Parasitology Research (2019) 118:2213–2221
https://doi.org/10.1007/s00436-019-06356-0

GENETICS, EVOLUTION, AND PHYLOGENY - ORIGINAL PAPER

Check for updates

Characterization of the complete mitochondrial genome of *Sphaerirostris picae* (Rudolphi, 1819) (Acanthocephala: Centrorhynchidae), representative of the genus *Sphaerirostris*

Nehaz Muhammad[1] · Suleman[1] · Jun Ma[1] · Mian Sayed Khan[2] · Liang Li[3] · Qing Zhao[3] · Munawar Saleem Ahmad[2] · Xing-Quan Zhu[1]

Received: 5 March 2019 / Accepted: 15 May 2019 / Published online: 10 June 2019

Abstract

The Centrorhynchidae (Acanthocephala: Palaeacanthocephala) is a cosmopolitan family commonly found in various avian and mammalian hosts. Within Centrorhynchidae, species of the genus *Sphaerirostris* Golvan, 1956 are usually parasitic in the digestive tract of various passerine birds. In the present study, adult specimens of *Sphaerirostris picae* (Rudolphi, 1819), the type species of this genus, were recovered from the small intestine of *Acridotheres tristis* (Sturnidae) and *Dendrocitta vagabunda* (Corvidae) in Pakistan. Molecular data from the nuclear or mitochondrial genome is either very limited or completely absent from this phylogenetically understudied group of acanthocephalans. To fill this knowledge gap, we sequenced and determined the internal transcribed spacers of ribosomal DNA (ITS rDNA) and the complete mitochondrial (mt) genome of *S. picae*. The ITS rDNA of *S. picae* was 95.2% similar to that of *Sphaerirostris lanceoides* which is the only member of the Centrorhynchidae whose ITS rDNA is available in GenBank. The phylogenetic tree based on the amino acid sequences of 12 mt protein-coding genes (PCGs) placed *S. picae* close to *Centrorhynchus aluconis* in a monophyletic clade of Polymorphida which also contain members of the families Polymorphidae and Plagiorhynchidae on separate branches. The mt gene arrangement, nucleotide composition and codon usage of 12 PCGs were discussed and compared with those of other acanthocephalan mt genomes. Within the Centrorhynchidae, *S. picae* and *C. aluconis* showed 67.7–86.8% similarity in the nucleotide sequences of 12 PCGs and 2 rRNAs, where *nad*4L is the most conserved gene while *atp*6 is the least conserved. The similarity in amino acid sequences ranged from 68.1 to 91.8%, where *cox*1 was recorded as the most conserved gene, while *atp*6 is highly variable among 12 PCGs. This novel mt genome of *S. picae* provides genetic resources for further studies of phylogenetics and molecular epidemiology of acanthocephalans.

Keywords *Sphaerirostris picae* · Mitochondrial genome · Nuclear ribosomal DNA · Phylogenetics · Acanthocephala

Introduction

The acanthocephalans, commonly known as spiny-headed or thorny-headed worms (Gazi et al. 2015), cause acanthocephaliasis in almost all classes of vertebrates (Amin 2013; Amin et al. 2013; Saini et al. 2018). They inhabit the digestive tract of hosts, penetrate the mucosa with their proboscis and cause inflammation and nodules (Tarello 2009; Halajian et al. 2011). The infected animals suffer from lethargy and emaciation (Tarello 2009). Species of the family Centrorhynchidae (class Palaeacanthocephala) mainly parasitize insectivorous and predatory birds and mammalian hosts worldwide (Dimitrova et al. 1997; Smales 2011; Hussein and El-Dakhly 2018).

Golvan (1956) erected the genus *Sphaerirostris* Golvan, 1956 as a subgenus of the genus *Centrorhynchus* Van Cleave, 1916. Later, Amin and Canaris (1997) raised the

Handling Editor: Una Ryan

✉ Jun Ma
dreamerjm@163.com

✉ Xing-Quan Zhu
xingquanzhu1@hotmail.com

[1] State Key Laboratory of Veterinary Etiological Biology, Key Laboratory of Veterinary Parasitology of Gansu Province, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou 730046, Gansu, People's Republic of China

[2] Department of Zoology, University of Swabi, Swabi, Khyber Pakhtunkhwa, Pakistan

[3] Key Laboratory of Animal Physiology, Biochemistry and Molecular Biology of Hebei Province, College of Life Sciences, Hebei Normal University, Shijiazhuang 050024, Hebei, People's Republic of China

subgenus *Sphaerirostris* Golvan, 1956 to the genus rank with the erection of another genus *Neolacunisoma* Amin & Canaris, 1997; thus, the family Centrorhynchidae currently includes a total of three genera (Amin 2013; Smales 2011). The genus *Sphaerirostris* is differentiated from other two genera by having short spindle-shaped trunk, short and globular anterior proboscis, 3–4 tubular cement glands and polydendritic lacunar vessels (Amin et al. 2015). Twenty-seven nominal species have been recorded in the genus *Sphaerirostris* (Amin 2013; Kang and Li 2018). *Sphaerirostris picae* (the type species of this genus) is characterized by its peculiar morphological features, most importantly the presence of receptacle process in its anterior proboscis (Amin et al. 2010; Kang and Li 2018).

The nuclear ribosomal DNA (rDNA) and mitochondrial (mt) DNA (mtDNA) sequences provide genetic markers which have been proven useful for species identification and phylogenetic analysis of acanthocephalans (Gazi et al. 2012, 2015, 2016; Pan and Nie 2013). Although the Centrorhynchidae is a large and important acanthocephalan family, the mt genome of only one species (*Centrorhynchus aluconis*) of this family was published recently (Gazi et al. 2016). Therefore, the research aims of the present study were to determine the complete mt genome sequences of *S. picae* and to assess its phylogenetic relationship with *C. aluconis* and other selected acanthocephalans using the amino acid sequences of 12 mt protein-coding genes (PCGs). In addition, the sequences of the internal transcribed spacers (ITS) of rDNA were also determined and compared with corresponding sequences of other centrorhynchid acanthocephalans.

Materials and methods

Parasites and DNA extraction

A total of ten adult specimens of acanthocephalans were collected from the small intestine of the common myna (*Acridotheres tristis*) and the rufous treepie (*Dendrocitta vagabunda*) from the district Swabi (34° 07′ 07.23″ N 72° 36′ 32.38″ E), KP, Pakistan. Specimens were kept for whole night in tap water until the proboscis were extruded and then stored in 70% ethanol at − 20 °C. The acanthocephalans were identified morphologically as *S. picae* based on existing Acanthocephala identification keys (Amin 1987; Amin 2013) and specific morphological descriptions (Amin et al. 2010). The total genomic DNA of each individual was extracted by treatment using sodium dodecyl sulfate (SDS)/ proteinase K and column purification (Wizard® SV Genomic DNA Purification System, Promega, Madison, USA) following the published protocol (Gasser et al. 2006) and instructions provided by the manufacturer.

The acquisition of *S. picae* ITS rDNA

The ITS rDNA region was amplified by PCR using universal primers BD1 (5′-GTCGTAACAAGGTTTCCGTA-3′) and BD2 (5′-TATGCTTAAATTCAGCGGGT-3′) (Král'ová-Hromadová et al. 2003). PCR reactions were conducted in a total volume of 25 μl using 12.5 μl of Ex *Taq*™ Version 2.0 DNA polymerase (Takara, Dalian, China), 25 pmol of each primer, 1 μl DNA template and H_2O. The cycling conditions were started with an initial denaturation 95 °C for 3 min followed by 35 cycles at 95 °C for 30 s (denaturation), 55 °C for 30 s (annealing), 72 °C for 90 s (extension) and a final extension step at 72 °C for 10 min. PCR amplicons were checked on 1.5% agarose gels, stained by ethidium bromide and photographed by GelDoc-It TS™ Imaging System (UVP, USA), and the positive amplicon was purified by using gel extraction kit (OMEGA, USA) according to the provided protocol.

The purified ITS PCR product was cloned into Plasmid pMD18-T vector (Takara, Dalian, China) and transformed into *Escherichia coli* strain DH5α. The recombinant bacteria were sequenced by Genewiz (Beijing, China). The ITS rDNA sequence was subjected to BLASTn in NCBI for homological searching. Clustal X 1.83 (Thompson et al. 1997) was used for sequence alignment. Sequence identity (%) was determined by using BioEdit 7.0.9.0 (Hall 1999).

PCR amplification and sequencing of *S. picae* mt genome

Eleven pairs of primers were used to amplify the complete mt genome of *S. picae* (Table 1). The primers were designed based on the conserved mt gene sequences of *C. aluconis* (KT592357) (Gazi et al. 2016). PCR amplification was initiated with 2 min denaturation at 94 °C, then 15 cycles of denaturation at 94 °C for 20 s, annealing at 55–65 °C for 30 s and extension at 65 °C for 3–5 min, followed by 94 °C denaturation for 2 min, plus 30 cycles of 94 °C for 20 s (denaturation), 55–65 °C for 30 s (annealing) and 68 °C for 3–5 min, and a final extension at 68 °C for 10 min.

The amplified fragments were cloned in pMD18-T vector, and the selected positive clones were sequenced by Genewiz (Beijing, China) using primer walking strategy. The chromatograms were checked using Chromas v.1.62. The sequences were assembled using the SeqMan program from DNASTAR (https://www.dnastar.com/). The nucleotide composition and codon usage were analysed using MEGA7 (Kumar et al. 2016). The boundaries of *rrn*L and *rrn*S, and 12 PCGs were identified by comparison with those of *C. aluconis* (KT592357) (Gazi et al. 2016). The sequences were aligned using Clustal X 1.83 (Thompson et al. 1997). The nucleotide sequences of PCGs were translated into amino acids using MEGA7 (Kumar et al. 2016), choosing invertebrate mt code.

Table 1 Primers used for amplification and sequencing of the complete mitochondrial genome of *Sphaerirostris picae*

Primers	Sequence (5′ → 3′)	Gene/region	Size (bp)	Reference
NLF1	GACTCCTTTACTGGTTTGATCG	*cox*1	602	This study
NLR1	GCACATAATGAAAATGAGCC			
NLF2	TTCCTCGAGGGCTTTATGGG	*cox*1-16S	1350	This study
NLR2	GCTTCCGGGGTCTTTCCGTC			
NLF3	CAGTAAGTTGACTGTGCTAAG	16S	335	This study
NLR3	CCGATCTAAACTCAGATCAC			
NLF4	CGGGGATAACAGGGCTATCG	16S-*nad*4	4834	This study
NLR4	CAAAGCCCACACCCACAACC			
NLF5	GGTTTGTGACGAATGCGG	*nad*4-*nad*5	1115	This study
NLR5	CCAATCCGATTCACCAACAAAG			
NLF6	TTGGGATGAGAGGTGTTAGG	*nad*5	302	This study
NLR6	TGAACTAACGCAGATACAGG			
NLF7	GAGTTATGTGGGGAGTGGAAG	*nad*5-*cyt*b	1806	This study
NLR7	CGCCTTAAAGTTCCAACAC			
NLF8	TTGGGTTATGTATTGCCGTG	*cyt*b	407	This study
NLR8	GTTTTCATGAACCTTACCTC			
NLF9	GGTATAAGGATGTTTTAGGGC	*cyt*b-12S	2004	This study
NLR9	ATCCTCCCTACACAGGTACC			
NLF10	CCAGCAGCTGCGGTTATACG	12S	371	This study
NLR10	GATATATTCATCCCATCTCC			
NLF11	GATTGACGTGCGATCTTAAGC	12S-*cox*1	3499	This study
NLR11	TACACTTCAGGATGTCCAA			

The boundaries of transfer RNA genes (tRNAs) and their secondary structures (data not shown) with specific anticodons were identified using the tRNAscanSE program (Lowe and Eddy 1997), DOGMA (Wyman et al. 2004), MITOS (available at http://mitos.bioinf.uni-leipzig.de) (Bernt et al. 2013) and ARWEN (Laslett and Canback 2008) or by comparing with the mt genome sequences of *C. aluconis*.

A comparative analysis of the nucleotide sequences of each PCG and 2 rRNAs, and the amino acid sequences of individual PCG between *S. picae* and *C. aluconis* were conducted to estimate the genetic similarity between these two centrorhynchids.

Phylogenetic analysis

The phylogenetic position of *S. picae* was assessed by phylogenetic analysis of 12 acanthocephalans representing four classes (Archiacanthocephala, Palaeacanthocephala, Eoacanthocephala and Polyacanthocephala) based on the concatenated amino acid sequences of 12 PCGs using Bayesian inference (BI) (Table 2). *Rotaria rotatoria* (Rotifera: Bdelloidea) (Min and Park 2009) was used as the out-group. The alignment of inferred amino acid sequences was done using MAFFT 7.130. The poorly aligned regions were excluded from the alignment using Gblocksv.0.91b (Talavera and Castresana 2007). The resulting file was converted into nexus format which was further subjected to MrBayes (version 3.1.1) (Ronquist and Huelsenbeck 2003) using the mixed model (Castoe and Parkinson 2006). BI analysis was run for 1,000,000 metropolis-coupled MCMC generations and sampling a tree with every 1000 generations with four independent Markov chain Monte Carlo (McMc) chains. The first 250 generations were treated as "burn-in". The analysis was finished when the potential scale reduction factor was close to 1 and the average standard deviation of split frequencies was less than 0.01. The phylograms were visualized using FigTree v. 1.42 (Chen et al. 2014).

Results and discussion

Comparison of ITS rDNA sequences with other species in Centrorhynchidae

The ITS (ITS1, 5.8S and ITS2) rDNA were 711 bp in length (GenBank accession no. MK875833). Currently, the *Sphaerirostris lanceoides* is the only member in the family Centrorhynchidae with ITS sequences available in GenBank (Kang and Li 2018). The similarity in ITS rDNA sequences between *S. picae* and *S. lanceoides* was 95.2%. The

Table 2 The species, taxonomy and GenBank accession numbers for acanthocephalan phylogenetic analysis in this study. Newly sequenced species is shown in bold

Species	Systematic position	GenBank accession no.
Sphaerirostris picae	**Palaeacanthocephala; Polymorphida**	**MK471355**
Centrorhynchus aluconis	Palaeacanthocephala; Polymorphida	KT592357
Southwellina hispida	Palaeacanthocephala; Polymorphida	KJ869251
Plagiorhynchus transversus	Palaeacanthocephala; Polymorphida	KT447549
Leptorhynchoides thecatus	Palaeacanthocephala; Echinorhynchida	NC_006892
Oncicola luehei	Archiacanthocephala; Oligacanthorhynchida	NC_016754
Macracanthorhynchus hirudinaceus	Archiacanthocephala; Oligacanthorhynchida	FR856886
Hebesoma violentum	Eoacanthocephala; Neoechinorhynchida	KC415004
Acanthosentis cheni	Eoacanthocephala; Neoechinorhynchida	KX108947
Paratenuisentis ambiguus	Eoacanthocephala; Neoechinorhynchida	FR856885
Pallisentis celatus	Eoacanthocephala; Gyracanthocephala	JQ943583
Polyacanthorhynchus caballeroi	Polyacanthocephala; Polyacanthorhynchida	KT592358

availability of molecular data from additional centrorhynchid species should shed more light on the interrelationship of *S. picae* with other members of this family.

General characteristics of *S. picae* mt genome

The complete mt genome of *S. picae* (GenBank accession no. MK471355) is 15,170 bp in length, 307 bp shorter than that of *Plagiorhynchus transversus* (15,477 bp) which is the longest mt genome among acanthocephalans sequenced so far, while it is slightly longer than that of *C. aluconis* (15,144 bp) (Gazi et al. 2016). All genes in *S. picae* mt genome are encoded in the same direction and on the same strand (Fig. 1), which is in accordance with those of many acanthocephalans, while in the *Acanthosentis cheni* mt genome, the tRNA-Gln is encoded on L-strand instead of H-strand (Song et al. 2016). The mt genome of *S. picae* encodes 36 genes, containing 12 PCGs (*cox*1–3, *nad*1–6, *nad4L*, *atp6*

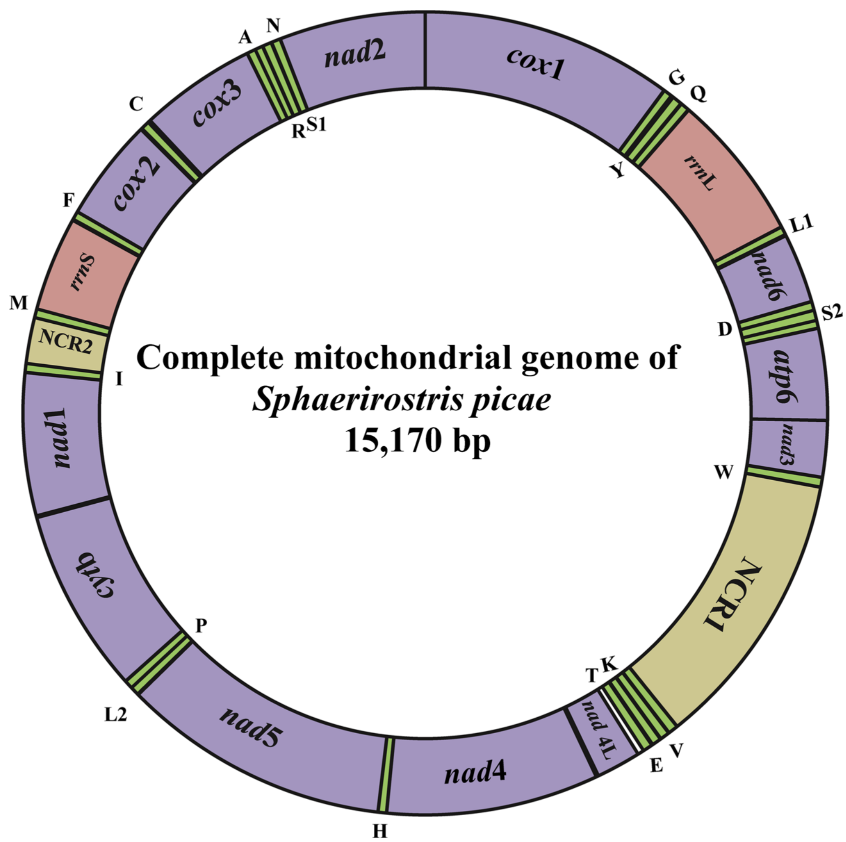

Fig. 1 Organization of the complete mitochondrial genome of *Sphaerirostris picae*. All genes are encoded in the clockwise direction on the same strand. Twenty-two tRNA genes are nominated by the one-letter code for the corresponding amino acid. NCR1 and NCR2 represent long non-coding and short non-coding regions, respectively

Table 3 Mitochondrial genome organization of *Sphaerirostris picae*

Gene/region	Positions	Size (bp)	No. of aa	Start and stop codons	Intergenic sequences
*cox*1	1–1542	1542	513	ATT/TAG	2
*trn*G	1545–1598	53			− 14
*trn*Q	1585–1645	60			− 4
*trn*Y	1642–1695	54			3
*rrn*L	1699–2604	906			1
*trn*L1	2606–2656	51			11
*nad*6	2668–3091	424	141	GTG/T	1
*trn*D	3093–3146	54			63
*trn*S2	3210–3257	48			− 5
*atp*6	3253–3819	567	188	ATG/TAA	0
*nad*3	3820–4165	346	115	ATG/T	1
*trn*W	4167–4225	59			− 1
NCR1	4225–5944	1720			1
*trn*V	5946–6006	61			3
*trn*K	6010–6067	58			− 4
*trn*E	6064–6120	57			6
*trn*T	6127–6186	60			39
*nad*4L	6226–6498	273	90	GTG/TAA	14
*nad*4	6513–7811	1299	432	ATG/TAG	1
*trn*H	7813–7863	51			− 1
*nad*5	7863–9503	1641	546	ATG/TAG	5
*trn*L2	9509–9561	53			0
*trn*P	9562–9614	53			− 1
*cyt*b	9614–10,747	1134	377	GTG/TAG	14
*nad*1	10,762–11,632	871	290	GTG/T	− 1
*trn*I	11,632–11,685	54			0
NCR2	11,686–11,967	282			0
*trn*M	11,968–12,021	54			0
*rrn*S	12,022–12,600	579			− 7
*trn*F	12,594–12,644	51			− 1
*cox*2	12,644–13,273	630	209	GTG/TAG	8
*trn*C	13,282–13,334	53			16
*cox*3	13,351–14,064	714	237	GTG/TAA	− 1
*trn*A	14,064–14,116	53			0
*trn*R	14,117–14,165	48			− 2
*trn*N	14,164–14,218	55			3
*trn*S1	14,222–14,276	55			− 1
*nad*2	14,276–15,170	895	298	GTG/T	0

Stop codons were not included

nt nucleotide, *aa* amino acid

and *cyt*b), 22 transfer RNAs and two ribosomal RNA genes (*rrnL* and *rrn*S) (Table 3). The organization and individual gene length of *S. picae* mt genome is very similar to those of *C. aluconis* (Gazi et al. 2016). The absence of *atp*8 gene is commonly seen in platyhelminths, nematodes (Park et al. 2007, 2011), Rotifera (Min and Park 2009) and

Table 4 Nucleotide composition of the mitochondrial genome of *Sphaerirostris picae*

Nucleotide sequence	Size (bp)	A (%)	C (%)	T (%)	G (%)	A + T (%)	G + C (%)
Overall	15,170	22.2	7.5	35.9	34.4	58.1	41.9
Protein coding	10,336	19.5	7.0	37.8	35.8	57.3	42.7
Codon position*							
1st	3445	18.4	8.2	39.2	34.2	57.6	42.4
2nd	3445	18.2	7.1	39.6	35.1	57.8	42.2
3rd	3445	21.9	5.6	34.6	37.9	56.5	43.5
Ribosomal RNA genes	1485	30.8	10.1	31.6	27.5	62.4	37.6
Transfer RNA genes	1195	25.8	9.3	35.2	29.6	61.0	38.9
Non-coding region 1	1720	27.3	7.6	28.5	36.6	55.8	44.2
Non-coding region 2	282	31.2	6.7	34.0	28.0	65.2	34.7

*Termination codons are excluded

Table 5 Genetic code and codon usage for the 12 mitochondrial protein-coding genes of *Sphaerirostris picae*

Codon	aa	No.	%	Codon	aa	No.	%
TTT	Phe	196	5.7	TAT	Tyr	90	2.6
TTC	Phe	26	0.8	TAC	Tyr	12	0.4
TTA	Leu	152	4.4	TAA	*	53	1.5
TTG	Leu	215	6.2	TAG	*	75	2.2
CTT	Leu	47	1.4	CAT	His	33	1.0
CTC	Leu	15	0.4	CAC	His	4	0.1
CTA	Leu	22	0.6	CAA	Gln	14	0.4
CTG	Leu	33	1.0	CAG	Gln	27	0.8
ATT	Ile	119	3.5	AAT	Asn	38	1.1
ATC	Ile	12	0.4	AAC	Asn	4	0.1
ATA	Met	57	1.7	AAA	Lys	18	0.5
ATG	Met	88	2.6	AAG	Lys	26	0.8
GTT	Val	141	4.1	GAT	Asp	73	2.1
GTC	Val	21	0.6	GAC	Asp	12	0.4
GTA	Val	83	2.4	GAA	Glu	53	1.5
GTG	Val	135	3.9	GAG	Glu	97	2.8
TCT	Ser	29	0.8	TGT	Cys	138	4.0
TCC	Ser	5	0.2	TGC	Cys	26	0.8
TCA	Ser	16	0.5	TGA	Trp	80	2.3
TCG	Ser	13	0.4	TGG	Trp	222	6.4
CCT	Pro	20	0.6	CGT	Arg	15	0.4
CCC	Pro	2	0.1	CGC	Arg	3	0.1
CCA	Pro	10	0.3	CGA	Arg	16	0.5
CCG	Pro	6	0.2	CGG	Arg	16	0.5
ACT	Thr	23	0.7	AGT	Ser	45	1.3
ACC	Thr	6	0.2	AGC	Ser	18	0.5
ACA	Thr	9	0.3	AGA	Ser	50	1.5
ACG	Thr	14	0.4	AGG	Ser	106	3.1
GCT	Ala	38	1.1	GGT	Gly	148	4.3
GCC	Ala	6	0.2	GGC	Gly	21	0.6
GCA	Ala	24	0.7	GGA	Gly	96	2.8
GCG	Ala	23	0.7	GGG	Gly	210	6.1

aa amino acid, *No.* number of copies

*Stop (termination) codon

acanthocephalans (Gazi et al. 2016), except for *Leptorhynchoides thecatus* (Rhadinorhynchidae: Palaeacanthocephala) which was suggested to have two putative *atp*8 genes (Steinauer et al. 2005).

The nucleotide composition of *S. picae* mt genome is 22.2% A (3369 bp), 34.4% G (5223 bp), 35.9% T (5440 bp) and 7.5% C (1138 bp) (Table 4). The overall A + T content (58.1%) of *S. picae* mt genome is higher than those of *C. aluconis* (55.6%) and *Polyacanthorhynchus caballeroi* (56.3%) and lower than those of *P. transversus* (61.1%), *Oncicola luehei* (60.2%), *Southwellina hispida* (63.9%), *Pallisentis celatus* (61.5%), *A. cheni* (65.3%) and *L. thecatus* (71.5%) (Steinauer et al. 2005; Gazi et al. 2012; Weber et al. 2013; Pan and Nie 2013; Gazi et al. 2015, 2016; Song et al. 2016).

The gene arrangement of ribosomal RNAs and PCGs in *S. picae* mt genome is the same as in other acanthocephalans reported so far (Wey-Fabrizius et al. 2013; Weber et al. 2013; Gazi et al. 2016). The length of 12 PCGs in *S. picae* mt genome is 10,336 bp in nucleotide which were inferred into 3436 amino acids and composed of 3445 codons. In the mt genome of *S. picae*, TGG (6.4%), TTG (6.2%) and GGG (6.1%) were the most frequently used codons, whereas CCC (0.1%), CGC (0.1%), CAC (0.1%), CCC and CGC were the least frequently used codons (Table 5). In PCGs of *S. picae* mt genome, leucine (14.0%) is the most abundant amino acid used, followed by glycine (13.8%) and valine (11.0%). The high content of leucine (15.0%), followed by valine (14.4%), was also found in *P. transversus* (Gazi et al. 2016). In the 12 PCGs of *S. picae*, GTG and ATG were used as start codons, except for *cox*1 where ATT acted as start codon. The *atp*6, *nad*3, *nad*4 and *nad*5 were inferred to use ATG as start codon, while *nad*6, *nad*4L, *cyt*b, *nad*1, *nad*2, *cox*2 and *cox*3 started with GTG, and *cox*1 used ATT as start codon. Complete stop codons were used in eight PCGs. Five genes (*cox*1, *nad*4, *nad*5, *cyt*b and *cox*2) were terminated with TAG, and three PCGs (*atp*6, *nad*4L and *cox*3) had TAA as stop codons. The incomplete stop codon T was used in four PCGs (*nad*6, *nad*3, *nad*1 and *nad*2) (Table 3). Incomplete stop codons are commonly used in mtDNAs of acanthocephalans including members of the order Polymorphida; for example, three genes (*nad*6, *nad*3 and *cox*2) of *Southwellina hispida* (Gazi et al. 2015), four genes (*nad*6, *nad*3, *nad*1 and *nad*2) of *C. aluconis* and four genes (*nad*6, *nad*3, *nad*4 and *nad*1) of *P. transversus* have been reported to terminate with an incomplete stop codon (T).

Two ribosomal RNAs, *rrn*L and *rrn*S, of the *S. picae* mt genome were 906 bp and 579 bp in length with A + T contents of 62.0% and 62.9%, respectively. The *rrn*L lies between *trn*Y and *trn*L1, the same as in other acanthocephalans reported so far, except for *Macracanthorhynchus hirudinaceus* where it lies between *trn*Y and *trn*L2 (Weber et al. 2013). The position of *rrn*S in the studied mt genome is also consistent with those of other acanthocephalans sequenced so far, i.e. between *trn*M and *trn*F, except for *L. thecatus* and *P. celatus* where it is located between *trn*S1 and *trn*F (Steinauer et al. 2005; Pan and Nie 2013) (Fig. 2). The lengths of 22 transfer RNAs (tRNAs) ranged from 48 bp (*trn*S2 and *trn*R) to 61 bp (*trn*V). Anticodons and secondary structures (data not shown) of tRNAs were also identified. Twenty tRNA genes can be folded into "cloverleaf" secondary structure. The dihydrouridine (DHU) arm is missing in the remaining two tRNAs (*trn*S1 and *trn*S2) which is consistently observed in the mt genomes of other invertebrates including acanthocephalans (Gazi et al. 2016). There are two non-coding regions

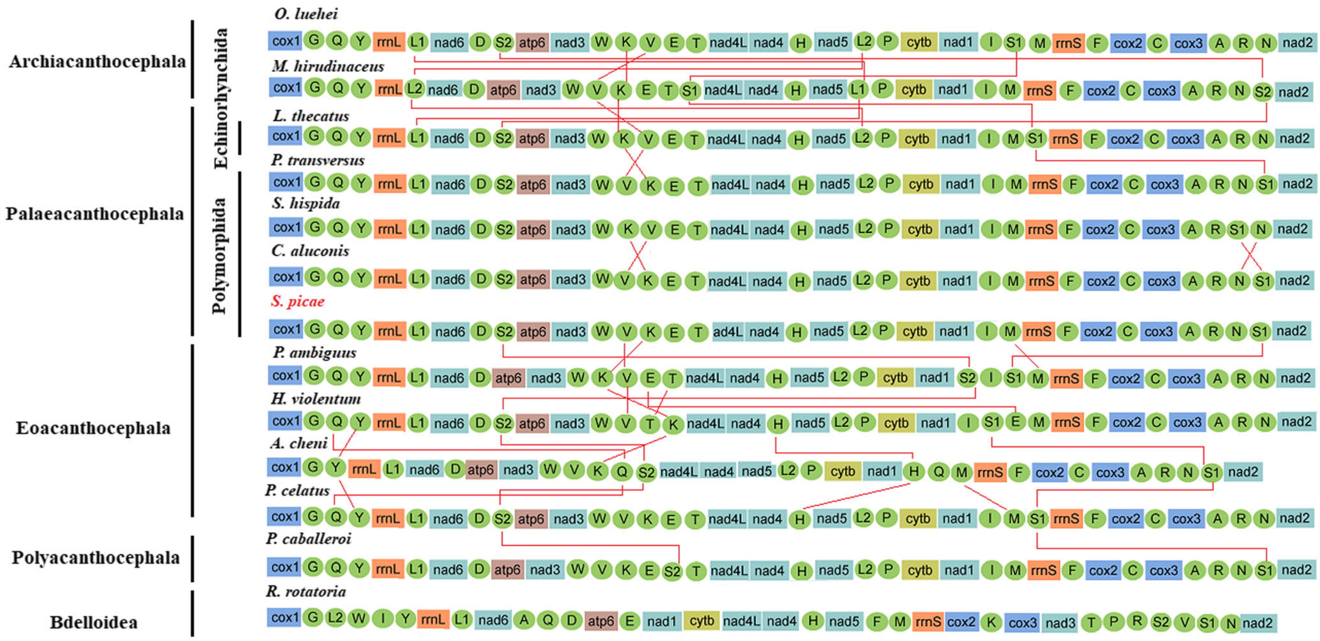

Fig. 2 Linearized comparison of the mitochondrial gene arrangement for 12 acanthocephalans and one rotifer species. Gene and genome size are not to scale. All genes are encoded in the same direction from left to right. The tRNAs are labelled by a single-letter code for the corresponding amino acid. The red colour indicates the newly sequenced species in this study

(NCRs) in the mt genome of *S. picae*. The NCR1 is 1720 bp in length, which is located between *trn*W and *trn*V. The NCR2 is 282 bp in length and located between *trn*I and *trn*M. The A + T contents of NCR1 and NCR2 of *S. picae* mt genome are 55.8% and 65.2%, respectively (Table 4).

The gene order in mt genomes of four polymorphid species including *S. picae* is almost the same except for a few translocations of tRNAs. The *trn*V–*trn*K and *trn*S1–*trn*N of *S. picae* and *C. aluconis* mt genomes are reciprocal in position to that of *S. hispida*. The mt gene arrangement of 12 acanthocephalans including a rotifer (*R. rotatoria*) are presented in Fig. 2.

Table 6 Identity (%) of nucleotide and predicted amino acid sequences between *Sphaerirostris picae* and *Centrorhynchus aluconis*

Gene/region	Nucleotide identity (%)	Amino acid identity (%)
*cox*1	80.3	91.8
*rrn*L	79.0	
*nad*6	69.4	77.3
*atp*6	67.7	68.1
*nad*3	73.6	76.5
*nad*4L	86.8	80.0
*nad*4	77.3	78.7
*nad*5	71.6	70.8
*cyt*b	78.8	81.2
*nad*1	75.5	82.4
*rrn*S	82.5	
*cox*2	79.5	85.2
*cox*3	72.5	68.8
*nad*2	70.5	73.8
PCGs	75.3	78.6
Overall	75.0	

PCGs protein-coding genes

Comparison of *S. picae* mt genome sequences with other centrorhynchid species

Currently, the mt genome sequence of only one species (*C. aluconis*) of the family Centrorhynchidae was available in GenBank. Hence, this was used for the comparative purposes. The similarity between *S. picae* and *C. aluconis* mt genomes across the concatenated PCGs was 75.3% in nucleotides and 78.6% in amino acids, respectively. The highest sequence similarity in nucleotides was observed in *nad*4L (86.8%) and the lowest in *atp*6 (67.7%) (Table 6). Although the similarity in *atp*6 gene between the two species was low in both nucleotides and amino acids, the average sequence similarity was higher in amino acids (68.1–91.8%) than in nucleotides (67.7–86.8%) in almost all PCGs. This high similarity in amino acid sequences than in nucleotides is probably due to the similarity of codons that encode a specific amino acid.

Phylogenetic analysis

In the present phylogenetic analysis, *Macracanthorhynchus hirudinaceus* and *Oncicola luehei* joined together forming a separate basal clade of Archiacanthocephala. As expected, *S. picae* grouped with *C. aluconis* (both are members of the

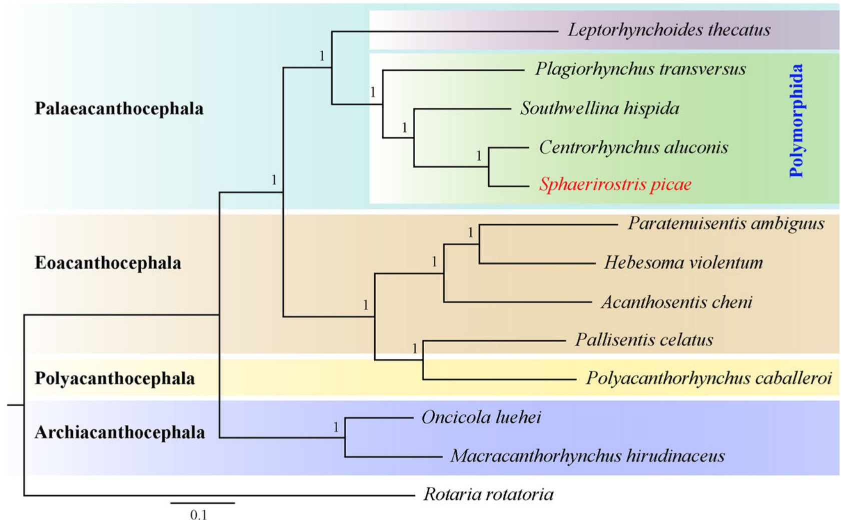

Fig. 3 Phylogenetic relationships of *Sphaerirostris picae* with other acanthocephalans constructed using Bayesian inference (BI) analysis based on the concatenated amino acid sequences of 12 protein-coding genes. *Rotaria rotatoria* (GQ304898.1) (Bdelloidea: Rotifera) served as out-group

Centrorhynchidae) with 100% nodal support in a clearly monophyletic clade Polymorphida (Gazi et al. 2015, 2016) which also contained other members of the family Polymorphidae and Plagiorhynchidae (each represented by a single mt genome) (Fig. 3). The tree placed Echinorhynchida (represented by *Leptorhynchoides thecatus*) and Polymophida together in a clade forming Palaeacanthocephala, the largest class of Acanthocephala (Amin 2013). The general tree topology is consistent with that of previous molecular studies including those based on mt genome sequences (Gazi et al. 2012, 2015, 2016; Weber et al. 2013).

Our results and that of Gazi et al. (2016) showed that polyacanthocephalan represented by *Polyacanthorhynchus caballeroi* (Polyacanthorhynchidae) grouped with Eoacanthocephala (*Pallisentis celatus*), although this was not consistent with earlier morphology-based classifications (Schmidt and Canaris 1967; Golvan 1956; Bullock 1969; Amin 1985) where the Polyacanthorhynchidae was placed within the class Palaeacanthocephala. A more comprehensive taxon sampling from both Polyacanthocephala and Eoacanthocephala is warranted to address the question of acanthocephalan monophyly.

Conclusions

This study determined the ITS rDNA sequences and the complete mt genome sequences of *S. picae*, the representative species of the genus *Sphaerirostris*, for the first time. Phylogenetic analysis based on the amino acid sequences of 12 mt PCGs supported the close genetic relationship between *S. picae* and *C. aluconis* within the family Centrorhynchidae. The obtained mt genome sequences of *S. picae* will be useful for further research of population genetics of *Sphaerirostris* spp. and for the systematic studies of the phylum Acanthocephala.

Funding information Project support was provided by National Natural Science Foundation of China (Grant No. 31702225), the International Science and Technology Cooperation Project of Gansu Provincial Key Research and Development Program (Grant No. 17JR7WA031), the Elite Program of Chinese Academy of Agricultural Sciences and the Agricultural Science and Technology Innovation Program (ASTIP) (Grant No. CAAS-ASTIP-2016-LVRI-03).

Compliance with ethical standards The collection of parasite specimens from birds was approved by the Animal Ethics Committee of the University of Swabi, Pakistan. All birds were handled in strict accordance with good animal practice according to the Animal Ethics Procedures and Guidelines of Pakistan.

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

References

Amin OM (1985) classification. In: Crompton DWT, Nickol BB (eds) Biology of the Acanthocephala. Cambridge University Press, London and New York, pp 27–72

Amin OM (1987) Key to the families and subfamilies of Acanthocephala, with the erection of a new class (Polyacanthocephala) and a new order (Polyacanthorhynchida). J Parasitol 73:1216–1219

Amin OM (2013) Classification of the Acanthocephala. Folia Parasitol 60:273–305

Amin OM, Canaris A (1997) Description of *Neolacunisoma geraldschmidti* gen. n., sp. n., (Acanthocephala: Centrorhynchidae) from South African shorebirds. J Helminthol Soc Wash 64:275–280

Amin OM, Heckmann RA, Halajian A, Eslami A (2010) Redescription of *Sphaerirostris picae* (Acanthocephala: Centrorhynchidae) from magpie, *Pica pica*, in northern Iran, with special reference to unusual receptacle structures and notes on histopathology. J Parasitol 96: 561–568

Amin OM, Evans P, Heckmann RA, El-Naggar AM (2013) The description of *Mediorhynchus africanus* n. sp. (Acanthocephala: Gigantorhynchidae) from galliform birds in Africa. Parasitol Res 112:2897–2906

Amin OM, Heckmann RA, Wilson E, Keele B, Khan A (2015) The description of *Centrorhynchus globirostris* n. sp. (Acanthocephala: Centrorhynchidae) from the pheasant crow, *Centropus sinensis*

(Stephens) in Pakistan, with gene sequence analysis and emendation of the family diagnosis. Parasitol Res 114:2291–2299

Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, Putz M, Middendorf M, Stadler PF (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol 69:313–319

Bullock WL (1969) Morphological features as tools and pitfalls in acanthocephalan systematics. In GD Schmidt (Ed.) Baltimore: University Park Press. Problems in systematics of parasites pp 9–43

Castoe TA, Parkinson CL (2006) Bayesian mixed models and the phylogeny of pitvipers (Viperidae: Serpentes). Mol Phylogenet Evol 39: 91–110

Chen YY, Chen S, Kang J, Fang H, Dao H, Guo W, Lai C, Lai M, Fan J, Fu L, Andrieu JM, Lu W (2014) Evolving molecular epidemiological profile of human immunodeficiency virus 1 in the southwest border of China. PLoS One 9:e107578

Dimitrova ZM, Georgiev BB, Genov T (1997) Acanthocephalans of the family Centrorhynchidae (Palaeacanthocephala) from Bulgaria. Folia Parasitol 44:22–232

Gasser RB, Hu M, Chilton NB, Campbell BE, Jex AJ, Otranto D, Cafarchia C, Beveridge I, Zhu XQ (2006) Single strand conformation polymorphism (SSCP) for the analysis of genetic variation. Nat Protoc 1:3121–3128

Gazi M, Sultana T, Min GS, Park YC, García-Varela M, Nadler SA, Park JK (2012) The complete mitochondrial genome sequence of *Oncicola luehei* (Acanthocephala: Archiacanthocephala) and its phylogenetic position within Syndermata. Parasitol Int 61:307–316

Gazi M, Kim J, Park JK (2015) The complete mitochondrial genome sequence of *Southwellina hispida* supports monophyly of Palaeacanthocephala (Acanthocephala: Polymorphida). Parasitol Int 64:64–68

Gazi M, Kim J, García-Varela M, Park C, Littlewood DTJ, Park JK (2016) Mitogenomic phylogeny of Acanthocephala reveals novel class relationships. Zool Scr 45:437–454

Golvan YJ (1956) Le genre Centrorhynchus Lühe 1911 (Acanthocephala-Polymorphidae). Révision des espèces européennes et description dune nouvelle espèce africaine parasite de rapace diurne. Bulletin de l'Institut Fondamental d'Afrique Noire 732–785

Halajian A, Eslami A, Mobedi I, Amin O, Mariaux J, Mansoori J, Tavakol S (2011) Gastrointestinal helminths of magpies (*Pica pica*), rooks (*Corvus frugilegus*) and carrion crows (*Corvus corone*) in Mazandaran Province, north of Iran. Iran J Parasitol 6:38–44

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser (Oxf) 41:95–98

Hussein NM, El-Dakhly KM (2018) Helminth community, with special reference to the acanthocephalans, *Sphaerirostris sp.* (Polymorphida) and *Mediorhynchus sp.* (Gigantorhynchidae), in the shorebird, *Calidris temminckii* (Aves: Charadriiformes). Qena, Upper Egypt. Asian J Anim Vet Adv 13:316–323

Kang J, Li L (2018) First report on cystacanths of *Sphaerirostris lanceoides* (Petrochenko, 1949) (Acanthocephala: Centrorhynchidae) from the Asiatic toad *Bufo gargarizans* Cantor (Amphibia: Anura) in China. Syst Parasitol 95:447–454

Král'ová-Hromadová I, Tietz DF, Shinn AP, Špakulová M (2003) ITS rDNA sequences of *Pomphorhynchus laevis* (Zoega in Müller, 1776) and *P. lucyi* Williams & Rogers, 1984 (Acanthocephala: Palaeacanthocephala). Syst Parasitol 56:141–145

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33: 1870–1874

Laslett D, Canback B (2008) ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24:172–175

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genome sequences. Nucleic Acids Res 25:955–964

Min GS, Park JK (2009) Eurotatorian paraphyly: revisiting phylogenetic relationships based on the complete mitochondrial genome sequence of *Rotaria rotatoria* (Bdelloidea: Rotifera: Syndermata). BMC Genomics 10:533

Pan TS, Nie P (2013) The complete mitochondrial genome of *Pallisentis celatus* (Acanthocephala) with phylogenetic analysis of acanthocephalans and rotifers. Folia Parasitol 60:181–191

Park JK, Kim KH, Kang S, Kim W, Eom KS, Littlewood DTJ (2007) A common origin of complex life cycles in parasitic flatworms: evidence from the complete mitochondrial genome of *Microcotyle sebastis* (Monogenea: Platyhelminthes). BMC Evol Biol 7:11

Park JK, Sultana T, Lee SH, Kang S, Kim HK, Min GS, Eom KS, Nadler SA (2011) Monophyly of clade III nematodes is not supported by phylogenetic analysis of complete mitochondrial genome sequences. BMC Genomics 12:392

Ronquist F, Huelsenbeck JP (2003) MrBayes 3. Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574

Saini J, Kumar H, Das P, Ghosh J, Gupta D, Chandra K (2018) Acanthocephala. In Faunal Diversity of Indian Himalaya: 135–138 (Published by the Director, Zool. Surv. India, Kolkata)

Schmidt GD, Canaris AG (1967) Acanthocephala from Kenya with descriptions of two new species. J Parasitol 53:634–637

Smales LR (2011) Centrorhynchidae (Acanthocephala) including the description of new species of *Centrorhynchus* from birds from the Côte d'Ivoire, Africa. Rev Suisse Zool 118:307–318

Song R, Zhang D, Deng S, Ding D, Liao F, Liu L (2016) The complete mitochondrial genome of *Acanthosentis cheni* (Acanthocephala: Quadrigyridae). Mitochondrial DNA Part B Resour 1:797–798

Steinauer ML, Nickol BB, Broughton R, Ortí G (2005) First sequenced mitochondrial genome from the phylum Acanthocephala (*Leptorhynchoides thecatus*) and its phylogenetic position within Metazoa. J Mol Evol 60:706–715

Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol 56:564–577

Tarello W (2009) Prevalence of acanthocephalans in birds of prey from the Middle East and report of two clinical cases. Parasite 16:73–74

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

Weber M, Wey-Fabrizius AR, Podsiadlowski L, Witek A, Schill RO, Sugár L, Hankeln T (2013) Phylogenetic analyses of endoparasitic Acanthocephala based on mitochondrial genomes suggest secondary loss of sensory organs. Mol Phylogenet Evol 66:182–189

Wey-Fabrizius AR, Podsiadlowski L, Herlyn H, Hankeln T (2013) Platyzoan mitochondrial genomes. Mol Phylogenet Evol 69:365–375

Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20:3252–3255