAATCAGTAAATTGA
	E S T - п о с л е д о в а т е л ь н о с т ь EPP005KGAA12S003959
L-285	CCGGAACCTTGCAAAATCAT
L-106 ^a	TGAAGAGCAGGGAAACTCTCA
L-3	CATTGTGCCTTGAAAATTCTG
R-432	TGACATTTGCATGCAGTGATCC
R-644 ^b	TCCTATCCGGCATAAGTTGG

П р и м е ч а н и е. а, б — наиболее эффективные пары праймеров.



(1, 2, 4, 5 и 6 — соответственно L133—R632, L395—R632, L90—R4834, L90—R408, L90—R415, 3 — маркер молекулярных масс). EST — expressed sequence tags; маркер молекулярных масс M25 100-3000 bp DNA Ladder (НПО «Сибэнзим», Россия).

ответствующих базах данных последовательностей, обладающих сколько-нибудь заметной гомологией с исследуемыми; в то же время транслируемые с этих последовательностей полипептиды проявляли высокую степень гомологии со всеми растительными DCL-белками, в частности наибольшую (81 %) —



DCL-белка домены (dsRNA bind, PAZ или RIBOc, рис. 3, А, Б, В). Анализ

Рис. 2. Электрофорерограммы ампликонов при выявлении экспрессии нуклеотидных последовательностей, кодирующих DCL-белки в растениях шалота, методом полимеразной цепной реакции с участием диагностических праймеров, сконструированных на основе базы данных GarlicESTdb:

А — праймерные EST-последовательности EPP005KGAA12S003959 (1 и 2 — соответственно пара L-106—R-644 и маркер молекулярных масс); Б — праймерные EST-последовательности EPP004KGAA12S005240 (1-8 — соответственно пары L38—R552, L38-F—R628, L102—R397, L102—R552, L102—R628, L378—R552, L378—R628 и маркер молекулярных масс); В — праймерные EST-последовательности EPP004KGAA12S004834

(1, 2, 4, 5 и 6 — соответственно L133—R632, L395—R632, L90—R4834, L90—R408, L90—R415, 3 — маркер молекулярных масс). EST — expressed sequence tags; маркер молекулярных масс M25 100-3000 bp DNA Ladder (НПО «Сибэнзим», Россия).

Рис. 3. Результаты компьютерного доменного анализа полипептидов, которые транслируются с нуклеотидных последовательностей, амплифицированных с помощью наиболее эффективных пар праймеров с тотальной ДНК растений шалота: А — L-106—R-644, Б — L-133—R-632, В — L-38—R-552; EST — expressed sequence tags.

с DCL-белками *Populus trichocarpa*. При этом аминокислотные последовательности у всех трех полипептидов не были идентичными, и в структуре каждого имелись различные специфичные для молекулы

DCL-белка домены (dsRNA bind, PAZ или RIBOc, рис. 3, А, Б, В). Анализ

аминокислотных последовательностей доменов RIBOc, обнаруженных в двух полипептидах, показал, что в каждом из трех характерных участков доменов (active site, metal binding site и dimerization interface) присутствуют по две аминокислотные замены, тогда как последовательности этих участков во всех других известных растительных DCL-белках идентичны.

Таким образом, впервые в мировой научной литературе мы представили доказательство экспрессии в здоровых и инфицированных X-вирусом шалота растениях *Allium cepa* L. var. *aggregatum* G. Don нуклеотидных последовательностей, кодирующих три домена различных DCL-белков — dsRNA bind, PAZ и RIBOc (RNase IIIa и RNase IIIb). Кроме того, полученные данные позволяют сделать вывод о высокой степени дивергенции DCL-белков у арабидопсиса и ряда других видов, с одной стороны, и растений рода *Allium* — с другой, что выявляется на уровне как нуклеотидных последовательностей матричных РНК, так и аминокислотных последовательностей исследуемых полипептидов. При этом эволюционные взаимосвязи между нуклеотидными последовательностями, кодирующими указанные домены, пока остаются неизвестными.

Л И Т Е Р А Т У Р А

1. Meister G., Tuschl T. Mechanisms of gene silencing by double-stranded RNA. *Nature*, 2004, 431: 343-349.
2. Ketting R.F. The many faces of RNAi. *Developmental Cell*, 2011, 20: 148-161.
3. Baulcombe D. RNA silencing in plants. *Nature*, 2004, 431: 356-363.
4. Garcia-Ruiz H., Takeda A., Chapman E.J. e.a. Arabidopsis RNA-dependent RNA polymerases and Dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *The Plant Cell*, 2010, 22: 481-496.
5. Liu Q., Ruiz-Ferrer V., Voynnet O. Roles of plant small RNAs in biotic stress responses. *Annu. Rev. Plant Biol.*, 2009, 60: 485-510.
6. Dunoyer P., Schott G., Himber C. e.a. Small RNA duplexes function as mobile silencing signals between plant cells. *Science*, 2010, 328: 912-916.
7. Vishnichenko V.K., Konareva T.N., Zavriev S.K. A new filamentous virus in shallot. *Plant Pathol.*, 1993, 42: 121-125.
8. Margis R., Fusaro A.F., Smith N.A., Curtin S.J., Watson J.M., Finegan E.J., Waterhouse P.M. The evolution and diversification of Dicers in plants. *FEBS Lett.*, 2006, 580: 2442-2450.
9. Kim D.-W., Jung T.-S., Nam S.-H. e.a. GarlicESTdb: an online database and mining tool for garlic EST sequences. *BMC Plant Biol.*, 2009, 9: 61.
10. Jackson J., Meyer J.D.F., Suzuki G. e.a. Pilot sequencing of onion genomic DNA reveals fragments of transposable elements, low gene densities, and significant gene enrichment after methyl filtration. *Mol. Genet. Genomics*, 2008, 280: 287-292.
11. Joseph C., Kuhl J.C., Cheung F. e.a. A unique set of 11,008 onion expressed sequence tags reveals expressed sequence and genomic differences between the monocot orders asparagales and poales. *The Plant Cell*, 2004, 16: 114-125.

ГНУ Всероссийский НИИ сельскохозяйственной
биотехнологии Россельхозакадемии,
127550 г. Москва, ул. Тимирязевская, 42,
e-mail: vish@iab.ac.ru

Посупила в редакцию
6 июня 2011 года

THE DICER-LIKE PROTEINS CODING NUCLEOTIDE SEQUENCES IN SHALLOT *Allium cepa* L. var. *aggregatum* G. Don

A.V. Arkhipov, V.K. Vishnichenko

S u m m a r y

For the first time in the world literature the authors presented the evidence of expression in shallot plants (*Allium cepa* L. var. *aggregatum* G. Don) of nucleotide sequences coding the polypeptides homologous to known plant Dicer-like protein (DCL), obtained by PCR, sequencing and computer analysis. The investigated polypeptides contain specific for DCL-protein the dsRNA bind, PAZ and RIBOc domains. The solution of assigned task is available only with use of diagnostic oligonucleotide primers, the design of which was accomplished on the basis of information on EST-sequences (expressed sequence tags) in *Allium sativum* (GarlicESTdb integrated database). The results permit to draw a conclusion about high extent of divergence of DCL-proteins in *Arabidopsis*, poplar, rice, on the one hand, and the plants of *Allium* genus, on the other hand, that is revealed on level both nucleotide sequences of mRNA, and amino acid sequences of studied polypeptides.