

The complete genome sequence of gill-associated virus of *Penaeus monodon* prawns indicates a gene organisation unique among nidoviruses*

Brief Report

J. A. Cowley and P. J. Walker

Cooperative Research Centre for Aquaculture, CSIRO Livestock Industries,
Long Pocket Laboratories, Indooroopilly, Australia

Received March 20, 2002; accepted May 17, 2002
Published online August 12, 2002 © Springer-Verlag 2002

Summary. We report here a 5596 nt sequence comprising the 3'-end of the (+) ssRNA genome of gill-associated virus (GAV), an invertebrate nidovirus of *Penaeus monodon* prawns. The sequence extends from a subgenomic RNA start site 35 nt upstream of the 4923 nt ORF3 gene to a 3'-poly(A) tail of the 26235 nt genome of GAV. The putative 1640 amino acid (aa) ORF3 protein (MW = 182049 Da, pI = 6.62) contains 15 potential N-linked glycosylation sites, 15 potential O-linked glycosylation sites and six highly hydrophobic regions predicted to represent transmembrane (TM) domains. Three of the predicted TM domains occur in the amino-terminal 228 aa, two in the central portion, and one near the carboxy-terminus of ORF3. Only one short (83 aa) open reading frame (ORF4) was identified between ORF3 and the 3'-poly(A) tail. Completion of the genome sequence of GAV has revealed a gene organisation unique among nidoviruses in there is no discrete membrane protein gene and that the putative ORF3 spike glycoprotein gene resides downstream of a gene (ORF2) encoding a structural protein associated with nucleocapsids.

*

Gill-associated virus (GAV) of *Penaeus monodon* prawns has been proposed as a type member of a new genus, *Okavirus*, within the *Nidovirales* [9]. Chronic GAV infection is highly prevalent in wild and farmed *P. monodon* in eastern Australia [8, 39, 43] and the virus has been associated with mortalities in farmed

*The sequence described in this paper has been deposited in GenBank (Acc: AY039647).

prawns [40]. GAV is morphologically identical to yellow head virus (YHV), which has caused significant losses of *P. monodon* farmed throughout Asia [25]. GAV and YHV nucleocapsids appear as helical tubular structures (~16 nm dia.) and their rod-shaped, enveloped virions (~34 nm × 192 nm) possess diffuse surface projections of ~8 nm [4, 6, 39, 40, 44]. YHV contains a >22 kb ssRNA genome [27, 45] with suggested positive polarity [41]. The GAV genome 5'-end contains a ~20 kb replicase gene encoding overlapping ORF1a and ORF1b polyproteins [9]. Translation of the ORF1a-1b polyprotein is facilitated by a AAAUUUU-1 ribosomal frameshift site and a predicted non-'H-type' RNA pseudoknot. Sequence similarity in the 3C-like protease domain of ORF1a and the 'SDD' polymerase, helicase and other domains of ORF1b indicate a distant evolutionary relationship to vertebrate viruses of the *Nidovirales* [9]. Conserved sequences in the GAV intergenic regions function as transcription start sites for two 3'-coterminal subgenomic (sg)RNAs that, in contrast to the sgRNAs of coronaviruses and arteriviruses, are not fused to 5'-genomic leader sequences [10]. Moreover, the ORF2 gene immediately downstream of ORF1a-1b encodes the 144 amino acid protein similar in size and charge to torovirus nucleocapsid (N) proteins [14, 34] and which immuno-electron microscopy indicates is associated with GAV nucleocapsids [JA Cowley et al. unpubl.], suggesting it is likely to represent the nucleocapsid protein.

In this paper, we report a 5596 nt sequence downstream of ORF2 to a 3'-poly(A) tail which, together with the upstream ORF2 sequence included in GenBank Acc: AY039647, completes the 26235 nt (+) ssRNA genome of GAV. The region contains a 4.9 kb ORF3 gene encoding a putative 1640 amino acid (182 kDa) glycoprotein and only one short (83 aa) potential ORF4 in the 638 nt region between ORF3 and the 3'-poly(A) tail. The GAV genome organisation is, therefore, markedly different to any currently known for viruses within the *Nidovirales*.

Random cDNA clones were generated from a ~22 kb dsRNA purified from lymphoid organ total RNA of GAV-infected *P. monodon* by RT-PCR [15] as described previously [7, 8]. Three 4.3 kb clones of p18/20 (i.e. 1, 9 and 16) and a truncated 3.0 kb clone (p33) that contained large portions of the GAV ORF3 gene were generated by RT-PCR using primers GAV18 and GAV20 as described previously [9].

Clones of the GAV genome 3'-terminus were generated using 3 different methods. (a) A random RT-PCR method [15] was modified to generate clones from oligo-dT-primed cDNA. Briefly, cDNA was reverse transcribed from the purified ~22 kb dsRNA using the anchored primer Uni-dT₁₆-C/A/G (5'GCCGGAGCTCTGCAGAATTC(T)₁₆-C/A/G^{3'}). Second strand cDNA was synthesised using Klenow DNA polymerase and primer Uni-dN₆ (5'GCCGGAGCTCTGCAGAATTC(N)₆^{3'}). Primer was removed following each reaction using a Sephadex 400 HR column (Pharmacia) and cDNA was amplified by PCR using Uni-primer (5'GCCGGAGCTCTGCAGAATTC^{3'}) and Taq DNA polymerase (Promega) [9]. PCR products of variable size were purified using a QIAquick column (QIAGEN) and cloned into pGEM-T (Promega) using standard procedures [31]. (b) A cDNA

clone was generated similarly using lymphoid organ total RNA from *P. monodon* experimentally-infected with GAV and cDNA synthesised using primer Uni-dT₁₆-C/A/G followed by PCR amplification using Uni-primer and the GAV sense primer GAV69 (5' TCATATTACATCGCTGGTGATCCAGG^{3'}) designed to a sequence in clone p18/20. A predominant 2.7 kb product was excised from an agarose gel, purified using a QIAquick column and cloned into pGEM-T. (c) cDNA clones of the GAV genome 3'-end were also obtained from a Uni-ZAP[®] XR (Stratagene) λ bacteriophage library kindly provided by Dr. Sigrid Lehnert, CSIRO Livestock Industries. Briefly, oligo-dT₁₈-primed cDNA was prepared from poly(A)+ RNA purified from a whole head homogenate of a *P. monodon* naturally-infected with GAV, ligated into the Uni-ZAP XR vector and packaged according to the ZAP-cDNA[®] Gigapack[®] III Gold cloning kit (Stratagene) instructions. The λ phage library in SM buffer (100 mM NaCl, 8 mM MgSO₄, 50 mM Tris-HCl pH 7.5, 0.01% gelatin) was diluted 1:10 in water, heated at 100 °C for 10 min and micro-centrifuged for 5 min. An aliquot (1 μ l) was amplified by PCR using Taq polymerase (Promega), an extended T7 primer (5' GTAATACGACTCACTATGGGC^{3'}) and the GAV sense primer GAV86 (5' GCTCCACACAGGCACTACCG^{3'}) using 40 cycles of 95 °C/45 sec, 60 °C/45 sec, 72 °C/1 min. PCR products derived from GAV-specific poly(A)+ RNA were cloned into pGEM-T. Plasmid inserts were sequenced using FS-Big Dye reagent (ABI), T7, SP6 or GAV-specific primers and ABI Model-377 sequencing apparatus at the Australian Genome Research Facility, Brisbane. Sequences were edited using SeqEd 3.0.1 (ABI) and analysed using MacVector 7 (Oxford Molecular).

A contig of a 5596 nt sequence from the GAV genomic 3'-poly(A) tail to an intergenic 5'-sgRNA start site 35 nt upstream of the ORF3 start codon [10] was deduced from overlapping cDNA clones (Figs. 1 and 2). The complete GAV genome sequence is deposited in GenBank Acc no: AF227196 and AY039647 and the ORF2 gene sequence will be described elsewhere. The 3'-terminal sequence contains a 4923 nt ORF3 gene encoding a putative 1640 amino acid (aa) protein with a deduced molecular weight of 182049 Da and pI of 6.62. No polypeptides with significant similarity to the putative ORF3 protein were identified in BLAST 2.0 [1] searches of GenBank. ORF3 contains 15 potential N-linked glycosylation sites and the NetOGly 2.0 program [18] predicted 15 potential O-linked glycosylation sites predominantly located in the N-terminal half of ORF3. Six highly hydrophobic regions (G²⁷-F⁵⁰, S¹³⁸-T¹⁵⁶, F²¹⁰-A²²⁸, A⁹⁹⁶-L¹⁰¹⁸, F¹⁰⁸¹-A¹¹⁰³ and Y¹⁶⁰⁵-I¹⁶²⁷), probably representing transmembrane (TM) domains, were predicted using TMpred [20] (Figs. 2 and 3). Prediction of membrane topology using TMHMM [38] identified three likely ORF3 ectodomains (Fig. 3). The putative ectodomains, ORF3a (4 Cys), ORF3b (24 Cys) and ORF3c (22 Cys) following TM domains 1, 3 and 5 contained even numbers of Cys residues reflecting the likelihood that they are involved in intra-domain disulphide-bond formation. No heptad-repeat patterns indicative of the coiled-coil structures common to the structural glycoproteins of coronaviruses [11] and toroviruses [35] were predicted in ORF3 using the COILS program [26]. The absence of

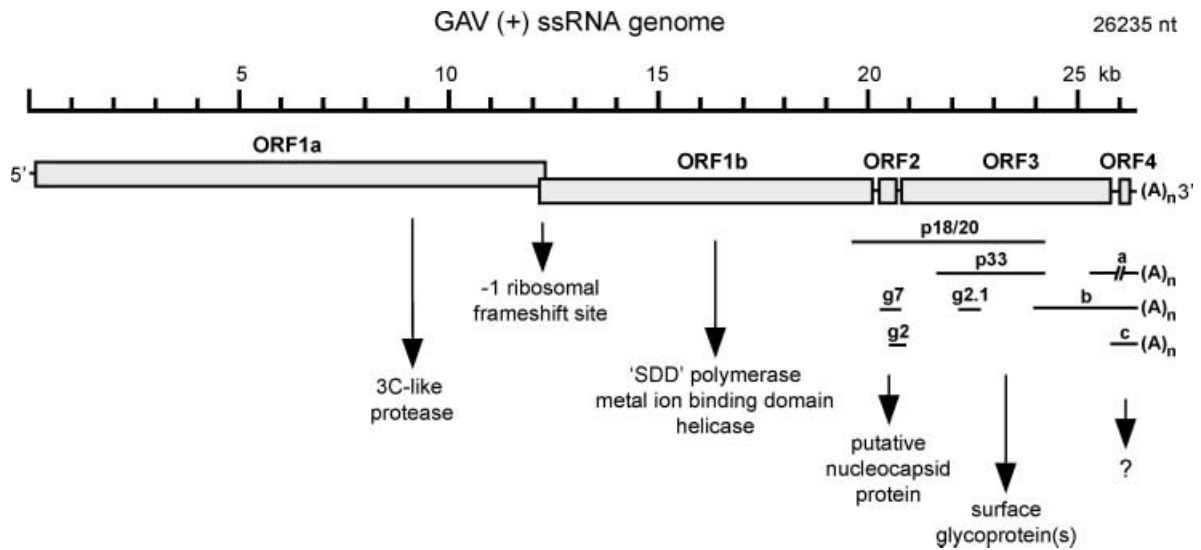


Fig. 1. Organisation of the 26235 nt (+) ssRNA GAV genome indicating translation features and deduced ORF functions. The relative positions of clones used to determine the sequence of the ~5.6 kb region from upstream of ORF3 to the 3'-poly(A) tail are also indicated. Clones were generated by random RT-PCR (*g*), primer-specific RT-PCR (*p*) or by 3 different PCR methods (*a*, *b*, *c*) employing oligo-dT primed cDNA, as described in the text

coiled-coil secondary structures in ORF3 probably accounts for GAV particles lacking the rigid stem-containing surface peplomers that are characteristic of the virions of coronaviruses and toroviruses.

The novel structure predicted for the putative GAV ORF3 glycoprotein suggests that it may be processed and function differently to the S glycoproteins of coronaviruses and toroviruses. At 1640 aa, ORF3 is the largest glycoprotein yet identified for any virus in the *Nidovirales*. The S glycoprotein gene of coronaviruses encodes polypeptides that range in size from 1160 aa for Infectious bronchitis virus [3] to 1452 aa for Feline infectious peritoneal virus [11], while the cognate genes of the toroviruses Berne virus and Breda virus [14, 35] encode polypeptides (1581–1583 aa) approaching the size of GAV ORF3. Moreover, while the S glycoproteins of these viruses contain hydrophobic regions representing a signal peptide and a single TM anchor, six hydrophobic TM domains are predicted in GAV ORF3 (Fig. 3). Although a cluster of basic amino acids (KHLKVHARHHK¹⁰⁶⁹) occurs in the central ORF3 region between TM domains

Fig. 2. Nucleotide and deduced amino acid sequence of the open reading frames (ORF3 and ORF4) in the 5596 nt region of GAV from the 5'-sgRNA start site in the intergenic region upstream of ORF3 to the 3'-poly(A) tail. The six hydrophobic ORF3 regions predicted to represent transmembrane domains are underlined, the 15 potential N-linked glycosylation sites are boxed and the cluster of basic amino acids between hydrophobic regions 4 and 5 are shaded. Potential cleavage sites predicted using SignalP [28] are indicated (↓)

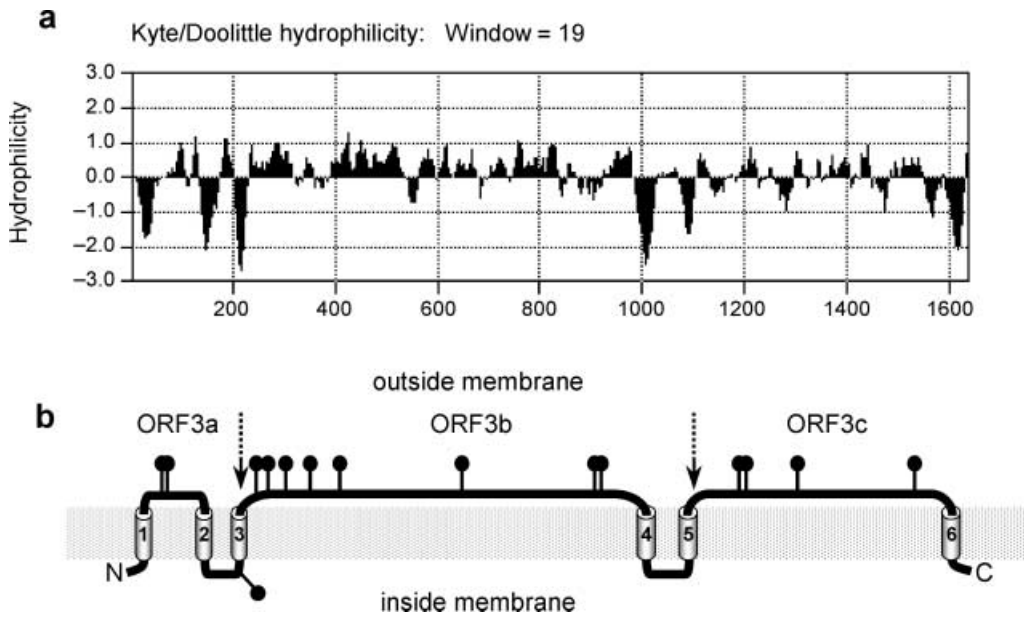


Fig. 3. **a** Hydropathy plot of the 1640 amino acid GAV ORF3 generated using the method of Kyte and Doolittle [22] using a window size of 19. **b** Schematic diagram of the membrane orientation predicted using TMHMM [38] showing the relative positions of the six predicted transmembrane domains, the 15 potential N-linked glycosylation sites (●) and the two cleavage sites (arrows) predicted using SignalP [28] with potential to generate ORF3a, ORF3b and ORF3c proteins

4 and 5, it is unlike the Arg-rich 'RR(F/S/H)RR'-type proteolytic cleavage motifs involved in generating the S1 and S2 glycoprotein components of the virion spikes of coronaviruses and toroviruses [5, 35]. Furthermore, the predicted internal membrane orientation would preclude enzymatic cleavage in this region of GAV ORF3.

The GAV genome contains conspicuously few genes (ORF2, ORF3 and possibly ORF4) other than the ORF1a-1b replicase gene and immuno-electron microscopy using antibodies to the ~20 kDa ORF2 protein indicates that it forms a component of nucleocapsids [JA Cowley et al. unpubl.], suggesting it is likely to be the structural nucleocapsid protein. Purified virions of YHV, which is genetically closely related to GAV [7], contain only three major structural proteins (M_r 110–135, 62–67 and 20–22 kDa), the largest of which is glycosylated [27, 44]. As the GAV ORF2 protein is comparable in size to the 20–22 kDa YHV protein, it can be deduced that ORF3 must encode the virus spike glycoprotein and that post-translational processing would be required to generate polypeptides of a size estimated for the two larger YHV proteins. The GAV ORF3 sequence gives few clues to mechanisms by which such proteins might be generated. Possible cleavage at sites (TFA²²⁸-KE and ASA¹¹⁰¹-LA) predicted after the third and fifth TM domains using SignalP [28] would generate three TM-anchored proteins, ORF3a (228 aa = 25.2 kDa), ORF3b (873 aa = 98.1 kDa) and ORF3c

(539 aa = 58.7 kDa). The theoretical GAV ORF3b and ORF3c proteins approach the size and, assuming that glycosylation could account for some additional mass, are plausibly equivalent to the two largest (i.e. 110–135 and 62–67 kDa) YHV structural proteins [27, 44]. However, this would require the use of these predicted TM domains as internal signal sequences as has been proposed for processing of the multiple membrane-spanning structural glycoproteins of alphaviruses [16, 19, 29] and phleboviruses [21, 30]. Identification of the terminal sequences of the processed GAV spike glycoproteins and experimental investigation of the six hydrophobic domains in ORF3 is required to determine their role in processing, transport and membrane anchoring of the GAV structural glycoproteins.

The presence of three, and in a few cases four, membrane-spanning motifs is a structural characteristic preserved across the integral membrane (M) proteins of coronaviruses [5], toroviruses [3, 14] and arteriviruses [37], even though these viruses possess distinct virion architectures. As GAV appears not to possess a discrete M protein gene, it is tempting to speculate that the putative tri-membrane-spanning ORF3a portion of the ORF3 protein might fulfil a similar function to the vertebrate nidovirus M proteins. However, no structural protein similar in size to the hypothesised 25.2 kDa ORF3a protein of GAV has been reported in purified YHV particles [27, 44]. We are currently preparing antisera to recombinant fusion proteins of components of the predicted ORF3 ectodomains to examine pro-polypeptide processing in GAV-infected prawn tissues and the association and location of processed components in mature virions.

The 638 nt sequence downstream of ORF3 to the GAV 3'-poly(A)-tail contains only one short (83 aa) ORF4 initiating 257 nt downstream of the ORF3 stop codon (Fig. 2). The region between ORF3 and ORF4 contains a 60 nt highly AU-rich sequence (A⁴⁹⁹⁹-A⁵⁰⁵⁸) in addition to a region (A⁵¹⁸⁵-C⁵¹⁹⁸) with limited homology (11/14 nt identical) to the highly conserved sequences encompassing the intergenic sgRNA start sites upstream of ORF2 and ORF3 [10]. However, the A residue corresponding to the start site of the ORF2 and ORF3 sgRNAs is replaced by G residue in the region upstream of ORF4 and we were unable to find evidence of an abundant sgRNA for ORF4 [10]. Further work is required to determine whether an ORF4 protein is translated or whether the entire 638 nt sequence downstream of ORF3 represents the GAV genome 3'-untranslated region.

Oligo-dT-primed cDNA prepared using three independent approaches was used to determine the 3'-terminal sequence of the GAV (+) ssRNA genome. Identification of the 3'-terminus using cDNA primed from the ~22 kb dsRNA purified from GAV-infected prawn tissue confirmed that this dsRNA represents a genomic replicative intermediate and that the 3'-end of the (+) sense strand is polyadenylated. Equivalent sequences identified in clones generated using cDNA primed similarly from either total or poly(A)-selected RNA also supported previous data that the (+) sense genomic and sgRNAs are 3'-polyadenylated. It should be noted, however, that 5 clones derived from a phage library produced using poly(A)+ RNA from the whole head of a GAV-infected *P. monodon* all possessed one to two point mutations (i.e. in lower case 5'..TGATA(T/g)GA(A/c)(A/g)A_n-3') within -1 to +3 nt of the poly(A) junction shown in Fig. 2. However, these

clones were generated using a non-anchored oligo-dT primer and significance of the point mutations is not yet clear. Each of the five clones also possessed upstream point mutations (5 to 8/306 nt) and two possessed single point deletions (C⁵⁵³⁷ and A⁵⁵⁴⁵) ~50 nt from the poly(A) junction. While RNA from *P. monodon* experimentally-infected with the prototype GAV isolate [40] was used to derive the sequence shown in Fig. 2, the more variable sequences occurred in clones derived from a prawn naturally infected with GAV. Thus, it is likely that the point mutations/deletions are attributable to RNA quasispecies and/or virus strain differences. The genomic (+) RNA 3'-untranslated regions of coronaviruses and arteriviruses contain promoter sequences for transcription of negative-sense genomic and sgRNAs [13, 17, 23, 24, 37]. Cell lines in which GAV will replicate and/or reverse genetics systems will be required to gain an understanding of the 3'-terminal sequences that are important to the genomic replication process in this crustacean nidovirus.

The data reported here demonstrates that the (+) ssRNA genome of GAV is 3'-polyadenylated similarly to genomes of all viruses of the *Nidovirales*. However, due to the apparent absence of discrete membrane (M) and nucleocapsid (N) protein genes downstream of the GAV ORF3 glycoprotein gene, independent approaches were employed to be confident that we had identified the true genomic 3'-end. Completion of the genome sequence of GAV has revealed a gene organisation that is unique among the nidoviruses currently described. Significantly, GAV contains no discrete gene encoding an integral M protein and the gene (ORF2) most likely to encode the viral N protein resides upstream of a gene (ORF3) predicted to encode a large complex glycoprotein(s). Only two subgenomic (sg)RNAs are transcribed in abundance in GAV-infected cells [10] compared to five to nine sgRNAs in vertebrate nidoviruses. Moreover, similarly to toroviruses [36], there is no evidence of the complex leader-primed discontinuous transcription mechanism employed by coronaviruses [32] and arteriviruses [42]. These characteristics, taken together with the fact that the 26235 nt GAV genome comprises relatively few genes, suggests that this crustacean nidovirus is evolutionary more primitive than its vertebrate counterparts. In this regard, it may be relevant to note that the natural penaeid host of GAV has changed little from ancestral crustaceans that date back at least to the Cambrian period >500 million years ago [33].

We have previously suggested that GAV, and the closely related YHV, should be considered as members of a new genus, *Okavirus*, within the order *Nidovirales* [9]. The number and organisation of genes, the characteristic rod-shaped virion morphology [39, 40] and the distinct transcription process [10] supports the placement of the genus *Okavirus* in a family distinct from the *Coronaviridae* and *Arteriviridae*, for which the name *Roniviridae* (sigla, rod-shaped nidovirus) is proposed.

Acknowledgments

We kindly thank Dr. Sigrid Lehnert for providing the prawn cDNA phage library and Dr. Ross Tellam for helpful discussions on glycoprotein processing.

References

1. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402
2. Binns MM, Bournsnel ME, Cavanagh D, Pappin DJ, Brown TD (1985) Cloning and sequencing of the gene encoding the spike protein of the coronavirus IBV. *J Gen Virol* 66: 719–726
3. Boon JA, Snijder EJ, Locker JK, Horzinek MC, Rottier PJ (1991) Another triple-spanning envelope protein among intracellularly budding RNA viruses: the torovirus E protein. *Virology* 182: 655–663
4. Boonyaratpalin S, Supamattaya K, Kasornchandra J, Direkbusaracom S, Aekpanithanpong U, Chantanachookin C (1993) Non-occluded baculo-like virus, the causative agent of yellow-head disease in the black tiger shrimp (*Penaeus monodon*). *Fish Path* 28: 103–109
5. Cavanagh D (1995) The coronavirus surface glycoprotein. In: Siddell SG (ed) *The Coronaviridae*. Plenum Press, New York, pp. 73–113
6. Chantanachookin C, Boonyaratpalin S, Kasornchandra J, Sataporn D, Ekpanithanpong U, Supamataya K, Sriurairatana S, Flegel TW (1993) Histology and ultrastructure reveal a new granulosis-like virus in *Penaeus monodon* affected by yellow-head disease. *Dis Aquat Org* 17: 145–157
7. Cowley JA, Dimmock CM, Wongteerasupaya C, Boonsaeng V, Panyim S, Walker PJ (1999) Yellow head virus from Thailand and gill-associated virus from Australia are closely related but distinct prawn viruses. *Dis Aquat Org* 36: 153–157
8. Cowley JA, Dimmock CM, Spann KM, Walker PJ (2000a) Detection of Australian gill-associated virus (GAV) and lymphoid organ virus (LOV) of *Penaeus monodon* by RT-nested PCR. *Dis Aquat Org* 39: 159–167
9. Cowley JA, Dimmock CM, Spann KM, Walker PJ (2000b) Gill-associated virus of *Penaeus monodon* prawns: an invertebrate virus with ORF1a and ORF1b genes related to arteri- and coronaviruses. *J Gen Virol* 81: 1473–1484
10. Cowley JA, Dimmock CM, Walker PJ (2002) Gill-associated nidovirus of *Penaeus monodon* prawns transcribes 3'-coterminal subgenomic mRNAs that do not possess 5'-leader sequences. *J Gen Virol* 83: 927–935
11. De Groot RJ, Luytjes W, Horzinek MC, van der Zeijst BAM, Spaan WJM, Lenstra JA (1987a) Evidence for coiled-coil structure in the spike proteins of coronaviruses. *J Mol Biol* 196: 963–966
12. De Groot RJ, Maduro J, Lenstra JA, Horzinek MC, van der Zeijst BAM, Spaan WJM, (1987b) cDNA cloning and sequence analysis of the gene encoding the peplomer protein of feline infectious peritonitis virus. *J Gen Virol* 68: 2639–2646
13. De Vries AAF, Horzinek MC, Rottier PJM, de Groot RJ (1997) The genome organization of the *Nidovirales*: similarities and differences between arteri-, toro-, and coronaviruses. *Seminars Virol* 8: 33–47
14. Duckmanton LM, Tellier R, Liu P, Petric M (1998) Bovine torovirus: Sequencing of the structural genes and expression of the nucleocapsid protein of Breda virus. *Virus Res* 58: 83–96
15. Froussard P (1992) A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. *Nucleic Acids Res* 20: 2900
16. Garoff H, Frischauf AM, Simons K, Lehrach H, Delius H (1980) Nucleotide sequence of cDNA coding for Semliki Forest virus membrane glycoproteins. *Nature* 288: 236–241
17. Godeny EK, Chen L, Kumar SN, Methven SL, Koonin EV, Brinton MA (1993) Complete genomic sequence and phylogenetic analysis of the lactate dehydrogenase elevating virus (LDV). *Virology* 194: 585–596

18. Hansen JE, Lund O, Tolstrup N, Gooley AA, Williams KL, Brunak S (1998) NetOglyc: Prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconjugate J* 15: 115–130
19. Hashimoto K, Erdel S, Keränen S, Saraste J, Kääriäinen L (1981) Evidence for a separate signal sequence for the carboxy-terminal envelope glycoprotein E1 of Semliki Forest virus. *J Virol* 38: 34–40
20. Hofmann K, Stoffel W (1993) TMbase – A database of membrane spanning protein segments. *Biol Chem Hoppe-Seyler* 374: 166
21. Ihara T, Smith J, Dalrymple JM, Bishop DHL (1985) Complete sequence of the glycoproteins and M RNA of Punta Toro phlebovirus compared to those of Rift Valley fever virus. *Virology* 144: 246–259
22. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157: 105–132
23. Lai MMC (1990) Coronavirus: organization, replication and expression of genome. *Ann Rev Microbiol* 44: 303–333
24. Lai MMC, Cavanagh D (1997) The molecular biology of coronaviruses. *Adv Virus Res* 48: 1–100
25. Limsuwan C (1991) Handbook for cultivation of black tiger prawns. Tansetakit Co. Ltd, Bangkok (in Thai)
26. Lupas A, Van Dyke M, Stock J (1991) Predicting coiled coils from protein sequences. *Science* 252: 1162–1164
27. Nadala ECB, Tapay LM, Loh PC (1997) Yellow-head virus: A rhabdovirus-like pathogen of penaeid shrimp. *Dis Aquat Org* 31: 141–146
28. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engng* 10: 1–6
29. Rice CM, Strauss JH (1981) Nucleotide sequence of the 26S mRNA of Sindbis virus and deduced sequence of the encoded virus structural protein. *Proc Natl Acad Sci USA* 78: 2062–2066
30. Rönnholm R, Pettersson RF (1987) Complete nucleotide sequence of the mRNA segment of Uukuniemi virus encoding the membrane glycoproteins G1 and G2. *Virology* 160: 191–202
31. Sambrook J, Fritsch EF, Maniatis T (1987) *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA
32. Sawicki SG, Sawicki DL (1998) A new model for coronavirus transcription. *Adv Exp Med Biol* 440: 215–219
33. Siveter DJ, Williams M, Waloszek D (2001) A phosphatocopid crustacean with appendages from the Lower Cambrian. *Science* 293: 479–481
34. Snijder EJ, Den Boon JA, Spaan WJM, Verjans GMGM, Horzinek M (1989) Identification and primary structure of the gene encoding the Berne virus nucleocapsid protein. *J Gen Virol* 70: 3363–3370
35. Snijder EJ, Den Boon JA, Spaan WJM, Weiss M, Horzinek MC (1990) Primary structure and post-translational processing of the Berne virus peplomer protein. *Virology* 178: 355–363
36. Snijder EJ, Horzinek MC, Spaan WJ (1990) A 3'-coterminal nested set of independently transcribed mRNAs is generated during Berne virus replication. *J Virol*: 331–338
37. Snijder EJ, Meulenberg JJM (1998) The molecular biology of arteriviruses. *J Gen Virol* 79: 961–979
38. Sonnhammer ELL, von Heijne G, Krogh A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. In: Glasgow J, Littlejohn T, Major F,

- Lathrop R, Sankoff D, Sensen C (eds) Proc Sixth International Conference on Intelligent Systems for Molecular Biology. AAAI Press, California, pp 175–182
39. Spann KM, Vickers JE, Lester RJG (1995) Lymphoid organ virus of *Penaeus monodon* from Australia. *Dis Aquat Org* 23: 127–134
 40. Spann KM, Cowley JA, Walker PJ, Lester RJG (1997) Gill-associated virus (GAV), a yellow head-like virus from *Penaeus monodon* cultured in Australia. *Dis Aquat Org* 31: 169–179
 41. Tang KF-J, Lightner DV (1998) A yellow head virus probe: application to *in situ* hybridization and determination of its nucleotide sequence. *Dis Aquat Org* 35: 165–173
 42. van Marle G, Dobbe JC, Gultyaev AP, Luytjes W, Spaan WJM, Snijder EJ (1999) Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc Natl Acad Sci USA* 96: 12056–12061.
 43. Walker PJ, Cowley JA, Spann KM, Hodgson RAJ, Hall MR, Withyachumnarnkul B (2001) Yellow head complex viruses: transmission cycles and topographical distribution in the Asia-Pacific region. In: Browdy CL, Jory DE (eds) *The New Wave: Proc Special Session on Sustainable Shrimp Culture, Aquaculture 2001*. The World Aquaculture Society, Baton Rouge, pp 227–237
 44. Wang Y-C, Chang P-S (2000) Yellow head virus infection in the giant tiger prawn *Penaeus monodon* cultured in Taiwan. *Fish Pathol* 35: 1–10
 45. Wongteerasupaya C, Sriurairatana S, Vickers JE, Akrajamorn A, Boonsaeng V, Panyim S, Tassanakajon A, Withyachumnarnjul B, Flegel TW (1995) Yellow-head virus of *Penaeus monodon* is an RNA virus. *Dis Aquat Org* 22: 45–50

Author's address: Jeff Cowley, CSIRO Livestock Industries, 120 Meiers Road, Indooroopilly, 4068, Australia; e-mail: jeff.cowley@csiro.au