

Evolution and Structural Variations of Chloroplast tRNAs in Gymnosperms

Yu-he Zhao

Northwest University

Tong Zhou

Northwest University

Ting-Ting Zhang

Northwest University

Peng-Bin Dong

Northwest University

Min-Feng Fang

Northwest University

Jian-Ni Liu

Northwest University

Zhong-Hu Li (✉ lizhonghu@nwu.edu.cn)

Northwest University

Research article

Keywords: Chloroplast tRNA, Conservation, Evolution, Phylogenetic relationships, Minimum free energy

Posted Date: November 24th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-112799/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Transfer RNAs can participate in a variety of vital life activities. The gymnosperms with important ecological and economic values were the dominant and constructive species in the forest ecosystems in the North hemisphere. However, evolution and structural changes of chloroplast tRNA in gymnosperms remain largely unclear.

Results: In this study, the nucleotide evolution, phylogenetic relationships and structural variations of 1,779 chloroplast tRNAs in gymnosperms were determined. The number and types of tRNA genes contained in the chloroplast genome of different gymnosperms were not much different, the average amount of tRNA contained was 33, and the occurrence frequency of various types of tRNA in gymnosperms was generally consistent. Nearly half of the anti-codons were absent. Molecular sequence variation analysis revealed the conservative secondary structure of tRNA. About a quarter of the tRNA genes were found to contain precoded 3' CCA tails. A few tRNAs underwent novel structural changes that were closely related to their minimum free energy, these structural changes affected the stability of tRNA. Phylogenetic analysis showed that tRNAs have evolved from multiple common ancestors and tRNA^{Leu} was the most primitive type of tRNA. The transition rate of chloroplast tRNAs was higher than the transversion rate in gymnosperms. The tRNAs continuously have experienced more loss than duplication events during the evolutionary process.

Conclusions: These findings provided novel insights into the molecular evolution and biological characteristics of the chloroplast tRNA in gymnosperms.

Background

Gymnosperm was a large group of seed plants which widely distributed around the world. Gymnosperms were the dominant and constructive species in the forest ecosystems in the North hemisphere, constituting 39% of the world's forests, and have great ecological and economic significance [1]. According to the Christenhusz gymnosperms system, the extant gymnosperms were divided into 12 families, 86 genera, and about 1,063 species [2]. The most abundant group of existing gymnosperms was conifers, which occupy a similar niche to the early stages of evolution because they have strong drought resistance [3]. The genetic relationship between gymnosperms and angiosperms gives them important phylogenetic status [4]. Furthermore, gymnosperm has a long extensive fossil record and can back to the Carboniferous (c. 290 million years ago (mya)). The five main lineages of gymnosperms (Cycads, Ginkgos, Cupressophytes, Pinaceae and Gnetophytes) separated from each other during the Late Carboniferous to the Late Triassic (311–212 Mya) [5].

Transfer RNA (tRNA) was among the most ancestral RNA in the world, tRNA was ubiquitous in all living organisms, from prokaryotes to eukaryotes [6]. tRNA was a class of microRNAs that carry and transport amino acids, it played a central role as the adaptor between mRNA and protein. During protein translation, tRNA pairs its anti-codon with a codon on mRNA and carries specific amino acids to ribosome sites to

mediate protein biosynthesis [7, 8]. tRNA was a multifunctional molecule, in addition to its translation function, it involved in multiple metabolic processes in cells, like aminoacyl-tRNA was a biosynthetic precursor and amino acid donor for other macromolecules [9]. Each tRNA can carry only one amino acid, however, one amino acid can be carried by multiple tRNAs called isoreceptor tRNA. In 1965, the first tRNA was sequenced for its primary structure, which was a tRNA^{Ala} in yeast [10]. Subsequently, the sequencing of other tRNA was gradually completed. The secondary structures of tRNA were almost all conservative clover-shaped, it consisted of an amino acid receiving arm, a D-arm, an anti-codon arm, a D-loop (a loop coupled to the D-arm), an anti-codon loop (a loop coupled to the anti-codon arm), and a TΨC loop (a loop coupled to the TΨC arm) [11]. The nucleotide sequence of tRNA was hydrogen-bonded to form a clover-shaped secondary structure, which then folded into an inverted L-type tertiary structure [12].

Chloroplasts were multi-copy organelles present in plant cells, responsible for photosynthesis and carbohydrate metabolism [13]. In addition, chloroplasts were also responsible for the synthesis of nucleotides, amino acids, fatty acids, vitamins, phytohormones, and several other metabolites [14, 15, 16]. It can be seen that chloroplasts play a vital role in the growth and development of plants. The chloroplast genome was very conservative, which was a double stranded circular molecule containing tRNA, rRNA, and a number of protein coding genes [17, 18]. The semi-autonomous and complete expression system of plant plastid genome made it a good material for evolutionary and genomics research [19, 20]. At the same time, tRNA acts as a bridge in the process of gene expression. Therefore, the analyses of tRNA genes in chloroplast can provide a theoretical basis for the further study of tRNA structure, function and evolutionary relationship, which has great biological significance.

At present, systematic studies on the chloroplast tRNA gene of gymnosperms are still lacking. Due to the global widespread distribution of gymnosperms and their important value in many aspects, they have important research significance. In this study, we selected 54 species belonging to 54 different genera in the gymnosperm phylum, and systematically analyzed their chloroplast tRNA. Based on the extraction and re-annotation of tRNA genes in the chloroplast genome of each species to discuss the differences in the composition, conservation and structural changes of tRNA in chloroplast of different plants, the evolutionary relationship and the main events that tRNA experienced in the evolutionary process. This study mainly aims at the following questions: (1) How about the distribution rules and conservativeness of different types of tRNA? (2) Why does certain tRNAs always contain precoded 3'CCA tails? (3) How does the minimum free energy affect the stability of tRNA's secondary structure? and (4) the main types of events that tRNA has experienced during its evolutionary history.

Results

Gene compositions of gymnosperm chloroplast tRNA

In 54 chloroplast genomes of gymnosperms (Table S1), 1,779 tRNA genes were totally annotated, encoding 20 essential amino acids. The chloroplast tRNA gene content of plants was relatively uniform [8]. In this study, the number of chloroplast tRNA genes in each species was approximately 33 in average.

Very few species encode only 27 tRNAs, and very few species encode up to 39 tRNAs. *Callitris rhomboidea*, *Dacrycarpus imbricatus* and *Pseudotaxus chienii* encoded 27 tRNAs; *Gnetum parvifolium* and *Macrozamia mountperriensis* encoded 39 tRNAs (Table 1).

In the chloroplast genome of each species, almost every tRNA was encoded, but some tRNAs were not encoded in some species (Table 1). tRNA^{Ala} was found to be missing in 8 species; tRNA^{Thr}, tRNA^{Glu}, tRNA^{Phe} and tRNA^{Leu} were found to be missing in 1 species; tRNA^{Val} was found to be missing in 2 species; tRNA^{Lys} were found to be missing in 14 species; tRNA^{Gln} were found to be missing in 5 species. There were more tRNA^{Ser}, tRNA^{Arg} and tRNA^{Leu} in the chloroplast genome of all species. tRNA^{Ser} appeared three times in most species and two or four times in some species, and appeared six times in *Nothotsuga longibracteata*; tRNA^{Arg} and tRNA^{Leu} commonly appeared in various species for 2-3 times, tRNA^{Leu} appeared six times in *Ephedra equisetina*. tRNA^{Gly}, tRNA^{Pro} and tRNA^{Thr} were the second largest groups, occurring twice in most species. In contrast, suppressor tRNA and selenocysteine were completely absent in the chloroplast genome of gymnosperm.

In addition, the length of the chloroplast tRNAs ranged from 56 nt to 90 nt while the average length was about 82 nt. tRNA^{Gly} (UCC) of *Cunninghamia lanceolata* is the smallest gene found in this study, only contains 56 nucleotides. The sequence length of tRNA^{Leu}, tRNA^{Ser} and tRNA^{Tyr} were all more than 80 nt. A few tRNA^{Ser} were found to reach 90 nt, and tRNA^{Gly} (UCC) of *Sequoia sempervirens* was also 90 nt in length. The length of other tRNA groups was all about 73 nt, but a few were shorter than 70 nt.

Gymnosperm chloroplast tRNAs contain 34 anti-codons

The genetic code was degenerate, in that 20 amino acids were encoded by 61 triplet codes [21], but analyses showed gymnosperm tRNA contain 34 different anti-codons in 1,779 tRNAs, the rest of the 27 anti-codons were not found in any of the tRNAs of the investigated gymnosperm chloroplast genomes (Table 2). The anti-codons encoded in this study include: tRNA^{Ala} (UGC), tRNA^{Gly} (GCC, UCC), tRNA^{Pro} (GGG, UGG), tRNA^{Thr} (GGU, UGU), tRNA^{Val} (GAC, UAC), tRNA^{Ser} (GGA, UGA, GCU), tRNA^{Arg} (ACG, CCG, UCU), tRNA^{Leu} (GAG, UAG, CAA, UAA), tRNA^{Phe} (GAA), tRNA^{Asn} (GUU), tRNA^{Lys} (UUU), tRNA^{Asp} (GUC), tRNA^{Glu} (UUC), tRNA^{His} (GUG), tRNA^{Gln} (UUG), tRNA^{Ile} (GAU, CAU), tRNA^{Met} (CAU), tRNA^{Tyr} (GUA), tRNA^{Cys} (GCA), tRNA^{Trp} (CCA). Among them, tRNA^{Leu} has the highest abundance of isoreceptor tRNA^{Leu} (GAG, UAG, CAA, UAA), followed by tRNA^{Ser} (GGA, UGA, GCU), tRNA^{Arg} (ACG, CCG, UCU) and tRNA^{Ile} (GAU, CAU, UAU). In addition, tRNA^{Leu} (GAG) present only in *Ephedra equisetina*; tRNA^{Lys} (CUU) present only in *Cunninghamia lanceolata*. tRNA^{Ile} (UAU) present only in *Taxus baccata*. tRNA^{Met} (CAU) presents at least twice in each species. tRNA^{Gly} (GCC), tRNA^{Pro} (UGG), tRNA^{Ser} (UGA, GCU), tRNA^{Arg} (ACG), tRNA^{Asn} (GUU), tRNA^{Asp} (GUC), tRNA^{His} (GUG), tRNA^{Ile} (CAU), tRNA^{Met} (CAU), tRNA^{Tyr} (GUA) and tRNA^{Trp} (CCA) present in every gymnosperms chloroplast genomes investigated.

Conservation of gymnosperm chloroplast tRNAs

Different tRNAs can transport different amino acids due to their different nucleotide compositions and structures. tRNA sequences were analyzed to find their conserved regions (Table 3). Through comparative analysis of the nucleotide composition of tRNA loops and arms, it was found that there were conserved nucleotides or nucleotide sequences in multiple positions. Analyses revealed that at the first position in acceptor arm, tRNA^{Ala} (UGC), tRNA^{Gly} (GCC, UCC), tRNA^{Thr} (GGU), tRNA^{Ser} (GCU, GGA, UGA), tRNA^{Arg} (ACG, CCG), tRNA^{Leu} (CAA, GAG, UAA, UAG), tRNA^{Lys} (UUU), tRNA^{Phe} (GAA), tRNA^{Asp} (GUC), tRNA^{Glu} (UUC), tRNA^{His} (GUG), tRNA^{Ile} (CAU, GAU), tRNA^{Tyr} (GUA) and tRNA^{Cys} (GCA) contain a conserved 5'G nucleotide; but tRNA^{Pro} (UGG), tRNA^{Met} (CAU) and tRNA^{Val} (GAC, UAC) contain conserved A nucleotide; tRNA^{Asn} (GUU) and tRNA^{Gln} (UUG) were found to contain U nucleotide. But tRNA^{Trp} (CCA), tRNA^{fMet} (CAU) and tRNA^{Thr} (UGU) were not very conservative at the first position in acceptor arm. And the content of G nucleotides was higher in the region of the acceptor arm. And tRNA^{Ser} (GCU, UGA) possess conserved nucleotide sequences G-G-A-G-A-G-A in the acceptor arm. At the first position of D-arm, tRNA^{Val} (GAC) and tRNA^{Lys} (UUU) contain a conserved A nucleotide; tRNA^{Tyr} (GUA) contain a conserved C nucleotide; tRNA^{Pro} (GGG) and tRNA^{Thr} (GGU) contain an A or G nucleotide at this position; tRNA^{Met} (CAU) contained no conserved nucleotides at this position; all the others tRNAs were all found to contain a conserved G nucleotide. At the same time, tRNA^{Ala} (UGC), tRNA^{Thr} (UGU), tRNA^{Val} (UAC), tRNA^{Arg} (ACG, CCG), tRNA^{Leu} (GAG), tRNA^{Phe} (GAA), tRNA^{Asn} (GUU) and tRNA^{Ile} (GAU) contain a conserved nucleotide sequence G-C-U-C; tRNA^{Cys} (GCA), tRNA^{His} (GUG) and tRNA^{Gln} (UUG) contain a conserved nucleotide sequence G-C-C. The D-loop was found to contain a conserved A nucleotide at the first position, except tRNA^{Gly} (GCC), tRNA^{Ser} (GCU, GGA), tRNA^{Leu} (UAA) and tRNA^{Ile} (CAU). And the last position of D-loop was very conserved, there was always a conserved A nucleotide except tRNA^{Gly} (GCC). Conservative degree of the anti-codon arms were lower since there was no conservative nucleotide in any position (Table 3). The second position of anti-codon loop was found to contain a conserved T nucleotide. The last position of anti-codon loop was mostly a conserved A nucleotide. In addition, it can be noted that the anti-codon loop had a high preference for uracil and adenine. Moreover, nucleotide conservation was very low in the variable region due to its structural variability. But at the last position of the variable region, there were still many tRNAs possess a conserved C nucleotide. The Ψ -arm and Ψ -loop were the most conservative regions whether in nucleotide number or composition of nucleotides. The Ψ -arms all contained five nucleotides, at the last two positions of this region, the vast majority of tRNAs were G nucleotides. The Ψ -loops all contain seven nucleotides, there was found a totally conserved sequence U-U-C; and vast majority of tRNAs had a conserved U nucleotide at the last position.

Possession of an intact CCA was a basic prerequisite for tRNAs to participate in the decoding process [22]. The 3' terminal of eukaryotic tRNAs generally did not have a CCA sequence, so adding a 3' CCA tail is an important step in tRNA biosynthesis. In the gymnosperms of this study, tRNA^{Ala}, tRNA^{Arg}, tRNA^{Glu}, tRNA^{Leu}, tRNA^{Tyr} and tRNA^{Lys} were found to contain a 3' CCA tail (Figure. 1), most tRNAs did not have 3' CCA tails.

Nucleotide variation in the arms and loops of tRNA

Conservation was also reflected in the number of nucleotides in each loop arm of tRNA. Out of all 1,779 tRNAs in this study, the number of nucleotides in the acceptor arm ranges from 0 to 8 (Table 4). Acceptor arms contained 7 nucleotides accounted for the vast majority (93.25%). And 58 (3.26%) tRNAs contained 6 nucleotides in the acceptor arm. The D-arm of most tRNAs contained 3 (34.23%) or 4 (65.65%) nucleotides. However, there were two tRNAs with a specific D-arm contained only one nucleotide, and both the two tRNAs were from *Pseudotsuga sinensis* var. *wilsoniana*. The D-loops had 6 to 26 nucleotides. Among the 1,779 tRNAs, 341 (19.17%) of the D-loops contained 7 nucleotides, 281 (15.8%) contained 8, 719 (40.42%) contained 9, 162 (9.11%) contained 10, 249 (14.00%) contained 11, 25 (1.41%) contained 12, one contained 6 nucleotides, and one contained 26 nucleotides. The anti-codon arm was identified as contained 4 or 5 nucleotides and none of the tRNAs possessed less than 4 or more than 5 nucleotides in the anti-codon arm. Most anti-codon loops contained 7 nucleotides (97.98%), the rest contained 9, 10 or 12 nucleotides. The number of nucleotides in the variable region changed significantly, most (1049, 58.97%) contained 5 nucleotides, the others contained 1 (0.17%), 2 (0.39%), 3 (5.45%), 4 (15.91%), 6 (12.65%), 7 (3.54%), 8 (0.06%), 11 (2.08%), 15 (0.06%), 16 (0.73%), 17 (0.06%) or 20 (0.06%). Among all the 1,779 tRNAs, only one tRNA^{Met} had 6 nucleotides in the Ψ -arm, 11 (0.62%) contained 4, while the remaining tRNAs contained 5 nucleotides. All tRNAs possess 7 nucleotides in the Ψ -loop.

Four types of structural changes in tRNA

The general structural characteristics of tRNA were: an amino acid receiving arm, a D-arm, a D-loop, an anti-codon arm, an anti-codon loop, a variable region loop, a T Ψ C arm and a T Ψ C loop. However, in this study, some novel tRNA structures were found. These new tRNA structures can be divided into four types (Table 5, Figure. 2). Type 1: lacked of acceptor arm; type 2: the 3'-end contained extra nucleotides; type 3: the variable region contained loops or arms; type 4: the 3'-end contained extra nucleotides while the variable region contained a loop or an arm. Among the tRNA structures that produce changes, type 3 shows obvious conservation. The variable region of tRNA^{Leu} (CAA, UAA), tRNA^{Ser} (GGA, UGA, GCU) and tRNA^{Tyr} (GUA) in all species had the same structure: extra loops and arms. tRNA^{Leu} also possessed an anti-codon UAG, but the variable region of this tRNA did not have such a structure. And in type 4, there were only two kinds of tRNAs: tRNA^{Tyr} (GUA) and tRNA^{Ser} (UGA).

We calculated the minimum free energy (ΔG) of each type of novel structural tRNA as well as some normal structural tRNAs (Table 5). The result shows, the minimum free energy of type 1 tRNAs was -12.6 kcal/mol on average. tRNA^{Gly} (UCC) of *Cunninghamia lanceolata* was very high ($\Delta G = -9.6$ kcal/mol), it shows that the absence of the acceptor arm leads to a very large impact on the stability of its structure. The minimum free energy of all tRNAs in type 1 was much higher than normal structural tRNAs ($\Delta G = -26.5$ kcal/mol). Therefore, the absence of the acceptor arm can reduce the structural stability of tRNA in

general. The minimum free energy of type 2 was around -19.3 kcal/mol. In type 2, tRNA^{Gly} (GCC) of *Sequoia sempervirens* possess the lowest value ($\Delta G = -28.3$ kcal/mol), indicates that the presence of extra nucleotides at the 3' end greatly improves the stability of the structure. On the contrary, tRNA^{Met} (CAU) of *Cephalotaxus oliveri* possess a highest value ($\Delta G = -11.8$ kcal/mol), the stability was greatly reduced due to the presence of atypical nucleotides at the 3' end. The minimum free energy of type 3 was -33.2 kcal/mol on average. The minimum free energy values of these tRNAs were basically below -30.0 kcal/mol. tRNA^{Tyr} (GUA) always possess a very low value. Therefore, the loops and arms that appear in the variable region work together with the structure of other regions to create an extremely stable tRNA structure. However, compared with other tRNAs in type 3, tRNA^{Leu} (CAA) was found to be special because of its higher minimum free energy value. The value of tRNA^{Leu} (CAA) was around -26.1 kcal/mol, it's much higher than the average of type 3 ($\Delta G = -32.8$ kcal/mol) and close to the normal structural tRNAs ($\Delta G = -26.5$ kcal/mol). So for tRNA^{Leu} (CAA), the structural changes of the variable region did not bring any obvious effects. The minimum free energy of type 4 was -28.3 kcal/mol on average, and the values of each tRNA were quite different, some were above the average value and some were below. Therefore, when the structural changed at the 3'end and the variable region coexist, multiple influences may be caused. Moreover, took the value of normal structural tRNAs ($\Delta G = -26.5$ kcal/mol) as a reference, type 1 and type 2 were much higher than the reference value; type 3 was lower than the reference value. As can be seen, various changes in the structure of tRNA can impact its stability.

Gymnosperm tRNA were evolved from multiple common ancestors

Constructed a phylogenetic tree used the maximum likelihood method to discuss the evolutionary relationship of all gymnosperm tRNAs (Figure. 3). Phylogenetic tree contained 2 large clusters and 32 small groups. Cluster I was much larger than cluster II as can be seen contained 28 groups while Cluster II contained 4. Not every type of anti-codon occupied a group, and those anti-codons that occurred less frequently often appear in the same branch as other anti-codons. Such as tRNA^{Lys} (CUU), it appeared only once in *Cunninghamia lanceolata*, then it grouped with tRNA^{Asn} (GUU) together. And tRNA^{Ile} (UAU) appeared only once in *Taxus baccata*, it grouped with tRNA^{Val} (GAC). The same thing happened to tRNA^{Leu} (GAG) which appeared twice in *Ephedra equisetina*, and it grouped with the branch of tRNA^{Ile} (GAU). In addition, tRNA^{Thr} (GGU) and tRNA^{Pro} (GGG) were grouped twice.

At the top clade of this phylogenetic tree, the branches of tRNA^{fMet} (CAU) and tRNA^{Trp} (CCA), tRNA^{Asn} (GUU), tRNA^{Arg} (ACG), and tRNA^{Arg} (CCG) together presented a stepwise evolutionary relationship. However, the other anti-codon UCU of tRNA^{Arg} did not appear with tRNA^{Arg} (ACG) and tRNA^{Arg} (CCG), it co-presented another stepwise evolutionary relationship with tRNA^{Glu} (UUC), tRNA^{Gly} (UCC) and tRNA^{Lys} (UUU). The three anti-codons (GGA, UGA and GCU) of tRNA^{Ser} occurred simultaneously and grouped together with tRNA^{Gln} (UUG) in the same branch. It suggested that tRNA^{Gln} and tRNA^{Ser} possessed a

common evolutionary lineage. The three anti-codons (CAA, UAA and UAG) of tRNA^{Leu} occurred simultaneously in one branch and they were at the bottom of the phylogenetic tree, indicated that they were the earliest evolved tRNA in the chloroplast genome of gymnosperms. The branches of tRNA^{Leu} and tRNA^{Pro} (GGG) together formed the second cluster. Therefore, tRNA^{Pro} and tRNA^{Leu} had a close relationship and were far from the first cluster of tRNA groups. Moreover, tRNA^{Thr} (UGU), tRNA^{Val} (UAC) and tRNA^{Ala} (UGC) were grouped together, indicated their common evolutionary lineage. Similarly, it can be seen the common evolutionary lineage of tRNA^{Met} (CAU), tRNA^{Thr} (GGU) and tRNA^{Val} (GAC) since they were in the same branch. And the branch of tRNA^{Tyr} (GUA), tRNA^{Pro} (UGG), tRNA^{Cys} (GCA) and tRNA^{His} (GUG) had the same case. We can also see from the phylogenetic tree that tRNA^{Phe} (GAA) and tRNA^{Ile} (GAU) were grouped separately, they each occupied a small branch instead of grouping together with the other types of tRNAs.

The rate of transition was higher than transversion

Transition refers to a change from one purine to another purine (A to G or G to A) or one pyrimidine to another pyrimidine (C to U/T or U/T to C). Transversion refers to a change from one purine to a pyrimidine (A or G to U/T or C) or the opposite (U/T or C to A or G) [23]. Analyzing patterns of base mutations can help understand the molecular basis of evolution. Table 6 showed the transition and transversion rates for each tRNA as well as the overall level of gymnosperm investigated. As can be seen from table 6, tRNA^{Asp} had the highest base transition rate (25.00). tRNA^{Phe} (22.41), tRNA^{Trp} (21.50) and tRNA^{Glu} (21.13) also had high transition rates. tRNA^{His} (11.40) had the lowest base transition rate. tRNA^{Ala} (12.53), tRNA^{Gly} (12.61) and tRNA^{Met} (13.35) also had low transition rates. In addition, in base transversion events, tRNA^{Ala} (6.23), tRNA^{Gly} (6.19) and tRNA^{His} (6.80) had relatively high values. In contrast tRNA^{Glu} (1.94), tRNA^{Phe} (1.29) and tRNA^{Trp} (1.75) had low transversion rates. The most special group was tRNA^{Asp}, which transversion rate was 0. Overall, the transition rate was higher than transversion rate, no transversion rate was found to be higher than the transition rate. Similarly, we calculated the values at the overall level of tRNA genes, and the result showed that the transition (18.3) rate was higher than transversion rate (3.19). Moreover, the transition rate was essentially inversely proportional to transversion rate. When a class of tRNA had a higher transition rate, it often had a lower transversion rate.

Duplication and loss events of gymnosperms chloroplast tRNA

After a gene duplication event, a copy of each replicated gene pair tends to happen a loss event. And the gene loss events have been happening all the time [24]. We calculated the duplication and loss events of gymnosperms chloroplast tRNA (Figure. 4). The results showed that 1,333 genes were duplicated whereas 3,657 genes were found to be lost. There were 314 genes happened conditional duplication

events. The loss events were far more than duplication events, majority of chloroplast tRNAs underwent loss events during the course of evolution.

Discussion

Distribution of tRNA

Analysis of gymnosperm chloroplast tRNAs revealed that tRNA genes were conserved whether in quantity or composition. The number of tRNA genes in the chloroplast genome of each species was not much different, and the rules of the frequency of each tRNA gene were basically the same, with only slight differences. Some tRNAs may be lost occasionally in a few species. For these missing tRNA genes, there may be tRNA genes from the nucleus or other organelles to replace their functions [25]. It has been found that tRNA^{Ser}, tRNA^{Arg} and tRNA^{Leu} always have higher occurrence frequency in chloroplast genome of gymnosperms. Also, the sequence length of tRNA^{Ser} and tRNA^{Leu} can be obviously longer due to the different nucleotides in the variable region. And it has been proven that tRNA^{Leu} had a large variable region [26]. The main function of the variable region of tRNA has not been thoroughly studied. However, it has been shown that larger variable regions can increase the affinity of tRNA for ribosomes and keep the tRNA-ribosomal complex stable in various environments to promote good interaction between tRNA and ribosome [27]. This may indicate that these types of tRNA were more commonly used in plant chloroplasts and were associated with more biological processes.

Suppressor tRNA was a mutated form of tRNAs, it read mRNA in a new way and allow insertion of appropriate amino acids at the mutation site of the protein-coding gene to suppresses the phenotypic effect of coding mutation, which affect the production of functional cellular protein [28,29,30,31]. So suppressor tRNA was not found in the chloroplast genome of gymnosperms. At the same time, selenocysteine was completely lacked. Selenocysteine was a kind of atypical amino acids [32]. It was the 21st amino acid in ribosome-mediated synthesis of proteins with a codon of UGA. And selenocysteine was found in both prokaryotes and eukaryotes [33]. However, selenocysteine was an oxygen-labile amino acid and had certain toxicity [34], so it wasn't found in chloroplast genome of gymnosperms. In addition, it was found that there was at least one tRNA^{Met} and one tRNA^{fMet} in each species, both corresponding to the anti-codon CAU. It is currently known that tRNA^{fMet} is necessary for the initiation of the protein translation process in prokaryotes [35,36,37,38]. The initiator tRNA was fairly conservative at all times [39]. And in plants, tRNA^{fMet} (CAU) and tRNA^{Met} (CAU) were both indispensable [40]. Interestingly, tRNA^{Met} and tRNA^{Ile} contained the same anti-codon CAU. There was a complex relationship between their identification and matching with codons. Previous studies [41,42,43] have shown that, in bacteria, when the C nucleotide in the anti-codon CAU was modified, tRNA^{Ile} can recognize isoleucine; while the unmodified tRNA^{Ile} with the CAU anti-codon will interact with methionine. And it was proved that the same situation exists in plant chloroplasts [40]. These events can be explained by the prokaryotic origin of the chloroplast.

Distribution of anti-codon

There were totally 64 codons in the genetic code, 61 of them can encode amino acids and 3 were stop codons, but they usually did not all show up together. In gymnosperm chloroplast tRNA genes, there were only 34 kinds of anti-codons found. Some anti-codons occurred in every species, and some occurred only occasionally in a few species. These 34 kinds of anticodons were able to perform all 61 kinds of anticodons that were responsible for all protein translation in the chloroplast. One reason is the degeneracy of genetic code, in other words, the “Wobble Hypothesis”. The first and second bases of the codon can be firmly paired with the anti-codon, but the third base can form a non-Watson-Crick base pair with the anti-codon [44]. So for some tRNA species, it corresponded to multiple anti-codon types, that is, one amino acid can be carried by multiple tRNAs. Substitutions in protein-coding genes were usually distributed according to the codon structure, and substitutions often occurred at the third position of the codon. Moreover, multiple anti-codons corresponding to one tRNA have the same “evolution potential” [45,21]. Another reason is the organism’s preference for codon usage. The use of synonymous codons was non-random and had a certain preference, which was mainly determined by the translation process [46]. There was a strong correlation between codon use and tRNA content, and the selection pattern of codon showed a high conservatism in the evolutionary process. Genes with high expression levels often have codons corresponding to more abundant tRNA species [47,48,49]. That is also to say gene expression levels were largely related to codon usage preference [50,51]. This is consistent with the results in Table 1 and Table 2: the total frequency of codons contained in tRNA^{Ser}, tRNA^{Arg}, and tRNA^{Leu} was higher. The selectivity of codons in organisms was due to the fact that when the tRNA usage was high, the presence of richer codons can greatly reduce the risk of depleting the translation mechanism [52].

tRNA secondary structure was very conservative

The secondary structure of tRNA was a shape of clover leaf, with the acceptor arm containing 7 nucleotides, the D-arm containing 3-4, the D-loop containing 4-12, the anti-codon arm containing 5, the anti-codon loop containing 3, the variable region containing 4-23, the Ψ -loop containing 5, and the Ψ -arm containing 7 [53,54]. However, a few of the chloroplast tRNAs in gymnosperms were different, and not all tRNAs fully conform to the traditional secondary structure. Moreover, the difference in the number of nucleotides in different regions of tRNA was largely related to the type of tRNA, and even varies with the corresponding anti-codon. For example, in the variable region, tRNA^{Gly} (GCC) contained 4 nucleotides, but tRNA^{Gly} (UCC) contained 5 nucleotides (Table 3). Not only the number of nucleotides, but also the composition of nucleotides in different regions, have similar rules. It can be seen that the sequence of tRNA was very conservative, there were common nucleotides or common sequences in almost every region, and the conservation was largely related to the type of tRNA. From the results of tRNA sequence alignment, the Ψ -loop was the most conserved without any changes, and the Ψ -arm was also extremely conserved, only a small part of the tRNA was mutated in this region. And in Ψ -loop there was a common

sequence U-U-C. And it was reported that in thermophilic bacteria, the conserved bases in the Ψ -loop were the determinant of the stability of the tRNA [55]. Anticodon loops were also highly conserved, with most of them contained seven nucleotides. Because the anticodon loop was the region that matches the codon of the mRNA, it required higher accuracy. Under the mediation of tRNA nucleotidyltransferase, adding a conservative C-C-A sequence to the 3' end of tRNA was necessary for tRNA maturation. Only when the CCA tail was obtained can it carry amino acids [56,57]. However, not all cells need to add CCA tails with the help of tRNA nucleotidyltransferase. It has been reported that in bacteria, CCA tails were sometimes included in tRNA gene templates. It has been reported that the templated 3'CCA sequence in bacteria was very common in the initial tRNA (tRNA^{fMet}), and it was also common in tRNA^{Tyr} [58]. Coincidentally, in our gymnosperm plant chloroplast tRNA gene, tRNA^{fMet} and tRNA^{Tyr} of each species all carried the encoded 3'CCA sequence. This suggests that during the evolution of chloroplast, it retained part of the prokaryotic translation mechanism. In addition, the biggest influencing factor of protein synthesis was the initiation of translation [59,60], so the templating of 3'CCA can greatly improve the rate of protein expression, because it accelerates the maturation of tRNA.

Phylogenetic relationships

In phylogenetic analysis of all tRNA genes, it was found that tRNA^{Leu} (CAA) was the evolutionary ancestor of them. The same results were obtained in the phylogenetic analysis of chloroplast tRNA in monocots [61]. And tRNA^{Pro} (GGG), tRNA^{Leu} (UAG, UAA, CAA) in the second cluster were primitive types of tRNAs. In addition, tRNA^{fMet} (CAU) appeared twice in the phylogenetic tree, once at the top of the tree, grouped with tRNA^{Trp} (CCA) together, and once in the lower part of the tree, grouped with tRNA^{Thr} (GGU) and tRNA^{Val} (GAC) together. This indicates that tRNA^{fMet} (CAU) was evolved from multiple common ancestors. It shows that tRNA^{fMet} (CAU) has experienced more frequent duplication events in the evolution process.

Effects of structural changes on the stability of tRNA

Structural changes of tRNA were very rare because it's extremely conserved [62]. The minimum free energy of a molecule was closely related to its structure. It was the guarantee of the thermodynamic stability of RNA [63]. The minimum free energy can be used to predict the secondary structure of RNA [64,65,66]. In thermophiles, tRNA folding undergoes adaptive changes to improve its stability. That is, changes in tertiary structure can have a certain effect on the stability of tRNA [67]. In this study, there were some changes in the structure of tRNA, which can be roughly divided into four categories, and there were obvious rules in the corresponding minimum free energy values. Compared with the normal structural tRNAs, the structural changes will have two effects on the minimum free energy of tRNA (make it higher or lower). It is reported that changes in the acceptor arm will increase the free energy of tRNA [68]. This is consistent with the results of this study: the free energy is higher when the acceptor arm is lacking or redundant nucleotides are present. tRNAs with large variable regions were not rare, and even large variable regions have evolved into conserved structures in some types of tRNAs, such as tRNA^{Leu}.

So it indicates this kind of structural change greatly reduced the free energy of tRNA, which improved the stability of tRNA structure. We have learned from table 1 that the occurrence frequency of tRNA^{Leu} was relatively high, which may indicate that this kind of structural change of tRNA^{Leu} was beneficial to its utilization by plants.

Evolution of substitution rate

There were 8 forms of transversion and 4 forms of transition. From the perspective of probability, transversion should be more frequent, but the statistical results show a high transition rate. This phenomenon was called “transition bias” [69]. The reason is that the transition has less effect on the protein than the transversion [70]. Because conversion was the substitution of bases of the same type, and inversion was the substitution of bases of different types. Within the respective family members of purines and pyrimidines, there was little structural difference, while the structural differences between purines and pyrimidines were large. So transition has less effect on protein structure. In addition, the transversion of tRNA^{Asp} was 0. One possibility is that it has not undergone transversion during its evolution. Another possibility is that if a transversion occurs in this gene, the synthesis of this tRNA will be terminated, resulting in an undetectable transversion rate. Overall, the chloroplast tRNA of gymnosperms mainly underwent base transversion during evolution.

Evolution of duplication and loss

In plant genomes, gene duplication and loss occur very frequently and they were important factors in evolution [71, 72, 73]. Throughout the evolutionary history of chloroplasts, the size of its genome has been shrinking, and it continues to experience gene loss events [74]. Also, it can be seen from Table 1 and Table 2 that some tRNAs were absent in certain species, and nearly half of the anti-codons were also absent. Combined with the analysis of tRNA gene duplication and loss, it can be concluded that the chloroplast tRNA gene of gymnosperm mainly experienced loss events during evolution.

Materials And Methods

Acquisition of chloroplast tRNA genes and analysis of secondary structure

Chloroplast genomes of 54 gymnosperms (Table S1) were downloaded from the public database of the National Center for Biotechnology Information (NCBI, [https:// www.ncbi.nlm.nih.gov/](https://www.ncbi.nlm.nih.gov/)). Annotation of the chloroplast genomes and extraction of tRNA genes were carried out in geneious [75]. All the sequences obtained of gymnosperms chloroplast tRNA genes were uploaded to the tRNAscan-Se server to predict the secondary structure and record other related results [76]. The free energy calculation of tRNAs with structural changes were performed using the RNAalifold webserver with default parameters.

Multiple sequence alignment

All tRNAs were classified according to their different types to find the consensus sequences of each regions. Similarly, the consensus sequences of each regions were observed at the overall level of tRNA. Multiple sequence alignment of tRNA gene was performed in the Multalin server [77]. All of the sequences, were used in the alignment analysis with the following parameters in FASTA format; sequence input format, auto; display of sequence alignment, colored; alignment matrix, Blosum61-12-2; gap penalty at opening and extension, default; gap penalty at extremities, none and one iteration only, none; highest consensus value, 90% (default); whereas, low consensus value, 50% (default). In the displayed alignments, red indicates a similarity/conservation of 90% or more; blue indicates a sequence conservation less than 90%. Alignments displayed in black indicates no conservation.

Construction of phylogenetic tree

To reveal the phylogenetic relationship of all the tRNAs, a phylogenetic tree was constructed using MEGA7 software [78]. After aligning all tRNA gene sequences, choose the model "K2 + G + I". The other parameters for constructing the phylogenetic tree were: analysis, phylogeny reconstruction; statistical model, maximum likelihood; test of phylogeny, bootstrap method; no. of bootstrap replicates, 1000; substitution type, nucleotide; rates among sites, Gamma distributed with invariant sites (G + I), No of discrete Gamma categories, 5; gaps/missing data treatment, partial deletion; site coverage cutoff, 95%; and branch swap filter, very strong.

Analysis of transition and transversion

The tRNA gene sequences used to construct the phylogenetic tree were also used to calculate the rate of transition and transversion. All tRNA gene sequences were classified according to different types, and then calculate the transition and transversion rates respectively. Similarly, the same calculations were performed at the overall level of tRNA. Use the following parameters to calculate the rates of transition and transversion: analysis, substitution pattern estimation (ML); tree to use, automatic (neighbor-joining tree); statistical method, maximum likelihood; substitution type, nucleotide; model/method, Kimura2-parameter model; rates among sites, Gamma distributed (G); no. of discrete Gamma categories, 5; gaps/missing data treatment, partial deletion, site coverage cutoff 95%, and branch swap filter, very strong.

Analysis of gene duplication and loss

To calculate duplication and loss events of tRNA genes, the phylogenetic tree of the tRNA genes and the species tree need to be reconciled. Species tree of 54 gymnosperms was constructed in NCBI taxonomy

server (<https://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi>). Reconcile the species tree and gene tree in software Notung2.9 [79].

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

All the genomic tRNA sequences used in this study can be found at NCBI with accession numbers in Table S1.

Declaration of interest

The authors declare no competing interest.

Funding

This work was financially supported by the National Natural Science Foundation of China (31970359), the Shaanxi Science and Technology Innovation Team (2019TD-012), the Public health specialty in the Department of Traditional Chinese Medicine (Grants no. 2017-66, 2018-43 and 2019-68) and the Open Foundation of Key Laboratory of Resource Biology and Biotechnology in Western China (Ministry of Education) (Grants no. ZSK2017007 and ZSK2019008).

Author contribution statement

ZHL conceived this study. YHZ, TZ, TTZ and PBD collected the original data and performed the analyses. MFF and JNL conducted the statistical analyses. YHZ wrote the manuscript draft. ZHL and YHZ revised the manuscript. All authors read and approved the manuscript.

Acknowledgements

Not applicable.

References

1. Armenise L, Simeone MC, Piredda R, Schirone B. Validation of DNA barcoding as an efficient tool for taxon identification and detection of species diversity in Italian conifers. *Eur J Forest Res.* 2012;131:1337–53.
2. Christenhusz MJM, Reveal JL, Farjon A, Gardner MF, Mill RR, Chase MW. A new classification and linear sequence of extant gymnosperms. *Phytotaxa.* 2011;19:55–70.
3. Leitch AR, Leitch IJ. Ecological and genetic factors linked to contrasting genome dynamics in seed plants. *New Phytol.* 2012;194(3):629–46.
4. Wang XQ, Ran JH. Evolution and biogeography of gymnosperms. *Mol Phylogenet Evol.* 2014;75:24–40.
5. Magallón S, Hilu KW, Quandt D. Land plant evolutionary timeline: gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *Am J Bot.* 2013;100:556–73.
6. Cook AG, Fukuhara N, Jinek M, Conti E. Structures of the tRNA export factor in the nuclear and cytosolic states. *Nature.* 2009;461(7260):60–5.
7. Hopper AK, Phizicky EM. tRNA transfers to the limelight. *Gene Dev.* 2003;17:162–80.
8. Michaud M, Cognat V, Duchêne AM, Maréchal-Drouard L. A global picture of tRNA genes in plant genomes. *Plant J.* 2011;66:80–93.
9. Francklyn CS, Minajigi A. tRNA as an active chemical scaffold for diverse chemical transformations. *FEBS Lett.* 2010;584:366–75.
10. Holley RW. Structure of an alanine transfer ribonucleic acid. *JAMA.* 1965;194(8):868–71.
11. Ishitani R, Nureki O, Nameki N, Okada N, Nishimura S, Yokoyama S. Alternative tertiary structure of tRNA for recognition by a posttranscriptional modification enzyme. *Cell.* 2003;113(3):383–94.
12. Smith JD. Nucleotide sequence and function of transfer RNA and precursor transfer RNA. *Basic Life Sci.* 1973;1:197–208.
13. Knorr W, Heimann M. Uncertainties in global terrestrial biosphere modeling: a comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme. *Global Biogeochem Cycles.* 2001;15:207–25.
14. Blee E, Joyard J. Envelope membranes from spinach chloroplasts are a site of metabolism of fatty acid hydroperoxides. *Plant Physiol.* 1996;110:445–54.
15. Noctor G, Arisi ACM, Jouanin L, Foyer CH. Manipulation of glutathione and amino acid biosynthesis in the chloroplast. *Plant Physiol.* 1998;118:471–82.
16. Spetea C, Hundal T, Lundin B, Heddad M, Adamska I, Andersson B. Multiple evidence for nucleotide metabolism in the chloroplast thylakoid lumen. *Proc. Natl. Acad. Sci. USA* 2004, 101, 1409-14.

17. Kolodner R, Tewari KK Inverted repeats in chloroplast DNA from higher plants. *Proc. Natl. Acad. Sci. USA* 1979, 76 (1), 41–45.
18. Meeker R, Tewari KK. Divergence of tRNA genes in chloroplast DNA of higher plants. *BBA - Gene Struct Expr.* 1982;696(1):66–75.
19. Giles KL, Taylor AO. The control of chloroplast division in *Funaria hygrometrica* I. Patterns of nucleic acid, protein and lipid synthesis. *Plant Cell Physiol.* 1971;12(3):437–45.
20. Zerges W. Translation in chloroplasts. *Biochimie.* 2000;82(6–7):583–601.
21. Agris PF, Vendeix FAP, Graham WD. tRNA's wobble decoding of the genome: 40 years of modification. *J Mol Biol.* 2007;366(1):1–13.
22. Shi PY, Maizels N, Weiner AM. CCA addition by tRNA nucleotidyltransferase: polymerization without translocation? *EMBO J.* 1998;17(11):3197–206.
23. Zhao H, Li Q, Li J, Zeng C, Hu S, Yu J. The study of neighboring nucleotide composition and transition/transversion bias. *Sci China Ser C.* 2006;49(4):395–402.
24. Schnable JC, Springer NM, Freeling M Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proc. Natl. Acad. Sci. USA* 2011, 108 (10), 4069–4074.
25. Pino P, Aeby E, Foth BJ, Sheiner L, Soldati T, Schneider A, Soldati-Favre D. Mitochondrial translation in absence of local tRNA aminoacylation and methionyl tRNA Met formylation in Apicomplexa. *Mol Microbiol.* 2010;76(3):706–18.
26. Dock-Bregeon AC, Westhof E, Giegé R, Moras D. Solution structure of a tRNA with a large variable region: yeast tRNA^{Ser}. *J Mo Biol.* 1989;206(4):707–22.
27. Curran JF, Poole ES, Tate WP, Gross BL. Selection of aminoacyl-tRNAs at sense codons: the size of the tRNA variable loop determines whether the immediate 3' nucleotide to the codon has a context effect. *Nucleic Acids Res.* 1995;23(20):4104–8.
28. Hatfield DL, Smith DWE, Lee BJ, Worland PJ, Oroszlan S. Structure and function of suppressor tRNAs in higher eukaryote. *Crit Rev Biochem Mol.* 1990;25(2):71–96.
29. Choisne N, Martin-Canadell A, Small I. Transactivation of a target gene using a suppressor tRNA in transgenic tobacco plants. *Plant J.* 1997;11(3):597–604.
30. Beier H, Grimm M. Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. *Nucleic Acids Res.* 2001;29(23):4767–82.
31. Mohanta TK, Bae H. Analyses of genomic tRNA reveal presence of novel tRNAs in *oryza sativa*. *Front Genet.* 2017;8:90.
32. Lukashenko NP. Expanding genetic code: amino acids 21 and 22, selenocysteine and pyrrolysine. *Russ J Genet+.* 2010;46(8):1013.
33. Stadtman TC, Selenocysteine. *Annu Rev Biochem.* 1996;65(1):83–100.
34. Plateau P, Saveanu C, Lestini R, Dauplais M, Decourty L, Jacquier A, Blanquet S, Lazard M. Exposure to selenomethionine causes selenocysteine misincorporation and protein aggregation in

- saccharomyces cerevisiae. Sci Rep-UK. 2017;7:44761.
35. Harvey RJ. Growth and initiation of protein synthesis in *Escherichia coli* in the presence of trimethoprim. J Bacteriol. 1973;114:309–22.
 36. Arnold HH. Initiation of protein synthesis in *Bacillus subtilis* in the presence of trimethoprim or aminopterin. Biochim Biophys Acta. 1977;476:76–87.
 37. Baumstark BR, Spremulli LL, RajBhandary UL, Brown GM. Initiation of protein synthesis without formylation in a mutant of *Escherichia coli* that grows in the absence of tetrahydrofolate. J Bacteriol. 1977;129(1):457–71.
 38. Guillon JM, Mechulam Y, Schmitter JM, Blanquet S, Fayat G. Disruption of the gene for Met-tRNA(fMet) formyltransferase severely impairs growth of *Escherichia coli*. J Bacteriol. 1992;174(13):4294–301.
 39. Grosjean H, Crécy-Lagard V, Marck C. Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. FEBS Lett. 2010;584(2):252–64.
 40. Alkatib S, Fleischmann TT, Scharff LB, Bock R. Evolutionary constraints on the plastid tRNA set decoding methionine and isoleucine. Nucleic Acids Res. 2012;40:6713–24.
 41. Köhrer C, Mandal D, Gaston KW, Grosjean H, Limbach PA, RajBhandary UL. Life without tRNA^{Ile}-lysine synthetase: Translation of the isoleucine codon AUA in *Bacillus subtilis* lacking the canonical tRNA^{Ile}. Nucleic Acids Res. 2013;42:1904–15.
 42. Mandal D, Kohrer C, Su D, Babu IR, Chan CTY, Liu YC, Söll D, Blum P, Kuwahara M, Dedon PC, RajBhandary UL. Identification and codon reading properties of 5-cyanomethyl uridine, a new modified nucleoside found in the anticodon wobble position of mutant haloarchaeal isoleucine tRNAs. J Fluid Mech. 2014;20(2):177–88.
 43. Tomikawa C, Auxilien S, Guérineau V, Yoshioka Y, Miyoshi K, Hori H, Fourmy D, Takai K, Yoshizawa S. Characterization of redundant tRNA^{Ile}s with CAU and UAU anticodons in *Lactobacillus plantarum*. J Biochem. 2018;163(3):233–41.
 44. Crick FH. Codon-anticodon pairing: the wobble hypothesis. J Mo Biol. 1966;19(2):548–55.
 45. McClellan DA. The codon-degeneracy model of molecular evolution. J Mol Evol. 2000;50(2):131–40.
 46. Fuglsang A. Estimating the “effective number of codons”: the Wright way of determining codon homozygosity leads to superior estimates. Genetics. 2006;172(2):1301–7.
 47. Ikemura T. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. J Mol Biol. 1981;146:1–21.
 48. Gouy M, Gautier C. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. 1982;10:7055–74.
 49. Ikemura T. Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol. 1985;2:13–34.
 50. Duret L, Mouchiroud D. Expression pattern and, surprisingly, gene length shape codon usage in *Caenorhabditis*, *Drosophila*, and *Arabidopsis*. Proc. Natl. Acad. Sci. USA 1999, 96 (8), 4482–4487.

51. Coghlan A, Wolfe KH. Relationship of codon bias to mRNA concentration and protein length in *Saccharomyces cerevisiae*. *Yeast*. 2000;16:1131–45.
52. Fuglsang A. The effective number of codons for individual amino acids: some codons are more optimal than others. *Gene*. 2003;320:185–90.
53. Kirchner S, Ignatova Z. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat Rev Genet*. 2015;16(2):98–112.
54. Wilusz JE. Controlling translation via modulation of tRNA levels. *Wiley Interdiscip Rev RNA*. 2015;6(4):453–70.
55. Shigi N, Suzuki T, Tamakoshi M, Oshima T, Watanabe K. Conserved bases in the T ϕ C loop of tRNA are determinants for thermophile-specific 2-thiouridylation at position 54. *J Biol Chem*. 2002;277(42):39128–35.
56. Vortler S, Morl M. tRNA-nucleotidyl transferases: highly unusual RNA polymerases with vital functions. *FEBS Lett*. 2010;584(2):297–302.
57. Betat H, Morl M. The CCA-adding enzyme: A central scrutinizer in tRNA quality control. *Bioessays* 2015, 37(9), 975–82.
58. Ardell DH, Hou YM. Initiator tRNA genes template the 3'CCA end at high frequencies in bacteria. *BMC Genom*. 2016;17:1003.
59. Hersch SJ, Elgamal S, Katz A, Ibba M, Navarre WW. Translation initiation rate determines the impact of ribosome stalling on bacterial protein synthesis. *J Biol Chem*. 2014;289(41):28160–71.
60. Pop C, Rouskin S, Ingolia NT, Han L, Phizicky EM, Weissman JS, Koller D. Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation. *Mol Syst Biol*. 2014;10:770.
61. Mohanta TK, Khan AL, Hashem A, Allah EFA, Yadav D, Al-Harrasi A. Genomic and evolutionary aspects of chloroplast tRNA in monocot plants. *BMC Plant Biol*. 2019;19:39.
62. Masta SE, Boore JL. Parallel Evolution of Truncated Transfer RNA Genes in Arachnid Mitochondrial Genomes. *Mol Biol Evol*. 2008;25(5):949–59.
63. Miklós I, Meyer IM, Nagy B. Moments of the Boltzmann distribution for RNA secondary structures. *B Math Biol*. 2005;67(5):1031–47.
64. Zarringhalam K, Meyer MM, Dotu I, Chuang JH, Clote P. Integrating chemical footprinting data into RNA secondary structure prediction. *PLoS One*. 2012;7(10):e45160.
65. Hofacker IL. Energy-directed RNA structure prediction. *Methods Mol Biol*. 2013;1097:71–84.
66. Will S, Jabbari H. Sparse RNA folding revisited: space-efficient minimum free energy structure prediction. *Algorithm Mol Biol*. 2016;11(7):1–13.
67. Dutta A, Chaudhuri K. Analysis of tRNA composition and folding in psychrophilic, mesophilic and thermophilic genomes: indications for thermal adaptation. *FEMS Microbiol Lett*. 2010;305(2):100–8.
68. Hisanori K, Kiyoshi A. Rchange: algorithms for computing energy changes of RNA secondary structures in response to base mutations. *Bioinformatics*. 2012;28(8):1093–101.

69. Li WH, Wu CI, Luo CC. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol Biol Evol.* 1985;2(2):150.
70. Stoltzfus A, Norris RW. On the causes of evolutionary transition: transversion bias. *Mol Biol Evol.* 2016;33(3):595–602.
71. Durand D, Halldórsson BV, Vernet B. A hybrid micro-macroevolutionary approach to gene tree reconstruction. *J Comput Boil.* 2006;13(2):320–35.
72. Charon C, Bruggeman Q, Thareau V, Henry Y. Gene duplication within the Green lineage: the case of TEL genes. *J Exp Bot.* 2012;63(14):5061–77.
73. Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R. Gene duplication as a major force in evolution. *J Genet.* 2013;92(1):155–61.
74. Scharff LB, Bock R. Synthetic biology in plastids. *Plant J.* 2014;78(5):783–98.
75. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 2012;28(12):1647–9.
76. Lowe TM, Chan PP. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* 2016;44(W1):W54–7.
77. Mitchell C. MultAlin—multiple sequence alignment. *Bioinformatics.* 1993;9(5):614.
78. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for vigger datasets. *Mol Biol Evol.* 2016;33(7):1870–4.
79. Chen K, Durand D, Farach-Colton M, NOTUNG. A program for dating gene duplications and optimizing gene family trees. *J Comput Biol.* 2000;7(3–4):429–47.

Tables

Due to technical limitations, tables 1-6 are only available as a download in the Supplemental Files section.

Figures

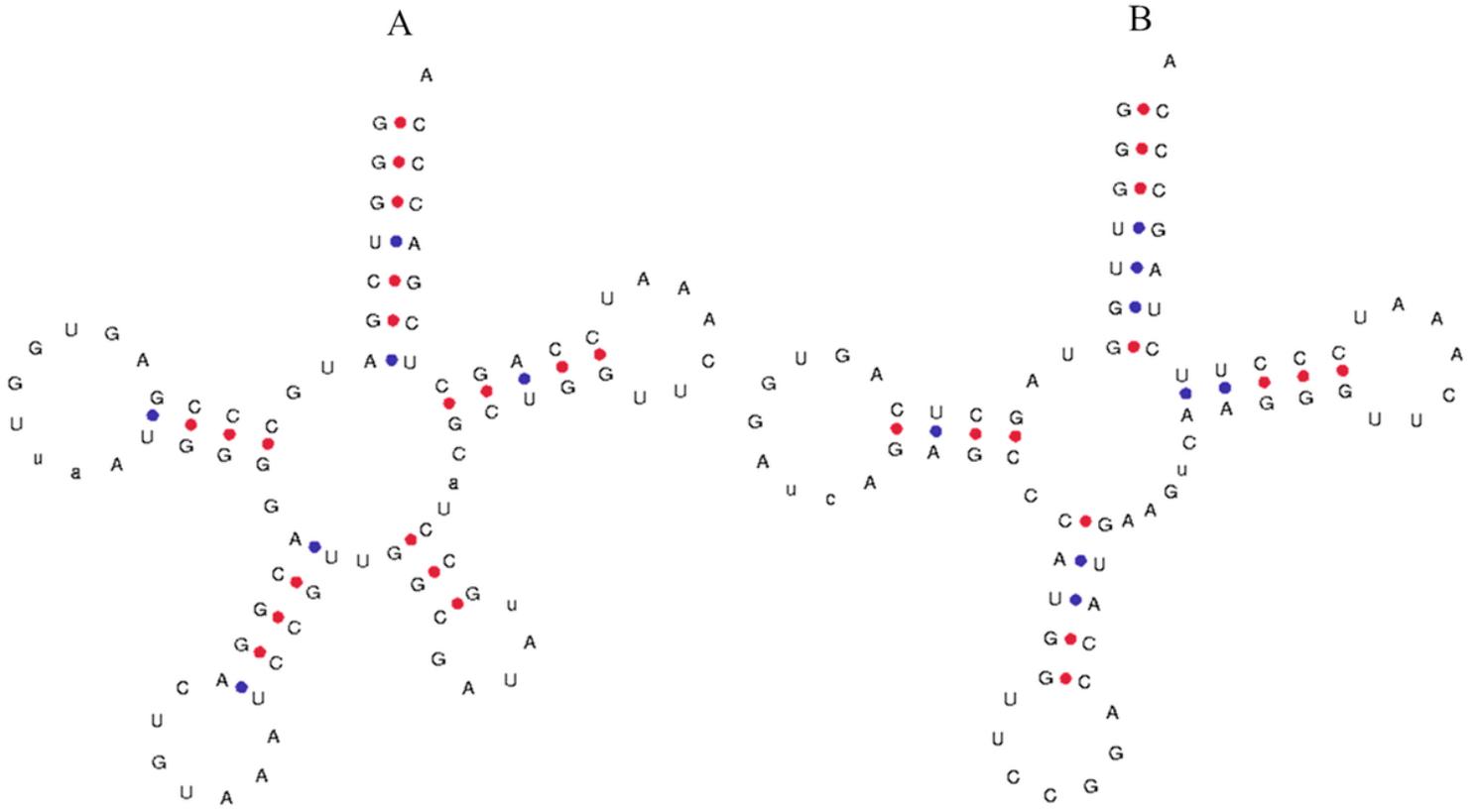


Figure 1

tRNAs with precoded 3' CCA tails. (A) tRNA^{Tyr} of *Abies koreana*. (B) tRNA^{Arg} of *Cedrus deodara*.

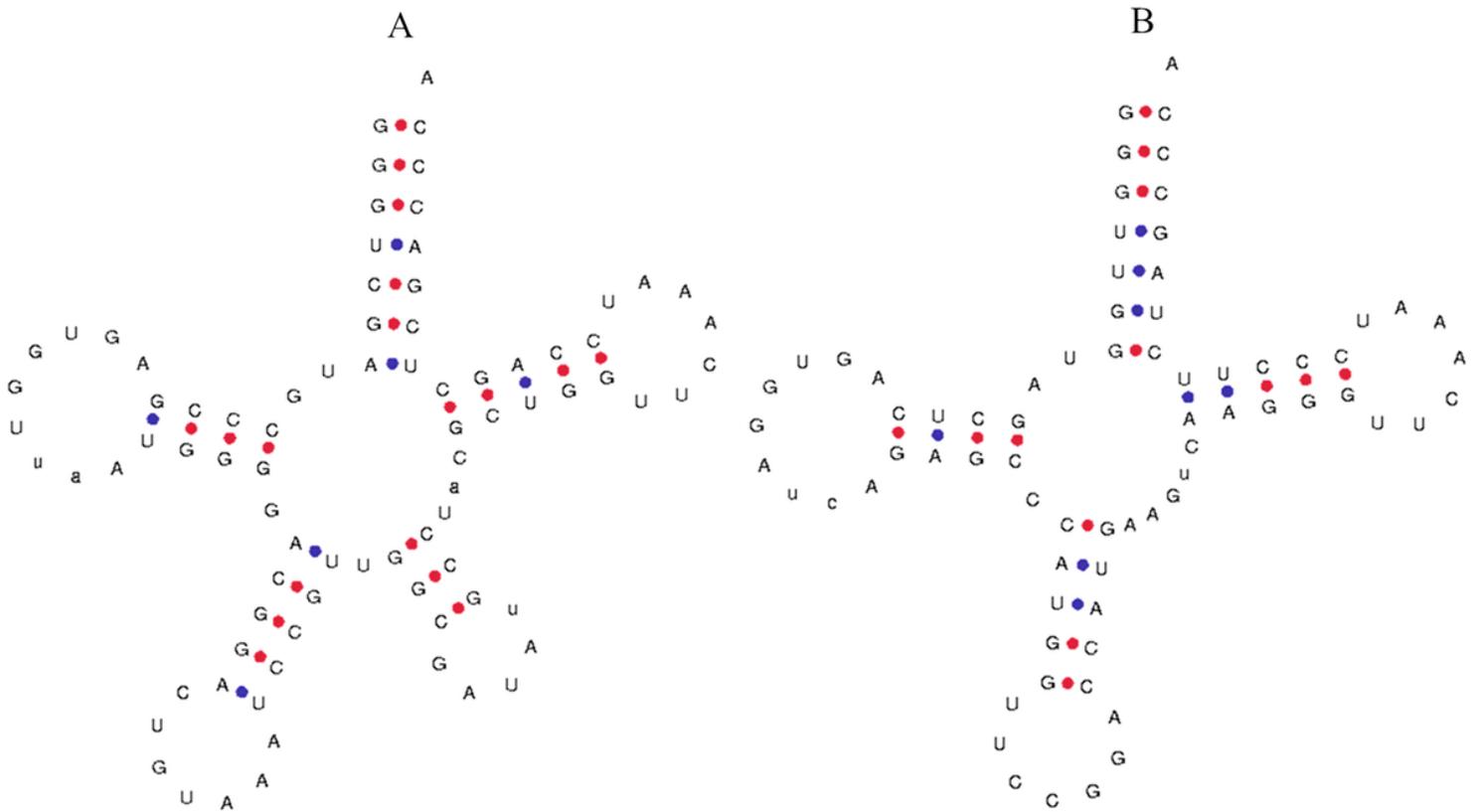


Figure 1

tRNAs with precoded 3' CCA tails. (A) tRNA^{Tyr} of *Abies koreana*. (B) tRNA^{Arg} of *Cedrus deodara*.

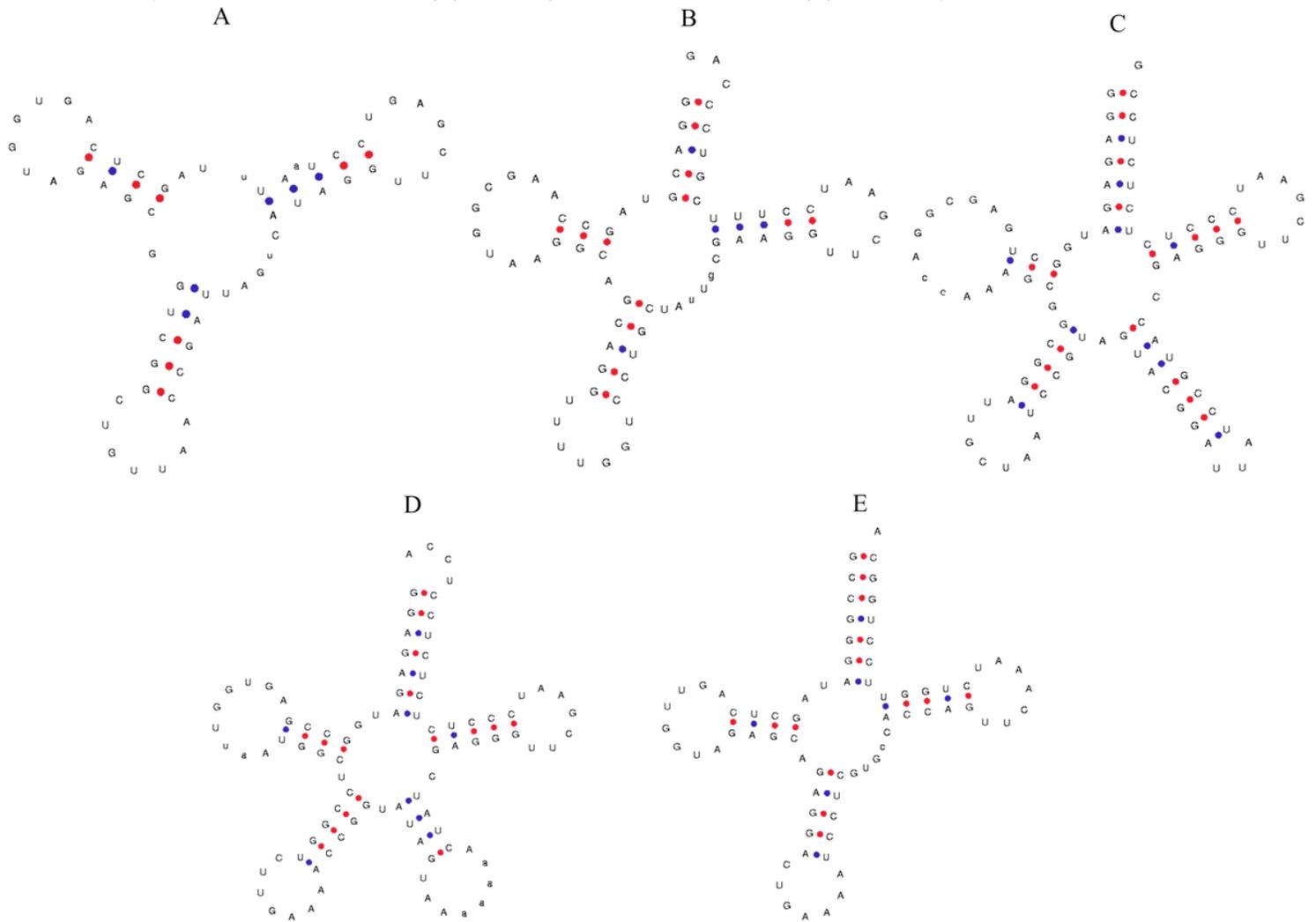


Figure 2

tRNAs with different structural changes. (A) type 1, lack of acceptor arm, tRNA^{Asn} (GTT) of *Nothotsuga longibracteata*, (B) type 2, the 3'-end contains extra nucleotides, tRNA^{Gln} (TTG) of *Sciadopitys verticillata*, (C) type 3, the variable region contains loops or arms, tRNA^{Ser} (GCT) of *Abies koreana*, (D) type 4, the 3'-end contains extra nucleotides while the variable region contains a loop and an arm, tRNA^{Ser} (TGA) of *Tsuga chinensis*, (E) normal structural tRNA^{Phe} (GAA) of *Cedrus deodara*.

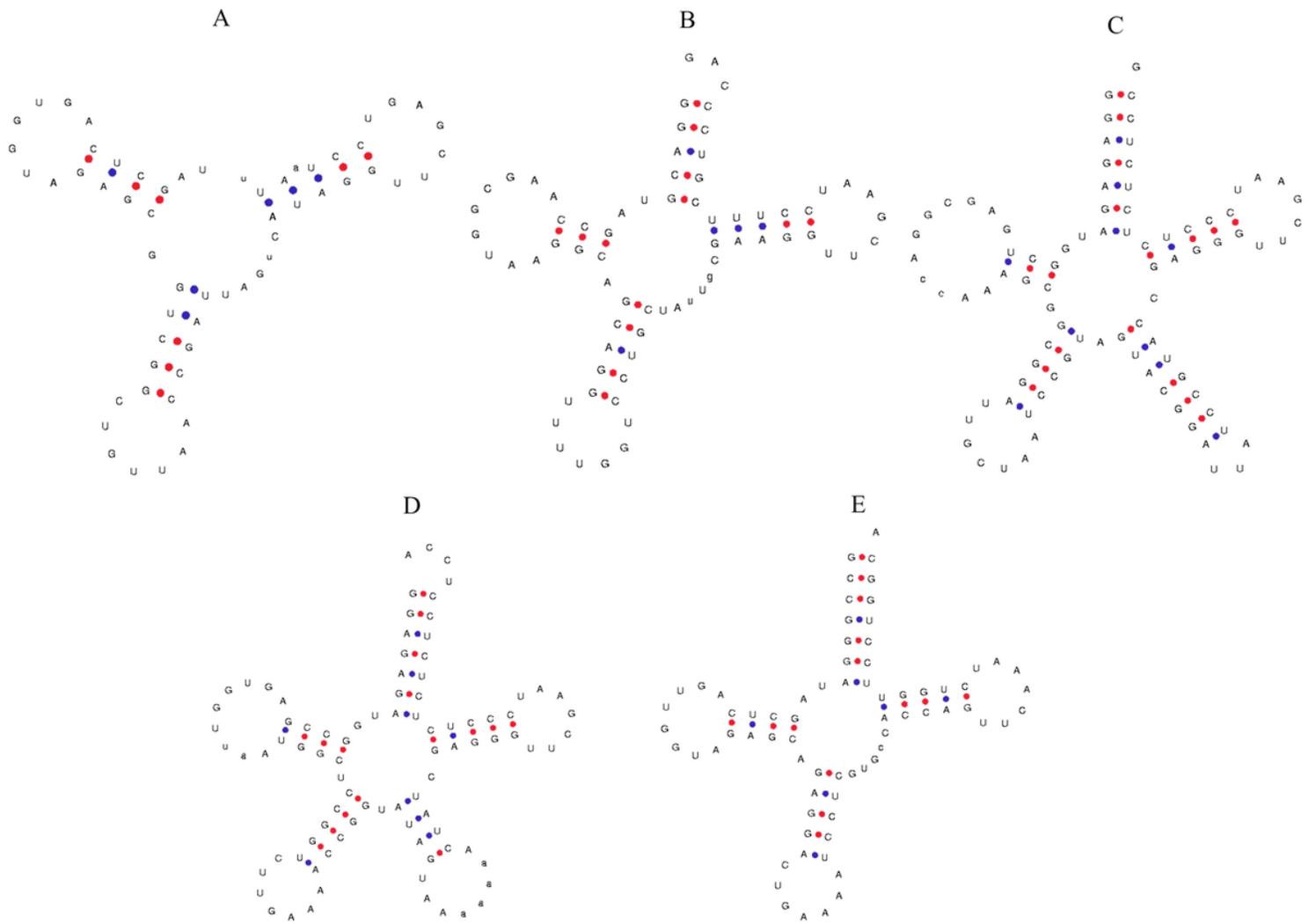


Figure 2

tRNAs with different structural changes. (A) type 1, lack of acceptor arm, tRNA^{Asn} (GTT) of *Nothotsuga longibracteata*, (B) type 2, the 3'-end contains extra nucleotides, tRNA^{Iln} (TTG) of *Sciadopitys verticillata*, (C) type 3, the variable region contains loops or arms, tRNA^{Ser} (GCT) of *Abies koreana*, (D) type 4, the 3'-end contains extra nucleotides while the variable region contains a loop and an arm, tRNA^{Ser} (TGA) of *Tsuga chinensis*, (E) normal structural tRNA^{Phe} (GAA) of *Cedrus deodara*.

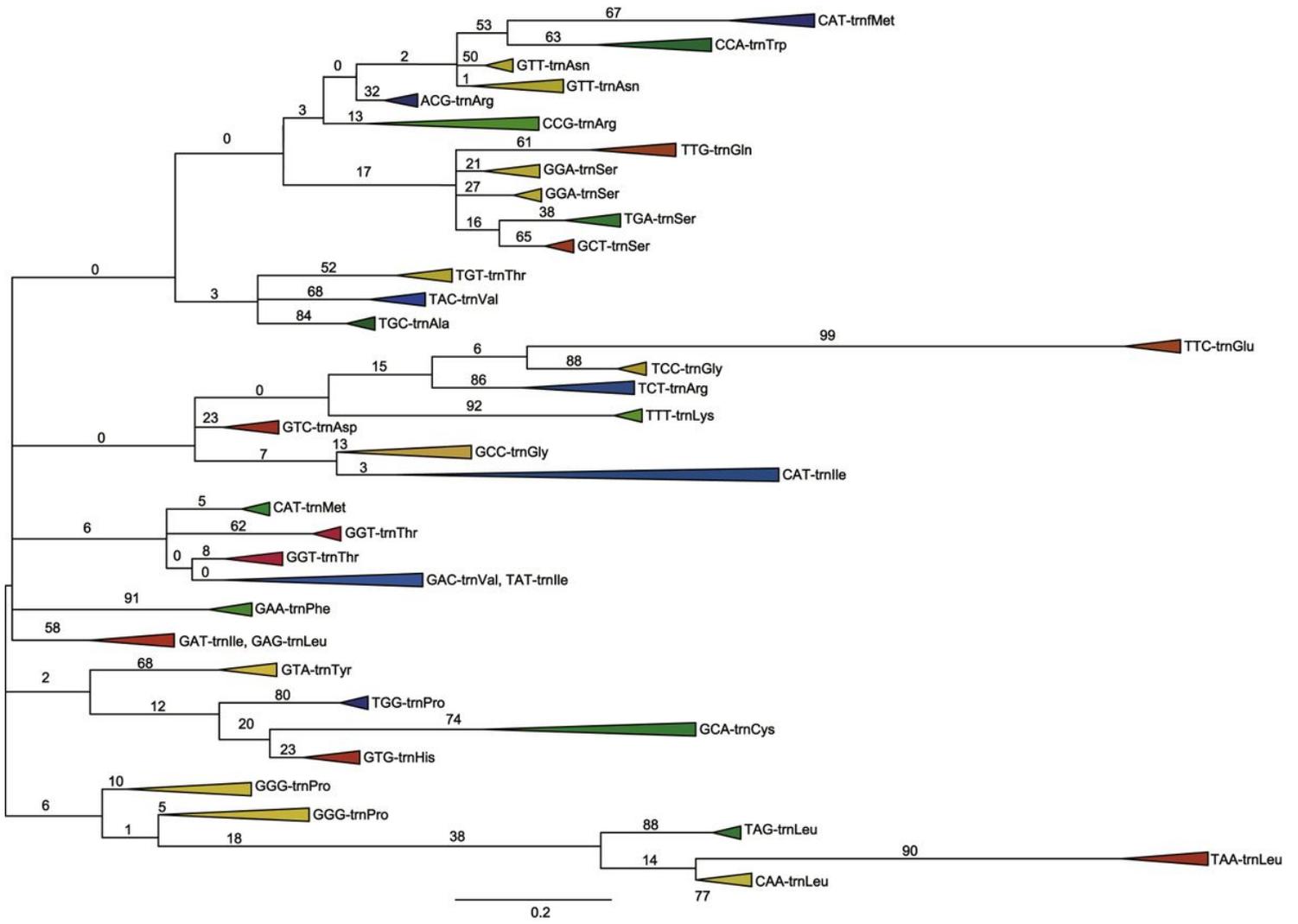


Figure 3

Phylogenetic relationship among all the chloroplast tRNA genes of gymnosperms.

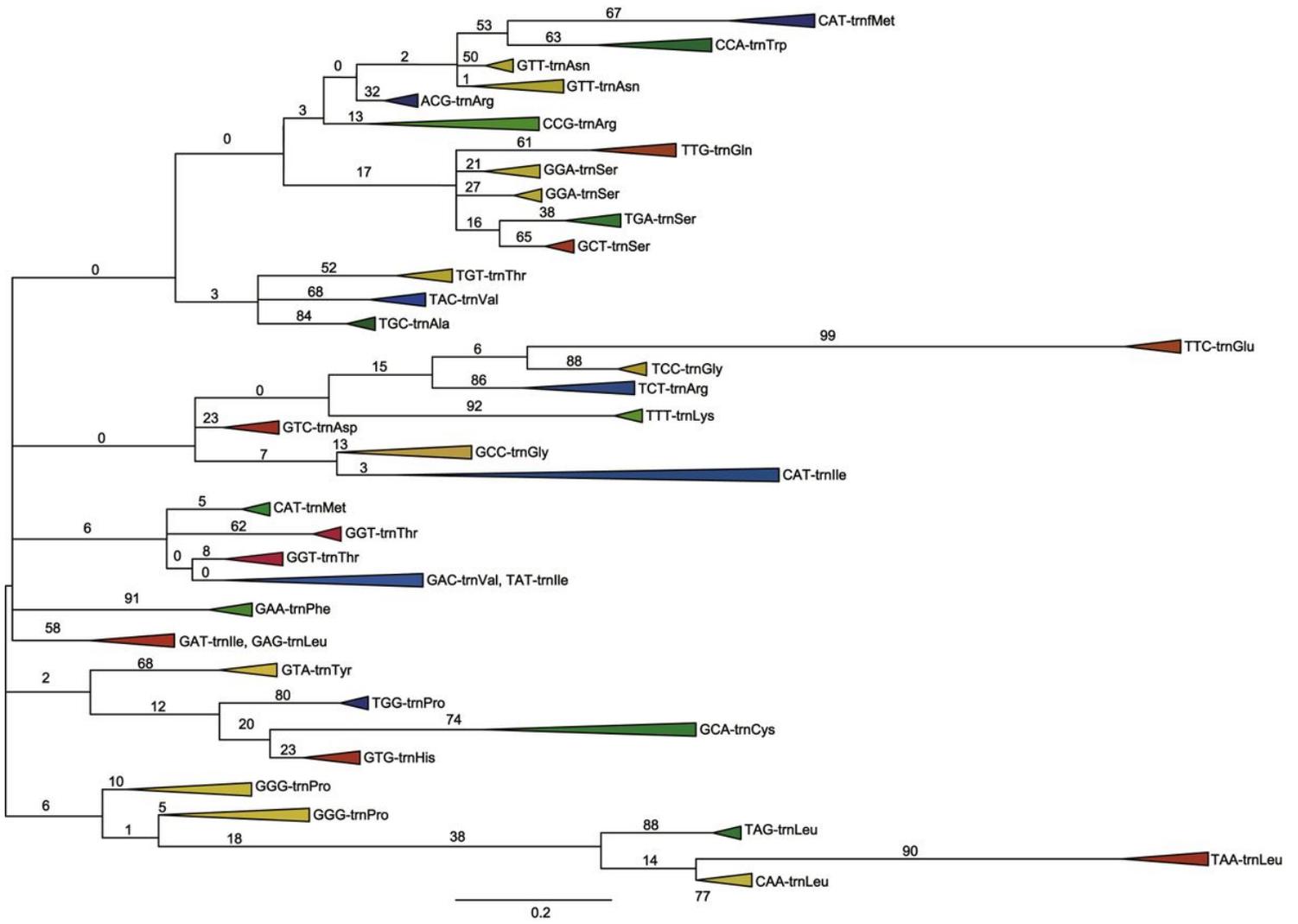


Figure 3

Phylogenetic relationship among all the chloroplast tRNA genes of gymnosperms.

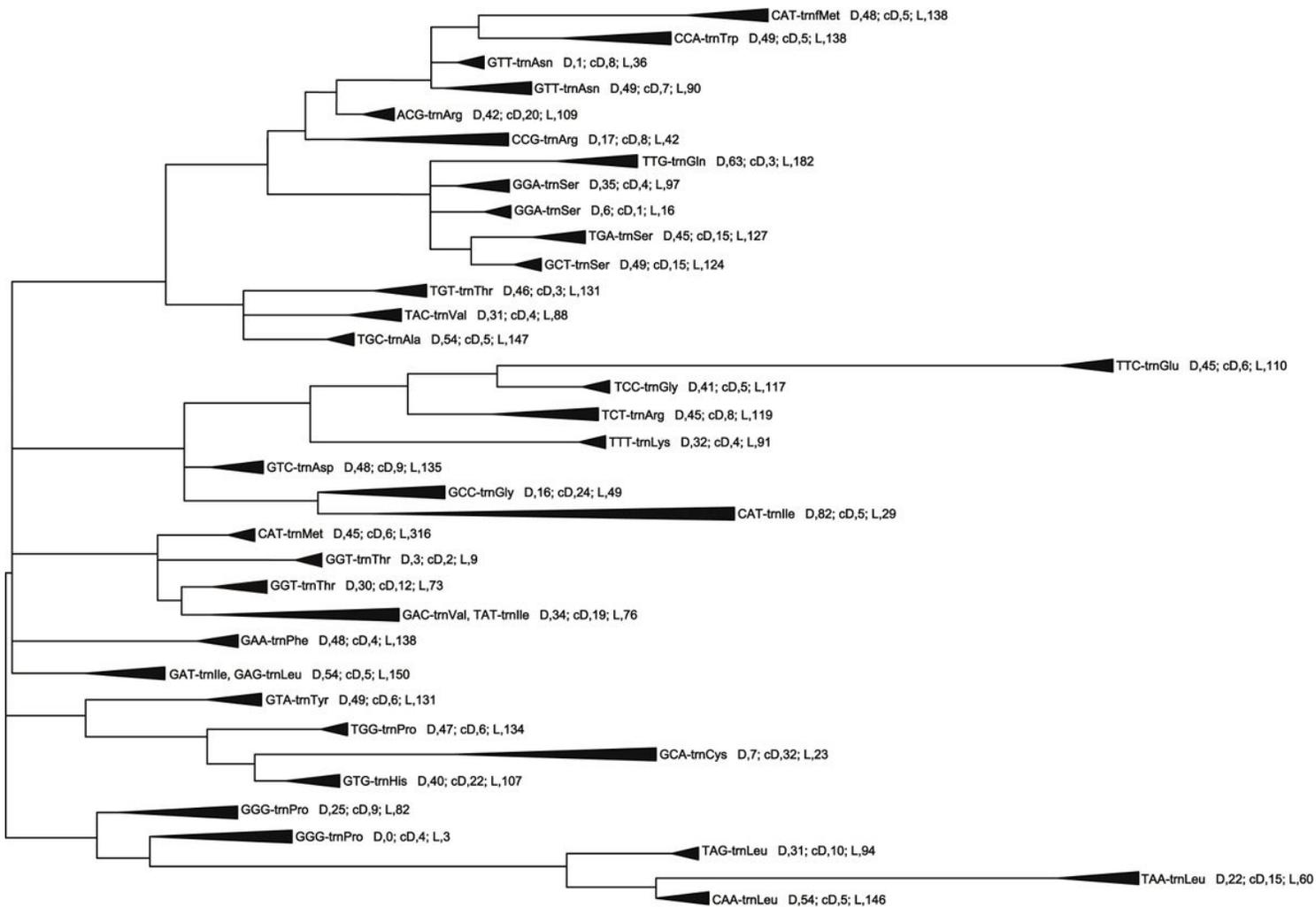


Figure 4

Duplication and loss events of gymnosperm chloroplast tRNAs. Analysis shows that tRNAs underwent mainly gene loss during evolution. D, duplication; cD, conditional duplication; L, loss.

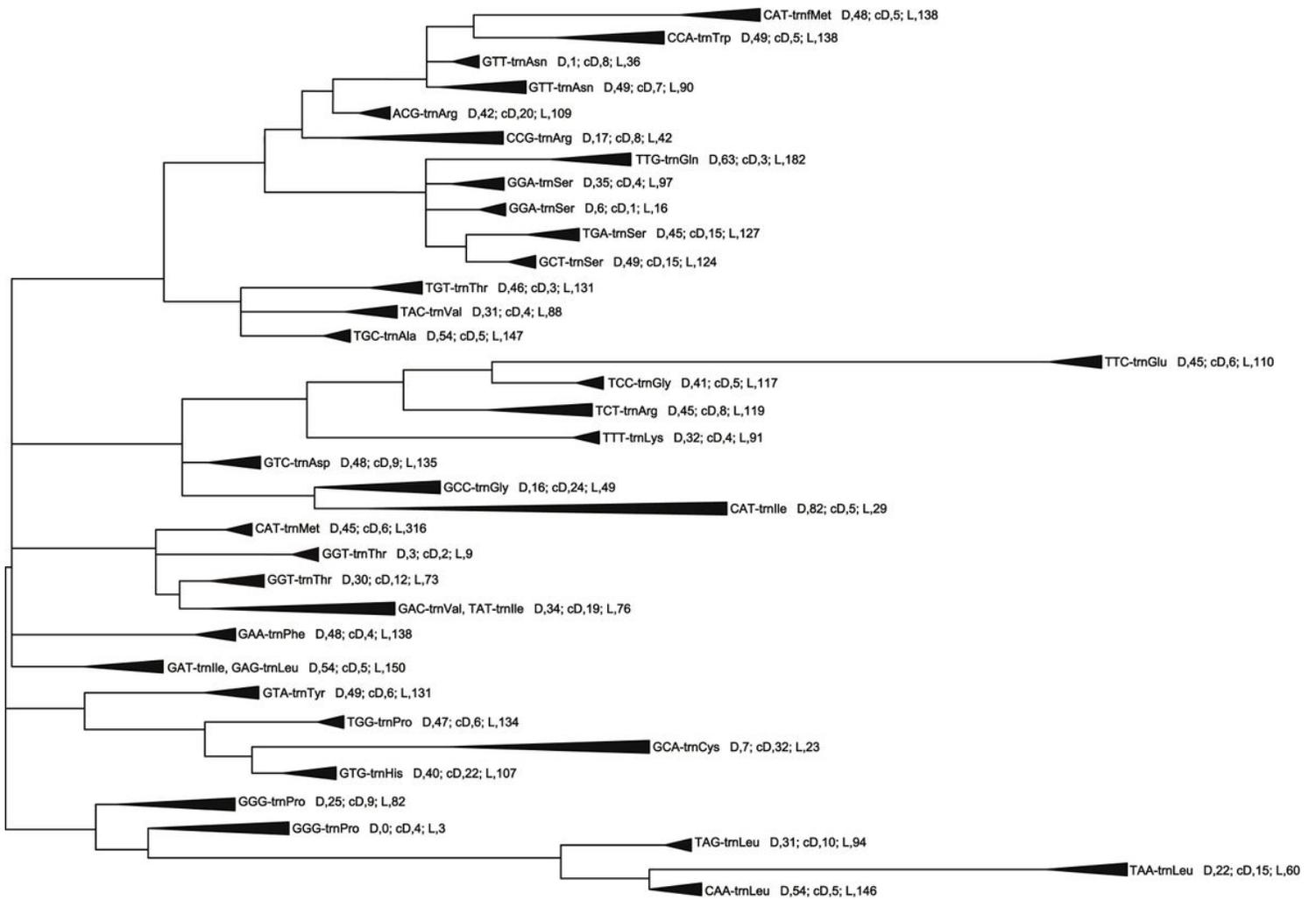


Figure 4

Duplication and loss events of gymnosperm chloroplast tRNAs. Analysis shows that tRNAs underwent mainly gene loss during evolution. D, duplication; cD, conditional duplication; L, loss.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Tables.docx](#)
- [Tables.docx](#)
- [FigureS1.tif](#)
- [FigureS1.tif](#)
- [FigureS2.png](#)
- [FigureS2.png](#)

- TableS1.docx
- TableS1.docx