

ГЕНЫ СИСТЕМЫ ДЕТОКСИКАЦИИ В ГЕНОМЕ ТЛИ *MACROSIPHUM ROSAE*

Н.В. Воронова-Барте, С.С. Левыкина, А.М. Шульгович

Белорусский государственный университет, Минск, Беларусь

E-mail: anyashulgovich@gmail.com

В работе представлены результаты аннотации ABC-транспортеров системы детоксикации в геноме *Macrosiphum rosae*. В геноме *M. rosae* найдено 67 генов ABC-транспортеров. Средняя длина гена – 5755 нуклеотидов, средняя длина экзонов – 1846 нуклеотидов, средняя длина интронов – 3909 нуклеотидов.

Ключевые слова: геномика; аннотация генома; тли.

ВВЕДЕНИЕ

Система детоксикации у насекомых включает в себя монооксигеназы цитохрома P450 (CYP450), глутатион S-трансферазы (GST), карбоксил-эстеразы (CSE), гликозилтрансферазы (UDT) и АТФ-связывающие каскадные транспортеры (ABC). Белки системы детоксикации метаболизируют не только защитные вещества растений, но также обеспечивают адаптацию насекомого к инсектицидам. Возникновение устойчивости у таких фитофагов может быть обусловлено увеличением копий этих генов и тем, что копии становятся уникальными для каждого метаболита.

ABC-транспортеры это суперсемейство интегральных мембранных белков, которые используют АТФ для перемещения различных субстратов через липидные мембраны. Они играют основную роль в заключительной фазе процесса детоксикации ксенобиотиков. В связи с этим, идентификация их генов является важной для разработки эффективных средств контроля численности, например, таких сельскохозяйственных вредителей как тли.

МАТЕРИАЛЫ И МЕТОДЫ

Целью исследования являлась аннотация генов ABC-транспортеров системы детоксикации *Macrosiphum rosae* фауны Беларуси.

Предварительно была проведена функциональная аннотация генома *Macrosiphum rosae*. Финальная модель была получена как консенсус между НММ-предсказаниями (скрытые марковские модели), белковой и РНК-гомологией, что подтверждает функциональность обнаруженных генов. Получение консенсуса проводилось в программе Evidence Modeller [1].

Поиск доменов в полученных моделях генов проводилась с помощью поиска IPR и Pfam-доменов в Interproscan и интерпретируемого языка про-

граммирования Python [2]. Каждый обнаруженный ген *M. rosae* был про-
 курирован в геномном браузере Apollo для уточнения границ экзонов: рас-
 положения старт- и стоп-кодонов в последовательностях, канонических
 сайтов сплайсинга [3].

РЕЗУЛЬТАТЫ

Результаты структурно-функционального описания идентифициро-
 ванных генов представлены в таблице.

Структурно-функционального описания идентифицированных генов

Семейство	Идентификационный номер	Локус (ориентация)	Длина гена, п.н.	Число экзонов	Длина экзонов, п.н.
A5	MACRO000011987	Scaffold_168444 (-)	3355	6	1650
B6	MACRO000000126	Scaffold_163101 (+)	7166	13	2523
C	MACRO000001134	Scaffold_168689 (-)	10796	20	3612
F2	MACRO000003103	Scaffold_165388 (+)	5353	11	1725
F4	MACRO000015501	Scaffold_152023 (-)	1198	2	1014
G4	MACRO000014479	Scaffold_156322 (-)	1825	6	663
G4	MACRO000005456	Scaffold_166244 (+)	8210	8	1356
G4	MACRO000016283	Scaffold_183958 (-)	1234	2	435
G4	MACRO000002655	Scaffold_185181 (+)	9098	12	1984
G8	MACRO000004229	Scaffold_169484 (+)	2339	9	1563
G9	MACRO000003028	Scaffold_153724 (+)	5314	13	1827
G20	MACRO000017447	Scaffold_159883 (+)	2745	2	447
G20	MACRO000002548	Scaffold_167411 (-)	3717	12	1709
G20	MACRO000002423	Scaffold_171574 (+)	12921	12	1512
G23	MACRO000000434	Scaffold_171604 (+)	9280	14	1833
G23	MACRO000000433	Scaffold_171604 (+)	5192	11	1479
G23	MACRO000003467	Scaffold_152778 (-)	2537	3	492
G23	MACRO000000996	Scaffold_153633 (-)	6610	15	2262
G23	MACRO000011734	Scaffold_157450 (+)	2246	3	414
G23	MACRO000005003	Scaffold_153904 (-)	2109	8	1067
G23	MACRO000008455	Scaffold_164900 (+)	1490	6	882
G23	MACRO000003449	Scaffold_168768 (-)	6732	10	1308
G23	MACRO000015911	Scaffold_207302 (+)	368	2	291
G23	MACRO000001104	Scaffold_165944 (+)	3774	11	1728

Продолжение табл.

G23	MACRO00002385	Scaffold_151776 (-)	7726	12	1530
G23	MACRO00002271	Scaffold_166891 (+)	7448	14	1947
H2	MACRO00015598	Scaffold_184754 (-)	259	2	198
mdr	MACRO00005670	Scaffold_157806 (+)	7493	22	3801
mdr 1	MACRO00001623	Scaffold_156325 (-)	7517	15	2409
mdr1	MACRO00005499	Scaffold_161591 (-)	3314	7	1173
mdr1	MACRO00006419	Scaffold_161800 (-)	2112	6	906
mdr1	MACRO00005876	Scaffold_164422 (+)	2621	8	1197
mdr1	MACRO00000343	Scaffold_164804 (-)	22212	28	4503
mdr 1	MACRO00002636	Scaffold_168537 (-)	6074	12	2061
mdr1	MACRO00005006	Scaffold_169760 (+)	4228	8	1212
mdr1	MACRO00003710	Scaffold_152153 (+)	2514	10	1812
mdr1	MACRO00017853	Scaffold_152132 (-)	2351	2	621
mdr 1A	MACRO00000777	Scaffold_170303 (+)	6600	21	3279
mdr1A	MACRO00002215	Scaffold_166270 (+)	3986	13	1995
mdr4	MACRO00000682	Scaffold_162556 (-)	19295	22	3840
mdr4	MACRO00000601	Scaffold_166690 (-)	7009	24	4026
mdr4	MACRO00000157	Scaffold_159141 (-)	7748	22	3798
mdr4	MACRO00000156	Scaffold_159141 (-)	11795	13	1794
mdr4	MACRO00003994	Scaffold_159646 (+)	4207	10	1548
mdr4	MACRO00000696	Scaffold_164506 (-)	5604	22	3714
mdr4	MACRO00000616	Scaffold_167669 (+)	8385	23	4017
mdr4	MACRO00000117	Scaffold_168154 (+)	23246	38	6378
mdr7	MACRO00014665	Scaffold_165398 (+)	852	3	711
mdr7	MACRO00007471	Scaffold_167420 (-)	1520	6	1080
mdr lethal	MACRO00005131	Scaffold_170856 (-)	1932	8	1128
mdr lethal	MACRO00000845	Scaffold_168251 (-)	10125	20	3420
mdr lethal	MACRO00000630	Scaffold_168346 (+)	12493	23	3951
sub A1	MACRO00011734	Scaffold_157494 (-)	12781	14	4080
subA3	MACRO00007998	Scaffold_163778 (-)	3557	5	1122
subA3	MACRO00004182	Scaffold_167170 (+)	9783	10	3177
sub A3	MACRO00011734	Scaffold_157611 (-)	4968	9	1866
subA5	MACRO00004861	Scaffold_161208 (+)	3700	9	2229

Окончание табл.

subB7	MACRO00005514	Scaffold_171561 (+)	7830	7	1236
subB8	MACRO00003066	Scaffold_165543 (+)	4350	11	1620
sub B10	MACRO00005670	Scaffold_157806 (+)	3343	7	1428
subD3	MACRO00011458	Scaffold_171034 (+)	1943	4	990
subE1	MACRO00008390	Scaffold_165801 (-)	2244	6	1803
subF1	MACRO00004295	Scaffold_169697 (-)	3374	9	1878
subG1	MACRO00004639	Scaffold_167900 (+)	8997	8	1461
subG1	MACRO00004814	Scaffold_168657 (-)	2430	8	1161
subG4	MACRO00003560	Scaffold_161549 (-)	3843	10	1566
subG4	MACRO00005106	Scaffold_169912 (-)	2796	9	1242

В результате было найдено 67 генов ABC-транспортеров. 25 из них относятся к семейству G, 3 – к семейству F, 1 – к семейству H, 1 – к семейству C, 1 – к подсемейству D, 6 – к подсемейству A, 1 – к подсемейству E, 8 – к генам множественной устойчивости 4 (multidrug resistance-associated protein 4), 4 – к летальным генам множественной устойчивости (probable multidrug resistance-associated protein lethal), 11 – к генам множественной устойчивости 1 (multidrug resistance-associated protein 1), 2 – к генам множественной устойчивости 7-подобные (multidrug resistance-associated protein 7-like), 4 – к подсемейству B.

Средняя длина гена – 5854 нуклеотидов, средняя длина экзонов – 1921 нуклеотидов, средняя длина интронов – 3933 нуклеотидов.

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the program to assemble spliced alignments. / Haas B.J., Salzberg S.L., Zhu W. [et al.] // Genome Biology. 2008. Vol. 9. P. R7. DOI: 10.1186/gb-2008-9-1-r7
2. InterProScan 5: genome-scale protein function classification. / Jones P., Binns D., Chang H., [et al.] // Bioinformatics. 2014. Vol. 30, № 9. P. 1236-1240.
3. Web Apollo: a web-based genomic annotation editing platform. / Lee E., Helt G., Reese J. [et al.] // Genome Biology. 2013. Vol. 14, № 8. P. R93. DOI: 10.1186/gb-2013-14-8-r93