RESEARCH ARTICLE

Phylogenomics analysis of *Scutellaria* (Lamiaceae) of the world

Yinghui Wang[1,2,3], Chao Xu[1,2], Xing Guo[4], Yan Wang[1,2,3], Yanyi Chen[1,2,3], Jie Shen[5], Chunnian He[6], Yan Yu[7] and Qiang Wang[1,2,3*]

Abstract

Background *Scutellaria*, a sub-cosmopolitan genus, stands as one of the Lamiaceae family's largest genera, encompassing approximately 500 species found in both temperate and tropical montane regions. Recognized for its significant medicinal properties, this genus has garnered attention as a research focus, showcasing anti-cancer, anti-inflammatory, antioxidant, and hepatoprotective qualities. Additionally, it finds application in agriculture and horticulture. Comprehending *Scutellaria*'s taxonomy is pivotal for its effective utilization and conservation. However, the current taxonomic frameworks, primarily based on morphological characteristics, are inadequate. Despite several phylogenetic studies, the species relationships and delimitations remain ambiguous, leaving the genus without a stable and reliable classification system.

Results This study analyzed 234 complete chloroplast genomes, comprising 220 new and 14 previously published sequences across 206 species, subspecies, and varieties worldwide. Phylogenetic analysis was conducted using six data matrices through Maximum Likelihood and Bayesian Inference, resulting in a robustly supported phylogenetic framework for *Scutellaria*. We propose three subgenera, recommending the elevation of Section *Anaspis* to subgeneric rank and the merging of Sections *Lupulinaria* and *Apeltanthus*. The circumscription of Subgenus *Apeltanthus* and Section *Perilomia* needs to be reconsidered. Comparative analysis of chloroplast genomes highlighted the IR/SC boundary feature as a significant taxonomic indicator. We identified a total of 758 SSRs, 558 longer repetitive sequences, and ten highly variable regions, including *trnK–rps16*, *trnC–petN*, *petN–psbM*, *accD–psaI*, *petA–psbJ*, *rpl32–trnL*, *ccsA–ndhD*, *rps15–ycf1*, *ndhF*, and *ycf1*. These findings serve as valuable references for future research on species identification, phylogeny, and population genetics.

Conclusions The phylogeny of *Scutellaria*, based on the most comprehensive sample collection to date and complete chloroplast genome analysis, has significantly enhanced our understanding of its infrageneric relationships. The extensive examination of chloroplast genome characteristics establishes a solid foundation for the future development and utilization of *Scutellaria*, an important medicinal plant globally.

Keywords Lamiaceae, *Scutellaria*, Phylogeny, Complete chloroplast genome

*Correspondence:
Qiang Wang
wangqiang@ibcas.ac.cn
Full list of author information is available at the end of the article

Wang *et al. BMC Biology* (2024) 22:185

Page 2 of 19

Background

The sub-cosmopolitan genus *Scutellaria*, one of the Lamiaceae family's largest genera, comprises approximately 500 species, as reported by the Plants of the World Online (POWO) [1]. These species are predominantly distributed in tropical montane and temperate regions. The Irano-Turanian regions, especially Central Asia and the Iranian plateau, represent the genus's maximum diversity center. Additionally, the eastern Mediterranean and the Andes serve as significant centers for its speciation [2, 3]. *Scutellaria* can be distinctly identified at the generic level by its calyx, which features two undivided lips and an appendage that folds into an erect, sail-like structure on the upper lip, known as the scutellum, giving rise to the genus's name. Alternatively, this structure may present as an appendage intumescence rather than a scutellum [2, 4]. Another notable characteristic is the corolla's upper lip, which is usually galeate except for some South American species placed in Sect. *Perilomia* (Kunth) Epling [2]. This genus is renowned for its medicinal properties, with many species being rich in flavonoids, such as *S. baicalensis* Georgi [5], *S. indica* L. [6, 7], and *S. barbata* D.Don [8]. These species have been utilized in medicinal practices across various countries, particularly in China for over 2000 years. As vital sources of raw materials for Chinese medicinal formulations,

they are extensively employed in treating conditions such as hepatitis, jaundice, tumors, leukemia, hyperlipidemia, arteriosclerosis, diarrhea, and inflammatory diseases [9, 10]. The medicinal significance of this genus remains a focal point of research, with modern pharmacology confirming its anticancer, anti-inflammatory, antioxidant, and hepatoprotective properties [8, 11–16]. Additionally, neoclerodane diterpenoids isolated from *Scutellaria* species exhibit significant antifeedant activity, offering a robust defense against herbivory by lepidopterous larvae, benefiting agricultural production [17–20]. Furthermore, numerous *Scutellaria* species hold economic value in the horticultural industry, attributed to their rich colors and broad adaptability, ranging from wetlands to rocky terrains and from coastlines to alpine regions (Fig. 1) [2].

Despite its significant economic value attracted the attention of many botanists and evolutionary biologists, the taxonomy of *Scutellaria* has remained poorly understood. Hamilton [21] first divided *Scutellaria* into three sections: Sect. *Lupulinaria* A. Ham., Sect. *Galericularia* A. Ham., and Sect. *Stachymacris* A. Ham. Following Hamilton, Bentham [22–24] merged Sect. *Galericularia* and Sect. *Stachymacris* into Sect. *Vulgare* Benth., and established Sect. *Heteranthes* Benth. Briquet [25] divided the genus into Subgenus *Scutellariopsis* Briq. and Subg. *Euscutellaria* Briq. including Bentham's three sections.
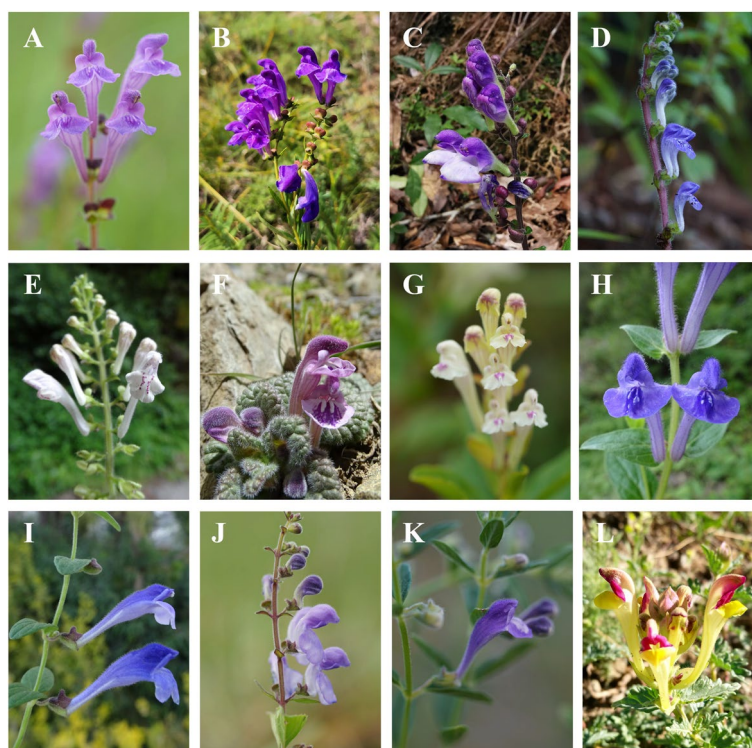

Fig. 1 The representative species of *Scutellaria*. **A** *S. amoena*, **B** *S. baicalensis*, **C** *S. hainanensis*, **D** *S. discolor*, **E** *S. lushuiensis*, **F** *S. kingiana*, **G** *S. likiangensis*, **H** *S. hypericifolia*, **I** *S. rehderiana*, **J** *S. tenax*, **K** *S. strigillosa*, **L** *S. sieversii*

Different taxonomies emphasize different morphological characters. Later, regional treatments mainly followed Briquet's classification system but various taxonomies appeared because the global variation was not taken into account [26–30]. The latest global taxonomy was proposed by Paton [2] based on more comprehensive morphological characters and samples. The number of species in *Scutellaria* is closer to 360 being lower than the 500 reported in POWO as Paton employed a broader species concept and believed that the species concept employed in the *Flora URSS* and *Flora Reipublicae Popularis Sinicae* was too narrow [2]. The genus was divided into Subg. *Scutellaria* and Subg. *Apeltanthus* (Nevski ex Juz.) Juz. *emend*. A.J.Paton. The former subgenus was further divided into five sections: Sects. *Scutellaria*, *Salazaria* (Torrey) A.J.Paton., *Perilomia* (Kunth) Epling *emend*. A.J.Paton, *Anaspis* (Rech.f.) A.J.Paton, and *Salviifoliae* (Boiss.) Edmondson. Subg. *Apelthanthus* was further divided into Sect. *Apeltanthus* Nevski ex Juz. and Sect. *Lupulinaria* A. Ham., and then Sect. *Lupulinaria* was subdivided into Subsect. *Lupulinaria* (A. Ham.) A.J.Paton and Subsect. *Cystaspis* (Juz.) A.J.Paton [2]. Meanwhile, Paton's analysis revealed that both Sect. *Scutellaria* and Sect. *Lupulinaria* Subsect. *Lupulinaria* are paraphyletic. The former encompasses approximately 240 species, tentatively organized into 34 species-groups, while the latter comprises about 120 species, categorized into four informal species-groups. Subsequent research on the classification and evolution of *Scutellaria* has predominantly relied on Paton's classification system.

Several molecular phylogenetic studies have provided insights into the taxonomy and phylogeny of *Scutellaria*. The phylogenetic study including 45 samples (34 taxa from China) by Zhao et al. [31] using ITS and ETS suggested that two clades were supported and clade I including *S. galericulata* L., *S. diffusa* Benth. and *S. nuristanica* Rech.f. at the base. Subgenus *Apeltanthus* was sister to *S. likiangensis* Diels allies and embedded into Subg. *Scutellaria*, so Subg. *Scutellaria* was considered to be paraphyletic. Although Subg. *Apeltanthus* clustered into a clade, the monophyly of the subgenus could be questioned because only five species of that taxon were investigated. In the study of Safikhani et al. [32], samples focused on Subg. *Apeltanthus*. Using ITS and *trnL-F* of 42 samples (36 samples from Iran), the phylogeny of *Scutellaria* presented by Safikhani et al. [32] showed that the members of Subg. *Apeltanthus* (31 samples) formed a clade, but Sect. *Lupulinaria* was non-monophyletic because *S. stocksii* Boiss. (Sect. *Apeltanthus*) was nested within it. Sect. *Scutellaria* was also non-monophyletic because *S. galericulata* was a sister to a clade including species of Sects. *Anaspis* and *Scutellaria*. However, the phylogenetic research based on 44 species of *Scutellaria* using

ITS and *trnL-F* by Seyedipour et al. [33] considered that Subg. *Apeltanthus* was paraphyletic because *S. repens* Buch.-Ham. ex D.Don (Sect. *Scutellaria*) was embedded within it. Subg. *Scutellaria* was also supported to be paraphyletic due to Subg. *Apeltanthus* nested within it. Also, Sect. *Scutellaria* and Sect. *Lupulinaria* were not monophyletic. The latest phylogeny of *Scutellaria* including 76 taxa from the sections of Paton also showed that Subg. *Scutellaria* was paraphyletic to Subg. *Apeltanthus* based on three plastid markers (*matK-trnK*, *rpl16*, *trnL-F*), and three clades were well-supported resolved [34]. *S. pontica* K. Koch and a clade including *S. baicalensis*, *S. amoena* C.H.Wright, *S. rehderiana* Diels, and *S. kingiana* Prain form an unresolved polytomy with Subg. *Apelthanthus*, and *S. schweinfurthii* Subsp. *paucifolia* (Baker) A.J.Paton is sister to that polytomy. Further, Sect. *Salviifoliae* was not monophyletic due to the different phylogenetic position between *S. diffusa* and *S. pontica*. Sect. *Lupulinaria* was also not monophyletic because *S. poecilantha* Nevski ex Juz. (Sect. *Apeltanthus*) was nested within it, and the members of Subsect. *Lupulinaria* were scattered into different subclades. The study also showed that *S. kingiana* belonging to Sect. *Anaspis* was sister to an East Asian lineage including *S. amoena*, *S. rehderiana*, and *S. baicalensis* (Sect. *Scutellaria*). *S. galericulata* placed in Sect. *Scutellaria* was sister to *S. minor* Huds. Recently, research on phylogenetic relationships within *Scutellaria* using the complete chloroplast genome (cpDNA) was also processed with a few species [35, 36]. The results consistently showed that *S. kingiana* belonging to Sect. *Anaspis* formed a clade with *S. baicalensis* and allies, and then was sister to Subg. *Apeltanthus*, which was consistent with the results by Salimov et al. [34] and Zhao et al. [31]. These studies have significantly enhanced our understanding of the phylogeny of *Scutellaria*. However, the phylogenetic relationships within the genus are still not well understood, primarily due to a lack of samples or molecular markers with insufficient phylogenetic loci. There is a critical need for a well-resolved phylogeny of the genus based on comprehensive global sampling.

Although unclear infrageneric relationships of *Scutellaria*, a series of phylogenetic studies strongly supported the monophyly of *Scutellaria* [37–41]. Despite some of the infrageneric taxa have been given generic rank in the past, such as *Anaspis* Rech.f., *Perilomia* Kunth, *Salazaria* Torrey, *Harlanlewisia* Epling, *Theresa* Clos, and *Cruzia* Phil., the morphological and molecular phylogenetic studies revealed that these taxa belong to the genus *Scutellaria* [2, 3, 27, 28]. Recent studies showed that six genera were circumscribed in the subfamily Scutellarioideae, which is characterized by abundant calyx fibers and pericarps with tuberculate or elongate outgrowths [37, 38, 41–47]. *Scutellaria* is the largest genus, followed

by *Tinnea* (including 19 species) which is endemic to Africa [43]. The other four genera including *Holmskioldia* Retz. [44], *Wenchengia* C.Y. Wu & S. Chow [45], *Renschia* Vatke [43, 46], and *Heliacria* Bo Li, C.L. Xiang, T.S. Hoang & Nuraliev [47] are all monospecific, which are distributed to the southern Himalayas, Hainan province of China and Vietnam, northern Somalia and Vietnam respectively. In the present study, the three genera (including *Tinnea*, *Holmskioldia*, and *Wenchengia*) closely related to *Scutellaria* were sampled as outgroups [37–41].

Since the first cpDNA of tobacco was reported by Shinozaki et al. [48], an increasing number of cpDNA sequences have been published. Until June 2023, over 3800 cpDNA sequences have been published in the National Center for Biotechnology Information (NCBI) organelle genome database. However, only 21 cpDNA sequences have been published in *Scutellaria*, and the number of species is too few to discuss the phylogenetic relationships based on omics data for the whole genus. It is reported that the published chloroplast genome of *Scutellaria* is a typical quadripartite structure containing two identical copies of inverted repeats (IRa and IRb, IRs) separated by a small single-copy region (SSC) and a large single-copy region (LSC), with a length from 151 to 154 kb, which is consistent with the characteristics of most photosynthetic land plant plastid chromosomes with a size from 120 to 160 kb [49–51]. The gene content and order of cpDNA within this genus are highly conserved comprising 112–114 distinct genes: 79–80 protein-coding genes, 29–30 transfer RNAs (tRNAs), and four ribosomal RNAs (rRNAs), similar to most terrestrial plants [51, 52]. The properties of stable maternal inheritance, highly conserved sequences, low mutation and recombination rates, as well as small genome make the cpDNA become an ideal model to study taxonomy and evolution, and many scholars have used it to study the evolutionary relationships of different taxa [52–57]. For *Scutellaria*, however, most studies based on a few molecular markers mainly focus on species identification and phylogeny of some taxonomic complexes in recent two decades [58–60]. The construction of phylogenetic trees was mainly based on ITS or chloroplast DNA fragments [31–34]. Phylogenetic relationships on the genus using the complete chloroplast genome were carried out with a few species by Zhao et al. [36] and Shan et al. [35]. Extensive sampling and more genomic data in improving phylogenetic inference are helpful [61, 62]. In this study, we acquired 234 cpDNA sequences encompassing 206 species, subspecies, and varieties of *Scutellaria* from regions across Eurasia, Americas, Africa, and Oceania. Additionally, four samples from three genera within the Scutellarioideae—*Tinnea aethiopica* Kotschy ex Hook. f.,

Wenchengia alternifolia C.Y. Wu & S. Chow, and *Holmskioldia sanguinea* Retz.—were included as outgroups. Utilizing this extensive dataset, our objectives were to (1) enhance the resolution of the phylogenetic framework for *Scutellaria*; (2) describe and compare the structure and gene content of *Scutellaria* cpDNA sequences; (3) identify candidate DNA markers within the cpDNA sequences for efficient and accurate species identification, with a focus on medicinal species and potential adulterants, and to facilitate future phylogenetic and phylogeographic studies.

Results

Chloroplast genome characteristics of *Scutellaria*

In the present study, a total of 220 samples representing 196 species, subspecies, and varieties of *Scutellaria* were sequenced successfully for the first time. Together with sequences obtained online, the size of the 234 cpDNA sequences ranged from 150,867 bp (*S. incana* Biehler) to 154,253 bp (*S. coerulea* Moc. & Sessé ex Benth.), exhibiting a variation of 3386 bp (Fig. 2). All chloroplast genomes showed a typical quadripartite structure, consisting of a pair of inverted repeats (IRa and IRb) regions (24,850–25,633 bp) separated by a large single-copy (LSC) region (83,245–86,276 bp) and a small single-copy (SSC) region (17,205–17,582 bp). The total guanine-cytosine (GC) content across all cpDNA sequences was relatively consistent, ranging from 38.19% (*S. arabica* Jaub. & Spach) to 38.54% (*S. scutellarioides* (Kunth) Harley), with IR regions exhibiting significantly higher GC content (43.49–43.72%) compared to the LSC region (36.17–36.65%) and the SSC region (31.89–32.95%) (Additional file 1: Table S1).

Typically, the *Scutellaria* chloroplast genomes contained 112–115 coding genes, including 79–81 protein-coding genes, 29–30 transfer RNA (tRNA), and four ribosomal RNA (rRNA) genes, arranged in the same sequence (Additional file 1: Table S1). Most protein-coding genes were present in a single copy, except for *ndhB*, *rpl2*, *rpl23*, *rps7*, *ycf2*, and *ycf15*, which had two copies. All rRNA genes were duplicated, and a relatively high proportion of tRNA genes (7/30) had two copies, including *trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnN-GUU*, *trnR-ACG*, and *trnV-GAC*. While most genes lacked introns, nine protein-coding genes (*ndhA*, *ndhB*, *petB*, *petD*, *atpF*, *rpl16*, *rpl2*, *rps16*, *rpoC1*) and six tRNA genes (*trnA-UGC*, *trnG-UCC*, *trnI-GAU*, *trnK-UUU*, *trnL-UAA*, *trnV-UAC*) contained one intron. Furthermore, the *clpP*, *ycf3*, and *rps12* genes had two introns, marking the highest intron count among all genes. Notably, the *rps12* gene underwent a trans-splicing event, with the 5'-*rps12* (including exon 1) located in the LSC
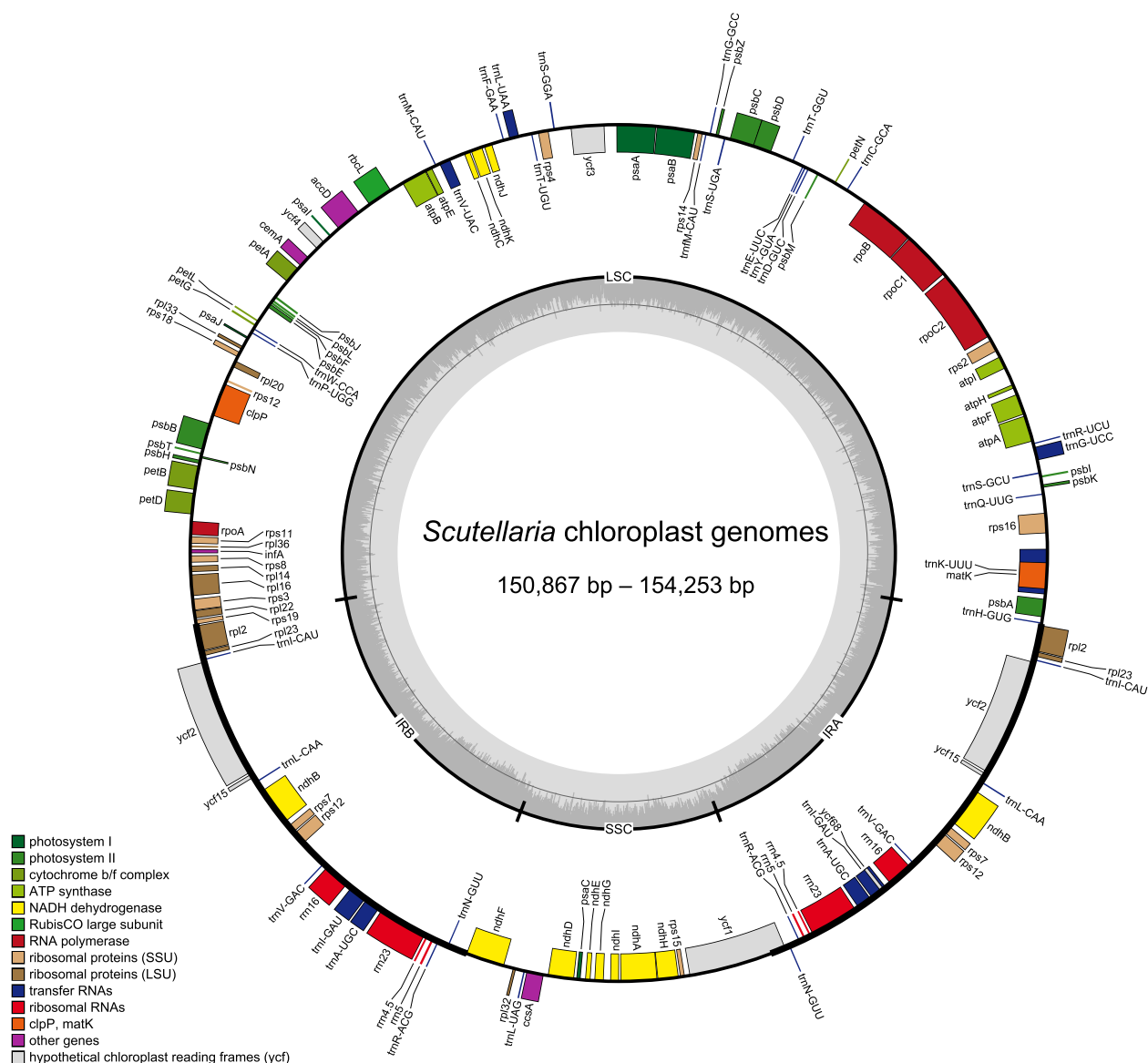
Fig. 2 Complete chloroplast genome map of *Scutellaria*. Genes inside the circle are transcribed clockwise, while those outside the circle are transcribed counterclockwise. The darker gray inner circle shows the GC content, while the lighter represents the AT content. Different colors represent different types of genes. LSC, large single copy; SSC, small single copy; IRA, inverted repeat region A; IRB, inverted repeat region B

region and the 3'-*rps12* (containing exon 2, exon 3, and an intron) generally situated in the IR region (Additional file 1: Table S2).

Phylogenetic analysis

Following alignment adjustments, the lengths of three sequence matrices were determined: 78,001 bp for 81 protein-coding genes, 30 tRNA genes, and four rRNA genes (PC), 59,245 bp for 134 non-coding regions (NC), and 137,246 bp for a concatenation matrix of 115 genes and 134 non-coding regions (PCN). Based on the unpartitioned and partitioned strategies (PC, NC, PCN, PC-p, NC-p, PCN-p) using Maximum Likelihood (ML) and Bayesian Inference (BI), 12 phylogenetic trees obtained a largely consistent topology and the relationships within eight clades were well-resolved with strong support (Fig. 3, Additional file 2: Figs. S1–S11). Clade 1 emerged as the earliest diverging group. Clades 2, 3, and 4 formed a monophyletic group and were identified as sister to another monophyletic branch that included Clades 5, 6, 7, and 8. Given comprehensive loci and higher resolution, the phylogenetic tree based on PCN-p combined
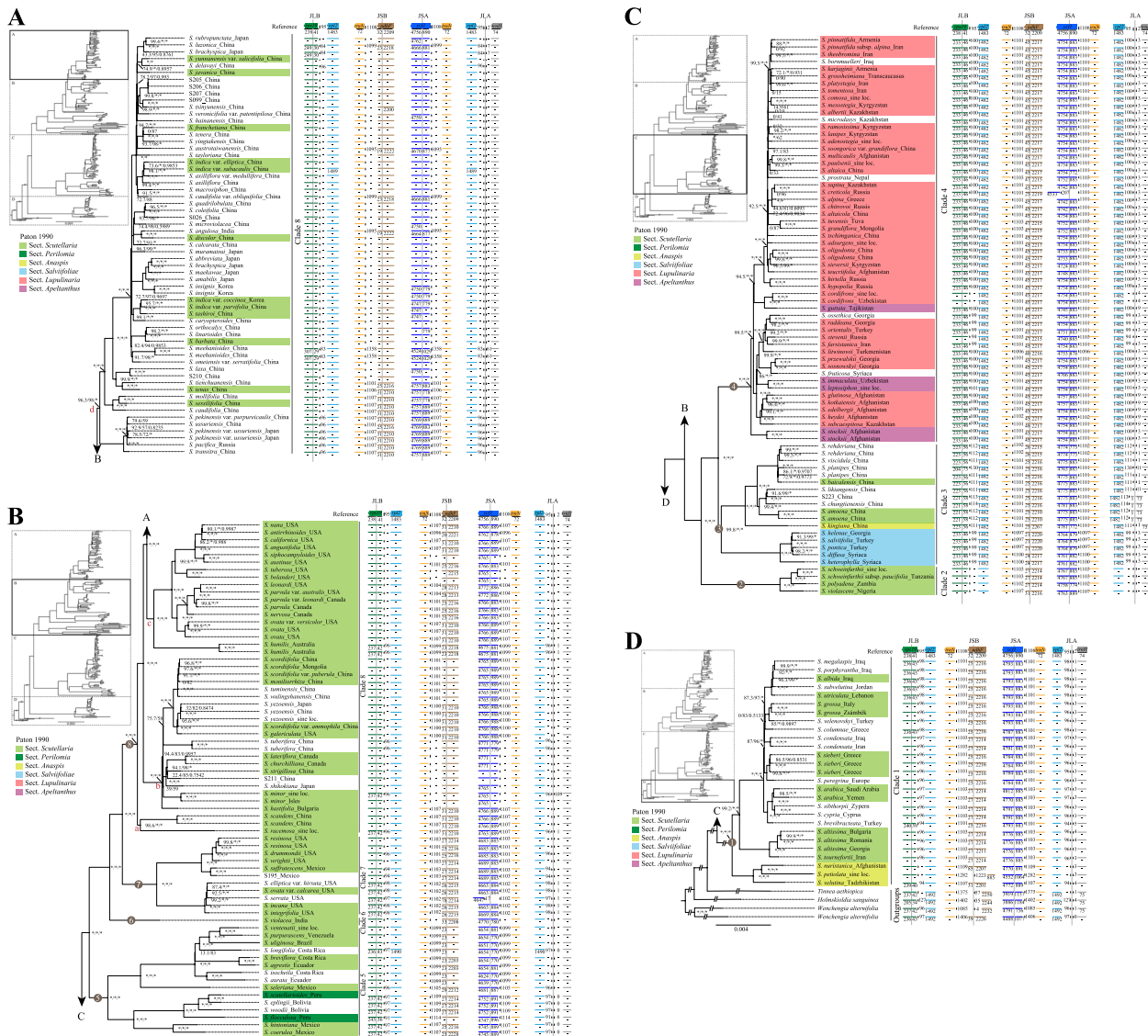
Wang *et al. BMC Biology* (2024) 22:185

Page 6 of 19

Fig. 3 The phylogenetic tree of *Scutellaria* combined Maximum likelihood and Bayesian inference using a partition matrix of 115 genes and 134 non-coding regions (PCN-p). The support values of bootstrap support (BS), SH-like approximate likelihood ratio (SH-aLRT), and posterior probabilities (PP) were shown in turn near the nodes. Only the support rates of BS and SH-aLRT were shown when the topologies of ML and BI trees were inconsistent. Values equal to 100% or 1 were replaced with asterisks. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. The boundary genes of different parts in each cpDNA were mapped to the topology. The "–" showed consistency with the length of the reference and the "↑" represented the distance between the boundary gene and the nearest boundary

ML and BI methods is presented in Fig. 3. The monophyletic *Scutellaria* defined by Paton [2] was strongly supported by all trees (BS = 100%, SH = 100%, PP = 1; Fig. 3, Additional file 2: Figs. S1–S11). *Tinnea* as an Africa endemic genus is closer related to *Scutellaria* with the maximum support (BS = 100%, SH = 100%, PP = 1) in Scutellarioideae. Members of the Subgenus *Scutellaria*, as defined by Paton [2], were distributed across all clades except Clade 4 belonging to Subg. *Apeltanthus*. The

phylogeny indicated that Subg. *Apeltanthus* was completely nested within Subg. *Scutellaria*, and suggested that Subg. *Scutellaria* is not monophyletic and requires significant reclassification. Furthermore, the largest section Subg. *Scutellaria* Sect. *Scutellaria* was identified as polyphyletic since its members were scattered across different clades, excluding Clade 4 (Fig. 3).

Clade 1 (27 accessions) was identified as the earliest diverging group. Within this clade, Section *Anaspis*,

excluding *S. kingiana*, was sister to a lineage of *S. megalaspis* Rech.f.–*S. tournefortii* Benth. from Sect. *Scutellaria*, receiving strong support (BS=100%, SH=100%, PP=1). Sect. *Anaspis* was not monophyletic because *S. kingiana* (Sect. *Anaspis*) was located in Clade 3. *S. altissima* L. was sister to *S. tournefortii*, both related to *S. brevibracteata* Stapf allies with high support (BS=99.2%, SH=100%, PP=1). The lineage of *S. megalaspis*–*S. peregrina* L. (Sect. *Scutellaria*) was resolved with moderately supported (BS=87%, SH=96%, PP=1).

Clades 2, 3, and 4 collectively formed a monophyletic group, serving as a sister group to another monophyletic group comprising Clades 5, 6, 7, and 8. Clade 2 (4 accessions), belonging to Section *Scutellaria*, featured species endemic to Africa, with well-resolved phylogenetic relationships among *S. schweinfurthii* Briq., *S. schweinfurthii* subsp. *paucifolia*, *S. polyadena* Briq., and *S. violascens* Gürke (BS=100%, SH=100%, PP=1). In Clade 3 (17 accessions), two lineages were disjunct in distribution between Eastern Mediterranean to Turkey (Sect. *Salviifoliae*) and China (*S. kingiana* and relatives), but the sister relationships between them were highly supported (BS=99.8%, SH=100%, PP=1). The members of Sect. *Salviifoliae* were strongly supported clustering into a monophyletic group (BS=100%, SH=100%, PP=1). *S. kingiana* (Sect. *Anaspis*) was placed as a sister to an eastern Asia lineage of *S. baicalensis* alliances (Sect. *Scutellaria*), so the location of the species needs to be reconsidered. Subgenus *Apeltanthus* mainly distributed in Central Asia was clustered into Clade 4 (56 accessions), which was resolved as monophyly. However, Sects. *Apeltanthus* and *Lupulinaria* were polyphyletic because *S. guttata* Nevski ex Juz. (Sect. *Apeltanthus*) was sister to *S. cordifrons* Juz. (Sect. *Lupulinaria*) with strong support (BS=100%, SH=100%, PP=1), and *S. immaculata* Nevski ex Juz. (Sect. *Apeltanthus*) as sister to *S. leptosiphon* Nevski (Sect. *Apeltanthus*) were embedded within Sect. *Lupulinaria*. These members of Sect. *Apeltanthus* appeared in different positions from *S. stocksii* (Sect. *Apeltanthus*) which is sister to the rest of Subg. *Apeltanthus* with high support (BS=100%, SH=100%, PP=1).

Within Section *Perilomia*, *S. scutellarioides* and *S. flocculosa* Epling & Mathias formed a clade with *S. eplingii* Legname and *S. woodii* J.M.Mercado unplaced by Paton [2], then together these species were sister to Mexico group of *S. hintoniana* Epling and *S. coerulea* of Sect. *Scutellaria* in Clade 5 (15 accessions). Two main subclades were branched in Clade 5. The one subclade where Sect. *Perilomia* was located as sister to the other subclade of *S. seleriana* Loes. and its relatives with the maximum support (BS=100%, SH=100%, PP=1). All members in this clade are mainly distributed in neotropical regions. Single species *S. violacea* B.Heyne ex Benth. was defined as Clade 6 and the sister relationships among Clades 6, 7, and 8 were strongly supported (BS=100%, SH=100%, PP=1). *S. violacea* is mainly distributed in the Indian Peninsula, Indochinese Peninsula to East Asia. There were two subclades in Clade 7 (11 accessions) and the phylogenetic relationships between them were strong support (BS=100%, SH=100%, PP=1). The members of this clade belonging to Sect. *Scutellaria* are mainly distributed in USA, except for *S. suffrutescens* S.Watson from Northeast Mexico. Clade 8 (103 accessions) included the most species of Sect. *Scutellaria* and was subdivided into Clades 8a, 8b, 8c, 8d. *S. scandens* D.Don mainly distributed in East Asia was sister to *S. racemosa* Pers. grown in the New World with high support in Clade 8a (BS=98.6%, SH=100%, PP=1). The members of Clade 8b are mainly distributed in East Asia and North America, except for *S. galericulata* widely found in the Northern Hemisphere, *S. minor* native to West Europe, and *S. hastifolia* L. widely distributed in Mid-West Europe. In Clade 8c, *S. humilis* R.Br. from Australia was sister to a North American lineage of *S. nana* A.Gray and its relatives with strong support (BS=100%, SH=100%, PP=1). The remaining members of Clade 8 were mostly from East Asia, particularly China. The lineage of *S. rubropunctata* Hayata–*S. meehanioides* C.Y.Wu was not fully resolved, appearing in a polytomy.

Although the phylogenetic trees largely agreed, the positions of some species within Clades 4 and 8 varied due to weakly supported relationships among several species, resulting in some species being grouped in polytomies (Fig. 3, Additional file 2: Figs. S1–S11). The position of *S. minor*, *S. hastifolia*, and *S. tuberifera* C.Y.Wu & C.Chen in Clade 8 differed in different analyses. *S. minor* was sister to *S. hastifolia* using the partitioned NC matrices with ML and BI methods (BS=92.2%, SH=73%, PP=0.9587), similar to the ML phylogeny of unpartitioned NC matrices with moderate support (BS=85.5%, SH=73%). The ML phylogeny based on PCN, PCN-p, and PC-p showed that *S. minor* was sister to *S. shikokiana* Makino allies with low support. However, the phylogenetic tree based on BI suggested *S. tuberifera* was sister to *S. shikokiana* with high support using PCN and PCN-p. Additionally, *S. immaculata* and *S. leptosiphon* formed a clade, being sister to a lineage of *S. ossethica* Kharadze–*S. fruticosa* P.Willemet based on NC, NC-p, PCN, and PCN-p matrices with ML or BI methods, while phylogenies based on PC and PC-p showed the lineage of *S. immaculata* and *S. leptosiphon* remained in a tritomy.

Comparative analyses of cpDNA sequences of the genus
Sequence variation analysis

Comparative chloroplast genomic analyses were conducted on 18 representative species spanning various branches. The genomic collinearity analysis revealed that the cpDNA sequences of these 18 species were highly conserved, with no gene inversion or rearrangement events detected (Additional file 3: Fig. S12). Multiple sequence alignment maps (Additional file 3: Fig. S13) and nucleotide polymorphism analysis consistently found that IR regions were relatively more conserved than LSC and SSC regions. The pi values in IR regions were significantly lower than 0.01, with a maximum value of 0.007, while nearly 49.2% of pi values in LSC were higher than 0.01, with the highest value reaching 0.038 and 79.3% pi values exceeding 0.01 in SSC with a maximum value of 0.048 (Additional file 1: Table S3). In addition, multiple sequence alignment maps showed that the cpDNA sequences were more highly divergent in the intergenic spacers compared to coding regions. The nucleotide polymorphism analysis identified ten notably hyper-variable regions, including *trnK–rps16*, *trnC–petN*, *petN–psbM*, *accD–psaI*, *petA–psbJ*, *ndhF*, *rpl32–trnL*, *ccsA–ndhD*, *rps15–ycf1*, and *ycf1*, with pi values exceeding 0.025. Among these, only two loci were located in coding regions (Fig. 4).

IR/SC boundary feature

All 234 cpDNA sequences of *Scutellaria* contained the same boundary genes that junction of LSC/IRb (JLB) within *rps19*, junction of IRb/SSC (JSB) within *ndhF* or *trnN–ndhF*, junction of SSC/IRa (JSA) within *ycf1* or *ycf1–trnN*, and junction of IRa/LSC (JLA) within *trnH* or *rpl2–trnH*. For JSB, 233 cpDNA sequences (99.6%) were located in *ndhF*, and discovered that this gene of *S. petiolata* Hemsl. ex Lace & Prain with only 885 bp was significantly shorter than others, which resulted in a distance of 1223 bp from the JSB. In addition, *ycf1* of *S. serrata* Andrews and *S. creticola* Juz. were slightly shorter, not crossing the JSA with a distance of 17 and 207 bp from the JSA, respectively. Only JLA was not within genes except for the cpDNA sequences of S223, *S. chungtienensis* C.Y.Wu, and *S. amoena* with *trnH* spanning IRa and LSC regions.

The location of the boundaries varied in different taxa. *Rpl2* in IRs had a distance of 95 bp from JLA or JLB in Clade 8 with 86 cpDNA sequences (83.5%), which were also present in Clades 2 and 6. However, the main boundary distance was 100 bp in Clade 4 with 48 cpDNA sequences (85.7%) (Fig. 3). *TrnN*, the other gene located in the IRs, was distant from JLA and JLB with 1101 bp in Clade 4 accounting for 89.3%, but only 12 cpDNA sequences (11.7%) in Clade 8. The *trnN* was distant from the boundaries with 1108 bp in 60 cpDNA sequences of Clade 8 (58.3%), followed by 1107 bp with 18 cpDNA sequences (17.5%). Additionally, in the LSC, the distance of 2 bp between *trnH* and JLA in all cpDNA sequences of Clades 2 and 6, was also present in 95 cpDNA sequences (92.2%) of Clade 8. But 3 bp was the most common boundary distance in Clade 4 with 45 cpDNA sequences (80.3%). Another gene, *rps19* spanned LSC and IRb regions and the length located in LSC was 233 bp in Clade 4 with 53 cpDNA sequences (94.6%), but it was mainly 238 bp in Clade 1 and Clade 8, having 19 cpDNA
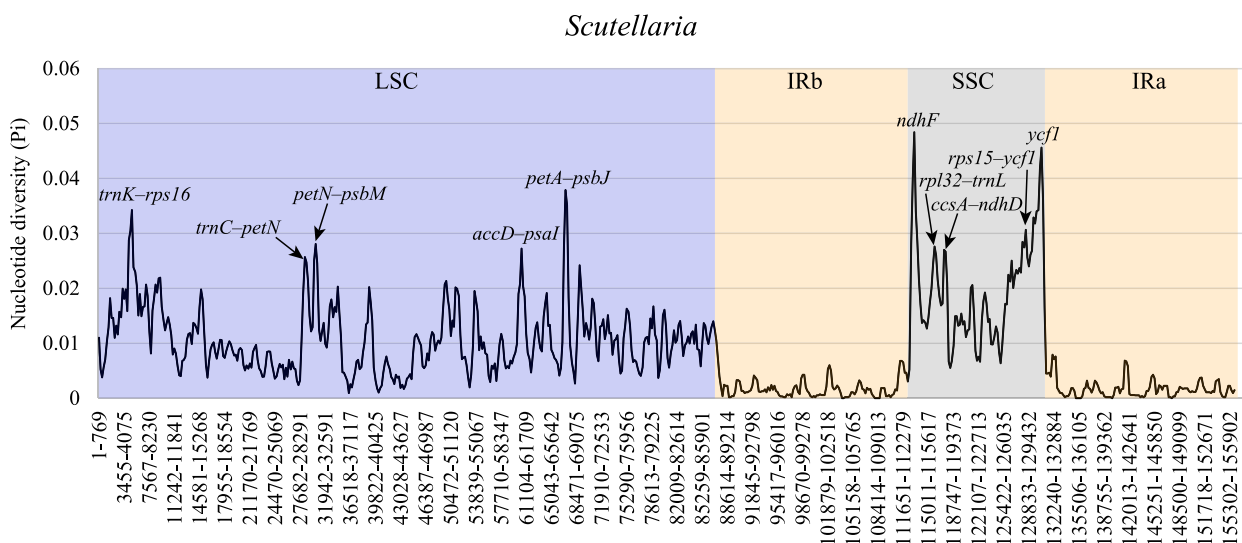

Fig. 4 Sliding window analysis of cpDNA sequences from 18 representative species randomly selected covering all clades of *Scutellaria*. X-axis: the length of the windows, Y-axis: nucleotide diversity within each window (window length: 600 bp, step size: 200 bp)

Wang *et al. BMC Biology* (2024) 22:185

Page 9 of 19

sequences (70.4%) and 95 cpDNA sequences (92.2%), respectively. Besides, *ndhF* also crossed two regions, SSC and IRb, and almost all of *ndhF* had 45 bp located in IRb of Clade 4 (87.5%), but the length in other clades varied greatly. Mapping the boundary feature of all cpDNA sequences to the phylogenetic trees, a detailed schematic diagram was drawn in Fig. 3.

SSRs and repetitive sequences

A total of 758 SSRs were identified, with 593 (78.2%) located in the LSC region, significantly outnumbering those in the SSC and IR regions, which contained 107 (14.1%) and 58 (7.7%) SSRs, respectively. Furthermore, 465 SSRs (61.3%) were found in intergenic spacers (IGS), markedly exceeding the numbers in introns and CDS, with 158 (20.8%) and 135 (17.8%) SSRs, respectively (Additional file 1: Table S4). Six types of SSRs were investigated including mononucleotide, dinucleotide, trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide. Mononucleotide repeats were the most abundant with 456 SSRs (60.2%), followed by tetranucleotide having 124 SSRs (16.4%), and pentanucleotide was the least with 5 SSRs only present in *S. parvula* Michx., *S. galericulata*, *S. violacea*, and *S. petiolata* but hexanucleotide was absent in all species (Fig. 5, Additional file 1: Table S5). In terms of repeat types, the majority

motif of mononucleotide repeats was A/T with 390 SSRs, accounting for 85.5%, and dinucleotide only had an AT/TA motif (Additional file 1: Table S6). These repeat types were present in all species. Additionally, repeat units of GAA/TTC, ATA/TAT, and AATA/TATT were also exited in all species. Conversely, AATG/CATT, AAGA/TCTT, and TTCT/AGAA of tetranucleotide repeats were unique motifs in *S. scandens*, *S. barbata*, and *S. scutellarioides*, respectively. And rare pentanucleotide repeats including AAAAG/CTTTT, GAATA/TATTC, CTTAT/ATAAG, AAATA/TATTT, and ATAAA/TTTAT were exclusive to *S. parvula*, *S. galericulata*, *S. petiolate*, and *S. violacea*, respectively (Additional file 1: Table S7).

A total of 558 repetitive sequences were identified, categorized into four types: palindromic (P), forward (F), reverse (R), and complement (C), among which the most abundant repeat type of P with 297 repeats (53.2%), followed by F with 244 repeats (43.7%) that were widely distributed in all species (Fig. 5, Additional file 1: Table S8). However, R type was detected in *S. scandens*, *S. incana*, *S. oligodonta* Juz., and *S. stocksii* and C type was only found in *S. petiolata* (Additional file 1: Tables S8–S9). The abundant repeats were located in LSC with 306 repeats (54.8%), followed by IRs with 140 repeats, but the repeats were the least in SSC (Additional file 1: Table S9). These sequences ranged from 30 to 1096 bp, with repeats of
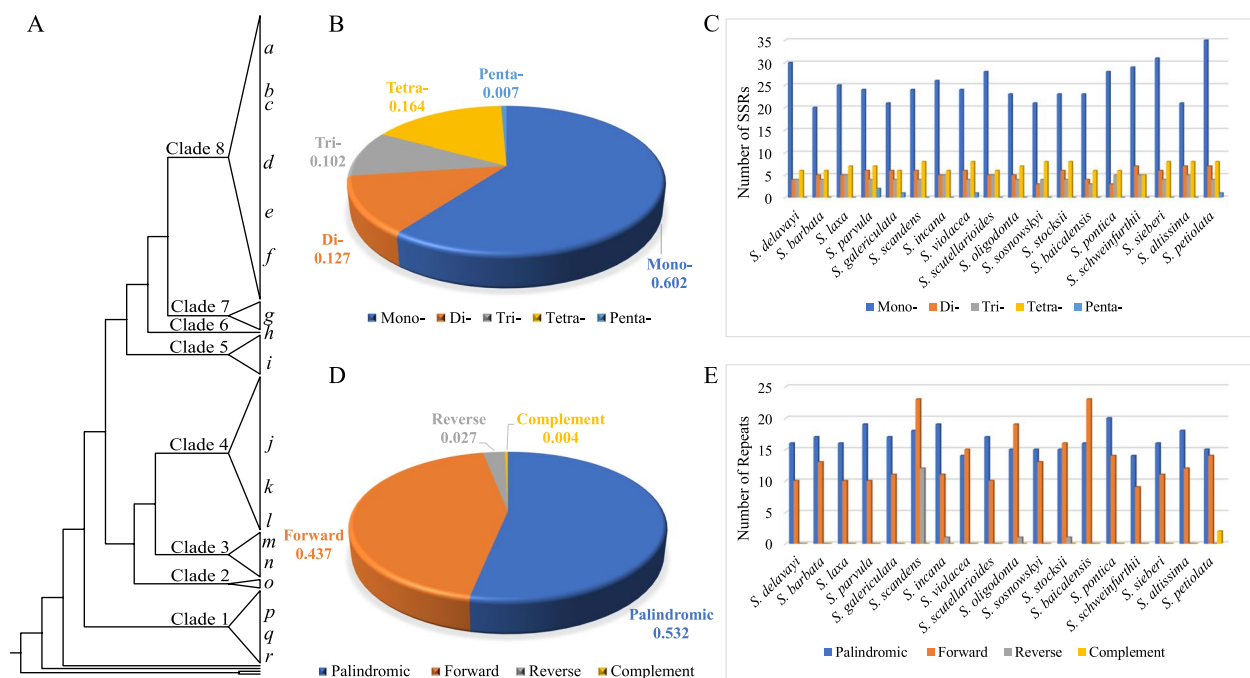
Fig. 5 The statistics of the simple sequence repeats (SSRs) and repetitive sequences of cpDNA sequences from 18 representative species randomly selected covering all clades of *Scutellaria*. **A** The location distribution of 18 representative species in the topology. **B** Proportion of five kinds of a total of SSRs. **C** The number of five types of SSRs in each species. mono-, mononucleotides; di-, dinucleotides; tri-, trinucleotides; tetra-, tetranucleotides; penta-, pentanucleotides. **D** Proportion of four kinds of repeats. **E** The number of four types of repetitive sequences in each species

Wang *et al. BMC Biology*   (2024) 22:185

Page 10 of 19

30, 39, and 60 bp present in all species (Additional file 1: Table S10).

Discussion

The phylogeny of *Scutellaria*

Our study represents the most comprehensive phylogenomic analysis of *Scutellaria* to date, employing the largest sample size and a substantial number of cpDNA sequences published for the first time. Compared to previous phylogenetic analyses, our research significantly enhanced tree resolutions and established a robust framework for understanding *Scutellaria*'s phylogeny utilizing six sequence matrices (PC, NC, PCN, PC-p, NC-p, PCN-p). Consistently, all phylogenetic trees affirmed that *Scutellaria*, as defined by Paton [2], is monophyletic, aligning with the findings of earlier studies [31–34, 37–41].

The monophyly of the Subgenus *Scutellaria* was not supported, as all phylogenetic trees indicated that Subg. *Apeltanthus* was nested within Subg. *Scutellaria*, with members of Subg. *Scutellaria* dispersed across all clades except for Clade 4 (Fig. 3, Additional file 2: Figs. S1–S11). This finding aligns with recent molecular research [31, 34, 36]. On the basis of morphology, *Scutellaria* was divided into Subg. *Scutellaria* and Subg. *Apeltanthus* by Paton [2], the former with one-sided or rarely spiral inflorescence and flowers subtended by leaves or leaf-like bracts, while 4-sided inflorescence with flowers subtended by cucullate bracts in the latter subgenus. However, the inflorescence of some species such as *S. linearis* Benth., *S. luteocaerulea* Bornm. & Sint., and *S. litwinowii* Bornm. & Sint. (Subg. *Apeltanthus* Sect. *Lupulinaria*) is one-sided, but bracts are still cucullate and nutlets have similar morphology to those of other species in Subg. *Apeltanthus*. Seyedipour et al. [33] considered that Subg. *Apeltanthus* was paraphyletic because *S. repens* (Subg. *Scutellaria*) was sister to *S. luteocaerulea* and its relatives (Subg. *Apeltanthus* Sect. *Lupulinaria*). The nutlet of *S. repens* is gray-black and completely covers the surface with hairs, also similar to the nutlets of Subg. *Apeltanthus* Sect. *Lupulinaria* [2]. However, the inflorescence is secund and the bracts are not cucullate. *S. repens* was placed into "*S. repens* species-group" (sect. *Scutellaria*) circumscribed by Paton based on axillary inflorescences subtended by leaf-like bracts. However, nutlet morphology showed that this species-group may not be a natural group [2, 63]. Nevertheless, the absence of *S. repens* in this and prior studies, underscores the need for further research with more comprehensive samples and loci information to investigate the placement of *S. repens*.

Subgenus *Apeltanthus*, divided into Sections *Apeltanthus* and *Lupulinaria* by Paton, exhibits varying calyx and nutlet characteristics across these sections

[2]. However, the monophyly of Sects. *Apeltanthus* and *Lupulinaria* was not supported by this analysis because *S. guttata* (Sect. *Apeltanthus*) was sister to *S. cordifrons* (Sect. *Lupulinaria*) with the strong support. *S. leptosiphon* and *S. immaculata* (Sect. *Apeltanthus*) were also nested into Sect. *Lupulinaria*, distant from the phylogenetic position of *S. guttata* and *S. stocksii*. The monophyly of Sect. *Lupulinaria* was also questioned in previous research based on ITS that *S. stocksii* of Sect. *Apeltanthus* was nested into Sect. *Lupulinaria* Subsect. *Lupulinaria* [32], which was consistent with the results of Seyedipour et al. [33] and Salimov et al. [34]. Sects. *Lupulinaria* and *Apeltanthus* should be merged as there is no evidence from this or previous molecular phylogenetic studies to maintain them as separate taxa [34]. The characteristics of boundary genes in this clade were evident. The distance of 100 bp between *rpl2* and JLA or JLB presents in 48 species (85.7%) of Subg. *Apeltanthus* and the 233 bp of *rps19* of 53 species (94.6%) located in LSC. Subg. *Apeltanthus* is distributed from the Caucasus eastward to the Pamir Mountains and southward to the Iranian plateau, though *S. alpina* L. occurs in western Europe. It is mainly concentrated in the western Irano-Turanian region which is the center of species diversity distribution [2, 34]. The presence of numerous short internal branches in Clade 4 further indicates rapid radiation within the Irano-Turanian region.

In the analyses of Salimov et al. [34], the Caucasian *S. pontica* (Section *Salviifoliae*), an East Asian lineage comprising Sect. *Scutellaria* (*S. baicalensis*, *S. amoena*, and *S. rehderiana*) and Sect. *Anaspis* (*S. kingiana*), along with Subg. *Apeltanthus* formed a tritomy [34]. Furthermore, Sect. *Salviifoliae* was not monophyletic, as *S. diffusa*'s phylogenetic position differed from *S. pontica*. However, in this study, the resolution among the aforementioned groups has been improved, and the species of Sect. *Salviifoliae* formed a monophyletic clade with strong support. Sect. *Salviifoliae* was strongly supported as a sister to the East Asian lineage (*S. baicalensis* and its relatives), together were sister to Subg. *Apeltanthus* with high support (Fig. 3, Additional file 2: Figs. S1–S11), which is consistent with Sect. *Salviifoliae* being morphologically intermediate between Sect. *Scutellaria* and Subg. *Apeltanthus*, as reported by Bentham [23] and Paton [2]. Sect. *Salviifoliae* was distinguished from the East Asian lineage by the apomorphy of hairs completely covering the nutlet surface, and from Sect. *Lupulinaria* Subsect. *Lupulinaria* due to the lack of the four-sided inflorescence, cucullate bracts, as well as strongly flattened pedicels. The analysis of boundary genes also indicated that the distance of 99 bp between *rpl2* and JLA or JLB, 3 bp between *trnH* and JLA, and the length of 233 bp of *rps19* in the LSC were both present in Sect. *Salviifoliae* and *S.*

ossethica, *S. raddeana* Juz., *S. orientalis* L., and *S. farsistanica* Rech.f. of Sect. *Lupulinaria*, inferencing that Sect. *Salviifoliae* was possibly viewed as the closest ancestor of Subg. *Apeltanthus*. The monophyletic clade containing Clades 2–4 was strongly supported, which is consistent with morphological characteristics in leaves with obtuse teeth or entire on each margin growing in arid upland or mountains, suggesting that Subg. *Apeltanthus* might include the larger clade containing Clades 2–4. Moreover, members of Sect. *Apeltanthus* were nested within Sect. *Lupulinaria* across various positions in Clade 4, suggesting that these two sections should be merged.

S. kingiana, previously classified under Subgenus *Anaspis* [28, 30] or Subgenus *Scutellaria* Section *Anaspis* [2], in this study, is shown to be sister to the East Asian lineage of the *S. baicalensis* alliance from Sect. *Scutellaria* with strong support (Fig. 3, Additional file 2: Figs. S1–S11). The phylogenetic relationships within this lineage are consistent with leaf morphology and habitat: species adapt to relatively arid zones with stem leaves dentate-serrate, crenate, to subentire or entire. *S. kingiana* and the members of Sect. *Anaspis* were clustered into different clades in this study, *S. kingiana* in Clade 3 and Sect. *Anaspis* in Clade 1. Thus, *S. kingiana* may not belong to Sect. *Anaspis* and Sect. *Anaspis* was non-monophyletic. All phylogenetic trees showed that Sect. *Anaspis* except for *S. kingiana* was sister to a lineage of *S. megalaspis–S. tournefortii* with strong support in Clade 1 (Fig. 3, Additional file 2: Figs. S1–S11). The lineage except for *S. utriculata* Labill. corresponding to "*S. albida* species-group" (Sect. *Scutellaria*) defined by Paton is similar to Sect. *Anaspis* (excluding *S. kingiana*) in nutlets type that gray-black nutlets with hairs partially covering the surface, the papillae with interior air space and lacking glands on the internal surface [2]. Species of "*S. albida* species-group" and *S. utriculata* probably should be transferred to Sect. *Anaspis*.

Subgenus *Scutellaria* Section *Scutellaria*, comprising approximately 240 species distributed across the Old and New Worlds, was divided into 34 informal species-groups by Paton [2]. Consistent with the morphological basis classifying Sect. *Scutellaria* as paraphyletic, this study also found that the monophyly of the section was not supported, aligning with previous analyses [31, 33–36]. As the largest section within Subg. *Scutellaria*, the sub-cosmopolitan Sect. *Scutellaria* was dispersed across various clades, yet exhibited clear morphological and geographic patterns within the section. The lineage of *S. megalaspis–S. tournefortii* was identified as sister to Sect. *Anaspis* excluding *S. kingiana*, based on similar nutlet anatomy. African species *S. schweinfurthii*, *S. polyadena*, and *S. violascens* were clustered into Clade 2 with strong support. *S. baicalensis* allies (Sect. *Scutellaria*),

primarily found in China, were determined to be sister to Sect. *Salviifoliae* in Clade 3. Neotropical species of Sect. *Scutellaria* in Clade 5 were identified as sister to Sect. *Perilomia* with strong support. Xerophytic species from America and Mexico, characterized by a woody rhizome, formed a distinct clade in Clade 7. A lineage of Sect. *Scutellaria* spanning from East Asia to North America was grouped into Clade 8, excluding Australian species *S. humilis* and West European species *S. minor* and *S. hastifolia*. The most comprehensive infrageneric classification of *Scutellaria* by Paton only included 13 species from China, despite China being a major center of diversity for the genus with 102 species, notably with most Chinese species being closely related to Sect. *Scutellaria* in Clade 8d. Therefore, a more detailed division of Sect. *Scutellaria*, based on denser sampling, is necessary.

The inferred close relationship between Clades 7 and 8 corresponds to their widespread distribution in East Asia and North America, suggesting a potential crossing between these regions via the Beringian land bridge, similar to the dispersion of the mint tribe Elsholtzieae (Lamiaceae) [64]. The widespread species *S. galericulata* in the Old and New World may be migrated by water and birds because it was grown near the waterside [3]. *S. minor* and *S. hastifolia* mainly distributed in West Europe were strongly supported as a sister to East Asia and North America species. Species migration in Europe and North America is likely to occur via Greenland in the Late Cretaceous and Early Tertiary and the Late Paleocene and Early Eocene, respectively [3, 65]. *S. violacea*, distributed across the Indian Peninsula, the Indo-China Peninsula, and East Himalaya, was positioned in Clade 6 as a sister to Clades 7 and 8, probably affected by the impact of the Indian Plate's collision with Asia [3, 65].

Within the neotropical clade (Clade 5), relationships are well-resolved. *Scutellaria scutellarioides*, the type species of Sect. *Perilomia*, and *S. flocculosa* from Peru are strongly supported as a sister to the neotropical species of Sect. *Scutellaria*. This result aligns with the research by Salimov et al. [34], which showed that *S. scutellarioides* and *S. volubilis* Kunth (Sect. *Perilomia*) were a sister to *S. costaricana* H.Wendl. and *S. incarnata* Vent. of Sect. *Scutellaria*. The inferred relationships correspond well with the floral morphology of a one-sided inflorescence, red or scarlet flowers, and a corolla tube that bends distally. Therefore, we recommend considering a wider circumscription of Sect. *Perilomia*. Intriguingly, *S. suffrutescens* primarily distributed in Northeast Mexico was sister to *S. wrightii* A. Gray, *S. drummondii* Benth., and *S. resinosa* Torr., which occur in America, with strong support in Clade 7. The formation of a connection between North and South America during the Eocene–Oligocene, named Gaarlandia, corresponding to the present-day

Caribbean islands [65], might have facilitated the dispersal of species in the New World.

Combining this study with other recent molecular phylogenetic studies, we suggest that updates to the taxonomy of the genus are urgently needed:

(1) Three main clades of *Scutellaria* were well-resolved, consistent with the inference by Salimov et al. [34] using chloroplast molecular fragments. Clade 1 in this study, equivalent to Clade B in Salimov's study, Clades 2–4 corresponding to Clade A, and Clades 5–8 equal to Clade C, are suggested to represent three subgenera. However, differences in phylogenetic relationships among them are evident. In Salimov's research [34], Clade A and Clade B formed a clade, but the relationship between the two clades was resolved with moderate support. In contrast, Clades 2–4 are strongly supported as sister to Clades 5–8 in this study. Thus, Clade 1 in this study is considered as the early-divergent branch of *Scutellaria*.

(2) Members of Section *Apeltanthus* were nested within Sect. *Lupulinaria* in different positions with high support in Clade 4, similar to the result by Salimov et al. and others [32–34]. Therefore, Sect. *Apeltanthus* and Sect. *Lupulinaria* should be merged. The monophyletic clade containing Clades 2–4 was consistent with morphological characteristics of leaves with obtuse teeth or entirely on each margin growing in arid upland or mountains, suggesting that the circumscription of Subg. *Apeltanthus* needs to be revised. Furthermore, African species *S. schweinfurthii relatives* (Sect. *Scutellaria*), Sect. *Salviifoliae*, *S. kingiana* (Sect. *Anaspis*) and *S. baicalensis* alliance (Sect. *Scutellaria*) should be transferred to Subg. *Apeltanthus* instead of Subg. *Scutellaria* whose type is *S. galericulata* within Clade 8. Subg. *Scutellaria* would then include Clades 5–8.

(3) Section *Anaspis*, excluding *S. kingiana*, was sister to the lineage of *S. tournefortii–S. megalaspis* in Clade 1, with similar nutlet anatomy, gray-black nutlets with hairs partially covering the surface, and papillae with interior air space but lacking glands on the internal surface. Thus, we suggest elevating Sect. *Anaspis*, except for *S. kingiana*, to subgeneric rank, as the name *Anaspis* was once considered at subgeneric rank in *Flora URSS* [28] and *Flora Reipublicae Populated Sinicae* [30]. The lineage of *S. tournefortii–S. megalaspis* should be separated from Sect. *Scutellaria* and transferred to the new subgenus. Further analysis is needed to confirm the phylogenetic position of *S. fedtschenkoi* Bornm., the type species of the section, due to the lack of DNA sequences.

(4) Based on the close phylogenetic relationships and similarities in morphological characters and geographical distribution between Sect. *Perilomia* and the neotropical population of Sect. *Scutellaria*, the circumscription of Sect. *Perilomia* should be expanded to include more neotropical species.
(5) The position of *S. kingiana* is distinct from the remaining members of Sect. *Anaspis*. However, *S. kingiana* is strongly supported as a sister to the *S. baicalensis* relatives with similar leaf morphology, habitat, and geographical distribution. Therefore, *S. kingiana* should be transferred to the section where the *S. baicalensis* alliance is located.

Unclear species identification

The identification of certain species complexes is problematic because the representatives of varieties of the same species do not form a clade, as observed in the *Scutellaria scordiifolia* Fisch. ex Schrank, *S. pekinensis* Maxim., and *S. indica* complexes. This issue highlights the need for extensive revision of a large number of species and variants. Establishing clear species boundaries is crucial for phylogenetic studies. In Flora of Pan-Himalaya [66], *S. nuristanica* was considered as a synonym of *S. petiolata* because there were no obvious differences in geographical disjunction and morphology of stem indumentum between the two species. Consistent with the present study, *S. nuristanica* was sister to *S. petiolata* with strong support. In addition, despite *S. altaicola* C.Y.Wu & H.W.Li being treated as a synonym for *S. altaica* Fisch. ex Sweet based on morphology [1], our analysis suggests it should remain an independent species due to their non-clustering in a clade (Fig. 3, Additional file 2: Figs. S1–S11). Species delimitation in *Scutellaria* often poses challenges, necessitating further revision work.

Potential molecular markers for *Scutellaria*

DNA barcoding, a tool for species identification and discovery of novel or cryptic species, supports food safety and identifies candidate exemplar taxa [67–71]. Since it was first proposed in 2003 [72], it has been applied to research on multiple taxa [70, 73, 74]. The technology has also been used for species identification in *Scutellaria*. *S. baicalensis*, the botanical origin of a well-known traditional Chinese medicine "*Huang Qin*", which is facing the rapid decline of natural resources due to high market demand, overexploitation, habitat destruction, and ecosystem degradation, resulting in high price in the market. This scarcity has led to a very common phenomenon of fraudulent adulteration with inherent differences that seriously affect the safety of medication [75, 76]. Some molecular markers, such as chloroplast fragments *rbcL*,

matK, *psbA–trnH*, *rpl16*, and *rpl16–rpl14*, and nuclear ribosomal ITS, have been used to discriminate *S. baicalensis* and its adulterants [59, 60, 75, 77]. However, these studies only focused on a few species. Chloroplast fragments *trnL* intron, *matK–trnK*, *trnL–trnF*, and ITS have been also applied to analyze intrageneric relationships of *Scutellaria* [32, 34]. However, resolution within related species was poor caused by insufficient phylogenetic signal of these markers with low diversity. In this study, 10 highly variable markers of cpDNA sequences with variation exceeding 0.025 were identified that can serve as potential DNA barcodes for species identification and phylogeny for *Scutellaria*, consistent with the result based on 11 species of *Scutellaria* by Zhao et al. [36]. Among them, candidate DNA barcodes including *trnH-psbA*, *trnK-rps16*, *petN-psbM*, *petA-psbJ*, and *ycf1* were identified based on three species of the genus by Jiang et al. [78]. While the reliability of these markers requires further validation, they provide a basis for future research in species identification, phylogeny, and population genetics.

SSRs, valuable for constructing high-density genome maps and establishing genetic and evolutionary relationships [79–82], are abundant in cpDNA sequences of *Scutellaria*, predominantly consisting of poly-adenine (A) and poly-thymine (T) repeats. SSRs tend to occur in intergenic spacer regions, although they are also found in coding regions where they can influence gene activity and protein expression. This study observed SSRs in regions that encode essential components of the chloroplast transcription machinery and detected mono- and tetranucleotide SSRs as the most abundant types within *Scutellaria*. The new SSRs identified offer potential molecular markers for medicinal species identification, genetic diversity estimation, and phylogenetic analysis.

Conclusions

Our study represents the most comprehensive phylogenomic analysis of *Scutellaria* to date, based on complete chloroplast genomes. It marks an improvement over previous studies, which were limited by either insufficient sampling or the range of genetic loci examined. Our results identified eight highly supported clades, with Clade 1 emerging as the earliest divergent branch and Clades 2–4 being sister to the monophyletic Clades 5–8. Through the integration of morphological and geographical distribution data, we proposed three subgenera and offered suggestions for updates to the genus's taxonomy. However, these proposals require validation through more extensive sampling and the inclusion of informative loci, including nuclear genes. The analysis of chloroplast genome characteristics provides a foundation

for studying the genetic structure and diversity of this resource-rich genus. Furthermore, comparative analyses of the chloroplast genomes of 18 representative species have identified potential molecular markers for taxonomy, genetic diversity, and species identification, particularly useful for distinguishing adulterants. Nevertheless, the rapid advancement of omics technologies and the need for an extensive multidisciplinary approach encompassing molecular systematics, phylogenomics, morphological anatomy, and taxonomy underscore the future directions for the phylogenetic study of *Scutellaria*.

Methods

Taxa sampling, DNA extraction, sequencing

Chloroplast genomes of 220 accessions, representing 196 species, subspecies, and varieties of *Scutellaria* spanning Eurasia, Americas, Africa, and Oceania, were newly sequenced. Additionally, we retrieved 14 cpDNA sequences corresponding to ten species and varieties from the NCBI GenBank database, including *S. altaica* (GenBank accession number: MN128387) [36, 83], *S. indica* var. *coccinea* S.Kim & S.Lee (MN047312) [84, 85], *S. orthocalyx* Hand.-Mazz. (MN128383) [36, 86], *S. mollifolia* C.Y. Wu & H.W. Li (MN128384) [36, 87], *S. calcarata* C.Y. Wu & H.W. Li (MN128385) [36, 88], *S. kingiana* (MN128389) [36, 89], *S. insignis* Nakai (NC028533 and KT750009) [90, 91], *S. rehderiana* (MT982397 and NC060314) [92, 93], *S. tuberifera* (MW376477 and NC059812) [94, 95], and *S. meehanioides* (MW381011 and NC057189) [96, 97]. Thus, a total of 234 cpDNA sequences were employed for phylogenetic and comparative genomic analyses, encompassing all sections except for Sect. *Salazaria* (Torrey) A.J.Paton. Additionally, four cpDNA sequences from three genera within the Scutellarioideae—*Holmskioldia sanguinea* (MN128388) [36, 98], *Tinnea aethiopica* (MN128380) [36, 99], and two sequences of *Wenchengia alternifolia* (MN128378 and MN128379) [36, 100, 101]—were selected as outgroups. Detailed information and voucher specimens for these accessions were cataloged in Additional file 1: Table S11.

High-quality total genomic DNA was extracted from about 1.5 g herbarium or silica gel-dried leaves using a modified cetyltrimethylammonium bromide (CTAB) protocol [102]. DNA quality was evaluated using horizontal electrophoresis with 1.5% agarose gel. Libraries construction was performed according to the manufacturer's manual (Illumina, San Diego, CA, USA) and sequenced on the Illumina NovaSeq 6000 platform with 2×150 bp paired-end reads by Tianjin Novogene Technology Co., ltd. Approximately 20 GB of raw data was generated for each sample. Fastp was used to trim the raw reads and adapter to obtain high-quality clean data by removing the connector sequences and the low-quality reads, and

Wang *et al. BMC Biology*   (2024) 22:185

Page 14 of 19

detailed information on sequencing data was placed in Additional file 1: Table S12 [103].

Chloroplast genome assembly and annotation

The complete chloroplast genomes were assembled with NOVOWrap v1.20 [104] and GetOrangelle v1.7.2a [105] and selected the best-quality genome for annotation. In this study, we chose the cpDNA of *Scutellaria barbata* (NC059814) as a reference when we used NOVOWrap v1.20. Afterward, we selected the best genome according to the schematic diagram of the collinearity analysis of the assembly results. In addition, we also used GetOrangelle v1.7.2a to assemble with default parameters: -F embplant_pt -R 15 -k 21,45,65,85,105. Then, PGA was used to annotate with *Amborella trichopoda* Baill. (NC005086) and *S. baicalensis* (NC027262) as references [106]. To improve the accuracy of annotation results, we have manually corrected the annotation results in Geneious v2021.1.1 [107]. A circular map of the cpDNA was visualized using Organellar Genome DRAW v1.3.1 (OGDRAW) [108].

Phylogenetic analyses

All of the protein-coding genes (CDS), tRNAs, rRNAs, and non-coding regions were automatically extracted based on annotation results by a custom Python script: get_annotated_regions_from_gb.py (https://github.com/Kinggerm/PersonalUtilities/). Each locus was individually aligned using MAFFT v7.471 with default settings except –localpair and –maxiterate 1000 [109]. Removal of divergent and ambiguous sequence alignments may improve topology [110]. All sequences were adjusted manually after being checked to improve positional homology following the rules of Löhne et al. [111] (Additional file 1: Table S13). We formed three sequence matrices after adjusting the alignment: (1) PC including 81 protein-coding genes, 30 tRNA genes, and four rRNA genes, (2) NC contained 134 non-coding regions, (3) PCN comprised 115 gene sequences and 134 non-coding regions. The phylogenetic trees were analyzed using an unpartitioned strategy on the three sequence matrices, respectively, because chloroplast genomes are often considered as a linked single locus [112, 113]. Empirical studies showed that different regions may not be regarded as a single locus due to experiencing divergent evolutionary forces [114, 115]. Thus, a total of six matrices with or without partitioning (PC, NC, PCN, PC-p, NC-p, PCN-p) were implemented for phylogenetic analysis using ML and BI methods, respectively.

For unpartitioned strategies, three unpartitioned super-matrices were used for phylogeny using ML and BI. The best-fit substitution model was identified based on the Bayesian Information Criterion (BIC) criterion using

ModelFinder embedded in IQ-TREE, with GTR + F + R3 model for PC, TVM + F + R10 model for NC, and GTR + F + R6 model for PCN, together with default settings except for 1000 Ultrafast bootstrap replicates and 1000 bootstrap replicates of SH-like approximate likelihood ratio test (SH-aLRT) to program ML analysis using IQ-TREE v 1.6.12 [116]. BI analysis was conducted with Mrbayes v 3.2.7a under the GTR + I + G model [117]. Four Markov Chain Monte Carlo chains with one cold and three hot chains were implemented for 1,000,000 generations and sampled every 500 generations, and the first 25% of all trees were discarded as 'burn-in' (the default settings of Mrbayes).

In terms of partitioned strategy, the appropriate partition schemes were obtained in the process of gene concatenation using a plug-in in PhyloSuite v1.2.2 [118]. Partitionfinder v2.1.1 was used to estimate the evolutionary model of each partition [119]. For ML analysis, the evolutionary model was obtained based on parameters branchlengths = linked, models = all, model_selection = aicc, and search = rcluster. The phylogenetic trees were inferred using IQ-TREE v 1.6.12 with 1000 Ultrafast bootstrap replicates and SH-aLRT, and the parament -spp to provide the most appropriate partition scheme and evolutionary model. In BI analysis, the evolutionary model of each partition was determined based on options branchlengths = unlinked, models = mrbayes, model_selection = aicc, and search = greedy, and the phylogeny was analyzed with Mrbayes v 3.2.7a. Four Markov Chain Monte Carlo chains (one cold and three hot chains) were run for 1,000,000 generations and sampled every 500 generations, and the first 25% of all trees were discarded as 'burn-in' (the default settings of Mrbayes). All phylogenetic relationships were visualized using FigTree v1.4.4 [120]. The Bayesian trees with posterior probability (PP) and ML trees with bootstrap support (BS) and SH-aLRT support (SH) were shown in all trees.

Comparative analyses of cpDNA sequences in the genus

CPJSdraw v1.0.0 and IRscope were used for boundary visualization analysis of IR/SC for the whole cpDNA sequences [121, 122].

In addition, we randomly selected 18 species covering eight clades for comparative genomic analyses. We used Mauve v2.4.0 for estimating the collinearity to analyze whether there was a large segment sequence rearrangement among cpDNA sequences [123]. For divergence analysis, the online comparison tool mVISTA was applied to compare and visualize the similarity with the shuffle-LAGAN mode, and *S. scutellarioides* newly sequenced in this study was used as a reference [124]. Additionally, to assess the nucleotide variabilities, cpDNA sequences of 18 representative species were aligned using MAFFT

v7.471 with default settings except for −maxiterate 1000 and −localpair [109]. Then, nucleotide diversity (pi) values were calculated using DnaSP v6 with a window length of 600 bp and a step size of 200 bp [125].

Simple sequence repeats (SSRs) or microsatellites were identified by the online tool MISA v2.1 with a settled minimum threshold of mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats set to 10, 5, 4, 3, 3, and 3, respectively, and set an option about the maximum length of sequence between two SSRs to register as compound SSR to 0 [126]. Then, the results were annotated to obtain the location of SSRs in the cpDNA sequences by the platform JSHYCloud (http://cloud.genepioneer.com:9929). The repeat sequences including forward, reverse, palindromic, and complementary formats were predicted using the online software REPuter with a minimal size of 30 bp and Hamming distance of 3 [127]. The results were also annotated to determine the location in the cpDNA sequences by the platform JSHYCloud (http://cloud.genepioneer.com:9929).

Abbreviations

POWO	Plants of the World Online
cpDNA	Complete chloroplast genome
IRs	Two identical copies of inverted repeats
SSC	A small single-copy region
LSC	A large single-copy region
JSB	Junction of IRb/SSC
JSA	Junction of SSC/IRa
JLA	Junction of IRa/LSC
JLB	Junction of LSC/IRb
PC	An unpartitioned matrix including 81 protein-coding genes, 30 tRNA genes, and four rRNA genes
PC-p	A partitioned matrix including 81 protein-coding genes, 30 tRNA genes, and four rRNA genes
NC	An unpartitioned matrix of 134 non-coding regions
NC-p	A partitioned matrix of 134 non-coding regions
PCN	An unpartitioned matrix of 115 genes and 134 non-coding regions
PCN-p	A partitioned matrix of 115 genes and 134 non-coding regions

Supplementary Information

Additional file 1: Table S1. Characteristics of 234 complete chloroplast genome of *Scutellaria*. Table S2. List of genes of cpDNA of *Scutellaria*. Table S3. Statistics of Pi values based on DnaSP. Table S4. The type and position of SSRs from cpDNA sequences of 18 representative species of *Scutellaria*. Table S5. Statistics about the number and proportion of different types of SSRs from cpDNA sequences of 18 representative species of *Scutellaria*. Table S6. The number and proportion of different repeat units of SSRs from cpDNA sequences of 18 representative species of *Scutellaria*. Table S7. Statistics on the presence or absence of repetitive units of SSRs in each representative species. Table S8. The type, length, and position of repetitive sequences from cpDNA sequences of 18 representative species of *Scutellaria*. Table S9. Statistics of the number of repetitive sequences for different types and regions from cpDNA sequences of 18 representative species of *Scutellaria*. Table S10. Statistics about the number of repeat sequences of different lengths in cpDNA sequences of 18 representative species of *Scutellaria*. Table S11. The origin of materials and accession numbers. Table S12. Sequencing data size and quality of the accessions. Table S13. The positions were masked during alignment adjustment.

Additional file 2: Fig. S1. The phylogeny of *Scutellaria* based on Maximum likelihood was inferred based on a tandem matrix of 115 genes (PC). The support values of bootstrap support (BS) and SH-like approximate likelihood ratio (SH-aLRT) were shown in turn near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 100 % were replaced with asterisks. Fig. S2. The phylogeny of *Scutellaria* based on Bayesian inference was inferred based on a tandem matrix of 115 genes (PC). The support values of posterior probabilities (PP) were shown near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 1 were replaced with asterisks. Fig. S3. The phylogeny of *Scutellaria* based on Maximum likelihood was inferred based on a partition matrix of 115 genes (PC-p). The support values of bootstrap support (BS) and SH-like approximate likelihood ratio (SH-aLRT) were shown in turn near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 100 % were replaced with asterisks. Fig. S4. The phylogeny of *Scutellaria* based on Bayesian inference was inferred based on a partition matrix of 115 genes (PC-p). The support values of posterior probabilities (PP) were shown near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 1 were replaced with asterisks. Fig. S5. The phylogeny of *Scutellaria* based on Maximum likelihood was inferred based on a tandem matrix of 134 non-coding regions (NC). The support values of bootstrap support (BS) and SH-like approximate likelihood ratio (SH-aLRT) were shown in turn near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 100 % were replaced with asterisks. Fig. S6. The phylogeny of *Scutellaria* based on Bayesian inference was inferred based on a tandem matrix of 134 non-coding regions (NC). The support values of posterior probabilities (PP) were shown near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 1 were replaced with asterisks. Fig. S7. The phylogeny of *Scutellaria* based on Maximum likelihood was inferred based on a partition matrix of 134 non-coding regions (NC-p). The support values of bootstrap support (BS) and SH-like approximate likelihood ratio (SH-aLRT) were shown in turn near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 100 % were replaced with asterisks. Fig. S8. The phylogeny of *Scutellaria* based on Bayesian inference was inferred based on a partition matrix of 134 non-coding regions (NC-p). The support values of posterior probabilities (PP) were shown near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 1 were replaced with asterisks. Fig. S9. The phylogeny of *Scutellaria* based on Maximum likelihood was inferred based on a tandem matrix including 115 genes and 134 non-coding regions (PCN). The support values of bootstrap support (BS) and SH-like approximate likelihood ratio (SH-aLRT) were shown in turn near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have

Wang *et al. BMC Biology*   (2024) 22:185

Page 16 of 19

not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 100 % were replaced with asterisks. Fig. S10. The phylogeny of *Scutellaria* based on Bayesian inference was inferred based on a tandem matrix including 115 genes and 134 non-coding regions (PCN). The support values of posterior probabilities (PP) were shown near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 1 were replaced with asterisks. Fig. S11. The phylogeny of *Scutellaria* based on Bayesian inference was inferred based on a partition matrix of 115 genes and 134 non-coding regions (PCN-p). The support values of posterior probabilities (PP) were shown near the notes. All species belonging to the classification system by Paton were marked with different colored rectangles for different sections. Some unlabeled species have not been classified yet. Clades 1–8 of *Scutellaria* were marked. Clade 8 was subdivided into Clades 8a–8d. Values equal to 1 were replaced with asterisks.

Additional file 3: Fig. S12. Collinearity analysis of cpDNA sequences from 18 representative species randomly selected covering all clades of *Scutellaria*. The white rectangle box represents CDS, the red box represents rRNA, and tRNAs are indicated by green boxes. The two pink elongated rectangular boxes present in each genome are IR regions. Fig. S13. Visualization for comparison of cpDNA sequences from 18 representative species randomly selected covering all clades of the genus with *S. scutellarioides* as the reference using mVISTA. The X-axis depicts the sequence length, and the Y-axis depicts the percentage identity of the reference. Genome regions are color-coded as exon, UTR, mRNA, and conserved noncoding sequences (CNS). Gray arrows at the top line show the direction and position of each gene. The vertical scale represents the percentage of sequence identity, ranging from 50% to 100%.

Acknowledgements

Authors' contributions
This study was conceived and designed by Y.H.W. and Q.W. DNA extraction was performed by C.X. Analyses were performed by Y.H.W. and manual correction of chloroplast genome annotation was performed by Y.H.W., Y.W., and Y.Y.C. The manuscript was written by Y.H.W. and Q.W. with contributions from Y.W., Y.Y.C., X.G., J.S., C.N.H., and Y.Y. The authors read and approved the final manuscript.

Funding

Availability of data and materials
All data generated in this study will be available on the NCBI. The GenBank accession numbers were listed in Additional file 1: Table S11. The plant materials collected for this study came from herbarium that voucher information in Additional file 1: Table S11. The information about the download data of a few species is included in the article. Sequence alignments underlying analyses and phylogenetic trees could be available from figshare (https://doi.org/10.6084/m9.figshare.26213495.v1) [128].

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]State Key Laboratory of Plant Diversity and Specialty Crops, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China. [2]China National Botanical Garden, Beijing 100093, China. [3]College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. [4]State Key Laboratory of Agricultural Genomics, Key Laboratory of Genomics, Ministry of Agriculture, BGI Research, Wuhan 430047, China. [5]School of Medical Laboratory, Shandong Second Medical University, Weifang 261053, China. [6]Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing 100193, China. [7]Key Laboratory of Bio-Resources and Eco-Environment of Ministry of Education, College of Sciences, Sichuan University, Chengdu 610065, China.

References
1. Govaerts R. World Checklist of Selected Plant Families Database in ACCESS: 1–216203. The Board of Trustees of the Royal Botanic Gardens, Kew. https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:30003498-2 (2003).
2. Paton A. A global taxonomic investigation of *Scutellaria* (Labiatae). Kew Bull. 1990:399–450.
3. Paton A. The phytogeography of *Scutellaria* L. Notes Roy Bot Gard Edinburgh. 1990;46(3):345–59.
4. Mercado JM, Paton AJ. *Scutellaria* L. (Lamiaceae) in Bolivia with observations on sect. *Perilomia*. Kew Bull. 2006;61:549–58.
5. Kudo Michiko, Kobayashi-Nakamura Kumiko, Tsuji-Naito K. Bifunctional effects of O-methylated flavones from *Scutellaria baicalensis* Georgi on melanocytes: Inhibition of melanin production and intracellular melanosome transport. PLoS One. 2017;12(2):e0171513.
6. Miyaichi Y, Imoto Y, Tomimori T, Lin C-C. Studies on the constituents of *Scutellaria* species. IX. On the flavonoid constituents of the root of *Scutellaria indica* L. Chem Pharm Bull. 1987;35(9):3720–5.
7. Miyaichi Y, Kizu H, Tomimori T, Lin CC. Studies on the constituents of *Scutellaria* species. XI. On the flavonoid constituents of the aerial parts of *Scutellaria indica* L. Chem Pharm Bull. 1989;37(3):794–7.
8. Sato Y, Suzaki S, Nishikawa T, Kihara M, Shibata H, Higuti T. Phytochemical flavones isolated from *Scutellaria barbata* and antibacterial activity against methicillin-resistant Staphylococcus aureus. J Ethnopharmacol. 2000;72(3):483–8.
9. Shen J, Li P, Liu S, Liu Q, Li Y, Sun Y, et al. Traditional uses, ten-years research progress on phytochemistry and pharmacology, and clinical studies of the genus *Scutellaria*. J Ethnopharmacol. 2021;2021(265):113198.
10. Wang Z, Qi F, Cui Y, Zhao L, Sun X, Tang W, et al. An update on Chinese herbal medicines as adjuvant treatment of anticancer therapeutics. Biosci Trends. 2018;12(3):220–39.
11. Dong Q, Chu F, Wu C, Huo Q, Gan H, Li X, et al. *Scutellaria baicalensis* Georgi extract protects against alcoholinduced acute liver injury in mice and affects the mechanism of ER stress. Mol Biol Evol. 2016;13:3052–62.
12. Yin X, Zhou J, Jie C, Xing D, Zhang Y. Anticancer activity and mechanism of *Scutellaria barbata* extract on human lung cancer cell line A549. Life Sci. 2004;75(18):2233–44.
13. Ye F, Xui L, Yi J, Zhang W, Zhang DY. Anticancer activity of *Scutellaria baicalensis* and its potential mechanism. J Altern Complement Med. 2002;8(5):567–72.
14. Lohani M, Ahuja M, Buabeid MA, Dean S, Dennis S, Suppiramaniam V, et al. Anti-oxidative and DNA protecting effects of flavonoids-rich *Scutellaria lateriflora*. Nat Prod Commun. 2013;8(10):1415–8.

Wang *et al. BMC Biology*   (2024) 22:185

Page 17 of 19

15. Chen Q-Y, Wang C-Q, Yang Z-W, Tang Q, Tan H-R, Wang X, et al. Differences in anti-inflammatory effects between two specifications of *Scutellariae* Radix in LPS-induced macrophages in vitro. Chin J Nat Med. 2017;15(7):0515–24.

16. Shang X, He X, He X, Li M, Zhang R, Fan P, et al. The genus *Scutellaria* an ethnopharmacological and phytochemical review. J Ethnopharmacol. 2010;128(2):279–313.

17. Kizu H, Sugita N, Tomimori T. Studies on Nepalese crude drugs. XXIV. Diterpenoid constituents of the leaves of *Scutellaria repens* Buch.-Ham. ex D. Don. Chem Pharm Bull. 1998;46(6):988–1000.

18. Bruno M, Piozzi F, Maggio AM, Simmonds MSJ. Antifeedant activity of neoclerodane diterpenoids from two Sicilian species of *Scutellaria*. Biochem Syst Ecol. 2002;30(8):793–9.

19. Bozova PI, Georgieva YP. Antifeedant Activity of Neo-clerodane Diterpenoids from *Scutellaria altissima* against Colorado Potato Beetle Larvae. Nat Prod Commun. 2017;12(3):327–8.

20. Rodríguez B,Torre MCdl, Rodríguez B, Bruno M, Piozzi F, Savona G, et al. Neo-clerodane insect antifeedants from *Scutellaria galericulata*. Phytochemistry. 1993;33(2):309–15.

21. Hamilton A. Esquisse d'une monographie du genre *Scutellaria* ou toque. Lyon: Louis Perrin; 1832.

22. Bentham G. *Scutellaria*. In: Candolle AP, editor. Paris: Prodromus Systematis Naturalis; 1848. p. 412–32.

23. Bentham G. *Scutellaria*, *Salazaria* and *Perilomia*. In: Bentham G, Hooker JD, editors. Genera Plantarum. London: Lovell Reeve & Co; 1876. p. 1201–3.

24. Bentham G. *Scutellaria* and *Perilomia*. Labiatarum Genera and Species. London: Ridgeway; 1834. p. 416–46.

25. Briquet J. *Scutellaria*, *Salazaria* and *Perilomia*. In: Engler A, Prantl KAE, editors. Die Natiirlichen Pflanzenfamilien. 1896. p. 224–33.

26. Edmondson JR. *Scutellaria*. In: Davis PH, editor. Flora of Turkey. Edinburgh: University Press; 1982. p. 78–100.

27. Epling C. American species of *Scutellaria*. Univ Cal Publ Bot. 1942;20:1–146.

28. Juzepczuk SV. *Scutellaria*. In: Shishkin BK, Juzepczuk SV, editors. Flora URSS. Moscow: Academiae Scientarum URSS; 1954. p. 71–225.

29. Rechinger KH. *Scutellaria*. In: Rechinger KH, editor. Flora Iranica. 1982. p. 48–84.

30. Wu CY, Li HW. *Scutellaria*. Beijing: Flora Reipublicae Popularis Sinicae; 1977. p. 124–248.

31. Zhao F, Liu E-D, Peng H, Xiang C-L. A new species of *Scutellaria* (Scutellarioideae, Lamiaceae) from Sichuan Province in southwest China. PeerJ. 2017;5:e3624.

32. Safikhani K, Jamzad Z, Saeidi H. Phylogenetic relationships in Iranian *Scutellaria* (Lamiaceae) based on nuclear ribosomal ITS and chloroplast *trn*L-F DNA data. Plant Syst Evol. 2018;304(9):1077–89.

33. Seyedipour S, Salmaki Y, Xiang C-L. Molecular phylogeny of *Scutellaria* (Lamiaceae; Scutellarioideae) in Iranian highlands inferred from nrITS and *trn*L-F sequences. Proc Biol Sci. 2020;7(1):169–81.

34. Salimov RA, Parolly G, Borsch T. Overall phylogenetic relationships of *Scutellaria* (Lamiaceae) shed light on the origin of the predominantly Caucasian and Irano-Turanian *S. orientalis* group. Willdenowia. 2021;51(3):395–427.

35. Shan Y, Pei X, Yong S, Li J, Yu J. Analysis of the complete chloroplast genomes of *Scutellaria tsinyunensis* and *Scutellaria tuberifera* (Lamiaceae). Mitochondrial DNA B Resour. 2021;6(9):2672–80.

36. Zhao F, Li B, Drew BT, Chen Y-P, Wang Q, Yu W-B, et al. Leveraging plastomes for comparative analysis and phylogenomic inference within Scutellarioideae (Lamiaceae). PLoS ONE. 2020;15(5):e0232602.

37. Li B, Cantino PD, Olmstead RG, Bramley GLC, Xiang C-L, Ma Z-H, et al. A large-scale chloroplast phylogeny of the Lamiaceae sheds new light on its subfamilial classification. Sci Rep. 2016;6:34343.

38. Li B, Xu W, Tu T, Wang Z, Olmstead RG, Peng H, et al. Phylogenetic position of *Wenchengia* (Lamiaceae): a taxonomically enigmatic and critically endangered genus. Taxon. 2012;61(2):392–401.

39. Bendiksby M, Thorbek L, Scheen A-C, Lindqvist C, Ryding O. An updated phylogeny and classification of Lamiaceae subfamily Lamioideae. Taxon. 2011;60(2):471–84.

40. Chen Y-P, Drew BT, Li B, Soltis DE, Soltis PS, Xiang C-L. Resolving the phylogenetic position of *Ombrocharis* (Lamiaceae), with reference to the molecular phylogeny of tribe Elsholtzieae. Taxon. 2016;65(1):123–36.

41. Chen Y-P, Li B, Olmstead RG, Cantino PD, Liu E-D, Xiang C-L. Phylogenetic placement of the enigmatic genus *Holocheila* (Lamiaceae) inferred from plastid DNA sequences. Taxon. 2014;63(2):355–66.

42. Zhao F, Chen Y-P, Salmaki Y, Drew BT, Wilson TC, Scheen A-C, et al. An updated tribal classification of Lamiaceae based on plastome phylogenomics. BMC Biol. 2021;19:2.

43. Vollesen K. A taxonomic revision of the genera *Tinnea* and *Renschia* (Lamiaceae, Ajugoideae). Bot Tidsskr. 1975;70(1):1–63.

44. Atkins S. *Holmskioldia sanguinea* Labiatae, formerly Verbenaceae. Bot Mag. 1996;13:79–81.

45. Paton A, Phillipson PB, Suddee S. Records of *Wenchengia* (Lamiaceae) from Vietnam. Biodivers Data J. 2016;4:e9596.

46. Bullock A. The Genera *Tinnea* and *Renschia*. Kew Bull. 1931;1931(9):455–7.

47. Li Bo, Hoang Thanh Son, Nuraliev Maxim S, Tai Vu Anh, Kuznetsov Andrey N, Kuznetsova Svetlana P, et al. *Heliacria* gen. nov. (Scutellarioideae, Lamiaceae) from Coastal Southern Vietnam. J Trop Subtrop Botany. 2023;31(5):727–35.

48. Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, et al. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J. 1986;5(9):2043–9.

49. Wang R-J, Cheng C-L, Chang C-C, Wu C-L, Su T-M, Chaw S-M. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. BMC Evol Biol. 2008;8:36.

50. Ruhlman TA, Jansen RK. The plastid genomes of flowering plants. Methods Mol Biol. 2014;1132:3–38.

51. Wicke S, Schneeweiss GM, dePamphilis CW, Muller KF, Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol. 2011;76:273–97.

52. Daniell H, Lin C-S, Yu M, Chang W-J. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biol. 2016;17(1):134.

53. Wu Y, Liu F, Yang D-G, Li W, Zhou X-J, Pei X-Y, et al. Comparative Chloroplast Genomics of *Gossypium* Species: Insights Into Repeat Sequence Variations and Phylogeny. Front Plant Sci. 2018;9:376.

54. Ahmed I. Chloroplast Genome Sequencing: Some Reflections. Next Gener Seq Appl. 2015;2(2):119.

55. Wang N, Chen S, Xie L, Wang L, Feng Y, Lv T, et al. The complete chloroplast genomes of three Hamamelidaceae species: Comparative and phylogenetic analyses. Ecol Evol. 2022;12:e8637.

56. Song W, Chen Z, He L, Feng Q, Zhang H, Du G, et al. Comparative Chloroplast Genome Analysis of Wax Gourd (*Benincasa hispida*) with Three Benincaseae Species, Revealing Evolutionary Dynamic Patterns and Phylogenetic Implications. Genes. 2022;13:461.

57. Lian C, Yang H, Lan J, Zhang X, Zhang F, Yang J, et al. Comparative analysis of chloroplast genomes reveals phylogenetic relationships and intraspecific variation in the medicinal plant Isodon rubescens. PLoS ONE. 2022;17(4):e0266546.

58. Hosokawa K, Minami M, Kawahara K, Nakamura I, Shibata T. Discrimination among three species of medicinal *Scutellaria* plants using RAPD markers. Planta Med. 2000;66:270–2.

59. Hosokawa K, Minami M, Nakamura I, Hishida A, Shibata T. The sequences of the plastid gene rpl16 and the rpl16–rpl14 spacer region allow discrimination among six species of *Scutellaria*. J Ethnopharmacol. 2005;99:105–8.

60. Chen S-y, Zhao R, Xu L, Feng Q-j, Wang B, Kang T-g. DNA Barcoding of Mongolian Medicinal Plant *Scutellaria scordifolia*. Zhong Yao Cai. 2016;39(7):1483–7.

61. Yao G, Zhang Y-Q, Barrett C, Xue B, Bellot S, Baker WJ, et al. A plastid phylogenomic framework for the palm family (Arecaceae). BMC Biol. 2023;21(1):50.

62. Pick KS, Philippe H, Schreiber F, Erpenbeck D, Jackson DJ, Wrede P, et al. Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. Mol Biol Evol. 2010;27(9):1983–7.

63. Paton A, Suddee S, Bongcheewin B. Two new species of *Scutellaria* (Lamiaceae) from Thailand and Burma. Kew Bull. 2016;71(1):3.

64. Li P, Qi Z-C, Liu L-X, Ohi-Toma T, Lee J, Hsieh T-H, et al. Molecular phylogenetics and biogeography of the mint tribe Elsholtzieae (Nepetoideae, Lamiaceae), with an emphasis on its diversification in East Asia. Sci Rep. 2017;7(1):2057.

Wang *et al. BMC Biology*   (2024) 22:185

Page 18 of 19

65. Morley RJ. Interplate dispersal paths for megathermal angiosperms. Perspect Plant Ecol Evol Syst. 2003;6(1–2):5–20.

66. Wang Q. *Scutellaria*. Flora of Pan-Himalaya. Beijing; 2019. p. 180–230.

67. Vassou SL, Kusuma G, Parani M. DNA barcoding for species identification from dried and powdered plant parts: A case study with authentication of the raw drug market samples of *Sida cordifolia*. Gene. 2015;559(1):86–93.

68. Mishra P, Kumar A, Nagireddy A, Mani DN, Shukla AK, Tiwari R, et al. DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. Plant Biotechnol J. 2015;14:8–21.

69. Singh A, Rai VP, Singh M, Singh AK, Sinha B. Molecular diversity analysis of *Asparagus racemosus* and its adulterants using random amplified polymorphic DNA (RAPD). J Med Plant Res. 2013;7(16):1050–6.

70. Gao T-X, Ji D-P, Xiao Y-S, Xue T-Q, Yanagimoto T, Setoguma T. Description and DNA Barcoding of a New Sillago Species, *Sillago sinica* (Perciformes: Sillaginidae), from Coastal Waters of China. Zool Stud. 2011;50(2):254–63.

71. Hosein FN, Austin N, Maharaj S, Johnson W, Rostant L, Ramdass AC, et al. Utility of DNA barcoding to identify rare endemic vascular plant species in Trinidad. Ecol Evol. 2017;7:7311–33.

72. Hebert PDN, Cywinska A, Ball SL, Dewaard JR. Biological identifications through DNA barcodes. Proc R Soc Lond B. 2003;270:313–21.

73. Saunders GW. Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. Phil Trans R Soc B. 2005;360:1879–88.

74. Sass C, Little DP, Stevenson DW, Specht CD. DNA Barcoding in the *Cycadales*: Testing the Potential of Proposed Barcoding Markers for Species Identification of *Cycads*. PLoS ONE. 2007;2(11):e1154.

75. Guo X, Wang X, Su W, Zhang G, Zhou R. DNA barcodes for discriminating the medicinal plant *Scutellaria baicalensis* (Lamiaceae) and its adulterants. Biol Pharm Bull. 2011;34(8):1198–203.

76. Irvin L, Vaidya BN, Arun A, Joshee N. DNA Barcoding Studies in the Medicinal Genus *Scutellaria*. 2016.

77. Xia Z, Feng C-y, Gao Z-m, Li H-m, Zhang H-r. Authentication of DNA barcoding of *Scutellaria baicalensis* and its related species. Zhong Cao Yao. 2014;45(1):107–12.

78. Jiang D, Zhao Z, Zhang T, Zhong W, Liu C, Yuan Q, et al. The Chloroplast Genome Sequence of *Scutellaria baicalensis* Provides Insight into Intraspecific and Interspecific Chloroplast Genome Diversity in *Scutellaria*. Genes. 2017;8(9):227.

79. Lambert P, Hagen LS, Arus P, Audergon JM. Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond Texas x peach Earlygold reference map for *Prunus*. Theor Appl Genet. 2004;108(6):1120–30.

80. Cai G, Yang Q, Yang Q, Zhao Z, Zhou Y. Identification of candidate genes of QTLs for seed weight in *Brassica napus* through comparative mapping among *Arabidopsis* and *Brassica* species. BMC Genet. 2012;13:105.

81. Matsuoka Y, Vigouroux Y, Goodman MM, Jesus SG, Buckler E, Doebley J. A single domestication for maize shown by multilocus microsatellite genotyping. PNAS. 2002;99(9):6080–4.

82. Neeraja CN, Maghirang-Rodriguez R, Pamplona A, Heuer S, Collard BCY, Septiningsih EM, et al. A marker-assisted backcross approach for developing submergence-tolerant rice cultivars. Theor Appl Genet. 2007;115(6):767–76.

83. Zhao F, Li B, Drew BT, Chen YP, Wang Q, Yu WB, et al. *Scutellaria altaica* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MN128387 (2020).

84. Lee Y, Kim S. *Scutellaria indica* var. *coccinea* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MN047312.1/ (2020).

85. Lee Y, Kim S. The complete chloroplast genome of *Scutellaria indica* var *coccinea* (Lamiaceae), an endemic taxon in Korea. Mitochondrial DNA B Resour. 2019;4(2):2539–40.

86. Zhao F, Li B, Drew BT, Chen YP, Wang Q, Yu WB, et al. *Scutellaria orthocalyx* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MN128383.1/ (2020).

87. Zhao F, Li B, Drew BT, Chen YP, Wang Q, Yu WB, et al. *Scutellaria mollifolia* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MN128384 (2020).

88. Zhao F, Li B, Drew BT, Chen YP, Wang Q, Yu WB, et al. *Scutellaria calcarata* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MN128385 (2020).

89. Zhao F, Li B, Drew BT, Chen YP, Wang Q, Yu WB, et al. *Scutellaria kingiana* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MN128389 (2020).

90. Son O, Park S. *Scutellaria insignis* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/NC_028533 (2023).

91. Son O, Park S. *Scutellaria insignis* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/KT750009 (2015).

92. Cai Z. *Scutellaria rehderiana* chloroplast, complete sequence. GenBank https://www.ncbi.nlm.nih.gov/nuccore/NC_060314 (2023).

93. Cai Z. *Scutellaria rehderiana* voucher SREHD20200816 chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MT982397 (2021).

94. Li J, Shan y, Pei X, Yong S, Yu J. *Scutellaria tuberifera* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MW376477 (2021).

95. Li J, Shan y, Pei X, Yong S, Yu J. *Scutellaria tuberifera* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/NC_059812 (2023).

96. Zhang C. *Scutellaria meehanioides* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/NC_057189 (2023).

97. Zhang C. *Scutellaria meehanioides* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MW381011 (2021).

98. Zhao F, Li B, Drew BT, Chen YP, Wang Q, Yu WB, et al. *Holmskioldia sanguinea* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MN128388 (2020).

99. Zhao F, Li B, Drew BT, Chen YP, Wang Q, Yu WB, et al. *Tinnea aethiopica* chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MN128380 (2020).

100. Zhao F, Li B, Drew BT, Chen YP, Wang Q, Yu WB, et al. *Wenchengia alternifolia* isolate VN chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MN128378 (2020).

101. Zhao F, Li B, Drew BT, Chen YP, Wang Q, Yu WB, et al. *Wenchengia alternifolia* isolate HN chloroplast, complete genome. GenBank https://www.ncbi.nlm.nih.gov/nuccore/MN128379 (2020).

102. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–5.

103. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90.

104. Wu P, Xu C, Chen H, Yang J, Zhang X, Zhou S. NOVOWrap: An automated solution for plastid genome assembly and structure standardization. Mol Ecol Resour. 2021;21(6):2177–86.

105. Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, et al. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 2020;21(1):241.

106. Qu XJ, Moore MJ, Li DZ, Yi TS. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. Plant Methods. 2019;15(1):50.

107. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28(12):1647–9.

108. Stephan G, Pascal L, Ralph B. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 2019;47(W1):W59–64.

109. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.

110. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 2007;56(4):564–77.

111. Löhne C, Borsch T. Molecular evolution and phylogenetic utility of the *petD* group II intron: a case study in basal angiosperms. Mol Biol Evol. 2005;22(2):317–32.

112. Green BR. Chloroplast genomes of photosynthetic eukaryotes. Plant J. 2011;66(1):34–44.

113. Zhang X, Sun Y, Landis JB, Lv Z, Shen J, Zhang H, et al. Plastome phylogenomic study of Gentianeae (Gentianaceae): widespread gene tree

Wang *et al. BMC Biology*   (2024) 22:185

Page 19 of 19

discordance and its association with evolutionary rate heterogeneity of plastid genes. BMC Plant Biol. 2020;20:340.

114. Schwarz EN, Ruhlman TA, Weng M-L, Khiyami MA, Sabir JSM, Hajarah NH, et al. Plastome-Wide Nucleotide Substitution Rates Reveal Accelerated Rates in Papilionoideae and Correlations with Genome Features Across Legume Subfamilies. J Mol Evol. 2017;84(4):187–203.

115. Gonalves DJP, Simpson BB, Ortiz EM, Shimizu GH, Jansen RK. Incongruence between gene trees and species trees and phylogenetic signal variation in plastid genes. Mol Phylogenet Evol. 2019;138:219–32.

116. Nguyen L-T, Schmidt HA, Haeseler Av, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol Biol Evol. 2014;32(1):268–74.

117. Ronquist F, Teslenko M, Mark PVD, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61(3):539–42.

118. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, et al. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Resour. 2020;20(1):348–55.

119. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Mol Biol Evol. 2017;34(3):772–3.

120. Rambaut A. Figtree v1.4.4. http://tree.bio.ed.ac.uk/software/figtree/ (2018).

121. Li H, Guo Q, Xu L, Gao H, Liu L, Zhou X. CPJSdraw: analysis and visualization of junction sites of chloroplast genomes. PeerJ. 2023;11:e15326.

122. Amiryousefi A, Hyvonen J, Poczai P. IRscope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics. 2018;34(17):3030–1.

123. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 2004;14(7):1394–403.

124. Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, et al. VISTA: visualizing global DNA sequence alignments of arbitrary length. Bioinformatics. 2000;16(11):1046–7.

125. Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 2003;19(18):2496–7.

126. Beier S, Thiel T, Munch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. Bioinformatics. 2017;33(16):2583–5.

127. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 2001;29(22):4633–42.

128. Wang YH, Xu C, Guo X, Wang Y, Chen YY, Shen J, He CN, Yu Y, Wang Q. Phylogenomics analysis of *Scutellaria* (Lamiaceae) of the world. figshare https://doi.org/10.6084/m9.figshare.26213495.v1 (2024).

Publisher's Note