

ПОИСК ГЕНОВ микроРНК В УЧАСТКАХ ГЕНОМА ВИРУСА ЯДЕРНОГО ПОЛИЭДРОЗА *Bombyx mori*, КОМПЛЕМЕНТАРНЫХ САМЫМ ПОЗДНИМ ГЕНАМ

Т. В. ШИРИНА, М. Т. БОБРОВСКАЯ, Э. А. КОЗЛОВ

Институт молекулярной биологии и генетики НАН Украины, Киев;
e-mail: e.a.kozlov@imbg.org.ua

Проведен поиск генов микроРНК, в участках генома вируса ядерного полиэдроза *Bombyx mori*, комплементарных самым поздним генам. Разработан алгоритм поиска *in silico* микроРНК в отдельно взятых участках генома бакуловирусов. Показано, что ген *orf4* вируса ядерного полиэдроза *Bombyx mori*, комплементарный гену *ph*, кодирует потенциальную микроРНК. Ген *p74*, комплементарный гену *p10*, кодирует зрелую микроРНК и потенциальную микроРНК. Ген *orf1629* кодирует две малые некодирующие РНК, комплементарные участкам мРНК полиэдрина в 5'-конце *orf*. На основании полученных результатов предполагается, что малые некодирующие РНК, комплементарные участкам мРНК полиэдрина, включаются в полиэдры.

Ключевые слова: вирус ядерного полиэдроза, *Bombyx mori*, биоинформативный метод, микроРНК.

Ранее биохимическими исследованиями было показано, что в полиэдры вируса ядерного полиэдроза (ВЯП) *Bombyx mori* включаются две малые РНК с молекулярной массой 17 и 21 кДа, в виде РНП комплексов [1]. Установлено, что с района генома ВЯП *B. mori*, содержащего самые поздние гены, считываются 24 альтернативных транскрипта, которые процессируются энзимом Droscha до 32-х малых РНК длиной 50–100 нуклеотидов (нт) [2]. Теоретически эти малые РНК могли бы включаться в полиэдры. Возникает вопрос, почему из 32-х малых РНК в полиэдры включаются только две. Мы предполагаем, что на этот вопрос можно ответить, опираясь на гипотезу Меклера [3]. Согласно этой гипотезе олигонуклеотиды, составленные из антикодонов, комплементарных кодонам, кодирующим соответствующие участки полипептидной цепи, будут взаимодействовать с этими участками, образуя прочные РНП-комплексы. На этом основании мы решили проверить, закодированы ли микроРНК на участке, комплементарном открытой рамке считывания полиэдрина (*orf ph*) генома ВЯП *B. mori*.

Настоящее исследование посвящено поиску микроРНК (microRNAs, miRNAs) в участках генома ВЯП *B. mori*, содержащих на минус цепи гены *orf 1629* и *p74*, комплементарные генам *ph* и *p10* соответственно.

Материалы и методы

Нуклеотидная последовательность генома ВЯП *B. mori* была взята из ICTVdB Management

(2006). 00.006.0.01. Nucleophyhedrovirus. In: ICTVdB – The Universal Virus Database, version 4. Büchen-Osmond, C. (Ed), Columbia University, New York, USA (<http://www.ncbi.nlm.nih.gov/ICTVdb/ICTVdb/00.006.0.01.htm>) [4]. Обратную комплементарную нуклеотидную последовательность получали при помощи программы Sequence Manipulation Suite (SMS) Reverse Complement [5]. Для поиска miRNAs использовали разработанный и описанный нами ранее алгоритм поиска miRNAs в отдельно взятых регионах генома ВЯП *Autographa californica* [6].

Результаты и обсуждение

Как известно, у ВЯП *A. californica* от позднего ТААГ промотора для гена *ph* транскрибируются три транскрипта длиной 1,16; 3,4 и 4,9 килобаз (kb). С минус цепи участка, комплементарного гену *ph*, синтезируется транскрипт длиной 3,2 kb, пересекающий ген *ph* в противоположном направлении, который содержит комплементарный гену *ph* ген *orf1629*, а также ген *orf603* [7]. У ВЯП *B. mori* сохраняется аналогичная конструкция этого участка генома, но отсутствует ген *orf603*. Данные по синтезу транскриптов с антипараллельных цепей ВЯП *B. mori* отсутствуют. Поэтому по аналогии с ВЯП *A. californica* мы полагаем, что у ВЯП *B. mori* также транскрибируются с плюс цепи три транскрипта с гена *ph* и с минус цепи – один транскрипт с гена *orf1629*, комплементарный гену *ph*. Величину транскрипта, включающего *orf 1629*, мы принимаем, анали-

зируя расположение 3,2kb-транскрипта у ВЯП *A. californica* [8, 9]. Вероятно этот транскрипт синтезируется от ТАТА промотора для *orf 1629* и заканчивается полиТ в непосредственной близости к концу *orf lef2* (длина такого транскрипта равна 3265 нт) [8]. По аналогии с ВЯП *A. californica* у ВЯП *B. mori* мы приняли длину транскрипта, включающего *orf 1629*, равного длине участка 2479-128290 (2530 нт) [8]. Этот транскрипт принят за мРНК *orf 1629*. Если принять за точку отсчета нуклеотид А в AUG кодоне, то расположение мРНК *orf 1629* записывается в координатах 2479-(-124). Поиск miRNAs в районе генома, комплементарном гену *ph*, осуществляли на участке 3430-128298 (участок 1), принятом нами ранее [2].

Второй исследуемый участок 110842-108248 (участок 2) на минус цепи включает ген *p74*, комплементарный гену *p10*. Если принять за точку отсчета нуклеотид А в AUG кодоне *orf p10*, то расположение выбранного участка записывается в координатах 2446-(-249).

На участке 1, содержащем *orf 1629*, расположено 11 ТАТА промоторов и 36 полиТ (не менее четырех). Исходя из этого, исследовали вторичную структуру 241-го альтернативного транскрипта (alternative transcripts, *alts*). Они содержали 29 шпичечно-петлевых структур (stem-loop structures, *sls*), из которых 18 *sls* (*sls1-sls18*) имели уникальные петли. Из нескольких *alts*, с *sls* уникальными петлями, за *pri-miRNA* (primary miRNA) мы приняли такой *alt*, который содержал одну или минимальное число *sls*. В табл. 1 приведены результаты исследования характеристик *sls1-sls18*.

На исследуемом участке, содержащем *orf p74*, расположено 14 ТАТА промоторов и 23 полиТ. Исследовали 123 *alts*, содержащих 24 *sls*, из которых 15 (*sls1-sls15*) имели уникальные петли. В табл. 2 приведены результаты исследования характеристик *sls1-sls15*.

Из 18 *sls*, обнаруженных на участке 1 (табл. 1), только одна (*sls6*) проходит фильтры всех использованных нами программ [6], кро-

Таблица 1. Характеристики шпичечных структур (*sls*) в *pri-miRNAs*, синтезируемых с участка 1 генома ВЯП *B. mori*, содержащего *orf 1629*

Sls	Локализация <i>pri-miRNA</i>	Старт	Длина, нт	Drosha	Real	E, -kcal/mol	Dicer
1	2	3	4	5	6	7	8
1	3425-2817	3411	77	-0,451	—	13,3	< -1,0
2	3425-3234	3383	62	-0,327	—	14,9	< -1,0
3	3415-3234	3318	51	0,007	—	5,8	< -1,0
4	3415-2825	3315	74	-0,331	—	9,5	< -1,0
5	3287-2000	3271	91	non	—	10,0	non
6	3425-2458	2870	78	-0,198	+	24,8	-0,847
7	2837-2450	2812	62	-0,511	—	15,0	< -1,0
8	3425-2458	2659	60	0,038	—	15,0	< -1,0
9	2628-2503	2584	60	-0,431	—	11,7	< -1,0
10	2837-2316	2491	60	-1,162	—	11,2	< -1,0
11	3201-2316	2472	50	-0,136	—	19,8	< -1,0
12	3201-2000	2257	48	0,121	—	11,0	< -1,0
13	2013-517	1969	65	-0,645	pseudo	22,1	< -1,0
14	1854-932	1434	51	-0,166	—	5,3	-0,488
15	1982-868	1381	79	-0,561	+	21,4	-0,603
16	2013-193	341	50	-0,657	pseudo	16,3	< -1,0
17	747-64	288	64	-0,417	pseudo	16,7	< -1,0
18	744-(-124)	129	53	-0,069	+	19,1	< -1,0

Жирным шрифтом выделены характеристики, проходящие фильтры соответствующих программ. Sls, выделенная жирным шрифтом, процессируется до *pri-miR-C*

Таблица 2. Характеристики шпилечных структур (sls) в pri-miRNA, синтезируемых с участка 2 генома ВЯП *B. mori*, содержащего orf p74

Sls	Локализация pri-miRNA	Старт	Длина, нт	Drosha	Real	E, -kcal/mol	Dicer
1	2	3	4	5	6	7	8
1	2446-597	2439	51	non	–	5,8	non
2	2446-1725	2426	69	0,316	–	10,8	-0,643
3	2345-1171	2303	75	-0,141	+	28,6	-0,632
4	2392-1533	2240	61	-0,464	+	21,0	-0,650
5	2392-1533	1768	79	non	pseudo	21,9	non
6	2392-1533	1619	70	0,012	pseudo	22,1	-0,836
7	1206-(-216)	977	106	-0,086	+	36,8	-0,112
8	1206-225	675	71	-0,297	–	14,1	-0,976
9	1206-474	592	62	non	–	15,7	non
10	414-151	407	103	0,146	+	19,2	<-1,0
11	326-(-37)	277	65	0,149	–	17,6	<-1,0
12	553-151	236	48	-0,330	–	19,4	<-1,0
13	326-(-216)	-7	88	-0162	–	15,7	<-1,0
14	1206-(-216)	-12	58	-0,759	–	3,1	<-1,0
15	-83-(-249)	-137	62	-0,125	–	8,9	<-1,0

Жирным шрифтом выделены характеристики, проходящие фильтры соответствующих программ. Sls, выделенные жирным шрифтом, процессируются до зрелой (sls7) и потенциальной miRNA (sls3). Фильтры программ (колонки 5–8) приведены в работе [6]

ме программы miRNA SVM, т.е. она не процессируется до зрелой miRNA. На основании этого мы считаем, что sls6 процессируется до предшественника (precursor, pre-miRNA), которую согласно программе Microprocessor SVM мы приняли за потенциальную (candidate, C) miRNA. Эта sls расположена в 3'-UTR мРНК orf 4 [9]. Поэтому предшественник обозначили как потенциальная pre-miR-2Corf4. Из 15 sls, обнаруженных на участке 2 (табл. 2) одна (sls7) процессируется в зрелую miR-4p74 и одна (sls3) процессируется только до pre-miR-3Cp74. Нумерация miRNAs является продолжением нумерации, описанной в работе [1]. Альтернативный транскрипт, содержащий зрелую miRNA или pre-miRNA, мы приняли за pri-miRNA.

Для поиска в исследуемых участках генома ВЯП *B. mori* возможных гомологов уже известных miRNAs использовали программу miRBase [10]. Поиск осуществляли в смысловых и антисмысловых цепях избранных нами участков. Из многочисленных результатов поиска, согласно программе, принимали только те, в которых найденная последовательность отличалась от известной miRNA не более чем на 4 нт. Таким путем мы обнаружили толь-

ко одну последовательность 109402-109424. В принятой нами системе координат это последовательность 906-928 на участке -249-2446 плюс цепи, включающем ген *p10*. Отсчет ведется от первого нуклеотида в AUG кодоне для orf p10. Найденная последовательность гомологична известной mdv-miR-M21 (рис. 1, а). С помощью программы RNAfold мы подобрали sls на участке 109389-109481 (893-985), в 5'-плече которой вписывается найденная последовательность (рис. 1, б). Подобранная sls проходит фильтр программ Microprocessor SVM (score -0,222) и RNAfold (E = -27,1 kcal/mol), но не проходит фильтры программ miPred и miRNA SVM (score -1,176). Кроме того, ни в одном из исследованных нами алтс в районе -249-2446, включающем участок 895-984, подобная sls не обнаружена. Согласно программе Microprocessor SVM один из центров расщепления sls 895-984 энзимом Drosha приходится на найденную последовательность (рис. 1, б). На основании вышесказанного мы считаем, что miRNA, найденная с помощью программы miRBase, на данном участке генома ВЯП *B. mori* не закодирована.

Таким образом, с учетом предыдущих исследований [1], в результате поиска miRNAs в комплементарных цепях на участках генома ВЯП *B. mori* 128298-3404 и 108248-110842 обнаружены 4 зрелые miRNAs и 3 потенциальные pre-miRNA-Cs. На рис. 2 приведены sls, из которых они процессируются. Нуклеотидные последовательности зрелых и потенциальных miRNAs приведены в табл. 3. На рис. 3 и рис. 4 приведены схемы участков генома ВЯП *B. mori*, с которых считываются pri-miRNAs с районов, содержащих гены *ph*, *orf1629* (рис. 3) и *p10*, *p74* (рис. 4). Очевидно, что мРНК, синтезируемые с этих участков, представляют собой мишени для соответствующих комплементарных miRNAs. Если время экспрессии miRNAs и комплементарных к ним мРНК совпадает, то мРНК должны расщепляться в центрах комплементарности. Для ВЯП *B. mori* нет данных по экспрессии рассматриваемых генов. Как мы показали в предыдущей публикации, у ВЯП *A. californica* действительно miRNAs, комплементарные к 3,2kb-транскрипту, расщепляют его [6].

Из 18-ти sls, приведенных в табл. 1, особый интерес вызывают sls17 и sls18 длиной 64 и 53 нт соответственно. Последовательности этих sls состоят из антикодонов, комплементарных кодонам, кодирующих участки 75Arg SerValLysProAspThrMetLysLeuIleValAsnTrpSer GlyLysGluPheLeuArg95 (sls17) и 26LeuIleLys

AsnAlaLysArgLysLysHisLeuValGluHisGluGln Glu42 (sls18) полипептидной цепи полиэдрина. Величины sls17 и sls18 коррелируют малыми РНК с молекулярной массой 21 и 17 кДа соответственно, обнаруженных в полиэдрах [1]. В контексте сказанного для таких олигонуклеотидов мы употребляем термин «комплементарный». Здесь нелишне будет вспомнить гипотезу Меклера о том, что «комплементарные» к участкам полипептидной цепи олигонуклеотиды образуют прочные специфические РНП-комплексы с ними [3]. Специфичностью можно объяснить, почему из множества малых РНК, процессируемых энзимом Drossha (табл. 1, 2), в полиэдр включаются только две, процессируемые из sls18 и sls17. Наиболее вероятно, что малые РНК из этих sls, образуют комплексы РНП50 и РНП31 соответственно [1]. Их мы обозначаем как малые некодирующие РНК (small non-coding, snc) sncRNA-1 и sncRNA-2 (табл. 1).

Интересно то, что согласно гипотезе структурно-функциональной организации полипептидной цепи полиэдрина [11] «комплементарная» к sncRNA-1 аминокислотная последовательность на участке 26-42 в N-концевом домене протеина, ответственном за протеин-нуклеиновые взаимодействия, содержит сайт ядерной локализации полиэдрина.

Ранее мы уже сравнивали нуклеотидные последовательности потенциальной ми-

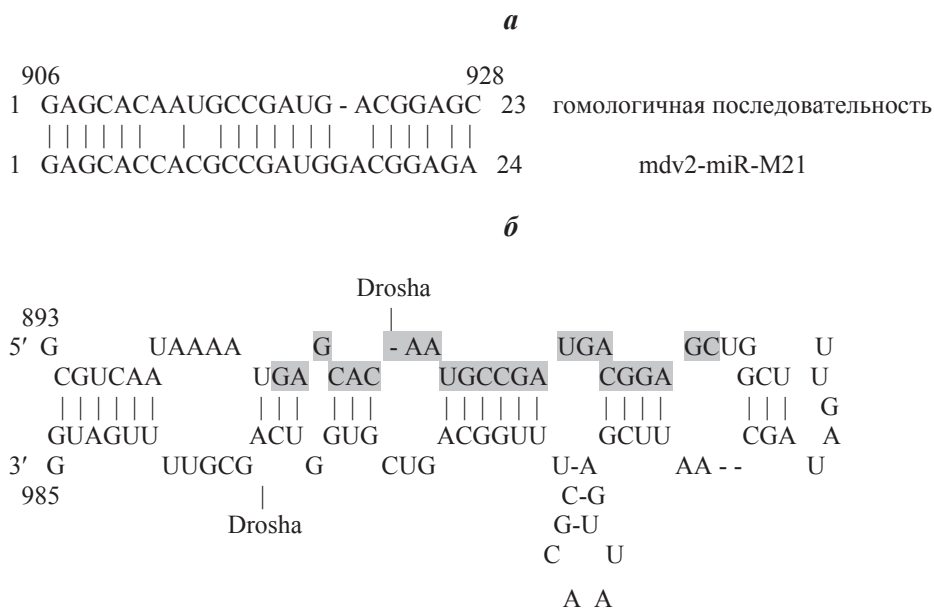


Рис. 1. а – Нуклеотидная последовательность 906-928 в гене *p10*, гомологичная нуклеотидной последовательности известной miRNA вируса болезни Марекка 2 – *mdv2-miR-M21*; б – шпилечная структура участка 893-985, включающая найденную гомологичную последовательность

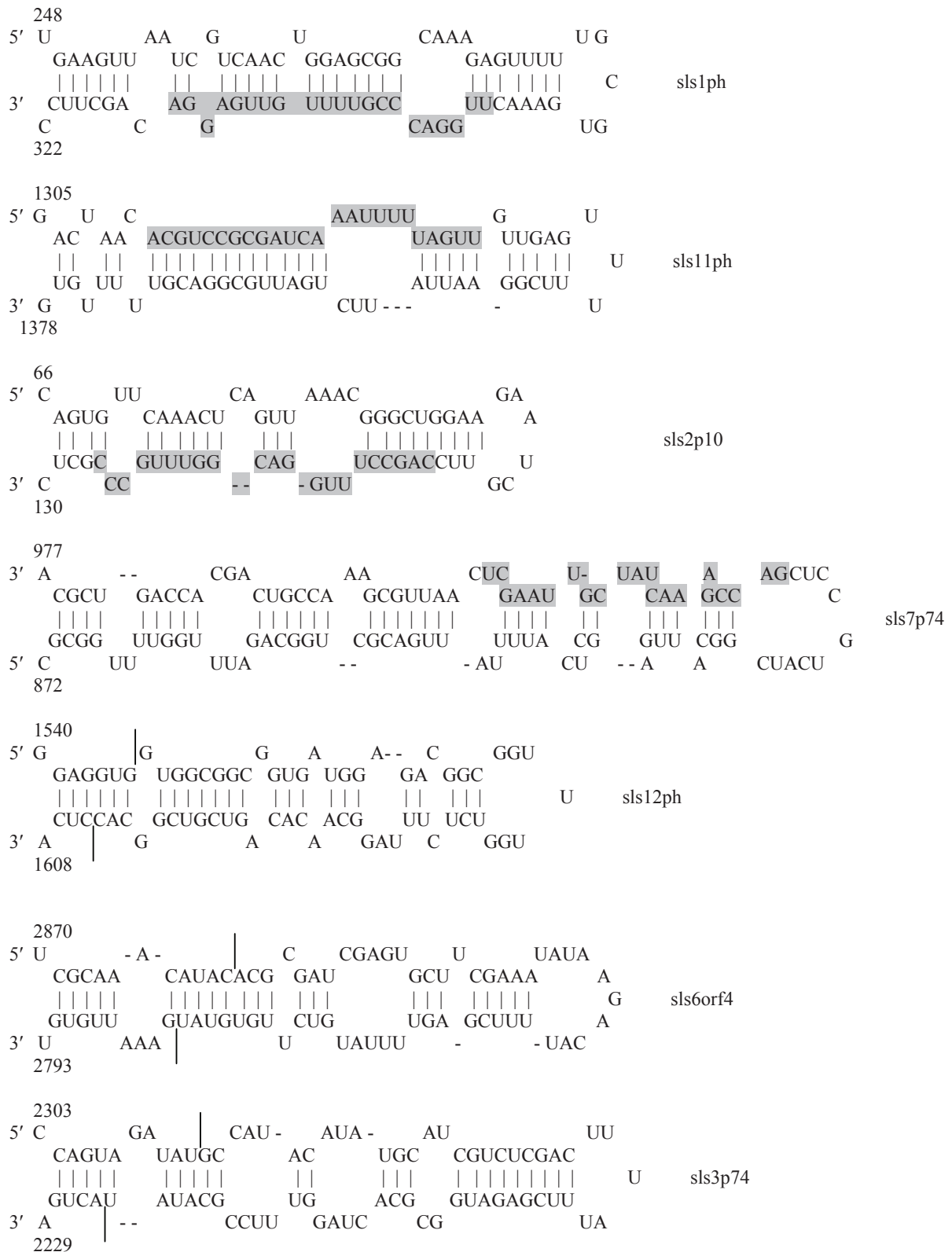


Рис. 2. Шпилечные структуры (sls) в составе соответствующих pri-miRNAs. Вертикальными черточками обозначены центры расщепления энзимом Droscha. Оттенены серым цветом зрелые miRNAs. Sls1ph, sls11ph, sls12ph и sls2p10 [1]

Таблиця 3. Нуклеотидні послідовності малих РНК (sRNA), закодированих в геномі ВЯП *B. mori*

sRNA	5'-3'-послідовність	Цепь	Локалізація в геномі
bmoNPV-mir-1ph	UUGGACCCGUUUUGUUGAGGA	+	294-314
-mir-2ph	UCCGCGAUCAAAUUUUUAGUU	+	1315-1335
-mir-3p10	CAGCCUUUGGACGGUUUGCCC	+	108602-108622
-mir-4p74	GACCGAAACUAUCGUUAAGCU	-	109444-109424
-pre-miR-1Cph	GUGGCGGCGGUGAUGCAGACGGCGGUUUG GGCUCUUUAGGCAACACAGUCGUCGGCAC	+	1547-1604
-pre-miR-2Corf4	ACGCGAUCGAGUGCUUCGAAAUUAAGAC AUUUUCGAGUUUUUAUGUCGUUGUGUAU	-	2858-2803
-pre-miR-3Cp74	GCCAUACAUAUGCAUCGUCUCGAGUUUAUU UCGAGAUGGCGCACUAGGUUUCGCAUUAU	-	110688-110630
sncRNA-1	GUCUUGUUCAUGUUCGACUAGGUGCUU CUUGCGCUUGGCGUUUUUGAUAAGAC	-	129-77
sncRNA-2	ACGCAAAAACUCUUUGCCGCUCCAGUUGACGA UUAACUUCAUGGUUUCGGUUUCACACUGCGA	-	288-225

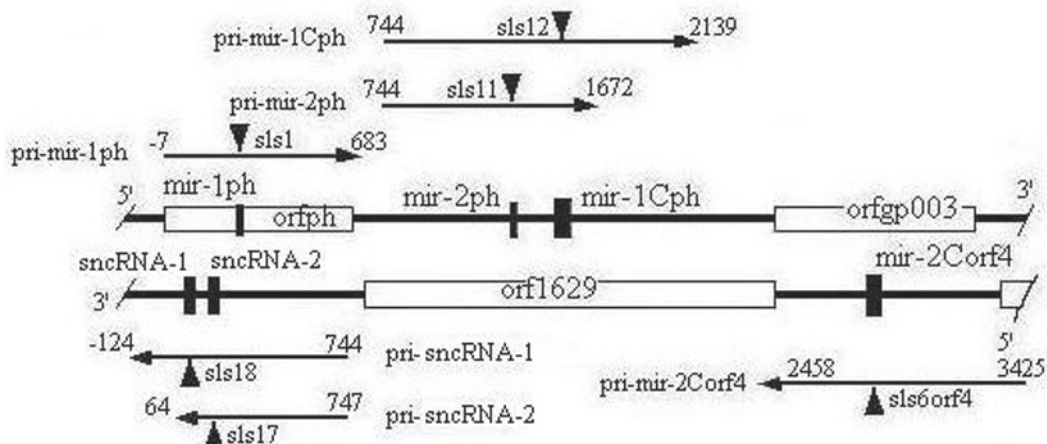


Рис. 3. Схема розташування miRNAs, miRNA-Cs та sncRNAs (■) на участку 1 геному ВЯП *B. mori*. Стрілками вказано розташування pri-miRNAs, що містять sls (▲), з яких процесуються відповідні малі РНК

кроРНК, розташованої в одному і тому ж районі гена *ph* ВЯП *B. mori* і *A. californica* [6]. При порівнянні отриманих даних нинішньої статті з аналогічними даними у ВЯП *A. californica* помічено більше відмінностей, ніж схожості як за кількістю закодированих зрілих і потенціальних miRNAs, так і за нуклеотидними послідовностями. Зрілі і потенціальні miRNAs, розташовані в різних районах одних і тих же ділянок 1 і 2, мають різні нуклеотидні послідовності. Це є підтвердженням загальновідомого факту неконсервативності

miRNAs у вірусів. Однак висока ступінь консервативності спостерігається між двома miRNAs, розташованими в одному і тому ж районі ділянки 2 (рис. 5). Можливо передбачити, що кожна з них регулює функцію загальну для двох вірусів.

Серед відмінностей інтригує одне — відсутність у ВЯП *A. californica* малих некодируючих РНК, комплементарних *orf ph* (sncRNA-1 і sncRNA-2). Цей факт передбачає відсутність малих РНК в поліедрах ВЯП *A. californica*. Дійсно, результати дослідження кристалічної структури поліедра

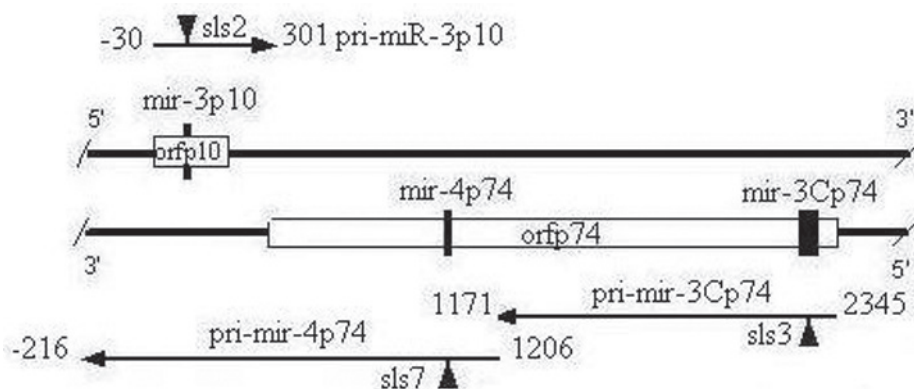


Рис. 4. Схема расположения miRNAs и miRNA-C (■) на участке 2 генома ВЯП *B. mori*. Стрелками указано расположение pri-miRNAs, содержащие sls (▲), из которых процессируются соответствующие малые РНК

UCGAAUUGCUAUCCAAGCCAG acaNPV-miR-6p74
UCGAAUUGCUAUCCAAGCCAG bmoNPV-miR-4p74

CAGCUUUUGGACGGUUUGCCC acaNPV-miR-3p10
CAGCUUUUGGACGGUUUGCCC bmoNPV-miR-3p10

Рис. 5. Сравнение нуклеотидных последовательностей зрелых miRNAs, расположенных в одинаковых районах участка 2 генома ВЯП *A. californica* и *B. mori*

ВЯП *A. californica* методом рентгеноструктурного анализа с разрешением 2,2 Å [12] показали, что полиэдры не содержат малых РНК.

Мы предполагаем, что обнаруженные, с помощью разработанного нами алгоритма, путем анализа вторичной структуры альтернативных транскриптов зрелые, потенциальные miRNAs, а также sncRNAs являются истинными. К такому же выводу пришли Ryu S. et al. [13], анализируя вторичную структуру pri-miRNA, содержащих 500 нт. Эти авторы считают предложенный ими анализ pri-miRNA реальным методом поиска bone fide miRNAs.

ПОШУК ГЕНІВ мікроРНК НА ДІЛЯНКАХ ГЕНОМУ ВІРУСУ ЯДЕРНОГО ПОЛІЕДРОЗУ *Bombyx mori* КОМПЛЕМЕНТАРНИХ НАЙПІЗНІШИМ ГЕНАМ

Т. В. Ширинa, М. Т. Бобровська,
Е. А. Козлов

Інститут молекулярної біології і
генетики НАН України, Київ;
e-mail: e.a.kozlov@imbg.org.ua

Проведено пошук генів мікроРНК у ділянках геному вірусу ядерного поліедрозу *Bombyx mori*, комплементарних найпізнішим генам. Розроблено алгоритм пошуку *in silico* мікроРНК в окремо взятих ділянках геному бакуловірусів. Показано, що ген *orf4* вірусу ядерного поліедрозу *Bombyx mori*, комплементарний гену *ph*, кодує потенційну мікроРНК. Ген *p74*, комплементарний гену *p10*, кодує зрілу мікроРНК і потенційну мікроРНК. Ген *orf1629* кодує дві малі некодуючі РНК, які комплементарні ділянкам мРНК поліедрина в 5'-кінці *orf*. На підставі одержаних результатів передбачається, що малі некодуючі РНК, які комплементарні ділянкам мРНК поліедрина, включаються в поліедри.

Ключові слова: вірус ядерного поліедрозу, *Bombyx mori*, біоінформативний метод, мікроРНК.

**THE SEARCH OF miRNA GENES
IN *Bombyx mori* NUCLEAR
POLYHEDROSIS VIRUS GENOMES
REGIONS COMPLEMENTARY
TO THE LATEST GENES**

*T. V. Shirina, M. T. Bobrovskaja,
E. A. Kozlov*

Institute of Molecular Biology and Genetics,
National Academy of Sciences of Ukraine, Kyiv;
e-mail: e.a.kozlov@imbg.org.ua

S u m m a r y

The search of miRNA genes in *Bombyx mori* nuclear polyhedrosis virus genome region complementary to very late genes has been carried out. The search miRNA algorithm *in silico* was developed by us. It was shown that NPV *B. mori* genome region containing *orf4* gene complementary to *ph* gene encodes the potential miRNA. NPV *B. mori* genome region containing *p74* gene complementary to *p10* gene encodes mature miRNA and potential miRNA. The genome region containing *orf1629* encodes two small non-coding RNAs complementary to *orf 5'*-end of polyhedrin miRNA. From obtained results it is proposed that two small non-coding RNAs complementary to regions of polyhedrin miRNA are included in polyhedra.

Key words: nuclear polyhedrosis virus, *Bombyx mori*, bioinformatic search, miRNAs.

1. *Ширина Т. В., Бобровская М. Т., Козлов Э. А.* // Укр. біохім. журн. 2010. — **82**, № 6. — С. 87–92.
2. *Ширина Т. В., Висловух А. А., Бобровская М. Т., Козлов Э. А.* // Biopolym. Cell. — 2009. — **25**, № 3. — Р. 226–223.
3. *Меклер Л. Б., Идлис Р. Г.* // Биофизика. — 1981. — **26**, № 3. — С. 574–575.
4. *Gomi S., Majima K., Maeda S.* // J. Gen. Virol. — 1999. — **80**, N 5. — Р. 1323–1337.
5. *Stothard P.* // Biotechniques. — 2000. — **28**. — Р. 1102–1104.
6. *Ширина Т. В., Бобровская М. Т., Козлов Э. А.* // Укр. біохім. журн. — 2011. — **83**, № 4. — С. 41–49.
7. *Ooi B. G., Miller L. K.* // J. Virol. — 1990. — **64**, N 6. — Р. 3126–3129.
8. *Ayres M. D., Howard S. C., Kuzio J. et al.* // Virology. — 1994. — **202**, N 2. — Р. 586–605.
9. *Kool M., Vlak J. M.* // Arch. Virol. — 1993. — **130**. — Р. 1–16.
10. *Griffiths-Jones S, Saini H. K., van Dongen S., Enright A. J.* // Nucleic Acids Res. — 2008. — **36**. — Р. D154–D158.
11. *Kozlov E. A., Levitina T. L., Gusak N. M.* // Curr. Top. Microbiol. Immunol. — 1986. — **131**. — Р. 135–164.
12. *Coulibaly F., Chiu E., Gutmann S. et al.* // Proc. Nat. Acad. Sci. USA. — 2009. — **106**, N 52 — Р. 22205–22210.
13. *Ryu S., Joshi N., McDonnell K. et al.* // PLoS One. — 2011. — **6**, N 2. — Р. e16403.

Получено 9.06.2011