ABSTRACT

Aedes albopictus, the Asian tiger mosquito, has emerged as an important threat for human health being one of the main vectors of deadly arboviruses, like Chikungunya, Dengue and Zika viruses. Nonretroviral integrated RNA virus sequences (NIRVS) with similarities to the non-retroviral RNA viruses of the Flaviviridae and Rhabdoviridae family have been found with high frequency in *Ae. albopictus* mosquitoes. Despite the abundance of NIRVS the biology, functional role and patterns of integration in wild mosquito populations are still relatively unexplored. NIRVS could be expressed and influence the phenotype of mosquitoes acting as host antiviral immunity factors or providing adaptive functions to environmental changes. In addition, NIRVS could serve as alternative genetic markers to monitor the origin and spread of *Ae. albopictus* and the intermixing of geographical mosquito populations to define likely invasion routes.

In this study, we integrated various approaches to study the genomic landscape of NIRVS in wild-collected samples of *Ae. albopictus*. The first objective was to examine NIRVS in distinct geographical populations of *Ae. albopictus* coming from locations in China, Thailand, La Reunion Island, United States and Italy. I examined 13 NIRVS selected for their similarity of Open Reading Frames (ORFs) in Flaviviruses or Rhabdoviruses and their distribution in distinct sites of the mosquito genome. The NIRVS were distributed differentially across the geographic populations and, based on the shared-allele distances calculated for either all NIRVS, NIRVS with similarity to Flaviviruses (F-NIRVS) or intergenic NIRVS, it was possible to match the connections between different mosquito populations with the historical record of their worldwide dissemination. This finding indicates that NIRVS could be useful genetic markers to examine invasion routes and distribution into new geographical areas.

Next, I focused on NIRVS-31N, a new viral integration found in *Ae. albopictus* mosquitoes from Mexico. NIRVS-31N was discovered initially by whole genome sequencing of *Ae. albopictus* mosquitoes from the town of Tapachula in the Mexican state of Chiapas. Testing additional mosquito populations for the presence of NIRVS-31N suggested that this viral integration is unique to mosquitoes from the Tapachula area. My specific objective was to define the frequency of NIRVS-31N in mosquitoes collected in six different locations in the Chiapas state. I found that the frequency of NIRVS-31N ranges from about 40% in Tapachula, the initial site of detection, to 100 % in the Panteon Union Juarez site. These differences in frequency of integration may indicate that the integration is a recent event in the Mexican population of *Ae. albopictus*. Furthermore, the presence of NIRVS-31N could give a fitness advantage to the mosquitoes within the region sites, as they widely differ in terms of geography and urbanization level.

I verified next whether NIRVS-31N was expressed in the mosquitoes that carry the integration. The results of RT-PCR experiments showed evidence of transcription of NIRVS-31N in RNA samples from adult mosquitoes, pupae and larvae of *Ae. albopictus*, suggesting a possible physiological role of NIRVS-31N transcripts. NIRVS-31N could confer novel phenotypes to the mosquito population harbouring it and it will be important to characterize better the NIRVS-31N transcripts and define their function.

Focusing on the Mexican population of the Chiapas region, I examined the origin of this specific population of *Ae. albopictus* by performing microsatellite genotyping and the relation between microsatellite polymorphism and NIRVS-31N integration. I analyzed five microsatellite loci in 36 individual mosquitoes from the Cacahoatan and Huehuetan sites in Chiapas with and without NIRVS-31N integration. Mosquitoes from both populations, with and without NIRVS-31N integration, were distant from each other based on principal coordinate analysis. Furthermore, the two NIRVS-31N positive populations were also distant in clustering analysis. Therefore, there was no apparent relationship between the genotypes defined by microsatellite polymorphisms and NIRVS-31N integration status.

Next, I performed principal coordinate analysis integrating the data from the Mexican population with microsatellite genotyping data from 288 individual *Ae. albopictus* mosquitoes from 11 different geographical worldwide populations. This analysis revealed that mosquitoes from the collection sites in Athens, Brescia and Virginia were the closest to the Mexican mosquitoes, suggesting that these populations might be connected. Therefore, the Mexican population in Chiapas could derive from these geographical areas. Applying the STRUCTURE software, I further analyzed the genetic structure of the worldwide collection of *Ae. albopictus* populations using the microsatellite genotyping data. The data fitted best a two- and six-clustering model. According to both models generated by STRUCTURE, the Mexican mosquitoes clustered together with the Virginia, Athens and Brescia populations, thus confirming a common origin or interconnection between the geographical groups. Furthermore, both analyses showed that the distribution of the Mexican mosquitoes was not affected by the NIRVS integration status, supporting a recent and local acquisition of the NIRVS-31N integration.