

## Sequencing and phylogenetic analysis of the complete chloroplast genome of *Cassia tora* Linn

Binxin Xie<sup>a\*</sup>, Yanan Gai<sup>b\*</sup>, Ziyang Zhu<sup>a</sup>, Mingzhi Li<sup>c</sup> and Yucheng Zhao<sup>a</sup>

<sup>a</sup>Department of Resources Science of Traditional Chinese Medicines and State Key Laboratory of Natural Medicines, School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing, Jiangsu, P. R. China; <sup>b</sup>Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, Jiangsu, P. R. China; <sup>c</sup>Biodata Biotechnology Co. Ltd., Hefei, Anhui, China

### ABSTRACT

*Cassia tora* Linn. is widely distributed in South-East Asia and South-West Pacific as an important weed. It has many pharmacological activities including anti-allergic, anti-hepatotoxic, and remedy in skin diseases. In this study, we assembled and characterized the complete chloroplast genome sequence of *C. tora* from high-throughput sequencing data. The chloroplast genome was 162,426 bp in length, consisting of large single-copy (LSC) and small single-copy (SSC) regions of 90,843 bp and 18,001 bp, respectively, which were separated by a pair of 26,791 bp inverted repeat (IR) regions. The genome is predicted to contain 131 genes, including 84 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. The overall GC content of the genome is 36.0%. A phylogenetic tree reconstructed by 32 chloroplast genomes reveals that *C. tora* is mostly related to *Senna occidentalis*. The work reported the firstly complete chloroplast genome of *C. tora* which may provide useful information to the evolution of Cassieae Bronn.

### ARTICLE HISTORY

Received 21 October 2019  
Accepted 27 October 2019

### KEYWORDS



Complete chloroplast genome; phylogenetic analysis; *Cassia tora*

*Cassia tora* is a medicinal plant that is widely distributed in South-East Asia and South-West Pacific as an important weed. The leaf of *C. tora* has been used to manage haemorrhoids, skin infections, cough, pneumonia, stomach ache, ulcer, and fever for many years (Adamu et al. 2005). Pharmacological analysis indicated that it has the ability to anti-hepatotoxic, anti-allergic, anti-mutagenic, radical scavenging, hypoglycaemic, and anti-microbial activities (Rejiya et al. 2009; Chethana et al. 2017). Alkaloids, phenols, anthraquinones, glycosides, flavonoids, and saponins are the major secondary metabolites have been reported in *C. tora* (Yen et al. 1998; Aldouri 2000). However, little is known in its biosynthesis mechanism. In addition, the phylogenetic position of *C. tora* and the genus Cassieae Bronn is still unresolved. In this study, we first reported the complete chloroplast (cp) genome of *C. tora*, and its phylogenetic analysis is also investigated which provides informatics data for the phylogeny of genus Cassieae Bronn.

The fresh leaves of *C. tora* from Fuyang, Anhui, China (32°35'N, 114°50'E) were used for genomic DNA extraction. Specimens were stored in the Department of Resources Science of Traditional Chinese Medicines of China pharmaceutical University with the accession number of JM220190715XBX-4. Total genomic DNA was extracted with a FastPure Plant DNA Isolation Mini Kit (Vazyme, Nanjing,

China). The whole-genome sequencing was conducted by Hefei Biodata Biotechnologies Inc. (Hefei, China) on the Illumina Hiseq platform. The filtered sequences were assembled using the program SPAdes assembler 3.10.0 (Anton et al. 2012). Annotation was performed using the DOGMA and BLAST searches (Wyman et al. 2004). The cp genome of *C. tora* was determined to comprise a 162,426 bp double-stranded, circular DNA (GenBank accession no. MN480300), which containing two inverted repeat (IR) regions of 26,791 bp, separated by large single-copy (LSC) and small single-copy (SSC) regions of 90,843 bp and 18,001 bp, respectively. The overall GC content of *C. tora* cp genome is 36.0% and the corresponding values in LSC, SSC, and IR regions are 33.3, 30.2, and 42.3%, respectively. The cp genome was predicted to contain 131 genes, including 84 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. Eleven genes contained two exons and four genes (*ycf3*, *clpP*, and two *rps12*) contained three exons.

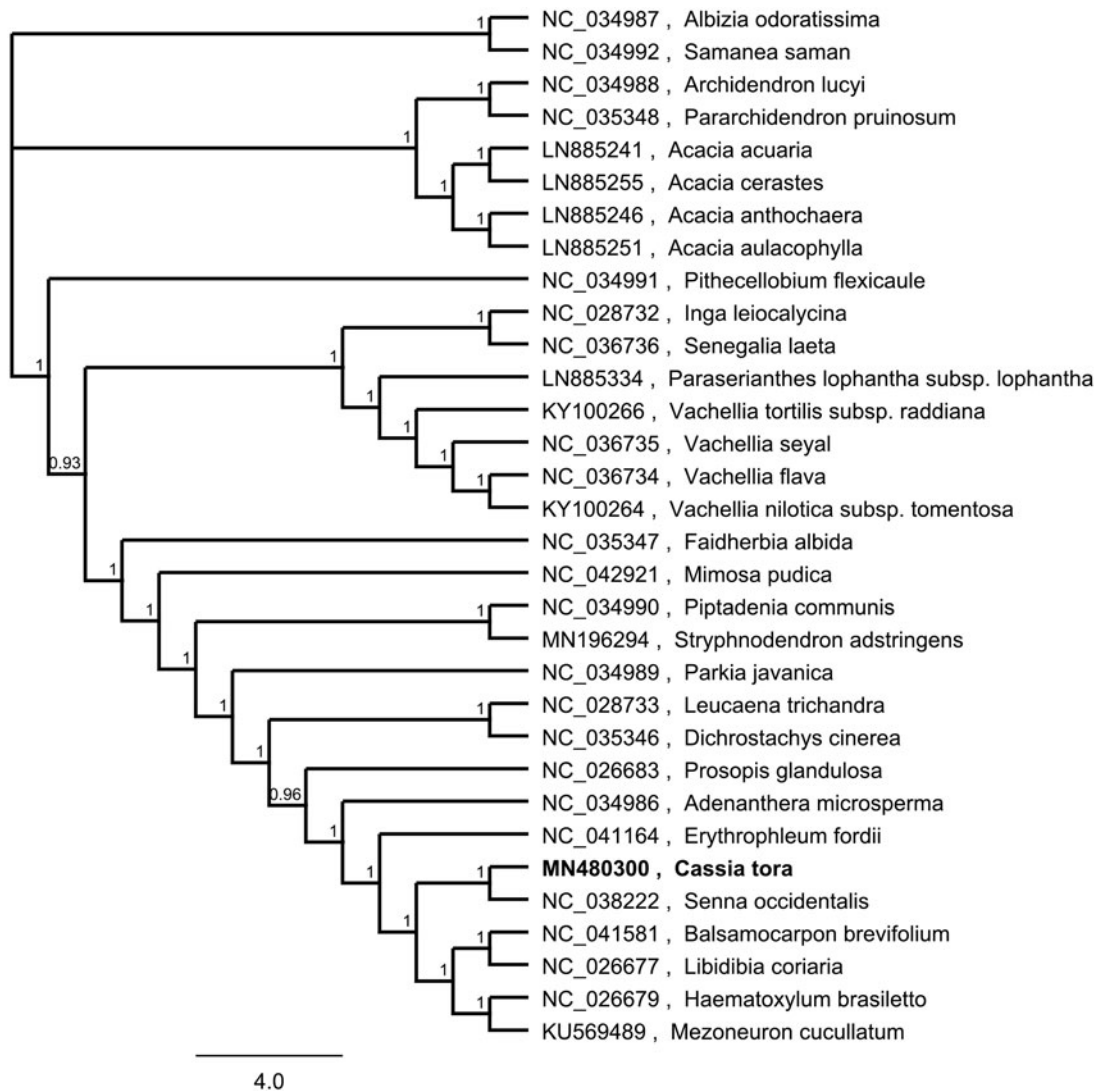
To investigate its taxonomic status, Alignment was performed on the 12 cp genome sequences using MAFFT v7.307, and a maximum likelihood (ML) tree was constructed by FastTree version 2.1.10 (Price et al. 2010; Kazutaka and Standley 2013). As expected, *Senna occidentalis* is the most related species to *C. tora*, with bootstrap support values of 100% (Figure 1). The complete cp genome sequence of

**CONTACT** Yucheng Zhao  [zhaoyucheng1986@126.com](mailto:zhaoyucheng1986@126.com)  Department of Resources Science of Traditional Chinese Medicines and State Key Laboratory of Natural Medicines, School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing, 210009, Jiangsu, P. R. China

\*These authors have contributed equally to this work.

© 2019 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.** Phylogenetic tree inferred by maximum likelihood (ML) method based on 32 representative species. A total of 1000 bootstrap replicates were computed and the bootstrap support values are shown at the branches. GenBank accession numbers were shown in Figure 1.

*C. tora* will provide a useful resource for the conservation genetics of this species as well as for the phylogenetic studies of Cassieae Bronn.

### Disclosure statement

No potential conflict of interest was reported by the authors.

### Funding

This project is funded by National Natural Science Foundation of China [81703637] and Natural Science Fund in Jiangsu Province [BK20170736].

### References

- Adamu HM, Abayeh OJ, Agho MO, Abdullahi AL, Uba A, Dukku HU, Wufem BM. 2005. An ethnobotanical survey of Bauchi State herbal plants and their antimicrobial activity. *J Ethnopharmacol.* 99(1):1–4.
- Aldouri NA. 2000. A survey of medicinal plants and their traditional uses in Iraq. *Pharm Biol.* 38:74–79.
- Anton B, Sergey N, Dmitry A, Gurevich AA, Mikhail D, Kulikov AS, Lesin VM, Nikolenko SI, Son P, Prjibelski AD. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19:455–477.
- Chethana KR, Senol FS, Orhan IE, Anilakumar KR, Keri RS. 2017. *Cassia tora* Linn.: a boon to Alzheimer's disease for its anti-amyloidogenic and cholinergic activities. *Phytomedicine.* 33:43–52.
- Kazutaka K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30:772–780.
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One.* 5(3):e9490.
- Rejiya CS, Cibir TR, Abraham A. 2009. Leaves of *Cassia tora* as a novel cancer therapeutic – an *in vitro* study. *Toxicol in Vitro.* 23(6): 1034–1038.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20(17):3252–3255.
- Yen GC, Chen HW, Duh PD. 1998. Extraction and identification of an antioxidative component from jue ming zi (*Cassia tora* L.). *J Agric Food Chem.* 46(3):820–824.