SHORT COMMUNICATION

Molecular comparison of Quebec and Newfoundland populations of the blackfly, *Simulium vittatum*, species complex

C. GAUDREAU¹, B. LARUE¹ and G. CHARPENTIER¹

¹Département de Chimie-biologie, Université du Québec à Trois-Rivières, Trois-Rivières, Quebec, Canada

Abstract. Specific haplotypes at five positions in the COI and COII mitochondrial genes allowed a partial differentiation of *Simulium vittatum* Zetterstedt (Diptera: Simuliidae) populations from Quebec–Ontario and Newfoundland, respectively. This geographical signature was superimposed on about 40 other polymorphic sites such that sequence divergence alone did not enable a clear-cut distinction between the two populations. Together with the sporadic occurrence of haplotypes intermediate to the Newfoundland and Quebec–Ontario consensus, this suggested that one peculiar sequence among many found in populations from the North American landmass predominates in Newfoundland as a result of a founder effect. The internal transcribed spacer (ITS1) sequence from the nuclear rDNA transcription unit was no more able to resolve populations along geographical lines than the COI/COII criteria.

Key words. *Simulium vittatum*, blackfly, COI, ITS1, mtDNA markers.

*Simulium vittatum* (Diptera: Simuliidae) is a complex of multivoltine blackfly species widely distributed across North America and Atlantic islands such as Newfoundland and Prince Edward Island (Adler & Kim, 1984). The immature stages occur in a wide range of habitats, from small streams to medium-sized rivers. The adult female, a pest of domestic fowl and many wild and farm-raised mammals, causes significant economic losses that result from severe scabbing, irritation and ear dermatitis (Adler et al., 2004). Feeding on humans remains uncommon (Wolfe & Peterson, 1959; Peterson, 1977), but *S. vittatum* s.l. can be a severe nuisance because it forms persistent swarms and enters the ears, eyes, nose and mouth. Although *S. vittatum* is a vector for the vesicular stomatitis virus and *Onchocerca sp.* (Adler et al., 2004) under laboratory conditions, this role has yet to be proven in the wild.

The *S. vittatum* species complex consists of at least two cytological forms, IS-7 and III-L-1 (Pasternak, 1964; Rothfels & Featherston, 1981), and the formal recognition of IS-7 as *Simulium vittatum* sensu stricto and III-L-1 as *Simulium tribulatum* has been recently proposed (Adler et al., 2004). This taxonomic revision left out the XoYo cytotype (undifferentiated sex chromosomes) which typifies the populations of Iceland and the northwest Atlantic islands (Rothfels & Featherston, 1981; Duncan et al., 2004; Minhas et al., 2005). On the North American continent, the distribution of the IS-7 and III-L-1 morphological twins correlates generally with ecology and geography. The III-L-1 type, which prefers warmer and less well oxygenated streams (Duncan et al., 2004), extends from the Atlantic coast to the Prairies, whereas the IS-7 type is usually found further north, reaching as far as Alaska in the west. III-L-1 and IS-7 coexist over a vast area covering the south of Quebec and Ontario and the northeastern U.S.A. In allopatry, larval populations of either one span early spring to early autumn. Under sympatric conditions, seasonal population peaks are somewhat shifted according to the preferred temperature range of each cytotype (Adler & Kim, 1984), which favours reproductive isolation.

The molecular data record is still ambiguous with respect to the existence of two distinct entities. Allozymes (Zhu et al., 1998) and directed heteroduplex analysis of 12S and 16S mt rRNA genes (Tang et al., 1996) failed to distinguish between the two siblings, whereas random amplification polymorphic
DNA (RAPD) procedures enabled a partial separation and also suggested western and eastern origins for IS-7 and IIII-L-1, respectively (Duncan et al., 2004). In an attempt to differentiate between the S. vittatum populations from Quebec and Ontario (IS-7/III-I-L-1) and Newfoundland (XoYo), we studied the sequence diversity of the COI and COII (cytochrome oxidase) mitochondrial genes and the intergenic spacer ITS1 located between the 18S and 5.8S nuclear rRNA genes. These targets were attractive as their variability often allows for distinguishing between siblings and characterizing population genetic structure (Day et al., 2008; Gaudreau et al., 2008).

Blackfly larvae were collected from five streams: two in Quebec (Rivière Richelieu, 45°25′ N, 73°15′ W and Lac Souris, 46°34′ N, 72°58′ W), one in Ontario (Rivière Rideau, 45°22′ N, 75°41′ W) and two in Newfoundland at the easternmost tip of North America (Broad Cove River, 47°35′ N, 53°05′ W and Bristol Cove Brook, 46°39′ N, 53°05′ W). Larvae fixed in 95% ethanol upon collection were first identified as S. vittatum (28 specimens overall) under a low-magnification microscope using the keys of Wood et al. (1963) and Adler et al. (2004) and secondly as of dark or light form according to head spot pattern (Adler & Kim, 1984). Head spots are confluent on the dark form but appear as separated on the light form.

DNA extraction and purification, amplification of COII or ITS1 and the recovery and sequencing of polymerase chain reaction (PCR) products were described in Gaudreau et al. (2008). The PCR for COI used the same conditions, except for an annealing temperature set at 52 °C for the custom-designed primers SCOII-1 (forward: ATAAATTGCATATTATAGCCGTTTGGTGATTATTAC) and SCOII-2 (reverse: AGATTTCATATTATAGCCGTTTGGTGATTATTAC). Alignment of DNA sequences was performed with GENEIOUS Version 3.8.5. (Drummond et al., 2007). COI and COII consensus sequences for Newfoundland and Quebec–Ontario populations, respectively, have been deposited in Genbank under accession numbers FJ231199–FJ231202 and the CA6C ITS1 variant (see below) has been deposited in Genbank under accession numbers FJ231203–FJ231204. The general time-reversible (GTR) phylogenetic tree for COI/COII was built with PHYLML (Guindon & Gascuel, 2003).

PCR products gave readable sequences of 252 and 651 nucleotides (nt) for COII and COI, respectively. Given that both are mtDNA genes with quite similar evolution rates (Gaudreau et al., 2008), COI and COII were considered as a 0.9-kb aggregate rather than separately. Forty-nine variable positions and 27 distinct haplotypes were noted within a sample including the 28 specimens from this study as well as the S. vittatum Iceland (AF083865, XoYo cytotype) and Nebraska (M76433, III-I-L-1) COI sequences (Pruess et al., 2000), which were not outstandingly distinct from the rest (Fig. 1). In agreement with the strong conservatism of the corresponding polypeptides (Pruess et al., 2000; C. Gaudreau, B. LaRue, G. Charpentier, unpublished data), mutations were all synonymous, except for one (COI position 628) producing an A to T amino acid change. Quite remarkably, more than half of them appeared only once. COI and COII sequence diversity was outstanding in the Quebec–Ontario data, with 16 distinct variants spread between only 17 specimens. Although these findings must be interpreted with caution, given the small sampling sizes, the average sequence divergence between individuals at each Quebec–Ontario location appear to be the same as for all the data clumped together (Table 1), indicating little or no local clustering of haplotypes. The Newfoundland population also showed less internal heterogeneity, especially as much of the average divergence value was contributed by a single specimen. By contrast with the Quebec–Ontario population, all other sequences were very closely related (up to two mutations) to the Newfoundland consensus sequence, which was indeed found as such in three specimens. Overall, the average and maximum divergence levels remained well within the usual limits of intraspecific COI sequence comparisons involving invertebrates (Carew et al., 2007). An attempt to relate all Quebec–Ontario haplotypes in a neighbour-joining tree (data not shown) indicated that many branches involved between two and five mutations. If we assume quite reasonably that COI/COII evolved through point mutations, this implies that many variants were either missed by the sampling in this study or simply became extinct.

Five positions enabled a distinction to be drawn along geographical lines because a consensus rule set separately for each area identified haplotypes ATAGG (from left to right in Fig. 1) and GCGTA as clearly predominant in Quebec–Ontario and Newfoundland, respectively. The distinction remained partial, with a Quebec–Ontario haplotype slipping into the Newfoundland sample and vice versa. A few sequences differing from either ATAGG or GCGTA at a single position were classified with the closest distinctive haplotype for calculation purposes and three others (* in Fig. 1) that did not adhere to the ‘four of five’ rule were left out. The statistical significance of these observations, assessed by a Monte Carlo simulation (105 individual runs) performed in Microsoft EXCEL software, showed the chance occurrence of such an asymmetric distribution between Newfoundland and Quebec–Ontario to be <0.1%. A few specimens showed haplotypes that were intermediate between the respective Newfoundland and Quebec–Ontario consensus, suggesting that the Newfoundland haplotype originated through migration and gene drift from one peculiar sequence found among a pool of pre-existing North American variants. Because of the lack of suitable habitats, any ensuing founder effect occurring under an insular context must obviously be no older than the end of the last glacial episode, which took place about 12 000 years ago. As the ice sheet retreated, Quebec and Ontario would also have been repopulated, but the gene flow within the North American landmass was more likely to maintain more genetic diversity and leave less opportunity for the development of local isolates, an interpretation to be eventually validated by molecular data from all over North America. Finally, a cursory examination of Fig. 1 reveals no differentiation of haplotypes with respect to head spot pattern. The light and dark morphs were found in both Quebec–Ontario and Newfoundland, coexisted at several locations and sometimes corresponded to identical or very similar sequences. This was not unexpected as Zettler et al. (1998) and Adler & Kim (1984) have already demonstrated that larval head morphology is influenced by environmental parameters such as ultraviolet exposure, diet and substrate background colour, but is not linked to cytotologial form or sex.

Further differentiation of DNA specimens through ITS1, located between the 18S and 5.8S genes and usually the
most variable part of the nuclear rDNA transcription unit, was attempted. As Fig. 2 shows, sequencing revealed a length of either 103 bp or 105 bp as a result of an oligoadenylate tract alternating between four (A4 variant) and six (A6) A residues in length. Two additional polymorphisms occurred elsewhere, leading to A/C and C/T dichotomies, respectively. This amounted theoretically to eight distinct sequences, with possible mixes within a same individual adding a further complication as a result of a heterozygous state or an internal heterogeneity of the rDNA cluster. A third of the specimens provided a pure sequence and, in another third, the presence of a single mixed-identity position led to an unequivocal interpretation about the two current sequences. For the remaining specimens, including exceptional ones containing up to four distinct variants, denaturing polyacrylamide gel electrophoresis allowed distinct heteroduplexes accumulated during the final amplification plateau to be distinguished from one another and independently sequenced (Gaudreau et al., 2008).

Among the eight possible variants, CA4C (polymorphic site identity listed 5′ to 3′ with respect to Fig. 2) predominated.

Fig. 2. Sense strand sequence of the 103–105-bp internal transcribed spacer-1 (ITS1). Flanking segments from 18S and 5.8S rDNA were omitted. The A4 or A6 microsatellite and the two other polymorphic sites are underlined. M, A/C; Y, C/T.
clearly in Quebec–Ontario and, in Newfoundland, coexisted at comparable frequency levels with CA6C. Sequences AA4C and AA4T were never detected and the four remaining variants all occurred as minor components (global frequency of 0.14) mixed under a submolar ratio with either CA6C or CA6C, an obvious clue to rDNA cluster heterogeneity. Given this heterogeneity and the lack of a clear geographical fragmentation, ITS1 was finally deemed to be insufficiently discriminating to be used here as a population marker, quite unlike the situation frequently observed with other diterans (Polanco et al., 1998; Rodríguez-Pérez et al., 2006; Gaudreau et al., 2008).

Because of ethanol fixation, cytotype could not be determined. However, under the assumption of a XoYo cytotype for the Newfoundland specimens (Rothfels & Featherston, 1981), COI/COII data alone could not distinguish them from the Quebec–Ontario sample, presumably IS-7/IIIL-1, because the average divergence level between the two populations (0.0083, see Table 1) was similar to that within Quebec and Ontario. Thus, the XoYo cytotype and the GCCTGA COI/COII haplotype would be incidental to recent genetic isolation and would not implicate enough characters for distinct species status to be granted at present.

Acknowledgements

We are grateful to the Université du Québec à Trois-Rivières, which provided C. Gaudreau with a scholarship and funding during her research. Newfoundland specimens were a gift from M. Colbo (Memorial University, St. John’s, NL, Canada).

References


Accepted 18 June 2009

© 2010 The Authors
Journal compilation © 2010 The Royal Entomological Society, Medical and Veterinary Entomology, 24, 214–217