Complete Mitochondrial Genome Sequence Analysis and Phylogenetic Location Determination of *Hydropsyche Fryeri* (Insecta: Trichoptera)

Yimin Li
Beibu Gulf University
Honglin Qin
Beibu Gulf University
Xifa Zhong
Beibu Gulf University
Jingcai Huang
Guangxi University
Yujun Wang
Beibu Gulf University
Hong Wang (3212641322@qq.com)
Beibu Gulf University
Yu Lan
Beibu Gulf University
Ruxia Zheng
Beibu Gulf University
Xiaochan Huang
Beibu Gulf University

Research Article

Keywords: Hydropsyche Fryeri, Trichopteridae, mitochondrial genome sequence

Posted Date: April 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-437236/v1

Abstract

Hydropsyche fryeri belongs to the Trichopteridae family and builds nests in clean and unpolluted streams using stones. It also can be used as an indicator of water quality. Here, we describe the complete mitochondrial genome sequence of *Hydropsyche fryeri*. The mitochondrial genome is 15,676 bp long and contains 13 protein-coding genes, 22 tRNAs, 2 rRNAs and an AT-rich control region. Phylogenetic tree analysis shows that *Hydropsyche fryeri* is more closely related to the family Hydroptera than other Trichoptera.

Introduction

Hydropsyche fryeri is a member of the Trichopteridae family. It has a fusiform head with white stripes and a brown or light-green body of about 1.9 ± 0.4 cm (Fig. 1). *Hydropsyche fryeri* live in clear rivers and can be used to detect water quality [1] and heavy metal [2] pollution. *Hydropsyche fryeri* has extremely strict requirements for water temperature and cannot survive below 20°C. However, *Hydropsyche Fryeri* can tolerate a turbulent water environment.

Most of the previous classification methods for insects of Trichoptera are morphological classification [3] and the phylogenetic data at the molecular level of the caddisfly still requires further research.

Mitochondrial DNA (mtDNA) is a circular structure DNA that exists in the mitochondria of eukaryotic cells [4]. Mitochondrial DNA is inherited by the maternal line and its primary structure shows significant inter- and intraspecies variation and thus can be used for molecular classification [5]. In general, metazoan mitochondrial genomes consist of a non-coding sequence called the control region (CR) and 37 genes, including 13 protein-coding genes (PCGs), 22 tRNAs and 2 rRNAs [6]. Mitochondrial cytochrome oxidase (COX 1) and (COX 2) [7]and the tRNA gene sequences[8] are used for the classification of species[9] and infraspecific category identification [10]. In this study, we determined the mitochondrial DNA sequence of *Hydropsyche fryeri* and analysed the phylogenetic relationships between this species and other related species.

Results

The complete mitochondrial genome is 15,676 bp long and includes 13 protein-coding genes, 2 rRNAs, 22 tRNAs and an AT-rich control region (D-loop). The genome consists of 42.29% A, 39.92% T, 11.16% C and 6.63% G bases. The control region, located between tRNA-Ile and 12S-rRNA, is 631 bp long and has an A + T content of 82.21% (Table 2). Of the PCGs and tRNAs, NAD5 (1720 bp) was the longest and tRNA-Ser (60 bp) was the shortest (Table 1). Thirteen protein-coding genes had ATN as the starting codon. NAD2 started with ATA, COX1, ATP6, COX3, NAD4, NAD4L and COB started with ATG, and COX2, ATP8, NAD3, NAD5 and NAD1 started with ATT. The stop codon of COX1 and NAD5 is the incomplete stop codon T, TAA is the stop codon for the remaining coding proteins (Table 1), and Mitochondrial genes are circular structures (Fig. 2). Based on the results obtained from the relative synonymous codon usage analysis

(RSCU), the values were higher for GCT (Ala), CGA (Arg), GGA (Gly), TTA (Teu), CCT (pro), TCA (Ser), AGA (Ser) and TCT (Ser) and lower for CTC (Leu), CTG (Leu), AGC (Ser), ACG (Thr) and TGG (Trp) (Table 3, Fig. 3).

Table 1
Hydropsyche Fryeri mitochondrial gene annotation.

Gene	Position (bp)	Size (bp)	A + T Percent	Intergenic nucleotide (bp)	Inferred Initiation Codon	Inferred Termination Codon
tRNA-Ile	1–69	69	84.06%	10		
tRNA-Gln	80–147	68	85.29%	7		
tRNA-Met	155–222	68	82.35%	17		
NAD2	240–1220	981	86.85%	3	ATA	TAA
tRNA-Trp	1224–1291	68	89.71%	-8		
tRNA-Cys	1284–1348	65	87.69%	0		
tRNA-Tyr	1349–1417	69	84.06%	7		
COX1	1425–2961	1537	73.52%	0	ATG	T
tRNA-Leu	2962–3030	69	81.16%	0		
COX2	3031–3711	681	76.51%	6	ATT	TAA
tRNA-Lys	3718–3787	70	72.86%	4		
tRNA-Asp	3792–3856	65	87.69%	0		
ATP8	3857–4024	168	89.88%	-7	ATT	TAA
ATP6	4018–4692	675	80.3%	13	ATG	TAA
COX3	4706–5497	792	76.01%	-1	ATG	TAA
tRNA-Gly	5497–5560	64	92.19%	0		
NAD3	5561–5914	354	82.49%	13	ATT	TAA

Gene	Position (bp)	Size (bp)	A + T Percent	Intergenic nucleotide (bp)	Inferred Initiation Codon	Inferred Termination Codon
tRNA-Ala	5928–5992	65	89.23%	-1		
tRNA-Arg	5992–6055	64	82.81%	-1		
tRNA-Asn	6055–6123	69	78.26%	0		
tRNA-Ser	6124–6183	60	85.00%	26		
tRNA-Glu	6210–6275	66	86.36%	1		
tRNA-Phe	6277–6342	66	84.85%	0		
NAD5	6343–8062	1720	81.22%	1	ATT	T
tRNA-His	8064–8126	63	88.89%	-1		
NAD4	8126–9451	1326	80.77%	-7	ATG	TAA
NAD4L	9445–9735	291	87.97%	2	ATG	TAA
tRNA-Thr	9738–9803	66	87.88%	30		
NAD6	9834–10343	510	88.24%	3	ATT	TAA
COB	10347–11477	1131	77.01%	-2	ATG	TAA
tRNA-Ser	11476–11542	67	82.09%	145		
tRNA-Pro	11688–11752	65	83.08%	4		
NAD1	11757–12692	936	79.38%	1	ATT	TAA
tRNA-Leu	12694–12760	67	86.57%	0		
16S-rRNA	12761–14165	1405	85.55%	0		

Gene	Position (bp)	Size (bp)	A + T Percent	Intergenic nucleotide (bp)	Inferred Initiation Codon	Inferred Termination Codon
tRNA-Val	14166–14231	66	86.36%	0		
12S-rRNA	14232–15045	814	88.21%	0		
D-loop	15046–15676	631	97.78%	0		

Table 2
Total gene and base content of *Hydropsyche fryeri*.

species	Total gene size (bp)	AT content(%)	A Size (bp)	T Size (bp)	C Size (bp)	G Size (bp)
Hydropsyche fryeri	15676	82.21%	6629	6258	1750	1039

Genome annotation showed that tRNA-Trp, ATP8, COX3, tRNA-Ala, tRNA-Arg, tRNA-His, NAD4 and COB overlapped with their adjacent genes, and tRNA-Trp and tRNA-Cys had the highest degree of overlap. There were intervals between tRNA-Ile, tRNA-Gln, tRNA-Met, NAD2, tRNA-Tyr, COX2, tRNA-Lys, ATP6, NAD3, tRNA-Ser, tRNA-Glu, NAD5, NAD4L, tRNA-Thr, NAD6, tRNA-Ser, tRNA-Pro, and NAD1, and the interval between tRNA-Ser and the adjacent tRNA-Pro was the largest (Table 1).

Of the 22 tRNAs, only tRNA-Ser (AGA) at position 6124–6183 had a structure without a TWC arm, and an A + T content of 85.0%, while the other 21 tRNAs had typical clover structures with an A + T content of 72.9–92.2% (Table 1). Eight of the

Table 3
Protein coding gene codons and relative synonymous codon usage (RSCU) of *Hydropsyche Fryeri*.

Amino acid	Codon	Number	RSCU	Amino acid	Codon	Number	RSCU
Ala	GCT	54	2.51	Lys	AAA	113	1.77
Ala	GCA	21	0.98	Lys	AAG	15	0.23
Ala	GCC	8	0.37	Met	ATA	270	1.84
Ala	GCG	3	0.14	Met	ATG	24	0.16
Arg	CGA	31	2.75	Phe	TTT	369	1.79
Arg	CGT	14	1.24	Phe	TTC	43	0.21
Arg	CGC	0	0.00	Pro	CCT	70	2.46
Arg	CGG	0	0.00	Pro	CCA	29	1.02
Asn	AAT	248	1.84	Pro	CCC	15	0.53
Asn	AAC	21	0.16	Pro	CCG	0	0.00
Asp	GAT	50	1.75	Ser	TCA	108	2.55
Asp	GAC	7	0.25	Ser	AGA	91	2.15
Cys	TGT	20	1.60	Ser	TCT	88	2.08
Cys	TGC	5	0.40	Ser	AGT	24	0.57
Gln	CAA	49	1.85	Ser	TCC	22	0.52
Gln	CAG	4	0.15	Ser	AGG	5	0.12
Glu	GAA	66	1.81	Ser	AGC	1	0.02
Glu	GAG	7	0.19	Ser	TCG	0	0.00
Gly	GGA	100	2.26	Stp	TAA	11	2.00
Gly	GGT	49	1.11	Stp	TAG	0	0.00
Gly	GGG	22	0.50	Thr	ACA	61	1.89
Gly	GGC	6	0.16	Thr	ACT	58	1.80
His	CAT	46	1.56	Thr	ACC	8	0.25
His	CAC	13	0.44	Thr	ACG	2	0.06
Ile	ATT	418	1.81	Trp	TGA	86	1.95
Ile	ATC	44	0.19	Trp	TGG	2	0.04

Amino acid	Codon	Number	RSCU	Amino acid	Codon	Number	RSCU
Leu	TTA	468	4.81	Tyr	TAT	147	1.83
Leu	CTT	42	0.43	Tyr	TAC	14	0.17
Leu	CTA	41	0.42	Val	GTT	71	2.12
Leu	TTG	27	0.28	Val	GTA	53	1.58
Leu	CTC	5	0.05	Val	GTC	5	0.15
Leu	CTG	1	0.01	Val	GTG	5	0.15

22 tRNAs had base mismatches: the D-arm of tRNA-Gln (TTG) had a T-G mismatch of two bases; the D-arm of tRNA-Tyr (GTA) had a G-A mismatch; a pair of T-T mismatches occurred on the forearm of tRNA-Leu (TAA); a pair of T-G base mismatches occurred on the D-arm of tRNA-Gly (TCC); a pair of T-T mismatches occurred on the anticodon arm of tRNA-Lys (AGA); a pair of T-T mismatches occurred on the D-arm of tRNA-His and tRNA-Val (TAC) (Fig. 4).

We compared other families of Trichoptera with Lepidoptera, which are closely related to Trichoptera. We selected eight families (Pryganeidae, Limnephilidae, Apataniidae, Uenoidae, Pryganopsychidae, Sericostomatidae, Leptoceridae and Hydropsychidae) from Trichoptera and 22 species of Hepialidae from Lepidoptera for the construction of an evolutionary tree. *Hydropsyche fryeri* was most closely related to *Hydropsyche orris*, *Hydropsyche simulans* and *Hydropsyche pellucidula* of the genus Arctopsyche of Hepialidae and was relatively closely related to *Potamyia flava* and *Hydromanicus wulaianus* of Hydropsychidae. However, *Hydropsyche fryeri* was distantly separated from *Sericostoma personatum* of Sericostomatidae and *Triaenodes tardus* of Leptoceridae.

Discussion

The basic composition of the *Hydropsyche fryeri* mitochondrial genome is consistent with the common composition of metazoans. The entire genome is 15,676 bp long and contains both overlapping and spaced gene segments. There is one structurally abnormal tRNA of 22 tRNAs while the remaining 21 have normal clover structures.

The evolutionary tree constructed by the maximum likelihood method showed that the genomic sequence and the protein-coding sequences were consistent (Fig. 5). Therefore, *Hydropsyche fryeri* was identified as being a member of Hydropsychidae, which forms a sister population with Pryganeidae and Limnephilidae.

Materials And Methods

Samples

This study was conducted without harming protected or endangered species, and all research activities were authorized. Samples were collected from of Shiwandashan River system (21°49' 33" N, 107°59' 119" E) in Fangchenggang City, Guangxi Zhuang Autonomous Region (China). The collected samples were transported through a cold chain to the Key Laboratory of Biodiversity, College of Marine Sciences, Beibu Gulf University (China). The samples were dissected in vivo under normal saline to remove the intestinal tract and head. The specimen has been deposited in Ocean college marine specimen showroom of Beibu Gulf University (Voucher No. BBGC 00014).

Mitochondrial DNA Extraction

Total mitochondrial DNA was extracted according to the method of Roehrdanz (1997) with partial modifications. Cells were disrupted, proteinase K and RNase were added for enzymatic digestion in a water bath for 5 h at 56 ° C. DNA was extracted in a phenol: chloroform: isoamyl alcohol (25:24:1) solution and then centrifuged with a cold isopropanol precipitate and 70% ethanol wash with a dissolved TE buffer.

Gel Electrophoresis

The mitochondrial genome was obtained by gel electrophoresis and sequenced by high-throughput sequencing [11]. A 1.0% agarose gel was prepared to separate the total mitochondrial genes, and the electrophoresis conditions were a voltage of 120 V and current of 40 Ma for 20 min. The gel was observed on a Tanon3500 gel imaging system (Shanghai Tianneng Technology Co., Ltd, China) and compared with a 10,000 bp DNA marker to preliminarily identify whether the DNA size was in the insect mtDNA size range.

DNA Recovery And Purification

DNA was cut from agarose gels, weighed in 2 ml centrifuge tubes, and the DNA was recovered according to the Tiangen universal DNA purification recovery kit protocol (Tiangen, Beijing, China).

Table 4
22 mitochondrial genome sequences downloaded from the NCBI.

ID	ORGANISM	ORDER	FAMILY	GENUS
KF717094	*Eubasilissa regina*	Trichoptera	Phryganeidae	Eubasilissa
NC_023374	*Eubasilissa regina*	Trichoptera	Phryganeidae	Eubasilissa
NC_039714	*Phryganea cinerea*	Trichoptera	Phryganeidae	Phryganea
NC_044710	*Limnephilus hyalinus*	Trichoptera	Limnephilidae	Limnephilus
NC_026219	*Limnephilus decipiens*	Trichoptera	Limnephilidae	Limnephilus
NC_036004	*Anabolia bimaculata*	Trichoptera	Limnephilidae	Anabolia
NC_043770	*Hydatophylax nigrovittatus*	Trichoptera	Limnephilidae	Hydatophylax
KF756944	*Apatania sp. YW-2014*	Trichoptera	Apataniidae	Apatania
KP455291	*Thremma gallicum*	Trichoptera	Uenoidae	Thremma
NC_043771	*Phryganopsyche latipennis*	Trichoptera	Phryganopsychidae	Phryganopsyche
KP455290	*Sericostoma personatum*	Trichoptera	Sericostomatidae	Sericostoma
NC_039659	*Triaenodes tardus*	Trichoptera	Leptoceridae	Triaenodes
NC_036951	*Hydropsyche orris*	Trichoptera	Hydropsychidae	Hydropsyche
NC_036950	*Hydropsyche simulans*	Trichoptera	Hydropsychidae	Hydropsyche
NC_029246	*Hydropsyche pellucidula*	Trichoptera	Hydropsychidae	Hydropsyche
NC_036955	*Cheumatopsyche analis*	Trichoptera	Hydropsychidae	Cheumatopsyche
NC_036954	*Cheumatopsyche campyla*	Trichoptera	Hydropsychidae	Cheumatopsyche
NC_036952	*Cheumatopsyche speciosa*	Trichoptera	Hydropsychidae	Cheumatopsyche
NC_043769	*Cheumatopsyche brevilineata*	Trichoptera	Hydropsychidae	Cheumatopsyche
NC_036953	*Potamyia flava*	Trichoptera	Hydropsychidae	Potamyia
NC_036156	*Hydromanicus wulaianus*	Trichoptera	Hydropsychidae	Hydromanicus
NC_044770	*Thitarodes damxungensis*	Lepidoptera	Hepialidae	Thitarodes

Phylogenetic trees

The mitochondrial genome of *Hydropsyche fryeri* was sequenced and assembled using Illumina high-throughput sequencing technology and Spades v.3.5.0 software.

The newly sequenced genomes and 22 complete mitochondrial genome sequences close to the *hydropsyche fryer* BLAST results were downloaded from the National Center for Biotechnology Information (Table 4). We used jModelTest2.1.7 (https://code.google.com/p/jmodeltest2/) for the selected sequences of DNA nucleic acid model test [12] and Prottest3.2 (https://code.google.com/p/prottest3/) for the amino acid model test [13]. We selected the AIC (Akaike Information Criterion) *Broussonetia papyrifera* minimum value as the best model and used RAXML 8.1.5 (https://sco.H–its.org/exelixis/web/software/raxml/index.html) and the Maximum Likelihood (ML) method to construct the phylogenetic tree with a bootstrap value set to 1000 [14].

Declarations

Data availability statement

The data that support the findings of this study are openly available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov/, reference number MW413803.

Acknowledgement

We are very grateful to Zhang hang and Li Hai-wei for helping us to collect samples in the waters of Shiwandashan. In addition, we would like to thank the National Natural Science Foundation of China (31760713; 31972873) for financial support.

Author contributions

Y. J. W presented the experimental protocols, J. C. H, X. F. Z performed the experimental work, Y. L processed the experimental data, H. L. Q prepared the picture, and Y. M. L, H. W, Y. L, R. X. Z wrote and revised the article, all authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

References

[1] Li Z-P, Zhou Y-M. An important indicator organism for water quality monitoring – Trichoptera larvae. *Environmental Bulletin of Heilongjiang Province*, 2001, (01): 77-78.

[2] Xue J, An Z-S, Niu C-Y, Lei C-L. Study on monitoring heavy metal pollution in water by aquatic diptera. *Entomological Knowledge*, 2008, (03): 378-383.

[3]Gao Y, Liu S-Y, Yang G, Yao Y-Z, Ren Dong. Research progress of Trichoptera fossils. *Acta Entomologica Sinica*, 2012, 49 (02): 543-555.

[4] Bridge D, Cunningham C W, Schierwater B, et al. Class-level relationships in the phyum cnidarian: evidence from mitochondrial genome structure. *Proceedings of the National Academy of Sciences of United States of America*, 1992, 89(18):8750-8753.

[5] Simon, C. et al. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annu. Rev. Ecol. Evol. Syst*, 2006, 37, 545–579.

[6]Boore J L. Animal mitochondrial genomes, *Nucleic Acids Research*, 1999, 27(8)1767-1780.

[7] Sperling, Felix A. H., Anderson, G. S. and Hickey, D. A. A DNA - based approach to the identification of insect species used for postmortem interval estimation. *Foren. Sci*., 1994, 39 (2):418- 427.

[8] Song, S W, Zhao, T Y, Dong, Y D et al. mtDNA - 16S rRNA sequence of the culex pipiens complex and their phylogenetic analysis. *Acta Zoot axonomica Si nica*, 2002, 27 (4):665-671.

[9] Zhao, MY., Huo, QB. & Du, YZ. Molecular phylogeny inferred from the mitochondrial genomes of Plecoptera with *Oyamia nigribasis* (Plecoptera: Perlidae). *Sci Rep*, 2020, 10, 20955. https://doi.org/10.1038/s41598-020-78082-y.

[10]Huang J-C, Wang X-X, Zhong X-F, et al. Characterization of the complete mitochondrial genome of *Stenopsyche angustata* (Trichoptera, Stenopsychidae). *Mitochondrial DNA B Resour,* 2020, 5(3):3114-3115. Published 2020 Aug 3. doi:10.1080/23802359.2020.1797564.

[11] Zou Y L, Ding Y R, Luo Q C, Chen B. Methods for extracting mosquito mitochondrial genomic DNA. *Chinese Journal of vector biology and control*, 2015, 26 (04): 333-336.

[12] David Posada. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, Volume 25, Issue 7, 1 July 2008, Pages 1253–1256.

[13] Diego Darriba, Guillermo L. Taboada, Ramón Doallo, David Posada; ProtTest 3: fast selection of best-fit models of protein evolution, *Bioinformatics*, Volume 27, Issue 8, 15 April 2011, Pages 1164–1165.

[14]Alexandros Stamatakis; RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, Volume 22, Issue 21, 1 November 2006, Pages 2688–2690.

Figures

Figure 1

Morphology of Hydropsyche fryeri, 0.25 cm.

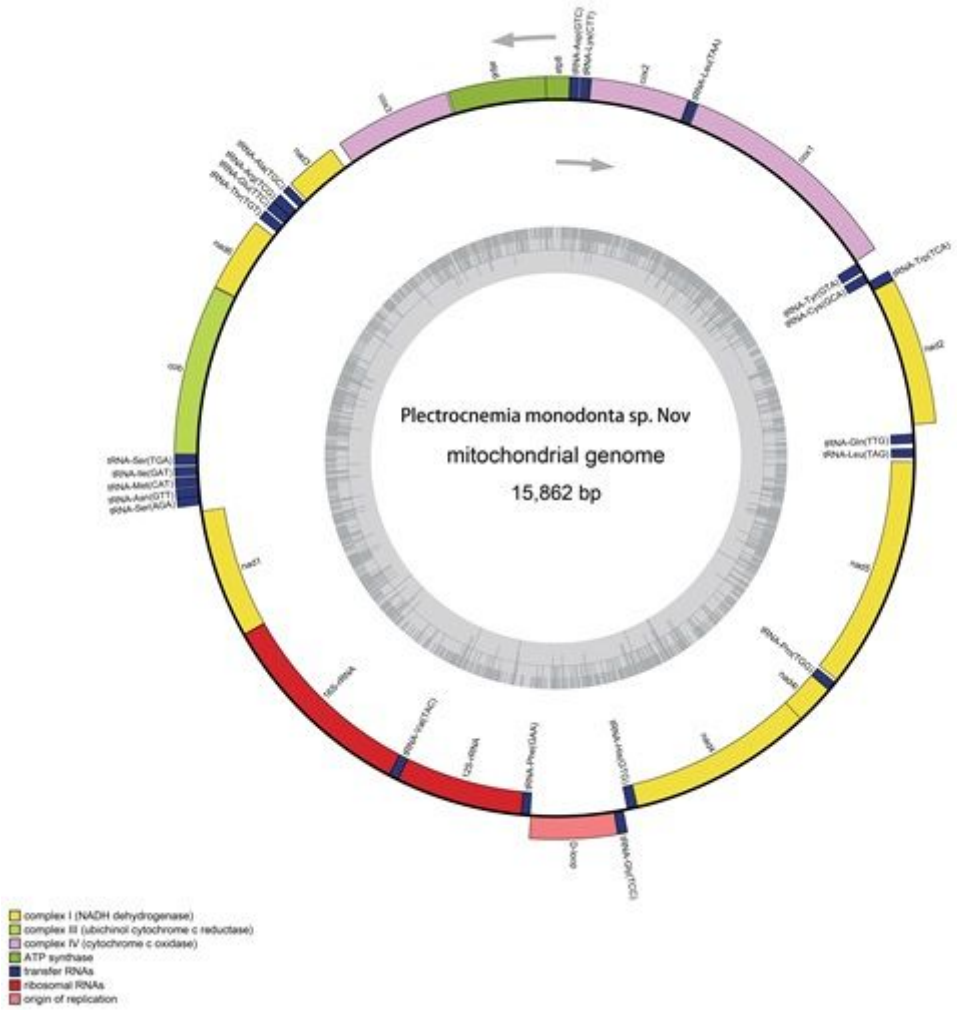

Figure 2

The mitochondrial ring structure of Hydroptyche fryeri.

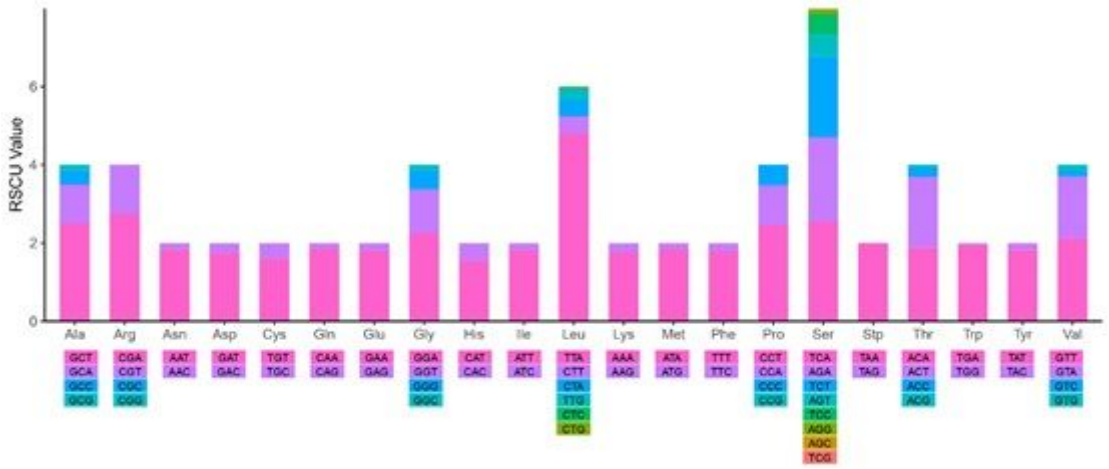

Figure 3

Relative synonymous codon usage analysis of Hydropsyche Fryeri.

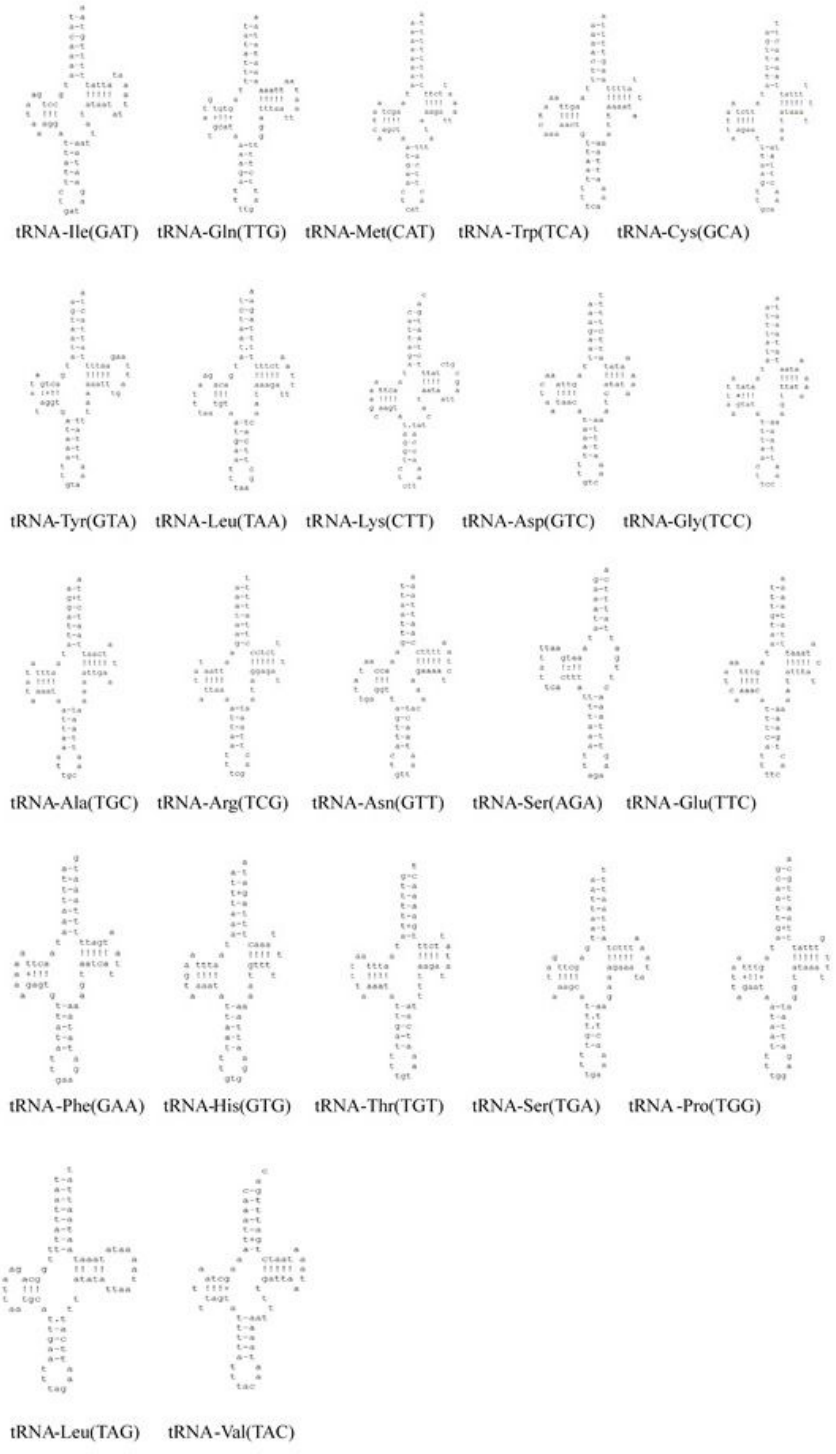

Figure 4

The clover structure of 22 tRNA.

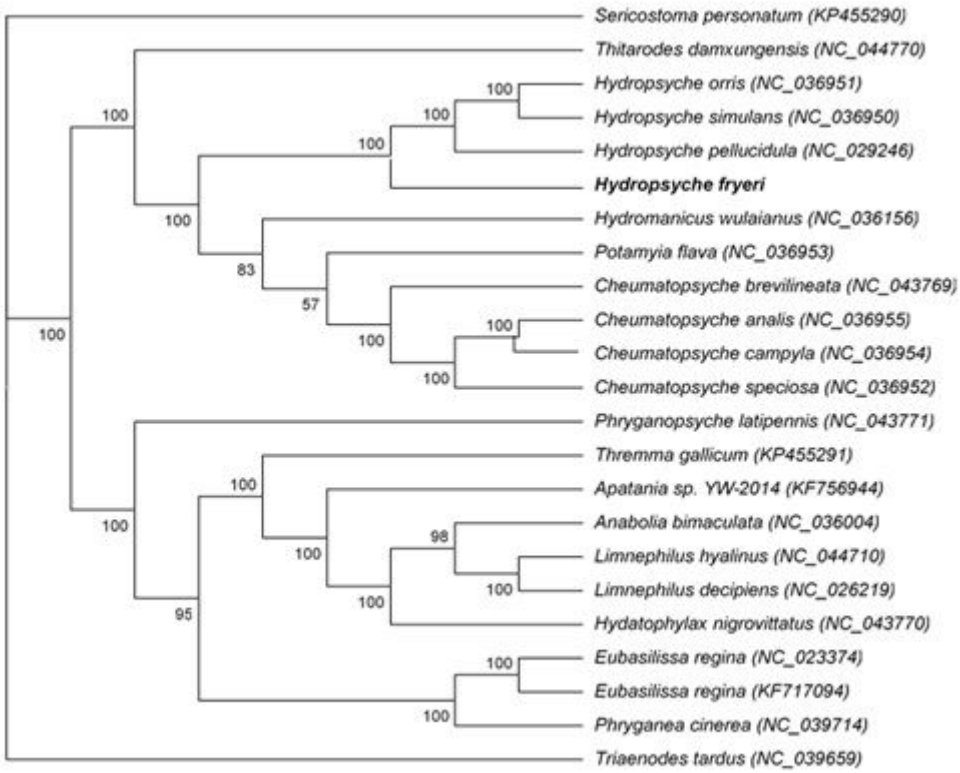

Figure 5

Hydropsyche Fryeri phylogenetic tree.