

GOPEN ACCESS

Citation: Yan L, Hou Z, Ma J, Wang H, Gao J, Zeng C, et al. (2022) Complete mitochondrial genome of *Episymploce splendens* (Blattodea: Ectobiidae): A large intergenic spacer and lacking of two tRNA genes. PLoS ONE 17(6): e0268064. https://doi.org/10.1371/journal.pone.0268064

Editor: Tzen-Yuh Chiang, National Cheng Kung University, TAIWAN

Received: May 9, 2021

Accepted: April 22, 2022

Published: June 2, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0268064

Copyright: © 2022 Yan et al. This is an open access article distributed under the terms of the <u>Creative</u> Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Our Data has been uploaded to NCBI Genbank database and is available at the GenBank Accession number RESEARCH ARTICLE

Complete mitochondrial genome of *Episymploce splendens* (Blattodea: Ectobiidae): A large intergenic spacer and lacking of two tRNA genes

Lin Yan¹, Zhenzhen Hou², Jinnan Ma¹, Hongmei Wang¹, Jie Gao³, Chenjuan Zeng³, Qin Chen¹, Bisong Yue¹, Xiuyue Zhang^{1,2*}

1 Key Laboratory of Bio-Resources and Eco-Environment, Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, China, 2 Sichuan Key Laboratory of Conservation Biology on Endangered Wildlife, College of Life Sciences, Sichuan University, Chengdu, China, 3 Sichuan Key Laboratory of Medicinal Periplaneta Americana, Sichuan Gooddoctor Pharmaceutical Group, Chengdu, China

* zhangxiuyue@scu.edu.cn

Abstract

The complete mitochondrial genome of *Episymploce splendens*, 15,802 bp in length, was determined and annotated in this study. The mito-genome included 13 PCGs, 20 tRNAs and 2 rRNAs. Unlike most typical mito-genomes with conservative gene arrangement and exceptional economic organization, *E. splendens* mito-genome has two tRNAs (tRNA-GIn and tRNA-Met) absence and a long intergenic spacer sequence (93 bp) between tRNA-Val and srRNA, showing the diversified features of insect mito-genomes. This is the first report of the tRNAs deletion in blattarian mito-genomes and we supported the duplication/random loss model as the origin mechanism of the long intergenic spacer. Two Numts, Numt-1 (557 bp) and Numt-2 (975 bp) transferred to the nucleus at about 14.15 Ma to 22.34 Ma, and 19.19 Ma to 24.06 Ma respectively, were found in *E. splendens*. They can be used as molecular fossils in insect phylogenetic relationship inference. Our study provided useful data for further studies on the evolution of insect mito-genome.

Introduction

The mitochondrial genome has the characteristics of maternal inheritance, rare genetic recombination, faster evolution and conservative gene arrangement [1]. The mito-genome commonly displays exceptional economy of organization, with overlapping genes or small intergenic spacers [2]. A typical insect mito-genome, with a compact circular molecule, generally encodes a fixed set of 37 genes including 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes, two ribosomal RNA (rRNA) genes and a control region [3]. However, exceptions have been reported, such as long gene intergenic spacer [4–6], gene rearrangement and gene loss [7, 8]. So far, mitochondrial gene rearrangement has been reported in about 17 orders of insects, and these genes include protein coding gene, rRNA gene and tRNA gene [8]. In addition tRNAs, the missing protein coding genes have also been reported in Mantodea,

OK094023 (https://www.ncbi.nlm.nih.gov/search/ all/?term=OK094023).

Funding: This study was funded by the Special funds for central government to guide local scientific and Technological Development (2020ZYD098) and National Natural Science Foundation of China (U21A20409). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Phthiraptera, and Psocoptera [9-13]. The changes of tRNAs are most various in no-typical insect mito-genomes, including tRNA translocation, tRNA loss, tRNA tandem duplication and tRNA conversion [4, 9-10, 14, 15]. Phthiraptera and Psocoptera insect mito-genomes may be most multipartite and most fast in evolution, that they could fission into fragmented mito-genomes and have numerous pseudo-genes and diverse gene rearrangements [7-10]. Thus, insect mitochondrial genomes are good models for researching mitochondrial genome evolution.

There are of 4,600 species of cockroaches [16] and 300 mitochondrial genomes of them have been sequenced and uploaded to the NCBI database. The blattarian mito-genomes seem to be conserved in evolution [16, 17]. The gene rearrangement was only reported in Cryptocercidae within blattarian mito-genomes [4]. *Episymploce splendens* belongs to Ectobiidae. Previous studies have mainly focused on the morphological features of *Episymploce* rather than molecular data [18, 19]. At present, no complete mitochondrial sequences of *Episymploce* have been recorded in the NCBI database. To obtain the sequence information and organization features of mito-genomes in *Episymploce*, we sequenced, annotated and described the complete mito-genome of *E. splendens*. Two tRNAs loss and a long intergenic region was found in *E. splendens* mito-genome and two pseudo-genes were also identified in the study. This study could deepen our understanding of mitochondrial genome of insects and contribute to the study of insect mito-genome evolution.

Materials and methods

Sample and DNA extraction

The cockroaches used in this study were collected on Mount Emei, Sichuan Province, China. The fresh material was placed in absolute ethanol and stored at -20°C. Total genomic DNA was extracted from the muscle tissue (legs) using TIANamp Genomic DNA kit (TIANGEN, Beijing, China). The concentration and purity of total DNA were detected by spectrophotometer. In addition, DNA was detected by agarose gel electrophoresis, and 1% agarose gel electrophoresis judged whether DNA was successfully extracted or not. Finally, DNA was stored at -20°C.

PCR amplification and sequencing

Primers were designed based on the conserved sequences of *Blattella germanica* and *Blattella bisignata*, and then specific primers were designed based on the amplified and sequenced sequences at both ends [20, 21]. The software Primer Premier 5.0 was used to designed primers and the primer details are listed in Table 1. Primers Es3, Es9 and Es10 were obtained from Xiao *et al* [10]. PCR was conducted using 2×Taq PCR Mix (Innovagene, Chengdu, China) and performed on a PTC-100 thermal cycler (BioRad, Hercules, CA) with the following cycling conditions: an initial denaturation for 5 min at 94°C, followed by 35 cycles of denaturation for 30s at 94°C, annealing for 30s at 50–62°C (depending on primer combinations), elongation for 1–4 min (depending on putative length of the fragments) at 72°C, and a final extension step of 72°C for 10 min. PCR products were estimated by 1.0% agarose gel electrophoresis and sent to Tsingke Biotechnology Company (Chengdu, China) for sequencing. All fragments were bidirectional sequences, and the unsuccessful sequenced fragments were redesigned with primers and sequenced again to complete the sequence.

The sequences which were transferred to nuclear DNA from mitochondrial genome were named Numts [22]. Primers that amplified mitochondrial genes may occasionally amplify Numts [23]. In this study, agarose gel electrophoresis of the amplification product (the primer Numt-1-1 and Numt-2-1, Table 1) appeared as two bands. We then redesigned primers

Primer name	Upstream primers sequences(5'-3')	Downstream primers sequences(5'-3')	Anneal temperature (°C)	Extension time (Second)
Es1	140-CCTCTCTTATCGCAATGTCCA	1609-CGTGGGAAAGCTATATCAGGA	58	90
Es2	1367-GGTCAACAAATCATAAAGATATTGG	2074-TAAACTTCAGGGTGACCAAAAAATCA	55	45
Es3	1559-CAACATTTATTTTGATTCTTTGG	3649-GTTTAAGAGACCACCACTTG	50	90
Es4	3401-GAAGACTTTCACCAACCATC	4391-TAGTACACTCATCTACTCTGGTAAC	52	60
Es5	4031-TGTAACAGCCCATGCTT	5797-AATCGCAATGATGGTAGG	51	110
Es6	4945-AGAAGACTTTCACCAACCAT	6997-GGATTCTCAAGATATTCGTT	51	120
Es7	6378-TCAACCGTTATCGAAAGACT	7328-CTCCTACTCCTGTATCTGCTT	52	60
Es8	7010-AACGAATATCCTGAGAATCC	8415-CACGGATTATGTTCTTCAGG	52	90
Es9	8290-GAAGGGGGTGCTGCTATATTAC	11151-ATTACTCCTCCTAATTTATTAGGAAT	62	180
Es10	10491-CAATGAGTATGAGGAGGATTTGCTGT	14755-TGTGCCAGCAGTCGCGGTTATACA	59.5	240
Es11	13822-CAGATTATATTGATTCGCACAAC	303-ATAGAACTGATGAAGCTAAGGC	55	75
Numt-1-1	GATTACGCTGTTATCCCTAAG	GGTGTAACTAGAATGATACAGGT	51	65
Numt-1-2	CGGTTTGAACTCAGATCATGTAAG	GAAGGTGTAACTAGAATGATACAGGT	57	78
Numt-2-1	TGCTACCTTTGCACGGTC	AGGTGAGATAAGTCGTAACATAGT	53	54
Numt-2-2	TAAACTCTATAGGGTCTTCTCG	CTAGAATGATACAGGTTAGGCT	53	70

Table 1. Primers for the PCR amplifications of *Episymploce splendens* (Es) mito-genome and Numts.

https://doi.org/10.1371/journal.pone.0268064.t001

(Numt-1-2 and Numt-2-2, <u>Table 1</u>) to amplify these regions and obtained the same result. These non-targeted bands were purified and sequenced, and obtained sequences belonging to mitochondria via the analysis of NCBI blast. Due to differences in mutation rates or freedom from selection pressure, they had some degree of base differences compared to corresponding mito-genome sequences.

Sequence analysis and annotation

DNA SeqMan program, which is included in the Lasergene software package (DNAStar Inc. Madison Wis), was used to assemble sequences to obtain the complete mitochondrial genome. PCGs and rRNAs were identified by comparing *E. splendens* to *B. germanica* and *B. bisignata* [20, 21]. Most of the tRNA and their secondary structure inferences were conducted using the online server ARWEN (http://mbio-serv2.mbioekol.lu.se/ARWEN/) [24]. The tRNA-Ile, tRNA-Phe and tRNA-Leu were not identified by ARWEN; they were manually checked by referring to secondary structural models of other blattaria insects. The mitogenomic map was depicted with SeqBuilder (http://www.dnastar.com). The A+T content of the nucleotide sequence and relative synonymous codon usage (RSCU) were calculated using MEGA 5.2 [25]. The AT skewness was calculated according to the following formula: AT skew = [A-T]/ [A+T], and the GC skewness was calculated according to the following formula: GC skew = [G-C]/[G+C] [26].

Divergence dating analysis

There were two Numts found in *E. splendens*, namely Numt-1 and Numt-2. Through sequence alignment by MEGA 5.2, Numt-1 corresponded to partial lrRNA, and Numt-2 was similar to partial lrRNA and its neighboring tRNA-Val of the *E. splendens* mito-genome. We performed divergence date analyses based on the aligned sequences of Blattodea, two mantises and two outgroups (S1 Table). The molecular clock was calibrated using three minimum age constraints based on cockroach fossils, as shown in Table 2. A relaxed molecular-clock model was used for this study with the program BEAST 1.6.1 [27]. Rate variation was modeled among branches using uncorrelated lognormal relaxed clocks [27]. A Yule speciation process was used for the tree prior and posterior distributions of parameters, including the tree, were

species	Age (Ma)	Calibration Group	Reference
Valditermes brennae	130.8	Cryptocercus + Isoptera	[29]
"Gyna" obesa	57.7	Blaberidae	[30]
Cratomastotermes wolfschwenningeri	113	termites	[31]

Table 2. Fossils used for estimation of divergence time of Numts in the analysis.

https://doi.org/10.1371/journal.pone.0268064.t002

estimated using MCMC sampling [28]. Two independent runs (each with 4 chains) of 100 million generations were sampled every 5000 generations based on the GTR model. The tree topology was then estimated using the combined sample from the last 50 million generations of each run.

Results

Genome content and organization

The mitochondrial genome of *E. splendens* was 15,802 bp in length with typical circular molecules. It contained 35 mitochondrial genes: 13 PCGs, 2 rRNAs (srRNA and lrRNA), 20 tRNAs and the A+T rich region (Fig 1, Table 3). Putative secondary structures of the 20 tRNAs were shown in S1 Fig. Similar to most species' mitochondria, the coding genes of *E. splendens* mitogenome are compact, with several genes overlapping. There were five overlaps totaling 21 bp with the two longest overlaps being 7 bp between *ATP8* and *ATP6*, and 7 bp between *NAD4* and *NAD4L*. There were 15 gene spacers totaling 190 bp within the entire mitochondrial genome. The longest spacer was 93 bp between tRNA-Val and srRNA, followed by 24 bp between tRNA-Ile and *NAD2* and 22 bp between *NAD1* and tRNA-Leu. The longest spacer had a similarity of 64.6% compared to its adjacent and corresponding srRNA (S2 Fig). There were 16 areas with neither gene overlap nor intergenic spacer. Additionally, there were 20 tRNAs in the mitochondrial genome of *E. splendens*, lacking two tRNAs usually located between tRNA-Ile and *NAD2*, tRNA-Gln and tRNA-Met (Fig 1).

Nucleotide composition and codon usage

We calculated the nucleotide composition of the mtDNA in *E. splendens* using MEGA5.2, refer to S2 Table for detailed results. The content of A+T (74.6%) was higher than G+C (25. 4%). It corresponded well to the AT bias generally observed in insect mito-genomes, which ranges from 69.5 to 84.9% [32, 33].

The high A+T content and nucleotide skewness of the mitochondrial genome were also reflected in the codon use of protein-coding genes. According to the relative synonymous codon usage (RSCU) value (<u>S3 Table, Fig 2</u>), the occurrence of synonymous codons ending in A or T was much higher than other synonymous codons, and they accounted for 89.7% (3,330) of the total codons. The third position was A or T for all codons with RSCU values greater than 1. The six most frequently used codons were also composed of A and T, namely TTT, TTA, ATT, ATA, TAT and AAT, accounting for 40.7% of the total number of codons.

Protein-coding genes

All PCGs of *E. splendens* used ATN as the start codon, except for *COX1*. The *COX1* gene in the *E. splendens* mito-genome used TTG as the starting codon, in agreement with other known cockroaches [5]. The stop codon was most commonly TAA in the *E. splendens* mito-genome, followed by TAG (*NAD1*). Four protein coding genes: *NAD4*, *NAD6*, *ATP6*, and *NAD3*, used incomplete TA as the stop codon.



Fig 1. Circular gene map of *Episymploce splendens* **mito-genome**. Genes coded in the J-strand are inside of the circle. Gene coded in the N-strand are outside of the circle. *COX1*, *COX2* and *COX3* refer to the cytochrome C oxidase subunits; *CytB* refers to cytochrome B; ATPase6 and ATPase8 refer to ATP synthase subunits 6 and 8 genes; and *NAD1-NAD6* and *NAD4L* refer to the NADH dehydrogenase subunit 1–6 and 4Lgenes.

https://doi.org/10.1371/journal.pone.0268064.g001

A+T rich region

The non-coding region of *E. splendens* mito-genome was located between srRNA and tRNA-Ile, and it was 290 bp long with 64.8% A+T content. Two 120 bp long repeated units separated by a 7 bp interval and four Poly-A structures were found in the A+T rich region of the *E. splendens* mito-genome (Fig 3).

Numts and its divergence time

Numts are originated from mt-genome. Numts have different mutation rates compared to their ancient mtDNA, but the pattern of their nucleotide substitution is similar to ancient mtDNA. So they can be called "fossil" markers. Numts can be used to solve some problems in

Gene	Grand	Location	Anticodon	Start codon	Stop codon	
tRNA-Ile	J	177	TAT			
NAD2	J	1021148		ATG	TAA	
tRNA-Trp	J	11491212	TCA			
tRNA-Cys	N	12081271	GCA			
tRNA-Tyr	N	12781347	GTA			
COX1	J	13502885		TTG	TAA	
tRNA-Leu(UUR)	J	28872954	TAA			
COX2	J	29563642		ATG	TAA	
tRNA-Lys	J	36453715	CTT			
tRNA-Asp	J	37173781	GTC			
ATPase8	J	37823940		ATT	TAA	
ATPase6	J	39344613		ATG	TA-	
COX3	J	46145402		ATG	TAA	
tRNA-Gly	J	54055468	TCC			
NAD3	J	54695821		ATG	TA-	
tRNA-Ala	J	58225886	TGC			
tRNA-Arg	J	58855950	TCG			
tRNA-Asn	J	59516020	GTT			
tRNA-Ser(AGN)	J	60216085	GCT			
tRNA-Glu	J	60886152	TTC			
tRNA-Phe	N	61626228	GAA			
NAD5	N	62297947		ATT	TAA	
tRNA-His	N	79638027	GTG			
NAD4	N	80289367		ATG	TA-	
NAD4L	N	93619642		ATG	TAA	
tRNA-Thr	N	96439709	TGT			
tRNA-Pro	N	97099774	TGG			
NAD6	J	977610275		ATT	TA-	
CytB	J	1027611409		ATG	TAA	
tRNA-Ser(UCN)	J	1140911477	TGA			
NAD1	N	1150012441		ATA	TAG	
tRNA-Leu(CUN)	N	1245112511	TAG			
lrRNA	N	1251213822				
tRNA-Val	N	1382313893	TAC			
srRNA	N	1398714792				
A+T-rich region		1479315082				

Table 3. Annotation of Episymploce splendens mito-genome.

'TA-' refer to incomplete stop codons.

https://doi.org/10.1371/journal.pone.0268064.t003

phylogeny, such as Zischler et al used the Numt as phylogenetic outgroup to prove the origin of man [34].

There were two Numts found in *E. splendens*, namely Numt-1 (557 bp) and Numt-2 (975 bp). Comparisons with aligned mitochondrial sequence showed 87.23% homologies in Numt-1 and 76.63% in Numt-2 (<u>S3</u> and <u>S4</u> Figs). Some characteristics in Numts such as the deletion mutation, base substitution and insertion mutation were also found in Numt-1 and Numt-2 [<u>35</u>, <u>36</u>].

The timescale for evolution of 25 species and Numt-1 diversification based on aligning sequences and calibrations based on three cockroach fossils is shown in Fig 4A while the timescale for Numt-2 is shown in Fig 4B. The divergence of the lineages leading to *Blattella* and



Fig 2. Relative Synonymous Codon Usage (RSCU) in *Episymploce splendens* **mito-genome.** A total of 3,711 codons for *E. splendens* mito-genome were analyzed, excluding stop codons. Leu, Leu*, Ser, and Ser* indicate trnL1 (CUN), trnL2 (UUR), trnS1 (AGN), and trnS2 (UCN), respectively.

https://doi.org/10.1371/journal.pone.0268064.g002

Episymploce was 22.60 Ma to 36.50 Ma (95% confidence interval [CI]) in Fig 4A while the estimated age of the split between them was 15.54 Ma to 34.27 Ma (95% confidence interval [CI]) in Fig 4B. Results were similar although different aligning sequences were used to calculate the divergence time in this study (Fig 4). Numt-1 transferred from the mitochondrion to the nucleus between 14.15 Ma to 22.35 Ma (95% confidence interval [CI]), and Numt-2 were estimated to have diverged between 19.19 Ma to 24.06 Ma (95% confidence interval [CI]).

Discussion

Intergenic spacer

Mito-genomes typically exhibited compact arrangements, such as small gene spacing, gene overlap, or incomplete stop codons. However, the *E. splendens* mito-genome had a long

two 120bp-long repeated segments, respectively. The red fonts represented the poly-A structures.

https://doi.org/10.1371/journal.pone.0268064.g003



Fig 4. a. Phylogenetic chronogram of Numt-1 and blattodean species based on the sequence aligned with Numt-1, reconstructed using BEAST. The best-fit evolution model was determined by PartitionFinder. Scale bar estimates age in millions of years and blue bars represent 95% highest posterior density intervals for the node ages. *Papilio protenor* and *Biston panterinaria* were employed to root the tree as outgroups. b. Phylogenetic chronogram of Numt-2 and blattodean species based on the sequence aligned with Numt-2, reconstructed using BEAST. The best-fit evolution model was determined by PartitionFinder. Scale bar estimates age in millions of years and blue bars represent 95% highest posterior density intervals for the node ages. *Papilio protenor* and *Biston panterinaria* were employed to root the tree as outgroups. *Papilio protenor* and *Biston panterinaria* were employed to root the tree as outgroups. *Papilio protenor* and *Biston panterinaria* were employed to root the tree as outgroups. *Papilio protenor* and *Biston panterinaria* were employed to root the tree as outgroups.

https://doi.org/10.1371/journal.pone.0268064.g004

intergenic spacer region (93 bp) located between tRNA-Val and srRNA, which is the longest intergenic spacer region in cockroach mito-genomes reported. The long intergenic spacers in mito-genomes have been reported in some Hymenopteran [37, 38], Hemipteran [39, 40], Dic-tyopteran [41] and Coleopteran insects [6, 42, 43]. There are two commonly posited evolutionary mechanisms for the origin of mitochondrial intergenic spacers, the duplication/random loss model and slipped-strand mispairing [6, 41]. We could not find a homologous sequence with both ends in this intergenic spacer, thus its formation is difficult to explain by slipped-strand mispairing [6, 44]. A similar long intergenic spacer was reported in a blattarian insect mito-genome, *Blaptica dubia* (71-bp between tRNA-Gln and tRNA-Met) [41], and the duplication/random loss model was used to explain the formation of this intergenic spacer [6, 41, 45, 46]. We suggested that this intergenic region may be derived from the replication of the 3' end of the srRNA when the DNA double helix unraveled, followed by random loss of partial duplicated gene, and then the residues formed the 93-bp remaining intergenic spacer in *E. splendens* (Fig 5).

Animal mito-genomes were generally considered to be economic and optimized for rapid replication and transcription [47]. Therefore, mitochondrial evolution had traditionally been regarded as favoring genome size reduction [3, 48, 49], possibly by eliminating intergenic spacers [50]. Eliminating nonfunctional intergenic spacers in mitochondrial evolution was important in the highly reduced and efficient mito-genomes [43]. But with the discovery of more large intergenic spacers in mito-genomes, as several containing additional origin of replication



the 93-bp remaining intergenic region

Fig 5. Putative formation mechanism of the intergenic spacer in *Episymploce splendens* **mito-genome.** The Randomly copied fragment was marked with *.

https://doi.org/10.1371/journal.pone.0268064.g005

(Apis mellifera, Triatoma dimidiata, Bombus ignitus) [33, 39, 51], tandem repeat units (Triatoma dimidiata, Pyrocoelia rufa) [39, 42], or even open reading frames retained (Triatoma *dimidiata*) [39], whether these long spacer regions were functional was controversial [39, 52]. We analyzed the large intergenic spacers in the mito-genomes of several insects, comparing the length and similarity of these large intergenic spacer sequences with their root sequences (S4 Table) [6, 40, 41]. We found these spacer sequences were usually shorter than their root sequences and the greater the length difference between these spacer and their root sequences, the lower the similarity between them. It implied that these mito-genomes lost partial duplicated nucleotides in mitochondrial evolution. In addition, large intergenic spacers found in several related species were not homologous, implying their origin occurred independently after species differentiation [6]. This suggests that these long intergenic spacers did not confer an evolutionary advantage, and they were gradually deleted during mitochondrial evolution. Consequently, we consider that the mito-genome may continue evolving towards compact arrangement and the long intergenic region may gradually decrease or even disappear during mito-genome evolution. However, these discoveries of large intergenic spacers in insect mitogenomes contribute to species identification, and also provide valuable information for the study of the evolution of insect mitochondrial genomes.

CAU (H)	AUU (I)	GenBank Accession number
1.69	1.86	OK094023
1.54	1.83	NC_012901.1
1.57	1.63	NC_018549.1
1.57	1.77	NC_030002.1
1.63	1.63	NC_030003.1
1.49	1.64	NC_035052.1
1.51	1.49	NC_037496.1
1.44	1.7	NC_034841.1
1.45	1.66	NC_034842.1
1.42	1.71	MG010455
1.55	1.54	NC_029224.1
1.36	1.71	NC_016956.1
1.45	1.42	NC_030001.1
1.43	1.73	NC_006076.1
1.35	1.6	NC_029225.1
1.21	1.57	NC_018132.1
1.33	1.6	NC_030191.1
0.93	1.64	NC_014274.1
	CAU (H) 1.69 1.54 1.57 1.57 1.63 1.49 1.51 1.44 1.45 1.42 1.55 1.36 1.43 1.35 1.21 1.33 0.93	CAU (H)AUU (I)1.691.861.541.831.571.631.571.771.631.631.491.641.511.491.441.71.451.661.421.711.551.541.361.711.451.661.361.711.351.61.211.571.331.60.931.64

Table 4. The RSCU values of CAU (H) and AUU (I) in known cockroaches.

https://doi.org/10.1371/journal.pone.0268064.t004

tRNAs deletion

Mitochondrial gene content, arrangement and composition were highly conserved, and mitogenomes typically contained 37 coding genes [3, 53]. However, some exceptions, such as gene duplications, deletions or rearrangements, were found in some species [54–56]. Although tRNA deletion is unusual, increasing cases have been reported, such as three amphibians species [54, 57], three reptile species [58, 59], one crustacean species [60], one Hemiptera insect and one Coleoptera insect [61], four Psocoptera insects [10, 62] and three Mantodea insects [9]. In this study, two tRNAs, tRNA-Gln and tRNA-Met, were absent in *E. splendens* mitogenome. The tRNA deletions was only found in the *E. splendens* mito-genome (this study) in all Blattaria mito-genomes reported and the tRNA deletion events reported in previous research also scattered in different clade branches, therefore, we consider tRNA deletions appear to be separate events occasionally occurring in some species or evolutionary branches. The deletion of functional genes was obviously disadvantageous for species, and the mechanism of deletion is still unclear.

The organism may have a functional replacement to cope with the loss of tRNAs. Two mechanisms were proposed for this functional compensation. The first mechanism where tRNAs from the cytosol are imported into mitochondria has been confirmed, as aminoacyl-tRNA synthetases being imported from the cytosol into mitochondria [63], and functional tRNA-Lys encoded in the nuclear genome being imported into marsupial mitochondria [64]. The second compensation mechanism for the missing tRNAs is 'superwobble', where a tRNA with an unmodified U in the wobble position reads all four nucleotides in the third codon position [65]. In our study, the loss of tRNA-Gln and tRNA-Met can be compensated with their first and second codons matching His and Ile, respectively, and the anticodon swing site U for His and Ile. We calculated the Relative Synonymous Codon Usage (RSCU) values of CAU (H) and AUU (I) in reported cockroach mito-genomes, and discovered that *E. splendens* mito-genome had the highest RSCU values (Table 4), indicating that *E. splendens* might be compensated through tRNA superwobble. Regardless, we cannot exclude tRNA import from

the cytosol in *E. splendens* mitochondrion. Therefore, more studies are needed to confirm the compensation mechanism for the absence of tRNAs in the mito-genome of *E. splendens*.

Conclusion

In this study, we sequenced and annotated the *E. splendens* mito-genome. Two tRNAs (tRNA-Gln, tRNA-Met) were lost and a long intergenic region between tRNA-Val and srRNA (93 bp, with a 64.6% similarity with its corresponding srRNA) was also found in *E. splendens* mito-genome. The duplication/random loss model may account for the origin of this long intergenic spacer. We also found two Numts, Numt-1 and Numt-2 transfered from mitochondrion to nucleus at about 14.15 Ma to 22.35 Ma, and 19.19 Ma to 24.06 Ma respectively.

Supporting information

S1 Table. GenBank accession numbers for divergence date analyses. (DOCX)

S2 Table. Nucleotide composition in different regions of *Episymploce splendens*. (DOCX)

S3 Table. Relative Synonymous Codon Usage (RSCU) for PCGs of *Episymploce splendens*. (DOCX)

S4 Table. The length and similarity of these large intergenic spacers with their original genes in some insects. The similarity was calculated by DNAMAN. (DOCX)

S1 Fig. Putative secondary structures of the 20 tRNA genes identified in *Episymploce splendens*.

(TIF)

S2 Fig. Aligement result of the long intergenic spacer and srRNA of *Episymploce splendens* by MEGA.

(TIF)

S3 Fig. Similarity comparison results of Numt-1 and mito-genome of *Episymploce splendens* by DNAMAN.

(TIF)

S4 Fig. Similarity comparison results of Numt-2 and mito-genome of *Episymploce splendens* by DNAMAN. (TIF)

Acknowledgments

We sincerely appreciate Natural History Museum of Sichuan University for the sample collection. Thanks Dr. Megan for language help.

Author Contributions

Conceptualization: Jinnan Ma, Xiuyue Zhang.

Funding acquisition: Xiuyue Zhang.

Investigation: Lin Yan, Zhenzhen Hou, Hongmei Wang, Jie Gao, Chenjuan Zeng, Qin Chen.

Methodology: Jinnan Ma.

Project administration: Hongmei Wang.

Resources: Lin Yan, Jie Gao, Chenjuan Zeng, Bisong Yue.

Software: Jinnan Ma.

Writing - original draft: Lin Yan, Zhenzhen Hou, Hongmei Wang, Jie Gao, Qin Chen.

Writing - review & editing: Lin Yan, Bisong Yue, Xiuyue Zhang.

References

- Wilson AJ, Xu J. Mitochondrial inheritance: diverse patterns and mechanisms with an emphasis on fungi. Mycology. 2012; 3(2): 158–166. https://doi.org/10.1080/21501203.2012.684361 PMID: 7219536.
- Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria. Nature. 1981; 290(5806): 470–474. https://doi.org/10.1038/290470a0 PMID: 7219536.
- Boore JL. Animal mitochondrial genomes. Nucleic Acids Res. 1999; 27(8): 1767–1780. https://doi.org/ 10.1093/nar/27.8.1767 PMID: 10101183.
- Li WJ, Wang ZQ, Che YL, The Complete Mitogenome of the Wood-Feeding Cockroach Cryptocercus meridianus (Blattodea: Cryptocercidae) and Its Phylogenetic Relationship among Cockroach Families. Int. J. Mol. Sci. 2017; 18, 2397. <u>https://doi.org/10.3390/ijms18112397</u> PMID: <u>29137151</u>. PMCID: PMC5713365
- Gong R, Guo X, Ma J, Song X, Shen Y, Geng F, et al. Complete mitochondrial genome of *Periplaneta brunnea* (Blattodea: Blattidae) and phylogenetic analyses within Blattodea. J Asia-Pac Entomol. 2018; 21(3): 885–895. https://doi.org/10.1016/j.aspen.2018.05.006
- Du C, Zhang L, Lu T, Ma J, Chen Z, Yue B, et al. Mitochondrial genomes of blister beetles (Coleoptera, Meloidae) and two large intergenic spacers in *Hycleus* genera. BMC Genomics. 2017; 18(1): 698. https://doi.org/10.1186/s12864-017-4102-y PMID: 28874137.
- Simon S, Hadrys H. A comparative analysis of complete mitochondrial genomes among Hexapoda. Molecular Phylogenetics and Evolution. 2013; 69(2): 393–403. https://doi.org/10.1016/j.ympev.2013. 03.033 PMID: 23598069
- Chen ZT, Du YZ. Rearrangement of mitochondrial genome in insects. Journal of Environmental Entomology. 2016; 38(4): 843–851. (In Chinese)
- Shi Y, Li LY, Liu QP, Ali MY, Yuan ZL, Guy S et al. Complete mitochondrial genomes of four species of praying mantises (Dictyoptera, Mantidae) with ribosomal second structure, evolutionary and phylogenetic analyses. PLoS ONE. 2021; 16(11):e0254914. https://doi.org/10.1371/journal.pone.0254914 PMID: 34735444. PMCID: PMC8568281.
- Shi Y, Chu Q, Wei DD, Qiu YJ, Shang F, Dou W, et al. The mitochondrial genome of booklouse, Liposcelis sculptilis (Psocoptera: Liposcelididae) and the evolutionary timescale of *Liposcelis*. Scientific RepoRts. 2016; 6:30660. https://doi.org/10.1038/srep30660 PMID: 27470659. PMCID: PMC4965752
- Shao RF, Zhu XQ, Barker SC, Herd K. Evolution of extensively fragmented mitochondrial genomes in the lice of humans. Genome Biol Evol. 2012; 4(11): 1088–1101. <u>https://doi.org/10.1093/gbe/evs088</u> PMID: 23042553. PMCID: PMC3514963.
- Chen SC, Wei DD, Shao RF, Shi JX, Dou W, Wang JJ. Evolution of multipartite mitochondrial genomes in the booklice of the genus Liposcelis (Psocoptera). BMC Genomics. 2014; 15(1):861. <u>https://doi.org/ 10.1186/1471-2164-15-861</u> PMID: 25282613. PMCID: PMC4197233.
- Zhang LP, Ma Y, Yu DN, Storey KB, Zhang JY. The mitochondrial genomes of *Statilia maculata* and *S. nemoralis* (Mantidae: Mantinae) with different duplications of *trnR* genes. Int J Biol Macromol. 2019; 121:839–845. https://doi.org/10.1016/j.ijbiomac.2018.10.038 PMID: 30340009
- Ye F, Lan XE, Zhu WB, You P. Mitochondrial genomes of praying mantises (Dictyoptera, Mantodea): rearrangement, duplication, and reassignment of tRNA genes. Scientific RepoRts. 2016;. 6:25634. https://doi.org/10.1038/srep25634 PMID: 27157299. PMCID: PMC4860592.
- Wang SS, Hou FX, Cao J, Peng C, Guo JL. The complete mitochondrial genome of the Statilia maculate. (Mantodea: Mantidae). Mitochondrial DNA Part B Resources. 2016; 1(1):860–861. <u>https://doi.org/ 10.1080/23802359.2016.1250134 PMID: 33490423</u>. PMCID: PMC7800993.
- 16. Ma JN, Du C, Zhou C, Shen YM, Fan ZX, Yue BS, et al. Complete mitochondrial genomes of two blattid cockroaches, *Periplaneta australasiae* and *Neostylopyga rhombifolia*, and phylogenetic relationships

within the Blattaria. PLoS One. 2017; 12(5): e0177162. https://doi.org/10.1371/journal.pone.0177162 PMID: 28486518.

- Bourguignon T, Tang Q, Ho Simon Y.W., Juna F, Wang ZQ, Arab D A., et al. Transoceanic Dispersal and Plate Tectonics Shaped GlobalCockroach Distributions: Evidence from Mitochondrial Phylogenomics. Molecular biology and evolution. 2018; 35(4):970–983. https://doi.org/10.1093/molbev/ msy013 PMID: 29420807.
- Li TT, Liu DX, Qiu DY, Yue QY. Two new species of Episymploce Bey-Bienko, 1950 (Blattodea, Ectobiidae, Blattellinae) from China. Zookeys. 2020; 954:31–45. <u>https://doi.org/10.3897/zookeys.954.49738</u> PMID: 32821203. PMCID: PMC7406546.
- Wang Z, Che Y. Three new species of cockroach genus Symploce Hebard, 1916 (Blattodea, Ectobiidae, Blattellinae) with redescriptions of two known species based on types from Mainland China. Zookeys. 2013; 337(337):1–18. https://doi.org/10.3897/zookeys.337.5770 PMID: 24146575.
- Xiao B, Chen AH, Zhang YY, Jiang GF, Hu CC, Zhu CD. Complete mitochondrial genomes of two cockroaches, *Blattella germanica* and *Periplaneta americana*, and the phylogenetic position of termites. Curr Genet. 2012; 58(2):65–77. https://doi.org/10.1007/s00294-012-0365-7 PMID: 22311390.
- **21.** Chen AH. Complete mitochondrial genome of the double-striped cockroach *Blattella bisignata* (Insecta: Blattaria: Blaberoidea). Mitochondrial DNA. 2013; 24(1):14–6. https://doi.org/10.3109/19401736.2012. 710228 PMID: 22897805.
- Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien SJ. Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. J Mol Evol. 1994; 39(2):174–90. https:// doi.org/10.1007/BF00163806 PMID: 7932781.
- Arctander P. Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. Proc Biol Sci. 1995; 262(1363):13–9. https://doi.org/10.1098/rspb.1995.0170 PMID: 7479989.
- Laslett D, Canbäck B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics. 2008; 24(2):172–175. https://doi.org/10.1093/bioinformatics/btm573 PMID: 18033792.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011; 28(10): 2731–2739. <u>https://doi.org/10.1093/molbev/msr121</u> PMID: 21546353. PMCID: PMC3203626.
- Perna NT, Kocher TD. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol. 1995; 41(3):353–358. https://doi.org/10.1007/BF00186547 PMID: 7563121.
- Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007; 7:214. https://doi.org/10.1186/1471-2148-7-214 PMID: 17996036
- Gernhard T. The conditioned reconstructed process. J Theor Biol. 2008; 253(4):769–778. <u>https://doi.org/10.1016/i.itbi.2008.04.005</u> PMID: 18538793.
- Wolfe JM, Daley AC, Legg DA, Edgecombe GD. Fossil calibrations for the arthropod Tree of Life. Earth-Science Reviews. 2016; 160:43–110.
- Evangelista DA, Wipfler B, Bethoux O, Donath A, Fujita M, Kohli MK, et al. An integrative phylogenomic approach illuminates the evolutionary history of cockroaches and termites (Blattodea). Proceedings of the Royal Society. Biological sciences. 2019; 286(1895). <u>https://doi.org/10.1098/rspb.2018.2076</u> PMID: 30963947
- 31. Vrs ansky P. Origin and the early evolution of mantises. AMBA Projekty. 2002; 6(1):1-16.
- Dotson EM, Beard CB. Sequence and organization of the mitochondrial genome of the Chagas disease vector, *Triatoma dimidiata*. Insect Mol Biol. 2001; 10(3):205–215. https://doi.org/10.1046/j.1365-2583. 2001.00258.x PMID: 11437912.
- Crozier RH, Crozier YC. The mitochondrial genome of the honeybee Apis mellifera: complete sequence and genome organization. Genetics. 1993; 133(1):97–117. https://doi.org/10.1093/genetics/133.1.97 PMID: 8417993. PMCID: PMC1205303.
- Zischler H, Geisert H, von Haeseler A, Pääbo S. A nuclear 'fossil' of the mitochondrial D-loop and the origin of modern humans. Nature.1995; 378(6556):489–492. https://doi.org/10.1038/378489a0 PMID: 7477404
- Tourmen Y, Baris O, Dessen P, Jacques C, Malthièry Y, Reynier P. Structure and chromosomal distribution of human mitochondrial pseudogenes. Genomics. 2002; 80(1):71–77. https://doi.org/10.1006/ geno.2002.6798 PMID: 12079285.
- Hazkani-Covo E, Sorek R, Graur D. Evolutionary dynamics of large numts in the human genome: rarity of independent insertions and abundance of post-insertion duplications. J Mol Evol. 2003; 56(2):169– 174. https://doi.org/10.1007/s00239-002-2390-5 PMID: 12574863.

- Wei SJ, Tang P, Zheng LH, Shi M, Chen XX. The complete mitochondrial genome of Evania *appendigaster* (Hymenoptera: Evaniidae) has low A+T content and a long intergenic spacer between atp8 and atp6. Mol Biol Rep. 2010; 37(4):1931–1942. https://doi.org/10.1007/s11033-009-9640-1 PMID: 19655273.
- Rodovalho CM, Lyra ML, Ferro M, Jr MB. The mitochondrial genome of the leaf-cutter ant Atta laevigata: a mitogenome with a large number of intergenic spacers. PLoS One. 2014; 9(5):e97117. https:// doi.org/10.1371/journal.pone.0097117 PMID: 24828084.
- Dotson EM, Beard CB. Sequence and organization of the mitochondrial genome of the Chagas disease vector, *Triatoma dimidiata*. Insect Mol Biol. 2001; 10(3):205–215. <u>https://doi.org/10.1046/j.1365-2583</u>. 2001.00258.x PMID: 11437912.
- 40. Hua J, Li M, Dong P, Cui Y, Xie Q, Bu W. Comparative and phylogenomic studies on the mitochondrial genomes of Pentatomomorpha (Insecta: Hemiptera: Heteroptera). BMC Genomics. 2008; 9(1):1–15. https://doi.org/10.1186/1471-2164-9-610 PMID: 19091056. PMCID: PMC2651891.
- Cheng XF, Zhang LP, Yu DN, Storey KB, Zhang JY. The complete mitochondrial genomes of four cockroaches (Insecta: Blattodea) and phylogenetic analyses within cockroaches. Gene. 2016; 586(1):115– 122. https://doi.org/10.1016/j.gene.2016.03.057 PMID: 27045773.
- **42.** Bae JS, Kim I, Sohn HD, Jin BR. The mitochondrial genome of the firefly, *Pyrocoelia rufa*: complete DNA sequence, genome organization, and phylogenetic analysis with other insects. Mol Phylogenet Evol. 2004; 32(3):978–985. https://doi.org/10.1016/j.ympev.2004.03.009 PMID: 15288070.
- Sheffield NC, Song H, Cameron SL, Whiting MF. A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. Mol Biol Evol. 2008; 25 (11):2499–2509. https://doi.org/10.1093/molbev/msn198 PMID: 18779259.
- Stanton DJ, Daehler LL, Moritz CC, Brown WM. Sequences with the potential stem-and-loop structures are associated with coding-region duplications in animal mitochondrial DNA. Genetics. 1994; 137 (1):233–241. https://doi.org/10.1093/genetics/137.1.233 PMID: 8056313.
- Moritz C, Dowling TE, Brown WM. Evolution of Animal Mitochondrial DNA: Relevance for Population Biology and Systematics. Annu Rev Ecol Evol S. 1987; 18:269–292. <u>https://doi.org/10.1146/annurev.ecolsys.18.1.269</u>
- Boore JL. Comparative Genomics: The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of Deuterostome animals. 1st ed. Dordrecht: Springer; 2000.
- Sayadi A, Immonen E, Tellgren-Roth C, Arnqvist G. The Evolution of Dark Matter in the Mitogenome of Seed Beetles. Genome Biol Evol. 2017; 9(10): 2697–2706. https://doi.org/10.1093/gbe/evx205 PMID: 29048527. PMCID: PMC5737749.
- Macey JR, Larson A, Ananjeva NB, Fang Z, Papenfuss TJ. Two novel gene orders and the role of lightstrand replication in rearrangement of the vertebrate mitochondrial genome. Mol Biol Evol. 1997; 14 (1):91–104. https://doi.org/10.1093/oxfordjournals.molbev.a025706 PMID: 9000757.
- **49.** McKnight ML, Shaffer HB. Large, rapidly evolving intergenic spacers in the mitochondrial DNA of the salamander family Ambystomatidae (Amphibia: Caudata). Mol Biol Evol. 1997; 14(11):1167–1176. https://doi.org/10.1093/oxfordjournals.molbev.a025726 PMID: 9364774.
- Burger G, Gray MW, Lang BF. Mitochondrial genomes: anything goes. Trends Genet. 2003; 19 (12):709–716. https://doi.org/10.1016/j.tig.2003.10.012 PMID: 14642752.
- Cha SY, Yoon HJ, Lee EM, Yoon MH, Hwang JS, Jin BR, et al. The complete nucleotide sequence and gene organization of the mitochondrial genome of the bumblebee, *Bombus ignitus* (Hymenoptera: Apidae). Gene. 2007; 392(1–2):206–220. https://doi.org/10.1016/j.gene.2006.12.031 PMID: 17321076.
- Jørgensen TE, Bakke I, Ursvik A, Andreassen M, Moum T, Johansen SD. An evolutionary preserved intergenic spacer in gadiform mitogenomes generates a long noncoding RNA. BMC Evol Biol. 2014; 14:182. https://doi.org/10.1186/s12862-014-0182-3 PMID: 25145347. PMCID: PMC4236577.
- Gemmell NJ, Janke A, Western PS, Watson JM, Pääbo S, Graves JA. Cloning and characterization of the platypus mitochondrial genome. J Mol Evol. 1994; 39(2):200–205. <u>https://doi.org/10.1007/</u> BF00163808 PMID: 7932783.
- Zhang JF, Nie LW, Wang Y, Hu LL. The complete mitochondrial genome of the large-headed frog, *Limnonectes bannaensis* (Amphibia: Anura), and a novel gene organization in the vertebrate mtDNA. Gene. 2009; 442(1–2):119–127. https://doi.org/10.1016/j.gene.2009.04.018 Epub 2009 May 3. PMID: 19397958.
- Hlaing T, Tun-Lin W, 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 Genet. 2009; 10:11. <u>https://doi.org/10.1186/1471-2156-10-11</u> PMID: <u>19267896</u>. PMCID: PMC2660364.

- von Nickisch-Rosenegk M, Brown WM, Boore JL. Complete sequence of the mitochondrial genome of the tapeworm *Hymenolepis diminuta*: gene arrangements indicate that Platyhelminths are Eutrochozoans. Mol Biol Evol. 2001; 18(5):721–30. https://doi.org/10.1093/oxfordjournals.molbev.a003854 PMID: 11319256.
- San MD, Gower DJ, Oommen OV, Wilkinson M, Zardoya R. Phylogeny of caecilian amphibians (Gymnophiona) based on complete mitochondrial genomes and nuclear RAG1. Mol Phylogenet. 2004; 33 (2):413–427. https://doi.org/10.1016/j.ympev.2004.05.014 PMID: 15336675
- Peng QL, Nie LW, Pu YG. Complete mitochondrial genome of Chinese big-headed turtle, *Platysternon megacephalum*, with a novel gene organization in vertebrate mtDNA. Gene. 2006; 380(1):14–20. https://doi.org/10.1016/j.gene.2006.04.001 PMID: 16842936.
- 59. Kumazawa Y, Miura S, Yamada C, Hashiguchi Y. Gene rearrangements in gekkonid mitochondrial genomes with shuffling, loss, and reassignment of tRNA genes. BMC Genomics. 2014; 15:1–13. https://doi.org/10.1186/1471-2164-15-930 PMID: 25344428. PMCID: PMC4223735
- Kilpert F, Podsiadlowski L. The complete mitochondrial genome of the common sea slater, Ligia oceanica (Crustacea, Isopoda) bears a novel gene order and unusual control region features. BMC Genomics. 2006; 7:241. https://doi.org/10.1186/1471-2164-7-241 PMID: 16987408. PMCID: PMC1590035.
- Dai Y, Li L, Jiang P, Song F, Ye Z, Yuan X, et al. Sequence and organization of the mitochondrial genome of an urostylidid bug, *Urochela quadrinotata* Reuter (Hemiptera: Urostylididae). Entomotaxonomia. 2012; 34, 613–623.
- Wei DD, Shao R, Yuan ML, Dou W, Barker SC, Wang JJ. The multipartite mitochondrial genome of Liposcelis bostrychophila: insights into the evolution of mitochondrial genomes in bilateral animals. PLoS One. 2012; 7(3):e33973. <u>https://doi.org/10.1371/journal.pone.0033973</u> PMID: 22479490. PMCID: PMC3316519.
- Duchêne AM, Pujol C, Maréchal-Drouard L. Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. Curr Genet. 2009; 55(1):1–18. <u>https://doi.org/10.1007/s00294-008-0223-9</u> PMID: 19083240.
- Dorner M, Altmann M, Paabo S, Morl M. Evidence for import of a lysyl-tRNA into marsupial mitochondria. Mol Biol Cell. 2001; 12(9):2688–2698. https://doi.org/10.1091/mbc.12.9.2688 PMID: 11553708
- Rogalski M, Karcher D, Bock R. Superwobbling facilitates translation with reduced tRNA sets. Nat Struct Mol Biol. 2008; 15(2):192–8. https://doi.org/10.1038/nsmb.1370 PMID: 18193063.