Population genetic structure and migration pattern of *Nilaparvata lugens* (Stål.) (Hemiptera: Delphacidae) populations in India based on mitochondrial *COI* gene sequences

Guru-Pirasanna-Pandi Govindharaj[1,*], Jaipal Singh Choudhary[2], Aashish Kumar Anant[1], C. Parameswaran[1], G. Basana-Gowda[1], Totan Adak[1], P. Paneerselvam[1], M. Annamalai[1], Naveenkumar Patil[1] and Prakash Chandra Rath[1]

[1]Division of Crop Protection, ICAR-National Rice Research Institute, Cuttack 753 006, India
[2]ICAR-Reseach Complex for Eastern Region, Farming Systems Research Centre for Hill and Plateau Region, Ranchi 834 010, India

Despite the economic and ecological impact of the brown planthopper, *Nilaparvata lugens* infestation associated with rice cultivation in India, studies on its genetic structure are lacking. Hence, the present study was conducted to assess the genetic variability of *N. lugens* in India. The study evaluated the diversity in *N. lugens* populations using mitochondrial cytochrome oxidase subunit I gene sequences from India, and compared them with the Bangladesh, China and Japan populations. In all, 47 unique haplotypes were identified and the haplotype number varied from 6 to 18 in the sampled populations. Genetic diversity indices like nucleotide diversity (0.004), average number of nucleotide differences (1.98), haplotype diversity (0.667) and haplotype number (47) of *N. lugens* populations from India revealed a low level of genetic diversity. A highly significant negative correlation of the demographic history of *N. lugens* populations along with no significant sum of square deviations indicated possible recent expansion of the brown planthopper in India. A non-significant correlation in isolation pattern by distance results indicated that geographic barriers present in the country are not sufficient for genetic differentiation among *N. lugens* from different migratory populations. In this study, the genetic diversity of *N. lugens* populations from India is compared with other Asian populations.

Keywords: Brown planthopper, genetic variability, migration pattern, mitochondrial gene sequences, rice cultivation.

RICE is the most prominent food for millions of people and plays a vital role in the food security of many nations. It occupies the second position in the world both in area and production. Brown planthopper (BPH), *Nilaparvata lugens* (Stål.) is a serious pest of rice and widely dispersed in most of the rice-growing ecologies[1,2]. This pest is a monophagous phloem-feeder of rice causing wilting and hopper burn. The infestation and hopper burn occurrence lead to yield losses of 70%–100% (refs 2, 3). Moreover, *N. lugens* plays a major role in grassy stunt and ragged stunt virus transmission to rice plants[4]. At present, *N. lugens* infestation in rice is reported from almost all major rice-cultivating states of India. BPH infestation and survivability prevail throughout the year in countries like Bangladesh, Philippines, Vietnam and Southern India, whereas its survival in the winter months has not been reported in parts of China, Japan, Korea and northern India[5–7]. Hence the migration of *N. lugens* from other parts of India to its north during the next rice-growing season is the only mode of regular BPH infestation.

Many Asian countries have reported severe damage from *N. lugens* in the last decade[8]. There are four *N. lugens* biotypes present globally, among which biotype 4 is the most destructive, prevailing in South Asian countries[9]. *N. lugens* is thermophilic in nature and has the tendency of long-distance migration. The characteristics of quick proliferation combined with strong migration ability makes this the most dreaded pest in the Indian subcontinent, where it causes severe yield loss and in some places total failure as well[2,3]. However, the development of biotypes against resistant sources is also a major concern for BPH management. As a result, effective *N. lugens* management strategies necessitate a thorough understanding of the biotypes of the migrant population as well as the level of resistance/susceptibility of rice varieties to biotypes[10–12]. Additionally, a critical aspect for the successful development of management strategies for this pest depends on the knowledge of population structure and genetic variability between and within populations present in distinct geographical locations with natural isolation.

*For correspondence. (e-mail: guru.g@icar.gov.in)

The analysis of mitochondrial DNA-based molecular markers has been widely used to decipher the intra- and inter-specific relationships, population structure, non-recombinant pattern, gene flow, copy numbers, the pattern of genetic variation, evolution rate, population size and simple maternal inheritance[13–17]. Various mitochondrial and nuclear genes like mitochondrial cytochrome c oxidase subunit I (*COI*), internal transcribed spacers (*ITSI*), *ITSII* and nuclear 12S-16S–18S ribosomal RNA genes have been widely used as genetic markers to differentiate intra and inter-species populations[15–18]. *COI*-based genetic variation was found to have useful information regarding planthopper evolution[19]. Population structure, gene flow and genetic variation would be helpful in designing and executing to management programmes of migratory insect pests on a large scale[20].

Despite the economic and ecological impact of *N. lugens* infestation associated with rice cultivation in India, studies on its population and genetic structure are lacking. Earlier reports suggest that adaptation differences of planthopper population among the localities depends on local environmental conditions prevails. In this background, the present study was conducted to determine the genetic diversity and population structure of *N. lugens* in India and compare them with other Asian populations on the basis of *COI* gene sequences.

Material and methods

Collection of N. lugens sample

N. lugens adults were fetched from different locations in India during 2017–19, which represent different climatic zones of the country. Sampling sites covered the major endemic hot spots of *N. lugens* infestation and rice cultivating areas of India[1,2,9] (Table 1 and Supplementary Figure 1). An aspirator was used to collect adult bugs from the paddy fields of their respective locations. Morphological characterization was done according to the standard method[21]. After identification, adults were stored in 95% ethanol at –20°C for further processing.

DNA extraction

DNA of two hind legs from each identified adult of *N. lugens* was extracted using the cetyltrimethyl ammonium bromide method[17]. In brief, sterile micro pestle was used to grind the preserved specimen of *N. lugens* in 500 µl of CTAB-solution (2% w/v) and incubated for 1 h at 65°C in a water bath. Further, 500 µl of 24 : 1 : 1 v/v ratio of phenyl : chloroform : isoamyl alcohol (PCI) was mixed and centrifuged for 10 min at 12,000 rpm, and the aqueous part was shifted to fresh micro tubes. Pre-chilled isopropanol (450 µl) was added to the aqueous phase to precipitate the DNA, stored at –20°C for 15–20 min and then centrifuged (4°C) at 12,000 rpm for 14 min. The resultant DNA pellets were washed three times with ice-cold ethanol (70%) for 10 min at 12,000 rpm in a refrigerated centrifuge, dried and re-suspended in 100 µl of *Tris*-EDTA buffer. The quantity and quality of the genomic DNA were determined by NanoDrop-One^{C} (Thermo fisher, USA) and 0.8% agarose gel electrophoresis respectively.

PCR amplification of mitochondrial COI gene

Forward (F) and reverse (R) primers with DNA sequences of 5′-GGTCAACAATCATAAAGATATTGG-3′ and 5′-TAAACTTCAGGGTGACCAAAAAATCA-3′ respectively, were used to amplify the targeted mitochondrial gene *COI* (ref. 22). The thermal conditions used for PCR amplification of the *COI* gene were as follows: 3 min initial denaturation at 94°C, followed by 35 cycles of denaturation at 94°C for 45 sec, annealing at 52°C for 45 sec and extension for 1 min at 72°C, with a final extension of 5 min at 72°C. Later, the PCR product was visualized in 2% (w/v) agarose gel using TAE buffer (40 mM *Tris*-acetate, 1 mM EDTA). The amplicon was freeze-dried and custom-sequenced using same *COI* primers (Agrigenome Labs Ltd, India). The sequences were corrected manually and the assembled unique sequences were submitted in GenBank with accession numbers MN560648–MN560765 and MW750696-MW750753.

Genetic data analysis

Forward and reverse sequencing results were combined and the contigs were prepared using the CAP3 sequence assembly program[23]. ClustalW in MEGA ver. 10.2.4 software was used to align the sequences[24]. Dnasp ver. 6.0 was used for unique haplotypes identification[25] and again reconfirmed using ARLEQUIN ver. 3.5 (ref. 26). Further, descriptive statistics details like nucleotide diversity (π), haplotype diversity (*Hd*), number of haplotypes (*H*) and average number of nucleotide differences (*k*) were evaluated using Dnasp ver. 6.0. Non-parametric permutation method (5000 permutations) was used to determine the statistical significance level of the derived indices with ARLEQUIN.

Geographical relationship was established among the haplotypes by drawing the median joining haplotype network of *COI* gene using NETWORK ver. 10.2 (ref. 27). According to the sampling sites, the major clusters of *N. lugens* haplotypes were colour-coded to infer the haplotype evolutionary relationships among the populations across the sampling sites and defined groups. Similarly, ARLEQUIN ver. 3.5 was used to partition total variance in the components within and among populations and defined groups employing the analysis of molecular variance (AMOVA) hierarchical method with 10,000 permutations. The spatial analysis of molecular variance (SAMOVA)

Table 1. Genetic diversity indices of *Nilaparvata lugens* collected from different locations

Country	Location	Sample number	*H*	*Hd*	$K \pm$ SD	$\pi \pm$ SD
China	Hangzhou (HGZ)	7	3	0.667	0.762 ± 0.629	0.002 ± 0.002
	Shanghai (SHG)	3	2	0.667	0.667 ± 0.667	0.002 ± 0.002
	Guizhou (GUZ)	2	1	n/c	n/c	n/c
Bangladesh	Dhaka (BGD)	3	1	n/c	n/c	n/c
Japan	Ibaraki (JPN)	29	12	0.786	1.308 ± 0.841	0.003 ± 0.002
India	Bengaluru (BKN)	7	1	n/c	n/c	n/c
	Sambalpur (SOD)	10	6	0.778	2.156 ± 1.304	0.005 ± 0.003
	Chiplima (CPOD)	13	3	0.378	0.800 ± 0.628	0.002 ± 0.002
	Cuttack (COD)	10	2	0.154	0.154 ± 0.229	0.001 ± 0.001
	Raipur (RCH)	11	3	0.345	0.727 ± 0.585	0.002 ± 0.001
	Bilaspur (BCH)	7	2	0.286	0.286 ± 0.341	0.001 ± 0.001
	Gariaband (GCH)	10	5	0.756	1.667 ± 1.067	0.004 ± 0.003
	Kawardha (KCH)	6	2	0.333	0.333 ± 0.380	0.001 ± 0.001
	Mungeli (MCH)	12	1	n/c	n/c	n/c
	New Delhi (PND)	10	2	0.400	0.800 ± 0.401	0.001 ± 0.001
	Karnal (KHR)	7	2	0.500	0.857 ± 0.682	0.003 ± 0.002
	Ludhiana (LPN)	9	3	0.417	0.400 ± 0.403	0.001 ± 0.001
	Coimbatore (CTN)	18	6	0.718	2.503 ± 1.415	0.005 ± 0.003
	Moncompu (MKL)	11	5	0.667	2.756 ± 1.592	0.006 ± 0.004
	Davangere (DKN)	13	7	0.731	6.410 ± 3.247	0.014 ± 0.008
	Shimoga (SKN)	10	3	0.600	0.667 ± 0.556	0.001 ± 0.001
	Bapatla (BAP)	6	2	0.333	0.333 ± 0.380	0.001 ± 0.001

H, Haplotypes number; *Hd*, Haplotypes diversity; *K*, Average number of nucleotide difference; π, Nucleotide diversity; SD, Standard deviation; n/c, Not calculated.

was completed using SAMOVA 1.0 (ref. 28). The SAMOVA groups formed from 2 to 10 was compared among groups based on fixation indices with 1000 permutations.

Tajimas' *D* and Fu's *Fs* were determined for neutrality among populations and defined groups[29,30]. DnaSP was used to calculate the demographic history with the mismatch distribution method. Further, this analysis (mismatch distribution) was also used to test the rapid population expansion model[31]. The statistics of the raggedness (R_g) index of the observed distributions and sum of square deviations (SSD) between the observed and the expected mismatch were also calculated using Arlequin. All parameters were tested against the expected values under the hypothesis of a recent population expansion based on 1000 bootstrap replicates. Kimura two-parameter method implemented in MEGA software was used to study the genetic distance among population pairs of *N. lugens*. Using Mantel test the relationship between geographic distance matrices (km in log scale) and genetic distance ($F_{st}/(1 - F_{st})$) were calculated to test the isolation-by-distance (IBD) model[32]. Alleles in Space (AIS) was used for Mantel test analysis with 10,000 randomizations[33]. Google Maps Distance Calculator was used to estimate the geographical distance between each populations pair (http://www.daftlogic.com/projects-google-maps-distance-calculator.htm).

To infer asymmetric dispersal and gene flow between different regions, three regions were defined on the basis of geographic location of sampling and distance. The defined groups were as follows: South Asia (BGD, BKN, SOD, COD, CPOD, RCH, BCH, GCH, KCH, MCH, PND, KHR, LPN, CTN, MKL, DKN, SKN and BAP), China (HGZ, SHG and GUZ) and Japan. Estimation of mutation-scaled active immigration rate for entering and leaving each region per generation (M; $M = m/\mu$, where m is the migration rate and μ is mutation rate per generation) and mutation-scaled population size (θ; $\theta = Ne\mu$, where *Ne* is the effective population size) was done using Bayesian inference applied in Migrate version 3.6.4 (ref. 34). Autonomous migrate runs of 20,000,000 generations with various random start seeds were executed to reach the dependability of the results, with the first 10,000 generations discarded as burn-in.

Results

Nucleotide information

A total of 214 sequences of *N. lugens*, including 51 sequences from NCBI Gen Bank, of other Asian countries were chosen for analysis (Supplementary Table 1). The final length of the submitted sequence in NCBI Gen Bank was 664 base pairs. Previous and the present study sequences were aligned, which showed no insertions and deletions (indels). Nucleotide sequence alignment revealed 64 overall variable sites. The total variable sites included 28 two-variants and two three-variants from 30 parsimony informative sites. A total of 34 singleton sites were observed, among which 31 were two-variants and three singletons were three-variants.

Table 2. Genetic diversity and demographic history parameters of *N. lugens* populations

Group	*H*	*Hd*	*K*	π	θ_0	θ_1	τ	Tajima's *D*	Fu's *Fs*	SSD
All	47	0.667	1.997	0.0043	0.04863	99999.0	0.8710	–1.29600	–25.314***	0.1060
South Asia	32	0.487	1.808	0.0039	0.0000	99999.0	0.0000	–2.56526***	–27.749***	0.2832***
China	4	0.742	1.002	0.0022	0.13184	99999.0	1.1660	0.02173	–19.987***	0.0288
Japan	12	0.786	1.307	0.0028	0.01406	99999.0	1.4472	–1.34449**	–28.207***	0.0058

θ_1, Effective population size after expansion; θ_0, Effective population size before expansion; τ, Population expansion time (million years); SSD, Sum of square deviation between expected and observed mismatch distribution under a sudden expansion model; $^{**}P < 0.01$; $^{***}P < 0.001$.

Table 3. Hierarchical analysis of molecular variance (AMOVA) of different populations and groups of *N. lugens*

Source of variation	d*f*	Sum of square	Variance components	Variation (%)	Fixation index
Two-level					
Among population	21	51.69	0.18611 Va	21.50	Φ_{ST} = 0.21501*
Within population	192	130.45	0.67946 Vb	78.50	
Three-level					
Among groups	1	27.60	0.50515 Va	40.58	Φ_{CT} = 0.40581*
Among populations within groups	20	24.09	0.06017 Vb	4.83	Φ_{SC} = 0.08135**
Within populations	192	130.48	0.67946 Vc	54.58	Φ_{ST} = 0.45415**

d*f*, Degree of freedom; $^{**}P < 0.001$; $^{*}P < 0.05$.

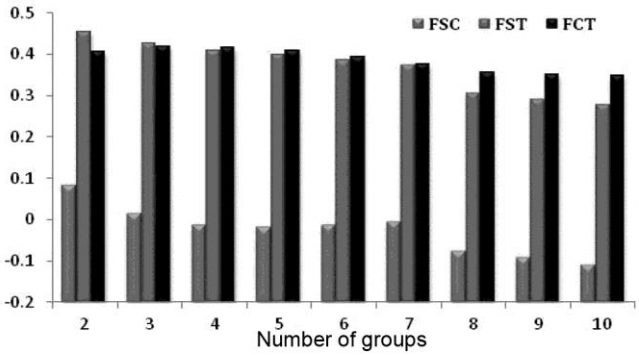

Figure 1. Number of groups defined with respect to fixation indices by SAMOVA analysis.

Population genetic diversity

Genetic diversity analysis of the sequences revealed 47 haplotypes identified from all the studied populations of *N. lugens* (Tables 1 and 2). Only 4 out of 47 haplotypes were common with at least two *N. lugens* populations and the remaining 43 were unique to single population as private haplotypes (Supplementary Table 2). Haplotype H13 was the most shared/common that comprises 134 individuals and accounted for 79.76% of all populations, except the Japanese populations of *N. lugens.*

Tables 1 and 2 present descriptive genetic diversity analysis results. Genetic diversity indices like haplotype diversity (*Hd*), average number of nucleotide diversity (*k*) and nucleotide diversity (π) varied from 0.154 to 0.778, 0.154 to 6.410 and 0.001 to 0.014 respectively. The result indicated that low values of genetic diversity ($Hd < 0.5$; $\pi < 0.5\%$) among *N. lugens* populations in the northern Indian states (New Delhi (PND) and Ludhiana, Punjab (LPN); whereas those in the southern (Coimbatore, Tamil Nadu (CTN); Davanagere, Karnataka (DKN) and Moncompu, Kerala (MKL)) and eastern (Sambalpur, Odisha (SOD) and Gariyaband, Chhattisgarh (GCH)) Indian states showed relatively high values of genetic diversity ($Hd > 0.5$; $\pi > 0.5\%$). Low values of π and high values of *Hd* were reported among the China (HGZ, SHG); Japan; Karnal, Haryana (KHR) and Shivamogga, Karnataka (SKN) populations. Overall, *N. lugens* populations across all the region included in the study had high values of *Hd* (0.667) and low values of π (0.0043).

Molecular variance hierarchical analysis and test to group definition

AMOVA revealed that most genetic variations were intra-population (78.50 at $P < 0.001$) than inter-population (21.50 at $P < 0.001$; Table 3). Significant genetic structure was observed among the populations (Φ_{ST} = 0.21501, $P < 0.05$). SAMOVA supported two-group structure with peak of F_{CT} at $K = 2$ (Figure 1). The three-level AMOVA showed significant structure between two groups (Φ_{CT} = 0.40581, $P < 0.05$) with 40.58% genetic variation (Table 3).

Genetic structure

Pairwise population fixation index (F_{st}) of *N. lugens* was highest (0.297) for four populations, of which two were from India (Bengaluru and Mungeli), and one each from Bangladesh (BGD) and China (Guangzhou). After Bonferroni correction, pairwise F_{st} test among 22 populations exhibited significant differences in 62 out of 214 pairwise comparisons. Pairwise difference of all the studies showed

Table 4. Population-specific F_{st} and pairwise differences among various populations of *N. lugens*

Population	HGZ	SHG	GZH	BGD	JPN	BKN	BAP	BCH	CPOD	COD	CTN	DKN	GCH	KCH	KHR	MCH	MKL	LPN	PND	RCH	SKN	SOD
HGZ (0.253)	–																					
SHG (0.267)	0.519*	–																				
GZH (0.297)	0.329	0.368	–																			
BGD (0.297)	0.408	0.500	0.000	–																		
JPN (0.212)	0.547*	0.494*	0.374*	0.415*	–																	
BKN (0.297)	0.555*	0.708*	0.000	0.000	0.480*	–																
BAP (0.278)	0.453*	0.460*	–0.304	–0.153	0.463*	0.027	–															
BCH (0.281)	0.528*	0.675*	0.000	0.000	0.469*	0.000	–0.180	–														
CPOD (0.249)	0.377*	0.289	–0.323	–0.184	0.470*	–0.039	–0.027	–0.020	–													
COD (0.287)	0.497*	0.526*	–0.317	–0.174	0.482*	–0.055	0.040	0.023	0.018	–												
CTN (0.138)	0.165*	–0.021	–0.304	–0.153	0.414*	0.012	–0.004	0.011	0.038	0.068	–											
DKN (0.101)	0.155*	0.019	–0.178	–0.061	0.377*	0.074	0.051	0.070	0.106*	0.147*	0.149	–										
GCH (0.196)	0.285*	0.155	–0.250	–0.119	0.451*	0.027	–0.073	–0.048	0.005	0.096*	0.055	0.088*	–									
KCH (0.278)	0.453*	0.460	–0.304	–0.153	0.463*	0.027	–0.000	0.002	–0.086	0.040	0.006	0.051	–0.012	–								
KHR (0.248)	0.340*	0.200	–0.263	–0.090	0.451*	–0.000	–0.011	0.000	0.002	0.060	0.029	0.067	0.035	–0.011	–							
MCH (0.297)	0.578*	0.734*	0.000	0.000	0.491*	0.000	0.124	0.082	–0.019	–0.006	0.063	0.140	0.098*	0.124	0.082	–						
MKL (0.130)	0.187*	0.036	–0.308	–0.170	0.407*	–0.025	–0.035	–0.020	–0.010	0.041	0.045*	0.096*	0.021	–0.054	–0.011	0.036*	–					
LPN (0.273)	0.437*	0.302	0.000	0.250	0.481*	–0.040	–0.006	–0.007	–0.000	0.011*	0.048	0.098	0.043	–0.006	0.018	0.019	0.002	–				
PND (0.273)	0.380*	0.299	–0.290	–0.132	0.452*	–0.040	–0.006	–0.007	–0.000	0.011	0.048	0.098	0.043	–0.006	0.018	0.019	0.002	–0.053	–			
RCH (0.253)	0.333*	0.216	–0.312	–0.166	0.453*	–0.046	–0.027	–0.022	–0.024	0.010	0.013	0.119*	0.043	–0.027	0.005	0.008	–0.014	–0.003	–0.036	–		
SKN (0.257)	0.453*	0.460	–0.304	–0.153	0.449*	0.114	0.091	0.101	0.083	0.163*	0.080*	0.122*	0.103*	0.091	0.087	0.195*	0.049	0.111	0.011	0.087	–	
SOD (0.167)	0.225*	0.080	–0.303	–0.166	0.425*	–0.021	–0.029	–0.015	–0.041	0.044*	0.047	0.097*	0.005	–0.053	–0.005	0.041*	–0.006	0.017	0.017	–0.005	0.060	–

Distance values below the diagonal are obtained by a bootstrap procedure (1000 replicates). Values within parenthesis are population specific F_{st} based on respective sequence.
*Values (distances) shown are significantly different.

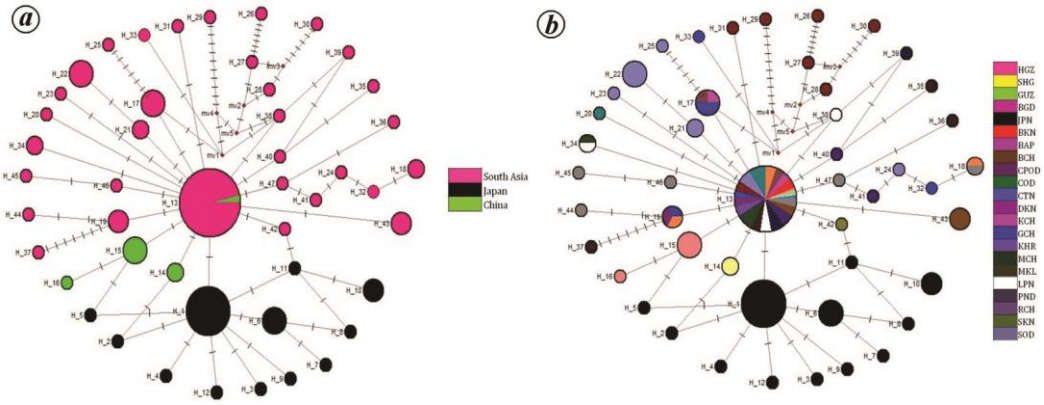

Figure 2. Median-joining network of mtDNA haplotypes of *Nilaparvata lugens*. ***a***, ***b***, Haplotype network with the median-joining of (***a***) grouped population and (***b***) individual population of different countries. Each circle represents a haplotype, and haplotype frequency is denoted by the diameter of the circles. Colours represent the geographic origin of specimens and size indicates the proportion of individuals sampled in different populations within the study area. Median vectors are represented by small red squares.

that the Japanese populations differed significantly from all other Asian populations, including the Indian populations (Table 4). Except the Japanese population, other Asian *N. lugens* populations had non-significant pairwise F_{st} values.

The median-joining (MJ) network was constructed among the haplotypes and two groups were formed (Figure 2 *a* and *b*). One group was the clustering of South Asia and China populations (pink and light green; Figure 2 *a*), which consisted of a large number of exclusive haplotypes and few common haplotypes. H13 was the most common and original/ancestral haplotype found in *N. lugens* populations of the Indian and China populations. Other major haplotypes identified were H1 and H17, which were connected through several mutational steps to widely distributed haplotypes. The MJ network also showed that the Japanese population formed a separate group of haplotypes (H1–H12) which connected to the Chinese population through several mutational steps.

Mantel test was performed predict the isolation pattern by distance, and results revealed a significant positive correlation between pairwise genetic and geographic distances of the *N. lugens* populations studied ($r = 0.412$ at $P < 0.05$). The findings suggest that the populations of *N. lugens* in the studied locations may either have reached equilibrium between migration and genetic drift, or geographic barriers are sufficient for genetic differentiation among *N. lugens* from different populations.

Demographic history and gene flow

To identify any deviation from neutrality, Tajima's *D* and Fu's *Fs* tests were performed using sequence variance in defined groups and all populations. Fu's *Fs* values for the defined groups and all the populations were found significantly negative. When two groups were defined, Tajima's *D* values were significantly negative for the South Asian and Japanese groups, but was not significant in the Chinese group (Table 2). When all samples were considered as one group, neutrality tests (Fu's *Fs* and Tajima's *D*) were statistically significant with insignificant SSD values, denoting *N. lugens* had undergone recent demographic expansion (Table 2). The mismatch distribution plots were left-skewed unimodel mismatch distributions, suggesting that *N. lugens* populations were undergoing population expansion (Figure 3). Non-significant SSD values with relatively small R_g values also supported the neutrality tests of sudden population expansion (Figure 3). The neutrality test values of South Asia indicate that of the *N. lugens* populations experienced recent demographic expansion and are still undergoing expansion.

Population expansion time (τ) in the groups varied from 0.00 to 1.44727. The difference in effective population size before ($\theta_0 = 0.000$–0.13184) and after ($\theta_1 = 9999.000$) expansion indicate large population growth of *N. lugens* (Table 3). When three formed population groups were analysed, effective population size (θ) estimates were constantly low and to the tune of 0.00072 for China to 0.08172 for South Asia which denoted scant gene flows ($\theta \times M$) among three regions (Figure 4). Further, estimates of migration rate between regions were relatively low and somewhat bidirectional from South Asia to China and China to Japan, which ranged from 0.00 to 848.7 (Figure 4). The MJ network analyses also demonstrate that the Japan haplotypes were separate from other groups, indicating restricted gene flow across the different geographically grouped populations over long distances.

Discussion

Genetic diversity analysis in different regions provides valuable insights in the evolution and understanding of migration dynamics of an organism. In the present study mtDNA sequences analysis proved low to high genetic

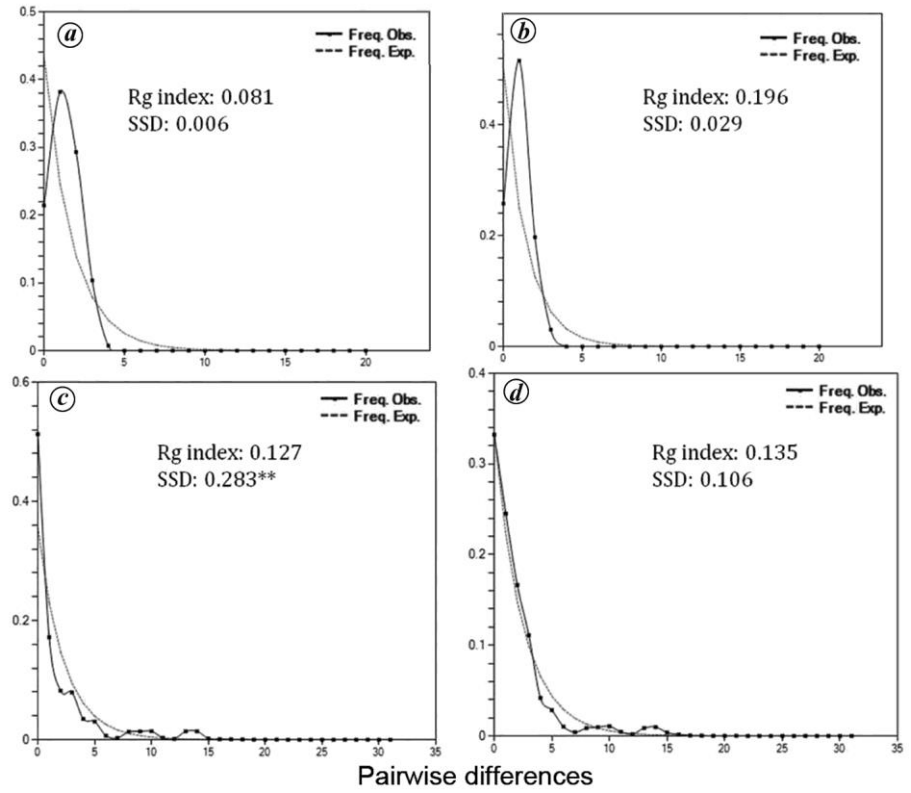

Figure 3. Observed and simulated mismatch distribution in (***a***) Japan, (***b***) China, (***c***) South Asia and (***d***) all samples of *N. lugens* populations. $^{**}P < 0.01$.

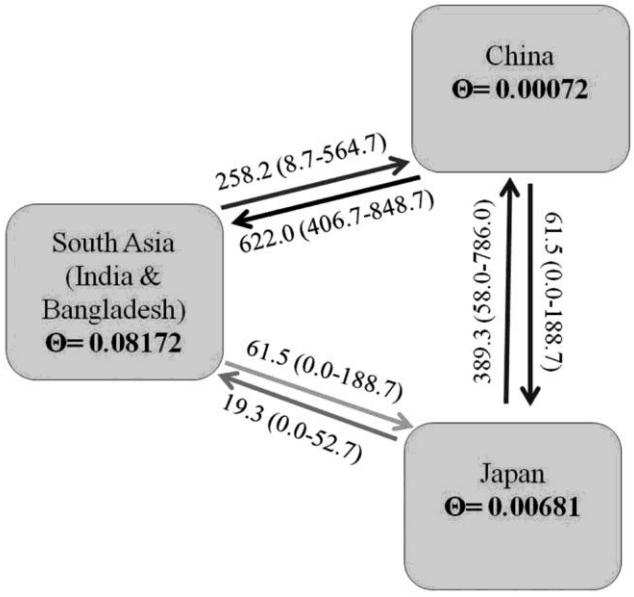

Figure 4. Estimation of migration rate and population size between populations of *N. lugens*. *M* denotes the mutation-scaled migration rate and Φ represents the mutation-scaled effective population size.

diversity of the Indian populations using diversity indices like *Hd*, *k* and π. Sample size did not influence the genetic diversity parameters, thereby suggesting them as appropriate diversity indices for *N. lugens* populations. The results of the present study strongly support that the Indian populations of *N. lugens* are not stable, with no evidence for neutral evolution in the present populations. However, most of Indian populations could have undergone expansion under the conditions of a recent bottleneck or founding effect substantiated with low haplotype ($Hd < 0.5$) and nucleotide ($\pi < 0.004$) diversity values compared to the Japanese and Chinese populations[35]. Besides, some Indian (LPN, KHR and PND) and China populations exhibited high *Hd* and low π values, suggesting that a small proportion of Indian populations might have undergone recent population expansion and divergence between geographically subdivided populations. High genetic diversity values ($Hd > 0.5$ and $\pi > 0.5\%$) of few populations indicated that they are stable with relatively long evolutionary history. The high level of genetic diversity in the southern and eastern parts of India may be due to abundance of suitable host plants (i.e. their throughout the year availability), which might have had little selective pressure in the past for the establishment and dispersal of the population. Earlier studies reported high level of genetic diversity when the species was regarded as throughout annual survival because the ancestors exhibited significantly more genetic diversity than the derived population due to founder effect[36]. Immigrant populations were low in genetic diversity owing to genetic drift and selection pressure, which may be the reason for low genetic diversity in the populations

in North India with very low temperature in winter and non-availability of host plant throughout the year[37,38]. Moreover, long-distance migration was previously reported in *N. lugens* which assisted in fresh colonization and rescue of a species from the status of genetic diversity loss[11]. A similar mechanism was reported earlier from *N. lugens* and *Bactrocera dorsalis*[15,37]. Demographic expansion of *N. lugens* was tested through mismatch distribution analysis with the expected unimodal shape and the results indicated that the population had experienced expansion[15,31]. The MJ reconstructed exhibited a star-like shape among haplotypes, which also in line of typical population expansion model[39]. The sharing of major haplotypes among the Chinese, Bangladeshi and Indian populations indicates recent gene mixing of *N. lugens* of various populations. The migration of *N. lugens* among South Asian populations is also exhibited by non-significant values of pairwise F_{st} and migration rate pattern of populations. The theory of demographic expansion is also in line with an explanation that *N. lugens* has higher migration rate and gene flow had occurred in the relatively recent past through migration[40].

Demographic analysis of all populations revealed significant negative Tajima's *D* and Fu's *F* statistical values, which in turn does not imply a simple model of selective neutrality and bottleneck analysis[41]. The reason for this might be either a recent selection pressure or recent increase in size experienced by the populations[42]. Moreover, the recent population expansion of *N. lugens* was also supported by unimodal mismatch distribution with non-significant SSD values and star-like structure of the MJ network[39]. Raggedness index and SSD were used to test the hypothesis of sudden expansion model[43]. Non-significant values for SSD and relatively small raggedness index advocated sudden expansion of populations in the past[15,16]. Besides, presence of low to medium genetic diversity in the region with weak genetic structure of populations indicated that *N. lugens* foundation may be inferred as a steady and range expansion of population with higher incidence and growth[44].

Further information on *N. lugens* migration route in India is necessary, while an earlier study from other Asian countries suggested that *N. lugens* flight duration varied from 12 to 36 h depending on tropical, subtropical and temperate conditions[7]. In the tropical and subtropical areas *N. lugens* follows a shorter migration distance in the range 4–30 km due to availability of food and other favourable weather conditions[45]. The dispersal pattern in the present study also supports the bi-directional migration of *N. lugens* from South Asia to China and China to Japan, corresponding to the Asian subtropical monsoon, that takes place from the southern to northern direction in the rainy seasons[6,45].

Many unique haplotypes in the Japanese populations which were not detected in Chinese and Indian populations, denoting geographical barriers and habitat quality variation mainly contributed to genetic separation between the Japanese and other Asian populations. Mantel test indicated a significant positive correlation between pairwise genetic and geographic distances of *N. lugens* populations studied ($r = 0.412$ at $P < 0.05$). Results advocate that *N. lugens* populations in the studied area might have either reached an equilibrium between migration and genetic drift or geographic barriers were sufficient for genetic differentiation among *N. lugens* populations from Japan and other Asian countries. This significant positive correlation in isolation pattern by distance was obtained due to inclusion of the Japanese populations. One of the possible reasons for these distinct haplotypes in Japan may be due to evolutionary association with the subspecies of rice cultivar, i.e. China and India mostly have *Oryza sativa indica* cultivar, whereas in Japan *Oryza sativa japonica* cultivar is mostly cultivated. Additionally, different levels of selection pressure and minor level of local extinction might cause phylogeographic breaks without geographic barriers[46], like the populations from Japan and southern India. These unique haplotypes were favoured in local habitation due to selection force; thus phylogeographic breaks could occur in spite of the gene flow. However, after the first report of BPH in the 1950s in India, spread and diffusion paths are largely unknown. The present study will help improve the *N. lugens* mitochondrial DNA database and be an important source for comprehensive phylogeography studies of BPH throughout the world.

Conclusion

Genetic diversity and expansion history were studied using mitochondrial DNA markers (*COI*) for Indian *N. lugens* populations along with those from other Asian countries. Results showed low to medium level of genetic diversity with a weak genetic structure among the *N. lugens* populations across Asia. The present findings will be valuable for better prediction and management of planthopper outbreaks in different regions of the Indian subcontinent in the future. Additionally, they suggest the occurrence of unique haplotypes from the Indian *N. lugens* populations, which would assist in the identification of their migration routes. In a nutshell, a combination of ecological factors (micro-climate and specific topography), historical factors (interglacial cycles) and anthropogenic factors (passive dispersal ability) might have helped form the existing population structure and dispersal routes of *N. lugens*.

Conflict of interest: The authors declare that there is no conflict of interest.

1. Pandi, G.G.P., Chander, S., Pal, M. and Pathak, H., Impact of elevated CO_2 and temperature on brown planthopper population in rice ecosystem. *Proc. Natl. Acad. Sci. India, Sect. B*, 2016, **88**(1), 57–64; doi:10.1007/s40011-016-0727-x.

2. Jena, M. *et al.*, Paradigm shift of insect pests in rice ecosystem and their management strategy. *Oryza*, 2018, **55**, 82–89; doi:10.5958/2249-5266.2018.00010.3.
3. Pandi, G. G. P., Chander, S., Pal, M. and Soumia, P. S., Impact of elevated CO_2 on *Oryza sativa* phenology and brown planthopper, *Nilaparvata lugens* (Hemiptera: Delphacidae) population. *Curr. Sci.*, 2018, **114**(8), 1767–1777; doi:10.18520/cs/v114/i08/1767-1777.
4. Li, S., Wang, H. and Zhoum, G. S., Synergism between Southern rice black-streaked dwarf virus and Rice ragged stunt virus enhances their insect vector acquisition. *Phytopathology*, 2014, **104**(7), 794–799; doi:10.1094/PHYTO-11-13-0319-R. PMID: 24915431.
5. Bottrell, D. G. and Schoenly, K. G., Resurrecting the ghost of green revolutions past: the brown planthopper as a recurring threat to high-yielding rice production in tropical. *J. Asia-Pac. Entomol.*, 2012, **15**(1), 122–140; https://doi.org/10.1016/j.aspen.2011.09.004.
6. Otuka, A., Migration of rice planthoppers and their vectored re-emerging and novel rice viruses in East Asia. *Front. Microbiol*., 2013, **4**, 309; doi:10.3389/fmicb.2013.00309.
7. Hu, G., Lu, M. H., Tuan, H. A. and Liu, W. C., Population dynamics of rice planthoppers, *Nilaparvata lugens* and *Sogatella furcifera* (Hemiptera, Delphacidae) in Central Vietnam and its effects on their spring migration to China. *Bull. Entomol. Res*., 2017, **107**, 369–381; https://doi.org/10.1017/S0007485316001024.
8. EPPO, European and Mediterranean Plant Protection Organization global database, 2021; https://gd.eppo.int/taxon/NILALU/distribution (accessed on 20 December 2021).
9. Anant, A. K. *et al.*, Genetic dissection and identification of candidate genes for brown planthopper, *Nilaparvata lugens* (Delphacidae: Hemiptera) resistance in farmers' varieties of rice in Odisha. *Crop Prot*., 2021, **144**, 105600; https://doi.org/10.1016/j.cropro.2021.105600.
10. Matsumura, M., Takeuchi, H., Satoh, M., Sanada-Morimura, S., Otuka, A., Watanabe, T. and Van, T. D., Species specific insecticide resistance to imidacloprid and fipronil in the rice planthoppers *Nilaparvata lugens* and *Sogatella furcifera* in East and South-east Asia. *Pest Manage. Sci.*, 2008, **64**(11), 1115–1121; doi:10.1002/ps.1641. PMID: 18803329.
11. Matsumoto, Y., Matsumura, M., Sanada-Morimura, S., Hirai, Y., Sato, Y. and Noda, H., Mitochondrial COX sequences of *Nilaparvata lugens* and *Sogatella furcifera* (Hemiptera, Delphacidae): low specificity among Asian planthopper populations. *Bull. Entomol. Res.*, 2013, **103**(4), 382–392; doi:10.1017/S000748531200082X. PMID: 23537548.
12. Naeemullah, M., Sharma, P. N., Tufail, M., Mori, N., Matsumura, M., Takeda, M. and Nakamura, C., Characterization of brown planthopper strains based on their differential responses to introgressed resistance genes and on mitochondrial DNA polymorphism. *Appl. Entomol. Zool*., 2009, **44**, 475–483; https://doi.org/10.1303/aez.2009.475.
13. Rollins, L. A., Woolnough, A. P., Sinclair, R., Mooney, N. J. and Sherwin, W. B., Mitochondrial DNA offers unique insights into invasion history of the common starling. *Mol. Ecol*., 2011, **20**(11), 2307–2317; doi:10.1111/j.1365-294X.2011.05101.x. PMID: 21507095.
14. Wan, X., Liu, Y. and Zhang, B., Invasion history of the oriental fruit fly, *Bactrocera dorsalis*, in the Pacific-Asia region: two main invasion routes. *PLOS ONE*, 2012, **7**(5), e36176; https://doi.org/10.1371/journal.pone.0036176.
15. Choudhary, J. S., Naaz, N., Prabhakar, C. S. and Lemtur, M., Genetic analysis of oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae) populations based on mitochondrial *COX 1* and *NAD 1* gene sequences from India and other Asian countries. *Genetica*, 2016, **144**, 611–623; doi:10.1007/s10709-016-9929-7. PMID: 27699519.
16. Choudhary, J. S., Naaz, N., Lemtur, M., Das, B., Singh, A.K., Bhatt, B. P. and Prabhakar, C. S., Genetic analysis of *Bactrocera zonata* (Diptera: Tephritidae) populations from India based on *COX 1* and *NAD 1* gene sequences. *Mitochondrial DNA A*, 2018, **29**(5), 727–736; https://doi.org/10.1080/24701394.2017.1350952.
17. Pandi, G. G. P. *et al.*, Molecular diversity of *Nilaparvata lugens* (Stål.) (Hemiptera: Delphacidae) from India based on internal transcribed spacer 1 (ITSI) gene. *Curr. Sci*., 2022, **122**(12), 1392–1400.
18. Watanabe, S. and Melzer, M. J., A multiplex PCR assay for differentiating coconut rhinoceros beetle (Coleoptera: Scarabaeidae) from oriental flower beetle (Coleoptera: Scarabaeidae) in early life stages and excrement. *J. Econ. Entomol*., 2017, **110**, 678–682; doi:10.1093/jee/tow299. PMID: 28115497.
19. Noda, H., How can planthopper genomics be useful for planthopper management? In *Planthoppers: New Threats to the Sustainability of Intensive Rice Production Systems in Asia* (eds Heong, K. L. and Hardy, B.), International Rice Research Institute, Philippines, 2009, pp. 429–446.
20. Roderick, G. K. and Navajas, M., Genes in new environments: genetics and evolution in biological control. *Nature Rev. Genet*., 2003, **4**, 889–899; https://doi.org/10.1038/nrg1201.
21. Wilson, M. R. and Claridge, M. F., *Handbook for the Identification of Leafhoppers and Planthoppers of Rice*, CAB International, London, UK, 1991, p. 143; ISBN 0-85198-692-7.
22. Mun, J. H., Song, Y. H., Heong, K. L. and Roderick, G. K., Genetic variation among Asian populations of rice planthoppers, *Nilaparvata lugens* and *Sogatella furcifera* (Hemiptera: Delphacidae): mitochondrial DNA sequences. *Bull. Entomol. Res*., 1999, **89**, 245–253; doi:10.1017/S000748539900036X.
23. Huang, X. and Madan, A., CAP3: a DNA sequence assembly program. *Genome Res*., 1999, **9**(9), 868–877; doi:10.1101/gr.9.9.868. PMID: 10508846.
24. Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S., MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol*., 2013, **30**(12), 2725–2729; doi:10.1093/molbev/mst197. Epub 2013 Oct 16. PMID: 24132122.
25. Librado, P. and Rozas, J., DnaSP V5: software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 2009, **25**(11), 1451–1452; https://doi.org/10.1093/bioinformatics/btp187.
26. Excoffier, L. and Lischer, H. E. L., Arlequin suite ver. 3.5, a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour*., 2010, **10**(3), 564–567; doi:10.1111/j.1755-0998.2010.02847.x. PMID: 21565059.
27. Bandelt, H., Forster, P. and Roehl, A., Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol*., 1999, **16**(1), 37–48; doi:10.1093/oxfordjournals.molbev.a026036. PMID: 10331250.
28. Dupanloup, I., Schneider, S. and Excoffier, L., A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol*., 2002, **11**, 2571e2581; doi:10.1046/j.1365-294x.2002.01650.x. PMID: 12453240.
29. Tajima, F., Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 1989, **123**(3), 585–595; doi:10.1093/genetics/123.3.585. PMID: 2513255.
30. Fu, Y. X., Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 1997, **147**, 915–925; doi:10.1093/genetics/147.2.915. PMID: 9335623.
31. Rogers, A. R. and Harpending, H., Population growth makes waves in the distribution of pairwise differences. *Mol. Biol. Evol*., 1992, **9**(3), 552–569; doi:10.1093/oxfordjournals.molbev.a040727.
32. Mantel, N., The detection of disease and a generalized regression approach. *Cancer Res*., 1967, **27**, 209–220. PMID: 6018555.
33. Miller, M. P., Alleles in space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. *J. Hered.*, 2005, **96**(6), 722–724; https://doi.org/10.1093/jhered/esi119.
34. Beerli, P. and Felsenstein, J., Maximum-likelihood estimation of a migration matrix and effective population sizes in subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. USA*, 2001, **98**(8), 4563–4568; https://doi.org/10.1073/pnas.081068098.

35. Rosetti, N. and Remis, M. I., Spatial genetic structure and mitochondrial DNA phylogeography of argentinean populations of the grasshopper *Dichroplus elongatus*. *PLoS ONE*, 2012, **7**(7), e40807; https://doi.org/10.1371/journal.pone.0040807.
36. Grant, W. S. and Bowen, B. W., Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *J. Hered*., 1998, **89**, 415–426; https://doi.org/10.1093/jhered/89.5.415.
37. Suarez, A. V. and Tsutsui, N. D., The evolutionary consequences of biological invasions. *Mol. Ecol*., 2008, **17**(1), 351–360; doi: 10.1111/j.1365-294X.2007.03456.x. PMID: 18173507.
38. Grapputo, A., Bisazza, A. and Pilastro, A., Invasion success despite reduction of genetic diversity in the European populations of eastern mosquito fish (*Gambusia holbrooki*). *Ital. J. Zool*., 2006, **73**, 67–73; https://doi.org/10.1080/11250000500502111.
39. Slatkin, M. and Hudson, R. R., Pairwise comparison of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, 1991, **129**(2), 555–562; doi:10.1093/genetics/129.2.555.
40. Hereward, J. P., Cai, X., Matias, A. M. A., Walter, G. H., Xu, C. and Wang, Y., Migration dynamics of important rice pest: the brown planthopper (*Nilaparvata lugens*) across Asia-insights from population genomics. *Evol. Appl*., 2020, **13**(9), 2449–2459; https://doi.org/10.1111/eva.13047.
41. Harpending, H. C., Batzer, M. A., Gurven, M., Jorde, L. B., Rogers, A. R. and Sherry, S. T., Genetic traces of ancient demography. *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 1691–1697; https://doi.org/10.1073/pnas.95.4.1961.
42. Slatkin, M., Gene flow in natural populations. *Annu. Rev. Evol. Syst*., 1985, **16**, 393–430; https://doi.org/10.1146/annurev.es.16.110185.002141.
43. Schneider, S. and Excoffier, L., Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, 1999, **152**(3), 1079–1089.
44. Wan, X., Nardi, F., Zhang, B. and Liu, Y., The oriental fruit fly, *Bactrocera dorsalis*, in China: origin and gradual inland range expansion associated with population growth. *PLoS ONE*, 2011, **6**(10), e25238; https://doi.org/10.1371/journal.pone.0025238.
45. Wei, S. J., Genetic structure and demographic history reveal migration of the diamondback moth *Plutellaxylostella* (Lepidoptera: Plutellidae) from the southern to northern regions of China. *PLoS ONE*, 2013, **8**, e59654; https://doi.org/10.1371/journal.pone.0059654.
46. Irwin, D. E., Phylogeographic breaks without geographic barriers to gene flow. *Evolution*, 2002, **56**(12), 2383–2394; https://doi.org/10.1554/0014-3820(2002)056[2383:PBWGBT]2.0.CO;2.

ACKNOWLEDGEMENTS. We thank the Director, ICAR-National Rice Research Institute, Cuttack for support. This study was funded by an ICAR-National Rice Research Institute in-house project.

Received 21 April 2022; accepted 18 May 2022

doi: 10.18520/cs/v123/i3/461-470