

~~Supplementary material~~

~~*Journal of Helminthology*~~

~~Sequencing, characterisation and phylogenomics of the complete mitochondrial genome
of *Dactylogyrus lamellatus* (Monogenea: Dactylogyridae)~~

~~Dong Zhang^{1,2}, Hong Zou¹, Shan G. Wu¹, Ming Li¹, Ivan Jakovlić³, Jin Zhang³, Rong Chen³, Gui T. Wang¹,
Wen X. Li¹~~

~~1. Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, P. R. China~~

~~2. University of Chinese Academy of Sciences, Beijing 100049, P. R. China.~~

~~3. Bio-Transduction Lab, Wuhan Institute of Biotechnology, Wuhan 430075, P. R. China~~

~~Corresponding author: Wen X. Li. E-mail: liwx@ihb.ac.cn~~

Table S1. Primers used to amplify and sequence the mitochondrial genome of *Dactylogyrus lamellatus*.

Product sizes are given in bp

Primer name	Primer sequence (5'-3')	Target gene	Size
TXCCOX1F	GGHTGAACHRTWTAYCCHCC	<i>cox1</i>	500
TXCCOX1R	TGRTGRGCYCAWACDAYAMAHCC	<i>cox1</i>	500
ZHCND5F	GARGCGAWGCGADCCMCTAHTCC	<i>nad5</i>	470
ZHCND5R	TGCTTASTADAGHAGABTCC	<i>nad5</i>	470
ZHCND1F	CGTMAGRGCCTADTAAKGTTGG	<i>nad1</i>	701
ZHCND1R	CGVATTYGHRGTHATGWTGYACG	<i>nad1</i>	701
ZHC12SF	CAGTGCBCAGSAATCTCKGWTA	<i>rrnS</i>	487
ZHC12SR	DDTGACGGBCGAHAYGKAC	<i>rrnS</i>	487
ZHCCOX2F	TWRITTDGACDGCRATVATWATG	<i>cox2</i>	271
ZHCCOX2R	CCRCDTWCATYAYKAMAAYATCC	<i>cox2</i>	271
ZHCCOX3F	GBTYTBHTAWTRYTKTTAC	<i>cox3</i>	303
ZHCCOX3R	CDADGRAATRTSAGBRTCA	<i>cox3</i>	303
ZHCCYTBF	TDCCTTVGCDTSAARTKTC	<i>cob</i>	430
ZHCCYTBR	AARAAATMTCRTTYTVKCTT	<i>cob</i>	430
ZHCND4F	CCTAHTBTTCMTKSTGWAGC	<i>nad4</i>	209
ZHCND4R	CTNARVAWTKCTHATCRTVGMFT	<i>nad4</i>	209
ZHCF1	TAGTCACATATGTCTTAG	<i>cox1-rrnS</i>	2167
ZHCR1	TAGCTGATCCAAAACACC	<i>cox1-rrnS</i>	2167
ZHCF2	ATGGTCAAGGTTTGCTAC	<i>rrnS-cox2</i>	563
ZHCR2	TAAATAAACTGAAGGGGC	<i>rrnS-cox2</i>	563
ZHCF3	GACGGATGCAATACCTGG	<i>cox2-nad5</i>	1581
ZHCR3	TGTCTTCAATTAACACACC	<i>cox2-nad5</i>	1581
ZHCF4	CCGGTTCCTATTCTAGCC	<i>nad5-cox3</i>	3341
ZHCR4	CTAAAATTGTAATGCTAC	<i>nad5-cox3</i>	3341
ZHCF5	TCACATGTTGTAATAGGG	<i>cox3-cob</i>	784
ZHCR5	CACACCAAATCCTCCAAC	<i>cox3-cob</i>	784
ZHCF6	TGCCCGGACTTAGTCCTGG	<i>cob-nad4</i>	1203
ZHCR6	AATATACCAATTACACCC	<i>cob-nad4</i>	1203
ZHCF7	GTAGCTTTGTTATTTGGG	<i>nad4-nad1</i>	2506
ZHCR7	ACTACGGTACTGCTGTAC	<i>nad4-nad1</i>	2506
ZHCF8	TTGTTGATTTAGCTCTAG	<i>nad1-cox1</i>	1464
ZHCR8	TTGCAGACTGTATAGTAC	<i>nad1-cox1</i>	1464

Table S2 Nucleotide composition of the protein-coding genes, tRNAs, rRNAs and non-coding region of mitochondrial genomes of *Dactylogyrus lamellatus* and *Tetrancistrum nebulosi*

Regions	Size (bp)	T(U)	C	A	G	AT(%)	GC(%)	AT skew	GC skew
<i>Dactylogyrus lamellatus</i>									
PCGs	9930	45.2	10.2	23.7	20.9	68.9	31.1	-0.312	0.344
1 st codon position	3310	38.3	10.1	27.1	24.5	65.4	34.6	-0.172	0.415
2 nd codon position	3310	47.3	14.7	17.8	20.2	65.1	34.9	-0.452	0.160
3 rd codon position	3310	50.0	5.8	26.2	18.0	76.2	23.8	-0.312	0.510
<i>rrnS</i>	725	34.5	12.6	31.9	21.1	66.4	33.7	-0.040	0.254
<i>rrnL</i>	947	36.9	10.0	32.9	20.2	69.8	30.2	-0.056	0.336
<i>atp6</i>	507	44.8	12.2	22.9	20.1	67.7	32.3	-0.324	0.244
<i>cox1</i>	1560	43.3	12.6	23.7	20.4	67.0	33.0	-0.292	0.235
<i>cox2</i>	570	39.1	11.8	24.2	24.9	63.3	36.7	-0.235	0.359
<i>cox3</i>	651	43.6	11.7	22.7	22.0	66.3	33.7	-0.315	0.306
<i>cob</i>	1080	43.8	11.4	24.1	20.7	67.9	32.1	-0.291	0.291
<i>nad1</i>	885	46.3	9.0	21.8	22.8	68.1	31.8	-0.360	0.433
<i>nad2</i>	825	46.4	8.7	23.6	21.2	70.0	29.9	-0.325	0.417
<i>nad3</i>	345	45.2	5.8	25.8	23.2	71.0	29.0	-0.273	0.600
<i>nad4</i>	1218	47.7	10.8	22.0	19.5	69.7	30.3	-0.369	0.285
<i>nad4L</i>	246	47.6	7.7	24.0	20.7	71.6	28.4	-0.330	0.457
<i>nad5</i>	1566	46.7	8.6	24.7	20.0	71.4	28.6	-0.308	0.397
<i>nad6</i>	444	48.9	6.8	25.7	18.7	74.6	25.5	-0.311	0.469
tRNAs	1412	38.2	10.7	30.0	21.0	68.2	31.7	-0.120	0.326
NCR	1926	41.4	6.9	41.2	10.5	82.6	17.4	-0.003	0.212
Full genome	15187	43.1	9.9	27.5	19.5	70.6	29.4	-0.220	0.328
<i>Tetrancistrum nebulosi</i>									
PCGs	9954	39.7	13.2	25.4	21.7	65.1	34.9	-0.220	0.243
1 st codon position	3318	33.3	13.2	28.3	25.2	61.6	38.4	-0.080	0.311
2 nd codon position	3318	46.2	14.7	18.2	20.9	64.4	35.6	-0.434	0.174
3 rd codon position	3318	39.5	11.8	29.6	19.1	69.1	30.9	-0.143	0.238
<i>rrnS</i>	727	32.5	13.8	32.6	21.2	65.1	35.0	0.002	0.213
<i>rrnL</i>	966	34.6	11.9	33.1	20.4	67.7	32.3	-0.021	0.263
<i>atp6</i>	507	42.6	16.2	22.1	19.1	64.7	35.3	-0.317	0.084
<i>cox1</i>	1557	39.4	15.1	24.4	21.1	63.8	36.2	-0.235	0.165
<i>cox2</i>	576	36.3	13.5	27.4	22.7	63.7	36.2	-0.139	0.254
<i>cox3</i>	645	39.7	12.6	26.2	21.6	65.9	34.2	-0.205	0.264
<i>cob</i>	1074	39.4	14.3	25.1	21.1	64.5	35.4	-0.221	0.192
<i>nad1</i>	879	41.1	11.6	25.7	21.6	66.8	33.2	-0.230	0.301
<i>nad2</i>	837	39.9	13.3	25.6	21.3	65.5	34.6	-0.219	0.232
<i>nad3</i>	351	40.5	8.8	26.5	24.2	67.0	33.0	-0.209	0.466
<i>nad4</i>	1215	39.5	13.5	23.8	23.2	63.3	36.7	-0.248	0.265
<i>nad4L</i>	246	43.9	13.8	24.0	18.3	67.9	32.1	-0.293	0.139

<i>nad5</i>	1581	38.4	11.4	27.5	22.7	65.9	34.1	-0.165	0.332
<i>nad6</i>	450	41.8	14.2	24.4	19.6	66.2	33.8	-0.262	0.158
tRNAs	1500	35.4	13.3	30.1	21.2	65.5	34.5	-0.081	0.228
NCR	239	31.8	10.5	40.6	17.2	72.4	27.7	0.121	0.242
Full genome	13392	38.2	13.1	27.2	21.5	65.4	34.6	-0.168	0.243

Table S3 General statistics (length and codons) for mitochondrial protein-coding genes and rRNAs of *Dactylogyrus lamellatus* and other 18 species used for phylogenomic analysis in this study. Abbreviations of species names are available in Table S4.

Gene	Species																		
	A_f	B_h	B_s	C_a	D_l	G_b	G_d	G_g	G_k	G_p	G_s	G_t	M_s	N_m	O_s	P_h	P_m	P_v	T_n
Length of PCGs (bp)																			
<i>atp6</i>	516	510	510	654	510	513	513	513	513	513	513	513	570	513	675	570	567	513	510
<i>cox1</i>	1530	1590	1590	1755	1563	1548	1548	1548	1548	1548	1548	1548	1581	1557	1737	1584	1599	1548	1560
<i>cox2</i>	576	582	582	720	571	582	582	582	582	582	582	582	642	576	780	639	636	582	579
<i>cox3</i>	645	648	648	798	654	639	639	639	639	639	639	639	756	648	792	756	756	639	648
<i>cob</i>	1101	1089	1089	1113	1083	1074	1074	1074	1074	1074	1074	1074	1161	1083	1182	1173	1161	1080	1077
<i>nad1</i>	894	894	885	825	888	888	888	888	888	888	888	888	916	891	894	915	918	891	882
<i>nad2</i>	825	861	861	972	828	857	858	858	858	858	856	857	894	867	999	885	888	876	840
<i>nad3</i>	363	354	351	357	348	351	349	348	348	351	351	351	282	354	342	267	279	351	354
<i>nad4</i>	1203	1197	1215	1356	1221	1209	1209	1209	1209	1206	1209	1209	1200	1218	1422	1233	1134	1209	1218
<i>nad4l</i>	255	255	228	234	249	249	249	249	249	249	249	249	264	249	234	264	267	255	249
<i>nad5</i>	1560	1530	1536	1650	1569	1551	1551	1551	1554	1548	1551	1551	1553	1551	1635	1581	1572	1569	1584
<i>nad6</i>	471	450	450	486	447	483	483	483	483	483	483	483	462	444	477	468	453	477	453
Length of rRNA genes (bp)																			
<i>rrnS</i>	729	729	752	633	725	710	706	710	707	707	711	710	721	729	662	724	736	712	727
<i>rrnL</i>	928	956	982	858	947	957	961	954	955	948	957	958	975	945	920	982	977	960	966
Putative start codon																			
<i>atp6</i>	ATG	ATG	ATG	TTG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	TTG	ATG	ATG	ATG
<i>cox1</i>	ATG	ATG	ATG	ATG	GTG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	GTG	ATG	GTG	GTG	ATG	ATG
<i>cox2</i>	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	TTG	ATG	ATG	TTG
<i>cox3</i>	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG
<i>cob</i>	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG
<i>nad1</i>	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	GTG
<i>nad2</i>	ATT	ATT	GTG	TTA	GTG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	TAT	ATG	ATG	ATG	ATG
<i>nad3</i>	ATG	ATG	ATG	GTG	GTG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	TTG	ATG	ATG	ATG
<i>nad4</i>	ATT	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATA	ATG	ATG	ATG	ATG	ATG	ATG
<i>nad4l</i>	ATG	ATA	ATG	TTG	GTG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	GTG	TTA	ATG	ATG	ATG

<i>nad5</i>	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	TTG	GTG	ATG	ATG	ATG
<i>nad6</i>	ATG	ATG	ATG	ATG	GTG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	TTG	ATG	ATG	ATG	ATG
Putative terminal codon																				
<i>atp6</i>	TAA	TAA	TAG	TAA	TAG	TAG	TAA	TAA	TAG	TAG	TAA	TAA	TAA	TAA	TAA	TAA	TAG	TAG	TAA	TAG
<i>cox1</i>	TAG	TAA	TAA	TAG	TAG	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAG	TAA	TAG
<i>cox2</i>	TAG	TAA	TAG	TAR	T	TAG	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAG	TAA	TAA	TAA	TAA	TAA	TAG
<i>cox3</i>	TAA	TAA	TAA	TAG	TAG	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAG	TAA	TAA	TAG
<i>cob</i>	TAG	TAA	TAG	TAG	TAA	TAA	TAA	TAA	TAA	TAA	TAG	TAA	TAG	TAG	TAG	TAG	TAA	TAG	TAA	TAG
<i>nad1</i>	TAG	TAG	TAG	TAA	TAA	TAG	TAA	TAA	TAG	TAG	TAA	TAA	T	TAA	TAA	TAG	TAA	TAA	TAA	TAG
<i>nad2</i>	TAG	TAG	TAA	TAG	TAG	TA	TAA	TAA	TAG	TAA	T	TA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAG
<i>nad3</i>	TAA	TAA	TAA	TAA	TAG	TAG	T	TAG	TAG	TAA	TAG	TAG	TAA	TAA	TAA	TAG	TAA	TAA	TAA	TAG
<i>nad4</i>	TAG	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAG	TAA	TAG
<i>nad4l</i>	TAA	TAA	TAA	TAG	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAG	TAA	TAG	TAG	TAA	TAA	TAA
<i>nad5</i>	TAA	TAA	TAA	TAG	TAA	TAA	TAA	TAG	TAG	TAG	TAG	TAG	TAG	TA	TAG	TAA	TAG	TAG	TAG	TAG
<i>nad6</i>	TAA	TAG	TAA	TAA	TAG	TAA	TAA	TAG	TAA	TAG	TAA	TAA	TAA	TAA	TAG	TAA	TAG	TAG	TAA	TAG

Table S4 The list of monogenean species used for comparative mitogenomic analyses

Species	Abbreviation	Accession No.	Full length (bp)	A+T content (%)
<i>Aglaiogyrodactylus forficulatus</i>	A_f	KU679421	14371	75.12
<i>Benedenia hoshinai</i>	B_h	NC_014591	13554	74.11
<i>Benedenia seriolae</i>	B_s	NC_014291	13498	75.60
<i>Crenobia alpina</i>	C_a	KP208776	16894	65.38
<i>Dactylogyrus lamellatus</i>	D_l	KR871673	15187	70.63
<i>Gyrodactylus brachymystacis</i>	G_b	NC_031337	14767	65.79
<i>Gyrodactylus derjavinoideis</i>	G_d	NC_010976	14741	68.18
<i>Gyrodactylus gurleyi</i>	G_g	KU659806	14771	72.07
<i>Gyrodactylus kobayashii</i>	G_k	NC_030050	14786	71.57
<i>Gyrodactylus parvae</i>	G_p	NC_031438	14702	73.51
<i>Gyrodactylus salaris</i>	G_s	NC_008815	14790	62.52
<i>Gyrodactylus thymalli</i>	G_t	NC_009682	14788	62.72
<i>Microcotyle sebastis</i>	M_s	NC_009055	14407	70.42
<i>Neobenedenia melleni</i>	N_m	JQ038228	13270	75.92
<i>Obama sp. MAP-2014</i>	O_s	NC_026978	14909	81.02
<i>Paragyrodactylus variegatus</i>	P_v	NC_024754	14517	76.28
<i>Polylabris halichoeres</i>	P_h	NC_016057	15527	72.65
<i>Pseudochauhanea macrorchis</i>	P_m	NC_016950	15031	71.48
<i>Tetrancistrum nebulosi</i>	T_n	NC_018031	13392	65.40

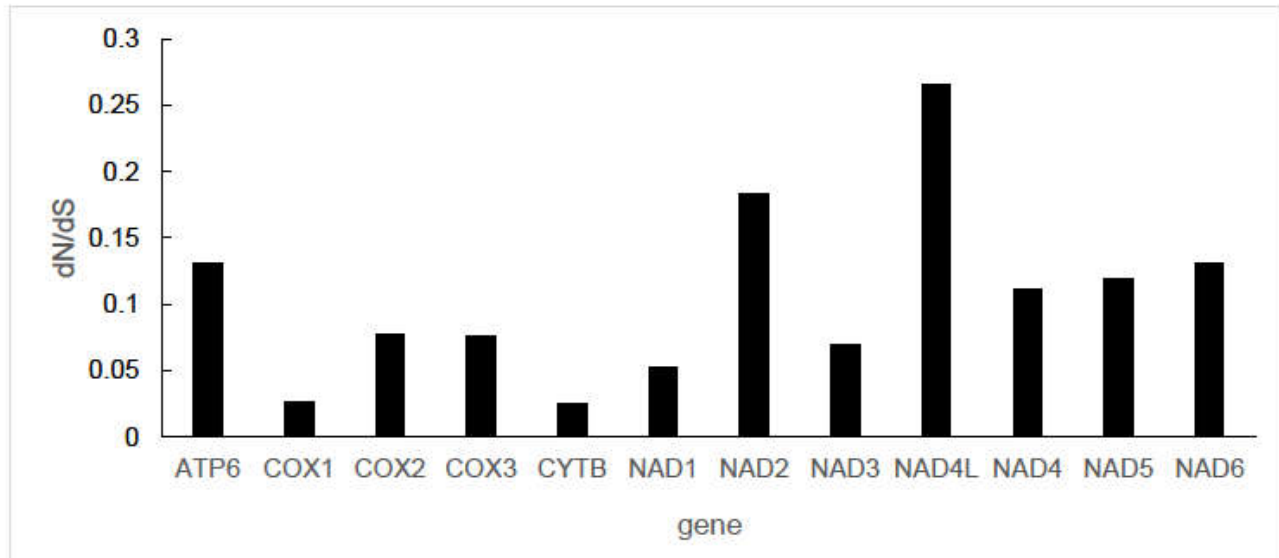
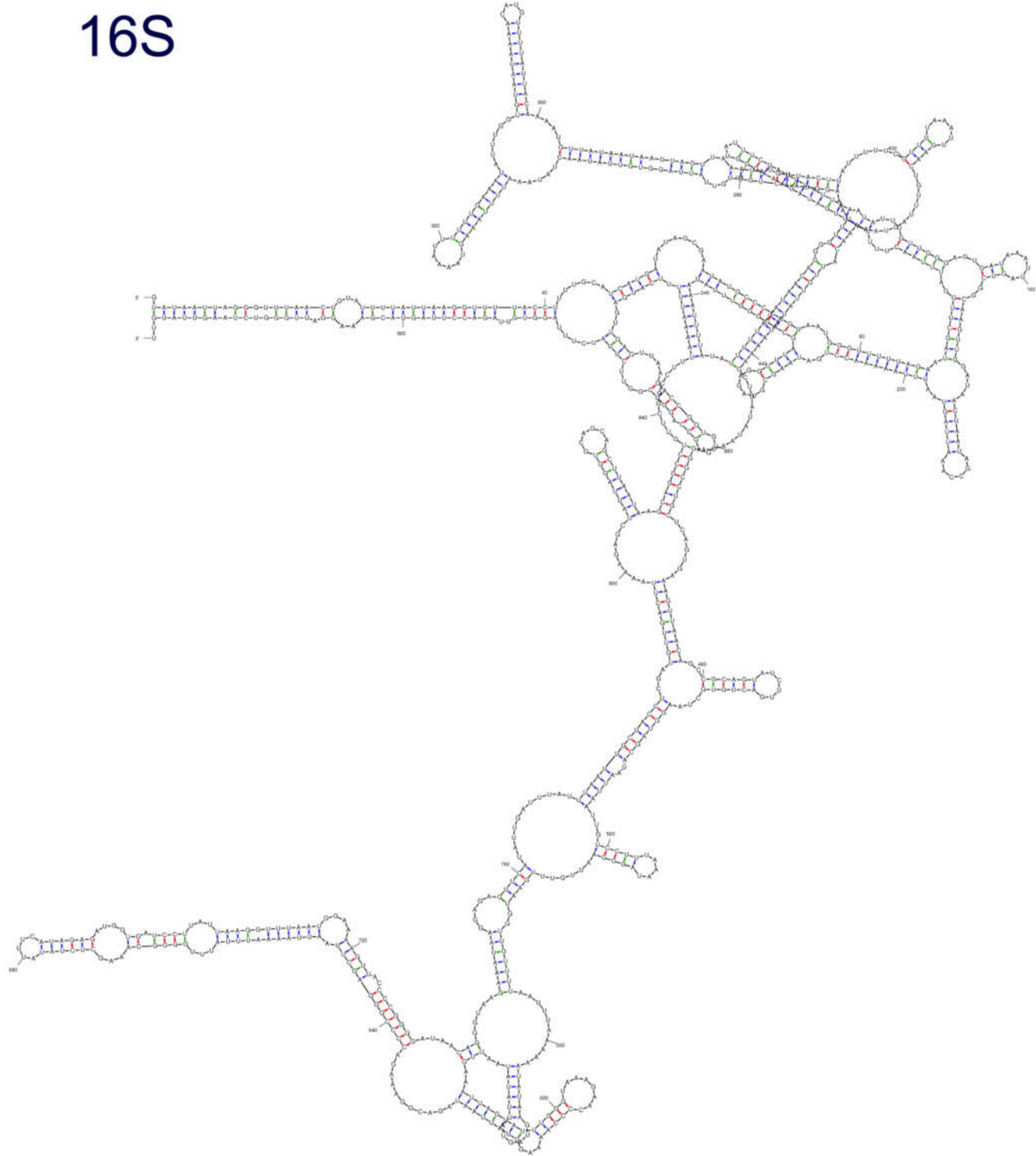


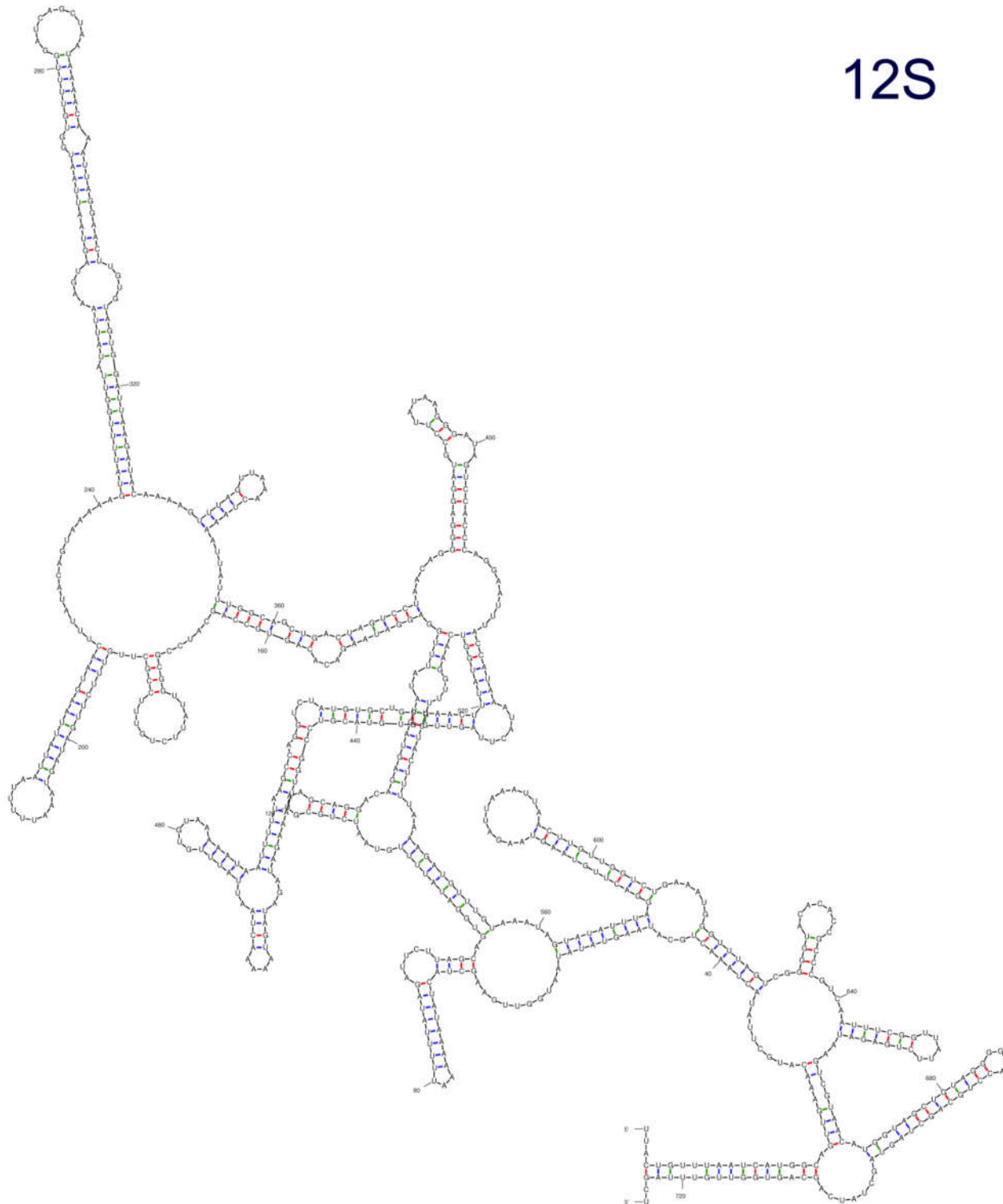
Fig. S1 Ratios of non-synonymous (dN) / synonymous (dS) nucleotide substitutions calculated from all 12 PCGs of *Dactylogyrus lamellatus* and *Tetrancistrum nebulosi* mitogenomes.

16S



$dG = -196.48$ [Initially -240.40] *Dactylogyrus_lamellatus_KR871673*

12S



$dG = -163.72$ [Initially -185.60] *Dactylogyrus_lamellatus_KR871673*

Fig. S2 Secondary structures of rrnL and rrnS predicted by Mfold.

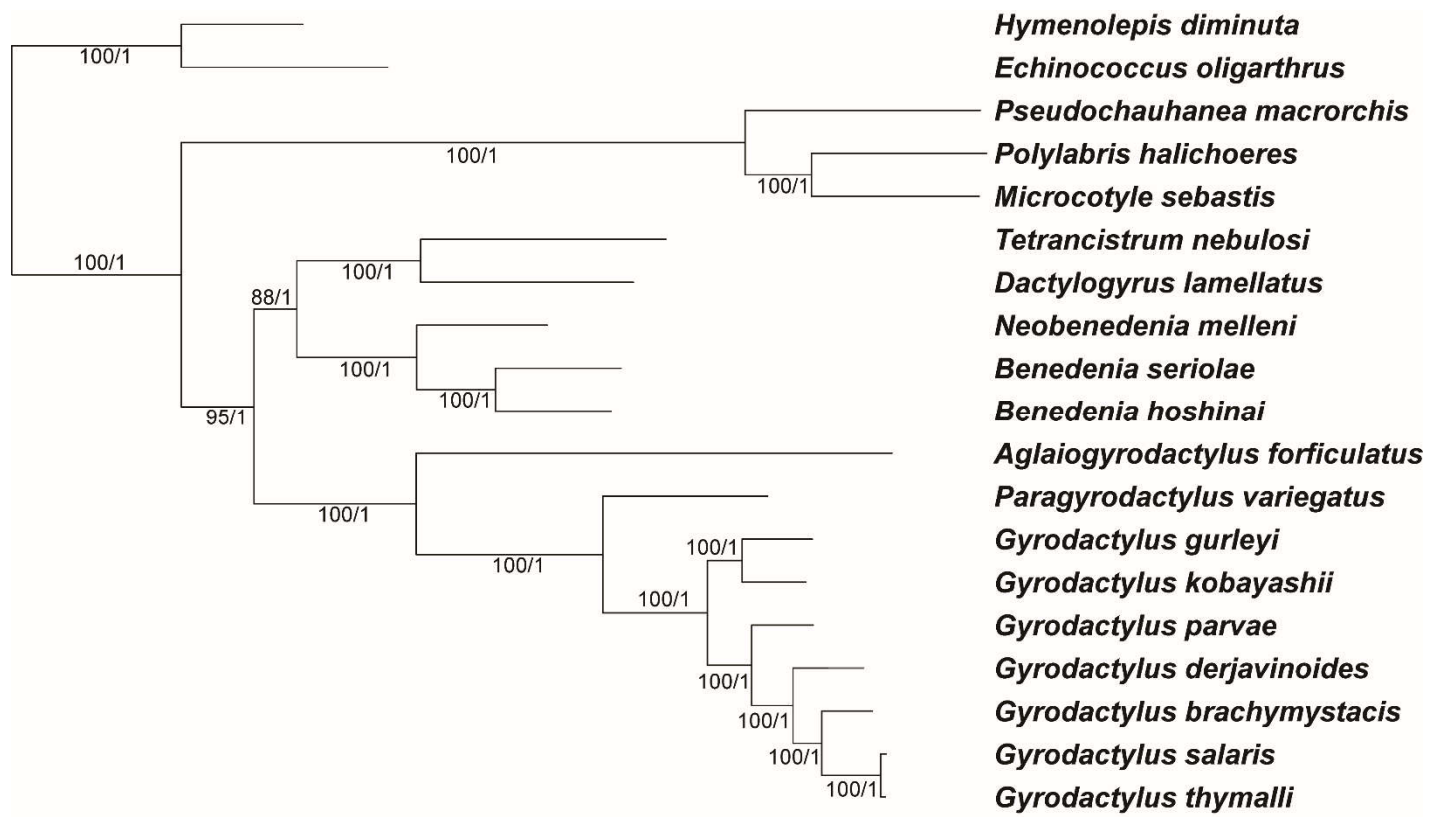
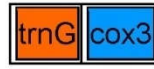


Fig. S3 Maximum Likelihood and Bayesian inference trees inferred using two Cestoda sequences as outgroups. Bootstrap/posterior probability support values of ML/BI analysis are shown above the nodes.

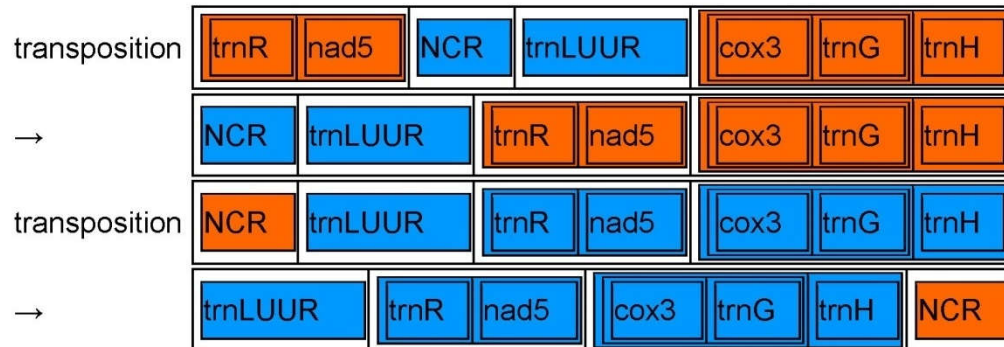
Dactylogyrus lamellatus → *Neobenedenia melleni* :

scenario:

transposition



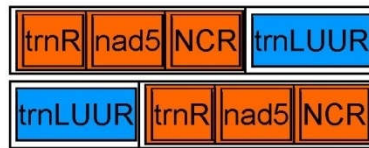
transposition



Dactylogyrus lamellatus → *Tetrancistrum nebulosi* :

scenario:

transposition :



Dactylogyrus lamellatus → *Paragyrodactylus variegatus* :

scenario:

transposition :

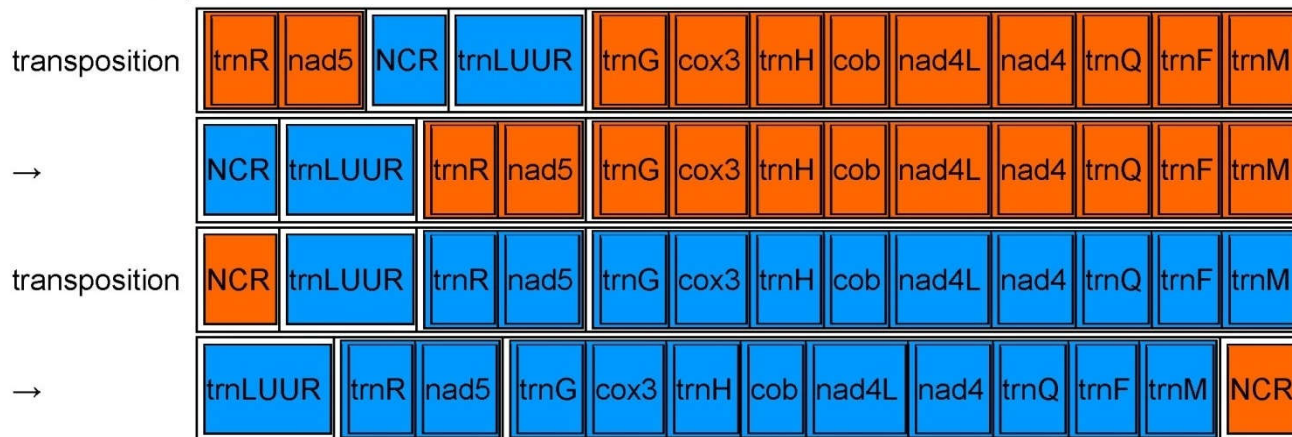


Fig. S4 Transformational pathways between mitogenomic gene orders in *Dactylogyrus lamellatus* and three selected monogenean species: *Neobenedenia melleni*, *Tetrancistrum nebulosi* and *Paragyrodactylus variegatus*.