


The complete mitochondrial genome sequence of *Boechera stricta*

Junji Li^{a,b} , Changwei Bi^a, Jing Tu^a and Zuhong Lu^a

^aState Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing, China; ^bMIIT Key Laboratory of Advanced Display Materials and Devices, School of Environmental and Biological Engineering, Nanjing University of Science and Technology, Nanjing, China

ABSTRACT

Boechera stricta (*B. stricta*) is a wild relative of *Arabidopsis*, occurring in mostly montane regions of western North America. In this article, we assembled the complete mitochondrial (mt) DNA sequence of *B. stricta* into a circular genome of length 271,601 bp, including 31 protein-coding genes, 21 tRNA genes, and 3 rRNA genes. From the neighbour-joining phylogenetic tree was constructed, based on the 23 conserved protein-coding genes of *B. stricta* and other 23 plant species, and the phylogenetic relationship and evolution position of *B. stricta* were determined. The complete mt genome would be useful for further investigation of the genotype-by-environment interactions in mitochondria of *Boechera*.

ARTICLE HISTORY

Received 29 June 2018
Accepted 8 July 2018

KEYWORDS

Boechera stricta;
mitochondrial genome;
phylogeny

Boechera stricta, a close relative of the model plant *Arabidopsis thaliana*, mostly grows in montane areas of western North America. It is an important plant to study the species' evolution in environments (Wagner et al. 2014) according to the genetic diversities introduced by self-fertilization of wild accessions (Schranz et al. 2007). In a recent study, a major chromosomal inversion in the *Boechera stricta* genome was identified to control ecologically important traits during the incipient speciation (Lee et al. 2017). However, the complete mitochondrial genome of *Boechera stricta*, which encodes the main generation of cellular ATP and plays a predominant role in metabolic regulation, has not been reported until now. Hence in this study, we assembled and primarily analyzed the entire mitochondrial genome sequence of *Boechera stricta*.

The *Boechera stricta* sample used in this study was harvested from a contact zone of its two differentiated subspecies in the northern Rocky Mountains (Geographic coordinate: 45°41'26" N, 113°54'29" W). The genomic DNA was extracted from etiolated seedlings according to the reference (Lee, Wang et al. 2017), and then was deposited in Department of Energy Joint Genome Institute (CA, USA). In this article, we assembled the complete mitochondrial DNA sequence of *Boechera stricta* into a free-of-gap genome of 271,601-bp length using the software of Newbler v3.0 from 454 GSFLX sequencing data (Wang et al. 2018), and then revised it with more abundant sequencing data from Illumina 2500 sequencing platform (Bi et al. 2016). Finally, the complete mt-genome

sequence was submitted to GenBank with the accession number of MH545496. We calculated that the GC content of this mitochondrial genome to be 44.93% (A: 27.53%, C: 22.38%, G: 22.55% and T: 27.54%), which is a common value in Brassicaceae.

We used MITOFY program to annotate genes in the mitochondrial genome of *Boechera stricta* (Alverson et al. 2010), and identified 55 genes, including 31 protein-coding genes, 21 tRNA genes, and 3 rRNA genes. Among these, the trnK-TTT gene contains two copies, trnY-GTA contains three copies, and trnS contains four copies. All the protein genes have an ATG start codon except for the *nad1* gene with ACG as its initiation codon and the *mttB* gene with an undetermined start codon. Additionally, we found four types of stop codons in the protein-coding genes: TAA (17 genes: *atp4*, *atp6*, *cox1*, *cox2*, *ccmB*, *ccmC*, *ccmFc*, *nad1*, *nad2*, *nad3*, *nad4L*, *nad5*, *nad6*, *nad9*, *rpl5*, *rpl6*, and *rps4*), TGA (8 genes: *atp8*, *atp9*, *ccmFn*, *cob*, *cox3*, *nad4*, *rpl2*, and *rps12*), TAG (5 genes: *atp1*, *matR*, *mttB*, *nad7* and *rps3*) and one undetermined stop codon in *sdh4*. Furthermore, a total number of 32 exons and 18 introns were identified in nine protein-coding genes (*rps3*, *cox2*, *nad1*, *nad2*, *nad4*, *nad5*, *nad7*, *ccmFc* and *rpl2*). In order to clarify the position of *Boechera stricta* in the phylogenetic tree, twenty three conserved protein-coding genes, including *atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFc*, *ccmFn*, *cob*, *cox1*, *cox2*, *cox3*, *matR*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, and *nad9*, were selected from 24 plant mitochondrial genomes. As shown in Figure 1, the neighbor-joining tree demonstrated that *Boechera stricta* was evolutionarily

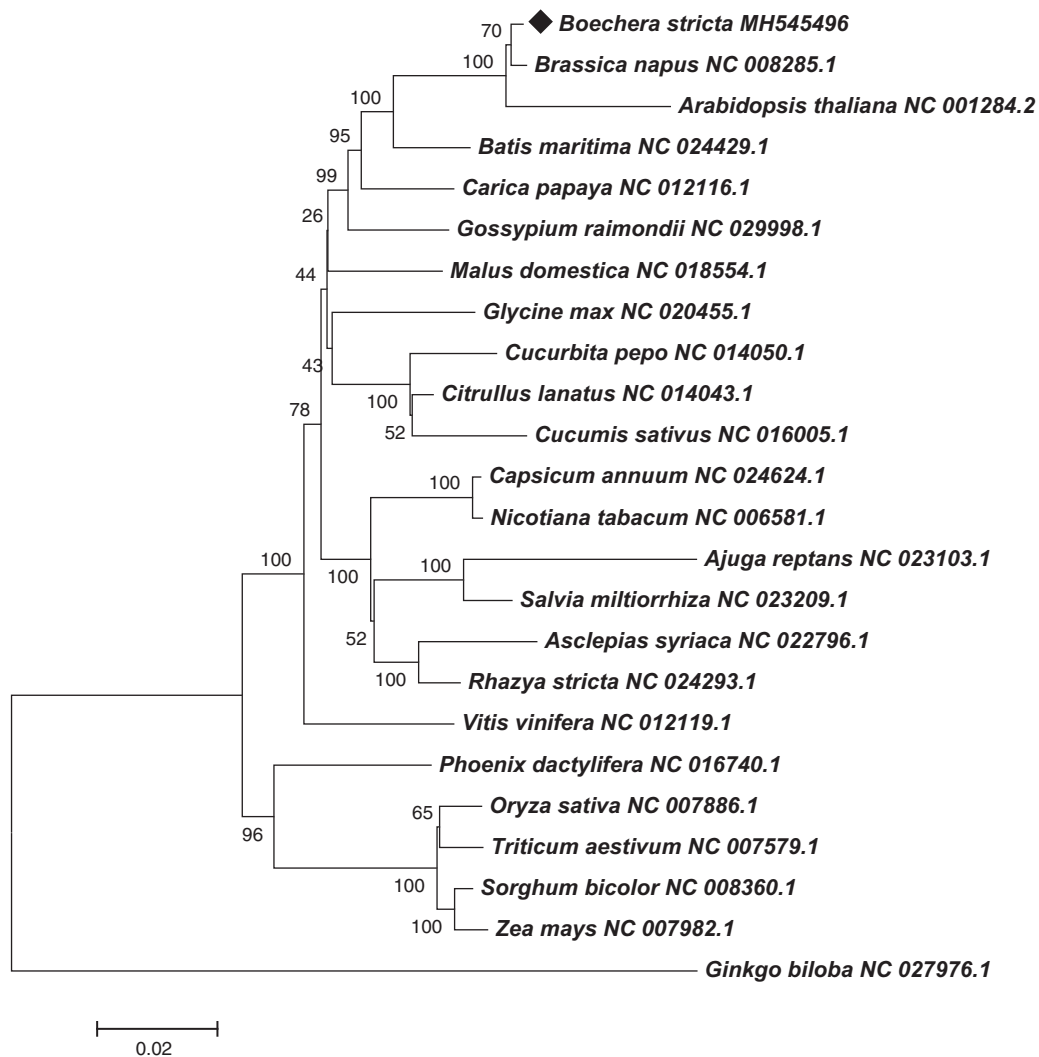


Figure 1. The Neighbor-Joining phylogenetic tree of 24 plant mt genomes based on 23 conserved mitochondrial genes. Bootstrap values are listed for each node. Accession numbers for tree construction are listed right to their scientific names.

close to *Brassica napus* than *Arabidopsis thaliana*, all belongs to Brassicaceae.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Acknowledgment

We would like to acknowledge the Department of Energy Joint Genome Institute for providing sequencing data of *Boechera stricta*.

Funding

This work was supported by National Key R&D Program of China [No. 2016YFA0501600] and the project 61571121 of the National Natural Science Foundation of China.

ORCID

Junji Li  <http://orcid.org/0000-0002-3104-6824>

Reference

- Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD. 2010. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol Biol Evol.* 27:1436.
- Bi, C., A. H. Paterson, X. Wang, Y. Xu, D. Wu, Y. Qu, A. Jiang, Q. Ye and N. Ye. 2016. Analysis of the complete mitochondrial genome sequence of the diploid cotton *Gossypium raimondii* by comparative genomics approaches. *BioMed Res Int.* 2016:3.
- Lee C-R, Wang B, Mojica JP, Mandáková T, Prasad KVS, Goicoechea JL, Perera N, Hellsten U, Hundley HN, Johnson J et al. 2017. Young inversion with multiple linked QTLs under selection in a hybrid zone. *Nat Ecol Evol.* 1:0119.
- Schranz ME, Windsor AJ, Song BH, Lawton-Rauh A, Mitchell-Olds T. 2007. Comparative genetic mapping in *Boechera stricta*, a close relative of *Arabidopsis*. *Plant Physiol.* 144:286–298.
- Wagner MR, Lundberg DS, Coleman-Derr D, Tringe SG, Dangl JL, Mitchell-Olds T. 2014. Natural soil microbes alter flowering phenology and the intensity of selection on flowering time in a wild *Arabidopsis* relative. *Ecol Lett.* 17:717–726.
- Wang X, Cheng F, Rohlsen D, Bi C, Wang C, Xu Y, Wei S, Ye Q, Yin T, Ye N. 2018. Organellar genome assembly methods and comparative analysis of horticultural plants. *Hortic Res.* 5:3.