



ELSEVIER

Contents lists available at ScienceDirect

## Data in brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)

## Data Article

Structural characteristics of a mitochondrial control region from *Myotis* bat (Vespertilionidae) mitogenomes based on sequence datasetsMd M. Rahman <sup>a, b</sup>, Kwang B. Yoon <sup>c</sup>, Yung C. Park <sup>b, \*</sup><sup>a</sup> Dept. of Biotechnology and Genetic Engineering, Islamic University, Kushtia, Bangladesh<sup>b</sup> Division of Forest Science, Kangwon National University, Chuncheon 24341, Republic of Korea<sup>c</sup> Inje County Office (186), Environment Protection Division, Inje-ro, Inje-eup, Inje-gun, Gangwon-do, Republic of Korea

## ARTICLE INFO

## Article history:

Received 3 January 2019

Received in revised form 7 February 2019

Accepted 6 March 2019

Available online 26 March 2019

## Keywords:

*Myotis*

Control region

Tandem repeat sequence

CSB

ETAS

## ABSTRACT

The datasets included sequences of a control region from *Myotis* bat mitogenomes. The control region (1706–2005 bp) of the *Myotis* mitogenomes was divided into three domains similar to that of other mammals, which included the common conserved blocks (ETAS domain, Central domain, and CSB domain). Several long tandem repeat sequences were present between the upstream of control regions and ETAS1. The size, base composition, and copy number of the long tandem repeat sequences differed between the *Myotis* species. Short tandem repeat sequences were also found between CSB1 and CSB2 in the CSB domain.

© 2019 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Data

The mitochondrial control region (CR), which is the most rapidly evolving region of the mitochondrial DNA, is one of the most commonly used molecular markers in population genetics and phylogenetic studies. We reported the structural characteristics of the CR from the genus *Myotis*, such as tandem repeat sequences and sequence composition, by analyzing the sequences extracted from five whole genomes. These characteristics may be responsible for the fast evolution of CR [2–6]. The

\* Corresponding author.

E-mail address: [parky@kangwon.ac.kr](mailto:parky@kangwon.ac.kr) (Y.C. Park).

## Specification table

Subject area	Biology
More specific subject area	Evolutionary Biology
Type of data	Table, figure, and word file
How data was acquired	Downloaded from NCBI GenBank
Data format	Analyzed DNA sequence
Experimental factors	Alignment of <i>Myotis</i> control region with Clustal W implemented in Geneious Pro 5.5.9, Tandem Repeats Finder program [1]
Experimental features	Features of domains and repeated sequences of the mitochondrial control region of the genus <i>Myotis</i> bats
Data source location	<i>Myotis formosus</i> [2] and <i>M. macrodactylus</i> [3] samples were collected in South Korea, <i>M. muricola</i> [4] in Malaysia, and <i>M. davidii</i> [5] and <i>M. brandtii</i> [6] in China
Data accessibility	<i>M. muricola</i> : <a href="https://www.ncbi.nlm.nih.gov/nuccore/NC_029422.1?from=15460&amp;to=17224&amp;report=fasta">https://www.ncbi.nlm.nih.gov/nuccore/NC_029422.1?from=15460&amp;to=17224&amp;report=fasta</a> <i>M. davidii</i> : <a href="https://www.ncbi.nlm.nih.gov/nuccore/NC_025568.1?from=15461&amp;to=17464&amp;report=fasta">https://www.ncbi.nlm.nih.gov/nuccore/NC_025568.1?from=15461&amp;to=17464&amp;report=fasta</a> <i>M. brandtii</i> : <a href="https://www.ncbi.nlm.nih.gov/nuccore/NC_025308.1?from=15451&amp;to=17400&amp;report=fasta">https://www.ncbi.nlm.nih.gov/nuccore/NC_025308.1?from=15451&amp;to=17400&amp;report=fasta</a> <i>M. formosus</i> : <a href="https://www.ncbi.nlm.nih.gov/nuccore/HQ184048.1?from=15454&amp;to=17159&amp;report=fasta">https://www.ncbi.nlm.nih.gov/nuccore/HQ184048.1?from=15454&amp;to=17159&amp;report=fasta</a> <i>M. macrodactylus</i> : <a href="https://www.ncbi.nlm.nih.gov/nuccore/KF440685.1?from=15462&amp;to=17466&amp;report=fasta">https://www.ncbi.nlm.nih.gov/nuccore/KF440685.1?from=15462&amp;to=17466&amp;report=fasta</a>
Related research article	F. Liu, Y. Song, S. Yan, J. Luo, F. Jiang, Structure and sequence variation of the mitochondrial DNA control region in <i>Myotis macrodactylus</i> , Chin. J. Zoo. 44 (2009) 19–27.

**Value of the data**

- These data will provide fundamental information to future molecular evolutionary studies of *Myotis* bats.
- The data will contribute to understanding rapid evolution in control regions of mammalian mitogenomes.
- The characteristics of the primary sequence of the control region will provide valuable information on population genetics, phylogeny, and phylogeography, which would be helpful in decision-making in bat ecological control and management programs.

datasets of CR sequences showed three common domains: extended terminal associated sequences (ETAS), central domain (CD), and conserved sequence block (CSB) (Fig. 1). Several copies of long tandem repeat sequences were located between the CR upstream and ETAS1 with the final copy positioned on ETAS1 (Fig. 1). Several short tandem repeat sequences were located between CSB1 and CSB2 (Fig. 2). The conserved sequence motifs (GYRCAT) were present in both regions of ETAS1 and ETAS2 within the ETAS domain. Highly conserved sequences were also found in the five regions of F to B boxes in the CD, and the three regions of CSB1 to CSB3 in the CSB domain (Fig. 2).

## 2. Experimental design, materials, and methods

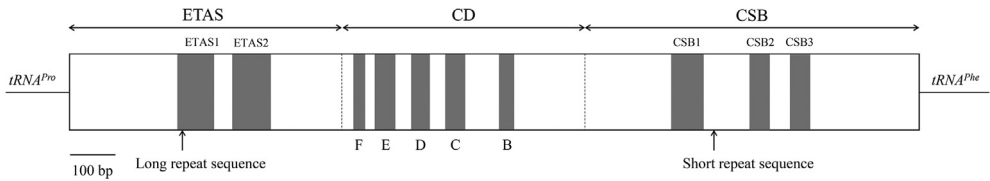
### 2.1. Sequence data collection

Whole mitogenome sequences of five species of the genus *Myotis* (*M. muricola*, *M. brandtii*, *M. formosus*, *M. macrodactylus*, and *M. davidii*) have been previously reported [2–6]. We extracted the CR sequences of these five mitogenomes from GenBank (Tables 1–2).

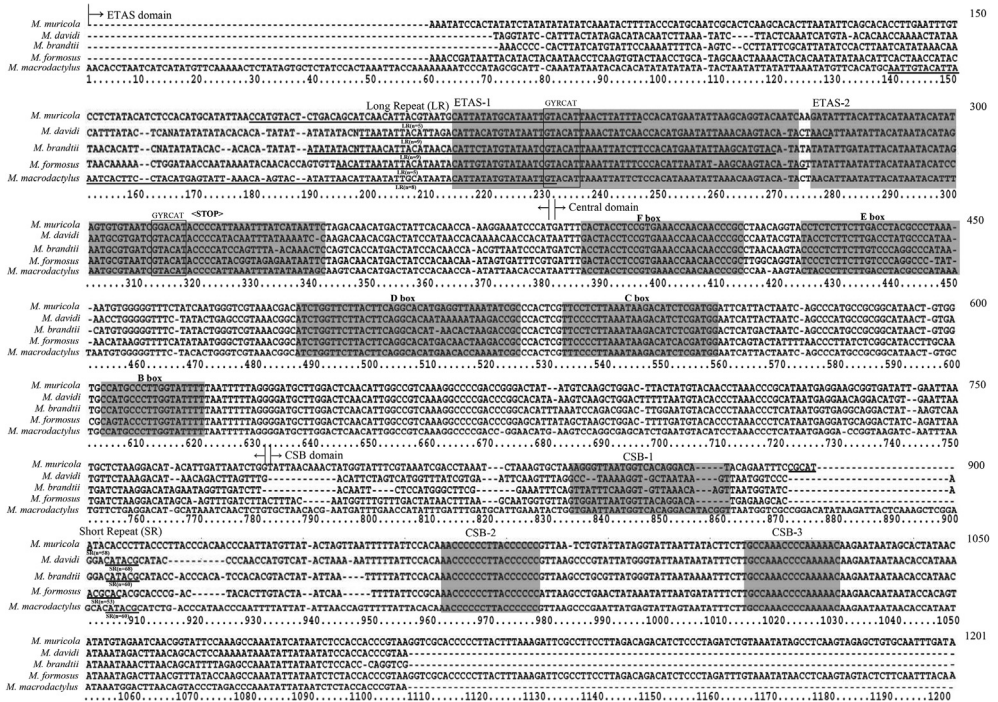
### 2.2. Processing and analysis of mitochondrial control region sequences of *Myotis* bats

#### 2.2.1. Sequence and primary structure of the mitochondrial control region

The CR sequences from *Myotis* bats were aligned using Clustal W implemented in Geneious Pro 5.5.9 (Auckland, New Zealand) and edited with BioEdit 7.02 (Tom Hall, USA). The CR sequences were annotated and characterized using other mammalian CR sequences [7–12] as references. The size of the



**Fig. 1.** Schematic diagram of the mitochondrial control region (CR) of the genus *Myotis*. The CR consists of three domains (ETAS, CD, and CSB). Highly conserved blocks were present in ETAS1-2, F–B, and CSB1-3 of the three domains. Long tandem repeats were present between CR upstream and ETAS1 and short tandem repeats were present between CSB1 and CSB2. Each copy of long and short repeat was marked in the figure.



**Fig. 2.** Dataset of the aligned CR sequences of *Myotis* bats. The datasets include only one copy of the tandem repeat sequences, which were underlined and indicated as long and short repeat sequences. The shaded areas indicate highly conserved sequences of ETAS1-2 blocks within ETAS, F-B blocks within the CD, and CSB1-3 blocks within CSB. The putative point of arrest of replication is indicated as (<STOP>). Regions of two Gyrcat (Y=C or T, R = A or G) motifs are indicated as boxes.

CR sequences ranged from 1706 (*M. formosus*) to 2005 bp (*M. mcrodactylus*) in length. The CR sequences were subdivided into three main domains: ETAS, CD, and CSB (Fig. 1).

**2.2.2. Detection and analysis of tandem repeat sequences**

The tandem repeat sequences in the CR region were investigated using Tandem Repeats Finder program [1], and we discovered the motif of the tandem repeat sequences, length of repeats, and copy number. Two type of tandem repeat sequences were found in the CR. Tandem repeats with long sequence motifs were present between CR upstream and ETAS-1 in the ETAS domain, and tandem repeats with short sequence motifs were present between CSB1 and CSB-2 (Fig. 1 and Table 2). The size of the long repeat sequence ranged from 80 bp in *M. formosus* to 82 bp in *M. muricola* and the copy

**Table 1**  
Tandem repeat sequences in mitochondrial control region of *Myotis* bats.

Species (Accession no.)	Long repeat sequence			Short repeat sequence			References for whole mitogenome sequences	
	Motif sequence	Length of a repeat sequence (bp)	Copy number	Total length of long repeat sequence region (bp)	Motif sequence	Copy number		Total length of short repeat sequence region (bp)
<i>M. muricola</i> (KT213444)	CCACATGAATATTAAGCAAGTACTTTAACAACATTAATATT ACATAATACATTATATGTATAAATTGTACATTAACCTATTTA	82	5.3	433	CGCATA	58.2	349	[4]
<i>M. davidii</i> (KM233172)	TTAATATTACATTAGACATTACATGTATAAATTGTACATTA CTATCAACCACATGAATATTAACAAGTACATACTAACA	81	9.1	741	CATACG	68.8	413	[5]
<i>M. brandtii</i> (KM199849)	ATATATATATTAACATTACATAACACATTCTATGTATAATCG TACATTAATTTATCTTCCACATGAATATTAAGCATGTAC	81	9.3	757	CATACG	60.8	365	[6]
<i>M. formosus</i> (HQ184048)	ATTAATATTACATAATACATTGTATGTATAATCGTACATT AAATTATTTCCACATTAATATAAGCAAGTACATAGTTAT	80	5.3	420	ACGCAT	53.8	323	[2]
<i>M. macrodactylus</i> (KF440685)	AATTGTACATTAATTAATTTTCCACATGAATATTAACAA GTACATACTAACATTAATATTACATAATACATTATATGTAT	81	8.9	721	CATACG	60.5	363	[3]

**Table 2**

Size (bp) of three domains in mitochondrial control region of five *Myotis* bats including each of one copy of long and short repeat sequences.

Species	ETAS	CD	CSB	Total size
<i>M. muricola</i>	317	392	388	1097
<i>M. davidii</i>	289	400	261	950
<i>M. brandtii</i>	293	392	263	948
<i>M. formosus</i>	313	395	365	1073
<i>M. macrodactylus</i>	368	393	323	1084
Mean $\pm$ SD	316 $\pm$ 31.5	394.4 $\pm$ 3.4	320 $\pm$ 57.9	1030.4 $\pm$ 74.8

number ranged from 5.3 copies in *M. muricola* and *M. formosus* to 9.3 copies in *M. brandtii*. The size of the short repeat sequence was 6 bp in all CRs, and the copy number ranged from 53.8 copies in *M. formosus* to 68.8 copies in *M. davidii* (Table 1). Excluding these tandem repeat regions, except for each copy of the long and short repeat motifs, the size of the CR ranged from 948 bp in *M. brandtii* to 1097 bp in *M. muricola* (Table 2).

### 2.2.3. Detection and analysis of the conserved region

We discovered sequence-conserved regions through the mitochondrial CRs based on comparative analysis of the datasets. The conserved sequences identified within the domains were in ETAS1–2, F–B boxes, and CSB1–3 (Fig. 2). ETAS1 was more conserved than ETAS2, suggesting that the former might be functionally more important in mtDNA replication than ETAS2 [7,8,13,14]. The conserved sequence motif 'GYRCAT' was present in two locations within ETAS: one of them within ETAS1 and the other within ETAS2 (Fig. 2). The putative point of arrest of D-loop synthesis, 'ACCCC', was situated within ETAS2, next to the 3' end of GYRCAT (Fig. 2). Similarly, putative points of arrest of D-loop synthesis were proposed to be ACCCC in rhinolophid bats [10] and ACCCC, GCCCC, or TCCCC in leaf-nosed bats [11].

### Acknowledgements

This study was supported by Basic Science Research Program through the National Research Foundation of Korea [NRF-2015R1D1A1A01061097] funded by the Ministry of Education and 2015 Research Grant from Kangwon National University [No. 520150254].

### Transparency document

Transparency document associated with this article can be found in the online version at <https://doi.org/10.1016/j.dib.2019.103830>.

### References

- [1] G. Benson, Tandem repeats finder: a program to analyze DNA sequences, *Nucleic Acids Res.* 27 (1999) 573–580.
- [2] K.B. Yoon, Y.C. Park, Complete mitochondrial genome and codon usage of the Nepalese whiskered bat *Myotis muricola* (Vespertilionidae), *Genet. Mol. Res.* 14 (2015) 14637–14645.
- [3] S.Q. Wang, Y.J. Li, A.G. Yin, W. Zhang, J.J. Jiang, W.L. Wang, M. Hu, The complete mitochondrial genome of David's myotis, *Myotis davidii* (Myotis, Vespertilionidae), *DNA Mapp. Seq. Anal.* 27 (2016) 1587–1588.
- [4] J. Jiang, S. Wang, Y. Li, W. Zhang, A. Yin, M. Hu, The complete mitochondrial genome of insect-eating brandt's bat, *Myotis brandtii* (Myotis, Vespertilionidae), *DNA Mapp. Seq. Anal.* 27 (2014) 1403–1404.
- [5] T.W. Nam, H.R. Kim, J.Y. Cho, Y.C. Park, Complete mitochondrial genome of a large-footed bat, *Myotis macrodactylus* (Vespertilionidae), *Mitochondrial DNA* 26 (2015) 661–662.
- [6] Y.M. Kim, E.H. Choi, S.K. Kim, K.H. Jang, S.H. Ryu, U.W. Hwang, Complete mitochondrial genome of the hodgson's bat *Myotis formosus* (mammalia, chiroptera, vespertilionidae), *Mitochondrial DNA* 22 (2011) 71–73.
- [7] A. Larizza, G. Pesole, A. Reyes, E. Sbisà, C. Saccone, Lineage specificity of the evolutionary dynamics of the mtDNA D-loop region in rodents, *J. Mol. Evol.* 54 (2002) 145–155.
- [8] A. Reyes, E. Nevo, C. Saccone, DNA sequence variation in the mitochondrial control region of subterranean mole rats, *Spalax ehrenbergi* superspecies, *Mol. Biol. Evol.* 20 (2003) 622–632.
- [9] F. Liu, Y. Song, S. Yan, J. Luo, F. Jiang, Structure and sequence variation of the mitochondrial DNA control region in *Myotis macrodactylus*, *Chin. J. Zool.* 44 (2009) 19–27.

- [10] K. Sun, J. Feng, L. Jin, Y. Liu, L. Shi, T. Jiang, Structure, DNA sequence variation and phylogenetic implications of the mitochondrial control region in horseshoe bats, *Mamm. Biol.* 74 (2009) 130–144.
- [11] K. Sun, L. Luo, Z. Zhang, J. Feng, Molecular characteristics and evolution of the mitochondrial control region in three genera (Hipposideridae: *hipposideros*, *Aselliscus*, and *Coelops*) of leaf-nosed bats, *Mitochondrial DNA* 24 (2013) 451–461.
- [12] B.M. Md- Zain, A. Abdul-Aziz, N.A. Rahman, N.S. Mohd-Yusof, J.R.R. Japning, N. Rosli, S. Yaakop, Sequence Variation Data of the Mitochondrial DNA D-Loop Region of the Captive Malayan Gaur (*Bos Gaurus hubbacki*), *Data in Brief* (2018) in press.
- [13] C.W. Matson, R.J. Baker, DNA sequence variation in the mitochondrial control region of red-backed voles (*Clethrionomys*), *Mol. Biol. Evol.* 18 (2001) 1494–1501.
- [14] A. Iyengar, F.M. Diniz, T. Gilbert, T. Woodfine, J. Knowles, N. Maclean, Structure and evolution of the mitochondrial control region in oryx, *Mol. Phylogenet. Evol.* 40 (2006) 305–314.