JSE Journal of Systematics and Evolution

Journal of Systematics and Evolution 51 (3): 353–364 (2013) doi: 10.1111/jse.12021

Research Article

Phylogenetic performance of mitochondrial protein-coding genes of *Oncomelania hupensis* in resolving relationships between landscape populations

[1]Shi-Zhu LI [1]Li ZHANG [2]Lin MA [1,3]Wei HU [1]Shan LV [1]Qin LIU
[1]Ying-Jun QIAN [1]Qiang WANG [1]Xiao-Nong ZHOU*

[1](*National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention; WHO Collaborating Centre for Malaria, Schistosomiasis and Filariasis*, Shanghai 200025, China)
[2](*Shaanxi Province Center for Disease Control and Prevention*, Xi'an 610032, China)
[3](*Fudan University*, Shanghai 200433, China)

Abstract *Oncomelania hupensis* is the unique intermediate host of *Schistosoma japonicum*, which plays a key role in the transmission of human blood fluke *Schistosoma*. The complete mitochondrial (mt) genome of *O. hupensis* has been characterized; however, the phylogenetic performance of mt protein-coding genes (PCGs) of the snail remain unclear. In this study, 11 whole mt genomes of snails collected from four different ecological settings in China and the Philippines were sequenced. The mt genome sizes ranged from 15 183 to 15 216 bp, with the G + C contents from 32.4% to 33.4%. A total of 15 251 characters were generated from the multiple sequence alignment. Of 2711 (17.8%) polymorphic sites, 56.22% (1524) were parsimony sites. The mt genomes' phylogenetic trees were reconstructed using minimum evolution, neighbor joining, maximum likelihood, maximum parsimony, and Bayesian tree estimate methods, and two main distinct clades were identified: (i) the isolate from mountainous regions; (ii) the remaining isolate which included three inner branches. All phylogenetic trees of the 13 PCGs were generated by running 1000 bootstrap replicates and compared with the complete mtDNA tree, the classification accuracy ranging from 21.23% to 87.87%, the topological distance of phylogenetic trees between PCGs ranging from 5 to 14. Therefore, the performance of PCGs can be divided into good condition (*COI*, *ND2*, *ND5*, and *ND3*), medium (*COII*, *ATP6*, *ND1*, *ND6*, *Cytb*, *ND4*, and *COIII*), poor (*ATP8* and *ND4L*). This study represents the first analysis of mt genome diversity of the *O. hupensis* snail and phylogenetic performance of mt PCGs. It presents clear evidence that the snail populations can be separated into four landscape genetic populations in mainland China based on whole mt genomes. The identification of the phylogenetic performance of PCGs provides new insight into the intensive genetic diversity study using mtDNA markers for the snail.

Key words mitochondrial genome, *Oncomelania hupensis*, phylogenetic performance, phylogeography, protein-coding genes.

Schistosomiasis japonica remains one of the most prevalent parasitic infections in China and has significant economic and health consequences (Li et al., 2009a; Wang et al., 2009; Collins et al., 2012). *Oncomelania hupensis* is the unique intermediate host of *Schistosoma japonicum* and plays a key role in the transmission of human blood fluke *Schistosoma*. At present, the snail mainly inhabits the Yangtze River valley and mountains and hills in southwest China (Davis et al., 1995; Zhou et al., 2007a). More recently, investigation on the distribution of *O. hupensis* has generated a lot of interest due to its relationship with the climate and environmental changes (Yang et al., 2006; Zhou et al., 2008; Butler, 2012). The snail has been found in different ecological settings in mainland China, which include the swamps and lakes in the Yangtze River basin (parts of the Anhui, Hubei, Hunan, Jiangsu, Jiangxi, and Zhejiang provinces), the mountainous region of Sichuan and Yunnan provinces, the hilly, littoral part of Fujian province, and the karst landscape of Guangxi autonomous region (Davis et al., 1997; Zhou et al., 1999; Li et al., 2009b) (Fig. 1). As the habitat ecology varies considerably among these areas, it is likely that the ecological features of the landscape have influenced and channeled the genetic variations of *O. hupensis* into the different subspecies that we see today. Such a situation deserves both an understanding of the ecology and a detailed

Received: 7 March 2013 Accepted: 10 April 2013
* Author for correspondence. E-mail: xiaonongzhou1962@gmail.com. Tel./Fax: 86-21-64746458.

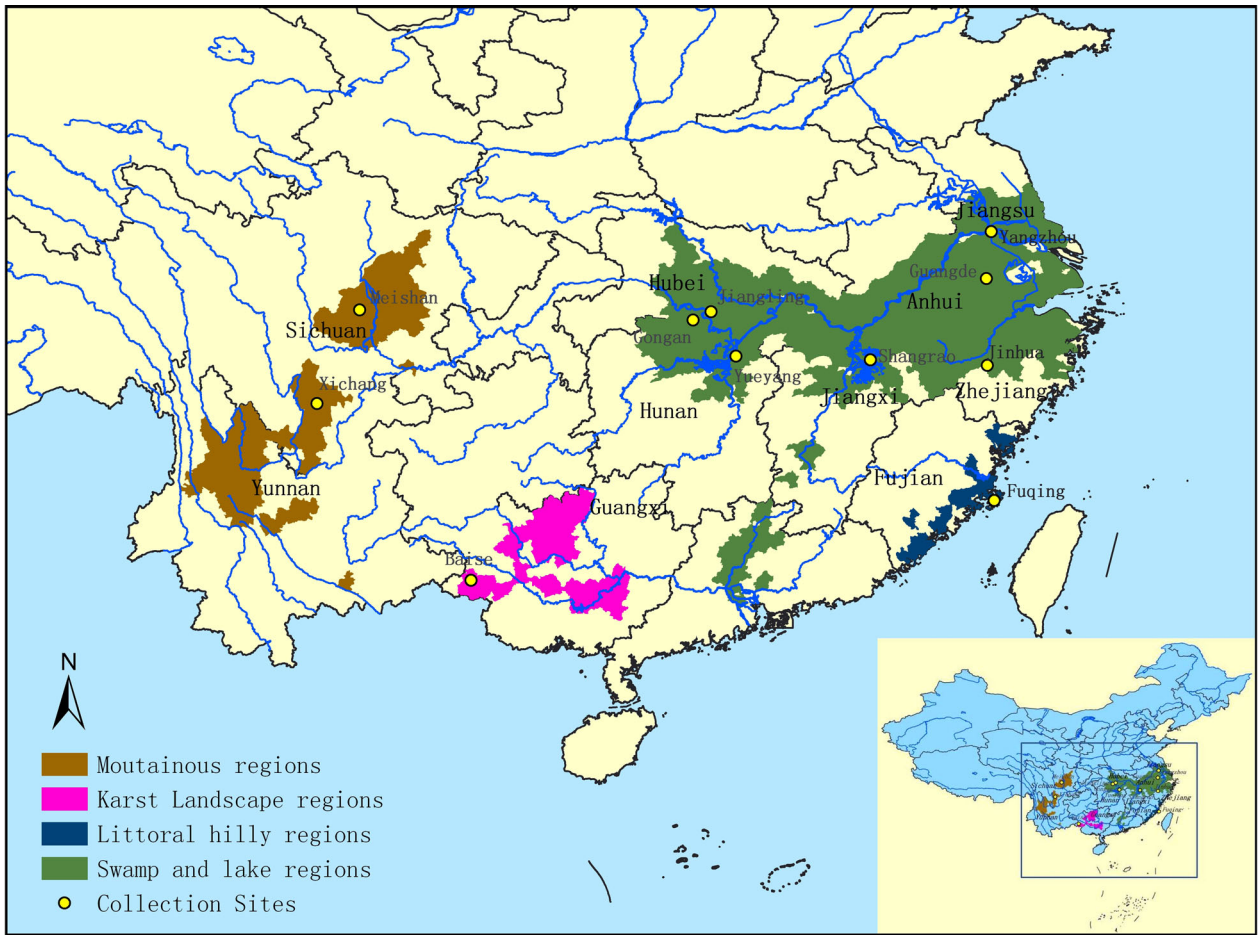

Fig. 1. Map of China showing the distribution of *Oncomelania hupensis* and collection sites. The map was constructed using ArcGIS 9.3.

knowledge of population genetics (Li et al., 2009). Accordingly, the genetic diversity and phylogeography of the snail in its different ecological settings as well as the co-evolution of *O. hupensis* and *S. japonicum* are of great interest because the branching patterns of snail diversity could be a map of the patterns of epidemiological diversity (Davis & Takada, 1969; Davis et al., 1995, 1997; Shi et al., 1999; Zhou et al., 2007b).

Oncomelania hupensis is recognized as a taxonomic group comprising seven subspecies worldwide, three of which are distributed in mainland China, namely *Oncomelania hupensis hupensis*, *Oncomelania hupensis robertsoni*, and *Oncomelania hupensis tangi* (NCBI Taxonomy Database, 2012). However, the geography-based subspecies remain controversial because of a lack of sufficient evidence from detailed geographical genetic studies. Initially, *Oncomelania hupensis guangxiensis* was considered to be the same subspecies as *O. h. tangi* (Davis et al., 1997), but it was found to have the highest genetic variation from *O. h. tangi* based on amplified fragment length polymorphisms. This genetic divergence is concordant with the geographic distribution of the subspecies, now considered a separate subspecies (Zhou et al., 1999, 2007a). Although *O. h. tangi* was also considered a discrete subspecies, its cross-breeding with *O. h. hupensis* has confused recognition of the snail (Li et al., 2009b).

Metazoan mitochondrial DNA (mtDNA) is a single circular duplex molecule ranging in size from 14 to 42 kb. The gene content of mtDNA is nearly identical, including 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes, and two ribsosomal RNA (rRNA) genes (Boore, 1999). Mitochondrial DNA has been a very useful marker for the study of genetic diversity (Bandyopadhyay et al., 2006), geographical population genetics (Zhao et al., 2012), and landscape genetics in recent years (Li et al., 2009b; Zhou et al., 2007b). A few mtDNA markers (mainly *COI*, but some studies focus on *Cytb* and *16S* genes) have been used to explore the population genetics structure or the relationship between geography-based populations of *O. hupensis* since the 1980s. The complete

mtDNA of *O. hupensis* has been reported and provides extensive possibilities for applying mtDNA markers (Li et al., 2007). However, there are few studies on snail population genetics using mtDNA molecular markers. Here, we report the complete mtDNA of *O. hupensis* snails from different geographical settings to better understand the relationships between geography-based snail population genetics and the evolution of *O. hupensis*, based on the comparison of these mtDNA sequences. The phylogenetic performance of *O. hupensis* mt PCGs was also analyzed.

1 Material and methods

1.1 Snail sampling

Oncomelania hupensis snails were collected from 10 sites in nine provinces in mainland China between 2007 and 2010, one sample from each collection site, representing four landscape populations, and incorporated the three isolates available in GenBank (NC013073, HNYYP; NC012899, HBGAP; and NC013187, SCXCP). We also included an *Oncomelania hupensis quadrasi* isolate (coded as FLB) collected from the Philippines as an outgroup for the phylogeny analysis. All the samples were bred in the laboratory for at least 1 week. The samples negative to *S. japonicum* were selected for subsequent study. Their geographical origins and locations are summarized in Table 1.

1.2 DNA preparation

Muscle tissue was isolated from snail soft body for each single isolate and digested with proteinase K (Amresco, Solon, OH, USA) for 3 h at 56 °C followed by DNA extraction (Parayre et al., 2007) using an Easy-DNA Kit (Qiagen, Hilden, Germany).

1.3 DNA amplification and sequencing

We used 16 pairs of primers, designed according to the alignment of three mt genomes of *O. hupensis* (Li et al., 2007; Zhao et al., 2010a) to amplify overlapping fragments of the entire mt genome (Table 2).

The gene sequences were amplified on a C1000 Thermal Cycler (Bio-Rad, Hercules, CA, USA). The reaction mixture contained 2.5 units Taq polymerase, 1× polymerase chain reaction (PCR) buffer (Mg^{2+} free), 0.33 mmol/L dNTPs, 2.0 mmol/L $MgCl_2$, 0.2 μmol/L primers, and 150 ng genomic DNA in a final volume of 25 μL. The PCR cycling was as follows: (i) initial denaturation at 94 °C for 5 min; (ii) 10 cycles of 94 °C for 30 s, annealing at an initial temperature 60 °C for 30 s, then declined by 1 °C per cycle, 72 °C elongation for 2 min; (iii) 25 cycles of 94 °C for 30 s, annealing at 55 °C for 30 s and 72 °C elongation for 2 min; and (iv) final elongation at 72 °C for 10 min. The resultant PCR products were resolved by electrophoresis in 1% agarose and the bands were cut out of the gel with a QIAQuick PCR purification kit (Qiagen). The sequences were identified by an ABI 377 automated sequencer (Shanghai Ding'an, Shanghai, China).

1.4 Sequence assembly and annotation

The DNA sequences were aligned and edited using CLUSTALX 1.83 (Thompson et al., 1997) and BioEdit 7.0.1 (Hall et al., 1999). The primers corresponding to two ends of the sequences were identified. The Editseq and Seqman programs in the DNAstar package (Almeida et al., 2004) were then used for sequence assembly and annotation. Protein-coding genes and rRNA genes were identified according to three reference mtDNA sequences published in NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The tRNAs were identified by tRNAscan-SE Search Server version 1.21 (Lowe & Eddy, 1997). The putative tRNAs that failed to be found by tRNAscan-SE were identified by sequence comparison. The complete mt genome of *O. hupensis* has been deposited in the GenBank database (Accession No. JF284686–JF284698, except JF284689 and JF284691).

Table 1 Geographical distribution of *Oncomelania hupensis* samples used in this study

Sequence ID	Site ID	Province	Country	Longitude (E)	Latitude (N)	GenBank Accession No.
AHGD1	Guangde	Anhui	China	119.4432	31.0675	JF284686
AHGD2	Guangde	Anhui	China	119.4432	31.0675	JF284687
JSYZ	Yangzhou	Jiangsu	China	119.5707	32.2651	JF284688
HBJL	Jiangling	Hubei	China	112.4206	30.2261	JF284690
HNYY	Yueyang	Hunan	China	113.0595	29.0861	JF284692
JXSR	Shangrao	Jiangxi	China	116.4887	28.9931	JF284693
ZJJH	Jinhua	Zhejiang	China	119.4688	28.8523	JF284694
FJFQ	Fuqing	Fujian	China	119.6471	25.4049	JF284695
GXBS	Baise	Guangxi	China	106.2997	23.3655	JF284696
SCMS	Meishan	Sichuan	China	103.4562	30.2701	JF284697
FLB	Philippines	na	Philippines	121.4422	12.5470	JF284698

na, Not available.

Table 2 Amplification primers for *Oncomelania hupensis* mitochondrial genome

Primer code	Primer sequence	Site	Length	PCR fragment size (bp)
MT.F1	AACAAATCATAAAGATATTGGGAC	18	24	912
MT.R1	GCAATAATTATCGTAGCCGC	929	20	
MT.F2	TCAGCTAAGAAAGAAACGTTTG	775	22	1546
MT.R2	CAATTGAGGCATTAAAGAATACT	2320	23	
MT.F3	CAATCATTCATTTATACCAATTGT	2158	24	1376
MT.R3	ACAAGCAGTGTTTAGGGCAC	3533	20	
MT.F4	CATTTGTTGGGGAGAATTAAC	3265	21	1307
MT.R4	CTTTTCAGCGAGAGCGAC	4571	18	
MT.F5	CGTCAAATCAAGGTACAGCC	4392	20	1216
MT.R5	GCTCGATAGGGTCTTCTTGTC	5607	21	
MT.F6	TACTCTGACCGTGCGAAGG	5447	19	778
MT.R6	GCGACTGCTAATAAAATACAGAT	6224	23	
MT.F7	GTGAGCCAGGTCAGTTTCT	5910	19	1280
MT.R7	AATATAATTACTGTCATTCAGGAC	7189	24	
MT.F8	TGACATGAAAAAGATTTTTACC	7028	22	1183
MT.R8	CGAGTTAATGTTGCATTATCA	8210	21	
MT.F9	TCTTATTGTGGTAGTAAAAATTTG	7610	24	1373
MT.R9	GTTAAAACTAAATAAAATGTTCTCAT	8982	26	
MT.F10	CTTGAAAATAAACTAAGAGTGCTC	8841	24	1007
MT.R10	AAGCTGTGAAAATTAAAAGAGTT	9847	23	
MT.F11	CCAACTCTCCTTATTATTTTAGG	9616	23	1074
MT.R11	AACATGGTTGGGTAAAATTAAA	10 689	22	
MT.F12	CTCCTTCACTGAATTCCTGC	10 552	20	744
MT.R12	GGTAGCCAGGCTGAAAAT	11 295	18	
MT.F13	ACTTACAATTATAGCAAATCGAAT	11 110	24	1328
MT.R13	ATCTTCAGTGCCATGCTCTA	12 437	20	
MT.F14	CAAGTATTATAAAATCAGTTGATCA	12 225	25	1306
MT.R14	CTCTACTTTACGGGCCGA	13 530	18	
MT.F15	TTCCTAAATGACTGAGAATAAGTG	13 329	24	1145
MT.R15	TGAAACGGAAAAATTCCAG	14 473	19	
MT.F16	TTGTCCTTTTTCTTATGTTTTCA	14 119	23	1249
MT.R16	CAAATGCATGTGCTGTAACA	181	20	

PCR, polymerase chain reaction.

1.5 Sequence and phylogeography analysis

Oncomelania hupensis quadrasi mtDNA was used as the outgroup in the phylogeography analysis. The analysis of the composition of the nucleic acid, the content of A + T, bias and frequency of PCG codons, nucleotide differentiation, and the calculation of the genetic distances between the different isolates were carried out in MEGA version 5.0 (Tamura et al., 2007, 2011).

The phylogenetic relationships were determined on the aligned sequences of the mt genome using the methods of minimum evolution (ME), neighbor joining (NJ), maximum likelihood (ML), and maximum parsimony (MP) algorithms in MEGA 5.03 and PAUP 4.0, and using Bayesian tree estimate in MrBayes 3.1.2. We applied the best-fit model of sequence evolution using the Akaike informative criterion implemented in jModeltest (Posada et al., 2008). All bootstrap support (BP) was assessed by 1000 replicates for ME, NJ, ML, and MP trees.

The phylogenetic trees based on complete mt genome sequences were considered as expected and ideal trees and compared with the 65 PCG trees derived from the 13 PCGs using the ME, NJ, ML, MP, and Bayesian methods. The consensus trees from the comparison pairs with BP values above 51% in all nodes were considered correct. Then the classification accuracy of all the phylogenetic trees derived from PCGs was analyzed by calculating the correct percentage (Galewski et al., 2006; Chen et al., 2008). The phylogenetic performance of PCGs at the geographical level was identified according to the formula $d_t = 2[\min(q_r, q_t) - r] + |q_r - q_t|$ where d_t represents the distance between the acquired phylogeny tree and the realistic phylogeny tree, q_r is the number of the clades of the acquired phylogeny tree, and q_t is the response for the realistic phylogeny tree. The value of r is the number of clades shared between the two trees (Rzhetsky & Nei, 1992).

2 Results

2.1 Sequence analysis

A total of 11 complete mt genomes of *Oncomelania hupensis* were characterized (GenBank Accession No. JF284686–JF284698, except JF284689 and JF284691). The mt genome sizes range from 15 183

to 15 215 bp, with the G + C content ranging from 32.4% to 33.4%. The mt genome contains all 37 genes typically present in mollusks, namely 13 PCGs, which include cytochrome *c* oxidase subunits I–III (*COI*, *COII*, *COIII*), ATPase subunit 6 and 8 (*ATP6*, *ATP8*) and NADH dehydrogenase subunits 1–6 and 4L (*ND1–6*, *ND4L*); 22 tRNA genes, two rRNA genes, and a short A + T-rich region. Eight of 22 tRNA genes ($tRNA^{Met}$, $tRNA^{Tyr}$, $tRNA^{Cys}$, $tRNA^{Trp}$, $tRNA^{Gln}$, $tRNA^{Gly}$, $tRNA^{Glu}$, and $tRNA^{Thr}$) were located in the light chain and the remainder were found in the other chain. The gene organization is the same as that formerly described (Li et al., 2007), except that the A + T-rich region is located between the $tRNA^{Phe}$ gene and the *COIII* gene (Fig. 2). The short overlap between contiguous genes varies from 1 to 30 bases; the longest is located between *ATP6* and $tRNA^{met}$ (Table 3).

Combined with the three reference mt genomes in GenBank, 15 251 characters were generated from the multiple sequence alignment, and 2711 (17.8%) were polymorphic. More than half (1524) of these polymorphic sites were informative with respect to parsimony. All the polymorphic sites were scattered in the sequences without any marked regions of concentration. There were 11 haplotypes among all the 11 mt genomes.

The genetic distances across the mt genomes ranged from 0.010 between JXSR isolate and HBJL isolate collected from the swamps and lakes region, to 0.106 between high-altitude region isolates (mountainous region of Sichuan and karst region of Guangxi), with an average genetic distance of 0.057. The population from the mountainous region (Sichuan) was found to be the most divergent. It differed from other isolates by a genetic distance ranging from 0.101 to 0.106 (Table 4).

2.2 Phylogenetic analysis

Phylogenetic trees based on 14 unique haplotypes were generated using ME, NJ, MP, ML algorithms, and Bayesian tree estimates with the *O. h. quadrasi* isolate as the outgroup. The same phylogenetic topology was obtained (Fig. 3). In comparison to all isolates, two main distinct clades in the phylogenetic trees were noted: (i) the isolates from the mountainous regions (Xichang and Meishan county of Sichuan province); and (ii) the remaining isolates. The latter included three inner clades, that is, the isolate from the Guangxi karst region, the isolate from Fuqing county of Fujian Province, and the isolates from the swamps and lakes region.

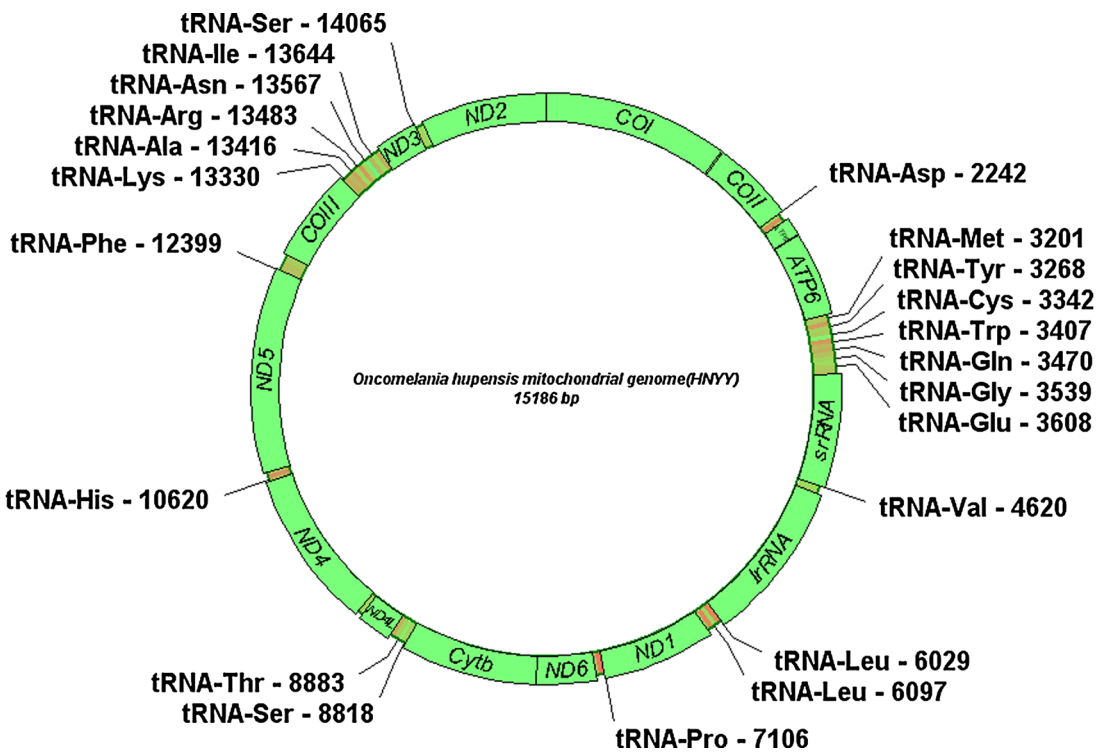

Fig. 2. Gene map for the mitochondrial genome of *Oncomelania hupensis*.

Table 3 Gene order and location of mitochondrial genome of *Oncomelania hupensis*

Gene/ Sequence ID	AHGD1	AHGD2	JSYZ	HBJL	HNYY	JXSR	ZJJH	FJFQ	GXBS	SCMS	FLB
COI	1–1536	1–1536	1–1536	1–1536	1–1536	1–1536	1–1536	1–1536	1–1536	1–1536	1–1536
COII	1553–2239	1553–2239	1553–2239	1553–2239	1553–2239	1553–2239	1553–2239	1553–2239	1553–2239	1553–2239	1553–2239
tRNA^{Asp}	2244–2310	2244–2310	2244–2310	2242–2308	2245–2311	2244–2310	2242–2308	2242–2308	2244–2310	2247–2313	2242–2307
ATP8	2309–2467	2309–2467	2309–2467	2309–2467	2309–2467	2309–2467	2309–2467	2309–2467	2309–2467	2310–2465	2309–2467
ATP6	2475–3170	2475–3170	2475–3170	2475–3170	2475–3170	2475–3170	2475–3170	2475–3170	2475–3170	2476–3171	2475–3170
tRNA^{Met}	3203–3268	3203–3268	3203–3268	3201–3266	3204–3269	3203–3268	3201–3267	3201–3266	3203–3268	3206–3270	3201–3266
tRNA^{Tyr}	3270–3334	3270–3334	3270–3334	3268–3332	3271–3335	3270–3334	3269–3333	3268–3332	3270–3334	3272–3336	3268–3332
tRNA^{Cys}	3344–3407	3344–3407	3344–3407	3342–3405	3345–3408	3344–3407	3334–3406	3342–3405	3344–3407	3345–3408	3342–3405
tRNA^{Sec}	3409–3474	3409–3474	3409–3474	3407–3472	3410–3475	3409–3474	3408–3473	3407–3472	3409–3474	3410–3475	3407–3471
tRNA^{Gln}	3470–3535	3473–3538	3473–3538	3470–3535	3470–3535	3470–3535	3474–3539	3470–3535	3470–3535	3469–3535	3469–3535
tRNA^{Gly}	3541–3606	3541–3606	3541–3606	3539–3604	3542–3607	3541–3606	3540–3605	3539–3604	3541–3606	3543–3608	3539–3604
tRNA^{Glu}	3610–3676	3610–3676	3610–3676	3608–3674	3611–3677	3610–3676	3609–3675	3608–3674	3610–3676	3614–3680	3608–3674
SsRNA	3675–4619	3675–4619	3677–4616	3675–4620	3675–4620	3675–4619	3677–4619	3675–4621	3675–4621	3677–4622	3675–4620
tRNA^{Val}	4622–4688	4622–4688	4620–4686	4621–4687	4623–4689	4622–4688	4621–4687	4622–4688	4624–4690	4627–4693	4621–4686
LsRNA	4687–6027	4687–6027	4687–6025	4688–6029	4688–6028	4689–6027	4690–6030	4688–6029	4689–6029	4790–6030	4687–6026
tRNA^{Leu}	6064–6131	6031–6098	6028–6095	6030–6097	6032–6099	6030–6097	6031–6098	6030–6097	6032–6099	6035–6102	6027–6094
tRNA^{Leu}	6132–6200	6099–6167	6096–6164	6098–6166	6100–6168	6098–6166	6099–6167	6098–6166	6100–6168	6103–6171	6095–6163
ND1	6199–7138	6166–7102	6163–7099	6167–7103	6166–7105	6165–7104	6168–7104	6167–7103	6167–7103	6168–7104	6164–7100
tRNA^{Pro}	7141–7207	7108–7174	7106–7171	7107–7173	7110–7175	7107–7173	7108–7174	7107–7173	7109–7175	7112–7178	7104–7170
ND6	7206–7706	7173–7673	7170–7670	7174–7674	7173–7673	7172–7672	7175–7675	7174–7674	7174–7674	7175–7675	7171–7677
Cytb	7711–8850	7678–8817	7675–8814	7679–8818	7678–8817	7677–8816	7680–8819	7679–8818	7679–8818	7680–8819	7676–8815
tRNA^{Ser}	8851–8915	8818–8883	8816–8883	8819–8883	8818–8882	8817–8881	8821–8888	8819–8883	8819–8888	8820–8884	8816–8880
tRNA^{Thr}	8918–8983	8885–8950	8882–8947	8884–8949	8886–8951	8884–8949	8885–8950	8884–8949	8886–8951	8889–8954	8946–8881
ND4L	8990–9286	8957–9253	8954–9250	8958–9254	8957–9253	8956–9252	8959–9255	8958–9254	8958–9254	8958–9254	8957–9253
ND4	9325–10 650	9292–10 617	9244–10 614	9293–10 618	9292–10 617	9291–10 616	9249–10 619	9293–10 618	9293–10 618	9293–10 618	9292–10 617
tRNA^{His}	10 655–10 719	10 622–10 686	10 619–10 683	10 621–10 685	10 623–10 687	10 621–10 685	10 622–10 686	10 621–10 685	10 623–10 687	10 625–10 689	10 620–10 684
ND5	10 718–12 430	10 685–12 397	10 682–12 394	10 686–12 398	10 685–12 397	10 684–12 396	10 687–12 399	10 686–12 398	10 686–12 398	10 686–12 398	10 685–12 397
tRNA^{Phe}	12 434–12 500	12 401–12 467	12 398–12 464	12 400–12 466	12 402–12 468	12 400–12 466	12 401–12 467	12 400–12 466	12 402–12 468	12 404–12 470	12 399–12 465
COIII	12 571–13 350	12 538–13 317	12 536–13 315	12 540–13 319	12 539–13 318	12 537–13 316	12 540–13 319	12 537–13 316	12 536–13 315	12 535–13 314	12 535–13 314
tRNA^{Lys}	13 364–13 432	13 331–13 399	13 329–13 398	13 331–13 400	13 333–13 401	13 330–13 398	13 331–13 399	13 328–13 396	13 329–13 397	13 330–13 397	13 326–13 394
tRNA^{Ala}	13 449–13 515	13 417–13 484	13 416–13 482	13 418–13 484	13 419–13 485	13 416–13 482	13 417–13 483	13 414–13 480	13 415–13 481	13 416–13 481	13 414–13 480
tRNA^{Arg}	13 516–13 584	13 485–13 553	13 483–13 551	13 485–13 553	13 486–13 554	13 483–13 551	13 484–13 552	13 481–13 549	13 482–13 550	13 482–13 550	13 481–13 549
tRNA ^{Asn}	13 600–13 665	13 569–13 635	13 567–13 633	13 569–13 635	13 570–13 636	13 567–13 633	13 568–13 634	13 565–13 631	13 566–13 632	13 568–13 634	13 565–13 631
tRNA^{Ile}	13 674–13 740	13 644–13 710	13 641–13 708	13 645–13 711	13 644–13 710	13 642–13 708	13 644–13 711	13 642–13 708	13 641–13 707	13 642–13 708	13 641–13 707
ND3	13 741–14 094	13 711–14 064	13 709–14 062	13 712–14 065	13 711–14 064	13 709–14 062	13 712–14 065	13 709–14 063	13 708–14 062	13 709–14 063	13 708–14 062
tRNA^{Ser}	14 095–14 162	14 065–14 132	14 063–14 130	14 066–14 133	14 065–14 132	14 063–14 130	14 066–14 133	14 063–14 130	14 062–14 129	14 063–14 130	14 062–14 129
ND2	14 163–15 215	14 133–15 185	14 131–15 183	14 134–15 186	14 133–15 185	14 131–15 183	14 134–15 186	14 131–15 183	14 130–15 179	14 131–15 183	14 130–15 182

Table 4 Genetic distance of *Oncomelania hupensis* based on mitochondrial genome

	AHGD1	AHGD2	JSYZ	ZJJH	HNYYP	HBJL	JXSR	HBGAP	HNYY	FJFQ	GXBS	SCMS	SCXCP	FLB
AHGD1														
AHGD2	0.022													
JSYZ	0.023	0.024												
ZJJH	0.023	0.025	0.017											
HNYYP	0.024	0.028	0.021	0.022										
HBJL	0.021	0.023	0.022	0.023	0.023									
JXSR	0.022	0.023	0.022	0.023	0.023	0.010								
HBGAP	0.023	0.026	0.026	0.027	0.025	0.013	0.016							
HNYY	0.022	0.024	0.024	0.023	0.020	0.016	0.016	0.019						
FJFQ	0.036	0.037	0.036	0.036	0.038	0.035	0.035	0.036	0.036					
GXBS	0.045	0.047	0.046	0.047	0.049	0.046	0.046	0.047	0.046	0.050				
SCMS	0.098	0.098	0.097	0.098	0.098	0.097	0.098	0.098	0.098	0.100	0.105			
SCXCP	0.101	0.103	0.100	0.103	0.102	0.101	0.101	0.103	0.101	0.102	0.106	0.065		
FLB	0.102	0.102	0.103	0.104	0.104	0.102	0.102	0.102	0.103	0.103	0.106	0.113	0.114	

Data of HNYYP, HBGAP, and SCXCP are from GenBank; for other abbreviations, refer to Table 1.

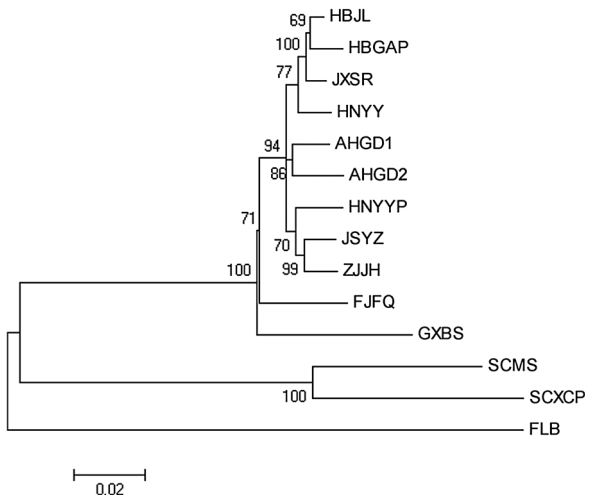

Fig. 3. Mitochondrial genome phylogeny tree of *Oncomelania hupensis* using minimum evolution method.

2.3 Phylogenetic performance analysis of PCGs

2.3.1 Differences analysis The entire PCG and rRNA gene locations were detected with complete mtDNA sequence alignment and remote Blast searching, then the PCG sequence alignments were analyzed separately. The A + T contents, and proportion of polymorphic sites and parsimony sites were calculated and compared with the complete mtDNA sequences. Except for *ATP8* and *ND4L* genes, all genes are similar to the complete mtDNA sequences, with A + T contents of 60.1%–69.3%, and the polymorphic sites ratio ranging from 17.6% to 24.3%. However, *ATP8* and *ND4L* genes were found to be considerably different from the complete mtDNA sequences. The size of the *ATP8* and *ND4L* genes are <300 bp with a higher A + T content. The *ATP8* gene looks relative conservative with only 11 haplotypes, whereas the *ND4L* gene has a higher difference in singleton variation (Table 5).

2.3.2 Phylogenetic analysis and classification accuracy The best-fit model of sequences was obtained (Table 6). The phylogenetic analysis of all 13 PCGs was carried out separately using ME, NJ, ML, and MP

Table 5 Genetic variation of protein-coding genes from *Oncomelania hupensis* mitochondrial genome

Gene	Length (bp)	GC content	Polymorphic sites (%)	Parsimony sites (%)	Singleton variation (%)
COI	1536	37.4%	295 (19.2)	158 (10.2)	137 (8.9)
COII	687	35.7%	131 (19)	75 (10.9)	56 (8.1)
ATP8	159	27. 6%	28 (17.6)	16 (10.1)	12 (7.5)
ATP6	696	31.9%	128 (18.3)	85 (12.2)	43 (6.1)
ND1	937	31.8%	183 (19.5)	91 (9.7)	92 (9.8)
ND6	501	30.8%	122 (24.3)	61 (12.1)	61 (12.1)
Cytb	1140	35.5%	248 (21.7)	142 (12.4)	106 (9.2)
ND2	1053	30.7%	237 (22.5)	140 (13.3)	97 (9.2)
ND4	1371	31.6%	304 (22.1)	188 (13.7)	116 (8.4)
ND5	1713	32.4%	400 (23.3)	256 (14.9)	144 (8.4)
COIII	780	39.9%	151 (19.4)	94 (12.1)	57 (7.3)
ND3	354	33.1%	76 (21.4)	40 (11.3)	36 (10.2)
ND4L	297	31.9%	52 (17.5)	36 (10.2)	16 (5.3%)

Table 6 Best-fit model of 13 protein-coding genes of *Oncomelania hupensis*

Gene	Best model
COI	TVM + I + G
COII	HKY + vG
COIII	K81uf + I + G
ATP6	HKY + I + G
ATP8	HKY + I
ND1	TIM + G
ND6	TrN + I + G
Cytb	K81uf + G
ND2	K81uf + I + G
ND3	HKY + G
ND4	GTR + I + G
ND4L	HKY + G
ND5	HKY + G

algorithms, and Bayesian tree estimate methods by running 1000 bootstrap replicates (Fig. 4). The complete mtDNA tree was regarded as an ideal topology and compared with the PCG trees.

According to the degree of similarity to the ideal tree, all the PCG trees were divided into good, medium, and poor gradations based on the topological structure and BP values. As a result, *COI*, *ND5*, *ATP6*, and *ND2* genes were classified as good with a classification accuracy ranging from 81.8% to 87.87%, followed by *ND3*, *COII*, *Cytb*, *ND4*, *COIII*, *ND1*, and *ND6* genes, with classification accuracy ranging from 60.57% to 78.77%. The lowest classification accuracy was found in *ATP8* and *ND4L* (<60%) (Table 7).

2.3.3 Phylogenetic performance The topological distances between PCG trees and the ideal tree were calculated at the nucleotide level. The d_t values were

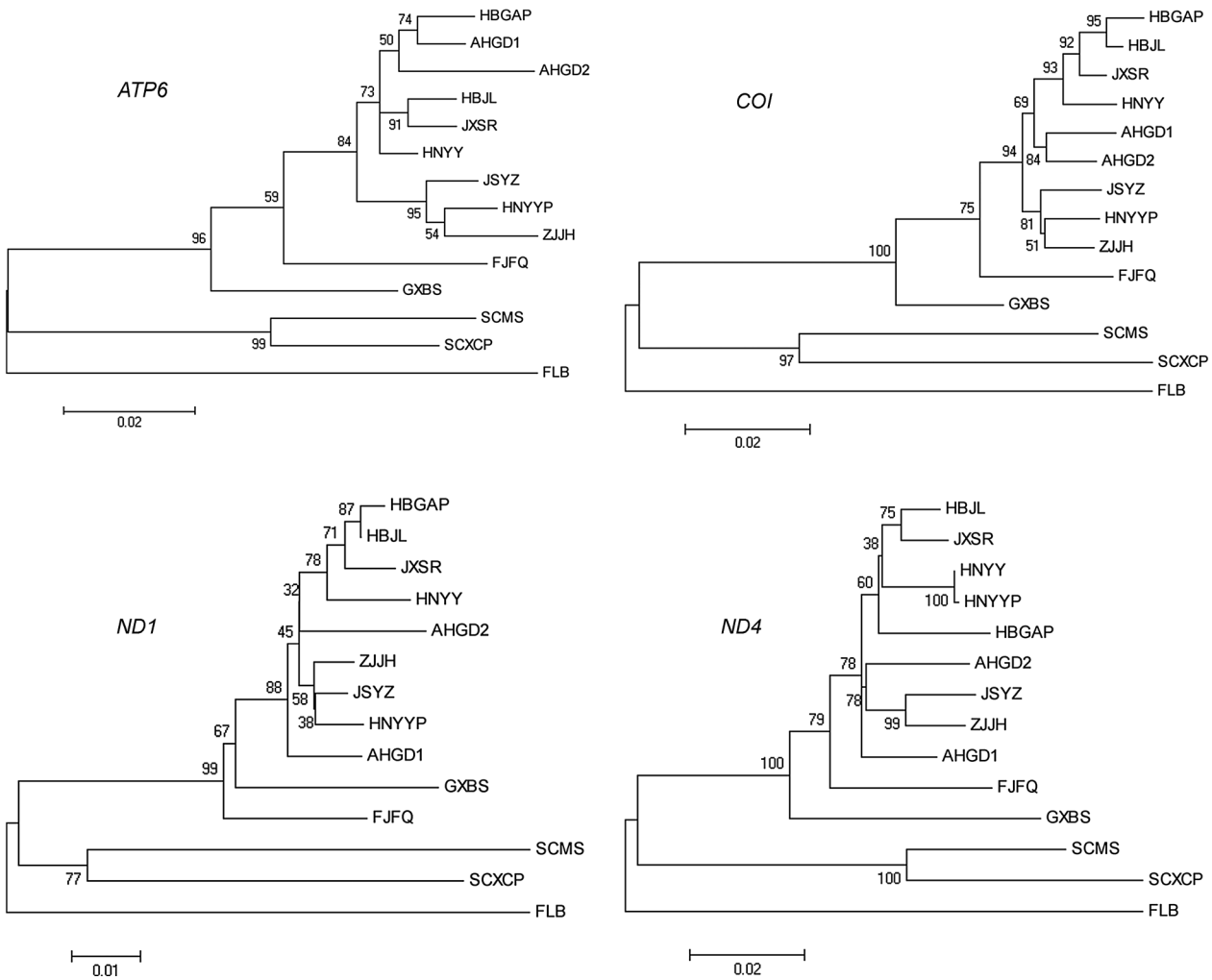

Fig. 4. Phylogeny trees of protein-coding genes from *Oncomelania hupensis* mitochondrial genome.

Table 7 Phylogeny information between 13 protein-coding genes with classification accuracy and topology distance (d_t value) of *Oncomelania hupensis* mitochondrial genome

Items\genes	*COI*	*COIII*	*ATP8*	*ATP6*	*ND1*	*ND6*	*Cytb*	*ND2*	*ND4*	*ND5*	*COIII*	*ND3*	*ND4L*
Phylogenetic tree													
NJ	G	M	P	M	M	G	P	G	M	G	M	G	P
ME	G	M	P	M	M	G	M	G	M	G	M	G	P
MP	G	M	P	M	P	M	P	G	M	G	G	G	P
Classification accuracy													
NJ	G (100%)	M (72.7)	P (27.3)	G (81.8)	M (72.7)	M (63.6)	G (81.8)	G (81.8)	M (72.7)	G (81.8)	G (81.8)	G (81.8)	P (54.5)
ME	G (81.8)	M (72.7)	P (18.2)	G (100)	M (63.6)	M (63.6)	M (72.7)	G (90.9)	M (72.7)	G (90.9)	G (81.8)	G (81.8)	P (45.5)
MP	G (81.8)	P (54.5)	P (18.2)	M (72.7)	P (45.5)	P (54.5)	M (72.7)	M (72.7)	M (63.6)	G (81.8)	M (72.7)	M (72.7)	P (45.5)
d_t value													
NJ	G (6)	M (12)	P (16)	M (9)	M (9)	M (8)	M (10)	G (6)	G (7)	M (10)	M (11)	M (10)	P (16)
ME	G (7)	M (12)	M (10)	M (9)	G (6)	M (8)	M (10)	G (4)	M (8)	M (10)	M (8)	G (6)	P (12)
MP	G (4)	M (14)	P (16)	P (12)	G (6)	M (8)	P (16)	G (4)	M (10)	M (8)	P (12)	M (8)	P (14)
General evaluation	G	M	P	M	M	M	M	G	M	G	M	G	P

Percentages are shown in parentheses. G, good; M, medium; ME, minimum evolution; MP, maximum parsimony; NJ, neighbor joining; P, poor.

classified into good ($d_t \leq 7$), medium ($7 < d_t \leq 14$), and poor condition ($d_t > 14$). The d_t values of 13 PCG trees listed in ascending order are *ND2*, *COI*, *ND1*, *ND6*, *ND3*, *ND4*, *ND5*, *ATP6*, *COIII*, *Cytb*, *COII*, *ATP8*, and *ND4L* (Table 7).

Integrating the phylogenetic information including diversity, classification accuracy, and the topological distance of phylogenetic trees between all the PCGs and the complete mt genome sequences, the performance of PCGs can be divided into good (*COI*, *ND2*, *ND5*, and *ND3*), medium (*COII*, *ATP6*, *ND1*, *ND6*, *Cytb*, *ND4*, and *COIII*), and poor (*ATP8* and *ND4L*) (Table 7).

3 Discussion

Oncomelania hupensis is the unique intermediate host snail of *S. japonicum*, which causes the most prevalent parasitic infection in mainland China (Zhou et al., 2005; Yang et al., 2010).

Therefore, it is of practical importance to construct a robust ecological and phylogenetic classification of this snail (Davis et al., 1997). However, the fine-scale geographical phylogeny, which can provide a firm base for surveillance of the source of infections and vector tracing, has not yet been established (Sun et al., 2011; Zhou et al., 2012). The phylogenetic analysis of landscape populations revealed by sensitive molecular markers will shed light on interspecific differentiation, which has resulted from the combined effect of various factors, including environmental influences, on the gene flow in *O. hupensis* populations (Zhao et al., 2010b). The mtDNA marker has been widely used as a powerful tool in population genetics, especially in the classification of relative species or geographical subspecies. However, there are few population genetics studies of *O. hupensis* using mtDNA markers (Simon et al., 2006; Wilke et al., 2006). We analyzed the phylogenetic performance of the complete mtDNA and 13 PCGs in order to provide more potential DNA markers.

In this study, 11 mt genomes of *O. hupensis* collected from four ecological landscape populations in mainland China and the Philippines were obtained. The isolates from Yunnan province were not included in the mountainous region population, although, as they share the same ecological settings and genetic structure (Li et al., 2009b), as well as close geographical distance to Sichuan province (Zhou et al., 2007c), the mountain population was adequately represented by the isolates from Sichuan province. The sequence lengths ranged from 15 183 to 15 216 bp, with GC contents of 32.3%–33.4%. The nucleotide composition, gene organization, and structure of tRNAs were the same as previously described (Li et al., 2007) and very similar to *Lophiotoma cerithiformis* (15 380 bp, A + T content of 67.9%) (Bandyopadhyay et al., 2006) and *Conus textile* (15 562 bp, 65.2%) (Bandyopadhyay et al., 2008), which are in the same suborder of Hypsogastropoda. The mtDNA gene codon has distinctive base bias compared to the other mt genomes sequenced so far, with higher A and T content in all genes. In particular, at the third codon position of PCGs, characterized by codon degeneracy and swinging, up to 76% higher A + T content was found.

The phylogenetic results showed that the isolates collected from four ecological settings were clearly divided into two major clades, one consisting of the isolates from the mountainous region, and the other consisting of the isolates from swamps and lakes in the Yangtze River basin, the littoral, hilly part of the Fujian province, and the karst landscape of Guangxi autonomous region. The geographical populations from the mountainous region showed genetic differences related

with geographical distance. In contrast, the populations from swamps and lakes in the Yangtze River basin were complicated because of their wide geographical spread and higher genetic variation. Therefore, it is difficult to draw a clear genetic relationship in this population; the Zhejiang population was particularly controversial in a study of rDNA markers (Li et al., 2009b). As for the obvious biogeographical separation, the populations from the littoral, hilly part of Fujian province (*O. h. tangi*) and the karst landscape of Guangxi autonomous region (*O. h. guangxiensis*) would be considered subspecies (Zhou et al., 2007a, 2007c). The *O. h. tangi* population shows a closer relationship to the population of the swamps and lakes region.

The phylogenetic performance of mt genes has been evaluated in a few species, including teleostean (Chen et al., 2008), tetrapods (Azuma et al., 2008; Yagishita et al., 2009) and some mammals (Zardoya & Meyer, 1996; Ari et al., 2012). To further evaluate the utility of single mt PCGs and to investigate potential causes for their low efficacy in phylogenetic analysis, the statistical support of the best tree yielded by each gene against the expected tree was estimated based on the classification accuracy and topological distance (d_t value). The classification accuracy denotes the phylogenetic information based on an individual gene and the d_t value represents the steps between the best tree yielded by each gene and the expected tree. The best tree would be close to the expected tree with more valid phylogenetic information if the classification accuracy was bigger or the d_t was smaller (Chen et al., 2008).

By combining the five methods and substitution models, the best tree yielded by each PCG was compared to the expected tree, which was produced based on the complete mtDNA sequences. Both tests yielded largely congruent results, indicating that all the single mt genes, with the exception of *ATP8* and *ND4L*, supported the expected tree with equal strength. The genes can be divided into three grades: the first consists of *COI*, *ND2*, *ND5*, and *ND3*; the second includes *COII*, *ATP6*, *ND1*, *ND6*, *Cytb*, *ND4*, and *COIII*; the remaining genes (*ATP8* and *ND4L*) fall into the third grade. As the recovered trees were not significantly different from the expected tree, most mt PCGs had the potential to recover the expected tree. Our results showed that the phylogenetic performance of *ND2*, *ND3*, and *ND5* genes in the first group, *COII* and *Cytb* in the second group, and *ATP8* and *ND4L* in the third group was consistent with the findings of previous studies (Chen et al., 2008). However, *COI* and *COIII* genes showed an unstable performance in different species. For example, the performance might be better in teleostean (Miya & Nishida, 1999, 2000), but average in other vertebrates. We suggest that the rates of evolution vary considerably between different genes, species, and classification orders.

The use of single mt genes for the phylogenetic analysis of *O. hupensis* has provided evidence of genetic divergence that defines the geographical subspecies. A clear phylogenetic relationship has been established between the populations from mountains and hills and those from other regions (Davis et al., 1997; Li et al., 2009b). However, the consensus regarding the phylogenetic localities of the Guangxi population and Fujian population, as well as the relationship with the swamps and lakes populations, has not yet been achieved (Zhou et al., 2006). This is partially because few proper DNA markers were selected for the taxonomic category of *O. hupensis*. Therefore, an assessment of the mtDNA markers for phylogenetic performance is required for further detecting the geographical divergence of *O. hupensis* regardless of analysis methods or substitution models. This study provides a primary definition and classification of the mtDNA gene markers for the phylogeny analysis of *O. hupensis* geographical populations, which can contract with the geographical distance to select or combine proper genes.

4 Conclusion

The phylogeography of *Oncomelania hupensis* based on the complete mt genome shows that these snails are divided into four ecological landscape clades. The clades from Guangxi and Fujian would be intergradations during the snail evolution from mountain populations to swamps and lakes populations. Taking into account the phylogenetic information including diversity, classification accuracy, and the topological distance of phylogenetic trees between all the PCGs and the complete mt genome sequences, the performance of PCGs can be divided into good (*COI*, *ND2*, *ND5*, and *ND3*), medium (*COII*, *ATP6*, *ND1*, *ND6*, *Cytb*, *ND4*, and *COIII*), and poor (*ATP8* and *ND4L*), which provides a robust basis for establishing the real evolutionary history of *O. hupensis*.

Acknowledgements This study was supported by the National Natural Science Foundation of China (Grant Nos. 81101280, 81101275), the National Project of Important Infectious Diseases (Grant Nos. 2008-ZX10004-011, 2012ZX10004-220, 2012ZX10004-201), and the Public Health Overseas Fund of Shanghai (Grant No. GWHW201216). The funders had no role in

study design, data collection, or analysis, decision to publish, or preparation of the manuscript.

References

Almeida CA, Tardiff DF, De LJP. 2004. An introductory bioinformatics exercise to reinforce gene structure and expression and analyze the relationship between gene and protein sequences. Biochemistry and Molecular Biology Education 32: 239–245.

Ari E, Ittzes P, Podani J, Thi QC, Jako E. 2012. Comparison of Boolean analysis and standard phylogenetic methods using artificially evolved and natural mt-tRNA sequences from great apes. Molecular Phylogenetics and Evolution 63: 193–202.

Azuma Y, Kumazawa Y, Miya M, Mabuchi K, Nishida M. 2008. Mitogenomic evaluation of the historical biogeography of cichlids toward reliable dating of teleostean divergences. BMC Evolutionary Biology 8: 215.

Bandyopadhyay PK, Stevenson BJ, Cady MT, Olivera BM, Wolstenholme DR. 2006. Complete mitochondrial DNA sequence of a Conoidean gastropod, *Lophiotoma* (*Xenuroturris*) *cerithiformis*: Gene order and gastropod phylogeny. Toxicon 48: 29–43.

Bandyopadhyay PK, Stevenson BJ, Ownby JP, Cady MT, Watkins M. 2008. The mitochondrial genome of *Conus textile*, coxI-coxII intergenic sequences and Conoidean evolution. Molecular Phylogenetics and Evolution 6: 215–223.

Boore JL. 1999. Animal mitochondrial genomes. Nucleic Acids Research 27: 1767–1780.

Butler CD. 2012. Disease emergence and global change: Thinking systemically in a shrinking world. Infectious Diseases of Poverty 1: 12.

Chen SJ, He CB, Mu YL, Liu WD, Zhou ZC. 2008. Informative efficiencies of mitochondrial genes in phylogenetic analysis of teleostean. Journal of Fishery Sciences of China 15: 12–21.

Collins C, Xu J, Tang S. 2012. Schistosomiasis control and the health system in China. Infectious Diseases of Poverty 1: 7.

Davis GM, Takada T. 1969. *Oncomelania hupensis* nosophora: Electrophoretic separation of foot proteins of laboratory-reared and field-collected specimens. Experimental Parasitology 25: 193–201.

Davis GM, Zhang Y, Guo YH, Spolsky CM. 1995. Population genetics and systematic status of *Oncomelania hupensis* (Gastropoda: Pomatiaop sidae) throughout China. Malacologia 37: 133–156.

Davis GM, Zhang Y, Guo YH, Spolsky CM. 1997. Systematic status of *Oncomelania hupensis* (Gastropoda: Pomatiopsidae) throughout China. Studia Marina Sinica 39: 89–95.

Galewski T, Tilak MK, Sanchez S, Chevret P, Paradis E. 2006. The evolutionary radiation of Arvicolinae rodents (voles and lemmings): Relative contribution of nuclear and mitochondrial DNA phylogenies. BMC Evolutionary Biology 6: 80.

Hall TA. 1999. BioEdit: User-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Research Symposium Series 41: 95–98.

Li SZ, Luz A, Wang XH, Xu LL, Wang Q. 2009a. *Schistosomiasis* in China: Acute infections during 2005–2008. Chinese Medical Journal 122: 1009–1014.

Li SZ, Wang YX, Liu Q, Lv S, Wang Q. 2007. Complete mitochondrial genome sequence of *Oncomelania hupensis* (Gastropoda: Rissooidea). Chinese Journal of Parasitology and Parasitic Disease 27: 291–296.

Li SZ, Wang YX, Yang K, Liu Q, Wang Q. 2009b. Landscape genetics: The correlation of spatial and genetic distances of *Oncomelania hupensis*, the intermediate host snail of *Schistosoma japonicum* in mainland China. Geospatial Health 3: 221–231.

Lowe TM, Eddy SR. 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Research 25: 955–964.

Miya M, Nishida M. 1999. Organization of the mitochondrial genome of a deep-sea fish, *Gonostoma gracile* (Teleostei: Stomiiformes): First example of transfer RNA gene rearrangements in bony fishes. Marine Biotechnology 1: 416–426.

Miya M, Nishida M. 2000. Use of mitogenomic information in teleostean molecular phylogenetics: A tree-based exploration under the maximum-parsimony optimality criterion. Molecular Phylogenetics and Evolution 17: 437–455.

NCBI Taxonomy Database. Bethesda: National Center for Biotechnology Information (US) [online]. Available from http://www.ncbi.nlm.nih.gov/Taxonomy/ [accessed 23 September 2012].

Posada D. 2008. jModelTest: Phylogenetic model averaging. Molecular Biology and Evolution 25: 1253–1256.

Rzhetsky A, Nei M. 1992. Statistical properties of the ordinary least-squares, generalized least-squares, and minimum-evolution methods of phylogenetic inference. Journal of Molecular Evolution 35: 367–375.

Shi CH, Wilke T, Davis GM, Xia MY, Qiu CP. 1999. Population genetics, micro-phylogeography, ecology, and infectivity of Chinese *Oncomelania hupensis hupensis* (Gastropoda: Rissooidea: Pomatiopsidae) in the Miao River system: Is there a relationship to shell sculpture? Malacologia 44: 333–347.

Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT. 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. Annual Review of Ecology, Evolution, and Systematics 37: 545–579.

Sun LP, Liang YS, Wu HH, Tian ZX, Dai JR. 2011. A Google Earth-based surveillance system for *Schistosomiasis japonica* implemented in the lower reaches of the Yangtze River, China. Parasites and Vectors 4: 223.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596–1599.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28: 2731–2739.

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 Research 25: 4876–4882.

Wang LD, Chen HG, Guo JG, Zeng XJ, Hong XL. 2009. A strategy to control transmission of *Schistosoma japonicum* in China. The New England Journal of Medicine 360: 121–128.

Wilke T, Davis GM, Qiu DC, Spear RC. 2006. Extreme mitochondrial sequence diversity in the intermediate schistosomiasis host *Oncomelania hupensis robertsoni*: Another case of ancestral polymorphism? Malacologia 48: 143–157.

Yagishita N, Miya M, Yamanoue Y, Shirai SM, Nakayama K. 2009. Mitogenomic evaluation of the unique facial nerve pattern as a phylogenetic marker within the percifom fishes (Teleostei: Percomorpha). Molecular Phylogenetics and Evolution 53: 258–266.

Yang GJ, Li W, Sun LP, Wu F, Yang K, Huang YX, Zhou XN. 2010. Molluscicidal efficacies of different formulations of niclosamide: Result of meta-analysis of Chinese literature. Parasites and Vectors 3: 84.

Yang GJ, Vounatsou P, Tanner M, Zhou XN, Utzinger J. 2006. Remote sensing for predicting potential habitats of *Oncomelania hupensis* in Hongze, Baima and Gaoyou lakes in Jiangsu province, China. Geospatial Health 1: 85–92.

Zardoya R, Meyer A. 1996. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. Molecular Biology and Evolution 13: 933–942.

Zhao QP, Jiang MS, Dong HF, Nie P. 2012. Diversification of *Schistosoma japonicum* in mainland China revealed by mitochondrial DNA. PLoS Neglected Tropical Diseases 6: e1503.

Zhao QP, Jiang MS, Littlewood DT, Nie P. 2010a. Distinct genetic diversity of *Oncomelania hupensis*, intermediate host of *Schistosoma japonicum* in mainland China as revealed by ITS sequences. PLoS Neglected Tropical Diseases 4: e611.

Zhao QP, Zhang SH, Deng ZR, Jiang MS, Nie P. 2010b. Conservation and variation in mitochondrial genomes of gastropods *Oncomelania hupensis* and *Tricula hortensis*, intermediate host snails of *Schistosoma* in China. Molecular Phylogenetics and Evolution 57: 215–226.

Zhou XN, Kristensen TK. 1999. Genetic and morphological variations in populations of *Oncomelania* spp in China. Southeast Asian Journal of Tropical Medicine and Public Health 30: 166–176.

Zhou XN, Wang LY, Chen MG, Wu XH, Jiang QW, Chen XY, Zheng J, Utzinger J. 2005. The public health significance and control of schistosomiasis in China – then and now. Acta Tropica 96: 97–105.

Zhou XN, Xu J, Chen HG, Wang TP, Huang XB. 2012. Tools to support policy decisions related to treatment strategies and surveillance of *Schistosomiasis japonica* towards elimination. PLoS Neglected Tropical Diseases 5: e1408.

Zhou XN, Yang GJ, Yang K, Wang XH, Hong QB, Sun LP, Malone JB, Kristensen TK, Bergquist NR, Utzinger J. 2008. Potential impact of climate change on schistosomiasis transmission in China. American Journal of Tropical Medicine and Hygiene 78: 188–194.

Zhou YB, Jiang QW, Zhao GM. 2005. Application of amplified fragment length polymorphism in the study of genetic diversity of *Oncomelania hupensis*. Chinese Journal of Shistosomiasis Control 17: 34–38.

Zhou YB, Jiang QW, Zhao GM, Yuan HC. 2007a. Subspecies differentiation of *Oncomelania hupensis* from mainland China. Chinese Journal of Shistosomiasis Control 19: 485–487.

Zhou YB, Yang MX, Zhao GM, Wei JG. 2007b. *Oncomelania hupensis* (Gastropoda:Rissooidea), intermediate host of *Schistosoma japonicum* in China: Genetics, molecular phylogeny based on amplified fragment length polymorphisms. Malacologia 49: 367–382.

Zhou YB, Zhao GM, Peng WX, Wei JG, Jiang QW. 2007c. Spatial genetic correlation analyses of *Schistosoma japonicum* intermediate hosts within *Oncomelania hupensis* (Gastropoda: Rissooidea) from mainland China based on amplified fragment length polymorphisms. Fudan University Journal of Medical Science 34: 207–212.

Zhou YB, Zhao GM, Wei JG, Jiang QW. 2006. Study on the genetic diversity among populations of *Schistosome* intermediate hosts within *Oncomelania hupensis* (Gastropoda: Rissooidea) in mainland China. Chinese Journal of Epidemiology 27: 865–870.