



# Evolutionary analysis of chloroplast tRNA of Gymnosperm revealed the novel structural variation and evolutionary aspect

Ting-Ting Zhang<sup>1,\*</sup>, Yi-Kun Hou<sup>1,\*</sup>, Ting Yang<sup>1</sup>, Shu-Ya Zhang<sup>1</sup>, Ming Yue<sup>1</sup>, Jianni Liu<sup>2</sup> and Zhonghu Li<sup>1</sup>

<sup>1</sup> Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Xi'an, China

<sup>2</sup> Early Life Institute, State Key Laboratory of Continental Dynamics, Department of Geology, Northwest University, Xi'an, China

\* These authors contributed equally to this work.

## ABSTRACT

Gymnosperms such as ginkgo, conifers, cycads, and gnetophytes are vital components of land ecosystems, and they have significant economic and ecologic value, as well as important roles as forest vegetation. In this study, we investigated the structural variation and evolution of chloroplast transfer RNAs (tRNAs) in gymnosperms. Chloroplasts are important organelles in photosynthetic plants. tRNAs are key participants in translation where they act as adapter molecules between the information level of nucleic acids and functional level of proteins. The basic structures of gymnosperm chloroplast tRNAs were found to have family-specific conserved sequences. The tRNA $\Psi$ -loop was observed to contain a conforming sequence, i.e., U-U-C-N-A-N<sub>2</sub>. In gymnosperms, tRNA<sup>Ile</sup> was found to encode a "CAU" anticodon, which is usually encoded by tRNA<sup>Met</sup>. Phylogenetic analysis suggested that plastid tRNAs have a common polyphyletic evolutionary pattern, i.e., rooted in abundant common ancestors. Analyses of duplication and loss events in chloroplast tRNAs showed that gymnosperm tRNAs have experienced little more gene loss than gene duplication. Transition and transversion analysis showed that the tRNAs are iso-acceptor specific and they have experienced unequal evolutionary rates. These results provide new insights into the structural variation and evolution of gymnosperm chloroplast tRNAs, which may improve our comprehensive understanding of the biological characteristics of the tRNA family.

Submitted 27 November 2019  
Accepted 15 October 2020  
Published 25 November 2020

Corresponding authors  
Jianni Liu, [eliljn@nwu.edu.cn](mailto:eliljn@nwu.edu.cn),  
[liujianni@126.com](mailto:liujianni@126.com)  
Zhonghu Li, [lizhonghu@nwu.edu.cn](mailto:lizhonghu@nwu.edu.cn)

Academic editor  
Genlou Sun

Additional Information and  
Declarations can be found on  
page 21

DOI 10.7717/peerj.10312

© Copyright  
2020 Zhang et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Evolutionary Studies, Genomics, Plant Science

**Keywords** tRNA, Chloroplast, Anti-codon, Evolution, Transition and transversion, Phylogeny

## INTRODUCTION

Gymnosperms originated in the Paleozoic Devonian Period (about 385 million years ago), and they are key groups in terms of the transformation from spore reproduction to seed reproduction in higher plants (*Gerrienne et al., 2004; Crisp & Cook, 2011*). According to the latest phylogenetic classification, gymnosperm species are divided into eight orders,

12 families, 84 genera, and more than 1,000 species (Wang & Ran, 2014). Gymnosperms include ginkgo, cycads, conifers, and gnetophytes, which are grown in forests as important timber species and they provide raw materials for human usage, such as fiber, resin, and tannin (Christenhusz et al., 2010). In addition, gymnosperms include some important threatened plants, where 40% are at high risk of extinction (Forest et al., 2018). Recent phylogenetic and evolutionary studies of gymnosperms have demonstrated the rapid evolution of mitochondrial (mt) genes and provided further evidence of sister relationship between conifers and Gnetales (Ran, Gao & Wang, 2010). The high levels of genetic diversity and population differentiation among the *Pinus* species in gymnosperms have been studied based on plastid DNA markers (Liu et al., 2014). Other studies have indicated patterns related to the physiological ecology, phylogenetic relationships, and population genetic structure of gymnosperm species (Yu et al., 2014; Li et al., 2015; Dong et al., 2016). However, these studies mainly considered the phylogeny and evolution at the whole populations level. Thus, the detailed evolutionary characteristics of gymnosperms still need to be elucidated.

Chloroplasts are the site of photosynthesis and of various essential metabolic pathways, e.g., fatty acid and amino acid biosynthesis and the assimilation of nitrogen, sulfur, and selenium (Hooper, 2006; Des Marais, 2000 Knorr & Heimann, 2001; Pilon-Smits et al., 2002; Guo et al., 2007; Kretschmer, Croll & Kronstad, 2017). It is generally recognized that chloroplasts are derived from proto-eukaryotic symbiotic cyanobacteria that internalized in eukaryotic cells (Hiroki & Daisuke, 2018) and evolved into central organelles. Chloroplasts have their own genome encoding about 100 proteins and they are maternally inherited organelles in most angiosperm plants (Abdallah, Salamini & Leister, 2000; Heuertz et al., 2004; Civan et al., 2014). Among gymnosperms, paternal plastid inheritance is the typical characteristic of conifers (Fauré et al., 1994; Kaundun & Matsumoto, 2011). Studies have shown that the chloroplast genome is quite conserved with an average evolutionary rate of  $0.2\text{--}1.0 \times 10^{-9}$  per site per year, which is only one-fifth of that for the nuclear genome (Drouin, Daoud & Xia, 2008; Duchene & Bromham, 2013). The chloroplast genome is a covalently closed circular structure with four parts comprising the large single copy region (LSC), small single copy region (SSC), inverted repeat region A (IRA), and inverted repeat region B (IRB). The two IRs have the same sequences but in the opposite direction (Wang et al., 2008; Logacheva et al., 2009; Hereward et al., 2018). Due to the independent evolution of the chloroplast genome, it is possible to construct a molecular phylogenetic tree using the chloroplast genome and without requiring any other data. Data analysis based on the conserved evolution of plastids is highly valuable for phylogenetic studies (Kim & Suh, 2013) because it can provide reliable and useful phylogenetic information. The relative completeness and independence of the chloroplast genome means that it can provide valuable material for research purposes.

Transfer RNAs (tRNAs) undergo numerous post-transcriptional nucleotide modifications and they exhibit abundant chemical diversity where the bases experience methylation, formylation, and other modifications (Suzuki & Suzuki, 2014). Chemical nucleotide modifications are frequent in tRNAs and they are important for the structure, stability, correct folding, aminoacylation, and decoding. For example, a previous analysis

of the chemically synthesized  $^{13}\text{C}$ -modified anticodon loop of human mt-tRNA<sup>Met</sup> showed that  $^{13}\text{C}$  contributes to the anticodon domain structure of the mt-tRNA (Lusic et al., 2008). tRNAs comprise sequences of less than 100 polynucleotides that fold into a clover-type secondary structure and then into an L-shaped tertiary structure (Wilusz, 2015). The secondary structure of tRNAs comprises different arms as well as loops, i.e., the D-arm, acceptor arm, anticodon arm, pseudouridine arm ( $\Psi$ -arm), D-loop, variable arm, anticodon loop, and pseudouridine loop ( $\Psi$ -loop) (Giegé, Puglisi & Florentz, 1993; Mizutani & Goto, 2000). This unique structure allows tRNA to act as important bridges between the information level of nucleic acids and functional level of proteins. The vital components of tRNAs comprise an anti-codon region that discerns the messenger RNA carried by the specific codons, a 3'-CCA tail for attaching to the cognate amino acid, the  $\Psi$ -arm, and a  $\Psi$ -loop that has a relationship with the ribosome machinery (Kirchner & Ignatova, 2014). Asymmetric combinations and the divided segments in tRNA genes allow us to understand the diversity of tRNA molecules. tRNA species fulfill various functions in cellular homeostasis, regulation of gene expression and epigenetics, biogenesis, and even biological disease (Ribas & Dedon, 2014; Kanai, 2015; Schimmel, 2017). The evolutionary relationships determined between cyanobacteria and monocots show that tRNAs evolved polyphyletically and they originated from multiple common ancestors with a high rate of gene loss (Mohanta et al., 2017; Mohanta et al., 2019). Nevertheless, the basic details of the tRNAs in plant chloroplasts still need to be elucidated and the diverse evolutionary features of gymnosperm tRNAs are still unclear.

In this study, we assessed all of the chloroplast genomes in 12 families of gymnosperms from eight orders. The main aims of this study were as follows: (1) to determine the diversification of nucleotides in the secondary structure of gymnosperm tRNAs; (2) to identify the detailed genomic features of chloroplast tRNAs; (3) to assess the evolutionary relationships among different chloroplast tRNAs; and (4) to evaluate the duplication or loss events that occurred in all of the tRNAs considered. Our findings provide important insights into the biological characteristics and evolutionary variation of the tRNA family.

## MATERIALS & METHODS

### Annotation and identification of chloroplast tRNA sequences in gymnosperms

We downloaded complete chloroplast genomes for 12 representative gymnosperms in eight orders from the National Center for Biotechnology Information database (NCBI, <https://www.ncbi.nlm.nih.gov/>). The gymnosperm species investigated were: *Cycas debaoensis* Y. C. Zhong & C. J. Chen (KM459003), *Dioon spinulosum* Dyer ex Eichler (NC\_027512), *Ginkgo biloba* L. (NC\_016986), *Cedrus deodara* (Roxb.) G. Don (NC\_014575), *Wollemia nobilis* W. G. Jones, K. D. Hill & J. M. Allen (NC\_027235), *Retrophyllum piresii* Silba C. N. (KJ017081), *Sciadopitys verticillata* (Thunb.) Sieb. et Zucc. (NC\_029734), *Cunninghamia lanceolata* (Lamb.) Hook. (NC\_021437), *Taxus mairei* (Lemee et Levl.) Cheng et L. K. Fu (KJ123824), *Welwitschia mirabilis* Hook.f. (EU342371), *Gnetum gnemon* L. (KR476377), and *Ephedra equisetina* Bge. (NC\_011954).

The gymnosperm tRNA genomes were annotated using GeSeq-Annotation of Organellar Genomes tool ([Tillich et al., 2017](#)) where the parameters were set as: circular sequence(s), chloroplast of sequence source, generate multi FASTA; BLAST protein search identity 25% for annotating plastid IR, 85% identity for BLAST rRNA, tRNA and DNA search, Embryophyta chloroplast (CDS+rRNA), third party tRNA annotator ARAGORN v1.2.38, ARWEN v1.2.3, tRNAScan-SE v2.0, and without Refseq choice.

### Structural analysis of chloroplast tRNAs

ARAGORN ([Laslett & Canback, 2004](#)) and tRNAScan-SE software ([Lowe & Eddy, 1997](#)) were employed to analyze the sequences and the secondary structure of tRNAs in the chloroplast genomes of the involved gymnosperm plants. The default parameters were set in ARAGORN software. The parameters for tRNAScan-SE were set as: sequence source, bacterial; search mode, default; query sequences, formatted (FASTA); and genetic code for tRNA isotype prediction, universal.

### Phylogenetic tree construction

A phylogenetic tree was constructed for all of the tRNAs using MEGA7.0 software ([Kumar et al., 2008](#); [Kumar, Stecher & Tamura, 2016](#)). To study the evolutionary details of chloroplast tRNAs in gymnosperm species, an alignment file for tRNAs was achieved by CLUSTAL Omega software before the phylogenetic tree was constructed. MEGA7 software was used to transform the alignment file into MEGA format. The phylogenetic tree was constructed with the following parameters: phylogeny reconstruction of analysis, maximum likelihood model, bootstrap method in phylogeny test, 1,000 bootstrap replicates, nucleotides type, gamma distributed with invariant sites (G+I) model, five discrete gamma categories, partial deletion for gaps/missing data treatment, 95% site coverage cut-off, and very strong for branch swap filter.

### Transition/transversion analysis

The sequences of the tRNA isotypes were aligned to determine the transition and transversion rates for chloroplast tRNAs in gymnosperm plants. The files covering all 20 types of tRNAs were transformed into the MEGA file format and analyzed separately using MEGA7.0 software ([Kumar, Tamura & Nei, 1994](#)). The transition and transversion rates were analyzed for tRNAs with the following parameters: substitution pattern estimation (ML) analysis, automatic (neighbor-joining tree), maximum likelihood statistical method, nucleotide substitution type, Kimura two-parameter model, gamma distributed (G) site rates, five discrete gamma categories, partial deletion of gaps/missing data treatment, 95% of site coverage cut-off, and very strong branch swap filter.

### Loss and duplication events analysis for tRNA genes

In order to investigate the duplication or loss events in tRNA genes, the NCBI taxonomy browser was utilized to construct the whole species tree for the 12 gymnosperm species considered. The phylogenetic tree conducted in the evolutionary study was employed as gene tree. The gene tree for the tRNAs and species tree for the gymnosperm species were submitted to Notung 2.9 software ([Chen, Durand & Farach-Colton, 2000](#)), and then

**Table 1** A view of the gymnosperms in analysis. Statistics of the 12 gymnosperms in the study.

Order	Family	Subfamily	Genus	Species	NCBI Locus
Cycadales	Cycadaceae		<i>Cycas</i>	<i>debaensis</i>	KM459003
	Zamiaceae	Diooideae	<i>Dioon</i>	<i>spinulosum</i>	NC_027512
Ginkgoales	Ginkgoaceae		<i>Ginkgo</i>	<i>biloba</i>	NC_016986
Pinales	Pinaceae	Abietae	<i>Cedrus</i>	<i>deodara</i>	NC_014575
Araucariales	Araucariaceae		<i>Wollemia</i>	<i>nobilis</i>	NC_027235
	Podocarpaceae		<i>Retrophillum</i>	<i>piresii</i>	KJ017081
Cupressales	Sciadopityaceae		<i>Sciadopitys</i>	<i>verticillata</i>	NC_029734
	Cupressaceae	Cunninghamia	<i>Cunninghamia</i>	<i>lanceolata</i>	NC_021437
	Taxaceae		<i>Taxus</i>	<i>mairei</i>	KJ123824
Welwitschiales	Welwitschiaceae		<i>Welwitschia</i>	<i>mirabilis</i>	EU342371
Gnetales	Gnetaceae		<i>Gnetum</i>	<i>gnemon</i>	KR476377
Ephedrales	Ephedraceae		<i>Ephedra</i>	<i>equisetina</i>	NC_011954

reconciled to discover duplicated and lost tRNA genes in the chloroplast genomes of gymnosperms.

## RESULTS

### Genomic features of gymnosperm chloroplast tRNAs

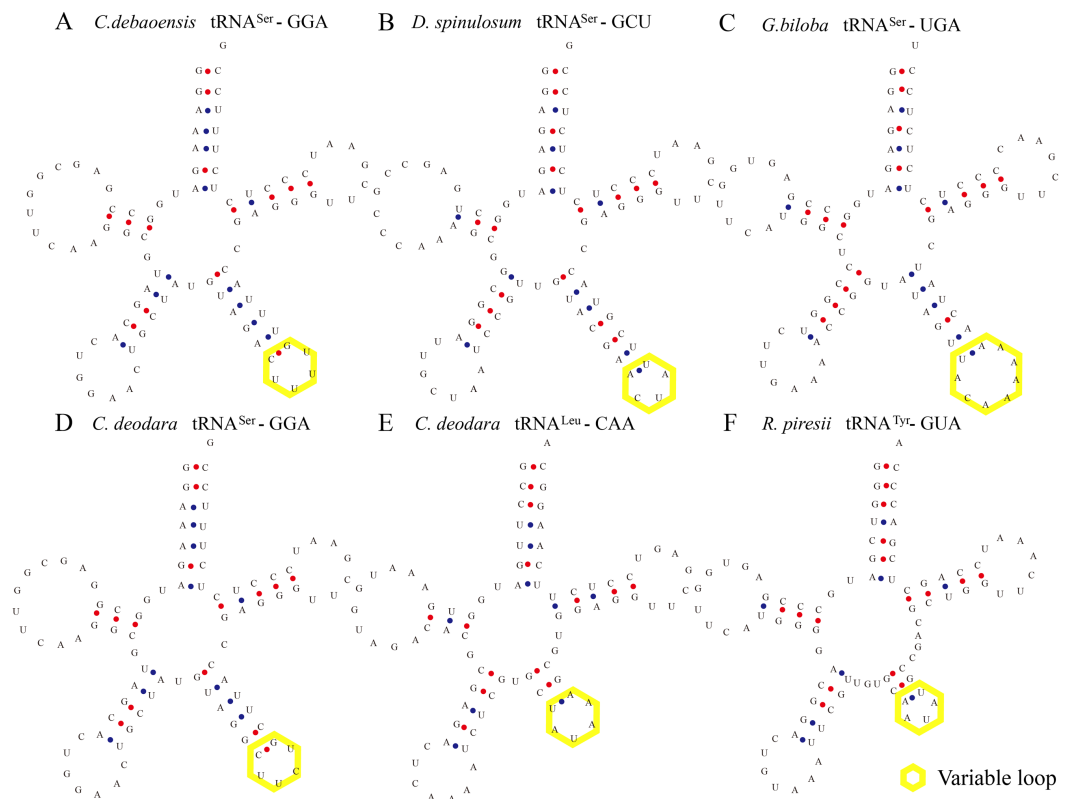
Sequences were analyzed to identify the genomic tRNAs in the chloroplast genomes of 12 gymnosperm species comprising *C. debaensis*, *D. spinulosum*, *G. biloba*, *C. deodara*, *W. nobilis*, *R. piresii*, *S. verticillata*, *C. lanceolata*, *T. mairei*, *W. mirabilis*, *G. gnemon*, and *E. equisetina*, which were obtained from the NCBI database (Table 1). The results showed that the length of the chloroplast tRNAs vary from the smallest with 64 nucleotides (nt) (tRNA<sup>Met</sup>-CAU in *T. mairei*) to the largest with 96 nt (tRNA<sup>Tyr</sup>-AUA in *W. nobilis*, *C. deodara*, and *G. biloba*) (Data S1). We found that the chloroplast genomes of gymnosperm plants encode 28 to 33 tRNAs (Table 2), where *D. spinulosum*, *C. deodara*, and *S. verticillata* encode 31 anticodons, *W. nobilis*, *R. piresii*, *C. lanceolata*, and *G. gnemon* encode 32 tRNA isotypes, *G. biloba*, and *W. mirabilis* encode 33 tRNAs. Other species comprising *T. mairei*, *E. equisetina* and *C. debaensis* encode 28, 28, 30 tRNA isotypes, respectively (Table 2). tRNA<sup>Ala</sup> was not found in *R. piresii* and *T. mairei*, and tRNA<sup>Val</sup> was not detected in *T. mairei* (Fig. S3). We also observed that all of the species do not encode selenocysteine and its suppressor tRNA (Table 2). Overall, tRNA<sup>Ser</sup> (in *W. nobilis*) and tRNA<sup>Arg</sup> (in *W. mirabilis*) are the most abundant (four types) followed by tRNA<sup>Leu</sup> (three types) (Table 2).

### Variations in structures of chloroplast tRNAs

Some tRNAs with a loop structure in the variable region were found to be encoded in the gymnosperm chloroplast genomes (Figs. 1 and 2). A novel tRNA lacking the D-arm was found in tRNA<sup>Gly</sup> in *W. nobilis* (Fig. 3). As shown in Figs. 1 and 2, tRNA<sup>Leu</sup>, tRNA<sup>Ser</sup>, and tRNA<sup>Tyr</sup> contain expanded variable stem/loops. In these tRNAs (except for tRNA<sup>Ser</sup>-GCU of *D. spinulosum*), the anticodon loop of tRNA<sup>Ser</sup> contains the conserved consensus

**Table 2** Distribution of tRNA isotypes in chloroplast genome of gymnosperms.

tRNA isotypes	Number of tRNAs											
	<i>C. debaensis</i>	<i>D. spinulosum</i>	<i>G. biloba</i>	<i>C. deodara</i>	<i>W. nobilis</i>	<i>R. piresii</i>	<i>S. verticillata</i>	<i>C. lanceolata</i>	<i>T. mairei</i>	<i>W. mirabilis</i>	<i>G. gnemon</i>	<i>E. equisetina</i>
Ala	1	1	1	1	1	0	1	1	0	1	1	1
Gly	1	1	2	2	2	2	1	1	1	2	2	1
Pro	2	2	2	2	2	2	1	2	2	2	2	1
Thr	1	2	2	2	2	2	2	2	2	2	2	1
Val	2	2	1	1	1	2	2	2	0	2	2	1
Ser	3	3	3	3	4	3	3	3	2	3	3	3
Arg	3	3	3	3	3	3	3	2	2	4	3	2
Leu	3	3	3	3	3	3	3	3	3	3	3	3
Phe	1	1	1	1	1	1	1	1	1	1	1	1
Asn	1	1	1	1	1	1	1	1	1	1	1	1
Lys	1	1	1	1	1	1	1	2	1	1	1	1
Asp	1	1	1	1	1	2	1	1	1	1	1	1
Glu	2	2	1	2	2	2	2	2	1	2	2	2
His	1	1	2	1	1	1	1	1	1	1	1	1
Gln	1	1	1	1	1	1	2	2	1	1	1	1
Ile	1	1	1	1	1	1	1	1	4	1	1	1
Met/fMet	2	2	2	2	2	2	2	2	2	2	2	2
Tyr	1	1	2	1	1	1	1	1	1	1	1	1
Cys	1	1	2	1	1	1	1	1	1	1	1	1
Trp	1	1	1	1	1	1	1	1	1	1	1	2
Selenocysteine	0	0	0	0	0	0	0	0	0	0	0	0
Suppressor	0	0	0	0	0	0	0	0	0	0	0	0
Total	30	31	33	31	32	32	31	32	28	33	32	28



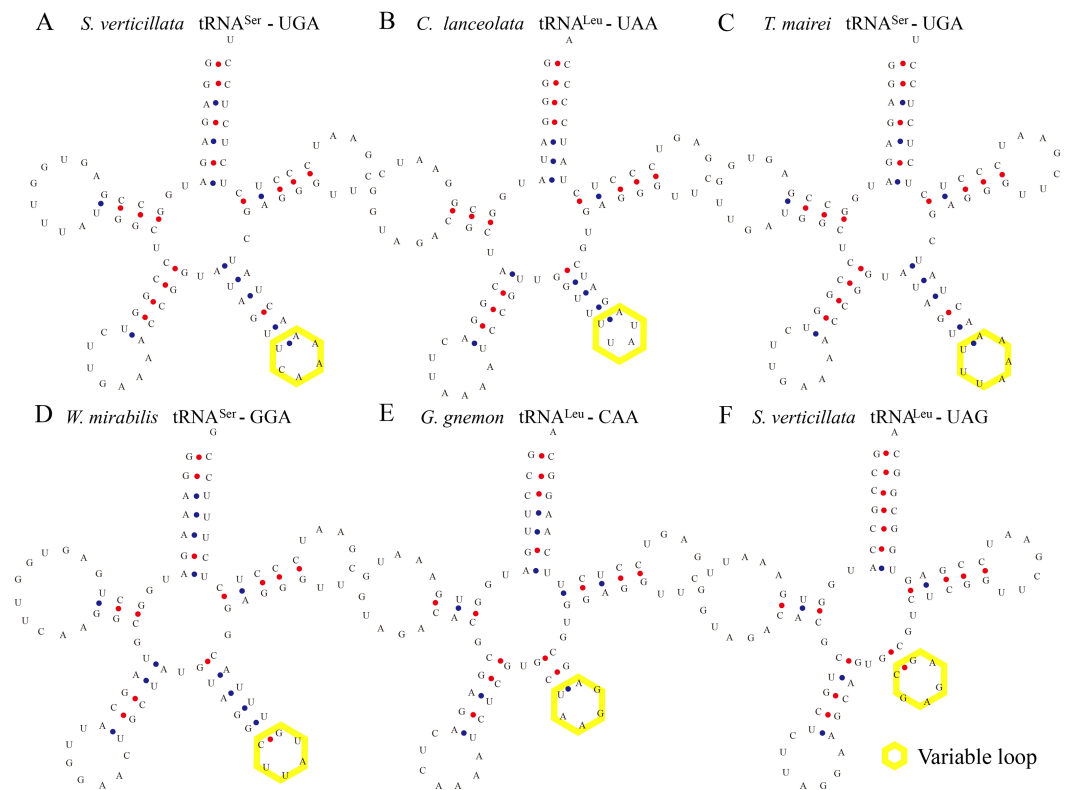
**Figure 1** Certain tRNAs in *C. debaensis*, *D. spinulosum*, *G. biloba*, *C. deodara*, and *R. piresii* contain expanded variable stem and loops. tRNA<sup>Ser</sup>, tRNA<sup>Leu</sup>, and tRNA<sup>Tyr</sup> from *C. debaensis* (A, tRNA<sup>Ser</sup>-GGA), *D. spinulosum* (B, tRNA<sup>Ser</sup>-GCU), *G. biloba* (C, tRNA<sup>Ser</sup>-UGA), *C. deodara* (D, tRNA<sup>Ser</sup>-GGA; E, tRNA<sup>Leu</sup>-CAA), *R. piresii* (F, tRNA<sup>Tyr</sup>-GUA) were observed to contain an expanded variable stem and variable loop (indicated by yellow box). The anti-codon loop of tRNA<sup>Ser</sup> (except for tRNA<sup>Ser</sup>-GCU of *D. spinulosum*) was made up of seven nucleotides with the conservative N-U-N-G-A-A-N consensus sequence.

Full-size DOI: 10.7717/peerj.10312/fig-1

sequence N-U-N-G-A-A-N, and tRNAs<sup>Leu</sup> have the consensus sequence C-U-N-A-N<sub>2</sub>-A. The variable loop region is predicted to fold into stem-loop structures with apical loops of 3 to 7 nt in tRNA<sup>Ser</sup> and several tRNA<sup>Leu</sup> variants. The stems contain up to 7 bp (Figs. 1 and 2). The expanded variable loop structures may play important functions during the protein translation process in chloroplasts.

### Chloroplast genomes contain 25 to 30 anticodon-specific tRNAs

The genomes of the species analyzed were found to code for at least two copies of tRNA<sup>Met</sup>-CAU/tRNA<sup>fMet</sup>-CAU. Each of the gymnosperm chloroplast genomes encodes 25 to 30 anticodon-specific tRNAs (Tables 2 and 3), where *E. equisetina* encodes 25 anticodons, *T. mairei* encodes 26 anticodons, *C. debaensis*, *S. verticillata*, and *C. lanceolata* encode 28 anticodons, and *D. spinulosum*, *C. deodara*, *W. mirabilis*, and *R. piresii* encode 29 anticodons. Other species comprising *W. nobilis*, *G. gnemon* and *G. biloba* encodes 30 anticodons (Table 3).



**Figure 2** Certain tRNAs in *S. verticillata*, *C. lanceolata*, *T. mairei*, *W. mirabilis*, and *G. gnemon* contain expanded variable stem and loops. tRNA<sub>Ser</sub>, tRNA<sub>Leu</sub> from *S. verticillata* (A, tRNA<sub>Ser</sub>-UGA; F, tRNA<sub>Leu</sub>-UAG), *C. lanceolata* (B, tRNA<sub>Leu</sub>-UAA), *T. mairei* (C, tRNA<sub>Ser</sub>-UGA), *W. mirabilis* (D, tRNA<sub>Ser</sub>-GGA), *G. gnemon* (E, tRNA<sub>Leu</sub>-CAA) were observed to contain a variable stem and variable loop (indicated by yellow box). The anti-codon loop of tRNA<sub>Ser</sub> was made up of seven nucleotides with the conservative N-U-N-G-A-A-N consensus sequence, and the consensus sequence was C-U-N-A-N2-A for tRNA<sub>Leu</sub>.

Full-size [DOI: 10.7717/peerj.10312/fig-2](https://doi.org/10.7717/peerj.10312/fig-2)

tRNA<sup>Arg</sup>-CCG was present in the genomes of nine gymnosperm species but absent from *C. lanceolata*, *T. mairei*, and *E. equisetina*, while tRNA<sup>Gly</sup>-UCC was lacking from *C. debaoensis*, *S. verticillata*, *D. spinulosum*, *C. lanceolata*, *T. mairei*, and *E. equisetina* (Table 3). The most abundant anticodons found in the chloroplast genomes were tRNA<sup>Gly</sup>-GCC, tRNA<sup>Pro</sup>-UGG, tRNA<sup>Ser</sup>-UGA, tRNA<sup>Ser</sup>-GCU, tRNA<sup>Arg</sup>-ACG, tRNA<sup>Arg</sup>-UCU, tRNA<sup>Leu</sup>-UAG, tRNA<sup>Leu</sup>-CAA, tRNA<sup>Phe</sup>-GAA, tRNA<sup>Asn</sup>-GUU, tRNA<sup>Lys</sup>-UUU, tRNA<sup>Asp</sup>-GUC, tRNA<sup>Glu</sup>-UUC, tRNA<sup>His</sup>-GUG, tRNA<sup>Gln</sup>-UUG, tRNA<sup>Ile</sup>-CAU, tRNA<sup>Met</sup>-CAU, tRNA<sup>Tyr</sup>-GUA, tRNA<sup>Cys</sup>-GCA, and tRNA<sup>Trp</sup>-CCA (Table 3). Two tRNA<sup>Trp</sup> iso-acceptors are present in *E. equisetina* chloroplasts, compared with a single one in the other gymnosperm species analyzed in this study.

### Conserved gymnosperm chloroplast tRNAs

The clover leaf-like secondary structure of a tRNA is shown in Fig. 4. In the study, we found that most tRNAs contain a “G” as the first nucleotide in the D-arm, except for tRNA<sup>Lys</sup>, tRNA<sup>Met</sup>, tRNA<sup>Pro</sup>, tRNA<sup>Thr</sup>, tRNA<sup>Tyr</sup>, and tRNA<sup>Val</sup>. “A” is present in the first and the last position of the D-loop apart from tRNA<sup>Gly</sup>, tRNA<sup>Ile</sup>, tRNA<sup>Leu</sup>, tRNA<sup>Met</sup>, and



**Table 3** Distribution of anti-codons in the chloroplast genome of gymnosperms. Each of the gymnosperm chloroplast genomes encodes 25 to 30 anticodon-specific tRNAs. *E. equisetina* encodes 25 anticodons, *T. mairei* encodes 26 anticodons, *C. debaoensis*, *S. verticillata*, and *C. lanceolata* encode 28 anticodons, and *D. spinulosum*, *C. deodara*, *W. mirabilis*, and *R. piresii* encode 29 anticodons. Other species comprising *W. nobilis*, *G. gnemon* and *G. biloba* encodes 30 anticodons.

tRNA Isotypes		Isoacceptors						tRNA Isotypes		Isoacceptors													
<i>C. debaoensis</i> (28)												<i>S. verticillata</i> (28)											
Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 1				Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 1											
Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 0				Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 0											
Pro	AGG: 0	GGG: 1	CGG: 0	UGG: 1				Pro	AGG: 0	GGG: 0	CGG: 0	UGG: 1											
Thr	AGU: 0	GGU: 0	CGU: 0	UGU: 1				Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 1											
Val	AAC: 0	GAC: 1	CAC: 0	UAC: 1				Val	AAC: 0	GAC: 1	CAC: 0	UAC: 1											
Ser	AGA: 0	GGA: 1	CGA: 0	UGA: 1	ACU: 0	GCU: 1		Ser	AGA: 0	GGA: 1	CGA: 0	UGA: 1	ACU: 0	GCU: 1									
Arg	ACG: 1	GCG: 0	CCG: 1	UCG: 0	CCU: 0	UCU: 1		Arg	ACG: 1	GCG: 0	CCG: 1	UCG: 0	CCU: 0	UCU: 1									
Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1		Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1									
Phe	AAA: 0	GAA: 1						Phe	AAA: 0	GAA: 1													
Asn	AUU: 0	GUU: 1						Asn	AUU: 0	GUU: 1													
Lys	CUU: 0	UUU: 1						Lys	CUU: 0	UUU: 1													
Asp	AUC: 0	GUC: 1						Asp	AUC: 0	GUC: 1													
Glu	CUC: 0	UUC: 2						Glu	CUC: 0	UUC: 2													
His	AUG: 0	GUG: 1						His	AUG: 0	GUG: 1													
Gln	CUG: 0	UUG: 1						Gln	CUG: 0	UUG: 2													
Ile	AAU: 0	GAU: 0	CAU: 1	UAU: 0				Ile	AAU: 0	GAU: 0	CAU: 1	UAU: 0											
Met	CAU: 2							Met	CAU: 2														
Tyr	AUA: 0	GUA: 1						Tyr	AUA: 0	GUA: 1													
Cys	ACA: 0	GCA: 1						Cys	ACA: 0	GCA: 1													
Trp	CCA: 1							Trp	CCA: 1														
Supressor	CUA: 0	UUA: 0	UCA: 0					Supressor	CUA: 0	UUA: 0	UCA: 0												
Sec	UCA: 0							Sec	UCA: 0														
<i>D. spinulosum</i> (29)												<i>C. lanceolata</i> (28)											
Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 1				Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 1											
Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 0				Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 0											
Pro	AGG: 0	GGG: 1	CGG: 0	UGG: 1				Pro	AGG: 0	GGG: 1	CGG: 0	UGG: 1											
Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 1				Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 1											
Val	AAC: 0	GAC: 1	CAC: 0	UAC: 1				Val	AAC: 0	GAC: 1	CAC: 0	UAC: 1											
Ser	AGA: 0	GGA: 1	CGA: 0	UGA: 1	ACU: 0	GCU: 1		Ser	AGA: 0	GGA: 1	CGA: 0	UGA: 1	ACU: 0	GCU: 1									

(continued on next page)

Table 3 (continued)

tRNA Isotypes		Isoacceptors					tRNA Isotypes		Isoacceptors				
Arg	ACG: 1	GCG: 0	CCG:1	UCG: 0	CCU: 0	UCU: 1	Arg	ACG: 1	GCG: 0	CCG: 0	UCG: 0	CCU: 0	UCU: 1
Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1	Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1
Phe	AAA: 0	GAA: 1					Phe	AAA: 0	GAA: 1				
Asn	AUU: 0	GUU: 1					Asn	AUU: 0	GUU: 1				
Lys	CUU: 0	UUU: 1					Lys	CUU: 1	UUU: 1				
Asp	AUC: 0	GUC: 1					Asp	AUC: 0	GUC: 1				
Glu	CUC: 0	UUC: 2					Glu	CUC: 0	UUC: 2				
His	AUG: 0	GUG: 2					His	AUG: 0	GUG: 1				
Gln	CUG: 0	UUG: 1					Gln	CUG: 0	UUG: 2				
Ile	AAU: 0	GAU: 0	CAU: 1	UAU: 0			Ile	AAU: 0	GAU: 0	CAU: 1	UAU: 0		
Met	CAU: 2						Met	CAU: 2					
Tyr	AUA: 0	GUA: 1					Tyr	AUA: 0	GUA: 1				
Cys	ACA: 0	GCA: 1					Cys	ACA: 0	GCA: 1				
Trp	CCA: 1						Trp	CCA: 1					
Supressor	CUA: 0	UUA: 0	UCA: 0				Supressor	CUA: 0	UUA: 0	UCA: 0			
Sec	UCA: 0						Sec	UCA: 0					
<i>G. biloba</i> (30)							<i>T. mairei</i> (26)						
Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 1			Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 0		
Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 1			Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 0		
Pro	AGG: 0	GGG: 1	CGG: 0	UGG: 1			Pro	AGG: 0	GGG: 1	CGG: 0	UGG: 1		
Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 1			Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 1		
Val	AAC: 0	GAC: 1	CAC: 0	UAC: 0			Val	AAC: 0	GAC: 0	CAC: 0	UAC: 0		
Ser	AGA: 0	GGA: 1	CGA: 0	UGA: 1	ACU: 0	GCU: 1	Ser	AGA: 0	GGA: 0	CGA: 0	UGA: 1	ACU: 0	GCU: 1
Arg	ACG: 1	GCG: 0	CCG:1	UCG: 0	CCU: 0	UCU: 1	Arg	ACG: 1	GCG: 0	CCG: 0	UCG: 0	CCU: 0	UCU: 1
Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 2	UAA: 0	Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1
Phe	AAA: 0	GAA: 1					Phe	AAA: 0	GAA: 1				
Asn	AUU: 0	GUU: 1					Asn	AUU: 0	GUU: 1				
Lys	CUU: 0	UUU: 1					Lys	CUU: 0	UUU: 1				
Asp	AUC: 0	GUC: 1					Asp	AUC: 0	GUC: 1				
Glu	CUC: 0	UUC: 1					Glu	CUC: 0	UUC: 1				
His	AUG: 0	GUG: 2					His	AUG: 0	GUG: 1				
Gln	CUG: 0	UUG: 1					Gln	CUG: 0	UUG: 1				
Ile	AAU: 0	GAU: 0	CAU: 1	UAU: 0			Ile	AAU: 1	GAU: 0	CAU: 2	UAU: 1		
Met	CAU: 2						Met	CAU: 2					
Tyr	AUA: 1	GUA: 1					Tyr	AUA: 0	GUA: 1				
Cys	ACA: 1	GCA: 1					Cys	ACA: 0	GCA: 1				
Trp	CCA: 1						Trp	CCA: 1					

(continued on next page)

Table 3 (continued)

tRNA Isotypes		Isoacceptors					tRNA Isotypes		Isoacceptors				
Suppressor	CUA: 0	UUA: 0	UCA: 0				Suppressor	CUA: 0	UUA: 0	UCA: 0			
Sec	UCA: 0						Sec	UCA: 0					
<i>C. deodara</i> (29)						<i>W. mirabilis</i> (29)							
Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 1		Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 1			
Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 1		Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 1			
Pro	AGG: 0	GGG: 1	CGG: 0	UGG: 1		Pro	AGG: 0	GGG: 1	CGG: 0	UGG: 1			
Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 1		Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 1			
Val	AAC: 0	GAC: 1	CAC: 0	UAC: 0		Val	AAC: 0	GAC: 1	CAC: 0	UAC: 1			
Ser	AGA: 0	GGA: 1	CGA: 0	UGA: 1	ACU: 0	GCU: 1	Ser	AGA: 0	GGA: 1	CGA: 0	UGA: 1	ACU: 0	GCU: 1
Arg	ACG: 1	GCG: 0	CCG: 1	UCG: 0	CCU: 0	UCU: 1	Arg	ACG: 1	GCG: 0	CCG: 2	UCG: 0	CCU: 0	UCU: 1
Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1	Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1
Phe	AAA: 0	GAA: 1					Phe	AAA: 0	GAA: 1				
Asn	AUU: 0	GUU: 1					Asn	AUU: 0	GUU: 1				
Lys	CUU: 0	UUU: 1					Lys	CUU: 0	UUU: 1				
Asp	AUC: 0	GUC: 1					Asp	AUC: 0	GUC: 1				
Glu	CUC: 0	UUC: 2					Glu	CUC: 0	UUC: 2				
His	AUG: 0	GUG: 1					His	AUG: 0	GUG: 1				
Gln	CUG: 0	UUG: 1					Gln	CUG: 0	UUG: 1				
Ile	AAU: 0	GAU: 04	CAU: 1	UAU: 0			Ile	AAU: 0	GAU: 0	CAU: 1	UAU: 0		
Met	CAU: 2					Met	CAU: 2						
Tyr	AUA: 0	GUA: 1					Tyr	AUA: 0	GUA: 1				
Cys	ACA: 0	GCA: 1					Cys	ACA: 0	GCA: 1				
Trp	CCA: 1					Trp	CCA: 1						
Suppressor	CUA: 0	UUA: 0	UCA: 0				Suppressor	CUA: 0	UUA: 0	UCA: 0			
Sec	UCA: 0						Sec	UCA: 0					
<i>W. nobilis</i> (30)						<i>G. gnemon</i> (30)							
Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 1		Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 1			
Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 1		Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 1			
Pro	AGG: 0	GGG: 1	CGG: 0	UGG: 1		Pro	AGG: 0	GGG: 1	CGG: 0	UGG: 1			
Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 1		Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 1			
Val	AAC: 0	GAC: 1	CAC: 0	UAC: 0		Val	AAC: 0	GAC: 1	CAC: 0	UAC: 1			
Ser	AGA: 0	GGA: 1	CGA: 1	UGA: 1	ACU: 0	GCU: 1	Ser	AGA: 0	GGA: 1	CGA: 0	UGA: 1	ACU: 0	GCU: 1
Arg	ACG: 1	GCG: 0	CCG: 1	UCG: 0	CCU: 0	UCU: 1	Arg	ACG: 1	GCG: 0	CCG: 1	UCG: 0	CCU: 0	UCU: 1
Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1	Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1
Phe	AAA: 0	GAA: 1					Phe	AAA: 0	GAA: 1				
Asn	AUU: 0	GUU: 1					Asn	AUU: 0	GUU: 1				
Lys	CUU: 0	UUU: 1					Lys	CUU: 0	UUU: 1				

(continued on next page)

Table 3 (continued)

tRNA Isotypes		Isoacceptors				tRNA Isotypes		Isoacceptors					
Asp	AUC: 0	GUC: 1				Asp	AUC: 0	GUC: 1					
Glu	CUC: 0	UUC: 2				Glu	CUC: 0	UUC: 2					
His	AUG: 0	GUG: 1				His	AUG: 0	GUG: 1					
Gln	CUG: 0	UUG: 1				Gln	CUG: 0	UUG: 1					
Ile	AAU: 0	GAU: 0	CAU: 1	UAU: 0		Ile	AAU: 0	GAU: 0	CAU: 1	UAU: 0			
Met	CAU: 2					Met	CAU: 2						
Tyr	AUA: 0	GUA: 1				Tyr	AUA: 0	GUA: 1					
Cys	ACA: 0	GCA: 1				Cys	ACA: 0	GCA: 1					
Trp	CCA: 1					Trp	CCA: 1						
Suppressor	CUA: 0	UUA: 0	UCA: 0			Suppressor	CUA: 0	UUA: 0	UCA: 0				
Sec	UCA: 0					Sec	UCA: 0						
<i>R. piresii</i> (29)						<i>E. equisetina</i> (25)							
Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 0		Ala	AGC: 0	GGC: 0	CGC: 0	UGC: 1			
Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 1		Gly	ACC: 0	GCC: 1	CCC: 0	UCC: 0			
Pro	AGG: 0	GGG: 1	CGG: 0	UGG: 1		Pro	AGG: 0	GGG: 0	CGG: 0	UGG: 1			
Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 1		Thr	AGU: 0	GGU: 1	CGU: 0	UGU: 0			
Val	AAC: 0	GAC: 1	CAC: 0	UAC: 1		Val	AAC: 0	GAC: 1	CAC: 0	UAC: 0			
Ser	AGA: 0	GGA: 1	CGA: 0	UGA: 1	ACU: 0	GCU: 1	Ser	AGA: 0	GGA: 1	CGA: 0	UGA: 1	ACU: 0	GCU: 1
Arg	ACG: 1	GCG: 0	CCG: 1	UCG: 0	CCU: 0	UCU: 1	Arg	ACG: 1	GCG: 0	CCG: 0	UCG: 0	CCU: 0	UCU: 1
Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1	Leu	AAG: 0	GAG: 0	CAG: 0	UAG: 1	CAA: 1	UAA: 1
Phe	AAA: 0	GAA: 1					Phe	AAA: 0	GAA: 1				
Asn	AUU: 0	GUU: 1					Asn	AUU: 0	GUU: 1				
Lys	CUU: 0	UUU: 1					Lys	CUU: 0	UUU: 1				
Asp	AUC: 0	GUC: 2					Asp	AUC: 0	GUC: 1				
Glu	CUC: 0	UUC: 2					Glu	CUC: 0	UUC: 2				
His	AUG: 0	GUG: 1					His	AUG: 0	GUG: 1				
Gln	CUG: 0	UUG: 1					Gln	CUG: 0	UUG: 1				
Ile	AAU: 0	GAU: 0	CAU: 1	UAU: 0			Ile	AAU: 0	GAU: 0	CAU: 1	UAU: 0		
Met	CAU: 2						Met	CAU: 2					
Tyr	AUA: 0	GUA: 1					Tyr	AUA: 0	GUA: 1				
Cys	ACA: 0	GCA: 1					Cys	ACA: 0	GCA: 1				
Trp	CCA: 1						Trp	CCA: 2					
Suppressor	CUA: 0	UUA: 0	UCA: 0				Suppressor	CUA: 0	UUA: 0	UCA: 0			
Sec	UCA: 0						Sec	UCA: 0					



**Table 4** Conserved sequence motifs in chloroplast tRNAs from gymnosperms. Small conserved consensus motifs were observed in the  $\Psi$  region. To be specific, except for tRNA<sup>Ser</sup>, the  $\Psi$ -loop in tRNAs was found to contain a conserved sequence comprising U-U-C-NA-N<sub>2</sub> according to a multiple sequence alignment of 20 members of the tRNA gene family.

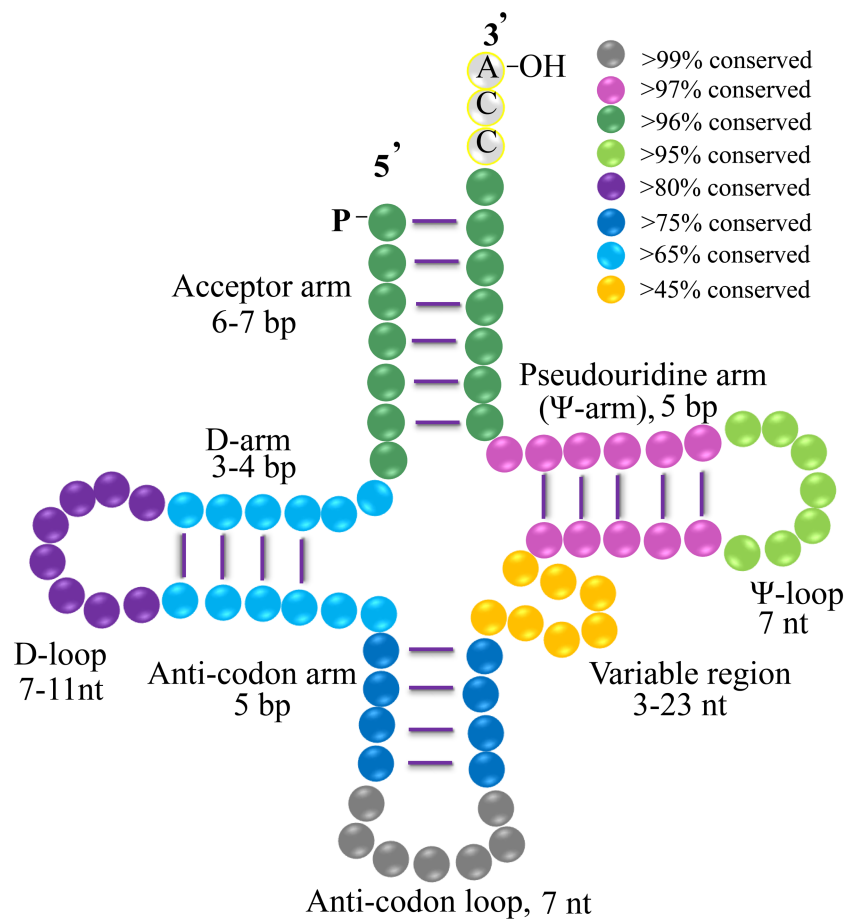
tRNA Isotypes	5' AC-arm	D-arm	D-loop	ANC-arm	ANC-loop	Variable region	$\Psi$ -arm	$\Psi$ -loop
Alanine	5'-G-G-G-A-U-A           3'-C-C-C-U-A-U	5'-G-C-U-C           3'-C-G-A-G	5'-A-G-U-U-G-G-U-A 	5'-C-C-G-C-U           3'-G-G-C-G-A	5'-C-U-U-G-C-A-U 	5'-A-U-G-U-C           3'-U-U-C-G-C-C	5'-A-G-C-G-G           3'-U-U-C-G-A-G-U	5'-U-U-C-G-A-G-U           3'-U-U-C-G-A-G-U
Arginine	N <sub>2</sub> -G-N <sub>2</sub> -G           N <sub>2</sub> -C-N <sub>2</sub> -C	G-N <sub>2</sub>           C-N <sub>2</sub>	A-N <sub>2</sub> -G-G-A-N <sub>12</sub> -A 	N <sub>2</sub> -GN           N <sub>2</sub> -CN	N-U-N-C-N-A-A 	**** 	N <sub>2</sub> -GN           N <sub>2</sub> -CN	U-U-C-N-A-A-U           U-U-C-N-A-A-U
Asparagine	U-C-C-N-C-A-N           A-G-G-N-G-U-N	G-C-U-C           C-G-A-G	A-G-N-G-G-U-A 	G-U-C-G-G           C-A-G-C-C	C-U-G-U-U-A-A 	U-N <sub>13</sub> G-U-C 	G-U-A-G-G           C-A-U-C-C	U-U-C-A-A-A-U           U-U-C-A-A-A-U
Aspartate	G-G-G-A-U-U-G           C-C-C-U-A-A-C	G-U-U-C           C-A-A-G	A-A-U-U-N-G-U-U-A 	C-C-G-N-C           G-G-C-N-G	C-U-G-U-C-A-A 	A-A-G-U-U 	C-C-G-G-G           C-C-C-C-C	U-U-C-G-A-G-N           U-U-C-G-A-G-N
Cysteine	G-G-C-G-N-C-A           C-C-G-C-N-G-U	G-C-C           C-G-G	A-A-G-N-G-G-U-A-A 	G-N <sub>23</sub> -A           C-N <sub>23</sub> -U	C-U-A-C-A-A-A 	**** 	G-G-G-N-C           G-G-G-N-C	U-U-C-G-A-A-U           U-U-C-G-A-A-U
Glutamate	U-G-G-G-G-C-G           A-C-C-C-G-C	G-C-C           C-G-G	A-A-G-N-G-G-U-A-A 	N-C-A-G-G           N-G-U-C-C	U-U-U-U-G-N-U 	U-A-N-U-N <sub>12</sub> 	N-A-A-G-G           N-U-U-G-C	U-U-C-G-A-A-U           U-U-C-G-A-A-U
Glutamine	G-N <sub>2</sub> -C-N <sub>3</sub>           C-N <sub>2</sub> -G-N <sub>2</sub>	G-N <sub>2</sub>           C-N <sub>2</sub>	A-G-N-G-G-N <sub>24</sub> 	**** 	C-U-U-U-C-A-N 	**** 	N <sub>2</sub> -G           N <sub>2</sub> -C	U-U-C-N-A-N-U           U-U-C-N-A-N-U
Glycine	G-C-G-G-N-U-A           C-G-C-C-N-A-U	G-U-U           C-A-A	N-A-N <sub>614</sub> 	U-C-U-C-N           A-G-A-G-N	U-U-G-C-C-A-N 	A-G-A-N 	G-C-G-G-G           C-G-C-C-C	U-U-C-G-A-U-N           U-U-C-G-A-U-N
Histidine	G-C-G-G-A-C-G           C-G-C-C-U-G-C	G-C-C           C-G-G	A-A-G-U-G-G-N <sub>24</sub> -A 	G-U-G-G-A           C-A-C-C-U	U-U-G-U-G-G-A 	C-A-C-G-C 	G-C-G-G-G           G-C-C-C-C	U-U-C-A-A-U-C           U-U-C-A-A-U-C
Isoleucine	G-C-A-U-C-C-A           C-G-U-A-G-G-U	G-C-U-C           C-G-A-G	G-A-A-U-G-G-U-A-A 	C-C-C-A-A           G-G-G-U-U	C-U-C-A-U-A-A 	A-A-N-U-C 	G-C-N-G-G           C-G-N-C-C	U-U-C-A-A-U-U           U-U-C-A-A-U-U
Leucine	G-N <sub>7</sub> -A           C-N <sub>4</sub> -U	G-N-G           C-N-C	N-A-A-U-N-G-U-A-G-A 	**** 	C-U-N-A-N <sub>7</sub> -A 	U-G-N-U-N <sub>19</sub> -G-C-N- U 	N <sub>7</sub> -G           N <sub>7</sub> -C	U-U-C-N-A-N-U           U-U-C-N-A-N-U
Lysine	G-G-G-U-U-G-C           C-C-C-A-A-C-G	A-C-U-C           U-G-A-G	A-A-U-G-G-U-A 	U-C-G-G           A-G-C-C	C-U-U-U-U-A-A 	N <sub>2</sub> -A-G-N-U 	C-C-G-G-G           G-G-C-C-C	U-U-C-N-A-N-U           U-U-C-N-A-N-U
Methionine	N-C-N <sub>14</sub>           N-G-N <sub>14</sub>	**** 	N <sub>15</sub> -G-G-U-N <sub>13</sub> 	**** 	N-U-C-A-U-A-N 	N <sub>2</sub> -U-C 	N <sub>7</sub> -G           N <sub>7</sub> -C	U-U-C-A-A-N-U           U-U-C-A-A-N-U
Phenylalanine	G-C-C-G-G-G-A           C-G-G-C-C-C-U	G-C-U-C           C-G-A-G	A-G-U-U-G-G-U-A 	G-A-G-G-A           C-U-C-C-U	C-U-G-A-A-A-A 	G-U-G-C-C 	A-C-C-A-G           U-G-G-U-C	U-U-C-A-A-A-U           U-U-C-A-A-A-U
Proline	N-G-G-N <sub>4</sub>           N-C-C-N <sub>4</sub>	N-C-G-C           N-G-C-G	A-G-N-U-U-G-G-U-A 	**** 	N-U-N-G-G-G-U 	A-N-G-U-C 	N <sub>7</sub> -G           N <sub>7</sub> -C	U-U-C-A-A-N-U           U-U-C-A-A-N-U
Serine	G-N <sub>7</sub> -A           C-N <sub>7</sub> -U	G-N <sub>23</sub>           C-N <sub>23</sub>	***** 	N <sub>2</sub> -G-N <sub>12</sub>           N <sub>2</sub> -C-N <sub>12</sub>	N-U-N <sub>7</sub> -A-N 	***** 	G-N-G-G-G           C-N-C-C-C	U-U-C-G-N <sub>7</sub> -U           U-U-C-G-N <sub>7</sub> -U
Threonine	G-C-N <sub>2</sub> -G-N <sub>2</sub>           C-G-N <sub>2</sub> -C-N <sub>2</sub>	N-C-U-C           N-G-A-G	A-G-N-G-G-U-N <sub>17</sub> -A 	N-G-C-G-N           N <sub>2</sub> -G-U	N-U-N <sub>7</sub> -A-A 	N <sub>2</sub> -G-U-C 	N-U-N-G-G           N-A-N-G-C	U-U-C-N-A-N <sub>2</sub>           U-U-C-N-A-N <sub>2</sub>
Tryptophan	C-G-G-C-U-C-U           C-G-C-G-A-G-A	G-U-U-N           C-A-A-N	A-G-N <sub>7</sub> -G-G-U-A 	N <sub>2</sub> -G-U           N <sub>2</sub> -C-C-A	C-U-C-C-A-A-A 	A-U-G-N-C 	G-U-A-G-G           C-A-U-C-C	U-U-C-A-A-A-U           U-U-C-A-A-A-U
Tyrosine	N-G-G-N-C-N-A           N-C-C-N-G-N-U	N-C-N <sub>1</sub>           N-G-N <sub>2</sub>	A-G-N <sub>7</sub> -A 	**** 	C-U-G-U-A-A-A 	**** 	N <sub>7</sub> -G           N <sub>7</sub> -C	U-U-C-N-A-N <sub>2</sub>           U-U-C-N-A-N <sub>2</sub>
Valine	A-G-G-G-N-U-A           U-C-C-C-N-A-U	N-C-U-C           N-G-A-G	A-G-N <sub>16</sub> -A 	U-C-N-C-N           A-G-N-G-N	U-U-N-A-C-N <sub>1</sub> 	N <sub>23</sub> -G-U-C 	N <sub>2</sub> -C-N-G           N <sub>2</sub> -G-N-C	U-U-C-N-A-N <sub>2</sub>           U-U-C-N-A-N <sub>2</sub>

#### Notes.

Note that the consensus sequences are shown from 5' to 3'. The asterisk mark (\*) show the absence of conserved nucleotide consensus sequence in respective region of chloroplast tRNAs. 5' AC-arm, 5' Acceptor arm; ANC-arm, Anti-codon arm; ANC-loop, Anti-codon loop;  $\Psi$ -arm, Pseudouridine arm;  $\Psi$ -loop, Pseudouridine loop. The short lines under the bases in anticodon loop of tRNA<sup>Ile</sup> are to indicate its possible modification.

## Diversification of tRNAs structures

The diverse arms and loops of tRNAs allow the regulation and control of protein translation. Each arm and loop has a specific nucleotide composition. Our analysis based on 373 tRNAs showed that the acceptor arm of chloroplast tRNAs contains 6 bp to 7 bp (Table S1). The D-arms were found to contain 3 or 4 bp generally, with a stable "G" in the initial position and "C" in the last position of the D-stem 5' strand in most tRNAs (such as tRNA<sup>Ala</sup>, tRNA<sup>Asn</sup>, tRNA<sup>Asp</sup>, tRNA<sup>Cys</sup>, tRNA<sup>Glu</sup>, tRNA<sup>His</sup>, tRNA<sup>Ile</sup>, and tRNA<sup>Phe</sup>). Most D-loops usually contain 7 to 11 nt with conserved "A" nucleotides at the two end locations. The anticodon arms of chloroplast tRNAs mainly contain 5 bp (90.4%). We found that 367



**Figure 4** Clover leaf-like structure of gymnosperms tRNA. The tRNA contains the Acceptor arm (6–7 bp, dark green, >96% conserved), D-arm (3–4 bp, light blue, >65% conserved), D-loop (7–11 nt, purple, >80% conserved), Anti-codon arm (5 bp, dark blue, >75% conserved), anti-codon loop (7 nt, gray, >99% conserved), variable region (3–23 nt, orange, >45% conserved), Ψ-arm (5 bp, light purple, >97% conserved), and Ψ-loop (7 nt, green, >95% conserved). “% conservation” means the conservative ratio of base identities in each stem and loop structure of the whole set of gymnosperm tRNAs. Several tRNAs harbor the nucleotides of C-C-A tail.

Full-size DOI: [10.7717/peerj.10312/fig-4](https://doi.org/10.7717/peerj.10312/fig-4)

(about 99%) tRNAs contain 7 nt in their anticodon loop, thereby indicating that the sequence of the anticodon loop is highly conserved (Table 4, Table S1). The variable loops of different tRNAs contain 3 to 23 nt, where those in tRNA<sup>Ala</sup>, tRNA<sup>Asp</sup>, tRNA<sup>His</sup>, tRNA<sup>Phe</sup>, and tRNA<sup>Pro</sup> contain 5 bp (Table S1). The Ψ-arm contains 5 bp in most of the gymnosperm chloroplast tRNAs, except for tRNA<sup>Ala</sup> and some of the tRNA<sup>Trp</sup>, tRNA<sup>Gly</sup>, tRNA<sup>Thr</sup>, and tRNA<sup>Arg</sup> in chloroplast. The Ψ-loops of most tRNAs contain 7 nt, apart from tRNA<sup>Ala</sup> and several of tRNA<sup>Cys</sup> and tRNA<sup>Thr</sup> (Table S1).

### Gymnosperm chloroplast tRNAs derived from multiple common ancestors

The phylogenetic tree demonstrated the presence of three major clusters covering 64 groups and the different types of all tRNAs (as shown by the different strings in Fig. S1).

We detected 37 groups in cluster I, five in cluster II, and 22 groups in cluster III. Cluster I contains tRNA<sup>Ser</sup>, tRNA<sup>Tyr</sup>, tRNA<sup>His</sup>, tRNA<sup>Gln</sup>, tRNA<sup>Thr</sup>, tRNA<sup>Pro</sup>, tRNA<sup>Gly</sup>, tRNA<sup>Met</sup>, tRNA<sup>Asp</sup>, tRNA<sup>Arg</sup>, tRNA<sup>Ala</sup>, tRNA<sup>Cys</sup>, tRNA<sup>Lys</sup>, tRNA<sup>Glu</sup>, tRNA<sup>Ile</sup>, tRNA<sup>Asn</sup>, tRNA<sup>Val</sup>, tRNA<sup>Leu</sup>, and tRNA<sup>Trp</sup>. Cluster II contains tRNA<sup>His</sup>, tRNA<sup>Ser</sup>, tRNA<sup>Tyr</sup>, and tRNA<sup>Leu</sup>. Cluster III contains tRNA<sup>Leu</sup>, tRNA<sup>Ile</sup>, tRNA<sup>Gly</sup>, tRNA<sup>Thr</sup>, tRNA<sup>Ser</sup>, tRNA<sup>Val</sup>, tRNA<sup>Glu</sup>, tRNA<sup>Lys</sup>, tRNA<sup>Cys</sup>, tRNA<sup>Gln</sup>, tRNA<sup>His</sup>, tRNA<sup>Arg</sup>, tRNA<sup>Phe</sup>, tRNA<sup>Ala</sup>, and tRNA<sup>Met</sup> (Fig. S1). tRNA<sup>Ser</sup>, tRNA<sup>His</sup>, and tRNA<sup>Leu</sup> are present in cluster I but also in cluster II and cluster III, thereby suggesting that these tRNAs evolved from multiple lineages. Most of the tRNAs were found to form more than one group in the phylogenetic tree. In cluster I, the tRNAs that formed two groups in the phylogenetic tree were identified as tRNA<sup>Tyr</sup>, tRNA<sup>Gln</sup>, tRNA<sup>Met</sup>, tRNA<sup>Asp</sup>, tRNA<sup>Ala</sup>, tRNA<sup>Lys</sup>, tRNA<sup>Ile</sup>, and tRNA<sup>Trp</sup>, whereas those that clustered to form three groups were determined as tRNA<sup>Ser</sup>, tRNA<sup>Pro</sup>, tRNA<sup>Arg</sup>, tRNA<sup>Glu</sup>, tRNA<sup>Asn</sup>, tRNA<sup>Val</sup>, and tRNA<sup>Leu</sup>. Moreover, tRNA<sup>Thr</sup> clustered into four groups. In cluster II, tRNA<sup>Ser</sup> was found to form two groups. In cluster III, tRNA<sup>Gly</sup> and tRNA<sup>Val</sup> were found to form two groups, whereas tRNA<sup>Thr</sup> formed three groups, tRNA<sup>Ile</sup> formed four groups. Some tRNAs in cluster III were found to group individually, where these tRNAs containing the anticodons C-G-A in tRNA<sup>Ser</sup>, U-U-C in tRNA<sup>Glu</sup>, U-U-U in tRNA<sup>Lys</sup>, G-C-A in tRNA<sup>Cys</sup>, U-U-G in tRNA<sup>Gln</sup>, G-U-G in tRNA<sup>His</sup>, U-C-U in tRNA<sup>Arg</sup>, G-A-A in tRNA<sup>Phe</sup>, U-G-C in tRNA<sup>Ala</sup>, and C-A-U in tRNA<sup>Met</sup> all grouped separately (Fig. S1). The multiple groupings of different tRNAs suggest that they evolved from multiple common ancestors. Furthermore, the tRNAs presented in cluster III, i.e., tRNA<sup>Met</sup> (CAU), tRNA<sup>Thr</sup> (UGU, GGU), tRNA<sup>Val</sup> (UAC), tRNA<sup>Ala</sup> (UGC), tRNA<sup>Phe</sup> (GAA), tRNA<sup>Arg</sup> (UCU), tRNA<sup>His</sup> (GUG), tRNA<sup>Gln</sup> (UUG), tRNA<sup>Cys</sup> (GCA), tRNA<sup>Lys</sup> (UUU), tRNA<sup>Glu</sup> (UUC), tRNA<sup>Ile</sup> (UAU), tRNA<sup>Val</sup> (GAC), tRNA<sup>Leu</sup> (CAA), tRNA<sup>Gly</sup> (UCC), tRNA<sup>Ser</sup> (CGA), tRNA<sup>Gly</sup> (GCC), and tRNA<sup>Ile</sup> (CAU), tended to be the most basic tRNAs and they had undergone gene duplication and diversification to generate other tRNA molecules.

### C-A-U anticodon in tRNA<sup>Ile</sup>

Our detailed genomic study showed that tRNA<sup>Ile</sup> also encodes a C-A-U anticodon in addition to the presence of this typical anticodon in tRNA<sup>Met</sup>. In general, the C-A-U anticodon is recognized as a typical characteristic of tRNA<sup>Met</sup> and there is only one iso-acceptor. In particular, we found that the tRNA<sup>Ile</sup> in *T. mairei* encodes two C-A-U anticodons, and *C. debaoensis*, *S. verticillata*, *D. spinulosum*, *C. lanceolata*, *G. biloba*, *C. deodara*, *W. mirabilis*, *G. gnemon*, *R. piresii*, *E. equisetina*, and *W. nobilis* also encode a C-A-U anticodon (Table 3, Data S1, Fig. S3).

### Transition/transversion of tRNAs

A previous study (Mohanta et al., 2019) showed that the evolutionary rates are almost equal for tRNAs with respect to transition and transversion despite the low probability of transition or transversion events in tRNAs. In this study, we identified several intriguing substitutions of gymnosperm chloroplast tRNAs. Overall, our analysis of the substitution rates detected using the whole set of chloroplast tRNAs showed that average transition rate (15.38) was significantly larger than the average transversion rate (4.81) with a ratio of 3:1



**Table 5 Transition and transversion rate of chloroplast tRNA.** In all of the chloroplast tRNAs, the average transition rate (shown in bold value) was slightly higher than the average transversion rate, thereby indicating that chloroplast tRNAs have unequal substitution rates.

From/To	A	U	C	G	From/To	A	U	C	G
<b>Alanine</b>					<b>lysine</b>				
A	–	12.50	12.50	<b>0.00</b>	A	–	1.47	1.47	<b>22.06</b>
U	12.50	–	<b>0.00</b>	12.50	U	1.47	–	<b>22.06</b>	1.47
C	12.50	<b>0.00</b>	–	12.50	C	1.47	<b>22.06</b>	–	1.47
G	<b>0.00</b>	12.50	12.50	–	G	<b>22.06</b>	1.47	1.47	–
<b>Arginine</b>					<b>Methionine</b>				
A	–	2.93	2.93	<b>19.13</b>	A	–	3.86	3.86	<b>17.27</b>
U	2.93	–	<b>19.13</b>	2.93	U	3.86	–	<b>17.27</b>	3.86
C	2.93	<b>19.13</b>	–	2.93	C	3.86	<b>17.27</b>	–	3.86
G	<b>19.13</b>	2.93	2.93	–	G	<b>17.27</b>	3.86	3.86	–
<b>Asparagine</b>					<b>Phenylalanine</b>				
A	–	1.12	1.12	<b>22.75</b>	A	–	1.21	1.21	<b>22.58</b>
U	1.12	–	<b>22.75</b>	1.12	U	1.21	–	<b>22.58</b>	1.21
C	1.12	<b>22.75</b>	–	1.12	C	1.21	<b>22.58</b>	–	1.21
G	<b>22.75</b>	1.12	1.12	–	G	<b>22.58</b>	1.21	1.21	–
<b>Aspartate</b>					<b>Proline</b>				
A	–	0.00	0.00	<b>25.00</b>	A	–	1.53	1.53	<b>21.95</b>
U	0.00	–	<b>25.00</b>	0.00	U	1.53	–	<b>21.95</b>	1.53
C	0.00	<b>25.00</b>	–	0.00	C	1.53	<b>21.95</b>	–	1.53
G	<b>25.00</b>	0.00	0.00	–	G	<b>21.95</b>	1.53	1.53	–
<b>Cysteine</b>					<b>Serine</b>				
A	–	2.75	2.75	<b>19.50</b>	A	–	5.16	5.16	<b>14.68</b>
U	2.75	–	<b>19.50</b>	2.75	U	5.16	–	<b>14.68</b>	5.16
C	2.75	<b>19.50</b>	–	2.75	C	5.16	<b>14.68</b>	–	5.16
G	<b>19.50</b>	2.75	2.75	–	G	<b>14.68</b>	5.16	5.16	–
<b>Glutamine</b>					<b>Threonine</b>				
A	–	3.83	3.83	<b>17.35</b>	A	–	3.91	3.91	<b>17.18</b>
U	3.83	–	<b>17.35</b>	3.83	U	3.91	–	<b>17.18</b>	3.91
C	3.83	<b>17.35</b>	–	3.83	C	3.91	<b>17.18</b>	–	3.91
G	<b>17.35</b>	3.83	3.83	–	G	<b>17.18</b>	3.91	3.91	–
<b>Glutamate</b>					<b>Tryptophan</b>				
A	–	4.59	4.59	<b>15.81</b>	A	–	2.02	2.02	<b>20.96</b>
U	4.59	–	<b>15.81</b>	4.59	U	2.02	–	<b>20.96</b>	2.02
C	4.59	<b>15.81</b>	–	4.59	C	2.02	<b>20.96</b>	–	2.02
G	<b>15.81</b>	4.59	4.59	–	G	<b>20.96</b>	2.02	2.02	–
<b>Glycine</b>					<b>Tyrosine</b>				
A	–	1.94	1.94	<b>21.13</b>	A	–	5.34	5.34	<b>14.33</b>
U	1.94	–	<b>21.13</b>	1.94	U	5.34	–	<b>14.33</b>	5.34
C	1.94	<b>21.13</b>	–	1.94	C	5.34	<b>14.33</b>	–	5.34
G	<b>21.13</b>	1.94	1.94	–	G	<b>14.33</b>	5.34	5.34	–

(continued on next page)

Table 5 (continued)

From/To	A	U	C	G	From/To	A	U	C	G	
		<b>Histidine</b>						<b>Valine</b>		
A	–	1.22	1.22	<b>22.56</b>	A	–	1.73	1.73	<b>21.54</b>	
U	1.22	–	<b>22.56</b>	1.22	U	1.73	–	<b>21.54</b>	1.73	
C	1.22	<b>22.56</b>	–	1.22	C	1.73	<b>21.54</b>	–	1.73	
G	<b>22.56</b>	1.22	1.22	–	G	<b>21.54</b>	1.73	1.73	–	
		<b>Isoleucine</b>						<b>Overall</b>		
A	–	4.72	4.72	<b>15.56</b>	A	–	4.81	4.81	<b>15.38</b>	
U	4.72	–	<b>15.56</b>	4.72	U	4.81	–	<b>15.38</b>	4.81	
C	4.72	<b>15.56</b>	–	4.72	C	4.81	<b>15.38</b>	–	4.81	
G	<b>15.56</b>	4.72	4.72	–	G	<b>15.38</b>	4.81	4.81	–	
		<b>Leucine</b>								
From/To	A	U	C	G						
A	–	4.40	4.40	<b>16.21</b>						
U	4.40	–	<b>16.21</b>	4.40						
C	4.40	<b>16.21</b>	–	4.40						
G	<b>16.21</b>	4.40	4.40	–						

(Table 5). The same transition: transversion ratio bias was found in all the set of tRNAs for tRNA<sup>Ser</sup>, tRNA<sup>Glu</sup>, tRNA<sup>Tyr</sup>, tRNA<sup>Ile</sup>, tRNA<sup>Met</sup>, tRNA<sup>Gln</sup>, tRNA<sup>Thr</sup>, and tRNA<sup>Leu</sup>. The ratio was over 6:1 for tRNA<sup>Cys</sup> and tRNA<sup>Arg</sup>. The transition rates for tRNA<sup>Trp</sup>, tRNA<sup>Val</sup>, and tRNA<sup>Gly</sup> were about 10 times higher than their transversion rates. These findings suggest that tRNA<sup>Ser</sup>, tRNA<sup>Glu</sup>, tRNA<sup>Tyr</sup>, tRNA<sup>Ile</sup>, tRNA<sup>Met</sup>, tRNA<sup>Gln</sup>, tRNA<sup>Thr</sup>, tRNA<sup>Leu</sup>, tRNA<sup>Cys</sup>, tRNA<sup>Arg</sup>, tRNA<sup>Trp</sup>, tRNA<sup>Val</sup> and tRNA<sup>Gly</sup> underwent transition substitutions more readily than transversion substitutions during their evolution in gymnosperm chloroplast genomes. In addition, the transition rates in tRNA<sup>Lys</sup> and tRNA<sup>Pro</sup> were about 15 times higher than their transversion rates. The transition rates in tRNA<sup>Asn</sup>, tRNA<sup>Phe</sup>, and tRNA<sup>His</sup> were about 20 times higher than their transversion rates. These results indicate that tRNAs are much more likely to have undergone transition events rather than transversion events. The highest transversion rate of 12.50 was found in tRNA<sup>Ala</sup> and the lowest transversion rate of 0.00 in tRNA<sup>Asp</sup> (Table 5). Correspondingly, tRNA<sup>Ala</sup> lacks any transitions (Table 5).

### tRNA duplication/loss events

In addition to transition and transversion events, gene duplication and loss events have played important roles in gene evolution. Our analysis of duplication and loss events indicated that 153 duplication events (duplication and conditional duplication) have occurred in all of the gymnosperm chloroplast tRNA genes investigated in this study (Fig. S2). In addition, 220 gymnosperm chloroplast tRNA gene loss events were detected (Table S2, Fig. S2). Thus, the loss of genes was slightly more frequent than their duplication for gymnosperm chloroplast tRNA genes.

## DISCUSSION

tRNAs are major genetic components of semi-autonomous chloroplasts and our analysis of gymnosperm chloroplast genomes showed that they have several basic conserved genomic features. The gymnosperm chloroplast genomes investigated in the present study were found to encode 28 to 33 tRNA isotypes, thereby indicating that there is substantial variation in the quantity of tRNAs in gymnosperm chloroplast genomes. The lack of tRNA<sup>Ala</sup> in *R. piresii* and *T. mairei*, and the absence of tRNA<sup>Val</sup> in *T. mairei* were interesting. Thus, it is necessary to understand how the translation process is conducted in chloroplasts without these crucial tRNAs. According to previous studies ([Treangen & Rocha, 2011](#); [Mohanta et al., 2019](#)), it is likely that the deficiency of these tRNAs is compensated for by the transfer of corresponding tRNAs from the nucleus or mitochondria. In addition to the absence of tRNA<sup>Ala</sup> and tRNA<sup>Val</sup>, all of the gymnosperm plants were shown to not encode selenocysteine tRNA and its suppressor tRNA in their chloroplast genomes ([Table 2](#)). Selenocysteine tRNA and its suppressor tRNA were also not detected in the chloroplast of *Oryza sativa* ([Mohanta & Bae, 2017](#)).

In addition to the presence of C-A-U anticodon in tRNA<sup>Met</sup>, we found that tRNA-CAU is present in tRNA<sup>Ile</sup> ([Table 3](#)). Similarly, the C-A-U anticodon was detected in tRNA<sup>Ile</sup> in *Bacillus subtilis* (Ehrenberg) Cohn and spinach ([Kashdan & Dudock, 1982](#); [Köhler et al., 2014](#)). The possible mechanism that governs the specificity of this amino acid may involve modification of the wobble position in the anticodon by a tRNA-modifying enzyme. Chloroplasts originate from bacteria so the tRNA modifications found in bacteria may also occur in chloroplast tRNAs. In bacteria, the tRNA-modifying enzyme TilS can convert the 5'-C residue in the CAU anticodon of specific tRNA<sup>Ile</sup> molecules into lysidine to decode 5'-AUA (Ile) codons instead of 5'-AUG (Met) codons ([Soma et al., 2003](#)). In addition, when lysidine decodes isoleucine, the tautomer form of lysidine provides compatible hydrogen bond donor-acceptor sites to allow base pairing with "A" and this may help to the recognition of the codon AUA instead of AUG ([Sonawane & Tewari, 2008](#); [Sambhare et al., 2014](#)). The absence of tRNA<sup>Ile</sup>-lysidine synthetase leads to a failure to modify C34 to lysidine in tRNA<sup>Ile</sup> (LAU) (i.e., the synthesis of CAU-tRNA<sup>Ile</sup>) and this inactivates the translation of AUA codons ([Köhler et al., 2014](#)).

During protein coding, a certain species or gene tends to use one or more specific synonym codons, which is referred to as codon usage bias ([Comeron & Aguadé, 1998](#); [Rota-Stabelli et al., 2012](#)). In the present study, tRNA<sup>Arg</sup>-CCG was found to be present in the genomes of nine species but absent from *C. lanceolata*, *T. mairei*, and *E. equisetina*. Similarly, tRNA<sup>Gly</sup>-UCC was shown to be absent from the chloroplast genomes of *C. debaoensis*, *S. verticillata*, *D. spinulosum*, *C. lanceolata*, *T. mairei*, and *E. equisetina* ([Table 3](#)). These results suggest that gymnosperm chloroplast tRNA genes are characterized by codon usage bias ([Wei & Jin, 2017](#); [Li et al., 2015](#)).

In general, the secondary structure of tRNAs is characterized as clover leaf-like, except for a few tRNAs with unusual secondary structures ([Jühling et al., 2018](#)). In our study, we identified clover leaf-like tRNAs with expanded variable loop regions ([Figs. 1 and 2](#)). Numerous tRNA<sup>Leu</sup>, tRNA<sup>Ser</sup>, and tRNA<sup>Tyr</sup> were found to have specific variable loop

configurations in terms of length and structure, suggesting significant structural variation among chloroplast tRNAs. It is interesting to note that there were also stem-loop structures in variable regions of certain tRNAs in cyanobacteria. This might indicate that similar structural variations exist between chloroplast tRNAs and cyanobacterial tRNAs (Mohanta *et al.*, 2017). Future studies will have to determine the biological importance of these variant tRNAs. The novel tRNA structure lacking the D arm might play some other significant functions in the translation process and additional research is necessary to elucidate its exact function and mechanisms. Most tRNAs have a clover-like structure formed by complementary base pairing between small segments (Hubert *et al.*, 1998; Florentz, 2002). Previous studies have showed that the acceptor arm of tRNAs in chloroplasts contain 7 bp to 9 bp, the D-arm contains 3 bp to 4 bp, the D-loop has 4 nt to 12 nt, the anticodon arm has 5 bp, the anticodon loop contains 7 nt, the variable region comprises 4 nt to 23 nt, and  $\Psi$ -arm contains 5 bp, and the  $\Psi$ -loop has 7 nt (Wilusz, 2015; Mohanta & Bae, 2017; Mohanta *et al.*, 2019). In the present study, we found that the acceptor arm of chloroplast tRNAs contains 6 bp to 7 bp in 373 tRNAs, where the D-arm has 3 bp or 4 bp and the D-loop usually contains 7 nt to 11 nt. The anticodon loop of gymnosperm chloroplast tRNAs generally contains 7 nt, and thus the sequence of the anticodon loop is typically conserved (Table 4, Table S1). The variable loop of different tRNAs contain 3 nt to 23 nt (Table S1). The  $\Psi$ -arm of gymnosperm chloroplast tRNAs generally contains 5 bp and the  $\Psi$ -loop has 7 nt (Table S1). Our results are consistent with previous findings (Wilusz, 2015; Mohanta & Bae, 2017) and they suggest that chloroplast RNAs are significantly conserved. The consensus sequence “U-U-C-N-A-N<sub>2</sub>” was found in the  $\Psi$  region (Table 4). Previous studies also reported the existence of a similar sequence in the  $\Psi$ -loop of tRNAs in *Oryza sariva* and Cyanobacteria (Mohanta & Bae, 2017; Mohanta *et al.*, 2017). This suggests that the consensus “U-U-C-N-A-N<sub>2</sub>” motif of the  $\Psi$  region, identified here and in previous analyses, is a general consensus motif of canonical tRNAs.

Our phylogenetic analysis detected three clear clusters and many tRNA groups. Some tRNAs (tRNA<sup>Ser</sup>, tRNA<sup>His</sup>, and tRNA<sup>Leu</sup>) in cluster I and cluster II were also in cluster III, thereby indicating that these tRNAs evolved from multiple lineages by gene duplication and gene divergence. Moreover, anticodon types comprising CGA, UUC, UUU, GCA, UUG, GUG, UCU, UGC, and CAU appeared several times in the phylogenetic tree, and thus the corresponding tRNAs evolved from multiple common ancestors. The overlapping of tRNAs groups demonstrates that these tRNAs might have diverse common ancestors in the evolutionary process (Mohanta & Bae, 2017). Phylogenetic analysis also showed that tRNA<sup>Met</sup> (CAU), tRNA<sup>Thr</sup> (UGU, GGU), tRNA<sup>Val</sup> (UAC), tRNA<sup>Ala</sup> (UGC), tRNA<sup>Phe</sup> (GAA), tRNA<sup>Arg</sup> (UCU), tRNA<sup>His</sup> (GUG), tRNA<sup>Gln</sup> (UUG), tRNA<sup>Cys</sup> (GCA), tRNA<sup>Lys</sup> (UUU), tRNA<sup>Glu</sup> (UUC), tRNA<sup>Ile</sup> (UAU), tRNA<sup>Val</sup> (GAC), tRNA<sup>Leu</sup> (CAA), tRNA<sup>Gly</sup> (UCC), tRNA<sup>Ser</sup> (CGA), tRNA<sup>Gly</sup> (GCC), and tRNA<sup>Ile</sup> (CAU) in cluster III tended to be the most basic tRNAs, whereas tRNA<sup>Met</sup> tended to be the most original tRNA. Overall, the results clearly indicate that the tRNAs encoded in gymnosperm chloroplast genomes have multiple common evolutionary ancestors.

Our results also provided insights into the gene substitution rates in gymnosperm chloroplast tRNAs. Overall, the average transition rate for tRNAs was greater than the

transversion rate, where the relationship was about 3:1 (Table 5). In all of the chloroplast tRNAs, the average transition rate was slightly higher than the average transversion rate, thereby indicating that chloroplast tRNAs have unequal substitution rates.

In addition to the transition and transversion events in tRNAs, loss and duplication events have played significant roles in the evolution of tRNAs in gymnosperm chloroplast genomes (He & Zhang, 2006; Magadum et al., 2013). In general, the gene loss events tended to occur after whole genome duplication events. We found 153 duplication events and 220 loss events in gymnosperm chloroplast tRNAs, and thus loss events have occurred slightly more frequently than duplication events (Table S2).

## CONCLUSIONS

Our basic structure analysis showed that gymnosperm chloroplast genomes encode 25 to 30 anticodon-specific tRNAs. The acceptor arm of chloroplast tRNA contains 6 bp to 7 bp, the D-arm has 3 bp or 4 bp, the D-loop contains 7 nt to 11 nt mainly, and the anticodon loop usually contains 7 nt. In different tRNAs, the variable loop contains 3 nt to 23 nt. The  $\Psi$ -arm contains a conserved sequence comprising U-U-C-N-A-N<sub>2</sub>. tRNA<sup>Ala</sup> was absent from *R. piresii* and *T. mairei*, and tRNA<sup>Val</sup> was lacking in *T. mairei*. Gymnosperm chloroplasts do not encode selenocysteine tRNA and its suppressor tRNA in their genomes. A CAU anticodon is encoded in tRNA<sup>Met</sup> as well as in tRNA<sup>Ile</sup>. A novel tRNA structure lacking the D arm was identified for the chloroplast tRNA<sup>Gly</sup> of *W. nobilis*. Numerous tRNA<sup>Leu</sup>, tRNA<sup>Ser</sup>, and tRNA<sup>Tyr</sup> types were found to have expanded variable regions. Phylogenetic analysis showed that tRNAs might have multiple common ancestors in the evolutionary process. Different tRNAs harbored their own transition/transversion rates, i.e., it was iso-acceptor specific. And the transition rate was generally higher than the transversion rate. Furthermore, gene loss events (220) have occurred slightly more frequently than gene duplication events (153) in gymnosperm chloroplast tRNAs. Our results provide new insights into the evolution of gymnosperm chloroplast tRNAs and their diverse roles.

## ACKNOWLEDGEMENTS

We thank Mr. Heng Liu for his kindly help for the evolutionary analysis of chloroplast tRNA. We also thank College of Life Sciences, Northwest University for supporting device platform for this study.

## ADDITIONAL INFORMATION AND DECLARATIONS

### 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 and 2018-43) and the Open Foundation of Key Laboratory of Resource Biology and Biotechnology in Western China (Ministry of Education) (Grants no. ZSK2017007

and ZSK2019008). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

National Natural Science Foundation of China: 31970359.

Shaanxi Science and Technology Innovation Team: 2019TD-012.

Public health specialty in the Department of traditional Chinese Medicine: 2017-66, 2018-43.

Open Foundation of Key Laboratory of Resource Biology and Biotechnology in Western China (Ministry of Education): ZSK2017007, ZSK2019008.

### Competing Interests

The authors declare there are no competing interests.

### Author Contributions

- Ting-Ting Zhang and Yi-Kun Hou performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Ting Yang and Ming Yue analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Shu-Ya Zhang analyzed the data, prepared figures and/or tables, and approved the final draft.
- Jianni Liu conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Zhonghu Li conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

### Data Availability

The following information was supplied regarding data availability:

tRNA sequences of gymnosperms chloroplast genome conducted in the study are available in the [Supplementary File](#).

### Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.10312#supplemental-information>.

## REFERENCES

- Abdallah F, Salamini F, Leister D. 2000.** A prediction of the size and evolutionary origin of the proteome of chloroplasts of *Arabidopsis*. *Trends in Plant Science* 5:141–142 DOI 10.1016/S1360-1385(00)01574-0.
- Chen K, Durand D, Farach-Colton M. 2000.** Notung: a program for dating gene duplications and optimizing gene family trees. *Journal of Computational Biology: a Journal of Computational Molecular Cell Biology* 7:429–447 DOI 10.1089/106652700750050871.

- Christenhusz MJM, Reveal JL, Farjon A, Gardner MF, Mill RR, Chase MW. 2010.** A new classification and linear sequence of extant gymnosperms. *Phytotaxa* **19**:55–70 DOI [10.11646/phytotaxa.19.1.3](https://doi.org/10.11646/phytotaxa.19.1.3).
- Civan P, Foster PG, Embley MT, Séneca A, Cox CJ. 2014.** Analyses of charophyte chloroplast genomes help characterize the ancestral chloroplast genome of land plants. *Genome Biology and Evolution* **6**:897–911 DOI [10.1093/gbe/evu061](https://doi.org/10.1093/gbe/evu061).
- Comeron JM, Aguadé M. 1998.** An evaluation of measures of synonymous codon usage bias. *Journal of Molecular Evolution* **47**:268–274 DOI [10.1007/PL00006384](https://doi.org/10.1007/PL00006384).
- Crisp MD, Cook LG. 2011.** Cenozoic extinctions account for the low diversity of extant gymnosperms compared with angiosperms. *The New Phytologist* **192**:997–1009 DOI [10.1111/j.1469-8137.2011.03862.x](https://doi.org/10.1111/j.1469-8137.2011.03862.x).
- Des Marais DJ. 2000.** When did photosynthesis emerge on earth? *Science* **289**:1703–1705 DOI [10.1126/science.289.5485.1703](https://doi.org/10.1126/science.289.5485.1703).
- Dong WL, Wang RN, Yan XH, Niu C, Gong LL, Li ZH. 2016.** Characterization of polymorphic microsatellite markers in *Pinus armandii* (pinaceae), an endemic conifer species to China. *Applications in Plant Sciences* **4**:Article 1600072 DOI [10.3732/apps.1600072](https://doi.org/10.3732/apps.1600072).
- Drouin G, Daoud H, Xia J. 2008.** Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Molecular Phylogenetics and Evolution* **49**:827–831 DOI [10.1016/j.ympev.2008.09.009](https://doi.org/10.1016/j.ympev.2008.09.009).
- Duchene D, Bromham L. 2013.** Rates of molecular evolution and diversification in plants: chloroplast substitution rates correlate with species-richness in the Proteaceae. *BMC Evolutionary Biology* **13**:65 DOI [10.1186/1471-2148-13-65](https://doi.org/10.1186/1471-2148-13-65).
- Fauré S, Noyer JL, Carreel F, Horry JP, Bakry F, Lanaud C. 1994.** Maternal inheritance of chloroplast genome and paternal inheritance of mitochondrial genome in bananas (*Musa acuminata*). *Current Genetics* **25**:265–269 DOI [10.1007/BF00357172](https://doi.org/10.1007/BF00357172).
- Florentz C. 2002.** Molecular investigations on tRNAs involved in Human mitochondrial disorders. *Bioscience Reports* **22**:81–98 DOI [10.1023/a:1016065107165](https://doi.org/10.1023/a:1016065107165).
- Forest F, Moat J, Baloch E, Brummitt NA, Bachman SP, Ickert-Bond S, Hollingsworth PM, Liston A, Little DP, Mathews S, Rai H, Rydin C, Stevenson DW, Thomas P, Buerki S. 2018.** Gymnosperms on the edge. *Scientific Reports* **8**:1–11 DOI [10.1038/s41598-018-24365-4](https://doi.org/10.1038/s41598-018-24365-4).
- Gerrienne P, Meyer-Berthaud B, Faison-Demaret M, StreeL M, Steemans P. 2004.** Runcaria, a middle devonian seed plant precursor. *Science* **306** DOI [10.1126/science.1102491](https://doi.org/10.1126/science.1102491).
- Giegé R, Puglisi JD, Florentz C. 1993.** tRNA structure and aminoacylation efficiency. *Progress in Nucleic Acid Research and Molecular Biology* **45**:129–206 DOI [10.1016/S0079-6603\(08\)60869-7](https://doi.org/10.1016/S0079-6603(08)60869-7).
- Guo SW, Zhou Y, Gao YX, Li Y, Shen QR. 2007.** New insights into the nitrogen form effect on photosynthesis and photorespiration. *Pedosphere* **17**:601–610 DOI [10.1016/S1002-0160\(07\)60071-X](https://doi.org/10.1016/S1002-0160(07)60071-X).

- He XL, Zhang JZ. 2006.** Transcriptional reprogramming and backup between duplicate genes: is it a genome wide phenomenon? *Genetics* **172**:1363–1367 DOI [10.1534/genetics.105.049890](https://doi.org/10.1534/genetics.105.049890).
- Hereward JP, Werth JA, Thornby DF, Keenan M, Chauhan BS, Walter GH. 2018.** Complete chloroplast genome of glyphosate resistant *Sonchus oleraceus* L. from Australia, with notes on the small single copy (SSC) region orientation. *Mitochondrial DNA Part B* **3**:363–364 DOI [10.1080/23802359.2018.1450682](https://doi.org/10.1080/23802359.2018.1450682).
- Heuertz M, Fineschi S, Anzidei M, Pastorelli R, Salvini D, Paule L, Frascaria-Lacoste N, Hardy OJ, Vekemans X, Vendramin GG. 2004.** Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. *Molecular Ecology* **13**:3437–3452 DOI [10.1111/j.1365-294X.2004.02333](https://doi.org/10.1111/j.1365-294X.2004.02333).
- Hiroki I, Daisuke S. 2018.** Bacterial heterologous expression system for reconstitution of chloroplast inner division ring and evaluation of its contributors. *International Journal of Molecular Sciences* **19**:544–556 DOI [10.3390/ijms19020544](https://doi.org/10.3390/ijms19020544).
- Hoober JK. 2006.** Chloroplast development: whence and whither. In: Wise RR, Hoober JK, eds. *The structure and function of plastids*. Springer Netherlands 27–51.
- Hubert N, Sturchler C, Westhof E, Carbon P, Krol A. 1998.** The 9/4 secondary structure of eukaryotic selenocysteine tRNA: more pieces of evidence. *RNA* **4**:1029–1033 DOI [10.1017/S1355838298980888](https://doi.org/10.1017/S1355838298980888).
- Jühling T, Duchardt-Ferner E, Bonin S, Wöhnert J, Pütz J, Florentz C, Betat H, Sauter C, Mörl M. 2018.** Small but large enough: structural properties of armless mitochondrial tRNAs from the nematode *romanomermis culicivorax*. *Nucleic Acids Research* **46**:9170–9180 DOI [10.1093/nar/gky593](https://doi.org/10.1093/nar/gky593).
- Kanai A. 2015.** Disrupted tRNA genes and tRNA fragments: a perspective on tRNA gene evolution. *Life* **5**:321–331 DOI [10.3390/life5010321](https://doi.org/10.3390/life5010321).
- Kashdan MA, Dudock BS. 1982.** The gene for spinach chloroplast isoleucine tRNA has a methionine anticodon. *The Journal of Biological Chemistry* **257**:11191–11194.
- Kaundun SS, Matsumoto S. 2011.** Molecular evidence for maternal inheritance of the chloroplast genome in tea, *Camellia sinensis* (L.) O. Kuntze. *Journal of the Science of Food and Agriculture* **91**:2660–2663 DOI [10.1002/jsfa.4508](https://doi.org/10.1002/jsfa.4508).
- Kim S, Suh Y. 2013.** Phylogeny of Magnoliaceae based on ten chloroplast DNA regions. *Journal of Plant Biology* **56**:290–305 DOI [10.1007/s12374-013-0111-9](https://doi.org/10.1007/s12374-013-0111-9).
- Kirchner S, Ignatova Z. 2014.** Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nature Reviews. Genetics* **16**:98–112 DOI [10.1038/nrg3861](https://doi.org/10.1038/nrg3861).
- Knorr W, Heimann M. 2001.** Uncertainties in global terrestrial biosphere modeling: 1. A comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme. *Global Biogeochemical Cycles* **15**:207–225 DOI [10.1029/1998gb001059](https://doi.org/10.1029/1998gb001059).
- Köhler C, Mandal D, Gaston KW, Grosjean H, Limbach PA, RajBhandary UL. 2014.** Life without tRNA<sup>Leu</sup>-lysine synthetase: translation of the isoleucine codon AUA in *Bacillus subtilis* lacking the canonical tRNA<sup>Leu</sup>. *Nucleic Acids Research* **42**:1904–1915 DOI [10.1093/nar/gkt1009](https://doi.org/10.1093/nar/gkt1009).



- Kretschmer M, Croll D, Kronstad JW. 2017.** Chloroplast-associated metabolic functions influence the susceptibility of maize to *Ustilago maydis*. *Molecular Plant Pathology* **18**:1210–1221 DOI [10.1111/mpp.12485](https://doi.org/10.1111/mpp.12485).
- Kumar S, Nei M, Dudley J, Tamura K. 2008.** MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics* **9**:299–306 DOI [10.1093/bib/bbn017](https://doi.org/10.1093/bib/bbn017).
- Kumar S, Stecher G, Tamura K. 2016.** MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology* **33**:1870–1874 DOI [10.1093/molbev/msw054](https://doi.org/10.1093/molbev/msw054).
- Kumar S, Tamura K, Nei M. 1994.** Mega: molecular evolutionary genetics analysis software for microcomputers. *Computer Applications in the Biosciences: CABIOS* **10**:189–191 DOI [10.1093/bioinformatics/10.2.189](https://doi.org/10.1093/bioinformatics/10.2.189).
- Laslett D, Canback B. 2004.** ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research* **32**:11–16 DOI [10.1093/nar/gkh152](https://doi.org/10.1093/nar/gkh152).
- Li J, Zhou J, Wu Y, Yang SH, Tian DC. 2015.** GC-content of synonymous codons profoundly influences amino acid usage. *G3: Genes - Genomes - Genetics* **5**:2027–2036 DOI [10.1534/g3.115.019877](https://doi.org/10.1534/g3.115.019877).
- Liu L, Hao ZZ, Liu YY, Wei XX, Cun YZ, Wang XQ. 2014.** Phylogeography of *Pinus armandii* and its relatives: heterogeneous contributions of geography and climate changes to the genetic differentiation and diversification of Chinese white pines. *PLOS ONE* **9**:e85920 DOI [10.1371/journal.pone.0085920](https://doi.org/10.1371/journal.pone.0085920).
- Logacheva MD, Penin AA, Valiejo-Roman CM, Antonov AS. 2009.** Structure and evolution of junctions between inverted repeat and small single copy regions of chloroplast genome in non-core Caryophyllales. *Molecular Biology* **43**:757–765 DOI [10.1134/s0026893309050070](https://doi.org/10.1134/s0026893309050070).
- Lowe TM, Eddy SR. 1997.** tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25**:955–964 DOI [10.1093/nar/25.5.0955](https://doi.org/10.1093/nar/25.5.0955).
- Lusic H, Gustilo EM, Vendeix FA, Kaiser R, Delaney MO, Graham WD, Moyer VA, Cantara WA, Agris PF, Deiters A. 2008.** Synthesis and investigation of the 5-formylcytidine modified, anticodon stem and loop of the human mitochondrial tRNA<sup>Met</sup>. *Nucleic Acids Research* **36**:6548–6557 DOI [10.1093/nar/gkn703](https://doi.org/10.1093/nar/gkn703).
- Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R. 2013.** Gene duplication as a major force in evolution. *Journal of Genetics* **92**:155–161 DOI [10.1007/s12041-013-0212-8](https://doi.org/10.1007/s12041-013-0212-8).
- Mizutani T, Goto C. 2000.** Eukaryotic selenocysteine tRNA has the 9/4 secondary structure. *FEBS Letters* **466**(00):359–362 DOI [10.1016/s0014-5793\(00\)01104-2](https://doi.org/10.1016/s0014-5793(00)01104-2).
- Mohanta TK, Bae H. 2017.** Analyses of genomic tRNA reveal presence of novel tRNAs in *Oryza sativa*. *Frontiers in Genetics* **8**:Article 90 DOI [10.3389/fgene.2017.00090](https://doi.org/10.3389/fgene.2017.00090).
- Mohanta TK, Khan AL, Hashem A, Abd\_Allah EF, Yadav D, Al-Harrasi A. 2019.** Genomic and evolutionary aspects of chloroplast tRNA in monocot plants. *BMC Plant Biology* **19**:39 DOI [10.1186/s12870-018-1625-6](https://doi.org/10.1186/s12870-018-1625-6).

- Mohanta TK, Syed AS, Fuad A, Hanhong B. 2017.** Novel genomic and evolutionary perspective of cyanobacterial tRNAs. *Frontiers in Genetics* **8**:Article 200 DOI [10.3389/fgene.2017.00200](https://doi.org/10.3389/fgene.2017.00200).
- Pilon-Smits EAH, Garifullina GF, Abdel-Ghany S, Kato SI, Mihara H, Hale KL, Burkhead JL, Esaki N, Kurihara T, Pilon M. 2002.** Characterization of a NifS-like chloroplast protein from *Arabidopsis*. Implications for its role in sulfur and selenium metabolism. *Plant Physiology* **130**:1309–1318 DOI [10.1104/pp.102.010280](https://doi.org/10.1104/pp.102.010280).
- Ran JH, Gao H, Wang XQ. 2010.** Fast evolution of the retroprocessed mitochondrial rps3 gene in Conifer II and further evidence for the phylogeny of gymnosperms. *Molecular Phylogenetics and Evolution* **54**:136–149 DOI [10.1016/j.ympev.2009.09.011](https://doi.org/10.1016/j.ympev.2009.09.011).
- Ribas PL, Dedon PC. 2014.** More than an adaptor molecule: the emerging role of tRNA in cell signaling and disease. *FEBS Letters* **588**:4267 DOI [10.1016/j.febslet.2014.10.022](https://doi.org/10.1016/j.febslet.2014.10.022).
- Rota-Stabelli O, Lartillot N, Philippe H, Pisani D. 2012.** Serine codon-usage bias in deep phylogenomics: pancrustacean relationships as a case study. *Systems Biology* **62**:121–133 DOI [10.1093/sysbio/sys077](https://doi.org/10.1093/sysbio/sys077).
- Sambhare SB, Kumbhar BV, Kamble AD, Bavi RS, Kumbhar NM, Sonawane KD. 2014.** Structural significance of modified nucleosides k<sup>2</sup>C and t<sup>6</sup>A present in the anticodon loop of tRNA<sup>lle</sup>. *RSC Advances* **4**:14176–14188 DOI [10.1039/c3ra47335](https://doi.org/10.1039/c3ra47335).
- Schimmel P. 2017.** The emerging complexity of the tRNA world: mammalian tRNA beyond protein synthesis. *Nature Reviews. Molecular Cell Biology* **19**:45–58 DOI [10.1038/nrm.2017.77](https://doi.org/10.1038/nrm.2017.77).
- Soma A, Ikeuchi Y, Kanemasa S, Kobayashi K, Ogasawara N, Ote T, Kato J, Watanabe K, Sekine Y, Suzuki T. 2003.** An RNA-modifying enzyme that governs both the codon and amino acid specificities of isoleucine tRNA. *Molecular Cell* **12**:689–698 DOI [10.1016/s1097-2765\(03\)00346-0](https://doi.org/10.1016/s1097-2765(03)00346-0).
- Sonawane KD, Tewari R. 2008.** Conformational preferences of hypermodified nucleoside lysidine (k<sup>2</sup>C) occurring at wobble position in anticodon loop of tRNA<sup>lle</sup>. *Nucleosides, Nucleotides and Nucleic Acids* **27**:1158–1174 DOI [10.1080/15257770802341475](https://doi.org/10.1080/15257770802341475).
- Suzuki T, Suzuki T. 2014.** A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs. *Nucleic Acids Research* **42**:7346–7357 DOI [10.1093/nar/gku390](https://doi.org/10.1093/nar/gku390).
- Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017.** GeSeq-versatile and accurate annotation of organelle genomes. *Nucleic Acids Research* **45**:W6–W11 DOI [10.1093/nar/gkx391](https://doi.org/10.1093/nar/gkx391).
- Treangen TJ, Rocha EPC. 2011.** Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLOS Genetics* **7**:e100128 DOI [10.1371/journal.pgen.1001284](https://doi.org/10.1371/journal.pgen.1001284).
- Wang RJ, Cheng CL, Chang CC, Wu CL, Su TM, Chaw SM. 2008.** Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evolutionary Biology* **8**:36 DOI [10.1186/1471-2148-8-36](https://doi.org/10.1186/1471-2148-8-36).
- Wang XQ, Ran JH. 2014.** Evolution and biogeography of gymnosperms. *Molecular Phylogenetics and Evolution* **75**:24–40 DOI [10.1016/j.ympev.2014.02.005](https://doi.org/10.1016/j.ympev.2014.02.005).

- Wei QK, Jin HY. 2017.** The complete chloroplast genome sequence of *Morus cathayana* and *Morus multicaulis*, and comparative analysis within genus *Morus* L. *PeerJ* 5:e3037 DOI [10.7717/peerj.3037](https://doi.org/10.7717/peerj.3037).
- Wilusz JE. 2015.** Controlling translation via modulation of tRNA levels. *Wiley Interdisciplinary Reviews. RNA* 6:453–470 DOI [10.1002/wrna.1287](https://doi.org/10.1002/wrna.1287).
- Yu F, Wang DX, Yi XF, Shi XX, Huang YK, Zhang HW, Zhang XP. 2014.** . Does animal-mediated seed dispersal facilitate the formation of *Pinus armandii-quercus* aliena var. *acuteserrata* forests? *PLOS ONE* 9:e89886 DOI [10.1371/journal.pone.0089886](https://doi.org/10.1371/journal.pone.0089886).