Studying the evolutionary relationships and phylogenetic trees of 21 groups of tRNA sequences based on complex networks

Fangping Wei · Bowen Chen

Received: 20 January 2011 / Accepted: 8 August 2011 /
Published online: 30 September 2011
© Springer Science+Business Media B.V. 2011

Abstract To find out the evolutionary relationships among different tRNA sequences of 21 amino acids, 22 networks are constructed. One is constructed from whole tRNAs, and the other 21 networks are constructed from the tRNAs which carry the same amino acids. A new method is proposed such that the alignment scores of any two amino acids groups are determined by the average degree and the average clustering coefficient of their networks. The anticodon feature of isolated tRNA and the phylogenetic trees of 21 group networks are discussed. We find that some isolated tRNA sequences in 21 networks still connect with other tRNAs outside their group, which reflects the fact that those tRNAs might evolve by intercrossing among these 21 groups. We also find that most anticodons among the same cluster are only one base different in the same sites when $S \geq 70$, and they stay in the same rank in the ladder of evolutionary relationships. Those observations seem to agree on that some tRNAs might mutate from the same ancestor sequences based on point mutation mechanisms.

Keywords tRNA sequences · Anticodon · Network · Phylogenetic tree

1 Introduction

It has been proven that BA models [1] and small-world models [2, 3] are effective methods to discuss complex natural systems, such as the Internet [4], the WWW [5, 6], protein networks [7], human disease networks [8], tRNA networks [9, 10], and sequence similarity

F. Wei (✉) · B. Chen
Physical Science & Technology College, Guangxi University,
Nanning, Guangxi 530004, China
e-mail: weipp@163.com

B. Chen
Key Laboratory of Information Photonics and Optical Communications,
Ministry of Education, Beijing University of Posts and Telecommunications,
Beijing 100876, China

networks [11]. Transfer RNA (tRNA) has a cloverleaf-shaped secondary structure, and it plays an essential role in protein synthesis. There are many theories and methods about studying evolutionary relationships of tRNA sequences, such as the study of the evolutionary history of tRNA by comparative sequence analysis of two specified tRNAs at various phylogenetic levels and of tRNA families within different species [12]. Some studies have found that tRNA genes can be recruited from one iso-accepting group to another by an anticodon's point mutation that concurrently changes tRNA amino acid identity and messenger RNA-coupling capacity [13]. The exact relationships between species might not only be found by sequence analysis, but could also reveal information about the constraints concerning the evolution of the tRNA molecule [14].

In a phylogenetic tree, each node represents the most recent common ancestor of the descendants, and the edge lengths in some trees correspond to time estimates. Phylogenetic tree theories have been used to study the evolution of tRNAs comprehensively [15]. The phylogenetic trees of the aminoacyl-tRNA synthetases of prokaryotes have been reconstructed by the Maximum Likelihood method [16]. Recently, Wei et al. proposed a new method in which the evolutionary relationship of modern tRNA sequences can be obtained by the analysis of tRNA networks [17].

In this paper, firstly, according to the 20 amino acids and one stop codon those anticodons carry, 3,420 tRNA sequences are divided into 21 groups, and then each group and the whole tRNAs are assembled into 22 networks with a given similarity degree. On one hand, we attempt to determine the relationships of the isolated tRNA anticodons which have no connections in their group but have many links with other groups. On the other hand, the phylogenetic trees of 20 amino acids and one stop codon can be generated by the neighbor-joining method under different similarity degrees. The evolutionary properties of the anticodons between isolated tRNAs and their connections with the entire tRNA sequences are discussed, as well as the associated phylogenetic trees.

2 Materials and methods

2.1 Experimental data

The tRNA sequences are derived from the database (<http://www.uni-bayreuth.de/departament/biochemie/sprinzl/trna/>) [18], which contains 3,420 tRNA sequences, including 59 different anticodon subsets and 429 different species that belong to three kingdoms: Archaea, Bacteria, and Eucarya. Each tRNA sequence has 99 bases in its length if its variable stem is considered. The number of tRNA sequences is updated periodically if new data are entered into this database [18–20].

2.2 The similarity degree

To compare these two different tRNAs simply, the similarity degree is defined as the total number of the identical base pairs at the corresponding site of two tRNA sequences,

Fig. 1 Similarity degree of two tRNAs' segments, $S = 11$

G	C	G	—	T	C	C	G	A	T	C	—	T	G	C	A	G	G	A	A	G	C
T	C	C	—	T	A	G	G	A	C	G	—	T	T	C	A	C	G	T	A	C	T

such as A–A, G–G, and is represented by S . For example, the similarity degree $S = 11$ in Fig. 1. Theoretically, if all present tRNA sequences evolved from common ancestors via mutation, the evolutionary relationship of two tRNA sequences is proportional to the similarity degree. Therefore, the similarity degree can reflect directly kindred evolutionary relationships among different tRNA sequences.

2.3 The degree and clustering coefficient

The most elementary character of a node is its degree, which tells us how many links a node has to other nodes. An undirected network with N nodes and M links is characterized by an average degree $\langle D \rangle$,

$$\langle D \rangle = \frac{2M}{N}. \quad (1)$$

One node i is selected in an undirected network, and it has K_i edges which connect it to K_i other nodes. If the nearest neighbors of the original node were part of a clique, there would be $K_i(K_i - 1)/2$ edges between them. The ratio between the number E_i of edges that actually exist between these K_i nodes and the total number $K_i(K_i - 1)/2$ gives the value of the clustering coefficient C of node i ,

$$C_i = \frac{2E_i}{k_i(k_i - 1)}. \quad (2)$$

Thus, the average clustering coefficient of a network is the ratio between the sum of all the C_i and the total nodes N :

$$\langle C \rangle = \frac{\sum C_i}{N}. \quad (3)$$

2.4 The method of constructing a tRNA network

To construct a tRNA network, the following steps are followed:

- (i) Each tRNA sequence is regarded as one node of a network.
- (ii) Comparing any one tRNA sequence with the others, two nodes of a network are connected when the similarity degree is larger than or equal to a given similarity degree S (connective condition).

Repeating these two steps, an undirected tRNAs network can be constructed by a given similarity degree. According to different tRNA anticodons, 3,420 tRNA sequences are divided into 21 groups, including Ala (210), Arg (177), Asn (145), Asp (114), Cys (136), Gln (143), Glu (134), Gly (167), His (135), Ile (163), Leu (290), Lys (139), Met (250), Phe (127), Pro (158), Ser (294), stop codon (98), Thr (172), Trp (54), Tyr (152), and Val (162) groups. The number in parentheses represents the number of tRNAs in the corresponding group. By using these two steps, 21 undirected networks are constructed.

For convenient comparison, we also construct a network with all 3,420 tRNAs. The isolated nodes of one amino acid or one stop codon network will be compared with the tRNAs in the whole network under the same similarity degree.

2.5 Phylogenetic tree construction

Many methods such as distance, parsimony and likelihood methods are usually used to construct a phylogenetic tree. In this paper, we construct tRNA phylogenetic trees by the neighbor-joining method [21]. To describe and discuss the evolutionary relationships among large numbers of tRNA sequences simply, a novel method based on alignment scores among different tRNA groups is used, which can be obtained by two important parameters, $\langle D \rangle$ and $\langle C \rangle$, of different networks based on two different groups of tRNAs rather than by comparing only two sequences. Thus, the tree can be constructed by the following steps:

- (i) We use i th and j th amino acid or stop codon group to combine a larger tRNA group and construct a network with it, and calculate the average degree $\langle D(i, j) \rangle$ and the average clustering coefficient $\langle C(i, j) \rangle$ of this larger network.
- (ii) It is supposed that the alignment scores of a distance matrix of a network are decided by the equation

$$\text{score}(i, j) = k_1 \langle D(i, j) \rangle + k_2 \langle C(i, j) \rangle. \quad (4)$$

In (4), the coefficients k_1 and k_2 are defined as $k_1 = \frac{S}{2 \times L(N-1)}$ and $k_2 = \frac{S}{2 \times L}$, where L , S , and N represent the length of a tRNA sequence ($L = 99$), the similarity degree, and the total nodes of a network, respectively. These two coefficients are introduced to eliminate the effect of different numbers of nodes and different similarity degree of networks and normalize the alignment score to the interval $[0, 1]$. Therefore, the alignment score of a distance matrix element can be expressed as

$$\text{score}(i, j) = \begin{cases} \frac{1}{2} \left(\frac{S}{L} \right) \times \left(\frac{\langle D(i, j) \rangle}{N(i, j) - 1} + \langle C(i, j) \rangle \right), & i \neq j \\ 0, & i = j \end{cases} \quad (5)$$

where i and j run from 1 to 21, and represent the i th and j th amino acid or stop codon group. It has been proven that all tRNAs have high similarity in sequences [10, 17]. So, in this paper, we use $\langle D(i, j) \rangle$ and $\langle C(i, j) \rangle$ to present the alignment scores of two different amino acids or stop codon groups, which can reveal the evolutionary affiliation of the tRNAs between two different amino acids and interpret the overall evolutionary tendency of tRNA sequences to a certain extent.

- (iii) We construct the phylogenetic trees using the software of Phylip3.67 and Tree-View1.6.6.

3 Results

As Table 1 shows, when similarity degree $S = 60$, which means two tRNA sequences have not less than 60% base identity at the same sites, there are only six groups of amino acids (Asp, Ile, Leu, Lys, Thr, and Val) that have isolated tRNAs and eight isolated tRNA sequences among these (corresponding anticodons are GUC, GUC, GAU, UAU, UAA, UUU, GGU, and UAC) have connections with other groups. One particular isolated tRNA with anticodon UAA, connects with the tRNAs in all groups except the His group, and nine groups have synonymous anticodons (which are marked by the red type in the table). However, another isolated tRNA sequence, with anticodon UAU, also connects with 17 different amino acid groups and one stop codon group, and three groups of these have

Table 1 The isolated tRNAs in one group connect with the tRNAs in other groups under the similarity degree $S = 60$

S = 60	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu
Asp	GUC	–	–	–	–	UUG-2	–	–	–	GAU-3	–
	GUC	–	–	–	GCA-1	–	–	–	GUG-2	–	–
Ile	GAU	–	UCU-1	–	–	–	–	–	–	–	–
	UAU	UGC-2	UCG-9	GUU-2	GUC-5	GCA-10	UUG-6	UUC-7	UCC-6	GUG-7	–
			UCC-4								UAA-2
Leu	UAA	AGC-2	ACG-18	GUU-16	GUC-13	GCA-2	CUG-3	CUC-1	GCC-8	–	GAU-58
		CGC-2	CCG-2				UUG-29	UUC-3	UCC-5		–
		GGC-2	CCU-1								UAG-7
		UGC-46	UCG-6								
			UCU-12								
Lys	UUU	–	–	–	–	UUG-4	–	–	–	–	–
Thr	GGU	–	–	–	–	–	–	–	–	–	–
Val	UAC	–	–	–	–	–	–	–	–	–	–
S = 60	Lys	Met	Phe	Pro	Ser	Stop codon	Thr	Trp	Tyr	Val	
Asp	GUC	CUU-1	–	GAA-7	–	GCU-26	UCA-2	–	–	GUA-1	–
		UUU-2	–	–	–	–	–	–	–	–	–
		–	–	UGG-2	–	UCA-1	–	–	–	–	–
Ile	GAU	–	–	–	GCU-52	–	–	–	–	–	–
	CUU-1	CAU-44	GAA-5	UGG-14	–	UCA-2	UGU-5	–	GUA-1	UAC-2	–
	UUU-1										
Leu	UAA	CUU-3	CAU-19	GAA-7	AGG-1	UGA-2	UCA-4	GGU-1	CCA-14	GUA-1	GAC-1
		UUU-13			CGG-1			GGU-2			GAC-4
					GGG-1			UGU-6			UAC-13
					UGG-11						
Lys	UUU	–	–	–	–	GUA-1	–	–	–	–	–
Thr	GGU	UUU-2	–	–	–	UCA-2	–	–	–	–	–
Val	UAC	–	–	–	–	ACU-2	–	–	–	–	–
					GCU-4						

The first row shows amino acids and stop codons, and the second column shows corresponding isolated tRNA anticodons. Numbers represent how many links each isolated tRNA makes with others among different groups in the first column. Anticodons are shown in red, and corresponding isolated tRNA anticodons in blue. Dashes represent a lack of connection between two groups

synonymous anticodons. The largest connected number of one isolated tRNA sequence is 58, accounting for 58/163 of the Ile group. In other words, this isolated tRNA ($tRNA_{UAA}^{Leu}$) still has 58 connections with tRNAs of the Ile group. The second and third connected numbers are 52 and 46, accounting for 52/294 in the Ser group and 46/210 in the Ala group.

With the growth of the similarity degree S , more and more tRNAs lose their connections and become isolated nodes in their group, but some of them are also connected with tRNAs in the other groups. In Table 2, when the similarity degree $S = 65$, there are 19 isolated tRNA sequences that still connect with other tRNAs of different groups, and the largest connection number is 55, which occurs in the Leu group and the Pro group. One isolated tRNA of the Asn group connects with 13 amino acids and one stop codon groups, and the other isolated tRNA connection numbers are all lower than 4. In Table 3, when the similarity degree $S = 70$, more and more tRNAs become isolated nodes in their group, but 18 isolated tRNA sequences still have connections with other tRNAs. Three isolated tRNA sequences, $tRNAs_{UUC}^{Glu}$, $tRNAs_{UAA}^{Leu}$, and $tRNAs_{GAU}^{Ile}$, have very high connections with other groups, which still connect with 43, 40, and 36 links to $tRNAs_{UUG}^{Gln}$, $tRNAs_{UGG}^{Pro}$, and $tRNAs_{GCU}^{Ser}$, respectively. Three isolated tRNA sequences have connections with tRNAs for synonymous anticodons.

Table 2 The isolated tRNAs in one group connect with the tRNAs in other groups under the similarity degree $S = 65$

$S = 65$	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu
Arg	UCU	–	–	–	GCA-1	–	–	–	–	–	–
Asn	GUU	UGC-2	–	GUC-3	GCA-2	UUG-2	UUC-2	UCC-2	–	–	UAA-2 UAG-2
Asp	GUC	–	–	–	–	–	–	–	–	–	–
	GUC	UGC-1	UCG-2	–	–	–	UUC-6	–	GUG-3	–	–
Gln	CUG	–	–	–	–	–	–	–	–	–	–
Glu	UUC	–	–	–	–	UUG-3	–	–	–	–	–
Gly	UCC	–	–	–	–	–	UUC-1	–	–	–	–
Ile	GAU	–	–	–	–	–	–	–	–	–	–
	UAU	–	UCG-1	–	GCA-2	–	UUC-1	–	GUG-1	–	–
Leu	UAA	UGC-4	–	–	GCA-1	UUG-1	UUC-2	UCC-4	GUG-4	–	–
	UAA	UGC-1	ACG-4	GUC-1	–	–	–	GCC-1	–	GAU-8	–
	UAG	–	–	–	–	UUG-1	–	–	GUG-1	–	–
Lys	CUU	–	–	–	–	–	–	–	–	–	–
Ser	GCU	–	UCU-1	–	–	–	–	–	–	–	–
	GCU	–	–	–	–	–	–	–	GUG-4	–	–
	GCU	–	–	–	GCA-1	–	–	–	–	–	–
Thr	UGU	–	UCU-1	–	–	–	–	UCC-2	–	–	UAG-2
	UGU	–	–	–	–	–	–	–	–	–	–
Val	UAC	–	–	–	–	–	–	–	–	–	–

$S = 65$	Lys	Met	Phe	Pro	Ser	Stop codon	Thr	Trp	Tyr	Val
Arg	UCU	–	–	–	GCU-4	–	–	–	–	–
Asn	GUU	UUU-1	CAU-1	–	UGG-3	–	UCA-2	UGU-2	–	GUA-2 UAC-2
Asp	GUC	–	–	–	GCU-3	–	–	–	–	–
	–	–	–	–	–	–	–	–	–	–
Gln	CUG	–	–	–	–	–	–	–	GUA-2	CAC-1
Glu	UUC	–	–	–	–	–	–	–	–	–
Gly	UCC	–	–	–	–	–	–	–	–	–
Ile	GAU	–	–	–	GCU-42	–	–	–	–	–
	UAU	–	CAU-5	–	–	–	UGU-1	–	–	–
Leu	UAA	–	CAU-1	–	UGG-55	UCA-3	–	–	GUA-3	UAC-1
	UAA	–	CAU-2	–	UGG-1	UCA-1	–	CCA-3	–	GAC-1
	UAG	–	–	–	–	–	UGU-1	–	–	UAC-1
Lys	CUU	–	–	–	UGG-1	–	UGU-1	–	–	UAC-2
Ser	GCU	–	–	–	–	–	–	–	–	–
	GCU	–	–	–	–	–	–	–	–	–
	GCU	–	–	–	–	–	–	–	–	–
Thr	UGU	UUU-1	CAU-3	–	UGG-1	UGA-1	–	–	–	UAC-2
	UGU	UUU-6	–	–	–	–	–	–	–	–
Val	UAC	–	–	–	ACU-2	–	–	–	–	–
	–	–	–	–	GCU-4	–	–	–	–	–

19 isolated tRNA sequences (their anticodons are UCU, GUU, GUC, GUC, CUG, UUC, UCC, GAU, UAU, UAA, UAA, UAG, CUU, GCU, GCU, GCU, UGU, UGU, and UAC) in the second column still have connections with other tRNAs among different groups. The highest connection is 55 between $tRNA_{UAA}^{Leu}$ and $tRNA_{UGG}^{Pro}$ and 2 isolated tRNA sequences GUU and UAC, have synonymous anticodons, which can be connected with UAA, UAG and ACU, GCU, respectively

When $S = 90$, as Table 4 shows, we find that there are only 8 isolated tRNAs which still connect with others and 75% of these differ in only one base.

Figure 2 illustrates the diverse relationships of the phylogenetic trees of 21 groups of tRNA families. To eliminate the influence of redundant information, we deleted the polyA tail of 3' terminus and the variable stem. Thus the length of each tRNA sequence is 76. Each branch of a tree represents one amino acid or stop codon group. As Fig. 2a and b show, when the similarity degree $S < 60$, the phylogenetic trees are mainly made up of clusters.

Table 3 The isolated tRNAs in one group connect with the tRNAs in other groups under the similarity degree $S = 70$

$S = 70$	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu
Asn	GUU	UGC-1	—	—	GUC-1	GCA-1	UUG-2	—	UCC-2	—	—
Cys	ACA	—	—	—	—	—	CUG-1	—	ACC-1	—	—
	GCA	—	UCU-1	—	—	—	—	—	ACC-1	GUG-2	—
								UCC-2			
Gln	CUG	—	—	GUU-1	—	ACA-1	—	—	ACC-1	GUG-3	—
						GCA-1			GCC-1		
Glu	UUC	—	—	—	—	—	UUG-43	—	GCC-1	—	—
His	GUG	—	—	—	GUC-1	—	—	—	—	—	—
	GUG	—	CCU-1	—	—	GCA-1	—	—	—	—	CAG-1
Ile	GAU	—	—	—	—	—	—	—	—	—	—
	GAU	UGC-2	UCU-1	GUU-6	—	—	—	—	—	GUG-1	—
	UAU	—	—	—	—	—	—	—	—	—	—
Leu	UAA	—	—	—	—	—	—	—	—	—	—
Met	CAU	—	—	—	—	—	—	—	—	—	UAA-1
Phe	GAA	UGC-1	—	—	—	—	—	—	—	—	GAU-6
Pro	UGG	—	—	GUU-1	GUC-1	GCA-1	UUG-2	—	UCC-1	UGU-1	—
Stop codon	UCA	UGC-1	UCG-16	—	—	—	—	—	UCC-2	—	—
Val	UAC	UGC-8	ACG-1	—	—	—	—	UUC-3	—	—	—
	UAC	—	—	—	—	—	—	—	UCC-2	—	—
	UAC	—	—	—	—	—	—	—	—	—	—

$S = 70$	Lys	Met	Phe	Pro	Ser	Stop codon	Thr	Trp	Tyr	Val
Asn	GUU	—	—	UGG-2	—	UCA-2	UGU-1	—	GUA-2	UAC-2
Cys	ACA	—	CAU-1	—	—	—	—	CCA-1	—	—
	GCA	—	—	GAA-1	—	—	UCA-4	UGU-1	—	—
Gln	CUG	CUU-9	CAU-2	GAA-10	—	—	—	AGU-1	CCA-3	—
		UUU-6						CGU-2		
								UGU-11		
Glu	UUC	—	—	—	UGG-1	—	—	—	—	—
His	GUG	—	—	—	—	—	—	—	—	—
	GUG	—	—	—	—	—	—	—	—	—
Ile	GAU	—	—	—	—	GCU-36	—	—	—	—
	GAU	—	CAU-7	GAA-2	—	—	UCA-1	—	—	—
	UAU	—	CAU-1	—	—	—	—	—	—	—
Leu	UAA	—	—	—	UGG-40	—	UAC-1	—	—	—
Met	CAU	—	—	—	—	—	—	—	—	—
Phe	GAA	—	—	—	—	—	—	—	—	—
Pro	UGG	—	—	—	—	—	—	—	GUA-2	UAC-2
Stop codon	UCA	—	—	—	—	—	—	—	—	—
Val	UAC	—	—	—	—	—	UCA-1	—	—	—
	UAC	—	—	—	—	—	—	—	—	—
	UAC	—	—	—	—	ACU-2	—	—	—	—
					GCU-3					

18 isolated tRNA sequences (their anticodons are GUU, ACA, GCA, CUG, UUC, GUG, GUG, GAU, GAU, UAU, UAA, CAU, GAA, UGG, UCA, UAC, UAC, and UAC) in the second column still connect with other tRNAs among different groups. The Asn, Cys, Gln, Ile, and Pro groups have relatively high connections with different amino acids and stop codon groups, and others have only several connections. Three isolated tRNA sequences have very high connections, namely 43, 40, and 36 corresponding to $tRNAs_{UUC}^{UUC}$, $tRNAs_{UAA}^{UAA}$, and $tRNAs_{GCU}^{GAU}$, respectively, and three isolated (GCA, CUG, and UAC) tRNA sequences have synonymous anticodons

For example, when $S = 50$, two clusters, ((Glu,(Pro,(Gln:His))):(Val,(Asp:Gly))) and (stop codon,(Cys:Trp)) form into a large cluster, and Ile, Ala, and Met stay in isolated branches. By comparing Fig. 2a with Fig. 2b, we find the clusters of the phylogenetic tree in Fig. 2b is more compact, and some amino acids add to the clusters and become larger, while amino

Table 4 The isolated tRNAs in one group connect with the tRNAs in other groups under the similarity degree $S = 90$

Anticodon	Connected anticodon	Total number
UGC(Ala)	U <i>AC</i> (Val)	3
GUC(Asp)	GCC(Gly)	1
GCC(Gly)	GUC(Asp)	1
UCC(Gly)	UGG(Pro)	4
UUU(Lys)	UU <i>A</i> (Stop codon)	2
CAU(Met)	GAU(Ile)	4
UGG(Pro)	UAA(Leu)	1
CAU(Met)	GAU(Ile)	3

Eight isolated tRNA sequences (UGC, GUC, GCC, UUU, CAU, UGG, and CAU) in the first column, which still connect with other tRNAs among different groups in the second column. Italics indicate a mutation site. The total number represents how many links one isolated tRNA sequence still connects with other tRNAs between the groups in the first column and the second column

Fig. 2 The phylogenetic trees of 20 amino acids and one stop codon are constructed by different similarity degree **a** $S = 50$, **b** $S = 56$, **c** $S = 60$, **d** $S = 64$, **e** $S = 66$, and **f** $S = 68$

acids alter their locations. With the similarity degree S increasing, Fig. 2c shows that more isolated branches emerge in the bigger cluster. By comparing Fig. 2a–f, we can find that: (1) the Ala occupies a basal clade and does not change with the similarity degree S in all phylogenetic trees; (2) when $68 > S \geq 60$, the phylogenetic trees show a regular ladder-like ranking; a few groups of amino acids still stay in the same rank, such as Leu and Met, stop codon, and Trp; (3) groups of Leu and Met and Ser, (Phe,(Asp,(Cys,(stop codon: Trp)))) always keep their evolutionary rank unchanged with all S ; (4) when $S > 68$, evolutionary distance tend to zero.

4 Discussions

We constructed 22 groups of networks and four phylogenetic trees of tRNAs and analyzed the anticodon feature of each isolated tRNA that connects with other groups, as well as the phylogenetic trees. We find that point mutation plays an important role in the evolution of two different amino acids, which is consistent with the Saks's [13] and Wei's [17, 22] results. We also find that some isolated tRNA sequences in the 21 networks still connect with other networks, which reflects "crossover" evolutionary relationships among these 21 groups, because they belong to different branches in the phylogenetic trees.

Phylogenetic trees of 20 amino acids and one stop codon family provide insights into the relationships between their lineages. We find that the evolutionary relationships of all tRNAs appear in "stages" because their relationships in the phylogenetic trees occur in a regular, ladder-like order and tend to stabilize when S reaches a relatively large value. A few groups of tRNAs retain close relationships, with many of them only one base apart at the same anticodon sites, which implies that their evolutionary relationships may be mutated by the same ancestor sequences by point mutation. In this work, the complementary duplication mechanism [22], another accepted evolutionary origin of modern tRNAs, is not evident.

Acknowledgements This paper was supported by Guangxi Natural Science Foundation (No. 0728003), National Natural Science Foundation of China (Nos. 10662002, 10865001), the Graduate Student Innovation Program of Guangxi Zhuang Autonomous Region (Grant No. T32070), and the 973 Program (2010CB328204).

References

1. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. *Science* **286**, 509–511 (1999)
2. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998)
3. Strogatz, S.H.: Exploring complex networks. *Nature* **410**, 268–276 (2001)
4. Vázquez, A., Pastor-Satorras, R., Vespignani, A.: Large-scale topological and dynamical properties of the Internet. *Phys. Rev. E* **65**, 066130 (2002)
5. Albert, R., Jeong, H., Barabási, A.L.: Diameter of the world-wide web. *Nature* **401**, 130 (1999)
6. Barabási, A.L., Albert, R., Jeong, H.: Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A* **281**, 69–77 (2000)
7. Jeong, H., Mason, S.P., Barabási, A.L., Oltvai, Z.N.: Lethality and centrality in protein networks. *Nature* **411**, 41–42 (2001)
8. Goh, K., Cusick, M.E., Valle, D., Childs, B., Vidal, M., Barabási, A.L.: The human disease network. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8685–8690 (2007)
9. Chen, B., Xian, Y., Wei, F.: The evolutionary network study of tRNA sequences. *Acta Biophysica Sinica* **25**, 117 (2009). In Chinese
10. Wei, F., Li, S., Ma, H.: Network of tRNA gene sequences. *J. Shanghai Jiaotong Univ. (Sci.)* **13**, 611–616 (2008)

11. Song, N., Joseph, J.M., Davis, G.B., Durand, D.: Sequence similarity network reveals common ancestry of multidomain proteins. *Comput. Biol.* **4**, 1–19 (2008)
12. Eigen, M., Winkler-Oswatitsch, R.: Transfer-RNA: the early adaptor. *Naturwissenschaften* **68**, 217–228 (1981)
13. Saks, M.E., Sampson, J.R., Abelson, J.: Evolution of a transfer RNA gene through a point mutation in the anticodon. *Science* **279**, 1665–1667 (1998)
14. Rodriguez-Vargas, A.M., Fajardo, J.E., Ramirez, B.C.: Phylogeny of transfer RNA. *Orig. Life* **14**, 547–555 (1984)
15. Soto, M.A., Toha, J.: Phylogenetic tree of tRNAs using a simple algorithm. *Orig. Life Evol. Biosph.* **20**, 161–166 (1990)
16. Farias, S.T., Guimaraes, R.C.: Aminoacyl-tRNA synthetase classes and groups in prokaryotes. *J. Theor. Biol.* **250**, 221–229 (2008)
17. Wei, F., Meng, M., Li, S., Ma, H.: Comparing two evolutionary mechanisms of modern tRNAs. *Mol. Phylogenet. Evol.* **38**, 1–11 (2006)
18. Sprinzl, M., Horn, C., Brown, M., Steinberg, S.: Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **26**, 148–153 (1998)
19. Sprinzl, M., Vassilenko, K.S.: Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **33**, 139–140 (2005)
20. Jühling, F., Mörl, M., Hartmann, R.K., Sprinzl, M., Stadler, P.F., Pütz, J.: tRNADB 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.* **37**, 159–162 (2009)
21. Saitou, N., Nei, M.: The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987)
22. Wei, F., Li, S., Ma, H.: Computer simulation of tRNA evolution. *J. Phys. A: Math. Theor.* **42**, 345101 (2009)