Genomic and evolutionary aspects of chloroplast tRNA in monocot plants

Tapan Kumar Mohanta*, Abdul Latif Khan†, Abeer Hashem‡,§, Elsayed Fathi Abd_ Allah¶, Dhananjay Yadav∥ and Ahmed Al-Harrasi†

Abstract

Background: Chloroplasts are one of the most indispensable organelles that make life forms on the earth possible by their capacity to photosynthesize. These organelles possess a circular genome with a number of coding genes responsible for self-regulation. tRNAs are an important evolutionary-conserved gene family that are responsible for protein translation. However, within the chloroplast genome, tRNA machinery are poorly understood.

Results: In the present study, the chloroplast genome of six monocot plants, Oryza nivara (NC_005973), Oryza sativa (NC_001320), Sachharum officinarum (NC_006084), Sorghum bicolor (NC_008602), Triticum aestivum (NC_002762), and Zea mays (NC_001666) were downloaded and analyzed to identify tRNA sequences. Further analysis of the tRNA sequences in the chloroplast genomes of the monocot plants resulted in the identification of several novel features. The length of tRNAs in the chloroplast genome of the monocot plants ranged from 59 to 155 nucleotides. Pair-wise sequence alignment revealed the presence of a conserved A-C-x-U-A-x-U-A-x-U-x 5-U-A-A nucleotide consensus sequence. In addition, the tRNAs in chloroplast genomes of the monocot plants also contain 21–28 anti-codons against 61 sense codons in the genome. They also contain a group I intron and a C-A-U anti-codon for tRNA Ile, which is a common anti-codon of tRNA Met. Evolutionary analysis indicates that tRNAs in the chloroplast genome have evolved from multiple common ancestors, and tRNA Met appears to be the ancestral tRNA that underwent duplication and diversification to give rise to other tRNAs.

Conclusion: The results obtained from the study of chloroplast tRNA will greatly help to increase our understanding of tRNA biology at a new level. Functional studies of the reported novel aspects of the chloroplast tRNA of the monocot plants will greatly help to decipher their roles in diverse cellular processes.

Keywords: tRNA, Chloroplast, Anti-codon, Evolution, Transition, Transversion, Phylogenetics

Background

Chloroplasts are multi-copy cellular organelle [1] which are responsible for photosynthesis and carbohydrate metabolism in photosynthetic plants which regulate our biosphere [2, 3]. They are an active metabolic center, and are responsible for sustaining the life on earth by converting solar energy into carbohydrates through the process of photosynthesis [4–6]. In addition to the major process of photosynthesis, chloroplasts also play an important role in various other molecular processes; including the synthesis of nucleotides, amino acids, fatty acids, vitamins, phytohormones, and several other metabolites [7–12]. Furthermore, they also contribute to the assimilation of nitrogen and sulphur [13–15]. In plants, these metabolites have been shown to play a critical role in the regulation of the physiology, growth, and development; as well as stress response. Therefore, chloroplasts can be regarded as the “metabolic center” of cellular reactions. Evolutionary studies indicate that chloroplasts have arisen from a cyanobacterial ancestor through internalization within a eukaryotic cell and have maintained an independent genome inside the plant cell [16–20]. The chloroplast genome (cpDNA) is a double stranded circular molecule containing tRNA, rRNA, and a number of protein coding genes [21]. The majority of the protein coding genes are

* Correspondence: nostoc.tapan@gmail.com; tapan.mohanta@unizwa.edu.om; dhanyadav16@gmail.com
†Natural and Medical Sciences Research Center, University of Nizwa, 616 Nizwa, Oman
‡Department of Medical Biotechnology, Yeungnam University, Gyeongsan 38541, Republic of Korea
Full list of author information is available at the end of the article

© The Author(s). 2019 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
associated with photosynthesis and bioenergetics [22, 23]. The chloroplast genome contains two large 6–76 Kb inverted repeats (IRs) that are divided into a large single copy (LSC) and small single copy region (SSC) [24–26]. The chloroplast genome is non-recombinant and inherited uniparentally through maternal inheritance [27, 28]. Therefore, the chloroplast genome is an excellent tool for genomic and evolutionary studies. It is very difficult, however, to detect polymorphisms in cpDNA due to a low level of substitutions [29, 30]. Recently, the advances in high-throughput genome sequencing technology have enabled rapid progress in the sequencing and analysis of chloroplast genomes. Specifically, these technological gains have enabled us to obtain and analyze the complete chloroplast genomes of several plants to better understand their molecular and genomic characteristics.

Since chloroplasts encode a complete and independent genome, it is important to study the chloroplast genomes; especially chloroplast tRNAs which are responsible for protein translation. Since the chloroplast genome is involved in the synthesis of nucleotides, amino acids and proteins, it is important to understand its organization to determine how these processes are regulated within the chloroplast genome. Protein translation within the chloroplast is regulated by tRNA and other associated genes. Thus, detailed analyses of chloroplast tRNAs can provide insight into the genomics and evolution of cyanobacterial tRNAs. In relative comparison to eucots, the monocot genome is more conserved than the eucots genome, and they have evolved from the eucot lineage [31–33]. In addition, several of the important agronomic crops species are monocots. Therefore, in the present study, we considered to study the chloroplast genome of six monocot plants to better understand the genomic and evolutionary characteristics of the chloroplast tRNA that can enable functional studies for the future.

tRNAs are one of the most important and versatile molecules responsible for sustaining and maintaining the protein translation machinery. They are characterized by the presence of a clover leaf like structure as proposed by Robert Holley [34]. This structure contains features such as an acceptor arm, D-arm, D-loop, anti-codon arm, anti-codon loop, variable arm, pseudouridine arm, and pseudouridine loop. The tRNAs are encoded within the nuclear genome and in the genome of sub-cellular organelles, including plastids and mitochondria. Over the years, detailed studies pertaining to the characterization of nuclear tRNA have gained considerable attention [35–37]. Structure and function of tRNAs and tRNA genes of chloroplast genome was previously described by Mareachal-Drouard et al., (1991) [38]. However, due to the lack of complete genome sequences of chloroplast genome, the study lacked the complete genomic details of tRNAs of plastid genome. Therefore, we attempted to understand the detailed genomic and molecular aspects of chloroplast tRNA in plants. Considering the conserved evolutionary lineages of monocots, six economically important monocots were investigated and reported within this study.

Results
Genomic of chloroplast tRNA
The whole chloroplast genome sequence of six monocot plants, Oryza nivara (NC_005973), Oryza sativa (NC_001320), Saccharum officinarum (NC_006084), Sorghum bicolor (NC_008602), Triticum aestivum (NC_002762), and Zea mays (NC_001666), were downloaded from the National Center of Biotechnology Information (NCBI) database. Subsequently, the sequences were annotated to identify the genomic tRNA sequences in these genomes (Fig. 1). The obtained genomic tRNA sequences were further analyzed using the tRNAscan-SE server to confirm their identity as tRNAs. Results indicated that O. nivara, O. sativa, S. officinarum, S. bicolor, T. aestivum, and Z. mays encode 38, 35, 37, 29, 39, and 39 tRNAs, respectively (Table 1). The length of the chloroplast tRNAs ranged from 59 nt [tRNA^Glu GGU, Sorghum bicolor, (20385)] to 155 nt [tRNA^Lys NNN, T. aestivum, (4982_TraeCt095)]. tRNA^Gly UCC of O. nivara (6129) was found to contain only 65 nt, whereas tRNA^Glu UUG of T. aestivum (4985) and tRNA^Leu(UAG) of T. aestivum (5086_TraeCt128) contained 118 nt and 100 nt, respectively. In the tRNA, tRNA^Gln (4985_TraeCt096), the tRNA begins at 46 nt and in tRNA^Leu(UAG) (5086_TraeCt128), it begins at 21 nt. Pairwise sequence alignment of 5′ nucleotide sequence of these two tRNAs revealed a 22.2% similarity (55.6% gaps) and the presence of a conserved A-C-x-U-A-x-U-A-x-U-A-A consensus sequence. On average, chloroplast tRNAs in the examined monocot plants contain 76 nucleotides. tRNA^Cys, tRNA^Asn, tRNA^Ala, tRNA^Asp, tRNA^Phe, and tRNA^Trp were found to contain 71, 72, 73, 74, 73, and 74 nucleotides, respectively. All of the sequences of the tRNA^Leu(UAG) and tRNA^Ser were found to contain 80 nt or more. tRNA^Lys was found to be absent from the chloroplast genome of O. sativa and S. bicolor (Table 1). Additionally, tRNA^Ala and tRNA^Ile were also found to be absent in S. bicolor (Table 1).

Chloroplast tRNAs of monocot plant encodes 21–28 anti-codons only
The chloroplast genomes of the investigated monocot plants, however, were found to encode only 21–28 anti-codons (Table 2). The chloroplast genome of O. nivara, O. sativa, S. officinarum, S. bicolor, T. aestivum, and Z. mays encoded 28, 25, 28, 21, 28, and 28 anti-codons, respectively (Table 2). The most common anti-codons found in the tRNA of chloroplast genome were UGC (tRNA^Ala), GCC (tRNA^Gly), UCC (tRNA^Gly), UGG (tRNA^Pro), GGU (tRNA^Thr), UGU (tRNA^Ile), GAC (tRNA^Asp), UAC
(tRNA\textsuperscript{Val}), GGA (tRNA\textsuperscript{Ser}), UGA (tRNA\textsuperscript{Ser}), GCU (tRNA\textsuperscript{Ser}), ACG (tRNA\textsuperscript{Arg}), UCU (tRNA\textsuperscript{Arg}), UAG (tRNA\textsuperscript{Leu}), CAA (tRNA\textsuperscript{Leu}), UUA (tRNA\textsuperscript{Leu}), GAA (tRNA\textsuperscript{Phe}), GUU (tRNA\textsuperscript{Ile}), UUU (tRNA\textsuperscript{Lys}), GUC (tRNA\textsuperscript{Asp}), UUC (tRNA\textsuperscript{Glu}), GUG (tRNA\textsuperscript{His}), UUG (tRNA\textsuperscript{Gln}), CAU (tRNA\textsuperscript{Ile}), GAU (tRNA\textsuperscript{Ile}), CUA (tRNA\textsuperscript{Met}), GUA (tRNA\textsuperscript{Tyr}), GCA (tRNA\textsuperscript{Cys}), and CCA (tRNA\textsuperscript{Trp})(Table 2). The UCC (tRNA\textsuperscript{Gly}), and UAC (tRNA\textsuperscript{Val}) anti-codons present in the genome of O. nivara were missing in the chloroplast genome of the related species, O. sativa (Table 2). Similarly, the anti-codons UCC (tRNA\textsuperscript{Gly}), and UAC (tRNA\textsuperscript{Val}) present in the genome of O. nivara, S. officinarum, T. aestivum, and Z. mays were found to be absent in the genome of S. bicolor (Table 2). In addition, the anti-codons GGU (tRNA\textsuperscript{Thr}) and UAA (tRNA\textsuperscript{Leu}) were also not present in S. bicolor; whereas, they were found in O. nivara, O. sativa, S. officinarum, T. aestivum and Z. mays. Outside of the above mentioned 28 anti-codons, the rest of the 33 anti-codons were not found in any of the tRNAs of the investigated monocot chloroplast genomes (Table 2).

**Conservation of chloroplast tRNA sequences is family specific**

Multiple sequence alignment analysis of all 20 tRNA gene family members of studied monocot species revealed small, highly conserved consensus sequences in...
the pseudouridine (Ψ) loop, but not in the other parts of the tRNA (Table 3). The Ψ-loop was found to possess a conserved U-U-C-x-A consensus nucleotide sequence (Table 3). The majority of the tRNAs contained a G nucleotide at the first position. tRNA Val, tRNA Met, and tRNA Pro, however, were found to possess an A nucleotide at the first position instead of a G (Table 3). tRNA Glu and tRNA Asn were found to possess a U nucleotide at the first position in the acceptor arm. Although no consensus sequence conservation was observed in the 5′-acceptor arm, the D-arm contained a conserved C nucleotide at the 4th position of the arm (13th position of the tRNA). In contrast, tRNA Glu, tRNA Gly, tRNA Met, tRNA Ser, tRNA Tyr, and all other tRNAs, possessed a C nucleotide at the 4th position of the D-arm. Nucleotide 7 to 16 of the canonical tRNA forms an A box, which has been reported to contain two conserved consensus sequences, 7GUGGCNNAGU16 and -G GU-AGNGC15 (– stands for gap & N stands for any nucleotide) [39]. Our analysis revealed that among the 20 tRNAs analyzed, only six of them possess a conserved G nucleotide at the 7th position (Table 3). The 7th position of the tRNA is instead occupied by an A, U, or C nucleotide (Table 3). The 14th position (1st nucleotide of D-loop) was found to be conserved in the majority of tRNA. Except for tRNA Arg, tRNA Asn, tRNA Gly, and tRNA Met, all other tRNAs were found to contain a conserved A nucleotide at the 14th position (Table 3). Similarly, the last nucleotide of the D-loop was found to be a conserved A nucleotide except tRNA Tyr (Table 3). The consensus sequence 52GGUUCGANUCC62, which starts from the 52nd position and ends at the 62nd position of tRNA, forms a B box [40]. Our analysis indicates that the conservation of box A and B nucleotide sequences in tRNA occurs in a family-specific manner. The G-G nucleotide at the 52nd and 53rd position was found to be conserved in the majority of tRNAs, except for tRNA Gin, tRNA Lys, and tRNA Val; whereas, the nucleotide sequence U-U-C-A-U-U was found to be conserved at the 54th, 55th, 56th, 58th, and 60th positions (Table 3). tRNA Met was found to contain a conserved U-U-C-A-U consensus sequence at the 54th, 55th, 56th, 58th, 59th, and 60th positions, instead of the U-U-C-A-U-U consensus sequence (Table 3). Similarly, tRNA Asp had a conserved U-U-C-G-A-G-C consensus sequence, while tRNA Val contained U-U-C-G-A-x-x conserved nucleotides. No conserved nucleotides were found at the 59th and 60th positions of tRNA Val. The anti-codon loop at the 32nd and 33rd positions were found to contain conserved C-U or U-U nucleotides. tRNA Gln, tRNA Pro, and tRNA Val contained conserved U-U nucleotides instead of the C-U nucleotides. In addition, in the majority of cases, the anti-codon loop at the 38th position had a conserved A nucleotide. tRNA Gln, tRNA Pro, and tRNA Val, however, possessed a conserved U nucleotide at the 38th position instead of nucleotide A (Table 3). The chloroplast genome encodes a predefined C-C-A tail in the gene of the tRNA. When the tRNA gene is transcribed, a C-C-A tail is included. The present study found that tRNA Ala, tRNA Arg, tRNA Leu, tRNA Lys, and tRNA Tyr contain C-C-A nucleotides in their 3′-end. A few of the encoded tRNA genes in the monocot chloroplast genomes also contain C-C-A tail in the 3′-end, however, the remaining tRNAs do not possess a C-C-A consensus sequence at their 3′-end.

### Nucleotide variation in the arms and loops of tRNA

In the present study, the acceptor arm of chloroplast tRNA was revealed to contain 1–7 nucleotides. Among the 213 tRNA sequences representing six species of monocot plants, only two were found to contain one nucleotide, one had five nucleotides, and one contained six nucleotides; while the rest of the 209 (98.12%) tRNAs had seven nucleotides. The D-arm was found to contain

---

**Table 1** Distribution of tRNA isotypes in the chloroplast genome of the monocot plants

| tRNA isotypes | O. nivara | O. sativa | S. officinarum | S. bicolor | T. aestivum | Z. mays |
|---------------|-----------|-----------|----------------|------------|-------------|---------|
| Alanine       | 2         | 2         | 2              | 0          | 2           | 2       |
| Glycine       | 2         | 1         | 2              | 1          | 2           | 2       |
| Proline       | 1         | 1         | 1              | 1          | 1           | 1       |
| Threonine     | 2         | 2         | 1              | 2          | 3           |         |
| Valine        | 3         | 2         | 3              | 2          | 3           | 3       |
| Serine        | 3         | 3         | 3              | 3          | 3           | 3       |
| Arginine      | 3         | 3         | 3              | 3          | 3           | 3       |
| Leucine       | 4         | 4         | 4              | 3          | 4           | 4       |
| Phenylalanine | 1         | 1         | 1              | 1          | 1           | 1       |
| Asparagine    | 2         | 2         | 1              | 2          | 2           | 2       |
| Lysine        | 1         | 0         | 1              | 0          | 1           | 1       |
| Aspartate     | 1         | 1         | 1              | 1          | 1           | 1       |
| Glutamate     | 1         | 1         | 1              | 1          | 1           | 1       |
| Histidine     | 2         | 2         | 2              | 2          | 2           | 2       |
| Glutamine     | 1         | 1         | 1              | 1          | 1           | 1       |
| Isoleucine    | 4         | 4         | 4              | 4          | 4           | 4       |
| Methionine    | 2         | 2         | 2              | 2          | 3           | 2       |
| Tyrosine      | 1         | 1         | 1              | 1          | 1           | 1       |
| Cysteine      | 1         | 1         | 1              | 1          | 1           | 1       |
| Tryptophan    | 1         | 1         | 1              | 1          | 1           | 1       |
| Selenocysteine| 0         | 0         | 0              | 0          | 0           | 0       |
| Suppressor    | 0         | 0         | 0              | 0          | 0           | 0       |
| Total         | 38        | 35        | 37             | 31         | 39          | 39      |
Table 2 Distribution of anti-codons in the chloroplast genome of the monocot plants

| tRNA Isotypes    | isoacceptors |
|------------------|--------------|
| *Oryza nivara* (29) |              |
| Alanine          | AGC          | GGC          | CGC          | UGC (2)      |
| Glycine          | ACC          | GCC (1)      | CCC          | UCC (1)      |
| Proline          | AGG          | GGG          | CGG          | UGG (1)      |
| Threonine        | AGU          | GGU (1)      | CGU          | UGU (1)      |
| Valine           | AAC          | GAC (2)      | CAC          | UAC (1)      |
| Serine           | AGA          | GGA (1)      | CGA          | UGA (1)      | ACU          | GCU (1)      |
| Arginine         | ACG (2)      | GCG          | CCG          | UCG          | CCU          | UCU (1)      |
| Leucine          | AAG          | GAG          | CAG          | UAG (1)      | CAA (2)      | UAA (1)      |
| Phenylnalanine   | AAA          | GAA (1)      |              |              |              |              |
| Asparagine       | AUU          | GUU (2)      |              |              |              |              |
| Lysine           | CUU          | UUU (1)      |              |              |              |              |
| Aspartate        | AUC          | GUC (1)      |              |              |              |              |
| Glutamate        | CUC          | UUC (1)      |              |              |              |              |
| Histidine        | AUG          | GUG (2)      |              |              |              |              |
| Glutamine        | CUG          | UUG (1)      |              |              |              |              |
| Isoleucine       | AAU          | GAU (2)      | CAU (2)      |              |              |              |
| Methionine       | CAU (4)      |              |              |              |              |              |
| Tyrosine         | AUA          | GUA (1)      |              |              |              |              |
| Cysteine         | ACA          | GCA (1)      |              |              |              |              |
| Tryptophan       | CCA (1)      |              |              |              |              |              |
| Suppressor       | CUA          | UUA          |              |              |              |              |
| Selenocysteine   | UCA          |              |              |              |              |              |
| *Oryza sativa* (26) |              |
| Alanine          | AGC          | GGC          | CGC          | UGC (2)      |
| Glycine          | ACC          | GCC (1)      | CCC          | UCC          |
| Proline          | AGG          | GGG          | CGG          | UGG (1)      |
| Threonine        | AGU          | GGU (1)      | CGU          | UGU (1)      |
| Valine           | AAC          | GAC (2)      | CAC          | UAC          |
| Serine           | AGA          | GGA (1)      | CGA          | UGA (1)      | ACU          | GCU (1)      |
| Arginine         | ACG (2)      | GCG          | CCG          | UCG          | CCU          | UCU (1)      |
| Leucine          | AAG          | GAG          | CAG          | UAG (1)      | CAA (2)      | UAA (1)      |
| Phenylnalanine   | AAA          | GAA (1)      |              |              |              |              |
| Asparagine       | AUU          | GUU (2)      |              |              |              |              |
| Lysine           | CUU          | UUU (1)      |              |              |              |              |
| Aspartate        | AUC          | GUC (1)      |              |              |              |              |
| Glutamate        | CUC          | UUC (1)      |              |              |              |              |
| Histidine        | AUG          | GUG (2)      |              |              |              |              |
| Glutamine        | CUG          | UUG (1)      |              |              |              |              |
| Isoleucine       | AAU          | GAU (2)      | CAU (2)      |              |              |              |
| Methionine       | CAU (4)      |              |              |              |              |              |
| Tyrosine         | AUA          | GUA (1)      |              |              |              |              |
| Cysteine         | ACA          | GCA (1)      |              |              |              |              |
| Tryptophan       | CCA (1)      |              |              |              |              |              |
Table 2 Distribution of anti-codons in the chloroplast genome of the monocot plants (Continued)

| tRNA Isotypes       | Isoacceptors |
|---------------------|--------------|
|                     |              |
| Suppressor          | CUA          |
| UUA                 |              |
| Selenocysteine      | UCA          |
| Saccharum officinarum (29) | | |
| Alanine             | AGC          |
| GGC                 |
| CGC                 |
| UGC (2)             |
| Glycine             | ACC          |
| GCC (1)             |
| CCC                 |
| UCC (1)             |
| Proline             | AGG          |
| GGG                 |
| CGG                 |
| UGG (1)             |
| Threonine           | AGU          |
| GGU (1)             |
| CGU                 |
| UGU (1)             |
| Valine              | AAC          |
| GAC (2)             |
| CAC                 |
| UAC (1)             |
| Serine              | AGA          |
| GGA (1)             |
| CGA                 |
| UGA (1)             |
| ACU                 |
| GCU (1)             |
| Arginine            | ACG (2)      |
| GCG                 |
| CCG                 |
| UCG                 |
| CCU                 |
| UCU (1)             |
| Leucine             | AAG          |
| GAG                 |
| CAG                 |
| UAG (1)             |
| CAA (2)             |
| UAA (1)             |
| Phenylalanine       | AAA          |
| GAA (1)             |
| Asparagine          | AUU          |
| GUU (2)             |
| Lysine              | CUU          |
| UUU (1)             |
| Aspartate           | AUC          |
| GUC (1)             |
| Glutamate           | CUC          |
| UUC (1)             |
| Histidine           | AUG          |
| GUG (2)             |
| Glutamine           | CUG          |
| UUG (1)             |
| Isoleucine          | AAU          |
| GAU (2)             |
| CAU (2)             |
| Methionine          | CAU (4)      |
| Tyrosine            | AUA          |
| GUA (1)             |
| Cysteine            | ACA          |
| GCA (1)             |
| Tryptophan          | CCA (1)      |
| Suppressor          | CUA          |
| UUA                 |
| Selenocysteine      | UCA          |
| Sorghum bicolor (23) |             |
| Alanine             | AGC          |
| GGC                 |
| CGC                 |
| UGC                 |
| Glycine             | ACC          |
| GCC (1)             |
| CCC                 |
| UCC                 |
| Proline             | AGG          |
| GGG                 |
| CGG                 |
| UGG (1)             |
| Threonine           | AGU          |
| GGU                 |
| CGU                 |
| UGU (1)             |
| Valine              | AAC          |
| GAC (2)             |
| CAC                 |
| UAC                 |
| Serine              | AGA          |
| GGA (1)             |
| CGA                 |
| UGA (1)             |
| ACU                 |
| GCU (1)             |
| Arginine            | ACG (2)      |
| GCG                 |
| CCG                 |
| UCG                 |
| CCU                 |
| UCU (1)             |
| Leucine             | AAG          |
| GAG                 |
| CAG                 |
| UAG (1)             |
| CAA (2)             |
| UAA                 |
| Phenylalanine       | AAA          |
| GAA (1)             |
| Asparagine          | AUU          |
| GUU (2)             |
| Lysine              | CUU          |
| UUU                 |
| Aspartate           | AUC          |
| GUC (1)             |
| Glutamate           | CUC          |
| UUC (1)             |
| Histidine           | AUG          |
| GUG (2)             |
| Glutamine           | CUG          |
| UUG (1)             |
| Isoleucine          | AAU          |
| GAU (2)             |
| CAU (2)             |
| Methionine          | CAU (4)      |
| Tyrosine            | AUA          |
| GUA (1)             |
### Table 2 Distribution of anti-codons in the chloroplast genome of the monocot plants (Continued)

| tRNA Isotypes | Isoacceptors |
|---------------|--------------|
| Cysteine      | ACA GCA (1)  |
| Tryptophan    | CCA (1)      |
| Suppressor    | CUA UUA      |
| Selenocysteine| UCA          |

**Triticum aestivum (28)**

| Alanine      | AGC GGC CGC UGC (2) |
| Glycine      | ACC GCC (1) CCC UCC (1) |
| Proline      | AGG GGG CGG UGG (1) |
| Threonine    | AGU GGU (1) CGU UGU (1) |
| Valine       | AAC GAC (2) CAC UAC (1) |
| Serine       | AGA GGA (1) CGA UGA (1) ACU GCU (1) |
| Arginine     | ACG (2) GCG CCG UCG CCU UCU (1) |
| Leucine      | AAG GAG CAG UAG (1) CAA (2) UAA (1) |
| Phenylalanine| AAA GAA (1) |
| Asparagine   | AUU GUU (2) |
| Lysine       | CUI UUU (1) |
| Aspartate    | AUC GUC (1) |
| Glutamate    | CUC UUC (1) |
| Histidine    | AUG GUG (1) |
| Glutamine    | CUG UUG |
| Isoleucine   | AAU GAU (2) CAU (2) |
| Methionine   | CAU (5) |
| Tyrosine     | AUA GUA (1) |
| Cysteine     | ACA GCA (1) |
| Tryptophan   | CCA (1) |
| Suppressor   | CUA UUA |
| Selenocysteine| UCA |

**Zea mays (29)**

| Alanine      | AGC GGC CGC UGC (2) |
| Glycine      | ACC GCC (1) CCC UCC (1) |
| Proline      | AGG GGG CGG UGG (1) |
| Threonine    | AGU GGU (2) CGU UGU (1) |
| Valine       | AAC GAC (2) CAC UAC (1) |
| Serine       | AGA GGA (1) CGA UGA (1) ACU GCU (1) |
| Arginine     | ACG (2) GCG CCG UCG CCU UCU (1) |
| Leucine      | AAG GAG CAG UAG (1) CAA (2) UAA (1) |
| Phenylalanine| AAA GAA (1) |
| Asparagine   | AUU GUU (2) |
| Lysine       | CUI UUU (1) |
| Aspartate    | AUC GUC (1) |
| Glutamate    | CUC UUC (1) |
| Histidine    | AUG GUG (2) |
| Glutamine    | CUG UUG (1) |
| Isoleucine   | AAU GAU (2) CAU (2) |
3 and 4 nucleotides and none of the tRNAs possessed less than three or more than four nucleotides in the D-arm. A total of 73 (34.25%) were had three nucleotides, while 140 (65.73%) were contained four nucleotides. The D-loop, that forms a part of the A box, had seven to eleven nucleotides. Among the 213 tRNAs, 45 (21.12%) of the D-loops contain seven, 38 (17.84%) contain 8, 75 (35.21%) contain nine, 22 (10.32%) contain 10, and 33 (15.49%) contain 11 nucleotides. The anti-codon arm of the chloroplast tRNAs had 4–5 nucleotides. Among the 213 tRNAs, 23 (10.79%) of the anti-codon arms contain four nucleotides, while 190 (89.20%) contain five nucleotides (Additional file 1: Table S1). All of the tRNAs, except for one, had seven nucleotides in the anti-codon loop. tRNA 6160_OrniCt018 of O. nivara contained nine nucleotides instead of seven (Additional file 1: Table S1). The variable loop was found to possess a diverse number of nucleotides with different tRNAs having 4 (9.38%), 5 (59.62%), 6 (3.75%), 7 (5.63%), 11 (2.34%), 12 (0.46%), 13 (6.1%), 14 (0.46%), 15 (1.87%), 16 (2.34%), 18 (2.34%), or 19 (5.63%) nucleotides. None of the chloroplast tRNAs were found to possess 8, 9, 10, 17, 20 or more nucleotides in the variable loop (Additional file 1: Table S1). tRNA Leu, tRNA Ser, and tRNA Tyr had 10 or more nucleotides, respectively, whereas the other tRNAs possessed less than 10 nucleotides in the variable loop (Additional file 1: Table S1). Among the 213 examined tRNA sequences, only three tRNA Gly genes had four nucleotides in the Ψ-arm, while the remaining tRNA sequences had five nucleotides. Similarly, the Ψ-loop region in all of the 213 tRNAs possessed seven nucleotides. Our study found 7 bp in the acceptor arm and 3–4 bp in the D-arm and considerable variation was observed in the other parts. The anti-codon arm was found to possess 4–5 bp, and the anti-codon loop 7 or 9 nucleotides. The number of

Table 2 Distribution of anti-codons in the chloroplast genome of the monocot plants (Continued)

| tRNA Isotypes | Isoacceptors |
|---------------|--------------|
| Methionine    | CAU (2)      |
| Tyrosine      | AUA          |
| Cysteine      | ACA          |
| Tryptophan    | CCA (1)      |
| Suppressor    | CUA          |
| Selenocysteine| UCA          |

Table 3 Multiple sequence alignment and the presence of isotype specific conserved nucleotide consensus sequence in chloroplast tRNA of monocot plant. Asterisks indicates no conserved sequence

| tRNA Isotypes | AC arm | D-arm | D-loop | ANC arm | ANC loop | Variable region | Ψ-arm | Ψ-loop |
|---------------|--------|-------|--------|---------|----------|-----------------|-------|--------|
| Alanine       | G-G-G-G-A-U-A | G-C-U-C | A-U-U-G-G-U-A | C-C-G-C-U | C-U-U-G-C-A | A-U-G-U-C | A-G-G-G | U-C-G-A-U-U |
| Arginine      | G-x-G    | G-x-U-C | G-x-G   |         |          |                 | G-x-G | U-C-x-A-U-U |
| Asparagine    | U-C-C-U-C-G | G-U-U-G-U-A |         |         |          |                 | G-U-G-C | U-G-U-A-U-U |
| Aspartate     | G-G-G-A-U-U-G | A-U-U-G-G-U-C | G-C-G-C-C | C-U-G-A-A-A | A-G-C-A-C | G-C-G-G-G | U-G-G-G |
| Cysteine      | G-G-G-G-C-G-A | G-C-U-C | A-A-G-x-G-G-U-A-A | G-G-G-G-A | G-C-U-A-A-A | U-A-x-C | C-C-A-G-G |
| Glutamate     | G-C-C-C-C-A | G-C-U-C | A-G-U-G-G-U-U-C-A | U-C-U-C | C-C-U-U-C-A | G-A-G-C | G-G-G-A-U-U |
| Glutamine     | U-G-G-G-G | G-C-U-C | A-G-U-G-U-A-A | U-U-G-U-G-U | U-A-U-C | G-G-G-G-G | G-G-G-U-U |
| Glycine       | G-C-G-A-G-G-G | G-U-U | C-x-U-G-G-U-A-A | U-C-U-C | U-G-U-C-A-A-A | A-G-A-U-A | C-C-G-G-G |
| Histidine     | G-G-C-G-G-U-G-C | G-C-U-C | A-A-G-U-G-G-U-C-A-A | G-U-G-G-A | U-G-U-G-G-A | C-A-G-G | U-G-G-G-U-U-U |
| Isoleucine    | G-G-G-C-U-A-U | G-U-U-G | A-G-U-G-G-U-U | G-C-C-C-C | C-U-G-A-U-A | A-G-G-G-G | U-C-U-U-G-U |
| Leucine       | G-x-G    | G-x-G | G-A-U-x-U-A-G-A |       |          |                 | C-U-x-A-x-A-A | U-x-A-U-G-G-U-U-A |
| Lysine        | G-G-G-U-G-C-U | A-U-U-C | A-A-G-U-G-G-U-A-A | U-C-G-G | C-U-U-U-U-A-A | C-U-A | U-C-G-A-G-U-G-U |
| Methionine    | G-C      | U-x-G-U-x-U-A | U-x-G-U-x-U-A-A | G-x-x-U-C-A-A | C-U-U-A-U-A-A | C-G-G-G | U-C-G-A-U-U |
| Phenylalanine | G-U-A-A-G-A | G-C-U-C | A-G-U-G-G-U-U-A-G | G-A-G-G-A | G-U-G-G-G | C-U-A-A-A-A-A | A-C-A-G-G-G |
| Proline       | A-G-G-A-G-G-C-G | A-G-C-U-G-U-U-G | I-U-U-U-G | I-U-U-G-G-G-U | A-G-U-G-G-G | A-C-A-G-G | U-G-G-G-G |
| Serine        | G-G-A-G-G-A-G-C-G | G-x-x-G | A-G-x-G-x-G-x-G-U-C | G-C-U-G-U-U-C-A-A | U-x-U-A-A-A | A-G-A-G-G-A-G-G | U-G-G-G-G |
| Threonine     | G-C-C-C-x-x-U-U-C | C-U-C | A-G-x-G-G-x-x-U-A-U | G-C | U-x-U-G-U-A-A | G-U-C | A-G-G-G-G-G-G-G |
| Tryptophan    | G-C-G-C-U-C-U | G-U-U-G | A-U-U-x-G-G-U-U-A | G-G-G-G | G-C-U-A-A-A-A-A | U-G-A-G-G | U-G-G-G-G-G-G |
| Tyrosine      | G-G-G-G-U-G-x-G-A | C-C-G-G | A-G-x-G-G-U-U-A-U | A-C-G-G-A | G-U-G-G-G-A-A | U-G-A-G-G-G | G-C-U-G-G-G |
| Valine        | A-G-G-G-x-U-A | C-U-C | A-G-x-x-G-G-U-U-A | C-C-C | C-U-U-A-A-A-A | A-G-G-G-G | U-C-G-A-x-x-x |

Mohanta et al. BMC Plant Biology (2019) 19:39 Page 8 of 24
nucleotides making up the variable loop ranged from 4 to 19 and none of the tRNAs had more than 19 nucleotides in the variable loop. Similar to the previous report, the ψ-arm possessed 4–5 nucleotides.

**Chloroplast tRNA contain group I intron**

In our study, however, chloroplast tRNA was found to contain introns. tRNA▼ of *T. aestivum* (4982_TraeCt095) was found to contain a group I intron located in the anti-codon loop region of tRNA▼ (Fig. 2). The intron was 84 nucleotides in length and began at nucleotide 37 and ended at nucleotide 120 of the tRNA▼ gene. The group I introns of chloroplast tRNA contain conserved U-U-x2-C and A-G-x2-U consensus sequences (Fig. 3). A phylogenetic tree was constructed to elucidate the evolution of the group I intron. The phylogenetic analysis indicated that the group I intron of chloroplast tRNA grouped with the group I intron of cyanobacteria (Fig. 4).

**Chloroplast tRNA encodes putative novel tRNAs**

In the present study, a few putative novel tRNAs were found to be encoded by the chloroplast genome (Fig. 5). tRNA▼ of *O. nivara* (6129_OrniCt007, ΔG = −18.10), and tRNA▼ of *S. bicolor* (20,385_trnM-CAU SobiCt011, ΔG = 14.7) did not contain an acceptor arm at the 5′-end (Fig. 5). Additionally, a few tRNA▼ in *O. nivara* (6152_OrniCt014, ΔG = −34.13), *O. sativa* (3720_OrsajCt137, ΔG = −34.13), *S. bicolor* (20,407_trnS-GGA SobiCt019, ΔG = −34.13), *S. officinarum* (6593), and *T. aestivum* (5020_TraeCt112, ΔG = −34.13) were found to contain a seven-nucleotide loop structure in the variable loop region, similar to the anti-codon loop of tRNA (Fig. 6). All of the loop structures comprising the variable loop region were found to be composed of A-C-U-U-U-U-G nucleotides. The tRNA▼ of *O. nivara* (6160_OrniCt018, ΔG = −25.20) was found to contain only four nucleotides in the anti-codon arm and nine nucleotides in the anti-codon loop (Fig. 7). Many similar tRNA structures have been found in the genomic tRNA of cyanobacteria, as well as plants (unpublished data).

**C-A-U anti-codon codes for tRNA▼ in chloroplast tRNAs**

The C-A-U anti-codon is a characteristic feature of tRNA▼ and has only one iso-acceptor. In addition to the presence of a C-A-U anti-codon in tRNA▼, we also found that the tRNA▼ of chloroplast tRNA also encodes a C-A-U anti-codon. The tRNA▼ in *O. nivara* (6206_OrniCp049, 6270_OrniCt035), *O. sativa* (3774_OrsajCt146, 3828_OrsajCt160), *S. officinarum* (officinarum_6644, officinarum_6710), *S. bicolor* (20,460, 20,502), *T. aestivum* (5069, 5108), and *Z. mays* (2069_trnI

![Fig. 2 Presence of group I intron in chloroplast tRNA. The intron was found to locate in the anti-codon loop of the tRNA](image-url)
ZemaCt144, 2131_trnI ZemaCt154) chloroplast genomes encode a C-A-U anti-codon. To our knowledge, this may be the first report to document the presence of a C-A-U anti-codon in chloroplast tRNA Ile.

**Chloroplast tRNAs have evolved from multiple common ancestors**

A phylogenetic tree was constructed using the tRNA sequences in the chloroplast genomes of all of the examined monocot plants. A phylogenetic analysis revealed the presence of two major clusters that consist of 30 groups. Cluster I contain tRNA Val, tRNA Ala, tRNA Arg, tRNA Thr, tRNA Met, tRNA Asp, tRNA Lys, tRNA Ile, tRNA Leu, tRNA Ser, tRNA Pro, tRNA Gin, tRNA His, tRNA Gly, tRNA Glu, and tRNA Arg. Cluster II contains tRNA Phe, tRNA Cys, tRNA Ile, tRNA Met, tRNA Tyr, tRNA Asn, tRNA Arg, tRNA Trp, and tRNA Leu. There are 21 groups in cluster I and 9 groups in cluster II (Fig. 8). In cluster I, tRNA Arg is grouped twice; once with tRNA Ala and once near to tRNA Met. Similarly, tRNA Met is also grouped twice; once near to the group containing tRNA Arg and once near the group containing tRNA Thr (Fig. 8). tRNA Asp, tRNA Ile, tRNA Leu, and tRNA Met present in cluster I are also found in cluster II of the phylogenetic tree. The tRNAs with the anti-codon G-A-C and U-A-C of tRNA Val, G-G-U and U-G-U of tRNA Thr, U-G-A, G-C-U, and G-G-A of tRNA Ser, G-C-C and U-C-C of tRNA Gly, U-A-A, U-A-G, and C-A-A of tRNA Leu, C-A-U of tRNA Ile, U-G-C, U-C-U, and A-C-G of tRNA Arg, all grouped separately (Fig. 8). tRNA Thr (CCA) is closely grouped with tRNA Arg (UCU) in cluster II, suggesting the evolution of tRNA Thr from tRNA Arg (Fig. 8). Similarly, tRNA Tyr (GUA) is closely grouped with tRNA Met (CAU) and tRNA Asp (CAU), suggesting the evolution of tRNA Tyr (GUA) and tRNA Asp (CAU) from tRNA Met (CAU). The grouping of tRNA Met (CAU) with tRNA Asp (CAU), and their similar anti-codon nucleotides, strongly suggests that tRNA Ile evolved directly from tRNA Met. In addition, the close grouping of tRNA Met (CAU) with tRNA Arg (ACG) further suggests that tRNA Arg has evolved from tRNA Met as well. The grouping of tRNA Glu (UUC) with tRNA Gly (GCC), tRNA His (GUG) with tRNA Gin (UUG), and tRNA Pro (UGG) suggests that these tRNAs may have evolved from a common ancestor or by a gene duplication event. tRNA Ser (GGA, GCU, UGA) grouped with tRNA Leu (UAA); which suggests that tRNA Ser evolved from tRNA Leu. Notably, tRNA Leu contains a C-A-A anti-codon, while tRNA Ser, which grouped with tRNA Ser, contains a U-A-A anti-codon. This suggests that tRNA Leu (CAA) has undergone a base substitution to give rise to tRNA Leu (UAA) and that further duplication and diversification resulted in tRNA Ser (GGA, GCU, UGA). The grouping of tRNA Ile (GAU), tRNA Cys (UUC), and tRNA Asp (GUC) together suggests their common evolutionary lineage. Further, grouping of tRNA Met with tRNA Thr (UGU and GGU) suggests that tRNA Thr (UGU and GGU) evolved from tRNA Met. Similarly, the close phylogenetic relationship of tRNA Met with tRNA Ala and tRNA Val in cluster I indicates that tRNA Ala and tRNA Val also evolved from tRNA Met. A disparity index test of substitution pattern homogeneity was conducted using Monte Carlo replications to determine if all of the substitutions and the rate of substitution of the nucleotides are
homogenous. Results indicated that the null hypothesis was rejected for tRNAArg, tRNAAsp, tRNAHis, tRNAIle, and tRNAPro; suggesting that the rate of substitution of nucleotides in these groups is homogenous. Outside of these six tRNA isotypes, 14 did not show pattern homogeneity, and hence, the substitution of nucleotides and evolution of tRNAArg, tRNAAsp, tRNAHis, tRNAIle, tRNAPhe, tRNAThr, tRNATyr, tRNAVal, and tRNAVal are not homogenous. To better understand the relationship of chloroplast tRNAs with the Archaea, we incorporated tRNA two Archaea species and the tRNA sequences of three cyanobacterial species were used as ingroups. The complementary DNA sequences of two Arabidopsis thaliana NAC transcription factors (AtNAC1 and AtNAC2) were used as out groups (Additional file 2: Figure S1). A phylogenetic analysis showed some overlapping relationship of Archaea tRNAs with the chloroplast tRNA. However, chloroplast tRNAs were much closer to cyanobacterial tRNA compared to the Archaea.

The rate of transition and transversion is isoacceptor specific

tRNAs are evolutionarily conserved molecules and the possibility of undergoing major transition or transversion events is very minimum. The rate of transition (8.33) and transversion (8.34) of tRNAAla, tRNAAsn, tRNAAsp, tRNAHis, tRNAPhe, and tRNAPro are almost equal. This indicates that, although the rate of transversion is slightly higher than the rate of transition, these tRNAs have evolved at almost an equal rate with respect to transition and transversion (Table 4). Additionally, the rate of transition (25.00) and transversion (0.00) of tRNAArg, tRNAAsn, tRNAAsp, tRNAHis, tRNAPhe, tRNAThr, tRNATyr, tRNAVal, and tRNAVal are not homogenous. Notably, however, tRNAArg, tRNAAsn, tRNAAsp, tRNAHis, tRNAPhe, tRNAThr, tRNATyr, tRNAVal, and tRNAVal in the chloroplast genome of monocot plants have undergone a high rate of transition but have not undergone any transversion. In contrast, the rate of transversion in tRNAArg (8.60), tRNAAsn (10.09), tRNAAsp (9.15), was found to be higher relative to the rate of transition for tRNAArg (7.80), tRNAAsn (4.82), and tRNAAsp (6.70), respectively (Table 4). A higher transition rate was also observed in tRNAArg (12.40), tRNAAsn (12.53), tRNAAsp (17.39), tRNAHis (11.88), tRNAMet (16.87), tRNAThr, and tRNAVal (Table 4). The highest rate of transition substitutions (25.00) was found in tRNAArg, tRNAAsn, tRNAAsp, tRNAHis, tRNAPhe, tRNAThr, tRNATyr, and tRNAVal. When all of the tRNAs are collectively examined, however, the average rate of transition (14.71) is greater than the average rate of transversion (5.15) (Table 4).

Duplication of chloroplast tRNA precedes over deletion

Plant genomes contain a greater abundance of duplicated genes and whole genome duplication events have occurred multiple times over the past 200 million years [41–44]. Given the cyanobacterial origin of the chloroplast...
genome, the rate of duplication and loss events could be different from genes within the nuclear-encoded genome. In the present study, duplication/loss analyses of chloroplast tRNA in monocot plants revealed that 101 genes experienced a duplication event and that 139 genes underwent losses; whereas, 80 genes underwent conditional duplication. The majority of chloroplast tRNAs underwent losses during the course of evolution. Although all of the tRNAs descended from the same lineage (monocot), the loss of genes was still greater than the duplicated genes (Fig. 9).

**Discussion**

tRNAs are conserved family genes responsible for conducting protein translation event. Their presence in the chloroplast genome is supplementary to the genome to make it semi-autonomous. Multiple sequence alignment of chloroplast tRNAs revealed several basic conserved genomic features. A few tRNAs were found to contain extended nucleotide sequences at the 5′-end. However, the tRNAscan-SE server was not able to confirm if these nucleotide sequences of the 5′-end were introns. As a result, it is highly possible that these sequences can be introns of the tRNAs. A previous study reported the presence of a group I intron in cyanobacterial tRNA [45]. Given the origin of the chloroplast genome from a cyanobacterial lineage, it is reasonable to consider that these sequences are most likely introns of the chloroplast tRNAs [45]. Analysis of each tRNA sequence revealed tRNA^{Leu} and tRNA^{Ser} encoded for longest tRNA sequences. A previous study also reported the presence of 80 or more nucleotides in tRNA^{Leu} and tRNA^{Ser} of *Oryza sativa* [45]. This indicates that tRNA^{Leu} and tRNA^{Ser} encode longer tRNA sequences as compared to the others. This study also revealed the absence of tRNA^{Lys}, tRNA^{Ala}, and tRNA^{Ile} genes in the chloroplast genome of these monocot plants. The absence of important tRNA encoding genes in the chloroplast genome
is quite intriguing and makes it important to understand how protein translation in these monocot plants is conducted in the absence of important tRNAs. Most likely, genomic tRNA compensate for the absence of plastidal tRNAs or it might be possible that other tRNAs from the organellar genome perform multiple functions to conduct protein translation. This is the first report regarding the absence of tRNA\textsuperscript{Lys}, tRNA\textsuperscript{Ala}, and tRNA\textsuperscript{Ile} in the chloroplast genome. In addition to the absence of tRNA\textsuperscript{Lys}, tRNA\textsuperscript{Ala}, and tRNA\textsuperscript{Ile}, the chloroplast genome of monocot plants also lacks selenocystein, pyrrolysine and suppressor tRNA (Table 1). Our analysis also revealed that the monocot chloroplast genome contains the highest number genes encoding tRNA\textsuperscript{Leu} and tRNA\textsuperscript{Met}; (4) followed by tRNA\textsuperscript{Arg}, and tRNA\textsuperscript{Ser} (3). The universal genetic table contains 64 codons; of which, 61 are sense and 3 are anti-sense codons. Therefore, it is possible that there will be tRNAs with 61 unique anti-codons to code for 61 sense codons. Approximately 33 anti-codons were found to be absent from the tRNAs of chloroplast genome. However, the absence of UCC anti-codons of tRNA\textsuperscript{Gly} is compensated by the presence of GCC anti-codons of tRNA\textsuperscript{Gly}, whereas the absence of anti-codon UAC of tRNA\textsuperscript{Val} is compensated by the presence of GAC anti-codons of tRNA\textsuperscript{Val}. Similarly, the anti-codon GGU of tRNA\textsuperscript{Thr} is compensated by the presence of the UGG anti-codon of tRNA\textsuperscript{Thr} and the anti-codon UAA of tRNA\textsuperscript{Leu} is compensated by the presence of anti-codon UAG and CAA. The complete absence of a tRNA gene for tRNA\textsuperscript{Lys} (UUU, CUU) in O. sativa and S. bicolor, and tRNA\textsuperscript{Ala} (AGC, GGC, CGC, and UGC) is difficult to understand. Nevertheless, it can be speculated that the deficiency created by the absence of these tRNAs in the chloroplast genome might be compensated by genomic tRNAs or other tRNAs of chloroplast or nuclear origin. The anti-codon CAU is encoded by tRNA\textsuperscript{Met} and tRNA\textsuperscript{fMet}. Our analysis indicated that chloroplast genome of the investigated monocot plants encodes tRNA\textsuperscript{Met} and tRNA\textsuperscript{fMet} as well. Previously, Howe (1985) and Hiratsuka et al., (1989) reported the presence of tRNA\textsuperscript{fMet} in chloroplast genome [46, 47]. All of the species were found to contain at least one tRNA\textsuperscript{Met} and one tRNA\textsuperscript{fMet}. O. nivara (6128_Orni Ct006), O. sativa (3694_OrsajCt127), S. officinarum (6569), S. bicolor (20382), T. aestivum (4994), and Z. mays (1994) each encode one tRNA\textsuperscript{Met}. In the prokaryotic genome, the initiation of protein translation is mediated by tRNA\textsuperscript{Met}, whereas subsequent addition of methionine to the polypeptide chain is mediated by tRNA\textsuperscript{Met} [48–50]. The presence of tRNA\textsuperscript{Met} and
tRNA\textsuperscript{Met} is a characteristic feature of prokaryotic and organellar genes [51] and the presence of tRNA\textsuperscript{Met} in the chloroplast genome of monocot plants suggests its prokaryotic origin.

tRNAs are an evolutionarily conserved multigene family due to their functional similarities across many species. The nucleotide composition of a tRNA is responsible for maintaining the tertiary structure of the translated tRNA. Thus, the common conserved functions of tRNA should also be reflected in conserved coding sequences. A previous study reported the presence of a conserved nucleotide consensus sequence in tRNAs which was confined to the Ψ-loop only [45]. In our study, we found the presence of U-U-C-x-A nucleotide consensus sequence in the Ψ-loop. However, no conserved consensus sequences were found in other parts of the tRNAs. Instead, they were found to contain some conserved nucleotides. The nuclear encoded tRNA\textsuperscript{Gln} and tRNA\textsuperscript{Asn} contain a U nucleotide at the first position (Table 3) [45]. However, a multiple sequence alignment study indicated that the sequence conservation present in chloroplast tRNAs is family specific (Table 3). During protein translation, polymerase binds with the promoter of the tRNA which is known as A and B box. These two boxes contain conserved consensus sequences. Box A starts at the +8 nucleotide of mature tRNA, whereas box B contains conserved 52GGUUCGANUC62 nucleotides consensus that constitutes a part of the Ψ-arm and whole Ψ-loop. Box A of chloroplast tRNA was not so conserved, whereas box B was highly conserved. Boxes A and B are considered to be the intragenic transcription promotor signal sequence for RNA polymerase III [52]. The signal sequence for transcription activation is not conserved in a universal manner in the tRNAs of the chloroplast genome. The anti-codon loop was reported to be conserved at the 32nd position [52]. However, in the present study, conservation of nucleotides was found at the 32nd and 33rd positions in the majority of cases. In addition, several tRNA sequences were found to contain 3'-C-C-A tail. The addition of a C-C-A tail to the 3’-end of a tRNA is facilitated by a tRNA nucleotidyltransferase. However, chloroplast genomes do not encode tRNA nucleotidyltransferases. Thus, adding a C-C-A tail to the 3’-end of the tRNA would be difficult in the absence of nucleotidyltransferases. The absence of a C-C-A tail at the 3’end of the few tRNAs reflect their recent evolution as the majority of nuclear tRNAs lacked a 3’ C-C-A tail.
| Alanine | U | C | G | Lysine | A | U | C | G |
|--------|---|---|---|--------|---|---|---|---|
| A      | 8.34 | 8.34 | 8.34 | 8.33 | A | 10.09 | 10.09 | 4.82 |
| U      | 8.34 | 8.33 | 8.34 | 8.34 | U | 10.09 | 4.82 | 10.09 |
| C      | 8.34 | 8.33 | 8.34 | 8.34 | C | 10.09 | 4.82 | 10.09 |
| G      | 8.33 | 8.34 | 8.34 | 8.34 | G | 4.82 | 10.09 | 10.09 |

| Arginine | Methionine | A | U | C | G | A | U | C | G |
|----------|------------|---|---|---|---|---|---|---|---|
| A        | 6.3 | 6.3 | 6.3 | 12.40 | A | 4.06 | 4.06 | 16.87 |
| U        | 6.30 | 12.40 | 6.30 | 4.06 | U | 16.87 | 4.06 |
| C        | 6.30 | 12.40 | 6.30 | 4.06 | C | 16.87 | 4.06 |
| G        | 12.40 | 6.30 | 6.30 | 4.06 | G | 16.87 | 4.06 | 4.06 |

| Asparagine | Phenylalanine | A | U | C | G | A | U | C | G |
|------------|---------------|---|---|---|---|---|---|---|---|
| A          | 8.34 | 8.34 | 8.34 | 8.33 | A | 8.34 | 8.34 | 8.33 |
| U          | 8.34 | 8.33 | 8.34 | 8.33 | U | 8.34 | 8.34 | 8.33 |
| C          | 8.34 | 8.33 | 8.34 | 8.34 | C | 8.34 | 8.33 | 8.34 |
| G          | 8.33 | 8.34 | 8.34 | 8.34 | G | 8.34 | 8.34 | 8.34 |

| Aspartate | Proline | A | U | C | G | A | U | C | G |
|-----------|---------|---|---|---|---|---|---|---|---|
| A         | 0.00 | 25.00 | 0.00 | 25.00 | A | 9.15 | 9.15 | 6.70 |
| U         | 0.00 | 25.00 | 0.00 | 25.00 | U | 6.70 | 9.15 |
| C         | 0.00 | 25.00 | 0.00 | 25.00 | C | 9.15 | 6.70 |
| G         | 25.00 | 0.00 | 0.00 | 25.00 | G | 9.15 | 9.15 |

| Cysteine | Serine | A | U | C | G | A | U | C | G |
|----------|--------|---|---|---|---|---|---|---|---|
| A        | 0.00 | 0.00 | 25.00 | 0.00 | A | 0.00 | 5.15 | 5.15 |
| U        | 0.00 | 25.00 | 0.00 | 25.00 | U | 5.15 | 14.70 |
| C        | 0.00 | 25.00 | 0.00 | 25.00 | C | 5.15 | 14.70 |
| G        | 25.00 | 0.00 | 0.00 | 25.00 | G | 5.15 | 5.15 |

| Glutamine | Threonine | A | U | C | G | A | U | C | G |
|-----------|----------|---|---|---|---|---|---|---|---|
| A        | 0.00 | 25.00 | 0.00 | 25.00 | A | 9.15 | 9.15 | 6.70 |
| U        | 0.00 | 25.00 | 0.00 | 25.00 | U | 6.70 | 9.15 |
| C        | 0.00 | 25.00 | 0.00 | 25.00 | C | 9.15 | 6.70 |
| G        | 25.00 | 0.00 | 0.00 | 25.00 | G | 9.15 | 9.15 |

| Glutamate | Tryptophan | A | U | C | G | A | U | C | G |
|-----------|------------|---|---|---|---|---|---|---|---|
| A        | 6.23 | 12.53 | 6.23 | 12.53 | A | 0.00 | 0.00 | 25.00 |
| U        | 6.23 | 12.53 | 6.23 | 12.53 | U | 25.00 | 0.00 |
| C        | 6.23 | 12.53 | 6.23 | 12.53 | C | 25.00 | 0.00 |
| G        | 12.53 | 6.23 | 6.23 | 12.53 | G | 25.00 | 0.00 |

| Glycine | Tyrosine | A | U | C | G | A | U | C | G |
|---------|----------|---|---|---|---|---|---|---|---|
| A        | 3.80 | 3.80 | 3.80 | 17.39 | A | 0.00 | 0.00 | 25.00 |
| U        | 3.80 | 17.39 | 3.80 | 17.39 | U | 25.00 | 0.00 |
| C        | 3.80 | 17.39 | 3.80 | 17.39 | C | 25.00 | 0.00 |
| G        | 17.39 | 3.80 | 3.80 | 17.39 | G | 25.00 | 0.00 |

| Histidine | Valine | A | U | C | G | A | U | C | G |
|-----------|-------|---|---|---|---|---|---|---|---|
| A        | 8.34 | 8.34 | 8.34 | 8.33 | A | 5.45 | 5.45 | 14.10 |
| U        | 8.34 | 8.34 | 8.34 | 8.33 | U | 5.45 | 14.10 |
| C        | 8.34 | 8.34 | 8.34 | 8.33 | C | 5.45 | 14.10 |

Table 4 Transition and transversion rate of the chloroplast tRNAs. Bold letter indicates transition.
Given the cyanobacterial origin of the chloroplast genome, it should be prokaryotic in nature, and in general, should be intron free. However, we found the presence of group I introns in the chloroplast tRNAs. Previous studies have also reported the presence of intron in tRNA$^{\text{Leu}}$ (UAA) and tRNA$^{\text{Met}}$ (UAC) of cyanobacterial tRNA [53, 54]. Additionally, a recent study conducted in our laboratory also reported the presence of introns in cyanobacterial tRNA$^{\text{Arg}}$, tRNA$^{\text{Gly}}$, and tRNA$^{\text{Lys}}$ [45]. Although the presence of introns in the cyanobacterial genome has been reported by several studies, the present study appears to be the first to report the presence of introns in chloroplast tRNA. The group I introns lack significant sequence conservation, however, the present analysis indicated that they contain short conserved consensus sequences. The group I intron of chloroplast tRNA grouped with the group I intron of cyanobacteria (Fig. 4), thus providing additional evidence to suggest that they evolved from a common cyanobacterial lineage.

As proposed by Robert Holley [34], tRNAs are characterized by a cloverleaf-like structure, although a few tRNAs vary in their secondary structure [35]. tRNAs contain various arms and loops that function in protein translation. Each arm and loop have their own unique nucleotide composition. A previous study reported that the acceptor arm contains seven base pairs 7 bp, the D-stem 3–4 bp, the D-loop 4–12 nucleotides, the anti-codon arm 5 bp, the anti-codon loop 7 nucleotides, the variable region 4–23 nucleotides, the Ψ-arm 5 bp, and the Ψ-loop seven nucleotides [37]. The previous report, along with the present study, suggests that significant variation exists in arms and loops of chloroplast tRNAs. The acceptor arm contains distinct information for tRNA-nucleotidyltransferases. However, the absence of an acceptor arm in tRNA$^{\text{Gly}}$ (UCC) of O. nivara and tRNA$^{\text{Thr}}$ (GGU) of S. bicolor is quite intriguing. The question arises as to how a tRNA without an acceptor arm can participate to carry an amino acid during the process of protein translation? Some tRNAs contain novel loops having A-C-U-U-U-G nucleotides. The stem of the novel loop allows the bonding of A to U and G to U nucleotides. The novel loop structures identified in the present study raises the question whether these loops mimic the anti-codon loop of the tRNA and play a critical role in the protein translation machinery within the chloroplast. Some of the tRNA were also found to contain nine nucleotides in the anti-codon loop; which may represent a novel phenomenon of tRNA. The functional impact of having nine nucleotides in the anti-codon loop remains to be determined. In addition to the presence of few putative novel tRNA structure, chloroplast tRNAs were found to contain a C-A-U anti-codon that codes for tRNA$^{\text{Ile}}$ as well. However, the presence of a C-A-U anti-codon in tRNA$^{\text{Ile}}$ was previously reported in Bacillus subtilis [55].

Phylogenetic analysis of chloroplast tRNA showed two distinct clusters and multiple groupings. Some of the tRNA members of cluster I also found to be present in cluster II; suggesting their evolution by duplication and divergence. However, anti-codon GAC, UAC, GGU, UGU, UGA, GCC, UCC, UAA, UAG, GAA, CCA, CAU UGC, UCU, and ACG fall independently in the phylogenetic tree; suggesting their evolution from multiple common ancestors. The overlapping grouping of tRNA family members suggests that the tRNAs with these anti-codon groups may have evolved from different common ancestors or may have arisen from duplication events. The presence of tRNA$^{\text{Met}}$ twice in cluster I and once in cluster II indicates that tRNA$^{\text{Met}}$ is one of the tRNA families that has undergone major duplication event(s) to give rise to other tRNAs. Phylogenetic analysis further revealed that tRNA$^{\text{Leu}}$ (CAA), tRNA$^{\text{Trp}}$ (CCA), tRNA$^{\text{Arg}}$ (UCU), tRNA$^{\text{Asn}}$ (GUU), tRNA$^{\text{Tyr}}$ (GUA), tRNA$^{\text{Met}}$ (CAU), tRNA$^{\text{Cys}}$ (GCA), and tRNA$^{\text{Phe}}$ (GAA) present in cluster II are the most primitive form of tRNAs with tRNA$^{\text{Leu}}$ as the most basal evolutionary

|        | A     | U     | C     | G     | A     | U     | C     | G     |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|
| Isoleucine | G 8.33 | 8.34  | 8.34  | –     | G 14.10 | 5.45  | 5.45  | –     |
| Leucine | A –   | 8.60  | 8.60  | 7.80  | A –   | 5.15  | 5.15  | 14.71 |
|         | U 8.60 | –     | 7.80  | 8.60  | U 5.15 | –     | 14.71 | 5.15  |
|         | C 8.60 | 7.80  | –     | 8.60  | C 5.15 | 14.71 | –     | 5.15  |
|         | G 7.80 | 8.60  | 8.60  | –     | G 14.71 | 5.15  | 5.15  | –     |

**Table 4** Transition and transversion rate of the chloroplast tRNAs. Bold letter indicates transition (Continued)
ancestor. The grouping of tRNA\(^{\text{Met}}\) (CAU) with tRNA\(^{\text{Ile}}\) (CAU), and their similar anti-codon nucleotides, strongly suggests that tRNA\(^{\text{Ile}}\) evolved directly from tRNA\(^{\text{Met}}\). The overall analysis clearly indicates that tRNA\(^{\text{Met}}\) is a major player in the evolution of tRNAs in the chloroplast genome. The distribution of tRNA\(^{\text{Met}}\) in two different clusters strongly suggests that tRNA\(^{\text{Met}}\) underwent several major substitution and duplication events to give rise to diverse tRNA families with distinct anti-codons. The rate of transition of chloroplast tRNAs were higher than the rate of transversion. tRNA\(^{\text{Cys}}\), tRNA\(^{\text{Gln}}\), tRNA\(^{\text{Trp}}\), and tRNA\(^{\text{Tyr}}\) belong to a polar R group and the rate of transversion is zero in tRNAs that carry polar amino acids. Polar amino acids are readily soluble in water and form strong hydrogen bonds with interacting molecules. This suggests that the evolution of chloroplast tRNA\(^{\text{Cys}}\), tRNA\(^{\text{Gln}}\), tRNA\(^{\text{Trp}}\), and tRNA\(^{\text{Tyr}}\) strongly favors transition substitutions rather than transversion substitutions and that some tRNA isoacceptors undergo transition more readily than transversion. A few tRNAs, however, underwent a higher rate of transversion than transition; suggesting that the rate of evolution and the rate of transition and transversion of tRNAs are isoacceptor-specific and that tRNAs have not undergone an equal rate of evolution.

In addition to the mutational event, gene duplication is also a major force in evolution and represents an important mechanism by which species acquire new genes [56]. The majority of novel gene functions have evolved through gene duplication events which can occur by genome duplication, retrotransposons, and unequal crossing over [57, 58]. Ancient duplication events coupled with the retention of extant pairs of duplicated genes have contributed enormously to the evolution of gene families and functional diversification [59]. Plant genomes tend to evolve at a high rate, leading to greater genome diversity relative to other organisms [60]. The study of chloroplast tRNAs showed the rate of deletion of tRNA is superior than the rate of duplication. This suggests that the maternal inheritance of the cyanobacterial-derived chloroplast genome is more intact than the nuclear-encoded plant genome. Therefore, although the species were part of the same lineage, some genes were still lost within each species. This provides further evidence that cyanobacterial tRNAs originated from polyphyletic common ancestors, and hence, loss events are more pronounced than duplication events. Almost all of the tRNAs experienced loss events in either of species studied (Table 5).
| tRNA Gene ID  | Species             | tRNA | Anti-codon | Lost in Species |
|--------------|---------------------|------|------------|-----------------|
| NC_002762.1  | T. aestivum         | Val  | GAC        | Z. mays         |
| NC_008602.1  | S. bicolor          | Val  | GAC        | Z. mays, T. aestivum |
| NC_006084.1  | S. officinarum      | Val  | GAC        | Z. mays, S. bicolor, T. aestivum |
| NC_001320.1  | O. sativa           | Val  | GAC        | O. nivara       |
| 3782         | O. sativa           | Val  | GAC        | O. nivara, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6258         | O. nivara           | Val  | GAC        | O. nivara, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6217         | O. nivara           | Val  | GAC        | O. nivara, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6160         | O. nivara           | Val  | UAC        | O. sativa       |
| 2026         | Z. mays             | Val  | UAC        | S. bicolor      |
| NC_002762.1  | T. aestivum         |      | GAC        |                 |
| 6601         | S. officinarum      |      | UAC        |                 |
| 6160         | O. nivara           |      | UAC        |                 |
| 2085         | Z. mays             | Ala  | UGC        | S. bicolor      |
| 6221         | O. nivara           | Ala  | UGC        |                 |
| 5101         | T. aestivum         | Ala  | UGC        |                 |
| 5076         | T. aestivum         | Ala  | UGC        |                 |
| 6691         | S. officinarum      | Ala  | UGC        |                 |
| 6662         | S. officinarum      | Ala  | UGC        |                 |
| 3815         | O. sativa           | Ala  | UGC        |                 |
| 3786         | O. sativa           | Ala  | UGC        |                 |
| 6221         | O. nivara           | Ala  | UGC        |                 |
| 2114         | Z. mays             | Ala  | UGC        |                 |
| 6253         | O. nivara           | Ala  | UGC        |                 |
| 2085         | Z. mays             | Ala  | UGC        | O. sativa       |
| 6221         | O. nivara           | Ala  | UGC        |                 |
| 5076         | T. aestivum         | Ala  | UGC        | Z. mays, O. nivara, O. sativa |
| 6662         | S. officinarum      | Ala  | UGC        | Z. mays, O. sativa, O. nivara, T. aestivum |
| 3815         | O. sativa           | Ala  | UGC        | O. nivara, S. officinarum, T. aestivum, Z. mays |
| 3786         | O. sativa           | Ala  | UGC        | O. nivara, S. officinarum, T. aestivum, Z. mays |
| 6221         | O. nivara           | Ala  | UGC        | O. sativa, S. officinarum, T. aestivum, Z. mays |
| 2114         | Z. mays             | Ala  | UGC        | O. sativa, O. nivara, S. officinarum, T. aestivum |
| 6253         | O. nivara           | Ala  | UGC        | O. sativa, S. officinarum, T. aestivum, Z. mays |
| NC_001320.1  | O. sativa           | Thr  | GGU        | O. nivara       |
| 6132         | O. nivara           | Thr  | GGU        | O. sativa, S. officinarum, T. aestivum, Z. mays |
| 1998         | Z. mays             | Thr  | GGU        | O. nivara, O. sativa, S. bicolor, S. officinarum, T. aestivum |
| NC_001320.1  | O. sativa           | Thr  | UGU        | O. nivara       |
| 6154         | O. nivara           | Thr  | UGU        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 3729         | O. sativa           | Met  | CAU        | O. nivara       |
| 6161         | O. nivara           | Met  | CAU        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| NC_001320.1  | O. sativa           | Asp  | GUC        | O. nivara       |
| 6135         | O. nivara           | Asp  | GUC        | O. nivara, O. sativa, S. bicolor, S. officinarum, T. aestivum |
| 6115         | O. nivara           | Lys  | UUU        | O. sativa       |
| 5075         | T. aestivum         | Ile  | GAU        | Z. mays         |
| 6660         | S. officinarum      | Ile  | GAU        | T. aestivum, Z. mays |
Table 5 Loss event of chloroplast tRNA in plants (Continued)

| tRNA Gene ID | Species       | tRNA | Anti-codon | Lost in Species                        |
|--------------|---------------|------|------------|----------------------------------------|
| 3816         | O. sativa     | Ile  | GAU        | O. nivara                              |
| 3784         | O. sativa     | Ile  | GAU        | O. nivara, S. officinarum, T. aestivum, Z. mays |
| 6255         | O. nivara     | Ile  | GAU        | O. sativa                              |
| 6219         | O. nivara     | Ile  | GAU        | O. sativa, S. officinarum, T. aestivum, Z. mays |
| 6556         | S. officinarum| Lys  | UUU        | S. bicolor                             |
| 6115         | O. nivara     | Lys  | UUU        |                                        |
| 1982         | Z. mays       | Lys  | UUU        |                                        |
| 4982         | T. aestivum   | Lys  | UUU        |                                        |
| 2116         | Z. mays       | Ile  | GAU        |                                        |
| 2083         | Z. mays       | Ile  | GAU        |                                        |
| 5102         | T. aestivum   | Ile  | GAU        |                                        |
| 5075         | T. aestivum   | Ile  | GAU        |                                        |
| 6693         | S. officinarum| Ile  | GAU        |                                        |
| NC_006084.1  | S. officinarum| Ile  | GAU        |                                        |
| 3816         | O. sativa     | Leu  | UAA        | O. sativa                              |
| 3816         | O. sativa     | Leu  | UAA        | O. sativa, S. officinarum              |
| 3784         | O. sativa     | Leu  | UAA        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6255         | O. nivara     | Leu  | UAA        | O. nivara                              |
| 6219         | O. nivara     | Leu  | UAA        | O. nivara                              |
| 3723         | O. sativa     | Leu  | UAA        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6155         | O. nivara     | Ser  | UGA        | O. nivara                              |
| NC_001320.1  | O. sativa     | Ser  | UGA        | O. nivara                              |
| 6125         | O. nivara     | Ser  | UGA        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 3686         | O. sativa     | Ser  | UGA        | O. nivara                              |
| 6121         | O. nivara     | Ser  | GCU        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 3720         | O. sativa     | Ser  | GGA        | O. nivara                              |
| 6152         | O. nivara     | Ser  | GGA        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 3748         | O. sativa     | Pro  | UGG        | O. nivara                              |
| 6178         | O. nivara     | Pro  | UGG        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 3683         | O. sativa     | Gln  | UUG        | O. nivara                              |
| 6118         | O. nivara     | Gln  | UUG        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6178         | O. nivara     | Gln  | UUG        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 5066         | T. aestivum   | His  | GUG        | Z. mays                                |
| NC_008602.1  | S. bicolor    | His  | GUG        | Z. mays, T. aestivum                   |
| NC_006084.1  | S. officinarum| His  | GUG        | S. bicolor, T. aestivum, Z. mays       |
| 3833         | O. sativa     | His  | GUG        | O. nivara                              |
| 3770         | O. sativa     | His  | GUG        | O. nivara, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6275         | O. nivara     | His  | GUG        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6129         | O. nivara     | Gly  | UCC        | O. sativa                              |
| 3701         | O. sativa     | Glu  | UUC        | O. nivara                              |
| 6133         | O. nivara     | Glu  | UUC        | O. sativa, T. aestivum                 |
| 3694         | O. sativa     | Met  | CAU        | O. nivara                              |
| 6128         | O. nivara     | Met  | CAU        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 5015         | T. aestivum   | Met  | CAU        | O. nivara, O. sativa, S. bicolor, S. officinarum, Z. mays |
| 5080         | T. aestivum   | Arg  | ACG        | Z. mays                                |
Table 5 Loss event of chloroplast tRNA in plants (Continued)

| tRNA Gene ID    | Species        | tRNA | Anti-codon | Lost in Species                      |
|-----------------|----------------|------|------------|--------------------------------------|
| NC_006084.1     | S. officinarum | Arg  | ACG        | T. aestivum, Z. mays                  |
| NC_001320.1     | O. sativa      | Arg  | ACG        | O. nivara                            |
| 3790            | O. sativa      | Arg  | ACG        | O. nivara, O. sativa, S. bicolor, S. officinarum, Z. mays |
| 6249            | O. nivara      | Arg  | ACG        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 3724            | O. sativa      | Phe  | GAA        | O. nivara                            |
| 6156            | O. nivara      | Phe  | GAA        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 3706            | O. sativa      | Cys  | GCA        | O. nivara                            |
| 6138            | O. nivara      | Cys  | GCA        | O. sativa, T. aestivum               |
| 5069            | T. aestivum    | Met  | CAU        | Z. mays                              |
| NC_008602.1     | S. bicolor     | Met  | CAU        | T. aestivum, S. bicolor              |
| NC_006084.1     | S. officinarum | Met  | CAU        | S. bicolor, T. aestivum, Z. mays      |
| 3828            | O. sativa      | Met  | CAU        | O. nivara                            |
| 3774            | O. sativa      | Met  | CAU        | O. nivara, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6270            | O. nivara      | Met  | CAU        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6206            | O. nivara      | Met  | CAU        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6134            | O. nivara      | Tyr  | GUA        | O. sativa                            |
| 3702            | O. sativa      | Tyr  | GUA        | O. nivara, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 5081            | T. aestivum    | Asn  | GUU        | Z. mays                              |
| NC_008602.1     | S. bicolor     | Asn  | GUU        | T. aestivum, Z. mays                  |
| NC_006084.1     | S. officinarum | Asn  | GUU        | S. bicolor, T. aestivum, Z. mays      |
| 3809            | O. sativa      | Asn  | GUU        | O. nivara                            |
| 3792            | O. sativa      | Asn  | GUU        | O. nivara, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6247            | O. nivara      | Asn  | GUU        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6228            | O. nivara      | Asn  | GUU        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 3715            | O. sativa      | Arg  | UCU        | O. nivara                            |
| 6147            | O. nivara      | Arg  | UCU        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 3747            | O. sativa      | Trp  | CCA        | O. nivara                            |
| 6177            | O. nivara      | Trp  | CCA        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 5070            | T. aestivum    | Leu  | CAA        | Z. mays                              |
| NC_008602.1     | S. bicolor     | Leu  | CAA        | T. aestivum, Z. mays                  |
| NC_006084.1     | S. officinarum | Leu  | CAA        | S. bicolor, T. aestivum, Z. mays      |
| 3825            | O. sativa      | Leu  | CAA        | O. nivara                            |
| 3777            | O. sativa      | Leu  | CAA        | O. nivara, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6263            | O. nivara      | Leu  | CAA        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 6212            | O. nivara      | Leu  | CAA        | O. sativa, S. bicolor, S. officinarum, T. aestivum, Z. mays |
| 2124            | Z. mays        | Leu  | CAA        | O. nivara, O. sativa, S. bicolor, S. officinarum, T. aestivum |

**Conclusion**

We conducted a tRNA analysis of the chloroplast genome of six monocot plants and found that the chloroplast genome in these plant species encode 28 to 39 tRNA genes. The numbers of tRNA Isoacceptors ranged from 23 to 29 and the majority of tRNAs were associated with only one Isoacceptor. The tRNAs in the chloroplast genome were also found to contain a group I intron in the anti-codon region and a phylogenetic analysis revealed that the chloroplast tRNAs in monocot plants evolved from multiple common ancestors. The chloroplast genomes of the examined monocot plant species were also found to contain putative, novel tRNAs which need to be further investigated to understand
their biological significance. An analysis of gene duplication and loss events revealed that gene loss events were more pronounced than duplication events in chloroplast tRNA.

**Methods**

**Identification and analysis of chloroplast tRNA of monocot plants**

The chloroplast genomes of the monocot species, *O. nivara* (NC_005973), *O. sativa* (NC_001320), *S. officinarum* (NC_006084), *S. bicolor* (NC_008602), *T. aestivum* (NC_002762), and *Z. mays* (NC_001666) were downloaded from the public database available at the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) [46, 61, 62]. The sequences were downloaded in FASTA format (Additional file 1: Table S1, Additional file 3: Data S1) and subsequently all of the chloroplast genomes were subjected to annotation. Annotation of all the chloroplast genomes was carried out using GeSeq-Annotation of Organellar Genomes (https://chlorobox.mpimp-golm.mpg.de/geseq.html) [63]. Parameters used to carry out the annotation process were circular sequence (s); sequence source, chloroplast; generate multi FASTA; annotate plastid IR, BLAT protein search identity 25%; BLAT tRNA, tRNA and DNA search 85% identity; HMMER profile search; Embryophyta chloroplast (CDS + rRNA); 3rd party tRNA annotator ARAGRON v1.2.38, ARWEN v1.2.3, tRNAscan-SE v2.0; and no Refseq selection party tRNA annotator ARAGRON v1.2.38, ARWEN v1.2.3, tRNAscan-SE v2.0; and no Refseq selection.

Annotated nucleotide sequences of the chloroplast tRNA genes in the six-monocot species were collected and used in the further sections of this study. The free energy calculation of predicted novel tRNAs was performed using the RNAalifold webserver with default parameters [64].

**Analysis of chloroplast tRNA of monocot plants**

The collected genomic tRNA sequences of chloroplast tRNAs of monocot plants were subjected to further analysis using ARAGRON and the tRNAscan-SE server [65]. Default parameters were used to analyze the genomic tRNA sequences in ARAGRON. In the tRNAscan-SE server, the following parameters were used to analyze the genomic tRNA; sequence source, bacterial; search mode, default; query sequences, formatted (FASTA); and genetic code for tRNA isotype prediction, universal. All of the tRNAs were analyzed using the same parameters and the number and composition of nucleotides in different arms and loops were recorded individually. The tRNAs that were found to have a different structure than the canonical clover leaf-like structure characteristic of tRNA were considered as putative novel tRNAs.

**Multiple sequence alignment**

To identify and analyze the conserved nucleotide sequences of tRNA isotypes, the nucleotide sequences of 20 isotypes were separately grouped. Later, tRNA isotypes were subjected to multiple sequence alignment using the Multalin server. All of the sequences, in FASTA format, were used in the alignment analysis with the following parameters; 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. The highest alignment consensus value was maintained at 90% (default); whereas, the lowest consensus value was kept at 50% (default). In the displayed alignments, red indicates a similarity/conservation of 90% or more; whereas, blue indicates a sequence conservation less than 90%. Alignments displayed in black indicates no conservation.

**Construction of phylogenetic tree**

To analyze the evolution of chloroplast tRNAs in monocot plants, a phylogenetic tree was constructed using MEGA6.0 software [66]. Prior to construction of the phylogenetic tree, a Clustal file of all the tRNAs was created using the Clustal omega server. The generated Clustal file of tRNAs was converted to a MEGA file format using MEGA6 software. Model selection was performed prior to the construction of the phylogenetic tree. Model selection was conducted by MEGA6 software using the following statistical parameters: analysis, model selection (ML); tree to use, automatic (neighbor-joining); statistical method, maximum likelihood; substitution type, nucleotide; gaps/missing data treatment, partial deletion and site coverage cutoff was 95%. The model selection analysis that resulted in the lowest Bayesian information criterion (BIC) was considered as the best model to construct the phylogenetic tree. The lowest BIC score was found to be 7785.682 for the Kimura2+ G + I model; as a result, the latter model was used to construct a phylogenetic tree. Other statistical parameters within the Kimura2+ G + I model were: analysis, phylogeny reconstruction; statistical model, maximum likelihood; test of phylogeny, bootstrap method; no. of bootstrap replicates, 1000; substitution type, nucleotides; 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 MEGA file format of tRNAs used to construct the phylogenetic tree was used to analyze the
transition/transversion rate for all of the tRNAs. Additionally, transition/transversion rates of all of the 20 tRNA isotypes were separately studied. The tRNA isotypes were also subjected to multiple sequence alignment using the Clustal omega server to generate a Clustal file for each individual isotype. The generated Clustal files of tRNA isotypes were converted to a MEGA file format and the rate of substitution was estimated using MEGA6 software. The following statistical parameters were used to study the transition/transversion rates in the chloroplast tRNAs of monocot plants: 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.

Disparity index analysis
To determine if all of the substitutions of nucleotides occurred homogeneously (equal rates) during evolution, a disparity index test of the pattern heterogeneity was conducted to determine the homogeneity of nucleotide substitutions. Statistical parameters used to analyze the pattern of homogeneity were: analysis, disparity index test of substitution pattern homogeneity; scope, in sequence pairs; no. of Monte Carlo Replications, 1000; substitution type, nucleotide; gaps/missing data treatment, partial deletion; and site coverage cutoff was 95%.

Analysis of gene duplication and loss
An all species tree was first constructed using the NCBI taxonomy browser (https://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi) to analyze the duplication and loss events of tRNA genes. Species used to construct the species tree were O. nivara, O. sativa, S. officinarum, S. bicolor, T. aestivum, and Z. mays. The phylogenetic tree used for the evolutionary analysis was utilized as the gene tree. Gene duplication/loss events were studied using Notung2.6 software. The gene tree was reconciled with the species tree during the analysis to obtain the duplication and loss nodes of the genes.

Additional file 1: Table S1. Nucleotide composition of acceptor arm, D-arm, D-loop, anti-codon arm, variable loop, pseudouridine arm and pseudouridine loop of chloroplast tRNA. (DOCX 26 kb)

Additional file 2: Figure S1. Phylogenetic tree of cyanobacterial tRNAs with tRNAs of Araboana cystidiana, Methanococcus maripaludis, Methanospirillum hungatei, Oscillatoria acuminata, and Thermococcus sibiricus. The tRNAs of these species were included as ingroup, whereas, ANAC1 and ANAC2 (NAC transcription factor) of Arabidopsis thaliana were used as out-groups. Phylogenetic tree was constructed using the Neighbor-joining method and 1000 bootstrap replicates using MEGA6 software. (PDF 114 kb)

Additional file 3: Data S1. tRNA sequences of studied chloroplast genome of the monocot plants. (TXT 24 kb)

Abbreviations
A: Adenine; C: Cytosine; G: Guanine; tRNA: Transfer RNA; U: Uracil; Ψ: Pseudouridine

Acknowledgements
Authors would like to extend their sincere thanks to the Chair of Natural and Medical Sciences Research Center, University of Nizwa for providing necessary support to carryout this research. The authors also like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for its funding this Research group NO (RGP- 271).

Availability of data materials
All the genomic tRNA sequences used during this study are provided as Additional file 3: Data S1.

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

Author’s contributions
TKM: Conceived the idea, performed the experiment and analysis, drafted and revised the manuscript. ALK: revised the manuscript, DY: Revised the manuscript; AH: drafted and revised the manuscript, EFA: drafted and revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Natural and Medical Sciences Research Center, University of Nizwa, 616 Nizwa, Oman. 2Botany and Microbiology Department, College of Science, King Saud University, Riyadh 11451, Saudi Arabia. 3Myology and Plant Disease Survey Department, Plant Pathology Research Institute, Agriculture Research Center, Giza, Egypt. 4Plant Production Department, College of Food and Agriculture Science, King Saud University, Riyadh 11451, Saudi Arabia. 5Department of Medical Biotechnology, Yeungnam University, Gyeongsan 38541, Republic of Korea.

Received: 12 June 2018 Accepted: 28 December 2018

Published online: 22 January 2019

References
1. Wise RR, Hooper J. The structure and function of plastids. Divers. Plast. In: Form Funct; 2006.
2. Wolfgang K, Martin H. Uncertainties in global terrestrial biosphere modeling: a comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme. Global Biogeochem Cycles Wiley-Blackwell. 2001;15:207–25.
3. Des Marais DJ. When did photosynthesis emerge on earth? Science. 2000; 289:1703 LP–1705.
4. Stern DB, Goldschmidt-Clermont M, Hanson MR. Chloroplast RNA metabolism. Annu Rev Plant Biol Annual Reviews. 2010;61:125–55.
5. Bolton JR, Hall DO. Photoschemical conversion and storage of solar energy. Annu Rev Energy Annual Reviews. 1979;4:353–401.
6. Jensen WA. Chloroplasts and photosynthesis. In: Jensen WA, editor. Plant Cell. London: Macmillan Education UK; 1973. p. 25–45.
61. Shahid Masood M, Nishikawa T, Fukuoka S, Njenga PK, Tsudzuki T, Kadowaki K. The complete nucleotide sequence of wild rice (Oryza nivara) chloroplast genome: first genome wide comparative sequence analysis of wild and cultivated rice. Gene. 2004;340:133–9.

62. Saski C, Lee SB, Fjellheim S, Guda C, Jansen RK, Luo H. Complete chloroplast genome sequences of Hordeum vulgare, Sorghum bicolor and Agrostis stolonifera, and comparative analyses with other grass genomes. Theor Appl Genet. 2007;115.

63. Tillich M, Lehward P, Pellizze T, Ulbricht-Jones ES, Fischer A, Bock R, et al. GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 2017;45:W6–11.

64. Bernhart SH, Hofacker IL, Will S, Gruber AR, Stadler PF. RNAalifold: improved consensus structure prediction for RNA alignments. BMC Bioinformatics. 2008;9:474.

65. Lowe TM, Chan PP. tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res. Oxford University Press; 2016;44(W5):W54–W57.

66. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary Genetics analysis version 6.0. Mol Biol Evol. 2013;30:2725–9.