Journal of Biogeography (J. Biogeogr.) (2010) 37, 170–178

ORIGINAL ARTICLE

Disjunct distributions in *Gerris* species (Insecta: Hemiptera: Gerridae): an analysis based on spatial and taxonomic patterns of genetic diversity

M.-C. Gagnon and J. Turgeon*

Département de Biologie et Centre d'Études Nordiques, Pavillon Alexandre-Vachon, 1045 Av. de la Médecine, Université Laval, Québec, QC, Canada G1V 0A6

*Correspondence: Julie Turgeon, Département de Biologie et Centre d'Études Nordiques, Pavillon Alexandre-Vachon, 1045 Av. de la Médecine, Université Laval, Québec, QC, Canada G1V 0A6.
E-mail: julie.turgeon@bio.ulaval.ca

ABSTRACT

Aim To perform a comparative analysis of distribution and genetic diversity in three closely related water strider species (*Gerris*) in order to shed light on a putative disjunct distribution in *Gerris gillettei*.

Location Canada and the western United States.

Methods Entomological collections from Canada and the United States were surveyed for records of *Gerris pingreensis*, *G. gillettei* and *Gerris incognitus* in order to establish the distribution range of each species. Using samples from present populations, mitochondrial and nuclear DNA sequence variation were used to construct minimum-spanning networks. Distribution patterns and genetic diversity were then compared among species.

Results Our results showed that *G. incognitus* is a genetically distinct species with an unsuspected disjunct distribution. *Gerris pingreensis* and *G. gillettei* were found to share genetic polymorphism and they displayed spatial differences only in terms of haplotype distribution, suggesting that they form a single species.

Main conclusions Distributional and molecular information uncover unusual distribution patterns and underline taxonomic uncertainty in a group of three closely related *Gerris* species. Vicariance and failure to recolonize following the last glaciation could explain the *G. incognitus* disjunction. Morphological and DNA-based species identifications suggest different post-glacial recolonization processes for *G. pingreensis* and *G. gillettei*. The putative discontinuous range of *G. gillettei* may be explained as disjunct phenotypes of a single species.

Keywords
Canada, comparative phylogeography, disjunct distributions, Gerridae, *Gerris*, mitochondrial DNA, morphological identification, nuclear DNA, western United States.

INTRODUCTION

Detailed information about species distributions is fundamental to understanding local, regional and global patterns of diversity and how these patterns change over time (Rosenzweig, 1995; Ricklefs, 2004; Lomolino *et al.*, 2006). Knowledge about the spatial distribution of species is also an essential tool in the planning and design of protected areas for conservation (Ferrier, 2002; Rondinini *et al.*, 2006) or in the detection of recent range shifts associated with global change (Parmesan & Yohe, 2003; Lenoir *et al.*, 2008).

Although distributions are well established for some taxa, our knowledge about the geographical ranges of most living species is incomplete (the Wallacean shortfall; Whittaker *et al.*, 2005). Moreover, biodiversity surveys are notoriously biased towards areas near biological stations, work centres, roads or the recorder's own home range (Dennis & Thomas, 2000; Hijmans *et al.*, 2000; Kadmon *et al.*, 2004), affecting reported distribution patterns for taxa from remote locations. Likewise, unusual distributions, such as apparent disjunctions, may conceal part of a species' distribution. Disjunctions have been documented at various taxonomic and spatial scales. A classic example is represented by lungless salamanders of the supergenus *Hydromantes* (family Plethodontidae), which can be found in California but also in southern Europe (Vieites *et al.*, 2007). On a smaller geographical scale, populations of many

www.blackwellpublishing.com/jbi
doi:10.1111/j.1365-2699.2009.02195.x

Journal of Biogeography

plant species have disjunct distributions in North America. A common pattern places the main range of 62 species on the Atlantic and Gulf coastal plain while small and localized populations are also found in the Great Lakes region (Reznicek, 1994).

Taxonomic errors, such as misdiagnosis and misidentification of species, changes in taxonomic nomenclature, as well as inadequate delineation of species boundaries, may also distort species distributions and biogeographical inferences. In a recent revision of the genus *Euscelidia* (Insecta: Diptera: Asilidae), for example, 83% of labelled specimens of museum collections were found to be incorrectly identified. Most specimens were misidentified (73%) while a smaller proportion (10%) of the incorrect identifications resulted from synonymy (Meier & Dikow, 2004). These two types of error seem to be widespread in natural history collections (Meier & Dikow, 2004). Species identification is also difficult when morphologically similar specimens belong to several cryptic species. For example, the marine copepod *Eurytemora affinis* was considered cosmopolitan across Northern Hemisphere estuaries until molecular investigations revealed that *E. affinis* comprised eight highly divergent and reproductively isolated species, each having a distinct and restricted geographical range (Lee, 2000).

In recent years, molecular tools have contributed significantly to revealing the historical and recent dynamics of species distributions. Phylogeography has clearly emerged as a powerful tool to better decipher post-glacial recolonization patterns (Bernatchez & Wilson, 1998; Taberlet *et al.*, 1998; Petit *et al.*, 2002; Hewitt, 2004). Population genetic analyses also contribute to the detection and understanding of recent range expansions and biological invasions, e.g. the significance of multiple introductions to the invasive potential of perennial grasses such as *Phalaris arundinacea* in North America (Lavergne & Molofsky, 2007). Finally, investigations of molecular variation can reveal unsuspected relationships among taxa and their associated ranges. In *Enallagma* damselflies, phylogenetic reconstructions revealed that two morphological species with Holarctic distributions each comprised two old lineages with distinct Palaearctic and Nearctic distributions, while many morphologically different species extensively shared genetic polymorphisms (Turgeon *et al.*, 2005).

In 2005, while investigating the diversity and distribution patterns of aquatic invertebrates in northern Quebec, many populations of water striders were sampled. All specimens belonged to the *gillettei* group, a group of closely related species comprising three Nearctic (*Gerris incognitus* Drake & Hottes, 1925, *Gerris pingreensis* Drake & Hottes, 1925 and *Gerris gillettei* Lethierry & Severin, 1896) and one Palaearctic species (*Gerris sphagnetorum* Gaunitz, 1947) (Andersen, 1993; Damgaard & Sperling, 2001). A small proportion of these specimens could unambiguously be identified as *G. incognitus*. However, for most specimens, morphology-based species identification was problematic; taxonomic identification was therefore entrusted to a gerrid specialist, Dr J. R. Spence (University of Alberta, Canada). The majority of specimens were identified as *G. pingreensis*. Along with *G. incognitus*, this species inhabits lakes and ponds and is commonly found in Canada and western United States (Drake & Harris, 1934; Smith, 1988; Andersen, 1993; Maw *et al.*, 2000). However, many individuals could not be clearly identified as belonging to *G. pingreensis*. Based on the presence of a pair of median longitudinal grooves on the sternum more distinctly separated by a fine ridge (Andersen, 1993), these individuals were tentatively identified as *G. gillettei*, the sister species of *G. pingreensis* (Damgaard & Sperling, 2001). This identification was totally unexpected given that the recognized distribution of *G. gillettei* is limited to mountainous regions of the western United States (Washington, Oregon, Missouri, Wyoming, Colorado, Texas, Utah, Nevada, California and New Mexico; Drake & Harris, 1934; Callahan, 1974; Smith, 1988; Andersen, 1993). If confirmed, the presence of *G. gillettei* in northern Quebec would reveal an important disjunction for this water strider species in North America.

This study combines distributional and molecular information to illustrate the impact of taxonomic decisions and identifications on distribution and biogeographical inferences. We use a comparative approach involving three *Gerris* species. First, distribution patterns of the three species are reappraised based on recent samples and entomological collections. Then, patterns of molecular variation in mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) are compared among species to assess the current taxonomy. Resulting biogeographical inferences and hypotheses are contrasted and discussed.

MATERIALS AND METHODS

Sampling and distribution survey

Specimen-based records of *G. pingreensis*, *G. gillettei* and *G. incognitus* were obtained from natural history museums and academic institutions across Canada and the United States. Existing natural populations of these species were also sampled in northern Quebec, Manitoba, Alberta, British Colombia and California (see Appendix S1 in Supporting Information). All records were mapped to establish the known distribution range for each species (Fig. 1). Sites reported were chosen in order to cover the extent of the distribution for each species that can be corroborated by existing specimens. When there were many records within a region, one locality was chosen and records within *c.* 150 km were not mapped.

DNA analyses

For 103 specimens collected from present populations (Table 1), total DNA was extracted from headless insects following classical phenol–chloroform purification (Sambrook & Russell, 2001) with isopropanol precipitation. Sequence information from GenBank was also included in the analyses, as presented in Table 1 and Fig. 2.

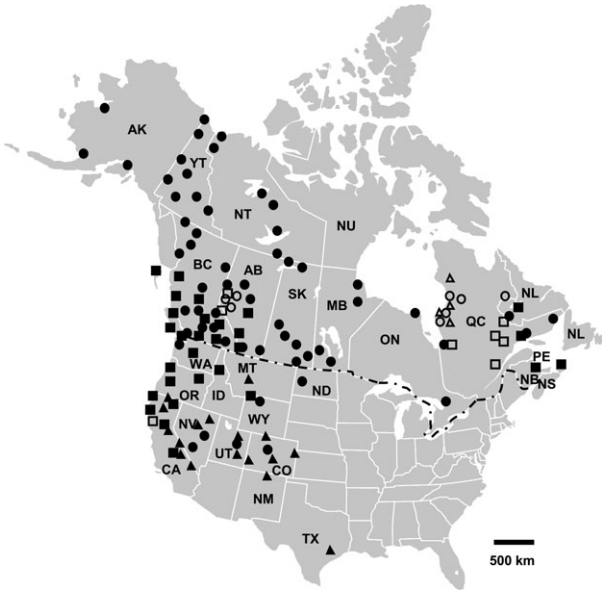

Figure 1 Distribution maps of *Gerris pingreensis* (circles), *Gerris gillettei* (triangles) and *Gerris incognitus* (squares) based on entomological collection records (filled symbols) and samples collected from natural populations (open symbols). Locations: AB, Alberta; AK, Alaska; BC, British Columbia; CA, California; CO, Colorado; ID, Idaho; MB, Manitoba; MT, Montana; NB, New Brunswick; ND, North Dakota; NL, Newfoundland; NM, New Mexico; NS, Nova Scotia; NT, Northwest Territories; NU, Nunavut; NV, Nevada; ON, Ontario; OR, Oregon; PE, Prince Edward Island; QC, Quebec; SK, Saskatchewan; TX, Texas; UT, Utah; WA, Washington; WY, Wyoming; YT, Yukon. Additional information on specimen locations is provided in Appendix S1.

Mitochondrial DNA variation was assessed using a 831-bp segment of the cytochrome *c* oxidase subunit I (COI) that was polymerase chain reaction (PCR)-amplified with primers C1-J-2183 and TL2-N-3014 (Simon *et al.*, 1994). For some amplifications, two internal primers were also used, namely C1-J-2578 (Galacatos *et al.*, 2002) with C1-J-2183 and C1-J-2441 (Simon *et al.*, 1994) with TL2-N-3014. Amplification reactions were carried out in 45 μL containing 4.5 μL of dNTP (2 mM), 3 μL of each primer (5 μM), 1.8 μL of $MgCl_2$ (50 mM), 0.3 μL of *Taq* DNA polymerase (5000 U mL^{-1}) and 2 μL of template DNA (50 ng μL^{-1}). Conditions were as follows: 1 cycle of 3 min at 95 °C, 35 cycles of 30 s at 95 °C, 30 s at 48 °C (annealing temperature, T_a), 1 min at 72 °C and 1 cycle of 10 min at 72 °C.

Nuclear DNA variation was assessed using a 485-bp segment of the elongation factor 1 alpha (EF1-α) amplified with primers EF1-F (5′-CTGGTGAATTCGAAGCTGGT-3′) and EF1-R (5′-GGCGAAGGTGACAACCATAC-3′) designed from a consensus sequence of several *Gerris* species. Amplification reactions were as above but with 1.35 μL of $MgCl_2$ (50 mM). The PCR conditions were also as above but with T_a = 55 °C and a shorter elongation time of 40 s.

All PCR products were cleaned using a QIAquick PCR purification Kit (Qiagen Inc., Mississauga, ON, Canada) and sequenced at the Plate-forme d'Analyses Biomoléculaires (Université Laval). Electrophoretograms were visually inspected to ensure proper base calling. DNA sequence alignment was performed using ClustalW (Thompson *et al.*, 1994) as implemented in mega version 4 (Tamura *et al.*, 2007). Most fragments were sequenced only in one direction. However, when there were ambiguities, the sequence was confirmed by sequencing in the opposite direction. For nDNA, nucleotide positions that showed double peaks were scored as heterozygous. Nuclear haplotypes were determined by subtracting known haplotypes (homozygous at all positions) from heterozygous genotypes (Clark, 1990; Olsen & Schaal, 1999; Abe *et al.*, 2005). This method was applied to sequences produced in this study and retrieved from GenBank. However, when it was not possible to determine which combination of haplotypes a heterozygous individual possessed, all of the possible haplotypes were represented in the results.

Genetic analyses

For each species, overall and within-population mtDNA and nDNA polymorphism was quantified by the number of haplotypes or genotypes, gene diversity (h; Nei, 1987) or nucleotide diversity (π; Nei, 1987), and the mean number of pairwise differences (Tajima, 1983). mtDNA and nDNA minimum-spanning networks were constructed using uncorrected pairwise difference among haplotype sequences observed in this and other studies. Two loops were resolved according to the rules of Crandall & Templeton (1993). Analyses were performed with Arlequin version 3.1 (Excoffier *et al.*, 2005).

RESULTS

Distribution range

Specimens of *G. incognitus* were documented from the western United States (Washington, Idaho, Wyoming, Oregon and California) and Canada (British Columbia and Alberta) and the easternmost part of Canada (Quebec, Nova Scotia, Prince Edward Island and Newfoundland) (Fig. 1 and Appendix S1). In contrast with previous work (Smith, 1988; Maw *et al.*, 2000), but in accordance with specific references for the regions involved (Brooks & Kelton, 1967; Cheng & Fernando, 1970), no specimens of *G. incognitus* were located from the central provinces of Canada (Saskatchewan, Manitoba and Ontario), revealing a disjunction in this species' distribution.

Information from entomological collections corresponded to the recognized range of *G. pingreensis* in North America (Drake & Harris, 1934; Smith, 1988; Andersen, 1993; Maw *et al.*, 2000). Specimens from the western United States (Alaska, Washington, Colorado, Utah, North Dakota, Wyoming and Nevada) and most of Canada, with the exception of Nunavut and the Maritimes, were located in collections (Fig. 1 and Appendix S1).

Specimens identified as *G. gillettei* were reported from the western United States (California, Colorado, Nevada, New

Table 1 Genetic diversity at cytochrome *c* oxidase subunit I (COI) and elongation factor 1 alpha (EF1-α) loci for *Gerris pingreensis*, *Gerris gillettei* and *Gerris incognitus*: gene diversity (h), nucleotide diversity (π), mean number of pairwise differences (pairwise diff.), numbers and haplotype/genotype codes.

			COI				EF1-α			
Species	Site	Prov./ State	*n*	*h*	Pairwise diff.	Haplotypes	*n*	π	Pairwise diff.	Genotypes
Gerris pingreensis	Overall		54	0.72 ± 0.03	1.10 ± 0.73	A, E, F, H, I, J, K	28	0.0014 ± 0.0013	0.54 ± 0.46	*aa, bb, ab, ad* or *bc, ae* or *bf*
	Schefferville	QC	5	0	0	A	4	0.0014 ± 0.0014	0.54 ± 0.49	*ab*(3), *bb*
	Umiujaq	QC	7	0	0	A	4	0	0	*aa*
	Lac Eau Claire	QC	3	1.00 ± 0.27	1.33 ± 1.10	A(1), E(1), F(1)	2	0	0	*aa*
	Kuujjuaraapik-1	QC	6	0.33 ± 0.22	0.33 ± 0.38	A(5), E(1)	2	0	0	*bb*
	Chisasibi	QC	6	0.73 ± 0.16	0.93 ± 0.74	A(3), F(2), H(1)	3	0.0014 ± 0.0015	0.53 ± 0.51	*aa, bb*(2)
	Churchill	MB	7	0.52 ± 0.21	0.57 ± 0.52	F(5), I(1), J(1)	4	0.0014 ± 0.0014	0.54 ± 0.49	*ab*(3), *bb*
	Switzer Lake	AB	6	0.33 ± 0.22	0.33 ± 0.38	F(1), K(5)	4	0.0015 ± 0.0015	0.57 ± 0.51	*aa*(2), *bb*(2)
	Grande Cache	AB	5	0.60 ± 0.18	0.60 ± 0.56	F(2), K(3)	2	0.0030 ± 0.0028	1.17 ± 0.93	*ab, ad* or *bc*
	Victor Lake	AB	8	0.25 ± 0.18	0.25 ± 0.31	F(1), K(7)	2	0.0013 ± 0.0016	0.50 ± 0.52	*ab, aa*
	GenBank*	AB	1	–	–	K	1	0	–	*ae* or *bf*
Gerris gillettei	Overall		20	0.73 ± 0.08	1.20 ± 0.80	A, B, C, D, F, G	11	0.0013 ± 0.0013	0.52 ± 0.46	*aa, bb, ab, gh*
	Boniface	QC	6	0.53 ± 0.17	1.60 ± 1.10	B(4), D(2)	3	0.0015 ± 0.0016	0.60 ± 0.55	*ab*
	Kuujjuaraapik-2	QC	5	0.40 ± 0.24	0.40 ± 0.44	A(4), F(1)	2	0	0	*bb*
	Kuujjuaraapik-3	QC	4	0.67 ± 0.20	0.67 ± 0.63	A(2), F(2)	2	0.0017 ± 0.0019	0.67 ± 0.63	*ab*
	Radisson	QC	4	0.50 ± 0.27	1.00 ± 0.83	A(3), G(1)	3	0.0014 ± 0.0015	0.53 ± 0.51	*aa*(2), *bb*
	GenBank*	NM	1	–	–	C	1	–	–	*gh*
Gerris incognitus	Overall		32	0.59 ± 0.07	1.70 ± 1.02	L, M, N, O, P, Q, R	16	0.0003 ± 0.0006	0.13 ± 0.20	*ii, jj*
	Fermont	QC	5	0.40 ± 0.24	0.40 ± 0.44	M(4), N(1)	2	0	0	*ii*
	Camp St-Pierre	QC	4	0	0	M	2	0	0	*ii*
	Camp Kruger	QC	3	0	0	M	2	0	0	*ii*
	Route 175	QC	1	–	–	O	1	–	–	*ii*
	Route 109	QC	6	0	0	M	2	0	0	*ii*
	Grande Cache	AB	6	0.33 ± 0.22	0.67 ± 0.59	L(5), M(1)	2	0	0	*ii*
	Mount Robson	BC	5	0.40 ± 0.24	2. 00 ± 1.34	L(4), Q(1)	3	0	0	*ii*
	San Francisco	CA	1	–	–	P	1	–	–	*jj*
	GenBank*	NM	1	–	–	R	1	–	–	*jj*
Total *n*			106				55			

QC, Quebec; MB, Manitoba; AB, Alberta; NM, New Mexico; BC, British Columbia; CA, California.
*Not used for calculation of diversity indices.

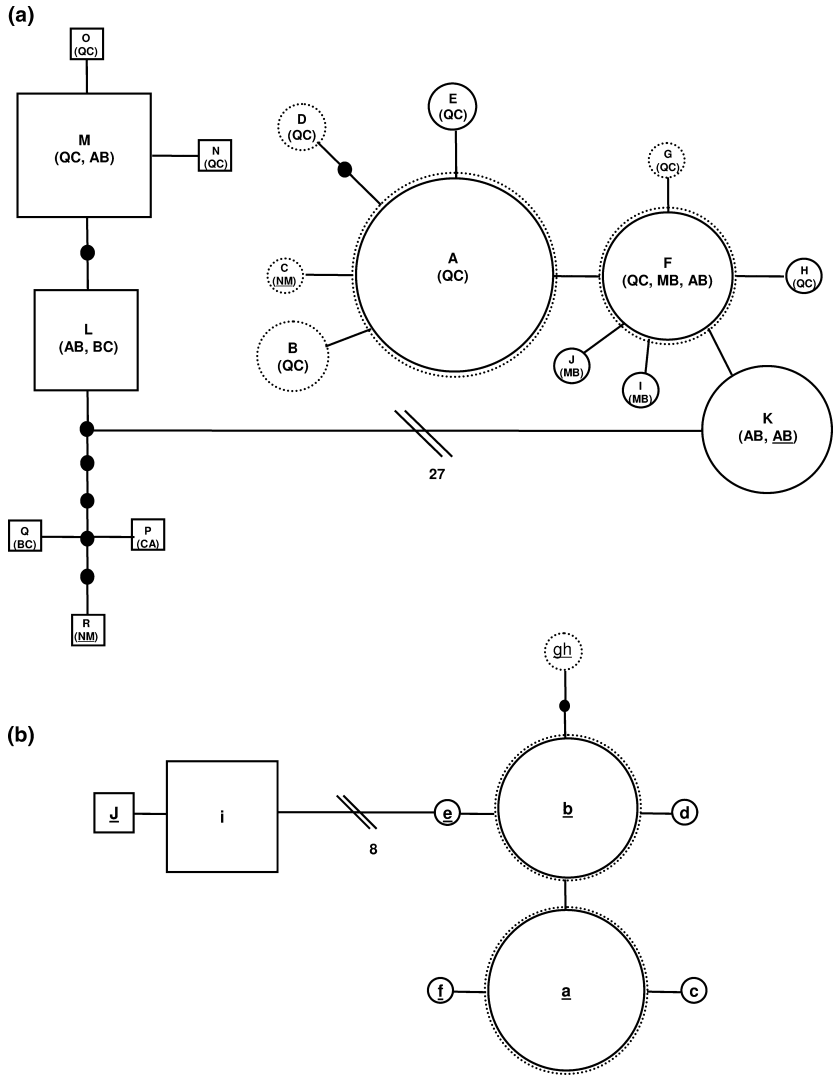

Figure 2 Minimum-spanning networks for (a) cytochrome *c* oxidase subunit I (COI; GenBank accession numbers GQ353353–GQ353367) and (b) elongation factor 1 alpha (EF1-α; GenBank accession numbers GQ353368–GQ353371) obtained for *Gerris pingreensis* (plain circles), *Gerris gillettei* (dotted circles) and *Gerris incognitus* (squares). Letters represent different haplotypes observed in this study or in GenBank [underlined, accession numbers AF251104 (R), AF251105 (C), U83345 (K), AF251084 (*jj*), AF251085 (*gh*) and AF200275 (*ae* or *bf*)]. Size is proportional to the relative abundance of each haplotype, ranging from 1 to 30 copies in (a) and 1 to 38 copies in (b). Connecting lines between haplotypes represent a mutational step while small black dots represent unsampled haplotypes.

Mexico, Texas, Utah and Montana). We also collected individuals putatively identified as *G. gillettei* in four localities from northern Quebec (Fig. 1 and Appendix S1). This is a considerable range extension and would represent a new addition to the Canadian fauna. In the western United States, records attributed to both *G. gillettei* and *G. pingreensis* were found from nearby localities in Colorado, Nevada and Utah. In northern Quebec, populations of *G. pingreensis* and of putative *G. gillettei* were also interspersed (Fig. 1).

mtDNA polymorphisms

Mitochondrial DNA variation clearly revealed that *G. incognitus* was genetically distinct from *G. pingreensis* and *G. gillettei* which, in contrast, hardly differed. The latter two species shared many haplotypes but there were differences in the spatial distribution of haplotypes.

A 748-bp COI sequence was obtained for 106 individuals. Seven haplotypes were found in *G. pingreensis* (A, E, F, H, I, J and K), six in putative *G. gillettei* (A, B, C, D, F and G) and seven in *G. incognitus* (L, M, N, O, P, Q and R) (Fig. 2a and Table 1). Overall, gene diversity was slightly higher for specimens identified as *G. pingreensis* and *G. gillettei* than for *G. incognitus*, and the mean number of pairwise differences was slightly greater for *G. incognitus* than for the other two species (Table 1).

The mtDNA minimum-spanning network of specimens attributed to all three species defined two haplotype clusters separated by 28 mutational steps (Fig. 2a). *Gerris incognitus* haplotypes formed one cluster, which further divided into two subclusters separated by four mutational steps. The most common haplotypes (L and M), along with rare haplotypes O and N, formed one subcluster. Haplotype L was found in western Canada (Alberta and British Columbia) while haplotype M was found mostly in Quebec, with one individual from Alberta (Fig. 2a and Table 1). The other subcluster comprised *G. incognitus* haplotypes restricted to western areas (New Mexico, California and British Columbia).

The 11 mtDNA haplotypes observed in specimens attributed to *G. pingreensis* and *G. gillettei* formed a single, tight network (Fig. 2a). Two very common haplotypes (A and F) were extensively shared between *G. pingreensis* and *G. gillettei*. These

haplotypes were found, respectively, in 30 and 15 of the 72 specimens of these two species sequenced. Haplotypes A and K were found only in northern Quebec and Alberta, respectively, while the other common and shared haplotype (F) had a large geographical distribution (Fig. 2a). The *G. pingreensis* haplotype from GenBank (K), collected in Alberta, was identical to the most common haplotype in western Canada. By contrast, GenBank *G. gillettei* haplotype (C), from New Mexico, was closely related to the most common haplotype found in both *G. pingreensis* and *G. gillettei* in northern Quebec (Fig. 2a and Table 1).

nDNA polymorphisms

Nuclear DNA variation indicates that *G. incognitus* is genetically distinct from *G. pingreensis* and *G. gillettei*. As for mtDNA, specimens attributed to these latter two species hardly differed and, in fact, shared the same haplotypes all across the continent.

A 394-bp sequence for EF1-α was obtained for 55 specimens identified as *G. pingreensis*, *G. gillettei* and *G. incognitus* (Table 1). Thirty-nine individuals were homozygous at all positions, 13 were heterozygous at one nucleotide position, two were heterozygous at two nucleotide positions while one was heterozygous at three nucleotide positions (Table 1), altogether defining eight nuclear haplotypes (*a*, *b*, *c* or *d*, *e* or *f*, *g*, *h*, *i* and *j*). Two haplotypes were restricted to *G. incognitus* (*i* and *j*), four were found in *G. pingreensis* (*a*, *b*, *c* or *d* and *e* or *f*) and four in *G. gillettei* (*a*, *b*, *g* and *h*). Again, common haplotypes (*a* and *b*) were shared between *G. pingreensis* and *G. gillettei*.

The minimum-spanning network distinguished two groups of haplotypes, one corresponding to *G. incognitus* and the other to the combined sample of *G. pingreensis* and *G. gillettei*. Overall and within-population haplotype diversity was greater for *G. pingreensis* and *G. gillettei* than for *G. incognitus*, where all individuals except one possessed the same (*i*) haplotype (Table 1). Haplotype *i* was observed in all specimens of *G. incognitus* from both eastern and western Canada (Fig. 2b and Table 1). Haplotype *j* was found only in one specimen from California and was identical to the *G. incognitus* sequence from GenBank from a specimen collected in New Mexico.

The nuclear haplotypes observed in specimens attributed to *G. pingreensis* and *G. gillettei* formed a single cluster (Fig. 2b). Haplotypes *a* and *b* were most common and were observed in all individuals of the two species collected in Quebec, Manitoba and Alberta. One individual of *G. pingreensis* from Alberta was heterozygous at two nucleotide positions. This individual possessed haplotypes *a* and *d*, or *b* and *c* (Fig. 2b and Table 1). Similarly, the *G. pingreensis* sequence from GenBank, collected in Alberta, was composed of haplotypes *a* and *e* or *b* and *f*. Inferred haplotypes *c*, *d*, *e* and *f* clearly clustered with haplotypes *a* and *b* in the network. The *G. gillettei* sequence from GenBank, collected in New Mexico, was heterozygous at three nucleotide positions, thus defining eight inferred haplotypes. Our reconstruction indicated that they are separated from haplotype *b* by at least two mutations. To avoid overloading the network, these are represented as '*gh*' in the network depiction (Fig. 2b).

DISCUSSION

Our results show that *G. incognitus* is a genetically distinct species, apparently with an interesting range disjunction (Figs 1 & 2). The findings of Smith (1988) and Maw *et al.* (2000) indicating that this species is found in central Canada could not be corroborated. No *G. incognitus* specimens were reported from Saskatchewan, Manitoba or Ontario from the many collections analysed, and in Alberta *G. incognitus* populations are rare, patchy, temporally unstable and confined to areas near the continental divide (J. R. Spence, University of Alberta, pers. comm.). It appears that reports of *G. incognitus* from central Canada were unsubstantiated and could have resulted from a typographical error during the editing of the catalogue (C. L. Smith, University of Georgia, pers. comm.). This catalogue provided the Saskatchewan, Manitoba and Ontario records of *G. incognitus* reported by Maw *et al.* (2000) (G. G. E. Scudder, University of British Columbia, pers. comm.). To our knowledge, this is the first mention of such an unusual east–west disjunction in North America.

The disjunct distribution of *G. incognitus* in North America could be explained as a vicariance event associated with Pleistocene glaciations and a failure of this species to recolonize the central provinces of Canada. Vicariance involving post-glacial recolonization from two distinct ice-free refugia is suggested by the geographically polarized patterns of mtDNA polymorphism (Fig. 2a). Given the absence of *G. incognitus* from Yukon and Alaska, it probably recolonized the western portion of its range from a Pacific refuge south of the ice sheet, and the eastern part from an Atlantic or, perhaps, Mississippian refuge. The absence of *G. incognitus* from the central provinces of Canada could also result from competitive exclusion by other water strider species following post-glacial recolonization. Indeed, it has been demonstrated that nymphs of *G. pingreensis* have a competitive advantage over other water strider species (Spence, 1983; Spence & Carcamo, 1991). Early hatching allows this species to monopolize resources and to feed on nymphs of species with slower development (Spence, 1983; Spence & Carcamo, 1991). This hypothesis, however, remains to be tested.

The extensively shared genetic polymorphism and the interspersed distribution patterns of *G. pingreensis* and *G. gillettei* suggest that these two taxa are in fact a single species. Indeed, the haplotype networks of *G. pingreensis* and *G. gillettei* overlapped completely and they were nearly always identical for nDNA (Fig. 2). These results do not stem from a lack of resolution as the markers used were adequate to distinguish reliably between *G. incognitus* and *G. pingreensis*/*G. gillettei*. Our results contrast with previous phylogenetic work that classified *G. pingreensis* and *G. gillettei* as distinct on the basis of mtDNA and nDNA, using a single individual of each species (Damgaard & Sperling, 2001). We suggest that the extent of shared genetic variation revealed by our dataset

outweighs the evidence provided by the limited dataset used by Damgaard & Sperling (2001).

The extreme morphological similarity between *G. pingreensis* and *G. gillettei* is also compatible with these two taxa forming a single species. Females of these species cannot be reliably distinguished and Andersen (1993) recognized that these two species must indeed be closely related. The only character that has been used to separate individuals of both sexes into these two species is the presence (*G. gillettei*) or absence (*G. pingreensis*) of silver spots on dorsal paratergites. The presence and extent of development of distinct spots varies among populations from the western United States (J. R. Spence, University of Alberta, pers. comm.). Separating males into species relies on the extent of development of a medial ridge separating grooves on sternum 6 and 7 of males. This ridge is said to be more developed in *G. gillettei* and less so in *G. pingreensis* (Andersen, 1993), but there is significant and overlapping variation in this trait within populations within the ranges presently defined for *G. pingreensis* and *G. gillettei* (M.-C. Gagnon, pers. obs.; J. R. Spence, University of Alberta, pers. comm.). Based on this information, we suggest that these two taxa should be synonymized as *G. gillettei*, based on the principle of priority. However, given that we were unable to obtain specimens of *G. pingreensis* and *G. gillettei* from the western United States, formal taxonomical changes should be put off until this ultimate analysis is undertaken.

A similar case of inadequate taxonomy is known for *Gerris sahlbergi* and *Gerris costae*, considered distinct on the basis of small morphological differences (Andersen, 1993). However, Damgaard (2006) showed that mtDNA and nDNA polymorphism observed in *G. sahlbergi* was nested within that of *G. costae*, suggesting that these two species were in fact the same species.

The distributions, external structure, molecular evidence and synonymy proposed for *G. gillettei* and *G. pingreensis* lead to new biogeographical inferences. First, distributional data and recent collections document the extensive, continuous range of *G. pingreensis* in North America (Fig. 1). Records of specimens fitting the description of *G. pingreensis* in northern Quebec only extend the north-eastern limit of the range previously understood for this species. Second, morphology-based identification revealed four putative populations of *G. gillettei* in northern Quebec, an observation completely at odds with the recognized distribution of *G. gillettei* being limited to the western United States. This would result in *G. gillettei* having an unusual disjunct distribution. The example of *G. incognitus* might suggest that a disjunction in the range of *G. gillettei* could also be explained in terms of vicariance and recolonization. However, if *G. pingreensis* and *G. gillettei* are more simply morphological variants with common genetic structure, populations in northern Quebec can be interpreted as having resulted from post-glacial colonization. This interpretation is supported by the fact that the two morphological types are often found in near sympatry in the western United States (Utah, Nevada and Colorado), with the latter seemingly confined to high altitudes, while populations of both types are interspersed in northern subarctic Quebec (Fig. 1). A post-glacial Beringian origin of this stock can be inferred given the continuous distribution of *G. pingreensis* from Beringia southwards along the Rockies and eastwards in northern Canada. A Pacific refuge followed by post-glacial recolonization of northern areas is also consistent with the present distribution of this species.

CONCLUSION

We used three water strider (*Gerris*) species to show how valid species identification is central to reliable inferences about distribution and diversity patterns. We suggest that morphology-based species identification can misinterpret natural intraspecific variation while, at the same time, DNA-based systematics can support premature designation of interspecific limits. When combined, traditional and molecular systematics, phylogeography and biogeography can be used to test hypotheses about diversification processes and the historical origin of species ranges. A major problem identified by this study is the surprising lack of accurate information about the distribution of even well-known insect taxa in Canada. This will only be resolved by support and encouragement for field biology.

ACKNOWLEDGEMENTS

We thank J.-F. Aubry, E. Boucher, E. Cardinal, S. Cyr, P. Etcheverry, A. Gonzalez, J.-F. Simard and J. Witt for samples of present-day insects; C. Barr (Essig Museum of Entomology), R. Blinn (North Carolina State University Insect Collection), S. Boucher (Lyman Entomological Museum), P. J. Clausen (University of Minnesota Insect Collection), D. C. Currie (Royal Ontario Museum), R. G. Foottit (Canadian National Collection of Insects, Arachnids and Nematodes), T. Galloway (J. B. Wallis Museum), M. F. O'Brien (University of Michigan Museum of Zoology), J. O'Donnell (University of Connecticut Insect Collection), V. Scott (University of Colorado, Museum of Natural History), K. B. Simpson (Enns Entomology Museum) and G. G. E. Scudder (Colbo Collection, Jamieson Collection, Royal British Columbia Museum, Saskatchewan Museum, Spencer Entomological Collection, University of Prince Edward Island Collection) for providing entomological collection records. We also thank J. R. Spence for specimens, discussions and taxonomic identification of specimens. This manuscript was improved by the comments from J. R. Spence and an anonymous referee. This work was supported by a NSERC scholarship to M.-C.G. and a NSERC research grant to J.T.

REFERENCES

Abe, T.A., Spence, J.R. & Sperling, F.A.H. (2005) Mitochondrial introgression is restricted relative to nuclear markers in

a water strider (Hemiptera: Gerridae) hybrid zone. *Canadian Journal of Zoology*, **83,** 432–444.

Andersen, N.M. (1993) Classification, phylogeny, and zoogeography of the pond skater genus *Gerris* Fabricius (Hemiptera: Gerridae). *Canadian Journal of Zoology*, **71,** 2473–2508.

Bernatchez, L. & Wilson, C.C. (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, **7,** 431–452.

Brooks, A.R. & Kelton, L.A. (1967) Aquatic and semiaquatic Heteroptera of Alberta, Saskatchewan, and Manitoba (Hemiptera). *Memoirs of the Entomological Society of Canada*, **51,** 1–92.

Callahan, J.R. (1974) Observations on *Gerris incognitus* and *Gerris gillettei* (Heteroptera: Gerridae). *Proceedings of the Entomological Society of Washington*, **76,** 15–21.

Cheng, L. & Fernando, C.H. (1970) *The water-striders of Ontario (Heteroptera: Gerridae)*. Life Science Miscellaneous Publications, Royal Ontario Museum, Toronto, ON, Canada.

Clark, A.G. (1990) Inference of haplotypes from PCR-amplified samples of diploid populations. *Molecular Biology and Evolution*, **7,** 111–122.

Crandall, K.A. & Templeton, A.R. (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics*, **134,** 959–969.

Damgaard, J. (2006) Phylogeny and mtDNA phylogeography of two widespread European pond skater species (Hemiptera–Heteroptera: Gerridae: *Gerris* Fabricius). *Insect Systematics and Evolution*, **37,** 335–350.

Damgaard, J. & Sperling, F.A.H. (2001) Phylogeny of the water strider genus *Gerris* Fabricius (Heteroptera: Gerridae) based on COI mtDNA, EF-1α nuclear DNA and morphology. *Systematic Entomology*, **26,** 241–254.

Dennis, R.L.H. & Thomas, C.D. (2000) Bias in butterfly distribution maps: the influence of hot spots and recorder's home range. *Journal of Insect Conservation*, **4,** 73–77.

Drake, C.J. & Harris, H.M. (1934) The Gerrinae of the western hemisphere. *Annals of the Carnegie Museum*, **23,** 179–240.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1,** 47–50.

Ferrier, S. (2002) Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? *Systematic Biology*, **51,** 331–363.

Galacatos, K., Cognato, A.I. & Sperling, F.A.H. (2002) Population genetic structure of two water strider species in the Ecuadorian Amazon. *Freshwater Biology*, **47,** 391–399.

Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359,** 183–195.

Hijmans, R.J., Garrett, K.A., Huaman, Z., Zhang, D.P., Schreuder, M. & Bonierbale, M. (2000) Assessing the geographic representativeness of Genebank collections: the case of Bolivian wild potatoes. *Conservation Biology*, **14,** 1755–1765.

Kadmon, R., Farber, O. & Danin, A. (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, **14,** 401–413.

Lavergne, S. & Molofsky, J. (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences USA*, **104,** 3883–3888.

Lee, C.E. (2000) Global phylogeography of a cryptic copepod species complex and reproductive isolation between genetically proximate 'populations'. *Evolution*, **54,** 2014–2027.

Lenoir, J., Gegout, J.C., Marquet, P.A., De Ruffray, P. & Brisse, H. (2008) A significant upward shift in plant species optimum elevation during the 20th century. *Science*, **320,** 1768–1771.

Lomolino, M.V., Riddle, B.R. & Brown, J.H. (2006) *Biogeography*, 3rd edn. Sinauer Associates, Sunderland, MA.

Maw, H.E.L., Foottit, R.G., Hamilton, K.G.A. & Scudder, G.G.E. (2000) *Checklist of the Hemiptera of Canada and Alaska*. National Research Press, Ottawa, ON, Canada.

Meier, R. & Dikow, T. (2004) Significance of specimen databases from taxonomic revisions for estimating and mapping the global species diversity of invertebrates and repatriating reliable specimen data. *Conservation Biology*, **18,** 478–488.

Nei, M. (1987) *Molecular evolutionary genetics*. Columbia University Press, New York.

Olsen, K.M. & Schaal, B.A. (1999) Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences USA*, **96,** 5586–5591.

Parmesan, C. & Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421,** 37–42.

Petit, R.J., Brewer, S., Bordacs, S. *et al.* (2002) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156,** 49–74.

Reznicek, A.A. (1994) The disjunct coastal-plain flora in the Great Lakes region. *Biological Conservation*, **68,** 203–215.

Ricklefs, R.E. (2004) A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, **7,** 1–15.

Rondinini, C., Wilson, K.A., Boitani, L., Grantham, H. & Possingham, H.P. (2006) Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology Letters*, **9,** 1136–1145.

Rosenzweig, M.L. (1995) *Species diversity in space and time*. Cambridge University Press, Cambridge.

Sambrook, J. & Russell, D. (2001) *Molecular cloning: a laboratory manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. & Flook, P. (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene-sequences and a compilation of conserved polymerase chain-reaction primers. *Annals of the Entomological Society of America*, **87,** 651–701.

Smith, C.L. (1988) Family Gerridae Leach, 1815: the water striders. *Catalog of the Heteroptera, or true bugs, of Canada and the continental United States* (ed. by T.J. Henry and R.C. Froeschner), pp. 140–151. E. J. Brill, Leiden.

Spence, J.R. (1983) Pattern and process in co-existence of water-striders (Heteroptera, Gerridae). *Journal of Animal Ecology*, **52,** 497–511.

Spence, J.R. & Carcamo, H.A. (1991) Effects of cannibalism and intraguild predation on pondskaters (Gerridae). *Oikos*, **62,** 333–341.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.G. & Cosson, J.F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7,** 453–464.

Tajima, F. (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics*, **105,** 437–460.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24,** 1596–1599.

Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22,** 4673–4680.

Turgeon, J., Stoks, R., Thum, R.A., Brown, J.M. & McPeek, M.A. (2005) Simultaneous Quaternary radiations of three damselfly clades across the Holarctic. *The American Naturalist*, **165,** E78–E107.

Vieites, D.R., Min, M.S. & Wake, D.B. (2007) Rapid diversification and dispersal during periods of global warming by plethodontid salamanders. *Proceedings of the National Academy of Sciences USA*, **104,** 19903–19907.

Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R.J., Watson, J.E.M. & Willis, K.J. (2005) Conservation biogeography: assessment and prospect. *Diversity and Distributions*, **11,** 3–23.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Distribution records of *Gerris pingreensis*, *Gerris gillettei* and *Gerris incognitus*.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCHES

Marie-Claude Gagnon is a PhD student working on historical and current aspects of dispersal in *G. pingreensis*/*G. gillettei*, notably on the importance of sexual conflict in the reproductive success of migrants.

Julie Turgeon is a professor at the Université Laval (Canada). Her research interests are centred on the phylogeography and population genetics of aquatic organisms.

Editor: Bradford Hawkins