

## Systematics, Morphology and Biogeography

# High-level phylogeographic structuring of *Neoleucinodes elegantalis* Guenée (Lepidoptera, Crambidae) in Brazil: an important tomato pest



André V. P. Maia<sup>a</sup>, Cícero Almeida<sup>b,\*</sup>, Kleber R. Santoro<sup>c</sup>, João L. A. Melo<sup>c</sup>, José V. Oliveira<sup>a</sup>, Raul N. C. Guedes<sup>d</sup>, César A. Badji<sup>c</sup>

<sup>a</sup> Universidade Federal Rural de Pernambuco, Departamento de Agronomia, Recife, PE, Brazil

<sup>b</sup> Universidade Federal de Alagoas, Laboratório de Recursos Genéticos, Arapiraca, AL, Brazil

<sup>c</sup> Universidade Federal Rural de Pernambuco, Unidade Acadêmica de Garanhuns, Garanhuns, PE, Brazil

<sup>d</sup> Universidade Federal de Viçosa, Departamento de Entomologia, Viçosa, MG, Brazil

## ARTICLE INFO

### Article history:

Received 15 December 2015

Accepted 8 March 2016

Available online 13 May 2016

Associate Editor: Héctor Vargas

### Keywords:

Insect pest

Phylogeography

Small tomato borer

## ABSTRACT

*Neoleucinodes elegantalis* is an important tomato pest in Brazil, occurring throughout the country and resulting in economic losses in agriculture. In several species, biogeographic studies in Brazil indicate the structuring of populations, following the refuge model, with a split between the populations of the northeast and the southeast regions of Brazil. The objective of this work was to analyze the phylogeography of *N. elegantalis* in Brazil, understanding its population structure and the demographic patterns. Larvae were collected from eight locations throughout Brazil, and the mitochondrial cytochrome c oxidase subunit 1 gene was analyzed. A total of 628 bp in 51 individuals were obtained, showing 12 haplotypes with a haplotype diversity of 0.836. Spatial analysis of molecular variance (SAMOVA) and cluster analysis showed two populations, indicating population structuring between individuals from the northeast (population 1) and southeast (population 2) regions of Brazil. Phylogenetic analysis indicated that the clades corresponding to the groups defined by SAMOVA have a divergence time of 0.2–0.5 million years, suggesting isolation during climatic events and a separation of the two populations coinciding with the predicted refuges to the Atlantic forest.

Published by Elsevier Editora Ltda. on behalf of Sociedade Brasileira de Entomologia. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introduction

The small tomato borer (*Neoleucinodes elegantalis* Guenée, 1984, Lepidoptera: Crambidae) is a pest that has had great economic impact in Brazil, Venezuela, and Colombia (Badji et al., 2003; Picanço et al., 2007). High infestation levels of this pest make the fruits unsuitable for consumption and industrial processing (Badji et al., 2003; Benvenega et al., 2010; Picanço et al., 2007). Losses resulting from the damage can be as high as 90% of the total production (Miranda et al., 2005).

In Brazil, *N. elegantalis* is distributed throughout the country, associated with human migrations in tomato-producing regions. It extends from the cold regions of the southeast to the dry regions of the northeast, particularly in areas of dry seasonal forests (Caatinga). These regions show differences in climate, topography, floristic composition, and demographic events that can model the population structure of *N. elegantalis*. *N. elegantalis* has been

controlled with the use of chemical insecticides and behavioral controls using sex pheromones (Badji et al., 2003). In controls using sex pheromones, population structuring is very important in determining the specificity of the pheromone types. For example, a variation in cuticular hydrocarbons associated with geographical distance (Bonelli et al., 2015) was detected in *Polistes biglumis* (Hymenoptera: Vespidae). In *N. elegantalis*, unpublished data suggest that pheromones have different effects on the populations of the southeast and the northeast of Brazil, indicating a geographical association with pheromone efficiency. These differences may be associated with population structure due to the influence of climate fluctuations as the biogeographic events.

Biogeographical events such as climate variations and changes in habitat are significant factors in explaining the geographical distribution of many species, particularly in climactic events in the Pleistocene. Using paleomodeling, Carnaval and Moritz (2008) showed the phylogeographic center of endemism in Brazil, suggesting distinct centers of endemism for different species, including butterflies. Thus, estimates of the spatial distribution of the populations of *N. elegantalis*, including demographic patterns and historical population parameters, enable an understanding of

\* Corresponding author.

E-mail: [cicerocarlos@pesquisador.cnpq.br](mailto:cicerocarlos@pesquisador.cnpq.br) (C. Almeida).

the dispersion of the species and contribute to pest management. Phylogeographic studies of insect species have been conducted using the cytochrome c oxidase subunit 1 (CO1) region, and it has become a standardized region for dispersion studies. The aim of this study was to analyze the phylogeographic structure of *N. elegantalis* in Brazil using the region of the CO1 gene. The results allow analyze the hypothesis about population structuring resulting from climate changes that may have occurred in the species, as well as contributing to a clearer understanding of distribution patterns of species resulting from successive climatic cycles during the Pleistocene.

## Material and methods

### Sampling and DNA extraction

Larvae were collected in eight locations (51 individuals) distributed from the southeast to the northeast of Brazil (Fig. 1). The samples were stored in 70% ethanol and conditioned at 4 °C. Two of the locations were in the southeast, in the states of São Paulo and Minas Gerais, and the other six locations were in the northeast of Brazil, in dry forest areas in the state of Pernambuco (Caatinga; map shown in Fig. 1A). The DNA was extracted using the CTAB protocol (Doyle and Doyle, 1987), quantified using spectrophotometry, and analyzed for quality using 1% agarose gel.

### Amplification and sequencing

The region of the mitochondrial CO1 gene was amplified using LepF1/LepR1 primers (Hebert et al., 2004). The reactions were performed in a total of 50 µL, containing 5 µL reaction buffer, 1.5 mM MgCl<sub>2</sub>, 0.2 mM dNTP, 1.25 U Taq DNA polymerase, 0.5 µM of each primer, and 100–150 ng of DNA. The amplification was performed with an initial denaturation at 94 °C for 4 min, followed by 30 cycles at 94 °C for 40 s, 55 °C for 35 s, and 72 °C for 1 min, and a final extension at 74 °C for 4 min. The PCR products were amplified using a BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems®) in electrophoresis in a 3500 Genetic Analyzer (Applied Biosystems Inc., Foster City).

### Analysis of the haplotypes

The forward sequences were edited using the software program Mega 5.2 and aligned with ClustalW and Muscle implemented in Mega 5.2. The alignment was performed using standard adjustments and manual optimization when necessary. The haplotype diversity (*h*) and nucleotide diversity ( $\pi$ ), along with Fu and Li's D and Tajima's D tests, were calculated using DnaSP 5.10.01 software (Librado and Rozas, 2009). Estimates with significant negative values are expected from the Fu and Li's D and Tajima's D tests in populations that have undergone recent demographic expansion.

The best adjusted nucleotide substitution model was obtained using the jModelTest 2.1.4 program (Darriba et al., 2012) to help select the molecular evolution model, and the HKY model was used for posterior phylogenetic inference using Bayesian Analysis and Maximum Likelihood (ML). The Bayesian analysis was performed using Beast v1.8.0 (Drummond and Rambaut, 2007) and the posterior distribution was approximated using Markov Chain Monte Carlo (MCMC) for 50 million steps. The convergence of the parameters was checked using Tracer 1.5 software (Rambaut, 2009). The time to most recent common ancestor (TMRCA) was calculated assuming a relaxed molecular clock (uncorrelated log-normal), following the parameters described by Papadopoulou et al. (2010): (ucl.stdev = 0.2571; a coefficient of variation = 0.2609; and a substitution rate of 0.0168 per million years ago, Ma) (outgroup Lepidoptera sp. – GenBank: JF843940). Genealogical relationships among the haplotypes were estimated using the median-joining

method, implemented in Network 4.613 software (Bandelt et al., 1999).

### Population structuring

Measurements of population differentiation ( $G_{ST}$  and  $N_{ST}$ ) were calculated using DnaSP 5.10.01 (Librado and Rozas, 2009). When the  $N_{ST}$  estimates were greater than those of the  $G_{ST}$ , phylogeographic structuring was assumed, with closely related haplotypes being detected more frequently in the same area than remotely correlated ones. This approach has been used in other studies to detect phylogeographical structure (Guicking et al., 2011; Liu et al., 2012; Chiu et al., 2013). A spatial analysis of molecular variance was conducted using the spatial analysis of molecular variation (SAMOVA) software program (Dupanloup et al., 2002). This software uses simulation to identify groups of populations (*k*) that are geographically homogeneous and those that maximize the differences between groups, allowing variation between the groups ( $F_{CT}$ ), between the locations within each group ( $F_{SC}$ ), and between the locations in relation to the total sample ( $F_{ST}$ ), to be obtained. SAMOVA analyses were conducted with 1000 interactions for  $k = \{2, \dots, 8\}$  groups.

A cluster analysis was constructed using the “Bayesian approach to phylogeographic clustering,” a Bayesian phylogeographic and ecological clustering (BPEC) package (Manolopoulou et al., 2011) implemented on R software (Team, 2012), using the parameters  $ds = 0$ , maximum number of migrations = 5, and 50 million steps in MCMC.

### Ecological niche modeling

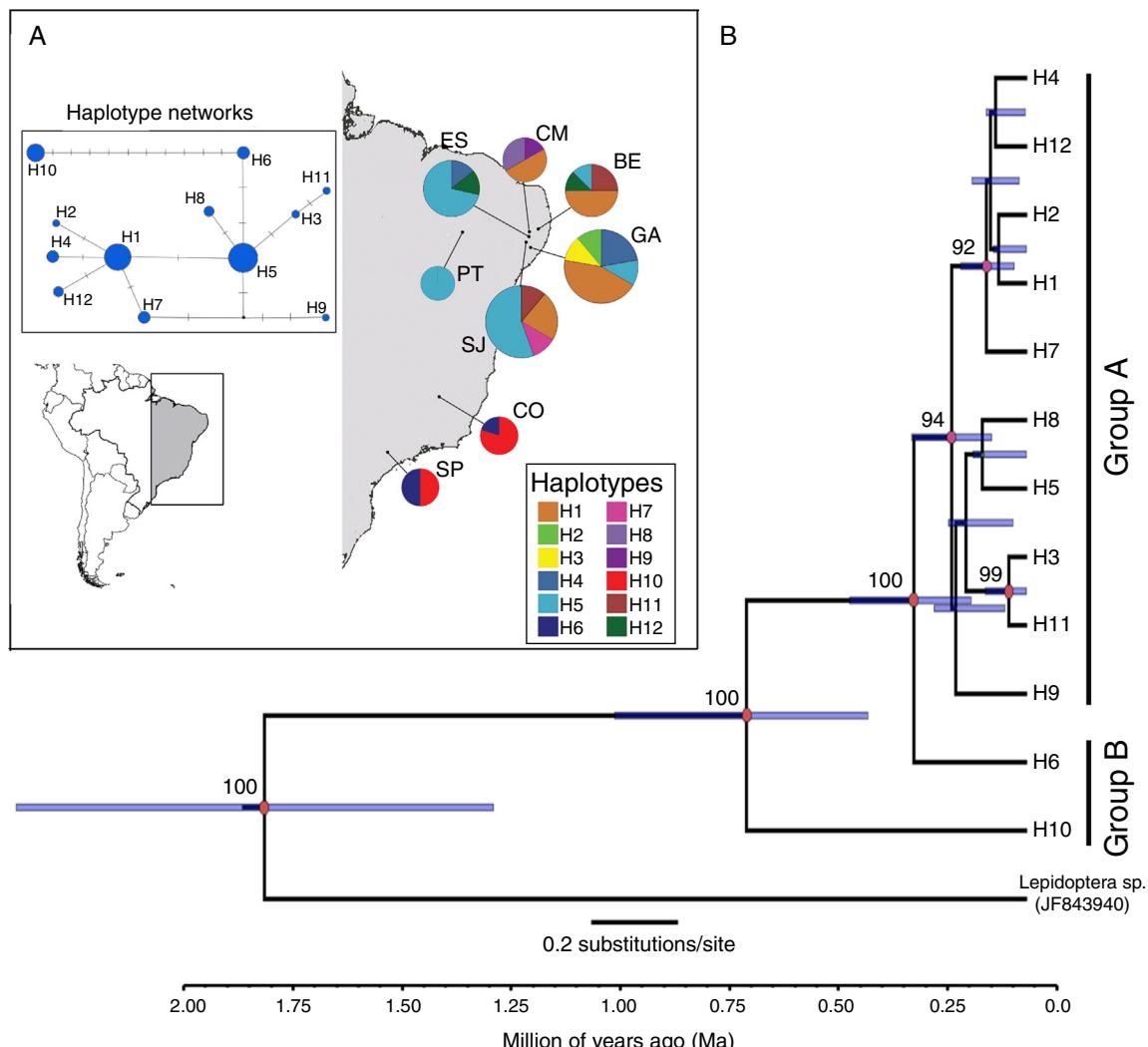
Ecological niche modeling was performed using MAXENT (Version 3.3.3k; Phillips et al., 2006). The climatic niches used were the 19 BIOCLIM variable available in the WorldSIM data base (<http://www.worldclim.org>). The environmental data contain three different periods: bioclim layers for the period from 1950 to 2000 at a resolution of 30 arc s, the last glacial maximum (LGM; <21,000 years BP) in the climatic conditions at a resolution of 2.5 arc min, and the last interglacial (LIG: <120,000–140,000 years BP) at a resolution of 30 arc s. To construct the ecological niches, runs were conducted with the parameters convergence threshold (0.00001), maximum iterations (500), and default prevalence (0.5). The figures were produced using the R software raster package (Team, 2012).

## Results

### Haplotype distribution and analysis

The region of the CO1 gene was sequenced in 51 individuals of *N. elegantalis*, and the analyzed fragments with 628 bp showed 17 polymorphic sites with a total of 12 different haplotypes (Fig. 1A),  $h = 0.836 \pm 0.032$  and  $\pi = 0.06608 \pm 0.00122$ . The haplotype distribution showed a clear distinction between the southeast and northeast region locations in that haplotypes H6 and H10 that occurred only in the populations from the southeast of Brazil, whereas the other haplotypes occurred only in the northeast region (Fig. 1A). In the northeast region locations, haplotypes H1 and H5 occurred with higher frequency, whereas haplotypes H2, H3, H9, and H12 were less common (Fig. 1A). According to the NETWORK results, haplotype H10 had the highest number of substitutions in relation to the most frequently occurring haplotypes (H1 and H5) (Fig. 1A).

The phylogenetic analyses showed four clades, two of them containing haplotypes from the southeast and two of them with only haplotypes from the northeast of Brazil (Fig. 1B). The analysis of the



**Fig. 1.** Distribution patterns of haplotypes for *N. elegantalis*. (A) Frequency of haplotypes for the cytochrome c oxidase subunit 1 (CO1) region by study sites and the haplotype network obtained by median-joining. The study sites are indicated by letters, and the size of the graphics is proportional to the sample size. The colors correspond to the haplotypes as described by the legend: Garanhuns (GA), Petrolina (PT), Coimbra (CO), Camocim (CM), Encruzilhada de São João (ES), São José do R. Pardo (SP), Bezerros (BE) and São João (SJ). (B) Phylogenetic analysis for the haplotypes using the Bayesian approach. This analysis shows the time of divergence between the haplotypes in millions of years (Ma) (95% confidences are indicated by bold horizontal bars) and the groups defined by Spatial Analysis of Molecular Variance (SAMOVA). The support values are estimated with posterior probabilities BY (here in percentages).

TMRCA showed that haplotype H10 had a differentiation of approximately 0.6 Ma, and haplotype H6 of around 0.3 Ma in relation to the other haplotypes (Fig. 1B).

#### Population structuring

The  $G_{ST}$  estimate (0.136) was lower than the  $N_{ST}$  estimate (0.669), indicating population structuring. The spatial analysis of molecular variance resulted in structuring of two groups, with statistical significance ( $F_{CT} = 0.77$  and  $F_{ST} = 0.78$ , Table 1), showing that 77.12% of the variation was between the groups. Group A

consisted of locations in the northeast of Brazil (Garanhuns, Petrolina, Camocim, Encruzilhada de São João, Bezerros, and São João) and Group B was exclusively consisted of locations in the southeast of Brazil (São Paulo e Minas Gerais) (Fig. 1A and Table 2). Haplotype analysis showed that for Group A,  $h = 0.78$  and  $\pi = 0.0026$ , whereas for Group B,  $h = 0.5$  and  $\pi = 0.0079$  (Table 2).

Cluster analysis using BPEC showed three clusters; one with southeast locations and other two with locations in the northeast of Brazil (Fig. 2). However, in this analysis approach, colors represent the clusters and the background or shaded colors show uncertainty about the respective clusters. In the results of this study,

**Table 1**

Spatial analyses of molecular variance (SAMOVA) between two the groups.

Source of variation	d.f.	Percentage of variation	p-Value
Among groups <sup>a</sup>	1	77.12	$F_{CT} = 0.77$ $<0.0001$
Among population within groups	6	1.07	$F_{SC} = 0.05$ $<0.0001$
Within populations	43	21.81	$F_{ST} = 0.78$ $<0.0001$
Total	50	100.00	

<sup>a</sup> Groups are showed in Fig. 1 and Table 2.

**Table 2**

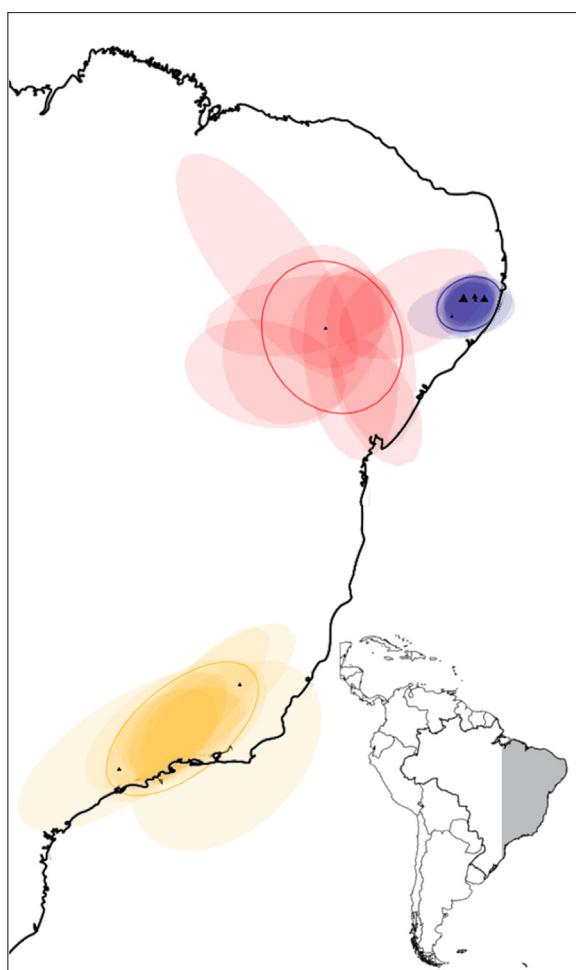
For each groups of populations, haplotype diversity ( $h$ ), nucleotide diversity ( $\pi$ ), Fu and Li's D, Tajima's D and numbers of haplotypes.

Groups	Sample size	$h$	$\pi$	Haplotypes no.	Fu and Li's D	Tajima's D
Group A (GA, PT, CM, ES, BE e SJ)	42	0.78	0.0026	10	-0.95	-1.31
Group B (CO e SP)	9	0.50	0.0079	2	1.47*	1.67
Total	51				0.272	-0.58

there is a lack of clarity in terms of the two clusters in the northeast region of Brazil; thus, they are considered together as a single cluster.

#### Ecological niche modeling

The distribution estimate shows good species distribution representation (Fig. 3). Comparing the predictions for the present and the LGM, we observe a decrease in distribution in the present with a loss of habitat suggesting contraction. On the other hand, comparing the LGM and LIG distributions, the results suggest an expansion of the habitat in the LGM. An analysis of the three ecological niche predictions shows an expansion from the LIG to the LGM and a retraction from the LGM to the present (Fig. 3).



**Fig. 2.** The cluster analysis was built using the "Bayesian approach to phylogeographic clustering." The colors represent the clusters, and the background or shaded colors indicate uncertainty about the respective clusters. The points mark the sampling sites.

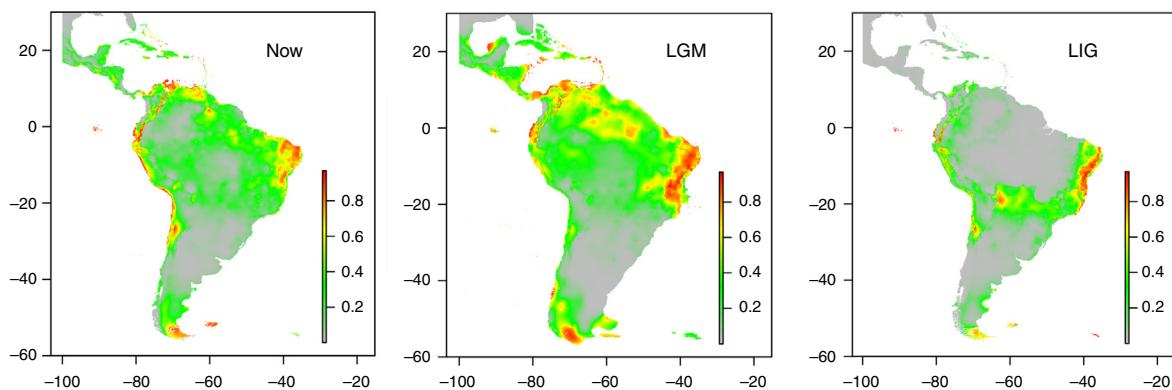
#### Discussion

*N. elegantalis* has been portrayed as an important tomato pest in Brazil, occurring throughout the country stretching from the south to the north. It is highly adaptable to different types of environments, living in cold Atlantic forest regions to semi-arid dry forest regions (Caatinga) of the northeast of Brazil. Accordingly, population studies are extremely important for understanding the population structures and demographic history of the species. In this study, the region of the mitochondrial CO1 gene was sequenced in individuals from different locations distributed throughout Brazil for analyze the hypothesis about population structuring resulting from climate changes, as well as contributing to a clearer understanding of distribution patterns of species resulting from successive climatic cycles during the Pleistocene.

The results of this study demonstrate a clear population structure, separating individuals originating in the southeast and northeast regions of Brazil and identifying the two populations. The haplotypes found in the southeast were not observed in the northeast nor were the haplotypes of the northeast observed in the southeast, showing that there has been no recent gene flow. This evidence is interesting because of the fact that *N. elegantalis* attacks the fruits, and commercialization could be a dispersion factor throughout Brazil, which was not observed in this study. It is possible that the insect is unable to complete its cycle under fruit storage conditions and/or that eliminating the damaged fruit prior to marketing impedes gene flow. Another explanation may be that differences in sex pheromones hinder reproduction between individuals from different regions. Unpublished data suggest that behavioral control using sex pheromones is not much effective when applied to individuals of the southeast and northeast populations, suggesting differences in distinct genetic groups.

The divergence time between the two haplotypes from the southeast and the ten haplotypes from the northeast of Brazil is 0.3–0.6 Ma, suggesting two refuges for the species. In theory, climatic fluctuations in the Pleistocene could have caused the fragmentation of habitats, creating isolated fragments and resulting in refuges for many species. The results of the TRMCA showed that clades containing haplotypes of the southeast diverged from haplotypes of the northeast during the Pleistocene, indicating that the population structure is a result of climatic variations. The hypothesis of the emergence of refuges is validated by the results of ecological niche modeling, in which the distribution shows that there was an expansion from the LIG to the LGM and a contraction from the LGM to the present, suggesting that a contraction may have occurred prior to the LIG.

The separation between populations of the southeast and the northeast of Brazil is observed in many studies. For example, studies by Martins (2011) in vertebrates show a separation between populations of the north and the south of the Atlantic forest between 0.4 and 0.5 Ma. The center of endemism for butterflies shows four regions; two in the southeast and one in the northeast (state of Pernambuco) (Carnaval and Moritz, 2008). A study of *N. elegantalis* in Colombia shows a population structure associated with the Andes Mountains, which act as a barrier to gene flow (Díaz-Montilla et al., 2013), a scenario similar to this study.



**Fig. 3.** Potential distribution as the probability of occurrence of *N. elegantalis* in the present (Now, 0 years BP), last glacial maximum (LGM; <21,000 years BP), and last interglacial (LIG, 120,000–140,000 years BP).

The movement or dispersion of insects across large geographical distances is very important to plant pest management strategies, including the identification of breeds that could have different pheromone compositions and in which genetic variations could differentiate the pest's response to control strategies. For example, the population structure of insect pests enables the definition of patterns of resistance associated with the geographical environment (Labbe et al., 2005). Insect studies have demonstrated an association between genetic variation and cuticular hydrocarbon variation (Nestmate recognition mediator, a common pheromone in social insects) (Dapporto et al., 2009; Dronnet et al., 2006). In this context, this study shows a population structure of *N. elegantalis* associated with geographical distance. Future studies are needed to study association between interpopulation variation and pest control strategies, such as behavioral and/or chemical insect control, because an understanding of dispersion patterns and genetic diversity is necessary for effective control, considering that the adaptation ability of an organism depends on its genetic variability.

### Conflicts of interest

The authors declare no conflicts of interest.

### Acknowledgements

To the Universidade Federal do Alagoas for the laboratories and scientific support and the Fundação de Apoio à Pesquisa de Alagoas for funding this project.

### References

- Badji, C.A., Eiras, A.E., Cabrera, A., Jaffe, K., 2003. Evaluation of sexual pheromone of *Neoleucinodes elegantalis* Guenée (Lepidoptera: Crambidae). *Neotrop. Entomol.* 32, 221–229.
- Bandelt, H.J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48.
- Benvenga, S.R., De Bortoli, S.A., Gravena, S., Barbosa, J.C., 2010. Monitoring the tomato fruit borer population for the control decision. *Hortic. Bras.* 28, 435–440.
- Bonelli, M., Lorenzi, M.C., Christides, J.-P., Dupont, S., Bagnères, A.-C., 2015. Population diversity in cuticular hydrocarbons and mtDNA in a mountain social wasp. *J. Chem. Ecol.* 41, 22–31.
- Carnaval, A.C., Moritz, C., 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *J. Biogeogr.* 35, 1187–1201.
- Chiu, Y., Bor, H., Tan, M., Lin, H., Jean, C., 2013. Phylogeography and genetic differentiation among populations of the moon turban snail *Lunella granulata* Gmelin, 1791 (Gastropoda: Turbinidae). *Int. J. Mol. Sci.* 14, 9062–9079.
- Dapporto, L., Liebert, A.E., Starks, P.T., Turillazzi, S., 2009. The relationships between cuticular hydrocarbon composition, faunal assemblages, inter-island distance, and population genetic variation in Tuscan Archipelago wasps. *Biochem. Syst. Ecol.* 37, 341–348.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772.
- Díaz-Montilla, A.E., Suárez-Barón, H.G., Gallego-Sánchez, G., Saldamando-Benjumea, C.I., Tohme, J., 2013. Geographic differentiation of Colombian *Neoleucinodes elegantalis* (Lepidoptera: Crambidae) haplotypes: evidence for solanaceae host plant association and holdridge life zones for genetic differentiation. *Ann. Entomol. Soc. Am.* 106, 586–597.
- Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19, 11–15.
- Dronnet, S., Lohou, C., Christides, J.P., Bagnères, A.G., 2006. Cuticular hydrocarbon composition reflects genetic relationship among colonies of the introduced termite *Reticulitermes santonensis* Feytaud. *J. Chem. Ecol.* 32, 1027–1042.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.*, 7.
- Dupanloup, I., Schneider, S., Excoffier, L., 2002. A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* 11, 2571–2581.
- Guicking, D., Fiala, B., Blattner, F.R., Slik, F., Mohamed, M., Weising, K., 2011. Comparative chloroplast DNA phylogeography of two tropical pioneer trees, *Macaranga gigantea* and *Macaranga pearsonii* (Euphorbiaceae). *Tree Genet. Genomes* 7, 573–585.
- Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 14812–14817.
- Labbe, P., Lenormand, T., Raymond, M., 2005. On the worldwide spread of an insecticide resistance gene: a role for local selection. *J. Evol. Biol.* 18, 1471–1484.
- Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452.
- Liu, Y., Yang, S., Ji, P., Gao, L., 2012. Phylogeography of *Camellia taliensis* (Theaceae) inferred from chloroplast and nuclear DNA: insights into evolutionary history and conservation. *BMC Evol. Biol.* 12, 92.
- Manolopoulou, I., Legarreta, L., Emerson, B.C., Brooks, S., Tavare, S., 2011. A Bayesian approach to phylogeographic clustering. *Interface Focus* 1, 909–921.
- Martins, F.D.M., 2011. Historical biogeography of the Brazilian Atlantic forest and the Carnaval-Moritz model of Pleistocene refugia: what do phylogeographical studies tell us? *Biol. J. Linn. Soc.* 104, 499–509.
- Miranda, M.M., Coutinho, P.M., Zanuncio, J.C., da Silva, E.M., 2005. Impact of integrated pest management on the population of leafminers, fruit borers, and natural enemies in tomato. *Cienc. Rural* 35, 204–208.
- Papadopoulou, A., Anastasiou, I., Vogler, A.P., 2010. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Mol. Biol. Evol.* 27, 1659–1672.
- Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190, 231–259.
- Picanço, M.C., Bacci, L., Crespo, A.L.B., Miranda, M.M.M., Martins, J.C., 2007. Effect of integrated pest management practices on tomato production and conservation of natural enemies. *Agric. For. Entomol.* 9, 327–335.
- Rambaut, A.D.A., 2009. Tracer v1.5. <http://tree.bio.ed.ac.uk/software/tracer/>.
- Team R.C., 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.