Title
Population genetics of the malaria vector Anopheles aconitus in China and Southeast Asia.

Journal
Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 12(8)

ISSN
1567-1348

Authors
Chen, Bin
Harbach, Ralph E
Walton, Catherine
et al.

Publication Date
2012-12-01

DOI
10.1016/j.meegid.2012.08.007

Published in final edited form as:
Infect Genet Evol. 2012 December ; 12(8): 1958–1967. doi:10.1016/j.meegid.2012.08.007.

Population genetics of the malaria vector *Anopheles aconitus* in China and Southeast Asia

Bin Chen[a,*], **Ralph E. Harbach**[b], **Catherine Walton**[c], **Zhengbo He**[a], **Daibin Zhong**[d], **Guiyun Yan**[d], and **Roger K. Butlin**[e]

[a]Institute of Entomology and Molecular Biology, College of Life Sciences, Chongqing Normal University, Chongqing 401331, China

[b]Department of Entomology, The Natural History Museum, London SW7 5BD, UK

[c]Faculty of Life Sciences, The University of Manchester, Manchester M13 9PT, UK

[d]College of Health Sciences, University of California at Irvine, CA 92697, USA

[e]Animal and Plant Sciences, The University of Sheffield, Sheffield S10 2TN, UK

Abstract

Anopheles aconitus is a well-known vector of malaria and is broadly distributed in the Oriental Region, yet there is no information on its population genetic characteristics. In this study, the genetic differentiation among populations was examined using 140 mtDNA *COII* sequences from 21 sites throughout southern China, Myanmar, Vietnam, Thailand, Laos and Sri Lanka. The population in Sri Lanka has characteristic rDNA D3 and ITS2, mtDNA COII and ND5 haplotypes, and may be considered a distinct subspecies. Clear genetic structure was observed with highly significant genetic variation present among population groups in Southeast Asia. The greatest genetic diversity exists in Yunnan and Myanmar population groups. All population groups are significantly different from one another in pairwise *F*st values, except northern Thailand with central Thailand. Mismatch distributions and extremely significant F_S values suggest that the populations passed through a recent demographic expansion. These patterns are discussed in relation to the likely biogeographic history of the region and compared to other *Anopheles* species.

Keywords

Anopheles aconitus; subspecies; population structure; demographic expansion; mtDNA; rDNA; China; Southeast Asia

1. Introduction

Anopheles aconitus was originally described from Sumatra, Indonesia (Dönitz, 1902) and is thought to be broadly distributed from Sri Lanka, India and Nepal eastward to Hainan Island

[*]Correspondence author at: Institute of Entomology and Molecular Biology, College of Life Sciences, Chongqing Normal University, Chongqing 401331, China. Tel: +862365918391; fax: +862365340559. c_bin@hotmail.com (B. Chen).

5. Conflict of interest

None.

in China, and south from southern China through Southeast Asia into Indonesia (Harrison, 1980; Chen et al., 2002). It has been incriminated as a vector of malarial parasites in Thailand (Gould et al., 1965) and is considered to be the primary vector in Malaysia and Indonesia (Damar et al., 1981; Rahman et al. 1993; Barcus et al. 2002). Besides malarial parasites, filariae of *Wuchereria bancrofti* have also been found in this species in Flores, Indonesia (Atmosoedjono and Dennis, 1977). Adults of *An. aconitus*, *An. varuna*, *An. minimus* and *An. harrisoni* are morphologically very similar to one another, with overlapping ranges of character variation, often resulting in their misidentification (Harrison, 1980; Harbach et al., 2007). Various molecular methods have been developed, as summarized by Van Bortel et al. (2000), to facilitate identification of these and other species of the Myzomyia Series of *Anopheles* subgenus *Cellia*, including single-strand conformation polymorphisms (Sharpe et al., 1999), restriction patterns of the ITS2 fragment of rDNA (Van Bortel et al., 2000) and random amplified polymorphic DNA markers (Kengne et al., 2001). Subsequently, Phuc et al. (2003) and Carros et al. (2004) each developed a multiplex PCR assay separately to identify the malaria vector *An. minimus* and four related species from Southeast Asia, and the major malaria vectors within the African *An. funestus* and the Oriental *An. minimus* groups, both using ITS2 locus. Morgan et al. (2011) investigated the impact of pleistocene environmental change in shaping genetic diversity across the Indo-Burma region, by combining *COII* data of *An. aconitus* and other 8 *Anopheles* species. However, they did not deliberately analyze the *COII* variation and population genetic characteristics of *An. aconitus*, nor did they sequence other DNA markers to investigate the diversity within the species. Phylogenetic analysis based on nuclear and mitochondrial DNA sequences shows that *An. aconitus*, *An. filipinae*, *An. mangyanus*, *An. pampanai* and *An. varuna* form a monophyletic group distinct from other members of the Funestus Group (previously Minimus Group, see Harbach, 2004) of the Myzomyia Series, and the name Aconitus Subgroup was applied to these five closely related species (Chen et al., 2003, Garros et al., 2005).

Currently, chemical measures remain central to the control of malarial parasites and their vectors despite the problems of parasite resistance to antimalarial drugs (Meshnick, 1997) and the spread of insecticide-resistance in vector populations (Collins and Paskewitz, 1995; Ranson et al. 2011). The insecticide resistance in *Anopheles* species has been widely spread in China and Mekong Region (Cui et al., 2006; Van Bortel, et al. 2008). The Mekong Region is still one of the most threatening foci of malaria (Cui et al. 2011), and the absence of knockdown resistance suggests metabolic resistance in the main malaria vectors of the region (Verhaeghen et al. 2009). Knowledge of vector population genetics may allow a more rational management of insecticides to defer the development of resistance and could help explain the role of vectors in relation to geographical variation in malaria transmission. However, the population genetic characteristics of *An. aconitus* are virtually unknown despite its very wide distribution and epidemiological importance. More importantly, different subspecies might be substantially different in their biology, ecology, behaviour and malaria vector status, and hence malaria control. Despite the wide distribution of this species, whether there is any subdivision at the subspecies level has never been investigated.

Because of its rapid evolution and maternal, non-recombining mode of inheritance, mitochondrial DNA (mtDNA) has been widely used to investigate population differentiation and the evolutionary history of populations and subspecies (Avise et al., 1988; Avise, 1989; Cronin, 1992; Arctander et al., 1996). In particular, the mitochondrial cytochrome oxidase II gene (*COII*) has been successfully used to study population genetic structure and population history for a wide range of insect species (Krafsur et al., 1999; Sperling et al., 1999; Clements et al., 2000; Franck et al., 2001; Rua et al., 2001; Chen et al., 2004; Chen et al., 2011). In the present study, we sampled populations of *An. aconitus* in China, Myanmar, Vietnam, Thailand, Laos and Sri Lanka, examined the pattern of variation within and among

populations using a large sample of *COII* sequences, and checked whether there is any subdivision of the morphologically defined species using two mtDNA genes (*COII* and *ND5*) and three rDNA sequences (D2, D3 and ITS2 regions) from a smaller set of samples.

2. Material and methods

2.1. Specimens

Mosquitoes were widely collected throughout southern China, Myanmar, Vietnam, Thailand, Laos and Sri Lanka in 1996–2003. Most localities visited were areas with relatively high malaria incidence. Adults resting near or landing on humans and tethered bovines were captured with aspirators between 2000 and 2300 h. Mosquitoes were killed by exposing them to ethyl acetate vapour, placed individually in Beem capsules (Agar Scientific Ltd, Stansted, UK) and stored in plastic bags with silica gel. Larvae and pupae were collected in rice fields and irrigation ditches at some sites and reared to adults. The keys of Harrison (1980) were used to identify adults and larvae to *An. aconitus* and related species.

2.2. Molecular identification and DNA sequencing

DNA was extracted from mosquitoes using a phenol-chloroform method (Chen et al., 2002). The single-strand conformation polymorphism (SSCP) method based on the 28S D3 region, developed for the identification of *An. minimus*, *An. harrisoni*, *An. aconitus*, *An. varuna*, *An. pampanai*, *An. fluviatilis*, *An. jeyporiensis* and *An. culicifacies* (Sharpe et al., 1999; Chen et al., 2002), was used to exclude samples of species indistinguishable from *An. aconitus* by morphology. The mtDNA *COII* gene of those individuals that were identified as *An. aconitus* by SSCP was sequenced for the study of population structure. A smaller set of these individuals, selected on the basis that they have different *COII* haplotypes and come from different geographical regions, was also sequenced for the rDNA D2, D3 and ITS2 regions and the mtDNA *ND5* gene to confirm the SSCP identification using the D3 locus sequence, and to check the monophyly and unity of the species. *Anopheles filipinae* and *An. pampanai* were used as outgroup taxa to analyse the variation of these loci within *An. aconitus*.

The primers used for PCR amplification and sequencing were taken from Sharpe et al. (2000) for *COII*, D3 and ITS2 and Krzywinski et al. (2001) for *ND5* and D2. Amplifications were performed in 50 μl volume overlaid with two drops of mineral oil on a HYBAID OmniGene cycler (Thermo Hybaid, Ashford, England). Each PCR included 1/100 DNA of a whole mosquito, 5 μl 10× ReddyMix™ buffer (ABgene, Epsom, England), 200 μM dNTPs, 2 mM $MgCl_2$, 600 nM of each primer and 1.3 units of Thermoprime Plus DNA Polymerase (ABgene). Reactions started with denaturation at 95ºC for 5 min, followed by 35 cycles, each cycle consisting of denaturation for 40 s at 95ºC, annealing for 40 s at 55ºC and extension for 1 min at 72ºC, with a final extension at 72ºC for 6 min. PCR products were electrophoresed through ethidium bromide-stained 1% agarose gels in 1× TBE and visualised under UV light to check for successful amplification. The PCR products were then purified using a spin column (Promega Wizard PCR Preps, Madison, Wisconsin, U.S.A.) and sequenced in both directions in an ABI 377 automated sequencer (PE Applied Biosystems, Warrington, England).

2.3. Analysis of mtDNA variation

The DNA sequences were edited manually and aligned using Clustal X (Thompson et al., 1997). The COII and ND5 sequences were translated into amino acids using the Drosophila mtDNA genetic code in Transeq (European Bioinformatics Institute: http://www.ebi.ac.uk/). Genetic diversity estimates of the *COII* locus, θ_π, based on the average pairwise number of

differences between sequences (Tajima, 1983) and θ_s, based on the number of segregating nucleotide sites per sequence (Watterson, 1975), were also estimated and then converted to values per base, at the within-population level and for geographically defined groups of populations (Fig. 1) using Arlequin, ver 3.5 (Excoffier et al., 2010). Genealogical relationships among haplotypes were constructed using TCS (Clement et al., 2010) with the method described by Templeton et al. (1992). Analysis of population genetic structure was carried out using analysis of molecular variance (AMOVA; Excoffier et al., 2010) also in Arlequin, ver 3.5. AMOVA estimates the proportion of total genetic variation attributable to different hierarchical levels based on the geographical distribution of haplotypes, taking into account the number of molecular differences through squared distances between haplotypes. Correlations among haplotypes are measured in terms of Φ-statistics in a manner similar to *F*-statistics. The significance of the Φ-statistics was tested by permutation on the original inter-individual squared distance matrix.

Pairwise *F*st values used as genetic distances between recently diverged population groups were estimated with the methods of Slatkin (1995) and tested for significance by permutation. Genetic isolation by geographic distance was assessed by a Mantel test. The frequency distributions of the numbers of segregating sites in all possible pairwise comparisons, known as mismatch distributions, were calculated using Arlequin. We also calculated Tajima's D (Tajima, 1989) and performed Fu's Fs test (Fu, 1997) using Arlequin for test of neutrality. LAMARC program (Kuhner 2006) was used to compute effective population size and population growth rate by using the likelihood method. The test involves comparisons of different estimates of the parameter θ, which is defined as $2N\mu$ for a haploid locus, where N is the effective population size and μ is the mutation rate per sequence per generation (Fu 1997).

Maximum likelihood (ML) phylogenetic analyses were conducted using the version of PAUP* (version 4.0b8) (Swofford, 2001). Each was performed using the heuristic search option employing step-wise addition with 100 random taxon addition sequence replicates and five trees held at each step. The model of DNA substitution for ML was determined using the program Modeltest (version 3.06) (Posada & Crandall, 1998). Modeltest uses hierarchical likelihood ratio tests to determine which of the General Time Reversible (GTR) family of substitution models (sixty-four in all) best fits the data. Node support for MP analysis was assessed using 1,000 bootstrap pseudo-replicates.

3. Results and analysis

3.1. COII sequence variation

Anopheles aconitus is difficult to collect because it resides in hilly or mountainous areas and earlier distribution reports are an unreliable guide due to the confusion in morphological identification. Thirty-four sites were visited at least once but only 79 individuals of *An. aconitus* from 17 sites were successfully collected, and these were sequenced for the entire *COII* gene with a length of 685 bp. Further 61 *COII* sequences with a length of 639 bp for *An. aconitus* were taken from NCBI, reported earlier in Morgan et al. (2011). Totally, 140 sequences with a length of 639 bp from 21 sites were used for the population genetic analysis (Table 1; Fig. 1).

The set of sequences had 46 haplotypes coding for 212 amino acids, and were very AT-rich (75.46% A and T bases). Forty-eight nucleotide substitutions were identified at 47 of the 639 sites, of which 44 were transitions (91.67%) and 4 were transversions (8.33%) (Table 2). Seven substitutions resulted in amino acid changes (direction of change based on outgroup comparison; Chen et al., 2003): those at site 16 (A→T) for haplotype 24, at site 44 (S→F) for haplotype 44, at site 395 (N→S) for haplotypes 9 and 27, at site 430 (V→I) for

haplotypes 6, 11, 15, 16 and 37, at site 463 (V→I) for haplotype 33, at site 496 (V→I) for haplotype 35, and at site 576 (L→F) for haplotypes 46. Twelve amino acid sequences resulted from these amino acid changes do not correspond completely to geographical groups. Sequences have been deposited in GenBank/EMBL (accession numbers in Table 2; AJ512745-AJ512746 for *An. filipinae* and *An. pampanai*).

Due to the relatively small sample size for each population, we pooled samples from the 21 collecting sites into 9 population groups based on geographical distribution, but the group N Vietnam with only 4 samples and relatively far geographical separation was not analyzed as a group. Table 1 gives the haplotype distribution in these population groups. Haplotype diversity was relatively low (compared with other mosquito species, see Discussion), with 46 haplotypes found among 140 individuals. This is in accord with the generally low levels of nucleotide diversity estimated by θ_π and θ_s. The population groups from Yunnan (θ_π = $5.80\pm3.58\times10^{-3}$ and $\theta_s = 6.96\pm3.24\times10^{-3}$ per base; Table 1) and from Myanmar (θ_π = $4.13\pm2.54\times10^{-3}$ and $\theta_s = 8.70\pm3.30\times10^{-3}$) are most diverse followed by the northern and central Thailand groups. The overall estimates of molecular diversity are θ_π = $4.82\pm2.79\times10^{-3}$ and $\theta_s = 12.89\pm3.48\times10^{-3}$, excluding the distinct Sri Lanka sample (see below).

3.2. Genealogical relationships among haplotypes

Nine samples with different *COII* haplotypes, selected to represent different population groups, had identical sequences for the D2 locus (502 bp, accession no. AJ626942; AJ626943-AJ626944 for outgroup *An. filipinae* and *An. pampanai*). Four haplotypes, from these 9 samples, varied at two sites for the D3 region (335 bp, AJ626936-AJ626939; AJ626940-AJ626941 for *An. filipinae* and *An. pampanai*) and 5 haplotypes varied at three sites for the ITS2 region (478 bp, AJ626945-AJ626947; AJ626948-AJ626949 for *An. filipinae* and *An. pampanai*). There are 7 different haplotypes in the 9 samples for the ND5 gene (570 bp, AJ626927-AJ626933; AJ626934-AJ626935 for *An. filipinae* and *An. pampanai*).

Modeltest selected the general time reversible model GTR+I model as the best fit for ML phylogenetic analysis for each of these 5 data. Maximum-likelihood trees of sequences from each locus were constructed under the GTR+I model using *An. filipinae* and *An. pampanai* as outgroup taxa and the Ln likelihood are all quite large (Fig 2). Samples from Sri Lanka form a distinct clade within the species based on *COII* and *ND5* sequence, and have ITS2 and D3 sequences that are not shared with Southeast Asian specimens. Other haplotypes do not show clear geographical differentiation. Figure 3 shows the genealogical network constructed from 46 haplotypes of the *COII* gene. The number of mutation differences (unit branches) on the network represents the maximum number of mutational connections between pairs of haplotypes justified by the 95% 'parsimony' criterion (Posada & Crandall, 2001). These haplotypes are linked via some missing intermediates that could represent sampling gaps or extinct intermediates. However, because the number of samples greatly exceeds the number of haplotypes, reflecting the low level of genetic variation, these haplotypes should adequately represent *COII* sequence variation of the species throughout the sampled distribution. Despite the low genetic diversity, a high level of homoplasy is apparent in the network, with some mutational events resulting in closed loops, as observed in other mosquito mtDNA studies (Walton et al., 2000). The network is centred on haplotypes 1, which has the highest outgroup weight (0.09), implying that it may be ancestral. Haplotype 1 is shared by 44 individuals and thus has the highest frequency in the sample, followed by haplotype 2 (20 individuals) and haplotype 3 (10 individuals). Importantly, all three haplotypes from Sri Lanka form a separate branch, well separated from other parts (but haplotype 37) of the network. This is consistent with the distinct D3, ITS2 and *ND5* sequences from the island. In addition, the haplotypes of Hainan and

southern Thailand were also on separate branches, but these two branches are closely linked to other haplotypes and the sequences of D3, ITS2 and *ND5* from these two places do not represent separate lineages.

3.3. Genetic structure and genetic distances

In the hierarchical AMOVA, pairwise comparisons were made between 7 population groups (excluding Sri Lanka) and 20 populations (sites). Geographic structure was strong at the large geographical scale of population groups with $P<0.001$ (Table 3). This is consistent with the strong geographical grouping of haplotypes seen in the network (Fig. 3).

Table 4 shows that percent pairwise sequence divergences within groups range from 0.069–0.580%. The each group is significantly different from any other by pairwise *F*st value, but the N Thailand from C Thailand. Sri Lanka is the most divergent with an *F*st value greater than 0.6 from any of the other groups. It is difficult to define the geographical distance between populations separated by sea. Therefore, a Mantel test for isolation by distance was carried out only for the 8 sites within mainland Southeast Asia with at least 7 samples, ML, MY, MD, CM, LO, VT, PK and NS. The correlation coefficient for the percentage of pairwise sequence divergence between populations with geographic distance was 0.452, but this is not significant based on 1000 permutations ($P=0.061$).

3.4 Pattern of segregating sites and tests of neutrality

The simulated mismatch distribution among haplotypes is closely roughly unimodal (Fig. 4), which is expected in population that have passed through a recent demographic expansion (Rogers and Harpending 1992). The unimodal distribution can also be generated by other processes, notably recombination and high variance in substitution rates. However, the recombination is not expected for these mtDNA sequences, and the possibility caused by high sequence variation is rather low, as the substitutions happen in 48 of 639 sites for the samples investigated. Moreover, the observed mismatch distribution for demographic equilibrium is reliably ragged and often multimodal as it reflects the highly stochastic shape of gene tree (Rogers and Harpending 1992; Harpending *et al.* 1998). Therefore, we suggest that the unimodal distribution pattern reflects a growing population expansion, and an underlying star-burst-like genealogy in which all of the coalescent events occurred in a narrow time window. No simple statistical test can be made for a fit of the distribution to a Poisson distribution, as the pairwise comparisons are not independent, and the alleles share a common history. Mean distance between haplotypes of the observed mismatch distribution is 3.077 and the variance is 3.934, without much difference with the mean. Despite this, there is a rightward deviation in which the observed distribution differs from a Poisson distribution. There is a deficit of observed values around the mode, a lowness of observed value (compared to the expected) at the lower value and an excess at the upper value (Fig. 4). Neutrality tests showed that Tajima's *D* and Fu' *F*s values are both negative for pooled samples including Sri Lanka (−1.92 and −26.13, respectively) and excluding Sri Lanka (−1.84 and −26.36), which indicates an excess of low-frequency mutations. The extremely significant F_s values ($P<0.01$) in the study most likely indicate a history of population growth or hitchhiking rather than background selection (Fu 1997). The extremely significant Tajima's *D* values ($P<0.01$) mean that deviation is more likely to come from selection effects than population expansion, bottleneck or heterogeneity of mutation rates (Tajima 1996). The F_s statistic is very sensitive to geographic expansion of populations, which may explain why the F_s values are largely negative and extremely significant ($P<0.01$) also for all population groups investigated.

3.5 Effective population size and exponential growth rate

Since evidence earlier in the study indicated population growth, we used the LAMARC program to make simultaneous, phylogenetic estimates of the present-day value of θ, θ_0, and the population growth rate, *g*, assuming an exponential modal of growth using a maximum likelihood approach (Kuhner 2006). The result showed that the effective population size, N_0 (4.95×10^6 female), is quite large, and indicates a much wider distribution. This agrees with the much wider distribution reported in the literature than represented by the populations studied here. The population growth rate, *g* (1488.83 ± 75.23), indicates rapid, exponential population growth.

4. Discussion

In this study, we obtained 140 mtDNA *COII* sequences of the mosquito species *An. aconitus* from 21 sites throughout southern China, Myanmar, Vietnam, Thailand, Laos and Sri Lanka (Fig.1). The species has been recorded from 5 provinces in southern China, Yunnan, Guizhou, Guangxi, Hainan and Zhejiang (Lu, 1997), and possibly in Guangdong Province (Baba, 1950). This study confirms its presence in southern Yunnan and Hainan Provinces but no specimens were collected from Guizhou, Guangxi and Guangdong despite visiting many sites there. The species is difficult to distinguish from the *An. minimus* complex on the basis of morphology. Therefore, it is likely that its reported occurrence in these more northern provinces is due to mistaken identification.

The samples investigated share a single D2 sequence. Despite some intraspecific variation, the sequences from each of the other loci (D3, ITS2, ND5 and COII) are monophyletic with respect to other members of the Myzomyia Series of *Anopheles* subgenus *Cellia* (Chen et al., 2003, and unpublished). This supports the current concept of *An. aconitus* in the area investigated. However, the samples from Sri Lanka have differentiated D3, ITS2, *ND5* and *COII* sequences that each forms a separate clade within *An. aconitus* (Fig. 2). In the *An. minimus* complex, the D3 differences between *An. minimus*, *An. harrisoni* and *An. yaeyamaensis*, *An. leesoni* and *An. fluviatilis* range between 3 and 5 bp, whereas the difference between Sri Lankan and Southeast Asian *An. aconitus* reaches 2 bp. Moreover, Sri Lanka population group is the most divergent with an *F*st value greater than 0.6 from any of the other groups. These observations suggest that the genetically distinct and geographically separated Sri Lankan populations could be recognised as a subspecies of *An. aconitus*, but there is a need in sample collecting to cover the distribution area of *An. aconitus* in further work. Significant isolation by distance was not observed in Southeast Asia based on 1000 permutations (*P*=0.061), so this pattern may be confirmed by wider sampling. The populations on the island of Hainan and in southern Thailand also show geographically distinct groups of haplotypes for the *COII* gene (Fig. 3), but the differences are smaller and their sequences for D3, ITS2 and *ND5* are not distinct.

The apparently ancestral haplotypes 1 (Fig. 3) has high frequency and occur in several populations. However, they are concentrated in northern and central Thailand, Myanmar and Yunnan, areas that also show higher diversity than any other area. This pattern suggests that it might have been a refuge for the species during the Pleistocene glaciations from which other areas have been colonised.

Interestingly, the Yunnan region has extraordinarily high haplotype diversity with θ_π 5.80±3.58 and θ_s 6.95±3.24. The samples of the group were collected from Xishuangbanna. This area is in a transitional position from the southern Himalayas to eastern Asia and from tropical Southeast Asia to subtropical China as well as at the junction of the Indian and Burmese plates of Gondwanaland and the Eurasian plate of Laurasia (Audley-Charles, 1987). The Hengduan mountains to the north of the region act as a huge barrier keeping out

the cold air from the north in winter, and in valleys and on lower hills below 900 m altitude there are pockets of tropical rain forest, as intermittent tracts or patches in local habitats with a tropical moist climate prevailing due to particular topography (Zhu, 1997). Due to poor access and remoteness from developed and highly populated regions, this area is seldom disturbed by human activity. All these characteristics make the area one of the richest for biodiversity in China, even higher than tropical rain forest of Southeast Asia (Jin and Ou, 1997). For example, 152 and 434 plant species were found to be condensed in 0.0025 and 0.04 km^2, respectively (Zhu et al., 2004; Jin, 1997), and the arbuscular mycorrhizal fungi spore density can reach 2550 per 100 g soil (Zhao et al., 2003). Previous studies on the biodiversity of these unique tropical rain forests have mainly focused on ecosystem and species biodiversity; however, the present research provides a case of high biodiversity within species. The high haplotype diversity in Yunnan, with no common haplotype, suggests long-term persistence of a locally structured population. Similarly, Sagaing region in Myanmar has also extraordinarily high haplotype diversity with θ_π 5.26±3.24 and θ_s 8.07±3.55, which was identified as putative refugial region as Yunnan (Morgan et al, 2011).

Hierarchical analysis of molecular variance shows significant variation among population groups, which makes up 36.09% of the total. In this respect, *An. aconitus* also differs from other Southeast Asian *Anopheles* species, which show little or no population structure (*An. dirus* and *An. baimaii* in Bangladesh, Myanmar and northern, central and southern Thailand, Walton et al., 2000; *An. maculatus* in northern, central and southern Thailand, Rongnoparut et al., 1999; *An. jeyporiensis*, Chen et al., 2004).

Although the observed mismatched distribution among haplotypes deviates little from a Poisson fit (Fig. 4), the closely unimodal distributions are consistent with some degree of population expansion, as simulations have shown that stable populations almost never produce this type of profile (Rogers and Harpending 1992; Harpending et al. 1998). The population expansion can be visualised in the star-like haplotype arrays (Fig 3). The extremely significant F_s value (P<0.01) also support the inference of population expansion (Table 1). Moreover, the estimate of exponential growth rate, g (1488.83 ± 75.23), in LAMARC indicates a large population expansion as well (Table 5). This result is consistent with that of Morgan et al (2011), which suggests similarly timed population expansions in *An. aconitus*, *An. jeyporiensis*, *An. baimaii* and *An. maculatus*.

Given an exponential population growth, the coalescence time can be expressed in terms of the number of pairwise differences between samples if a constant mutation rate is assumed (Slatkin and Hudson 1991). The pairwise difference between samples per site are estimated to be 4.82×10^{-3} in ARLEQUIN, assuming 10 generations per year and a mutation rate of 10^{-8} per site per generation (Powell *et al.* 1986). From this, it can be inferred that population expansion started approximately 482,000 years ago. The present effective population size, N_0, is estimated to be 2.47×10^7 females (Table 5).

The immature stages of this species mainly live in rice fields, and therefore Harrison (1980) suggests that the widespread distribution and general abundance of this species were probably due primarily to the spread of the human-rice monoculture system in the Orient. However, the earliest and most convincing archaeological evidence, discovered at Non Nok Tha in northern Thailand, can only trace back to 4000 B.C. for domestication of rice in Southeast Asia (http://www.riceweb.org/history.htm). Originally the larvae and pupae of the species probably occurred primarily in grassy marshes with slow clear running water and along open streams and rivers with grassy margins like *An. minimus*. In recent history, man has greatly expanded these habitats by cutting forests and exposing streams, by digging ditches for irrigation and by creating artificial grassy marshes in the form of rice fields.

In conclusion, the genetic differentiation analysis of *An. aconitus* suggested clear genetic structure with highly significant genetic variation present among population groups in Southeast Asia, the closely unimodal mismatch distributions and extremely significant F_s values suggest a recent demographic expansion, and the population in Sri Lanka may be considered a distinct subspecies.

Acknowledgments

This work was supported by grant from National Natural Science Foundation of China (No. 31071968) and grants from the National Institute of Health (Nos. 1R01AI095184, R03 TW008237, U19 AI089672). B. Chen was supported by a Wellcome Trust Travelling Research Fellowship during his stay in UK. We would like to thank three anonymous reviewers for the comments on early version of this manuscript. A number of people provided field and laboratory assistance, mosquito samples and information used in the study. We are especially grateful to Yongxin Fu, Indira S. Weerasinghe, Sha Zhang, Jianren Huang, Weiben Li, Zhigang Liao, Jianming Peng, Michael Pocock, and Xuezhong Wang.

References

Arctander P, Kat PW, Aman AR, Siegismund HR. Extreme genetic differences among populations of Grant's gazelle *Gazella granti* in Kenya. Heredity. 1996; 76:465–475. [PubMed: 8666544]

Atmosoedjono S, Dennis DT. *Anopheles aconitus* and *Anopheles subpictus* naturally infected with *Wuchereria bancrofti* in Flores, Indonesia. Mos News. 1977; 37:529–529.

Audley-Charles, MG. Dispersal of Gonwanaland: relevance to evolution of the angiosperms. In: Whitemore, TC., editor. Biogeographical Evolution of the Malay Archipelago. Clarendon Press; Oxford: 1987. p. 5-21.

Avise JC. Gene trees and organismal histories: a phylogenetic approach to population biology. Evolution. 1989; 43:1192–1208.

Avise, JC. Molecular markers, natural history, and evolution. Chapman & Hall; New York: 1994.

Avise JC, Ball RM, Arnold J. Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA polymorphism and inbreeding theory for neutral mutations. Mol Biol Evol. 1988; 5:331–344. [PubMed: 3405076]

Baba K. A study of anopheline mosquitoes in relation to the epidemiology of malaria in Canton Delta, South China. Trans Kansai Ent Soc. 1950; 15:1–11.

Barcus MJ, Laihad F, Sururi M, et al. Epidemic malaria in the Menoreh Hills of Central Java. Am J Trop Med Hyg. 2002; 66:287–292. [PubMed: 12139222]

Chen B, Butlin RK, Harbach RE. Molecular phylogenetics of the Oriental members of the Myzomyia Series of *Anopheles* subgenus *Cellia* Diptera: Culicidae) inferred from nuclear and mitochondrial DNA sequences. Sys Ent. 2003; 27:57–69.

Chen B, Harbach RE, Butlin RK. Molecular and morphological studies on the *Anopheles minimus* group of mosquitoes in southern China: taxonomic review, distribution and malaria vector status. Med Vet Ent. 2002; 16:253–265.

Chen B, Harbach RE, Butlin RK. Genetic variation and population structure of the mosquito *Anopheles jeyporiensis* in southern China. Mol Ecol. 2004; 13:3051–3056. [PubMed: 15367119]

Chen B, Pedro MP, Harbach RE, Somboon P, Walton C, Butlin RK. Mitochondrial DNA variation in the malaria vector *Anopheles minimus* across China, Thailand and Vietnam: evolutionary hypothesis, population structure and population history. Heredity. 2011; 106:241–252. [PubMed: 20517346]

Clement M, Posada D, Crandall KA. TCS: a computer program to estimate gene genealogies. Mol Ecol. 2000; 9:1657–1659. [PubMed: 11050560]

Clements KM, Wiegmann BM, Sorenson CE, et al. Genetic variation in the *Myzus persicae* complex Homoptera: Aphididae: evidence for a single species. Ann Ent Soc Am. 2000; 93:31–46.

Collins FH, Paskewitz SM. Malaria: Current and future prospects for control. Ann Review Ent. 1995; 40:195–219.

Cronin MA. Intraspecific variation in mitochondrial DNA of North American cervids. J Mammol. 1992; 73:70–82.

Damar T, Fleming GA, Gandahusada S, Bang YH. Nocturnal indoor resting heights of the malaria vector *Anopheles aconitus* and other anopheles (Diptera: Culicidae) in Central Java, Indonesia. J Med Ent. 1981; 18:362–365.

Dönitz W. Beitrage zur Kenntniss der *Anopheles*. Z Hyg Infektionskr. 1902; 41:15–88.

Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 1992; 131:479–491. [PubMed: 1644282]

Excoffier L, Lischer HEL. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Eco Res. 2010; 10:564–567.

Franck P, Garnery L, Loiseau A, et al. Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data. Heredity. 2001; 86:420–430. [PubMed: 11520342]

Fu YX. Statistical tests or neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 1997; 147:915–925. [PubMed: 9335623]

Fu YX, Li WX. Statistical test of neutrality of mutations. Genetics. 1993; 133:693–709. [PubMed: 8454210]

Garros C, Harbach RE, Manguin S. Morphological assessment and molecular phylogenetics of the Funestus and Minimus Groups of *Anopheles* (*Cellia*). J Med Ent. 2005; 43:522–536.

Gould DJ, Esah S, Pranith U. Relation of *Anopheles aconitus* to malaria transmission in the central plain of Thailand. Trans Roy Soc Trop Med Hyg. 1965; 59:441–442. [PubMed: 14347463]

Harbach RE. The classification of genus *Anopheles* (Diptera: Culicidae): a working hypothesis of phylogenetic relationships. Bul Ent Res. 2004; 95:537–553.

Harbach RE, Garros C, Manh ND, Manguin S. Formal taxonomy of species C of the *Anopheles minimus* sibling species complex (Diptera: Culicidae). Zootaxa. 2007; 1654:41–54.

Harpending HC, Batzer MA, Gurven M, et al. Genetic traces of ancient demography. P Nat Acad Sci USA. 1998; 95:1961–1967.

Harrison BA. Medical entomology studies - XIII. The Myzomyia Series of *Anopheles* (*Cellia*) in Thailand, with emphasis on intra-interspecific variations (Diptera: Culicidae). Cont Am Ent Inst. 1980; 17:1–195.

Jin ZZ. The diversity features of the floristic elements in the tropical rain forest of Xishuangbanna, Yunnan. Acta Bot Yunnanica (suppl). 1997; 1:129–148.

Jin ZZ, Ou XK. The diversity features of plant community types in the tropical rain forest vegetation of Xishuangbanna, Yunnan. Acta Bot Yunnanica (suppl). 1997; 1:1–30.

Kengne P, Trung HD, Baimai V, et al. A multiplex PCR-based method derived from random amplified polymorphic DNA (RAPD) markers for the identification of species of the *Anopheles minimus* group in Southeast Asia. Ins Mol Biol. 2001; 10:427–435.

Krafsur ES, Wohlford DL. Breeding structure of *Glossina pallidipes* populations evaluated by mitochondrial variation. J Heredity. 1999; 90:635–642.

Krzywinski J, Wilkerson RC, Besansky NJ. Toward understanding Anophelinae (Diptera, Culicidae) phylogeny: Insights from nuclear single-copy genes and the weight of evidence. Sys Biol. 2001; 50:540–556.

Kuhner MK. LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics. 2006; 22:768–770. [PubMed: 16410317]

Li, WH.; Graur, D. Fundamentals of Molecular Evolution. Sinauer Associates; Sunderland, MA: 1991.

Lu, BL. Diptera: Culicidae II. Vol. 9. Science Press; Beijing: 1997. Fauna Sinica, Insecta.

Meshnick SR. Why does quinine work after 350 years of use? Parasitol Today. 1997; 13:89–90. [PubMed: 15275109]

Morgan K, O'Loughlin SM, Chen B, Linton YM, Thongwat D, Somboon P, Fong MY, Butlin R, Verity R, Prakash A, Htun PT, Hlaing T, Nambanya S, Socheat D, Dinh TH, Walton C. Comparative phylogeography reveals a shared impact of pleistocene environmental change in shaping genetic diversity within nine *Anopheles* mosquito species across the Indo-Burma biodiversity hotspot. Mol Ecol. 2011; 20:4533–4549. [PubMed: 21981746]

Nei, M. Molecular Evolutionary Genetics. Columbia University Press; New York: 1987.

Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. Bioinformatics. 1998; 14:817–818. [PubMed: 9918953]

Posada D, Crandall KA. Intraspecific gene genealogies: Trees grafting into networks. Trends Ecol Evol. 2001; 16:37–45. [PubMed: 11146143]

Powell JR, Caccone A, Amato GD, Yoon C. Rates of nucleotide substitution in *Drosophila* mitochondrial DNA and nuclear DNA are similar. Proceeding of National Academy of Science of USA. 1986; 83:9090–9093.

Rahman WA, Abuhassan A, Adanan CR. Seasonality of *Anopheles aconitus* mosquitoes, a secondary vector of malaria, in an endemic village near the Malaysia-Thailand border. Acta Tropica. 1993; 55:263–265. [PubMed: 8147283]

Rogers AR, Harpending H. Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol. 1992; 9:552–569. [PubMed: 1316531]

Rongnoparut P, Sirichotpakorn N, Rattanarithikul R, et al. Estimates of gene flow among *Anopheles maculatus* populations in Thailand using microsatellite analysis. Am J Trop Med Hyg. 1999; 60:508–515. [PubMed: 10466986]

Rua PG, Serrano J, Moritz RFA. Genetic structure and distinctness of *Apis mellifera* L. populations from the Canary Islands. Mol Ecol. 2001; 10:1733–1742. [PubMed: 11472540]

Sharpe RG, Harbach RE, Butlin RK. Molecular variation and phylogeny of members of the Minimus Group of *Anopheles* subgenus *Cellia* (Diptera: Culicidae). Sys Ent. 2000; 25:263–272.

Slatkin M. A measure of population subdivision based on microsatellite allele frequencies. Genetics. 1995; 139:457–462. [PubMed: 7705646]

Slatkin M, Hudson RR. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics. 1991; 129:556–562.

Sperling FAH, Raske AG, Otvos IS. Mitochondrial DNA sequence variation among populations and host races of *Lambdina fiscellaria* (Gn) (Lepidoptera: Geometridae). Insect Mol Biol. 1999; 8:97–106. [PubMed: 9927178]

Swofford, D. PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods), Version 4.0b8. Sinauer Associates, Inc; Sunderland, Massachusetts: 2001.

Tajima F. Evolutionary relationship of DNA sequences in finite populations. Genetics. 1983; 123:437–460. [PubMed: 6628982]

Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989; 123:585–595. [PubMed: 2513255]

Tajima F. The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. Genetics. 1996; 143:1457–1465. [PubMed: 8807315]

Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. Mol Biol Evol. 2007; 24:1596–1599. [PubMed: 17488738]

Templeton AR, Crandall KA, Sing CF. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III Cladogram estimation. Genetics. 1992; 132:619–633. [PubMed: 1385266]

Thompson JD, Gibson TJ, Plewniak F, et al. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nuc Acids Res. 1997; 24:4876–4882.

Van Bortel W, Trung HD, Roelants P, et al. Molecular identification of *Anopheles minimus s.1* beyond distinguishing the members of the species complex. Insect Mol Biol. 2000; 9:335–340. [PubMed: 10886418]

Walton C, Handley JN, Tun-Lin W, et al. Population structure and population history of *Anopheles dirus* mosquitoes in Southeast Asia. Mol Biol Evol. 2000; 17:962–974. [PubMed: 10833203]

Watterson GA. On the number of segregation sites in genetic models without recombination. Theor Pop Biol. 1975; 7:256–276. [PubMed: 1145509]

Zhao ZW, Wang GH, Yang I. Biodiversity of arbuscular mycorrhizal fungi in a tropical rainforest of Xishuangbanna, southwest China. Fung Div. 2003; 13:233–242.

Zhu H. Ecological and biogeographical studies on the tropical rain forest of south Yunnan, SW China with a special reference to its relation with rain forests of tropical Asia. J Biogeog. 1997; 24:647–662.

Zhu H, Xu ZH, Wang H, Li BG. Tropical rain forest fragmentation and its ecological and species diversity changes in southern Yunnan. Biodiv Cons. 2004; 13:1355–1372.

Phuc HK, Ball AJ, Son L, Hanh NV, Tu ND, Lien NG, Verardi A, Townson H. Multiplex PCR assay for malaria vector *Anopheles minimus* and four related species in the Myzomyia Series from Southeast Asia. Medical and Veterinary Entomology. 2003; 17:423–428. [PubMed: 14651657]

Garros C, Koekemoer LL, Coetzee M, Coosemans M, Manguin S. A single multiplex assay to identify major malaria vectors within the African *Anopheles funestus* and the Oriental *An. minimus* groups. American Journal of Tropical Medicine and Hygiene. 2004; 70:583–590. [PubMed: 15210996]

Ranson H, N'Guessan R, Lines J, Moiroux N, Nkuni Z, Corbel V. Pyrethroid resistance in African anopheline mosquitoes: what are the implications for malaria control? Trends in Parasitology. 2011; 27:91–98. [PubMed: 20843745]

Cui L, Yan G, Sattabongkot J, Cao Y, Chen B, Chen X, Fan Q, Fang Q, Jongwutiwes S, Parker D, Sirichaisinthop J, Kyaw MP, Su XZ, Yang H, Yang Z, Wang B, Xu J, Zheng B, Zhong D, Zhou G. Malaria in the Greater Mekong Subregion: Heterogeneity and complexity. Acta Tropica. 2011; 121:227–239. [PubMed: 21382335]

Van Bortel W, Trung HD, Thuan LK, Sochantha T, Socheat D, Sumrandee C, Baimai V, Keokenchanh K, Samlane P, Roelants P, Denis L, Verhaeghen K, Obsomer V, Coosemans M. The insecticide resistance status of malaria vectors in the Mekong region. Malaria Journal. 2008; 7:15. [PubMed: 18205915]

Van Bortel W, Trung HD, Sochantha T, Coosemans M. Absence of knockdown resistance suggests metabolic resistance in the main malaria vectors of the Mekong region. Malar J. 2009; 8:84. [PubMed: 19400943]

Cui, Feng; Raymond, M.; Qiao, CL. Insecticide resistance in vector mosquitoes in China. Pest Management Science. 2006; 62:1013–1022. [PubMed: 16953491]

Sharpe RG, Hims MM, Harbach RE, Butlin RK. PCR-based methods for identification of species of the *Anopheles minimus* group: allele-specific amplification and single-strand conformation polymorphism. Med Vet Ent. 1999; 13:265–273.

The population in Sri Lanka of *An. aconitus* may be considered a distinct subspecies.

Clear genetic structure was observed with highly significant genetic variation present among population groups in China and Southeast Asia.

Mismatch distributions and extremely significant F_s values suggest a recent demographic expansion.

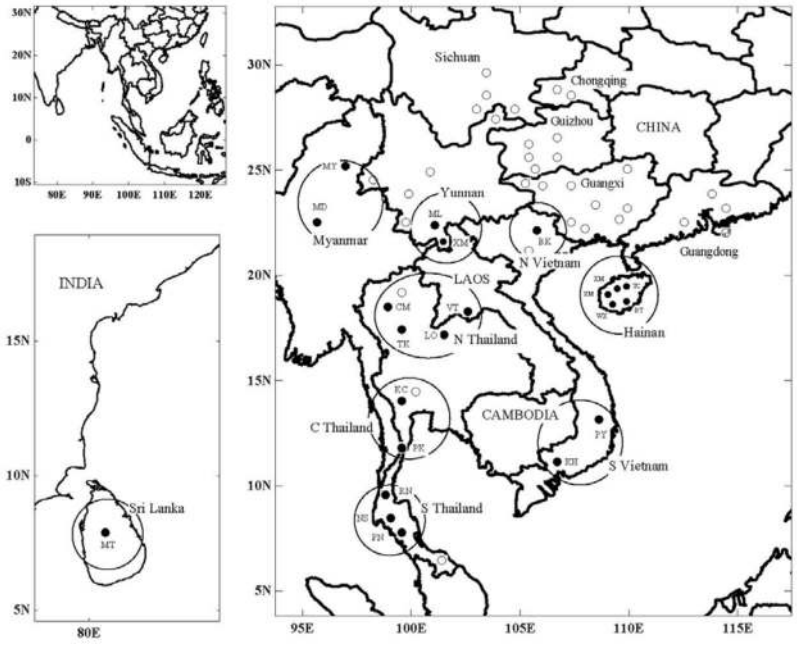

Fig. 1.

Localities collected for populations of *An. aconitus* in China, Myanmar, Vietnam, Thailand, Laos and Sri Lanka. The 22 populations with samples collected are shown with filled circles and divided into 9 population groups (in separate circles) based on their geographical distributions. The sites without samples encountered are shown with empty circles. The top-left frame shows the previously suggested distribution area of the species.

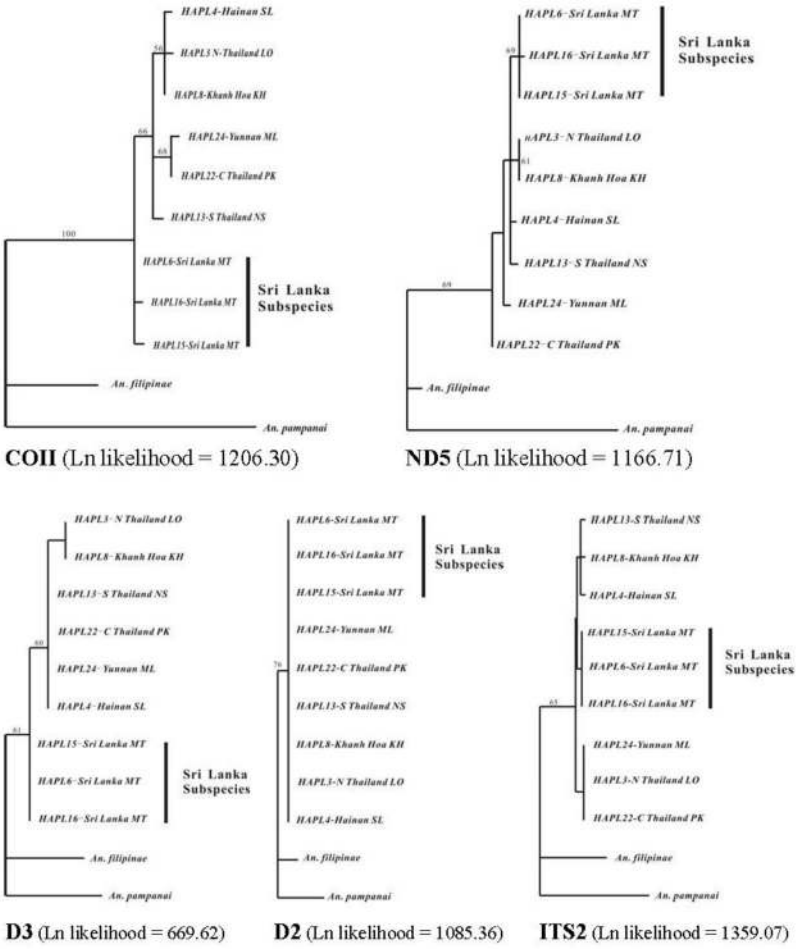

Fig. 2.

Maximum-likelihood trees of sequences of mtDNA *COII* and *ND5* genes and rDNA D3, D2 and ITS2 loci (*An. filipinae* and *An. pampanai* as outgroup taxa), showing their variation within *An. aconitus* and the subdivision of the Sri Lankan population to subspecies. The sample names indicate their *COII* haplotypes and localities. Bootstrap percentages of 1000 replicates calculated with the maximum parsimony method are shown above the branches where they exceed 50%. Branch lengths are proportional to the number of character differences.

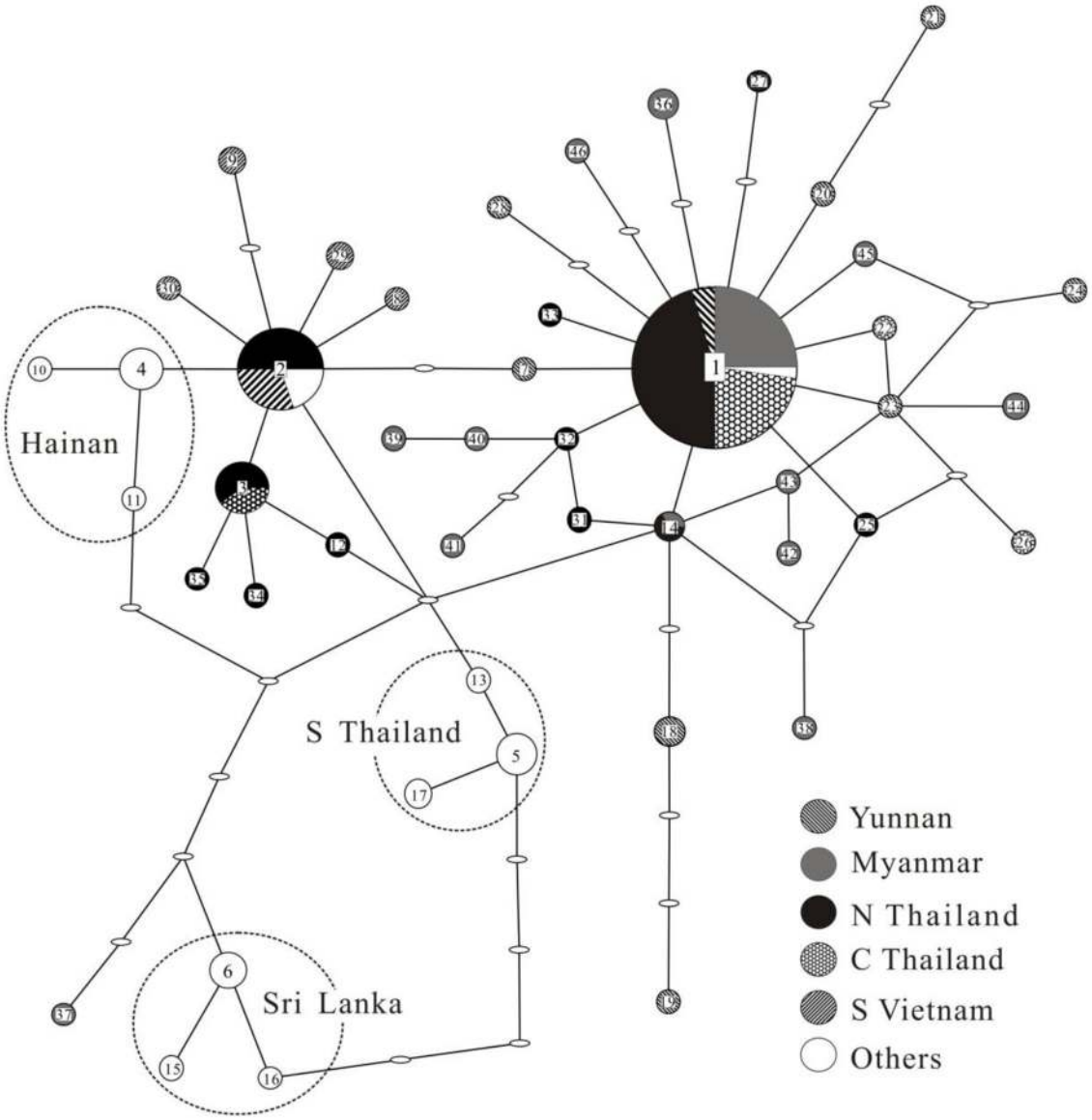

Fig. 3.

Genealogical relationships among 46 *COII* haplotypes of *An. aconitus* estimated by TCS (Clement et al., 2000). The size of a circle corresponds to the *COII* haplotype frequency, and a unit branch represents one mutation. Small ovals indicate *COII* haplotypes that were not observed.

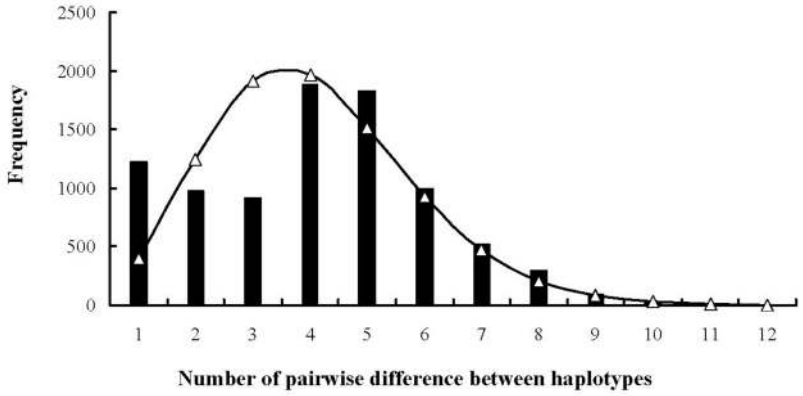

Fig. 4.

Observed mismatch distributions among pairwise differences among haplotypes for all populations except Sri Lanka, which is closely unimodal and fits to a Poisson distribution (χ^2: $s^2 = 1.32$ and $P < 0.001$).

Table 1

Summary of populations, haplotypes, nucleotide diversity and neutrality tests of *An. aconitus. S* is the number of segregating sites, and θ_π and θ_s are the estimates of nucleotide diversity per base.

Population group[a]	Collection Site[b]	N	Haplotypes[c]	*S*	θ_π±SD (×10^{-3})	θ_s ± SD (×10^{-3})	Tajima's D	Fu's *Fs*
Hainan (China)	Baoting (BT, 1), Shilu (SL, 2), Wangxia (WX, 2), Danzhou (DZ, 3), Tunchang (TC, 1)	9	**4**(7), 10, 11	2	0.70±0.77	1.15±0.88	−1.36	−18.17**
Yunnan (China)	Menglun (ML, 9), Xiangming (XM, 2)	11	**1**(2), 7, **18**(2), 19, 20, 21, 23, 24, 28	13	5.80±3.58	6.95±3.24	−0.72	−8.56**
Myanmar		25		21	4.13±2.54	8.70±3.30	−1.91*	−26.54**
Kachin	Myitkyina (MY, 12)	12	**1**(7), **36**(3), 38, 46	7	2.58±1.84	3.63±1.88	−1.13	−15.73**
Sagaing	Maung Daung (MD, 13)	13	**1**(4), **14**, 37, 39, 40, 41, 42, 43, 44, 45	16	5.26±3.24	8.07±3.55	−1.46*	−12.09**
N Vietnam	Bac Kan (BK, 4)	4	**2**(4)					
S Vietnam	Khanh Hoa (KH, 7), Phu Yen (PY, 5)	12	**2**(6), 8, **9**(2), **29**(2), 30	5	1.94±1.49	2.59±1.46	−0.92	−18.09**
N Thailand		46		10	3.38±2.12	3.56±1.47	−0.15	−26.97**
N Thailand	Chiang Mai (CM, 7), Loei (LO, 12), Tak (TK, 3)	22	**1**(8), **2**(4), **3**(5), 12, **14**(2), 25, 27	7	3.50±2.23	3.01±1.46	0.52	−26.89**
Vientiane (Laos)	Vientiane (VT, 24)	24	**1**(12), **2**(6), **3**, 31, 32, 33, 34, 35	8	3.32±2.14	3.35±1.56	−0.03	−26.95**
C Thailand	Kanchanaburi (KC, 4), Prachuap Khiri Khan (PK, 12)	16	**1**(10), **3**(4), 22, 26	7	3.27±2.16	3.30±1.65	−0.03	−21.91**
S Thailand	Nakhon Si Thamarat (NS, 7), Phang Nga (PN, 2), Ranong (RN, 1)	10	**1**, **5**(6), 13, **17**(2)	5	2.05±1.58	2.77±1.59	−1.04	−13.24**
Sri Lanka	Matale (MT, 7)	7	**6**(5), 15, 16	2	0.89±0.93	1.28±0.99	−1.24	−11.15**
All Populations (without Sri Lanka)		133		45	4.82±2.79	12.89±3.48	−1.84**	−26.36**
All Populations		140		47	5.31±3.03	13.34±3.56	−1.92**	−26.13**

[a] N Vietnam with only 4 samples and relatively far geographical separation is not analyzed as a group, and Vientiane in Laos is combined into N Thailand;

[b] Code and individual number at each site is in parentheses;

[c] Haplotypes in bold occur in more than one population and the number in parentheses indicates frequency of the haplotype at the site;

* P<0.05;

** P<0.01.

Table 2

Variation in the 639 bp sequence of the *COII* gene of *An. aconitus*.

Haplotype	Sites 1111222222233333333344444444445555555512234778567701345790122335590123667799111225677 6780425195141044897684703475240038175606928 2706	GenBank accession no.
HAPL1(44)	GACACTTTCTACTAACTTAATTTATATATTTGGTATA	AJ626951
HAPL2(20)	GAAATTTAAA	AJ626952
HAPL3(10)	.....G.....G...C...	AJ626953
HAPL4(7)	..T.....G.....G...C...	AJ626954
HAPL5(6)	.....C.....G.....G...C...	AJ626955
HAPL6(5)	.....T.....C.....G...C...	AJ626956
HAPL7(1)	.....C.....G.....C..A.....G...C...	AJ626957
HAPL8(1)	.....G.....	AJ626958
HAPL9(2)	.....G.....G..CC...	AJ626959
HAPL10(1)	.....G.....G.....G.....G...C...	AJ626960
HAPL11(1)	.....C...T.....G.....G...C...	AJ626961
HAPL12(1)	.....C.....G.....A.....G...C...	AJ626962
HAPL13(1)	..T.....G...C...	AJ626963
HAPL14(3)	.....C.....G...C...	AJ626964
HAPL15(1)	.....C...	AJ626965
HAPL16(1)	.....C.....G...G....C..A.....G...C...	AJ626966
HAPL17(2)	.....T.C.....G.....C..A.....G...C...	AJ626967
HAPL18(2)	.....T.....C..C.....G...C...	AJ626968
HAPL19(1)	.....T.....G.....C...	AJ626969
HAPL20(1)	...G....T.....CC..G.....C...	AJ626970
HAPL21(1)	.....C.....	AJ626971
HAPL22(1)	.....CCC.....	AJ626972
HAPL23(1)	.....G.....	AJ626973
HAPL24(1)	.....C.....	AJ626974
HAPL25(1)	A.....C.....C.....	AJ626975
HAPL26(1)	.G.....	AJ626976
HAPL27(1)	.G.....C.....C.....	AJ626977
HAPL28(1)	.....G..C.....	AJ626978
HAPL29(2)	.....CG.....	HQ403792.
HAPL30(1)	.....G.....C...G...C...	1
HAPL31(1)	.....C.....G.....G...C...	HQ403794.
HAPL32(1)	.....C.G.	1
HAPL33(1)	.....G.	HQ403776.
HAPL34(1)	.....A.....	1
HAPL35(1)	.GT.....G.....G...C...	HQ403771.
HAPL36(3)	..T.....G.....A.G...C...	1
HAPL37(1)	.....G.C.....	HQ403765.
HAPL38(1)	.....C.....C..A.....GGG..C...	1
HAPL39(1)	.G.....C.....C...	HQ403757.
HAPL40(1)	.....G.G.....G.	1
HAPL41(1)	.....G.....G.	HQ403772.
HAPL42(1)	.....T.....C.....G.	1
HAPL43(1)	.....C.....CG..	HQ403808.
HAPL44(1)	.....C.....C...	1
HAPL45(1)	....T.....C.....	HQ403804.
HAPL46(1)	.....C.....	1
	.....G.....T	HQ403812.

1
HQ403799.
1
HQ403795.
1
HQ403796.
1
HQ403800.
1
HQ403802.
1
HQ403806.
1
HQ403807.
1
HQ403814.
1
HQ403799.
1
HQ403795.
1
HQ403796.
1
HQ403800.

Frequencies of haplotypes are in parentheses and variable positions are shown in the first row. Identity with the first sequence is denoted by a dot and substitution by a different base letter.

Table 3

Hierarchical analysis of molecular variance based on pairwise difference for *An. aconitus*.

Source of variation	d. f.[a]	Sum of squares	Variance components	% total variation	Φ-statistics	*P*[b]
Among groups	6	69.987	0.606	36.09	Φ_{CT} = 0.361	<0.001
Among populations within groups	13	15.927	0.036	2.13	Φ_{SC} = 0.033	<0.05
Within populations	126	113.063	1.037	61.78		

[a]Degrees of freedom.

[b]*P* indicates the probability of having a more extreme variance component tested under 1023 random permutations.

Table 4

Genetic distances within and between population groups.

Group	Hainan	Yunnan	Myanmar	S Vietnam	N Thailand	C Thailand	S Thailand	Sri Lanka
Hainan	0.069							
Yunnan	0.606**	0.580						
Myanmar	0.627**	0.059*	0.413					
S Vietnam	0.533**	0.515**	0.543**	0.194				
N Thailand	0.504**	0.179**	0.153**	0.340**	0.338			
C Thailand	0.637**	0.083*	0.051*	0.501**	0.014	0.327		
S Thailand	0.783**	0.519**	0.538**	0.662**	0.495**	0.565**	0.205	
Sri Lanka	0.920**	0.689**	0.702**	0.833**	0.704**	0.756**	0.837**	0.089

On the diagonal are the percent pairwise differences within groups. Below the diagonal are the genetic distances between groups calculated as pairwise F_{st} values, with those $P < 0.05$ and $P < 0.01$ based on 1023 permutations shown by * and **.

Table 5

FLUCTUATE estimates of effective population size and exponential growth rate for *An. aconitus*.

θ_0 ± SD	N_0	g ± SD
0.99 ± 0.145	2.47×10^7	1488.83 ± 75.23

Point estimates of θ_0 and g are shown with approximate 95% confidence intervals. The exponential growth rate, g, is in units of $1/\mu$ per generation ($\theta = 2N\mu$ in haploid organism). N_0, the current effective population size of females, is estimated from θ_0 assuming μ of 1×10^{-8} per site per generation (Powell *et al.* 1986).