Genomics 102 (2013) 63–71

Contents lists available at SciVerse ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Genomic diversity of *Bombyx mori* nucleopolyhedrovirus strains

Yi-Peng Xu [a,b], Ruo-Lin Cheng [b], Yu Xi [b], Chuan-Xi Zhang [b,*]

[a] Zhejiang Provincial Key Laboratory of Biometrology and Inspection & Quarantine, College of Life Sciences, China Jiliang University, Hangzhou 310018, China
[b] Institute of Insect Science, Zhejiang University, Hangzhou 310058, China

ARTICLE INFO

Article history:
Received 8 March 2013
Accepted 18 April 2013
Available online 29 April 2013

Keywords:
Bombyx mori nucleopolyhedrovirus
Genome diversity
Evolution
Comparative genomics

ABSTRACT

Bombyx mori nucleopolyhedrovirus (BmNPV) is a baculovirus that selectively infects the domestic silkworm. In this study, six BmNPV strains were compared at the whole genome level. We found that the number of *bro* genes and the composition of the homologous regions (hrs) are the two primary areas of divergence within these genomes. When we compared the ORFs of these BmNPV variants, we noticed a high degree of sequence divergence in the ORFs that are not baculovirus core genes. This result is consistent with the results derived from phylogenetic trees and evolutionary pressure analyses of these ORFs, indicating that ORFs that are not core genes likely play important roles in the evolution of BmNPV strains. The evolutionary relationships of these BmNPV strains might be explained by their geographic origins or those of their hosts. In addition, the total number of hr palindromes seems to affect viral DNA replication in Bm5 cells.

1. Introduction

Baculoviruses are arthropod-specific enveloped rod-shaped viruses that contain large double-stranded circular DNA genomes ranging from 80 to 180 kb in length. They are characterized by a biphasic life cycle, forming two types of virions, occlusion-derived virions (ODVs) and budded virions (BVs) [1]. During primary oral infection, occlusion bodies (OBs), which contain ODVs, dissolve in the alkali insect midgut and release the ODVs to infect the midgut cells. After this stage, the systemic infection is initiated by BVs, and the virus spreads from cell to cell.

To date, hundreds of baculoviruses have been described, and approximately 60 species (strains) have been completely sequenced. Among these baculoviruses, there are 31 common genes, called core genes [2]. Comparisons of these core genes indicate that the family Baculoviridae comprises four genera: Alphabaculovirus (lepidopteran-specific NPV), Betabaculovirus (lepidopteran-specific Granuloviruses), Gammabaculovirus (hymenopteran-specific NPV), and Deltabaculovirus (dipteran-specific NPV) [2]. Alphabaculovirus can be further divided into two groups, Group I and Group II. New genome-sequencing technologies have laid the foundation for a boom in sequencing, and more and more baculovirus genomes are being sequenced. However, only a few reports of comparisons among strains of the same viral species based on whole genome sequences have been published [3–5], and many comparative genomic analyses have been based on only the restriction endonuclease mapping [4,6–8]. Recently, we reported a variant of *Bombyx mori* NPV (BmNPV), *Bombyx mandarina* NPV S1 (BomaNPV S1), and published the genomes of two BmNPV strains (the BmNPV Indian strain and BmNPV Cubic strain) [9,10]. All these strains are members of the Alphabaculovirus genus. In the current study, the Indian strain and Cubic strain, along with two other newly sequenced strains (the BmNPV Guangxi strain and BmNPV Zhejiang strain), were compared with BomaNPV S1 and BmNPV T3 to characterize the diversity and phylogenic evolution of BmNPV strains at the whole genome level. Additional analyses of the other characteristics of these viruses are also presented.

2. Results

2.1. Comparison of genomes and ORFs

The genomes of four BmNPV strains, T3 (GenBank accession L33180.1), Boma S1 (GenBank accession FJ882854.1), Cubic (GenBank accession JQ991009) and Indian (GenBank accession JQ991010), were sequenced previously [9–12], and the genomes of two additional BmNPV strains, Zhejiang (GenBank accession JQ991008) and Guangxi (GenBank accession JQ991011), were successfully sequenced in this study. These six viral genomes have similar sizes (from 126,125 bp for BmNPV Zhejiang to 126,879 bp for BmNPV Indian) and G + C contents (~40.3%). The size difference is mostly due to the divergence in the *bro* genes and the homologous repeat regions (hrs) (Table 1). Among these six virus strains, the T3 strain, Cubic strain and Guangxi strain contain three subgroups of *bro* genes, whereas the others only have two (Table 2), according to the classification by Kang et al. [13]. Interestingly, the Guangxi strain and Cubic strain only have one *bro* gene in subgroup C, whereas the T3 strain has two. All six virus strains have eight hr regions but only the number of palindrome units in hr4b and hr4c is conserved among all the viruses.

These six virus strains have 134 common ORFs that encode at least 60 amino acids, with the exception of *bro-b* in T3 and Guangxi, *bro-e*

* Corresponding author. Fax: +86 571 88982991.
E-mail address: chxzhang@zju.edu.cn (C.-X. Zhang).

0888-7543/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ygeno.2013.04.015

Table 1
Genomic comparison of six BmNPV strains.

BmNPV-T3 ORFs	AcMNPV ORFs	Gene name	Length (aa)						PAACS (%)	dN/dS	Number of positively selected sites (P > 95%)		Number of mutations
			Boma S1	Cubic	Guangxi	Indian	T3	Zhejiang			NEB	BEB	
1	8	*polyhedrin*	245	245	245	245	245	245	100.00				0
2	9	*orf1629*	542	545	541	542	542	542	94.71	0.3985	6	6	14
3	10	*Pk1*	274	274	275	275	275	275	98.18	0.3159	1	1	5
4	11	*Ac11*	340	340	340	340	340	340	97.35	0.4435	9	Non	9
5	13	*Ac13*	331	331	331	331	331	331	96.68	0.3739	11	Non	11
		Bm5RC	153	153	153	153	153	153	94.12				
6	14	*lef-1*	270	270	270	270	270	270	98.89	0.2218	3	Non	3
7	15	*egt*	506	506	506	506	506	506	97.23	0.4407	14	Non	14
8	16	*bv/odv-e26*	229	229	229	229	229	229	97.82	0.2172	5	Non	5
9	17	*Ac17*	210	210	210	211	210	210	96.19	0.3496	8	Non	8
10	18	*Ac18*	356	356	358	357	356	356	96.91	1.2681	2	1	11
11	19	*Ac19*	110	110	111	110	110	110	99.09	0.3236	Non	Non	1
12	20/21	*arif-1*	443	447	447	448	440	447	96.59	0.3101	3	1	16
13	22	*pif-2*	382	382	382	382	382	382	97.91	0.2728	8	Non	8
14	23	*f*	686	673	686	673	673	673	98.66	0.3388	1	1	9
		Bm14RC	135	135	135	135	135	135	97.78				
15	24	*pkip*	169	169	169	169	169	169	98.82	0.5922	3	Non	2
16	25	*dbp*	317	317	317	317	317	317	100.00				0
17	26	*Ac26*	129	129	129	129	129	129	99.22	0.089	Non	Non	1
18	27	*iap1*	289	288	292	289	292	292	96.53	0.2263	Non	Non	14
19	28	*lef-6*	173	173	173	173	173	173	98.84	0.2722	2	Non	2
20	29	*Ac29*	71	71	71	71	71	71	100.00				0
21	30	*Ac30*	472	472	472	472	472	472	98.73	0.2701	6	Non	6
22		*bro-a*	330	316	330	330	317	316	82.91	0.384	2	2	54
23	31	*sod*	151	151	151	151	151	151	96.69	0.232	Non	Non	5
		hr2L (bp)	874	1027	70	885	867	231					
24	32	*fgf*	182	182	182	182	182	182	97.25	0.2633	5	Non	5
		hr2R (bp)	288	287	287	188	287	278					
25	34	*Ac34*	215	215	215	215	215	215	96.28	0.3653	Non	Non	8
26	35	*ubiquitin*	77	77	77	77	77	77	100.00				0
27	36	*39k*	277	277	277	278	277	277	98.92	0.0938	Non	Non	3
28	37	*lef-11*	112	112	112	112	112	112	98.21	0.3934	Non	Non	2
29	38	*bv-e31*	217	217	217	217	217	217	100.00				0
30	39	*p43*	362	362	362	362	362	362	95.86	0.9835	15	Non	15
31	40	*p47*	399	399	399	399	399	399	98.50	0.3865	6	Non	6
32	41	*lef-12*	177	177	177	177	183	177	90.96	1.8457	1	2	16
33	42	*gta*	506	506	506	506	506	506	98.62	0.1706	7	Non	7
		Bm33RC	197	197	197	197	197	197	98.98				
34	43	*Ac43*	78	78	78	78	78	78	93.59	1.1037	5	Non	5
35	44	*Ac44*	131	131	131	131	131	131	99.24	0.0226	Non	Non	1
36	45	*Ac45*	195	194	194	192	193	193	97.40	0.6672	5	Non	5
37	46	*odv-e66*	709	708	708	708	702	708	98.15	0.2918	Non	3	13
38	47	*ets*	89	102	89	89	89	89	94.38	0.2858	Non	Non	5
39	50	*lef-8*	877	877	877	877	877	877	99.32	0.055	6	Non	6
40	51	*Ac51*	319	319	319	319	319	319	98.75	0.2611	1	Non	4
41	52	*Ac52*	194	194	194	194	194	194	99.48	0.0422	Non	Non	1
42	53	*Ac53*	139	139	139	139	139	139	100.00				0
42a	53a	*lef-10*	78	78	78	78	78	78	98.72	Out	D2	D2	1
43	54	*vp1054*	365	365	365	365	365	365	100.00				
44	55	*Ac55*	77	75	77	77	77	77	97.33	0.7534	2	Non	2
45	56	*Ac56*	84	84	84	84	84	84	100.00				0
46	57	*Ac57*	161	161	161	161	161	161	98.76	0.4126	2	Non	2
47	58/59	*Ac58/59*	171	171	171	171	171	171	98.83	0.1167	Non	Non	2
48	60	*Ac60*	83	83	83	82	83	83	98.78	2.0994	1	1	1
49	61	*fp25*	214	214	214	214	214	214	98.13	0.4531	Non	Non	4
50	62	*lef-9*	490	490	490	490	490	490	99.59	0.0458	Non	Non	2
51	63	*Ac63*	155	155	155	155	155	155	98.71	0.4424	2	2	2
52	64	*gp37*	294	294	294	294	294	294	97.28	0.8107	3	3	8
53	65	*dnapol*	989	987	989	988	986	988	98.98	0.2822	2	2	10
54	66	*Ac66*	805	802	804	804	805	805	97.63	0.1921	19	Non	19
55	67	*lef-3*	385	385	385	385	385	385	99.22	0.0837	3	Non	3
56	68	*odv-nc42*	134	134	134	134	134	134	98.51	0.1138	Non	Non	2
57	69	*Ac69*	262	262	262	262	262	262	98.47	0.4725	4	Non	4
58	71	*iap2*	249	249	249	249	249	249	97.19	0.3534	7	2	7
58a	72	*Ac72*	60	60	60	60	60	60	98.33	Out	D2	D2	1
59	73	*Ac73*	99	99	99	99	99	99	92.86	4.2811	8	2	8
60	74	*Ac74*	267	269	268	269	268	268	98.50	0.8366	4	Non	4
61	75	*Ac75*	133	133	133	133	133	133	99.25	0.3176	Non	Non	1
62	76	*Ac76*	85	85	85	85	85	85	100.00				0
63	77	*vlf-1*	379	381	379	377	379	379	99.47	0.079	Non	Non	2
64	78	*Ac78*	110	110	110	110	110	110	97.27	0.7427	Non	Non	3

Table 1 (*continued*)

BmNPV-T3 ORFs	AcMNPV ORFs	Gene name	Length (aa)						PAACS (%)	dN/dS	Number of positively selected sites (P > 95%)		Number of mutations
			Boma S1	Cubic	Guangxi	Indian	T3	Zhejiang			NEB	BEB	
65	79	*Ac79*	104	104	104	104	104	104	99.04	2.0777	1	Non	1
66	80	*gp41*	403	403	403	401	403	403	97.01	0.4372	1	2	12
67	81	*Ac81*	234	234	234	235	234	234	99.57	0.457	1	1	1
68	82	*Ac82*	181	181	181	182	181	181	98.34	0.258	Non	Non	3
69	83	*p95 = vp91*	837	837	837	835	839	*837*	*97.99*	0.3016	16	Non	16
		Bm69RC	135	135	135	135	135	135	97.78				
		hr3 (bp)	568	402	568	569	570	555					
70	87	*vp15*	126	126	126	126	126	126	100.00				0
71	88	*cg30*	267	266	267	267	267	267	98.12	0.3051	Non	Non	5
72	89	*vp39*	347	350	348	348	350	350	97.41	0.5355	9	1	9
73	90	*lef-4*	465	465	465	465	465	465	99.35	0.1753	3	Non	3
		Bm73RC	154	154	154	154	154	154	98.70				
74	91	*Ac91*	158	157	161	153	154	154	94.12	0.2546	9	Non	9
75	92	*p33*	259	259	259	259	259	259	100.00				0
76	93	*Ac93*	161	161	161	161	161	161	100.00				0
77	94	*odv-e25*	228	228	228	228	228	228	99.12	0.0825	Non	Non	2
78	95	*dnahel*	1222	1222	1222	1222	1222	1222	99.51	0.0441	Non	Non	6
79	96	*odv-e28*	182	182	182	173	182	173	97.11	0.2097	4	Non	5
80		*bro-b*			241		239						17
81		*bro-c*	325	331	327	322	318	318	87.33	Out	Out	Out	39
82	98	*38k*	320	320	320	320	320	320	98.13	0.2588	Non	Non	6
83	99	*lef-5*	265	265	265	265	265	265	99.62	0.0413	Non	Non	1
84	100	*p6.9*	65	65	65	65	65	65	95.38	1.0108	65	Non	3
85	101	*p40*	362	362	360	363	362	362	98.06	0.2313	5	Non	7
86	102	*p12*	125	125	123	125	123	123	99.19	0.1381	Non	Non	1
87	103	*p45*	387	387	387	387	387	387	99.22	0.1599	Non	Non	3
88	104	*vp80*	693	694	692	694	692	692	98.70	0.3027	9	5	9
89	105	*he65*	289	289	289	289	289	289	99.31	1.3808	1	1	8
		hr4a (bp)	481	480	480	624	337	553					
90	106/107	*Ac106/107*	249	249	249	249	249	249	97.99	0.1899	Non	Non	5
91	108	*Ac108*	105	105	105	105	105	105	98.10	Out	Out	Out	2
92	109	*Ac109*	391	391	391	391	391	391	99.49	0.1123	2	Non	2
93	111	*Ac111*	67	67	67	67	67	67	100.00				0
		hr4b (bp)	527	527	427	527	527	527					
94	114	*Ac114*	424	424	424	424	424	424	98.82	0.2548	5	Non	5
95	115	*pif-3*	204	204	204	204	204	204	97.06	0.2906	Non	Non	6
96	117	*Ac117*	95	95	95	95	95	95	98.95	0.0749	Non	Non	1
97	119	*pif-1*	527	527	527	527	527	527	98.86	0.7869	26	Non	26
		Bm97RC	142	142	142	142	142	142	92.96				
98	120	*Ac120*	82	82	82	82	82	82	98.78	Out	D2	D2	1
		hr4c (bp)	111	111	111	110	111	111					
99	122	*Ac122*	61	61	61	61	61	61	98.36	0.3471	Non	Non	1
100	123	*pk2*	225	225	225	225	225	225	98.22	0.4474	1	1	4
101	124	*Ac124*	244	244	244	244	244	244	96.31	0.6961	9	Non	9
102	125	*lef-7*	227	227	227	227	227	227	95.15	0.5057	11	Non	11
103	126	*chitinase*	551	551	552	552	552	552	98.37	0.1705	9	Non	9
104	127	*v-cath*	323	323	323	323	323	323	98.76	0.1302	4	Non	4
105	128	*gp64/67*	530	529	530	530	530	530	99.24	0.0839	1	Non	4
		Bm105RC	209	209	209	209	209	209	93.30				
106	129	*p24*	195	195	195	195	195	195	98.46	0.3195	Non	Non	3
107	130	*gp16*	106	106	106	106	106	106	100.00				0
108	131	*pp34*	315	314	314	315	315	315	98.73	0.1371	Non	Non	4
		Bm108-109RC	149	148	148	149	149	149	97.97				
109	132	*Ac132*	220	220	220	220	220	220	97.27	0.4724	Non	Non	6
110	133	*alk-exo*	420	420	420	420	420	420	98.57	0.1261	5	Non	6
111§			56	58	<50	<50	70	71					
112	135	*p35*	229	299	299	299	299	299	97.66	0.5219	7	Non	7
		hr5 (bp)	776	600	600	885	775	600					
113	136	*p26*	240	240	240	240	240	240	97.92	0.4195	5	Non	5
114	137	*p10*	70	70	70	70	70	70	98.57	0.1164	Non	Non	1
115	138	*p74 = pif*	645	645	645	645	645	645	98.60	0.4628	9	Non	9
116	139	*me53*	452	453	451	451	451	451	96.45	0.8594	15	Non	16
		Bm116-117RC	67	67	67	67	67	67	98.51				
117	141	*ie-0*	261	261	261	261	261	261	99.23	0.1512	Non	Non	2
		Bm117-118RC	136	136	136	136	136	136	100.00				
118	142	*bv/odv-nc50*	476	476	476	476	476	476	99.79	0.0257	Non	Non	1
119	143	*odv-e18*	101	101	101	102	101	101	98.02	Out	D2	D2	2
120	144	*odv-ec27*	290	290	290	290	290	290	99.66	0.0332	Non	Non	1
121	145	*Ac145*	95	95	95	95	95	95	100.00				0
122	146	*Ac146*	201	201	201	201	201	201	98.51	0.1261	Non	Non	3
123	147	*ie-1*	584	584	584	584	584	584	98.46	0.1848	9	Non	9

(*continued on next page*)

Table 1 (*continued*)

BmNPV-T3 ORFs	AcMNPV ORFs	Gene name	Length (aa)						PAACS (%)	dN/dS	Number of positively selected sites (P > 95%)		Number of mutations
			Boma S1	Cubic	Guangxi	Indian	T3	Zhejiang			NEB	BEB	
124	148	*odv-e56*	375	374	375	377	375	375	96.79	0.6842	Non	Non	12
125	149	*Ac149*	106	106	106	106	106	106	98.11	2	D2	D2	2
126	150	*Ac150*	115	113	115	105	115	115	97.14	0.3875	Non	Non	3
127	151	*ie-2*	422	422	422	421	422	422	96.44	1.2762	1	1	15
128	153	*pe38*	309	309	309	309	309	309	96.76	0.2214	10	Non	10
129	154	*Ac154*	77	78	77	78	77	77	98.70	5.7072	1	Non	1
		hr1 (bp)	661	611	666	678	676	611					
130	1	*ptp*	168	168	168	168	168	168	95.83	0.3325	7	Non	7
131	2	*bro-d*	348	348	349	348	349	349	88.18	0.4429	17	11	41
132		*bro-e*		241			241						11
133	4	*Ac4*	151	151	151	151	151	151	96.69	0.2671	Non	Non	5
134	5	*Ac5*	109	109	109	109	109	109	97.25	0.1661	Non	Non	3
135	6	*lef-2*	210	210	210	210	210	210	98.57	0.2084	3	Non	3

PAACS, NEB and BEB are abbreviations of "percentage of amino acid constant sites", "Naive Empirical Bayes analysis" and "Bayes Empirical Bayes analysis".
"Out" in dN/dS analysis means its value can't not be calculated by Codeml of PAML software, because no synonymous substitution exists in homologs or homologous proteins are not high identical. "D2" means only two homologs exist within six viruses and the value of dN/dS is calculated manually if synonymous substitutions exist.

in Cubic and T3, and *orf111* in T3 and Zhejiang, which is truncated in the other four strains. Using the percentage of amino acid constant sites (PAACS) as a criterion for the evaluation of ORF conservation, most of the ORFs show high identity. A total of 15 ORFs (*polyhedrin*, *dbp*, *Bm20*, *ubiquitin*, *bv-e31*, *Bm42*, *vp1054*, *Bm45*, *Bm62*, *vp15*, *p33*, *Bm76*, *Bm93*, *gp16*, and *Bm121*) are 100% identical. However, the PAACS of 9 ORFs (*orf1629*, *bro-a*, *lef-12*, *Bm34*, *ets*, *Bm59*, *Bm74*, *bro-c*, and *bro-d*) are under 95%, and the PAACS of *bro-a* (82.91%), *bro-c* (87.33%), and *bro-d* (88.18%) are especially low. Predicted ORFs encoding fewer than 60 amino acids were not used in this comparison because these small ORFs are often not present in all six virus strains and may be disrupted in the middle of the sequence or truncated at one terminus.

Analysis with Softberry online service revealed ten additional new ORFs in these six virus strains (Table 1). These ORFs are present in the reverse complement sequences of *Bm5* (*Bm5RC*), *f* (*Bm14RC*), *gta* (*Bm33RC*), *p95* (*Bm69RC*), *lef-4* (*Bm73RC*), *pif-1* (*Bm97RC*), *gp64* (*Bm105RC*), *pp34-Bm109* (*Bm108-109RC*), *me53-ie-0* (*Bm116-117RC*), and *ie-0-bv/odv-nc50* (*Bm117-118RC*). *Bm33RC* has an early promoter motif (a TATA box followed by CAGT) in its upstream region, and *Bm97RC* has a late promoter motif (GTAAG), whereas *Bm105RC* and *Bm117-118RC* possess an enhancer-like motif (CGTGC). The existence of transcripts corresponding to these ORFs was confirmed by RT-PCR (reverse transcription-polymerase chain reaction) (data not shown). This phenomenon was also observed in the transcriptome sequencing that we performed in this study. Several of these novel predicted ORFs encode more than 100 aa and have transcriptional regulatory motifs, suggesting that they may be functional and should be characterized.

2.2. Evolution of the six virus strains

The genomic evolutionary tree of the six virus strains was constructed using MEGA5 software with the maximum likelihood method based on the cascading DNA comparison of 132 common ORFs, using AcMNPV (which lacks *bro-a* and *bro-c*) as an outgroup (Fig. 1). BomaNPV S2 (Boma S2), which was isolated from the same host and the same location as Boma S1 but shares the host ranges of *B. mori* NPV and *Autographa californica* MNPV in cultured cells [14], was also included in this tree. As shown in the phylogenetic tree, the Guangxi strain clusters with the Zhejiang stain near the T3 strain, and this cluster are far from the others. The Boma S1 and Cubic strains, which were isolated from the same location, are more closely related to each other than to the other strains. Boma S2 is more closely related to the Cubic strain and the Boma S1 strain than to the other strains. The Indian strain is positioned between the T3 strain and Boma S1 but is more closely related to the T3 strain than to Boma S1.

To assess the evolutionary pressure on each ORF, the dN/dS value was calculated using PAML (Table 1). As shown in Fig. 2, 11 genes (*Bm10*, *lef-12*, *Bm34*, *Bm48*, *Bm59*, *Bm65*, *p6.9*, *he65*, *Bm125*, *ie-2*, and *Bm129*) had dN/dS values greater than 1, demonstrating that these genes have been positively selected during evolution. Among these genes, only *p6.9*, which encodes a protein that binds to viral DNA [15], belongs to the core set of baculovirus genes. The homologs of *Bm59* and *Bm65* were found to encode a BV-associated protein and ODV-associated protein in AcMNPV, respectively [16,17]. It is noteworthy that *lef-12*, *Bm34* and *Bm59* had identities of less than 95%. Therefore, these three genes are not appropriate to use as reference genes for the analysis of the evolution of different viral species given their high levels of variability in a single species. Positively selected sites (P > 95%) in each ORF were also analyzed using the Naive Empirical Bayes (NEB) and Bayes Empirical Bayes (BEB) methods (Table 1).

Table 2
Distribution of *bro* genes in six BmNPV strains.

Virus strains	Subfamily		
	Subgroup A	Subgroup B	Subgroup C
Boma S1	*bro-d*	*bro-a*, *bro-c*	
Cubic	*bro-d*	*bro-a*, *bro-c*	*bro-e*
Guangxi	*bro-d*	*bro-a*, *bro-c*	*bro-b*
Indian	*bro-d*	*bro-a*, *bro-c*	
T3	*bro-d*	*bro-a*, *bro-c*	*bro-b*, *bro-e*
Zhejiang	*bro-d*	*bro-a*, *bro-c*	

Subfamily was referred to Kang et al. [13].

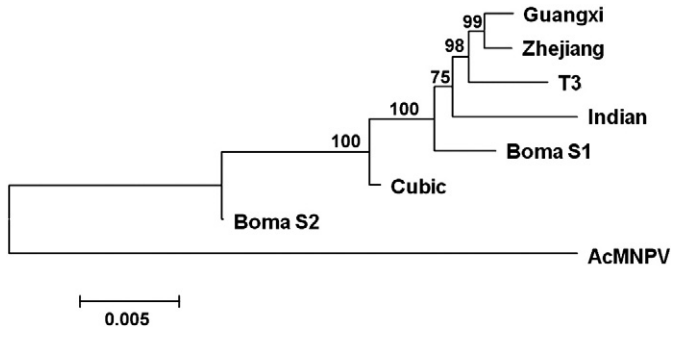

Fig. 1. A phylogenetic tree of various BmNPV strains based on a cascading DNA comparison of 132 common ORFs. Two closely related NPVs (AcMNPV and Boma S2) were included. This tree was constructed in MEGA5 using the maximum likelihood method and Tamura–Nei substitution model with 1000 bootstrap replicates.

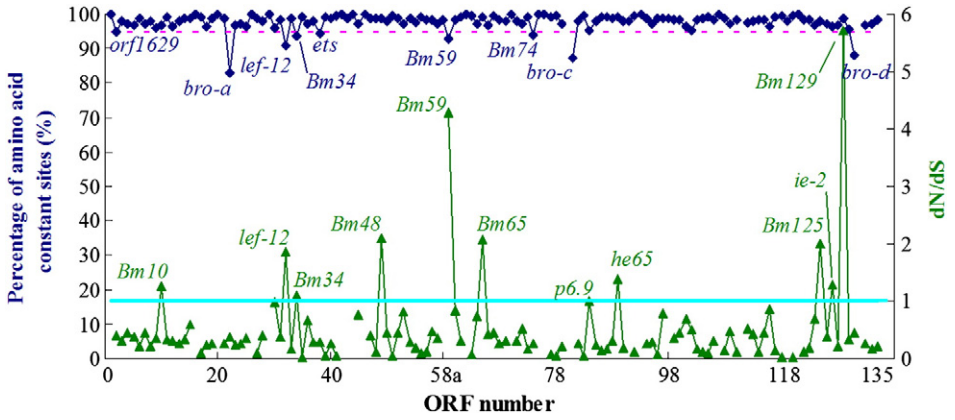

Fig. 2. Amino acid sequence homology and evolutionary pressure on different ORFs. The ORF order of the BmNPV T3 strain was used as the reference. *bro-b*, *Bm111§*, *bro-e* and ORFs encoding small peptides of less than 60 aa were not analyzed. ORFs that encoded the same predicted peptide in all six virus strains were excluded from the dN/dS analysis. *Bm125* has only two homologs in the six virus strains, and they possess two nonsynonymous substitutions and one synonymous substitution, for a dN/dS = 2. *bro-a*, *bro-c* and *bro-d* are so different between the six virus strains that the value of dN/dS cannot be calculated. ORFs (*Bm42a*, *Bm58a*, *Bm91*, *Bm98* and *Bm119*) without synonymous substitutions have no dN/dS values.

Twenty-two genes were found to harbor positively selected sites in the encoded proteins: *orf1629*, *Pk1*, *Bm10* (*Ac18*), *arif-1*, *f*, *lef-12*, *Bm48* (*Ac60*), *Bm51* (*Ac63*), *gp37*, *dnapol*, *iap2*, *Bm59* (*Ac73*), *gp41*, *Bm67* (*Ac81*), *vp39*, *vp80*, *he65*, *pk2*, *ie-2*, *bro-a*, *bro-d* and *bro-c* (out of calculation). Most of these are not core genes. Taken together, these results indicate that variation in non-core genes greatly affect BmNPV evolution.

2.3. Comparative analysis of hr regions

In this study, cAMP response elements (CREs) and TPA response elements (TREs), in addition to 30 bp palindromes, were taken into consideration when searching for hr regions. Some mutations in the *Eco*R I site (GAATTC) of the palindromes or the CRE/TRE (-like) sites were allowed. Palindrome variants were identified as mutants of the core sequence, 'GTTTTACACGTAGAATTCTACTCGTAAAGC', or its reverse complement. The boxes represent sites where single nucleotide polymorsphisms (SNPs) often occur, strengthening or weakening the self-complementarity of the palindrome. As a result, another hr region (hr4c), which is located downstream of hr4L and hr4R, was also discovered in BmNPV strains. The "hr4L" and "hr4R" sites are also referred to as "hr4a" and "hr4b". Therefore, all six virus strains possess eight hr regions: hr2L, hr2R, hr3, hr4a, hr4b, hr4c, hr5, and hr1 (Table 1, Fig. 3).

Hrs exhibit large differences between BmNPV strains, not only with respect to the number of inner palindromes (Table 3) but also with respect to their structure, with differences of one to three CREs (-like) or TREs (-like) between each pair of hr palindromes, except hr4c (Fig. 3). Different hrs exhibit different degrees of conservation. Apparently, hr4c is the most conserved hr, followed by hr4b and hr2R, for which only one virus has a unique structure of the palindrome and CRE/TRE (-like) sites. Meanwhile, hr2L seems to be the most unstable hr, and the Cubic strain exhibits more complexity at this locus than do the Guangxi and Zhejiang strains. Moreover, the Indian strain, with a total of 46 palindromes in various hrs, stands out from the other viruses. Boma S1, with 42 palindromes, is the same as the T3 strain. The total number of palindromes in the Guangxi strain is similar to that in Zhejiang strain.

2.4. Analysis of viral DNA replication in Bm5 cells

To investigate DNA replication in different viruses, DNA from infected Bm5 cells was analyzed by qPCR. Bm5 cells infected with six virus strains (M.O.I. = 5) were harvested at 6, 12, and 24 h.p.i. Then, using the Universal Genomic DNA Extraction Kit Ver 3.0 (Takara), the total intracellular DNA, including the viral DNA, was extracted for real-time PCR analysis, and the DNA concentration of each sample was measured by Thermo scientific NanoDrop 2000 spectrophotometer. The PCR product of the partial *ie1* sequence was shown to be specific and unique by melting curve analysis, and the standard curve for qPCR had characteristics appropriate for quantification ($y = -3.474x + 37.553$, $y = CT$, $x = \text{Log (Copies)}$, $R^2 = 0.9997$). According to the standard curve, the average number of viral genome copies in each sample was quantified, and viral genome copies per 1 ng total DNA was calculated for comparison between viruses at specific times. As shown in Fig. 4, the Indian strain shows a remarkably high level of DNA replication at the early stage of infection, but the DNA replication rates of the Boma S1 and T3 strains were faster than those of the other viruses from 6 to 12 h.p.i. These faster rates were associated with greater amounts of these two strains at 12 h.p.i. At 24 h.p.i., the differences among the viruses were not significant, except that the amount of Boma S1 DNA was significantly greater than that of the Guangxi strain. This pattern does not seem to be related to the phylogeny. However, based on the statistical analysis, these six virus strains can be divided into two general groups. The Boma S1 strain, Indian strain and T3 strain have more palindromes in their hrs and form the high-replication group, while the Cubic strain, Zhejiang strain, and Guangxi strain, with fewer palindromes, make up the low-replication group. Moreover, the Guangxi strain, with only 34 total palindromes, clearly produces the lowest amount of DNA. Then the bivariate correlation between palindrome number and viral DNA replication efficiency was analyzed by Spearman's rank correlation coefficients with one-tailed test. The *p* values of the correlation between palindrome number and viral DNA replication at 6 h.p.i., 12 p.h.i., and 24 p.h.i. are 0.212, 0.087, and 0.042, respectively, the significance of this correlation increasing over time. These results suggested that the DNA replication rate of the virus may be related to the total number of palindromes in the hrs, consistent with previous findings that have implicated hrs as origins of DNA replication [18–20].

2.5. OB infectivity in B. mori larvae

Second instar silkworm larvae (Dazao strain) were used to test the OB pathogenicity of these six virus strains. By examining larval response under the pressure of 10^7 OBs/mL (10^5 OBs each larva on average), the OB pathogenicity of these six virus strains was compared. As shown in Fig. 5, the Indian strain exhibited the highest mortality rate, followed by T3, whereas Boma S1 had the lowest mortality rate. The pathogenicities of the Guangxi strain, Zhejiang strain, and Cubic strain were not significantly different. These results are to some extent consistent with the evolutionary relationships of these six virus strains. The strains with close relationships were associated with similar mortality rates.

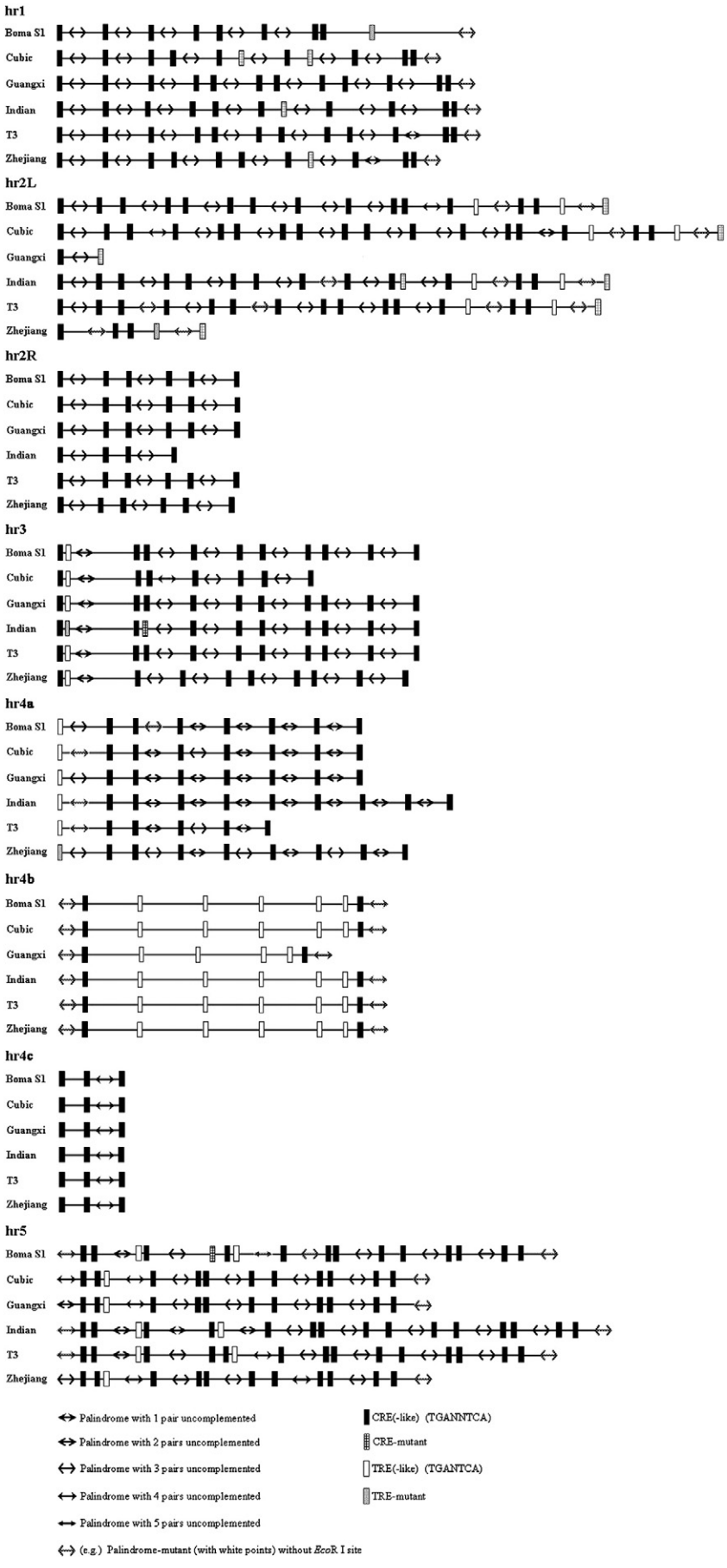

Fig. 3. Palindromes and CRE/TREs in the hrs of six virus strains.

Table 3
Distribution of palindromes in the *hrs* of six virus strains.

	BomaS1	Cubic	Guangxi	Indian	T3	Zhejiang
hr2L	9	11	1	9	9	2
hr2R	3	3	3	2	3	3
hr3	6	4	6	6	6	6
hr4a	6	6	6	8	4	7
hr4b	2	2	2	2	2	2
hr4c	1	1	1	1	1	1
hr5	9	7	7	10	9	7
hr1	6	7	8	8	8	7
Total	42	41	34	46	42	35

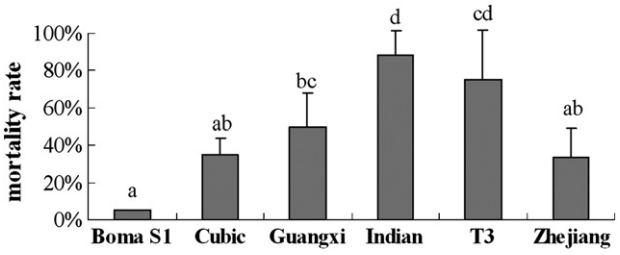

Fig. 5. Comparison of the OB infectivity of different BmNPV strains. Different lowercase letters following the figures in the columns indicate significant differences ($P < 0.05$) as analyzed using Duncan's test (one-way ANOVA).

3. Discussion

3.1. Diversity and evolution of BmNPVs

Many RNA viruses and some DNA viruses exist as multiple variants due to their high levels of adaptability and rapid evolution [21–24]. Baculoviruses, especially those that are pathogenic to lepidopteran larvae, also exhibit considerable genetic diversity [25–29]. Geographic isolates often exhibit large differences in genotype and pathogenicity, suggesting that there is a correlation between these differences and geographic location [25,30,31]. This hypothesis was also supported in our study. According to the phylogenic tree of the six strains of BmNPV, the Cubic strain and Boma S1 are closely related, possibly because they are from the same location and may have originated from a common ancestor, even though the Cubic strain produces unique cubic OBs, and Boma S1 was originally isolated from *B. mandarina*. This same is true for the Indian strain and T3 strain. Both the Indian strain and the T3 strain are from regions that experience tropical monsoons and in which these viruses' hosts, silkworms, are multivoltine. Surprisingly, the Zhejiang strain and the Guangxi strain, which are from different provinces of China, are closely related and appear to be the youngest strains according to the phylogenetic tree. This result may be attributed to the facts that the Guangxi and Zhejiang provinces both have a history of silk production spanning thousands of years and that these provinces are two of Chinese biggest silk-producing area, where it is popular to exchange good silkworm strains.

Because baculoviruses exhibit such high diversity, the evolution of baculoviruses was analyzed using phylogenetic trees based on certain conserved genes, such as *polh*/*gran*, *dnapol*, *lef-8*, *pif-2*, and *lef-9* [32–34]. Indeed, these single genes can sometimes be used to generate a phylogenic tree that represents the evolutionary relationships between large branches in the baculovirus tree [35]. However, single gene phylogenies often conflict with each other [35–37] because these genes do not meet several criteria [38]. As the complete genome sequences of viruses become available, whole-genome comparison becomes a powerful approach with which to infer the phylogeny of baculoviruses. To date, a common set of 31 genes, called core genes, has been identified based on the comparison of the known genome sequences of approximately 60 baculoviruses [2]. These core genes seem to be irreplaceable, and they participate in basic biological functions, such as RNA transcription, DNA replication and the formation of the virion structure. Based on the alignment of the amino acid sequences of the core genes, baculoviruses may be divided into four genera [2,39]. In our study, the core genes of six strains of BmNPV were found to be conserved, while the non-core genes appear to play important roles in BmNPV evolution, as determined via dN/dS analysis. This result suggests that the core genes are not suitable for use in the evolutionary analysis of various strains of a specific baculovirus.

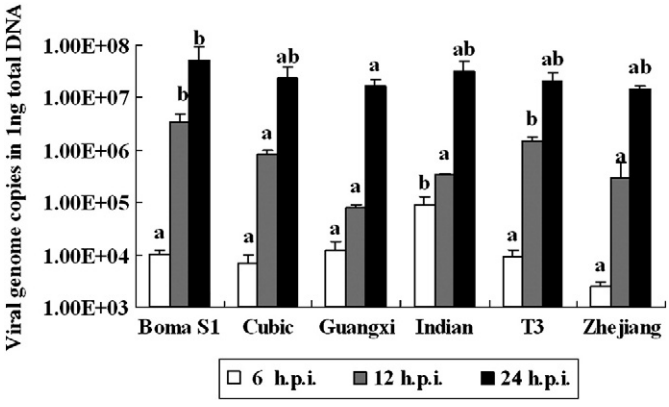

Fig. 4. DNA replication of different BmNPV strains in Bm5 cells. The results were analyzed using Duncan's test (one-way ANOVA). Different lowercase letters following the figures in the same color column indicate significant differences ($P < 0.05$).

3.2. Baculovirus homologous repeat regions (hrs)

The baculovirus hrs have been implicated both as origins of DNA replication and as transcriptional enhancers [18–20]. A single complete palindrome can initiate AcMNPV infection-dependent plasmid replication, and this ability can be affected by the number of palindromes in a specific hr [18–20]. Although each hr of AcMNPV can stimulate the replication of an hr-containing plasmid in Sf9 cells at different levels (Leisy and Rohrmann 1993), no single hr is absolutely essential for virus replication in cell culture as the deletion of either a single hr or two hrs fails to affect virus replication [40]. In this study, viruses containing more total palindromes in all hrs exhibited enhanced DNA replication capabilities and Spearman's rank correlation coefficients analysis showed there is a positive correlation between the palindrome number and viral DNA replication. This provided a powerful evidence for the impact of these regions on the DNA replication of this virus.

The Zhejiang and Guangxi strains have many fewer palindromes in hr2L than do the other viruses. This low number of palindromes may have an adverse effect on the expression of nearby genes, such as *fgf*, *Bm25* (Ac34), and *ubiquitin* (Ac35), which have been thought to be related to the infectivity of baculoviruses, thus negatively affecting the pathogenicity of the Zhejiang and Guangxi strains.

3.3. Variations in the OB infectivity of BmNPV strains

The result of baculovirus evolution is ultimately reflected in the infectivity of the virus through the oral route because baculoviruses coevolve with their hosts. Therefore, the OB infectivities of the six BmNPV variants were compared. The results showed that closely related strains are associated with similar mortality rates. However, PIFs (per os infectivity factors) (e.g., P74-PIF, PIF-1, PIF-2, PIF-3, PIF-4, and ODV-E56) are also important because they have been shown to be essential for oral infection of baculoviruses in insects [41–46].

Amino acid polymorphisms within the PIFs of different BmNPV variants are listed in Table S1. First, no particular amino acid change was found in the PIF-4 of any virus strain. Second, no amino acid

change was found in the Guangxi strain. Third, the Indian strain has the most amino acid changes in the PIF genes, indicative of its great divergence from the other viruses. Fourth, a greater number of particular amino acid changes were found in the non-transmembrane domains than in the transmembrane domains of the PIFs, and the former amino acid changes should be given more attention because they are located outside of the membrane and may participate in interactions with proteins. Amino acid mutations in these proteins may affect their modification or structures and consequently affect their functions. This may to some extent explain why the Indian strain showed the greatest infectivity via the oral route, whereas Boma S1 had the weakest infectivity. In short, the variations in the OB infectivity of these six BmNPV strains may be related to amino acid polymorphisms in these PIFs. Further studies are necessary to identify which amino acid changes are responsible for these variations.

4. Materials and methods

4.1. Cell line, insects, and viruses

Bm5 cells were cultured at 27 °C in TC-100 medium supplemented with 10% fetal bovine serum. *B. mori* larvae (Dazao strain) were reared on mulberry leaves at 25 °C. In addition to BmNPV T3 (from Thailand) and BomaNPV S1 (Boma S1, from Jiangsu province, China), BmNPV Cubic (from Jiangsu province, China), BmNPV Guangxi (from Guangxi province, China), BmNPV Zhejiang (Zhejiang province, China) and BmNPV Indian (from India) were originally isolated from diseased *B. mori* larvae and plaque-purified in Bm cells. These viruses were propagated by infecting Bm cells or 5th instar silkworm larvae.

4.2. Genome sequence determination

Viral DNA was extracted with phenol-chloroform from the purified occlusion bodies of different strains from infected silkworm larvae, as described previously [12]. The viral genomic DNA was amplified using the REPLI-g® Mini kit (QIAGEN), sheared into 200- to 300-bp fragments by ultrasonication and sequenced using Solexa technology with the Solexa 1 G Genome Analyzer. Average sequencing depth of BmNPV Guangxi or Zhejiang strains was ~100×. Ambiguous regions were identified by PCR with specific primers (Table S2), and gaps were bridged. Finally, the whole genome was assembled.

4.3. DNA sequence analysis

The nucleotide compositions of the genomic DNA and the predicted ORFs were analyzed using ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html), Softberry GENE Finding in Viral Genomes (http://linux1.softberry.com/berry.phtml?topic=virus&group=programs&subgroup=gfindv) and Lasergene software. The alignments of predicted homologous ORFs were performed with ClustalX2, and evolutionary trees were generated using MEGA5 with the maximum likelihood method [47]. The dN/dS value of relevant homologous ORFs was calculated using PAML (Codeml), and positively selected sites were also analyzed using the Naive Empirical Bayes (NEB) and Bayes Empirical Bayes (BEB) methods [48,49].

4.4. qPCR analysis of viral DNA replication in Bm5 cells

Bm5 cells were cultured in 6-well plates, infected with viruses (M.O.I. = 5) and harvested at 6, 12, and 24 h.p.i. Each sample was performed in triplicate. Total intracellular DNA was extracted using the Universal Genomic DNA Extraction Kit Ver 3.0 (Takara) and treated with RNase A. The DNA (50 μL) was diluted ten-fold for quantitative real-time PCR analysis. Real-time PCR was carried out using the BIOER Fluorescent Quantitative Detection System (China) with SYBR® *premix Ex* Taq™ (Takara) under the following conditions: 95 °C for 30 s, 40 cycles of 95 °C for 5 s, and 56–60 °C for 34 s, followed by a heat-dissociation process (65 to 90 °C). Each 20 μL reaction contained 2 μL of template. For the absolute quantification of viral genomes, *ie1* was used as the marker of viral DNA. The PCR products of the partial *ie1* sequence were used as the standard DNA samples (Table S3), and a standard curve was generated using serial dilutions of the PCR products from 10^9 to 10 copies/μL.

4.5. Bioassay of OBs

Varying doses of OBs of different virus strains were diluted in PBS containing kanamycin and ampicillin (10^7 OBs/mL, 10^6 OBs/mL, 10^5 OBs/mL, 10^4 OBs/mL, 10^3 OBs/mL, and 0 OBs/mL) and were evenly applied on fresh mulberry leaves (4 cm × 4 cm, 200 μL each piece). Each treated leaf was fed to 20 2nd instar silkworm larvae (Dazao strain) that had been starved for 3 h after molting. Each concentration of each virus was tested in triplicate. After 4 days, the number of dead larvae was recorded.

4.6. Statistical analysis

The statistical analysis was performed using DPS Statistics software [50]. One-way ANOVA followed by Duncan's test ($P < 0.05$) was used to identify significant differences, and Spearman's rank correlation coefficients with one-tailed test was applied to analyze bivariate correlations.

Acknowledgments

This work was supported by grants from the National Science Foundation of China (grants 30271008 and 30570074).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ygeno.2013.04.015.

References

[1] B.A. Keddie, G.W. Aponte, L.E. Volkman, The pathway of infection of Autographa californica nuclear polyhedrosis virus in an insect host, Science 243 (1989) 1728.
[2] S.A. Miele, M.J. Garavaglia, M.N. Belaich, P.D. Ghiringhelli, Baculovirus: molecular insights on their diversity and conservation, Int. J. Evol. Biol. 2011 (2011) 379424.
[3] C.X. Zhang, X.C. Ma, Z.J. Guo, Comparison of the complete genome sequence between C1 and G4 isolates of the Helicoverpa armigera single nucleocapsid nucleopolyhedrovirus, Virology 333 (2005) 190–199.
[4] J.G. Ogembo, B.L. Caoili, M. Shikata, S. Chaeychomsri, M. Kobayashi, M. Ikeda, Comparative genomic sequence analysis of novel Helicoverpa armigera nucleopolyhedrovirus (NPV) isolated from Kenya and three other previously sequenced Helicoverpa spp. NPVs, Virus Genes 39 (2009) 261–272.
[5] L. Li, Q. Li, L.G. Willis, M. Erlandson, D.A. Theilmann, C. Donly, Complete comparative genomic analysis of two field isolates of Mamestra configurata nucleopolyhedrovirus-A, J. Gen. Virol. 86 (2005) 91–105.
[6] K. Kamiya, J. Zhu, M. Murata, B.A. Lavina-Caoili, M. Ikeda, M. Kobayashi, S. Kawamura, Cloning and comparative characterization of three distinct nucleopolyhedroviruses isolated from the common cutworm, Spodoptera litura (Lepidoptera: Noctuidae) in Japan, Biol. Control. 31 (2004) 38–48.
[7] J.G. Ogembo, S. Chaeychomsri, K. Kamiya, H. Ishikawa, Y. Katou, M. Ikeda, M. Kobayashi, Cloning and comparative characterization of nucleopolyhedroviruses isolated from African bollworm, Helicoverpa armigera,(Lepidoptera: Noctuidae) in different geographic regions, J. Insect Biotechnol. Sericol. 76 (2007) 1.
[8] H.K. Hong, S.D. Woo, J.Y. Choi, H.K. Lee, M.H. Kim, Y.H. Je, S.K. Kang, Characterization of four isolates of Bombyx mori nucleopolyhedrovirus, Arch. Virol. 145 (2000) 2351–2361.
[9] R.L. Cheng, Y.P. Xu, C.X. Zhang, Genome sequence of a Bombyx mori nucleopolyhedrovirus strain with cubic occlusion bodies, J. Virol. 86 (2012) 10245.
[10] H.W. Fan, X.C. Zhang, Y.P. Xu, X.W. Cheng, C.X. Zhang, Genome of a Bombyx mori nucleopolyhedrovirus strain isolated from India, J. Virol. 86 (2012) 11941.
[11] S. Gomi, K. Majima, S. Maeda, Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus, J. Gen. Virol. 80 (Pt 5) (1999) 1323–1337.
[12] Y.P. Xu, Z.P. Ye, C.Y. Niu, Y.Y. Bao, W.B. Wang, W.D. Shen, C.X. Zhang, Comparative analysis of the genomes of Bombyx mandarina and Bombyx mori nucleopolyhedroviruses, J. Microbiol. 48 (2010) 102–110.
[13] W. Kang, M. Suzuki, E. Zemskov, K. Okano, S. Maeda, Characterization of baculovirus repeated open reading frames (bro) in Bombyx mori nucleopolyhedrovirus, J. Virol. 73 (1999) 10339–10345.

[14] Y.P. Xu, L.Z. Gu, Y.H. Lou, R.L. Cheng, H.J. Xu, W.B. Wang, C.X. Zhang, A baculovirus isolated from wild silkworm encompasses the host ranges of Bombyx mori nucleopolyhedrosis virus and Autographa californica multiple nucleopolyhedrovirus in cultured cells, J. Gen. Virol. 93 (2012) 2480–2489.
[15] M.E. Wilson, T.H. Mainprize, P.D. Friesen, L.K. Miller, Location, transcription, and sequence of a baculovirus gene encoding a small arginine-rich polypeptide, J. Virol. 61 (1987) 661–666.
[16] R. Wang, F. Deng, D. Hou, Y. Zhao, L. Guo, H. Wang, Z. Hu, Proteomics of the Autographa californica nucleopolyhedrovirus budded virions, J. Virol. 84 (2010) 7233–7242.
[17] S.C. Braunagel, W.K. Russell, G. Rosas-Acosta, D.H. Russell, M.D. Summers, Determination of the protein composition of the occlusion-derived virus of Autographa californica nucleopolyhedrovirus, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 9797–9802.
[18] M. Pearson, R. Bjornson, G. Pearson, G. Rohrmann, The Autographa californica baculovirus genome: evidence for multiple replication origins, Science 257 (1992) 1382–1384.
[19] D.J. Leisy, G.F. Rohrmann, Characterization of the replication of plasmids containing hr sequences in baculovirus-infected Spodoptera frugiperda cells, Virology 196 (1993) 722–730.
[20] S.M. Rodems, P.D. Friesen, The hr5 transcriptional enhancer stimulates early expression from the Autographa californica nuclear polyhedrosis virus genome but is not required for virus replication, J. Virol. 67 (1993) 5776–5785.
[21] E. Domingo, E. Martinez-Salas, F. Sobrino, J.C. de la Torre, A. Portela, J. Ortin, C. Lopez-Galindez, P. Perez-Brena, N. Villanueva, R. Najera, et al., The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance — a review, Gene 40 (1985) 1–8.
[22] M.D. Bowen, P.E. Rollin, T.G. Ksiazek, H.L. Hustad, D.G. Bausch, A.H. Demby, M.D. Bajani, C.J. Peters, S.T. Nichol, Genetic diversity among Lassa virus strains, J. Virol. 74 (2000) 6992–7004.
[23] A. López-Bueno, M.G. Mateu, J.M. Almendral, High mutant frequency in populations of a DNA virus allows evasion from antibody therapy in an immunodeficient host, J. Virol. 77 (2003) 2701–2708.
[24] G.J. Smith, T.S. Naipospos, T.D. Nguyen, M.D. de Jong, D. Vijaykrishna, T.B. Usman, S.S. Hassan, T.V. Nguyen, T.V. Dao, N.A. Bui, Y.H. Leung, C.L. Cheung, J.M. Rayner, J.X. Zhang, L.J. Zhang, L.L. Poon, K.S. Li, V.C. Nguyen, T.T. Hien, J. Farrar, R.G. Webster, H. Chen, J.S. Peiris, Y. Guan, Evolution and adaptation of H5N1 influenza virus in avian and human hosts in Indonesia and Vietnam, Virology 350 (2006) 258–268.
[25] R.R. Gettig, W.J. McCarthy, Genotypic variation among wild isolates of Heliothis spp nuclear polyhedrosis viruses from different geographical regions, Virology 117 (1982) 245–252.
[26] R.I. Graham, W.I. Tyne, R.D. Possee, S.M. Sait, R.S. Hails, Genetically variable nucleopolyhedroviruses isolated from spatially separate populations of the winter moth Operophtera brumata (Lepidoptera: Geometridae) in Orkney, J. Invertebr. Pathol. 87 (2004) 29–38.
[27] G.W. Blissard, G.F. Rohrmann, Baculovirus diversity and molecular biology, Annu. Rev. Entomol. 35 (1990) 127–155.
[28] J.S. Cory, B.M. Green, R.K. Paul, F. Hunter-Fujita, Genotypic and phenotypic diversity of a baculovirus population within an individual insect host, J. Invertebr. Pathol. 89 (2005) 101–111.
[29] H.L. Williams, I.S. Otvos, Genotypic variation and presence of rare genotypes among Douglas-fir tussock moth multicapsid nucleopolyhedrovirus (OpMNPV) isolates in British Columbia, J. Invertebr. Pathol. 88 (2005) 190–200.
[30] M. Shapiro, J.L. Robertson, Natural variability of three geographic isolates of gypsy moth (Lepidoptera: Lymantriidae) nuclear polyhedrosis virus, J. Econ. Entomol. 84 (1991) 71–75.
[31] M.C. Del Rincón-Castro, J.E. Ibarra, Genotypic divergence of three single nuclear polyhedrosis virus (SNPV) strains from the cabbage looper, Trichoplusia ni, Biochem. Syst. Ecol. 25 (1997) 287–295.
[32] G.F. Rohrmann, M.N. Pearson, T.J. Bailey, R.R. Becker, G.S. Beaudreau, N-Terminal polyhedrin sequences and occluded Baculovirus evolution, J. Mol. Evol. 17 (1981) 329–333.
[33] P.M. Zanotto, B.D. Kessing, J.E. Maruniak, Phylogenetic interrelationships among baculoviruses: evolutionary rates and host associations, J. Invertebr. Pathol. 62 (1993) 147–164.
[34] E.A. Herniou, J.A. Olszewski, D.R. O'Reilly, J.S. Cory, Ancient coevolution of baculoviruses and their insect hosts, J. Virol. 78 (2004) 3244–3251.
[35] E.A. Herniou, T. Luque, X. Chen, J.M. Vlak, D. Winstanley, J.S. Cory, D.R. O'Reilly, Use of whole genome sequence data to infer baculovirus phylogeny, J. Virol. 75 (2001) 8117–8126.
[36] E.E. Clarke, M. Tristem, J.S. Cory, D.R. O'Reilly, Characterization of the ecdysteroid UDP-glucosyltransferase gene from Mamestra brassicae nucleopolyhedrovirus, J. Gen. Virol. 77 (Pt 11) (1996) 2865–2871.
[37] W. Kang, M. Tristem, S. Maeda, N.E. Crook, D.R. O'Reilly, Identification and characterization of the Cydia pomonella granulovirus cathepsin and chitinase genes, J. Gen. Virol. 79 (Pt 9) (1998) 2283–2292.
[38] D.L. Swofford, G.J. Olsen, P.J. Waddell, D.M. Hillis, Phylogenetic inference, in: D.M. Hillis, C. Moritz, B.K. Mable (Eds.), Molecular Systematics, Second edition, Sinauer Associates, Sunderland, Mass, 1996, pp. 407–514.
[39] E.A. Herniou, J.A. Jehle, Baculovirus phylogeny and evolution, Curr. Drug Targets 8 (2007) 1043–1050.
[40] E.B. Carstens, Y. Wu, No single homologous repeat region is essential for DNA replication of the baculovirus Autographa californica multiple nucleopolyhedrovirus, J. Gen. Virol. 88 (2007) 114–122.
[41] J. Kuzio, R. Jaques, P. Faulkner, Identification of p74, a gene essential for virulence of baculovirus occlusion bodies, Virology 173 (1989) 759–763.
[42] I. Kikhno, S. Gutierrez, L. Croizier, G. Croizier, M.L. Ferber, Characterization of pif, a gene required for the per os infectivity of Spodoptera littoralis nucleopolyhedrovirus, J. Gen. Virol. 83 (2002) 3013–3022.
[43] G.P. Pijlman, A.J. Pruijssers, J.M. Vlak, Identification of pif-2, a third conserved baculovirus gene required for per os infection of insects, J. Gen. Virol. 84 (2003) 2041–2049.
[44] M. Fang, Y. Nie, Q. Wang, F. Deng, R. Wang, H. Wang, J.M. Vlak, X. Chen, Z. Hu, Open reading frame 132 of Helicoverpa armigera nucleopolyhedrovirus encodes a functional per os infectivity factor (PIF-2), J. Gen. Virol. 87 (2006) 2563–2569.
[45] X. Li, J. Song, T. Jiang, C. Liang, X. Chen, The N-terminal hydrophobic sequence of Autographa californica nucleopolyhedrovirus PIF-3 is essential for oral infection, Arch. Virol. 152 (2007) 1851–1858.
[46] M. Fang, Y. Nie, S. Harris, M.A. Erlandson, D.A. Theilmann, Autographa californica multiple nucleopolyhedrovirus core gene ac96 encodes a per os infectivity factor (PIF-4), J. Virol. 83 (2009) 12569–12578.
[47] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, Mol. Biol. Evol. 28 (2011) 2731–2739.
[48] Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood, Mol. Biol. Evol. 24 (2007) 1586–1591.
[49] Z. Yang, W.S. Wong, R. Nielsen, Bayes empirical bayes inference of amino acid sites under positive selection, Mol. Biol. Evol. 22 (2005) 1107–1118.
[50] Q.Y. Tang, C.X. Zhang, Data processing system (DPS) software with experimental design, statistical analysis and data mining developed for use in entomological research, Insect Sci. 20 (2013) 254–260.