Virology 380 (2008) 213–225

Contents lists available at ScienceDirect

Virology

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

Genetic diversity in the yellow head nidovirus complex

Priyanjalie K.M. Wijegoonawardane [a,1], Jeff A. Cowley [a], Thuy Phan [a], Richard A.J. Hodgson [a,2], Linda Nielsen [c,d,3], Wansika Kiatpathomchai [c,d], Peter J. Walker [a,b,*]

[a] CSIRO Livestock Industries, Queensland Bioscience Precinct, 306 Carmody Road, St. Lucia, Queensland 4067, Australia
[b] CSIRO Livestock Industries, Australian Animal Health Laboratory (AAHL), 5 Portarlington Road, Geelong, Victoria 3220, Australia
[c] Centex Shrimp, Faculty of Science, Mahidol University, Rama 6 Road, Bangkok 10400, Thailand
[d] National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Thailand Science Park, Klong Luang, Patumthani, 12120, Thailand

ARTICLE INFO

Article history:
Received 18 April 2008
Returned to author for revision 23 May 2008
Accepted 8 July 2008
Available online 2 September 2008

Keywords:
Yellow head virus
Gill-associated virus
Phylogeny
Replicase gene
Genotypic variation
Genotypes

ABSTRACT

Penaeus monodon shrimp collected from across the Indo-Pacific region during 1997–2004 were screened for the presence of yellow head-related viruses. Phylogenetic analyses of amplified ORF1b gene segments identified at least six distinct genetic lineages (genotypes). Genotype 1 (YHV) was detected only in shrimp with yellow head disease. Genotype 2 (GAV) was detected in diseased shrimp with the less severe condition described as mid-crop mortality syndrome and in healthy shrimp from Australia, Thailand and Vietnam. Other genotypes occurred commonly in healthy shrimp. Sequence comparisons of structural protein genes (ORF2 and ORF3), intergenic regions (IGRs) and the long 3′-UTR supported the delineation of genotypes and identified both conserved and variant structural features. In putative transcription regulating sequences (TRSs) encompassing the sub-genomic mRNA 5′-termini, a core motif (5′-GUCAAUU*ACAAC*-3′) is absolutely conserved. A small (83 nt) open reading frame (ORF4) in the 3′-UTR of GAV is variously truncated in all other genotypes and a TRS-like element preceding ORF4 is invariably corrupted by a *A*>*G/U* substitution in the central core motif (5′-UU(*G/U*)*CAAC*-3′). The data support previous evidence that ORF4 is a non-functional gene under construction or deconstruction. The 3′-UTRs also contain predicted 3′-terminal hairpin-loop structures that are preserved in all genotypes by compensatory nucleotide substitutions, suggesting a role in polymerase recognition for minus-strand RNA synthesis.

Introduction

Yellow head disease (YHD) emerged in 1990 in farmed black tiger shrimp (*Penaeus monodon*) in Thailand (Limsuwan, 1991). It is a devastating disease that can cause total crop losses within a few days following the appearance of gross signs (Chantanachookin et al., 1993). The causative agent, yellow head virus (YHV), is a bacilliform, enveloped, (+) ssRNA virus classified in the genus *Okavirus*, family *Roniviridae* within the order *Nidovirales* (Walker et al., 2005). A related nidovirus, gill-associated virus (GAV), occurs commonly as a chronic infection in healthy wild and farmed *P. monodon* in eastern Australia (Spann et al., 1995; Cowley et al., 2000a; Walker et al., 2001). GAV has been associated with a disease described as mid-crop mortality syndrome (MCMS) in which mortalities progressively accumulate from the mid-late juvenile stage onwards (Spann et al., 1997; Walker et al., 2001). A variant YHV genotype has also been detected in healthy P. *monodon* broodstock from Thailand (Soowannayan et al., 2003) and YHV or related viruses have been reported in *P. monodon* and *Penaeus japonicus* farmed in Taiwan (Wang et al., 1996; Wang and Chang, 2000). However, these viruses were not associated with typical YHD and their relationships to YHV or GAV remain unconfirmed. There are also reports of YHD or YHV in *P. monodon* in several other countries in the Asian region including the Philippines, India, Indonesia, Sri Lanka, Malaysia, Vietnam and China (Walker et al., 2001) but these have rarely been confirmed by laboratory analysis.

Like other nidoviruses, the polyadenylated, (+) RNA genomes of YHV (26,662 nt) and GAV (26,235 nt) are expressed from a nested set of genomic and subgenomic mRNAs (Cowley and Walker, 2002; Cowley et al., 2002a; Sittidilokratna et al., 2008). The genomes are organized similarly (5′-ORF1a/ORF1b-ORF2-ORF3-(ORF4)-polyA-3′). ORF1a encodes a long polyprotein (pp1a) containing 3C-like and papain-like proteases required for auto-processing, and overlaps ORF1b which encodes a 'SDD' RNA-dependent RNA polymerase, helicase, metal-ion-binding, exonuclease, uridylate-specific endoribonuclease and ribose-2′-*O*-methyl transferase domains (Cowley et al., 2000b; Ziebuhr et al., 2003; Sittidilokratna et al., 2002, 2008). A -1

* Corresponding author. Fax: +61 3 5227 5000.
E-mail address: Peter.Walker@csiro.au (P.J. Walker).
[1] National Aquatic Resources Research and Development Agency, Crow Island, Mattakkuliya, Colombo 15, Sri Lanka.
[2] Faculty of Health Science & Medicine, Bond University, Gold Coast, QLD 4229 Australia.
[3] International School Bangkok, 39/7 Nichadathani, Samaki Road, Pakred, Nonthaburi 11120, Thailand.

0042-6822/$ – see front matter.
doi:10.1016/j.virol.2008.07.005

ribosomal frame-shift element at the ORF1a/ORF1b overlap facilitates translation of the complete pp1ab polyprotein (Cowley et al., 2002a). ORF2 encodes the nucleoprotein (p20) and ORF3 encodes a polyprotein that is processed to generate two envelope glycoproteins (gp116 and gp64) and a small N-terminal triple-membrane-spanning (TMS) fragment of unknown function (Cowley and Walker, 2002; Jitrapakdee et al., 2003; Cowley et al., 2004a; Sittidilokratna et al., 2006). ORF4 is a short open reading frame that, although possibly expressed in GAV (Cowley and Walker, 2008), is severely truncated in YHV and unlikely to encode a functional protein (Sittidilokratna et al., 2008). Intergenic regions (IGRs) upstream of ORF2 and ORF3 contain conserved transcription regulatory sequences (TRSs) that facilitate transcription of two sub-genomic mRNAs (sgmRNAs) which, together with the genome-length mRNA, form a 3′-coterminal nested set characteristic of nidoviruses (Cowley et al., 2002a; Sittidilokratna et al., 2008). The genome sequences of YHV and GAV are ~79% identical, with amino acid sequence identity ranging from ~73% in gp116 to ~84% in pp1ab (Sittidilokratna et al., 2008). Based on levels of sequence divergence, YHV and GAV have been regarded as closely related variants that are likely to represent geographic topotypes in a larger complex (Cowley et al., 1999; Walker et al., 2001). Currently, they are each classified as the species *Gill-associated virus* (Walker et al., 2005).

This paper reports the analysis of tissue samples from healthy and diseased *P. monodon* collected from several countries in the Indo-Pacific region. Analysis of partial ORF1b sequences identified at least six distinct genetic lineages (genotypes) of yellow head-related viruses of which only YHV (genotype 1) was detected in shrimp from YHD-outbreak ponds. Sequencing of the structural gene region of representatives of genotypes 3, 4 and 5 identified significant diversity, particularly in IGRs and the N-terminal gp116 region, but preservation of TRS-like elements and a 3′-terminal hairpin structure likely to be involved in negative-strand RNA synthesis.

Results

A new genotype (genotype 3) detected in healthy P. monodon broodstock and postlarvae

In an initial study, a semi-nested RT-PCR targeting a 279 nt ORF1b sequence located just downstream of helicase domain motif VI was used to detect YHV-related viruses in 38 apparently healthy *P. monodon* broodstock sampled from commercial hatcheries in Central Thailand in March 2000. PCR products were amplified from 25 (66%) of the broodstock (Fig. 1A). Nucleotide sequences were determined for six of these in addition to PCR products amplified from juvenile shrimp from YHD outbreaks in Thailand and Taiwan and healthy broodstock and postlarvae from hatcheries in Australia, Vietnam and Sarawak in Malaysia (Table 1). Using 25 sequences, ClustalX was used to align the 231 nt segment bounded by the nested PCR primers (equivalent to the sequence A^{17492} to A^{17722} in the GAV reference strain; GenBank AF227196). An unrooted neighbour-joining phylogenetic tree generated from this alignment segregated the sequences into three clusters (Fig. 1B). One cluster (genotype 1) included the YHV reference strain (GenBank AY052786) and all eight viruses from juvenile shrimp from YHD outbreaks. Whilst slight nucleotide variation occurred between some of these viruses, the sequences of one Thai (THA-00-D11) and three of the four Taiwanese viruses (TWN-00-D1, TWN-00-D2, TWN-00-D3) were identical. The second cluster (genotype 2) included the GAV reference strain, the two other Australian viruses from either healthy or MCMS-affected shrimp, and one of the five Vietnamese viruses from healthy postlarvae (VNM-01-H65). The third phylogenetic cluster (genotype 3), which has not been described previously, comprised all viruses from healthy broodstock from Thailand and Sarawak and the other four Vietnamese viruses detected in healthy postlarvae. The Vietnamese viruses appeared to form a sub-group within genotype 3.

Other genotypes detected in P. monodon from various Indo-Pacific regions

A more extensive study was conducted using a larger set of *P. monodon* sampled over a wider geographic range. A new RT-nested PCR designed to amplify a 722 nt ORF1b gene segment encompassing the semi-nested RT-PCR sequence was used in initial analyses. This test was further modified to utilize degenerate primer pairs YH30-F2/R2 and YHV31-F2/R2 to better accommodate sequence variations among the known genotypes and was applied to ~200 *P. monodon* sampled between 1997 and 2004 from Indonesia, the Philippines, Taiwan, Vietnam, Thailand, Malaysia, India, Sri Lanka, Mozambique and Australia (Table 1). Samples included whole postlarvae as well as gill, lymphoid organ, hepatopancreas or whole head tissues of juvenile farmed shrimp or adult broodstock from hatcheries. Most samples were from healthy shrimp but some originated from shrimp either displaying typical signs of YHD or that were moribund and collected from MCMS-affected ponds in Australia.

In total, 57 of the ~200 samples generated nested PCR amplicon yields suitable for sequencing. Sequences of these 57 viruses, together with YHV and GAV reference strain sequences, were aligned using ClustalX. An unrooted neighbour-joining tree generated from the alignment of the 688–671 nt segment (equivalent to the GAV reference strain sequence G^{17259} to A^{17929}) segregated the viruses into six major clusters well supported by bootstrap values (>70%) (Fig. 2). The clustering of viruses in genotypes 1, 2 and 3 substantiated initial findings with the 231 nt sequence in that all 13 viruses clustering with YHV (genotype 1) originated from juvenile YHD-affected shrimp and all seven viruses from healthy and diseased Australian shrimp clustered with GAV (genotype 2). Ten Thai and Vietnamese viruses from healthy shrimp also clustered with GAV. The 17 genotype 3 viruses displayed the widest geographic distribution, being identified in healthy broodstock and postlarvae from hatcheries in Thailand, Vietnam, Taiwan, Indonesia, Malaysia and Mozambique. The three viruses from Malaysian shrimp also appeared to form a subcluster in genotype 3.

In addition to these genotypes, the tree delineated three new clusters. One lineage (genotype 4) was clearly separated from all other genotypes and comprised viruses detected in three healthy postlarvae batches sampled from hatcheries in Nellore, India. Another lineage (genotype 6) was most closely related to genotype 2 (GAV) and comprised viruses detected in five of the six broodstock from Mozambique. The other Mozambique virus clustered in genotype 3. However, as Mozambique shrimp were supplied from a commercial breeding facility in Malaysia, it is possible this sample was mislabeled. Another lineage (genotype 5) comprised viruses derived from a healthy sub-adult from Malaysia (MYS-03-H4), a sub-adult from Thailand displaying slower than normal growth (THA-03-SG21), and a batch of healthy postlarvae sampled from the Philippines (PHL-03-H8). The relationships among the three genotype 5 viruses were more disparate than viruses clustering within the other lineages. Pair-wise alignments indicated that THA-03-SG21 and MYS-03-H4 share 97.3% sequence identity and that the levels of identity of these two viruses to PHL-03-H8 (92.8% and 93.4%) was less than that distinguishing viruses clustered in genotypes 2 and 6 (~96.5%) and comparable to that distinguishing viruses clustered in genotypes 2 and 3 (~93.3%). PHL-03-H8 was the only YH-related virus detected among 18 *P. monodon* sampled from the Philippines and it is likely that genotype 5 will resolve into a two separate lineages once more viruses are analyzed. Phylogenetic analysis using the maximum parsimony method clustered the viruses similarly into six genotypes (data not shown).

Translation of the 57 amplified ORF1b gene fragments revealed 35 unique 223 aa sequences that extended from the start of the pp1ab helicase domain motif V to a locus 173 aa downstream of motif VI (Cowley et al., 2000b; Sittidilokratna et al., 2002). An unrooted phylogenetic tree was generated from a ClustalX alignment of the