Gene 519 (2013) 326–334

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Characterization of the complete mitochondrial genome of *Bombyx mori* strain H9 (Lepidoptera: Bombycidae)

Li-Shang Dai [1], Bao-Jian Zhu [1], Qiu-Ning Liu, Guo-Qing Wei, Chao-Liang Liu *

College of Life Science, Anhui Agricultural University, 130 Changjiang West Road 230036, PR China

ARTICLE INFO

Article history:
Accepted 6 February 2013
Available online 20 February 2013

Keywords:
Mitochondrial genome
Bombyx mori H9
Lepidoptera
Bombycidae
Phylogenetic analysis

ABSTRACT

The complete mitochondrial genome (mitogenome) of *Bombyx mori* strain H9 (Lepidoptera: Bombycidae) is 15,670 base pairs (bp) in length, encoding 13 protein-coding genes (PCGs), two rRNA genes, 22 tRNA genes and a control region. The nucleotide composition of the genome is highly A + T biased, accounting for 81.31%, with a slightly positive AT skewness (0.059). The arrangement of 13 PCGs is similar to that of other sequenced lepidopterans. All the PCGs are initiated by ATN codons, except for the cytochrome c oxidase subunit 1 (*cox1*) gene, which is proposed by the TTAG sequence as observed in other lepidopterans. Unlike the other PCGs, the *cox1* and cytochrome c oxidase subunit 2 (*cox2*) genes have incomplete stop codons consisting of just a T. All tRNAs have typical structures of insect mitochondrial tRNAs, which is different from other sequenced lepidopterans. The structure of A + T-rich region is similar to that of other sequenced lepidopterans, including non-repetitive sequences, the ATAGA binding domain, a 18 bp poly-T stretch and a poly-A element upstream of transfer RNA M (*trnM*) gene. Phylogenetic analysis shows that the domesticated silkmoth *B. mori* originated from the Chinese *Bombyx mandarina*.

1. Introduction

The mitochondrial genome (mitogenome) forms a unit of genetic information and evolves independently from the nuclear genome (Boore, 1999). Mitochondrial DNA (mtDNA) can be used as molecular marker for evolutionary studies since it is maternally inherited DNA with a rapid evolutionary rate and lacks genetic recombination. The insect mitogenome is generally ranged from 14 kilobases to 20 kilobases (kb) in length and has a remarkably conserved set of 37 genes, including 13 protein-coding genes (PCGs); ATPase subunits 6 and 8 (*atp6* and *atp8*); cytochrome c oxidase subunits1–3 (*cox1–cox3*), cytochrome b (*cob*); NADH dehydrogenase subunits 1–6 and 4L (*nad1–nad6* and *nad4L*), ribosomal RNA genes L and S (*rrnL* and *rrnS*), 22 transfer RNA (*trn*) genes and a variable A + T-rich region containing some initiation sites for transcription and replication of the genome (Wolstenholme, 1992).

To date, more than 240 species of the complete or nearly complete mitogenome sequences from insects have been determined based on the GenBank database (http://www.ncbi.nlm.nih.gov/). Although the Lepidoptera is one of the largest groups of insects, only about 30 complete or near-complete mitogenomes are currently available in GenBank. The complete mitogenomes of three species of Bombycidae and six species of Saturniidae had been sequenced (Cameron and Whiting, 2008; Jiang et al., 2009; Kim et al., 2008, 2011; Liu et al., 2008, 2012; Pan et al., 2008; Yukuhiro et al., 2002). Sequence analysis of the other lepidopteran mitogenomes will provide further insight into our understanding of evolutionary relationships among these species. In this study, we determined the complete mitogenome sequence of *Bombyx mori* strain H9, and compared it with other lepidopterans. The phylogenetic analysis was performed using neighbor-joining (NJ) method based on the mitogenome sequences from various species.

2. Materials and methods

2.1. Insect and mtDNA extraction

Silkworm H9 strain maintained in the Key Laboratory of Sericultural Biotechnology, Anhui Agricultural University (Hefei, China) was used in this experiment. Whole genomic DNA was extracted from single specimens using the Aidlab Genomic DNA Extraction Kit (Aidlab Co., Beijing, China) according to the manufacturer's instructions.

2.2. Primer design, PCR, cloning, and sequencing

According to the known mitogenomes of lepidopterans, nine pairs of primers (Table 1) were designed and synthesized (Beijing

Abbreviations: PCR, polymerase chain reaction; ATPase, adenosine triphosphatase; NADH, nicotinamide adenine dinucleotide hydrogen; dNTP, deoxyribonucleoside triphosphate.

* Corresponding author. Tel./fax: +86 551 5786201.
E-mail address: cyschx@163.com (C.-L. Liu).

[1] These authors contributed to this work equally.

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.02.002

Table 1
The primers used in this study.

Primer pair	Primer sequence (5′→3′)
1F	GCTTTTGGGCTCATACCTCA
1R	GATGAAATACCTGCAAGATGAAG
2F	TGGAGCAGGAACAGGATGAAC
2R	GAGACCADTACTTGCTTTCAG
3F	ATTTGTGGAGCTAATCATAG
3R	GGTCAGGGACTATAATCTAC
4F	TCGACCTGGAACTTTAGC
4R	GCAGCTATAGCCGCTCCTACT
5F	TAAAGCAGAAACAGGAGTAG
5R	ATTGCGATATTATTTCTTTTG
6F	ACATTCTTAGGTGGATTA
6R	GTTAAAGTGGCATTATCT
7F	GGAGCTTCTACATGAGCTTTTGG
7R	GTTTGCGACCTCGATGTTG
8F	GGTCCCTTACGAATTTGAATATATCCT
8R	AAACTAGGATTAGATACCCTATTAT
9F	CTCTACTTTGTTACGACTTATT
9R	TCTAGGCCAATTCAACAACC

Sunbiotech Co., Ltd., Beijing, China). The PCR were carried out in a 50 μl mixture containing 1 U of Aidlab long taq (Aidlab Co., Beijing, China), 1 μl (about 15 ng) DNA, 5 μl 10× long Taq buffer (Mg^{2+} plus), 200 μM dNTPs, and 10 pmol each primer. PCRs were done under the following conditions: 4 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 45 s at 48–60 °C, 1–3 min at 72 °C and a final extension at 72 °C for 10 min. The PCR products were analyzed by agarose gel electrophoresis (1% w/v) and purified using a DNA gel extraction kit (Aidlab Co., Beijing, China). The purified PCR fragments were ligated into the T-vector (TaKaRa Co., Dalian, China) and sequenced at least three times (Invitrogen Co., Ltd., Shanghai, China).

2.3. Sequence assembly and gene annotation

Sequence annotation was performed using the DNAStar package (DNAStar Inc. Madison, USA). The nucleotide sequences of the PCGs were translated on the basis of the invertebrate mtDNA genetic code. The PCG sequences were initially identified using BLAST searches in GenBank and compared with the other mitogenome sequences of lepidopterans using ClustalX version 2.0 (Larkin et al., 2007). The base composition of nucleotide sequences was described by skewness and measured according to the formulas (AT skew = [A − T]/[A + T], GC skew = [G − C]/[G + C]) (Junqueira et al., 2004). Identification of tRNA genes was verified using the tRNAscan-SE program software available online at http://lowelab.ucsc.edu/tRNAscan-SE/ (Lowe and Eddy, 1997). The secondary structures of tRNA genes were analyzed by comparison with the nucleotide sequences of other insect tRNA sequences. The entire A + T-rich region was subjected to a search for the tandem repeats using the Tandem Repeats Finder program (http://tandem.bu.edu/trf/trf.html) (Benson, 1999).

2.4. Phylogenetic analysis

The complete nucleotide sequences of other 24 lepidopterans (Table 2) were downloaded from GenBank to illustrate the phylogenetic relationship among lepidopterans. The mitogenomes of *Locusta migratoria* (JN858212) (Ma et al., 2012), *Drosophila yakuba* (NC_001322) (Clary and Wolstenholme, 1985) and *Anopheles gambiae* (NC_002084) (Beard et al., 1993) were used as outgroups. Alignment of amino acid sequences of each of individual 13 PCGs was performed through Clustal X (Thompson et al., 1997) and the phylogenetic analysis was carried out using neighbor-joining (NJ) method with MEGA version 5.0 program (Tamura et al., 2011).

Table 2
The complete mitochondrial genome of Lepidoptera.

Species	Length (bp)	Accession number	References
Bombyx mori H9	15,670		This study
B. mori C108	15,656	AB070264	Yukuhiro et al. (2002)
B. mori Xiafang	15,664	AY048187	Lu et al. (2002)
Chinese *B. mandarina*	15,682	AY301620	Pan et al. (2008)
Japanese *B. mandarina*	15,928	NC_003395	Yukuhiro et al. (2002)
Actias selene	15,236	NC_018133	Liu et al. (2012)
Caligula boisduvalii	15,360	NC_010613	Hong et al. (2008)
Eriogyna pyretorum	15,327	FJ685653	Jiang et al., (2009)
Antheraea pernyi	15,575	AY242996	Liu et al. (2008)
Antheraea yamamai	15,338	EU726630	Kim et al. (2008)
Samia cynthia ricini	15,366	JN215366	Kim et al. (2012)
Helicoverpa armigera	15,374	GU188273	Yin et al. (2010)
Manduca sexta	15,516	EU286785	Cameron and Whiting (2008)
Adoxophyes honmai	15,680	DQ073916	Lee et al. (2006)
Coreana raphaelis	15,314	NC_007976	Kim et al. (2006)
Artogeia melete	15,140	EU597124	Hong et al. (2009)
Hipparchia autonoe	15,489	GQ868707	Kim and Kim (2010)
Acraea issoria	15,245	GQ376195	Hu et al. (2010)
Hyphantria cunea	15,481	GU592049	Liao et al. (2010)
Diatraea saccharalis	15,490	FJ240227	Li et al. (2011)
Spilonota lechriaspis	15,368	HM204705	Zhao et al. (2011)
Pieris rapae	15,157	HM156697	Mao et al. (2010)
Fabriciana nerippe	15,140	JF504707	Kim et al. (2011)
Grapholita molesta	15.717	HQ392511	Son and Kim (2011)
Chilo suppressalis	15,456	HQ860290	Yin et al. (2011)
G. molesta	15,776	HQ116416	Gong et al., (2011)
Helicoverpa armigera	15,347	GU188273	Yin et al. (2010)
Ochrogaster lunifer	15,593	AM946601	Salvato et al. (2008)
Phthonandria atrilineata	15,499	EU569764	Yang et al. (2009)

3. Results and discussion

3.1. Genome structure, organization and composition

The complete mitogenome of *B. mori* strain H9 is found to be circular molecules with 15,670 bp in size (Fig. 1), which is well within the range detected in the completely sequenced lepidopterans (Table 2). It presents the typical gene content observed in metazoan mitogenomes (Table 3), including 13 PCGs (*cox1–3*, *nad1–6*, *nad4L*, *cob*, *atp6* and *atp8*), 22 tRNA genes, two mitochondrial ribosomal RNAs (*rrnS* and *rrnL*), and a major non-coding region known as the "A + T-rich region" detected in other insects (Fig. 1). The order of genes and the orientation of the *B. mori* strain H9 are identical to those of sequenced lepidopteran mitogenomes (Liu et al., 2012; Yukuhiro et al., 2002). The nucleotide composition of the mitogenome of *B. mori* strain H9 is as follows: A = 6745 (43.04%), T = 5997 (38.27%), G = 1156 (7.38%), C = 1772 (11.31%) and the AT skewness is 0.059, indicating the occurrence of more As than Ts. Similar results are observed in Chinese *Bombyx mandarina* (0.057), Japanese *B. mandarina* (0.055), *Ostrinia furnicalis* (0.032) and *Ostrinia lunifer* (0.030). In addition, the entire mitochondrial genome of *B. mori* strain H9 is biased to A and T, with an A + T content of 81.31%, similar to other sequenced lepidopteran mitogenomes (Table 4). The AT skewness of tRNA and rRNA genes in the *B. mori* H9 mitogenome is 0.032 and 0.037, respectively. But the AT skew in the A + T-rich region is negative (−0.064), indicating the occurrence of more Ts than As. In the PCGs, the bias is toward As over Ts, and the AT skewness is 0.077. The GC skewness values in all sequenced lepidopteran mitogenomes are negative, ranging in size from −0.158 to −0.318.

3.2. Protein-coding genes

All the PCGs are initiated by ATN codons, except for the *cox1* gene (Table 3). A tetranucleotide, TTAG, has been designated as a *cox1* initiator. This noncanonical putative start codon is also found in *Artogeia*