# Re-evaluation of nuclear mitochondrial pseudogenes (NUMTs) and heteroplasmy in the Japanese spiny lobster Panulirus japonicus

# Re-evaluation of nuclear mitochondrial pseudogenes (NUMTs) and heteroplasmy in the Japanese spiny lobster *Panulirus japonicus*

Seinen Chow[1,2], Motoshige Yasuike[3], Takashi Yanagimoto[3]

[1]Fisheries Technology Institute, Japan Fisheries Research and Education Agency, Fukuura 2-12-4, Yokohama, Kanagawa 236-8648, Japan, [2]Aquos Institute, Motohachioji 3-2153-79, Hachioji, Tokyo 193-0826, Japan, [3]Fisheries Resources Institute, Japan Fisheries Research and Education Agency, Fukuura 2-12-4, Yokohama, Kanagawa 236-8648, Japan.

Abstract

The Japanese spiny lobster *Panulirus japonicus* has been reported to harbor a numbers of nuclear mitochondrial pseudogenes (NUMTs) and heteroplasmy. However, distinguishing phylogenetically young NUMTs, heteroplasmy, and PCR-cloning artefacts may be challenging. In addition, greater degradation for mtDNA than nuclear DNA in elderly tissue specimens may promote amplification of NUMTs. In this study, we performed clone library-based nucleotide sequence analysis of the partial mtDNA COI gene using genomic DNA and cDNA obtained from fresh tissues of the Japanese spiny lobster and genomic DNA obtained from three crustacean and three fish species. Minor nucleotide substitutions between clones in an individual were ubiquitously observed in all species examined including the lobster cDNA, suggesting that most of these were artefacts. Rarely, a few clones were most likely to have originated from heteroplasmic copies, as they had skewed nucleotide substitutions at the third codon. The Japanese spiny lobster is more likely than others to detect NUMTs, while the detection of NUMTs may be somewhat suppressed using genomic DNA obtained from fresh tissue.

Keywords: Japanese spiny lobster, NUMTs, mitochondrial DNA, heteroplasmy, sample degradation, cDNA

## Introduction

The number of reports on nuclear mitochondrial pseudogenes (NUMTs) and mitochondrial heteroplasmy in eukaryotes is drastically increasing (Lopez et al. 1994; Sorenson and Quinn 1998; Bensasson et al. 2001; Frey and Frey 2004; Kmiec et al. 2006; Pamilo et al. 2007; Song et al. 2008; Hazkani-Covo et al. 2010; Rodríguez-Pena et al. 2020; Tan et al. 2020; Wei et al. 2022). Co-amplification of these "noises" with genuin mtDNA can not only complicate the interpretation of mtDNA data, but also have a significant negative impact on the quality of the chromatogram obtained by direct nucleotide sequencing (Song et al. 2008; Buhay 2009). Chow et al. (2021) performed clone library-based nucleotide sequence analysis for PCR amplified mtDNA fragments of the Japanese spiny lobster *Panulirus japonicus* and found that approximately 20 % of the COI clones examined were NUMTs. They also claimed that some variant copies were heteroplasmy, but the criteria for distinguishing phylogenetically young NUMTs, heteroplasmic copies, and PCR-cloning artefacts were ambiguous. Since Chow et al. (2021) used relatively old samples (leg muscle preserved in ethanol for 24 years), amplification of NUMTs may be enhanced in older tissues with degraded mtDNA. DNA degradation depends on temperature and time of storage and tissue type, but in general the half-life of nuclear DNA is longer than that of mtDNA (Higgins et al. 2015). In addition, *Taq* polymerase errors may become more obvious when the amplified products are cloned and sequenced (Pääbo and Wilson 1988), strongly suggesting that the variant copies with minor nucleotide substitutions from the genuine haplotype may be artefacts.

This study aims to re-evaluate the presence and distribution of NUMTs and heteroplasmy in the Japanese spiny lobster *Panulirus japonicus*. Here, we report the results of clone library-based nucleotide sequence analysis on partial mtDNA COI gene using genomic DNA and cDNA obtained from fresh tissues of the Japanese spiny lobster and genomic DNA obtained from other crustacean and fish species.

## Materials and Methods

Four crustacean and three fish species used in the present study are presented in Table 1. Genomic DNA was extracted from the muscle or gonad using a DNA extraction kit (QuickGene DNA tissue kit, DT-S, KURABO). Nucleotide sequences of family or genus specific primers to amplify partial mitochondrial DNA COI regions (c. a. 900 bp) of these species are presented in Table 2. PCR amplification was performed in 12 μL reaction mixture containing 1 μL of template DNA (1–10 ng/μL), 1.2 μL of 10×reaction bufer, 1.2 μL of dNTP (2.5 mM each), 0.7 μL of each primer (10 μM), 0.3 μL of EX Taq HS polymerase (5 units) (Takara Bio, Inc.), and 7.6 μL of distilled water. The reaction mixtures were preheated at 94 °C for 5 min, followed by 35 amplification cycles (denaturation at 94 °C for 0.5 min, annealing at 58 °C for 0.5 min and extension at 72 °C for 1 min), with a final extension at 72 °C for 7 min. The PCR products were treated with ExoSAP-IT (GE Healthcare) to remove PCR primers, and direct nucleotide sequencing was performed

in all species except for the Japanese spiny lobster using PCR primers. The PCR products were cloned using a DynaExpress TA PCR Cloning Kit (BioDynamics Laboratory Inc.). Colony-direct PCR was performed using M13 primers, and clones having c. a. 1000 bp insert were selected using 1.5 % agarose gel electrophoresis. The PCR products were treated with ExoSAP-IT (GE Healthcare) to remove PCR primers and subjected to nucleotide sequence analysis using M13 primers. Pieces of leg muscle and gonad (testis) were dissected from two alive individuals of the Japanese spiny lobster (PJK10 and PJK11) and preserved in RNAlater (ThermoFisher Scientific, Waltham, MA). Total RNA was extracted using Direct-zol™ RNA MiniPrep (Zymo Research, Irvine, CA, USA). The quality and purity of the isolated total RNA were measured by a Nanodrop™ One spectrophotometer (Thermo Scientific, Waltham, MA). First-strand cDNA was synthesized from 500 ng total RNA with a PrimeScript II 1st Strand cDNA Synthesis kit (Takara Bio Inc., Shiga, Japan) according to the manufacturer's instruction. These cDNA samples were treated as described above for genomic DNA to obtain nucleotide sequences. Electropherograms obtained from clone libraries were carefully checked, and those with ambiguous peaks were discarded as they likely resulted from picking multiple colonies. Most frequent haplotype in an individual or haplotype obtained by direct nucleotide sequencing were determined to be genuine mtDNA of the individual. Translation to deduced amino acid sequence and nucleotide sequence alignment using the ClustalW algorithm was performed by GENETYX ver. 12 (GENETYX Co., Tokyo). Counting the number of nucleotide substitutions and calculation of Kimura's two-parameter distance (K2P) between variant haplotypes and the corresponding genuine haplotypes, model selection for nucleotide substitutions, and construction of maximum likelihood phylogenetic tree were performed using MEGA 6 (Tamura et al. 2013). MEGA files of aligned nucleotide sequences are available upon request to the corresponding author. Nucleotide sequences of the genuine haplotypes and NUMTs obtained in the present study were deposited to the International Nucleotide Sequence Database under accession numbers of LC781989–LC782008. To date, complete mtDNA sequences of 12 *Panulirus* species are available in the database, and COI sequences of these species were included for phylogenetic tree construction. However, one sequence (NC_052750) registered as *Panulirus penicillatus* was not used because the sequence of the COI region did not match many of the reported *P. penicillatus* COI sequences. Instead, partial COI sequences of *P. penicillatus* from western and eastern Pacific populations (Chow et al. 2011) were used.

Chi-square analysis was conducted using the monte carlo simulation of Roff and Bentzen (1989) with 1000 randomizations of the data to test the heterogeneity of haplotype distributions among samples, to which the Bonferroni correction was applied to the multiple comparisons. *G*–test of goodness-of-fit was used to test significant deviation from equal numbers of nucleotide differences (1:1:1) for each codon site.

**Results**

**Japanese spiny lobster**

Nucleotide sequences of the genuine haplotypes of the Japanese spiny lobster determined in the present study were 807 bp, and those of the variant haplotypes ranged from 799 to 809 bp. All 95 variant haplotypes were unique in nucleotide sequence. Fig. 1 is a phylogenetic tree constructed using genuine and variant haplotype sequences of the Japanese spiny lobster obtained in the present (PJK 10 and 11) and previous (PJK1 to 3) (Chow et al. 2021) studies and COI sequences of 12 *Panulirus* species derived from the database. Variant haplotypes that differed by up to two nucleotides from the corresponding genuine haplotypes were not involved for constructing phylogenetic tree, as these could be due to PCR-cloning error (Chow et al. 2021). Among the variant haplotypes, 25 haplotypes enclosed by the dotted rectangle were determined to be obvious NUMTs of the Japanese spiny lobster. Of these 25 haplotypes, numerous stop and nonsense codons were observed in 17 haplotypes and a number of nonsynonymous amino acid substitutions were observed in the remaining eight haplotypes. The other 23 variants were determined to be minor variant haplotypes, of which stop codons were observed in only two haplotypes. The K2P distances between the NUMTs and the corresponding genuine haplotypes ranged from 6.13 to 23.91 %, and those between the minor variant haplotypes and the corresponding genuine haplotypes ranged from 0.12 to 1.38 %. K2P distances between the genuine haplotypes of *P. japonicus* and the closest relatives (*P. longipes* and *P. cygnus*) ranged from 19.63 to 20.32 %. The clone numbers of genuine haplotypes, minor variant haplotypes categorized by the number of nucleotide difference from the genuine haplotypes, and NUMTs observed in the Japanese spiny lobster are summarized in Table 3. The number of nucleotide differences (including indels) between NUMTs and the corresponding genuine haplotypes ranged from 47 to 169, and those between the minor variant haplotypes and the corresponding genuine haplotypes ranged from one to 11. NUMTs were observed in four out of five individuals, whereas these were not detected in clones obtained from cDNA samples (PJK10 cDNA and PJK11 cDNA). There were significant differences in the frequencies of genuine haplotypes, minor variant haplotypes, and NUMTs among several samples (chi-square test, $P < 0.002$), to which presence of NUMTs may be responsible.

Since nucleotide substitutions occurred during PCR-cloning and those observed in NUMTs are expected to be independent of codon position, nucleotide substitutions between the variant and the corresponding genuine haplotypes were investigated in terms of the codon position. In the minor variant haplotypes, a significant bias in nucleotide substitution at the third codon was observed in three samples (PJK1, PJK3 and PJK11) (Fig. 2) (*G*-test, $P < 0.05$). All 10 nucleotide substitutions in a clone PJK1 (LC571547) were at the third codon. All three nucleotide substitutions in a clone PJK3 (OK429342), three out of four in a clone PJK3 (LC571559), and three out of five in a clone PJK3 (LC571571) were at the third codon. Five out of seven nucleotide substitutions in a clone PJ11M-C3-1-5, all 11 in a clone PJ11M-C3-2-12, and nine out of 11 in a clone PJ11-G-C4-2-12 were at the third codon. These clones, with the exception of PJK3, were outliers from the corresponding genuine haplotype and may have originated from heteroplasmic

copies, while no outlier clone was obtained from cDNA samples (Fig. 1). Surprisingly, a highly significant bias of nucleotide substitutions toward the third codon was observed in NUMTs of all four samples (Fig. 3) (*G*-test, $P < 0.001$).

**Other aquatic animals**

The number of clones having genuine haplotypes and variant haplotypes categorized by the number of nucleotide difference from the genuine haplotypes in six aquatic animals are summarized in Table 4. The number of nucleotide difference (including indels) between the variant haplotypes and the corresponding genuine haplotypes was at most five. Obvious NUMTs, as observed in the Japanese spiny lobster, were not detected in these samples. Individual data within a species were merged for statistical analysis, and no significant difference in the frequencies of the genuine and variant haplotypes was observed among species (chi-square test, $P > 0.009$). The number of clones examined in these species is not large enough to investigate codon position and nucleotide substitution. Although very preliminary and not significant (*G*-test, $P > 0.05$), four out of five nucleotide substitutions in a clone of the Japanese Kuruma shrimp (*Penaeus japonicus*) and all three in a clone of the Japanese mitten crab (*Eriocheir japonica*) were at the third codon.

**Discussion**

Given that Nuclear Mitochondrial DNA Segments (NUMTs) are not expressed, the occurrence of variant haplotypes in the cDNA of the Japanese spiny lobster is likely attributed to erroneous nucleotide incorporation during PCR-cloning or the presence of heteroplasmic copies. In previous studies on *Alpheus* spp., which employed similar methodologies as in this study, variant haplotypes differing by up to five nucleotides from their genuine counterparts were attributed to *Taq* polymerase errors (Williams et al. 2001; Williams and Knowlton 2001). Since the sequences analyzed in this study are longer compared to those in *Alpheus* spp., we must apply a slightly higher error threshold. Therefore, our previous threshold of up to two nucleotide differences (Chow et al. 2021) is insufficient. Most of the minor variant haplotypes in the Japanese spiny lobster, as well as all variant haplotypes in the six aquatic species examined in this study, fall within this error threshold. We have proposed that some of the minor variants in the Japanese spiny lobster might be derived from heteroplasmic copies, but this remains inconclusive. Nonetheless, even if these variants are of heteroplasmic origin, their frequency appears to be quite low and has minimal impact on direct nucleotide sequencing.

In a study by Williams et al. (2017), double peaks in sequence chromatograms of the blue crab *Callinectes sapidus* mitochondrial DNA were attributed to extensive heteroplasmy. However, the nucleotide sequence differences between many variant haplotypes and the dominant (genuine) haplotype in the blue crab fell within the error threshold, and some variant haplotypes even exhibited greater differences than those between different individuals, suggesting the presence of NUMTs in the blue crab. While heteroplasmy cannot be ruled out, the contribution of NUMTs to the double peaks in the sequence chromatograms of the blue crab, as demonstrated in the Japanese spiny lobster (Chow et al. 2021), is plausible.

Nucleotide sequence differences between NUMTs and their corresponding genuine haplotypes in the Japanese spiny lobster significantly exceed the range of erroneous nucleotide incorporation during PCR-cloning. The minimum K2P distance between NUMTs and genuine haplotypes was 6.13%, while the maximum K2P distance between minor variant haplotypes and their corresponding genuine haplotypes was 1.38%. Interestingly, no NUMTs have been identified that bridge this gap, despite the ongoing integration of mitochondrial DNA into the nuclear genome (Gaziev and Shaikhaev 2010; Wei et al. 2022). After integration events, NUMTs are typically amplified in tandem fashion, requiring time to form a large tandem array (Lopez et al. 1994). Therefore, the most likely explanation for this gap is that phylogenetically very young NUMTs are too few in number to be detected by the methods used in this study. Conversely, since NUMTs generally become progressively shorter in length and less similar to the original mtDNA sequences (Pamilo et al. 2007; Hlaing et al. 2009), older and shorter NUMTs may have been eliminated during the initial size screening or could have been poorly amplified due to priming site degeneration. According to Li et al. (1981), time since divergence ($t$) between and mtDNA and a nuclear pseudogene is estimated as $t = \delta/(\mu_1 + \mu_2)$, where $\delta$ is nucleotide sequence divergence between mtDNA and a nuclear pseudogene and $\mu$ is substitutions per site per year. Given $\mu_1 = 2.5 \times 10^{-8}$ for mtDNA (Hasegawa et al. 1985) and $\mu_2 = 4.7 \times 10^{-9}$ for nuclear pseudogene (Li et al. 1981), the time elapsed since integration is estimated to be 2 to 8 Myr. The maximum estimate is close to the divergence between *Panulirus japonicus* and its closest relatives (*P. longipes* and *P. cygnus*) in mid-Miocene (George 2006), suggesting that the oldest NUMTs observed in the present study were integrated during speciation events of these species. A highly significant bias of nucleotide substitutions toward the third codon was observed in all NUMTs regardless the age, although mitochondrial genes should cease to function once integrated into the nuclear genome. Explanations for this result are rather difficult and await further study.

Although the Japanese spiny lobster is more likely to detect NUMTs than the other aquatic animals examined in the present study, it remains unclear whether this is due to the choice of primers or the diverse nature of NUMTs among taxa. It is worth noting that difficulties in obtaining high-quality chromatograms are more common in crustaceans, where NUMTs are suspected to play a role. Buhay (2009) presented crustacean "COI-like" sequences, many of which are suspected to originate from NUMTs. NUMTs have been identified in various crustaceans, including crayfish (Nguyen et al. 2002; Song et al. 2008; Tan et al.

2020), spiny lobster (Chow et al. 2021; this study), snapping shrimp (Williams and Knowlton 2001; Williams et al. 2001), vent crabs (Kim et al. 2013), stone crabs (Schneider-Broussard and Neigel 1997), kelp crabs (unpublished data from this study), water fleas (Kowai et al. 2020), and copepods (Bucklin et al. 1999; Suyama et al. 2019, 2021). Moreover, as previously mentioned, the blue crabs studied by Williams et al. (2017) may also be included in this list.

If direct nucleotide sequencing proves successful for one taxon but not for another, a likely cause is significant co-amplification of NUMTs in the latter. To mitigate this issue, Song et al. (2008) and Calvignac et al. (2011) have proposed several strategies, including the use of mitochondria-rich tissues, mtDNA enrichment, taxon-specific primers, amplification of longer fragments, pre-PCR dilution, and RT-PCR. The results of this study indicate that using fresh tissue may provide some degree of NUMT avoidance, although this method is not entirely efficient. Techniques such as long PCR, mtDNA enrichment, and RT-PCR, while effective, are tedious, time-consuming and expensive, and co-amplification of NUMTs cannot be avoided using mtDNA rich tissue and mtDNA enrichment (Song et al. 2008). On the other hand, the use of taxon-specific primers designed to circumvent amplification of NUMTs is a convenient and cost-effective approach, that may be feasible, especially for the Japanese spiny lobster.

**Acknowledgements**

The authors are greatly indebted to Y. Hayashi, Fisheries Resources Institute, for her dedicated laboratory work. We also thank M. Sekino and H. Sakaji, Fisheries Resources Institute, for sharing their invaluable specimens of the Japanese eel *Anguilla japonica* and kuruma shrimp *Penaeus japonicus*. This work was supported by the Japan Fisheries Research and Education Agency.

**References**

Bensasson, D., Zhang, D. X., Hartl, D. L., Hewitt, G. M. (2001). Mitochondrial pseudogenes, evolution's misplaced witnesses. Trends Ecol. Evol. 16: 314–321.

Bucklin, A., Guarnieri, M., Hill, R. S., Bentley, A. M., Kaartvedt, S. (1999). Taxonomic and systematic assessment of planktonic copepods using mitochondrial COI sequence variation and competitive, species-specific PCR. Hydrobiologia 401: 239–254.

Buhay, J. E. (2009). "COI-like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. J. Crust. Biol. 29: 96–110.

Calvignac, S., Konecny, L., Malard, F., Douady, C. J. (2011). Preventing the pollution of mitochondrial datasets with nuclear mitochondrial paralogs (numts). Mitochondrion 11: 246–254.

Chow, S., Jeffs, A., Miyake, Y., Konishi, K., Okazaki, M., Suzuki, N., Abdullah, M. F., Imai, H., Wakabayashi, T., Sakai, M. (2011). Genetic isolation between the western and eastern Pacific populations of pronghorn spiny lobster *Panulirus penicillatus*. PLOS ONE 6: e29280.

Chow, S., Yanagimoto, T., Takeyama, H. (2021). Detection of heteroplasmy and nuclear mitochondrial pseudogenes in the Japanese spiny lobster *Panulirus japonicus*. Sci. Rep. 11: 21780.

Gaziev, A. I., Shaikhaev, G. O. (2010). Nuclear mitochondrial pseudogenes. Mol. Biol. 44: 358–368.

George, R. W. (2006). Tethys Sea fragmentation and speciation of *Panulirus* spiny lobsters. Crustaceana 78: 1281–1309.

Hasegawa, M., Kishino, H., Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22: 160–174.

Hazkani-Covo, E., Zeller, R. M., Martin, W. (2010). Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. PLOS Genet. 6: e1000834.

Higgins, D., Rohrlach, A. B., Kaidonis, J., Townsend, G., Austin, J. J. (2015). Differential nuclear and mitochondrial DNA preservation in post-mortem teeth with implications for forensic and ancient DNA studies. PLOS ONE 10: e0126935.

Hlaing, T., Tun-Lin, W., Somboon, P., Socheat, D., Setha, T., Min, S., Chang, M. S., Walton, C. (2009). Mitochondrial pseudogenes in the nuclear genome of *Aedes aegypti* mosquitoes: implications for past and future population genetic studies. BMC Genetics 10: 11.

Kim, S-J., Lee, K. Y., Ju, S-J. (2013). Nuclear mitochondrial pseudogenes in *Austinograea alayseae* hydrothermal vent crabs (Crustacea: Bythograeidae): effects on DNA barcoding. Mol. Ecol. Resour. 13: 781–787.

Kmiec, B., Woloszynska, M., Janska, H. (2006). Heteroplasmy as a common state of mitochondrial genetic information in plants and animals. Curr. Genet. 50, 149–159.

Kowai, K., Tkaczyk, A., Pierzchala, M., Bownik, A., Slaska, B. (2020). Identification of mitochondrial DNA (NUMTs) in the nuclear genome of *Daphnia magna*. Int. J. Mol. Sci. 21: 8725.

Li, W. H., Gojobori, T., Nei, M. (1981). Pseudogenes as a paradigm of neutral evolution. Nature 292: 237–239.

Lopez, J. V., Yuhki, N., Masuda, R., Modi, W., O'Brien, S. J. (1994). Numt, a recent transfer and tandem amplifcation of mitochondrial DNA to the nuclear genome of the domestic cat. J. Mol. Evol. 39: 174–190.

Nguyen, T. T. T., Murphy, N. P., Austin, C. M. (2002). Amplifcation of multiple copies of mitochondrial Cytochrome b gene fragments in the Australian freshwater crayfsh, *Cherax destructor* Clark (Parastacidae: Decapoda). Anim. Genet. 33: 304–308.

Pääbo, S. and Wilson, A. (1988). Polymerase chain reaction reveals cloning artefacts. Nature 334: 387−388.

Pamilo, P., Viljakainen, L., Vihavainen, A. (2007). Exceptionally high density of NUMTs in the honeybee genome. Mol. Biol. Evol. 24: 1340–1346.

Rodríguez-Pena, E. et al. (2020). High incidence of heteroplasmy in the mtDNA of a natural population of the spider crab *Maja brachydactyla*. PLOS ONE 15: e0230243.

Roff, D. A., Bentzen, P. (1989). The statistical analysis of mitochondrial DNA polymorphisms: χ2 and the problem of small sample sizes. Mol. Biol. Evol. 6: 539–545.

Song, H., Buhay, J. E., Whiting, M. F., Crandall, K. A. (2008). Many species in one, DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc. Natl. Acad. Sci. USA 105: 13486–13491.

Sorenson, M. D., Quinn, T. W. (1998). NUMTs, a challenge for avian systematics and population biology. Auk 115: 214–221.

Suyama, S., Masuda, Y., Yanagimoto, T., Chow, S. (2019). Genetic and morphological variation in *Pennella* sp. (Copepoda: Siphonostomatoida) collected from Pacific saury, *Cololabis saira*. Mar. Biodivers. 49: 1233–1245.

Suyama, S., Yanagimoto, T., Nakai, K., Tamura, T., Shiozaki, K., Ohshimo, S., Chow, S. (2021). A taxonomic revision of *Pennella* Oken, 1815 based on morphology and genetics (Copepoda: Siphonostomatoida: Pennellidae). J. Crust. Biol. 41: 1–12.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. (2013). MEGA6, molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30: 2725–2729.

Tan, M. H., Gan, H. M., Lee, Y. P., Grandjean, F., Croft, L. J., Austin, C. M. (2020). A giant genome for a giant crayfsh (Cherax quadricarinatus) with insight into cox1 pseudogenes in decapod genomes. Front. Genet. 11: 201.

Wei, W., Schon, K. R., Elgar, G., Orioli, A., Tanguy, M., Giess, A., Tischkowitz, M., Caulfield, M. J., Chinnery, P. F. (2022). Nuclear-embedded mitochondrial DNA sequences in 66,083 human genomes. Nature 611: 105−114.

Williams, S. T., Knowlton, N. (2001). Mitochondrial pseudogenes are pervasive and often insidious in the snapping shrimp genus *Alpheus*. Mol. Biol. Evol. 18: 1484–1493.

Williams, S. T., Knowlton, N., Weigt, L. A., Jara, J. A. (2001). Evidence for three clades within the snapping shrimp genus *Alpheus* inferred from nuclear and mitochondrial gene sequence data. Mol. Phyl. Evol. 20: 375–389.

Williams, E. P., Feng, X., Place, A. R. (2017). Extensive heteroplasmy and evidence for fragmentation in the *Callinectes sapidus* mitochondrial genome. J. Shellfsh Res. 36: 263–272.

Figure legends

Fig. 1. Maximum likelihood phylogenetic tree showing relationships among the variant haplotypes of COI

from the Japanese spiny lobster (*Panulirus japonicus*) and COI sequences of 12 congeneric species derived from the database. T92+G was selected as the optimal substitution model. Bootstrap values of > 60 % (from 1000 replicates) are shown at each node. Genuine haplotypes of five Japanese spiny lobsters are shown in bold. Haplotypes obtained from cDNA are marked with dagger, and those with stop and/or nonsense codons are marked with asterisk. Haplotypes determined to be obvious NUMTs are surrounded by dotted squares. The number of nucleotide differences in variant haplotypes from the corresponding genuine haplotypes is shown in square brackets. GenBank accession numbers are shown in parenthesis.

Fig. 2. Percentage of nucleotide substitutions at first (shaded bar), second (closed bar), and third (open bar) codons of minor variant haplotypes against the corresponding genuine haplotypes in the Japanese spiny lobster (*Panulirus japonicus*). n: total number of nucleotide substitutions. Asterisk indicates significant deviation from equal number of nucleotide differences (1:1:1) for each codon site.

Fig. 3. Percentage of nucleotide substitutions at first (shaded bar), second (closed bar), and third (open bar) codons of NUMTs against the corresponding genuine haplotypes in the Japanese spiny lobster (*Panulirus japonicus*). n: total number of nucleotide substitutions. Asterisk indicates significant deviation from equal number of nucleotide differences (1:1:1) for each codon site.