Occurrence of mitochondrial CO1 pseudogenes in *Neocalanus plumchrus* (Crustacea: Copepoda): Hybridization indicated by recombined nuclear mitochondrial pseudogenes

Ryuji J. Machida*, Ya-Ying Lin

Biodiversity Research Center, Academia Sinica, Nankang, Taipei, Taiwan

* ryujimachida@gmail.com

Abstract

A portion of the mitochondrial cytochrome *c* oxidase I gene was sequenced using both genomic DNA and complement DNA from three planktonic copepod *Neocalanus* species (*N. cristatus*, *N. plumchrus*, and *N. flemingeri*). Small but critical sequence differences in CO1 were observed between gDNA and cDNA from *N. plumchrus*. Furthermore, careful observation revealed the presence of recombination between sequences in gDNA from *N. plumchrus*. Moreover, a chimera of the *N. cristatus* and *N. plumchrus* sequences was obtained from *N. plumchrus* gDNA. The observed phenomena can be best explained by the preferential amplification of the nuclear mitochondrial pseudogenes from gDNA of *N. plumchrus*. Two conclusions can be drawn from the observations. First, nuclear mitochondrial pseudogenes are pervasive in *N. plumchrus*. Second, a mating between a female *N. cristatus* and a male *N. plumchrus* produced viable offspring, which further backcrossed to a *N. plumchrus* individual. These observations not only demonstrate intriguing mating behavior in these species, but also emphasize the importance of careful interpretation of species marker sequences amplified from gDNA.

Introduction

A growing body of literature is documenting the presence of nuclear copies of mitochondrial sequences, known as nuclear mitochondrial pseudogenes (reviewed in Bensasson et al. 2001 [1]). Occurrences of the nuclear mitochondrial pseudogenes are observed in various taxonomic groups in Eukaryotes, and Crustaceans are no exception [2–7]. The paralogous nature of nuclear mitochondrial pseudogenes have profound implications for studies which use mitochondrial DNA sequence as a genetic marker, including studies in the fields of phylogeny, phylogeography, species identification, metagenetics and metabarcoding.

Neocalanus copepods are some of the most dominant pelagic copepods in the Pacific [8–10], and certain species of this genus undergo extensive ontogenetic vertical migrations,

residing and growing in the euphotic layer from winter to summer, and descending to the meso- and bathypelagic layers from summer to winter (or autumn) for maturation and spawning [8]. During the course of a phylogenetic study of *Neocalanus* copepods [11], we repeatedly obtained sequences of unknown identity, with many polymorphic sites (double peaks) observed in sequence electropherograms. By cloning those PCR products, we detected several haplotypes within single organisms.

In this study we compared the patterns of sequence variations between genomic DNA (gDNA) and complement DNA (cDNA). Genomic DNA is composed of not only mitochondrial but also nuclear chromosomal DNA, which directly extracted from individuals. In contrast, cDNA is reverse transcribed DNA, which synthesized from extracted RNA as template. Therefore, sequences obtained from gDNA include not only active gene sequences, but also non-active pseudogene sequences. In contrast, sequences obtained from cDNA analyses are only actively transcribed gene sequences. From the comparative analyses of gDNA and cDNA, we detected multiple nuclear mitochondria pseudogenes in *N. plumchrus* and evidence for a recombination event between these pseudogenes. One pseudogene was a chimera of sequences from *N. plumchrus* and *N. cristatus*, which demonstrated the production of viable hybrid offspring from a mating between the two species.

Materials and methods

Animal samples

Specimens of *Neocalanus* species (*N. cristatus*, *N. plumchrus*, and *N. flemingeri*) were collected off the Pacific coast of southeastern Hokkaido, Japan, using a Norpac net (North Pacific Standard Plankton Net). Specimens were either immediately preserved in RNAlater (Ambion Inc.) or brought back to laboratory alive. An additional sample was also collected off the Pacific coast of Tohoku, northern Japan. No specific permissions were required for these activities and work did not involve endangered or protected species.

RNA and DNA extraction

At least 3 individuals were examined from each species. Bodies of each individual were divided into two parts using sterilized scissors and served for gDNA and messenger RNA (mRNA) extraction. Although only gDNA and not mRNA was extracted from a *N. plumchrus* individual caught off the Pacific coast of Tohoku (Table 1), we included this organism in the analysis as it contains a putative chimera sequence (discussed later).

Total gDNA was extracted from a portion of each individual using a DNeasy (Qiagen Inc.) kit, in accordance with the manufacturer's protocol. Messenger RNA was selectively extracted from the other half of each individual by oligo-(dT) cellulose chromatography using the QuickPrep Micro mRNA Purification Kit (Amersham Biosciences Inc.).

Amplification from gDNA

Mitochondrial DNA (ca. 4 kbp) was amplified by long PCR with the H13842-12S (5′-TGT GCC AGC ASC TGC GGT TAK AC-3′) and L12168-16S (5′-CGT CGA TTT KAA CTC AAA TCA TGT-3′) primers (Fig 1 [11,12]). Thermal condition of the long PCR was as follows. Shuttle long PCR thermal profile: 10 seconds denaturation at 98°C, followed by annealing and extension combined at the same temperature (68°C) for 6 minutes. Temperature for the annealing and extension was decreased by– 0.5°C each cycle during the first 16 cycles from 68°C to 60°C, followed by an additional 21 cycles with fixed temperature (60°C). The reactions contain the following reagents: a total of 25 μl volume with 12.75 μl of sterile, distilled H_2O,

Table 1. Summary of mitochondrial cytochrome *c* oxidase I gene sequences from three *Neocalanus* species.

Neocalanus species	Individual number	gDNA			cDNA		
		Number of clones sequenced	Number of substitutions Non-synonymous / synonymous	Accession number (INSDC)	Number of clones sequenced	Number of substitutions Non-synonymous / synonymous	Accession number (INSDC)
N. cristatus	381	5	3 / 3	AB099151-AB099155	7	5 / 1	AB099175-AB099181
	382	8	7 / 8	AB099156-AB099163	5	0 / 0	AB099182-AB099186
	383	7	4 / 2	AB099164-AB099170	7	1 / 3	AB099187-AB099193
	384	4	8 / 27	AB099171-AB099174	7	2 / 1	AB099194-AB099200
N. flemingeri	387	8	8 / 28	AB099201-AB099208	6	0 / 1	AB099221-AB099226
	388	8	6 / 4	AB099209-AB099216	8	3 / 2	AB099227-AB099234
	389	4	1 / 6	AB099217-AB099220	4	1 / 0	AB099235-AB099238
N. plumchrus	407	6	28 / 61	AB099239-AB099244	6	6 / 0	AB099267-AB099272
	408	8	14 / 96	AB099245-AB099252	6	3 / 3	AB099273-AB099278
	409	8	29 / 149	AB099253-AB099260	7	6 / 1	AB099279-AB099285
	168	6	7 / 180	AB099261-AB099266			

doi:10.1371/journal.pone.0172710.t001

2.5 μl of 10 X buffer, 4 μl of dNTPs (2.5 mM each), 1.0 μl of each primer (5 μM), 0.25 μl of 1.25 units of LA Taq (Takara Bio Inc.), and 1.0 μl of template. The long PCR products were electro-phoresed on a 1.0% L 03 agarose gel (Takara Bio Inc.), which was subsequently stained with ethidium bromide for band characterization under ultraviolet transillumination.

Next, nested PCR using a 50-fold dilution of the long PCR products as a templates was per-formed to amplify a segment of the mitochondrial cytochrome *c* oxidase I (CO1) gene. L1384-COI (5′-GGT CAT GTA ATC ATA AAG ATA TTG G-3′) and H2088-COI (5′-GAT GTG KGA GAT TAG TCC GAA-3′) primers were used for the *N. plumchrus*, and L1384-COI (5′-GGT CAT GTA ATC ATA AAG ATA TTG G-3′) and H2061-COI (5′-GGA ATT AGA ATG TAT GCC TC-3′) primers were used for *N. cristatus* and *N. flemingeri* [11,12]. The expected length of the PCR products was 704 and 677 bp for *N. plumchrus* and *N. flemingeri*, respectively. The thermal condition of the PCR was as follows: 30 cycles of 5 sec-ond denaturation at 94°C, followed by 5 second annealing at 50°C, and 60 second extension at 72°C. The reaction contained the following reagents: a total of 15-μl reaction volume with 9.8 μl of sterile, distilled H$_2$O, 1.5 μl of 10 X buffer, 1.0 μl of dNTPs (2.5 mM each), 1.0 μl of each primer (5 μM), 0.1 μl of 0.25 units of EX *Taq* (Takara Bio Inc.), and 1.0 μl of template. The PCR products were electrophoresed on a 1% L 03 agarose gel (Takara Bio Inc.), which was then stained with ethidium bromide for band characterization under ultraviolet transillumination.

All PCR were performed using a Model 9700 thermal cycler (Applied Biosystems Inc.).
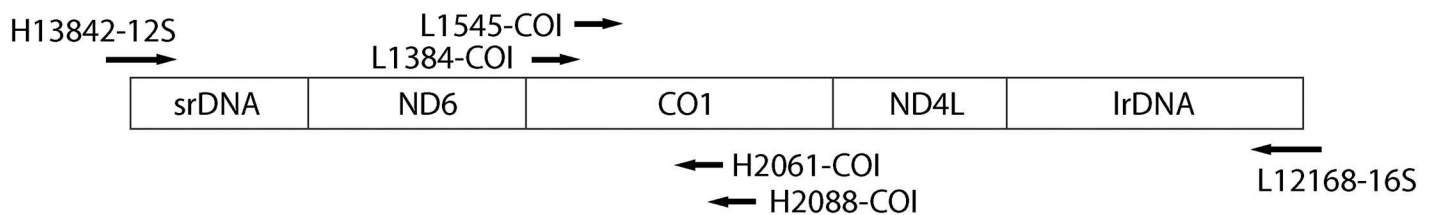

Fig 1. A portion of the mitochondrial genome amplified in the present study. Primers used for the amplification were indicated as arrows. Refer to the text for the primer sequences.

doi:10.1371/journal.pone.0172710.g001

Reverse transcription of the CO1 gene

First-strand synthesis of cDNA was performed using the 3´-Full Race Core Set (Takara Bio Inc.), in accordance with the manufacturer's instructions. A 2 μl aliquot of the resulting cDNA was mixed with 33.6 μl of sterile, distilled H$_2$O, 5.0 μl of 10 X buffer, 5.0 μl of dNTP (2.5 mM each), 2.0 μl of each primer (5 μM), and 0.4 μl of 0.25 units of EX *Taq* (Takara Bio Inc.), for a 50-μl total reaction volume. L1384-COI and 3sites Adaptor Primer (3´-Full Race Core Set; Takara Bio Inc.) pairs were used. Thermal cycle parameters were the same as stated above for nested PCR, with the sole exception of the extension time, which was increased to 120 seconds. Thirty cycles were performed. Since clear single bands were not observed in PCR products from *N. cristatus* and *N. plumchrus*, a second round of amplification was carried out in each case, using 1.0 μl of the previous PCR product as a template. L1545-COI (5′-GCT CAT GCW TTT GTC ATG ATT TTT TTT ATG G-3′ [11]) and 3sites Adapter Primer (3´-Full Race Core Set; Takara Bio Inc.) pairs were used in this step. The PCR procedure was the same as that of the first round of amplification.

Cloning

Prior to cloning, PCR amplicons from gDNA and cDNA were gel-purified on a 1.0% L 03 agarose gel (Takara Bio Inc.). The bands were excised and treated with the Wizard SV Gel and PCR Clean-Up System (Promega Inc.). Purified products were cloned into DH5 α cells using the pGEM-T Easy Vector System (Promega Inc.), following the manufacturer's instructions. Individual bacterial colonies were picked from the bacterial plate using sterilized toothpicks, and then resuspended in a PCR mix containing SP6 and T7 primers. Reaction concentrations and thermal cycle parameters were the same as those of the nested PCRs (described above), except for the volume, which was scaled down to 10 μl. PCR products with the correctly sized insert were subjected to enzymatic clean up using a Pre-Sequencing kit (USB Inc.).

Sequencing

Purified PCR products were directly subjected to Sanger sequencing (Applied Biosystems Inc.). Both primers were used for each sequencing reaction, and adenine stretches were confirmed in the case of cDNA analyses. All sequencing reactions were performed according to the manufacturer's instructions. Labeled fragments were analyzed on an Applied Biosystems 3100 DNA sequencer. Four to eight clones were sequenced for both gDNA and cDNA from each individual (Table 1).

Phylogenetic analysis

We estimated phylogenetic relationships of the obtained clone sequences using *N. tonsus* as the outgroup (INSDC: AB093146). First, the sequences were aligned using the MAFFT server with G-INS-1 [13]. Next, Maximum-likelihood (ML) analysis was performed using the RAxML 7.7.1 server [14] with the following options: -f a -m GTRCATI. The best-score ML tree was selected for examination in the following discussion.

Results

Most of the sequences from different species formed different clades, with three long branches (Fig 2). An exception was clone P7945, which was amplified from the gDNA of *Neocalanus plumchrus* but clustered in the *N. cristatus* clade (discussed below). For *N. flemingeri* and *N. cristatus*, there were very minor sequence differences between clones from gDNA and cDNA, which lead to a flat-top topology. Very small differences were also observed between clone
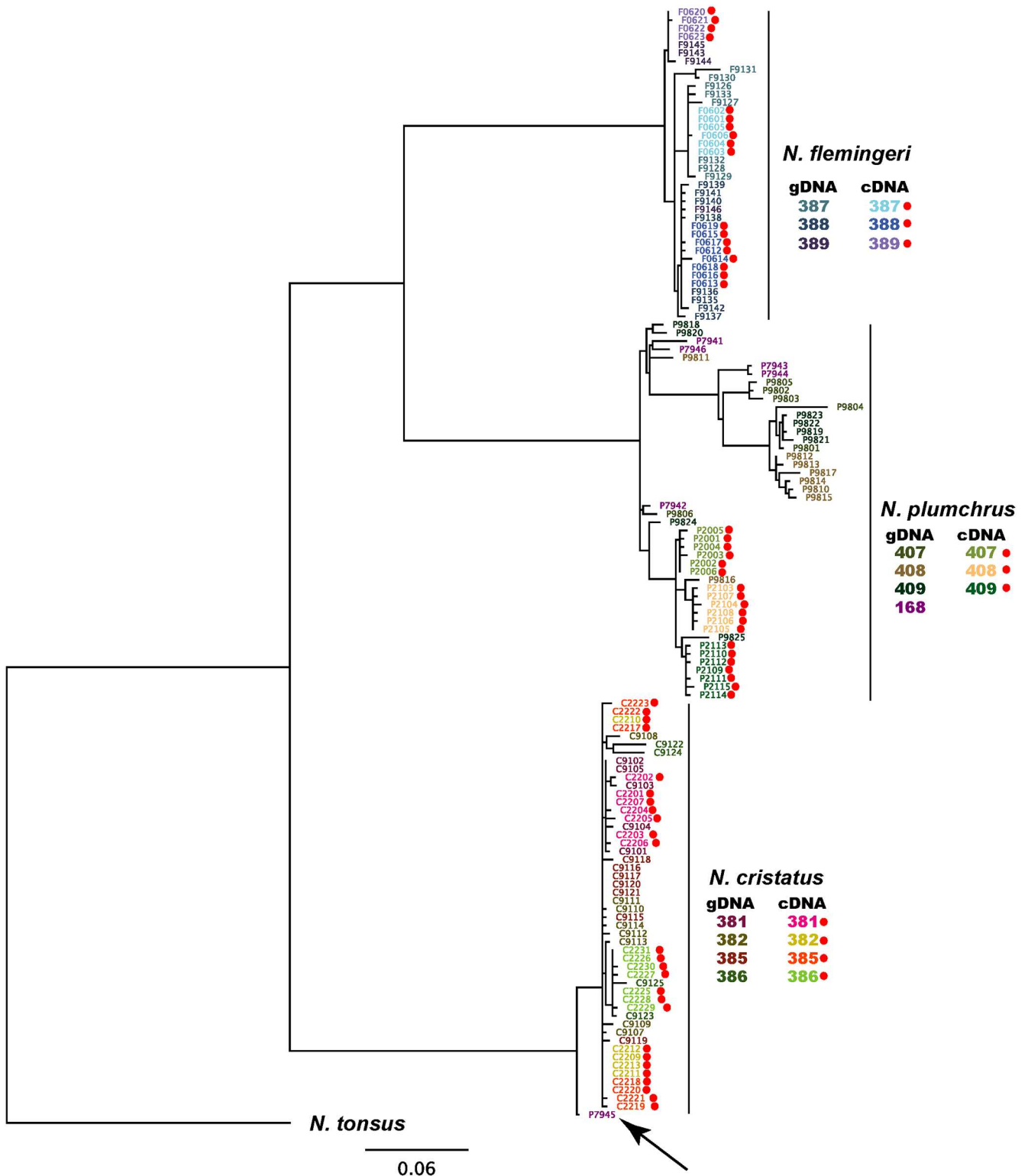
Fig 2. The best-score Maximum Likelihood tree of the mitochondrial cytochrome *c* oxidase I gene clone sequences obtained from *Neocalanus* copepods. Clone number P7945, which was amplified from the genomic DNA of *N. plumchrus* but clustered in the clade of *N. cristatus*, is indicated by the arrow. Each clone is shown in a different color to enable comparison between individuals and between genomic DNA and complement DNA. Red small circle indicates the clone, which was amplified from complement DNA.

doi:10.1371/journal.pone.0172710.g002

```
                          Variable sites
                 [0000000000000000011111111111111]
                 [00122446677889990022566778899]
                 [4502569474625247061473923173 69]


Colony number (Ind. num.)


                              N. plumchrus
P7942 (168)     CCCGTTACCTAATCTTGTTTG.CTACCTCT
P7943 (168)     CCCGTTACCTAATCTTGCTTGTCTACCCCT
P7945 (168)     CTCGTGTTGCCGATACTTTAACTATTTATC
                              N. cristatus
C9101 (381)     TTTTCGTTGCCGGTACTTCAACTATTTATC
C9102 (381)     TTTTCGTTGCCGGTACTTCAACTATTTATC
C9103 (381)     TTTTCGTTGCCGGTACTTCAACTATTTATC
```

Fig 3. Alignment of mitochondrial cytochrome *c* oxidase I gene sequences from *Neocalanus cristatus* and *N. plumchrus* individuals. Only sites with variations in more than two individuals were extracted. Numbers at the top of the figure refer to the 30 variable sites. Clone number P7945 is the sequence that was amplified from genomic DNA of *N. plumchrus*, but fell in the *N. cristatus* cluster (as shown in Fig 2). The grey box indicates the sequence region in which recombination is believed to have occurred. Underlined letters indicate the putative source segment of the recombination.

doi:10.1371/journal.pone.0172710.g003

sequences from cDNA of *N. plumchrus*. In contrast, the extent of sequence differences varied considerably between clones amplified from gDNA of *N. plumchrus*, with a considerable number of synonymous substitutions (Table 1 and Fig 2).

One sequence amplified from gDNA of *N. plumchrus*, which fell in the *N. cristatus* clade (Fig 2: P7945), was aligned with sequences obtained from *N. cristatus* and *N. plumchrus* (Fig 3). This alignment revealed that the first 25 base pairs of P7945 were identical to the sequence from *N. plumchrus*, but the sequence from position 46 to the 3' end (indicated by the grey box) was similar to that of *N. cristatus*.

We also aligned all of the sequences from both cDNA and gDNA of *N. plumchrus* (Fig 4). Three types of haplotype were obtained from cDNA, which correspond with each individual. In contrast, much higher sequence variations were observed for the sequences obtained from gDNA, which loosely fall into three groups. Careful comparisons of the aligned sequences confirmed that the three regions indicated by grey boxes (Fig 4) were recombinants; underlined letters indicated the putative source segment of the recombination.

Discussion

In all *Neocalanus* species, sequencing of cDNA clones generally reveals one haplotype for each individual, with putative PCR and cloning miscopies [15]. In contrast, various types of mitochondrial CO1 gene sequence were obtained from gDNA of *N. plumchrus* individuals, with a high number of synonymous substitutions (Table 1).

```
                                          Variable sites
                      [000000000011111111111111111111122222222223333333333333333444444444444555555]
                      [11233466780011112345667778888800012356681112334477888900124666888901233]
                      [39038458834735678468173892456813421910311451294727679647982146258292924009]
Consensus              TGTTTTTGCATGACCTTGTGAACACGCTCAGTCAGAGATTATGTTTTCCACTGTTTTTCAGAGATATGCCTA

Colony number (Ind. num.)

cDNA
───────────────────────────────────────────────────────────────────────────────────────
P2001 (407)           ................................C...CTG............T....C.C...G..C.......A
P2002 (407)           ................................C...CTG............T....C.C..-G..C.......A
P2003 (407)           C...............................C...CTG............T....C.C.-.G..C.......A
P2004 (407)           ........T.......................C...CTG............T....C.C...G..C.......A
P2005 (407)           ................................C...CTG............T....C.C...G..C.......A
P2006 (407)           ................................C...CTG............T....C.C...G..C.......A

P2103 (408)           .........................G......C...C.G......C.....T....CCC...GA.C.......
P2104 (408)           .........................G......C...C.G......C.....T....CCC...GA.C.......
P2105 (408)           .........................G......C...C.G......C.....T....CCC...GA.C.......
P2106 (408)           ...............A...G......C...C.G......C.....T....CCC...GA.C.......
P2107 (408)           ....................C..G........C...C.G......C.....T....CCC...GA.C.......
P2108 (408)           .........................G......C...C.G......C.....T....CCC...GA.C.......

P2109 (409)           ...............................C.G.C.G............T....C.C...G..C......G.
P2110 (409)           ...............................C.G.C.G............T....C.C...G..C......G.
P2111 (409)           ...............................C.G.C.G............T....C.C...G..C......G.
P2112 (409)           ...............................C.G.C.G............T....C.C...G..C......G.
P2113 (409)           ...............................C.G.C.G............T....C.C...G..C......G.
P2114 (409)           ...............................C.G.C.G............T....C.C...G..C......G.
P2115 (409)           C..............................C.G.C.G............T....C.C...G..C......G.

gDNA
───────────────────────────────────────────────────────────────────────────────────────
P7941 (168)           ..........C........TT....C.............................T.......CT...A
P7942 (168)           ....C.....................T....................................A
P9806 (407)           ....C...T...........T.....T.................................C.......A
P9824 (408)           ..............G................................................A
P9818 (409)           .A..C............T.....A...............C.......................A
P9820 (409)           .A..C............T.....A...............C...................CT....
P7946 (168)           ....C.....C.......TT....A..............................._____
P8911 (408)           ....C...........A.TT....A......................................A

P9816 (408)           ...................G.........C...C.G......C......T....CCC..--...G.TT...A
P9825 (408)           _____.........C.G.C.G..........T....C.CA......GC.T.A.

P7943 (168)           ..CC.G.A...A..TCC.CTT.T....C......A....CG....CCT.G..ACG.........C...CT...A
P7944 (168)           ..C..G.A...A..TCC.CTT.TG...C......A....CG....CCT.G..ACG.........C...CT...A
P9802 (407)           .....A.....A..T.C..TC.T.TAT.....A.A...CCG....CCT.G..ACG..............TT..A
P9805 (407)           .....A.....A..T.C..TC.T.TAT.....A.A...CCG....CCT.G..ACG..............TT..A
P9803 (407)           .....A.....A..T.C..TC.TGTAT.....A.A...CCG....CCT.G..ACG..............TT..A
P9801 (407)           .....GC..G.AGTT.C..TTGT.T....GC...A.C.CCGC...CCTTG.CACG........GC.T.A..
P9819 (409)           .....GC..G.AGTT.C..TTGT.T....GC...A.C.CCGC...CCTTG.CACG..A_____GC.T.A..
P9822 (409)           .....GC..G.AGTT.C..TTGT.T....GC...A.C.CCGC...CCTTG.CACG..A......GC.T.A..
P9823 (409)           .....GC..G.AGTT.C..TTGT.T....GC...A.C.CCGC...CCTTG.CACG...A......GC.T.A..
P9821 (409)           .....GC..G.AGTT.C..TTGT.T....GC...A.C.CCGC...CCTTG.CACG...A......GC.T.A..
P9810 (408)           .....GC.TG.AGTT.C..TTGT.T....G....A.C.CCGCA..CCTTG.CACG.........GC.T....
P9814 (408)           .....GC.TG.AGTT.C..TTGT.T....G....A.C.CCGCA..CCTTG.CACG.........GC.T...A
P9815 (408)           .....GC.TG.AGTT.C..TTGT.T....G....A.C.CCGCA..CCTTG.CACG.........GC.T....
P9812 (408)           .....GC.TG.AGTT.C..TTGT.T....G....A.C.CCGCA..CCTTG.CACG.........GC.T....
P9813 (408)           ...C.GC.TG.AGTT.C..TTGT.T....G....A.C.CCGCA..CCTTG.CACG.........GC.T....
P9817 (408)           .....GC.TG.AGTT.C..TCGT.T....G.C..A.C.CCGCA..CCTTG.CACG...............
P9804 (407)           ....................GT.T...GC...A.C.CCGC...CCTTG.CACG.........GC.T.A..
```

Fig 4. Alignment of all of the mitochondrial cytochrome *c* oxidase I gene sequences from both complement DNA and genomic DNA of *Neocalanus plumchrus*. Only the sites in which variations were observed in more than two individuals were extracted. Numbers at the top of the figure refer to the 72 variable sites. The grey box indicates the region in which recombination is believed to have occurred. Underlined letters indicate the putative source segment of the recombination.

doi:10.1371/journal.pone.0172710.g004

These results observed in the present studies are best explained by the presence of mito-chondrial pseudogenes in nuclear genomes because of following reasons. First, we have observed large number of recombined sequences. In general, recombination occurs in the nuclear genome, but not in the mitochondrial genome. Second, we observed only one sequence from cDNA but multiple sequences from gDNA. Differences in nuclear and mito-chondrial genetic codes mean that mitochondrial DNA within the nuclear genome is nonfunc-tional. Moreover, there is only a single promoter for transcription of either strand in the mitochondrial genome, and not individual promoters for each gene, as is observed in nuclear DNA [16]. For these two reasons, it is expected that nuclear mitochondrial pseudogenes will not be transcribed into mRNA [17].

One might suspect that the recombination took place during PCR. However, two lines of evidence indicate that it is very unlikely. First, in general, PCR recombination is considered as a rare event compared to other PCR errors [18], unless the starting template has been seriously damaged [19]. In the present study, four out of 28 cloned sequences exhibited recombination (Figs 3 and 4). Second, we have obtained chimera sequence of *N. plumchrus* and *N. cristatus* from *N. plumchrus* gDNA (Fig 3). If this sequence is created during PCR, template contamina-tion and PCR recombination, both of which are rare events, need to occur simultaneously. Furthermore, we did not observe non-recombined sequence of *N. cristatus*, which expected to be observed if template contamination took place. These lines of evidence indicated it is very unlikely that those recombined products were created during the PCR. In general, animal mitochondrial DNA sequences evolve faster than nuclear DNA sequences. As such, once mito-chondrial DNA inserts into the nuclear genome, it becomes a 'fossil' sequence [20]. Repeated transfer of the mitochondrial DNA sequence to the nuclear genome generates multiple haplo-types with a predominance of silent substitutions.

The genomes of calanoid copepods are very large [21], and there is a strong relationship between the occurrence of nuclear mitochondrial pseudogenes and nuclear genome size [22,23]. In addition, increasing amounts of evidence suggest a high rate of transfer of mito-chondrial genes to the nuclear genome [24–31], and pseudogene copies of mitochondrial genes have been observed in a wide range of organisms [32–35], including crustaceans [2–7].

The above lines of observation have important implications for the use of mitochondrial genetic markers in species identification (such as DNA barcoding) and estimation of commu-nity diversity (metagenetics). First, most of the sequences from gDNA of *N. plumchrus* did not have indels. Therefore, these sequences can be informatically translated into full proteins, which make it impossible to ascertain whether they are pseudogenes or functional protein cod-ing regions from their nucleotide sequences. Second, one organism contained a sequence from a different species (*N. cristatus* in this study). Third, we observed prevalent nuclear mitochon-drial pseudogene sequences and recombination between some pseudogenes. Several thought-ful discussions regarding avoidance of possible contamination of nuclear mitochondrial pseudogenes for analyses have been previously published [1,5,7]. Furthermore, more and more studies are using massively-parallel sequencers, which can detect rare variants. The results of the present study further demonstrate the importance of careful translation of data obtained from analyses of mitochondrial marker regions amplified from gDNA.

One clone sequence amplified from the gDNA of *N. plumchrus* was a chimera sequence of *N. plumchrus* and *N. cristatus* (clone num. P7945; Figs 2 and 3). Three genetic events need to have occurred to explain this phenomenon. First, hybridization has taken place between a female *N. cristatus* individual and a male *N. plumchrus* individual, followed by backcrossing to a *N. plumchrus* individual. Second, the mitochondrial sequence has been transferred to the nuclear genome. Third, recombination had occurred between the nuclear mitochondrial pseu-dogenes from *N. plumchrus* and *N. cristatus*. To our knowledge, while recombination of

mitochondrial pseudogenes has been previously reported in human and bird [28,36], our study is the first to detect hybridization based on recombination of mitochondrial pseudogenes. Although, all of the *Neocalanus* species analyzed in the present study constitute a single monophyletic group in the genus [11], body size of *N. cristatus* (Female prosome length: 7.74 ± 0.41 mm; [9]) is much longer than that of the other two species (*N. flemingeri*: 4.09 ± 0.32 mm, *N. plumchrus*: 4.06 ± 0.18 mm; [37,38]). Therefore, the present study has demonstrated hybridization of morphologically dissimilar species (*N. cristatus* and *N. plumchrus*). One explanation for this unlikely mating is the prolonged reproductive duration observed in *N. cristatus* [8,9], which overlaps with that of *N. plumchrus*. In contrast, *N. plumchrus* and *N. flemingeri* exhibit similar body lengths, but have limited mating durations, which do not overlap with each other [10]. To our knowledge, only one other study has demonstrated hybridization of copepods in a natural environment [39]. Together with the present study, these observations indicate that hybridization between species in which mating times overlap is more common than previously thought.

Acknowledgments

Author Contributions

Conceptualization: RM.

Funding acquisition: RM.

Investigation: RM YL.

Methodology: RM.

Resources: RM.

Writing – original draft: RM.

References

1. Bensasson D, Zhang DX, Hartl DL, Hewitt GM. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends Ecol Evol. 2001; 16: 314–321. PMID: 11369110

2. Schneider-Broussard R, Neigel JE. A large-subunit mitochondrial ribosomal DNA sequence translocated to the nuclear genome of two stone crabs (*Menippe*). Mol Biol Evol. 1997; 14: 156–165. PMID: 9029793

3. Williams ST, Knowlton N. Mitochondrial pseudogenes are pervasive and often insidious in the snapping shrimp genus *Alpheus*. Mol Biol Evol. 2001; 18: 1484–1493. PMID: 11470839

4. Song H, Buhay JE, Whiting MF, Crandall KA. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc Nat Acad Sci USA. 2008; 105: 13486–13491. doi: 10.1073/pnas.0803076105 PMID: 18757756

5. Buhay JE. "COI-like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. J Crust Biol. 2009; 29: 96–110.

6. Baeza JA, Fuentes MS. Exploring phylogenetic informativeness and nuclear copies of mitochondrial DNA (numts) in three commonly used mitochondrial genes: mitochondrial phylogeny of peppermint, cleaner, and semi-terrestrial shrimps (Caridea: *Lysmata*, *Exhippolysmata*, and *Merguia*). Zoolog J Linnean Soc. 2013; 168: 699–722.

7. Kim SJ, Lee KY, Ju SJ. Nuclear mitochondrial pseudogenes in *Austinograea alayseae* hydrothermal vent crabs (Crustacea: Bythograeidae): effects on DNA barcoding. Mol Ecol Res. 2013; 13: 781–787.

8. Miller CB, Frost BW, Batchelder HP, Clemons MJ, Conway RE. Life histories of large, grazing copepods in a subarctic ocean gyre: *Neocalanus plumchrus*, *Neocalanus cristatus*, and *Eucalanus bungi* in the Northeast Pacific. Prog Oceanog. 1984; 13: 201–243.

9. Kobari T, Ikeda T, Vertical distribution, population structure and life cycle of *Neocalanus cristatus* (Crustacea: Copepoda) in the Oyashio region, with notes on its regional variations. Mar Biol. 1999; 134: 683–696.

10. Tsuda A, Saito H, Kasai H. Life histories of *Neocalanus flemingeri* and *Neocalanus plumchrus* (Calanoida: Copepoda) in the western subarctic Pacific. Mar Biol. 1999; 135: 533–544.

11. Machid RJ, Miya MU, Nishida M, Nishida M. Molecular phylogeny and evolution of the pelagic copepod genus *Neocalanus* (Crustacea: Copepoda). Mar Biol. 2006; 148: 1071–1079.

12. Machida RJ, Miya MU, Nishida M, Nishida S. Large-scale gene rearrangements in the mitochondrial genomes of two calanoid copepods *Eucalanus bungii* and *Neocalanus cristatus* (Crustacea), with notes on new versatile primers for the srRNA and COI genes. Gene. 2004; 332: 71–78. doi: 10.1016/j.gene.2004.01.019 PMID: 15145056

13. Katoh K, Standley DM. MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013; 30: 772–780. doi: 10.1093/molbev/mst010 PMID: 23329690

14. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML web servers. Syst Biol. 2008; 57: 758–771. doi: 10.1080/10635150802429642 PMID: 18853362

15. Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science. 1988; 239: 487–491. PMID: 2448875

16. Taanman JW. The mitochondrial genome: structure, transcription, translation and replication. BBA-Bioenergetics. 1999; 1410: 103–123. PMID: 10076021

17. Collura RV, Auerbach MR, Stewart CB. A quick, direct method that can differentiate expressed mitochondrial genes from their nuclear pseudogenes. Curr Biol. 1996; 6: 1337–1339. PMID: 8939570

18. Kebschull JM, Zador AM. Sources of PCR-induced distortions in high-throughput sequencing data sets. Nucl Acid Res. 2015; 43: e143.

19. Pääbo S, Irwin DM, Wilson AC. DNA damage promotes jumping between templates during enzymatic amplification. J Biol Chem. 1990; 265: 4718–4721. PMID: 2307682

20. Zischler H, Geisert H, Vonhaeseler A, Pääbo S. A nuclear fossil of the mitochondrial d-loop and the origin of modern humans. Nature. 1995; 378: 489–492. doi: 10.1038/378489a0 PMID: 7477404

21. Wyngaard GA, Rasch EM. Patterns of genome size in the copepoda. Hydrobiologia. 2000 417: 43–56.

22. Graur D, Shuali Y, Li WH. Deletions in processed pseudogenes accumulate faster in rodents than in humans. J Mol Evol. 1989; 28: 279–285. PMID: 2499684

23. Petrov DA. Evolution of genome size: new approaches to an old problem. Trends Genet. 2001; 17: 23–28. PMID: 11163918

24. Adams KL, Daley DO, Qiu YL, Whelan J, Palmer JD. Repeated, recent and diverse transfer of a mitochondrial gene to the nucleus in flowering plants. Nature. 2000; 408: 354–357. doi: 10.1038/35042567 PMID: 11099041

25. Bensasson D, Petrov DA, Zhang DX, Hartl DL, Hewitt GM. Genomic gigantism: DNA loss is slow in mountain grasshoppers. Mol Biol Evol. 2001; 18: 246–253. PMID: 11158383

26. Berg OG, Kurland CG. Why mitochondrial genes are found most often in nuclei. Mol Biol Evol. 2000: 17: 951–961. PMID: 10833202

27. Hlaing T, Willoughby TL, Somboon P, Socheat D, Setha T, Min S, et al. Mitochondrial pseudogenes in the nuclear genome of *Aedes aegypti* mosquitoes: implications for past and future population genetic studies. BMC Genetics. 2009; 10: 11. doi: 10.1186/1471-2156-10-11 PMID: 19267896

28. Simone D, Calabrese FM, Lang M, Gasparre G, Attimonelli M. The reference human nuclear mitochondrial sequences compilation validated and implemented on the UCSC genome browser. BMC Genomics. 2011; 12: 517. doi: 10.1186/1471-2164-12-517 PMID: 22013967

29. Antunes A, Pontius J, Ramos MJ, O'Brien SJ. Johnson WE. Mitochondrial introgressions into the nuclear genome of the domestic cat. J Heredity. 2007; 98: 414–420.

30. Kim JH, Antunes A, Luo SJ, Menninger J, Nash WG, O'Brien SJ, et al. Evolutionary analysis of a large mtDNA translocation (numt) into the nuclear genome of the *Panthera* genus species. Gene. 2006; 366: 292–302. doi: 10.1016/j.gene.2005.08.023 PMID: 16380222

31. Hazkani-Covo E, Zeller RM, Martin W. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. PLoS Genetics. 2010; 6(2): e1000834. doi: 10.1371/journal.pgen.1000834 PMID: 20168995

32. Kamimura N, Ishii S, Liandong M, Shay JW. Three separate mitochondrial DNA sequences are contiguous in human genomic DNA. J Mol Biol. 1989; 16: 111–120.

33. Lopez JV, Yuhki N, Masuda R, Modi W, Obrien SJ. Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. J Mol Evol. 1994; 39: 174–190. PMID: 7932781

34. Sunnucks P, Hales DF. Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Mol Biol Evol. 1996; 13: 510–524. PMID: 8742640

35. Bensasson D, Zhang DX, Hewitt GM. Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. Mol Biol Evol. 2000; 17: 406–415. PMID: 10723741

36. Nielsen KK, Arctander P. Recombination among multiple mitochondrial pseudogenes from a passerine genus. Mol Phylogenetics Evol. 2001; 18; 362–369.

37. Kobari T, Ikeda T. Life cycle of *Neocalanus flemingeri* (Crustacea: Copepoda) in the Oyashio region, western subarctic Pacific, with notes on its regional variations. Mar Ecol Prog Ser. 2001; 209: 243–255.

38. Kobari T, Ikeda T. Ontogenetic vertical migration and life cycle of *Neocalanus plumchrus* (Crustacea: Copepoda) in the Oyashio region, with notes on regional variations in body sizes. J Plankton Res. 2001; 23: 287–302.

39. Parent GJ, Plourde S, Turgeon J. Naturel hybridization between *Calanus finmarchicus* and *C. glacialis* (Copepoda) in the Arctic and Northwest Atlantic. Limnol Oceanog. 2012; 57: 1057–1066.