Letter to the Editor

DNA Barcoding Can Successfully Identify *Penaeus monodon*, Associate Life Cycle Stages, and Generate Hypotheses of Unrecognised Diversity
(Pengekodan Bar DNA Boleh Digunakan untuk Mengenal Pasti *Penaeus monodon*, Perkaitan Peringkat Kitaran Hayat dan Menjana Hipotesis daripada Kepelbagaian yang Tidak Dikenali)

JOHN JAMES WILSON* & KONG WAH SING

DNA barcoding is the technique of sequencing a short fragment of cytochrome c oxidase subunit I (COI) mtDNA – a 'DNA barcode' - from a taxonomically unknown organism. This DNA barcode is then matched against a reference library of barcodes of known species origin to make a species identification. Since 2004 an international consortium has been promoting DNA barcoding as a global standard for taxonomic identifications (Savolainen et al. 2005). While we applaud Kumar et al. (2012) for bringing DNA barcoding to the attention of researchers in Malaysia, we have concerns about their understanding of the theory and methods of DNA barcoding which we discuss below.

One major advantage of DNA barcoding over traditional species identification methods is the ability to correctly identify organisms from any life cycle stage to species. DNA barcoding has this capacity because intraspecific variation in DNA barcodes is generally much lower than interspecific variation between DNA barcodes (Savolainen et al. 2005), and crucially, because an organism's DNA barcode does not change as the organism progresses from one life stage to the next. The integrity of DNA sequences through replication and organism growth is fundamental to their role as the "blueprint of life" (Hill 2006). Kumar et al. (2012) sequenced DNA barcodes from giant tiger prawns and report the pattern typically seen in DNA barcoding studies: "The average pair-wise distance within the sequences of different stages of *P. monodon* was two times lesser than the pair-wise distance of the out-group used [*sic*]". Given the reported success of DNA barcoding at effectively identifying all individuals to species, the title of the article seems incongruous; what or where is the "Limit"? Kumar et al. (2012)'s conclusion points to the source of this incongruence: "However, it (*COI*) failed to segregate the different developmental stages to its corresponding life history stages [*sic*]". If Kumar et al. (2012) intended to use a molecular technique to segregate individuals of *P. monodon* into different life stages, as implied in their objectives and conclusion, they took a fundamentally wrong approach; DNA characters are independent of life stage. We suggest instead, they focus on RNA and differential gene expression (e.g. Pérez-Porro et al. 2013). We can state confidently that increasing the sample size will not "reveal the efficiency of COI gene in pinpointing various life history stages of any organism" as Kumar et al. (2012) postulated.

It is curious that Kumar et al. (2012) did not comment on the intraspecific distances observed within their *P. monodon* DNA barcodes (Figure 1(a)). We downloaded the barcodes from GenBank (GQ461913-GQ461918) and discovered the distances were an artefact of poor sequence editing and alignment. After removal of questionable nucleotides at the sequence edges and careful alignment (see Wilson 2012 for details on this procedure), Kimura 2-parameter (K2P) distances between the most distant barcodes were 1.3% (Figure 1(b)). Besides, the values from Kumar et al. (2012) are nonsensical as the "outgroup" sequences are a different fragment of COI and cannot be aligned to the "ingroup" (see alignment file at http://www.ukm.my/jsm/english_journals/vol42num12_2013/pg 1827-1829.html.).

'Blasting' the *P. monodon* barcodes against BOLD (boldsystems.org) revealed two distinct clusters (Figure 1(c)). The second cluster contained barcodes from India and Africa which matched closely with all the barcodes of Kumar et al. (2012). Dendograms of COI sequences (neighbour-joining trees are dendograms not a "phylograms" or "phylogenetic trees") consistently showed two clusters within *P. monodon*, separated by K2P distance of >7% (Figure 1(b)-1(c)). Studies of COI variation in crustaceans have considered K2P 'intraspecific' distance of >1.3% (Lefebure et al. 2006) to be suggestive of unrecognised species. Lavery et al. (2004) previously suggested the presence of cryptic species within *P. monodon* based on mtDNA.

We suggest that rather than focussing on a misconceived application of DNA barcoding, a more fruitful line of investigation may be sequencing (with careful editing and alignment) of nuclear regions to corroborate the existence of unrecognised species within *P. monodon*.
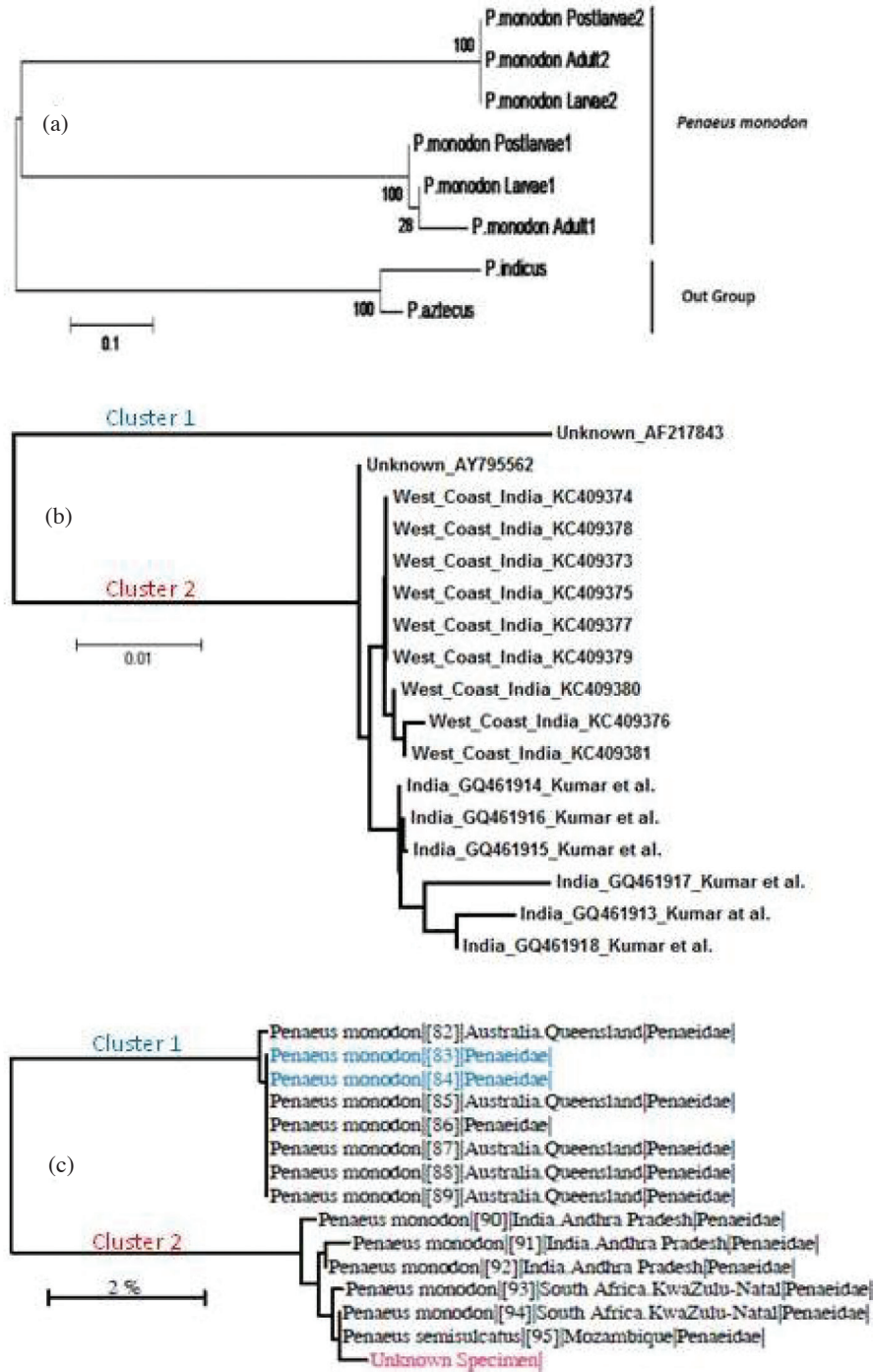
FIGURE 1. Neighbour-joining (NJ) trees showing Kimura 2-parameter distances between *Penaeus monodon* DNA barcodes. A) Figure 3 from Kumar et al. (2012). B) NJ showing two distinct clusters of *P. monodon* barcodes from GenBank. C) NJ from BOLD when 'blasting' sequence GQ461918. The name *semisulcatus* has been used incorrectly by some authors for *Penaeus monodon* explaining its appearance here. BOLD analysis groups AF217843 with Cluster 1

REFERENCES

Hill, E. 2006. Righting the wrongs: Structural insights into replicating damaged DNA. *PLoS Biology* 4(1): e32. doi:10.1371/journal.pbio.0040032.

Kumar, C.P., Akbar John, B., Ajmal Khan, S., Lyla, P.S. & Jalal, K.C.A. 2012. Limit of DNA barcode in delineating *Penaeus monodon* and in its developing stages. *Sains Malaysiana* 41(12): 1527-1533.

Lavery, S., Chan, T.Y., Tam, Y.K. & Chu, K.H. 2004. Phylogentic relationship and evolutionary history of genus *Penaeus s.l.* derived from mitochondrial DNA. *Molecular Phylogenetics and Evolution* 31: 39-49.

Lefébure, T., Douady, C.J., Gouy, M. & Gibert, J. 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: Proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution* 40: 435-447.

Pérez-Porro, A.R., Navarro-Gómez, D., Uriz, M.J. & Giribet, G. 2013. A NGS approach to the encrusting Mediterranean sponge *Crella elegans* (Porifera, Demospongiae, Poecilosclerida): transcriptome sequencing, characterization and overview of the gene expression along three life cycle stages. *Molecular Ecology Resources*. doi: 10.1111/1755-0998.12085.

Savolainen, V., Cowan, R.S., Vogler, A.P., Roderick, G.K. & Lane, R. 2005. Towards writing the encyclopaedia of life: An introduction to DNA barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360: 1805-1811.

Wilson, J.J. 2012. DNA barcodes for insects. *Methods in Molecular Biology* 858(2): 17-46.

John James Wilson* & Kong Wah Sing
Museum of Zoology
Faculty of Science
University of Malaya
50603 Kuala Lumpur
Malaysia

John James Wilson*
Institute of Biological Sciences
Faculty of Science
University of Malaya
50603 Kuala Lumpur
Malaysia

*Email: johnwilson@um.edu.my