NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents

ISBN 0-315-59144-7

Genetic population structure of the stream dwelling waterstrider, *Gerris remigis* (Hemiptera: Gerridae)

Richard F. Preziosi

A Thesis

in

The Department

of

Biology

Presented in Partial Fulfilment of the Requirements
for the Degree of Master of Science at
Concordia University
Montreal, Québec, Canada

May 1990

ABSTRACT

Genetic population structure of the stream dwelling waterstrider, *Gerris remigis* (Hemiptera: Gerridae).

Richard F. Preziosi

Gerris remigis populations were genetically characterized at 15 loci using starch gel electrophoresis. Sampling over two years was designed for a hierarchical analysis of population structure incorporating variation among sites within streams, streams within watersheds, watersheds within regions, and regions within North America. Hierarchical F statistics indicated a high level of genetic isolation of subpopulations. Only sites within streams maintained enough gene flow to prevent differentiation through drift. Comparison of genotype frequencies between years indicated a possible bias produced by the sampling of sibling groups. Number of alleles per locus and expected heterozygosity did not differ between the group of sites containing long-winged individuals (California region) and the group of sites that did not contain long-winged individuals. However, F statistics calculated to compare sites with winged individuals and sites without showed a lower level of genetic differentiation for the former. Neither group maintained enough gene flow to prevent genetic differentiation by drift. Overall the genetic population structure of *G. remigis* seems to be dominated by founder effects and population bottlenecks occurring within the context of highly restricted gene flow between local

population subgroups. Previous assignment of subspecific status to Californian *G. remigis* is not supported by genetic distances between those populations and other populations in North America. Previous assignment of specific status to south-eastern *G. remigis* is supported by genetic distances between North Carolina populations and other populations in North America, and a high proportion of region specific alleles in the North Carolina populations.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Daphne Fairbairn, without whose help and support in myriad ways, this thesis would not have been possible. I greatly appreciated the support and criticisms of colleagues, friends and teachers, especially Bobbi Brown, Margaret Cooke, Nancy Ennis, Luc-Alain Giraldeau, Gaby Grad, Sharon Harney, Suzanne Plante, and Jennifer Templeton. Dr. Derek Roff assisted in the collection and shipping of samples from California and North Carolina. Dr. D. L. Swofford provided a copy of the BIOSYS-1 program and Dr. R. S. Waples provided a copy of the TEMPTEST program. The comments of Dr. E. Maly and Dr. L-A. Giraldeau have helped to improve this thesis. Special thanks go to my parents for moral and financial support, to Deb Turnbull for having been there, and to my sister Anne for enduring many late night phone calls about nothing in particular.

Thanks also to the McGill University Gault Estate, the Iowa Department of Conservation, the Parc National du Mont-Saint-Bruno, and Parc Mont-Tremblant, all of whom provided permits for sampling.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

"It is clear that descriptions of the genetic variation in populations are the fundamental observations on which evolutionary genetics depends."

Lewontin, 1974, p.19.

INTRODUCTION.

Theoretical background

The genetic structure of populations is defined by the frequencies of alleles in population subunits and the rate at which genetic material is transferred between these subunits (gene flow or migration). An initial problem in determining the genetic structure of a population is to define the size of group that is able to become genetically different from other such groups. Such a group may be termed a neighborhood. Gene frequencies of neighborhoods may be altered by two major forces, selection and genetic drift (random sampling of the gametes transferred from generation to generation). If selection is acting locally on the population, the neighborhood is the smallest group able to track this selection (Endler, 1979). If there is no selection acting on the population then the neighborhood is the smallest group that can randomly diverge from other such groups (Endler, 1979). The neighborhood is therefore the smallest unit in which differentiation, by selection or genetic drift, can take

place. As a result, neighborhood size governs the amount of microgeographic variation (Coyne et al, 1987). Thus any science that has an interest in the adaptation or evolution of organisms has a basis in the genetic structure of populations (Barrowclough, 1980), especially in the sense of defining the size of population subunits that may evolve independantly.

Assessment of genetic population structure is often made in terms of population subunits (sub-populations, demes) that may not necessarily be directly equivalent to neighborhoods. While methods exist to determine neighborhood size in continuous populations, populations that consist of groups of isolated or semi-isolated demes have no equivalent measure and population structure is reported in terms of the population subunits themselves.

The assessment of genetic population structure involves the quantification of both allele frequencies and the rate of transfer of genetic material between population subunits (gene flow). Starch gel electrophoresis provides a relatively easy method of quantifying allele frequencies using large enough sample sizes for statistical testing. The quantification of gene flow is a much more difficult task. Gene flow is often measured in terms of N_em, the product of the effective number of individuals in a subunit (N_e) and the proportion of migrants to the subunit per generation (m). The effective population number may be defined as "the number of individuals that would give rise to the calculated sampling variance (of

the existing population) ... if they bred in the manner of the idealized population" (Falconer, 1989, p.70), and is affected by "...any factor that puts constraints on the random sampling of gametes from the parental population or on their random union..." (Chepko-Sade et al, 1987, p.288). A migrant to the population subunit is defined as an emigrant from another subunit that successfully contributes genetic material to the next generation. Since both migration and effective population number are difficult to measure, an estimate is often made of the product ($N_e m$), which is the number of migrants per generation.

The most direct way to measure migration of individuals between population subunits is by mark-recapture methods. These methods may overestimate gene flow by including the movement of individuals who do not contribute genetic material to the next generation, or may underestimate gene flow through the movement of unmarked individuals or life stages (Slatkin, 1981). All other methods of measuring gene flow are indirect and are based on some measure of the similarity of the genetic material of population subunits. Often these methods depend on biochemical data obtained by electrophoresis. An advantage of indirect methods is that they provide information on historic as well as current gene flow and thus may provide more relevant estimates for evolutionary studies (Johnson et al, 1988).

Allele frequencies obtained from electrophoretic data are

treated in several ways to obtain information about genetic population structure. Wright's F statistics are commonly used measures of the genetic divergence of a group of population subunits, or more correctly, the degree to which the subunits have been fixed (reduced to a single allele at a locus). F statistics can be calculated from estimates of heterozygosity (Hartl, 1987, p. 78). The fixation index is the reduction in heterozygosity in a population subunit due to drift: $F_{ST} = (H_T - H_S) / H_T$, where H_T, the total heterozygosity, is the expected frequency of heterozygotes in a panmictic population, and H_S, the subunit heterozygosity, is the average expected frequency of heterozygotes in panmictic subunits. F_{ST} can also be thought of as "the amount of genetic differentiation among sub-populations relative to a hypothetical group of sub-populations, each homozygous, but having the same overall average allele frequency as the real sub-populations" (Hartl, 1980, p.164). Thus F_{ST} gives a measure of " the degree of completion of the process of fixation, not absolute differentiation" (Wright, 1978).

The degree of fixation of alleles in population subunits is related to the amount of gene flow among population subunits, and it can be shown that $F_{ST} = 1 / (1 + 4N_em)$ (Wright, 1969, p.291). This relationship, based on an island model of population structure, assumes a large number of demes and a small mutation rate (Maynard Smith, 1989, p.160). The island model is one in which several population subunits exist and

the probability of exchange of migrants is the same for all pairs of subunits. Wright (1931) has shown that for an island model, genetic divergence due to drift can occur if $Nm < 1$.

Two other models developed by Wright (1969) are more applicable to natural populations. The stepping stone model is one in which gene flow is most likely to occur between geographically close population subunits. The continuous model is one in which individuals are evenly distributed and not confined to discrete habitats. When discussing the continuous model, Wright (1969, p. 295) replaces effective population size with effective neighborhood size, defined as "an area from which the parents of central individuals may be treated as if drawn at random". For two dimensional versions of the stepping stone and continuous models, $Nm > 1$ still maintains high genetic similarity. Populations structured as linear models are more likely to diverge genetically under this level of gene flow (Maynard Smith, 1989, p.161). In considering several types of population structures, Slatkin (1985) has stated that, in general, exchange of an individual every second generation should prevent genetic differentiation by drift.

An alternate method for estimating the level of gene flow between populations has been developed by Slatkin (1981,1985). This method is based on the frequencies of alleles found exclusively in one deme (referred to as private alleles). This method is based on a regression of private allele

frequencies on gene flow. Slatkin used simulations of populations with known gene flow levels to generate a regression of private allele frequencies on gene flow. The resulting regression equation is $\ln(p(1)) = -0.505 \ln(Nm) - 2.440$, where $p(1)$ is the average frequency of private alleles. Since the simulations were all run with a sample size of 25, Slatkin (1985) suggests using a correction factor of the sample size over 25. Slatkin and Barton (1989) found that F_{ST} and private allele methods of estimating Nm are equivalent under a wide variety of conditions but that practical problems make F_{ST} the more useful measure.

F statistics, while important for examining the genetic structure of populations, are not suitable for measuring the degree of genetic differentiation between pairs of population subunits since the formulae used for F statistics would provide coefficients for pairwise comparisons that would not sum to the total coefficient for all groups (Wright, 1978, p.89). Several measures have been developed to estimate the genetic divergence of pairs of population subunits. The most common of these measures is Nei's genetic distance (Nei, 1972), which is based directly on the allele frequencies in each population subunit. If an allele is found at a frequency of p_x in population subunit X, and at frequency p_y in population subunit Y, then the probability of getting 2 randomly chosen alleles which are the same is p_x^2 if both are drawn from population subunit X, p_y^2 if both are drawn from

population subunit Y, and $p_x p_y$ if they are drawn from different population subunits. The normalized identity of alleles from X and Y is Nei's index of similarity, $I = p_x p_y / SQRT (p_x^2 * p_y^2)$. The equation can be expanded to cover all loci if each probability term is averaged over all loci. Nei's corresponding distance measure is $D = -\log_e I$ (Nei, 1972). A second common measure is Roger's genetic distance ($D=[\frac{1}{2}\Sigma(q_x-q_y)^2]^{\frac{1}{2}}$) (in Wright, 1978). Roger's g[illegible] distance is better suited for the construction of cluster analysis dendrograms, used to group population subunits based on their relative genetic distances. The reason for this is that for Roger's distance, but not Nei's distance, subunits can be plotted in Euclidian hyperspace with axes based on the allele types present. The distance between any two population subunits can be determined from the scores on the axes by means of the expanded Pythagorean theorum (Wright, 1978, p.90).

Literature review

It is clear from the literature that the extent of genetic differentiation among population subunits varies greatly (for a table of F_{ST} values for insects see McCauley and Eanes, 1987). Gene flow between populations can maintain genetic similarity but the degree to which this actually occurs in nature is still unknown. Ehrlich and Raven (1969) suggest that gene flow is rare and that selection acts on each part of a species range independently. Mayr (1963) suggested that

gene flow is common and maintains genetic coherence within a species. Most likely, as Slatkin (1981) says "...both of these views are probably correct for some species...".

Regardless of which force is dominating the genetic population structure, the most important population parameters are population subunit size and level of gene flow. The effect of these parameters can best be seen in studies where comparisons are made between subdivided and continuous populations. Pounds and Jackson (1981) examined the effect of division of populations of the eastern fence lizard (*Sceloporus undulatus*) by rivers. They found greater morphological differentiation among populations divided by rivers than among equivalent populations not divided by rivers. King (1987) found that populations of the beetle, *Collops georgianus* on 'islands' of granite outcrops that were isolated from other nearby outcrop 'islands' did not have significantly larger genetic distances than undivided populations. However the genetic distance between 'far disjunct pairs' of outcrops (average distance 520m) was significantly greater than genetic distance between 'disjunct' pairs of outcrops (average distance 85m).

While it may be possible to estimate effective population size and gene flow levels in continuous populations it is certainly easier for isolated or semi-isolated populations. McCauley and Eanes (1987) found that the milkweed beetle (*Tetraopes tetraopthalmus*) has a partially isolated population

structure because of specificity for its host plant (*Asclepias syriaca*) and low vagility of individuals. The genetic differentiation for these beetles is much larger on a macrogeographic scale (F_{ST}=.172 for north-eastern and north-central U.S.) than on a local scale (average F_{ST}=.026 within states). Selander and Kaufman (1975) studied the genetic population structure of the brown snail (*Helix aspersa*) on two adjacent residential blocks (area approx. 10 acres each) in Bryan, Texas, and found inter-block variation (F_{ST}=.11613) to be far greater than intra-block variation (F_{ST}=.03367). This would indicate that the subunit size able to differentiate genetically in this species is a city block or smaller in the area studied.

There is evidence that the apparent dispersal ability of an organism will be reflected in the genetic population structure (Murray and Clarke, 1984; Waples, 1989). Other studies suggest that apparent dispersal ability may be misleading as far as gene flow is concerned (Liebherr, 1986). Baker (1981) examined the spread of an introduced allele in a population of house mice. Previous studies, genetic and otherwise, had suggested that house mice lived in small, isolated groups with very little or no gene flow between these groups. Baker introduced mice with an allele not present on the existing population into two coops on a poultry farm. Two generations later the introduced allele was still present in the coops where introductions had been made and was present in two

other coops as well (Baker, 1981). Baker's work emphasizes the importance of studies that use both genetic and non-genetic data in evaluating gene flow.

Organisms which inhabit lotic (flowing water) habitats are an ideal example of isolated or semi-isolated populations. Animals may move freely within a watershed, but would not be expected to move between watersheds. This type of genetic population structure has been found in newts (Hedgecock, 1978) and fish (Parkinson, 1984; Foote et al, 1989).

Waterstriders (Hemiptera: Gerridae) are semiaquatic true bugs that live on the water surface of ponds, pools, lakes or streams (Andersen, 1982). The advantage of using waterstriders for studies of genetic population structure is that these species come very close to theoretical designs. Population sites are discrete and well defined, and proportions of migrants can be predicted from the proportion of long-winged individuals (Vepsalainen, 1978). From a practical aspect, locating population sites is relatively easy as ponds and streams are often recorded on local area maps, and the capture of the insects is not difficult.

There have been several European studies of the genetic population structure of waterstrider species. Varvio-Aho (1981) related the degree of allelic variation to 'ecological differences' in Finnish waterstrider species. The 'ecological differences' were gene flow and population size as estimated by abundance, habitat stability, and dispersal ability.

Varvio-Aho (1981) found Finnish waterstriders to have 'maximal' values of genetic distance. Other studies in Finland have found varying degrees of population differentiation for waterstrider species and have indicated different population structures for the same species in different geographic areas (Varvio-Aho, 1979; Varvio-Aho and Pamilo, 1980).

Several aspects of the dispersal ability of the waterstrider *Gerris remigis* have been examined. *G. remigis* is found on the surface of streams and is the most common and abundant waterstrider in North America (Drake and Harris, 1928; Polhemus and Chapman, 1979). In Canada *G. remigis* are univoltine or partially bivoltine, overwinter as adults and reproduce in spring (Matthey, 1974; Galbraith and Fernando, 1977; Fairbairn, 1985a). In most populations, *G. remigis* is almost completely wingless with less than 1% winged individuals (Calabrese, 1979, for exceptions see Froeschner, 1962 and Fairbairn, 1985a). Fairbairn (1986) found that long-winged *G. remigis* in populations where winged individuals are rare "do not disperse using any means unavailable to the apterous (wingless) morph", although the possibility of rare long distance dispersal by flight was not ruled out (estimated as 0.03%/population/generation). Fairbairn and Desranleau (1987) found that *G. remigis* in the lab have a very high flight threshold and that the majority of overwintered, long-winged individuals histolyse their flight muscles in the early

spring. They concluded that the long-winged *G. remigis* fly only rarely, if at all. Bowdan (1978) examined locomotion in *G. remigis* and determined that, while *G. remigis* has become very coordinated at rowing, they have lost their coordination in walking and, in fact, attempt to row even on hard dry surfaces. This implies that *G. remigis* must move primarily along the water surface. Thus, for the majority of *G. remigis*, dispersal can only occur between sites connected by water. Gene flow along a stream is restricted to active dispersal of adults or passive downstream drift of adults and nymphs (Fairbairn, 1985a).

Fairbairn (1986) conducted an intensive mark-recapture study of movement of adult *G. remigis* at a site in Québec. The greatest individual displacement occurred in spring and, even at this time, the net displacement of an individual was unlikely to be more than 100 meters. Although individuals moved upstream and downstream with equal frequency, distances moved were greater in the upstream direction. Of 4828 *G. remigis* marked by Fairbairn (1986) only 2 individuals moved between streams. This indicates that movement of *G. remigis* between streams, even within the same watershed, is quite low.

Examination of the genetic population structure of G. remigis is confined to a single study. Zera (1981) compared the genetic population structure of *G. remigis* and *Limnoporus canaliculatus* in the eastern U.S. *L. canaliculatus* is wing polymorphic and showed much less genetic differentiation than

G. remigis. In *G. remigis*, five of six polymorphic loci showed significant differences in allele frequencies among populations. At two of these loci, some populations were fixed for different alleles. Zera concluded that *G. remigis* is a highly isolated, 'island' species, able to diverge genetically by selection or drift. None of Zera's (1981) sites were connected by water and thus no information is available on genetic variation within a watershed for *G. remigis*.

If migration between population subunits is restricted, then differentiation among subunits may be enhanced by local extinction and recolonization or by bottlenecks caused by overwinter mortality (Zera, 1981) which may be as large as 90% for *G. remigis* (Matthey, 1974). Yearly recolonization of sites within streams is likely to cause the loss of rare alleles at a site through founder effects (or less likely, may inflate the frequency of rare alleles). The effect of these factors was best described by Sewall Wright:

> "This bottle neck effect is greatest in cases in which the total population consists of small demes, each likely to become extinct after a few generations but, if so, always replaced sooner or later by a few stray migrants from populations that have persisted. In this way, every deme at any given time has a history of passage through a great number of bottlenecks of small numbers on being

> traced back from place to place, and since a few momentarily flourishing demes may be the source from which many new colonies are founded, large areas or even the whole species may, in the course of time, trace to a single deme that has passed through many bottlenecks..." (Wright, 1969).

It should be noted that if yearly extinctions and recolonizations occur, then Slatkin's (1985) method of estimating gene flow (Nm) may estimate N(m+e) where e is the extinction rate (Slatkin, 1985).

Gerris remigis varies greatly in morphology, aptery, and voltinism across its range (Drake and Harris, 1928; Calabrese, 1974; Polhemus and Chapman, 1979). Calabrese (1974) has suggested that western *G. remigis* are a separate subspecies (*G. remigis caloregon*) from eastern *G. remigis*. This division is based on differences in the proportion of individuals that are long-winged and in the morphology of the male genitalia (Michel, 1961). Schaefer and Calabrese (1980) have shown that the other member of the amphi-atlantic species pair, *G. najas*, is morphologically closer to eastern *G. remigis* than eastern *G. remigis* is to western *G. remigis*. Polhemus and Chapman (1979) do not agree with the assignment of subspecific status to western *G. remigis* and indicate that if such an assignment were made it should be *G. remigis orba*, as named by Stal who first recorded such a separate type for California in 1859.

G. orba was later regrouped with *G. remigis* by Drake and Harris (1928). Michel (1961) found *remigis* from Virginia, Texas and Arizona, to be distinguishable in size, color, and male genitalia from other *G. remigis*, and suggested specific status for populations in the south. Calabrese (1974) also notes distinguishable morphological characteristics for Virginia populations of *G. remigis*.

Objectives

The purpose of this study is to assess the genetic population structure of *G. remigis* on a local and regional scale using electrophoretic analysis. The expected population structure in eastern North America is one in which allelic variation will exist within a watershed but will be much smaller than variation among watersheds or regions. The genetic structure of the population depends, at least partially, on migration. Waterstriders from any one region are expected to be more similar to each other than to waterstriders from another region. Since wingless *G. remigis* move almost exclusively along the water surface, watersheds are the next most probable level of isolation below that of regional isolation. Zera (1981) has determined that *G. remigis* populations in the east are highly differentiated between drainages but no information is available on genetic variation within drainages. A high degree of isolation at the level of streams within watersheds was implied by the

dispersal distances of *G. remigis* measured by Fairbairn (1985a). Based on Fairbairn's (1985a,b;1986) mark-recapture work, we might expect genetic differentiation to occur even within a stream if distances are large enough. If differences do exist among sites within a stream, they are expected to be much smaller than differences at any other level.

The isolation of populations may depend on the degree of aptery as it affects dispersal ability. Calabrese (1974), has shown that aptery is not always the case for *G. remigis* populations, especially in the southwest. Thus variation among *G. remigis* populations in California is expected to be smaller than variation among eastern *G. remigis* populations. Comparison of variation in eastern and western populations provides a direct test of the hypothesis that the high level of population differentiation seen by Zera (1981) in eastern *G. remigis* is due to reduced dispersal by flight.

METHODS.

Sampling was designed to determine the degree of genetic differentiation at 4 levels of expected isolation: regions within North America, watersheds within regions, streams within watersheds, and sites within streams (figs. 1 to 4 and table 1). With the exception of California, regions were not chosen to reflect any ecological or selective patterns within or between regions but simply as areas of geographic isolation by distance. The California region contained distances between sites that were of the same order as distances among eastern regions. In this case the region was based on an expected genetic differentiation similar to the ecological and morphological differentiation of *G. remigis* in the California region. Figures 2 to 4 illustrate the watersheds of Mont-St-Hilaire, Mont-Tremblant and Mont-St-Bruno respectively. The Thames river watershed sites are on streams that feed into opposite sides of the Thames river at points approximately 28 km apart. The Northern California watershed sites are on streams that feed into opposite sides of the Klamath river at points approximately 25 km apart.

Thirteen sites were sampled for two consecutive years (table 1). Changes in gene frequencies between years provide information on the stability of gene frequencies as well as providing increased sample sizes without decimating populations.

Sampling was conducted in the summer or fall when

Figure 1. Map of North America showing locations of watersheds sampled in this study (black triangles). Site names can be found in table 1.

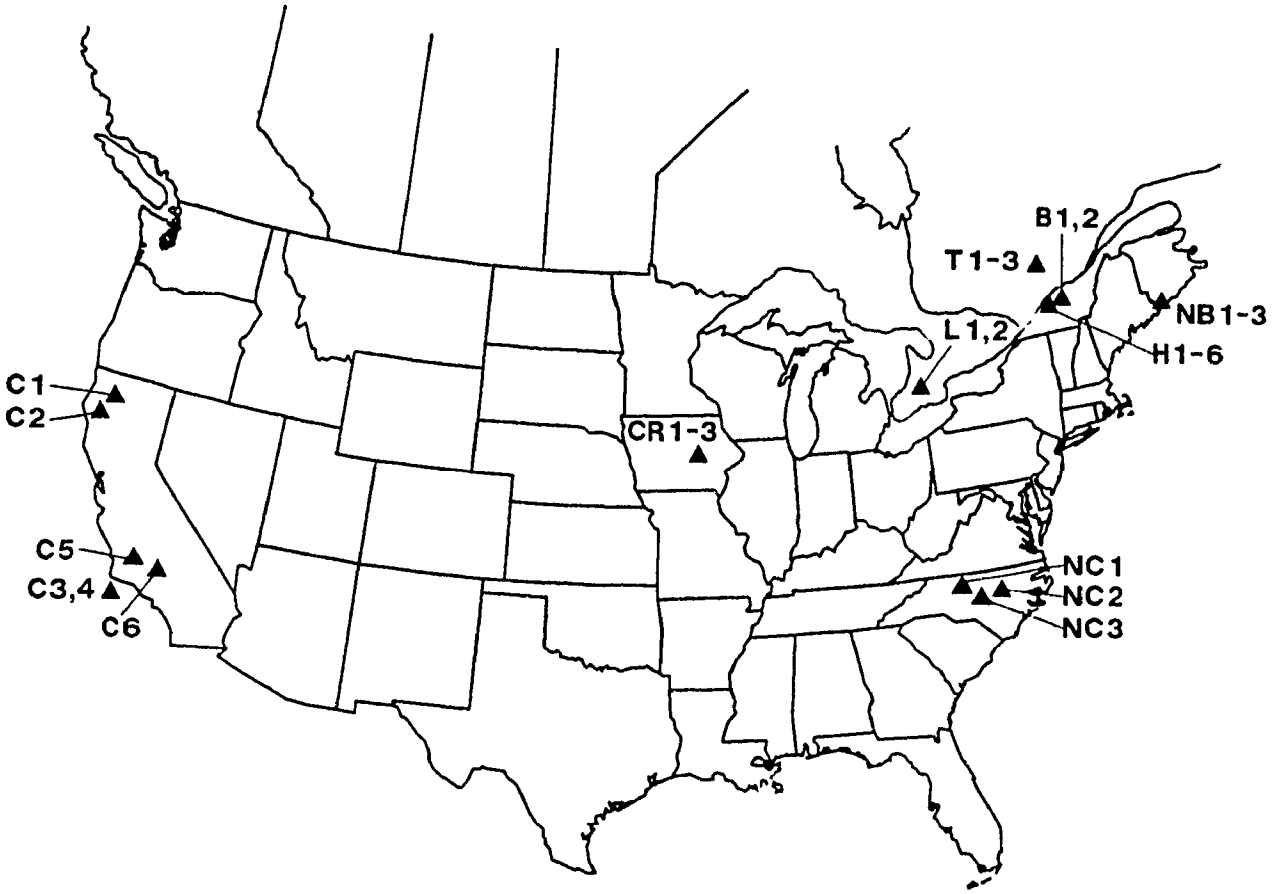

B1,2
T1-3
NB1-3
H1-6
L1,2
C1
C2
CR1-3
C5
C3,4
C6
NC1
NC2
NC3

Figure 2. Map of the Mont-St-Hilaire watershed. Sampling sites are indicated by black triangles. Site names can be found in Table 1.

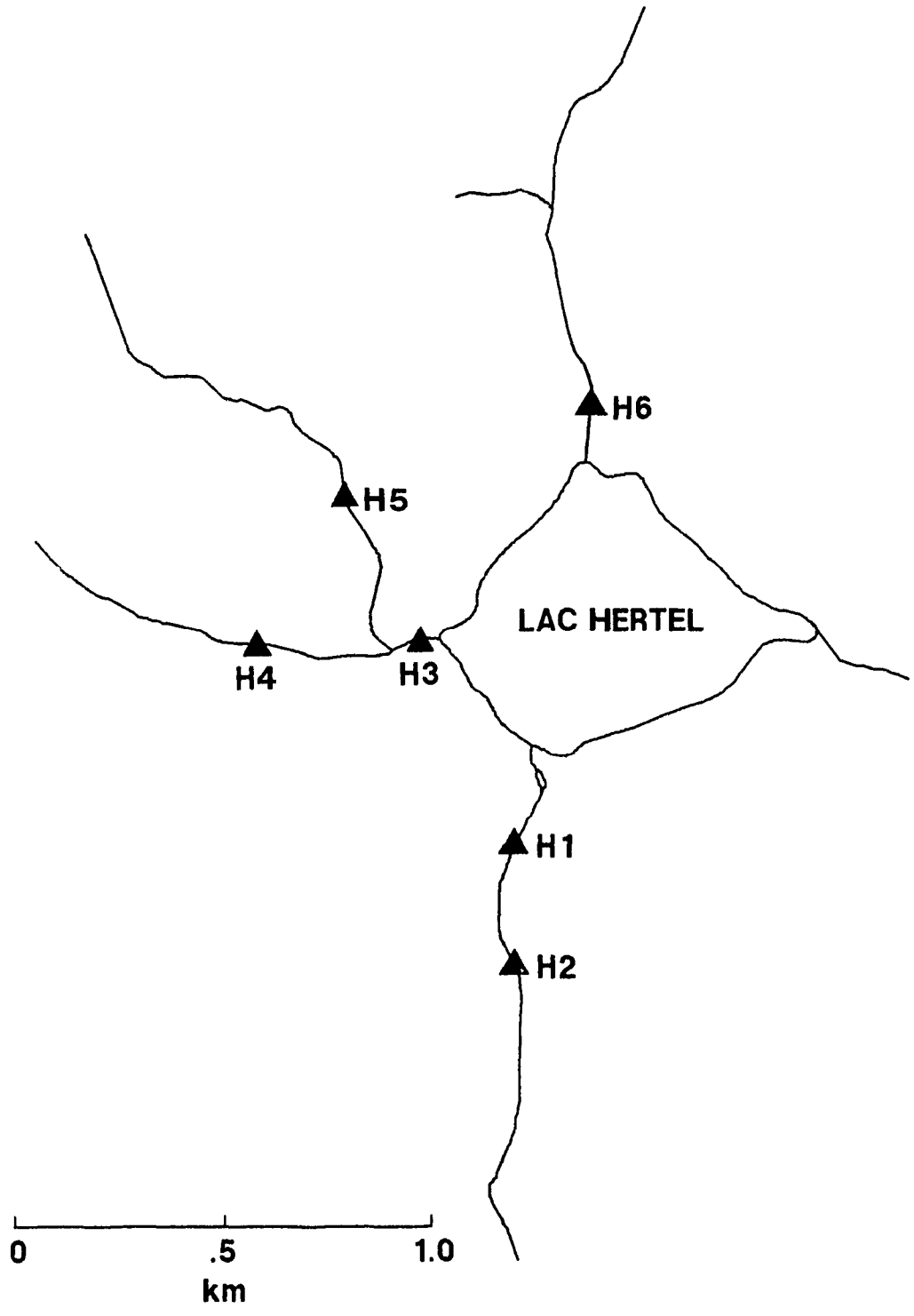

H6
H5
LAC HERTEL
H4
H3
H1
H2
0
.5
1.0
km

Figure 3. Map of Mont-Tremblant watershed. Sampling sites are indicated by black triangles. Site names can be found in Table 1.

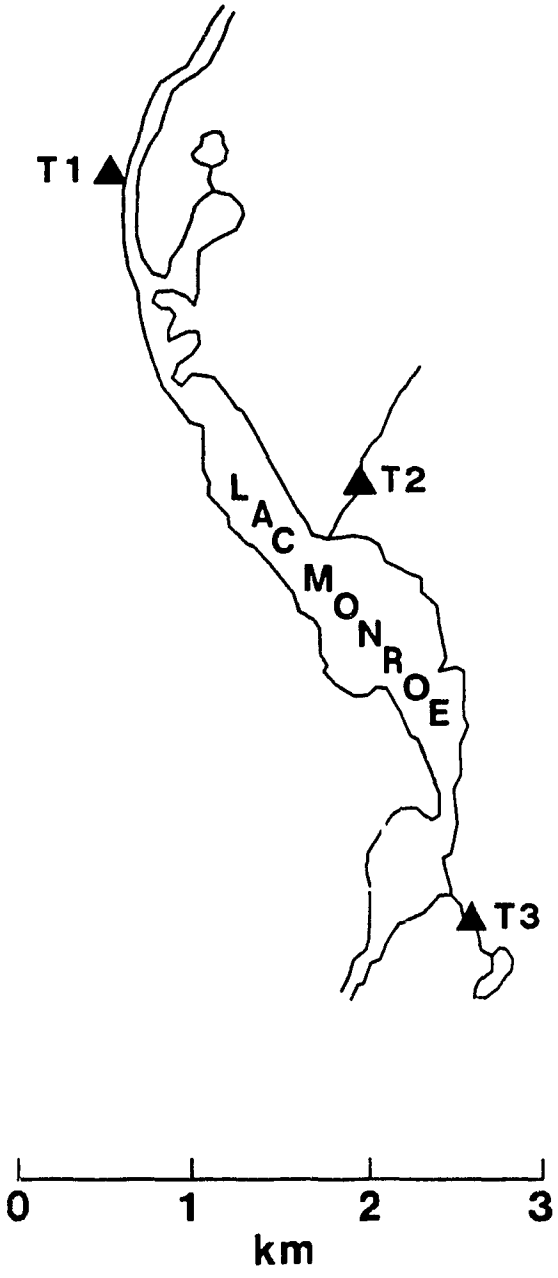

T1
T2
T3
LAC MONROE
0
1
2
3
km

Figure 4. Map of Mont-St-Bruno watersheds. Sampling sites are indicated by black triangles. Site names can be found in Table 1.

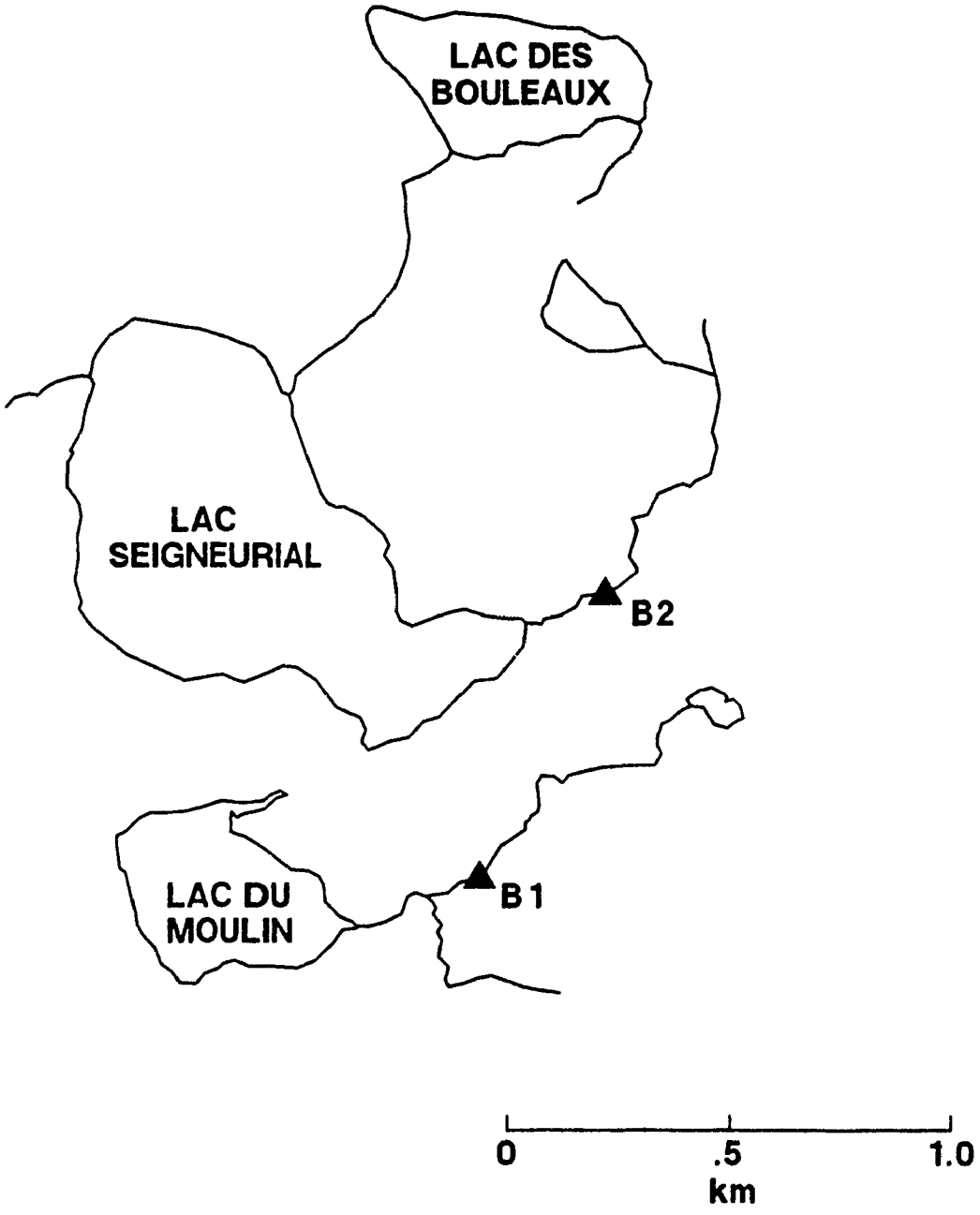

LAC DES
BOULEAUX
LAC
SEIGNEURIAL
B2
LAC DU
MOULIN
B1
0
.5
1.0
km

Table 1. Site names, sample sizes and collection dates of all samples in the hierarchical sampling design.

Region	Watershed	Stream[a]	Site	Sampling Date	Sample Size
Québec	Mont-St-Hilaire	South Creek	H1	08/87	157
				08/88	151
			H2	08/87	201
				08/88	184
		West Creek	H3	08/87	119
				08/88	170
			H4	08/87	167
				08/88	110
			H5	08/87	66
				08/88	160
		North Creek	H6	08/87	124
				08/88	116
	Mont-St-Bruno #1	(Lac Du Moulin)	B1	09/87	38
	Mont-St-Bruno #2	(Lac Seigneurial)	B2	09/87	207
				09/88	157
	Mont-Tremblant	Chutes Croche	T1	09/87	179
				08/88	241
		Ruis. des Aulnes	T2	09/87	226
				08/88	213
		Ruis. Lac des Femmes	T3	09/87	200
Ontario	Thames River	Pottersburg Creek	L1	07/87	109
				08/88	68
		Dingman Creek	L2	08/88	70

Table 1. Continued

Region	Watershed	Stream[a]	Site	Sampling Date	Sample Size
Iowa	Cedar Rapids #1	(Beaver Park)	CR1	07/87	29
				07/88	28
	Cedar Rapids #2	(Ellis Park)	CR2	07/87	98
				07/88	28
	Cedar Rapids #3	Squaw Creek	CR3	07/87	35
New Brunswick	St. Andrew's #1	(Caithness)	NB1	09/87	168
				09/88	165
	St. Andrew's #2	Lelands Creek #1	NB2	09/87	75
		Lelands Creek #2	NB3	09/88	377
North Carolina	Deep River	(Asheboro)	NC1	06/87	21
	Yadkin River	(Tanglewood)	NC2	06/87	15
	Blue Ridge	(Doughton Park)	NC3	06/87	16
California	Klamath River	Beaver Creek	C1	09/88	47
		Scott's River	C2	09/88	133
	Santa Cruz Is. #1	Prisoner's Creek	C3	09/88	142
	Santa Cruz Is. #2	San Justiano Creek	C4	09/88	100
	Los Padres Forest	Sespe Creek	C5	11/88	100
	Angeles Forest	San Gabriel Creek	C6	11/88	100

[a]Where stream names could not be determined a nearby town or landmark is given in parentheses.

population sizes are the largest, because the sampling method was destructive. Adult *G. remigis* were collected from streams using hand nets and kept alive on ice until returned to Concordia University where they were placed individually in .5ml micro-centrifuge tubes and frozen at -60°C. Animals from California were frozen at -60°C in California and shipped to Montreal on dry ice. A goal of 200 individuals per site per year was set but was not always met (table 1). Samples of this size were collected to increase the power of the chi-square tests of goodness-of-fit to Hardy-Weinberg, used to test the underlying genetic models in the absence of breeding data (Fairbairn and Roff, 1980). Large sample sizes also increased the power of tests of independence of allele frequencies between sites and years (but see Waples 1989).

Horizontal starch gel electrophoresis was conducted using the 0.135M tris-citrate buffer of Shaw and Prasad (1970) according to the methods of Zera (1981). The pH of the Tris-citrate buffer was adjusted for each enzyme system (values are given in appendix A). The following 10 enzyme systems were used to detect 15 loci: Alkaline phosphatase (ALP-1,ALP-2, E.C. 3.1.3.1); Malate dehydrogenase (MDH-1, E.C. 1.1.1.37); Phosphogluconate dehydrogenase (PGD-1, E.C. 1.1.1.44); Glutamate-oxaloacetate transaminase (GOT-1,GOT-2, E.C. 2.6.1.1); Glycerol-3-phosphate dehydrogenase (GPD-1, E.C. 1.1.1.8); Glucose-6-phosphate dehydrogenase (GDH-1, E.C. 1.1.1.49); Isocitrate dehydrogenase (ICD-1,ICD-2, E.C.

1.1.1.42); Lactate dehydrogenase (LDH-1,LDH-2, E.C. 1.1.1.27); Malic enzyme (MEZ-1,MEZ-2, E.C. 1.1.1.40), Superoxide dismutase (SOD-1, E.C. 1.15.1.1). Gels were stained for enzymes according to Shaw and Prasad (1970) except for Malic enzyme gels which were stained according to Harrison (1977).

Genotypic frequencies were tested for goodness-of-fit to Hardy-Weinberg equilibrium values by χ^2. Genotypic classes were pooled if the expected value for any class was less than one. If more than 2 alleles were present at a locus, observations were pooled into 3 classes: common homozygotes, heterozygotes containing a common allele, and rare homozygotes and other heterozygotes (Swofford and Selander, 1981). For sites sampled both years, allele frequencies were compared between years using a χ^2 contingency test (SYSTAT, Wilkinson, 1988). If allele frequencies were significantly different between years, Waples (1989) modified χ^2 program (TEMPTEST) was used to determine if changes in allele frequencies could be explained by a combination of drift and sampling error. Since N_e affects the expected change due to drift, a minimum estimate of neighborhood size was calculated from data in Fairbairn (1985a,b). Data from both years were then pooled for further analysis.

Mean heterozygosity and percent polymorphic loci were calculated for each site. Percent polymorphic loci was calculated under the criterion of the most common allele having a frequency not larger than .95. Nei's genetic

distance was used for ease of comparison to other studies. Nei's genetic distances were calculated for all pairs of sites and for all pairs in each level of the hierarchy outlined in table 1. Roger's genetic distance was calculated for all pairs of sites for use in constructing a dendrogram through cluster analysis. Contingency tests of allele frequencies among sites were conducted in a hierarchical manner. Where more than one polymorphic locus was present, χ^2 values and degrees of freedom were combined to obtain an overall probability (Daniel, 1978, p.339) for significant differences of allele frequencies between the sites. Hierarchical F statistics were calculated for each of the levels in table 1. F_{ST} (watersheds within total) values were calculated separately for long-winged and wingless populations for comparison. Cluster analysis was performed on a matrix of Roger's genetic distance between sites using the UPGMA clustering technique. All calculations for the above mentioned analysis, except where noted, were performed using the BIOSYS-1 package of Swofford and Selander (1981).

Matrices of Nei's genetic distances and geographic distances were compared using a Mantel test (Mantel, 1967) following the methods of Sokal (1979) and Manly (1986, pp. 53-57), modified to produce a two tailed probability. A plot of all pairs of genetic and geographic distances is also provided as suggested by Sokal (1979). A Mantel test was also used to compare a matrix of Nei's genetic distance and a binary

connectivity matrix (King, 1988), modified so that tied distances were each given a value of 1. A binary connectivity matrix contains a value of 1 in cells where the two sites are closest neighbors and a value of zero in all other cells. Comparison between a binary connectivity matrix and a matrix of genetic distance tests whether sites are genetically more similar to their closest neighbor than to any other site.

Multidimensional scaling (MDS) (Wilkinson, 1988) was used as an alternate method of examining the association between genetic and geographic distance. MDS is a method of constructing maps from matrices of some measure of similarity or distance (in this case Nei's genetic distance). If an association exists between genetic and geographic distance then the map produced by the multidimensional scaling procedure applied to genetic distance, should resemble the geographic map. While the fit of the multidimensional scaling map to the geographic map is not easily quantified, visual comparison of the two maps provides an assessment of the association between the two distances. This method was sucessfully used by McDonald, Krysan and Johnson (1985) to recreate an approximate map of geographic distribution from a matrix of genetic distances for the adult northern corn rootworm.

RESULTS.

Allele frequencies for all sites are given in appendix B. MDH data were not obtained for site T1 in 1987. Four out of 15 (26.67%) loci examined were monomorphic at all sites. Out of 53 alleles observed over all loci, 24 (45.28%) were unique to a region. In all there were 308 site-locus combinations (15 loci * 28 sites). Of these, 263 (85.39%) were fixed for a single allele, 13 (4.22%) were fixed for an allele found only in that region, and 5 (1.62%) were fixed for an allele found only at that site. Percent polymorphic loci and mean heterozygosity for each site are given in table 2. Percent polymorphic loci did not differ significantly among regions ($F = 1.505$; d.f. = 4,21; $p = 0.237$; based on arcsin transformed data). Expected heterozygosity did differ significantly among regions (Kruskal-Wallis ANOVA: $H = 12.185$, d.f. = 5, $p = 0.032$) but did not differ significantly when the Québec region was excluded from the analysis ($H = 2.245$, d.f. = 4, $p = 0.691$). The highest expected heterozygosities for the Québec region all came from the Tremblant watershed (table 2). If the Tremblant watershed was not included in the Québec region, an analysis of all regions was no longer significant ($H = 9.042$, d.f. = 5, $p = 0.107$), indicating that, in the Québec region, variability in heterozygosity is found at the level of watersheds within the region.

Tests of deviation from Hardy-Weinberg equilibrium values

Table 2. Percentage of loci polymorphic and heterozygosity estimates.

Site	Percentage of loci polymorphic[a]	Mean heterozygosity (SE)	
		Direct count	H-W expected[b]
H1	13.3	.027 (.027)	.028 (.028)
H2	20.0	.024 (.023)	.024 (.022)
H3	13.3	.041 (.033)	.044 (.035)
H4	13.3	.046 (.033)	.050 (.036)
H5	20.0	.053 (.036)	.054 (.038)
H6	13.3	.045 (.034)	.045 (.034)
B1	6.7	.017 (.017)	.020 (.020)
B2	20.0	.018 (.017)	.052 (.037)
T1	20.0	.073 (.043)	.079 (.046)
T2	20.0	.063 (.042)	.070 (.047)
T3	20.0	.069 (.046)	.070 (.047)
L1	20.0	.024 (.019)	.023 (.019)
L2	20.0	.010 (.016)	.009 (.006)
CR1	20.0	.006 (.005)	.041 (.034)
CR2	0.0	.000 (.000)	.000 (.000)
CR3	6.7	.010 (.010)	.009 (.009)
NB1	20.0	.012 (.007)	.021 (.012)
NB2	0.0	.000 (.000)	.000 (.000)
NB3	6.7	.008 (.008)	.008 (.008)
NC1	0.0	.000 (.000)	.000 (.000)
NC2	13.3	.022 (.018)	.027 (.022)
NC3	40.0	.076 (.037)	.210 (.071)
C1	26.7	.020 (.011)	.019 (.011)
C2	26.7	.040 (.031)	.041 (.030)
C3	13.3	.027 (.022)	.025 (.020)
C4	6.7	.006 (.006)	.007 (.007)
C5	20.0	.018 (.017)	.016 (.015)
C6	26.7	.031 (.019)	.032 (.019)

[a]Loci are considered polymorphic if the frequency of the most common allele does not exceed .95.
[b]Mean heterozygosity based on expected Hardy-Weinberg frequencies.

for 1987 data were significant for 5 out of 40 comparisons. Three of these 5 deviants were from site NC3 and showed a heterozygote deficiency. In 1988, 1 out of 54 comparisons was significant. Out of 94 comparisons, approximately 5 would be expected to be significant due to sampling error. Only 6 of 94 comparisons were significant, and since no overall trend in heterozygote deficiency or excess was found for any locus, the basic genetic model of inheritance for the systems analyzed was not rejected.

Comparisons of allele frequencies between years showed significant differences for 8 of 26 comparisons (Table 3). Waples (1989) modified χ^2 test determines if differences in allele frequencies between years can be explained by a combination of sampling error and genetic drift. However the χ^2 statistic produced by Waples test is greatly affected by the neighborhood size. A minimum estimate of neighborhood size in a linear habitat can be calculated from data in Fairbairn (1985a,b) using the formula $N_e = 2 \sqrt{\pi}\ D\ \sigma$ (Wright, 1969, p.303), where D is the density of individuals per unit distance and σ is the standard deviation of movement of individuals along the habitat. Fairbairn (1985a) estimated a minimum density of 1.5 individuals per meter for *G. remigis*. Matthey (1974) found a similar density of .95 individuals per square meter in areas with surface current for beaver ponds in streams in southern Alberta. Fairbairn (1985b) estimated the standard deviation of movement of overwintered individuals at

32 meters. This latter estimate does not include movement in the fall or over the winter and thus is a minimum estimate. The neighborhood size as estimated from these parameters is 170 individuals or approximately 113 meters along a stream.

Waples (1989) states that the modified χ^2 test is applicable only to comparisons where significant differences have been found using an χ^2 contingency test. Thus, Waples (1989) modified χ^2 was conducted for significant comparisons using a minimum estimate of N_e of 170 individuals. All comparisons were still significant (Table 3), indicating that differences could not be explained by the combined effects of sampling error and genetic drift. As an example, the variability of allele frequencies for the PGD locus at Mont-St-Hilaire can be seen in figure 5.

Results of χ^2 contingency tests of independence of allele frequencies among sites at all hierarchical levels indicate that significant differences exist among sites at all levels, even at the level of sites within a stream (Table 4). Examples of the variability among sites for Mont-St-Hilaire and Mont-St-Bruno can be seen in figures 5 and 6 respectively.

Estimates of Nei's genetic distance and Roger's genetic distance between all sites are presented in table 5. Cluster analysis of Roger's genetic distances using a UPGMA algorithm produced the dendrogram shown in figure 7. Since only 15 loci were examined, the length of the branches are not assumed to be precise, however the final clustering pattern is assumed to

Table 3. Summary of allele frequency comparisons between years.

Site[a]	Locus	χ^2	p[b]	p[c]
H1	ALP-1	2.086	.149	-
	PGD-1	1.040	.595	-
H2	ALP-1	0.004	.950	-
	PGD-1	0.554	.457	-
H3	ALP-1	1.061	.304	-
	PGD-1	1.388	.500	-
H4	ALP-1	11.526	.001	<.05
	PGD-1	0.099	.753	-
H5	ALP-1	0.532	.466	-
	PGD-1	8.992	.011	<.05
H6	ALP-1	1.496	.221	-
	PGD-1	3.593	.166	-
B2	PGD-1	15.662	.000	<.05
T1	ALP-1	10.886	.001	<.05
	PGD-1	30.261	.000	<.05
T2	ALP-1	1.365	.243	-
	ALP-2	1.319	.251	-
	PGD-1	1.111	.574	-
L1	ALP-1	1.476	.224	-
	PGD-1	53.399	.000	<.05
	GOT-2	1.888	.169	-
CR1	ALP-1	104.000	.000	<.05
	GPD-1	6.560	.010	<.05
NB1	ALP-1	2.013	.156	-
	ALP-2	2.043	.153	-
	PGD-1	0.315	.575	-

[a]The 13th site, CR2, was fixed for all loci in both years.
[b]Probability level for χ^2 contingency test.
[c]Probability level for Waples (1989) adjusted χ^2 (see text).

Figure 5. Allele frequencies at the PGD-1 locus from Mont-St-Hilaire sites for 1987 and 1988. * $p < 0.05$.

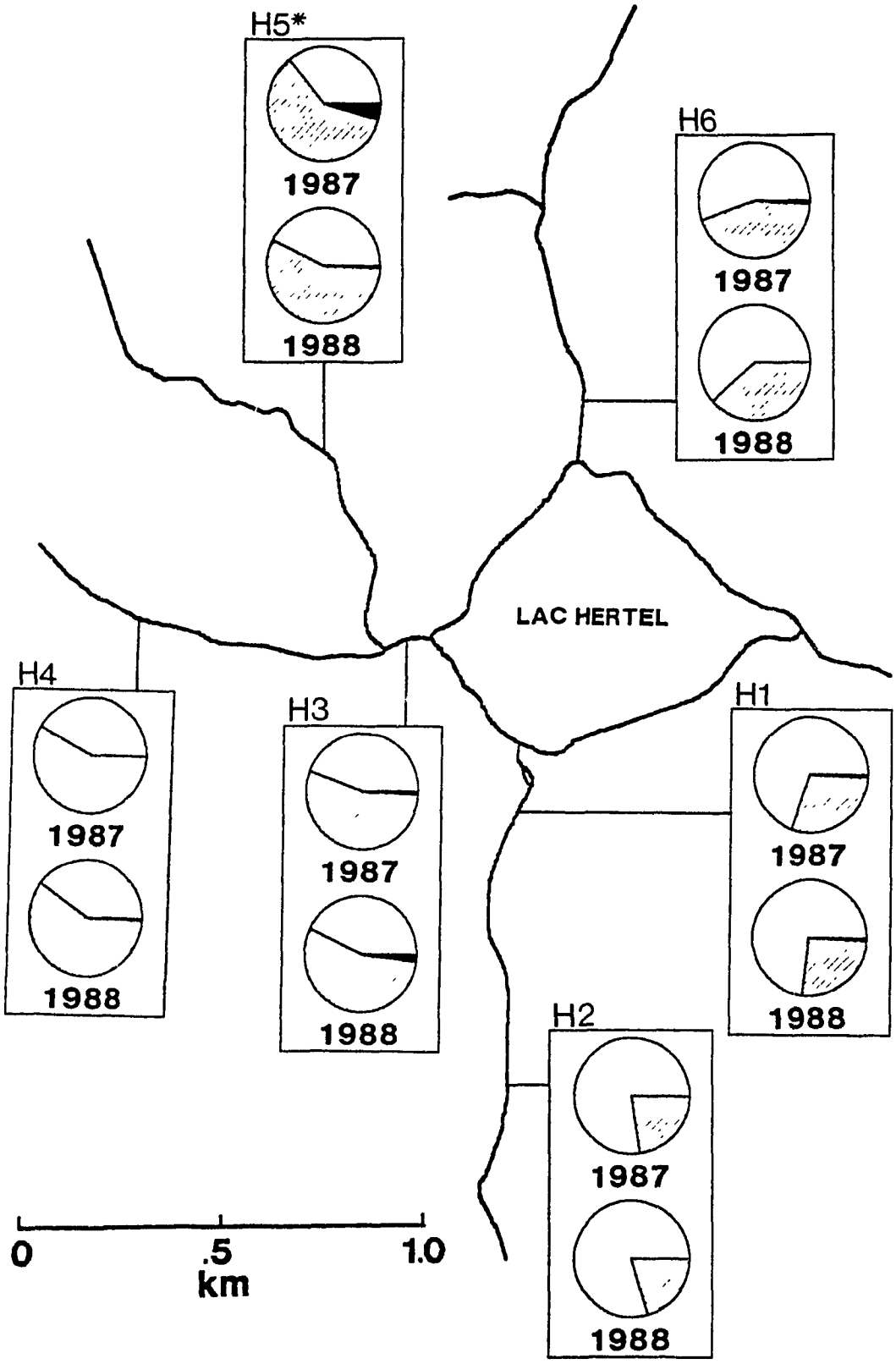
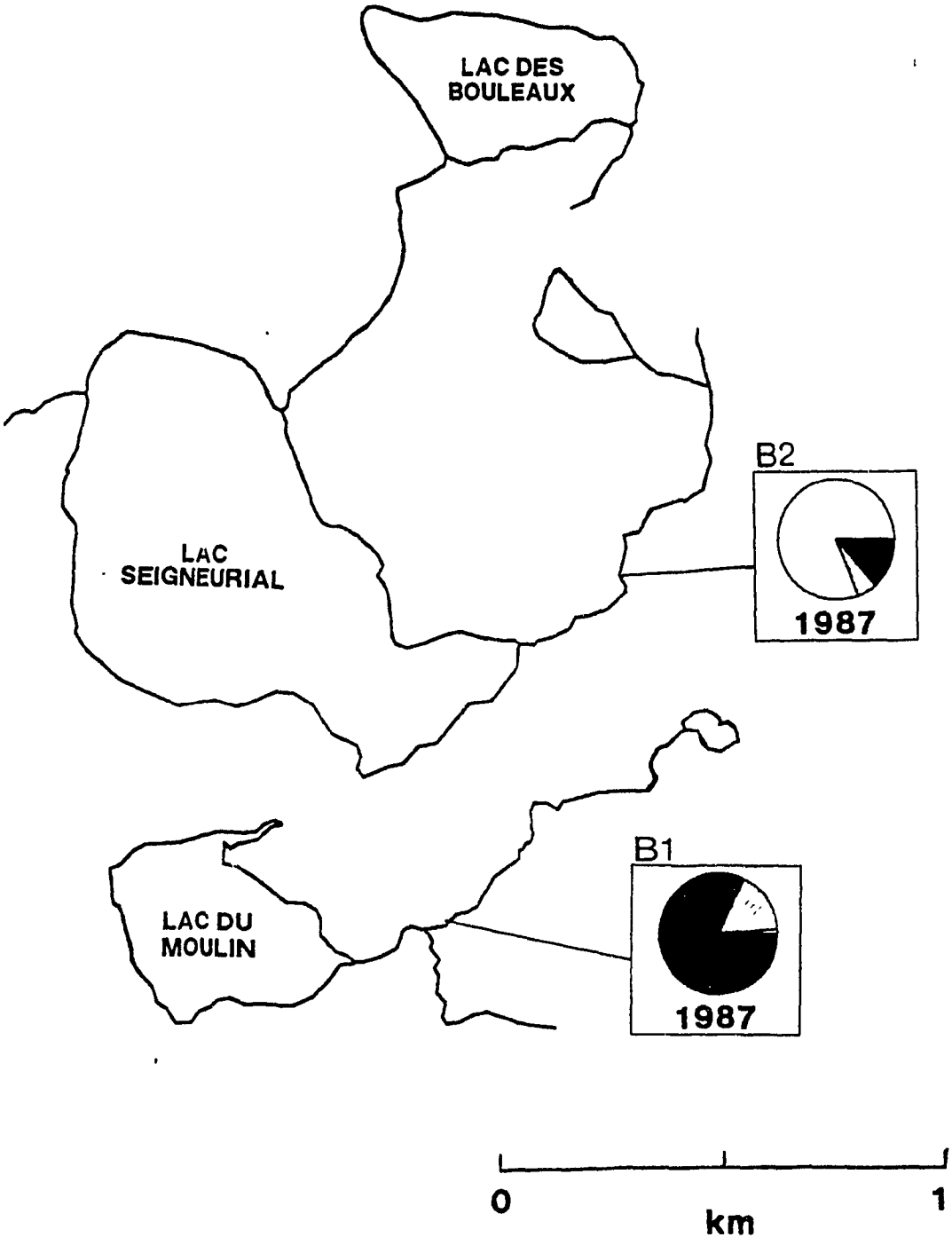

H5*
1987
1988
H6
1987
1988
LAC HERTEL
H4
1987
1988
H3
1987
1988
H1
1987
1988
H2
1987
1988
0
.5
1.0
km

Table 4. χ^2 contingency tests of allele frequencies among sites within areas.

Area	χ^2	d.f.	p
	Overall		
Total	149297.600	999	<.001
	by Region		
Québec	21617.770	90	<.001
Ontario	43.713	6	<.001
Iowa	177.881	8	<.001
New Brunswick	5810.202	18	<.001
North Carolina	474.870	26	<.001
California	1127.390	60	<.001
	by Watershed		
Mont-St-Hilaire	578.622	25	<.001
Mont-Tremblant	1845.320	14	<.001
Thames River	43.713	6	<.001
St. Andrew's #2	1351.664	3	<.001
Klamath River	40.712	7	<.001
	by Stream		
South Creek	9.957	1	.002
West Creek	30.110	6	<.001

Figure 6. Allele frequencies at the PGD-1 locus from Mont-St-Bruno sites for 1987 and 1988.


LAC DES
BOULEAUX
LAC
SEIGNEURIAL
B2
1987
LAC DU
MOULIN
B1
1987
0
1
km

be correct (cophenetic correlation = .907) (Nei et al, 1985). In some cases sites within a region have clustered together (North Carolina (NC1,2,3), Iowa (CR1,2,3)). In other cases sites from different regions have been clustered together. For example the cluster diagram indicates that sites H1 and H2 from the south stream of the Mont-St-Hilaire watershed are more closely related to sites from Ontario, New Brunswick, and Santa Cruz Island, CA than they are to other sites from the Mont-St-Hilaire watershed. This type of clustering indicates that the presence of particular alleles at a site is not necessarily associated with the region in which the site is found.

One clear feature of the dendrogram is the separation of the North Carolina sites from all other sites. This separation supports the specific status assigned to southeastern populations by Michel (1961).

Averaged estimates of Nei's genetic distances among sites in different watersheds (Table 6) and regions (Table 7) are given. With the exception of St. Andrew's #2, watersheds (table 6) which contained more than 1 site, had genetic distances between sites within a watershed which were less than the genetic distances between sites within that watershed and sites from different watersheds. The values in table 7 indicate that while the average genetic distance between sites within a region is generally smaller than the average genetic distance between sites in that region and sites in other

Table 5. Nei's genetic distances (above diagonal) and Roger's genetic distance (below diagonal) between all pairs of sites.

SITE	H1	H2	H3	H4	H5	H6	B1	B2	T1	T2	T3	L1	L2	CR1	CR2
H1	****	.000	.006	.008	.009	.002	.114	.093	.015	.015	.087	.001	.005	.098	.076
H2	.005	****	.009	.011	.012	.003	.118	.092	.016	.016	.088	.000	.003	.095	.072
H3	.023	.028	****	.000	.001	.002	.108	.108	.015	.012	.084	.011	.022	.111	.088
H4	.031	.035	.007	****	.000	.002	.111	.113	.012	.009	.082	.014	.025	.110	.086
H5	.032	.037	.009	.003	****	.003	.111	.113	.011	.008	.081	.015	.026	.108	.085
H6	.015	.020	.012	.015	.017	****	.112	.100	.010	.008	.081	.005	.012	.099	.076
B1	.118	.120	.118	.124	.126	.122	****	.218	.118	.123	.201	.121	.132	.240	.212
B2	.114	.110	.136	.143	.145	.128	.217	****	.114	.114	.114	.091	.090	.152	.168
T1	.042	.045	.047	.044	.042	.036	.140	.152	****	.001	.076	.016	.023	.085	.062
T2	.038	.042	.037	.034	.032	.027	.142	.148	.011	****	.074	.016	.024	.086	.064
T3	.104	.108	.101	.098	.096	.092	.206	.149	.080	.071	****	.089	.097	.169	.142
L1	.011	.006	.031	.038	.040	.023	.123	.109	.048	.045	.111	****	.002	.092	.070
L2	.023	.019	.046	.054	.056	.039	.131	.113	.063	.060	.127	.016	****	.092	.069
CR1	.124	.119	.140	.140	.139	.129	.233	.149	.125	.122	.191	.115	.110	****	.019
CR2	.085	.081	.101	.101	.101	.091	.194	.176	.086	.084	.152	.077	.071	.039	****
CR3	.081	.076	.097	.096	.096	.086	.191	.173	.082	.079	.148	.072	.075	.043	.005
NB1	.156	.151	.171	.179	.180	.163	.263	.111	.188	.185	.184	.146	.147	.237	.204
NB2	.152	.148	.176	.183	.185	.168	.261	.134	.193	.189	.189	.146	.138	.195	.200
NB3	.023	.018	.038	.045	.047	.030	.132	.113	.055	.052	.118	.013	.009	.104	.065
NC1	.484	.484	.489	.489	.488	.483	.509	.475	.471	.475	.476	.484	.490	.411	.423
NC2	.349	.346	.362	.362	.361	.353	.387	.335	.339	.345	.346	.343	.350	.266	.279
NC3	.219	.215	.242	.250	.252	.235	.261	.242	.250	.256	.255	.213	.205	.299	.267
C1	.085	.080	.100	.100	.099	.090	.193	.174	.085	.083	.151	.076	.073	.107	.069
C2	.081	.077	.097	.097	.096	.086	.181	.171	.076	.078	.145	.072	.068	.103	.065
C3	.029	.024	.046	.049	.050	.036	.138	.119	.054	.051	.117	.020	.019	.103	.065
C4	.022	.018	.042	.048	.050	.034	.131	.113	.058	.054	.121	.014	.008	.102	.063
C5	.082	.077	.098	.097	.097	.087	.190	.172	.082	.080	.148	.073	.068	.102	.064
C6	.079	.074	.094	.094	.094	.084	.187	.168	.081	.078	.146	.070	.067	.108	.069

Table 5. continued.

SITE	CR3	NB1	NB2	NB3	NC1	NC2	NC3	C1	C2	C3	C4	C5	C6
H1	.073	.152	.152	.006	.620	.422	.233	.071	.061	.007	.006	.068	.056
H2	.071	.149	.148	.003	.617	.418	.230	.068	.058	.005	.003	.065	.053
H3	.083	.170	.172	.021	.631	.437	.256	.082	.073	.021	.022	.080	.068
H4	.081	.174	.176	.024	.630	.436	.262	.080	.071	.024	.025	.079	.067
H5	.079	.174	.177	.025	.628	.435	.263	.079	.070	.024	.025	.077	.066
H6	.072	.159	.161	.012	.620	.424	.244	.070	.061	.012	.012	.068	.056
B1	.207	.294	.299	.133	.671	.486	.299	.205	.185	.133	.133	.204	.188
B2	.168	.091	.129	.091	.598	.399	.252	.163	.154	.093	.091	.160	.147
T1	.059	.172	.176	.019	.599	.402	.254	.057	.046	.019	.021	.054	.047
T2	.060	.172	.176	.020	.607	.411	.263	.058	.049	.020	.023	.056	.048
T3	.138	.172	.176	.093	.608	.413	.261	.137	.127	.094	.096	.134	.126
L1	.069	.148	.147	.002	.615	.415	.228	.065	.055	.003	.002	.062	.050
L2	.070	.146	.144	.000	.618	.415	.224	.065	.055	.002	.000	.062	.050
CR1	.020	.246	.207	.087	.461	.290	.334	.087	.077	.082	.084	.084	.083
CR2	.000	.221	.223	.065	.485	.316	.310	.065	.055	.060	.062	.062	.061
CR3	****	.222	.225	.065	.485	.316	.312	.065	.055	.060	.062	.062	.061
NB1	.203	****	.226	.146	.609	.411	.226	.217	.207	.149	.146	.214	.202
NB2	.205	.210	****	.144	.614	.413	.310	.220	.210	.147	.144	.216	.203
NB3	.069	.140	.138	****	.610	.409	.225	.060	.050	.002	.000	.058	.046
NC1	.420	.483	.490	.487	****	.149	.348	.609	.595	.604	.603	.606	.606
NC2	.277	.344	.346	.343	.214	****	.148	.410	.394	.404	.404	.405	.407
NC3	.271	.210	.267	.204	.358	.149	****	.307	.292	.229	.224	.303	.290
C1	.071	.201	.202	.066	.484	.345	.269	****	.005	.044	.061	.000	.001
C2	.068	.199	.198	.062	.479	.335	.258	.025	****	.037	.051	.003	.006
C3	.066	.149	.148	.013	.483	.339	.215	.057	.055	****	.002	.042	.032
C4	.068	.143	.137	.004	.486	.342	.204	.067	.063	.013	****	.058	.047
C5	.068	.200	.197	.061	.485	.342	.263	.008	.019	.055	.062	****	.002
C6	.071	.194	.197	.061	.481	.345	.263	.013	.029	.051	.062	.017	****

Figure 7. UPGMA dendrogram of Roger's genetic distance among sites.

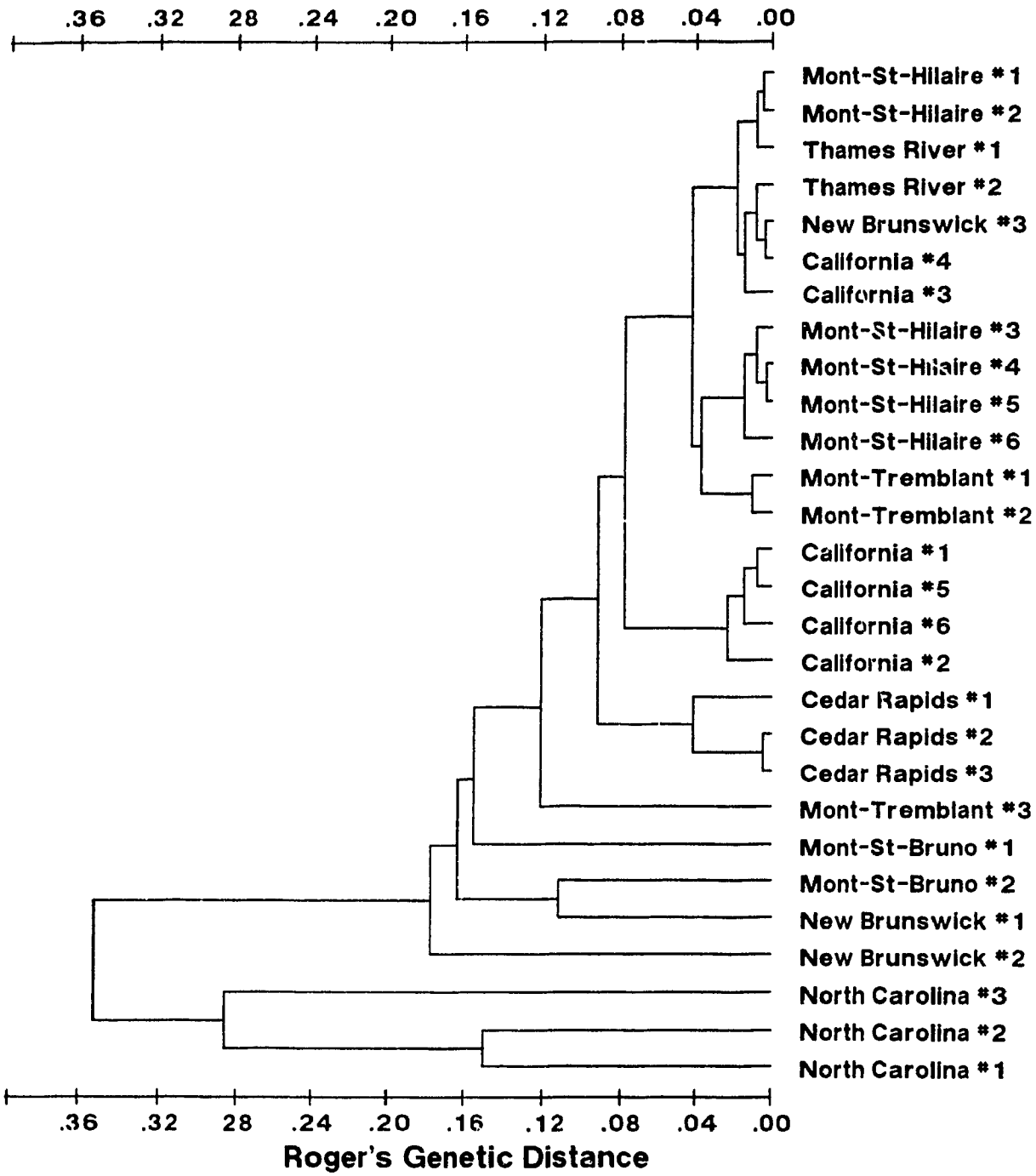

.36 .32 28 .24 .20 .16 .12 .08 .04 .00
Mont-St-Hilaire #1
Mont-St-Hilaire #2
Thames River #1
Thames River #2
New Brunswick #3
California #4
California #3
Mont-St-Hilaire #3
Mont-St-Hilaire #4
Mont-St-Hilaire #5
Mont-St-Hilaire #6
Mont-Tremblant #1
Mont-Tremblant #2
California #1
California #5
California #6
California #2
Cedar Rapids #1
Cedar Rapids #2
Cedar Rapids #3
Mont-Tremblant #3
Mont-St-Bruno #1
Mont-St-Bruno #2
New Brunswick #1
New Brunswick #2
North Carolina #3
North Carolina #2
North Carolina #1
.36 .32 28 .24 .20 .16 .12 .08 .04 .00
Roger's Genetic Distance

Table 6. Nei's genetic distance between sites in different watersheds. Values on the diagonal are mean genetic distances among sites within a watershed.

WATERSHED	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.Mont-St-Hilaire	.004																	
2.Mont-St-Bruno #1	.112	***																
3.Mont-St-Bruno #2	.103	.218	***															
4.Mont Tremblant	.036	.147	.114	.050														
5.Thames River	.012	.126	.091	.044	.002													
6.Cedar Rapids #1	.104	.240	.152	.113	.092	***												
7.Cedar Rapids #2	.080	.212	.168	.089	.069	.019	***											
8.Cedar Rapids #3	.077	.207	.168	.086	.069	.020	.000	***										
9.St. Andrew's #1	.163	.294	.091	.172	.147	.246	.221	.222	***									
10.St. Andrew's #2	.090	.216	.110	.110	.073	.147	.144	.145	.186	.144								
11.Deep River	.624	.671	.598	.605	.616	.461	.485	.485	.609	.612	***							
12.Yadkin River	.429	.486	.399	.409	.415	.290	.316	.316	.411	.411	.149	***						
13.Blue Ridge	.248	.299	.252	.259	.226	.334	.310	.312	.226	.267	.348	.148	***					
14.Klamath River	.070	.195	.159	.079	.060	.082	.060	.060	.212	.135	.602	.402	.299	.005				
15.Santa Cruz Is. #1	.016	.133	.093	.044	.003	.082	.060	.060	.149	.075	.604	.404	.229	.040	***			
16.Santa Cruz Is. #2	.015	.133	.091	.046	.001	.084	.062	.062	.146	.072	.603	.404	.224	.056	.002	***		
17.Los Padres Forest	.073	.204	.160	.082	.062	.084	.062	.062	.214	.137	.606	.407	.303	.002	.042	.058	***	
18.Angeles Forest	.061	.188	.147	.074	.050	.083	.061	.061	.202	.125	.606	.407	.290	.004	.032	.047	.002	***

Table 7. Nei's genetic distance between sites in different regions with ranges given in brackets. Values on the diagonal are for sites within a region.

REGION	QUEBEC	ONTARIO	IOWA	NEW BRUNSWICK	NORTH CAROLINA	CALIFORNIA
QUEBEC	.058 (.000- .218)					
ONTARIO	.038 (.000- .132)	.002 (.002- .002)				
IOWA	.108 (.059- .240)	.077 (.069- .092)	.013 (.000- .020)			
NEW BRUNSWICK	.129 (.003- .299)	.098 (.000- .148)	.174 (.065- .246)	.172 (.144- .226)		
NORTH CAROLINA	.434 (.230- .671)	.419 (.224- .618)	.368 (.290- .485)	.425 (.225- .614)	.215 (.148- .348)	
CALIFORNIA	.074 (.003- .205)	.039 (.000- .065)	.068 (.055- .087)	.138 (.000- .220)	.427 (.224- .609)	.026 (.000- .061)

regions, the range of genetic distances for sites within a region greatly overlaps the range of genetic distances between sites within that region and sites in other regions. Since only the Mont-St-Hilaire watershed had more than 2 sites the pattern of overlap found for sites within regions is not as reliable for sites within watersheds. Overall the highest levels of genetic divergence are between North Carolina sites and all other sites, supporting the specific status assigned by Michel (1961)

F statistics and gene flow estimates are given in table 8 and show the same pattern of site differentiation as found in the matrices of averaged genetic distances. The F statistics indicate that within a stream there is very little genetic differentiation. Above the level of sites within a stream subgroups are highly genetically divergent. Each of the F statistics can be thought of as representing the proportion of allelic variance within the group that is due to variance among the sub-groups. The proportion of variance within the sub-groups within the group is 1 - F. Thus 0.9% of the allelic variance within streams is found among sites and 99.1% is found within the sites. Similarly, 46.3% of the allelic variance within watersheds is found among streams and 53.7% is found within streams. When considering variability at each level relative to the level above it, it is important to remember that each level incorporates the variability at all levels below it.

Table 8. Hierarchical F statistics and gene flow for all sampling levels.

Sub-group	Total group	Variance	F_{ST}	Nm[a]
Site	Stream	.00497	.009	27.53
Site	Watershed	.47845	.468	.28
Site	Region	1.00602	.649	.14
Site	Total	1.78132	.766	.08
Stream	Watershed	.47348	.463	.29
Stream	Region	1.00105	.646	.14
Stream	Total	1.77635	.764	.08
Watershed	Region	.52756	.340	.49
Watershed	Total	1.30287	.560	.20
Region	Total	.77530	.333	.50

[a]Nm calculated from $F_{ST} = 1 / (4N_em + 1)$.

F statistics may also be interpreted as the proportion of the total variance, or gene diversity, found at each level of the hierarchy. For the lowest level, sites, the proportion of total variance is calculated as 1 - F (sites within total). Thus, 23.4% of the total variance is found within sites. For the highest level, among regions, the proportion of total variance is simply F (regions within total) which is 33.3%. For intermediate levels the proportion of total variance is calculated as the difference between the F values, relative to the total, for the level in question and the level above the level in question. Thus 0.2% of the total variance is found among sites, 20.4% is found among streams, 22.7% is found among watersheds. These results indicate that, except for among sites, the proportion of total variance is relatively similar at all levels. The low proportion of total variance at the level of among sites shows that there is very little genetic isolation among sites when these sites are considered in the context of highly genetically isolated streams.

Genetic divergence of sites is expected to be greatly influenced by dispersal ability, in this case the presence of long-winged individuals. Genetic divergence can not be compared among populations grouped by presence or absence of long-winged individuals using the F values presented above since none of the levels represents a grouping of this type. Thus, separate F statistics were calculated based on groupings of sites with or without long-winged individuals. $F_{s'}$ values

were 0.795 among watersheds containing only wingless individuals and 0.496 among watersheds containing long-winged individuals (California), indicating a higher level of genetic isolation for populations containing only wingless individuals. The North Carolina sites were not used in the calculation of these F_{ST} values because of the high level of differentiation of these sites from all other sites and their possible specific status.

Gene flow estimates (table 8) indicate that only sites within streams exchange enough individuals to avoid genetic differentiation through genetic drift (Nm > 1). Gene flow among sites within streams was approximately two orders of magnitude larger than gene flow between streams within a watershed. If Slatkin's alternate estimate of an individual exchanged every second year being sufficient to prevent genetic differentiation through genetic drift (Nm > .5) is used, the divergence of watersheds within a region, and among regions themselves, may not occur.

The two-tailed Mantel test for a correlation between matrices of Nei's genetic distance and geographic distance was not significant (r=-0.0265, p=.844) indicating that sites that are geographically close are not necessarily genetically more similar. The plot of all pairs of genetic and geographic distances supports the lack of association between these measures (Figure 8). The two-tailed Mantel test for a correlation between Nei's genetic distance and the binary

connectivity matrix was negative and significant (r=-0.2033,p=.002). This indicates that sites are less related to their closest neighbors than they are to other more distant neighbors. The alternative method of visually examining the relation between genetic and geographic distance was multidimensional scaling (MDS). Figure 9 shows the map produced by the multidimensional scaling procedure applied to the matrix of genetic distances between watersheds. It is clear that the MDS map produced from Nei's genetic distances bears no resemblance to a geographic map of all watersheds sampled.

Figure 8. Plot of Nei's genetic distance against geographic distance for all pairs of sites.

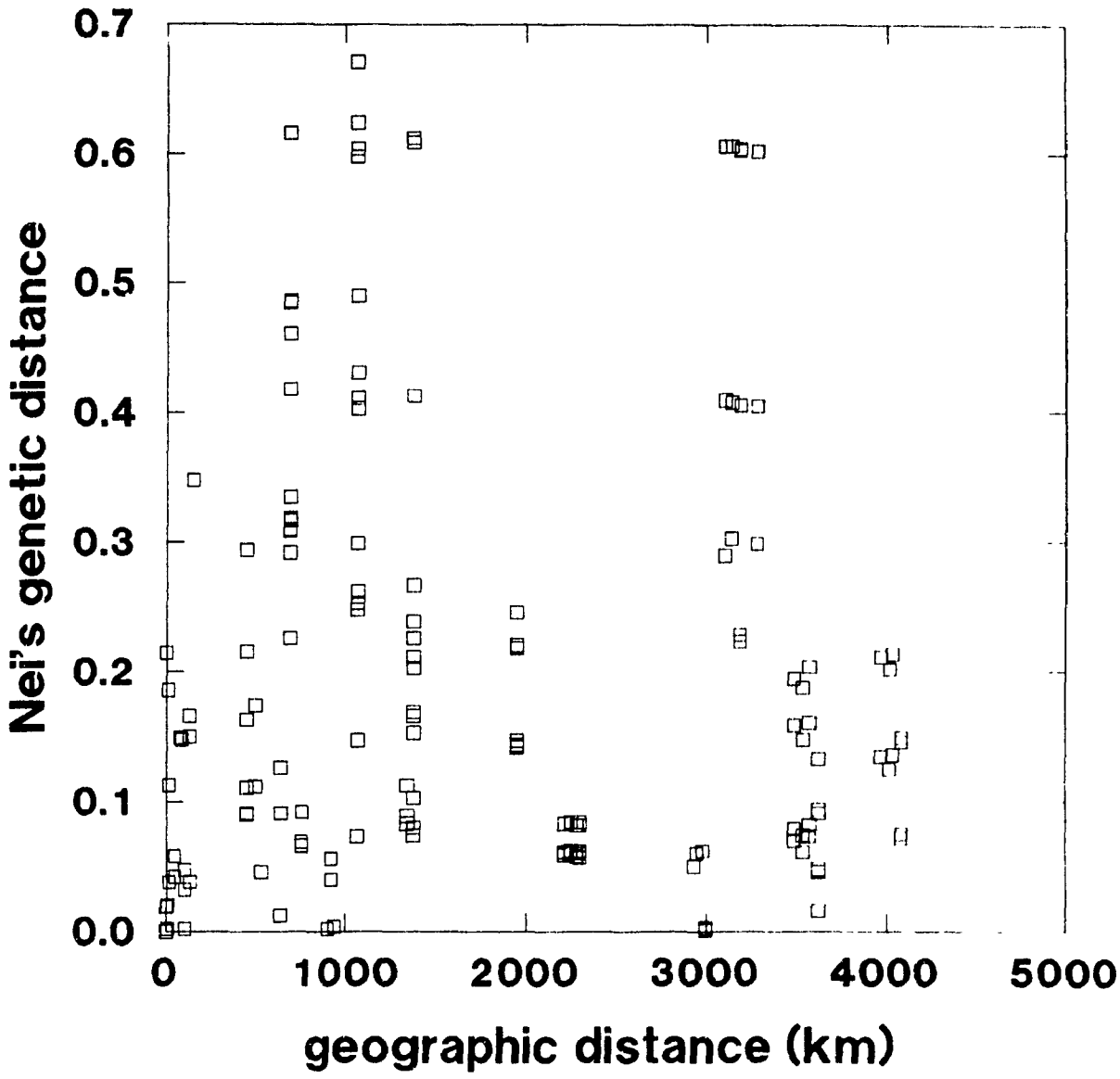
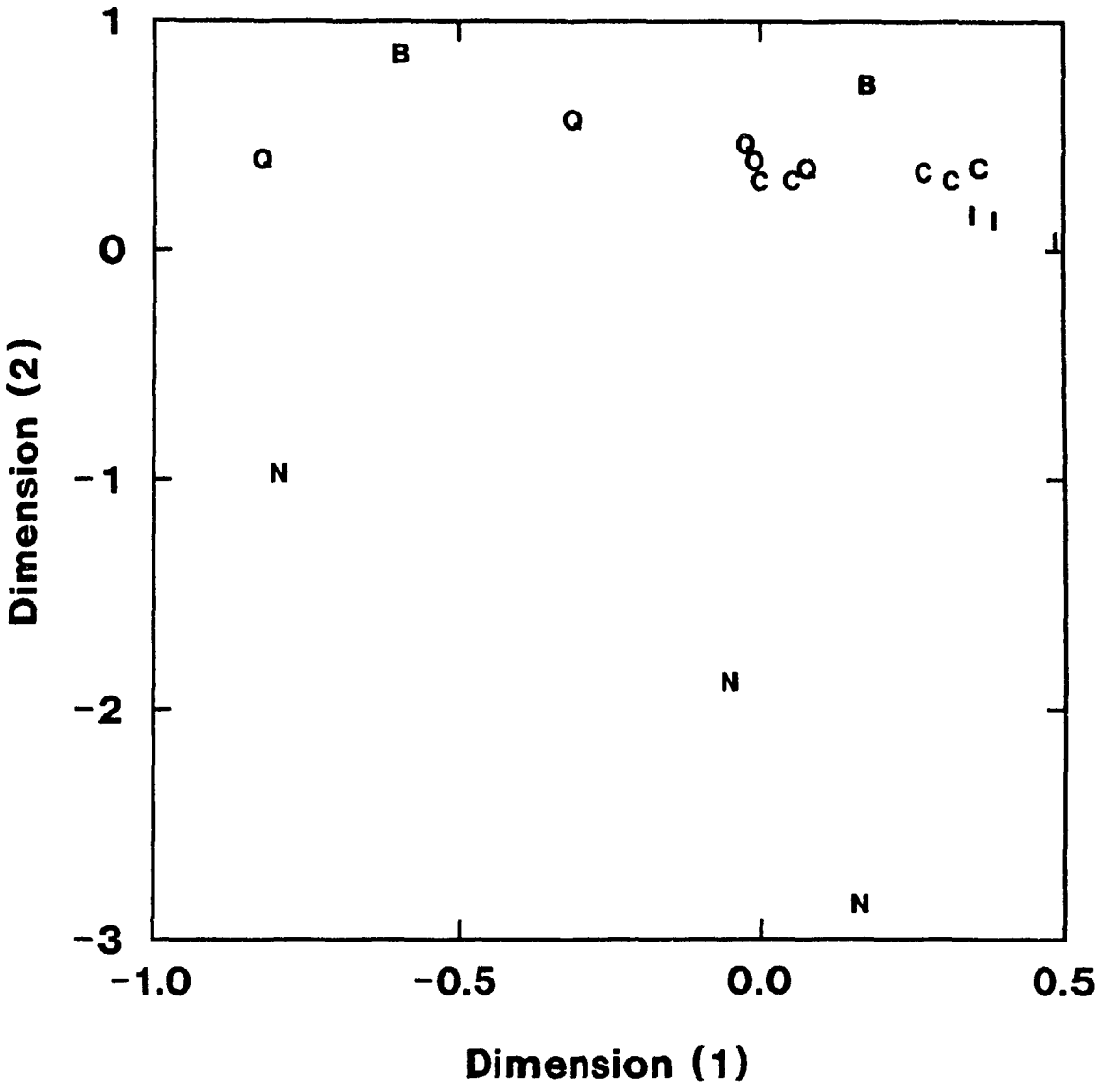

0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Nei's genetic distance
0
1000
2000
3000
4000
5000
geographic distance (km)

Figure 9. Map produced by MDS analysis of Nei's genetic distance (see text). Symbols represent watersheds and are coded by region (Q=Québec, O=Ontario, I=Iowa, B=New Brunswick, C=California, N=North Carolina).


1
0
−1
−2
−3
−1.0
−0.5
0.0
0.5
Dimension (2)
Dimension (1)
B
B
Q
Q
Q
Q
C
C
C
C
C
C
I
I
N
N
N

DISCUSSION.

In this study I show that gene flow may be highly restricted among local populations. For *Gerris remigis*, the population subdivision level at which gene flow may prevent differentiation of subdivisions is that of sites within streams. All levels of subdivision above sites within streams (i.e. streams within watersheds, watersheds within regions, regions within North America) have greatly reduced gene flow and may differentiate from each other through the effects of drift. In addition I have examined the effects of apparent dispersal ability on genetic differentiation of population subdivisions by comparing the degree of fixation among groups with and without long-winged individuals. The degree of fixation among subdivisions with winged individuals was lower than among subdivisions without winged individuals but in neither group was gene flow high enough to prevent genetic differentiation by drift. Each aspect of the effect of population subdivision is examined in detail below.

Neighborhood size and genetic variation within a stream.

There is almost no genetic differentiation of sites within a stream relative to the total genetic variation of all sites combined. F values for sites within the total and streams within the total differed by only .002. Both of these values were larger than F for watersheds within the total or regions within the total, indicating that, while there is

differentiation of watersheds and regions, the majority of genetic differentiation is found at the level of streams. Similarly, sites within a stream had a much lower F than streams within a watershed while watersheds within a region and regions within the total were similar and slightly smaller than streams within a watershed. These results support the initial hypothesis of high levels of genetic isolation among streams within a watershed. The largest level of genetic divergence was expected to be found at the level of watersheds but was found at the even lower level of streams.

Gene flow estimates calculated from these F statistics indicate that only sites within a stream exchange a sufficient number of individuals to prevent genetic differentiation through the effects of drift, the difference between gene flow estimates for sites within a stream and streams within a watershed being approxmately two orders of magnitude. This sharp cutoff in gene flow above the level of sites within streams indicates that the neighborhood size of *G. remigis* is at the level of a stream or smaller. Isolation by distance was expected within a stream only if streams were long enough to include more than one neighborhood. For the Mont-St-Hilaire streams this appears to be the case since allele frequencies differed significantly among sites within a stream. Since sample sites were more than 113 meters apart these results agree well with the neighborhood size estimates made from Fairbairn's data.

Comparison of allele frequencies between years.

Comparisons of allele frequencies between years were significant for 8 out of 26 comparisons at 4 loci and 6 sites. These comparisons remain significant even when χ^2 values are corrected by Waples (1989) technique. Rejection of the null hypothesis of Waples test indicates that sampling error and drift combined cannot explain the changes in allele frequencies between years. The χ^2 values generated by Waples test are greatly affected by effective population size. For most of the populations where χ^2 contingency tests revealed significant differences in allele frequencies, the effective population sizes would have to be quite low (e.g. less than 50 individuals) for a significant difference not to be found using Waples (1989) test. An N_e of 50 is small relative to the minimum estimated effective neighborhood size for *G. remigis*, but considering the high overwinter mortality and founder effects that would result from limited dispersal ability, and the fact that these effects may be compounded several times, it is possible that neighborhood sizes are much smaller than can be estimated from mark-recapture data. If we use the neighborhood size calculated from Fairbairn's data (170 individuals), some force in excess of sampling error and drift must be invoked to explain the changes in gene frequencies observed between years.

Considering the results of comparisons between years from Mont-St-Hilaire, it is unlikely that selection of such

intensity would be acting at only one site out of six in such close proximity. This pattern is found for two loci, the site showing significant differences being different for each locus. Even if selection is occurring differentially at sites in such close proximity, the high levels of gene flow found within streams would be expected to eliminate the effects of selection.

An alternative explanation for apparent changes in allele frequencies between years is the sampling of sibling groups. This problem has been reported in several studies (Varvio-Aho, 1979; Parkinson, 1984; and Guttman and Weigt, 1989; Reisenbichler and Phelps, 1989). Varvio-Aho (1979) examined seasonal and yearly changes in allele frequencies in *G. odontogaster* and *G. lacustris* in Finland, and found that simple binomial sampling could not explain changes in gene frequencies of the magnitude observed. Varvio-Aho suggested that, in waterstriders, sampling error may not only be an error of non-random sampling of individuals by the investigator but may also include a sampling error in the development of nymphal stages. Fairbairn (1985b) estimated that the net movement of summer born adult *G. remigis* before overwintering was only 7.2 meters, compared with a distance of 39.5 meters for *G. remigis* in the spring after overwintering. Thus, sibling groups may remain together until late fall when they leave the water surface to diapause. Lack of mixing of sibling groups may be especially true in small streams where

the surface area may be reduced during the summer because of drying. If individuals do not disperse along the stream, samples of individuals taken in the late summer or fall may consist of a small number of sibling groups. Samples with this type of error could be expected to show differences in allele frequencies between years since it is a small number of sibling groups and not really the neighborhood that have been sampled. If this type of sampling error is present, allele and genotype frequencies from a single sample may not accurately reflect frequencies of the neighborhood. Therefore, where two years of data were available for a site, the data were combined for all other analysis on the assumption that the combined data, being representative of more sib-groups, would be more representative of the neighborhood allele frequencies. Data were combined by summing the data from the two years to maintain a large sample size.

Macrogeographic variation.

In a comparison of genetic population structure in the eastern U.S. between the long-winged *L. canaliculatus* and the almost wingless *G. remigis*, Zera (1981) found that *G. remigis* populations were highly genetically divergent and suggested that this was due to a lack of gene flow between populations, small population size, bottlenecks, and founder effects. My results also show high levels of genetic isolation and

differentiation for all levels above sites within streams. At none of these levels is the estimated gene flow high enough to prevent genetic differentiation by drift. McCauley and Eanes (1987), using the data from Zera (1981), calculated an F value of 0.082 for *L. canaliculatus*. Varvio-Aho (1979) examined the genetic divergence of long-winged *G. odontogaster* populations (G_{ST} = 0.055). Studies of the wing polymorphic *G. lacustris* (Varvio-Aho and Pamilo; 1979,1981) have shown a high degree of genetic isolation among populations (G_{ST}=.2832). All of these values are small when compared to the F_{ST} for streams within the total (F = 0.764, the most nearly comparable F statistic) for *G. remigis*. The increased F value for *G. remigis* indicates that, among Gerridae species, genetic divergence of population subunits is higher for species with decreased dispersal ability.

For *G. remigis*, a high proportion of the total allelic variance may be found at each of the levels above sites within streams. Since such a small proportion of the total alleleic variance is found among sites within watersheds, the maximum level to which sites within watersheds can diverge is quite limited relative to the levels of divergence available to watersheds within regions or regions within North America.

In continuous populations genetic drift should lead to a correlation between geographic and genetic distance among population subunits. This would also be true for subdivided populations if the gene flow between neighboring subunits is

sufficiently high to prevent significant genetic differentiation of the subunits ($Nm > 1$) but gene flow between distant populations is more restricted ($Nm < 1$). If no such association exists between genetic and geographic distance selection may be acting independently on different subunits (overriding the effects of any gene flow that may exist) or there may be very low or no gene flow between neighboring subunits (making the neighborhood size equal to or less than the subunit examined). For selection to explain such a pattern would require that it is acting in a different manner in different population subunits and that patterns of selection are themselves uncorrelated with geographic distance. For my data, when comparisons are made between genetic and geographic distance at the level of watersheds, no association is found. Lack of a significant association reflects the high degree of genetic isolation at all levels above sites within streams.

Guttman and Weigt (1989) found no correlation between genetic distance and geographic distance for populations of treehoppers on individual trees in the *Enchenopa binotata* complex, a species complex of treehoppers which are highly host specific. King (1987) also found no correlation for the beetle *Collops georgianus*, found on granite outcrop 'islands' in the south-eastern U.S. King also tested for an association between genetic distance and a binary connectivity matrix to determine if the closest site in a pair is genetically more

similar than all other sites. This method is based on the idea that when individuals disperse between islands by flight, they travel until they reach the next island, regardless of the distance. When such a comparison is made for my data, a significant negative association is found, indicating that watersheds are genetically less similar to the geographically closest watershed than they are to all other watersheds. While this relationship indicates that there is no positive association, the negative correlation does not have any obvious explanation.

The matrix of genetic distances averaged by region indicates that genetic distances are smaller between sites within a region than between sites between regions, but that the range of genetic distances within a region has a large overlap with the range of genetic distances between regions. Approximately 85% of all site-locus combinations were fixed for a single allele and approximately 45% of all alleles were unique to a region. This high level of fixation of alternate alleles without geographic pattern indicates the effects of drift. Cluster analysis of Roger's genetic distances grouped some sites by region but the general pattern of clustering of sites appears to be random. My results at this level agree well with Zera's proposal that populations are highly genetically isolated and divergent due to the effects of drift.

Calabrese (1974) has suggested that western *G. remigis* from

California and Oregon are a separate subspecies on the basis of morphology and proportion of long-winged individuals. The average genetic distances between regions indicate that California is not obviously genetically distinct from any regions except North Carolina and possibly New Brunswick. Based on these results *G. remigis* from the south-western U.S. do not form a genetic sub-species distinct from north-central and northeastern *G. remigis*. However, North Carolina has genetic distances from other regions which would suggest divergence at the species level as suggested by Michel (1961) on the basis of morphological data. The North Carolina region had high variability between sites and sample sizes were small, yet out of 28 alleles found in North Carolina populations, 11 were unique to that region.

Influence of dispersal by flight on genetic structure.

Liebherr (1986) found similar levels of genetic differentiation for two species of carabid beetles (F_{ST}=.26 and F_{ST}=.27) in spite of the fact that one species was fully winged and the other had only vestigial wings and could not fly. *G. odontogaster* is monomorphic for long wings and inhabits much more temporary sites than does *G. lacustris*, which is polymorphic for wing length. Varvio-Aho (1979) concluded that the difference between the genetic population structures of the two species was due to the higher degree of isolation of *G. lacustris* populations because of reduced dispersal ability.

For *G. remigis*, the F value for streams within the total is larger than the G_{ST} value given for *G. lacustris* (Varvio-Aho and Pamilo, 1979). This relationship supports the hypothesis that, at the species level, population subdivision is related to dispersal ability.

Fairbairn and Desranleau (1987) have previously estimated long distance dispersal in *G. remigis* by flight at .03%/population/generation. Both Zera and I have shown that *G. remigis* populations are highly isolated and highly genetically divergent. I have also shown that isolation occurs at the level of streams within a watershed, and that although a significant proportion of the total variability may be found at higher levels, differentiation of populations at lower levels is so large that no associations between genetic and geographic distances, or genetic distances between contiguous and non-contiguous sites, are found. However, Zera's (1981) conclusions, Fairbairn and Desranleau's (1987) conclusions and, for the most part, my conclusions are based on examination of almost completely wingless populations of *G. remigis*. Although Fairbairn (1986) found that long-winged *G. remigis* in the Mont-St-Hilaire watershed were no more dispersive within the watershed than wingless individuals, populations of *G. remigis* in the south-western U.S. often have a majority of long-winged individuals which do disperse by flight (Calabrese, 1974; Fairbairn, pers. comm.), a factor that would be expected to influence the degree of genetic

divergence between populations.

In my results, the number of alleles per locus for sites was not significantly different between regions. Significance of differences in expected heterozygosity of sites among regions depended upon a single watershed in Québec with high values. If this watershed is removed, expected heterozygosity is not significantly different among regions. This result indicates that neither the level of heterozygosity nor the number of alleles per locus is related to the dispersal capacity of populations.

It was expected that the presence of long-winged individuals in the California region would result in a lower level of population differentiation, and, as anticipated, the F_{ST} for sites with long-winged individuals was much lower than the F_{ST} for sites with no long-winged individuals. While the degree of population differentiation is reduced in California, it remains high enough to indicate that gene flow is rare. In the east, population sizes are large in the fall but high overwinter mortality reduces population size up to 90% (Matthey, 1974). The number of individuals on the stream in spring will increase greatly by the fall, and be reduced once again over winter (Fairbairn, 1985a). Genetic drift would play a large role in the evolution of such populations. This, in combination with the poor dispersal ability of *G. remigis*, supplies a sufficient mechanism for high genetic divergence of population subunits. In much of California, populations

experience both winter and prolonged drought. The combination of these two seasons may have the same effect on population size in California as winter does in eastern populations (Fairbairn, pers. obs.) subjecting these populations to the same effects of drift which occur in the east despite the increased dispersal ability of population subunits.

SUMMARY.

G. remigis populations are highly subdivided. Genetic differentiation is highest at the level of streams and occurs within the context of lower levels of genetic differentiation among watersheds and regions. Gene flow is greatly reduced among streams within a watershed and only sites within a stream maintain high enough levels of gene flow to prevent genetic differentiation by drift. Neighborhood size was estimated to be smaller than a stream, approximately 170 individuals. This estimate is supported by the low level of differentiation among sites within a stream relative to the differentiation among sites at higher levels. However, the possibility of isolation by distance within a stream is supported by significant differences in allele frequencies among sites within streams. The overall pattern of genetic variation in *G. remigis* appears to be dominated by the effects of random drift, bottlenecks and founder events. These factors occur within the context of a lack of gene flow due to winglessness for eastern populations. For western populations, the increased dispersal ability of individuals has not provided sufficient gene flow to prevent differentiation by drift. The genetic population structure of *G. remigis* supports the views of Ehrlich and Raven (1969) that gene flow is highly reduced between populations and that the local population, in this case streams or possibly subunits of streams, is the unit of evolutionary importance.

Subspecific status of Californian *G. remigis* as suggested by Calabrese (1974), is not supported by genetic distances between California populations and other populations in North America. Specific status of south-eastern *G. remigis* as suggested by Michel (1961), is supported by genetic distances between North Carolina populations and other populations in North America, and a high proportion of region specific alleles in the North Carolina populations.

REFERENCES

Andersen, N.M. 1982. The Semiaquatic bugs (Hemiptera, Gerromorpha). Phylogeny, adaptations, biogeography, and classification. Entomonograph Vol. 3, Scandanavian Science Press Ltd., Klampenborg, Denmark.

Baker, A.E.M. 1981. Gene flow in house mice: introduction of a new allele into free-living populations. Evolution 35:243-258.

Barrowclough, G. F. 1980. Gene flow, effective population size, and genetic variance components in birds. Evolution 34(4):789-798.

Bowdan, E. 1978. Walking and rowing in the water strider, Gerris remigis: I. A cinematographic analysis of walking. J. Comp. Physiol. 123:43-49.

Calabrese, D. 1974. Population and subspecific variation in Gerris remigis Say. Ent. News 85:27-28.

Chepko-Sade, B.D. et al. 1987. The Effects of Social Structure on Population Genetics. University of Chicago Press, Chicago.

Coyne, J.A., S.H.Bryant, and M. Turelli. 1987. Long-distance migration of Drosophila. 2. presence in desolate sites and dispersal near a desert oasis. Am. Nat. 129:847-861.

Crow, J.F. 1986. Basic Concepts in Population, Quantitative, and Evolutionary Genetics. W.H. Freeman and Co., New York.

Daniel, W.W. 1978. Biostatistics: A Foundation for Analysis in Health Sciences, 2nd ed. John Wiley and Sons, New York.

Drake, C.J. and H.M. Harris. 1928. Concerning some North American water-striders with descriptions of three new species. Ohio Jour. Sci. 28:269-275.

Ehrlich, P.R. and P.H. Raven. 1969. Differentiation of populations. Science 163:1228-1232.

Endler, J.A. 1979. Gene flow and life history patterns. Genetics 93:263-284.

Fairbairn, D.J. 1985a. Comparative ecology of *Gerris remigis* (Hemiptera: Gerridae) in two habitats: a paradox of habitat choice. Can. J. Zool. 63:2594-2603.

Fairbairn, D.J. 1985b. A test of the hypothesis of compensatory upstream dispersal using a stream-dwelling waterstrider, *Gerris remigis* Say. Oecologia 66:147-153.

Fairbairn, D.J. 1986. Does alary polymorphism imply dispersal polymorphism in the waterstrider, *Gerris remigis*? Ecol. Ent. 11:355-368.

Fairbairn, D.J. and L. Desranleau. 1987. Flight threshold, wing muscle histolysis, and alary polymorphism: correlated traits for dispersal in the Gerridae. Ecol. Ent. 12:13-24.

Fairbairn, D.J. and D.A. Roff. 1980. Testing genetic models of isozyme variability without breeding data: Can we depend on the chi-square? Can. J. Fish. Aquat. Sci. 37:1149-1159

Falconer, D.S. 1989. Introduction to Quantitative Genetics, 3rd ed. John Wiley and Sons, New York.

Foote, C.J., C.C. Wood, and R.E Withler. 1989. Biochemical genetic comparison of sockeye salmon and kokanee, the anadromous and nonanadromous forms of *Oncorhynchus nerka*. Can. J. Fish. Aquat. Sci. 46:149-158.

Froeschner, R.C. 1962. Contributions to a synopsis of the Hemiptera of Missouri, Part V. Am. Mid. Nat. 67:208-240.

Galbraith, D.F. and C.H. Fernando. 1977. The life history of *Gerris remigis* (Heteroptera: Gerridae) in a small stream in southern Ontario. Can. Ent. 109:221-228.

Guttman, S.I. and L.A. Weigt. 1989. Macrogeographic genetic variation in the *Enchenopa binotata* Complex (Homoptera: Membracidae). Ann Entomol. Soc. Am. 82:156-165.

Harrison, R.G. 1977. Patterns of variation and genetic differentiation in closely related species: The field crickets of eastern North America. Thesis - Cornell University

Hartl, D.L. 1980. Principles of Populations Genetics. Sinauer Associates, Sunderland, Mass.

Hartl, D.L. 1987. A Primer of Population Genetics. pp. 69-141. Sinauer Associates, Sunderland, Mass.

Hedgecock, D. 1978. Population subdivision and genetic divergence in the red-bellied newt, *Taricha rivularis*. Evolution 32:271-286.

Johnson, M.S., B. Clarke, and J. Murray. 1988. Discrepancies in the estimation of gene flow in Partula. Genetics 120:233-238.

King, P. 1987. Macro- and microgeographic structure of a spatially subdivided beetle species in nature. Evolution 41:401-416.

King, P. 1988. Distribution and genetic structure of two allopatric beetle (Coleoptera: Melyridae) species on rock outcrops in the Southeast. Ann. Entomol. Soc. Am. 81:890-898.

Lewontin, R.C. 1974. The genetic basis of evolutionary change. Columbia University Press, New York.

Liebherr, J.K. 1986. Comparison of genetic variation in two Carabid beetles (Coleoptera) of differing vagility. Ann. Entomol. Soc. Am. 79:424-433.

Manly, B.J.F. 1986. Multivariate Statistical Methods: A Primer. Chapman and Hall, New York.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27:209-220.

Matthey, W. 1974. Contribution a l'ecologie de Gerris remigis Say sur deux etangs des Montagnes Rocheuses. Bull. de la Soc. Entomol. Suisse 47:85-95.

Maynard Smith, J. 1989. Evolutionary genetics. Oxford University Press, New York.

Mayr, E. 1963. Populations, Species, and Evolution. Harvard University Press. Cambridge, Mass.

McCauley, D.E. and W.F. Eanes. 1987. Hierarchical population structure analysis of the milkweed beetle, *Tetraopes tetraopthalmus* (Forster). Heredity 58:193-201.

McDonald, I.C., J.L. Krysan, and O.A. Johnson. 1985. Genetic variation within and among geographic populations of *Diabrotica barberi* (Coleoptera: Chrysomelidae). Ann. Entomol. Soc. Am. 78:271-278.

Michel, F.A. 1961. The taxonomic value of the male genitalia of the genus *Gerris* Fabricus (Hemiptera: Heteroptera: Gerridae). Thesis- Oregon State University.

Murray, J. and B. Clarke 1984. Movement and gene flow in *Partula taeniata*. Malacological 25:343-348.

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283-292.

Nei, M. 1975. Molecular Population Genetics and Evolution. North Holland, New York.

Nei, M., J.C. Stephens, and N. Saitou. 1985. Methods for computing the standard errors of branching points in an evolutionary tree and their application to molecular data from humans and apes. Mol. Biol. Evol. 2:66-85.

Parkinson, E.A. 1984. Genetic variation in populations of steelhead trout (*Salmo gairdneri*) in British Columbia. Can. J. Fish. Aquat. Sci. 41:1412-1420.

Polhemus, J.T. and H.C. Chapman. 1979. Family Gerridae-Water Striders, Pond Skaters, Wherrymen. in A.S. Menke (ed.) The Semiaquatic and Aquatic Hemiptera of California. Bull. California Insect Surv. 21.

Pounds, J.A. and J.F. Jackson. 1981. Riverine barriers to gene flow and the differentiation of fence lizard populations. Evolution 35:516-528.

Riesenbichler, R.R. and S.R. Phelps. 1989. Gentic variation in steelhead (*Salmo gairdneri*) from the north coast of washington. Can. J. Fish. Aquat. Sci. 46:66-73

Schaefer, C.W. and D.M. Calabrese. 1980. Amphi-atlantic species-pairs in two genera of water striders (Hemiptera: Gerridae). Entomologia Generalis 6:271-280.

Selander, R.K. and D.W.Kaufman. 1975. Genetic structure of populations of the brown snail (*Helix aspersa*) I. Microgeographic variation. Evolution 29:385-401.

Shaw, C.R. and R. Prasad. 1970. Starch gel electrophoresis of enzymes - A compilation of recipes. Biochem. Genet. 4:297-320.

Slatkin, M. 1981. Estimating levels of gene flow in natural populations. Genetics 99:323-335.

Slatkin, M. 1985. Rare alleles as indicators of gene flow. Evolution 39:53-65.

Slatkin, M. and N.H. Barton. 1989. A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43:1349-1368.

Sokal, R.R. 1979. Testing statistical significance of geographic variation patterns. Syst. Zool. 28:227-232.

Swofford, D.L. and R.B. Selander. 1981. BIOSYS-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. J. Heredity 72:281-283.

Varvio-Aho, S. 1979. Genic differentiation of *Gerris odontogaster* populations. Hereditas 91:207-214.

Varvio_aho, S. 1981. On the causes of seasonal genetic changes in *Gerris lacustris*. Hereditas 94:139-142.

Varvio_aho, S. and P. Pamilo. 1979. Genic differentiation of *Gerris lacustris* populations. Hereditas 90:237-249.

Varvio-Aho, S. and P. Pamilo. 1981. Spatio-temporal microdifferentiation of water-strider (*Gerris*) populations. Ent. Res. Camb. 37:253-263.

Vepsalainen, K. 1978. Wing dimorphism and diapause in Gerris: determination and adaptive significance. Evolution of Insect Migration and Diapause. (ed. H. Dingle) pp. 218-253. Springer, New York.

Waples, R.S. 1982. A multispecies approach to the analysis of gene flow in marine shore fishes. Evolution 41:385-400.

Waples, R.S. 1989. Temporal variation in allele frequencies: Testing the right hypothesis. Evolution 43:1236-1251.

Wilkinson, L. 1988. SYSTAT: The system for statistics. Evanston, IL.

Wright, S. 1931. Evolution in Mendalian populations. Genetics 16:97-159.

Wright, S. 1969. Evolution and the Genetics of Populations: 2. The Theory of Gene Frequencies. University of Chicago Press, Chicago.

Wright, S. 1978. Evolution and the Genetics of Populations: 4. Variability Within and Among Natural Populations. pp. 79-103. University of Chicago Press, Chicago.

Zera, A.J. 1981. Genetic structure of two species of waterstriders (Gerridae:Hemiptera) with differing degrees of winglessness. Evolution 35:218-225.

APPENDIX A

pH values for 0.135M Tris-citrate buffer (Shaw and Prasad 1970) used as both gel and chamber buffer for all enzyme systems.

Enzyme system	pH
ALP	7.3
PGD	8.3
MDH	7.3
GOT	7.3
GPD	7.3
GDH	7.3
LDH	8.3
MEZ	7.3
ICD	8.3
SOD	8.3

APPENDIX B

Allele frequencies of all populations. If populations were sampled For two years the allele frequencies given are based on the combined data for both years.

	Population								
Locus	H1	H2	H3	H4	H5	H6	B1	B2	T1
ALP-1									
(N)	308	392	237	230	195	167	25	289	471
A	.997	.992	.920	.841	.813	.892	1.00	.998	.541
B	.003	.008	.080	.159	.187	.108	.000	.002	.459
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
E	.000	.000	.000	.000	.000	.000	.000	.000	.000
F	.000	.000	.000	.000	.000	.000	.000	.000	.000
G	.000	.000	.000	.000	.000	.000	.000	.000	.000
ALP-2									
(N)	301	361	220	270	226	189	38	309	420
A	1.00	1.00	1.00	1.00	.998	1.00	1.00	.510	1.00
B	.000	.000	.000	.000	.002	.000	.000	.490	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
PGD-1									
(N)	296	372	288	246	224	222	36	316	486
A	.713	.788	.436	.413	.406	.588	.014	.842	.632
B	.280	.212	.549	.585	.576	.408	.167	.033	.291
C	.007	.000	.016	.002	.018	.005	.819	.125	.077
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
E	.000	.000	.000	.000	.000	.000	.000	.000	.000
F	.000	.000	.000	.000	.000	.000	.000	.000	.000
G	.000	.000	.000	.000	.000	.000	.000	.000	.000
MDH-1									
(N)	224	145	289	67	122	232	38	360	64
A	1.00	1.00	1.00	1.00	1.00	1.00	.000	1.00	.906
B	.000	.000	.000	.000	.000	.000	.000	.000	.094
C	.000	.000	.000	.000	.000	.000	1.00	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
E	.000	.000	.000	.000	.000	.000	.000	.000	.000

	Population								
Locus	H1	H2	H3	H4	H5	H6	B1	B2	T1
GOT-1									
(N)	308	385	289	277	226	240	38	364	420
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
GOT-2									
(N)	308	385	289	277	226	240	38	364	420
A	1.00	.999	1.00	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.001	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
GPD-1									
(N)	308	385	289	277	226	240	38	364	420
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
SOD-1									
(N)	308	385	289	277	226	240	38	323	420
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.000	1.00
B	.000	.000	.000	.000	.000	.000	.000	1.00	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
E	.000	.000	.000	.000	.000	.000	.000	.000	.000
GDH-1									
(N)	308	385	289	277	226	240	38	364	420
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
LDH-1									
(N)	308	385	289	277	226	240	19	364	420
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
LDH-2									
(N)	308	385	289	277	226	240	19	237	420
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

	Population								
Locus	H1	H2	H3	H4	H5	H6	B1	B2	T1
MEZ-1									
(N)	308	385	289	277	226	240	38	364	420
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
MEZ-2									
(N)	308	385	289	277	226	240	38	364	420
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ICD-1									
(N)	308	385	289	277	226	240	38	364	420
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ICD-2									
(N)	308	385	289	277	226	240	38	364	420
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

	Population								
Locus	T2	T3	L1	L2	CR1	CR2	CR3	NB1	NB2
ALP-1									
(N)	304	191	177	70	23	116	34	364	54
A	.551	.579	.972	1.00	.000	.000	.000	.935	1.00
B	.449	.421	.028	.000	.000	.000	.000	.054	.000
C	.000	.000	.000	.000	.000	.000	.000	.011	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
E	.000	.000	.000	.000	1.00	1.00	1.00	.000	.000
F	.000	.000	.000	.000	.000	.000	.000	.000	.000
G	.000	.000	.000	.000	.000	.000	.000	.000	.000
ALP-2									
(N)	172	136	128	70	53	67	35	332	36
A	.994	1.00	1.00	1.00	.472	1.00	1.00	1.00	.000
B	.006	.000	.000	.000	.528	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	1.00
D	.000	.000	.000	.000	.000	.000	.000	.000	.000

	Population								
Locus	T2	T3	L1	L2	CR1	CR2	CR3	NB1	NB2
PGD-1									
(N)	435	193	166	70	56	126	35	301	75
A	.583	.565	.837	.993	.991	1.00	.929	.934	1.00
B	.346	.368	.145	.000	.000	.000	.071	.022	.000
C	.071	.067	.012	.007	.000	.000	.000	.045	.000
D	.000	.000	.006	.000	.009	.000	.000	.000	.000
E	.000	.000	.000	.000	.000	.000	.000	.000	.000
F	.000	.000	.000	.000	.000	.000	.000	.000	.000
G	.000	.000	.000	.000	.000	.000	.000	.000	.000
MDH-1									
(N)	411	129	165	70	57	126	6	333	75
A	1.00	.988	1.00	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.012	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
E	.000	.000	.000	.000	.000	.000	.000	.000	.000
GOT-1									
(N)	439	200	177	70	57	126	35	333	75
A	1.00	1.00	1.00	.964	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.036	.000	.000	.000	.000	.000
GOT-2									
(N)	439	200	177	70	57	126	35	333	75
A	1.00	1.00	.992	.971	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.008	.029	.000	.000	.000	.000	.000
GPD-1									
(N)	439	200	177	70	57	126	35	333	75
A	1.00	1.00	1.00	1.00	.947	1.00	1.00	.000	1.00
B	.000	.000	.000	.000	.000	.000	.000	1.00	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.053	.000	.000	.000	.000
SOD-1									
(N)	411	26	177	41	16	102	35	333	18
A	1.00	.000	1.00	1.00	1.00	1.00	1.00	.000	.000
B	.000	.000	.000	.000	.000	.000	.000	1.00	.000
C	.000	1.00	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	1.00
E	.000	.000	.000	.000	.000	.000	.000	.000	.000

	Population								
Locus	T2	T3	L1	L2	CR1	CR2	CR3	NB1	NB2
GDH-1									
(N)	439	200	177	70	53	126	35	333	75
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
LDH-1									
(N)	439	200	177	70	46	122	35	333	75
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
LDH-2									
(N)	439	200	177	70	25	126	35	333	75
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
MEZ-1									
(N)	439	200	177	70	57	126	35	333	75
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.965	1.00
B	.000	.000	.000	.000	.000	.000	.000	.035	.000
C	.000	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000	.000
MEZ-2									
(N)	439	200	177	70	57	126	35	333	75
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ICD-1									
(N)	439	200	177	70	57	126	35	333	75
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ICD-2									
(N)	439	200	177	70	57	126	35	333	75
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

	Population							
Locus	NB3	NC1	NC2	NC3	C1	C2	C3	C4
ALP-1								
(N)	344	6	15	21	45	80	142	100
A	.933	.000	.000	1.00	.000	.000	.827	.945
B	.067	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000
E	.000	1.00	1.00	.000	.000	.000	.000	.055
F	.000	.000	.000	.000	.922	.669	.173	.000
G	.000	.000	.000	.000	.078	.331	.000	.000
ALP-2								
(N)	241	6	15	21	47	133	142	100
A	1.00	.000	.000	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	1.00	1.00	.000	.000	.000	.000	.000
PGD-1								
(N)	377	16	15	21	47	133	44	100
A	1.00	.500	.800	1.00	.968	.981	.955	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.031	.000	.000	.011	.019	.000	.000
D	.000	.125	.200	.000	.000	.000	.000	.000
E	.000	.094	.000	.000	.021	.000	.045	.000
F	.000	.250	.000	.000	.000	.000	.000	.000
G	.000	.000	.000	.000	.000	.000	.000	.000
MDH-1								
(N)	279	14	15	21	47	63	142	100
A	1.00	.000	.000	.000	1.00	.937	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.063	.000	.000
D	.000	.286	.000	.000	.000	.000	.000	.000
E	.000	.714	1.00	1.00	.000	.000	.000	.000
GOT-1								
(N)	377	16	14	21	47	133	142	100
A	1.00	.750	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.250	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000
GOT-2								
(N)	377	10	15	21	47	133	142	100
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000

	Population							
Locus	NB3	NC1	NC2	NC3	C1	C2	C3	C4
GPD-1								
(N)	377	16	15	21	47	133	142	100
A	1.00	.000	.000	.000	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.500	.033	.000	.000	.000	.000	.000
D	.000	.500	.967	1.00	.000	.000	.000	.000
SOD-1								
(N)	377	6	15	5	47	133	142	100
A	1.00	.000	.000	.000	1.00	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.000	.000	.000	.000
E	.000	1.00	1.00	1.00	.000	.000	.000	.000
GDH-1								
(N)	377	16	15	21	47	133	142	100
A	1.00	.344	1.00	1.00	1.00	1.00	1.00	1.00
B	.000	.219	.000	.000	.000	.000	.000	.000
C	.000	.438	.000	.000	.000	.000	.000	.000
LDH-1								
(N)	377	14	15	21	47	133	142	100
A	1.00	.000	1.00	1.00	.989	1.00	1.00	1.00
B	.000	.000	.000	.000	.000	.000	.000	.000
C	.000	1.00	.000	.000	.000	.000	.000	.000
D	.000	.000	.000	.000	.011	.000	.000	.000
LDH-2								
(N)	377	16	8	21	47	133	142	100
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
MEZ-1								
(N)	377	13	15	21	47	133	142	100
A	1.00	.654	1.00	1.00	.968	.996	1.00	1.00
B	.000	.346	.000	.000	.000	.000	.000	.000
C	.000	.000	.000	.000	.011	.000	.000	.000
D	.000	.000	.000	.000	.021	.004	.000	.000
MEZ-2								
(N)	377	16	15	21	47	133	142	100
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ICD-1								
(N)	377	16	15	21	47	133	142	100
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

	Population							
Locus	NB3	NC1	NC2	NC3	C1	C2	C3	C4
ICD-2								
(N)	377	16	15	21	47	133	142	100
A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

	Population	
Locus	C5	C6
ALP-1		
(N)	100	42
A	.000	.155
B	.000	.000
C	.000	.000
D	.000	.000
E	.000	.000
F	.875	.845
G	.125	.000
ALP-2		
(N)	100	100
A	1.00	1.00
B	.000	.000
C	.000	.000
D	.000	.000
PGD-1		
(N)	100	100
A	.995	.960
B	.000	.000
C	.000	.015
D	.000	.000
E	.000	.025
F	.000	.000
G	.005	.000
MDH-1		
(N)	100	100
A	.995	1.00
B	.005	.000
C	.000	.000
D	.000	.000
E	.000	.000

Locus	Population C5	C6
GOT-1		
(N)	100	100
A	1.00	1.00
B	.000	.000
C	.000	.000
GOT-2		
(N)	100	100
A	1.00	1.00
B	.000	.000
C	.000	.000
GPD-1		
(N)	100	100
A	1.00	1.00
B	.000	.000
C	.000	.000
D	.000	.000
SOD-1		
(N)	100	100
A	1.00	1.00
B	.000	.000
C	.000	.000
D	.000	.000
E	.000	.000
GDH-1		
(N)	100	100
A	1.00	1.00
B	.000	.000
C	.000	.000
LDH-1		
(N)	100	100
A	1.00	.995
B	.000	.005
C	.000	.000
D	.000	.000
LDH-2		
(N)	100	100
A	1.00	1.00

	Population	
Locus	C5	C6
MEZ-1		
(N)	100	101
A	1.00	.936
B	.000	.000
C	.000	.000
D	.000	.064
MEZ-2		
(N)	100	100
A	1.00	1.00
ICD-1		
(N)	100	100
A	1.00	1.00
ICD-2		
(N)	100	100
A	1.00	1.00