

Transcriptional Analysis of the Ribonucleotide Reductase Genes of Shrimp White Spot Syndrome Virus

Meng-Feng Tsai,* Chu-Fang Lo,*, Mariëlle C. W. van Hulten,† Huey-Fen Tzeng,* Chih-Ming Chou,‡ Chang-Jen Huang,‡ Chung-Hsiung Wang,§ Jung-Yaw Lin, Just M. Vlak,† and Guang-Hsiung Kou*, Jung-Yaw Lin, Just M. Vlak,† and Guang-Hsiung Kou*,

*Department of Zoology, §Department of Entomology, and ¹Institute of Biochemistry, National Taiwan University, Taipei, and ‡Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan, Republic of China; and †Laboratory of Virology, Wageningen University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands

Received May 30, 2000; returned to author for revision June 27, 2000; accepted August 17, 2000

The causative agent of white spot syndrome (WSS) is a large double-stranded DNA virus, WSSV, which is probably a representative of a new genus, provisionally called Whispovirus. From previously constructed WSSV genomic libraries of a Taiwan WSSV isolate, clones with open reading frames (ORFs) that encode proteins with significant homology to the class I ribonucleotide reductase large (RR1) and small (RR2) subunits were identified. WSSV rr1 and rr2 potentially encode 848 and 413 amino acids, respectively. RNA was isolated from WSSV-infected shrimp at different times after infection and Northern blot analysis with rr1- and rr2-specific riboprobes found major transcripts of 2.8 and 1.4 kb, respectively. 5' RACE showed that the major rr1 transcript started at a position of -84 (C) relative to the ATG translational start, while transcription of the rr2 gene started at nucleotide residue -68 (T). A consensus motif containing the transcriptional start sites for rr1 and rr2 was observed (TCAc/tTC). Northern blotting and RT-PCR showed that the transcription of rr1 and rr2 started 4-6 h after infection and continued for at least 60 h. The rr1 and rr2 genes thus appear to be WSSV "early genes." © 2000 Academic Press Key Words: Penaeus monodon; white spot syndrome virus; Taiwan WSSV isolate; ribonucleotide reductase gene; transcription analysis.

INTRODUCTION

White spot syndrome (WSS) is an economically significant shrimp disease which causes high shrimp mortalities and severe damage to shrimp cultures. The disease is caused by a virus called white spot syndrome virus (WSSV) (Takahashi et al., 1994; Chou et al., 1995). WSSV is an enveloped ovoid-shaped virus with a rod-shaped nucleocapsid with flat ends (Wang et al., 1995; Wongteerasupaya et al., 1995). On the basis of restriction enzyme analysis, the genome of this virus was determined to be a double-stranded DNA molecule (Wang et al., 1995; Wongteerasupaya et al., 1995) with an estimated length of 300 kb (Yang et al., 1997; Lo and Kou, unpublished results). Comparison of different geographical isolates has indicated that limited differences exist among them (Lo et al., 1999).

WSSV has been collected and purified from infected Penaeus monodon (Wang et al., 1995) and was used to construct WSSV genomic libraries (Lo et al., 1996a). Early DNA sequence data suggested that WSSV is probably a unique virus (Lo et al., 1997). This supposition is now

¹ To whom correspondence and reprint requests should be addressed at National Taiwan University, Department of Zoology, Taipei 106, Taiwan, ROC. Fax: 886-2-23638179. E-mail: ghkou@ccms. ntu.edu.tw or gracelow@ccms.ntu.edu.tw.

further supported by sequence analysis, including that of WSSV ribonucleotide reductase (van Hulten et al., 2000a). WSSV was previously classified by Wang et al. (1995) as a member of the genus Non-occluded baculovirus (NOB) (Francki et al., 1991), but at present this genus is no longer accepted into the Baculovirus family (Murphy et al., 1995) as a result of a lack of molecular information. Genome analysis and data on WSSV replication and transcription are therefore needed to conclusively establish WSSV's taxonomic position (Lo et al., 1996b). In addition to sequence analysis, another important aspect of the molecular analysis is the regulation of transcription.

Recently, we have identified genes (rr1 and rr2) for both the large (RR1) and the small (RR2) subunit of ribonucleotide reductase in WSSV (van Hulten et al., 2000a). This enzyme is involved in nucleotide metabolism and reduces ribonucleotides into deoxyribonucleotides as immediate precursors of DNA (Jordan and Reichard, 1998). The open reading frame (ORF) of WSSV rr1 is 2547 nucleotides (nt) long, and that of rr2 is 1242 nt. Phylogenetically, the WSSV rr1 and rr2 genes are distinct not only from the baculovirus rr genes but also from rr genes of other organisms and viruses (van Hulten et al., 2000a).

In this investigation we report the first study on the transcription of WSSV. Northern blot analysis was used



to determine the size of the WSSV rr1 and rr2 transcripts, and the 5' and 3' ends of the messenger RNA were analyzed by 5' and 3' RACE. Northern blot analysis and RT-PCR were then used to investigate the temporal expression of the rr1 and rr2 genes in diseased shrimp. The results show that the transcriptional regulation of these genes involves unique promoters that are distinct from those of known baculoviruses. This observation further supports the view that WSSV may be a representative of a new genus, tentatively named Whispovirus.

RESULTS

Sequence analysis of WSSV rr1 and rr2

The genomic fragments used in this study came from plasmid libraries constructed from WSSV that was isolated from *Penaeus monodon* collected in Taiwan in 1994 (Wang *et al.*, 1995). The four libraries (pmh, pms, pme, and pmbh, where "pm" indicates *Penaeus monodon*) are named for their *HindIII*, *Sal*I, *EcoRI*, and *HindIII*/BamHI enzymes. Genomic fragments pmh21 (1.6 kb), pms54b (2.7 kb), and pms420 (2.0 kb) were found to contain the large subunit of ribonucleotide reductase gene *rr1*; the cloned 2.9-kb WSSV genomic *EcoRI* fragment (pme257) and the overlapping 4.9-kb *HindIII*/BamHI fragment (pmbh32) encompassed the gene (*rr2*) for the small subunit of ribonucleotide reductase. The *rr1* and *rr2* genes are arranged in a head-to-head configuration but separated by 5.7 kb of sequence.

The *rr1* and *rr2* sequences from the Taiwan isolate were compared to the same genes from a strain from Thailand (van Hulten *et al.*, 2000a). The *rr1* sequences were found to be 100% identical. In *rr2*, a C in the Taiwan strain was replaced by a T in the Thailand isolate at position 164 downstream of A (+1) of the *rr2* start codon, and there were other differences at position –48 (C to G), –85 (C to G), –95 (A to G), and –120 (deletion of a C nucleotide). The downstream difference resulted in an F55S change. Thus, the RR1 and RR2 amino acid sequences of Taiwan WSSV isolate, respectively, have 100 and 99.5% identity with those of the Thailand WSSV isolate. This confirms a previous observation made on the basis of limited sequencing, that there is little genetic variation in WSSV isolates (Lo *et al.*, 1999).

Number of copies of rr1 and rr2 ORFs on the WSSV genome

Using Southern blot analysis, the *rr1*-specific probe hybridized with a 2.7-kb *Sal*l and a 1.6-kb *Hin*dIII WSSV genomic fragment (Fig. 1a), whereas the *rr2*-specific probe hybridized with a 15-kb *Sal*l, an 8-kb *Hin*dIII, and a 2.8-kb *Eco*RI WSSV genomic DNA fragment (Fig. 1b). These data suggest that the WSSV genome contains only a single copy of each of the *rr* genes.

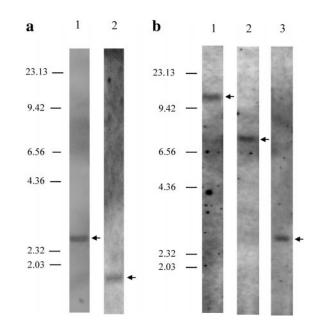


FIG. 1. (a) Hybridization of a DIG-labeled WSSV rr1-specific probe to Southern blots of WSSV DNA digested with SalI (lane 1) and HindIII (lane 2) restriction endonucleases. (b) Hybridization with a DIG-labeled WSSV rr2-specific probe. Southern blots of WSSV DNA digested with SaII (lane 1), HindIII (lane 2), and EcoRI (lane 3) restriction endonucleases. The size standards are indicated using λ HindIII DNA marker (Promega).

Determination of the termini of the *rr1* and *rr2* transcripts

The sequences of the genomic clones that harbor the rr1 and rr2 genes were analyzed by NNPP (see Materials and Methods), which predicted a potential promoter element for the rr1 gene between -122 and -73 nt, where +1 is the putative translation start codon (Fig. 2). Experimental results from 5' RACE confirmed this computer prediction. The locations of the clones obtained from 5' RACE and 3' RACE are shown in Fig. 3. The 5' RACE analysis (based on three clones, 5'rr1-B2, 5'rr1-B4, and 5'rr1-B5) revealed that transcription started predominantly at a C residue 84 nt upstream of the translation initiation codon of the rr1 ORF and 28 nt downstream of a TATA box (TATAAA; Fig. 2). At the 3' end of the rr1 ORF in WSSV DNA there is a polyadenylation signal AATAAA 3 to 8 nt downstream of the translational stop codon (Fig. 2). A 3' RACE analysis of the rr1 mRNA revealed an addition of poly (A) at a site 12 nt downstream of the polyadenylation signal.

For the rr2 gene, the NNPP program identified two high-probability promoter regions spanning from -166 to -117 nt for element I, and from -109 to -60 nt for element II, where +1 is the putative translation start codon. A predicted transcriptional initiation site at -69 nt (T) was confirmed with 5' RACE (clones 5'rr2-B6 and 5'rr2-F17), which suggests that element II is the actual promoter (Fig. 4). Another potential transcriptional initiation site (revealed by clone 5'rr2-F15) was located at -63

94 TSAI ET AL.

TCGACCCAGCGGTCCACCCTCGGAACTTGACATCAGGCCTAGACAGCGGTCCACCCCTAAACTCGAGTGATCCAGAAAAATTTTTAAAAAGTTTTTTGAGGAACAGTTTTGCACCTGCAC -241 ATGGGTTCTAACCAGCAACAATCATTCATCTCAAAGAGGAATGGCACTAAGCAAGAAATTAGTCTTGAAAAGATAATCAAGAGGATTGAAAATGCCTGTCTTCCAGTCAACCAGTATGTG 120 M G S N O O O S F I S K R N G T K O E I S L E K I I K R I E N A C L P V N O Y V CCCAAGCTTGACAAGAACGCAATTAACCCTCAAGAACTTGCATCTCACATCATGGACCGTCTTCCCGCTACCATCCTCCCTTCCAAGAAATGGACGATTTTCTGGCCGATTATGCAAAGACA 240 P K L D K N A I N P O E L A S H I M D R L P A T I S F O E M D D F L A D Y A K T AAAATTGTTGACCACCCTGATTTTGGAAAACTGGCAGGAAGATTCATCTGTTCGAACATCCACAAAAACACCAAAGAGTGGAATAGTTTTAGTGCAACAACACCAGAAATTGAGGCACGCA K I V D H P D F G K L A G R F I C S N I H K N T K E W N S F S A T T O K L R H A ATCCACCCTGGAACTGGTAAACCAGCATCAGTAGTTAATGATACCTACTATGAAAATGTTATGGCCAATGCTGAAATTCTCGATGCTGTTATCGATTACCAAAATGGATTATCTCTTCACC 480 I H P G T G K P A S V V N D T Y Y E N V M A N A E I L D A V I D Y K M D Y L F 600 C F G L R T L E Y S Y L I K I G S P T D R K K R I L V E R P Q D M I M R V A V G ATTCACGGATCAGACATCAGATCTGTCATTGAAACATATGATCTCATGTCGAGGCACTATTTTACCCCATGCTTCCCCCCACACTGTTTAATTGTGGAACAGTCACACCCCAACTTTCCTCG 720 I H G S D I K S V I E T Y D L M S R H Y F T H A S P T L F N C G T V T P O L S S TGCTTCCTTCTGGGCCTTCAAGATGATAGCATTGAGGGTATTTATGATACTCTTAAGGAGGCGGCAATCATCTCCAAGACTGCTGGAGGACTCGGCATCCACTTTCATGATTTGAGGGCA 840 C F L L G L Q D D S I E G I Y D T L K E A A I I S K T A G G L G I H F H D L R A 960 K G S P I S S W S G T H P G L M A F L Q I F N V S V K K V S Q G G D K R R G A A I Y I S D W H L D V K D F I D C R K N A G N E D L R T R D L F P A I W V S D L TTCATGGAGAGAGTGAAGGCTGGGAAGAATTGGTCCCTGATGTGCCCCCACGAGTGCCCTTGCCCTTTCCGACGTCCATGGCGAAGAGTTTAAGGCCGTTGTACGAGAAATATGAGGCTGAA 1200 F M E R V K A G K N W S L M C P H E C P G L S D V H G E E F K A L Y E K Y E A E 1320 G K G K E V V K A R A L F D Q I N S A R I E T G T P Y V C F K D T I N R K S N Q GAAAATGTCGGCATCATCAAGTCTCAAATTTGTGCACTGAAATTGTCCAGTACAGTGATTCGGAGGAAACTGCAGTGTGCAATTTGGCATCTATCGCAGTCAACAAGTTTTGTGAAGTAT ENVGIIKS SNLCTEIVOYS DSEETAV CNLASIAVNK FVKY 1560 S P I P S L R P Y V D Y R E M K R V V K I M T R N L D K V I D V N F Y A V D K T CGCATTTCCAATATGAAAACTAGGCCAATGGGATTGGGTGTGCAGGGACTAGCAGATTTGTTCTTCAAACTCAGAATCCCCTTCGAATCTGAAGAAGCGGCACTAATTAACAAGAGGGATT 1680 R I S N M K T R P M G L G V Q G L A D L F F K L R I P F E S E E A A L I N K R I TTIGAAACTATATACTATGGTGCCTTGGAAGCTTCATGTGAAATTGCCAAAGAAAAGGGAGAAACATATGACCTGTTTGAGGGTAGTCCTTTGACAAGGGATTTTCCAATTTGACATG 1800 F F T I V V G A I F A S C F I A K F K G F T V F I F F G S P I S K G I F O F D M GGGAAGGAAAATATTAAGAATAGGGACATATATTTCAACTCTTTGCCAATTCACGATTGGGAGCAATTGAGAAGGGACATTATGAAGTATGGTGTTCACAATTCAATGTTTTGTTGCTCCC 1920 G K E N I K N R D I Y F N S L P I H D W E Q L R R D I M K Y G V H N S M F V A P ATGCCTACTGCATCCACTGCACAGATCCTGGGCCACTCTGAATCCTTTGAGCCTTTAACGTCCAACATGTATAATCGTAATCTTTCAGGATCATTCCAAGTGGTGAACGAATATCTG 2040 M P T A S T A Q I L G N S E S F E P L T S N M Y N R N V L S G S F Q V V N E Y V ATTAGAGAGCTCATAAAACTGGGAGAATGGAATTCAGTAACTAAACAGAGGGATTATGGCAAGTGGTGGTCGATCTATTCAGACGCTTCCTAATATCCCTAAATCGACCAAGGAACTATTCAAA 2160 I R F I I K I G F W N S V T K O R I M A S G G S I O T I P N I P K S T K F I F K 2280 T V W E I N P R T T L D M A I Q R G M F V D Q A Q S L N L F V E E P E L S K V R TCAATGACTATGTATOCATOGGAGAAGGOGATCAAGACTCTCTATTATCTACGCACAAAGGGCGCAGCTAGAGCTGTCCAGTTCACGGTCGACAAGAATGTACTCCAAGAAGTCAAGAAG 2400 S M T M Y A W E K G I K T L Y Y L R T K G A A R A V Q F T V D K N V L Q E V K K 2520 E A P S P V A A F S A P V R E E E E E K K S S I V V P D P A A A L L C S I N N P 2640 GACEMCSS* Poly A addition site

FIG. 2. Nucleotide sequence of WSSV rr1 containing the 5' and 3' terminal region. The deduced amino acid sequence is indicated below the nucleotide sequence. The NNPP program predicted a potential promoter element between -122 and -73 nt. The transcriptional start point is indicated by a bent arrow; the predicted promoter elements are shaded; poly (A) signal and TATA are underlined and boxed, respectively.

ACTITATOGCATTCATTCCCTCGTTCAACATGGAAGGTGTAACATCCATAGTTGCAGCTGCTGTCCCAGAAGTTGCAATCTTGATCACCGACCTAATGG

nt; however, this alternative 5' terminus of transcript may only have resulted from early termination during cDNA synthesis in 5' RACE. The TATA box is located at -99 to -94 nt of the initiation ATG. A potential polyadenylation signal AATAAA was at 60 to 65 nt downstream of the translational stop codon (Fig. 4). Sequence analysis of 3' RACE products revealed that poly (A) was added at a site 13 nt downstream of this polyadenylation signal.

Analysis of the upstream sequence of the rr genes

The predicted major transcriptional start points with their surrounding sequences are CTTCACTCA for *rr1* (Fig. 2) and CAGCATCAT and CCCCCCTCC for *rr2* (Fig. 4). The predicted transcriptional initiation sites (at positions -85, -69, and -121, respectively) are underlined. The transcriptional initiation points identified by 5' RACE were in the middle of the predicted sequence CTTCACTCA for *rr1* (Fig. 2) and CAGCATCAT for *rr2* (Fig. 4);

the second predicted sequence for *rr2*, CCCCCTCC, did not contain a transcriptional initiation point in the 5' RACE analysis.

2732

The putative TATA box of rr1 is located at -108 to -113 nt upstream of the initiation ATG. There is a general putative binding sequence for TATA-binding proteins located between -114 to -100.

Temporal expression of *rr1* and *rr2* transcription in WSSV-infected shrimp

Using rr1- and rr2-specific riboprobes generated by in vitro transcription, the Northern blot hybridization analysis revealed a major transcript of rr1 of approximately 2.8 kb (Fig. 5a) and a major transcript of rr2 of about 1.4 kb (Fig. 5b). Neither of these transcripts was detected in uninfected (0 h postinfection [p.i.]) shrimp. The rr1 transcripts were first detected at 4 h p.i. and the rr2 transcript first appeared at 6 h p.i. A larger transcript (\sim 4.5 kb) for

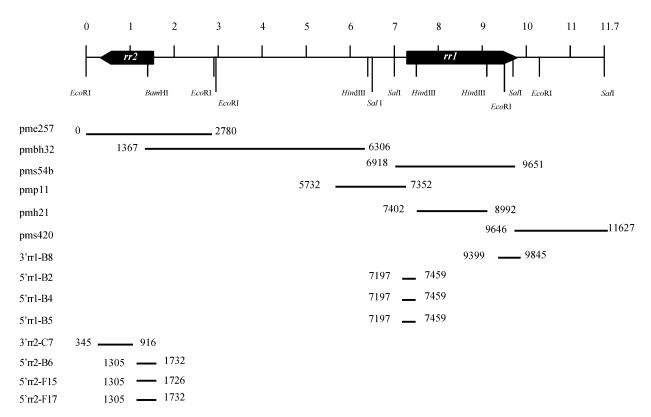


FIG. 3. Schematic alignment of cloned WSSV genomic fragments (pme257, pmbh32, pms54b, pmp11, pmh21, pms420) and cDNA fragments generated by 3' RACE (3'rr1-B8; 3'rr2-C75) and 5' RACE (5'rr1-B2, -B4, -B5; 5'rr2-B6, -F15, -F17). Location of *rr1* and *rr2* on the 11.7-kb WSSV genomic fragment is shown at the top.

rr2 was detected at 18 and 60 h p.i. (Fig. 5b), but the nature of this late transcript was not further investigated. In confirmation of the Northern blot analysis, RT-PCR

(Figs. 5 and 6) first detected both the *rr1* and *rr2* transcripts at 4 h p.i. and they continued to be present up to 60 h p.i. These results indicated that both the *rr1* and *rr2*

CTTGCAAGACCAGAAGGGTGTATTTGAGACCGTCTGCAAAATAAAGAGCAGAGATGGTTTGTGAATTGGCCACACTGATTCCGATACCGACTAAAAAAGGCTGATTGAT	121
CTCCAGACTICAGCCGCATGAGTATAAATATIGGCCCACTTCTCACACCACAGCATCATTCCCTCGTCATCACTGCCGTCAACTTCCACTACTACTACCAACAAAATACCAACAACACCCAGAA	-1
ATGGAGTICAATCAAACTGTTCACCGTTGCTCGGATATGGAGCAAGCCAAGCCAAGTCGGTTGAGAAATCAAGTTCAGAATATAAAACCGAGGAAAAGGAGAATGCCCAGGAAGTTG	120
M E S I K L F T V A G L N M E O A N O V A E E I K S E Y K T E E E K R I A O E V	120
TTTGACAAATTCACCAAAAACTCATTATGCAAGTAGATACGTCTAAACACTTACTT	240
F D K F T K K L I M Q V D T S K H L L T R E N P N R F V S R P I V H E D L W E M	
TACAAAAAAGGGTTGCCTGTTTTTGGACATTGGAAGAGATTGATT	360
Y K K E V A C F W T L E E I D F E R D P K D W E K L T O D E K D F I L O I L A F	
TTTGCATCCTCTGACGGAATTGTAATTGAAAATCTTACAACACGTCTTCGTCAAGTGGCGCAGATTCCAGAAGCGAGGAGTTTCTTTTGACTTCCAAGTTTGGAATGGAGAGTATTCATGGC	480
FASSDGIVIENLTTRLROVAOIPEARSFFDFOVGMESIHG	
AACGTCTACGGAGAACTGATTGATAGACTGGTGCCCGACGAAAAAGACAAGGCTATCTTGTTTAACGCTGCACAACACTTCCCCGCCATCAAGAAGAAGAAGGAGCAGTGGGCTATTAATTGG	600
N V Y G E L I D R L V P D E K D K A I L F N A A O H F P A I K K K E O W A I N W	
ATGCAAAGCAATAACGATTTGGCGGAACTAATTGTTGCCTTTGCTGCAGTTGAAGGAATCTTCTTTTAGTGGTGCATTCCCATCCAT	720
M O S N N D L A E L I V A F A A V E G I F F S G A F A S I F W I K N R G I L P G	
CTCACCTCCTCCAATGAGTTCATTTCTAGGGACGAAGGTCTTCATCGCGACTTTGCATGCA	840
L T S S N E F I S R D E G L H R D F A C M L L K K G F V D T P S R E R I L E I V	
ACTIGAAGCCGTCCGAATTIGAACAAGAATTTCTCACAGTTTTCCCTGCTCGAAATTAGTCGGGAATGAACTGCAAGTTGATGAGCCAGTACATTTGAATTTTTGGCCAGATTAGAACTATTTGGTT	960
T E A V R I E O E F L T V S L P V K L V G M N C K L M S O Y I E F V A D K L L V	
	080
EMGLEKHYNVTNPFPFMDNISLENKTNFFEKRVAEYORAO	
GTCATGGCTTCTATCAATAGATCAGAAGAGGCCAACAAACCCAAGAACTGGTTCTCCTCCCCAATTCTGACTGCACCTCCTCCCAGTCTCTCCTCATCATCATCATCATCATCATCATCATCATCA	200
V M A S I N K I K K D O O T O E T G S P L P I L T A P P P V S S S S S E O E D V	
	320
E D G V G D Y I S Y D D F * Poly A addition site	
Toy in the same	1440
	560
GTCAGGAAAAGGTCCTTGT 1	1579

FIG. 4. Nucleotide sequence of WSSV rr2 containing the 5' and 3' terminal region. The deduced amino acid sequence is indicated below the nucleotide sequence. The NNPP program identified two predicted promoter elements spanning from -166 to -117 nt for element I, and from -109 to -60 nt for element II. The transcriptional start point is indicated by a bent arrow; the predicted promoter elements are shaded; poly (A) signal and TATA are underlined and boxed, respectively.

96 TSAI ET AL.

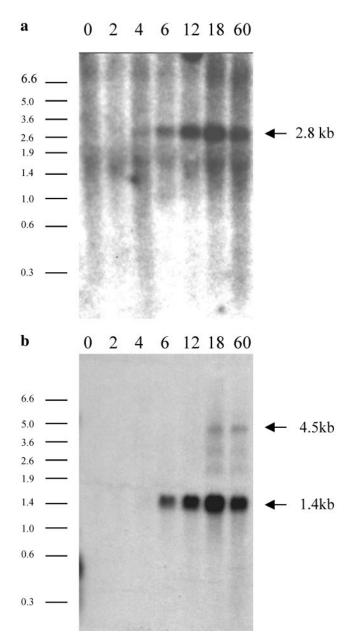


FIG. 5. Northern blot analysis of total RNA isolated from the pereiopods of WSSV-infected *P. monodon* using *rr1*-specific (a) and *rr2*-specific (b) riboprobes. RNA was isolated at 0, 2, 4, 6, 12, 18, and 60 h p.i. as indicated by the lane headings. The size standards are indicated using RNA markers (Promega).

genes were expressed early in the course of a WSSV infection of shrimp and suggested that transcription of both genes continued for at least 60 h.

DISCUSSION

In the present study, a comparison of the sequence of the ribonucleotide reductase of the Taiwan WSSV isolate with the sequence of the Thailand WSSV isolate reported by van Hulten *et al.* (2000a) has shown that the identity of WSSV *rr1* and *rr2* in these two different geographical

isolates approximates 100%. Sequencing of two of the major virion protein genes (VP26 and VP28; GenBank accession numbers AF272980 and AF272979, respectively) of the Taiwan strain confirmed this conclusion, as they were 100% identical to the corresponding genes in the Thailand WSSV isolate (van Hulten et al., 2000b). Some earlier studies, which used PCR and restriction fragment length polymorphism (RFLP) analysis of specific genomic DNA fragments instead of a nucleotide sequence comparison, have also shown that there is little genetic variation among many WSSV isolates from different crustacean hosts collected from Taiwan (Lo et al., 1996b) and cultured shrimp from around the world (Lo et al., 1999). To date, however, only a very small fraction (maybe \sim 3%) of the ORFs in the WSSV genome have been reported; moreover a study by Nadala et al. (1998) reported that when certain restriction enzymes are used, different geographical isolates of WSSV (from China, Indonesia, the United States, and Japan) could be distinguished by RFLP analyses. Clearly it is too soon to conclude that all WSSV geographical isolates are genetically similar, and further genetic studies will be needed to establish genetic markers that are able to identify WSSV geographical isolates for investigations into WSSV epidemiology.

Since some DNA viruses carry two rr2 genes (e.g., see Kuzio et al., 1999), Southern hybridization was used here to check the number of copies in the entire genome, which to date has not yet been completely sequenced. As Fig. 1 shows, the rr1- and rr2-specific probes indicated that only a single copy of each of these genes is present in the WSSV genome.

The location and arrangement of *rr1* and *rr2* in viral genomes differ from virus to virus (Boursnell *et al.*, 1991; Kuzio *et al.*, 1999; IJkel *et al.*, 1999; Ahrens *et al.*, 1997). In WSSV, *rr1* and *rr2* are 5.7 kb apart and are arranged in a head-to-head opposite orientation. Despite the distance

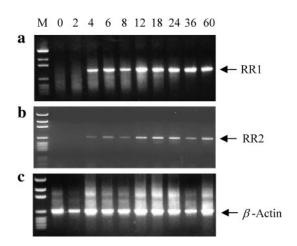


FIG. 6. RT-PCR with (a) rr1-specific primers rr1-F1/rr1-R1; (b) rr2-specific primers rr2-F1/rr2-R1; (c) internal control, Actin-F1/Actin-R1. M represents a pGEM DNA marker (Promega). Lane headings show hours p.i.

between them, the initiation of *rr1* and *rr2* gene expression in WSSV appears to occur in a coordinated manner (cf. Figs. 5 and 6), but it is not clear how this is achieved.

At present, little is known about promoters in WSSV. The TATA boxes of the *rr1* and *rr2* genes are located at -103 to -108 nt and -99 to -94 nt, respectively. More evidence for the functional significance of this box is provided by the fact that the region immediately surrounding this sequence also contains a predicted binding site for a putative TATA-binding protein (TBP). Thus, WSSV *rr1* and *rr2* may be transcribed by the host RNA polymerase II, a transcription mode, which is also used by the early genes of baculoviruses (Kuzio *et al.*, 1999).

In this study, 5' RACE revealed that the major transcriptional start sites of WSSV rr1 and rr2 were CT-TCACTCA and CAGCATCAT. The sequences of these major transcriptional start sites showed no homology to the early (CAGT) or late (TAAG) consensus transcriptional start site of the baculoviruses (Blissard and Rohrmann, 1990). The WSSV rr1 and rr2 transcriptional start sites also lacked homology to two other previously characterized transcriptional initiation sites for the rr1 gene of a baculovirus, Spodoptera exigua multicapsid nucleopolyhedroviruses (SeMNPV) (van Strien et al., 1997). Nor was any identity observed in the transcriptional initiation site between WSSV rr1 (CTTCACTCA) and rr2 (CAGCATCAT) per se. However, in the vicinity of the start point of rr1 (CTTCACTCA) and rr2 (CAGCATCATTC), there is a consensus sequence (shown in bold; TCAc/ tTC) which has only one nucleotide difference at the fourth position in this stretch of six nucleotides. The significance of this (if any) is not clear at present. Nevertheless, the consensus sequence (TCAc/tTC) may have a function in the coordinated initiation of gene expression in WSSV-infected cells. Future DNase I footprinting experiments could demonstrate the importance of these regions in WSSV rr transcriptions.

MATERIALS AND METHODS

Virus and plasmid clones

The virus used in this study was isolated from a batch of WSSV-infected *Penaeus monodon* collected in Taiwan in 1994 (Wang *et al.*, 1995), which is now known as WSSV Taiwan isolate (Lo *et al.*, 1999). From this virus, plasmid libraries (pmh and pms) of WSSV *HindIII* and *SalI* genomic fragments were constructed (Wang *et al.*, 1995). These libraries contain WSSV genomic DNA fragments with an ORF showing high homology to *rr1*, but no ORF for *rr2*. Plasmid libraries of *Eco*RI and *HindIII/BamHI* genomic fragments (pme and pmbh libraries) of WSSV Taiwan isolate were therefore constructed from virus purified from the same batch of frozen *P. monodon* collected in 1994. A DIG-labeled PCR product amplified from WSSV DNA template and a degenerate primer set *rr2*-

DF1/rr2-DR1 (5'-CABRTWCTKGCKTTCTTTGC-3'/5'-RTC-DGCMACAAAYTCAATGTAY-3') for rr2 were then used to identify clones that contained the rr2 ORF.

DNA sequencing and computer analysis

Plasmid clones from the Sall library (pms54b, pms420), the HindIII library (pmh21), and the EcoRI library (pme257) carrying rr1 and rr2 were sequenced on both DNA strands by using universal M13 forward and reverse primers. The internal sequences of the cloned fragments were obtained by automatic sequence walking (Mission Biotech, Taiwan) using custom synthesized primers. All of the sequences were confirmed by sequencing both strands completely. Sequence data were compiled and analyzed using three computer programs: GeneWorks (IntelliGenetics, Campbell, CA), UWGCG (release 9.0; Genetics Computer Group, Madison, WI), and Neural Network for promotor prediction (NNPP) (Reese, 1994; Reese and Eeckman, 1995; Reese et al., 1996). The DNA and the deduced amino acid sequences were compared with the latest GenBank/EMBL, SWISSPORT, and PIR databases using FASTA and BLAST. Alignments of amino acid sequences were made in CLUSTAL X (Thompson et al., 1997) and edited in GeneDoc (Nicholas et al., 1997a,b)

Southern blot analysis

Southern blot analysis (Southern, 1975) was used to determine the number of *rr1* and *rr2* copies in WSSV. The probes for this analysis were generated by PCR with the primer sets *rr1*-F1/*rr1*-R1 (5'-GACATGATTATGCGTGTGG-3'/5'-CTCCCGGTAATCAACATAG-3') for *rr1* and *rr2*-F1/*rr2*-R1 (5'-ATGCAAGTAGATACGTTTAAACAC-3'/5'-CAATTCGGACGGCTTCAGTGAC-3') for *rr2*. WSSV DNA was digested with *Sal*I, *Hind*III, or *Eco*RI restriction enzymes, separated in 0.7% agarose gel, transferred to Hybond-N⁺ membrane (Amersham, Arlington Heights, IL), and then hybridized with the DIG-labeled WSSV *rr1* or *rr2* probe. DIG-labeled nucleotides in the blots were detected as described previously (Lo *et al.*, 1999).

RNA isolation

Healthy (two-step WSSV diagnostic PCR negative) (Lo et al., 1996a) subadult P. monodon (15–20 g) were infected with WSSV by injection using the method described previously by Tsai et al. (1999). At various times after injection, two or three shrimps were randomly selected and their pereiopods were excised. The pereiopods were immediately frozen and stored in liquid nitrogen. Total RNA was extracted from the frozen pereiopods as described in Tsai et al. (2000) and stored in 75% ethanol at $-20\,^{\circ}$ C.

98 TSALET AL.

Determination of the 5' terminal region of the *rr1* and *rr2* transcripts

The 5' regions of the rr1 and rr2 transcripts were obtained by rapid amplification of the cDNA 5' ends (5' RACE) (Frohman et al., 1988) using a commercial 5'/3' RACE kit (Boehringer Mannheim, Indianapolis, IN) according to the instructions provided by the manufacturer. Total RNA was prepared from pereiopods of shrimp that had been injected with WSSV inoculum 18 h previously. The appropriate gene-specific primers (rr1-R2 [5'-CCGAGTCCAGCAGTCTTG-3'] for the rr1 gene and rr2-R1 for the rr2 gene) were then used for cDNA synthesis. Before being subjected to PCR, a "poly A head" with terminal transferase was added to the cDNA products in the presence of dATP. The firstround PCR for rr1 was performed using the primer rr1-R3 (5'-GTTCGAACAGATGAATCTTCCTGC-3') and an oligo dT-anchor. The PCR product after the first round was used as the template for the second round of amplification using primer rr1-R4 (5'-GACGGTCCAT-GATGTGAGATGC-3') and anchor primer (5'/3' RACE kit). For rr2 transcription analysis, primers rr2-R2 (5'-CACTTGACGAAGACGTGTTGTAAG-3') and rr2-R3 (5'-GAACGCCAGAATC TGGAGAATG-3') were used in the first and second rounds of amplification, respectively. The final products were cloned into pGEM-T Easy vector (Promega, Madison, WI) and sequenced; the resulting sequences were compared with the genomic sequence.

Determination of the 3' terminal region of *rr1* and *rr2* transcripts

The 3' regions of the *rr1* and *rr2* transcripts were determined by 3' RACE. First-strand cDNA was synthesized using the oligo dT-anchor primer. The resulting cDNA was amplified with the anchor and the appropriate primer (*rr1*-F2 [5'-GGCAAGTGGTGGATCTATTCA-3'] for *rr1* transcript, and *rr2*-F1 for *rr2* transcript). The final products were subcloned and sequenced, and the resulting sequences were compared with the genomic sequence.

RT-PCR

Total RNA was subjected to RT-PCR as described in Tsai *et al.* (2000), except that only the rr1-F1/rr1-R1 (5'-GACATGATTATGCGTGTGG-3'/5'-CTCCCGGTAATCAA-CATAG-3') and rr2-F1/rr2-R1 (5'-ATGCAAGTAGATACGTT-TAAACAC-3'/5'-CAATTCGGACGGCTTCAGTGAC-3') primer sets were used. The β -actin gene served as an internal control for RNA quality and amplification efficiency (Actin-F1/Actin-R1 [5'-GAYGAYATGGAGAA-GATCTGG-3'/5'-CCRGGGTACATGGTGGTRCC-3']).

Northern blot hybridization analysis with *rr1-* and *rr2-* specific riboprobes

The rr1- and rr2-specific [α - 32 P]rCTP-labeled riboprobes were generated by $in\ vitro$ transcription (Sambrook $et\ al.$, 1989) using T7 RNA polymerase (Boehringer Mannheim). The RNA polymerase promoter addition kit Lig'nScribe (Ambion Inc., Austin, TX) was used in accordance with the manufacturer's instructions to generate templates from rr1- and rr2-specific PCR product for the $in\ vitro$ transcription. T7 RNA polymerase promoter was added to the PCR fragment by the Lig'nScribe reaction. The rr1-specific PCR product was amplified by rr1-R1/rr1-F1 primers, and the rr2-specific PCR product was amplified by rr2-R1/rr2-F1 primers. Approximately 10 μ g total RNA were used for Northern blot hybridization; protocols followed Sambrook $et\ al.$ (1989).

ACKNOWLEDGMENTS

This work was supported by the Council of Agriculture Grant 88-BT-2.1-FID-01(4-1) and National Science Council Grants NSC88-2311-B-002-022-B20 and NSC89-2311-B-002-040. We are indebted to Paul Barlow for his helpful criticism of the manuscript.

REFERENCES

- Ahrens, C. H., Russell, R. L., Funk, C. J., Evans, J. T., Harwood, S. H., and Rohrmann, G. F. (1997). The sequence of the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus genome. *Virology* 229, 381–399.
- Blissard, G. W., and Rohrmann, G. F. (1989). Location, sequence transcriptional mapping, and temporal expression of the gp64 envelope glycoprotein gene of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 170, 537–555.
- Boursnell, M., Shaw, K., Yáñez, R. J., Viñuela, E., and Dixon, L. (1991). The sequence of the ribonucleotide reductase genes from African swine fever virus show considerable homology with those of the orthopoxvirus, vaccinia virus. *Virology* **184**, 411–416.
- Chou, H. Y., Huang, C. Y., Wang, C. H., Chiang, H. C., and Lo, C. F. (1995). Pathogenicity of a baculovirus infection causing white spot syndrome in cultured penaeid shrimp in Taiwan. *Dis. Aquat. Org.* 23, 165–173.
- Frohman, M. A., Dush, M. K., and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
- Francki, R. I. B., Fauquet, C. M., Knudson, D. L., and Brown, F. (1991). "Classification and Nomenclature of Viruses: Fifth Report of the International Committee on Taxonomy of Viruses." Springer-Verlag, New York.
- IJkel, W. F. J., van Strien, E. A., Heldens, J. G. M., Broer, R., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1999). Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome. *J. Gen. Virol.* 80, 3289–3304.
- Jordan, A., and Reichard, P. (1998). Ribonucleotide reductase. *Ann. Rev. Biochem.* **67**, 71–98.
- Kuzio, J., Pearson, M. N., Harwood, S. H., Funk, J. T., Slavicek, J. M., and Rohrman, G. F. (1999). Sequence and analysis of the genome of a baculovirus pathogenic for *Lymantria dispar. Virology* **253**,17–34.
- Lo, C. F., Ho, C. H., Chen, C. H., Liu, K. F., Chiu, Y. L., Yeh, P. Y., Peng, S. E., Hsu, H. C., Liu, H. C., Chang, C. F., Su, M. S., Wang, C. H., and Kou, G. H. (1997). Detection and tissue tropism of white spot syndrome baculovirus (WSBV) in captured brooders of *Penaeus mon-*

- odon with a special emphasis on reproductive organs. *Dis. Aquat. Org.* **30,** 53-72.
- Lo, C. F., Ho, C. H., Peng, S. E., Chen, C. H., Hsu, H. C., Chiu, Y. L., Chang, C. F., Liu, K. F., Su, M. S., Wang, C. H., and Kou, G. H. (1996b). White spot syndrome baculovirus (WSBV) detected in cultured and captured shrimp, crab and other arthropods. *Dis. Aquat. Org.* 27, 215–225.
- Lo, C. F., Hsu, H. C., Tsai, M. F., Ho, C. H., Peng, S. E., Kou, G. H., and Lightner, D. V. (1999). Specific genomic fragment analysis of different geographical clinical samples of shrimp white spot syndrome virus. *Dis. Aquat. Org.* **35**, 175–185.
- Lo, C. F., Leu, J. H., Ho, C. H., Chen, C. H., Peng, S. E., Chen, Y. T., Chou, C. M., Yeh, P. Y., Huang, C. J., Chou, H. Y., Wang, C. H., and Kou, G. H. (1996a). Detection of baculovirus associated with white spot syndrome (WSBV) in penaeid shrimps using polymerase chain reaction. *Dis. Aquat. Org.* **25**, 133–141.
- Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D. (1995). "Virus Taxonomy—The Classification and Nomenclature of Viruses: Sixth Report of the International Committee on Taxonomy of Viruses." Springer-Verlag, New York.
- Nadala, E. C. B., Jr., and Loh, P. C. (1998). A comparative study of three different isolates of white spot virus. *Dis. Aquat. Org.* **33**, 231–234.
- Nicholas, K. B., Nicholas, H. B., Jr., and Deerfield II, D. W. (1997a). GeneDoc: Analysis and visualization of genetic variation, *EMBNEW*. NEWS 4, 14.
- Nicholas, K. B., and Nicholas, H. B., Jr. (1997b). GeneDoc: Analysis and visualization of genetic variation. http://www.cris.com/~Ketchup/genedoc.shtml.
- Reese, M. G. (1994). Erkennung von Promotoren in pro- und eukaryontischen DNA-Sequenzen durch k stliche Neuronale Netze. Master Thesis. German Cancer Research Center, Heidelberg, Germany.
- Reese, M. G., and Eeckman, F. H. (1995). New neural network algorithms for improved eukaryotic promoter site recognition. *In* "Proceedings of the Seventh International Genome Sequencing and Analysis Conference." Hilton Head Island, SC.
- Reese, M. G., Harris, N. L., and Eeckman, F. H. (1996). Large scale sequencing specific neural networks for promoter and splice site recognition. *In* "Biocomputing: Proceedings of the 1996 Pacific Symposium" (L. Hunter and T. E. Klein, Eds.). World Scientific Publishing, Singapore.

- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503–517.
- Takahashi, Y., Itami, T., Kondo, M., Maeda, M., Fujii, R., Tomonaga, S., Supamattaya, K., and Boonyaratpalin, S. (1994). Electron microscopic evidence of bacilliform virus infection in Kuruma shrimp. *Fish Pathol.* 29, 121–125.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882.
- Tsai, M. F., Kou, G. H., Liu, H. C., Liu, K. F., Chang, C. F., Peng, S. E., Hsu, H. C., Wang, C. H., and Lo, C. F. (1999). Long-term presence of white spot syndrome virus (WSSV) in a cultured shrimp population without disease outbreaks. *Dis. Aquat. Org.* 38, 107–114.
- van Hulten, M. C. W., Tsai, M. F., Schipper, C. A., Lo, C. F., Kou, G. H., and Vlak, J. M. (2000a). Analysis of a genomic segment of white spot syndrome virus of shrimp containing ribonucleotide reductase genes, and repeat regions. *J. Gen. Virol.* 81, 307–316.
- van Hulten, M. C. W., Westenberg, M., Goodall, S. D., and Vlak, J. M. (2000b). Identification of two major virion protein genes of white spot syndrome virus of shrimp. *Virology* **266**, 227–236.
- van Strien, E. A., Faktor, O., Hu, Z. H., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1997). Baculoviruses contain a gene for the large subunit of ribonucleotide reductase. *J. Gen. Virol.* **78**, 2365–2377.
- Wang, C. H., Lo, C. F., Leu, J. H., Chou, C. M., Yeh, P. Y., Chou, H. Y., Tung, M. C., Chang, C. F., Su, M. S., and Kou, G. H. (1995). Purification and genomic analysis of baculovirus associated with white spot syndrome (WSBV) of *Penaeus monodon. Dis. Aquat. Org.* 23, 239–242.
- Wongteerasupaya, C., Vickers, J. E., Sriurairatana, S., Nash, G. L., Akarajamorn, A., Boonsaeng, V., Panyim, S., Tassanakajon, A., Withyachumnarnkul, B., and Flegel, T. W. (1995). A non-occluded, systemic baculovirus that occurs in cells of ectodermal and mesodermal origin and causes high mortality in the black tiger prawn *Penaeus monodon. Dis. Aquat. Org.* 21, 69–77.
- Yang, F., Wang, W., Chen, R. Z., and Xu, X. (1997). A simple and efficient method for purification of prawn baculovirus DNA. *J. Virol. Methods* **67,** 1–4.