Supplementary Table S1. Primers used in this study.

Name	Sequence (5'-3')
PCR and RACE for sequencing	
Pmcox1F1	TTCAACGGTGATGTGGTATTATAAC
Pmcox1R3	CTAAACCTACTGTATATATATGATGAGC
cox1-3f	ACNGGITGRACIYTITAYCCNCCNYT
Pmcox1-5RACE	AAGTGTTGATATAATACTACATCACC
Pmcox1-3RACE	ATGGTGATGTAGGTATTATATCAACAC
Pmcox1fullF	CTAATCAGTAATCGTGATACGCTAACC
Pmcox1fullR	TGAACCAATAGATGATATTAAATTCCATAC
Probe amplification for southern hybridization	
Pmcox1pF	GCCTAGGTTTATATGGTGGTATACC
Pmcox1pR	CTTCTAGGGGCATTACATTGAAACC
nuLSU-7f	TCCTGAGGGAAACTTCGGAGG
nuLSU-7r	GATAGCAACAAGTACCGTGAGG
PmNumt1F	GTCGAGGCAGCACAGTACTG
PmNumt1R	CGATTCTCGACATCACAGTCACC
PmNumt2F	CCAAGAGACATGCATGAGTAGC
PmNumt2R	GGTGACTGTGATACGACTGCC

Supplementary Table S2. Codon usage of *Pmcox1* and *Perkinsus* nuclear genes.

Codon	Amino acid	*Pmcox1*	*ispC*	*sod1*	*sod2*	*actin*
UUU	F	14.9	5.6	4.4	7.1	0
UUC		25.6	18.6	53.1	25	39.9
UUA	L*	119.4	0	4.4	0	0
UUG		0	9.3	31	32.1	23.9
CUU	L*	0	11.1	4.4	7.1	8
CUC		0	27.8	13.3	46.4	18.6
CUA		0	9.3	0	7.1	2.7
CUG		0	39	13.3	28.6	13.3
AUU	I	74.6	16.7	26.5	10.7	26.6
AUC		6.4	37.1	22.1	17.9	42.6
AUA		185.5	14.8	4.4	10.7	0
AUG	M	23.5	29.7	17.7	25	47.9
GUU	V	8.5	18.6	13.3	17.9	18.6
GUC		0	24.1	31	17.9	29.3
GUA		49	13	13.3	7.1	5.3
GUG		0	16.7	4.4	21.4	16
UCU	S	25.6	13	13.3	25	26.6
UCC		4.3	9.3	0	3.6	16
UCA		23.5	7.4	4.4	3.6	2.7
UCG		2.1	20.4	4.4	21.4	2.7
CCU	P	10.7	20.4	17.7	17.9	18.6
CCC		0	9.3	0	10.7	23.9
CCA		14.9	11.1	13.3	14.3	5.3
CCG		0	11.1	17.7	10.7	2.7
ACU	T	2.1	14.8	31	17.9	42.6
ACC		0	9.3	22.1	14.3	18.6
ACA		27.7	9.3	8.8	3.6	2.7
ACG		0	11.1	4.4	10.7	10.6
GCU	A*	14.9	22.3	66.4	25	39.9
GCC		0	39	17.7	17.9	23.9
GCA		0	27.8	8.8	28.6	2.7
GCG		0	31.5	8.8	10.7	0

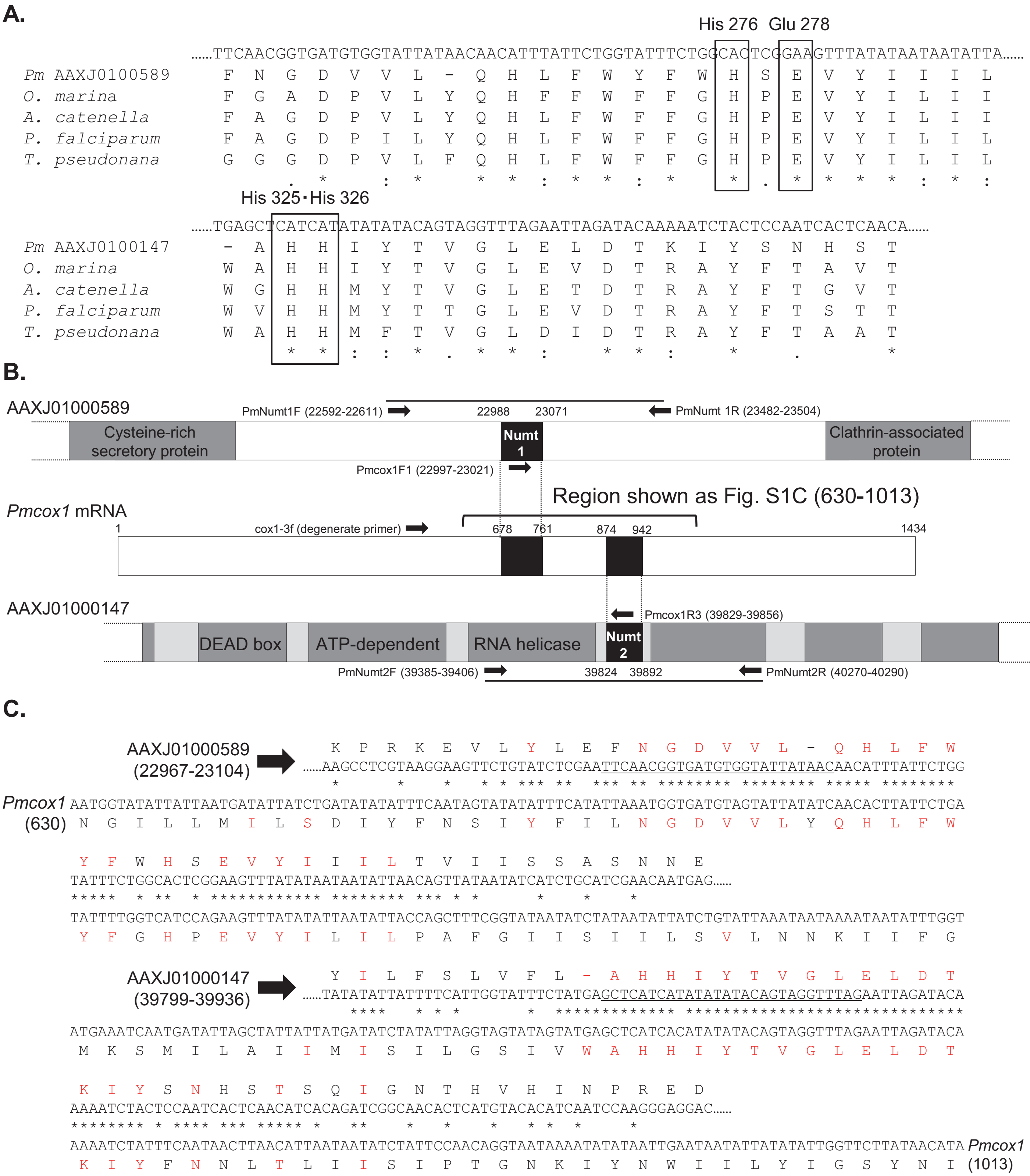
Codon	Amino acid	*Pmcox1*	*ispC*	*sod1*	*sod2*	*actin*
UAU	Y	72.5	1.9	22.1	7.1	21.3
UAC		10.7	14.8	17.7	32.1	13.3
UAA	Stop	0	0	0	0	0
UAG		0	0	0	0	0
CAU	H	12.8	7.4	35.4	21.4	13.3
CAC		4.3	9.3	8.8	10.7	10.6
CAA	Q*	8.5	9.3	13.3	0	2.7
CAG		0	24.1	17.7	14.3	26.6
AAU	N	70.4	7.4	17.7	17.9	10.6
AAC		6.4	14.8	53.1	17.9	13.3
AAA	K*	25.6	3.7	4.4	7.1	13.3
AAG		0	44.5	53.1	28.8	47.9
GAU	D*	23.5	29.7	17.7	28.6	31.9
GAC		0	24.1	31	25	26.6
GAA	E*	10.7	16.7	0	14.3	16
GAG		0	27.8	48.7	28.6	58.5
UGU	C^{Ψ}	0	3.7	0	3.6	2.7
UGC		0	14.8	8.8	3.6	8
UGA	Stop/W	10.7	0	0	0	0
UGG	W	4.3	7.4	22.1	32.1	10.6
CGU	R*	0	9.3	8.8	10.7	13.3
CGC		0	5.6	8.8	3.6	18.6
CGA		0	9.3	0	25	2.7
CGG		0	11.1	0	21.4	0
AGU	S	21.3	14.8	17.7	25	10.6
AGC		0	26	13.3	3.6	5.3
AGA	R*	12.8	13	4.4	3.6	5.3
AGG		0	5.6	8.8	21.4	8
GGU	G	64	16.7	26.5	10.7	42.6
GGC		4.3	31.5	26.5	21.4	16
GGA		4.3	13	17.7	10.7	13.3
GGG		0	24.1	0	32.1	5.3

	Pmcox1	*ispC*	*sod1*	*sod2*	*actin*
gene length (bp)	1410	1617	678	840	1128
total codon number	470	539	226	280	376
codon kind used	35	60	53	59	56
A+T content	81%	43%	50%	45%	49%

The numbers represent codon frequencies per thousand. Note that the eight As and two CCs which are hypothetically skipped in translation were manually eliminated.

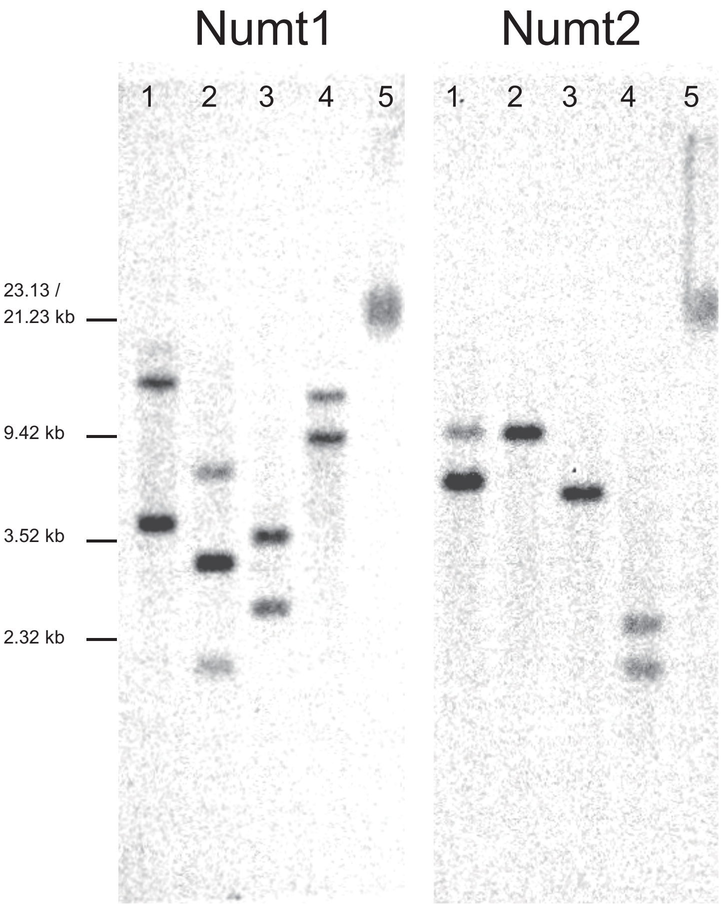
* The amino acid coded only by one codon species in PmCOX1.

$^{\Psi}$ The amino acid which is not contained in PmCOX1.



Supplementary Figure S1.

A, *cox1*-like Numt sequences (represented as NCBI entry number due to the unavailability of TIGR database) and alignment of deduced COX1 amino acid sequences with those of related species. Symbols are in the same way as in Figure 4; B, Schematic model of Numt-containing contigs and *Pmcox1* (not scaled). Dark and light gray boxes indicate exons and introns, respectively, and black boxes indicate *cox1*-like Numts. Arrows stand for primers, and regions used as probe for southern hybridization are indicated by bars; C, Pairwise alignment of each of Numt sequences and *Pmcox1,* and their corresponding amino acid sequences. Identical nucleotides and amino acid residues are represented by asterisks and red characters, respectively.



Supplementary Figure S2.
Southern hybridization images using Numt1 (left) and Numt2 (right) probes. Lanes 1-4, *P. marinus* genomic DNA digested with BamHI (1), HindIII (2), PstI (3), and SalI (4); Lane 5, uncut genomic DNA.

```
Omar            ---------------------CACAGUUCUAUUCGCAUUUCCUCUUUAAACCAAUUUCUACGGGCGAAUCAUAAAAGGCUUGCUUGCUAUUAUCUUAUCUCUGCUAUUCUUUUCGGAGUU
Acat            ------------UACAGAUUUUUGAAUUCUUUUUCCUUAUCUUUUCUAUCACAUGUAAAAAAUUGUAAUCAUAAACGCUUAGGAAUCUAUUAUUUAUUAUCAGCAUUCAUCUUUGGGAUC
Pfal            ------------------------AUUUUUAUUGUUUUAAAUAGAUAUUCACUUAUUACAAAUUGUAACCAUAAAACUUUAGGAUUAUACUAUUUAUGGUUUUCAUUUUUAUUUGGUAGU
Pmar            UAAUCUAAUCAGUAAUCGUGAUACGCUAACCAAUAUAAAUAUAAUAAAUUCUAUAUAUUAGUAAAUAAUAAUAAAAGAAUAGGAAUAUAUUACAUUAAAUUAUCAAUAAUUAUUGGUAUA
                                          *                            *   *      **  *****    * *     ** **  *    *   *  *  *  *  * **

Omar            UCUGCUACACUCUUCUCUCUUCUCUUACGAAUAGAAUUAUAUAGCUCUGGAAACAGGAUCAUUCCUUUCGAAAAUCAAAACUUCUACAACUUAAGCAUUACUUUACAUGGAUUACUUAUG
Acat            UCCGGUACAAUAGUCUCAGUCCUUAUGAGAAUAGAAUUAUAUUCAUCAGGAAAUAGGAUUAUAUCUCCAGAAAACCAGAACUUCUAUAAUGUAAGCAUAACCUUGCAUGGGUUUCUUAUG
Pfal            UAUGGAUUUUUAUUAUCAGUAAUACUACGUACUGAAUUAUAUUCUUCAUCUUUAAGAAUAAUUGCACAAGAAAAUGUAAAUCUAUAUAAUAUGAUAUUUACAAUUCACGGAAUAAUUAUG
Pmar            UUAGGUAUAGUAUUAUCUUAUAUAAUUAGAGUUGAAUUAUAUAAUAGUGGUAAUAGAAUUAUUAAAUAUGAUAAUGUAAAUUACUAUAAUAUGGUUAUAACAUUACAUGGUUUAUUAAUG
                *  *      *  * **     *  *  *    *********            ** ** **       ** **    **    ** **  *     * **  * ** **  *  * ***

Omar            AUAUUCUUUCUUGUUAUGCCU-GGUCUCUAUGGAGCCUU--UGGUAAUUAUUUUCUUCCAGUCUUUCUUGGCAGUCCUGAAGUUGCUUUUCCCAGAGUAAACAGCUUUUCUUUCUUAUUA
Acat            AUCUUCUUUUUAGUAAUGCCC-GGCUUGUUUGGAGGAUU--UGGAAAUUAUUUUGCACCUAUCUUUCAAGGGUCUCCAGAAGUGGUAUCCUAGAGUCAAUGAUUUUUCGAUCUUAGUU
Pfal            AUUUUUUUCAAUAUAAUGCCA-GGAUUAUUCGGAGGAUU--UGGUAAUUACUUUCUACCUAUUUUAUGUGGAUCUCCAGAAUUAGCAUAUCCUAGAAUUAAUAGUAUAUCUUUACUGUUA
Pmar            AUAUUUUAUAUUAUAAUGCCUAGGUUUAUAUGGUGGUAUACCCCUAAUUAUAUAUUACCAAUAUUAAGUGUAAUAACAGAUAUUGUAUUACCAAGAAUAAAUAAUAUAUCUAUAAUUAUU
                ** ** *      * *****  **  * *  ** *   *      *****  *    **  * **    *      * **  * *  *  ** *** * **     * **  *  *  *

Omar            CUGCCAGUAUCUUAUGCUUUUGUCAUUCUAUCGACUGCUUCCGAAUUUGGUGCAGC-AUUAGGUUGGACUUUAUAUCCACCAUUGAGCACUUCUUUGAUGAAUCUUUCAGCAGCUGCUGU
Acat            CUUUUUCUUUCAUAUCUUUUCGUAAUCCUUUCUUUAAUCUCAGAGUUUGGAGGUGG-AACAGGAUGGACAUUAUAUCCACCAUUAAGCACUUCUUUCAUGAGUUUAUCACCUUCAAGUAC
Pfal            CAACCAAUUGCUUUUGUUUUAGUUAUAUUAUCUACUGCAGCAGAAUUUGGUGGUGG-AACUGGAUGGACUUUAUAUCCACCAUUAAGUACAUCUUUAAUGUCAUUAUCUCCUGUAGCUGU
Pmar            GUAUUAAUAUCAUAUAUAGUAGUAAUAAAUAGUAUUGUAAUAGAAUACAAUAUAGGUACUAGGCUGAACAUUAUAUCCUCCAUUA------UCAAUUAUAGGU------ACAGUAAUAGU
                       *  * * *    * ** **                ** *        *  *   ** ** ** ******** *****       **  * **           *

Omar            UGAUCUUAUUAUCGAUGGUUUAAUUAUUAGU-GGUGUUUCAUCUUUAUUCGGCUCAUUGAACUUCUUAGCUACUAUUGCAGCUUAUGGA----GUUCUCCGUUCUU-----CUUUUCCUC
Acat            AGCAAAUCUUAUCUUUGGAUUAUUAAUCUCA-GGUAUAUCCUCAUGUCUCACAUCUCUUAACUUUUGGGUAACAAUUCUAAAUCUGAGAUCUUAUUGUCUGACAUUAAAGACUAUGCCAU
Pfal            AGAUGUAAUAAUUUUUGGUUUAUUAGUAUCU-GGAGUCGCUAGUAUUAUGUCUUCAUUAAAUUUUAUUACUACAGUAAUGCAUUUAAGAGCAAAAGGAUUAACACUUGGUAUAUUAAGUG
Pmar            UAAUAUGAUAUUAUAUGGUUUAAUAAUUAUUAGGCAUAUCCUCUAUAAUAUCUGCUAUUAAUUUCAUGAAUAUAUUAAUAGUAAUUG-----------AUGGAAU-------AAUAUAUG
                        *  *   *** *** *  *     **  *  *        *     *  * ** **       *   *                                      *

Omar            UAUUUACUGCUUCAAUACUUAUCACGGCUAUCUUACUGAUUUUAACCUUGCCUGUAUUAUCUGCUGCUUUAUUAAUCUUAUUAGCUGAUCUUCAUUUUAACACAGUUUAUUUCAACCCAG
Acat            UAUUCCCUUGGGCUCUCUUGAUUACAGGAGGAAUGCUUUUAUUAACAUUACCAAUCUUAUCAGGAGCUUUUCUAAUGGUCUUGGCUGAUCUUCAUUCUAAUACACUUUUCUUUGAUCCAA
Pfal            UUUCUACAUGGUCAUUGAUCAUUACAUCAGGAAUGUUAUUGCUAACACUACCGGUUUUAACUGGAGGAGUAUUAAUGUUAUUAUCAGACUUACAUUUUAAUACUUUAUUUUUUGACCCAA
Pmar            UAUAUAUAUGGAGUAUUAUUAUAACAAGUGUAUUAUUAAUUAUUUCAUUACCAAUAUUAAAUGGUAUAUUAUUAAUGAUAUUAUCUGAUAUAUAUUUCAAUAGUAUAUAUUUC------A
                * *            *  * ** **        *  *  *  *  *  * **  * ***   *      *  ****  *  * **  * **  *  ***  ** *   *   *  **

Omar            CUUUUGCUGCUGAUCCUGUUCUAUAUCAACAUUUCUUCUGGUUCUUUGGUCACCCUGAAGUUUAUAUUCUCAUAAUUCCUGCUUUCGGUAUUAUAAGUGAAGGAAUCUCUCUUUUUAGUC
Acat            UCUUUGGAGGAGAUCCUGUACUCUAUCAACACUUAUUUUGGUUUUUUGGACAUCCAGAAGUUUACAUCUUAAUAAUUCCAGCAUUUGGGGUCAUUAGCAUAGUAAUUUCUGGAAUUUCAC
Pfal            CAUUUGCAGGAGAUCCAAUAUUAUAUCAACAUUUAUUCUGGUUUUUUGGACAUCCUGAAGUAUACAUUUUAAUAUUACCUGCUUUUGGAGUAAUUAGUCAUGUAAUUUCUACUAAUUAUU
Pmar            UAUUAAAUGGUGAUGUAGUAUUAUAUCAACACUUAUUCUGAUAUUUUGGUCAUCCAGAAGUUUAUAUAUUAAUAUUACCAGCUUUCGGUAUAAUAUCUAUAAUAUUAUCUGUAUUAAAUA
                  **    *  ***    *  * ******** ** ** ** *  ***** ** ** ***** ** **  * *** * ** ** ** **  * **         * * ***

Omar            AAAAGAUUAUUCUUGCUUUUGAUUCUAUGAUCCUUGCUUUGCUUUGUCUUUCUAUUCUCGGAUCAGGUGUAUGGGCACAUCAUAUUUAUACUGUUGGGUUAGAAGUAGAUACAAGAGCAU
Acat            AAAAAAUAAUCUUUGGGAACCAAUCAAUGAUCUUUGCCAUGAGCUGUAUUUCUCUUCUUGGAACUGUUGUUUGGGGACAUCAUAUGUAUACUGUAGGUUUAGAAACUGAUACAAGAGCUU
Pfal            GCAGAAAUCUAUUUGGUAAUCAAUCUAUGAUACUUGCUAUGGGAUGUAUAGCUGUUUUAGGAAGCUUAGUAUGGGUACAUCAUAUGUACACUACUGGUUUAGAAGUUGAUACUAGAGCUU
Pmar            AUAAAAUAAUAUUUGGUAUGAAAUCAAUGAUAUUAGCUAUUAUUAUGAUAUCUAUAUUAGGUAGUAUAGUAUGAGCUCAUCACAUAUAUACAGUAGGUUUAGAAUUAGAUACAAAAAUCU
                  *  *   *  ***      * ** *****  * **  *        *  ** *  *  * **       ** ** *  ***** ** ** **    ** ******   ***** * *   *

Omar            ACUUUACAGCUGUUACGAUUAUGAUCUCUUUACCAACAGGAACGAAAGUCUUUAAUUGGUUCUGCUCUUAUCUUGGUAAUUCAAUGCAAUUGAGCAAUAUAACA---UCAAUAAAGUGGA
Acat            AUUUUACAGGAGUUACAAUCUUGAUAUCCUUACCGACUGGUACAAAAAUCUUUAAUUGGUUAAGUACAUAUCUUGGGAAUCCGCCAUUGUUACACCUCAAAACUAGUUCAGCAUUCUUUG
Pfal            AUUUUACUUCGACUACCAUUUUAAUAUCAAUACCUACCGGUACAAAAGUAUUUAACUGGAUAUGUACAUAUAUGAGUAGU--AAUUUUGGUAUGAUACACAGCU----CUUCAUUAUUGU
Pmar            AUUUCAAUAACUUAACAUUAAUAAUAUCUAUUCCAACAGGUAAUAAAAUAUAUAAUUGAAUAAUAUUAUAUAUUGGUUCUUAUAACAUUAUUAUAUAAUGGUUAUCAAUCAUUAAUAUUCA
                * ** *        **  *  *  ** **  * ** ** ** *  *** * * *** **  *       *** *  *   *          *                 *   *   *

Omar            UUAUCUUAUUCUUGCUUACCUUUACACUUGGU-GGUUCUACUGGAGUAAUCUUGGCAAACUCUGCAACAGAU-UUGGCUCUUCAUGACACCUAUUAUGUUGUAGCACAUUUUCACUUUGU
Acat            CACUUCUCUUUUUAUUAAUGUUUACUAUAGGU-GGCUCAACAGGAGUAAUUCUUGGAAAUGCUGCUGUGGAU-CUAGGAUUACAUGAUACAUAUUAUGUUGUAGCUCAUUUUCAUUUUGU
Pfal            CAUUAUUAUUUAUAUGUACAUUUACAUUUGGA-GGUACUACUGGAGUUAUAUUAGGUAAUGCUGCCAUUGAU-GUAGCAUUACAUGACACAUUAUUAUGUUAUUGCUCAUUUCCAUUUUGU
Pmar            GUAUAAUGUUUAUUAUAAUAUUUAUAAUAGGUAGGUAUAACAGGUAUAAUUAUAAGUAUAGAUAUCAUAGAUAUUAGGUUUACAUGAUACAUAUUAUAUAGUAUCUCAUUUCCAUUAUAU
                   *  * **  *    *  ****   *  **   **    ** **  * **  *    *    *      ***  * *   * ***** ** ****** *  *  * ***** ** * * *

Omar            UCUUAGUCUUGGUGCAGUAAUAGCCCUUUUCUCUUCUUUAAUCUUAUAUCAAAAGGUUAUCUUUGCUUCUGA---UUUAUCAUUUUA---UUCAAGUACAGUCUUUC------AUUUCUU
Acat            UCUUUCUUUAGGAGCUGUAAUUGCUAUCUUCUCUGGAAUAAUCUUCAAUGGAGAAAAGAUUCUUGGCUCUAAGAGUUUAUUACCUUCACCUUCCAGUACUCUCUCUCUUUAUCAUUUGGU
Pfal            ACUAUCAAUUGGUGCAAUUAUUGGAUUAUUUACAACUGUAAGUGCAUUUCAAGAUAAUUUCUUUGGUAAAAA------CUUACGUGAAAAUUCUAUUGUAAUACUAU------GGUCAAU
Pmar            AUUAUCAAUUGGUGCUGUAAUAUCAUUAUUAGCUAGGU--AUAUUAUUAUUAAAAGAUAUUAUUGGUUAUUA-------UAAUGUAAUAAUUAAAAUAAAUAAAUACUUUA  GGUU--U
                  *     * ** **  * **     * **  *       *          * *     *  ***      *       * *  *     **  * *                  *   *

Omar            UCUUUCAUUUAUCGCUGUUUUAAUGACUUUUACGCCAAUGCAUUUC-UUGGGCUUUAAUGUUAUGCC--AAGACGCAUCUCUGAUAUUACAGAUAAUUUUAACUCAUGGAAUUAUGUGUC
Acat            UUUAACAUUUGUUGGUAUUCUUCUAACCUUUUCCCCAAUGCAUUUC-UUAGGAUUUAAUGUUAUGCC--AAGAAGAAUCCCGGACUUUCCAGAUUCUUUUCAUUCCUGGAAUUUUCUGUC
Pfal            GUUAUUUUUUGUAGGUGUAAUAUUAACAUUUUUACCUAUGCAUUUU-UUAGGAUUUAAUGUAAUGCC--UAGACGUAUUCCUGAUUAUCCAGACGCUUUAAAUGGAUGGAAUAUGAUUUG
Pmar            AUUAUUAUUCAUAAAUAUUAAUAUAAUAUUCACACCUCAAUUUAUAAUUAGGUUUCAAUGUAAUGCCCCUAGAAGAAUAUUAGAAUAUAGUGAUAAUAUCAUUGUAUGGAAUUUAAUAUC
                  *    **  *   * *     * *  **    **      * * ** ** ** ***** *****   *** * **    **   *   **   *  * *       ******    * *

Omar            GUCUGUAGGCUCAAUCUUCACUCUUGUAUCUGCAGCUUUAUUAAAAAAAAAAAAA
Acat            UUCUAUUGGAUCAGGAAUAACUCUCUUUUCCUUUGCUAUUUU-------------
Pfal            UUCUAUUGGGUCAACAAUGACUUUAUUUGGUUUACUAAUUUUUAAAUAA------
Pmar            AUCUAUUGGUUCAAUAUCAACAAUAUUAAUAUUAUUAUCCAUA------------
                 *** * ** ***      **  *  *              *
```

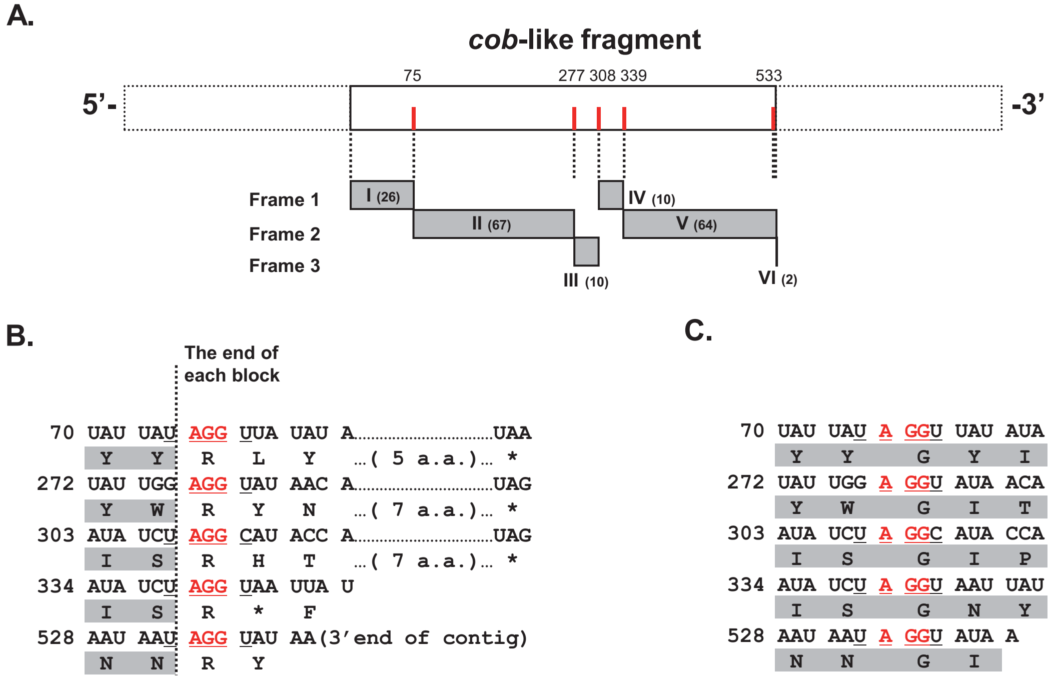
Supplementary Figure S3.

Alignment of several *cox1* sequences and *Pmcox1* mRNA sequence. Asterisks denote conserved nucleotides among four sequences. One- or two-base indels specific in *PmcoxI* are indicated with red squares. UAGGY and CCCCUA motifs in the vicinities of the indels are highlightened in black.

```
    ************ *********** ********************************* ***** ** ******** *********** ***
Ph  AUCAAUUAUAGGAACAGUAAUAGUAAAUAUGAUAUUAUAUGGUUUAAUAAUUAUUAGGUAUAUCAUCUAUAAUAUCUGCUAUUAAUUUCAUG
Pm  AUCAAUUAUAGGUACAGUAAUAGUUAAUAUGAUAUUAUAUGGUUUAAUAAUUAUUAGGCAUAUCCUCUAUAAUAUCUGCUAUUAAUUUCAUG
Po  AUCAAUUAUAGGUACAGUAAUAGUUAAUAUGAUAUUAUAUGGUUUAAUAAUUAUUAGGUAUAUCAUCAAUAAUAUCAGCUAUUAAUUUUAUG
aa    S  I  I  G  U  V  I  V  N  M  I  L  Y  G  L  I  I  I   G  I  S  S  I  I  S  A  I  N  F  M

    ************** ******** ******************************************************************
Ph  AAUAUAUUAAUAGUAAUUGAUGGUAUAAUAUAUGUAUAUAUAUGGAGUAUUAUUAUAACAAGUGUAUUAUUAAUUAUUUCAUUACCAAUA
Pm  AAUAUAUUAAUAGUAAUUGAUGGAAUAAUAUAUGUAUAUAUAUGGAGUAUUAUUAUAACAAGUGUAUUAUUAAUUAUUUCAUUACCAAUA
Po  AAUAUAUUAAUAGUUAUUGAUGGUAUAAUAUAUGUAUAUAUAUGGAGUAUUAUUAUAACAAGUGUAUUAUUAAUUAUUUCAUUACCAAUA
aa   N  I  L  I  V  I  D  G  I  I  Y  V  Y  I  W  S  I  I  I  T  S  V  L  L  I  I  S  L  P  I

    ******************************************************************** *********************
Ph  UUAAAUGGUAUAUUAUUAAUGAUAUUAUCUGAUAUAUAUUUCAAUAGUAUAUAUUUCAUAUUAAAUGGUGAUGUAGUAUUAUAUCAACAC
Pm  UUAAAUGGUAUAUUAUUAAUGAUAUUAUCUGAUAUAUAUUUCAAUAGUAUAUAUUUCAUAUUAAAUGGUGAUGUAGUAUUAUAUCAACAC
Po  UUAAAUGGUAUAUUAUUAAUGAUAUUAUCUGAUAUAUAUUUCAAUAGUAUAUAUUUCAUAUUAAAUGGAGAUGUAGUAUUAUAUCAACAC
aa   L  N  G  I  L  L  M  I  L  S  D  I  Y  F  N  S  I  Y  F  I  L  N  G  D  V  V  L  Y  Q  H

    ************** ******************************* *************************************** ***
Ph  UUAUUCUGAUAUUUCGGUCAUCCAGAAGUUUAUAUAUUAAUAUUACCUGCUUUCGGUAUAAUAUCUAUAAUAUUAUCUGUAUUAAAUAAU
Pm  UUAUUCUGAUAUUUUGGUCAUCCAGAAGUUUAUAUAUUAAUAUUACCAGCUUUCGGUAUAAUAUCUAUAAUAUUAUCUGUAUUAAAUAAU
Po  UUAUUCUGAUAUUUCGGUCAUCCAGAAGUUUAUAUAUUAAUAUUACCAGCUUUCGGUAUAAUAUCUAUAAUAUUAUCUGUAUUAAACAAU
aa   L  F  W  Y  F  G  H  P  E  V  Y  I  L  I  L  P  A  F  G  I  I  S  I  I  L  S  V  L  N  N

    *********************** ************************** ********************
Ph  AAAAUAAUAUUUGGUAUGAAAUCUAUGAUAUUAGCUAUUAUUAUGAUAUCAAUAUUAGGUAGUAUAGUAUGG
Pm  AAAAUAAUAUUUGGUAUGAAAUCAAUGAUAUUAGCUAUUAUUAUGAUAUCUAUAUUAGGUAGUAUAGUAUGA
Po  AAAAUAAUAUUUGGUAUGAAAUCUAUGAUAUUAGCUAUUAUUAUGAUAUCUAUAUUAGGUAGUAUAGUAUGA
aa   K  I  I  F  G  M  K  S  M  I  L  A  I  I  M  I  S  I  L  G  S  I  V  W
```

Supplementary Figure S4.

Alignment of partial *Pmcox1* mRNA (Pm) and its orthologous sequences obtained from *P. olseni* (Po) and *P. honshuensis* (Ph), and their corresponding amino acid sequence deduced based on the frameshift model. Asterisks show the conserved residues.The UAGGY motifs which seem to be involved in frameshift are written in black background.



Supplementary Figure S5.

A, COB-like blocks dispersed in three reading frames of *P. marinus cob*-like fragment found in the database; B, translation with standard codon table; C, translation with the frameshift at in-frame AGG. Detail descriptions are found in the legend of Figure 3. Note that U of the UAGGY motif is substituted by G in block II.