

Accepted Manuscript

Title: Population genetics of *Oncomelania hupensis* snails, intermediate hosts of *Schistosoma japonicum*, from emerging, re-emerging or established habitats within China

Authors: Chen Qiu, Da-Bing Lu, Yao Deng, Hui-Ying Zou, You-Sheng Liang, Joanne P. Webster



PII: S0001-706X(18)31348-2
DOI: <https://doi.org/10.1016/j.actatropica.2019.105048>
Article Number: 105048

Reference: ACTROP 105048

To appear in: *Acta Tropica*

Received date: 23 October 2018
Revised date: 19 April 2019
Accepted date: 30 May 2019

Please cite this article as: Qiu C, Lu D-Bing, Deng Y, Zou H-Ying, Liang Y-Sheng, Webster JP, Population genetics of *Oncomelania hupensis* snails, intermediate hosts of *Schistosoma japonicum*, from emerging, re-emerging or established habitats within China, *Acta Tropica* (2019), <https://doi.org/10.1016/j.actatropica.2019.105048>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Title

Population genetics of *Oncomelania hupensis* snails, intermediate hosts of *Schistosoma japonicum*, from emerging, re-emerging or established habitats within China

Chen Qiu^{a,b,c*}, Da-Bing Lu^{a,b,c*}, Yao Deng^{a,b,c}, Hui-Ying Zou^{a,b,c}, You-Sheng Liang^{b,c,d}, & Joanne P. Webster^{d,e}

^aDepartment of Epidemiology and Statistics, School of Public Health, Soochow University, Suzhou 215123, China;

^bKey Laboratory of National Health and Family Planning Commission on Parasitic Disease Control and Prevention, Wuxi 214064, China;

^cJiangsu Provincial Key Laboratory on Parasites and Vector Control Technology, Wuxi 214064, China;

^dJiangsu Institute of Parasitic Diseases, Wuxi 214064, China;

^eCentre for Emerging, Endemic and Exotic Diseases (CEEED), Department of Pathology and Population Sciences, Royal Veterinary College, University of London, Herts AL9 7TA, United Kingdom.

*The two authors contributed equally to this work: Chen Qiu and Da-Bing Lu

Authors for Correspondence: Da-Bing Lu, email: Ludabing@suda.edu.cn; You-Sheng Liang, email: liangys58@163.com

Highlights

- Low genetic diversities of snails were found in new or re-emerged snail habitats
- The largest N_e was in the new snail habitat, rather than in the persistent habitat
- No signs of genetic bottleneck effects were detected in the new or re-emerged habitats
- Snails from three sites (65 to 450 km) were clearly separated and low gene flow were detected

Abstract

Schistosomiasis remains one of the world's most significant neglected tropical diseases, second only to malaria in terms of socioeconomic impact. In 2014, China proposed the goal of schistosomiasis japonicum elimination by 2025. However, one major challenge is the widely distributed, and in certain cases potentially increasing, habitats of *Oncomelania hupensis*, the snail intermediate hosts of *S. japonicum*. Therefore, an understanding of population genetics of *O. hupensis* in new or re-emerged habitats, together with that of the established habitats with snail persistence, would be valuable in controlling and predicting the future transmission dynamics of schistosomiasis in China. Using nine microsatellite loci, we conducted population genetic analyses of snails sampled from one habitat where snails were detected for the first time, one (previously eliminated) habitat with re-emerged snails, and one habitat with established snail persistence. Results showed lower diversities, in terms of number of observed alleles per locus (N_a), number of effective alleles per locus (N_eA), observed (H_o) and expected heterozygosity (H_e), in snails from new or re-emerged snail habitats than from the habitat with snail persistence. The smallest effective population size was inferred in the re-emerged snail habitat, but the largest was in the new habitat rather than in the habitat with snail persistence. No bottleneck effects were detected in new or re-emerged habitats. No or low sub-structure was inferred in new and persistent snail habitats. Snails from the three sites were clearly separated and low gene flow was estimated between sites. We propose that snails at the new habitat may have been introduced through immigration, whereas snails at the re-emerged habitat may be the consequence of those few snails remaining subsequently expanding through reproduction. We discuss our results in terms of their theoretical and applied implications.

Key words: population genetics; *Oncomelania hupensis*; schistosomiasis; snail habitats; bottleneck effect; effective population size; disease control.

1. Introduction

Schistosomiasis, caused by blood flukes of the genus *Schistosoma*, is one of the world's most significant neglected tropical diseases. Over 240 million people are currently infected worldwide (WHO, 2015), including a recent outbreak in developed regions (Boissier *et al.*, 2016). *Schistosoma japonicum* is mainly prevalent in China, the Philippines and isolated parts of Indonesia. In China, after decades of major efforts in schistosomiasis control, significant progress has been made. For example, the estimated number of infected humans was reduced from approximately 11.6 million in 1950s to 325,824 in 2010 (Collins *et al.*, 2012). By the end of 2015, the number of infected humans was further reduced to 77,194, and out of 453 endemic counties (city or district), 343 (75.72%) reached the criteria of transmission interruption (Zhang *et al.*, 2016). Therefore the central government of China, in line with the revised World Health Organization targets, proposed the goal of schistosomiasis elimination/interruption of transmission by 2025 (Lei and Zhou, 2015).

However, despite such progress and ambitious goals, several major challenges remain, one of which is the widely distributed habitats of *Oncomelania hupensis*, the intermediate host snails of *S. japonicum* (Zhang *et al.*, 2016). Furthermore, anthropogenic change, particularly in terms of the increasing number of projects aimed to develop natural wetlands and/or water resources, are predicted to create more potential habitats for snails (Campbell *et al.*, 2010; Xu *et al.*, 2000). Such expansion may be further facilitated through potential rising or shifting temperatures in these regions (Stensgaard *et al.*, 2018; Zhou *et al.*, 2008). Snails have the potential to spread and be introduced into new habitats either by flooding along the Yangtze River and its tributaries (Zhou *et al.*, 2002) or via carrier water birds (Boag, 1986) (facilitating long distance dispersal). The fact that, in 2015, *O. hupensis* snails were first detected in 31 villages where no snails had ever been detected since records began in 1950s, and that these new snail habitats covered a total area of 6.66 million m² (Zhang *et al.*, 2016), suggests that the habitat range for this important intermediate host species is expanding within China. Such extensive and emerging/re-emerging snail habitats, together with the broad range of mammalian species able to serve as definitive host reservoirs for *S. japonicum* (He *et al.*, 2001; Lu *et al.*, 2010b; Rudge *et al.*, 2013), and the increasing numbers of infections amongst mobile people reported each year (Qian *et al.*, 2015), all have the potential to threaten China's goal to achieve complete interruption of transmission (Lei and Zhou, 2015). Therefore, a deep understanding of the population dynamics of *O. hupensis* snails in new or re-emerged snail habitats, as well as established habitats, is of great importance, particularly in the context of snail control via focal mollusciciding and/or environment modification (Wang *et al.*, 2009).

Population genetic analyses of snails across contrasting levels of habitat establishments is an ideal method in which to help achieve this. In particular, elucidating the population genetic structure of natural snail populations may be predicted to prove highly useful in interpreting prevailing patterns of distribution, and hence infection risk, and to evaluate and predict the impact of ongoing control efforts and/or anthropogenic change in general. *O. hupensis* has been documented to show genetic divergence among different regions of China, even by using less powerful (i.e. in terms of their ability to distinguish between individuals) molecular markers including sequencing COI (Wilke *et al.*, 2000), Cytb (Spolsky *et al.*, 1996), rDNA (Hauswald *et al.*, 2011; Wilke *et al.*, 2006; Zhao *et al.*, 2010), and RFLP (Hope and McManus, 1994) or AFLP (Zhou *et al.*, 2007a). In addition, genetic monitoring using more sensitive neutral markers such as microsatellites can identify, for instance, clusters of genetically related organisms, or bottleneck effects, indicative of both historical and contemporary population level changes. A number of *O. hupensis* microsatellite markers have thus been developed and evaluated (Zhang *et al.*, 2012; Zhang *et al.*, 2010). Application of these markers on *O. hupensis* has revealed geographical variation (Guan *et al.*, 2016; Guo *et al.*, 2008; Li *et al.*, 2009), distinct clusters (Zhou *et al.*, 2007b), or distance or proportion of migrant individuals (Head *et al.*, 2016). However, to our knowledge, no investigation into the population genetic structure of *O. hupensis* snails in new or re-emerged snail habitats have been performed.

Therefore, based on nine microsatellite loci, we conducted population genetic analyses of *O. hupensis* snails sampled from three contrasting snail habitats across mainland China. Our main aims were to: 1) to test whether there were lower levels of genetic diversity, effective population sizes and/or evidence of bottleneck effects within snails from the newly emerged or re-emerged snail habitats relative to the established habitat; and 2) estimate any potential relationship and/or evidence and extent of recent migration between snail populations. Our results are discussed in terms of their theoretical and applied implications for ongoing schistosomiasis control and the potential for interruption of transmission within China.

2. Materials and methods

2.1. Sampling sites

Snails were collected from three contrasting habitat sites (Fig. 1). One was from Shitai county (ST) of Anhui province where established transmission for *S. japonicum* has not been interrupted. Due to findings of rodents

serving as the main reservoir for *S. japonicum* there (Lu *et al.*, 2010b), intensive research has been performed on the transmission dynamics of the parasite in this region (Lu *et al.*, 2010a; Shi *et al.*, 2014). In April of 2015, a snail survey was performed and a total of 245 snails were collected from a habitat with an area of 1200 m². The other two sites were from Suzhou city, once a serious epidemic area, of Jiangsu province. In 1995 the city (and its rural areas) reached the criteria to schistosomiasis elimination, and since then snail surveillance in all suspected habitats has been performed annually or biennially. In April of 2015, field surveys showed that snails re-emerged or were first found in few habitats (Suzhou Centers for Disease Control and Prevention). We then sampled a total of 243 snails from one re-emerged snail habitat (with an area of 1510 m²) in Hengtang district of Suzhou city (HT), where snails were last reported ten years ago, and a total of 220 snails from one newly emerging snail habitat (with an area of 1324 m²) in Taichang district of Suzhou city (TC). All snails collected from the field were checked for potential cercarial shedding within the laboratory following standard protocols (Dabo *et al.*, 2015), and all uninfected adult snails were stored in alcohol prior to DNA extraction. See Table 1 for detailed information on sample sites.

2.2. DNA extraction and microsatellite genotyping

DNA from each snail was extracted using an EZgene™ Mollusc gDNA Kit (Biomiga, USA) according to the manufacturer's protocols, and then stored at -25°C. Each snail was screened by nine microsatellite loci with seven previously reported (Guan *et al.*, 2016; Zhang *et al.*, 2012) and two novel microsatellite loci explicitly developed for this study (Appendix A). The forward primer for each pair was labeled with 6-FAM, HEX, TAMRA or ROX. Two separate multiplex reactions were performed with markers T6-17, B14, C22, DH01 and DH02 in multiplex one and markers T1-10, T4-25, T4-22 and D11 in multiplex two. Each PCR amplification was carried out in 15µl reactions containing 1.5µl snail DNA, 0.15-0.3 µM of each primer, 7.5µl Master Mix, and adding water to the volume, using the QIAGEN Multiplex PCR Kit. Thermo cycling (Thermo Scientific) was performed with the following PCR profile: 95°C for 5min, followed by 30 cycles of 30s at 95°C, 60s at 60°C, and 30s at 72°C, with a final extension at 65°C for 30min. PCR products were genotyped using an ABI3100 automated sequencer (Applied Biosystems) in Sangon Biotech (Shanghai, China). Alleles were scored using GeneMarker HID (SoftGenetics LLC), coupled with visual editing of the automatically scored peaks. To make a robust estimation, we genotyped a random sample of about 70 snails per site as recommended by Davis and colleagues (Davis *et al.*, 1995).

2.3. Genotypic analyses

2.3.1. Genetic diversities and substructure

We calculated the genetic diversity of *O. hupensis* snails at each site with Genalex V6.5 (Peakall and Smouse, 2012). The polymorphism of each locus was measured with the total number of alleles and of effective alleles. The descriptive statistics measured, based on multi-locus, are number of observed alleles per locus (N_a) and number of effective alleles per locus (N_{eA}), observed (H_o) and expected heterozygosity (H_e), unbiased expected heterozygosity (uH_e) and inbreeding coefficient (F_{IS}). Deviations from Hardy-Weinberg Equilibrium (HWE) for each locus and linkage disequilibrium between loci were evaluated in FSTAT (Goudet, 2001). The sequential Bonferroni's procedure was applied to correct for multiple tests (Holm, 1979).

To infer the most likely number of genetic clusters (K) present within each site, STRUCTURE v2.2 (Pritchard *et al.*, 2000) was used to conduct a Bayesian multi-locus clustering analysis. Simulation series were run with the values of K from two to ten. For each K value, a series of ten independent runs was performed under assumption of mixture models and correlated allele frequencies. A burn-in period of 0.5 million iterations was used in each run, followed by a 1.5 million Markov Chain Monte Carlo iterations. The optimal value of K , according to Evanno *et al.* (Evanno *et al.*, 2005), was identified via calculation of the statistic Delta K (ΔK), which is based on the change rate in the calculated log probability between successive K values. The peak of the ΔK graph corresponds to the inferred number of subpopulations. The online STRUCTURE HARVESTER (Earl and vonHoldt, 2012) was applied to the analyses of the result files from the STRUCTURE and output the Delta K and the estimates.

2.3.2. Bottleneck effect and effective population size

A test for the existence or not of potential recent bottlenecking events in each site was performed using the program BOTTLENECK V1.2 (Piry *et al.*, 1999). A bottleneck effect is revealed by the existence of heterozygosity excess in subsequent generations (Cornuet and Luikart, 1996) as the number of alleles (rare alleles mainly) decreases faster than heterozygosity; whereas the existence of heterozygosity deficiency may indicate a recent population expansion. Tests were performed under three different mutational models including the Stepwise Mutational Model (SMM), the Infinite Allele Model (IAM) and the Two-Phase model (TPM). The SMM is consensually the accepted mutation model for microsatellites, which assumes that STRs evolve by addition or subtraction of one or more repeat units per mutation. The IAM is additionally considered given that

sometimes the SMM is violated. The TPM is intermediate to the SMM and IAM. In the TPM, we assumed 70% of mutations at one step and 30% at a multiple-step with a variance of 30 (default in the software). As only the available nine loci (< 20 loci) were used, a Wilcoxon's sign-rank test was applied to assess the significance (Piry *et al.*, 1999). The calculations were based on 0.2 million replications.

The effective population size (N_e) refers to the number of parental individuals that effectively contribute to the next generation. We estimated the N_e values and their 95% confidence intervals (CIs) for each site with two software NeEstimator (NeEst) (Peel *et al.*, 2004) and LDNe (Waples and Do, 2008). The calculation was based on linkage disequilibrium (LD). Alleles with frequency more than or equal to 0.05 were included in the estimation progress.

2.3.3. Genetic relationship between populations

The relationship or gene flow of snails between three sites was assessed using three complementary programs. With Genalex V6.5 genetic distances between snail individuals were calculated, based on which a Principal Coordinate Analysis (PCoA) was conducted for visualization of genetic relatedness of each other. FSTAT (Goudet, 2001) was used to estimate pairwise F_{ST} values and their significance tested by permutations at the adjusted nominal level. Migration rates between three sites were estimated using a maximum likelihood coalescent approach implemented in MIGRATE V3.0 (Beerli and Felsenstein, 2001). We estimated Θ and M (immigration rate/mutation rate). We ran four short chains with a total of 100,000 genealogy samples and one long chains with 10,000,000 samples, following a burn-in of 5,000 samples.

3. Results

3.1. Genetic diversities

The nine microsatellite loci, scored in a total of 211 snails, displayed a high level of polymorphism in terms of number of different alleles, ranging from 12 alleles at locus T6-17 to 35 alleles at locus T4-25, and of number of effective alleles, varied between 5.420 at locus T4-22 and 11.354 at locus T4-25. This suggests the nine loci each are sufficiently informative to assess the degree of diversity and structuring of snail populations (Nyakaana *et al.*, 2013).

The multi-locus diversities of each site are provided in Table 2. In general, based on a qualitative comparison,

the genetic diversities, in terms of N_a , N_eA , H_o and H_e , were higher in snails from the established habitat (ST), as compared to those from the re-emerged (HT) or newly emerging (TC) habitats. Out of nine microsatellite loci, three loci (C22, D11 and T1-10) showed a significant deviation, at the adjusted P-value, from HWE at each site. Five out of 144 pairs of nine loci showed a significant Linkage Disequilibrium but the LD patterns were observed between seven loci rather than between any particular pair of loci, indicating the independence of the loci used. Snails from each site was not in HWE (all $P = 0.0056$). This was similar to the previous studies (Davis *et al.*, 1999a; Davis *et al.*, 1999b), in which all snails were sampled from localities along the Yangtze River.

The difference in genetic diversities between three sites were also evident from the STRUCTURE analyses. According to the highest likelihood and the ΔK , the potential co-ancestry (and/or subpopulation structure) appeared to be seven for ST population, and four for both HT and TC, although the bar plots showed substructures only in HT (Fig 2).

3.2. Bottleneck effect and effective population size

Bottleneck analyses showed a significant number of microsatellites had a heterozygosity excess for ST under IAM model ($P = 0.004$), indicating evidence of a recent bottleneck. ST also presented a heterozygosity deficiency under SMM ($P = 0.0019$). It was noted that HT snails showed a heterozygosity deficiency under both TPM and SMM models, and TC presented a heterozygosity deficiency under SMM. The sign of heterozygosity deficiency may indicate a potential expansion of snail populations (Table 3).

N_e based on LD estimates for TC snails were either infinite or negative, the later of which was also considered as infinite (Waples and Do, 2008). However, both NeEst and LDNe analyses revealed comparable and congruent estimates of N_e for either ST (i.e, 179.9 with NeEst and 284.9 with LDNe) or HT snails (i.e. 17.2 with NeEst and 46.1 with LDNe), with the higher values estimated for ST snails than for HT (Table 4).

3.3. Genetic relationships between three sites

The PCoA plot showed that snail individuals tended to cluster together within site (Fig. 3). Values of pairwise F_{ST} s estimated with FSTAT were 0.217 between ST and HT, 0.273 between ST and TC and 0.370 between HT and TC (all $P < 0.016$), indicating a significantly low gene flow between sites. Gene flow estimates calculated with MIGRATE also indicated the low level of gene flow between sites, with evidence of asymmetrical gene

flow in both HT and TC (Table 5). The effective number of migrants entering and leaving each site per generation (M) ranged between 1.80 (HT to TC) and 5.07 (TC to HT).

4. Discussion

Our results demonstrate that, despite similar habitat range sizes and snail densities between sites, as may be predicted, lower genetic diversities of snails were observed in both the newly emerged (TC) and re-emerged (HT) habitats relative to the persistent established habitat (ST), with STRUCTURE indicating a different co-ancestry of snails between sites. Surprisingly however, using NeEst, the smallest estimate of effective population size was inferred for the re-emerged HT and the highest for the newly emerged TC, rather than as may have been predicted for the established ST, and with neither of the former indicating any significant bottleneck effects.

The considerably higher genetic diversity observed in ST was from an endemic area where *S. japonicum* transmission has not been interrupted since records began. Habitats with the persistence of the intermediate host snails like this remains substantial across China, where, for example, by 2015 within 110 counties with uninterrupted and ongoing schistosomiasis transmission, the total snail habitats was estimated to cover 3.563 billion m² (Zhang *et al.*, 2016). The observed higher diversity could be reflective of long-term host-schistosomes co-evolutionary interactions in these areas, as snails with higher genetic diversities have been reported within *Schistosoma* infection hotspot zones across different species, for example, not only *Oncomelania hupensis* with *S. japonicum* (Shi *et al.*, 2002), but also *Biomphalaria pfeifferi* with *S. mansoni* (Webster *et al.*, 2001) and *Bulinus globosus* with *S. haematobium* (Davies *et al.*, 1999). Furthermore, it is possible that the observed evidence of recent bottlenecks amongst this established snail population only could reflect the efficacy of the ongoing snail control programs targeted here, in contrast to that within the other two sites where snails were either believed to have been eliminated (HT) or not present (TC). Indeed, previous studies have revealed an impact of molluscicide exposure on the genetic composition of snails even at the individual level (Zhao *et al.*, 2015), in which about 254 genes showed significant differential expression after exposure.

Low host diversity, as observed in the snails from newly emerging and re-emerging sites here, can be associated with lower fitness and parasite resistance in general (Coltman *et al.*, 1999; Jarne and Theron, 2001). Furthermore, reduced intra population diversity promoted by population genetic perturbations may precipitate unexpectedly severe disease outbreaks, as appears to be a factor in the major *S. mansoni* outbreak in Senegal

(Campbell et al., 2010). Therefore, we may predict that these recently established *Oncomelania* snail populations may be highly susceptible to schistosomes if exposed. It is true that such exposure remains likely across China due to the highly zoonotic nature of *S. japonicum* and hence, even if not from human contamination directly, but from animals, both livestock and in particular rodents. Indeed rodents have been demonstrated to serve as the main reservoir for the parasite in some areas (Lu et al., 2010a; Rudge et al., 2013) and such wildlife reservoirs are notoriously difficult to control globally.

We inferred very low levels of gene flow between our sites, partly due to geographical distances. The upstream ST is about 450 km away from Suzhou city where two habitats of HT and TC are located. HT is 65 km apart from TC. Both ST and TC are indirectly connected with the Yangtze River, but HT is almost completely isolated due to its central location within the city. The lack of gene flow here, in terms of isolation (for example for HT), could also have implications in for future schistosomiasis outbreaks, since curtailment of gene flow and enhanced inbreeding may similarly make intermediate host populations more vulnerable to parasite challenge (Gandon et al., 1996; Gow et al., 2004).

We inferred substructuring most clearly in HT (the re-emerged habitat) but potentially less or none (depending on the measure used) in either ST (the established) or TC (the newly emerged habitat). Combined with the location of the site, we suppose that the site HT might be a *refugia*, where some snails may have survived (and occasionally few snails may also have been imported), which subsequently reproduced, resulting in substructures. In contrast, for the site TC, although here we only detected relatively low levels of gene flow between this site and its upstream ST, mainly due to the long distance, numerous other snail habitats including marshlands and islands remain along the Yangtze River between the two sites (within Anhui and within Jiangsu provinces). Therefore, due to the connection with the Yangtze River via the small canal Liujiang, large numbers of nearby and upstream snails might have been brought into TC by an annual flooding and then colonized. The sperm a female *Oncomelania* snail received at one time is sufficient for the production of fertilized eggs during the rest of her lifetime for the female (Zhou, 2005). A potential influence of the upstream and nearby snails could be the main explanation for the estimated infinite effective population size in TC snails.

It is generally considered that most microsatellite data sets better fit the TPM than the SMM or IAM (Di Rienzo et al., 1994). Under the TPM, an expansion sign was detected in HT and under the SMM in all three

sites. The reason why snails of three sites each showed a potential population expansion was because of either the specific reproduction capability of these kinds of snails, or the Wahlund effect of ‘the aggregation of snails’ caused by the River flooding (Wilke et al., 2000). This may warrant further research.

Molecular and ecological investigations are pertinent in understanding and forecasting the likely co-evolutionary trajectory of host-parasite systems. We do acknowledge, however, that the fact that only three contrasting habitats were sampled and analyzed here, without multiple site replicates within each group, is a potential limitation of the current study, although approximately 70 snails were genotyped within each group and the effects observed were strong. Secondly, because population genetic research on *O. hupensis* has severely lagged behind on *Biomphalaria glabrata* (Ittiprasert et al., 2013) and *B. globosus* (Mkize et al., 2016), the intermediate hosts for *S. haematobium* and for *S. mansoni* respectively, and hence, in addition to the novel microsatellite loci developed here, we were restricted to only those nine microsatellite loci currently available. Finally, microsatellites may be one of the most powerful mendelian markers found, but the mutation rate varies across loci (Jarne and Lagoda, 1996). Therefore, further research with microsatellites based on the locus-specific model would be inspired. However, this work is the first study to provide in-depth information on genetic diversities of snails within each of three different habitats.

5. Conclusion

Our study revealed significant differences in *O. hupensis* population genetics between contrasting snail habitats, with the expected lower genetic diversities in snails from new or re-emerged snail habitats relative to those from the long-term established habitat. To our initial surprise, we did not see any signs of bottleneck effects in snails from the new or the re-emerged snail habitat, but we did see the difference in population substructure between the new and the re-emerged habitats. Moreover, an infinite effective population size was estimated from the new snail habitat. We conclude that at the new habitat snails were imported at large scale, whereas at the re-emerged habitat snails were largely left before and subsequently reproduced.

Currently there are few intermediate hosts with the potential for such explosive increases in range expansion, population numbers, and perhaps most importantly genetic perturbation, as gastropod molluscs (Dybdahl and Lively, 1998). In future, if public health costs associated with disease control activities, but also other forms of anthropogenic change such as water projects or climate change are to be mitigated, a thorough understanding of

the population genetics underpinning the causes, progression and consequences of range expansion and parasite compatibility in *O. hupensis* intermediate hosts is imperative. We propose that more surveys should be now undertaken in newly emerged habitats and more intense control measures should be performed in the re-merged habitats. Ultimately, until (if ever) China can actually declare interruption of *S. japonicum* transmission, monitoring of potential snail habitats should be continued, even in areas where past records have not observed their occurrence, particularly in our time of substantial anthropogenic change.

Appendix A. Primer sequences and characteristics of nine microsatellites on *Oncomelania hupensis* snails.

Acknowledgements

The work was supported by the Chinese National Research and Development Plan (YSL, grant number 2016YFC1200500) and by the Jiangsu Provincial Project of Invigorating Health Care through Science, Technology and Education (DBL).

References

- Beerli, P., Felsenstein, J., 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc Natl Acad Sci USA* 98, 4563-4568.
- Boag, D.A., 1986. Dispersal in pond snails: a potential role of waterfowl. *Canadian Journal of Zoology* 64, 904-909.
- Boissier, J., Grech-Angelini, S., Webster, B.L., Allienne, J.F., Huyse, T., Mas-Coma, S., Toulza, E., Barre-Cardi, H., Rollinson, D., Kincaid-Smith, J., Oleaga, A., Galinier, R., Foata, J., Rognon, A., Berry, A., Mouahid, G., Henneron, R., Mone, H., Noel, H., Mitta, G., 2016. Outbreak of urogenital schistosomiasis in Corsica (France): an epidemiological case study. *Lancet Infect Dis* 16, 971-979.
- Campbell, G., Noble, L.R., Rollinson, D., Southgate, V.R., Webster, J.P., Jones, C.S., 2010. Low genetic diversity in a snail intermediate host (*Biomphalaria pfeifferi* Krass, 1848) and schistosomiasis transmission in the Senegal River Basin. *Mol Ecol* 19, 241-256.
- Collins, C., Xu, J., Tang, S., 2012. Schistosomiasis control and the health system in P.R. China. *Infect Dis Poverty* 1, 8.
- Coltman, D.W., Pilkington, J.G., Smith, J.A., Pemberton, J.M., 1999. Parasite-Mediated Selection against Inbred Soay Sheep in a Free-Living Island Populaton. *Evolution* 53, 1259-1267.

- Cornuet, J.M., Luikart, G., 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144, 2001-2014.
- Dabo, A., Diarra, A.Z., Machault, V., Toure, O., Niambele, D.S., Kante, A., Ongoiba, A., Doumbo, O., 2015. Urban schistosomiasis and associated determinant factors among school children in Bamako, Mali, West Africa. *Infect Dis Poverty* 4, 4.
- Davies, C.M., Webster, J.P., Kruger, O., Munatsi, A., Ndamba, J., Woolhouse, M.E., 1999. Host-parasite population genetics: a cross-sectional comparison of *Bulinus globosus* and *Schistosoma haematobium*. *Parasitology* 119, 295-302.
- Davis, G.M., Wilke, T., Zhang, Y., Xu, X., Qiu, C.P., Spolsky, C.M., Qiu, D., Li, Y.S., Xia, M.Y., Zheng, F., 1999a. Snail-*Schistosoma*, *Paragonimus* interactions in China: population ecology, genetic diversity, coevolution and emerging diseases. *Malacol* 41, 355-377.
- Davis, G.M., Yi, Z., Hua, G.H., Spolsky, C., 1995. Population genetics and systematics of *Oncomelania hupensis* (Gastropods:Pomatiopsidae) throughout China. *Malacol* 37, 133-156.
- Davis, G.M., Zhang, Y., Xu, X.J., Yang, X.X., 1999b. Allozyme analyses test the taxonomic relevance of ribbing in Chinese *Oncomelania* (Gastropoda: Rissoacea: Pomatiopsidae). *Malacol* 41, 297-317.
- Di Rienzo, A., Peterson, A.C., Garza, J.C., Valdes, A.M., Slatkin, M., Freimer, N.B., 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proc Natl Acad Sci USA* 91, 3166-3170.
- Dybdahl, M.F., Lively, C.M., 1998. Host-parasite coevolution: evidence for rare advantage and time-lagged selection in a natural population. *Evolution* 52, 1057-1066.
- Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genet Resour* 2012, 359-361.
- Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14, 2611-2620.
- Gandon, S., Michalakis, Y., Ebert, D., 1996. Temporal variability and local adaptation. *Trends Ecol Evol* 11, 431.
- Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Updated from Goudet (1995). Available from <http://www.unil.ch/izea/software/fstat.html>.
- Gow, J.L., Noble, L.R., Rollinson, D., Mimpfundi, R., Jones, C.S., 2004. Breeding system and demography shape population genetic structure across ecological and climatic zones in the African freshwater snail, *Bulinus forskalii* (Gastropoda, Pulmonata), intermediate host for schistosomes. *Mol Ecol* 13, 3561-3573.
- Guan, W., Li, S.Z., Abe, E.M., Webster, B.L., Rollinson, D., Zhou, X.N., 2016. The genetic diversity and

- geographical separation study of *Oncomelania hupensis* populations in mainland China using microsatellite loci. *Parasit Vectors* 9, 28.
- Guo, J.T., Zhou, Y.B., Wei, J.G., 2008. Sequencing on products of *Oncomelania hupensis* through simple sequence repeat anchored polymerase chain reaction amplification. *Chin. J. Epidemiol* 29, 1119-1122.
- Hauswald, A.K., Remais, J.V., Xiao, N., Davis, G.M., Lu, D., Bale, M.J., Wilke, T., 2011. Stirred, not shaken: genetic structure of the intermediate snail host *Oncomelania hupensis robertsoni* in an historically endemic schistosomiasis area. *Parasit Vectors* 4, 206.
- He, Y.X., Salafsky, B., Ramaswamy, K., 2001. Host-parasite relationships of *Schistosoma japonicum* in mammalian hosts. *Trends in Parasitology* 17, 320-324.
- Head, J.R., Chang, H., Li, Q., Hoover, C.M., Wilke, T., Clewing, C., Carlton, E.J., Liang, S., Lu, D., Zhong, B., Remais, J.V., 2016. Genetic evidence of contemporary dispersal of the intermediate snail host of *Schistosoma japonicum*: movement of an NTD host is facilitated by land use and landscape connectivity. *PLoS Negl Trop Dis* 10, e0005151.
- Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scand J Stat* 6, 65-70.
- Hope, M., McManus, D.P., 1994. Genetic variation in geographically isolated populations and subspecies of *Oncomelania hupensis* determined by a PCR-based RFLP method. *Acta Trop* 57, 75-82.
- Ittiprasert, W., Miller, A., Su, X.Z., Mu, J., Bhusudsawang, G., Ukoskit, K., Knight, M., 2013. Identification and characterisation of functional expressed sequence tags-derived simple sequence repeat (eSSR) markers for genetic linkage mapping of *Schistosoma mansoni* juvenile resistance and susceptibility loci in *Biomphalaria glabrata*. *Int J Parasitol* 43, 669-677.
- Jarne, P., Lagoda, P.J.L., 1996. Microsatellites, from molecules to populations and back. *Trends in Evolution and Ecology* 11, 424-429.
- Jarne, P., Theron, A., 2001. Genetic structure in natural populations of flukes and snails: a practical approach and review. *Parasitology* 123 (Suppl), S27-40.
- Lei, Z.L., Zhou, X.N., 2015. Eradication of schistosomiasis: a new target and a new task for the National Schistosomiasis Control Programme in the People's Republic of China. *Chin J Schisto Control* 27, 1-4.
- Li, S.Z., Wang, Y.X., Yang, K., Liu, Q., Wang, Q., Zhang, Y., Wu, X.H., Guo, J.G., Bergquist, R., Zhou, X.N., 2009. Landscape genetics: the correlation of spatial and genetic distances of *Oncomelania hupensis*, the intermediate host snail of *Schistosoma japonicum* in mainland China. *Geospat Health* 3, 221-231.
- Lu, D.B., Rudge, J.W., Wang, T.P., Donnelly, C.A., Fang, G.R., Webster, J.P., 2010a. Transmission of

- Schistosoma japonicum* in marshland and hilly regions of China: parasite population genetic and sibship structure. PLoS Negl Trop Dis 4, e781.
- Lu, D.B., Wang, T.P., Rudge, J.W., Donnelly, C.A., Fang, G.R., Webster, J.P., 2010b. Contrasting reservoirs for *Schistosoma japonicum* between marshland and hilly regions in Anhui, China - a two-year longitudinal parasitological survey. Parasitology 137, 99-110.
- Mkize, L.S., Mukaratirwa, S., Zishiri, O.T., 2016. Population genetic structure of the freshwater snail, *Bulinus globosus*, (Gastropoda: Planorbidae) from selected habitats of KwaZulu-Natal, South Africa. Acta Trop 161, 91-99.
- Nyakaana, S., Stothard, J.R., Nalugwa, A., Webster, B.L., Lange, C.N., Jorgensen, A., Rollinson, D., Kristensen, T.K., 2013. *Bulinus globosus* (Planorbidae; Gastropoda) populations in the Lake Victoria basin and coastal Kenya show extreme nuclear genetic differentiation. Acta Trop 128, 226-233.
- Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics 28, 2537-2539.
- Peel, D., Ovenden, J.R., Peel, S.L., 2004. NeEstimator: software for estimating effective population size, Version 1.3. Queensland Government, Department of Primary Industries and Fisheries.
- Piry, S., Luikart, G., Cornuet, J.M., 1999. Bottleneck: a computer program for detecting recent reductions in the effective population size using allele frequency data. J Hered 90, 502-503.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945-959.
- Qian, Y.J., Li, S.Z., Xu, J., Yang, K., Huang, Y.X., Cao, Z.G., Miu, F., Dang, H., Zhang, L.J., Zhang, L., Wang, Q., Bergquist, R., Zhou, X.N., 2015. Potential schistosomiasis foci in China: a prospective study for schistosomiasis surveillance and response. Acta Trop 141, 342-348.
- Rudge, J.W., Webster, J.P., Lu, D.B., Wang, T.P., Fang, G.R., Basanez, M.G., 2013. Identifying host species driving transmission of schistosomiasis japonica, a multihost parasite system, in China. Proc Natl Acad Sci USA 110, 11457-11462.
- Shi, C.H., Wilke, T., Davis, G.M., Xia, M.Y., Qiu, C.P., 2002. Population genetics, microphylogeography, ecology and susceptibility to schistosome infection of Chinese *Oncomelania hupensis hupensis* (Gastropoda: Rissooidea: Pomatiopsidae) in the Miao river system. Malacol 44, 333-347.
- Shi, H.P., Lu, D.B., Shen, L., Shi, T., Gu, J., 2014. Single- or mixed-sex *Schistosoma japonicum* infections of intermediate host snails in hilly areas of Anhui, China. Parasitol Res 113, 717-721.

- Spolsky, C.M., Davis, G.M., Zhang, Y., 1996. Sequencing methodology and phylogenetic analysis: cytochrome b gene sequence reveals significant diversity in Chinese populations of *Oncomelania* (Gastropoda: Pomatiopsidae). *Malacol* 38, 213-221.
- Stensgaard, A.S., Vounatsou, P., Sengupta, M.E., Utzinger, J., 2018. Schistosomes, snails and climate change: Current trends and future expectations. *Acta Trop*.
- Wang, L.D., Guo, J.G., Wu, X.H., Chen, H.G., Wang, T.P., Zhu, S.P., Zhang, Z.H., Steinmann, P., Yang, G.J., Wang, S.P., Wu, Z.D., Wang, L.Y., Hao, Y., Bergquist, R., Utzinger, J., Zhou, X.N., 2009. China's new strategy to block *Schistosoma japonicum* transmission: experiences and impact beyond schistosomiasis. *Trop Med Int Health* 14, 1475-1483.
- Waples, R.S., Do, C., 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour* 8, 753-756.
- Webster, J.P., Davies, C.M., Hoffman, J.I., Ndamba, J., Noble, L.R., Woolhouse, M.E., 2001. Population genetics of the schistosome intermediate host *Biomphalaria pfeifferi* in the Zimbabwean highveld: implications for co-evolutionary theory. *Ann Trop Med Parasitol* 95, 203-214.
- WHO, 2015. Schistosomiasis: number of people treated worldwide in 2013. *Wkly Epidemiol Rec* 90, 25-32.
- Wilke, T., Davis, G.M., Cui, E.C., Xiao-Nung, Z., Xiao Peng, Z., Yi, Z., Spolsky, C.M., 2000. *Oncomelania hupensis* (Gastropoda: rissoidae) in eastern China: molecular phylogeny, population structure, and ecology. *Acta Tropica* 77, 215-227.
- Wilke, T., Davis, G.M., Qiu, D., Spear, R.C., 2006. Extreme mitochondrial sequence diversity in the intermediate schistosomiasis host *Oncomelania hupensis robertsoni*: another case of ancestral polymorphism? *Malacol* 48, 143-157.
- Xu, X.J., Wei, F.H., Yang, X.X., Dai, Y.H., Yu, G.Y., Chen, L.Y., Su, Z.M., 2000. Possible effects of the Three Gorges dam on the transmission of *Schistosoma japonicum* on the Jiang Han plain, China. *Ann Trop Med Parasitol* 94, 333-341.
- Zhang, L., Li, S., Wang, Q., Qian, Y., Liu, Q., Yang, P., Zhou, X., 2012. Isolation and characterization of 15 new microsatellite markers in *Oncomelania hupensis*, the snail Intermediate host of *Schistosoma japonicum* in mainland China. *Int J Mol Sci* 13, 5844-5850.
- Zhang, L.J., Xu, Z.M., Qian, Y.J., Dang, H., Lv, S., Xu, J., Li, S.Z., Zhou, X.N., 2016. Endemic status of schistosomiasis in People's Republic of China in 2015. *Chin J Schisto Control* 28, 611-617.
- Zhang, S.H., Zhao, Q.P., Jiao, R., Gao, Q., Nie, P., 2010. Identification of polymorphic microsatellites for the

- intermediate host *Oncomelania hupensis* of *Schistosoma japonicum* in China. *Malacol* 53, 147-153.
- Zhao, Q.P., Jiang, M.S., Littlewood, D.T., Nie, P., 2010. Distinct genetic diversity of *Oncomelania hupensis*, intermediate host of *Schistosoma japonicum* in mainland China as revealed by ITS sequences. *PLoS Negl Trop Dis* 4, e611.
- Zhao, Q.P., Xiong, T., Xu, X.J., Jiang, M.S., Dong, H.F., 2015. *De Novo* transcriptome analysis of *Oncomelania hupensis* after molluscicide treatment by next-generation sequencing: implications for biology and future snail interventions. *PLoS One* 10, e0118673.
- Zhou, X., Dandan, L., Huiming, Y., Honggen, C., Leping, S., Guojing, Y., Qingbiao, H., Brown, L., Malone, J.B., 2002. Use of landsat TM satellite surveillance data to measure the impact of the 1998 flood on snail intermediate host dispersal in the lower Yangtze River Basin. *Acta Trop* 82, 199-205.
- Zhou, X.N., 2005. Science on *Oncomelania* Snail (in Chinese). Science Press, Beijing, China.
- Zhou, X.N., Yang, G.J., Yang, K., Wang, X.H., Hong, Q.B., Sun, L.P., Malone, J.B., Kristensen, T.K., Bergquist, N.R., Utzinger, J., 2008. Potential impact of climate change on schistosomiasis transmission in China. *Am J Trop Med Hyg* 78, 188-194.
- Zhou, Y.B., Yang, M.X., Zhao, G.M., Wei, J.G., Jiang, Q.W., 2007a. *Oncomelania hupensis* (Gastropoda: Rissooidea), intermediate host of *Schistosoma japonicum* in China: genetics and molecular phylogeny based on amplified fragment length polymorphisms. *Malacol* 49, 367-382.
- Zhou, Y.B., Zhao, G.M., Wei, J.G., Jiang, Q.W., 2007b. Genetic diversity in 19 Chinese populations of *Oncomelania hupensis* (Gastropoda: *Rissoidea*) detected by simple sequence repeat-anchored polymerase chain reaction amplification. *Chin J Epidemiol* 28, 859-862.

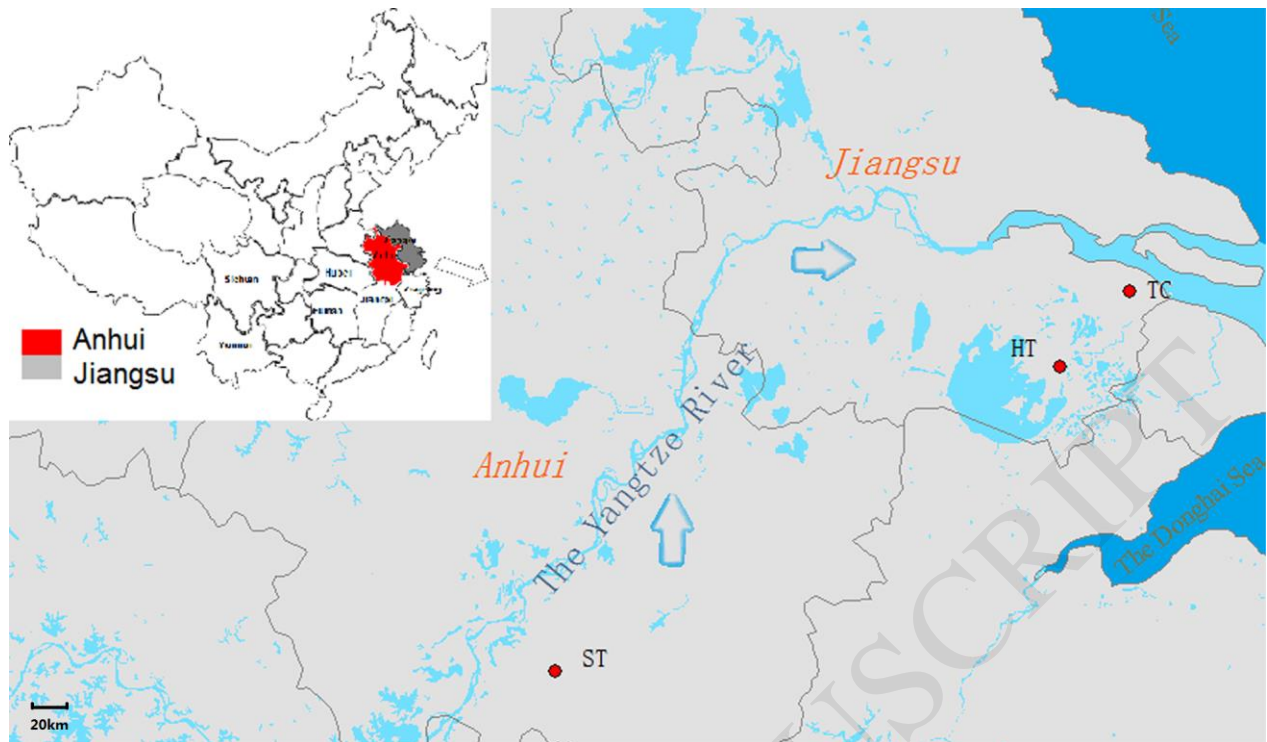


Fig. 1. Map of geographical locations of three snail populations sampled. ST for Shitai of Anhui province, HT for Hengtang of Jiangsu province and TC for Taicang of Jiangsu province. Arrow refers to the direction of the River flow. Map was created using Eric ArcGIS Pro (trial) and data sources were from <http://bbs.pinggu.org/thread-3582623-1-1.html>.

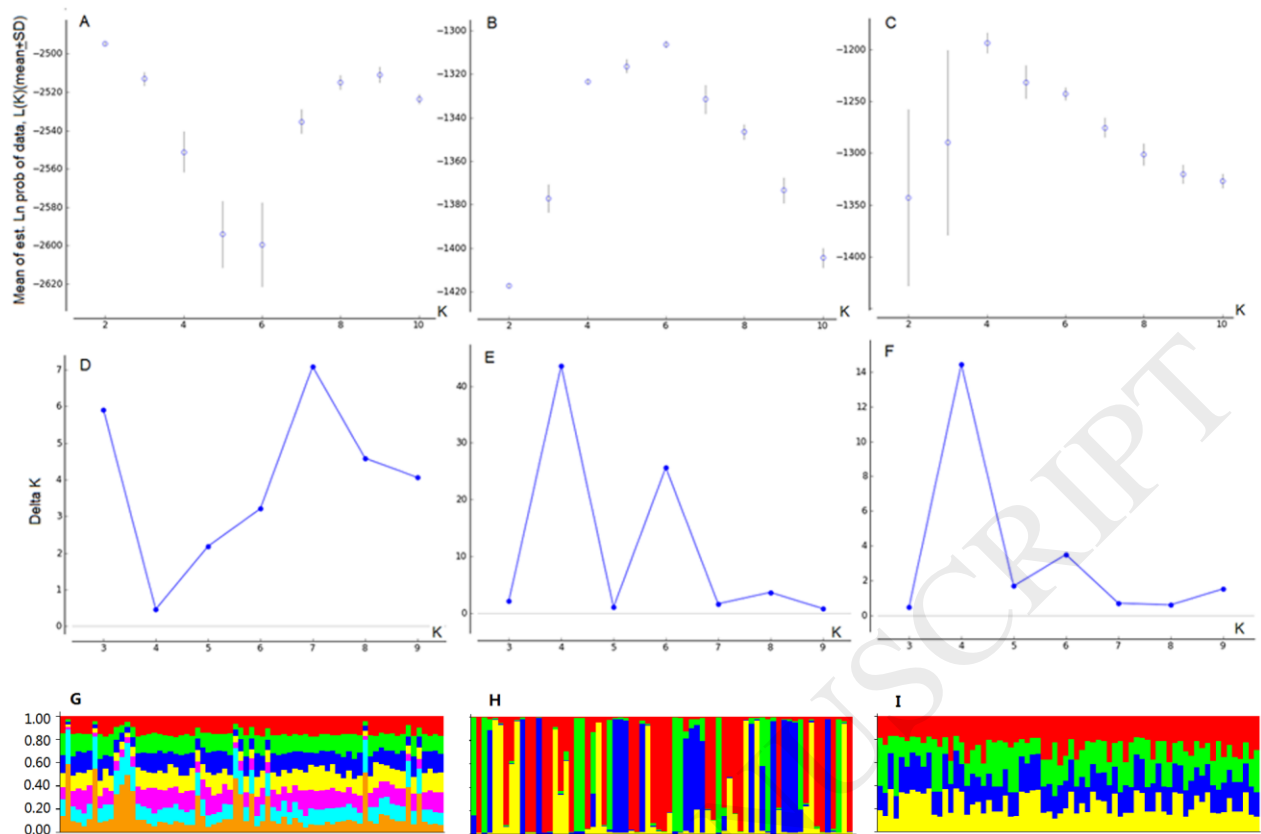


Fig. 2. Estimated number of clusters for *O. h. hupensis* using STRUCTURE software. A to C: Mean (\pm SD) natural logarithm of the likelihood of the data [$L_nP(X | K)$] over ten replicated runs for each value of assumed clusters (K) in A for ST, B for HT and C for TC. D to F: delta K values for ST, HT and TC respectively, with inferred $K=7$ in D for ST, $K=4$ in E for HT and $K=4$ in F for TC. G to I: the proportional membership of each individual in the inferred K genetic clusters (shown in different colors), with G for ST, H for HT and I for TC.

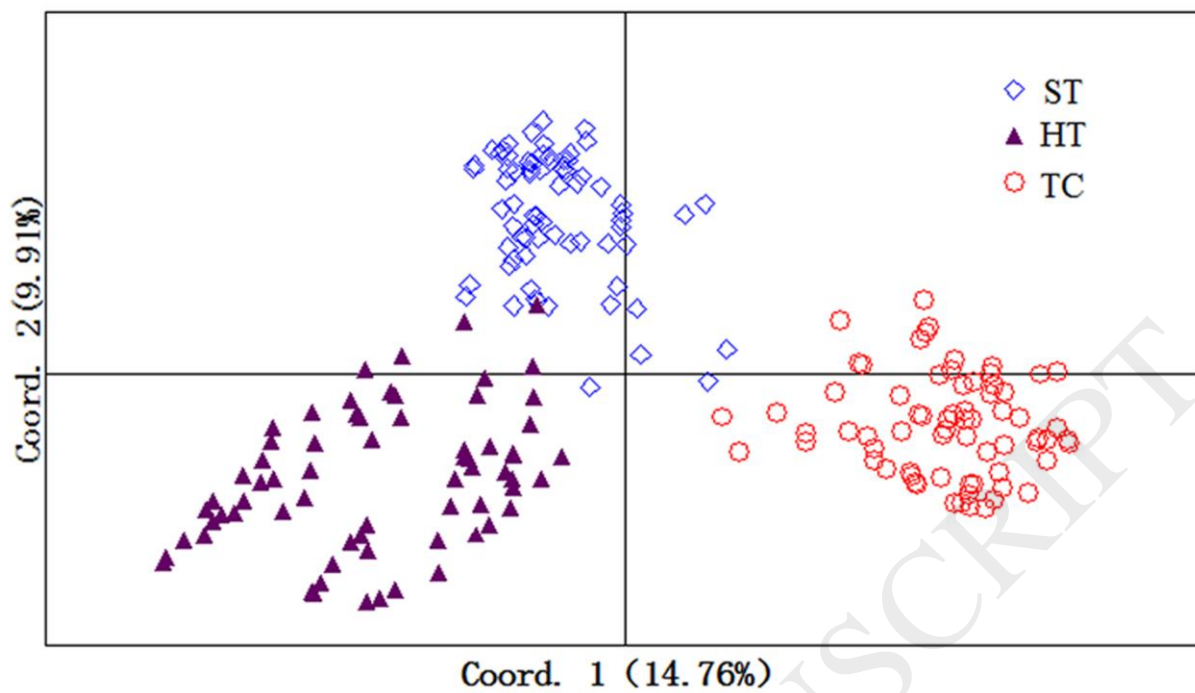


Fig. 3. PCoA illustrates the relatedness of snail individuals among sites.

TablesTable 1. Information on sampling sites and *O. hupensis* snails

Site (code)	Habitat	Coordinate	Area (m ²)	No. of frames investigated (0.11 m ² /frame)	No. of frames with snails	No. of total snails collected	Snail density (No. snails per frame)
Shitai, Chizhou, Anhui (ST)	Snail persistence	117.357767E 30.198765W	1200	30	25	245	8.17
Hengtang, Suzhou, Jiangsu (HT)	Snails re-emerged	120.576252E 31.249472W	1510	54	23	243	4.50
Taicang, Suzhou, Jiangsu (TC)	Snails first found in history	121.073147E 31.563253W	1324	32	26	220	6.88

Table 2. Average genetic diversity (SE) of *O. hupensis* analyzed at nine microsatellite loci with Genealex

		No. of snails					
Site	genotyped	Na	NeA	Ho	He	uHe	F _{IS}
		14.333	7.028	0.502	0.816	0.823	0.394
ST	71	(1.965)	(1.542)	(0.084)	(0.030)	(0.030)	(0.096)
		9.444	3.392	0.314	0.688	0.696	0.551
HT	70	(0.818)	(0.276)	(0.063)	(0.026)	(0.026)	(0.085)
		5.444	2.553	0.345	0.498	0.503	0.443
TC	70	(1.029)	(0.373)	(0.115)	(0.099)	(0.099)	(0.153)

Abbreviations: Na, no. of different alleles per locus; NeA, no. of effective alleles per locus; Ho, observed heterozygosity; He, expected heterozygosity; uHe, unbiased expected heterozygosity; F_{IS}, the inbreeding coefficient, bold for significance.

Table 3. Test for heterozygosity excess or deficiency with the Wilcoxon sign-rank test in the software Bottleneck

Snail population	Probability for heterozygosity excess			Probability for heterozygosity deficiency		
	IAM	TPM	SMM	IAM	TPM	SMM
ST	0.00488	0.58984	0.99902	0.99707	0.45508	0.00195
HT	0.54492	0.99316	0.99902	0.50000	0.00977	0.00195
TC	0.42188	0.80859	0.98633	0.62891	0.23047	0.01953

Table 4. Effective population size (N_e) and its 95% confidence intervals (CI) of *O. hupensis* snails with two software based on Linkage disequilibrium

Snail populations	with NeEst	with LDNe
ST	179.9 (91.6, 1070.9)	284.9 (113.5, inf)
HT	17.2 (12.7, 23.4)	46.1 (27.4, 101.0)
TC	inf (173.5, inf)	-78.0 (-316.1, inf)

Abbreviations: inf, infinite. The lowest allele frequency used was 0.05.

ACCEPTED MANUSCRIPT

Table 5. Immigration rates (M) into each of the three sites from every other site as estimated by MIGRATE

Site, i	Theta	M(ST→i)	M(HT→i)	M(TC→i)	M (Total→i)
ST	2.23		1.19	3.58	4.77
HT	0.78	2.33		5.07	7.40
TC	2.17	2.44	1.80		4.24
Total		4.77	3.00	8.65	

ACCEPTED MANUSCRIPT