


The complete mitochondrial genome sequence of an alpine plant *Arabis alpina*

Yiqing Xu^a  and Changwei Bi^b

^aSchool of Computer Science and Engineering, Southeast University, Nanjing, Jiangsu, China; ^bSchool of Biological Science and Medical Engineering, Southeast University, Nanjing, Jiangsu, China

ABSTRACT

Arabis alpina (*A. alpina*) is an arctic-alpine flowering plant in the family Brassicaceae, naturally growing in the tundra of arctic regions and in mountains at southern latitudes. In this study, we first report the assembly of the complete *A. alpina* mitochondrial (mt) genome using the next-generation sequencing technologies. The *A. alpina* mt circular genome is 323,159 bp in length and contains 33 protein-coding genes, 18 tRNA genes and 3 rRNA genes. To analyze the phylogenetic and evolutionary relationship of *A. alpina*, a neighbour-joining phylogenetic tree was reconstructed based on the mt genome of *A. alpina* and other 27 plants. The complete *A. alpina* mt genome will be helpful in population studies or investigations of mt functions of these alpine plant species.

ARTICLE HISTORY

Received 7 May 2018
Accepted 29 May 2018

KEYWORDS

Arabis alpina; mitochondrial genome; alpine plant; phylogeny

The alpine plant *Arabis alpina* is a thoroughly studied plant, which is well suited to analyze the flowering regulation of perennial plants (Bergonzi et al. 2013), plant phylogeography (Karl et al. 2012) and local adaptation to harsh alpine environmental conditions (Zulliger et al. 2013). Recently, there is growing interest to develop *A. alpina* as an emerging model for the study of genetics, population genetics and molecular biology (Lobréaux et al. 2014).

In most eukaryotic organisms, mitochondria exist as a double-membrane-bound organelle, which generate most of the cell's supply of ATP, and used as a source of chemical energy (Ye et al. 2017). It has been demonstrated that plant mt genomes vary considerably in their structure, length and gene order (Bi et al. 2016). Therefore, it is difficult to de novo assemble a plant mt genome. In this study, we assembled the complete *A. alpina* mt DNA sequence into a circular genome, which may contribute to improving our understanding of the population studies or investigations of mt functions in alpine plant organisms.

The plant sample of *A. alpina* used for extraction and genomic sequencing was collected in the French Alps site 11 in the Vercors mountains (Geographic coordinate: 44°52'47"N, 5°31'21"E, Elevation: 2010 m). Genomic DNA was extracted from 20 mg of dried material ground into a fine powder following the manufacturer's instructions, and then the DNA was deposited in Laboratoire d'Ecologie Alpine of Université Joseph Fourier (Grenoble, France). All genomic DNA samples were sequenced using Illumina technology and Roche/454 GSFLX Titanium, and the generated sequencing reads were a mixture of nucleus, chloroplast and mitochondrion. In order to assemble a complete *A. alpina* mt genome, we first assembled the mixed Roche/454 generated reads using

Newbler 3.0 (Roche Diagnostics company, Indianapolis, IN) with default parameters. After that, we filtered the mt contigs with the depth between 20 and 50, and linked these filtered contigs based on the Newbler generated file '454AllContigGraph.txt' using in-house Perl scripts. Gaps and errors in assembled contigs were filled up and corrected by Illumina generated reads with the software BWA (Li 2013) and MacVector. Finally, a total of 323,159 bp nucleotide mt genome were completely assembled, and then submitted to GenBank with the accession number NC_037070.1. The overall base composition of the *A. alpina* mt genome is A: 27.7%, G: 22.53%, C: 22.41%, T: 27.36%, and the GC content is 44.94%.

The *A. alpina* mt genome was annotated with an online program MITOFY (Alverson et al. 2010), and a total of 54 genes were identified, composed of 33 protein-coding genes, 18 tRNA genes and 3 rRNA genes. Among the 33 protein-coding genes, 9 genes (*ccmFc*, *cox2*, *nad1*, *nad2*, *nad4*, *nad5*, *nad7*, *rpl2* and *rps3*) were identified to contain a total of 18 introns and 32 exons. Additionally, we analyzed the codon usage of these protein-coding genes, and the results showed that most of the genes shared the same start codon ATG, except *nad1* uses ACG as start codon because of C to U RNA-editing, and two genes (*matR* and *mttB*) have no start codons probably due to the base mutation in evolution. Three types of stop codons were also identified in these protein-coding genes: TAG (*atp1*, *nad7*, *rps3* and *matR*), TGA (*atp8*, *atp9*, *cox3*, *ccmC*, *ccmFn*, *cob*, *nad4*, *rps12*, *rpl2* and *mttB*), TAA (*atp4*, *atp6*, *cox1*, *cox2*, *ccmB*, *ccmFc*, *nad1*, *nad2*, *nad3*, *nad4L*, *nad5*, *nad6*, *nad9*, *rps4*, *rps7*, *rpl5* and *rpl16*). In order to confirm the phylogenetic position of the *A. alpina* mt genome, 23

CONTACT Yiqing Xu  yiqingxu@njfu.edu.cn  School of Computer Science and Engineering, Southeast University, Nanjing, Jiangsu 211189, China

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

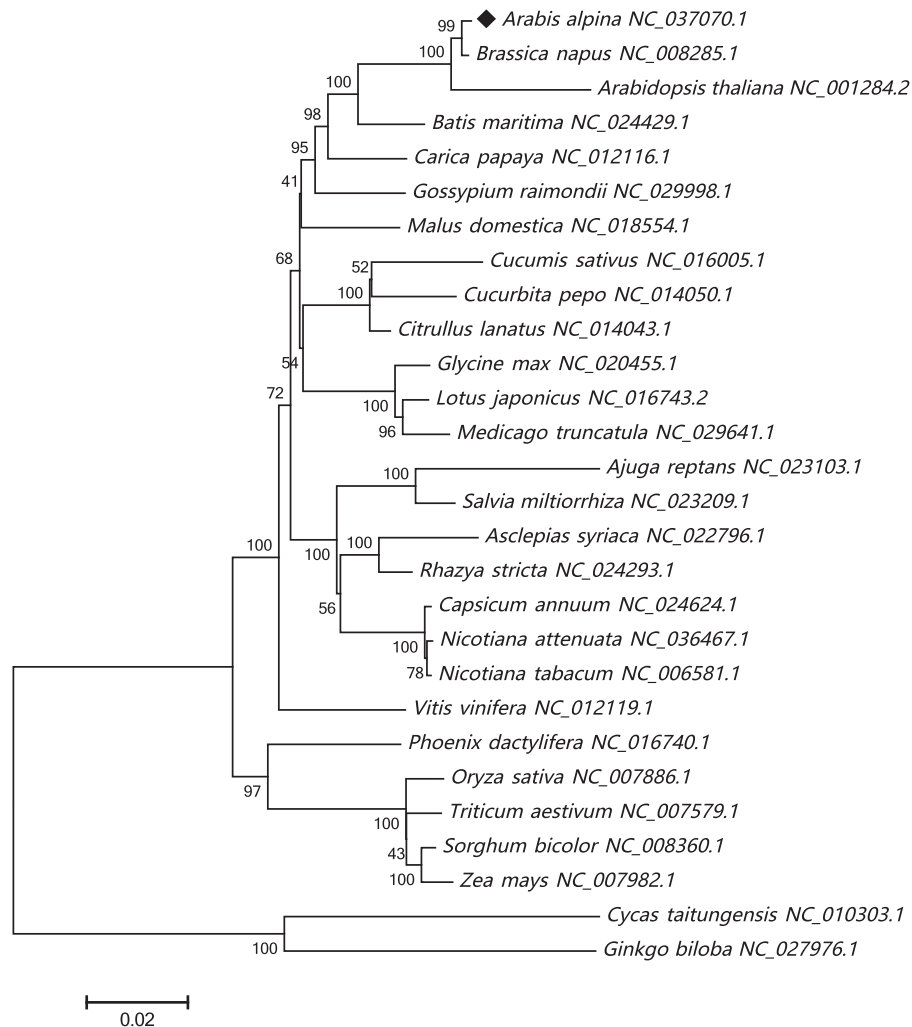


Figure 1. The neighbour-joining phylogenetic tree of 28 plant mt genomes based on 23 conserved mt genes. Bootstrap values are listed for each node. Accession numbers for tree construction are listed right to their scientific names.

conserved protein-coding genes (*atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFc*, *ccmFn*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9* and *matR*) were extracted from 28 plant mt genomes to reconstruct the phylogenetic tree. *Cycas taitungensis* and *Ginkgo biloba* were set as out-group. As illustrated in Figure 1, the neighbour-joining phylogenetic tree exhibited that *A. alpina* mt genome was evolutionarily close to that of *Brassica napus* and then *Arabidopsis thaliana* in Brassicaceae family. The complete mt genome of *A. alpina* would contribute to the further biological study of mt genomes in Brassicaceae and provide vital functional information of *A. alpina* in ecology.

Acknowledgments

We would like to acknowledge the Laboratoire d'Ecologie Alpine for their assistance in sampling the material of *Arabis alpina*.

Disclosure statement

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

Funding

This work was supported by the National Key Research and Development Plan of China (No. 2016YFD0600101).

ORCID

Yiqing Xu  <http://orcid.org/0000-0002-7962-6668>

References

- 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.
- Bergonzi S, Albani MC, Loren van Themaat EV, Nordstrom KJV, Wang R, Schneeberger K, Moerland PD, Coupland G. 2013. Mechanisms of age-dependent response to winter temperature in perennial flowering of *Arabis alpina*. *Science.* 340:1094–1097.
- Bi C, Paterson AH, Wang X, Xu Y, Wu D, Qu Y, Jiang A, Ye Q, Ye N. 2016. Analysis of the complete mitochondrial genome sequence of the diploid cotton *Gossypium raimondii* by comparative genomics approaches. *BioMed Res Int.* 2016:1.
- Karl R, Kiefer C, Ansell SW, Koch MA. 2012. Systematics and evolution of Arctic-Alpine *Arabis alpina* (Brassicaceae) and its closest relatives in the eastern Mediterranean. *Am J Bot.* 99:778.

- Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv, pp. 1303–1399.
- Lobréaux S, Manel S, Melodelima C. 2014. Development of an *Arabis alpina* genomic contig sequence data set and application to single nucleotide polymorphisms discovery. Mol Ecol Resour. 14:411.
- Ye N, Wang X, Li J, Bi C, Xu Y, Wu D, Ye Q. 2017. Assembly and comparative analysis of complete mitochondrial genome sequence of an economic plant *Salix suchowensis*. PeerJ. 5:e3148.
- Zulliger D, Schnyder E, Gugerli F. 2013. Are adaptive loci transferable across genomes of related species? Outlier and environmental association analyses in *Alpine Brassicaceae* species. Mol Ecol. 22:1626–1639.