Genetic variability and divergence of *Neutrodiaptomus tumidus* Kiefer 1937 (Copepoda: Calonida) among 10 subpopulations in the high mountain range of Taiwan and their phylogeographical relationships indicated by mtDNA COI gene

Shuh-Sen Young¹*, Yan-Ying Lee¹ and Min-Yun Liu²

**Abstract**

**Background:** In the mountain area of Taiwan, we investigated 10 subpopulations of *Neutrodiaptomus tumidus* Kiefer 1937 living in isolated alpine ponds or lakes. We used mitochondrial DNA cytochrome C oxidase subunit I (COI) sequence as molecular marker to investigate the population genetic structure and their phylogeographical relationships.

**Results:** We obtained 179 sequences from 10 subpopulations and found 94 haplotypes. Nucleotide composition was AT-rich. Haplotype diversity (Hd) and nucleotide diversity (π) indicated significant genetic differences between subpopulations (Hd = 0.131 ~ 0.990; π = 0.0002 ~ 0.0084); genetic differentiation index (Fst) and gene flow index (Nm) also exhibited significant genetic diversification between subpopulations (Fst = 0.334 ~ 0.997; Nm = 0 ~ 1). Using Tajima’s D and Fu and Li’s D* and F* as neutrality tests, we found that the nucleotide variation within the population was consistent with the neutral theory except in the JiaLuoHu subpopulation. The JiaLuoHu subpopulation significantly deviated from the neutral theory and was speculated to have experienced a bottleneck effect. According to the phylogenetic tree, these alpine lake subpopulations could be divided into four phylogroups (northern region, Xueshan group, central region, and southwestern region). Xueshan group contains one subpopulation, DuRongTan, which is a unique group relative to other groups. It is close to northern group geographically but closer to southwestern group genetically. According to AMOVA, the major genetic variation came from different geographical distribution of subpopulations. Molecular clock estimates that the northern and southern regional divergence time was about 2.2 ~ 3.9 MYA, when the Central Mountain Range uplift (3 ~ 5 MYA) caused the population of *N. tumidus* to be segregated into northern and southern parts.

**Conclusions:** Significant genetic divergence between each subpopulation of *N. tumidus* was found in this study. This result indicated the low dispersion ability of planktonic copepods with limited gene flow between each subpopulation.

**Keywords:** Metapopulation; Population genetics; Phylogeography; Copepoda; mtDNA; COI

* Correspondence: shuh@mail.nhcue.edu.tw
¹Department of Applied Science, National Hsinchu University of Education, 521, Nan-Da Road, Hsinchu city 30014, Taiwan
Full list of author information is available at the end of the article

© 2014 Young et al.; licensee Springer. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.
Background
Based on the geological record, Taiwan, an island separated from mainland China by the Taiwan Strait, emerged above sea level not more than 500 million years ago (MYA). About 300 MYA, the first uplifted land appeared as the earliest stage of the Central Mountain Range (Teng 2002). After the formation of the Central Mountain Range, other series of mountain formations formed Xueshan Range, Yushan Range, and Alishan Range (Figure 1). Today these four major mountain systems cover the two thirds of the land area of Taiwan; the Central Mountain Range stretches along the central area of island from north to south and off to the east. Along the west side of central mountain ridge are the Xueshan Range and Yushan Range. West of Alishan and slightly further south is the Yusan Range. The northern part of Xueshan Range is more isolated from the main range, and the LanYang River system flows through the deep valley between Xueshan Range and Central Mountain Range and northward into Ilan Bay. DaJia River system runs across middle region of Xueshan Range then through the western plain into the Taiwan Strait, with its headwaters near the origin of LanYang River. Zhuoshui
River, originated from the Central Mountain Range (HeHuanShan), also flows westward and through the mountains between Xueshan and Yushan Ranges into the Taiwan Strait.

Besides rivers, there are small lakes and ponds scattered in remote areas of the mountains. According to Chen and Wang (1997), the high-mountain water body of Taiwan can be divided into two types: type I is above the forest line on the alpine grassland and type II is at the lower altitudes with dense forest around the water body. Lakes or ponds above the forest line usually are small in size and isolated from the river system without clear connection, their volume of water is much fluctuated seasonally, and some of them are dried out during long period without rain. Lakes and ponds in the forest are normally larger in size and deeper than those of the type I and always have clear connection to the river system; their volume of water also fluctuated seasonally but never dried out. Disregarding the types, lakes and ponds are important natural habitats for all kinds of wildlife as well as human beings.

Some general ecological survey of highland lakes and ponds of Taiwan had been conducted in the past (Kano and Yoshimizu 1934; Otsu et al. 1988, 1989; Otsu et al. 1992a, b; Lin et al. 2006, 2011a, b). Otsu et al. (1988) reported the presence of Eudiaptomus formosus from DaGuiHu (Big Ghost Lake) (Taromalin in Japanese) in southern Taiwan. However, Young and Shih (2011) surveyed Taromalin in 1997 and 2008 and only found Neutrodiaptomus tumidus. Besides Taromalin, Young and Shih (2011) also found N. tumidus in freshwater bodies within the mountain area, with altitudes ranging from 1,300 to 3,500 m and remote from human activity. All the taxonomic surveys on the mountain area also found many species of Cladocera, cyclopoid, and harpacticoid copepods (Young 2010; Young and Tuo 2011; Tuo and Young 2011). Some populations of N. tumidus inhabit in small ponds and lakes higher than the forest line and have absolutely no water connection among each other. Some populations live in type II water bodies and are isolated by different river systems. To investigate the isolation and dispersal of these scattered populations of N. tumidus, studies on their metapopulation genetics can be useful.

In the metapopulation model, each subpopulation is considered as a small pool in a large area, and the population dynamics is based on colonization-extinction equilibrium between each subpopulation (Hanski 1998, 1999). Freshwater invertebrates, including copepods, living in lakes or ponds not always connected to each other by water way in high mountain area are strong candidates for metapopulations. According to Bohonak and Jenkins (2003), if there is a metapopulation in freshwater invertebrate populations, dispersal studies play a crucial role in quantifying the effects of individual migration. Dispersal of individuals can inhibit extinction of subpopulation and maintain genetic diversity. The genetic divergence studies between each subpopulation can provide basic evidence of dispersion and estimation of the gene flow rate.

In this study, we report the levels of genetic variability and divergence within and among 10 subpopulations of N. tumidus isolated in lakes and ponds in the high mountain area. We screened sequence of mtDNA coding for a fragment of cytochrome C oxidase subunit I (COI). This gene has been proven informative at the interspecific and intraspecific levels in crustacean (Henzler 2006; Kelly et al. 2006; Ketmaier et al. 2010; Pavesi et al. 2011; Young et al. 2012). More specifically, first we tested whether the geographical isolation of lakes and ponds will prevent gene flow or not, and second we constructed the phylogeographical relationships of the subpopulations.

**Methods**

Ten samples were taken using a plankton net with mesh size of 55 μm. All the collection sites were isolated from each other by long distance and absence of water connection (Figure 1). Samples were fixed in 70% ethanol (EtOH) immediately after collection then transferred to 95% EtOH and stored at a low temperature (less than ~20°C). Within 72 h, each sample was sorted and identified under a stereomicroscope. Based on sampling sites, 10 populations were collected (Table 1). Individuals in each population were randomly selected for the analysis of their genetic structure by COI gene marker.

**DNA extraction, amplification, and sequencing**

Total genomic DNA was extracted from single animals using Chelex (InstaGene Matrix BIO-RAD 7326030, Bio-Rad Laboratories, Hercules, CA, USA). Each specimen from 95% EtOH was taken and placed in pure water for 1 h for cleaning. Each of the cleaned specimen was placed at the bottom of a 0.5-ml centrifuge tube for 30 min to dry in a speed vacuum-drying system. Dried specimens were then crushed by needles and incubated in 30 μl of 5% Chelex solution at 56°C for 2 to 3 h to extract the DNA then continued for 8 min at 90°C. For each polymerase chain reaction (PCR), 5 μl of upper cleaning solution was used as the DNA template after centrifugation at 10,000 rpm (9,168 × g) for 3 min.

We employed the universal primers, LCO1490 (5′-GGTCAACAATCTAATAGATATGG-3′) and HCO2198 (5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) (Folmer et al. 1994), to amplify the mitochondrial COI gene by a PCR. For the population of DaGuiHu, HCO2198 and LCO1490 could not be successfully applied. We designed another primer DGF (5′-TGGTTGATAGATTGGATCTC-3′) and DGR (5′-GACTTTGTTATTTTA TGCGCCG-3′) based on the sequence amplified from HCO2198 and LCO1490. Each PCR sample had a total
volume of 50 μl and consisted of pH 9.2 buffer solution (50 mM Tris-HCl, 16 mM ammonium sulfate, 2.5 mM MgCl₂, and 0.1% Tween 20), 5 pM of each primer, 50 μM of dNTPs, 2 units of Taq DNA polymerase (super Therm DNA polymerase, Bio-Taq, BioKit Biotechnology, Miaoli, Taiwan), and 10 to 50 ng of genomic DNA. The PCRs were performed in an Eppendorf Mastercycler gradient 384 machine (Eppendorf, Hamburg, Germany). Thermocycling began with 5 min of preheating and continued with 35 cycles at 94°C for 30 s, primer annealing at 51°C for 45 s, and extension at 72°C for 45 s, followed by incubation at 72°C for 10 min for full extension of the DNA and ended with 4°C holding. PCR products were electrophoresed in 2% agarose gels, after which the gels were stained with ethidium bromide (EtBr) and photographed under an ultraviolet light box. DNA fragments were excised from the gel and extracted using a 1-4-3 DNA extraction kit (Gene-Spin, Protech, Taipei, Taiwan) to obtain purified DNA. Sequences of DNA fragments were resolved on an ABI3730 automated sequencer (Applied Biosystems, Carlsbad, CA, USA) using 20 to 50 ng of template with 5 pM of the LCO1490 and HCO2198 primers read from both directions.

Alignment, genetic diversity, and phylogeny

COI gene sequences were aligned by eye using the SeqMan (DNASTAR, Lasergene version 7.1, 2006). We calculated the haplotype diversity (Hd, Nei 1987), nucleotide diversity (π, Nei 1987), genetic differentiation indices (Fst, Hudson et al. 1992), gene flow index (Nm, Hudson et al. 1992), genetic distance (Dxy, Nei 1987; average genetic distances between each pair of species), and neutrality test (including Tajima D test (Tajima 1989), Fu and Li’s D* and Fu and Li’s F* (Fu and Li 1993)), using DNAsp version 5.10 (Rozas 2009). The index of neutrality test which equals to 0 indicated that the population genetic structure reaches to neutral expectation. The positive value indicates the population genetic structure under some type of balancing selection or diversifying selection. The negative value indicates that an excess of rare alleles is selected against some genotypes or population under growing condition with excess rare alleles (Hartl and Clark 2007). The COI sequences of Mongolodiaptomus birulai (GenBank accession number AB592995) and Neodiaptomus schmackeri (GenBank accession number AB592987) from Lin (2004) and Young et al. (2013) were used as out groups, and the phylogeographical tree was derived using all sequences by the maximum likelihood (ML), maximum-parsimony (MP), and Bayesian inference (BI). The maximum likelihood analysis was run by MEGA version 6 (Tamura et al. 2013) using model GTR + G with 100 runs and found the best ML tree by comparing the likelihood scores. The robustness of the ML tree was evaluated by 1,000 bootstraps. Maximum-parsimony methods (Saitou and Nei 1987) were based on Kimura two-parameter (K2P) distances with 1,000 bootstraps using MEGA version 6 (Tamura et al. 2013).

We also used MrBayes version 3.1.2 (Huelsenbeck et al. 2001) to run BI. Before BI analyses, the model was selected by MrModeltest 3.7 (Posada and Crandall 2001), and the optimal evolutionary parameters was selected by Akaike information criterion (AIC) from MrModeltest 3.7. The obtained best models were GTR + I + G and HKY + I + G, respectively. The search was run with four chains for 10 million generations and four independent runs, with trees sampled every 1,000 generations. The convergence of the chains was determined by the effective sample size (ESS) (>200 as recommended) in Tracer (v. 1.5, Rambaut and Drummond 2009), and the first 1,000 trees were discarded as the burnin (determined by the average standard deviation

Table 1 Sampling location, sample size, and collection date of 10 populations of Neutrodiaptomus tumidus

| Geographic distribution (mountain range) | Population site (code) | Altitude (m) | Sample size (N) | Collection date |
|-----------------------------------------|------------------------|-------------|----------------|----------------|
| Xueshan                                 | SongLuoHu (SL)         | 1,300       | 15             | February 2009  |
| Xueshan                                 | DuRongTan (DR)         | 2,350       | 18             | May 2009       |
| Central mountain                        | JiaLuoHu (JL)          | 2,320       | 30             | April 2005     |
| Central mountain                        | TaiwanChi (TC)         | 2,900       | 15             | November 2009  |
| Central mountain                        | BaiShiChi (BS)         | 2,770       | 15             | November 2009  |
| Central mountain                        | TunLuChi (TL)          | 2,850       | 15             | November 2009  |
| Central mountain                        | ChiTsaiHu (CT)         | 2,900       | 15             | February 2010  |
| Central mountain                        | JiaMingHu (JM)         | 3,310       | 18             | September 2011 |
| Central mountain                        | ShihShanSiouHu (SS)    | 2,300       | 25             | April 2011     |
| Central mountain                        | DaGuiHu (DG)           | 2,150       | 13             | January 2008   |
| Total                                   |                        |             | 179            |                |

The locality names are phonetically translated by the common use romanization system.
of split frequency values below the recommended 0.01; Ronquist et al. 2005).

The minimum spanning network among each population was constructed by TCS version 1.21 (Clement et al. 2000). Between populations and geographic groups, the analyses of molecular variance (AMOVA) were executed by Arlequin version 3.1 (Excoffier and Schneider 2005) to verify the variance between each subpopulations in the same geographic groups and different geographic groups. For each subpopulation, the diversification level was estimated by the index of Harpending (1994) and Ramos-Onsins and Rozas (2002). Harpending’s raggedness index was estimated to test whether the sequence data from each population deviated from what is expected under a sudden expansion, which is based on mismatch distributions of pairwise differences among all haplotypes in a sample. Ramos-Onsins and Rozas’s R2 neutrality tests by comparison of the difference between the number of singleton mutation and the average number of nucleotide differences were conducted.

Results

Analysis of genetic diversity

We had 179 sequences (individuals) in total from 10 subpopulations: 13 sequences from DaGuiHu (DG) with length of 608 bp and the other 166 sequences from nine other subpopulations with length of 658 bp. The mean nucleotide composition was adenine 23.7% (22.9% to 24.5%), thymine 35.7% (34.5% to 37.0%), cytosine 17.3% (16.3% to 18.5%), and guanine 23.3% (22.3% to 24.0%). The AT-rich COI segment contained 59.4% of nucleotides as A + T. Based on 608 bp comparison, the average ratio of transition/transversion was 3.38 and not approximated to saturation. The 608 bp of COI segments were composed of 453 sites without variation and 155 polymorphic sites. Out of 155 polymorphic sites, 123 were parsimony informative sites and the other 32 were singleton variable sites. The variable sites included 21 on the first codon (15.4%); the second codon is without variable site, and variable sites on third codon are 115 (86.6%).

Among the 10 subpopulations, 179 individuals belonged to 94 haplotypes (accession number AB921860-AB921953); the Hd was 0.966 ± 0.00006 (range 0.990 to 0.131) and π was 0.0782 (range 0.084 to 0.0002) (Table 2). Each subpopulation had unique haplotype that was not shared with other subpopulations. The haplotype diversity of BaiShiChi (BS) (13 haplotypes) and ChiTsaiHu (CT) (14 haplotypes) are 0.990 ± 0.028 higher than other subpopulations, and their nucleotide diversity was 0.0072 and 0.0084, respectively. The haplotype diversity of JiaLuoHu (JL) (three haplotypes) and TaiwanChi (TC) (two haplotypes) populations was much lower than other subpopulations, with haplotype diversity of 0.131 and 0.133 and nucleotide diversity of 0.0003 and 0.0002, respectively.

Population genetic structure of N. tumidus

Among the 10 subpopulations, ranges of genetic differentiation indices (Fst) between each subpopulation pair were from 0.334 to 0.997 (with 3 subpopulation pairs less than 0.9 and mean of 0.948), which indicated high differentiation between the subpopulations. According to Fst, the mean number of genes (Nm) exchanged between each subpopulation pair was 0.03 (range 0 to 1), and between the closest subpopulation, BS and TC only had one copy of gene exchanged from one generation. Based on Kimura’s two-parameter model estimation, the range of intersubpopulation genetic distance between each pair was from 3.3 ± 0.7% to 14.8 ± 1.8%, and the intrasubpopulation genetic distance between each individuals was less than 1% (Table 3). Across different species and genera, the genetic distance, e.g., 27% to 30% between N. tumidus and outgroup Neodiaptomus schmackeri and 23% to 36% between N. tumidus and outgroup Mongolodiaptomus birulai, was much larger than interpopulation differences.

Table 2 Sample size, haplotype number, Hd, π, and polymorphic sites of populations of N. tumidus

| Site (sample size) | No. of haplotype (accession number) | Haplotype diversity (Hd) ± SD | Nucleotide diversity (π) | Polymorphic site (S) | Singletons |
|-------------------|-----------------------------------|-------------------------------|-------------------------|---------------------|-----------|
| SL (15)           | 10 (AB921916-AB921925)            | 0.95 ± 0.040                  | 0.0042                  | 11                  | 5         |
| DR (18)           | 5 (AB921908-AB921912)             | 0.60 ± 0.113                  | 0.0013                  | 5                   | 4         |
| JL (30)           | 3 (AB921913-AB921915)             | 0.31 ± 0.082                  | 0.0003                  | 3                   | 3         |
| TC (15)           | 2 (AB921952-AB921953)             | 0.13 ± 0.112                  | 0.0002                  | 1                   | 1         |
| BS (15)           | 13 (AB921860-AB921872)            | 0.99 ± 0.028                  | 0.0072                  | 21                  | 14        |
| TL (15)           | 11 (AB921941-AB921951)            | 0.95 ± 0.040                  | 0.0071                  | 17                  | 8         |
| CT (15)           | 14 (AB921887-AB921900)            | 0.99 ± 0.028                  | 0.0084                  | 29                  | 19        |
| JM (18)           | 14 (AB921873-AB921886)            | 0.94 ± 0.003                  | 0.0045                  | 21                  | 22        |
| SS (25)           | 15 (AB921926-AB921940)            | 0.95 ± 0.001                  | 0.0042                  | 18                  | 14        |
| DG (13)           | 7 (AB921901-AB921907)             | 0.81 ± 0.113                  | 0.0068                  | 19                  | 11        |
| Total (179)       | 94 (AB921860-AB921953)            | 0.97 ± 0.000                  | 0.0078                  | 155                 | 32        |
The genetic distance (GD) between each subpopulation pair was correlated significantly with geographical distance (ID) (GD = 0.000395 ID + 0.062, R = 0.62, F = 26.4; p < 0.000). Geographical distance of more than 100 km separation can reach more than 10% genetic difference. Based on the result of neutrality test (Table 4), the genetic variation of whole populations was under the random effect of mutation or genetic drift (Tajima's D). Geographical distance of more than 100 km separation can reach more than 10% genetic difference.

### Table 3 Pairwise comparisons of genetic distance (below diagonal and diagonal) and standard error (above diagonal) between populations of *N. tumidus*

|         | SL (%) | DR (%) | JL (%) | TC (%) | BS (%) | TL (%) | CT (%) | JM (%) | SS (%) | DG (%) |
|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| SL      |        | 1.7    | 0.7    | 1.6    | 1.6    | 1.6    | 1.7    | 1.5    | 1.8    | 1.7    |
| DR      | 11.4   |        | 1.6    | 1.2    | 1.5    | 1.5    | 1.6    | 1.4    | 1.5    | 1.5    |
| JL      | 3.3    | 11.8   |        | 0.0    | 1.6    | 1.5    | 1.5    | 1.6    | 1.7    | 1.8    |
| TC      | 11.7   | 7.1    | 11.8   |        | 0.8    | 0.8    | 0.9    | 0.8    | 1.6    | 1.8    |
| BS      | 12.6   | 9.8    | 11.4   | 4.4    |        | 0.7    | 0.3    | 0.5    | 0.9    | 1.5    |
| TL      | 12.1   | 9.8    | 11.1   | 4.0    | 1.1    |        | 0.7    | 0.5    | 0.9    | 1.5    |
| CT      | 13.3   | 11.1   | 12.4   | 5.2    | 2.0    | 2.0    |        | 0.9    | 1.0    | 1.6    |
| JM      | 11.6   | 8.8    | 12.0   | 4.3    | 5.7    | 5.6    | 6.7    |        | 0.5    | 1.6    |
| SS      | 14.7   | 9.5    | 14.1   | 11.4   | 11.3   | 11.3   | 11.5   | 12.6   |        | 11.3   |
| DG      | 14.5   | 11.0   | 14.3   | 13.3   | 13.4   | 14.0   | 14.8   | 14.1   | 7.5    |        |

The genetic distance was calculated using Kimura two-parameter model.

### Table 4 Neutrality tests of each population of *N. tumidus*

| Population | Tajima's D | Fu and Li's D* | Fu and Li's F* |
|------------|------------|----------------|----------------|
| SL         | -0.98      | -0.76          | -0.94          |
| DR         | -1.27      | -1.92          | -2.00          |
| JL         | -1.73      | -2.69*         | -2.80*         |
| TC         | -1.16      | -1.43          | -1.54          |
| BS         | -1.10      | -1.60          | -1.69          |
| TL         | -0.62      | -0.80          | -0.86          |
| CT         | -1.78      | -1.71          | -2.00          |
| JM         | -2.34**    | -3.37**        | -3.56**        |
| SS         | -1.80      | -2.66*         | -2.80*         |
| DG         | -1.39      | -1.02          | -1.27          |
| Total      | 1.05       | 0.32           | 0.80           |

*p < 0.05; **p < 0.01.

### Analysis of phylogeographical relationships

Phylogeographical relationships between each subpopulation pair can be divided into four phylogeographical groups based on ML analysis, MP analysis, and BI. Three analysis methods reach similar dendrograms (Figure 2). Statistic support (ML/MP/BI) on each node was represented by bootstrap values with 1,000 replicates. Outgroups were *Mongolodiaptomus birulai* (MB) and *Neodiaptomus schmackeri* (NS). The northern phylogeographical group included SongLuoHu (SL) and JL (supported value, ML 100%; MP 99%; BI 97%), in both lakes located in northern Taiwan but of different mountain system. DuRongTan (DR) is a unique population among all subpopulations and is located near the northern group geographically.

The major difference of the relationship between phylogeographical groups derived by ML, MP, and BI is the linkage of the northern group. The northern group of ML and MP tree was an independent clade (supported value 100% and 99%), and BI analysis result with northern group as a base of tree could not divide it into an independent clade.

The minimum spanning network analysis showed no significant result for all the subpopulations. Of all subpopulations, only the genetically closely related ones, BS, TL, and CT, can share a significant minimum spanning network with BS as the expanding center. These
Figure 2 Maximum likelihood tree based on COI partial sequence variation among all haplotypes of *N. tumidus*. 
three subpopulations, geographically from north to south and with less than 10 km apart between two neighboring subpopulations, are all in the field of the Yushan bamboo, *Yushania niitakayamensis*, of the Central Mountain Range.

### Analysis of molecular variance and molecular clock

According to the four phylogroups derived by ML, MP, and BI methods, the analysis of molecular variance between each pair of subpopulation indicated that 60.35% of the genetic variation belonged to intergroup difference, 36.07% variation to the population within group, and 3.58% variation within the populations (Table 5). The genetic differentiation was significant among phylogroups ($\Phi = 0.603$, $p < 0.0001$), between each subpopulation within a phylogroup ($\Phi = 0.964$, $p < 0.0001$), and within a population ($\Phi = 0.909$, $p < 0.0001$).

By analysis of pairwise differences among all DNA sequences in a population, the mismatch distribution indicated that JL, TW, and DR subpopulations are still in the state of range expansion (Harpending’s raggedness index = 0.65, 0.56, 0.12) (Table 6). The other subpopulations are in stationary state (Harpending’s raggedness index less than 0.08), and the mismatch distribution as single peak indicated ancient separation and expansion of population. The mismatch distribution of $R^2$ statistics that was closer to 0 indicated that the population process matched with the predicted expansion model with big population size.

We use the gene diversification rate of marine *Sesarma* (1.66% per million year) and terrestrial *Sesarma* (2.33% per million year) as a baseline to estimate the molecular clock of *N. tumidus* evolution. The diversification between DG-DR and SL-JL was the oldest, at 2.8 to 3.9 MYA, DG and DR at 2.2 to 3.0 MYA, and separation of SL-JL and central group at about the same time of first diversification between DG and SL.

### Discussion

The moisture from Pacific Ocean and South China Sea intercepted by mountain plays a crucial role of water supply for lakes or ponds on forest edge or higher than timberline. Even in long-time dry season without rain, the shrank water body has still water enough for aquatic organisms year after year. *N. tumidus* lives in highland waters with water temperature reaching 4°C during the 4- to 5-month cold period and not higher than 15°C throughout the year. Based on our laboratory observation, its individuals still can actively swim when the water temperature was subzero prior to ice formation. During our transportation of samples from collecting sites in the mountain to laboratory, they can survive in wet mud with limited water for many days under low temperature. Most lakes and ponds of our sampling sites are acidic, with pH value less than 6 (Chen and Wang 1997), mainly due to decomposition of organic matter and acid rain. The most acidic waters were in SL, JL, and CT. DG has the largest water volume and a pH value close to 7. The aquatic organisms that live in high elevation and shallow water face another challenge, the UV radiation. The red color in *N. tumidus* was from carotenoid pigments deposited in oil drop and protects the animal from UV damages (Hairston 1980). A relatively harsh abiotic environment, the highland water bodies are devoid of fish and other larger predators and, therefore, provide an environment with relatively low predation pressure. Consequently *N. tumidus* is able to maintain a high population density in these waters throughout the year. During our survey, our plankton net was frequently clogged by the reddish copepods in any season.

Mitochondria DNA evolved in rapid way; the gene order of copepods is very different from other invertebrates (Machida et al. 2002; Rawson and Burton 2006). Some studies also present high variation of COI sequence in interpopulation level in different geographical scales (Lee 1999; Rawson and Burton 2006; Chen and Hare 2008). These properties of mitochondria gene may explain why Folmer’s primers fail to work on DG subpopulation.

Based on the water volume filtered by plankton net, we could find 10 or even more individuals of *N. tumidus* per liter of water and estimated that each subpopulation

### Table 5 Analysis of molecular variance (AMOVA) of *N. tumidus* among different populations

| Source of variation | df | Percentage of variation | $\Phi$ value | $p$ value |
|---------------------|----|-------------------------|-------------|----------|
| Among groups        | 3  | 60.35                   | 0.60346     | <0.0001  |
| Among populations within groups | 6  | 36.07                   | 0.96417     | <0.0001  |
| Within populations  | 169| 3.58                    | 0.90964     | <0.0001  |

*df*, degrees of freedom.

### Table 6 Analysis of pairwise differences among all DNA sequences in a population and mismatch distribution for each population

|                  | Harpending’s raggedness index ($r$) | Ramos-Onsins and Rozas ($R^2$) statistic |
|------------------|-------------------------------------|-----------------------------------------|
| SL               | 0.074                               | 0.091                                   |
| DR               | 0.123                               | 0.107                                   |
| JL               | 0.651                               | 0.132                                   |
| TC               | 0.556                               | 0.249                                   |
| BS               | 0.039                               | 0.085                                   |
| TL               | 0.036                               | 0.110                                   |
| CT               | 0.044                               | 0.068                                   |
| JM               | 0.041                               | 0.067                                   |
| SS               | 0.056                               | 0.067                                   |
| DG               | 0.074                               | 0.093                                   |
size to be over millions and even billions. Grant and Bowen (1998) suggested that if $H_d > 0.5$ and $\pi > 0.005$, the population was stable with a huge population size. The whole population of *N. tumidus* in Taiwan with $H_d = 0.966 > 0.5$ and $\pi = 0.0078 > 0.005$ fits with this category. The subpopulations such as BS, TL, CT, and DG also have $H_d > 0.5$ and $\pi > 0.005$. The subpopulations SL, DR, JM, and SS with $H_d > 0.5$ and $\pi < 0.005$ might, according to Avise (1994) and Grant and Bowen (1998), experience bottleneck effect first, then accumulated genetic variation and increased haplotype diversity but has limited mutation to increase nucleotide diversity through expansion of population. The subpopulations JL and TW with $H_d < 0.5$ and $\pi < 0.005$ might experience long-term bottleneck effect and founder effect to expedite evolution without accumulating too much variation even with large population size.

Among our sampling sites, only two were connected to river, DG as one of the headwaters of LaoNong River at southern Taiwan and SL as one of the headwaters of Nan-Shih-Si River, upstreams of Tanshui River system at northern Taiwan. Except DG and SL, other sampling sites were fully isolated on the mountain ridge. The river system may be a two-way traffic for some aquatic animals, such as fishes and aquatic insects, but a one-way traffic (down the stream) for zooplanktons. Geographical isolation, limited dispersal ability, and low gene flow between subpopulations have been the major cause of the independent evolution among subpopulations. The most geographical distant group between DG-SS and SL with genetic distance greater than 14% is subjected to different species level as suggested by many authors for different crustacean groups (Adamowicz and Purvis 2005; Adamowicz et al. 2007; Costa et al. 2007; Markow and Pfeiler 2010).

According to Bofkin and Goldman (2007), the evolution rate of the second codon was much lower than the third codon for most families. Base on the codon redundancy, the second-codon position is the most functionally constrained; any change to the second-codon position causes a nonsynonymous change in the coding sequence. The third-codon position is the least functionally constrained. This may explain why the COI coding sequence of *N. tumidus* has no second-codon variation. In the COI coding sequence, *N. tumidus* has 86.6% variable sites on the third codon, and the translated AA sequences only have less than 1.2% differences between each pair of subpopulations, similar to their limited difference between external morphology (Young and Shih 2011). The diversification of external morphology less than their COI sequence may reflect their harsh living environment with similar abiotic factors such as low temperature, acidity of water, and high UV radiation. The genetic differences based on COI sequences could be an appropriate evidence for long-term isolation between each population without gene flow. If we justify the species identification only by sequences differences without morphological comparisons, some subpopulations will treat as a subspecies or a new species.

JL subpopulation with genetic differentiation deviates from neutral state based on neutrality test (Tajima's $D$, Fu and Li's $D^*$ and Fu and Li's $F^*$) and still in the state of range expansion as indicated by Harpending's raggedness index and Ramos-Onsins and Rozas statistic. According to the satellite image, JL was a group (11 pools) of water bodies in different sizes which may or may not connect each other. We could only sample two of them, and some pools are even untouched by human until now. The population ecology of these sites still needs more study to understand its relationship between pools. This site also maybe the best place to study the metapopulation process of zooplankton without frequent human disturbances.

The large genetic divergence found among subpopulations of *N. tumidus* suggests that these subpopulations have begun their diversification about 2.2 to 3.9 MYA, matching the geological formation of mountain system in Taiwan about 500 to 300 MYA (Teng 1990; Liu et al. 2000). Based on the geological formation of mountain system in Taiwan, the landmass was first lifted in the north then extended southward (Suppe 1981), and the expanding population of *N. tumidus* should follow the same pattern of geological formation. The phylogeographic relationships in this study also represent the history of metapopulation evolution. The phylogenetic tree morphology indicates that the northern and central phylogeographic group shared a common ancestor and their sister group, DR, shared a common ancestor with the southern phylogeographic group. DR population might be the descendants of an ancient population which was genetically close to DG and SS. We speculate that diversification of the *N. tumidus* population began at the formation of Xueshan Range. One lineage of this ancient population extended southward along the Central mountain ridge and another lineage northward along Xueshan Range. Prior to the end of the last glacial period, extinction of subpopulations occurred in high elevation site, and Xueshan Range and the south end of the central mountain became a refugium for *N. tumidus*. After the glacial epoch, the present subpopulations at Central Mountain Range, excluding DG and SS, originated from north end of Xueshan. Subpopulations DG and SS were not disperse northward, and all the central area was inhabited by populations dispersed southward from the northern refugium. Some studies of plants with limited mobility also consider Xueshan Range and south end of central mountain as a refugium during the Ice Age (Hwang et al. 2003; Cunninghami konishii;
Huang et al. 2004: *Trochodendron aralioides*; Cheng et al. 2006: *Castanopsis carlesii*).

The taxonomic status of *N. tumidus* and their possible relationships with *Neutrodiaptomus formosus* has been discussed by Young and Shih (2011). The close-related *N. formosus* is an endemic species of Japan that lives in lakes with low water temperature. Some other *Neutrodiaptomus* could be found in more northern part and as a dominant species (Kurenkov 1970: *Neutrodiaptomus angustilobus*, Kamchatka). *Neutrodiaptomus* in China also not a common species, it could found only in some isolated lake on remote area (Shen et al. 1979). The distribution scenario of *Neutrodiaptomus* is more or less like the distribution of *N. tumidus* subpopulation in Taiwan. Our study only can unravel part of history of *N. tumidus* just in Taiwan; the origin of this species in Taiwan may have more ancient phylogeographical story connected to other *Neutrodiaptomus* species living in Japan and Eurasia continent.

**Conclusions**

In this study, we found significant genetic divergence between each subpopulation of *N. tumidus* isolated in high mountain ponds or lakes. This result indicated the low dispersion ability of planktonic copepods with limited gene flow between each subpopulation for a long time. The formation of these in the water body and maintaining a high population density for a long period of time was a great mystery and needed to be studied more. The phylogeographical relationship between each subpopulation indicated by COI gene sequences presents multiple divisions along chronological scale. The oldest subpopulation is located at northern Taiwan, which shared the close relationship with the southern subpopulation by early division. Both old lineages were all located at lower elevation considered to be the *refugium* during a glacial period: the high-latitude northern population at about 1,500 m and the low-latitude southern population up to 2,000 m. Subpopulation colonization from north to south on higher elevation at about 3,000 m happened more recently with complicated division pattern. More geological studies for lakes and ponds on high mountain range about their formation times are necessary to match the pattern of phylogeographical relationship of *N. tumidus*. Some other invertebrates live in the same sites with different dispersion ability could also be candidates for study in the future to result more solid facts about this page.

**Competing interests**
The authors declare that they have no competing interests.

**Authors’ contributions**

SSY designed the experiment, collecting samples, data analysis and drafted the manuscript. YYL and MYL carried out the molecular genetic studies, participated in the sequence alignment and phylogenographical analysis. All authors read and approved the final manuscript.

**Acknowledgements**

We thank Dr. Chang-tai Shih for his kind help and encouragement. Deep into the remote mountain region of Taiwan, we got a lot help from many friends with precision guiding and heavy loading of food and sampling gear. Without them we never could walk so far on rugged trial and reach many beautiful ponds and lakes hidden in the valley and forest. Our thanks also extend to A-Dao U-Pa-S, a great man of Atayal tribe; who we could always find warm help about the mountain.

**Author details**

1Department of Applied Science, National Hsinchu University of Education, 521, Nan-Da Road, Hsinchu city 30014, Taiwan. 2National Applied Research Laboratories, Taiwan Ocean Research Institute, 3F, 106, Hoping Road, Section 2, Taipei 10622, Taiwan.

Received: 15 November 2013 Accepted: 16 April 2014

**References**

Adamowicz SJ, Purvis A (2005) How many branchiopod crustacean species are there? Quantifying the components of underestimation. Glob Ecol Biogeogr 14:455–468

Adamowicz SJ, Marque SM, Hebert PON, Purvis A (2007) Molecular systematics and patterns of morphological evolution in the Centropagidae (Copepoda: Calanoida) of Argentina. Biol J Linn Soc 90:279–292

Avise JC (1994) Molecular Marker: Natural History and Evolution. Chapman and Hall, New York, p 511

Bokfin L, Goldman N (2007) Variation in evolutionary processes at different codon positions. Mol Biol Evol 24(2):513–521; doi:10.1093/molbev/msl178

Bohonak AJ, Jenkins DG (2003) Ecological and evolutionary significance of dispersal by freshwater invertebrates. Ecol Lett 6:783–796

Chen G, Hare MP (2011) Cryptic ecological diversification of a planktonic estuarine copepod, Acartia tonsa. Mol Ecol 17:1451–1468

Chen CT, Wang BJ (1997) The lakes and reservoirs of Taiwan. Bor-Hae-Tarng, Taipei, p 504 (in Chinese)

Cheng YP, Hwang SY, Chiou WL, Lin TP (2006) Allozyme variation of populations of *Castanopsis carlesii* (Fagaceae) corroborating diversity centers and potential divergence areas in Taiwan. Ann Bot 98:601–608

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9(10):1567–1600

Costa FG, deWard JR, Bouillon J, Ratnasingham S, Dooch RT, Hajibabaei M, Hebert PON (2007) Biological identifications through DNA barcodes: the case of crustacean. Can J Fish Aquat Sci 64:272–295

Excoffier LGL, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299

Fu Y-X, Li W-H (1993) Statistical tests of neutrality of mutations. Genetics 133:693–709

Grant W, Bowen BW (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. J Hered 89:415–425

Hastion NG, Jr (1980) The vertical distribution of diaptomid copepods in relation to body pigmentation. In: Kerfoot WC (ed) Evolution and ecology of zooplankton communities. University Press of New England, Hanover, New Hampshire, pp 98–110

Hanski I (1998) Metapopulation dynamics. Nature 396(5):41–49

Hanski I (1999) Metapopulation ecology. Oxford University, Oxford, p 309

Harpending H (1994) Signature of ancient population growth in a low resolution mitochondrial DNA distribution. Hum Biol 66:591–600

Hart DL, Clark AG (2007) Principles of population genetics, 4th edition. Sinauer Associates Inc., Sunderland, p 652

Hendler CM (2006) Rafting and refugia in North Atlantic amphipods: understanding the effects of Pleistocene climate change on low-dispersal marine species. PhD Dissertation. Duke University, Durham, NC

Huang SF, Hwang SY, Wang JC, Lin TP (2004) Phylogeography of *Trochodendron aralioides* (Trochodendraceae) in Taiwan and its adjacent areas. J Biogeogr 31:1–9

Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. Genetics 132:583–589
