# Complete chloroplast genome assembly and phylogenetic analysis of blackcurrant (*Ribes nigrum*), red and white currant (*Ribes rubrum*), and gooseberry (*Ribes uva-crispa*) provide new insights into the phylogeny of Grossulariaceae

Xinyu Sun[1], Ying Zhan[1], Songlin Li[1], Yu Liu[1], Qiang Fu[1], Xin Quan[1], Jinyu Xiong[1], Huixin Gang[1,2,3], Lijun Zhang[2,4], Huijuan Qi[2,4], Aoxue Wang[1,3], Junwei Huo[1,2,3], Dong Qin[1,2,3] and Chenqiao Zhu[1,2,3]

[1] College of Horticulture & Landscape Architecture, Northeast Agricultural University, Harbin, Heilongjiang, China
[2] National-Local Joint Engineering Research Center for Development and Utilization of Small Fruits in Cold Regions, National Development and Reform Commission, Harbin, Heilongjiang, China
[3] Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture and Rural Affairs, Harbin, Heilongjiang, China
[4] Heilongjiang Institute of Green Food Science, Harbin, Heilongjiang, China

## ABSTRACT

**Background:** Blackcurrant (*Ribes nigrum*), red currant (*R. rubrum*), white currant (*R. rubrum*), and gooseberry (*R. uva-crispa*) belong to Grossulariaceae and are popular small-berry crops worldwide. The lack of genomic data has severely limited their systematic classification and molecular breeding.

**Methods:** The complete chloroplast (cp) genomes of these four taxa were assembled for the first time using MGI-DNBSEQ reads, and their genome structures, repeat elements and protein-coding genes were annotated. By genomic comparison of the present four and previous released five *Ribes* cp genomes, the genomic variations were identified. By phylogenetic analysis based on maximum-likelihood and Bayesian methods, the phylogeny of Grossulariaceae and the infrageneric relationships of the *Ribes* were revealed.

**Results:** The four cp genomes have lengths ranging from 157,450 to 157,802 bp and 131 shared genes. A total of 3,322 SNPs and 485 Indels were identified from the nine released *Ribes* cp genomes. Red currant and white currant have 100% identical cp genomes partially supporting the hypothesis that white currant (*R. rubrum*) is a fruit color variant of red currant (*R. rubrum*). The most polymorphic genic and intergenic region is *ycf1* and *trnT-psbD*, respectively. The phylogenetic analysis demonstrated the monophyly of Grossulariaceae in Saxifragales and the paraphyletic relationship between Saxifragaceae and Grossulariaceae. Notably, the *Grossularia* subgenus is well nested within the *Ribes* subgenus and shows a paraphyletic relationship with the co-ancestor of *Calobotrya* and *Coreosma* sections, which challenges the dichotomous subclassification of the *Ribes* genus based on morphology (subgenus *Ribes* and

subgenus *Grossularia*). These data, results, and insights lay a foundation for the phylogenetic research and breeding of *Ribes* species.

# INTRODUCTION

*Ribes* is the sole genus in the Grossulariaceae family and comprises approximately 150 to 200 known species. *Ribes* species are deciduous perennial shrubs that are native to temperate areas of Eurasia, North America, the northwestern part of Africa, and the Andes (*Cortez & de Mejia, 2019*). Cultivated *Ribes* species, commonly referred to as "currant" and "gooseberry", are popular and old small berry crops in Europe and Asia (*Lukša et al., 2018*; *Orsavová et al., 2019*). Blackcurrant (*R. nigrum*), red currant (*R. rubrum*), white currant (*R. rubrum*), and gooseberry (*R. uva-crispa*) (Fig. 1) are dominant cultivated taxa due to their high yield, fruit quality, and ecological adaptability (*Hummer & Dale, 2010*). In 2021, the global cultivation area of currants and gooseberries was 148,275 hectares, with a fruit production of 0.82 million tons, reaching a gross value of 1,503 million USD (FAOSTAT, https://www.fao.org/faostat/en/#data). Currants and gooseberries are commonly processed into juice, wine, candy, and jam owing to their high concentrations of functional metabolites and distinctive flavors (*Määttä, Kamal-Eldin & Törrönen, 2001*). In recent years, they have been reported to contain multi-functional bioactive metabolites, such as anthocyanins, polyphenols, phenolic acids, and flavonoids, and are characterized by potential antioxidant, anti-inflammatory, and antibacterial bioactivities that are involved in antidiabetic, antineoplastic, anti-tumor, anti-cardiovascular, and anti-neurodegenerative functions (*Lukša et al., 2018*; *Orsavová et al., 2019*). Despite their great economic importance and health potential, there has been little research on the *Ribes* species to date, which not only lags far behind that on major fresh fruits such as apple, citrus, and grape, but also falls behind that on emerging fruit crops that are mainly supplied for the processing market, such as sea-buckthorn (*Wu et al., 2022*; *Yu et al., 2022*), blue honeysuckle berry (*Zhu et al., 2023*), and wolfberry (*Cao et al., 2021*). In particular, the phylogeny of Grossulariaceae and the infrageneric relationships of the *Ribes* genus remain highly controversial (*Weigend, 2007*).

Based on morphology, *Ribes* was once included in the polyphyletic family Saxifragaceae (*Lingdi & Alexander, 2001*) but is now more accepted as the sole genus in the Grossulariaceae family (*Cronquist & Takhtadzhian, 1981*). However, because some literature still places Grossulariaceae under Saxifragaceae, the delineation of the Grossulariaceae family remains blurry (*Messinger, Hummer & Liston, 1999*), and its monophyletic origin has been rarely discussed in previous studies because only a limited number of *Ribes* species were used (*Dong et al., 2013*). In addition, the infrageneric classification of *Ribes* remains unclear due to its wide geographical distribution, unknown interspecific hybridization, and high phenotypic diversity (*Weigend, 2007*; *Zhang et al., 2023*). From the first monograph of Janczewski on *Ribes* (*Janczewski, 1907*), the
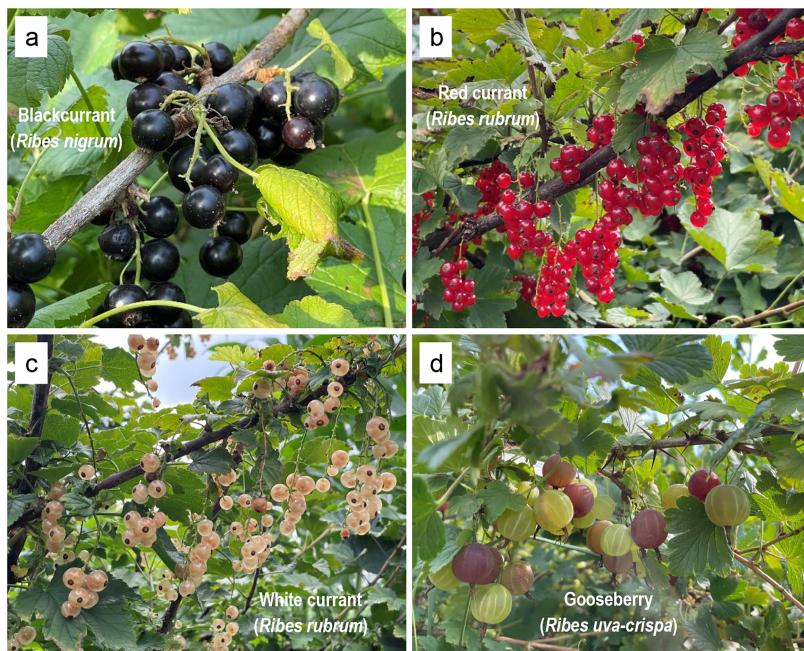
**Figure 1 Morphology of the four *Ribes* taxa.** (A) Blackcurrant (*R. nigrum*). (B) Red currant (*R. rubrum*). (C) White currant (*R. rubrum*). (D) Gooseberry (*R. uva-crispa*). Photo credit: by Xinyu Sun.

Full-size ◰ DOI: 10.7717/peerj.16272/fig-1

infrageneric classification of *Ribes* has been constantly proposed and revised (*Coville et al., 1908*; *Berger, 1924*; *Poyarkova, 1939*; *Rheder, 1940*; *Sinnott, 1985*; *Weigend, 2007*), resulting in a highly confusing taxonomy. In the work of Janczewski, the *Ribes* genus was divided into six subgenera with eleven sections (*Grossularioides*, *Parilla* (*Parilla* and *Andina*), *Berisia* (*Diacantha* and *Euberisia*), *Coreosma* (*Coreosma*, *Symphocalyx*, *Heritiera*, *Cerophyllum* and *Calobotrya*), *Ribes*, and *Grossularia* (*Grossularia* and *Robsonia*)) (*Janczewski, 1907*). Moreover, Berger divided Grossulariaceae into the *Ribes* genus (currants, including eight subgenera) and the *Grossularia* genus (gooseberries, including four subgenera). This classification was later demoted to two corresponding subgenera with 12 sections, which is the most consensus and widely accepted classification of *Ribes* so far (*Coville et al., 1908*; *Berger, 1924*; *Poyarkova, 1939*; *Sinnott, 1985*; *Messinger, Hummer & Liston, 1999*). Namely, the *Ribes* subgenus includes eight sections: *Berisia* (alpine currant), *Calobotrya* (red-flowering currant), *Coreosma* (blackcurrant), *Grossularioides* (spiny currant), *Heritiera* (skunk currant), *Parilla* (Andine currant), *Ribes* (red currant), *Symphocalyx* (golden currant), and the *Grossularia* subgenus includes four sections: *Grossularia* (true gooseberry), *Robsonia* (Sierra gooseberry), *Hesperia* (ornamental gooseberry), and *Lobbia* (desert gooseberry). Although internal transcribed spacer, nontranscribed spacer, and chloroplast markers have been recently applied to resolve the infrageneric phylogeny of *Ribes* (*Messinger, Hummer & Liston, 1999*; *Senters & Soltis, 2003*; *Schultheis & Donoghue, 2004*), inconsistent results have been reported regarding the monophyly of gooseberries (subgenus *Grossularia*) and the phylogenetic relationship among the above-mentioned 12 sections. Additionally, the genomic data of *Ribes*

have been exclusively used in studies of the evolution of Saxifragales (*Dong et al., 2013*; *Wang et al., 2021*) and the genetic diversity of endemic *Ribes* populations (*Zhang et al., 2023*).

Plastids are semi-autonomous organelles in the plant cell, carrying out photosynthesis, and various metabolic and signaling functions (*Daniell et al., 2016*; *Dobrogojski, Adamiec & Lucinski, 2020*). Due to the haploidy, uniparental inheritance, no genetic recombination, conserved genomic structure, and relatively small genome size of plastids, the polymorphic loci of plastids have been increasingly applied to plant phylogenetics, taxonomy, population genetics, and marker development for molecular breeding and DNA barcoding (*Olmstead & Palmer, 1994*; *Kuang et al., 2011*; *Dong et al., 2012*; *Gao et al., 2022*). Meanwhile, the possibility that a lack of variation in the modest number of plastid genes may hamper species discrimination, suggesting improved resolutions based on complete plastome sequences (*Parks, Cronn & Liston, 2009*; *Nock et al., 2011*). Through comparative analysis based on the complete chloroplast (cp) sequences, the phylogeny of some plants, such as *Cardiocrinum*, *Ficus*, Orchidaceae, and Myrtales, has been delineated from the species level to the order level (*Zhang et al., 2021*; *Chen, Hu & Zhang, 2022*; *Huang et al., 2022*; *Jiang et al., 2022*). However, only five *Ribes* cp genomes have been released in NCBI to date (Jan. 2023). Among them, *R. glaciale* and *R. fasciculatum* var. *chinense* (*Dong et al., 2018*) are wild species indigenous to East Asia; and Sierra currant (*R. nevadense*), Sierra gooseberry (*R. roezlii*) (*Folk et al., 2020*), and clove currant (*R. odoratum*) (*Wang et al., 2021*) are endemic species of America. There has been no report about the cp genomes of cultivated *Ribes* species and their phylogenies, which has severely limited the germplasm exploration, genetic conservation, cross-breeding, and genetic improvement of *Ribes*.

Herein, complete cp genomes of blackcurrant (*R. nigrum*), red currant and white currant (*R. rubrum*), and gooseberry (*R. uva-crispa*) were assembled and characterized, laying a molecular foundation for the research on *Ribes* taxa. The study also identified the highly polymorphic sites across all the nine released *Ribes* cp genomes, providing practical tools for molecular breeding and DNA barcoding. Moreover, phylogenetic trees of Saxifragales were constructed by using 31 complete cp genomes, which can shed new light on the phylogeny of the Grossulariaceae family and *Ribes* species.

## MATERIALS AND METHODS

### Plant materials and DNA extraction

The fresh leaves of blackcurrant 'Liangye houpi' (*R. nigrum*), red currant 'Red Cross' (*R. rubrum*), white currant 'Witte Parel' (*R. rubrum*), and gooseberry 'Pixwell' (*R. uva-crispa*) were collected at the Horticultural Station of Northeast Agricultural University (126.73°E, 45.74°N), Harbin, China. Samples were frozen immediately in liquid nitrogen and kept at an ultra-low temperature freezer (−80 °C) until DNA extraction. Total genomic DNA was isolated using the Hi-Fast Plant Genomic DNA Kit (GeneBetter BioTech Co. Ltd., Beijing, China) following the manufacturer's instructions (http://www.gene-better.cn/).

## Sequencing, assembly, and annotation

The DNA library preparation and next-generation sequencing were accomplished at *Frasergen Bioinformatics* Co., Ltd. (Wuhan, China) following the manufacturer's instructions (MGI Tech Co., Ltd., Shenzhen, China) of the DNBSEQ-T7 platform. The adaptors of the raw reads were removed using Trimmomatic v0.93 (*Bolger, Lohse & Usadel, 2014*) and then the reads were filtered using SOAPnuke v2.X (*Ashraf et al., 2018*) with parameters of "–lowQual=20, –nRate=0.005, –qualRate=0.5". The four cp genomes were assembled using GetOrganelle v1.7.6.1 (*Jin et al., 2020*) with parameters of "-R 15 -k 21,45,65,85,105" and with the reference of *R. fasciculatum* var. *chinense* (MH191388). Genomic annotation was performed using CPGAVS2 (*Shi et al., 2019*) and further examined manually according to *R. nevadense* (MN496075) and *R. roezlii* (MN496076) (*Folk et al., 2020*). The cp feature map was constructed using OGDRAW v1.3.1 (*Greiner, Lehwark & Bock, 2019*) with default parameters.

## Comparative genomic analysis

The five previously reported *Ribes* cp genomes were re-annotated using CPGAVS2 with the reference of *R. fasciculatum* var. *chinense* (MH191388). The nine *Ribes* cp genomes were aligned using MAFFT v7.471 (*Katoh & Standley, 2013*). The pairwise genetic divergence was calculated using BioEdit v7.2.6 (*Hall, Biosciences & Carlsbad, 2011*). The IR boundary comparison of the nine *Ribes* was performed using IRSCOPE (*Amiryousefi, Hyvönen & Poczai, 2018*). mVISTA program (*Frazer et al., 2004*) was used to visualize the alignments of the nine *Ribes* cp genomes with the reference of *R. fasciculatum* var. *chinense*.

## Repeat, SNP, and indel identification

Tandem Repeats Finder (*Benson, 1999*) was used to determine tandem repeats with default parameters. REPuter (*Kurtz et al., 2001*) was used to determine forward (F), palindromic (P), reverse (R), and complement (C) repeats with the parameters of "Hamming Distance: 3; Maximum Computed Repeats: 5,000 bp; Minimum Repeat Size: 30 bp". SSR loci were identified using Misa-web (*Beier et al., 2017*) (http://pgrc.ipk-gatersleben.de/misa/) with default parameters. SNPs and Indels among the nine cp genomes were identified using DnaSP v6.12 (*Rozas et al., 2017*) and the *Pi* values were calculated with a window length of 600 bp and a step size of 200 bp.

## Phylogenetic analysis

The complete cp genomes of 20 representative species in Saxifragales, four focal taxa in the present study, five previously released *Ribes* species, and two outgroup species (Rosales) were selected for phylogenetic analysis and their Genbank accession numbers are shown in Table S10. The multi-sequence alignment was formalized using trimAL (*Capella-Gutiérrez, Silla-Martínez & Gabaldón, 2009*). The maximum-likelihood (ML) phylogenetic tree was constructed using MEGA v11.0.13 in the multithread mode with 1,000 bootstrap (BS) replications (*Tamura, Stecher & Kumar, 2021*) and further annotated using iTOL v6 online (https://itol.embl.de/) (*Letunic & Bork, 2021*). The Bayesian tree was constructed

using BEAST 2.7.5 (*Bouckaert et al., 2019*) following the standard workflow (https://beast2.
blogs.auckland.ac.nz/tutorials/) and referenced the literature reported by *Nascimento, dos
Reis & Yang (2017)*. The confidences of the clades were evaluated by using posterior
probabilities (PP).

## RESULTS

### Organizations and gene features of chloroplast genomes of four *Ribes*

The cp genomes of the four *Ribes* taxa ranged from 157,450 bp (*R. nigrum*) to 157,802 bp
(*R. rubrum*) in length (Fig. 2). The large single copy (LSC), small single copy (SSC), and
inverted repeat regions (IRa/IRb) of the four cp genomes ranged from 87,086 bp
(*R. nigrum*) to 87,369 bp (*R. uva-crispa* and *R. rubrum*), 18,288 bp (*R. uva-crispa*) to 18,423
bp (*R. rubrum*), and 26,001 bp (*R. uva-crispa*) to 26,007 bp (*R. nigrum*), respectively
(Table 1). The coding sequence (CDS) length of the four cp genomes was from 80,253 bp
(*R. nigrum*) to 80,439 bp (*R. rubrum*). The overall GC content of the four cp genomes
ranged from 38.21% (*R. uva-crispa*) to 38.08% (*R. rubrum*). Moreover, the GC content of
LSC, SSC, and IRs ranged from 36.14% (*R. rubrum*) to 36.30% (*R. uva-crispa*), 33.15%
(*R. nigrum*) to 33.41% (*R. uva-crispa*), and 43.08% (*R. rubrum*) to 43.13% (*R. uva-crispa*),
respectively. Notably, red currant and white currant exhibited an identical cp structure,
including total length, length of the quadruplicate parts, and the GC content. A total of 131
genes were commonly annotated in the four cp genomes, including 86 protein-coding
genes, 37 tRNA genes, and eight rRNA genes (Table 2). Among them, seven
protein-coding genes, seven tRNA genes, and four rRNA genes were duplicated in the IRs.
A total of 19 genes (*petB*, *petD*, *atpF*, *ndhA*, two *ndhB*, two *rpl2*, two *rps12*, *rpl16*, *rpoC1*,
*rps16*, two *trnA-UGC*, *trnG-GCC*, two *trnI-GAU*, *trnK-UUU*, *trnL-UAA*, and *trnV-UAC*)
harbored one intron (Table S1); two genes (*ycf3* and *clpP*) harbored two introns; and *rps12*
underwent trans-splicing (Fig. S1).

### Identification of repeat elements among the four *Ribes* cp genomes

Aiming to provide useful information for marker development, the repeat elements in the
four cp genomes were annotated. A total of 219 long repeats were identified (Fig. 3A),
including tandem, forward, and palindromic repeats but not reverse and complement
repeats. Tandem repeats accounted for the largest proportion (44.7%), followed by
palindromic repeats (30.1%) and then forward repeats (25.1%) (Tables S2 and S3).
For blackcurrant (*R. nigrum*), red/white currant (*R. rubrum*), and gooseberry (*R. uva-
crispa*), 53, 51, and 64 long repeats were identified, respectively. A total of 176 SSR loci
were identified in the four cp genomes, including mononucleotides (97.2%) and
dinucleotides (2.8%). In addition, 54, 40, and 42 SSRs were identified in blackcurrant
(*R. nigrum*), red/white currant (*R. rubrum*), and gooseberry (*R. uva-crispa*), respectively
(Fig. 3B and Table S4). SSRs showed the highest distribution frequency in the LSC (32–40),
followed by the SSC (4–7), and then the IRs (2–4) (Fig. 3C). Moreover, SSRs had the
highest distribution frequency in the intergenic region (28–33), followed by the intron
region (6–9), and the CDS (3–17) (Fig. 3D).
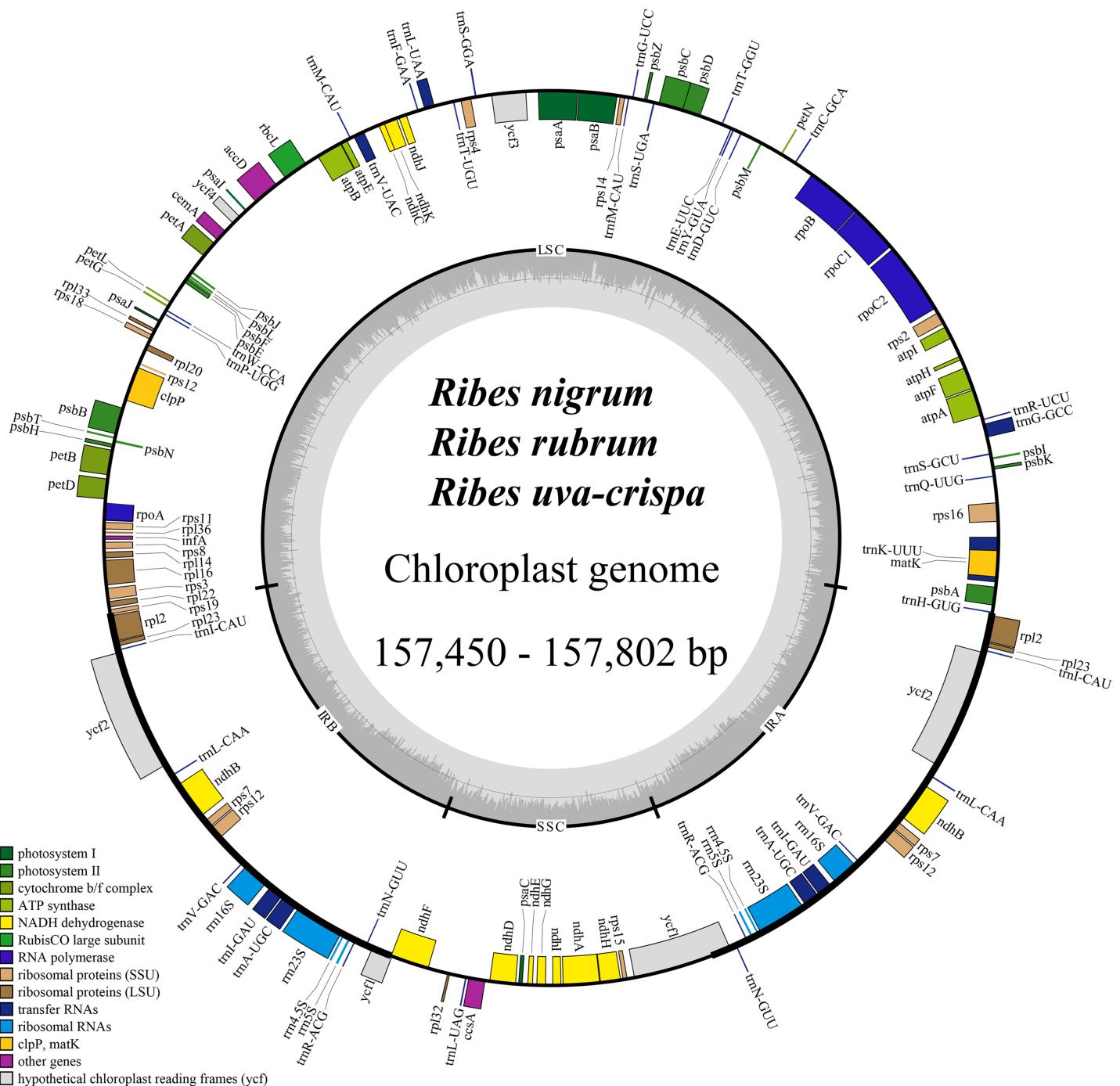
**Figure 2 Consensus feature map of cp genomes for blackcurrant (*Ribes nigrum*), red and white currant (*R. rubrum*), and gooseberry (*R. uva-crispa*).** Genes outside and inside the circle are transcribed clockwise and counterclockwise, respectively. Gray bars in the inner ring show the GC content percentage. The inverted repeat regions (IRa and IRb) are denoted with thick lines, which are separated by the large single copy (LSC) region and the small single copy (SSC) region. Genes belonging to different functional groups are color-coded accordingly at the bottom left.

Full-size ☒ DOI: 10.7717/peerj.16272/fig-2

**Table 1 Summary of cp genome subunits in the four *Ribes* taxa.**

| Species | Blackcurrant (*R. nigrum*) | Red currant (*R. rubrum*) | White currant (*R. rubrum*) | Gooseberry (*R. uva-crispa*) |
|---|---|---|---|---|
| Total size | 157,450 bp | 157,802 bp | 157,802 bp | 157,659 bp |
| LSC length | 87,086 bp | 87,369 bp | 87,369 bp | 87,369 bp |
| SSC length | 18,350 bp | 18,423 bp | 18,423 bp | 18,288 bp |
| IR length | 26,007 bp | 26,005 bp | 26,005 bp | 26,001 bp |
| CDS length | 80,253 bp | 80,439 bp | 80,439 bp | 80,340 bp |
| LSC GC content | 36.19% | 36.14% | 36.14% | 36.30% |
| SSC GC content | 33.15% | 33.19% | 33.19% | 33.41% |
| IR GC content | 43.12% | 43.08% | 43.08% | 43.13% |
| Total GC content | 38.13% | 38.08% | 38.08% | 38.21% |

**Table 2 Gene annotation of the four *Ribes* cp genomes.**

| Function | Gene names | Number |
|---|---|---|
| Photosystem I | *psaA*; *psaB*; *psaC*; *psaI*; *psaJ* | 5 |
| Photosystem II | *psbA*; *psbB*; *psbC*; *psbD*; *psbE*; *psbF*; *psbH*; *psbI*; *psbJ*; *psbK*; *psbL*; *psbM*; *psbN*; *psbT*; *psbZ* | 15 |
| Cytochrome b/f complex | *petA*; *petB**; *petD**; *petG*; *petL*; *petN* | 6 |
| ATP synthase | *atpA*; *atpB*; *atpE*; *atpF**; *atpH*; *atpI* | 6 |
| NADH-dehydrogenase | *ndhA**; *ndhB* (×2); *ndhC*; *ndhD*; *ndhE*; *ndhF*; *ndhG*; *ndhH*; *ndhI*; *ndhJ*; *ndhK* | 12 |
| Large subunit of ribosome | *rpl2** (×2); *rpl14*; *rpl16**; *rpl20*; *rpl22*; *rpl23* (×2); *rpl32*; *rpl33*; *rpl36* | 11 |
| DNA-dependent RNA polymerase | *rpoA*; *rpoB*; *rpoC1**; *rpoC2* | 4 |
| Small subunit of ribosome | *rps2*; *rps3*; *rps4*; *rps7* (×2); *rps8*; *rps11*; *rps12*$^{* \text{ T}}$ (×2); *rps14*; *rps15*; *rps16**; *rps18*; *rps19* | 14 |
| Transfer RNAs | *trnA-UGC** (×2); *trnC-GCA*; *trnD-GUC*; *trnE-UUC*; *trnF-GAA*; *trnfM-CAU*; *trnG-GCC**; *trnG-UCC*; *trnH-GUG*; *trnI-CAU* (×2); *trnI-GAU** (×2); *trnK-UUU**; *trnL-CAA* (×2); *trnL-UAA**; *trnL-UAG*; *trnM-CAU*; *trnN-GUU* (×2); *trnP-UGG*; *trnQ-UUG*; *trnR-ACG* (×2); *trnR-UCU*; *trnS-GCU*; *trnS-GGA*; *trnS-UGA*; *trnT-GGU*; *trnT-UGU*; *trnV-GAC* (×2); *trnV-UAC**; *trnW-CCA*; *trnY-GUA* | 37 |
| Ribosomal RNAs | *rrn4.5S* (×2); *rrn5S* (×2); *rrn16S* (×2); *rrn23S* (×2) | 8 |
| Conserved open reading frames | *ycf1* (×2); *ycf2* (×2); *ycf3***; *ycf4* | 6 |
| Subunit of acetyl-CoA-carboxylase | *accD* | 1 |
| C-type cytochrome synthesis gene | *ccsA* | 1 |
| Envelop membrane protein | *cemA* | 1 |
| Protease | *clpP*** | 1 |
| Translational initiation factor | *infA* | 1 |
| Maturase | *matK* | 1 |
| Rubisco | *rbcL* | 1 |

**Note:**
* indicates one intron; ** indicates two introns; (×2) indicates that the gene contains two copies in IRs; $^{\text{T}}$ indicates trans-spliced genes.
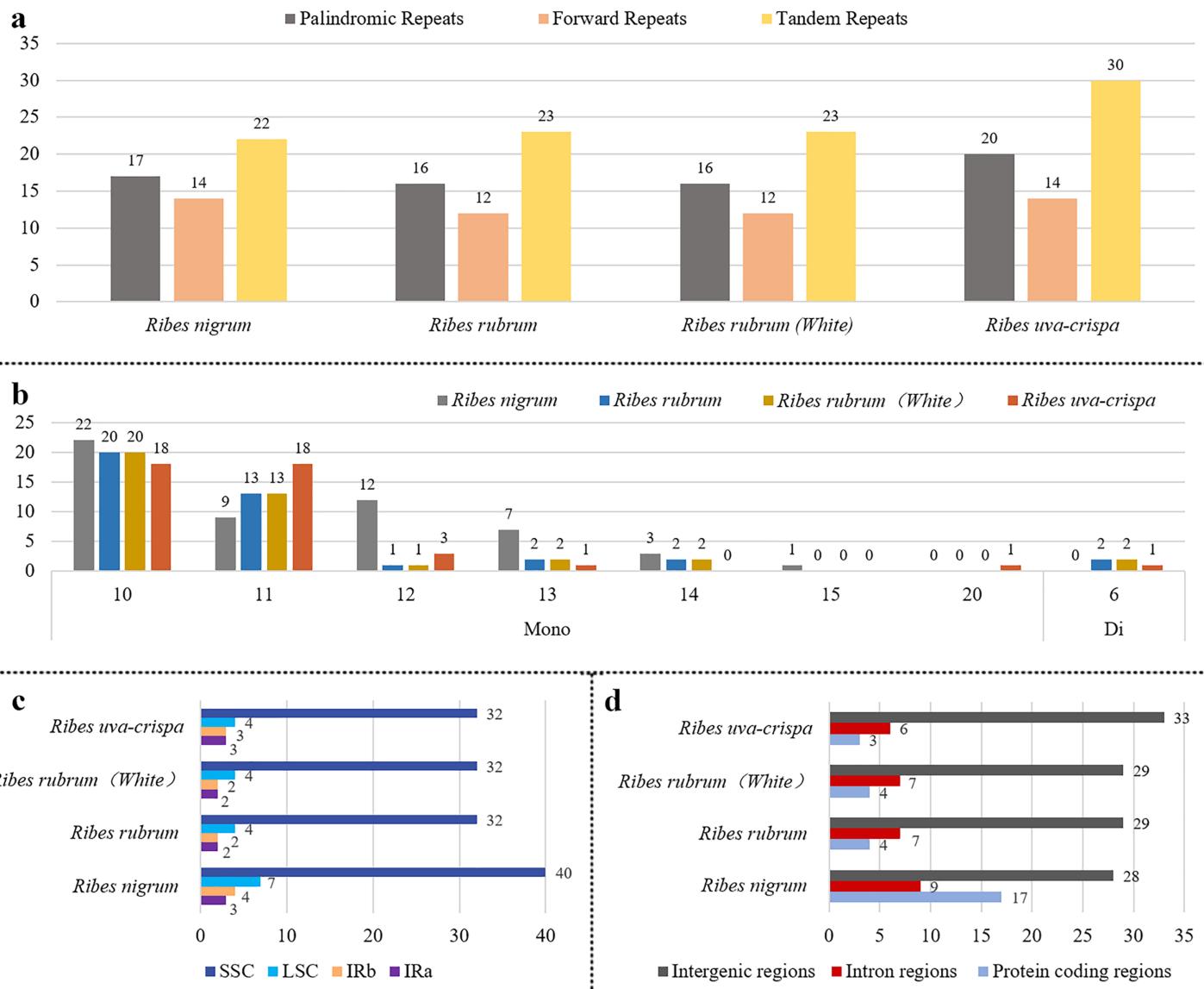
**Figure 3 Characterization of repeats in the four *Ribes* cp genomes.** (A) Summary of long repeats in the four *Ribes* cp genomes. (B) Types and repetitions of SSRs identified in the four cp genomes. Mono: mononucleotide-repeat; Di: dinucleotide-repeat. The numbers above the bars are the counts of the SSRs and those below the bars are the repetitions. (C) Counts of SSRs in the LSC, SSC, and IR regions. (D) Counts of SSRs in intergenic, intron, and protein-coding regions.

Full-size ☑ DOI: 10.7717/peerj.16272/fig-3

## Genomic comparison of the nine *Ribes* cp genomes

We performed a genomic comparison of nine *Ribes* cp genomes, including the four assembled in the present study and five cp genomes released previously (*R. glaciale, R. fasciculatum* var. *chinense, R. nevadense, R. roezlii,* and *R. odoratum*). The nine *Ribes* cp genomes showed identical genomic features, gene orders, and gene numbers (Fig. 4). As expected, the coding regions exhibited lower levels of divergence than non-coding regions. The pairwise genetic divergence between any two of the nine cp genomes was lower than 2.90% (Table S5). The highest sequence divergence was detected between *R. fasciculatum* var. *chinense* and *R. odoratum*. Notably, the cp genomes of red currant and
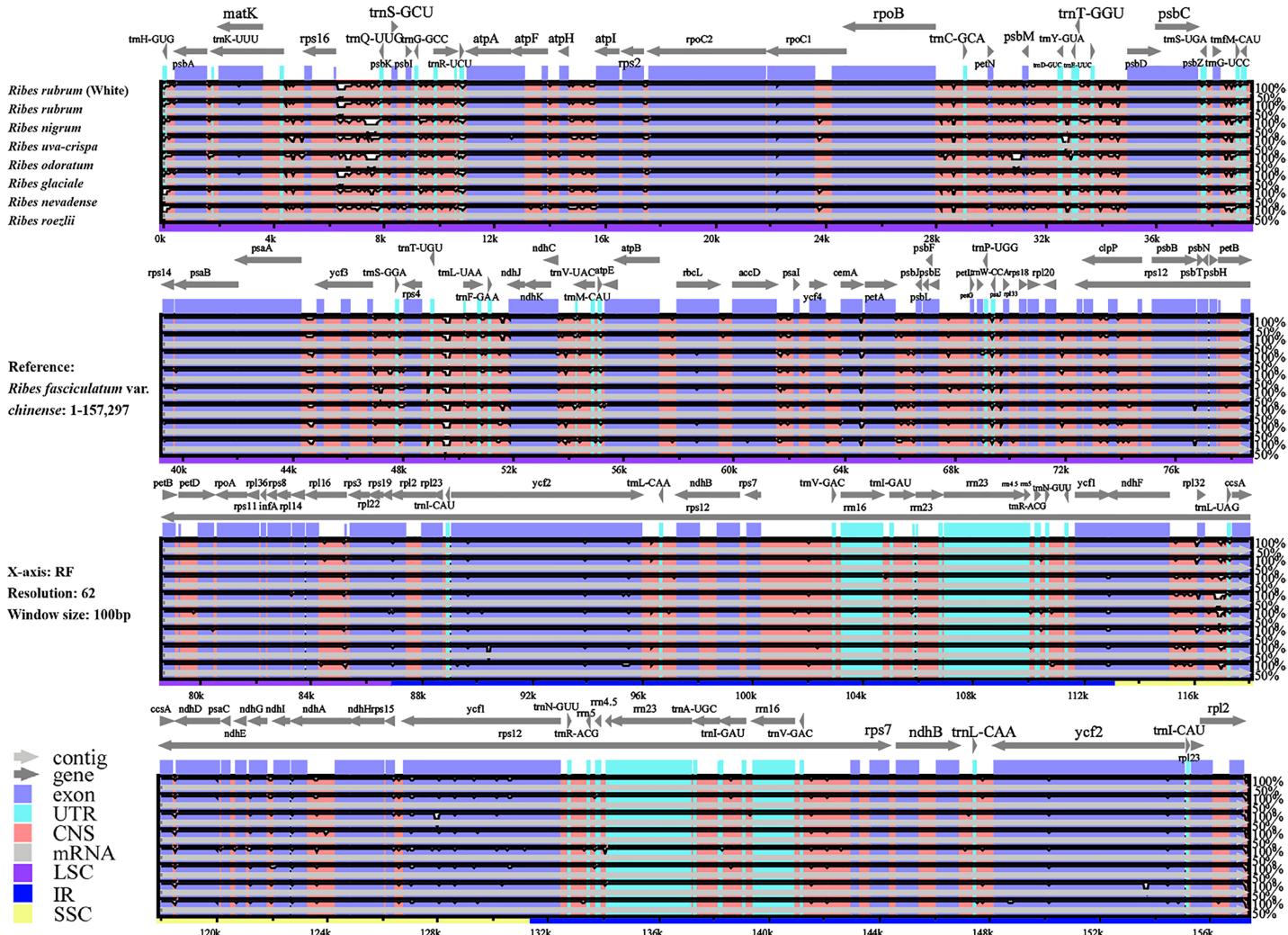
**Figure 4 Global alignment of the nine *Ribes* cp genomes.** The x-axis represents aligned base sequences, and the y-axis represents percent pairwise identity within 50–100%. (Coding regions are shown in blue, and non-coding regions are indicated by pink. The location and direction of each gene are indicated by gray arrows above the alignment. CNS, conserved noncoding sequences; UTR, untranslated region).

white currant showed 100% sequence identity (Dataset S1). The IR/SC boundary shifts among the nine cp genomes were investigated (Fig. 5). Generally, 16 types of IR/SC shift were identified among the nine *Ribes* cp genomes. The *rps*19 gene spanned the LSC-IRb boundary (JLB) with five types of shifts, including "219 bp + 60 bp" (*R. uva-crispa*, *R. nigrum*, *R. roezlii*, and *R. nevadense*), "212 bp + 70 bp" (*R. rubrum*), "209 bp + 70 bp" (*R. glaciale*), "207 bp + 72 bp" (*R. odoratum*), and "202 bp + 77 bp" (*R. fasciculatum* var. *chinense*). The pseudogene gene *ycf1* (Ψ) spanned the IRb-SSC boundary (JSB) with four types of shifts, including "1,131 bp + 11 bp" (*R. uva-crispa* and *R. roezlii*), "1,141 bp + 16 bp" (*R. nigrum*, *R. nevadense*, and *R. glaciale*), "1,131 bp + 98 bp" (*R. rubrum* and *R. fasciculatum* var. *chinense*), and "1,141 bp + 94 bp" (*R. odoratum*). The *ycf1* gene spanned the SSC-IRa boundary (JSA) with four types of shifts, including "4,530 bp + 1,131 bp" (*R. uva-crispa* and *R. roezlii*), "4,406 bp + 1,141bp" (*R. nigrum*), "4,257 bp +
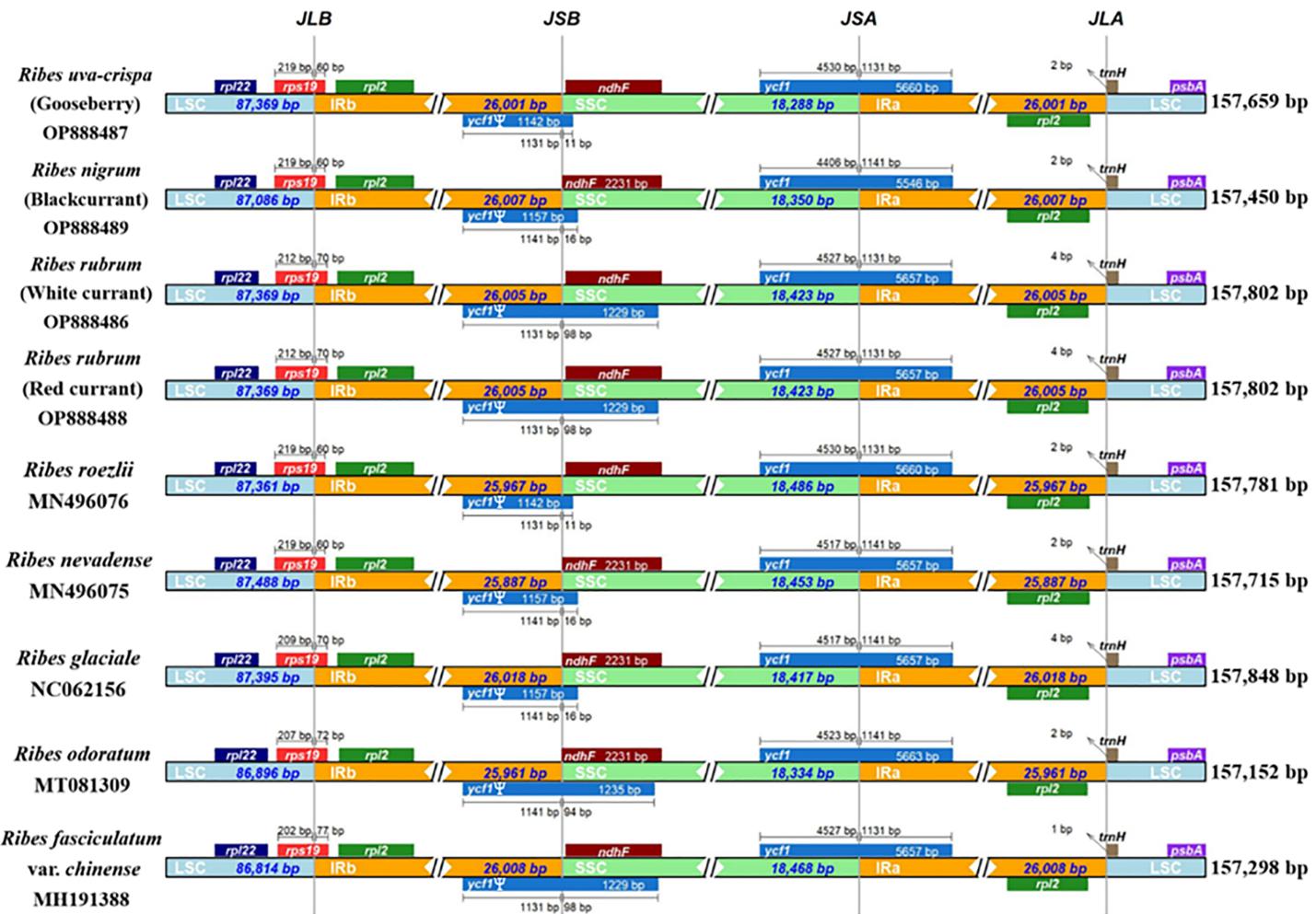
**Figure 5 Comparison of LSC, SSC, and IR region boundaries in the nine *Ribes* cp genomes.** Distance in the figure is not to scale. The gene names are indicated in the colored box, and the length of the corresponding region is shown next to the box. Ψ represents the pseudogene. JLB, LSC-IRb boundary; JSB, IRb-SSC; JSA, SSC/IRa; JLA, IRa-LSC. LSC, large single copy; IRa, IRb, two IR regions that are identical but in opposite orientations; SSC, small single copy. Full-size ☑ DOI: 10.7717/peerj.16272/fig-5

1,131 bp" (*R. rubrum* and *R. fasciculatum* var. *chinense*), "4,517 bp + 1,141 bp" (*R. nevadense* and *R. glaciale*), and "4,523 bp + 1,141 bp" (*R. odoratum*). The *trnH* gene spanned the IRa-LSC boundary (JLA) with three types of IRa residue, including "1 bp" (*R. fasciculatum* var. *chinense*), "2 bp" (*R. uva-crispa, R. nigrum, R. roezlii, R. nevadense,* and *R. odoratum*), and "4 bp" (*R. rubrum* and *R. glaciale*).

## Identification of SNPs and Indels in the *Ribes* cp genomes

A total of 3,322 SNPs and 485 Indels were identified in the nine *Ribes* cp genomes (Table 3). The LSC region comprised the largest number of SNPs (2,337) and Indels (361), followed by the SSC region (701 SNPs and 65 Indels). Generally, the CDS is more conserved than non-coding regions. Here, a total of 1,594 SNPs and 377 Indels were detected in the intergenic spacer (IGS); 424 SNPs and 80 Indels were found in the intron; and 1,304 SNPs and 28 Indels were detected in the CDS (Table S6). The most polymorphic

**Table 3 Summary of SNPs and Indels identified in the nine *Ribes* cp genomes.**

| Genomic region | SNPs | Indels |
|---|---|---|
| IRa | 141 | 36 |
| IRb | 142 | 35 |
| LSC | 2,337 | 348 |
| SSC | 701 | 65 |
| LSC-IRb | 1 | 0 |
| IRb-SSC | 0 | 1 |
| Total | 3,322 | 485 |
| Intergenic regions | 1,594 | 377 |
| Intron | 424 | 80 |
| CDS | 1,304 | 28 |
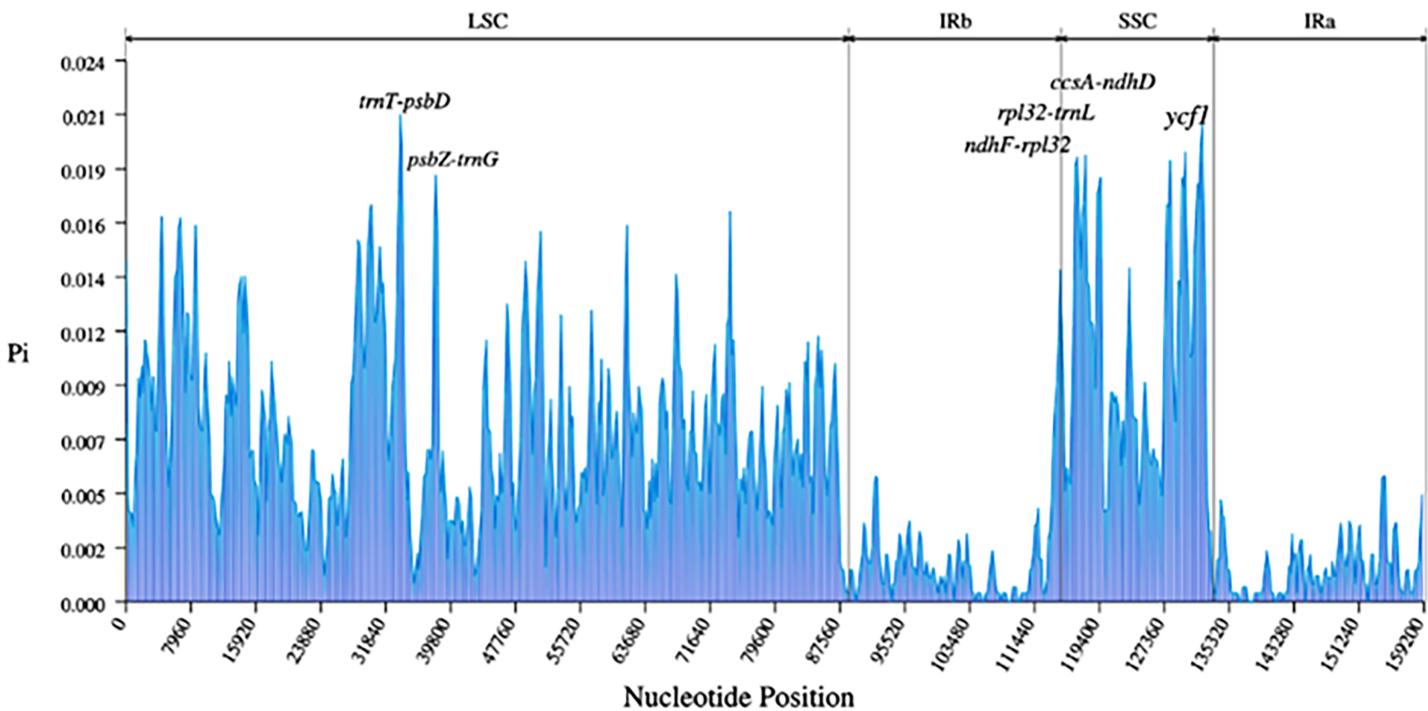


**Figure 6 Nucleotide diversities (*Pi*) of the nine *Ribes* chloroplast genomes presented in a sliding window approach.** X-axis represents the position of the window midpoint; Y-axis represents the nucleotide diversity within each window (window size: 600 bp; step size: 200 bp). Six regions with relatively high *Pi* values were marked based on the alignment and annotation. Full-size 🖼 DOI: 10.7717/peerj.16272/fig-6

intergenic region was *trnT-psbD*, which contains 86 mutation sites, followed by *ndhF-rpl32* (76) and *atpH-atpI* (71). The most variable gene *ycf1*, whose coding-sequence harbors 248 mutation sites, followed by *rpoC2* (104), *ndhF* (56), and *ycf2* (56). In terms of introns, the introns of *trnK* and *matK* contained the most variable mutation sites (57), followed by introns of *rpl16* (38) and *ndhA* (57) (Tables S7 and S8). In the pairwise comparison, the most variation sites were found between *R. roezlii* and *R. fasciculatum* var. *chinense*
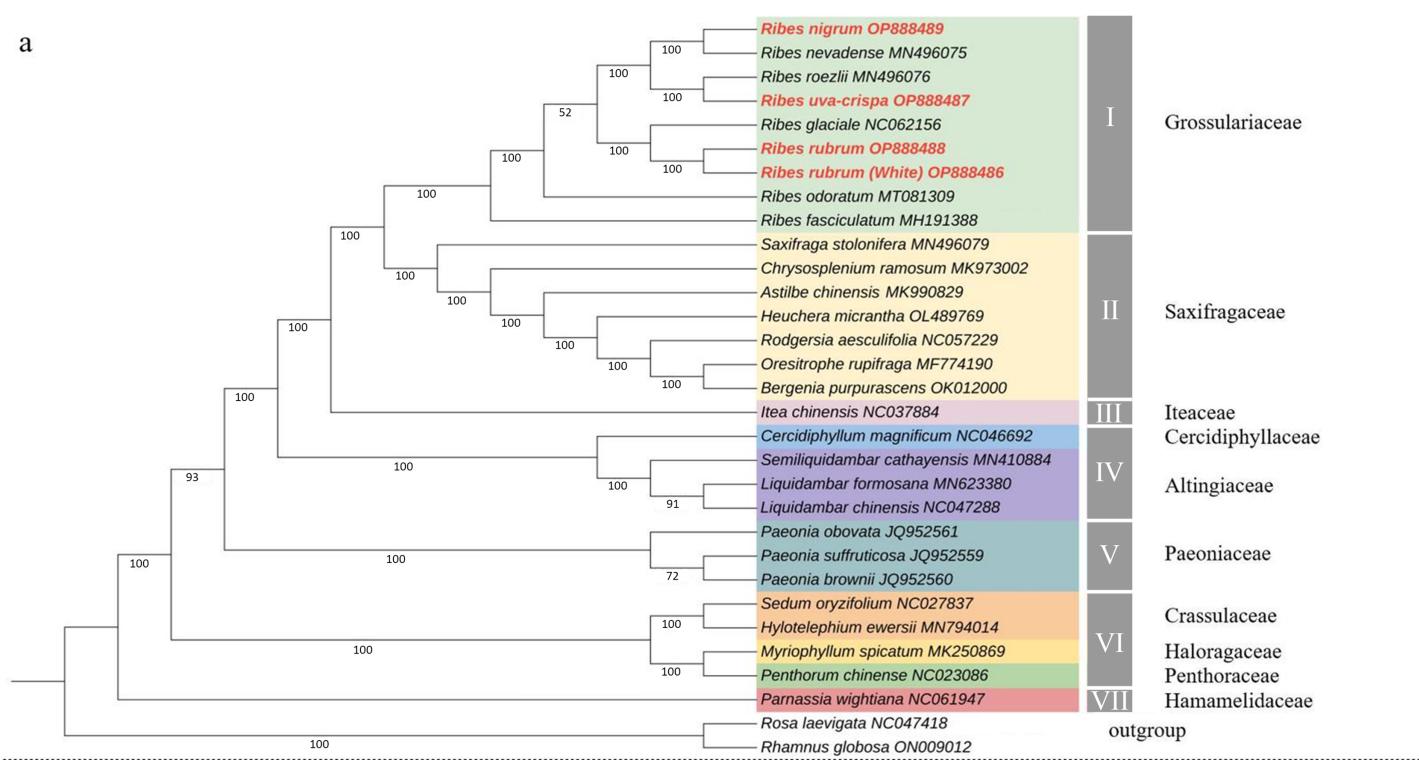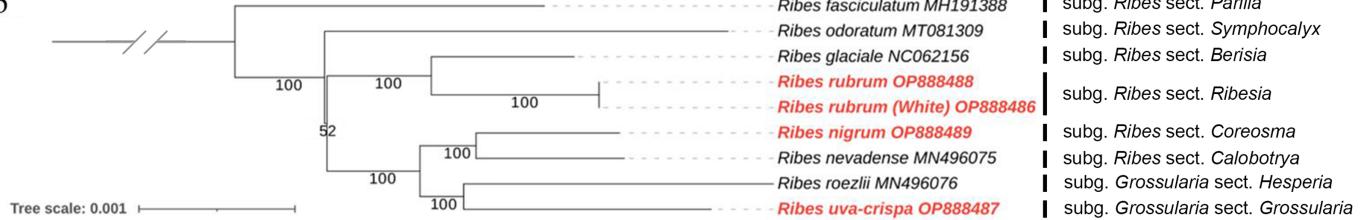
**Figure 7 Chloroplast phylogeny of Saxifragales and *Ribes*.** (A) Maximum likelihood tree constructed using 29 complete cp genomes from Saxifragales and two outgroup species from Rosales with 1,000 bootstrap replications. (B) Maximum likelihood tree inferred from nine complete cp genomes of *Ribes* with 1,000 bootstrap replications. Vertical bars at the right label the section classification of these nine taxa according to the consensus revision of *Berger (1924)*, *Sinnott (1985)*, *Messinger, Hummer & Liston (1999)*, and *Senters & Soltis (2003)*.

Full-size ☒ DOI: 10.7717/peerj.16272/fig-7

(1,863), and the fewest variation sites (507) were detected between blackcurrant (*R. nigrum*) and *R. nevadense* (Table S9). The nucleotide diversity (*Pi*) of the nine *Ribes* cp genomes was visualized using a window length of 600 bp and a step size of 200 bp (Fig. 6). The average *Pi* of whole sequences, SSC, LSC, and IR regions was 0.0060, 0.011, 0.0077, and 0.0015 respectively. Five IGS (*trnT-psbD*, *psbZ-trnG*, *ndhF-rpl32*, *rpl32-trnL*, and *ccsA-ndhD*), and one gene (*ycf1*) exhibited relatively high genetic diversities (*Pi* ≥ 0.018). *TrnT-psbD* was the most polymorphic region with a *Pi* of 0.021.

## Phylogenomic analysis of Grossulariaceae

To determine the phylogenetic position of Grossulariaceae in Saxifragales, ML and Bayesian phylogenetic trees were constructed based on 31 complete cp genomes, including the above-mentioned nine *Ribes*, 20 representative species from Saxifragales and two

outgroup species from Rosales (Table S9). The ML and Bayesian tree topology consensually revealed seven well-supported main clades in Saxifragales (BS ≥ 93% and PP = 100%), including clade I, Grossulariaceae; clade II, Saxifragaceae; clade III, Iteaceae; clade IV, Cercidiphyllaceae and Altingiaceae; clade V, Paeoniaceae; clade VI, Crassulaceae, Haloragaceae, and Penthoraceae; and clade VII, Hamamelidaceae (Figs. 7A and S2). Notably, the nine *Ribes* taxa were clustered independently to form clade I with maximum BS and PP values instead of nesting within Saxifragaceae (clade II), indicating the monophyly of Grossulariaceae. In terms of the *Ribes* genus (Fig. 7B), *R. fasciculatum* (subgenus *Ribes*, section *Parilla*) was first differentiated from other species, followed by *R. odoratum* (subgenus *Ribes*, section *Symphocalyx*). However, the ancestral clade of the remaining seven species was not well-resolved (BS = 52%). In addition, paraphyletic relationships were found between the *Berisia* and *Ribesia* sections, between the *Coreosma* and *Calobotrya* sections, and between the *Hesperia* and *Grossularia* sections. Notably, the *Grossularia* subgenus was nested within the *Ribes* subgenus instead of being parallel to it at the basal of the tree as expected.

## DISCUSSION

The present study for the first time assembled and characterized the complete cp genomes of four cultivated *Ribes* taxa, including blackcurrant (*R. nigrum*), red currant and white currant (*R. rubrum*), and gooseberry (*R. uva-crispa*), which enriches the genomic foundation for *Ribes*. The Grossulariaceae showed relatively high chloroplast conservativeness in terms of sequence length, structural organization, gene content, order, and arrangement than other plant families such as Gentianaceae, Asteraceae, Caprifoliaceae, and the genetic close family, Saxifragaceae (*Walker, Zanis & Emery, 2014*; *Sun et al., 2018*; *Wang et al., 2020*; *Chen et al., 2022*). Intriguingly, the results showed that the IR/SC shift types of the nine *Ribes* cp genomes are in accordance with the phylogenetic groups (Figs. 5, 7B and S2), suggesting the shrinkage and expansion of the IR/SC boundaries might serve as an additional confirmation for phylogenetic results that calculated from the algorithms based on sequences. Notably, the cp genomes of red currant (OP888488) and white currant (OP888486) are 100% identical, which supports the hypothesis: white currant is an albino cultivar of red currant (*Reisch & Pratt, 1996*; *Prange, 2002*; *Lim, 2012*) and calls for further verification based on nuclear data. Repeat elements play an important role in sequence divergence and are involved in plastome rearrangement (*Timme et al., 2007*; *Weng et al., 2014*). Herein, 51/40, 64/42, and 53/54 long-repeats/SSR-loci was respectively identified in *R. rubrum*, *R. uva-crispa*, and *R. nigrum* cp genomes, whereas that of the wild *R. odoratum* is 49/56 (*Wang et al., 2021*). These differences not merely display the characteristic evolutionary trace of these plastomes but also provide effective information for species discrimination and population genetics. Moreover, a total of 3,322 SNPs and 485 Indels were identified in the nine released *Ribes* cp genomes. In particular, the intergeneric spacer of *trnT-psbD*, the CDS of *ycf1*, and the introns of *trnK* and *matK* showed high-resolution potentials. The molecular markers developed from these loci may contribute to genetic evaluations for germplasm collections, conservation genetic research for endangered species or enigmatic taxa, DNA barcoding for endemic

species or important cultivars, and hybrid identification in cross-breeding of *Ribes* in the future.

In terms of the phylogeny of Grossulariaceae, we for the first time used the genomic data of *Ribes* species to clarify the monophyly of Grossulariaceae and the paraphyletic relationship between Grossulariaceae and Saxifragaceae, which are consistent with the results of previous studies about the phylogeny of Saxifragales (*Dong et al., 2013*, *2018*). Moreover, in both ML and Bayesian trees, the highly supported co-ancestral clade of Saxifragaceae and Grossulariaceae and the relatively later divergence between them were revealed, suggesting their close phylogenetic relationship, which is in accordance with the previous studies based on various data types (*Dong et al., 2013*; *Soltis et al., 2013*; *Han et al., 2022*). In terms of the infrageneric relationships, our results generally support the consensus classification of *Ribes* proposed by *Berger (1924)* and revised by *Sinnott (1985)*, that is, the *Ribes* genus includes 12 sections of *Berisia*, *Calobotrya*, *Coreosma*, *Grossularioides*, *Heritiera*, *Parilla*, *Ribesia*, *Symphocalyx*, *Grossularia*, *Hesperia*, *Lobbia* and *Robsonia*, which is also supported by recent studies based on molecular markers (*Messinger, Hummer & Liston, 1999*; *Weigend, Mohr & Motley, 2002*; *Senters & Soltis, 2003*; *Schultheis & Donoghue, 2004*). Although our results revealed a well-supported clade of the subgenus *Grossularia*, only two species (*R. roezlii* and *R. uva-crispa*) in this subgenus were used. Therefore, the monophyletic origin of *Grossularia* cannot be asserted here. However, in any case, the clade of the *Grossularia* subgenus is nested within the *Ribes* subgenus and shows a paraphyletic relationship with the co-ancestor of the *Calobotrya* and *Coreosma* sections either in the present study or in previous studies (*Senters & Soltis, 2003*; *Schultheis & Donoghue, 2004*), which is incongruent to that of the generally accepted subclassification based on morphological characteristics of the *Ribes* genus, regardless of whether gooseberry (*Grossularia*) is monophyletic. However, according to the limited samples and only cp genomic data used in the present study, the thorough clarification of intrageneric phylogeny and infrageneric resolution, evolutional scenario, and diversification history of *Ribes* still call for comprehensive studies based on large-scale sample collection, nuclear genomic data, biogeographic information, and fossil evidence.

## CONCLUSIONS

The complete cp genomes of blackcurrant (*R. nigrum*), red currant and white currant (*R. rubrum*), and gooseberry (*R. uva-crispa*) have lengths of 157,450–157,802 bp and share a total of 131 genes (86 protein-coding genes, 37 tRNA genes, and eight rRNA genes). Notably, a 100% sequence identity was observed between red currant and white currant, supporting the inference that the latter is an albino cultivar of the former. A total of 485 Indels and 3,322 SNPs were identified in the nine released *Ribes* cp genomes. *Ycf1* ($Pi = 0.020$) and *trnT-psbD* ($Pi = 0.021$) are the most polymorphic genic and intergenic regions, respectively. The ML and Bayesian phylogenetic trees provide strong support for the monophyly of Grossulariaceae in Saxifragales and the paraphyletic relationship between Saxifragaceae and Grossulariaceae. The tree topology of the nine *Ribes* taxa shows that the *Grossularia* subgenus is nested within the *Ribes* subgenus, and has a paraphyletic

relationship with the co-ancestor of the *Calobotrya* and *Coreosma* sections. Our findings lay a foundation for the phylogenetic research and molecular breeding of *Ribes* species.

## ACKNOWLEDGEMENTS

## ADDITIONAL INFORMATION AND DECLARATIONS

### Competing Interests

The authors declare that they have no competing interests.

### Author Contributions

- Xinyu Sun analyzed the data, prepared figures and/or tables, and approved the final draft.
- Ying Zhan performed the experiments, prepared figures and/or tables, and approved the final draft.
- Songlin Li analyzed the data, prepared figures and/or tables, and approved the final draft.
- Yu Liu performed the experiments, prepared figures and/or tables, and approved the final draft.
- Qiang Fu performed the experiments, prepared figures and/or tables, and approved the final draft.
- Xin Quan performed the experiments, prepared figures and/or tables, and approved the final draft.
- Jinyu Xiong analyzed the data, prepared figures and/or tables, and approved the final draft.
- Huixin Gang performed the experiments, prepared figures and/or tables, and approved the final draft.

- Lijun Zhang performed the experiments, prepared figures and/or tables, and approved the final draft.
- Huijuan Qi conceived and designed the experiments, prepared figures and/or tables, and approved the final draft.
- Aoxue Wang conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Junwei Huo conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Dong Qin conceived and designed the experiments, authored or reviewed drafts of the article, funding, and approved the final draft.
- Chenqiao Zhu conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, funding, and approved the final draft.

## DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The four complete chloroplast assembly sequences are available at GenBank: OP888489 (blackcurrant), OP888488 (red currant), OP888486 (white currant), and OP888487 (gooseberry).

## Data Availability

The following information was supplied regarding data availability:

The sequencing data is available at NCBI SRA: SRR23008392 (blackcurrant, *R. nigrum*), SRR22955278 (red currant, *R. rubrum*), SRR23008393 (white currant, *R. rubrum*) and SRR22956007 (gooseberry, R. uva-crispa).

## Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.16272#supplemental-information.

## REFERENCES

**Amiryousefi A, Hyvönen J, Poczai P. 2018.** IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **34(17)**:3030–3031 DOI 10.1093/bioinformatics/bty220.

**Ashraf J, Zuo D, Wang Q, Malik W, Zhang Y, Abid MA, Cheng H, Yang Q, Song G. 2018.** Recent insights into cotton functional genomics: progress and future perspectives. *Plant Biotechnology Journal* **16(3)**:699–713 DOI 10.1111/pbi.12856.

**Beier S, Thiel T, Münch T, Scholz U, Mascher M. 2017.** MISA-web: a web server for microsatellite prediction. *Bioinformatics* **33(16)**:2583–2585 DOI 10.1093/bioinformatics/btx198.

**Benson G. 1999.** Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27(2)**:573–580 DOI 10.1093/nar/27.2.573.

**Berger A. 1924.** A taxonomic review of currants and gooseberries. *Bulletin of the New York State Agricultural Experimental Station* **109**:1–118.

**Bolger AM, Lohse M, Usadel B. 2014.** Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30(15)**:2114–2120 DOI 10.1093/bioinformatics/btu170.

Bouckaert R, Vaughan TG, Barido-Sottani J, Duchene S, Fourment M, Gavryushkina A, Heled J, Jones G, Kuhnert D, De Maio N, Matschiner M, Mendes FK, Muller NF, Ogilvie HA, du Plessis L, Popinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu CH, Xie D, Zhang C, Stadler T, Drummond AJ. 2019. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology* **15(4)**:e1006650 DOI 10.1371/journal.pcbi.1006650.

Cao YL, Li YL, Fan YF, Li Z, Yoshida K, Wang JY, Ma XK, Wang N, Mitsuda N, Kotake T, Ishimizu T, Tsai KC, Niu SC, Zhang DY, Sun WH, Luo Q, Zhao JH, Yin Y, Zhang B, Wang JY, Qin K, An W, He J, Dai GL, Wang YJ, Shi ZG, Jiao EN, Wu PJ, Liu XD, Liu B, Liao XY, Jiang YT, Yu X, Hao Y, Xu XY, Zou SQ, Li MH, Hsiao YY, Lin YF, Liang CK, Chen YY, Wu WL, Lu HC, Lan SR, Wang ZW, Zhao X, Zhong WY, Yeh CM, Tsai WC, Van de Peer Y, Liu ZJ. 2021. Wolfberry genomes and the evolution of *Lycium* (Solanaceae). *Communications Biology* **4(1)**:671 DOI 10.1038/s42003-021-02152-8.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25(15)**:1972–1973 DOI 10.1093/bioinformatics/btp348.

Chen Q, Hu HS, Zhang DQ. 2022. DNA barcoding and phylogenomic analysis of the genus *Fritillaria* in China based on complete chloroplast genomes. *Frontiers in Plant Science* **13**:764255 DOI 10.3389/fpls.2022.764255.

Chen ZYF, Yu XL, Yang YJ, Wei P, Zhang WC, Li XZ, Liu CL, Zhao SQ, Li XY, Liu X. 2022. Comparative analysis of chloroplast genomes within Saxifraga (Saxifragaceae) takes insights into their genomic evolution and adaption to the high-elevation environment. *Genes* **13(9)**:1673 DOI 10.3390/genes13091673.

Cortez RE, de Mejia EG. 2019. Blackcurrants (*Ribes nigrum*): a review on chemistry, processing, and health benefits. *Journal of Food Science* **84(9)**:2387–2401 DOI 10.1111/1750-3841.14781.

Coville FV, Britton NL, Gleason HA, Small JK, Pollard CL, Rydberg PA. 1908. *Grossulariaceae*. New York: New York Botanical Garden.

Cronquist A, Takhtadzhian AL. 1981. *An integrated system of classification of flowering plants*. Columbia: Columbia University Press.

Daniell H, Lin C-S, Yu M, Chang W-J. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biology* **17(3)**:1–29 DOI 10.1111/j.1755-0998.2008.02319.x.

Dobrogojski J, Adamiec M, Lucinski R. 2020. The chloroplast genome: a review. *Acta Physiologiae Plantarum* **42(6)**:1–13 DOI 10.1007/s11738-020-03089-x.

Dong WP, Liu J, Yu J, Wang L, Zhou SL. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLOS ONE* **7(4)**:e35071 DOI 10.1371/journal.pone.0035071.

Dong WP, Xu C, Cheng T, Lin K, Zhou SL. 2013. Sequencing angiosperm plastid genomes made easy: a complete set of universal primers and a case study on the phylogeny of Saxifragales. *Genome Biology and Evolution* **5(5)**:989–997 DOI 10.1093/gbe/evt063.

Dong WP, Xu C, Wu P, Cheng T, Yu J, Zhou SL, Hong DY. 2018. Resolving the systematic positions of enigmatic taxa: manipulating the chloroplast genome data of Saxifragales. *Molecular Phylogenetics and Evolution* **126**:321–330 DOI 10.1016/j.ympev.2018.04.033.

Folk RA, Sewnath N, Xiang C-L, Sinn BT, Guralnick RP. 2020. Degradation of key photosynthetic genes in the critically endangered semi-aquatic flowering plant *Saniculiphyllum guangxiense* (Saxifragaceae). *BMC Plant Biology* **20(1)**:1–11 DOI 10.1186/s12870-020-02533-x.

Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Research* **32(Web Server)**:W273–W279 DOI 10.1093/nar/gkh458.

Gao Y, Shen GH, Yuan GH, Tian ZH. 2022. Comparative analysis of whole chloroplast genomes of three common species of *Echinochloa* (Gramineae) in paddy fields. *International Journal of Molecular Sciences* **23(22)**:13864 DOI 10.3390/ijms232213864.

Greiner S, Lehwark P, Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research* **47(W1)**:W59–W64 DOI 10.1093/nar/gkz238.

Hall T, Biosciences I, Carlsbad C. 2011. BioEdit: an important software for molecular biology. *GERF Bulletin of Biosciences* **2**:60–61.

Han SY, Ding HW, Bi D, Zhang SJ, Yi R, Gao JM, Yang JK, Ye YX, Wu LH, Kan XZ. 2022. Structural diversities and phylogenetic signals in plastomes of the early-divergent angiosperms: a case study in Saxifragales. *Plants-Basel* **11(24)**:3544 DOI 10.3390/plants11243544.

Huang YY, Li J, Yang ZR, An WL, Xie CZ, Liu SS, Zheng XS. 2022. Comprehensive analysis of complete chloroplast genome and phylogenetic aspects of ten *Ficus* species. *BMC Plant Biology* **22(1)**:1–15 DOI 10.1186/s12870-022-03643-4.

Hummer KE, Dale A. 2010. Horticulture of *Ribes*. *Forest Pathology* **40(3–4)**:251–263 DOI 10.1111/j.1439-0329.2010.00657.x.

Janczewski E. 1907. Monographie des groseilliers, *Ribes* L. In: *Mémoires de la Société de Physique et d'Histoire Naturelle de Genève*, Vol. 35. Genéve: Impr. W. Kündig & fils, 119–517.

Jiang H, Tian J, Yang JX, Dong X, Zhong ZX, Mwachala G, Zhang CF, Hu GW, Wang QF. 2022. Comparative and phylogenetic analyses of six *Kenya Polystachya* (Orchidaceae) species based on the complete chloroplast genome sequences. *BMC Plant Biology* **22(1)**:1–21 DOI 10.1186/s12870-022-03529-5.

Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biology* **21(1)**:241 DOI 10.1186/s13059-020-02154-5.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30(4)**:772–780 DOI 10.1093/molbev/mst010.

Kuang DY, Wu H, Wang YL, Gao LM, Zhang SZ, Lu L. 2011. Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): implication for DNA barcoding and population genetics. *Genome* **54(8)**:663–673 DOI 10.1139/G11-026.

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids research* **29(22)**:4633–4642 DOI 10.1093/nar/29.22.4633.

Letunic I, Bork P. 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* **49(W1)**:W293–W296 DOI 10.1093/nar/gkab301.

Lim TK. 2012. *Ribes rubrum*. In: *Edible Medicinal and Non-Medicinal Plants*. Dordrecht, Netherlands: Springer, 43–50.

Lingdi L, Alexander C. 2001. Saxifragacea. In: *Flora of China*. Beijing: Science Press, 428–452.

Lukša J, Vepštaitė-Monstavičė I, Yurchenko V, Serva S, Servienė E. 2018. High content analysis of sea buckthorn, black chokeberry, red and white currants microbiota—a pilot study. *Food Research International* **111(8)**:597–606 DOI 10.1016/j.foodres.2018.05.060.

**Määttä K, Kamal-Eldin A, Törrönen R. 2001.** Phenolic compounds in berries of black, red, green, and white currants (*Ribes* sp.). *Antioxidants & Redox Signaling* **3(6)**:981–993 DOI 10.1089/152308601317203521.

**Messinger W, Hummer K, Liston A. 1999.** *Ribes* (Grossulariaceae) phylogeny as indicated by restriction-site polymorphisms of PCR-amplified chloroplast DNA. *Plant Systematics and Evolution* **217(3–4)**:185–195 DOI 10.1007/bf00984364.

**Nascimento FF, dos Reis M, Yang ZH. 2017.** A biologist's guide to Bayesian phylogenetic analysis. *Nature Ecology & Evolution* **1(10)**:1446–1454 DOI 10.1038/s41559-017-0280-x.

**Nock CJ, Waters DLE, Edwards MA, Bowen SG, Rice N, Cordeiro GM, Henry RJ. 2011.** Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnology Journal* **9(3)**:328–333 DOI 10.1111/j.1467-7652.2010.00558.x.

**Olmstead RG, Palmer JD. 1994.** Chloroplast DNA systematics: a review of methods and data analysis. *American Journal of Botany* **81(9)**:1205–1224 DOI 10.1002/j.1537-2197.1994.tb15615.x.

**Orsavová J, Hlaváčová I, Mlček J, Snopek L, Mišurcová L. 2019.** Contribution of phenolic compounds, ascorbic acid and vitamin E to antioxidant activity of currant (*Ribes* L.) and gooseberry (*Ribes uva-crispa* L.) fruits. *Food Chemistry* **284(4)**:323–333 DOI 10.1016/j.foodchem.2019.01.072.

**Parks M, Cronn R, Liston A. 2009.** Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biology* **7(1)**:1–17 DOI 10.1186/1741-7007-7-84.

**Poyarkova A. 1939.** Ribesioideae. *Flora of USSR* **9**:226–270.

**Prange RK. 2002.** Currant, gooseberry and elderberry. *Available at https://www.researchgate.net/publication/237333435_Currant_Gooseberry_and_Elderberry*.

**Reisch B, Pratt C. 1996.** *Fruit breeding: vine and small fruit crops.* Vol. II. New York: John Wiley and Sons.

**Rheder A. 1940.** *Manual of cultivated trees and shrubs.* New York: MacMillan Publishing Co., Inc.

**Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sanchez-Gracia A. 2017.** DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34(12)**:3299–3302 DOI 10.1093/molbev/msx248.

**Schultheis LM, Donoghue MJ. 2004.** Molecular phylogeny and biogeography of *Ribes* (Grossulariaceae), with an emphasis on gooseberries (subg. *Grossularia*). *Systematic Botany* **29(1)**:77–96 DOI 10.1600/036364404772974239.

**Senters AE, Soltis DE. 2003.** Phylogenetic relationships in *Ribes* (Grossulariaceae) inferred from ITS sequence data. *Taxon* **52(1)**:51–66 DOI 10.2307/3647301.

**Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019.** CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Research* **47(W1)**:W65–W73 DOI 10.1093/nar/gkz345.

**Sinnott QP. 1985.** A revision of *Ribes* L. subg. *Grossularia* (Mill.) Pers. Sect. *Grossularia* (Mill.) Nutt.(Grossulariaceae) in North America. *Rhodora* **87**:189–286.

**Soltis DE, Mort ME, Latvis M, Mavrodiev EV, O'Meara BC, Soltis PS, Burleigh JG, de Casas RR. 2013.** Phylogenetic relationships and character evolution analysis of Saxifragales using a supermatrix approach. *American Journal of Botany* **100(5)**:916–929 DOI 10.3732/ajb.1300044.

**Sun SS, Fu PC, Zhou XJ, Cheng YW, Zhang FQ, Chen SL, Gao QB. 2018.** The complete plastome sequences of seven species in *Gentiana* sect. *Kudoa* (Gentianaceae): insights into plastid gene loss and molecular evolution. *Frontiers in Plant Science* **9**:493 DOI 10.3389/fpls.2018.00493.

**Tamura K, Stecher G, Kumar S. 2021.** MEGA11 molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* **38(7)**:3022–3027 DOI 10.1093/molbev/msab120.

**Timme RE, Kuehl JV, Boore JL, Jansen RK. 2007.** A comparative analysis of the *Lactuca* and *Helianthus* (Asteraceae) plastid genomes: identification of divergent regions and categorization of shared repeats. *American Journal of Botany* **94(3)**:302–312 DOI 10.3732/ajb.94.3.302.

**Walker JF, Zanis MJ, Emery NC. 2014.** Comparative analysis of complete chloroplast genome sequence and inversion variation in *Lasthenia Burkei* (Madieae, Asteraceae). *American Journal of Botany* **101**:722–729 DOI 10.3732/ajb.1400049.

**Wang L, Liang J, Sa W, Wang L. 2021.** Sequencing and comparative analysis of the chloroplast genome of *Ribes odoratum* provide insights for marker development and phylogenetics in *Ribes*. *Physiology and Molecular Biology of Plants* **27(1)**:81–92 DOI 10.1007/s12298-021-00932-4.

**Wang HX, Liu H, Moore MJ, Landrein S, Liu B, Zhu ZX, Wang HF. 2020.** Plastid phylogenomic insights into the evolution of the Caprifoliaceae s.l. (Dipsacales). *Molecular Phylogenetics and Evolution* **142**:106641 DOI 10.1016/j.ympev.2019.106641.

**Weigend M. 2007.** Grossulariaceae. In: *Flowering Plants · Eudicots*. Berlin, Heidelberg: Springer, 168–176.

**Weigend M, Mohr O, Motley TJ. 2002.** Phylogeny and classification of the genus *Ribes* (Grossulariaceae) based on 5S-NTS sequences and morphological and anatomical data. *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie* **124(2)**:163–182 DOI 10.1127/0006-8152/2002/0124-0163.

**Weng M-L, Blazier JC, Govindu M, Jansen RK. 2014.** Reconstruction of the ancestral plastid genome in Geraniaceae reveals a correlation between genome rearrangements, repeats, and nucleotide substitution rates. *Molecular Biology and Evolution* **31(3)**:645–659 DOI 10.1093/molbev/mst257.

**Wu ZF, Chen HY, Pan Y, Feng H, Fang DM, Yang J, Wang YY, Sahu SK, Liu JL, Xing YE, Wang XL, Liu M, Luo XY, Gao P, Li LF, Liu ZJ, Yang HM, Liu X, Xu X, Liu H, Wang ET. 2022.** Genome of *Hippophae rhamnoides* provides insights into a conserved molecular mechanism in actinorhizal and rhizobial symbioses. *New Phytologist* **235(1)**:276–291 DOI 10.1111/nph.18017.

**Yu LY, Diao SF, Zhang GY, Yu JG, Zhang T, Luo HM, Duan AG, Wang JP, He CY, Zhang JG. 2022.** Genome sequence and population genomics provide insights into chromosomal evolution and phytochemical innovation of *Hippophae rhamnoides*. *Plant Biotechnology Journal* **20(7)**:1257–1273 DOI 10.1111/pbi.13802.

**Zhang XF, Landis JB, Wang HX, Zhu ZX, Wang HF. 2021.** Comparative analysis of chloroplast genome structure and molecular dating in Myrtales. *BMC Plant Biology* **21(1)**:1–19 DOI 10.1186/s12870-021-02985-9.

**Zhang BS, Yu ZY, Xu ZC, Zheng BJ. 2023.** A phylogenetic and morphological evolution study of *Ribes* L. in China using RAD-seq. *Plants* **12(4)**:829 DOI 10.3390/plants12040829.

**Zhu CQ, Sun XY, Fu Q, Zhan Y, Li SL, Liu Y, Yu M, Qin D, Zhang LJ, Huo JW. 2023.** Complete chloroplast genome and phylogenetic analysis of *Lonicera caerulea* var. *edulis* (Caprifoliaceae). *Mitochondrial DNA Part B-Resources* **8(2)**:314–318 DOI 10.1080/23802359.2023.2180309.