A clementine (Citrus clementina) is a citrus fruit hybrid between a mandarin orange and a sweet orange (Wu et al. 2014; Shimizu et al. 2016). Clementines have a looser skin that makes them easier to peel than oranges, though they are smaller than oranges. Additionally, they are sweeter and have more vitamin C than other mandarins and oranges. The scaffold-level genome (size: 301.4 Mb, scaffolds: 1398, N50: 110 kb) of C. clementina has been assembled using a Sanger whole genome shotgun approach in 2013 (Wu et al. 2014), but the chloroplast genome has not been published so far. The sequence and structure of different chloroplast genomes are very conserved, which will provide important genomic resources for further species identification and evolutionary study (Wang et al. 2018). In this study, the complete chloroplast genome of Citrus clementina was assembled and characterized, which will facilitate the taxonomic and evolutionary studies of the genus Citrus.

Leaf sample of C. clementina was collected from Citrus Research Institute of Chinese Academy of Agricultural Sciences (CAAS), Chongqing, China (29°45′36.2″N, 106°22′40.5″E), and the voucher specimen was deposited at the Herbarium of Nanjing Forestry University under the voucher number 20171221_CC01 (Haifeng Lin, haifeng.lin@njfu.edu.cn). Genomic DNA was extracted from leaves using hexadecyl trimethyl ammonium bromide (CTAB) DNA extraction protocol (Bubner et al. 2004) and then sequenced using Illumina HiSeq 4000 platform. The DNA quantity and quality were assessed by NanoDrop spectrophotometer and 1% agarose gel electrophoresis, respectively. The raw sequencing data were first filtered and trimmed by fastp program (Chen et al. 2018), and then fed into NOVOPlasty v4.2.1 (Dierckxsens et al. 2017) for assembly using the reference chloroplast genome sequence of Citrus sinensis (GenBank accession no.: NC_008334.1). The assembled genome was then annotated using PGA (Qu et al. 2019) and corrected manually using Macvector v18.1 (Bi et al. 2016), and finally submitted to NCBI GenBank under the accession number MW207298.

The chloroplast genome of C. clementina displayed a typical quadripartite structure of most angiosperm chloroplast genomes, composing of one large single-copy region (LSC: 87,879 bp), one small single-copy region (SSC: 18,779 bp), and a pair of inverted repeats (IRs: 27,022 bp). The overall GC content of the chloroplast genome was 38.42%, higher than that of LSC (36.79%) and SSC (33.12%), but lower than IRs (42.92%). The chloroplast genome of C. clementina encoded a total of 134 functional genes, including 90 protein-coding genes, 36 tRNA genes, and eight rRNA genes. Among these, 20 genes contained one intron (12 protein-coding genes and eight tRNA genes), and two genes contain two introns (ycf3 and clpP). To determine the phylogenetic relationships of C. clementina, we downloaded 16 other chloroplast sequences of genus Citrus and two species in Sapindales from the NCBI GenBank database. Anacardium occidentale and Dimocarpus longan were chosen as outgroup. A maximum-likelihood (ML) tree was constructed based on 76 conserved protein-coding genes of 19 cp genomes by RAxML (Stamatakis 2014) in CIPRES Science Gateway (Miller et al. 2010) with 1000 bootstrap replicates. The phylogenetic tree strongly supported that C. clementina was closely related to Citrus sunki (Figure 1). The complete chloroplast genome of C. clementina will provide important genetic and genomic resources to facilitate the molecular taxonomy and breeding of the genus Citrus.
Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Jiangsu Basic Research Program (Natural Science Foundation) Project No. [BK20191393], Jiangsu Agriculture Science and Technology Innovation Fund (JASTIF) [CX(20)3038], and Nanjing Forestry University Higher Education Research Project ‘Research on the First Class Curriculum Construction Standard of Programming Language for Agriculture and Forestry Majors’ [2020 [14]].

Data availability statement

The genome sequence data that support the findings of this study are openly available in NCBI GenBank at https://www.ncbi.nlm.nih.gov/nuccore/MW2072968. The associated BioProject, SRA, and Biosample numbers are PRJNA422948, SRR6415870, and SAMN08222600, respectively.

References

Bi C, Wang X, Xu Y, Wei S, Shi Y, Dai X, Yin T, Ye N. 2016. The complete mitochondrial genome of Medicago truncatula. Mitochondrial DNA Part B. 1(1):122–123.

Bubner B, Gase K, Baldwin IT. 2004. Two-fold differences are the detection limit for determining transgene copy numbers in plants by real-time PCR. BMC Biotechnol. 4:14.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 34(17):i884–i890.

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.

Miller M, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE 2010); New Orleans (LA). p. 1–8.

Qu XJ, Moore MJ, Li DZ, Yi TS. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. Plant Methods. 15:50.

Shimizu T, Kitajima A, Nonaka K, Yoshioka T, Ohta S, Goto S, Toyoda A, Fujiyama A, Mochizuki T, Nagasaki H, et al. 2016. Hybrid origins of citrus varieties inferred from DNA marker analysis of nuclear and organelle genomes. PLOS One. 11(1):e0166969.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9): 1312–1313.

Wang X, Cheng F, Rohlsen D, Bi C, Wang C, Xu Y, Wei S, Ye Q, Yin T, Ye N. 2018. Organelar genome assembly methods and comparative analysis of horticultural plants. Hortic Res. 5:3.

Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, Perrier X, Ruiz M, Scalabrini S, Terol J, et al. 2014. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. Nat Biotechnol. 32(7):656–662.